Analysis of promoter-associated chromatin interactions reveals biologically relevant candidate target genes at endometrial cancer risk loci

Tracy A. O'Mara [1], Endometrial Cancer Association Consortium, Amanda B. Spurdle [1], and Dylan M. Glubb [1,*]

[1] Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia

[*] Correspondence: dylan.glubb@qimrberghofer.edu.au; Tel.: +61-7-3845-3725

Abstract: The identification of target genes at genome-wide association study (GWAS) loci is a major obstacle for GWAS follow-up. To identify candidate target genes at the 16 known endometrial cancer GWAS risk loci, we performed HiChIP chromatin looping analysis of endometrial cell lines. To enrich for enhancer-promoter interactions, a mechanism through which GWAS variation may target genes, we captured loops associated with H3K27Ac histone, characteristic of promoters and enhancers. Analysis of HiChIP loops contacting promoters revealed enrichment for endometrial cancer GWAS heritability and intersection with endometrial cancer risk variation identified 103 HiChIP target genes at 13 risk loci. Expression of four HiChIP target genes (*SNX11*, *SRP14*, *HOXB2* and *BCL11A*) was associated with risk variation, providing further evidence for their regulation. Network analysis functionally prioritized a set of proteins that interact with those encoded by HiChIP target genes, and this set was enriched for pan-cancer and endometrial cancer drivers. Lastly, HiChIP target genes and prioritized interacting proteins were over-represented in pathways related to endometrial cancer development. In summary, we have generated the first global chromatin looping data from endometrial cells, enabling analysis of all known endometrial cancer risk loci and identifying biologically relevant candidate target genes.

Keywords: endometrial cancer risk; GWAS; HiChIP; H3K27Ac, chromatin looping; enhancer; promoter

1. Introduction

To date, 16 loci have been found to robustly associate with endometrial cancer risk by genome wide association studies (GWAS) [1-3]. As for other GWAS, the vast majority of credible variants (CVs; i.e. the lead variant and other correlated variants) at these loci are non-coding and likely mediate their effects through gene regulation (as reviewed in [4]). Indeed, we previously found that a majority of endometrial cancer risk CVs from ten recently discovered GWAS loci were coincident with promoter- or enhancer-associated epigenetic features in relevant cell lines or tissues [1]. Notably, the overlap between CVs and these elements was significantly greater for features observed in endometrial cancer cell lines after stimulation with estrogen [1], one of the most established risk factors for endometrial cancer [5-7].

The follow-up of GWAS is challenging because the target genes of CVs are generally not obvious; particularly as CVs located in enhancers can regulate promoter activity over long distances through chromatin looping [2,8-11] and enhancers do not necessarily loop to the nearest gene [12]. Fortunately, an array of chromatin conformation capture (3C) based techniques are now available to explore long-range chromatin looping [13] and identify candidate target genes at GWAS loci. These include a local low-throughput 3C approach, which we have previously used to identify chromatin looping between CV containing regions and *KLF5* at the 13q22.1 endometrial cancer risk

locus [2]. Ideally, what is needed is a high-throughput approach that can simultaneously identify functional elements and interrogate looping between elements containing CVs and all genes located at the 16 endometrial cancer risk loci. Hi-ChIP, a technique recently developed from the Hi-C global chromatin looping analysis, appears to fit these requirements. HiChIP can be used to assess chromatin loops associated with specific protein-bound regions, generating high resolution interactions using fewer cells and fewer sequencing reads than Hi-C [14]. HiChIP has thus been used to enrich for enhancer-promoter loops by capturing chromatin interactions associated with H3K27Ac [12], a histone mark characteristic of active promoters and enhancers. Moreover, the ability of H3K27Ac HiChIP to drive the interpretation of GWAS, through the identification of candidate target genes, has been demonstrated in two recent studies [12,15].

As global chromatin looping analyses have not yet been performed in endometrial cancer, we have used H3K27Ac HiChIP to identify likely regulatory promoter-associated chromatin interactions in normal and al endometrial cells. Using these data, we have assessed promoter-associated chromatin loops for enrichment of endometrial cancer heritability and identified candidate target genes of CVs at endometrial cancer risk loci. Lastly, we have performed network and pathway analyses of the candidate genes to aid the biological interpretation of endometrial cancer GWAS risk variation.

2. Results

2.1. H3K27Ac HiChIP analysis identifies promoter-associated chromatin loops in endometrial cell lines

To assess chromatin looping at endometrial cancer risk loci, we sequenced and analysed H3K27Ac HiChIP libraries from a normal immortalized endometrial cell lines (E6E7hTERT) and three endometrial cancer cell lines (ARK1, Ishikawa and JHUEM-14) for valid chromatin interactions (**Supplementary Table S1**). We identified 66,092 to 449,157 cis HiChIP loops (5 kb – 2 Mb in length) per cell line, with a majority involving interactions over 20 kb in distance (**Table 1**). Of the total loops, 35-40% had contact with a promoter and these promoter-associated loops had a median span >200 kb (**Table 1**), indicating that they may be involved with long-range gene regulation. BED files for promoter-associated loops can be found in Supplementary Files S1-4.

Table1. Characteristics of HiChIP loops in the endometrial cell lines.

Cell line	Total loops	Loops <20 kb	Loops >20 kb	Promoter-associated loops	Median span of promoter-associated loops (kb)
E6E7hTERT	162,476	25,133 (15.5%)	137,343 (84.5%)	59,658 (36.7%)	206
ARK1	449,157	45,932 (10.2%)	403,225 (89.8%)	155,08 (34.5%)	282
Ishikawa	219,067	29,954 (13.7%)	189,113 (86.3%)	79,309 (36.2%)	259
JHUEM-14	66,092	10,254 (15.5%)	55,838 (84.5%)	26,492 (40.0%)	209

2.2. HiChIP promoter loops are enriched for endometrial cancer heritability

To determine if promoter-associated HiChIP loops (i.e. those potentially involved with gene regulation) are enriched for heritability of endometrial cancer at a genome-wide level, we applied stratified linkage disequilibrium (LD) score regression analysis to the GWAS summary statistics from the largest study of endometrial cancer performed to date [1]. All four endometrial cell lines demonstrated an enrichment of endometrial cancer heritability in the anchors of promoter-associated loops (**Table 2**), although the enrichment in Ishikawa did not reach statistical significance (Bonferroni threshold, p <0.0125).

Table2. HiChIP promoter loops in endometrial cell lines are enriched for endometrial cancer heritability.

Cell line	Enrichment (standard error)	p-value
E6E7hTERT	6.92 (1.70)	1.30×10^{-04}
ARK1	4.08 (0.84)	2.50×10^{-04}
JHUEM-14	9.61 (3.11)	5.00×10^{-03}
Ishikawa	3.23 (1.18)	0.07

2.3. HiChIP promoter looping reveals 103 candidate target genes at endometrial cancer risk loci

To identify HiChIP target genes at the 16 known endometrial cancer GWAS loci, we intersected CVs with HiChIP promoter loops from the four endometrial cell lines. Through this analysis, we identified 103 HiChIP target genes (81 protein coding and 22 non-coding; **Table 3** and **Supplementary Table S2**) at 13 endometrial cancer GWAS loci. Ten of the non-coding HiChIP target genes encoded anti-sense transcripts (Supplementary Table 2), with these genes often sharing a promoter with a coding HiChIP target gene e.g. *CDKN2A* and *CDKN2B-AS1* (**Figure 1**).
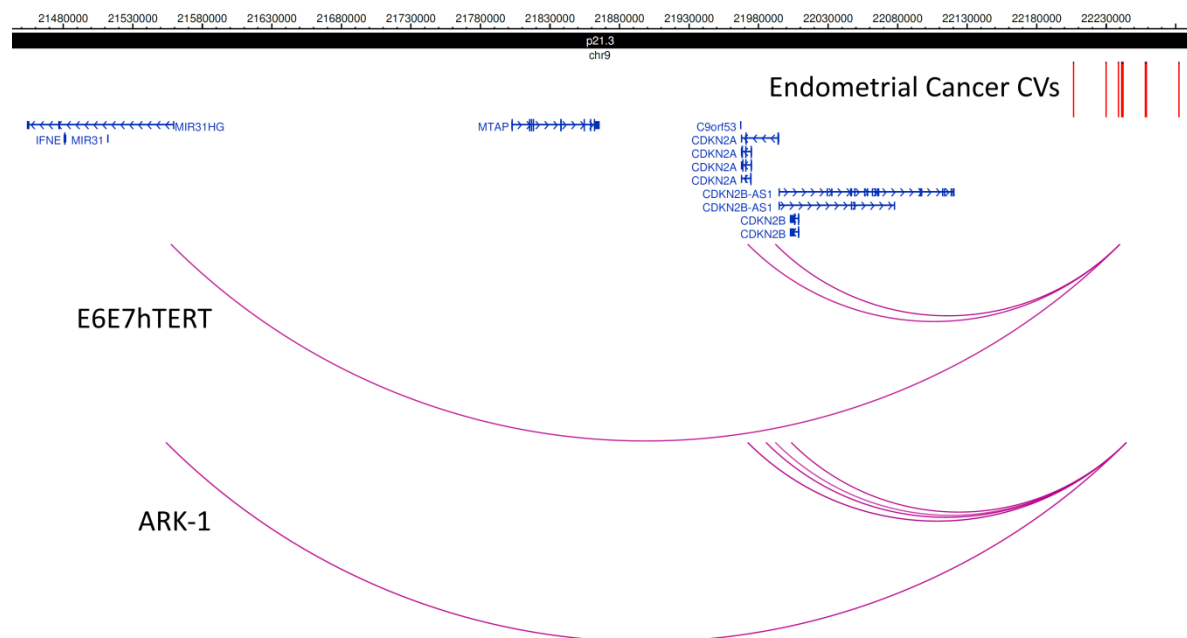

Figure 1. Promoter-associated chromatin looping identifies HiChIP target genes at the 9p21.3 locus. Promoter-associated loops were intersected with endometrial cancer risk CVs (coloured red), revealing loops (purple) that interact with the promoters of *MIR31HG* (in E6E7hTERT and ARK-1 cells), *CDKN2A* (in E6E7hTERT and ARK-1 cells), *CDNK2B-AS1* (in E6E7hTERT and ARK-1 cells) and *CDKN2B (*in ARK-1 cells*)*. *CDK2NA* and the non-coding anti-sense gene *CDKN2B-AS1* are encoded on opposite DNA strands and share a promoter.

The number of HiChIP target genes at a locus ranged from 1 to 38, with a median of 4. The HiChIP target genes included three genes (*WT1*, *WT1-AS* and *GNL2*) that had CVs located in a HiChIP looping contact at a promoter region, but for which there was no looping from an element containing a distal CV. These findings suggest that *WT1*, *WT1-AS* and *GNL2* may be directly regulated by a promoter CV in endometrial cells. In total, only 18 genes were the nearest gene to a CV (**Table 1**). More than one third (36%) of the HiChIP target genes were identified using looping data from at least two endometrial cell lines (underlined in **Table 3; Supplementary Table S2**), providing additional evidence for their targeting.

Table3. HiChIP target genes at endometrial cancer risk loci.

Risk locus	HiChIP target genes	Nearest gene(s) to CVs [1]
1p34.3	*GNL2*, *C1orf122*	*GNL2, RSPO1*
2p16.1	*BCL11A*	*BCL11A*
8q24.1	*MIR1207, PVT1, LINC00824*	*LINC00824*
9p21.3	*CDKN2A*, CDKN2B, *CDKN2B-AS1*, *MIR31HG*	*CDKN2B-AS1*
11p13	*WT1, WT1-AS, CD59, PAX6, RCN1*	*WT1-AS*
12p12.1	*BHLHE41*, PTHLH, SSPN, LRMP	*SSPN*
12q24.11	*SH2B3*, PHETA1, ACAD10, ARPC3, BRAP, IFT81, LINC02356	*SH2B3, ATXN2*
12q24.21	*TBX3*	*TBX3*
15q15.1	*SRP14*, *SRP14-AS1*, *BMF*, *BAHD1*, *CCDC9B*, *GPR176*, *KNSTRN*, *PAK6*, *PLCB2*, *PLCB2-AS1*, THBS1, EIF2AK4, CHST14, DISP2, FSIP1, *INAFM2*, PLA2G4B, RASGRP1, SPINT1, *ANKRD63*, PHGR1, SPINT1-AS1, *C15orf56*	*SRP14, SRP14-AS1, EIF2AK4*
15q21.2	*DMXL2, TRPM7, TNFAIP8L3*	*CYP19A1*
17q11.2	*RAB11FIP4*, *MIR193A*, TEFM, *RNU6ATAC7P*	*RAB11FIP4, NF1, EVI2A, EVI2B*
17q12	*HNF1B, DUSP14, MRM1, MRPL45, SRCIN1, TBC1D3, C17orf78*	*HNF1B*
17q21.32	*SNX11*, *MIR1203*, *SKAP1-AS1*, *SKAP1*, *CBX1*, HOXB1, HOXB2, *HOXB3*, *HOXB4*, HOXB5, *HOXB6*, *HOXB7*, *HOXB8*, *HOXB9*, HOXB13, HOXB-AS1, HOXB-AS3, *HOXB-AS4*, *PRR15L*, CDK5RAP3, LRRC46, MRPL10, NFE2L1, SCRN2, CALCOCO2, COPZ2, DLX3, KPNB1, PNPO, SNF8, SP2, SP2-AS1, SP6, *MIR10A*, MIR152, *MIR196A1*, MIR3185, PHOSPHO1	*SNX11, MIR1203, SKAP1-AS1, SKAP1, CBX1*

[1] At some loci, CVs are coincident with multiple genes

Underlined candidate target genes are supported by HiChIP data from multiple cell lines

2.4. HiChIP target genes are enriched for potential targets of a miRNA encoded by the HiChIP target gene MIR196A1

ToppFun bioinformatic analysis revealed that the predicted targets of 86 miRNAs were over-represented among the HiChIP target genes ($p_{Bonferroni}<0.05$, **Supplementary Table S3**). hsa-mir-196a-5p was one of these miRNAs and is encoded by *MIR196A1*, itself one of the seven miRNA genes among the HiChIP identified targets (**Table 3**). hsa-mir-196a-5p is predicted to bind to transcripts from six HiChIP target genes: four of which (*HOXB1, HOXB6, HOXB7* and *HOXB8*) are located at the same locus as *MIR196A1* (17q21.32), with the remaining two (*BRAP* and *RASGRP1*) encoded at other endometrial cancer risk loci (12q24.11 and 15q15.1, respectively).

2.5. HiChIP target genes are differentially expressed in endometrial tumors

Differential gene expression data derived from endometrial tumor and paired normal samples (n=13; The Cancer Genome Atlas Project (TCGA) [16]) were evaluated and 36 HiChIP target genes were found to be differentially expressed (**Supplementary Table S2**). Assessment of HiChIP target genes revealed a more than two fold over-representation of genes that were differentialy expressed in endometrial tumor (OR=2.39, 95% CI 1.59-3.59, $p=6.08×10^{-05}$).

2.6. HiChIP target gene expression associates with CVs

To aid prioritisation of HiChIP target genes, we interrogated expression quantitative trait locus (eQTL) data from the largest study of whole blood gene expression [17] and TCGA endometrial tumors [18]. Using these data, we evaluated the overlap between endometrial cancer risk CVs and the top eQTL variants for each HiChIP target gene. From the blood eQTL data, we found that the lead CV at the 17q21.32 risk locus, rs882380, was one of the top eQTLs for *SNX11* and *HOXB2*, and the lead CV at the 15q15.1 risk locus, rs937213, was one of the top eQTLs for *SRP14* (**Table 4** and

Supplementary Table 4). From the endometrial tumor eQTL data, we found that the top eQTL for the HiChIP target gene *BCL11A* was rs7579014, a CV at the 2p16.1 risk locus (**Table 4** and **Supplementary Table 4**).

2.7. Protein-protein interaction network of HiChIP target genes reveals enrichment for endometrial cancer driver genes

The HiChIP target genes included three known pan-cancer driver genes (*CDKN2A*, *TBX3* and *WT1*) identified by Bailey et al [19], but no known drivers of endometrial cancer from lists compiled by Bailey et al or Gibson et al [20]. Pathway analysis was performed by ToppFun using the 103 candidate target genes to gain biological insights but no pathways were found to be enriched after Bonferroni correction. To explore protein-protein interaction networks involving the candidate target genes, we used the ToppGenet bioinformatic tool. Mining of protein-protein interaction databases by ToppGenet revealed 2,135 proteins that interacted with those encoded by HiChIP target genes (**Supplementary Table S5**). Prioritisation was then performed by ToppGenet to identify those proteins with the most similar functional features to the HiChIP target gene set i.e. a "guilt by association" approach. Using this method, 387 of the interacting proteins had significant similarity scores at a 5% false discovery rate (FDR) (**Supplementary Table S5**). The protein with the most statistically significant similarity score was encoded by *TP53*, an endometrial and pan-cancer driver gene. Indeed, many other proteins encoded by known cancer driver genes were observed in the prioritised set of proteins. Of the 85 pan-cancer driver genes encoding interacting proteins, 55 encoded proteins in the prioritised set, a significant enrichment (OR=9.49, 95% CI 6.06-14.80; p<1×10^{-09}; **Supplementary Table S5**). The two available lists of endometrial cancer driver genes [19,20] were combined and of the 28 encoding interacting proteins, 19 encoded proteins in the prioritised set (**Table 4**; **Supplementary Table S5**), also a significant enrichment (OR=9.98, 95% CI 4.64-22.58; p=1.8×10^{-08}).

Table 4. Endometrial cancer drivers interacting with proteins encoded by HiChIP target genes.

Protein encoding gene	Similarity score	p-value	FDR [1] value
TP53	0.60	3.65E-09	4.00E-06
ESR1	0.54	5.95E-07	1.27E-04
FOXA2	0.57	1.86E-06	1.99E-04
EP300	0.41	8.40E-06	4.27E-04
CTNNB1	0.47	1.35E-05	5.54E-04
PTEN	0.46	1.98E-05	7.05E-04
CCND1	0.49	2.12E-05	7.42E-04
FGFR2	0.44	3.97E-05	1.10E-03
RB1	0.50	8.21E-05	1.91E-03
MYCN	0.44	1.15E-04	2.51E-03
ERBB2	0.39	4.15E-04	6.28E-03
AKT1	0.35	5.24E-04	7.31E-03
ERBB3	0.39	1.12E-03	0.01
MAX	0.31	1.75E-03	0.02
NRIP1	0.32	1.82E-03	0.02
ATM	0.31	2.05E-03	0.02
CHD4	0.34	2.67E-03	0.02
FBXW7	0.38	3.75E-03	0.03
DICER1	0.33	4.44E-03	0.03
KRAS	0.33	9.91E-03	0.05
TAF1	0.27	0.03	0.11
ATR	0.29	0.04	0.13
PIK3R2	0.19	0.06	0.17
POLE	0.26	0.07	0.17
CHD3	0.20	0.14	0.26
TAB3	0.22	0.36	0.45
METTL14	0.21	0.40	0.49
KANSL1	0.09	0.67	0.67

[1] False discovery rate (FDR)
Bolded proteins have a statistically significant similarity score (FDR < 0.05)

2.8. HiChIP target genes and interacting proteins are over-represented in relevant biological pathways

Pathway analysis using the combined list of 103 HiChIP target genes and 387 prioritised interacting proteins found 462 pathways to be significantly enriched after Bonferroni correction (**Supplementary Table S6**). Many of these pathways were related to gene regulation (e.g. "transcriptional misregulation in cancer") and cancer (e.g. "pathways in cancer"), including hallmarks of cancer identified by Hanahan and Weinberg [21] (**Table 5**). A KEGG "endometrial cancer" pathway and pathways related to endometrial cancer risk factors such as obesity (e.g. "signaling by leptin"), insulinemia (e.g. "insulin receptor signalling cascade") and estrogen exposure ("plasma membrane estrogen receptor signalling") were also found among the significantly enriched pathways.

Table 5. Examples of enriched pathways related to hallmarks of cancer.

Cancer hallmark	Related pathway (source)	$p_{Bonferroni}$
Evading growth suppressors	Regulation of TP53 activity (REACTOME)	1.44E-07
Avoiding immune destruction	Innate immune system (REACTOME)	2.06E-06
Enabling replicative immortality	Regulation of telomerase (Pathway Interaction Database)	1.43E-12
Tumor-promoting inflammation	Inflammation mediated by chemokine and cytokine signalling pathway (PantherDB)	0.03
Activating invasion and metastasis	Focal adhesion (KEGG)	5.61E-15
Inducing angiogenesis	VEGFA-VEGFR2 pathway (REACTOME)	9.10E-08
Genome instability and mutation	RB Suppressor/Checkpoint Signaling in response to DNA damage (MSigDB C2 BIOCARTA)	1.05E-04
Resisting cell death	Apoptosis signaling pathway (Panther DB)	1.78E-08
Deregulating cellular energetics	Choline metabolism in cancer (KEGG)	2.10E-04
Sustaining proliferative signalling	PI3K-Akt signalling pathway (KEGG)	1.15E-18

3. Discussion

We performed H3K27Ac HiChIP in endometrial cell lines to enrich for enhancer-promoter chromatin looping interactions and found more than a third of identified HiChIP chromatin loops interacted with a promoter. The anchors of promoter-associated loops from two endometrial cancer cell lines and an immortalised normal endometrial cell line were significantly enriched for endometrial cancer heritability, highlighting the potential importance of these loops in mediating the effects of endometrial cancer risk variation. Intersection of the promoter-associated loops with endometrial cancer GWAS CVs revealed 103 HiChIP target genes at 13 loci, 36% of which were identified from loops in multiple cell lines. At only two loci (2p16.1 and 12q24.21) was the nearest gene the only HiChIP target identified (*BCL11A* and *TBX3*, respectively). Similar to another HiChIP study that integrated chromatin interaction data with GWAS findings, the majority of HiChIP target genes (83%) involved a CV-promoter looping interaction that skipped the gene(s) closest to the CVs at that locus [12]. These findings further highlight the potential for long-range regulation and the pitfalls of mapping GWAS variation to the nearest gene (as we have previously discussed [22]). The median lengths of the promoter-associated loops in the four HiChIP cell lines were all greater than 200 kb, consistent with gene skipping by the putative enhancers at the endometrial cancer risk loci. Furthermore, the observed rate of gene skipping was similar that found by HiChIP analyses of other disease risk loci [12].

The HiChIP target genes were enriched for genes that are differentially expressed in endometrial s, providing evidence that these genes may be involved with endometrial cancer development and that the HiChIP data can help identify biologically relevant genes. Moreover, of the 25 candidate target genes previously identififed from endometrial cancer GWAS (reviewed in [4]), the targeting of ten genes (*MIR1207*, *WT1-AS*, *RCN1*, *SH2B3*, *BMF*, *GPR176*, *SRP14-AS1*, *SRP14*, *HNF1B* and *SNX11*) was supported by the HiChIP data. Anoher of the previously identified candidate target genes was *KLF5* at the 13q22.1 risk locus. Endometrial cancer GWAS risk variation has been found to loop to *KLF5* in data generated by a local 3C-based technique [2], an interaction which we did not observe in our HiChIP analyses. Our HiChIP approach had much higher resolution (due to the use of a restriction enzyme producing smaller fragments) and closer examination of the HiChIP data at 13q22.1 revealed a *KLF5* promoter interaction in JHUEM-14 cells that looped to an anchor 23 bp from endometrial cancer risk CV rs9600103.

Assessment of the other two loci without HiChIP target genes, 6q22.3 and 6q22.31, did not reveal other promoter loops in such close proximity to CVs. It is possible at these and other loci that additional genes may be targeted through chromatin looping occurring in cell types (or settings) not studied here, or through regulatory mechanisms not associated with H3K27Ac chromatin looping. Functional genomic approaches using alternative chromatin looping analyses or CRISPR genome/epigenome editing may thus prove useful in identifying further candidate target genes (particularly at 6q22.3, 6q22.31 and 13q22.1) and would aid prioritisation and validation of candidate target genes identified in this study.

To provide further evidence to support the HiChIP target genes, we integrated available eQTL data with endometrial cancer risk CVs and found that CVs associated with the expression of four HiChIP target genes. Three of these associations were observed in whole blood (*SNX11*, *HOXB2* and *SRP14*) and one in endometrial s (*BCL11A*). *SNX11* encodes a member of the sorting nexin family and is involved in endosomal intracellular trafficking [23] and may prevent degradation of its endosomal cargo [24]. *HOXB2* is a homeobox B gene and encodes a transcription factor that is involved in development in mice [25]. Expression of *HOXB2* (along with the HiChIP target genes *HOXB5* and *PTHLH*) has been found to be downregulated in a rare syndrome that is characterised by abnormal development of the uterus and vagina [26]. In our study, we found that reduced *HOXB2* expression in whole blood was associated with endometrial cancer risk variation, compatible with reports that *HOXB2* has a tumor suppressor function [27,28]. *SRP14* is involved in the formation of stress granules [29], which can also be initiated by phosphorylation of protein encoded by the *EIF2AK4* HiChIP target gene [30]. Consistent with our observation that endometrial cancer risk variation was associated with increased *SRP14* expression, stress granules promote cell survival and cancer cell fitness, and their components are upregulated in tumors (reviewed in [31]). Finally, *BCL11A* encodes a zinc finger protein transcription factor and plays an important role in lymphocyte development [32]. In cancer, the role of BCL11A appears to be context dependent. In some cancers, it has oncogenic effects [33,34], whereas in T cell leukaemia it may act as a tumor suppressor [32]. Further supporting a tumor suppressor role are findings that down-regulation of *BCL11A* increases the resistance of cancer cells to radiation [35] and loss of function of *BCL11A* is associated with genome instability in lung cancer [36]. Concordant with an anti-cancer function, we found that endometrial cancer risk variation was associated with decreased *BCL11A* expresssion in endometrial .

HiChIP target genes were enriched for miRNA targets of *MIR196A1*, itself a HiChIP target gene. *MIR196A1* miRNA was predicted to target six HiChIP target genes and has also been shown to regulate another HiChIP target *HOXB9* [37]. Expression of *MIR196A1* miRNA has been correlated with expression of the endometrial cancer candidate target gene *KLF5* in breast tumors, and has been associated with poor outcome in breast and ovarian cancer patients [38,39]. *MIR196A1* miRNA inhibits a range of cancer cell phenotypes including apoptosis, proliferation, migration and invasion [37,40-42]. Consistent with these findings, *MIR196A1* miRNA levels are lower in endometrial tumors compared with healthy endometrial tissue [42]. There may also be a link between *MIR196A1* and the endometrial cancer risk factors of obesity and insulinemia [43,44]: *MIR196A1* miRNA is upregulated in gluteofemoral fat, which is associated with lower risk of diabetes [45]; and forced expression of the mouse homologue of *MIR196A1* has been found to make mice resistant to obesity and prevent them from developing insulin resistance [46].

Lastly, bioinformatic analysis was used to functionally prioritise 387 proteins that interact with those encoded by the HiChIP target genes. This prioritised set had nearly a ten-fold over-representation of proteins encoded by pan-cancer or endometrial cancer driver genes compared to the remaining interacting proteins. Further analysis demonstrated that the combined set of prioritised proteins and candidate target genes were enriched in pathways relevant to hallmarks of cancer and endometrial cancer risk factors. Taken together, these observations suggest that proteins encoded by HiChIP target genes may mediate their effects through interactions with cancer drivers and other proteins that are involved in endometrial cancer development pathways.

4. Materials and Methods

4.1. Cell culture

Ishikawa, JHUEM-14, ARK-1 and E6E7hTERT cells were a gift from Prof PM Pollock (Queensland University of Technology). Cell lines were authenticated using STR profiling and confirmed to be negative for mycoplasma contamination. For routine culture, Ishikawa, ARK-1 and E6E7hTERT cells were grown in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum (FBS) and antibiotics (100 IU/ml penicillin and 100 μg/ml streptomycin). JHUEM-14 cells were cultured in DMEM/F12 medium with 10% FBS and antibiotics. All cell lines were cultured in a humidified incubator (37°C, 5% CO_2).

4.2. Cell fixation

For fixation, cells on 10-cm tissue culture plates (~80% confluence) were washed with PBS and fixed at room temperature in 1% formaldehyde in PBS. After 10 min, cells were placed on ice and the formaldehyde was quenched by washing twice with 125 mM glycine in PBS. Cells were removed from the dish with a cell scraper and washed with PBS before the storage of cell pellets at -80°C. As we had previously observed greater enrichment of endometrial cancer GWAS variation in epigenomic features observed after estrogen stimulation, including those found in Ishikawa and JHUEM-14 endometrioid endometrial cancer cell lines [1], these two cell lines were stimulated with 10 nM estradiol for 3 h prior to fixation, as previously [1]. Normal endometrium is considered estrogen responsive [47], so E6E7hTERT (a normal immortalised endometrial cell line) cells were also stimulated with estradiol. ARK-1 cells (derived from a serous endometrial tumor) were not stimulated with estradiol as serous endometrial tumors are not considered estrogen responsive [48].

4.3. HiChIP library generation

HiChIP libraries were generated as per the method of Mumbach et al. [14] with modifications. Briefly, cell nuclei were extracted from fixed cell pellets and digested overnight with 375U of DpnII to improve resolution [49]. After digestion, nuclei were resuspended in NEB buffer 2 (New England Biolabs) and restriction fragment overhangs were filled in with biotin-dATP using the DNA polymerase I, large Klenow fragment (incubated at 30°C for 2h). Proximity ligation was performed for 4 h at 16°C, then nuclei lysed and chromatin sheared for 9 min using the Covaris S220 Sonicator as per Mumbach et al. For each sample, sheared chromatin was split into two tubes and incubated overnight with 4.6 μg of H3K27Ac antibody (Abcam, EP16602). The next day Protein A beads were used to capture H3K27Ac-associated chromatin which was eluted and purified with Zymo Research concentrator columns (columns were washed twice with 10 μl of water). As per Mumbach et al, the DNA concentration of the purified chromatin was to estimate the amount of TDE1 enzyme (Illumina) needed for tagmentation which was performed with biotin-labelled chromatin captured on streptavidin beads. Sequencing libraries were then generated by PCR of tagmented samples using the Nextera DNA preparation kit (Illumina) as per the manufacturer's instructions. Afterwards, size selection was performed using Ampure XP beads to capture 300-700 bp fragments. For each cell line, at least two independent sequencing libraries were pooled together to provide 25 μl of library at ≥10 nM for Illumina HiSeq4000 (AGRF, Brisbane, QLD, Australia) paired-end sequencing.

4.4. HiChIP bioinformatic analyses

HiChIP reads (fastq files) were aligned to the human reference genome (hg19) using HiC-Pro version 2.9.0 [50] and default settings used to remove duplicate reads, assign reads to DpnII restriction fragments and filter for valid interactions. The hichipper pipeline version 0.7.0 [51] was used to process all valid reads from HiC-Pro, with the HiChIP reads used to identify H3K27Ac peaks using the standard MACS2 background model. Chromatin interactions were filtered using a minimum distance of 5 kb and a maximum of 2 Mb. The final set of chromatin loops used for further investigation were interactions which were supported by a minimum of two unique paired end tags and with a Mango [52] q-value < 5%.

4.5. Stratified LD score regression analysis

We used stratified LD score regression [53,54] to quantify enrichment of endometrial cancer risk variation in HiChIP promoter-associated loops. Stratified LD score regression calculates enrichment as the proportion of genetic heritability attributable to a particular set of variants (e.g. variants located within HiChIP promoter-associated loops) divided by the proportion of total genetic variants annotated to that set. Enrichment for each cell line HiChIP promoter-associated loops annotation categories were assessed individually conditional on a "full baseline model" of 53 overlapping categories as used previously (https://data.broadinstitute.org/alkesgroup/LDSCORE/) [54]. For regression, variants were pruned to the HapMap3 variant list (~1 million variants) and the 1000 Genomes Project Phase 3 European population variants were used for the LD reference panel. The major histocompatibility complex (MHC) region was removed from this analysis because of its complex LD structure.

4.6. Identification and analysis of HiChIP target genes

Promoter-associated loops were defined as HiChIP loops with an anchor within 3 kb of a transcription start site (GRCh37; accessed May 2019). To identify candidate target genes, HiChIP promoter loops were intersected with endometrial cancer risk CVs (n=457) which had been determined using a 100:1 log likelihood ratio with the p-value for the lead variant at each GWAS locus. Differential gene expression from TCGA endometrial tumors was obtained from GEPIA2 [55] using limma analysed data with a $\log_2$ fold cut-off of 1 and $q<0.01$ for statistical signifcance. All bioinformatic analysis of HiChIP target genes was performed using the ToppGene Suite of tools (accessed June 7 2019) [56]. ToppFun was used to detect enrichment of gene lists based on miRNA binding sites and pathways using hypergeometric distribution analysis. ToppGenet was used to identify and prioritise genes in protein-protein interaction networks based on functional similarity to the HiChIP target genes. Analyses to identify over-representation of genes in different sets was performed using Fisher's exact test in GraphPad Prism 8.1.2.

5. Conclusions

Here we present the first global study of chromatin looping in endometrial cell lines, using an H3K27Ac HiChIP approach to enrich for enhancer-promoter interactions. These data will provide an extremely useful resource for genetic studies of not only endometrial cancer but also other diseases that involve the endometrium. Through these data, we have found that promoter-associated HiChIP loops are significantly enriched for endometrial cancer heritability and used these loops to identify a set of candidate target genes at endometrial cancer GWAS loci, which contains an over-representation of genes differentially regulated in endometrial tumors. Integration of eQTL data provided evidence to prioritise candidates for functional studies and further supports the hypothesis that endometrial cancer GWAS variation regulates gene expression through long-range regulatory interactions. Previous reports from the literatures suggests there is interplay among the products of HiChIP target genes and that proteins encoded by HiChIP target genes interact with cancer drivers. Finally, bioinformatic analysis provides data demonstrating that the HiChIP target genes and their interacting protein-protein networks belong to pathways that are relevant to endometrial cancer development..

In summary, this study has identified candidate endometrial cancer GWAS target genes for future studies and furthers our understanding of the genetic basis of endometrial cancer development.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, File S1: Ishikawa promoter-associated loops, File S2: E6E7hTERT promoter-associated loops, File S3: ARK-1 promoter-associated loops, File S4: JHUEM-14 promoter-associated loops, File S5: ECAC acknowledgements, Table S1: HiC-Pro QC results, Table S2: HiChIP target genes, Table S3: miRNAs with an enrichment of putative targets among HiChIP target genes, Table S4: Endometrial cancer risk CVs that associate with gene expression, Table S5: Similarity of interacting proteins to HiChIP target genes, Table S6: Pathways enriched for HiChIP target genes

and interacting proteins. Raw and processed sequencing data from HiChIP experiments will be submitted to the Gene Expression Omnibus (GEO) data repository to allow access to the full data from these studies.

References

1.   O'Mara, T.A.; Glubb, D.M.; Amant, F.; Annibali, D.; Ashton, K.; Attia, J.; Auer, P.L.; Beckmann, M.W.; Black, A.; Bolla, M.K., et al. Identification of nine new susceptibility loci for endometrial cancer. *Nat Commun* **2018**, *9*, 3166, doi:10.1038/s41467-018-05427-7.

2.   Cheng, T.H.; Thompson, D.J.; O'Mara, T.A.; Painter, J.N.; Glubb, D.M.; Flach, S.; Lewis, A.; French, J.D.; Freeman-Mills, L.; Church, D., et al. Five endometrial cancer risk loci identified through genome-wide association analysis. *Nat Genet* **2016**, *48*, 667-674, doi:10.1038/ng.3562.

3.   Spurdle, A.B.; Thompson, D.J.; Ahmed, S.; Ferguson, K.; Healey, C.S.; O'Mara, T.; Walker, L.C.; Montgomery, S.B.; Dermitzakis, E.T.; Fahey, P., et al. Genome-wide association study identifies a common variant associated with risk of endometrial cancer. *Nat Genet* **2011**, *43*, 451-454, doi:10.1038/ng.812.

4.   O'Mara, T.A.; Glubb, D.M.; Kho, P.F.; Thompson, D.J.; Spurdle, A. Genome-wide association studies of endometrial cancer: Latest developments and future directions. *Cancer Epidemiol Biomarkers Prev* **2019**, *28*, 1095-1102.

5.   Thompson, D.J.; O'Mara, T.A.; Glubb, D.M.; Painter, J.N.; Cheng, T.; Folkerd, E.; Doody, D.; Dennis, J.; Webb, P.M.; Group, A.N.E.C.S., et al. CYP19A1 fine-mapping and Mendelian randomization: estradiol is causal for endometrial cancer. *Endocr Relat Cancer* **2016**, *23*, 77-91, doi:10.1530/erc-15-0386.

6.   Key, T.J.; Pike, M.C. The dose-effect relationship between 'unopposed' oestrogens and endometrial mitotic rate: its central role in explaining and predicting endometrial cancer risk. *Br J Cancer* **1988**, *57*, 205-212.

7.   Antunes, C.M.; Strolley, P.D.; Rosenshein, N.B.; Davies, J.L.; Tonascia, J.A.; Brown, C.; Burnett, L.; Rutledge, A.; Pokempner, M.; Garcia, R. Endometrial cancer and estrogen use. Report of a large case-control study. *N Engl J Med* **1979**, *300*, 9-13, doi:10.1056/nejm197901043000103.

8.   Glubb, D.M.; Johnatty, S.E.; Quinn, M.C.J.; O'Mara, T.A.; Tyrer, J.P.; Gao, B.; Fasching, P.A.; Beckmann, M.W.; Lambrechts, D.; Vergote, I., et al. Analyses of germline variants associated with ovarian cancer survival identify functional candidates at the 1q22 and 19p12 outcome loci. *Oncotarget* **2017**, 10.18632/oncotarget.18501, doi:10.18632/oncotarget.18501.

9.   Glubb, D.M.; Maranian, M.J.; Michailidou, K.; Pooley, K.A.; Meyer, K.B.; Kar, S.; Carlebur, S.; O'Reilly, M.; Betts, J.A.; Hillman, K.M., et al. Fine-Scale Mapping of the 5q11.2 Breast Cancer Locus Reveals at Least Three Independent Risk Variants Regulating MAP3K1. *Am J Hum Genet* **2015**, *96*, 5-20.

10.   Michailidou, K.; Lindstrom, S.; Dennis, J.; Beesley, J.; Hui, S.; Kar, S.; Lemacon, A.; Soucy, P.; Glubb, D.; Rostamianfar, A., et al. Association analysis identifies 65 new breast cancer risk loci. *Nature* **2017**, *551*, 92-94, doi:10.1038/nature24284.

11.   Gallagher, M.D.; Chen-Plotkin, A.S. The Post-GWAS Era: From Association to Function. *Am J Hum Genet* **2018**, *102*, 717-730, doi:10.1016/j.ajhg.2018.04.002.

12.   Mumbach, M.R.; Satpathy, A.T.; Boyle, E.A.; Dai, C.; Gowen, B.G.; Cho, S.W.; Nguyen, M.L.; Rubin, A.J.; Granja, J.M.; Kazane, K.R., et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat Genet* **2017**, *49*, 1602-1612, doi:10.1038/ng.3963.

13.   Davies, J.O.; Oudelaar, A.M.; Higgs, D.R.; Hughes, J.R. How best to identify chromosomal interactions: a comparison of approaches. *Nat Methods* **2017**, *14*, 125-134, doi:10.1038/nmeth.4146.

14.   Mumbach, M.R.; Rubin, A.J.; Flynn, R.A.; Dai, C.; Khavari, P.A.; Greenleaf, W.J.; Chang, H.Y. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods* **2016**, *13*, 919-922, doi:10.1038/nmeth.3999.

15.   Jeng, M.Y.; Mumbach, M.R.; Granja, J.M.; Satpathy, A.T.; Chang, H.Y.; Chang, A.L.S. Enhancer Connectome Nominates Target Genes of Inherited Risk Variants from Inflammatory Skin Disorders. *J Invest Dermatol* **2019**, *139*, 605-614, doi:10.1016/j.jid.2018.09.011.

16.   Kandoth, C.; Schultz, N.; Cherniack, A.D.; Akbani, R.; Liu, Y.; Shen, H.; Robertson, A.G.; Pashtan, I.; Shen, R.; Benz, C.C., et al. Integrated genomic characterization of endometrial carcinoma. *Nature* **2013**, *497*, 67-73, doi:10.1038/nature12113.

17.   Võsa, U.; Claringbould, A.; Westra, H.-J.; Bonder, M.J.; Deelen, P.; Zeng, B.; Kirsten, H.; Saha, A.; Kreuzhuber, R.; Kasela, S., et al. Unraveling the polygenic architecture of complex traits using blood eQTL meta-analysis. *bioRxiv* **2018**, 10.1101/447367, 447367, doi:10.1101/447367.

18.   Lim, Y.W.; Chen-Harris, H.; Mayba, O.; Lianoglou, S.; Wuster, A.; Bhangale, T.; Khan, Z.; Mariathasan, S.; Daemen, A.; Reeder, J., et al. Germline genetic polymorphisms influence gene expression and immune cell infiltration. *Proc Natl Acad Sci U S A* **2018**, *115*, E11701-e11710, doi:10.1073/pnas.1804506115.

19.   Bailey, M.H.; Tokheim, C.; Porta-Pardo, E.; Sengupta, S.; Bertrand, D.; Weerasinghe, A.; Colaprico, A.; Wendl, M.C.; Kim, J.; Reardon, B., et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **2018**, *173*, 371-385.e318, doi:https://doi.org/10.1016/j.cell.2018.02.060.

20.   Gibson, W.J.; Hoivik, E.A.; Halle, M.K.; Taylor-Weiner, A.; Cherniack, A.D.; Berg, A.; Holst, F.; Zack, T.I.; Werner, H.M.; Staby, K.M., et al. The genomic landscape and evolution of endometrial carcinoma progression and abdominopelvic metastasis. *Nat Genet* **2016**, *48*, 848-855, doi:10.1038/ng.3602.

21.   Hanahan, D.; Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **2011**, *144*, 646-674, doi:10.1016/j.cell.2011.02.013.

22.   Pritchard, J.E.; O'Mara, T.A.; Glubb, D.M. Enhancing the Promise of Drug Repositioning through Genetics. *Front Pharmacol* **2017**, *8*, 896, doi:10.3389/fphar.2017.00896.

23.   Liu, T.; Li, J.; Liu, Y.; Qu, Y.; Li, A.; Li, C.; Zhang, Q.; Wu, W.; Li, J.; Liu, Y., et al. SNX11 Identified as an Essential Host Factor for SFTS Virus Infection by CRISPR Knockout Screening. *Virologica Sinica* **2019**, 10.1007/s12250-019-00141-0, doi:10.1007/s12250-019-00141-0.

24. Joyal, J.S.; Nim, S.; Zhu, T.; Sitaras, N.; Rivera, J.C.; Shao, Z.; Sapieha, P.; Hamel, D.; Sanchez, M.; Zaniolo, K., et al. Subcellular localization of coagulation factor II receptor-like 1 in neurons governs angiogenesis. *Nat Med* **2014**, *20*, 1165-1173, doi:10.1038/nm.3669.

25. Barrow, J.R.; Capecchi, M.R. Targeted disruption of the Hoxb-2 locus in mice interferes with expression of Hoxb-1 and Hoxb-4. *Development* **1996**, *122*, 3817-3828.

26. Nodale, C.; Ceccarelli, S.; Giuliano, M.; Cammarota, M.; D'Amici, S.; Vescarelli, E.; Maffucci, D.; Bellati, F.; Panici, P.B.; Romano, F., et al. Gene Expression Profile of Patients with Mayer-Rokitansky-Küster-Hauser Syndrome: New Insights into the Potential Role of Developmental Pathways. *PLOS ONE* **2014**, *9*, e91010, doi:10.1371/journal.pone.0091010.

27. Lindblad, O.; Chougule, R.A.; Moharram, S.A.; Kabir, N.N.; Sun, J.; Kazi, J.U.; Ronnstrand, L. The role of HOXB2 and HOXB3 in acute myeloid leukemia. *Biochem Biophys Res Commun* **2015**, *467*, 742-747, doi:10.1016/j.bbrc.2015.10.071.

28. Boimel, P.J.; Cruz, C.; Segall, J.E. A functional in vivo screen for regulators of progression identifies HOXB2 as a regulator of growth in breast cancer. *Genomics* **2011**, *98*, 164-172, doi:10.1016/j.ygeno.2011.05.011.

29. Berger, A.; Ivanova, E.; Gareau, C.; Scherrer, A.; Mazroui, R.; Strub, K. Direct binding of the Alu binding protein dimer SRP9/14 to 40S ribosomal subunits promotes stress granule formation and is regulated by Alu RNA. *Nucleic Acids Research* **2014**, *42*, 11203-11217, doi:10.1093/nar/gku822.

30. Donnelly, N.; Gorman, A.M.; Gupta, S.; Samali, A. The eIF2α kinases: their structures and functions. *Cellular and Molecular Life Sciences* **2013**, *70*, 3493-3511, doi:10.1007/s00018-012-1252-6.

31. El-Naggar, A.M.; Sorensen, P.H. Translational control of aberrant stress responses as a hallmark of cancer. *J Pathol* **2018**, *244*, 650-666, doi:10.1002/path.5030.

32. Liu, P.; Keller, J.R.; Ortiz, M.; Tessarollo, L.; Rachel, R.A.; Nakamura, T.; Jenkins, N.A.; Copeland, N.G. Bcl11a is essential for normal lymphoid development. *Nat Immunol* **2003**, *4*, 525-532, doi:10.1038/ni925.

33. Khaled, W.T.; Choon Lee, S.; Stingl, J.; Chen, X.; Raza Ali, H.; Rueda, O.M.; Hadi, F.; Wang, J.; Yu, Y.; Chin, S.F., et al. BCL11A is a triple-negative breast cancer gene with critical functions in stem and progenitor cells. *Nat Commun* **2015**, *6*, 5987, doi:10.1038/ncomms6987.

34. Lazarus, K.A.; Hadi, F.; Zambon, E.; Bach, K.; Santolla, M.F.; Watson, J.K.; Correia, L.L.; Das, M.; Ugur, R.; Pensa, S., et al. BCL11A interacts with SOX2 to control the expression of epigenetic regulators in lung squamous carcinoma. *Nat Commun* **2018**, *9*, 3327, doi:10.1038/s41467-018-05790-5.

35. Park, S.Y.; Lee, S.-J.; Cho, H.J.; Kim, J.-T.; Yoon, H.R.; Lee, K.H.; Kim, B.Y.; Lee, Y.; Lee, H.G. Epsilon-Globin HBE1 Enhances Radiotherapy Resistance by Down-Regulating BCL11A in Colorectal Cancer Cells. *Cancers* **2019**, *11*, 498.

36. Huang, H.T.; Chen, S.M.; Pan, L.B.; Yao, J.; Ma, H.T. Loss of function of SWI/SNF chromatin remodeling genes leads to genome instability of human lung cancer. *Oncol Rep* **2015**, *33*, 283-291, doi:10.3892/or.2014.3584.

37. Darda, L.; Hakami, F.; Morgan, R.; Murdoch, C.; Lambert, D.W.; Hunter, K.D. The role of HOXB9 and miR-196a in head and neck squamous cell carcinoma. *PLoS One* **2015**, *10*, e0122285, doi:10.1371/journal.pone.0122285.

38. Milevskiy, M.J.G.; Gujral, U.; Del Lama Marques, C.; Stone, A.; Northwood, K.; Burke, L.J.; Gee, J.M.W.; Nephew, K.; Clark, S.; Brown, M.A. MicroRNA-196a is regulated by ER and is a prognostic biomarker in ER+ breast cancer. *Br J Cancer* **2019**, *120*, 621-632, doi:10.1038/s41416-019-0395-8.

39.   Meng, X.; Jin-Cheng, G.; Jue, Z.; Quan-Fu, M.; Bin, Y.; Xu-Feng, W. Protein-coding genes, long non-coding RNAs combined with microRNAs as a novel clinical multi-dimension transcriptome signature to predict prognosis in ovarian cancer. *Oncotarget* **2017**, *8*, 72847-72859, doi:10.18632/oncotarget.20457.

40.   Hao, Y.; Wang, J.; Zhao, L. The effect and mechanism of miR196a in HepG2 cell. *Biomed Pharmacother* **2015**, *72*, 1-5, doi:10.1016/j.biopha.2014.10.032.

41.   Xu, H.; Li, G.; Yue, Z.; Li, C. HCV core protein-induced upregulation of microRNA-196a promotes aberrant proliferation in hepatocellular carcinoma by targeting FOXO1. *Molecular medicine reports* **2016**, *13*, 5223-5229, doi:10.3892/mmr.2016.5159.

42.   Chen, H.; Fan, Y.; Xu, W.; Chen, J.; Meng, Y.; Fang, D.; Wang, J. Exploration of miR-1202 and miR-196a in human endometrial cancer based on high throughout gene screening analysis. *Oncol Rep* **2017**, *37*, 3493-3501, doi:10.3892/or.2017.5596.

43.   Nead, K.T.; Sharp, S.J.; Thompson, D.J.; Painter, J.N.; Savage, D.B.; Semple, R.K.; Barker, A.; Perry, J.R.; Attia, J.; Dunning, A.M., et al. Evidence of a Causal Association Between Insulinemia and Endometrial Cancer: A Mendelian Randomization Analysis. *J Natl Cancer Inst* **2015**, *107*, doi:10.1093/jnci/djv178.

44.   Painter, J.N.; O'Mara, T.A.; Marquart, L.; Webb, P.M.; Attia, J.; Medland, S.E.; Cheng, T.; Dennis, J.; Holliday, E.G.; McEvoy, M., et al. Genetic Risk Score Mendelian Randomization Shows that Obesity Measured as Body Mass Index, but not Waist:Hip Ratio, Is Causal for Endometrial Cancer. *Cancer Epidemiol Biomarkers Prev* **2016**, *25*, 1503-1510, doi:10.1158/1055-9965.epi-16-0147.

45.   Lotta, L.A.; Wittemans, L.B.L.; Zuber, V.; Stewart, I.D.; Sharp, S.J.; Luan, J.; Day, F.R.; Li, C.; Bowker, N.; Cai, L., et al. Association of Genetic Variants Related to Gluteofemoral vs Abdominal Fat Distribution With Type 2 Diabetes, Coronary Disease, and Cardiovascular Risk Factors. *Jama* **2018**, *320*, 2553-2563, doi:10.1001/jama.2018.19329.

46.   Mori, M.; Nakagami, H.; Rodriguez-Araujo, G.; Nimura, K.; Kaneda, Y. Essential role for miR-196a in brown adipogenesis of white fat progenitor cells. *PLoS Biol* **2012**, *10*, e1001314, doi:10.1371/journal.pbio.1001314.

47.   Hewitt, S.C.; Deroo, B.J.; Hansen, K.; Collins, J.; Grissom, S.; Afshari, C.A.; Korach, K.S. Estrogen receptor-dependent genomic responses in the uterus mirror the biphasic physiological response to estrogen. *Mol Endocrinol* **2003**, *17*, 2070-2083, doi:10.1210/me.2003-0146.

48.   Gründker, C.; Günthert, A.R.; Emons, G. Hormonal Heterogeneity of Endometrial Cancer. In *Innovative Endocrinology of Cancer*, Berstein, L.M., Santen, R.J., Eds. Springer New York: New York, NY, 2008; 10.1007/978-0-387-78818-0_11pp. 166-188.

49.   Belaghzal, H.; Dekker, J.; Gibcus, J.H. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods* **2017**, *123*, 56-65, doi:10.1016/j.ymeth.2017.04.004.

50.   Servant, N.; Varoquaux, N.; Lajoie, B.R.; Viara, E.; Chen, C.-J.; Vert, J.-P.; Heard, E.; Dekker, J.; Barillot, E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biology* **2015**, *16*, 259, doi:10.1186/s13059-015-0831-x.

51.   Lareau, C.A.; Aryee, M.J. hichipper: a preprocessing pipeline for calling DNA loops from HiChIP data. *Nat Methods* **2018**, *15*, 155-156, doi:10.1038/nmeth.4583.

52.   Phanstiel, D.H.; Boyle, A.P.; Heidari, N.; Snyder, M.P. Mango: a bias-correcting ChIA-PET analysis pipeline. *Bioinformatics* **2015**, *31*, 3092-3098, doi:10.1093/bioinformatics/btv336.

53.   Bulik-Sullivan, B.K.; Loh, P.R.; Finucane, H.K.; Ripke, S.; Yang, J.; Patterson, N.; Daly, M.J.; Price, A.L.; Neale, B.M. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **2015**, *47*, 291-295, doi:10.1038/ng.3211.

54.   Finucane, H.K.; Bulik-Sullivan, B.; Gusev, A.; Trynka, G.; Reshef, Y.; Loh, P.R.; Anttila, V.; Xu, H.; Zang, C.; Farh, K., et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **2015**, *47*, 1228-1235, doi:10.1038/ng.3404.

55.   Tang, Z.; Kang, B.; Li, C.; Chen, T.; Zhang, Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Research* **2019**, *47*, W556-W560, doi:10.1093/nar/gkz430.

56.   Chen, J.; Bardes, E.E.; Aronow, B.J.; Jegga, A.G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res* **2009**, *37*, W305-311, doi:10.1093/nar/gkp427.