

Genomic landscape of endometrial stromal sarcoma of uterus

Youn Jin Choi^{1,*}, Seung-Hyun Jung^{2,3,*}, Min Sung Kim¹, In-Pyo Baek^{2,3}, Jae-Keun Rhee⁴, Sung Hak Lee⁵, Soo Young Hur⁷, Tae-Min Kim⁴, Yeun-Jun Chung^{2,3,6}, Sug Hyung Lee^{1,2}

¹Department of Pathology, College of Medicine, The Catholic University of Korea, Seoul 137-701, Republic of Korea

²Department of Cancer Evolution Research Center, College of Medicine, The Catholic University of Korea, Seoul 137-701, Republic of Korea

³Department of Integrated Research Center for Genome Polymorphism, College of Medicine, The Catholic University of Korea, Seoul 137-701, Republic of Korea

⁴Department of Medical Informatics, College of Medicine, The Catholic University of Korea, Seoul 137-701, Republic of Korea

⁵Department of Hospital Pathology, College of Medicine, The Catholic University of Korea, Seoul 137-701, Republic of Korea

⁶Department of Microbiology, College of Medicine, The Catholic University of Korea, Seoul 137-701, Republic of Korea

⁷Department of Obstetrics/Gynecology, College of Medicine, The Catholic University of Korea, Seoul 137-701, Republic of Korea

*These authors have contributed equally to this work

Correspondence to:

Yeun-Jun Chung, **e-mail:** yejun@catholic.ac.kr

Sug Hyung Lee, **e-mail:** suhulee@catholic.ac.kr

Keywords: endometrial stromal sarcoma, mutation, genome, whole exome, copy number

Received: July 07, 2015

Accepted: September 18, 2015

Published: September 30, 2015

ABSTRACT

Although recurrent gene fusions such as *JAZF1-JJAZ1* are considered driver events for endometrial stromal sarcoma (ESS) development, other genomic alterations remain largely unknown. In this study, we performed whole-exome sequencing, transcriptome sequencing and copy number profiling for five ESSs (three low-grade ESS (LG-ESS) and two undifferentiated uterine sarcomas (UUSs)). All three LG-ESSs exhibited either one of *JAZF1-SUZ12*, *JAZF1-PHF1* and *MEAF6-PHF1* fusions, whereas the two UUSs did not. All ESSs except one LG-ESS exhibited copy number alterations (CNAs), many of which encompassed cancer-related genes. In UUSs, five CNAs encompassing cancer-related genes (*EZR*, *CDH1*, *RB1*, *TP53* and *PRKAR1A*) accompanied their expressional changes, suggesting that they might stimulate UUS development. We found 81 non-silent mutations (35 from LG-ESSs and 46 from UUSs) that included 15 putative cancer genes catalogued in cancer-related databases, including *PPARG* and *IRF4* mutations. However, they were non-recurrent and did not include any well-known mutations, indicating that point mutations may not be a major driver for ESS development. Our data show that gene fusions and CNAs are the principal drivers for LG-ESS and UUS, respectively, but both may require additional genomic alterations including point mutations. These differences may explain the different biologic behaviors between LG-ESS and UUS. Our findings suggest that ESS development requires point mutations and CNAs as well as the gene fusions.

INTRODUCTION

Endometrial stromal sarcoma (ESS) is a malignant tumor arising from the stroma of endometrium and accounts for approximately 10% of uterine sarcomas [1]. The current World Health Organization recognizes four

categories of the ESS: endometrial stromal nodule (ESN), low-grade endometrial stromal sarcoma (LG-ESS), high-grade endometrial stromal sarcoma (HG-ESS) and undifferentiated uterine sarcoma (UUS) [2]. Both HG-ESS and UUS behave more aggressively than LG-ESS and they are categorized by their molecular features [3]. HG-ESS is

defined as an ESS with t(10;17)(q22;p13) rearrangement leading to *YWHAE-FAM22* gene fusion, while UUS is a ‘wastebasket category’ that does not harbor any specific chromosomal translocation [4]. A recurrent gene fusion (*JAZF1-JJAZ1*, t(7;17)(p15;q21)) [5–8] and other less common gene fusions (e.g., *PHF1-JAZF1*, *PHF1-EPC1*, *MEAF6-PHF1*, *ZC3H7-BCOR* and *MBTD1-CXorf67*) have been reported in LG-ESS [9–13]. Although the gene rearrangements are main mechanisms for ESS development, it is possible that somatic mutations might stimulate the development as well. Somatic mutation is a major driving force for the development of most tumors, but there has been no report on somatic mutation status in ESSs at whole-exome or whole-genome level. Also, there are only a few analyses for copy number alterations (CNA) and gene-expression profiling for ESSs [14–16].

To further characterize ESS genomes and extend the knowledge on genetic mechanisms for ESS development, the following questions were investigated in this study: (i) whether LG-ESS, HG-ESS and UUS genomes have driver mutations or CNAs; (ii) whether there are any previously uncharacterized gene fusions in either LG-ESS or HG-ESS or UUS; and (iii) whether there are any differences in expression profiling among LG-ESS, HG-ESS and UUS. For these, we analyzed genomes of LG-ESSs and UUSs (Table 1) by whole-exome sequencing (WES), whole-transcriptome sequencing and microarray-comparative genomic hybridization (a-CGH) in this study.

RESULTS

Copy number alterations and gene expression profiles

A total of 40 CNAs (12 gains and 28 losses) were identified in the five ESSs (LG-ESSs: cases 1–3 and UUSs: cases 4 and 5) by a-CGH (Supplementary Table S1). Of note, one LG-ESS (case 1) did not harbor any CNAs while the other four ESSs harbored CNAs (Figure 1A). To address whether the mutations found in our study could be causally implicated in ESS development, we queried the cancer Gene Census, the CHASM analysis and the pan-cancer databases [17–19]. All four ESS genomes with CNAs harbored one or more CNAs encompassing

the cancer-related genes (Table 2). LG-ESSs (cases 2 and 3) harbored CNA losses with the cancer-related genes *PMS2* and *PTEN*, while UUSs (cases 4 and 5) harbored CNA gains of *POU5F1*, *EZR*, *GNBQ* and *SYK*, and CNA losses of *ARID1A*, *DMD*, *RBI*, *DICER1*, *CYLD*, *CDH1*, *TP53* and *PRKARIA*. We also observed recurrent CNAs ($n \geq 2$) on 11q21 (CNA loss) and 16q11.2 - q24.3 (CNA loss) (Figure 1B). Of note, all the UUSs harbored CNA loss on 16q, where tumor suppressor genes *CYLD* and *CDH1* reside. In addition, we found five genes with CNA losses (*DMD*, *ARID1A*, *RBI*, *RAD51B* and *PTEN*) that overlapped pan-cancer CNA drivers (Figure 1C) [19].

We analyzed gene expression profiles of the five ESSs using the whole-transcription data, which were subsequently compared to their CNA profiles. A normal (non-malignant) endometrial tissue was used to compare the gene expression level with ESS tissues. Five cancer-related genes in the UUSs exhibited a positive correlation between CNAs and gene expressions (down-regulation of genes with CNA loss: *CDH1*, *RBI*, *TP53* and *PRKARIA*; up-regulation of a gene with CNA gain: *EZR*) (Figure 2). Compared to the normal endometrium, the ESSs showed increase of *EZR* expression (5.63 fold-change) and decreases of *RBI* (0.15 fold-change), *CDH1* (0.15 and 0.13 fold-changes in case 4 and case 5, respectively), *TP53* (0.04 fold-change) and *PRKARIA* (0.11 fold-change) expressions. All of the copy number losses except *RBI* were heterozygous deletion while *RBI* loss was homozygous deletion (Figure 2).

Gene fusions

In this study, a total of 536 putative gene fusions (mean, 107; range, 77–150) were identified by whole-transcriptome sequencing (Supplementary Table S2) from the five ESS samples. Of them, we focused on the previously reported genes with translocations based on the Mitelman Database of Chromosome Aberrations (<http://cgap.nci.nih.gov/Chromosomes/Mitelman>) and found three gene fusions (*JAZF1-SUZ12*, *JAZF1-PHF1* and *MEAF6-PHF1*). All of them, however, were well-known fusions in ESS. These fusions were detected in the LG-ESSs (cases 1–3), but not in the UUSs (Table 3). The fusions were subsequently validated by reverse transcription (RT)-polymerase chain reaction (PCR) and Sanger sequencing (Supplementary Figure S1).

Table 1: Clinical and histologic characteristics of five endometrial stromal sarcomas

	Age	Histopathology grade	Diagnosis	Specimen status	TNM
Case 1	32	Low-grade	LG-ESS	Primary	T1bN0M0
Case 2	34	Low-grade	LG-ESS	Primary	T1bN1M0
Case 3	57	Low-grade	LG-ESS	Metastatic	TxNxM1
Case 4	65	High-grade	UUS	Primary	T2aN0M0
Case 5	57	High-grade	UUS	Metastatic	TxNxM1

LG-ESS: low-grade ESS, UUS: undifferentiated uterine sarcoma, TNM: tumor, lymph node and metastasis

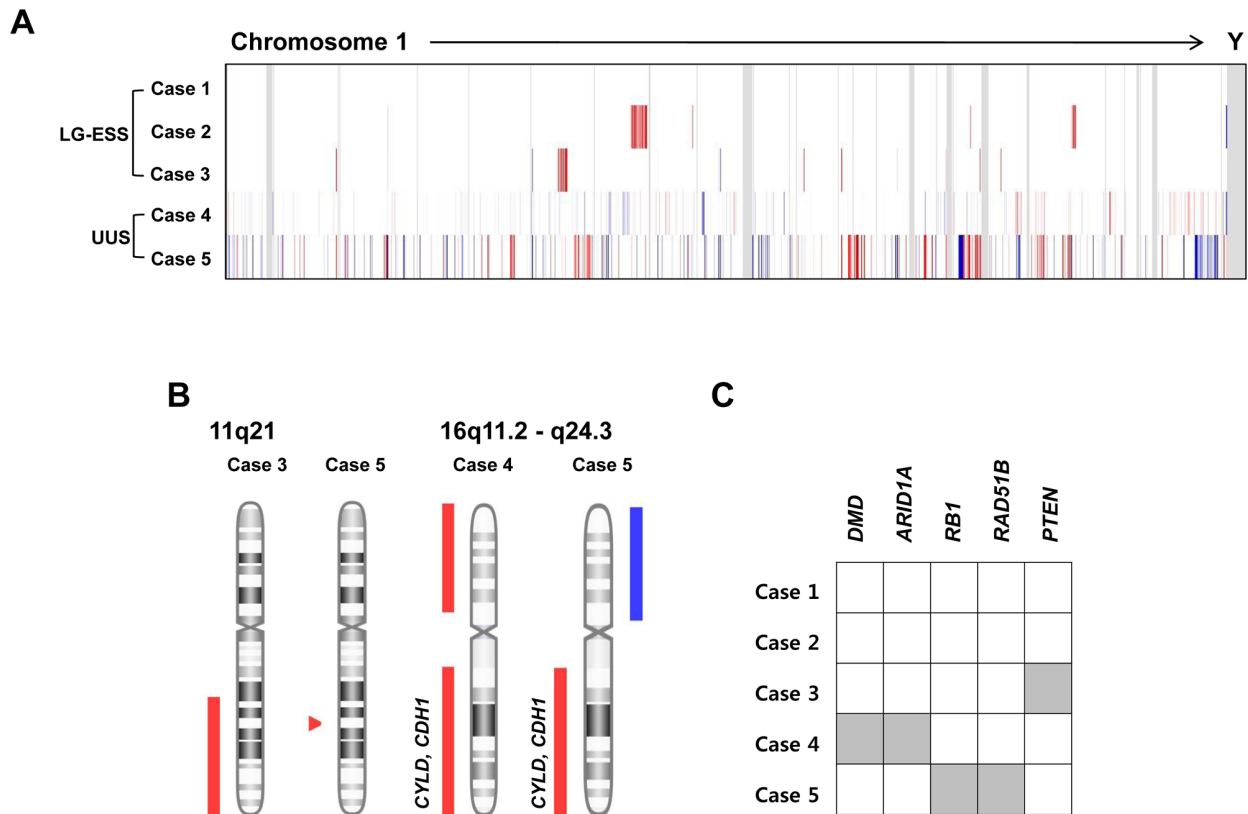


Figure 1: Copy number alteration (CNA) profiles of five endometrial stromal sarcomas. **A.** A heatmap for probe-level signal intensities (\log_2 ratios) of five endometrial stromal sarcomas. (Blue: CNA gains, Red: CNA losses) **B.** Recurrent CNAs ($n \geq 2$) on 11q21 (CNA loss) and 16q11.2 - q24.3. The CNA loss on 16q encompasses candidate tumor suppressor genes *CYLD* and *CDH1*. (Blue: CNA gains, Red: CNA losses) **C.** Five genes (grey) that are catalogued in pan-cancer CNA driver database [19].

Whole-exome sequencing profiles

The five ESSs (three LG-ESSs and two UUSs) were also analyzed by WES. Mean coverage of the sequencing depth was 116X (range: 88–143X) for normal samples and 142X (range: 128–155X) for tumor samples. A total of 81 non-silent point mutations and indels (median: 10, range: 6–36) were identified in the five ESSs (Supplementary Table S3). Both distribution of sequence composition and relative fraction of mutation spectra were not significantly different among the five ESSs ($P > 0.05$). Missense mutation with C:G > T:A transition was the most common type (Figures 3A and 3B).

To address whether the mutations found in our study could be causally implicated in ESS development, we queried the cancer Gene Census and found four mutations (*SRGAP3*, *EBF1*, *IRF4* and *PPARG*). Also, by analyzing the CHASM to distinguish driver and passenger mutations, we identified ten putative cancer-related mutations (*THSD7A*, *TBC1D14*, *TTK*, *SLC38A1*, *CTNNA2*, *SUPT6H*, *DCC*, *MDGA2*, *CNNM1* and *MYO19*) [18]. Finally, in the pan-cancer database, we detected *SRGAP3* and *ZFP36L2* as putative cancer-related genes [19]. Together, we detected 15 putative cancer-related genes with somatic mutations that could be causally

implicated in ESS development (Figure 3C). Of them, *SRGAP3* was co-detected in the cancer Gene Census and the pan-cancer database (Figure 3C). Sanger validation of the mutations including four putative cancer-related genes (*SRGAP3*, *PPARG*, *DCC* and *ZFP36L2*) is illustrated in Supplementary Figure S2. Of the mutated genes identified, three genes, which included a cancer-related gene *MYO19*, were found to have expressional changes in the same cases by the transcriptome analysis (Supplementary Figure S3).

DISCUSSION

The aim of this study was twofold. First, we attempted to disclose any somatic genetic alterations other than the gene fusions in ESSs. Second, we for the first time attempted to disclose genetic features of UUSs that by definition do not harbor any of the ESS-specific fusions. We found that the ESSs analyzed in this study harbored 6–36 non-silent somatic mutations per genome, but they did not include well-known mutations (e.g., *TP53*, *KRAS* and *PIK3CA*). As for the CNAs, all except one ESS harbored CNAs, many of which encompassed well-known driver genes. Importantly, some cancer-related genes in UUSs showed a close correlation

Table 2: Summary of gains and losses in five endometrial stromal sarcomas detected by microarray-CGH

	Diagnosis	Gains	Losses	Cancer-related genes
Case 1	LG-ESS	none	none	none
Case 2	LG-ESS	none	7p22.3 - p14.2, 18p11.32 - p11.21	<i>PMS2</i>
Case 3	LG-ESS	none	5q22.3 - q34, 10q22.3 - q24.2, 11q21, 15q22.1 - q22.2	<i>PTEN</i>
Case 4	UUS	6p21.33 - p21.32, 6q21, 6q24.3 - q27, 8q12.1 - q13.1, 13q14.2 - q14.3, 13q21.2, 14q32.33,	1p36.33 - p35.3, 9p24.1 - p21.3, 9p21.3 - p13.1, 16p13.3 - p11.2, 16q11.2 - q24.3, 19q11 - q13.43, Xp22.33 - p11.21, Xq21.32 - q22.1	<i>ARID1A, POU5F1, EZR, DMD, CYLD*</i> , <i>CDHI*</i>
Case 5	UUS	9q21.11 - q31.3, 14q12 - q22.1, 16p13.3 - p11.2, 18p11.32, Xq13.1 - q28	2q36.3 - q37.3, 4q21.1, 4q24 - q25, 4q25, 4q34.1 - q35.2, 6p25.3 - p12.1, 11q14.1 - q25, 13q14.2 - q14.3, 13q21.33, 14q22.1 - q32.33, 16q11.2 - q24.3, 17p13.1, 17q21.33 - q25.3, 22q13.33	<i>GNAQ, SYK, RBI, RAD51B, DICER1, CYLD*</i> , <i>CDHI*</i> , <i>TP53, PRKARIA</i>

*The cancer-related genes that were recurrently harbored in CNAs in UUSs ($n \geq 2$).

LG-ESS: low-grade ESS, UUS: undifferentiated uterine sarcoma

between CNAs and gene expression changes, strongly suggesting their implications on ESS tumorigenesis. The UUSs showed bigger median values of CNAs, non-silent mutations and putative cancer-related genes than the LG-ESSs (Table 4), but the differences were not significant, probably due to the small number of the ESSs analyzed. Together, this study identified that ESSs harbored not only ESS-specific fusions but also somatic mutations and CNAs encompassing driver genes. Our findings suggest a possibility that gene fusions alone may not fully develop ESSs as identified in other tumors [20].

We identified five genes in UUSs that shared CNA and gene expression changes (down-regulation and CNA loss: *PRKARIA*, *CDHI*, *RBI* and *TP53*, up-regulation and CNA gain: *EZR*). *CDHI* is a tumor suppressor gene that encodes E-cadherin and is known to be associated with various malignancies [21, 22]. Loss of *CDHI* results in dysfunction of cell-cell adhesion, allowing for abnormal

cell-cell interaction such as epithelial-mesenchymal transition (EMT) [23]. In addition, *CDHI* behaves as a negative regulator in Wnt signaling [24]. ESSs are related to up-regulation of Wnt signaling, for example, exhibiting a down-regulation of *SFRP4*, a negative regulator in Wnt signaling [22, 25]. In our study, we also found that both *SFRP4* and *CDHI* expressions were decreased in the ESSs, further suggesting the importance of Wnt signaling in ESSs. Recent research advances yielded a number of US Food and Drug Administration (FDA)-approved drugs that may change Wnt signaling [26]. Our study may provide further rationale for performing future studies that may explore the use of Wnt modulators in ESSs. *TP53* is the most common tumor suppressor gene that is frequently inactivated in most cancers [27]. Tumor suppressor genes generally follow “two-hit hypothesis”, being bi-allelically inactivated by point mutation, deletion and promoter hypermethylation [28]. In spite of the earlier report on

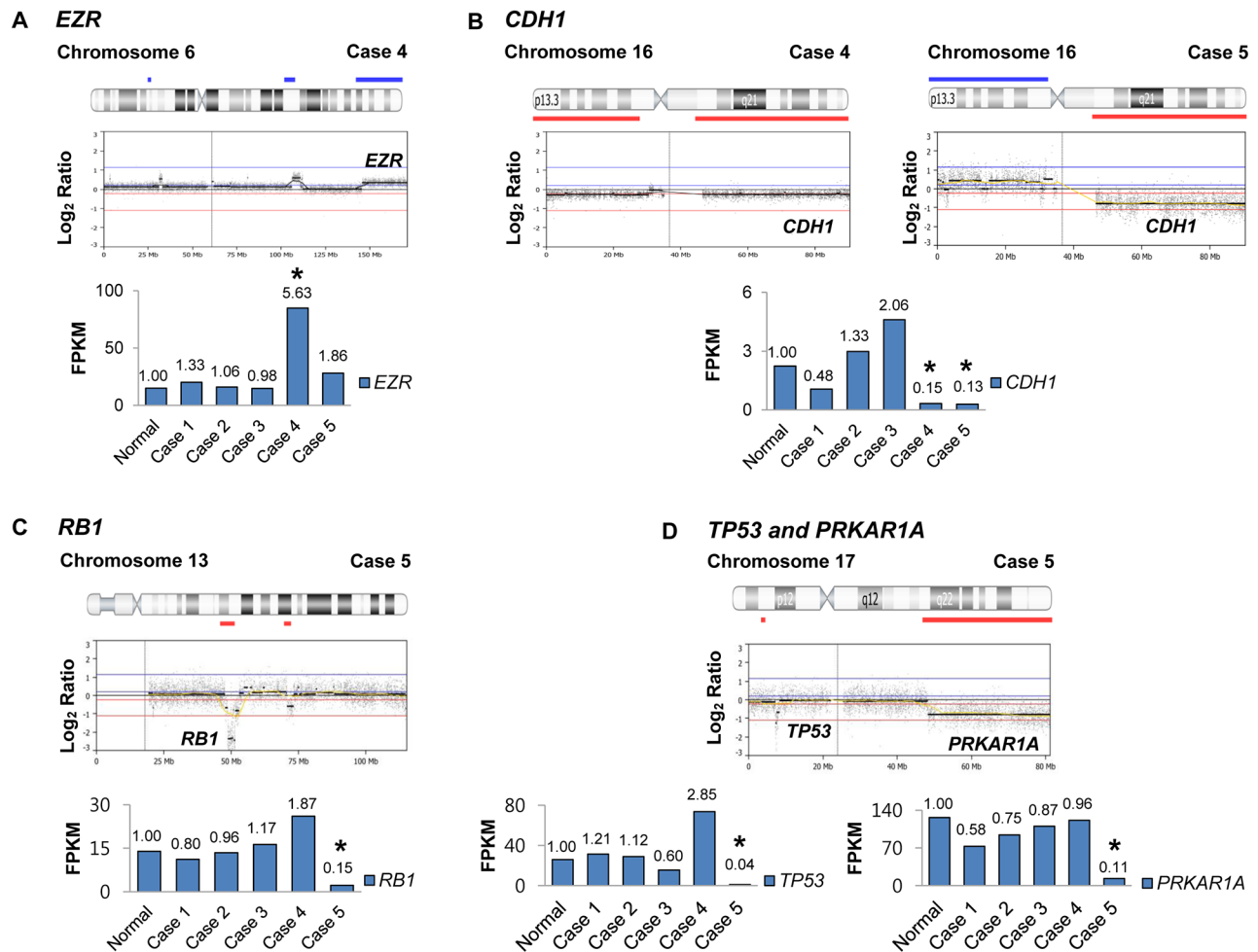


Figure 2: The loci maps of cancer-related genes with copy number alterations and correlated gene expressions. Each of the copy number loci is denoted with the corresponding gene. The graph depicts gene expression levels of five endometrial stromal sarcoma tissues as compared to the normal endometrial tissue (fold-changes are shown above the graph). The cases where expression changes accompany CNAs are marked with asterisk (*). (FPKM: fragments per kilobase of exon per million fragments mapped) **A.** Case 4 harbors a copy number alteration (CNA) gain on *EZR* locus that exhibits expressional increase. **B.** Cases 4 and 5 harbor CNA losses on *CDH1* locus that exhibits expressional decrease. **C.** Case 5 harbors a homozygous CNA loss on *RB1* locus that exhibits expressional decrease. **D.** Case 5 harbors CNA losses on *TP53* and *PRKAR1A* loci that exhibit expressional decreases.

Table 3: Gene fusions detected in endometrial stromal sarcomas

Diagnosis	Sample ID	5'-partner gene					3'-partner gene				
		Gene	Reference sequence	Chromosome	Break position	Strand	Gene	Reference sequence	Chromosome	Break position	Strand
LG-ESS	Case 1	<i>MEAF6</i>	ENSG00000163875	1	37967405	+	<i>PHF1</i>	ENSG00000112511	6	33380025	-
LG-ESS	Case 2	<i>JAZF1</i>	ENSG00000153814	7	27934839	+	<i>SUZ12</i>	ENSG00000178691	17	30267305	-
LG-ESS	Case 3	<i>JAZF1</i>	ENSG00000153814	7	27934839	+	<i>PHF1</i>	ENSG00000112511	6	33380027	-

LG-ESS: low-grade ESS

positive *TP53* mutations in ESSs [29], we did not identify any *TP53* point mutations in the ESSs by WES. In a UUS (case 5), *TP53* locus was deleted heterozygously and *TP53* expression was decreased. Together, these data indicate that *TP53* inactivation in ESSs may result from point mutation or deletion and suggest a possibility that

haploinsufficiency of *TP53* might inactivate its tumor suppressor roles. Similarly, chronic lymphocytic leukemia exhibited both bi-allelic and mono-allelic *TP53* gene [30]. By contrast, we observed that another tumor suppressor gene *RB1* exhibited homozygous deletion in another UUS (case 5).

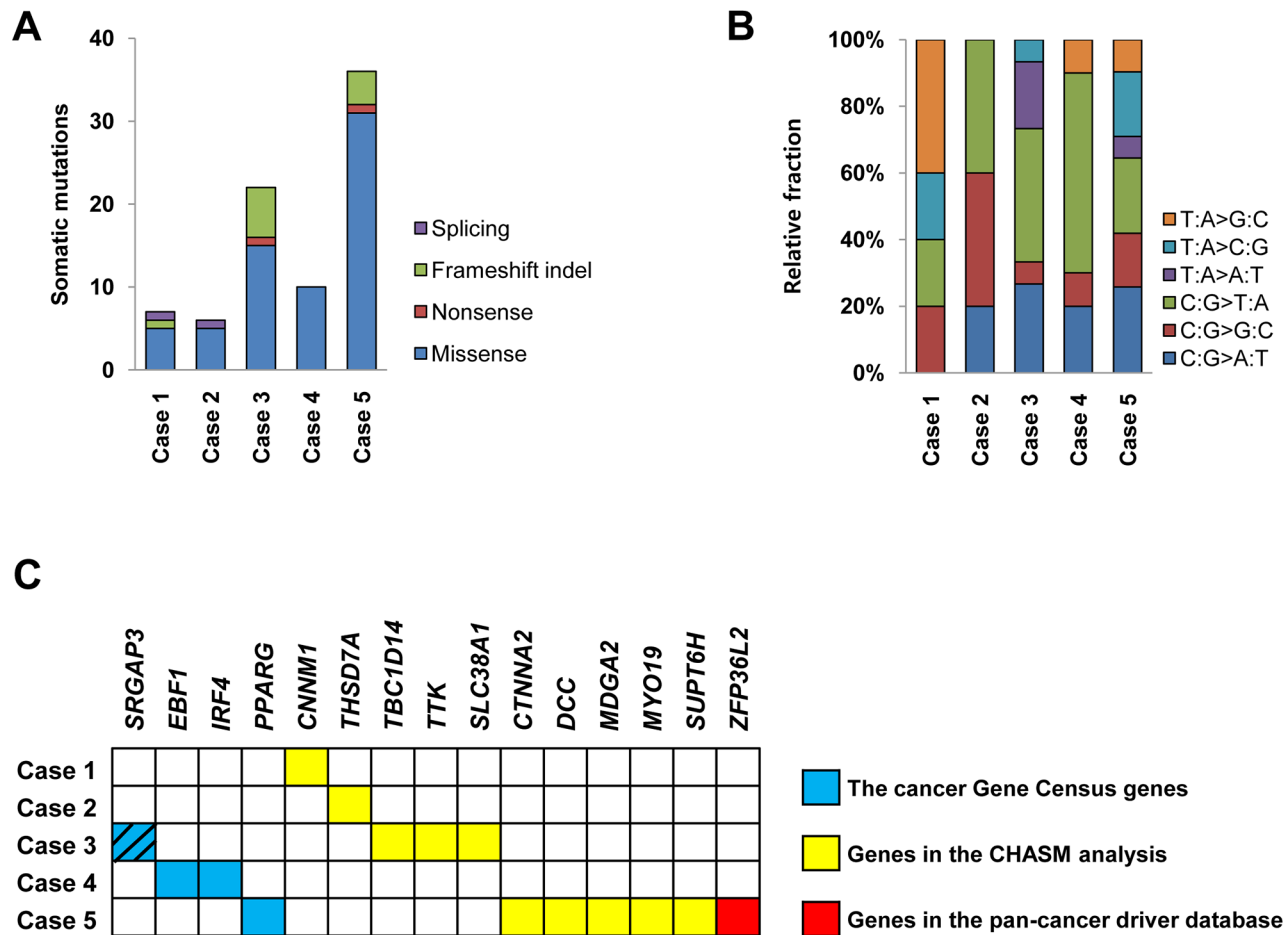


Figure 3: Somatic mutations of five endometrial stromal sarcomas. A. and B. The numbers of somatic mutations and their relative fractions of sequence spectra of five endometrial stromal sarcomas are shown. C. Fifteen putative cancer-related genes with somatic mutations. Blue box denotes the genes that overlap the cancer Gene Census genes [17], yellow box denotes the genes that were detected by the CHASM analysis [18] and red box denotes the gene catalogued in the pan-cancer driver database [19]. Of them, *SRGAP3* was detected in both of the cancer Gene Census and the pan-cancer driver database (hatched).

Table 4: Summary of comparison data between LG-ESS and UUS genomes

	LG-ESS (n = 3)	UUS (n = 2)	P value
Number of CNA	6	34	0.08
Length of CNA	129 Mb	746 Mb	0.08
Cancer-related genes in which copy number and gene expression are correlated	0	5	0.05
Gene fusions	3	0	0.05
Non-silent somatic mutation numbers	35	46	0.25
Putative driver genes (15 somatic mutations and five CNAs)	9	11	0.56

CNA: copy number alteration, LG-ESS: low-grade endometrial stromal sarcoma, UUS: undifferentiated uterine sarcoma

Recurrent fusion transcripts are frequently detected in ESSs [4, 5, 9–13]. *JAZF1-SUZ12*, *JAZF1-PHF1* and *MEAF6-PHF* are gene-fusions that have been detected most frequently in LG-ESSs [5, 12, 13] and *YWHAE-FAM22* gene fusion is a key factor in defining HG-ESS

[31]. By definition, UUSs are high-grade ESSs without known gene fusions [2]. As expected, we were able to find *JAZF1-SUZ12*, *JAZF1-PHF1* and *MEAF6-PHF* fusions in each of the three LG-ESSs. For the two ESSs without any known ESS-specific fusions, we were not able to detect

any novel fusions, identifying that they were UUSs per se. Our data suggest that the UUS with a driver fusion might, if any, be very rare.

We found *PPARG* and *IRF4* mutations in the ESSs. *PPARG* encodes peroxisome proliferator-activated receptor gamma (PPAR- γ /PPARG) [32] that has tumor suppressor functions in many endocrine organs including breast, prostate and pituitary gland [33–35]. In uterus, PPARG activation inhibits growth and survival of human endometriosis cells by suppressing estrogen biosynthesis [36]. *IRF4* encodes a transcription factor in interferon regulatory factor family. A chromosomal translocation involving *IRF4* and the IgH locus, t(6;14)(p25;q32) is considered a cause of multiple myeloma [37]. *IRF4* is required for endometrial decidualization [38]. Together, these data suggest a possible rationale that *PPARG* and *IRF4* mutations might be involved in ESS development.

In this study, ESSs harbored at least one genetic alteration (somatic mutations or CNAs or gene fusions) that may stimulate ESS tumorigenesis (Figure 4). Also, we observed that all of the ESSs carried either somatic mutations or CNAs-harboring driver genes, albeit not recurrent. It suggests that non-recurrent alterations may cooperate together for the ESS tumorigenesis.

Our data are based on the analysis of five ESS genome pairs (three LG-ESSs and two UUSs). The small sample size is due to the rarity of ESS [1, 2]. Further investigation with larger sample size across diverse ethnic groups would reveal additional information, e.g.,

discovery of potential additional driver mutations in ESSs and additional novel fusions in LG-ESSs and HG-ESSs. In addition, a larger cohort would provide genomic features of metastatic ESSs compared to the primary ESSs.

In summary, our data for the first time attempted the integrative analyses of whole-exome, whole-transcriptome and a-CGH in ESS genomes. Previous studies on ESS genomes mainly focused on gene fusions. However, our data indicate that fusions are not the only genetic alteration occurred during ESS development. Somatic mutations, CNAs as well as gene fusions alone or together might contribute to ESS development. Our data also suggest that CNAs may be a major genomic alteration for UUS development while gene fusions are the major genomic alteration for LG-ESS. Our findings may provide a useful resource for understanding this heterogeneous disease and identifying genomic clues for differential diagnosis and therapy options for ESS.

MATERIAL AND METHODS

Endometrial stromal sarcoma tissues

Normal and tumor tissues from five ESS patients were obtained from the tissue banks of Korean Gynecologic Cancer Bank (Seoul, Korea), Guro Hospital of Korea University (Seoul, Korea) and Busan National University Hospital (Busan, Korea). We also used normal endometrial tissue from a healthy woman. All of the six samples were

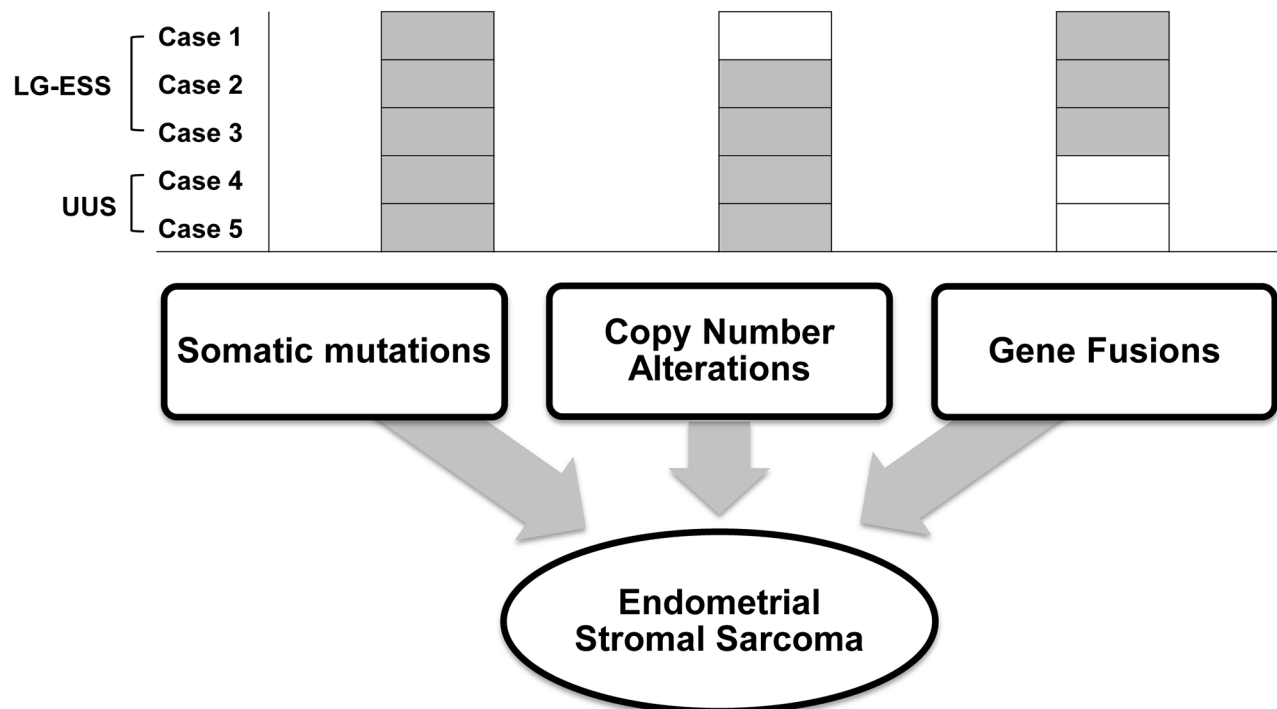


Figure 4: Schematic representation of suggested genetic alterations in endometrial stromal sarcoma. Somatic mutations and copy number alterations as well as gene fusions are found in the endometrial stromal sarcomas. Compositions of these genomic alterations are summarized for each ESS.

from Koreans. Approval for this study was obtained from the institutional review board at the Catholic University of Korea, College of Medicine. Clinicopathologic features of the five ESS patients are summarized in Table 1. Initially, frozen tissues from the tissue banks were cut, stained with the hematoxylin/eosin and examined under microscope by a pathologist, who identified areas rich in ESS tumor cells in the frozen tissues. In order to procure matched normal tissues from each ESS patient, we used peripheral blood lymphocytes or normal tissues that were confirmed to be free of tumor cells by microscopic examination. Purities of the tumor cells were approximately 70%. For genomic DNA and RNA extraction from the frozen tissues, we used the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) and mirVanamiRNA Isolation Kit (Invitrogen, Carlsbad, CA, USA), respectively.

Whole-exome sequencing

Using genomic DNA from five ESSs and matched normal cells, we performed exome-capture using the Agilent SureSelect Human All Exome 50Mb kit (Agilent Technologies) according to the manufacturer's instructions. DNA libraries were constructed according to the protocol provided by the manufacturer and whole exome-sequencing (WES) was performed using an Illumina HiSeq2000 platform to generate 101bp paired-end reads. Burrows-Wheeler aligner was used to align the sequencing reads onto the human reference genome (hg19). The aligned sequencing reads were evaluated by using Qaulimap [39]. Supplementary Table S4 shows the information of sequencing alignments (e.g., the number of reads and sequencing coverage). Processing and the management of the sequencing data were performed as described elsewhere [20]. In brief, somatic variants were identified by using MuTect [40] and SomaticIndelDetector [41] for point mutations and indels. ANNOVAR package was used to select somatic variants located in coding sequences and to predict their functional consequences, such as silent or non-silent variants [42]. Then, we used the CHASM analysis program with 'uterus' option for cancer tissue type (FDR < 0.3) in order to identify the putative cancer-related mutations [18]. In order to validate noticeable somatic mutations, genomic DNA from tumor areas and matched normal tissues from each case were amplified by PCR and sequence analyses were performed.

Transcriptome sequencing for gene fusion and expression profiling analyses

The mRNA of five ESSs and normal proliferative endometrium of a woman was converted into a 272bp to 289bp-sized cDNA library using TruSeq RNA Library Preparation Kit (Illumina). Whole-transcriptome sequencing was performed using an Illumina HiSeq2000 platform. Sequencing reads were mapped onto the human

reference genome (GRCh37, hg19). Gene fusions were identified by searching for the spanning reads and split reads by using the deFuse program [43]. Transcriptome sequencing data were analyzed using TopHat (<http://ccb.jhu.edu/software/tophat/index.shtml>) for alignment, Cufflinks for assembly [44] and a known set of reference transcripts from Ensembl v. 65 (<http://www.ensembl.org>) for expression estimation. The expression levels were quantified as fragments per kilobase of exon per million fragments mapped (FPKM).

Reverse transcription-polymerase chain reaction and sequencing

RT was performed using oligo-(dT) primer and SuperScript III reverse transcriptase (Invitrogen). PCR was performed with Pfu DNA polymerase (Promega) according to the manufacturer's instruction. The thermal cycling included one cycle at 95°C for 2 min followed by 35 cycles of 95°C at 0.5 min, 55–61°C for 0.5 min, 72°C for 1 min and a final extension of 72°C for 5 min. Details of the primer pairs and corresponding genes are available in Supplementary Table S5. PCR products were visualized on 2% agarose gel and subsequently analyzed by direct DNA sequencing.

DNA copy number profiling

DNA copy number profiling was performed using the Agilent Sure Print G3 Human CGH Microarray 180K. The genomic DNA of five ESS tissues and matched normal tissues was hybridized onto the array according to the manufacturer's instructions. Background correction and normalization for array images was performed using Agilent Feature Extraction Software v10.7.3.1. The RankSegmentation statistical algorithm in NEXUS software v7.5 (Biodiscovery Inc., El Segundo, CA) was used to define the CNAs of each sample; a log₂ ratio larger than 0.3 was identified as gain and lower than -0.3 as loss.

ACKNOWLEDGEMENTS AND FUNDING

This study was supported by a grant from National Research Foundation of Korea (2012R1A5A2047939).

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

Authors' contribution

YJC and SugHL conceived of the study. YJC, SHJ, SugHL and YJC wrote the article. SHJ, SHL and MSK performed experiments. TMK, JKR and IPB performed computation analyses. YJC and SYH collected the

specimen and performed clinical review. All authors have read and approved the manuscript for publication.

REFERENCES

1. D'Angelo E, Prat J. Uterine sarcomas: a review. *Gynecol Oncol.* 2010; 116:131–139.
2. Ali RH, Rouzbahman M. Endometrial stromal tumours revisited: an update based on the 2014 WHO classification. *J Clin Pathol.* 2015; 68:325–332.
3. Chang KL, Crabtree GS, Lim-Tan SK, Kempson RL, Hendrickson MR. Primary uterine endometrial stromal neoplasms. A clinicopathologic study of 117 cases. *Am J Surg Pathol.* 1990; 14:415–438.
4. Lee CH, Ou WB, Mariño-Enriquez A, Zhu M, Mayeda M, Wang Y, Guo X, Brunner AL, Amant F, French CA, West RB, McAlpine JN, et al. 14-3-3 fusion oncogenes in high-grade endometrial stromal sarcoma. *Proc Natl Acad Sci U S A.* 2012; 109:929–934.
5. Koontz JI, Soreng AL, Nucci M, Kuo FC, Pauwels P, van Den Berghe H, Dal Cin P, Fletcher JA, Sklar J. Frequent fusion of the JAZF1 and JJAZ1 genes in endometrial stromal tumors. *Proc Natl Acad Sci U S A.* 2001; 98:6348–6353.
6. Nucci MR, Harburger D, Koontz J, Dal Cin P, Sklar J. Molecular analysis of the JAZF1-JJAZ1 gene fusion by RT-PCR and fluorescence *in situ* hybridization in endometrial stromal neoplasms. *Am J Surg Pathol.* 2007; 31:65–70.
7. Huang HY, Ladanyi M, Soslow R. Molecular detection of JAZF1-JJAZ1 gene fusion in endometrial stromal neoplasms with classic and variant histology: evidence for genetic heterogeneity. *Am J Surg Pathol.* 2004; 28:224–232.
8. Hrzenjak A, Moinfar F, Tavassoli FA, Strohmeier B, Kremser ML, Zatloukal K, Denk H. JAZF1/JJAZ1 gene fusion in endometrial stromal sarcomas: molecular analysis by reverse transcriptase-polymerase chain reaction optimized for paraffin-embedded tissue. *J Mol Diagn.* 2005; 7:388–395.
9. Micci F, Gorunova L, Gatus S, Matias-Guiu X, Davidson B, Heim S, Panagopoulos I. MEAF6/PHF1 is a recurrent gene fusion in endometrial stromal sarcoma. *Cancer Lett.* 2014; 347:75–78.
10. Dewaele B, Przybyl J, Quattrone A, Finalet Ferreiro J, Vanspauwen V, Geerdens E, Gianfelici V, Kalender Z, Wozniak A, Moerman P, Sciort R, Croce S, et al. Identification of a novel, recurrent MBTD1-CXorf67 fusion in low-grade endometrial stromal sarcoma. *Int J Cancer.* 2014; 134:1112–1122.
11. Panagopoulos I, Thorsen J, Gorunova L, Haugom L, Bjerkehagen B, Davidson B, Heim S, Micci F. Fusion of the ZC3H7B and BCOR genes in endometrial stromal sarcomas carrying an X;22-translocation. *Genes Chromosomes Cancer.* 2013; 52:610–618.
12. Panagopoulos I, Micci F, Thorsen J, Gorunova L, Eibak AM, Bjerkehagen B, Davidson B, Heim S. Novel fusion of MYST/Esa1-associated factor 6 and PHF1 in endometrial stromal sarcoma. *PLoS One.* 2012; 7:e39354.
13. Micci F, Panagopoulos I, Bjerkehagen B, Heim S. Consistent rearrangement of chromosomal band 6p21 with generation of fusion genes JAZF1/PHF1 and EPC1/PHF1 in endometrial stromal sarcoma. *Cancer Res.* 2006; 66:107–112.
14. Davidson B, Abeler VM, Hellesylt E, Holth A, Shih IeM, Skeie-Jensen T, Chen L, Yang Y, Wang T-L. Gene expression signatures differentiate uterine endometrial stromal sarcoma from leiomyosarcoma. *Gynecol Oncol.* 2013; 128:349–355.
15. Halbwedl I, Ullmann R, Kremser ML, Man YG, Isadi-Moud N, Lax S, Denk H, Popper HH, Tavassoli FA, Moinfar F. Chromosomal alterations in low-grade endometrial stromal sarcoma and undifferentiated endometrial sarcoma as detected by comparative genomic hybridization. *Gynecol Oncol.* 2005; 97:582–587.
16. Flicker K, Smolle E, Haybaeck J, Moinfar F. Genomic characterization of endometrial stromal sarcomas with array comparative genomic hybridization. *Exp Mol Pathol.* 2015; 98:367–374.
17. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR. A census of human cancer genes. *Nat Rev Cancer.* 2004; 4:177–183.
18. Carter H, Samayoa J, Hruban RH, Karchin R. Prioritization of driver mutations in pancreatic cancer using cancer-specific high-throughput annotation of somatic mutations (CHASM). *Cancer Biol Ther.* 2010; 10:582–587.
19. Rubio-Perez C, Tamborero D, Schroeder MP, Antolín AA, Deu-Pons J, Perez-Llamas C, Mestres J, Gonzalez-Perez A, Lopez-Bigas N. In silico prescription of anticancer drugs to cohorts of 28 tumor types reveals targeting opportunities. *Cancer Cell.* 2015; 27:382–396.
20. Kim TM, Jung SH, Baek IP, Lee SH, Choi YJ, Lee JY, Chung YJ, Lee SH. Regional biases in mutation screening due to intratumoural heterogeneity of prostate cancer. *J Pathol.* 2014; 233:425–435.
21. Van Roy F. Beyond E-cadherin: roles of other cadherin superfamily members in cancer. *Nat Rev Cancer.* 2014; 14:121–134.
22. Herzig M, Savarese F, Novatchkova M, Semb H, Christofori G. Tumor progression induced by the loss of E-cadherin independent of beta-catenin/Tcf-mediated Wnt signaling. *Oncogene.* 2007; 26:2290–2298.
23. Howard S, Deroo T, Fujita Y, Itasaki N. A positive role of cadherin in Wnt/ β -catenin signalling during epithelial-mesenchymal transition. *PLoS One.* 2011; 6:e23899.
24. Orsulic S, Huber O, Aberle H, Arnold S, Kemler R. E-cadherin binding prevents beta-catenin nuclear localization and beta-catenin/LEF-1-mediated transactivation. *J Cell Sci.* 1999; 112:1237–1245.

25. Hrzenjak A, Dieber-Rotheneder M, Moinfar F, Petru E, Zatloukal K. Molecular mechanisms of endometrial stromal sarcoma and undifferentiated endometrial sarcoma as premises for new therapeutic strategies. *Cancer Lett.* 2014; 354:21–27.
26. Anastas JN, Moon RT. WNT signalling pathways as therapeutic targets in cancer. *Nat Rev Cancer.* 2013; 13:11–26.
27. Cho Y, Gorina S, Jeffrey PD, Pavletich NP. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science.* 1994; 265:346–355.
28. Knudson AG. Two genetic hits (more or less) to cancer. *Nat Rev Cancer.* 2001; 1:157–162.
29. Kurihara S, Oda Y, Ohishi Y, Iwasa A, Takahira T, Kaneki E, Kobayashi H, Wake N, Tsuneyoshi M. Endometrial stromal sarcomas and related high-grade sarcomas: immunohistochemical and molecular genetic study of 31 cases. *Am J Surg Pathol.* 2008; 32:1228–1238.
30. Malcikova J, Smardova J, Rocnova L, Tichy B, Kuglik P, Vranova V, Cejkova S, Svitakova M, Skuhrova Francova H, Brychtova Y, Doubek M, Brejcha M, et al. Monoallelic and biallelic inactivation of TP53 gene in chronic lymphocytic leukemia: selection, impact on survival, and response to DNA damage. *Blood.* 2009; 114:5307–5314.
31. Isphording A, Ali RH, Irving J, Goytain A, Nelnyk N, Hoang LN, Gilks CB, Huntsman DG, Nielsen TO, Nucci MR, Lee CH. YWHAE-FAM22 endometrial stromal sarcoma: diagnosis by reverse transcription-polymerase chain reaction in formalin-fixed, paraffin-embedded tumor. *Hum Pathol.* 2013; 44:837–843.
32. Lehrke M, Lazar MA. The many faces of PPARgamma. *Cell.* 2005; 123:993–999.
33. Mueller E, Sarraf P, Tontonoz P, Evans RM, Martin KJ, Zhang M, Fletcher C, Singer S, Spiegelman BM. Terminal differentiation of human breast cancer through PPAR gamma. *Mol Cell.* 1998; 1:465–470.
34. Mueller E, Smith M, Sarraf P, Kroll T, Aiyer A, Kaufman DS, Oh W, Demetri G, Figg WD, Zhou XP, Eng C, Spiegelman BM, et al. Effects of ligand activation of peroxisome proliferator-activated receptor gamma in human prostate cancer. *Proc Natl Acad Sci U S A.* 2000; 97:10990–10995.
35. Heaney AP, Fernando M, Yong WH, Melmed S. Functional PPAR-gamma receptor is a novel therapeutic target for ACTH-secreting pituitary adenomas. *Nat Med.* 2002; 8:1281–1287.
36. Lebovic DI, Kavoussi SK, Lee J, Banu SK, Arosh JA. PPARγ activation inhibits growth and survival of human endometriotic cells by suppressing estrogen biosynthesis and PGE2 signaling. *Endocrinology.* 2013; 154:4803–4813.
37. Iida S, Rao PH, Butler M, Corradini P, Boccadoro M, Klein B, Chaganti RS, Dalla-Favera R. Deregulation of MUM1/IRF4 by chromosomal translocation in multiple myeloma. *Nat Genet.* 1997; 17:226–230.
38. Vasquez YM, Mazur EC, Li X, Kommagani R, Jiang L, Chen R, Lanz RB, Kovanci E, Gibbons WE, DeMayo FJ. FOXO1 is required for binding of PR on IRF4, novel transcriptional regulator of endometrial stromal decidualization. *Mol Endocrinol.* 2015; 29:421–433.
39. García-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Götz S, Tarazona S, Dopazo J, Meyer TF, Conesa A. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics.* 2012; 28:2678–2679.
40. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol.* 2013; 31:213–219.
41. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011; 43:491–498.
42. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010; 38:e164.
43. McPherson A, Hormozdiari F, Zayed A, Giuliany R, Ha G, Sun MG, Griffith M, Heravi Moussavi A, Senz J, Melnyk N, Pacheco M, Marra M, et al. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. *PLoS Comput Biol.* 2011; 7:e1001138.
44. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc.* 2012; 7:562–578.