Gynecologic Oncology xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Gynecologic Oncology

journal homepage: www.elsevier.com/locate/ygyno

Gene expression profiling of low-grade endometrial stromal sarcoma indicates fusion protein-mediated activation of the Wnt signaling pathway

Joanna Przybyl [a,*], Lukasz Kidzinski [b], Trevor Hastie [c], Maria Debiec-Rychter [d], Roel Nusse [e], Matt van de Rijn [a]

[a] Department of Pathology, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, 94305, CA, USA
[b] Department of Bioengineering, Stanford University, 443 Via Ortega, Stanford, 94305, CA, USA
[c] Department of Statistics, Stanford University, 390 Serra Mall, Stanford, 94305, CA, USA
[d] Department of Human Genetics, KU Leuven and University Hospitals Leuven, UZ, Herestraat 49, 3000 Leuven, Belgium
[e] Department of Developmental Biology, Stanford University School of Medicine, 265 Campus Drive, Stanford, 94305, CA, USA

HIGHLIGHTS

- LGESS-specific fusion proteins may disrupt the repressive function of PRC2 complex.
- SUZ12 targets are consistently deregulated in LGESS.
- LGESS tumors show overexpression of multiple genes implicated in the Wnt signaling.
- LGESS tumors demonstrate concordant nuclear expression of β-catenin and Lef1.
- Therapeutic targeting of Wnt pathway may prove beneficial to LGESS patients.

ARTICLE INFO

Article history:
Received 22 January 2018
Received in revised form 3 March 2018
Accepted 7 March 2018
Available online xxxx

Keywords:
Endometrial stromal sarcoma
Uterine sarcoma
Gene expression profiling
Wnt signaling pathway
Chromosomal translocations

ABSTRACT

Objective. Low-grade endometrial stromal sarcomas (LGESS) harbor chromosomal translocations that affect proteins associated with chromatin remodeling Polycomb Repressive Complex 2 (PRC2), including SUZ12, PHF1 and EPC1. Roughly half of LGESS also demonstrate nuclear accumulation of β-catenin, which is a hallmark of Wnt signaling activation. However, the targets affected by the fusion proteins and the role of Wnt signaling in the pathogenesis of these tumors remain largely unknown.

Methods. Here we report the results of a meta-analysis of three independent gene expression profiling studies on LGESS and immunohistochemical evaluation of nuclear expression of β-catenin and Lef1 in 112 uterine sarcoma specimens obtained from 20 LGESS and 89 LMS patients.

Results. Our results demonstrate that 143 out of 310 genes overexpressed in LGESS are known to be directly regulated by SUZ12. In addition, our gene expression meta-analysis shows activation of multiple genes implicated in Wnt signaling. We further emphasize the role of the Wnt signaling pathway by demonstrating concordant nuclear expression of β-catenin and Lef1 in 7/16 LGESS.

Conclusions. Based on our findings, we suggest that LGESS-specific fusion proteins disrupt the repressive function of the PRC2 complex similar to the mechanism seen in synovial sarcoma, where the SS18-SSX fusion proteins disrupt the mSWI/SNF (BAF) chromatin remodeling complex. We propose that these fusion proteins in LGESS contribute to overexpression of Wnt ligands with subsequent activation of Wnt signaling pathway and formation of an active β-catenin/Lef1 transcriptional complex. These observations could lead to novel therapeutic approaches that focus on the Wnt pathway in LGESS.

1. Introduction

Endometrial stromal sarcoma (ESS) and leiomyosarcomas (LMS) are the two most frequent types of uterine sarcomas, accounting for approximately 20% and 60% of cases, respectively. Low grade endometrial stromal sarcomas (LGESS) harbor two major molecular

* Corresponding author at: Department of Pathology, Stanford University School of Medicine, 300 Pasteur Drive, Lane Bldg., L209, 94305 Stanford, CA, USA.
E-mail address: jprzybyl@stanford.edu (J. Przybyl).

https://doi.org/10.1016/j.ygyno.2018.03.007
0090-8258/

hallmarks: specific chromosomal translocations and nuclear accumulation of β-catenin [1]. Chromosomal translocations are detected in the vast majority of LGESS and nuclear expression of β-catenin, an indication of activated Wnt signaling pathway, was described in over 60% (83/137) of LGESS cases in five previously published studies [2–6]. Despite the high frequency of these alterations, the relation between these two molecular findings and the manner in which they may contribute to LGESS tumorigenesis have not been thoroughly studied.

The majority of LGESS carry subtype-specific translocations that affect several proteins associated with the Polycomb Repressive Complex 2 (PRC2). These include SUZ12, PHF1 and EPC1 [7–10], suggesting a shared pathogenetic mechanism between different rearrangements. The most common translocation in LGESS is t(7;17)(p15; q21) that results in a *JAZF1-SUZ12* fusion. *SUZ12* is one of the core components of the PRC2 that mediates gene repression via trimethylation of H3K27. It was recently shown by Ma et al. that overexpression of JAZF1-SUZ12 fusion protein in 293T cells destabilized the PRC2 complex and decreased its histone methyl transferase activity [11]. This disruption of PRC2 resulted in upregulation of two genes involved in the Wnt pathway, *HOXA9* and *WNT11*, which are normally repressed by this complex [11].

The purpose of this study was to explore a broader spectrum of molecular targets affected by fusion oncoproteins and abnormal Wnt signaling in LGESS. Based on the combined analysis of multiple previously published gene expression datasets, we show for the first time that roughly half of the genes consistently deregulated in LGESS constitute direct targets of the SUZ12. These genes are normally repressed by the PRC2 complex and their upregulation point to a possible disruption of this complex by abnormal SUZ12. This suggests a similar mechanism as has been demonstrated in synovial sarcoma where SS18-SSX fusion proteins disrupt mSWI/SNF (BAF) complex and contribute to upregulation of genes controlled by this complex [12]. We also demonstrate overexpression of a high number of genes implicated in Wnt signaling. In addition, we further confirm the involvement of the Wnt pathway in LGESS by demonstrating concordant expression of β-catenin and Lef1 in these tumors. Our findings suggest that the activation of Wnt signaling pathway in LGESS may be triggered directly by the fusion proteins and offer a further rationale for exploring drugs that target Wnt pathway in this tumor.

2. Materials and methods

2.1. Immunohistochemistry

Five-micrometer sections of formalin-fixed, paraffin-embedded (FFPE) tissues were used for hematoxylin and eosin (H&E) and immunohistochemical (IHC) staining by avidin-biotin-peroxidase complex method. Expression of β-catenin and Lef1 was evaluated using following monoclonal antibodies: for β-catenin — clone 14, 1:25 dilution, Cell Marque, catalog no. 224M-16, and for Lef1 — clone EPR2029Y, 1:200 dilution, Abcam, catalog no. ab137872. Appropriate positive and negative controls were run in parallel. The percentage of cells showing β-catenin and Lef1 staining in the nuclei were estimated in each core at 400× magnification. TA-349 comprised duplicate cores from each LGESS and LMS tumor specimen, and TA-166 and TA-201 comprised a single core of each LMS and LGESS tumor specimen, respectively. TA-201 comprises 131 cores from primary and metastatic LMS tumors from 89 patients.

Tissue microarrays containing LGESS and LMS were scored blinded to diagnosis (MvdR). The images from tissue microarrays described in this study are publicly available through the Stanford Tissue Microarray Database (https://tma.im, TA-166, TA-201 and TA-349).

2.2. Gene expression dataset analysis

2.2.1. Microarray data analysis

Two gene expression datasets of Przybyl et al. (GSE85383) and Davidson et al. (raw data obtained from the authors) were used in this study. Expression values from different probes mapped to the same gene were collapsed using avereps function from limma package. Differential expression analyses of microarray data were performed in R Studio (built: R 3.0.2). Differential expression analysis was performed using significance analysis of microarrays (SAM) (samr package) [25] on quantile-normalized expression values filtered for SD > 100 and $\log_2$ transformed (with seed value 1,234,567 and estimated false detection rate of 0.05). SAM output was filtered for genes with fold change <−2 and >2.

2.2.2. 3SEQ data analysis

Differential expression analysis of the data published by Lee et al. (LGESS and LMS samples included in GSE54734) was performed in R Studio (built: R 3.0.2). Raw reads counts were used as the input for the SAMseq function in the samr package (with seed value 1,234,567 and estimated false detection rate of 0.05) [26]. SAMseq output was filtered for genes with fold change <−2 and >2.

Functional annotation of genes identified as differentially expressed across different datasets was performed using ToppGene Suite and Enrichr. *p* values in ToppGene are calculated using the hypergeometric probability mass function. *p* values in Enrichr are calculated from the Fisher's exact test, which is a proportion test that assumes a binomial distribution and independence for probability of any gene belonging to any set [27,28].

2.3. Statistical methods

The code applied for calculation of significance of the overlapping genes identified in LGESS is included in the Supplementary File 1 and at https://github.com/kidzik/geneoverlap. We derived a custom statistical method to evaluate the significance of the number of genes overlapping between at least 2 of 3 datasets. For the genes measured in each dataset, we first calculated the probability of each gene being identified as differentially expressed, assuming binomial sampling in each experiment and taking into account different numbers of genes measured by each platform. Next, we calculated the probability of each gene being identified as differentially expressed in at least two of three independent datasets. Based on these calculations, in LGESS, the number of overlapping genes expected by chance would be 98. Since the observed number of overlapping genes was 352, the resulting *Z* statistic is equal 26.11. Two-tailed *p*-value of this *Z* score is ≈ 0 (2.5×10^{-150}).

The analysis of contingency tables was performed using Fisher's exact test with two-tailed *p* value.

3. Results

3.1. Consistent gene expression profiles of LGESS across three independent studies

LGESS are rare and only a limited number of these tumors have been studied to date. Three previously published gene expression datasets collectively included 20 LGESS and 20 uterine LMS cases [13–15]. Here we report a cross-platform meta-analysis of these datasets to compare gene expression profiles of LGESS to LMS. We re-analyzed these three datasets using the samr package [we applied SAM for microarray data published by Przybyl et al. [13] and Davidson et al. [15], and SAMseq for RNA sequencing data published by Lee et al. [14]] to identify differentially expressed genes called at the same fold change and false detection rate cut-offs.

When compared to uterine LMS, LGESS tumors were characterized by 2262 differentially expressed genes in the dataset of Przybyl et al.

Please cite this article as: J. Przybyl, et al., Gene expression profiling of low-grade endometrial stromal sarcoma indicates fusion protein-mediated activation of the Wnt signali..., Gynecol Oncol (2018), https://doi.org/10.1016/j.ygyno.2018.03.007