Macrophage infiltration and genetic landscape of undifferentiated uterine sarcomas

Joanna Przybyl,1,2,3 Magdalena Kowalewska,2,4 Anna Quattrone,3 Barbara Dewaele,3 Vanessa Vanspauwen,1 Sushama Varma,1 Sujay Vennam,1 Aaron M. Newman,5,6 Michal Swierniak,7 Elwira Bakul-Walewska,6 Janusz A. Siedlecki,2 Mariusz Bidzinski,9,10 Jan Cools,11 Matt van de Rijn,1 and Maria Debiec-Rychter3

1Department of Pathology, Stanford University School of Medicine, Stanford, California, USA. 2Department of Molecular and Translational Oncology, Maria Sklodowska-Curie Memorial Cancer Center and Institute - Oncology Center, Warsaw, Poland. 3Department of Human Genetics, KU Leuven and University Hospitals Leuven, Leuven, Belgium. 4Department of Immunology, Biochemistry and Nutrition, Medical University of Warsaw, Warsaw, Poland. 5Institute for Stem Cell Biology and Regenerative Medicine, 6Department of Medicine, Division of Oncology, Stanford Cancer Institute, Stanford University, Stanford, California, USA. 7Human Cancer Genetics, Center of New Technologies, CENT, University of Warsaw, Warsaw, Poland. 8Department of Pathology, 9Department of Gynecologic Oncology, Maria Sklodowska-Curie Memorial Cancer Center and Institute - Oncology Center, Warsaw, Poland. 10The Faculty of Medicine and Health Sciences, Jan Kochanowski University, Kielce, Poland. 11KU Leuven and Flanders Interuniversity Institute for Biotechnology (VIB), Leuven, Belgium.

Endometrial stromal tumors include translocation-associated low- and high-grade endometrial stromal sarcomas (ESS) and highly malignant undifferentiated uterine sarcomas (UUS). UUS is considered a poorly defined group of aggressive tumors and is often seen as a diagnosis of exclusion after ESS and leiomyosarcoma (LMS) have been ruled out. We performed a comprehensive analysis of gene expression, copy number variation, point mutations, and immune cell infiltrates in the largest series to date of all major types of uterine sarcomas to shed light on the biology of UUS and to identify potential novel therapeutic targets. We show that UUS tumors have a distinct molecular profile from LMS and ESS. Gene expression and immunohistochemical analyses revealed the presence of high numbers of tumor-associated macrophages (TAMs) in UUS, which makes UUS patients suitable candidates for therapies targeting TAMs. Our results show a high genomic instability of UUS and downregulation of several TP53-mediated tumor suppressor genes, such as NDN, CDH11, and NDRG4. Moreover, we demonstrate that UUS carry somatic mutations in several oncogenes and tumor suppressor genes implicated in RAS/PI3K/AKT/mTOR, ERBB3, and Hedgehog signaling.

Introduction

The most common uterine sarcomas are leiomyosarcomas (LMS) that originate in the muscular wall of the uterus and tumors derived from endometrial stroma, the tissue supporting the endometrial glands. The classification of endometrial stromal sarcomas has evolved over the past decade and currently distinguishes 3 tumor subtypes: low- and high-grade endometrial stromal sarcomas (LG and HG ESS, respectively) that harbor specific chromosomal translocations and undifferentiated uterine sarcomas (UUS). UUS is a poorly defined entity that is often seen as a diagnosis of exclusion after ESS and leiomyosarcoma (LMS) have been ruled out. The least frequent of all pathological subtypes of uterine sarcomas, UUS have been barely studied. The largest series describing solely the clinical features of UUS included only 21 patients, while most of the remaining series reported only isolated cases, usually observed in a single institution over long periods of 20–30 years. Previous molecular studies of UUS included no more than 4 tumors. Surgery is the therapy of choice for the treatment of all subtypes of endometrial stromal tumors. UUS are associated with very poor prognosis and, while most UUS patients present with high-stage disease, even
those with stage I disease usually die within 2 years from diagnosis (2, 3). LG ESS frequently express estrogen and progesterone receptors (ER and PR, respectively) and may respond to adjuvant hormonal therapy. HG ESS may also express ER and PR in LG compartments of the tumor. In contrast, UUS are mostly negative for ER and PR, and they are often treated with combinations of gemcitabine/docetaxel and ifosfamide/doxorubicin that result only in partial responses, usually of very short duration (2). Little is known about the genetics of UUS beyond their complex chromosomal aberrations and the presence of $TP53$ mutations in a subset of cases (7, 8, 11). Recent studies evaluating the expression of selected tyrosine kinase receptors in ESS and UUS suggest that these tumors are unlikely to respond to tyrosine kinase inhibitors (12–14). Thus, identification of new therapeutic options remains a priority to improve the outcomes for UUS patients.

Here, we describe the first, to our knowledge, integrated analysis of gene expression, copy number variation, point mutations, and immune cell infiltrates in the largest series of 4 different subtypes of uterine mesenchymal tumors. Our study provides a comprehensive insight into the genetic landscape of UUS and shows that this group should not be seen as a mixture of heterogeneous tumors. Through a combination of genomic and gene expression profiling of 8 UUS and further validation of selected findings in an independent group of 11 UUS cases, our study identifies therapeutic targets in these aggressive tumors.

**Results**

**Gene expression profiling identifies UUS as tumors with a distinct molecular profile from other uterine sarcomas.** Gene expression profiling by microarrays was performed on a training cohort of 8 UUS, 9 LG ESS, 4 HG ESS, and 4 uterine LMS (Table 1). Classification of these tumors was based on extensive histological sampling, evaluation of multiple immunohistochemical markers (i.e., CD10, ER, PR, SMA, h-caldesmon, and desmin), and fluorescence in situ hybridization (FISH) analysis for the presence of subtype specific chromosomal rearrangements. All LG ESS cases carried $JAZF1$, $PHF1$, or $MBTD1$ rearrangements; all HG ESS cases had $YWHAE-NUTM2A/B$ translocations; and UUS tumors did not reveal either LG ESS– or HG ESS–specific translocations.

Unsupervised hierarchical clustering and principal component analysis (PCA) revealed that UUS, ESS, and LMS formed distinct subgroups (Figure 1, A and B). PCA also showed that the first principal component separates UUS from the 2 subtypes of ESS and LMS, indicating that UUS have a markedly distinct gene expression profile from the other uterine sarcomas (Figure 1, B and C). These results contradict the previous notion that UUS likely encompass a heterogeneous mixture of dedifferentiated or poorly differentiated homologous sarcomas (including LMS and ESS).

Next, we performed a cross-platform validation of the microarray-based findings by RNA sequencing (RNA-seq). Gene expression profiles of LMS have been described previously in several studies (15–17). Therefore, for the in-depth gene expression analysis combining microarray and RNA-seq platforms, we focused solely on the 3 tumor types derived from endometrial stroma. RNA-seq was performed on the same cases profiled by gene expression microarrays that had sufficient RNA yield (i.e., 8 UUS, 8 LG ESS,
Unsupervised hierarchical clustering based on RNA-seq data confirmed that UUS form a cluster separate from LG and HG ESS (Supplemental Figure 1; supplemental material available online with this article; https://doi.org/10.1172/jci.insight.94033DS1). Similar to the microarray analysis, LG and HG ESS clustered together and separately from UUS and LMS; therefore, for the subsequent analyses, we treated these cases as one group of translocation-associated ESS.

Gene expression profile of UUS is strongly driven by M2 macrophage infiltration. In order to characterize gene expression profiles of UUS and ESS, we employed highly stringent analysis criteria to identify differentially expressed genes overlapping between RNA-seq (using DESeq2, edgeR, and significance analysis of sequencing data [SAMseq]) and microarray data (using significance analysis of microarrays [SAM]) (18–21). We identified 406 genes that were differentially expressed between UUS and ESS in both microarray data and RNA-seq data (Supplemental Figure 2 and Supplemental Tables 1 and 2). Genes encoding immune cell markers and cytokines were among the most overrepresented gene families showing upregulated expression in UUS (according to the gene family resource in the MSigDB [Molecular Signatures Database]; http://software.broadinstitute.org/gsea/msigdb) (Figure 2A, Supplemental Table 3, and Supplemental
In addition, ToppGene Suite analysis of the 216 genes overexpressed in UUS pointed to a strong overrepresentation of genes expressed by myeloid cells, with 31 of 50 top coexpression annotations being associated with myeloid subsets.

To systematically evaluate the presence of tumor-infiltrating immune cells in UUS and ESS, we applied CIBERSORT, a machine learning approach to enumerate the proportions of distinct cell types in bulk tissue expression profiles. As input, CIBERSORT requires expression profiles that differentiate each cell type of interest, collectively termed a signature matrix. In this study, we used LM22, a signature matrix distinguishing 22 subsets of mature human hematopoietic subsets (22).

CIBERSORT deconvolution in microarray data reached statistical significance in all 8 UUS samples but in only 2 of 13 ESS samples ($P<0.05$) (Supplemental Table 4), indicating that UUS but not ESS may contain significant numbers of immune cells. Among hematopoietic cell subsets, CIBERSORT revealed higher fractions of myeloid cells in UUS compared with ESS, with a significant enrichment for M2 macrophages ($P = 0.004$).

**Figure 2. Identification of M2 macrophage infiltration in UUS.** (A) Top 3 gene families overrepresented in UUS include numerous macrophage-associated differentiation markers and chemokines (gene family classification based on Molecular Signatures Database, MSigDB). (B) CIBERSORT analysis of microarray data indicates significantly higher fractions of M2 macrophages in UUS ($n=8$) compared with ESS specimens ($n=13$). (C) Forty of 216 genes overexpressed in UUS in both microarray and RNA-seq data belong to the M2 macrophage–related CSF1 response signature. (D) Heatmap of log$_e$ (TPM+1) expression values of the 40 CSF1-response genes that distinguish UUS and ESS specimens. TPM, transcript per million.
We also observed a high correlation between M2 macrophage fractions estimated by CIBERSORT in microarray and RNA-seq datasets (Pearson’s $r = 0.79$, $P = 6 \times 10^{-5}$). Numerous genes encoding M2-specific markers (including $CD163$, $CD206$, $CD209$, $STAB1$, and $CCR2$) and genes involved in TAM-driven tumor progression, including cell invasion ($MMP9$, $MMP12$) and immunosuppression ($CCL13$, $CCL18$), were significantly overexpressed in UUS specimens (Figure 2A and Supplemental Table 1). Validation of gene expression profiling results was performed by quantitative reverse transcription PCR (qRT-PCR) for $CCL18$ and $LILRA6$ (Supplemental Figure 4).

One of the top upregulated genes in UUS encodes the protein tyrosine kinase receptor $CSF1R$. We have previously identified a CSF1-induced immune response signature in tenosynovial giant cell tumors that was later applied to study immune cell infiltrates in breast carcinoma (23, 24). This signature is driven by overexpression of CSF1 by tumor cells, resulting in a massive invasion of TAMs. We found that 36% (40 of 112) of the core CSF1 response signature genes were overexpressed in UUS (hypergeometric test, $P = 5 \times 10^{-65}$) (Figure 2, C and D).

To validate gene expression data, IHC was performed on the same tumors as were used for the gene expression profiling studies, using $CD163$, $CD68$, and $CCR2$ as TAM-specific markers (Figure 3A and Supplemental Figure 5). Staining for $CD163$ and $CD68$ confirmed heavy infiltration by M2 macrophages in UUS compared with ESS (Figure 4, G–I). With the exception of 1 tumor with equivocal staining (UUS_3), $CD163$ and $CD68$ expression in UUS specimens was seen solely in macrophages and not in the tumor cells. $CCR2$ staining also showed a significant difference between UUS and ESS; however, the counts of $CCR2^+$ cells were much lower than $CD163^+$ cells (Figure 4, G–L, and Supplemental Table 5). LG and HG ESS showed only scattered M2 macrophages expressing $CD163$ and $CCR2$ (Figure 4, G, H, J, and K). We further confirmed $CD163^+$ M2 macrophage infiltration in UUS in an independent cohort of 21 patients (11 UUS and 10 ESS) (Figure 3B).

Taken together, these data demonstrate a dense infiltration by TAMs in UUS, which strongly contributes to the gene expression profile of these tumors.

Genomic complexity and mutational landscape in UUS. Having characterized distinct transcriptional signatures in UUS and ESS, we next analyzed copy number and mutational profiles of endometrial stromal tumors. Based on the array comparative genomic hybridization (aCGH) results from 19 tumors (9 LG ESS, 3 HG ESS, and 7 UUS), we calculated the genomic index (GI), a measure of the extent of copy number alterations in each sample. GI was calculated as $A^2/C$, where $A$ is the total number of segmental gains and losses, and $C$ is the number of involved autosomal chromosomes (as described in ref. 25). LG and HG ESS were characterized by relatively simple genomic profiles (median GI = 0 and 5, respectively), while UUS presented with a very high genomic complexity (median GI = 57, range 20–81). The most frequent DNA copy number alterations in UUS are summarized in Figure 5 and Table 2.

In addition to copy number analysis, we performed whole exome sequencing (WES) on 3 UUS and 1 LG ESS, for which normal tissue adjacent to the tumor was also banked after surgery. In UUS,
we identified a median of 45 somatic single nucleotide variants (SNVs) and indels (range 37–45) per exome (Figure 6A). On average 33% (range 27%–38%) of these mutations were expressed and confirmed in RNA-seq data from corresponding UUS specimens (Supplemental Table 6). In the single LG ESS case analyzed by WES, we identified only 2 somatic variants in exonic regions that were not present in RNA-seq data, indicating that these variants were not expressed.

In UUS, the expressed variants were detected in 9 genes known to be frequently mutated in cancer (Figures 6, B and C), including TP53, PIK3CA, KRAS, and ERBB3 oncoproteins, as well as FBXW7, PTCH1, and ASXL1 tumor suppressor genes. Mutations in TP53, PIK3CA, KRAS, ERBB3, and FBXW7 were confirmed by Sanger sequencing in all 7 UUS cases with DNA available for validation. A potential role of phosphatidylinositol 3-kinase (PI3K) and RAS pathways was further supported by the functional enrichment analysis of genes upregulated in UUS, which showed overexpression of 15 and 25 genes regulated by PI3K complex and KRAS oncogenic signaling, respectively (Supplemental Tables 7 and 8). In addition, we found DNA copy number loss and/or loss of heterozygosity (LOH) together with low gene-expression level in 3 TP53-associated tumor suppressor genes, i.e., NDN, CDH11, and NDRG4 (Figure 6C and Supplemental Tables 9 and 10). For NDN and CDH11, mRNA downregulation was statistically significantly correlated with the presence of genomic aberrations in TP53 in UUS specimens (Fisher’s exact test, \( P = 0.003 \) and \( P = 0.013 \), respectively; Supplemental Tables 9 and 11).

This comprehensive analysis of somatic mutations, copy number alterations, and gene expression shows a high chromosomal instability and the possible implication of several known oncogenes and suppressor genes in the tumorigenesis of UUS.

**Discussion**

In this study, we describe molecular and biologic features of the largest series of UUS cases to date analyzed in a training/validation approach. Our results show that UUS harbor clinically relevant genetic aberrations and present with high infiltration of TAMs.

**M2 macrophages as therapeutic targets in UUS.** TAMs are known to exhibit predominantly M2 polarization and to orchestrate tumor initiation, invasion, metastasis, immunosuppression, and angiogenesis. In the vast majority of tumor types, TAMs are associated with poor outcome (26). Given the contribution of TAMs to nearly all stages of tumor progression, they are currently seen as a target for anticancer treatment. Most macrophage-targeted therapies are directed to either inhibit their recruitment to the tumor site,
suppress their survival, redirect M2 polarization to antitumor M1 phenotype, or block tumor-promoting activities. In an alternative approach, interfering with the CD47-SIRPα axis has been shown to improve the phagocytosis of tumor cells by macrophages and lead to tumor regression (27). In preclinical models, macrophage-targeted therapies were effective as a monotherapy or in combination with angiogenic inhibitors, adoptive cell transfer, and immune checkpoint inhibitors (reviewed in ref. 26). It has also been demonstrated that depletion of macrophages may improve response to chemo- and radiotherapy in preclinical models (28–30). These promising preclinical data have led to numerous clinical trials evaluating TAM-targeted therapies in various tumor types (26, 31).

Here, we demonstrate for the first time to our knowledge that UUS are densely infiltrated by TAMs, which opens new therapeutic venues in this aggressive disease. We suggest that UUS patients can be treated with already-registered macrophage-targeted drugs or may enter selected clinical studies. Chemokines such as CCL2 and cytokines such as CSF1 are major determinants of macrophage recruitment and functional polarization in tumors (32). In addition, MMP9 that is known to be involved in TAM-driven cell invasion was also highly expressed in UUS specimens. Therefore, based on our gene expression data, we suggest that UUS patients may benefit from therapies targeting MMP9-expressing macrophages and inhibiting CCL2-CCR2 and CSF1-CSF1R axes (Table 3). Based on the general properties of TAMs, therapies disrupting CD47-SIRPα axis may also be considered in UUS patients (Table 3).

While we found very high counts of CD163+ and CD68+ TAMs in UUS, less abundant CCR2+ cells were noted in the same specimens. CD163 is a M2 macrophage-specific marker, while CD68 is considered to be expressed both in M1 and M2 macrophages. There was a very high correlation between the numbers of CD68+ and CD163+ macrophages in ESS and UUS specimens (Pearson’s r = 0.98, P < 0.00001, based on 12 specimens evaluated for both markers); therefore, we assume that M2 was the dominant phenotype in UUS. It has been proposed that, within the M2 macrophage group, there are several different cell subpopulations (M2a, M2b, M2c) that express different markers and secrete different chemokines. CD163 expression has been associated with both M2a and M2c phenotypes, whereas CCR2 appears to be associated only with M2c macrophages (32, 33). Our data suggest that the TAMs in UUS may predominantly belong to the M2a subpopulation.

Another interesting property of TAMs is their involvement in angiogenesis in selected tumor types. We previously showed that, in nongynecological LMS, there is a correlation between increased angiogenesis and the expression of CSF1 in the tumor cells and the presence of CD163+ TAMs (34). However, no such correlation was found in uterine LMS, possibly due to already high vascularity in the uterus (34). Similarly, in UUS patients, we did not observe a correlation between CD34 protein levels and CD163 expression (data not shown).

Table 2. Minimal overlapping regions of the most frequent copy number changes in 7 UUS cases analyzed by aCGH+SNP microarrays

<table>
<thead>
<tr>
<th>% UUS with aberration (n = 7)</th>
<th>Chr</th>
<th>Cytoband</th>
<th>Gain/Loss</th>
<th>Start</th>
<th>Stop</th>
<th>Size (Mb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>57%</td>
<td>chr1</td>
<td>q21.1 - q31.3</td>
<td>Gain</td>
<td>144009907</td>
<td>196309360</td>
<td>52.30</td>
</tr>
<tr>
<td>57%</td>
<td>chr1</td>
<td>q32.1 - q32.2</td>
<td>Gain</td>
<td>204069700</td>
<td>208980594</td>
<td>4.91</td>
</tr>
<tr>
<td>57%</td>
<td>chr3</td>
<td>q22.3 - q29</td>
<td>Gain</td>
<td>136732261</td>
<td>197837049</td>
<td>61.10</td>
</tr>
<tr>
<td>71%</td>
<td>chr3</td>
<td>q26.2 - q29</td>
<td>Gain</td>
<td>168381936</td>
<td>197837049</td>
<td>29.46</td>
</tr>
<tr>
<td>57%</td>
<td>chr6</td>
<td>q21 - q22.31</td>
<td>Gain</td>
<td>108529570</td>
<td>125530734</td>
<td>17.00</td>
</tr>
<tr>
<td>57%</td>
<td>chr7</td>
<td>q31.31 - q32.1</td>
<td>Gain</td>
<td>11836452</td>
<td>128624248</td>
<td>10.26</td>
</tr>
<tr>
<td>57%</td>
<td>chr8</td>
<td>q11.1 - q24.3</td>
<td>Gain</td>
<td>47735940</td>
<td>142909763</td>
<td>95.17</td>
</tr>
<tr>
<td>57%</td>
<td>chr12</td>
<td>p12.3 - p11.21</td>
<td>Loss</td>
<td>5714477</td>
<td>32799215</td>
<td>17.08</td>
</tr>
<tr>
<td>71%</td>
<td>chr16</td>
<td>q11.2 - q24.3</td>
<td>Loss</td>
<td>46441545</td>
<td>9019248</td>
<td>43.66</td>
</tr>
<tr>
<td>71%</td>
<td>chr16</td>
<td>q22.1 - q24.1</td>
<td>Loss</td>
<td>691849</td>
<td>85003869</td>
<td>15.25</td>
</tr>
<tr>
<td>57%</td>
<td>chr20</td>
<td>q11.21 - q13.33</td>
<td>Gain</td>
<td>29652452</td>
<td>62908674</td>
<td>33.26</td>
</tr>
<tr>
<td>71%</td>
<td>chr20</td>
<td>q13.12 - q13.33</td>
<td>Gain</td>
<td>46288190</td>
<td>62895927</td>
<td>16.61</td>
</tr>
<tr>
<td>71%</td>
<td>chrX</td>
<td>p22.33 - p22.31</td>
<td>Loss</td>
<td>316344</td>
<td>6013484</td>
<td>5.70</td>
</tr>
<tr>
<td>57%</td>
<td>chrX</td>
<td>p22.33 - p11.3</td>
<td>Loss</td>
<td>316344</td>
<td>42759090</td>
<td>42.44</td>
</tr>
</tbody>
</table>

Start/stop positions refer to hg19. Minimum number of probes in region set to 500. Chr, chromosome; Mb, megabase.
Clinically relevant oncogenic pathways in UUS. Our data from genomic and gene expression profiling of UUS indicate that there are clinically actionable molecular alterations in UUS. According to the Broad Institute’s TARGET (Tumor Alterations Relevant for GEnomics-driven Therapy) database (version 3), the presence of somatic mutations in PIK3CA, KRAS, ERBB3, FBXW7, and PTCH1 may qualify UUS patients for selected targeted treatments. While this work may be seen as an exploratory study, our results point to a need for an extended screening for clinically relevant mutations in UUS that may translate into a direct clinical benefit for these patients.

Our results show that the PI3K pathway may be activated in UUS patients through mutations and/or copy number gains of PIK3CA or copy number loss/LOH in PTEN. Such alterations in the components of PI3K pathway have been associated with poor prognosis and HG histology in several types of carcinoma (35–39). The possible activation of the PI3K pathway in a subset of UUS is further supported in our data by increased mRNA expression of numerous PI3K-regulated genes. Furthermore, 4 UUS patients carried homozygous missense mutation or had copy number loss with concordant LOH of FBXW7, a tumor suppressor gene that regulates mTOR signaling (40). Biallelic inactivation of FBXW7 may predict sensitivity to mTOR inhibitors, as well as resistance to anti-tubulin chemotherapy. These findings suggest that UUS may be sensitive to PI3K/AKT/MTOR inhibitors.

Hotspot mutations and copy number gains in the KRAS locus were also frequently detected in our group of UUS patients. Coexisting KRAS point mutations and copy number gains have been described in a limited subset of tumors, either in advanced stage or in association with poor survival and/or undifferentiated histology (41–43). UUS tumors showed increased mRNA expression of multiple targets activated by oncogenic KRAS signaling, as well as downregulation of KRAS-associated tumor suppressor PAWR (44). Activation of this pathway may predict sensitivity to MEK inhibitors in a subset of UUS patients.

Our findings also implicate a potential role of Hedgehog signaling in UUS based on a decreased expression of PTCH1 tumor suppressor gene or a possible activation of UBR5 gene (45, 46). According to TARGET database, PTCH1 mutations may predict sensitivity to Hedgehog inhibitors, such as vismodegib.

In addition, 2 UUS patients carried point mutations and/or copy number gains in ERBB3 locus, which may indicate sensitivity to pertuzumab and other ERBB3 inhibitors.

Figure 5. The most frequent chromosomal copy number alterations in UUS. Only chromosomes altered in more than 50% of cases are shown.
These findings point to previously unrecognized therapeutic venues for UUS patients. Unfortunately, preclinical validation of these therapies in UUS is currently problematic due to lack of in vitro or in vivo models of this malignancy. Nevertheless, our results show that UUS patients harbor clinically actionable mutations and could benefit from genomic testing offered by increasing number of hospitals and diagnostic companies.

**Spectrum of genomic alterations in UUS.** In this study, we observed genomic aberrations in *TP53* in 86% (6 of 7) of UUS cases. While *TP53* alterations have been previously reported in a subset of UUS (7), our study is the first to identify putative tumor suppressor genes that may be affected by *TP53* in UUS, i.e., a DNA damage response gene *NDN* (47) and *CDH11* that is epigenetically inactivated in many types of carcinomas (48). Since *NDN* loss has been reported in aggressive serous ovarian cancer cell lines carrying *TP53* mutations (49), we hypothesize that *NDN* is a *TP53*-target gene in UUS, as well. It has been previously described that *CDH11* downregulation may also be induced upon expression of mutant *TP53* in mammary epithelial cells (50). In addition, *NDRG4* (N-Myc downstream-regulated gene 4 protein) is another *TP53*-dependent tumor suppressor gene (51) that was downregulated in UUS compared with ESS; however, it was not significantly associated with an aberrant *TP53* in the analyzed tumor specimens.

Our analysis confirmed a high genomic complexity of UUS tumors (Figure 5 and Table 2), which was previously reported in 2 studies (8, 11). However, copy number aberrations (CNAs) identified in UUS are not specific to this tumor type and are also frequently present in tumors considered in differential diagnosis of UUS, including uterine carcinosarcomas (e.g., gains of 1q22, 3q26.2, 8q11.23, 20q11.21) (52) and uterine LMS (e.g., loss of 16q) (3). Similarly, SNVs and indels found in UUS affect genes that are frequently mutated across a wide spectrum of high-grade tumors. In UUS, we identified a median of 45 somatic SNVs and indels per exome, which is a mutation load comparable with other types of high-grade gynecological tumors (53, 54).

**Advances in understanding genetic background of UUS.** The histological classification of uterine sarcomas has evolved in recent years. Based on the discovery of a *YWHAE-NUTM2A/B* chromosomal translocation, the latest WHO classification identified a new subtype of ESS (i.e., HG ESS) that was previously described...
as “undifferentiated endometrial sarcomas with nuclear uniformity” (7). Our study provides further evidence showing that HG ESS significantly differ from UUS in their genetic profile and immune cell infiltration.

There are 4 previously published reports that used high-throughput methods to characterize UUS, but these studies used low numbers of UUS cases (2–4 cases) (8–11) and often used nonstandardized techniques, such as home-made gene expression microarrays covering 3,600 genes (9) or comparative genomic hybridization (11). Moreover, none of the previous molecular studies directly compared UUS to all 3 major types of uterine mesenchymal tumors, including LG ESS, HG ESS, and LMS, despite the hypothesis that UUS may be derived from these tumors. The prior studies also failed to show the clinically relevant features that characterize UUS. By comparing the largest series of UUS cases to date with the other 3 main classes of uterine sarcomas, we provide an insight into molecular features of UUS.

In conclusion, we performed a comprehensive genomic and gene expression profiling of endometrial stromal tumors, which led to the identification of clinically relevant features of UUS. A combination of differential gene expression analysis and computational approaches for inferring immune cell fractions demonstrated that the genetic profile of UUS is strongly driven by infiltrating TAMs. Most importantly, our study prompts reconsideration of the prior “diagnosis of exclusion” view on UUS. We demonstrate that gene expression profile of UUS is distinct from other uterine mesenchymal tumors and propose therapeutic avenues for UUS patients, by targeting TAMs and oncogenic pathways activated in these tumors.

<table>
<thead>
<tr>
<th>Therapy</th>
<th>Macrophage-related target</th>
<th>Ref</th>
<th>Drug mechanism</th>
<th>Current indication/Development phase</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trabectedin</td>
<td>CCR2-CCL2 axis</td>
<td>76</td>
<td>Inhibit macrophage recruitment</td>
<td>Europe: Advanced soft tissue sarcoma and ovarian carcinoma USA: Selected soft tissue sarcomas - unresectable or metastatic liposarcoma and leiomyosarcoma</td>
<td>77 78</td>
</tr>
<tr>
<td>Dasatinib</td>
<td>MMP9</td>
<td>79</td>
<td>Suppress TAM survival</td>
<td>Chronic myeloid leukemia (CML) and Philadelphia chromosome-positive acute lymphoblastic leukemia (Ph+ ALL)</td>
<td>80 81</td>
</tr>
<tr>
<td>Zoledronic acid</td>
<td>MMP9</td>
<td>82, 83</td>
<td>Suppress TAM survival</td>
<td>Hypercalcemia of malignancy, multiple myeloma, bone metastases from solid tumors</td>
<td>84 85</td>
</tr>
<tr>
<td>PLX3397 (small molecule inhibitor)</td>
<td>CSF1-CSF1R axis</td>
<td>86, 87</td>
<td>Suppress TAM survival, Inhibit macrophage recruitment</td>
<td>Phase I studies in tenosynovial giant cell tumors completed</td>
<td>87</td>
</tr>
<tr>
<td>RG7155 emactuzumab (mAb)</td>
<td>CSF1-CSF1R axis</td>
<td>86, 87</td>
<td>Suppress TAM survival, Inhibit macrophage recruitment</td>
<td>Phase I/II, solid tumors</td>
<td>NCT02452424</td>
</tr>
<tr>
<td>LY3022855 (IMC-CS54) (mAb)</td>
<td>CSF1-CSF1R axis</td>
<td>86, 87</td>
<td>Suppress TAM survival, Inhibit macrophage recruitment</td>
<td>Phase II trial in prostate cancer completed</td>
<td>NCT00992186</td>
</tr>
<tr>
<td>AMG820 (mAb)</td>
<td>CSF1-CSF1R axis</td>
<td>86, 87</td>
<td>Suppress TAM survival, Inhibit macrophage recruitment</td>
<td>Phase I, solid tumors</td>
<td>NCT02265536, NCT01346358</td>
</tr>
<tr>
<td>CNTO 888 carlumab (mAb)</td>
<td>CCR2-CCL2 axis</td>
<td>76</td>
<td>Inhibit macrophage recruitment</td>
<td>Phase II trial in metastatic cancer completed</td>
<td>NCT01015560</td>
</tr>
<tr>
<td>MLN1202 (mAb)</td>
<td>CCR2-CCL2 axis</td>
<td>76</td>
<td>Inhibit macrophage recruitment</td>
<td>Phase II study in metastatic cancer completed</td>
<td>NCT024345408</td>
</tr>
<tr>
<td>CCX872 (small molecule inhibitor)</td>
<td>CCR2-CCL2 axis</td>
<td>76</td>
<td>Inhibit macrophage recruitment</td>
<td>Phase I, pancreatic cancer</td>
<td>NCT02216409</td>
</tr>
<tr>
<td>HuSF9-G4 (mAb)</td>
<td>CD47-SIRPα axis</td>
<td>27, 89</td>
<td>Induction of specific tumor cell phagocytosis</td>
<td>Phase I, solid tumors and acute myeloid leukemia</td>
<td>NCT02641002, NCT02367196</td>
</tr>
<tr>
<td>CC-90002 (mAb)</td>
<td>CD47-SIRPα axis</td>
<td>27, 89</td>
<td>Induction of specific tumor cell phagocytosis</td>
<td>Phase I, acute myeloid leukemia and myelodysplastic syndromes and hematologic neoplasms</td>
<td>NCT02641002, NCT02367196</td>
</tr>
<tr>
<td>TTI-621 (SIRPαFc)</td>
<td>CD47-SIRPα axis</td>
<td>27, 89</td>
<td>Induction of specific tumor cell phagocytosis</td>
<td>Phase I, hematologic neoplasms</td>
<td>NCT02663518</td>
</tr>
</tbody>
</table>
Methods

Patients

A total of 42 endometrial stromal tumors (EST) from 41 patients and 4 uterine LMS were included in the study. Twenty patients were treated at the Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, 18 patients were treated at the KU Leuven and University Hospitals, and 7 patients were enrolled at Stanford University Medical Center. The study included frozen and formalin-fixed paraffin embedded (FFPE) tissues from 19 UUS, 16 LG ESS, 7 HG ESS, and 4 LMS collected with the approval of the respective institutional research boards. Studied material included 33 tumors collected before chemotherapy and 6 tumors collected after chemotherapy (1 LG ESS, 2 LMS, and 3 UUS), and the treatment status at the time of sampling was not specified for 7 tumor specimens.

Detailed clinical follow-up was available for 21 EST patients included in the genomic and gene expression studies (training cohort, Table 1 and Supplemental Table 12). Median follow-up was 47 months (range 1–276 months). Studied material included 17 primary and 4 recurrent tumor samples. Tumor grade was assigned according to the International Federation of Gynecology and Obstetrics (FIGO) 2009 staging criteria for uterine sarcomas (55). FFPE sections from another 21 ESTs were used only for immunohistochemical studies (validation cohort, Table 1 and Supplemental Table 12).

Forty-one of 42 EST were examined for the presence of JAZF1, YWHAE, and BCOR rearrangements by FISH and/or RT-PCR. Selected LG ESS cases were also examined for the presence of PHF1 rearrangements. All LG and HG ESS cases included in this study carried subtype-specific translocations, and none of the UUS cases disclosed any gene rearrangements.

Microarray gene expression data from 11 cases described in our study (4 LG ESS, 3 HG ESS, and 4 UUS) have been previously included in a report by Dewaele et al. (ref. 56, GSE46285, SubSeries GSE45783), to demonstrate that LG ESS cases with novel MBTD1-CXorf67 fusion clustered with the cases carrying classical JAZF1-SUZ12 fusions. It was indicated in that report that a detailed gene expression profiling analysis would be a subject of a separate study (56). Eight cases (3 LG ESS and 5 HG ESS) have been previously examined using different analytical methods and included in the study of Lee et al. (57). Selected patients were previously described in case reports that did not include any molecular studies, as indicated in Supplemental Table 12.

Histopathology and IHC

Histopathological examination was performed on FFPE tissues. Sections (5 μm) were used for routine H&E and IHC staining by avidin-biotin-peroxidase complex method. For IHC, the following monoclonal antibodies were utilized: ER (clone SP1, 1:60, Thermo Fisher or clone SP1, prediluted, Ventana), desmin (1:20, ICN Pharmaceuticals or clone D33, 1:40, Leica), CD10 (clone 56C6, 1:10, Leica), h-caldesmon (clone h-CD, 1:25, Dako), SMA (clone 1A4, 1:200, Cell Marque), CD163 (clone 10D6, 1:50, Leica), CD68 (clone KP1, 1:1600, Dako), and CCR2 (clone 48607, 1:200, R&D Systems). Appropriate positive and negative controls were run in parallel. CD163+, CD68+, and CCR2+ cells were counted in 10 randomly selected fields at ×400 magnification.

The morphological features of the 3 types of endometrial stroma-derived tumors were consistent with those described in the literature (3–5), with LG ESS showing a monomorphic population of cells, with little nuclear pleomorphism. The cells in HG ESS were more pleomorphic in appearance, whereas the UUS cases showed at least focal HG nuclear changes and tumor necrosis. Representative examples of each histologic subtype are shown in Figure 4. In 31 of 42 investigated EST, the histologic diagnosis was supported by the presence or absence of characteristic translocations reported previously in ESS.

FISH

FISH analysis was performed on 18 UUS specimens and 23 LG and HG ESS specimens. Bacterial artificial chromosome (BAC) DNA probes differentially labeled with SpectrumGreen (SG) or SpectrumOrange (SO) (Abbott Molecular) were used for dual-color interphase FISH for detection of gene fusion/rearrangement on the 4- to 5-μm paraffin tumor sections. BACs were obtained from the BACPAC Resource Center (Supplemental Table 13; http://bacpac.chori.org). BAC clones were selected based on their location in the NCBI Map Viewer build 36.3 (http://www.ncbi.nlm.nih.gov/projects/mapview), and the UCSC Human Genome Browser Gateway GRCh37/hg19 (https://genome.ucsc.edu/cgi-bin/hgGateway). DNA
isolation, probe labeling, and hybridization were performed as previously described (58). For analysis, 100 nuclei were scored manually for each set of probes using an Axioplan fluorescence microscope and analyzed with CYTOVISION software. A cut-off of 20% was established for a positive result.

**Total RNA extraction**

Total RNA was extracted from frozen EST and LMS specimens using the miRNeasy Mini Kit (Qiagen), including DNase treatment using RNase-free DNase Set (Qiagen). RNA quality was verified using Bio-Rad Experion system.

**Gene expression microarrays**

Gene expression analysis was performed on 21 EST (9 LG, 4 HG ESS, and 8 UUS) and 4 uterine LMS using SurePrint G3 Human GE 8x60K microarrays (G4851A), following one-color microarray-based expression analysis protocol (Agilent Technologies). Total RNA (25 ng) was reverse transcribed into cDNA by incorporating a T7 oligo-dT promoter primer prior to the generation of fluorescent complementary RNA (cRNA) using an Agilent LowInput Quick Amp Labeling Kit (Agilent Technologies). Microarray slides were scanned using the Agilent G2565A DNA Microarray Scanner (Agilent Technologies). Image analysis was done using the Feature Extraction V10.10.1.1 software (Agilent Technologies).

**RNA-seq**

Paired-end whole-transcriptome sequencing was performed on 19 EST (8 LG, 3 HG ESS and 8 UUS). Samples were prepared with the Illumina TruSeq RNA Sample Preparation Kit. The insert sizes of the libraries were assessed using Agilent Technologies 2100 Bioanalyzer. Validated DNA libraries were pooled, clustered on the cBot cluster generation station (Illumina), and sequenced on a HiSeq 2000 platform (Illumina). Prepared libraries were sequenced using HiSeq 2000 (Illumina) operated in paired-end 2 × 100 bp mode. Reads were quality-filtered using a standard Illumina pipeline.

The paired-end sequence reads were aligned to the human genome (hg19) with Subread (version 1.5.0) using default settings (59). Reads were summarized using default settings of the featureCounts program for paired-end reads (60). The reads were annotated to 60,308 coding and noncoding genes using a Gencode v24lift37 GTF file. Single nucleotide polymorphisms were identified using default settings of the exactSNP program within the Subread package. Variants were annotated with wANNOVAR web interface (http://wannovar.wglab.org) (61). Mutations were filtered for < 0.1% presence in any population included in 1,000 Genomes or ExAC studies, exonic location, and nonsynonymous or stopgain effect. For variants identified in both WES and RNA-seq data, the presence on forward and reverse strands was further verified using integrative genomics viewer (IGV).

**qRT-PCR**

For the qPCR analysis, samples of total RNA extracted from 19 EST were used. cDNA synthesis was carried out from 0.5 μg of RNA with the RT² First Strand Kit (Qiagen). Custom PCR arrays (SA Biosciences) were used to simultaneously examine the mRNA levels of CCL18 (chemokine [C-C motif] ligand 18, NM_002988) and LILRA6 (leukocyte immunoglobulin-like receptor, subfamily A [with TM domain], member 6, NM_024318), according to the manufacturer’s protocol. The arrays also included primers for 2 housekeeping genes. qPCR analysis was performed with the RT² SYBR Green ROX qPCR Master Mix (Qiagen) in the 7500 Fast Real-Time PCR System (Applied Biosystems) according to the manufacturer’s instructions (Qiagen).

**aCGH**

Genomic DNA was extracted from 19 frozen tumor specimens (9 LG, 3 HG ESS, and 7 UUS) using Blood & Tissue DNAeasy kit (Qiagen) or High Pure PCR Template Preparation Kit (Roche Diagnostics) with the kit-specific RNase treatment. Genomic DNA (1 μg) was hybridized to 4x180K CGH+SNP whole-genome microarrays (G4890A) (Agilent Technologies), according to the manufacturer’s protocol. Microarray slides were scanned using Agilent G2565A DNA Microarray Scanner, and image analysis was performed using Feature Extraction V10.10.1.1 software (Agilent Technologies).

aCGH data were analyzed using Agilent Genomic Workbench 7.0.4.0 software (Agilent Technologies). The ADM-2 algorithm was applied to identify DNA CNAs. A copy number loss was defined as a log, ratio < −0.25, and a copy number gain was defined as a log, ratio > 0.25, with at least 500 probes in a
region. An amplification was defined as log<sub>2</sub> ratio > 1.5. A female HapMap genotype (European Female, NA12878_V1) was used as a reference for the detection of LOH.

GI for each tumor specimen was calculated as follows: GI = A<sup>2</sup>/C, where A is the total number of alterations (segmental gains and losses of at least 500 probes in a region), and C is the number of involved autosomal chromosomes (25).

Whole exome sequencing
Genomic DNA from 1 LG ESS and 3 UUS tumors was extracted from frozen material using the same method as for aCGH analysis. Fragments of normal uterine tissue, distant from the tumor site, sampled from the same patients were used as control specimens. Genomic DNA was sheared by sonication, and libraries were prepared using the TruSeq DNA Library Preparation Kit (Illumina). DNA sequencing libraries were enriched with the SeqCap EZ Human Exome Library v3.0 (Roche, NimbleGen). Validated DNA libraries were pooled and sequenced using 2 × 100 bp paired-end mode on the HiSeq2000 platform (Illumina). Median target coverage was 55× (range 47×–66×).

The paired-end sequence reads were aligned to the human genome (GRCh37) with the Burrows-Wheeler Aligner (BWA; version 0.7.13) using BWA-MEM algorithm with default settings (62). SAMtools (version 1.3) (63) was used for converting SAM to BAM format, and sorting and indexing alignments. Picard (version 1.96) was used for duplicate reads removal. The GATK framework (version 3.3-0) (64) was used for the local realignment and base call recalibration. Somatic variants were identified using an automated ensemble approach requiring at least 2 of 4 variant callers to identify each SNV and indel. The variant callers used were Mutect (version 2.7-1) (65), VarScan2 (version 2.3.7) (66), FreeBayes (version 0.9.21-7) (67), and VarDictJava (version 1.4.1) (68). Variants were annotated with wANNOVAR web interface (http://wannovar.wglab.org) (61).

Somatic mutations were identified as alternate allele frequency > 20% in tumor specimen and < 5% in adjacent normal tissue. Somatic variants were additionally filtered for < 1% frequency in any population in 1000 Genomes and ExAC databases, exonic location and predicted effect being nonsynonymous SNVs, frameshift deletions/insertions, or stopgain. For variants identified in both WES and RNA-seq data, the presence on forward and reverse strands was further verified using IGV.

Somatic variant validation
Primers were designed using Primer3 software (Supplemental Table 14 and refs. 69, 70). PCR products were obtained using AmpliTaq Gold DNA Polymerase (Invitrogen), purified using QIAquick PCR Puriﬁcation Kit (Qiagen), and sequenced using BigDye Terminator v3.1 chemistry (Invitrogen) on the 3130xl Genetic Analyzer (Invitrogen). Chromas Lite 2.01 software (Technelysium Pty Ltd.) and BLAST software were used for the analysis of sequencing data (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Data availability
Raw gene expression microarray data is available in Gene Expression Omnibus (GEO), series GSE85383. Raw RNA-seq data is available in GEO, series GSE87581. Raw aCGH+SNP data is available in GEO, series GSE85382. Raw WES data is available in the Sequence Read Archive (SRA), study accession number SRP090691.

Statistics
Microarray gene expression data analysis. Expression values from different probes mapped to the same gene were collapsed using averegps function from limma package. Exploratory and differential expression analyses of microarray data were performed in R Studio (build: R 3.0.2). For the exploratory analysis, raw signal values were normalized to median equal 0 and log<sub>2</sub> transformed. PCA was performed using pcomp function (“stats” R package, version 3.0.2) and visualized using ClustVis (71). Unsupervised hierarchical clustering was performed using Euclidean distance and complete linkage rule on genes with SD > 1 (function hclust in “stats” R package, version 3.0.2). Differential expression analysis was performed using SAM (21) on quantile-normalized expression values filtered for SD > 100 and log<sub>2</sub> transformed (with seed value 1234567 and adjusted P value cut-off 0.05). SAM output was filtered for genes with fold change < –2 and > 2. CIBERSORT analysis was performed on quantile-normalized raw signal values with number of permutations set to 500 using CIBERSORT web interface (https://cibersort.stanford.edu) (22).
Functional enrichment/gene set analysis of differentially expressed genes was performed using Gene Set Enrichment Analysis (GSEA) (MSigDB, Molecular Signatures Database v5.1) (72, 73), ToppGene Suite (74), and TSGene database (75). Hypergeometric distribution was calculated using \( n = 45,956 \) genes in gene universe, following the GSEA methodology. Representation factor was calculated as \( \frac{x}{[n \times D]/N} \) where \( x \) is a number of overlapping genes, \( n \) is a number of genes in group 1, \( D \) is a number of genes in group 2, and \( N \) is a number all genes in gene universe.

**RNA-seq gene expression analysis.** Differential expression analysis and exploratory analysis of RNA-seq data was performed in R Studio (built: R 3.0.2), using read counts quantified by featureCounts (28–67.4 million reads per sample, median 43.2 million reads per sample).

Exploratory analysis was performed on genes with mean read counts > 10. Filtered read counts were median centered and loge transformed after adding 1 to each read count, and the unsupervised hierarchical clustering was performed using Euclidean distance and ward.D linkage rule.

Differential gene expression analysis of RNA-seq data was performed using parametric methods DESeq2 (package version 1.2.10) (18) and edgeR (package version 3.4.2) (19), which were applied to genes with mean raw read counts > 10. SAMseq (samr package version 3.4.2) (20), a nonparametric method, was performed with the number of permutations set to 100 and seed value 1234567. The outputs of DESeq2, edgeR, and SAMseq were filtered for adjusted \( P \) value < 0.05 and fold change < –2 and > 2.

Supervised hierarchical clustering based on loge(TPM+1) expression values for 406 genes differentially expressed between UUS and ESS was performed on genes and samples using Euclidean distance and complete linkage rule.

**qRT-PCR data analysis.** qPCR data were analyzed using Ct values with the SDS 2.1 software (Applied Biosystems). Normalization was performed based on the mean values of 2 housekeeping genes, APP (amyloid \( \beta \) [A4] precursor, NM_000484) and HPRT1 (hypoxanthine phosphoribosyltransferase 1, NM_000194), and the relative amounts of RNA for each gene were calculated by the 2 \( ^{ΔΔCt} \) method using DataAssist software v. 3.0 (Applied Biosystems) and compared between UUS and ESS using 2-tailed Student’s \( t \) test. \( P < 0.05 \) was considered statistically significant.

In the box and whisker plots, the middle line represents median, the bounds of the boxes represent first and third quartiles, the whiskers above and below the box represent data range, and the circles above or below whiskers identify outliers. All statistical tests were 2-tailed. \( P < 0.05 \) was considered significant.

**Study approval**
Specimens were collected prospectively at the Maria Sklodowska-Curie Memorial Cancer Center and Institute - Oncology Center after written informed consent had been obtained. Archival tumor specimens from KU Leuven and University Hospitals and Stanford University Medical Center originate from patient care and were requalified for research. The study was approved by Maria Sklodowska-Curie Institute - Oncology Center, KU Leuven, and Stanford University IRBs.

**Author contributions**
JP, MK, JC, MVDR, and MDR designed the study. Acquisition of data was performed by JP, MK, AQ, VV, S. Varma, S. Vennam, EBZ, JAS, MB, MVDR, and MDR. Analysis and interpretation of data were performed by JP, MK, BD, AMN, MS, MVDR, and MDR. All authors participated in writing the manuscript.

**Acknowledgments**
This study was supported by KU Leuven Concentrated Action grant number 2011/010 and the grant of the Polish Ministry of Science and Higher Education no. NN407 134 039. JP was supported by the International PhD Projects Program of the Foundation for Polish Science and by the European Union Regional Development Fund. We thank Robert West and Joseph Foley for valuable suggestions and critical reading of the manuscript. We thank Kamil Zalewski for assistance with collection of clinical data.

Address correspondence to: Joanna Przybyl, Department of Pathology, Stanford University School of Medicine, 300 Pasteur Drive, Lane Bldg., L209, Stanford, California 94305, USA. Phone: 650.725.7742; E-mail: jprzybyl@stanford.edu.
37. Agrell L, et al. PI3K signaling pathway is activated by PIK3CA mRNA overexpression and copy gain in prostate tumors, but PIK3CA, BRAF, KRAS and AKT1 mutations are infrequent events. Mod Pathol. 2011;24(3):443–452.


