Clinically Relevant Molecular Subtypes in Leiomyosarcoma

Xiangqian Guo1, Vickie Y. Jo2, Anne M. Mills3, Shirley X. Zhu1, Cheng-Han Lee4, Inigo Espinosa5, Marisa R. Nucci2, Sushama Varma1, Erna Forgó1, Trevor Hastie6, Sharon Anderson1, Kristen Ganjoo7, Andrew H. Beck8, Robert B. West1, Christopher D. Fletcher2, and Matt van de Rijn1

Abstract

Purpose: Leiomyosarcoma is a malignant neoplasm with smooth muscle differentiation. Little is known about its molecular heterogeneity and no targeted therapy currently exists for leiomyosarcoma. Recognition of different molecular subtypes is necessary to evaluate novel therapeutic options. In a previous study on 51 leiomyosarcomas, we identified three molecular subtypes in leiomyosarcoma. The current study was performed to determine whether the existence of these subtypes could be confirmed in independent cohorts.

Experimental Design: Ninety-nine cases of leiomyosarcoma were expression profiled with 3′end RNA-Sequencing (3SEQ). Consensus clustering was conducted to determine the optimal number of subtypes.

Results: We identified 3 leiomyosarcoma molecular subtypes and confirmed this finding by analyzing publically available data on 82 leiomyosarcoma from The Cancer Genome Atlas (TCGA). We identified two new formalin-fixed, paraffin-embedded tissue-compatible diagnostic immuno-

histochemical markers: LMOD1 for subtype I leiomyosarcoma and ARL4C for subtype II leiomyosarcoma. A leiomyosarcoma tissue microarray with known clinical outcome was used to show that subtype I leiomyosarcoma is associated with good outcome in extraterine leiomyosarcoma while subtype II leiomyosarcoma is associated with poor prognosis in both uterine and extraterine leiomyosarcoma. The leiomyosarcoma subtypes showed significant differences in expression levels for genes for which novel targeted therapies are being developed, suggesting that leiomyosarcoma subtypes may respond differentially to these targeted therapies.

Conclusions: We confirm the existence of 3 molecular subtypes in leiomyosarcoma using two independent datasets and show that the different molecular subtypes are associated with distinct clinical outcomes. The findings offer an opportunity for treating leiomyosarcoma in a subtype-specific targeted approach.

Clin Cancer Res; 21(15): 3501–11. ©2015 AACR.

Introduction

Leiomyosarcoma is reported to represent as many as 24% of all sarcomas (1, 2). Currently, no targeted therapy exists for leiomyosarcoma and the tumor responds poorly to conventional chemotherapy or radiotherapy. A significant proportion of leiomyosarcoma originates in the uterus, and the remainder of leiomyosarcoma originates in various soft tissue sites where it is thought to often arise from smooth muscle cells in vessel walls.

The successful stratification of several tumors (including breast cancer, lung cancer, and colon cancer) into molecular subtypes in the past decades has significantly improved our knowledge of these malignancies and has led to changes in the therapeutic approach to these cancers (3–12). In a previous study, our group proposed the existence of three molecular subtypes of leiomyosarcoma by using microarray-based gene expression profiling on 51 fresh frozen samples of leiomyosarcoma (13). To validate this finding, we analyzed 99 leiomyosarcoma cases collected at different institutions from the years 1991 to 2012 as an independent cohort, and performed expression profiling with 3′ end RNA sequencing (3SEQ; refs. 14–20). In addition, we analyzed the publically available expression profiling dataset from TCGA on 82 cases. Our analysis confirms the existence of three molecular subtypes in leiomyosarcoma. Using two novel markers, we could distinguish the three subtypes by immunohistochemistry and correlate the subtypes with clinical outcome. The recognition of these molecular subtypes in leiomyosarcoma may have clinical significance as the subtypes vary in their expression of a number of genes for which novel targeted therapies either already exist or are being developed.
Translational Relevance

Leiomyosarcoma is a malignant neoplasm with varying degrees of smooth muscle differentiation and complex genetic abnormalities. It can occur in a wide range of sites but is usually managed as a single disease in conventional therapy and in clinical trials. The recognition of molecular subtypes can lead to the identification of novel therapies that may affect one of these subtypes preferentially. We confirmed the results of an earlier study and determined the existence of 3 molecular subtypes in leiomyosarcoma using two independent cohorts. We identified two new formalin-fixed, paraffin-embedded (FFPE) tissue-compatible diagnostic immunohistochemical markers for two of these subtypes: LMO1D1 for subtype I and ARL4C for subtype II, and showed that these molecular subtypes are associated with distinct clinical outcomes. Our study offers an opportunity for a leiomyosarcoma subtype–specific targeted treatment.

Materials and Methods

3SEQ library construction and bioinformatics analysis

Paraffin blocks of 99 leiomyosarcomas, 4 myometrium, 3 leiomyoma, and 6 undifferentiated pleomorphic sarcomas from 1991 to 2012 from 9 hospitals were obtained with Institutional review board approval and a waiver of consent due to the archival nature of the specimens. Multiple 2 mm-diameter cores were re-embedded in paraffin blocks longitudinally and sectioned again to ensure the purity of tumor cells. After RNA isolation, 3SEQ libraries for next-generation sequencing–based expression profiling were sequenced and analyzed as described previously (14–20). All gene expression profiling data used for this study have been deposited in the Gene Expression Omnibus (GEO) and are publicly accessible through GSE45510, GSE53844 and GSE54511.

After filtering data to genes with a SD greater than 100, transforming the data by log2 and gene-based centering, the Consensus Clustering (R package ConsensusClusteringPlus; ref. 21) was used to determine the number of subtypes and to assign the subtype for each leiomyosarcoma case. This was run over 1,000 iterations with the “Distance” (1–Pearson correlation), a 80% sample resampling, 80% gene resampling, a maximum evaluated k of 12, and agglomerative hierarchical clustering algorithm. Silhouette width (R package cluster; ref. 22) was then calculated on the basis of the assignment from ConsensusClusteringPlus to measure the accuracy of assignments from ConsensusClusteringPlus. Separately, RNASeq data of 82 leiomyosarcoma cases were downloaded from the TCGA database. To compare with 3SEQ data, the TCGA data were normalized into TPM and analyzed with ConsensusClusteringPlus and Silhouette analysis as performed for 3SEQ data. Subclass mapping (23) was performed to determine the common leiomyosarcoma subtypes identified in 3SEQ and TCGA RNASeq datasets.

To get subtype specific genes, SAMSeq (24) was performed on both datasets between each subtype and all other subtypes with a false discovery rate (FDR) of 0.05. Kaplan–Meier plots were used to compute the survival curves, and Log-rank (Mantel–Cox) test was used to determine the statistical significance of survival between groups by using GraphPad Prism5 software. Significance of contingency analysis between one subtype and the other subtypes was measured by two-tailed Fisher exact test and χ2 test with GraphPad Prism5 software. Univariate and multivariate analysis by the Cox proportional hazard method was performed using the survival package in R. Analysis of gene ontology (GO) was done using DAVID Bioinformatics Resources version 6.7 (25, 26). CSF1-signature (27) positive, and CINSARC signature (28) positive cases were identified as those cases that coordinate highly expressed these signature genes with ≥0.3 correlation with the centroid as performed previously (19, 27, 29, 30). Principal component analysis (PCA) was performed on square root transformed TPM with R package pcaMethods, ellipse contour of principal components (PC) was computed and drawn with R package car.

Results

Consensus clustering identifies three molecular subtypes of leiomyosarcoma based on gene expression data

We analyzed 99 cases of leiomyosarcoma by 3SEQ, a next-generation sequencing–based approach that quantifies the number of 3’ ends for each polyadenylated mRNA or IncRNA and that performs well on RNA isolated from FFPE material (14–20). Clinical information of the cases is shown in Table 1 and Supplementary Table S1. The median age for this cohort is 56 years, with 49 cases originating in the uterus and 49 cases in extrauterine sites. For one case the origin was unknown. A female: male ratio of 3.3:1 was noted in this cohort. Half of the extrauterine leiomyosarcomas (24 cases) were from the extremities, while the rest were from thoracic, abdominal, or retroperitoneal regions.

3SEQ yielded an average of 18 million total reads after quality filtering and an average of 2 million uniquely mapping reads against human transcriptome (refRNA) per sample. All reads that uniquely mapped to an individual gene (UCSC genome browser, hg19) were combined to obtain a quantification of the expression level for that gene. The genes with the highest degree of variation in expression levels across all samples were identified by filtering the dataset for genes with a SD ≥100, yielding 1,300 genes. These were used to perform Consensus Clustering (21), a method that uses data resampling followed by iterative reclustering to estimate cluster stability.
Molecular Subtypes in Leiomyosarcoma

Table 1. Patient characteristics (N = 99)

<table>
<thead>
<tr>
<th></th>
<th>Patients, n (%)</th>
<th>Subtype I</th>
<th>Subtype II</th>
<th>Subtype III</th>
<th>Other LMS*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)</td>
<td></td>
<td>56</td>
<td>55</td>
<td>55</td>
<td>60</td>
</tr>
<tr>
<td>Median</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>76 (77%)</td>
<td>22</td>
<td>18</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>Male</td>
<td>23 (23%)</td>
<td>13</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uterine</td>
<td>49 (49%)</td>
<td>9</td>
<td>13</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>TAR*</td>
<td>25 (25%)</td>
<td>12</td>
<td>8</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Extremities</td>
<td>24 (24%)</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (1%)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>18 (18%)</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Intermediate</td>
<td>19 (19%)</td>
<td>10</td>
<td>4</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>High</td>
<td>62 (63%)</td>
<td>15</td>
<td>15</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>Relapse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>62 (63%)</td>
<td>22</td>
<td>9</td>
<td>12</td>
<td>19</td>
</tr>
<tr>
<td>Local recur</td>
<td>27 (27%)</td>
<td>9</td>
<td>12</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Metastasis</td>
<td>10 (10%)</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Therapy#</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>11 (11%)</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>No</td>
<td>86 (87%)</td>
<td>29</td>
<td>18</td>
<td>12</td>
<td>27</td>
</tr>
<tr>
<td>Unknown§</td>
<td>2 (2%)</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>99</td>
<td>55</td>
<td>38</td>
<td>38</td>
<td>25</td>
</tr>
</tbody>
</table>

*Other leiomyosarcoma, leiomyosarcoma with negative silhouette value.

**TAR, thoracic/abdominal/retroperitoneal sites.

#Any prior chemotherapy and radiotherapy.

and to help determine the optimal number of clusters in a dataset. As shown in Fig. 1A and B, the greatest increase in the area under the CDF curve (empirical cumulative distribution function) is seen when 3 molecular subtypes are assumed. Assuming 4 molecular subtypes showed a lesser degree of increase in the area under the CDF curve and showed no significant increase in the ability to classify the cases (Supplementary Fig. S1). To keep the number of subtypes to a manageable number, we focused on the existence of three molecular subtypes of leiomyosarcoma for this study. A consensus matrix of the probability that 2 cases of leiomyosarcoma will belong to the same subtype is shown in Fig. 1C. Silhouette analysis (22) was performed to measure the consensus matrix of the probability that 2 cases of leiomyosarcoma will belong to the same subtype.

Molecular subtypes of leiomyosarcoma can be reproduced in an independent dataset

To further explore the reproducibility of the three leiomyosarcoma subtypes, we downloaded gene expression data for 82 leiomyosarcoma cases from the TCGA database. Of these samples, 80 cases were primary leiomyosarcoma while 2 cases were recurrences. The origin of the tumors differed from the 99 cases analyzed by 3SEQ with 19 leiomyosarcoma originating from the uterus, 13 cases from the extremities and 32 cases from other extrauterine sites. The remaining 18 cases had an unknown primary site. In addition to the differences in the site of origin for the tumors examined there was a significant difference in the percentage of samples derived from primary lesions versus recurrences between the two cohorts. In the 3SEQ cohort 37% of the cases were recurrences or metastases while only 2% of TCGA leiomyosarcoma cases were recurrent and none were metastatic (Supplementary Table S2). The two datasets differed in two additional aspects. First, the TCGA dataset used fresh frozen material whereas our 3SEQ analysis was performed on FFPE material. Second, whole transcriptome RNASeq was used for the TCGA dataset while 3SEQ determines the gene expression level based on quantification of 3′ end reads. Despite these differences, analysis of the TCGA dataset produced results very similar to those found in the 3SEQ cohort. The TCGA dataset yielded 1,174 genes when filtered for a SD ≥100 (1,300 genes in the 3SEQ set). By Consensus Clustering and Silhouette analysis, the TCGA dataset also showed the greatest increase in area under the CDF curve when three subtypes were used (Fig. 2A). Compared with the cases studied by 3SEQ in our cohort the homogeneity of the subtypes in the TCGA cohort was greater as determined by consensus clustering (Fig. 2B), possibly due to the fact that frozen material and not FFPE material was used for the cases in the TCGA cohort. To measure the reproducibility between the 3SEQ and TCGA cohorts the cases with positive Silhouette values from both sets were analyzed by Subclass mapping (23). The 3SEQ-subtype I and -subtype II types were significantly reproduced as subtype-C3 and -C1 in the TCGA cohort with FDR corrected P values of 0.0090 for both. The 3SEQ-subtype III did not reach statistical significance when it was compared with the remaining TCGA-subtype C2 (Fig. 2C; P = 0.0959). A substantial number of genes identified by SAMSeq in the 3SEQ and TCGA datasets showed overlap between 3SEQ subtypes I and II and the corresponding TCGA...
subtypes (Supplementary Table S3). Importantly, five genes for which we previously used immunohistochemical markers (13, 33) to identify subtype I leiomyosarcoma (CASQ2, ACTG2, MYLK, SLMAP, and CFL2) were shown to be highly expressed on the mRNA level in TCGA subtype C3, as was the new subtype I marker LMOD1. In addition, the new marker ARL4C, which identifies 3SEQ subtype II cases, was highly expressed in the corresponding TCGA subtype C1 (see below and Supplementary Table S4). Together, these data indicate that the existence of 3 molecular subtypes in leiomyosarcoma can be reproduced on the mRNA level across these different datasets.

Clinical features of leiomyosarcoma molecular subtypes in three datasets

In the 3SEQ dataset, approximately 26%, 59%, and 92% of subtype I, subtype II, and subtype III leiomyosarcoma were derived from the uterus with uterine leiomyosarcoma being...
significantly associated with subtype III (χ² test; \( P = 0.0005 \)). However, when evaluating only those 34 cases arising in the uterus, we found that uterine leiomyosarcoma had a roughly equal chance of belonging to each of the molecular subtypes with 9, 13, and 12 cases belonging to subtype I, II, and III, respectively. In contrast, the extrauterine leiomyosarcoma were overrepresented in subtype I and II, with 25, 9, and 1 cases belonging to subtype I, II, and III, respectively. These results indicate that a subset of uterine leiomyosarcomas behave as independent molecular subtype (subtype III) while another subset of uterine leiomyosarcoma together with extrauterine leiomyosarcoma belong to subtype I and subtype II leiomyosarcoma. In the TCGA dataset, we observed similar demographic characteristics as in the 3SEQ cohort, with TCGA C2–like 3SEQ subtype III being significantly

Figure 2.
Identification of molecular subtypes of leiomyosarcoma (LMS) in the TCGA data. A, empirical cumulative distribution plots as a function of the number of molecular subtypes in TCGA cohort. B, consensus clustering matrix of leiomyosarcoma samples using 3 molecular subtypes. C, subclass association (SA) matrix between 3SEQ data (subtype I, subtype II, subtype III) and TCGA data (subtype C1, subtype C2, subtype C3). \( P \) value is corrected with FDR. D, principal component analysis (PCA) of 70 core 3SEQ leiomyosarcoma cases with other diagnoses. Subtype I leiomyosarcoma, black circle; subtype II leiomyosarcoma, red triangle; subtype III leiomyosarcoma, green triangle; leiomyoma, cyan solid circle; myometrium, orange solid circle; undifferentiated pleomorphic sarcomas, blue solid pyramid. The elliptical contours represent the normal probability of 0.75.
associated with uterine leiomyosarcoma \( (P < 0.0001) \), while TCGA C3 (subtype I) and TCGA C1 (subtype II) were overrepresented in extraterine sites (Supplementary Table S2). These results correlate well with our findings on 51 leiomyosarcoma cases from our previous dataset analyzed by gene microarrays, where the ratio of uterine leiomyosarcoma (13) also was higher in subtype III (10/24) than in subtype I (1/13) and subtype II (3/12), but did not reach significance \((\chi^2 \text{ test}; P = 0.1178)\).

Within the 3SEQ cohort, 18% (18) of tumors are low grade, 19% (19) are intermediate grade, and 63% (62) are high grade by histologic analysis. Low-grade lesions were more frequent in subtype I leiomyosarcoma, with 10 of 35 cases showing low-grade histology while 3 of 22 of subtype II and 0 of 13 of subtype III leiomyosarcoma cases are low grade, respectively; however, there was no statistically significant correlation between grade and molecular subtype (Table 1, \(\chi^2 \text{ test}; P = 0.0989\)). The median age for subtype I, subtype II, and subtype III was 55, 55, and 60 years, respectively. One-way ANOVA analysis showed that there was no significant difference in age between the three subtypes \((P = 0.3769)\).

Functional annotation of leiomyosarcoma molecular subtypes based on subtype-specific genes in the 3SEQ dataset

SAMSeq (24) was used to identify genes that were significantly over/underexpressed in each subtype of leiomyosarcoma analyzed by 3SEQ, again using only those cases with a positive Silhouette value. In subtype I leiomyosarcoma, 5,900 genes are overexpressed while 1,900 genes are underexpressed compared with the other two subtypes (Supplementary Table S5). The overexpressed genes were enriched in Gene Ontology (GO) Biology processes that included muscle contraction, muscle system processes, and cytoskeleton organization. In addition, a number of genes (LMOD1, MYOCD, CALD1, DES, CNN1, ACTG2, SLMAP, and CASQ2) known to be involved in smooth muscle function were identified in this list, indicating that subtype I was most associated with normal smooth muscle function (Supplementary Table S6). Consistent with this finding, 4 samples of normal myometrium and 3 cases of benign uterine leiomyoma were grouped with subtype I leiomyosarcoma in principal component analysis (Fig. 2D) and clustered with subtype I when analyzed by unsupervised hierarchical clustering with the 70 leiomyosarcoma core samples (Supplementary Fig. S2A).

As stated above, the overexpression of genes associated with muscle function was also noted on the mRNA level in corresponding subtype C3 from the TCGA dataset, while the subtype I identified previously through gene microarray analysis had high levels of expression for these genes at both the mRNA and protein level (13).

In 3SEQ data on subtype II leiomyosarcoma, 5,100 genes were significantly overexpressed and 2,087 genes were expressed at lower levels as compared with the other two subtypes of leiomyosarcoma. These subtype II leiomyosarcoma-specific genes are enriched in GO biologic processes, including translation, translational elongation, and protein localization (Supplementary Table S6). Subtype II showed significantly less muscle-specific genes than subtype I. Undifferentiated pleomorphic sarcoma is a tumor type that fails to express differentiation markers and that can be difficult to distinguish from leiomyosarcoma. Six cases of undifferentiated pleomorphic sarcomas grouped with subtype II leiomyosarcoma in principal component analysis (Fig. 2D) and by unsupervised hierarchical clustering (Supplementary Fig. S2B). A search for the CSF1 signature, which we previously reported to be associated with poor outcome for both gynecologic and non-gynecologic leiomyosarcoma (27), in the 3SEQ data showed that the cases with a centroid correlation > 0.3 for the CSF1 signature were overwhelmingly subtype II leiomyosarcoma \((\chi^2 \text{ test}; P < 0.0001)\).

A total of 1,320 genes are significantly overexpressed and 7,544 genes are underexpressed in subtype III leiomyosarcoma when compared with the other two subtypes. Subtype III overexpressed genes are enriched in GO biologic processes, including metabolic process, ion transport, and regulation of transcription (Supplementary Table S6). When considering only the genes with a mean level of expression 50 reads (TPM) or more in any subtype in the Stanford 3SEQ cohort the number of overexpressed genes identified by SAMseq will decrease from 11,125 to 3,350. The same approach will decrease the number of genes in the TCGA cohort from 11,068 to 3,845.

Identification of novel immunohistochemical markers for leiomyosarcoma subtype I and II

Through gene microarray analysis, we previously identified 5 IHC markers (ACTG2, SLMAP, MYLK, CFL2, and CASQ2) that could identify subtype I leiomyosarcoma in FFPE material (13, 33). The quantitative nature of gene expression profiling by 3SEQ allowed for a more detailed search for additional immunohistochemical markers for leiomyosarcoma subtypes. LMOD1 mRNA was highly expressed in subtype I leiomyosarcoma. A tissue microarray (TA-381) was generated that contained 58 of the 70 leiomyosarcoma cases with positive Silhouette values. Immunohistochemistry showed strong staining of a subset leiomyosarcoma (Fig. 3A). LMOD1 stained positive in 31 leiomyosarcoma cases, 19 of which were subtype I leiomyosarcoma as assigned by 3SEQ analysis, in contrast, 21 of 27 LMOD1-negative leiomyosarcomas were subtype II or subtype III leiomyosarcoma. LMOD1 protein expression therefore showed a significant association with subtype I leiomyosarcoma \((\chi^2 \text{ test}; P = 0.0085; r = 0.8964, P < 0.0001, \text{Spearman correlation}; \text{Supplementary Table S7})\). Comparison of the LMOD1 staining results with the five previously identified subtype I biomarkers (13, 33) showed that the highest correlation of IHC staining (0.65) was obtained between genes ACTG2, SLMAP, and LMOD1. CASQ2 had the lowest correlation with the 5 other genes \((r = –0.22)\) and we refined our panel of subtype I biomarkers to include ACTG2, SLMAP, LMOD1, CFL2, and MYLK, while omitting CASQ2. Using this panel, the correlation for the new panel of markers was 0.47 (Supplementary Fig. S3). Finally, staining of TA-381 showed a significant association \((P < 0.0001)\) between LMOD1 immunostaining and mRNA levels as determined by 3SEQ LMOD1 mRNA level (Fig. 3B).

Analysis of 3SEQ data showed high levels of mRNA expression for AR1AC in subtype II leiomyosarcoma. Using a FFPE compatible antibody, strong staining was seen in a subset of leiomyosarcoma cases (Fig. 3C) with positive staining in 30 leiomyosarcoma cases, 17 of which were subtype II leiomyosarcoma. Negative staining was seen in 28 leiomyosarcoma cases, 23 of which were subtype I or subtype III leiomyosarcoma. This IHC result validates ARL4C to be a subtype II leiomyosarcoma biomarker not only at the mRNA level (as determined by SAMseq) but also at the protein level \((\chi^2 \text{ test}, P = 0.0033; r = 0.5277, P < 0.0001, \text{Spearman correlation};\)
Supplementary Table S7). Comparison between 3SEQ mRNA levels and ARL4C immunohistochemistry showed a good association ($P = 0.0046$; Fig. 3D). Comparison between LMOD1 and ARL4C protein expression and mRNA expression levels varied as expected across the leiomyosarcoma subtypes as defined by 3SEQ (Supplementary Fig. S4).

While in individual cases these markers may not definitely identify a case as belonging to a specific subtype, when combined these markers do allow us to distinguish large numbers of cases represented on a TMA in distinct molecular groups. No outcome data were available for the leiomyosarcoma samples used for 3SEQ analysis. To correlate the assignment of leiomyosarcoma subtypes with outcome, we used immunostaining data on TA-381 that contains 58 cases of leiomyosarcoma analyzed by 3SEQ; leiomyosarcoma cases positive for anti-LMOD1 staining have significantly higher mRNA expression levels (TPM) than leiomyosarcoma cases negative for anti-LMOD1 ($P < 0.0001$). The $P$ value was derived from unpaired $t$ test with Welch correction.

Subtype II leiomyosarcoma cases were defined as those reacting for ARL4C on the outcome TMA. Staining for ARL4C was seen in 25 of 111 available leiomyosarcoma cases (23%) and these cases were associated with a worse DSS when all leiomyosarcoma were combined (log-rank test, Fig. 4D; $P = 0.0046$). This finding was seen in both extrauterine ($P = 0.0352$) and uterine ($P = 0.0461$) leiomyosarcoma. While subtype II was associated with poor outcome in a statistically significant manner in univariate analysis ($P = 0.0030$, Wald test), this significance was also lost in the multivariate analysis when prognosticators ROR2 and the CSF1 response protein
signature were included (refs. 27, 34; Supplementary Tables S8 and S9). For correlation with outcome in the IHC analysis, subtype III was defined as those cases that failed to classify as either subtype I or II. When thus defined, subtype III had an intermediate outcome (Supplementary Fig. S5) but failed to reach prognostic significance in univariate analysis (Supplementary Table S8).

Potential clinical applications of leiomyosarcoma subtyping

To explore the potential for clinical implications in the recognition of leiomyosarcoma subtypes, we compared the genes that are specifically overexpressed in each leiomyosarcoma subtype with genes from the TARGET V2 database known to be activated by mutations or amplifications (35). This database contains genes for which targeted therapy is either currently available or under development. A large number of these genes were found to be expressed at different levels between the 3 molecular subtypes (Table 2). This suggests that leiomyosarcoma subtypes may respond differentially to those targeted therapies, an observation that may play a role in determining the success rate of novel therapies in clinical trials. For several of these targets there may be an incomplete correlation between mRNA levels for the mutated or amplified target gene and response to drug, but this analysis nevertheless forms a first step towards personalization of leiomyosarcoma patient care.

In addition to the genes identified from the TARGET V2, we studied the expression level of ROR2, a receptor tyrosine kinase that plays a role in tumor progression and that has a low
level of expression in normal human tissues. This molecule has been proposed as a novel therapeutic target in leiomyosarcoma (34) and other tumors (36–38) and was found to be more frequently expressed in subtype II leiomyosarcoma than in the other subtypes (FDR < 0.05). Immunostaining for the ROR2 protein on the TMA containing cases analyzed by 3SEQ similarly showed a correlation with subtype II leiomyosarcoma (χ² test, P = 0.0347).

Discussion

Leiomyosarcoma is a malignant soft tissue tumor with complex genetic abnormalities. In clinical practice, the tumors are treated differently depending on whether they originate from the uterus or from extraterine sites. However, this distinction is not based on a molecular rationale and appears based, at least in part, on clinical subspecialization. Recently Italiano and colleagues demonstrated a molecular difference between leiomyosarcoma occurring in the retroperitoneum and extremities (39), but this article does not include uterine lesions. Leiomyosarcomas are tumors that exhibit varying degrees of smooth muscle differentiation and often appear to be derived from smooth muscle cells in the myometrium. They can also arise from the walls of blood vessels or connective tissues throughout the body. Conventional chemotherapy and radiotherapy only have limited effects and surgical resection remains the best option for treatment. Characterization of molecular subtypes in malignancies can lead to better individualization of treatment. In a prior study on 51 cases of leiomyosarcoma for which frozen tissue was available, we used 44,000 spotted cDNA arrays for gene expression profiling and obtained data that indicated that at least 3 molecular subtypes exist in leiomyosarcoma (13). This study relied on the availability of fresh frozen tissues and as a result only small numbers of leiomyosarcoma cases could be studied. The 3SEQ approach can use FFPE materials and renders a digital rather than an analog readout for gene expression levels with a higher signal to noise ratio and allows for accurate measurement of genes expressed at a wide range of levels, including those with low levels of expression. 3SEQ analysis on 99 samples of leiomyosarcoma and data on a third cohort of 82 leiomyosarcoma cases from the TCGA study were obtained. Despite the differences in methodology and sample selection between these three cohorts, remarkable similarities were apparent and indicated the existence of three molecular subtypes in leiomyosarcoma.

Subtype I leiomyosarcoma expresses most genes associated with muscle function by Gene Ontology annotation. The similarity of subtype I leiomyosarcoma to smooth muscle differentiation was also supported by the fact that subtype I cases clustered with samples of leiomyoma and myometrium. Finally, we observed high expression levels for LMOD1 in subtype I leiomyosarcoma, a smooth muscle cell-restricted gene that is preferentially expressed in differentiated smooth muscle cells (40). The expression of smooth muscle markers in subtype I suggest a tumor subtype that is closer to the smooth muscle cells than subtype II and III. However this similarity is not captured by the objective measurement of histologic grade, as no correlation between grade and molecular subtype was found. Others have noticed the existence of a subset of leiomyosarcoma that express muscle function related genes at higher levels than other leiomyosarcoma cases. Francis and colleagues (41) identified a 200 gene signature for leiomyosarcoma by expression profiling different subtypes of 177 soft tissue sarcomas; this signature was characterized by overexpression of muscle-specific genes and was shown to be highly expressed in 12 of 40 leiomyosarcoma cases by supervised clustering. Villacic and colleagues (42) performed gene expression profiling and supervised clustering with 587 genes expressed differentially between undifferentiated pleomorphic sarcomas and leiomyosarcoma, and found that a subset of leiomyosarcoma that highly expressed muscle function associated genes clustered separately from undifferentiated pleomorphic sarcomas and other leiomyosarcoma cases.

In contrast to subtype I leiomyosarcoma, subtype II leiomyosarcoma showed no significant smooth muscle differentiation by GO analysis and therefore represents a less differentiated form of leiomyosarcoma. Undifferentiated pleomorphic sarcoma is a type of sarcoma that lacks any indication of differentiation by histologic examination and by immunohistochemical analysis. Hierarchical clustering of the core leiomyosarcoma cases with 6 cases of undifferentiated pleomorphic sarcomas showed coclustering of undifferentiated pleomorphic sarcomas with subtype II leiomyosarcoma. Others have studied the relation between leiomyosarcoma and undifferentiated pleomorphic sarcomas by molecular studies but did not take into account the existence of leiomyosarcoma subtypes (43). Villacic and colleagues performed gene expression profiling of 22 leiomyosarcoma and 22 undifferentiated pleomorphic sarcomas, and did not identify a clear distinction of expression profiles of undifferentiated pleomorphic sarcomas and leiomyosarcoma by either unsupervised clustering or even supervised clustering using a gene list derived from SAM analysis. However, when performing supervised clustering with 587 genes expressed differentially between undifferentiated pleomorphic sarcomas and leiomyosarcoma, a subset of leiomyosarcoma that consisted predominantly of retroperitoneal leiomyosarcoma clustered separately from undifferentiated pleomorphic sarcomas and the remaining leiomyosarcoma, in part based on higher expression of muscle function associated genes in the retroperitoneal leiomyosarcoma (42). These findings partially overlap with our results; in our study 6 retroperitoneal leiomyosarcoma were analyzed, 4 of which grouped in subtype I with 2 cases grouping in subtype II. However, this subtype most associated with undifferentiated pleomorphic sarcomas.

Subtype III leiomyosarcoma is the only subtype that shows a preference for a specific anatomic site and was more likely to be from uterus (Fisher exact test: P = 0.005). Only one of 13 subtype III leiomyosarcoma did not originate from the uterus and presented in the scrotum. However, uterine leiomyosarcoma as a group was evenly distributed over the 3 subtypes.

Neither uterine nor extraterine leiomyosarcoma have clinically accepted prognostic molecular markers, and histologic grading to predict clinical outcome is controversial in uterine leiomyosarcoma. Several molecular prognosticators (28, 39, 41, 44, 45) have previously been reported in leiomyosarcoma. By genomic and expression profiling of 183 sarcomas, Chibon and colleagues established a prognostic signature called Complexity Index in Sarcoma (CINSARC), that includes 67 genes related to mitosis and chromosome management. This signature was shown to be associated with metastatic outcome in sarcomas, including leiomyosarcoma, and even in lymphomas and breast carcinoma. The CINSARC signature is a
powerful predictor of tumor metastasis and could improve the patient selection for chemotherapy (28). To investigate the significance of leiomyosarcoma subtypes on clinical outcome, we stained a TMA containing cores from 127 leiomyosarcoma cases with clinical outcome with the IHC markers for subtype I and II leiomyosarcoma. Subtype I leiomyosarcoma was associated with good outcome in extratherine leiomyosarcoma, while subtype II was associated with poor outcome in univariate analysis. In this context, it is interesting to note that the association between muscle gene expression and better clinical outcome is a finding that was also reported in various ways by other immunohistochemical studies (18, 45). While subtyping of leiomyosarcoma by immunohistochemistry can help predict clinical outcome, in multivariate analysis the subtype I and II biomarkers were outperformed by the previously described markers ROR2 and the “CSF1 protein response signature” (27, 34). The association between immunohistochemically defined subtype II leiomyosarcoma with poor outcome was consistent with previously published gene expression profiling data. The CINSARC signature was significantly associated with subtype II leiomyosarcoma but not with the other subtypes (P = 0.0279 by Fisher exact test).

At this time, no targeted therapies exist for leiomyosarcoma. The recognition of molecular subtypes in malignancy has led to the identification of targets that are unique to a subset of the cases in a variety of tumors that include breast cancer, colon cancer, bladder cancer and glioblastoma, and others (3–12). In previous studies, we have identified ROR2, a receptor tyrosine kinase (34) and the presence of intratumoral macrophages (27) as prognostic markers in both uterine and extratherine leiomyosarcoma. On the basis of our finding, subtype II leiomyosarcoma, which has high levels of ROR2 expression, would benefit most from drugs targeting this molecule either as a single drug or in combination with CD47-targeted therapy (34, 46, 47).

Here we report that the molecular subtypes in leiomyosarcoma show differential expression of a large number of genes for which therapeutics either currently exist or are under development. These subtype specific targets may not show a significant response when leiomyosarcoma cases are considered as a single homogenous entity in a clinical trial. Recognition of molecular subtypes in leiomyosarcoma may highlight preferential response in that subtype for which the target gene is expressed at higher levels.

In summary, we used gene expression profiling and immunohistochemical assays to demonstrate the existence of 3 distinct molecular subtypes in leiomyosarcoma in two independent sets of cases and showed that they are associated with distinct clinical outcomes. These findings indicate distinct biologic subclasses in leiomyosarcoma that may respond differently to novel therapeutic approaches.

Disclosure of Potential Conflicts of Interest

C-H. Lee is a member of the AstraZeneca Ovarian Cancer Advisory Board. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions

Conception and design: X. Guo, R.B. West, M. van de Rijn
Development of methodology: X. Guo, S. Zhu, R.B. West, M. van de Rijn
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): X. Guo, V.Y. Jo, A.M. Mills, S.X. Zhu, C.-H. Lee, I. Espinosa, M.R. Nucci, S. Anderson, C.D. Fletcher, M. van de Rijn
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): X. Guo, T. Hastie, K. Ganjoo, A.H. Beck, R.B. West, C.D. Fletcher, M. van de Rijn
Writing, review, and/or revision of the manuscript: X. Guo, V.Y. Jo, C.-H. Lee, I. Espinosa, E. Forgo, K. Ganjoo, A.H. Beck, R.B. West, C.D. Fletcher, M. van de Rijn
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): X. Guo, V.Y. Jo, S.X. Zhu, S. Varma, M. van de Rijn
Study supervision: X. Guo, M. van de Rijn

Acknowledgments

The authors thank the Stanford Center for Genomics and Personalized Medicine with help on Illumina sequencing and Alayne Brunner, Patricia Marian Robinson, Joanna Pezibyl, and other labmates for sample collection, helpful discussions and experimental design.

Grant Support

This study was supported by the NIH (grant no. CA 112270; to M. van de Rijn) and the Fund for Health Research (FIS PI12/01645; to I. Espinosa). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 17, 2014; revised April 7, 2015; accepted April 11, 2015; published OnlineFirst April 20, 2015.

References

Molecular Subtypes in Leiomyosarcoma


