Gene Expression Profiling in Leiomyosarcomas and Undifferentiated Pleomorphic Sarcomas: SRC as a New Diagnostic Marker

Rolando A. R. Villacis1, Sara M. Silveira1, Mateus C. Barros-Filho1, Fabio A. Marchi2, Maria A. C. Domingues3, Cristovam Scapulatempo-Neto4, Samuel Aguiar Jr5, Ademar Lopes5, Isabela W. Cunha6, Silvia R. Rogatto1,7

1 Neogene Laboratory, Research Center (CIPE), A. C. Camargo Cancer Center, São Paulo, Brazil, 2 Inter-Institutional Grad Program on Bioinformatics, Mathematics and Statistics Institute, USP - University of São Paulo, São Paulo, Brazil, 3 Department of Pathology, School of Medicine, UNESP - São Paulo State University, Botucatu, São Paulo, Brazil, 4 Molecular Oncology Research Center, Barretos Cancer Hospital, Barretos, São Paulo, Brazil, 5 Department of Pelvic Surgery, A. C. Camargo Cancer Center, São Paulo, Brazil, 6 Department of Pathology, A. C. Camargo Cancer Center, São Paulo, Brazil, 7 Department of Urology, School of Medicine, UNESP - São Paulo State University, Botucatu, São Paulo, Brazil

Abstract

Background: Undifferentiated Pleomorphic Sarcoma (UPS) and high-grade Leiomyosarcoma (LMS) are soft tissue tumors with an aggressive clinical behavior, frequently developing local recurrence and distant metastases. Despite several gene expression studies involving soft tissue sarcomas, the potential to identify molecular markers has been limited, mostly due to small sample size, in-group heterogeneity and absence of detailed clinical data.

Materials and Methods: Gene expression profiling was performed for 22 LMS and 22 UPS obtained from untreated patients. To assess the relevance of the gene signature, a meta-analysis was performed using five published studies. Four genes (BAD, MYOCD, SRF and SRC) selected from the gene signature, meta-analysis and functional in silico analysis were further validated by quantitative PCR. In addition, protein-protein interaction analysis was applied to validate the data. SRC protein immunolabeling was assessed in 38 UPS and 52 LMS.

Results: We identified 587 differentially expressed genes between LMS and UPS, of which 193 corroborated with other studies. Cluster analysis of the data failed to discriminate LMS from UPS, although it did reveal a distinct molecular profile for retroperitoneal LMS, which was characterized by the over-expression of smooth muscle-specific genes. Significantly higher levels of expression for BAD, SRC, SRF, and MYOCD were confirmed in LMS when compared with UPS. SRC was the most value discriminator to distinguish both sarcomas and presented the highest number of interaction in the in silico protein-protein analysis. SRC protein labeling showed high specificity and a positive predictive value therefore making it a candidate for use as a diagnostic marker in LMS.

Conclusions: Retroperitoneal LMS presented a unique gene signature. SRC is a putative diagnostic marker to differentiate LMS from UPS.

Citation: Villacis RAR, Silveira SM, Barros-Filho MC, Marchi FA, Domingues MAC, et al. (2014) Gene Expression Profiling in Leiomyosarcomas and Undifferentiated Pleomorphic Sarcomas: SRC as a New Diagnostic Marker. PLoS ONE 9(7): e102281. doi:10.1371/journal.pone.0102281

Editor: Adriano Angelucci, University of L’Aquila, Italy

Received February 14, 2014; Accepted June 17, 2014; Published July 16, 2014

Copyright: © 2014 Villacis et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by grants from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Conselho Nacional de Pesquisa (CNPq) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: silvia.rogatto@cipe.accamargo.org.br
The morphological and histopathological similarities between high-grade LMS and UPS are also observed at molecular level. Although copy number alterations have been described as similar in both tumors [9,10], gains of 1q21.3 have been reported as an independent prognostic marker for shorter survival in patients with LMS [11]. Hierarchical clustering analysis of transcriptomic data for a large series of STS failed to discriminate LMS and UPS. However, tumor subgroups with similar gene expression profiles have been reported [12-14]. Studies that combine gene-expression and DNA copy number alterations have also failed to differentiate between these neoplasms [15], however, three novel molecular subtypes of LMS have been described [6], [16], in addition to a genetic signature capable of predicting metastasis in UPS and LMS [17]. More recently, retroperitoneal LMS (LMS-R) has been described as different from LMS of extremities at the molecular level, demonstrating a distinct clinical outcome [18].

In this study, gene expression profiles were evaluated in 44 STS samples obtained from untreated patients prior to surgery, aiming to identify molecular biomarkers.

Materials and Methods

Patients

Forty-four fresh frozen tissue samples (22 UPS and 22 LMS) were obtained prior to chemotherapy or radiotherapy from 43 patients, at the A.C. Camargo Cancer Center (Sa˜o Paulo, Brazil) and Barretos Cancer Hospital (Barretos, Sa˜o Paulo, Brazil), between 2002 and 2010. Two primary tumors (UPS8 and UPS18) were obtained from the same patient. Ninety formalin-fixed paraffin-embedded (FFPE) samples (52 LMS and 38 UPS, which included 12 LMS and 7 UPS evaluated by oligoarrays) were analyzed using IHC assays. Study design is shown in Figure 1A.

Figure 1. Flow diagrams showing the study design (A) and meta-analysis (B).

doi:10.1371/journal.pone.0102281.g001

Transcriptional profiling (oligoarrays)

Total RNA was extracted using Trizol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer’s instructions. The RNA quality was assessed using the RNA 6000 Nano Kit on the Agilent 2100 Bioanalyzer platform (Agilent Technologies, Palo Alto, CA, USA). Only samples with RIN (RNA integrity number) >7 were considered. The Two-Color Human GE 4X44K Microarrays platform (Agilent Technologies) was used [21]. A combination of equal amounts of total RNA obtained from 15 different cell lines was used as reference [22]. Hierarchical clustering analysis (HCL) of the most variant probes (standard deviation, SD >1.5) was used to visualize cluster samples. Significance Analysis of Microarray (SAM) was applied to


**Table 1.** Pathological features of 22 LMS and 22 UPS samples evaluated by large-scale expression analysis.

| Variable                  | LMS (N=22) | UPS (N=22) |
|---------------------------|------------|------------|
| **Sample source**         |            |            |
| Primary tumour            | 17 (77%)   | 17 (77%)   |
| Recurrence                | 5 (23%)    | 5 (23%)    |
| **Topology**              |            |            |
| Extremity                 | 9 (41%)    | 11 (50%)   |
| Retropertioneum           | 10 (45%)   | 6 (27%)    |
| Trunk                     | 2 (9%)     | 2 (9%)     |
| Head and neck             | 0          | 2 (9%)     |
| Pelvis                    | 1 (5%)     | 1 (5%)     |
| **Tumor size**            |            |            |
| <5 cm                     | 5 (23%)    | 3 (14%)    |
| >5 cm                     | 17 (77%)   | 19 (86%)   |
| **Histological grade**    |            |            |
| I                         | 4 (18%)    | 0          |
| II                        | 5 (23%)    | 0          |
| III                       | 13 (59%)   | 22 (100%)  |
| **Local recurrence**      |            |            |
| Yes                       | 8 (36%)    | 10 (45%)   |
| No                        | 14 (64%)   | 12 (55%)   |

*not considered surgery extension of surgical margins.*

doi:10.1371/journal.pone.0102281.t001

...continue...
Supervised hierarchical clustering (SHCL) using SAM analysis revealed 587 differentially expressed genes in LMS compared to UPS, being 580 over-expressed and 7 down-expressed, but failed to discriminate between the neoplasms. However, over again the retroperitoneal LMS cases demonstrated similar expression profiles (Figure 2B). Only two LMS of the extremities (LMS14 and LMS15) clustered with the LMS-R samples. The IPA network analysis generated three main interactions networks (Table S2). The first network was associated with skeletal and muscular system development and function, tissue morphology, and cellular assembly and organization, including ACTA2, MYLK, MYOCD and SRF (Figure S1A). The second was related to cellular movement, cell morphology and cellular assembly and organization, and encompassed ILK, CXCL1, LAMA5 and SRC (Figure S1B). The third network harbored genes associated with cell death, DNA replication, recombination, repair and gene expression, including FLNA, MEF2C, CHEK1 and BAD (Figure S1C).

Meta-analysis and protein-protein interaction network analysis

Aiming to validate the signature obtained in the comparison between LMS and UPS from the SAM analysis, a meta-analysis including five studies was conducted (Table 2) and revealed 316 of 587 differentially expressed genes obtained by comparison between LMS and UPS. A total of 193 differentially expressed genes (191 over-expressed and 2 down-expressed) in LMS compared with UPS exhibited the same expression pattern in all studies, including several genes related to muscular function (Table S3). Protein-protein interaction analysis involving the 193 validated meta-analysis genes revealed a network of 59 proteins with more than 10 interaction partners (Figure 3). SRC protein presented the highest number of interactions, in particular with ACTN1, AR, FLNA and MUC1.

Real-time quantitative RT-PCR of BAD, SRC, SRF, and MYOCD

Based on the gene signature, gene network analysis, gene function and meta-analysis, BAD, SRC, SRF and MYOCD were selected to be further investigated by RT-qPCR. SRF, SRC and BAD were detected as central node genes in the main networks generated by IPA software (Figure S1), supporting their potential involvement in LMS development.

Similarly to the gene expression microarray, all genes showed significant over-expression \( (P<0.01) \) in LMS when compared to UPS (Figure 4A). Although not significant, MYOCD presented over-expression in LMS-R compared to non-retroperitoneal LMS.
Of all genes tested, **SRC** was revealed to be the most valuable discriminator for distinguishing between LMS and UPS (area under the ROC curve, AUC = 0.897 and accuracy = 92.0%) (Table S4). Two samples were classified incorrectly (LMS24 and UPS8) when using SRC as the discriminator.

**SRC** protein expression by immunohistochemistry

**SRC** protein revealed negative/weak immunostaining (score 0–1) in 30 LMS and 36 UPS while positivity (scores 2–3) was detected in 22 LMS and two UPS (P<0.001) (Figure 4B). Although most LMS-R presented positivity for **SRC** (5/7), no significant difference was observed between LMS-R and LMS-NR (Table 3). **SRC** positivity exhibited high specificity (94.7%) and a positive predictive value of 91.7% in discriminating LMS from UPS.

**Figure 3.** Protein-protein interaction network of 59 genes derived from the meta-analysis. Only proteins with more than 10 interactions partners were considered. The triangle size is proportional to the number of interactions for each protein. **SRC** presented the largest number of interactions, including four strong physical interactions found in the network. The interaction partners for each protein were obtained from the Interologous Interaction Database (I2D) 2.0 and the network was visualized and analyzed with the NAViGaTOR 2.3 software.

doi:10.1371/journal.pone.0102281.g003

**Table 2.** Soft-tissue sarcoma profiling studies included in the meta-analysis.

| GEO accession | Study | Platform | Samples | Tumors included in the current study |
|---------------|-------|----------|---------|-------------------------------------|
| GSE3443       | Nielsen et al. [12] | Custom cDNA array | 46      | 11 LMS vs. 6 UPS (MFH)             |
| GSE2719       | Detwiller et al. [25] | Affymetrix HG-U133A | 54      | 6 LMS vs. 9 UPS (MFH)             |
| GSE6481       | Nakayama et al. [13] | Affymetrix HG-U133A | 105     | 6 LMS vs. 21 UPS (MFH)           |
| GSE21122      | Barretina et al. [26] | Affymetrix HG-U133A | 158     | 26 LMS vs. 3 UPS (pleomorphic MFH) |
| GSE21050      | Chibon et al. [17]  | Affymetrix HG-U133A Plus 2 | 310   | 85 LMS vs.136 UPS                |

GEO: Gene Expression Omnibus; LMS: Leiomyosarcoma; MFH: malignant fibrous histiocytoma; UPS: undifferentiated pleomorphic sarcoma.

doi:10.1371/journal.pone.0102281.t002
Discussion

Large-scale gene expression studies in STS have many limitations, including evaluation of only a small number of fresh-frozen samples, a large number of histological subtypes included within the same report, a small number with the same histological subtype and the inclusion of cases treated with chemotherapy or radiotherapy prior to sample collection [3,31]. Overall, these limitations reduce the probability of identifying recurrent genetic alterations with clinical significance. In this study, 22 UPS and 22 LMS samples were collected prior to treatment with chemotherapy or radiotherapy, with their diagnosis confirmed using the markers recommended in the literature.

Although the UHCL analysis was not able to accurately discriminate LMS from UPS, a distinct molecular profile for LMS-R was established, independent of histological grade (cluster I). Cluster II was divided in two subgroups and included the high-grade LMS-NR and UPS from different anatomical sites. It is important to highlight that one patient developed two UPS (UPS8 and UPS18), which clustered together and demonstrated a similar gene expression profile, suggesting that these samples were not two primary tumors as previously diagnosed. Comparison of the clusters with the clinicopathological findings and overall survival revealed no significant difference, probably due to the fact that most of the tumors presented with a high histological grade and similar prognosis.

Several reports have shown a strong similarity between UPS and LMS, suggesting that they share common oncogenic pathways and may correspond to different stages of the same tumor entity [9,14,15]. However, this study has demonstrated that LMS-R clusters separately. Both UHCL and SHCL analysis suggested that high-grade LMS-NR demonstrates a gene expression profile with a closer proximity to UPS than to LMS-R. Almost all genes that distinguished LMS-R from other sarcomas were over-expressed (580 of 587 genes).

The function and interaction network analysis revealed a large number of genes associated with muscle structure and function, including ACTG2, CALD1, DMD, MYOCD, MILK, SRF and TAGLN. A meta-analysis with datasets from five studies supported these findings. In all studies, 193 of the 316 genes evaluated were over-expressed in LMS compared to UPS. A considerable number of genes related to muscle function were found to be over-expressed in all studies, including ACTC1, ACTA2, ACTN1, CALD1, CNN1, MILK, MYL9, DMD, SGCA, TAGLN, TPM1, TPM2 and SMTN.

In LMS, over-expression of genes associated with muscle structure has been previously reported in the literature [12,14,16,32]. Recently, Italiano et al. [18] used genomic and transcriptomic analysis for 73 LMS and described a distinct clinical and molecular profile for LMS-R compared to LMS of the extremities. This finding was based mainly on over-expression of genes associated to muscle differentiation. MYOCD, a transcriptional cofactor of SRF, was found to be associated with muscular differentiation in well-differentiated retroperitoneal LMS [14,18,32].

In this study, overexpression of MYOCD, SRF, BAD and SRC in LMS was found, predominantly in LMS-R. These genes were tested for their potential as diagnostic markers to distinguish LMS from UPS. Combined accuracy, sensitivity and specificity were calculated for these genes, and SRC was identified as the most valuable discriminator. In addition, a higher level of SRC protein expression was detected in both LMS-R and LMS-NR when compared with UPS. Despite the use of SRC as a unique marker not being highly sensitive (<50%), its high specificity and positive

![Figure 4](https://example.com/figure4.png)
predictive value (90%) supports its potential use when in combination with other conventional muscular markers. It is important to highlight that in silico analysis of protein-protein interaction using the genes obtained in the meta-analysis showed a large number of interactions involving SRC, therefore suggesting its important role in the pathogenesis of LMS.

Conclusions

A distinct molecular profile was demonstrated for LMS-R, mediated by over-expression of genes involved in muscular development and function. Both well-differentiated LMS-R [14,20] and high grade LMS-R presented over-expression of MYOCD. Meta-analysis confirmed the involvement of 193 differentially expressed genes in the comparison between LMS and UPS. Our findings suggested that LMS-R is a new molecular entity of sarcomas whose alterations could be useful to stratify patients on specific therapeutic protocols. In addition, in-silico analysis revealed SRC as a central gene associated with LMS. Furthermore, SRC protein expression could be useful as a diagnostic marker to differentiate between LMS and UPS, especially in cases where muscular markers are negative.

Supporting Information

Figure S1 Graphic representation of the three interaction networks with the genes over-expressed (red) and down-regulated (green) in LMS compared to UPS. Genes were associated with skeletal and muscular system development and function, tissue morphology, cellular assembly and organization in first network (A); related to cellular movement, cell morphology and cellular assembly and organization in second network (B); and associated with cell death, DNA replication, recombination, repair and gene expression in third network (C). The red and green colors tones are proportional to intensity of expression for each gene. The genes selected for validation are indicated in blue circles. Image adapted from Ingenuity Pathway Analysis (IPA) software.

Table S1 Primer sequences and properties of the transcripts evaluated by RT-qPCR.

Table S2 Description of differentially expressed genes in LMS and UPS categorized in three main interaction networks according to IPA software.

Table S3 Meta-analysis of 316 genes from the supervised analysis present in five array datasets categorized in three main interaction networks according to IPA software.

Table S4 Diagnostic criteria and IHC antibody panel used to define LMS and UPS.

Table S1 Diagnostic criteria and IHC antibody panel used to define LMS and UPS.

Table S1 Diagnostic criteria and IHC antibody panel used to define LMS and UPS.

Table S1 Diagnostic criteria and IHC antibody panel used to define LMS and UPS.

Table S1 Diagnostic criteria and IHC antibody panel used to define LMS and UPS.

Table S1 Diagnostic criteria and IHC antibody panel used to define LMS and UPS.
Author Contributions

Conceived and designed the experiments: SRR IWC RARV. Performed the experiments: RARV SMS. Analyzed the data: RARV SMS MCBF.

References

1. Guillou L, Aurias A (2010) Soft tissue sarcomas with complex genomic profiles. Virchows Arch 456: 201–217. doi: 10.1007/s00428-009-0553-4.

2. Nielsen TO, West RB (2010) Translating gene expression into clinical care: sarcomas as a paradigm. J Clin Oncol 28: 1796–1805. doi: 10.1200/JCO.2009.26.1917.

3. West RB (2010) Expression profiling in soft tissue sarcomas with emphasis on synovial sarcoma, gastrointestinal stromal tumor, and leiomyosarcoma. Adv Anat Pathol 17: 366–373. doi: 10.1097/PAP.0b013e3181e74728.

4. Toro JR, Travis LB, Wu HJ, Zha K, Fletcher CD, et al. (2006) Incidence of molecular subtypes in leiomyosarcoma through integrative molecular profiling. J Surg Pathol 31: 502–510. doi: 10.1002/mod.20120.

5. Fletcher CD (2014) The evolving classification of soft tissue tumours: an update manuscript: RARV SMS MCBF FAM MACD CSN IWC. Collected the tissue samples and performed the histopathological analysis: MACD CSN IWC. Contributed reagents/materials/analysis tools: SA AL FAM. Wrote the paper: RARV SSR. Performed the histopathological analysis: MACD CSN IWC. Collected the tissue samples and clinical information: SA AL. Read and approved the final version of the manuscript: RARV SMS MCBF FAM MACD CSN SA AL IWC SRR.

16. Beck AH, Lee CH, Witten DM, Gleason AG, Edris B, et al. (2010) Discovery of molecular subtypes in leiomyosarcoma through integrative molecular profiling. Oncogene 29: 645–654. doi: 10.1038/onc.2009.301.

17. Chibon F, Lagarde P, Salas S, Pérot G, Brouste V, et al. (2010) Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. Nat Med 16: 781–787. doi: 10.1038/nm.2174.

18. Italiano A, Lagarde P, Brulard C, Terrier P, Lae M, et al. (2013) Genetic profiling identifies two classes of soft-tissue leiomyosarcomas with distinct clinical characteristics. Clin Cancer Res 19: 1190–1196. doi: 10.1158/1078-0432.CCR-12-2970.

19. Boccardo E, Manzini Baldi CV, Carvalho AF, Rabachini T, Torres C, et al. (2010) Expression of human papillomavirus type 16 E7 oncoprotein alters keratinocytes expression profile in response to tumor necrosis factor-alpha. Carcinogenesis 31: 511–518. doi: 10.1093/carcin/bgp333.

20. Guillou L, Coindre JM, Bonichon F, Nguyen BB, Terrier P, et al. (1997) Comparative study of the National Cancer Institute and French Federation of Cancer Centers Sarcoma Group grading systems in a population of 410 adult patients with soft tissue sarcoma. J Clin Oncol 15: 330–362. doi: 10.1111/j.1553-2712.1997.tb06790.x.

21. Cirilo PD, Marchi FA, Barros Filho MdeC, Rocha RM, Domingues MA, et al. (2013) An integrative genomic and transcriptomic analysis reveals potential targets associated with cell proliferation in uterine leiomyomas. PLoS One 8: e57901. doi: 10.1371/journal.pone.0057901.

22. Weigelt B, Baehner FL, Reis-Filho JS (2010) The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. J Pathol 220: 263–280. doi: 10.1002/path.2648.

23. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98: 115–121. doi: 10.1073/pnas.091062498.

24. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6:e1000097. doi: 10.1371/journal.pmed.1000097.

25. Detwiller KY, Fernando NT, Segal NH, Ryeom SW, D’Amore PA, et al. (2005) Analysis of hypoxia-related gene expression in sarcomas and effect of hypoxia on RNA interference of vascular endothelial cell growth factor A. Cancer Res 65: 5881–5889. doi: 10.1158/0008-5472.CAN-04-4078.

26. Barretina J, Taylor BS, Banerji S, Ramos AH, Lagos-Quintana M, et al. (2010) Comparison of the National Cancer Institute and French Federation of Sarcoma bone. Lyon: IARC Press. 427 p.

27. Italiano A, Lagarde P, Brulard C, Terrier P, Lae M, et al. (2013) An integrative genomic and transcriptomic analysis reveals potential targets associated with cell proliferation in uterine leiomyomas. PLoS One 8: e57901. doi: 10.1371/journal.pone.0057901.

28. Brown KR, Otasek D, Ali M, McGuffin MJ, Xie W et al. (2009) NAViGaTOR: Comparative study of the National Cancer Institute and French Federation of Sarcoma bone. Lyon: IARC Press. 427 p.

29. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3: research0034.1. doi: 10.1186/gb-2002-3-7-research0034.

30. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Bioinformatics 17: 402–408. doi: 10.1093/bioinformatics/17.4.309.

31. Weigelt B, Baehner FL, Reis-Filho JS (2010) The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. J Pathol 220: 263–280. doi: 10.1002/path.2648.

32. Weigelt B, Baehner FL, Reis-Filho JS (2010) The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. J Pathol 220: 263–280. doi: 10.1002/path.2648.

33. Weigelt B, Baehner FL, Reis-Filho JS (2010) The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. J Pathol 220: 263–280. doi: 10.1002/path.2648.