Genomics 99 (2012) 275–281

Contents lists available at SciVerse ScienceDirect

Genomics

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

Characterization of the uterine leiomyoma microRNAome by deep sequencing

Boryana Georgieva [a,1], Ivan Milev [a,1], Ivan Minkov [a], Irina Dimitrova [b], Andrew P. Bradford [b], Vesselin Baev [a,*]

[a] Department of Plant Physiology and Molecular Biology, University of Plovdiv, 24 TsarAssen St, 4000 Plovdiv, Bulgaria
[b] Department of Obstetrics and Gynecology, University of Colorado School of Medicine, Aurora, Colorado 80045, USA

ARTICLE INFO

Article history:
Received 6 January 2012
Accepted 9 March 2012
Available online 16 March 2012

Keywords:
miRNAs
Small RNAs
Deep sequencing
Uterine leiomyoma
Myometrium

ABSTRACT

MicroRNAs (miRNAs) are a class of small non-coding RNAs, which are negative regulators of gene expression. Many genes in human uterine leiomyoma (ULM) are aberrantly expressed and in some cases this can be due to dysregulation of miRNAs. Here we present the first study to determine genome-wide miRNA expression patterns in uterine leiomyoma and myometrium using Solexa high-throughput sequencing. We found more than 50 miRNAs, which were differentially expressed, and furthermore we extend the list of putative new miRNA genes. The top five significantly de-regulated miRNAs in ULMs that we found in our libraries were miR-363, miR-490, miR-137, miR-217 and miR-4792. We also observed "isomiRs" with higher copy number than referenced mature miRNA specific for the leiomyoma libraries, which have a potential role in tumorigenesis. The microRNA transcriptomes obtained in this study deliver insights and further expand our understanding the role of small RNAs in uterine leiomyoma development.

1. Introduction

MicroRNAs (miRNAs) are a new class of negative regulators that repress gene expression by binding with their target messenger RNAs (mRNAs). Through this pairing to the 3′UTR of mRNAs, miRNAs trigger degradation of the target transcript or inhibition of protein translation [1,2]. Biogenesis of miRNA involves several steps. Firstly, miRNA genes are transcribed by RNA polymerase II into primary RNA (pri-miRNA) molecule. The pri-miRNA is subsequently cleaved in the nucleus by an RNase III endonuclease, Drosha, forming a 70–100 nucleotide (nt) hairpin loop structure known as precursor miRNA. Next, the pre-miRNA is exported into the cytoplasm and further processed, by Dicer, into a double stranded RNA duplex with 3′-overhangs [3–5]. Based on the thermodynamic stability of this duplex, one of the strands is selected as the biologically active mature miRNA whilst the other, known as the passenger strand or miRNA star, is typically degraded [6].

The mature miRNAs are short (20–24 nt) single stranded RNAs that, together with Argonaute proteins, mediate direct post-transcriptional regulation. miRNAs are essential for normal mammalian development and regulate genes involved in cell division and differentiation, metabolism, stress response and apoptosis [7]. miRNAs have been shown to regulate oncogenes, tumor-suppressors and a number of cancer-related genes, therefore their deregulation can predispose to disease and malignancies [8,9]. Most importantly, different cancer types, and stages exhibit unique miRNA expression patterns, suggesting that miRNA genes can function as novel biomarkers for cancer diagnosis [8,10,11]. Moreover, miRNA signatures are also altered in variety of benign tumors such as human uterine leiomyomas (ULMs) [12–14].

Leiomyomata uteri or fibroids are the most common neoplasm of the female genital tract developing primarily during the reproductive years and becoming symptomatic during perimenopause [15–18]. Tumors occur in approximately 80% of women, and approximately 25% of Caucasians have clinically significant, symptomatic lesions. Most affected women have multiple tumors with the average number of tumors per uterus estimated to be 6.5 [19]. The relative risk of fibroids is two to threefold greater in black women than white women and clinical disease is more severe in the former [20]. The biological sequelae of fibroids manifest as a spectrum of clinical symptoms, primarily pain and excessive menstrual bleeding. Reproductive issues are also a concern, as fibroids are associated with infertility and, if present during pregnancy, may contribute to second trimester pregnancy losses, premature labor, fetal malpresentations and/or distocia. Despite the significant impact on women's health, the etiology of fibroids remains poorly understood and only a few specific genes have been identified to be associated with development of this tumor.

Little is known about the repertoire and function of miRNA in human uterine leiomyoma. Several studies have assessed levels of subsets of miRNAs through microarray expression analysis [12–14], demonstrating that many miRNAs are deregulated in leiomyoma compared to normal tissue. To profile the small RNAome in leiomyomas and myometrium issues and to reveal deeper insights into miRNAs expression changes between diseased and normal states, we used Solexa high-throughput sequencing technology. This method has significant advantages over microarray and PCR-based assays in characterizing the miRNA

* Corresponding author.
E-mail address: vebaev@gmail.com (V. Baev).
[1] Contributed equally to this work.

0888-7543/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2012.03.003

transcriptome [21–23]. Unlike microarray, deep-sequencing is not limited by array content and thus allows identification of novel miRNAs. Moreover, parallel sequencing has the ability to identify low abundance miRNAs and affords an opportunity for detection of small differences in miRNAs expression between samples [21].

To the best of our knowledge, the differential expression of miRNAs has not been studied via deep-sequencing in ULMs. In this manuscript we used next-generation sequencing to characterize small RNA transcriptome, to compare expression signatures of known miRNAs in leiomyoma versus matched normal myometrium and to discover novel miRNAs. We were able to identify more than 50 miRNAs exhibiting altered expression and 25 genomic loci that encode putative novel miRNA genes, thus expanding the list of miRNAs representing the unique signature of uterine leiomyomas.

2. Results

2.1. Overview — sequencing statistics, lengths and reads

By using Solexa technology, a newly developed method for high-throughput sequencing producing highly accurate and quantitative read-outs of small RNA, four libraries of small RNAs from myometrium (B25N and B27N) and leiomyoma (B25L and B27L) were established and sequenced. After applying reads filters and discarding low quality reads, adaptor sequences were trimmed. Additionally, reads attributed to ligation contaminants or self-ligation of the adaptors were also discarded. We obtained 7,284,165 (B25N) and 5,506,769 (B27N) clean reads from normal myometrium and −7,629,120 (B25L) and 7,423,663 (B27L) from leiomyoma libraries. For the small RNA analysis we further clustered all raw reads into unique sequences with associate copy numbers generating 74,547 (B25N), 48,114 (B27N), 45,538 (B25L) and 59,148 (B27L) unique tags from normal myometrium and leiomyoma samples, respectively.

The length distribution of the high quality tags was consistent among the four RNA libraries. In our analysis of small RNAs we selected sequences in the range of 15 to 30 nt. Results showed that this resulted in enrichment of miRNAs since the fraction of reads within the range 20–22 nt (considered as a typical miRNA length) comprised the majority (80–90%) in all four samples (Fig. 1). Interestingly, a significant 23 nt peak representing 1,237,791 of 7,423,663 reads (~17%) was observed in sample B27L but not in the other libraries.

2.2. Mapping and annotation of small RNAs

The 15–30 nt clean reads from four libraries were subsequently aligned and mapped to the human genome allowing at most one mismatch. Total mappable reads were approximately 99% in all sequencing libraries. After clustering these mappable reads to unique sequences they represented 67.51% (B25N), 66.58% (B27N), 59.34% (B25L) and 61.96% (B27L) of the sequences in myometrium and leiomyoma samples, respectively. Based on the UCSC human genome annotation and additional well-characterized RNA datasets (see Materials and Methods) all small RNA sequences were annotated as known miRNAs, degradation fragments from non-coding RNAs (tRNA, rRNA, snRNA/snoRNA), genomic repeats, mRNAs (exons, introns) or unclassified. The fraction of small RNAs derived from UTR regions of mRNA was very small and therefore not included in our results.

As expected, the most abundant RNA category from the four libraries was known miRNAs: 95% (B25N and B27N); 96% (B25L and B27L). The remaining, less abundant categories were non-coding RNAs and repeat elements. A minor fraction of small RNAs were mapped to coding-sequences, which likely represent degradation products from mRNAs. Read counts of the classified categories are summarized in Table 1.

Recent studies showed that some genomes can generate repeat associated small RNAs, which can regulate retrotransposon expression. Yang et al., demonstrated that small interfering RNAs, which are endogenously coded, can suppress the L1 retrotransposon in human cell cultures [24]. Therefore, we additionally explored the repertoire of small RNAs that are aligned to genome repeats, and represent a set of putative RNAs associated with transposable elements. In the mammalian genome, transposon elements are mostly composed of three major classes: LTR, LINE and SINE retrotransposons. Our results revealed that the majority of reads mapping to repeat elements were

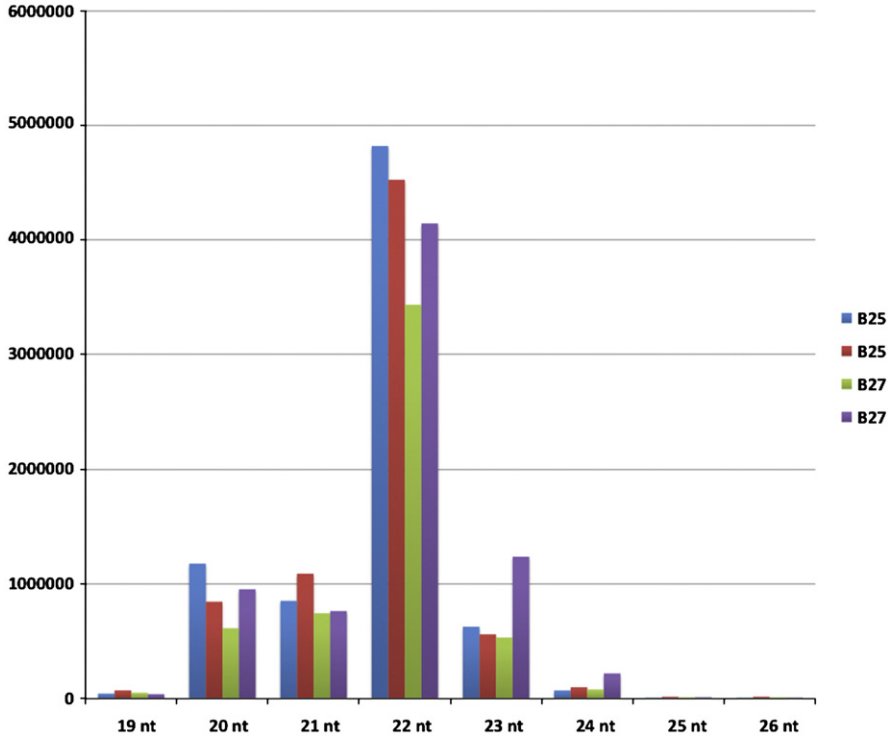

Fig. 1. The length distribution of small RNAs in leiomyoma (B25L and B27L) and myometrial (B25N and B27N) libraries.