MDS Diagnosis Using cfRNA Profiling by Targeted Enrichment Next Generation Sequencing

#4268

Cynthie Wong[1] • Vincent Funari[1] • Wanlong Ma[1] • Ivan De Dios[1] • Forrest Blocker[1] • Babak Shahbaba[2] • Maher Albitar[1]

[1]NeoGenomics Laboratories, Aliso Viejo, CA • [2]University of California, Irvine, CA

Introduction

The confirmation of a diagnosis of myelodysplastic syndrome (MDS) can be very difficult based on morphology alone. The detection of cytogenetic abnormalities or the presence of mutations in one or more of the myeloid-related genes can provide significant help in confirming this diagnosis. However, some patients with MDS may lack mutations or cytogenetic abnormalities. Cell free DNA (cfDNA) and microRNA are used extensively in testing for molecular abnormalities in hematologic and solid tumors, but little is known about the value of cell free mRNA (cfRNA) in liquid biopsy. We explored the use of targeted enrichment and Next Generation Sequencing (NGS) of cfRNA in the RNA profiling and diagnosis of MDS.

Methods

We used NucliSenS EasyMAG (bioMérieux, Marcy-l'Étoile, France) automated platform for extracting total nucleic acid from peripheral blood plasma collected in EDTA. We used the TruSight RNA Pan-Cancer Panel (Illumina, San Diego, CA) for detecting fusion, expression, and mutations in 1385 genes. The Pan-Cancer Panel is a hybridization based targeted panel, therefore alignments and gene expression are calculated only to the 1385 genes as per manufacturer recommendations. Each sample was deeply sequenced (>55M reads) using Illumina paired-end 75 bp sequencing and duplicate reads were removed before fragment per kilobase of transcript per million mapped reads (FPKMs) were calculated for each gene. Samples from 23 patients were analyzed, including 19 patients confirmed by molecular, cytogenetic, and morphologic methods to have MDS or chronic myelomonocytic leukemia (CMML) and 4 individuals who had one or more cytopenia, but MDS was ruled out by molecular mutation analysis, cytogenetics, as well as morphology. Expression levels were first normalized using *ABL1* and *PAX5* (i.e., regressing each gene on *ABL1* and *PAX5* jointly and using the residuals instead of the raw data) and models for distinguishing MDS from non-MDS were developed.

Mapped Reads to *CBFB*, *RASGRF2*, & *SP-1* Genes

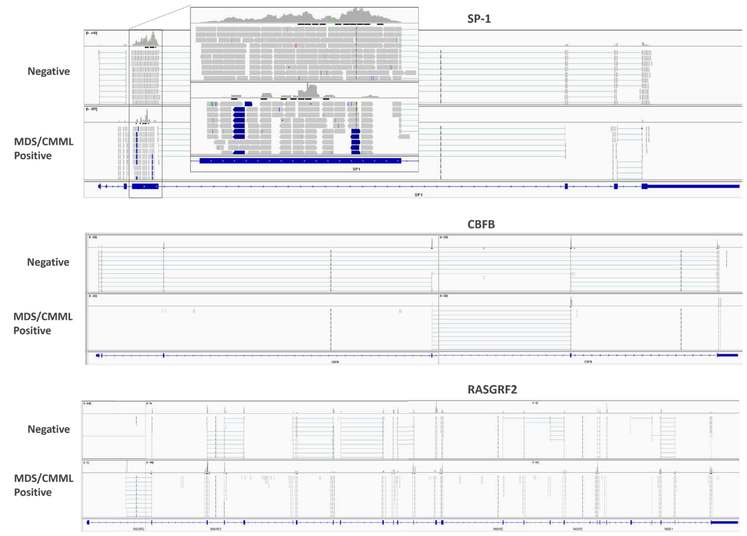

Figure 1: Number of aligned reads to exon regions of *SP-1*, *CBFB*, and *RASGRF2* from MDS/CMML and normal samples corroborate statistical findings. Note: Representative MDS/CMML sample had three times as many reads as the negative sample.

- Read alignments for MDS/CMML samples illustrate the relatively intact mRNA found in plasma for these genes. Figures show relatively uniform distribution of aligned reads to the gene and confirm good capture and sequencing of the entire transcript.
- Total number of reads aligned to the *CBFB* and *SP-1* genes were significantly less than the number aligned for the normal samples. For the samples shown, each MDS sample was sequenced ~5X deeper than the normal sample making this finding more pronounced. The statistical significance of these differences are shown in Figure 2.
- In general, samples had ca. 75% bases aligned to coding regions, 15% aligned to UTR, 5% aligned to intron regions, and 5% were aligned to non-targeted regions.
- Majority of the reads from MDS/CMML samples results from good quality mRNA and highlights our ability to extract intact cfRNA from plasma specimens.

Standardized Expression of 3 Significant Genes: *CBFB*, *RASGRF2*, & *SP-1*

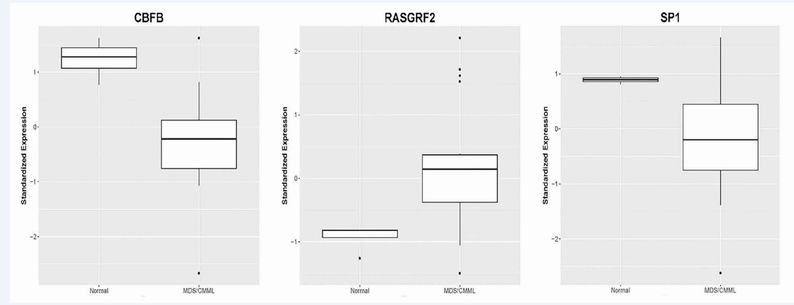

Figure 2: Standardized expression of three genes: *CBFB*, *RASGRF2*, and *SP-1*. As shown, there is significant difference between normal control and MDS patients with minimal overlap.

- Statistical significance analysis only considered expression levels and not the mutational profile.
- T-tests were used to examine the statistical significance of each gene separately. Forty-nine genes with p-values < 0.05 were determined to be significant. After adjusting for multiple hypothesis testing, only three genes remained significant with a false discovery rate set at 0.1. The three genes were: *CBFB*, *RASGRF2*, and *SP-1*.
- Logistic regression model based on these three genes obtained an accuracy rate of 87.0% and AUC (area under the ROC curve) of 92.1% based on the leave-one-out testing. This model identified all the normal cases correctly, but misclassified 3 MDS cases as normal.
- An alternate method using top 3 principal components based on the selected 49 significant genes in a linear regression model can improve prediction and AUC to 95.6% and 96.7%, respectively, based on the leave-one-out testing. In this model, only one MDS case (out of 19) was misclassified as normal, while all the normal cases were classified correctly.

Conclusion

Although additional studies and validations are needed, RNA profiling using targeted enrichment NGS of cfRNA is potentially useful in the classification and diagnosis of MDS. This model can be expanded to incorporate an entire mutational profile assessment. This approach has the potential to be used not only for MDS diagnosis, but also in the diagnosis and predictive response to therapy of a variety of myeloid and lymphoid neoplasms.

Overexpression of 3 Proteins in Cytogenetically Normal MDS Patients

- *CBFB* is expressed in all myeloid cells and T lymphocytes, but the expression level is normally low in adult B cells. (Kundu M, Chen A, Anderson S, et al. *Blood.* 2002;100(7):2449-2456.)
- *RASGRF2* is expressed in T cells and activates Ras. (Ruiz S, Santos E, Bustelo XR. RasGRF2, *Mol Cell Biol.* 2007;27(23): 8127-8142.) *RASGRF2* is implicated as a tumor suppressor and is suppressed in many cancers. (Fernández-Medarde A, Santos E. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer.* 2011;1815(2):170-188.)
- *SP-1* is over-expressed in many cancers, (Beishline K, Azizkhan-Clifford J. *FEBS.* 2015;282(2):224-258), activates cKIT, and shares binding sites with *CBFB*.