Transcriptome-wide analysis demonstrated significant differences in gene expression variability between WM and IgM-MGUS BM B cell clones

Alessandra Trojani[1], Milena Lodola[1], Alessandra Tedeschi[1], Antonino Greco[2], Barbara Di Camillo[3], Tiziana Sanavia[3], Anna Maria Frustaci[1], Maddalena Mazzucchelli[1], Chiara Villa[4], Daniela Boselli[4], Enrica Morra[1], Roberto Cairoli[1]

[1]Department of Hematology, ASST Grande Ospedale Metropolitano Niguarda, Milano, Italy; [2]Ospedale Cardinale Panico, Tricase, Italy; [3]Department of Information Engineering, University of Padova, Italy; [4] San Raffaele Scientific Institute, Milano, Italy.

In this study we focused on determining the transcriptome differences between Waldenström's Macroglobulinemia (WM) and IgM Monoclonal Gammopathy of Undetermined Significance (IgM-MGUS) by gene expression profiling (GEP) considering all the different transcript isoforms of genes that map the human transcriptome (coding transcripts, non-coding transcripts). We performed the analysis on BM B cell clones (CD45+,CD38+,CD19+,LAIR-1-,CD27dim,IgM+,CD22dim,CD25+) from WM (*n*=21) and IgM-MGUS (*n*=13) patients. These populations were identified with a 8-colors panel by flow cytometry and successively isolated by cell sorting. GEP of WM vs. IgM-MGUS BM B cell clones was performed using Affymetrix Gene Chip HTA 2.0.

Data was first pre-processed using Affymetrix Expression Console and then normalized using ComBat [1] and quantile normalization. We investigated both differential expression using SAM test [2] and differential variability using F-test to compare means and variances between groups, respectively. In particular, testing the variability is useful to investigate a heterogeneous disease like Waldenström's Macroglobulinemia as well as IgM-MGUS, since B clonal cells proliferation and growth are driven by different mutations acting as perturbations in different molecular pathways, and these perturbations vary from individual to individual. False Discovery Rate (FDR) [3] p-values adjusted for multiple testing below 5% were considered significant. “Genomic Regions Enrichment of Annotations Tool” (GREAT) [4] was used to annotate the selected probe sets and perform biological pathway enrichment analysis.

We considered 67,529 probe sets for the analyses. There were no differentially expressed probe sets in means after the correction for multiple comparisons, whereas 446 probe sets showed differential variability between IgM-MGUS and WM samples. Figure 1 shows how the selected probe sets map on the human genome according to GREAT.

Enrichment analysis performed on these 446 probe sets showed after correction for multiple testing (FDR threshold set at 5%), significant enrichment for apoptosis, B cell receptor signaling

pathway, chemokine signaling pathway, ERBB signaling pathway, PI3K-AKT signaling pathway and WNT signaling pathway. Of note, *BCL2, RAF1, MAPK1, GRB2, GSK3B, NRG1, SOS1, WNT5A, NLK, PTK2B* genes belonging to these pathways, demonstrated significant different expression variability (table 1). We found that IgM-MGUS showed significantly increased variability of expression of all the selected genes (a part from *SOS1* and *NLK*) across patients.

We could speculate that IgM-MGUS B cell clones showed increased expression variability in the identified genes in their developmental stage, indicating the likely presence of cells at different step of differentiation whereas the expression of the same genes was more stable in WM patients.

In summary, we found that IgM-MGUS was characterized by higher variability in gene expression across patients which could be related to higher intra-patient variability suggesting the possible link between expression variability and genetic heterogeneity. Important functions showing increased variability in IgM-MGUS compared to WM were related to apoptosis, B cell receptor signaling pathway, chemokine signaling pathway, ERBB signaling pathway, PI3K-AKT signaling pathway and WNT signaling pathway. Larger datasets and clinical evolution of IgM-MGUS individuals would provide a deeper insight into the functional context of the pathways and the differential variable genes highlighted by the comparison between IgM-MGUS and WM.

References

[1] Johnson WE, Rabinovic A, Li C. Adjusting batch effects in microarray expression data using Empirical Bayes methods. *Biostatistics.* 2007 Jan;8(1):118-127.

[2] Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A.* 2001 Apr24;98(9):5116–5121.

[3] Storey JD. A direct approach to false discovery rates. *J R Stat Soc.* 2002 Aug12;64(3):479–498.

[4] McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* 2010 May;28(5):495-501.

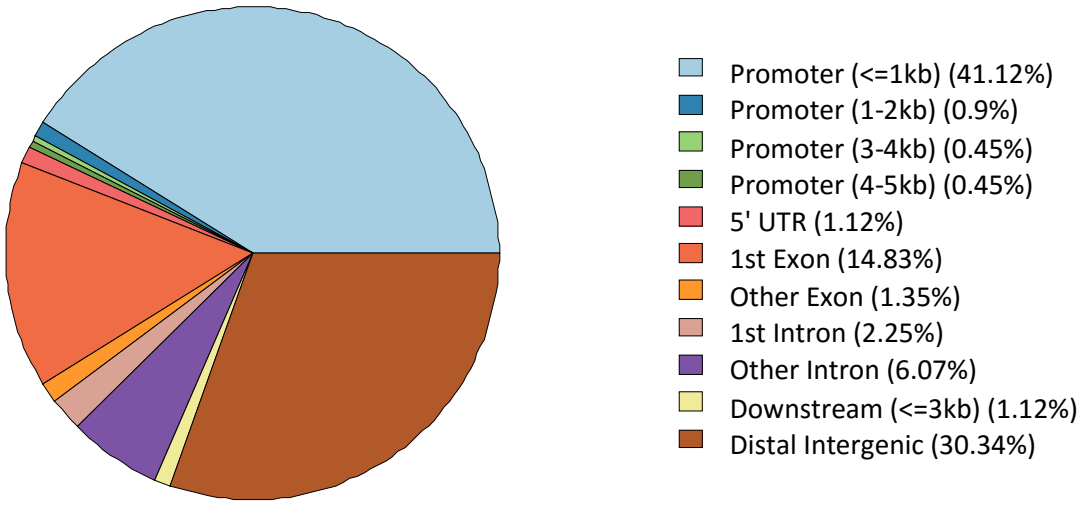

Figure 1. Distribution of the 446 probes showing significant differential variability in IgM-MGUS vs. WM samples.

Pathway (KEGG)	**Genes annotated to the pathway with significant different expression variability**
Apoptosis	*BCL2, RAF1, MAPK1*
B cell receptor signaling pathway	*GRB2, SOS1, RAF1, MAPK1, GSK3B*
Chemiokine signaling pathway	*GRB2, SOS1, RAF1, MAPK1, GSK3B, PTK2B*
ERBB signaling pathway	*NRG1, GRB2, RAF1, MAPK1*
PI3K-Signaling pathway	*GRB2, SOS1, RAF1, MAPK1*
WNT Signaling pathway	*WNT5A, GSK3B, NLK*

Table 1. Pathways significantly enriched with genes showing significant different expression variability in IgM-MGUS vs. WM.