Characterization of small RNAs from Extracellular Vesicles

SUZANNE DUFAULT, NIMA HEJAZI, IVANA MALENICA, & HECTOR ROUX DE BEZIEUX UC Berkeley

School of Public Health
UNIVERSITY OF CALIFORNIA, BERKELEY

INTRODUCTION & OVERVIEW

- The discovery of ribonucleic acid (RNA) inside extracellular vesicles (EV) has raised significant interest in the biological function and biomarker potential of small RNAs, miRNAs in particular.
- This project represents a first-pass attempt to use sequencing to compare EV small RNAs from vesicles and non-vesicles from the same individuals.

DATA

- Sequencing data was generated by Illumina HiSeq2000 and aligned to the reference using miRDeep2 software package.
- The data contained 42 matched subjects with sequenced vesicle and non-vesicle (flowthrough) material. These samples also had recorded background characteristics and cognitive measures, in particular diagnosis, Age, Sex, APOE genotype, MMSE and DRS score.

CLUSTERING: METHODS & RESULTS

- Consensus matrices were generated with PAM and hierarchical clustering, using 1 - Spearman correlation distance on log transformed RLE normalized counts with 80 percent sample resampling.

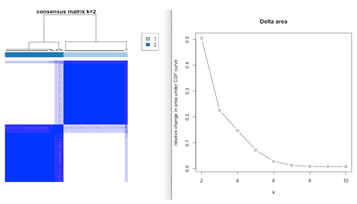


Figure 7: Consensus Matrix plot with k=2 and the relative change in area under the CDF for k from 1 to 10.

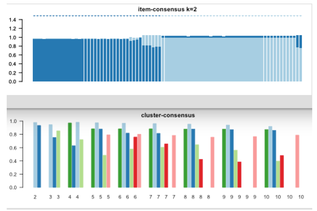


Figure 8: Item Consensus for k=2 and Cluster Consensus for k from 1 to 10.

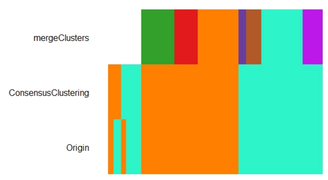


Figure 9: Comparison of Consensus Clustering and Combine Many method of clusterExperiment.

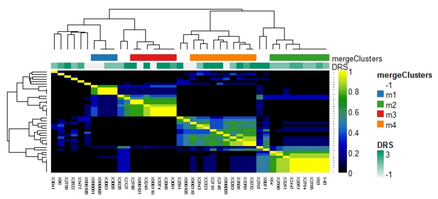


Figure 10: clusterExperiment heatmap within vesicle dataset with diagnosis as group.

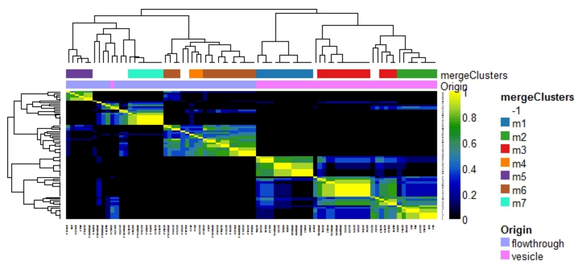


Figure 11: clusterExperiment heatmap with flowthrough vs. vesicle as groups.

EDA AND NORMALIZATION: METHODS & RESULTS

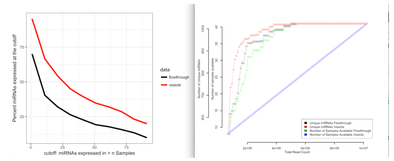


Figure 1: First graph depicts the percent of miRNAs expressed at a certain percent of samples in both flowthrough and vesicle datasets. Graph on the right shows the number of unique miRNAs and samples with at least 1 count for different total read sums.

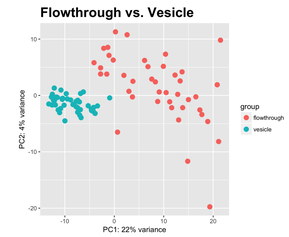


Figure 2: First two principal components depict differences between Flowthrough and Vesicle samples

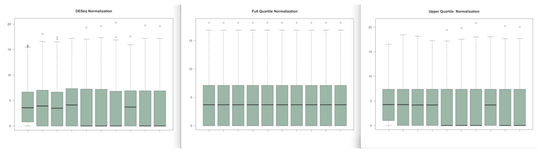


Figure 3: Boxplots for 10 randomly selected samples using 3 different normalization methods: DESeq, Full Quartile and Upper Quartile.

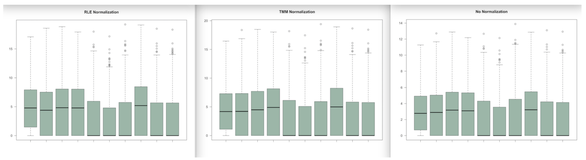


Figure 4: Boxplots for 10 randomly selected samples using 2 different normalization methods: RLE and TMM, as well as no normalization.

DIFFERENTIAL EXPRESSION: METHODS & RESULTS

Figure 5: estim. fold change

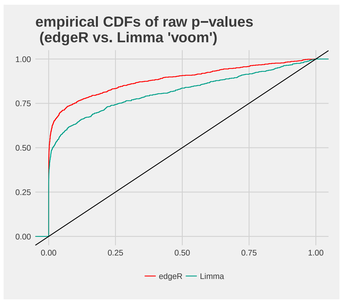
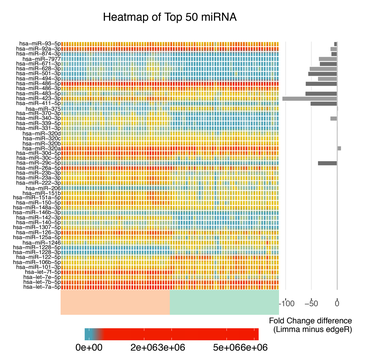


Figure 6: raw p-values



- There is some obvious discord between the set of differentially expressed genes.
- "Limma voom" controls the false positive rate more strictly than does "edgeR".

PRINCIPAL REFERENCES

[1] Daniel Enderle, Alexandra Spiel, Christine M Coticchia, Emily Berghoff, Romy Mueller, Martin Schlumpberger, Markus Sprenger-Haussels, Jonathan M Shaffer, Eric Lader, Johan Skog, et al. Characterization of RNA from exosomes and other extracellular vesicles isolated by a novel spin column-based method. *PloS one*, 10(8), 2015.

[2] Kevin P. McCormick, Matthew R. Willmann, and Blake C. Meyers. Experimental design, preprocessing, normalization and differential expression analysis of small rna sequencing experiments. *Silence*, 2(1):2, 2011.

ACKNOWLEDGEMENTS

We thank TGEN for access to their data as well as the direction on this project. We would also like to thank Sandrine Dudoit and Kelly Street for their consultation and advice.