Synopsis

Gencode has a shitload of cell-lines and compartments. The compartmentalization si the most interesting one.

Main points

1) Sensitivity (we only get high-rpkm genes) diagram 2) Coverage (Venn diagram) 3) Differences in high RPKM genes 4) Cytoplasm/nucleus differences in polyadenylation (compared to whole cell?)

TODO 1) the rpkm-poly(A)tail connection and 3)

We found signals of polyadenylation events in the GENCODE cell lines by trimming and remapping those reads that were originally unmappable and ended in a stretch of As or beginning with a stretch of Ts. We subsequently merged the polyadenylation sites into clusters in a similar way to Tian et. al [3]. Figure X shows the relationship between the 3UTR rpkm of transcripts and how many poly(A) sites are found in the 3UTR. As can be seen, we identify poly(A) sites for highly expressed transcripts, and the number of poly(A) sites in these genes average to XX, similar to what has been previosuly reported XXX. Of a total XYX clusters in he 3UTRs, X land at annotated poly(A) sites, while XUU may be novel ones. XYX pcnt of the clusters have one of the 8 polyadenylation signals within 40 nt downstream (XY for an, and 42 for non-an), also in line with what has been reported previously, showing that the method is highly specific for detecting polyA sites.

We wanted to compare poly(A) sites in the nucleus and cytoplasm. Poly(A) sites in the nucleus are expected to originate from stable mRNA that may exist in multiple copies. The poly(A) sites identified in the nucleus are expected to stem from mRNA diffusing toward the cytoplasm as well as mRNA undergoing processing, and some might possibly be marks of RNA undergoing degradation [1, 2]

References

[1] Natalia Shcherbik, Minshi Wang, Yevgeniya R Lapik, Leena Srivastava, and Dimitri G Pestov. Polyadenylation and degradation of incomplete RNA polymerase i transcripts in mammalian cells. *EMBO Rep*, 11(2):106–111, February 2010.

[2] Shimyn Slomovic, Ella Fremder, Raymond H. G. Staals, Ger J. M. Pruijn, and Gadi Schuster. Addition of poly(A) and poly(A)-rich tails during RNA degradation in the cytoplasm of human cells. *Proceedings of the National Academy of Sciences*, 107(16):7407 –7412, April 2010.

[3] Bin Tian, Jun Hu, Haibo Zhang, and Carol S. Lutz. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Research*, 33(1):201 –212, January 2005.