Replicates, and between inputs that vary by as much as 100-fold.

Intergenic, intronic, and other low-confidence regions compose only 0.1% of transcript length (5’ to 3’). Gene coverage is uniform across the entire transcript.

Figure 2: CONSISTENT PERFORMANCE ACROSS A RANGE OF INPUTS

Figure 3: AnyDeplete is a method for targeted depletion of virtually any unwanted transcript from an mRNA-Seq transcriptome. As a demonstration, the top 4 most abundant transcripts from the ERCC (External RNA Controls Consortium) RNA were targeted for depletion. A. log (FPKM) correlation between ERCC samples with and without AnyDeplete. Depletion of the 4 most abundant ERCC transcripts (H3B, 16, 2 and 74, in red) increases the number of reads of lower expression (gray). B. For each of the ERCC transcripts, the fold change in FPKM was calculated by dividing the FPKM of the With AnyDeplete library by the FPKM of the Without AnyDeplete library. Transcripts were sorted by reads mapping to each transcript in the Without AnyDeplete library, from low to high. Each symbol represents one transcript, displayed in rank order along the X-axis. Y-axis is the fold change in FPKM achieved by targeted removal of the top 4 most abundant transcripts (green). Fold changes of the remaining transcripts are generally consistent, although noise increases as read count decreases. Transcripts with less than 100 reads (blue) show more variability compared to transcripts with more than 100 reads (red).

Figure 4: TRANSCRIPT CONCORDANCE WITH AND WITHOUT ANYDEPLETE

Figure 5: CASE STUDY 1: GLOBIN READ REDUCTION OF mRNA-SEQ LIBRARIES

Figure 6: Figure 5: Glibin transcripts typically represent 30-80% of reads from blood mRNA-Seq studies, requiring deeper sequencing or global reduction to obtain the necessary data. In the case of whole blood from newborns, both adult and fetal globin transcript reduction was necessary. Bar plot illustrates successful depletion of globin reads from sequencing libraries prepared from newborn human whole blood total RNA.

Figure 7: CASE STUDY 2: MITOCHONDRIAL READ REDUCTION OF mRNA-SEQ LIBRARIES

Figure 8: Table 2: Sequencing metrics for 4 cardiac samples with AnyDeplete (to mitochondrial and ribosomal transcripts) and without AnyDeplete. All datasets were normalized to the same number of uniquely mapping reads before calculating FPKM values for RefSeq gene annotations with Cufflinks. By reducing mitochondrial reads from above 10% to below 1%, an increase in the fraction of reads mapping to exons, and a concomitant increase in library complexity was achieved.

Figure 9: INCREASE IN INFORMATIVE EXON READS FOLLOWING MITO DEPLETION

Figure 10: Table 2: Sequencing metrics for 4 newborn whole blood samples with AnyDeplete, with AnyDeplete to adult globin, and with AnyDeplete to adult and fetal globin. For each of the 4 newborn whole blood samples, both adult and fetal globin transcription reduction was necessary. Bar plot illustrates successful depletion of globin reads from sequencing libraries prepared from newborn human whole blood total RNA.

Figure 11: MORE USABLE DATA AFTER GLOBIN ANYDEPLETE OF HUMAN BLOOD SAMPLES

Figure 12: Table 1: Sequencing metrics for 4 newborn whole blood samples with AnyDeplete (to adult globin, fetal globin, and ribosomal transcript) and without AnyDeplete. Data illustrates successful depletion of globin reads from sequencing libraries prepared from multiple variable quality samples. All datasets were normalized to the same number of uniquely mapping reads before calculating FPKM values for RefSeq gene annotations with Cufflinks.

Table 1: Sample | AnyDeplete | mRNA | Globin | Total Aligned | RefSeq Exon | Genes FPKM >1
--- | --- | --- | --- | --- | --- | ---
1 | No | 2.6% | 76.5% | 90.5% | 80.6% | 8326
2 | Yes | 1.1% | 10.4% | 98.2% | 78.6% | 9388
3 | No | 3.2% | 79.2% | 88.9% | 87.2% | 955
4 | Yes | 1.8% | 15.2% | 98.2% | 83.2% | 9217
5 | No | 1.7% | 72.2% | 92.6% | 83.4% | 8791
6 | Yes | 0.5% | 82.9% | 89.9% | 82.9% | 9655
7 | No | 2.4% | 67.7% | 95.0% | 86.4% | 10003
8 | Yes | 0.3% | 79.9% | 95.1% | 85.9% | 10527

Table 2: Sequencing metrics for 4 newborn whole blood samples with AnyDeplete (to adult globin, fetal globin, and ribosomal transcript) and without AnyDeplete. Data illustrates successful depletion of globin reads from sequencing libraries prepared from multiple variable quality samples. All datasets were normalized to the same number of uniquely mapping reads before calculating FPKM values for RefSeq gene annotations with Cufflinks.

CONCLUSIONS

- We demonstrate here a flexible, modular workflow which allows customized depletion of any poly-A transcripts from sequencing libraries. Processing libraries with AnyDeplete can reduce uninformative reads by ~95%.

- The data from libraries prepared using AnyDeplete retain strong correlation to non-depleted libraries.

- This targeted depletion method enables more efficient detection of low copy transcripts which may be biologically relevant to the experimental hypothesis.

ACKNOWLEDGEMENTS

Our special thanks to Drs. Ashesh Saraiya and Stephanie Huelga for additional analysis and graphics.