



**Additional file 2: Uniqueness of contigs.**

Distribution of transcriptome contigs by percent of 100-mers that are  $\geq 10$  mismatches away from the nearest other 100-mer in the entire transcriptome.