Supplementary Figure 1. Nucleotide preference at the start positions by multiple RNA-seq methods

Nucleotide preferences as a function of relative position within genes was examined for the dsDNALigSeq (A) ssRNALigSeq (B) and CircLigSeq (C) methods. The histograms show the relative abundance of each nucleotide in the alignment position normalized to the nucleotide abundance in all aligned cDNA. For each histogram, we also examined positions extending five nucleotides upstream of the beginning of the alignment (positions -5 through -1). To allow this extension, the template cDNA rather then the sequence tag was used for the nucleotide preference calculation.

Supplementary Figure 2. RNA-seq tags are distributed evenly in the aligned exon-junctions

The histograms depict the distribution of the RNA-seq tags on the aligned exon-exon junctions. An exemplary schematic for tags that cover 11 bases from one exon and 21 bases from the second exon is presented. The number of nucleotides from each junction is shown on the X-axis, with the percentage of the aligned sequences on the Y- axis.

Supplementary Figure 3. RNA-seq by the ssRNALigSeq method can reveal transcripts that reside in gene regions

The value of ssRNALigSeq methods for identifying transcripts that could be mistakenly identified as exons or as regions in existing genes. RNA-seq by ssRNALigSeq method (black bars) shows a non-annotated transcript at *B0025.2* gene region, while the

dsDNALigSeq RNA-seq method (non-directional, green bars) suggests an exon in the gene. Recent wormbase releases also annotate this gene as *B0025.5*. The arrows in the bars show the directionality of the sequence alignment. The orange track is the non-annotated genomic tags dataset, the blue track is the potential non-annotated exon junction, and the green track is the putative polyadenylated sites. This intronic transcript is present in polysome-sized compartment, as suggested by examination of the polysome read map (pink track).

Supplementary Figure 4. Detecting mis-annotated 3' and 5' transcript regions

A dataset of non-annotated genomic tags was constructed from RNA-seq samples described in Table 1 and contains RNA-seq tags that align to the genome and not to the annotated WS190 cDNA reference. This dataset was uploaded to the UCSC genome browser (orange track in A and C). The size of the bar in each position in the track correlates with the number of sequences that align to that position. The non-annotated genomic tags dataset suggest mis-annotated 5' region of *cnb-1* gene (A) and mis-annotated 3' region of the *flp-4* gene (C). Both *cnb-1* (B) and *flp-4* (D) non-annotated transcript regions were confirmed by PCR, RT-PCR and Sanger sequencing. The Sanger sequencing alignment of the PCR and RT-PCR is shown in (A) and (C). A polyadenylation site was also detected in *flp-4*'s 3' UTR indicated by the green bars in the polyA tags track. An SL1 site was detected in the *cnb-1* 5'-UTR indicated by the purple bars in the SL1 tags track. The blue track in (A and C) displays the potential non-annotated exon junctions.

Supplementary Figure 5. Annotating transcripts by RNA-seq

Examples of multi-modal RNA-seq display in transcriptome annotations. (A) the annotated pseudogene *H40L08.2* is apparently transcribed (non-annotated genomic tags, splice junctions, and polyadenylation sites) and confirmed by RT-PCR (B). *T07A5.5* transcript (C) contains non-annotated splice junctions that were confirmed by PCR, RT-PCR, and Sanger sequencing (D). Examination of the read map constructed from ssRNALigSeq RNA-seq on polysomes sedimentation (pink track in A and C) suggests that both transcripts are present in polysome fractions.

Supplementary Figure 6. RNA-seq alignment to *let-7* microRNA precursor

An example is shown in which RNA-seq data can help annotating microRNA precursors more accurately. Browser shot of the *let-7* microRNA loci (C05G5.6) shows the microRNA precursor as suggested from the non-annotated genomic tags dataset (orange track), splice leader sites (purple track), and putative polyadenylation site (green track).

Supplementary Figure 7. Lack of coverage at the extreme 5' region of genes by the dsDNALigSeq method

Transcript coverage on the 5' region was determined by mapping SRR006511 (black), SRR006519 (red) RNA-seq datasets (Hillier et al. 2009) to the well annotated start sites of the *myo-1*, *myo-2* and *unc-54* genes (Dibb et al. 1989; Okkema et al. 1993).Sequence tags were mapped using BLAT (Kent 2002). The plots depict transcript coverage from 30 bases before the annotated start sites to 70 bases after. The bar on the X-axis presents the 1-5 bases ambiguity and variation in natural start sites (Dibb et al. 1989; Okkema et al.

1993). L1 larval stage constructed by CircLigSeq method (green) is shown for comparison. The overall number of sequences that aligned to the assayed region are indicated in parenthesis.