Wang et al. Figure S1

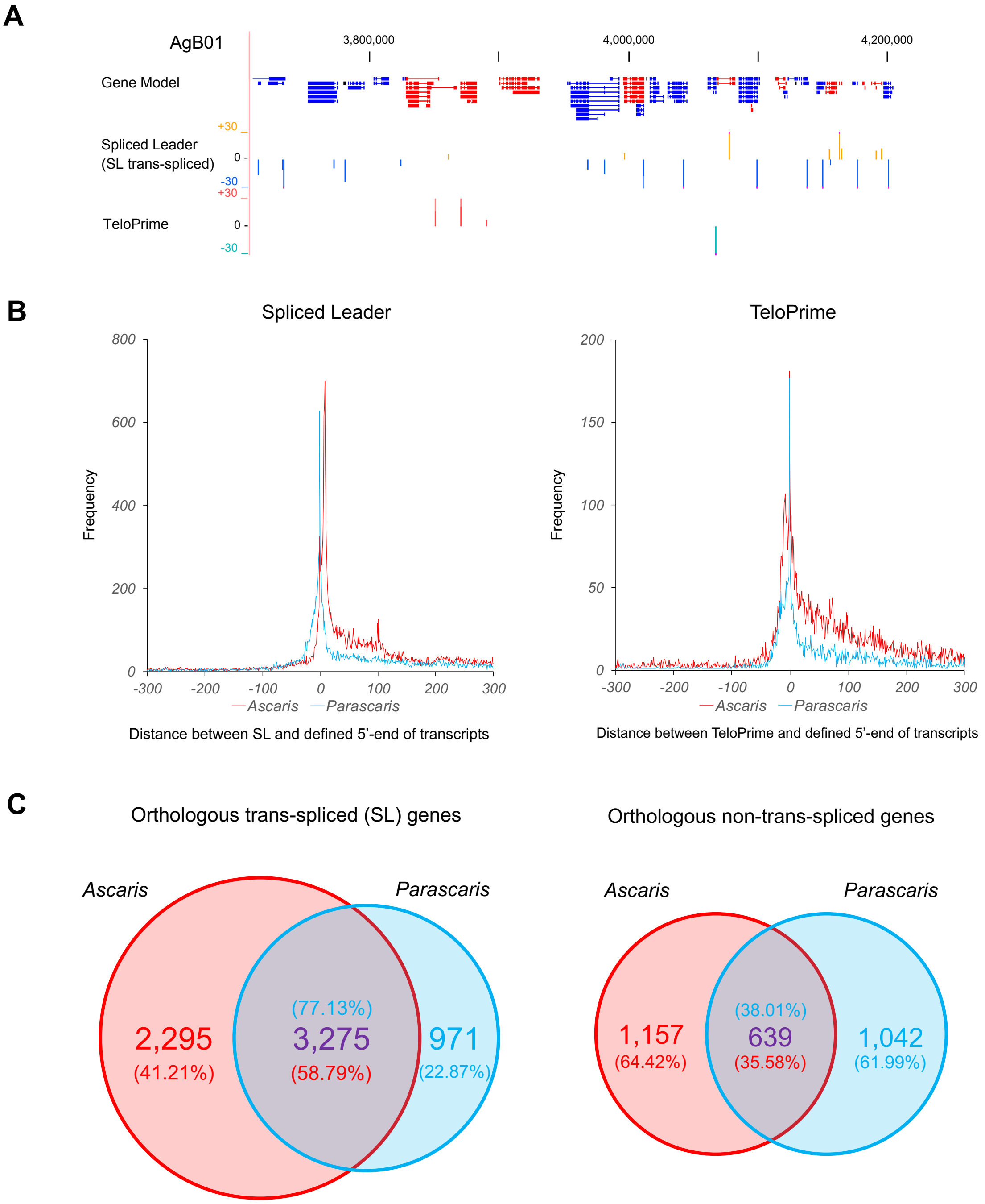

**Figure S1. Characterization of the 5' ends of *Ascaris* mRNAs. A.** Genome browser capture illustrating 5' mRNA reads for spliced leader (SL) and TeloPrime RNA libraries. Transcripts (gene model track) are colored as red for the plus strand and blue for the minus strand. The frequency (height) of the observed SL and TeloPrime sites is normalized (count per million mapped reads or CPM). We used a CPM cutoff >= 5 to define transcripts with an SL or TeloPrime priming sites (mRNA 5' end). For both the SL and TeloPrime tracks, red indicates reads mapped to plus strand and blue to minus strand. Note that the 5' end analysis was done on mixed RNA samples corresponding to all stages for which RNA-seq was carried out. The approximate ratio of SL to TeloPrime sites was approximately 4:1. **B.** SL and TeloPrime defined 5' ends of mRNAs map to and overlap with our assembled transcriptome transcripts. Illustrated is the distribution of the distance between 5' ends of the mRNAs (SL and TeloPrime sites) and the defined 5' end of the transcripts for both *Ascaris* and *Parascaris*. Note mRNA 5' ends are mostly within ~50 nts of the assembled transcripts. Upstream of 0 (negative number) represents the defined transcripts has yet to reach the 5'-end of the mRNAs while downstream of 0 (positive number) suggesting the 5'-end is within the defined transcripts. Note that TeloPrime can identify both non-trans-spliced mRNAs and trans-spliced mRNAs, but appears to have a preference for non-trans-spliced mRNAs. **C.** Conservation of trans-spliced and non-trans-spliced genes between *Ascaris* and *Parascaris.* Of 9,327 reciprocal-best-hit orthologous genes between *Ascaris* and *Parascaris*, we found 5,570 *Ascaris* and 4,246 *Parascaris* genes that are trans-spliced and 3,275 of them are shared. For the non-trans-spliced mRNAs, we found 1,796 genes in *Ascaris* and 1,681 genes in *Parascaris*, and 639 of them are shared.