**Supplemental Figure S1.** Venn diagram showing the overlap of evidence supporting the protein-coding capacities of the filtered gene models. The category ‘GLEAN’ means the RNA-seq model exons overlap with the GLEAN model exons. The category ‘Swprot’ means the RNA-seq model has significant matches (e-value < 1e-10) with protein sequences in the Swiss-Prot database. The category ‘ORF’ means the RNA-seq model has a long (≥500 bp) open reading frame and has more than 1 exon.

**Supplemental Figure S2.** Genome browser views of typical inconsistences between the RNA-seq models and the GLEAN models.

**Supplemental Table S1.** Types of inconsistences for 405 transcription factors derived from comparisons between the RNA-seq models and the GLEAN models.

**Supplemental Table S2.** The function class ontology compiled in this study covering 8,954 annotations of 7,000 GLEAN models. Each GLEAN model was assigned to a function class organized up to three hierarchical levels.