Figure S1

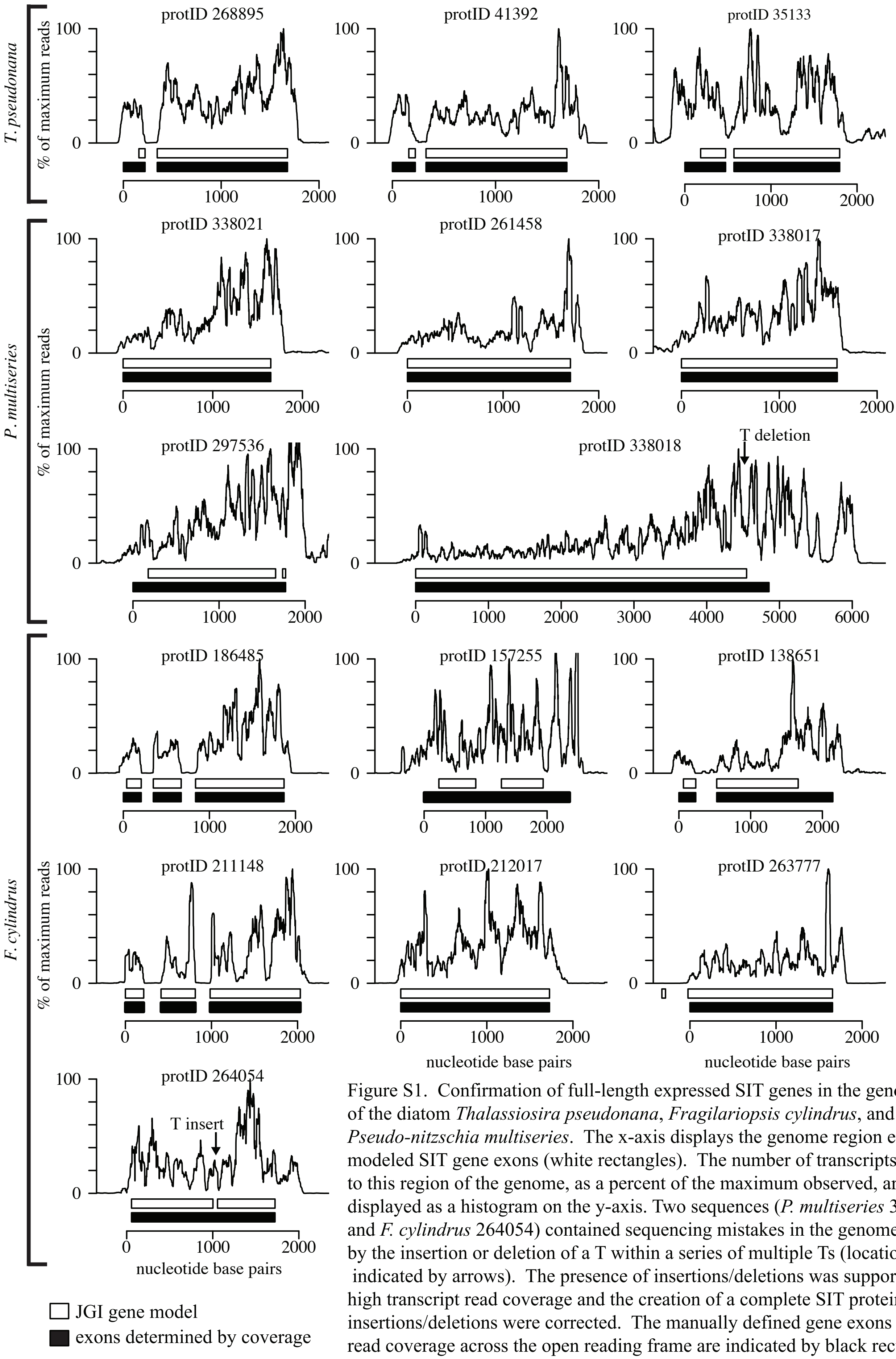

Figure S1. Confirmation of full-length expressed SIT genes in the genomes of the diatom *Thalassiosira pseudonana*, *Fragilariopsis cylindrus*, and *Pseudo-nitzschia multiseries*. The x-axis displays the genome region encoding modeled SIT gene exons (white rectangles). The number of transcripts that map to this region of the genome, as a percent of the maximum observed, are displayed as a histogram on the y-axis. Two sequences (*P. multiseries* 338018 and *F. cylindrus* 264054) contained sequencing mistakes in the genome, caused by the insertion or deletion of a T within a series of multiple Ts (locations indicated by arrows). The presence of insertions/deletions was supported by the high transcript read coverage and the creation of a complete SIT protein if insertions/deletions were corrected. The manually defined gene exons based on read coverage across the open reading frame are indicated by black rectangles.