



Figure S2. Phylogenetic tree of SIT proteins, as in Figure 2, displaying all individual branches and all sequences labeled by genus, species, and their sequence identifier (NCBI accession number, JGI protein ID, or MMETSP sequence ID). Branches are color coded by clade assignment (black = non-diatom outgroup and clade B, red = clade A, blue = clade C, orange = clade D, green = clade E).