**Table S17.** Presence of fused genes in essential amino acid biosynthesis in representative bacteria, cyanobacteria, green algae, diatoms, Nannochloropsis and higher plants

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Gene pairs | *E. coli* | Synechocystis*sp.* | *C. reinhardtii* | *P. tricornutum* | *T. pseudonana* | *N. oceanica* | Arabidopsis |
| *AK-HSDH* | + | - | + | + | + | + | + |
| *DHQDH-SDH* | - 1 | - | + | + | + | + | + |
| *ADT-ADH* | - | - | - | + | + | + | - |
| *ASA-ASB* | - 2 | - 3 | + | + | + | + | + |
| *IGPS-PAI* | + | - | - | + | + | + | - |
| *TSA-TSB* | - | - | - | + | + | + | - |

 1Although fused *DHQDH-SDH* was not found in *E. coli*, it is present in other bacteria, such as *Acidobacterium sp.* and *Planctomyces maris.*

 2Although fused *ASA-ASB* was not found in *E. coli*, it is present in other bacteria, such *Legionella longbeachae* and *Sorangium cellulosum*.

 3Although fused *ASA-ASB* was not found in Synechocystis *sp.*, it is present in other cyanobacteria, such as *Anabaena variabilis* and *Nostoc sp*.*.*

The accession numbers for AK-HSDH homologs in *E. coli*, *C. reinhardtii*, *P. tricornutum*, *T. pseudonana*, and Arabidopsis are ZP\_08372323, XP\_001695256, XP\_002182284, XP\_002296299, At1g31230, and At4g19710. The accession numbers for DHQDH-SDH homologs in *Acidobacterium sp.*, *P. maris*, *C. reinhardtii*, *P. tricornutum*, *T. pseudonana*, and Arabidopsis are ZP\_07032873, ZP\_01854559, XP\_001694346, XP\_002179655, XP\_002289031, and At3g06350. ADT-ADH homologs in *P. tricornutum* and *T. pseudonana* were obtained using tBLASTn at [www.jgi.doe.gov](http://www.jgi.doe.gov); the corresponding accession numbers are estExt\_fgenesh1\_pg.C\_chr\_130260 and estExt\_fgenesh1\_pg.C\_chr\_10420. The accession numbers for ASA-ASB homologs in *L. longbeachae*, *S. cellulosum*, *A. variabilis*, *Nostoc sp.* *C. reinhardtii*, *P. tricornutum*, *T. pseudonana*, and Arabidopsis are YP\_003455732, YP\_001615687, YP\_325382, NP\_484458, XP\_001702943, XP\_002177062, XP\_002287912, and At2g28880. The accession number for IGPS-PAI homolog in *E. coli* is ZP\_06938008. IGPS-PAI homologs in *P. tricornutum* and *T. pseudonana* were obtained using tBLASTn at [www.jgi.doe.gov](http://www.jgi.doe.gov); the corresponding accession numbers are estExt\_fgenesh1\_pg.C\_chr\_60121 (from DOE-JGI) and fgenesh1\_pg.C\_chr\_10000275. The accession numbers for TSA-TSB homologs in *P. tricornutum* and *T. pseudonana* are XP\_002176877 and XP\_002294706.