**Table S16**. Predicted genes in the biosynthetic pathways of Asp-derived, aromatic and branched-chain amino acids and in nitrogen assimilation in CCMP1779

|  |  |  |
| --- | --- | --- |
| **Description** | **Name** | **ID** |
| **Biosynthesis of Asp-derived amino acids** |  |  |
| Asp kinase | AK | CCMP1779\_3150 |
| Asp kinase - homoserine dehydrogenase | AK-HSDH1 | CCMP1779\_10446 |
| Homoserine dehydrogenase | HSDH | augustus\_masked-nanno\_5085-abinit-gene-0.2 5 |
| Asp semialdehyde dehydrogenase | ASD | CCMP1779\_11914 |
| Asp semialdehyde dehydrogenase | ASD | CCMP1779\_6330 |
| Dihydrodipicolinate synthase | DHDPS | CCMP1779\_11357 |
| Dihydrodipicolinate reductase | DHDPR2 | CCMP1779\_4610 |
| Diaminopimelate aminotransferase  | DAPAT | augustus\_masked-nanno\_2338-abinit-gene-0.3 5 |
| Diaminopimelate epimerase | DAPE2 | CCMP1779\_10579 |
| Diaminopimelate decarboxylase | DAPDC | CCMP1779\_2505 |
| Diaminopimelate decarboxylase | DAPDC | CCMP1779\_8817 |
| Homoserine kinase | HSK | CCMP1779\_8751 |
| Cystathionine γ synthase | CGS | CCMP1779\_10900 |
| Cystathionine γ synthase | CGS | CCMP1779\_4094 |
| Cystathionine β lyase | CBL | CCMP1779\_9018 |
| Cystathionine β lyase | CBL | CCMP1779\_9158 |
| Cystathionine β lyase | CBL | CCMP1779\_3085 |
| Homocysteine S-methyltransferase | HMT | CCMP1779\_1690 |
| Cobalamin-dependent Met synthase | MetH | CCMP1779\_2985 |
| Cobalamin-independent Met synthase | MS | CCMP1779\_5185  |
| Cobalamin-independent Met synthase | MS | CCMP1779\_8229  |
| Thr synthase | TS3 | CCMP1779\_10068 |
| Thr synthase | TS2 | CCMP1779\_4990 |
| **Biosynthesis of aromatic amino acids** |  |  |
| 3-Deoxy-D-arabino-heptulosonate-7-phosphate synthase  | DAHPS | CCMP1779\_7898 |
| Dehydroquinate synthase | DHQS | augustus\_masked-nanno\_505-abinit-gene-0.2-mRNA-1 5 |
| Dehydroquinate dehydratase - shikimate dehydrogenase | DHQDH-SDH1 | CCMP1779\_5703 |
| Shikimate kinase | SK | CCMP1779\_10591 |
| 5-Enolpyruvylshikimate-3-phosphate synthase | EPSPS | CCMP1779\_6038 |
| Chorismate synthase | CS3 | CCMP1779\_8941 |
| Chorismate synthase | CS3 | CCMP1779\_2899 |
| Chorismate mutase | CM | CCMP1779\_234 |
| Arogenate dehydratase - arogenate dehydrogenase | ADT-ADH1 | CCMP1779\_11837 |
| Anthranilate synthase α and β subunits | ASA-ASB  | CCMP1779\_4984 |
| Anthranilate phosphoribosyltransferase | PRT | CCMP1779\_4197 |
| Indole-3-glycerol phosphate synthase - phosphoribosylanthranilate isomerase | IGPS-PAI1 | CCMP1779\_5546 |
| Indole-3-glycerol phosphate synthase | IGPS4 | CCMP1779\_398 |
| Trp synthase α and β subunits | TSA-TSB1 | CCMP1779\_102 |
| **Biosynthesis of branched-chain amino acids**  |  |  |
| Thr deaminase | TD | CCMP1779\_7883 |
| Acetohydroxyacid synthase | AHAS | CCMP1779\_3427 |
| Ketolacid reductoisomerase | KARI | CCMP1779\_6007 |
| Dihydroxyacid dehydratase | DHAD | CCMP1779\_4633 |
| Dihydroxyacid dehydratase | DHAD3 | CCMP1779\_7539 |
| Isopropylmalate synthase | IPMS3 | CCMP1779\_4567 |
| Isopropylmalate isomerase large subunit | IPMIL | CCMP1779\_2773 |
| Isopropylmalate isomerase small subunit | IPMIS3 | CCMP1779\_5262 |
| Isopropylmalate dehydrogenase | IMD | CCMP1779\_7869 |
| Branched-chain aminotransferase | BCAT | CCMP1779\_8799 |
| Branched-chain aminotransferase | BCAT3 | CCMP1779\_5665 |
| Branched-chain aminotransferase | BCAT2 | CCMP1779\_10881 |
| **Nitrogen Assimilation** |  |  |
| High affinity nitrate transporter | NRT2:1 | CCMP1779\_438 |
| Nitrite transporter | NAR1 | CCMP1779\_6536 |
| Ammonium transporter | AMT1:1 | CCMP1779\_5503 |
| NAD(P)H:Nitrite reductase | NIR | CCMP1779\_6877 |
| Ferredoxin nitrite reductase | NII | CCMP1779\_9262 |
| NAD(P)H-Nitrate reductase | NIA | CCMP1779\_443 |

Genes in the biosynthetic pathways of Asp-derived, aromatic and branched-chain amino acids were manually annotated. For each *N. oceanica* locus, the gene function was predicted based on BLAST similarity to homologous genes in Arabidopsis. The accuracy of gene function prediction was evaluated by PFAM domain support and reverse blast searches of Arabidopsis homologs with *N. oceanica* genes.

 1These genes are predicted to be fused with other genes in the same pathway in *N. oceanica*.

 2These predicted genes are missing the start codon.

 3These predicted genes are missing the stop codon.

 4This IPGS gene appears to be fused with some unrelated gene(s).

 5this gene model is from augustus or snap gene annotation and was found superior to the final maker annotation after manual examination.