## Supplementary Information

## Bioinformatic analysis

alr2269N TEPRVLVSEVLVRPQSGQLTPELETQVYNVIRTQPGRTTTRSQLQEDINAIFGTGFFSNV ..... 60
2QDZ ..... 0
ptToc75 ..... 0
alr2269N QASPEDTPLGVRVSFIVQPNPVLSKVEIQANPGTNVPSVLP---QATADEIFRAQYGKIL ..... 117
2QDZ ..... 02QDZ
34
alr2269N NLRDLQEGIKELTKRYQDQGYVLANVVGAPQVSENGVVTLQVAEGVVENISVRFRNKEGQ ..... 177
2QDZ ..... 0
ptToc75 QSTSIQSFAKSLKEWYIRKGYILHSVTGATLQSETGTAIITVQEPAVSKDPVGITFCKEM ..... 94
alr2269N DVNEQGQPIRGRTQDYIITREVELKPGQVFNRNT-V-QKDLQRVFGTGLFE-D-V-NVSL ..... 232
2QDZ ..... 0
ptToc75 VVDTTFVPTAGHTRPNRIASALGLSPGNPFQ---W-DGIRWQQIASSGIFAQ-V-LRATP ..... 148
alr2269N DPGTDPTKVNVVVNVVERS ..... 251
2ODZ --------------------QLLPGARDLNRIDDRQRKEQLQAAAAAAATSGHTVTVHAVD ..... 41
ptToc75 QP-MQDGTVQLHILATEA ..... 165
alr2269N ..... 251
2QDZ LDFGVEGRLFDPAPLVQDYLNRPLDNEQLFLLVKALSAALYDRGYATSIVTFVPPGVVDG ..... 101
ptToc75 ..... 165
alr2269N ..... 251
2QDZ VLKLKVEWGRIKGWLIDGKPLEGTRDRMMVFSAMPGWQDKVLNVFDIDQAIYNINNGGKT ..... 161
ptToc75 ..... 165
alr2269N ..... 251
$2 Q D Z$ GNITIVPADEYGYSYLDLQLQRRALPRVSLGMDNSGPGTPENG----RYKYNASVTANDL ..... 217
ptToc75 -----------------------PTRHLEYGLGKS----------LYTGSWEGELDFEHINL ..... 194
alr2269N ..... 251
2QDZ LGLNDTLGLYIGNRYY--RDAGHD-AERNYDLMYSVPL--GR--TRLDLQTGYSTYRNLL ..... 270
ptToc75 LGGGESLGLSVRRG--TK------DAEPSIRLSFTDSKFG--LGGGYDLEVFSDFIGDQP ..... 244
alr2269N ..... 251
$2 Q D Z$ KTRYGQ-------YQSAGNSRSFGLKATRLLYRDTRSQFSVYGGLKLRQNKNYLAGTRLD ..... 323
ptToc75 EEMQEGELPPDYDHDSILDRKGATVRLRNPICSRIVRNSVASGSLERTATKTGLH---ET ..... 301
alr2269N ..... 251
2QDZ VSSKHYSDVTVGMQYSTQRGANAYFGDLSFTRGVGV ..... 366
ptToc75 IGSTTLAVGPFVKRLPYDARS-SIDAKCSLGTRLFAASSTTEEANALLPPLKLKPYTSAT ..... 360
alr2269N ..... 251
2 QDZ GSLAWTRYMALA--GQPIQWASQLGFQYSRQQ------LLNSYQITAGAAAAAAGAAAAA ..... 418
ptToc75 ATTRQIFPLLASFGPGGRPLILALKHSI----TTATPNLPRHEAKAQGIAHNIRGSSLNG ..... 416
alr2269N ..... 251
$2 Q D Z$ AASGDSGVYLSNTLTVPVQFSLLGKQASVAPFVGADVGALKSNHPDARTIRMAGLAAGVR ..... 478
ptToc75 RVA--SALTGTTELRVPTDTPVVQTRODTNVVFFGDWLFAAKDS-QSSFRRKSCVGVGVR ..... 473
alr2269N ..... ---------------------------------------------------- ..... 251
2QDZ FDLPYARMSFTYSK------PVGAQPGGAPRAPVWLYINAGLSF 516ptToc75 KSVQGIPLKYDLCYSSDDGK-----------FQSAFGLGADFDF 506

## Supplemental Figure 1. Structural analysis of ptOmp85

Multiple alignment of the N-terminal part of alr2269, the crystal structure of FhaC (PDB:2QDZ) and Toc75 from Phaeodactylum tricornutum constructed with MAFFT v6.624b (38), manually adjusted and fed to Modeller v8.2 for constructing the homology model.

Nnen/1-797
10
2032

Fnuc/1-678 1 MKR-3

Anal/1-833
1 MVNLPIKSVEVVNMQOVPASLIKNTLKLKEGAKFSTEALLADFNALKETGYFEDVILOPVSYDGGVRIVVDVVEKENVVDLLKEKG- ..... 86

AthaV/1-732 AthaIII/1-818 Atham/1-524 Xlae/1-468 Ptri/1-690 1 MAAVAITAPLSSSLTANAQTPTNSEQTLEAIAPATMQQPEQDVRSESPEDFTIVKALADMQSPSVEF .....-STDKSTKSAATTKKDVIVPTLETLVATIMP96 1 ….............................................. 73

1 MAAFSVIGQLIPTATSSTASTSLSSR
26
1 MENPAEKPD ..... $-9$
1 MGTVHARS ..... 81 MKTSSLV7Cmer/1-8401 MCVLGFLAGDPLRVGAHALKICTRSIASRERRRNAPRRQFPADRTGQRWHRSTTWLSKPGEPLTCSWQSQMLD-73
Tpse/1-8041 MVS-3
Ehux/1-636
Tgon/1-14251 MITFQTTRRKYPLLGGLACCLCLRAFVHISGD-- 32
1 MLVLLLST17
Nnen/1-797

Ecol/1-810
Clar/1-738
Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Anal/1-833
AthaV/1-732
AthaIII/1-818
Athar //1-524
Xlae/1-468
Ptri/1-690
Cmer/1-840
Cmer_Sam50/1-541
Tpse/1-804
Ehux/1-636
Tgon/1-1425

Mnen/1-797
Ecol/1-810
Clar/1-738
Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Ana1/1-833
AthaV/1-732 AthaIII/1-818 AthaM/1-524 Xlae/1-468 Ptri/1-690
Cmer/1-840
Cmer_Sam50/1-541
Tpse/1-804
Ehux/1-636
Tgan/1-1425

Mnen/1-797
Ecol/1-810
Clar/1-738
Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Anal/1-833
AthaV/1-732 AthaIII/1-818

AthaM/1-524
Xlae/1-468
Ptri/1-690
Cmer/1-840

8

PLLLLALIVPECQAFLPPSNSPFGRIGGLFSOPFTDGRNSGDIGVPPSDDI 58

33 ............................................................................................
18 -LKLFLCLLQFLFCVSTPTLLLSLAALYLVLPLFGDSPNAGFLGFVFADASRLPASGVFSSFPGLYYAQRQEASH-


PRPSHASHAASLSTVSSSLSSVEPSDGCKRL STAFSIRPLPDSSDARSSSL SHSSPPL 14

| 310 | 320 | 330 | 349 | 350 | 360 | 37 | 389 | 390 | 400 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  |  |  |  |  |
| 60 SNTIRALFA |  |  |  |  |  |  |  | IASITFS |  |
| 3 DTAIKNLF |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 78 DRSLKNLFETGLFADVSMRHE INRIAFE -NRRIKD125 <br> 84 PRTIEYLWKQRFFSDIQANONISGT -- -TVNLOL $\qquad$ IVKEL IVKELPL $\square$ LDRVVFD---G $\qquad$ -- NDIFDD135 |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 136 EDAQRRLLA |  |  |  |  |  |  |  | VKSVIIT |  |
|  |  |  |  |  |  |  |  |  |  |
|  210 QKELETLATCGMFEKVDLEGKTKPDGTLGVTISFAESTWQSADRFRCIMGLMVQSKPI日MDSDMTDKEKLEYYRSLEKDYKRRIDRA $-\ldots$ R-.-PCLLPA 394 |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| 101 DRTLFGGGGFGGISYVDGVAGGFPDLPNPGSFLEWLTLCWHTVQWLRAPTALALSVPTAATEEEAAEIYAEHPNDOTKGPLROWVLR---GPVDRLLLPR198 |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
| SSAPVAFAGPRRADPLSVEQIGSGDVDTPEAAL-- - - - - - - - - - - - - - - - |  |  |  |  |  |  |  |  |  |

Nmen/1-797
Ecol/1-810
Clar/1-738 Ddes/1-893 Rrub/1-778 Cpha/1-821
Tthe/1-822
Fnuc/1-678
Anal/1-833
AthaV/1-732 AthaIII/1-818 Atham/1-524
Xlae/1-468
Ptri/1-690
Cmer/1-840
Cmer_Sam50/1-541
Tpse/1-804
Ehux/1-636
Tgon/1-1425

Nnen/1-797
Ecol/1-810
Clar/1-738
Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Anal/1-833
AthaV/1-732
AthaIII/1-818
Atham/1-524
Xlae/1-468
Ptri/1-690
Cmer/1-840
Cmer_Sam50/1-541
Tpse/1-804
Ehux/1-636
Tgon/1-1425
7 DAIKMLESFGLAQSQYFMQATLMQAVAGLKEEYLGFGKLNIQITPKVTKLA-RURVDIDITIDEGKSAKITDIEFEGMQVYSDRKLMRQMS ..... TEGSIW 204
108 DMLKMLEASGVRVGESLDRTTIADIEKGLEDFYYSVGKYSASVKAVVTPLP-RNRVDLKLVFQEGVSAEIOQINIVGMHAFTTDELISHFQ--LRDEVPN 205101 KQIE-- SLVGLKPGILYDENSAKEAAEKIKLFYQAKGFYDTVVEIKDEKLSNSSSLKLTFWVNFGENIIIEKVHLSGAKNLSYYSDIEPAVANIQREALGN198251 DEIL-- SMMSTKTGSMMEKLLAQDIMKVTDLYRKKGYYLADVSHRVVEGS-GGAASLVLSVQEGDKLYIKAVRIVGAEQLDEDDVLDELA - -LRPRNMI 345
126 PQLE---TEIQLRPRVVYTRAKVQEDVQRILELYRRSGRYAVSVEPKVIALD-QNRVDLAFEINEGPSTTVRRISFVGNKVYDDSALREVVE--TSEEAWY 220
13 KELN-- NIASLYAGRHVDEQSLLTAASKIRTAYTKKYLRAEVEYTLKLQ-NIRVSARFSIDEKSKVVIENITFHGNTAFKDGKLFGVFK-ETSQNAWN 23199 RLLTFLEQNFAIGKGATYMPVRAQEAAEALARAYRQAGFPFTPKVAVEAKED-KEGIALTFRVEEGPE--VKAVRLLGVSLLPEEELRKGLE-18
188 STIMSELTT ..... 196
319 ATADEIFRA ..... 327
257 KFIHEAFRD ..... 265
314
305 PVYGEMMML
29
26 DLVE ..... 182
199 DRLEQAVNRILQAMV ..... 213
199 EVITQSSKQANL ..... 210
50 PVLAQAVRESGV ..... 61
207 RLVPQLEKELLSKHVSPASTSSASSTSSSMIMPRVSRRSASPSPGSPAASSKSS ..... 259
510 520 530 540 550 ..... 570
580 590 ..... 600205 TWILTRSMOFNEQKFAQDMEKVTDFYQNNGYF206 WINVGDRKYQKOKLAGDLETLRSYYLDFGYA199 MUGFNDGKLKIFDLANDSSRISDVYLKEGYL346 SWITGTGVLREEYLERDSAAIGAYYLNPGPM221 RFLSSTDTYDPDRLLYDRELLRRHYLKHGYA232 KTIFGQPKLDVEKFNADKKLLVEFYRENGYR-188 ․-ALKGSFDFAKYQEALRAIAKRYEEAGYRFSGPDPEASTLEEGVLTVRVRELKVARVEGEGLDLEGFPLKPGDYLNYAKLLEGVQDLSRKLSRVVNFTL 285

Mnen/1-797
Ecol/1-810 Clar/1-738 Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Anal/1-833
AthaV/1-732 AthaIII/1-818 AthaM/1-524 Xlae/1-468 Ptri/1-690
Cmer/1-840 Cmer_Sam50/1-541 Tpse/1-804
Ehux/1-636
Tgon/1-1425
$\begin{array}{llllllllll}610 & 620 & 630 & 640 & 650 & 660 & 670 & 689 & 690 & 700\end{array}$
 246 SLTPDKKGIVVTVNITEGDOYKLSGVEVSGNLAGHSAE--IEQLTKI -239 NTYTDTYQADLTYFINEGEVYKVKGIKIFNPIFSDEENEALANDLKL385 QYEED--GIVVSFAVKEGTRYKVGMIAFGGDLIDTDEA--LLININMDEQA 251 ELSTDRQAFPNTFTVSEGERYRFGPVGLEITLAELDPE-ALFGVITF272 SYTNDLKGVIIDLYIDEGAVYYVRNVTWMGNSKNFATTDILHSVFSI 295 LPOGDEVVVRLELGPEGG-- RIERVEVSGITA -- LPTEALLGLLRL197 $\qquad$
$\qquad$
EPGEL YNGTKVTMMED-SVGKIVNIEKLREDIKQAADEEY FIYDVVLOKDTKKDGDNY SADKVEAAVQ-.....-KLTDEVGNFGFAFVDVVKPRV KKGDRYNSKKIDEKLNFSKDNRDVASLYLDRGYLSFRSSLDE-
 kPGEVMTPALAGEDAAKPGSVQIYYNILREDRD QYGKILILRDLOEGIK-GFGKIINIKRLEEAIT-RDOGKVSARLLQRIRD-EEEEYEEEDD-- VEPEKMEILE-VGHPLOSTSIDSFAKIGGTLTPETMERTAS - LGRPASDPAVGALSD-ALSAEENEALIRTFVD-

KILGL MAOGYTLVIITIDMS ELTKFY QDOGYVLANVVGAP-SINGNMEFG-LFGIVSDIRYQKM HDEGYACAOVVNFG.
SLKENYIRKGYILHSVTGAT- SLIMAFYGEHGYPLALVRRAP.
LIMQWWIMGGYMNSVTGATRL EHV YRRSGY LFARVGKRV--IVNAWYLENGYLFAHLVPRPHFLPSSVP 393

Mnen/1-797
Ecol/1-810
Clar/1-738
Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Anal/1-833
Athav/1-732 AthaIII/1-818
AthaM/1-524
Xlae/1-468
Ptri/1-690 Cmer/1-840 Cmer_Sam50/1-541 Tpse/1-804 Ehux/1-636 Tgon/1-1425


Nmen/1-797 Ecol/1-810 Clar/1-738 Ddes/1-893 Rrub/1-778 Cpha/1-821 Tthe/1-822 Fnuc/1-678 Anal/1-833 AthaV/1-732 AthaIII/1-818 AthaM/1-524 Xlae/1-468 Ptri/1-690 Cmer/1-840 Cmer_Sam50/1-541 Tpse/1-804 Ehux/1-636 Tgan/1-1425

Mnen/1-797 Ecol/1-810 Clar/1-738 Ddes/1-893 Rrub/1-778 Cpha/1-821 Tthe/1-822 Fnuc/1-678 Anal/1-833 AthaV/1-732 AthaIII/1-818 Athar //1-524 Xlae/1-468 Ptri/1-690 Cmer/1-840 Cmer_Sam50/1-541 Tpse/1-804 Ehux/1-636 Tgon/1-1425

592 VSGVSAPLPATPPQTPARLLVPAGHFRVGFAASGERRLGAAGTMMGEATIERRIGLYVNLGLAGSASRQGSDSSRLPSSLAEASDTSHVGEKRNEDAPHRK692

1420
1439
1440
$1450 \quad 1460$
1470
1499
1490
1500
1510

Mnen/1-797
Ecol/1-810 Clar/1-738 Ddes/1-893 Rrub/1-778 Cpha/1-821 Tthe/1-822 Fnuc/1-678 Ana1/1-833 AthaV/1-732 AthaIII/1-818 AthaM/1-524 Xlae/1-468 Ptri/1-690 Cmer/1-840 Cmer_Sam50/1-541 Tpse/1-804
Ehux/1-636
Tgon/1-1425

|  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 413 LLSAVAPRAAGLLVGGQLS-................-AEHVPASGGGGGGGGSGGGEGPGGGPGDELPVLLEARATA--GALRPPEAVAEAELRQR-............. 488 792 YTPYTAADAAAAFAPLTSSGSASTRPGSRPSRVSRDRPLPPLPGLAGFSPPRPHVGTARSFSLSSELVPSRVSAFAAKLVPRPLRDESRFSQRCREVKETVS 892 |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 1520 | 1530 | 1540 | 1550 | 1560 | 1570 |  |  | , |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 560 ..................................................................................................................................... 577 |  |  |  |  |  |  |  |  |  |  |
| 554 … ................................................................................................................................. 568 |  |  |  |  |  |  |  |  |  |  |
| 414 ---DSILFHEIDTI <br> GFFTNIGKGLGKNFTLS- $-441$ |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 550 PGVEEVPPIWVDRA GVKANITENFTRQSKFT-- 589 |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 503 <br>  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |

Mnen/1-797
Ecol/1-810
Clar/1-738
Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Anal/1-833
AthaV/1-732
AthaIII/1-818
AthaM/1-524
Xlae/1-468
Ptri/1-690
Cmer/1-840 Cmer_Sam50/1-541

Tpse/1-804
Ehux/1-636
Tgan/1-1425

Mnen/1-797
Ecol/1-810
Clar/1-738 Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Anal/1-833
AthaV/1-732
AthallI/1-818
AthaM/1-524
Xlae/1-468
Ptri/1-690 Cmer/1-840 Cmer_Sam50/1-541

Tpse/1-804
Ehux/1-636
Tgon/1-1425

Mnen/1-797
Ecol/1-810
Clar/1-738
Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Anal/1-833
AthaV/1-732
AthaIII/1-818 AthaM/1-524
Xlae/1-468
Ptri/1-690
Cmer/1-840
Cmer_Sam50/1-54
Tpse/1-804
Ehux/1-636
Tgon/1-1425

Nnen/1-797
Ecol/1-810
Clar/1-738
Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Ana1/1-833
AthaV/1-732
AthaIII/1-818
AthaM/1-524
Xlae/1-468
Ptri/1-690
Cmer/1-840
Tpse/1-804
Ehux/1-636
Tgan/1-1425

Nnen/1-797
Ecol/1-810
Clar/1-738
Ddes/1-893
Rrub/1-778
Cpha/1-821
Tthe/1-822
Fnuc/1-678
Anal/1-833
AthaV/1-732 AthaIII/1-818

AthaM/1-524
Xlae/1-468
Ptri/1-690
Cmer/1-840 Cmer_Sam50/1-541

Tpse/1-804
Ehux/1-636
Tgon/1-1425
620
1630
1640
1650
1660
1670
1689
1690


$1820 \quad 1830$
$1850 \quad 1870$
1890
1890
1900
1910

$\begin{array}{llllllllll}1920 & 1930 & 1940 & 1950 & 1960 & 1970 & 1990 & 1990 & 2000 & 2010\end{array}$

-LTVNTYNKAPKHYADFIKK
-LTVNTYNKAPKHYADFIKK
5554

29 $\frac{29}{76}$ 555 95
$\stackrel{1890}{1810}$



|  | 2030 | 2040 | 2050 | 2060 | 2070 | 2089 | 2090 | 2100 | 2110 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Nnen/1-797 | 715-...- -GSWDGK |  |  | DNSSSA | RVQNIYG | THKSTFT |  |  |  | 757 |
| Ecol/1-810 | 742 ------GTWDTI |  |  |  |  | PDVSD |  |  |  | SA 779 |
| Clar/1-738 | 683 -....- GAIGESS |  |  |  |  |  |  |  |  | 51698 |
| Ddes/1-893 | 833 -.....- GFVIDPL |  |  |  | - | --HENS |  |  |  | VV852 |
| Rrub/1-778 | 715-...-- - - - |  |  |  |  | EALMDGA |  |  |  | SA 738 |
| Cpha/1-821 | 756 --- ----GNLWGST |  |  |  |  | -IM |  |  |  | 54774 |
| Tthe/1-822 |  | GGVW |  |  |  |  |  |  |  | GA 781 |
| Fnuc/1-678 | 611 --....-GRAWMQI | IGRDPSYTRDI | HIIG |  |  |  |  |  |  | TA638 |
| Ana1/1-833 | 776 -....--GSDLGT |  |  |  | SGSG |  |  |  |  | G) 798 |
| AthaV/1-732 | 671 --...-- - - 1 DMGSCS |  |  | VPGDPAC | KPGSGY |  |  |  |  | GY697 |
| AthalII/1-818 | $760 \text {-......-GNDLGSS }$ |  |  | VkGIPTA | RTGOGS |  |  |  |  | 5Y 786 |
| Atham/1-524 | 457 --...--GMMELS |  |  |  |  | FTAPKL |  |  |  | 5V483 |
| Xlae/1-468 | 399 -....--GNLCNLI |  |  |  |  | RAHLQRL |  |  |  | 5Y425 |
| Ptri/1-690 | 636 --...--LFAAKDS | SQSSFRRKS |  |  |  |  |  |  |  | Q652 |
| Cmer/1-840 | 729 -- --- - -GSEIQPP | SHLVRFPKL | HGHR | DSASAES | gVAPAS | SIQRSSTA |  |  |  | 5 V 91 |
| Cmer_Sam50/1-541 | 479 -.....- GDCREWS |  |  |  |  | --VAW |  |  |  | SA498 |
| Tpse/1-804 | 742 -- .-. - SI SQAQT | TQSSSHQDOYD |  |  |  |  |  |  |  | S1767 |
| Ehux/1-636 | $617 \text { - }$ |  |  |  |  |  |  |  |  | $\text { AS } 618$ |
| Tgon/1-1425 | 1285 VSPFLLPGIESSSP | SPAAAAASAL | AIA | 5 S | GS | RFL | RFG |  | TG | V1385 |
|  | 2130 | $2140$ | $2150$ | $\underset{2160}{ }$ |  |  |  |  |  |  |
| Nnen/1-797 | 758GGAVTWL --.-5P. | -LGPMIFFPYAY | PLKKKPP |  | $\therefore \text { QFO }$ | T-TF-- |  |  |  | 797 |
| Ecol/1-810 | 771 GLALOWM - - - SP- | LGPLVFSYAO | PFKKYDC | AE--QF | --QFI | K-TW-- |  |  |  | 810 |
| Clar/1-738 | 699 GVGFENL--. | -LGALNLVFAK | PFITNS | S--1F | - EF | A-RF-- |  |  |  | 738 |
| Ddes/1-893 | 853 GLEVFWQ--- - | LGILRFAYGY | PLNEVCC | KDG- KF | - EFT | Q-TF-- |  |  |  | 893 |
| Rrub/1-778 | $739 \text { GAGVTWV-.. SP- }$ | VGPISVDLGY | PLVKEKF | TE-VF | $\cdots \text { RLS }$ | -RF-- |  |  |  | 778 |
| Cpha/1-821 | $775 \mathrm{GLGLRLY}-\ldots \text { LPI }$ | IIGLIGLDYGYG | DPVDSNPG | VGWikF | - IF | --FAO |  |  |  | 821 |
| Tthe/1-822 | 782 GIGVOLDLDVFGAL | LPSLRLDYAF | ---SPE | TG-RI |  | P-MF-- |  |  |  | 822 |
| Fnuc/1-678 | 639 GVGIRLN--- TP. | IGPLRFDFGW | PVGILMM | M--kF | - YFI | -SF-- |  |  |  | 678 |
| Anal/1-833 | 799 GLGVRVQ----5P. | -LGPIRIDYGI | --NI | S--RI | - NFG | -RF-- |  |  |  | 833 |
| AthaV/1-732 | 698 GLGVRVD --- | LGPLRLEYAF | -- ND | MG--RF | - HFO | -RN-- |  |  |  | 732 |
| AthaIII/1-818 | 787 GAGVK - .-. | LGLVPAEYAV | --DH1 | T--AL | FFF | - $\mathrm{Pl}^{\text {l }}$ - |  |  |  | 818 |
| AthaM/1-524 | 484 GAGIVVP---TS | LFFWEL ${ }^{\text {a }}$ CH | ILKKQEL | AKS-GF | - P M | T-SS- |  |  |  | 524 |
| Xlae/1-468 | $426 \text { GAGLVLRLG- } N-$ | IARLELIVCI | FMGVQS | ICD-GV- | $\cdots Q F$ | $I-R F-L$ |  |  |  | 468 |
| Ptri/1-690 | $653 \text { GVGVRKSVQ }$ | GIPLKYDLCY | $-\mathrm{SSD}$ | KFQS | ….-AFG | $\mathrm{A}-\mathrm{DFDF}$ |  |  |  | 690 |
| Cmer/1-840 | 792 GLGLRV | LGALFMECGW | ILDRAPG | 3G--GFP1 | SRDGYW | VDVSY |  |  |  | 840 |
| Cmer_Sam50/1-541 | 499 GLGLACC----TS- | LGRFEVMWAR | ILRFGDG | DE-GF- | ---SI | -SFSP |  |  |  | 541 |
| -Tpse/1-804 | 768 GIGFRKRVQ-- | GIPLIMDACI | ILR | AR--GV | - FFG | --DFGV |  |  |  | 894 |
| Ehux/1-636 | 619 GGAPAVE-- |  |  |  | - AEF | GDGSVP |  |  |  | 635 |
| Tgon/1-1425 | 1385 GVCVR-- | LaLlsVIFAR | PIAFRIGG | LPGRF | --FVI | DHSF-- |  |  |  | 1425 |

Supplemental Figure 2. Alignment of Omp85 sequences for the calculation of the phylogenetic tree presented in Figure 1B
Representation of the alignment of all Omp85 sequences used in the phylogenetic analysis (Fig. 1B) created with JalView (Waterhouse, A.M., Procter, J.B., Martin, D.M.A, Clamp, M. and Barton, G. J. (2009) Bioinformatics 25, 1189-1191) and colored with the "Clustal-X" color scheme.

## A

>gi|30687572|ref|NP_189102.2| PL-SP1 [Arabidopsis thaliana]
MMVMISLHFSTPPLAFLKSD $\overline{S N}^{\prime}$ SRFLKNPNPNFIQFTPKSQLLFPQRLNFNTGTNLNRRTLSCYGIKDSSETTKS APSLDSGDGGGGDGGDDDKGEVEEKNRLFPEWLDFTSDDAQTVFVAIAVSLAFRYFIAEPRYIPSLSMYPTFDVG DRLVAEKVSYYFRKPCANDIVIFKSPPVLQEVGYTDADVFIKRIVAKEGDLVEVHNGKLMVNGVARNEKFILEPP GYEMTPIRVPENSVFVMGDNRNNSYDSHVWGPLPLKNIIGRSVFRYWPPNRVSGTVLEGGCAVDKQ
>estExt_gwp_gw1.C_chr_10974| Phatr2:17972 [Phaeodactylum tricornutum] MPTISDHTHACRAANLASSPQRRMVLLLALCLSLVAPSVTAFRPSTPPAVRSSVFVSSRTKISWADASSSRRKQI LGMS DNKNDS PVDRKPASEPGLKASLRNKWEDETPASEPDAGFLEGIKRWFNSDEGREDVKTYFISLFLALLLRF TIIEPRFIPSLSMYPTFEVGDQLAVEKVTKRIKPFYRTEVVVFQPPQAFRDIVENQYGDKSKGKEALIKRIVAVE GDKVEIKNGKLLINDIEQEEAYTAEDAQYAFGPVRVPPENVLVLGDNRNHSLDGHIWGFLPTKNVIGRAVFVYWP PWRVGNGGMF

## B

|  | 20 * 40 |  |
| :---: | :---: | :---: |
| A.thaliana : <br> P.tricornu : | MMVMISLHESTPELAFLKSDSN------SRFLKNPNENFIQE | : 38 |
|  | -MPTISDHTHACRAANLASSPQRRMVLLLALCLSLVAESVTAERESTPPA | 49 |
|  | 60 * 80 * 100 |  |
| A.thaliana : <br> P.tricornu : | FPQRLNENTGTNLNRRTLSCYGIKDSSE--TTKSAESLDSEDGG | 85 |
|  | VRSSVEVSSRTKISWADASSSRRKQILGMSDNKNDS PVDRKEASERGLKA | 99 |
|  | * 120 * 140 * |  |
| A.thaliana : <br> p.tricornu : | GGDGGDDDKGEVEEKNRLEPENLD--ETSD----DAQTVEVATAVSLAFR | 129 |
|  | SLRNKWEDET PASEPDAGELEGIKRWENSDEGREDVKTYEISLFLALLLR | 149 |
|  | 160 * 180 * 200 |  |
| A.thaliana <br> p.tricornu | YEIAEPRYIPSLSMYPTEDVGDRLVAEKVSYYFRKPCANDIVIEKSEPVL | 179 |
|  | GTIIEPREIPSLSMYPTEEVGDQLAVEKVTKRIKPEYRTEVVVEQPEQAF | 199 |
|  | * 220 * 240 * |  |
| A.thaliana P.tricornu | QEVGY-------TDADVEIKRIVAKEGDLVEVHNGKLMVNGVARNEKEIL | 222 |
|  | RDIVENQYGDKSKGKEALIKRIVAVEGDKVEIKNGKLLINDIEQEEAYTA | 249 |
|  | 260 * 280 * 300 |  |
| A.thaliana <br> P.tricornu | EPPGYEMT PIRVEENSVEVMGDNRNNSYDSHVWGPLPLKNIIGRSVERYw | 272 |
|  | EDAQYAF GPVRVEPENVLVLGDNRNHSLDGHIWGFLPTKNVIGRAVEVYW | 299 |

A.thaliana : PENRVSGTVLEGCCAVDKQ : 291
P.tricornu : REWRVG----NGGMF---- : 310

Supplemental Figure 3. Identification of the putative PL-SP1 homologue in $P$. tricornutum
$A$. The two sequences which were compared are shown. The predicted signal peptide is shown in blue and the predicted transit peptide in red. $B$. The alignment of $A$. thaliana plastidic type I signal peptidase 1 and the respective homolog from $P$. tricornutum is shown. Note that the conserved functional domain starts at $\sim$ aa120.

