Table S1. Protein spot identification.

The MS/MS spectra of each protein spot were searched against the NCBI non-redundant protein database by Mascot. Protein name and NCBI accession number were extracted from positive hits. When there were multiple protein hits for homologous proteins, or proteins belonging to the same family, only the one with highest protein score (Prot_score) was listed. Protein spots with multiple hits belonging to different protein families or without any positive hits were listed as “not identified”. Theoretical molecular weight (Mw^{T}) and isoelectric point (pI^{T}) of hit proteins were calculated by ExPASy Compute pI/Mw Tool (http://web.expasy.org/compute_pi/). Because most of the hit proteins were unverified hypothetical proteins, the computed Mw^{T} and pI^{T} may differ from the actual protein. Assembled *T. oceanica* transcripts obtained from an mRNA sequencing project[1] were used to revise gene models predicted by Lommer et al.[2] Revised Mw and pI were listed as Mw^{M} and pI^{M}.

Spot No.	**F-ratio**	**p-value**	**Accession No.**	**Protein name**	**Species**	**Mw^{T} (kDa)**	**Mw^{M} (kDa)**	**pI^{T}**	**pI^{M}**	**Prot score**	**Prot matches (sig.)**	**Prot sequences (sig.)**	**Coverage**
1	2.07	4.80E-02	YP_004072574	cytochrome c550	Thalassiosira oceanica	17.97	14.98	7.62	5.91	459	17 (13)	5 (5)	40%
2	6.56	8.10E-07	EJK49624	hypothetical protein THAOC_31479	Thalassiosira oceanica	15.42		5.70		906	38 (38)	8 (8)	68%
3	2.16	1.00E-04	XP_002296360	acyl-CoA dehydrogenase	Thalassiosira pseudonana	88.23	86.13	6.51	6.00	605	30 (23)	11 (10)	14%
4	2.09	1.30E-06		not identified									
5	1.54	5.60E-04	EJK67362	hypothetical protein THAOC_11616	Thalassiosira oceanica	78.54	89.50	6.27	5.88	1222	63 (52)	21 (19)	32%
7	1.88	3.40E-03	EJK67362	hypothetical protein THAOC_11616	Thalassiosira oceanica	78.54	89.50	6.27	5.88	552	78 (28)	26 (16)	34%
8	1.58	4.30E-03	EJK51095	hypothetical protein THAOC_29771	Thalassiosira oceanica	43.00		5.17		2400	38 (25)	7 (6)	21%
9	1.75	3.00E-03	EJK51095	hypothetical protein THAOC_29771	Thalassiosira oceanica	43.00		5.17		788	45 (295)	8 (7)	28%
10	1.58	1.20E-03	EJK51095	hypothetical protein THAOC_29771	Thalassiosira oceanica	43.00		5.17		84	6 (4)	3 (2)	7%
11	1.96	2.30E-03	EJK62503	hypothetical protein THAOC_16878	Thalassiosira oceanica	71.72		7.11		204	8 (6)	5 (3)	10%
12	1.57	1.30E-02		not identified									
13	1.60	3.10E-03		not identified									

Table S1. Continued.

Spot No.	F-ratio	p-value	Accession No.	Protein name	Species	Mw^T (kDa)	Mw^M (kDa)	pI^T	pI^M	Prot score	Prot matches (sig.)	Prot sequences (sig.)	Coverage
14	1.55	2.30E-04		not identified									
15	1.63	2.00E-03		not identified									
16	1.55	8.20E-03		not identified									
18	1.60	4.50E-03	EJK47576	hypothetical protein THAOC_33694	Thalassiosira oceanica	18.07	32.88	6.12	6.65	242	10 (6)	2 (2)	13%
19	1.51	4.40E-03	EJK56070	hypothetical protein THAOC_24110	Thalassiosira oceanica	22.01	30.68	4.85	5.86	568	19 (19)	5 (5)	27%
20	1.73	4.20E-04	EJK74364	hypothetical protein THAOC_03961	Thalassiosira oceanica	123.90	23.06	7.47	8.80	62	38 (3)	9 (2)	8%
21	1.64	3.30E-05	EJK63584	hypothetical protein THAOC_15749	Thalassiosira oceanica	39.26	35.61	8.33	6.03	165	8 (8)	4 (4)	10%
22	1.72	2.50E-03	EJK68249	hypothetical protein THAOC_10589	Thalassiosira oceanica	25.82		6.53		341	23 (10)	8 (4)	37%
23	1.53	5.30E-03	EJK74743	hypothetical protein THAOC_03563	Thalassiosira oceanica	19.50		4.95		167	8 (6)	4 (2)	19%
24	-2.03	3.60E-04	EJK53643	hypothetical protein THAOC_26871	Thalassiosira oceanica	54.34	49.47	6.47	6.10	256	5 (5)	3 (3)	8%
25	-2.12	9.10E-07	EJK53643	hypothetical protein THAOC_26871	Thalassiosira oceanica	54.34	49.47	6.47	6.10	2088	74 (72)	13 (13)	33%
26	-2.90	1.30E-02	EJK47917	hypothetical protein THAOC_33331	Thalassiosira oceanica	22.22	36.58	6.19	5.76	1354	41 (31)	7 (7)	38%
27	-2.39	1.50E-02		not identified									
28	-2.04	1.10E-05	EJK55038	hypothetical protein THAOC_25270	Thalassiosira oceanica	21.31	18.67	4.52	4.38	377	11 (8)	2 (2)	14%
29	-3.01	3.90E-07	EJK71623	plastocyanin precursor	Thalassiosira oceanica	27.49	23.35	4.92	4.87	2051	43 (40)	7 (6)	49%
30	-2.03	5.20E-06		not identified									
31	-2.34	1.10E-05	EJK47429	hypothetical protein THAOC_33854	Thalassiosira oceanica	22.37		4.57		187	27 (7)	6 (3)	18%
32	-2.37	6.40E-04	EJK66277	hypothetical protein THAOC_12811	Thalassiosira oceanica	12.66		5.18		66	21 (4)	4 (1)	38%
33	-3.74	5.00E-04		not identified									
34	-1.89	2.20E-02	EJK78098	nitrite reductase	Thalassiosira oceanica	68.16	65.75	5.72	5.65	1187	57 (31)	23 (16)	41%
35	-1.56	8.30E-04	YP_004072542	ATP synthase CF1 subunit beta	Thalassiosira oceanica	51.17		4.69		3823	166 (144)	20 (19)	73%
36	-1.64	3.70E-04	EJK53643	hypothetical protein THAOC_26871	Thalassiosira oceanica	54.34	49.47	6.47	6.10	247	4 (4)	2 (2)	6%
39	-1.68	1.40E-05	EJK54757	hypothetical protein THAOC_25588	Thalassiosira oceanica	44.97	43.09	6.03	5.88	960	40 (39)	13 (13)	33.70%
40	-1.59	4.90E-04		not identified									
41	-1.51	2.60E-04	EJK70534	hypothetical protein THAOC_08095	Thalassiosira oceanica	23.48	20.89	5.04	4.90	1355	118 (69)	8 (6)	26%
44	-1.69	7.90E-04	EJK48286	hypothetical protein THAOC_32932	Thalassiosira oceanica	22.59	19.85	4.66	4.43	146	12 (3)	3 (1)	10%

Table S1. Continued.

Spot No.	F-ratio	p-value	Accession No.	Protein name	Species	Mw^T (kDa)	Mw^M (kDa)	pI^T	pI^M	Prot score	Prot matches (sig.)	Prot sequences (sig.)	Coverage
45	-1.57	1.10E-04	EJK75762	hypothetical protein THAOC_02510	Thalassiosira oceanica	23.88	21.08	4.81	4.57	674	69 (58)	7 (7)	31%
46	-1.60	5.60E-04	EJK47511	hypothetical protein THAOC_33758	Thalassiosira oceanica	13.66	18.15	4.52	5.16	214	13 (8)	5 (5)	46%
49	-1.90	2.00E-02	EJK70163	hypothetical protein THAOC_08500	Thalassiosira oceanica	22.54	19.7	7.66	5.66	158	27 (14)	7 (5)	17%
50	-1.72	8.70E-05	EJK72668	hypothetical protein THAOC_05777	Thalassiosira oceanica	20.93	18.12	4.75	4.54	172	34 (21)	5 (5)	21%
52	-1.82	3.50E-03	YP_004072584	RuBisCO small subunit	Thalassiosira oceanica	16.12		4.95		1071	34 (33)	5 (5)	25.20%
53	-1.82	1.30E-02	EJK45858	hypothetical protein THAOC_35505	Thalassiosira oceanica	94.85		5.99		100	24 (4)	14 (3)	17%
54	1.59	3.60E-02	EJK55940	hypothetical protein THAOC_24261	Thalassiosira oceanica	69.63		5.23		248	4 (4)	3 (3)	8%
55	2.08	2.30E-03	EJK77039	hypothetical protein THAOC_01158	Thalassiosira oceanica	77.30		5.64		611	31 (21)	13 (11)	22%
58	3.04	1.30E-02	EJK49989	hypothetical protein THAOC_31088	Thalassiosira oceanica	74.31	76.75	5.75	5.61	208	12 (8)	8 (6)	11%
59	-1.77	1.60E-02	YP_004072585	RuBisCO large subunit	Thalassiosira oceanica	54.67		6.10		51	1 (1)	1 (1)	3.50%
62	-1.92	2.20E-05	XP_002293787	phosphoadenosine-phosphosulphate reductase	Thalassiosira pseudonana	49.41	47.27	4.99	4.94	200	21 (9)	7 (4)	13%
63	-1.65	2.70E-04		not identified									
65	-1.70	1.20E-04	EJK46456	hypothetical protein THAOC_34872	Thalassiosira oceanica	35.61	33.54	5.63	5.51	160	20 (10)	6 (4)	20%
66	2.10	1.10E-02		not identified									
71	-1.79	1.40E-04	XP_002294845	fucoxanthin chlorophyll a/c protein	Thalassiosira pseudonana	21.79	18.99	4.71	4.53	323	22 (11)	4 (3)	18%
72	1.64	1.70E-03	EJK67933	hypothetical protein THAOC_10956	Thalassiosira oceanica	17.17		5.56		146	11 (4)	5 (2)	28%

Spot 1: A 26-amino acid signal peptide were removed.[3]

Spot 3: A new protein was predicted from RNAseq transcript c16960_g1 to replace the *T. pseudonana* protein acyl-CoA dehydrogenase (XP_002296360).

Spots 5, 7: A new protein was predicted from RNAseq transcript c4619_g1 to replace the partial hypothetical protein THAOC_11616.

Spot 18: A new protein was predicted from RNAseq transcript c13191_g1 to replace the partial hypothetical protein THAOC_33694.

Spot 19: A new protein was predicted from RNAseq transcript c5774_g1 to replace the partial hypothetical protein THAOC_24110.

Spot 20: A new protein was predicted from RNAseq transcript c9230_g1 to replace the incorrect gene model of hypothetical protein THAOC_03961.

Spots 24, 25, 36: A new protein was predicted from RNAseq transcript c13958_g1_i1 to replace the incorrect gene model of hypothetical protein THAOC_26871.

Spot 26: A new protein was predicted from RNAseq transcript c15229_g1 to replace the incorrect gene model of hypothetical protein THAOC_33331.

Spot 29: A new protein was predicted from RNAseq transcript c1262_g1 to replace the incorrect gene model of EJK71623.

Spot 46: A homologous protein from a closely related species, *Fragilariopsis cylindrus* (OEU15753), was used to replace the partial hypothetical protein THAOC_ 33758.

Spot 58: A new protein was predicted from RNAseq transcript c15129_g1 to replace the partial hypothetical protein THAOC_ 31088.

Spots 28, 41, 44, 45, 49, 50, 71: Signal and transit peptides were removed.[4]

Spots 21, 24, 25, 34, 36, 39, 46, 49, 62, 65: Signal peptides (16 to 29 amino acids) predicted by SignalP were removed.[5]

Spot 1

> YP_004072574 cytochrome c550

~~MFKRYSKFCACILFCIFNLFVVSASA~~IDLDEATRTVTVDSSGKTTVLTPEQVKRGKRLYNATCGACHTGGITKTNPNVGLDPEALSLATPRRDNIE
ALVDYLKNPTTYDGLESIAEIHPSIKSADIYPRMRSLTDEDLYSIAGHIMLQPKIVAEKWGGGKIYF

Spot 3
>c16960_g1
MLSSRGAISRTAGARAAAARPLAAPRGGDRRGQARGILGVSHAIDKRVYRWAKGVLPPISKTENIALGCGTIGFDRDIFGGSPSLQHLIDTYTPAL
SDEEQSFLDVQVHHLCSILSDHDVATDKDFSREAWDYMRDERFFGMKIPREWGGLGFSTHAVSTILAKLATQCFDANATVAVPNSLGPGELLVRY
GTPDQKEYFLPRLADGTLIPCFGLTGPHSGSDATSLIGSDCVVSRHPDTGELGVVASFRKRYITLAPVAGVVGLGLNLRDPEGLLGGEGEEGFTVA
LLERGHEGLRMGPRHIPLSAAFMNGTVEGDDVWIPMDSILGGQGRCGFGWNMFVECLAEGRGVSLPAGSIGAARSVVAGVGAYSRVRKQFRVP
IAEFGGIQEAMAKAGSDGLITIAGGDVMNAIVDNHEAPMVISSVMKQNCTERGRRIVECGMDVAAGSAICRGERNYIGNAYMSLPIAITVEGANI
MTRSFQIIGQGLTRCHPHMSDLLRALQRPPSEEKEATAVFVRQFYKIVGHGVSNFFGSVGRGVTSSVSSAARSKTAYKNGDDLLAYHEKQLLRLS
NNFALTADLCFTLGGRLKFEELLMGRLADALGAIYLGYATLHHYHRRRGVDGLEALTEHAMLRLETEAQDALYAASENFPGPLGPLASAVMKV
GCFPLGGLTRPYSPPGDDLTKEVSRLLTTPSEIRDMFEENIYRAGEGAGPHQMTDLIDALPVCVEADRAASSVRREKREPTAQEADLIARADALR
DALIQVDVFEHATAAEAAPGYVRPALQGTEDRFAALERTVFREAA

Spots 5, 7
>c4619_g1
MAEIPEDVAKLWASLDPSVKAALISSEEANGTNGGAPKKVGNHNQGRRAIQSDGPLMESKLIGGDCFAKREANPSWLKTRTEVYDAVKAKRDE
ELAAKVPVDIEVALPDGKVLSEDKEGNKYQSWRTSPFDVACTISRGLADNSSVARVTYASYVSDYDLAMDGVGGGDVDGMAEAMEDLSLEEE
NKSEMTHLWDMTRPLVGNVSKIEFLKFDDDGDAKTTFWHSSAHILGEALEHVFGSRLTIGPPLAGGFYYDSFMGDAESGGALTESDYKIVEQEF
QKIVKSKQKFERLVVTKDEALEMFEGNPFKTQIISTKVPDGSRTTVYRCGDLVDLCRGPHLPHTGRVKAFAATRHSATNWLGDTDNDMLQRMY
GISFPDKKMLKVWKENQEKAKERDHRRIAAKQDLIMFHDLSAGSAFWLPHGARIYNKLCDFIRSHYWKRGFTEVITPNVYNLDLWNTSGHAQ
HYKDAMFCFDVEGKEWAMKPMNCPGHCLMFGSKLRSYRDLPIRYADFGVLHRNELSGALSGLTRVRRFQQDDGHIFCREDQIEQEVLGALDF
MSSCYKTFGMSYKLELSTRPKKALGELEMWNRAEAALERALNDFAGKGNWKINPEDGAFYGPKIDIKVMDAMDRIHQCATIQLDFQLPIRFNL
EYNTGTKENPFARPVIVHRAVLGSVERSFAVLCEHFGGKWPLWLSPRQVMLIPIHADFNDYCQQVRDRLHDEGFYADVDLSKATFSKKVRTAQV
DQYNFQLVVGKNEVENGTVNIRTRDNKQEGEKKIDDLVAYLKKLEAEYQ

Spot 18
>c13191_g1
MSKPSKPMSASKKRKAGPGVAAASGTTLDDLTALTTAQLYSLSSRRHRDESLGNVRRAQEERDRAYLSLEQSQRWLERASEHLERMEENAKEA
HKEYLDAKGLLTRVRLASSSAVDMMGGKVAAEELSDDEAGLEAIIAKKKKSATSSAKSRKIPFIQDVIIDMAKRGELVDGKKLTTAESEMVAKK

DRTHFVQTMMLIEELWTEEEELFLRSPVDEILESKEELQVIVETIALRCLTKLNEWEGRVDPTPTAHQKPSYISLGTRARRIFHSRKKASKAENETE
AEEVTKVDV

Spot19
>c5774_g1
MFSRRLPSIARRAFSTPSIAPARTTPSAKALALKDFDAEVLQSQAHLEDQKSEIQKYREEVSAIDSHAVYAVTEPLPEPDLPENKSEVAALDPAYKN
QIPLNPDGSEKWVVITQSQAKWPGQAPLGKESEWVISFQDNGETAETWSNPLMGWVSSADPMANNMRLQMSFETAEEAKYFAEKRGWKFTIE
RPIIRRGRDDDAQYQDVFLPQSVAGKVKRDGHKCDHWYRDQAGASHYFRPLKYHGDGTVRQHGPNMEQASEKDTEGYYKMR

Spot 20
>c9230_g1
MSSSDEGIDHLESEEEVVAPKPKRQRKPKKKKDPNAPKRNQSAFFLYSNANRNRVKAENPDAKFGDIAKLISVEFKALSEHERAKWDKLAAED
KERYQREMEDYEPPSDLEDDAPKRKAKKDPNAPKRNQSAFFLYSNATRNDVKEANPEAKFGDIAKIISTHFKALPVEERAYWDNKAAEDKARY
QQQLQMYKETGHF

Spots 24, 25, 36
>c13958_g1
MKLAIASLCVGSTTAFSSFMGQNVAHAPATSSSALSMKYKVAVVGGGPSGACAAEIFAQEKNIDTVLFERKMDNAKPCGGAIPLCMIGEFDIPET
TVDRKVRRMKLISPTNVEVDIGDTLQPNEYIGMCRREIMDKFLRDRAISYGAEPINALVTAIDVPADHVENPDAKYNIKYSEFVEGSRTGTAKEM
MVDLIVGADGANSRVAKAMDAGEYNFAIAFQERIKISDEKLKFYEEMAEMYVGDDVSPDFYGWVFPKYDHVGVGTGTVVNRPAIKQYQKAIR
DRAGDKIAGGKIIKVEAHPIPEHYRPRRVQGRMALVGDAAGYVTKCSGEGIYFAAKSGRMAAEAIVKLMDGGRRLPTQADIERTYIADYDKLY
GPTYTVLDILQKVFYSNNGAREAFVELCNSKYVQQVTFDSYLYKKVQGNNPLDDIKLLGETIGCLIKGYSIAKPDAEFSNPVESMKRL

Spot 26
>c15229_g1
MTVKLGINGFGRIGRLVCRAALEHEGDVMPVAVNDPFLSLDYAAYLFQYDSVHGKYPGTVTADADSNSLIIDDGKTKVSIKFFAERNPSDIPWSS
VDASYVCESTGVFTTTEKAKAHLGGGAKKVIISAPSADAPMYVVGVNHKKYDGSADVVSNASCTTNCLAPLAKVINEVYGIQEGLMTTVHAS
TATQLVVDGPARGGKDWRGGRAAVANLIPSSTGAAKAVGKVIPELNGVLTGMAVRCPTPDVSMVDLTVKLKKGCTKDEMLATLKAASEGDEL
KGVLGYTDHAVVSQDFVHDNRSSIVDGTACIALNDTFHKVISWYDNEWGYSNRLVDLAVFMSTVDK

Spot 29

>c1262_g1
MKFVLAACLAAAVSAEGLRKAEPEIESNHLDLELEINGERQLFPLLPGTKCPTGHTCRTRAVEGGVSPMINSLKRNIKTPLAMSMDWIDMNGEL
ETVVMSDNFCTRRNAMAKAAGLAAGLSMAAVSAPAYAAQTVEVKMGADSGLLVFEPAKVTVCKGDTVKWINNKAGPHNVVFDEDNIPDGV
DQEKISMDDQLGEPGDTFEMKFDTAGTYGYYCEPHRGAGMQATLVVQ

Spot 46
>OEU15753 hypothetical protein FRACYDRAFT_187021 [Fragilariopsis cylindrus CCMP1102]
MAAVAAFALVSTLSFTEPAFAVEGGGATTAANAKITTGGASTLQSGRTIAITRGVNLDRSDFSNQNLKGVAFQQSIVRDSNFKGCNLYGSSFFDAT
LDGSNFEDADMSLSNVEMAQFNRANLHNTIMREVYVSGSTLFEGIKDIEGSDWSETYLRADQKKLLCEHPTSKGTNPVTGVNTRESLMCKD

Spot49
>EJK70163.1 hypothetical protein THAOC_08500
~~MMKIALCAALVTSAAAFAPQANNARPSVA~~VQAENSRREALSSIAAAAAVIAPAAANAAAGESPRFSVFGVIGDGTSYSEGAAYGSDQSSAVYSP
YSVYGNVGSSDALYSANNAGEVERKKGFIAESQKRLSKLPAYVERKEWFNVKDELTRYMYETRGAVRGLAETPEQKKIAKSFFQAIEEASLQAT
LKNQDKCAAASADSAKLLDQFVASL

Spot 58
>c15129_g1
MSAEAETKPTPAEAPSAPAPAAQSAAPGPSASGGAPNSFTSASLYVGDLLPEVNEGLLYEIFNAVGPVASIRVCRDAVTRRSLGYAYVNYHQAAD
AERALDSMNFTDIKGKPCRIMWSQRDPSVRRSGVGNIFVKNLHEGIDNKQLYDTFSLFGNILSCKVVCDRETGLSKGYGYVHYETNEAAASAID
KLDGMLIDGKEVQVGVFMRRDTRPDQEVYTNLFIKNMPYEWDDKRLEAEFAEFGEIVSASVKMGKRKKFAKKGKKAEAKKDEDDKEDGAED
KPAEEEKPEEEEAKPAEEEAKPESTEPESLGFGFVNFATHEAAAAAVEAMNDKVYKVTEDGDEVEKALFVGRAQKKSERQAELRSKYEAEKME
RIAKFQGVNLYVKNLDDAVTDDMLRDEFGGMGTITSAKIMRDAKTNNSRGFGFVCFSTPEDATRAVNEMSGKIVAGKPIYVSLAQRREVRRAQ
LEAQHAGGRGGGPGGPGMMRGGPMGGPPMGGGYPGQVPMYMPRPGPGMQPSYPMGGPMMGGRGGYPGRGGMQPMQPGYGMMGGRGGP
GGRGQPGPYGRGPGRGAPYGRGRGPMGPGPGRGGMPGQPPVQFHGGVRNAGMPPGPGGPMPGGPGDVPQGPPPPQQQQEAPPSANEQLTP
AALASATPEIQKNMIGERLYPLIHQTQPDLAGKITGMLLEMDNSELLHLLESPEALGAKIQEALQVLDAHNAADK

References

1. L. Kong, Molecular and physiological responses of an oceanic diatom to copper deficiency, PhD thesis, McGill University, 2019.
2. M. Lommer, M. Specht, A.-S. Roy, L. Kraemer, R. Andreson, M. A. Gutowska, et al., Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation, *Genome Biol*., 2012, **13**, R66, DOI: 10.1186/gb-2012-13-7-r66.
3. M. Roncel, D. Kirilovsky, F. Guerrero, A. Serrano and J. M. Ortega, Photosynthetic cytochrome c550, *Biochim. Biophys. Acta Bioenerg*., 2012, **1817**, 1152-1163.
4. M. Lang and P. G. Kroth, Diatom fucoxanthin chlorophyll a/c-binding protein (FCP) and land plant light-harvesting proteins use a similar pathway for thylakoid membrane insertion, *J. Biol. Chem*., 2001, **276**, 7985-7991.
5. J. J. A. Armenteros, K. D. Tsirigos, C. K. Sønderby, T. N. Petersen, O. Winther, S. Brunak, et al., SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotech*., 2019, **37,** 420-423.