

Supplementary Information

How haptophytes microalgae mitigate vitamin B₁₂ limitation

Charlotte Nef*, Sébastien Jung, Francis Mairet, Raymond Kaas, Dominique Grizeau and Matthieu Garnier

*Corresponding author

Supplementary Table 1. Reference of cobalamin-dependent methionine synthase (*metH*) sequences compiled for phylogenetic analysis. Identity and similarity percentages with *T. lutea* sequence are provided.

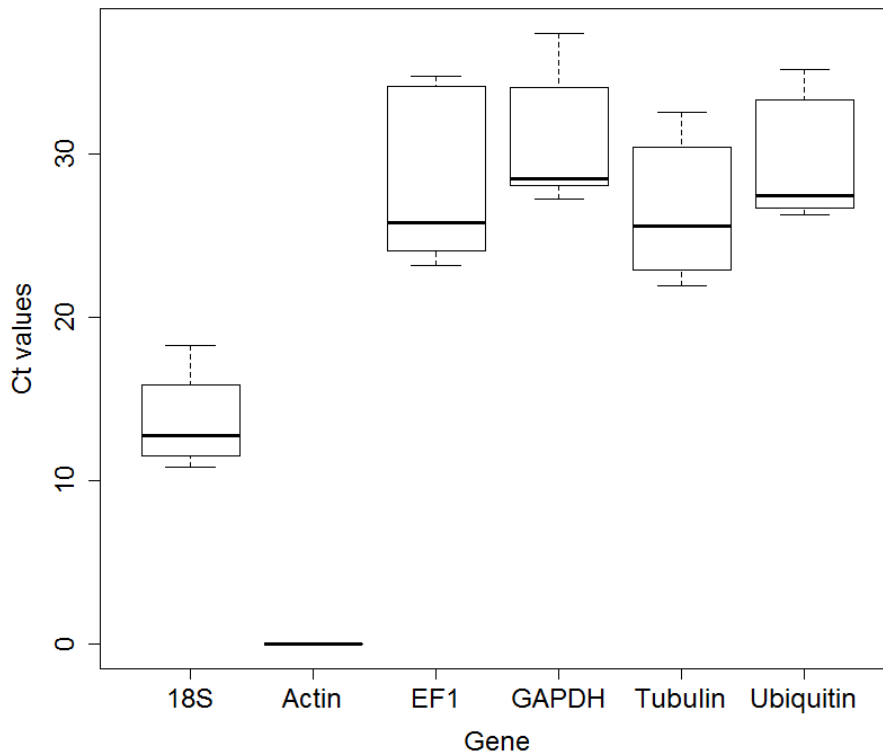
Data	Family	Species	Strain	iMicrobe sample reference	iMicrobe sample collection site	Sequence name	Evalue	Score	Identity/Similarity (%)
Genomic	Isochrysidaceae	<i>Tisochrysis lutea</i> *	CCAP 927/14			1602	-	-	-
Transcriptomic	Isochrysidaceae	<i>Isochrysis</i> sp.	CCMP 1244	MMETSP1 090	North Atlantic	CAMNT_004201 9453	0	2407	86.8/95.8
Transcriptomic	Isochrysidaceae	<i>Isochrysis</i> sp.	CCMP 1323	MMETSP0 944	Irish Sea	CAMNT_002168 4687	0	3001	92.7/97.9
Transcriptomic	Prymnesiaceae	<i>Chrysochromulina rotalis</i>	UIO 044	MMETSP0 287	North Sea	CAMNT_000329 9591	0	2098	76.2/88.9
Transcriptomic	Prymnesiaceae	<i>Chrysochromulina brevifilum</i>	UTEX LB 985	MMETSP1 094	English Channel	CAMNT_001593 7587	0	2052	74.9/89.5
Transcriptomic	Prymnesiaceae	<i>Prymnesium parvum</i>	Texoma1	MMETSP0 006	Lake Texoma, Oklahoma	CAMNT_002493 4099	10 ⁻¹⁰³	379	77.7/91.0
Transcriptomic	Pavlovophyceae	<i>Pavlova gyrans</i>	CCMP 608	MMETSP1 466	Helford river, Falmouth, UK	CAMNT_005340 6913	0	1917	70.9/87.2
Transcriptomic	Pavlovophyceae	<i>Pavlova lutheri</i>	RCC 1537	MMETSP1 463	North Sea	CAMNT_005333 6917	0	1897	71.8/86.6
Transcriptomic	Phaeocystaceae	<i>Phaeocystis antarctica</i>	CCMP 1374	MMETSP1 444	Ross Sea	CAMNT_004384 2521	10 ⁻¹¹⁸	424	77.1/89.4
Transcriptomic	Noëlarhabdaceae	<i>Gephyrocapsa oceanica</i>	RCC 1303	MMETSP1 364	-	CAMNT_002797 1079	0	2387	86.7/95.7
Transcriptomic	Noëlarhabdaceae	<i>Emiliana huxleyi</i>	379	MMETSP0 994	English Channel	CAMNT_004161 3577	0	2407	86.8/95.8

Transcriptomic	Noëlarhabdaceae	<i>Emiliana huxleyi</i>	PLY M219	MMETSP1 150	-	CAMNT_002393 7657	0	2409	86.8/95.7
Transcriptomic	Noëlarhabdaceae	<i>Emiliana huxleyi</i>	CCMP 370	MMETSP1 154	-	CAMNT_002425 9279	0	2408	86.8/95.7
Transcriptomic	Calcidiscaceae	<i>Calcidiscus leptoporus</i>	RCC 1130	MMETSP1 334	South Atlantic	CAMNT_000736 5605	2×10^{-98}	361	77.4/90.9
Transcriptomic	Pontosphaeraceae	<i>Scyphosphaera apsteinii</i>	RCC 1455	MMETSP1 333	Spanish coast, Mediterranean Sea	CAMNT_000731 3019	6×10^{-86}	319	79.6/91.8
Transcriptomic	Pleurochrysidaceae	<i>Pleurochrysis carterae</i>	CCMP 645	MMETSP1 136	-	CAMNT_001011 6929	10^{-145}	517	78.9/91.0
Transcriptomic	Coccolithaceae	<i>Coccolithus pelagicus</i> ssp. <i>braarudi</i>	PLY 182g	MMETSP0 164	English Channel	CAMNT_002553 2985	3×10^{-90}	333	76.5/89.5
Proteomic	Prymnesiaceae	<i>Chrysochromulina</i> sp.	CCMP 291	-		KOO28302.1	0	2006	74.9/88.4
Proteomic	Noëlarhabdaceae	<i>Emiliana huxleyi</i>	CCMP 1516	-		XP_005791493.1	0	2228	82.0/91.4
Proteomic	Thalassiosiraceae	<i>Thalassiosira pseudonana</i>	CCMP 1335	-		XP_002293120.1	0	1352	55.8/79.9
Proteomic	Phaeodactylaceae	<i>Phaeodactylum tricorutum</i>	CCAP 1055/1	-		XP_002184461.1	0	1320	56.3/78.7

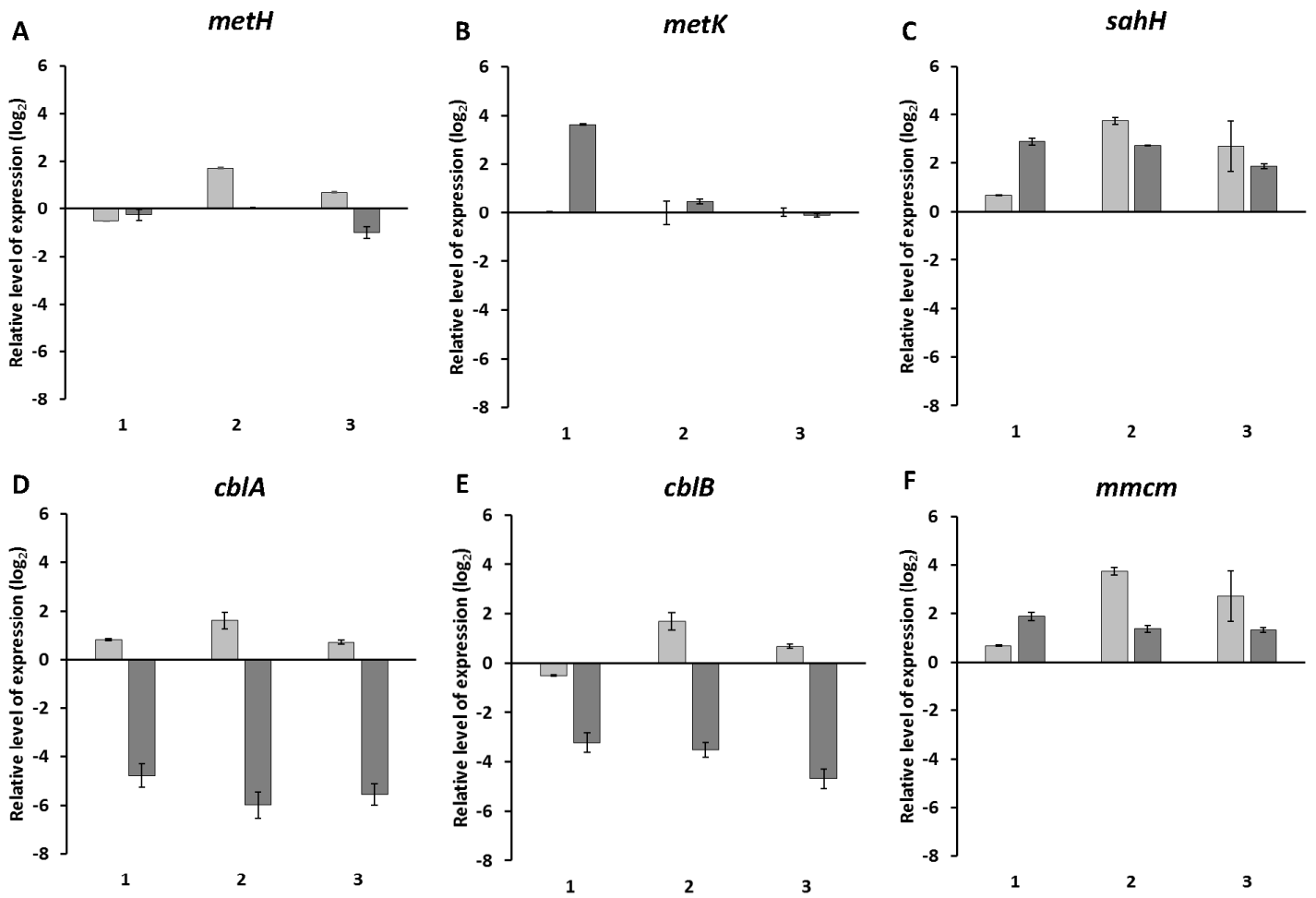
*See Carrier et al. (2018) and Bertheliet et al. (2018) for *T. lutea* genome.

Supplementary Table 2. Primer sequences used for RT-qPCR analyses.

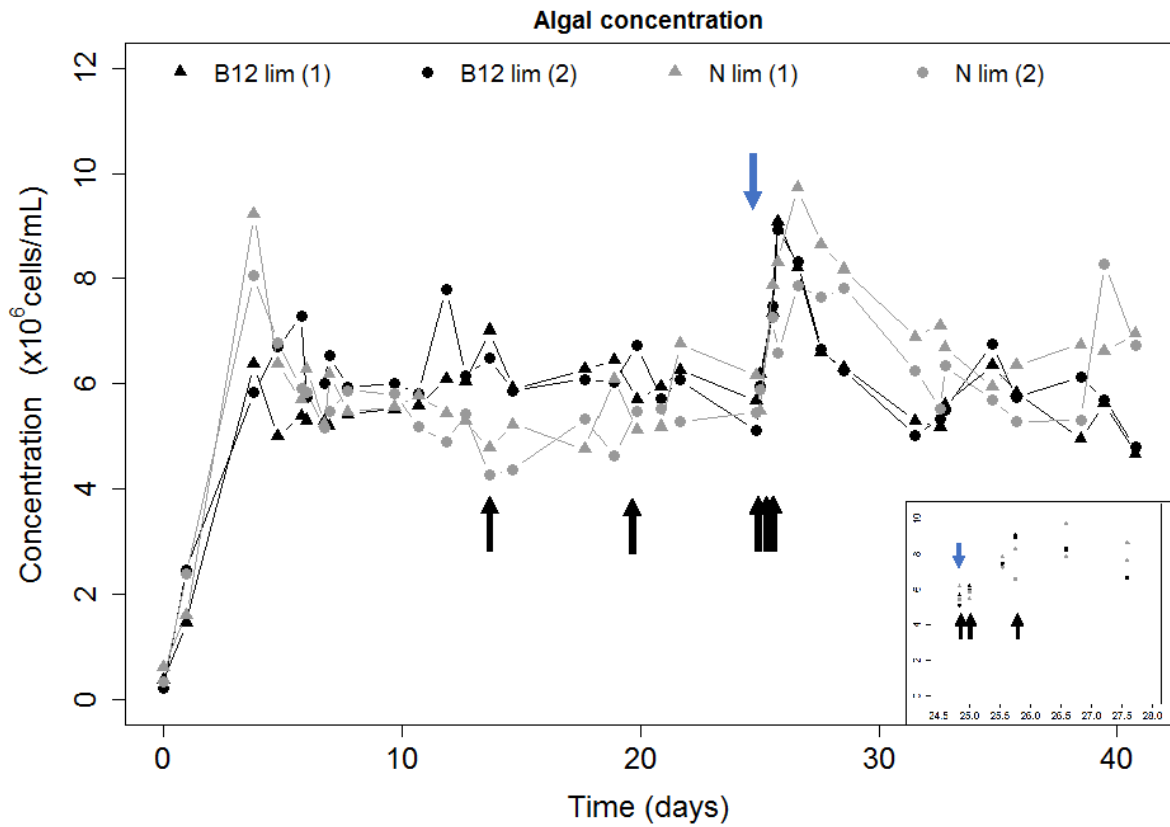
Gene	Primer name	Primer sequence (5' → 3')
<i>metH</i>	METH-F	CGAACTGCTCCATCACAAGA
	METH-R	GACGGATGCCTTGGTACTTG
<i>metK</i>	METK-F	TATTGCCGTTCTTCAGCTTG
	METK-R	AAGTACACTGGCCCGGAGAT
<i>sahH</i>	SAHH-F	GCTGGAGGAGTACTGGTGGT
	SAHH-R	CTCATGGAGGAGGAGGGTTG
<i>cblA</i>	CBLA-F	GTGCAGCTTGTCCAGCACT
	CBLA-R	CTCAGTCATGTGACCCAACG
<i>cblB</i>	CBLB-F	CGTCGATCCATCCACCAC
	CBLB-R	TGAGCCAACATCAAGTAGTCG
<i>mmcm</i>	MMCM-F	CACGCAGTCTGTGCAAG
	MMCM-R	CGAGAAGACCAAGGGAGACT



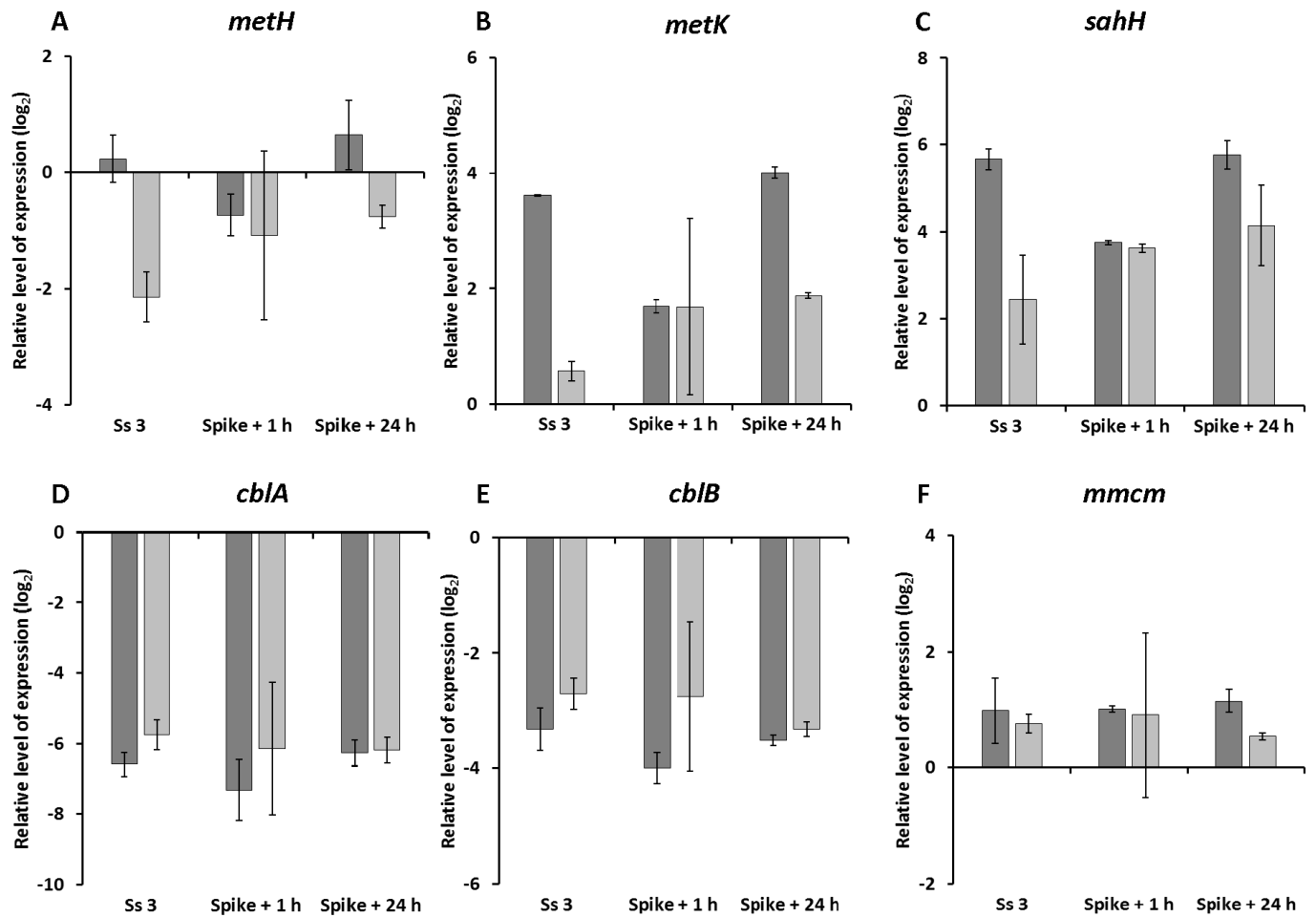
Supplementary Figure 1. Boxplot of Ct variation for reference genes tested with bold line indicating median (n = 18).



Supplementary Figure 2. Genes expression in B₁₂-limited batch cultures of *T. lutea*. Relative levels of expression of (A) *METH*, (B) *METK*, (C) *SAHH*, (D) *CBLA*, (E) *CBLB* and (F) *MMCM* genes. Values represent mean expression level at exponential phase (gray) and stationary phase (black) quantified and normalized using *GAPDH* as reference gene. Data are log₂ normalized. Values are shown for each biological replicate (1, 2 and 3). Bars indicate means of technical triplicate measurements and error bars represent one standard deviation.



Supplementary Figure 3. Algal concentration for chemostat cultures of *Tisochrysis lutea*. Blue arrows represent nutrient spike, black arrows indicate sampling points (Ss 1, 2, 3, spike + 1 h and spike + 24 h). Data are for B₁₂-limited (black) and nitrogen-limited (gray) chemostats.



Supplementary Figure 4. Genes expression in chemostat cultures of *T. lutea* during steady-state (Ss), 1 and 24 hours after nutrient spike for (A) *METH*, (B) *METK*, (C) *SAHH*, (D) *CBLA*, (E) *CBLB* and (F) *MMCM* genes. Data are for B₁₂-limited (black) and nitrogen-limited (gray) chemostats and are log₂ normalized. Values represent mean expression level quantified and normalized using *GAPDH* as reference gene. Bars indicate means of biological duplicates and error bars represent the range.

References

Carrier, G. C. *et al.* Draft genome and phenotypic characterization of *Tisochrysis lutea* strains. Toward the production of domesticated strains with high added value. *Algal Res.* **29**, 1–11 (2018).

Berthelie, J. *et al.* A transposable element annotation pipeline and expression analysis reveal potentially active elements in the microalgae *Tisochrysis lutea*. *BMC Genomics* **19** (2018).