

## Supplement

Phytoplankton pigments, group composition and absorption coefficients:

The following diagnostic pigments were used to identify seven phytoplankton groups: fucoxanthin (Fuco), peridinin (Peri), alloxanthin (Allo), 19'-hexanoyloxyfucoxanthin (19HF), 19'-butanoyloxyfucoxanthin (19BF), zeaxanthin (Zea), and total chlorophyll-b (TChl-b, i.e. the sum of monovinylchl-b, chl-b, and divinylchl-b (div-b)). According to Hirata *et al.* (2011) the weighted relationships of these diagnostic pigments (DPw) were calculated by multiple regression analysis as follows:

$$\% \text{ pico } (<2 \mu\text{m}) = 100*(0.86 \text{ Zea})/\text{DPw}$$

$$\% \text{ nano } (2 -20 \mu\text{m}) = 100*(1.27 \text{ 19HF} +1.01 \text{ TChl-b} +0.35 \text{ 19BF} +0.6 \text{ Allo})/\text{DPw}$$

$$\% \text{ micro } (>20 \mu\text{m}) = 100*(1.41 \text{ Fuco} +1.41\text{Peri})/\text{DPw}$$

$$\% \text{ diatoms} = 100*(1.41 \text{ Fuco})/\text{DPw}$$

$$\% \text{ dinoflagellates} = 100*(1.41 \text{ Peri})/\text{DPw}$$

$$\% \text{ haptophytes} = 100*(1.27\text{19HF})/\text{DPw}$$

$$\% \text{ chrysophytes} = 100*(0.35\text{19BF})/\text{DPw}$$

$$\% \text{ cryptophytes} = 100*(0.6\text{Allo})/\text{DPw}$$

$$\% \text{ chlorophytes} = 100*(1.01\text{TChl-b})/\text{DPw}$$

$$\% \text{ all cyanobacteria} = 100*(0.86\text{Zea})/\text{DPw}$$

where,

$$\text{DPw} = 0.86 \text{ Zea} +1.01 \text{ TChl-b} +1.27 \text{ 19HF} +0.35 \text{ 19BF} +0.6 \text{ Allo} +1.41 \text{ Fuco} +1.41 \text{ Peri}.$$

By multiplying the total chl-a concentration (TChl-a) (i.e. the sum of monovinylchl-a, chl-a, and divinylchl-a, div-a) with %-values for each group, the chl-a concentration for each group was derived. In addition, TChl-a concentration of prochlorophytes, a subgroup of cyanobacteria which is characterized by very low size (~0.5  $\mu\text{m}$ ), and the pigments of div-a and div-b, was calculated from div-a/(div-a+chl-a). The chl-a concentration of all other cyanobacteria was calculated by subtracting prochlorophytes chl-a from all cyanobacteria chl-a concentration.

Identifying phytoplankton assemblages with hierarchical cluster analysis:

The HCA clustering based on the hyperspectral phytoplankton absorption coefficients provided a cluster tree in which the input data set was partitioned into subsets or clusters with no previous information regarding membership of input data objects to predefined classes. Each cluster tree was obtained based on a selected linkage algorithm that considers a previously calculated similarity distance between all samples included in the input data set. Because we focus in this study on spectral signatures related to the specific pigment composition, when computing the similarity between pairs of phytoplankton absorption spectra, an angular distance was utilized. This type of distance reflects better the differences in the spectral shape of optical data. The cluster partition obtained based on the pigment data served as a reference for partitioning the entire data set into distinct groups - clusters, each characterized by a different phytoplankton pigment composition (shown in Fig. 1 in the supplemental material). To minimize variability in pigment composition associated with changes in phytoplankton biomass, the input to the cluster analysis was represented by the ratio of individual pigment concentrations to the surface TChl-a. In that sense, the information regarding the dominance of pigments for each station can be better assessed. Otherwise, the analysis would be mainly driven by the amount of total pigments concentration. In addition, a Euclidean distance was utilized to generate the pigment-based cluster partition in order to indicate differences in magnitude of ratios of concentrations of individual pigments to TChl-a rather than differences in shape.

The feasibility of using hyperspectral phytoplankton absorption data for identifying phytoplankton assemblages (see clustering in Fig. 2 in the supplemental material) was tested by comparison of the clustering based on the pigment data using the cophenetic index (see more details in Torrecilla et al., 2011). This index is an objective criterion of cluster similarity ranging from 0 (for no similarity) to 1 (for maximum similarity).

Statistical analysis:

Multiple linear regression models (MLRM) computed with RStudio<sup>TM</sup> were used. The terms were added and removed from a MLRM based on their statistical significance. At each step an F-test was performed to test the regressions with and without certain terms. A term was added to the model if it contributed significantly at the 95%-confidence level or removed from the model if it did not contribute at the 95%-confidence level. To identify the simplest MLRM with the best

explanatory power, each model was compared with the previous one using Analysis of Variances (ANOVA) to test if the latest and simpler model showed no significant differences to the former model. The influence of the interaction of two or more parameters on the sulphur compound concentrations was additionally tested with the same procedure as described above.

Different diagnostic tests were performed to determine if the assumptions made to perform the regression model calculations were valid. All response variables were tested for normal distribution with the Shapiro-Wilk normality test or the Anderson-Darling test (for more than 100 data points) and got transformed if necessary. A tree model was used to obtain an overview of the interactions between the predictor variables prior the calculation of the MLRM. After each calculation, the model was tested for multicollinearity by computing the variance inflation factor (VIF). The heteroscedascity of the models as well as the normal distribution of the residuals were examined graphically. A Durbin-Watson Test was used to find auto-correlations in the residuals. Data points that had a strong influence on the models were identified graphically with the Cook's distance. The model was compared with the next to last simplest model using the Akaike's Information Criterion to check if the simplest model with the best prediction was selected. The entire outputs for all MLRM are given in Table I between the different sulphur compounds and in Table II between sulphur compounds and phytoplankton marker pigments.

Table I: Significant multiple linear regressions between DMS, DMSP and DMSO (d=dissolved, p=particulate, t=total) for the whole data set and within the cluster 2 and 4. Single coefficients, estimates, standard Errors, t and p values of the different independent variables in each multiple linear regression model as well as  $R^2$ , F-statistic and p-value of each whole model are given. Under model number is the response variable given. Variable square showed quadratic relationship to response variable. The complete output of all models is given in the supplements.

model no.	single coefficients	Estimate	Std. Error	t value	p-value	R <sup>2</sup> , F-st., p-value (e. m.)	model no.	single coefficients	Estimate	Std. Error	t value	p-value	R <sup>2</sup> , F-st., p-value (e. m.)	
<b>a</b> DMS e. d. s.	Intercept	-0.20	0.04	-4.71	7.65E-06	0.32	<b>g</b> DMSPP e. d. s.	Intercept	1.49	0.21	7.11	3.92E-10	0.4	
	DMSPP	-0.11	0.02	-6.54	2.36E-09	24.57		DMS	-1.19	0.50	-2.35	0.02	7.92	
	DMSOp	0.02	0.00	5.64	1.49E-07	1.83E-09		DMSOp	0.02	0.01	4.57	1.73E-05	2.57E-07	
<b>b</b> DMS cluster 2	DMSOp	0.02	0.005	3.34	0.002	5.27E-05	DMSOp	0.02	0.01	4.57	1.73E-05	9.09E-03		
	DMSPd	0.54	0.29	1.85	0.07		DMSOp^2	0.53	0.30	1.79	0.08			
	DMSOp^2	-0.02	0.01	-1.82	0.08		DMSPd^2	0.06	0.02	2.67	0.009			
	DMSPd^2	-0.18	0.10	-1.78	0.08		DMSOp^2	0.0005	0.0002	1.96	0.05			
	DMSPP:DMSOp	-0.04	0.01	-3.08	0.004		DMSOp:DMSPd	-0.01	0.004	-2.78	6.78E-03			
							DMSPd:DMSOp	-0.02	0.01	-1.83	0.07			
							<b>h</b> DMSPP cluster 2	Intercept	1.82	0.31	5.94		2.7E-09	0.38
								DMSOp	0.09	0.03	3.47		0.001	10.09
<b>c</b> DMS cluster 4	Intercept	-0.60	0.78	-0.78	0.45	0.58	<b>i</b> DMSOp e.d.s.	Intercept	2.93	0.42	6.91	4.28E-10	0.19	
	DMSPd	-0.56	0.16	-3.53	0.002	5.62		DMSOp	0.72	0.36	2.04	0.04	8.05	
	DMSOp	0.73	0.37	1.99	0.06	0.002		DMSOp^2	1.04	0.25	4.23	5.13E-05	7.26E-05	
	DMSPP	-1.04	0.24	-4.34	0.0003			DMSPd^2	-0.13	0.06	-2.23	0.03		
	DMSOp^2	-0.09	0.04	-2.13	0.05									
DMSPd:DMSPP	0.68	0.19	3.63	0.002										
<b>d</b> DMSPd e. d. s.	Intercept	0.027	0.08	0.32	0.75	0.32	<b>j</b> DMSOp cluster 2	Intercept	2.66	0.82	3.24	0.002	0.28	
	DMSOp^2	0.002	0.0003	5.71	1.14E-07	16.22		DMSPp	0.96	0.42	2.26	0.03	4.82	
	DMSOp^2	0.01	0.003	2.32	0.02	1.084e-08		DMSPd	1.40	0.49	2.84	0.01	0.002	
	DMSPp:DMSOp	-0.01	0.002	-3.72	3.27E-04			DMSPp^2	-0.09	0.04	-2.01	0.05		
<b>e</b> DMSPd Cluster 2	Intercept	0.65	0.18	3.53	9.81E-04	0.47	<b>k</b> DMSOp cluster 4	Intercept	6.31	0.63	10.05	1.91E-10	0.35	
	DMSPp	0.23	0.08	2.99	0.005	5.01		DMSOp	-1.56	0.43	-3.65	0.001	4.59	
	DMSOp	0.06	0.02	2.55	0.014	1.77E-04		DMSOp	-0.26	0.08	-3.14	0.004	0.01	
	DMSOp	-0.04	0.01	-3.12	0.003			DMSPp:DMSOp	0.18	0.05	3.43	0.002		
	DMS	0.38	0.15	2.63	0.01									
	DMSPp^2	-0.015	0.008	-1.99	0.05									
	DMSOp^2	0.003	0.0006	4.094	1.7E-04			<b>l</b> DMSOp e. d. s.	Intercept	1.57	0.10	16.24	2.00E-16	0.43
	DMSPp:DMSOp	-0.004	0.002	-1.69	0.097			DMSPd	0.26	0.05	5.36	5.61E-07	36.53	
	DMSPp:DMSOp:DMS	-0.04	0.013	-3.15	0.0029			DMSPp	0.13	0.02	5.85	6.72E-08	1.49E-12	
<b>f</b> DMSPd cluster 4	Intercept	-7.51	2.3	-3.26	0.004	0.52	<b>m</b> DMSOp cluster 2	Intercept	1.29	2.54	0.51	0.614	0.56	
	DMS	4.5	1.77	2.55	0.019	3.54		DMS	7.39	2.87	2.57	0.014	8.761	
	DMSOp	3.39	0.98	3.46	0.002	0.015		DMSOp^2	0.25	0.14	1.74	0.089	3.33E-06	
	DMSOp^2	-0.24	0.11	-2.25	0.04			DMSPp^2	7.14	1.43	5.01	1.03E-05		
	DMSPp^2	0.4	0.09	4.48	2.3E-04			DMSPp:DMSOp	0.7	0.15	4.55	4.56E-05		
	DMS:DMSOp	-1.28	0.43	-2.94	0.008			DMSPp:DMSPd	-1.6	0.48	-3.28	0.002		
	DMSOp:DMSPP	-0.24	0.05	-4.44	2.5E-04			DMSOp:DMSPd	-2.9	0.84	-3.46	0.0013		
								<b>n</b> DMSOp cluster 4	Intercept	0.40	2.27	0.18	0.86	0.46
						DMS	2.75		1.40	1.97	0.06	7.23		
							DMSPd	3.01	1.46	2.06	0.05	0.001		
							DMSPp	3.09	0.68	4.53	1.26E-04			

Abbr.: st.: statistic; e. m.: entire model; e. d. s.: entire data set; a – n: number of models

Tabel II: Significant multiple linear regressions between DMS, DMSP and DMSO (d=dissolved, p=particulate, t=total) and phytoplankton marker pigments for the whole data set and within the cluster 2 and 4. Single coefficients, estimates, standard Errors, t and p values of the different independent variables in each multiple linear regression model as well as R<sup>2</sup>, F-statistic and p-value of each whole model are given. Under model number is the response variable given. Variable square showed quadratic relationship to response variable.

model no.	single coefficients	Estimate	Std. Error	t value	p-value	R <sup>2</sup> , F-st., p-value (e. m.)	model no.	single coefficients	Estimate	Std. Error	t value	p-value	R <sup>2</sup> , F-st., p-value (e. m.)
<b>a</b> DMS cluster 2	Intercept	0.56	0.21	2.66	0.01	0.32 3.66 0.005	<b>e</b> DMSPp cluster 2	Intercept	-3.68	0.95	-3.87	4.41E-04	0.73
	fuco	-0.18	0.06	-3.04	0.004			fuco	0.48	0.12	4.09	2.32E-04	11.94
	hex	0.05	0.02	2.59	0.01			diato	-0.64	0.20	-3.26	2.46E-03	4.02E-08
	peri^2	0.01	0.002	3.09	0.003			zea	0.08	0.02	3.80	5.32E-04	
	fuco^2	0.01	0.004	2.98	0.005			hex	-0.02	0.01	-2.19	3.51E-02	
	diato^2	-0.02	0.01	-2.07	0.04			fuco^2	-0.03	0.01	-3.49	1.28E-03	
	hex^2	-0.001	0.0003	-2.63	0.01			diato^2	0.31	0.08	4.19	1.75E-04	
<b>b</b> DMSPd e. d. s.	Intercept	0.11	0.14	0.77	0.44	0.44 11.34 2.36E-09	<b>f</b> DMSPt e. d. s.	Intercept	-5.66	1.17	-4.84	6.10E-06	0.75
	but	0.13	0.03	4.69	1.01E-05			dia	0.29	0.12	2.50	1.46E-02	16.2
	peri	0.12	0.03	3.77	2.96E-04			fuco	1.10	0.16	6.68	2.65E-09	2.20E-16
	zea	-0.01	0.002	-5.09	2.01E-06			zea	0.13	0.02	5.35	7.70E-07	
	but^2	-0.01	0.002	-4.74	8.21E-06			peri^2	0.05	0.02	3.28	1.53E-03	
	but:peri:fuco	-8.32E-04	3.21E-04	-2.59	0.01			dia^2	-0.02	0.01	-2.79	6.59E-03	
	but:zea:fuco	7.31E-05	1.78E-05	4.11	8.82E-05			fuco^2	-0.03	0.01	-4.17	7.59E-05	
<b>c</b> DMSPd cluster 2	Intercept	0.89	0.07	13.47	2.00E-16	0.61 5.93 1.15E-05	zea^2	-0.0004	0.0001	-3.57	6.10E-04		
	fuco	0.04	0.02	2.58	0.01		viola^2	-0.25	0.06	-4.33	4.22E-05		
	diato	0.26	0.10	2.60	0.01		fuco:peri	-0.09	0.01	-6.21	2.07E-08		
	but^2	-0.002	0.001	-3.32	1.91E-03		dia:fuco	0.02	0.01	1.90	6.16E-02		
	fuco:peri	-0.01	0.005	-2.46	0.02		fuco:zea	-0.01	0.002	-3.12	2.47E-03		
	diato:peri	0.08	0.02	4.35	8.83E-05		dia:viola	-0.18	0.05	-3.42	9.70E-04		
	fuco:diato	-0.06	0.01	-4.29	1.07E-04		fuco:viola	0.20	0.05	3.63	4.94E-04		
	diato:zea	-0.01	0.001	-5.40	3.04E-06		zea:viola	-0.02	0.004	-4.58	1.62E-05		
	peri:but	0.01	0.004	3.87	3.79E-04		dia:zea:viola	0.002	0.0003	6.18	2.38E-08		
	diato:but	0.03	0.008	3.99	2.63E-04		<b>g</b> DMSPt cluster 2	Intercept	-1.34	0.78	-1.72	0.09	0.78
	fuco:diato:zea	0.001	0.0001	5.02	1.04E-05			hex	-0.18	0.05	-3.80	4.46E-04	16.98
	diato:peri:but	-0.01	0.001	-5.82	7.87E-07			but	0.99	0.36	2.79	7.84E-03	1.72E-11
	<b>d</b> DMSPp e. d. s.	Intercept	0.28	0.13	2.08			0.04	0.37 9.3 5.01E-08	fuco	1.28	0.32	3.96
peri		0.10	0.04	2.63	9.88E-03	peri^2		0.07		0.02	3.14	3.08E-03	
but		0.09	0.02	4.08	9.23E-05	but^2		-0.02		0.01	-2.10	0.04	
fuco		0.05	0.02	2.01	0.05	fuco^2		-0.07		0.02	-3.80	4.49E-04	
fuco^2		-0.001	0.0002	-4.69	8.97E-06	ant^2	-0.67	0.24		-2.79	7.75E-03		
peri:but		-0.02	0.004	-5.33	6.41E-07	fuco:peri	-0.08	0.02		-4.74	2.35E-05		
peri:but:fuco		0.0002	0.00004	5.77	9.87E-08	fuco:ant	0.26	0.06		4.03	2.22E-04		

model no.	single coefficients	Estimate	Std. Error	t value	p-value	R <sup>2</sup> , F-st., p-value (e. m.)	model no.	single coefficients	Estimate	Std. Error	t value	p-value	R <sup>2</sup> , F-st., p-value (e. m.)
<b>h</b>	Intercept	0.83	0.49	1.67	0.11	0.36	<b>l</b>	Intercept	3.93	3.60	1.09	0.28	0.84
DMSPt cluster 4	but	0.54	0.13	17.1	4.14	17.1		DMSOp cluster 2	peri	-4.72	1.66	-2.84	7.63E-03
						0.0003		diato	-5.37	1.71	-3.14	3.56E-03	1.93E-09
<b>i</b>	Intercept	4.04	0.20	19.76	2.00E-16	0.42		but	1.76	0.81	2.16	0.04	
DMSOd	diato	0.49	0.22	2.20	0.03	7.55		peri^2	0.33	0.07	4.52	7.65E-05	
e. d. s.	hex^2	0.01	0.002	4.06	1.11E-04	1.65E-07		but^2	-0.12	0.05	-2.67	0.01	
	but^2	-0.08	0.02	-4.10	9.68E-05			peri:fuco	0.61	0.22	2.76	9.41E-03	
	fuco:peri	-0.09	0.04	-2.24	0.03			peri:diato	1.37	0.59	2.35	0.03	
	fuco:hex	-0.05	0.01	-4.04	1.20E-04			diato:fuco	0.87	0.17	4.99	1.88E-05	
	fuco:but	0.19	0.05	3.90	1.96E-04			peri:zea	0.08	0.02	4.93	2.28E-05	
	peri:but	0.10	0.04	2.64	9.95E-03			peri:but	-0.94	0.24	-4.00	3.37E-04	
	diato:but	-0.07	0.03	-2.80	6.43E-03			peri:diato:fuco	-0.22	0.07	-3.07	4.27E-03	
								peri:fuco:zea	-0.01	0.002	-4.24	1.71E-04	
<b>j</b>	Intercept	2.67	0.91	2.94	5.01E-03	0.45	<b>m</b>	peri:but:fuco	0.10	0.03	3.85	5.18E-04	
DMSOd cluster 2	peri	0.33	0.07	4.86	1.24E-05	10.1		DMSOt cluster 2	Intercept	1.32	4.41	0.30	0.77
	dia	0.36	0.17	2.07	4.38E-02	4.81E-06		but	-1.28	0.36	-3.54	9.55E-04	9.82
	but	-0.11	0.04	-2.74	8.61E-03			hex	0.69	0.32	2.14	0.04	2.15E-06
	dia^2	-0.02	0.01	-2.41	1.98E-02			fuco	1.53	0.37	4.12	1.61E-04	
								peri^2	0.14	0.05	2.84	6.69E-03	
<b>k</b>	Intercept	0.33	0.45	0.73	0.468082	0.54	<b>n</b>	hex^2	-0.01	0.004	-2.09	0.04	
DMSOp	Fuco	0.51	0.11	4.83	6.83E-06	9.18		DMSOt cluster 4	Intercept	1.94	0.92	2.12	0.04
e. d. s.	Diato	-0.82	0.24	-3.39	1.09E-03	8.46E-10		hex	0.30	0.13	2.34	0.03	4.65
	Zea	0.03	0.01	5.18	1.76E-06			zea	-0.04	0.01	-3.22	3.54E-03	0.006
	peri^2	0.05	0.01	4.62	1.50E-05			hex^2	-0.01	0.005	-2.21	0.04	
	fuco:diato	0.11	0.04	3.18	2.15E-03			zea^2	0.0003	0.0001	3.41	2.23E-03	
	fuco:zea	-0.01	0.002	-5.08	2.54E-06								
	diato:zea	0.01	0.003	2.92	4.59E-03								
	peri:diadino	-0.04	0.01	-4.78	8.16E-06								
	fuco:diato:diadino	-0.01	0.001	-3.88	2.21E-04								
	fuco:zea:diadino	0.0002	0.0001	4.31	4.87E-05								

Abbr.: st.: statistic; e. m.: entire model; e. d. s.: entire data set; a – n: number of models, fuco: fucoxanthin, hex: 19'-hexanoyloxyfucoxanthin, peri: peridinin, diato: diatoxanthin, dia: diadinoxanthin, diato: diatoxanthin, but: 19'-butanoyloxyfucoxanthin, zea: zeaxanthin, ant: anthreaxanthin, viola: violaxanthin

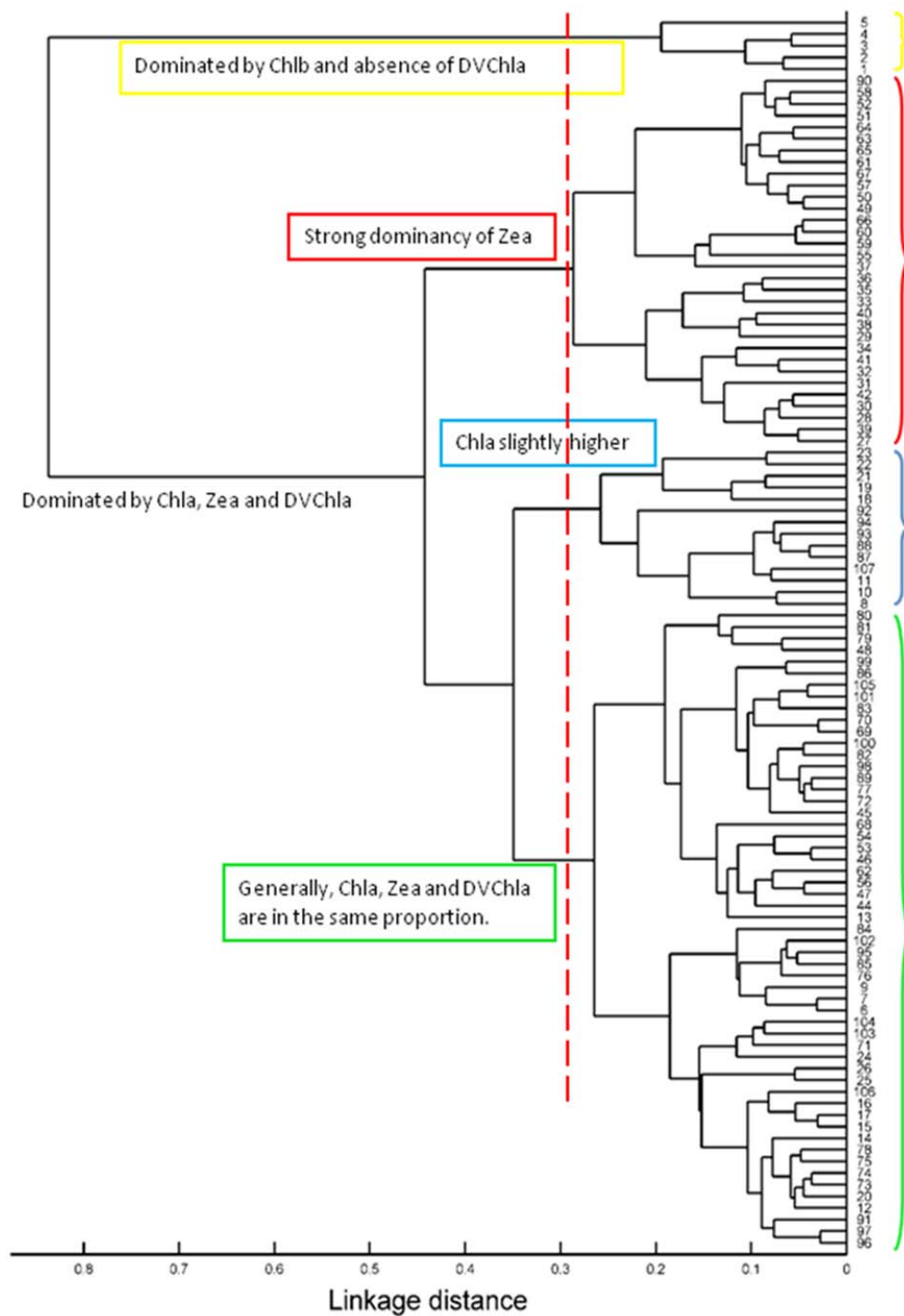


Figure 1: HCA cluster tree based on pigment data obtained during TransBrom Sonne (stations are labeled with consecutive numbers increasing with latitude). Different clusters are identified with different colors: cluster 1 (yellow), cluster 2 (green), cluster 3 (blue) and cluster 4 (red). It is noted that clusters 2, 3 and 4 have as dominant pigments Chla, Zea and Diva. Nevertheless, for cluster 4 Zea contributes the largest, for cluster 3 Chla and for cluster 2 all pigments equally.

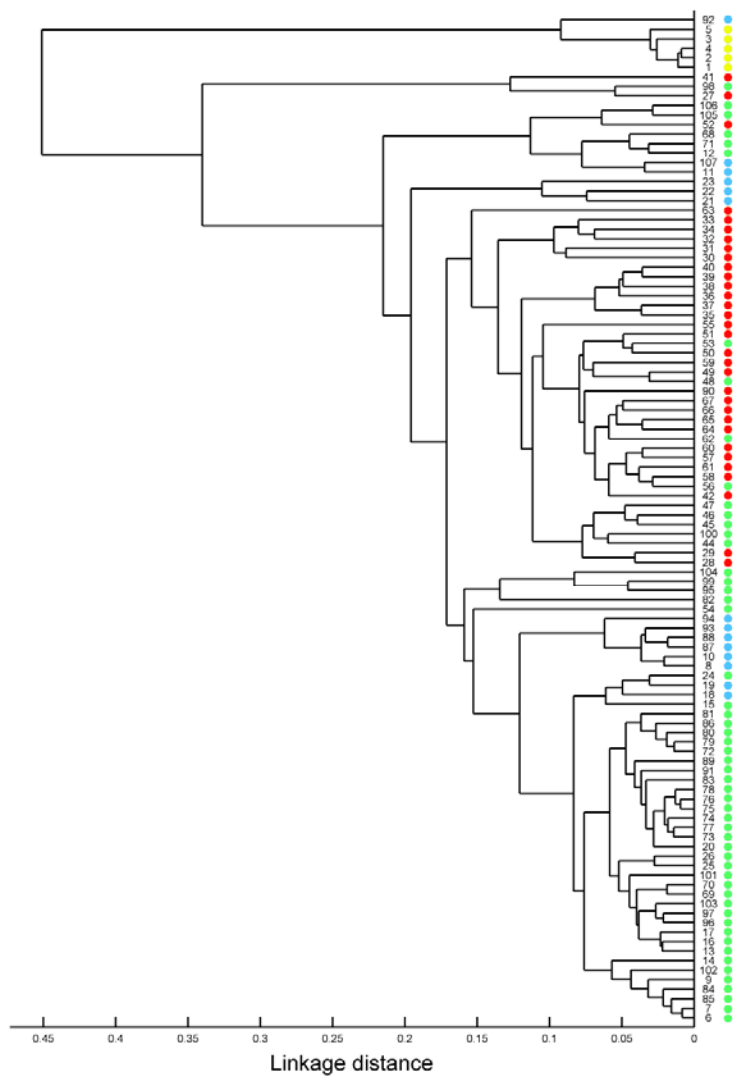


Figure 2: HCA cluster tree based on hyperspectral phytoplankton absorption coefficients measured at different stations during TransBrom Sonne. Stations are labeled with consecutive numbers increasing with latitude and identified with different colors in accordance with the pigment-based clusters shown in Fig. 1.