IDENTIFICATION AND QUANTIFICATION OF PHYTOPLANKTON BY IMAGE ANALYSIS

Karsten Rodenacker, Peter Gais[1]*, Uta Jütting and Burkhard A. Hense*

GSF-National Research Centre for Environment and Health,
Institute of Biomathematic and Biometry, [1]Institute of Pathology,
85764 Neuherberg, Germany
Email: [redacted] karo13de@googlemail.com

1. INTRODUCTION

For a number of questions in aquatic environment it is essential to obtain detailed information about the qualitative and/or quantitative structure of the biocenosis. These comprise purely ecological aspects as well as the estimation of the impact of anthropogenic stressors like supply of nutrients or xenobiotica or acidification. Therefore, e.g. longtime monitoring of biocenosis in natural waters or investigations on biocenosis in artificial, selective affected waters are conducted. The latter allows to control the interesting parameters, and because of that to receive more reproducible information about anthropogenic effects on ecosystems.

As ecosystems normally reveal a functional redundancy it is often necessary to get information on as detailed taxonomic level as possible to make sure to detect potential changes. This is essential to be able to protect relevant structural features of the ecosystem. Microorganisms (organisms too small to be seen without microscopes) represent by far the main part of the numbers of individuals in aquatic ecosystems and, despite their small sizes, provide a major part of the total biomass [1]. Among this the phytoplankton holds a key position as an important, in many waters the most important primary producer. For a comprehensive assessment of the biocenosis or of effects on it it is therefore indispensable to consider the phytoplankton, which is consequently a part of a huge number of studies dealing with e.g. water quality or judgement of chemicals. However, identification and quantification of the microorganism cells are very time consuming and require concentrated work of a well qualified person making it expensive.

Therefore an at least partial automation of it saving money and increasing efficiency is desirable. The developmental state of a semi-automatic recognition system of algal cells by image analysis is presented. It is meant as a step towards semi-automated quantification of at least some of the occurring species (e.g. the most numerous or difficult to quantify).

Recently, interest focuses on the impact of endocrine disruptors, i.e. environmental chemicals that interfere with the endocrine (hormonal) systems of organisms. As there are a number of studies dealing with the effects on single species, but very few on entire ecosystems, aquatic microcosm studies have been conducted with different endocrine disruptors within

the cross-sectional topic *ecotoxicology* in the GSF. The samples used for the presented paper are part of the study with ethinylestradiol.

2. MATERIAL

12 cylindrical stainless steel microcosms (80 cm high, 60 cm wide) were filled with littoral water (230 l) and sediment (10 cm) of an oligo-mesotrophic natural lake (Ammersee, Bavaria). After five weeks of relaxation six microcosms were continuously exposed to different concentrations of ethinylestradiol for six weeks. Six microcosms served as controls. Samples were taken once a week before, during and for six weeks after exposure, giving a total number of 210 samples. The samples were fixed with about 12 drops Lugol per 100 ml and sedimented in a plankton chamber using the Utermöhl method [2]. The chambers were analyzed with an inverse microscope. For more details see Hense et al. [3]. A short description of the microorganisms to be detected is given in Tab. 3 and Fig. 6.

3. METHODS

3.1. Acquisition

For scanning the plankton chambers, images were taken following two stripes, crossing vertical in the middle of the chamber (Fig. 1 left). Other acquisition walks are tested and possible like meander, spiral and random walk to obtain reliable representation of the distribution of organisms. The images were taken with small overlapping areas of neighboring sites, so that the complete stripes were imaged, allowing to determine almost the same area as when calculating the chambers directly with the microscope as usual till now, i.e. without acquiring images. Bright-field microscopy was used. Two magnifications were chosen to be able to analyze smaller cells as well as the larger, mostly more seldom ones statistically.

Image acquisition was done with an automatic inverse microscope[1] DM IBRE and a 3-CCD colour TV camera[2] KY-F58 . For the two magnifications the objectives 20x N-Plan, n.a. 0.4 and the 40x N-Plan, n.a. 0.55 were used. The objective change as well as the automatic focus, the lamp voltage and the stage are controlled by a self-developed procedure, running under the Leica software QWIN. The digitized images were stored on the computer disk each with a size of 760 x 576 pixel. The pixel size results for the two magnifications to 0.316 μm and 0.632 μm respectively. One data set comprises 270 images from 40x objective and 136 from 20x objective (see Fig. 1).

3.2. Image analysis

Image analysis is performed by means of IDL[3] which is a platform independent data analysis tool with broad range of mathematical and graphical capabilities. A graphical user interface (GUI) has been designed for all analysis, display and supervision tasks.

[1]Leica, Bensheim, Germany
[2]JVC, Tokyo, Nippon
[3]Research System Inc., Boulder, USA

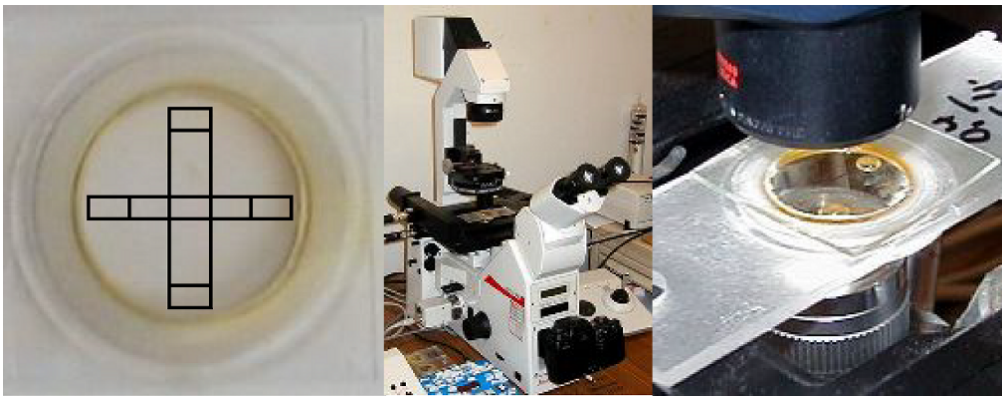

Fig. 1. Plankton chamber with sketched image gathering field, microscope and scanning table with plankton chamber

Image quality

Automatic acquisition necessitates a rejection scheme for bad images. This concerns mainly the choice of the field of view, the adjustment of focal position, possible movement artifacts during image acquisition. As a first step from every image a set of global features including a segmentation threshold is calculated (Tab. 1), which is then subject for quality evaluation and further analysis. The threshold value is the position of mode of the transmission intensity histogram reduced by the right hand variance multiplied by an adjustable appropriate factor.

Segmentation

For further analysis of an image organisms or more general objects have to be *segmented*. Each point (pixel) in the image gets marked to be background or object. The segmentation steps are illustrated in Fig. 2. First a threshold is applied to the luminance image, resulting in the marked image. It is cleaned and slightly smoothed by morphological operations [4] to preserve the original object shapes. Finally holes are filled and objects with area less than a threshold are deleted. The final mask image is labelled for object identification. For later analysis and evaluation for each object the list of contour points is stored.

Featuring

From each detected and identified object a set of features (descriptive numbers) is calculated [5] (listed in Tab. 2). These features comprise of mostly morphological and some intensity (calculated from extinction values) [6] features. As an example some features are illustrated with a *Cryptomonas marsonii* in Fig. 3. Structure or texture features are implemented up to now only in a limited degree. A relative large set of quantitative features used for cyto- and histometry is described in the internet[4].

Classification

To relate found objects to taxa names a tree-based classification system is used. A threefold strategy is applied. First the verbal description is used to design a simple so-called *interactive* box classification scheme (Each object inside this box gets one name). Parallel to this with

[4] http://www.gsf.de/ILIAD/DIC/Features.html

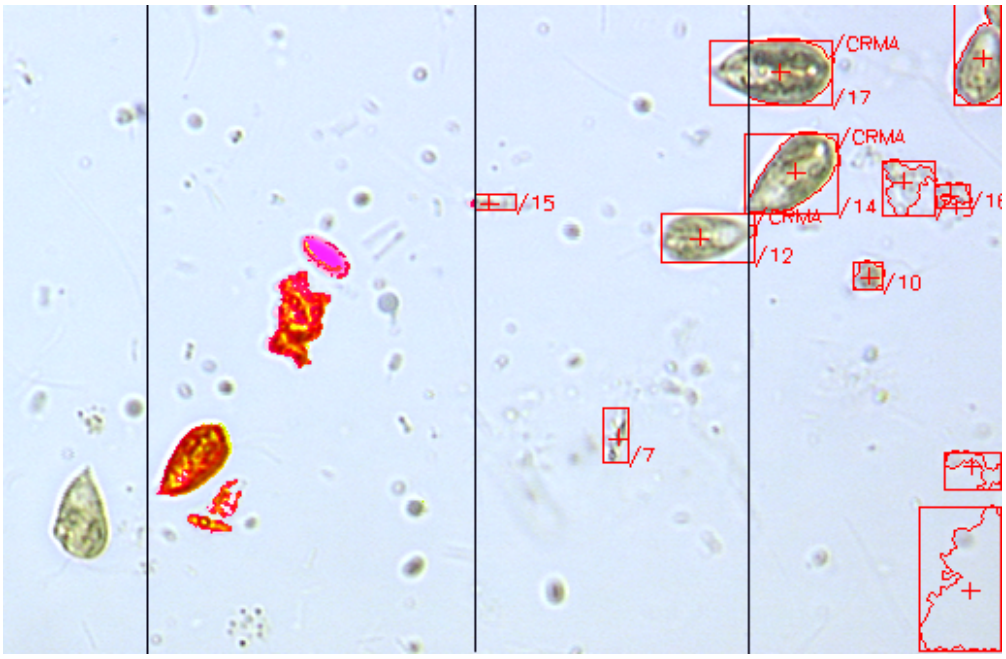

Fig. 2. Segmentation of one image, central field with masks, right field with marked featured objects

stepwise linear discrimination analysis of the feature data a linear hierarchical classification scheme is designed based on proven results of the interactive classification scheme and on interactively selected and controlled training data. The first step of the classification scheme is shown in Fig. 5. Finally we plan to design a neuronal net based classification scheme on the base of the previous results. This strategy is used to reduce the necessary time for design of training sets and to reach a highly adaptable, extensible and improvable scheme for biologists.

In Fig. 6 and Tab. 3 the list of actually examined organisms is outlined. The classification focuses on algae species occurring in most of the microcosms studies of the GSF in relevant numbers. Additionally some data used for the interactive classification scheme is listed.

4. RESULTS

Automatic specimen acquisition

The software controlling the microscope developed in our lab till now enables the automated acquisition of an image set allowing a representative evaluation of the phytoplankton via image analysis.

The microscope soft- and hardware system available allows following the specifications a wide range of functionality. However the details deliver the problems. Especially the automatic focusing system has to be improved. Adding modes of image acquisition e.g. fluorescent acquisition mode will render more difficulties beside the one of moving organisms.

Automatic image analysis

The segmentation of the objects is visually sufficient, organisms found would have been counted visually. However actually the segmentation is done by only one global threshold per image. This might have to be changed by breaking the segmentation up into a detection

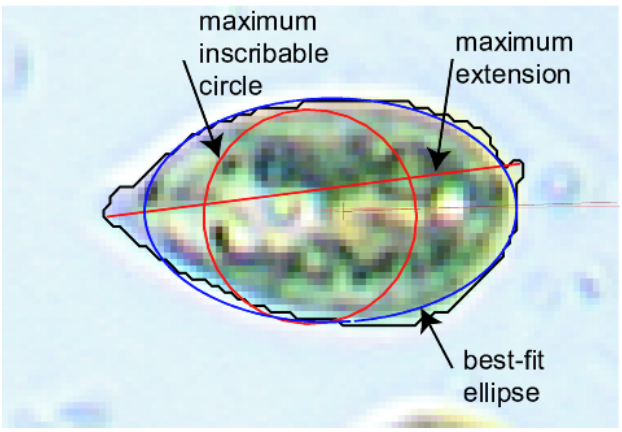

Fig. 3. *Cryptomonas marsonii* with morphological features

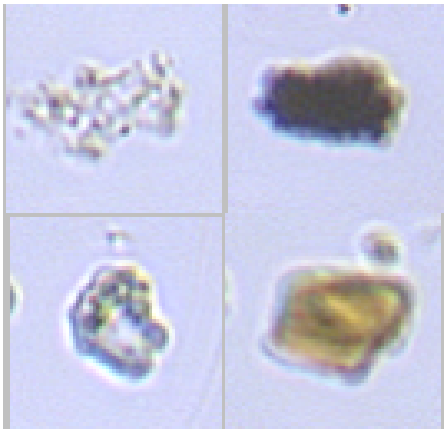

Fig. 4. Some artifacts

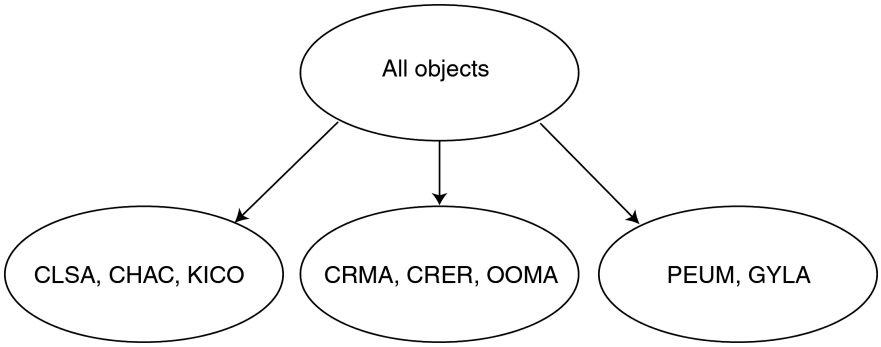

Fig. 5. First level hierarchical classifier with mnemonic taxa names (see Tab. 3)

step and the object segmentation. Also touching and occluding objects are up to now not processed. The manifold shapes existent do not allow general cutting rules.

Classification

A first estimate of the recognition functionality presents a distinctive result for the species. Seven sample sets were analyzed. We compared the manual counting results with the results from the automatic procedure. To minimize statistical variations only organisms with at least 20 occurrences in at least one of the counting procedures were used.

The deviation of *Cryptomonas marsonii* (-15 %) is rather low and can mainly be explained by remaining problems concerning the automated focus of the microscope, overlapping objects, organisms at the edges of the chamber or the pictures. *Oocystis marsonii* shows a higher, but relatively constant reduction (-65 %). Possibly the thresholds of one or more features are set too stringent, excluding a certain, constant portion of the organisms. The value for *Chroomonas acutae* (5-times higher) seems to be dissatisfying. Problems are caused primarily by detritus being false classified as *Chroomonas acutae* and by objects cut by the image border. Nevertheless, in samples with a high number of this species, the respective fault is considerably smaller. Even more difficult to analyze seems *Chlorococcus spec*, a small organism which just characterized by its round, unspecific shape and size. Detritus, small drops of oil coming from destroyed organisms and insufficiently cleaned slides complicate the correct recognition of this species.

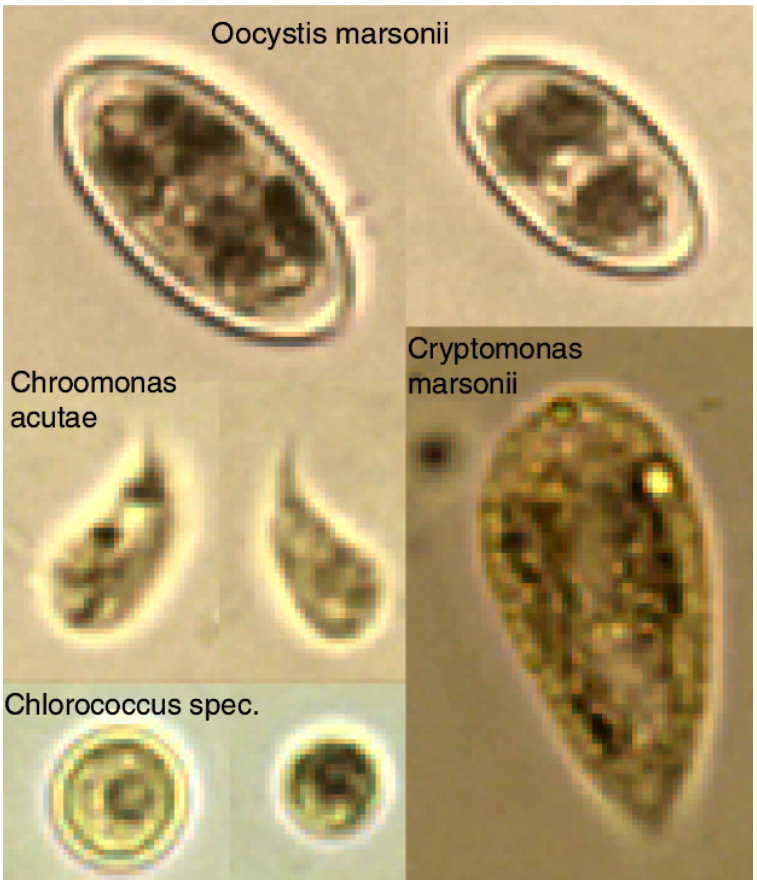

Fig. 6. Some microorganisms to classify

The amount of artifacts that are difficult to reject is reasonable (see Fig. 4). This results from occlusion, overlap, out of focus and touching objects beside detritus of organisms in various states. However the frequencies of several automatically recognized species correspond with manually counted data very well and allow even in this preliminary state of development a considerable reduction of manual counting time.

5. DISCUSSION AND OUTLOOK

An image processing system is presented for semi-automatic counting of microorganisms. The chosen strategy of separated steps of image gathering and image analysis results from the limited resources of highly automated microscopes and the relatively cheap computing power available. This allows the full-time use of the image gathering device without disturbance by possible interactions during and after the image analysis process. This is of course necessary for constant extension and improvement of the programs.

As far as the system is developed until now, every step of the whole process has reduced the amount of human interaction time, the detection (segmentation) for reduction of the time of visual orientation, the selection of the well recognized objects to allow the user to focus on the rare and possibly more difficult events.

The image processing system is at the current status able to identify a high percentage of the investigated species, depending on species structure and size. However, its ability to rejects false objects has to be improved, especially for small cells. Nevertheless, it enables even now a automatically preliminary selection of optical structures and thus can accelerate

identification and quantification of smaller species.

At the moment, different aspects of the system are developed. For a further improvement of the classification the up to now ignored intensity or density features will be used. Additionally dominant contour features [8] are implemented. Also consideration of additional relevant species is done to improve classification.

Concerning image acquisition, beside an optimization of achieving statistically usable image sets of the samples, an integration of other optical technics like fluorescence microscopy will provide more information to distinguish different algal cells from zooplankton and detritus.

The use of image analysis for identification and quantification of plankton samples has shown a high potential for easying and speeding up and thus lowering the costs of environmental investigations, although the diversity of the species and the structure of the samples received from aquatic ecosystems make great demands on image analysis. In future, the presented image processing system will be validated both in the microcosms project of the GSF and in external cooperations

Acknowledgement: We thank W. Jaser and Dr. G.F. Severin, Institute of Ecological Chemistry, GSF, for preparing the phytoplankton samples, and Dr. K.-W. Schramm, chairman of the crosssectional topic *ecotoxicology*, within the scope of which the ethinylestradiol study was conducted.

6. REFERENCES

[1] G. P. Harris, *Phytoplankton ecology*, Chapman and Hall, London, 1986.

[2] H. Utermöhl, "Zur Vervollkommnung der quantitativen Phytoplanktonmethodik," *Mitt.Int.Ver.Limnol.*, vol. 9, pp. 1–38, 1958.

[3] B. A. Hense, I. Jüttner, G. Severin, G. Pfister, G. Welzl, K.-W. Schramm, and A. Kettrup, "Investigations on primary production in microcosms continuously exposed to the endocrine disruptor nonylphenol," *Organohalogen compounds*, vol. 49, pp. 454–456, 2000.

[4] Jean Serra, *Image analysis and mathematical morphology – Volume 2: Theoretical Advances*, vol. two, Academic Press Inc., London u.a., 1988.

[5] Karsten Rodenacker, Paul Bischoff, and Bidyut B. Chaudhuri, "Featuring of toplogical characteristics in digital images," *Acta Stereol*, vol. 6, no. Suppl III, pp. 945–950, 1987.

[6] Georg Burger, Michaela Aubele, B. Clasen, Uta Jütting, Peter Gais, and Karsten Rodenacker, "Malignancy associated changes in squamous epithelium of the head and neck region," *Anal. Cell. Pathol.*, vol. 7, pp. 181–193, 1994.

[7] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. IT*, vol. 8, pp. 179–187, 1962.

[8] Michaela Aubele, Uta Jütting, Karsten Rodenacker, Peter Gais, Georg Burger, and U. Hacker-Klom, "Quantitative evaluation of radiation-induced changes in sperm morphology and chromatin distribution," *Cytometry*, vol. 11, pp. 586–594, 1990.

Table 1. Global image features

0	Bild	Image number
1	Thr	Threshold
2	TotalA	Total area below threshold
3	Number	Total number of objects
4	MaxA	Maximum object area
5	Class	Classification
7	MeanG	Mean of gradient intensity image
8	VarianceG	Variance
9	SkewnessG	Skewness
10	KurtosisG	Kurtosis
12	ModeI	Mode of intensity image
13	SDI	SD
14	LowerBI	Minimum
15	UpperBI	Maximum

Table 2. Object features

0	Bild	Image number
1	Obj	Object number
2	Type	Object type
3	Im_1-Im_7	Invariant moments [7] of grey scale image
10	M_{00}	Total extinction
11	(K_x,K_y)	Centroid coordinates of grey scale image
13	ϑ	Angle of main axis of grey scale image
15	(KM_x,KM_y)	Centroid coordinates of mask
17	ϑ_M	Angle of main axis of mask
18	γ	Multiplicative correction term of inv. moments
19	α	Additive correction term
20	ImM_1-ImM_7	Invariant moments [7] of the mask image
27	X_{min},X_{max}	Box coordinates in x
29	Y_{min},Y_{max}	and y
31	M_1-M_4	Mean extinction, SD, Skewness, Kurtosis
35	A	Area
36	P	Perimeter
37	P^2A	Shape factor
38	mx_Ext	Maximal extension
39	(k_{x1},k_{y1})	Coordinates of contour points
41	(k_{x2},k_{y2})	of maximal extension
43	rad	Radius of smallest inscribable circle
44	(rk_x,rk_y)	Coordinates of smallest inscribable circle
46	$spec_{(-6,...,6)}$	Some frequencies of Fourier spectrum
		. of parameterized contour
52	np	Length of spectrum

Table 3. List of microorganisms with their mnemonic abbreviations and some size and shape criteria

CRMA	*cryptomonas marsonii*	similar to *CHAC* but larger
	$96\mu m^2 < A < 250\mu m^2$	$1.3 \leq P^2A \leq 1.7$
	$0.1 \leq M_1 \leq 0.25$	$0.22 \leq rad/mx\_Ext \leq 0.32$
	$16\mu m \leq mx\_Ext \leq 26\mu m$	$4.8\mu m \leq rad \leq 6.8\mu m$
CRER	*cryptomonas erosa*	?
	$192\mu m^2 < A < 384\mu m^2$	$1.2 \leq P^2A \leq 1.7$
	$0.18 \leq M_1 \leq 0.25$	$0.26 \leq rad/mx\_Ext \leq 0.35$
	$22\mu m \leq mx\_Ext \leq 30\mu m$	$6\mu m \leq rad \leq 8\mu m$
CLSA	*chlorococcus spec 1*	nearly exact spherical
	$16\mu m^2 < A < 64\mu m^2$	$1. \leq P^2A \leq 1.15$
	$0.05 \leq M_1 \leq 0.3$	$0.4 \leq rad/mx\_Ext \leq 0.5$
	$4\mu m \leq mx\_Ext \leq 20\mu m$	$2\mu m \leq rad \leq 10\mu m$
CHAC	*chroomonas acutae*	sim. mouse sperm heads[8]
	$28\mu m^2 < A < 64\mu m^2$	$1.3 \leq P^2A \leq 2.5$
	$0.1 \leq M_1 \leq 0.25$	$0.2 \leq rad/mx\_Ext \leq 0.3$
	$8\mu m \leq mx\_Ext \leq 14\mu m$	$2\mu m \leq rad \leq 4\mu m$
OOMA	*oocystis marsonii*	elliptic, regular shape, smooth border
	$80\mu m^2 < A < 800\mu m^2$	$P^2A \leq 1.5$
	$0.1 \leq M_1 \leq 0.3$	$0.28 \leq rad/mx\_Ext \leq 0.4$
	$12\mu m \leq mx\_Ext \leq 44\mu m$	$4\mu m \leq rad \leq 14\mu m$
GYLA	*gymnodinium lacustris*	elliptic spherical
	$160\mu m^2 < A < 720\mu m^2$	$P^2A \leq 1.8$
	$0.18 \leq M_1 \leq 0.35$	$0.36 \leq rad/mx\_Ext \leq 0.5$
	$16\mu m \leq mx\_Ext \leq 32\mu m$	$8\mu m \leq rad \leq 14\mu m$
PEUM	*peridinium umbonatum*	similar *GYLA* but smaller
	$200\mu m^2 < A < 720\mu m^2$	$P^2A \leq 2$
	$0.18 \leq M_1 \leq 0.38$	$0.35 \leq rad/mx\_Ext \leq 0.5$
	$18\mu m \leq mx\_Ext \leq 60\mu m$	$7.5\mu m \leq rad \leq 14\mu m$
KICO	*kirchneriella contorta*	small, like half moon
	$8\mu m^2 < A < 32\mu m^2$	$12.3 \leq P^2A \leq 3.8$
	$0.08 \leq M_1 \leq 0.2$	$0.15 \leq rad/mx\_Ext \leq 0.3$
	$4.4\mu m \leq mx\_Ext \leq 9.2\mu m$	$0.8\mu m \leq rad \leq 1.8\mu m$