4

Natural hybrids in *Pseudo-nitzschia pungens*: genetic and morphological evidence[1]

Hybridization between genetically distinguishable taxa provides opportunities for investigating speciation. While hybridization is a common phenomenon in various macro-organisms, natural hybridization among micro-eukaryotes is barely studied. Here we used a nuclear and a chloroplast molecular marker and morphology to demonstrate the presence of natural hybrids between two genetically and morphologically distinct varieties of the marine planktonic diatom *Pseudo-nitzschia pungens* (vars. *pungens* and *cingulata*) in a contact zone in the northeast Pacific. Cloning and sequencing of the rDNA internal transcribed spacer region revealed strains containing ribotypes from both varieties, indicating hybridization. Both varieties were found to also have different chloroplast encoded *rbc*L sequences. Hybrid strains were either hetero- or homoplastidial, as demonstrated by Denaturing Gradient Gel Electrophoresis, which is in accordance with expectations based on the mode of chloroplast inheritance in *Pseudo-nitzschia*. While most hybrids are probably first generation, there are also indications for further hybridization. Morphologically, the hybrids resembled var. *pungens* for most characters rather than having an intermediate morphology. Further research should focus on the hybridization frequency, by assessing the spatial and temporal extent of the contact zone, and hybrid fitness, to determine the amount of gene flow between the two varieties and its evolutionary consequences.

---

[1] Published as: Casteleyn G, Adams NG, Vanormelingen P, Debeer AE, Sabbe K ,Vyverman W (2009) Natural hybrids in the marine diatom *Pseudo-nitzschia pungens* (Bacillariophyceae): genetic and morphological evidence. Protist, doi:10.1016/j.protis.2008.11.002

Introduction

Hybridization is generally defined as the crossing of genetically distinguishable taxa leading to viable individuals of mixed ancestry (Mallet 2005). Hybridization plays an important role in creating new diversity either by introgression of single alleles, or by the establishment of recombinant genotypes as a new species (Barton 2001). Regions where divergent populations meet and interbreed are referred to as hybrid zones. They are useful in studying the process of speciation as they provide natural examples of, through their recognizable and limited geneflow, the continuum between a single species with multiple populations and multiple reproductively isolated species. Hybridization events have traditionally been studied through morphological comparisons, under the assumption that various hybrid offspring are predictably intermediate to parents (Arnold 1997). However heterospecific crosses sometimes produce offspring that more closely resemble the phenotype of one of the parental taxa (Byrne & Anderson 1994, Harper & Hart 2007, Lamb & Avise 1987). In these cases genetic markers independent of morphological traits can often reveal allele or haplotype sharing that is consistent with hybridization and introgression. Therefore, when studying natural hybridization, joint comparisons of morphology and genotypic constitution are recommended (Lamb and Avise 1987).

The importance of hybridization and introgression has long been recognized in plants (reviewed in Arnold 1997) but only relatively recently in animals (Mallet 2005). While the majority of hybridization examples are from terrestrial species, Gardner (1997) suggested that hybridization might be a common phenomenon in marine environments as well. Well-studied examples of hybridization among marine organisms are those of the blue mussel *Mytilus edulis* species complex (Bierne et al. 2003), the brown seaweed *Fucus* (Coyer et al. 2007) and various corals (Willis et al. 2006, Slattery et al. 2008). In contrast, there are hardly any studies investigating natural hybridization in aquatic micro-eukaryotes. One reason is the problematic delineation of species, often based solely on morphology, which leads to hybrids being undetected. Studies on natural hybridization are also hampered by difficulties in the experimental manipulation of microscopic organisms together with the need to establish clonal cultures if one wants to study different aspects of a clone (e.g. different molecular markers and morphological features). There are indications for natural inter-specific hybridizations in dinoflagellates based on rDNA markers but the results are not decisive (Edvardsen et al. 2003, Hart et al. 2007). In diatoms, one of the most speciose taxa of aquatic protists, there are a number of studies on experimental hybridization (overview given by Vanormelingen et al. 2008), but only a single recent study addressed intraspecific hybridization in the field (D'Alelio et al. 2009b).

The pennate diatom *Pseudo-nitzschia pungens* is a common component of marine phytoplankton assemblages worldwide (Hasle 2002). In a recent study, at least three distinct but closely related genetic entities were found within *P. pungens* based on nrDNA internal transcribed spacers (ITS) sequences (Casteleyn et al. 2008, chapter 3 of this thesis). These three ITS clades showed slight but consistent morphological differences. Strains of clade I (referred to hereafter as ribotype I) were found to correspond morphologically to *P. pungens* var. *pungens* while strains of clade II (ribotype II) fit *P. pungens* var. *cingulata* (Casteleyn et al. 2008). Recently, members of the third clade were described as a new variety, var. *aveirensis* (Churro et al. in press, appendix of this thesis). *P. pungens* var. *pungens* is distributed in temperate waters worldwide while var. *cingulata* seems to be restricted to the NE Pacific Ocean where it co-exists with the former variety. Both varieties are sexually compatible in the laboratory and produce viable offspring. Despite the fact that apparently few or no intrinsic prezygotic barriers between both varieties exist, no signs of mixtures of ribotypes or strains with intermediate morphologies in nature have previously been recovered from the NE Pacific (Casteleyn et al. 2008), possibly because the number of strains studied was too low. Recently, a microsatellite study using a more extensive set of *P. pungens* strains from the Washington coast revealed two populations of *P. pungens* (Adams 2006). Interestingly, some strains with an intermediate probability of belonging to one or the other population were recovered.

In the present study, we demonstrate natural hybridization between *Pseudo-nitzschia pungens* var. *pungens* and var. *cingulata* by analysing molecular markers in two different genomic compartments (nucleus and chloroplast) and examining morphological features for a subset of strains used by Adams (2006). We first investigated whether the two populations found by Adams (2006) correspond to the two *P. pungens* varieties and if natural hybridization between the two varieties does indeed occur in nature. We used the ITS region as a diagnostic marker to identify the two varieties. It was already shown by Casteleyn et al. (2008) that intragenomic ITS variation is negligible so that each variety contains a single characteristic sequence type. Since 18S–26S rDNA arrays reside in the nuclear genome, ITS sequences are biparentally inherited. A cross between the two varieties will therefore produce an F_1 hybrid with the two sequence types. In addition, we investigated whether the varieties were also differentiated with respect to a chloroplast marker (*rbc*L) and characterized hybrid strains using this chloroplast marker. Plastid DNA is known to have a different pattern of inheritance than nuclear markers. Although plastid inheritance in *Pseudo-nitzschia* is also biparental, F_1 cells contain two maternal or two paternal plastids, or one from both parents, due to random plastid inheritance in the F_1 progeny (Levialdi Ghiron et al. 2007). We also examined whether putative hybrids had intermediate morphologies or if they more closely resembled one or the other of the parental varieties.

Material and methods

Cultures

A subset of 57 strains from a study of *Pseudo-nitzschia pungens* (Grunow ex Cleve) Hasle in the NE Pacific (Adams 2006) was used for this investigation. We included a large random sample of strains that were clearly assignable (with probability > 0.9) to one of the two populations revealed in a cluster analysis (STRUCTURE) and all 12 still available strains with an intermediate probability that did not allow them to be confidently assigned to either population and thus potential hybrids (NA116, NA169, NA177, NA178, NA179, NA192, NA200, NA204, NA231, NA233, NA235, NA242). Living material for these cultures is no longer available because of size reduction in diatoms, but DNA and ITS clones for these strains are available upon request. For 37 additional strains from other geographic areas and the Pacific from a previous study (Casteleyn et al. 2008), the *rbc*L type was determined (Table S1). Cultures were maintained in f/2 medium (Guillard, 1975) based on natural seawater (~30-32 psu) at 18 °C, 60-80 µmol photons m^{-2} s^{-1} irradiance and a 12:12 hour light:dark cycle.

DNA extraction, ITS cloning and sequencing

Genomic DNA of the 57 strains was extracted from pelleted cells using a bead-beating method with phenol extraction and ethanol precipitation (Zwart et al. 1998). The internal transcribed spacer region (ITS1 – 5.8S – ITS2) of the rDNA was amplified using the primers 1800F (Friedl 1996) and a newly designed reverse primer, ITSRpsn (5'- TCC TCT TGC TTG ATC TGA GAT CCG -3'). PCR mixtures of 25 µL were as in Casteleyn et al. (2008). Thirty cycles (1 min at 94 °C, 1 min at 52 °C, 1 min at 72 °C) were performed, with an initial denaturation step of 7 min at 94 °C and a final elongation step of 10 min at 72 °C. Amplified products were purified using the Qiaquick PCR purification kit (Qiagen, Hilden, Germany) and directly sequenced in both directions using the PCR primers. PCR products that gave ambiguities on certain nucleotide positions through direct sequencing, were cloned into plasmids using the pGEM-T Kit (Promega, Madison, WI, USA). Also PCR products of five strains that gave no ambiguities were cloned as a control (Table 2). Three to ten plasmids per strain were sequenced using the universal vector primers T7 and SP6. Sequencing was conducted in ABI 3130xl DNA (Applied Biosystems, Foster City, CA, USA) automated sequencer with the aid of a Big Dye™ Terminator Cycle Sequencing Ready reaction Kit (Applied Biosystems, Foster City, USA). The sequences were assembled, edited and aligned using BioNumerics (Applied Maths, Belgium) software. The boundaries of the ITS sequences were determined by comparison with Casteleyn et al. (2008).

*rbc*L sequencing and DGGE analysis

Sequences of the plastid gene that encodes for the large subunit of RuBisCo (*rbc*L) were obtained for six strains, two of each ITS ribotype and two hybrids. The *rbc*L gene was amplified using the primers DP*rbc*L1 and DP*rbc*L7 (Daugbjerg and Andersen 1997). PCR mixtures were the same as for the ITS region. Twenty-five cycles (1 min at 94 °C, 1 min at 52 °C, 1.5 min at 72 °C) were performed, with an initial denaturation step of 3 min at 94 °C and a final elongation step of 5 min at 72 °C. PCR products were sequenced using amplification primers and two internal primers 15R and 17R (Jones et al. 2005). Sequences were assembled and aligned with Bionumerics. Inspection of the first, second and third codon positions was conducted using MEGA 3.0 (Kumar et al. 2004) and compared with published sequences (Amato et al. 2007, Kowallik et al. 1995).

Two *rbc*L haplotypes were revealed, differing in four point mutations. Primers *rbc*L-PNp-F (5'- TTA TTA CAC TTA CAC CGT GCT GG -3') and *rbc*L-PNp-R (5'- TTT ACC TAC AAC TGT TCC AGC G -3') were designed to amplify a short part of the *rbc*L (135 bp) comprising two of these mutations. Denaturing gradient gel electrophoresis (DGGE) of PCR amplified *rbc*L fragments was developed to separate the different haplotypes. This allowed a rapid screening of all Pacific strains (Casteleyn et al. 2008 and this study) along with strains of other geographic areas (North Sea and adjoining waters, Denmark, Ireland, Spain, Canada, Japan and New Zealand, Table S1, online supplementary material), without the need to clone and sequence PCR products. The *rbc*L fragments were amplified using a nested PCR protocol: the first PCR was as for sequencing of the *rbc*L gene and the second PCR used the designed primers with the forward primer having a GC-clamp. The PCR program started with a denaturation step of 5 min at 94 °C, followed by 30 cycles. Cycle step times were 30s each for denaturation (94 °C), annealing (53 °C) and extension (72 °C). A final extension step was performed for 10 min at 72 °C. PCR samples were loaded onto 8% polyacrylamide gels, 1 mm thick, in 1×TAE (20 mM Tris–acetate (pH 7.4), 10 mM acetate, 0.5 mM disodium EDTA). The denaturing gradient contained 35-45% denaturant (100% denaturant corresponded to 7 M urea and 40% formamide). Equal amounts of PCR products were applied to the DGGE gel. Electrophoresis was performed for 16 h at 75 V and a temperature of 60 °C. DGGE gels were stained with ethidium bromide and photographed on a UV trans-illumination table with a CCD camera.

The production of PCR artefacts (e.g. chimeras and heteroduplexes) is a potential risk when mixed templates of related sequences are amplified by PCR (Wang and Wang 1996, Thompson et al. 2002). This leads to an overestimation of the genetic variation through the amplification of non-existent sequences (Hugenholtz & Huber 2003, von Wintzingerode et al. 1997). To detect PCR artefacts, DNA of two strains with one

haplotype each (confirmed by sequencing of complete *rbc*L gene) were mixed before PCR amplification (to mimic strains with both *rbc*L haplotypes) and loaded on the DGGE-gel. As a negative control the separate PCR products of these 2 strains were mixed after PCR.

Morphology and morphometrics

In order to evaluate morphological variation and differentiation between *P. pungens* var. *pungens* and *P. pungens* var. *cingulata* Villac and their putative hybrids, frustules of 16 strains (Table 2) were studied. Six strains of ribotype/haplotype I and six strains of ribotype/haplotype II and four hybrids (all with both ribotypes and three with both haplotypes and one with a single haplotype) were studied in detail. The strains were randomly chosen within each group and were all grown in the same culture conditions for the same period before morphological measurements were carried out. Therefore morphological differentiation by phenotypic responses to different environmental conditions can be ruled out. Frustules were cleaned by oxidation with hydrogen peroxide and glacial acetic acid, filtered onto 2 µm pore-size, 25 mm diameter polycarbonate filters (Whatman Inc., Florharm Park, NJ, USA) and repeatedly washed with distilled water. Filters were dried and fixed onto coverslips, which were attached to aluminium stubs using conductive tape, coated with gold, and examined with a JEOL JSM5600LV scanning electron microscope (JEOL, Ltd., Tokyo, Japan) (slides available upon request). Photographs were taken by the SEM videograbber and measurements were done on the digital images using ImageJ software version 1.34s (http://rsb.info.nih.gov/ij/). For morphometric analysis, valve length, width, stria density, fibula density, poroid density and band stria density of the valvocopulae were measured for at least 10 valves per strain. For poroid densities a minimum of 150 measurements for 10 valves were made to reduce measurement errors. In addition, the structure of the valvocopula, which differs between the two varieties (Villac and Fryxell 1998), was studied.

Principal component analysis (PCA) on average valves was carried out to investigate patterns of morphological variation. PCA does not employ any pre-existing classification of objects, but can be used to discover discontinuities between groups of objects. For PCA of morphological measurements, we used CANOCO vs. 4.5 for Windows (Biometrics — Plant Research International, Wageningen, The Netherlands). Prior to analysis, the morphological variables were standardized using the center and standardize option for these variables in CANOCO and the linearity of the response curves was checked by looking at the lengths of gradient of a detrended correspondence analysis which were well below a threshold value of 3 (Leps & Smilauer 2003). Analysis

of Variance (ANOVA) of PCA sample scores on the first and second axes was carried out with the ribotypes used as a priori groups (type 1, type 2, hybrid). Post hoc Tukey's honest significant difference (HSD) tests were used to identify significant differences in PCA scores between the three groups. Since valve length varies considerably through the life cycle in diatoms, correlations between morphological variables and length can confound differences between groups. To include the effects of length on the different morphological variables, we used a "homogeneity of slopes" design to test whether length and ribotype interacted in influencing the putative differences in morphology. Post hoc Tukey's HSD tests were used to identify characters showing significant differences between means of the three groups. In all analyses, $P < 0.05$ was considered significant. All statistical analyses were performed using STATISTICA version 6.0 for Windows (StatSoft, Tulsa, USA).

Results

Molecular analyses

Examination of the electropherograms obtained from direct sequencing of the ITS region of 57 NE Pacific strains of *Pseudo-nitzschia pungens* did not reveal any ambiguous nucleotides for 47 of these strains. The sequences could be readily attributed to either ribotype I (30 strains) or ribotype II (17 strains), differing in three point mutations in ITS1 and six point mutations in ITS2 (Table 1). Uniformity of ITS within each variety was already shown by Casteleyn et al. (2008). For the remaining ten strains ambiguities were encountered on these positions, indicating the presence of intragenomic variation in ITS in these strains. Cloning of these strains revealed three types of sequences within clones: ITS ribotype I and II, and several recombinant ribotypes of the two

Table 1. Nucleotide variable sites between the different rDNA ITS ribotypes found in *P. pungens* var. *pungens* (ribotype I) and var. *cingulata* (ribotype II).

	Nucleotide variable sites											EMBL accession numbers
	ITS1				ITS2							
position	49	100	202	219	452	554	611	643	649	669	684	
ribotype I	G	T	G	C	C	A	G	A	T	T	A	FM207591
ribotype II	G	A	A	T	T	A	A	C	C	C	C	FM207590
recombinant types	G	A/T	A/G	C/T	C/T	A	G/A	A/C	C/T	C/T	A/C	FM207592 - FM207611
novel type II	A	A	A	T	T	T	A	C	C	C	C	FM207612

Table 2. ITS ribotype and *rbc*L haplotype composition of 57 *P. pungens* strains. ITS ribotypes were determined by direct sequencing or after cloning (when ambiguities were revealed through direct sequencing + 5 control strains). For cloned strains, number of sequences type I / number of sequences type II and-or novel type II / number of recombinant sequences, are provided. Complete sequences of *rbc*L were obtained for six strains (two of each ITS ribotype and two hybrids). For the remaining strains the *rbc*L haplotype was determined by DGGE. In bold: strains used for morphometric analysis. Light grey shaded cells show hybrid strains based on ITS and *rbc*L. Dark grey shaded cell shows the strain that was not recognized as intermediate by Adams (2006) but that showed some recombinant ribotypes indicating a past hybridization event.

	ITS		*rbc*L	
Strains	Sequencing	ITS ribotype	Method of screening	rbcL haplotype
NA2, NA19, NA21, NA98, NA99, NA100, NA102, **NA104**, NA110, NA175, NA180, NA183, NA185, NA203, NA207, NA208, **NA212**, NA236, NA237, NA240, NA243, NA245, NA246, NA247	directly	I	DGGE	I
NA22	after cloning[b]	3/0/0	DGGE	I
NA205	after cloning[b]	7/0/0	DGGE	I
NA169[a]	after cloning[b]	10/0/0	DGGE	I
NA235[a]	after cloning[b]	10/0/0	DGGE	I
NA95, NA213	directly	I	sequencing	I
NA241	after cloning	4/0/2	DGGE	I
NA1, **NA9**, **NA11**, NA14, NA15, **NA16**, **NA172**, NA174, NA199, NA206, **NA211**, NA216, NA234, NA250	directly	II	DGGE	II
NA108, **NA244**	directly	II	sequencing	II
NA177[a]	after cloning[b]	0/7/3	DGGE	both
NA116[a]	after cloning	4/3/3	DGGE	both
NA178[a]	after cloning	2/5/3	DGGE	both
NA179[a]	after cloning	3/6/1	sequencing	both
NA192[a]	after cloning	1/7/2	DGGE	both
NA200[a]	after cloning	4/2/4	DGGE	both
NA204[a]	after cloning	3/7/0	DGGE	I
NA231[a]	after cloning	8/2/0	DGGE	both
NA233[a]	after cloning	3/5/2	DGGE	I
NA242[a]	after cloning	3/7/0	sequencing	I

[a] strains that could not be confidently assigned to one of the two populations found in a STRUCTURE analysis (i.e. with probability of < 0.9 of belonging to one of the two populations) (Adams 2006)

[b] no ambiguities through direct sequencing (control strains)

parental types (Tables 1 and 2). The existence of both ribotypes in these strains indicates a hybrid origin, and we will refer to them as "hybrids" from here on. Apart from some autapomorphies in sequences of both ribotypes (which are typically revealed through cloning, Thornhill et al. 2007), some novel synapomorphic base pair changes were found in ribotype II sequences: two new point mutations (on position 49 in ITS1 and on position 554 in ITS2, Table 1) were found in six strains (17 sequences in total). The

cloned sequences of five control strains (no ambiguities revealed through direct sequencing) were identical to those of the direct sequencing results except for one strain (NA177) for which three sequences of recombinant origin were recovered after cloning of the PCR product (Table 2).

Direct sequencing of the *rbc*L chloroplast gene of six strains revealed the presence of two haplotypes and one sequence (NA179) with ambiguities. The two haplotypes differed in four bases (Table 3), all situated on third codon positions and resulting in synonymous substitions. With DGGE, the five strains having one haplotype gave a single clear band while strain NA179 showed both bands plus some additional bands (Fig. 1, arrowheads). To verify that these additional bands resulted from PCR artefacts, two strains with different haplotypes were mixed together and then amplified by PCR. These PCR products gave the same banding pattern on DGGE as the strain with both haplotypes. When the PCR products of separate haplotypes were mixed and analysed with DGGE no additional bands appeared, confirming that the additional bands were PCR amplification artefacts and could be ignored. The remaining strains were screened with DGGE to determine which haplotype they possess (Fig. 1). Ribotype groups observed in the nuclear ITS sequences corroborated the *rbc*L haplotype groupings (Table 2), except for the hybrid strains. Hybrid genotypes, containing two ribotypes, had either both haplotypes (six strains) or only haplotype I (three strains). Strain NA177, which harboured three mixed ribotype sequences also contained two haplotypes. The 30 var. *pungens* strains from other geographic areas (North Sea and adjoining waters, Denmark, Ireland, Spain, Canada, Japan and New Zealand) were found to consistently contain only haplotype I plastids (Table S1). The seven Pacific strains from a previous study (Casteleyn et al. 2008) showed the haplotype corresponding to their ribotype (Table S1).

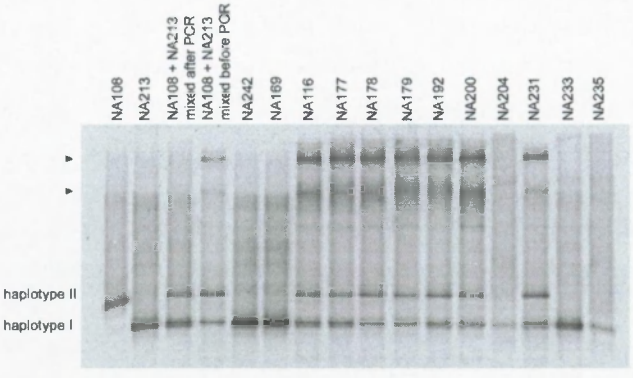

Fig. 1. DGGE gel of *rbc*L fragments of some *P. pungens* strains. Bands represent *rbc*L haplotypes I and II. Arrowheads indicate PCR chimeras.

Table 3. Nucleotide variable sites in the two *rbc*L haplotypes found in *P. pungens* var. *pungens* (haplotype I) and var. *cingulata* (haplotype II). The position numbering corresponds to the complete *rbc*L sequence of *Odontella sinensis* (GenBank accession number NC001713).

	Nucleotide variable sites				EMBL accession numbers
Position	312	660	912	984	
Haplotype I	C	G	T	C	FM207548
Haplotype II	T	A	C	T	FM207547

Morphological analyses

Results of principal components analysis of morphometric data showed that the first and second axes account for 66% and 17% of the total variation respectively. Analysis of variance showed a significant separation ($P = 0.009$) of two groups along the first axis: var. *pungens* and hybrid strains on the left side and var. *cingulata* strains on the right side (Fig. 2). Cell width, stria, fibula, poroid and band striae density contributed most to the separation along the first axis. Length was associated with the second axis but did not result in further differentiation of groups. A homogeneity of slopes model was used to account for size-dependent variations among the three groups (var. *pungens*, var. *cingulata* and hybrids). Univariate tests of significance showed that the effect of length on stria, fibula and band stria densities was different for the three groups (Table 4). In var. *cingulata* length was significantly correlated with stria, fibula and band stria densities, whereas for var. *pungens* strains and hybrid strains no significant correlations were found (P-values > 0.05). Post hoc Tukey HSD tests pointed out that while there was no significant difference in length among the three groups, cell width did significantly differ among them. For the other morphological characters, var. *pungens* and var. *cingulata* are significantly different but only var. *cingulata* could be discriminated from the hybrids (Table 4).

Detailed examination of the band structure of the valvocopula, showed that strains of var. *pungens* had one row of simple large oval poroids (Fig. 3C, see also fig. 4A and B in Casteleyn et al. 2008), while var. *cingulata* strains had square or rectangular poroids which were always occluded by rotae (Fig. 3F, see also fig. 4C in Casteleyn et al. 2008). The bands of the hybrid strains showed poroids with an irregular form, in between the oval and square to rectangular poroids characteristic of each variety. Rotae were usually present except in a few cases but they were less pronounced than in var. *cingulata* strains (Figs. 3I-L).

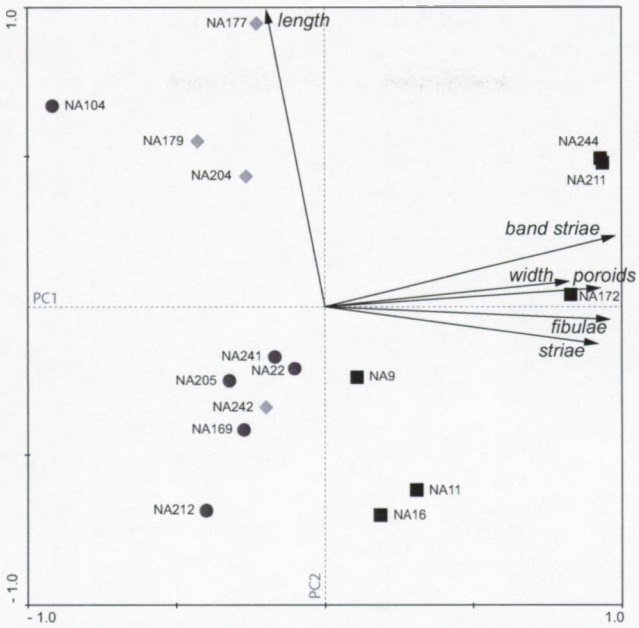

Fig. 2. PCA of morphological measurements on *Pseudo-nitzschia pungens* var. *pungens* (circles), *P. pungens* var. *cingulata* (squares) and hybrids strains (diamonds). The direction of each arrow indicates increasing values of the corresponding morphological parameter. The first two axes are shown, explaining 66% and 17% of total variation respectively.

Table 4. Homogeneity of slopes model (HSM) and Tukey HSD tests for morphometric trait variation between *P. pungens* var. *pungens* (type I), *P. pungens* var. *cingulata* (type II) and hybrids. In bold: statistically significant at $P = 0.05$

	Type I	Type II	Hybrids	HSM *P*-values			Probabilities for Tukey HSD-tests		
	Average ± SD	Average ± SD	Average ± SD	type	length	type * length	between type I and II	between type I and hybrids	between type II and hybrids
Length (µm)	109.5 ± 15.3	105.0 ± 16.3	127.9 ± 18.3				0.86	0.18	0.081
Width (µm)	2.9 ± 0.1	3.8 ± 0.2	3.3 ± 0.3	0.71	0.99	0.91	**0.0004**	**0.043**	**0.033**
Striae density (in 10µm)	11.0 ± 0.6	11.8 ± 0.4	11.0 ± 0.1	0.073	0.83	**0.033**	**0.017**	1.00	**0.028**
Fibulae density (in 10µm)	11.4 ± 0.6	12.2 ± 0.6	11.4 ± 0.2	**0.023**	0.91	**0.0089**	**0.012**	0.95	**0.014**
Poroid density (in 1µm)	3.2 ± 0.1	4.3 ± 0.6	3.2 ± 0.1	0.91	0.42	0.52	**0.0012**	1.00	**0.0026**
Band striae density (in 10µm)	14.4 ± 1.1	19.8 ± 3.8	15.0 ± 0.5	0.10	0.092	**0.026**	**0.0011**	0.85	**0.0052**

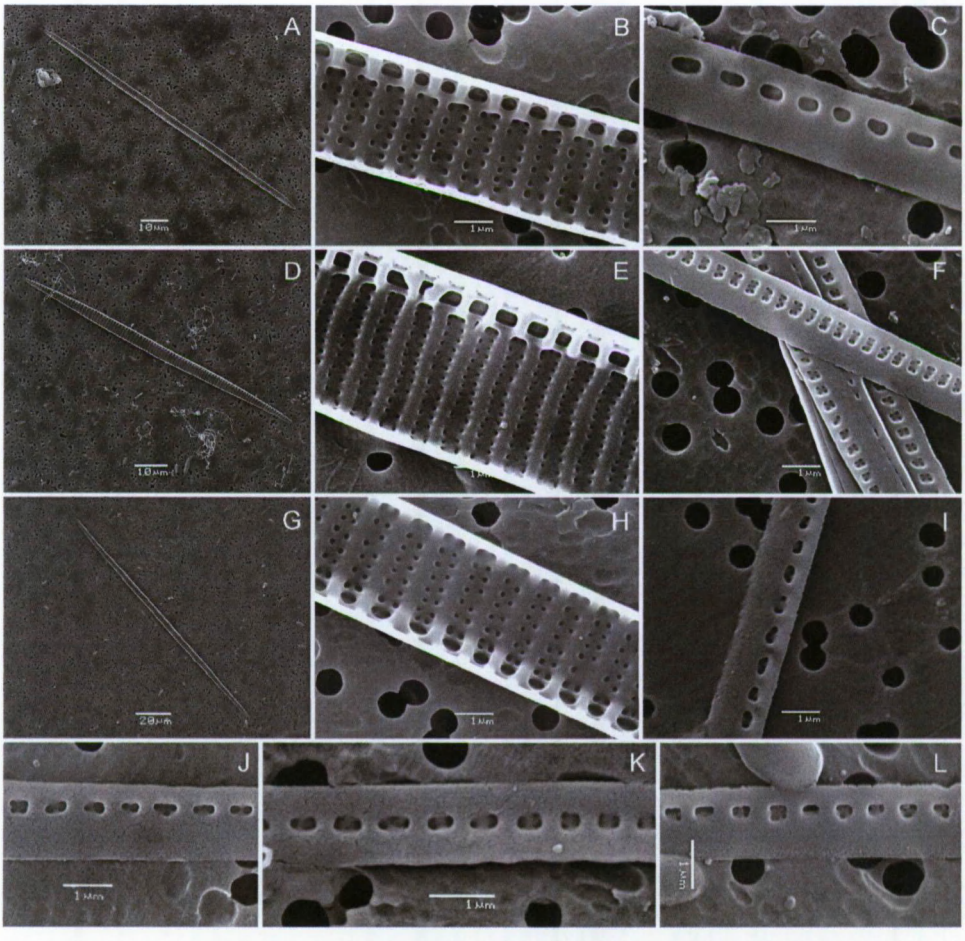

Fig. 3. *Pseudo-nitzschia pungens* frustules and valvocopulae, SEM. **(A–C)** *P. pungens* var. *pungens*. Complete frustule (strain NA205) (A), valve face in internal view (strain NA212) (B) and valvocopula with one row of simple large oval poroids (strain NA104) (C). **(D–F)** *P. pungens* var. *cingulata*. Complete frustule (strain NA11) (D), valve face in internal view (strain NA211) (E) and valvocopula with square or rectangular poroids occluded by rotae (strain NA244) (F). **(G–L)** Hybrid strains. Complete frustule (strain NA177) (G), valve face in internal view (NA179) (H) and valvocopula with more irregular poroids (strains NA177, NA179, NA204 and NA242) (I–L).

Discussion

Genetics

The notion that the two *P. pungens* varieties represent genetically well-differentiated entities (Casteleyn et al. 2008) is reinforced by the observation that, next to having divergent ITS rDNA sequences, they can also be distinguished by diagnostic base pair

differences in the plastid encoding *rbc*L gene. Moreover, the two *P. pungens* populations outlined in a microsatellite study by Adams (2006) using STRUCTURE (Pritchard et al. 2000) also correspond to the two varieties, as was shown in the present study by determining the ITS ribotypes of the same strains, thereby demonstrating that the varieties are also differentiated for these microsatellite loci. In *Pseudo-nitzschia,* the *rbc*L and ITS loci are not always genealogically concordant within species. Amato et al. (2007) demonstrated that in *Pseudo-nitzschia delicatissima* strains two distinct *rbc*L haplotypes could be found whereas all strains have a single ITS genotype. In contrast, *Pseudo-nitzschia pseudodelicatissima*-like strains with different ITS sequences can share the same *rbc*L haplotype (Amato et al. 2007).

The fact that we found strains with both ribotypes together with the uniformity of ITS within each variety (Casteleyn et al. 2008), clearly demonstrates the occurrence of natural hybrids between both *P. pungens* varieties. In a recent study on *Pseudo-nitzschia multistriata*, strains with different ITS types were also shown to hybridize in the field (D'Alelio et al. 2009b). However, in *P. multistriata* there were no indications that the strains with a different ITS type belonged to separate lineages. We can expect that with a better understanding of species boundaries based on sequence data, natural hybrids are likely to be found increasingly in the future.

The hybrid strains in our study also contain novel ITS ribotypes, either resulting from recombination of both ribotypes or characterized by two new point mutations in ribotypes II. The recombinant ribotypes found in our hybrids could either be real recombinants resulting from crossing-over during gamete formation in F_1 or mere PCR artefacts (e.g. chimeras or heteroduplexes) or a combination of both. On the basis of the available data we were not able to distinghuish between these possibilities. However we suspect most recombinants to be PCR artefacts. It was shown that up to 30% of sequences could be chimeras when mixed templates of closely related sequences are amplified by PCR (Wang and Wang 1996) and 9 out of 20 of our recombinants showed mosaic patterns typical for heteroduplexes (Thompson et al. 2002). Yet it is not known whether parental ribotypes after formation of F_1 hybrids are maintained without recombination or whether they are maintained and undergo various degrees of recombination, both possibilities having been demonstrated in plants (Alvarez & Wendel 2003). It is also possible that one ribotype rises to dominance within a genome as a consequence of concerted evolution (Dover 1982, Liao 1999); this may occur as fast as one generation after the combination of two parental ITS types in plants (Aguilar et al. 1999). A way to study the fate of the ITS repeats in hybrids would be by creating artificial hybrids to explore if and how fast concerted evolution acts on the ITS region of early generation hybrids and introgressants. Despite the fact that Casteleyn et al. (2008) demonstrated sexual compatibility within and between the two *P. pungens* varieties, we

have been unable to induce sexual reproduction in our cultures despite multiple trials to cross ribotype I, ribotype II and hybrid strains in all possible combinations under various culture conditions (Tables S2 and S3).

Chloroplast loci are inherited in a different way than nuclear DNA. Based on cytological observations, plastid inheritance has been shown to be biparental in several centric and in all studied pennate diatoms (Jensen et al. 2003, Round et al. 1990). In *P. pungens*, the vegetative cells possess two plate-like chloroplasts. During mitosis they divide clonally and segregate over the daughter cells. In the meiotic cycle, each gamete contains two plastids and the two parental plastids are retained in the zygote, so all four plastids arrive in the initial cell and then segregate two by two during the first mitotic division of the initial cell into the two daughter cells (Chepurnov et al. 2005, chapter 2 of this thesis). The origin of plastids in the F_1 generation was studied in *P. delicatissima* using *rbc*L haplotypes by Levialdi Ghiron et al. (2007). They showed that the F_1 cells inherit either two maternal plastids or two paternal ones, or one paternal and one maternal one in a ratio of 1/6, 1/6 and 2/3 respectively. Our results are in agreement with their findings: hybrid strains (containing two ribotypes) with either 2 haplotypes or only one haplotype were found although no strains with only haplotype II (var. *cingulata*) chloroplasts were recovered. This might be attributed to the low number of hybrid strains examined. Although Levialdi Ghiron et al. (2007) showed that cells with different *rbc*L haplotypes interbreed freely in the laboratory and F_1 hybrids are viable and fertile, no strains with both haplotypes were found in their field samples. This could point to the existence of prezygotic mating barriers in the field. In the present study, heteroplastidial strains were recovered in our field samples from the NE Pacific, indicating that prezygotic barriers are weak in *P. pungens*. Since chloroplasts are biparently inherited we could not use the *rbc*L marker to distinguish the paternal and maternal species.

Analysis of a combination of ITS and *rbc*L did not allow us to discriminate between hybrid classes (F_1, F_2, backcrosses, …), because strains containing both ribo- and one or two haplotypes can obviously be found in each hybrid class and the uncertain origin of the recombinant ribotypes complicates proving the presence of further hybridization and recombination. Nevertheless, comparison with a population genetic study of *P. pungens* strains from the NE pacific using four microsatellite markers (Adams 2006) yields some further insights. That study revealed sixteen strains that could not be confidently assigned to one of the two populations found in a STRUCTURE analysis (i.e. with probability of < 0.9 of belonging to one of the two populations) (Adams 2006). For ten out of twelve of these strains used in the present study, we have now shown that they are the result of hybridization between the two *P. pungens* varieties. Eight of these strains had a probability between 0.4-0.6 of belonging to one or the other population in the STRUCTURE analysis, suggesting that they are first generation hybrids (e.g. Coyer et al.

2007). However, one strain had only ribotype I (through direct sequencing) and both *rbc*L haplotypes, indicative of further hybridization since F_1 hybrids should have both ribotypes. The four other strains had a probability of 0.8-0.9 of belonging to one or the other cluster which points to backcrosses to either one or the other population (e.g. Coyer et al. 2007). Two of them (NA231 and NA242) were recognised as hybrids in our study and according to the STRUCTURE analysis they could be backcrosses to var. *cingulata*. This suggests that not all hybrids are F_1's and that further hybridization is taking place. The presumable backcrosses to var. *pungens* (NA169 and NA235) were not recognised as hybrids in our analysis. We also found one strain that was not recognised as intermediate in the microsatellite analysis but that harboured some recombinant ribotypes in the nucleus which may indicate a past hybridization. These results corroborate the notion that a relatively large number of microsatellite loci (12 at least) are needed to clearly distinguish between backcrosses, F_1 hybrids and purebred parental individuals (Vaha and Primmer 2006).

Morphology

Various morphological features of the frustules were compared between the hybrid strains and the two *P. pungens* varieties. Of the six morphometric characters studied, only valve width was found to be intermediate between the parents in the hybrids. In addition, a detailed examination of the structure of the valvocopula also revealed intermediate characteristics. For the other morphometric characters, the hybrids could not be discriminated from var. *pungens* although both varieties could be readily distinguished. Artificial F_1 progeny between genetically and morphologically closely related species of the diatom *Eunotia bilunaris* were intermediate with respect to all measured morphological characters (Vanormelingen et al. 2008). Despite the fact that most of our hybrids are presumably F_1's, we did not observe an intermediate morphology. It thus appears that factors involved in valve morphology do not cooperate in an additive way in the hybrids. Instead, genetic factors with dominant and/or epistatic effects appear to be involved to a varying degree in the genetic architecture of valve morphology when parental genomes are combined in hybrids. To date, however, it remains an open question how particular genes or interactions between genes effect the development of a morphological/phenotypic character in diatoms.

Conclusions and perspectives

The present study clearly demonstrates the occurrence of natural hybridization between two genetically well-differentiated varieties of *P. pungens* in the NE Pacific. The evolutionary outcome of such hybridization events depends on the hybridization

frequency and the fitness of the resulting progeny, as this will determine the amount of gene flow between lineages (Barton and Hewitt 1985). Further analysis of the spatial and temporal range of the contact zone by extensive sampling along the NE Pacific coast, where both varieties co-occur, and the use of multiple biparentally inherited nuclear markers is necessary to investigate the hybridization frequency and the extent of the hybrid zone. In addition, physiological experiments and detailed studies of population dynamics should be conducted to determine hybrid fitness. Together, these data will allow defining the amount of gene flow between the *P. pungens* varieties and its evolutionary consequences.

Acknowledgements

Financial support for this research was provided by the Research Programmes G.0292.00 and G.0197.05 of the Flemish Fund for Scientific Research (FWO) and the BOF-project GOA 12050398 of Ghent University. G.C. is funded by the Flemish Institute for the Promotion of Innovation by Science and Technology (IWT). The *P. pungens* isolates used in this study were collected as part of the Ecology and Oceanography of Harmful Algal Blooms in the Pacific Northwest (ECOHAB-PNW) project that was funded by the National Science Foundation's (NSF) ECOHAB project OCE-0234587 and NOAA ECOHAB grant NA16OP1450. We thank V. Creach, N. Lundholm, M. Lastra, S. Bates, Y. Kotaki and L. Rhodes for providing strains from other areas.

Supplementary tables

Table S1. Strains of *P. pungens* from other geographic areas for which the *rbc*L haplotype is determined by DGGE. ITS sequences and EMBL accession numbers are from a previous study (Casteleyn et al. 2008).

strain	sampling location	sampling date	sample supplier and initial strain designation	*P. pungens* variety	EMBL accession number for ITS	*rbc*L haplotype determined by DGGE
18	North Sea, The Netherlands, Noordwijk	22 May 2002	L. Peperzak, NB-P4 (2002)	var. pungens	AM778738	I
57	North Sea, northern France, Ambleteuse	13 Sep. 2003	F. Leliaert, Ambl - 57	var. pungens	AM778746	I
28	North Sea, Belgium, VLIZ station 130	30 Jun. 2003	K. Muylaert, ZL130 - 28	var. pungens	AM778742	I
V215(3)2	North Sea, Belgium, VLIZ station 215	24 Mar. 2004	Flemish Marine Institute, VLIZ215(24/03/04) - 2	var. pungens	AM778748	I
V700(7)5	North Sea, Belgium, VLIZ station 700	19 Jul. 2004	Flemish Marine Institute, VLIZ700(19/07/04) - 5	var. pungens	AM778758	I
V7(5)10	North Sea, Belgium, VLIZ station 700	4 May 2007	Flemish Marine Institute, VLIZ700(04/05/07)-10	var. pungens		I
V330(3)4	North Sea, Belgium, VLIZ station 330	24 Mar. 2004	Flemish Marine Institute, VLIZ330(24/03/04) - 4	var. pungens		I
V421(10)4	North Sea, Belgium, VLIZ station 421	19 Oct. 2004	Flemish Marine Institute, VLIZ421(19/10/04) - 4	var. pungens		I
S(7)1	Oostend lagoon, Belgium, Spuikom Oostende	13 Jul. 2004	Flemish Environment Agency, Spuikom(13/07/04) - 1	var. pungens	AM778768	I
1	Westerscheldt, The Netherlands, Terneuzen	24 May 2002	K. Sabbe, Trz1	var. pungens	AM778733	I
W4(7)2	Westerscheldt, The Netherlands, Terneuzen	14 Jul. 2004	Netherlands Institute of Ecology, WS1(14/07/04) - 2	var. pungens	AM778775	I
W1(7)1	Westerscheldt, The Netherlands, Breskens	14 Jul. 2004	Netherlands Institute of Ecology, WS1(14/07/04) - 1	var. pungens	AM778769	I
Vlissingen 1	Westerscheldt, The Netherlands, Vlissingen	15 May 2007	G. Casteleyn, vlissingen 1	var. pungens		I
Terneuzen 1	Westerscheldt, The Netherlands, Terneuzen	15 May 2007	G. Casteleyn, terneuzen 1	var. pungens		I
Hfdplaat 1	Westerscheldt, The Netherlands, Hoofplaat	15 May 2007	G. Casteleyn, hfdplaat 1	var. pungens		I
schore 23	Westerscheldt, The Netherlands, Schore	15 May 2007	G. Casteleyn, schore 23	var. pungens		I

Table S1. Continued.

zijpe 1	Oosterscheldt, The Netherlands, Zijpe	15 May 2007	G. Casteleyn, zijpe 1	var. pungens		I
Neeltje Jans 1	Oosterscheldt, The Netherlands, Neeltje Jans	4 July 2007	G. Casteleyn, neeltje jans 1	var. pungens		I
Zeedijk 1	Lake Grevelingen, The Netherlands, Zeedijk	15 May 2007	G. Casteleyn, zeedijk 1	var. pungens		I
Osse 1	Lake Grevelingen, The Netherlands, Osse	15 May 2007	G. Casteleyn, osse 1	var. pungens		I
Denmark 1	Limfjord, Denmark, Løgstær	27 Aug. 1997	N. Lundholm, L7	var. pungens		I
Ierse Zee H8-1	Irish Sea, Ireland, Belfast Harbour	Jul. 2007	V. Creach, HAB8-1	var. pungens		I
Vigo-1	E Atlantic, Spain, Bay of Vigo	1 Apr. 2004	M. Lastra	var. pungens	AM778776	I
Cn-218	NW Atlantic, Canada, Boughton River, Prince Edward Island	7 Sep. 2004	S. Bates, CL-218	var. pungens	AM778794	I
Cn-2	NW Atlantic, Canada, March Water, New London Bay, Prince Edward Island	21 Nov. 2005	S. Bates, IS-2	var. pungens		I
Cn-254	NW Atlantic, Canada, Miramichi Bay, New Brunswick	18 Sep. 2007	S. Bates, CL-254	var. pungens		I
Jp-01	NW Pacific, Japan, Ofunato Bay, Iwate Prefecture	7 Aug. 2000	Y. Kotaki, OFPp 001	var. pungens	AM778812	I
OFPp06-54	NW Pacific, Japan, Ofunato Bay, Iwate Prefecture	13 Nov. 2006	Y. Kotaki, OFPp06-54	var. pungens		I
NZ-74	SW Pacific, New Zealand, Taylor's Mistake, South Island	2004	L. Rhodes, CAWB74	var. pungens	AM778817	I
NZ25	SW Pacific, New Zealand, Gisborne Harbour, North Island	16 Apr. 2007	L. Rhodes, NZ25	var. pungens		I
US-94	NE Pacific, North America, Sequim Bay State Park, WA	8 Aug. 2002	B. Bill, NWFSC-094	var. pungens	AM778798	I
US-96	NE Pacific, North America, Sequim Bay State Park, WA	8 Aug. 2002	B. Bill, NWFSC-096	var. pungens	AM778803	I
US-135	NE Pacific, North America, ECOHAB I (coastal Washington)	Jun. 2003	B. Bill, NWFSC-135	var. cingulata	AM778811	II
US-132	NE Pacific, North America, ECOHAB I (coastal Washington)	Jun. 2003	B. Bill, NWFSC-132	var. cingulata	AM778810	II
US-115	NE Pacific, North America, ECOHAB I (coastal Washington)	Jun. 2003	B. Bill, NWFSC-115	var. cingulata	AM778804	II
US-93	NE Pacific, North America, Eld Inlet, Mud Bay, WA	17 Jun. 2002	B. Bill, NWFSC-093	var. cingulata	AM778796	II
US-77	NE Pacific, North America, La Push, WA	18 Sep. 2001	B. Bill, NWFSC-077	var. cingulata	AM778795	II

Table S2. Culture conditions used to induce sexual reproduction in all pairwise combinations within and between ribotype I and ribotype II strains (for a list of strains used see Table S3).

conditions	
light	10-25-50-100 µmol photons m^{-2} s^{-1} Natural daylight – window on the north 24h or 48h dark arrest continous light
Nutrient concentrations	100%, 50%, 25%, 10% f/2 filtered seawater without addition of nutrients L1 medium (contains additional trace metals in comparison with f/2)
temperature	8-22°C
salinity	20 – 30 – 35 - 40 psu
pH	7,5 – 7,8 – 8 – 8,3
Culture density	Low versus high
Culture volume	96-, 24-, 6- well plates (0.025 – 3 -15 ml), 50 ml culture flasks

Table S3. Cell size of *P. pungens* strains at first crosses. Trials to cross strains were done until the cultures were below the viable size, also for the not measured cultures (nm). Length measurements were done at 14.02.2006. Monthly attempts were conducted to cross ribotype I, ribotype II and hybrid strains in all possible combinations in the same culture conditions successfully applied previously (Chepurnov et al. 2005, Casteleyn et al. 2008). Upper sexual size threshold for North Sea strains = 115 µm (Chepurnov et al. 2005).

ribotype	strain	Mean cell length in µm (n=5)
I	NA2	92.5
	NA19	90
	NA21	67.5
	NA22	102.5
	NA95	137.5
	NA98	130
	NA99	137.5
	NA100	140
	NA102	142.5
	NA104	135
	NA110	100
	NA169	96
	NA175	122
	NA180	100
	NA183	135
	NA185	128.5
	NA203	130
	NA205	105
	NA207	117.5
	NA208	137.5
	NA212	90
	NA213	127.5
	NA235	nm
	NA236	107.5
	NA237	135

Table S3. Continued.

I	NA240	97.5
	NA243	135
	NA245	140
	NA246	nm
	NA247	nm
hybrid?	NA241	102.5
II	NA1	67.5
	NA9	110
	NA11	82.5
	NA14	77.5
	NA15	82.5
	NA16	85
	NA108	120
	NA172	100
	NA174	117.5
	NA199	nm
	NA206	nm
	NA211	117.5
	NA216	nm
	NA234	nm
	NA244	121
	NA250	nm
hybrids	NA116	nm
	NA177	142.5
	NA178	nm
	NA179	118.5
	NA192	nm
	NA200	nm
	NA204	127.5
	NA231	nm
	NA233	nm
	NA242	96