

## Molecular species discovery in the diatom *Sellaphora* and its congruence with mating trials

Pieter VANORMELINGEN<sup>1,\*</sup>, Katharine M. EVANS<sup>2,3</sup>, Victor A. CHEPURNOV<sup>1</sup>, Wim VYVERMAN<sup>1</sup> & David G. MANN<sup>2</sup>

<sup>1</sup>Laboratory of Protistology and Aquatic Ecology, Ghent University, Krijgslaan 281 – S8, 9000 Gent, Belgium; \*Corresponding author e-mail: pieter.vanormelingen@UGent.be, Tel.: +32 9 264 85 05, fax: +32 9 264 85 99

<sup>2</sup>Royal Botanic Garden, Edinburgh EH3 5LR, United Kingdom

<sup>3</sup>School of Geosciences, University of Edinburgh, EH9 3JW, United Kingdom

**Abstract:** Many diatom and other microbial eukaryote morphospecies consist of a variable number of (pseudo)cryptic species, with obvious consequences for such fields as biogeography and community ecology. Here, we investigated the species limits of morphologically similar small-celled strains of the model diatom *Sellaphora* from the United Kingdom and Australia, using *cox1* mitochondrial and *rbcL* chloroplast gene sequences. Based on *cox1* sequence data, the sequenced strains belonged to six closely related lineages, presumably species, of which one corresponds to the previously described *S. auldreekie* D.G. MANN & S.M. McDONALD. Although *rbcL* displayed less sequence variation, the same six lineages were also recovered in an *rbcL* phylogeny of the genus. Molecular species discovery was compared to mating trials involving three of the lineages, showing that they were reproductively isolated. Incomplete evidence from a fourth lineage suggested that it too was reproductively isolated. *A posteriori* examination of light microscope morphology revealed no simple metrics or presence/absence characters that could consistently separate all species of the *auldreekie* complex, even though some do differ in pole width or stria density. While it is premature to make conclusions about their biogeography, it is obvious that a number of cryptic *Sellaphora* species thus far undetected in the UK are easily found at several localities in warm-temperate Australia.

**Key words:** *cox1*, cryptic species, diatoms, *rbcL*, reproductive isolation, *Sellaphora*, species delimitation

## INTRODUCTION

Species taxonomy in microbial eukaryotes has been based almost exclusively on morphology and has given the impression of a low number of generalist cosmopolitan species. More recently, the application of molecular tools in taxonomy has revealed a large number of previously unrecognized species (e.g. COLEMAN 2001; MONTRESOR et al. 2003; ŠLAPETA et al. 2005; DARLING & WADE 2008; EVANS et al. 2008), with potentially restricted geographical ranges (PAWLOWSKI & HOLZMANN 2008) and niche breadths (DE VARGAS et al. 1999; LEWIS & FLECHTNER 2004; POULÍČKOVÁ et al. 2008; VANELSLANDER et al. 2009; KAERIYAMA et al. 2011). Sometimes, molecular evidence for lineage splitting has been backed up by breeding data and detailed morphological investigations. Reasons why the use of morphology can be deficient for species delimitation and identification include deviations from original species descriptions, force-fitting of morphologically variable individuals to known species, unrecognized phenotypic plasticity, a failure to detect subtle morphological discontinuities between species,

and a large overlap or even total lack of distinguishing morphological characters (e.g. TRAINOR 1998; MANN 1999; KIM et al. 2004; COLEMAN 2005; ŠLAPETA et al. 2005; VANORMELINGEN et al. 2008a). Species that cannot be distinguished morphologically are generally termed cryptic, while the term ‘pseudocryptic’ has been proposed for species with subtle morphological differences and ‘semicryptic’ for species whose morphological ranges overlap so that identification of some individuals is impossible unless extra, non-morphological evidence is available (MANN & EVANS 2008). While such hidden species diversity has been found in many different organisms, its prevalence may be especially high in microbial eukaryotes due to the often limited number of available morphological characters, and the presence of non-visual reproductive signals (BICKFORD et al. 2007).

One way to avoid difficulties associated with morphology as the basis for delimiting species is the use of molecular phylogenies. The primary evidence for different species is then the existence of distinct lineages, i.e. clusters of sequences showing little intracluster sequence divergence subtended by a

relatively long and highly supported branch. Correlation between different molecular markers, and with other properties such as morphology, reproductive isolation, or ecology, provides further evidence for the presence of distinct species (e.g. GÓMEZ et al. 2007; BURNS et al. 2008; JARGEAT et al. 2010; TRONHOLM et al. 2010), based on the general notion of species as separately evolving metapopulation lineages (DE QUEIROZ 2007). In groups with a large proportion of (pseudo)cryptic or unstudied species, the availability of model species complexes with well-understood species limits based on a range of properties can be highly useful, for instance to investigate the performance of different molecular markers for large-scale phylogenetic species delimitation (such as applied by e.g. MONAGHAN et al. 2009), the discriminatory power of potential DNA barcoding regions (HAMSHER et al. 2011; ZIMMERMAN et al. 2011), or the congruence between different methods used for species delimitation (AMATO et al. 2007).

Diatoms are the most species-rich group of microbial eukaryotes and are characterized by the presence of an intricately ornamented siliceous cell wall, which provides a wealth of morphological characters for taxonomy (ROUND et al. 1990). Despite this, many described morphospecies actually comprise several to many species, which are difficult or impossible to distinguish routinely, based on morphology (e.g. AMATO et al. 2007; BESZTERI et al. 2007; LI et al. 2007; VANORMELINGEN et al. 2008b, 2013; ELLEGAARD et al. 2008; TROBAJO et al. 2009; SOUFFREAU et al. 2012). Perhaps the most intensively studied diatom genus concerning (pseudo)cryptic species diversity is *Sellaphora* (MANN 1989a, 1999, 2008; MANN et al. 1999, 2004, 2008, 2009; BEHNKE et al. 2004; EVANS et al. 2007, 2008; POULÍČKOVÁ et al. 2008; EVANS & MANN 2009; MANN et al. 2011). While traditionally only four larger-celled species have been recognized (*S. americana*, *S. bacillum*, *S. pupula* and *S. laevissima*: e.g. KRAMMER & LANGE-BERTALOT 1986, who treated all within *Navicula sensu lato*), a recent survey of diversity for the British Isles, based on detailed light microscopical examination and some molecular data, came up with no fewer than 54 so-called phenodemes (MANN et al. 2008) that seemed likely to be worthy of species status. Thirteen of these correspond to demes or species outlined previously based on a combination of valve morphology and molecular data, and some mating trials (BEHNKE et al. 2004; MANN et al. 2004; EVANS et al. 2007, 2008; JAHN et al. 2008). This indicates high species diversity in the genus. Species identification and especially discovery based exclusively on valve morphology are difficult, however, due to the often very subtle morphological differences between species (BEHNKE et al. 2004; MANN et al. 2008). A molecular approach to explore *Sellaphora* species diversity has therefore been taken, and the sampling effort extended to geographical areas other than the UK (EVANS et al. 2007, 2008; MANN et

al. 2011).

Here we investigate the species limits for a number of small-celled *Sellaphora* strains from the United Kingdom and Australia that morphologically resemble the previously described *S. auldreekie* D.G. MANN & S.M. McDONALD (MANN et al. 2004). The aim was to determine whether an extension of sampling using gene sequences would reveal extra diversity beyond that already known from morphology (MANN et al. 2008) and, if so, whether this diversity would be manifested *a posteriori* also in crossability and morphology. Molecular phylogenies based on *cox1* mitochondrial and *rbcl* chloroplast gene sequences were constructed and lineage separation compared, mating trials conducted among selected strains to check for the presence of intrinsic reproductive isolation between lineages, and morphological comparisons made using simple metrics routinely used in diatom taxonomy.

## MATERIAL AND METHODS

**Culture establishment and maintenance.** Clonal cultures were established from natural samples taken between 2003 and 2008 in the UK, Czech Republic and Australia (Table S1). Epipelagic communities were sampled and harvested as described by POULÍČKOVÁ et al. (2007). Strains were isolated as described by EVANS et al. (2007) and grown in WC medium (GUILLARD & LORENZEN 1972; at the Laboratory of Protistology and Aquatic Ecology [PAE] without pH adjustment), at 18 °C, 20–30  $\mu\text{mol photons. m}^{-2} \text{ s}^{-1}$ , and a 12:12h light:dark cycle. Under these conditions, cultures were re-inoculated every two to three weeks. Long-term maintenance of strains at the Royal Botanic Garden Edinburgh (RBGE) was achieved by keeping them in 90 mm Petri dishes at a lower temperature (12–15 °C) and irradiance (c. 5–20  $\mu\text{mol photons. m}^{-2} \text{ s}^{-1}$ ). This way, strains could be kept for up to four months without re-inoculation. A list of the strains used is provided in Table S2.

**Morphology.** Cells for valve morphology were usually harvested from cultures in late exponential phase, grown in Petri dishes (RBGE) under the same conditions as detailed above. Voucher preparations were made by cleaning harvested cells in concentrated nitric acid and mounting them in Naphrax (Brunel Microscopes: <http://www.brunelmicroscopes.co.uk/>). All voucher slides (Table S2) are kept in the diatom herbarium of the RBGE. Bright-field light microscope (LM) observations of valve morphology were carried out with planapochromat lenses (nominal numerical apertures 1.32 or 1.4), using either (i) a Reichert Polyvar 2 photomicroscope fitted with a Polaroid DMC2 digital camera (1600×1200 pixels; images were captured via Optimas image analysis software version 6.2: MediaCybernetics, Silver Spring, MD 20910, USA); or (ii) a Zeiss AxioImager.M2 with a Zeiss AxioCam MRc5 digital camera (images captured and managed via Zeiss Axiovision software). Background noise and specks were removed digitally either by using an inbuilt background correction facility (Zeiss Axiovision) or by image division (BAYER et al. 2001). In addition, for length measurements of initial cells in selected crosses (see

below), photos were taken using a Zeiss Axiovert 40 inverted microscope equipped with a digital camera (Powershot G3, Canon), and measurements performed using the open-source program ImageJ (<http://rsb.info.nih.gov/ij/index.html>). Representative valves are shown in Fig. 1, grouped by molecular clade.

Lengths and widths of valves were measured either directly, using a calibrated scale viewed via a Reichert drawing attachment, or from digital images. Pole width was measured 2.5, 2 or 1.5  $\mu\text{m}$  from the end of the valve, in post-initial F1 valves (which had a valve length of at least 23.5  $\mu\text{m}$ ), medium-sized valves ( $< 23.5 \geq 15 \mu\text{m}$ ) or the smallest valves ( $< 14 \mu\text{m}$ ), respectively, reflecting the fact that the point of inflection of the valve outline moves progressively closer to the valve apex as size reduction proceeds in these rostrate or subcapitate diatoms. Stria densities (# in 10  $\mu\text{m}$ ) were determined adjacent to the raphe, by measuring the distance parallel to the apical axis occupied by the first 10 complete striae (i.e. those running from the raphe–sternum to the margin) out from the central area, on the primary side of the valve. In small-celled strains, there were fewer than 10 complete striae towards either pole, above or below the central area, and we therefore measured 8 or 6 striae instead.

#### **Cox1 and rbcL sequencing and phylogenetic analysis.**

Cells for DNA extraction were harvested from cultures in late exponential growth phase in 90 mm Petri dishes (RBGE) or 50 ml culture flasks (PAE) grown under the conditions detailed above. Cells were harvested by centrifugation and DNA extracted using the Qiagen Plant DNeasy kit (Qiagen, Crawley, UK) (RBGE) or as described in ZWART et al. (1998), but without the last purification step (PAE). The partial *cox1* gene was amplified using primers GazF2 and KEtmR (EVANS et al. 2007). The *rbcL* chloroplast gene was amplified using primers DPrbcL1 and DPrbcL7 (JONES et al. 2005). At PAE, PCR reaction volumes were 50  $\mu\text{l}$  and contained 2  $\mu\text{l}$  of template DNA, 1  $\times$  buffer (including 1.5 mM  $\text{MgCl}_2$ : ‘Buffer II’, Applied Biosystems, Foster City, USA), 200  $\mu\text{M}$  of dNTPs, 0.5  $\mu\text{M}$  of each primer, bovine serum albumen (BSA) at a final concentration of 0.4  $\mu\text{g.}\mu\text{l}^{-1}$ , and 2.5 units of Taq for *cox1* and 5 units for *rbcL* (AmpliTaq, Perkin–Elmer, Wellesley, USA). At RBGE, reaction volumes for *cox1* were 30  $\mu\text{l}$  and contained 1.5  $\mu\text{l}$  of template DNA, 1  $\times$   $\text{NH}_4$  buffer (Bioline, London, UK), 2.5 mM  $\text{MgCl}_2$ , 200  $\mu\text{M}$  of dNTPs, 0.3  $\mu\text{M}$  of each primer and 0.75 units of Taq (Bioline). PCR conditions were as described in EVANS et al. (2007). *RbcL* PCR reactions were as described for *cox1*, except total volumes were 50  $\mu\text{l}$  and 2  $\mu\text{l}$  of DNA template and 1.25 units of Taq were added. PCR conditions were as described in JONES et al. (2005).

PCR products were purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) at PAE, and the GFX PCR purification kit or ExoSAP-IT (GE Healthcare Life Sciences, Little Chalfont, UK) at RBGE. At PAE, the sequencing reaction was performed using an ABI Prism BigDye V 3.1 Terminator Cycle Sequencing kit (Applied Biosystems) with the following cycle sequencing profile: initial step of 1 min at 96  $^\circ\text{C}$ , 30 cycles of 96  $^\circ\text{C}$  for 10 s, 50  $^\circ\text{C}$  for 10 s and 60  $^\circ\text{C}$  for 1 min 15 s. The resulting products were analysed on a Perkin–Elmer ABI Prism 3100 sequencer (Applied Biosystems). At RBGE, sequencing was conducted as described in EVANS et al. (2007) or as described above, but with the following cycle sequencing profile: 25 cycles of 95  $^\circ\text{C}$  for 30 s, 50  $^\circ\text{C}$  for 20 s and 60  $^\circ\text{C}$  for 4 min and analysis on an ABI 3730 sequencer (Applied Biosystems).

For *cox1* sequencing, both PCR primers were used and internal primers, KEint2F and KEintR, if necessary (EVANS et al. 2007). For *rbcL* sequencing, amplification primers plus a selection of internal primers were used (NDRbcL5, NDRbcL6, NDRbcL11, NDRbcL12, NDRbcL13, 15R, and 17R: DAUGBJERG & ANDERSEN 1997, JONES et al 2005, EVANS et al. 2008). Sequences were obtained for all strains attempted, except mm178 and Str13, where *cox1* sequencing failed.

All *cox1* and *rbcL* sequences obtained for this study are listed in Table S2, together with their corresponding GenBank accession number. These were complemented with *Sellaphora cox1* and *rbcL* sequences published in EVANS et al. (2007, 2008). (Provisional) species and deme names were given to the sequences following EVANS et al. (2007, 2008) and MANN et al. (2009). (Partial) sequences of some strains have already been published in a study on diatom barcoding (HAMSHER et al. 2011). Sequences were aligned automatically using ClustalW (THOMPSON et al. 1994) as implemented in BioEdit version 7.0.3 (HALL 1999), and checked visually. The ends of the alignment were cut off as these were not available for all sequences, resulting in alignments of 611 and 1245 positions for *cox1* and *rbcL*, respectively. The *cox1* sequences of Cr44 and BB83 and the *rbcL* sequences of Sel492D and mm111 were omitted from the alignment, as a larger stretch of sequence was missing for these strains than for the others.

Tree reconstruction methods were Bayesian Inference (BI) using MrBayes version 3.1.1 (RONQUIST & HUELSENBECK 2003) and Maximum Likelihood (ML) using TreeFinder version of October 2008 (JOB 2008). The base composition was checked to be largely independent of species and location on the gene. Model selection was performed in TreeFinder using the Akaike Information Criterion for each codon position separately. The models considered were HKY, TN, J1, J2, J3 (TIM), TVM, and GTR, with or without a gamma shape parameter (G) with four rate categories and/or a proportion of invariable sites (I). The models selected were J3 + G + I, TVM + G + I, and GTR + G for *rbcL*, and GTR + G, TN + G + I, and GTR + G + I for *cox1* for the first, second and third codon position, respectively. As the J3, TN and TVM model are not available in MrBayes, and Bayesian Analysis is robust to slight overparameterization of the model, we applied the GTR model instead. No initial values were assigned to the model parameters. Two runs of four Markov Chains (one cold and three heated) were run for 11 and 12 million generations for *rbcL* and *cox1* respectively and sampled every 400 generations. This yielded a posterior probability (PP) distribution of 27,501 trees. After exclusion of 7,500 and 10,000 ‘burn-in’ trees for *rbcL* and *cox1* respectively, PPs were calculated by constructing a 50% majority-rule consensus tree. For the ML phylogeny reconstruction, a Maximum Parsimony (MP) tree, obtained in PAUP\* version 4.0b.10 as described in VANORMELINGEN et al. (2007), was used as starting tree. The resulting ML tree was used to generate 200 new starting trees, half of which were 10 steps and the other half 50 steps away from the original tree. The resulting trees, including the original ML tree, were then used as starting trees for the final ML phylogeny reconstruction.

Finally, uncorrected pairwise *p*-distances, a *cox1* MP tree including Cr44 and BB83 and an *rbcL* MP tree including Sel492D were obtained using PAUP\* version 4.0b.10.

**Mating behaviour.** The mating system of *S. auldreekie* was determined in late 1996, using strains isolated in October

1996 by V.A.C. from Blackford Pond and Dunsapie Loch in Edinburgh (Tables S1 and S3). Vouchers for this material are kept in the diatom herbarium of RBGE as accessions E2636–2646 and E3006–3018. Between April 2008 and January 2009, mating trials involving *S. auldreekie* and other *auldreekie*-like species were performed by P.V. using strains from three lineages with a sufficient number of available strains, namely *S. auldreekie* (five strains used out of five sequenced), ‘southern auldreekie’ (nine out of twelve), and ‘coarse auldreekie’ (six out of six) (Table S2). One strain that was not sequenced, Sel686K, was also included. Crossing experiments were carried out in the same conditions used for growing strains, except for the first crossing experiment which was carried out at 16 °C. Because sexual reproduction was not vigorous at that temperature, it was increased to 18 °C. Crossing experiments were performed by mixing cells from exponentially growing cultures (re-inoculated into fresh medium two to three days before the start of the crossing experiment) in pairwise combinations in wells of 24-well plates. They were checked for the presence of sexual reproduction two to three times during the next five to eleven days. Negative controls in each crossing experiment consisted of monoclonal cultures of the different strains. Each well was visually scanned for approximately one minute using an inverted microscope at low magnification (200×), and the intensity of sexual reproduction noted via a simple scoring system (no sexualized cell pairs, 1–5 cases, 5–10 cases, 10s to 100s of cases) as well as the presence of large cells close to the maximal cell size (and thus resulting from previous auxosporulation) and dividing initial cells. To ensure that the absence of sexual reproduction in a particular mixture was due to the strains being incompatible and not to the inability of one of the strains to mate, a negative result was only accepted when both strains in the cross displayed sexual reproduction with at least one of the other strains in the other mixtures of the same crossing experiment. When cells of the two strains in a mixture differed in size, we checked whether the cells in mating pairs differed in size and determined the sex (male or female) of the different strains by observing in which of the parental frustules the zygote or resulting auxospore was found.

Some additional crosses were performed with isolates LE J9 and LE D35, the two available strains of lineage ‘Czech auldreekie’. They were crossed twice with each other and against two strains of each mating type of lineages *S. auldreekie* (Sel776INV, Sel781INV, Cr44, Sel775INV), ‘southern auldreekie’ (KewA3, Sel624K, Sel635K, Sel618K) and one strain of each mating type of ‘coarse auldreekie’ (Sel620K, Sel642K). As positive controls, crosses were made between compatible strains from the same lineage (*S. auldreekie*, ‘southern auldreekie’ and ‘coarse auldreekie’), and negative controls again consisted of the monoclonal cultures.

Initial cell lengths were measured from Sel620K × Sel627K (‘coarse auldreekie’), Sel776INV × Sel775INV (*S. auldreekie*), and Sel629K × Sel618K (‘southern auldreekie’) to determine the maximal cell length ( $n = 20$  for each cross).

**A posteriori examination of morphological differences among clades.** Following the discovery that the *S. auldreekie* complex contained six species, we re-examined the pattern of morphological variation for characters that might be diagnostic for species in LM. Visual inspection of photographs of valves from voucher slides suggested that two characters might have some value, these being stria

density and the width of the valve pole. The width of the valve at the centre was rejected as a character, because the valves of some species and strains were sufficiently flimsy to collapse and spread during mounting; the poles were less affected. Because valve characteristics can change with decreasing size during vegetative growth, which may blur species differences, we analysed pole width and stria density separately for three sets of strains having more or less similar valve length. Given that size ranges of different *auldreekie* species are very similar (see Results section), length can be taken throughout as a proxy for life cycle stage. The first set consisted of nine F1 strains having almost maximal cell size, three derived from the progeny of each of the three following crosses: *S. auldreekie* Sel776INV × Sel775INV, ‘southern auldreekie’ Sel629K × Kew21, and ‘coarse auldreekie’ Sel620K × Sel642K. The measurements were made from material collected some weeks after the initiation of the strains through auxosporulation (Fig. 2). A second subset comprised 12 original strains with medium-sized valves (c. 20 µm, except for mm120 and Cr44: Table S4) and the other comprised six strains with small valves (c. 11 µm: Table S4); an extra criterion for selection of strains was to include strains from different localities, where this was possible (*S. auldreekie* and ‘southern auldreekie’). 10 valves were measured for each strain.

For the F1 subset, nested ANOVA (random factor strain, nested within species) was used to test for species-level differences in pole width and stria density, in the two other subsets one-way ANOVA was applied. Post-hoc Tukey tests were used to outline statistically significant strain groupings, and their correspondence to species limits was evaluated. Statistical analyses were performed using Minitab 16 (www.minitab.com). The significance level was 0.05 throughout.

## RESULTS

### General morphology

All investigated *Sellaphora* strains had a very similar valve morphology (Fig. 1). Valves were narrowly elliptical to elliptical, up to c. 34 µm long (in the longest initial cells), and up to 8 µm wide. Most valves were clearly rostrate, but this characteristic was lost in the smallest valves found in culture (Fig. 1). The largest (post-initial) cells often had slightly capitate poles (Fig. 2). The striae were strongly radiate at the centre, becoming ± parallel at the poles, and the central area was bow-tie-shaped. Stria densities were between 20.5 and 26.0 in 10 µm. The variation observed fell within the limits described for *S. auldreekie* by MANN et al. (2004, 2008), except that the range of stria density was wider (MANN et al. 2004, 2008 reported a maximum of 24 striae in 10 µm).

### Molecular phylogenies

In the *cox1* ML phylogeny of *Sellaphora* (Fig. 3), *cox1* sequences of the *auldreekie*-like *Sellaphora* strains grouped into five well-separated clades with almost no intraclade variation (0 or 1 base pairs [bp] or 0–0.16%), and 4.3–7.9% sequence divergence (15–

46 bp) between them. This is similar to the sequence divergence between other closely related *Sellaphora* species (Fig. 4). A possible sixth clade was represented by a single strain (BB94) and was closely related to the ‘southern auldreekie’ clade (2.5% divergence). One of the clades contained the barcode sequence previously designated for *S. auldreekie* (GenBank EF164932: EVANS et al. 2007; MANN et al. 2008) and can therefore be identified as *S. auldreekie sensu stricto*. The other five clades, presumably corresponding to different species, were given provisional names (*S. pupula* agg. ‘southern auldreekie’, ‘coarse auldreekie’, ‘mumblin auldreekie’, ‘another auldreekie’, and ‘Czech auldreekie’). We considered the recovered part of the *cox1* sequence of Cr44 (415 base pairs) and BB83 (404 bp) as too short to be included in the phylogeny. However, a MP analysis including these sequences showed that they were identical to other sequences from the *S. auldreekie* and ‘southern auldreekie’ clades, respectively, and well-separated from the other clades (not shown). Together, the six clades were monophyletic, albeit with limited posterior probability (PP) and bootstrap support (Fig. 3).

The *rbcL* ML phylogeny (Fig. 5) largely confirmed the patterns seen in the *cox1* phylogeny. The *rbcL* sequences were divided into six clades corresponding to the clades outlined on the basis of *cox1*. Intraspecific variability was limited in five of the clades (0–2 base pairs or 0–0.16%), but extraordinarily large between the Czech and both Scottish strains of ‘Czech auldreekie’ (6 bp, 0.48%), which is even a bit larger than the difference between the two most closely related other *Sellaphora* species (*S. bisexualis* and *S. pupula* agg. ‘upland elliptical’ differ by 4 or 5 bp or 0.32–0.40%). The sequence divergence between the clades was 0.6–2.8% (7–35 bp), again similar to differences found between other closely related *Sellaphora* species (Fig. 4). The recovered parts of the *rbcL* sequence of Sel492D (802 base pairs) and mm111 (748 bp) were too short to be included in the phylogeny. However, a MP analysis including both sequences showed that they were identical to other sequences from the *S. auldreekie* (the same as EF143320) respectively ‘mumblin auldreekie’ clade, and well-separated from the other clades (not shown). Together, the six clades formed a monophyletic group, supported by PP, but not by bootstrap analysis.

### Crossing experiments

Sexual reproduction in *S. auldreekie* was first investigated in natural populations (MANN 1984, as the ‘small’ deme) and then experimentally in 1996 by V.A.C., using strains from Blackford Pond and Dunsapie Loch in Edinburgh. The *rbcL* sequence of one of the strains used in the 1996 work, strain 1B, was sequenced by Evans et al. (2008, EF143320; see also Fig. 5), confirming its identification as *S. auldreekie*. The sexual process proved to be similar to that of

other allogamous *Sellaphora* species (e.g. MANN 1989b; MANN et al. 1999; see Figs. S1 and S2) and the experimental crosses revealed that *S. auldreekie* was heterothallic (Table S3). This information was used in 2008 to design mating experiments among a selection of the new British and Australian strains sequenced during the present study.

All strains used in 2008 were sexually active upon the first cross, except ‘southern auldreekie’ strains Sel618K and Sel630K, even though they were not larger than the largest sexually active ‘southern auldreekie’ strains (Table S2). In a second crossing experiment, respectively two and four weeks later, they proved to be sexually active. The largest sexually inducible strains had average valve lengths of 18.7, 22.1, and 20.9  $\mu\text{m}$  for ‘southern auldreekie’, ‘coarse auldreekie’, and *S. auldreekie*, respectively. Initial cell lengths were rather similar for the three species, being  $29.7 \pm 2.0$  (25.3–33.8)  $\mu\text{m}$  for *S. auldreekie* Sel776INV  $\times$  Sel775INV,  $29.9 \pm 0.9$  (27.9–31.3)  $\mu\text{m}$  for ‘southern auldreekie’ Sel629K  $\times$  Sel618K, and  $31.9 \pm 1.1$  (30.1–33.9)  $\mu\text{m}$  for ‘coarse auldreekie’ Sel620K  $\times$  Sel627K. ‘Coarse auldreekie’ had a somewhat higher mean initial cell length, but this might (in part) be caused by a larger parental cell length: combined parental cell lengths were on average 32.4  $\mu\text{m}$  for ‘coarse auldreekie’, compared to 30.0  $\mu\text{m}$  for *S. auldreekie* and 26.5  $\mu\text{m}$  for ‘southern auldreekie’.

A summary of the crossing experiments between strains of *S. auldreekie* (Fig. S2A), ‘coarse auldreekie’ (Fig. S2B), and ‘southern auldreekie’ (Fig. S2, C and D) is given in Table 1. As in the earlier experiments with *S. auldreekie*, the mating systems were heterothallic and strains behaved consistently as males or females in different crosses. Dividing initial cells or large cells resulting from heterothallic sexual reproduction were seen in several combinations of compatible strains in each of the three lineages. Some strains were capable of rare intraclonal reproduction, as evidenced by the presence of sexual reproduction both in monoclonal controls and between similar-sized cells in mixtures of strains with different cell sizes. The female strains Sel635K and BolinA10 were the only ones in which intraclonal sexual reproduction (and the resulting progeny) could be confidently seen, producing some large cells in monoclonal culture, albeit never in more than a few cases. This contrasts with the 10s to 100s of sexualized cell pairs that were observed in mixtures of compatible strains. In only a few crosses with compatible strains, between Sel642K and BolinA2 or BolinA4, was a somewhat lower number of sexualized cell pairs observed (between 5 and 10 after a short visual scan) (Table 1).

Despite the large number of interlineage crosses (at least six different crosses for each combination of lineage and mating type were made), not a single certain successful interlineage cross was observed (Table 1). At the same time, the strains

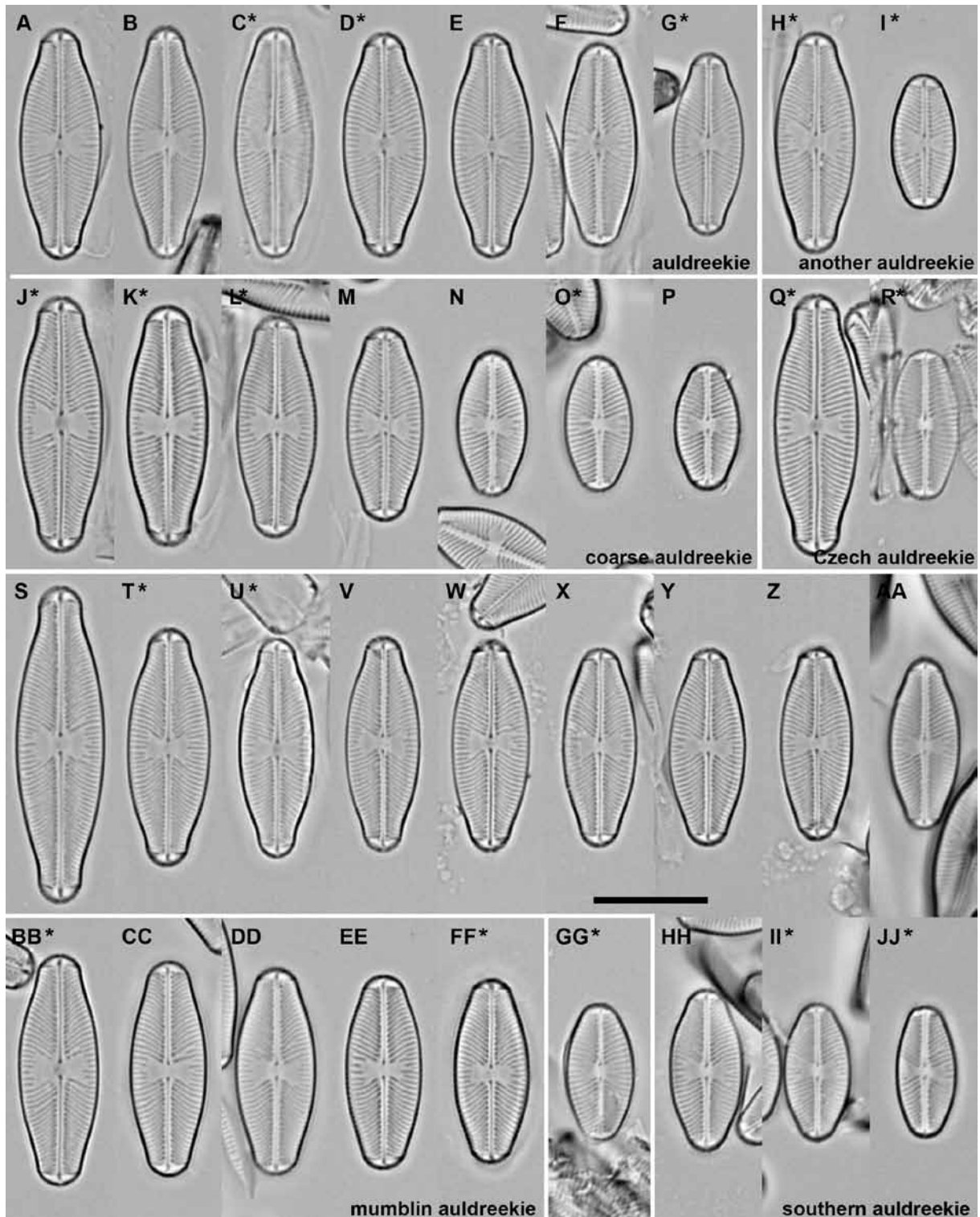


Fig. 1. Valve morphology of *Sellaphora* strains used in the present study. Strains are grouped according to molecular clade, except for the small-celled (later) voucher of Sel492D (Fig. 1, GG), which is placed near small-celled strains of 'southern auldreekie' for comparison. (A) Sel7811NV. (B) LE D41. (C) Sel492D (voucher E4822). (D) Sel7751NV. (E) Sel7761NV. (F) LE F66. (G) Cr44. (H) Str13. (I) BB94. (J) Sel642K. (K) Sel627K. (L) Sel620K. (M) Sel686K. (N) Bolin-A10. (O) BolinA4. (P) BolinA2. (Q) LE D35. (R) LE J9. (S) Kew-27. (T) Kew-21. (U) Str06. (V) Sel630K. (W) Sel629K. (X) Sel635K. (Y) Sel618K. (Z) Sel624K. (AA) Kew-A7. (BB) mm111. (CC) mm185. (DD) mm135. (EE) mm178. (FF) mm120. (GG) Sel492D (voucher E5255). (HH) KewA6. (II) KewA3. (JJ) BB83. Scale bar 10  $\mu$ m.

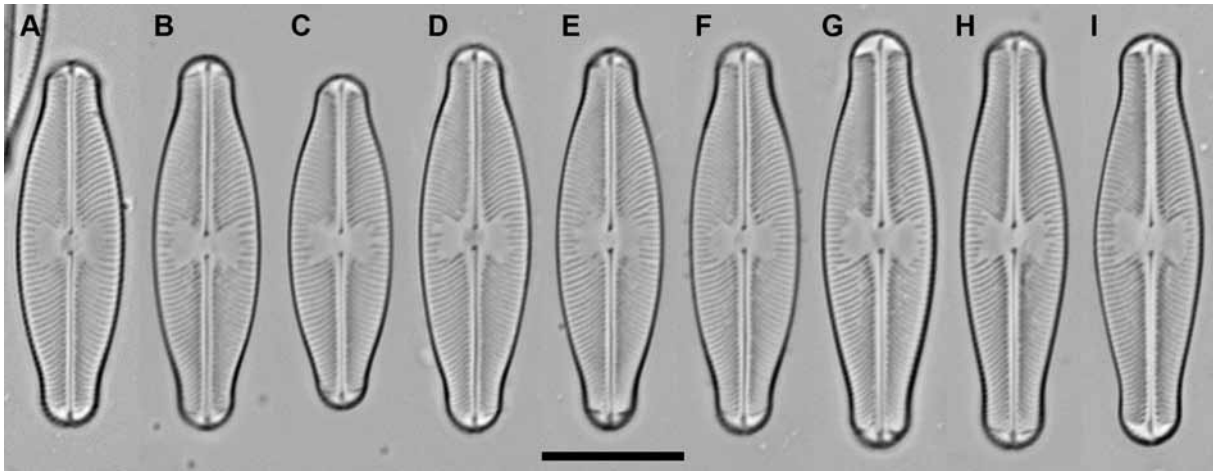


Fig. 2. Valve morphology of *Sellaphora* F1 progeny strains. (A–C) *Sellaphora auldreekie* sensu stricto: progeny of Sel775INV × Sel775INV, F1 clones 1, 2 and 3 respectively (left to right). (D–F) *Sellaphora* ‘southern auldreekie’: progeny of SEL629K × Kew 21, F1 clones 1, 2 and 3 respectively. (G–I) *Sellaphora* ‘coarse auldreekie’: progeny of Sel620K × Sel642K, F1 clones 1, 2 and 3 respectively; note the greater width of the valve poles relative to (A–C) and (D–F). Scale bar 10  $\mu\text{m}$ .

involved mated vigorously with compatible strains from the same lineage. In three combinations, a rare sexualized cell pair could be observed in interlineage crosses but, based on the parental valve lengths, in two cases this could be ascribed to intraclonal sexual reproduction. The third case concerned three pairs of gametangia in the gamete stage in a cross between the similar-sized strains Sel775INV (*S. auldreekie*) and Sel620K (‘coarse auldreekie’) (Table 1). No zygote formation could be detected however. Finally, in some interlineage combinations (involving all three lineages used), several individuals to a few hundreds of unpaired gametangia, i.e. single cells that had become sexualized and undergone gametogenesis, were seen (Table 1). Gamete fusion was never observed in these crosses.

Additional crosses with the two available strains of ‘Czech auldreekie’ (LE J9 and LE D35) showed that they did not mate with strains of the lineages *S. auldreekie*, ‘southern auldreekie’ and ‘coarse auldreekie’, which were at the same time shown to be capable of sexualization using control crosses (data not shown). They also failed to mate with each other, however, possibly because they belong to the same mating type. Their cell sizes were not very large (Table S2) and, by analogy with other *auldreekie*-like species, it is unlikely that the strains were outside the sexual size range. Also here, unpaired gametangia were observed in some crosses (those between LE J9 or LE D35 and Sel776INV).

### A posteriori examination of morphological differences among clades

Differences among lineages in pole width and stria density (Fig. 6, Table S4) were investigated *a posteriori* for three subsets of strains representing different stages in the life cycle, namely F1 progeny strains with near-

maximal cell size, original strains having medium cell size, and original small-celled strains.

For the F1 strains, both strain and species were highly significant sources of variation in pole width (nested ANOVA,  $F = 4.56$ ,  $p \leq 0.001$ , resp.  $F = 33.78$ ,  $p \leq 0.001$ ), with species accounting for 78.6% of total variance. Pairwise Tukey tests showed that ‘southern auldreekie’ and *S. auldreekie* were not significantly different from each other but that both differed from ‘coarse auldreekie’. There was only limited overlap between pole widths of individual valves of ‘coarse auldreekie’ and the other species (Fig. 6A). With respect to stria density, species was a significant source of variation, though accounting for only 28.5% of total variance, whereas strain identity made no contribution (nested ANOVA,  $F = 22.75$ ,  $p \leq 0.01$ , resp.  $F = 0.55$ ,  $p > 0.05$ ). In contrast to pole width, stria density grouped ‘coarse auldreekie’ with *S. auldreekie*, these being statistically inseparable, but separated both from ‘southern auldreekie’, even though stria densities of individual valves overlapped widely between ‘southern auldreekie’ and the other two species (Fig. 6B).

In the Medium Subset, ANOVA of pole width indicated highly significant differences among strains ( $F = 106.43$ ,  $p \leq 0.001$ ), with four statistically separate groupings ( $p < 0.05$ ), which did not completely correspond to the different species (Fig. 6C). In agreement with the F1 strain results, ‘coarse auldreekie’ strains belonged to different groups than *S. auldreekie* and ‘southern auldreekie’, with strains of the latter two species being intermixed. The other species complicate the picture, as the ‘another auldreekie’ strain Str13 had a pole width similar to some *S. auldreekie* and ‘southern auldreekie’ strains, and ‘mumblin auldreekie’ and ‘Czech auldreekie’ strains resembled those of ‘coarse auldreekie’. Strains of the Medium Subset showed highly significant differences with respect to

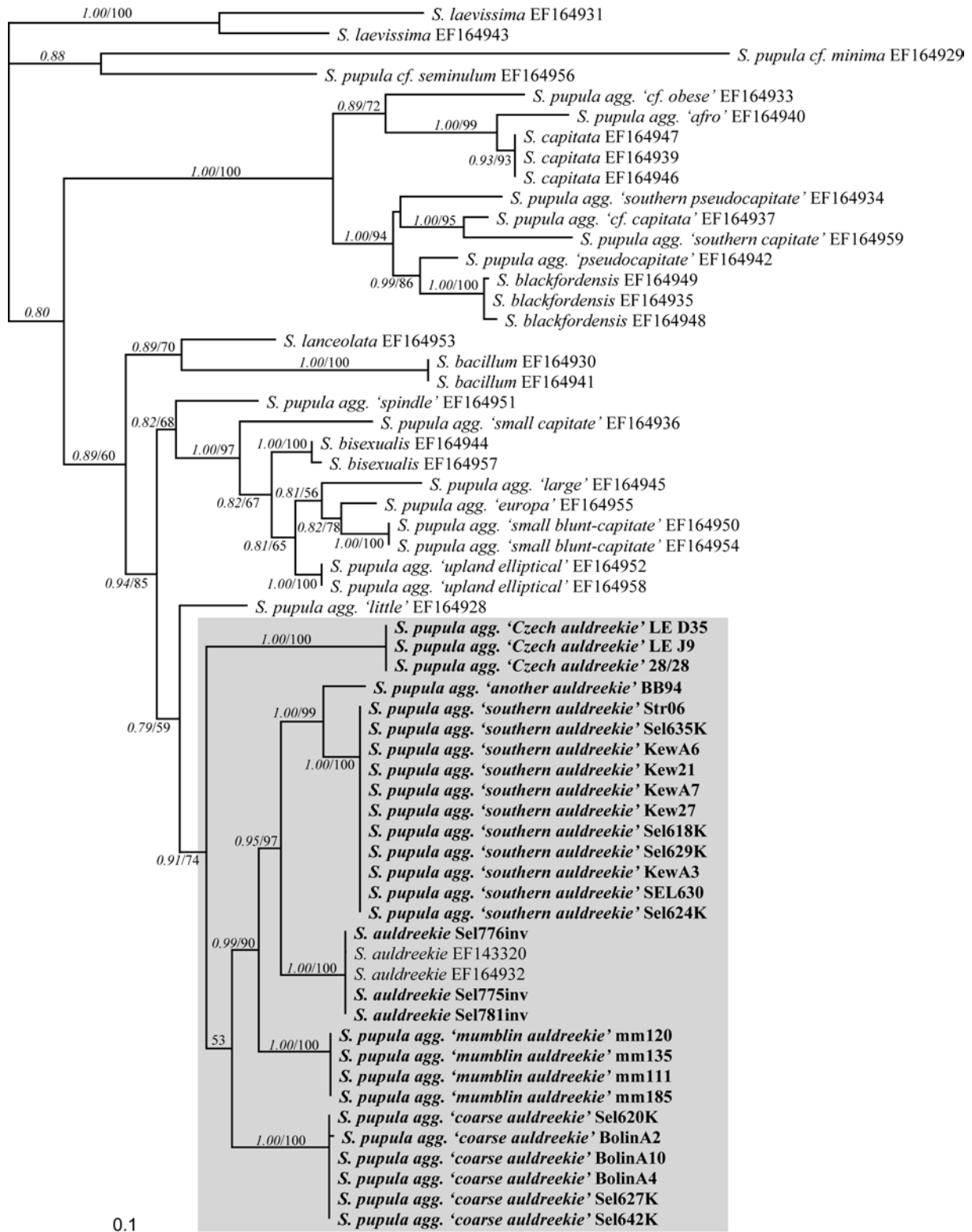


Fig. 3. *Sellaphora* *cox1* Maximum Likelihood (ML) phylogeny. The clade investigated in the present study is indicated with a gray box. Bayesian posterior probabilities (italic) and ML bootstrap values (regular) higher than 0.5 or 50 respectively are indicated at the respective nodes.



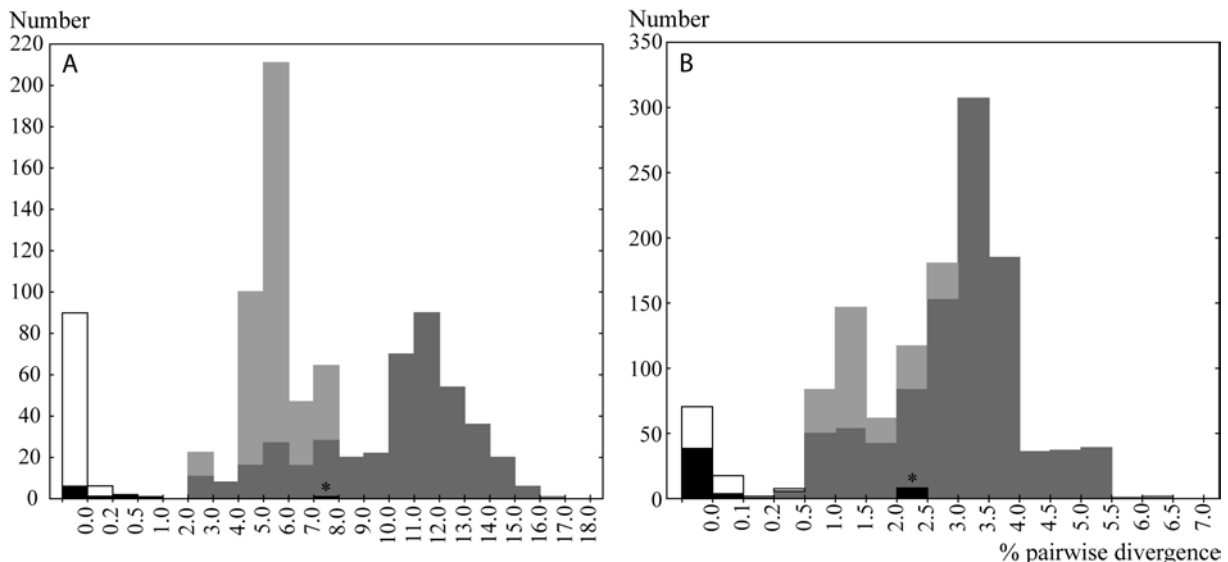


Fig. 4. Histograms of pairwise divergences between *Sellaphora cox1* (A) and *rbcL* (B) sequences. White bars indicate intraspecific differences and light grey interspecific differences between sequences from the six *auldreekie*-like species; black indicates intraspecific differences and dark grey interspecific differences between sequences from other *Sellaphora* species based on Evans et al. (2007, 2008). The extraordinarily large intraspecific differences between sequences of \* *S. laevissima* reflect the fact that this taxon is heterogeneous, since it contains two relatively distantly related lineages.

stria density ( $F = 6.22$ ,  $p \leq 0.001$ ), but the variation was clinal, with only the extremes being statistically separable. The two ‘southern auldreekie’ strains had the highest mean stria densities though.

In the Small Subset similar patterns were seen: three distinct groupings were present with respect to pole width (ANOVA,  $F = 19.56$ ,  $p \leq 0.001$ ), but the order of the species (arranged according to mean pole width) differed from that in the Medium Subset, with ‘another auldreekie’ having the widest poles, rather than ‘Czech auldreekie’ or ‘coarse auldreekie’, which had lower pole widths than in the medium-sized strain set. Stria densities differed among strains (ANOVA,  $F = 4.14$ ,  $p \leq 0.01$ ) but again varied clinally.

## DISCUSSION

The discovery of five new *cox1* lineages among the *auldreekie*-like strains examined, with low intraclade compared to interclade variation (Fig. 3), suggests the presence of five additional species, even though some of the lineages were represented by only a few strains. That the six *cox1* lineages correspond to different species is supported by a phylogeny derived from a chloroplast gene, *rbcL*, in which the same lineages were recovered. Mating trials between three of the lineages, and much more incomplete trials involving a fourth lineage, showed that they are reproductively isolated, providing additional support to the DNA-based species delimitation.

Some residual sexual interactions were seen between strains belonging to different lineages,

gametogenesis being stimulated in unpaired cells, causing the deaths of the cells involved. These were not investigated in detail and it is unknown whether the phenomenon is restricted to particular mating types or certain species (though it must occur in at least two of three species investigated: Table 1). It is noteworthy that the interactions occur between species that are sympatric [cf. crosses between the Kew Billabong clones Sel642K (‘coarse auldreekie’) and Sel629K or Sel624K (‘southern auldreekie’) and could therefore be potentially important in the dynamics of their life cycles in nature. In mating trials between two other *Sellaphora* species with a genetic divergence similar to the species studied here, namely *S. blackfordensis* and *S. pupula* agg. ‘pseudocapitata’, hybridization was found (BEHNKE et al. 2004). The more distantly related species *S. capitata* and *S. blackfordensis*, which occur together in Blackford Pond and Dunsapie Loch, also show some reciprocal stimulation of activity in crosses between male and female clones, but again surface recognition and bonding do not occur between sexualized cells (MANN et al. 1999) and there is no transition to gametogenesis. While some of the *auldreekie*-like species discovered here may be allopatric and thus never come into direct contact with each other, the presence of intrinsic reproductive isolation nevertheless means that the lineages would be incapable of interbreeding even if they did ever come into secondary contact.

All the strains studied are morphologically highly similar, resembling *S. auldreekie* as previously described by MANN et al. (2004, 2008), and also *S. paenepupula*, *S. rostrata* and *S. pupula* ‘tiny’ (see discussion in MANN et al. 2008; pp. 49, 50), for which no

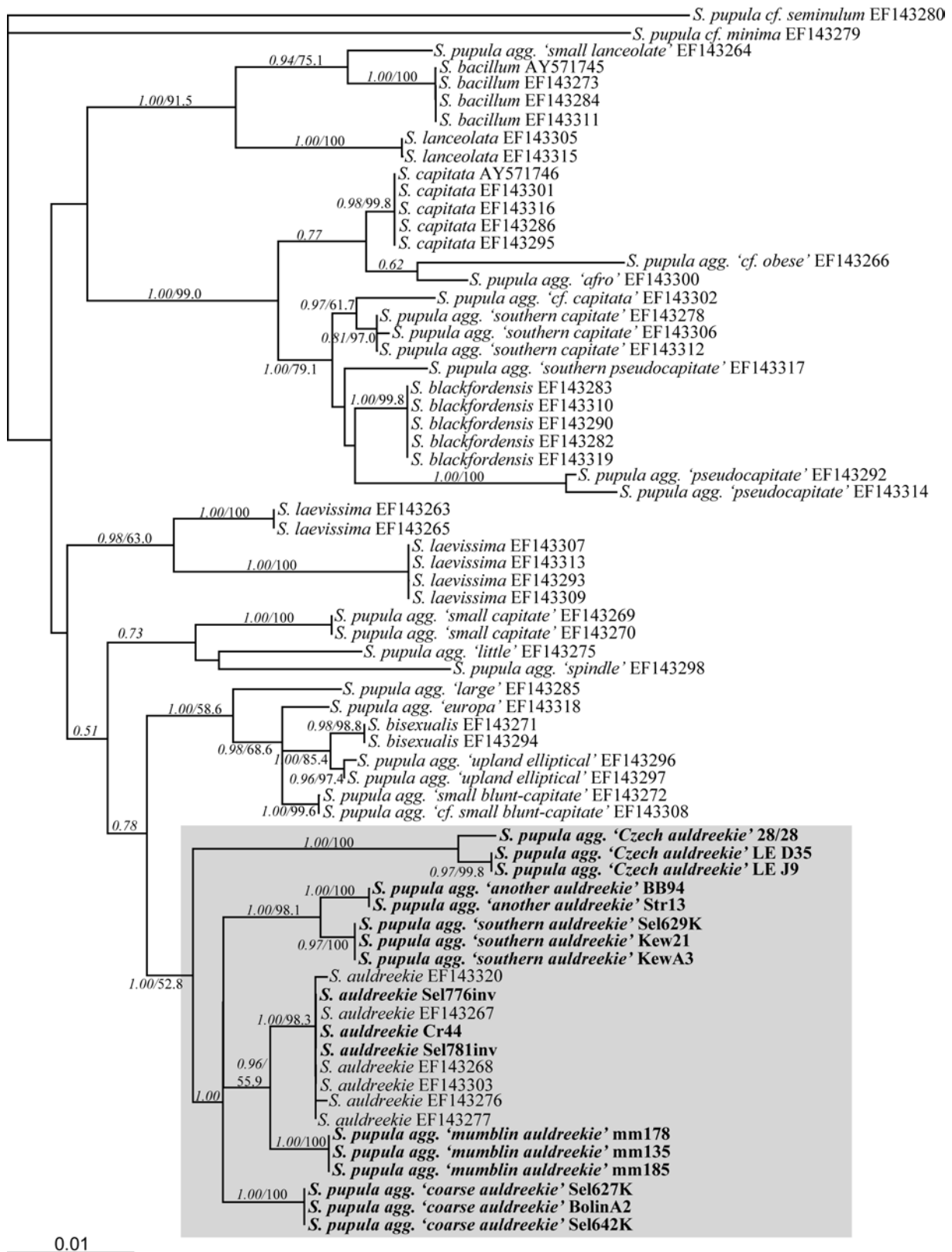


Fig. 5. *Sellaphora rbcL* Maximum Likelihood (ML) phylogeny. The clade investigated in the present study is indicated with a gray box. Bayesian posterior probabilities (italic) and ML bootstrap values (regular) higher than 0.5 or 50 respectively are indicated at the respective nodes.

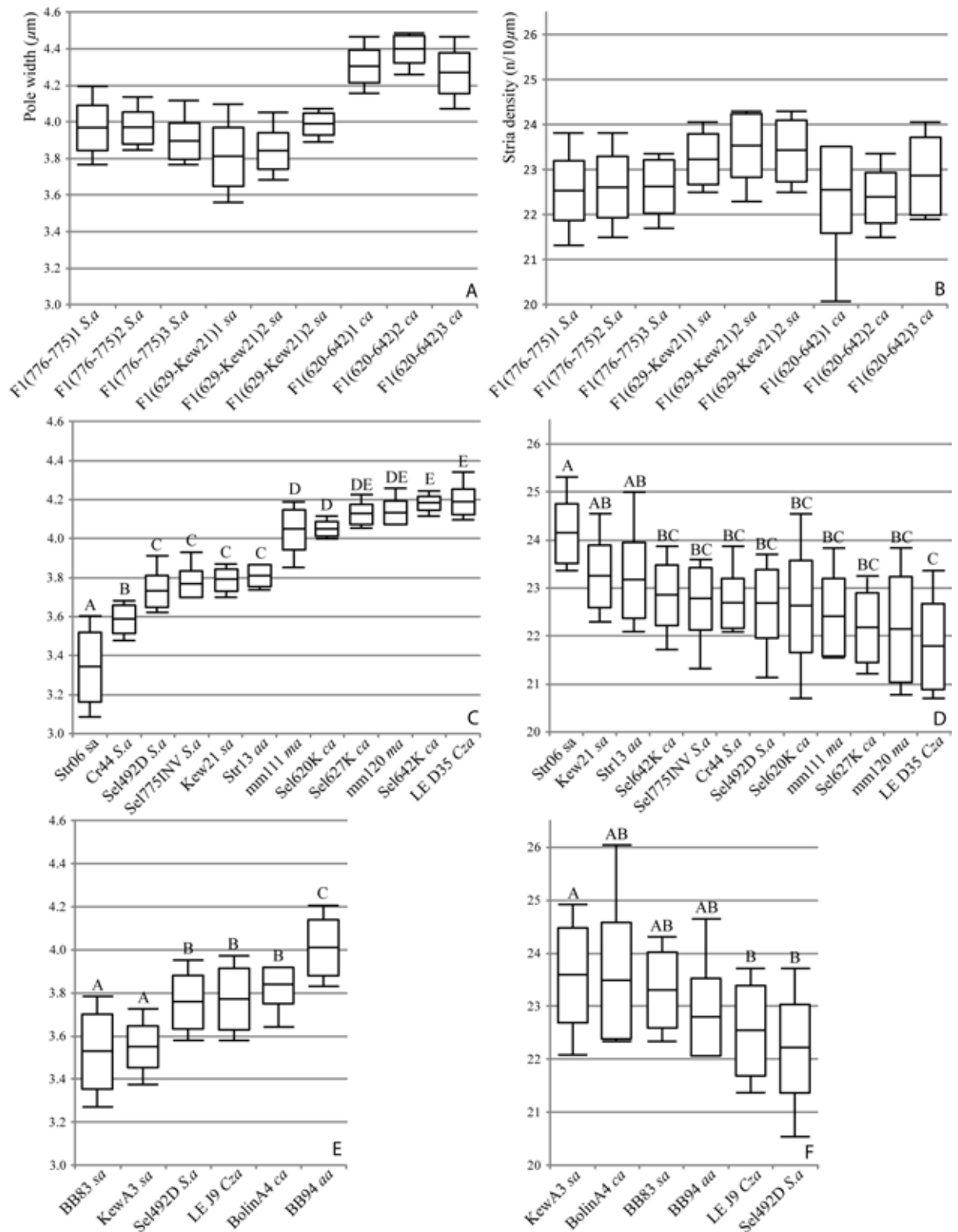


Fig. 6. Pole widths and stria densities of *auldreekie*-like *Sellaphora* F1 strains of maximal cell size (A resp. B), medium (C and D) and small original strains (E and F). For each strain, the mean value, standard deviation (box) and minimum and maximum (whiskers) are given. The results of a post-hoc Tukey test after one-way ANOVA for both groups of original strains are summarized above the box plots of the medium- and small-sized strains using capital letters. Strains that do not share a capital letter are significantly different from each other ( $p \leq 0.05$ ). Species identity of the strains is indicated behind each strain name using a two-letter abbreviation of the species name (*S. auldreekie*) or the provisional name (the other species).

DNA sequence data are available. The size ranges and critical size thresholds of *auldreekie*-like species are very similar in *S. auldreekie*, 'southern auldreekie' and 'coarse auldreekie', the maximum length of initial cells being ~30 µm and the maximum length of sexualizable cells probably c. 22 µm. Based on observations of valve morphology and statistical comparisons of stria density and pole width, there appear to be subtle differences in valve shape and pattern among at least some of the *auldreekie*-like lineages and it is possible that all of the species would be distinguishable using more sophisticated morphometrics than we employed here (DU BUF & BAYER 2002 and MANN et al. 2004 give examples of computer-aided classification and identification of *Sellaphora*, using contour and pattern features). However, it is clear that identification by eye, or with the simple metrics generally used in diatoms (length, width, striation density), is impossible: the ranges of variation among *auldreekie*-like species overlap considerably, even though there may be highly significant differences between mean values for some species. While it is possible to correctly identify strains using mean values for one or a few morphological characters when only two to three species are present, adding more species prevents separation of all species and causes variation to be clinal, as is shown here for pole width and stria density. This is also exemplified in the *Nitzschia palea* species complex (TROBAJO et al. 2009, 2010), where both stria density and valve width vary more or less continuously, despite the presence of genetic discontinuities revealed by sequence analysis and mating data. Another complicating factor is the gradual change in valve morphology with decreasing cell size, the nature of which may differ between species. For instance, while 'coarse auldreekie' and 'Czech auldreekie' clearly have wider poles than *S. auldreekie* when strains have medium-sized to large cells, pole width in the first two species decreases to such an extent with size that small-celled strains cannot be distinguished any more from those of *S. auldreekie*. Morphological analysis of many more strains than only the ones used here, or of natural populations of which the valves can confidently be assigned to species, would be needed to fully understand and compare the changes in valve morphology with decreasing cell size for all *auldreekie*-like species.

As a result, discovery and subsequent routine identification of species of genera such as *Sellaphora* and *Nitzschia*, if based only on valve morphology, is problematic. Indeed, except in one case, we did not realize that additional diversity was present among the *auldreekie*-like strains until molecular data had been obtained; even in the UK, which is relatively well-studied (MANN et al. 2008), an extra species ('Czech auldreekie') was found among strains isolated at RBGE as '*S. auldreekie*'. The exception was 'coarse auldreekie', since one of us (D.G.M), working without knowledge of molecular data or mating success,

picked out three larger-celled strains of this species (Sel642K, Sel627K and Sel620K) as being unlikely to be conspecific with the sympatric clones from Kew Billabong later identified as 'southern auldreekie'. The basis for distinguishing the three strains from other *auldreekie*-like strains from Kew was their greater robustness and coarser striation (compare Fig. 1, J–L, with 'southern auldreekie' in Fig. 1, T, V and W), though measurements of striation density subsequently showed that the differences are actually tiny (compare the Kew clones of 'coarse auldreekie' and 'southern auldreekie' in Table 6). Perhaps not surprisingly, therefore, a further 'coarse auldreekie' clone (Sel686K), with smaller cells than Sel642K, Sel627K and Sel620K, was incorrectly classified by D.G.M. as 'southern auldreekie' until molecular and mating data became available.

The molecular phylogenies obtained here and by EVANS et al. (2007, 2008), as well as their congruence with mating trials, show that *cox1* alone, or in combination with *rbcL*, is suitable for species discovery in *Sellaphora*. Recently, automated phylogenetic species delimitation has been developed, based on a strong increase in branching rates at the transition from species- to population-level evolutionary processes (PONS et al. 2006; MONAGHAN et al. 2009). While this could potentially speed up species delimitation in species-rich groups, it is at the moment not very useful in *Sellaphora* given the near-absence of intraspecific variation with the molecular markers and numbers of strains per species used. Additional markers and/or strains should be sequenced for this purpose.

The pairwise distances between intraspecific and interspecific sequences in *Sellaphora* show almost no overlap, both for *cox1* and *rbcL*, implying that assignment of individuals to species in *Sellaphora* could potentially be based on mere distances, on which for instance the identification engine of the barcode database is currently based ([www.boldsystems.org](http://www.boldsystems.org)). This also seems to be possible for dinoflagellate identification based on the ITS rDNA region (LITAKER et al. 2007). One exception where a very high 'intraspecific' distance was found in *Sellaphora* by EVANS et al. (2007: *S. laevissima*) has proved to reflect heterogeneity of the species concerned (MANN et al. unpublished), like that shown here in *auldreekie*-like diatoms. It remains to be seen whether the so-called barcode gap will still exist in *Sellaphora* when adding more *rbcL* and *cox1* sequences of the same species, especially from distant geographical localities and recently diverged species. In this context it should be noted that the *cox1* and *rbcL* sequences from the single *S. auldreekie* strain from Australia (Cr44) were identical to sequences recovered from the UK. On the other hand, even though their *cox1* sequences were identical, there was an extraordinary large *rbcL* sequence divergence between the Czech and



and ‘southern auldreekie’, ‘another auldreekie’ and ‘coarse auldreekie’ in Bolin Bolin Billabong). Thus, the additional *Sellaphora* diversity encountered here and by EVANS et al. (2007, 2008), when extending the geographical sampling from the UK to Australia, and the diversity present even within individual lakes, suggest that many more *Sellaphora* species can be expected in other geographical areas with different climates and habitat types. Whether this is driven primarily by geographical isolation or by habitat-related differences is uncertain. On the other hand, some species are widely distributed. Most *S. auldreekie* clones have been isolated from Scottish lochs, but a single *S. auldreekie* strain, Cr44, was isolated from Australia, and *S. capitata* and *S. bisexualis* have also been found in both the UK and Australia (EVANS et al. 2009; MANN et al. 2009). The extent to which human-mediated introductions have altered the geographical distribution of *Sellaphora* species remains unknown, but numerous cases of invasive aquatic species are known in Australasia, including some microalgae (HARPER 1994; WELLS et al. 1999; KILROY et al. 2009). New sequencing technologies applied directly to natural samples or community fingerprinting techniques in combination with a DNA barcode database for species identification might in the future be used to tackle the ecology and geographical distribution of these species.

#### ACKNOWLEDGEMENTS

This research was supported by a EU framework 6 SYNTHESYS grant (GB-TAF-4253) to Pieter Vanormelingen. We would like to thank Caroline Souffreau for her help with the molecular analyses. P.V. is a postdoctoral research fellow with the Research Foundation – Flanders (FWO). K.M.E. was supported by a Natural Environment Research Council Fellowship (NE/C518373/1). The first studies of *S. auldreekie* reproduction were supported by INTAS grants 93–3605 and 93–3605-ext to V.A.C. and D.G.M.

#### REFERENCES

- AMATO, A.; KOOISTRA, W.H.C.F.; LEVIALDI GHIRON, J.H.; MANN, D.G.; PRÖSCHOLD, T. & MONTRESOR, M. (2007): Reproductive isolation among sympatric cryptic species in marine diatoms. – *Protist* 158: 193–207.
- BAYER, M.M.; DROOP, S.J.M. & MANN, D.G. (2001): Digital imaging as a tool in phycological research, with especial reference to microalgae. – *Phycol. Res.* 49: 263–274.
- BESZTERI, B.; JOHN, U. & MEDLIN, L.K. (2007): An assessment of cryptic genetic diversity within the *Cyclotella meneghiniana* species complex (Bacillariophyta) based on nuclear and plastid genes, and amplified fragment length polymorphisms. – *Eur. J. Phycol.* 42: 47–60.
- BEHNKE, A.; FRIEDL, T.; CHEPURNOV, V.A. & MANN, D.G. (2004): Reproductive compatibility and rDNA sequence analyses in the *Sellaphora pupula* species complex (Bacillariophyta). – *J. Phycol.* 40: 193–208.
- BICKFORD, D.; LOHMAN, D.J.; SODHI, N.S.; NG, P.K.L.; MEIER, R.; WINKER, K.; INGRAM, K.K. & DAS, I. (2007): Cryptic species as a window on diversity and conservation. – *Trends Ecol. Evol.* 22: 145–155.
- BURNS, J.M.; JANZEN, D.H.; HAJIBABAEI, M.; HALLWACHS, W. & HEBERT, P.D.N. (2008): DNA barcodes and cryptic species of skipper butterflies in the genus *Perichares* in Area de Conservacion Guanacaste, Costa Rica. – *P. Natl. Acad. Sci. U.S.A.* 105: 6350–6355.
- COLEMAN, A.W. (2001): Biogeography and speciation in the *Pandorina/Volvulina* (Chlorophyta) superclade. – *J. Phycol.* 37: 836–851.
- COLEMAN, A.W. (2005): *Paramecium aurelia* revisited. – *J. Eukaryot. Microbiol.* 52: 68–77.
- DARLING, K.F. & WADE, C.M. (2008): The genetic diversity of planktonic foraminifera and the global distribution of ribosomal RNA genotypes. – *Mar. Micropaleontol.* 67: 216–238.
- DAUGBJERG, N. & ANDERSEN, R.A. (1997): Phylogenetic analyses of *rbcL* sequences from haptophytes and heterokont algae suggest their chloroplasts are unrelated. – *Mol. Biol. Evol.* 14: 1242–1251.
- DE QUEIROZ, K. (2007): Species concepts and species delimitation. – *Syst. Biol.* 56: 879–886.
- DESALLE, R.; EGAN, M.G. & SIDDALL, M. (2005): The unholy trinity: taxonomy, species delimitation and DNA barcoding. – *Philos. T. R. Soc. B* 360: 1905–1916.
- DE VARGAS, C.; NORRIS, R.; ZANINETTI, R.; GIBB, S.W. & PAWLOWSKI, J. (1999): Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanographic provinces. – *P. Natl. Acad. Sci. U.S.A.* 96: 2864–2868.
- DU BUF, H. & BAYER, M.M. (2002): Automatic diatom identification. – 316 pp., World Scientific, Singapore.
- ELLEGAARD, M.; GODHE, A.; HÄRNSTRÖM, K. & MCQUOID, M. (2008): The species concept in a marine diatom: LSU rDNA-based phylogenetic differentiation in *Skeletonema marinoi/dohrnii* (Bacillariophyceae) is not reflected in morphology. – *Phycologia* 47: 156–167.
- EVANS, K.M. & MANN, D.G. (2009): A proposed protocol for nomenclaturally effective DNA barcoding of microalgae. – *Phycologia* 48: 70–74.
- EVANS, K.M.; WORTLEY, A.H. & MANN, D.G. (2007): An assessment of potential diatom “barcode” genes (*cox1*, *rbcL*, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). – *Protist* 158: 349–364.
- EVANS, K.M.; WORTLEY, A.H.; SIMPSON, G.E.; CHEPURNOV, V.A. & MANN, D.G. (2008): A molecular systematic approach to explore diversity in the *Sellaphora pupula* species complex (Bacillariophyta). – *J. Phycol.* 44: 215–231.
- GÓMEZ, A.; WRIGHT, P.J.; LUNT, D.H.; CANCINO, J.M.; CARVALHO, G.R.; HUGHES, R.N. (2007): Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozoans taxon. – *P. Roy. Soc. B – Biol. Sci.* 274: 199–207.
- GUILLARD, R.R.L. & LORENZEN, C.L. (1972): Yellow-green algae with chlorophyllide c. – *J. Phycol.* 8: 10–14.
- HALL, T.A. (1999): BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. – *Nucleic Acids Symp. Ser.* 41:95–98.
- HAMSHER, S.E.; EVANS, K.M.; MANN, D.G.; POULÍČKOVÁ,

- A. & SAUNDERS, G.W. (2011): Barcoding diatoms: exploring alternatives to COI–5P. – *Protist* 162: 405–422.
- HARPER, M.A. (1994): Did Europeans introduce *Asterionella formosa* Hassall to New Zealand? – In: KOCIOLEK, J.P. (ed.): Proceedings of the 11th International Diatom Symposium. – pp. 479–484, California Academy of Sciences, San Francisco.
- JAHN, R.; MANN, D.G.; EVANS, K.M. & POULÍČKOVÁ, A. (2008): The identity of *Sellaphora bacillum* (Ehrenberg) D.G. Mann. – *Fottea* 8: 121–124.
- JARGEAT, P.; MARTOS, F.; CARRICONDE, F.; GRUYTA, H.; MOREAU, P.–A. & GARDES, M. (2010): Phylogenetic species delimitation in ectomycorrhizal fungi and implications for barcoding: the case of the *Trichosoma sculpturatum* complex (Basidiomycota). – *Mol. Ecol.* 19: 5216–5230.
- JOBB, G. (2008): *TREEFINDER* version of October 2008. Munich, Germany. Distributed by the author at [www.treefinder.de](http://www.treefinder.de).
- JONES, H.M.; SIMPSON, G.E.; STICKLE, A.J. & MANN, D.G. (2005): Life history and systematics of *Petronopsis* (Bacillariophyta) with special reference to British waters. – *Eur. J. Phycol.* 40: 61–87.
- KAERIYAMA, H.; KATSUKI, E.; OTSUBO, M.; YAMADA, M.; ICHIMI, K.; TADA, K. & HARRISON, P.J. (2011): Effects of temperature and irradiance on growth of strains belonging to seven *Skeletonema* species isolate from Dokai Bay, southern Japan. – *Eur. J. Phycol.* 45: 113–124.
- KILROY, C.; LARNED, S.T. & BIGGS, B.J.F. (2009): The non-indigenous diatom *Didymosphenia geminata* alters benthic communities in New Zealand rivers. – *Freshwater Biol.* 54: 1990–2002.
- KIM, E.; WILCOX, L.; GRAHAM, L. & GRAHAM, J. (2004): Genetically distinct populations of the dinoflagellate *Peridinium limbatum* in neighboring northern Wisconsin lakes. – *Microb. Ecol.* 48: 521–527.
- KRAMMER, K. & LANGE–BERTALOT, H. (1986): Bacillariophyceae 1. Teil: Naviculaceae. – In: Ettl, H.; Gerloff, J.; Heynig, H. & Moltenhauer, D. (eds): Süßwasserflora von Mitteleuropa 2/1. – 876 pp., Gustav Fischer Verlag, Stuttgart.
- LEWIS, L.A. & FLECHTNER, V.R. (2004): Cryptic species of *Scenedesmus* (Chlorophyta) from desert soil communities of western North America. – *J. Phycol.* 40: 1127–1137.
- LI, H.; YANG, G.; SUN, Y.; WU, S. & ZHANG, X. (2007): *Cylindrotheca closterium* is a species complex as was evidenced by the variations of *rbcL* gene and SSU rDNA. – *J. Ocean Univ. China* 6: 167–174.
- LITAKER, R.W.; VANDERSEA, M.W.; KIBLER, S.R.; REECE, K.S.; STOKES, N.A.; LUTZONI, F.M.; YONISH, B.A.; WEST, M.A.; BLACK, M.N.D. & TESTER, P.A. (2007): Recognizing dinoflagellate species using ITS rDNA sequences. – *J. Phycol.* 43: 344–355.
- MANN, D.G. (1984): Observations on copulation in *Navicula pupula* and *Amphora ovalis* in relation to the nature of diatom species. – *Ann. Bot. – London* 54: 429–438.
- MANN, D.G. (1989a): The species concept in diatoms: evidence for morphologically distinct, sympatric gamodemes in four epipelagic species. – *Plant Syst. Evol.* 164: 215–237.
- MANN, D.G. (1989b): The diatom genus *Sellaphora*: separation from *Navicula*. – *British Phycol. J.* 24: 1–20.
- MANN, D.G. (1999): The species concept in diatoms. – *Phycologia* 38: 437–495.
- MANN, D.G. (2008): The genus *Sellaphora*: an addition and corrections. – *Fottea* 8: 117–119.
- MANN, D.G.; CHEPURNOV, V.A. & DROOP, S.J.M. (1999): Sexuality, incompatibility, size variation and preferential polyandry in natural populations and clones of *Sellaphora pupula* (Bacillariophyceae). – *J. Phycol.* 35: 152–170.
- MANN, D.G. & EVANS, K.M. (2008): The species concept and cryptic diversity. – In: MOESTRUP, Ø. (ed.): Proceedings of the 12<sup>th</sup> International Conference on Harmful Algae. – pp. 262–268, International Society for the Study of Harmful Algae and Intergovernmental Oceanographic Commission of UNESCO, Copenhagen.
- MANN, D.G.; EVANS, K.M.; CHEPURNOV, V.A. & NAGAI, S. (2009): Morphology and formal description of *S. bisexualis* sp. nov. (Bacillariophyta). – *Fottea* 9: 199–209.
- MANN, D.G.; McDONALD, S.M.; BAYER, M.M.; DROOP, S.J.M.; CHEPURNOV, V.A.; LOKE, R.E.; CIOBANU, A. & DU BUF, J.M.H. (2004): The *Sellaphora pupula* species complex (Bacillariophyceae): morphometric analysis, ultrastructure and mating data provide evidence for five new species. – *Phycologia* 43: 459–482.
- MANN, D.G.; POULÍČKOVÁ, A.; SATO, S. & EVANS, K.M. (2011): Scaly incunabula, auxospore development and girdle polymorphism in *Sellaphora marvanii* sp. nov. (Bacillariophyceae). – *J. Phycol.* 47: 1368–1378.
- MANN, D.G.; THOMAS, S.J. & EVANS, K.M. (2008): Revision of the diatom genus *Sellaphora*: a first account of the larger species in the British Isles. – *Fottea* 8: 15–78.
- MONAGHAN, M.T.; WILD, R.; ELLIOT, M.; FUJISAWA, T.; BALKE, M.; INWARD, D.J.; LEES, D.C.; RANAIVOSOLO, R.; EGGLETON, P.; BARRACLUGH, T.G. & VÖGLER, A.P. (2009): Accelerated species inventory on Madagascar using coalescent–based models of species delineation. – *Syst. Biol.* 58: 298–311.
- MONIZ, M.B.J. & KACZMARSKA, I. (2009): Barcoding diatoms: Is there a good marker? – *Mol. Ecol. Notes* 9: 65–74.
- MONTRESOR, M.; SGROSSO, S.; PROCACCINI, G. & KOOISTRA, W.H.C.F. (2003): Intraspecific diversity in *Scrippsiella trochoidea* (Dinophyceae): evidence for cryptic species. – *Phycologia* 42: 56–70.
- PAWLOWSKI, J. & HOLZMANN, M. (2008): Diversity and geographical distribution of benthic foraminifera: a molecular perspective. – *Biodivers. Conserv.* 17: 317–328.
- PONS, J.; BARRACLUGH, T.G.; GOMEZ–ZURITA, J.; CARDOSO, A.; DURAN, D.P.; HAZELL, S.; KAMOUN, S.; SUMLIN, W.D. & VÖGLER, A.P. (2006): Sequence–based species delimitation for the DNA taxonomy of undescribed insects. – *Syst. Biol.* 55: 595–609.
- POULÍČKOVÁ, A.; MAYAMA, S.; CHEPURNOV, V.A. & MANN, D.G. (2007): Heterothallic auxosporulation, incunabula and perizonium in *Pinnularia* (Bacillariophyceae). – *Eur. J. Phycol.* 42: 367–390.
- POULÍČKOVÁ, A.; ŠPAČKOVÁ, J.; KELLY, M.G.; DUCHOSLAV, M. & MANN, D.G. (2008): Ecological variation within *Sellaphora* species complexes (Bacillariophyceae): specialists or generalists? – *Hydrobiologia* 614:

- 373–386.
- RONQUIST, F. & HUELSENBECK, J.P. (2003): MRBAYES 3: Bayesian phylogenetic inference under mixed models. – *Bioinformatics* 19: 1572–1574.
- ROUND, F.E.; CRAWFORD, R.M. & MANN, D.G. (1990): The diatoms. Biology and morphology of the genera. – 747 pp., Cambridge University Press, Cambridge, UK.
- ŠLAPETA, J.; LÓPEZ-GARCÍA, P. & MOREIRA, D. (2005): Global dispersal and ancient cryptic species in the smallest marine eukaryotes. – *Mol. Biol. Evol.* 23: 23–29.
- SOUFFREAU, C.; VANORMELINGEN, P.; VAN DE VIJVER, B.; ISHEVA, T.; VERLEYEN, E.; SABBE, K. & VYVERMAN, W. (2012): Molecular evidence for distinct Antarctic lineages in the cosmopolitan terrestrial diatoms *Pinnularia borealis* and *Hantzschia amphioxys*. – *Protist* 164: 101–115.
- THOMPSON, J.D.; HIGGINS, D.G. & GIBSON, T.J. (1994): CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. – *Nucleic Acids Res.* 22: 4673–4680.
- TRAINOR, F.R. (1998): Biological aspects of *Scenedesmus* (Chlorophyceae) – phenotypic plasticity. – *Nova Hedwigia, Beih.* 117: 1–367.
- TROBAJO, R.; CLAVERO, E.; CHEPURNOV, V.A.; SABBE, K.; MANN, D.G.; ISHIHARA, S. & COX, E.J. (2009): Morphological, genetic, and mating diversity within the widespread bioindicator *Nitzschia palea* (Bacillariophyceae). – *Phycologia* 48: 443–459.
- TROBAJO, R.; MANN, D.G.; CLAVERO, E.; EVANS, K.M.; VANORMELINGEN, P. & MCGREGOR, R.C. (2010): The use of partial *cox1*, *rbcL* and LSU rDNA sequences for phylogenetics and species identification within the *Nitzschia palea* species complex (Bacillariophyceae). – *Eur. J. Phycol.* 45: 413–425.
- TRONHOLM, A.; STEEN, F.; TYBERGHEN, L.; LELIAERT, F.; VERBRUGGEN, H.; SIGUAN, M.A.R. & DE CLERCK, O. (2010): Species delimitation, taxonomy, and biogeography of *Dictyota* in Europe (Dictyotales, Phaeophyceae). – *J. Phycol.* 46: 1301–1321.
- VANELSLANDER, B.; CRÉACH, V.; VANORMELINGEN, P.; ERNST, A.; CHEPURNOV, V.A.; SAHAN, E.; MUYZER, G.; STAL, L.J.; VYVERMAN, W. & SABBE, K. (2009): Ecological differentiation between sympatric pseudocryptic species in the estuarine benthic diatom *Navicula phyllepta* (Bacillariophyceae). – *J. Phycol.* 45: 1278–1289.
- VANORMELINGEN, P.; CHEPURNOV, V.A.; MANN, D.G.; SABBE, K. & VYVERMAN, W. (2008b): Genetic divergence and reproductive barriers between morphologically heterogeneous sympatric clones of *Eunotia bilunaris* sensu lato (Bacillariophyta). – *Protist* 159: 73–90.
- VANORMELINGEN, P.; HEGEWALD, E.; BRABAND, A.; KITSCHKE, M. & FRIEDL, T. (2007): The systematics of a small spineless *Desmodesmus* species, *D. costato-granulatus* (Sphaeropleales, Chlorophyceae), based on ITS2 rDNA sequence analyses and cell wall morphology. – *J. Phycol.* 43: 378–396.
- VANORMELINGEN, P.; VERLEYEN, E. & VYVERMAN, W. (2008a): The diversity and distribution of diatoms: from cosmopolitanism to narrow endemism. – *Biodivers. Conserv.* 17: 393–405.
- VANORMELINGEN, P.; VANELSLANDER, B.; SATO, S.; GILLARD, J.; TROBAJO, R.; SABBE, K. & VYVERMAN, W. (2013): Heterothallic sexual reproduction in the model diatom *Cylindrotheca*. – *Eur. J. Phycol.* 48: 93–105.
- WELLS, R.D.S.; HALL, J.A.; CLAYTON, J.S.; CHAMPION, P.D.; PAYNE, G.W. & HOFSTRA, D.E. (1999): The rise and fall of water net (*Hydrodictyon reticulatum*) in New Zealand. – *J. Aquat. Plant Manage.* 37: 49–55.
- ZIMMERMAN J.; JAHN R. & GEMEINHOLZER B. (2011): Barcoding diatoms; evaluation of the V4 subregion on the 18S rRNA gene, including new primers and protocols. – *Org. Diversity Evol.* 11: 173–192.
- ZWART, G.; HUISMANS, R.; VAN AGTERVELD, M.P.; VAN DE PEER, Y.; DE RIJK, P.; EENHOORN, H.; MUYZER, G.; VAN HANNEN, E.J.; GONS, H.J. & LAANBROEK, H.J. (1998): Divergent members of the bacterial division Verrucomicrobiales in a temperate freshwater lake. – *FEMS Microb. Ecol.* 25: 159–165.

#### Supplementary material

the following supplementary material is available for this article:

Fig. S1. The sexual process in *Sellaphora* cf. *auldreekie* in natural populations from Blackford Pond and Lochend Loch, Edinburgh.

Fig. S2. Interclonal crosses of auldreekie-like *Sellaphora* species. (A) *Sellaphora auldreekie* sensu stricto, Se-1775INV (larger cells) × Cr44: early stage in pairing, before chloroplast rearrangement.

Table S1. Source localities from which *Sellaphora* strains used in this study were isolated.

Table S2. *Sellaphora* strain information. Strains indicated in gray were used in the crossing experiments 2008–2009; n/a not applicable, n/d no data. Accession numbers in bold are new to this study, others were published in Hamsher et al. (2011).

Table S3. Results of crosses made on 31 October 1996 between strains of *Sellaphora auldreekie* from Blackford Pond (clones with –B suffix) and Dunsapie Loch (clones with –D suffix), Edinburgh.

Table S4. Valve length, midwidth, pole width, and striation density of auldreekie-like *Sellaphora* strains used in this study. Values are mean ± standard deviation (range), n = 10 per strain. The column ‘subset’ indicates the belonging of the strain to the different size groupings made for the a posteriori morphological analysis.

This material is available as part of the online article (<http://fottea.czechphycology.cz/contents>)