The Journal of
Eukaryotic Microbiology

Published by the International Society of Protistologists

Journal of Eukaryotic Microbiology ISSN 1066-5234

ORIGINAL ARTICLE

Evidence for Gene Duplication and Allelic Codominance (not Hierarchical Dominance) at the Mating-Type Locus of the Ciliate, *Euplotes crassus*

Adriana Vallesi[a], Claudio Alimenti[a], Sergio Federici[b], Graziano Di Giuseppe[b], Fernando Dini[b], Graziano Guella[c] & Pierangelo Luporini[a]

a Laboratory of Eukaryotic Microbiology and Animal Biology, University of Camerino, 62032, Camerino, MC, Italy
b Department of Biology, University of Pisa, 56100, Pisa, Italy
c Department of Physics, University of Trento, 38123, Povo, TN, Italy

Keywords
Cell chemical signals; ciliate mating types; conjugation; macronuclear gene-size DNA molecules; protein pheromones.

Correspondence
P. Luporini, Laboratory of Eukaryotic Microbiology and Animal Biology, via Gentile III da Varano, University of Camerino, 62032, Camerino, MC, Italy
Telephone number: +39 0737 403229;
FAX number: +39 0737 403290;
e-mail: piero.luporini@unicam.it

Received: 22 January 2014; revised 25 March 2014; accepted May 3, 2014.

doi:10.1111/jeu.12140

ABSTRACT

The high-multiple mating system of *Euplotes crassus* is known to be controlled by multiple alleles segregating at a single locus and manifesting relationships of hierarchical dominance, so that heterozygous cells would produce a single mating-type substance (pheromone). In strain L-2D, now known to be homozygous at the mating-type locus, we previously identified two pheromones (E*c*-α and E*c*-1) characterized by significant variations in their amino acid sequences and structure of their macronuclear coding genes. In this study, pheromones and macronuclear coding genes have been analyzed in strain POR-73 characterized by a heterozygous genotype and strong mating compatibility with L-2D strain. It was found that POR-73 cells contain three distinct pheromone coding genes and, accordingly, secrete three distinct pheromones. One pheromone revealed structural identity in amino acid sequence and macronuclear coding gene to the E*c*-α pheromone of L-2D cells. The other two pheromones were shown to be new and were designated E*c*-2 and E*c*-3 to denote their structural homology with the E*c*-1 pheromone of L-2D cells. We interpreted these results as evidence of a phenomenon of gene duplication at the *E. crassus* mating-type locus, and lack of hierarchical dominance in the expression of the macronuclear pheromone genes in cells with heterozygous genotypes.

EUPLOTES species, like other species of hypotrichs and stichotrichs, evolved high-multiple (virtually open) mating-type systems (Dini and Luporini 1985; Dini and Nyberg 1993; Nobili et al. 1978; Phadke and Zufall 2009; Valbonesi et al. 1992). In studying the Mendelian genetics of these systems, *Euplotes patella* and *Euplotes crassus*, together with its phylogenetically closely allied *Euplotes minuta* and *Euplotes vannus*, have traditionally served as reference species. Based on results of mating tests—directed to detect the presence or absence of conjugant pairs in cell mixtures—as a unique criterion for determining the phenotype "mating type", it was concluded that, (i) the multiple mating types (Mt I, Mt II, Mt III and so forth) of these species are equally controlled by series of alleles at a single locus (annotated as *mt*, or *mat* locus) in the cell germinal, transcriptionally inert micronucleus (Kimball 1942; Heckmann 1964, 1967; Nobili 1966) and (ii) each *mat* allele (*mat-1*, *mat-2*, *mat-3,* and so forth) encodes, through its transcriptionally active copies in the cell macronucleus, a mating-type specific factor, or signaling pheromone. However, while the *E. patella* model implicates relationships of codominance among the *mat* alleles (Kimball 1942), the *E. crassus* model implicates relationships of hierarchical (or serial) dominance, which has usually been indicated by writing either *mat-n* > *mat-3* > *mat-2* > *mat-1*, or *mat-1* > *mat-2* > *mat-3* > *mat-n* (Dini and Nyberg 1993; Heckmann 1964; Nobili et al. 1978).

At the origin of these implications are clear differences between *E. patella* and *E. crassus* in the mating behavior of cells with heterozygous genotypes. The *E. patella*

codominance pattern accounts for the capacity of heterozygous cells to form mating pairs with *both* their homozygous counterparts, which implies that they produce *both* the pheromones that are encoded by their two *mat* alleles. Instead, the *E. crassus* pattern of hierarchical dominance accounts for the fact that cells with a heterozygous genotype form mating pairs with *only one* of the two homozygous counterparts, thus behaving like homozygous cells. This implies that the pheromone production by *E. crassus* heterozygous cells is restricted to *only one* pheromone, i.e. the pheromone that is specified by the dominant *mat* allele.

Direct analysis of these implications was made possible by *Euplotes* pheromone isolation and structural characterization, first, in *E. raikovi* (Miceli et al. 1983) and *E. octocarinatus* (Schulze-Dieckhoff et al. 1987; Weischer et al. 1985) which are species characterized by an *E. patella*-like mating pattern, then in *E. crassus* itself (Alimenti et al. 2011). The *E. patella* pattern was fully validated by the demonstration, in *E. raikovi* in particular, that cells with heterozygous genotypes effectively secrete, usually in eccentric proportions, two structurally homologous protein pheromones, in full accord with their determination by alleles at the same *mat* locus (Luporini and Miceli 1986; Luporini et al. 1986). Instead, the *E. crassus* pattern was challenged by observing that two water-borne pheromones (designated E*c*-α and E*c*-1) isolated from cell filtrates of the wild-type strain L-2D (carrying an undetermined mating-type genotype) were characterized by marked differences not only in their amino acid sequences, but also in the structure of their macronuclear coding genes that, in the form of individual gene-sized DNA molecules, represent the transcriptionally active copies of the transcriptionally silent micronuclear *mat* alleles (Alimenti et al. 2011). On the basis of these differences, it was argued that the E*c*-α and E*c*-1 pheromone determination did not involve two alleles of the same gene, but it involved two alleles of distinct genes.

The validity of this hypothesis is supported here by a pheromone and macronuclear pheromone gene analysis in *E. crassus* cells of the wild-type strain POR-73 characterized by a heterozygous mating-type genotype, as deduced from studying the Mendelian segregation of the mating-type phenotype among offspring clones of a cross between these cells and the strongly mating compatible L-2D cells. From this analysis it appeared that POR-73 cells contain three distinct macronuclear pheromone genes, each responsible for the production of a distinct water-borne pheromone. Since one of the three genes showed structural identity with the E*c*-α pheromone gene previously characterized in L-2D cells and the other two genes appeared to be allelic forms of the E*c*-1 pheromone gene of L-2D cells, it was concluded that *E. crassus* is characterized by a gene duplication at the mating-type locus and, as is the case in the *E. patella* model, cells with heterozygous genotypes express their macronuclear pheromone genes without relationships of serial dominance.

MATERIALS AND METHODS

Cell cultures

The *E. crassus* POR-73 and L-2D strains are cultivated in the laboratory since July 1984 and March 1978, respectively. They are vigorous remnants of a large group of more than 100 wild-type strains originally used in a study designed to show direct and indirect genetic connections among natural populations of ciliate species characterized, such as *E. crassus*, by high-multiple (virtually open) mating systems (Valbonesi et al. 1992). The *E. raikovi* strain 13 is cultivated since June 1979 (Miceli et al. 1981), and is the strain originally used to isolate and characterize *Euplotes* pheromones (Miceli et al. 1983). It is deposited at the ATCC Center with the catalog # PRA-327. All the three strains were expanded from single isolates collected from shallow sandy bottom sites rich in organic sediment, located in Porto Recanati (Eastern coast of Italy, 43°26′N, 13°40′E) in the case of strains POR-73 and 13, and Leghorn (Western coast of Italy, 43°33′N, 10°19′E) in the case of strain L-2D. Their cultures (available upon request) are periodically recloned and maintained, at 20–22 °C, under a natural cycle of weak light and dark on green algae *Dunaliella tertiolecta* (or *D. salina*) grown in natural, or artificial sterilized sea water enriched with Walne medium (Walne 1966).

Pheromone preparation and molecular mass determination

Pheromones were purified from cell-free filtrates of 5-liter cultures grown for 1–2 wk in the continuous presence of food at concentrations of approximately 5,000 cells/ml and then deprived of food for 2–3 d. The purification was based on a slightly modified version of a routine chromatographic procedure originally devised for the purification of *E. raikovi* pheromones (Raffioni et al. 1987). It involved: (i) protein adsorption onto Sep-Pak C_{18} cartridges (Waters, Milford, CA) and cartridge elution with 2 ml of 30% (v/v) 2-propanol; (ii) protein fractionation on a Superdex-Peptide column (GE Healthcare, Little Chalfont, UK) with 50 mM Tris-HCl, pH 7.8, containing 0.5 M NaCl, at a constant flow rate of 0.25 ml/min; (iii) pooling of the fractions containing 4–12 kDa proteins; and (iv) fractionation of the pooled material by reverse-phase high performance liquid chromatography (RP-HPLC) on a 4.6 × 250 mm C_{18} column (Supelco, Bellefonte, PA) eluted with a discontinuous acetonitrile gradient in presence of 0.1% (v/v) trifluoroacetic acid at a constant flow rate of 0.6 ml/min.

MALDI-TOF measurements were performed on a Bruker Daltonics Ultraflex MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany). The acceleration voltage was set at 20 kV. For desorption of the components, a nitrogen laser beam (λ = 337 nm) was focused on the template. The laser power level was adjusted to obtain high signal-to-noise ratios, while ensuring minimal fragmentation of the parent ions. All measurements were

carried out in the delayed extraction mode. After crystallization at ambient conditions, positive-ion spectra were acquired in the linear mode, giving mainly singly protonated molecular ions ([M + H]+). Samples were directly applied onto the stainless-steel spectrometer plate as 1-μl droplets, followed by the addition of 1 μl of synapinic acid-matrix solution (0.5 M in methanol/water 1:1 containing 0.1% TFA). Every mass spectrum represents the average of about 100 single laser shoots.

Mating induction assays

Purified pheromone preparations were assayed for their mating inducing activity after dilution in 0.5 ml volumes of sterilized sea water, at serial concentrations from 5×10^{-5} to 5×10^{-8} M. Assays were carried out in three-spot depression slides on 100-μl aliquots of L-2D cells, or *E. raikovi* cells of strain 13. Test cells were used at concentrations of $4–5 \times 10^{3}$ cells/ml and the formation of mating pairs was analyzed at intervals of 2–3 h for 1 d.

DNA extraction and polymerase chain reaction amplification

DNA was prepared from 1–2 d food-deprived cultures following a standard protocol (La Terza et al. 2009), and used as a template for polymerase chain reaction (PCR) amplifications with primers synthesized by Invitrogen (Van Allen Way, Carlsbad, CA). Amplifications were performed in an Eppendorf Mastercycler (Eppendorf AG, Hamburg, Germany), using 50-μl reaction mixtures containing 0.5 μg of DNA with 0.2 μM of each primer, 1 U of Platinum Taq DNA Polymerase High Fidelity (Invitrogen), 0.2 mM dNTP, and 2 mM $MgSO_4$. The amplification profile included an initial 5-min denaturation step at 94 °C, 35 cycles, and a final 5-min elongation step at 68 °C. Each cycle consisted of a 30-s denaturation step at 94 °C, a 40-s annealing step, and 1-min elongation step at 68 °C. The temperature of the annealing step varied 55–60 °C, depending on the primer G + C content. Amplified products were purified with Quantum Prep PCR Kleen Spin columns (Bio-Rad, Hercules, CA), and cloned into the pCR2.1-TOPO vector of the TOPO TA Cloning kit (Invitrogen) following the manufacturer's recommendations. PCR products and five different clones for each cloning reaction were sequenced at the BMR Genomics Center of the University of Padova (Italy).

Enzymatic digestion

Aliquots of 5 μl of PCR products were digested with 5 U of the enzyme *Sma*I (Fermentas, Milan, Italy) in a total 10-μl reaction volume, according to the manufacturer's instructions. Fragmented DNA was separated by electrophoresis on 1.0% agarose gel, then bands were stained with ethidium bromide (0.1 μg/ml) and photographed with a documentation gel system (Gel-Doc 2000; Bio-Rad).

RESULTS

Macronuclear pheromone gene identification and characterization

Preliminarily to the analysis of their pheromones and macronuclear pheromone genes, POR-73 cells were analyzed to establish their genotype on the basis of the Mendelian segregation of the phenotype "mating type". This segregation was assessed among a family of 50 offspring clones raised from a cross of POR-73 cells with L-2D cells, characterized (as outlined in the Introduction) by strong mating and breeding compatibility with POR-73 cells (Alimenti et al. 2011). Mating-type tests were carried out at the end of the so-called immaturity and adolescence periods (operatively identified with the cell inability and unstable ability to mate, respectively), that in *E. crassus* usually take more than 30–40 cell fissions (in temporal terms, more than 1 mo) following the cell reorganization from a sexual process (Dini and Nyberg 1994). Twenty-eight clones showed the same mating type as the L-2D cells (they mated with POR-73 cells and did not mate with L-2D cells), while the other 22 clones showed the same mating type as the POR-73 cells (they mated with L-2D cells and did not mate with POR-73 cells). This 1:1 segregation between the two parental mating types thus indicated an origin of the progeny from a cross of a heterozygous genotype with a homozygous one, with one *mat* allele of the heterozygous parent dominant and the second allele recessive or identical to the allele of the homozygous parent.

Having ascertained that one of the two strains L-2D and POR-73 contained a heterozygous genotype for the mating type, and having previously characterized the macronuclear pheromone genes Ec-α and Ec-1 from the L-2D cells (Alimenti et al. 2011), we next analyzed POR-73 cells for their macronuclear pheromone genes of which we obtained the full-length nucleotide sequences by a two-step PCR approach. Based on the notion that other *Euplotes* species, such as *E. octocarinatus* (Brünen-Nieveler et al. 1998; Möllenbeck and Heckmann 1999; Schulze-Dieckhoff et al. 1987), *E. nobilii* (Vallesi et al. 2009, 2012) and *E. raikovi* (Ricci 2012; Vallesi, A., unpubl. results), are characterized by a strict intraspecific sequence conservation of the 5′ and 3′ noncoding regions of their macronuclear pheromone genes, oligonucleotides (Table 1) specific to subtelomeric sequences of the Ec-α and Ec-1 pheromone genes were synthesized and used in the first step of DNA amplifications. In the second step, oligonucleotides specific to pheromone gene sequences were then used in combination with an oligonucleotide equivalent to the C_4A_4 repeats that form the telomeric ends of all the *Euplotes* pheromone genes. The complete pheromone gene sequences were finally reconstructed from the overlapping of all the individual sequences.

A single nucleotide sequence of 913 base pairs (bp) was obtained from the DNA amplifications run with primers specific to the Ec-α pheromone gene, and two distinct sequences of 1,179 and 1,183 bp were obtained from the

Table 1. PCR primers used in the amplifications of the macronuclear pheromone genes of POR-73 cells

PCR steps	Designations	Nucleotide sequences
1st	5′-FWα[a]	GGACTGTGAACACTAAGTTGAATAATTAG
1st	3′-RVα[a]	GAGTGTCAAAGATAATATTGATTTAATTTTTATTTGG
2nd	FWα[a]	GTATAACCAAGGAAACTATGAAGA
2nd	RVα[a]	TCATGATGGTAAGGCTTTTGCGT
1st	5′-FW1[b]	GCTCAACCTCACCTCCAA
1st	3′-RV1[b]	CTGCAGCACTACTTGACGA
2nd	FW1[b]	CGCAGAGAAAGAAGTTAGCC
2nd	RV1[b]	TTTCGTGCCCTTATGGCTTCT
2nd	TEL[c]	CCCCAAAACCCCAAAACCCC

[a]Primers for the E*c*-α pheromone gene.
[b]Primers for the E*c*-2 and E*c*-3 pheromone genes.
[c]Primer for telomeric repetitions.

DNA amplifications run with primers specific to the E*c*-1 pheromone gene.

The 913-bp sequence fully overlapped with the E*c*-α pheromone gene sequence of L-2D cells throughout the open reading frame (ORF) specifying the E*c*-α pheromone in a precursor form (i.e. prepro-E*c*-α) and the entire 3′ noncoding region, implying that the same E*c*-α pheromone gene is carried in common between POR-73 and L-2D cells (Fig. 1). Variations between the E*c*-α pheromone gene sequences of POR-73 and L-2D cells were limited only to a five-nucleotide (AGGGA) deletion and four nucleotide substitutions in the 5′ noncoding region, rich in AGGA and AGGGA repetitions of unknown functional significance and lacking putative regulatory elements for the initiation of transcription.

On the contrary, the other two sequences of 1,179 and 1,183 bp appeared to be new and specific to POR-73 cells for the synthesis of two new pheromones (annotated as E*c*-2 and E*c*-3) structurally homologous to the E*c*-1 pheromone of L-2D cells (Fig. 2). The ORF of the 1,179-bp sequence specific for the synthesis of the E*c*-2 pheromone precursor (i.e. prepro-E*c*-2) extended from $A_{799}TG$ to $TAA_{1053,}$ while the ORF of the 1,183-bp sequence of the E*c*-3 pheromone precursor (i.e., prepro-E*c*-3) extended from $A_{803}TG$ to TAA_{1057} as in the case of the E*c*-1 pheromone gene sequence. The unusually long (more than half the entire gene sequence) 5′ noncoding regions of the E*c*-2 and E*c*-3 pheromone genes varied with respect to their counterpart in the E*c*-1 pheromone gene sequence for six nucleotide substitutions, one insertion (in the E*c*-3 pheromone gene), and two short deletions (of three bp in the E*c*-2 pheromone gene, and one bp in both genes) (Fig. S1). Of greater functional importance there appeared to be full conservation of three inverted repeats and one palindromic motif that, as observed in other systems (Brázda et al. 2011), might be involved in the regulation of transcription and/or translation. With regard to the 3′ regions, alike that of the E*c*-1 pheromone sequence, they are quite short (only 98 bp) and lack a typical AAATAA polyadenilation signal that is probably replaced with a TAATAT motif.

Pheromone amino acid sequences

Based on the above macronuclear pheromone gene characterization of POR-73 cells, it appears (Fig. 3) that the amino acid sequences of the two new pheromones E*c*-2 and E*c*-3 overlap with one another and with the E*c*-1 pheromone sequence only throughout the pre and pro segments (extending for 19 and 20 amino acids, respectively) of their cytoplasmic precursor forms. The three sequences also include common proteolytic sites (Choo and Ranganathan 2008), represented by the di-peptides Ala-Glu for the removal of the presegment and Thr-Gly/Leu for the removal of the prosegment. Instead, significant variations distinguish the 45-amino acid sequences of the mature proteins, with the majority of these variations concentrated in the amino-terminal half of the molecules. However, the positions of all the 10 Cys residues appear to be strictly conserved, implying that the three pheromones all have similar three-dimensional conformations and form, like *E. raikovi* and *E. nobilii* pheromones (Luporini et al. 2005, 2014), a structurally homologous protein family.

With regard to the E*c*-α pheromone, its variations in amino acid sequence with respect to the E*c*-2 and E*c*-3 pheromones closely reflect those previously described in relation to the E*c*-α and E*c*-1 pheromone characterization in L-2D cells (Alimenti et al. 2011). They involve: (i) a new prosegment processing site represented by the Gly-Asp dipetide; (ii) the conservation of only six residues (over 19) in the presegment and only five residues (over 16) in the prosegment; (iii) a substantially hydrophilic and strong basic organization of the pro segment; (iv) an extension of the mature protein sequence from 45 to 56 residues with a parallel reduction from 10 to 8 Cys residues; and (v) the inclusion of two unusually extended and largely hydrophilic domains in middle and C-terminal regions of the mature protein.

Pheromone isolation and analysis

The effective expression of the E*c*-α, E*c*-2 and E*c*-3 pheromone genes of POR-73 cells was verified by chromatographic analysis of cell-free filtrates prepared from

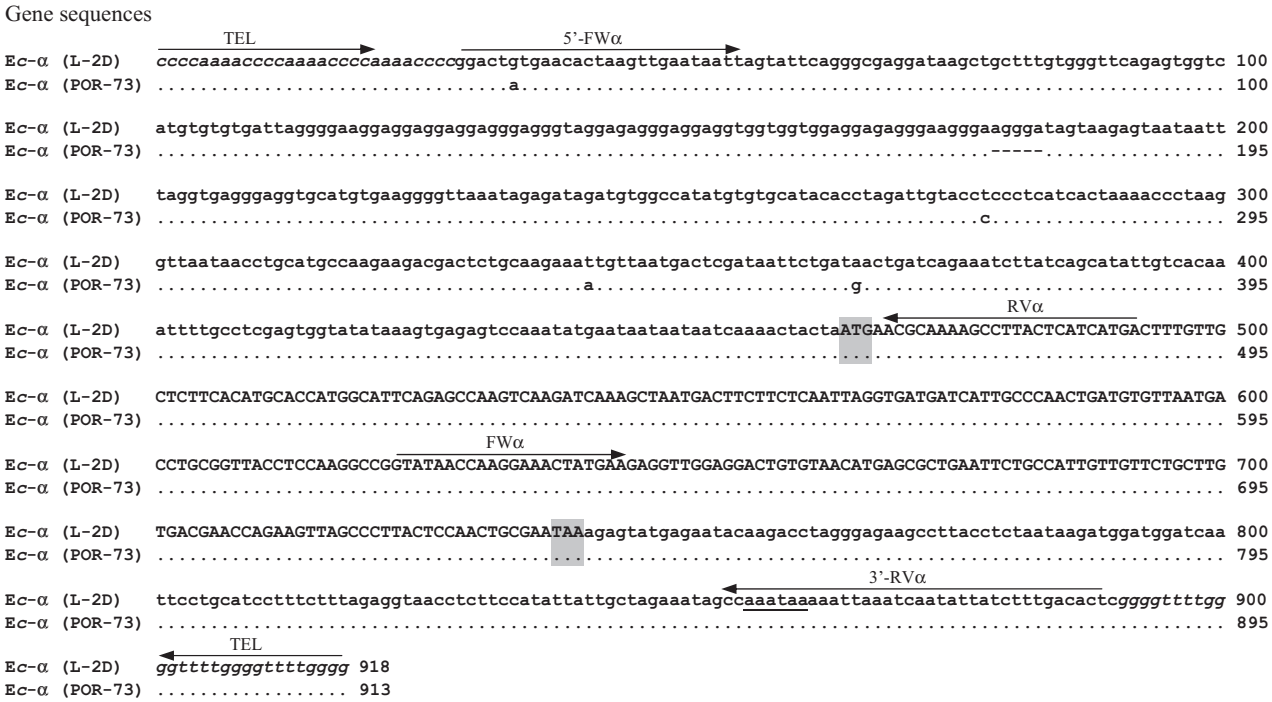

Figure 1 Alignment of the Ec-α pheromone gene sequence of POR-73 cells (GenBank accession number, KF905232) with the Ec-α pheromone gene sequence previously determined (Alimenti et al. 2011) in L-2D cells (GenBank accession number, HM072429). Gaps were inserted to maximize the alignments and dots indicate identical nucleotides. ATG start and TAA stop codons of the coding region (capital letters) are shadowed. Telomeric repeats are in italics and a putative polyadenylation signal is underlined. The numbers at the right indicate the sequence lengths. Arrows indicate positions, directions, and denominations of primers used in PCR amplifications (Table 1).

cultures of these cells grown for 1–2 wk in the presence of food at concentrations of approximately 5,000 per ml and then deprived of food for 2–3 d. Three major peaks of different dimensions were systematically well resolved (Fig. 4). Their association with the pheromones Ec-α, Ec-2 and Ec-3 was assessed by measuring the molecular mass of each protein content by MALDI-TOF/TOF mass spectrometry. The greatest of the three peaks showed full superimposition with the Ec-α pheromone peak previously detected in the chromatographic analysis of cell-free filtrate preparations of L-2D cell cultures (Alimenti et al. 2011) and revealed a protein of 6,183.8 Da, while the other two peaks of similar dimensions revealed proteins of 5,049.1 and 4,801.7 Da. These three molecular mass determinations thus closely matched the theoretical masses of 6,189.7, 5,049.7, and 4,813.5 Da predicted for the Ec-α, Ec-2, and Ec-3 pheromones by the respective coding gene sequences, implying that POR-73 cells effectively secrete all three Ec-α, Ec-2, and Ec-3 pheromones.

Purified preparations of the new Ec-2 and Ec-3 pheromones being available, we analyzed aliquots for their biological activity using only traditional mating induction assays. (Assays to detect activity as autocrine growth factors, like *E. raikovi* pheromones, Vallesi et al. 1995, were left to future research as they require more complex and long-term procedures.) L-2D cells served as intraspecific test cells and *E. raikovi* cells, previously found to be responsive to the Ec-α and Ec-1 pheromones (Alimenti et al. 2011), served as interspecific test cells. Even if used at concentrations of 10^{-5} M, pheromones Ec-2 and Ec-3 were equally unable (like pheromone Ec-1) to induce stable cell pairing at the intraspecific level; only loose, ciliary cell–cell sticking was observed. At the interspecific level, instead, 10^{-6} M concentrations of both of them were usually sufficient to induce approximately 10% of *E. raikovi* test cells to form effective mating pairs.

Pheromone gene and pheromone analysis in offspring clones of a cross between POR-73 and L-2D cells

To verify that the Ec-2 and Ec-3 pheromone genes were expressed independently of their coexistence in POR-73 cells, we determined the pheromone gene combinations of the 50 offspring clones previously obtained from crossing POR-73 cells with L-2D cells and analyzed clones carrying the Ec-2 and Ec-3 pheromone genes in new combinations with the Ec-1 pheromone gene for their pheromone secretion.

The pheromone gene combinations of the offspring clones were determined on products of DNA amplification obtained from each clone by PCRs run with primers specific to conserved sequence stretches located close to the 5′ and 3′ telomeric ends of the Ec-1, Ec-2, and Ec-3 pheromone gene sequences (Table 1). Amplified products were

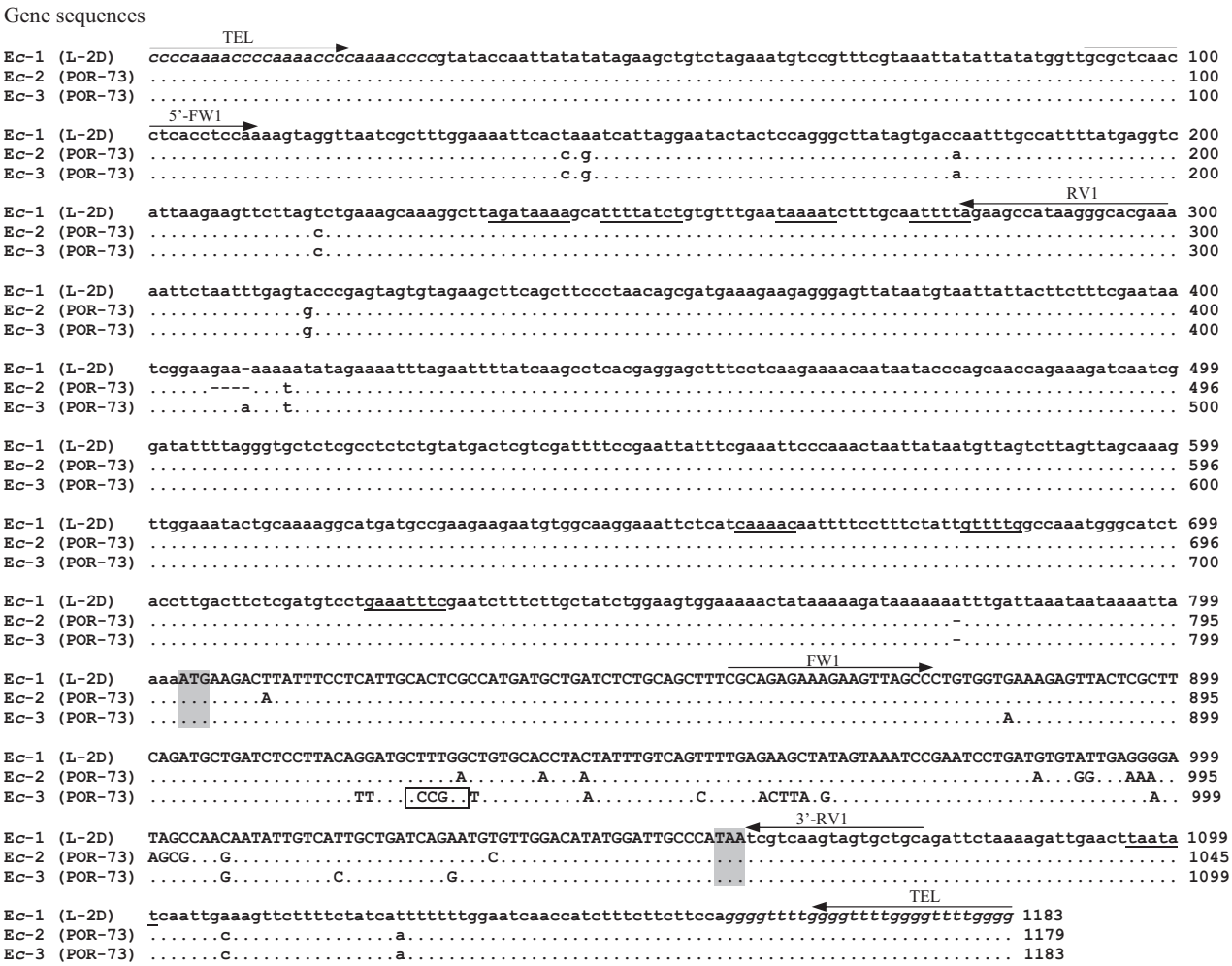

Figure 2 Alignment of the E*c*-2 and E*c*-3 pheromone gene sequences of POR-73 cells (GenBank accession numbers, KF905230 and KF905231, respectively) with the E*c*-1 pheromone gene sequence previously determined (Alimenti et al. 2011) in L-2D cells (GenBank accession number, HM072430). Gaps were inserted to maximize the alignments and dots indicate identical nucleotides. ATG start and TAA stop codons of the coding region (capital letters) are shadowed. Telomeric repeats are in italics. The numbers at the right indicate the sequence lengths. Conserved inverted repeats, a conserved palindromic motif, and putative polyadenylation signal are underlined. The *Sma*I restriction site distinctive of the E*c*-3 gene sequence is boxed. Arrows indicate positions, directions, and denominations of primers used in PCR amplifications (Table 1).

then incubated with the restriction enzyme *Sma*I which cuts only the E*c*-3 pheromone gene sequence, unique for the inclusion of a *Sma*I-specific (CCCIGGG) restriction site, into two segments. It leaves intact the sequences of the E*c*-1 and E*c*-2 pheromone genes. As shown in Fig. 5A, the 28 offspring clones previously found to express the same mating type as the parental L-2D cells all revealed an E*c*-1/E*c*-2 pheromone gene combination, while the other 22 clones expressing the same mating type as the parental POR-73 cells all revealed an E*c*-3/E*c*-1 pheromone gene combination.

Two clones representative of each new pheromone gene combination were grown into massive cultures, and preparations of cell-free filtrates of these cultures were then analyzed by chromatography and MALDI-TOF/TOF mass spectrometry to determine the pheromones present in solution. As shown in Fig. 5B, the four clones all revealed the secretion of three pheromones. One was always pheromone E*c*-α, while the other two pheromones were pheromones E*c*-1 and E*c*-2 in one case, and E*c*-1 and E*c*-3 in the other case, thus proving that the macronuclear coding genes of these three pheromones are always expressed independently of the heterozygous combinations in which they are involved.

DISCUSSION

From a characterization of macronuclear pheromone genes and analyses of pheromone secretion we obtained results that conflict with the traditional paradigm positing, on the unique basis on Mendelian analyses of mating-type inheritance, that the mating-type determination in

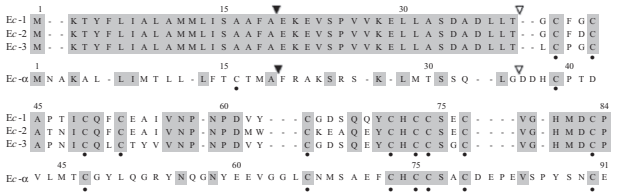

Figure 3 Alignment of the prepro-E*c*-1, prepro-E*c*-2 and prepro-E*c*-3 sequences with the prepro-E*c*-α sequence. Identical residues are shadowed and dots mark the positions of Cys residues. Filled and light arrowheads indicate the extension of the pre and proregions, respectively. Numbers indicate the progressive amino acid positions in the sequences.

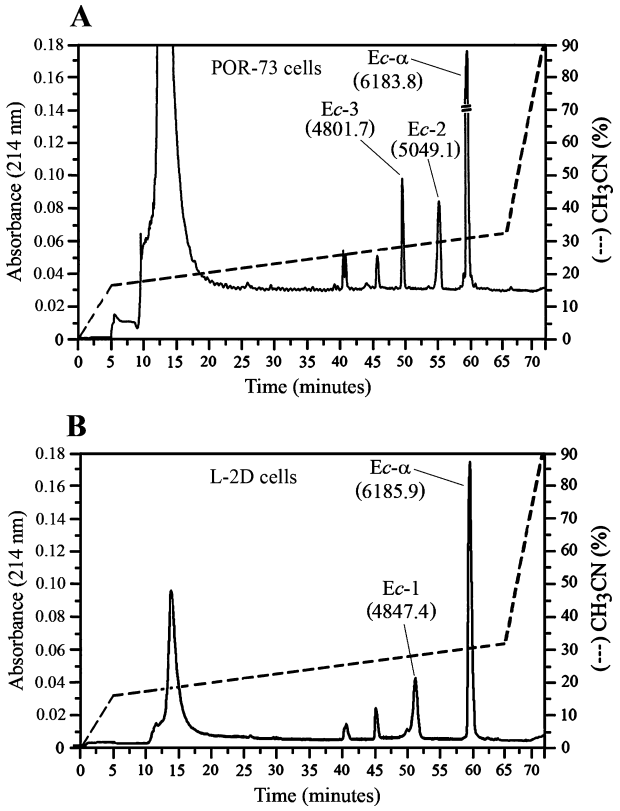

Figure 4 Elution profiles from RP-HPLC chromatography. **A.** Elution profile of pheromones E*c*-α, E*c*-2, and E*c*-3 isolated from culture filtrates of POR-73 cells. **B.** Elution profile of pheromones E*c*-α and E*c*-1 previously isolated from culture filtrates of L-2D cells (Alimenti et al. 2011) and analyzed for comparison. The elution peak of each pheromone is indicated together with the molecular mass, as determined by MALDI-TOF/TOF measurements, of the associated protein. The applied acetonitrile gradient is indicated by a dashed line.

E. crassus involves a multiple series of alleles of the same gene which are expressed with relationships of hierarchical dominance (Heckmann 1964; Nobili et al. 1978). We proved that cells with a heterozygous genotype contain three distinct macronuclear pheromone genes and secrete three distinct pheromones, and that cells with a

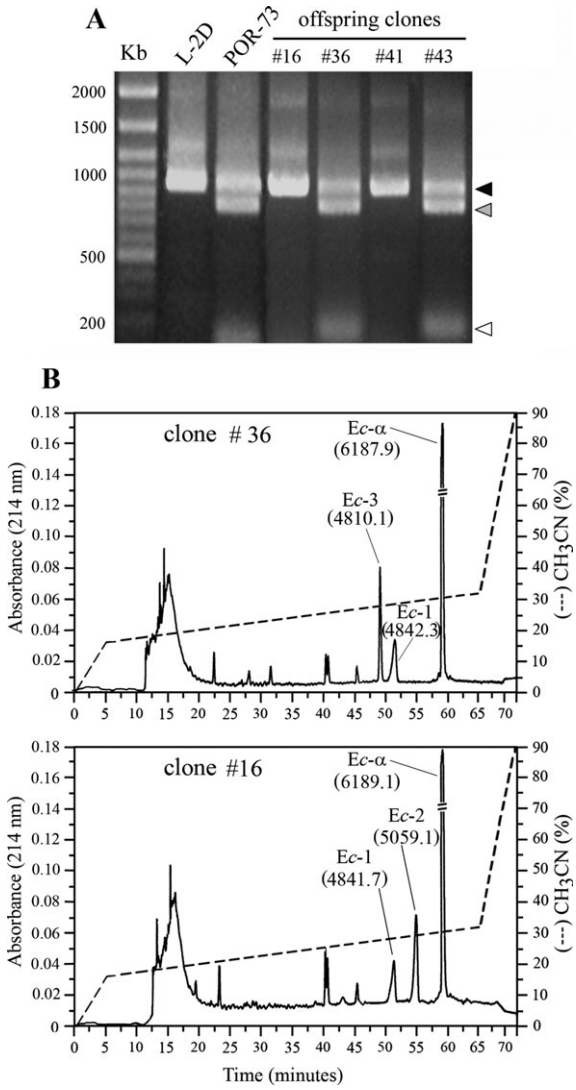

Figure 5 Pheromone gene and pheromone secretion analyses in offspring clones raised by crossing POR-73 cells with L-2D cells. **A**. Digestion, by the restriction enzyme *Sma*I, of pheromone genes of the POR-73 and L-2D parental cells, and of four clones (#16, #36, #41, and #43) representative of 50 offspring clones in all examined. The E*c*-2/E*c*-3 pheromone gene combination in POR-73 parental cells and E*c*-1/E*c*-3 pheromone gene combination in offspring clones #36 and #43 are revealed by cleavage of the PCR products of the E*c*-3 pheromone gene into two fragments of approximately 820 and 150 bp, which are indicated by gray and light arrowheads, respectively. The E*c*-1/E*c*-1 pheromone gene combination in L-2D parental cells and E*c*-1/E*c*-2 pheromone gene combination in offspring clones #16 and #41 are revealed by undigested PCR products of approximately 970 bp, which are indicated by a filled arrowhead. Molecular weight markers are shown on the left. **B**. RP-HPLC chromatographic separation of pheromones synthesized by the offspring clone #36 taken as representative of the E*c*-3/E*c*-1 pheromone gene combination, and the offspring clone #16 taken as representative of the E*c*-2/E*c*-1 pheromone gene combination. The elution peak of each pheromone is indicated together with the molecular mass, as determined by mass spectrometry, of the associated protein. The applied acetonitrile gradient is indicated by a dashed line.

homozygous genotype contain two distinct macronuclear pheromone genes and secrete two distinct pheromones. We interpret these results as evidence that a phenomenon of gene duplication characterizes the micronuclear mating-type locus of *E. crassus*, and most likely of the entire *E. crassus-minuta-vannus* sibling species complex.

One gene would be orthologous to (i.e. diverged by evolution from) the gene that Mendelian and molecular analyses have concordantly indicated to be unique in *E. raikovi* (Luporini et al. 1986), *E. octocarinatus* (Brünen-Nieveler et al. 1998; Heckmann and Kuhlmann 1986; Möllenbeck and Heckmann 1999) and *E. nobilii* (Di Giuseppe et al. 2011; Vallesi et al. 2012). This gene analogously shows multiple alleles which encode the cell-type specific pheromones E*c*-1, E*c*-2, E*c*-3 and so forth, and are expressed *not with*, *but without* relationships of hierarchical dominance in the transcriptionally active macronucleus of the cell. As shown in Fig. 6, these *E. crassus* pheromones represent a structurally homologous protein family clearly correlated, by various degrees of conservation of the Cys residues and of the prepro sequences of the cytoplasmic pheromone precursors, with the pheromone families that have so far been characterized in *E. raikovi*, *E. octocarinatus*, and *E. nobilii* (Alimenti et al. 2009; Luporini et al. 2005; Möllenbeck and Heckmann 1999).

Of the second paralogous gene, generated by duplication, we currently know only the allele encoding the E*c*-α pheromone; it will be necessary to identify other alleles to be able to reliably determine structural correlations with the other *Euplotes* pheromones. Considered individually, pheromone E*c*-α appears to branch off early from the other *E. crassus* cell-type specific pheromones (Fig. 6); it differs from them in numerous sequence variations that distinguish not only the secreted region but also the pre and proregions of the cytoplasmic precursor.

With regard to the E*c*-α pheromone activity, the fact that this same pheromone is equally synthesized by cells of different and mutually compatible mating types would *a priori* exclude that it evolved for duplicating the activity of the cell-type specific pheromones. More likely, its activity is directed to supplement, or assist that of the cell-type specific pheromones. Our current hypothesis is that pheromone E*c*-α behaves like an adaptor, or a scaffold protein that mediates the binding of the cell-type specific pheromones to their cell-membrane receptors which, as is the case in *E. raikovi* (Miceli et al. 1992; Ortenzi et al. 2000; Vallesi et al. 1995, 2005), we presume to be represented by membrane-bound pheromone isoforms characterized by an extracellular binding domain structurally identical to the partner soluble forms.

This hypothesis is supported by ongoing research showing that the E*c*-α pheromone is unique for a strong propensity to associate in solution with the cell-type specific pheromones (Fig. S2), and to oligomerize when is purified by chromatographic separation. It starts forming irreversible precipitates even at concentrations as low as 5–6 μg/μl, whereas the cell-type specific pheromones remain stable in solution at concentrations even 10-fold higher (unreported data). Consistently with

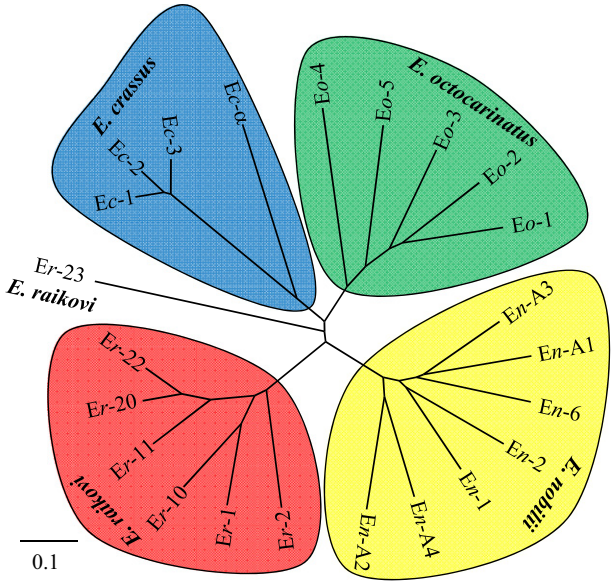

Figure 6 Structural relationships among *Euplotes* pheromones. Unrooted tree, based on the neighbor-joining method using MEGA v. 4.0 (Tamura et al. 2007), linking the amino acid sequences of *Euplotes crassus* pheromones, considered in their precursor forms (prepropheromones), with those of *E. nobilii*, *E. octocarinatus,* and *E. raikovi*. Because of its longer polypeptide chain, higher number of Cys residues, and NMR solution structure with two additional irregular α-helices, pheromone E*r*-23 is regarded as a deviant member of the *E. raikovi* pheromone family (Di Giuseppe et al. 2002; Zahn et al. 2001). The abbreviation "E*o*" stands for *E. octocarinatus* pheromones. The scale bar corresponds to a distance of 10 variations per 100 amino acid residues. The GenBank accession numbers of the pheromone sequences included in this analysis are the following ones. *E. crassus* pheromones: see text. *E. octocarinatus* pheromones: E*o*-1, O15823; E*o*-2, O15825; E*o*-3, P28717; E*o*-4, P28716; E*o*-5, CAA76771. *Euplotes nobilii* pheromones: E*n*-1, ACQ66088; E*n*-2, ACQ66089; E*n*-6, ABK15649; E*n*-A1 ACQ66090; E*n*-A2, ACQ66091; E*n*-A3, ACQ66092; E*n*-A4, ACQ66093. *Euplotes raikovi* pheromones: E*r*-1, P10774; E*r*-10, P12350; E*r*-2, P26886; E*r*-11, P26887; E*r*-20, P26888; E*r*-22, P58548; E*r*-23, P58547. The last four accession numbers refer to the amino acid sequences of the secreted proteins. The pre and proregion sequences of pheromone E*r*-23 have been taken from Di Giuseppe et al. (2002), and those of pheromones E*r*-11, E*r*-20, and E*r*-22 from Ricci (2012) and our unpublished data.

these observations, the E*c*-α pheromone structure includes not only an extended domain (spanning from Cys_{57} to Cys_{66} counting from the Met_1 residue of the prepro-E*c*-α sequence) rich in Gly residues and lacking Cys residues but also two amyloidogenic domains (one spanning from Val_{43} to Tyr_{49} and one from Phe_{72} to Ser_{77}) as identified by the computer programs "FoldAmyloid" (Conchillo-Sole et al. 2007) and "AGGRESCAN" (Garbuzynskiy et al. 2010). The Gly-rich and Cys-free domain likely forms a random-coil region suitable for increasing the structural plasticity and ability of the E*c*-α pheromone to interact with other molecules. The two amyloidogenic domains are potential sites for phenomena of protein-protein interactions and aggregation.

Capability of the E*c*-α pheromone to interact with the other cell-type specific pheromones may also explain why *E. crassus* cells with heterozygous genotypes respond to mating tests in the same way as cells with homozygous genotypes, thus mimicking relationships of hierarchical dominance among the *mat* alleles. Assumed that the activity of the cell-type specific pheromones depends on their interactions with the E*c*-α pheromone and these interactions involve different degrees of protein-protein binding affinity, one of the two cell-type specific pheromones secreted by heterozygous cells would regularly appear as more active than the other one causing a situation functionally analogous to the production of a single pheromone by homozygous cells.

As previously observed in the study of the E*c*-α and E*c*-1 pheromones purified from L-2D cells (Alimenti et al. 2011), also the new E*c*-2 and E*c*-3 pheromone preparations from POR-73 cells appeared unable to induce *in vitro* formation of homotypic (selfing) mating pairs in *E. crassus* itself. To show their mating induction activity, it was necessary to carry out assays on *E. raikovi* cells. While this interspecific activity of *E. crassus* pheromones can be explained by the close conservation of common structural and functional domains that is manifested within and between the pheromone families of different *Euplotes* species (Luporini et al. 2005, 2014), the why *E. crassus* cells are so reluctant to form homotypic mating pairs in response to *in vitro* pheromone interactions is more difficult to understand. Differently from the *E. patella*-like species, which all indistinctly form homo- and heterotypic mating pairs regardless of the homo- or heterozygous conditions of the mating types that are involved in a mating mixture (Akada 1986; Kimball 1942; Luporini and Miceli 1986), *E. crassus* mating mixtures usually result in a largely prevalent, or exclusive formation of heterotypic pairs (Dini and Nyberg 1994; Heckmann 1964; Nobili et al. 1978). Considering that the E*c*-α pheromone is specific to *E. crassus* and unknown to all the *E. patella*-like species, it seems reasonable to correlate the evolution of this pheromone in *E. crassus* with a mechanism directed to halt the formation of the homotypic pairs and favor that of the heterotypic ones, which are the only suitable to promote an effective gene pool reshuffling of the species. Knowledge of how this mechanism works calls for specific analyses of the E*c*-α pheromone interactions with the cell-type specific pheromones and their membrane receptors.

ACKNOWLEDGMENTS

This study was financially supported by a contribution provided by the Ministero dell'Istruzione, Università e Ricerca in relation to PRIN research projects. The expert advice of Sheila Beatty in editing the English usage of the text is gratefully acknowledged.

LITERATURE CITED

Alimenti, C., Vallesi, A., Federici, S., Di Giuseppe, G., Dini, F., Carratore, V. & Luporini, P. 2011. Isolation and structural characterization of two water-borne pheromones from *Euplotes crassus*, a ciliate commonly known to carry membrane-bound pheromones. *J. Eukaryot. Microbiol.*, 58:234–241.

Alimenti, C., Vallesi, A., Pedrini, B., Wüthrich, K. & Luporini, P. 2009. Molecular cold-adaptation: comparative analysis of two homologous families of psychrophilic and mesophilic signal proteins of the protozoan ciliate *Euplotes*. *IUBMB Life*, 61:838–845.

Akada, R. 1986. Spontaneous selfing (intraclonal conjugation) in the ciliate *Euplotes patella* syngen 2 takes place without a change in mating type expression. *J. Protozool.*, 33:235–237.

Brázda, V., Laister, R. C., Jagelská, E. B. & Arrowsmith, C. 2011. Cruciform structures are a common DNA feature important for regulating biological processes. *BMC Mol. Biol.*, 12:33.

Brünen-Nieveler, C., Weiligmann, J. C., Hansen, B., Kuhlmann, H. W., Möllenbeck, M. & Heckmann, K. 1998. The pheromones and pheromone genes of new stocks of *E. octocarinatus* species complex. *Eur. J. Protistol.*, 34:124–132.

Choo, K. H. & Ranganathan, S. 2008. Flanking signal and mature peptide residues influence signal peptide cleavage. *BMC Bioinformatics*, 9(Suppl. 12):S15. doi:10.1186/1471-2105-9-S12-S15.

Conchillo-Sole, O., de Groot, N. S., Aviles, F. X., Vendrell, J., Duara, X. & Ventura, S. 2007. AGGRESCAN: a server for the prediction and evaluation of "hot spot" of aggregation in polipeptides. *BMC Bioinformatics*, 8:65–81.

Di Giuseppe, G., Erra, F., Dini, F., Alimenti, C., Vallesi, A., Pedrini, B., Wüthrich, K. & Luporini, P. 2011. Antarctic and Arctic populations of the ciliate *Euplotes nobilii* show common pheromone-mediated cell-cell signaling and cross-mating. *Proc. Natl Acad. Sci. USA*, 108:3181–3186.

Di Giuseppe, G., Miceli, C., Zahn, R., Damberger, F., Wüthrich, K. & Luporini, P. 2002. A structurally deviant member of the *Euplotes raikovi* pheromone family, E*r*-23. *J. Eukaryot. Microbiol.*, 49:86–92.

Dini, F. & Luporini, P. 1985. Mating-type polymorphic variations in *Euplotes minuta* (Ciliophora: Hypotrichida). *J. Protozool.*, 32:111–117.

Dini, F. & Nyberg, D. 1993. Sex in ciliates. *In*: Jones, J.G. (ed.), Advances in Microbial Ecology, vol. 13. Plenum Press, New York, NY. p. 85–153.

Dini, F. & Nyberg, D. 1994. Adolescence and the reversibility of maturity in *Euplotes crassus*. *J. Eukaryot. Microbiol.*, 41:373–380.

Garbuzynskiy, S. O., Lobanov, M. Y. & Galzitskaya, O. V. 2010. FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence. *Bioinformatics*, 26:326–332.

Heckmann, K. 1964. Experimentelle Untersuchungen an *Euplotes crassus* I. Paarungssystem, Konjugation und Determination der Paarungstypen. *Z. VererbLehre*, 95:114–124.

Heckmann, K. 1967. Age-dependant intraclonal conjugation in *Euplotes crassus*. *J. Exp. Zool.*, 165:269–278.

Heckmann, K. & Kuhlmann, H.-W. 1986. Mating types and mating inducing substances in *Euplotes octocarinatus*. *J. Exp. Zool.*, 237:87–96.

Kimball, R. F. 1942. The nature and inheritance of mating types in *Euplotes patella*. *Genetics*, 27:269–285.

La Terza, A., Dobri, N., Alimenti, C., Vallesi, A. & Luporini, P. 2009. The water-borne protein signals (pheromones) of the Antarctic ciliated protozoan *Euplotes nobilii*: structure of the gene coding for the E*n*-6 pheromone. *Can. J. Microbiol.*, 55:57–62.

Luporini, P., Alimenti, C., Ortenzi, C. & Vallesi, A. 2005. Ciliate mating types and their specific protein pheromones. *Acta Protozool.*, 44:89–101.

Luporini, P., Alimenti, C. & Vallesi, A. 2014. Ciliates mating types and pheromones. *In*: Hausmann, K. & Radek, R. (ed.), Cilia/Flagella – Ciliates/Flagellates. Schweizerbart Science Publishers, Stuttgart. p. 95–118.

Luporini, P. & Miceli, C. 1986. Mating pheromones. *In*: Gall, J. G. (ed.), The Molecular Biology of Ciliated Protozoa. Academic Press, New York, NY. p. 263–299.
Luporini, P., Raffioni, S., Concetti, A. & Miceli, C. 1986. The ciliate *Euplotes raikovi* heterozygous at the *mat* genetic locus co-releases two individual species of mating pheromones: genetic and biochemical evidence. *Proc. Natl Acad. Sci. USA*, 83:2889–2893.
Miceli, C., Concetti, A. & Luporini, P. 1983. Isolation of the mating-inducing factor of the ciliate *Euplotes*. *Exp. Cell Res.*, 149:593–598.
Miceli, C., La Terza, A., Bradshaw, R. A. & Luporini, P. 1992. Identification and structural characterization of a cDNA clone encoding a membrane-bound form of the polypeptide pheromone E*r*-1 in the ciliate protozoan *Euplotes raikovi*. *Proc. Natl Acad. Sci. USA*, 89:1988–1992.
Miceli, C., Luporini, P. & Bracchi, P. 1981. Morphological description, breeding system, and nuclear changes during conjugation of *Euplotes raikovi* Agamaliev from Mediterranean sea. *Acta Protozool.*, 20:215–224.
Möllenbeck, M. & Heckmann, K. 1999. Characterization of two genes encoding a fifth so far unknown pheromone of *Euplotes octocarinatus*. *Eur. J. Protistol.*, 35:225–230.
Nobili, R. 1966. Mating types and mating type inheritance in *Euplotes minuta* Yocom (Ciliata Hypotrichida). *J. Protozool.*, 13:38–41.
Nobili, R., Luporini, P. & Dini, F. 1978. Breeding systems, species relationships and evolutionary trends in some marine species of Euplotidae (Hypotrichida, Ciliata). *In*: Battaglia, B. & Beardmore, J. (ed.), Nato Conference Series. Marine Organisms. Genetics, Ecology and Evolution, vol. 1. Plenum Press, New York, NY & London. p. 591–616.
Ortenzi, C., Alimenti, C., Vallesi, A., Di Pretoro, B., La Terza, A. & Luporini, P. 2000. The autocrine mitogenic loop of the ciliate *Euplotes raikovi*: the pheromone membrane-bound forms are the cell binding sites and potential signalling receptors of soluble pheromones. *Mol. Biol. Cell*, 11:1445–1455.
Phadke, S. S. & Zufall, R. A. 2009. Rapid diversification of mating systems in ciliates. *Biol. J. Linnean Soc.*, 98:187–197.
Raffioni, S., Miceli, C., Concetti, A., Barra, D. & Luporini, P. 1987. Purification and characterization of new mating pheromones of the ciliate *Euplotes raikovi*. *Exp. Cell Res.*, 172:417–424.
Ricci, F. (2012). Structure and expression of the *Euplotes raikovi* macronuclear pheromone genes. Ph.D. Thesis. University of Camerino, XXIV Cycle.
Schulze-Dieckhoff, H., Freiburg, M. & Heckmann, K. 1987. The isolation of gamones 3 and 4 of *Euplotes octocarinatus*. *Eur. J. Biochem.*, 168:89–94.
Tamura, K., Dudley, J., Nei, M. & Kumar, S. 2007. MEGA4: Molecular Evolutionary Genetic Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.*, 24:1596–1599.
Valbonesi, A., Ortenzi, C. & Luporini, P. 1992. The species problem in a ciliate with a high-multiple mating type system, *Euplotes crassus*. *J. Protozool.*, 39:45–54.
Vallesi, A., Alimenti, C., La Terza, A., Di Giuseppe, G., Dini, F. & Luporini, P. 2009. Characterization of the pheromone gene family of an Antarctic and Arctic protozoan ciliate, *Euplotes nobilii*. *Mar. Genomics*, 2:27–32.
Vallesi, A., Alimenti, C., Pedrini, B., Di Giuseppe, G., Dini, F., Wüthrich, K. & Luporini, P. 2012. Coding genes and molecular structures of the diffusible signaling proteins (pheromones) of the polar ciliate, *Euplotes nobilii*. *Mar. Genomics*, 8:9–13.
Vallesi, A., Ballarini, P., Di Pretoro, B., Alimenti, C., Miceli, C. & Luporini, P. 2005. The autocrine, mitogenic phromone-receptor loop of the ciliate *Euplotes raikovi*: pheromone-induced receptor internalization. *Euk. Cell*, 4:1221–1227.
Vallesi, A., Giuli, G., Bradshaw, R. A. & Luporini, P. 1995. Autocrine mitogenic activity of pheromones produced by the protozoan ciliate *Euplotes raikovi*. *Nature*, 376:522–524.
Walne, P. R. 1966. Experiments in the large-scale culture of larvae of *Ostrea edulis* L. *Fish Invest. London, Ser. 2*, 25:1–10.
Weischer, A., Freiburg, M. & Heckmann, K. 1985. Isolation and partial characterization of gamone 1 of *Euplotes octocarinatus*. *FEBS Lett.*, 191:176–180.
Zahn, R., Damberger, F., Ortenzi, C., Luporini, P. & Wüthrich, K. 2001. NMR structure of the *Euplotes raikovi* pheromone E*r*-23 and identification of its five disulfide bonds. *J. Mol. Biol.*, 313:923–931.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Sequence specificities of *Euplotes crassus* pheromone genes. **A**. Chromatograms of the E*c*-2 and E*c*-3 pheromone genes of POR-73 cells showing the sequence segment, spanning between nucleotides 395 and 420/424, that differs between the two genes for a four-nucleotide (GAAA) deletion. **B**. Alignment of the E*c*-2 and E*c*-3 pheromone gene sequence segments (**A**) with the matching segment of the E*c*-1 pheromone gene previously characterized from L-2D cells.

Figure S2. Binding reactions of *Euplotes crassus* pheromones analyzed by means of Surface Plasmon Resonance on a BIACORE X apparatus. **A**. Representative sensorgrams of concentration-dependent binding of pheromone E*c*-α to itself, and to the cell-type specific pheromones E*c*-1, E*c*-2, and E*c*-3. A purified preparation of pheromone E*c*-α was cross-linked to a sensor chip (CM5) and used as ligand of purified preparations of the E*c*-α, E*c*-1, E*c*-2, and E*c*-3 pheromones, that were injected as analyte into the flow chamber of the apparatus at increasing concentrations. Each concentration was exposed to the sensor chip surface for 300 s (association phase) and removed by 300 s flow of running buffer (dissociation phase). The curves obtained from different injections were superimposed using the BIAevaluation 4.1 software, and their fit was verified in accord with a bivalent analyte model which describes the binding of a bivalent analyte molecule to one or two immobilized ligand molecules with two different sets of rate constants (Wells 1992). The application of this model is supported by the knowledge that the crystal structure of the *E. raikovi* E*r*-1 pheromone involves a molecule cooperative packing resulting from the formation of two distinct types of dimers (Weiss et al. 1995). The experimental and calculated curves are represented by colored and black lines, respectively, and plotted in resonance units (RU) versus time. The plotted data are differences in the RU signals between the flow cell carrying the immobilized ligand and the reference (ligand-free) flow cell; they reflect the formation of specific ligand-analyte complexes. **B.** Summary of the kinetic constants calculated for each ligand-analyte interaction. K_a, association constant; K_d, dissociation constant; K_D, equilibrium dissociation constant.