



Article

Primary Structure and Coding Genes of Two Pheromones from the Antarctic Psychrophilic Ciliate, *Euplotes focardii*

Claudio Alimenti, Annalisa Candelori, Yaohan Jiang, Pierangelo Luporini and Adriana Vallesi *

Laboratory of Eukaryotic Microbiology and Animal Biology, School of Biosciences and Veterinary Medicine, University of Camerino, 62032 Camerino, MC, Italy; claudio.alimenti@unicam.it (C.A.); annalisa.candelori@unicam.it (A.C.); yaohan.jiang@unicam.it (Y.J.); piero.luporini@unicam.it (P.L.)

* Correspondence: adriana.vallesi@unicam.it

Abstract: In ciliates, diffusible cell type-specific pheromones regulate cell growth and mating phenomena acting competitively in both autocrine and heterologous fashion. In *Euplotes* species, these signaling molecules are represented by species-specific families of structurally homologous small, disulfide-rich proteins, each specified by one of a series of multiple alleles that are inherited without relationships of dominance at the *mat*-genetic locus of the germinal micronuclear genome, and expressed as individual gene-sized molecules in the somatic macronuclear genome. Here we report the 85-amino acid sequences and the full-length macronuclear nucleotide coding sequences of two pheromones, designated Ef-1 and Ef-2, isolated from the supernatant of a wild-type strain of a psychrophilic species of *Euplotes*, *E. focardii*, endemic to Antarctic coastal waters. An overall comparison of the determined *E. focardii* pheromone and pheromone-gene structures with their homologs from congeneric species provides an initial picture of how an evolutionary increase in the complexity of these structures accompanies *Euplotes* speciation.



Citation: Alimenti, C.; Candelori, A.; Jiang, Y.; Luporini, P.; Vallesi, A. Primary Structure and Coding Genes of Two Pheromones from the Antarctic Psychrophilic Ciliate, *Euplotes focardii*. *Microorganisms* **2022**, *10*, 1089. <https://doi.org/10.3390/microorganisms10061089>

Academic Editor: Juan Carlos Gutiérrez

Received: 19 April 2022

Accepted: 23 May 2022

Published: 25 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: protein pheromones; macronuclear gene structures; polar microbiology; water-borne chemical signals

1. Introduction

Like many other organisms, from bacteria [1–4] to animals [5,6], ciliates synthesize cell-type specific signaling pheromones in functional association with their binary or multiple ‘mating type systems’, which control the cell switching from the vegetative (growth) stage to the sexual (mating) stage of the life cycle [7,8], and thus play a central role in speciation by driving the interpopulation gene flow [9,10]. Pheromonal activity has been identified in the cell-culture supernatant of various ciliates, either on the basis of mating induction assays in species of *Blepharisma*, *Dileptus*, *Euplotes* and *Oxytricha* ([11–14] as reviews), or observations of context-specific changes in cell morphology and behavior in species of *Ephelota* [15] and *Tokophrya* [16]. However, a successful pheromone isolation and chemical characterization has been carried out in only two cell types (I and II) of *Blepharisma japonicum* [11], and multiple cell types of five species of *Euplotes*, first of *E. raikovi* and then, *E. octocarinatus*, *E. nobilii*, *E. petzi* and *E. crassus* [12–14].

The two *B. japonicum* pheromones (originally designated ‘gamone 1’ or blepharhormone’, and ‘gamone 2’ or ‘blepharismone’) are uniquely to be represented by chemically unrelated molecules, one being a glycoprotein of 272 amino acids and seven sugars, and the second a tryptophan derivative that has been reported to be synthesized also by other species of *Blepharisma* in addition to *B. japonicum* [17,18]. On the other hand, *Euplotes* pheromones have all been revealed to be acidic, cysteine-rich proteins varying in sequence length from 37 amino acids in *E. petzi* [19] to 105 in *E. octocarinatus* [20]. In *E. octocarinatus* and *E. crassus*, from which pheromones are purified in minimal amounts (0.4–0.5 µg of protein/liter of culture supernatant in *E. octocarinatus*), the pheromone structural knowledge has remained

limited to the primary amino acid sequences [20–23]. The three-dimensional conformations have instead been determined, by NMR spectroscopy and/or X-ray crystallography, for seven *E. raikovi* pheromones (Er-1, Er-2, Er-10, Er-11, Er-13, Er-22 and Er-23), four (En-1, En-2, En-6 and En-1) of *E. nobilii*, and one (Ep-1) of *E. petzi*, that are overall purified in larger amounts (up to 200–300 µg of protein/liter of culture supernatant in *E. raikovi*) [19,24–30]. Sharing a common helical fold tightly stabilized by densely spaced and tightly conserved disulfide bonds while differing from one another in local structural specificities, these conformations provided definitive evidence of the close relationships of structural homology that make *Euplotes* pheromones representative of species-specific protein families, and capable of competing with one another to bind target cells in both autocrine (self) fashion to promote cell growth and heterologous (non-self) fashion to induce cell mating [30,31]. This structural homology is the result of a genetic determination operated by multiple series of alleles that are inherited in Mendelian fashion at the so-called *mat* (mating-type) genetic locus of the chromosomal micronuclear (germinal) genome [32–34]. These allelic genes are then expressed with no mutual relationship of dominance in the sub-chromosomal macronuclear (somatic) genome [23], where they take the form of linear gene-sized DNA molecules including a single central coding region flanked by 5'-leader and 3'-trailer regions capped with telomeric C₄A₄ repeats [35,36]. Through a combined biochemical and genetic approach, we have determined and report the amino acid sequences and the structures of the macronuclear coding genes of two pheromones that are secreted in tiny amounts by a rather uncommon *Euplotes* species, *E. focardii*, which is characterized by a strictly psychrophilic behavior being endemic to the Antarctic coastal waters [37], and a position in the *Euplotes* phylogenetic tree distant from the basal clades including *E. petzi*, *E. raikovi* and *E. nobilii*, and close to the late clades including *E. octocarinatus* and *E. crassus* [38,39]. An overall comparison of the pheromone and pheromone–gene structures between *E. focardii* and the other *Euplotes* species suggests an initial picture of how the evolution of these structures accompanies *Euplotes* speciation.

2. Materials and Methods

2.1. Pheromone Source

The pheromone source was the *E. focardii* strain TN₁, which was cultivated on the green alga *Dunaliella tertiolecta*, at 2–4 °C, under a rhythm of 16 h of darkness and 8 h of very weak light. This strain (deposited at the Culture Collection of Algae and Protozoa, CCAP, Scottish Marine Institute, Argil PA37 1QA, Scotland, under the accession number 1624/17) was established in the laboratory starting from one of several specimens isolated from a sample of sandy sediments collected in January 1988 from a small cove close to the Italian research station at Terra Nova Bay (Ross Sea, Antarctica) [37]. Supernatant preparations for the pheromone isolation were obtained from cultures that were expanded by continuous additions of moderate concentrations of food, to reach a final volume of 15–20 L and a cell density of approximately 2×10^3 cells/mL, and were then left to starve for 1 week. Supernatant was prepared by removing cells and debris in suspension through filtration.

2.2. Pheromone Purification and Analysis

The pheromone purification was carried out according to a standard three-step chromatographic procedure, previously set up for the pheromone purification from other marine species of *Euplotes* [22,40]. Proteins were adsorbed from supernatant preparations onto reverse-phase Sep-Pak C₁₈ cartridges (Waters, Milford, CA, USA) and separated, first, by gel filtration on a Superdex peptide HR10/30 column (GE Healthcare, Little Chalfont, UK), equilibrated in Tris-HCl 20 mM and NaCl 0.4 M and eluted at 0.4 mL/min; then, by reverse-phase-high-performance chromatography [RP-HPLC] on a 4.6-mm × 250-mm C₁₈ column (Supelco, Bellefonte, PA, USA). Elution was carried out with a discontinuous acetonitrile gradient in the presence of 0.06% (v/v) trifluoroacetic acid. The molecular mass measurements were performed by matrix-assisted laser desorption/ionization [MALDI] time-of-flight [TOF] mass spectrometry, using a Voyager DE-PRO spectrometer (Percip-

tive Biosystems, Framingham, MA, USA), equipped with a nitrogen laser operated at 337 nm. The protein automated Edman degradation was carried out in a Procise 492 protein sequencer (Applied Biosystem, Perkin Elmer, Foster City, CA, USA).

2.3. Assaying Pheromone Activity

Ciliate pheromones are conventionally detected in solution via mating induction assays, which entail the capacity of genetically identical cells to form homotypic (or selfing) mating pairs in response to the suspension with heterologous pheromone solutions [11,13]. Among *Euplotes* species, this capacity is, however, native to only some species and alien to others including *E. focardii* which, in addition, is rather impracticable for mating studies due to the enormous lengthening that its life in freezing waters imposes on the dynamics of the mating process [41]. The assays for the identification of the chromatographic peaks containing pheromone activity were thus carried out, in a heterospecific manner, on *E. raikovi* cells, which are constitutively committed to practice homotypic pairing and were previously used to identify pheromones from another species, *E. crassus*, which, like *E. focardii*, is refractory to mate in homotypic fashion [22]. Samples of 2×10^3 cells from *E. raikovi* type-I cultures, previously deprived of food for a couple of days, were resuspended with 500- μ L of fresh seawater and incubated with three increasing volumes (10, 20 and 40- μ L) from the chromatographic fractions of interest. The mixtures were then inspected over an interval of 1–2 days to detect the cell mating response. This response was also taken as positive in the case of the formation of unstable cell–cell unions in mating pairs.

2.4. Gene Amplification, Cloning and Sequencing

Gene amplification was carried out by applying a Rapid Amplification of Telomeric Ends [RATE]-Polymerase Chain Reaction [PCR] protocol to macronuclear DNA preparations obtained from cell pellets following a standard procedure [42]. PCR primer designations and sequences are listed in Table 1. The reactions were run in an Eppendorf Ep-gradient Mastercycler (Eppendorf AG, Hamburg, Germany), using 0.5 μ g DNA aliquots in 50 μ L-reaction mixtures containing 0.5 μ M of each primer, 0.25 mM dNTP and 0.02 U/ μ L of Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific Inc., Waltham, MA, USA). As a rule, 35 amplification cycles were carried out. Each cycle consisted of a 98 °C denaturation step for 30 s, a 30 s annealing step, and a 72 °C elongation step for 30–90 s. The temperature of the annealing step varied from 55 to 65 °C, depending on the primer G + C content. A final incubation step, at 72 °C for 5 min, was added to the last cycle. Amplified products were cloned in the pJET1.2/blunt cloning vector (Thermo Fisher Scientific Inc., Waltham, MA, USA), following the procedure suggested by the manufacturer.

Table 1. PCR primer designation and sequences.

Name	Nucleotide Sequence (5'–3')
dFW1 ^a	GATTGCCAYGGWGATACNGA
dFW2 ^a	GATACTGAATAYTTMATYGAYGART
TEL	CCCCAAAACCCCAAAACCCC
RV1	TCAATACAAACTTGACAGCAGTTACA
RV2	TAATCACCAAAAACCGTCAGATCCCAT
5'-FW1	TGGCTAAAGCATAGAATTCTAACAT
5'-FW2	TGTTTATAAATGAGGAAGTGCTTAAG
3'-RV	ACACTATTGAACCAGAATATTCCTCT

^a Y, W, M and R, alternatives between C and T, A and T, A and C, and A and G, respectively. N, any nucleotide.

3. Results

3.1. Pheromone Isolation and Structural Characterization

The protein material extracted from the supernatant preparations of the *E. focardii* cultures and initially separated on gel filtration revealed pheromonal activity, as detected by heterospecific mating induction assays, in association with only four fractions in the

range from 9 to 12 kDa (Figure 1a). On reverse-phase chromatography, the protein pool of these fractions was next resolved into two major peaks (Figure 1b), each containing a single protein species as assessed by mass spectrometry analysis (Figure 1c,d). One species was measured to be 9245.5 Da, and was identified with one pheromone designated *Ef-1*; the second species was measured to be 9201.7 Da, and identified with a second pheromone designated *Ef-2*. The two proteins were quantified in the cell-culture supernatant at a concentration of 10–20 µg/L, and their co-existence provided direct evidence that their source TN₁ cells carried a heterozygous allelic combination of the *mat* locus of the micronuclear genome, as in general is the case in wild-type *Euplotes* strains [9,34].

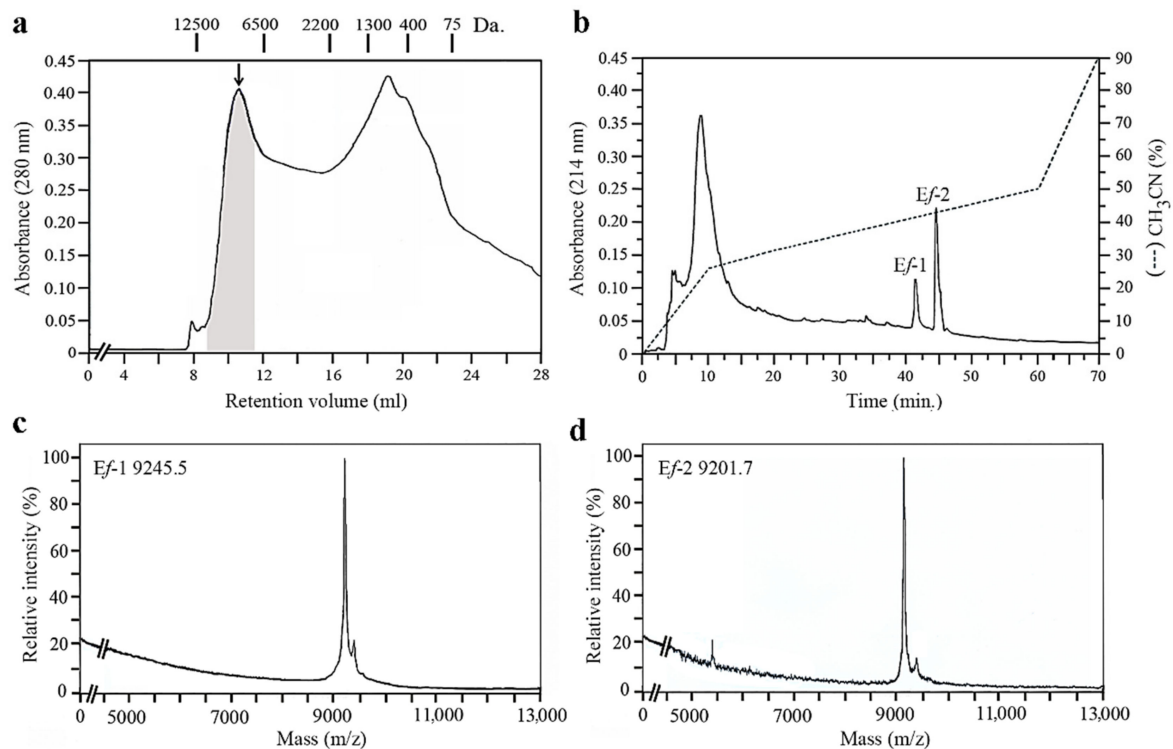


Figure 1. Purification and mass spectrometry of pheromones *Ef-1* and *Ef-2*. (a) Elution profile of the gel filtration chromatography of the protein material adsorbed onto reverse-phase cartridges from cell-culture supernatant preparations. In the chromatogram, the shadowed area indicates the pheromone-containing fractions. The points of elution of proteins with standard molecular masses are indicated above the chromatogram. (b) Elution profile of the protein material pooled from the shadowed fractions and separated on reverse-phase chromatography. The applied acetonitrile gradient is indicated by a dotted line and the elution peaks of the two proteins identified as pheromones are labelled *Ef-1* and *Ef-2*. (c) Molecular mass determination of pheromone *Ef-1*. (d) Molecular mass determination of pheromone *Ef-2*.

The determination of the *Ef-1* and *Ef-2* amino acid sequences involved a combined chemical and genetic approach. By subjecting aliquots of the two purified proteins to automated Edman degradation, a N-terminal 15-amino acid sequence, Ser₁-Asp-Cys-His-Gly-Asp-Thr-Glu-Tyr-Leu-Ile-Asp-Asp-Glu-Ser₁₅, was preliminarily obtained. This sequence was identical for the two proteins and included one Cys residue indirectly identified as ‘no signal’. The complete *Ef-1* and *Ef-2* amino acid sequences were then deduced from the nucleotide sequences of the respective macronuclear coding genes, designated *mac-ef-1* and *mac-ef-2*, which were amplified and cloned via a nested RATE-PCR procedure.

The 3′-trailer region and a 218-bp portion of the coding region of the *mac-ef-1* and *mac-ef-2* genes were obtained by sequencing amplicons of a similar 400-bp size generated by a first round of DNA amplifications, in which two degenerate oligonucleotides (“dFW1” and “dFW2”, Table 1) designed on the determined N-terminal amino acid sequence were

used as forward primers in combination with one oligonucleotide (“TEL”, Table 1) specific to the telomeric repeats distinctive of the 5′ and 3′ telomeric ends of *Euplotes* macronuclear genes [35,36].

The 5′-leader region and the remaining portion of the coding region of the two pheromone genes were next obtained by sequencing 1650-bp and 850-bp amplicons generated from a second round of DNA amplifications run with oligonucleotides (“RV1” and “RV2”, Table 1) designed on the 400-bp amplicon sequence and used as reverse primers in combination with the TEL oligonucleotide. The 1650-bp amplicon sequence, terminating with 5′ telomeric repeats, allowed a direct reconstruction of a full-length 1882-bp gene sequence (telomeres excluded), which was recognized to be specific of pheromone *Ef*-2 in relation to its open reading frame (ORF). It encoded a 120-amino acid pre-pro-pheromone precursor in which the calculated molecular mass of the secreted 85-amino acid protein closely matched the 9201.7 Da mass previously measured for the isolated *Ef*-2 protein.

At the opposite of the 1650-bp amplicon, the 850-bp amplicon lacked 5′-end telomeric repeats, and efforts to amplify the whole 5′-leader region necessary to reconstruct the second full-length pheromone gene sequence were unsuccessful. To identify this sequence, the *E. focardii* genome assembled from the same cells of the TN₁ strain used in this work [43] was helpful. Queried on this genome, the 850-bp amplicon sequence localized within a 2318-bp gene sequence (MJUV02005382.1), which was thus recognized to be specific of pheromone *Ef*-1. Its ORF encoded a 126-amino acid pre-pro-pheromone precursor in which the calculated molecular mass of the secreted 85-amino acid protein closely matched the 9245.5 Da mass previously measured for the isolated *Ef*-1 protein. The *mac-ef-1* gene sequence was then validated by sequencing the 2253-bp fragments obtained by DNA amplifications run with two oligonucleotides (“5′-FW1” and “3′-RV” in Table 1) designed near the 5′ and 3′ telomeric ends of the identified genomic sequence (Figure 2b).

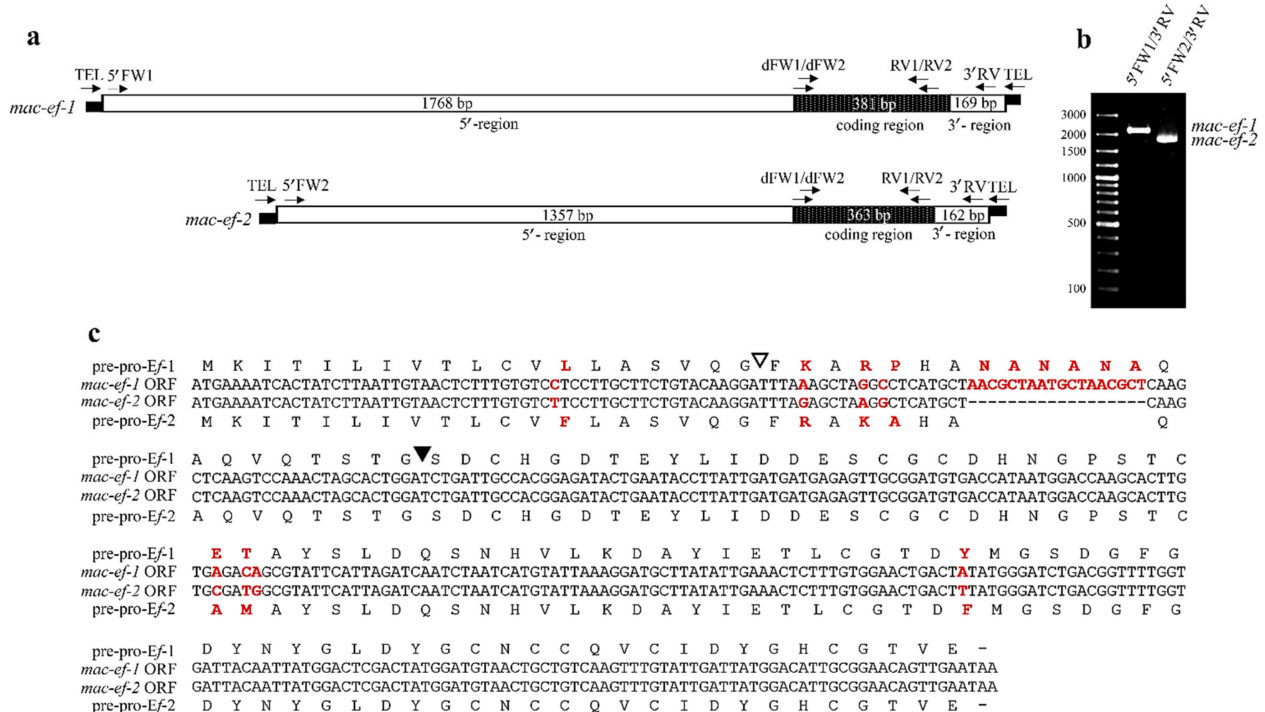


Figure 2. Structure of the *mac-ef-1* and *mac-ef-2* genes. (a) Schematic representations of the two genes, in which bars represent the telomeric ends consisting of C₄A₄ and G₄T₄ repetitions, boxes represent the 5′-leader, coding and 3′-trailer regions of different lengths, and arrows indicate the directions and positions along with the denominations of primers used in the PCR gene amplifications. (b) Agarose gel electrophoresis of PCR fragments obtained by amplifications of *mac-ef-1* and *mac-ef-2*

genes, using the indicated primer combinations. (c) Alignment of the *mac-ef-1* and *mac-ef-2* ORF nucleotide sequences, and deduced amino acid sequences of pre-pro-*Ef-1* and pre-pro-*Ef-2*. Nucleotide and amino acid variations are highlighted in red. In the amino acid sequences, the light and filled arrowheads indicate the putative cleavage sites between the signal-peptide and pro-segment, and between the pro-segment and secreted pheromone, respectively.

Instead, the query of the 1882-bp *mac-ef-2* sequence on the *E. focardii* genome identified only a partial 1355-bp sequence (MJUV02014221.1) equivalent to the gene 5'-leader region. The gene validation was obtained by sequencing 1778-bp fragments from a DNA amplification run with two oligonucleotides ("5'FW2" and "3'RV" in Table 1) designed close to the gene 5' and 3' telomeric ends (Figure 2b).

3.2. Pheromone Gene Structure

The ORFs spanning from A₁₇₆₉TG to TAA₂₁₄₉ in the *mac-ef-1* gene, and from A₁₃₅₈TG to TAA₁₇₂₀ in the *mac-ef-2* gene, were unequivocally identified as specific for the synthesis of cytoplasmic (immature) pheromone precursors, pre-pro-*Ef-1* and pre-pro-*Ef-2* of 126 and 120 amino acids, respectively, destined to be proteolytically processed to remove the signal-peptide and pro-segment before the secretion of the mature pheromone forms (Figure 2c). The coding sequences of these pheromone precursors vary only for a deletion of an 18-bp segment responsible for a shorter pro-segment in the *mac-ef-2* gene (16 amino acids vs. 22 in the pro-segment of pre-pro-*Ef-1*), and eight nucleotide mutations that are responsible for seven amino acid substitutions. One localizes in the 19-amino acid pre-segment, three in the 16- or 22-amino acid pro-segment, and three in the 85-amino acid secreted protein.

At the level of the two ORF flanking regions, the *mac-ef-1* and *mac-ef-2* genes are characterized by a tight conservation (96% sequence identity) of their relatively short 3'-trailer regions (169 and 162 bp, respectively), lacking a canonical AATAAA polyadenylation motif, which is likely replaced with a TTATTT motif located 15 bp downstream the ORF TAA stop codon. On the other hand, the 5'-leader regions are poorly conserved (42% sequence identity) and markedly extended (1768 bp in *mac-ef-1*, and 1357 bp in *mac-ef-2*). This poor conservation, that has no equal in allelic pheromone genes characterized from other *Euplotes* species [19–21,28,44], is also due to the loss in the *mac-ef-2* gene of five segments of 73, 90, 93, 146 and 169 bp from the 5'-leader region (Supplementary Figure S1). The remarkable extension of this region likely reflects the inclusion of intron sequences, as suggested by the presence of multiple canonical GTA-TAG splicing sites.

3.3. Pheromone Primary Structure

The 85-amino acid *Ef-1* and *Ef-2* sequences, which result from the cytoplasmic precursors by proteolytic removals of the signal peptide and pro-segment at predicted Gly-Phe and Gly-Ser sites, respectively, are closely similar (Figure 3a). They differ for only three amino acid substitutions, Glu₂₇/Ala, Thr₂₈/Met and Tyr₅₂/Phe, the first two of which credit more structural relevance, since the substitution of charged and polar residues with hydrophobic residues locally alters the molecule physico-chemical properties [45]. Both sequences start with a Ser residue, are strongly acidic with an isoelectric point of 3.8, and include ten unevenly spaced Cys residues, destined to pair into intra-chain disulfide bridges warranting a global tight stabilization of the *Ef-1* and *Ef-2* molecular structures. Four cluster into a CX₁₋₃CCX₂₋₃C motif (X stands for any residue) that finds a precise counterpart in *E. crassus*, *E. nobilii* and *E. petzi* pheromones (Supplementary Figure S2) [19,23,28].

In line with cold-adapted *E. nobilii* pheromones [28,46] and psychrophilic proteins in general [47], *Ef-1* and *Ef-2* amino acid composition strongly reflects a molecular adaptation to low temperatures, which in the first place requires weakening the packing of the molecular structure and an improved molecular backbone flexibility. Compared with the pheromones of the mesophilic *Euplotes* species, the very hydrophobic residues (Ala, Ile, Leu, Met, Phe, Pro, Val and Trp), which preferentially bury into the protein structural core, are under-represented (or not represented, Trp), overall accounting for only 24.7% of the amino acid composition. In contrast, charged and polar residues are largely preferred, in

particular Asp, Thr, Tyr and Ser residues, which, together with 11 Gly residues (neutral), constitute exactly half the amino acid composition. Their overrepresentation, Asp residues (amounting to 12!) in particular, finds a close parallel in psychrophilic proteins of bacteria [47], and is justified by a key role in improving the backbone structural flexibility by avoiding inclusion into helices, or favoring their destabilization [48,49].

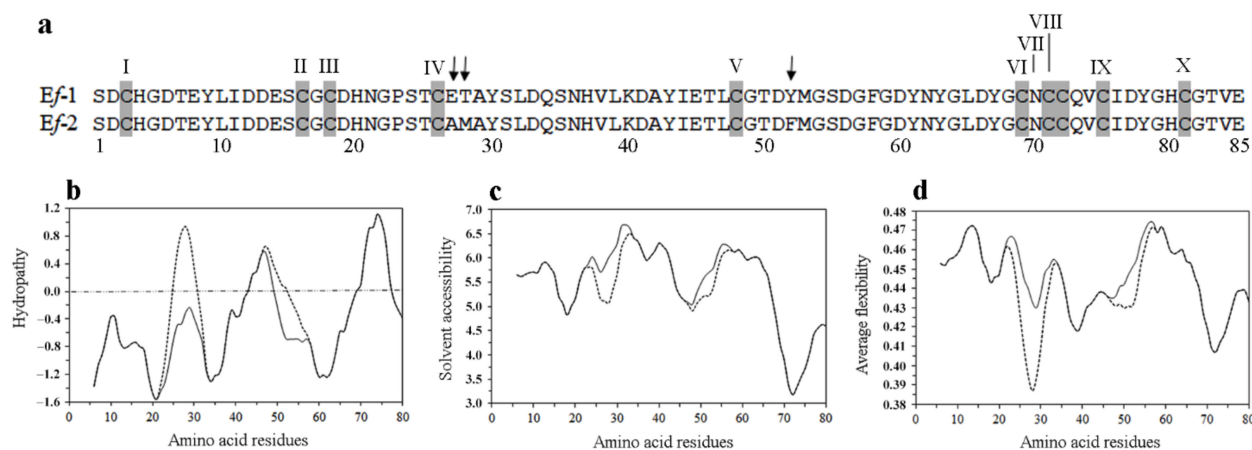


Figure 3. Amino acid sequence alignment and physico-chemical traits of Ef-1 and Ef-2 pheromones. (a) Sequence alignment with Cys residues, progressively numbered I to X, highlighted in grey and amino acid substitutions indicated by arrows. (b) Hydropathy, (c) solvent accessibility and (d) average flexibility plots of the Ef-1 (solid line) and Ef-2 (dotted line) pheromones, as generated by predictive algorithms [50–52].

The hydropathy plot (Figure 3b) of the Ef-1 and Ef-2 amino acid sequences unequivocally identifies the amino-terminal and central regions as mostly characterized by three strongly hydrophilic and poorly structured districts. The major one spanning 20 residues from Cys-V to Cys-VI includes six of the 11 Gly residues, and each is associated with a local increase in solvent accessibility and molecular backbone flexibility (Figure 3c,d). The carboxy-terminal region is instead identified as the major hydrophobic district preceded by one (Ef-1 sequence) or two (Ef-2 sequence) minor hydrophobic sites, all coincident with a reduced solvent accessibility and a more rigid structure. The likely organization of these hydrophobic sites into helical secondary structures is strongly supported in particular for the major one. Its counterparts in the carboxy-terminal region of the determined *Euplotes* pheromone molecular structures are all shown to be organized into a regular α -helix, which is tightly conserved at both intra- and inter-specific levels, serving a key function in the pheromone/receptor interactions [30].

4. Discussion

The structural characterization of *E. focardii* pheromones and pheromone coding genes is an initial step to outlining a picture of how these molecular structures evolve in concert with *Euplotes* speciation. Using a simplified version of the *Euplotes* phylogenetic tree commonly proposed with seven distinct clades, annotated from I to VII starting from the earliest branching one [39], Figure 4 delineates the phylogenetic relationships and ecological diversification of *E. focardii* and the five other *Euplotes* species (*E. crassus*, *E. nobilii*, *E. octocarinatus*, *E. petzi* and *E. raikovi*) with known pheromone and pheromone coding gene structures [12,13,19]. It appears that *E. petzi* (marine, polar) and *E. crassus* (marine, temperate) represent the two more distantly related species, the former splitting into the basal clade I and the latter into the terminal clade VII. The four other species instead separate into in-between branching clades: *E. raikovi* (marine, temperate) and *E. nobilii* (marine, polar) into clade IV; *E. octocarinatus* (lacustrine, temperate) into clade V including only freshwater species; and *E. focardii* (marine, polar) into the widest and least resolved clade VI.

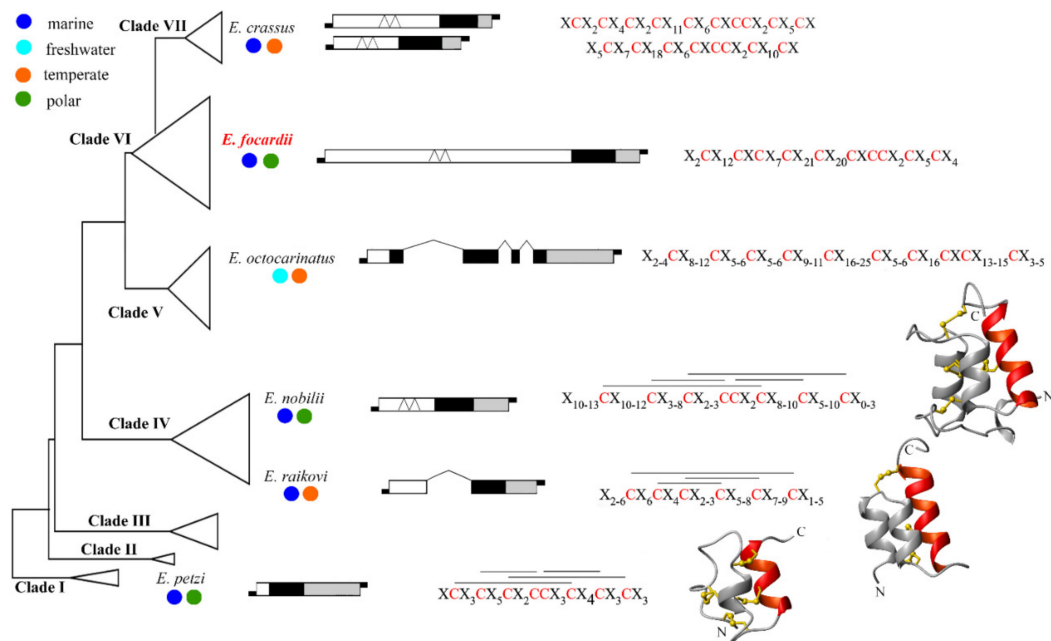


Figure 4. *Euplotes* pheromone and pheromone-gene structural evolution. The *Euplotes* phylogenetic tree, articulated on seven clades, is reported in a very simplified version to highlight the positions of the species with known pheromone and pheromone-gene structures. Indicated for each species is: (i) ecology, (ii) the pheromone gene structure, (iii) the pheromone amino acid consensus sequence, and (iv) the pheromone native three-dimensional structure, known from *E. petzi*, *E. raikovi* and *E. nobilii*. In the pheromone gene diagrams, drawn to scale except the telomeric extremities represented by bars, the 5'-leader, coding, and 3'-trailer regions are represented by empty, filled, and grey boxes, respectively. Inter-box zig-zag lines indicate intron-sequence splicing, as determined in the 5'-leader and coding regions of *E. raikovi* [44] and *E. octocarınatus* [21] pheromone genes, respectively. The likely presence of intron sequences (as deduced from intron-splicing motifs) in the 5'-leader region of the *E. nobilii*, *E. focardii* and *E. crassus* pheromone genes is indicated by intra-box zig-zag lines. In the amino acid consensus sequences, the conserved Cys residues are highlighted in red, X stands for any residue, and lines connect Cys residues according to their disulfide bridges. In *E. petzi*, the consensus sequence is pertinent to pheromones *Ep*-1, *Ep*-2, *Ep*-3 and *Ep*-4 [19]; in *E. raikovi*, to pheromones *Er*-1, *Er*-2, *Er*-4, *Er*-5, *Er*-6, *Er*-7, *Er*-8, *Er*-10, *Er*-11, *Er*-13, *Er*-20, *Er*-21, and *Er*-22 (the 51-amino acid eccentric family member *Er*-23 excluded) [26,44]; in *E. nobilii*, to pheromones *En*-1, *En*-2, *En*-6, *En*-A1, *En*-A2, *En*-A3, and *En*-A4 [28]; in *E. octocarınatus*, to pheromones *Phr*1, *Phr*1*, *Phr*2, *Phr*2*, *Phr*3, *Phr*3*, *Phr*⁻⁵, and *Phr*⁻² (the *Phr*4-amino acid eccentric family member excluded) [20,21,53]; in *E. focardii*, to pheromones *Ef*-1 and *Ef*-2 [this work]; in *E. crassus*, to pheromones *Ec*-1, *Ec*-2 and *Ec*-3 representing one sub-family, and to pheromone *Ec*- α representing a second sub-family [23]. The pheromone three dimensional configurations, shown in frontal view and ribbon diagrams, are specific of pheromones *Ep*-1 [19], *Er*-1 [24] and *En*-6 [54] taken as representative of *E. petzi*, *E. raikovi* and *E. nobilii* pheromone families, respectively. The disulfide bridges are in yellow, the structurally more conserved helix 3 is highlighted in red, and N and C indicate the molecule amino- and carboxyl terminus, respectively.

The overall interspecific comparison of the pheromone and pheromone gene structures (which in the figure are typified alongside each species) provides evidence for an evolutionary tendency of these structures to jointly increase in extension and complexity.

In the structural evolution of the pheromone genes, the most manifest variations coincide with the transition from *E. petzi* to the other species and primarily involve the 5'-leader region or, uniquely in *E. octocarınatus*, the coding region, while the 3'-trailer region is left practically untouched. The 5'-leader region of minimal length (72 bp) in *E. petzi*, at odd with much more extended coding and 3'-trailer regions (219 and 368 bp, respectively), abruptly lengthens from 5 to 7 folds in *E. raikovi*, *E. nobilii* and *E. crassus*—which is distinguished by a duplicated

micronuclear *mat*-gene locus [23]—and up to 24 folds in *E. focardii*. The result is that it largely overcomes in length the coding and 3'-trailer regions that remain substantially unvaried. Only in *E. octocarinatus*, the 5'-leader region eccentrically maintains a minimal length as in *E. petzi*, and its contribution to the gene sequence lengthening is replaced by the coding region that becomes 3 folds longer than in all the other species.

These increases in the extension of the pheromone genes are explained by the findings [21,44] that the 5' leader region of the *E. raikovi* pheromone genes and the coding region of the *E. octocarinatus* pheromone genes include intron sequences (up to 514 bp long in *E. octocarinatus*). They thus assume strong functional relevance considering the central role that the splicing of these intron sequences plays in the mechanism of pheromone gene expression. Although intron sequences have not directly been identified in the 5'-leader region of the *E. nobilii*, *E. focardii* and *E. crassus* pheromone genes, their presence is strongly supported by the inclusion of multiple canonical GTA/TAG splicing motifs which, instead, are completely lacking throughout the *E. petzi* pheromone gene sequences [19].

Similar to the pheromone gene structural evolution being shown to primarily operate by widening the 5'-leader region or the coding region with the inclusion of intron sequences, the pheromone structural evolution appears to primarily operate by extending one or more sequence segments spanning preferentially between conserved Cys residues lying in the sequence middle (Supplementary Figure S2). By including new amino acid combinations, these extensions clearly provide *Euplotes* pheromones with the necessary structural substrate to adapt and optimize their activity in any new environment.

No inter-cysteine segment includes more than four residues in *E. petzi* pheromones; a maximum of eight or nine residues are included in two inter-cysteine segments of *E. raikovi* pheromones, and systematically more than 10 (up to 25) residues are included in one or more inter-cysteine segments of pheromones of all the other species. These local sequence extensions have functionally been related to cold-adaptation in *E. focardii* on the basis of predictive algorithms applied to the determined Ef-1 and Ef-2 amino acid sequences. This association was previously more directly supported by comparing the pheromones of the two phylogenetically closely related marine species, the polar *E. nobilii* and the temperate *E. raikovi*, for their three-dimensional structures [46] and the dynamics of thermal unfolding and refolding [55]. Most likely, however, the functions of the local pheromone sequence extensions are multiple and varied, and their identification requires the determination of the pheromone three-dimensional structures from a wider spectrum of ecologically well-distinct *Euplotes* species.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms10061089/s1>, Figure S1: Nucleotide sequence alignment of *mac-ef-1* and *mac-ef-2* macronuclear pheromone genes. Figure S2: Amino acid sequence alignment of *E. focardii* pheromones Ef-1 and Ef-2 with other *Euplotes* pheromones carrying the CX₁₋₃CCX₂₋₃ C motif.

Author Contributions: Conceptualization, P.L., C.A. and A.V.; methodology, C.A. and A.C.; investigation, C.A., A.C., A.V. and Y.J.; resources, P.L.; data curation, A.V.; writing—original draft preparation, P.L., A.V. and C.A.; writing—review and editing, P.L. and A.V.; funding acquisition, C.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Programma Nazionale di Ricerca in Antartide of the Italian Ministry of Research, grant number PNRA18_00077.

Data Availability Statement: Macronuclear pheromone gene sequences are deposited in GenBank database under the accession numbers: *mac-ef-1*, ON428185; *mac-ef-2*, ON428186.

Acknowledgments: Authors thank Graziano Guella (University of Trento) for mass spectrometry analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stephens, K. Pheromones among the prokaryotes. *Crit. Rev. Microbiol.* **1986**, *13*, 309–334. [[CrossRef](#)] [[PubMed](#)]
2. Cook, L.C.; Federle, M.J. Peptide pheromone signaling in *Streptococcus* and *Enterococcus*. *FEMS Microbiol. Rev.* **2014**, *38*, 473–492. [[CrossRef](#)] [[PubMed](#)]
3. Monnet, V.; Juillard, V.; Gardan, R. Peptide conversations in Gram-positive bacteria. *Crit. Rev. Microbiol.* **2016**, *42*, 339–351. [[CrossRef](#)] [[PubMed](#)]
4. Stabb, E.V. Could positive feedback enable bacterial pheromone signaling to coordinate behaviors in response to heterogeneous environmental cues? *MBio* **2018**, *9*, e00098-18. [[CrossRef](#)] [[PubMed](#)]
5. Wyatt, T.D. *Pheromones and Animal Behaviour*; Cambridge University Press: Cambridge, UK, 2003; Volume 626.
6. Shorey, H.H. *Animal Communication by Pheromones*; Academic Press: New York, NY, USA, 2013.
7. Luporini, P.; Miceli, C.; Ortenzi, C.; Vallesi, A. Ciliate Pheromones. In *Signaling Mechanisms in Protozoa and Invertebrates*; Progress in Molecular and Subcellular Biology Series; Csaba, G., Müller, W.E.G., Eds.; Springer: Berlin/Heidelberg, Germany, 1996; Volume 17, pp. 80–104.
8. Vallesi, A.; Giuli, G.; Bradshaw, R.A.; Luporini, P. Autocrine mitogenic activity of pheromones produced by the protozoan ciliate *Euplotes raikovi*. *Nature* **1995**, *376*, 522–524. [[CrossRef](#)]
9. Nobili, R.; Luporini, P.; Dini, F. Breeding system, species relationships and evolutionary trends in some species of Euplotidae (Hypotrichida, Ciliata). In *Marine Organisms*; Battaglia, B., Beardmore, J., Eds.; Plenum Press: New York, NY, USA, 1978; pp. 591–616.
10. Dini, F.; Nyberg, D. Sex in ciliates. In *Advances in Microbial Ecology*; Jones, J.G., Ed.; Plenum Press: New York, NY, USA, 1993; Volume 13, pp. 85–153.
11. Miyake, A. Fertilization and sexuality in ciliates. In *Ciliates, Cells and Organisms*; Hausmann, K., Bradbury, P.C., Eds.; Gustav-Fisher Verlag: Stuttgart, Germany, 1996; pp. 244–290.
12. Luporini, P.; Alimenti, C.; Ortenzi, C.; Vallesi, A. Ciliate mating types and their specific protein pheromones. *Acta Protozool.* **2005**, *44*, 89–101.
13. Luporini, P.; Pedrini, B.; Alimenti, C.; Vallesi, A. Revisiting fifty years of research on pheromone signaling in ciliates. *Eur. J. Protistol.* **2016**, *55*, 26–38. [[CrossRef](#)]
14. Luporini, P.; Alimenti, C.; Pedrini, B.; Vallesi, A. Ciliate communication via water-borne pheromones. In *Biocommunication of Ciliates*; Witzany, G., Nowacki, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 159–174.
15. Grell, K.G. Die Konjugation von *Ephelota gemmipara* R. Hertwig. *Arch Protistenk* **1953**, *98*, 287–326.
16. Sonneborn, T.M. Genetics of cell–cell interaction in ciliates. In *The Molecular Basis of Cell–Cell Interaction*; Lerner, R.A., Bergsma, D., Eds.; Alan R. Liss: New York, NY, USA, 1978; pp. 417–427.
17. Kubota, T.; Tokoroyama, T.; Tsukuda, Y.; Koyama, H.; Miyake, A. Isolation and structure determination of blepharismine, a conjugation initiating gamone in the ciliate *Blepharisma*. *Science* **1973**, *179*, 400–402. [[CrossRef](#)]
18. Sugiura, M.; Harumoto, T. Identification, characterization, and complete amino acid sequence of the conjugation-inducing glycoprotein (blepharone) in the ciliate *Blepharisma japonicum*. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 14446–14451. [[CrossRef](#)]
19. Pedrini, B.; Suter-Stahel, T.; Vallesi, A.; Alimenti, C.; Luporini, P. Molecular structures and coding genes of the water-borne protein pheromones of *Euplotes petzi*, an early diverging polar species of *Euplotes*. *J. Eukaryot. Microbiol.* **2017**, *64*, 164–172. [[CrossRef](#)]
20. Dieckhoff, H.S.; Freiburg, M.; Heckmann, K. The isolation of gamones 3 and 4 of *Euplotes octocarinatus*. *Eur. J. Biochem.* **1987**, *168*, 89–94. [[CrossRef](#)]
21. Brünen-Nieveler, C.; Weiligmann, J.C.; Hansen, B.; Kuhlmann, H.W.; Möllenbeck, M.; Heckmann, K. The pheromones and pheromone genes of new stocks of *Euplotes octocarinatus* species complex. *Eur. J. Protistol.* **1998**, *34*, 124–132. [[CrossRef](#)]
22. Alimenti, C.; Vallesi, A.; Federici, S.; Di Giuseppe, G.; Dini, F.; Carratore, V.; Luporini, P. Isolation and structural characterization of two water-borne pheromones from *Euplotes crassus*, a ciliate commonly known to carry membrane-bound pheromones. *J. Eukaryot. Microbiol.* **2011**, *58*, 234–241. [[CrossRef](#)]
23. Vallesi, A.; Alimenti, C.; Federici, S.; Di Giuseppe, G.; Dini, F.; Guella, G.; Luporini, P. Evidence for gene duplication and allelic codominance (not hierarchical dominance) at the mating-type locus of the ciliate, *Euplotes crassus*. *J. Eukaryot. Microbiol.* **2014**, *61*, 620–629. [[CrossRef](#)]
24. Mronga, S.; Luginbühl, P.; Brown, L.R.; Ortenzi, C.; Luporini, P.; Bradshaw, R.A.; Wüthrich, K. The NMR solution structure of the pheromone Er-1 from the ciliated protozoan *Euplotes raikovi*. *Protein Sci.* **1994**, *3*, 1527–1536. [[CrossRef](#)]
25. Luginbühl, P.; Ottiger, M.; Mronga, S.; Wüthrich, K. Structure comparison of the pheromones Er-1, Er-10, and Er-2 from *Euplotes raikovi*. *Protein Sci.* **1994**, *3*, 1537–1546. [[CrossRef](#)]
26. Zahn, R.; Damberger, F.; Ortenzi, C.; Luporini, P.; Wüthrich, K. NMR structure of the *Euplotes raikovi* pheromone Er-23 and identification of its five disulfide bonds. *J. Mol. Biol.* **2001**, *313*, 923–931. [[CrossRef](#)]
27. Placzek, W.J.; Etezady-Esfarjani, T.; Herrmann, T.; Pedrini, B.; Peti, W.; Alimenti, C.; Luporini, P.; Wüthrich, K. Cold-adapted signal proteins: NMR structures of pheromones from the Antarctic ciliate *Euplotes nobilii*. *IUBMB Life* **2007**, *59*, 578–585. [[CrossRef](#)]
28. Vallesi, A.; Alimenti, C.; Pedrini, B.; Di Giuseppe, G.; Dini, F.; Wüthrich, K.; Luporini, P. Coding genes and molecular structures of the diffusible signaling proteins (pheromones) of the polar ciliate, *Euplotes nobilii*. *Mar. Genom.* **2012**, *8*, 9–13. [[CrossRef](#)]

29. Weiss, M.S.; Anderson, D.H.; Raffioni, S.; Bradshaw, R.A.; Ortenzi, C.; Luporini, P.; Eisenberg, D. A cooperative model for ligand recognition and cell adhesion: Evidence from the molecular packing in the 1.6 Å crystal structure of the pheromone Er-1 from the ciliate protozoan *Euplotes raikovi*. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 10172–10176. [[CrossRef](#)]
30. Pedrini, B.; Finke, A.D.; Marsh, M.; Luporini, P.; Vallesi, A.; Alimenti, C. Crystal structure of the pheromone Er-13 from the ciliate *Euplotes raikovi*, with implications for a protein–protein association model in pheromone/receptor interactions. *J. Struct. Biol.* **2002**, *214*, 107812. [[CrossRef](#)]
31. Ortenzi, C.; Luporini, P. Competition among homologous polypeptide pheromones of the ciliate *Euplotes raikovi* for binding to each other’s receptors. *J. Eukaryot. Microbiol.* **1995**, *42*, 242–248. [[CrossRef](#)]
32. Kimball, R.F. The nature and inheritance of mating types in *Euplotes patella*. *Genetics* **1942**, *27*, 269–285. [[CrossRef](#)]
33. Heckmann, K. Paarungssystem und genabhängige paarungstypdifferenzierung bei dem hypotrichen ciliaten *Euplotes vannus* O. F. Müller. *Arch. Protistenk.* **1963**, *106*, 393–421.
34. Dini, F.; Luporini, P. Mating-type polymorphic variation in *Euplotes minuta* (Ciliophora: Hypotrichida). *J. Protozool.* **1985**, *32*, 111–118. [[CrossRef](#)]
35. Hoffman, D.C.; Anderson, R.C.; DuBois, M.L.; Prescott, D.M. Macronuclear gene-sized molecules of hypotrichs. *Nucleic Acids Res.* **1995**, *23*, 1279–1283. [[CrossRef](#)]
36. Jahn, C.L.; Klobutcher, L.A. Genome remodeling in ciliated protozoa. *Annu. Rev. Microbiol.* **2002**, *56*, 489–520. [[CrossRef](#)]
37. Valbonesi, A.; Luporini, P. A new marine species of *Euplotes* (Ciliophora, Hypotrichida) from Antarctica. *Bull. Br. Mus. Nat. Hist. Zool.* **1990**, *56*, 57–61.
38. Zhao, Y.; Yi, Z.; Warren, A.; Song, W.B. Species delimitation for the molecular taxonomy and ecology of the widely distributed microbial eukaryote genus *Euplotes* (Alveolata, Ciliophora). *Proc. R. Soc. B* **2018**, *285*, 20172159. [[CrossRef](#)]
39. Valbonesi, A.; Di Giuseppe, G.; Vallesi, A.; Luporini, P. Two new species of *Euplotes* with cirrotype-9, *Euplotes foissneri* sp. nov. and *Euplotes warreni* sp. nov. (Ciliophora, Spirotrichea, Euplotida), from the coasts of Patagonia: Implications from their distant, early and late branching in the *Euplotes* phylogenetic tree. *Int. J. Syst. Evol. Microbiol.* **2021**, *71*, 004568.
40. Raffioni, S.; Miceli, C.; Vallesi, A.; Chowdhury, S.K.; Chait, B.T.; Luporini, P.; Bradshaw, R.A. Primary structure of *Euplotes raikovi* pheromones: Comparison of five sequences of pheromones from cells with variable mating interactions. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 2071–2075. [[CrossRef](#)] [[PubMed](#)]
41. Valbonesi, A.; Luporini, P. Biology of *Euplotes focardii*, an Antarctic ciliate. *Polar Biol.* **1993**, *13*, 489–493. [[CrossRef](#)]
42. Vallesi, A.; Di Pretoro, B.; Ballarini, P.; Apone, F.; Luporini, P. A novel protein kinase from the ciliate *Euplotes raikovi* with close structural identity to the mammalian intestinal and male-germ cell kinases: Characterization and functional implications in the autocrine pheromone signaling Loop. *Protist* **2010**, *161*, 250–263. [[CrossRef](#)]
43. Mozzicafreddo, M.; Pucciarelli, S.; Swart, E.C.; Piersanti, A.; Emmerich, C.; Migliorelli, G.; Ballarini, P.; Miceli, C. The macronuclear genome of the Antarctic psychrophilic marine ciliate *Euplotes focardii* reveals new insights on molecular cold adaptation. *Sci. Rep.* **2021**, *11*, 18782. [[CrossRef](#)]
44. Ricci, F.; Candelori, A.; Brandi, A.; Alimenti, C.; Luporini, P.; Vallesi, A. The sub-chromosomal macronuclear pheromone genes of the ciliate *Euplotes raikovi*: Comparative structural analysis and insights into the mechanism of expression. *J. Eukaryot. Microbiol.* **2019**, *66*, 376–384. [[CrossRef](#)]
45. Pace, C.N.; Scholtz, J.M.; Grimsley, G.R. Forces stabilizing proteins. *FEBS Lett.* **2014**, *588*, 2177–2184. [[CrossRef](#)]
46. Alimenti, C.; Vallesi, A.; Pedrini, B.; Wüthrich, K.; Luporini, P. Molecular cold-adaptation: Comparative analysis of two homologous families of psychrophilic and mesophilic signal proteins of the protozoan ciliate, *Euplotes*. *IUBMB Life* **2009**, *61*, 838–845. [[CrossRef](#)]
47. Metpally, R.P.R.; Reddy, B.V.B. Comparative proteome analysis of psychrophilic versus mesophilic bacterial species: Insights into the molecular basis of cold adaptation of proteins. *BMC Genomics* **2009**, *10*, 11. [[CrossRef](#)]
48. Lewis, P.N.; Go, N.; Go, M.; Kotelchuck, D.; Scheraga, H.A. Helix probability profiles of denatured proteins and their correlation with native structures. *Proc. Natl. Acad. Sci. USA* **1970**, *65*, 810–815. [[CrossRef](#)]
49. Lee, D.Y.; Kim, K.A.; Yu, Y.G.; Kim, K.S. Substitution of aspartic acid with glutamic acid increases the unfolding transition temperature of a protein. *Biochem. Biophys. Res. Commun.* **2004**, *320*, 900–906. [[CrossRef](#)]
50. Kyte, J.; Doolittle, R.F. Hydrophobicity. *J. Mol. Biol.* **1982**, *157*, 105–132. [[CrossRef](#)]
51. Janin, J. Free energy of transfer from inside to outside of a globular protein. *Nature* **1979**, *277*, 491–492. [[CrossRef](#)]
52. Bhaskaran, R.; Ponnuswamy, P.K. Average flexibility index. *Int. J. Pept. Protein Res.* **1988**, *32*, 242–255.
53. Meyer, F.; Schmidt, H.J.; Heckmann, K. Pheromone 4 gene of *Euplotes octocarinatus*. *Dev. Genet.* **1992**, *13*, 16–25. [[CrossRef](#)]
54. Pedrini, B.; Plackzek, W.J.; Koculi, E.; Alimenti, C.; La Terza, A.; Luporini, P.; Wüthrich, K. Cold-adaptation in seawater-borne signal proteins: Sequence and NMR structure of the pheromone En-6 from the Antarctic ciliate *Euplotes nobilii*. *J. Mol. Biol.* **2007**, *372*, 277–286. [[CrossRef](#)]
55. Geralt, M.; Alimenti, C.; Vallesi, A.; Luporini, P.; Wüthrich, K. Thermodynamic stability of psychrophilic and mesophilic pheromones of the protozoan ciliate *Euplotes*. *Biology* **2013**, *2*, 142–150. [[CrossRef](#)]