Genetic Diversity in the *Paramecium aurelia* Species Complex

Francesco Catania, François Wurmser,†‡ Alexey A. Potekhin,§ Ewa Przyboś,‖ and Michael Lynch**

*Department of Biology, Indiana University; †Laboratory Evolution, Genomes and Speciation; CNRS-UPR9034, Gif sur Yvette Cedex, France; ‡Universitè Paris-Sud 11, 91405 Orsay Cedex, France; §Laboratory of Protozoan Karyology, Faculty of Biology and Soil Science, St Petersburg State University; Oranienbaumskoye sh., 2, Stary Peterhof, 198504 Saint Petersburg, Russia; and ‖Department of Experimental Zoology, Institute of Systematics and Evolution of Animals, Polish Academy of Sciences, Kraków, Poland

Current understanding of the population genetics of free-living unicellular eukaryotes is limited, and the amount of genetic variability in these organisms is still a matter of debate. We characterized—reproductively and genetically—worldwide samples of multiple *Paramecium* species belonging to a cryptic species complex, *Paramecium aurelia*, whose species have been shown to be reproductively isolated. We found that levels of genetic diversity both in the nucleus and in the mitochondrion are substantial within groups of reproductively compatible *P. aurelia* strains but drop considerably when strains are partitioned according to their phylogenetic groupings. Our study reveals the existence of discrepancies between the mating behavior of a number of *P. aurelia* strains and their multilocus genetic profile, a controversial finding that has major consequences for both the current methods of species assignment and the species problem in the *P. aurelia* complex.

Introduction

Protists are among the most abundant and diverse eukaryotic groups (Patterson 1999), yet current knowledge of these organisms is surprisingly sparse, compared with that for animals, plants, and fungi. Two aspects of microbial eukaryotes that still remain controversial are the level of species diversity and their geographic distribution (Finlay and Fenchel 1999; Foissner 1999). One school of thought maintains that protist species are relatively few, cosmopolitan, and ubiquitous (Fenchel et al. 1997); another supports essentially the opposite view, that is, protist species are numerous and tend to be endemic and isolated (Foissner 2006). The currently limited understanding of the protist world (Weisse 2008) makes it difficult at this time to support one hypothesis or the other, in part due to the paucity of population-genetic studies that could shed light on the global (effective) population sizes of free-living species.

In a species, the genetic effective population size (N_e) (Wright 1931)—which modulates the efficiency of natural selection via its effects on random genetic drift (Ohta 1992)—reflects the number of breeding individuals as well as aspects of the breeding system, population structure, and degree of linkage, and $4N_e\mu$ (where μ is the neutral mutation rate) is the expected neutral genetic variation (average nucleotide heterozygosity per site) of a diploid population at mutation–drift equilibrium. As estimates of μ are available for a number of species (Lynch 2006a), measures of neutral genetic variation (e.g., diversity at synonymous sites in protein-coding genes) allow N_e to be inferred indirectly from estimates of variation at silent sites in natural populations.

Although large-scale polymorphism studies of protist pathogens (e.g., Grigg et al. 2001; Mu et al. 2002) have provided measures of neutral genetic variation (Lynch and Conery 2003; Lynch 2006b), the neutral genetic variation of pathogenic organisms is expected to be biased downward and not directly applicable to free-living taxa. On the other hand, a number of small-scale polymorphism surveys of protein-coding loci in free-living microbes, freshwater ciliates in particular, have produced conflicting estimates of intraspecific variation (Gerber et al. 2002; Barth et al. 2006; Katz et al. 2006; Snoke et al. 2006), thus impeding wider conclusions.

Not only can polymorphism surveys based on single or a few genes produce estimates of $4N_e\mu$ that are potentially misleading, but poorly defined species boundaries also make the estimation of intraspecific variation problematic. The latter issue may well represent a confounding factor in the computation of genetic diversity for microbial organisms (Daubin and Moran 2004), as the process of "species" identification is a difficult task, given the frequent existence of cryptic species (Waugh 2007), and may require a combination of approaches that include molecular, morphological, and ecological studies (Weisse 2008).

In this sense, the free-living *Paramecium* is a very suitable system, as the "species problem" in this ciliate group, and in *P. aurelia* in particular, has been intensely investigated. *Paramecium aurelia* is a complex that contains 15 species (Sonneborn 1975; Aufderheide et al. 1983). Despite being morphologically identical, these species are genetically isolated and, to quote Sonneborn (1975), "the distinguishing characters of each of the [then 14] species include its mating reactions, breeding relations, and mode of inheritance of mating type." Specifically, each of the *P. aurelia* species contains two mating types—E (for "even") and O (for "odd")—that determine intraspecific sexual reproduction by conjugation (Sonneborn 1937). When two clones that have either the same mating type or belong to different species are mixed, no conjugation occurs. Although weak cross-reactions between distinct mating types of different species may be observed (e.g., between *Paramecium tetraurelia* and *Paramecium octaurelia* or between *Paramecium primaurelia* and *Paramecium pentaurelia*) (Beale and Preer 2008), these reactions result in the death of the conjugating cells or in the sterility of their F1s. The reproductive cohesion of *P. aurelia* species (Sonneborn 1975), the available macronuclear genome sequence of *P. tetraurelia* (Aury et al. 2006), and a variety of worldwide samples collected during both the past century and in very recent years,

Key words: *Paramecium*, genetic diversity, effective population size, speciation, cryptic species, mating group switching.

E-mail: fcatania@indiana.edu.

Mol. Biol. Evol. 26(2):421–431. 2009
doi:10.1093/molbev/msn266
Advance Access publication November 20, 2008

make this species complex an appealing candidate for a population-genetic study.

In the attempt to shed light on extant controversial results (Katz et al. 2006; Snoke et al. 2006) and to expand our knowledge on the levels of genetic diversity in free-living unicellular eukaryotes, we performed a polymorphism survey of multiple nuclear and mitochondrial loci in a large set of *P. aurelia* strains. Our analysis aims to overcome potential limitations of previous studies on free-living unicellular eukaryotes, such as limited sample sizes and both restricted number of genes and sampled locations. Along with a number of insights to the population genetics of multiple species of the *P. aurelia* complex, we found a surprising incongruence between mating behavior and strains' genetic profiles, which complicates definitive conclusions on intraspecific levels of genetic variability.

Material and Methods

Paramecium Strains

We examined a total of 80 *Paramecium* strains, 79 members of the *P. aurelia* species complex (Sonneborn 1975; Aufderheide et al. 1983) (table 1), and 1 *Paramecium multimicronucleatum* strain. A total of 26 strains were kindly provided by J. Cohen, H. Schmidt, M. Simon, T. Berendonk, and K. Aufderheide, whereas the remaining 54 strains are maintained in the Collection of *Paramecium* of the Laboratory of Protozoan Karyology, St Petersburg State University. All strains are fully homozygous, following multiple rounds of autogamy.

Culturing and Identification of *Paramecium* Strains

Paramecia were cultured in a medium made of dried lettuce (or Yeast extract, Cerophyl, Na_2HPO_4, and Stigmasterol) and distilled water inoculated with *Enterobacter aerogenes* and identified according to the methods of Sonneborn (1970). Clones matured for conjugation were mated with the reactive complementary mating types of standard strains of all species of the *P. aurelia* complex.

In intrastrain and interstrain crosses, the F_1 generation was obtained by conjugation and F_2 by autogamy (using the method of daily isolation lines). The occurrence of the desired stage of autogamy (specimens at the stage of two macronuclear anlagen) was examined on preparations stained with aceto-carmine. Survival of clones in both generations was estimated as percentages. According to Chen (1956), clones can be considered as surviving after passing six to seven fissions during 72 h after separation of partners of conjugation or postautogamous caryonides. The methods are described in detail in Przyboś (1975).

DNA Extraction and Gene Sequencing

Genomic DNA extraction was performed by incubating *Paramecium* cells for 20 min at 90 °C with 5% Chelex 100 Resin. We surveyed genetic variability at 10 nuclear and 5 mitochondrial protein-coding loci (table 2). In addition, we sequenced three rDNA fragments: a portion of the nuclear small rDNA subunit (SSU), the two internal transcribed spacers (ITS1 and ITS2), and a mitochondrial fragment encompassing the 5.8S and large rDNA subunit (LSU) (table 2).

We used the *P. tetraurelia* macronuclear (http://paramecium.cgm.cnrs-gif.fr/) (Aury et al. 2006) and mitochondrial (Pritchard et al. 1990) genome sequences and the freely available online software "Primer3" (Rozen and Skaletsky 2000) to design DNA oligo pairs to coding regions for polymerase chain reaction (PCR) amplifications (table 2 and supplementary table 3 in Supplementary Material online). More than a single pair of oligos was often required to amplify a locus across most or all strains surveyed. The nuclear loci studied were random single-copy genes in the *P. tetraurelia* macronuclear genome. We amplified the gene regions using the following PCR conditions: 3 min at 94 °C, 40 cycles of 40 s at 94 °C, 40 s at 55 °C, and 1 min at 72 °C. Amplification products were subsequently checked on 1% agarose gels, and reactions giving unique products were sequenced in both directions. We verified the quality of the sequenced gene fragments both automatically, using the software Codon Code Aligner, and visually. We used ClustalW (Thompson et al. 1994), as implemented in the software Bioedit (Hall 1999), to align the consensus sequences.

Data Analysis

Sequences were exported to the software MEGA 4.0 (Nei and Kumar 2000) for data analysis. For each protein-coding locus, we measured intraspecies and interspecies diversity at both silent sites and replacement sites, following the Kumar method (Nei and Kumar 2000). Diversity values were multiplied by $n/(n-1)$, where n is the appropriate number of strains for each species, to correct for small sample size, and the average values were weighted by the number of base pairs sequenced for each locus. The Maximum Composite Likelihood model implemented in MEGA 4.0 and bootstrap statistics (1,000 replicates) were used to build Neighbor-Joining (NJ) trees using all the sites of the regions examined. We also measured diversity in single or concatenated intronic regions (the latter when genes contained more than one intron) contained in each of the surveyed loci, after removing the conserved GT–AG ending sites. The package GENECONV (Sawyer 1999) was employed to estimate both the number and the location of possible recombinant alleles.

All the sequences generated in this study were submitted to GenBank, accession numbers FJ002921–FJ004151.

Results and Discussion

High Levels of Silent Genetic Variation within Reproductively Compatible *P. aurelia* Strains

Levels of intraspecific diversity at nuclear synonymous sites (π_s) in the *P. aurelia* species complex average 0.1493 (SEM = 0.0382) (table 3 and Supplementary Material online). Diversity at silent sites is ~13-fold higher than at nonsynonymous sites (π_a = 0.0117 [SEM = 0.0031]) but is comparable with that calculated for intronic

Table 1
List of *Paramecium* Strains/Species Surveyed

Strain	Species	Origin
	Paramecium primaurelia	
16		Woodstock, MD, United States
168		Japan
33		Unknown
60		United States
61		Unknown
90		Pennsylvania, United States
AZ15-8		Astrakhan Nature Reserve, Russia
AZ9-3		Astrakhan Nature Reserve, Russia
Ir4-1		The Irtysh river, Omsk, Russia
KK2-7		Kaliningrad region, Russia
TB10		Rejkjavik, Iceland
V7-6		Volgograd region, Russia
223-7		The Volga river, Tver' region, Russia
112-1		Vologda Region, Russia
227		Unknown
	Paramecium biaurelia	
1038s		Unknown
114		Bloomington, IN, United States (uncertain)
AB7-16		Boston, United States
AZ25-2		Astrakhan Nature Reserve, Russia
F2		Unknown
GA62-1		Altai Mountains, Russia
Kr133-2		Krasnoyarsk, Russia
P2		Unknown
PK2		Kraków, Poland
RR2-1		Saint-Petersburg region, Russia
TB15		Teneriffa, Canary Islands, Spain
UM2		Unknown
V1-4		Volgograd region, Russia
W7		Unknown
	Paramecium triaurelia	
152alpha		New Haven, United States
EP_2/3		Natural reserve Complex Volga-Ahtuba, Russia
kr128		Krasnoyarsk, Russia
Kr149-1		Krasnoyarsk, Russia
TB17		Greece
TB18		Santuhalm near Deva, Romania
V10-7		Volgograd region, Russia
	Paramecium tetraurelia	
172		Perú
298s		Unknown
32		Unknown
51s		Indiana, United States
d4-2		—
nr7-1		Novorossiysk, Russia
PF15		Paris, France
S		Sydney, Australia
ST		Tatry, Slovakia
BGD19		Unknown
	Paramecium pentaurelia	
AZ6-24		Astrakhan Nature Reserve, Russia
GA1-9		Altai Mountains, Russia
Nr1-10		Novorossiysk, Russia
RA81-8		Altai Forelands, Russia
V2-7		Volgograd region, Russia
	Paramecium sexaurelia	
159		Unknown
AZ8-4		Astrakhan Nature Reserve, Russia
CB16-2		Beijing, China
	Paramecium septaurelia	
AZ5-2		Astrakhan Nature Reserve, Russia
AZ8-3		Astrakhan Nature Reserve, Russia
GFg-1		Freiburg, Germany
V5-13		Volgograd region, Russia
	Paramecium octaurelia	
138		Florida, United States
565s		Unknown
IEE		Ein Efek, Israel
K8		Unknown
K9		Unknown
	Paramecium novaurelia	
AB8-22		Boston, United States
Iz16		Izmail, Ukraine
V9-6		Volgograd region, Russia
VL4-8		Vladimir, Russia
	Paramecium decaurelia	
223		Florida, United States
93-21		Yaroslavl region, Russia
GA1-12		Altai Mountains, Russia
JN		Nara, Japan
	Paramecium undecaurelia	
219		United States
	Paramecium dodecaurelia	
246		Mississippi, United States
256-2		Yaroslavl region, Russia
UV1-3		Vinnitsa, Ukraine
	Paramecium tredecaurelia	
209		Paris, France
	Paramecium quadecaurelia	
328		Australia
AN1		Namibia
	Paramecium sonneborni	Texas, United States
	Paramecium multimicronucleatum	Krasnoyarsk, Russia

sites ($\pi_i = 0.0944$ [SEM = 0.0246]). No intron presence or absence polymorphisms were detected across species, and levels of diversity in coding and intronic regions show significant positive correlations when estimates for single loci are examined (π_a vs. π_i, Spearman's rho = 0.643, $P < 0.01$; π_s vs. π_i, Spearman's rho = 0.540, $P < 0.01$). In the mitochondria, loci also showed a substantial degree of variation ($\pi_s = 0.3257$ [SEM = 0.0674] and $\pi_a = 0.0173$ [SEM = 0.0057]), with silent sites having on average ~2.2 times the nucleotide diversity of nuclear genes (table 4 and Supplementary Material online). Levels of net interspecific divergence (i.e., in excess of within-species divergence) at silent sites average 0.3307 (SEM = 0.0209) for the nuclear loci (supplementary table 1 in Supplementary Material online) and 0.5214 (SEM = 0.0184) for mitochondrial loci (supplementary table 2 in Supplementary Material online). At amino acid replacement sites, divergence drops to 0.0239 (SEM = 0.0012) and 0.0175 (SEM = 0.0175) for nuclear and mitochondrial loci, respectively. The ratio of mitochondrial to nuclear divergence at (neutrally evolving) silent sites reflects the ratio of the mutation rates in the two cellular compartments and averages 4.0223 (SEM = 2.6470) when all species are examined but decreases to 1.2504 (SEM = 0.2451) after the removal of *Paramecium sexaurelia*, a species previously reported to show unusually divergent genotypes across different geographic locations and where F2 generations of interstrain crosses show low rates of survival (Stoeck et al. 1998; Przybos, Rautian, et al. 2007).

Table 2
***Paramecium* Genes Surveyed and Primers Used for PCR Amplification**

Nuclear Loci	Molecular Function	PCR Primers (5′–3′)
	DNA mismatch repair	gagagacatctaaactgtgcgtttt
GSPATG00000151001	system Msh2	tggcattttgatccatcttc
	Calcium binding	ttccctgaccgaataggattt
GSPATG00000223001	protein	ggccataagcatccaagatt
		gcagcagctgttggtcataa
GSPATG00000363001	tRNA methyltransferase	tccagtgtgtccataagcttaattt
	GTP-binding protein	tgaacatgatgcaactattggag
GSPATG00000454001	RAB2 homolog	gtggatttgcaggtttgcta
	GrpE protein	tcccttgatgaatctcacga
GSPATG00000546001	homolog	tgttcctggttctttttcagg
	3-Oxoadipate	ttggcaggatggtctttagg
GSPATG00027725001	enol-lactonase	tggcatatgaccaacacctg
	Serine	ttacatcattagtgtttgtcaggtg
GSPATG00028095001	carboxypeptidase	cctgcataactctccccaaa
		agttggcgataatggaggac
GSPATG00028110001	Hypothetical protein	aaccatcagctccatcaacc
	Ubiquitin-like 1 activating	tggcattggataatgcagaa
GSPATG00022332001	enzyme E1B	gtagaggcaatggcgtgaat
	Hypothetical	aagccacctccagattttca
GSPATG00020758001	protein	tttcccagatcgaacctttg
Mitochondrial loci		
		tgggtagtatgacattgcttttc
NADH 3		ggccagcgagttctggac
		gccctctcagagctcaacat
ORF189-2 & yegR		acaccgggtggacgtaga
		ccaaaattggggttacgatg
COX I		attgcgtttttgaggacgac
		gcatagaagtagtcatcagcaacc
COX II		ttttgtgaggttgtatttccaca
rDNA loci		
		gcaagtctggtgccagcagcc
SSU[a] (nuclear)		cttccgtcaattcctttaag
ITS–5.8S rDNA–ITS[a]		tcctccgcttattgatatgc
(nuclear)		ggaagtaaaagtcgtaacaagg
5.8S-LSU		ttggttgagggcgtaaatct
(mitochondrial)		ctaccccgcacaaaagaaaa

[a] Primer sequences match NS3, NS4, ITS4, and ITS5 in White et al. (1990).

Low Levels of Genetic Variation within Monophyletic Taxa

The analysis of the gene genealogies across all nuclear and mitochondrial loci revealed that 13 "discordant" strains consistently group with species that are different from that to which they were originally assigned by mating tests (table 5 and supplementary fig. 1 in Supplementary Material online). We verified this unexpected finding by further studying the macronuclear SSU and ITSs and the mitochondrial 5.8 and LSU sequences (Supplementary Material online) and reperforming mating experiments on this set of suspect strains. We conclude that although these strains contain a multilocus genetic profile that is highly similar to that of a given *P. aurelia* species, they preferentially mate with strains belonging to a different species within the species complex. As it is legitimate to question whether or not these discordant strains ought to be included in the calculation of intraspecific genetic variability, we also calculated levels of polymorphism after their exclusion. We found that this leads to a considerable decrease in within-species genetic variability, with an average π_s of

Table 3
Intraspecific Genetic Diversities at Nonsilent Sites (π_a), Silent Sites (π_s), and Intronic Sites (π_i) with Corresponding Standard Errors (SE), Estimated for All Strains across 10 Nuclear Loci

	All Strains					Only Nondiscordant Strains				
Species	Average No. of Strains	π_a (SE)	π_s (SE)	π_a/π_s	π_i (SE)	Average No. of Strains	π_a (SE)	π_s (SE)	π_a/π_s	π_i (SE)
Paramecium primaurelia	12	0.0011 (0.0003)	0.0248 (0.0058)	0.0444	0.0249 (0.0142)	11	0.0010 (0.0003)	0.0228 (0.0055)	0.0439	0.0240 (0.0144)
Paramecium biaurelia	10	0.0131 (0.0047)	0.1683 (0.0695)	0.0778	0.1302 (0.1010)	10	0.0004 (0.0003)	0.0038 (0.0015)	0.1053	0.0075 (0.0095)
Paramecium triaurelia	7	0.0029 (0.0007)	0.0344 (0.0069)	0.0843	0.0167 (0.0083)	6	0.0003 (0.0003)	0.0066 (0.0025)	0.0455	0.0010 (0.0013)
Paramecium tetraurelia	10	0.0076 (0.0013)	0.0908 (0.0116)	0.0837	0.0446 (0.0130)	9	0.0007 (0.0004)	0.0079 (0.0028)	0.0886	0.0040 (0.0034)
Paramecium pentaurelia	4	0.0108 (0.0056)	0.1208 (0.0555)	0.0891	0.0416 (0.0351)	4	0.0009 (0.0012)	0.0017 (0.0021)	0.5294	0.0036 (0.0046)
Paramecium sexaurelia	3	0.0003 (0.0006)	0.0038 (0.0079)	0.0789	0.0058 (0.0166)	3	0.0003 (0.0006)	0.0038 (0.0079)	0.0789	0.0057 (0.0163)
Paramecium septaurelia	4	0.0197 (0.0040)	0.2748 (0.0470)	0.0717	0.1514 (0.0505)	3	0.0004 (0.0003)	0.0034 (0.0024)	0.1176	0.0053 (0.0060)
Paramecium octaurelia	5	0.0057 (0.0014)	0.0743 (0.0081)	0.0767	0.0498 (0.0211)	4	0.0003 (0.0002)	0.0036 (0.0016)	0.0833	0.0085 (0.0069)
Paramecium novaurelia	3	0.0166 (0.0037)	0.2561 (0.0336)	0.0648	0.2012 (0.1128)	2	0	0.0091 (0.0130)	0.0000	0
Paramecium decaurelia	3	0.0294 (0.0094)	0.2940 (0.0711)	0.1000	0.2451 (0.1141)	2	0.0023 (0.0020)	0.0104 (0.0068)	0.2212	0.0109 (0.0184)
Paramecium dodecaurealia	3	0.0310 (0.0064)	0.4154 (0.0740)	0.0746	0.1971 (0.0849)	2	0.0008 (0.0009)	0.0078 (0.0061)	0.1026	0.0109 (0.0158)
Paramecium quadecaurelia 1	2	0.0019 (0.0040)	0.0345 (0.0115)	0.0551	0.0241 (0.0811)	2	0.0019 (0.0040)	0.0345 (0.0115)	0.0551	0.0245 (0.0823)
Average		0.0117 (0.0031)	0.1493 (0.0382)	0.0751 (0.0043)	0.0944 (0.0246)		0.0008 (0.0002)	0.0096 (0.0028)	0.1226 (0.0401)	0.0088 (0.0023)

Table 4
Intraspecific Genetic Diversities at Nonsilent (π_a) and Silent Sites (π_s), with Corresponding SE, Estimated for All Strains across Five Mitochondrial Loci

	All Strains				Only Nondiscordant Strains			
Species	Average No. of Strains	π_a (SE)	π_s (SE)	π_a/π_s	Average No of Strains	π_a (SE)	π_s (SE)	π_a/π_s
Paramecium primaurelia	15	0.0022 (0.0005)	0.1074 (0.0279)	0.0205	14	0.0020 (0.0005)	0.0986 (0.0257)	0.0203
Paramecium biaurelia	13	0.0123 (0.0057)	0.2207 (0.0417)	0.0557	11	0.0002 (0.0001)	0.0043 (0.0013)	0.0465
Paramecium triaurelia	7	0.0039 (0.0016)	0.2124 (0.0378)	0.0184	6	0	0	
Paramecium tetraurelia	10	0.0100 (0.0052)	0.1255 (0.0309)	0.0797	9	0.0006 (0.0006)	0.0157 (0.0046)	0.0382
Paramecium pentaurelia	5	0.0123 (0.0075)	0.1688 (0.0833)	0.0729	4	0	0	
Paramecium sexaurelia	2	0	0.1563 (0.1260)	0	2	0	0.1563 (0.1260)	0
Paramecium septaurelia	4	0.0225 (0.0096)	0.6567 (0.1080)	0.0343	3	0	0.0052 (0.0041)	0
Paramecium octaurelia	5	0.0112 (0.0079)	0.2904 (0.0477)	0.0386	4	0	0.0115 (0.0040)	0
Paramecium novaurelia[a]	7	0.0308 (0.0164)	0.5744 (0.0848)	0.0536	5	0.0016 (0.0018)	0.1100 (0.1118)	0.0145
Paramecium decaurelia	3	0.0695 (0.0530)	0.7804 (0.1740)	0.0891	1	0	0	
Paramecium dodecaurealia	2	0.0314 (0.0307)	0.4729 (0.2252)	0.0664	2	0	0.0080 (0.0129)	0
Paramecium quadecaurelia	2	0.0016 (0.0080)	0.1420 (0.0599)	0.0113	1.7	0.0016 (0.0080)	0.1420 (0.0599)	0.0113
Average		0.0173 (0.0057)	0.3257 (0.0674)	0.0450 (0.0083)		0.0005 (0.0002)	0.0460 (0.0177)	0.0145 (0.0059)

[a] COX I sequences of 21 *P. novaurelia* strains (Przybos, Tarcz, and Skoblo 2007; Tarcz S., unpublished data) were retrieved from NCBI and analyzed along with the strains surveyed in this study.

0.0096 (SEM = 0.0028) and 0.0460 (SEM = 0.0177) in the nucleus and in the mitochondrion, respectively (tables 3 and 4 and Supplementary Material online). The net interspecific average divergence at silent sites increases to 0.4034 (SEM = 0.0198) and to 0.6717 (SEM = 0.0203) at nuclear and mitochondrial loci, respectively (supplementary tables 1 and 2 in Supplementary Material online), but it remains similar at nonsynonymous sites (0.0293 [SEM = 0.0012] and 0.0269 [SEM = 0.0016]). In addition, levels of diversity at intronic sites and at amino acid replacement sites are no longer significantly correlated (π_a vs. π_i, Spearman's rho = 0.032, $P = 0.772$; π_s vs. π_i, Spearman's rho = 0.292, $P < 0.01$). The ratio of mitochondrial-to-nuclear divergence at silent sites averages 3.8049 (SEM = 2.5900) and decreases to 1.1597 (SEM = 0.4212) after the removal of *P. sexaurelia*.

Discordant and Phylogenetically Distinct *P. aurelia* Lines

The existence of distinct and reproductively isolated species (or "varieties" or "syngens" as originally reported [Sonneborn 1937, 1957]) within the *P. aurelia* complex has been convincingly described (Sonneborn 1975) and correlated with molecular makers by many workers (e.g., Stoeck and Schmidt 1998). Yet, as reported above, our polymorphism survey reveals that a number of *P. aurelia* species are not monophyletic. Allele sharing between *P. aurelia* species has been detected by previous studies (Coleman 2005; Hori et al. 2006; Snoke et al. 2006; Tarcz et al. 2006; Barth et al. 2008), but little significance has typically been given to this finding. The limited number of species, strains, or loci in these earlier investigations leads to a number of nonmutually exclusive interpretations of the data, including 1) the occurrence of recent speciation events; 2) a large N_e, whereby strains still contain segregating alleles that predate species divergence (i.e., transspecies polymorphisms); and 3) incomplete reproductive isolation.

Our analysis crucially shows that allele sharing between strains with different mating behavior is not sporadic but rather extends across all macronuclear and mitochondrial loci surveyed. Also, allele sharing does not involve the whole set of strains for a species but often concerns all nondiscordant strains for one species and only a single strain for a second species. Finally, distinct *P. aurelia* species tend to form separate clades—after the exclusion of the discordant strains—with very short internal branches.

Consistent with a number of previous findings (e.g., Haggard 1974), our observations seem to rule out the possibility of recurrent gene flow among *P. aurelia* species, as the genetic profile of the surveyed strains is never discontinuous, that is, we do not find evidence of introgressive hybridization or episodes of cross-species recombination. The hypothesis of transspecies polymorphisms due to a large N_e in *Paramecium* also seems unlikely, given the modest level of intraspecific genetic variability ($= 4N_e\mu$) calculated after the removal of the discordant strains. Finally, the hypothesis of interspecific allele sharing due to recent speciation could well explain the intimate phylogenetic relationship between *Paramecium septaurelia* and-*Paramecium octaurelia*, whose strains tend to consistently group together across all gene genealogies but hardly

Table 5
List of Discordant Strains

Strain	Species According to Mating Tests	Species According to mt rRNAs and Nuclear ITS	Species According to Multiple Nuclear and mt Loci
61	*Paramecium primaurelia*	*P. primaurelia/Paramecium pentaurelia*	*P. pentaurelia*
PK2	*Paramecium biaurelia*	*Paramecium septaurelia/Paramecium octaurelia*	*P. septaurelia/P. octaurelia*
GA62-1	*P. biaurelia*	*Paramecium triaurelia*	*P. triaurelia*
TB18	*P. triaurelia*	*P. primaurelia*	*P. primaurelia*
ST	*Paramecium tetraurelia*	*P. primaurelia*	*P. primaurelia*
GA1-9	*P. pentaurelia*	Unclear	Unclear
AZ8-3	*P. septaurelia*	*P. primaurelia*	*P. primaurelia*
IEE	*P. octaurelia*	*P. tetraurelia*	*P. tetraurelia*
AB8-22	*Paramecium novaurelia*	*P. primaurelia*	*P. primaurelia*
V9-6	*P. novaurelia*	*P. primaurelia*	*P. primaurelia*
JN	*Paramecium decaurelia*	*P. primaurelia*	*P. primaurelia*
223	*P. decaurelia*	Unclear	Unclear
246	*Paramecium dodecaurelia*	*P. septaurelia/P. octaurelia*	*P. septaurelia/P. octaurelia*

explains the allele sharing between multiple strains of one species and an isolated strain of a second species.

A more parsimonious explanation for our observations is that the locus (loci) that is responsible for mating behavior in *Paramecium*—whose identity is still unknown—has undergone changes in the discordant strains that switch their mating preference. These changes would only alter the mating preference of these strains, which would otherwise normally react with genetically similar conspecific strains, leaving intact the overall genetic profile. It is notable that all 13 discordant strains show apparently normal viability in their matings to genetically distinct species (supplementary table 4 in Supplementary Material online).

Although not reported for any other ciliate or, to our knowledge, eukaryotic species, the incongruence we observe between mating behavior and genetic profile may not be limited to the *P. aurelia* species complex, and "species switches" might at least extend to other ciliates with cryptic species (e.g., *P. multimicronucleatum*, *Tetrahymena pyriformis*). Under this hypothesis, it would be informative to study the genetic profile of species other than *P. aurelia* and examine whether conspecific strains (following successful mating reactions) carry alleles that tend to be highly divergent at multiple loci while similar to alleles of a related species, an analysis that is currently hampered in ciliates by the dearth of available molecular markers and/or large-scale population-genetic surveys. If the discordance between mating preferences and genetic profile is confirmed to be caused by an alteration of the mating locus (loci), then mating locus switches could be considered a novel mechanism of speciation that, for its effects, could be comparable with other known mechanisms, such as, for example, cytoplasmic incompatibility (Bordenstein et al. 2001).

If the above hypothesis is correct, the genetic and ecological effects of two distinct genetic profiles, combined in one nucleus, remains an issue to be addressed in future investigations. Future studies will determine whether the F1 generation obtained from such a cross is fertile (they do survive autogamy) and less or more vigorous than the parental strains. In addition, it is currently not known if the pronounced interspecific sequence divergence would disfavor homologous recombination—thus gene flow and introgression—as has been observed for other organisms (Datta et al. 1997; Lukacsovich and Waldman 1999; Opperman et al. 2004), and/or if the set of genes of one of the two conjugating strains might be invariably discarded (gynogenesis or hybridogenesis) (Macgregor and Uzzell 1964; Schultz 1969; Mantovani and Scali 1992; Ragghianti et al. 2007). The latter event emerges as a possibility from earlier studies, where a cross between two *P. aurelia* species, *P. tetraurelia*, and *P. octaurelia* was reported to produce survivor clones that are not true hybrids and behave as one or the other species (the true F1 hybrids were sterile) (Levine 1953; Haggard 1974). Also, unless gynogenesis or hybridogenesis occurs, these discordant species would lead to introgression (which we have not observed) if they generated fertile progeny.

Finally, it is worth noting that even after the exclusion of discordant strains, *P. aurelia* species do not always form monophyletic groups. This can be seen, for example, in the case of *P. primaurelia,* where strain 227 branches off alone, making this taxon paraphyletic. Levels of sequence divergence between this *P. primaurelia* strain and all other conspecific strains are higher, both for macronuclear and mitochondrial loci, than the divergence estimated between *P. primaurelia* and *P. pentaurelia*. Strain 227 also diverges considerably from its other conspecific strains at rRNA loci, where it tends to cluster with *Paramecium triaurelia* strains (data not shown). It is tempting to conjecture that this *P. primaurelia* strain—of unknown geographic origin—may represent an example of incipient speciation in the *P. aurelia* complex.

Identification of Paralogous Loci

Paramecium tetraurelia has undergone at least three whole genome duplications (WGDs), the most recent of which has been suggested to coincide with the burst of speciation events that led to the emergence of the *P. aurelia* species complex (Aury et al. 2006). An erroneous sampling of WGD-associated paralogs in our study could have inflated the true genetic variation at macronuclear loci (but not at mitochondrial loci), affecting both intraspecific and interspecific estimates of diversity. The strain we used to design PCR primers (d4-2, *P. tetraurelia*) does not contain macronuclear paralogs for the loci we have sequenced. Nevertheless, we cannot rule out the possibility that either *P. tetraurelia* strains other than d4-2 or the remaining *P. aurelia* species are not

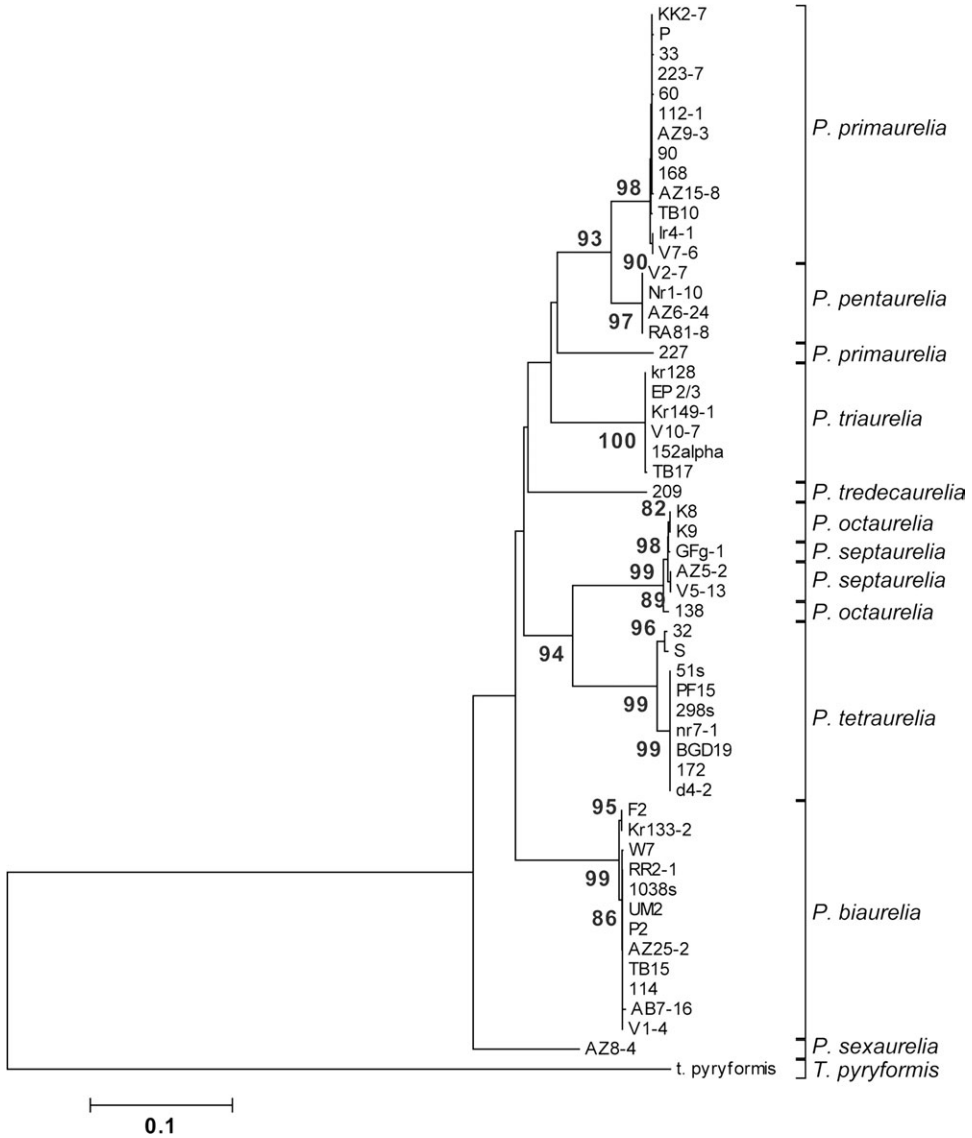

FIG. 1.—NJ tree built by concatenating three mitochondrial protein-coding loci, NADH dehydrogenase subunit 3 (NADH3), cytochrome c oxidase subunit I (COXI) and II (COXII) (1,057 bp). Only nondiscordant strains are included in the analysis. Bootstrap values equal to or higher than 75 are shown. *Tetrahymena pyriformis* used as outgroup.

paralog-free. Also, divergent resolution of paralogous genes following polyploidization and speciation (Lynch and Conery 2000; Lynch and Force 2000b) might facilitate such an erroneous sampling, but whether or not such a model applies to the *P. aurelia* complex is yet unknown.

We found two cases in which we PCR amplified paralogous loci and we removed these cases from our analyses. At locus GSPATG00022332001—the only nuclear intronless gene we surveyed—DNA sequencing revealed the existence of multiple double peaks in single-band PCR products obtained for the *P. biaurelia* species. We scored double peaks for all but five of the *P. biaurelia* strains, and we only included these five strains in the final analyses. As single-gene duplications appear to be uncommon in *Paramecium* (Aury et al. 2006) and no other double-peaked DNA sequences were observed for other species at this locus, this additional gene copy in *P. biaurelia* may be in the process of being lost from the genome. Alternatively, this gene could be a young duplicate that arose in *P. biaurelia*. GSPATG00022332001 is also the only gene for which we could not obtain any PCR product in *P. primaurelia* (notably, strain 227 was the only one producing a band in *P. primaurelia*). Overall, our observations suggest that GSPATG00022332001, which codes for a protein that is involved in the ubiquitin cycle, is an evolutionarily labile gene in *P. aurelia*.

We detected the second case of paralogy after inspecting possible inconsistencies among gene genealogies. In particular, a phylogenetic tree constructed with five concatenated mitochondrial protein-coding loci strongly supports a close relatedness between *P. primaurelia* and *P. pentaurelia* and this relationship was supported by a tree constructed from three concatenated nuclear loci (fig. 2) and is observed for all nuclear loci but one, namely, GSPATG00028095001. At this locus, strains belonging to *P. pentaurelia* still cluster

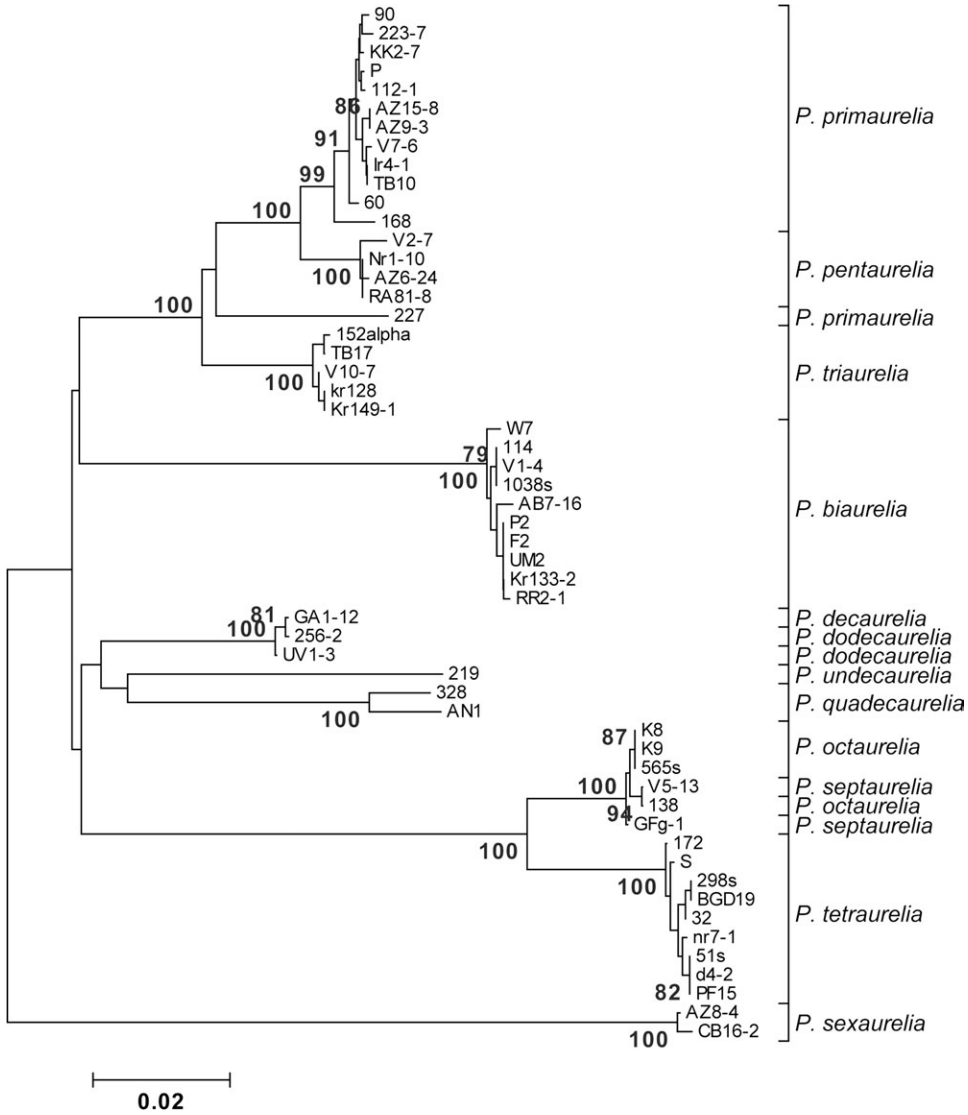

FIG. 2.—Unrooted NJ tree built by concatenating the nuclear loci GSPATG00000363001, GSPATG00027725001, and GSPATG00000546001 (1,233 bp). Only nondiscordant strains are included in the analysis. Bootstrap values equal to or higher than 75 are shown.

together (they share identical sequences) but are isolated from all other species in the tree. The level of divergence at silent sites between *P. primaurelia* and *P. pentaurelia* is as high as 80% at GSPATG00028095001 but drops to an across-loci average of <7% when this locus is removed from the analysis. We observed no PCR band at this locus for *P. pentaurelia* when we used an additional PCR primer that is external to the former pair. Although further experiments are needed to unequivocally show that the orthologous locus to the *P. tetraurelia*'s GSPATG00028095001 is not present in *P. pentaurelia*, our observations are consistent with a differential retention of paralogous loci between the two *P. aurelia* species at this locus.

Molecular Phylogeny, Mating Features, and Speciation in the *P. aurelia* Species Complex

We obtained well-supported phylogenetic relationships only for the *P. aurelia* species that have diverged most recently (figs. 1 and 2). In these cases, the groupings broadly reflect the tendency that species have to cross-react during mating experiments (Sonneborn 1975). In particular, *P. primaurelia* and *P. pentaurelia* as well as *P. tetraurelia* and *P. octaurelia*, which pair closely in all gene genealogies, are known to cross-react (the latter pair more strongly than the former) and conjugate, without however giving rise to viable or fertile hybrids (Sonneborn 1974; Beale and Preer 2008). For either combination of mating types, 40% of cells will form true conjugants in a mixed culture of *P. primaurelia* and *P. pentaurelia* species, and the frequency rises to 90% when mating type E of *P. tetraurelia* is mixed with mating type O of *P. octaurelia*. *Paramecium primaurelia* mating type E also weakly reacts with type O of *P. triaurelia*. One would also expect a tendency for a cross-reaction between the closely related *P. septaurelia* and *P. octaurelia* or between *P. septaurelia* and *P. tetraurelia*. In fact, no such reactions are reported in the literature, but both *P. septaurelia* and *P. octaurelia* can—depending

on their mating type—weakly react with *P. primaurelia* (Sonneborn 1975). It is worth noting that the genetic profile of most of the detected discordant strains corresponds to that of *P. primaurelia*, which can cross-react with more species within the *P. aurelia* complex than any other, an observation that remains unexplained.

We consistently find that the internal nodes of the molecular phylogeny of the *P. aurelia* species complex are only supported by low bootstrap values. A similar poor phylogenetic resolution has also been observed by previous studies (e.g., Coleman 2005; Hori et al. 2006; Barth et al. 2008). The virtual lack of phylogenetic resolution does not depend on the type of marker examined, as neither nuclear nor mitochondrial, and neither coding nor noncoding markers appear to be sufficiently informative. A plausible explanation for the latter observation is that speciation in *P. aurelia* may not have always proceeded as a branching process, followed by genetic isolation, but may have been the result of an initial burst of speciation events, perhaps accompanied by extensive hybridization. As the average divergence at silent sites between recent paralogs in *P. tetraurelia* is equal to 2.17 (Aury et al. 2006)—>2 times higher than the maximum species divergence per silent site, we estimate in this study (0.9804, between *P. tetraurelia* and *P. sonneborni*)—our observations indicate that the *P. aurelia* radiation postdates the most recent WGD and that speciation events may be associated with differential retain of gene copies across isolated populations (Lynch and Force 2000a).

Species Diversity and Geographic Distribution

Although our results do not suggest that huge global populations of the *P. aurelia* species exist, we find the same alleles in strains isolated from multiple continents, consistent with regular (or recent) global dispersal of most of the species. These observations intriguingly suggest that *P. aurelia* species might both have a limited (effective) population size but also be cosmopolitan. In an attempt to provide an explanation for this counterintuitive conclusion, we suggest that individual *Paramecia* inhabiting relatively few and durable habitats, which provide the long-term persistence of the species, would be dispersed to multiple geographically separated (and often ephemeral) environments, generating clonal populations whose life span is limited and depends on the features of the newly colonized habitats. The proposed scenario suggests a possibly cyclical and rapid transition from limited geographic distribution to an effectively, temporary, and cosmopolitan distribution of *Paramecium* species, a dynamic that is similar to that proposed by source–sink or source–pseudosink ecological models (Pulliam 1988; Watkinson and Sutherland 1995) and that may lead to a decrease in N_e (Whitlock and Barton 1997).

Additional and nonmutually exclusive explanations for allele sharing between strains inhabiting widely separated geographical areas would be strong selection against mutations at silent sites and/or an exceptionally low mutation rate in *Paramecium*. These factors would influence the accumulations of nucleotide changes after geographical separation of conspecific strains, and the extensive interspecific sequence divergence we observe would indicate fairly old events of speciation in the *P. aurelia* species complex. Whereas some support for selection against mutation at silent sites has been already provided (Salim et al. 2008), our ongoing study on the rate of spontaneous mutations in the genome of *P. tetraurelia* will determine the likelihood of the low-mutation-rate hypothesis.

Conclusions

The average estimate of neutral genetic variability (π_s) in a species is expected to reflect the global effective population size of the species (Lynch 2006a). We found that within the *P. aurelia* species complex, estimates of π_s are considerably higher when strains are grouped according to their reproductive compatibility but decrease greatly when the grouping criterion is the similarity of the strains' genetic profiles. This effect is due to the exclusion from the analysis of a small set of strains with apparently aberrant mating behavior.

Overall, our study reveals that the average level of nuclear genetic diversity in the free-living ciliate *Paramecium* may not be as exceptionally high as previously reported (Snoke et al. 2006) but is, on average, not as low as the degree of neutral genetic diversity reported for another ciliate, *Tetrahymena thermophila* (~0.003) (Katz et al. 2006). Depending on the criteria that we use to tell true species apart, values of macronuclear π_s within *P. aurelia* may be either comparable with or higher than the average diversity reported for a set of mostly pathogenic unicellular eukaryotes (0.0573 [SD = 0.0777]) (Lynch 2006b) and for the unicellular green alga *Chlamydomonas reinhardtii* (0.032) (Smith and Lee 2008). Similar conclusions apply to the average genetic diversity observed at mitochondrial loci, where the π_s within the *P. aurelia* complex could be one of the highest values ever reported, together with *P. multimicronucleatum* (π_s = 0.308), a close outgroup of the *P. aurelia* complex (Snoke et al. 2006) or comparable with the π_s of five other unicellular (mostly pathogenic) eukaryotes (0.011 [SEM = 0.004]) and *C. reinhardtii* (0.0085) (Smith and Lee 2008), after the removal of discordant strains.

Although this study cannot provide a definitive answer to the amount of genetic variability in *Paramecium*, our and previous observations (Katz et al. 2006) seem to suggest that levels of genetic diversity in free-living species may be relatively low and comparable to those of pathogenic taxa. Our findings support the proposal that speciation in *Paramecium* may not have always followed a cladistic process and highlight how the species problem in *P. aurelia*, an issue extensively addressed by Sonneborn during the last century (Sonneborn 1937, 1957, 1974, 1975), may prove even more complex than anticipated.

Supplementary Material

Supplementary tables 1-4, supplementary figure 1, and supplementary material are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank T.G. Doak and C.L. McGrath for insightful comments and suggestions for the improvement of the

manuscript. We are grateful to E. Choi and V. Cloud for the maintenance of the *Paramecium* strains and to J.R. Preer Jr for his patience and for sharing with us his own experience and knowledge of *Paramecium*. We thank the Russian Ministry of Education and Science (project RNP 2.2.3.1.4148) for supporting the maintenance of Collection of *Paramecium* strains of the Laboratory of Protozoan Karyology, St Petersburg State University, from which originate most of the strains examined in this study. This work was supported by the MetaCyte funding from the Lilly Foundation to Indiana University.

Literature Cited

Aufderheide KJ, Daggett PM, Nerad TA. 1983. *Paramecium-sonneborni* N-Sp, a new member of the *Paramecium-aurelia* species-complex. J Protozool. 30:128–131.

Aury JM, Jaillon O, Duret L, et al. (42 co-authors). 2006. Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. Nature. 444:171–178.

Barth D, Krenek S, Fokin SI, Berendonk TU. 2006. Intraspecific genetic variation in *Paramecium* revealed by mitochondrial cytochrome C oxidase I sequences. J Eukaryot Microbiol. 53:20–25.

Barth D, Przybos E, Fokin SI, Schlegel M, Berendonk TU. 2008. Cytochrome *b* sequence data suggest rapid speciation within the *Paramecium aurelia* species complex. Mol Phylogenet Evol. 49:669–673.

Beale G, Preer JR. 2008. Paramecium: genetics and epigenetics. Boca Raton (FL): CRC Press.

Bordenstein SR, O'Hara FP, Werren JH. 2001. *Wolbachia*-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. Nature. 409:707–710.

Chen TT. 1956. Varieties and mating types in *Paramecium Bursaria*. 2. Variety and mating types found in China. J Exp Zool. 132:255–268.

Coleman AW. 2005. *Paramecium aurelia* revisited. J Eukaryot Microbiol. 52:68–77.

Datta A, Hendrix M, Lipsitch M, Jinks-Robertson S. 1997. Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. Proc Natl Acad Sci USA. 94:9757–9762.

Daubin V, Moran NA. 2004. Comment on "The origins of genome complexity". Science. 306:978; author reply 978.

Fenchel T, Esteban GF, Finlay BJ. 1997. Local versus global diversity of microorganisms: cryptic diversity of ciliated protozoa. Oikos. 80:220–225.

Finlay BJ, Fenchel T. 1999. Divergent perspectives on protist species richness. Protist. 150:229–233.

Foissner W. 1999. Protist diversity: estimates of the near-imponderable. Protist. 150:363–368.

Foissner W. 2006. Biogeography and dispersal of micro-organisms: a review emphasizing protists. Acta Protozool. 45:111–136.

Gerber CA, Lopez AB, Shook SJ, Doerder FP. 2002. Polymorphism and selection at the SerH immobilization antigen locus in natural populations of *Tetrahymena thermophila*. Genetics. 160:1469–1479.

Grigg ME, Bonnefoy S, Hehl AB, Suzuki Y, Boothroyd JC. 2001. Success and virulence in *Toxoplasma* as the result of sexual recombination between two distinct ancestries. Science. 294:161–165.

Haggard BW. 1974. Interspecies crosses in *Paramecium aurelia* (syngen 4 by syngen 8). J Protozool. 21:152–159.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser. 41:95–98.

Hori M, Tomikawa I, Przybos E, Fujishima M. 2006. Comparison of the evolutionary distances among syngens and sibling species of *Paramecium*. Mol Phylogenet Evol. 38:697–704.

Katz LA, Snoeyenbos-West O, Doerder FP. 2006. Patterns of protein evolution in *Tetrahymena thermophila*: implications for estimates of effective population size. Mol Biol Evol. 23:608–614.

Levine M. 1953. The interactions of nucleus and cytoplasm in the isolation and evolution of species of *Paramecium*. Evolution. 7:366–385.

Lukacsovich T, Waldman AS. 1999. Suppression of intra-chromosomal gene conversion in mammalian cells by small degrees of sequence divergence. Genetics. 151:1559–1568.

Lynch M. 2006a. Streamlining and simplification of microbial genome architecture. Annu Rev Microbiol. 60:327–349.

Lynch M. 2006b. The origins of eukaryotic gene structure. Mol Biol Evol. 23:450–468.

Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. Science. 290:1151–1155.

Lynch M, Conery JS. 2003. The origins of genome complexity. Science. 302:1401–1404.

Lynch M, Force A. 2000a. Gene duplication and the origin of interspecific genomic incompatibility. Am Nat. 156:590–605.

Lynch M, Force A. 2000b. The probability of duplicate gene preservation by subfunctionalization. Genetics. 154:459–473.

Macgregor HC, Uzzell TM Jr. 1964. Gynogenesis in salamanders related to *Ambystoma Jeffersonianum*. Science. 143:1043–1045.

Mantovani B, Scali V. 1992. Hybridogenesis and androgenesis in the stick-insect *Bacillus-Rossius Grandii-Benazzii* (Insecta, Phasmatodea). Evolution. 46:783–796.

Mu J, Duan J, Makova KD, Joy DA, Huynh CQ, Branch OH, Li WH, Su XZ. 2002. Chromosome-wide SNPs reveal an ancient origin for *Plasmodium falciparum*. Nature. 418:323–326.

Nei M, Kumar S. 2000. Molecular evolution and phylogenetics. New York: Oxford University Press.

Ohta T. 1992. The nearly neutral theory of molecular evolution. Annu Rev Ecol Syst. 23:263–286.

Opperman R, Emmanuel E, Levy AA. 2004. The effect of sequence divergence on recombination between direct repeats in Arabidopsis. Genetics. 168:2207–2215.

Patterson DJ. 1999. The diversity of eukaryotes. Am Nat. 154:S96–S124.

Pritchard AE, Seilhamer JJ, Mahalingam R, Sable CL, Venuti SE, Cummings DJ. 1990. Nucleotide sequence of the mitochondrial genome of *Paramecium*. Nucleic Acids Res. 18:173–180.

Przybos E. 1975. Genetic studies of *Paramecium jenningsi* strains (Diller, Earl, 1958). Folia Biol (Krakow). 23:425–471.

Przybos E, Rautian M, Greczek-Stachura M, Potekhin A. 2007. Polymorphism within *Paramecium sexaurelia* (Ciliophora, Oligohymenophorea) and description of a new stand of the species in China. Folia Biol (Krakow). 55:121–125.

Przybos E, Tarcz S, Skoblo I. 2007. First American stand of *Paramecium novaurelia* and intra-specific differentiation of the species. Folia Biologica-Krakow. 55:53–63.

Pulliam HR. 1988. Sources, sinks, and population regulation. Am Nat. 132:652–661.

Ragghianti M, Bucci S, Marracci S, Casola C, Mancino G, Hotz H, Guex GD, Plotner J, Uzzell T. 2007. Gametogenesis of intergroup hybrids of hemiclonal frogs. Genet Res. 89:39–45.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 132:365–386.

Salim HMW, Ring KL, Cavalcanti ARO. 2008. Patterns of codon usage in two ciliates that reassign the genetic code: *tetrahymena thermophila* and *Paramecium tetraurelia*. Protist. 159:283–298.

Sawyer SA. 1999. GENECONV: a computer package for the statistical detection of gene conversion. Distributed by the author, Department of Mathematics, Washington University in St. Louis, available at http://www.math.wustl.edu/~sawyer.

Schultz RJ. 1969. Hybridization, unisexuality, and polyploidy in teleost *Poeciliopsis* (Poeciliidae) and other vertebrates. Am Nat. 103:605–619.

Smith DR, Lee RW. 2008. Nucleotide diversity in the mitochondrial and nuclear compartments of *Chlamydomonas reinhardtii*: investigating the origins of genome architecture. BMC Evol Biol. 8:156.

Snoke MS, Berendonk TU, Barth D, Lynch M. 2006. Large global effective population sizes in *Paramecium*. Mol Biol Evol. 23:2474–2479.

Sonneborn TM. 1937. Sex, sex inheritance and sex determination in *Paramecium Aurelia*. Proc Natl Acad Sci USA. 23:378–385.

Sonneborn TM. 1957. Breeding systems, reproductive methods, and species problems in protozoa. In: Mayr E, editor. The species problem. Washington DC: AAAS.

Sonneborn TM. 1970. Methods in *Paramecium* research. Methods in cell physiology. New York, London: DMPA Press. p. 241–339.

Sonneborn TM. 1974. Paramecium aurelia. Handbook of genetics. New York: R.C.K.P. Press.p. 469–594.

Sonneborn TM. 1975. The *Paramecium-aurelia* complex of 14 sibling species. Trans Am Microsc Soc. 94:155–178.

Stoeck T, Przybos E, Schmidt HJ. 1998. A combination of genetics with inter- and intra-strain crosses and RAPD-fingerprints reveals different population structures within the *Paramecium aurelia* species complex. Eur J Protistol. 34: 348–355.

Stoeck T, Schmidt HJ. 1998. Fast and accurate identification of European species of the *Paramecium aurelia* complex by RAPD-fingerprints. Microb Ecol. 35:311–317.

Tarcz S, Przybos E, Prajer M, Greczek-Stachura M. 2006. Intraspecific variation of diagnostic rDNA genes in *Paramecium dodecaurelia*, *P. tredecaurelia* and *P. quadecaurelia* (Ciliophora: oligohymenophorea). Acta Protozool. 45: 255–263.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.

Watkinson AR, Sutherland WJ. 1995. Sources, sinks and pseudo-sinks. J Anim Ecol. 64:126–130.

Waugh J. 2007. DNA barcoding in animal species: progress, potential and pitfalls. Bioessays. 29:188–197.

Weisse T. 2008. Distribution and diversity of aquatic protists: an evolutionary and ecological perspective. Biodivers Conserv. 17:243–259.

White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenies. In: Innis MA, Gelford DH, Sninsky JJ, White TJ, editors. CR protocols: a guide to methods and applications. New York: Academic Press. p. 315–322.

Whitlock MC, Barton NH. 1997. The effective size of a subdivided population. Genetics. 146:427–441.

Wright S. 1931. Evolution in Mendelian populations. Genetics. 16:97–159.

Laura Katz, Associate Editor

Accepted November 11, 2008