Supplementary Information for Zoller *et al.*

Table S1. Primer sequences. A list of the primers used to PCR amplify full nanochromosome sequences to each telomere. All new protein-coding sequences are accession numbers HQ432909-HQ432953. Other accession numbers are GenBank sequences that were used to design primers.

Taxon (initials)	**Accession Number**	**Round 1 primers** (forward/reverse, 5'-3')	**Round 2 primers** (nested PCR)	**Program used**[1]
elongation factor-1 (ef-1)				
O.t.	HQ432919	CGATGCCCCAGGTCACAGAGA TTTCATCAAGAAC	CGATGCCCCAGGTCACAGAG ATTTCATCAAGAAC	A
		ACAGTTCCGATACCACCGATCT TGTAG	ACAGTTCCGATACCACCGAT CTTGTAG	A
S.l.	HQ432920	CCACTCAGACTCCCACTCCAAG ACG	CAAGATCGGTGGTATCGGAA CTGTCCC	A
		ATTTGCAGGCAATATGGGCAG TGTGG	GACCTGGATGGTTAAGAACG ATAACTTGAGC	A
S.h.	HQ432921	CCACTCAGACTCCCACTCCAAG ACG	CAAGATCGGTGGTATCGGAA CTGTCCC	A
		ATTTGCAGGCAATATGGGCAG TGTGG	GACCTGGATGGTTAAGAACG ATAACTTGAGC	C
S.n.	HQ432922	CCACTCAGACTCCCACTCCAAG ACG	CAAGATCGGTGGTATCGGAA CTGTCCC	A
		ATTTGCAGGCAATATGGGCAG TGTGG	GACCTGGATGGTTAAGAACG ATAACTTGAGC	C
N.o.	HQ432923, AM897492	TCAGGTCACTGGTCAGAAGAG AGATTCAACG	CCCAGACAAGATTCCATTCA TCCCCATCTCAG	A
		CAGTTCCGATACCACCGATCTT GTAGACG	GGAGTGGGAGTCTGAGTGG CTTATCCTTGG	A
2-oxoglutarate dehydrogenase (sucA)				
O.t.	HQ432934	AGAAAGATGGGACACAACGAA CTTGATGCTCC	AGAAAGATGGGACACAACG AACTTGATGCTCC	B
		ACAGCGTGTCTGTGTGAGAAA GTACC	ACAGCGTGTCTGTGTGAGAA AGTACC	B
S.l.	HQ432935	AGAAAGATGGGACACAACGAA CTTGATGCTCC	AGAAAGATGGGACACAACG AACTTGATGCTCC	C
		ACAGCGTGTCTGTGTGAGAAA GTACC	ACAGCGTGTCTGTGTGAGAA AGTACC	A
S.h.	HQ432936	AGAAAGATGGGACACAACGAA CTTGATGCTCC	AGAAAGATGGGACACAACG AACTTGATGCTCC	A
		ACAGCGTGTCTGTGTGAGAAA GTACCTCTCTC	ACAGCGTGTCTGTGTGAGAA AGTACCTCTCTC	A
S.n.	HQ432937	AGAAAGATGGGACACAACGAA CTTGATGCTCC	AGAAAGATGGGACACAACG AACTTGATGCTCC	C
		ACAGCGTGTCTGTGTGAGAAA GTACCTCTCTC	ACAGCGTGTCTGTGTGAGAA AGTACCTCTCTC	A
N.o.	HQ432938	GGGAAGCHYAATTCGGTGATT TCGC[2]	GGGAAGCHYAATTCGGTGA TTTCGC	A
		GTAAGGGTCGTCGTCGGTCATC TGC	CGTACCCGTGGGGAAGTAGC AGC	A

[1] See Methods for a description of each program type for telomere-specific PCR.

ribosomal protein L10 (rL10)				
O.t.	HQ432929	GCAAGGTATGAGAGGTGCCTTCGGAAAGC	CATTTCCATCAGAACCACTGCCCCCAACG	A
		CACCGTCTGGAATAATCTTGCCCTCTCTTTCG	GAACTCTTGCTTAGTGAACTTGGTGAAGCCG	A
S.l.	HQ432930	GGAGCCGATAGACTTCAGCAAGGTATGAGAGG	ATCATTTCCATCAGAACCACTGCTCCCAACG	A
		GTCGGGGACGATTTTACCCTCTTTCTCG	TGGAATTCTTGTCTGGTGAACTTAGTGAATCCG	A
S.h.	HQ432931, AY883854	GCAAGGTATGAGAGGTGCCTTCGGAAAGC	CATTTCCATCAGAACCACTGCCCCCAACG	A
		TAACTCCGTCAGAGATGATCTTGCCTTCTCTCTCG	GTGGAACTCTTGGAAGGTATACTTGGTGAAACC	A
S.n.	HQ432932	GCAAGGTATGAGAGGTGCCTTCGGAAAGC	CATTTCCATCAGAACCACTGCCCCCAACG	A
		TAACTCCGTCAGAGATGATCTTGCCTTCTCTCTCG	GTGGAACTCTTGGAAGGTATACTTGGTGAAACC	A
N.o.	HQ432933, AM899590	TGCCGGAGCTGATAGATTACAGACAGGAATG	CGGACAGATAATAATGTCAGTTCGCTGCAAGG	A
		CCATATCGCTCCAACTGTCCATCCTTCTTTAG	CGTGGGATATGGCTAAATCCGTGGG	A
turgor pressure sensor (turgor)				
O.t.	HQ432944	AGAACTATTCTTGAACTTGGAGGAAACAATGCTGC	CAACCGTTGCACAACTTTGCGTAGAATCATTATTCATG	A
		GTCTTGAACTTAATGACGTATACAACTGGAGCGAAGAG	TTTCCTTAACCTTCTACGATATTACCTCCAACAAGCAC	A
S.l.	HQ432945	ACCATTCTTGAACTAGGAGGAAACAATGCTGCTG	CCAGATGCTGATCTAGAACTTGCTTTGAAGGC	A
		CTAGAGTCTTGAACTTGATAACATATACAACNGGTGCG	GTCCCTCTACAACTGCTCCTCCGAATAGG	A
S.h.	HQ432946, AY883858	GTGGTAACAATGCTGCAGTTATTATGCCAGATGCTG	CGTTGCACCACACTCCGTAGAATTATTATTCACG	A
		CCTTCACCAGGTACGATATTTCCACCTACTAGG	GGACTTTACCACCTTGACTCTTAATAGCTTCGATACC	A
S.n.	HQ432947	GTGGTAACAATGCTGCAGTTATTATGCCAGATGCTG	CGTTGCACCACACTCCGTAGAATTATTATTCACG	A
		CACCTTGGCTCTTAATAGCTTCGATACCCTC	CCAAGTAATGTATCTGGGTGTAATGGATCGCC	A
N.o.	HQ432948	GGAAGCTTATGTAGAGTATCAGCCAACCGG	AAACCTTCACCAACTACTCCTCAATCTGGAATTATC	A
		CTCTTGCTGATGCTCCTAATTTAGCCGCCTATTC	GCGTCGTGAAGAACAACAAAAGGATCTGAGC	A
ubiquitin conjugating enzyme E2 (Ubc)				
O.t.	HQ432949	TAGTTGTTGGTATTTATGCTGAGCCAAGCGAGAAC	TTCTTTGTTAAAATCGCAGGACCAGAGGGAACC	B
		TCCACTTGTCCTTCAAGATATCAAGACAGATTCTGCC	GTCTTAATTACAATGCACATACCTAGCTTATCAATATTTGGG	B
S.l.	HQ432950	GATCAGAAATGGCTGGCATTCCTCCCAG	GGAAACTCAACGTCTTGCTTAAGAACCAGG	B
		CATCTGTTTGCATGTTCACTACTCCTATCACTGG	GATGGCAGCTTCTGGGTCTTTCTTCCA	B
S.h.	HQ432951, AY883855	TAGTTGTTGGTATTTATGCTGAGCCAAGCGAGAAC	GTGAAAATTGCTGGTCCAGAGGGCACTACC	A
		GGGCTCCATTTATCCTTCAAAATATCGAGGC	GGCAAATTCTTCCTATAAATTAATGTGGGATTAGTAAGAGG	A
S.n.	HQ432952	TAGTTGTTGGTATTTATGCTGAGCCAAGCGAGAAC	CACAGACATTTCTTTGTGAAGATTGCAGGC	B
		ATTTTGAAGGACAAGTGGAGCCCCGC	GTTCTGGAAAATATAAACCAATAGTGGTCGGGTTAG	B
N.o.	HQ432953	GAACTCATGGGCATCGGCAAGGACC	TGGTCCTGCCAACGAAAAAGAAATGTACC	A
		TACTTCCTGCACCACTCTCTTGCTGTCG	CGGAACCAAAGGATCGTCAGGATTAGG	A

[2] Degenerate PCR primers derive from a consensus of *O.t., S.l., S.h., T. thermophila,* and *P.*

heat shock protein 70 (hsp70)				
O.t.	HQ432924	ACGTGTTAAGGTAGTATATTTAAG	-	51°C anneal[3]
		TTAAGAAAAAAACTATTTATAATAGGC	-	
S.l.	HQ432925	CATCAACAAAGGTATTCAATAAG	-	51°C anneal[4]
		GGTCTTCGTGGTCTTCTTC	-	
S.h.	HQ432926	GCCAGTCTTGCTAAAGGTCAAGTTCACG	CGGAGGTTCCACTAGAATCCCCAAGG	A
		CGGGGTTGATAGATCTGTTGAGGGTCTTTCC	CCTTGGGGATTCTAGTGGAACCTCCG	A
S.n.	HQ432927, U37280	CGATCTCGGTGGTGGTACTTTCGACG	CGATCTCGGTGGTGGTACTTTCGACG	A
		GATTTGTCCCTTGGCAAGACTGGCG	GATTTGTCCCTTGGCAAGACTGGCG	A
N.o.	HQ432928, AJ871325	GCGGGATCGTGCTTTGTCATTTGC	GGGTGCGAAGTTGGAGGAGTTGATATTG	A
		CAATATCAACTCCTCCAACTTCGCACCC	GCAAATGACAAAGCACGATCCCGC	-
actin I				
O.t.	HQ432909	CATTCAACCCGTCCATAGCC	-	55°C anneal[3]
		GTGACAAATATAATTTAAGTCAAGACTG	-	
S.l.	HQ432910	GCAAGGATATTTAAGTAAGGGCATATC	-	56°C anneal[4]
		ACCAACCTCATTCAACCTGTCC	-	
S.h.	HQ432911	CGTAGTCCAAAGCGACGAAGCAAAGC	GCTTTTCCTTGATGTCTCTGACGATTTCCATCTC	A
		GATGACCCAGATCATGTTCGAGACCTTCAATG	CCTCTCCCTCTACTCCGCTGGTAGAAC	A
S.n.	HQ432912, M22480	CGTAGTCCAAAGCGACGAAGCAAAGC	GCTTTTCCTTGATGTCTCTGACGATTTCCATCTC	A
		GATGACCCAGATCATGTTCGAGACCTTCAATG	CCTCTCCCTCTACTCCGCTGGTAGAAC	A
N.o.	HQ432913, AM898066	CTGTCTTCCCTGAAGTTATTGGCAGACCC	CTGGTGTTCTCATTGGAATAGGCACAAAGG	A
		GATGGCAACATAGAATGTCGGAACTTTGAAGG	GGATCATTCTTTCACGATTCATCTTTGGATTCAGTGG	A

tetraurelia.

[3] Designates that standard PCR was used to confirm a database sequence from shotgun genome sequencing. Thermocycler program for all standard PCR reactions was: 95°C hold for manual hot start; (95°C 30", Anneal Temp° 30", 65°C 2 min) x 25; 65°C 5 min.; 4°C hold.

[4] Designates that standard PCR was used to verify a sequence from a partial 454 genome assembly. Thermocycler program for all standard PCR reactions was: 95°C hold for manual hot start; (95°C 30", Anneal Temp° 30", 65°C 2 min) x 25; 65°C 5 min.; 4°C hold.

choline-phosphate cytidylyltransferase-γ (cct-γ)				
O.t.	HQ432914	ATGTAACAAAAAGAATTGGAATATG	-	51°C anneal[3]
		TATAAAGTTCAATATGTCAAACAG	-	
S.l.	HQ432915	GGCAACTCCGAGGGCATCATG	CGTTCAAGACATCCTTTGAGGCTCCTC	A
		GAGGAGCCTCAAAGGATGTCTTGAACG	CATGATGCCCTCGGAGTTGCC	A
S.h.	HQ432916	CTGAAACACCTTTCTCAGTGATAACAACATCTGGC[5]	CCATGTTGGTCATGGACTCTCCC	A
		CTTGCATTGGAACTAAATTCTCATCCAAGTGGGG	GCAGTCAAGACTGTCTACAAGAAGGACGG	A
S.n.	HQ432917	AGGACGGCAACATCCTGGGG[5]	CTGGAGTACAACTACAACAGCCACAACG	A
		GCACTCTCGAGTACAAAAAGGGAGAGTCC	GCACTCTCGAGTACAAAAAGGGAGAGTCC	A
N.o.	HQ432918, AM898381	TCTCTATCCTCCTTCTCATCTGTGGATGCG	CATCAAGTACAACACAATCTTCTATTGTTCCTCCTG	A
		CCCAACCATTATTGTGTCAGCCTACTTCC	CCCTTAAAGTAGTGGATGAAATAGCCACAGAC	A
tata box binding protein (tbp)				
O.t.	HQ432939	GCTGAATACAACCCAAAGCGTTTCGCTGC	GTCATTATGCGTATTAGAGATCCAAAAACAACTGC	A
		CTGTAAATTAGACCTGGGAAGATTTCRGG	CTGTAAATTAGACCTGGGAAGATTTCRGG	A
S.l.	HQ432940	GCTGAATACAACCCAAAGCGTTTCGCTGC	GTGATCATGCGTATCAGAGAACCAAAGACTACAGC	A
		CTGTAAATTAGACCTGGGAAGATTTCRGG	ACCTTCTTGATTGCCTTTGAATACTTCTTTGCAGC	C
S.h.	HQ432941	GTTCTGGCAAAATGGTCTGTACTGGAGC	GTAGAGATGCTGCCAAGAAATACTCAAAGGC	A
		TCTAAAGAAATGGGGAAGCCGACATCG	TCTAAAGAAATGGGGAAGCCGACATCG	A
S.n.	HQ432942	GTCATTATGCGTATTAGAGATCCAAAAACAACTGC	GTCATTATGCGTATTAGAGATCCAAAAACAACTGC	C
		GCAGCGAAACGCTTTGGGTTGTATTCAGC	GCAGCGAAACGCTTTGGGTTGTATTCAGC	C
N.o.	HQ432943, AM890164	GTGAATAGCGGGAACAGTATAGTGTGGCG	CCTGTAACACCAAAAAGAAGTACAGAACCAAGATTAGAG	A
		TTAATTCCAAACTCAACTCCACATGATGCTACC	GCCTAATTTCTTAATAGTTTTAGCATATTTCATAGCAGC	A

[5] *S.n* and *S.h.* primers designed against a consensus of *O.t*., *S.l.* and *O. granulifera*, accession Y11967.

Collections and Names of Stocks

Over 500 individual cell isolations were made from 30 different aquatic sites (streams, rivers, ponds and lakes) during a one-year period (1985-1986) that included all major seasons of the year. Individual isolates (clones) were considered separate stocks and assigned a stock number based on the aquatic source, sampling period, the collection site and the isolate (clone) number. For example, JRB310 refers to the Jordan River, sample B, at site three, clone 10. Three major aquatic sources provided the bulk of the related strains of *Oxytricha*-like organisms investigated further. These included the Jordan River (JR) on the campus of Indiana University, Bloomington, Indiana, Twin Lakes (TL) on the west edge of Bloomington, Indiana, and White River (WR) in Muncie, Indiana. Interestingly, all *Oxytricha* isolates that showed any hybridization to cloned *O. fallax* macronuclear gene probes or micronuclear TBE1 probes were isolated from winter samples, and none of the late spring, summer or early fall samples showed any probe hybridization under the conditions employed. The original *O. fallax* (*unifallax*) was also isolated in winter – February 1966. This suggests the intriguing possibility that species abundance may be seasonal or climate related. Sample A was collected late November 1985, and B in early January 1986. Most of the stocks that have been characterized as related to the original *O. fallax* by hybridization of DNA probes – including all *O. trifallax* – have been cryogenically preserved and stored in liquid nitrogen. Cysts of the original *O. fallax* (now *O. unifallax*) have been successfully excysted after 25 years. The cryogenic maintenance of all of these individual isolates (especially of all the *O. trifallax* stocks) provides a large source of natural genetic diversity for future research.

Length and Width Measurements

Length and width measurements of the original *O. fallax* (*unifallax*), *O. trifallax* (JRB510), *O. trifallax* (JRB310), and *S. histriomuscorum* (BA) were made at 1000x using an Olympus BH-2 microscope equipped with an ocular micrometer. Table S2 lists the sample size (n), length, width, mean, standard deviation, minimum and maximum for each of these cell lines. While *O. trifallax* is generally slightly larger than *O. unifallax,* the overlap in these measurements makes these features unreliable for taxonomy amongst the oxytrichids, but significantly distinguishes them from *S. histriomuscorum* (BA). All measurements were made on lightly-fed cells that had not entered into cortical morphogenesis in preparation for cell division or encystment.

Cell Type (Stock #)	N	Length Mean	Std. Dev.	Min.	Max.	Width Mean	Std. Dev.	Min.	Max.
O. unifallax (9D1)	30	65.05	1.27	55.50	86.58	32.82	4.09	25.53	42.18
O. trifallax (JRB510)	15	66.53	5.10	57.72	76.59	36.56	1.60	33.30	38.85
O. trifallax (JRB310)	29	72.42	8.43	57.72	92.13	35.10	3.10	31.08	41.07
S. histriomuscorum (BA)	15	102.40	10.22	91.02	134.31	51.68	3.89	45.51	61.05

Table S2. Length and width comparisons for *O. unifallax* , *O. trifallax* and *S. histriomuscorum* BA

Size measurements for length and width for the original *O. fallax* (*O. unifallax*), two stocks (common lab strains) of *O. trifallax* and *S. histriomuscorum* BA. Sample size (N), length, width, mean, standard deviation, minimum and maximum (all in μm) for each cell line are provided.

Mating Compatibility Groups and Mating Types

All cell lines ultimately assigned to the *Oxytricha trifallax* sibling species group were so assigned because of their ability to mate with one or more other cell lines in this complex. Extensive mating tests have been conducted over a period of years and have allowed individual isolates to be characterized for their mating type and assigned to one of eleven specific mating compatibility groups (manuscript in preparation); each of the matings has been repeated several times with the same result. Note that mating groups should not be equated to mating types, because different individual lines within a group differ somewhat in their mating capacity. Mating type characterization involves not only whether a cell line can mate with another individual line, but also the distinctive pattern of cell lines that display mating compatibility. This is consistent with how Valbonesi, Ortenzi, and Luporini (1992) define mating types for the *Euplotes crassus* group. The mating capacity of 41 isolates has been looked at extensively and with the exception of two cell lines, all cell lines within a mating group show somewhat different patterns of mating, and hence are presumably of different mating types. While one could argue that the differences among members in a group are due to the environment, such as temperature, or age of the clone, we have no evidence to support this prospect and, in addition, the reproducibility of the mating patterns over years argue against such an impact. Cell lines **JRB27** and **JRB51** are the "promiscuous maters", compatible with virtually all other mating groups. A total of 22 different cell lines can mate with each of these cell lines, and collectively **JRB27** and **JRB51** mate with 33 of the 41 cell lines surveyed; **JRB27** and **JRB51** can also mate with each other. Consequently, each of these lines is assigned its own mating group (Group V and VI respectively). If cell line JRA52 (Group VIII) is included in a mating survey, these 3 cell lines (JRB27, JRB51, and JRB52) can mate with almost 90% of all isolates.

History of strains and publications related to *O. trifallax*

We are not aware of any published studies of *O. trifallax* that named it exclusively *S. histriomuscorum*; however, from the time of introduction of that taxon, many authors have provided both names. Papers where *S. histriomuscorum* has been used as the exclusive name, notably elegant work on encystment from A. Baroin-Tourancheau and collaborators (*e.g.*, Baroin-Tourancheau *et al* 1999; Grisvard *et al*. 2008; Fryd-Versavel *et al*. 2010) have used the *S. histriomuscorum* stock BA, isolated by Baroin-Tourancheau *et al*. (1999).

Furthermore, work on *O. trifallax*, other than from our labs, has been reported repeatedly in the literature, but only in a few cases are we confident that the work was with the species we describe in this paper. Adl and Berger obtained *O. trifallax* JRB310 and 510 from the Herrick lab to perform the seminal work reported in 1997 & 2000, but give the name as "S. *histriomuscorum*". In 1989 the Prescott lab (Greslin *et al*. 1989) report studying the two clones of *O. trifallax* supplied by one of us, RLH, and in 1992 (Prescott and Greslin 1992) again report data from JRB 310, one of those clones. But in 1991 they isolated a new ciliate they named "*O. trifallax* WR" from a Colorado lake (DuBois and Prescott 1995). Except for Prescott & Greslin (1992), who reported data from *O. trifallax* JRB310, work reported on "*O. trifallax*" after 1991 was stated to have used DNA or cells provided by Prescott (e.g. Lingner, Hendrick and Cech, 1994; Bryan *et al*. 1998; Inagaki and Doolittle, 2000, 2001), and was not further designated as JRB310, JRB510 nor WR.

Especially confusing are two reports from Foissner and colleagues (Foissner *et al*., 2004; Schmidt *et al*., 2007) in which they investigate the relationships of various ciliates using sequenced rDNA from *Sterkiella histriomuscorum*. In the 2004 paper the *S. histriomuscorum* sequence AF164121 (from "Indiana, USA, Jordan River", though the source of this sequence is not clear, see footnote "f" of their Table I) falls almost exactly on their tree like *S. histriomuscorum* strain BA does in the current study, but in the tree in the 2007 paper (Schmidt *et al*. 2007) the *S. histriomuscorum* (AF508770) groups very similarly to our *O. trifallax* in the current study. AF508770 refers to the GenBank accession number [not a strain] for the SSU rRNA sequence used by Hewitt *et al*. (2003), who give its source as "*Sterkiella histriomuscorum* (formerly named *Oxytricha trifallax*)…isolated from the Jordan River, Bloomington, Indiana." Because Prescott received cultures of *O. trifallax* strains JRB310 and JRB510 prior to 2003, the DNA used in Hewitt *et al*. may have been from the original *O. trifallax* that was later misnamed *S. histriomuscorum*.