

Supplementary Figures and Tables

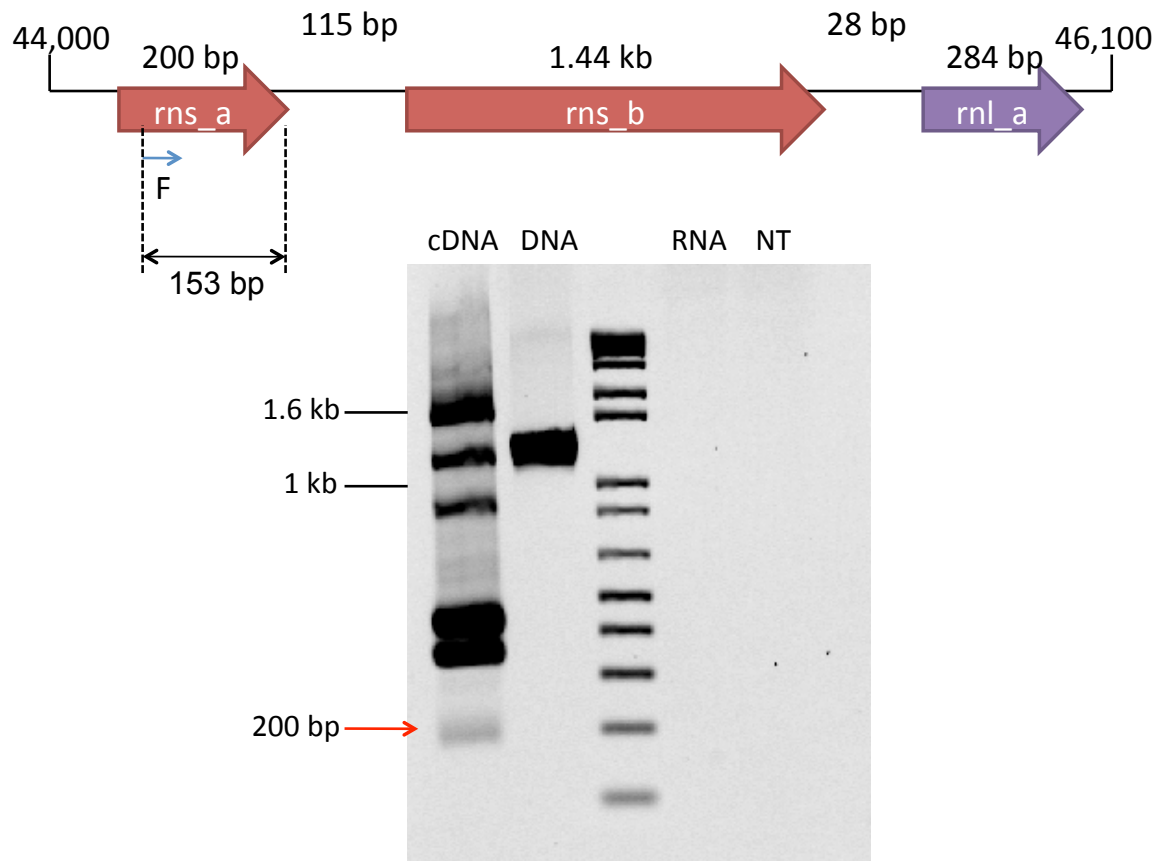
Supplementary Table 1. Annotation of unknown ciliate mitochondrial ORFs

Gene	Annotation description
<i>rps2</i>	We found convincing matches to <i>rps2</i> in both <i>Oxytricha</i> and <i>Nyctotherus</i> (HHSearch p-value 7e-07 and 3e-07, respectively). No transmembrane helices are predicted for the putative <i>rps2</i> and the open reading frames appear to be homologs based on the alignment of this region. The potentially syntenic region in the <i>Euplotes</i> mitochondrial genome appears to be unsequenced (sequencing ends in the middle of <i>nad5</i> , which is the gene preceding <i>rps2</i> in <i>Oxytricha</i>).
<i>rps4</i>	The <i>ymf76</i> gene found in <i>Tetrahymena</i> (<i>ymf81</i> and <i>ymf85</i> in <i>Paramecium</i>), previously classified as "putative ribosomal" (Brunk <i>et al.</i> 2003), is homologous to <i>Euplotes rps4</i> (de Graaf <i>et al.</i> 2009), with the <i>Oxytricha rps4</i> providing a transitive link between spirotrichs and oligohymenophorans.
<i>rps7</i>	<i>rps7</i> has HHpred hits (p-value 6.6e-4) for both <i>Oxytricha</i> and <i>Tetrahymena</i> (<i>ymf63</i>) and is syntenic to <i>Euplotes ORF170</i> .
<i>rps8</i>	We were able to classify an additional ribosomal superfamily S8 protein in <i>Oxytricha</i> (<i>rps8</i>) using HHpred (this is also detectable by BLAST: the best hits are to <i>rps22/15</i> proteins, which are part of this superfamily). We found a weak tblastn hit to <i>T. pyriformis ymf74</i> , which is syntenic to <i>Paramecium ymf84</i> (Burger <i>et al.</i> 2000) (e-value 5e-03). The <i>T. pyriformis</i> entry was previously classified as 'putative non-ribosomal', but had no predicted transmembrane region like the other mitochondrial ribosomal proteins (Brunk <i>et al.</i> 2003). We were also able to find a nuclear-encoded <i>rps8</i> homolog in <i>Tetrahymena</i> , <i>Paramecium</i> (gi:118360340 and 145483251) and an unpublished <i>Oxytricha</i> Mac-encoded protein. These homologs are more similar to other eukaryotic <i>rps8</i> proteins than to the mitochondrially-encoded <i>rps8</i> . The <i>Euplotes ORF125</i> protein appears to be syntenic with the <i>Oxytricha rps8</i> .
<i>rps10</i>	We discovered <i>rps10</i> in <i>Oxytricha</i> using HHpred. This gene is syntenic with <i>Euplotes ORF111</i> and appears to be homologous to the <i>Tetrahymena</i> putative ribosomal gene <i>ymf59</i> and <i>Paramecium ORF124</i> .

Supplementary Figures and Tables

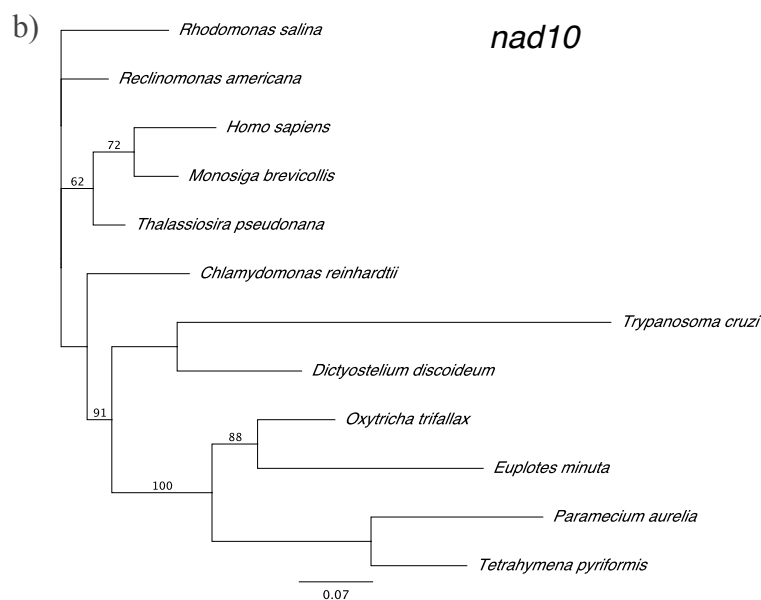
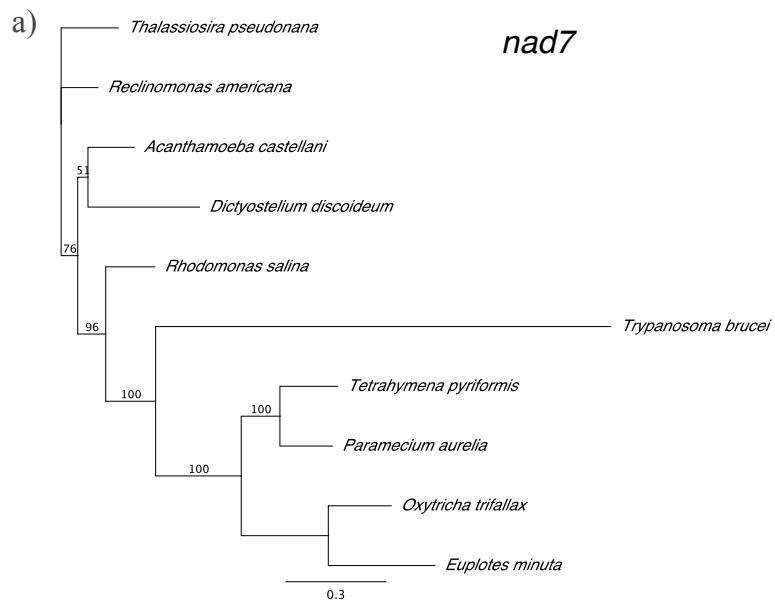
Gene	Annotation description
<i>rps13</i>, <i>rps14</i>, <i>rps19</i>	BLAST searches of <i>Tetrahymena</i> and <i>Paramecium</i> mitochondrial proteins indentified <i>Oxytricha rps13</i> and <i>rps19</i> . These genes are syntenic in <i>Euplotes</i> . In <i>Euplotes</i> , <i>ORF102</i> is <i>rps19</i> , while <i>ORF155</i> plus an N-terminal portion of <i>Euplotes</i> “ <i>rps4</i> ” corresponds to <i>rps13</i> . We were also able to identify <i>rps14</i> (synteny block Ib) in <i>Oxytricha</i> and <i>Euplotes</i> ; in the latter, <i>rps14</i> corresponds to <i>ORF49</i> plus a portion of unannotated sequence preceding it (before <i>nad7</i>).
<i>nad6</i>	We were also able to find <i>nad6</i> in the <i>Oxytricha</i> and <i>Euplotes</i> mitochondrial genomes. The <i>Oxytricha</i> ortholog shows weak sequence similarity to the <i>Paramecium</i> ortholog ORF256 (a BLAST hit to the NCBI nrdb: 22% identity, e-value of 0.005), which has previously been shown to be <i>nad6</i> (Brunk <i>et al.</i> 2003). The gene is either an annotation error or a fusion with the C-terminus of the <i>cox1</i> gene in <i>Euplotes</i> , which was previously noted as an atypical 267 amino acid C-terminal extension of <i>cox1</i> (de Graaf <i>et al.</i> 2009) – the <i>Oxytricha nad6</i> gene is immediately preceded by <i>cox1</i> and is ~27% identical to this "extension". This gene is predicted to have between 5 (<i>Oxytricha</i> , <i>Paramecium</i> , <i>Tetrahymena</i>) and 6 transmembrane helices (<i>Euplotes</i>) using THMM2 (Krogh <i>et al.</i> 2001). We also noticed that the <i>Nyctotherus</i> hydrogenosome genome fragment has a gene (<i>ORF236</i> from accession AJ871267.1) of similar length (237 aa) in a region that, with the exception of the loss of <i>cox1</i> , appears to be syntenic to <i>Oxytricha</i> and <i>Euplotes</i> , and also has 5 transmembrane helices predicted by THMM2.

Supplementary Figures and Tables



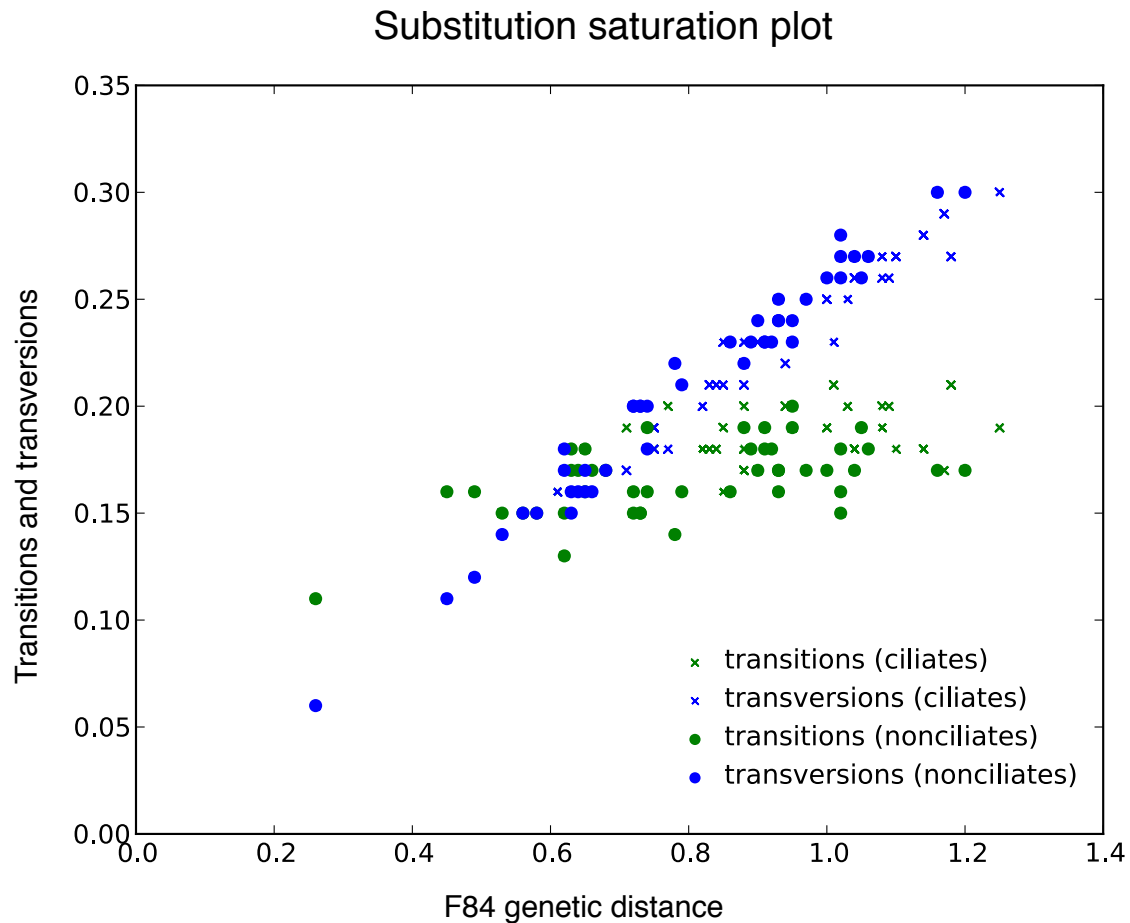
Supplementary Figure 1. Assessment of the putative split *rns* gene by RACE. The RACE products suggests that the *rns_a* transcript exists alone. The expected PCR product size including gene, tail, and primer was ~190 bp and observed size is 200 bp and was confirmed by DNA sequencing. Non-specific bands in cDNA and DNA lanes were due to mis-priming of the anchor primer. cDNA = RACE products; DNA = whole cell DNA, shows no ~190 bp product since no tailing present; RNA = diluted RNA used in RT reaction, which shows no DNA contamination; NT = no template control.

Supplementary Figures and Tables



Supplementary Figure 2. a) PhyML maximum likelihood phylogeny of *nad7* (*T. pseudonana* - YP_316619.1; *R. americana* - NP_044782.1; *A. castellanii* - NP_042561.1; *D. discoideum* - NP_050085.1; *R. salina* - NP_066485.1; *T. brucei* - P21301.2; *T. pyriformis* - NP_049579.1; *P. aurelia* - NP_059408.1; *O. trifallax* - JN383842; *E. minuta* - ACX30966.1); b) PhyML maximum likelihood phylogeny of *nad10* (*R. salina* - NP_066482.1; *R. americana* - NP_044784.1; *H. sapiens* - EAW69508.1; *M. brevicollis* - XP_001749993.1; *D. discoideum* - XP_638331.1; *T. pseudonana* - XP_002292597.1; *C. reinhardtii* - XP_001700585.1; *T. cruzi* - EFZ26582.1; *O. trifallax* - JN383842; *E. minuta* - ACX30960.1; *P. aurelia* - NP_059405.1; *T. pyriformis* - NP_049576.1). Bootstrap values ≥ 60 for 100 replicates are indicated on the branch labels.

Supplementary Figures and Tables



Supplementary Figure 3. Transitions, transversions and F84 genetic distance calculated using DAMBE (Xia and Xie, 2001), over a mitochondrial SSU alignment (ClustalW2 alignment using the IUB matrix, gap opening and extension penalties of 15 and 6.66) for the following eukaryotes: *Schizosaccharomyces pombe* (X15738), *Acanthamoeba castellanii* (U12386), *Reclinomonas americana* (AF007261), *Dictyostelium discoideum* (AB000109), *Cyanidioschyzon merloae* (D89861), *Malawimonas jakobiformi* (AF295546), *Rhodomonas salina* (AF288090), *Monosiga brevicollis* (AF538053), *Emiliana huxley* (AY342361), *Thalassiosira pseudonana* (DQ186202), and the ciliates *Tetrahymena pyriformis* (AF160864), *Paramecium tetraurelia* (NC_001324) and *Oxytricha trifallax* (JN383842). Transitions and transversions for pairwise comparisons involving at least one ciliate are indicated by crosses. It can clearly be seen that transitions have reached saturation.