Article

Current Biology

The Function and Evolution of Motile DNA Replication Systems in Ciliates

Highlights

- The replication envelope is a new form of motile DNA replication system
- Calcium signaling and cytoskeletal regulation may coordinate the replication band
- Phylogenomics resolves the deep phylogeny of the ciliate phylum

Authors

Nicholas A.T. Irwin, Alexandros A. Pittis, Varsha Mathur, LeAnn J. Howe, Patrick J. Keeling, Denis H. Lynn, William A. Bourland

Correspondence

nickatirwin@gmail.com (N.A.T.I.), willbour@me.com (W.A.B.)

In Brief

The fragmented nature of ciliate chromosomes complicates genome replication. These issues can be overcome using motile DNA replication systems that systematically copy DNA while traversing the nucleus. Irwin et al. reveal the diversity of these systems and highlight a role of calcium signaling and cytoskeletal regulation in their function.

Irwin et al., 2021, Current Biology *31*, 66–76
January 11, 2021
https://doi.org/10.1016/j.cub.2020.09.077

CellPress

CellPress

Current Biology

Article

The Function and Evolution of Motile DNA Replication Systems in Ciliates

Nicholas A.T. Irwin,[1,7,*] Alexandros A. Pittis,[1] Varsha Mathur,[1] LeAnn J. Howe,[2] Patrick J. Keeling,[1] Denis H. Lynn,[3,4,6] and William A. Bourland[5,*]

[1]Department of Botany, University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[2]Department of Biochemistry and Molecular Biology, Life Sciences Institute, University of British Columbia, Vancouver, BC V6T 1Z3, Canada
[3]Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[4]Department of Integrative Biology, University of Guelph, Guelph, ON N1G 2W1, Canada
[5]Department of Biological Sciences, Boise State University, Boise, ID 83725, USA
[6]Deceased
[7]Lead Contact
*Correspondence: nickatirwin@gmail.com (N.A.T.I.), willbour@me.com (W.A.B.)
https://doi.org/10.1016/j.cub.2020.09.077

SUMMARY

DNA replication is a ubiquitous and conserved cellular process. However, regulation of DNA replication is only understood in a small fraction of organisms that poorly represent the diversity of genetic systems in nature. Here we used computational and experimental approaches to examine the function and evolution of one such system, the replication band (RB) in spirotrich ciliates, which is a localized, motile hub that traverses the macronucleus while replicating DNA. We show that the RB can take unique forms in different species, from polar bands to a "replication envelope," where replication initiates at the nuclear periphery before advancing inward. Furthermore, we identify genes involved in cellular transport, including calcium transporters and cytoskeletal regulators, that are associated with the RB and may be involved in its function and translocation. These findings highlight the evolution and diversity of DNA replication systems and provide insights into the regulation of nuclear organization and processes.

INTRODUCTION

DNA replication is the process by which genomic DNA is duplicated in anticipation of cell division. To ensure that the genome is replicated accurately and only once, a variety of mechanisms are employed that coordinate the assembly and initiation of replication complexes.[1] In eukaryotes, hundreds to thousands of dispersed DNA-encoded replication origins serve as docking sites upon which the pre-replication machinery can assemble.[2,3] However, the varying accessibility of eukaryotic genomes constrains *cis* regulators because replication origin firing is strongly influenced by local chromatin environments.[1,4–6] Although open, transcriptionally active chromatin is permissive to replication, silenced heterochromatin restricts or at least delays initiation.[2,7,8] In turn, accessibility to replication origins generates temporal patterns in DNA replication as cells establish replication timing profiles in which certain sites are consistently replicated before others.[9,10] This has been proposed to have functional implications for the establishment of transcriptional and chromatin states following cell division.[9,11] Indeed, replication patterns have been demonstrated in yeast and animals, and replication timing profiles can be evolutionarily conserved, as in mammals.[9,12] In addition to temporal control, chromatin-dependent DNA replication results in spatial coordination because replication origins in the interior of the nucleus fire earlier than those found at the nuclear periphery, a pattern that correlates with the distribution of euchromatin and heterochromatin and is independent of the position of specific chromosomes.[2,13] However, detailed investigations of DNA replication patterns and mechanisms have only been carried out in a small number of eukaryotic model systems despite the degree of nuclear diversity observed across the tree of eukaryotes. These diverse and underexplored systems can hold important clues for understanding nuclear mechanisms, such as whether replication timing profiles and spatial organization are functionally significant or, instead, are mostly byproducts of chromatin structure and transcriptional activity.

One system that can provide insights into the function and mechanisms governing the regulation of DNA replication is the macronuclear replication band (RB) in spirotrich ciliates. Spirotrichs are a large and diverse group of ciliated protozoa that includes model organisms such as *Oxytricha* and *Euplotes*.[14] Like all ciliates, spirotrichs have a transcriptionally quiescent germline micronucleus and a transcriptionally active somatic macronucleus, but they have also evolved a number of additional unusual nuclear characteristics. For instance, spirotrich macronuclear genomes are encoded on tens of thousands of endoreplicated gene-sized chromosomes that typically contain single genes flanked by short untranscribed regions and telomeres.[15,16] The abundance, small size, and distribution of these chromosomes within the macronucleus likely pose challenges for coordinating DNA replication that appear to be circumvented, at least in part, by the physical and temporal coordination of DNA replication in an RB. Spirotrich RBs are disk-like structures that

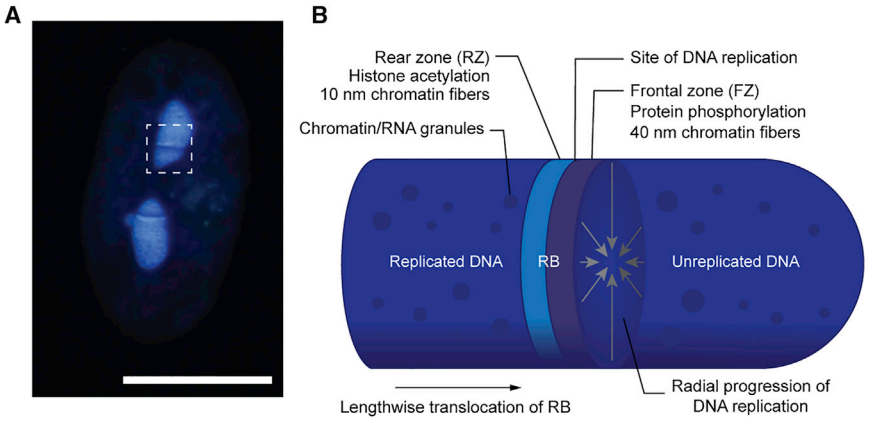

Figure 1. Representation of the replication band in *Oxytricha trifallax*

(A) A fluorescent micrograph depicting the replication band (RB) in *O. trifallax* following DAPI staining. Note the two lobes of the macronucleus and the stratified appearance of the RB. The bands will travel lengthwise along the linear axis of the elongated macronuclear lobes before merging and then disappearing at the nuclear midpoint. Scale bar, 50 μm.

(B) A diagrammatic representation of the spirotrich RB with key features labeled. As the RB progresses, DNA replication advances radially from the nuclear periphery toward the interior.

typically form at each pole of a long, ellipsoidal macronucleus during initiation of S phase and then migrate toward one another, generating a wave of chromatin modifications and replicated DNA before merging at the macronuclear midpoint (Figure 1).[17,18] The structure of the RB is stratified into multiple zones defined by differential chromatin states.[19–21] The frontal zone (FZ) is characterized by abundant protein phosphorylation and chromatin packaged in 40-nm fibers, whereas in the rear zone (RZ), histones become acetylated, and fiber size is reduced to 10 nm (Figure 1B).[17,22] At the interface between the FZ and RZ, replication proteins are localized, and DNA replication occurs at hundreds of foci that initially fire from the nuclear envelope before proceeding toward the nuclear interior (Figure 1B).[23–25] So far, the molecular machinery organizing the RB and controlling its orderly translocation is unknown. The altered chromatin states in the FZ and RZ suggest that chromatin modifiers play a role; however, the mechanisms coordinating the motility of the band and its capacity for initiating replication are more enigmatic, perhaps involving signaling cascades.[23,26] Given the vectorial nature of DNA replication in this system, the RB is a unique but so far largely unexplored model for investigating the underlying chromatin dynamics, nuclear changes, and spatiotemporal patterns associated with DNA replication.

To better understand the function and evolution of the RB and how it relates to spatial control over DNA replication in nuclear genomes, we aimed to reconstruct the evolutionary history of the RB and to identify processes related to RB function using phylogenomics and comparative genomics. Using transcriptomics data from two phylogenetically significant ciliate lineages related to spirotrichs, *Licnophora macfarlandi* and *Phacodinium metchnikoffi*, we show that the RB has a paraphyletic distribution because of the phylogenetic position of *P. metchnikoffi*. Rather than possessing a typical RB, *P. metchnikoffi* uses a unique system, here called a replication envelope, where DNA replication initiates simultaneously at the entire nuclear periphery before progressing into the nuclear interior in a concerted wave. Based on the phylogenetic distribution of RBs, we identified genes correlated with the canonical RB, many of which were involved in intracellular transport and some of which were required for DNA replication and progression through the cell cycle in *Oxytricha trifallax*, a model spirotrich exhibiting typical RBs. These data support a model suggesting a role for calcium signaling and cytoskeletal elements in coordinating DNA replication and nuclear processes in general.

RESULTS

Phylogenomics Show the Paraphyly of the RB

To reconstruct the evolutionary history of the RB, we generated transcriptomics data from two ciliates thought to be divergent relatives of the well-studied spirotrichs, *Licnophora macfarlandi* and *P. metchnikoffi*. *Licnophora macfarlandi* is a sessile marine ciliate known to exhibit RBs that inhabits the respiratory tree of the California sea cucumber (*Parastichopus californicus*) (Figures S1A–S1D).[27,28] In contrast, *P. metchnikoffi* is a free-living ciliate that inhabits fresh water pools and appears to lack an RB (Figures S1E–S1S).[29] Current phylogenetic placement of *P. metchnikoffi* is ambiguous, with different molecular phylogenies placing it within or sister to spirotrichs, indicative of a paraphyletic or monophyletic RB, respectively.[30–33] To reconcile these topologies, we sought to construct a phylogenomic dataset incorporating *L. macfarlandi* and *P. metchnikoffi*. We combined transcriptomics data from *L. macfarlandi* and *P. metchnikoffi* with 52 previously published genomic and transcriptomic datasets from other ciliates (n = 40) and outgroup taxa (n = 12) to construct a matrix encompassing 231 concatenated genes and 73,533 amino acid sites (Table S1). Maximum likelihood analyses of the concatenated alignment produced a robustly supported phylogeny that was consistent when using either free rate heterogeneity or empirical profile mixture substitution models (Figure 2A). Similarly, Bayesian analyses recovered a nearly identical phylogeny with all four chains converging on similar, equally well supported topologies (chain bipartition discrepancies: max difference = 0.175857, mean difference = 8.37×10^{-4}) (Figure 2A).

The resulting trees fully resolved the deep phylogeny of the ciliate phylum and reconciled the placement of *L. macfarlandi* and *P. metchnikoffi*. All analyses separated the ciliates into four well-supported and previously defined clades: SAL (Spirotrichea, Armophorea, and Litostomatea), CONThreeP (Colpodea, Oligohymenophorea, Nassophorea, Plagiopylea, Prostomatea, and Phyllopharyngea), Protocruziaea, and the subphylum Postciliodesmatophora (Figure 2A).[34] Inclusion of the karyorelictid ciliate *Loxodes* sp. reaffirmed the independent branching of *Protocruzia*; previous investigations had placed it sister to

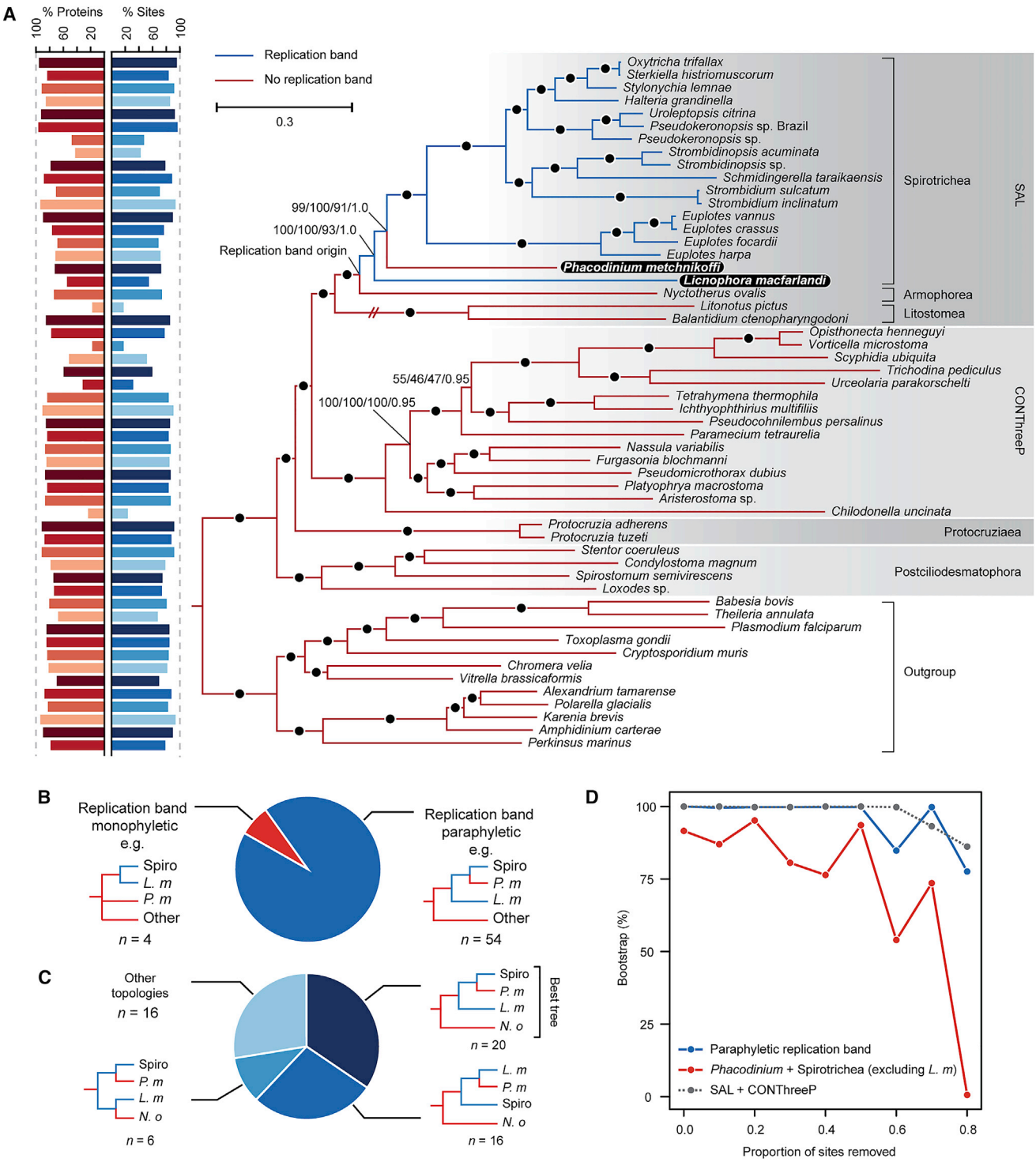

Figure 2. Phylogenomics Resolve the Paraphyly of the RB and the Deep Phylogeny of the Ciliates

(A) Maximum likelihood phylogeny based on 231 concatenated proteins comprising 73,533 amino acid sites generated using the LG+C40+F+G4 substitution model as implemented in IQ-Tree. Statistical support is displayed at the nodes and was assessed using 1,000 ultrafast bootstraps with the LG+C40+F+G4 model, 1,000 ultrafast and non-parametric bootstraps using the LG+F+R8 model, and posterior probabilities determined using the GTR+CAT+G4 model in PhyloBayes. Black circles represent full statistical support. The percentages of proteins and sites present for each taxon are shown on the left, and the presence and absence of the RB is denoted by blue and red branches, respectively. The length of the hashed branch leading to Litostomea was reduced by half for simplicity. *P. metchnikoffi* and *L. macfarlandi* are bolded and highlighted.

(legend continued on next page)

hypotrichs, albeit with poor support, or branching independently in the absence of karyorelictid representation.[34–36] Furthermore, despite previous ambiguity in the topology of SAL,[30,34,36] our analyses recovered Spirotrichea being sister to Armophorea, with Litostomatea branching at the base of the clade with full statistical support (Figure 2A), consistent with recent work.[37] Within SAL, we found that *L. macfarlandi* and *P. metchnikoffi* branched with spirotrichs where they formed the two deepest branches: *L. macfarlandi* branching as sister to all other spirotrichs, and *P. metchnikoffi* branching next as sister to all remaining spirotrichs (Figure 2A).

Based on this topology, the RB would be paraphyletic because of its probable absence in *P. metchnikoffi* (Figure 2A). To assess whether this topology could be a result of gene sampling bias, we sub-sampled the 231 gene set into 5% (12 gene) concatenations by jackknifing (i.e., random sampling without replacement; n = 58) and used the resulting alignments for phylogenetics analysis. The majority of the resulting phylogenies showed a paraphyletic RB (93%), and many recovered the best tree topology (34%), suggesting that these phylogenies are not strongly influenced by gene selection (Figures 2B and 2C). Similarly, sequential removal of fast-evolving sites did not influence RB paraphyly or the sister relationship between SAL and CONThreeP, suggesting that these topologies are not biased by long branch attraction (Figure 2D). However, the exact relationship between *P. metchnikoffi* and spirotrichs aside from *L. macfarlandi* was dependent on gene sampling and site removal, making this placement tentative (Figures 2B–2D) but not in any way consistent with a monophyletic RB.

Characterization of the Replication Envelope in *P. metchnikoffi*

Phylogenomics analyses indicated that the RB had been lost in *P. metchnikoffi* or that previous studies had failed to observe it. To discern between these possibilities, we investigated nuclear morphology in *P. metchnikoffi* using a variety of cytological techniques that have been used to observe the RB in other species. In agreement with previous work, staining with methyl green-rhodamine B (Figure 3A), Feulgen (Figure 3B), and Protargol (Figure 3C) did not reveal an RB in *P. metchnikoffi*. We next sought to visualize replication directly using EdU (5-ethynyl-2′-deoxyuridine) incorporation followed by Cy3-azide labeling. As a positive control, Cy3-EdU labeling and Hoechst staining were applied to the spirotrich *Euplotes* sp. and revealed a fluorescent pattern characteristic of the RB, where the nuclear regions through which the band had passed exhibited Cy3 fluorescence consistent with replicated DNA (Figures 3D and 3E). Cy3-EdU labeling in *P. metchnikoffi* also distinguished replicated DNA but showed an alternative pattern where fluorescence appeared initially throughout the nuclear periphery before progressing inward toward the nuclear interior in a single concerted wave (Figures 3F–3I, Figures 4A–4C; Figure S2). Re-evaluation of the Protargol staining preparations revealed cells with a comparable nuclear structure, suggesting that these observations are independent of staining procedures (Figure 3C). Unlike the RB, the replication envelope lacked defined or stratified structures (i.e., the FZ and RZ were not observed). Similarly, replication appeared to initiate globally at the nuclear envelope, contrasting with the focality observed with the RB. To distinguish this from a canonical RB, we called this a replication envelope.

Transport and Cytoskeleton-Related Genes Are Associated with the Function of the RB and DNA Replication

Given the divergence between the canonical spirotrich RB and the replication envelope of *P. metchnikoffi*, we next sought to leverage RB paraphyly to identify genes that could be associated with its structure and function. In particular, we hypothesized that the differential presence/absence and expansion/duplication rates of protein families between related lineages with and without the RB may suggest functional relatedness. To this end, we clustered proteins from diverse ciliate genomics and transcriptomics datasets into protein families (orthologous groups) and compared the frequency of proteins within each family (Table S1). Of the 37,318 protein families examined, 406 and 23 were found to be enriched or depleted in RB-containing clades, respectively ($p < 0.05$, permutation test, n = 10,000) (Figure 5A; Table S2). Of those, 186 enriched families were found in *L. macfarlandi* and the model spirotrich *O. trifallax*, which had, on average, 1.8 and 2.7 times more homologs per family compared with *P. metchnikoffi*. Consistent with this, principal-component analysis based on differentially present families produced discernible RB-containing and RB-lacking species clusters (Figure 5B), and the RB-association analysis produced distinct results compared with a spirotrich association analysis (Figure S3).

Functional annotation and Gene Ontology (GO) analysis of protein families enriched in RB-containing lineages revealed an association with functions such as ion transport and regulation of the cytoskeleton (Figure 5C; Table S3). For example, biological processes such as cellular ion homeostasis (GO: 0006873; $p < 1 \times 10^{-6}$), ion transport (GO: 0006811; $p = 1 \times 10^{-3.19}$), and regulation of transport (GO: 0051049; $p = 1 \times 10^{-3.10}$) were significantly enriched because of the prevalence of calcium transporters (e.g., polycystin-2, sodium/calcium exchangers, and calcium channels) (Figures 5C and 5D; Figure S4A). Moreover, genes involved in cytoskeletal organization (e.g., GO: 0007010; $p = 1 \times 10^{-2.10}$), including nuclear distribution protein homologs, which regulate nuclear microtubules,[38] and gelsolin repeat proteins, which are calcium-dependent actin regulators,[39] were also correlated with the RB (Figures 5C and 5D; Figure S4B; Table S2). Other functions, including responses to chemicals (GO: 0042221; $p = 1 \times 10^{-2.39}$) and histone modifications (GO: 0016570, $p = 1 \times 10^{-1.30}$) were also enriched because of the presence of proteins containing, for example, cyclic AMP/guanosine monophosphate (GMP) phosphodiesterase domains and acetylated lysine-binding bromodomains, respectively (Figures

(B and C) The proportion of phylogenomic jackknives (12 gene subsamples, n = 58) that recovered a paraphyletic RB (B) and individual tree topologies (C).
(D) Non-parametric bootstrap support (n = 500) for given topologies following removal of the fastest-evolving sites as inferred using the LG+F+R8 model in IQ-Tree.
Spiro, Spirotrichia; *P. m*, *P. metchnikoffi*; *L. m*, *L. macfarlandi*; *N. o*, *Nyctotherus ovalis*. See Figure S1 for morphological details of *L. macfarlandi* and *P. metchnikoffi* and Table S1 for a list of datasets used in the phylogenomics analysis.

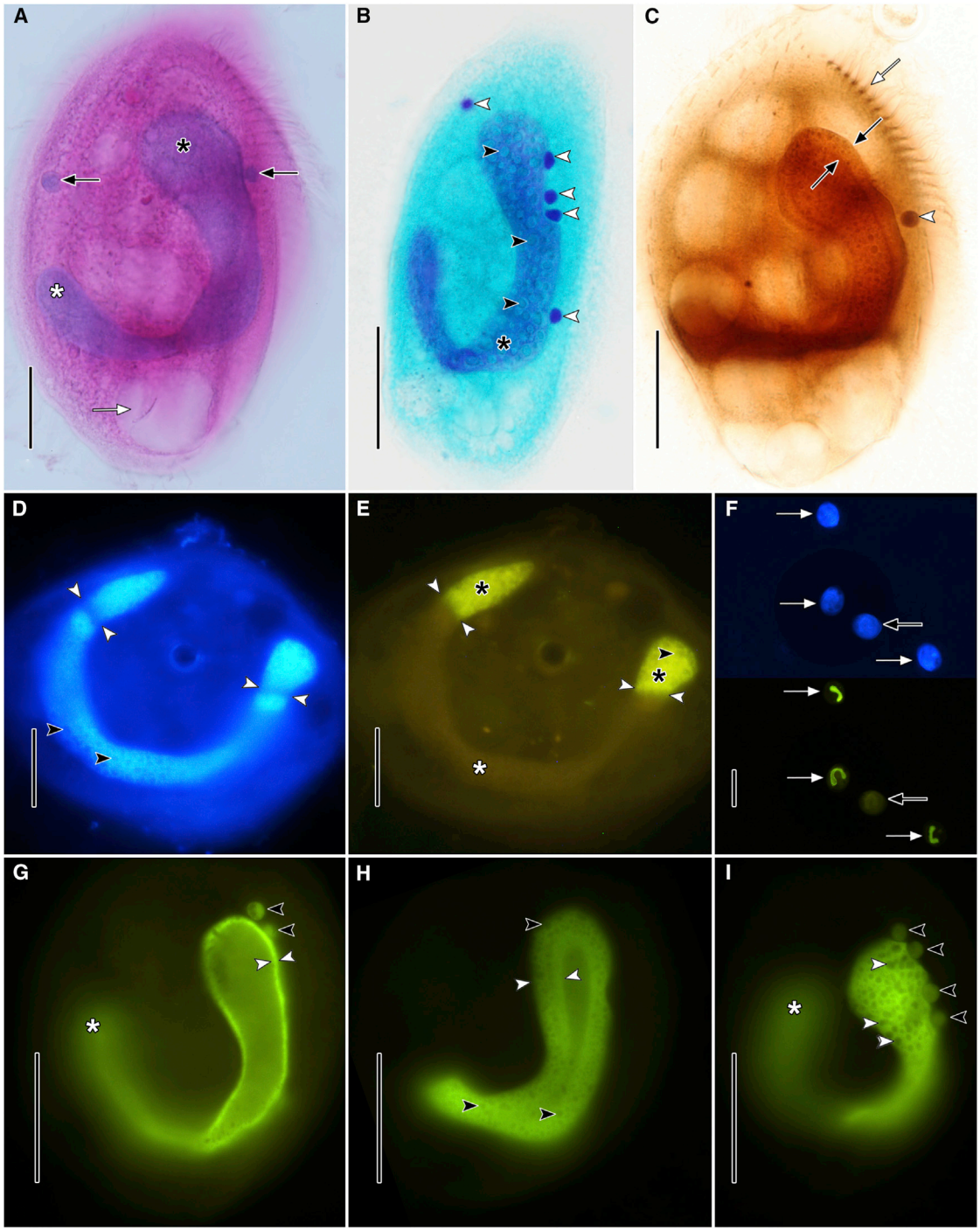

Figure 3. *P. metchnikoffi* Exhibits a Unique DNA Replication Morphology

(A) Methyl green-rhodamine B stain of *P. metchnikoffi*, showing the anterior and posterior lobes of the macronucleus (black and white asterisks, respectively), several micronuclei (black arrows), and the contractile vacuole (white arrow).

(B) *P. metchnikoffi* after Feulgen stain for DNA, showing the macronucleus (asterisk) with many small nucleoli (black arrowheads) and multiple micronuclei (white arrowheads).

(C) *P. metchnikoffi* (ventral view, probable early S phase, cf. Figure 3H) showing a putative replication envelope (black arrows), one micronucleus (white arrowhead), and the adoral zone of ciliary membranelles (white arrow).

(D) *Euplotes* sp. in early S phase stained with Hoechst 33342, showing RBs (white arrowheads) at each end of the macronucleus. Nucleoli (black arrowheads) appear as small lucencies.

(E) Same cell as in Figure 3D after a 13-h pulse with EdU and Cy3-azide labeling. RBs (white arrowheads) advance in front of regions with EdU-labeled replicated DNA (black asterisks). The unreplicated DNA is unlabeled (white asterisk).

(legend continued on next page)

5C and 5D). In agreement with the biological process terms, molecular functions, including transmembrane transporter activity (GO: 0022857; $p = 1 \times 10^{-3.60}$), enzyme binding (GO: 0019899; $p = 1 \times 10^{-2.62}$), cytoskeletal protein binding (GO: 0008092; $p = 1 \times 10^{-1.47}$), and phosphatase activity (GO: 0016791; $p = 1 \times 10^{-1.49}$), were also connected to RB-related proteins (Table S3).

Given the correlative nature of the preceding analysis, the list of RB-associated protein families naturally contains some families that are enriched in RB-containing species but not functionally related to the RB. To investigate which of these RB-associated proteins are required for RB function, we used RNA interference (RNAi)-induced gene knockdown and assessed cell cycle states as well as RB progression and morphology in the model spirotrich *O. trifallax*. We selected 10 of the most highly enriched protein families ($p < 1 \times 10^{-5}$) (Figure 5D; Table S2), most with functions related to significant GO terms, and generated RNAi constructs capable of knocking down expression of all RB-associated paralogs for each group (for a total of 64 proteins knocked down). Compared with the non-specific RNAi, knockdown of a positive control, the essential S phase-initiating protein CDC6, significantly decreased the proportion of cells with RBs (Figure 6A), and when bands were observed, CDC6 RNAi resulted in impaired RB progression (measured as the distance of the RB from its origin divided by the length of the nuclear lobe) and altered RB morphology where the FZ and RZ were difficult to discern (Figures 6B and 6C). This confirmed that at least serious cell cycle and DNA replication defects can be detected by the experimental system. Knockdown of protein families identified as correlating with RBs confirmed that several are required for normal DNA replication and cell cycle progression: protein families 280 (polycystin calcium transporter), 1,006 (sodium/calcium exchanger), and 12,334 (nuclear distribution protein) demonstrated a decrease in the proportion of cells exhibiting RBs (Figure 6A). Knockdown of 280 also resulted in altered RB morphology comparable with knockdown of CDC6, and depletion of 12,334 appeared to impair band progression (Figures 6B and 6C). Knockdown of families 1,008 (EF-hand calcium binding protein) and 2,113 (no annotation) increased the proportion of cells exhibiting RBs (Figure 6A). Depletion of 1,008 also coincided with altered band progression and morphological discrepancies; specifically, the FZ was often thicker and more clearly defined than in the wild type, whereas knockdown of 2,113 had no apparent effect on band movement or appearance (Figures 6B and 6C). Depletion of protein families 259, 482, 1,120, 1,197, and 1,336 did not result in significant phenotypic changes, which might indicate that they are not involved in the RB but could equally reflect technical limitations of the RNAi screen (e.g., insufficient knockdown, dose dependence, redundancy, or lack of precision). Overall, these results confirm the prediction that RB-associated proteins, particularly those comprising families 280, 1,008, 2,113, and 12,334, are important for DNA replication, RB structure and function, or both.

DISCUSSION

Despite the central importance of DNA replication, nuclear genomes exhibit a variety of mechanisms that regulate how the genome is duplicated. However, regardless of the mechanism, the function of the system is conserved, and divergent cases can offer unique perspectives into how genome duplication is achieved. In the case of spirotrich ciliates, the fragmentation of the genome into hundreds of thousands of small, single-gene chromosomes alleviates the common problem of quickly replicating large chromosomes (which generally require multiple replication origins) but elevates the challenge of "bookkeeping." The spirotrich macronucleus contains so many chromosomes that ensuring they are all replicated once and only once likely requires a novel degree of control over the spatial organization of DNA replication, manifested in the RB.

The *P. metchnikoffi* replication envelope shows that motile DNA replication systems are evolutionarily malleable, but it probably evolved from a more canonical RB. This is likely because phylogenomic analyses indicate that the last common ancestor of all spirotrichs, including *P. metchnikoffi*, contained an RB. Moreover, both motile DNA replication systems adhere to the core function of moving the DNA replication zone through the nucleus in a spatially and temporally coordinated fashion so that the replication status of any chromosome is evident simply by its physical location in the nucleus. At a finer scale, other similarities between the RB and envelope are evident. In the RB, the DNA replication hub progresses lengthwise along the axis of the macronucleus as replication proceeds radially from the nuclear envelope toward the interior (Figure 1).[23] Accordingly, the replication envelope may reflect a loss of this lengthwise movement along the axis of the macronucleus. Given the relative simplicity of the replication envelope, it is also tempting to speculate that this system could reflect a reversion to the ancestral state of the RB. However, Cy3-EdU labeling in spirotrich relatives, including armophoreans and litostomateans, which similarly lack RBs but contain nanochromosomes, has been inconclusive, limiting ancestral character reconstruction.

Regardless, identification of the replication envelope in *P. metchnikoffi* provides an opportunity to exploit natural diversity, which may facilitate mechanistic dissection of motile DNA

(F) Low-magnification view of *P. metchnikoffi* cells labeled with Hoechst 33342 (blue, top) and after an 18-h pulse with EdU (green, bottom), showing three cells in various stages of S phase (white arrows) and one non S-phase cell (black arrow).
(G) *P. metchnikoffi* macronucleus (larger anterior lobe) in early S phase showing the very thin peripheral envelope of replicating DNA (white arrowheads) surrounding the relatively lucent central volume of unreplicated DNA. The posterior lobe (asterisk) is out of the focal plane. Micronuclei (black arrowheads) are also beginning to replicate DNA.
(H) Progression of the replication envelope (white arrowheads) toward the center of the macronucleus as the volume of unreplicated DNA diminishes. Nucleoli (black arrowheads) are visible only within regions of replicated DNA.
(I) Completion of macronuclear DNA replication. The replication envelope has disappeared, and nucleoli (white arrowheads) are now distributed throughout the nucleoplasm. Micronuclei (black arrowheads) appear to be completing S phase.
Scale bars: 100 μm (F), 25 μm (A–C and G–I), and 10 μm (D and E). (G)–(I) show three cells from the same slide at 300-ms exposure time. See also Figure S2 for merged fluorescent micrographs of *P. metchnikoffi* during early S phase.

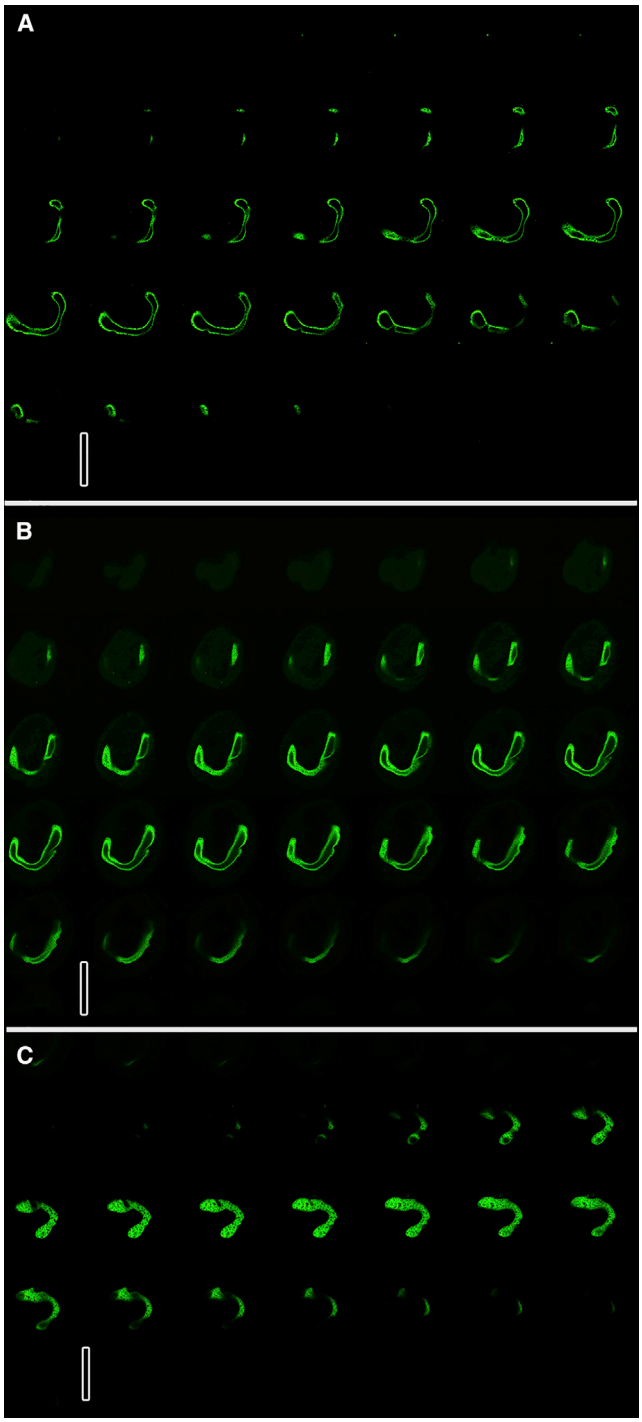

Figure 4. Confocal Fluorescence Imaging of the Replication Envelope in *P. metchnikoffi*

(A–C) Confocal imaging montages of three separate cells.
(A) Early S phase.
(B) Slightly later S phase than in (A).
(C) Completed S phase.
Scale bars, 50 μm; step size, 0.5 μm.

replication systems. Based on this, we identified a number of protein families that were significantly enriched in RB-containing species and hypothesize that many of these proteins may be involved in the lengthwise progression of the RB (as opposed to being involved in motile DNA replication systems in general because of the presence of the replication envelope in *P. metchnikoffi*). RNAi knockdown corroborated this for many proteins by demonstrating their requirement for normal S phase progression, RB movement, or morphology. The overrepresentation within these protein families of calcium transporter paralogs and cytoskeletal regulators, such as calcium-influenced gelsolin repeat proteins and nuclear distribution protein homologs, alludes to an overall mechanism whereby calcium signaling waves trigger translocation of the band along a nuclear cytoskeleton. A role of calcium waves in band progression has indeed been hypothesized previously,[26] and the spirotrich nucleus has also been shown to possess calcium-dependent contractility[40] and RB sensitivity to cyclic AMP (cAMP)-phosphodiesterase and calmodulin inhibitors.[41] Additionally, transmission electron microscopy studies have identified the presence of 10-nm non-chromatin filaments that run through the FZ in *Euplotes eurystomus*.[42] A comparable mechanism was recently described in human fibroblasts, where G-protein-coupled receptor signaling and calcium release trigger linear F-actin polymerization radially from the inner nuclear membrane toward the nuclear interior.[43] Furthermore, nuclear actin dynamics have been implicated in DNA replication control and regulation of other nuclear processes, partly through enzyme relocalization and clustering.[44–47] This not only argues for the feasibility of calcium-dependent motility driving spatial and temporal coordination of a DNA replication system but also suggests that the RB could be a derivation of a broadly conserved eukaryotic nuclear trait. Congruent with this, the protein family enrichment analysis indicated that RB innovation is largely a product of gene duplication.

Ultimately, the spirotrich ciliates provide a unique case from which functional, regulatory, and evolutionary insights into the nature of DNA replication systems can be derived. Although motile DNA replication systems are currently perceived to be a ciliate-specific trait, the flexibility of these systems, as observed in *P. metchnikoffi*, and the mechanistic similarities to processes observed in human cells,[43] suggest that related, perhaps morphologically subtler systems may exist in other eukaryotes. This work also highlights the potential for the RB to serve as a model to not only dissect the consequences of replication timing but the mechanisms underpinning DNA replication and the nuances defining nuclear organization and regulation.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - *Phacodinium metchnikoffi*
  - *Licnophora macfarlandi*
  - *Oxytricha trifallax*

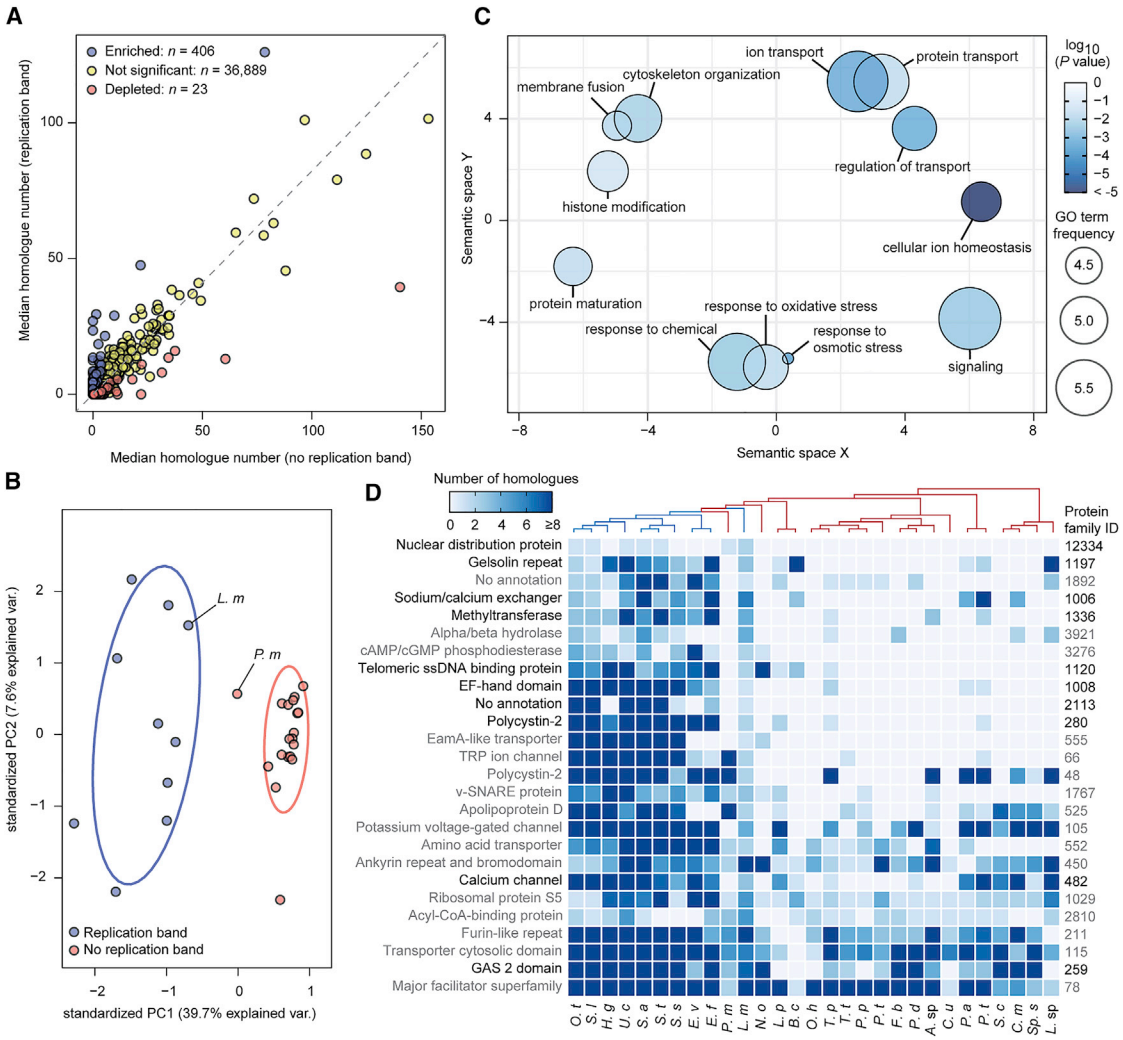

Figure 5. Comparative Genomics Identify RB-Associated Protein Families

(A) Scatterplot depicting the median number of homologs per protein family for clades containing and lacking the RB. Significantly enriched and depleted families are denoted in blue and red, respectively ($p < 0.05$, permutation test, n = 10,000).

(B) Principal-component analysis based on the number of homologs in enriched and depleted protein families for taxa included in the analysis.

(C) Semantic similarity-based scatterplot generated using REVIGO that displays significant gene ontology (GO) terms for RB-associated protein families following summarization. Point color and size are dependent on the p value (permutation test, n = 100,000) and the frequency of the GO term in UniProt.

(D) Heatmap displaying the frequency of the 26 most highly enriched RB-associated protein families ($p < 1 \times 10^{-5}$, permutation test, n = 10,000) across examined taxa that had representation from *L. macfarlandi* and *O. trifallax*. The phylogeny at the top is based on Figure 2, and abbreviated species names can be found at the bottom. Protein identifiers and functional annotations are listed on the right and left, respectively. Protein families highlighted in bold were further examined with RNAi (Figure 6).

See Figure S3 for a comparable analysis identifying spirotrich-associated protein families and Figure S4 for lists and distributions of protein families with GO annotations relating to ion homeostasis and cytoskeletal organization. See also Tables S2 and S3 for the complete set of RB-associated protein families and GO terms, respectively.

- METHOD DETAILS
  - RNA-sequencing and transcriptome assembly
  - Phylogenomic dataset assembly
  - Phylogenomic tree building and analysis
  - Comparative genomics and gene ontology
  - RNAi construct generation and knockdowns
  - Light microscopy and protargol staining
  - Scanning electron microscopy
  - Fluorescence microscopy and EdU labeling
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2020.09.077.

ACKNOWLEDGMENTS

We gratefully acknowledge Laura Landweber, Sheela George, and Talya Yerlici for providing strains of *Oxytricha trifallax* and detailed culturing instructions, Catharine Rankin and Alvaro Luna for supplying HT115 *Escherichia coli* strains and L4440 vectors, Jae-Hyeok Lee and Moyan Jia for providing

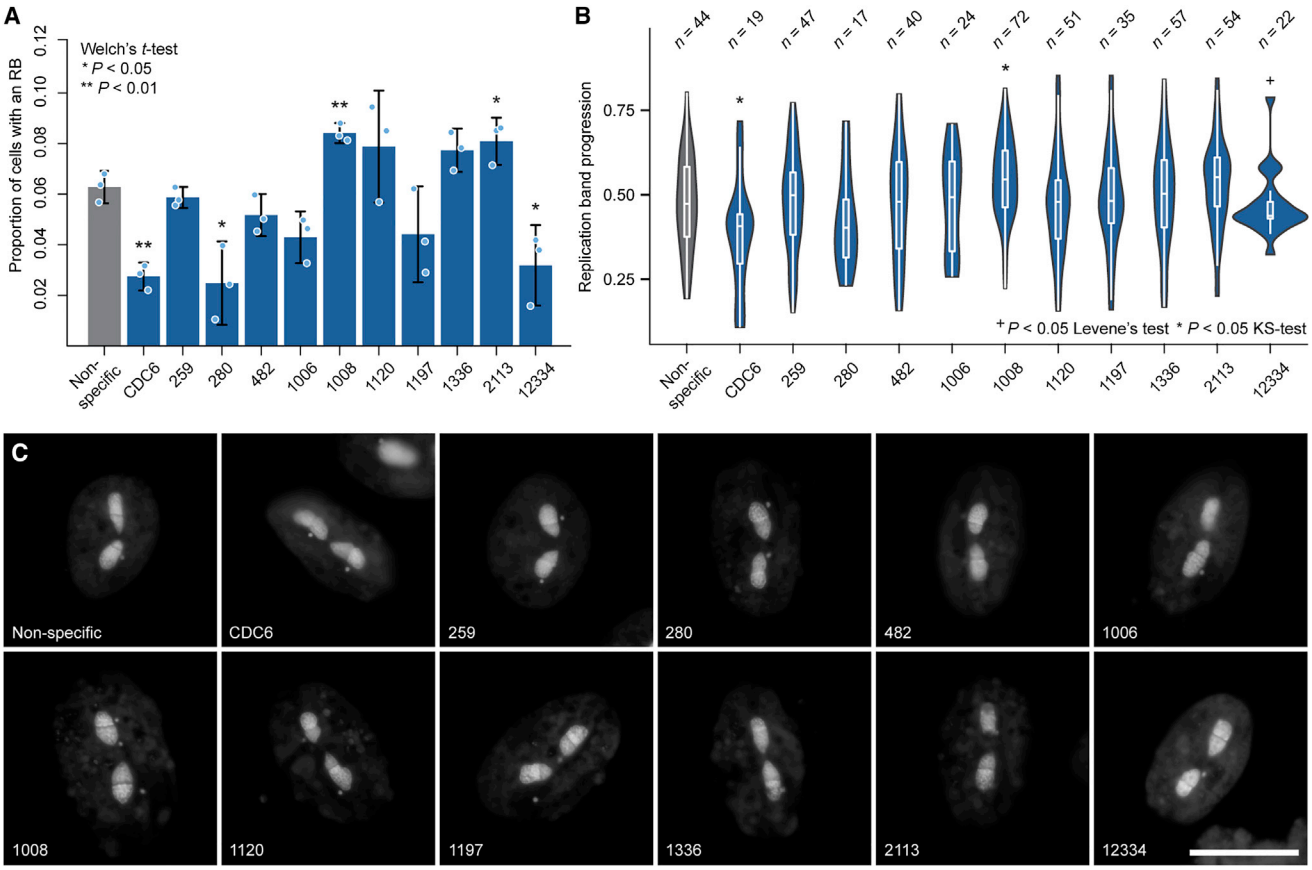

Figure 6. RB-Associated Protein Families Are Required for DNA Replication and Cell Cycle Progression

(A) The proportion of *O. trifallax* cells with RBs following RNAi knockdown of specified protein families (n = 3 biological replicates). For each replicate, an average of 238.7 cells were observed (151–325, standard deviation = 37.4). Error bars represent 95% confidence intervals, and significance was calculated using Welch's t tests.

(B) RB progression in cells following RNAi knockdown of specified protein families. Progression was calculated as the distance of the RB from the nuclear pole divided by the total length of the nuclear lobe. For each RB-containing cell, progression within a single lobe was measured. Significance was assessed using Levene's test and Kolmogorov-Smirnov (KS) tests for variance and distribution equality, respectively.

(C) Fluorescence micrographs taken following DAPI staining, showing representative RB morphology in cells treated with RNAi. Scale bar, 50 μm. The targeted protein families are labeled on the bottom of each plot along with the non-specific control RNAi with which samples were compared. The corresponding protein functions are listed in Figure 5D.

Chlamydomonas strains, Thomas Richards and David Milner for sharing their non-specific RNAi control vector, and Francis Lynn for critically reading the manuscript. We also thank Elizabeth Cooney and Emma George for culturing assistance and Lumiprobe Corp. for the gift of Cy3-azide. This work was supported by grants from the Natural Sciences and Engineering Research Council (NSERC) (RGPIN-2014-03994 to D.H.L. and P.J.K.). N.A.T.I. was supported by an NSERC Canadian graduate scholarship, and A.A.P. was supported by a European Molecular Biology Organization (EMBO) long-term fellowship.

AUTHOR CONTRIBUTIONS

N.A.T.I., D.H.L., and W.A.B. conceived the study. N.A.T.I., A.A.P., V.M., D.H.L., and W.A.B. performed the experiments and analyzed the data. L.J.H., P.J.K., D.H.L., and W.A.B. provided supervision, financial support, and reagents. N.A.T.I., A.A.P., P.J.K., and W.A.B. wrote the manuscript with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 7, 2020
Revised: September 9, 2020
Accepted: September 25, 2020
Published: October 29, 2020

REFERENCES

1. Bleichert, F., Botchan, M.R., and Berger, J.M. (2017). Mechanisms for initiating cellular DNA replication. Science *355*, eaah6317.
2. Fragkos, M., Ganier, O., Coulombe, P., and Méchali, M. (2015). DNA replication origin activation in space and time. Nat. Rev. Mol. Cell Biol. *16*, 360–374.
3. Prioleau, M.-N., and MacAlpine, D.M. (2016). DNA replication origins-where do we begin? Genes Dev. *30*, 1683–1697.
4. Kurat, C.F., Yeeles, J.T.P., Patel, H., Early, A., and Diffley, J.F.X. (2017). Chromatin controls DNA replication origin selection, lagging-strand synthesis, and replication fork rates. Mol. Cell *65*, 117–130.
5. Devbhandari, S., Jiang, J., Kumar, C., Whitehouse, I., and Remus, D. (2017). Chromatin constrains the initiation and elongation of DNA replication. Mol. Cell *65*, 131–141.

6. Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B.J., and Grunstein, M. (2002). Histone acetylation regulates the time of replication origin firing. Mol. Cell *10*, 1223–1233.
7. McNairn, A.J., and Gilbert, D.M. (2003). Epigenomic replication: linking epigenetics to DNA replication. BioEssays *25*, 647–656.
8. Schwaiger, M., Kohler, H., Oakeley, E.J., Stadler, M.B., and Schübeler, D. (2010). Heterochromatin protein 1 (HP1) modulates replication timing of the *Drosophila* genome. Genome Res. *20*, 771–780.
9. Rhind, N., and Gilbert, D.M. (2013). DNA replication timing. Cold Spring Harb. Perspect. Biol. *5*, a010132.
10. Hiratani, I., Ryba, T., Itoh, M., Yokochi, T., Schwaiger, M., Chang, C.W., Lyou, Y., Townes, T.M., Schübeler, D., and Gilbert, D.M. (2008). Global reorganization of replication domains during embryonic stem cell differentiation. PLoS Biol. *6*, e245.
11. Lande-Diner, L., Zhang, J., and Cedar, H. (2009). Shifts in replication timing actively affect histone acetylation during nucleosome reassembly. Mol. Cell *34*, 767–774.
12. Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T.C., Robins, A.J., Dalton, S., and Gilbert, D.M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. Genome Res. *20*, 761–770.
13. Fox, M.H., Arndt-Jovin, D.J., Jovin, T.M., Baumann, P.H., and Robert-Nicoud, M. (1991). Spatial and temporal distribution of DNA replication sites localized by immunofluorescence and confocal microscopy in mouse fibroblasts. J. Cell Sci. *99*, 247–253.
14. Lynn, D.H. (2008). The ciliated protozoa: Characterization, classification, and guide to the literature, Third Edition (Springer).
15. Swart, E.C., Bracht, J.R., Magrini, V., Minx, P., Chen, X., Zhou, Y., Khurana, J.S., Goldman, A.D., Nowacki, M., Schotanus, K., et al. (2013). The *Oxytricha trifallax* macronuclear genome: a complex eukaryotic genome with 16,000 tiny chromosomes. PLoS Biol. *11*, e1001473.
16. Riley, J.L., and Katz, L.A. (2001). Widespread distribution of extensive chromosomal fragmentation in ciliates. Mol. Biol. Evol. *18*, 1372–1377.
17. Olins, D.E., and Olins, A.L. (1994). The replication band of ciliated protozoa. Int. Rev. Cytol. *153*, 137–170.
18. Gall, J.G. (1959). Macronuclear duplication in the ciliated protozoan *Euplotes*. J. Biophys. Biochem. Cytol. *5*, 295–308.
19. Olins, A.L., Olins, D.E., Franke, W.W., Lipps, H.J., and Prescott, D.M. (1981). Stereo-electron microscopy of nuclear structure and replication in ciliated protozoa (Hypotricha). Eur. J. Cell Biol. *25*, 120–130.
20. Olins, D.E., Olinsl, A.L., Herrmann, A., Lin, R., Allis, C.D., and Robert-Nicoud, M. (1991). Localization of acetylated histone H4 in the macronucleus of *Euplotes*. Chromosoma *100*, 377–385.
21. Arikawa, M., Watanabe, A., Watanabe, K., and Suzaki, T. (2000). High-resolution scanning electron microscopy of chromatin bodies and replication bands of isolated macronuclei in the hypotrichous ciliate *Euplotes aediculatus*. Eur. J. Protistol. *36*, 40–45.
22. Fauré-Fremiet, E., Rouiller, C., and Gauchery, M. (1957). La réorganisation macronucléaire chez les Euplotes; étude au microscope électronique. Exp. Cell Res. *12*, 135–144.
23. Postberg, J., Alexandrova, O., Cremer, T., and Lipps, H.J. (2005). Exploiting nuclear duality of ciliates to analyse topological requirements for DNA replication and transcription. J. Cell Sci. *118*, 3973–3983.
24. Olins, D.E., Olins, A.L., Cacheiro, L.H., and Tan, E.M. (1989). Proliferating cell nuclear antigen/cyclin in the ciliate *Euplotes eurystomus*: localization in the replication band and in micronuclei. J. Cell Biol. *109*, 1399–1410.
25. Fang, G., and Cech, T.R. (1995). Telomerase RNA localized in the replication band and spherical subnuclear organelles in hypotrichous ciliates. J. Cell Biol. *130*, 243–253.
26. Jaffe, L.F. (1999). Organization of early development by calcium patterns. BioEssays *21*, 657–667.
27. Balamuth, W. (1941). Microscopic anatomy of *Licnophora macfarlandi*. J. Morphol. *68*, 241–277.
28. da Silva-Neto, I.D., da Silva Paiva, T., Pedroso Dias, R.J., Alexandre Campos, C.J., and Migotto, A.E. (2012). Redescription of *Licnophora chattoni* Villeneuve-Brachon, 1939 (Ciliophora, Spirotrichea), associated with *Zyzzyzus warreni* Calder, 1988 (Cnidaria, Hydrozoa). Eur. J. Protistol. *48*, 48–62.
29. Fernández-Galiano, D., and Calvo, P. (1992). Redescription of *Phacodinium metchnikoffi* (Ciliophora, Hypotrichida): general morphology and taxonomic position. J. Protozool. *39*, 443–448.
30. Gao, F., Warren, A., Zhang, Q., Gong, J., Miao, M., Sun, P., Xu, D., Huang, J., Yi, Z., and Song, W. (2016). The all-data-based evolutionary hypothesis of ciliated protists with a revised classification of the phylum Ciliophora (Eukaryota, Alveolata). Sci. Rep. *6*, 24874.
31. Lynn, D.H., and Strüder-Kypke, M. (2002). Phylogenetic position of *Licnophora, Lechriopyla*, and *Schizocaryum*, three unusual ciliates (phylum Ciliophora) endosymbiotic in echinoderms (phylum Echinodermata). J. Eukaryot. Microbiol. *49*, 460–468.
32. Boscaro, V., Santoferrara, L.F., Zhang, Q., Gentekaki, E., Syberg-Olsen, M.J., Del Campo, J., and Keeling, P.J. (2018). EukRef-Ciliophora: a manually curated, phylogeny-based database of small subunit rRNA gene sequences of ciliates. Environ. Microbiol. *20*, 2218–2230.
33. Shin, M.K., Hwang, U.W., Kim, W., Wright, A.D.G., Krawczyk, C., and Lynn, D.H. (2000). Phylogenetic position of the ciliates *Phacodinium* (Order Phacodiniida) and *Protocruzia* (Subclass Protocruziidia) and systematics of the spirotrich ciliates examined by small subunit ribosomal RNA gene sequences. Eur. J. Protistol. *36*, 293–302.
34. Gentekaki, E., Kolisko, M., Boscaro, V., Bright, K.J., Dini, F., Di Giuseppe, G., Gong, Y., Miceli, C., Modeo, L., Molestina, R.E., et al. (2014). Large-scale phylogenomic analysis reveals the phylogenetic position of the problematic taxon *Protocruzia* and unravels the deep phylogenetic affinities of the ciliate lineages. Mol. Phylogenet. Evol. *78*, 36–42.
35. Chen, X., Zhao, X., Liu, X., Warren, A., Zhao, F., and Miao, M. (2015). Phylogenomics of non-model ciliates based on transcriptomic analyses. Protein Cell *6*, 373–385.
36. Gao, F., and Katz, L.A. (2014). Phylogenomic analyses support the bifurcation of ciliates into two major clades that differ in properties of nuclear division. Mol. Phylogenet. Evol. *70*, 240–243.
37. Rotterová, J., Salomaki, E., Pánek, T., Bourland, W., Žihala, D., Táborský, P., Edgcomb, V.P., Beinart, R.A., Kolísko, M., and Čepička, I. (2020). Genomics of new ciliate lineages provides insight into the evolution of obligate anaerobiosis. Curr. Biol. *30*, 2037–2050.e6.
38. Vergnolle, M.A.S., and Taylor, S.S. (2007). Cenp-F links kinetochores to Ndel1/Nde1/Lis1/dynein microtubule motor complexes. Curr. Biol. *17*, 1173–1179.
39. Sun, H.Q., Yamamoto, M., Mejillano, M., and Yin, H.L. (1999). Gelsolin, a multifunctional actin regulatory protein. J. Biol. Chem. *274*, 33179–33182.
40. Arikawa, M., Momokawa, N., Saito, A., Omura, G., Khan, S.M.M.K., Suetomo, Y., Kakuta, S., and Suzaki, T. (2003). Ca^{2+}-dependent contractility of isolated and demembranated macronuclei in the hypotrichous ciliate *Euplotes aediculatus*. Cell Calcium *33*, 113–117.
41. Olins, D.E., and Olins, A.L. (1993). Inhibition of DNA synthesis in the macronuclear replication band of *Euplotes eurystomus*. J. Eukaryot. Microbiol. *40*, 459–467.
42. Olins, A.L., and Olins, D.E. (1990). Identification of 10 nm non-chromatin filaments in the macronucleus of *Euplotes eurystomus*. Chromosoma *99*, 205–211.
43. Wang, Y., Sherrard, A., Zhao, B., Melak, M., Trautwein, J., Kleinschnitz, E.-M., Tsopoulidis, N., Fackler, O.T., Schwan, C., and Grosse, R. (2019). GPCR-induced calcium transients trigger nuclear actin assembly for chromatin dynamics. Nat. Commun. *10*, 5271.
44. Hurst, V., Shimada, K., and Gasser, S.M. (2019). Nuclear actin and actin-binding proteins in DNA repair. Trends Cell Biol. *29*, 462–476.
45. Parisis, N., Krasinska, L., Harker, B., Urbach, S., Rossignol, M., Camasses, A., Dewar, J., Morin, N., and Fisher, D. (2017). Initiation of

DNA replication requires actin dynamics and formin activity. EMBO J. *36*, 3212–3231.

46. Percipalle, P., and Visa, N. (2006). Molecular functions of nuclear actin in transcription. J. Cell Biol. *172*, 967–971.
47. Wei, M., Fan, X., Ding, M., Li, R., Shao, S., Hou, Y., Meng, S., Tang, F., Li, C., and Sun, Y. (2020). Nuclear actin regulates inducible transcription by enhancing RNA polymerase II clustering. Sci. Adv. *6*, eaay6515.
48. Picelli, S., Faridani, O.R., Björklund, Å.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. *9*, 171–181.
49. Timmons, L., and Fire, A. (1998). Specific interference by ingested dsRNA. Nature *395*, 854.
50. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.
51. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. Nat. Methods *12*, 59–60.
52. Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., and Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. Mol. Biol. Evol. *34*, 2115–2122.
53. Huerta-Cepas, J., Serra, F., and Bork, P. (2016). ETE 3: Reconstruction, analysis, and visualization of phylogenomic data. Mol. Biol. Evol. *33*, 1635–1638.
54. Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
55. Rambaut, A. (2012). FigTree, a graphical viewer of phylogenetic trees. http://tree.bio.ed.ac.uk/software/figtree/.
56. Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005). InterProScan: protein domains identifier. Nucleic Acids Res. *33*, W116-20.
57. Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. *32*, 268–274.
58. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780.
59. Enright, A.J., Van Dongen, S., and Ouzounis, C.A. (2002). An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res. *30*, 1575–1584.
60. Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics *30*, 614–620.
61. Lartillot, N., Lepage, T., and Blanquart, S. (2009). PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics *25*, 2286–2288.
62. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.
63. Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS ONE *6*, e21800.
64. Roure, B., Rodriguez-Ezpeleta, N., and Philippe, H. (2007). SCaFoS: a tool for selection, concatenation and fusion of sequences for phylogenomics. BMC Evol. Biol. *7 (Suppl 1)*, S2.
65. Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. *8*, 1494–1512.
66. Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973.
67. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.
68. Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. *29*, 644–652.
69. da Silva Neto, I.D. (1993). Structural and ultrastructural observations of the ciliate *Phacodinium metchnicoffi* certes, 1891 (Heterotrichea, Phacodiniida). Eur. J. Protistol. *29*, 209–218.
70. Pringsheim, E.G. (1964). Pure cultures of algae, their preparation and maintenance (Hafner Publishing Co.).
71. Bracht, J.R., Perlman, D.H., and Landweber, L.F. (2012). Cytosine methylation and hydroxymethylation mark DNA for elimination in *Oxytricha trifallax*. Genome Biol. *13*, R99.
72. UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. *43*, D204–D212.
73. Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.C., Estreicher, A., Gasteiger, E., Martin, M.J., Michoud, K., O'Donovan, C., Phan, I., et al. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. *31*, 365–370.
74. Burki, F., Kaplan, M., Tikhonenkov, D.V., Zlatogursky, V., Minh, B.Q., Radaykina, L.V., Smirnov, A., Mylnikov, A.P., and Keeling, P.J. (2016). Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista. Proc. Biol. Sci. *283*, 20152802.
75. Mathur, V., Kolísko, M., Hehenberger, E., Irwin, N.A.T., Leander, B.S., Kristmundsson, Á., Freeman, M.A., and Keeling, P.J. (2019). Multiple independent origins of apicomplexan-like parasites. Curr. Biol. *29*, 2936–2941.e5.
76. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods *14*, 587–589.
77. Quang, S., Gascuel, O., and Lartillot, N. (2008). Empirical profile mixture models for phylogenetic reconstruction. Bioinformatics *24*, 2317–2323.
78. Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. *35*, 518–522.
79. El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., et al. (2019). The Pfam protein families database in 2019. Nucleic Acids Res. *47* (D1), D427–D432.
80. Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., Rattei, T., Mende, D.R., Sunagawa, S., Kuhn, M., et al. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. *44* (D1), D286–D293.
81. Harris, M.A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., Eilbeck, K., Lewis, S., Marshall, B., Mungall, C., et al.; Gene Ontology Consortium (2004). The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res. *32*, D258–D261.
82. Paschka, A.G., Jönsson, F., Maier, V., Möllenbeck, M., Paeschke, K., Postberg, J., Rupprecht, S., and Lipps, H.J. (2003). The use of RNAi to analyze gene function in spirotrichous ciliates. Eur. J. Protistol. *39*, 449–454.
83. Pan, X., Bourland, W.A., and Song, W. (2013). Protargol synthesis: an in-house protocol. J. Eukaryot. Microbiol. *60*, 609–614.
84. Foissner, W. (2014). An update of 'basic light and scanning electron microscopic methods for taxonomic studies of ciliated protozoa'. Int. J. Syst. Evol. Microbiol. *64*, 271–292.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
Escherichia coli RNAi feeding strain	*Caenorhabditis* Genetics Center	HT115
Klebsiella sp. for feeding *O. trifallax*	[15]	N/A
Chemicals, Peptides, and Recombinant Proteins		
1,4-Dithiothreitol (DTT)	Invitrogen	Cat#80204
Acetic acid	Fisher Scientific	Cat#A38S-500
Ampure XP beads (for Smart-Seq2)	Beckman Coulter	Cat#A63881
Ascorbic acid ($C_6H_8O_6$)	Sigma	Cat#A0278
Betaine (for Smart-Seq2)	Sigma	Cat#61962
Copper(II) sulfate hydrate ($CuSO_4 \cdot 5H_2O$)	Sigma	Cat#209201
DAPI (4′,6-diamidino-2-phenylindole)	Sigma	Cat#D9542
dNTP mix (for Smart-Seq2)	Fermentas	Cat#R0192
EdU (5-Ethynyl-2′-deoxyuridine)	Lumiprobe	Cat#10540
Ethanol 99.5% (vol/vol) (for Smart-Seq2)	Kemethyl	Cat#SN366915-06
Euparal Microscope Slide Mounting Medium	Hempstead Halide	N/A
First-strand buffer (for Smart-Seq2)	Invitrogen	Cat#18064-014
Glutaraldehyde 25% ($OCHCH_2CH_2CH_2CHO$)	Electron Microscopy Sciences	Cat#16220
Hoechst 3342 (bisBenzimide H 33342 trihydrochloride)	Sigma	Cat. No.14533
IPTG (Isopropyl β- d-1-thiogalactopyranoside)	Sigma	Cat#I6758
KAPA HiFi HotStart ReadyMix (2x) (for Smart-Seq2)	KAPA Biosystems	Cat#KK2601
Magnesium chloride (for Smart-Seq2)	Sigma	Cat#M8266
Methanol	Fisher Scientific	Cat#A412P-4
Osmium tetroxide 2%	Electron Microscopy Sciences	Cat#19152
Phusion High-Fidelity PCR Master Mix with HF Buffer	New England Biolabs	Cat#M0531S
Protargol (silver proteinate) see comment in Method Details	Polysciences	Cat#01070
Recombinant ribonuclease inhibitor (for Smart-Seq2)	Invitrogen	Cat#10777019
SacII	New England Biolabs	Cat#R0157S
Sulfo-Cyanine3 azide	Lumiprobe	Cat#A1330
Superscript II reverse transcriptase (for Smart-Seq2)	Invitrogen	Cat#18064-014
T4 DNA ligase	New England Biolabs	Cat#M0202S
Tris (Tris (Hydroxymethyl) Aminomethane)	Electron Microscopy Sciences	Cat#11720
Triton X-100	Sigma	Cat#x100
UltraPure DNase/RNase-free distilled water (for Smart-Seq2)	Thermofisher	Cat#10977023
XbaI	New England Biolabs	Cat#R0145S
Critical Commercial Assays		
Nextera XT	Illumina	FC-131-1024
Deposited Data		
Phylogenomic dataset	This paper	https://doi.org/10.5061/dryad.zcrjdfn83
Protein family clusters	This paper	https://doi.org/10.5061/dryad.zcrjdfn83
RNA sequencing reads	This paper	NCBI SRA SRR12376981-SRR12376982

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Transcriptome assemblies	This paper	https://doi.org/10.5061/dryad.zcrjdfn83
Experimental Models: Organisms/Strains		
Chlamydomonas reinhardtii CC-125 mt+	https://www.chlamycollection.org/product/cc-125-wild-type-mt-137c/	137c
Licnophora macfarlandi	This paper	N/A
Oxytricha trifallax JRB310	[15]	N/A
Parastichopus californicus	This paper	N/A
Phacodinium metchnikoffi	This paper	N/A
Oligonucleotides		
5′–AAGCAGTGGTATCAAC GCAGAGTACT30VN-3′	[48]	Oligo-dT30VN
5′–AAGCAGTGGTATCAACG CAGAGT–3′	[48]	IS-PCR oligo
5′–AAGCAGTGGTATCAACGCAGAGT ACATrGrG+G–3′	[48]	TSO
Recombinant DNA		
RNAi constructs	This paper	https://doi.org/10.5061/dryad.zcrjdfn83
RNAi expression vector backbone	[49]	Addgene L4440 (#1654)
Software and Algorithms		
BLAST	[50]	http://blast.ncbi.nlm.nih.gov//blast.ncbi.nlm.nih.gov/Blast.cgi
Diamond	[51]	https://github.com/bbuchfink/diamond
eggNOG-Mapper	[52]	https://github.com/eggnogdb/eggnog-mapper
ETE3	[53]	https://github.com/etetoolkit/ete
FastQC	[54]	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
FigTree	[55]	https://beast.community/figtree
Gene ontology enrichment script	This paper	https://doi.org/10.5061/dryad.zcrjdfn83
InterproScan	[56]	https://github.com/ebi-pf-team/interproscan
IQ-Tree	[57]	http://www.iqtree.org/
MAFFT	[58]	https://mafft.cbrc.jp/alignment/software/
Map2Slim	https://github.com/owlcollab/owltools/wiki/Map2Slim	N/A
MCL	[59]	https://micans.org/mcl/
PEAR	[60]	https://github.com/tseemann/PEAR
PhyloBayes	[61]	http://www.atgc-montpellier.fr/phylobayes/
Protein family association script	This paper	https://doi.org/10.5061/dryad.zcrjdfn83
RAxML	[62]	https://cme.h-its.org/exelixis/web/software/raxml/
REVIGO	[63]	http://revigo.irb.hr/index.jsp?error=expired
SCaFoS	[64]	https://megasun.bch.umontreal.ca/Software/scafos/scafos.html
Transdecoder	[65]	https://github.com/TransDecoder/TransDecoder/wiki
trimAl	[66]	http://trimal.cgenomics.org/
Trimommatic	[67]	http://www.usadellab.org/cms/?page=trimmomatic
Trinity	[68]	https://github.com/trinityrnaseq/trinityrnaseq/wiki

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Nicholas Irwin (nickatirwin@gmail.com).

Materials Availability

All materials used in this study including cultures, plasmids, and other reagents are available from the authors upon request.

Data and Code Availability

The accession numbers for the sequence data reported in this paper are Short Read Archive: SRR12376982, SRR1237681. Transcriptome assemblies, phylogenomic datasets, protein family clusters, RNAi construct sequences, and analysis scripts, including those used to conduct protein family association and gene ontology analyses, are available from Dryad (Dryad: https://datadryad.org/stash/dataset/doi:10.5061/dryad.zcrjdfn83) or from the authors upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Phacodinium metchnikoffi

Phacodinium metchnikoffi was isolated from soil percolates from a grass lawn at Julia Davis Municipal Park in Boise, Idaho, USA (43°36'22.7"N 116°11'43.2"W, elev. 816 m) on 22 November 2017. Enriched non-clonal cultures were established in a medium composed of 10% v/v soil extract (1% w/v sterile commercial garden soil boiled in distilled water for 10 minutes, filtered with Grade 102 filter paper after cooling, and then autoclaved and buffered through the addition of 0.5 g K_2HPO_4/L [2.9 mM]) in 0.22 μm filtered site water with a rice grain added to support growth of bacteria as food for the ciliates. Cultures were maintained at room temperature (about 19.4°C) out of direct sunlight and sub-culturing was performed once weekly. The population was verified as *P. metchnikoffi* according to morphological characteristics[69] and a partial 18S rRNA gene sequence that was > 99 % identical to *P. metchnikoffi* (GenBank: AJ277877).[33]

Licnophora macfarlandi

Licnophora macfarlandi was collected from the respiratory trees of the California sea cucumber (*Parastichopus californicus*) as described previously.[31] The body wall of the sea cucumber was opened and the respiratory trees were transferred to a Petri dish and inspected for the presence of *L. macfarlandi*. Cells were flushed from the host tissue using filtered sea water and isolated using a glass micropipette. The identity of *L. macfarlandi* was confirmed based on morphology, the host species, and a partial 18S rRNA gene which was > 99% identical to *L. macfarlandi* (GenBank: AF527758).[31]

Oxytricha trifallax

Oxytricha trifallax (strain JRB310) was cultured in Pringsheim media (0.11 mM Na_2HPO_4, 0.08 mM $MgSO_4$, 0.85 mM $Ca(NO_3)_2$, 0.35 mM KCl, pH 7.0). at 20°C on a 12:12 hour light:dark cycle in disposable Petri dishes.[70,71] Cultures were fed daily with variable amounts of *Chlamydomonas reinhardtii* (typically around 0.05 OD_{600}/mL) that had been washed with Pringsheim media. Every second day, *Klebsiella* sp., that had been cultivated overnight, was added to the cultures (1:1,000 of the culture volume) and the media was transferred to a fresh dish. *Chlamydomonas reinhardtii* (strain CC-125) was cultured in rich media (12 mM $NaCH_3COOH$, 0.1% (w/t) beef extract, 0.2% tryptone, 0.2% yeast extract, and 0.07 mM $CaCl_2 \cdot 2H_2O$), sub-cultured every second day, and fed to *O. trifallax* three days post-inoculation.

METHOD DETAILS

RNA-sequencing and transcriptome assembly

Phacodinium metchnikoffi and *L. macfarlandi* cells were isolated manually using drawn glass micropipettes, washed three successive times in 0.22 μm-filtered water, and placed in 2 μL of cell lysis buffer (0.2% Triton X-100 and RNase inhibitor).[48] RNA was extracted and cDNA was synthesized from pools of 10 cells according to the Smart-Seq2 protocol.[48] cDNA quantification was achieved using a Qubit 2.0 fluorometer and library assembly was conducted using a Nextera XT DNA library preparation kit. Libraries generated from *P. metchnikoffi* and *L. macfarlandi* cDNA were sequenced on an Illumina MiSeq using either 250 or 300 bp paired-end reads.

Raw paired-end Illumina reads were merged using PEAR v0.9.6 and read quality was assessed with FastQC.[54,60] Adapters and low-quality regions were removed using Trimmomatic v0.36 prior to *de novo* transcriptome assembly using Trinity v2.4.0.[67,68] Bacterial contamination was identified and removed by comparing transcriptome contigs to UniProt reference proteomes using Diamond BLASTx (e-value < $1x10^{-25}$).[51,72] Proteins were then predicted with TransDecoder v5.1.0 following BLASTp searches against the SWISS-PROT database.[50,65,73] The genetic codes selected for translation were based on information provided in the NCBI (National Center for Biotechnology Information) Taxonomy database (http://www.ncbi.nlm.nih.gov/Taxonomy).

Phylogenomic dataset assembly

In order to generate a phylogenomic dataset encompassing ciliate diversity, ciliate transcriptomes and genomes (Table S1) were searched for a set of 263 homologs that have been used previously to reconstruct the phylogeny of alveolates and eukaryotes in general.[74,75] Previously curated protein sequences, derived from a diversity of eukaryotes, were used as queries for BLASTp searches (e-value < $1x10^{-20}$, query coverage > 50%) against databases of transcriptomic and genomic protein predictions.[74] The resulting hits were combined with the original protein sequences and single gene phylogenies were generated in RAxML v8.2.11 using the PROTGAMMALG model and 100 rapid bootstraps following alignment with MAFFT L-INS-i v7.222 and subsequent trimming with trimAl v1.2 using a gap threshold of 80%.[58,62,66] Paralogs and contaminants were identified and removed manually after visually inspecting the phylogenies in FigTree v1.4.2.[55] The curated alignments were then filtered to remove those with less than 60% of species present and the remaining 231 alignments were concatenated in SCaFoS v1.2.5.[64] The final concatenation comprised 231 proteins spanning 73,533 amino acid sites from 52 taxa.

Phylogenomic tree building and analysis

Maximum likelihood (ML) phylogenies were generated from the concatenated alignment in IQ-Tree v1.5.5 using either the free-rate heterogeneity model, LG+F+R8, as selected based on Bayesian information criteria in ModelFinder, or the empirical profile mixture model, LG+C60+F+G4.[57,76,77] Statistical support was assessed using both non-parametric bootstrap (n = 1,000, LG+F+R8) and ultrafast bootstrap pseudoreplicates (n = 1,000, LG+F+R8 and LG+C60+F+G4).[57,78] Bayesian analyses were conducted using PhyloBayes MPI v1.4 using the GTR matrix with the CAT infinite mixture model and four gamma rate categories after the removal of constant sites.[61] Four MCMC chains were run for at least 10,000 generations with every second tree saved. Chain convergence was assessed in PhyloBayes using a burn-in of 20%.

Topology robustness was assessed using alignment jackknifing and fast site removal to test for gene sampling bias and long branch attraction, respectively. Proteins comprising the concatenated alignment were randomly sampled without replacement to generate 12 gene concatenations (5% of the 231 genes, n = 58) from which ML phylogenies were generated using IQ-Tree v1.5.5 and the LG+F+R8 substitution model.[57] Topologies were inspected in FigTree v1.4 and the frequencies of recovering a paraphyletic replication band, the best tree, and other topological variations were assessed.[55] Moreover, site-specific rates were calculated across the 231 gene alignment using IQ-Tree v1.5.5 and the fastest sites were removed in 10% increments. Following fast site removal, ML phylogenies and non-parametric bootstraps (n = 500) were calculated in IQ-Tree v1.5.5.[57] Bootstrap support for the paraphyly of the replication band and the sister relationship between SAL and CONThreeP as well as *P. metchnikoffi* and previously sampled spirotrichs was assessed using ETE3.[53]

Comparative genomics and gene ontology

To identify RB-associated homologs, we first compared all proteins predicted from ciliate transcriptomes and genomes (Table S1) to one another using BLASTp (e-value < $1x10^{-5}$, query coverage[3] 50%).[50] With the resulting comparison, we clustered proteins into orthologous groups (protein families) using the Markov clustering (MCL) algorithm with an inflation value of 1.4.[59] Clusters only containing proteins from a single species were discarded. To identify protein families that were significantly associated with the RB, we calculated the median number of homologs within each protein family for clades with (Spirotrichs, *L. macfarlandi*) and without the RB (*P. metchnikoffi*, *Nyctotherus ovalis*, Litostomatids, CONThreeP, Protocruziaea, and Postciliodesmatophora), and then in RB containing and lacking groups, generally. For each protein family, the difference in the median count between groups with and without the RB was then examined, and significance was assessed using two-sided permutation tests with a test distribution generated by randomizing the families 10,000 times by randomly sampling ciliate proteins without replacement. Proteins within enriched families ($p < 0.05$), that had representation from *O. trifallax* and *L. macfarlandi*, were subsequently annotated using InterproScan v5.36, eggNOG-Mapper v2, and BLASTp (e-value < $1x10^{-5}$, maximum target sequences = 1) in combination with Pfam, EggNOG, and SWISS-PROT databases, respectively.[50,52,56,73,79,80] Protein annotation was achieved by assigning the predominant annotation from the constituent proteins to the group as a whole. Principal component analysis was conducted in R using differentially present protein families that had representation in *O. trifallax* and *L. macfarlandi*. Importantly, these RB-associated protein families are only correlated with the RB and must be experimentally validated to demonstrate a functional connection with the RB itself (see the RNAi analyses detailed below).

To explore functional enrichments among RB-associated proteins, gene ontology (GO) terms were assigned to individual proteins using eggNOG-Mapper v2.[52,81] GO annotations connected to individual proteins were then related to the protein's respective family before being simplified by mapping the terms to the *Saccharomyces cerevisiae* GO-slim subset using Map2Slim (see https://github.com/owlcollab/owltools/wiki/Map2Slim). GO enrichment was assessed by comparing the frequency of the terms in RB-associated families (that contained sequences from *O.trifallax* and *L. macfarlandi*) to their frequency across all ciliate protein families. *P*-values were calculated using permutation tests wherein the test distributions were produced by generating 100,000 randomized, equally sized, test sets through random selection of families without replacement. Significantly enriched GO terms ($p < 0.05$) were summarized and visualized using REVIGO.[63]

RNAi construct generation and knockdowns

RNAi plasmids were assembled using L4440 vectors and synthesized RNAi constructs.[49] RNAi inserts comprised variable numbers of tandem 100-150 bp fragments with homology to RB-associated paralogs in *O. trifallax*. Fragment specificity was confirmed using

BLASTn against *O.trifallax* coding sequences (no identical matches over 17 bp) and inserts were synthesized by Integrated DNA Technologies. RNAi constructs capable of knocking down protein families CDC6 (2 paralogs), 259 (12), 280 (11), 482 (13), 1006 (3), 1008 (8), 1120 (6), 1197 (2), 1336 (3), 2113 (9), and 12334 (1) were synthesized (Table S2). These families were selected as they had association *P*-values < $1x10^{-5}$ and functions related to enriched gene ontology terms (Table S2). The RNAi constructs were amplified by polymerase chain reaction (PCR) using a high fidelity Phusion polymerase (New England BioLabs, NEB) and SacII and XbaI restriction sites were integrated at the 5′ and 3′ ends, respectively. RNAi constructs were inserted into the L4440 multiple cloning site following digestion with SacII and XbaI and ligation using T4 DNA ligase (NEB). The sequence composition of the resulting RNAi plasmids was confirmed by Sanger sequencing conducted by GeneWiz.

RNAi-expressing bacteria were produced as previously described.[82] Assembled plasmids were initially transformed into HT115 *Escherichia coli* and transformants were inoculated into lysogeny broth (LB) supplemented with ampicilin (100 μg/mL) and tetracycline (10 μg/mL), and grown over night at 37°C. Inoculants were then diluted to 0.02 OD_{600} in 2xYP media with antibiotics[49] and cultures were incubated at 37°C until an OD_{600} of 0.4 was reached. RNAi-expression was then induced through the addition of 0.4 mM IPTG (Isopropyl β- d-1-thiogalactopyranoside) and induction was allowed to proceed over four hours. Cultures were harvested by centrifugation (3000xg for 10 minutes) and washed twice with Pringsheim media before being resuspended to an OD_{600} of 4.0 in Pringsheim media containing 2% glycerol. Prior to storage at −20°C, *E. coli* were heat killed following a five minute incubation at 65°C.

Feeding-based RNAi knockdowns were performed on *O. trifallax* by diluting stationary cultures to 2,000 cells/mL before growing them as described above with a reduced amount of *C. reinhardtii* (∼95% reduction) and the addition of RNAi expressing *E. coli*. Cultures were slowly acclimatized to the presence of *E. coli* over three days by sequentially increasing the amount of *E. coli* in the cultures daily from an OD_{600} of 0.02, to 0.035, and finally to 0.05. After three days of pre-treatment, the cultures were again diluted to 2,000 cells/mL and each subsequent day, RNAi-expressing *E. coli* was added to 0.05 OD_{600}. After 36 hours, during log-phase growth, *O. trifallax* cultures were collected for RB analysis. Cultures were concentrated by centrifugation at 200xg for 2 minutes prior to the addition of 0.2 μg/mL DAPI (4′,6-diamidino-2-phenylindole). After a 10 minute incubation, the cells were fixed with 10% 3:1 methanol:acetic acid and were stored at room temperature in the dark. Replication band presence and morphology were inspected using a Zeiss Axioplan 2 microscope fitted with an X-Cite120LED fluorescence illuminator and micrographs were acquired using a Sony A7RIII digital camera attached with an LMScope digital SLR universal adaptor. Replication band progression was measured in ImageJ and was calculated as the distance between the nuclear pole and the RB divided by the total length of the nuclear lobe. Band progression was measured in one nuclear lobe from each RB-containing cell.

Light microscopy and protargol staining

The morphology of living and protargol-impregnated cells of *Phacodinium metchnikoffi* (see Figure S1E–N) were examined under an Olympus BX51 microscope equipped with a Canon EOS 6 digital camera and cell measurements were made using an ocular micrometer. Differential interference contrast (DIC) and brightfield illumination (BF) were used to observe living and protargol-impregnated cells, respectively. For protargol staining, the following protocol was employed: Cells were selected with micropipettes under the dissecting microscope, fixed in 10% v/v formalin (final concentration about 4 –5%) for 15 min, washed with tap water, placed in a few drops of 1% Mayer's albumin (1:1 egg albumin:glycerol), dropped on glass slides, and air-dried for at least 2 h. Dried slides were submerged in 95% isopropanol, 70% isopropanol, and tap water in staining jars for 5 min each. Slides were then placed in 0.2% potassium permanganate for 2 min, washed in tap water for 30 s, followed by 3 min in 2.5% oxalic acid then transferred to tap water for two 3 min soaks followed by 3 min in distilled water. Washed slides were then placed in pre-warmed 0.4% protargol solution (Polysciences) at 60°C in a warming oven for 20 min (Note: Since 2013, protargol effective in the impregnation of ciliates is no longer commercially available. Commercial vendors may still offer ''Strong Protargol'' or ''Protargol-S'' which are ineffective for the demonstration of the ciliature. If stocks of an effective older product are unavailable, a protocol for synthesis of protargol effective for ciliate impregnation has been published[83]). Dieckmann's acetone developer was added to several drops of warm protargol on the slides to differentiate the ciliature and development was controlled under a compound microscope. When sufficiently developed, the slides were promptly washed in tap water for 15 s and then placed in 2.5% $Na_2S_2O_3$ for 3 s before being washed for 5 min in tap water, dehydrated for 10 min in both 70% and 100% isopropanol, and finally mounted in euparal (Hempstead Halide Inc.).

Scanning electron microscopy

For scanning electron microscopy, cells were selected under the dissecting microscope, fixed for 30–60 min (1% osmium tetroxide and 2.5% glutaraldehyde), and thoroughly washed at least three times in tap water under the dissecting microscope to remove debris. Cells were then placed in a small drop on 10 μm plankton net filters in custom-made brass chambers.[84] Cell containing chambers were subsequently transferred through a graded ethanol series (30%, 50%, 70%, 95%, 100%, 100% for 5 min each) for dehydration, and then processed in an EMS 850 critical point dryer. Dried cells were scattered onto adhesive carbon tabs on aluminum and sputtered with gold (argon pressure 0.15 Pa, energy 6–8 mA, ten 90 s sputtering exposures between 60 s rests) in a manual sputter coater (Agar Scientific). Processed samples (see Figure S1O–S) were examined using an SU3500 scanning electron microscope (Hitachi, USA) at an accelerating voltage of 10 kV in the secondary electron imaging mode.

Fluorescence microscopy and EdU labeling

To label cells with EdU (Invitrogen), cells were grown in 3 mL of medium to log phase in sterile 10 × 35 mm polystyrene culture dishes (Corning), 1 μL/mL of EdU (20mM) was added to the dish (final EdU conc. 20μM) and dishes were left for the desired interval (2–48 hr) before selecting cells under the stereo microscope for labeling with sulfo-Cy3-azide (Lumioprobe). Cells were collected in the wells of glass depression-slides before being washed three times in 0.22 μm-filtered EdU-free medium. Washed cells were then fixed for 15 min in 10% formalin, washed once in 100 mM Tris buffer, and then placed in fresh Tris for 5 min. Tris was removed and replaced with 0.01% Triton X-100 for 10 min to permeabalize the cells. The cells were then washed three times in tap water and the Cy3-labeling mixture was prepared (See notes below). Water was removed leaving cells in as small a volume as possible, before the cells were incubated in the dark for 30 min after the addition of 200 μL of labeling solution. Following the incubation, cells were washed three times and then stained with 0.5 μg/mL Hoechst 33342 (Sigma Aldrich) for 5 min. After staining, the cells were washed three times in tap water, and several drops of 1% Mayer's albumin were added before the cells were dropped onto clean glass slides and air-dried for at least 2 hr in the dark. Lastly, slides were dehydrated and mounted as in the protargol staining protocol. Slides were examined with a Zeiss Aksioskop 2 plus epifluorescence microscope with a 100x oil immersion objective (N.A. 1.3) and Zeiss filter sets 20 HE and 01 for Cy3 and Hoechst 3342, respectively. Images were taken with a Flex CCD camera (Diagnostic Instruments). Confocal fluorescent microscopy was done using a Zeiss LSM 510 Meta confocal imaging system using a rhodamine laser line (excitation 543 nm/emission 560 nm) and a Plan-Apochoromat 63X (N.A. 1.4) oil immersion objective (X-Y voxel size 0.8 μm, Z voxel size 0.5 μm).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical support for phylogenies was obtained using 1,000 non-parametric bootstraps (using the LG+F+R8 substitution model), 1,000 ultrafast bootstraps (using the LG+C60+F+G4 and LG+F+R8 models), and Bayesian posterior probabilities calculated after at least 10,000 generations (GTR+CAT model, four chains, chain bipartition discrepancies: max differencert = 0.175857, mean difference = 8.37×10^{-4}) (Figure 2A). Non-parametric bootstrap support for clades following fast site removal was calculated from 500 non parametric bootstraps using the LG+F+R8 substitution model (Figure 2D). Statistical support for the correlation between protein families and defined phenotypes was assessed using permutation tests and 10,000 permutations (Figure 5A; Figure S3A) whereas gene ontology enrichment was assessed using 100,000 permutations (Figure 5C; Figures S3C and S3D). Principal component analyses were conducted in R (Figure 5B; Figure S3B). Comparisons between replication band prevalence following RNA interference were made using Welch's t tests (n = 3, where each biological replicate represents the proportion of cells with replication bands assessed from an average of 238.7 cells (151-325, standard deviation = 37.4)) (Figure 6A). Replication band progression (Figure 6B) was compared between RNA interference treatments using both Levene's tests and Kolmogorov-Smirnov (KS) tests for variance and distribution equality, respectively (sample size is denoted in Figure 6B, where each replicate represents a single measured cell. Cell measurements were pooled from three biological replicates (see Figure 6A)).