Lia1p, a Novel Protein Required during Nuclear Differentiation for Genome-Wide DNA Rearrangements in \textit{Tetrahymena thermophila}

Charles H. Rexer and Douglas L. Chalker*

\textit{Department of Biology, Campus Box 1137, Washington University, St. Louis, Missouri 63130}

Received 2 May 2007/Accepted 11 June 2007

Extensive genome-wide rearrangements occur during somatic macronuclear development in \textit{Tetrahymena thermophila}. These events are guided by RNA interference-directed chromatin modification including histone H3 lysine 9 methylation, which marks specific germ line-limited internal eliminated sequences (IESs) for excision. Several genes putatively involved in these developmental genome rearrangements were identified based on their proteins’ localization to differentiating somatic nuclei, and here we demonstrate that one, \textit{LIA1}, encodes a novel protein that is an essential component of the genome rearrangement machinery. A green fluorescent protein-Lia1 fusion protein exhibited dynamic nuclear localization during development that has striking similarity to that of the dual chromodomain-containing DNA rearrangement protein, Pdd1p. Coimmunoprecipitation experiments showed that Lia1p associates with Pdd1p and IES chromatin during macro-nuclear development. Cell lines in which we disrupted both the germ line and somatic copies of \textit{LIA1} (\textit{ΔLIA1}) grew normally but were unable to generate viable progeny, arresting late in development just prior to returning to vegetative growth. These mutant lines failed to properly form Pdd1p-containing nuclear structures and eliminate IESs despite showing normal levels of H3K9 methylation. These data indicate that Lia1p is required late in conjugation for the reorganization of the \textit{Tetrahymena} genome.

Differentiation of the somatic genome during development from a totipotent germ line involves extensive genetic reprogramming. Genes are activated or silenced as necessary, and chromosomal domains are established to enforce these actions. The unique nuclear dualism of the ciliated protozoan \textit{Tetrahymena thermophila} provides an ideal context in which to examine such chromosomal reorganization during development. This organism, like other ciliates, contains two structurally and functionally distinct nuclei within the cytoplasm of a single cell (see reference 30). The diploid micronucleus, while transcriptionally silent during vegetative growth, harbors the germ line genome for genetic propagation during the next sexual generation (14). The polyploid macronucleus contains 45 to 50 copies of the genome and is the site of all somatic gene expression.

During the vegetative life cycle, the micronucleus divides mitotically, while the polyploid macronucleus divides via a poorly understood amitotic mechanism (reviewed in reference 18). It is during the sexual stages that these different nuclei are formed. Conjugation of cells of complementary mating types induces both meiosis of micronuclei and loss of parental somatic macronuclei and the genomes within. The zygotic genome, the combination of meiotic products derived from the micronucleus of each mating partner, then replicates and undergoes differentiation to give rise to the new germ line and somatic nuclear precursors. The new germ line genome is preserved transcriptionally silent within micronuclei while the new somatic genome becomes transcriptionally active within developing macronuclear anlagen. Specific chromatin modifications are established within the macronuclear chromosomes that are believed to enforce proper genetic regulation (reviewed in reference 23).

In addition to the establishment of the active chromosomes, the developing somatic genome undergoes DNA rearrangements on a massive scale (reviewed in reference 40). Two major types of programmed genome rearrangements occur within the macronuclear anlagen during conjugation, resulting in the elimination of \~15% of the germ line-derived DNA (41). Chromosomal breakage at \~280 sites, coupled with de novo addition of telomeres, fragments the five original chromosomes. This occurs at a highly conserved 15-bp chromosome breakage sequence, creating macronuclear chromosomes that range in size from 20 kbp to \textgreater 2 Mbp (12, 16, 42). Furthermore, specific DNA deletion coordinately eliminates an estimated 6,000 DNA segments from the developing macronuclear chromosomes (39). These excised segments are known as internal eliminated sequences (IESs) and vary in length from 0.6 to greater than 20 kbp. Upon the completion of chromosome breakage and IES excision, the remaining DNA is amplified to the final polyploid state (45C) in the newly formed macronucleus.

Given the abundance and diversity of IESs, it has been challenging to understand how \textit{Tetrahymena} targets these germ line-limited DNAs for elimination, as they possess no common structure nor contain identifiable sequence motifs (see reference 40). Recent research has provided important insight by uncovering an RNA interference (RNAi)-related mechanism that is critical for IES excision (reviewed in references 23, 27, and 38). In meiotic micronuclei, bidirectional transcripts of IESs form double-stranded RNA and are processed into small RNAs (called scan RNAs) by a Dicer-like protein, Dcl1p (9, 22, 25). In the cytoplasm the scan RNAs assemble into a protein complex containing Tw1p, a member of the PPD (PAZ-Piwi domain)/Argonaute protein family, whose mem-
bers are key actors in the RNA-induced silencing complexes of RNAi-related processes. Twi1p is required for both stabilization of scan RNAs and DNA elimination (24). As conjugation progresses, these putative scan RNA-containing RNA-induced silencing complexes search the newly developed macronuclear genome to identify the homologous micronuclear-derived sequence. An interaction at homologous loci establishes histone H3 lysine 9 (H3K9) methylation on the targeted sequences (35). Thus, IES recognition appears to be mechanistically related to the establishment of heterochromatin domains in other eukaryotes.

The proteins that comprise the DNA rearrangement machinery are incompletely understood. The two proteins discussed above, Dcl1p and Twi1p, are clearly involved in the early events of RNAi-directed IES recognition. Three additional proteins, Pdd1p, Pdd2p, and Pdd3p (for programmed DNA degradation proteins), are key players in the DNA deletion process (20, 21, 28, 33, 34). Pdd1p is the most extensively studied of these and has roles in the initial recognition of IESs and likely in their removal as well. All three Pdd proteins are found in macronuclear anlagen, where they associate with proteinaceous structures containing both H3K9 methylated chromatin and IESs (20, 28, 33). Gene knockout studies have shown that Pdd1p and Pdd2p are essential for genome rearrangement. In both cases, knockout of the copies from the parental macronucleus is sufficient to cause the DNA rearrangement defect (10, 29). Pdd1p contains two chromodomains and Pdd3p contains one chromodomain, all of which have been shown to specifically bind to methylated H3K9 peptides (35). The chromodomains of Pdd1p and Pdd3p associate with the modified chromatin on IESs and likely recruit additional components of the genome rearrangement machinery.

These genome rearrangements also appear to involve large-scale chromosome reorganization. After the IES chromatin is modified and bound by Pdd proteins, these sequences appear to coalesce into distinct nuclear foci (20, 33). By the time that IES excision occurs, thousands of loci are concentrated into dozens of Pddp-containing structures. Thus, given the dynamic nature of these events, it is likely that more than the three identified Pdd proteins are needed to carry out the removal of the 15 to 20 Mbp of germ line-limited DNA from the developing somatic macronucleus. We recently developed a cytological screen to identify additional participants in this process by virtue of their localization within these DNA rearrangement foci (37). The proteins identified, encoded by the LIA (localized in macronuclear anlagen) genes, were found to be primarily novel proteins that were expressed exclusively during conjugation. Transcription of each gene was highly induced very near the time that the anlagen first emerge.

In this report, we describe detailed studies of Lia1p. Analysis of a green fluorescent protein (GFP)-Lia1p fusion revealed that this protein exhibits dynamic redistribution within macronuclear anlagen during conjugation similar to that of Pdd1p. Immunoblot and chromatin immunoprecipitation (ChIP) experiments demonstrate that Lia1p expression is restricted to stages of Tetrahymena development when genome rearrangements occur. Furthermore, the Lia1 protein forms part of a common complex with Pdd1p enriched for specific germ line-limited IESs targeted for elimination but not with macronuclear location-destined sequences. Although disruption of Lia1 leads to failure of DNA rearrangement, H3K9 methylation occurs normally. These data indicate that Lia1p is a novel component of the protein machinery that associates with modified IES chromatin, leading to its removal from the developing macronuclear genome.

**MATERIALS AND METHODS**

**Strains and culture conditions.** All Tetrahymena strains were maintained as described previously (1). Standard wild-type laboratory strains B2086 (II), CU428 (mpr1-1/mpr1-1 [VII, mp-s]), and CU427 (htsh1-1/htsh1-1 [VI, cy-s]) and micronuclear-defective strains B*VI (VI) and B*VII (VII), originally obtained from Peter J. Bruns (Cornell University, Ithaca, NY), or their transformed progeny were used in all experiments described. Cell growth was carried out in 1× SPP or 1× Neff medium at 30°C. Strains were starved overnight in 10 mM Tris-HCl (pH 7.4) prior to mating. Conjugation was induced by mixing populations of complementary mating types at equal cell numbers.

**Cloning of the LIA1 gene.** We initially cloned a partial LIA cDNA, GenBank accession no. EF219411 (37). The entire LIA1 coding sequence plus flanking genomic DNA was cloned using modified vectorate PCR (31). Tetrahymena genomic DNA was digested with BglII, Xhol, or Sau3A ligated to the indicated oligonucleotide (VKT7r, 5′-TCTCTTCTCTGTGACAGCCGCAAGCTCGGAAATACCCCTCTCTGCC-3′) that had been annealed to one of the following partner oligonucleotides creating 4-bp, 5′ overhangs that are compatible “sticky ends” to join with cut genomic DNA (VKT7-Xba, 5′-CTAGGCGCAGAAGACACGCAGGTACGATAGCT-3′, or VKT7-Bam, 5′-GATCGCAGGAACGGCGCAGGGGTTATACGACAGGAGG-3′). Ligation products were used as templates in PCRs using nested LIA1-specific primers and nested vectorate anchor primers, Bubl3 (5′-CCCTGACAGACCGCAAGCCTTGGA-3′) and T3 (5′-CGCCAAGCTGCTAATTACCAACTATTATG-3′). Amplified PCR products were then cloned into pCR2.1 using the dye terminators (Applied Biosystems). Coding region fragments were amplified from genomic DNA or random-primed cDNA using oligonucleotides 5′-AACCTCGAGATGACAGGAAAAAATGACGACAGGAGGAGGAGGAGG-3′ and 5′-AAAGGGGCAAATATTTAATTATATTGTTGC-3′, which added an Xhol site immediately upstream of the ATG start codon and an Apal site to the downstream end of the PCR fragment. After TA cloning of the amplified product into pCR2.1, these restriction sites were used to fuse the cDNA downstream of, and in frame with, GFP in pCRF-1 (37). Subsequently, LIA1 coding sequence amplified from genomic DNA was cloned into pENTR-D (Invitrogen) and recombined into pGF-gtw to create an inducible GFP-Lia1 fusion under the control of the MTT1 promoter (37).

**Construction of the LIA1 germ line knockout.** The LIA1 targeting knockout construct consisted 1.1 kb of upstream genomic sequence (CH454662; nucleotide [nt] 333246 to nt 334349) and 0.9 kbp of downstream sequence (CH454662; nt 335672 to nt 336561) flanking each side of the NEO3 cassette (32); the bacterial resistance gene, which confers paromomycin resistance (pmr1-1), was tagged with the cadmium (Cd2+)-inducible MTT1 promoter. The LIA1 knockout construct, linearized by Xhol and SacI digestion, was used to coat 0.6 mM gold particles and introduced into mating B2086 and CU428 cells between 2.5 and 3.5 h postmating by particle bombardment as described elsewhere (6, 7). Transfectants with incomplete replacement of macronuclear Lia1 copies with the Δlia1 allele were subcloned into medium containing 1 μg/ml G418 and increasingly higher concentrations of paromomycin until all macronuclear chromosomes contained the mutant allele due to phenotypic assortment. These heterozygous (nia1::neo3/Δlia1 [Cd2+::pm-r]) transfomers were crossed to star strains B*VI and B*VII to generate one homokaryon Δlia1 strain (nia1::neo3lia1::neo3 [Cd2+::pm-r]) and two homzygous germ line-knockout heterokaryons, BVI Δlia1 (nia1::neo3lia1::neo3 [VI, Lia1+ Cd2+::pm-s]) and BVI Δlia1 (nia1::neo3lia1::neo3 [VII, Lia1+ Cd2+::pm-s]). BVI and BVI Δlia1 germ line-knockout heterokaryons, which produced viable progeny due to sufficient expression of LIA1 from the parental macronucleus, were crossed to obtain additional complete Δlia1 homozygous homokaryon strains (Δlia1 1.1, 1.2, 3.4, 4.7, and 4.12). These were used for phenotypic analyses with the initially generated complete micronuclear and macronuclear knockout lines described above. Complete elimination of the LIA1 gene was confirmed by genomic PCR and Southern blot hybridization analysis.

**Construction of the HA-LIA1 strain.** The hemagglutinin (HA) epitope sequence was added immediately after the LIA1 initiation codon by employing an overlapping PCR strategy. The primers used for a first round of PCR were Lia1-309-ApaI (5′-TGGGCCCCATCTCTTTGATTTACTC-3′) with Lia1-
HA1473r (5'-TATTCGCAAGACTGATAGGAGATAcTAAAACCAGATGCTAAA TTAA') and L1aa-HA1474 (5'-CTCTTATATGCCCTATGATGCTAGCTACAGGA AAAAAAATACACCT-3') with L1aa-2604Xc (5'-ACCTGAAAAAGAAAATA TGATAGATAAAG -3'). The HA coding sequence is shown in italics, the L1aa start codon is given in lowercase, and Apal and Xhol sites used for subsequent cloning are underlined. The two PCR fragments were denatured and mixed in an overlap PCR using the distal primers L1aa-380-Apal and L1aa-2604Xc. The amplified product was digested with Apal and Xhol and inserted into plasmid pT2 upstream of the histone H4 promoter-driven NEO cassette (13). Primers L1aa-2501B (TGGTATCTTCATCTCCTAATATTAG-3') and L1aa-3'SacI (5'-CGGAGCTCTAGATAGATAGATAGAACAAGATAGGATC C-3') were used to amplify L1aa downstream genomic sequence (BamHI and SacI sites are underlined). This amplified product was digested with BamHI and SacI and inserted in the poly linker sequence downstream of the NEO gene in the pT2-HA-L1aa vector to create the pHA-L1aa. Tetranyctena strains were generated by targeting the pHA-L1aa construct into the macronuclear L1aa locus by biolistic transformation. The construct, digested with Apal and SacI, was bombarded on gold particles separately into starved populations of strains B2086 and CU428 in 10 mM Tris-HCl (pH 7.4), followed by selection in paromomycin-containing medium. After phenotypic assortment, complete replacement of somatic copies of the L1aa gene with the HA-L1aa allele in both strains was confirmed by Southern blot analysis.

Fluorescence microscopy. Visualization of GFP fluorescence was performed as previously described (37). Conjugating cells were concentrated to 10^6 cells/ml by low-speed (1,000 x g) centrifugation, stained with DAPI (4',6'-diamidino-2-phenylindole; 1 μg/ml) for 10 to 30 min, and immobilized under 22-22.5 mm coverslips in 5 to 6% 2% v/v 10% dimethylformamide. Live and fixed-cell images were captured using a Nikon E600 epifluorescence microscope equipped with a Qimaging Retiga EX charge-coupled device camera driven by Openlab software (Improvision). If necessary, brightness and/or contrast of images was uniformly adjusted using Adobe Photoshop.

Protein analysis and immunoprecipitation assays. Immunoblotting was performed as previously described (21). B2086 × CU428 and HA-L1aa × HA-L1aa mating cells (2 x 10^6 cells/ml) were collected by centrifugation at the desired time points of conjugation. Concentrated cell samples were resuspended in lysis buffer (50 mM Tris-HCl, pH 8.5, 150 mM NaCl, 20 mM EDTA; 1% Triton X-100; and 1% sodium deoxycholate) with the addition of Complete Protease Inhibitor (Roche) and 2 μM phenylmethylsulfonyl fluoride (Sigma) (26, 37). After cell lysis, insoluble materials were sedimented by centrifugation (15,000 x g) for 15 min, and the supernatant was incubated with either (i) diluted (1/500) mouse monoclonal anti-HA.11 antibody (gift from C. D. Allis, Rockefeller University); rabbit monoclonal anti-Lia1 peptide MTGKQQVPVKGNFVKGM [Proteintech Group, Inc.]) antibodies. Pierce hors eradish peroxidase-conjugated goat anti-rabbit or goat anti-mouse secondary antibodies and SignalWest Pico or Dura kits were used to visualize immunoreactivity of selected antibodies with blotted proteins, followed by autoradiography.

Purification of immunoprecipitation. 1 x 10^6 mating pairs at 9 h of conjugation were collected by centrifugation and homogenized in 1 ml of lysis solution (50 mM Tris-HCl, pH 8.5, 150 mM NaCl, 20 mM EDTA; 1% Triton X-100; and 1% sodium deoxycholate) with the addition of Complete Protease Inhibitor (Roche) and 2 μM phenylmethylsulfonyl fluoride (Sigma) (26, 37). After cell lysis, insoluble materials were sedimented by centrifugation (15,000 x g) for 15 min, and the supernatant was incubated with either (i) diluted (1/500) mouse monoclonal anti-HA.11 antibodies or (ii) rabbit polyclonal anti-Pdd1p antibodies (1/10,000) overnight at 4°C. After overnight incubation, 50 μl of protein A agarose beads (Kirkgaard & Perry Laboratories) was added to each experimental sample and rotated for 3 h at 4°C to precipitate the target and interacting proteins. Beads with bound protein complexes were rinsed three times in wash buffer (50 mM Tris at pH 7.5, 150 mM NaCl, 2 mM MgCl2) and then eluted in a buffer containing 1% sodium dodecyl sulfate and 100 mM NaHCO3 prior to immunoblotting with either anti-Pdd1p or anti-HA antibodies, respectively, agarose beads, and DNA was deproteinized by phenol-chloroform (1:1) extraction followed by ethanol precipitation. PCR analysis was performed as described previously (22, 35) using primer pairs specific for the MIES or RIES or the macronucleus-retained sequences with recovered DNA (30 ng per reaction), PCR products were resolved on a 1.8% agarose gel and stained with 0.5% ethidium bromide. Kodak 1D Image analysis software was used to measure fluorescence intensities of PCR products. Quantification of each experimental sample was normalized to products amplified using primers for the BTU1 locus.

Nucleotide sequence accession numbers. Assembled DNA sequences are available under GenBank accession number EF645677 for the combined genomic clones and EF645678 for the cDNA clones.

RESULTS

L1aa encodes a small, basic protein that dynamically localizes to macronuclear anlagen during development. DNA rearrangement is a highly regulated process that involves extensive remodeling of the developing macronuclear genome. To find candidate genes whose proteins participate in these genome rearrangements, we utilized a GFP-based cytological screen to identify proteins that localize specifically to developing macronuclei when and where this process occurs (37). The localized in macronuclear anlagen (LIA) genes that were specifically expressed during development were obvious candidates for further investigation. Of the five candidates initially recovered, the protein encoded by L1aa was first selected for focused studies as the recovered GFP-Li1 fusion protein exhibited punctate localization in developing macronuclear similar to that of the essential genome rearrangement protein, Pdd1p (20, 21, 37) (data not shown) (Fig. 1A and B).

To begin further investigation, the L1aa coding sequence and flanking genomic DNA were cloned by employing a ligation-mediated, anchor PCR strategy (31). Subsequent mapping of L1aa to the recently published macronuclear genome sequence (11) revealed that it is the last predicted gene (TTHERM_00677500) on the right end of macronuclear chromosome CH445662, with the telomere positioned 1,779 nt downstream of the stop codon. Between the end of L1aa and the telomere, a −120-bp, 4-nucleotide repeat expansion of unknown function is found that is repeated elsewhere in the macronuclear genome (Fig. 1C). Comparison of the L1aa genomic sequence (GenBank accession no. EF645677) to cloned cDNAs (GenBank accession no. EF645678) revealed that the coding region contains two small introns of 58 (I1) and 67 (I2) nt, the second of which was miscalled by automated gene prediction. Annotated cDNA sequences recovered from the Tetrahymena Genome Database genome browser (www.ciliate.org) indicate that L1aa may contain an additional 62-bp intron in the 5′ untranslated region of some or all transcripts, but this has not been otherwise verified. The poly(A) addition site was verified by RNA-ligation-mediated PCR (9) and is located in a track of six adenines beginning 238 nt downstream of the UGA stop codon.

L1aa encodes a predicted protein of 233 amino acids (27 kDa) (Fig. 1C). Comparison of the predicted amino acid sequence with the nonredundant GenBank database or other protein databases (Pfam, SMART, etc.) did not show significant similarity to any known proteins or conserved domains. The coding sequence is rich in basic amino acids (22% lysine plus arginine), consistent with its nuclear localization, and could facilitate an association with nucleic acids or chromatin.

To more closely examine its localization, we generated an amino-terminal GFP fusion to a cloned L1aa cDNA. We observed uniform distribution of the GFP-Li1 fusion protein within the developing macronuclei of conjugating cells coinci-
Fig. 1. Lia1p exhibits dynamic localization within the developing macronuclei. (A) An N-terminal GFP-LIA1 cDNA fusion produced from the expression cassette shown was visualized in conjugating cells by epifluorescence microscopy. Differential interference contrast (DIC) light microscopy is displayed adjacent to GFP-Lia1p fluorescence and DAPI-stained imaging of cells for a representative mating pair at 7 to 8 h and in an exconjugate at 12 to 13 h of conjugation as indicated. The GFP-Lia1p fusion protein localizes within the developing macronuclear anlagen (large white arrowheads) and to the conjusome (white arrows). (B) Threefold enlargement of developing macronuclei of the exconjugate in panel A highlights GFP-Lia1p foci (white arrows). (C) A schematic representation of the LIA1 chromosomal locus shows the start and stop codons and the location of one upstream intron (IU) and two coding-region introns (11 and 12) of 62, 58, and 67 nt, respectively. The proximal telomere and C1AT nucleotide repeat shown are other features of the locus. Northern blot analysis was performed using total RNA (~12 μg/lane) from log-phase, vegetatively growing (G), starved (S), and synchronously conjugating cells at the times indicated (in hours). Arrowheads indicate transcript hybridization with the probes specific for LIA1 (shown in schematic above) and the loading control rPL21 mRNA.

Lia1p associates with the essential DNA deletion protein, Pdd1p. While our initial studies circumstantially linked Lia1p to structures associated with genome reorganization, we wanted to further connect this protein to the DNA rearrangement machinery by assessing its association with Pdd1p. To this end, we engineered two mating-compatible strains that expressed the Lia1 protein tagged with a single HA epitope, thus allowing us to detect the protein with antibodies specific to the HA tag. These strains were generated by introducing a cloned Lia1 allele (HA-LIA1), with the HA tag sequence added to its amino terminus, into the macronuclei of strains B2086 and CU428, replacing the endogenous LIA1 copies with the mod-
ified allele via homologous recombination (Fig. 2A). The introduced HA-LIA1 construct also contained a neomycin cassette downstream of the LIA1 polyadenylation site that permitted selection of the transformants by growth in paromomycin-containing medium. After phenotypic assortment, replacement of all macronuclear copies of LIA1 with the tagged allele was confirmed using genomic DNA PCR (data not shown), followed by Southern blot analysis (Fig. 2A). The HA-LIA1 strains were shown to mate without delay and produce viable progeny (data not shown). A Western blot of whole-cell protein lysates isolated from synchronously mating HA-LIA1 strains showed undetectable levels of Lia1 protein in vegetatively growing (V), starved (S), and early conjugating cells (4.5 h). The Lia1 protein (~27 kDa) starts to accumulate during conjugation at approximately 7.5 h, with protein expression continuing through the end of conjugation after the time at which DNA rearrangement occurs (Fig. 2B). Subsequent creation and use of Lia1p-specific peptide antibodies in wild-type cells showed that the endogenous protein and the tagged allele had identical expression patterns (Fig. 2B). Thus, while LIA1 transcript levels decrease before DNA rearrangement initiates, the protein is maintained at constant levels from 7.5 h until the completion of conjugation. As Lia1p does not accumulate until the time that anlagen appear, it is clear that this protein must act exclusively in postzygotic steps of development.

To test whether Lia1p and Pdd1p physically interact or are present within a common complex as suggested by their similar localization patterns, coimmunoprecipitation experiments were performed using the HA-LIA1 Tetrahymena strains. Whole-cell protein lysates were isolated from mating populations of wild-type (B2086 × CU428) and HA-LIA1 strains and immunoprecipitated separately with anti-Pdd1p and anti-HA antibodies. Western blot assays were performed on immunoprecipitates by using the reciprocal antibodies. Immunoprecipitation with HA-specific antibodies coprecipitated Pdd1p from extracts of conjugating HA-LIA1 cells but not from lysates containing the untagged Lia1p (Fig. 2C). Likewise, precipitation of Pdd1p from extracts coprecipitated HA-Lia1p. To ensure that detection of Pdd1p in our HA immunoprecipitates was not just due to its high abundance, we tested our immunoprecipitates for the presence of another very abundant protein, actin, using a Drosophila melanogaster-specific antiactin antibody that recognizes the Tetrahymena protein and did not detect such nonspecific precipitation (data not shown). In addition, separate treatment of lysates prior to immunoprecipitation with RNase A and DNase I indicated that the Lia1p-

FIG. 2. Lia1p coimmunoprecipitates with Pdd1p. (A) An HA-epitope tagged allele was recombined into the LIA1 locus. Inclusion of a histone H4 promoter-driven NEO gene in the HA-LIA1 replacement cassette allowed selection of transformants. The indicated probe corresponding to the 3’ nontranscribed sequence downstream of the LIA1 stop codon was used for the Southern blot analysis of three wild-type and five transformant lines. The hybridizing fragment corresponding to the wild type and HA-LIA1 allele is indicated. The specific transformant subclones analyzed are indicated by the numbers above each lane. (B) Protein lysates from conjugating HA-LIA1 × HA-LIA1 and wild-type (B2086 × CU428) cells, harvested every 1.5 h, were used for Western blot analysis using antisera to the HA epitope, an N-terminal Lia1p peptide, and Pdd1p (21). V, vegetatively growing cells; S, starved cells. Numbers above the lanes are times in hours. (C) Protein lysates from mating Tetrahymena cells expressing HA-LIA1 were collected at 9 h and immunoprecipitated separately with antisera recognizing the HA epitope (Lia1p) or Pdd1p. Immunoblot analysis of precipitated proteins was performed with the reciprocal antibodies indicated to the right of each autoradiogram.
Pdd1p interaction did not require any nucleic acid intermediate (data not shown).

**Lia1p associates with germ line-limited sequences (IESs).** If Lia1p is involved in DNA rearrangements, it should associate with sequences that are eliminated. It has been shown previously through in situ hybridization and ChiP studies that Pdd1p-associated chromatin is enriched for repetitive (Tt2512) and unique (M and R element) germ line-limited sequences, relative to macronucleus-retained sequences (28, 33, 35). To assess whether Lia1p is associated with IESs prior to their removal from the developing somatic genome, we performed ChiP experiments using HA-LIA1 strains and HA-specific antibodies (Fig. 3). The quantity of immunoprecipitated chromatin from two well-characterized IESs, the M and R IESs, relative to the macronucleus-retained sequence between them was determined by PCR. All PCRs also contained primers to the macronuclear BTU1 locus as a normalization control. ChiP with anti-Pdd1p, anti-histone H3 dimethyl K9, or anti-HA (Lia1p) antibodies resulted in a four- to fivefold enrichment of IES chromatin over the macronucleus-retained region. Additional control immunoprecipitations with general anti-histone H4 or anti-histone H3 trimethyl K4 antibodies showed no enrichment of IES chromatin. These findings indicate that Lia1p is specifically associated with germ line-limited sequences during macronuclear development.

**LIA1 is an essential gene for the completion of nuclear development and IES excision.** The association of Lia1p with IESs and Pdd1p clearly implicates this protein in the DNA rearrangements occurring in the somatic genome. To determine whether this novel protein is critical for this process, we disrupted the gene and assessed the loss of function phenotype (Fig. 4A). Initially we knocked out LIA1 from the parental macronuclei of two mating-compatible strains. After confirming complete replacement of macronuclear copies of LIA1 with the NEO3 selection cassette (32), we crossed these strains and examined their progression through the events of conjugation. LIA1 macronuclear knockout lines completed events through the formation of developing anlagen but appeared to be somewhat delayed in the late stages of nuclear differentiation. Eventually, the strains completed conjugation and produced progeny at the same frequency as did the wild type. When we examined LIA1 expression during conjugation in these knockout lines by Northern blot analysis (Fig. 4B), we observed a 3- to 4-h delay in transcript accumulation, which correlated with the time delay that we observed in completing conjugation, as assessed by timing of pair separation (data not shown). This suggested to us that Lia1p may have an essential late function in nuclear differentiation but that zygotic expression from the newly forming macronuclei rescued the loss of parental gene expression.

To determine whether LIA1 is essential for conjugation, we disrupted the gene from both the macronucleus and the micronucleus of mating-compatible strains. Complete loss of LIA1 gene copies was verified by Southern blot analysis for the five transformant lines (Fig. 4C). We crossed lines lacking all LIA1 gene copies and found that these cells were unable to complete conjugation and arrested late in development, just prior to elimination of one of the two micronuclear precursors (Fig. 4D). Upon subcloning mating pairs into growth medium, we obtained no viable progeny lines. The only surviving pairs had aborted mating without forming new macronuclei. Thus, LIA1 is an essential gene for *Tetrahymena* development. The conjugation arrest phenotype appears identical to that of other knockouts (e.g., *PDD1, DCL1*, and *TWI1*) (10, 22, 24) that fail to complete development, arresting as exconjugants with two new micronuclei and two new macronuclei (wild-type cells terminate conjugation after eliminating one of their new micronuclei and will divide once returned to growth medium). These arrested LIA1 knockout lines appeared to incompletely amplify the DNA within their developing macronuclei as...
judged by DAPI staining to determine fluorescence intensity, relative to micronuclei, and reduced recovery of total DNA from postconjugative cells (data not shown).

To ascertain which steps in nuclear development were perturbed by loss of LIA1 expression, we assessed the establishment of histone H3K9 methylation, distribution of Pdd1p in anlagen, and rearrangement of known IESs. Lia1p normally begins accumulating near the time at which H3K9 methylation appears in new macronuclei. In knockout lines, this modification occurred normally; thus, Lia1p appears to act downstream of the establishment of this chromatin mark (Fig. 4E).

Pdd1p binds IES-associated, H3K9 methylated chromatin and organizes into distinct nuclear foci prior to excision of these germ line-limited sequences (20, 35). As H3K9 methylation was not disrupted in conjugating ΔLIA1 cells (Fig. 4E), Pdd1p should still be recruited to IESs. We therefore examined whether Pdd1p can still be organized into nuclear foci where DNA rearrangement is believed to occur. We fixed both conjugating wild-type (CU427 × CU428) and ΔLIA1 cells 13 h after mixing and detected Pdd1p in both populations by indirect immunofluorescence localization (Fig. 5). DNA elimination structures containing Pdd1p were readily apparent in wild-type conjugants as ring-like foci as previously described (20). In ΔLIA1 cells, Pdd1p was abundant within macronuclear anlagen but was dispersed or concentrated in relatively large ring-like structures that have the appearance of very large DNA elimination structures. Thus, it appears that LIA1 is required for the normal organization of Pdd1p-containing, DNA elimination structures.

To determine whether LIA1 knockout lines fail to excise IESs, we examined two loci that contain germ line-limited sequences by Southern blot analysis of genomic DNA isolated from populations of postconjugative wild type or LIA1 knockouts. The first of these loci is the well-characterized L/M/R locus of micronuclear chromosome 4 that contains three IESs for which the locus is named (2–5). Analysis of DNA from wild-type matings (B2086 × CU428) showed that the majority of the DNA is in the macronuclear rearranged form. Due to alternative rearrangement boundaries for the M IES in the middle of this region, a doublet of 6.0- and 6.3-kbp fragments was observed and was the size expected for the fully rearranged locus with all three IESs removed (Fig. 6A). The unrearranged DNA observed as a 9.3-kbp fragment in the analysis of wild-type cells is certainly the copies remaining in the new micronuclei. In contrast, analysis of DNA from postconjugative ΔLIA1 knockout cells predominantly detected this unrearranged form. The DNA of the size expected for the rearranged locus likely represents the DNA from the ~4 to 5% of unmated cells analysis shown. Genomic DNA was analyzed from three wild-type strains and five ΔLIA1 strains. The specific transformant subclones analyzed are indicated by the numbers above each lane. (D) Fluorescence microscopy of unfed, DAPI-stained wild-type (WT) (B2086 × H11003 CU428) and ΔLIA1 strains collected H11011 32 h postmixing. Percentages given below report the numbers of wild-type and ΔLIA1 cells at indicated stages of development. (E) Whole-cell protein isolated from conjugating wild-type and ΔLIA1 strains at indicated time points postmixing was analyzed by Western blot analysis with antibodies recognizing histone H3 dimethyl K9 (Upstate Biotechnology, NY).
(Fig. 4D) in the population of ∆LIA1 cells that retain the rearranged locus in their macronuclei, as they have not gone through new macronuclear development.

We observed similar results for the IES upstream of the calmodulin locus (19). While the predominant form in wild-type crosses was the size expected for elimination of this 1.4-kbp IES, this rearranged form was underrepresented in the DNA of postconjugative LIA1 knockout lines (Fig. 6B). The unrearranged form was not as apparent for this locus. It is unclear if this is due to underamplification of this locus during macronuclear development or to technical issues related to the probe used in detection of this locus, but this result has been reproducible in multiple crosses of the LIA1 knockouts and crosses of DCL1 knockouts (22). Regardless, it is clear that DNA rearrangement of this IES is also severely or completely abrogated in these knockout lines.

We introduced the GFP-LIA1 expression construct into these knockout lines and found that expression of the tagged protein was sufficient to rescue both the conjugation arrest phenotype (data not shown) and the DNA rearrangement defect (Fig. 6). This result demonstrates that the GFP-Lia1 protein is functional and supports the significance of the localization described (Fig. 1). Together these results show that Lia1p is a key player in the reorganization of the Tetrahymena genome.

**DISCUSSION**

We have identified a novel protein encoded by LIA1 and show here that it is an essential component of the DNA rearrangement machinery. It is assembled on the H3K9 methylated chromatin of IESs along with Pdd1p, an abundant protein that is also required for the elimination of these germ line-limited sequences (Fig. 3). Disruption of LIA1 in both micronuclei and macronuclei resulted in conjugative arrest and failure of detectable IES excision (Fig. 4 and 5). The timing of LIA1 expression and our finding that H3K9 methylation occurs normally in LIA1 knockout lines demonstrate that this protein is critical for a downstream step of DNA rearrangement. Identification of LIA1 provides new insight into the types of pro-
teins necessary to eliminate the more than 15 Mbp of DNA from developing macronuclei.

We believe that Lia1p is directly involved in these genome rearrangements based on its association with IESs and Pdd1p and the fact that disruption of the gene blocks this process. The determination of the exact role of Lia1p in DNA rearrangement will likely require the development of biochemical assays for the steps of DNA rearrangement. Analysis of the sequence of this protein offers few obvious clues. The protein has a strong positive charge, as it is rich (22%) in lysine and arginine. This offers the potential that it could engage in nucleic acid binding, but overall the protein sequence exhibits low complexity, making bioinformatics challenging. After H3K9 methylation, which occurs shortly after anlagen are formed, IESs must begin their assembly into foci to be later excised from the genome with the flanking DNA rejoined. It is possible that Lia1p directly participates in the excision-rejoining events, but this is just one of the downstream steps of the process in which this protein might act. As we found that LLAI1 is required to properly form the Pdd1p-containing, DNA elimination structures (Fig. 5), Lia1p may be involved in the maturation of these structures prior to IES excision.

It is unclear, though, why impeding DNA rearrangement leads to the conjugation arrest phenotype observed. $\Delta L AI 1$ cells, like all mutant lines that abolish DNA rearrangement ($\Delta P D D 1$, $\Delta P D D 2$, $\Delta T W 1$, and $\Delta D C L 1$), arrest at the same stage before final elimination of one micronuclear precursor, even though they likely act at different times and steps in the pathway (10, 22, 24, 29). Thus, it would appear that a checkpoint exists at the end of conjugation that monitors the initiation and/or completion of genome rearrangement. This putative checkpoint arrests cells before the elimination of one of the two new micronuclei, an event that normally would occur prior to the first postconjugative cell division. In $\Delta L AI 1$ cells, it is possible that IESs begin to be excised, but as the process cannot be completed, double-strand breaks accumulate that trigger a DNA damage checkpoint. If DNA rearrangement is blocked prior to IES excision and unrepairred breaks are not the resulting lesion triggering the arrest, it is less obvious how cells might sense that IESs have not been removed.

Recent studies have focused on the finding that DNA rearrangement is directed by an RNAi-related mechanism. These efforts have reinforced the idea that this unique process in Tetrahymena is related more generally to the establishment of heterochromatin. The generation of IES-specific small RNAs and targeting of H3K9 methylation occur relatively early in conjugation and micronuclear differentiation (22, 24, 25, 35). The events that occur after RNAi-directed chromatin modification also exhibit parallels to the packaging of heterochromatin in other eukaryotes. The most striking similarity is the condensation of these IESs into punctate foci prior to their eventual excision. It seems that DNA rearrangement involves extensive reorganization of dispersed loci into structures in the nucleus where their excision occurs. It is compelling to speculate that many of the events up to the final removal of these sequences are evolutionarily related to the silencing and organization of heterochromatin at the nuclear periphery in unparsable eukaryotes (15). In this way, Tetrahymena DNA rearrangements serve as a model for understanding the partitioning of chromosomal domains during development.

Lia1p is the first DNA rearrangement protein identified for which removal of parental copies of its gene is not sufficient to abolish the process. Pdd3p and the other Lia proteins may also act similarly, as each is expressed relatively late into conjugation, near or after the time that macronuclear anlagen are formed, but knockout phenotypes for these genes have not been described. We are intrigued by our observation that macronuclear LLAI1 knockout strains exhibited delayed conjugation (data described above but not shown) but completed development once zygotic expression of the gene was induced (Fig. 4B). This suggests to us that the latter stages may involve a stepwise assembly of the DNA rearrangement machinery. It will be informative to generate mutants in each of these putative components to elucidate their roles in performing this remarkable genome restructuring.

ACKNOWLEDGMENTS

We recognize Meng-Chao Yao for his initial role in the identification of LLAI1. We also thank Rajat Jain for his contributions to this study and Deborah Frank for reading of the manuscript.

This work was supported by National Institutes of Health research grant GM069593 and a grant from the American Cancer Society, IRO-58-010-45, to D.L.C.

REFERENCES

1. Asai, D. J., and J. D. Forney (ed.). 1999. Methods in cell biology, vol. 62. Tetrahymena thermophila. Academic Press, San Diego, CA.
2. Austerberry, C. F., C. D. Allis, and M. C. Yao. 1984. Specific DNA rearrangements in synchronously developing nuclei of Tetrahymena. Proc. Natl. Acad. Sci. USA 81:7383–7387.
3. Austerberry, C. F., R. O. Snyers, and M. C. Yao. 1989. Sequence microheterogeneity is generated at junctions of programmed DNA deletions in Tetrahymena thermophila. Nucleic Acids Res. 17:7263–7272.
4. Austerberry, C. F., and M. C. Yao. 1987. Nucleotide sequence structure and consistency of a developmentally regulated DNA deletion in Tetrahymena thermophila. Mol. Cell. Biol. 7:445–445.
5. Austerberry, C. F., and M. C. Yao. 1988. Sequence structures of two developmentally regulated, alternative DNA deletion junctions in Tetrahymena thermophila. Mol. Cell. Biol. 8:3947–3950.
6. Bruns, P. J., and D. Cassidy-Hanley. 2000. Biochemical transformation of macronuclear and micronuclear. Methods Cell Biol. 62:501–512.
7. Cassidy-Hanley, D., J. Bowen, J. H. Lee, E. Cole, L. A. VerPlank, J. Gaertig, M. A. Gorovsky, and P. J. Bruns. 1997. Germline and somatic transformation of mating Tetrahymena thermophila by particle bombardment. Genetics 146:135–147.
8. Chalker, D. L., and M.-C. Yao. 1996. Non-Mendelian, heritable blocks to DNA rearrangement are induced by loading the somatic nucleus of Tetrahymena thermophila with germ line-limited DNA. Mol. Cell. Biol. 16:3658–3667.
9. Chalker, D. L., and M. C. Yao. 2001. Nongenetic, bidirectional transcription precedes and may promote developmental DNA deletion in Tetrahymena thermophila. Gene Dev. 15:1287–1298.
10. Coyne, R. S., M. A. Nikiforov, J. F. Smothers, C. D. Allis, and M. C. Yao. 1999. Parenteral expression of the chromodom protein Pdd1p is required for completion of programmed DNA elimination and nuclear differentiation. Mol. Cell 4:460–472.
11. Eisen, J. A., R. S. Coyne, M. Wu, D. Wu, M. Thigaarajan, J. R. Wortman, J. H. Badger, Q. Ren, P. Amedo, K. M. Jones, L. J. Tallon, A. L. Delcher, S. L. Salzberg, C. J. Silva, B. J. Haas, W. H. Majoros, M. Farzad, J. M. Carlson, R. K. Smith, Jr., J. Garg, R. E. Pearlman, K. M. Karrer, L. Sun, G. Manning, N. C. Eldred, A. P. Turkewitz, D. J. Asai, E. D. Wilkes, Y. Wang, H. Cai, K. Collins, B. A. Stewart, S. R. Lee, K. Wilamowska, Z. Weinberg, W. L. Ruzzo, D. Wloga, J. Gaertig, J. Frankel, C. C. Tsao, M. A. Gorovsky, P. J. Keeling, R. F. Waller, N. J. Patrun, J. M. cherry, N. A. Stover, C. J. Krieger, C. del Toro, H. F. Ryder, S. C. Williamson, R. A. Barbeau, E. P. Hamilton, and E. Orias. 2000. Macronuclear genome sequence of the ciliate Tetrahymena thermophila, a model eukaryote. PLoS Biol. 4(12):2628–2682.
12. Orf, D., and M.-C. Yao. 2000. A long stringent sequence for programmed chromosome breakage in Tetrahymena thermophila. Nucleic Acids Res. 28:895–900.
13. Gaertig, J., L. Gu, B. Hai, and M. A. Gorovsky. 1994. High frequency vector-mediated transformation and gene replacement in Tetrahymena. Nucleic Acids Res. 22:5391–5398.
14. Gorovsky, M. A., and J. Woodard. 1969. Studies on the nuclear structure and function in Tetrahymena pyriformis. J. Cell Biol. 42:673–682.
29. Nikiforov, M., J. Smothers, M. Gorovsky, and C. Allis. 2000. Tetrahymena genetics: two nuclei are better than one. Methods Cell Biol. 62:127–186.

30. Prescott, D. M. 1994. The DNA of ciliated protozoa. Microbiol. Rev. 58: 233–267.

31. Riley, J., R. Butler, D. Ogilvie, R. Finnear, D. Jenner, S. Powell, R. Anand, J. G. Smith, and A. F. Markham. 1990. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. Nucleic Acids Res. 18:2887–2890.

32. Shang, Y., X. Song, J. Bowen, R. Corstanje, Y. Gao, J. Gaertig, and M. A. Gorovsky. 2002. A robust inducible-repressible promoter greatly facilitates gene knockouts, conditional expression, and overexpression of homologous and heterologous genes in Tetrahymena thermophila. Proc. Natl. Acad. Sci. USA 99:3734–3739.

33. Smothers, J. F., M. T. Madireddi, F. D. Warner, and C. D. Allis. 1997. Programmed DNA degradation and nuclear biogenesis occur in distinct organelles during macronuclear development in Tetrahymena. J. Eukaryot. Microbiol. 44:79–88.

34. Smothers, J. F., C. A. Mizzen, M. M. Tubbert, R. G. Cook, and C. D. Allis. 1997. Pdd1p associates with germline-restricted chromatin and a second novel anlagen-enriched protein in developmentally programmed DNA elimination structures. Development 124:4537–4545.

35. Taverna, S. D., R. S. Coyne, and C. D. Allis. 2002. Methylation of histone h3 at lysine 9 targets programmed DNA elimination in tetrahymena. Cell 110: 701–711.

36. Williams, N. E. 2000. Preparation of cytoskeletal fractions from Tetrahymena thermophila. Methods Cell Biol. 62:441–447.

37. Yao, M.-C., C.-H. Yao, L. M. Halasz, P. Fuller, C. H. Rexer, S. H. Wang, R. Jain, R. S. Coyne, and D. L. Chalker. 2007. Identiﬁcation of novel chromatin-associated proteins involved in programmed genome rearrangements in Tetrahymena. J. Cell Sci. 120:1976–1989.

38. Yao, M. C., and J. L. Chao. 2005. RNA-guided DNA deletion in Tetrahymena: an RNAi-based mechanism for programmed genome rearrangements. Annu. Rev. Genet. 39:537–559.

39. Yao, M. C., J. Choi, S. Yokoyama, C. F. Austerberry, and C. H. Yao. 1984. DNA elimination in Tetrahymena: a developmental process involving extensive breakage and rejoining of DNA at deﬁned sites. Cell 36:433–440.

40. Yao, M. C., S. Dubarcourt, and D. L. Chalker. 2002. Genome-wide rearrangements of DNA in ciliates, p. 790–798. In N. Craig, R. Craigie, M. Gelbert, and A. Lambowitz (ed.), Mobile DNA II. Academic Press, New York, NY.

41. Yao, M. C., and M. A. Gorovsky. 1974. Comparison of the sequences of macro- and micronuclear DNA of Tetrahymena pyriformis. Chromosoma 48:1–18.

42. Yao, M. C., K. Zheng, and C. H. Yao. 1987. A conserved nucleotide sequence at the sites of developmentally regulated chromosomal breakage in Tetrahymena. Cell 48:779–788.