A Functional Role for Transposases in a Large Eukaryotic Genome

Mariusz Nowacki[1], **Brian P. Higgins**[1], **Genevieve M. Maquilan**[1], **Estienne C. Swart**[1], **Thomas G. Doak**[2], and **Laura F. Landweber**[1,*]

[1]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA

[2]Department of Biology, Indiana University, Bloomington, IN 47405, USA

Abstract

Despite comprising much of the eukaryotic genome, few transposons are active, and they usually confer no benefit to the host. Through an exaggerated process of genome rearrangement, *Oxytricha trifallax* destroys 95% of its germline genome during development. This includes the elimination of all transposon DNA. We show that germline-limited transposase genes play key roles in this process of genome-wide DNA excision, which suggests that transposases function in large eukaryotic genomes containing thousands of active transposons. We show that transposase gene expression occurs during germline-soma differentiation and that silencing of transposase by RNA interference leads to abnormal DNA rearrangement in the offspring. This study suggests a new important role in *Oxytricha* for this large portion of genomic DNA that was previously thought of as junk.

Oxytricha trifallax, a ciliated protozoan, is a unicellular eukaryote with two types of nuclei in the same cytoplasm. Diploid micronuclei are transcriptionally silent during vegetative growth but transmit the germline genome through sexual conjugation (fig. S1). After sex, the gene-rich macronuclei that govern somatic gene expression degrade, and new macronuclei develop from germline micronuclei. Macronuclear development in *Oxytricha* (also called *Sterkiella*) involves massive chromosome fragmentation and the deletion of thousands of copies of germline transposons (selfish DNA elements that are capable of self-replication in their host genomes) plus the destruction of intergenic DNA and noncoding DNA sequences that interrupt genes [internally eliminated sequences (IESs)]. This leads to a 95% reduction in sequence complexity, compressing the 1-Gb germline genome into a gene-dense macronuclear genome of only ~50 Mb. Sorting and reordering the hundreds of thousands of remaining short pieces [macronuclear-destined segments (MDSs)] produces mature macronuclear sequences. The resulting ~2-kb nanochromosomes typically bear only a single gene bound by short telomeres and amplified to a high copy number (1). The remarkable degree of specificity and reproducibility of genome reorganization in ciliates relies on RNA template–guided trans-nuclear comparison of germline and somatic genomes

---

[*]To whom correspondence should be addressed. lfl@princeton.edu.

(2–5). However, the molecular machinery responsible for this DNA excision and recombination remains unknown.

O. trifallax carries thousands of germline-limited transposable elements in its diploid micronucleus (6). These DNA sequences are restricted to the micronucleus and absent from the macronucleus because they are deleted during genome differentiation. Because they are ~4-kb repetitive sequences flanked by blocks of *Oxytricha* telomeric repeats $(G_4T_4)_4$, they have been designated telomere-bearing elements (TBEs) (7). Several lines of evidence suggest that TBEs transpose through a DNA intermediate: Apart from the presence of 3-bp insertion site duplications and inverted terminal repeats, TBEs code for a 42-kD member of a superfamily of transposases, characterized by a DDE catalytic motif (fig. S2) (8, 9). Moreover, phylogenetic evidence supports a recent origin of TBE insertions (10). *Oxytricha* TBEs comprise three known types: TBE1, -2, and -3, whose transposase genes differ 26 to 29% from each other at the nucleotide level (table S1A) while sharing high protein similarity (≥83%). Previous examination of *Oxytricha* TBE1 trans-posase genes identified no copies that were interrupted by stop codons or frame shifts and provided evidence of purifying selection acting on their encoded proteins (6, 8), which suggests that the encoded transposase might be responsible for the excision of TBE1 elements during macro-nuclear development (11, 12).

In order to test whether germline-encoded TBE transposases function in somatic DNA deletion and rearrangement, we examined the expression and function of TBE transposases, given their ability to cut and join DNA (11). First, we asked whether TBE transposase genes are expressed during macronuclear development (Fig. 1). We examined the expression of transposases from both vegetative and conjugating cells [at 11, 24, and 34 hours after the initiation of mating (mixing of mating types)], which revealed that they were expressed exclusively during conjugation; the highest expression level occurred 24 hours after mixing [most likely coinciding with polytene chromosome formation and TBE1 elimination (11)]. The presence of polyadenylated transcripts (Fig. 1G) at this stage was confirmed, suggesting that TBE expression occurs during DNA rearrangement and that transposase could therefore play a role in their elimination.

In order to determine whether transposases influence developmental DNA rearrangements, we silenced all three transposase types by micro-injecting double-stranded RNA (dsRNA) into the cytoplasm at three different stages during conjugation, including at (i) early pair formation, 3 to 5 hours after the mixing of different mating types; (ii) 10 hours after mixing, right before pair separation; and (iii) 12 to 15 hours after mixing, right after pair separation (table S2). We observed the greatest effect when dsRNA was injected 12 to 15 hours after mixing because few cells survived and those that survived showed delayed growth. DNA that was extracted from the progeny of cells injected ~10 hours and 12 to 15 hours after mixing and cultured for only 1 week sometimes contained incorrect DNA rearrangements as well as increased levels of nonprocessed DNA in the new macronucleus, as revealed by polymerase chain reaction (PCR) examination of three genes: *telomere end-binding protein α subunit* (*TEBPα*), *telomere end-binding protein β subunit* (*TEBPβ*), and *DNA polymerase α* (*pol-α*) (Fig. 2, fig. S3, and table S2). The cells that were injected 12 to 15 hours after mixing and harvested after 2 weeks of growth, on the other hand, evidenced no partial rearrangements, even in clones from the region of the gel corresponding to partially processed forms (in between the visible precursor- and product-sized bands) or by direct cloning of all amplified DNA, presumably because of reversion to the wild type during extended culture (5).

At 12 to 15 hours after mixing, we silenced each transposase type individually (TBE1, TBE2, or TBE3) as well as all three possible pairs and all three types together. Because

triple silencing at 12 to 15 hours after mixing reliably inhibited rearrangements (Fig. 3), we repeated the triple silencing and the combination of TBE1 and TBE2 at this time point and also analyzed transposase mRNA levels (Fig. 4F). None of the single or other paired treatments had a strong effect. PCR at several loci in the progeny of cells triply silenced at 12 to 15 hours revealed a substantial accumulation of precursor nonrearranged versions of each gene (Fig. 3 and table S2, B and C), indicating stalled rearrangement. A weak effect was also observed after TBE1 and TBE2 were simultaneously silenced (fig. S4A). Normally the precursor, micro-nuclear DNA is rare and hard to detect (untreated lanes in Fig. 3 and fig. S4, A and B) relative to the macronuclear products, which have roughly a 1000-fold amplification. The abundance of precursor-sized product in TBE-silenced cells strongly suggests an increase in the ratio of nonrearranged to rearranged DNA. This was confirmed in the progeny of triply silenced versus control cells, with nonrearranged versions of *TEBPa* displaying a 6.6- or 3.2-fold greater abundance in the two independent experiments performed at 12 to 15 hours after mixing (Fig. 3 and fig. S4).

Unprocessed germline chromosomes in *Oxytricha* are much longer than somatic molecules. Therefore, we confirmed that the progeny of triply silenced and TBE1- and TBE2-silenced cells contain more high–molecular weight DNA (limited by mobility in the gel) than do control cells (Fig. 4A and fig. S4D) and noted a fourfold increase when all TBE genes were silenced and a twofold increase when TBE1 and TBE2 were silenced (Fig. 4B and fig. S4D). In addition, we observed an increased hybridization of the high–molecular weight DNA to transposase DNA (Fig. 4, C and D, and fig. S4D), which is consistent with the hypothesis that these cells had a reduced ability to eliminate these transposons, which represent approximately 2% of the germline genome. Furthermore, we also observed no hybridization of transposase DNA to macronuclear-sized (<12 kb) molecules (Fig. 4C and fig. S4D), which is consistent with prior observations that TBE transposons are absent from macronuclear chromosomes (7, 11, 13). The silencing of the three transposase types (and to some extent TBE1 and TBE2 together) alters the efficiency of developmental DNA rearrangements, including chromosome fragmentation, and this leads to the accumulation of long nonprocessed DNA molecules. Our results suggest that germline-limited transposase genes in *Oxytricha* function in genome rearrangement.

These transposons might not merely be parasitic invaders that reduce host fitness or have little phenotypic effect (14, 15) but instead mutualists directly contributing a useful function for the organism (12), such as genomic DNA processing. This would also account for the evidence that strong purifying selection (6, 16) acts on TBE-encoded proteins (fig. S5 and table S3). Particularly striking is that all three TBE transposase groups need to be silenced to achieve the greatest effect. Given the massive level of DNA deletion and reorganization in *Oxytricha*, it is possible that the three types have either specialized or overlapping functions or that their contributions are additive. A high level of transposase activity may be necessary to facilitate thousands of simultaneous DNA rearrangements, affording an indispensable role to the large quantity of germline-limited transposable elements, which were previously considered selfish DNA or part of the dispensable genome (17). This mutualism (18) is uniquely different from other known cases of transposon domestication, such as RAG recombinase (19), piggy*Bac* (20), or Mariner (21), in which a single transposase copy has assumed a host function. The situation in *Oxytricha* could also be viewed as a case of massive transposon domestication, in which thousands of unmodified transposons may have been recruited for their transposase as part of the catalytic machinery of RNA-guided genome rearrangements just before they are eliminated from the macronuclear genome. In support of this model, we found that IES excision can precede transposon elimination (fig. S6), refuting a simple indirect effect of genome rearrangements occurring only after transposon processing. Transient TBE transposase expression begins after meiosis (Fig. 1 and fig. S1) and coincides with the production of maternal RNA templates that guide

rearrangement (5) as well as all events of DNA processing, including TBE transposon elimination, IES excision, and segment-reordering in the developing nucleus (22).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References and Notes

1. Prescott DM. Microbiol Rev. 1994; 58:233. [PubMed: 8078435]

2. Chalker DL, Yao MC. Mol Cell Biol. 1996; 16:3658. [PubMed: 8668182]

3. Duharcourt S, Keller AM, Meyer E. Mol Cell Biol. 1998; 18:7075. [PubMed: 9819394]

4. Mochizuki K, Fine NA, Fujisawa T, Gorovsky MA. Cell. 2002; 110:689. [PubMed: 12297043]

5. Nowacki M, et al. Nature. 2008; 451:153. [PubMed: 18046331]

6. Witherspoon DJ, et al. Mol Biol Evol. 1997; 14:696. [PubMed: 9214742]

7. Herrick G, et al. Cell. 1985; 43:759. [PubMed: 3000614]

8. Doak TG, Doerder FP, Jahn CL, Herrick G. Proc Natl Acad Sci USA. 1994; 91:942. [PubMed: 8302872]

9. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

10. Seegmiller A, et al. Mol Biol Evol. 1996; 13:1351. [PubMed: 8952079]

11. Williams K, Doak TG, Herrick G. EMBO J. 1993; 12:4593. [PubMed: 8223469]

12. Klobutcher LA, Herrick G. Prog Nucleic Acid Res Mol Biol. 1997; 56:1. [PubMed: 9187050]

13. Hunter DJ, Williams K, Carinhour S, Herrick G. Genes Dev. 1989; 3:2101. [PubMed: 2560753]

14. Doolittle WF, Sapienza C. Nature. 1980; 284:601. [PubMed: 6245369]

15. Orgel LE, Crick FH. Nature. 1980; 284:604. [PubMed: 7366731]

16. Doak TG, Witherspoon DJ, Doerder FP, Williams K, Herrick G. Genetica. 1997; 101:75. [PubMed: 9465400]

17. Morgante M, De Paoli P, Radovic S. Curr Opin Plant Biol. 2007; 10:149. [PubMed: 17300983]

18. Kidwell MG, Lish AJ. Evolution. 2001; 55:1. [PubMed: 11263730]

19. Kapitonov VV, Jurka J. PLoS Biol. 2005; 3:e181. [PubMed: 15898832]

20. Sarkar A, et al. Mol Genet Genomics. 2003; 270:173. [PubMed: 12955498]

21. Robertson HM, Zumpano KL. Gene. 1997; 205:203. [PubMed: 9461395]

22. Möllenbeck M, et al. PLoS One. 2008; 3:e2330. [PubMed: 18523559]

23. Materials and methods are available as supporting material on *Science* Online.
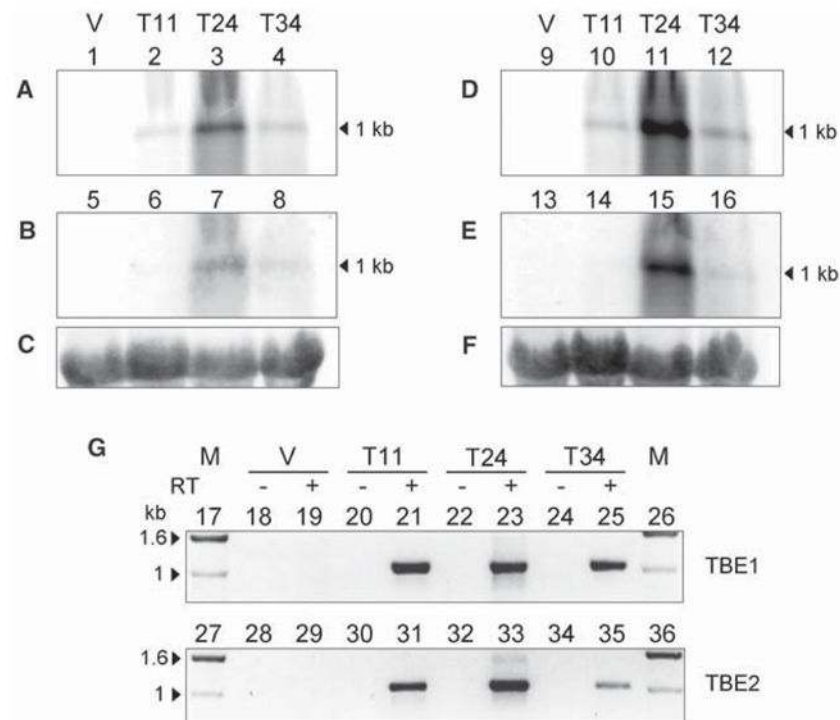
Fig. 1.
Expression of TBE transposase genes. (**A** to **F**) Northern hybridization against total RNA extracted from vegetative cells (V) as well as 11-, 24-, and 34-hour time points (T) after mixing of mating types. TBE transposase transcripts are most abundant 24 hours after mixing, when DNA rearrangements occur (22). 10 μg of RNA was loaded per lane and hybridized to ~1-kb DNA probes (23) for (A) TBE1, (B) TBE3, (D) TBE1, -2, and -3 mixed together (1:1:1 ratio), (E) TBE2 transposase, and (C and F) small-subunit ribosomal RNA (rRNA); (C) is the control for (A) and (B), and (F) is the control for (D) and (E). (**G**) Reverse transcription PCR of TBE1 (lanes 17 to 26) and TBE2 (lanes 27 to 36) transposase polyadenylate-containing mRNA from vegetative cells (V) and three time points (T) during conjugation (23). Plus and minus symbols indicate the presence or absence of reverse transcriptase (RT), respectively. Arrowheads indicate the size of relevant markers [1-kb ladder (Invitrogen, Carlsbad, CA)].
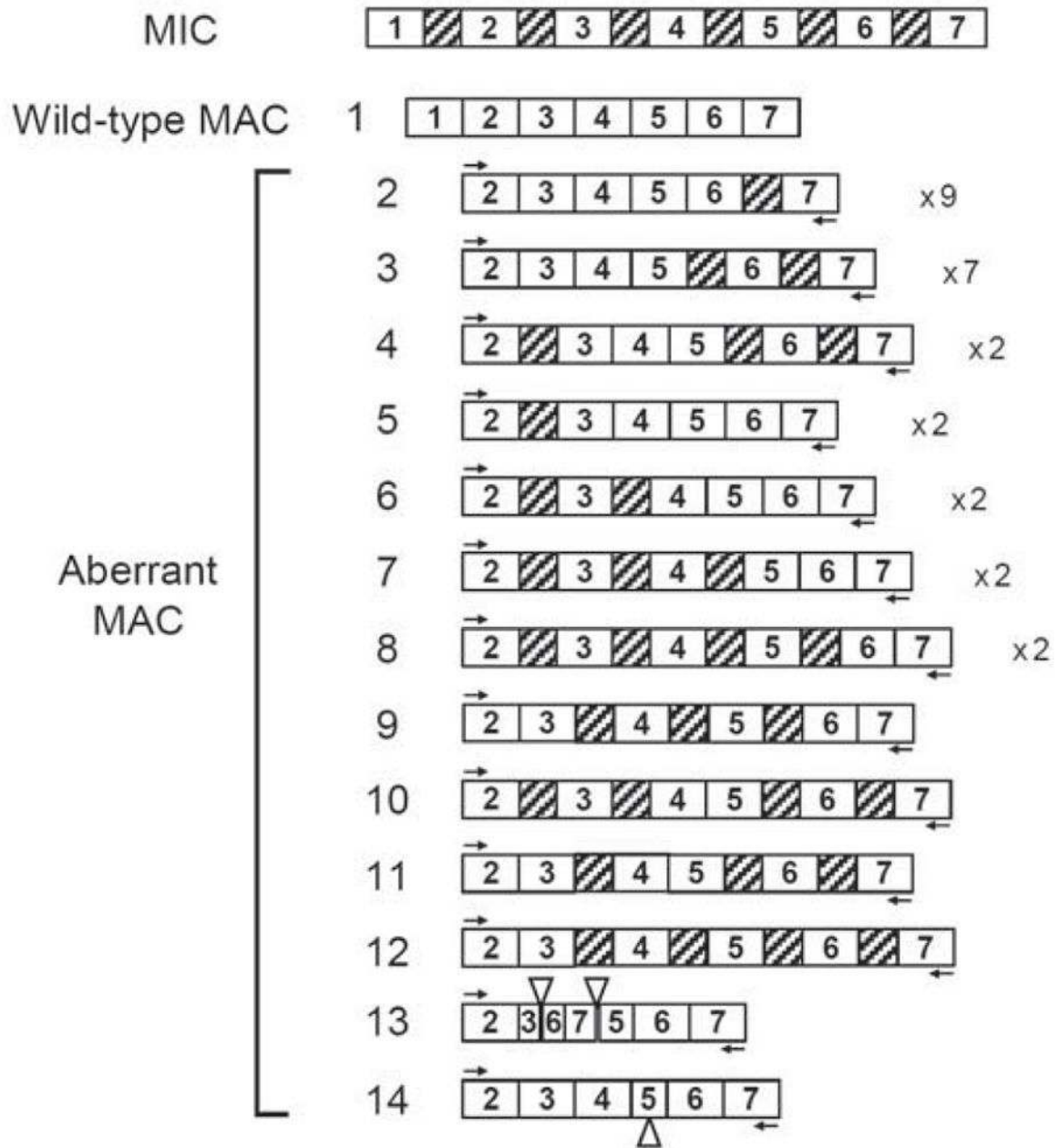
Fig. 2.
Silencing of TBE transposases leads to accumulation of aberrantly processed DNA. MDS segments (white boxes) and IESs (hatched boxes) are shown schematically and not to scale. Shown are the sequences of PCR products for *TEBPβ* between segments 2 and 7 in the progeny of triply silenced cells treated 12 to 15 hours after mixing; shown are all 32 examples of aberrantly rearranged chromosomes found among 37 sequenced clones [redundancy is indicated by the number after an x (table S2)]. Open triangles indicate locations of cryptic junctions between neighboring (pointing up) and nonneighboring (pointing down) segments on the basis of precursor micronuclear order. MIC, nonrearranged micronuclear gene sequences; MAC, expected rearranged macronuclear gene sequences.
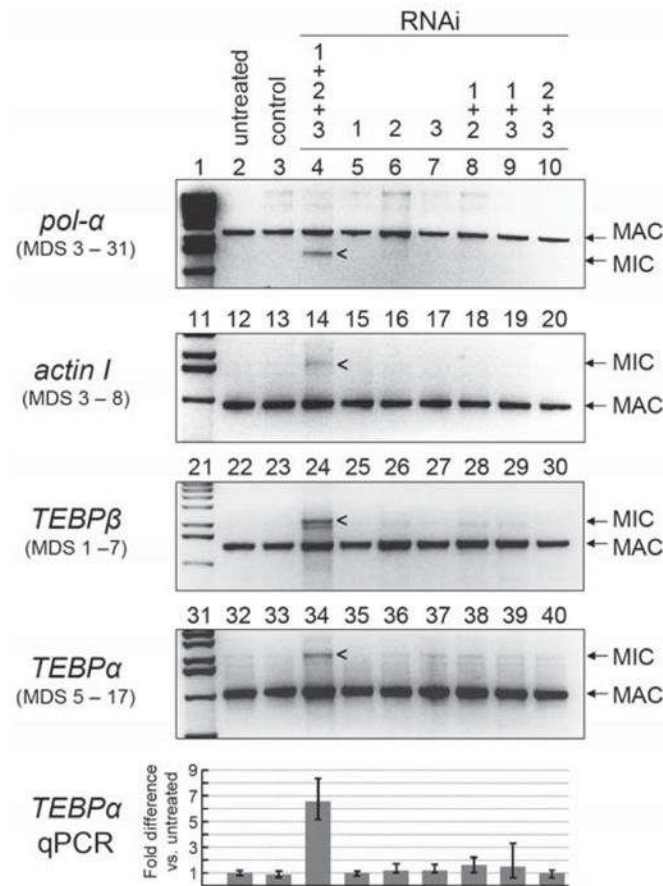
Fig. 3.
RNA interference (RNAi) against TBE transposases leads to accumulation ofnonprocessedDNA.PCR amplification of gene regions between two distant MDSs is shown (*pol-α* between MDSs 3 to 31; *actin I* between MDSs 3 to 8; *TEBPβ* between MDSs 1 to 7; and *TEBPα* between MDSs 5 to 17) (23). Samples that were triply silenced (12 to 15 hours after mixing) against all three TBE transposases (lanes 4, 14, 24, and 34) show accumulation of nonrearranged (MIC) gene versions. Controls were non-treated cells (lanes 2, 12, 22, and 32) and cells injected with 184-nucleotide vector polylinker dsRNA (lanes 3, 13, 23, and 33); MAC, macronuclear; MIC, micronuclear. Quantitative PCR (QPCR) with *TEBPα* primers shows 6.55-fold more nonrearranged DNA as compared with that of the untreated control in the triple silencing example (bottom).
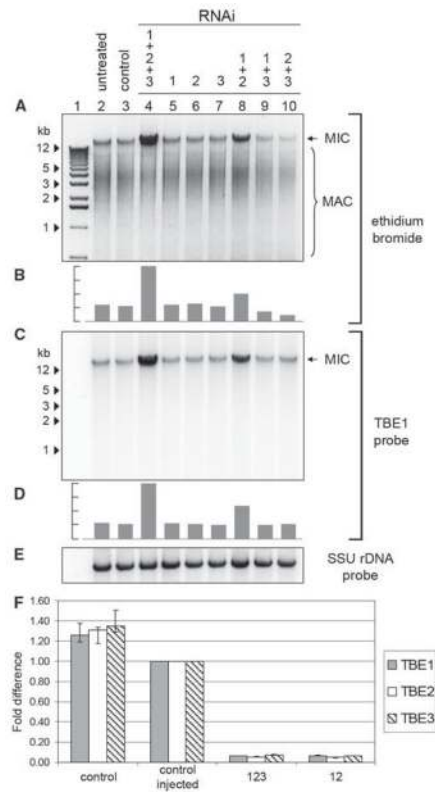
Fig. 4.
Transposase silencing leads to increased levels of nonprocessed high–molecular weight DNA and transposon retention. (**A**) 1 µg total DNA from vegetative progeny of cells treated with RNAi against TBE transposases (lanes 4 to 10) as well as untreated and control cells (23); ethidium bromide staining was used. (**C**) DNA from (A) hybridized to a TBE1 transposase DNA probe (23). Samples that were treated with RNAi (12 to 15 hours after mixing) against all three transposases or TBE1 and -2 together (lanes 4 and 8) show increased levels of (A) high–molecular weight DNA and (C) TBE1 hybridization to high–molecular weight DNA as well as decreased levels of transposase mRNA (F). (**B** and **D**) Gray bars indicate relative amount of ethidium bromide (B) and TBE1 probe signals (D) in high–molecular weight DNA. (**E**) Control hybridization with a small-subunit rDNA probe. Arrowheadsindicate size of relevant markers (1-kb ladder). (**F**) Relative transposase mRNA levels for TBE1, TBE2, and TBE3 in injected (12 to 15 hours after mixing) versus noninjected cells in an independent experiment (23). Absolute values for each transcript were obtained with QPCR (in femtograms) and normalized against the transcript level for the mitochondrial large-subunit rRNA. Transcript levels of the control injected sample were set to 1.