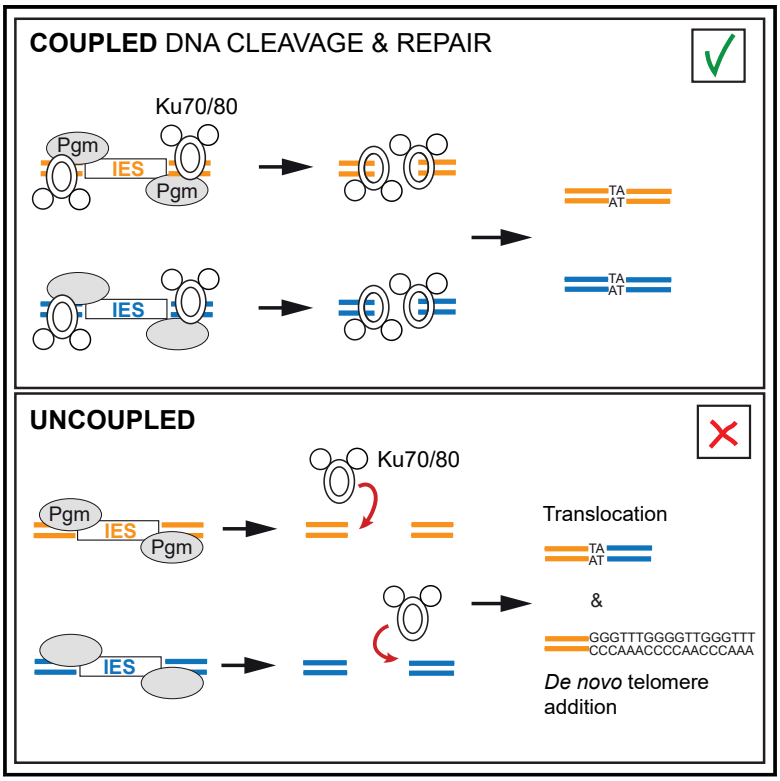
Uncoupling programmed DNA cleavage and repair scrambles the Paramecium somatic genome

Julien Bischerour, Olivier Arnaiz, Coralie Zangarelli, Vinciane Régnier, Florence Iehl, Virginie Ropars, Jean-Baptiste Charbonnier, Mireille Bétermier

Cell Reports

Uncoupling programmed DNA cleavage and repair scrambles the *Paramecium* somatic genome

Graphical abstract



Highlights

- A DNA-binding-deficient Ku mutant allows programmed DNA cleavage in *Paramecium*
- In the repair-deficient cells, broken DNA ends are trimmed or healed by telomere addition
- Uncoupling DNA cleavage and repair favors chromosome translocations
- Tight coupling ensures faithful genome assembly during programmed rearrangements

Authors

Julien Bischerour, Olivier Arnaiz, Coralie Zangarelli, ..., Virginie Ropars, Jean-Baptiste Charbonnier, Mireille Bétermier

Correspondence

julien.bischerour@i2bc.paris-saclay.fr (J.B.), mireille.betermier@i2bc.paris-saclay.fr (M.B.)

In brief

Bischerour et al. show the importance of the link between DNA double-strand break introduction and repair during programmed DNA elimination in *Paramecium*.When the two steps are uncoupled, proper somatic genome assembly and offspring survival are compromised by DNA translocations and *de novo* telomere addition.

Bischerour et al., 2024, Cell Reports *43*, 114001
April 23, 2024
https://doi.org/10.1016/j.celrep.2024.114001

CellPress

Cell Reports

CellPress
OPEN ACCESS

Report

Uncoupling programmed DNA cleavage and repair scrambles the *Paramecium* somatic genome

Julien Bischerour,[1,3,*] Olivier Arnaiz,[1] Coralie Zangarelli,[1] Vinciane Régnier,[1,2] Florence Iehl,[1] Virginie Ropars,[1] Jean-Baptiste Charbonnier,[1] and Mireille Bétermier[1,*]

[1]Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France
[2]Université Paris Cité, UFR Sciences du vivant, 75205 Paris Cedex 13, France
[3]Lead contact
*Correspondence: julien.bischerour@i2bc.paris-saclay.fr (J.B.), mireille.betermier@i2bc.paris-saclay.fr (M.B.)
https://doi.org/10.1016/j.celrep.2024.114001

SUMMARY

In the ciliate *Paramecium*, precise excision of numerous internal eliminated sequences (IESs) from the somatic genome is essential at each sexual cycle. DNA double-strands breaks (DSBs) introduced by the PiggyMac endonuclease are repaired in a highly concerted manner by the non-homologous end joining (NHEJ) pathway, illustrated by complete inhibition of DNA cleavage when Ku70/80 proteins are missing. We show that expression of a DNA-binding-deficient Ku70 mutant (Ku70-6E) permits DNA cleavage but leads to the accumulation of unrepaired DSBs. We uncoupled DNA cleavage and repair by co-expressing wild-type and mutant Ku70. High-throughput sequencing of the developing macronucleus genome in these conditions identifies the presence of extremities healed by *de novo* telomere addition and numerous translocations between IES-flanking sequences. Coupling the two steps of IES excision ensures that both extremities are held together throughout the process, suggesting that DSB repair proteins are essential for assembly of a synaptic precleavage complex.

INTRODUCTION

The genome must be preserved in its integrity throughout the life of an organism but must also adapt to its environment during evolution through changes in structure and sequence. Many somatic cell differentiation processes are accompanied by programmed DNA rearrangements that are driven by active "mutagenic" mechanisms, such as V(D)J recombination of immunoglobulin genes or programmed DNA elimination in nematodes.[1,2] This type of extreme genomic regulation is well illustrated in ciliates, in which two kinds of nuclei co-exist in the same cytoplasm[3]: the somatic macronucleus (MAC) is essential for gene expression but is destroyed at each sexual cycle, while the germline micronucleus (MIC) undergoes meiosis and transmits its genome to the zygotic nucleus. New MICs and MACs of sexual progeny differentiate from copies of the zygotic nucleus. Extensive genome rearrangements, together with multiple rounds of genome endoreplication, take place in the new MAC.

During this amplification, several thousands of short sequences and large genomic regions are eliminated via programmed genome rearrangements (PGRs).[4–6] Eliminations, accounting for ~30% of the germline genome content in species of the *Paramecium aurelia* group (up to 98% in some other ciliates), are initiated by the introduction of DNA double-strand breaks (DSBs), generally followed by the ligation of both flanking ends through the classical non-homologous end joining (cNHEJ) pathway.[4,7,8] Alternatively, telomeric repeats are added at the heterogeneous borders of large and imprecisely eliminated repeated sequences (mobile genetic elements, mini-satellites, etc.).[9]

The fidelity of rearrangements is essential for the formation of a functional new MAC. This relies on the introduction of precise DNA cuts at the nucleotide level and controlled repair of the resulting DNA extremities. The PiggyMac enzyme (Pgm), a domesticated transposase of the PiggyBac family, in combination with five Pgm-like proteins, catalyzes DSBs centered on the TA dinucleotide invariably present at all internal eliminated sequence (IES) boundaries.[8,10,11] After cleavage, the two DNA extremities generated on the genome have the same geometry (4-base 5′ overhangs) but are rarely fully compatible for direct rejoining. The final joining step involves prior processing of the ends, which includes the removal of one base at the 5' extremity, pairing the TA dinucleotides of each overhang, and filling the nucleotide gap.

The *Paramecium* genome encodes NHEJ repair proteins, such as Ku70, Ku80, Ligase4 (Lig4), and Xrcc4, as well as putative homologs of DNAPKcs and Xlf,[4,12,13] which form the core of the NHEJ machinery in eukaryotes.[14] The number of breaks and the fidelity of repair have raised questions about how the maturation steps are controlled to restore the sequence of the excision junction. How translocations of the ends to be repaired are proscribed, while thousands of breaks are generated concomitantly, remains to be addressed.

Depletion of Ku70/80 DNA repair proteins by RNAi has shown that in their absence, DSBs are no longer introduced,

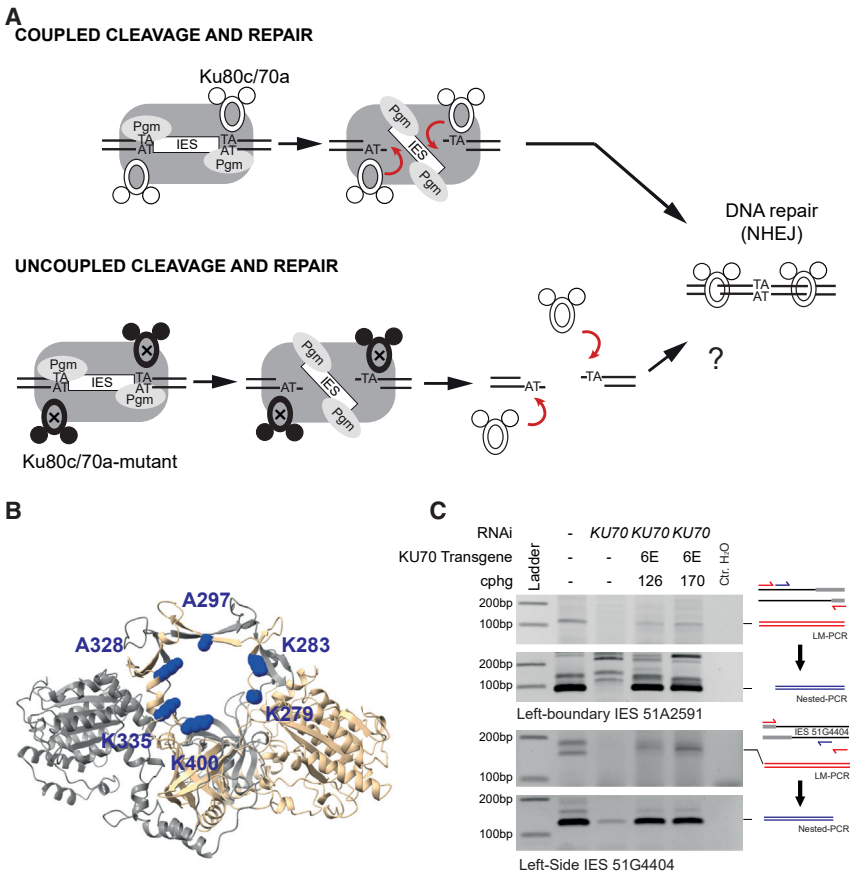


Figure 1. A Ku70 DNA-binding-deficient mutant to uncouple DNA cleavage and repair during PGR

(A) Strategy to uncouple DNA cleavage and repair.

(B) AlphaFold model of the Ku80c/Ku70a heterodimer highlighting the position of the 6E mutations (in blue) in the ring domain of Ku.

(C) Detection of DSBs by ligation-mediated PCR (LMPCR) using genomic DNA isolated from early (T5–T10) autogamous cells. The genomic left boundary of IES 51A2591 and the left extremity of IES 51G4404 are analyzed using adaptor oligonucleotides (gray) and PCR primers (in red) (top), followed by nested PCR (blue primers) (bottom).[15] See also Figures S1 and S2.

strongly suggesting that Ku proteins are an integral part of the excision machinery.[15] This phenotype is accompanied by the loss of Pgm anchoring in the nucleus.[16] This coupling, which conditions the introduction of breaks to the prior presence of repair actors, relies on a development-specific Ku80c paralog.[16] Although such a tight coupling has not been described in any other system, an interaction between Ku and proteins encoded by mobile genetic elements has been reported several times.[17–19] Ku proteins are also associated with the Rag1/Rag2 complex made of domesticated transposases that ensure the rearrangement of immunoglobulin loci during development, a process that involves the NHEJ pathway.[20]

Before conducting this study, we proposed that the coupling observed in *Paramecium* could favor faster recruitment of repair proteins and maintenance in close proximity of the correct ends to be ligated.[16,21] *In fine*, this would allow faithful assembly of chromosomes in the context of numerous DSBs throughout the genome.

To test this hypothesis, we addressed a simple question: what are the consequences of uncoupling DNA cleavage from its repair steps during genome rearrangements? This would correspond to the more usual situation of repair following accidental DSBs where breaks are introduced first and repair proteins are recruited in response to the presence of DNA ends. How will the quality of the junctions and, more generally, the quality of PGRs be affected in these conditions?

RESULTS

Design of a DNA-binding-deficient Ku70a mutant to uncouple DNA repair and DNA cleavage

The current model of coupling during IES excision postulates that Ku proteins, present within the excision complex when DSBs are introduced, are the molecules that are immediately loaded onto the free DNA ends (Figure 1A). Our strategy to uncouple DNA cleavage and DSB repair was to poison the process by expressing a second version of Ku, which is repair deficient but still able to activate Pgm. Under these conditions, the mutant Ku will activate DNA cleavage but fail to directly load on the DNA ends. If present in the cell, wild-type Ku proteins will then bind the ends and initiate the repair process.

A DNA-binding-deficient mutant of human Ku70 (Ku70-Mut6E) was characterized previously.[22] The *Paramecium* Ku70a sequence (PtKu70a), which we will call Ku70 for simplicity, was aligned with the human Ku70 sequence (Figure S1). PtKu70a K279, K283, A297, K328, K335, and K400 were identified as counterparts of the K282, K287, T300, K331, K338, and R403 residues of human Ku70. Their positions in the DNA-binding ring domain are supported by the predicted modeled structure of PtKu70/80c (Figure 1B). Recombinant Ku70 or Ku70-6E was expressed in insect cells together with the development-specific Ku80c subunit, and the DNA-binding deficiency of the mutant was confirmed by electrophoretic mobility shift assay using a 35-bp DNA substrate (Figure S2).

We then performed an *in vivo* complementation assay and demonstrated that cells transformed by an RNAi-resistant wild-type *FLAG-KU70* construct provide viable sexual progeny after autogamy when subjected to RNAi against endogenous *KU70*. In contrast, microinjection of the *FLAG-KU70-6E* mutant fails to complement depletion of the endogenous wild-type protein, resulting in death of the sexual progeny,

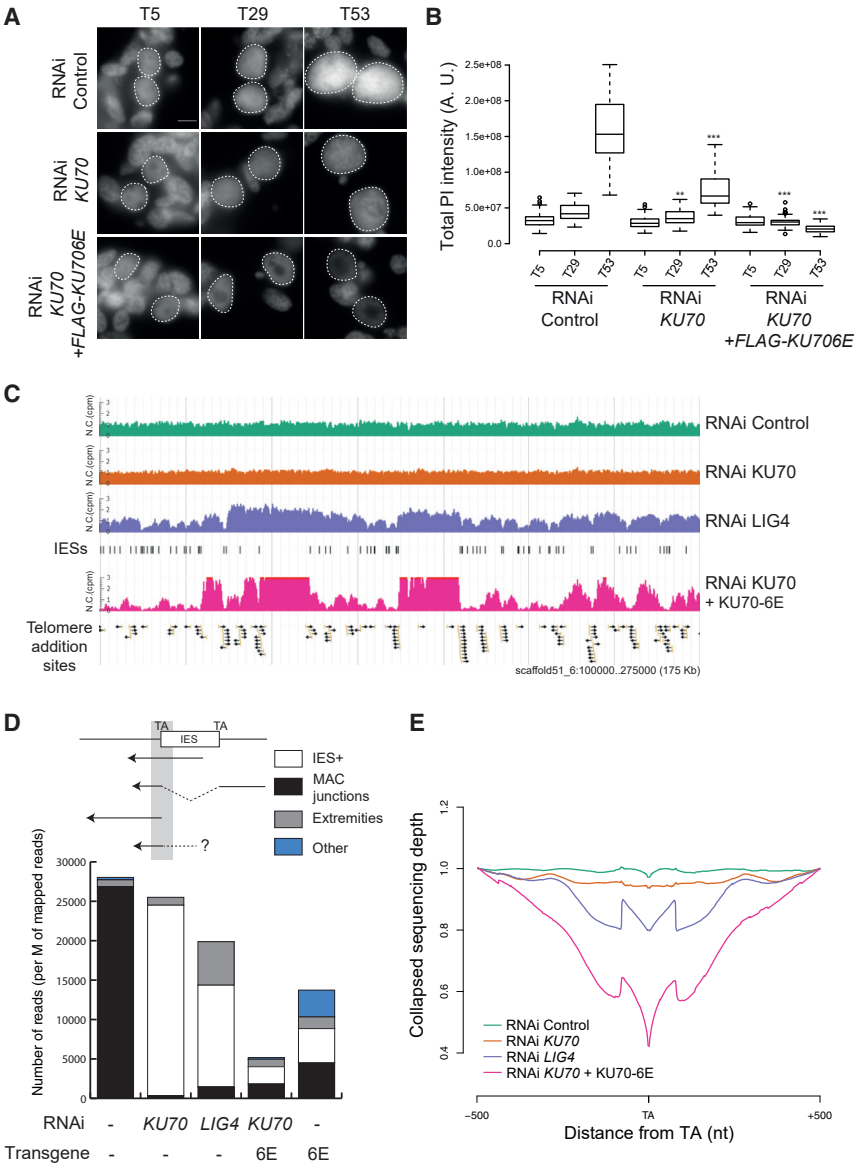


Figure 2. Ku70-6E is defective for DNA repair

(A) Propidium iodide (PI) staining of genomic DNA content in the developing MACs of control cells, *KU70*-silenced cells, and FLAG-KU70-6E transformants subjected to *KU70* RNAi. Scale bar: 5 μm.

(B) Quantification of total PI intensity in developing MACs. 44–59 developing MACs were analyzed for each condition. The statistical test (Mann-Whitney-Wilcoxon) compares the *KU70* RNAi and FLAG-Ku70-6E conditions at T5, T29, and T53 to their respective control RNAi at T5, T29, and T53. $^{**}p < 0.01$ and $^{***}p < 0.001$. The analysis of anlagen sizes is presented in Figure S3. T5 = 5 hours after begining of autogamy

(C) Screenshot of read coverage along the *Paramecium* MAC genome (position bp 100,000–275,000 on scaffold51_6). Genomic DNA was purified from sorted new developing MACs at T25–T30 after the beginning of autogamy. IES excision sites are shown as vertical bars. Telomere addition sites are represented by small arrows. In each condition, the sequencing depth was normalized by the total number of mapped reads (cpm = counts per million).

(D) Analysis of the reads covering IES excision sites (using the MEND module of ParTIES). Arrows represent 75 bp *Illumina* reads from paired-end sequencing data, which can be mapped on the positions of IES excision sites (TA dinucleotide). They correspond to (1) the genomic loci with an IES (IES+), (2) an already excised IES (MAC junction), (3) reads starting at the TA (+ or −2 bp) (extremities), and (4) reads stopping on the TA dinucleotide (+ or −2 bp) (other). See Figure S4 for analysis of replicate.

(E) Analysis of collapsed sequencing depth in the 1 kb region surrounding IES excision sites.

See also Figure S3.

even when expressed at similar or higher levels relative to wild type (Figure S2).

Ku70-6E supports Pgm-dependent DNA cleavage

The expression of Ku70/Ku80c is essential for anchoring the excision machinery in the new developing MAC and for the DNA cleavage activity of Pgm.[16] We found that whereas the Pgm signal vanished from the new MAC, when Ku70 was depleted by RNAi as previously reported,[16] the expression of Ku70-6E restored the nuclear Pgm signal to a level close to control conditions (Figure S2). To test the DNA cleavage activity of Pgm in the presence of Ku70-6E, total genomic DNA was isolated from cells 5 h after the beginning of autogamy (T5) when DSBs were introduced.[10,23] The presence of cleaved DNA ends was first tested by ligation-mediated PCR (LMPCR) for two previously well-characterized excision sites.[24] In wild-type cells, but not in Ku70-depleted cells, cleaved DNA extremities were detected as intermediate products of IES elimination (Figure 1C). In contrast, DSBs were detected at a similar level to the wild type in Ku70-6E-expressing cells, which indicates that the Ku70-6E mutant supports Pgm activity and that Ku's ability to bind DNA extremities is dispensable for DNA cleavage.

KU70-6E expression impairs DNA endoreplication in the new MAC

Using propidium iodide (PI) staining, we previously observed that unrepaired DSBs accumulating in Lig4-depleted cells lead to a severe defect in DNA amplification.[12] In control cells, the development of the new MACs until T53 (T53 = 53 h after begining of autogamy) correlates with intense PI staining, as expected for a high level of endoreplication[23] (Figures 2A and 2B). In contrast, when Ku70 is replaced by the DNA-binding-deficient Ku70-6E mutant, the intensity of the PI signal does not increase at T29 and T53. Concomitantly, growth of the new MACs is also severely impaired, and large DNA-depleted regions are observed in these nuclei (Figures 2A and S3). Our data also

indicate that Ku70 depletion leads to the retention of all IESs and a significant reduction of DNA amplification in the new MAC at late stages. The latter phenotype has not been documented in other conditions of IES retention, suggesting that it might be linked to a Ku-related replication phenotype. Altogether, our observations are consistent with the inability of the Ku70-6E mutant to repair DSBs after licensing of Pgm DNA cleavage, resulting in DNA amplification failure.

To gain global insight into the molecular consequences of Ku70-6E expression on genome assembly, we purified the new developing MACs of cells collected at T25–T30 using a fluorescence-assisted nuclear sorting (FANS) approach.[23] Cells were exposed to control, *LIG4*, or *KU70* RNAi in the presence or absence of the *KU70-6E* transgene. Quantification of the PI signal in the FANS-purified new developing MACs confirms that both *LIG4*-silenced cells and Ku70-6E-expressing cells have lower DNA content than control and *KU70*-silenced cells (Figure S3). Sequencing of the new developing MAC genome of Ku70-6E-expressing cells shows that the read coverage along the MAC reference genome is highly heterogeneous (Figures 2C and S3). IES-free fragments of the genome were highly covered compared to IES-dense regions. Moreover, IES excision sites were invariably found at the boundaries of amplified regions. To extend this observation to the whole genome, we calculated the average sequencing depth for each inter-IES fragment (Figure S3). Above 2 kb, the level of endoreplication is correlated with fragment size: the larger the inter-IES fragment, the more it is amplified. This mirrors a previous observation that was made for injected transgenes in *Paramecium*: they must be long enough to be efficiently maintained as mini-chromosomes through vegetative divisions.[25,26] In *Paramecium*, although the mechanism of DNA replication initiation is unknown, longer fragments appear to provide more opportunities for replication.

Ku70-6E is defective for DNA DSB repair

To better decipher the origin of read coverage heterogeneity, we developed MEND, a new module of the ParTIES pipeline,[27] to analyze in more detail the DNA reads overlapping the TA dinucleotide at IES excision sites (Figure 2D). Whereas most reads correspond to MAC junctions in the control, *KU70*-knockdown cells—in which Pgm activity is abolished—yield almost exclusively IES-containing reads. In *LIG4*-knockdown cells or in cells expressing the KU70-6E mutant, the normalized numbers of reads mapping to the TA are reduced by 30% and 80%, respectively, confirming the specific loss of the genomic sequences surrounding IES excision sites. The disappearance of these sequences correlates with an increase in the number of reads starting at the TA position (+ or – 2 bases, "extremities" in Figure 2D), in agreement with the presence of unrepaired DSBs specifically at this position. We extended the analysis of read coverage in the vicinity of IESs by collapsing all the DNA reads mapping −500 to +500 bases from the TA dinucleotides at the excision sites (Figure 2E). Compared to control, the read coverage decreases progressively in Lig4-depleted and Ku70-6E-expressing cells, suggesting possible trimming of DNA extremities after DSBs are introduced at these positions. The two peaks of read coverage positioned 75 bp (= size of DNA reads) away from the TA disappeared when extremities reads were omitted, suggesting they are due to the overrepresentation of sequenced fragments starting exactly at the TA ("extremities" in Figures 2D and S3). We conclude that the heterogeneity of genome coverage and the strong deficit in IES-containing regions observed in *LIG4*-silenced or Ku70-6E-expressing cells are the consequences of a replication defect and the partial degradation of genomic sequences consecutive to unrepaired DSBs.

Together with LMPCR results, this supports the notion that expressing Ku70-6E in place of wild-type Ku70 enables Pgm-dependent cleavage at IES boundaries but not DSB repair.

Chromosome healing by *de novo* telomere deposition

De novo telomere addition is part of PGR as a consequence of imprecise DNA elimination at chromosome fragmentation sites.[9,28] Unrepaired IES excision has been proposed to provide another opportunity for *de novo* telomere addition,[29] an alternative DSB repair pathway in many organisms.[30] Therefore, we looked for *de novo* telomere addition in the new developing MAC genome by counting the reads containing MAC DNA sequences linked to telomeric repeats (Figures 3A and S4). These telomeric DNA reads are strongly enriched in Ku70-6E-expressing cells and in Lig4-depleted cells; they are distributed all along the scaffolds and are overrepresented at the borders of amplified segments, which correspond to IES excision sites (Figure 2C). Only 17% of telomeres are added at a short distance from the TA (<10 nucleotides), indicating that the ends generated by Pgm cleavage tend to be trimmed before telomere addition when wild-type Ku is depleted (Figures 3A and S4). We also observed telomere addition at the −1, T, and A positions, suggesting that the 3′ recessive ends generated by initial Pgm cleavage (−2 position) have been filled in by a DNA polymerase before telomere addition (Figure 3B). In Lig4-depleted cells, we observed that telomeres tend to branch closer to the TA (Figure 3A) with a preponderant addition site at −2 and almost no telomeres added at the −1, T, and A positions (Figure 3B). These observations suggest that the DNA-binding activity of Ku is dispensable for the recruitment or the activation of the NHEJ DNA polymerase at cleaved extremities. In contrast, Lig4 appears to play an important role in the control of 3′ end processing.

Uncoupling DNA cleavage and repair induces numerous translocations

To uncouple the two steps of IES excision, we permitted the expression of both wild-type and mutant Ku70 proteins (*KU70-6E* transgene without RNAi). We first analyzed the survival of the progeny of transformants harboring increasing levels of the *FLAG-KU70-6E* transgene (Figure 4A). At low injection levels (<100 cphg, copies per haploid genome), we observed no phenotype with regard to progeny survival, whereas higher levels led to the death of the entire progeny. We selected the two most injected transformants (253 and 470 cphg) and one transformant (142 cphg) that provided 50% progeny survival after autogamy. For the three transformants, we deep-sequenced total genomic DNA from cells collected 50 h after the beginning of autogamy (total DNA in Figure S4 and Table S3). In addition, we sequenced DNA from sorted new MACs of the 253 cphg transformant. In the

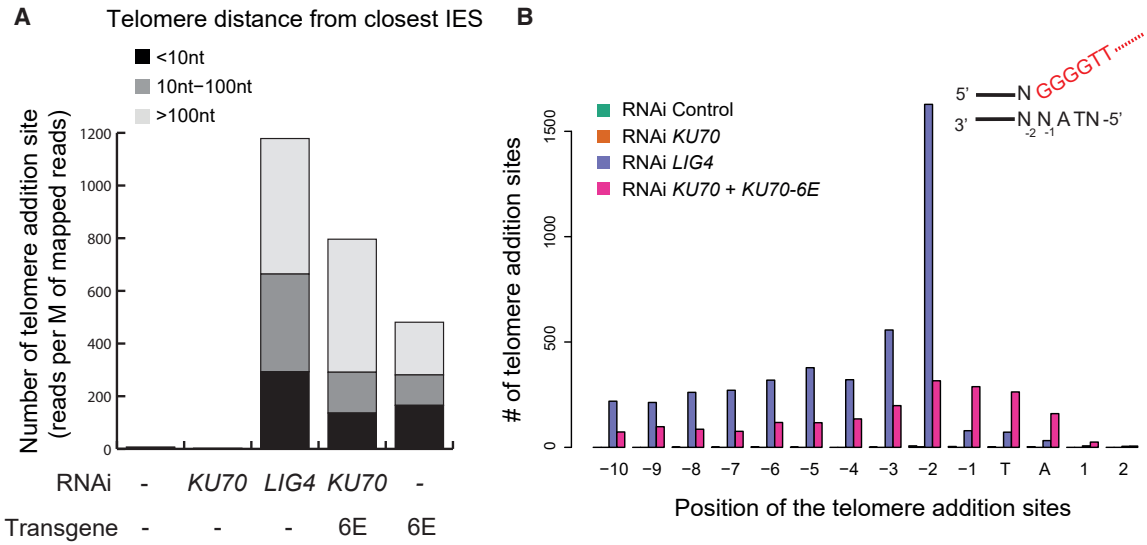


Figure 3. DSB repair by *de novo* telomere addition

(A) Abundance of *de novo* telomere addition sites. A telomere addition site was pinpointed whenever a read alignment on the MAC sequence was interrupted, and the sequence proceeds with telomeric repeats (G_4T_2 or G_3T_3).[23] The number of reads was normalized by the total number of mapped reads (in millions [M]).
(B) Position of telomere addition sites. The position corresponds to the last mapped nucleotide on the genome sequence. To avoid ambiguity, IESs flanked by a guanine at the −1 or −2 positions were not considered in this analysis.
See also Figure S4.

presence of both wild-type Ku70 and the 6E mutant, the presence of correctly assembled MAC junctions demonstrates that IESs are partially excised (in black in Figure 2D). Interestingly, a significant fraction of reads (classified as "other") were found to map only partially on the MAC genome, and the mapping stopped on the TA of IES excision sites. Using an improved version of the MILORD module of ParTIES,[27] the unmapped sequences of these reads can be remapped successfully on another part the genome, suggesting that they result from chromosome translocation events. In control non-injected cells (no RNAi), only rare translocations are detected (Figures 4B and S4). They implicate sequences that are distant from IES excision sites (>100 bp) and are characterized by frequent microhomologies of up to 15 bases at the junction (Figure S5). They are also detected in *KU70* and *LIG4* RNAi samples, indicating that they are likely caused by a pathway distinct from cNHEJ (Figure S5).

The numerous translocations generated in uncoupling conditions involved IES-flanking sequences (Figures 4B and S5). In contrast to the above control, these translocations rely on the cNHEJ pathway since they are not observed in the absence of endogenous Ku70. The most frequent translocations bring together two sequences flanking different IESs joined on a TA dinucleotide (dark blue) or one IES-flanking sequence joined to an excised IES (light blue) (Figure 4B). Alternatively, the second extremity may have been trimmed before the ligation (gray). Other rare translocations with differently processed DNA ends could be detected at low frequency (Figure S5). These events were also observed in other transformants (Figure S5).

Taken together, these results demonstrate that coupling between DNA cleavage and DSB repair is crucial to maintain the two flanking MAC-destined DNA extremities in close proximity throughout IES excision. Moreover, the existence of translocations with imprecise junctions suggests that coupling the two reactions may limit the processing of DNA extremities prior to ligation.

DISCUSSION

In this study, we have addressed the necessity of the coupling that renders the DNA cleavage activity of the IES excision complex entirely dependent on the presence of the DNA repair proteins Ku70 and Ku80 by measuring the consequences of its abolition. We first validated that a Ku70 mutant unable to bind free DNA ends can activate the cleavage machinery but fails to repair DSBs. In cells depleted of endogenous Ku70, the expression of this mutant results in heterogeneous amplification of a fragmented version of the somatic genome. Consequently, we observed the addition of multiple telomeres and the amplification of extended IES-free subregions of the genome that are not exposed to programmed DSBs.

Distinct recruitment of NHEJ partners in Lig4-depleted cells and Ku70-6E-expressing cells

Both Lig4 depletion and the replacement of Ku70 by its 6E mutant lead to the accumulation of unrepaired DSBs across the genome. However, DNA extremities are subsequently processed in two different DNA repair backgrounds formally designated as Lig4−/Ku+ and Lig4+/Ku70−, respectively, which could explain the observed differences in genome amplification (Figure 2) and telomere addition pattern (Figure 3). In the Ku-defective background, we noted more loss of read coverage in the vicinity of DSBs, correlating with telomere deposition at a greater distance from the breaks. This suggests that free DNA ends are exposed to extensive trimming in the absence of

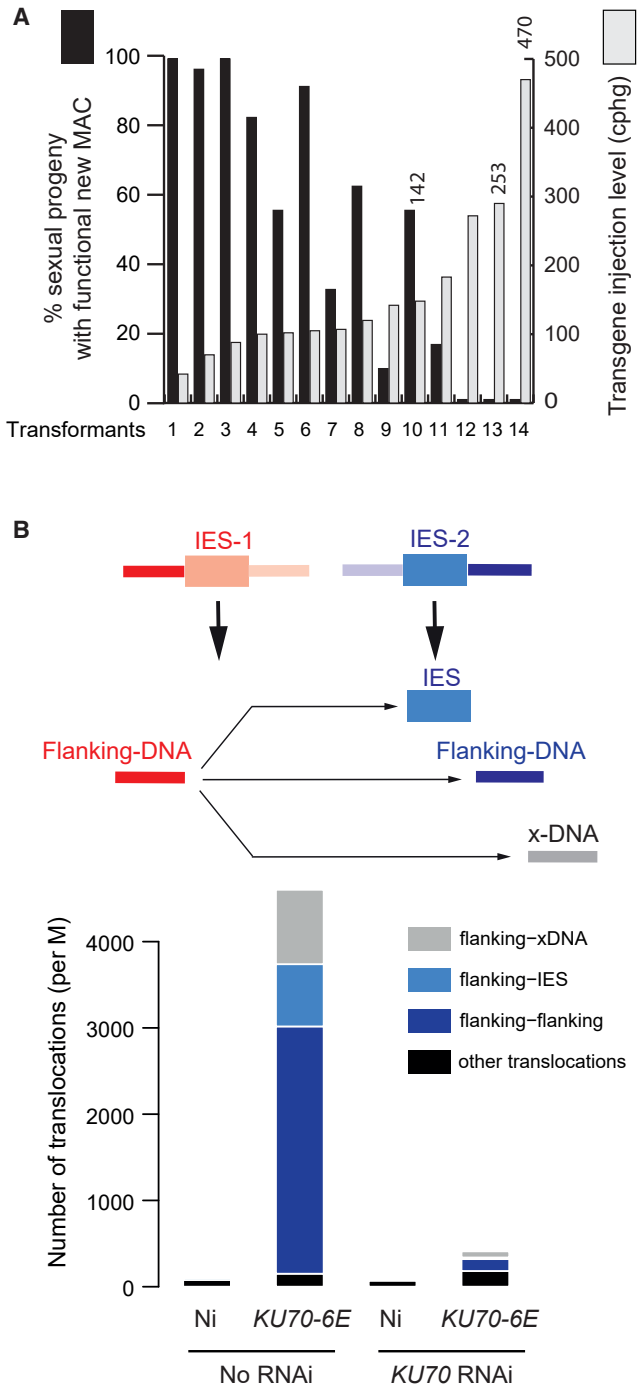


Figure 4. Uncoupling of DNA cleavage and repair induces genomic translocations

(A) Sexual progeny survival assays of *FLAG-KU70a-6E* transformants under control (no RNAi) conditions. For each transformant, the percentage of progeny with a functional new MAC (black bars) and the qPCR quantification of the number of transgenes (cphg) (white bars) are presented (see Table S1 for details).

(B) Quantification of genomic translocation events and scheme of the most frequent translocations observed upon uncoupling. Translocations implicating sequences distant by more than 2 nucleotides away from the TA are considered as xDNA (gray).

See also Figure S5.

DNA-bound Ku. Precise analysis of the positions of telomere addition on untrimmed ends points to differential recruitment of NHEJ factors at the extremities. In both conditions, only a fraction of telomeric repeats are added to the 3′OH at position −2, which is the extremity generated by Pgm cleavage. Telomere repeats branched at the −1, T, and A positions indicate that filling of the 3′ recessive extremities has preceded telomere addition. Their absence in the Lig4 background is consistent with the proposed role of the Lig4/Xrcc4 complex in recruiting the NHEJ-dedicated DNA polymerase.[12]

Moreover, IES retention is greater upon *LIG4* RNAi. We cannot formally rule out that Lig4 plays a role in the activation of DNA cleavage, but an indirect effect is more likely. Indeed, the unrepaired DNA extremities that accumulate in the absence of Lig4 may consequently trap endogenous Ku70/80 complexes and reduce the available pool of unbound Ku that is necessary to activate further DNA cleavages. In agreement with this hypothesis, we observed that late excised IESs are more retained than the early ones[23] (Figure S5). In contrast, the Ku70-6E protein, which is defective for binding to DNA extremities, will not be trapped on DNA and will remain available to license Pgm-mediated cleavage.

Coupling preserves synapsed DNA ends from the DNA cleavage step to DSB repair

During IES excision, both the cleavage of DNA termini and the subsequent NHEJ-mediated repair of the two broken ends, rely on the formation of a synaptic complex. PiggyBac transposases, and probably all cut-and-paste DNA transposases, including Rag1/2, establish contacts between transposon extremities before DSB introduction.[2,31,32] Similarly, IES excision in *Paramecium* also involves a crosstalk between the two IES ends before DNA cleavage.[33] Interaction of transposases with NHEJ factors like Ku70/80 has been documented for many transposons, suggesting that NHEJ is a favored DNA repair pathway during transposition.[17–20]

We propose that the main function of coupling the two steps of the IES excision reaction is to ensure continuity between the synaptic complex that is established between both IES ends before DSB introduction and the synaptic complex that is formed between the two flanking broken ends during DNA repair (Figure 5). Pgm on its own might be deficient for the assembly of the DNA cleavage synaptic complex and might have become dependent on DNA repair factors to properly hold together and cleave IES boundaries. Although this model is attractive, it implies that Ku70/80 are recruited to the presynaptic complex before DSB introduction: this may be mediated by a direct interaction with Pgm or other partners of the cleavage complex[15] or through binding to uncleaved DNA. How then would Ku stimulate the bridging of the two IES ends during cleavage? In DNA repair synaptic complexes, Ku recruits partners, including DNAPKcs, Xlf, Xrcc4/Lig4, and Paxx, that bridge the two DNA extremities.[14,34–38] Interestingly, two *Paramecium* paralogs called Die5a and Die5b (for deficient in IES excision) share structural homology with human Paxx.[39] Their corresponding genes are induced during the sexual cycle, and RNAi-mediated *DIE5* knockdowns lead to death of sexual progeny and to the retention of several IESs examined by molecular approaches.[40]

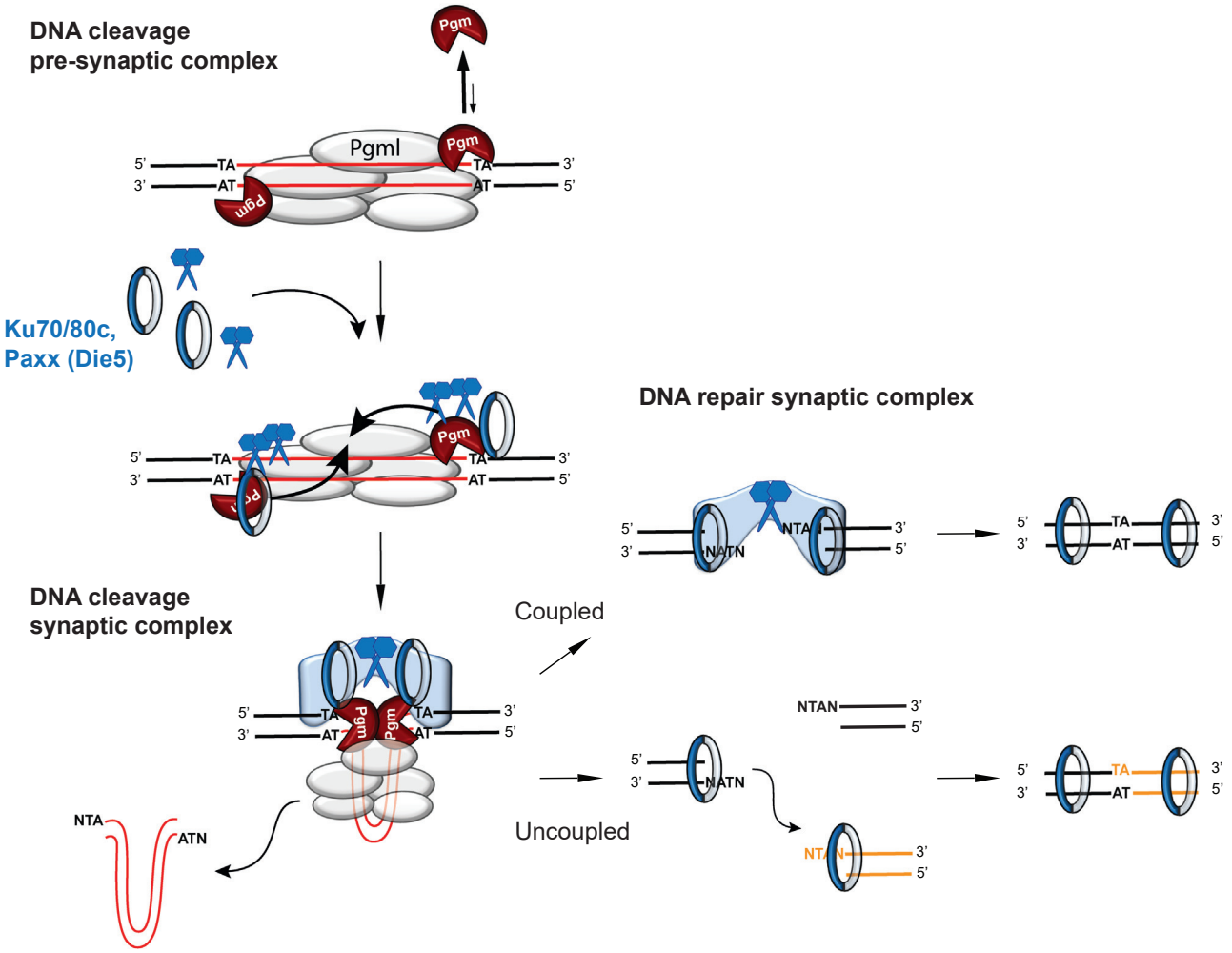


Figure 5. Model of DNA cleavage synaptic complex assembly with DNA repair proteins

Whether Die5 is the functional *Paramecium* homolog of Paxx, and whether its depletion induces IES retention genome-wide, remains to be explored. However, its properties support the hypothesis that a DNA repair complex is partially assembled before DSB introduction and may contribute to licensing DNA cleavage by bridging the two IES boundaries in the DNA cleavage synaptic complex (Figure 5).

The contribution of repair factors to DNA cleavage during IES elimination contrasts with V(D)J recombination. In the latter, the Rag1/2 recombinases introduce DNA breaks independently of DNA repair factors. Furthermore, they help to maintain the two broken DNA ends within the DNA repair synaptic complex, as indicated by the observation that destabilizing the Rag1/2 complex via the deletion of Rag2 subdomains renders the DNA repair step sensitive to the absence of Xlf.[41] Due to the massive scale of programmed DSBs during MAC development in *Paramecium*, the interaction between the domesticated transposase and the NHEJ factors has been pushed to an extreme by rendering the repair factors indispensable for the activation of DNA cleavage. Comparing the two systems will further highlight the multifaceted interactions that have been established during the co-evolution of DSB-inducing enzymes and DNA repair machineries.

Limitations of the study

We have uncoupled the two steps of the DNA elimination reaction by co-expressing wild-type Ku with a DNA-repair-deficient mutant. Since the wild-type proteins alone are capable of ensuring the normal course of events (break introduction and repair), we anticipated that only a fraction of the observed events would correspond to the uncoupling scenario. Therefore, our results likely underestimate the actual effect of complete uncoupling. Furthermore, because wild-type and mutant Ku70 proteins required Ku80c as a partner, overexpression of the mutant will reduce the available amount of Ku80c; the more mutant Ku70 proteins are expressed, the less wild-type Ku70/Ku80c complexes are assembled. As a consequence, DNA repair is abolished. Whether DNA damage response is induced or not during PGR in *Paramecium* is unknown. One can imagine that because the system has evolved toward coupling, the DNA damage response has become ineffective.

The present study does not provide elements that might support this hypothesis.

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 STUDY PARTICIPANT DETAILS
  - *Paramecium* strains and culture conditions
- METHOD DETAILS
  - Gene knockdowns during autogamy
  - Transgene construction, micro-injection and protein expression analysis
  - Cell staining
  - Purification and DNA binding assays using recombinant Ku70/80 proteins expressed in insect cells
  - Detection of DNA extremities by ligation mediated PCR (LMPCR)
  - Purification of new developing MACs, sequencing of genomic DNA and mapping on the *paramecium* genomes
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Statistical analysis
  - Bioinformatic analyses
  - MEND analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114001.

ACKNOWLEDGMENTS

The authors thank Linda Sperling, Joël Acker, and Christophe Possoz for their critical reading of the manuscript and Florian Meurisse for technical help. This work was supported by the Centre National de la Recherche Scientifique (CNRS) (intramural funding to M.B.), the French National Agency for Research (ANR) (grant no. ANR-21-CE12-0019 "CURE" to M.B. and J.-B.C.), the Fondation pour la Recherche Médicale (grant FRM EQU202103012766 to M.B.), the ARC Foundation for Cancer Research (grant no. PJA_A2020_CA2020121 to J.B.), and the I2BC (grant « SUBOKU » to J.B.). We acknowledge the sequencing and bioinformatics expertise of the I2BC High-Throughput Sequencing facility, supported by France Génomique (funded by the French National Program "Investissement d'Avenir" ANR-10-INBS-09). FANS experiments benefited from the expertise of the Imagerie Gif core facility, supported by the Agence Nationale de la Recherche (ANR-11-EQPX-0029/Morphoscope, ANR-10-INBS-04/FranceBioImaging, and ANR-11-IDEX 0003-02/Saclay Plant Sciences). J.B.C. thanks FRISBI for insect cell expression facilities (ANR-10-INBS-0005).

AUTHOR CONTRIBUTIONS

Conceptualization, J.B.; methodology, J.B. and O.A.; investigation, J.B., O.A., C.Z., V. Régnier, F.I., and V. Ropars; software, O.A.; data curation, O.A.; writing – original draft, J.B. and M.B.; writing – review & editing, J.B., O.A., V. Régnier, J.-B.C., and M.B.; funding acquisition, J.B., J.-B.C., and M.B.; resources, J.B., F.I., V. Ropars, J.-B.C., and M.B.; supervision, J.B., V. Ropars, and M.B.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 6, 2023
Revised: January 24, 2024
Accepted: March 11, 2024
Published: March 27, 2024

REFERENCES

1. Estrem, B., and Wang, J. (2023). Programmed DNA elimination in the parasitic nematode Ascaris. PLoS Pathog. *19*, e1011087. https://doi.org/10.1371/journal.ppat.1011087.
2. Schatz, D.G. (2004). V(D)J recombination. Immunol. Rev. *200*, 5–11. https://doi.org/10.1111/j.0105-2896.2004.00173.x.
3. Prescott, D.M. (1994). The DNA of ciliated protozoa. Microbiol. Rev. *58*, 233–267. https://doi.org/10.1128/mr.58.2.233-267.1994.
4. Betermier, M., and Duharcourt, S. (2014). Programmed Rearrangement in Ciliates: Paramecium. Microbiol. Spectr. *2*. https://doi.org/10.1128/microbiolspec.MDNA3-0035-2014.
5. Yao, M.-C., Chao, J.-L., and Cheng, C.-Y. (2014). Programmed Genome Rearrangements in Tetrahymena. Microbiol. Spectr. *2*. https://doi.org/10.1128/microbiolspec.MDNA3-0012-2014.
6. Yerlici, V.T., and Landweber, L.F. (2014). Programmed Genome Rearrangements in the Ciliate Oxytricha. Microbiol. Spectr. *2*. https://doi.org/10.1128/microbiolspec.MDNA3-0025-2014.
7. Sellis, D., Guérin, F., Arnaiz, O., Pett, W., Lerat, E., Boggetto, N., Krenek, S., Berendonk, T., Couloux, A., Aury, J.-M., et al. (2021). Massive colonization of protein-coding exons by selfish genetic elements in Paramecium germline genomes. PLoS Biol. *19*, e3001309. https://doi.org/10.1371/journal.pbio.3001309.
8. Arnaiz, O., Mathy, N., Baudry, C., Malinsky, S., Aury, J.-M., Denby Wilkes, C., Garnier, O., Labadie, K., Lauderdale, B.E., Le Mouël, A., et al. (2012). The Paramecium germline genome provides a niche for intragenic parasitic DNA: evolutionary dynamics of internal eliminated sequences. PLoS Genet. *8*, e1002984. https://doi.org/10.1371/journal.pgen.1002984.
9. Le Mouël, A., Butler, A., Caron, F., and Meyer, E. (2003). Developmentally regulated chromosome fragmentation linked to imprecise elimination of repeated sequences in paramecia. Eukaryot. Cell *2*, 1076–1090. https://doi.org/10.1128/EC.2.5.1076-1090.2003.
10. Baudry, C., Malinsky, S., Restituito, M., Kapusta, A., Rosa, S., Meyer, E., and Bétermier, M. (2009). PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements in the ciliate Paramecium tetraurelia. Genes Dev. *23*, 2478–2483. https://doi.org/10.1101/gad.547309.
11. Bischerour, J., Bhullar, S., Denby Wilkes, C., Régnier, V., Mathy, N., Dubois, E., Singh, A., Swart, E., Arnaiz, O., Sperling, L., et al. (2018). Six domesticated PiggyBac transposases together carry out programmed DNA elimination in Paramecium. Elife *7*, e37927. https://doi.org/10.7554/eLife.37927.
12. Kapusta, A., Matsuda, A., Marmignon, A., Ku, M., Silve, A., Meyer, E., Forney, J.D., Malinsky, S., and Bétermier, M. (2011). Highly precise and developmentally programmed genome assembly in Paramecium requires ligase IV-dependent end joining. PLoS Genet. *7*, e1002049. https://doi.org/10.1371/journal.pgen.1002049.
13. Lees-Miller, J.P., Cobban, A., Katsonis, P., Bacolla, A., Tsutakawa, S.E., Hammel, M., Meek, K., Anderson, D.W., Lichtarge, O., Tainer, J.A., and Lees-Miller, S.P. (2021). Uncovering DNA-PKcs ancient phylogeny, unique

sequence motifs and insights for human disease. Prog. Biophys. Mol. Biol. *163*, 87–108. https://doi.org/10.1016/j.pbiomolbio.2020.09.010.

14. Pannunzio, N.R., Watanabe, G., and Lieber, M.R. (2018). Nonhomologous DNA end-joining for repair of DNA double-strand breaks. J. Biol. Chem. *293*, 10512–10523. https://doi.org/10.1074/jbc.TM117.000374.
15. Marmignon, A., Bischerour, J., Silve, A., Fojcik, C., Dubois, E., Arnaiz, O., Kapusta, A., Malinsky, S., and Bétermier, M. (2014). Ku-mediated coupling of DNA cleavage and repair during programmed genome rearrangements in the ciliate Paramecium tetraurelia. PLoS Genet. *10*, e1004552. https://doi.org/10.1371/journal.pgen.1004552.
16. Abello, A., Régnier, V., Arnaiz, O., Le Bars, R., Bétermier, M., and Bischerour, J. (2020). Functional diversification of Paramecium Ku80 paralogs safeguards genome integrity during precise programmed DNA elimination. PLoS Genet. *16*, e1008723. https://doi.org/10.1371/journal.pgen.1008723.
17. Izsvák, Z., Stüwe, E.E., Fiedler, D., Katzer, A., Jeggo, P.A., and Ivics, Z. (2004). Healing the wounds inflicted by sleeping beauty transposition by double-strand break repair in mammalian somatic cells. Mol. Cell *13*, 279–290. https://doi.org/10.1016/s1097-2765(03)00524-0.
18. Jin, Y., Chen, Y., Zhao, S., Guan, K.-L., Zhuang, Y., Zhou, W., Wu, X., and Xu, T. (2017). DNA-PK facilitates piggyBac transposition by promoting paired-end complex formation. Proc. Natl. Acad. Sci. USA *114*, 7408–7413. https://doi.org/10.1073/pnas.1612980114.
19. Anisenko, A.N., Knyazhanskaya, E.S., Zalevsky, A.O., Agapkina, J.Y., Sizov, A.I., Zatsepin, T.S., and Gottikh, M.B. (2017). Characterization of HIV-1 integrase interaction with human Ku70 protein and initial implications for drug targeting. Sci. Rep. *7*, 5649. https://doi.org/10.1038/s41598-017-05659-5.
20. Raval, P., Kriatchko, A.N., Kumar, S., and Swanson, P.C. (2008). Evidence for Ku70/Ku80 association with full-length RAG1. Nucleic Acids Res. *36*, 2060–2072. https://doi.org/10.1093/nar/gkn049.
21. Bétermier, M., Borde, V., and de Villartay, J.-P. (2020). Coupling DNA Damage and Repair: an Essential Safeguard during Programmed DNA Double-Strand Breaks? Trends Cell Biol. *30*, 87–96. https://doi.org/10.1016/j.tcb.2019.11.005.
22. Britton, S., Coates, J., and Jackson, S.P. (2013). A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. J. Cell Biol. *202*, 579–595. https://doi.org/10.1083/jcb.201303073.
23. Zangarelli, C., Arnaiz, O., Bourge, M., Gorrichon, K., Jaszczyszyn, Y., Mathy, N., Escoriza, L., Bétermier, M., and Régnier, V. (2022). Developmental timing of programmed DNA elimination in Paramecium tetraurelia recapitulates germline transposon evolutionary dynamics. Genome Res. *32*, 2028–2042. https://doi.org/10.1101/gr.277027.122.
24. Gratias, A., and Bétermier, M. (2003). Processing of double-strand breaks is involved in the precise excision of paramecium internal eliminated sequences. Mol. Cell Biol. *23*, 7152–7162.
25. Kim, C.S., Preer, J.R., and Polisky, B. (1992). Bacteriophage lambda DNA fragments replicate in the Paramecium macronucleus: absence of active copy number control. Dev. Genet. *13*, 97–102. https://doi.org/10.1002/dvg.1020130202.
26. Gilley, D., Preer, J.R., Aufderheide, K.J., and Polisky, B. (1988). Autonomous replication and addition of telomerelike sequences to DNA microinjected into Paramecium tetraurelia macronuclei. Mol. Cell Biol. *8*, 4765–4772. https://doi.org/10.1128/mcb.8.11.4765-4772.1988.
27. Denby Wilkes, C., Arnaiz, O., and Sperling, L. (2016). ParTIES: a toolbox for *Paramecium* interspersed DNA elimination studies. Bioinformatics *32*, 599–601. https://doi.org/10.1093/bioinformatics/btv691.
28. Bétermier, M., Klobutcher, L.A., and Orias, E. (2023). Programmed chromosome fragmentation in ciliated protozoa: multiple means to chromosome ends. Microbiol. Mol. Biol. Rev. *87*, e0018422. https://doi.org/10.1128/mmbr.00184-22.
29. Amar, L. (1994). Chromosome End Formation and Internal Sequence Elimination as Alternative Genomic Rearrangements in the Ciliate Paramecium. J. Mol. Biol. *236*, 421–426. https://doi.org/10.1006/jmbi.1994.1154.
30. Putnam, C.D., Pennaneach, V., and Kolodner, R.D. (2004). Chromosome healing through terminal deletions generated by de novo telomere additions in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA *101*, 13262–13267. https://doi.org/10.1073/pnas.0405443101.
31. Hickman, A.B., and Dyda, F. (2016). DNA Transposition at Work. Chem. Rev. *116*, 12758–12784. https://doi.org/10.1021/acs.chemrev.6b00003.
32. Bischerour, J., Lu, C., Roth, D.B., and Chalmers, R. (2009). Base flipping in V(D)J recombination: insights into the mechanism of hairpin formation, the 12/23 rule, and the coordination of double-strand breaks. Mol. Cell Biol. *29*, 5889–5899. https://doi.org/10.1128/MCB.00187-09.
33. Gratias, A., Lepère, G., Garnier, O., Rosa, S., Duharcourt, S., Malinsky, S., Meyer, E., and Bétermier, M. (2008). Developmentally programmed DNA splicing in Paramecium reveals short-distance crosstalk between DNA cleavage sites. Nucleic Acids Res. *36*, 3244–3251. https://doi.org/10.1093/nar/gkn154.
34. Stinson, B.M., and Loparo, J.J. (2021). Repair of DNA Double-Strand Breaks by the Nonhomologous End Joining Pathway. Annu. Rev. Biochem. *90*, 137–164. https://doi.org/10.1146/annurev-biochem-080320-110356.
35. Seif-El-Dahan, M., Kefala-Stavridi, A., Frit, P., Hardwick, S.W., Chirgadze, D.Y., Maia De Oliviera, T., Andreani, J., Britton, S., Barboule, N., Bossaert, M., et al. (2023). PAXX binding to the NHEJ machinery explains functional redundancy with XLF. Sci. Adv. *9*, eadg2834. https://doi.org/10.1126/sciadv.adg2834.
36. Chen, S., Vogt, A., Lee, L., Naila, T., McKeown, R., Tomkinson, A.E., Lees-Miller, S.P., and He, Y. (2023). Cryo-EM visualization of DNA-PKcs structural intermediates in NHEJ. Sci. Adv. *9*, eadg2838. https://doi.org/10.1126/sciadv.adg2838.
37. Graham, T.G.W., Walter, J.C., and Loparo, J.J. (2016). Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. Mol. Cell *61*, 850–858. https://doi.org/10.1016/j.molcel.2016.02.010.
38. Chen, S., Lee, L., Naila, T., Fishbain, S., Wang, A., Tomkinson, A.E., Lees-Miller, S.P., and He, Y. (2021). Structural basis of long-range to short-range synaptic transition in NHEJ. Nature *593*, 294–298. https://doi.org/10.1038/s41586-021-03458-7.
39. Khan, H., and Ochi, T. (2023). Plant PAXX has an XLF -like function and stimulates DNA end joining by the Ku-DNA ligase IV/XRCC4 complex. Plant J. *116*, 58–68. https://doi.org/10.1111/tpj.16359.
40. Matsuda, A., Shieh, A.W.-Y., Chalker, D.L., and Forney, J.D. (2010). The conjugation-specific Die5 protein is required for development of the somatic nucleus in both Paramecium and Tetrahymena. Eukaryot. Cell *9*, 1087–1099. https://doi.org/10.1128/EC.00379-09.
41. Lescale, C., Abramowski, V., Bedora-Faure, M., Murigneux, V., Vera, G., Roth, D.B., Revy, P., de Villartay, J.-P., and Deriano, L. (2016). RAG2 and XLF/Cernunnos interplay reveals a novel role for the RAG complex in DNA repair. Nat. Commun. *7*, 10529. https://doi.org/10.1038/ncomms10529.
42. Dubois, E., Mathy, N., Régnier, V., Bischerour, J., Baudry, C., Trouslard, R., and Bétermier, M. (2017). Multimerization properties of PiggyMac, a domesticated piggyBac transposase involved in programmed genome rearrangements. Nucleic Acids Res. *45*, 3204–3216. https://doi.org/10.1093/nar/gkw1359.
43. Beisson, J., Bétermier, M., Bré, M.H., Cohen, J., Duharcourt, S., Duret, L., Kung, C., Malinsky, S., Meyer, E., Preer, J.R., and Sperling, L. (2010). Maintaining clonal Paramecium tetraurelia cell lines of controlled age through daily reisolation. Cold Spring Harb. Protoc. *2010*, pdb.prot5361. https://doi.org/10.1101/pdb.prot5361.
44. Beisson, J., Bétermier, M., Bré, M.H., Cohen, J., Duharcourt, S., Duret, L., Kung, C., Malinsky, S., Meyer, E., Preer, J.R., and Sperling, L. (2010). Silencing specific Paramecium tetraurelia genes by feeding double-stranded

RNA. Cold Spring Harb. Protoc. *2010*, pdb.prot5363. https://doi.org/10.1101/pdb.prot5363.

45. Beisson, J., Bétermier, M., Bré, M.H., Cohen, J., Duharcourt, S., Duret, L., Kung, C., Malinsky, S., Meyer, E., Preer, J.R., and Sperling, L. (2010). DNA microinjection into the macronucleus of paramecium. Cold Spring Harb. Protoc. *2010*, pdb.prot5364. https://doi.org/10.1101/pdb.prot5364.
46. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.
47. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.
48. Arnaiz, O., Meyer, E., and Sperling, L. (2020). ParameciumDB 2019: integrating genomic data across the genus for functional and evolutionary biology. Nucleic Acids Res. *48*, D599–D605. https://doi.org/10.1093/nar/gkz948.
49. Galvani, A., and Sperling, L. (2002). RNA interference by feeding in Paramecium. Trends Genet. *18*, 11–12. https://doi.org/10.1016/S0168-9525(01)02548-3.
50. Nemoz, C., Ropars, V., Frit, P., Gontier, A., Drevet, P., Yu, J., Guerois, R., Pitois, A., Comte, A., Delteil, C., et al. (2018). XLF and APLF bind Ku80 at two remote sites to ensure DNA repair by non-homologous end joining. Nat. Struct. Mol. Biol. *25*, 971–980. https://doi.org/10.1038/s41594-018-0133-6.
51. Arnaiz, O., Mathy, N., Baudry, C., Malinsky, S., Aury, J.-M., Denby Wilkes, C., Garnier, O., Labadie, K., Lauderdale, B.E., Le Mouël, A., et al. (2012). The Paramecium Germline Genome Provides a Niche for Intragenic Parasitic DNA: Evolutionary Dynamics of Internal Eliminated Sequences. PLoS Genet. *8*, e1002984. https://doi.org/10.1371/journal.pgen.1002984.
52. Guérin, F., Arnaiz, O., Boggetto, N., Denby Wilkes, C., Meyer, E., Sperling, L., and Duharcourt, S. (2017). Flow cytometry sorting of nuclei enables the first global characterization of Paramecium germline DNA and transposable elements. BMC Genom. *18*, 327. https://doi.org/10.1186/s12864-017-3713-7.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-Ku70a (588–602) (Rabbit polyclonal)	This paper	N/A
anti-Flag antibody α-FLAG M2	Sigma-Aldrich	AB_262044
α-Pgm 2659-GP (Polyclonal guinea Pig)	(Dubois et al.)[42]	N/A
Alexa Fluor 568 goat anti-mouse IgG (H + L) antibody	ThermoFisherScientific LifeTech	A-11004
Goat anti-guinea pig IgG (H + L) highly cross adsorbed secondary antibody AF488	ThermoFisherScientific LifeTech	A-11073
Bacterial and virus strains		
Klebsiella pneumoniae (Cosson strain)	(Beisson et al.)[43]	N/A
Escherichia coli HT115	(Beisson et al.)[44]	N/A
Chemicals, peptides, and recombinant proteins		
Luna Qpcr Mix	New England Biolabs	M3003S
DAPI	Sigma-Aldrich	D9542
Propidium iodide	Sigma-Aldrich	P4170
Deposited data		
See Table S1-C for deposited sequencing data	This paper	https://www.ebi.ac.uk/ena/browser/view/PRJEB66127
Statistical data and R scripts used to generate the bioinformatic images and proteomic data	This paper	https://zenodo.org/doi/10.5281/zenodo.10047597
Experimental models: Organisms/strains		
P. tetraurelia wild-type 51 new	(Besson et al.)[43]	N/A
P. tetraurelia wild-type 51 *ND7-1*	(Besson et al.)[45]	N/A
Oligonucleotides		
See Table S1-B for oligonucleotides	This paper	N/A
Recombinant DNA		
See Files S1 for plasmid sequence	This paper	N/A
Software and algorithms		
bcl2fastq2 Conversion Software (v2.18.12)	Illumina	https://emea.support.illumina.com/
Cutadapt (v1.15/v1.12)	Cutadapt	https://cutadapt.readthedocs.io/
CASAVA (v1.8.2)	Illumina	https://jp.support.illumina.com/
Bowtie2 (v2.2.9)	(Langmead and Salzberg)[46]	http://bowtie-bio.sourceforge.net/
samtools (v1.9)	(Li et al.)[47]	http://www.htslib.org/
ParTIES (v1.06)	(Denby Wilkes et al.)[27]	https://github.com/oarnaiz/ParTIES https://zenodo.org/doi/10.5281/zenodo.6347088
R statistical computing platform (v4.0.4)	R	https://www.r-project.org/
ggplot2 v3.3.5; GenomicRanges v1.42; rtracklayer v1.50	R packages	https://bioconductor.org/
ParameciumDB website	(Arnaiz et al.)[48]	https://paramecium.i2bc.paris-saclay.fr/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Julien Bischerour (Julien.bischerour@i2bc.paris-saclay.fr).

Materials availability

Plasmids and antibodies used in this study are available from the lead contact upon reasonable request.

Data and code availability

- Sequencing data have been deposited at ENA and are publicly available as of the date of publication. Accession numbers are listed in the key resources table and Table S3. Microscopy data reported in this paper will be shared by the lead contact upon request. Sequencing and mapping data, as well as technical information, are available on the ParameciumDB website (https://paramecium.i2bc.paris-saclay.fr).
- Statistical data and R scripts used to generate the bioinformatic images have been deposited at Zenodo and are publicly available as of the date of publication. The DOI is listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Paramecium strains and culture conditions

P. tetraurelia wild-type 51 new or its mutant derivative 51 *nd7-1* were grown in a standard medium made of a wheat grass infusion inoculated with *Klebsiella pneumoniae* and supplemented with 0.8 μg/mL β-sitosterol and 100 μg/mL ampicillin.

METHOD DETAILS

Gene knockdowns during autogamy

Autogamy was carried out as described[42] and the progression of old MAC fragmentation and new MAC development was monitored using a Zeiss Lumar.V12 fluorescence stereo-microscope, following quick fixation and staining of cells in 0.2% paraformaldehyde 20 μg/mL DAPI. The T0 time-point for each experiment is defined as the stage where 50% of cells in the population harbor a fragmented old MAC.

RNAi was achieved using the feeding procedure, as described.[11,49] *Paramecium* cells grown for 10 to 15 vegetative fissions in plasmid-free *Escherichia coli* HT115 bacteria were transferred to medium containing non-induced HT115 harboring each RNAi plasmid and grown for ∼4 additional divisions. Cells were then diluted into plasmid-containing HT115 induced for dsRNA production and allowed to grow for ∼8 additional vegetative divisions before the start of autogamy. Final volumes were 50–100 mL for middle-scale experiments (western blotting, immunostaining and DNA extraction) and 0.5 to 1 L for large-scale experiments (whole-genome sequencing). The presence of a functional new MAC in the progeny was tested after four days of starvation as described.[12] Control experiments were performed using the L4440 vector. RNAi plasmids were L4440 derivatives carrying the following inserts: KU70a-1 (bp 514–813 from *KU70a*)[15] and pLIG4b-L(bp 521–1774 from LIG4b) for LIG4.[12]

Transgene construction, micro-injection and protein expression analysis

The Ku70-6E mutations were introduced into a *Paramecium* Ku70a-expressing vector carrying the *KU70a* gene fused at its 5′ end to sequences encoding the 3xFLAG peptide. Silent mutations were introduced into the *KU70a-6E* nucleic acid sequence to make it insensitive to RNAi. As a control, a wildtype version of *KU70a* was cloned and modified similarly to become RNAi-resistant. Plasmid sequences of plasmids are presented in File S1. All transgene-bearing pUC18 derivatives were linearized with appropriate restriction enzymes (BsaI) and co-injected with an *ND7*-complementing plasmid into the MAC of vegetative 51 *nd7-1*. Sequences of the *FLAG-KU70a* transgenes encoding N-terminal fusions of the 3X FLAG tag (YKDHDGDYKDHDIDYKDDDDKT) to Ku70a are displayed in File S1. Transgene injection level (copy per haploid genome or cphg) was determined by qPCR on genomic DNA extracted from vegetative transformants, using a LightCycler 480 and the Luna Qpcr Mix (New England Biolabs). Oligonucleotide primers for the *KU70a* transgene and the genomic reference locus are listed in Table S2.

A peptide corresponding to amino acid sequence 588 to 602 of *P. tetraurelia* Ku70a was used for rabbit immunization to yield anti-Ku70a (588–602) antibodies (Eurogentec). Polyclonal antibodies were purified by antigen affinity purification. The FLAG tag was revealed using monoclonal anti-Flag antibody α-FLAG M2 (Sigma-Aldrich). Protein extracts used for western-blot were prepared as previously.[16]

Cell staining

Cell staining and Immunostaining of fixed cells using polyclonal anti-Pgm guinea pig antibody α-Pgm 2659-GP or monoclonal anti-Flag antibody α-FLAG M2 (Sigma-Aldrich) were performed as described previously.[16] For propidium iodide staining, autogamous cells were permeabilized with ice-cold PHEM (60mM PIPES, 25 mM HEPES, 10mM EGTA, 2 mM MgCl2 pH 6.9) with 1% Triton X-100 during 4min, and fixed for 10min in PHEM +2% formaldehyde. Cells were further washed twice in TBST (10mM Tris pH 7.4, 0.15 M NaCl, 0.1% Tween 20) + 3% BSA, incubated for 5min in PI/RNase staining buffer (BD Pharmingen), and mounted in Citifluor AF2 (Biovalley). Observations were made with an Olympus BX63 epifluorescence microscope with a 63x oil objective or an Olympus BX63 epifluorescence microscope with a 60x oil objective, focusing on the maximal area section of new developing MACs. Quantification of new MAC sizes, fluorescence intensity, boxplot representation and statistical analysis were performed as described.[16]

Purification and DNA binding assays using recombinant Ku70/80 proteins expressed in insect cells

Ku70a/80c and Ku70a-6E/80c were expressed and purified as previously described, using a 6His tag at the N terminus of Ku80c.[15,50] Electrophoretic mobility shift assays were performed in the presence of 50nM of double stranded DNA substrate made of LE35Top

and LE35Bot-Cy3 oligos (Table S2), in a buffer containing 50 mM Tris pH 7.4, 100 mM NaCl, 5% glycerol and 0.1 mg/mL BSA. Complexes were assembled on ice for 15min before loading on 5% Acrylamide gel (29/1) and 0.5X TBE at room temperature. Gels were analyzed using a Typhoon scanner.

Detection of DNA extremities by ligation mediated PCR (LMPCR)

DNA double-strand breaks at the left genomic boundary of the IES 51A2591 and at the extremity of the excised IES 51G4404 were detected using the LMPCR technique.[12,24] PCR was performed with total genomic DNA that was purified from cells 5 h after the beginning of autogamy.

Purification of new developing MACs, sequencing of genomic DNA and mapping on the *paramecium* genomes

New developing Macs (anlagen) were purified by FANS using anti-Pgml1 antibodies and genomic DNA was extracted as previously described.[23] The total genomic DNA of late autogamous cells (4 days of starvation) was extracted from large-scale cultures (1L) using the NucleoSpin Tissue extraction kit (Macherey Nagel) as previously described.[16] All genomic DNAs were sequenced at a 76 to 160X coverage by a paired-end strategy using Illumina HiSeq (paired-end read length: 2 × 100 nt) or NextSeq (paired-end read length: 2x∼75 nt) sequencers (Table S1-C). DNA-seq data were filtered on expected contaminants (ribosomal DNA, mitochondrial and bacterial genomes) then mapped on the *P.tetraurelia* MAC (ptetraurelia_mac_51.fa) or MAC+IES (ptetraurelia_mac_51_with_ies.fa) reference genome using Bowtie2 (v2.2.9 –local -X 500).[51,52]

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

The number of replicates and the statistical methods used are indicated in the figure legend. A summary of the experiments carried out is presented in Table S1.

Bioinformatic analyses

Coverage around or between IES excision sites: Only IES excision sites on scaffolds greater than 30 kb (98% of IESs) were considered. The mean sequencing depth was calculated for inter-IES fragments (>300 nt) using SAMTools (depth v1.9 -q 30 -Q 30) then normalized by the number of mapped reads to compare samples. The coverage around IES excision sites (+/− 500 nt of the TA) was calculated using SAMTools (view v1.9). A home-made perl script wrapper (coverage_TA.pl) allowed to consider both mappings on the MAC and the MAC+IES genomes, to take into account all reads around IES excision sites. The collapsed coverage on all IESs was then normalized by the number of mapped reads. Telomere detection: the detection of *de novo* telomere addition sites is described in.[23] The determination of the distances to the closest IES and the positions of telomere addition sites around IES excision sites was performed with R GenomicRanges functions.

MEND analysis

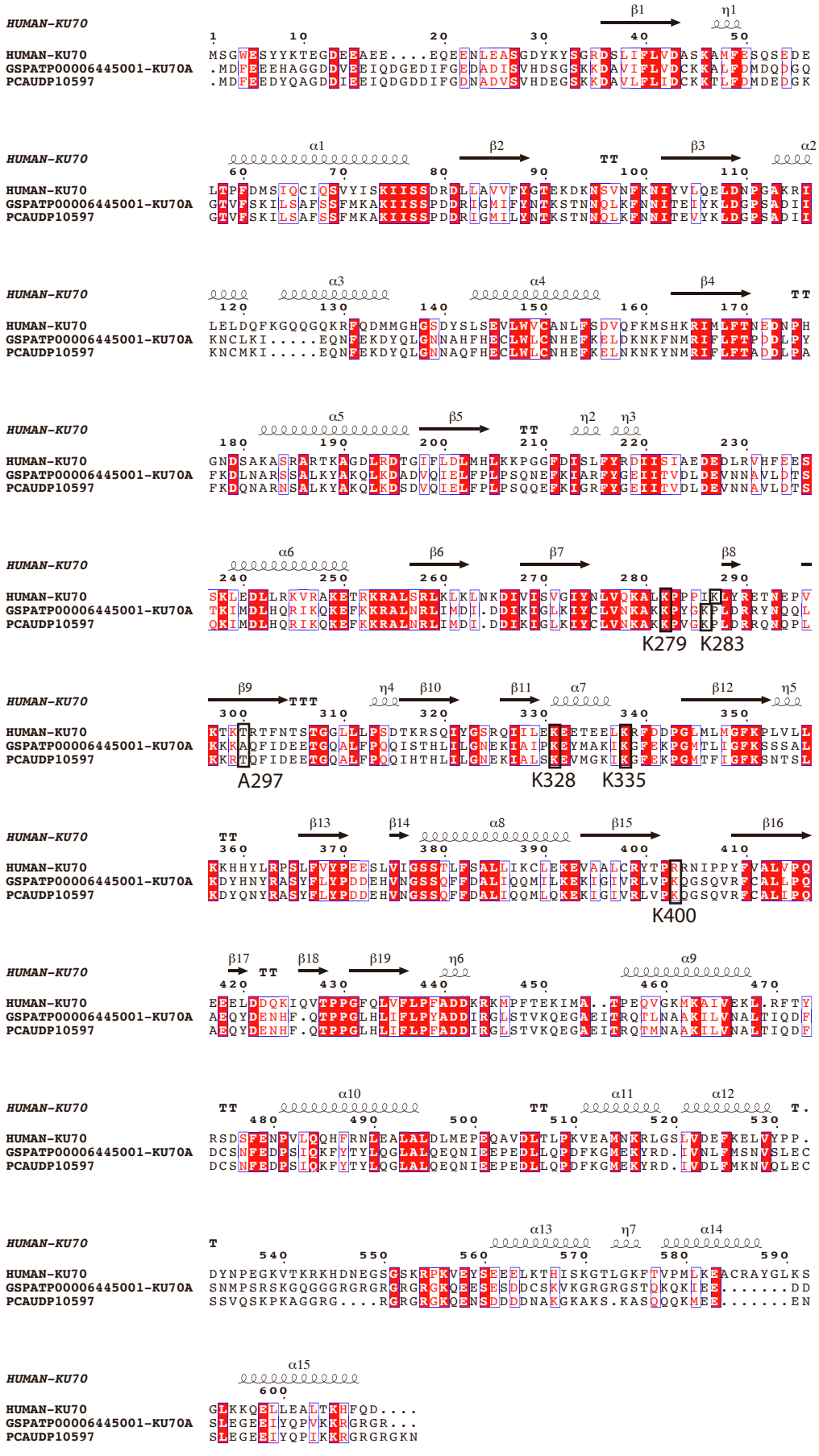
The IES annotation v1 (internal_eliminated_sequence_PGM_ParTIES.pt_51.gff3) was used for this study.[51] These datasets are available from the ParameciumDB download section.[48] A new version of ParTIES (v1.06 https://github.com/oarnaiz/ParTIES) was developed for this study. It is now possible to analyze translocation events (MILORD module, parameters: -consider_translocation) and to analyze and classify the reads around IES junctions (MEND module default parameters). The number of events was normalized by the number of mapped reads. R (v4.0.4) packages were used to manipulate annotations and to generate images (ggplot2 v3.3.5; GenomicRanges v1.42; rtracklayer v1.50).

Cell Reports, Volume *43*

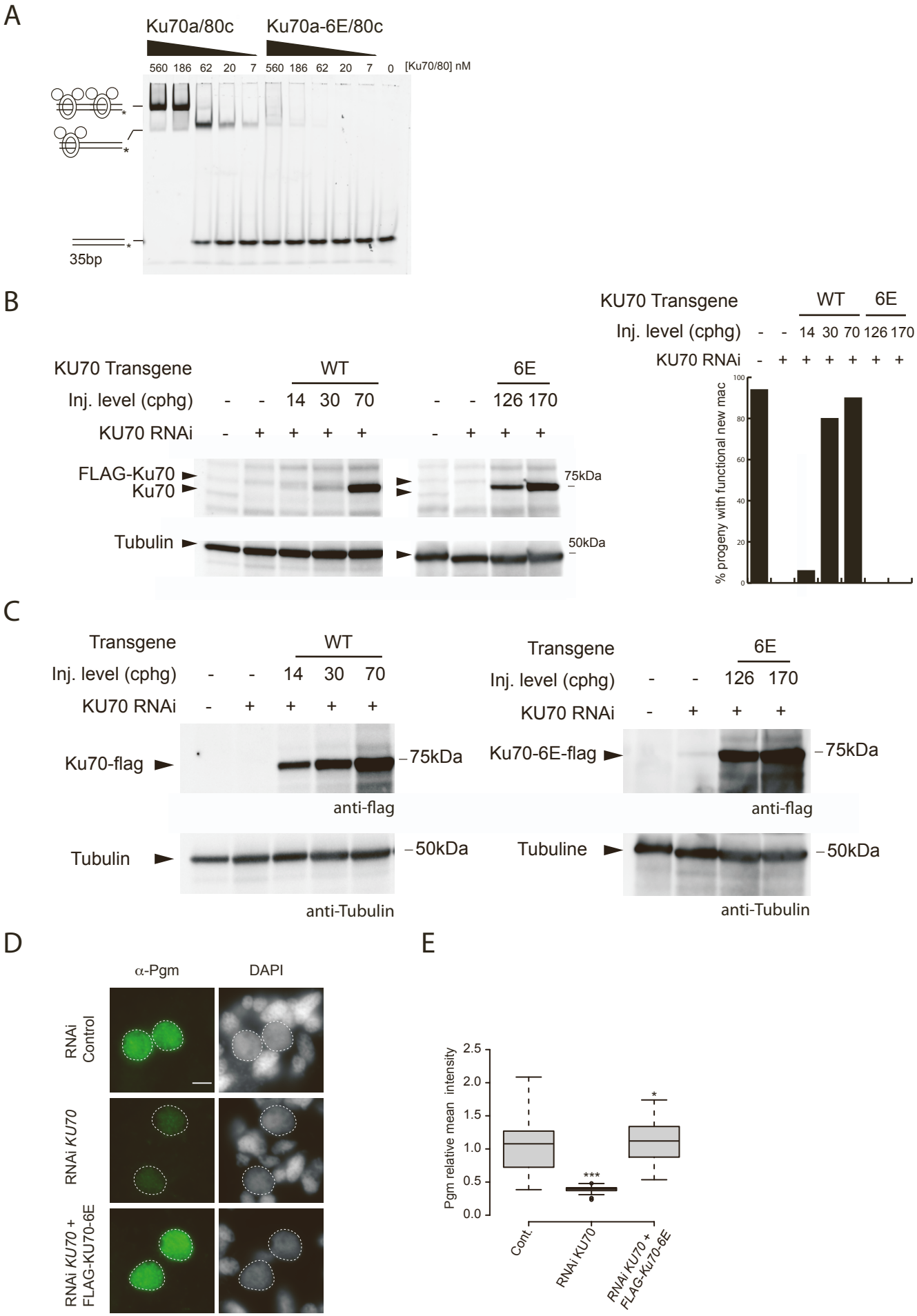
Supplemental information

Uncoupling programmed DNA cleavage and repair scrambles the *Paramecium* somatic genome

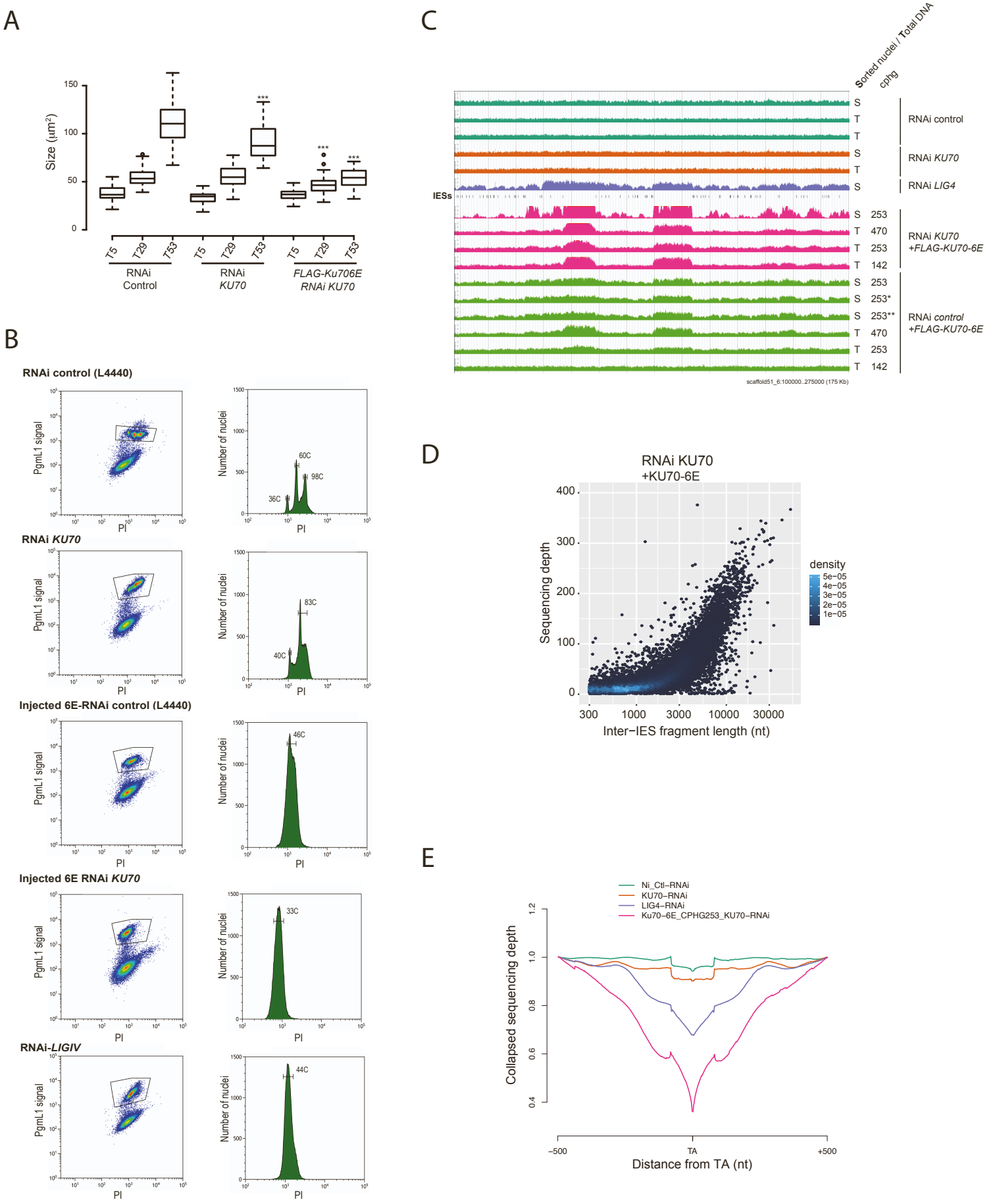
Julien Bischerour, Olivier Arnaiz, Coralie Zangarelli, Vinciane Régnier, Florence Iehl, Virginie Ropars, Jean-Baptiste Charbonnier, and Mireille Bétermier



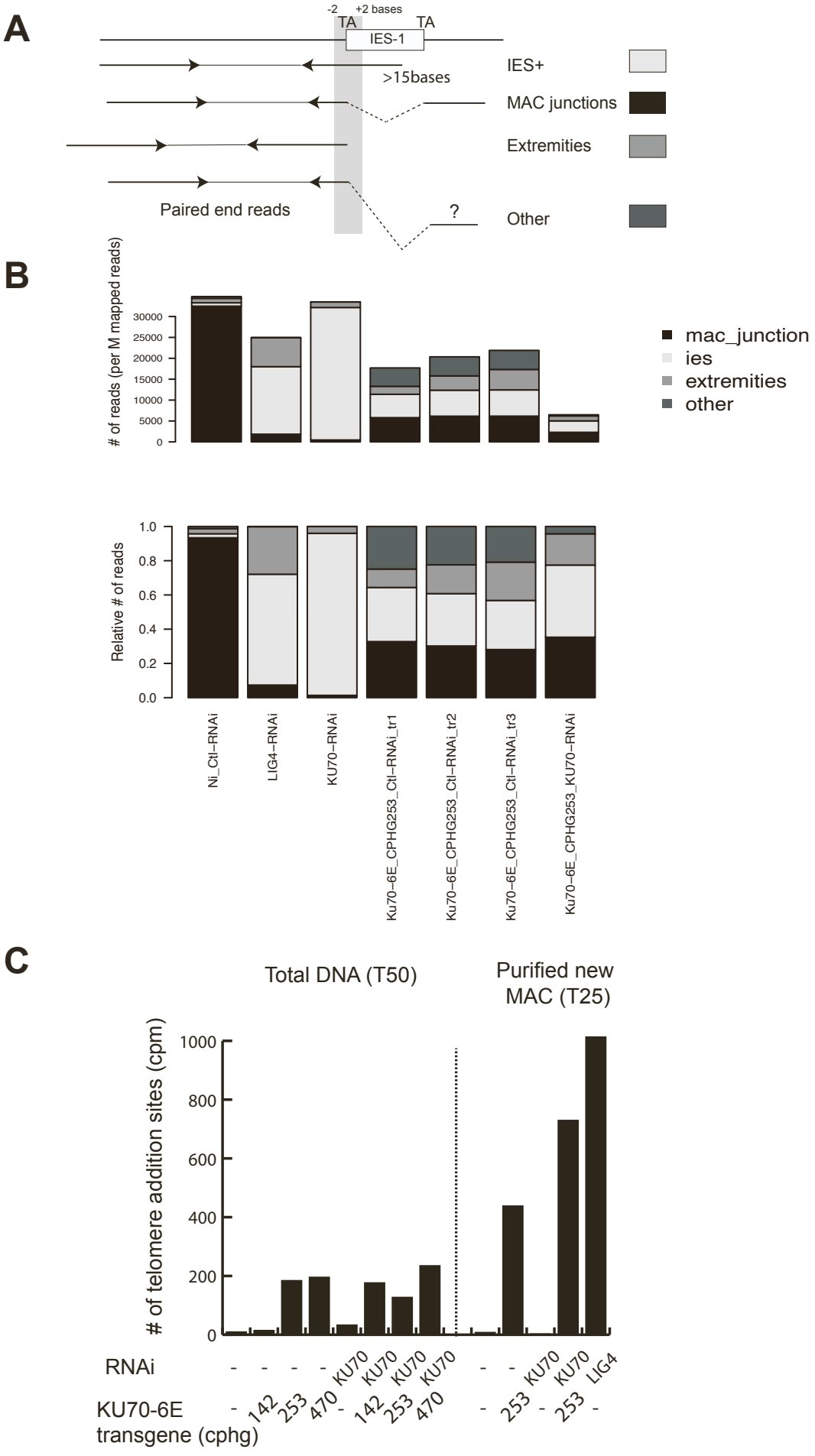
Bischerour et al, Figure S1



Bischerour et al, Figure S2

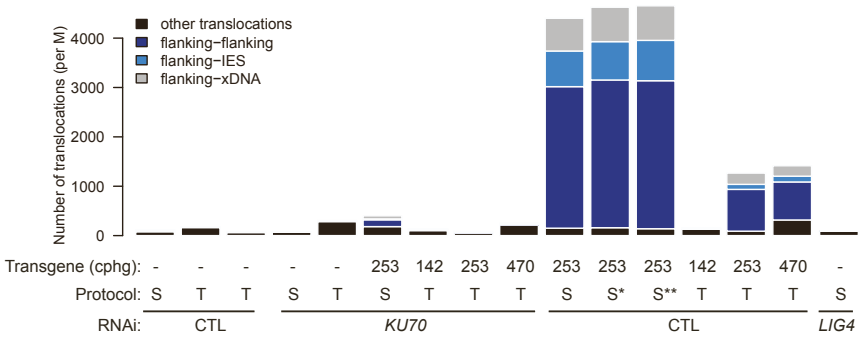


Bischerour et al, Figure S3

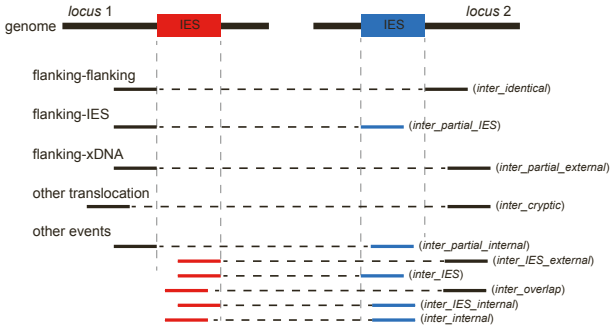


Bischerour et al, Figure S4

A

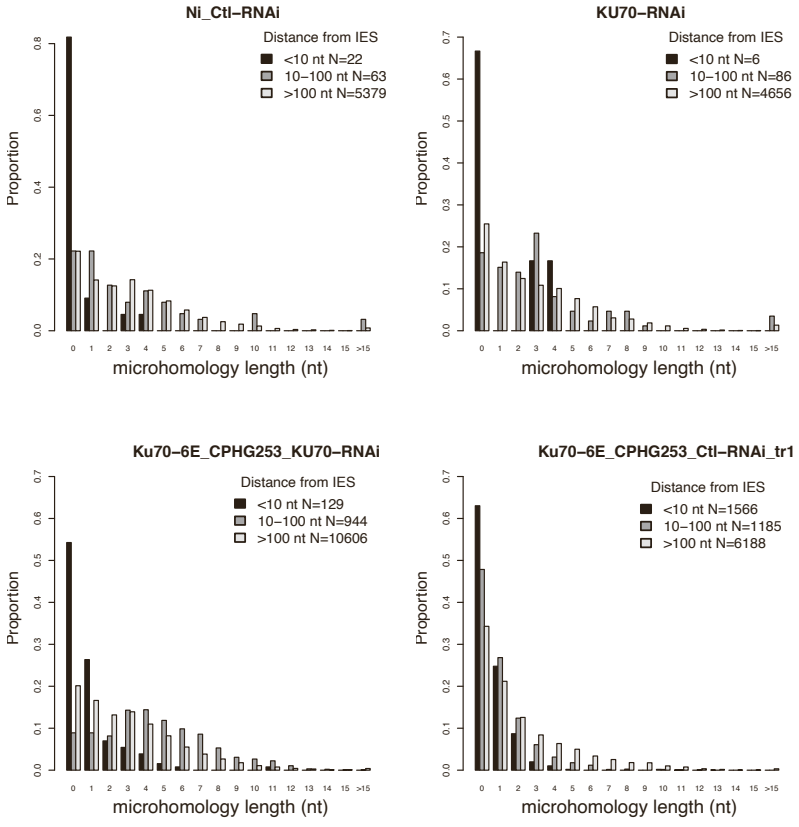


B

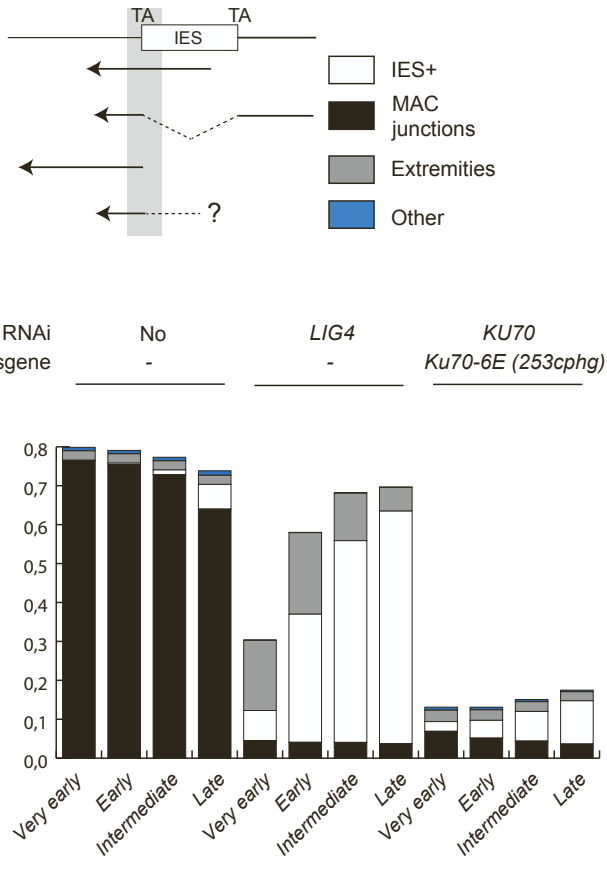


Transgene	RNAi	Protocol	inter_identical	inter_partial_IES	inter_partial_external	inter_cryptic	inter_partial_internal	inter_IES_external	inter_IES	inter_overlap	inter_IES_internal	inter_internal
Ni	Ctl	S	21.86	4.89	9.49	78.12	1.83	2.59	23.98	1.44	1.86	0.77
Ni	Ctl	T	6.01	0.15	6.92	162.79	0.58	0.11	0.03	1.33	0.03	0.12
Ni	Ctl	T	5.40	0.15	6.05	61.49	0.67	0.11	1.75	1.39	0.09	0.11
Ni	KU70	S	2.17	0.55	1.65	72.40	0.23	1.63	0.08	6.79	0.09	0.41
Ni	KU70	T	6.85	3.93	8.49	283.63	0.83	1.56	9.48	4.81	0.48	0.74
Ni	LIG4	S	0.41	0.22	6.28	90.94	0.17	0.95	0.15	4.51	0.13	0.33
Ku70-6E (253)	KU70	S	137.83	18.62	61.41	181.95	9.94	5.76	2.91	7.46	1.04	0.76
Ku70-6E (142)	KU70	T	20.83	2.23	13.32	101.26	1.27	0.85	0.13	2.00	0.15	0.29
Ku70-6E (253)	KU70	T	22.22	1.63	12.84	48.33	1.71	0.69	0.69	1.52	0.18	0.21
Ku70-6E (470)	KU70	T	23.74	1.82	15.08	210.97	1.41	0.86	0.37	2.25	0.14	0.18
Ku70-6E (253)	Ctl	S	2866.59	721.90	662.37	150.42	200.21	94.03	61.27	38.11	26.45	6.14
Ku70-6E (253)*	Ctl	S	2992.28	776.38	696.93	159.80	217.70	104.42	65.35	43.51	29.73	6.56
Ku70-6E (253)**	Ctl	S	2998.03	817.28	699.44	139.07	219.03	107.50	74.83	42.87	30.46	6.65
Ku70-6E (142)	Ctl	T	11.53	2.83	7.68	130.29	0.90	0.31	0.09	1.49	0.07	0.12
Ku70-6E (253)	Ctl	T	846.23	101.80	227.65	89.67	58.69	22.10	9.91	15.95	4.39	2.00
Ku70-6E (470)	Ctl	T	769.86	117.07	210.29	316.57	50.60	26.55	4.55	15.36	4.19	1.37

C



D



Bischerour et al, Figure S5

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Sequence alignment of *Paramecium* Ku70a and human Ku70 and modeling using ESPript3.0. (Related to Figure 1)

Alignment was performed using ClustalW, with the following sequences for Human *KU70 (CAG47015.1 G22P1)*, *Paramecium tetraurelia* Ku70a (PTET.51.1.P0150242), and *Paramecium caudatum* (PCAU.43c3d.1.P00420064) and used for ESPript3.0 analysis [1].

The 6E positions are represented by a black square, identical residues in red squares and homologs are written in red. Alfa helices (α), Beta strands (β) and Turns (T) are depicted.

Figure S2. Characterization of the Ku70-6E mutant activity. (Related to Figure 1)

(A) Analysis of the DNA binding activity of Ku80c/70a and Ku80c/Ku70a-6E mutant by electrophoretic mobility shift assay (EMSA).

(B) (Left Panel) Western blot analysis of FLAG-Ku70a and FLAG-Ku70a-6E expression levels in early autogamous cells (T5, *i.e.* 5 hrs after the T0 time-point) subjected to *KU70a* RNAi, using α-Ku70a antibodies. FLAG-Ku70a-expressing transgenes are RNAi-resistant, thanks to the modification of the nucleotide sequence. The injection level of each transformant is indicated as cphg (copies par haploid genome). (Right Panel) Survival test for the sexual progeny of *FLAG-KU70a* and *FLAG-KU70a-6E*-complemented cells.

(C) Western blot analysis of FLAG-Ku70a and FLAG-Ku70a-6E expression levels in early autogamous transformed cells (T5) subjected or not to *KU70* RNAi. Total protein extracts were prepared at T5 after the beginning of autogamy. FLAG-Ku70a proteins were revealed on western blots using α-Flag (M2-Sigma) antibodies.

(D) Immunolabelling of Pgm in early developing MACs (T5) of control cells, *KU70*-silenced cells and cells transformed with FLAG-*KU70-6E* (280 cphg) subjected to *KU70* RNAi. Scale bar is 5 μm.

(E) Boxplot of Pgm mean fluorescent intensity in 36-70-μm^2 developing MACs under the different conditions. This window of new MAC sizes corresponds to the maximal Pgm signal in the control. 73 to 81 developing MACs were analyzed for each condition. * for p-value <0.05 and *** for p-value <0.001 in a Mann-Whitney-Wilcoxon statistical test, in comparison with the control.

Figure S3. Ku70-6E mutant impairs new developing MAC assembly. (Related to Figure 2)

(A) Boxplot of the sizes (μm^2) of developing MACs from the different conditions. 44 to 59 developing MACS were analyzed for each condition. The statistical test (Mann-Whitney-Wilcoxon) compares the *KU70* RNAi and FLAG-Ku70-6E/*KU70* RNAi conditions at T5, T29 and T53, with their respective control RNAi. ** for p-value <0.01 and *** for p-value <0.001.

(B) Flow cytometry sorting of αPgmL1-immunolabeled nuclei in RNAi control (L4440), *KU70a, LIG4* RNAi and in *FLAG-KU70-6E* injected cells (*KU70* RNAi and L4440 control conditions). Sorted nuclei correspond to the gated events on the PgmL1 labelling vs PI plots. The PI histogram shows the DNA content of the sorted population. Estimation of the C-level for the indicated peaks was performed as described using tomato nuclei as an internal standard [2].

(C) Screenshot of read coverage over a 175kb window along 51_6 scaffold (position 100000 to 275000). Purified new developing MACs (S= sorted) collected at T25 and total genomic DNA (T=total) collected at T53 from the different conditions were sequenced using the

Illumina technology. Sequencing libraries were prepared using enzymatic digestion ("westburg library prep kit"), Nebnext DNA fragmentase (*) or mechanical fragmentation (covaris) (**) (Table S3). IES positions are indicated by small vertical bars. The injection level of the *Flag-KU70-6E* transgene is indicated on the right (cphg = copies per haploid genome).

(D) Sequencing depth analysis of Inter-IES fragments in a *FLAG-KU70-6E* transformant (253cphg) in *KU70* RNAi conditions. The average sequencing coverage was calculated for individual inter IES fragments along the genome and reported according to their size. Fragments shorter than 300bp were not considered, since they were possibly counter-selected during the preparation of the sequencing libraries.

(E) Analysis of collapsed sequencing depth at the proximity of IES excision sites (-500 to +500nt). The figure is equivalent to figure 2E, except that reads "extremities" (MEND analysis) were omitted.

Figure S4. Analysis of uncoupling in supplementary samples. (Related to Figure 3 and Figure 4)

(A) Scheme of events considered in MEND analysis. Sequencing reads overlapping the TA dinucleotide which is present at the extremities of each IES were classified in the four depicted categories. Categories of reads, IES+, Mac junction, Extremities and "other". Reads must map over a minimal 15 nucleotides region to be considered.

(B) Top Panel, MEND analysis of purified new developing MACs. Samples were collected at T25 and sequenced using the Illumina technology. Sequencing libraries were prepared using enzymatic digestion ("westburg library prep kit"), Nebnext DNA fragmentase (*) or mechanical fragmentation (covaris) (**) (Table S3). Bottom Panel, relative quantification of MEND analysis of purified new developing MACs (S= sorted) collected at T25.

(C) Quantification of *de novo* telomere addition. As previously described[2], a telomere addition site was pinpointed wherein a read alignment on the MAC sequence stops and the read sequence proceeds with telomeric repeats (G_4T_2 or G_3T_3). Then the number of events was normalized by the total number of mapped reads (cpm).

Figure S5. Detailed analysis of translocations in all genomic DNA samples. (Related to Figure 4)

(A) Quantification of translocation events. The quantification of the number of translocations was performed using the same procedure as in Figure 4B. The conditions of RNAi, the level of injection (cphg), and the protocol of preparation of the genomic DNA samples (S= sorted nuclei at T25, T= total genomic DNA at T53) are indicated below the histogram.

(B) (Left Panel) Scheme of all subcategories of translocations. The four main types of translocations, according to the nature of DNA extremities and to the processing of ends, are presented at the top. Other types of translocations indicated as "other events", were not considered since their frequency was very low or not relevant for this study, such as the inter-IES junctions that correspond to the assembly of excised IESs (concatemers) (see Table S3). The * and ** correspond to technical replicates of the KU70-6E injected sample (253 cphg) with different methods of genomic DNA library preparation (see Table S3). (Right Panel) Numerical values of translocation frequencies (number of events normalized per million mapped reads) for each sample.

(C) Analysis of the length of microhomologies. Translocations classified as "other" were analyzed. They are grouped in three classes according to the distance to the nearest IES (< 10 nucleotides (black), < 100 nucleotides (dark grey) or > 100 nucleotides (light grey)). The frequency of translocations is represented as a function of the size of the microhomology that is detected at the junction. The number of translocations for each category is presented at the

right top part of the figure. In control and *KU70* RNAi, most of the translocations implicated fragments of the genome that are distant from IES. They are characterized by frequent longer microhomologies at the junction. In *KU70-6E* injected cells, translocations are found in the three categories of distance. Long microhomologies are more often detected when the translocations implicate junctions that are distant from an IES. This suggests that translocation induced by errors during IES excision are likely processed by cNHEJ. In contrast, when limited end trimming happens, another DNA repair pathway requiring longer microhomology is ongoing. All the analyses were performed with sorted new developing MACs.

(D) Analysis of reads coverage as a function of IES excision time. IESs were clustered in four group of excision timing, (very early, early, intermediate and late) as described previously[2]. The analysis was performed using the MEND module of PARties and normalized by the number of IESs in each group. The different configurations of IES-flanking sequences are presented at the top. They are classified in four categories; sequencing reads that correspond to the genomic loci with an IES (IES+ in white), or to an already excised IES (MAC junction in black), reads that start at the TA (+ or-2 bp) (Extremities in grey), and reads that map on the flanking DNA but cannot be aligned on adjacent MIC or MAC beyond the TA dinucleotide (+ or-2 bp) (Other). Analysis was performed on purified (sorted nuclei) new developing MAC genomic DNA. The data show the accumulation of IES+ reads on the latest excised IES in *LIG4* RNAi and in *KU70-6E* expressing cells, suggesting an indirect effect of unrepaired DSBs generated during early IES excision, on the excision of the late IESs.

SUPPLEMENTAL FILES

File S1. Sequence of plasmids

Sequence of pUC-3FLAG-KU70a and pUC-3FLAG-KU70a-6E

> pUC-3FLAG-KU70a (5229 bp)

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCA
GAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGC
GCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAA
GGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTG
CCAAGCTTGCATGCTAAGGACAAAAATATCTAATACTTTTTTTTTGGACTTTTTCAAATTATCATTCCAATAATTATAATGAT
ATTAATTAATATAATAATTCTGTAGAGTTTCTGTTCATTAAATTAATAAAATAAATGAATTCTAGAGACTACAAAGACCAT
GACGGTGATTATAAAGATCATGACATCGACTACAAGGATGACGATGATAAGGGATCCATGGATTTTGAAGAGGAGCACG
CAGGAGGAGACGACGTTGAAGAAATCTAAGACGGAGAAGATATATTTGGTGAAGACGCTGATATATCAGTACACGATTC
AGGGAGCAAAAAAGATGCAGTCATATTTTTGGTAGATTGCAAAAAAGCATTGTTTGACATGGATCAAGATGGACAAGGCA
CAGTCTTTTCTAAAATATTATCTGCCTTTTCTAGCTTTATGAAGGCAAAAATCATATCAAGTCCAGATGATCGCATTGGTAT
GATATTCTACAACACAGTAATTAATATCTATAAAATTAGAAATCCACCAATAATCAATTAAAATTCAATAACATCACAGA
AATTTATAAATTAGATGGACCATCTGCAGATATTATCAAAAATTGTTTAAAAATAGAGCAAAATTTTGAGAAAGATTATC
AATTAGGCAACAATGCTCATTTTCATGAATGTTTATGGCTCTGCAATCATGAATTCAAAGAATTGTAAACATATTATCAA
TATTTATTAGAGATAAGAACAAATTCAATATGCGTATATTTTTATTCACTCCAGATGATTTACCATATTTTAAAGATTTG
AATGCCAGATCATCAGCTTTGAAATACGCCAAGTAATTGAAGGATGCAGATGTTTAGATTGAATTGTTTCCATTGCCCAG
TTAAAATGAATTCAAAATTGCTAGATTTTACGGTGAGATTATAACTGTTGATTTGGATGAAGTTAACAATGCTGTACTTG
ATACTTCAACTAAAATTATGGATTTGCATTAAAGAATTAAATAAAAGGAATTTAAGAAAAGAGCTTTGAATAGACTTATT
ATGGATATCGATGATATCAAGATAGGTTTGAAGATTTACTGTCTAGTTAACAAAGCTAAGAAACCTTACGGAAAGCCATT
GGACAGAAGATATAATTAACAATTAAAGAAAAAGGCTTAATTCATAGACGAGGAAACTGGATAAGCTCTATTCCCATAAC
AAATTTCAACACATTTAATTTTAGGTAACGAGAAGATTGCTATACCTAAAGAATACATGGCAAAGATTAAGGGATTCGAG
AAACCAGGAATGACATTAATTGGATTCAAGTCATCATCTGCACTAAAAGACTACCACAACTACAGGGCTTCATACTTTCT
ATACCCAGACGACGAGCATGTTAACGGTTCATCTCAATTCTTTGACGCTTTAATATAACAAATGATTTTAAAGGAAAAGA
TTGGAATTGTACGACTAGTTCCAAAGTAAGGTTCATAAGTTAGATTCTGTGCTTTACTACCTTAAGCAGAACAATATGAT
GAAAATCATTTTTAGACACCTCCAGGATTACATCTGATATTTTTACCTTATGCTGATGATATTAGAGGACTGTCTACAGT
TAAGTAAGAAGGAGCCGAGATTACAAGATAGACCTTGAATGCAGCTAAGATTTTAGTAAATGCTCTGACAATCCAAGATT
TTGATTGCTCAAACTTTGAAGATCCTTCAATATAGAAGTAATAATAGATTATTAATATAGATTTTACACATATTTACAAG
GGTTGGCATTGTAAGAGTAAAACATCGAAGAACCTGAAGATTTACTTTAGCCAGACTTTAAAGGTATGGAAAAATATAGA
GATATTGTGAACTTATTTATGAGTAATGTCAGTCTGGAATGCTCTAATATGCCCTCAAGATCTAAAGGATAAGGAGGGGG
AAGAGGTAGAGGCAGAGGAAGAGGAAGAGGCAAACAAGAAGAAAGTGAGAGTGACGATTGTTCTAAAGTAAAGGGGA
GAGGCAGAGGGTCCACTTAGAAGCAAAAGATTGAAGAGGATGACAGTTTGGAAGGAGAGGAGATCTATTAACCAGTAAA
AAAGAGAGGAAGAGGACGTTGATTATTGTTTTTACAAACTCATATTAGGATTATATAAAAACTTAAATATCAATTGGTTTAT
TCCACATTTTTAAAATAAAATCATTTCATTTGTCGCTTTGGCAACCCCGTTGGAGGTTCTTTTATCAACTAAGACATATGTT
TTAGATTATTAATATTTATTGTATGATAGATATTTGGATTCATGAAAATTCTGAGATAAATTCTATATTAAAATAATTAG
CATTATAAAATCTAGAAAAATGGGTAATTGTAAGAGCTCAAGCTTGAGTTGGGATTTGACATAATCGGTGAAATGAAGCT
GTTCTTTTTCTTTCGTGTAATACTGTAAAATAAAAACAATTTATCTTAGGTAATCATCAGATCCAATCATTAAATGCTAT
GAATTCAAACTCCGATATAATATACGGTTTGAATGCAAATATTCAAGAAAAAGAACAACCTAGGCAGCATGGAATTCAGC
GTGTTCTTGGTTTAATTAGAGCATGCGAGCTCGAATTCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCC
GCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACAT
TAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCG
GGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG
AGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCA
TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCT
CCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTT
CTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC
AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGC
TACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC
GGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA
AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTAT
CAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG
TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACT
CCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT

CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC
TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCAT
TGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTA
CATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTG
TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGA
GTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCG
CGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTG
TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTG
AGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTT
TTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA
ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA
AAATAGGCGTATCACGAGGCCCTTTCGTC

> pUC-3FLAG-KU70a-6E

TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCA
GAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGC
GCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAA
GGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTG
CCAAGCTTGCATGCTAAGGACAAAAATATCTAATACTTTTTTTTTGGACTTTTTCAAATTATCATTCCAATAATTATAATGAT
ATTAATTAATATAATAATTCTGTAGAGTTTCTGTTCATTAAATTAATAAAATAAATGAATTCTAGAGACTACAAAGACCAT
GACGGTGATTATAAAGATCATGACATCGACTACAAGGATGACGATGATAAGGGATCCATGGATTTTGAAGAGGAGCACG
CAGGAGGAGACGACGTTGAAGAAATCTAAGACGGAGAAGATATATTTGGTGAAGACGCTGATATATCAGTACACGATTC
AGGGAGCAAAAAAGATGCAGTCATATTTTTGGTAGATTGCAAAAAAGCATTGTTTGACATGGATCAAGATGGACAAGGCA
CAGTCTTTTCTAAAATATTATCTGCCTTTTCTAGCTTTATGAAGGCAAAAATCATATCAAGTCCAGATGATCGCATTGGTAT
GATATTCTACAACACAGTAATTAATATCTATAAAATTAGAAATCCACCAATAATCAATTAAAATTCAATAACATCACAGA
AATTTATAAATTAGATGGACCATCTGCAGATATTATCAAAAATTGTTTAAAAATAGAGCAAAATTTTGAGAAAGATTATC
AATTAGGCAACAATGCTCATTTTCATGAATGTTTATGGCTCTGCAATCATGAATTCAAAGAATTGTAAACATATTATCAA
TATTTATTAGAGATAAGAACAAATTCAATATGCGTATATTTTTATTCACTCCAGATGATTTACCATATTTTAAAGATTTG
AATGCCAGATCATCAGCTTTGAAATACGCCAAGTAATTGAAGGATGCAGATGTTTAGATTGAATTGTTTCCATTGCCCAG
TTAAAATGAATTCAAAATTGCTAGATTTTACGGTGAGATTATAACTGTTGATTTGGATGAAGTTAACAATGCTGTACTTG
ATACTTCAACTAAAATTATGGATTTGCATTAAAGAATTAAATAAAAGGAATTTAAGAAAAGAGCTTTGAATAGACTTATT
ATGGATATCGATGATATCAAGATAGGTTTGAAGATTTACTGTCTAGTTAACAAAGCTAAGgAACCTTACGGAgAGCCATT
GGACAGAAGATATAATTAACAATTAAAGAAAAAGGaaTAATTCATAGACGAGGAAACTGGATAAGCTCTATTCCCATAAC
AAATTTCAACACATTTAATTTTAGGTAACGAGAAGATTGCTATACCTgAAGAATACATGGCAAAGATTgAGGGATTCGAG
AAACCAGGAATGACATTAATTGGATTCAAGTCATCATCTGCACTAAAAGACTACCACAACTACAGGGCTTCATACTTTCT
ATACCCAGACGACGAGCATGTTAACGGTTCATCTCAATTCTTTGACGCTTTAATATAACAAATGATTTTAAAGGAAAAGA
TTGGAATTGTACGACTAGTTCCAgAGTAAGGTTCATAAGTTAGATTCTGTGCTTTACTACCTTAAGCAGAACAATATGAT
GAAAATCATTTTTAGACACCTCCAGGATTACATCTGATATTTTTACCTTATGCTGATGATATTAGAGGACTGTCTACAGT
TAAGTAAGAAGGAGCCGAGATTACAAGATAGACCTTGAATGCAGCTAAGATTTTAGTAAATGCTCTGACAATCCAAGATT
TTGATTGCTCAAACTTTGAAGATCCTTCAATATAGAAGTAATAATAGATTATTAATATAGATTTTACACATATTTACAAG
GGTTGGCATTGTAAGAGTAAAACATCGAAGAACCTGAAGATTTACTTTAGCCAGACTTTAAAGGTATGGAAAAATATAGA
GATATTGTGAACTTATTTATGAGTAATGTCAGTCTGGAATGCTCTAATATGCCCTCAAGATCTAAAGGATAAGGAGGGGG
AAGAGGTAGAGGCAGAGGAAGAGGAAGAGGCAAACAAGAAGAAAGTGAGAGTGACGATTGTTCTAAAGTAAAGGGGA
GAGGCAGAGGGTCCACTTAGAAGCAAAAGATTGAAGAGGATGACAGTTTGGAAGGAGAGGAGATCTATTAACCAGTAAA
AAAGAGAGGAAGAGGACGTTGATTATTGTTTTTACAAACTCATATTAGGATTATATAAAAACTTAAATATCAATTGGTTTAT
TCCACATTTTTAAAATAAAATCATTTCATTTGTCGCTTTGGCAACCCCGTTGGAGGTTCTTTTATCAACTAAGACATATGTT
TTAGATTATTAATATTTATTGTATGATAGATATTTGGATTCATGAAAATTCTGAGATAAATTCTATATTAAAATAATTAG
CATTATAAAATCTAGAAAAATGGGTAATTGTAAGAGCTCAAGCTTGAGTTGGGATTTGACATAATCGGTGAAATGAAGCT
GTTCTTTTTCTTTCGTGTAATACTGTAAAATAAAAACAATTTATCTTAGGTAATCATCAGATCCAATCATTAAATGCTAT
GAATTCAAACTCCGATATAATATACGGTTTGAATGCAAATATTCAAGAAAAAGAACAACCTAGGCAGCATGGAATTCAGC
GTGTTCTTGGTTTAATTAGAGCATGCGAGCTCGAATTCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCC
GCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACAT
TAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCG
GGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCG
AGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCA
AAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCA

TCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCT
CCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTT
CTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGC
AGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGC
TACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCC
GGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGA
AGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTAT
CAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGG
TCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACT
CCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC
TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCAT
TGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTA
CATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTG
TTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGA
GTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCG
CGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTG
TTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTG
AGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTT
TTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA
ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAA
AAATAGGCGTATCACGAGGCCCTTTCGTC

SUPPLEMENTAL TABLES

Table S1. Summary of RNAi experiments

Experiment	Clone number	Injected Transgene	CPHG (PCRq)	Feeding conditions	caryonide	% of progeny with functional new mac	Experimental reference of the clone	Comments
01/01/21	Ni	-		L4440	6a	93	Ni-A	Figure S2, 1C
01/01/21	Ni	-		KU70	6a	0	Ni-A	Figure S2, 1C
01/01/21	6E-C1	3F-KU70-6E*	126	KU70	6a	0	Clone F	Figure S2, 1C
01/01/21	6E-C2	3F-KU70-6E*	170	KU70	6c	0	CloneM	Figure S2, 1C
May-21	Ni	-		L4440	5b1	93	Ni	Figure S2
May-21	Ni	-		L4440	5b1	0	Ni	Figure S2
May-21	Wt-c1	3F-KU70*	14	KU70	5b1	7	clone H	Figure S2
May-21	Wt-c2	3F-KU70*	30	KU70	5b1	80	clone C	Figure S2
May-21	Wt-c3	3F-KU70*	70	KU70	5b1	90	clone D	Figure S2
01/07/17	Ni	-		L4440	2c	96	Ni	Figure S3
01/07/17	Ni	-		KU70	2c	0	Ni	Figure S3
01/07/17	Inj6E	3F-KU70-6E*	280	KU70	2c	0	2c1	Figure S3
01/07/17	Inj6E	3F-KU70-6E*	42	L4440		100		Figure 4A
01/07/17	Inj6E	3F-KU70-6E*	100	L4440		83		Figure 4A
01/07/17	Inj6E	3F-KU70-6E*	280	L4440	2c	10	2c1	Figure 4A
01/07/20	Ni	-		L4440	7a3	100	Ni	Figure 2&3&4 DNA sequencing of anlagen
01/07/20	Ni	-		KU70	7a3	0	Ni	Figure 2&3&4 DNA sequencing of anlagen
01/07/20	InjKU70-6E	3F-KU70-6E*	253	L4440	7a3	0	Clone A	Figure 2&3&4 DNA sequencing of anlagen
01/07/20	InjKU70-6E	3F-KU70-6E*	253	L4440	7a3	0	Clone A	Figure 2&3&4 DNA sequencing of anlagen
01/07/20	InjKU70-6E	3F-KU70-6E*	253	L4440	7a3	0	Clone A	Figure 2&3&4 DNA sequencing of anlagen
01/07/20	InjKU70-6E	3F-KU70-6E*	253	KU70	7a3	0	Clone A	Figure 2&3&4 DNA sequencing of anlagen
01/03/17	InjKU70-6E	3F-KU70-6E*	4	L4440	-	90	H2	Figure 4A
01/03/17	InjKU70-6E	3F-KU70-6E*	9	L4440	-	93	D1	Figure 4A
01/03/17	InjKU70-6E	3F-KU70-6E*	88	L4440	-	100	C3	Figure 4A
01/03/17	InjKU70-6E	3F-KU70-6E*	102	L4440	-	56	G1	Figure 4A
01/03/17	InjKU70-6E	3F-KU70-6E*	107	L4440	-	33	B1	Figure 4A
01/03/17	InjKU70-6E	3F-KU70-6E*	272	L4440	-	0	F2	Figure 4A
01/09/17	InjKU70-6E	3F-KU70-6E*	30	L4440	-	0	3A	Figure 4A
01/09/17	InjKU70-6E	3F-KU70-6E*	120	L4440	-	0	3B	Figure 4A
01/12/17	Ni	-		L4440	-	90	Ni	DNA sequencing Total DNA
01/12/17	Ni	-		KU70	-	0	Ni	DNA sequencing Total DNA
01/12/17	InjKU70-6E	3F-KU70-6E*	470	L4440	-	0	R	DNA sequencing Total DNA
01/12/17	InjKU70-6E	3F-KU70-6E*	142	L4440	-	56	V	DNA sequencing Total DNA
01/12/17	InjKU70-6E	3F-KU70-6E*	470	KU70	-	0	R	DNA sequencing Total DNA
01/12/17	InjKU70-6E	3F-KU70-6E*	142	KU70	-	0	V	DNA sequencing Total DNA
01/09/19	InjKU70-6E	3F-KU70-6E*	105	L4440	-	92	I-23	Figure 4A
01/09/19	InjKU70-6E	3F-KU70-6E*	148	L4440	-	56	I-31	Figure 4A
01/09/19	InjKU70-6E	3F-KU70-6E*	183	L4440	-	17	I-30	Figure 4A
01/08/19	Ni	-	0	L4440	-	97		Figure 2A
01/08/19	Ni	-	0	KU70	-	0		Figure 2A
01/08/19	InjKU70-6E	3F-KU70-6E*	70	L4440	-	87		
01/08/19	InjKU70-6E	3F-KU70-6E*	70	KU70	-	0		Figure 2A
Printemps 2021	RNAi LIG4	-	-	LIG4	-	0		Figure 2&3 DNA sequencing of anlagen
	RNAi Control	-	-	L4440	-	96		

Table S2. List of oligonucleotides

-

Unique Name	Other name	Sequence (5' -> 3')
Reference locus for normalization (IES *51A* 1835)		
OMB063	51A1835-3'(4)	AGACAAGTAGGGAATCCACTTCTAGTAATC
OMB064	51A1835-5'	TAATGTATTGATAAGGCTTGCTCTACAGCC
KU70a locus		
OMB1174		GGAGGAGACGACGTTGAAGA
OMB1175		TACCAATGCGATCATCTGGA
DNA substrat EMSA		
OMB1125	LE35Top	CCCTAGAAAGATAGTCTGCGTAAAATTGACGCATG
OMB1126	LE35Bot-Cy3	CATGCGTCAATTTTACGCAGACTATCTTTCTAGGG

Table S3. Summary of sequencing data with ENA accession numbers

	Sorted nuclei / Total DNA	CPHG	Sample	Label	ENA Accession	Number of reads	Aligned reads on the MAC	%	Aligned reads on the MIC	%	Library preparation
RNAi control	S		PTET_Ni_L4440_T25_AlgFACS_DNA_BET43_S1	Ni_Ctl-RNAi_AlgT25	ERS16373096	87482632	78677828	90%	86674146	99%	westburg library prep kit
	T		PTET_BET2_ADN_WT_L4440_run2	Ni_Ctl-RNAi_TotD3_br1	ERS16373097	70438642	66955643	95%	68455794	97%	Nextera
	T		PTET_Ni_L4440_d3_Total_DNA_BET96	Ni_Ctl-RNAi_TotD3_br2	ERS16373098	70875700	67681714	95%	70029896	99%	Nebnext DNA fragmentase
RNAi *KU70*	S		PTET_Ni_KU70-RNAi_T25_AlgFACS_DNA_BET44_S2	KU70-RNAi_AgT25	ERS16373100	86926210	71633719	82%	86085354	99%	westburg library prep kit
	T		PTET_BET8_ADN_Ku70-RNAi	KU70-RNAi_TotD3	ERS16373101	48879906	43960790	90%	46844881	96%	westburg library prep kit
RNAi *LIG4*	S		PTET_LIG4-RNAi_T25_AlgFACS_DNA_BET55	LIG4-RNAi_AlgT25	ERS16373099	86510142	71215396	82%	86109565	100%	Nebnext DNA fragmentase
RNAi control +	S	253	PTET_Ku70-6E_L4440_T25_AlgFACS_DNA_BET45_S3	Ku70-6E_CPHG253_Ctl-RNAi_AlgT25_tr1	ERS16373102	77176836	67613556	88%	76498954	99%	westburg library prep kit
FLAG-KU70-6E	S	253	PTET_Ku70-6Er2_L4440_T25_AlgFACS_DNA_BET63	Ku70-6E_CPHG253_Ctl-RNAi_AlgT25_tr2	ERS16373103	96474916	83980671	87%	96094249	100%	Nebnext DNA fragmentase
	S	253	PTET_Ku70-6Er2_L4440_T25_AlgFACS_DNA_Covaris_BET63	Ku70-6E_CPHG253_Ctl-RNAi_AlgT25_tr3	ERS16373104	81340270	70952245	87%	80967489	100%	Covaris
	T	142	PTET_BET3_ADN_Ku70-6E_L4440_run2	Ku70-6E_CPHG142_Ctl-RNAi_TotD3	ERS16373105	61899886	59076729	95%	60256311	97%	Nextera
	T	253	PTET_Ku70-6E_L4440_d3_Total_DNA_BET47_S5	Ku70-6E_CPHG253_Ctl-RNAi_TotD3	ERS16373106	72489602	69304913	96%	71171707	98%	westburg library prep kit
	T	470	PTET_BET6_ADN_Ku70-6E_L4440_cloneR	Ku70-6E_CPHG470_Ctl-RNAi_TotD3	ERS16373107	49372454	47339826	96%	48728115	99%	westburg library prep kit
RNAi *KU70+*	S	253	PTET_Ku70-6E_KU70-RNAi_T25_AlgFACS_DNA_BET46_S4	Ku70-6E_CPHG253_KU70-RNAi_AlgT25	ERS16373108	81039144	71894849	89%	80320514	99%	westburg library prep kit
FLAG-KU70-6E	T	142	PTET_BET4_ADN_Ku70-6E_Ku70-RNAi_run2	Ku70-6E_CPHG142_KU70-RNAi_TotD3	ERS16373109	59830644	56325343	94%	57956869	97%	Nextera
	T	253	PTET_Ku70-6E_KU70-RNAi_d3_Total_DNA_BET48_S6	Ku70-6E_CPHG253_KU70-RNAi_TotD3	ERS16373110	88676292	84750173	96%	86967913	98%	westburg library prep kit
	T	470	PTET_BET7_ADN_Ku70-6E_Ku70-RNAi_cloneR	Ku70-6E_CPHG470_KU70-RNAi_TotD3	ERS16373111	54556504	52046453	95%	53556574	98%	westburg library prep kit

REFERENCES

[1].Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res *31*, 3320–3323. 10.1093/nar/gkg556.

[2].Zangarelli, C., Arnaiz, O., Bourge, M., Gorrichon, K., Jaszczyszyn, Y., Mathy, N., Escoriza, L., Bétermier, M., and Régnier, V. (2022). Developmental timing of programmed DNA elimination in Paramecium tetraurelia recapitulates germline transposon evolutionary dynamics. Genome Res *32*, 2028–2042. 10.1101/gr.277027.122.