3204–3216 *Nucleic Acids Research, 2017, Vol. 45, No. 6* *Published online 19 January 2017*
doi: 10.1093/nar/gkw1359

Multimerization properties of PiggyMac, a domesticated *piggyBac* transposase involved in programmed genome rearrangements

Emeline Dubois[1], Nathalie Mathy[1], Vinciane Régnier[1,2], Julien Bischerour[1], Céline Baudry[1], Raphaëlle Trouslard[1] and Mireille Bétermier[1,*]

[1]Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France and [2]Université Paris Diderot, Sorbonne Paris Cité, Paris, France

Received July 08, 2016; Revised December 23, 2016; Editorial Decision December 26, 2016; Accepted December 28, 2016

ABSTRACT

During sexual processes, the ciliate *Paramecium* eliminates 25–30% of germline DNA from its somatic genome. DNA elimination includes excision of ∼45 000 short, single-copy internal eliminated sequences (IESs) and depends upon PiggyMac (Pgm), a domesticated *piggyBac* transposase that is essential for DNA cleavage at IES ends. Pgm carries a core transposase region with a putative catalytic domain containing three conserved aspartic acids, and a downstream cysteine-rich (CR) domain. A C-terminal extension of unknown function is predicted to adopt a coiled-coil (CC) structure. To address the role of the three domains, we designed an *in vivo* complementation assay by expressing wild-type or mutant Pgm-GFP fusions in cells depleted for their endogenous Pgm. The DDD triad and the CR domain are essential for Pgm activity and mutations in either domain have a dominant-negative effect in wild-type cells. A mutant lacking the CC domain is partially active in the presence of limiting Pgm amounts, but inactive when Pgm is completely absent, suggesting that presence of the mutant protein increases the overall number of active complexes. We conclude that IES excision involves multiple Pgm subunits, of which at least a fraction must contain the CC domain.

INTRODUCTION

Active eukaryotic DNA transposons are mobilized through the action of their self-encoded transposase, an enzyme that catalyzes DNA cleavage at its cognate transposon ends and promotes their integration into genomic target sites (1,2). Transposon dissemination drives host genome plasticity by introducing mutations, triggering ectopic recombination, modifying gene expression patterns or establishing new gene regulatory networks (3). As a consequence of mobility, transposase-encoding genes are largely over-represented in extant sequenced genomes and metagenomes relative to other genes (4). This observation led to the suggestion that transposases may have provided a selective advantage to their hosts, not only through their impact on genome diversification, but also through exaptation and emergence of novel cellular functions, a process referred to as transposase domestication (5–7). Only a few known domesticated transposases, however, have retained a catalytic function related to that of their ancestor. Rag1, exapted from a *Transib* transposon together with its cofactor Rag2, catalyzes V(D)J recombination during the assembly of vertebrate immunoglobulin genes (8,9). Mating type switching in the yeast *Kluyveromyces lactis* is initiated by two different domesticated transposases: α3 and Kat1, which derive from *hAT* and *Mutator-like* transposable elements, respectively (10–12). DNA cleavage activity was also reported for a human domesticated *piggyBac* transposase, Pgbd5, whose cellular function is still unclear (13,14), and for SETMAR/Metnase, a primate-specific domesticated *mariner* transposase possibly involved in DNA double-strand break repair (15,16). As exemplified by SETMAR, which is a fusion between a histone methyl-transferase and a partially active transposase, domesticated transposases may have acquired additional domains that contribute to their cellular function.

The study of developmentally programmed genome rearrangements in ciliates has provided a novel example of catalytically functional domesticated transposases (17,18). Ciliates are unicellular eukaryotes, in which two distinct types of nuclei coexist in the same cytoplasm (19). The highly polyploid somatic macronucleus (MAC, ∼800 n in *Paramecium*), responsible for gene expression, is essential for cell growth and survival, while the diploid germline micronucleus is transcriptionally silent during vegetative growth. During sexual processes (conjugation between sexually complementary partners or autogamy, a self-fertilization

*To whom correspondence should be addressed. Tel: +33 169823164; Email: mireille.betermier@i2bc.paris-saclay.fr

process that can take place in some *Paramecium* species), the micronucleus undergoes meiosis and transmits the germline genome to the zygotic nucleus, while the old MAC is progressively destroyed and eventually lost. To ensure progeny survival, a functional new MAC has to differentiate from a copy of the zygotic nucleus, a process that involves large-scale genome rearrangements. In *Paramecium*, assembly of the functional new MAC genome includes elimination of 25–30 Mbp of DNA, representing 25 to 30% of the germline genome (20). DNA elimination consists in imprecise removal of repeated sequences (21) and precise excision of 45 000 short single-copy internal eliminated sequences (IES, a total of ∼3.5 Mbp), at least a fraction of which originate from ancestral *Tc/mariner* transposons (20,22). *Paramecium* IESs are flanked by two TA dinucleotides, one at each end, which are targets for programmed DNA double-strand breaks (DSB) (23). Even though IESs have derived from *Tc/mariner* elements, DNA cleavage at their boundaries requires the presence of PiggyMac (Pgm), a domesticated transposase from the distinct *piggyBac* family, which very likely plays a catalytic role in this step (17). The resulting DSBs (at chromosomal IES excision sites and on excised molecules) are repaired precisely by the classical non-homologous end-joining pathway (24), which involves a development-specific Ku70/Ku80 heterodimer (25). In addition to being essential for DSB repair, the Ku proteins interact with Pgm and are required to license DNA cleavage, indicating that IES excision depends upon pre-assembly of a multiprotein complex (25).

Pgm is a 1065-amino acid protein, in which several domains can be highlighted (Figure 1A). A core domain, characteristic of *piggyBac* transposases (Pfam domain PF13843, or DDE_Tnp_1_7), harbors three conserved aspartic acids ($D_{401}D_{491}D_{609}$) that were proposed to constitute a functional catalytic triad (17). This was further supported by studies on Tpb2p, a close homolog of Pgm that carries out IES elimination in a related ciliate, *Tetrahymena thermophila*: a triple mutation in the conserved Tpb2p DDD triad abolishes DNA cleavage *in vitro* (18) and blocks DNA elimination *in vivo* (26). Both Pgm and Tpb2p carry a cysteine-rich (CR) domain downstream of their core transposase domain, which was shown, for the *Tetrahymena* protein, to interact with heterochromatin-specific histone marks (26). Finally, a long C-terminal extension, predicted to adopt a coiled-coil structure, seems to have been acquired by ciliate-specific domesticated transposases, since it is not present in classical *piggyBac* transposases nor in their domesticated counterparts found in other species (Bouallègue, M., Rouault, J.D., Hua-Van, A., Makni, M. and Capy, P. (2016) Molecular evolution of piggyBac superfamily: from selfishness to domestication. Genome Biol. Evol., in press). The present study was initially designed to address the role of the different Pgm domains in *Paramecium tetraurelia*. For this purpose, we developed an *in vivo* complementation assay by expressing a functional Pgm-GFP fusion in cells, in which expression of endogenous *PGM* has been knocked-down, either through RNAi or somatic gene deletion. Using this assay, we established that the catalytic DDD triad and CR domain are essential for Pgm activity. Moreover, we present evidence that the C-terminal coiled-coil region is essential for productive MAC development, and that a mutant lacking this region can still form active complexes with full-length Pgm *in vivo*. Taken together, our data indicate that programmed IES excision involves assembly of multiple active Pgm subunits, among which at least some must carry an intact C-terminal domain.

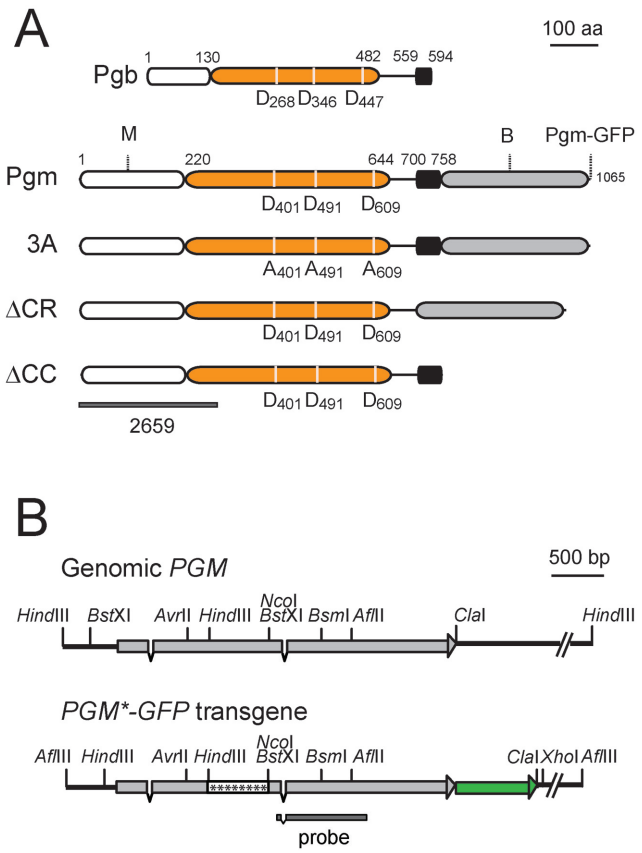

Figure 1. PiggyMac is a domesticated PiggyBac transposase. (**A**) Domain organization of the PiggyBac transposase from *Trichoplusia ni* (Pgb) and of the different derivatives of the PiggyMac (Pgm) domesticated transposase from *P. tetraurelia* that were analyzed in this study. The region presenting homology to Pfam domain DDE_Tnp_1_7, characteristic of PiggyBac-related transposases, is shown in orange (E-value = 7.8e-72 Bit score 242.4 for Pgm), with the respective position of each conserved aspartic acid (D). The Pgm_{3A} mutant (3A) harbors D to A substitutions at the three conserved positions. $Pgm_{\Delta CR}$ (ΔCR) lacks the cysteine-rich domain (black box), and $Pgm_{\Delta CC}$ (ΔCC) is deleted of the ciliate-specific C-terminal extension (in grey). The three available GFP fusions are shown: M and B represent the positions of previously published in-frame insertions of the GFP tag (17), Pgm-GFP is described in the present study. The N-terminal peptide (2659) represented by a grey bar was used to raise α-Pgm antibodies. (**B**) Restriction maps of endogenous genomic *PGM* and the plasmid carrying the *PGM*-GFP* transgene. The *PGM* gene is drawn as a grey arrow interrupted by introns and the GFP gene is in green. The synthetic mutagenized *Hind*III-*Bst*XI fragment is labeled with stars. The *PGM* probe used for Southern blot hybridization is represented as a dark grey box at the bottom (17).

MATERIALS AND METHODS

Paramecium strains and cell culture

P. tetraurelia 51 *nd7-1* was obtained from a cross between d4-2 *nd7-1* mt7 (27) and 51 new mt8 (23), followed by one round of autogamy: one viable homozygous *nd7-1*

post-autogamous F2 clone was submitted to two additional backcross/autogamy cycles to obtain F6 stock 51 *nd7-1* (25). Standard culture conditions were at 27°C in a wheat grass infusion (WGP; Pines International Inc.) inoculated with *Klebsiella pneumoniae* and supplemented with 0.8 μg/ml β-sitosterol**.** Autogamy and conjugation were carried out as described (28). During autogamy, total genomic DNA was extracted from ∼200 000 cells for each time-point as described in (17) and quantified using Qubit dsDNA assay kits (Thermo Fisher Scientific).

Molecular procedures

Oligonucleotides were purchased from Eurofins Genomics (Supplementary Table S1). PCR amplifications were performed in a final volume of 25 μl, with 10 pmol of each primer, 5 nmol of each dNTP and 1 U of DyNAzyme II or DreamTaq DNA polymerases (Thermo Scientific) according to the enzyme supplier's recommendations, and using an Eppendorf Mastercycler personal thermocycler. PCR products were analyzed on 0.8–1.5% UltraPure agarose gels (Invitrogen) or 3% NuSieve GTG agarose gels (Lonza). Sanger DNA sequencing was performed by GATC Biotech. Restriction enzymes and T4 DNA ligase were used as recommended by their supplier (New England Biolabs or Promega). Southern blot hybridization with ^{32}P-labeled probes was carried out as described (see Supplementary Data and (17)).

Excision of nested IES *51A*6649Δ29 was monitored by PCR on 20 ng total genomic DNA, using primers OMB099 and OMB100 (Supplementary Table S1) and 1.25 U DreamTaq DNA polymerase (Thermo Scientific). Unless otherwise stated, 25 cycles were performed as follows: 30 s at 95°C, 1 min at 52°C, 30 s at 72°C, with a final extension at 72°C for 5 min. PCR products were separated by electrophoresis in a 3% NuSieve GTG agarose gel (Lonza) and visualized following ethidium bromide staining. Ethidium bromide-stained gels were washed in H_2O for 10 min and acquisition of the fluorescence signal was carried out using a Typhoon scanner (GE Healthcare Life Sciences).

Microinjection of *GFP* fusion transgenes

Plasmid *PGM-GFP* and its derivatives (see Supplemental Data) were linearized with *Afl*III and microinjected into the MAC of vegetative 51 *nd7-1* cells together with a *BsaA*I-restricted *ND7*-complementing plasmid, as described previously (25). To evaluate the copy number of each injected transgene, Southern blot hybridization with ^{32}P-labeled PBL49-1/2 probe was performed on *Hind*III-restricted total genomic DNA of individual injected clones (17). All hybridization signals were collected using a Typhoon phosphorimager and quantified using the ImageQuant TL software (GE Healthcare Life Sciences). We used the *GSPATG00009109001* gene (bp 1 to 522 relative to ATG start codon, see (29)) to normalize the injected copy number of the transgene relative to the haploid genome equivalent (cphg, see Supplementary Data).

Localization of Pgm using GFP fusions or immunofluorescence

During autogamy, cells were permeabilized and fixed as described (25), then washed twice in TBST (10 mM Tris pH 7.4, 0.15 M NaCl, 0.1% Tween20) + 3% BSA (bovine serum albumin, Sigma-Aldrich). For the localization of GFP fusion proteins, fixed cells were directly transferred into TBST + 3% BSA + 0.2 μg/ml DAPI (4′,6-diamidino-2-phenylindole, Sigma Aldrich) for 5 min before being mounted in Citifluor AF2 (Biovalley). For immunofluorescence analysis, cells were incubated for 2 h at room temperature with α-Pgm 2659 primary antibodies (1:300 in Signal + solution A for immunostaining, GenTex), and washed with TBST + 3% BSA prior to 40-min incubation with Alexa fluor 488 goat anti-rabbit IgG secondary antibody (1:300, ThermoFisher Scientific), followed by DAPI staining. Epifluorescence microscopy imaging was performed using a Zeiss Axioplan 2 Imaging epifluorescence microscope using a 63x/1.4 Plan-Apochromate oil objective. Unless otherwise indicated, fluorescence signals were acquired with the same exposure time for AF488 or GFP, and identical window settings were applied to the image display using both Photoshop and ImageJ softwares, in order to compare the fluorescence intensities. Confocal imaging was carried out using a Yokogawa CSU-X1-A1 Spinning Disk coupled with a Nikon Eclipse Ti E microscope. We used a Plan-Apochromat 100x/1.40 oil objective and images were collected by an Evolve 512 Delta EM-CCD camera mounted behind a 2x magnification lens. Z-series were done with a step of 0.2 μm. Quantifications were performed on the optical section that maximizes the surface of each new MAC, using the ImageJ public software (https://imagej.nih.gov/ij/). The fluorescence intensity of each new MAC was obtained after subtracting an estimate of the local background to its Integrated Density.

Gene silencing experiments

Plasmid PGM-1, which carries the *Hind*III-*Nco*I fragment of the *PGM* gene, was used to trigger RNAi against endogenous *PGM* (17). Plasmids p0ND7c (30) and pICL7a (31) carry control RNAi inserts targeting non-essential *ND7* and *ICL7a* genes, respectively.

RNAi during autogamy was achieved as described (25), by transferring cells grown for 10 to 15 vegetative fissions in plasmid-free *Escherichia coli* HTT115 bacteria (32) to medium containing the same bacteria harboring each RNAi plasmid and induced for dsRNA production. Survival of post-autogamous progeny was tested by transferring 30 individual starved autogamous cells to standard growth medium. Cells with a functional new MAC were identified as normally-growing survivors that were unable to undergo a novel round of autogamy if starved after ∼8 vegetative divisions.

Recovery of somatic *PGM* deletions following conjugation

The RNAi-resistant *PGM*-GFP* transgene was microinjected into both 51 *nd7-1* mt7 and 51 *nd7-1* mt8 parents before mating. Conjugating pairs were transferred to *PGM* RNAi medium and, following their separation, individual

exconjugants were isolated and grown for two days in the same medium before transfer to standard growth medium. Efficient *PGM* silencing was confirmed by the failure to recover viable progeny from a control mating of non-injected parents. Among the viable F1 exconjugants obtained from injected parents, somatic *PGM* deletion was screened for by PCR and by monitoring lethality levels in their post-autogamous F2 progeny.

Anti Pgm antibodies

Peptide 2659 corresponding to Pgm amino acid sequence 2 to 288 and carrying a N-terminal His tag was expressed and purified by affinity chromatography, before rabbit immunization and purification of sera through protein A-affinity purification. The resulting α-Pgm 2659 antibodies (ProteoGenix) were further purified by antigen affinity against peptide 2659 to yield α-Pgm 2659(AF) antibodies (Covalab). Polyclonal antibodies were also raised in guinea pig against peptide 2659 and purified by antigen affinity to obtain highly specific α-Pgm 2659-GP antibodies (Proteogenix).

Protein extraction from *Paramecium* cells and Western blot analysis

Between 2 and 7.5 × 10^5 cells were collected by centrifugation at each autogamy time-point and washed in 10 mM Tris pH 7.4 before concentrated cells were frozen in liquid nitrogen. Prior to SDS-PAGE analysis, ∼1 to 2 × 10^5 cells were quickly transferred into 50 to 100 μl of boiling 10% sodium dodecyl sulphate containing 1X Protease inhibitor cocktail Set 1 (Calbiochem), and incubated at 100°C for 2 min for complete lysis. Cell lysates were kept at −20°C before gel electrophoresis. Unless otherwise indicated, electrophoresis was carried out in 4–15% pre-cast tris-glycine polyacrylamide gels (Biorad) in tris/glycine/sodium dodecyl sulphate buffer.

After transfer to 0.45 μm NC Protran nitrocellulose blotting membranes (Amersham), Pgm derivatives were detected using either α-Pgm 2659 (1:2000), α-Pgm 2659(AF) (1:2500), α-Pgm 2659-GP (1:500) polyclonal antibodies or α-GFP mouse monoclonal antibodies (1:500, Roche Diagnostics). For α-Pgm 2659 and α-Pgm 2659(AF), incubations were performed in Signal+ solution for Western blot (GenTex), whereas incubations with α-Pgm 2659-GP and α-GFP antibodies were performed in TBST (50 mM Tris-HCl, pH8; 150 mM NaCl and 0.5% Tween) containing 4% low fat milk powder. Monoclonal Anti-αβtubulin TAP952 (α-Tub) antibody (1:300) or anti-αTubulin TEU435 (α-alpha Tub) antibodies (1:300 or 1:1000) were used as loading controls. Primary antibody was detected with the species-appropriate HRP-conjugated secondary antibody (anti-rabbit IgG, anti-mouse IgG, 1/2500, Promega or anti-guinea pig IgG, 1:5000, Thermo Scientific) using the ECL detection system (WesternBright, Advansta). Signal was visualized with the LAS-3000 Image Reader (Fujifilm) and densitometric analyses were performed with the Image-Quant TL software (GE Healthcare Life Science).

Protein expression in insect cells and co-precipitation assays

Plasmid pVL1392-MBP-PGM, harboring a synthetic Pgm-coding sequence adapted to the universal genetic code (25), was used to construct pVL1392-MBP-PGM_{3A}, pVL1392-MBP-$PGM_{\Delta CR}$ and pVL1392-MBP-$PGM_{\Delta CC}$, which encode N-terminal MBP fusions to the same Pgm derivatives as those expressed in microinjected *Paramecium* cells. Plasmid pVL1392-PGM-HA was constructed by introducing the synthetic Pgm-coding sequence into the pVL-1392 vector (BD Biosciences) and adding an HA-tag to the C-terminus of Pgm. Each plasmid was transfected individually into High Five cells together with the BD BaculoGold Linearized Baculovirus DNA (BD Biosciences) to produce recombinant baculoviruses.

Recombinant baculovirus-infected High Five cells were lysed in buffer A (50 mM HEPES pH 7.5, 500 mM NaCl, 0.5 mM ethylenediaminetetraacetic acid, 1 mM dithiothreitol, 5% glycerol and 0.5% Tween-20), using a Dounce homogenizer. For the purification of MBP-tagged proteins, soluble protein extracts were incubated for 2 h with 25 μl of amylose resin (New England Biolabs). For the immunoprecipitation of HA-tagged proteins, 1.5 μg of monoclonal α-HA antibodies (HA-7 from Sigma Aldrich) were incubated overnight on a rotating wheel at 4°C with 10 μl of protein A sepharose beads (GE Healthcare), before incubating the coated beads with cell extracts for 2 h at 4°C. In both experiments, the beads were washed 3 times with 1 ml of lysis buffer A, and re-suspended in Laemmli buffer (33) before electrophoresis in polyacrylamide gels.

For the detection of Pgm-HA, western blots were incubated with HA-7 monoclonal α-HA primary antibodies (1:3000 in TBST) and anti-mouse HRP-coupled secondary antibodies (Promega). Derivatives of MBP-Pgm were detected using α-Pgm 2659 primary antibodies (1:2500 in TBST) and anti-rabbit HRP-coupled secondary antibodies (Promega).

RESULTS

A functional *PGM*-GFP* transgene resistant to RNAi against endogenous *PGM*

To set up a functional complementation assay, we used a tagged version of *PGM* that allowed us to monitor production of Pgm from the complementing transgene, follow its localization and distinguish it from endogenous Pgm. We constructed a *PGM-GFP* transgene, in which the GFP coding sequence was fused to the 3′ end of the *PGM* coding sequence, just upstream of the stop codon (Supplementary Figure S1). To obtain an RNAi-resistant transgene, the original *Hind*III-*Bst*XI fragment of *PGM* was replaced by an extensively mutagenized synthetic DNA fragment, designed to maximize nucleotide sequence divergence with the endogenous genomic locus without modifying the amino acid sequence of the encoded Pgm protein (Supplementary Figure S2). The resulting transgene was named *PGM*-GFP* (Figure 1B) and encodes a wild-type Pgm protein fused to GFP at its C-terminal end.

Following microinjection of *PGM*-GFP* into the MAC of vegetative cells, GFP fluorescence accumulated preferentially in the developing new MACs and distinct foci were

observed over a granular nucleoplasmic signal in fixed autogamous cells (Figure 2A, Supplementary Figure S3), as previously described for other GFP-tagged Pgm derivatives (17). A similar pattern was detected in live cells (Supplementary Figure S5A), indicating that foci formation is not an artefact of the fixation procedure used for sample preparation. We confirmed that GFP foci are not attributable to overexpression from the microinjected transgene, because a punctuated pattern was also observed for endogenous Pgm in fixed non-injected cells, as revealed by immunofluorescence staining using specific α-Pgm rabbit antibodies (Supplementary Figure S5B). Confocal microscopy analysis of the same slides showed a nucleoplasmic granular pattern with variably-sized foci (Supplementary Figure S5C). A dynamic localization pattern was observed for Pgm relative to large DAPI-poor regions inside the developing MAC: prominent Pgm foci were detected inside these regions at early developmental stages, before splitting into smaller foci and eventually becoming dispersed at late stages, until a diffuse Pgm signal filled the whole compartment (Supplementary Figure S5C). These nuclear DAPI-poor compartments are reminiscent of prominent spherical DNA-poor structures that were described previously in the developing and mature MAC of ciliates (34–36) and exhibit similar features to those of putative nucleoli (37). The biological significance of Pgm localization in these compartments will have to be explored in future studies. At the cell population level (Supplementary Figure S6), Western blot analysis of whole proteins extracted during an autogamy time-course of the same non-injected clone confirmed the progressive accumulation of endogenous Pgm, with a T5-T10 peak corresponding to the time when DNA DSBs are introduced at IES ends (17), and a decrease at later time-points.

In injected cells, the expression patterns of Pgm and Pgm-GFP were shown to be quite similar, as judged by Western blot analysis (Figure 2B). Fluorescence microscopy of fixed cells confirmed that Pgm-GFP is detected preferentially in early developing MACs and disappears at later stages (Supplementary Figure S3). For the particular clone analyzed in Figure 2B, the ratio between Pgm-GFP relative to endogenous Pgm (between 0.3 and 0.5) was consistent with the measured copy-number of injected *PGM*-GFP* (copy-number per haploid genome or cphg = 0.44). Within a wide-range of cphg (0.44 to 79), no toxicity was observed during vegetative growth nor following autogamy of injected cells grown under control conditions (Figure 2C, black bars), indicating that the Pgm-GFP fusion has no dominant-negative effect. To check whether Pgm-GFP is functional, injected cells were submitted to RNAi targeting the endogenous *PGM* sequence (Figure 2C, grey bars). In non-injected cells (cphg = 0), no viable progeny was obtained, confirming that the endogenous *PGM* gene is efficiently silenced. At low cphg (0.44), partial recovery of viable sexual progeny was observed (60% progeny with functional new MACs), while full rescue was obtained for higher cphg levels. To confirm that the rescue is actually due to RNAi complementation by the *PGM*-GFP* transgene rather than simple release of endogenous *PGM* silencing, we checked for the presence of Pgm and Pgm-GFP using specific α-Pgm antibodies. Endogenous Pgm was found to disappear following *PGM* RNAi (Figure 2D), while the Pgm-GFP fusion was still detectable both on western blots (Figure 2D) and by epifluorescence microscopy analysis of fixed cells (Figure 2A bottom panels). Taken together, these data show that the synthetic *PGM*-GFP* transgene is resistant to RNAi against endogenous *PGM* and encodes a functional protein.

Complementation of a *PGM* knockdown by mutant *PGM* transgenes reveals strict requirement for the DDD catalytic triad and the CR domain

Using the above functional complementation assay, we tested the activity of three different mutant derivatives of Pgm-GFP (Figure 1A). The D401A D491A D609A triple mutant transgene, designated as *PGM*$_{3A}$-GFP* or *3A*, harbored an alanine residue in place of each of the three conserved aspartic acids of the putative catalytic triad. We also constructed two deleted transgenes: *PGM*$_{\Delta CR}$-GFP* (or *ΔCR*) encodes a Pgm$_{\Delta CR}$ variant carrying a deletion of the CR domain, from residues 701 to 758, while *PGM*$_{\Delta CC}$-GFP* (or *ΔCC*) encodes a deletion of the C-terminal coiled-coil (CC) extension, from residues 759 to 1065.

Our strategy was to complement an endogenous *PGM* knockdown (KD) using each mutant transgene, with *PGM*-GFP* as a positive control. Gene KD in *P. tetraurelia* is routinely achieved through two different methods: somatic gene deletions (30) or feeding-induced RNA interference (38). To ensure that all complementing transgenes would be tested under similar *PGM* KD conditions, we opted for a somatic deletion of the endogenous *PGM* gene. In *Paramecium*, somatic deletions can be induced experimentally during sexual processes—either conjugation or autogamy—following RNAi-mediated silencing of a target gene (30 and Saudemont *et al.*, in preparation). Because Pgm is essential for the recovery of viable sexual progeny, we applied RNAi against endogenous *PGM* to conjugating cells harboring the RNAi-resistant *PGM*-GFP* transgene in their parental MAC. As described in Materials and Methods, we obtained viable exconjugants, which harbored different deletion levels of endogenous *PGM* in their new MAC and had completely lost the parental *PGM*-GFP* transgene. We selected one deleted clone (*ΔPGM*), which gave the lowest level of viable progeny after autogamy (Supplementary Figure S7B), and confirmed the somatic deletion of *PGM* by PCR, with only few residual copies of endogenous *PGM* (Supplementary Figure S7C). A significantly reduced level of endogenous Pgm in this clone (∼25% relative to a wild-type non-deleted control) was confirmed on Western blots (Figure 3A).

The *3A*, *ΔCC* and *ΔCR* transgenes, and a wild-type *PGM*-GFP* (WT) control were microinjected separately into the vegetative MAC of *ΔPGM* cells. We selected transformed clones harboring medium-range copy numbers of each transgene (cphg = 21–68), and followed their behavior during an autogamy time-course in standard growth medium. At given time-points, the localization of each GFP fusion was monitored by fluorescence microscopy, the total amount of fluorescent protein was checked on Western blots using α-GFP antibodies, and total genomic DNA was extracted to follow IES excision at the molecular level. After autogamy, the fraction of viable progeny bearing a

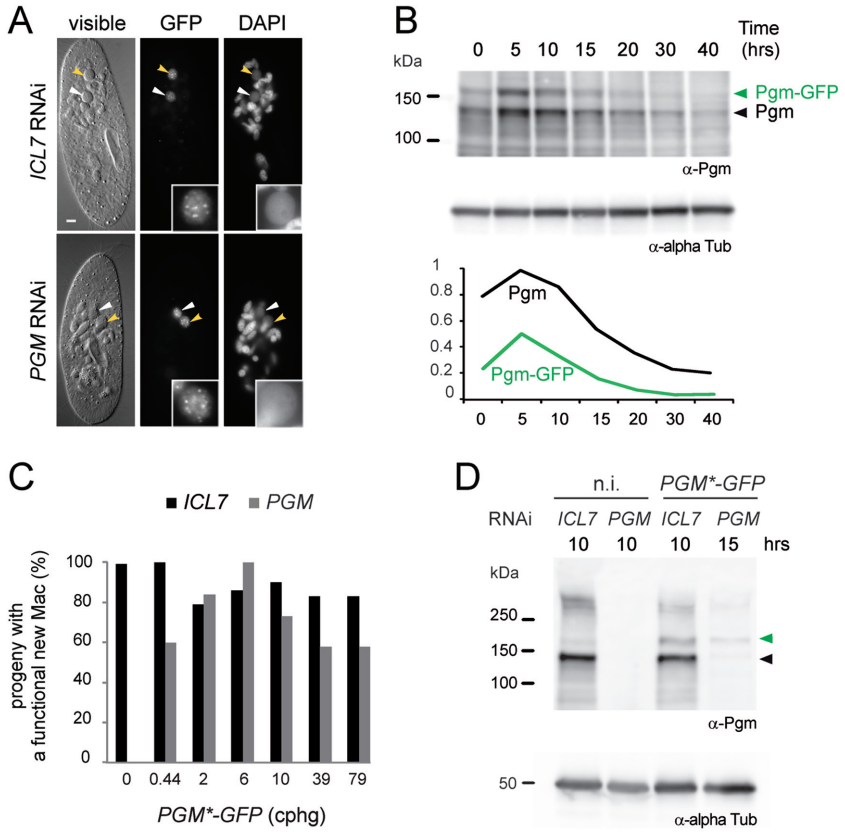

Figure 2. The Pgm-GFP fusion expressed from the *PGM*-GFP* transgene is functional. (**A**) Localization of Pgm-GFP expressed in wild-type cells following microinjection of the RNAi-resistant *PGM*-GFP* transgene (cphg = 0.44). Autogamy was monitored following RNAi against a control gene (*ICL7*) or endogenous *PGM* (Supplementary Figure S4A). The GFP fusion is expressed under the control of *PGM* endogenous transcription signals. Scale bar = 5 μm. (**B**) Expression pattern of Pgm-GFP during autogamy, relative to endogenous Pgm. *PGM*-GFP* was injected into wild-type cells (cphg = 0.44). Expected sizes: 129 kD for Pgm, 156 kD for Pgm-GFP. Top panel: Western blot of total proteins extracted during an autogamy time-course of cells grown in a control RNAi medium (*ICL7,* Supplementary Figure S4A). Pgm was revealed using purified α-Pgm 2659(AF) rabbit antibodies. The α-tubulin signal obtained with anti-αTub TEU435 antibodies (α-alpha Tub, middle panel) was used to normalize the quantification (middle panel). (**C**) Complementation of *PGM* RNAi following microinjection of the *PGM*-GFP* transgene into wild-type cells. Following microinjection, copy-numbers of the *PGM*-GFP* transgene (expressed in cphg, for copies per haploid genome) were measured for each injected clone as indicated in Materials and Methods. For each injection level, autogamy was induced in RNAi medium against a non-essential control gene (*ICL7*, except for the clone corresponding to cphg = 39, which was submitted to RNAi against *ND7*) or endogenous *PGM*. For each RNAi, the percentage of post-autogamous progeny with a functional new MAC is plotted as a function of transgene copy number (cphg). (**D**) Expression of the *PGM*-GFP* transgene is resistant to RNAi against endogenous *PGM*. Top panel: Western blot of whole proteins extracted from autogamous non-injected cells (n.i.), or from autogamous cells harboring *PGM*-GFP* (cphg = 0.44), following RNAi against control *ICL7* gene or *PGM* (see Supplementary Figure S4A for details about autogamy time-points). α-Pgm 2659-GP antibodies were used for Pgm detection. Bottom panel: Western blot of the same membrane, following stripping and incubation with α-alpha Tub antibodies.

functional new MAC was measured by transferring individual autogamous cells to fresh growth medium and letting them resume vegetative growth. Non-injected cells gave rise to 17% viable sexual progeny (Figure 3B), consistent with the presence of residual Pgm in the Δ*PGM* clone (see above). Injection of the WT transgene restored 67% viable progeny, confirming that expression of Pgm-GFP complements a *PGM* KD. Intermediate levels of viable sexual progeny (37%) were obtained in the presence of $Pgm_{\Delta CC}$-GFP, which suggests that deletion of the C-ter coiled-coil region still allows partial activity. In contrast, Pgm_{3A}-GFP and $Pgm_{\Delta CR}$-GFP were unable to complement a *PGM* KD, as indicated by the complete absence of viable sexual progeny following autogamy of microinjected cells. Strikingly, the *3A* and *ΔCR* transgenes completely prevented Δ*PGM* cells from producing viable sexual progeny, suggesting that Pgm_{3A}-GFP and $Pgm_{\Delta CR}$-GFP even have dominant-negative properties. To confirm this observation, we repeated the microinjection of *3A*, *ΔCC* and *ΔCR* into wild type (WT) cells harboring normal levels of endogenous Pgm (Figure 3C). Strong lethality was observed in the post-autogamous progeny of cells harboring *3A* or *ΔCR*, for copy-numbers above cphg = 19. In contrast, little or no lethality was observed in the progeny of cells that had been transformed with *ΔCC* or the WT *PGM*-GFP* construct.

An important control was to check whether Pgm_{3A}-GFP, $Pgm_{\Delta CC}$-GFP and $Pgm_{\Delta CR}$-GFP are produced in the micro-injected Δ*PGM* clones used for the complementation assay, and whether they localize in the new MACs during sexual processes. Western blot analysis of total protein extracts using α-GFP antibodies showed that all three mutant Pgm-GFP derivatives are indeed produced during au-

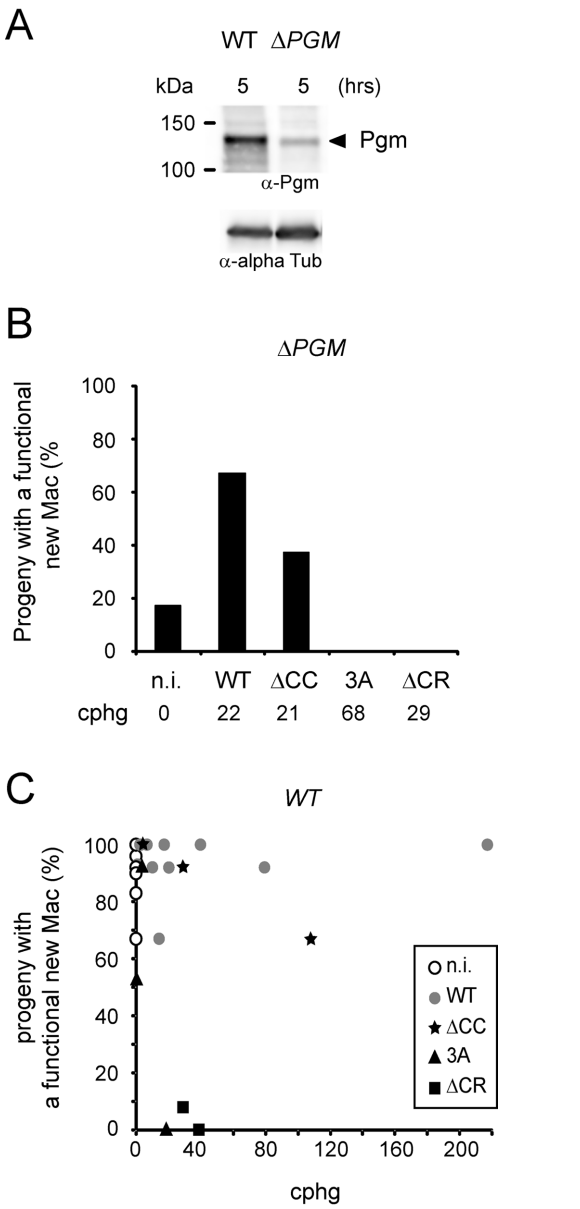

Figure 3. Complementation assay in Δ*PGM* cells using the *PGM*-GFP* transgene or its mutant *PGM*$_{\Delta CC}$-GFP*, *PGM*$_{3A}$-GFP* or *PGM*$_{\Delta CR}$-GFP* derivatives. (**A**) Western blot detection of residual endogenous Pgm in Δ*PGM* cells. Total protein extracts were prepared at comparable autogamy stages (T5, see Supplementary Figure S4B) from non-injected wild type (WT) and Δ*PGM* cells submitted to *ICL7* RNAi. Endogenous Pgm was detected using α-Pgm 2659-GP antibodies. Bottom panel: Western blot of the same membrane, following stripping and incubation with α-alpha Tub antibodies. (**B**) Complementation of a *PGM* knock-down in Δ*PGM* cells using *PGM*-GFP* (WT), *PGM*$_{\Delta CC}$-GFP* (ΔCC), *PGM*$_{3A}$-GFP* (3A) or *PGM*$_{\Delta CR}$-GFP* (ΔCR) transgenes. Each bar represents the percentage of post-autogamous progeny with a functional new MAC, for control non-injected Δ*PGM* cells (n.i.), or cells harboring each construct in their parental MAC. Transgene copy-numbers are indicated at the bottom. (**C**) Dominant-negative effect of Pgm$_{3A}$-GFP and Pgm$_{\Delta CR}$-GFP in wild-type cells. *PGM*$_{3A}$-GFP* (*3A*), *PGM*$_{\Delta CR}$-GFP* (Δ*CR*) and *PGM*$_{\Delta CC}$-GFP* (Δ*CC*) transgenes were microinjected into wild-type cells. *PGM*-GFP* (*WT*) was used as a control. Injected cells were allowed to undergo autogamy in standard culture medium. Each symbol represents, for each individual injected clone, the percentage of post-autogamous progeny harboring a functional new MAC, as a function of transgene copy-number (cphg).

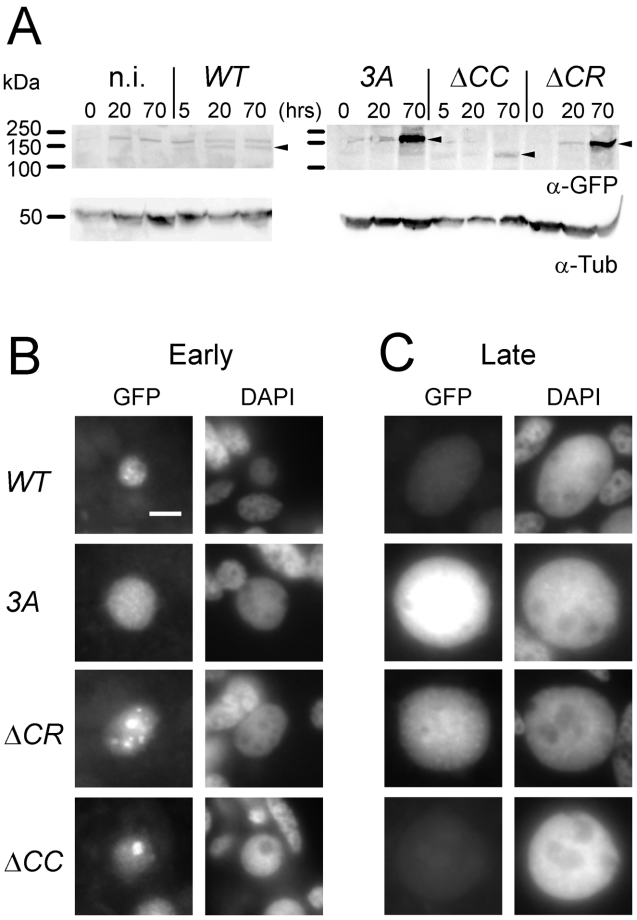

Figure 4. Expression and localization of Pgm-GFP and its mutant derivatives in Δ*PGM* cells. (**A**) Western blots showing expression of WT and mutant Pgm-GFP fusions, and the aberrant accumulation of Pgm$_{3A}$-GFP and Pgm$_{\Delta CR}$-GFP at late autogamy stages. For each microinjected clone, total proteins were extracted at different time-points during autogamy (see Supplementary Figure S4B) and loaded on 0.1% sodium dodecyl sulphate, 8% polyacrylamide gels. Western blots were first incubated with α-GFP antibodies for the detection of variant Pgm-GFP fusions (top panel) and stripped before using α-Tub monoclonal antibody. Black arrowheads indicate the position of each GFP fusion. (B) Localization of GFP fusion proteins expressed following microinjection of *PGM*-GFP* (*WT*), *PGM*$_{3A}$-GFP* (*3A*), *PGM*$_{\Delta CR}$-GFP* (Δ*CR*) or *PGM*$_{\Delta CC}$-GFP* (Δ*CC*) into Δ*PGM* cells (same injected cells as in Figure 3B). GFP signals were acquired with the same exposure time (1000 ms) using an epifluorescence microscope. Scale bar = 5 μm. The size of each developing new MAC was measured at its maximal area section using ImageJ: 18 μm^2 for *WT*, 60 μm^2 for *3A*, 55 μm^2 for Δ*CR* and 38 μm^2 for Δ*CC*. (**C**) Accumulation of Pgm$_{3A}$-GFP and Pgm$_{\Delta CR}$-GFP in late developing MACs of injected Δ*PGM* cells. GFP fluorescence signals were acquired with the same exposure time for Pgm-GFP and Pgm$_{\Delta CC}$-GFP (1000 ms), shorter acquisition times (500 ms) were used for Pgm$_{3A}$-GFP and Pgm$_{\Delta CR}$-GFP to avoid saturation (scale bar = 5 μm). Developing new MAC size was measured at the maximal area section: 122 μm^2 for *WT*, 181 μm^2 for *3A*, 172 μm^2 for Δ*CR* and 167 μm^2 for Δ*CC*.

togamy (Figure 4A and Supplementary Figure S8A). Moreover, epifluorescence microscopy confirmed the localization of GFP fluorescence in developing MACs for each fusion construct (Figure 4B). Therefore, the failure of Pgm$_{3A}$-GFP and Pgm$_{\Delta CR}$-GFP to complement a Δ*PGM* somatic deletion cannot be attributed to defective protein expression or nuclear import. Taken together, our data indicate

that Pgm_{3A}-GFP and $Pgm_{\Delta CR}$-GFP are inactive and have a dominant-negative effect in the presence of endogenous Pgm. We conclude that the putative catalytic DDD triad and the CR domain are essential during MAC development.

Different localization patterns could be noted, however, for the mutant proteins relative to wild-type Pgm-GFP (Figure 4B). The WT fusion formed nuclear foci, while Pgm_{3A}-GFP displayed a more homogeneous pattern without conspicuous foci, suggesting that Pgm foci are associated with DNA cleavage. A variable localization pattern was observed for $Pgm_{\Delta CR}$-GFP in young developing MACs, with a homogeneous GFP nuclear signal in some cells and accumulation of GFP fluorescence in DAPI-poor regions in others. Similarly, $Pgm_{\Delta CC}$-GFP tended to accumulate in DAPI-poor regions. At later stages, the global intensity of Pgm_{3A}-GFP and $Pgm_{\Delta CR}$-GFP fluorescence was consistently higher in late developing MACs than for WT Pgm-GFP, despite a twofold shorter acquisition time (Figure 4C). Strong accumulation of Pgm_{3A}-GFP and $Pgm_{\Delta CR}$-GFP at late autogamy stages was confirmed by western blot analysis (Figure 4A and Supplementary Figure S8B). The accumulation of inactive mutant Pgm derivatives at late time-points is reminiscent of our previous observation that Pgm is overexpressed at late autogamy stages in cells depleted for Ku, an essential factor involved in IES excision (25), and confirms the correlation between aberrant late Pgm accumulation and the failure to produce a functional new MAC.

The ΔCC mutant is partially active in IES excision, but only in the presence of endogenous Pgm

Expression of $Pgm_{\Delta CC}$-GFP in Δ*PGM* cells resulted in a twofold increase in the yield of viable sexual progeny relative to non-injected cells (Figure 3B). To gain molecular insight into the ability of $Pgm_{\Delta CC}$-GFP to carry out IES excision, we monitored excision of a 29-bp IES nested inside the 370-bp IES *51A*6649 from the A^{51} surface antigen gene (Figure 5A). We previously showed that quick and efficient excision of the nested IES precedes that of IES *51A*6649 in WT cells, resulting in the release of 341-bp circular molecules when the large encompassing IES eventually gets excised (39). During autogamy, excision of the nested IES is detected by PCR using primers hybridizing to the ends of encompassing IES *51A*6649: two PCR products are observed, a 370-bp band (IES+) corresponding to the non-excised germline form and a 341-bp band (Exc) resulting from excision and amplified from chromosomal or extra-chromosomal molecules (Figure 5B). For each time-point, the percentage of Exc forms relative to all PCR products provides a quantitative estimate of excision efficiency for the short nested IES. As expected for a *PGM* KD, non-excised IES molecules (IES+) were predominant throughout autogamy in the non-injected (n.i.) Δ*PGM* clone. However, consistent with the presence of residual endogenous Pgm, a low level of IES excision (∼30%) was detected at late time-points (Figure 5B). In the presence of the *WT* complementing transgene, IES excision levels reached ∼60%. In contrast, almost exclusive accumulation of IES+ PCR products was observed during autogamy of Δ*PGM* cells expressing Pgm_{3A}-GFP and, to a slightly lesser extent, $Pgm_{\Delta CR}$-GFP, indicating that both mutant proteins are inactive for IES excision. Inhibition of IES excision was even more severe in the presence Pgm_{3A}-GFP or $Pgm_{\Delta CR}$-GFP than in non-injected Δ*PGM* cells (Figure 5B), consistent with the dominant-negative effect of the two constructs on sexual progeny survival (Figure 3). A 341-bp PCR product was clearly detected in the presence of $Pgm_{\Delta CC}$-GFP (Figure 5B), indicating that this mutant supports excision of the nested 29-bp IES in the Δ*PGM* clone, with a comparable efficiency (64%) relative to the WT construct.

The Δ*PGM* cells used in the above complementation experiments expressed residual levels of endogenous Pgm (∼25% relative to wild-type, see Figure 3A), making it difficult to evaluate whether the mild phenotypes of cells harboring Δ*CC* result from the ability of $Pgm_{\Delta CC}$-GFP alone to support IES excision or whether partial complementation activity of $Pgm_{\Delta CC}$-GFP depends upon the presence of residual endogenous Pgm. To investigate this further, we took advantage of the higher efficiency of gene-specific RNAi to knock down *PGM* expression, relative to the somatic Δ*PGM* deletion: indeed, Western blots failed to detect endogenous Pgm in non-injected wild-type cells submitted to *PGM* RNAi (Figure 2D and Supplementary Figure S9). We repeated the complementation experiment by microinjecting the Δ*CC* transgene into wild-type cells, and subsequently applying RNAi against endogenous *PGM* to two injected clones carrying different copy-numbers of the transgene (cphg = 5 or 29) (Figure 5C). Under these conditions, neither of the two ΔCC-expressing clones gave rise to viable sexual progeny following *PGM* RNAi, while good survival was observed in a control RNAi. The inability of $Pgm_{\Delta CC}$-GFP to complement a *PGM* RNAi indicates that $Pgm_{\Delta CC}$ is defective in the absence of endogenous Pgm. We then diluted *PGM* RNAi-inducing bacteria in a control RNAi medium and selected a dilution factor (0.33–0.35x) that allowed detection of low residual amounts of endogenous Pgm (Supplementary Figure S9A). Using this dilution strategy, we induced full or partial *PGM* KDs in two independent clones harboring Δ*CC* and in control non-injected cells. PCR analysis of excision of the 29-bp IES internal to IES *51A*6649 confirmed that $Pgm_{\Delta CC}$ restores excision of the nested IES in partial *PGM* knockdowns, while no rescue was observed in complete *PGM* KDs (Figure 5D). Restoration of IES excision was only partial, however, since no obvious increase in the survival of sexual progeny was observed in partial PGM KDs for cells harboring ΔCC relative to the non-injected control (Supplementary Figure S9B) and the larger IES *51A*6649 was clearly retained in the developing MAC of injected cells in all *PGM* KDs (Supplementary Figure S9C). We conclude from these experiments that the C-ter coiled coil plays an essential role during MAC development, but that the truncated $Pgm_{\Delta CC}$ protein exhibits partial activity in the presence of endogenous Pgm.

Multimerization properties of Pgm

The above experiments suggest that $Pgm_{\Delta CC}$ interacts with Pgm to form an active complex, consistent with the known ability of many cut-and-paste transposases to assemble into oligomers during transposition (1). To check whether Pgm

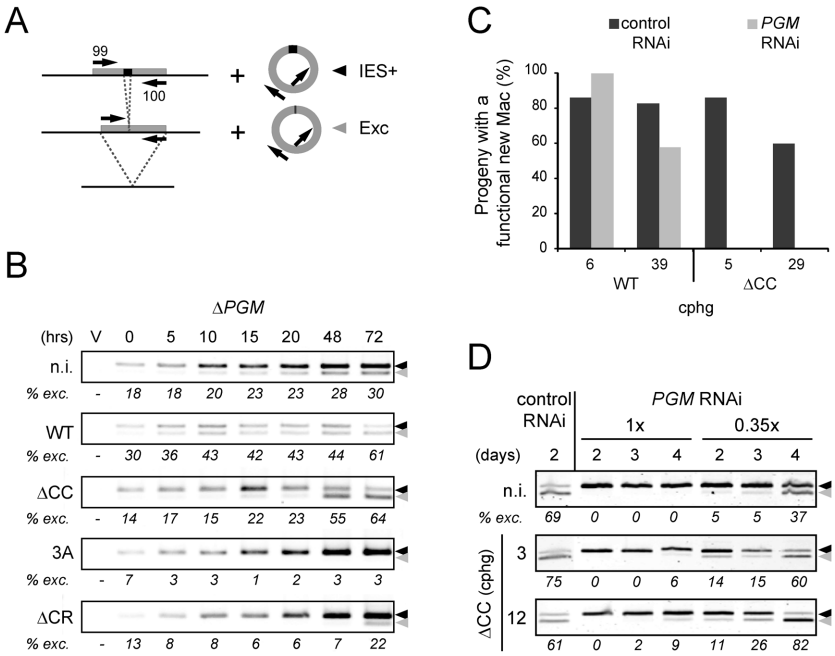

Figure 5. Partial activity of $Pgm_{\Delta CC}$-GFP depends upon the presence of endogenous *PGM*. (**A**) PCR analysis of excision of a nested 29-bp IES (*51A*6649Δ29), embedded within larger IES *51A*6649 (370 bp). Expected excision products are diagrammed, with PCR primers OMB099 and OMB100 represented by arrows. Black arrowheads: chromosomal and circular forms of IES *51A*6649 carrying the non-excised 29-bp internal IES (IES+); grey arrowheads: excision products following removal of the internal 29-bp IES (Exc). (**B**) $Pgm_{\Delta CC}$-GFP partially restores excision of the nested IES in Δ*PGM* cells. PCR reactions were performed on total genomic DNA from autogamous non-injected cells (n.i.), or cells expressing each complementing GFP fusion (see Figure 3 for injection levels and Supplementary Figure S4B for details on autogamy time-courses). The percentage of Exc forms relative to all PCR products is indicated in italics below each lane. (**C**) *PGM**$_{\Delta CC}$*-GFP* fails to restore survival of sexual progeny following *PGM* RNAi. *PGM*-GFP* (WT) or *PGM**$_{\Delta CC}$*-GFP* (Δ*CC*) transgenes were microinjected into WT cells and two injection levels (cphg) were selected for each construct. Each injected clone was submitted to RNAi against endogenous *PGM* or a control non-essential gene (*ICL7*, except for the clone harboring 39 copies of the WT transgene, which was submitted to RNAi against *ND7*). The histogram represents, for each condition, the percentage of post-autogamous progeny with a functional new MAC. (**D**) $Pgm_{\Delta CC}$-GFP partially restores excision of the nested 29-bp IES in cells submitted to inefficient *PGM* RNAi. The Δ*CC* transgene was micro-injected into the somatic MAC of individual vegetative cells from a wild-type clonal population. n.i. and transformed clones harboring different transgene copy-numbers (cphg) were submitted to control RNAi (0), or full (1x) or partial *PGM* RNAi (0,35x dilution of *PGM* RNAi-inducing bacteria in *ND7* RNAi medium). In this experiment, T0 autogamy time-points (50% of cells with fragmented old MAC) were observed following 1.5 to 2 days of starvation: day 4 is therefore equivalent to T60-T72 in the experiment shown in B. PCR was performed as indicated in Materials and Methods, with 30 cycles for day 2 samples and 35 cycles for day 3 (or day 4) samples. The percentage of IES excision is indicated in italics below each lane.

indeed forms oligomers, we co-expressed Pgm-HA with an MBP-Pgm fusion in a heterologous insect cell expression system. We found that both proteins consistently co-purify on amylose beads, while only background levels of Pgm-HA are recovered with the MBP tag alone or in the absence of an MBP-tagged partner (Figure 6A). Reciprocally, MBP-Pgm co-immunoprecipitates with Pgm-HA following incubation with anti-HA antibodies (Figure 6B). This shows that Pgm can oligomerize in insect cell extracts, as expected for a transposase-related protein. We obtained the same results when a MBP-$Pgm_{\Delta CC}$ mutant fusion (ΔCC) was co-expressed with Pgm-HA (Figure 6A and B). In the anti-HA immunoprecipitation experiments, similar amounts of MBP-tagged protein and Pgm-HA were recovered in the co-precipitate, even though exceeding amounts of MBP fusions were present in input cell extracts (Figure 6B). We conclude from these experiments that $Pgm_{\Delta CC}$ has retained the ability to interact with Pgm and that the resulting complexes are close to equimolar. The nearly 2-fold increase in sexual progeny survival observed upon injection of moderately high copy-numbers (cphg = 21) of the *ΔCC* transgene into Δ*PGM* cells expressing low but detectable levels of endogenous Pgm (Figure 3), suggests that active equimolar Pgm/$Pgm_{\Delta CC}$ complexes form instead of endogenous Pgm/Pgm multimers, therefore increasing the number of available DNA elimination complexes. Using the same experimental strategy, we observed that N-terminal MBP fusions with Pgm_{3A} and $Pgm_{\Delta CR}$ were also each able to co-precipitate with Pgm-HA in heterologous cell extracts (Figure 6A and B). The ability of the two inactive mutants, Pgm_{3A} and $Pgm_{\Delta CR}$, to associate with Pgm provides a possible explanation for their dominant-negative properties (Figure 3), because this association would give rise to inactive complexes and compete with assembly and/or function of the WT DNA elimination complex (see Discussion).

DISCUSSION

Transgene-mediated complementation of gene knockdowns in *P. tetraurelia*: a tool for functional domain analysis

Gene functional analyses are often carried out through siRNA-mediated silencing of endogenous genes and complementation with RNAi-resistant transgenes (see 40–42). In *P. tetraurelia*, RNAi is generally induced by feeding cells with dsRNA-producing bacteria (38). Complementation of

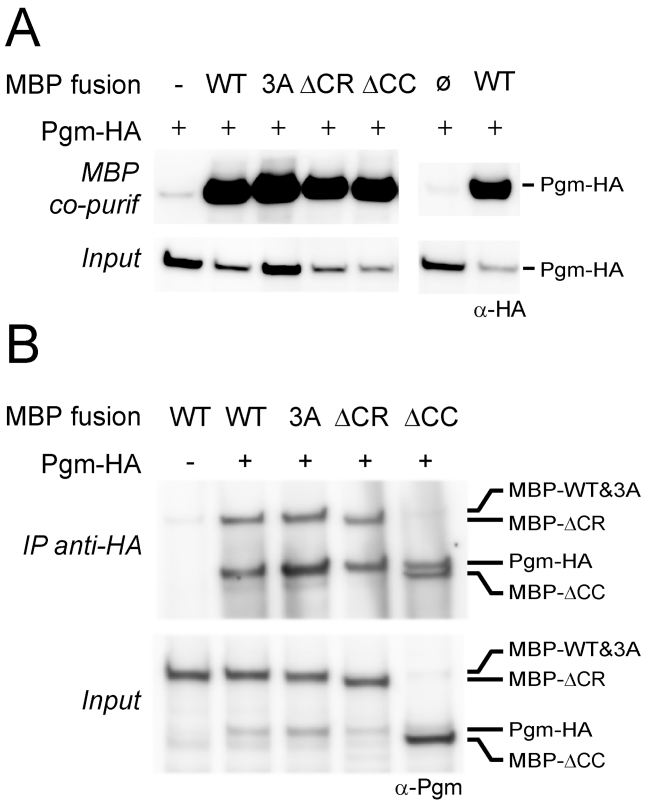

Figure 6. Multimerization properties of Pgm and its mutant derivatives. (**A**) Co-precipitation of Pgm-HA with N-terminal MBP fusions of Pgm or its mutant variants (3A, ΔCR, ΔCC) on amylose beads. The presence of Pgm-HA in the precipitate was revealed on Western blots using α-HA antibodies. Top right panel: control co-precipitation of Pgm-HA with MBP alone (*Ø*) or the MBP-Pgm fusion (WT). Bottom panels: immunodetection of Pgm-HA in input cell extracts. (**B**) Co-immunoprecipitation of N-terminal MBP fusions of Pgm and its variants with Pgm-HA, using α-HA antibodies. All proteins were revealed on Western blots using α-Pgm 2659 antibodies. Top panel: co-precipitated fractions; bottom panel: input cell extracts.

RNAi was achieved for the first time in *Paramecium* by expressing the human ortholog of a silenced gene (43). Here, we show that a *Paramecium* transgene (*PGM**) carrying multiple silent base substitutions within the region targeted by bacterial dsRNAs is resistant to RNAi against endogenous *PGM* and can be used for functional complementation assays. Our strategy might have been complicated by the previous observation that feeding-induced RNAi triggers the production of secondary siRNAs complementary to the target mRNA outside of the region homologous to the dsRNAs produced from the feeding insert (44). Secondary siRNAs might have been able to anneal to *PGM** mRNA outside of the mutagenized region. However, we observed efficient complementation, confirming that secondary siRNAs have little or no impact on RNAi (44). We also conclude that the Pgm-GFP fusion used in this study is functional. Maximal complementation levels were observed for 1 to 10 copies of the transgene per haploid genome. Moreover, we found that higher copy-numbers of the *PGM*-GFP* transgene do not strongly inhibit MAC development, which suggests that no overproduction inhibition, or OPI (45,46), takes place when Pgm is present in excess. Our observation is consistent with the previously reported absence of OPI for the *T. ni* PiggyBac transposase in human cells (47).

Complementation was also observed in the absence of dsRNA-producing bacteria, following microinjection of *PGM*-GFP* into cells harboring a *PGM* somatic deletion. Use of the same Δ*PGM* clone for all microinjections allowed accurate comparison of the relative activities of complementing mutant transgenes. Moreover, the presence of residual levels of endogenous Pgm in Δ*PGM* cells allowed us to detect the dominant-negative properties of the Pgm_{3A} and $Pgm_{\Delta CR}$ complementing constructs. We found that Pgm_{3A} is inactive, which shows that the conserved DDD triad is essential for IES excision and that Pgm plays an essential catalytic role in DNA cleavage at IES ends. A similar finding was reported previously for the *Tetrahymena* Pgm homolog, Tpb2p (18,26), highlighting the functional importance of catalytically active domesticated *piggyBac* transposases in ciliates. The inability of $Pgm_{\Delta CR}$ to support IES excision indicates that the CR domain is also essential. Further investigations will clarify its exact function. It may interact with the H3K27me3 heterochromatin mark involved in the control of programmed DNA elimination in *Paramecium* (35), as reported for Tpb2p (26). The CR domain may also interact with other targets, such as nucleic acids, as demonstrated for some CR domains (48). Moreover, the dominant-negative properties of Pgm_{3A} and $Pgm_{\Delta CR}$ indicate that either mutant inhibits the normal action of WT Pgm, similarly to what was proposed in other systems for inactive transposase variants that act as natural repressors (46,49). Alternatively, Pgm_{3A} and $Pgm_{\Delta CR}$ may associate with and titrate out essential IES excision factors, such as the Ku proteins (25), or compete with Pgm for binding to its natural chromatin targets.

The ciliate-specific C-ter coiled-coil extension is essential for activity of the Pgm-associated IES excision complex

DNA 'cut-and-paste' transposition mediated by DD(E/D) transposases involves assembly of a highly ordered protein–DNA complex, in which both transposon ends, each bound by one or more transposase subunits, are synapsed together through protein–protein interactions to assemble an active transpososome (1). Whatever the number of transposase subunits present in the synaptic complex, a single active catalytic site carries out all DNA cleavage and strand transfer reactions at each end. *PiggyBac*-related transposons are extensively used as tools for genome engineering (50). However, still little is known about the structural and functional mechanisms that underlie transpososome assembly for transposons of the *piggyBac* family, and information is lacking on the stoichiometry of the transposition reaction. The present work shows that the *Paramecium* domesticated transposase Pgm acts as a multi-subunit complex *in vivo*, as described for canonical transposases. Indeed, the ability of $Pgm_{\Delta CC}$ to rescue IES excision only when residual amounts of Pgm are present supports a mechanistic model, in which productive genome rearrangements rely on assembly of an active excision complex including several Pgm subunits. Co-precipitation experiments (Figure 6) show that Pgm oligomerizes in solution and that Pgm and $Pgm_{\Delta CC}$

can also interact, indicating that one oligomerization domain lies within the transposase-related portion of the protein. We propose, therefore, that a complex including both Pgm and $Pgm_{\Delta CC}$ is formed *in vivo* and that a Pgm/$Pgm_{\Delta CC}$ multimer is active for IES excision, at least partially. Alternatively, individual subunits within the complex may each carry out one step of the excision reaction without directly interacting with each other, and $Pgm_{\Delta CC}$ may only be active for one step. Further studies will establish the exact number of Pgm subunits present in the IES excision complex and whether direct Pgm multimerization is actually required for catalytic activity. Similar activation of a wild-type transposase by inactive truncated mutants was reported for other transposons, such as the plant *Ac* element (51), suggesting that truncated transposases may sometimes retain the ability to form an active transposition complex with their cognate full-length counterpart. Interestingly, $Pgm_{\Delta CC}$ is defective if full-length Pgm is absent: therefore, an active IES excision complex requires the presence of the C-terminal domain in at least some Pgm subunits.

The C-terminal domain of Pgm was proposed to adopt a coiled-coil structure (17). Coiled coils are generally involved in protein–protein interactions, either protein oligomerization or interaction with other partners (52,53). According to the Multicoil sequence prediction software (54), the Pgm coiled-coil domain may even preferentially form trimeric structures. Interestingly, the coiled coil has been acquired by ciliate-specific domesticated *piggyBac* transposases (Bouallègue *et al*., in press). IES excision in *Paramecium* is a transposition-related mechanism, during which both ends of a given IES establish a long-range interaction, as reported for cut-and-paste transposons (55). However, because of their short size (30% are 26–30 bp in length) (20), *Paramecium* IESs are probably not flexible enough to bend easily and form a DNA loop. Instead, higher-order interaction between the coiled coils of Pgm subunits bound at each IES boundary may contribute to assemble a catalytically active 'excisome'. We also note that programmed DNA elimination in ciliates differs from standard cut-and-paste transposition in two other aspects. First of all, no recognizable conserved DNA motif was identified that may constitute a specific Pgm binding site on germline eliminated sequences (56). This may be linked to general properties of transposases from the *piggyBac* family or their domesticated derivatives. Indeed, previously published work demonstrated that binding of the purified *piggyBac* transposase from *Trichoplusia ni* (PB) to DNA fragments carrying the transposon left or right terminal inverted repeats induces multiple shifted bands, but sequence-specific binding motifs were not determined experimentally (57). Recently, the vertebrate domesticated *piggyBac* transposase, Pgbd5 (58), which is very distant from PB and Pgm (Bouallègue *et al*., in press), was shown to mobilize the *piggyBac* transposon from *T. ni* (14) and to trigger genomic rearrangements when overexpressed in human cells (59). Pgbd5-specific signal sequences were identified at the rearranged sites, but have little to do with the inverted repeats of the original *piggyBac* transposon, suggesting that relaxed sequence specificity may be a common property of domesticated *piggyBac* transposases. A second difference between IES excision and cut-and-paste transposition is the strong coupling that exists between DNA cleavage and DSB repair at IES excision sites (25). Taking these differences into account, an alternative function may be proposed for the Pgm coiled-coil domain, i.e. interacting with other components of the DNA elimination machinery, perhaps involved in IES recognition or DSB repair. Future studies will provide more insight into the composition of the DNA elimination complex and how the Pgm C-ter coiled coil is involved in interacting with other components of the complex.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank Anne Aubusson for the generous gift of monoclonal anti-Tubulin antibodies and her help in confocal microscopy, Kamila Malizsewska-Olejniczak for rescuing our stock of 51 *nd7-1* mt7 cells, Baptiste Saudemont and Eric Meyer for sharing their unpublished method to induce somatic deletions during mating of *Paramecium* cells, and Maryem Bouallègue and Pierre Capy for sharing their data before publication. The authors are grateful to Arthur Abello and Marc Guérineau for stimulating discussions, and to Sophie Malinsky and Linda Sperling for critical reading of the manuscript. The present work has benefited from the facilities and expertise of the light microscopy facilities of Imagerie-Gif, (http://www.i2bc.paris-saclay.fr/spip.php?article278). This core facility is supported by the French national Research Agency under Investments for the Future programs 'France-BioImaging', and the Labex 'Saclay Plant Science' (ANR-10-INSB-04-01 and ANR-11-IDEX-0003-02, respectively).

FUNDING

Intramural CNRS and Agence Nationale pour la Recherche [BLAN08-3_310945 'PARADICE', ANR-10-BLAN-1603 'GENOMAC', ANR-14-CE10-0005-01 'PIGGYPACK']; ARC Foundation for Cancer Research [#SFI20121205487 to M.B.]. Funding for open access charge: Agence Nationale pour la Recherche [ANR-14-CE10-0005-01 'PIGGYPACK'].
Conflict of interest statement. None declared.

REFERENCES

1. Hickman,A.B. and Dyda,F. (2015) Mechanisms of DNA transposition. *Microbiol. Spectr.*, **3**, MDNA3-0034-2014.
2. Wicker,T., Sabot,F., Hua-Van,A., Bennetzen,J.L., Capy,P., Chalhoub,B., Flavell,A., Leroy,P., Morgante,M., Panaud,O. *et al.* (2007) A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.*, **8**, 973–982.
3. Levin,H.L. and Moran,J.V. (2011) Dynamic interactions between transposable elements and their hosts. *Nat. Rev. Genet.*, **12**, 615–627.
4. Aziz,R.K., Breitbart,M. and Edwards,R.A. (2010) Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Res.*, **38**, 4207–4217.
5. Sinzelle,L., Izsvak,Z. and Ivics,Z. (2009) Molecular domestication of transposable elements: from detrimental parasites to useful host genes. *Cell Mol. Life Sci.*, **66**, 1073–1093.
6. Volff,J.N. (2006) Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. *Bioessays*, **28**, 913–922.

7. Feschotte,C. and Pritham,E.J. (2007) DNA transposons and the evolution of eukaryotic genomes. *Annu. Rev. Genet.*, **41**, 331–368.
8. Carmona,L.M., Fugmann,S.D. and Schatz,D.G. (2016) Collaboration of RAG2 with RAG1-like proteins during the evolution of V(D)J recombination. *Genes Dev.*, **30**, 909–917.
9. Teng,G. and Schatz,D.G. (2015) Regulation and Evolution of the RAG Recombinase. *Adv. Immunol.*, **128**, 1–39.
10. Barsoum,E., Martinez,P. and Astrom,S.U. (2010) Alpha3, a transposable element that promotes host sexual reproduction. *Genes Dev.*, **24**, 33–44.
11. Chiruvella,K.K., Rajaei,N., Jonna,V.R., Hofer,A. and Astrom,S.U. (2016) Biochemical characterization of Kat1: a domesticated hAT-transposase that induces DNA hairpin formation and MAT-switching. *Sci. Rep.*, **6**, 21671.
12. Rajaei,N., Chiruvella,K.K., Lin,F. and Astrom,S.U. (2014) Domesticated transposase Kat1 and its fossil imprints induce sexual differentiation in yeast. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 15491–15496.
13. Sarkar,A., Sim,C., Hong,Y.S., Hogan,J.R., Fraser,M.J., Robertson,H.M. and Collins,F.H. (2003) Molecular evolutionary analysis of the widespread *piggyBac* transposon family and related 'domesticated' sequences. *Mol. Genet. Genomics*, **270**, 173–180.
14. Henssen,A.G., Henaff,E., Jiang,E., Eisenberg,A.R., Carson,J.R., Villasante,C.M., Ray,M., Still,E., Burns,M., Gandara,J. *et al.* (2015) Genomic DNA transposition induced by human PGBD5. *Elife*, **4**, e10565.
15. Liu,D., Bischerour,J., Siddique,A., Buisine,N., Bigot,Y. and Chalmers,R. (2007) The human SETMAR protein preserves most of the activities of the ancestral Hsmar1 transposase. *Mol. Cell. Biol.*, **27**, 1125–1132.
16. Shaheen,M., Williamson,E., Nickoloff,J., Lee,S.H. and Hromas,R. (2010) Metnase/SETMAR: a domesticated primate transposase that enhances DNA repair, replication, and decatenation. *Genetica*, **138**, 559–566.
17. 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.
18. Cheng,C.Y., Vogt,A., Mochizuki,K. and Yao,M.C. (2010) A domesticated piggyBac transposase plays key roles in heterochromatin dynamics and DNA cleavage during programmed DNA deletion in *Tetrahymena thermophila*. *Mol. Biol. Cell*, **21**, 1753–1762.
19. Prescott,D.M. (1994) The DNA of ciliated protozoa. *Microbiol. Rev.*, **58**, 233–267.
20. Arnaiz,O., Mathy,N., Baudry,C., Malinsky,S., Aury,J.M., Denby-Wilkes,C., Garnier,O., Labadie,K., Lauderdale,B.E., Le Mouel,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.
21. Le Mouël,A., Butler,A., Caron,F. and Meyer,E. (2003) Developmentally regulated chromosome fragmentation linked to imprecise elimination of repeated sequences in *Paramecium*. *Eukaryot. Cell*, **2**, 1076–1090.
22. Dubois,E., Bischerour,J., Marmignon,A., Mathy,N., Régnier,V. and Bétermier,M. (2012) Transposon invasion of the *Paramecium* germline genome countered by a domesticated PiggyBac transposase and the NHEJ pathway. *Int. J. Evol. Biol.*, **2012**, 436196.
23. Gratias,A. and Bétermier,M. (2003) Processing of double-strand breaks is involved in the precise excision of *Paramecium* IESs. *Mol. Cell. Biol.*, **23**, 7152–7162.
24. Kapusta,A., Matsuda,A., Marmignon,A., Ku,M., Silve,A., Meyer,E., Forney,J., 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.
25. Marmignon,A., Bischerour,J., Silve,A., Fojcik,C., Dubois,E., Arnaiz,O., Kapusta,A., Malinsky,S. and Betermier,M. (2014) Ku-mediated coupling of DNA cleavage and repair during programmed genome rearrangements in the ciliate paramecium tetraurelia. *PLoS Genet.*, **10**, e1004552.
26. Vogt,A. and Mochizuki,K. (2013) A domesticated PiggyBac transposase interacts with heterochromatin and catalyzes reproducible DNA elimination in *Tetrahymena*. *PLoS Genet.*, **9**, e1004032.
27. Skouri,F. and Cohen,J. (1997) Genetic approach to regulated exocytosis using functional complementation in *Paramecium*: identification of the *ND7* gene required for membrane fusion. *Mol. Biol. Cell*, **8**, 1063–1071.
28. Bétermier,M., Duharcourt,S., Seitz,H. and Meyer,E. (2000) Timing of developmentally programmed excision and circularization of *Paramecium* internal eliminated sequences. *Mol. Cell. Biol.*, **20**, 1553–1561.
29. Arnaiz,O. and Sperling,L. (2011) ParameciumDB in 2011: new tools and new data for functional and comparative genomics of the model ciliate *Paramecium tetraurelia*. *Nucleic Acids Res.*, **39**, D632–D636.
30. Garnier,O., Serrano,V., Duharcourt,S. and Meyer,E. (2004) RNA-mediated programming of developmental genome rearrangements in *Paramecium tetraurelia*. *Mol. Cell. Biol.*, **24**, 7370–7379.
31. Gogendeau,D., Klotz,C., Arnaiz,O., Malinowska,A., Dadlez,M., de Loubresse,N.G., Ruiz,F., Koll,F. and Beisson,J. (2008) Functional diversification of centrins and cell morphological complexity. *J. Cell Sci.*, **121**, 65–74.
32. Timmons,L., Court,D.L. and Fire,A. (2001) Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in *Caenorhabditis elegans*. *Gene*, **263**, 103–112.
33. Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
34. Jurand,A., Beale,G.H. and Young,M.R. (1964) Studies on the Macronucleus of *Paramecium aurelia*. II. Development of Macronuclear Anlagen. *J. Protozool.*, **11**, 491–497.
35. Lhuillier-Akakpo,M., Frapporti,A., Denby Wilkes,C., Matelot,M., Vervoort,M., Sperling,L. and Duharcourt,S. (2014) Local effect of Enhancer of zeste-like reveals cooperation of epigenetic and cis-acting determinants for zygotic genome rearrangements. *PloS Genet.*, **10**, e1004665.
36. Postberg,J., Alexandrova,O., Cremer,T. and Lipps,H.J. (2005) Exploiting nuclear duality of ciliates to analyse topological requirements for DNA replication and transcription. *J. Cell Sci.*, **118**, 3973–3983.
37. Sleeman,J.E. and Trinkle-Mulcahy,L. (2014) Nuclear bodies: new insights into assembly/dynamics and disease relevance. *Curr. Opin. Cell Biol.*, **28**, 76–83.
38. Galvani,A. and Sperling,L. (2002) RNA interference by feeding in *Paramecium*. *Trends Genet.*, **18**, 11–12.
39. Gratias,A. and Bétermier,M. (2001) Developmentally programmed excision of internal DNA sequences in *Paramecium aurelia*. *Biochimie*, **83**, 1009–1022.
40. Palatnik,J.F., Allen,E., Wu,X., Schommer,C., Schwab,R., Carrington,J.C. and Weigel,D. (2003) Control of leaf morphogenesis by microRNAs. *Nature*, **425**, 257–263.
41. Kittler,R., Pelletier,L., Ma,C., Poser,I., Fischer,S., Hyman,A.A. and Buchholz,F. (2005) RNA interference rescue by bacterial artificial chromosome transgenesis in mammalian tissue culture cells. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 2396–2401.
42. Schulz,J.G., David,G. and Hassan,B.A. (2009) A novel method for tissue-specific RNAi rescue in Drosophila. *Nucleic Acids Res.*, **37**, e93.
43. Laligne,C., Klotz,C., de Loubresse,N.G., Lemullois,M., Hori,M., Laurent,F.X., Papon,J.F., Louis,B., Cohen,J. and Koll,F. (2010) Bug22p, a conserved centrosomal/ciliary protein also present in higher plants, is required for an effective ciliary stroke in Paramecium. *Eukaryot. Cell*, **9**, 645–655.
44. Carradec,Q., Gotz,U., Arnaiz,O., Pouch,J., Simon,M., Meyer,E. and Marker,S. (2015) Primary and secondary siRNA synthesis triggered by RNAs from food bacteria in the ciliate *Paramecium tetraurelia*. *Nucleic Acids Res.*, **43**, 1818–1833.
45. Bire,S., Casteret,S., Arnaoty,A., Piegu,B., Lecomte,T. and Bigot,Y. (2013) Transposase concentration controls transposition activity: myth or reality? *Gene*, **530**, 165–171.
46. Lohe,A.R. and Hartl,D.L. (1996) Autoregulation of mariner transposase activity by overproduction and dominant-negative complementation. *Mol. Biol. Evol.*, **13**, 549–555.
47. Wilson,M.H., Coates,C.J. and George,A.L. Jr (2007) PiggyBac transposon-mediated gene transfer in human cells. *Mol. Ther.*, **15**, 139–145.

48. Meyer,S. and Kieffer,B. (2015) Protein motifs: Zinc fingers. In: *eLS*. John Wiley & Sons, Ltd, Chichester, doi:10.1002/9780470015902.a0020395.
49. de la Cruz,N.B., Weinreich,M.D., Wiegand,T.W., Krebs,M.P. and Reznikoff,W.S. (1993) Characterization of the Tn5 transposase and inhibitor proteins: a model for the inhibition of transposition. *J. Bacteriol.*, **175**, 6932–6938.
50. Yusa,K. (2015) piggyBac Transposon. *Microbiol. Spectr.*, **3**, MDNA3-0028-2014.
51. Kunze,R., Behrens,U., Courage-Franzkowiak,U., Feldmar,S., Kuhn,S. and Lutticke,R. (1993) Dominant transposition-deficient mutants of maize Activator (Ac) transposase. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 7094–7098.
52. Hohl,M., Kwon,Y., Galvan,S.M., Xue,X., Tous,C., Aguilera,A., Sung,P. and Petrini,J.H. (2011) The Rad50 coiled-coil domain is indispensable for Mre11 complex functions. *Nat. Struct. Mol. Biol.*, **18**, 1124–1131.
53. Sibanda,B.L., Critchlow,S.E., Begun,J., Pei,X.Y., Jackson,S.P., Blundell,T.L. and Pellegrini,L. (2001) Crystal structure of an Xrcc4-DNA ligase IV complex. *Nat. Struct. Biol.*, **8**, 1015–1019.
54. Wolf,E., Kim,P.S. and Berger,B. (1997) MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci.*, **6**, 1179–1189.
55. 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.
56. Bétermier,M. and Duharcourt,S. (2014) Programmed rearrangement in ciliates: *Paramecium*. *Microbiol. Spectr.*, **2**, MDNA3-0035-2014.
57. Mitra,R., Fain-Thornton,J. and Craig,N.L. (2008) *piggyBac* can bypass DNA synthesis during cut and paste transposition. *EMBO J.*, **27**, 1097–1109.
58. Pavelitz,T., Gray,L.T., Padilla,S.L., Bailey,A.D. and Weiner,A.M. (2013) PGBD5: a neural-specific intron-containing piggyBac transposase domesticated over 500 million years ago and conserved from cephalochordates to humans. *Mob. DNA*, **4**, 23.
59. Henssen,A.G., Jiang,E., Zhuang,J., Pinello,L., Socci,N.D., Koche,R., Gonen,M., Villasante,C.M., Armstrong,S.A., Bauer,D.E. *et al.* (2016) Forward genetic screen of human transposase genomic rearrangements. *BMC Genomics*, **17**, 548.