

Title: Nuclease proteins CifA and CifB promote spermatid DNA damage associated with symbiont-induced cytoplasmic incompatibility

Rupinder Kaur^{1,2*}, J. Dylan Shropshire^{3,4}, Brittany A. Leigh^{1,2}, Seth R. Bordenstein^{1,2,5,6*}

Affiliation

¹Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA

²Vanderbilt Microbiome Innovation Center, Vanderbilt University, Nashville, TN, USA

³Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA

⁴Department of Biological Sciences, Lehigh University, Bethlehem, PA 18015, USA

⁵Department of Pathology, Microbiology & Immunology, Vanderbilt University Medical Center, Nashville, TN, USA

⁶Vanderbilt Institute for Infection, Immunology and Inflammation, Vanderbilt University Medical Center, Nashville, TN, USA

*Correspondence:

Rupinder Kaur, Nashville, TN, 37235, E-Mail: rupinder.kaur@vanderbilt.edu

Seth R. Bordenstein, Nashville, TN, 37235, E-Mail: s.bordenstein@vanderbilt.edu

Keywords: *Wolbachia*, cytoplasmic incompatibility, Cif protein, nuclease activity, spermatid DNA damage

Summary

The worldwide endosymbiosis between arthropods and *Wolbachia* bacteria is an archetype for reproductive parasitism. This parasitic strategy rapidly increases the proportion of symbiont-transmitting mothers, and the most common form, cytoplasmic incompatibility (CI), impacts insect evolution and arboviral control strategies. During CI, sperms from symbiotic males kill embryos of aposymbiotic females via two nuclear-targeting proteins, CifA and CifB, that alter sperm chromatin organization in *Drosophila melanogaster*. Here we hypothesize that Cif proteins metabolize nucleic acids of developing sperm to initiate genome integrity changes. Using *in vitro* and *in situ* transgenic, mutant, enzymatic, and cytochemical assays, we show that CifA is a previously-unrecognized DNase and RNase, and CifB is a DNase. Notably, *in vitro* nuclease activity translates to *in situ* spermatid DNA damage at the canoe stage of spermiogenesis. Evolution-guided mutations ablate Cif enzymatic activity. Nucleic acid metabolism by Cif enzymes expands a fundamental understanding of the mechanism of symbiont-mediated reproductive parasitism.

Introduction

Wolbachia are maternally inherited, obligate intracellular bacteria that occur worldwide in 40-65% of arthropod species (Hilgenboecker et al., 2008; Kaur et al., 2021; Weinert et al., 2015; Zug and Hammerstein, 2012). Many *Wolbachia* strains selfishly alter host reproduction to increase the relative number of symbiotic females that transmit the bacteria to the next generation (Hoffmann et al., 1990; Hoffmann and Turelli, 1997; Kriesner et al., 2013). Cytoplasmic incompatibility (CI) is the most studied reproductive phenotype with major impacts on arthropod evolution and vector control. Specifically, CI results in embryonic death when symbiotic males mate with aposymbiotic females. Nullification of death, and thus rescue of CI, occurs when transmitting females harbor the same strain of *Wolbachia* (Shropshire et al., 2020). CI-causing mosquitoes are released in both population suppression (Caputo et al., 2020; Mains et al., 2016; Puggioli et al., 2016; Zheng et al., 2019) and population replacement strategies (Flores and O'Neill, 2018; Indriani et al., 2020; Nazni et al., 2019; Ryan et al., 2020) to curb arbovirus transmission of Dengue, Zika, and Chikungunya (Caragata et al., 2016; Crawford et al., 2020; Dobson et al., 2002; Hoffmann et al., 2011; Moreira et al., 2009; O'Neill, 2018; O'Neill et al., 2018). Thus, CI is at the forefront of controlling mosquito-borne diseases globally (Bourtzis et al., 2014; O'Neill et al., 2018; Utarini et al., 2021).

CI is caused by the germline expression of two cytoplasmic incompatibility factor genes, *cifA* and *cifB*, in *wMel*-infected *Drosophila melanogaster* testes (Beckmann et al., 2017; LePage et al., 2017), and *cifA* expression in ovaries rescues CI (Shropshire et al., 2018). Thus, CI is governed by the Two-by-One genetic model demonstrated in various systems (Beckmann et al., 2019; Chen et al., 2019; LePage et al., 2017; Shropshire et al., 2021b; Shropshire and Bordenstein, 2019). *cifB* alone can also cause *cifA*-rescuable CI under a One-by-One genetic model (Adams et al., 2021; Sun et al., 2021). Therefore, deciphering the Cif molecular mechanism(s) and potentially universal features of CI and rescue remains a central goal for understanding how sexual reproduction serves as a battleground for host-*Wolbachia* interactions.

CifA and CifB proteins from *wMel* of *D. melanogaster* and *wPip* of *Culex pipiens* were recently found to invade the nuclei of developing sperm (Horard et al.,

2022; Kaur et al., 2022). Specifically, dual expression of wMel CifA and CifB alters sperm chromatin integrity before fertilization via modulating the histone-protamine transition of spermiogenesis (Kaur et al., 2022), and wPip CifB expression impacts sperm DNA stress in the embryo (Horard et al., 2022). Sperm nuclear access by wMel Cifs indicates that Cifs may directly interact with sperm DNA and/or RNA to modify sperm genome integrity and contribute to the CI mechanism. Thus, we hypothesized that Cifs may metabolize sperm nucleic acids by their enzymatic properties.

Enzymatic functions of CifA have not been tested, while *in vitro* enzymatic properties of some of the CifB variants have been evaluated. T1 wPip CifB is a deubiquitinase (DUB) that cleaves poly-ubiquitin chains *in vitro* to cause CI (Beckmann et al., 2017). However, mutating the catalytic residue in CifB's DUB domain ablated CI in one study (Beckmann et al., 2017) and maintains CI in another (Horard et al., 2022), suggesting that DUB enzymatic function may not be essential for CI. T4 wPip CifB is a DNase with active PDDEXK nuclease domains that cleave DNA substrates (Chen et al., 2019). Presumptive activation of PDDEXK nuclease sites with artificial amino acid substitutions in T1 wPip CifB did not degrade DNA *in vitro* (Chen et al., 2019), suggesting these domains are inactive in T1 CifB variants. Notably, the enzymatic activity of wild-type T1 wMel CifB remains untested to date. Mutating conserved amino acid residues in the nuclease and other domains of wMel CifB ablates CI (Shropshire et al., 2020), suggesting that nuclease domains and various other sites throughout CifB are crucial for CI, thus warranting *in vitro* and *in situ* enzymatic characterization.

CifA is weakly predicted to encode a domain of unknown function (DUF3243) with homology to a Puf-family RNA-binding domain found throughout eukaryotes, and mutating conserved residues in this domain notably ablates CI but not rescue (Shropshire et al., 2020). In contrast, mutating the conserved residues of the sterile-like transcription factor (STE) domain of CifA had no impact on CI (Shropshire et al., 2020), establishing the hypothesis that these domains and residues may possess a molecular function relevant to CI. Moreover, a nuclear localization sequence is necessary for CifA-mediated CI, rescue and maintaining sperm genome integrity (Kaur et al., 2022), thus motivating interest to understand CifA's enzymatic roles in nucleotide metabolism.

Here, we performed *in vitro* and *in situ* biochemical assays with wild type, transgenic, and mutant proteins to determine the nuclease capabilities of T1 *wMel* CifA and CifB proteins. We demonstrate that (i) both CifA and CifB are *in vitro* nucleases, where CifA cleaves single-stranded (ss) DNA, double-stranded (ds) DNA and ssRNA substrates, and CifB cleaves both DNA substrates, (ii) Evolutionary-guided substitutions and truncations ablate nuclease functions, (iii) *wMel Wolbachia* and transgenic *cif* expression causes *in situ* spermatid DNA fragmentation in *D. melanogaster*. We discuss the importance of these findings to the mechanism of the Cif proteins in inducing CI.

Results

Type 1 CifB protein is an *in vitro* DNase.

To test nuclease activity of T1 *wMel* CifB, we first generated CifB recombinant proteins for *in vitro* enzyme assays against DNA and RNA oligonucleotide substrates. The full CifB protein is too large to recombinantly express in *E. coli* (Wang et al., 2022; Xiao et al., 2021). As such, we generated proteins spanning the unannotated amino terminus (A), the N-terminal nuclease domain (NTND) and C-terminal nuclease domain (CTND) (labeled as CifB_AD, 1-796 aa), both nuclease domains only (CifB_{A&D}, 277-796 aa) (**Fig. 1A**), and A-terminus alone (CifB_{A&C&D}, 1-276 aa) (**Fig. S1**). We previously demonstrated that mutations in highly conserved residues across CifB, including both nuclease domains (labeled CifB₂ and CifB₃), ablate transgenic CI (Shropshire et al., 2020), suggesting these domains are crucial for the CI mechanism. Therefore, we generated CifB recombinants with the same substitutions in the NTND (CifB_{2;A}D) and CTND (CifB_{3;A}D) (**Fig. 1A**) to test if these same mutants ablate nuclease activity *in vitro*. We report four key outcomes:

First, CifB_AD with the amino terminus and both nuclease domains cleaved ssDNA and dsDNA substrates similar to the control DNase enzyme (**Fig. 1B**). As expected, ethylenediaminetetraacetic acid (EDTA) ablated the nuclease activity since it chelates divalent cations required for nuclease activity. We, therefore, conclude that T1 CifB_AD (referred to as CifB hereafter) encodes DNase enzymatic function, which is contrary to previous claims that tested recombinant protein with

presumptive mutations thought to recapitulate catalytic activity (Beckmann et al., 2017; Chen et al., 2019).

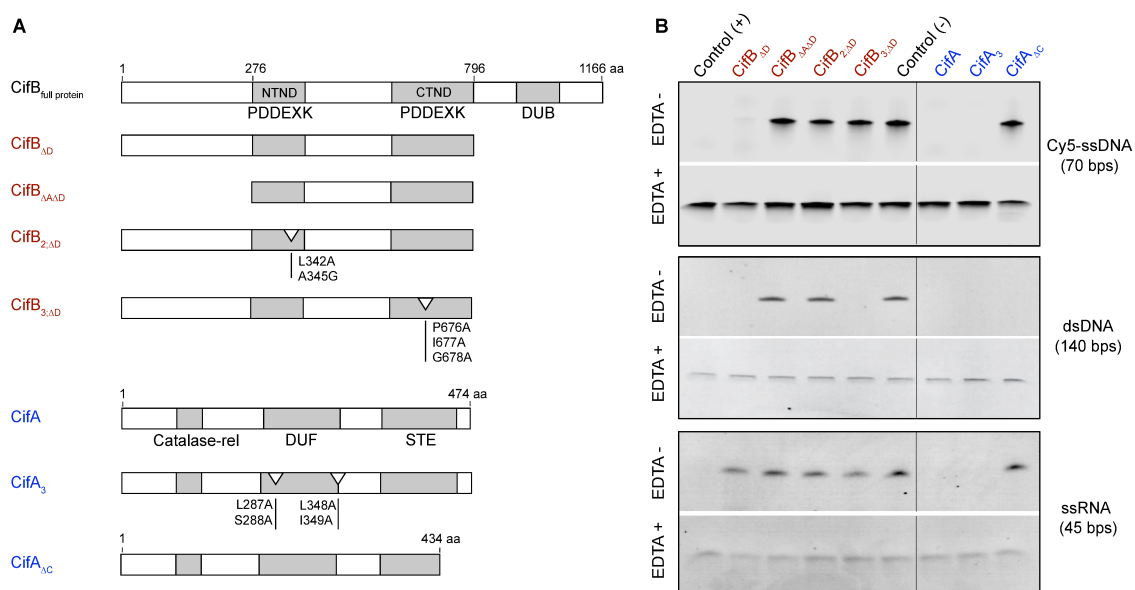


Figure 1. CifA and CifB proteins are *in vitro* nucleases in which CifA is a DNase and RNase, and CifB is a DNase. (A) Schematic representation showing the Cif recombinant proteins used in the study. Since the full CifB protein is too large to be recombinantly expressed in *E. coli*, purified His-GST tagged CifB variants were generated with a mixture of the amino terminus, N-terminal nuclease domain (NTND), and C-terminal nuclease domain (CTND) along with various engineered substitutions in conserved residues. The full His-tagged CifA protein was generated with and without substitutions in conserved residues and a truncation in the C-terminus. (B) Nuclease activities of the Cif proteins against DNA and RNA substrates. CifB shows both ss- and ds-DNase activity. While ss-DNase activity is ablated by mutations in both NTND (CifB_{2AD}) and CTND (CifB_{3AD}), ds-DNase activity is ablated by substitutions in NTND but not in CTND. CifA is both a DNase and RNase. Substitutions in conserved residues of the DUF domain (CifA₃) do not alter CifA nuclease activity, while a truncation at the C-terminus (CifA_{ΔC}) ablates ss-DNase and RNase action of CifA. DNase and RNase enzymes were used as positive controls, and no CifA or CifB proteins were added to the reaction mixtures for negative controls. EDTA was added to a 20x molar excess over Mg²⁺ to inhibit the reactions. Cy5-labeled ssDNA samples were run in a 10% polyacrylamide/TBE gel. Non-labeled dsDNA and ssRNA samples were run in 10% polyacrylamide/TBE gel and stained with GelRed. Assays were replicated three times independently. Full uncropped gel images are submitted in source data files. Related reagents can be available upon request from Bordenstein lab.

Second, the truncated CifB without the amino terminus (CifB_{ΔAΔD}) failed to cleave ssDNA or dsDNA, suggesting the amino terminus plays a crucial role in nuclease activity. The amino terminus, for instance, may affect proper structural folding of the protein or binding to the target substrate. The amino terminus alone (CifB_{ΔNΔCΔD}) also did not exhibit cleaving activity (**Fig. S1**), reinforcing the importance

of the intact amino terminus together with both nuclease domains. Intact CifB's DNase activity was substantiated once more against a ssDNA substrate using a range of protein concentrations and reaction incubation periods in which enzymatic activity decreases with lower enzyme concentration and shorter incubation time (**Fig. S2**), as expected.

Third, mutant CifB proteins with substitutions in each nuclease domain (CifB_{2;ΔD} and CifB_{3;ΔD}) that ablate transgenic CI in flies (Shropshire et al., 2020) also abolished ssDNA cleavage here *in vitro* (**Fig. 1B**). Findings suggest that substitutions in these negatively-selected residues (Shropshire et al., 2020) are important for the ss-DNase activity of CifB. Moreover, incubation of these proteins with dsDNA revealed that CifB_{2;ΔD} ablates ds-DNase activity, whereas CifB_{3;ΔD} did not.

Fourth, we investigated if CifB is a RNase and found that CifB variants did not cleave ssRNA oligonucleotide substrates (**Fig. 1B and SI**). We also performed mass spectrometry of the purified CifB protein to rule out the alternative explanation that co-purified *E. coli* nucleases with the CifB protein confound the CifB nuclease activity observed. We detected no contaminating DNase enzymes (**Table S1**).

Taken together, CifB is a DNase, and the CifB₂ sites in the NTND are crucial for both ssDNA and dsDNA cleavage, whereas CifB₃ sites in CTND are substrate-specific as the mutated residues only ablate cleavage of ssDNA. These data are semi-consistent with crystal structures of Cif proteins (Xiao et al., 2021) that showed the NTND of T4 CifB from *wPip* strain is the main catalytic center, and the CTND may be involved in specific substrate binding.

Type 1 CifA protein is an *in vitro* DNase and RNase.

Purified full-length CifA was incubated with the same DNA and RNA substrates as above. We report that CifA cleaved the ssDNA, dsDNA, and ssRNA substrates in the absence of EDTA (**Fig. 1B, Fig. S2**). Notably, CifA's nuclease activity is never reported before. Mass spectrometry of the CifA protein once again indicated there was no co-purified DNase from *E. coli* that could alternatively explain CifA's DNase function (**Table S1**). However, there was a very low abundance of ribonuclease peptides from *E. coli* in the purified CifA construct

(**Table S1**). To test if the *E. coli* RNase contaminant contributed to CifA's RNase activity, we performed nuclease activity assays by incubating RNA substrate with various concentrations of CifA protein diluted up to 16-fold to eliminate the rare contaminant. CifA RNase activity persisted upto 8-fold enzyme dilutions (**Fig. S2A**). Moreover, the same low abundant *E. coli* RNase contaminant was also detected in the purified CifB sample (**Table S1**) that notably lacked RNase activity. These results support the RNase properties of CifA and that the minute traces of randomly distributed *E. coli* contaminants are irrelevant. CifA's RNase function is consistent with predictive CifA annotations for RNA binding and transcription.

CifA has low sequence similarity to a domain of unknown function (DUF3243) with structural homology to a RNA-binding domain (Lindsey et al., 2018). We previously showed that mutating conserved residues in the T1 CifA DUF (labeled CifA₃) ablated CI but not rescue in transgenic flies (Shropshire et al., 2020). Thus, we tested here whether CifA₃ (**Fig. 1A**) cleaves DNA and/or RNA. CifA₃ retained nuclease activity (**Fig. 1B**), suggesting that the conserved residues in the DUF are crucial to CI but do not contribute to CifA's *in vitro* nuclease activity. In contrast, CifA_C protein with a 40 amino acid truncation inclusive of the predicted STE transcription factor at the C-terminus ablated ss-DNase and ss-RNase function, but dsDNA cleavage remained (**Fig. 1B**). These findings suggest that the C-terminus of CifA contributed to substrate specificity for ssDNA and ssRNA. Notably, this region is proposed to be disordered based on the crystal structure of CifA (Wang et al., 2022; Xiao et al., 2021). Disordered regions can regulate ssDNA binding in *Escherichia coli* (Kozlov et al., 2015), facilitate post-translational modifications, and often regulate key protein functions (Sharma and Schiller, 2019).

In *D. melanogaster*, dual expression of wMel CifA with CifB is required to induce CI (Beckmann et al., 2017; Kaur et al., 2022; LePage et al., 2017; Shropshire and Bordenstein, 2019). To examine if T1 wMel CifA and CifB have any impact on their independent nuclease activities, we co-incubated the proteins with dsDNA or ssRNA substrates as described earlier (Chen et al., 2019). We found that DNA and RNA cleavage continued upon co-incubation of CifA variants with different active and inactive CifB variants (**Fig. 2**).

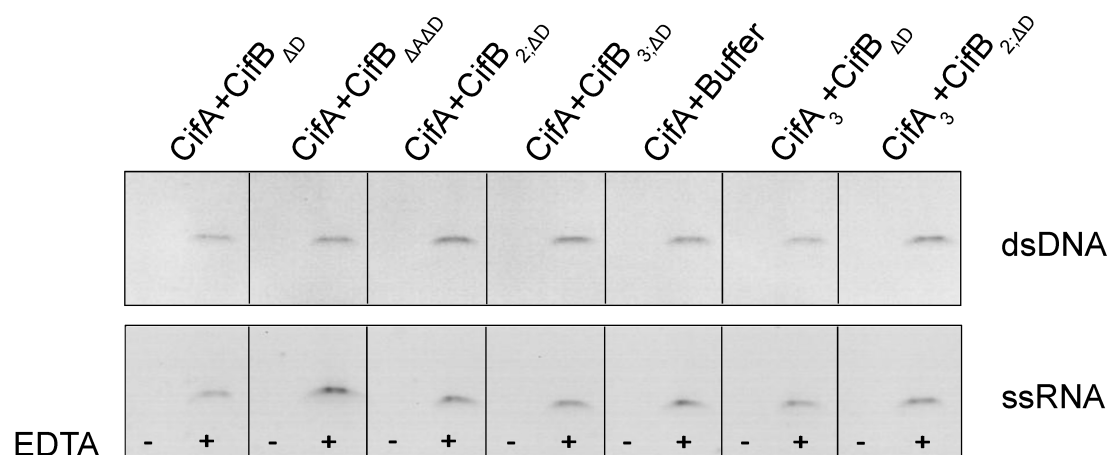


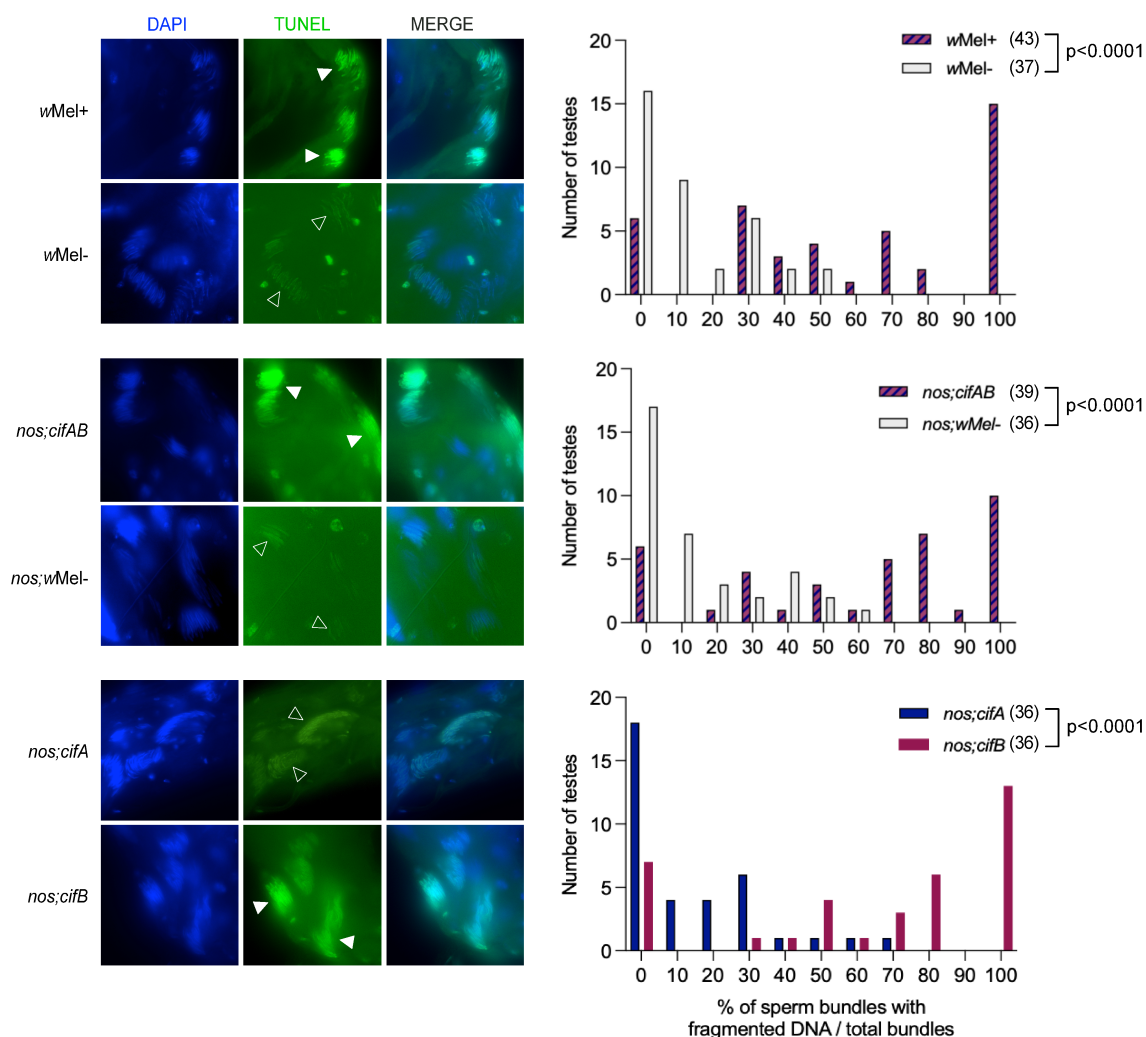
Figure 2. Nuclease activity persists when CifA and CifB variants are co-incubated. Prior to adding enzymes to the dsDNA and ssRNA substrates, 10 μ M of CifA and 1 μ M CifB variants were co-incubated on ice for 30 min to allow complex formation. To serve as a negative control, CifA was also co-incubated with the storage buffer only, in which CifB purified proteins were stored. The reactions were carried out for 120 min. EDTA was added to a 20x molar excess over Mg^{2+} to inhibit the reactions. Samples were run in 10% polyacrylamide/TBE gel and stained with GelRed. Full uncropped gel images are submitted in source data files. Related reagents can be available upon request from Bordenstein lab.

wMel *Wolbachia* and transgenic Cif expression promotes *in situ* spermatid DNA damage

We next sought to evaluate if *in vitro* Cif nuclease activity translates to *in situ* activity by dissecting whole testes from <8 hrs old wild type (wMel+ and wMel-) and transgene-expressing *Drosophila* males. We performed TUNEL (terminal deoxynucleotidyl transferase dUTP nick-end labeling) assays to examine DNA breaks in spermatid nuclei isolated from testes squashes. DNA breaks and repair are typical features accompanying the histone-to-protamine transition during the canoe stage of spermiogenesis to tightly pack the chromatin in the sperm nucleus (Rathke et al., 2007). However, excessive DNA breaks cause sperm chromatin abnormalities and adversely impact male fertility status in mammals (Agarwal and Said, 2003). Therefore, we investigated CifA- and CifB-induced spermatid DNA damage with TUNEL signals upon transgenic and wild-type expression of the Cifs relative to negative controls.

TUNEL signals occurred in the elongating spermatids at the canoe stage and were no longer detected in the individualized and mature sperms in all treatment groups, as expected (Rathke et al., 2007) (**Fig. S3**). In both CI-inducing wMel and

292



293

294

295 **Figure 3. *wMel* Wolbachia and transgenic Cif expression cause *in situ* DNase activity in**

296 **canoe-stage spermatids.** TUNEL staining on testes squashes from <8 hrs old males was

297 performed to visualize and quantify sperm bundles with DNA breaks across the genotypes tested.

298 Compared to the negative controls *wMel-* (empty arrow heads), DNA breaks marked by TUNEL

299 staining (green) and DNA (blue) are abundant in the wild type *wMel* Wolbachia and dually

300 expressed *cifAB* transgenic lines at the canoe stage of spermiogenesis (solid arrow heads). Upon

301 individual expression, CifB induces DNA damage, while CifA does not. Total sperm bundles and

302 those with fragmented DNA were manually counted. The numbers of testes investigated are shown

303 in parentheses next to the genotype. The experiment was performed with two independent

304 biological replicates and samples were blind-coded for the first run. P-value significance was

305 calculated by Mann-Whitney pairwise comparison test. P-values are reported in Table S2. CI-based

306 hatch rate assay was conducted in parallel as shown in Figure S4.

307

308 *cifAB*-expressing males, there were significantly higher percentages of sperm

309 bundles with strong TUNEL intensity relative to their aposymbiotic controls (**Fig. 3;**

310 Mann-Whitney test, $p < 0.0001$, **Table S2**), indicating the Cifs caused the increased

DNase activity. CifB initiated significantly higher DNA damage in single expressing lines than CifA and negative control lines (**Fig. 3**; Kruskal-Wallis test, $p < 0.0001$, **Table S2**), suggesting CifB with and without the DUB domain can cleave DNA *in situ* and *in vitro*, respectively. Notably, transgenic expression of individual CifB from wMel does not cause CI in *D. melanogaster* (**Fig. S4, Table S3**) (LePage et al., 2017; Shropshire and Bordenstein, 2019). CifB-mediated *in situ* DNase activity alone is therefore insufficient to induce CI. CifA did not cause *in situ* DNase activity, which contrasted with the *in vitro* data above. We hypothesize that CifA's ability to cleave spermatid nuclear DNA may depend on the co-expression of CifB, or CifA's RNase activity in the testes (not observable by TUNEL here) may be relevant to CI (**Fig. 5**).

Most, but not all, of the conserved residues that were mutated in CifA and CifB ablated or decreased CI strength in flies (Shropshire et al., 2020). To check if these same variants contribute to *in situ* DNA damage of the developing spermatids, we first performed TUNEL assays using transgenic expression of CifA mutants alone or together with intact CifB in the testes. Similar to individual CifA mutant expressing lines, there were very few or no sperm bundles with DNA damage in CifA mutant lines upon the individual (**Fig. S5A**) or dual expression with CifB (**Fig. S5D**; $p > 0.05$, Kruskal-Wallis test, **Table S2**). This result indicated that CifA protein mutants inhibit the *in situ* DNA cleavage activity mediated by CifB because dual expression of intact CifAB cleaved spermatid DNA at the canoe stage (**Fig. 3**). Inhibition of such activity could occur because of mutated CifA residues hampering proper CifB binding, which may impact downstream spermatid DNA cleavage. For example, the CifA₃ mutant residue is near the binding interface for CifA and CifB (Xiao et al., 2021). Interestingly, despite mutated residues in CifA₄B (with substituted residues in the predicted STE domain) that did not apparently cleave spermatid DNA *in situ* (**Fig. S5D, S6**), this mutant caused strong transgenic CI (**Fig. S5F**), suggesting the CifA₄ mutant sites are important for DNase activity in the canoe-stage spermatids but not CI. Similar to the CifA mutant findings, single expression of CifB mutants and dual expression of CifB mutants with wild type CifA generally ablated spermatid DNase activity (**Fig. S5B; S5E** $p < 0.0001$, Kruskal-Wallis test, **Table S2**) as well as CI (**Fig. S5C, S5F, Table S3**). Overall, amino acid substitutions throughout the length of the T1 CifA and CifB

proteins ablated CI and *in situ* DNase function at the canoe stage of spermiogenesis.

QxxxY motif associates with CifA and CifB nuclease activity.

T1 CifB lacks the canonical PDDEXK catalytic residues (Beckmann et al., 2017; LePage et al., 2017; Lindsey et al., 2018; Wang et al., 2022; Xiao et al., 2021), and CifA lacks any similarity to PDDEXK nuclease domains. As such, the nuclease data above lead to a hypothesis that there may be an alternative catalytic site for nuclease activity of the wMel Cif proteins. Either the PDDEXK-like domain in CifB is active, or the Cif proteins encode alternative catalytic sites or motifs to cleave DNA and RNA. Both explanations are equally viable, though several mutants within and outside the PDDEXK domains can ablate *in vitro* and *in situ* DNase function (**Fig. 1B**, **Fig. S1**, **Fig. S5**). Interestingly, T1 CifB of wMel possesses a QxxxY motif within a region of predicted α -helices at the CTND (**Fig. 4A**). The QxxxY motif is characteristic of RecB-family nucleases with HsdR subunits of *Escherichia coli* that also encode a canonical PDDEXK domain (Singleton et al., 2004). Mutating Q and Y residues impairs EcoR124I DNA cleavage, possibly by stabilizing the catalytic pocket of the PDDEXK domain or modulating the binding efficacy of this domain to the DNA (Šišáková et al., 2008).

Notably, the QxxxY motif is also present in T4 CifB of wPip characterized as a DNase, present in T3 CifB of wNo (Adams et al., 2021; Sun et al., 2021) and missing in T1 CifB of wPip previously deemed to lack nuclease function (Chen et al., 2019). Thus, it is tentatively possible that the QxxxY motif may assist CifB DNase activity since it is localized in CTND, and the sites in CTND are proposed to be involved in substrate binding or catalytic regulation (Xiao et al., 2021). Notably, CifA also possesses the QxxxY motif in one of its predicted α -helices (**Fig. 4B**), emphasizing the potential association of this motif in mediating nuclease activity. While this motif's presence and predicted structure correlated with the *in vitro* DNase activity data of CifA and CifB homologs tested thus far, it is not possible to prescribe relative importance over the PDDEXK-like motifs since various mutants beyond these sites ablated DNase function.

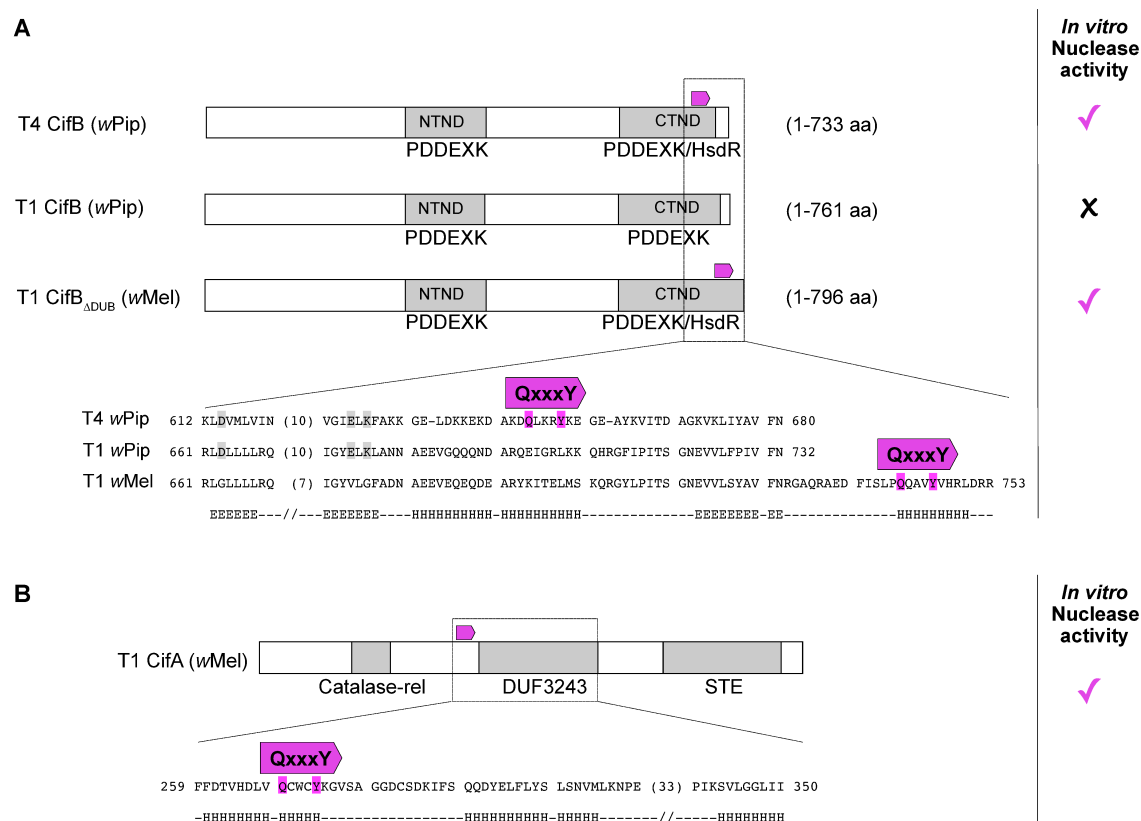


Figure 4. Nuclease activity of Type 4 CifB and Type 1 CifA and CifB associates with a QxxxY motif. (A) Schematic of CifB proteins and amino acid sequence showing the CTND location of QxxxY motifs that belong to the HsdR subunit of RecB nuclease family (Singleton et al., 2004). The motif and glutamine (Q) and tyrosine (Y) residues are highlighted in magenta, whereas the previously characterized aspartate (D), glutamate (E) and lysine (K) residues belonging to the PDDEXK nuclease family (Chen et al., 2019) are highlighted in grey. The QxxxY motif is present in wPip Type 4 (T4) CifB known as an *in vitro* nuclease (Chen et al., 2019) and T1 CifB from wMel *Wolbachia* characterized as an *in vitro* nuclease (this study). The QxxxY motif is absent in the non-nuclease wPip T1 CifB. (B) T1 CifA protein from wMel *Wolbachia* shown to be an *in vitro* nuclease (this study) also contains a QxxxY motif in a predicted α -helical region. Predicted α -helical residues are labeled “H” and residues predicted to be part of β -sheets are labeled “E.” The numbers of excluded residues are shown in parentheses. The last residue numbers are shown at the end of each sequence.

Discussion

CifA and CifB proteins are both required for CI induction in *D. melanogaster* (LePage et al., 2017; Shropshire et al., 2018; Shropshire and Bordenstein, 2019) induced by the wMel strain, the same strain released in *Aedes aegypti* field trials by the World Mosquito Program (Bourtzis et al., 2014; O'Neill et al., 2018; Utarini et al., 2021). At the cellular level, both CifA and CifB target the nuclei of developing

sperms in flies; and CifA harbors a nuclear localization sequence that is functionally required for nuclear entry in developing fly sperm cells and the recapitulation of CI and rescue (Kaur et al., 2022). Based on these premises, we hypothesized that the two Cif proteins metabolize nucleic acids. We thus set out to evaluate their enzymatic properties using *in vitro* and *in situ* assays. We describe five key results: (i) T1 CifB is a DNase, (ii) T1 CifA is both a DNase and RNase, (iii) *Wolbachia* and transgenic CifAB expression causes DNA damage at the canoe stage of spermiogenesis, but *in situ* DNase activity of CifB alone is insufficient to cause CI, (iv) truncations and conserved site-directed mutagenesis establish the dependency of Cif nuclease functions across the length of the proteins, and (v) a QxxxY motif in both CifA and CifB associates with observed nuclease activity. Notably, this is the first report to demonstrate nuclease functions of the CifA protein. Below, we discuss the relevance of Cif nuclease findings to the molecular bases of CI, the Host Modification model, and *cif* gene nomenclature.

Previous work showed that the T1 and T4 phylogenetic clades of CifB biochemically act as *in vitro* deubiquitinases (DUBs) and DNases, respectively. These findings led to a proposal of an enzyme-dependent gene nomenclature and mechanistic narrative of CI (Beckmann et al., 2017; Chen et al., 2019). Structural homology-based annotations reveal that DUB-encoding T1 CifBs also contain PDDEXK-like nuclease dimers (Beckmann et al., 2017; LePage and Bordenstein, 2013; Lindsey et al., 2018) - however, due to the absence of canonical catalytic residues, T1 CifBs were deemed non-nucleases (Chen et al., 2019). Notably, PDDEXK nuclease family enzymes vary substantially in sequence and structure while retaining DNase activity (Knizewski et al., 2007; Steczkiewicz et al., 2012). Indeed, CifB nuclease domains across all CI Types show homology to not only the PDDEXK family but also to the restriction endonucleases HsdR (Lindsey et al., 2018). HsdR is a subunit of type I restriction-modification enzymes (Obarska-Kosinska et al., 2008) that carries a unique and conserved QxxxY motif sequence located in the α -helix immediately C-terminal to the PDDEXK motifs (Šišáková et al., 2008). Mutated Q and Y residues hamper the DNA cleaving efficiency in the EcoR124I restriction enzyme, indicating their contribution to DNase activity (Šišáková et al., 2008). Interestingly, both CifA and CifB of wMel possess the QxxxY motif in a predicted α -helix region, and in CifB, they are present in the C-

terminus preceded by the PDDEXK motif. Thus, future CI studies may consider association of the QxxxY motif with DNase activity by the CifA and CifB proteins.

In situ assays at the canoe stage of development notably revealed spermatid DNA damage in both *Wolbachia* and dual CifAB expressing lines, thus demonstrating the Cif enzymes are responsible for metabolizing DNA. DNA damage and repair are typical features facilitating the histone-to-protamine transition during the canoe stage of spermiogenesis to tightly pack the sperm chromatin (Rathke et al., 2007). However, excessive DNA breaks result in abnormal sperm chromatin packaging that can impact male fertility and embryonic viability in mammals (Agarwal and Said, 2003; Hosen et al., 2015; Sakkas and Alvarez, 2010). *Wolbachia*-induced oxidative DNA damage was previously reported in an early spermatogenesis stage of *D. simulans* (Brennan et al., 2012) suggesting that modifications to developing sperm contribute to CI. Therefore, detection of Cif-mediated DNA cleavage of spermatids is noteworthy here as DNA damage can adversely impact Histone modifications that hampers the histone removal process from chromatin (Bannister and Kouzarides, 2011; Tjeertes et al., 2009). Indeed, the nuclear-targeting CifA and CifB proteins alter sperm chromatin integrity by increased histone retention and decreased protamine deposition during sperm development (Kaur et al., 2022). Moreover, mutant males with protamine knockouts enhance wild type *Wolbachia* CI levels (Kaur et al., 2022). Taken together, our data suggests that Cifs promote spermatid DNA damage to alter sperm genome integrity and paternally bestow a long lasting modification which underpins CI.

While dual expression of intact CifA and CifB promotes DNA damage, CifA mutants expressed with intact CifB inhibit spermatid DNA cleavage. These findings suggest nuclear-targeting CifA is central to metabolizing spermatid nucleotides, and mutations and deletions in CifA may hinder the proper binding and/or localization of CifA and CifB (Horard et al., 2022; Kaur et al., 2022; Wang et al., 2022) in testes so that they cannot access target substrates to mediate cleavage. Ablation of spermatid DNase action by dual expression of CifB and CifA₄ (with substituted residues in the predicted Sterile transcription factor domain (STE)) is particularly notable as it induces CI. CifA's STE domain is highly conserved across the phylogenetic Types (Lindsey et al., 2018). Crystal structure of CifA predicts this region to be disordered (Wang et al., 2022; Xiao et al., 2021). Disordered regions

may facilitate post-translational modifications and regulate key protein functions (Sharma and Schiller, 2019), in this case, DNase activity that is independent of CI.

Upon individual Cif expression, we observed that CifA alone does not recapitulate *in situ* DNase activity despite its ability to function as a DNase *in vitro*. At the histone-to-protamine transition stage of spermiogenesis, wMel CifA is not highly abundant, and wPip CifA reportedly vanishes (Horard et al., 2022; Kaur et al., 2022). Thus, one explanation is that CifA levels are below the threshold to induce spermatid DNA nicks or it requires co-expression with the CifB binding partner (Wang et al., 2022; Xiao et al., 2021). In contrast, single CifB promotes DNA fragmentation at this stage. T4 wPip CifB is also an *in vitro* DNase, which is proposed to contribute to the CI mechanism independent of CifA via DNA cleavage (Chen et al., 2019), although the host DNA substrate that T4 CifB acts on remains unknown.

In conclusion, we report previously-unrecognized nuclease functions of both Cif proteins from wMel *Wolbachia* and establish a new understanding that T1 CifA is a DNase and RNase, and both T1 and T4 CifBs are DNase enzymes, which is consistent with general homology between their core nuclease domains. Based on these findings, we suggest that a durable gene nomenclature for Cifs will be grounded in evolutionary genetics and stable phylogenetic Types (T1-T5) that are less subject to shifting discoveries based on Cif enzymatic functions, as revealed here. Finally, we link *in vitro* and *in situ* enzymatic activity of the Cif proteins and highlight their impact on the nuclei of developing host sperm, a finding that is consistent with the Host Modification model of CI in this system (Kaur et al., 2022). This work highlights the significance of replicating *in vitro* enzymatic assays with *in situ* and/or *in vivo* measurements to confirm the biological functions of proteins in the arthropod host systems and their relation to the CI mechanism.

Ideas and Speculation

Given that individual expression of wMel CifA and CifB is insufficient to establish CI (LePage et al., 2017) and modify sperm genome integrity (Kaur et al., 2022), we propose that Cif-mediated spermatid DNA and/or RNA damage may be central to the incipient CI-defining sperm modifications that alter chromatin integrity (**Fig. 5**). For example, CifA RNase activity may regulate *cifB* transcripts and thus gene expression that is known to be lower than that of *cifA* (Beckmann and Fallon,

2013; LePage et al., 2017; Lindsey et al., 2018), and/or CifA RNase may degrade host mRNAs. In flies, RNA depletion in the spermatocytes can lead to sperm chromatin abnormalities and infertility (Mills et al., 2019). In addition, CifB's DNase activity may work alone or together with CifA's RNase or DNase functions to alter spermatid chromatin. Each of these hypotheses can impact CI (**Fig. 5**) and hypothetically, it is fair to speculate that divergent Cif homologs and host genetic backgrounds may impact the outcomes. Thus, some *Wolbachia* strains may cause CI by variations on a theme around the Cif proteins. The relative importance of CifA's RNase activity to CI will be significant for future research.

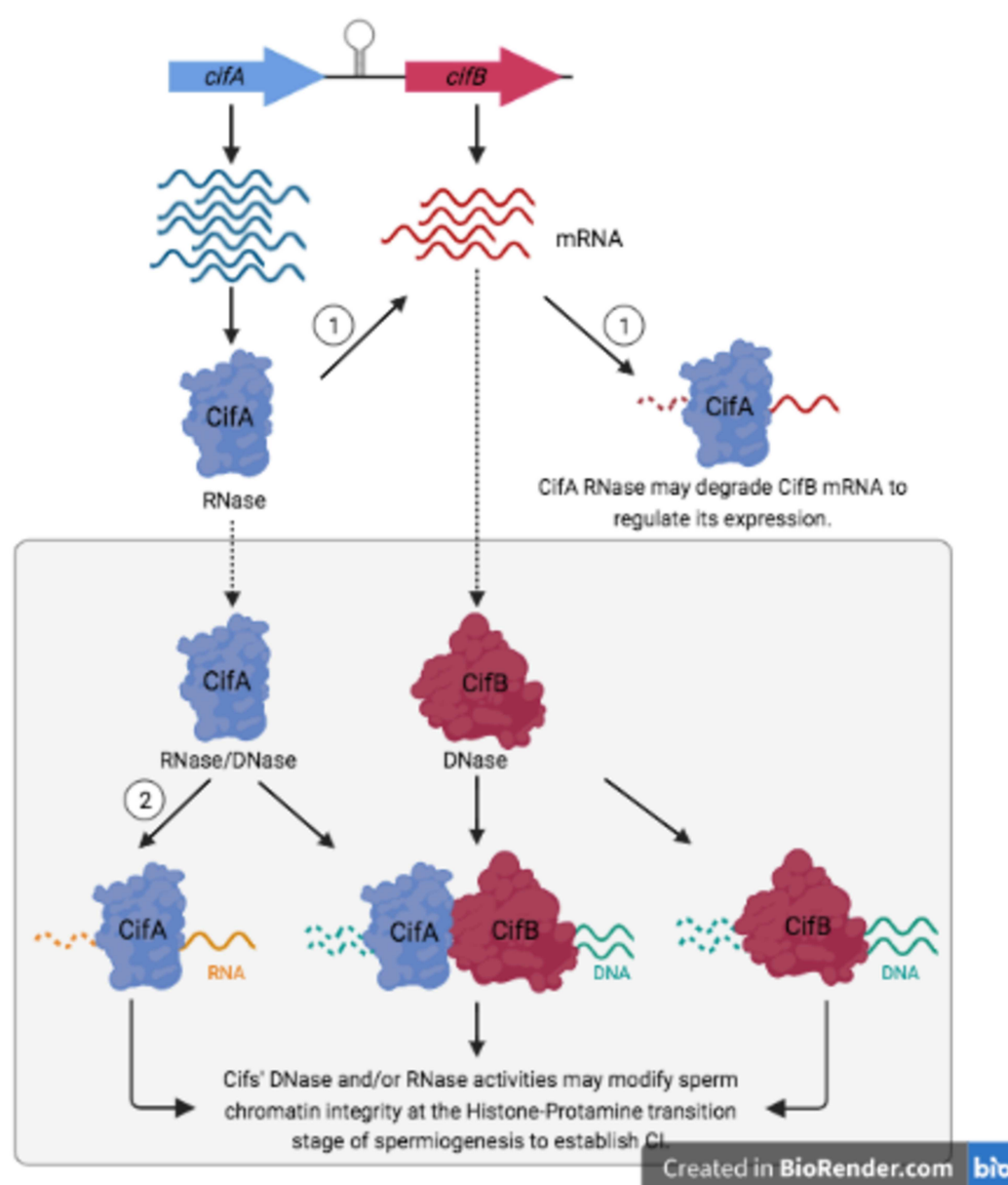


Figure 5. Illustrative summary and hypotheses on CifA and CifB nuclease functions under the Host Modification model of CI. *cifA* and *cifB* genes are shown in blue and red arrows, respectively, with an annotated hairpin terminator. *cifA* is highly transcribed compared to *cifB* (Kaur et al., 2022). CifA enzyme encodes *in vitro* DNase and RNase activities. 1. Under the type V toxin-antitoxin system (52), the CifA RNase may degrade mRNA of the *cifB* gene and suppress its toxicity against the host. 2. Under the Host Modification model of CI, CifA RNase may target spermatogenesis-associated mRNA(s) to modify sperm genome integrity and induce CI. CifB's DNase activity alone or combined with CifA's DNase and/or RNase functions may induce DNA fragmentation in canoe spermatids, the stage when histones are replaced by protamines to generate tightly condensed chromatin (Rathke et al., 2014). CifB's DNase action is necessary for CI in some systems (30), though the DNA substrate in the host remains uncharacterized. On the other hand, we conversely show here that CifB spermatid DNase activity from *wMel* *Wolbachia* is insufficient to cause CI, which may be contingent upon CifA's RNase and/or DNase activity that could hypothetically modify sperm genome integrity. CifAB-induced sperm modifications at this stage may pave the incipient steps to establish CI.

Materials and Methods

1. Fly rearing and strains

D. melanogaster stocks *y¹w** (BDSC #1495), *nos*-GAL4:VP16 (BDSC #4937), and UAS transgenic (TG) lines homozygous for *cifA*, *cifB*, *cifA;B*, and *cif* mutants (LePage et al., 2017; Shropshire et al., 2020) were maintained at 12:12 light:dark at 25°C and 70% relative humidity on 50 ml of standard cornmeal- and molasses-based food medium. Lines without *Wolbachia* were previously generated (LePage et al., 2017) through tetracycline treatment for three generations. *Wolb_F* and *Wolb_R3* primers were used to confirm symbiont presence. Virgin flies were collected and stored at room temperature.

2. In vitro Nuclease assay

Codon optimization, gene synthesis, cloning, and protein expression/purification were outsourced to GenScript Biotech (New Jersey, USA). Briefly, *cifA* and *cifB* genes were codon-optimized for translation and expression in *E. coli*, *de novo* synthesized, cloned into pGS-21a expression vector, and transformed into *E. coli* BL21 (DE3) competent cells. Cultures were grown in 1 L TB medium containing ampicillin and were incubated in 37°C at 200rpm. Once cell density reached to 1.2 O.D. at 600 nm, 0.5 mM isopropyl-b-D-thiogalactoside (IPTG) was introduced for induction at 15°C for 16 hours and then centrifuged at 8000 *g* for 20 min.

For CifA, cell pellets were lysed on ice using buffer A (50 mM Tris, 150 mM NaCl, 1 mM TECP, pH 8.0) by sonication. The lysate was clarified by centrifugation at 12000 g for 30 min at 4°C. The clarified extracts of target CifA proteins carrying His tags at the C-terminus were filtered with a 0.45 µm filter and loaded onto Ni-nitrilotriacetic acid (NTA) agarose resin preequilibrated in buffer A. The column was then washed with buffer B (50 mM Tris, 150 mM NaCl, pH 8.0) 10 column volumes (CV) to wash away impurities from the target proteins. The protein was eluted using buffer B supplemented with 20/50/300 mM imidazole. All proteins were sterilized by passing through 0.22 µm filter before being stored in aliquots. The concentration was determined by BCATM protein assay with BSA as a standard. SDS-PAGE and Western blot were used to confirm protein purity and molecular weight (**source data files**).

For CifB, cell pellets were lysed on ice using buffer A (50 mM Tris, 150 mM NaCl, pH 8.0). The lysate was clarified by centrifugation at 12000 g for 30 min at 4°C. The inclusion body pellet was solubilized in denature buffer A (7M Gu-HCl, 50 mM Tris-HCl, 150 mM NaCl, pH 8.0) by sonication. The cell precipitate was spun down at 13,000 rpm for 30 min at 4°C, then the supernatant including target CifB proteins carrying His-GST tag at the N-terminus was filtered with a 0.45 µm filter and loaded onto NTA agarose resin preequilibrated in buffer A. The column was then washed with buffer B (8 M Urea, 50 mM Tris-HCl, pH 8.0) 10 CV to wash away impurities from the target proteins. The protein was eluted using buffer B supplemented with 20/50/300 mM imidazole. All proteins were sterilized by passing through 0.22 µm filter before being stored in aliquots. The concentration was determined by BCATM protein assay with BSA as a standard. SDS-PAGE and Western blot were used to confirm protein purity and molecular weight (**source data files**).

Additionally, mass spectrometry-based protein identification was performed to ensure no contaminant nuclease from the *E. coli* expression system was co-purified. Briefly, 20 µg of purified proteins (CifA, CifA_{ΔC}, CifB_{ΔD}, and CifB_{2;ΔD}) were prepared for analysis using Suspension trap technology (Casiraghi et al., 2001) using the manufacturer protocol. The resulting peptides were analyzed by a 70-minute data-dependent LC-MS/MS analysis. Briefly, peptides were auto sampled onto a 200 mm by 0.1 mm (Jupiter 3 micron, 300A), self-packed analytical column coupled directly to an LTQ (ThermoFisher) using a nanoelectrospray source and

resolved using an aqueous to organic gradient. A single full-scan mass spectrum followed by 5 data-dependent tandem mass spectra (MS/MS) was collected throughout the run and dynamic exclusion was enabled to minimize the acquisition of redundant spectra. The Resulting MS/MS spectra were searched via SEQUEST against a database containing the expressed proteins, an *E. coli* background proteome and reversed version for each of the entries. Identifications were filtered and collated at the protein level using Scaffold Proteome Software.

In vitro nuclease activity assays were performed as previously described (Chen et al., 2019). For DNase and RNase activity measures, 1 μM of individual Cif proteins were incubated in a reaction buffer containing 20 mM Hepes (pH 8.0), 5 mM MgCl₂, 2.5% sucrose, 150 mM NaCl, 0.001% Triton X-100, and 2 mM DTT with 500 nM single-stranded (ss) Cy5-labeled DNA [70-mer: Cy5-GCAATTCGATCGTTGACATCTCGCGTGCTCGGTCAATCGGCAGATGCGGAGTGAAGTTCCAACGTTTCGGC-3] as previously used (Chen et al., 2019); 15nM double-stranded (ds) 154bp PCR purified fragment of *rp49* gene (Shropshire et al., 2018); or 100nM of synthetic RNA [45-mer: 5' GGGUCAACGUGGGCAAAGAUGUCCUAGCAAGCCAGAAUUCGGCAG -3'] generated by Sigma. In reactions where CifA was co-present with CifB, 10 μM CifA was used as previously described (Chen et al., 2019). All reactions were carried out at 25°C for 120 min and quenched by adding EDTA to a final concentration of 100 mM unless otherwise noted. Samples were run in 10% TBE polyacrylamide urea gels at 180 V for 60 min. For reactions using Cy5-labeled ssDNA, gels were imaged on the Odyssey CLx imaging system. Unlabeled dsDNA and ssRNA sample gels were post-stained with GelRed (Biotium) stain and imaged with Alpha innotech imager.

3. *In situ* TUNEL assay

Terminal deoxynucleotidyl transferase-mediated dUTP nick-end labeling (TUNEL) assays are based on detecting single- and double-stranded DNA nicking and fragmentation, which are characteristic of apoptotic cells (Vasudevan and Ryoo, 2016). To perform the assay and detect sperm DNA fragmentation, we first set up the flies as previously described (Shropshire et al., 2020). Briefly, virginity-controlled wild type (*wMel*⁺ and *wMel*⁻) and TG (*nos*-Gal4:VP16) females were aged 9–11 days and mated with males (Layton et al., 2019). We used *nos*-

Gal4:VP16 line since it was previously shown to drive *cifA*; *cifB* expression sufficient to induce near-complete embryonic death (Shropshire and Bordenstein, 2019). We collected <8 hours old wild type and TG Gal4-UAS males as the young hatched males induce strong CI levels (Shropshire et al., 2021a; Yamada et al., 2007), anesthetized on ice to stop their movement, and dissected whole testes in ice-cold 1X PBS solution. Dissected tissues were treated with 2 mM dithiothreitol to stabilize cellular proteins for 45 min at room temperature, followed by fixation in 2% paraformaldehyde on ice for 15 min. After washing in 1X PBS for 2 min, samples were permeabilized in 0.1% TritonX-100 in sodium citrate (10 mg sodium citrate, 10 μ l Triton, 10 ml milliQ H₂O) for 2 min on ice. After washing in 1X PBS for 2 min, samples were incubated with 50 μ l mix of 5 μ l enzyme and 45 μ l labeling solution (TUNEL *In situ* Cell Death Detection Kit, TMR Red' from Roche) for 1.5 h at 37°C in a dark humid chamber. After washing in 1X PBS for 2 min, samples were finally incubated with 50 μ L of DAPI staining solution (0.2 μ g/ml), mounted on a glass slide, squashed with a coverslip, and stored overnight at 4°C. Imaging was performed using green fluorescence filter excited at 488 nm laser for TUNEL and blue at 359 nm for DAPI stain at 100x magnification in All-in-one Keyence BZ-X700 fluorescence microscope. Image exposure settings were kept constant throughout the treatment groups and images were analyzed using ImageJ software. The total number of sperm bundles and TUNEL-positive bundles with damaged DNA were manually counted per testes.

4. Hatch rate assays

Male siblings from TUNEL assays were used to measure CI hatch rate levels as previously described (LePage et al., 2017). Briefly, males and females were paired in 8 oz bottles affixed with a grape-juice agar plate smeared with yeast. Bottles were incubated at 25°C for 24 h at which time the plates were replaced with freshly smeared plates and again stored for 24 h. Plates were then removed from bottles, and the numbers of eggs on each plate were counted. Any crosses with fewer than 25 eggs laid were discarded from the count. After another 30 h incubation at 25°C, the remaining unhatched eggs were counted. The percent of eggs hatched into larvae was calculated by dividing the number of hatched eggs by the total egg count and multiplying by 100.

5. *In silico* prediction analysis

Protein sequence alignment of CifB orthologs from different *Wolbachia* strains (NCBI accession number, CifB T1 wMel - WP_010962721.1, T1 wPip - WP_012481788.1, T4 wPip - WP_007302979.1) and CifA T1 wMel (WP_010962721.1) were performed using the MUSCLE plugin (Edgar, 2004) in Geneious Prime v2021.0.3 (Kearse et al., 2012). Secondary structure predictions of CifA and CifB protein sequences were made using the PSIPRED Protein Sequence Analysis Workbench program (Jones, 1999). We manually curated the presence of the QxxxY motif based on its localization within a region of predicted α -helices in the nuclease domains in CifB and throughout the length of CifA proteins.

6. Statistical analysis

All statistical analyses were performed using GraphPad Prism 9 software. While comparing *in situ* TUNEL data between two groups, we used a two-tailed, non-parametric Mann-Whitney *U*-test. For comparisons between more than two data sets, we used a non-parametric Kruskal-Wallis one-way analysis of variance test followed by a Dunn's multiple correction. This allowed robust testing between all data groups while correcting for multiple test bias. For CI hatch rate assays, statistical significance was determined by Kruskal-Wallis and Dunn's multiple correction tests. All P-values are reported in **Table S2 and S3** and raw data files related to each experiment are included in the source data files.

Acknowledgments

We thank members of the Bordenstein lab, especially, Sarah Bordenstein, Luis Mendez, and Luis Martinez for helpful discussions during the writing process. We thank Cell & Development Biology Equipment Resources at Vanderbilt core facility for their help in using Odyssey CLx imaging system. We also thank the Proteomics Core of the Mass Spectrometry Research Center at Vanderbilt University for performing protein identification service.

Author contributions

R.K., J.D.S., B.L., and S.R.B. designed the research; R.K. and B.L. performed research; R.K. and S.R.B. analyzed data; R.K. wrote original draft of the study. J.D.S., B.L. and S.R.B. edited and provided feedback on the paper. R.K. and S.R.B. wrote the final version of the paper.

Declaration of interest

S.R.B. and J.D.S. are authors on pending patents related to this work.

Supplementary figure titles and legends

Supplementary figure 1. CifB amino terminus alone does not induce DNase or RNase activity. Schematic representation showing the location of truncated domains of CifB recombinant proteins used in the study. 1 μ M of Cif protein was incubated with 500 nM Cy5-labeled ssDNA, 15nM dsDNA and 100nM ssRNA for 120 min. DNase and RNase enzymes were used as positive controls. To stop the reactions, EDTA was added to a 20x molar excess over Mg^{2+} . Cy5-labeled ssDNA samples were run in a 10% polyacrylamide/TBE gel. Non-labeled dsDNA and RNA samples were run in 10% polyacrylamide/TBE gel and stained with GelRed.

Supplementary figure 2. CifA and CifB are *in vitro* nucleases. (A) Dilution-based nuclease assay shows CifA's DNase and RNase, and CifB's DNase activities persist at higher protein concentrations and diminish with dilutions. (B) Time-course assay shows nuclease activity of Cifs diminishes at shorter incubation time points.

Supplementary figure 3. DNA breaks are detected at the canoe stage of *Drosophila* spermiogenesis. TUNEL staining on testes squashes of 0-8 hrs old males was performed to visualize DNA breaks during different stages of spermatogenesis in the same animal. DNA breaks marked by TUNEL staining are detectable only in canoe-stage spermatids, as expected (Rathke et al., 2007). During individualization, DNA breaks are no longer detectable in needle stage spermatids and the mature individualized sperms (white arrowheads).

Supplementary figure 4. Single transgenic expression of *cifB* does not induce CI in *D. melanogaster*. Hatch rate assays were conducted to test if wild type *Wolbachia*-carrying symbiotic males and transgenic, aposymbiotic males expressing single and dual *cifA* and *cifB* genes (unfilled sex symbols) induce CI. Each dot represents the percent of embryos that hatched from a single male and female pair. Letters to the right indicate significant differences based on $p = 0.05$ calculated by Kruskal-Wallis and Dunn's test for multiple comparisons between all groups. This experiment was conducted three times in parallel to the *in situ* TUNEL assay of Figure 3. P-values are reported in Table S3.

Supplementary figure 5. Single and dual expression of various *cifA* and *cifB* mutants ablates *in situ* nuclease activity and CI phenotype. (A, B, D, E) TUNEL assays on testes squashes of <8 hrs old males was performed to quantify sperm bundles with DNA breaks across the genotype treatment groups. Single and dual expression of *cifA* and *cifB* mutants in transgenic lines fails to induce spermatid DNA fragmentation. Total sperm bundles and those with fragmented DNA were manually counted from the images acquired. The numbers of testes investigated are shown in parentheses next to the genotype. The experiment was performed in two independent biological replicates in the same setup as

experiments in Figure 3. P-value significance was calculated by Kruskal-Wallis and Dunn's test for multiple comparisons between all groups. P-values are reported in Table S2. (C, F) Hatch rate assays were conducted to test if single and dual expression of *cifA* and *cifB* mutants can induce CI when transgenically expressed in aposymbiotic males (unfilled sex symbols). Each dot represents the percent of embryos that hatched from a single male and female pair. Letters to the right indicate significant differences based on $p = 0.05$ calculated by Kruskal-Wallis and Dunn's test for multiple comparisons between all groups. The experiment was conducted twice in parallel to the TUNEL assay. P-values are reported in Table S3.

Supplementary figure 6. Spermatid DNase activity is ablated upon dual expression of *cifA₄* and *cifB*. Image shows DNA nuclei (blue, DAPI) and TUNEL signal (green) in sperm bundles with less DNA breaks (empty arrow heads) compared to *cifAB* in Figure 3. The experiment was performed with two biological replicates in the same setup as that in Figure 3.

References

- Adams KL, Abernathy DG, Willett BC, Selland EK. 2021. Wolbachia cifB induces cytoplasmic incompatibility in the malaria mosquito. doi:10.1101/2021.04.20.440637
- Agarwal A, Said TM. 2003. Role of sperm chromatin abnormalities and DNA damage in male infertility. *Hum Reprod Update*. doi:10.1093/humupd/dmg027
- Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell Res* **21**:381–395. doi:10.1038/cr.2011.22
- Beckmann JF, Fallon AM. 2013. Detection of the Wolbachia protein WPIP0282 in mosquito spermathecae: implications for cytoplasmic incompatibility. *Insect Biochem Mol Biol* **43**:867–78. doi:10.1016/j.ibmb.2013.07.002
- Beckmann JF, Ronau JA, Hochstrasser M. 2017. A Wolbachia deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat Microbiol* **2**. doi:10.1038/nmicrobiol.2017.7
- Beckmann JF, Sharma GD, Mendez L, Chen H, Hochstrasser M. 2019. The wolbachia cytoplasmic incompatibility enzyme CIDB targets nuclear import and protamine-histone exchange factors. *Elife* **8**. doi:10.7554/eLife.50026
- Bourtzis K, Dobson SL, Xi Z, Rasgon JL, Calvitti M, Moreira LA, Bossin HC, Moretti R, Baton LA, Hughes GL, Mavingui P, Gilles JRL. 2014. Harnessing mosquito-Wolbachia symbiosis for vector and disease control. *Acta Trop* **132**. doi:10.1016/j.actatropica.2013.11.004
- Brennan LJ, Haukedal JA, Earle JC, Keddie B, Harris HL. 2012. Disruption of redox homeostasis leads to oxidative DNA damage in spermatocytes of Wolbachia-infected Drosophila simulans. *Insect Mol Biol*. doi:10.1111/j.1365-2583.2012.01155.x
- Caputo B, Moretti R, Manica M, Serini P, Lampazzi E, Bonanni M, Fabbri G, Pichler V, della Torre A, Calvitti M. 2020. A bacterium against the tiger: preliminary evidence of fertility reduction after release of Aedes albopictus males with manipulated Wolbachia infection in an Italian urban area. *Pest Manag Sci* **76**:1324–1332. doi:10.1002/ps.5643
- Caragata EP, Dutra HLC, Moreira LA. 2016. Exploiting Intimate Relationships: Controlling Mosquito-Transmitted Disease with Wolbachia. *Trends Parasitol* **32**:207–218. doi:10.1016/j.pt.2015.10.011
- Casiraghi M, Anderson TJ, Bandi C, Bazzocchi C, Genchi C. 2001. A phylogenetic analysis of filarial nematodes: comparison with the phylogeny of Wolbachia endosymbionts. *Parasitology* **122 Pt 1**:93–103. doi:10.1017/S0031182000007149
- Chen H, Ronau JA, Beckmann JF, Hochstrasser M. 2019. A Wolbachia nuclease and its binding partner provide a distinct mechanism for cytoplasmic incompatibility. *Proc Natl Acad Sci U S A* **116**:22314–22321. doi:10.1073/pnas.1914571116
- Crawford JE, Clarke DW, Criswell V, Desnoyer M, Cornel D, Deegan B, Gong K, Hopkins KC, Howell P, Hyde JS, Livni J, Behling C, Benza R, Chen W, Dobson KL, Eldershaw C, Greeley D, Han Y, Hughes B, Kakani E, Karbowski J, Kitchell A, Lee E, Lin T, Liu J, Lozano M, MacDonald W, Mains JW, Metlitz M, Mitchell SN, Moore D, Ohm JR, Parkes K, Porshnikoff A, Robuck C, Sheridan M, Sobecki R, Smith P, Stevenson J, Sullivan J, Wasson B, Weakley AM, Wilhelm M, Won J, Yasunaga A, Chan WC, Holeman J, Snoad N, Upson L, Zha T, Dobson SL, Mulligan FS, Massaro P, White BJ. 2020. Efficient production of male Wolbachia-infected Aedes aegypti mosquitoes enables

816 large-scale suppression of wild populations. *Nat Biotechnol* **38**:482–492.
817 doi:10.1038/s41587-020-0471-x

818 Dobson SL, Fox CW, Jiggins FM. 2002. The effect of Wolbachia-induced
819 cytoplasmic incompatibility on host population size in natural and manipulated
820 systems. *Proc Biol Sci* **269**:437–45. doi:10.1098/rspb.2001.1876

821 Edgar RC. 2004. MUSCLE: A multiple sequence alignment method with reduced
822 time and space complexity. *BMC Bioinformatics*. doi:10.1186/1471-2105-5-
823 113

824 Flores HA, O'Neill SL. 2018. Controlling vector-borne diseases by releasing
825 modified mosquitoes. *Nat Rev Microbiol*. doi:10.1038/s41579-018-0025-0

826 Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH. 2008.
827 How many species are infected with Wolbachia? - A statistical analysis of
828 current data. *FEMS Microbiol Lett* **281**:215–220. doi:10.1111/j.1574-
829 6968.2008.01110.x

830 Hoffmann AA, Montgomery BL, Popovici J, Iturbe-Ormaetxe I, Johnson PH, Muzzi
831 F, Greenfield M, Durkan M, Leong YS, Dong Y, Cook H, Axford J, Callahan a
832 G, Kenny N, Omodei C, McGraw E a, Ryan P a, Ritchie S a, Turelli M, O'Neill
833 SL. 2011. Successful establishment of *Wolbachia* in *Aedes* populations to
834 suppress dengue transmission. *Nature* **476**:454–7. doi:10.1038/nature10356

835 Hoffmann AA, Turelli M. 1997. Cytoplasmic incompatibility in insects. *Influ*
836 *Passengers Inherit Microorg Arthropod Reprod* **140**:42–80.

837 Hoffmann AA, Turellit M, Harshman LG. 1990. Factors Affecting the Distribution of
838 Cytoplasmic Incompatibility in.

839 Horard B, Terretaz K, Gosselin-Grenet A-S, Sobry H, Sicard M, Landmann F,
840 Loppin B. 2022. Paternal transmission of the Wolbachia CidB toxin underlies
841 cytoplasmic incompatibility. *Curr Biol*. doi:10.1016/j.cub.2022.01.052

842 Hosen MB, Islam MR, Begum F, Kabir Y, Howlader MZH. 2015. Oxidative stress
843 induced sperm DNA damage, a possible reason for male infertility. *Iran J*
844 *Reprod Med*.

845 Indriani C, Tantowijoyo W, Rancès E, Andari B, Prabowo E, Yusdi D, Ansari MR,
846 Wardhana DS, Supriyati E, Nurhayati I, Ernesia I, Setyawan S, Fitriana I,
847 Arguni E, Amelia Y, Ahmad RA, Jewell NP, Dufault SM, Ryan PA, Green BR,
848 McAdam TF, O'Neill SL, Tanamas SK, Simmons CP, Anders K, Utarini A.
849 2020. Reduced Dengue Incidence Following Deployments of *Wolbachia*-
850 Infected *Aedes aegypti* in Yogyakarta, Indonesia. *SSRN Electron J*.
851 doi:10.2139/ssrn.3544812

852 Jones DT. 1999. Protein secondary structure prediction based on position-specific
853 scoring matrices. *J Mol Biol*. doi:10.1006/jmbi.1999.3091

854 Kaur R, Leigh BA, Ritchie IT, Bordenstein SR. 2022. The Cif proteins from
855 Wolbachia prophage WO modify sperm genome integrity to establish
856 cytoplasmic incompatibility. *bioRxiv*. doi:10.1101/2022.01.15.476471

857 Kaur R, Shropshire JD, Cross KL, Leigh B, Mansueto AJ, Stewart V, Bordenstein
858 Sarah R., Bordenstein Seth R. 2021. Living in the endosymbiotic world of
859 Wolbachia: A centennial review. *Cell Host Microbe*.
860 doi:10.1016/j.chom.2021.03.006

861 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S,
862 Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond
863 AA, Heled J. 2012. Geneious. *Bioinformatics*.

864 Knizewski L, Kinch LN, Grishin N V., Rychlewski L, Ginalska K. 2007. Realm of PD-
865 (D/E)XK nuclease superfamily revisited: Detection of novel families with

modified transitive meta profile searches. *BMC Struct Biol.* doi:10.1186/1472-6807-7-40

Kozlov AG, Weiland E, Mittal A, Waldman V, Antony E, Fazio N, Pappu R V., Lohman TM. 2015. Intrinsically disordered C-terminal tails of E. coli single-stranded DNA binding protein regulate cooperative binding to single-stranded DNA. *J Mol Biol* **427**:763–774. doi:10.1016/j.jmb.2014.12.020

Kriesner P, Hoffmann A a, Lee SF, Turelli M, Weeks AR. 2013. Rapid Sequential Spread of Two Wolbachia Variants in Drosophila simulans. *PLoS Pathog* **9**:e1003607. doi:10.1371/journal.ppat.1003607

Layton EM, On J, Perlmuter JI, Bordenstein SR, Shropshire JD. 2019. Paternal grandmother age affects the strength of Wolbachia-induced cytoplasmic incompatibility in Drosophila melanogaster. *MBio* **10**. doi:10.1128/mBio.01879-19

LePage D, Bordenstein SR. 2013. Wolbachia: Can we save lives with a great pandemic? *Trends Parasitol* **29**:385–93. doi:10.1016/j.pt.2013.06.003

LePage DP, Metcalf JA, Bordenstein Sarah R., On J, Perlmuter JI, Shropshire JD, Layton EM, Funkhouser-Jones LJ, Beckmann JF, Bordenstein Seth R. 2017. Prophage WO genes recapitulate and enhance Wolbachia-induced cytoplasmic incompatibility. *Nature* **543**:243–247. doi:10.1038/nature21391

Lindsey ARI, Rice DW, Bordenstein Sarah R., Brooks AW, Bordenstein Seth R., Newton ILG. 2018. Evolutionary Genetics of Cytoplasmic Incompatibility Genes cifA and cifB in Prophage WO of Wolbachia. *Genome Biol Evol* **10**:434–451. doi:10.1093/gbe/evy012

Mains JW, Brelsfoard CL, Rose RI, Dobson SL. 2016. Female adult aedes Albopictus suppression by Wolbachia-infected male mosquitoes. *Sci Rep* **6**. doi:10.1038/srep33846

Mills WK, Lee YCG, Kochendoerfer AM, Dunleavy EM, Karpen GH. 2019. RNA from a simple-tandem repeat is required for sperm maturation and male fertility in drosophila melanogaster. *Elife* **8**. doi:10.7554/eLife.48940

Moreira LA, Iturbe-Ormaetxe I, Jeffery J a, Lu G, Pyke AT, Hedges LM, Rocha BC, Hall-Mendelin S, Day A, Riegler M, Hugo LE, Johnson KN, Kay BH, McGraw E a, van den Hurk AF, Ryan P a, O'Neill SL. 2009. A Wolbachia symbiont in Aedes aegypti limits infection with dengue, Chikungunya, and Plasmodium. *Cell* **139**:1268–78. doi:10.1016/j.cell.2009.11.042

Nazni WA, Hoffmann AA, NoorAfizah A, Cheong YL, Mancini M V., Golding N, Kamarul GMR, Arif MAK, Thohir H, NurSyamimi H, ZatilAqmar MZ, NurRuqqayah M, NorSyazwani A, Faiz A, Irfan FRMN, Rubaaini S, Nuradila N, Nizam NMN, Irwan SM, Endersby-Harshman NM, White VL, Ant TH, Herd CS, Hasnor AH, AbuBakar R, Hapsah DM, Khadijah K, Kamilan D, Lee SC, Paid YM, Fadzilah K, Topek O, Gill BS, Lee HL, Sinkins SP. 2019. Establishment of Wolbachia Strain wAlbB in Malaysian Populations of Aedes aegypti for Dengue Control. *Curr Biol.* doi:10.1016/j.cub.2019.11.007

O'Neill SL. 2018. The use of wolbachia by the world mosquito program to interrupt transmission of aedes Aegypti transmitted virusesAdvances in Experimental Medicine and Biology. pp. 355–360. doi:10.1007/978-981-10-8727-1_24

O'Neill SL, Ryan PA, Turley AP, Wilson G, Retzki K, Iturbe-Ormaetxe I, Dong Y, Kenny N, Paton CJ, Ritchie SA, Brown-Kenyon J, Stanford D, Wittmeier N, Anders KL, Simmons CP. 2018. Scaled deployment of Wolbachia to protect the community from Aedes transmitted arboviruses. *Gates Open Res* **2**:36. doi:10.12688/gatesopenres.12844.1

Obarska-Kosinska A, Taylor JE, Callow P, Orlowski J, Bujnicki JM, Kneale GG. 2008. HsdR Subunit of the Type I Restriction-Modification Enzyme EcoR124I: Biophysical Characterisation and Structural Modelling. *J Mol Biol*. doi:10.1016/j.jmb.2007.11.024

Puggioli A, Calvitti M, Moretti R, Bellini R. 2016. wPip Wolbachia contribution to Aedes albopictus SIT performance: Advantages under intensive rearing. *Acta Trop* **164**:473–481. doi:10.1016/j.actatropica.2016.10.014

Rathke C, Baarends WM, Awe S, Renkawitz-Pohl R. 2014. Chromatin dynamics during spermiogenesis. *Biochim Biophys Acta - Gene Regul Mech* **1839**:155–168. doi:10.1016/j.bbagr.2013.08.004

Rathke C, Baarends WM, Jayaramaiah-Raja S, Bartkuhn M, Renkawitz R, Renkawitz-Pohl R. 2007. Transition from a nucleosome-based to a protamine-based chromatin configuration during spermiogenesis in Drosophila. *J Cell Sci* **120**:1689–1700. doi:10.1242/jcs.004663

Ryan PA, Turley AP, Wilson G, Hurst TP, Retzki K, Brown-Kenyon J, Hodgson L, Kenny N, Cook H, Montgomery BL, Paton CJ, Ritchie SA, Hoffmann AA, Jewell NP, Tanamas SK, Anders KL, Simmons CP, O'Neill SL. 2020. Establishment of wMel Wolbachia in Aedes aegypti mosquitoes and reduction of local dengue transmission in Cairns and surrounding locations in northern Queensland, Australia. *Gates Open Res*. doi:10.12688/gatesopenres.13061.2

Sharma S, Schiller MR. 2019. The carboxy-terminus, a key regulator of protein function. *Crit Rev Biochem Mol Biol*. doi:10.1080/10409238.2019.1586828

Shropshire J, Leigh B, Bordenstein SR, Initiative M. 2020. Symbiont-mediated cytoplasmic incompatibility : What have we learned in 50 years? *Elife* 1–51. doi:10.20944/preprints202008.0350.v1

Shropshire JD, Bordenstein SR. 2019. Two-by-one model of cytoplasmic incompatibility: Synthetic recapitulation by transgenic expression of cifa and cifB in drosophila. *PLoS Genet* **15**. doi:10.1371/journal.pgen.1008221

Shropshire JD, Hamant E, Cooper BS. 2021a. Male age and Wolbachia dynamics: Determining how fast and why bacterial densities and cytoplasmic incompatibility strengths vary. *bioRxiv* 2021.06.01.446638. doi:10.1128/mBio.02998-21

Shropshire JD, Kalra M, Bordenstein SR. 2020. Evolution-guided mutagenesis of the cytoplasmic incompatibility proteins: Identifying CifA's complex functional repertoire and new essential regions in CifB. *PLoS Pathog* **16**:e1008794. doi:10.1371/journal.ppat.1008794

Shropshire JD, On J, Layton EM, Zhou H, Bordenstein SR. 2018. One prophage WO gene rescues cytoplasmic incompatibility in Drosophila melanogaster. *Proc Natl Acad Sci U S A* **115**:4987–4991. doi:10.1073/pnas.1800650115

Shropshire JD, Rosenberg R, Bordenstein SR. 2021b. The impacts of cytoplasmic incompatibility factor (cifA and cifB) genetic variation on phenotypes. *Genetics*. doi:10.1093/GENETICS/IYAA007

Singleton MR, Dillingham MS, Gaudier M, Kowalczykowski SC, Wigley DB. 2004. Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature*. doi:10.1038/nature02988

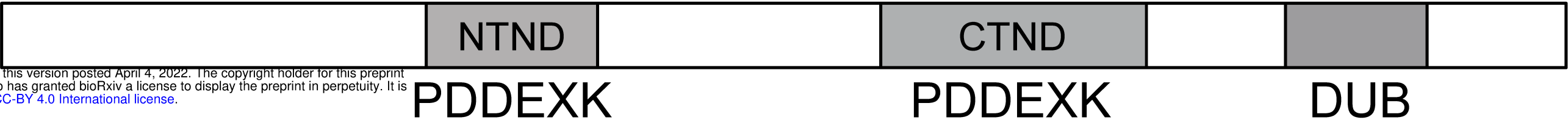
Šišáková E, Stanley LK, Weiserová M, Szczelkun MD. 2008. A RecB-family nuclease motif in the Type I restriction endonuclease EcoR124I. *Nucleic Acids Res*. doi:10.1093/nar/gkn333

Steczkiwicz K, Muszewska A, Knizewski L, Rychlewski L, Ginalski K. 2012. SURVEY AND SUMMARY: Sequence, structure and functional diversity of

- 966 PD-(D/E)XK phosphodiesterase superfamily. *Nucleic Acids Res.*
- 967 doi:10.1093/nar/gks382
- 968 Sun G, Zhang M, Chen H, Hochstrasser M. 2021. The Wolbachia CinB Nuclease is
- 969 Sufficient for Induction of Cytoplasmic Incompatibility. *bioRxiv*
- 970 2021.10.22.465375.
- 971 Tjeertes J V., Miller KM, Jackson SP. 2009. Screen for DNA-damage-responsive
- 972 histone modifications identifies H3K9Ac and H3K56Ac in human cells. *EMBO*
- 973 *J* **28**:1878–1889. doi:10.1038/emboj.2009.119
- 974 Utarini A, Indriani C, Ahmad RA, Tantowijoyo W, Arguni E, Ansari MR, Supriyati E,
- 975 Wardana DS, Metika Y, Ernesia I, Nurhayati I, Prabowo E, Andari B, Green
- 976 BR, Hodgson L, Cutcher Z, Ryan PA, O'Neill SL, Dufault SM, Tanamas SK,
- 977 Jewell NP, Anders KL, Simmons CP, for the AWED study group. 2021.
- 978 Efficacy of Wolbachia-infected mosquito deployments for the control of
- 979 dengue. *N Engl J Med.* doi:10.1056/NEJMoa2030243
- 980 Vasudevan D, Ryoo HD. 2016. Detection of cell death in Drosophila
- 981 tissuesMethods in Molecular Biology. doi:10.1007/978-1-4939-3581-9_11
- 982 Wang H, Xiao Y, Chen X, Zhang M, Sun G, Wang Feng, Wang L, Zhang H, Zhang
- 983 X, Yang X, Li W, Wei Y, Yao D, Zhang B, Li J, Cui W, Wang Fenghua, Chen
- 984 C, Shen W, Su D, Bai F, Huang J, Ye S, Zhang L, Ji X, Wang W, Wang Z,
- 985 Hochstrasser M, Yang H. 2022. Crystal Structures of Wolbachia CidA and
- 986 CidB Reveal Determinants of Bacteria-induced Cytoplasmic Incompatibility
- 987 and Rescue. *Nat Commun* **13**. doi:10.1038/s41467-022-29273-w
- 988 Weinert LA, Araujo-Jnr E V., Ahmed MZ, Welch JJ. 2015. The incidence of
- 989 bacterial endosymbionts in terrestrial arthropods. *Proc R Soc B Biol Sci* **282**.
- 990 doi:10.1098/rspb.2015.0249
- 991 Xiao Y, Chen H, Wang H, Zhang M, Chen X, Berk JM, Zhang Lilin, Wei Y, Li W,
- 992 Cui W, Wang F, Wang Q, Cui C, Li T, Chen C, Ye S, Zhang Lei, Ji X, Huang J,
- 993 Wang W, Wang Z, Hochstrasser M, Yang H. 2021. Structural and mechanistic
- 994 insights into the complexes formed by Wolbachia cytoplasmic incompatibility
- 995 factors . *Proc Natl Acad Sci* **118**:e2107699118. doi:10.1073/pnas.2107699118
- 996 Yamada R, Floate KD, Riegler M, O'Neill SL. 2007. Male development time
- 997 influences the strength of wolbachia-induced cytoplasmic incompatibility
- 998 expression in Drosophila melanogaster. *Genetics* **177**:801–808.
- 999 doi:10.1534/genetics.106.068486
- 1000 Zheng X, Zhang D, Li Y, Yang C, Wu Y, Liang X, Liang Y, Pan X, Hu L, Sun Q,
- 1001 Wang X, Wei Y, Zhu J, Qian W, Yan Z, Parker AG, Gilles JRL, Bourtzis K,
- 1002 Bouyer J, Tang M, Zheng B, Yu J, Liu J, Zhuang J, Hu Zhigang, Zhang M,
- 1003 Gong JT, Hong XY, Zhang Z, Lin L, Liu Q, Hu Zhiyong, Wu Z, Baton LA,
- 1004 Hoffmann AA, Xi Z. 2019. Incompatible and sterile insect techniques combined
- 1005 eliminate mosquitoes. *Nature* **572**:56–61. doi:10.1038/s41586-019-1407-9
- 1006 Zug R, Hammerstein P. 2012. Still a host of hosts for Wolbachia: analysis of recent
- 1007 data suggests that 40% of terrestrial arthropod species are infected. *PLoS*
- 1008 *One* **7**:e38544. doi:10.1371/journal.pone.0038544
- 1009

CifB full protein

bioRxiv preprint doi: <https://doi.org/10.1101/2022.04.04.487029>; this version posted April 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



CifB_{ΔD}



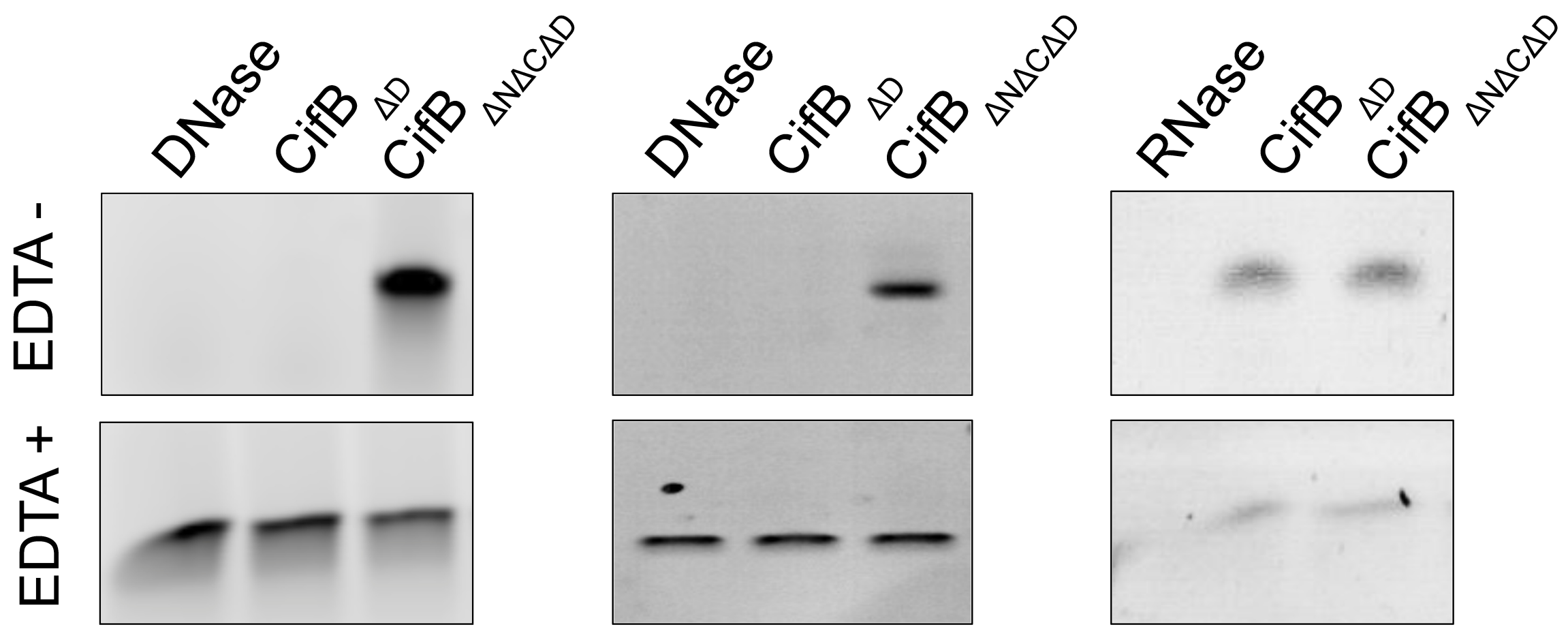
CifB_{ΔNΔCAD}

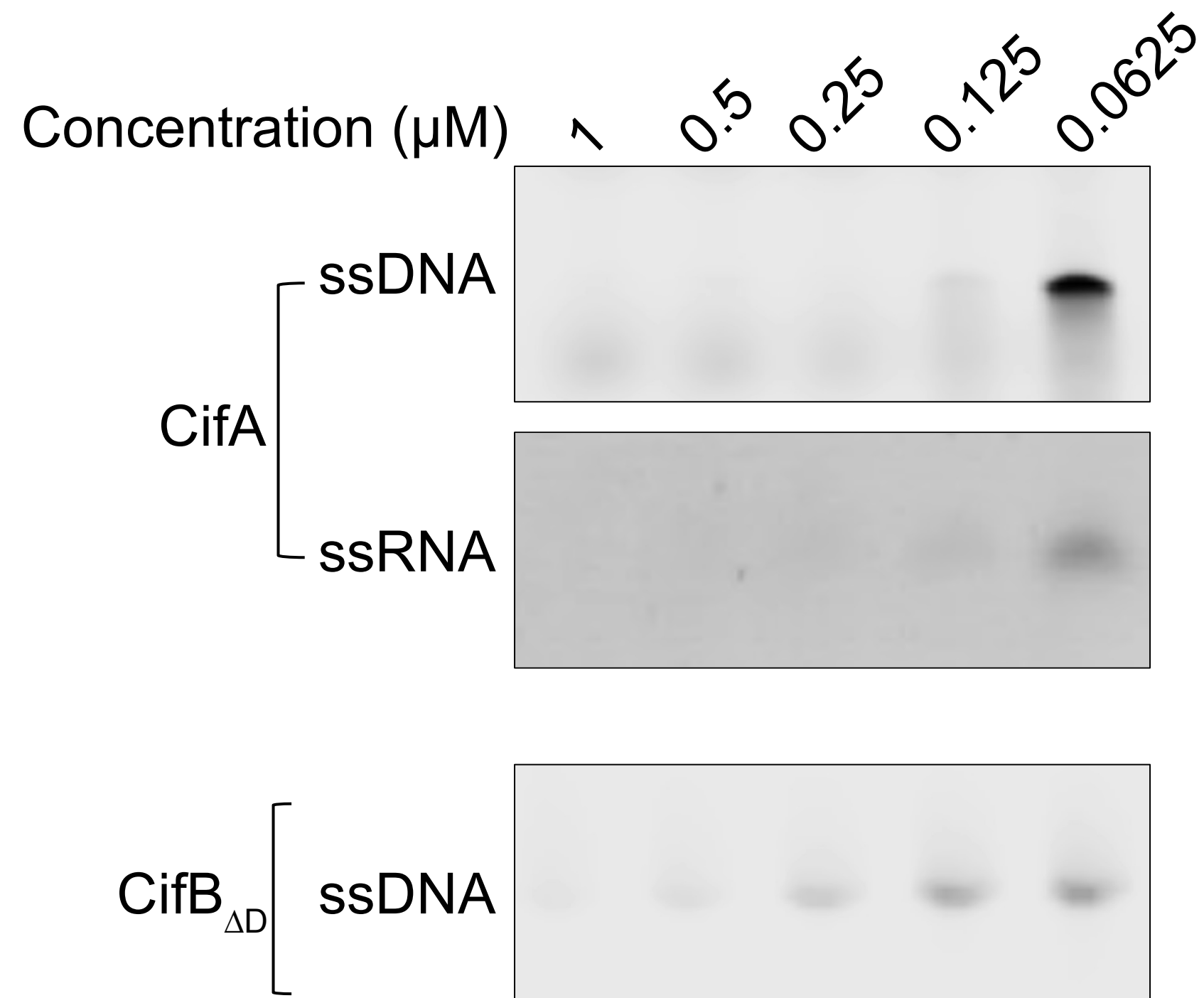
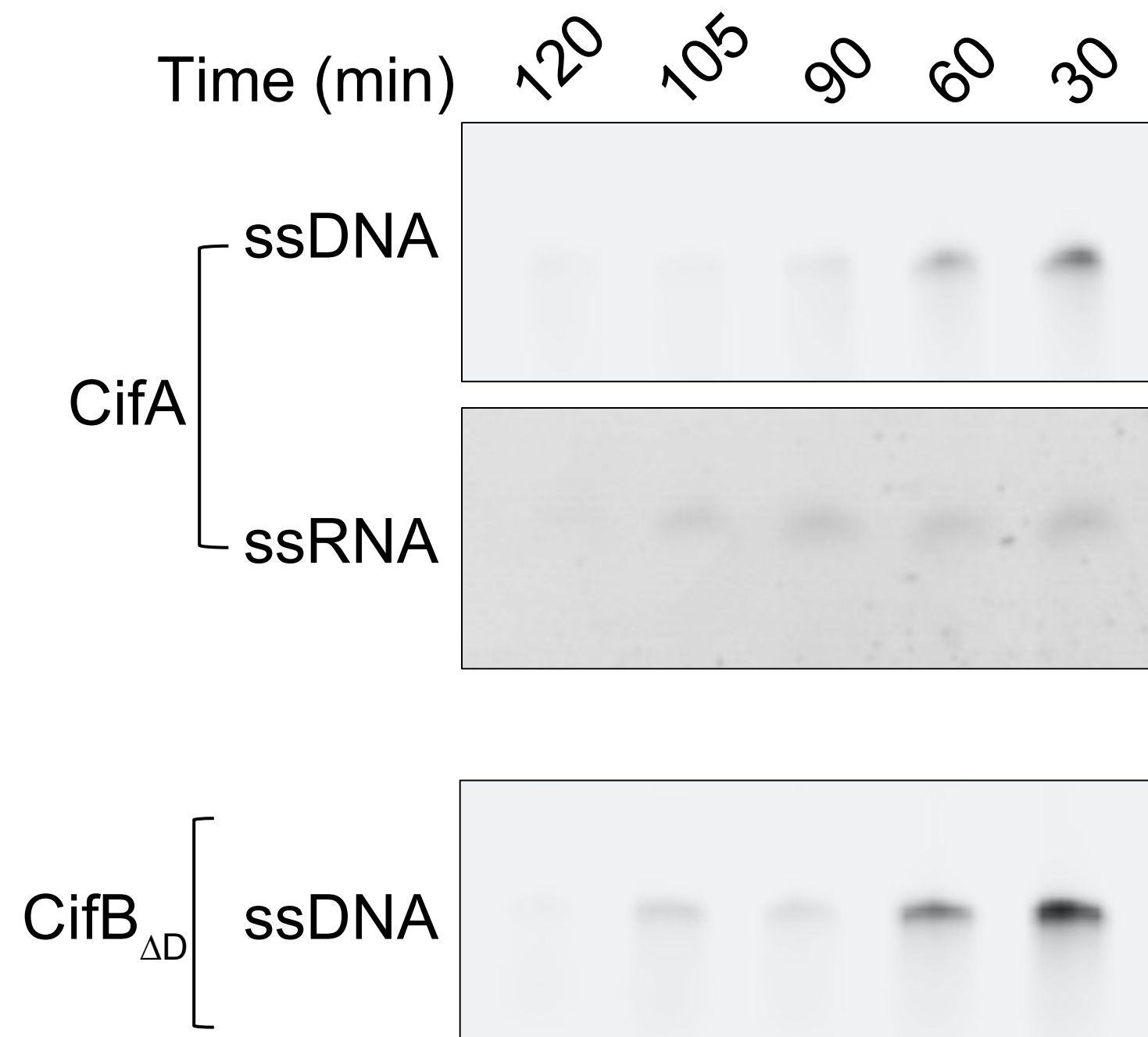


ssDNA
(70 bps)

dsDNA
(140 bps)

ssRNA
(45 bps)



A**B**

Canoe

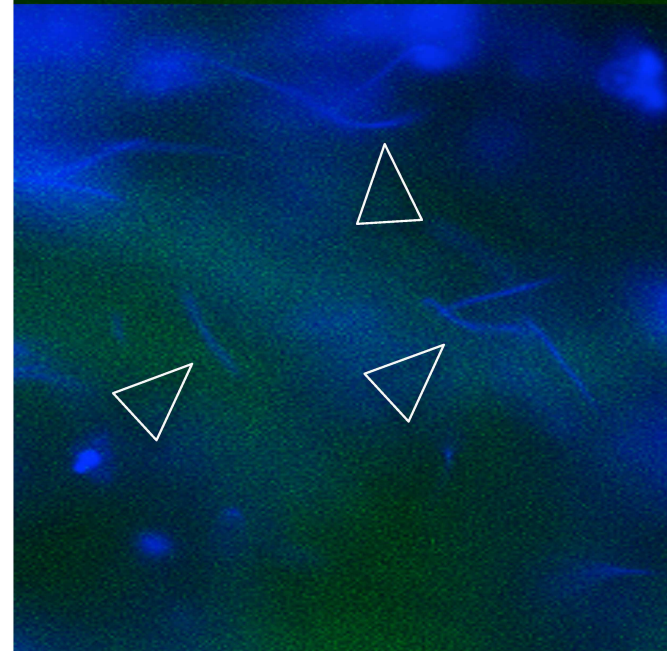
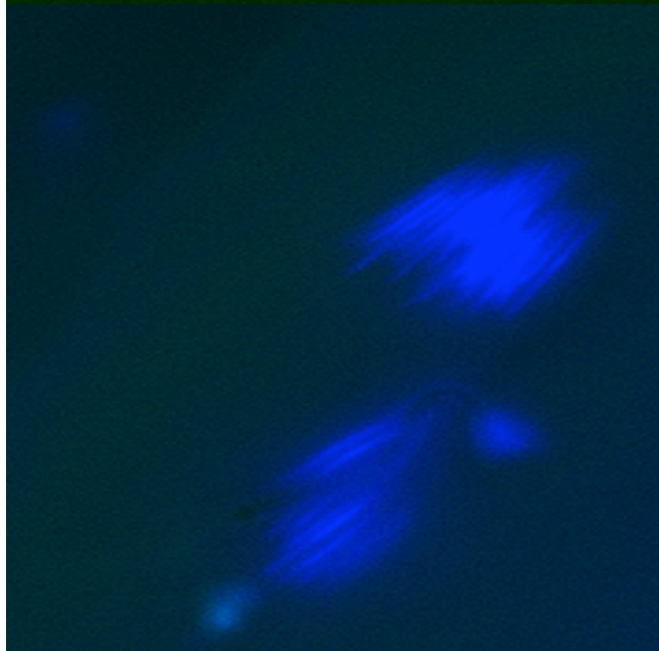
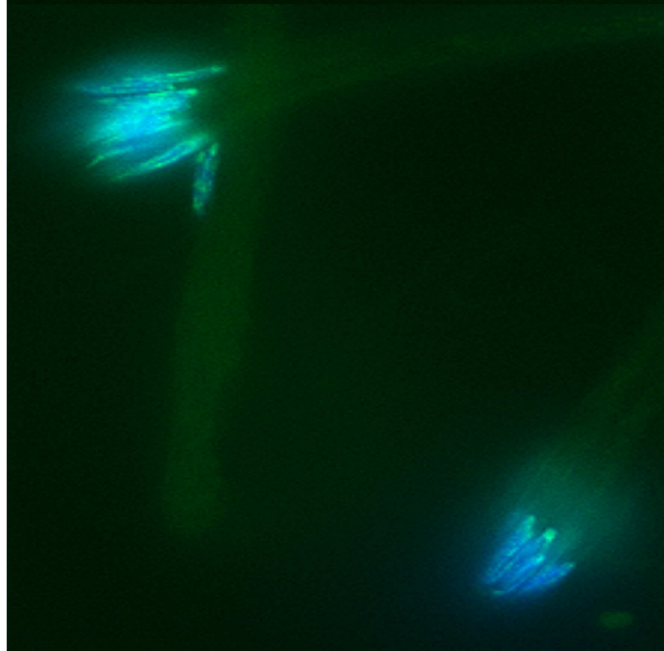
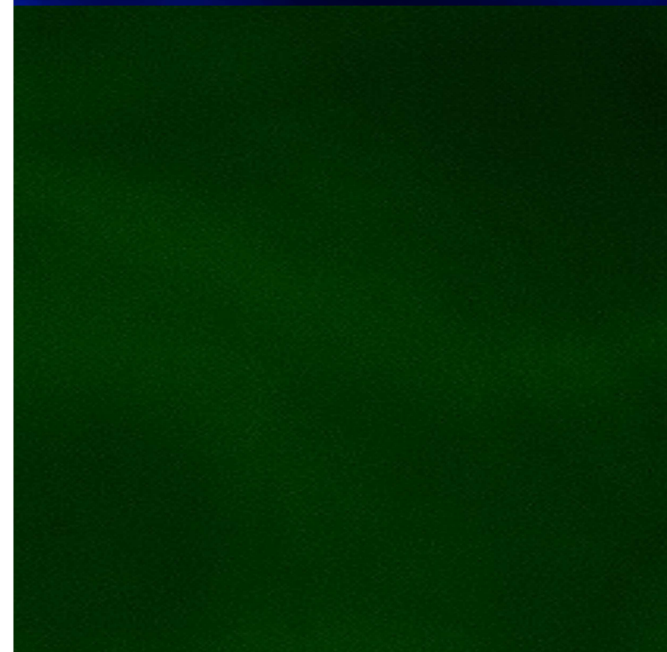
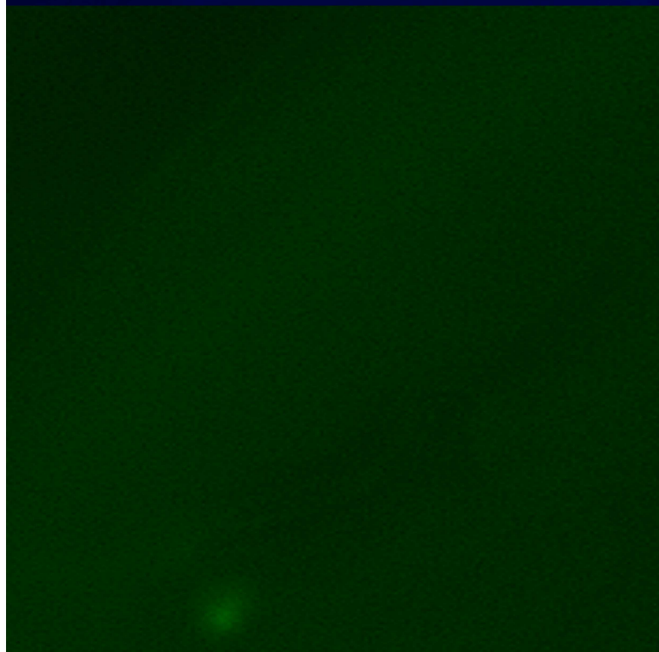
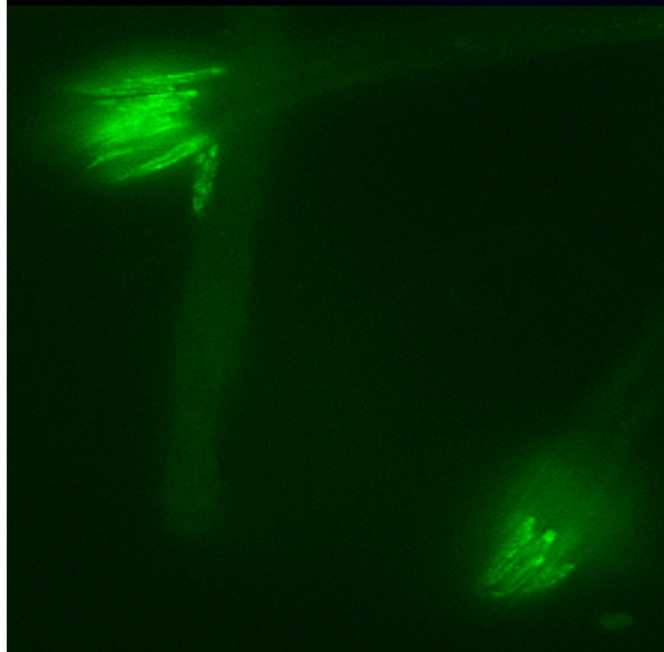
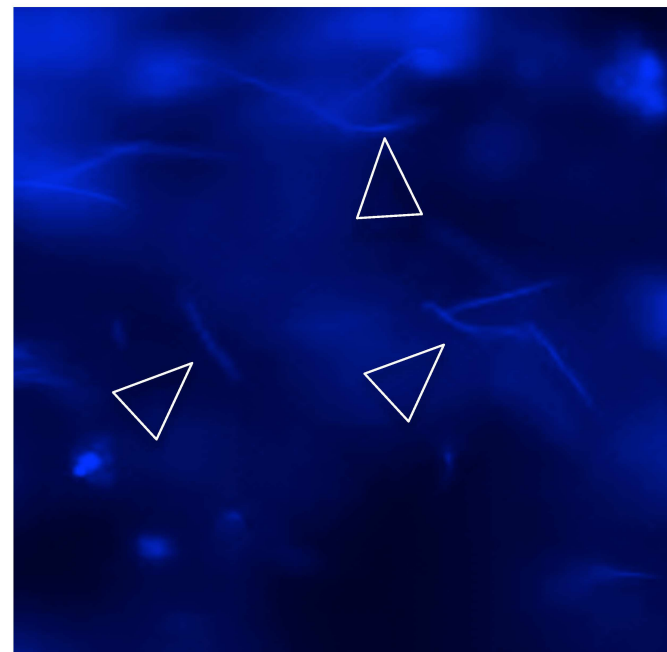
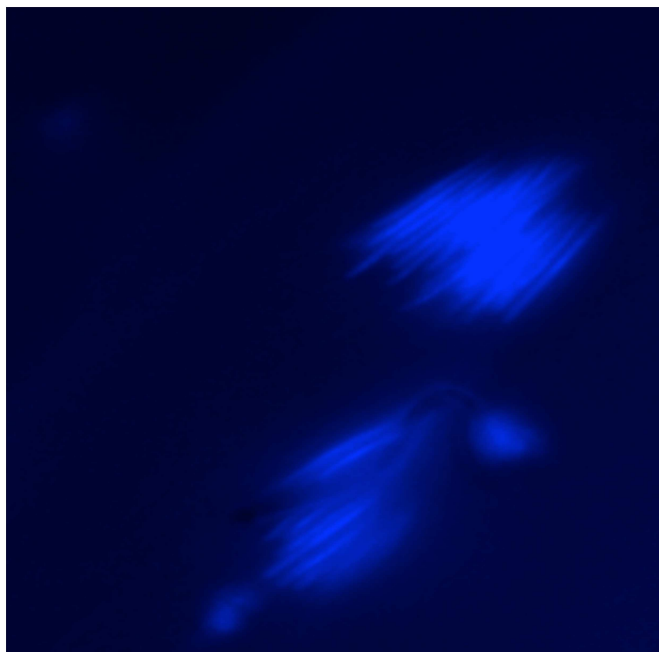
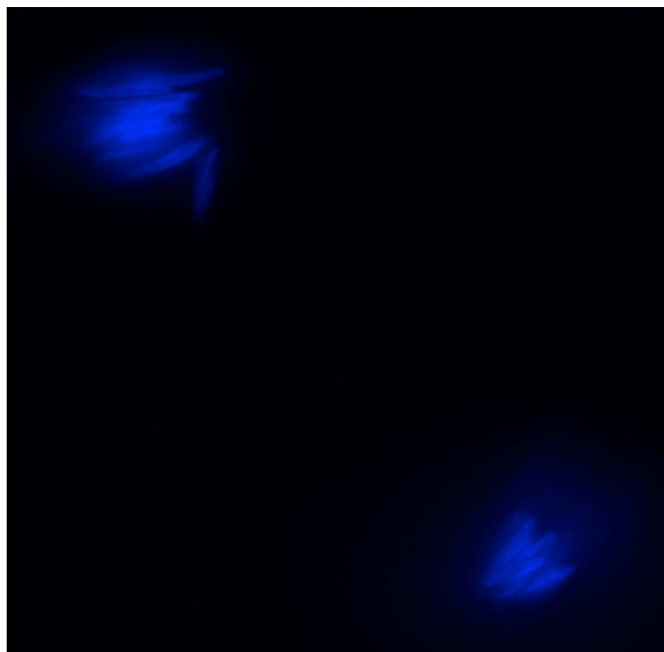
Needle

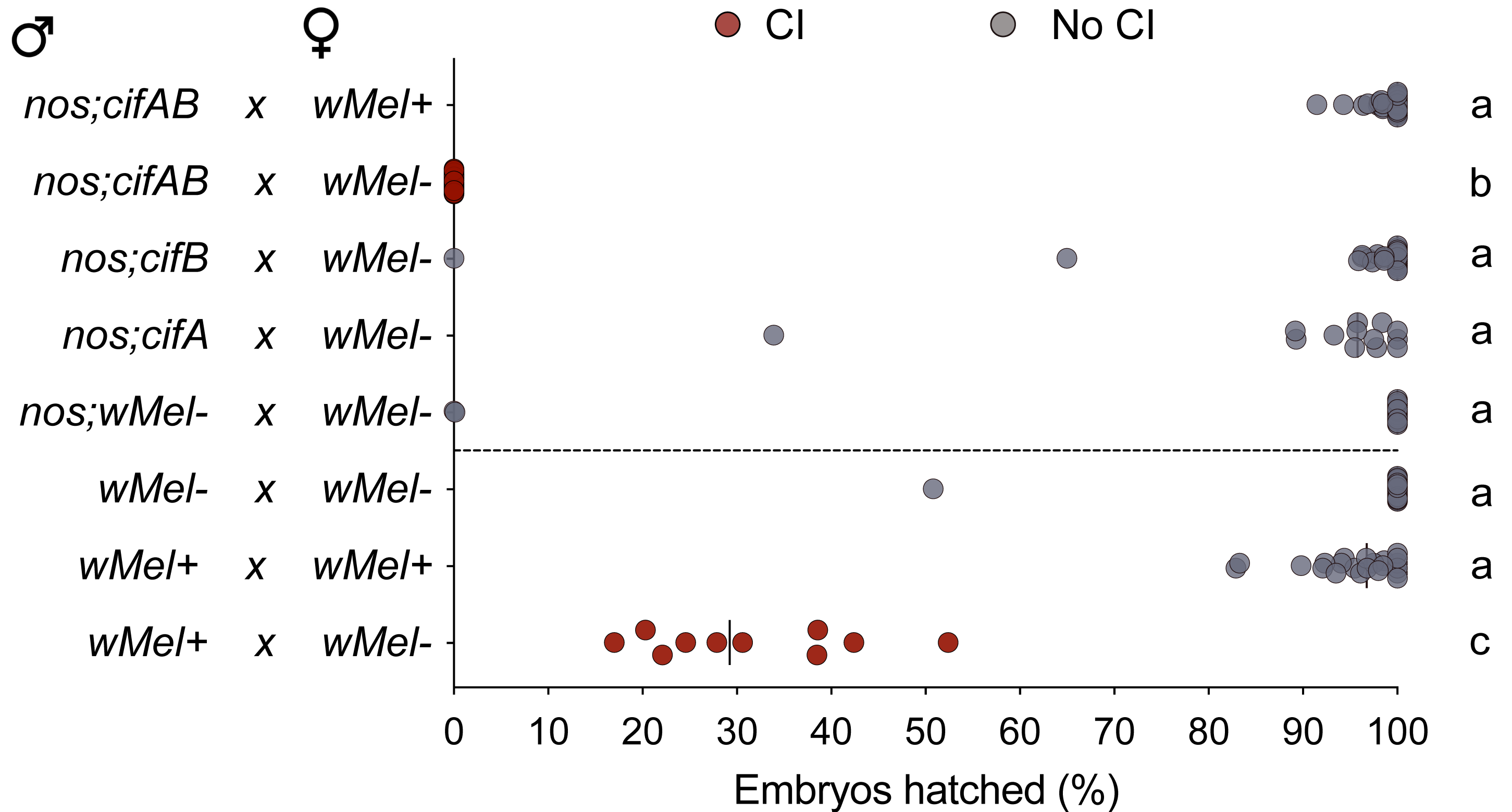
Mature sperms

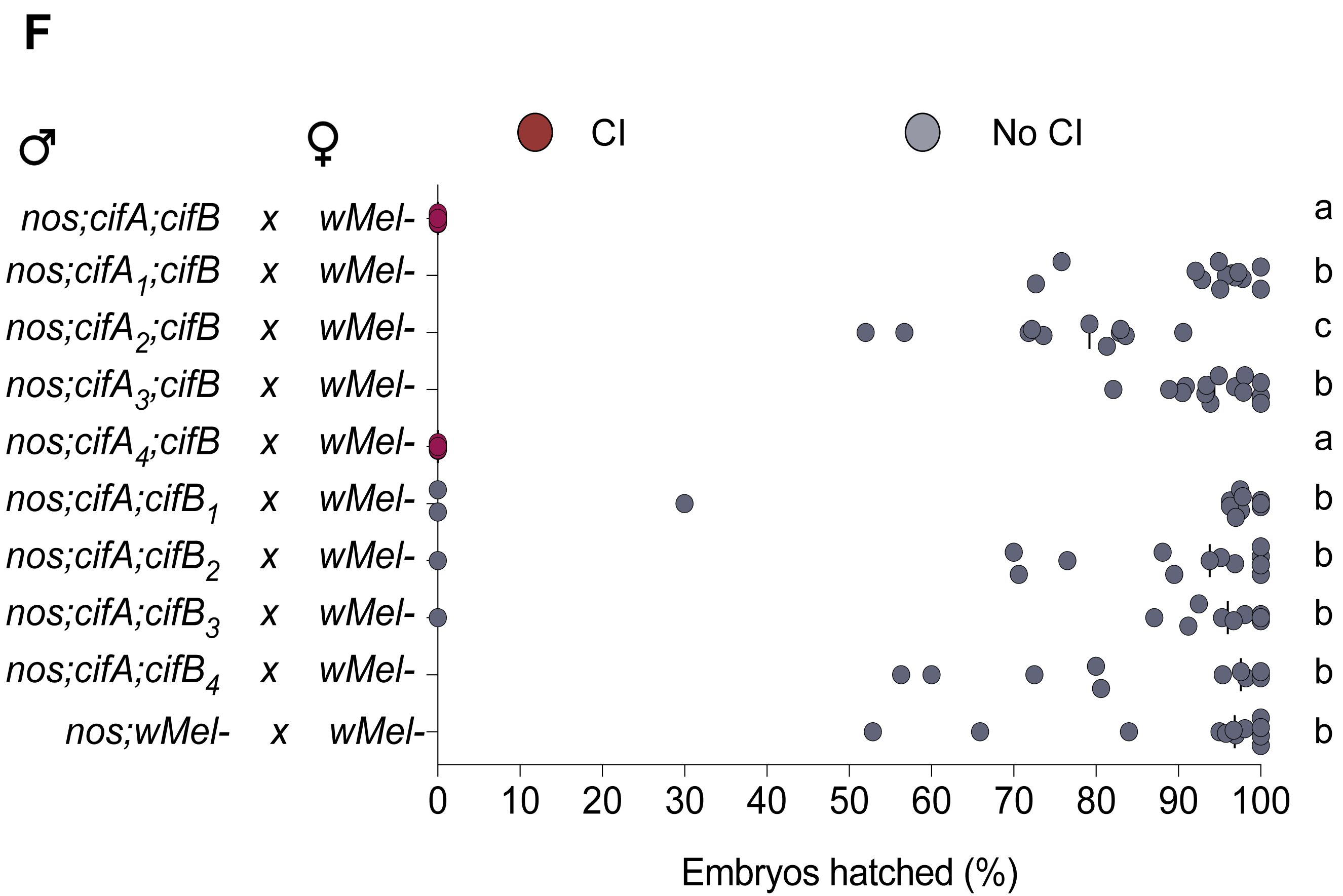
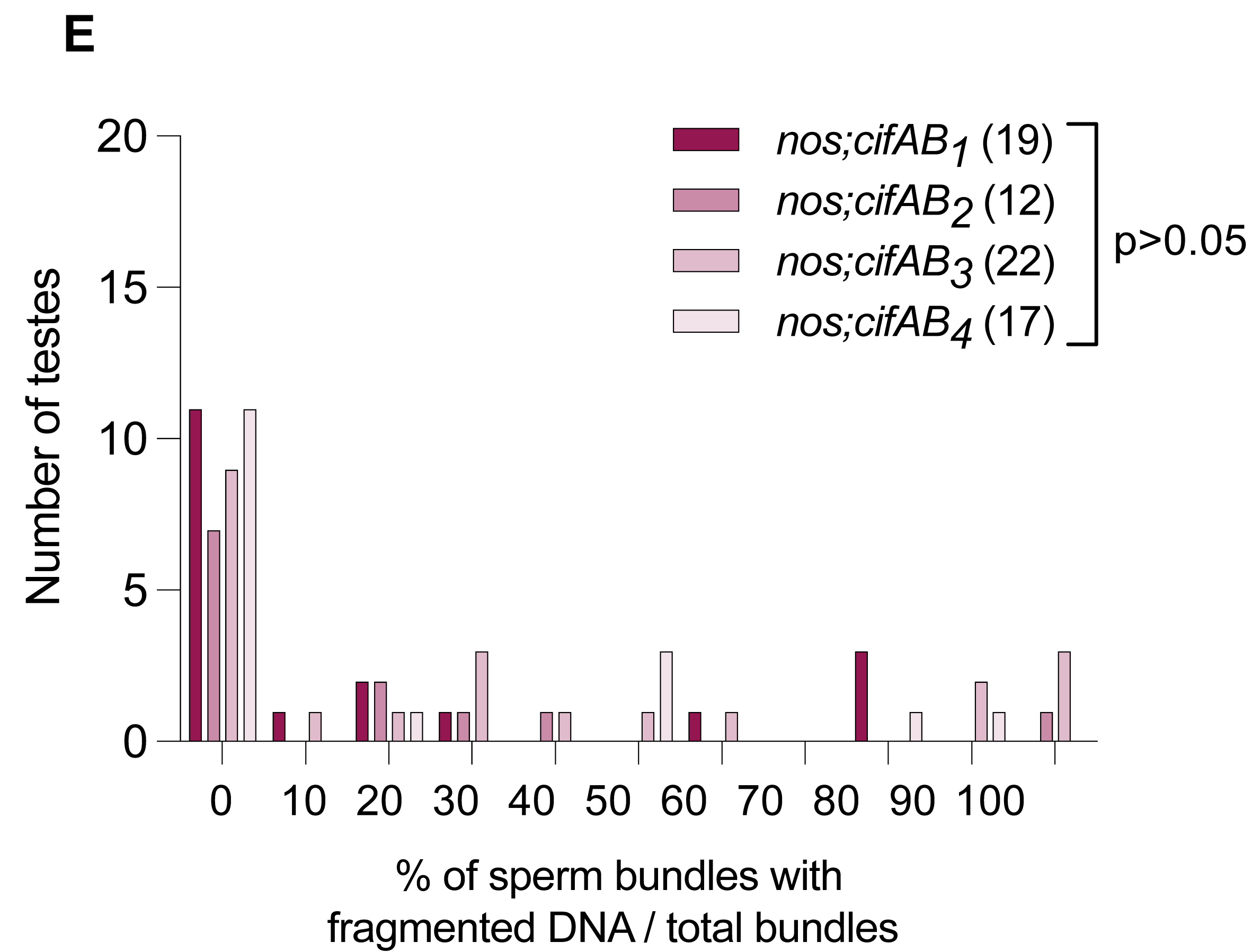
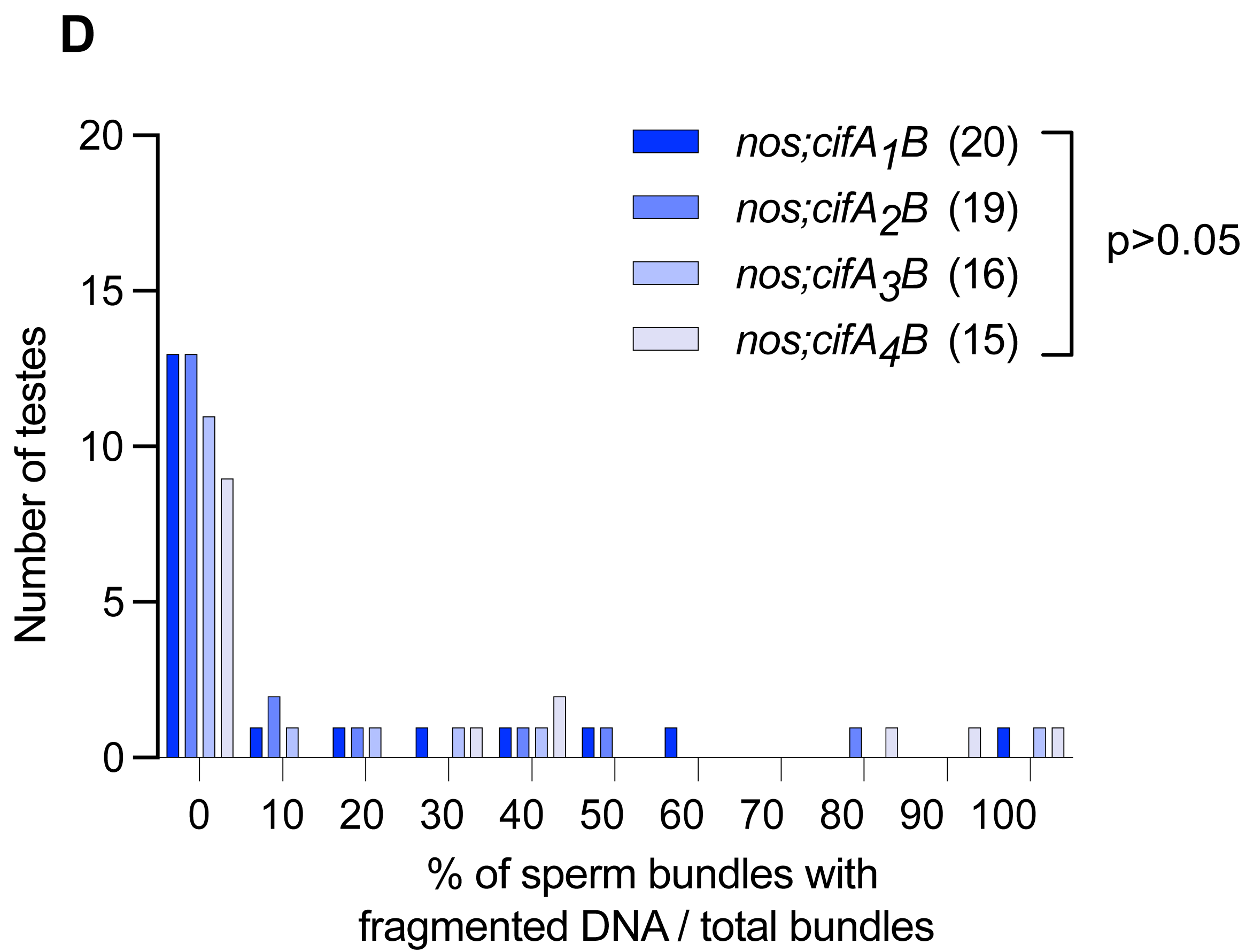
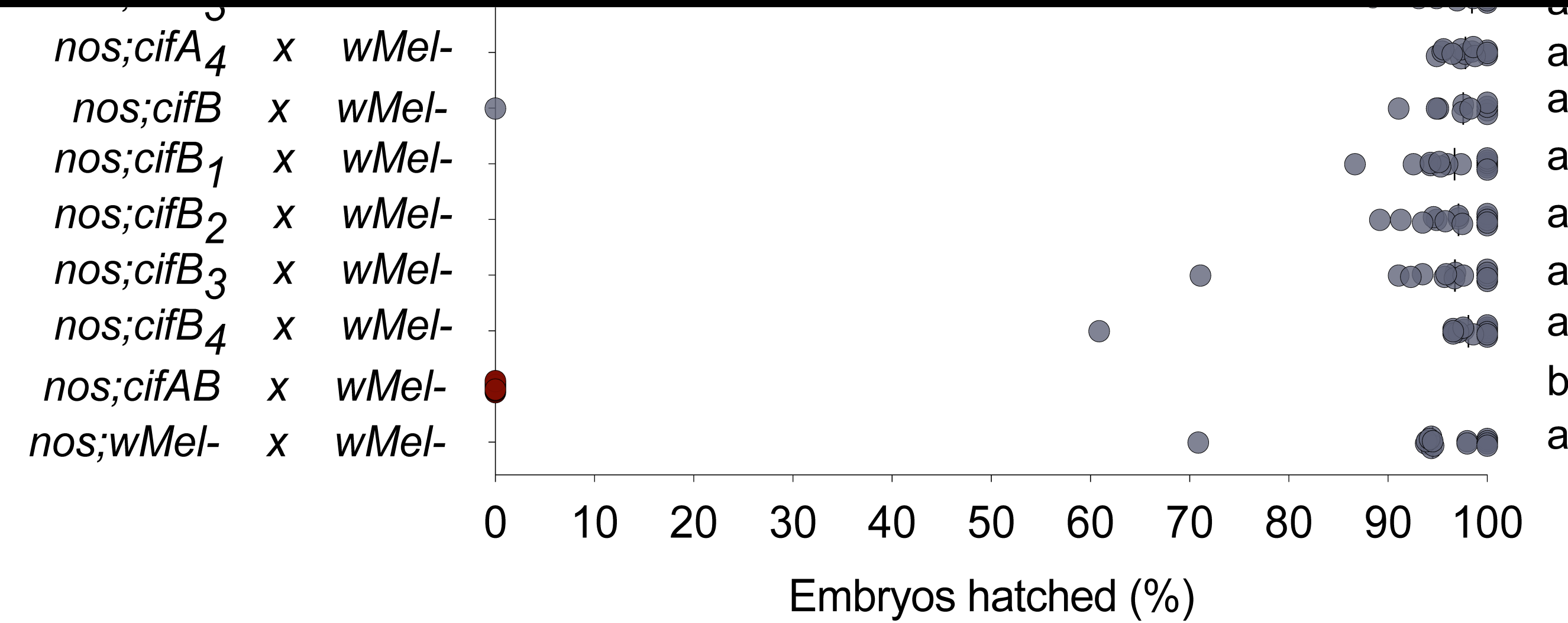
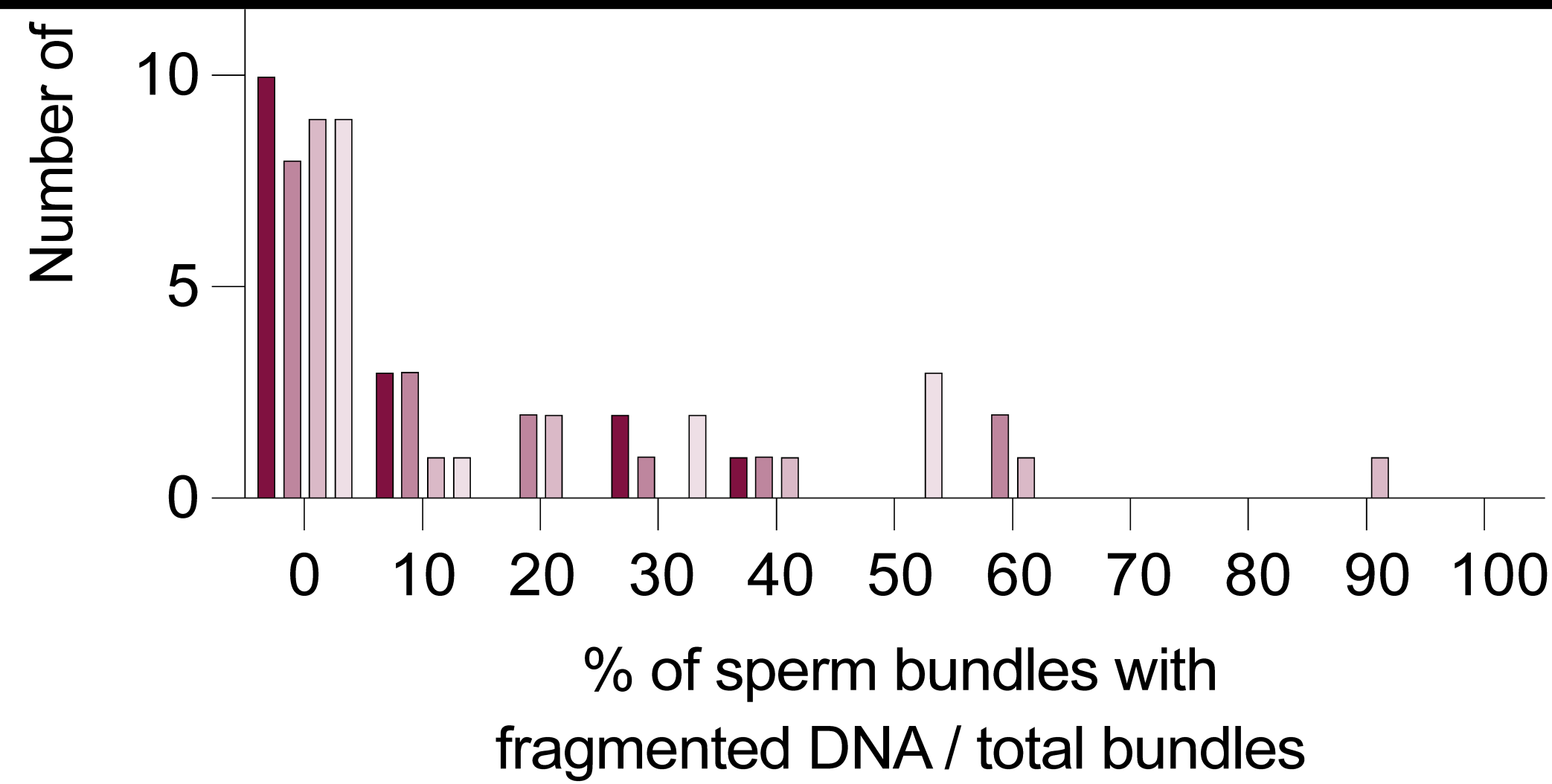
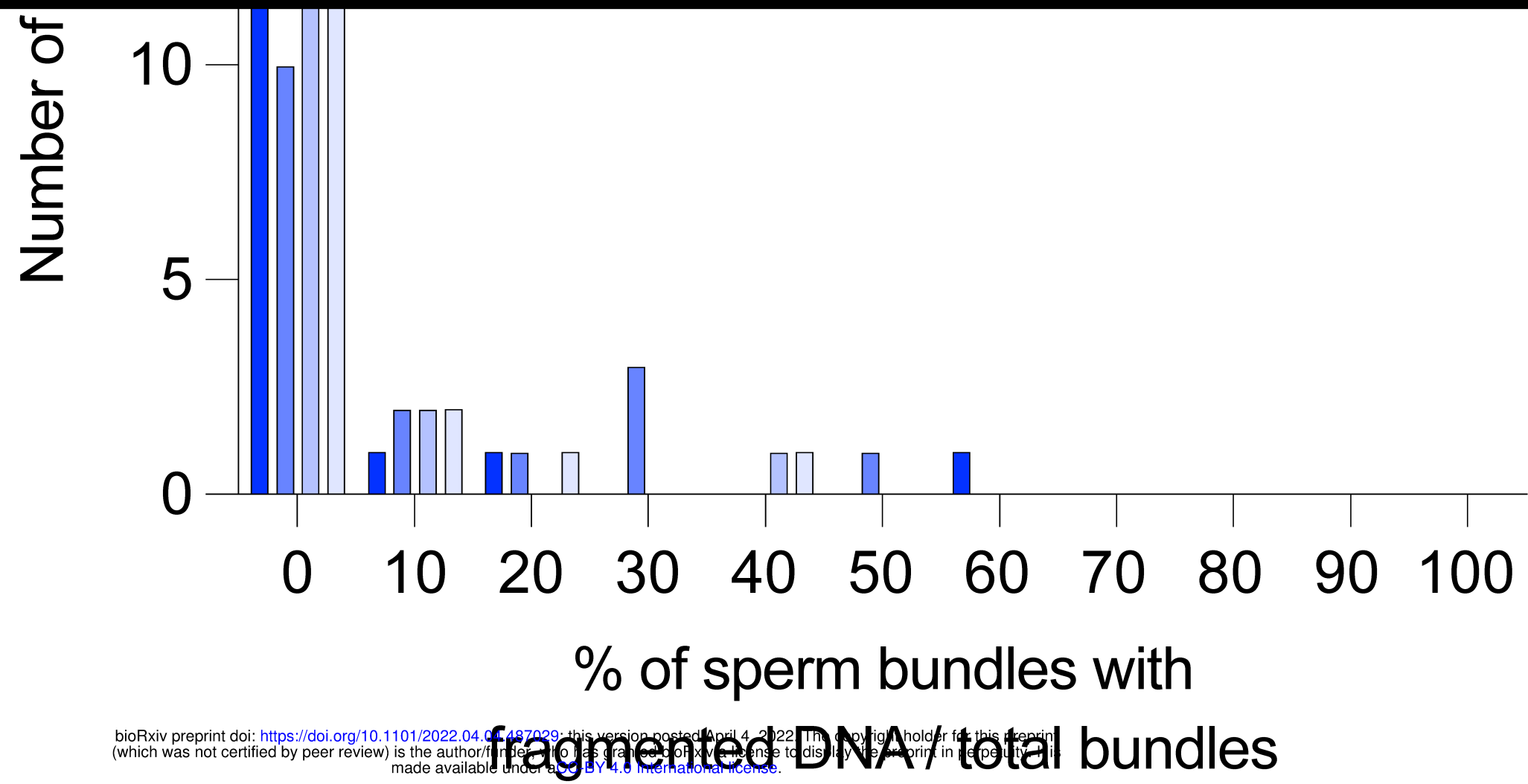
DAPI

TUNEL

MERGE



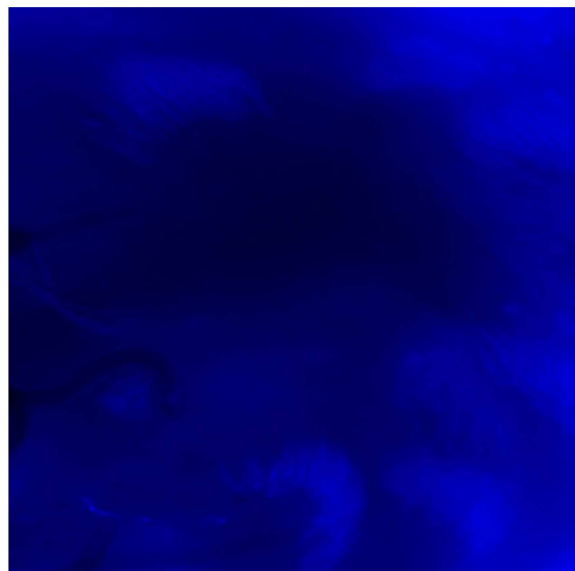




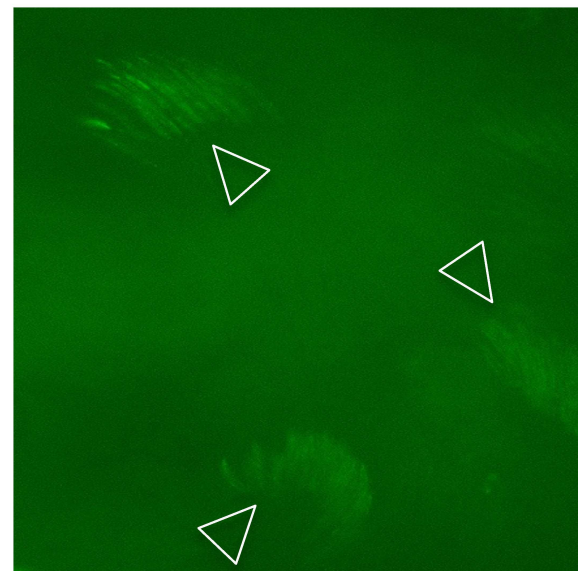
nos;cifA₄B

canoe-stage spermatids

DAPI



TUNEL



MERGE

