CellPress
REVIEWS

Letter

Models and Nomenclature for Cytoplasmic Incompatibility: Caution over Premature Conclusions – A Response to Beckmann *et al.*

J. Dylan Shropshire,[1,2,*] Brittany Leigh,[1,2] Sarah R. Bordenstein,[1,2] Anne Duplouy,[3] Markus Riegler,[4] Jeremy C. Brownlie,[5] and Seth R. Bordenstein[1,2,6,7,*]

Recent studies have identified two genes in bacteriophage WO, *cifA* and *cifB*, that contribute to the induction of cytoplasmic incompatibility (CI) [1,2], and one of these two genes, *cifA*, rescues it [3]. These findings underpin a two-by-one genetic model (Figure 1A) that reflects current understanding of CI genetics and embraces various functional models [3] (Figure 1B). A recent article by Beckmann *et al.* [4] provides interesting ideas about the mechanism and evolutionary history of the CI genes. Therein, they claim that it is 'clearer than ever that the CI induction and rescue stem from a toxin–antidote (TA) system', and that disputes regarding the operon status of the *cif* genes are semantic. They also propose a new nomenclature to describe the genes. It is important to test hypotheses and develop nomenclature carefully in the context of current data because misconceptions can sometimes become a narrative for those unfamiliar with the evidence. Here, we present and evaluate three points of criticism of the arguments related to the TA model, the operon hypothesis, and the proposed gene nomenclature. We recommend caution and nuance in interpreting current data (and lack thereof). As we will frequently note, more research will be necessary before a functional narrative should be prescribed for CI.

The TA Model

The proposed TA model [4] assumes that male-derived CifB (the presumed toxin) is transferred to the host embryo during fertilization and that its associated defects are rescued upon binding to embryo-derived CifA (the presumed antidote). Although CifA and CifB bind to each other *in vitro* [3], there is no evidence for transfer of CifB to the embryo. In fact, there is evidence to the contrary that was mentioned by the authors. Mass spectrometry and SDS-PAGE analyses indicate that CifA, but not CifB, is present in the spermatheca of females mated to infected males [5]. There are numerous technical explanations for why CifB is absent – for example, CifB protein expression levels may be below the threshold of detection – but the current evidence is consistent with a model wherein CifA, but not CifB, reaches the female reproductive tract. Therefore, it cannot be assumed that CifB from the male directly interacts with CifA in the embryo. For this reason the essential premise of the TA model is unfounded. One of several alternative hypotheses is that, instead of CifB transferring with the sperm, a host product modified by CifA and/or CifB leads to CI induction, and the host modification is then reversed by CifA in the embryo [3]. It is notable that, if supported, this model and others (Figure 1B) would contradict the proposed TA model.

The *cif* Operon Hypothesis

cifA and *cifB* transcriptional regulation is proposed to occur as an operon [4], defined as a set of genes that are coregulated by a single promoter [6]. However, the number of *cif* promoters is crucially unknown. Moreover, there is a marked ninefold reduction in gene expression of *cifB* relative to *cifA* as well as a predicted hairpin termination element in the short intergenic region [7,8]. This evidence is consistent with either one or two promoters because the hairpin termination element could contribute to the transcriptional differences. The genes may have two separate promoters with two separate functional transcripts, but aberrant cotranscription could occur owing to an imperfect hairpin terminator. In such a model, *cifA* and *cifB* would not form an operon. Alternatively, *cifA* and *cifB* may have a single promoter upstream of *cifA*, and the imperfect terminator would provide a mechanism to control the large transcriptional differences between *cifA* and *cifB*. Therefore, we do not see this as a 'semantic debate' but as a hypothesis that requires further testing, and conclusions can only be drawn once firm evidence of a single promoter is established.

Gene Nomenclature

Useful gene nomenclature should accurately and conservatively describe the phenotype(s) of mutants of the corresponding gene [9]. The CI factor (*cif* and *cif*-like) nomenclature conservatively names two separate and codiverging genes involved in CI into at least four clades designated types I–IV (Figure 1C). Only type I *cif* genes (CI deubiquitylases, *cid* in the competing nomenclature) have been shown to be involved in CI [1,4] and rescue [3]. Homologous genes in distant clades (types II–IV) are therefore denoted *cif*-like to reflect the fact that CI function has not been established [1]. By contrast, the CI nuclease (*cin*) and CI nuclease/deubiquitylase (*cnd*) nomenclatures prematurely assigns CI function despite the absence of evidence for a causal role in CI. Moreover, the *cid*, *cin*, and *cnd* nomenclature inaccurately describes gene A (Figure 1D) based on the purported function of gene B (Figure 1E), and is not conservative to the unresolved functions of these putatively polyvalent proteins. It would also group types II–IV into one gene category, and type I in another, without any phylogenetic rationality with respect to divergence levels.

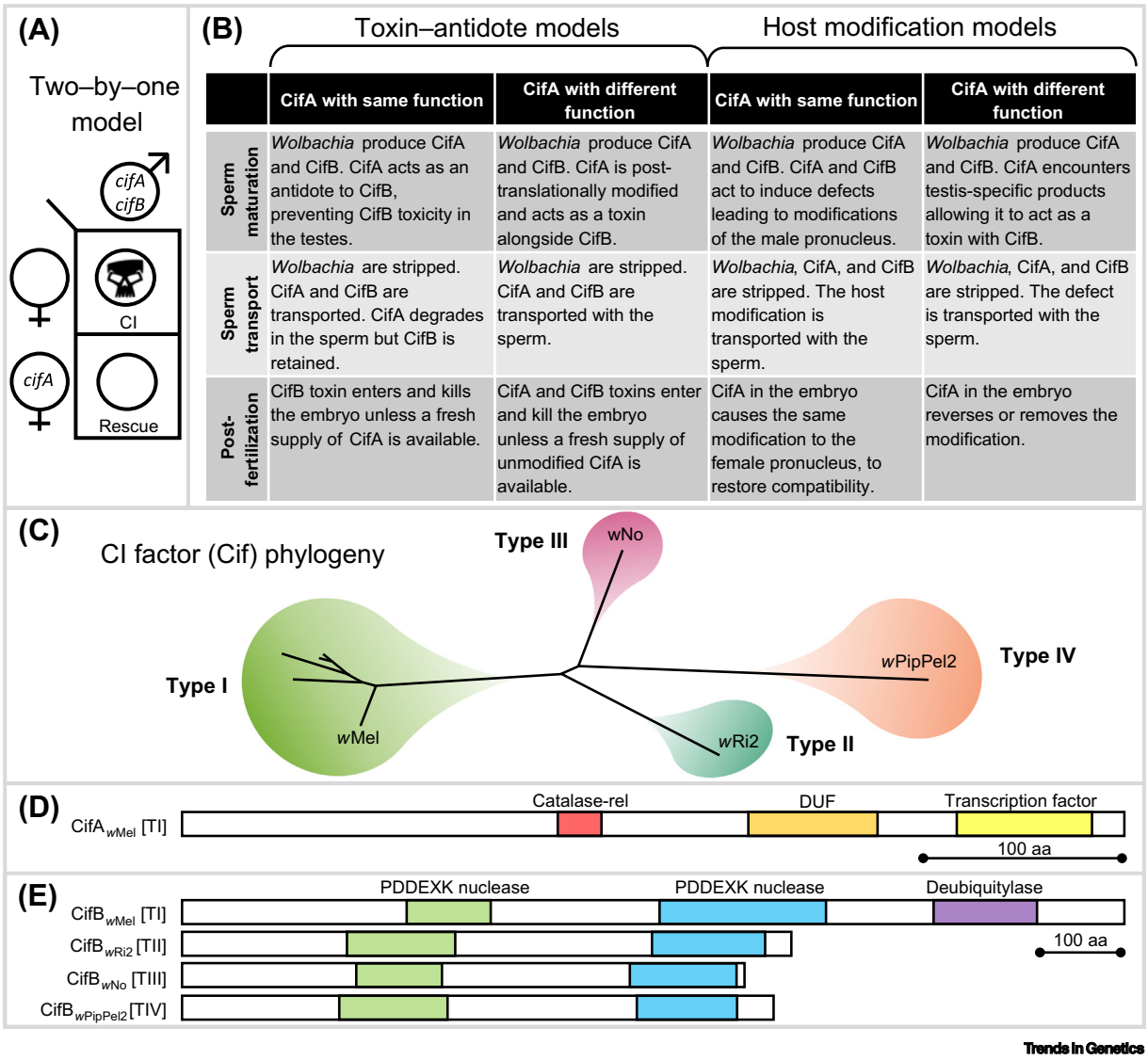

Figure 1. Models for Cytoplasmic Incompatibility (CI), Phylogenetics, and Annotated Cif Protein Architecture. (A) The two-by-one model posits that *cifA* and *cifB* expression in males causes CI that can be rescued if *cifA* is expressed in the female/embryo. (B) This genetic model remains conservatively agnostic towards the functional characterization of the CI genes, encompassing multiple toxin–antidote (TA) models and models wherein the host is modified but the Cif products are not transferred from the male to the embryo. We highlight here two examples of TA and host modification models wherein CifA may have the same function in both the testes and embryo, or different functions. The functional description of the models is separated into what takes place during sperm maturation (before individualization), during sperm transport, and after fertilization of the embryo. Importantly, these models do not represent a comprehensive set of possibilities. For example, the ability of CifA ability to act in both instances could be the result of differential localization, post-translational modifications, or comparable functions in both the testes and embryo [3]. (C) The phylogeny of Cif proteins (adapted from [8]) reveals at least four monophyletic clades (types I–IV). Representative alleles are labeled for each clade. (D) Architecture for the CifA protein of *w*Mel in the type I clade is shown with previously annotated domains. (E) Representative CifB proteins from types I–IV clades are shown. All clades have two putative PDDEXK nuclease domains based on structural homology, and type I also has a deubiquitylase domain. Amino acid scale bars are shown. White spaces in protein schematics are unannotated regions. Abbreviations: Catalase-rel, catalase-related; DUF, domain of unknown function.

Similar Gene Names for Similar Functions

It would be inaccurate to describe one gene based on the phenotype of another gene unless the genes 'govern related functions' (https://jb.asm.org/sites/default/files/additional-assets/JB-ITA.pdf) or are in an operon. Because CifA has three predicted domains that are completely unrelated to deubiquitylase activity (Figure 1D) [8], does not influence the deubiquitylase activity of CifB [3], and functions independently to rescue CI [3], it

should not be designated as a Cid. Although the Cid nomenclature is based on the operon claim [3], the evidence for the operon hypothesis, as explained above, is insufficient and should not be applied to the CI gene nomenclature. The same issue applies to CinA, which does not have nuclease annotations or confirmed phenotypes. Although further research is necessary, the *cif* nomenclature is based on current evidence that identifies both genes as being CI factors, and is therefore accurate because it makes no premature claims about protein functions across the diversity of alleles in the phylogeny (Figure 1C).

Concluding Remarks

We conclude that, first, there is no evidence for transfer of CifB protein (the presumed toxin of the TA model) from males to the embryo. For this reason, the essential premise of the proposed TA model is unfounded. Second, evidence for the operon hypothesis is equivocal, and consequently it remains unclear whether CI induction and rescue stem from an operon system. Third, the proposed *cid*, *cin*, and *cnd* gene nomenclature has several weaknesses that are solved with the *cif* nomenclature.

[1]Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA
[2]Vanderbilt Microbiome Initiative, Vanderbilt University, Nashville, TN 37235, USA
[3]Department of Biology, Lunds Universitet, Sölvegatan 37, 223 62, Lund, Sweden
[4]Hawkesbury Institute for the Environment, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia
[5]School of Environment and Science, Griffith University, Brisbane, QLD, 4111, Australia
[6]Department of Pathology, Microbiology, and Immunology, Vanderbilt University, Nashville, TN 37235, USA
[7]Vanderbilt Institute of Infection, Immunology and Inflammation, Vanderbilt University, Nashville, TN 37235, USA

*Correspondence:
shropxp@gmail.com (J.D. Shropshire) and
s.bordenstein@vanderbilt.edu (S.R. Bordenstein).

https://doi.org/10.1016/j.tig.2019.03.004

© 2019 Elsevier Ltd. All rights reserved.

References

1. Beckmann, J.F. *et al.* (2017) A *Wolbachia* deubiquitylating enzyme induces cytoplasmic incompatibility. *Nat. Microbiol.* 2, 17007
2. LePage, D.P. *et al.* (2017) Prophage WO genes recapitulate and enhance *Wolbachia*-induced cytoplasmic incompatibility. *Nature* 543, 243–247
3. Shropshire, J.D. *et al.* (2018) One prophage WO gene rescues cytoplasmic incompatibility in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U. S. A.* 115, 4987–4991
4. Beckmann, J.F. *et al.* (2019) The toxin–antidote model of cytoplasmic incompatibility: genetics and evolutionary implications. *Trends Genet.* 35, 175–185
5. Beckmann, J.F. and Fallon, A.M. (2013) Detection of the *Wolbachia* protein WPIP0282 in mosquito spermathecae: implications for cytoplasmic incompatibility. *Insect Biochem. Mol. Biol.* 43, 867–878
6. Jacob, F. *et al.* (1960) The operon: a group of genes with expression coordinated by an operator. *C. R. Biol.* 250, 1727–1729
7. Gutzwiller, F. *et al.* (2015) Dynamics of *Wolbachia pipientis* gene expression across the *Drosophila melanogaster* life cycle. *G3* 5, 2843–2856
8. Lindsey, A. *et al.* (2018) Evolutionary genetics of cytoplasmic incompatibility genes cifA and cifB in prophage WO of *Wolbachia*. *Genome Biol. Evol.* 10, 434–451
9. Demerec, M. *et al.* (1966) A proposal for a uniform nomenclature in bacterial genetics. *Genetics* 54, 61–76

Letter

Caution Does Not Preclude Predictive and Testable Models of Cytoplasmic Incompatibility: A Reply to Shropshire *et al.*

John F. Beckmann,[1,7] Manon Bonneau,[2,7] Hongli Chen,[3] Mark Hochstrasser,[3] Denis Poinsot,[4] Hervé Merçot,[5] Mylène Weill,[2] Mathieu Sicard,[2,*] and Sylvain Charlat[6,*]

Scientists often face a dilemma: should they produce explicit, predictive models to explain a body of incomplete data, at the risk of missing some critical aspects, or should they accumulate additional observations, allowing more objective and realistic models to emerge. There is a genuine trade-off between these two positions, which tend to be given different weights by different scientific disciplines, from quantum physics to anthropology. The comments of Shropshire *et al.* [1], who we thank for having given attention to our recent Opinion paper [2], illustrate that, in the fields of molecular and evolutionary genetics, there are also different views on where one should stand with respect to this trade-off. Our colleagues argue that caution should prevent us from stating that cytoplasmic incompatibility (CI) induction and rescue most likely stem from a toxin–antidote (TA) system encoded by *Wolbachia* endosymbionts. We can only agree that caution is always advisable. However, the understanding of CI, with its long theoretical and empirical history, has, in our view, come to a stage where explicit and testable models can and should be formulated.

Let us first summarize the list of predictions and empirical data supporting our claim [2]. The TA model was first proposed as a theoretical possibility. It was later evaluated in light of a variety of empirical observations and found to be more parsimonious and flexible than other available explanations, some of which were similar to the models proposed by our commentators in their Figure 1, where a direct interaction between toxin and antidote factors is not predicted [1]. A potential *Wolbachia* CI gene was then identified in infected *Culex pipiens* mosquitoes through sperm proteomics, and this gene, later named *cidA*, happened to occur right next to another conserved gene in CI-inducing *Wolbachia* strains; this striking synteny was consistent with a putative TA-like genetic structure [3]. The two genes were subsequently found to encode proteins acting in a typical TA fashion in yeast that, importantly, were also shown to form a protein complex [4]. One of the two genes (the first in the putative operon, as is typical in TA systems) was confirmed to act as a necessary and sufficient antidote to CI in *Drosophila* embryos [5].

Arguably, some observations did not match the initial TA predictions: (i) the two gene products both appear to be required for CI induction in *Drosophila* [4,6]; and