

CAN MULTIPLE COPIES MAKE NUCLEAR-LOCUS CYTOPLASMIC INCOMPATIBILITY (NLCI) GENE DRIVE FEASIBLE?

Jason L. Rasgon

The W. Harry Feinstone Department of Molecular Microbiology and Immunology, Johns Hopkins Malaria Research Institute, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205 USA

Strategies to control vector-borne diseases by release of transgenic mosquitoes require transgene spread to high frequency in populations. Due to lower expected fitness of transgenic individuals, transgenes must be actively and rapidly driven into populations in spite of fitness costs. One potential gene drive mechanism is Cytoplasmic Incompatibility (CI), caused by the bacterial endosymbiont *Wolbachia*. CI causes matings between infected males and uninfected females to be sterile, while the reciprocal mating is fertile. In a mixed population, CI confers a reproductive advantage to infected females and can allow *Wolbachia* infection to spread rapidly to high frequency in the population. Thus, linked transgenes either inserted into the *Wolbachia* genome or carried on a separate maternally-inherited construct will theoretically hitchhike into the population along with the symbionts.

One important caveat of this strategy is that transgenes must be maternally inherited with 100% fidelity. It may not be possible to insert functional transgene copies into the *Wolbachia* genome, and there are no current maternally-inherited constructs that satisfy the perfect transmission requirement. Even slightly imperfect maternal transgene inheritance will result in disassociation between the transgene and the *Wolbachia* driver, and will result in elimination of the transgene from the population.

To get around this obstacle, some have suggested that if the *Wolbachia* genes responsible for CI were identified and cloned, they could be inserted into the host nuclear chromosomes and spread in a manner similar to an under-dominant trait. If tightly linked to the anti-parasite gene, this would negate the need for a perfectly-inherited maternal construct. However, this potential advantage comes at a cost. The predicted efficiency of a nuclear-locus CI (NLCI) gene drive system is much lower than maternally-inherited CI (MICI) gene drive, due to large reductions in transgene frequency caused by CI expression in the initial generations following a release. The number of transgenic insects that must be released (Introduction Threshold) and the time for the transgene to spread into the population are much greater for NLCI versus MICI drive.

One potential method to improve the efficiency of NLCI drive is to place multiple independently-assorting (unlinked) copies of the construct in the insect genome. It has been suggested that assortment of multiple transgene copies would allow the transgene introduction threshold to be more easily exceeded, and would increase the speed of transgene spread into the population. A model was constructed to explore the theoretical feasibility of this multi-locus strategy, simulating up to 4 unlinked loci (8 total copies/genome). We assumed that CI expression was dominant; i.e., matings between wild-type females and any type of transgenic male were incompatible, while matings between any type of transgenic female and transgenic male were compatible (regardless of the number of inserts carried). Wild-type males were compatible with all females. CI-induced sterility ranged from 0 ($H = 1.0$) to complete ($H = 0$). Fitness costs of transgene copies were multiplicative, and ranged from neutral ($F = 1.0$) to 10%/insert ($F = 0.9$). In addition, we assumed that (1) insects were diploid, (2) CI modification and rescue functions were encoded by the same transgene construct, (3) random mating and (4) discrete generations.

We simulated releases of mosquitoes homozygous for the transgene at 1, 2, 3, or 4 loci (initially carrying 2, 4, 6, or 8 transgene copies, respectively). In contrast to MICI, introduction thresholds are lower for NLCI when CI is weaker rather than stronger (as long as $H < 1$). If constructs induce a fitness cost, there is an equilibrium value for H ($0 > H > 1$) that results in the lowest introduction threshold. Equilibrium prevalence levels follow a simpler pattern, where infection prevalence reaches a higher level as a function of lower fitness costs and stronger CI. For all fitness levels, speed of spread is slower for weaker CI values.

If fitness costs are low ($F > 0.99$), placing CI genes on multiple constructs results in only negligible increases in the efficiency of NLCI gene drive. Even in the case of small-to-no fitness costs, it is unlikely that the negligible increase in efficiency gained warrants the technical difficulties involved in creating a multi-locus CI mosquito strain. If fitness costs are moderate to high ($F < 0.99$), the multi-copy strategy makes NLCI efficiency worse due to the greater negative fitness effects experienced by released insects. This decrease in efficiency results in higher introduction thresholds, lower equilibrium prevalence levels and slower transgene spread. This pattern holds true across a wide range of values for F and H . Increasing the number of loci will not improve the situation, as the 4 loci simulated here are already past the point of diminishing returns. These results indicate that insertion of CI genes into the mosquito nuclear genome is unlikely to be an effective gene drive system for transgenic control of vector-borne diseases, and that improving the theoretical efficiency of such a gene drive system will be very difficult.