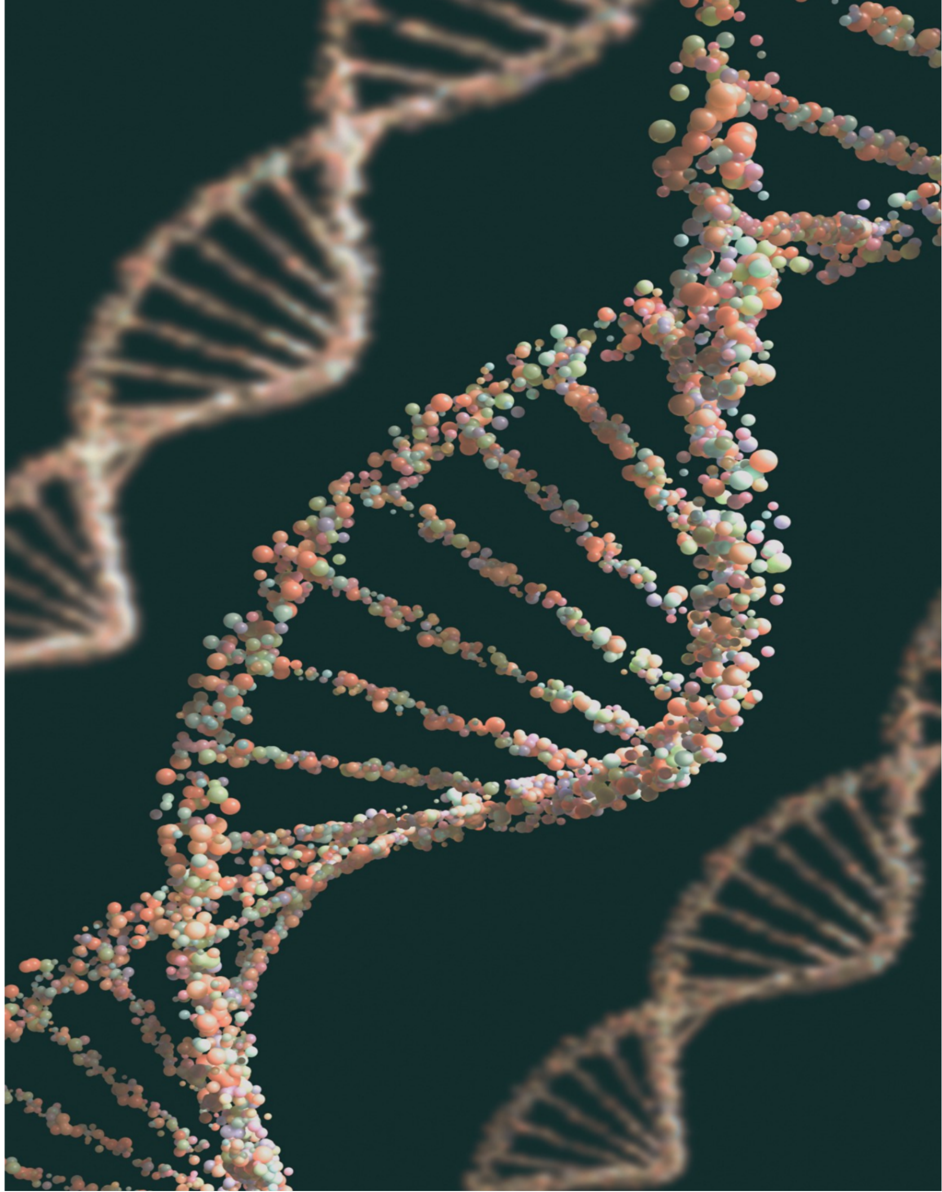
New approaches for the area-wide genetic control of insect populations

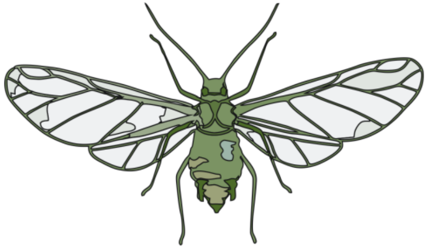
nikolai.windbichler@imperial.ac.uk

Imperial College London

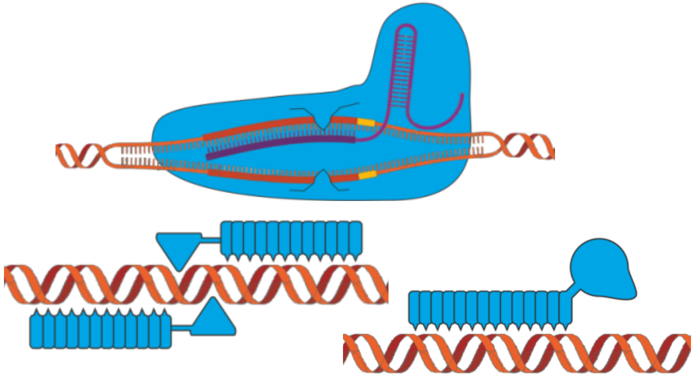


genome engineering to control insects pests and vectors of disease

Agricultural pests
15% of crops worldwide
are currently lost to insects



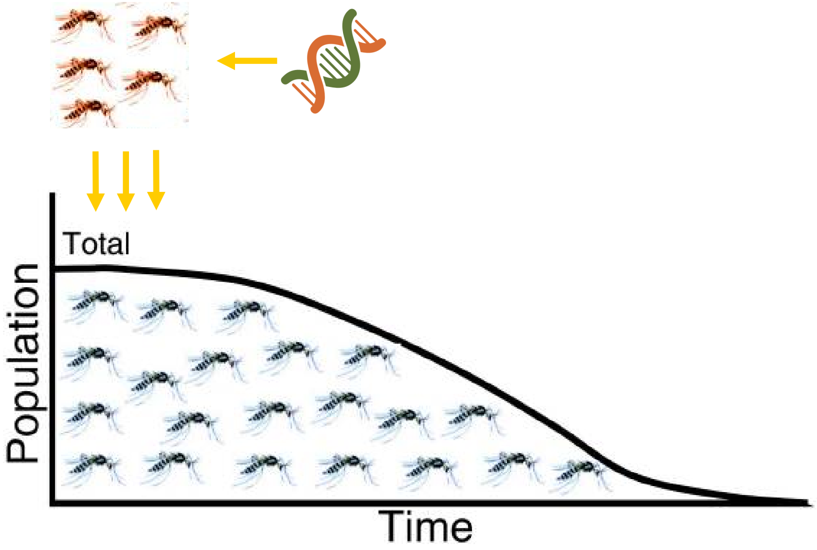
Mosquito borne disease
500 million malaria cases
>1 million deaths per year
20 million dengue cases
each year in >100 countries



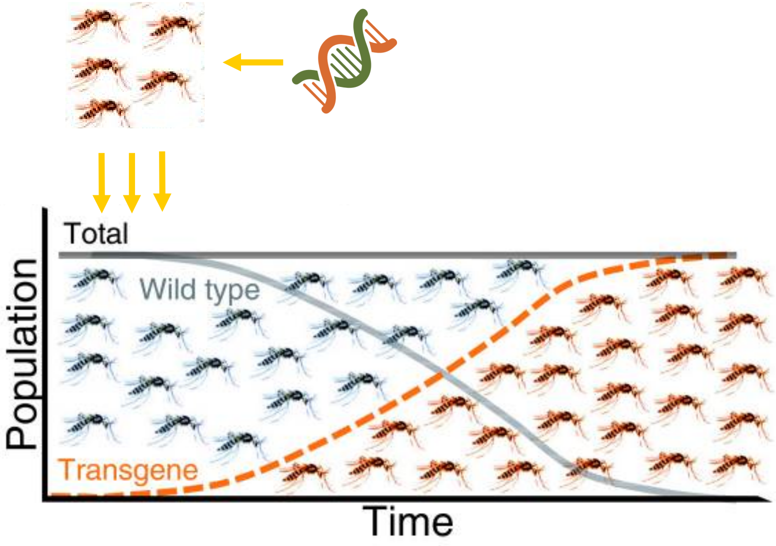
Mutate, tag, modify, activate, silence, replace any gene
TALEs, CRISPRe, CRISPRa...

genetic control

genetic control



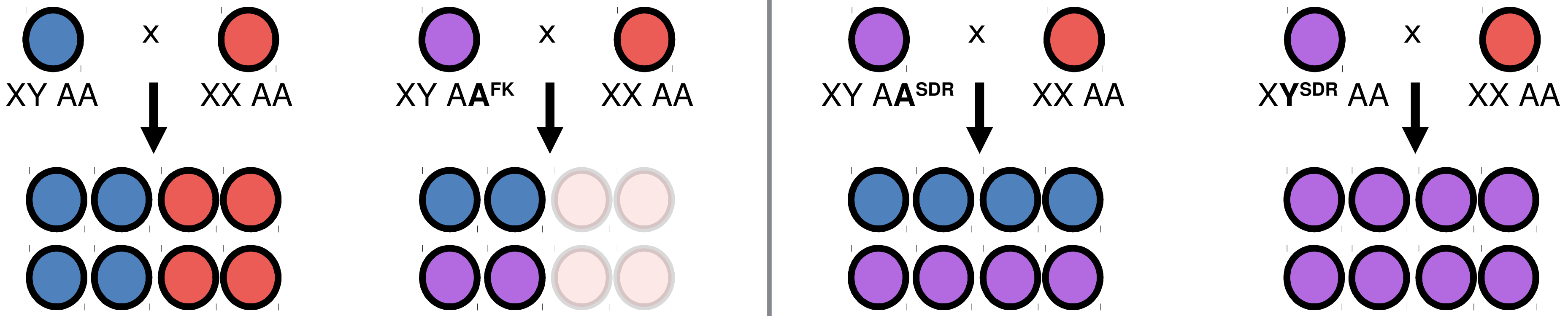
Population suppression



Population replacement

How many do we have to release?
Classic interventions e.g. sterile insect technique → 10-100x males released
New genetic interventions rely on biological replication → Release few: achieve a large effect

what is sex ratio distortion? (SRD)



Normal

50% males

100% fertility

Female Killer (FK)

100% males
50% inherit FK

50% fertility

SRD

100% males
50% inherit SRD

100% fertility

driving SRD

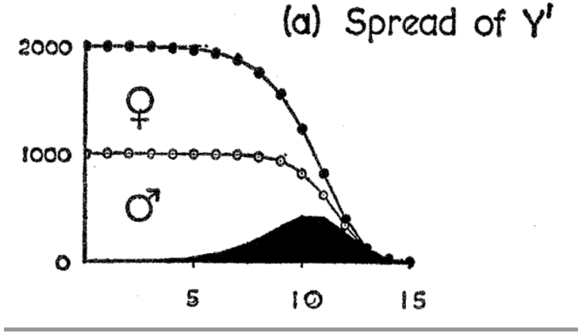
100% males
100% inherit SRD

100% fertility

sex ratio distortion as strategy for eliminating pest populations

”Suppose the Y chromosome has mutated in a way which causes it always to win in the race to fertilize. A male with the Y mutant then produces nothing but sons.”

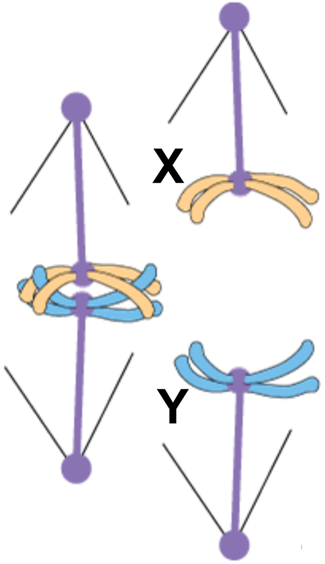
W.D. Hamilton



“[A] method of biological control [that] is in theory very powerful, since the mere seeding of a population with a few prepared males could cause its extermination.”

“Extraordinary Sex Ratios” *Science* **156**, 477 (April 28, 1967)

sex ratio distorters occur naturally in *Aedes* and *Culex* mosquitoes

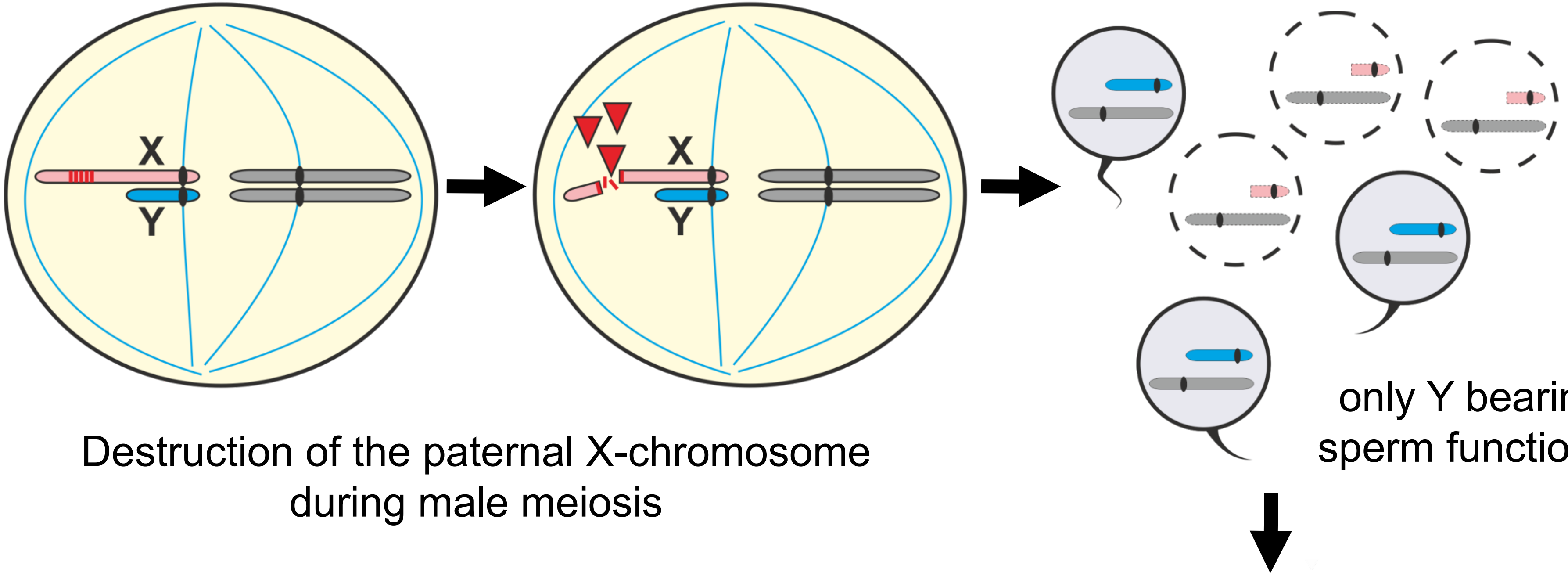


Newton et.al.
Genetica **46**, 297, 1976

sex bivalents during metaphase I

Bias towards male gamete production associated with preferential breakage of the X chromosome during male meiosis

building a synthetic sex ratio distortion system from scratch



Destruction of the paternal X-chromosome during male meiosis

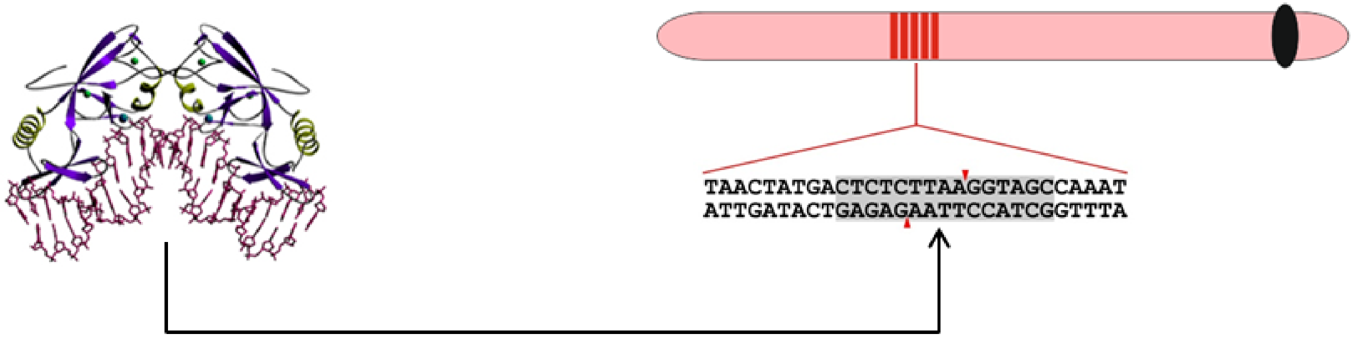
only Y bearing sperm functional

male only progeny ♂

realizing this idea in the malaria mosquito

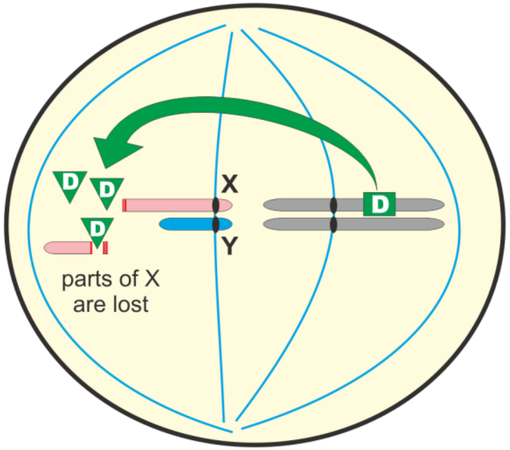
I-PpoI endonuclease

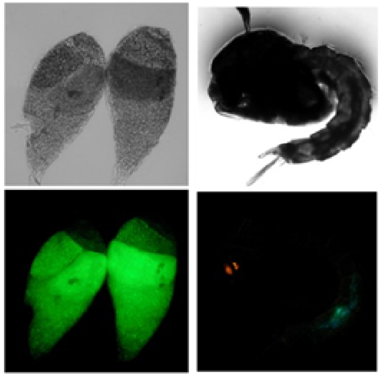
the mosquito X chromosome
contains all 400 rDNA repeats



it cuts a site present in all eukaryotic rDNA genes

engineered I-PpoI acts as a synthetic sex ratio distorter

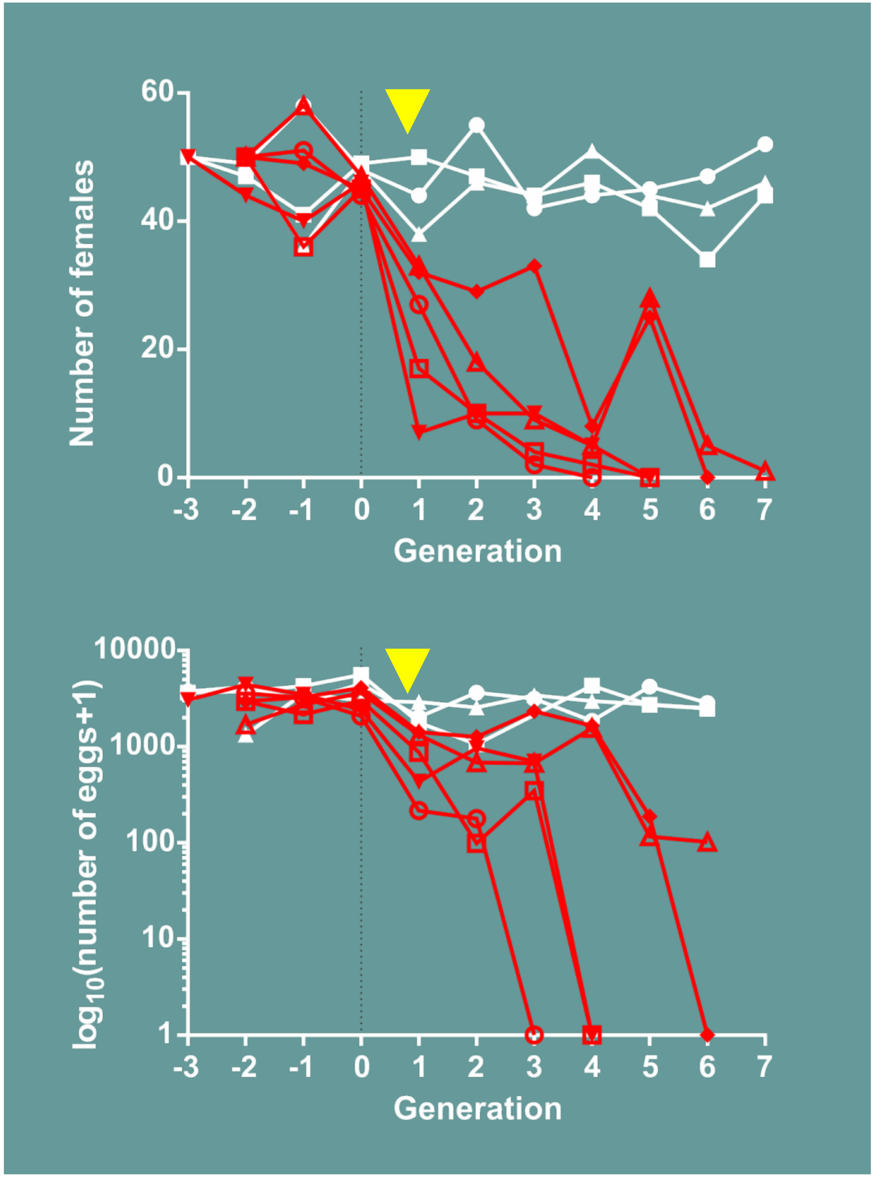




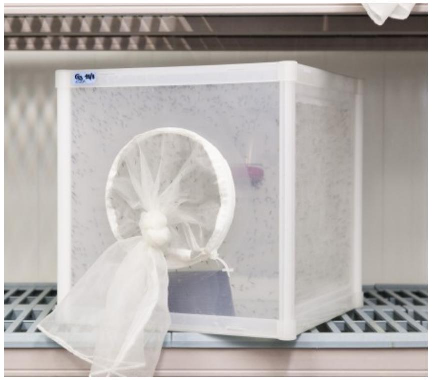
Transgenic mosquitoes expressing I-PpoI
during spermatogenesis

The best transgenic strains showed:

>95% male offspring
and no reduction in fertility



the release of distorter males eliminates caged populations of the malaria mosquito

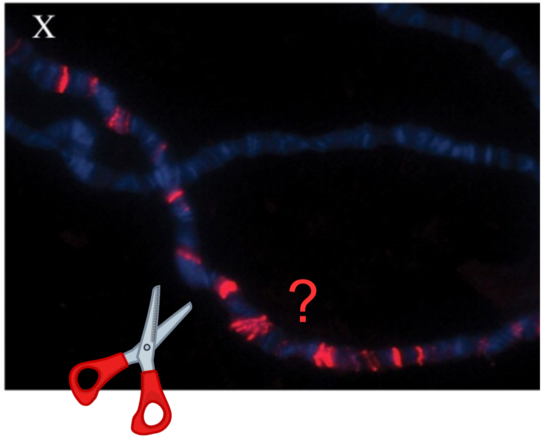


Initial population size: 100

Release rate per generation:
3X Sterile males (white cages, control)
3X hemizygous ^{gfp}111A-2 males (red cages)

how can we apply this technology to other species?

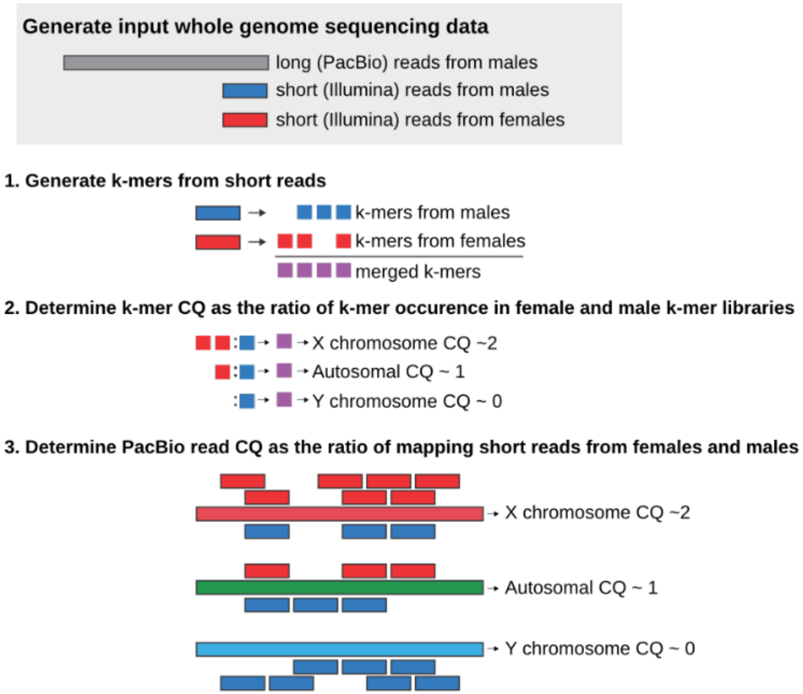
Species	Expert collaborator	Type	Host
Drosophila melanogaster	-	MO	-
Ceratitis capitata	Giuseppe Saccone	AP	Fruit
Bactrocera olea	Francesca Scolari	AP	Fruit
Anastrepha ludens	Marc Schetelig	AP	Fruit
Bactrocera dorsalis	Anna Malacrida	AP	Fruit
Drosophila suzukii	Omar Akbari	AP	Fruit
Lucilia cuprina	Max Scott	AP	Livestock
Cochliomyia hominivorax	Max Scott	AP	Livestock
Aedes albopictus	Philippos Papathanos	DV	Humans
Aedes aegypti	Omar Akbari	DV	Humans
Mayetiola destructor	-	AP	Grain



1. Need a method to identify sequences on X-chromosome that are both abundant (i.e. repeats) and X-specific

2. Need engineerable endonucleases to target and cut such X-chromosome sequences

redkmer genomic pipeline identifies X-linked repeats



Generate input whole genome sequencing data

long (PacBio) reads from males
short (Illumina) reads from males
short (Illumina) reads from females

1. Generate k-mers from short reads

k-mers from males
k-mers from females
merged k-mers

2. Determine k-mer CQ as the ratio of k-mer occurence in female and male k-mer libraries

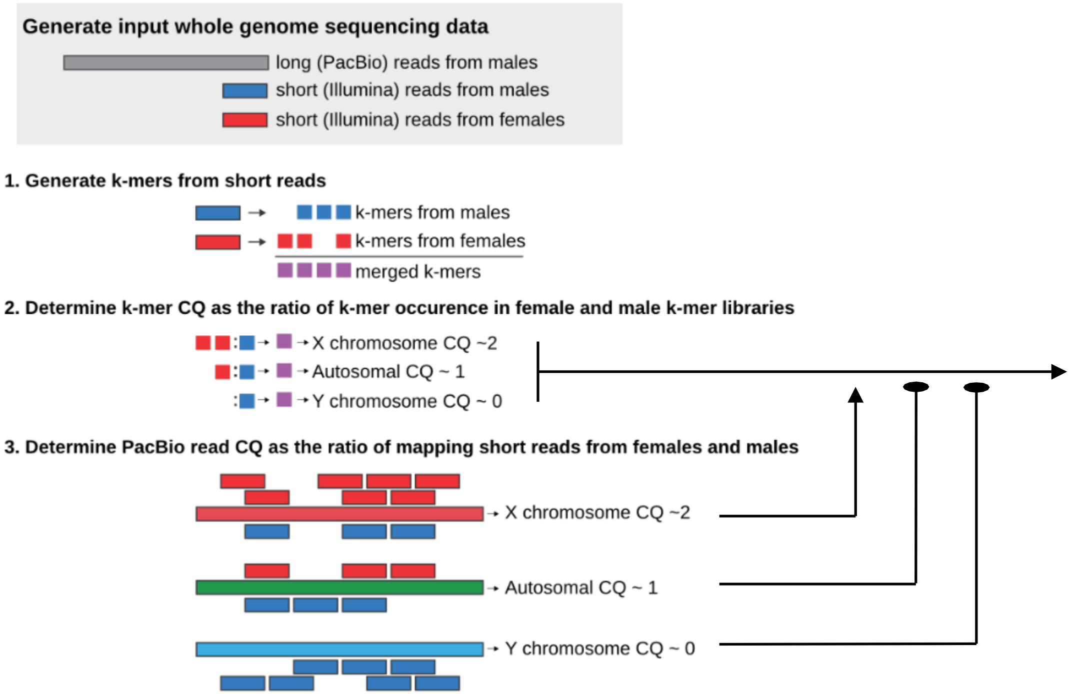
→ X chromosome CQ ~2
→ Autosomal CQ ~ 1
→ Y chromosome CQ ~ 0

3. Determine PacBio read CQ as the ratio of mapping short reads from females and males

→ X chromosome CQ ~2
→ Autosomal CQ ~ 1
→ Y chromosome CQ ~ 0

GitHub genome-traffic/redkmer-hpc

redkmer genomic pipeline identifies X-linked repeats



Generate input whole genome sequencing data

long (PacBio) reads from males

short (Illumina) reads from males

short (Illumina) reads from females

1. Generate k-mers from short reads

k-mers from males

k-mers from females

merged k-mers

2. Determine k-mer CQ as the ratio of k-mer occurence in female and male k-mer libraries

X chromosome CQ ~2

Autosomal CQ ~ 1

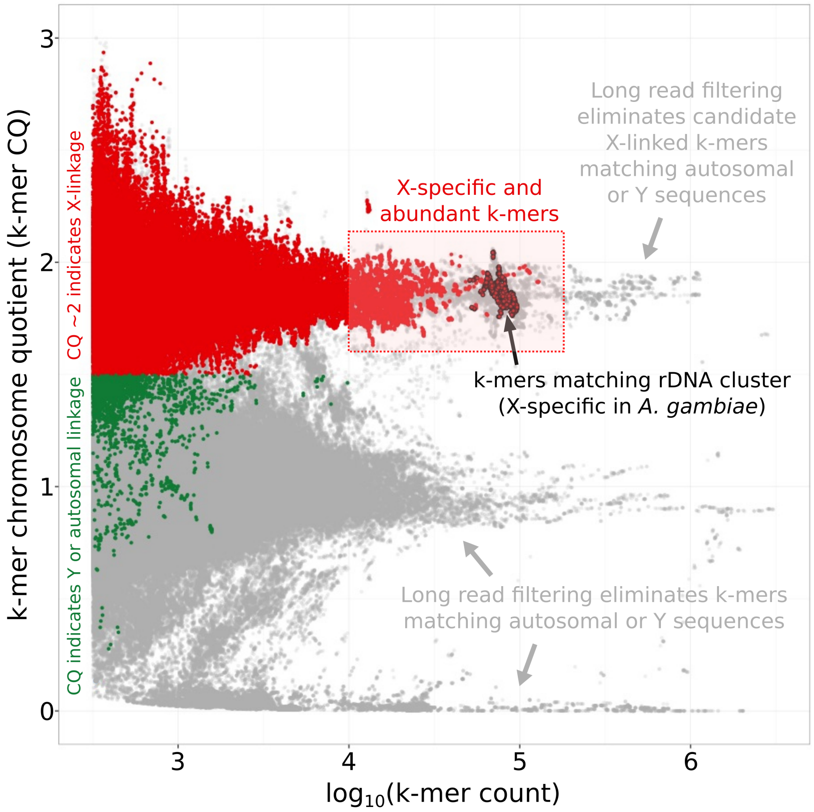
Y chromosome CQ ~ 0

3. Determine PacBio read CQ as the ratio of mapping short reads from females and males

X chromosome CQ ~2

Autosomal CQ ~ 1

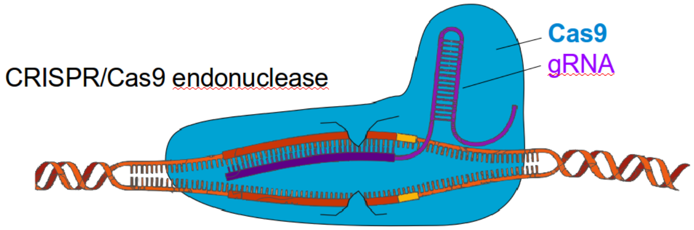
Y chromosome CQ ~ 0

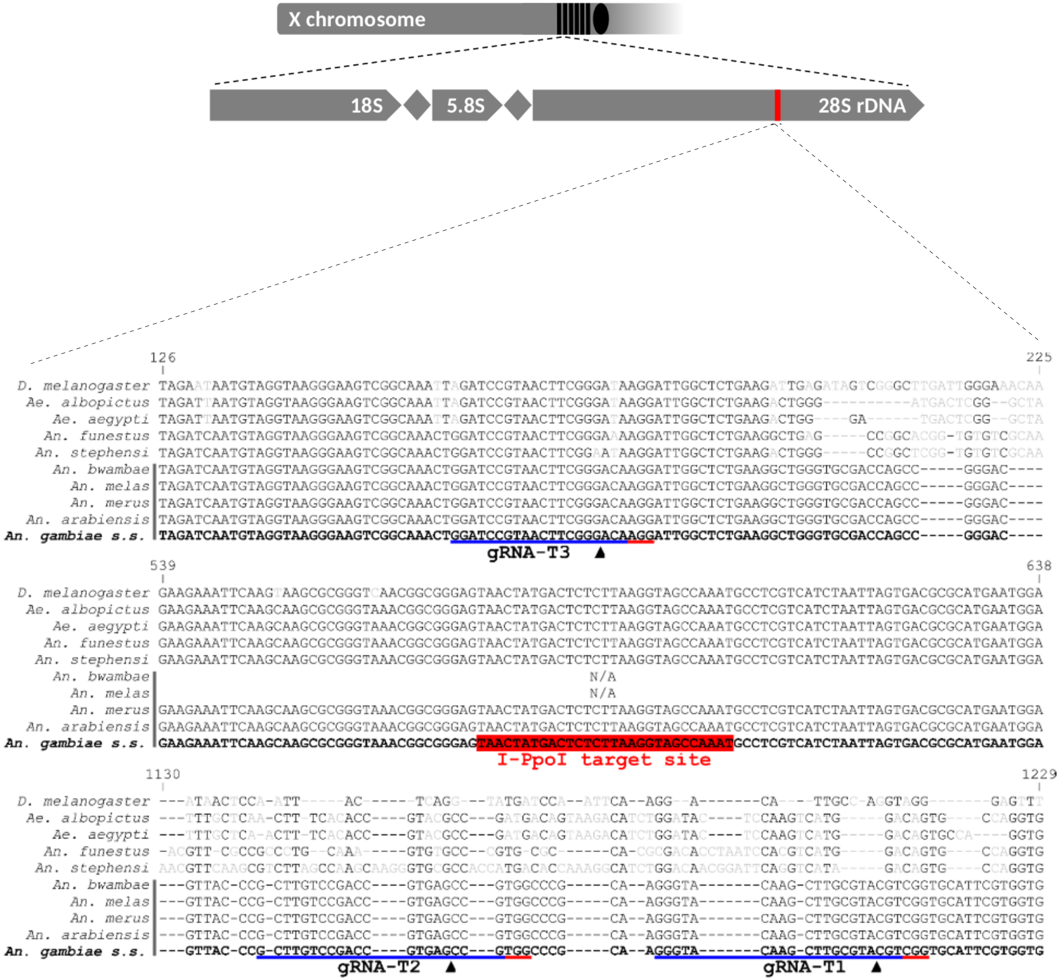


correctly identifies the abundant, X-specific and experimentally validated rDNA cluster in *Anopheles gambiae*

GitHub genome-traffic/redkmer-hpc

CRISPR/Cas9 as a reprogrammable X-shredding endonuclease

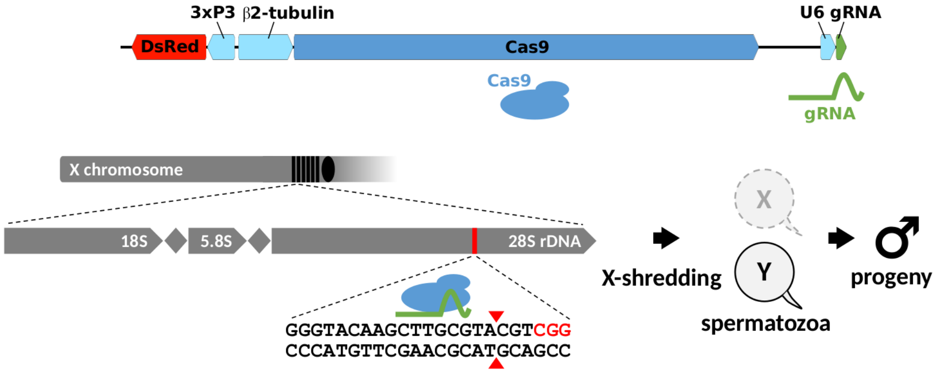


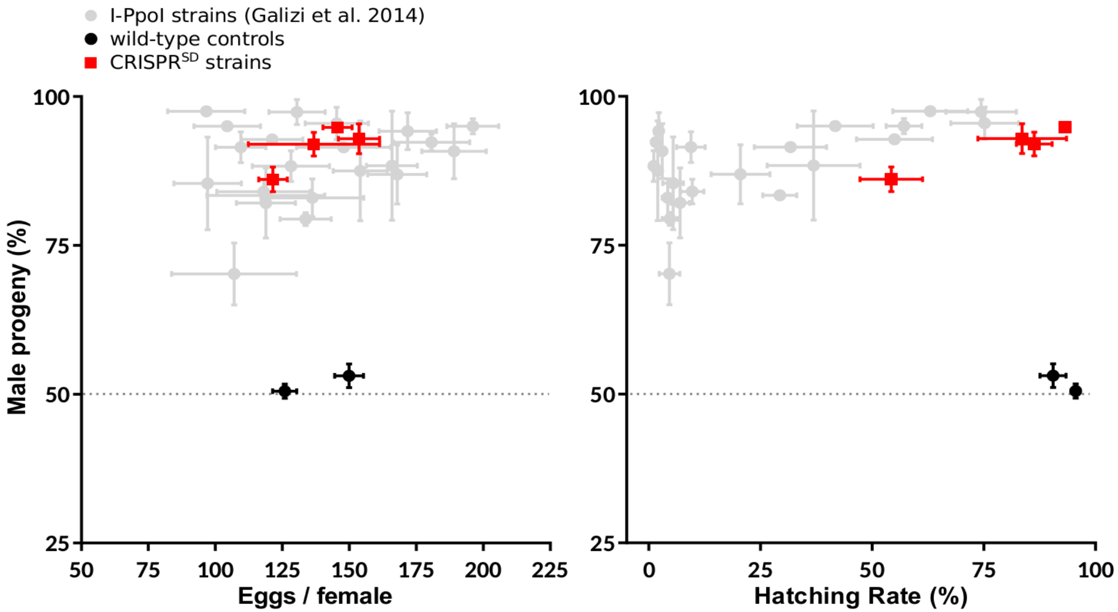


```
                 126                                                                                              225
D. melanogaster  TAGAATAATGTAGGTAAGGGAAGTCGGCAAATTAGATCCGTAACTTCGGGATAAGGATTGGCTCTGAAGATTGAGATAGTCGGGCTTGATTGGGAAACAA
  Ae. albopictus TAGATTAATGTAGGTAAGGGAAGTCGGCAAATTAGATCCGTAACTTCGGGATAAGGATTGGCTCTGAAGACTGGG----------ATGACTCGG--GCTA
     Ae. aegypti TAGATTAATGTAGGTAAGGGAAGTCGGCAAATTAGATCCGTAACTTCGGGATAAGGATTGGCTCTGAAGACTGG----GA------TGACTCGG--GCTA
    An. funestus TAGATCAATGTAGGTAAGGGAAGTCGGCAAACTGGATCCGTAACTTCGGGAAAAGGATTGGCTCTGAAGGCTGAG-----CCGGCACGG-TGTGTCGCAA
   An. stephensi TAGATCAATGTAGGTAAGGGAAGTCGGCAAACTGGATCCGTAACTTCGGAATAAGGATTGGCTCTGAAGACTGGG-----CCGGCTCGG-TGTGTCGCAA
     An. bwambae TAGATCAATGTAGGTAAGGGAAGTCGGCAAACTGGATCCGTAACTTCGGGACAAGGATTGGCTCTGAAGGCTGGGTGCGACCAGCC-----GGGAC----
       An. melas TAGATCAATGTAGGTAAGGGAAGTCGGCAAACTGGATCCGTAACTTCGGGACAAGGATTGGCTCTGAAGGCTGGGTGCGACCAGCC-----GGGAC----
       An. merus TAGATCAATGTAGGTAAGGGAAGTCGGCAAACTGGATCCGTAACTTCGGGACAAGGATTGGCTCTGAAGGCTGGGTGCGACCAGCC-----GGGAC----
  An. arabiensis TAGATCAATGTAGGTAAGGGAAGTCGGCAAACTGGATCCGTAACTTCGGGACAAGGATTGGCTCTGAAGGCTGGGTGCGACCAGCC-----GGGAC----
An. gambiae s.s. TAGATCAATGTAGGTAAGGGAAGTCGGCAAACTGGATCCGTAACTTCGGGACAAGGATTGGCTCTGAAGGCTGGGTGCGACCAGCC-----GGGAC----
                                                     gRNA-T3  ▲
                 539                                                                                              638
D. melanogaster  GAAGAAATTCAAGTAAGCGCGGGTCAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGA
  Ae. albopictus GAAGAAATTCAAGCAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGA
     Ae. aegypti GAAGAAATTCAAGCAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGA
    An. funestus GAAGAAATTCAAGCAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGA
   An. stephensi GAAGAAATTCAAGCAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGA
     An. bwambae                                                   N/A
       An. melas                                                   N/A
       An. merus GAAGAAATTCAAGCAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGA
  An. arabiensis GAAGAAATTCAAGCAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGA
An. gambiae s.s. GAAGAAATTCAAGCAAGCGCGGGTAAACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAATGGA
                                                     I-PpoI target site
                1130                                                                                              1229
D. melanogaster  ---ATAACTCCA-ATT-----AC-----TCAGG--TATGATCCA--ATTCA--AGG--A-------CA---TTGCC-AGGTAGG-------GAGTTT
  Ae. albopictus ---TTTGCTCAA-CTT-TCACACC----GTACGCC---GATGACAGTAAGACATCTGGATAC----TCCAAGTCATG-----GACAGTG----CCAGGTG
     Ae. aegypti ---TTTGCTCA-ACTT-TCACACC----GTACGCC---GATGACAGTAAGACATCTGGATAC----TCCAAGTCATG-----GACAGTGCCA----GGTG
    An. funestus -ACGTT-CGCCGCCTG--CAAA-----GTGTGCC--CGTG-CGC------CA-CGCGACACCTAATCCACGTCATG-----GACAGTG----CCAGGTG
   An. stephensi AACGTTCAAGCGTCTTAGCCAAGCAAGGGTGCGCCACCATGACACCAAAGGCATCTGGACAACGGATTCAGGTCATA-----GACAGTG----CCAGGTG
     An. bwambae ---GTTAC-CCG-CTTGTCCGACC----GTGAGCC---GTGGCCCG-----CA--AGGGTA-------CAAG-CTTGCGTACGTCGGTGCATTCGTGGTG
       An. melas ---GTTAC-CCG-CTTGTCCGACC----GTGAGCC---GTGGCCCG-----CA--AGGGTA-------CAAG-CTTGCGTACGTCGGTGCATTCGTGGTG
       An. merus ---GTTAC-CCG-CTTGTCCGACC----GTGAGCC---GTGGCCCG-----CA--AGGGTA-------CAAG-CTTGCGTACGTCGGTGCATTCGTGGTG
  An. arabiensis ---GTTAC-CCG-CTTGTCCGACC----GTGAGCC---GTGGCCCG-----CA--AGGGTA-------CAAG-CTTGCGTACGTCGGTGCATTCGTGGTG
An. gambiae s.s. ---GTTAC-CCG-CTTGTCCGACC----GTGAGCC---GTGGCCCG-----CA--AGGGTA-------CAAG-CTTGCGTACGTCGGTGCATTCGTGGTG
                          gRNA-T2  ▲                                        gRNA-T1  ▲
```

CRISPR-gRNA target site, PAM sequence, ▲ Cut site

CRISPR/Cas9 as a reprogrammable X-shredding endonuclease





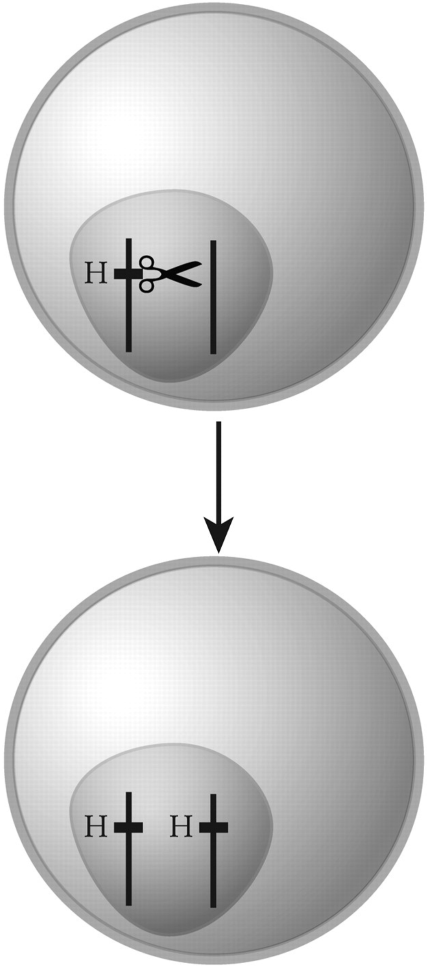
X-shredding checklist for your pest/vector species:

X/Y chromosomal system

Long read data from males
Short read data from males & females
→ identify X-specific & abundant targets

transgenesis capabilities

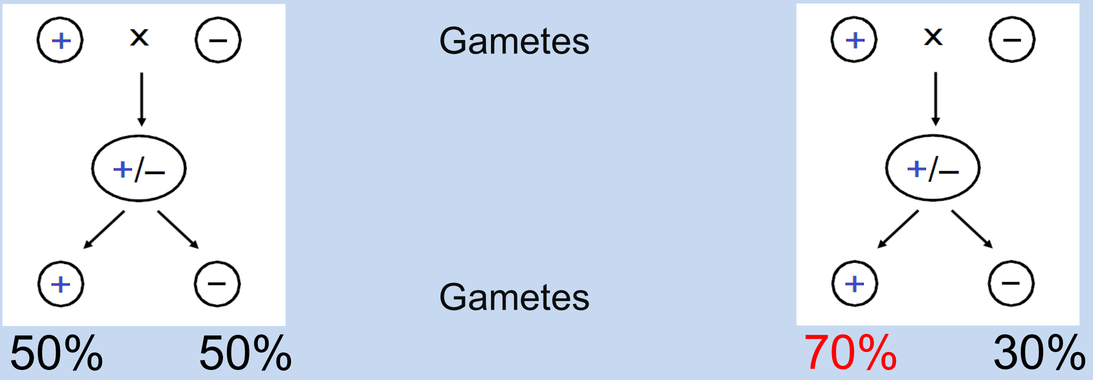
beta2 tubulin and U6 promoters for CRISPR expression during spermatogenesis gene drive of homing endonucleases



what is gene drive?

INDIVIDUAL LEVEL

Inheritance of allele (+)

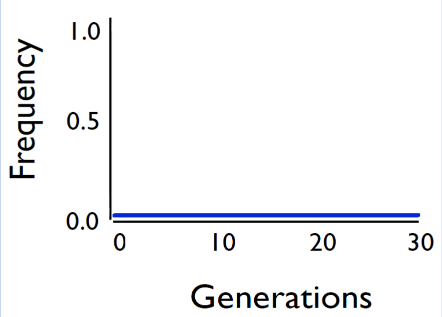


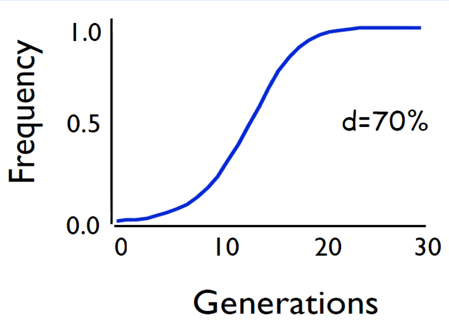
mendelian inheritance

gene drive

POPULATION LEVEL

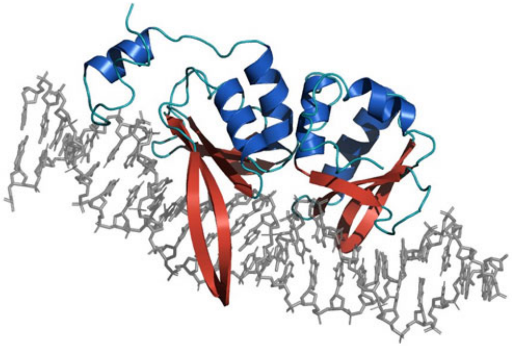
Frequency of a rare allele (+) in a population



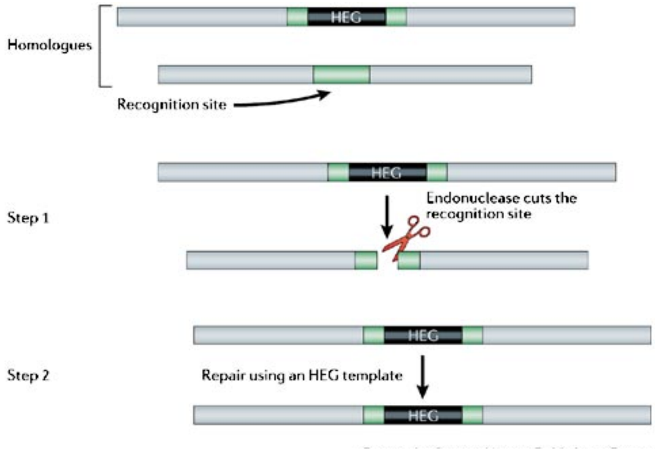


Driving genes can spread in a population even if they decrease the fitness of their host (that's why they are often called selfish genes)

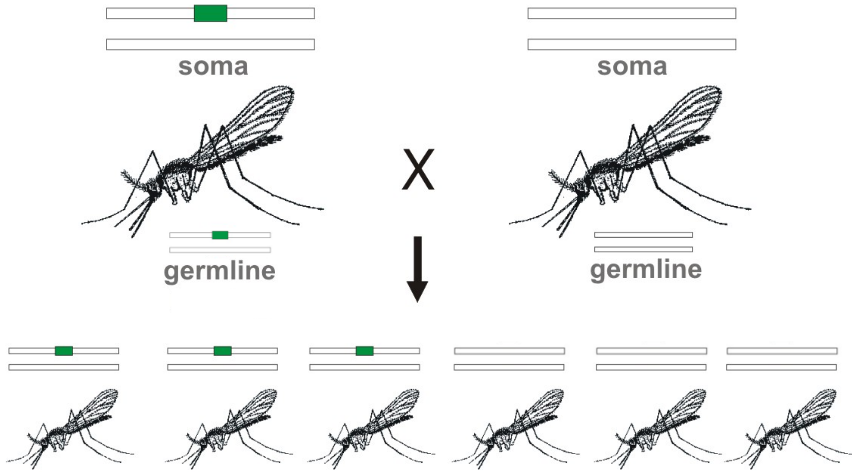
Driving (Homing) Endonuclease Genes



HEGs are highly specific DNA endonucleases
They cut DNA only at unique target sites

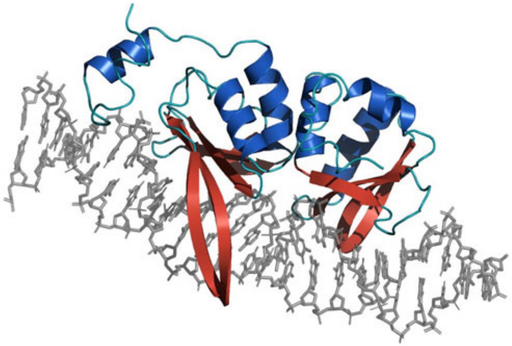


DNA breaks are repaired using the HEG+ allele as template
Thus the HEG is copied from one chromosome to another

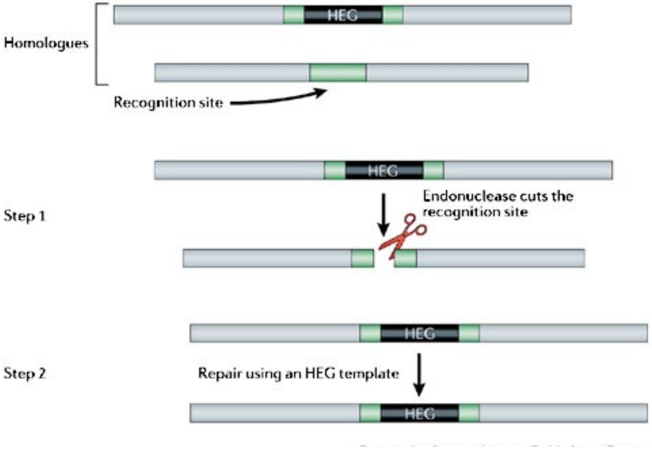


HEG transmitted to a high proportion of the progeny

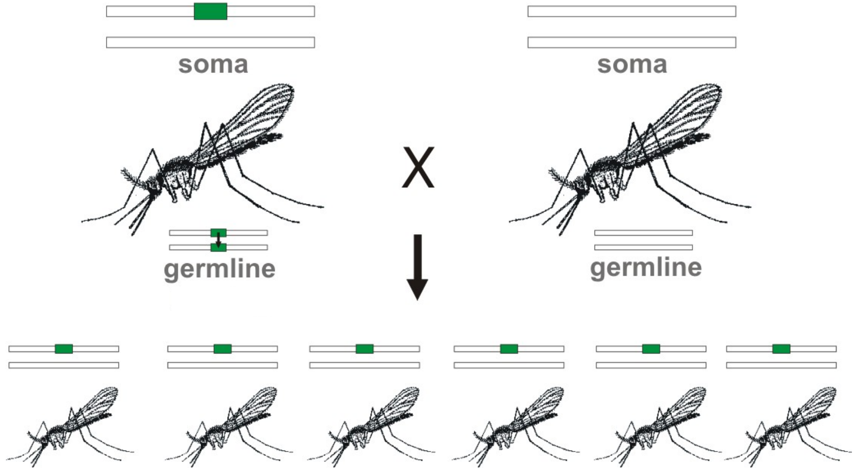
Driving (Homing) Endonuclease Genes



HEGs are highly specific DNA endonucleases
They cut DNA only at unique target sites



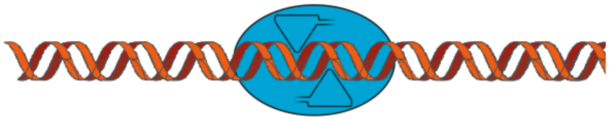
DNA breaks are repaired using the HEG+ allele as template
Thus the HEG is copied from one chromosome to another

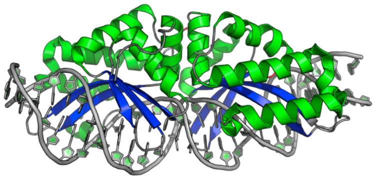


HEG transmitted to a high proportion of the progeny

Driving (Homing) Endonuclease Genes – CRISPR is a game changer

classic homing endonuclease

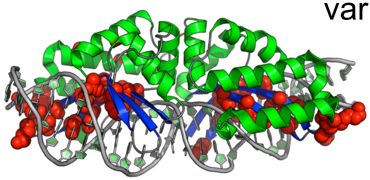




TTTCCACTTATTCAACCTTTTA

original target site

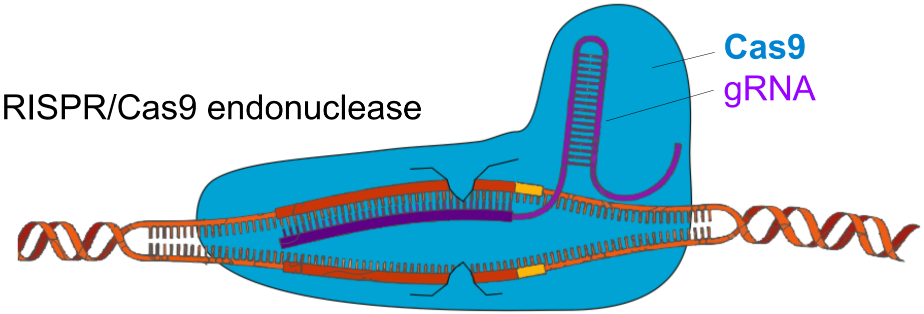
How to change specificity: Reassembly of compatible domains from huge libraries of variants. Months of work, for a large team.



CCTCCTCACTTTCTTCCTCACC

mosquito target gene AGAP007280

CRISPR/Cas9 endonuclease

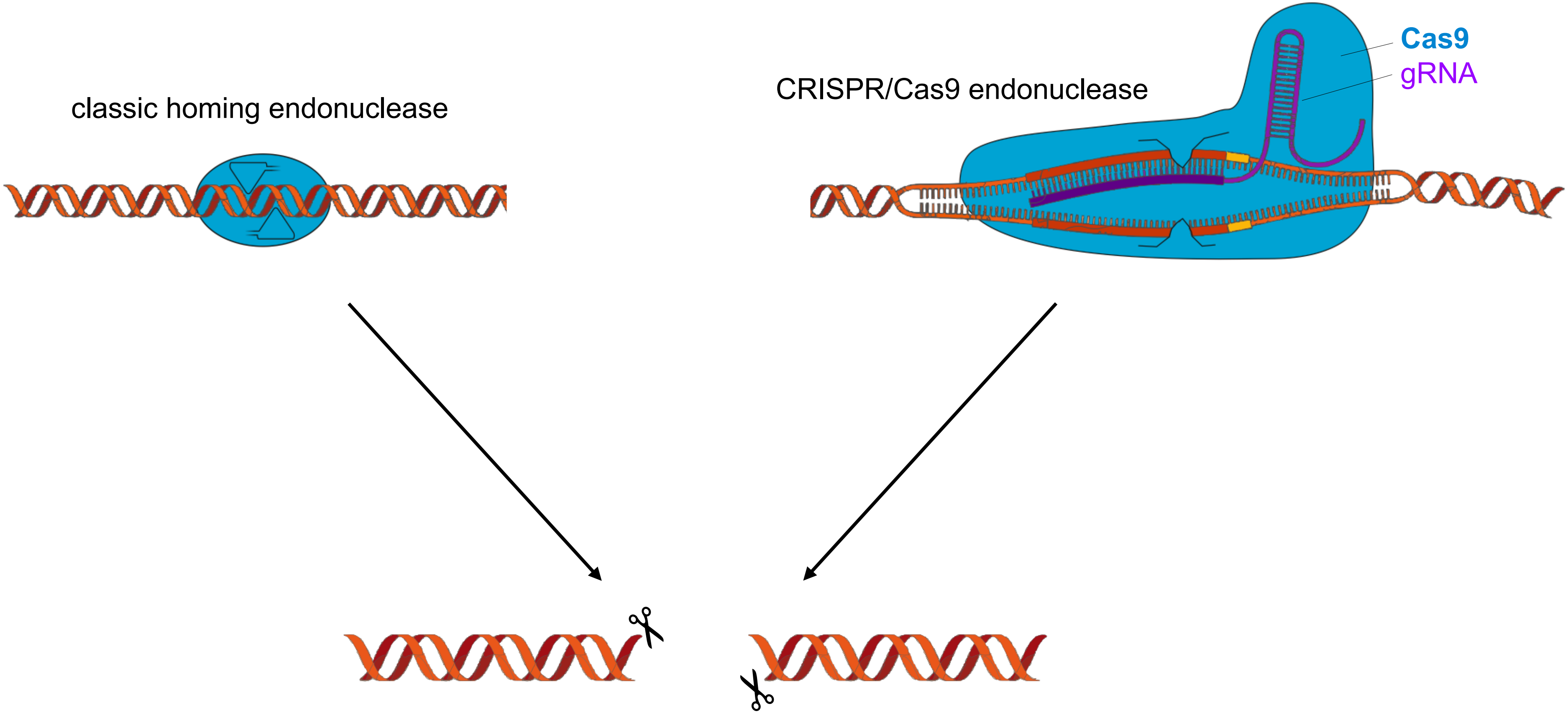


How to change specificity:
Clone a new gRNA by ordering two primers!

TGCTGGAAGAAAGTGAGGAGGA
AAACTCCTCCTCACTTTCTTCC

Primers ordered until 10:30 a.m. CET are shipped on the same day!

Driving (Homing) Endonuclease Genes – CRISPR is a game changer



classic homing endonuclease

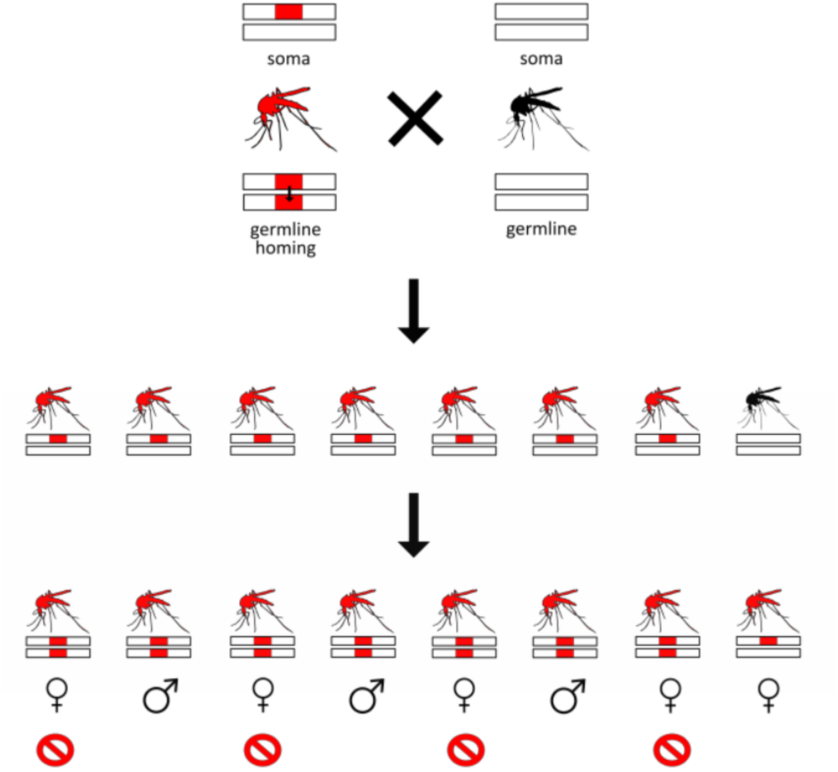
CRISPR/Cas9 endonuclease

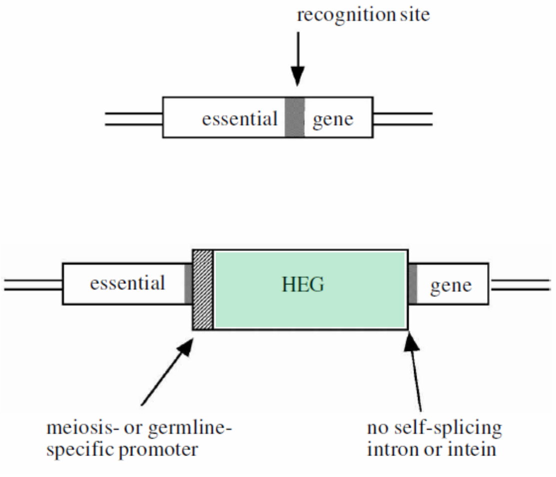
Cas9

gRNA

DNA double strand break

Population suppression using gene drive





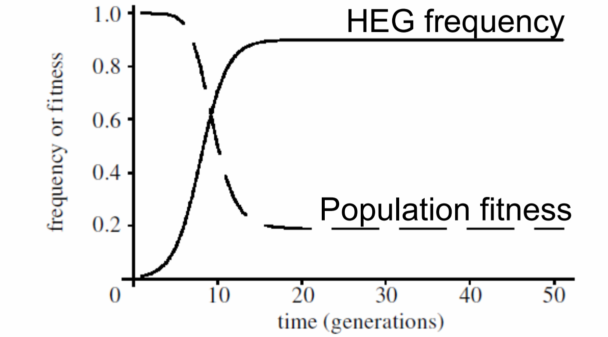
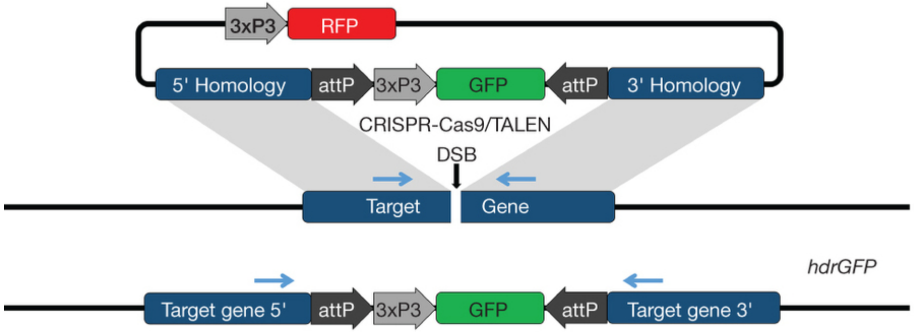


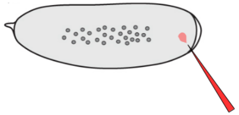
Figure 2. Frequency of the HEG (solid curve) and population mean fitness (dashed curve) assuming $e = 0.9$ and an initial release frequency of 1%. These results, and all others in the paper, are for an idealized population, from which all real populations will deviate in some way. They should, therefore, be taken as rough indications, not precise predictions.

A driving endonuclease targeting an essential female fertility gene (disruption is recessive) will increase in frequency. Once HEG individuals are common homozygotes (sterile females) are generated

knock-out of 3 female fertility genes with GFP replacement cassette



inject into embryos

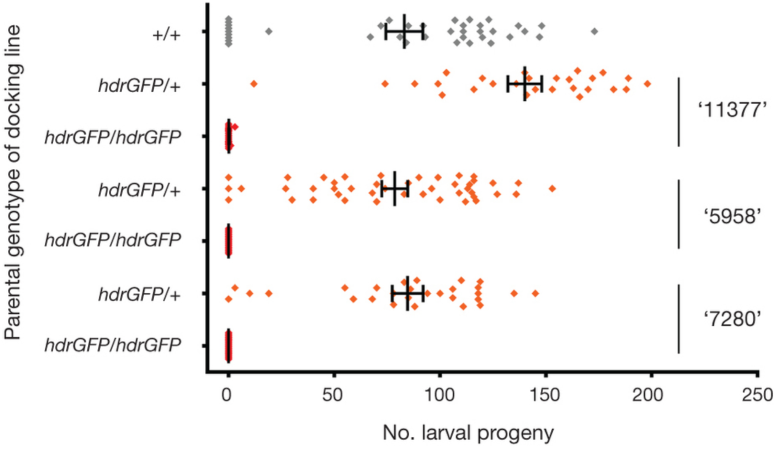


genes with high ovary expression and tissue specificity were chosen:

AGAP005958 an ortholog of Drosophila yellow-g, a haplosufficient female-fertility gene expressed in somatic follicle cells

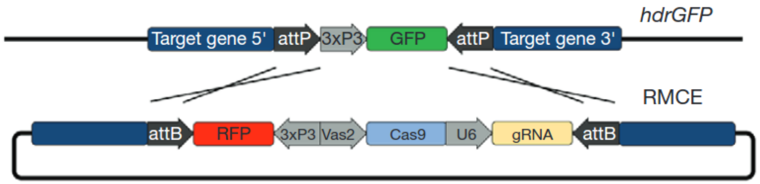
AGAP007280 an ortholog of Drosophila nudel, a haplosufficient female-fertility gene expressed in somatic follicle cells involved in dorsoventral patterning of the embryo

AGAP011377 contains a probable chitin binding domain).



homozygous sterility observed in all 3 cases

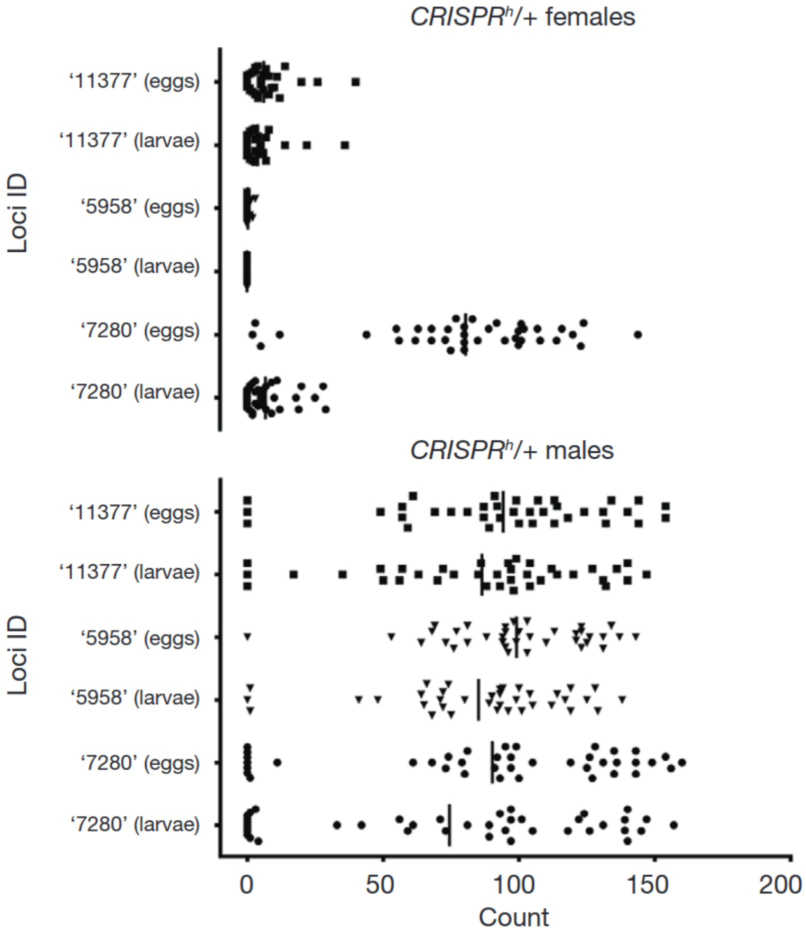
Cassette exchange introduce the driving CRISPR allele locus



RMCE using the PhiC31 integrase inserting a Cas9 gene driven by the vasa promoter and a U6 driven gRNA

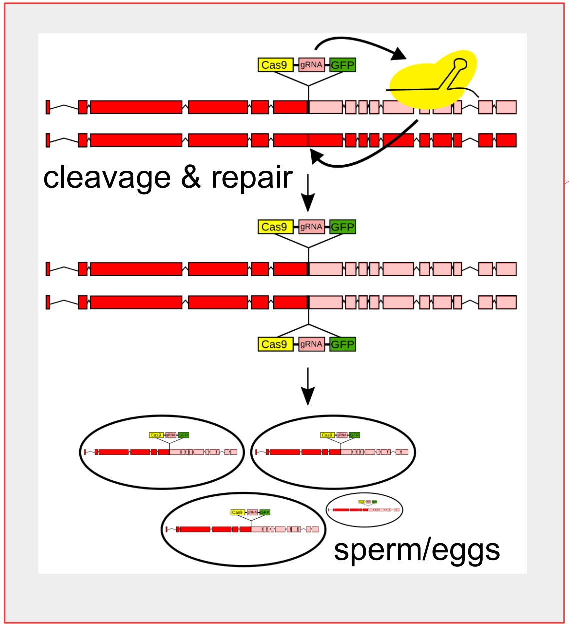
strong fertility effect in heterozygous CRISPRh/+ females observed for 2 genes

AGAP007280 shows full heterozygous fertility

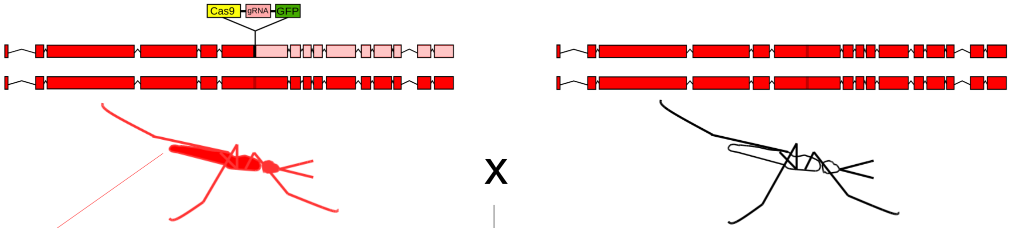


is the CRISPR construct targeting AGAP007280 homing?

germline (ovaries/testes)



soma



progeny

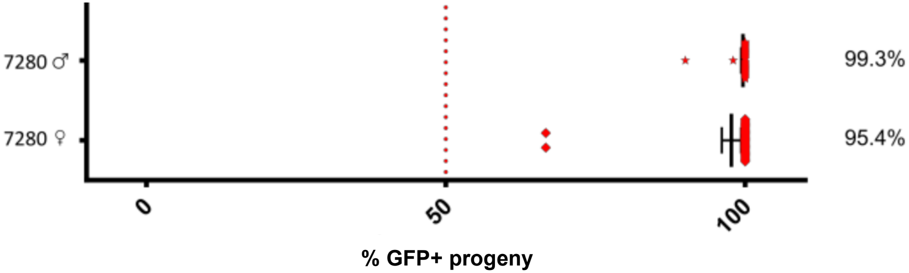
expectation if endonuclease inactive

50% GFP +
50% GFP-

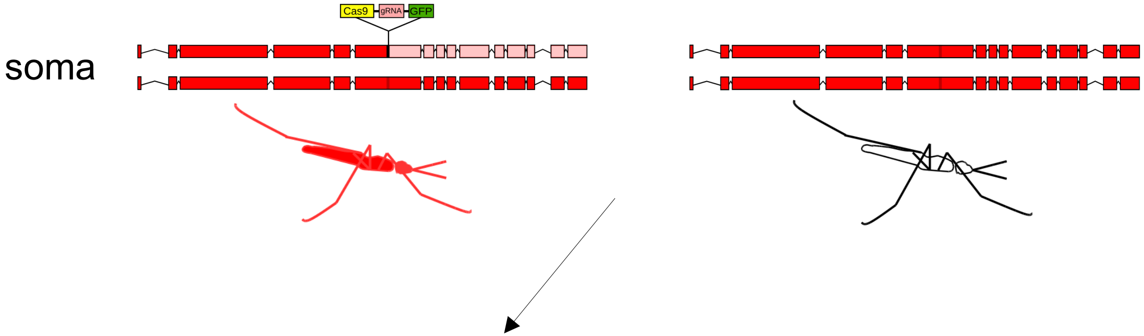
mendelian inheritance

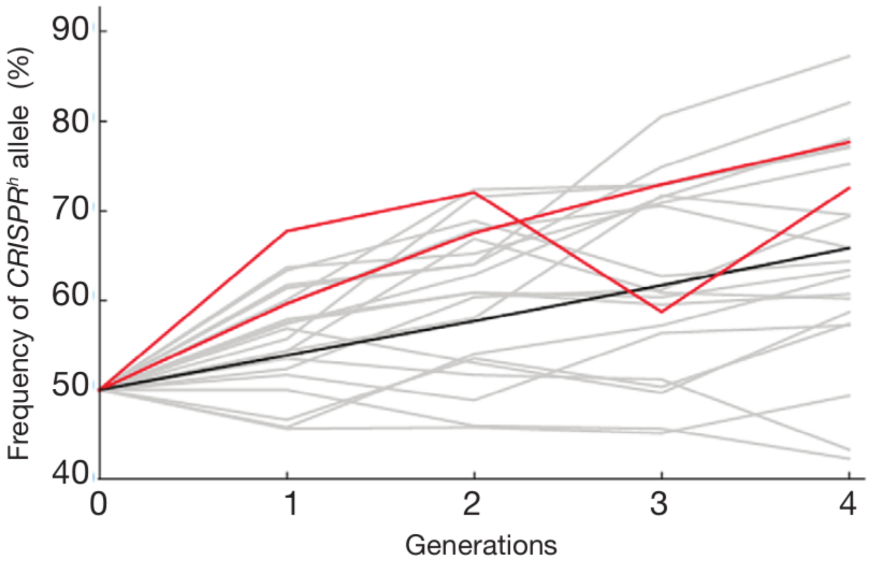
observed (AGAP007280 target gene)

gene drive!



gene drive dynamics of construct targeting 7280 in population cages





An equal number of CRISPR h/+ and WT individuals
were used to start a population of size 600
black line shows deterministic prediction
gray lines show results from 20 stochastic simulations
red lines show results from two replicate cages

Conclusions

A range of new genetic tools that could work in the field are now available or under development

→ gene drives
→ sex ratio distorters

They rely on biological amplification
→ few released insects can have
a large effect on a whole population

They can be used for both population replacement and suppression

Will they be applied?
A question of the actual and perceived safety, of efficacy and of social acceptability

Acknowledgements

Roberto Galizi
Drew Hammond
Phillipos Aris Papathanos
Barbara Fasulo
Tony Nolan
Barry Stoddard
Austin Burt
Andrea Crisanti
All members of the Crisanti
& Windbichler labs

BBSRC bioscience for the future
erc
BILL & MELINDA GATES foundation
FNIH Foundation for the National Institutes of Health