

Strategies to overcome Meiotic Sex Chromosome Inactivation
in the human malaria vector *A. gambiae*.

Dario Meacci

Department of Life Sciences
Imperial College London

A thesis submitted for the award of the degree of
Doctor of Philosophy – PhD.

August 2022

"It is not the critic who counts; not the man who points out how the strong man stumbles, or where the doer of deeds could have done them better. The credit belongs to the man who is actually in the arena, whose face is marred by dust and sweat and blood; who strives valiantly; who errs, who comes short again and again, because there is no effort without error and shortcoming; but who does actually strive to do the deeds; who knows great enthusiasms, the great devotions; who spends himself in a worthy cause; who at the best knows in the end the triumph of high achievement, and who at the worst, if he fails, at least fails while daring greatly, so that his place shall never be with those cold and timid souls who neither know victory nor defeat."

T. Roosevelt

Acknowledgements

I want to thank Prof. Andrea Crisanti and Dr Tony Nolan for the opportunities they gave me. First, hiring me as Research Technician and then giving me the possibility to pursue my desire to achieve a PhD. Thanks also to Barbara Fasulo for her guidance during the last year of my PhD

Then I want to thank Roberto and Alekos for always being available and willing to help me in any possible way. Alekos is also responsible for my passion for climbing, which I spread afterwards inside the lab. Thanks also to Fede and Chrissa for all the weird and funny chats we had.

I want to thank Matt for his friendship, the Monday football and all football-related events we attended. Chris, for all the good moments and the laughs we had together. Giulia, for always being there for me in my bright and dark moments. Drew, Antonios, and Ioanna for being fantastic friends, colleagues and flatmates. Kyros for just being himself, meaning funny, kind and generous.

A special thanks to my PhD and Technicians partners in crime Nace, Matteo, Ignacio, Anna, Nicole, Silvia, Nayomi and Mariana.

Thanks to all the Crisanti lab members for being the best team I could have hoped to find when I moved to London.

Finally, I'm grateful to have met Alberto, my eternal flatmate, and Mattia, a sincere and generous friend that was always there for me, literally holding my rope at every fall.

1	Introduction.....	12
1.1	A brief history of Malaria.....	12
1.2	Malaria today	13
1.3	Mosquito life cycle	15
1.4	Strategies for vector control	17
1.5	The successful Italian eradication and the African situation today	18
1.6	Alternative strategies for vector control	21
1.6.1	<i>Wolbachia</i> as a tool for population replacement and suppression	21
1.6.2	Incompatible Insect Technique (IIT) and Sterile Insect Technique (SIT)	22
1.6.3	The Meiotic Sex Distorter and Release of Insects Carrying a Dominant Lethal (RIDL).....	24
1.6.4	A new hope for malaria eradication: the gene drive.	25
1.7	The CRISPR/Cas9 system	28
1.7.1	A brief CRISPR chronology	29
1.8	The Homing Gene Drive.....	31
1.9	Self-limiting gene drives and anti-gene drive systems.....	33
1.10	The Y-linked meiotic Gene Drive.....	34
1.11	Gametogenesis in mosquitoes	36
1.12	Meiotic Sex Chromosome Inactivation.....	37
1.13	The Importance of Epigenetics.....	39
1.13.1	Acetylation and Methylation, two histone modifications.....	41
1.14	The gypsy retro-transposable element as a chromatin insulator	43
1.15	Activating transcription via Gal4 and catalytically defective Cas9 fused to HAT.....	45
1.16	The “Pcaf” Histone acetyltransferase.....	47
2	Results	49
2.1	Strategy 1 – Developing a Gal4-UAS system to block heterochromatin spreading	49
2.2	Strategy 2 - The design of a CRISPR system to counteract MSCI	50
2.3	The dCas9 kinetics during gametogenesis	54
2.3.1	Generation of the dCas9 transgenic lines	54
2.3.2	Assessing the dCas9 binding stability during gametogenesis	55
2.3.3	A Cas9/dCas9 competition experiment	56
2.3.4	Testing dCas9 misexpression and deposition in <i>A. gambiae</i>	61
2.3.5	The engineering and testing of the dCas9-mCherry lines	63
2.4	Strategy 3 - Engineering of the gypsy retrotransposable element as a Y-chromatin insulator	65
2.5	The dCas9-Pcaf fusion failed to activate the transgene on the Y chromosome.	66
2.6	Characterisation of the Y-linked I-Ppol cleavage activity in the gypsy line.....	70

3	<i>Discussions and future perspectives.....</i>	74
3.1	The dCas9/Cas9 competition experiment provides insights on dCas9 kinetics during gametogenesis.....	74
3.2	The dCas9 is actively deposited in the embryo and interferes with dsx expression.	78
3.3	The dCas9-Pcaf failed to enhance gene expression during meiosis, but there might be room to improve.....	79
3.4	The gypsy I-PpoI line might be tuned to achieve the Y-drive.	82
3.5	Delaying the expression of a gene with RNA-targeting Cas proteins.....	84
3.6	The Y-linked meiotic sex distorter is one of the most powerful gene drive strategies but poses several challenges.	86
3.7	The absence of a Y-chromosome assembly limits Y-linked strategies for pest control. ..	88
3.8	Conclusions	89
4	<i>Methods</i>	91
4.1	PCR amplification and gel extraction	91
4.2	Gibson Assemblies and Bacterial Transformations.....	91
4.3	Restriction digestions	92
4.4	Colony-PCR.....	93
4.5	Genomic DNA extraction	93
4.6	RNA extraction and qRT-PCR	94
4.7	Statistical analysis	95
4.8	Bioinformatics tools	96
4.9	The assembly of the Gal-UAS plasmids	96
4.10	The assembly of the dCas9-mCherry plasmids	99
4.11	Amplification of the Pcaf and fusion with the dCas9.....	102
4.12	DNA microinjections in mosquito embryos	104
4.13	Molecular confirmation of transgene integration	105
4.14	Mosquitoes' microdissections	105
4.15	Phenotype assays	106
4.16	Amplicon sequencing of females from gypsy Y-Ppo I line.	106
5	<i>Bibliography.....</i>	108

List of Figures

Figure 1-1 Malaria incidence in 2000 vs 2020.	14
Figure 1-2 Mosquito Life Cycle.	16
Figure 1-3 Wolbachia-induced Cytoplasmic incompatibility.	23
Figure 1-4 Mendelian inheritance vs Gene Drive.	26
Figure 1-5 Gene editing and cell repairing pathways.	27
Figure 1-6. The CRISPR/Cas9 immune system in <i>S. pyogenes</i>	30
Figure 1-7 Meiotic sex distortion.	35
Figure 1-8. Meiotic stages and details of MSCI.	38
Figure 1-9. Epigenetic levels.	39
Figure 1-10 The gypsy retrotransposable element as chromatin insulators.	44
Figure 1-11 Structure and conservation score of the Pcaf histone acetyltransferase.	48
Figure 2-1 Structure of plasmid injected, Pcaf domain structure and B2tub target sites.	51
Figure 2-2 Expression patterns of the vasa-2 and B2tub regulatory regions.	53
Figure 2-3 The PiggyBac integration system.	54
Figure 2-4 Crossed performed for the dCas9-Cas9 competition and experimental setup.	58
Figure 2-5. Data collected from the Cas9-dCas9 competition experiment.	59
Figure 2-6 The Cas9-dCas9 competition experiment in the 7280 gene drive line.	60
Figure 2-7 Testing the dCas9 misexpression and deposition.	62
Figure 2-8 Vasa2 promoter drives the expression of a GFP on the antisense strand (Scale bar = 200 μ m).	63
Figure 2-9 The dCas9 alone failed to produce sex bias.	64
Figure 2-10 Site-specific integration on the <i>A. gambiae</i> Y chromosome	65
Figure 2-11 Crosses and results of the Pcaf delivery on the Y-linked B2tub-I PpoI transgene.	67
Figure 2-12 Primers positions for I-PpoI qRT-PCR	68
Figure 2-13 qRT-PCR expression values for the I-PpoI.	69
Figure 2-14 Amplicon sequencing analysis of the 28S ribosomal RNA-coding sequence in single daughters of Y-linked gypsy-I-PpoI fathers.	72
Figure 2-15 Significance level of the mutations detected from amplicon sequencing.	73
Figure 3-1 Visualisation of the model hypothesised to explain the probability of dCas9 binding during spermatogenesis.	76
Figure 3-2: A model that tries to explain the dynamics of dCas9/Cas9 competition when expressed under different promoters. The homing rate inhibition in this experiment was 25%. This percentage might be explained if the two proteins start to compete at the same levels after the second cycle of replication. On the right is the relative expression strength of the vasa and zpg regulatory regions, as reported by Taxiarchi et al. Made with Biorender.com	77

List of Tables

<i>Table 1-1 Histone tails modification and their know effects.</i>	<i>42</i>
<i>Table 4-1 Primers employed for assembling the GAL-UAS plasmids.</i>	<i>98</i>
<i>Table 4-2 Primers employed for the mutagenesis of the Cas9 and the cloning of the plasmids with the dCas9 fused with the mCherry fluorescent protein.</i>	<i>101</i>
<i>Table 4-3 Oligonucleotides employed for cloning the gRNA spacers into the dCas9 plasmids.</i>	<i>101</i>
<i>Table 4-4 Oligos used to amplify the Pcaf histone acetyltransferase and its truncate version (CORE).</i>	<i>103</i>

List of Acronyms

bp	base pairs
Cas	CRISPR-associated protein
CB	Cytoblast
CDS	Coding sequence
CI	Cytoplasmic Incompatibility
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
<i>dsx</i>	<i>Anopheles gambiae doublesex gene</i>
GB	Gonioblast
GFP	Green Fluorescent Protein
GSC	Germline Stem Cells
gRNA	guide RNA
HDR	Homology-Directed Repair
HDG	Homing gene drive
HEG	Homing Endonuclease Gene
MMEJ	Microhomology-Mediated End-Joining
NHEJ	Non-Homologous End-Joining
ORF	Open Reading Frame
RFP	Red Fluorescent Protein
trRNA	Trigger RNA
SD	standard Deviation
SEM	Standard Error of Mean

Statement of Originality

I hereby declare that the scientific work presented in this thesis is entirely my own. The project, the experimental design and the data analysis are the results of grouped discussions with my supervisor and my progress review panel. All else was appropriately referenced.

Copyright Statement

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a [Creative Commons Attribution-NonCommercial 4.0 International Licence](#) (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

Mosquito-borne diseases are responsible for around 700,000 deaths every year. Malaria alone is responsible for at least 400,000 deaths yearly, with the African continent accounting for more than 90% of those deaths. In recent years, new outbreaks of pathogens focused international attention on developing new methodologies to curb mosquito diffusion. Among those methodologies, the generation of genetically modified mosquitoes to reduce the size of vector populations seemed the most promising. The development of CRISPR systems has facilitated the engineering and application of several strategies to achieve a super-mendelian inheritance of desired genes, also known as gene drive.

A sex-distorter meiotic drive developed in *A. gambiae* allowed to heavily male-bias the vector progeny, lowering the population's reproductive capacity. To achieve the super-mendelian inheritance of this transgene, the solution would be to place it on the Y chromosome. Unfortunately, when placed on the Y, the sex-distorter gene is completely silenced. The silencing is attributed to Meiotic Sex Chromosome Inactivation (MSCI), an epigenetic silencing mechanism that is responsible for the silencing of the two sexual chromosomes during meiosis.

Here, I am testing different strategies to rescue the activity of this Y-linked gene. I am employing a nuclease-dead Cas9 to deliver an epigenetic activator on the target gene and testing *Drosophila*'s gypsy-retrotransposable elements as chromatin insulators.

Although the delivery of the epigenetic editor failed to increase the transcription levels of the Y-linked sex distorter gene, this study casts some light on the binding kinetics of the dCas9 during meiosis. Moreover, I discovered that the gypsy insulators were actually able to mildly increase the transcript levels of the transgene from the Y-chromosome, even if not high enough to produce a significant sex distortion. Finally, I discuss some ideas for future developments that might be promising for the final achievement of the Y-linked meiotic drive.

1 Introduction

1.1 A brief history of Malaria

Malaria is one of the most ancient human diseases. Its traces are dated back to the Neolithic, and its prevalence is widespread around the globe. Malaria imposed its burden on all the great human civilisations, probably spreading from subtropical Africa across the Nile and then, through Mesopotamia, reached Asia and southern Europe. During the expansion of the Roman Empire, the disease spread across the entire European continent, reaching England and Denmark (Arrow, Panosian and Gelband, 2004). There are no records of Malaria in America before the European colonisation. Hence, it is highly probable that the *Plasmodium* parasite crossed the Atlantic Ocean on European ships. By 1750, the disease was diffused across the entire continent except in colder regions. But the African continent is where malaria imposed, by far, the higher death cost. The Portuguese traders across the fifteenth and sixteenth centuries were the first Europeans to face the deadly disease. Since then, any attempt to colonise the continent had to face the consequences of the epidemic.

The disease imposed such a high cost in terms of human lives that Africa's western and central regions earned the title of "White man's grave". The burden imposed on the local population was even worse. From an evolutionary point of view, the fitness cost set by malaria in some regions of Africa was so high that it fostered the diffusion of a rare genetic disease, Sickle Cell Disease (SCD), that confers protection against *Plasmodium falciparum*. SCD is a genetic disorder characterised by a modified form of haemoglobin: the haemoglobin S. An individual who inherits both copies of the haemoglobin S allele has meagre chances of survival. However, the inheritance of a single allele confers protection against *Plasmodium* with mild or minor consequences for the carrier. In affected individuals, the polymerisation of the haemoglobin S causes a morphological change in the erythrocytes that resemble a sickle moon (hence, the name of the disease) instead of being biconcave. The new environment is hostile to the development of the *Plasmodium*, and therefore the SCD heterozygous individuals are more resistant to its infection. The "selection" for the SCD allele amounts to 20% in some regions of Africa compared to levels of around 0.5% in others, with a geographical distribution that matches quite accurately that of malaria (Piel et al., 2010; Grosse et al., 2011).

1.2 Malaria today

Malaria is still one of the world's most diffuse and deadly diseases. Its impact on human life is significant, considering that in 2020 it killed about 627.000 people over an estimated 241 million cases worldwide (WHO, 2021). Furthermore, 67% of malaria-related deaths were children under the age of 5. The African continent alone accounted for 96% of malaria-related deaths in 2020 (*Figure 1-1*). Malaria can also have severe consequences on pregnant women. In 2018, about 11 million pregnant women were exposed to malaria, delivering around 872.000 low birthweight children. Moreover, in countries where malaria is prevalent, children are also exposed to the risk of developing severe anaemia. In 2018, in sub-Saharan Africa, 1.8 million children developed severe anaemia due to malaria infection (White, 2018). Despite the present considerable number of deaths, the situation in the past was even more frantic, with malaria-related deaths that peaked at around 900.000 in the first 2000s. (WHO, 2019).

Malaria's economic impact on the African continent is estimated to be worth billions of US dollars, thus severely constraining the efforts put in place by the African countries to elevate the quality of life of their citizens (WHO, 2021). Despite the difficult situation, the global efforts to reduce the malaria burden were actually successful. After the alarming situation of the early 2000s, malaria deaths reached a low of 560.000 in 2018, after following a negative trend the years before. This negative trend accurately and inversely matches the amount of funds allocated to fight malaria, which has been in constant growth since the beginning of the new millennium. Therefore, it is not surprising that in 2020 malaria-related deaths increased for the first time after a twenty-year negative trend. The COVID pandemic redirected a substantial amount of money normally allocated to fight malaria. Moreover, the world's governments' travel restrictions made all the operations to control and prevent malaria more complex and difficult (WHO, 2021).

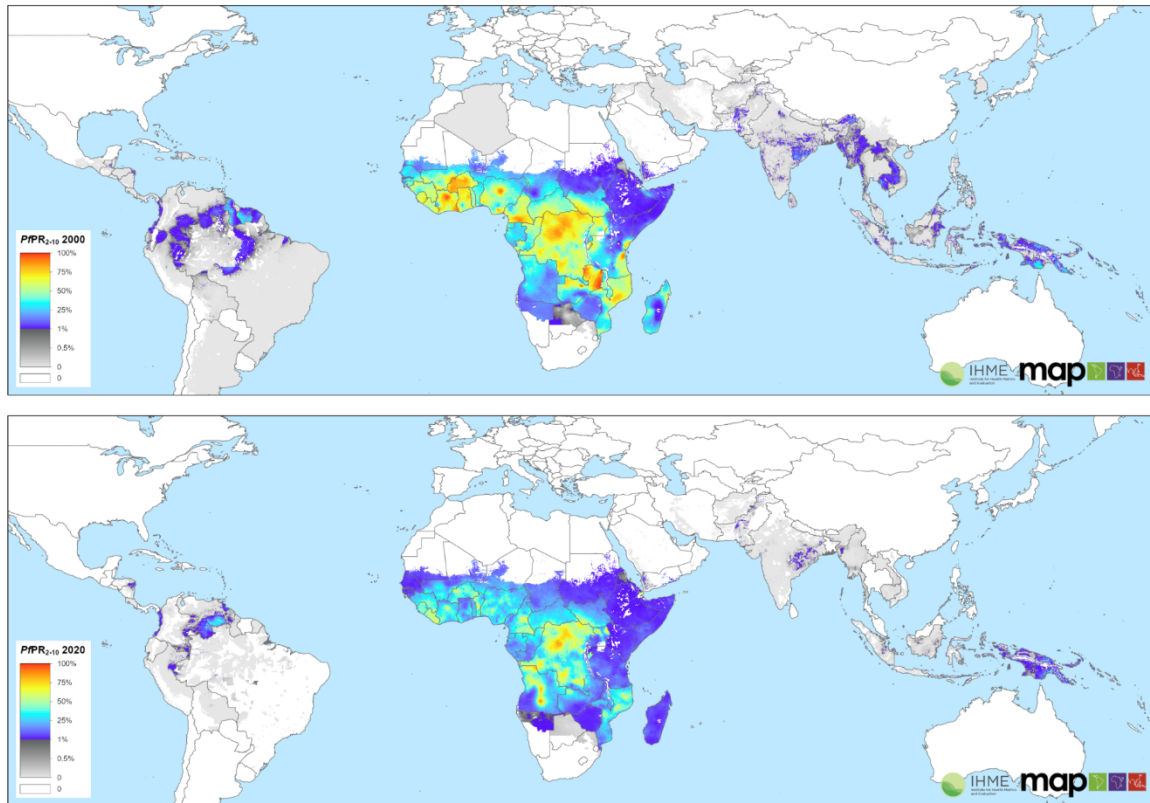


Figure 1-1 Malaria incidence in 2000 vs 2020.

The efforts put in place to control the diffusion of the Plasmodium parasite were successful in reducing the overall burden imposed by the disease, but its complete eradication still looks far away. Credits: The malaria atlas project (www.malariaatlas.org)

1.3 Mosquito life cycle

The *Anopheles* genus belongs to the order Diptera, an order that encompasses insects characterised by the presence of two wings (from the Greek “di-pterón”, literally “two wings”). Their morphology allows them to be highly efficient in flying. Moreover, they are equipped with advanced mechanosensory organs that can rapidly detect changes in several environmental parameters like temperature, air pressure or the presence of carbon dioxide (Tauxe et al., 2013).

Understanding the mosquito life cycle is crucial for planning strategies aimed at population control. Knowing the different developmental stages help to identify how and when to intervene. For example, being aware of the existence of aquatic stages in the mosquito life cycle allowed us to carry out strategies for the treatment and removal of breeding sites.

The life cycle of a mosquito starts, like any other animal with sexual reproduction, with mating (Figure 1-2). The mated female then starts looking for a blood meal to provide the nutrients for egg development. The ovaries are in a suspended state at this stage, and their development can restart only upon a blood meal. To locate her host, she uses her sophisticated sensory system. *Anopheles* mosquitoes are considerably anthropophilic, but they can also feed on other animals if needed (Adugna, Yewhelew and Getu, 2021).

Once fed, the female is in a highly vulnerable state. The extra weight from the blood meal heavily limits her ability to fly. Therefore, during the first hour after the blood meal, the female mosquito rests and concentrates the ingested blood, expelling over 40% of the water intake accumulated with the blood meal (Liu et al., 2011). About three days after the meal, the development of the eggs is completed, and the female seeks out suitable oviposition sites, preferably stagnant fresh water bodies. This is the first step of the aquatic stages of the mosquito life cycle. About two days after the deposition, the eggs hatch and a tiny larva emerge. The larva will start immediately looking for food and begin growing, undergoing four main stages characterised by the progressive increase of larval size (L1-L4). After about seven days from the hatching, the larva is ready for the first metamorphic stage of the life cycle, becoming a pupa. The pupal stage lasts for about two days, in which the pupa positions herself at the water surface to facilitate the emergence of the mosquito adult form. The pupa doesn't have a mouth; hence it will not feed during this stage. Finally, two days after, the adult mosquito emerges from the dorsal portion of the pupa exposed to air, leaving behind an empty exuvia. After a few moments spent on the water surface, thanks to legs specialised in taking advantage of the water surface tension, the adult mosquito is ready to fly and restart the cycle.

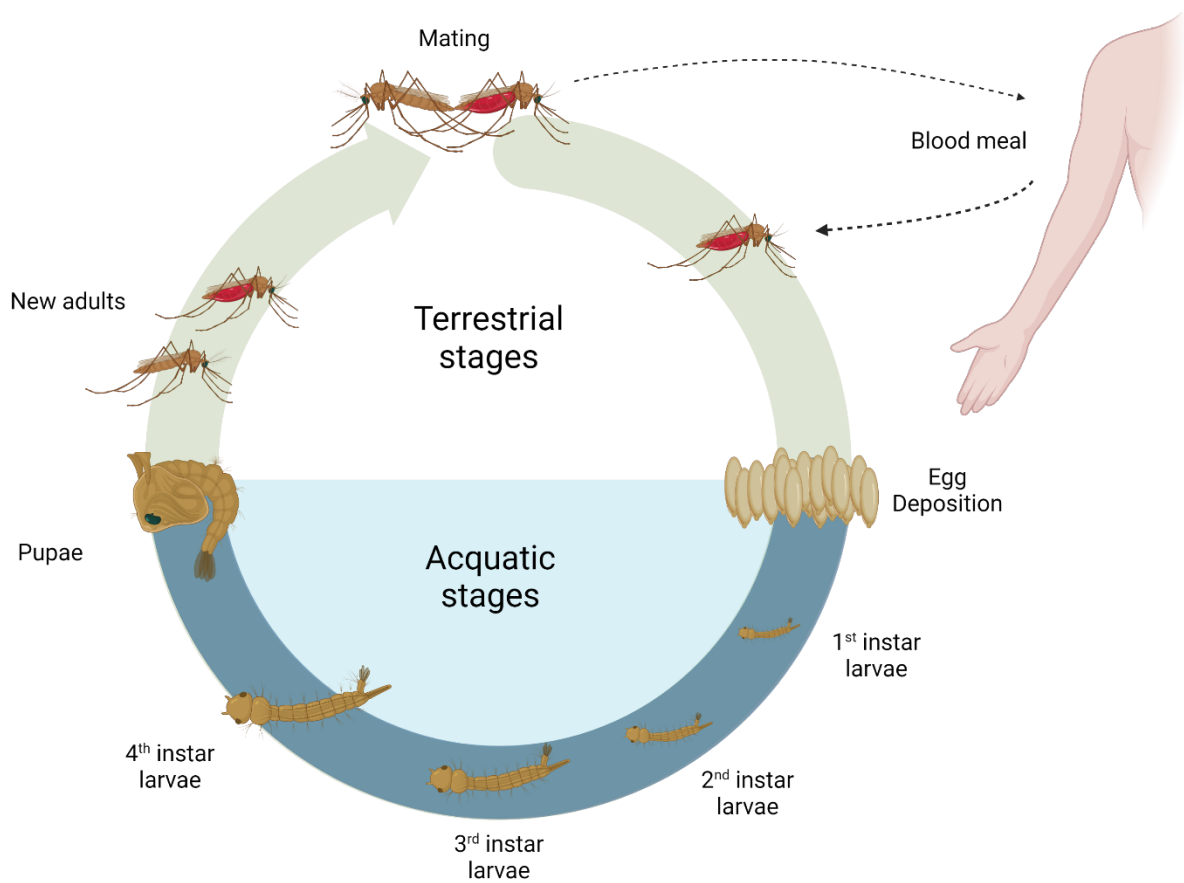


Figure 1-2 Mosquito Life Cycle.

The mosquito life cycle is divided into terrestrial and aquatic stages. The terrestrial stages are usually considered “adult stages” and include mating, blood meal and egg deposition. The aquatic stages start with the eggs, followed by the larvae’s hatching and development through four size-dependent stages. The last aquatic stage is the pupa, from which a new adult mosquito will emerge. Made with Biorender.com

1.4 Strategies for vector control

Nowadays, in the fight against malaria, two main strategies are employed:

- 1) targeting the human malaria parasite *Plasmodium falciparum* with drugs or vaccines, and
- 2) targeting the vectors responsible for its transmission with insecticides or bed nets.

So far, targeting the parasite has been extremely difficult and produced mixed results. The extensive use of drugs like Artemisinin to target *Plasmodium* in malaria-infected patients led to the development of drug-resistant parasites. The selective pressure exerted by pharmacological treatment fostered the selection of mutations (particularly the PfKelch13 polymorphisms that confer artemisinin resistance) and made the parasite refractory or less susceptible to further drug treatments. The alternate use of different drugs is now in place to avoid the selective pressure exerted by the extensive use of a single one. Still, an increasing rate of treatment failure casts doubts on the long-term efficacy of this pharmacological approach (WHO, 2021).

Tremendous efforts were also made to develop a vaccine. But after more than thirty years of research, the only vaccine that underwent phase III of clinical trials, the vaccine RTS, S/AS01 (Mosquirix™), was able to prevent only 30% of severe cases (Mahmoudi and Keshavarz, 2017). Moreover, its diffusion was also hampered by safety issues related to cases of meningitis and an increased mortality rate in young girls (Klein et al., 2016). Nevertheless, a new potential candidate named R21/Matrix-M started the clinical trials in 2019 and showed a promising overall efficacy of around 75%. It seemed also better tolerated and did not raise any particular concern about its safety (Dattoo et al., 2021). Despite these significant results, a full-scale implementation of this vaccine in Africa still seems far in the future. And even if a mass vaccination campaign starts, concerns remain about the possibility that the parasite will evolve to counteract the vaccine as with artemisinin. Moreover, the COVID pandemic has shifted the world's vaccine manufacturing capacity towards anti-COVID vaccines, probably furtherly delaying the full-scale production of the R21.

1.5 The successful Italian eradication and the African situation today

Vectors from the *Anopheles* genus play a pivotal role in spreading malaria infection. The four main species that transmit the disease are *Anopheles gambiae*, *Anopheles arabiensis*, *Anopheles coluzzi* and *Anopheles funestus*. Conventional strategies used so far to reduce these vector populations are mainly chemical-based. Insecticide-impregnated bed nets (ITN) and indoor residual spraying (IRS) are the two most used tools to mitigate the growth of vector populations, as long as removing mosquito breeding habitats. These tools are the main ones responsible for reducing malaria death counts in these last three decades and showed that targeting the vector is the best strategy we have so far against malaria.

Italy represents a successful case of malaria eradication achieved mainly through vector control. At the beginning of the 20th century, malaria was present in almost one-third of the Italian peninsula. It caused nearly two million deaths yearly in a population of around thirty million people. As a result, the Italian government put in place efforts to remove mosquito breeding sites and produced and distributed quinine, mainly thanks to The Italian School of Malariology's guide. The first attempts to fight mosquitoes with insecticides were made before the second world war using the “Paris Green” insecticide. From 1924 to 1934, thanks to the contributions of the Rockefeller Foundation, Italian researchers showed that targeting the vector population was the most efficient strategy for reducing malaria morbidity. After the second world war, a further push to remove mosquito breeding sites and large-scale use of DDT led to the total eradication of malaria from the Italian peninsula, which was declared malaria-free in 1970 (Majori, 2012).

Unfortunately, the approaches followed to eradicate malaria in Italy are not applicable to the African continent. The smaller surface of the Italian peninsula and its peculiar geographical conformation created a clustered distribution of vector populations that were relatively easier to target with insecticides. In Africa, the lesser presence of geographical obstacles to mosquito migration led to a pretty homogeneous distribution of vector populations and the disease. Except for the Sahara desert, the four main malaria vector populations are spread throughout the African continent. The logistical and technical problems in employing insecticides on such a vast territory are unsurmountable. Furthermore, the environmental impact of intensive insecticide use on such a large scale will be catastrophic for many other insect species, including pollinators. Moreover, due to their intensive use in recent years, mosquitoes have

started to develop resistance to insecticides, posing new challenges to malaria eradication (Hemingway et al., 2004).

Although the successful use of chemical-based strategies like ITNs and IRS in malaria control measures, in regions where these tools have been used to achieve high vector control coverage, entomological surveillance and monitoring revealed a return in malaria transmission (Brady et al., 2016; Sougoufara, Ottih and Tripet, 2020), numerous mosquito characteristics, such as behaviour, genetics, and physiology, have been impacted by the selective pressures exerted by pesticide exposure, locally changing malaria's transmission patterns. The fast evolution and spread of pesticide resistance in the primary malaria vectors in many locations of the African continent poses a threat to the effectiveness of indoor chemical control techniques (Brady et al., 2016; Sougoufara, Ottih and Tripet, 2020).

There have been initiatives to create new vector control techniques to overcome these issues. The appealing toxic sugar bait (ATSB) approach is one such technique that has been successful in eradicating local anopheline and culicine mosquito populations in Israel (Müller et al., 2010). It employs fruit or flower aroma as an attractant, sugar solution as a feeding stimulant, and an oral poison. Three field studies were carried out also in Mali to test the attractiveness of *Anopheles gambiae* s.l. to native fruits/seedpods and flowering plants. The study shows that male and female *A. gambiae* are attracted to different natural items, which can help to develop more refined ATSB malaria vector control strategies in Africa (Müller et al., 2010; Sougoufara, Ottih and Tripet, 2020).

Also, using biolarvicides to target the immature stages of malaria vectors in their outdoor aquatic habitats, is becoming increasingly important for achieving malaria elimination in sub-Saharan Africa. This method has successfully reduced the number of invasive populations of *Anopheles* mosquitoes in Egypt and Brazil. Biolarvicides, such as *Bacillus thuringiensis israelensis* (Bti) and *Bacillus sphaericus* (Bs), are usually preferred over chemical compounds due to their natural origin and low environmental toxicity. Large-scale application of Bti in Burkina Faso resulted in a significant decrease in exposure to bites, and the combination of biolarvicides with ITNs led to significant decreases in malaria transmission in some areas (Sougoufara, Ottih and Tripet, 2020). Biolarvicides can specifically target anophelines and have fewer effects on non-target organisms than chemical larvicides, making them a promising option for an integrated vector control approach in sub-Saharan Africa (Sougoufara, Ottih and Tripet, 2020).

The use of ivermectin, an anti-filarial medication, is another tool used to manage anopheline vectors. This strategy showed promising results against indoor and outdoor mosquito

populations. The administration of ivermectin to human hosts and/or their domestic animals lowered the mosquito post-blood meal survivability. *Anopheles* mosquitoes feeding on cattle or people treated with ivermectin have been shown to have higher mortality rates and lower fertility in studies conducted in Burkina Faso and Senegal (Alout et al., 2014). Ivermectin has been widely administered in three countries in West Africa, and results have indicated a substantial decline in sporozoite rates in both indoor resting and outdoor host-seeking *Anopheles* mosquitoes. Although the ivermectin effect seemed to last just a couple of weeks, new trials are currently being conducted to develop substitute longer-lasting substances that can be employed for mass-drug administration methods while minimising potential side effects (Alout et al., 2014).

It is clear that there is no “silver bullet” for malaria eradication. The current indoor chemical control tools endorsed by WHO are progressively losing effectiveness as malaria transmission is now significantly sustained by outdoor biting vectors. The spread of insecticide resistance and the shift in vector composition and feeding patterns need additional control tools that deal specifically with these new phenotypes. New emerging tools, such as synthetic attractants, larvicides, sugar baits, and genetic engineering, may be effective but only if employed in an integrated manner. The hope is that a more diverse toolbox will facilitate increased versatility and integration of vector control management, as well as adopt more responsible and sustainable use of classic chemical control tools.

1.6 Alternative strategies for vector control

Malaria represents only one of the many diseases carried by vector populations. With climate change, the rapid growth of the human population and its invasion of previously wild territories, the fraction of people potentially exposed to vector-carried pathogens might account for almost half of the entire human population. The African continent is not the only one threatened by mosquito-carried diseases. The economic and scientific efforts to fight the spread of the most dangerous vector species are now on a global scale, and several strategies, alternatives to the classic insecticide approach, are already in place trying to control the increasing transmission of vector-carried diseases (Alphey et al., 2010; Wang et al., 2021).

The two main strategies the scientific community pursues to curb the diffusion of these diseases aim to achieve population replacement or suppression. The population replacement strategies seek to provide the vector with some characteristics that make it refractory to pathogens. On the other hand, with population suppression, the target is to lower the vector population size under a threshold that will stop pathogen transmission. Both these approaches are pursued mainly via the generation of transgenic mosquitoes, but there are also non-transgenic alternatives.

1.6.1 *Wolbachia* as a tool for population replacement and suppression

The bacteria strain *Wolbachia* is one of the most promising non-transgenic tools for population suppression and population replacement. This intracellular parasite of insects (and some nematodes) is maternally transmitted and produces what is known as Cytoplasmic incompatibility (CI). *Wolbachia*-infected females mated with wild type males will produce viable progeny that will inherit the bacteria infection. This will also happen if both parents are infected with the same *Wolbachia* strain. On the other hand, when an infected male mates with a wild type female, the female will produce no offspring. The same incompatibility arises if the two parents are infected with different *Wolbachia* strains (*Figure 1-3*) (Alphey et al., 2013; Consortium and Ching, 2021; Wang et al., 2021).

Interestingly, *Wolbachia* also seems to reduce mosquito vector competence against several viruses, including dengue, Zika, West Nile and the malaria parasite *Plasmodium* (Alphey et al., 2013; Gomes and Barillas-Mury, 2018; Pereira et al., 2018; Gesto et al., 2021). This made *Wolbachia* suitable also for population replacement strategies. For example, a first trial in Singapore aimed to implement a *Wolbachia*-IIT (Incompatible Insect Technique) strategy

showed that establishing *Wolbachia*-infected mosquito populations significantly reduced the overall morbidity of the targeted area (Consortium and Ching, 2021). This trial was carried out in *Aedes aegypti* and was the first successful example of *Wolbachia*-based vector control in a natural environment, with a reduction of the local *Aedes* population of up to 98% and a reduction of dengue incidence of up to 88%.

Wolbachia infections were also reported in *Anopheles* (Wong et al., 2020). However, till recently, no *Wolbachia* strain seemed to be able to induce CI in *Anopheles*. Recent findings characterised two *Wolbachia* genes (*cifA* and *cifB*) that were able to cause conditional sterility when expressed in the *A. gambiae* germline. The embryonic lethality is mainly induced by *cifB* and can be rescued by the co-expression of *cifA* in females. In males, the *cifB*-induced sterility is rescued by *cifA* to a much lower extent (Adams et al., 2021).

Wolbachia is an up-and-coming tool for population replacement and population suppression strategies. Nevertheless, doubts persist about the possible selection of pathogens that might adapt to the *Wolbachia*-infected host environment.

1.6.2 Incompatible Insect Technique (IIT) and Sterile Insect Technique (SIT)

The use of CI in insects to suppress vector populations is known as Incompatible Insect Technique (IIT) and could be considered a variant of the Sterile Insect Technique (SIT) (Alphey et al., 2013; Wang et al., 2021). Both concepts rely on releasing sterile (or infected) males in wild type populations to lower their reproductive capacity. Female mosquitoes mate only once in their lifetime. They store the sperm in an organ called spermatheca that will act as a sperm reservoir throughout a female's lifespan. Hence, if a sterile male mates with a female, the latter will never be able to produce viable progeny. Male sterility can be achieved in different ways. The first SIT approach mainly relied on the irradiation of the males to induce sterility. Unfortunately, this approach has the drawback of reducing the mosquitoes' overall fitness and hence their mating competitiveness with wild type rivals. Moreover, the practicality of this method is limited by the need to produce and irradiate up to several billion insects per week, according to the area of the territory involved and the size of the target wild type population. This is, in fact, a non-self-sustained approach that will require multiple and constant releases to obtain results (Alphey et al., 2010; Wise De Valdez et al., 2011). IIT and genetically modified sterile males partially addressed the issues related to the use of irradiation. Still, they could not solve the problem related to insect mass production and sex sorting. Using

mechanised automatic systems to sex-sort mosquitoes is not fail-proof. In some cases, like with the *Wolbachia*-derived IIT trial in Singapore, the accidental release of a few females compromised the entire IIT strategy and established a *Wolbachia*-infected colony that was not targetable anymore by *Wolbachia*-infected males (Consortium and Ching, 2021). These new challenges were partially addressed with the development of techniques that enable genetic sex sorting.

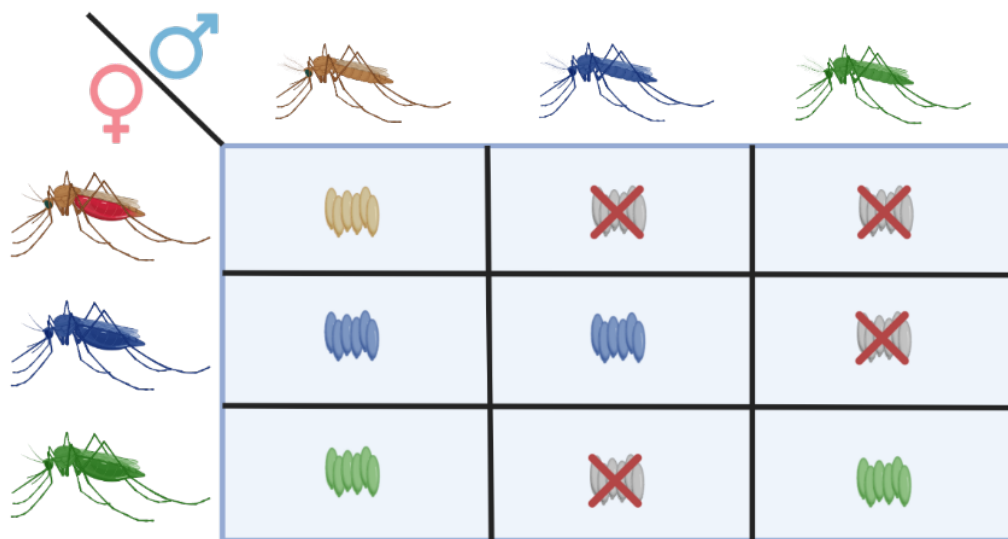


Figure 1-3 *Wolbachia*-induced Cytoplasmic incompatibility. Infected males mated with wild type females or females infected with a different *Wolbachia* strain produce no progeny. Infected females only give progeny when mated with wild type males or males infected with the same *Wolbachia* strain. All progeny generated by an infected female will inherit the infection. Made with Biorender.com

1.6.3 The Meiotic Sex Distorter and Release of Insects Carrying a Dominant Lethal (RIDL)

Recent advances in genetic engineering and DNA sequencing fostered the development of genetic-based strategies to generate heavily sex-biased progenies in mosquitoes.

A transgenic mosquito strain that can produce a heavily male-biased progeny was developed in *A. gambiae* (Galizi et al., 2014, 2016). The efficiency of the generation of a male-biased progeny was extremely high, peaking at >98%. This result was mainly achieved through the meiotic expression of an enzyme that was able to destroy the X chromosome during spermatogenesis, leading to the incapacitation of all the X-bearing sperm (*Figure 1-7*).

In *Aedes*, a female-specific version of the Release of Insects carrying a Dominant Lethal (fsRIDL) was able to generate progeny in which females were incapacitated to fly and hence unable to bite or reproduce. In the standard RIDL system, the mosquitoes carry a lethal transgene that is suppressed by an antidote. During the aquatic stages of development, the larvae are bred in the presence of it. Upon antidote removal, it is possible to induce lethality at any mosquito developmental stage. In fsRIDL, the toxic effector is linked to a gene expressed only in the flight muscles of female mosquitoes. Therefore, once the antidote is removed, the gene starts to damage the wings of the females selectively. If the same strategy was applied to an essential sex-specific gene, it would produce lethality in a sex-specific fashion (Fu et al., 2010; Wise De Valdez et al., 2011; Labbé et al., 2012; Alphey et al., 2013).

A step further in genetic sex-sorting applied for population suppression was developed with the precision-guided Sterile Insect Technique (pgSIT). Exploiting the precision of the programmable nuclease Cas9, this system targets two genes at once: one important for male fertility (*beta-2-tubulin*) and one female-specific sex determination gene (*doublesex* female isoform, *sex lethal* or *transformer*) (Kandul et al., 2019). With this approach, it was possible to generate a progeny in which the males were sterile, and the females exhibited an intersex phenotype unable to bite and reproduce.

All these strategies are extremely promising for future application in the field, but none of them is completely fail-proof, and still, they do not fully address the problems related to mosquito mass rearing. For these reasons, the use of transgenic modifications able to self-propagate within a population or Gene Drive (GD) is drawing more and more attention among the scientific community.

1.6.4 A new hope for malaria eradication: the gene drive.

A gene drive is a genetic system that allows a gene to spread in a population by defying the traditional rules of Mendelian inheritance (*Figure 1-4*) (Burt, 2003; Galizi et al., 2016; Kyrou et al., 2018). A gene is usually present in two copies or alleles in diploid organisms. Each allele is present on one of two sister (homologous) chromosomes. When the two alleles are identical, the individual is said to be homozygous for that gene, and its entire progeny will inherit one of the two identical alleles. Therefore, the transmission rate of that gene will be 100%. When a transgene is inserted into an individual's genome (creating a Genetically Modified Organism - GMO), it is always in a single copy. Therefore, it is inherited only by half of the progeny. If the transgene-induced phenotype does not give a competitive advantage to the organism or is deleterious, the genetic construct will be progressively or immediately selected against and disappear over time from the population (Burt, 2003; Hammond et al., 2017; Kyrou et al., 2018). A gene drive system allows the desired transgene to be always inherited by the progeny, even if it produces a less competitive phenotype. Almost two decades ago, Austin Burt theorised the “gene drive” concept, exploiting the idea of engineering “selfish” genetic elements able to copy and multiply themselves into an individual genome. He theorised that by using these engineered selfish elements to disrupt essential haplosufficient genes, it could be possible to suppress entire wild type populations of pest or vector species. An essential haplosufficient gene produces a normal phenotype if one of the two copies is disrupted but can be lethal or highly harmful if both are affected.

It was observed that these “selfish” genetic elements could recognise and cut defined DNA sequences and self-copying during the repairing process if placed in the locus opposite the cut. In fact, to repair the double-strand break, the cell may use a process named Homology-Directed Repair (HDR) that uses the same locus on the unbroken homologous chromosome as a template for repair (*Figure 1-5*) (Liang et al., 1998). A programmable HEG could then be employed to bind and cut specific DNA sequences and then be copied using HDR (*Figure 1-5*) (Burt, 2003). HEGs are genetic elements often associated with self-splicing introns (groups I, II and III). This class of nucleases may promote the insertion of introns into an intronless allele recognising and cutting sequences that are just upstream and downstream of the intron sequence. Their preferred insertion sites are inside or nearby essential or highly conserved genes that are usually less subjected to recombination. In this way, HEGs are more likely to be inherited throughout generations. Another critical characteristic of HEGs is their ability to

change their recognition sequence easily. Thus, they can adapt in the event of a mutation or disruption of their original target sequence (Edgell, 2009).

It was also predicted that some form of genetic resistance might arise. HDR is only one of the different possible mechanisms of DNA repair. Microhomology-Mediated End Joining (MMEJ) and Non-homologous End Joining (NHEJ) are two other mechanisms involved in DNA double-strand break repair (*Figure 1-5*) (Truong et al., 2013). These repairing processes often result in DNA mutations, insertions or deletions. Still, due to the degenerated nature of the DNA codon code, the gene sequence might mutate without compromising protein functionality. In the context of a gene drive, this will create a resistant allele that may hinder the nuclease homing potential. If we consider this in a population-dynamics context, even if the disruption of one allele of the target gene has a minimal fitness cost for a single individual (or for the population as a whole), the functional nuclease-resistant allele will be positively selected because it will relieve the population from that fitness cost. Therefore, it will quickly spread across the population until fixation, erasing the gene drive effects (Hammond et al., 2017).

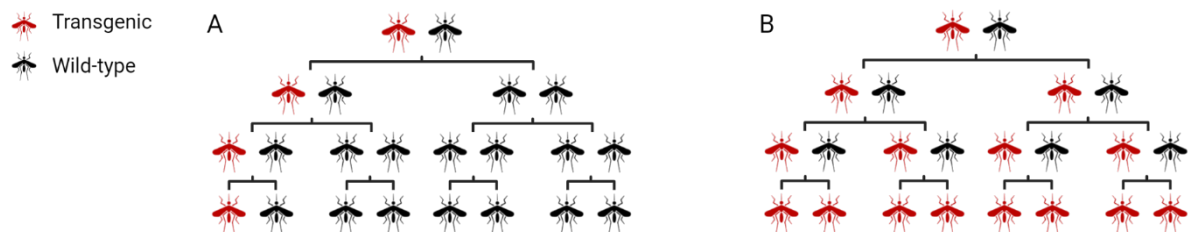


Figure 1-4 Mendelian inheritance vs Gene Drive.

A: The classic Mendelian Inheritance for a given transgene. B: How gene drive transgene spread across a wild type population.

Thus, in an ideal gene drive system, a programmable nuclease engineered to recognise and cut the coding sequence of a target gene on the homologous chromosome would be able to spread into a wild type population exploiting the HDR pathway. Over time, this would allow the endonuclease to produce a double knock-out of the target gene generating unviable individuals and eventually suppressing the population in the face of natural selection (*Figure 1-4*).

Hence, the characteristics of such an ideal gene drive target are:

- I) Haplosufficiency, to keep the individual competitive with the wild type counterpart even with one allele knocked out
- II) Lethality, meaning that the complete knock-out of the gene should produce a lethal or at least a sterile phenotype
- III) Intolerance to any mutation to hamper the generation of resistant alleles.

The endonuclease expression can be restricted to gametogenesis using an appropriate promoter, thus ensuring that all gametes inherit the transgene. With such a setup, the “selfish” gene would rapidly spread throughout the target population until heterozygous individuals start mating, generating homozygous non-viable progeny and leading to population suppression. The highest efficiency in lowering the population’s growth would be achieved by targeting females. Females bear most of the population’s reproductive capacity, while targeting males would only slow down the gene drive's transmission whilst not effectively lowering the population growth rate. Moreover, female mosquitoes are the only ones who spread pathogens. Hence lowering their numbers will also slow the spread of the carried diseases.

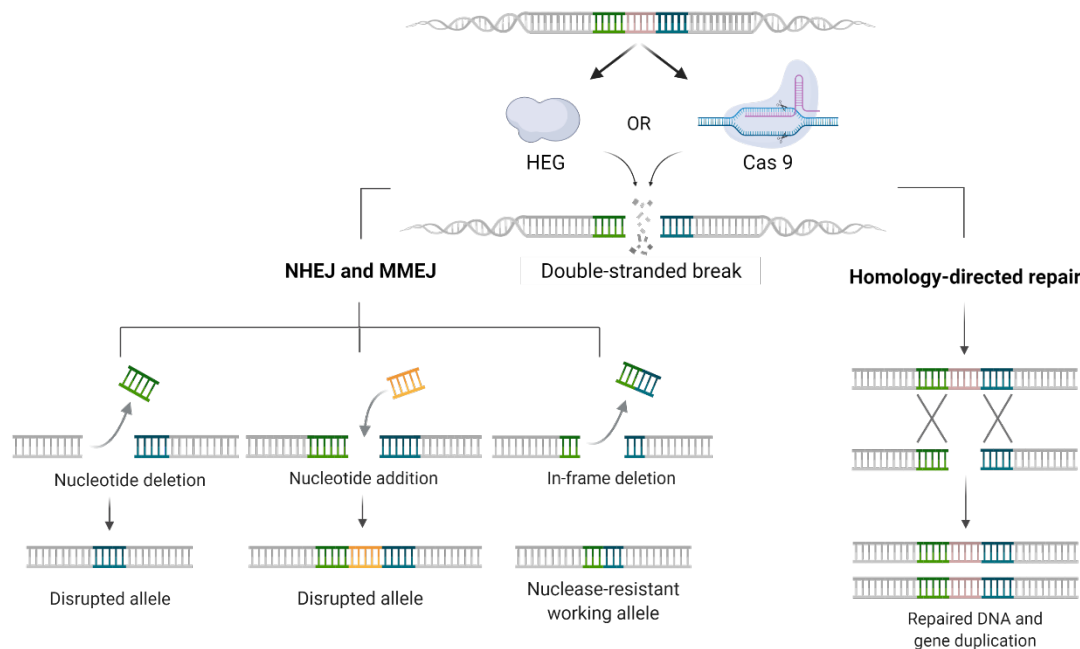


Figure 1-5 Gene editing and cell repairing pathways.

A gene expressing a programmable nuclease is employed to cut the opposite chromosome. Suppose the cell uses the HDR pathway (on the right) to repair the double-strand break. In that case, the gene expressing the nuclease will be copied into the opposite chromosome, disrupting both copies. If the cell repairs the double-strand break using NHEJ or MMEJ pathways (on the left), the gene might be disrupted, or the mutation might keep the gene functional whilst making it resistant to further cleavage due to the change of the nucleases' target sequence. Made with Biorender.com

1.7 The CRISPR/Cas9 system

The Clustered Repeated Interspaced Short Palindromic Repeats (CRISPR) is the name given to a series of short repeated DNA sequences found in Bacteria and Archaea. These repeated arrays are separated by non-repetitive sequences called “spacers”, portions of a viral genome incorporated into the bacteria DNA during previous infections (*Figure 1-6*) (Mojica et al., 2009; Mojica and Rodriguez-Valera, 2016). A bacterial nuclease can employ the CRISPR sequences and the spacers to recognise and cleave new incoming phage genomic material. Recently, this system was modified and employed as a powerful DNA editing tool. The Cas9, an endonuclease found in the *Streptococcus pyogenes*, is highly suitable for tunings and has become one of the most used tools for DNA editing. The CRISPR/Cas9 system in *S.pyogenes* contains the CRISPR array, the endonuclease Cas9 and two RNA molecules: a crRNA (CRISPR RNA) and a tracrRNA (transactivator crRNA). The two molecules combine to form an RNA duplex that contains the “spacer”. Once embedded into the Cas9 nuclease, the complementary spacer allows the ribonucleoprotein to bind to the foreign genomic material. Once bonded, the Cas9 uses its two nucleasic sites to produce a blunt double-strand break, making it the most effective tool for site-specific DNA editing. The only limitation to the programmability of the Cas9 is the presence of a sequence called Protospacer Adjacent Motif (PAM) preceding the spacer sequence. The PAM for Cas9 is NGG, but other Cas proteins may have different PAM sequences. The PAM allows the bacteria to discriminate between their CRISPR library and foreign DNA. In fact, the “spacers” present in the bacteria genome lack the PAM sequence. Hence the Cas proteins cannot bind and cut them, preserving the bacteria's defensive genetic library (Jinek et al., 2012; Mojica and Montoliu, 2016).

1.7.1 A brief CRISPR chronology

The first observations of CRISPR repeats span a decade, from the beginning of the 1990s to the early 2000s (Ishino et al., 1987; Mojica et al., 2000; Jansen et al., 2002; Pourcel, Salvignol and Vergnaud, 2005; Mojica and Montoliu, 2016; Mojica and Rodriguez-Valera, 2016). It was speculated that the short, regularly spaced repeats found in bacteria and archaea were part of an evolutionary immune response against phages. In 2002 Ruud Jansen et al. observed that the CRISPR clusters were usually associated with another set of genes named CRISPR Associated Systems or CAS. They also found that the protein encoded in these CAS genes contained helicase and nuclease activity domains. In the mid-2000s, the study of CRISPR intensified, and several independent groups concluded that the “spacers” that they observed among the CRISPR arrays were portions of foreign genomic material, mainly coming from phages that probably tried to infect the cell in the past (Mojica et al., 2000, 2005; Marraffini, 2015). In 2005, Francisco Mojica and other independent groups proved that the CRISPR/Cas system was an adaptive immune response developed by prokaryotes to prevent further phage infections (Bolotin et al., 2005; Mojica et al., 2005; Pourcel, Salvignol and Vergnaud, 2005).

Moreover, it was observed that the repeats in the bacteria genome lack a 3 bp sequence present in the viral counterpart. That sequence was named the Protospacer Adjacent Motif or PAM and proved essential for binding the CAS protein. All the interest around the CRISPR/Cas system led to a major breakthrough in 2010 when Jennifer Doudna and Emmanuelle Charpentier reengineered the two RNAs needed from the Cas9 to cleave DNA into one “single-guide RNA”. Their work massively increased the usability of Cas9 as a programmable endonuclease, reducing the CRISPR/Cas system to only two components: the Cas9 nuclease and the single guide RNA (sgRNA). Furthermore, this newly engineered component could be easily edited to direct the Cas nuclease to cleave any target, with the only limitation being the presence of the PAM motif. For their work, Jennifer Doudna and Emmanuelle Charpentier were awarded the Nobel Prize for Chemistry 2020 “for the development of a method for genome editing”.

CRISPR-Cas9 Adaptive Immune System of *Streptococcus pyogenes* against bacteriophages

PHASE 1: Immunization / Acquisition:
Information to "remember" the virus

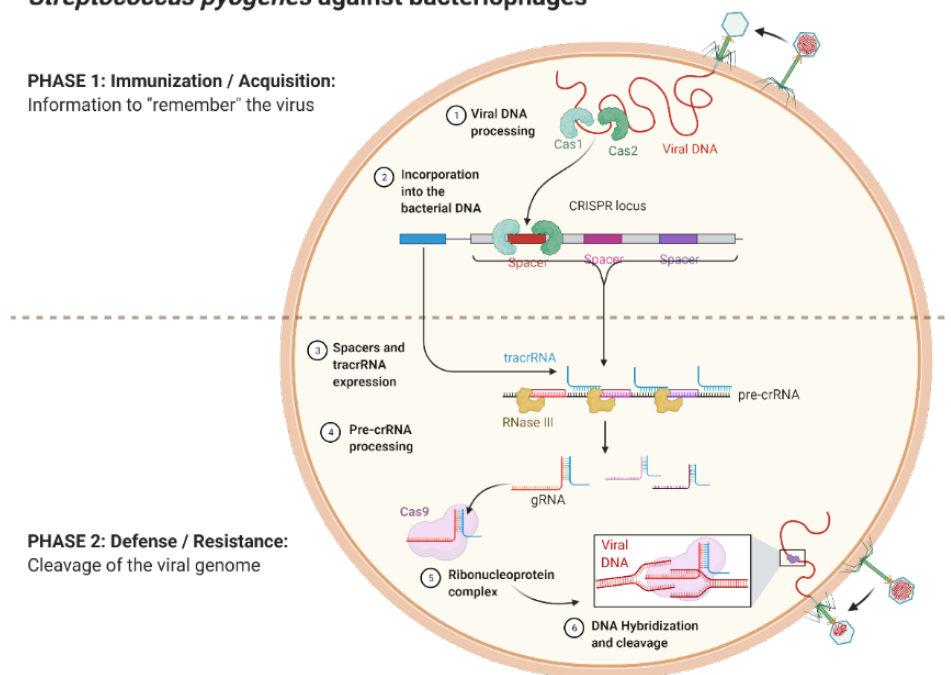


Figure 1-6. The CRISPR/Cas9 immune system in *S. pyogenes*.
Made with Biorender.com

1.8 The Homing Gene Drive

With the advent of the CRISPR/Cas technology, the ideal programmable nuclease for the genetic control of vector populations becomes available, as well as the possibility of developing a homing gene drive system (HGD). One of the first proofs-of-concept of these theories was achieved in the human malaria vector *A. gambiae*.

In 2016, Hammond et al. showed that it is possible to spread infertility into a mosquito population by targeting female fertility-related haplosufficient genes (Andrew Hammond, Galizi, et al., 2016). Employing the programmable Cas9 nuclease and relying on the HDR as a DNA-repairing mechanism, the gene drive construct, initially inserted into only one of the alleles of these genes, was copied in the opposite locus, disrupting both the copies of the gene. Using appropriate promoters to restrict the nuclease expression in the germline resulted in the allele being copied in all the gametes and inherited with a frequency close to 100% (Burt, 2003; Galizi et al., 2016). Evolutionary speaking, spreading infertility is an oxymoron since the diffusion of any gene into a population is proportional to the fertility of the individuals carrying that gene. Targeting these haplosufficient female fertility genes was the key to defying this concept. Heterozygous females with only one allele disrupted were largely fertile. The copy (homing) of the CRISPR/Cas transgene was restricted to the germline and resulted in all gametes carrying a disrupted allele containing the gene drive construct. Moreover, males remained unaffected by the disruption of one or both the alleles, contributing to the spreading of the construct into the wild type population. This gene drive system proved to be highly effective for the first generations of mosquitoes, spreading in caged populations with homing rates that peaked at 99%. Despite that, the limitations related to the selection of nuclease-resistant alleles and some fertility issues due to the leakiness of the promoter in somatic cells led to the failure of population suppression (Andrew Hammond, Kyrou, et al., 2016; Galizi et al., 2016; Hammond et al., 2017). In 2018 the same group published a paper with an improved version of their gene drive. Focusing on a better and less leaky promoter and, most importantly, placing their genetic construct into the ultraconserved *doublesex* gene, they suppressed mosquito-caged populations in eleven to twelve generations (Krzywinska et al., 2016; Kyrou et al., 2018).

Doublesex (*dsx*), first discovered in *D. melanogaster* (Hildreth, 1965), is a gene located at the bottom of the sex determination pathway. *Dsx* products consist of two sex-specific isoforms of a DNA-binding protein that regulate male and female sex differentiation. The female-specific transcript contains an exon (exon 5) strictly conserved across the *Anopheles* genus and absent

in the final male isoform. The boundary between intron 4 and exon 5, responsible for the splicing of exon 5 in females, is highly conserved and is an ideal target for a CRISPR-based gene drive. Any homozygous mutation in this exon produces sterile females with an intersex phenotype, while homozygous males remain unaffected. Consequently, intersex ‘females’ cannot spread the mutated allele in the population. This was the first proof that a gene drive could successfully wipe out caged vector populations (Kyrou et al., 2018).

Despite the importance of this success, some doubts remain about the consequences of releasing fully functional gene drive mosquitoes in the wild. Some of these concerns also involve the molecular structure of the genetic construct. Despite its efficiency, Cas9 can have off-target activity and potentially generate mutations in random regions of the mosquito genome (Zhang et al., 2015; Naeem et al., 2020). Spontaneous mutations in the genome happen constantly and are the main driving force for evolution. Even if a random Cas9 off-target activity might happen, it is improbable it would produce detectable effects on the population. Even more unlikely, it would generate some unwanted mosquito phenotype due to the vast majority presence of non-coding DNA in genomes (Dunham et al., 2012). Other concerns about gene drive applications are related to the possibility that a gene drive construct might invade closely related species to the one carrying the construct. Once released in the wild, a fully self-sustainable gene drive might be complicated to reverse back. In addition, mosquitoes do not know borders or boundaries. Hence a gene drive released in one nation might spread to neighbourhood countries that might not want it on their territory. Therefore, it is essential to minimise any possible risk when it comes to releasing the first lab-engineered gene drive mosquito in the wild. Thus, alternative strategies are constantly taken into consideration.

1.9 Self-limiting gene drives and anti-gene drive systems

The safety concerns about a possible uncontrolled spreading of fully self-sustained gene drive fostered the scientific community to develop strategies to limit or possibly eliminate gene drive constructs from a population if needed. Systems that constrain the spreading of the genetic trait could also help contain the geographical spread of GMO mosquitoes. For example, split versions of the HGD were developed (Li et al., 2020; Terradas et al., 2021). The two main components of the HGD, the Cas9 nuclease and the gRNA were unlinked and placed in two different genomic locations. As a result, only one of the two components will “drive”, while the other will exhibit the classical Mendelian inheritance. Thus, the split drive will not spread as fast as a full HGD, and the two components will be progressively removed from the population without new releases of transgenic mosquitoes. This strategy will self-limit the persistence of the GD in a vector population both in space and time and represent a safer option for field releases and trials.

Another approach to limit the spreading of a gene drive is the employment of proteins that compete with or disable the Cas9 nuclease.

Literature reports several proteins able to inhibit or interfere with the nuclease activity of Cas proteins (Marino et al., 2020). Most of them are systems developed by bacteriophages to hack and disable bacteria CRISPR defensive system.

An interesting study on the applications of these anti-Cas proteins proved that their employment could disable the super-mendelian inheritance proper of gene drive systems and avoid the suppression of mosquito-caged populations even if at the cost of reduced fitness in mosquitoes carrying the anti-Cas gene (Taxiarchi et al., 2021a).

Other anti-gene drive systems rely on employing alternative gRNAs that target and damage or excise the gene drive transgene and, in some cases, can also restore the wild type allele.

All these systems aim to increase the safety, and in some cases a potential commercial value, of gene drives applications (Wu, Luo and Gao, 2016; Marino et al., 2020).

1.10 The Y-linked meiotic Gene Drive

An alternative strategy to the HGD is the so-called Y-linked meiotic drive. During the first phases of meiosis, in mosquitoes belonging to the *Aedes* and *Culex* genera, the Y chromosome induces, through some yet unknown mechanism, breaks into the X chromosome (Burt, 2003). Mimicking this mechanism with a Y-linked spermatogenesis-expressed nuclease might lead to a self-sustainable genetic sex bias. The nuclease should target a highly repeated sequence present on the X chromosome. The high number of cuts and the lack of a homologous partner, usually used as a repair template, could make the repair process unattainable. In this way, only the Y-bearing spermatozoa will survive the process, and the Y chromosome bearing the endonuclease will exhibit a super-mendelian inheritance (*Figure 1-7*). Therefore, after the mating, the vast majority of the eggs will be fertilised by Y-bearing spermatozoa shifting the sex bias of the progeny toward a theoretical 100% of males (Burt, 2003). A male sex bias actively hampers the reproductive potential of a vector population. In addition, a lower number of female mosquitoes corresponds to fewer individuals who actively spread diseases. One of the main challenges of this strategy is that, in some species, the expression of genes mapping on the Y chromosome is inhibited during meiosis, a phenomenon known as Meiotic Sex Chromosome Inactivation (MSCI).

The first attempt to sex-bias a vector population was reported in R. Galizi et al. manuscript (Galizi et al., 2014). Here, the authors were able to shift the sex bias of the malaria mosquito *A. gambiae*. They identified an X-linked multicopy gene expressing the 28S subunit ribosomal RNA. This multi copies gene maps only on the X chromosome in *Anopheles gambiae* and therefore represents an ideal target for the X-shredding. They engineered a DNA construct expressing a modified version of the I-PpoI bacterial nuclease capable of recognising and cutting a 15 bp sequence on the multicopy gene. This destabilised version of the I-PpoI nuclease contained aimed mutations to reduce its thermal stability, half-life and specific activity. Mosquitoes expressing the wild type version of the nuclease were completely infertile due to the persistence of the protein in the spermatozoa that later on kept cutting the maternally-inherited X chromosome in the zygote (Galizi et al., 2014).

Finally, they used a portion of the 5' UTR of the male-specific gene, *beta-2 tubulin* (*B2tub*), to drive the nuclease expression. This testis-specific tubulin gene was well characterised in *Drosophila*. The production of the testis-specific tubulin starts from the late stages of larval development, before the beginning of meiosis in the evolving testes (Michiels et al., 1989; Smith et al., 2007). In mosquitoes, *B2tub* drove the expression of fluorescent markers from the

latter larval stage (L4-Pupa) (Catteruccia, Benton and Crisanti, 2005). Its expression is sustained throughout the mosquito's adult life, and it's estimated to account for almost 90% of the testes tubulin (Smith et al., 2007). In *Drosophila*, a 53-bp core region of the gene's 5' UTR located just upstream of the Transcription Starting Site (TSS) granted a robust and testis-specific expression with the tissue-specificity depending on a short core sub-region spanning just 14 bp from the TSS (Michiels et al., 1989; Santel et al., 2000).

The *B2tub-I* PpoI construct was inserted on an autosome to avoid the MSCI. The progeny produced by the modified male mosquitoes showed a male sex bias of >95%. Even if the transgene inheritance still abided by Mendelian rules, this was the first time such a result was achieved in a vector population. The next step to obtain a Y-linked gene drive (Y-drive) was placing the same transgene on the *A. gambiae* Y chromosome. Unfortunately, the X-shredding nuclease expression cassette placed on the Y chromosome was silenced entirely, in agreement with data showing the timing for the activation of the testis-specific promoter *beta-2-tubulin* lies within the time frame of MSCI (Landeem et al., 2016).

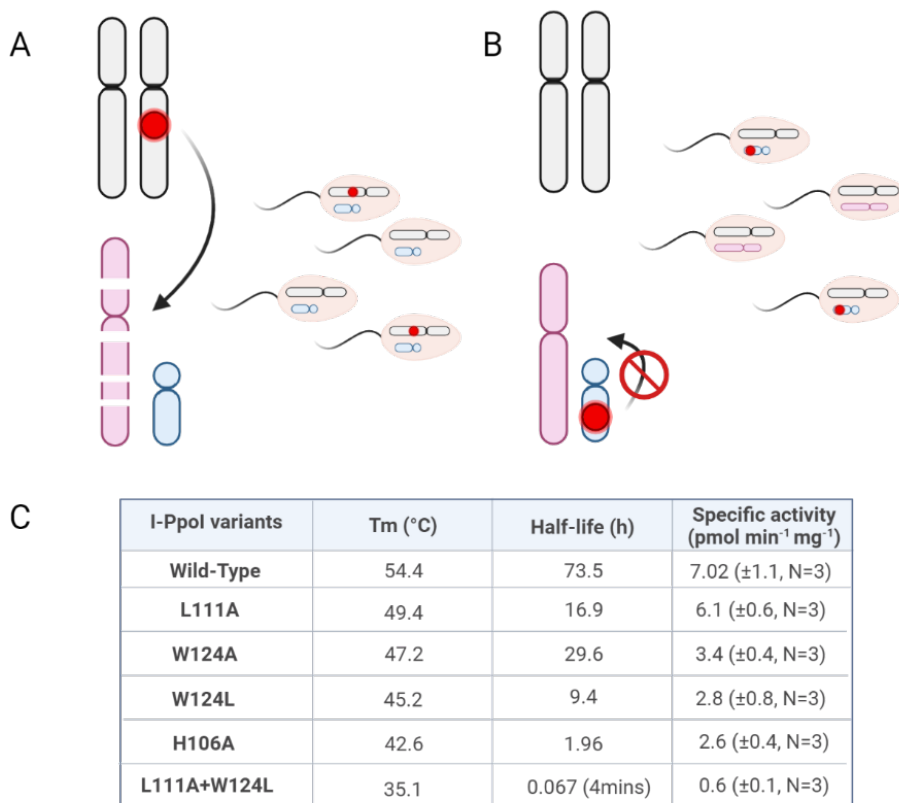


Figure 1-7 Meiotic sex distortion.

A: Representation of the sex distorter line published in Galizi et al. B: The non-working Y-linked sex distorter line. C: The different variants of the I-PpoI nuclease employed to achieve the male bias. The successful one is the W124L. Stronger variants of the nuclease produced sterility. Made with Biorender.com

1.11 Gametogenesis in mosquitoes

Meiosis is a crucial part of gametogenesis and occurs only in the germline. It is a variant of mitosis, and instead of producing two daughter diploid cells from a single diploid progenitor cell, it ends with four haploid cells. After fertilisation, the haploid cells (gametes) will form a diploid zygote. This process is needed to transmit genetic information while maintaining genetic variability across a population, ensuring fitness and adaptability to environmental changes.

In male mosquitoes, gametogenesis starts from the apical region of the testes, where germline stem cells (GSC) undergo a first cell division. This division allows the cell to self-renew, generating one new GSC and a gonoblast (GB) that will go through gametogenesis. Each GB undergoes a series of four mitotic divisions, generating sixteen diploid spermatogonia that will then start the meiotic divisions ending up to mature sperm (Klowden, 2013). The succession of all the meiotic stages starts from the apical tip of the testis, which matches the germ cells niche and proceeds across the longitudinal axle to the bottom part containing the mature sperm and *vas efferens* (Figure 1-8). In females, the GSC produce a cytoblast (CB) that will follow a similar path as the GB, but after the divisions, only one cell will be selected as an oocyte. In contrast, the others will develop in seven nurse cells that will provide support for the oocyte. At this point, the gametogenesis is arrested until the blood meal, which will provide the nutrients to develop the egg's yolk and complete the process (Hansen et al., 2014).

In males, meiosis starts with a diploid spermatocyte that undergoes a duplication of its karyotype. Then, two meiotic divisions will follow, each characterised by five phases: prophase, metaphase, anaphase, telophase and cytokinesis. Prophase I can be further divided into five substages: leptotene, zygotene, pachytene, diplotene and diakinesis. In the leptotene substage, chromosomes condense and become distinguishable. Next, during the zygotene, chromosomes start pairing, preparing for the crossing-over events except for the X and Y that, due to their lack of homology, cannot entirely pair. In the following substages, chromosomes exchange genetic material by crossing over, and prophase I end with the disgregation of the nuclear membrane. The transition to metaphase I mark the meiotic spindle formation that attaches to the chromosome's centromere. In the following anaphase I, chromosomes segregate, and the meiotic spindle dissolves. Telophase I will complete the first meiotic division by forming two new nuclear membranes and separating the two newborn cells through cytokinesis. Then, each of these diploid cells undergoes a second meiotic division that, lacking the first karyotype duplication, generates four haploid gametes.

1.12 Meiotic Sex Chromosome Inactivation

Previous studies in mammals have shown that MSCI is a specific form of transcriptional repression, part of a more general mechanism called Meiotic Silencing of Unsynapsed Chromatin (MSUC). This form of transcriptional silencing represses the chromosomes that fail to pair with their homologous partner during meiosis. Thus, transcriptional silencing occurs on both the X and Y chromosomes during the meiotic stage of spermatogenesis as they lack a homologous partner to pair with (Turner, 2007). The primary mechanism that silences the unsynapsed chromosomes is large-scale chromatin reshaping, triggered by histone marks modifications, hot-swap of histones with histone variants and DNA-repairing proteins (Turner, 2007; Landeen et al., 2016). The transcriptional activity of the sex chromosomes varies throughout gametogenesis and is tightly regulated. When germ cells initiate the meiotic process, the sex chromosomes are still transcriptionally active. They remain active during the Leptotene and Zygotene when chromosomes begin to pair, and DNA double-strand breaks trigger meiotic recombination (J. M. A. Turner et al., 2005; Turner, 2007). The silencing starts shortly after the transition to the Pachytene stage and affects the vast majority of the sex chromosomes, except for small regions that might show homology and hence pair to produce a local XY recombination (Helena Mangs and Morris, 2007; Kauppi et al., 2011; Otto et al., 2011). One of the main factors involved in sex chromosome silencing is the swap of histone H2A with histone H2AX, a variant of the H2A histone family. H2AX is involved in the recruitment of DNA-repairing proteins and becomes rapidly phosphorylated in response to a Double Strand Brake (DSB). Intriguingly, a high concentration of the H2AX variant was observed in mammalian testes compared to other tissues. Moreover, during male meiosis, its phosphorylation notably increases specifically on the X and Y chromosomes, just after the transition to the Pachytene stage, a time frame that accurately matches the start of MSCI (J. M. A. Turner et al., 2005; Turner, 2007). Once shut down, the XY couple is segregated in a separate nucleus compartment and forms the so-called sex body. The XY silencing is maintained throughout the end of spermatogenesis with MSCI-related modifications that persist even when the dephosphorylation of histone H2AX occur. The maintaining of the sex chromosome silencing after the Diplotene is considered a separate process from MSCI since different factors are involved. The chromatin at this stage is referred to as “post-meiotic sex chromatin” (Namekawa et al., 2006; Turner et al., 2006).

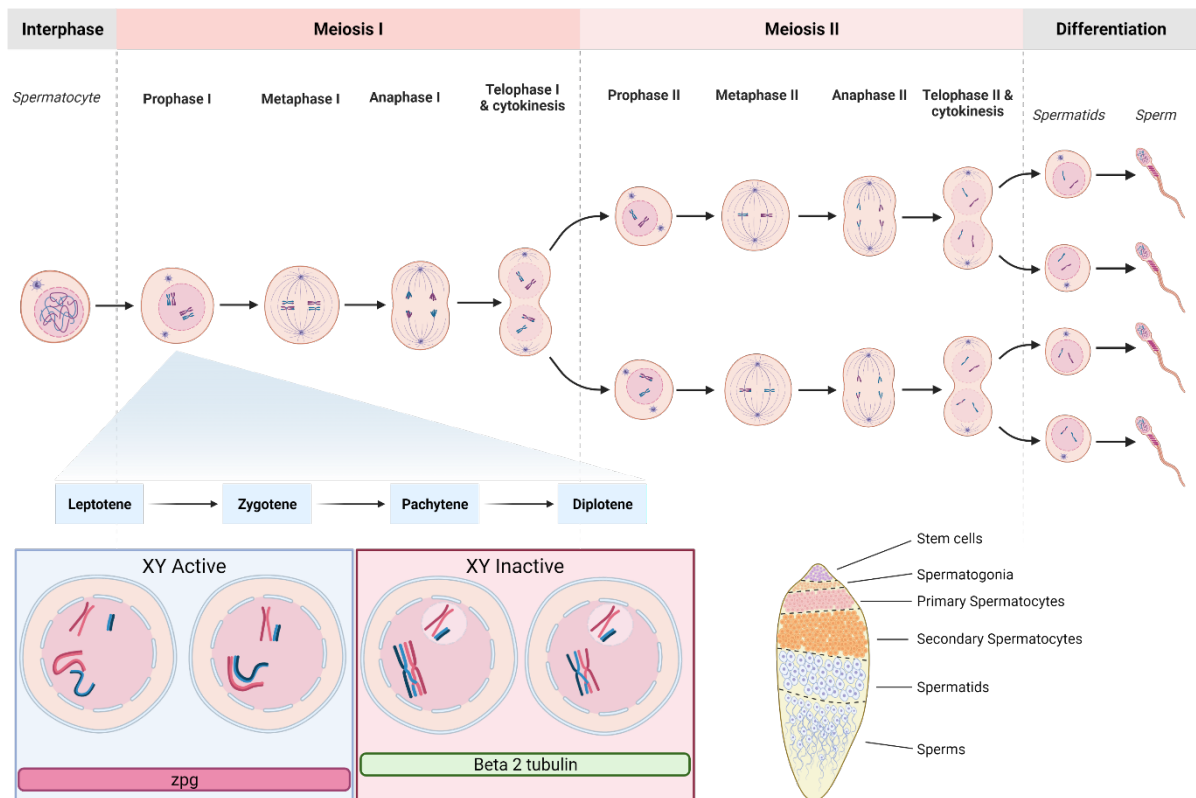


Figure 1-8. Meiotic stages and details of MSCI. The MSCI starts during the transition from Zygotene to Pachytene. Then, the two sexual chromosomes fail to pair entirely and become silenced in a separate nucleus compartment, forming the so-called “sex body”. On the bottom right: a representation of the distribution of the meiotic stages in the *A. gambiae* testes. Made with Biorender.com

1.13 The Importance of Epigenetics.

In 1942 Conrad Hal Waddington proposed the term “Epigenetics” to describe the mechanisms “by which the genes of the genotype bring about phenotypic effects” (Waddington, 2012). Essentially, epigenetics tries to explain how populations of cells with identical genomic sequences can differentiate into various types with even more different tasks. The library of genes in any eukaryotic genome is tightly regulated to produce significant variation in cell structures and functions. Epigenetics is also correlated with adaptation to environmental changes. To a certain extent, genes can be regulated in response to different environmental stimuli, including changes in environmental parameters like temperature, humidity or salinity. This ability to regulate gene expression “on the run” was named phenotypic plasticity (Kelly, Panhuis and Stoehr, 2012; Xue and Leibler, 2018) and is linked to changes in physiology, behaviour and morphology. The three main mechanisms of epigenetic regulation are DNA methylation, histone tail modifications and noncoding RNAs.

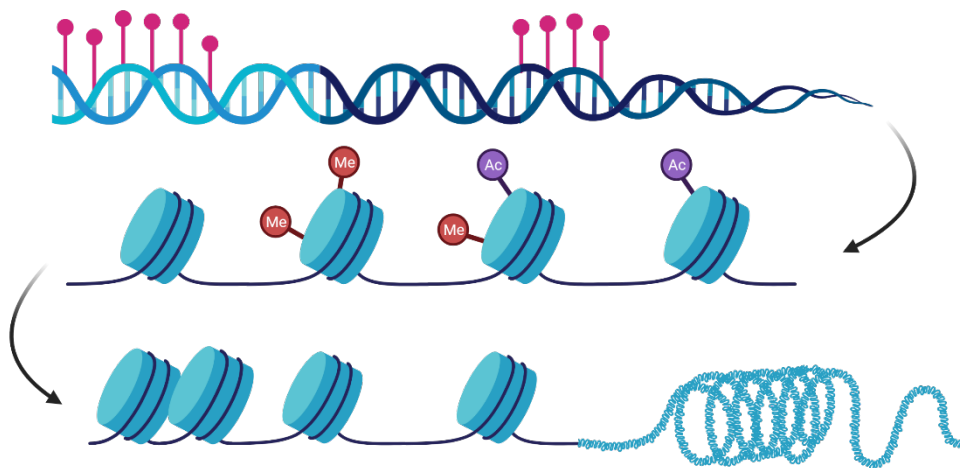


Figure 1-9. Epigenetic levels.

From the top: DNA CpG methylation (red dots); histone tail modifications; chromatin higher structural organisations. Made with Biorender.com

The most studied mechanism of epigenetic regulation is DNA methylation (Moore, Le and Fan, 2013; Messerschmidt, Knowles and Solter, 2014; Chiarella, Lu and Hathaway, 2020). The addition of a methyl group on the DNA strand usually occurs on the 5'-cytosine-phosphate-guanine-3' (CpG) dinucleotides. These dinucleotides may be disproportionately present in

some DNA regions, forming the so-called CpG islands. A methylated CpG island usually correlates with transcriptional repression (Pikaart, Recillas-Targa and Felsenfeld, 1998; Messerschmidt, Knowles and Solter, 2014; Provataris et al., 2018). Gene repression might be achieved by inhibiting transcription facilitators' binding or recruiting repression factors. The enzymes that catalyse the transfer of the methyl group to the C5 of the cytosine (5mC) belong to two classes: *de novo* DNA methyl-transferase (DNMT3) and maintenance methyl-transferase (DNMT1). Diptera seems to lack DNMT3 but shows methylation patterns inherited by their progeny. Hence, it was speculated that the maintenance activity of DNMT1 is enough to sustain and transfer the information coded in the Diptera methylome or that DNMT1 might have acquired some “*de novo*” methylation activity (Glastad, Hunt and Goodisman, 2018; Provataris et al., 2018).

Histone tail modifications are probably the broader and most unknown mechanism of epigenetic regulation. Histones are the primary component of chromatin; they are globular proteins rich in Lysine and Arginine with an N-terminal tail that protrude from the globular domain. Their positive charges interact with the negatively charged DNA scaffold through non-covalent bonds. Although there are many types of histones, the “standard” histones are H2A, H2B, H3 and H4. H2A and H2B form dimers, while H3 and H4 form tetramers. Two H2A/H2B dimers and an H3/H4 tetramer assemble to form an octamer named nucleosome.

Chromatin is a complex structure made of DNA and proteins. Nucleosomes constitute the building blocks of chromatin, representing the basic level of chromatin organisation. They align in a beads-on-a-string array, where 147 bp of DNA surrounds each nucleosome. The stretch of DNA connecting two nucleosomes is called linker DNA, and it's usually bound by another and slightly smaller histone, the H1, that connects and compacts the nucleosome array (Luger, Dechassa and Tremethick, 2012; Manning and Manning, 2018). The nucleosome array represents the first order of chromatin organisation, also known as the 10nm fibre. Chromatin has at least two higher-order of organisations: 1) the 30nm fibre produced by nucleosome-nucleosome interactions and 2) interactions between 30nm fibres, including folding, pairing and looping (Luger, Dechassa and Tremethick, 2012). This organisation allows the storage of meters-long DNA in the cell's nucleus, which spans a few micrometres (*Figure 1-9*)

Moreover, the level of DNA compaction in the nucleus can be locally adjusted to limit or enhance the access to transcription factors and or polymerases, resulting in efficient gene regulation (Ahmad and Henikoff, 2001; Berger, 2007; Bowman and Poirier, 2015; Dugar et al., 2018; Glastad, Hunt and Goodisman, 2018). However, the link between DNA sequence and nucleosome position is still relatively unknown, considering that the nucleotide sequence

is not a strict requirement for nucleosome positioning. On the contrary, nucleosomes can be very mobile and dynamic elements, binding and releasing DNA every 25ms in a process known as “nucleosome breathing” (Berger, 2007; Bowman and Poirier, 2015; Isaac et al., 2016). Furthermore, the hot-swap of some histones with specific alternative variants might reduce the affinity of the nucleosomes for the DNA, promoting nucleosome mobilisation or destabilising the nucleosome itself and leading to its dismantling. Nevertheless, some Nucleosome Positioning DNA Sequence patterns (NPS) were identified, implying some underlying and yet unclear mechanisms connecting nucleosomes and DNA sequence (Pranckeviciene et al., 2020). Histone’s tails stick out of octamer and are the substrate for the additions of chemical groups that form the so-called “histone code”. There are at least twenty known histone tail modifications (*Tab 1*) (Zhao and Garcia, 2015). Histone-modifying enzymes can add chemical groups to multiple amino acids of the same histone tail. More than a chemical group can be added to the same amino acid, exponentially increasing the complexity of the epigenetic code (Kouzarides, 2007; Bannister and Kouzarides, 2011).

1.13.1 Acetylation and Methylation, two histone modifications

I started focusing on two of the most studied histone tail modifications for this study: methylation and acetylation. The artificial modification of these marks proved to be effective in altering chromatin structure and affecting gene expression (Ali et al., 2017; Li et al., 2017; O’Geen et al., 2017; Pulecio et al., 2017; Waters et al., 2018). Up to three methyl groups can be placed on the same tail’s residue, creating mono-, di- and tri-methylations (*Tab 1*). The modification can result in active or repressed gene regions according to the histone targeted, and the number of methyl groups added (*Tab 1*). On the other hand, acetylation is a modification exclusively linked with transcriptionally active regions independently of the histone tail residue targeted. The transfer of an acetyl group from an acetyl-Coenzyme A to the histone tails neutralizes the positive charges of the histone, reducing the affinity for the negatively charged DNA phosphate scaffold. These acetylated portions of chromatin show less compacted nucleosome arrays and increased transcription activity (Oki et al., 2004; Hilton et al., 2015). Moreover, it was shown that the artificial modification of histone marks through the addition of acetylations groups positively regulated gene expressions in plants and animals, including arthropods (Hilton et al., 2015; Ali et al., 2017; Manning and Manning, 2018; Roca Paixão et al., 2019).

Histone modifications and dynamic chromatin regulations involve a plethora of enzymes, including methyl- and acetyl-transferase, histone chaperones and ATP-dependent chromatin remodelers. Once the histone code is “written” on one nucleosome, it can be read by specific “reader-writer” enzymatic complexes that can “write” nearby nucleosomes spreading the modification throughout the adjacent chromatin from nucleation points across the chromosome in a wave-like fashion (Bowman and Poirier, 2015; Allshire and Madhani, 2018).

Modification	Histone Targeted	Known effect
Acetylation	H2A,H2B,H3,H4	Transcriptional activation
Mono-ADP-ribosylation	H1	Involved in neurotrophic activity
Biotinylation	H2A,H4	Cell proliferation, gene silencing, cellular response to DNA damage
Butyrylation	H2A	
Citrullination	H1	Cellular reprogramming / nucleosome binding
Crotonylation	H1,H2A,H2B,H3,H4	
Formylation	H1,H2A,H2B,H3,H4	
Glutathionylation	H3	
2-hydroxyisobutyrylation	H1,H2A,H2B,H3,H4	
Isomerization	H3	
Malonylation	H2A,H2B,H3	
Monometylation	H1,H2A,H2B,H3,H4	Gene activation / repression,
Dimethylation	H1,H2A,H2B,H3,H4	Gene activation / repression,
Trimethylation	H1,H2A,H2B,H3,H4	Gene activation / repression,
O-GlcNAcylation	H2A,H2B,H3,H4	
Hydroxylation	H1,H2A,H2B,H4	
Oxidation	H4	
Phosphorylation	H1,H2A,H2AX,H2B,H3,H3.3,CEN-H3/CENP-A,H4	DNA damage response, telomere silencing, chromosomal stability, apoptosis, transcriptional activation,
SUMOylation	H1,H2A,H2B,H4	Transcriptional repression,
Ubiquitination	H2A,H2AX,H2B	DNA damage response, Polycomb silencing, gene activation, telomere silencing

*Table 1-1 Histone tails modification and their know effects.
As reported by (Tan et al., 2011)*

1.14 The gypsy retro-transposable element as a chromatin insulator

In eukaryotes, chromosomes are organised in heterochromatic and euchromatic chromatin domains, usually characterised by different sets of nucleosome modifications. This organisation reflects the cell's need to regulate temporally and spatially the expression of genes on the same chromosome (West, Gaszner and Felsenfeld, 2002; Hong WEI, Pei LIU and Chuan LIANG, 2005; Dixon et al., 2012; Ghirlando et al., 2012). At the boundaries of these domains lie sequences that keep this organisation discrete. These insulator sequences may act differently, from blocking the interaction between enhancer and promoter to behaving like a barrier that stops heterochromatin expansion (Bi and Broach, 1999; West, Gaszner and Felsenfeld, 2002; Oki et al., 2004; Gaszner and Felsenfeld, 2006; Dhillon et al., 2009; Yang and Corces, 2011). Several DNA sequences with insulator activity were characterised across different species, including in the malaria mosquito *A. stephensi* and *A. gambiae* (Tubio et al., 2011; Carballar-Lejarazú, Brennock and James, 2016). In insects, the *Drosophila* retro-transposable element, known as *gypsy*, is one of the most studied insulator sequences. It consists of a series of repeated motives targeted by the zinc-finger protein “suppressor of Hairy Wings” Su(Hw). Su(Hw) recruits another protein, the “modifier of mobile dispersed genetic element 4” (mod(mdg4)). Once bound to Su(Hw), mod(mdg4) can self-associate, setting the ground for DNA loop formation (Gerasimova, Byrd and Corces, 2000; Byrd and Corces, 2003; Carballar-Lejarazú, Jasinskiene and James, 2013) (*Figure 1-10*). These loops may co-localise on the nuclear matrix, and when in the same spot, they form the so-called insulator bodies bringing together chromosomal regions usually far apart (Chen and Corces, 2001; Byrd and Corces, 2003; Schoborg et al., 2013). However, the *gypsy* array is not the only sequence that su(Hw) recognises. The zinc finger protein was linked to other binding sites and also involved in recruiting other partners, including CP190, known to open chromatin through nucleosome depletion, counteracting the spreading of the heterochromatin and halting gene silencing (Ali et al., 2017; Melnikova et al., 2019). In the *A. gambiae* genome, the *gypsy* elements constitute about 1-1.5% of the entire genome, even if a good share of them results degenerated and mainly heterochromatic (Tubio et al., 2011). Interestingly, the distribution of the *gypsy* elements on the *A. gambiae* sexual chromosomes is disproportionate compared with the autosomes. The Ty3/*gypsy* transposable element is present on the *gambiae* X chromosome in more than 5.6 elements per megabases (Mb) against 2.9 to 3.7 elements per Mb on the autosomes. Its density increases in the pericentric region of each chromosome, particularly in the pericentric heterochromatin, probably due to the lower recombination rate in the premises of the

centromere. Only two *gypsy* elements were detected on the Y chromosome, but it was impossible to locate them correctly due to the lack of an assembly for the Y chromosome in *A. gambiae* (Tubio et al., 2011).

Despite being widely adopted in *Drosophila* to mitigate the positional effect variegation of transgenes, few papers reported using the *gypsy* elements as transgene insulators in the *Anopheles* genus (Lynd and Lycett, 2012; Carballar-Lejarazú, Jasinskiene and James, 2013). From the data available, the employment of the exogenous *Drosophila gypsy* element in *Anopheles* seems to stabilise and increase gene expressions whilst reducing the positional effect variegation (Carballar-Lejarazú, Jasinskiene and James, 2013). Moreover, orthologues of the three main *gypsy*-binding proteins were identified in the *Anopheles* and *Culex* genera as further proof that a *Drosophila*-like *gypsy* transposon system is in place also in the mosquito genome and probably was inherited from a common ancestor (Carballar-Lejarazú, Jasinskiene and James, 2013; Carballar-Lejarazú, Brennock and James, 2016). To our knowledge, the *gypsy* insulator elements were never employed on the Y chromosome to stabilise transgene insertions and hence were never tested to counteract MSCI.

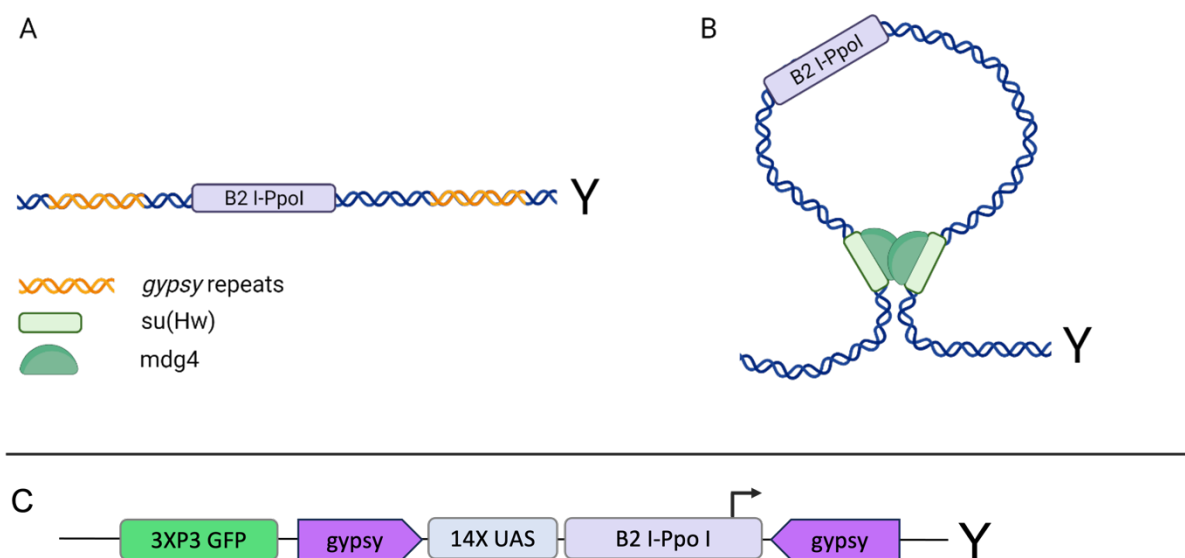


Figure 1-10 The *gypsy* retrotransposable element as chromatin insulators. A and B show how two *gypsy* motives can form loops of DNA to isolate a given transgene. C: schematic representation of the construct inserted on the *A. gambiae* Y chromosome. Made with Biorender.com

1.15 Activating transcription via Gal4 and catalytically defective Cas9 fused to HAT

Following the discovery of the CRISPR/Cas9 system, the employment of Cas9 as a DNA site-specific nuclease was just one of the many possibilities. Further studies showed that the two Cas9 catalytic domains (RuvC and HNH) could be inactivated by replacing just three nucleotides. The resulting nuclease-dead Cas9 (dCas9) loses its nuclease activity but retains the ability to search and bind target sequences when provided with the appropriate sgRNA. This feature massively expanded the versatility of the dCas9, now used to interfere with gene expression (Gilbert et al., 2013; Larson et al., 2013), to deliver transcriptional activators or repressors (Gilbert et al., 2013; Maeder et al., 2013; Chavez et al., 2016), to visually highlight target genes by tethering it to one or more fluorescent proteins (Tanenbaum et al., 2014; Deng et al., 2015), and to deliver epigenetic editors to edit the “histone code” and reshape chromatin (Hilton et al., 2015; Braun et al., 2017; Brocken, Tark-Dame and Dame, 2017; Pulecio et al., 2017; Waters et al., 2018).

A study performed on the Y chromosome of *D. melanogaster* showed that a small DNA-binding protein like Gal4 could interfere with nucleosome formation and keep active a usually silenced reporter gene (Ahmad and Henikoff, 2001). In *Drosophila* imaginal discs, the Gal4 was placed under the control of several promoters acting in different larval development moments and with varying expression strengths. The Gal4 targeted a UAS-HSP70-GFP reporter gene set on two distinct loci of the *Drosophila* Y chromosome characterised as differently heterochromatic. When under the control of the earlier (embryonic) promoter, the Gal4 was able to bind the reporter gene and keep it active in both reporter lines during the entire larval development. On the contrary, when the Gal4 was driven by promoters expressed in the later larval stages, the expression of the GFP was sustained only in the euchromatic reporter line. The authors pointed out that during the third larval stage, the *Drosophila* eye cells start to differentiate and then multiply. The Gal4 expressed at this stage is less likely to find its cognate binding site available due to the heavy nucleosome deployment triggered by differentiation and mitosis.

On the contrary, when expressed in the embryo, the Gal4 finds a more favourable environment and the probability of finding its binding site increases. Once bound to its target, the Gal4 seems to keep the reporter gene accessible throughout the larval development. This suggests that the presence of DNA-binding proteins interferes with mitosis-induced heterochromatin spreading. The study points out that the occupancy on the DNA strand by the Gal4 physically

interferes with nucleosome formation. The efficacy of the derepression of the reporter genes was described as a function of the strength of the promoter employed to drive Gal4 expression and its activation time (Ahmad and Henikoff, 2001). In detail, this study highlighted that to keep the reporter gene functional, the binding of the Gal4 should happen before chromatin condensation. Studies performed with dCas9 reached similar conclusions, suggesting that dCas9 produces a steric hindrance to nucleosome assembly and defines a nucleosome-depleted region spanning around one hundred base pairs on both flanks of the binding site (Barkal et al., 2016; Horlbeck et al., 2016).

Even if dCas9 seemed to be able to interfere with nucleosome formations, its presence alone might not be enough to activate the transcription of strictly silenced genomic regions. Previous studies showed that the local delivery of transcriptional activators like the virus-derived VP16 transcription factor or epigenetic editors like histone acetyltransferases or demethylase was highly effective in activating genes normally repressed (Maeder et al., 2013; Hilton et al., 2015; Choudhury et al., 2016; Ali et al., 2017; Chiarella et al., 2019).

1.16 The “Pcaf” Histone acetyltransferase

In literature, there are plenty of examples of epigenetic editors that, once locally delivered, can activate or silence genes, and nowadays, most of them use a dCas9 as the primary carrier (Chavez et al., 2016; Ali et al., 2017; Pulecio et al., 2017). Ali et al. showed how the local delivery of the histone acetyltransferase Gcn5 activates a polycomb-repressed gene in *Drosophila* and depletes the surrounding nucleosomes of H3K27me3 marks (Ali et al., 2017). In *A. gambiae*, the closest orthologue of Gcn5 is the annotated gene AGAP004434 coding for the histone acetyltransferase Pcaf (p300/CBP Associated Factor). The Pcaf is a histone acetyltransferase that primarily targets lysines (mainly H3K14, H4K8 and H4K16) and that has been characterised, in mice and other vertebrates, as active during embryogenesis. It contains three main domains, other than the acetyltransferase domain. There is one highly conserved Bromodomain that recognises and binds lysines, and a third domain (Pcaf-N), that is not retained in lower organisms, with a ubiquitin ligase activity (**Error! Reference source not found.**) (Poux and Marmorstein, 2003; Koutelou, Farria and Dent, 2021). The overall structure of the Pcaf gene is extremely conserved throughout the *Anopheles* genus (**Error! Reference source not found.**).

This work primarily aims to activate transcription from the highly heterochromatic *A. gambiae* Y chromosome. The target gene contains the before-mentioned expression cassette for the X-shredding nuclease (I-PpoI). A precisely timed activation of this transgene is crucial to obtain a sex bias. Early activation of the nuclease will shred the X chromosome during the pre-meiotic stages, leading to the meiotic arrest and making male mosquitoes sterile (unpublished). At the moment, the only promoter that could drive the expression of the X-shredding nuclease in the appropriate time frame is a portion of the 5'UTR of the late meiotic *beta-2 tubulin* gene (Galizi et al., 2014). Hence, the strategies tested in this study aim to allow the *B2tub* promoter to be active in its own time frame. To ultimately achieve the Y-linked gene drive, any gene expressing an “activator molecule” able to rescue *B2tub* activity should be placed on the Y chromosome and expressed before the meiotic silencing. The *vasa-2* regulatory region seems the most promising candidate for this task. In our lab, it could drive both a GFP (Bernardini et al., 2014) and the I-PpoI (unpublished) from the Y chromosome. But as already anticipated, the expression of the I-PpoI under the *vasa-2* leads to male sterility. Therefore, placing an “activator” under the control of this regulatory region poses a challenge in how to avoid a premature expression of the *B2tub* I-PpoI construct. The pre-meiotic delivery of polymerase recruiters like VP16/VP64 will most likely lead to early activation of the X-shredder cassette

since they directly recruit the main elements of the transcriptional machinery. To avoid that, choosing an epigenetic editor that keeps the Y-linked transgene accessible to transcription factors during the MSCI seems more appropriate. If it is possible to keep the Y chromosome region encompassing our X-shredder expression cassette accessible to transcription factors, the *B2tub* promoter might be able to drive the expression of the I-PpoI nuclease within its own time frame.

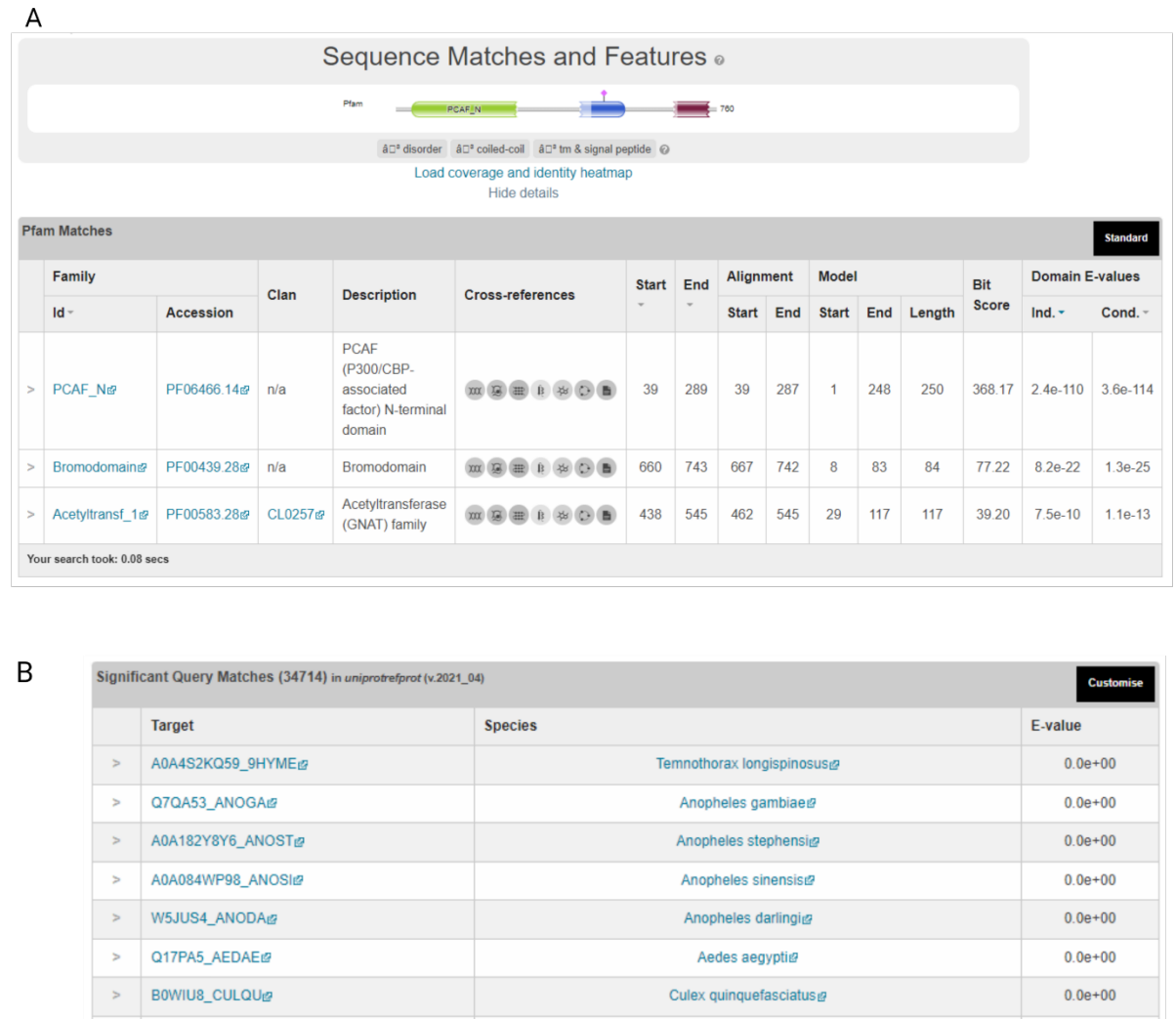


Figure 1-11 Structure and conservation score of the *Pcaf* histone acetyltransferase.
A: *Pcaf* domains structure. B: Examples of the high conservation of the *Pcaf* across several species as reported on HMMER

2 Results

2.1 Strategy 1 – Developing a Gal4-UAS system to block heterochromatin spreading

To develop the Gal4-UAS reporter system, I started with cloning the constructs, and once ready, I immediately started the injection process. Initially, I designed the plasmid with an attB/attP system for site-specific integration in one of the autosomal docking lines available in our lab. After a few months of injections, I could not retrieve any transgenic mosquitoes. I then decided it would have been more productive to shift to random integration in the genome employing the PiggyBac transposable system. The PiggyBac system employs a transposase that promotes the insertion of genetic elements flanked by specific sequences named Inverted Terminal Repeats (ITR). ITRs usually span from 15bp to 30bp. They are found in transposon systems and are recognised by transposases promoting the mobilisation of the transposon. The transposase catalyses the excision of the transposon and its insertion in the genome at any TTAA site. In the same way, a genetic element flanked by ITR can be excised by a donor plasmid and integrated in any suitable TTAA site in mosquito genome with an increased probability compared to site specific integration. Unfortunately, before the start of the new cloning phase, the Covid-19 pandemic caused the stop of all scientific activities in our laboratory. Just before the pandemic started, I retrieved the first transgenic mosquitoes from the other two strategies presented in this work. So I decided to put the Gal4 strategy on hold until the end of the pandemic. Once I was allowed to restart my scientific work, I still had to face restrictions to access the lab and the insectary and I had to prioritise my work. Hence I decided to focus on the other two strategies that already gave me transgenic mosquitoes and abandon the development of the Gal4-UAS system in *A. gambiae*.

2.2 Strategy 2 - The design of a CRISPR system to counteract MSCI

The aim of this second strategy was to design a CRISPR system for the site-specific delivery of an epigenetic editor. The first element of this system is the carrier, the nuclease-dead Cas9 (dCas9). Our lab has already used with success and for several purposes a human codon-optimised Cas9 in *A. gambiae* (Andrew Hammond, Kyrou et al., 2016; Hammond et al., 2017; Kyrou et al., 2018). There's an abundance of human codon-optimized dCas9 on the market, but most of them have a different codon code from the one we have used. To rule out the possibility that a different codon code of the dCas9 available in literature could affect the dead nuclease expression in *A. gambiae*, I mutagenised the Cas9 already tested in our lab. The standard Cas9 produces double-strand breaks on the DNA helix using two catalytic domains (RuvC and HNH). Therefore, I introduced two amino acid substitutions (D10A and H840A) to deactivate both the catalytic domains and disable the nuclease (Brocken 2018). To drive the expression of the dCas9 in the testis, I used the pre-meiotic promoter *vasa-2*. This regulatory element drives the expression of an RNA helicase in the pole plasm and germ cells. Hence, it has an early gonads-specific expression (*Figure 2-2*) (Papathanos et al., 2009). Having the dCas9 early expressed during gametogenesis is crucial to allow its binding to the target sites on the Y chromosome before the beginning of MSCI. The *vasa-2* promoter was already used in our lab to create the first generation of gene drive mosquitoes (Hammond et al., 2017). However, it was not the best promoter to drive a Cas9 in a gene drive system due to a leaky expression in the soma and maternal deposition in the embryo (Andrew Hammond, Kyrou, et al., 2016). This misexpression of the Cas9 led to the production of unwanted cuts in the genome that compromised fertility and generated gene drive-resistant alleles (Andrew Hammond, Galizi, et al., 2016; Andrew Hammond, Kyrou, et al., 2016).

The dCas9 cannot cut and produce damage. Moreover, it targets a gene (the *beta-2 tubulin*) endogenously expressed at extremely high levels (*Figure 2-2*). Therefore even if the dCas9 will most likely bind and deliver the epigenetic activator to the endogenous *B2tub*, it is unlikely to upregulate a gene already expressed at such high levels. Moreover, the maternal deposition produced by the *vasa-2* promoter proves that the dCas9 will be present in the cell throughout all the meiotic stages, particularly during the later ones, in which it should bind and activate the transgene on the Y-chromosome.

The next step was to tether the dCas9 with the chosen epigenetic editor. I selected the Pcaf histone acetyltransferase using the information from the Vectorbase database (Giraldo-Calderón et al., 2015). First, I looked for orthologues of the *Drosophila* histone

acetyltransferase Gcn5 in the *A. gambiae* genome (Ali et al., 2017), and I identified the gene AGAP004434 annotated as Pcaf histone acetyltransferase. As expected, the protein exhibited the three-domain structure proper of high Eukaryotes. I extracted the total RNA from our *A. gambiae* G3 strain and retro-transcribed it to obtain the Pcaf spliced coding sequence. I designed a few sets of primers to amplify the 2280 bp codifying for the wild type protein and then verified it through Sanger sequencing. I identified only a few small nucleotide polymorphisms (SNPs) compared to the annotated genome, but none of them affected the amino acid sequence. To accurately identify the location of the three Pcaf domains on the peptide chain, I compared the results of the Sanger sequencing with the annotations on Vectorbase (Giraldo-Calderón et al., 2015). I then checked the protein structure on HMMER, a software designed to compare and detect sequence homologs against several databases (www.hmmer.org), which showed very high conservation of the HAT across several mosquito species (**Error! Reference source not found.11**). Once I assessed the correct position of the three main domains, I amplified the “CORE” sequence of the protein, containing the two main domains responsible for the acetyltransferase activity and excluding the first 1272 bp encompassing the PCAF-N domain (Figure 2-1). Portions of the polypeptide chain of these HATs proteins that retained only these “core” domains (Bromodomain and acetyltransferase domain) have proved to be more efficient in the acetyltransferase activity (Hilton et al., 2015).

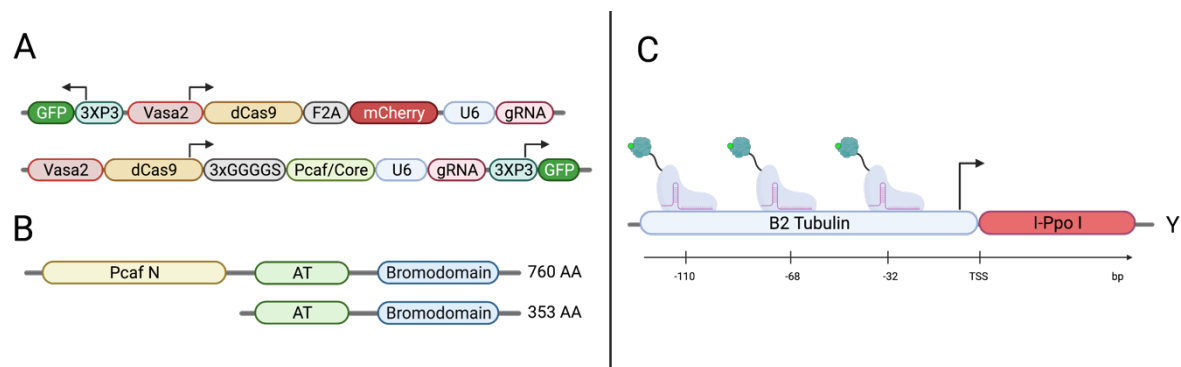


Figure 2-1 Structure of plasmid injected, Pcaf domain structure and B2tub target sites.

A: The two constructs expressing the dCas9 and the binding site on the B2tub. The one on the top produced the ectopic expression of the GFP in the testes. The bottom one shows the dCas9-Pcaf fusion. B: The long (wild type) and short (core) versions of the Pcaf. C: schematics of the Y-linked I-Ppo I transgene and the binding sites for the dCas9 on the B2tub promoter. Made with Biorender.com

I then fused the dCas9 to the two versions of the PCAF, the wild type and the truncated (CORE), through a 3X(GGGGS) linker. This small peptide should keep the dCas9 and the Pcaf components apart, avoiding the risk of miss-folding due to their proximity and providing some extra degrees of freedom to Pcaf, potentially increasing its range of activity (Chen, Zaro and Shen, 2013). Each of these two main plasmids contained a Golden Gate cloning site for the insertion of an appropriate gRNA that will direct the dCas9-Pcaf and dCas9-Pcaf^{CORE} fusions in a target-specific manner. As a marker for the visual identification of the transgenic mosquito, I employed a Green Fluorescent Protein (GFP) under the control of an array of three P3 promoters (3xP3 or *pax*). The *pax* promoter drives the expression in the nerve cells. Therefore, the transgenic larvae and the pupae show fluorescence in the eyes and the spinal cord. Finally, to choose the binding sites for the dCas9 on the late meiotic *B2 tubulin* (*B2tub*) regulatory region, I compared the *A. gambiae* and the *D. melanogaster* promoters. In *D. melanogaster*, it was shown that the essential portion for testis-specific expression of this regulatory region is enclosed in the last 53 bp (Michiels et al., 1989). The regulatory region that successfully drove the I-PpoI in the *A. gambiae* autosomal sex-distorter line included a 496 bp portion of the 5'UTR of the endogenous *beta-2 tubulin* gene, annotated as AGAP008622 (Catteruccia, Benton and Crisanti, 2005; Galizi et al., 2014). When tested on the Y chromosome to drive the X-shredder expression cassette, the same regulatory sequence failed to show any sex bias (unpublished). I located four potential binding sites for the dCas9 on this truncated *B2tub* 5' UTR sequence at -32 bp, -68 bp, -110 bp and 511 bp from the Transcription Starting Site (TSS). All the binding sites had a low off-target score and were detected using Benchling (<http://www.benchling.com>). Testing different gRNA would inform on the strength of the transactivation compared to the distance from the TSS. An alternative approach would be using more than one gRNA at a time. This might allow a more robust opening of the chromatin, creating a small array of dCas9-Pcaf bound to the promoter, but also increasing the risk of interfering with transcription factors and potentially hampering the activity of the RNA polymerase. For these reasons, I chose the single gRNA approach.

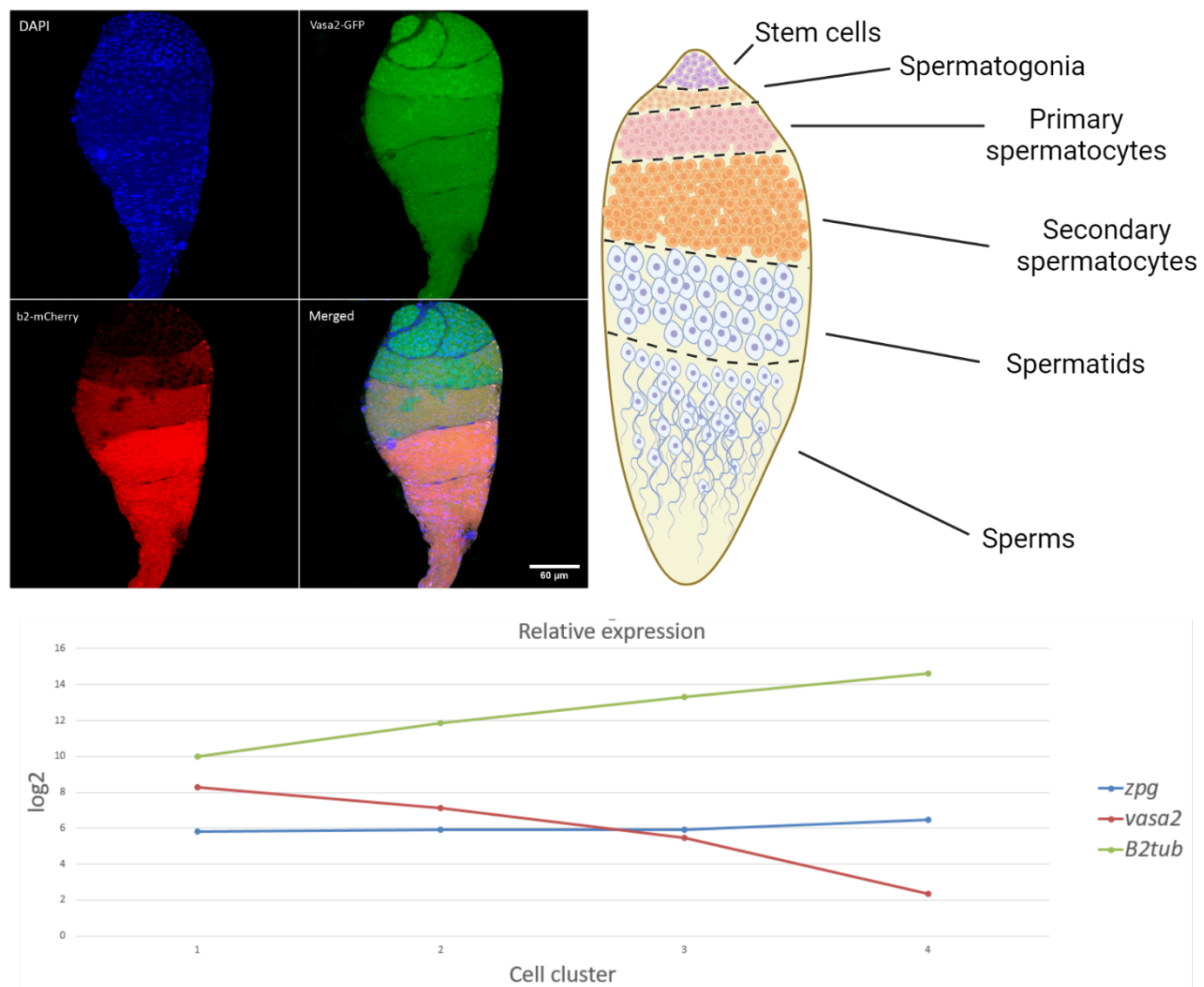


Figure 2-2 Expression patterns of the *vasa-2* and *B2tub* regulatory regions. The picture on the left shows the expression pattern of the *vasa-2* (GFP) and *B2tub* (mCherry) regulatory regions. It is evident how *vasa-2* is an early meiotic promoter while the *B2tub* is active in the later stages of meiosis. The localisation of the expression of the *zpg* promoter overlaps with the *vasa-2* (not shown in the picture), as reported by (Terradas et al., 2022). The graph on the right highlights the relative expression strengths of the three regulatory regions employed in this study. The expression was measured in cells obtained from whole mosquito testes and sorted into four main cell clusters (via FACS sorting) relative to spermatogenic stages (from 1 early spermatogenesis to 4 mature sperm), as reported by (Taxiarchi et al., 2019). The *vasa-2* is significantly expressed in the GSC. Its expression progressively declines with the progression of the meiotic process. On the other hand, the *zpg* shows a consistent expression throughout the process. Finally, the *B2tub* shows the most robust expression of the three, constantly increasing its yield till the latter meiotic stages. Credits: Matteo Vitale, Biorender.com

2.3 The dCas9 kinetics during gametogenesis

2.3.1 Generation of the dCas9 transgenic lines

In this project, I generated ten transgenic lines in total, nine of which express a dCas9. The first three lines contained a simple dCas9 fused with a mCherry fluorescent protein. These lines targeted the *B2tub* regulatory region that drove the expression of the “X-shredder” cassette from the Y-chromosome at -62 bp, -110 bp and -511 bp from the TSS. These lines were named D162, D1110 and D1511, respectively. I also created an extra line without gRNA as a dCas9 donor for testing purposes named D0. The lines carrying the dCas9-3X(GGGGS)-Pcaf were named using the “D2” prefix followed by the number of base pairs from the TSS of their target region. For the generation of these lines, I focused only on target sites within the *B2tub* 5’UTR portion employed to drive the I-PpoI on the Y chromosome. Therefore I didn’t employ the gRNA -511. Instead, I designed an extra gRNA targeting a binding site at -32 bp to test the system on the core sequence of the *B2tub* UTR described in the literature (*Figure 2-1*) (Michiels et al., 1989; Santel et al., 2000). Hence, these new lines were named D232 and D232^{CORE}, and they expressed a dCas9-3X(GGGGS)-Pcaf and dCas9-3X(GGGGS)-Pcaf^{CORE}, respectively. I used the same nomenclature system for the lines expressing the Pcaf and targeting the *B2tub* at -62 bp and -110 bp. Hence, these lines were named D262, D262^{CORE} and D2110^{CORE}. The injections for the generation of line D2110 took place in the very last part of this project. Therefore I could generate only the line carrying the “core” version of the Pcaf, and the line itself was tested only via qRT-PCR. All the plasmids injected contained Inverted Terminal Repeats (ITR), flanking the constructs for random integration in the genome in the presence of a PiggyBac transposase (*Figure 2-3*).

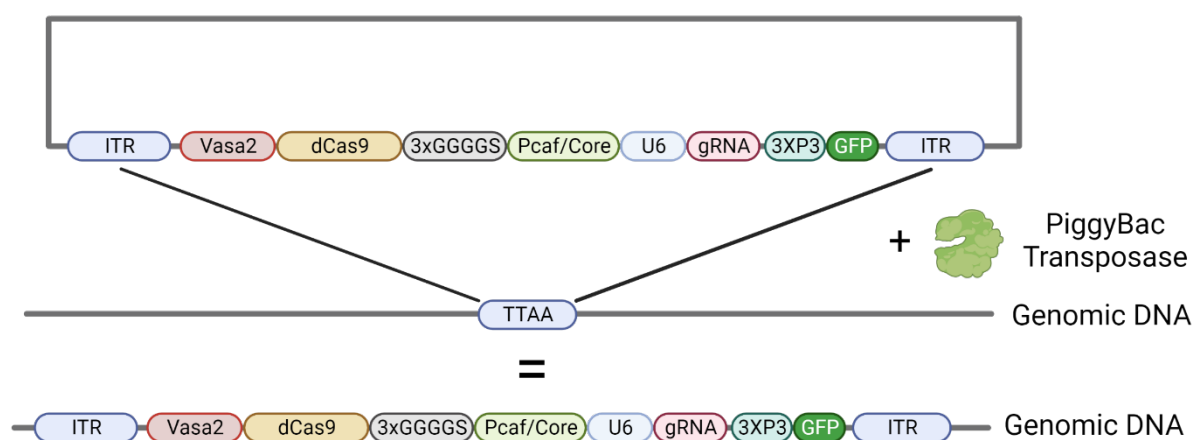


Figure 2-3 The PiggyBac integration system.

Two ITRs flanking the construct and the presence of a PiggyBac Transposase favour its integration in any TTAA sequence in the genome. Made with Biorender.com

2.3.2 Assessing the dCas9 binding stability during gametogenesis

The dCas9 is known to stably binds the DNA both *in vivo* and *in vitro* (Larson et al., 2013; Boyle et al., 2017; Xu and Qi, 2019). The bond with the DNA lasts for a very long time until cellular replication. Usually, the signal for DNA release in a nuclease-active Cas9 is represented by the cut, which doesn't happen in a nuclease-dead Cas9. Nevertheless, few studies tested dCas9 in a highly dynamic environment, such as meiosis (Kuhl et al., 2020).

During gametogenesis and throughout meiosis, chromatin is heavily reshaped and rearranged. Every time a germ cell replicates, the DNA-bonded proteins are stripped away to allow chromosome replication. This should also happen for dCas9, which otherwise will hamper the DNA polymerase activity. Immediately after DNA replication, the alpha helix is re-bound by histones forming nucleosomes. Other DNA-binding proteins condense and compact the structure to allow chromosome pairing, crossing over and segregation. In this turmoil of events, it's unclear how long a dCas9 would remain bonded to the DNA. In the Crisanti's lab, we have broadly used nuclease-active Cas9 during meiosis and tested its efficacy in producing double-strand breaks (DSBs) throughout gamete production (Galizi et al., 2016; Hammond et al., 2017; Kyrou et al., 2018).

Many papers have shown how the presence of nucleosomes and Cas9 simultaneously in the same section of the DNA are mutually exclusive due to the steric hindrance that each molecule exerts on the other one (Barkal et al., 2016; Isaac et al., 2016). Nevertheless, Cas9 has proved to be active in many situations, even when driven by late meiotic promoters, during meiotic stages typically characterised by chromosome condensation and increased nucleosome deployment on the DNA. For example, a Cas9 under the control of the late meiotic promoter *B2tub* was able to bind and cut the ribosomal repeats present on the X chromosome of *A. gambiae* during MSCI in a very similar way to the I-Ppo I nuclease (Galizi et al., 2016). This probably happens due to the dynamic nature of chromatin. Nucleosomes are not fixed elements of the DNA strand. Instead, they release and rebind the DNA every 25 milliseconds in a phenomenon known as "nucleosome breathing" (Isaac et al., 2016). Moreover, they are often shifted and repositioned to allow the epigenetic regulations of genes, providing opportunities for Cas9 binding. This should also be true for the dCas9. Nonetheless, it's easier to prove the activity of a Cas9 than a dCas9. This is because the Cas9 binds and cleaves the DNA introducing mutations and deletions or generating an altered phenotype, while dCas9 *in vivo* doesn't leave detectable "scars" on the alpha helix.

2.3.3 A Cas9/dCas9 competition experiment

To test the dCas9 binding activity, I designed an experiment to assess Cas9/dCas9 competition. As a dCas9 competitor, I used the gene drive line named QFS1, containing the Cas9 expression cassette targeting the *dsx* locus developed in our lab (Kyrou et al., 2018). The QFS1 line is the first gene drive line that effectively suppressed mosquito-caged populations. The Cas9 expression cassette is placed in one of the two *dsx* alleles with a gRNA that targets the opposite *dsx* locus. The line uses the homing gene drive system (HGD); hence, it exploits HDR to copy the transgene into the opposite chromosome (*Figure 2-4*). The Cas9 expressed by QFS1 is controlled by the regulatory region of the pre-meiotic gene “zero population growth” (*zpg*), which restricts its expression in the gonads during gametogenesis. The employment of the *zpg* promoter increased the efficacy of this HGD compared to the *vasa-2*, thanks to a more restricted expression in the germline and almost no deposition in the embryo (Andrew Hammond, Kyrou, et al., 2016).

Moreover, the transgene insertion in the ultraconserved *dsx* gene made it almost impossible for the line to develop resistant alleles (Kyrou et al., 2018). Even if the cut produced by the Cas9 would be repaired via NHEJ or MMEJ, any mutation of the nucleotidic sequence on the *dsx* allele targeted by the Cas9 will disable the gene producing an intersex phenotype, incapable of propagating the resistant allele in the population. QFS1 showed homing rates close to 99% both in males and females and fertility rates comparable to the wild type counterpart (Kyrou et al., 2018) (*Figure 2-4*).

In this experiment, the QFS1 gene drive line was crossed with the D0 transgenic line. The D0 line expresses a simple dCas9 with no gRNA. In individuals containing both transgenes, Cas9 and dCas9 bond to the *dsx* gRNA should compete for the same binding site. Stable binding of dCas9 ahead of Cas9 should protect the target site from the nuclease activity, reducing the gene drive's homing rate (*Figure 2-4*).

I used the *vasa-2* promoter to drive the dCas9 while the Cas9 was under the *zero population growth* (*zpg*) regulatory region. Both these promoters have a pre-meiotic activity with an overlapping expression pattern. *Vasa2* shows much higher levels of expression than the *zpg*, particularly during the first phases of gametogenesis (*Figure 2-2*). This condition should favour the dCas9 over the Cas9 binding (Taxiarchi et al., 2019; Terradas et al., 2022). On the other hand, dCas9 must remain on the double helix for a long time to protect the binding site, while Cas9 needs only a few seconds to bind and cut the DNA. Therefore, anytime the dCas9 releases the alpha helix, the binding site becomes available again for cleavage, eventually triggering the

homing of the Cas9 transgene. Once cut, the binding site could not be bound again due to the changes in its sequence produced by any of the repairing mechanisms employed by the cell (NHEJ, MMEJ or HDR).

After crossing the two transgenic strains, I selected transheterozygous males and females from the progeny. These individuals were crossed with their wild type counterparts, and their progeny was screened for the presence of the Red Fluorescent Protein (RFP) marker, indicating the presence of the QFS1 transgene (*Figure 2-4*).

When studying a gene drive line that exhibits a super-mendelian inheritance rate, is appropriate to consider the total inheritance rate as the sum of the Mendelian inheritance and the homing-related inheritance. In a gene drive line, 50% of the transgene counts are due to Mendelian inheritance, and only the quota exceeding this percentage could be considered homing inheritance. Therefore, I only contemplated the latter one to calculate the homing rate. To express it as a total percentage of the homing-related inheritance quota, I doubled the values exceeding the Mendelian inheritance rates. For example, a total inheritance rate of 75% can be considered as 50% of Mendelian inheritance and 25% of homing inheritance. Therefore, the calculated homing rate will be 50%.

For completion and to have a more straightforward comparison with the papers previously published describing HGD lines (Andrew Hammond, Kyrou, et al., 2016; Kyrou et al., 2018), I present in this work both the total transgenic inheritance and the corresponding homing rate. The inheritance values of the QFS1 transgene collected in the work by (Kyrou et al., 2018) showed that, in the absence of dCas9, the transgene inheritance corresponded to $96.6\% \pm 1.4$ with a homing rate of $93\% \pm 2.7$ for males and $98.5\% \pm 0.8$ with a homing rate of $96.7\% \pm 1$, for females respectively. With a single copy of the dCas9 transgene, these values dropped to a total inheritance of $86.6\% \pm 2.4$ with a homing rate of $73.2\% \pm 4.8$ for males and $87.6\% \pm 2.4$ with a homing rate of $75.3\% \pm 4.8$ for females (*Figure 2-5*).

I then tested the dCas9 inhibitory effect on QFS1 transgene homing by doubling the copies of the dCas9 transgene. An increase in the amount of dCas9 presumably would result in a further decrease in QFS1 inheritance.

Transheterozygous for QFS1 and dCas9 transgenes were crossed again to homozygous D0 individuals (dCas9 donor) (*Figure 2-4*). These crosses will not guarantee that all the progeny will be homozygous for the dCas9 allele, but they should considerably increase the fraction of the homozygous individuals. From this newly generated offspring, I screened the transheterozygous individuals that showed the greater GFP intensity (I used a 3xP3::GFP marker for the dCas9 transgene), assuming that a stronger expression of the GFP would

correspond to the presence of a double allele. Finally, I crossed those individuals to their wild type counterparts. The resulting progeny was screened for the presence of the red fluorescence marker as an indication of the QFS1 transgene inheritance. As expected, the inhibition of the homing rate was more robust, with a total inheritance rate of $64.6\% \pm 2.5$ and a homing rate of 29.2 ± 5 for the progeny generated from transgenic males and $72.2\% \pm 3.7$ and $44.5\% \pm 7.5$ for females (Figure 2-5). All the transgenic lines presented in this work were bred and tested in the same facility with the same conditions. Despite not being ideal to use previously collected data as a control for these experiments (the data from QFS1 crossed with wild type were from (Kyrrou et al., 2018)), the experimental conditions were identical, ensuring a good grade of robustness.

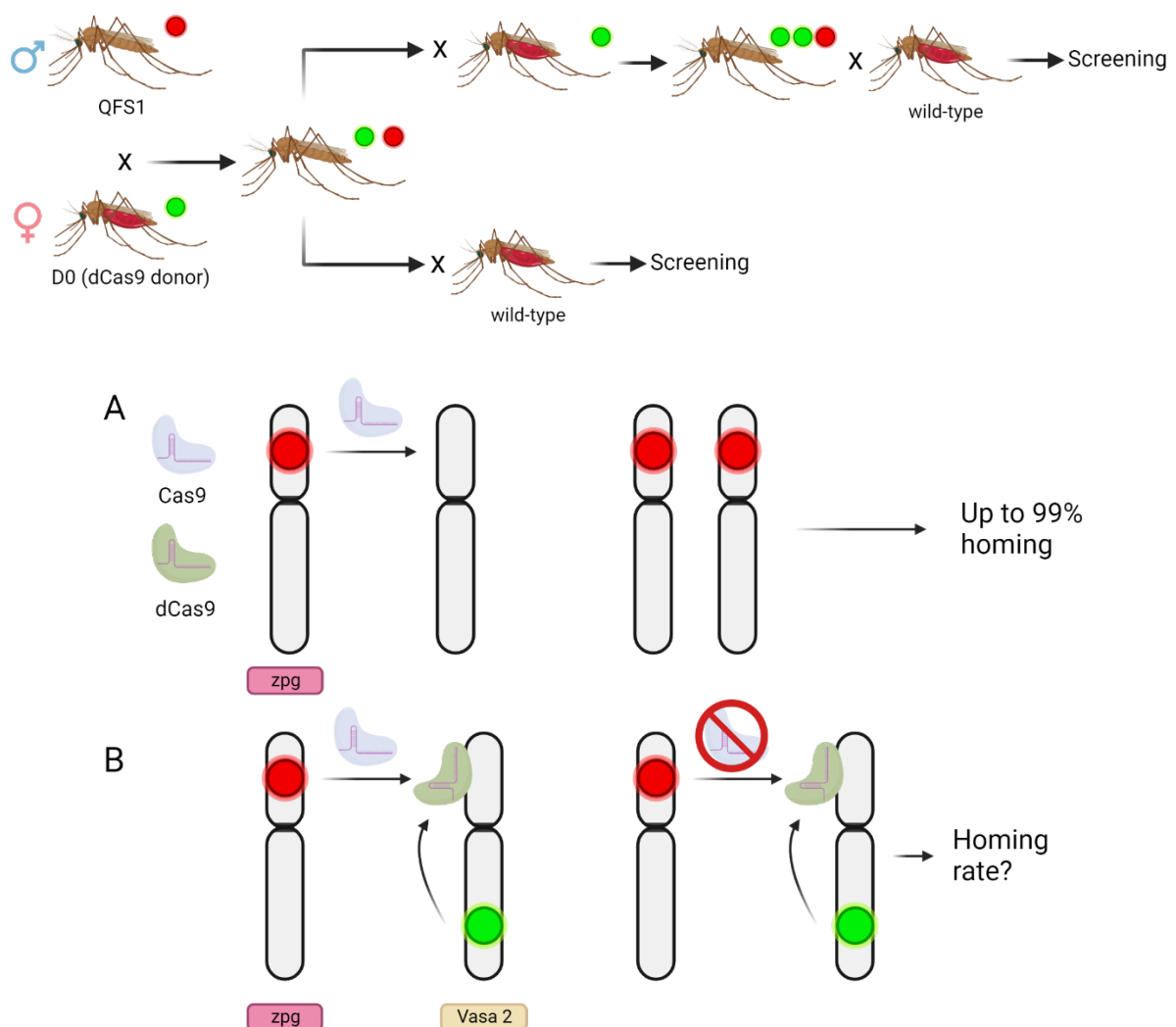


Figure 2-4 Crossed performed for the dCas9-Cas9 competition and experimental setup. The top of the figure shows how the crosses were performed. The red dot indicates the QFS1 transgene, while the green represents the dCas9 allele. Two dots indicate the presence of a double allele. The bottom part of the figure: A: how the QFS1 gene drive line works B: The Cas9/dCas9 competition experiment. The dCas9 expressed under the vasa-2 promoter should exploit the gRNA provided by the QFS1 transgene and bind the target site shielding it from the nuclease activity. Made with Biorender.com

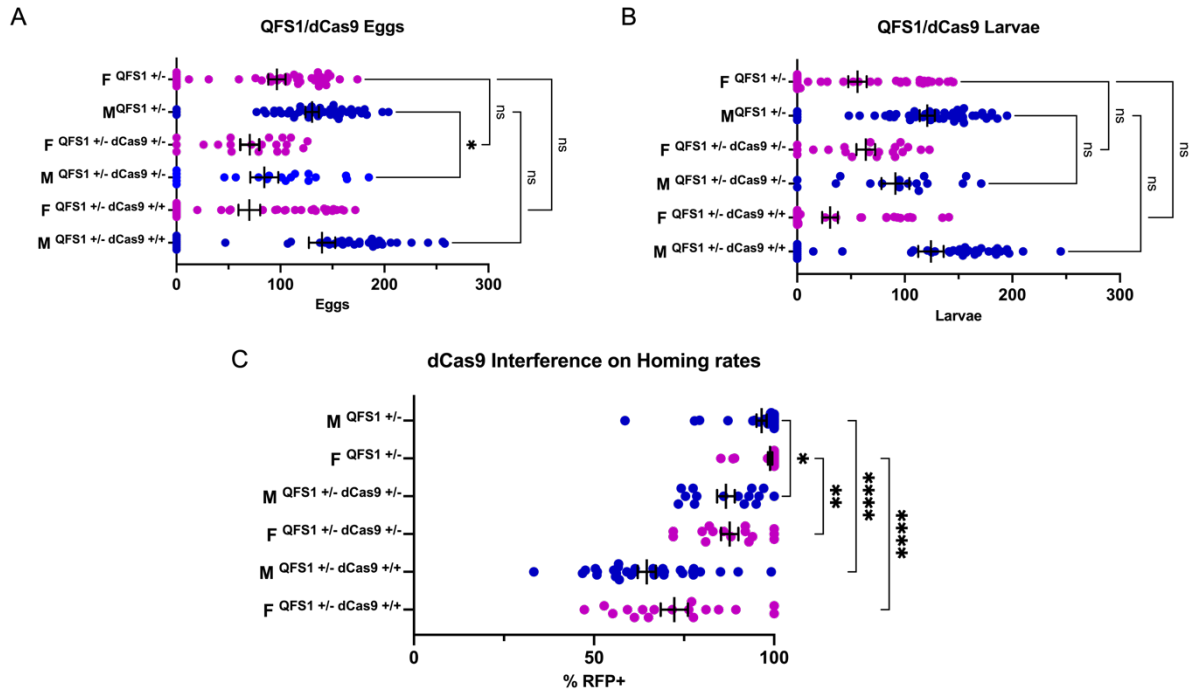


Figure 2-5. Data collected from the Cas9-dCas9 competition experiment.

A and B show the eggs and larvae output for each line assessed. No significant differences were detected in the number of eggs laid and the larvae produced. C shows the homing inheritance without the dCas9 interference and in the presence of one or two alleles expressing the dCas9. The data regarding egg deposition, larvae production and homing rate of the QFS1 line are the ones published by (Kyrou et al., 2018). Statistical test employed: Kruskal-Wallis, P values: * = ~0.015, ** = ~0.0026, **** < 0.0001

To better understand the dynamics of the Cas9/dCas9 competition, I thought it would be interesting to assess their dynamics when driven by the same regulatory region. Therefore, I run another test with a different gene drive line that targets the essential female fertility gene AGAP007280 and expresses the Cas9 under the *vasa-2* promoter (Andrew Hammond, Galizi, et al., 2016). This line works in the same way described for the QFS1 line; the only differences are the gene disrupted by the transgene insertion and the promoter employed to drive the Cas9. The experiment was conducted only with males because females with one allele disrupted from the gene drive transgene were almost entirely infertile. As expected, the co-expression of the two proteins under the same regulatory region led to lower efficiency in inhibiting the homing process, with the control line that exhibited a total inheritance of $99.4\% \pm 0.3$ and a homing rate of $98.9\% \pm 0.7$ while these values in the presence of a single dCas9 allele were $96\% \pm 1.6$ for the total inheritance and $92\% \pm 3.2$ for the homing rate (Figure 2-6). The statistical analysis didn't consider the differences between the homing rates significant (P values > 0.99). Nevertheless, in this case, it was expected to see little difference in the homing rate. In the previous experiment, under the control of two regulatory regions with a significant difference in expression strength (the *vasa-2* expresses almost four times more transcript than the *zpg* in the GSC, according to Taxiarchi et al.), the competition between the two proteins gave a

homing rate reduction of the 25% at best. When under the expression of the same promoter, the fast kinetics of the nuclease-active Cas9 seems to grant an overwhelming advantage over the deactivated counterpart. Despite this, the previous experiment proved that competitive dynamics exist between the two Cas9 proteins, even if they lie below the statistical significance threshold, like in this case.

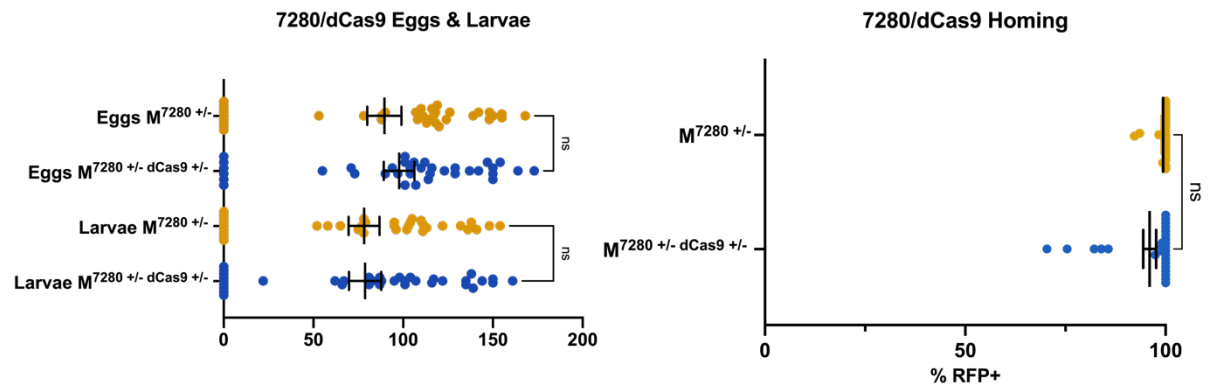


Figure 2-6 The Cas9-dCas9 competition experiment in the 7280 gene drive line. The left graph shows the eggs and larvae output for the gene drive line with and without the presence of the dCas9 allele. On the right are the differences in the homing rate induced by the dCas9 interference. The statistical analysis was performed using the Kruskal-Wallis test. Ns = non-significant, P values > 0.99

2.3.4 Testing dCas9 misexpression and deposition in *A. gambiae*

To corroborate previous results and further investigate dCas9 binding efficacy, I tested whether the dCas9 could affect the *dsx* expression in a way that resembles CRISPR-interference (CRISPRi) (Larson et al., 2013). Previous experiments have shown that the *vasa-2* regulatory region employed in this study, in combination with an active Cas9 targeting genes important for female fertility, could induce female sterility probably due to a combination of leaky expression (i.e. expression of the protein also in somatic cells) and embryonic deposition of the protein or its mRNA (A Hammond et al., 2016; Kyrou et al., 2018). The idea behind this experiment is that if the dCas9 is also misexpressed in other tissues and/or deposited in the embryo, it might be able to interfere with the expression of the *dsx* gene and potentially induce some kind of mosaic phenotype that can lead to reduced fertility or sterility (*Figure 2-7*). In females, the *dsx* gene is expressed mainly during embryonic development and larval and pupal stages, while in males, the gene is expressed primarily during the pupal stage. The main tissues in which is expressed in females are the ovaries, oviducts, and accessory glands, as well as in the brain and fat body. In contrast, in males, *dsx* expression is primarily detected in the testes, accessory glands, and brain (Dean A Baker, Tony Nolan, Bettina Fischer, Alex Pinder, 2005; Marinotti et al., 2013; Taxiarchi et al., 2019).

In our lab, we developed a transgenic line with a gRNA cassette expression that targets the *dsx* locus placed in one of the *dsx* alleles. The gRNA spacer is the same one that was employed to develop the first fully working gene drive in *Anopheles* (Kyrou et al., 2018). I crossed this line with the dCas9 donor line (D0), which only carried the dCas9 transgene under the control of the *vasa-2* regulatory region. From the progeny of this cross, I screened for transheterozygous, thanks to the presence of two different fluorescent markers for the two transgenes. Males and females transheterozygous were then crossed with their wild type complement, and females from both crosses were placed in single cups for deposition. Then, eggs and larvae were counted for each single deposition to calculate the hatching rate (*Figure 2-7*).

In transgenic females, the expression of dCas9 in the germline might impair the expression of the female *dsx* isoform in the ovaries, hampering overall fertility. Additionally, the leaky expression reported for the *vasa-2* promoter (A Hammond et al., 2016) might also affect other fertility-related tissues like oviducts and accessory glands. Similarly, if dCas9 is deposited in the embryo, it might interfere with the expression of the *dsx* gene during embryonic development, further reducing the hatching rate.

We observed an overall reduction in egg output in the transheterozygous females (even if it was not statistically significant), with approximately only half of them laying eggs (21 out of 40) and, in six cases laying less than 15 eggs in total. If we consider only the females that did lay eggs, the average hatching rate was $40\% \pm 8.4$, which was significantly lower than the hatching rate of $84.7\% \pm 4.5$ (P value < 0.0001) in females carrying only the *dsx* gRNA cassette and $88.1\% \pm 3.3$ in the wild type control group.

When wild type females were mated with transheterozygous males, we also observed a reduction in egg output, even if to a lower extent. The females who laid eggs were 27 out of 40, and none of them laid less than 100 eggs. Considering only the females that laid eggs, the average hatching rate was $70.3\% \pm 5.7$, compared to $84.8\% \pm 4.6$ when mated with males carrying only the gRNA cassette (P value = 0.0063).

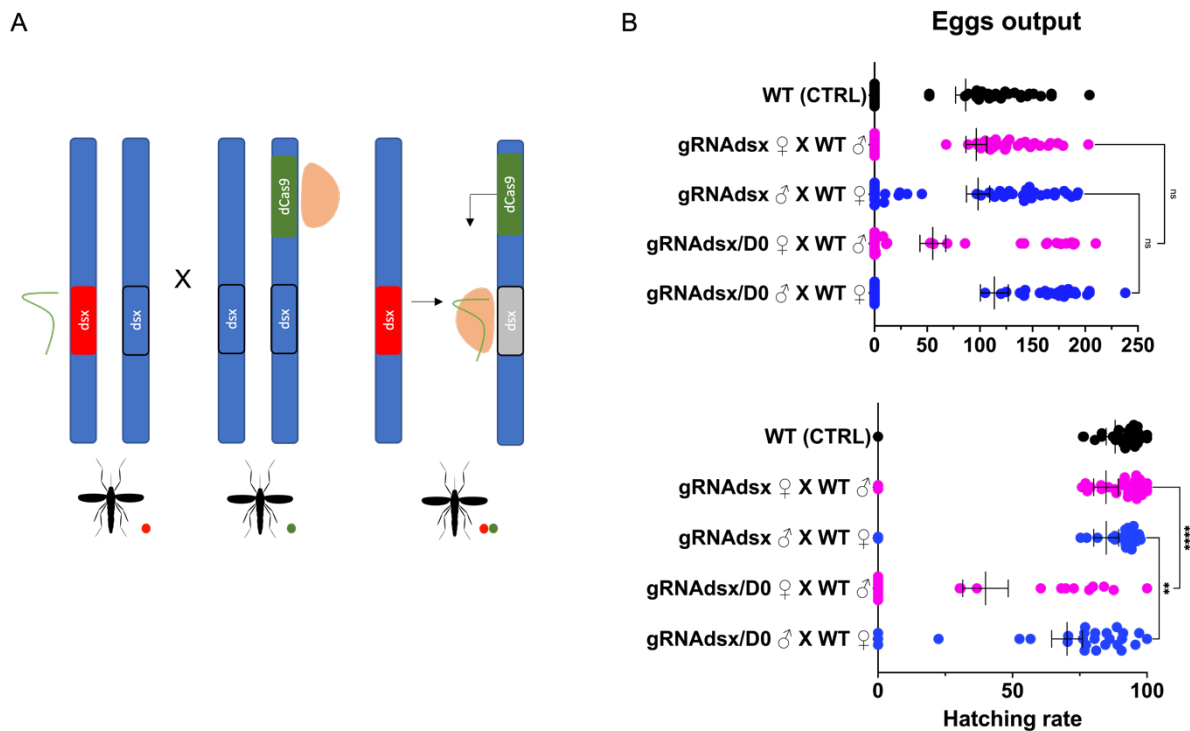


Figure 2-7 Testing the dCas9 misexpression and deposition.

A) Crosses performed between transgenic mosquitoes to produce transheterozygotes carrying a *dsx*-targeting gRNA in the *dsx* locus and a dCas9 (randomly inserted in the genome). B) A significant reduction in the hatching rate was detected mainly in crosses involving transheterozygote females. The statistical analysis was performed using the Kruskal-Wallis test. Ns P Value > 0.05 , ** P value = ~ 0.0063 , **** P Value = < 0.0001

2.3.5 The engineering and testing of the dCas9-mCherry lines

Once I assessed the binding of the dCas9, I proceeded with testing the protein as a potential chromatin barrier. It was previously shown that dCas9 alone could create a nucleosome-depleted region that spans 100-150 bp from the binding site (Barkal et al., 2016). The first lines generated expressed a dCas9-F2A-mCherry fusion and a gRNAs that targeted three selected regions at -511 bp, -110 bp and -62 bp on the Y-linked *B2tub-I* PpoI expression cassette, respectively (lines D1511, D1110 and D162 respectively). I first verified the presence of the dCas9 in the testis by detecting the mCherry fluorescence. The fluorescence was extremely faint, probably due to the mCherry position after the self-cleavable F2A epitope along the transcription unit. It is known that transcripts downstream of the F2A are less translated (Liu et al., 2017). Strangely, I also detected the presence of GFP that was not supposed to be expressed in the testes, given that the only GFP marker in the transgene was regulated by the 3xP3 (*pax*) promoter specific to nerve cells. The GFP cassette was placed before the *vasa-2* regulatory region on the antisense strand (Figure 2-8). If we exclude autofluorescence of the testes (not detected on the wild type control testes), a possible explanation is that the *vasa-2* promoter could drive its expression bi-directionally even on genes placed on the opposite strand. The presence of the dCas9 and the mCherry transcripts were also confirmed via RT-PCR.

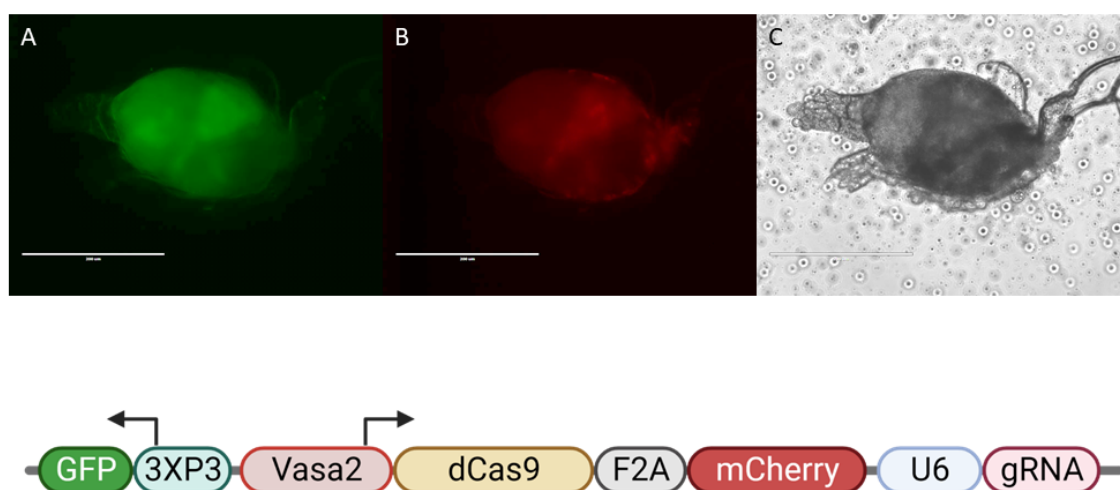


Figure 2-8 *Vasa2* promoter drives the expression of a GFP on the antisense strand (Scale bar = 200 μ m).

A: The unexpected fluorescence detected on the testes is probably due to the *vasa-2*-driven activation of the 3xP3-GFP cassette. B: mCherry fluorescence C: the testis in normal light. Below is the construct that gave the miss-expression of the GFP.

To activate the transcription of the silent Y-linked I-PpoI endonuclease, I crossed the line containing it (Y-I-PpoI line) to each of the dCas9-bearing lines. I was trying to answer whether the simple binding of the dCas9 on the meiotic *B2tub* promoter would be enough to create a nucleosome-depleted region and allow transcription (Barkal et al., 2016). In this case, I should have found a male bias progeny due to the I-PpoI targeting the X chromosome and eliminating X-bearing gametes. I used the different fluorescent markers on each transgene to select transheterozygous individuals. Females from both lines expressing the dCas9-F2A-mCherry were crossed with males of the Y-I-PpoI line. Males transheterozygous were selected and crossed with wild type females. Six hundred larvae were randomly selected from each progeny and screened for the presence of the RFP marker on the Y chromosome to assess the sex ratio. No significant sex distortion was detected in the progenies of these crosses despite the three different gRNA employed (*Figure 2-9*).

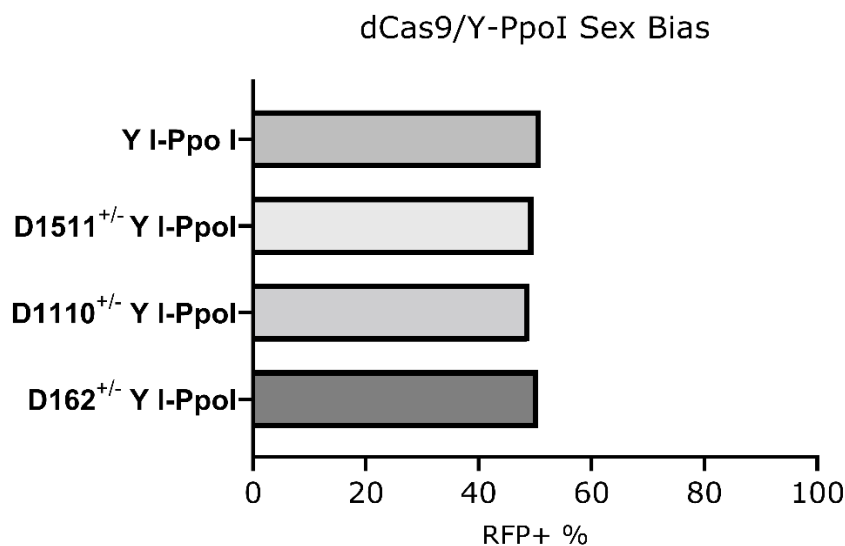


Figure 2-9 The dCas9 alone failed to produce sex bias. For this experiment, 600 larvae were randomly selected from each cross and screened for the presence of the Y-linked RFP+ protein. No sex distortion was detected.

2.4 Strategy 3 - Engineering of the gypsy retrotransposable element as a Y-chromatin insulator

To block the spreading of the epigenetic silencing along the Y chromosome during MSCI, I decided to surround the Y-linked I-PpoI nuclease with the chromatin insulator *gypsy*. Similarly to the autosomal sex distorter line described in (Galizi et al., 2014), I engineered a transgene bearing the I-PpoI nuclease driven by the late male germline-specific *B2tub* promoter and flanked it with two *gypsy* repeats (*Figure 1-10*). The *gypsy* elements were placed in opposite orientations (5'→3' and 3'→5') upstream and downstream of the transgene (*Figure 1-10*). This seems to be the most effective configuration to promote chromatin looping and effective insulations (Chen and Corces, 2001; West, Gaszner and Felsenfeld, 2002; Byrd and Corces, 2003). The construct also contained 14 UAS repeats (Upstream Activation Sequence) placed before the *B2tub* regulatory region as a docking site for GAL4 activation.

Moreover, the 14 UAS repeats can also be used as target sites for a dCas9 since each contains a PAM sequence. Fourteen dCas9 could potentially bind at once with only one gRNA. Finally, I added an attB site to the construct for site-specific integration. To generate the lines bearing the Y-linked I-PpoI nuclease surrounded by the *gypsy* elements, I co-injected the plasmid with a source of $\phi 31$ integrase driven by the *vasa-2* promoter. This promoted, *via* attB/P recombination, the specific insertion of the cassette on a Y-chromosome docking line already available in our lab (*Figure 2-10*) The line contains a Y-linked 3xP3-attP-RFP cassette, as described in (Bernardini et al., 2014). In this study, I directed the dCas9/gRNA complex to the sites within the *B2tub* promoter to activate transcription. Hence the UAS repeats were not utilised. However, the value of this transgenic line remains, as it could be employed for future tests on a broader range of activator molecules, including the Gal4.

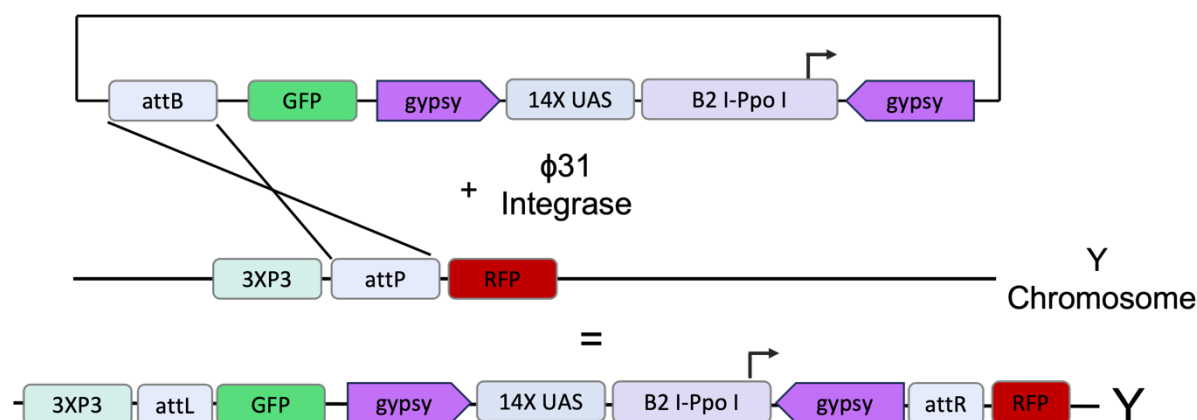


Figure 2-10 Site-specific integration on the *A. gambiae* Y chromosome

2.5 The dCas9-Pcaf fusion failed to activate the transgene on the Y chromosome.

Given the first negative results in rescuing the activity of the Y-linked *B2tub*-I-PpoI expression cassette employing the dCas9 alone, I decided to focus only on the generation of the dCas9-Pcaf lines and to use gRNA targeting binding sites on the *B2tub* regulatory region closer to the TSS. I also re-engineered the constructs moving the GFP cassette at the end of the dCas9 expression cassette to avoid the trans-activation driven from the *vasa-2* promoter (*Figure 2-1*). To test the ability of the dCas9-Pcaf and the dCas9-Pcaf^{CORE} fusion proteins to activate the I-PpoI transcription, I engineered plasmids expressing gRNAs targeting regions in the *B2tub* promoter upstream of the *I-PpoI* TSS. As a result, four new transgenic lines were generated. Two lines express the dCas9-Pcaf fusion protein targeting the -32 bp and -68 bp binding sites (lines D232, D262), and the other two express the dCas9-Pcaf^{CORE} fusion targeting the same two loci (lines D232^{CORE} and D262^{CORE}). Transgenic females from these transgenic lines were crossed to Y-linked I-PpoI males.

Males transheterozygous were selected from the progeny. Fifty of them were placed in a cage and crossed with fifty wild type females (*Figure 2-11*). I also tested the I-PpoI transcription activation by doubling the dose of the dCas9-Pcaf fusion protein. Increasing the dCas9 concentration would probably increase its binding ratio to the target site.

For space availability and feasibility reasons, I had to choose only two lines to be tested this way. I decided on the two lines carrying the gRNA targeting the binding site at -62 bp because it looked close enough to the TSS without being within the 53 bp of the *B2tub* core region (Michiels et al., 1989). Therefore, fifty transheterozygous males carrying the transgenes from the lines D262 and D262^{CORE} were placed in separate cages and crossed once more with females having the same transgenes to increase the fraction of homozygosity for their respective alleles (*Figure 2-11*). New transheterozygous males were screened from the progeny, and only those with the higher GFP intensity were selected, assuming that a higher intensity of the GFP corresponded to the presence of a double transgenic allele. The selected transheterozygous males from all the crosses were placed in cages and crossed with wild type females (*Figure 2-11*). I waited five days to allow mating, then blood-fed the cages. This time I opted for a single deposition assay to collect data on the lines' overall fertility. Forty females from each line were placed in single cups for single-cup deposition three days after the blood meal. Eggs and larvae were counted from each cup to assess the egg production and hatching rate. No significant differences were detected in eggs productions and hatching rate when

comparing the Pcaf and *gypsy*-I-PpoI transgenic lines with the Y-I PpoI control. Then the larvae were screened for the presence of the Y-linked RFP to determine the overall sex distortion. No significant sex distortion was detected in any of the progenies (Figure 2-11).

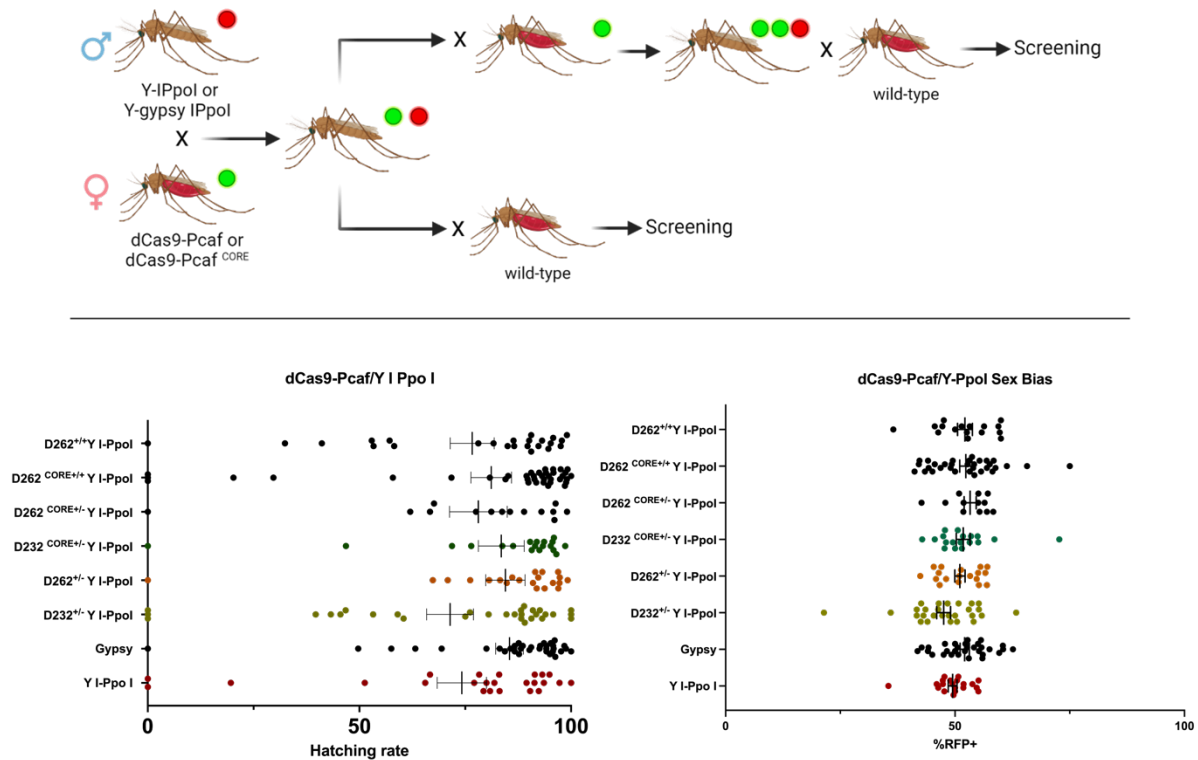


Figure 2-11 Crosses and results of the Pcaf delivery on the Y-linked *B2tub*-I PpoI transgene. The upper half of the picture is a schematic representation of the crosses performed. The red dots represent the Y-linked I-PpoI transgene, while the green dots represent the dCas9-Pcaf. Two green dots represent mosquitoes homozygous for the dCas9-Pcaf transgene. The bottom half of the picture shows the data points of the hatching rate and sex distortion detected. No significant difference was detected among the data sets (*P* Values > 0.05, tested via Kruskal-Wallis test). Made with Biorender.com.

To understand if the absence of male bias was correlated with the inability of the *B2tub* to activate transcription of the I-PpoI, I analysed the expression levels of the I-PpoI via qPCR. For this analysis, I also included crosses between the dCas9-Pcaf lines and the *gypsy*-I-PpoI line to see if the two combined strategies might have a synergistic outcome. Given the high number of crosses and lines to be maintained, I focused mainly on testing lines containing what I think are the most balanced gRNAs in terms of distance from the TSS (the -62 and -110) and the HAT that should have the highest acetyltransferase activity, the Pcaf^{CORE}. Therefore I tested the lines D262^{CORE}, D2110^{CORE} crossed with the Y-linked I-PpoI line with and without *gypsy* insulators. In addition, I tested one line only expressing the dCas9-Pcaf fusion (the

D262), and also, in this case, I crossed it with the Y-linked I-PpoI line with and without *gypsy* insulators.

I designed four pairs of primers to amplify different fragments of the I-PpoI transcript and tested them on the autosomal sex distorter line (*B2tub*-I-PpoI on the 3R chromosome) (Galizi et al., 2014). All amplifications resulted in fragments of the right size, so I chose the pair that showed a stronger signal. The primers pair targets a region that spans 86 bp, from codon 53 to 81 of the I-PpoI open reading frame (*Figure 2-12*).

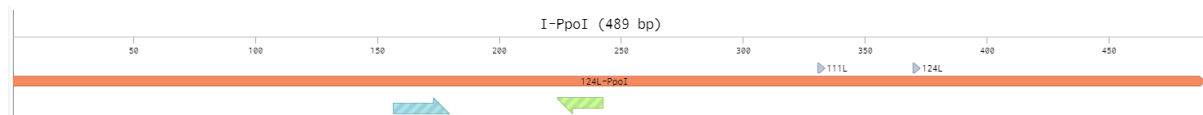


Figure 2-12 Primers positions for I-PpoI qRT-PCR

The mRNA levels from abdomens (containing the testes) and carcasses (rest of the body) were compared (only the value for abdomens are shown in *Figure 2-13*).

The qRT-PCR was performed using the RPL19 as a housekeeping gene for normalisation. Despite the absence of sex distortion, the I-PpoI transcript was detected in several lines (*Figure 2-13*). Also, the Y-linked I-PpoI line, always considered completely silenced, showed a very faint but detectable expression of I-PpoI in the abdomen. I decided to use this value to compare the I-Ppo I expression in the other lines.

Compared to the Y-linked I-PpoI line, the PMB1 strain (the only line with strong sex bias) expresses the transcript 153 folds more. Here, the I-PpoI transcript was also detectable in the carcasses. This might imply some leakage of the *B2tub* promoter activity in tissues other than the testes. However, all the other lines tested showed a level of expression in the carcasses below the reference value, which was no different from the wild type control.

The expression in the abdomens of the other lines showed mostly values in line or below the reference value. The only lines showing increased I-PpoI transcription were two lines containing the *gypsy* insulators. In the Y-linked *gypsy*-I-PpoI line, the transcript levels were almost 3.3 folds higher. Also, the line bearing the Y-linked *gypsy*-I PpoI and the dCas9-Pcaf^{CORE}, targeting the binding site -62 bp on the *B2tub* promoter, showed an increase in I-PpoI transcripts levels of about 3.2 folds, while the same line carrying the gRNA targeting the binding site -110 bp showed an expression of 0.61 folds the reference value (*Figure 2-13*).

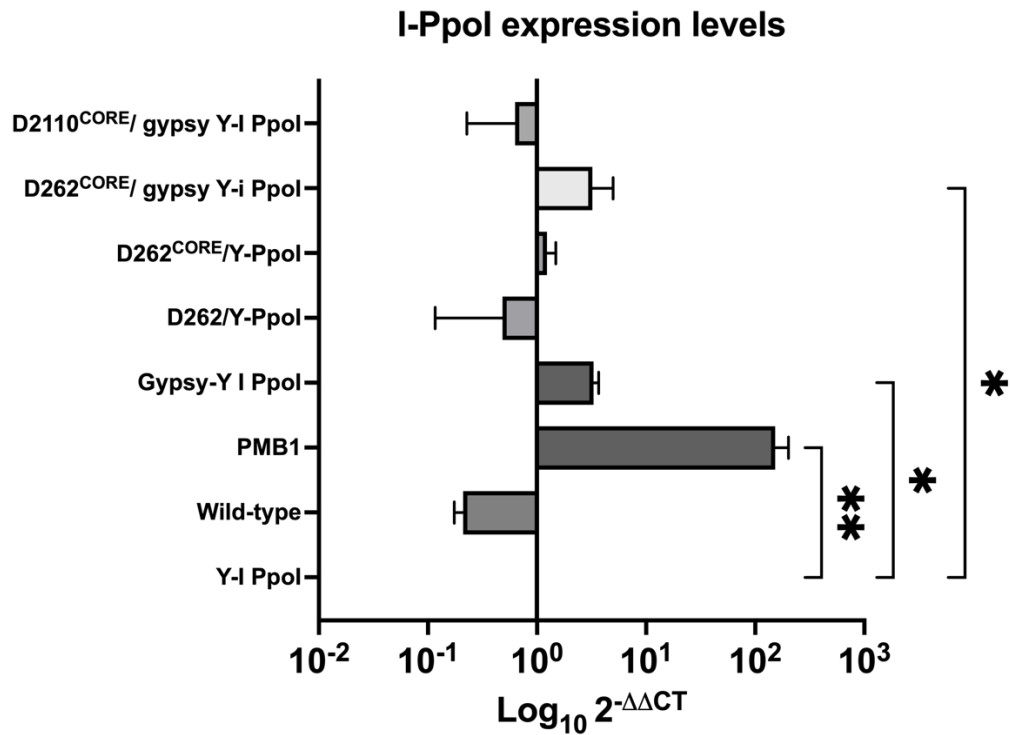


Figure 2-13 qRT-PCR expression values for the I-PpoI.

All the measurements were performed on male abdomens containing testes. The values in this graph refer to I-PpoI expression levels in the Y-linked I-Ppo I line. Two of three lines containing the Y-linked I-PpoI transgene flanked by gypsy insulators showed increased transcription levels of approximately 3.2 folds. The autosomal sex distorter line PMB1 showed an increased transcription level of 153 folds. The lines containing the Y-linked I-PpoI without the gypsy insulators showed comparable transcription levels to the reference value even in the presence of the dCas9-Pcaf/ dCas9Pcaf^{CORE} transgenes. The wild type showed an expression level of 0.2 folds that can be considered background noise. All the experiments were performed in triplicates. The statistical significance was calculated using a Kruskal-Wallis test. P values: * = ~0.02, ** = 0.0019

2.6 Characterisation of the Y-linked I-PpoI cleavage activity in the gypsy line

I-PpoI transcripts were detected in the abdomens of the Y-gypsy-I-PpoI transgenic lines indicating a possible activity of the nuclease in the testis. However, the overall amount of transcripts produced was extremely low, especially compared to the levels produced in the autosomal sex-distorter line. I decided to investigate if the transcript level in the *gypsy* line was enough to induce cleavages on the X chromosome rDNA repeats.

Daughters of Y-gypsy-I-PpoI males are essentially wild type since the transgene is placed on the father's Y chromosome that they do not inherit. Nonetheless, the low transcription level of the I-PpoI in the parental testes could still induce some damage to the paternally inherited X-chromosomes. After a few generations of inbreeding, these mutations might accumulate and be quantifiable. Therefore, I inbred the Y-gypsy-I-PpoI line for three generations. After extracting the genomic DNA from ten single females, I amplified a 453 bp fragment of the 28S ribosomal subunit repeats (there are ~300 repeats on the X chromosome) containing the nuclease target site. If the region was cleaved by the I-Ppo I, mutations at the cut site would be detected by digesting the DNA fragment with the same nuclease as reported by Galizi et al. when they analysed the autosomal sex distorter line (Galizi et al., 2014). Incomplete digestion of the fragment might indicate possible mutations in the I-PpoI binding site proximal to the cleavage site. The I-PpoI cleavage activity was compared to that of a single wild type female as a negative control. It is important to recall that this is a multi-copy gene and that there are hundreds of repeats on a single X-chromosome; if the expression levels of the I-Ppo I are low, its effect on the multiple repeats could get diluted.

No differences were detected between the samples compared to the wild type control (*Figure 2-14*). I then decided to analyse these samples via amplicon sequencing to rule out the possibility that some mutations might not be detectable on an agarose gel. The sequencing showed the presence of several deletions of variable size around the I-Ppo I cut site. As expected, the percentage of the mutated reads was extremely low compared to the total reads, which explains why they could not be detected on the agarose gel (*Figure 2-14*). Nevertheless, all the reads with a deletion or an insertion (indel) mapped around the target site in all ten *gypsy* females analysed. In contrast, indels were completely absent in the wild type control. I compared these data with the amplicon sequencing data previously collected from the few females generated by the autosomal sex-distorter line PMB1. The data referred to individuals that inherited the transgene and those that didn't. The size and distribution of the deletions

around the I-PpoI target site matched those observed in daughters of *gypsy*-I-PpoI fathers. The striking difference between the two lines was the number of reads containing deletions. The multicopy DNA of the 28S ribosomal RNA-coding sequence targeted by the I-PpoI is an essential gene necessary for ribosome assembly, protein synthesis and cell survival. Hence, females surviving the numerous cuts inflicted by the autosomal I-PpoI still have enough undamaged copies of the rDNA to fulfil protein synthesis. In these females, > 99.5% of the reads were wild type. Only 0.42% of the reads in females that didn't inherit the PMB1 transgene and 0.24% in those that inherited it contained deletions. In comparison, reads from G3 females of the Y-linked *gypsy*-I-PpoI line had deletion values ranging from 0.0018% to 0.0068%. In the wild type sample, no deletions were detected around the target site (*Figure 2-14*). Despite the small numerical difference, the indel values from the *gypsy* line significantly differed from the wild type when tested statistically with the non-parametric Wilcoxon test (P value = 0.0039). It is important to notice that the values of the wild type sample were all “zeroes” and that with a bigger sample, it could be possible to detect variations that do not appear here. Moreover, the survived females from the PMB1 line were not included in this analysis as they represent a minor and biased set (the one with the lowest amount of cuts on the X chromosome) of the potential sample that could not be analysed due to the lethality imposed by the “X-shredder” transgene.

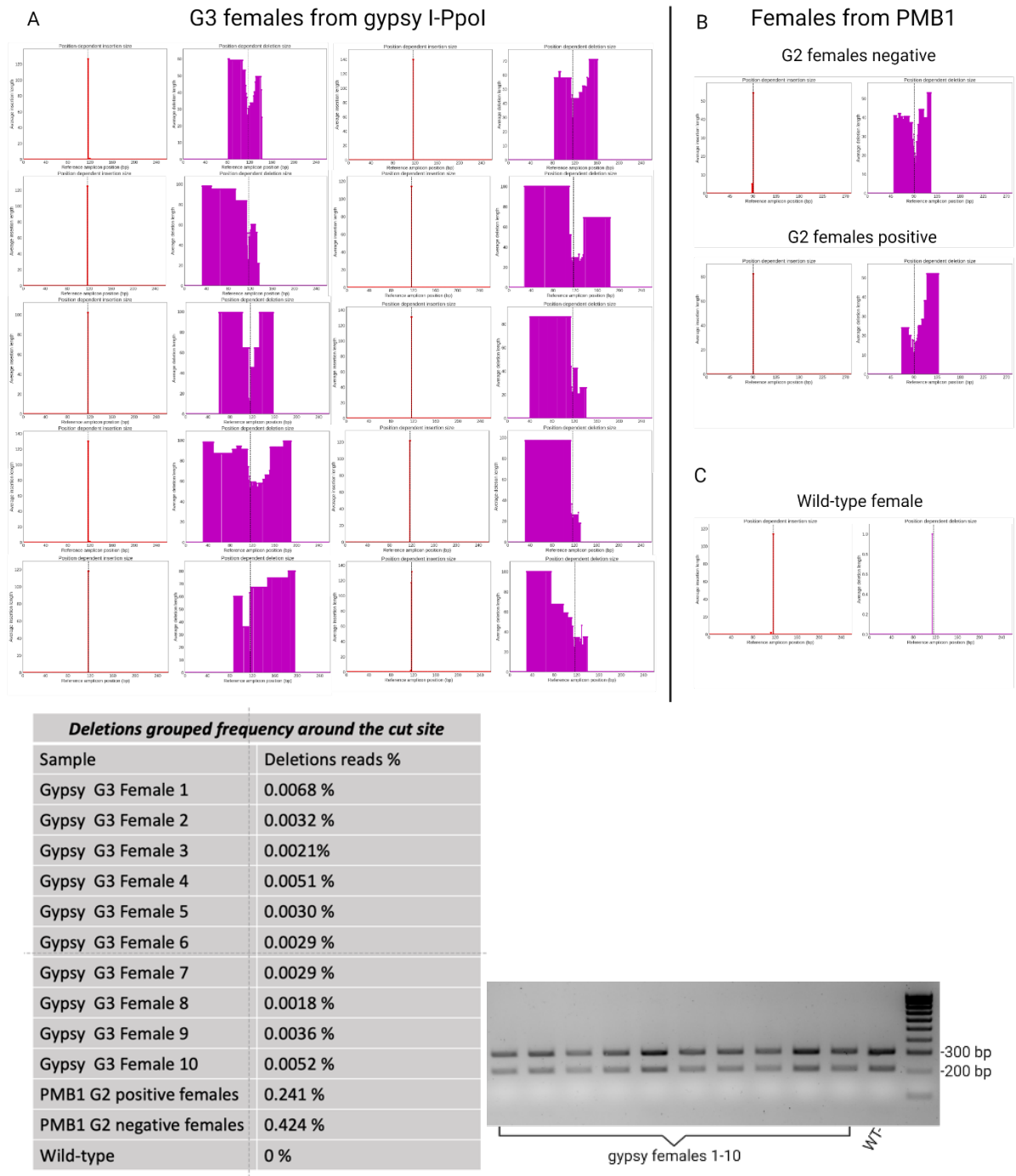


Figure 2-14 Amplicon sequencing analysis of the 28S ribosomal RNA-coding sequence in single daughters of Y-linked gypsy-I-PpoI fathers.

A) deletions distribution around the I-Ppo I cut site. The graphs on the left with the red bars show the insertion detected. Those on the right with the purple bars show the deletions. The height of the bars represents the size of the insertions/deletions. All the females from the gypsy line had a similar profile to those from PMB1 (B). No deletions were detected in the wild type (C). Tab on the right: deletions grouped frequency compared to the total of the reads. The picture on the bottom shows the digestion of the amplicons of the 28S ribosomal subunit. No undigested band was detected on the gel.

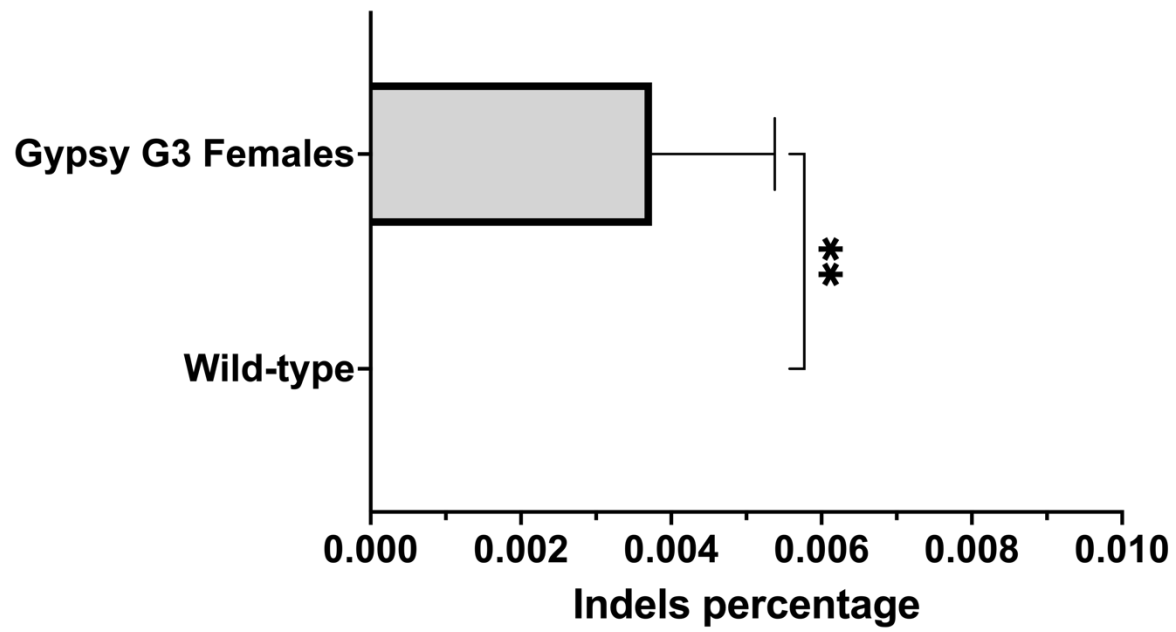


Figure 2-15 Significance level of the mutations detected from amplicon sequencing. The graph shows the significance of the indels detected in females from the Y-gypsy-Ppo I line compared to wild type females, for which the indel value is zero. P value = 0.0039 (Calculated using the Wilcoxon test).

3 Discussions and future perspectives

3.1 The dCas9/Cas9 competition experiment provides insights on dCas9 kinetics during gametogenesis.

The first strategy employed to rescue the Y-linked I-PpoI activity relied on the local delivery of a histone acetyltransferase fused with Cas9 protein devoid of nucleolytic activity (dead Cas9 protein; dCas9). Its targeted delivery aims to modify the epigenetic marks surrounding the transgene to activate its expression. To prove that the dCas9 was functional, I tested it against Cas9. The intent was to interfere with the homing rate of the Cas9-dependent gene drive (QFS1 line) to provide evidence that the dead nuclease could compete for either the binding site or the gRNA source. My findings suggest that the dCas9 may be an efficacious tool to control the homing of the gene drive QFS1 line. To drive the expression of the interfering dCas9, I used the *vasa-2* promoter.

The *vasa-2* transcripts are localised at the pole cells of the *A. gambiae* embryo (Figure 2-2) (Papathanos et al., 2009). Early in the zygote, the *vasa-2* promoter becomes transcriptionally active in the Germline Stem Cells (GSC) (a differentiation product of the pole cells) and will stay effective until the adult stage and for its entire duration.

Each cell, after an asymmetric division, gives rise to a new stem cell and a gonoblast (GB) in males or a cytoblast (CB) in females (Kahney, Snedeker and Chen, 2019). In *Anopheles*, the *vasa-2* transcript is found in the apical section of testes and ovaries, corresponding to the niche that contains the GSC (Figure 2-2) (Kahney, Snedeker and Chen, 2019). Here, it is expressed at its maximum levels, but moving down the gonad, the transcript concentration progressively fades (Figure 2-2). The *zpg* promoter has a similar expression pattern to *vasa-2* but is transcribed throughout gametogenesis at significantly lower levels (Figure 2-2) (Papathanos et al., 2009; Terradas et al., 2022).

If this is the case, it is reasonable to assume that, in transheterozygous mosquitoes bearing a *zpg*-driven Cas9 and a *vasa-2*-driven dCas9, the dCas9 pre-emptively occupies the vast majority. The prolonged binding of the dCas9 on the cleavage site should protect it from the Cas9 nuclease activity, and consequently, no “homing” should be recorded. In this study, this

was clearly not the case. As previously noted, dCas9 persists on the DNA strand until cellular replication starts (Boyle et al., 2017). During the four rounds of mitotic divisions of GB and CB, any protein bound to the DNA is challenged and removed from it by the DNA replication machinery apparatus or other mechanisms related to cell division. This is probably what happened in the dCas9/Cas9 competition experiment. A hypothesis to explain the relatively high homing rate could be the removal of dCas9 from the DNA at every round of replication, hence exposing the binding site to the activity of the Cas9 (*Figure 3-1*).

Moreover, Cas9 needs just a few seconds to bind the helix and perform the cut that will trigger HDR. This gives it a substantial advantage over its deactivated counterpart. The data presented here seems to point to this scenario. When I tested dCas9 and Cas9 under the control of the same regulatory region (*vasa-2*), I extrapolated a prediction of homing inhibition. Let's assume that after each round of replication, the two proteins have the same probability (1/2) of binding the DNA strand. The reduction in homing should equal the probability of the dCas9 to re-bind the DNA after each replication cycle of the four mitotic divisions. The likelihood of these events is:

$$P_i = 1/2 * 1/2 * 1/2 * 1/2 = 1/16 = 6.25\%$$

The homing rates detected during this experiment were $98.9\% \pm 0.7$ for the control (no dCas9 present) and $92\% \pm 3.2$ for the progeny generated by males carrying both the transgenes, a reduction of $6.9\% \pm 3.9$, in range with what was calculated in the model. This is obviously a very simplistic model and, particularly in this scenario, takes into account data that were below the threshold of statistical significance. Nevertheless, the following competition experiments with different promoters showed a more robust homing inhibition, validating the idea that the two proteins actually compete.

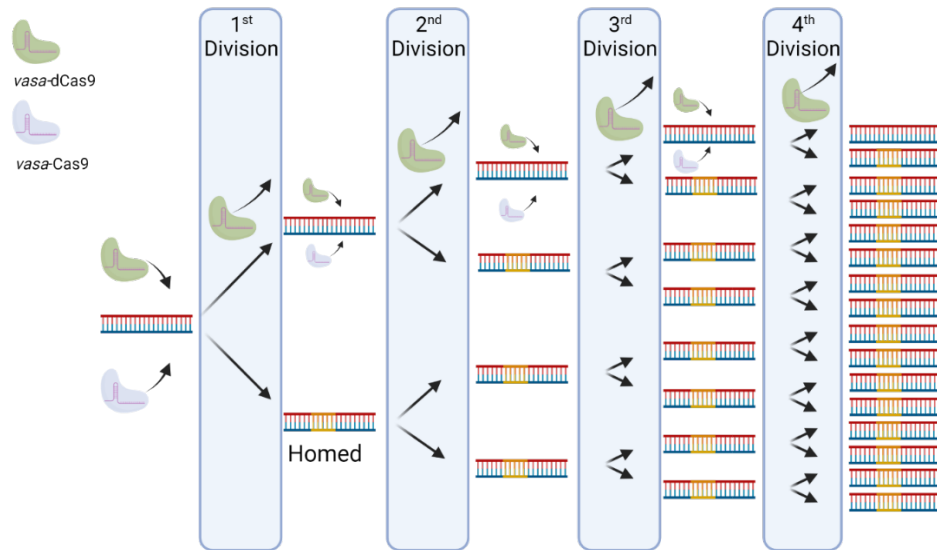


*Figure 3-1 Visualisation of the model hypothesised to explain the probability of dCas9 binding during spermatogenesis. Both dCas9 and Cas9 expressed from the GSC stage have the same possibility of binding the target site since they are driven by the same promoter (*vasa-2*). Every time Cas9 prevail, the target site is cut and becomes unavailable for further binding or cleavage. Therefore, during the four mitotic divisions, the dCas9 should re-bind the DNA after each replication cycle to protect it from cleavage. In this case, only one out of sixteen alleles will remain uncleaved at the end of the four cycles. Made with Biorender.com*

Assuming this model is correct, it should be possible to apply it with the appropriate approximations to explain the results of the competition experiment between the Cas9 and dCas9 proteins driven by two different promoters. The differences to consider in this case are the different expression strengths of the two promoters employed. From the data collected, Cas9 seems less competitive than dCas9 when under the control of the *zpg* promoter. This makes sense if we consider that the *zpg* promoter is relatively weaker than the *vasa* even if the *vasa* strength decreases over time (Figure 3-2) (Taxiarchi et al., 2019; Terradas et al., 2022). Starting from the model hypothesised for the previous experiment, it should be possible to try to infer at what replication cycle the two molecules start competing at the same rate. The average homing rate reduction detected in this experiment was around 25%. Hence, the two proteins should begin to compete actively after the second cycle of GB and CB replication (Figure 3-2). According to the model, the probability for the dCas9 to achieve such a level of homing inhibition would be:

$$P_1 = \frac{1}{2} * \frac{1}{2} = \frac{1}{4} = 25\%$$

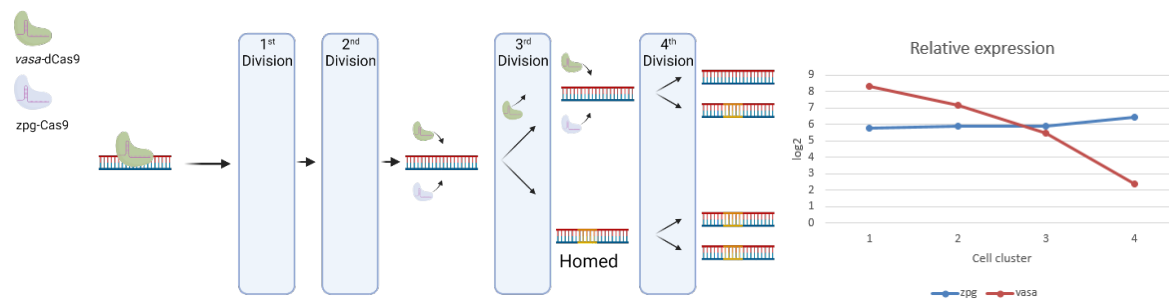


Figure 3-2: A model that tries to explain the dynamics of dCas9/Cas9 competition when expressed under different promoters. The homing rate inhibition in this experiment was 25%. This percentage might be explained if the two proteins start to compete at the same levels after the second cycle of replication. On the right is the relative expression strength of the *vasa* and *zpg* regulatory regions, as reported by Taxiarchi et al. Made with Biorender.com

However, the model was more difficult to apply in mosquitoes with two copies of dCas9 and one copy of the Cas9 transgenes and two different promoters driving the proteins. If the same promoter drove them, the ratio of dCas9/Cas9 would be 2:1, implying that dCas9 binds the target site twice as much as the Cas9, with a $2/3$ probability at any replication event. If that's the case and if we consider, as in the previous model, that the *zpg* promoter becomes competitive with the *vasa* after two cycles of replication, then the probability for the dCas9 to consecutively bind the target site will be:

$$P_i = 2/3 * 2/3 = 4/9 = 44.4\%.$$

In this case, the data collected from the experiment showed a more significant homing inhibition than what was calculated. The progeny from the males had a homing rate reduction of around 70%, and the females had approximately 55%. Multiple factors might concur to the discrepancy between what the model predicted and what was detected, starting from the difference in the strength of the two promoters employed, which might lead to a more significant early accumulation of dCas9 in the cell. A further possible explanation for this discrepancy would be the limiting effect imposed by the unchanged gRNA resources (only one copy of the expression cassette) that is now shared by an increased concentration of proteins since dCas9 doubled. Therefore, the probability of the gRNA binding the dCas9 is higher than that of binding the Cas9. This might result in the dCas9 having an advantage and generating an enhanced homing rate inhibition.

3.2 The dCas9 is actively deposited in the embryo and interferes with dsx expression.

When tested for embryo deposition, dCas9 seemed active and able to interfere with *dsx* expression, probably also in the somatic tissues, due to the leakiness of the *vasa-2* promoter. Both the fertility reductions detected in transheterozygous males and females are consistent with this scenario. In the case of the females, the higher level of infertility detected could be explained if we consider three factors: the target site bound from the dCas9, the leakiness of the *vasa-2* promoter, and the maternal deposition of the dCas9 in the embryo. The gRNA employed, targeted specifically the border between intron 4 and exon 5 of the *dsx* gene, probably compromising the female isoform for which exon 5 is fundamental (Kyrou et al., 2018). The overall transheterozygous female infertility might also be connected with the miss expression of the dCas9 in somatic tissues that are important for the egg development (i.e. oviducts or other ovaries-related tissues). Besides the previous factors mentioned, the lower hatching rate detected could have also been enhanced by the maternal deposition of the dCas9 in the embryo, something that was previously linked to the employment of the *vasa-2* promoter (Andrew Hammond, Galizi, et al., 2016; Andrew Hammond, Kyrou, et al., 2016). Indeed, the expression of the dCas9 in the ovaries and subsequently in the eggs lead to a major deposition in the embryo compared to what may be deposited paternally. And the interference with the *dsx* expression in the embryo, when it is normally highly expressed, is likely to compromise the whole sex-differentiation pathway and possibly result in embryo lethality.

On the other hand, the fertility reduction detected in transheterozygous males could be explained if we consider the impairing mechanism of the dCas9 on the *dsx* gene. The reason why males are not affected by damages produced by an active Cas9 on *dsx* exon 5 is that exon 5 is absent in the male-specific mRNA splice form. Unlike the active Cas9, which introduces mutations at the DNA level that can eventually be removed from the RNA during the splicing phase, the dCas9 interferes with the transcription, probably producing truncated mRNAs. In *Anopheles gambiae*, the *dsx* male isoform contains a small exon downstream of exon 5 that might be excluded from the final mRNA due to dCas9 binding on the DNA during the transcription phase. The effect of the removal of this exon from the final form of the male *dsx* isoform remains unknown, and we do not have enough data to prove that the fertility reduction detected in males can be explained only or mainly with the expression of a truncated male *dsx* mRNA. Nevertheless, it is reasonable to contemplate this as a possible explanation. Finally, if we compare the differences in the hatching rate detected between transheterozygous males and

females, we should also consider the differences in the deposition of the dCas9 in the embryo. Even if transheterozygous males might be able to deposit dCas9 in the embryo through the sperm (as protein or mRNA), it is highly unlikely that they were able to deposit the same amount that can be deposited in the eggs from the females. Hence, it is reasonable to conclude that this experiment furtherly proved the efficacy of the binding of the dCas9 in the germline and probably in the embryo, producing what can be considered a CRISPR-interference (CRISPRi) on the *dsx* gene.

3.3 The dCas9-Pcaf failed to enhance gene expression during meiosis, but there might be room to improve.

Once assessed that the dCas9 effectively binds the DNA target site during gametogenesis, I re-engineered it to deliver a histone acetyltransferase cargo to activate gene transcription. The task for the dCas9 fused to the Pcaf epigenetic editor was to modify the state of the chromatin surrounding the Y-linked transgene and start the expression of the silent I-PpoI nuclease during the later stages of meiosis. Following its activation, the I-PpoI would repeatedly cleave the X chromosomes generating Y-bearing sperm and only male progeny. Moreover, the acetylation of the histone tails on the target transgene is a process that might not necessarily require a long time. Once the histone code is edited on one nucleosome, the modification should spread across the DNA strand (Bowman and Poirier, 2015; Allshire and Madhani, 2018). However, the delivery of the Pcaf HAT to activate the Y-linked I-PpoI did not result in a male bias progeny. Also, quantifying the I-PpoI transcript levels by qRT-PCR didn't show any significant increase. Unfortunately, I could not accurately analyse the acetylation levels around the target gene. A ChIP-Seq analysis requires many cells that would be difficult to collect from tiny mosquito testes. New techniques like CUT&RUN (Cleavage Under Target & Release Using Nuclease) and CUT&TAG (Cleavage Under Target & Tagmentation) have the potential to perform an accurate epigenetic analysis with a relatively small quantity of starting material (Janssens et al., 2018; Kaya-Okur et al., 2019). However, the scarcity of information in the scientific literature and the need to set up and optimise a protocol for *Anopheles* mosquitoes limited their application because of time constraints.

Nevertheless, this was just a first attempt, and there is room for improvement. Together with an accurate epigenome analysis, it should be possible to tailor this system for a more robust effect. Choosing a more appropriate epigenetic editor and or transcriptional activator, and the

use of more than one gRNA are just two of the possible solutions. Recently, a paper published interesting results on several CRISPRa molecules in *Drosophila* (Waters et al., 2018). Particularly encouraging results were achieved using the HAT “Nejire”, an orthologue of the human P300. In *A. gambiae*, this gene is annotated as AGAP000029, with a structure closely resembling the one seen in *Drosophila*.

In recent years, many new activator molecules have been employed to activate genes using dCas9 as a carrier (Lin et al., 2015; Chavez et al., 2016; Ewen-Campen et al., 2017; Roca Paixão et al., 2019). This approach is generally known as CRISPR activation (CRISPRa) and proved highly effective in many organisms. Even if meiotic cells remain a challenging environment, the constant discovery and testing of new molecular tools to activate genes give hope for future applications.

Another attractive employment of dCas9 as a tool for gene activation aimed to use the nuclease-dead protein to dislocate portions of the DNA alpha helix in nuclear compartments that might be more transcriptionally active (Wang et al., 2018; Lin et al., 2019). This strategy might be particularly interesting in studying one of the factors that probably contribute to meiotic silencing; the segregation of the X-Y couple (the XY body or sex body) in a separate nuclear compartment (J. M. Turner et al., 2005; Turner et al., 2006). It is unknown if this particular nucleus compartment is depleted in transcriptional factors and or transcriptional machinery. Still, given the fact that only the silenced X and Y chromosomes are segregated in this nuclear partition, it is reasonable to assume that it might contribute to their silencing. It was shown that the cell nuclear organisation in eukaryotes is highly heterogeneous and actively contributes to gene regulation, with regions near nuclear pore complexes (NPC) that seems more transcriptionally active (Guo et al., 2014). Therefore, displacing targeted portions of DNA in more transcriptionally active nuclear compartments might help overcome some of the transcriptional limitations imposed by MSCI.

One of the main challenges in testing all these molecular tools for developing meiotic drives is the need to create transgenic mosquito lines. The efficiency in generating transgenic lines remains quite low when using systems for transposase-induced integration (like PiggyBac) that have the advantage of multiple targets for insertion. This efficiency drops dramatically when using systems for site-specific integration (Bernardini et al., 2014). The Y-linked *gypsy-I* PpoI line generated for this study required more than six months of injections. This long time is probably also due to the highly heterochromatic nature of the Y chromosome, which poses additional challenges for site-specific integration. Nevertheless, autosomal-linked attP docking

lines also required additional time and effort to achieve a successful integration compared to the transposase-induced counterpart.

Recently, the development of the first male-specific cellular line of *A. gambiae* opened a wide range of possibilities for testing Y-chromosome-linked molecular systems (Krzywinska, Ferretti and Krzywinski, 2022). The establishment of this line not only represents a valuable tool for the *in vitro* study of the *A. gambiae* male-specific transcriptome but can also be a useful *in vitro* platform to perform an accurate screening on Y-chromosome-targeting CRISPR systems before their transposition *in vivo*.

3.4 *The gypsy I-PpoI line might be tuned to achieve the Y-drive.*

The employment of the *gypsy* insulators produced, so far, the best results in increasing the transcripts levels of the Y-linked I-PpoI nuclease, and it also produced signs of activity of the I-PpoI. Moreover, the Y-linked expression of the I-PpoI transcript was also detected in the line without the *gypsy* insulators, even if at extremely low levels. This news slightly changes the paradigm of complete inactivation of the Y chromosome induced by MSCI and opens a range of new possibilities to achieve the Y-drive.

When R. Galizi et al. developed the autosomal sex distorter, they tested several variants of the I-PpoI nuclease, each with a specific activity and half-life (*Figure 1-7*). Considering that our transgene is able to give some expression from the Y chromosome, even if at a level that is not sufficient to achieve a relevant sex bias, it should be possible to compensate for the low transcript production with a more active variant of the I-PpoI nuclease. The variant employed in the autosomal sex-distorter line PMB1 is the W124L. The other two variants tested (W124A and L111A) showed improved thermal stability, higher half-life and specific activity, while the wild type nuclease outshines all the engineered variants on all three parameters. These more active and thermally-stable versions of the nuclease produced sterility or decreased fertility when expressed from the autosomes (Galizi et al., 2014). But as shown from our qRT-PCR, the expression level of the *B2tub* from the autosomes is 153 folds its expression from the Y-chromosome. It would then be possible to employ a more active version of the I-PpoI to compensate for the reduced activity of the *B2tub* from the Y and maybe achieve some sex distortion.

Moreover, we have employed just a portion of the B2 tubulin 5' UTR so far. Its employment was enough to express the I-PpoI nuclease at appropriate levels to give sex distortion from the autosomes. But the employment of a bigger 5' flanking region, including the full 5' UTR (as made for the *vasa* promoter) (Papathanos et al., 2009) might also improve the activity of the regulatory region when placed on the Y chromosome.

Another consideration comes from the females produced from the autosomal sex distorter line. In those females, the maximum deletion frequency detected from amplicon sequencing was slightly more than 0.4%. This might be considered a threshold above which lethality arises. Considering that the target gene is a ribosomal subunit, it is not difficult to hypothesise that if too many repeats are damaged, the zygote will not survive due to ribosome shortage.

Maybe the strong expression achieved from the autosome is not even necessary to achieve a sound sex distortion. Therefore, combining these two elements, an improved *B2tub* regulatory

region and a more active version of the I-PpoI nuclease might be enough to achieve a self-sustainable sex distortion driven by the Y chromosome.

3.5 Delaying the expression of a gene with RNA-targeting Cas proteins

One of the main problems in overcoming MSCI is that it is not available, at the moment, a regulatory region able to trigger the expression of the X-shredder nuclease (I-PpoI or Cas9) just before the beginning of the meiotic silencing. The *vasa* or *zpg* promoters are active too early during the gametogenesis, and any nuclease driven by those promoters targeting the X chromosome it's likely to destroy it when the cells are still at the stage of GSC, causing meiotic arrest. Any attempt to express these nucleases under the control of these promoters led to mosquito infertility. On the other hand, the *B2tub* is active when MSCI is already in place and, hence, is extremely downregulated when placed on the Y chromosome.

Nature provides plenty of examples of transcribed genes, with delayed translation regulated at the RNA level. One of those happens in *Drosophila* spermatocytes, which accumulate translationally-repressed mRNAs, which are required to produce important proteins once the DNA is heavily compacted in the acrosome (Barckmann et al., 2013)

There are already groups working on the spatiotemporal regulation of genes using a mix of CRISPR-interference (CRISPRi), and re-engineering of gRNAs and riboswitches (Galizi and Jaramillo, 2019). All these approaches work at the DNA level. For example, hampering the transcription of a gene with a dCas9 until a specific trigger interferes with the protein or the gRNA, eventually freeing the alpha helix. An opposite example is a Cas protein complexed with a gRNA that forms a loop at the spacer level, blocking the protein from binding to the DNA. Also, in this case, an opportunely designed trigger RNA (trRNA) may pair with one side of the loop that will then unfold, exposing the portion containing the spacer sequence and allowing DNA binding (Galizi and Jaramillo, 2019).

The constant research for new CRISPR-associated molecular tools has produced some interesting discoveries about Cas proteins able to bind the RNA (Abudayyeh et al., 2017; Dugar et al., 2018; Strutt et al., 2018; Terns, 2018). For example, *S. pyogenes* Cas9 and other Cas9s discovered in other prokaryotes (i.e. *S. aureus* Cas9 and *C. jejuni* Cas9) can cleave ssRNA providing the presence of an ssDNA oligonucleotide that match the PAM sequence (named PAMmer). Also, proteins like Cas13 have naturally evolved to target RNA, and their deactivated variants are already employed for CRISPR-RNAi (Abudayyeh et al., 2017).

A study performed *in vitro* on dCas9 highlights how it is possible to modulate the time of residency of the protein on the alpha helix through the insertion of specific mismatches in key points of the gRNA (Boyle et al., 2017). According to the distance of the mismatch(s) from the PAM, and the nucleotide substituted, it would be possible to put a virtual timer on the dCas9

to release the binding site. If this is also applicable when employing RNA-binding Cas proteins, it would be possible to create a time-controlled CRISPR-RNAi mimicking the natural mechanism of accumulating translationally repressed mRNA. For example, it would be possible to express the transcript of the I-PpoI nuclease under the pre-meiotic *vasa* promoter from the Y-chromosome before MSCI and repress its translation using a nuclease-dead Cas13. Providing a gRNA tuned with sequence-specific mismatches should trigger the release of the mRNA in a time-dependent manner allowing the synthesis of the I-PpoI nuclease in later meiotic stages. Moreover, this approach would not necessitate the expression of any external trigger that, in the context of developing a Y-linked gene drive, should also be placed on the Y chromosome to be co-inherited and will face the limitations imposed by MSCI.

3.6 The Y-linked meiotic sex distorter is one of the most powerful gene drive strategies but poses several challenges.

In this study, I described the attempt to develop a system to rescue and enhance the activity of a Y-linked transgene in the human malaria vector *A. gambiae*. This transgene contains an expression cassette for the I-PpoI nuclease under the control of a portion of 496 bp of the 5'UTR of the *beta-2 tubulin* gene (*B2tub*). When expressed from the autosomes, this transgene shredded the *A. gambiae* X-chromosome and granted a Y-specific bias in the production of spermatozoa (Galizi et al., 2014). Therefore transgenic males carrying this gene gave a heavily male-biased progeny when mated with wild type females. This is an extremely promising strategy to be implemented as genetic sex sorting, which might help to reduce the number of potential pathogen-carrying females in wild type populations. Moreover, the 28S ribosomal subunit repeats on the *A. gambiae* X chromosome grant a reservoir of cleavable targets for the I-PpoI nuclease (or eventually for a Cas9), making this strategy virtually immune to any selection for resistant alleles. Furthermore, when we analysed the mutations produced around the I-PpoI cut site in the few females produced from the autosomal sex distorter, the presence of mutations was extremely low ($< 0.5\%$), meaning that few cuts might be enough to obtain female lethality.

Nevertheless, this strategy would require multiple releases of transgenic mosquitoes with all the problems connected to mosquito mass rearing. The presence of the transgene on a non-sexual chromosome grants a simple Mendelian inheritance, with only half of the male-biased progeny inheriting the transgene. Hence this strategy is not self-sustainable, and the transgene will be lost over time in a wild type population if not constantly replenished with new transgenic individuals.

Repositioning the transgene on the Y-chromosome would grant a self-sustained gene drive that will spread within wild type populations without requiring new releases of transgenic mosquitoes. Several mathematical models showed how the Y-drive sex-distorter would be among the most effective gene drive strategy in multiple scenarios (Beaghton and Burt, 2022). The only limitations predicted would arise in systems characterised by an inbred low-density population, in which the very transgenic males responsible for the male bias would struggle to find a female partner (Beaghton and Burt, 2022). Effectively expressing genes from the Y chromosome would be crucial to opening the way to other Y-linked pest control strategies, potentially orders of magnitude more effective than standard SIT (Geci, Willis and Burt, 2022).

Unfortunately, the editing and employment of the Y-chromosome as an effective transgene docking site pose several challenges. Particularly in *A. gambiae*, the absence of a complete genome assembly for the Y (due to its highly repetitive and heterochromatic structure) makes it almost impossible to locate loci suitable for stable and effective transgene insertion. Moreover, when transgenes were actually inserted on the Y (Bernardini et al., 2014) they showed an extremely variable expression even among individuals belonging to the same transgenic line. Finally, the expression of genes during meiosis, like the meiotic sex distorter, looked almost completely shut down, most likely due to MSCI.

For these reasons, approaches that try to edit the Y-chromosome epigenome locally, like the site-specific delivery of epigenetic editors, or that insulate transgenes from the nearby chromatin regulation might represent a valuable solution to unlock the potential of Y-chromosome-based pest strategies.

3.7 The absence of a Y-chromosome assembly limits Y-linked strategies for pest control.

Efficient engineering of the Y-chromosome might have an enormous impact on pest and vector control strategies (Geci, Willis and Burt, 2022).

Particularly regarding mosquito control, in which the females are responsible for diseases diffusion, being able to insert and control genes from the only chromosome that is always male-inherited would open a wide array of possibilities (Alphey et al., 2010; Wang et al., 2021; Geci, Willis and Burt, 2022)

Unfortunately, the engineering of the Y Chromosome is heavily hampered by the lack of a complete assembly. Sex chromosomes are usually difficult to sequence and assemble due to their highly repetitive sequences (Rhie et al., 2021). The *A. gambiae* Y-chromosome was particularly challenging, and only recently, some light was shed on a few genes and small regions of the Y-chromosome (Hall et al., 2016; Zamyatin et al., 2021). Nevertheless, the vast majority of the chromosome remain unknown. It is known that *A. gambiae* X and Y chromosomes have repeated sequences that share some homology and might recombine (Hall et al., 2016). If recombination happens, this means that X and Y have to pair to a certain extent. Pairing is the condition for a sexual chromosome to escape MSCI (Turner, 2007); therefore, having a full assembly of the Y chromosome would potentially allow to locate portions of the chromosome that are active during MSCI.

Thanks to sequencing technologies that allow megabase-sized long reads, some progress was made in sequencing highly repeated regions. Nanopore, PacBio and Hi-C are among those technologies that promise to shed light on chromosomal regions still largely unknown in different species (Rhie et al., 2021). Sexual chromosomes are not the only ones difficult to read. The same highly repetitive structure is present in autosome pericentromeric regions and telomeres.

The most recent *Anopheles* complex assembly refers to *A. arabiensis* and *A. coluzzi* species. They employed a mix of long-reads technologies and an improved bioinformatics pipeline to increase the coverage of the previously generated assemblies (Zamyatin et al., 2021).

The fight against malaria must focus on achieving a sound and reliable assembly for the *Plasmodium* vector species. Genetic engineering strategies can still make a difference, but to

completely eradicate vector-borne diseases, we must be able to look at completely assembled genomes.

3.8 Conclusions

The work hereby presented tried to address the problem of the silencing of heteromorphic sexual chromosomes from an epigenetic perspective. Despite the overall unsuccess in effectively counteracting MSCI, several experiments were carried out to investigate why the strategies employed here were unable to fully rescue the activity of Y-linked transgenes during spermatogenesis. Moreover, some of these experiments shed some light and provided new ideas on the dynamics and the possible uses of the molecular tools employed here.

The most relevant molecular tool employed in this work was the nuclease-dead Cas9. This protein proved to bind the DNA alpha helix during the first stages of meiosis and to be able to compete for the binding site with an active Cas9, effectively reducing the cutting potential of the latter. The patterns of homing interference also revealed that, most likely, the dCas9 is removed from the DNA at every replication cycle of the GSC, a factor to be considered for future application of the nuclease-dead protein in contexts subjected to high cellular replication. Moreover, despite its early expression during spermatogenesis due to the control of the *vasa 2* promoter, the dCas9 was effectively maternally deposited, and it was able to reduce female fertility, interfering with the embryonic expression of the *dsx* gene. These findings increase the knowledge of how dCas9 behave in during cellular divisions and open the way to new possible applications. One of these new applications could be using the dCas9 to counteract gene-drive systems when employing other anti-CRISPR proteins is not feasible or toxic for the carrier (Marino et al., 2020; Taxiarchi et al., 2021b). Another possible example could be the employment of an embryonically expressed dCas9 to curb the Cas9 activity in the embryo to reduce the generation of resistant alleles or embryo lethality (Andrew Hammond, Galizi, et al., 2016).

Unfortunately, the site-specific delivery of the Pcaf histone acetyltransferase failed to counteract the MSCI on *A. gambiae* Y chromosome. Even the truncated version of the protein that, according to similar experiments carried out in the literature (Ali et al., 2017; Waters et al., 2018), should have had an increased acetylation activity, was unable to significantly increment the activity of the Y-inserted “X shredder” expression cassette. This does not mean,

that other HATs couldn't be successfully employed for the same task, and given the importance of such a task, it might be worth it to keep investigating and testing other HATs.

The employment of the *gypsy* retrotransposable elements as chromatin insulators were able to modestly increase the transcript levels of the Y-linked transgene. Even if the increment in the transcriptional level was not enough to produce a detectable sex bias in the progeny of modified mosquitoes, it was enough to be detected via q-PCR and to produce cuts in the X-linked ribosomal repeats. Considering the fact that the nuclease employed to target the repeats on the X-chromosome was considerably destabilised to reduce its half-life and specific activity, it could be possible to keep working, in conjunction with the employment of *gypsy* insulators, on the protein thermal stability to finally achieve a full-working Y-linked gene drive.

Finally, an important point to highlight is that the Y-linked gene drive is, at the moment, one of the most powerful gene drive strategies available (Burt, 2003). Any small step toward achieving such a goal would get us closer to eradicating malaria. Moreover, achieving a consistent transgene expression from sexual chromosome is also a landmark for more conservative self-limiting strategies (Burt and Deredec, 2018; Geci, Willis and Burt, 2022) that can potentially be employed in other vector or pest species and overcome regulatory processes and public concern more easily.

4 Methods

4.1 PCR amplification and gel extraction

Standard PCRs were performed using the Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB, #M0531S). PCR reactions were incubated in a Veriti™ Thermal Cycler following the manufacturer's instructions. The initial denaturation step was set at 98°C for 30 seconds, followed by a 10-second denaturation step at 98°C, a 30-second annealing step and an elongation step at 72°C for 35 cycles. The final elongation step was set at 72°C for 10 minutes. As DNA templates, I used plasmid DNA (0.5 to 1 ng), genomic DNA (up to 50 ng) and PCR products (10 to 20 ng) per reaction. The resulting products were loaded on 0.7% agarose gel (5, 25 or 50 µL) with an appropriate volume of 5X loading dye (Bioline, #BIO-37045). The gels were set to run at 100 V for 60 minutes and then visualised on an Azure 200 Gel Visualization System (Azure Biosystem). After the visualisation, the appropriate bands were excised using a UV transilluminator and purified using a Monarch DNA gel extraction kit (NEB, #T1020L). Finally, the DNA yield was quantified with a NanoDrop spectrophotometer (ThermoFisher).

4.2 Gibson Assemblies and Bacterial Transformations

The Gibson Assembly Cloning Kit (NEB, #E5510S) was used to perform Gibson assemblies according to the manufacturer's instructions. DNA fragments with 16 to 40 bp overhangs were combined in a reaction, with the longest fragment acting as the "vector" and all molar ratio calculations referred to it. Approximately 100 ng of the vector was combined with one or more inserts in molar ratios that ranged from 1:1 to 1:3 in a final volume of 10 µL. The 10 µL of vector-fragment mixes were combined with 10 µL of Gibson Assembly master mix and incubated at 50°C for one hour. Next, 2 µL of the Gibson reactions were added to 50 µL of NEB 5-alpha competent cells (NEB, #C2987H) that had been freshly defrosted. This mixture was left on ice for 30 minutes before the cells were heat-shocked in a pre-warmed water bath at 42°C for 30 seconds and immediately placed on ice for 2 minutes. Then, 950 µL of SOC

medium were added to each tube and incubated on a shaking incubator at 37°C for one hour. After a gentle 2-minute spin at 500 g, most of the cells were collected at the bottom of the vial, and approximately 800 µL of the supernatant was removed from each tube, leaving behind 200 µL of the medium, used afterwards to resuspend the pellet. Finally, 100 to 200 µL of the culture were plated on agarose Petri dishes containing the appropriate antibiotic and incubated overnight at 37°C. The resulting colonies were screened via colony PCR (see below).

For cell transformations that did not involve Gibson assemblies, OneShot TOP10 competent cells (ThermoFisher, #C404010) were used in accordance with the manufacturer's protocol. In brief, 2 µL of plasmid (1 to 10 ng) was added to 25 µL of freshly thawed cells, and the mixture was incubated on ice for 30 minutes. The cells were then heat-shocked in a water bath at 42°C for 30 seconds and immediately placed on ice. 250 µL of SOC medium was added to each tube, and the tubes were incubated at 37°C in a shaking incubator for 1 hour. Afterwards, 50 to 100 µL of the culture was plated on agar Petri dishes with the appropriate antibiotic and incubated at 37°C overnight. Bacterial colonies were picked from the Petri dishes using sterile 10 µL pipette tips and placed in 5 mL (for a Miniprep) or 100 mL (for a Maxiprep) of LB media with the proper antibiotic and incubated at 37°C overnight on a shaking incubator. The cultures were then pelleted via centrifugation at 4000 g for 15 minutes and stored at -20°C or immediately used for plasmid preparation. Plasmids were isolated and purified using the QIAprep Spin Miniprep kit (Qiagen, #27104) or ZymoPURE II Plasmid Maxiprep Kit (Zymo Research, #D4202).

4.3 Restriction digestions

Restriction digestions were carried out using FastDigest Restriction Enzymes (ThermoFisher) following the manufacturer's instructions. Up to 4 µg of DNA were incubated with 1 to 2 µL of the appropriate enzyme for 1 hour at 37°C, and then inactivated at 80°C for 20 minutes. The reaction mixture was then separated via electrophoresis on a 0.7% agarose gel at 110V for 60 minutes and visualised on a Gel Visualization System (Azure Biosystem). The desired bands were excised from the gel using either the QIAquick Gel Extraction Kit (Qiagen, #28706X4) or Monarch DNA Gel Extraction Kit (NEB, #T1020L), following the manufacturer's protocols. The digestions of the X-linked ribosomal repeats amplicons were performed using the I-Ppo I restriction enzyme (Promega, #R7031). Approximately 500 ng of PCR-amplified DNA was

digested at 37°C for 2 hours and inactivated at 80°C for 20 mins. The digestion reaction was run on a 1% agarose gel at 100V for 1 hour and visualised on an Azure 200 Gel Visualization System (Azure Biosystem).

4.4 Colony-PCR

To verify the success of the Gibson assemblies, colony-PCRs were performed on individual colonies. Sterile 10 µL pipette tips were used to transfer each colony to a new Petri dish with a numerical grid to allow colony identification. After transferring the colonies to the grid plate, the same pipette tip was used to transfer a small number of cells into a PCR tube containing 5 µL of water, which served as the template for the PCR reaction.

The PCR reactions were performed using a Qiagen Fast Cycling PCR kit (Qiagen, #203743), with 3 µL of the template used in a reaction with a final volume of 10 µL. The reactions were incubated in a Veriti™ Thermal Cycler, starting with an initial denaturation at 95°C for 5 minutes to break down bacterial cells and release DNA into the reaction. Denaturation was then performed at 96°C for 5 seconds, followed by annealing for 5 seconds and elongation at 68°C for 35 cycles. Finally, a 72°C elongation step was performed for one minute to complete the reaction. The PCR products were run on a 1% agarose gel at 100 V for one hour and visualised using a Gel Visualization System (Azure Biosystem).

4.5 Genomic DNA extraction

Mosquito DNA was extracted using the Wizard Genomic DNA Purification kit (Promega, #A1120) according to the manufacturer's specifications. Mosquitoes were homogenised with micro pestles in 200 µL Lysis solution in 1.5 mL Eppendorf tubes and incubated at 65°C for 30 minutes. The lysate was then mixed with 70 µL of protein precipitation solution, incubated on ice for 5 minutes, and then vortexed and centrifuged for one minute at 13000 g on a benchtop centrifuge (Eppendorf, #2231000909). The supernatant was transferred to a fresh Eppendorf tube, mixed with 200 µL of isopropanol (Sigma-Aldrich, #I9516), and then centrifuged at 4°C for 10 minutes to allow the precipitation of the DNA. The supernatant was removed, and the DNA pellet was washed with 200 µL of 70% molecular biology grade ethanol (Fisher

Scientific, #64-17-5) and then newly centrifuged at 4°C for one minute. The ethanol was then carefully discarded to avoid the loss of the DNA pellet, and the tube was left to air-dry for 15 minutes. The genomic material was resuspended in 30 µL of UltraPure DNase/RNase-free distilled water (Invitrogen, #10977035) and quantified using a NanoDrop spectrophotometer (ThermoFisher).

4.6 RNA extraction and qRT-PCR

All the RNA extractions were performed in sterile conditions using the Monarch Total RNA Miniprep kit (NEB, #T2010S) according to the manufacturer's recommendations. RNA extractions from wild type mosquitoes were performed as follows. Briefly, a pool of three pupae was mechanically homogenised in 300 µL of DNA/RNA protection buffer. Then, 30 µL of Proteinase K reaction buffer were mixed with 15 µL of Proteinase K in a separate tube. The 45 µL of the proteinase K solution were added to the 300 µL of pupae lysate and incubated for 30 mins at 55°C. Following the incubation, I briefly vortexed the lysate and spun it at 16000 g for 1 minute to pellet the debris, and I transferred the supernatant to a clean 1.5 mL tube. I added an equal volume of RNA Lysis Buffer and proceeded exactly as described in the kit's manual Part 2: RNA binding and Elution. The RNA was also treated with the DNase provided with the kit to remove eventual residual genomic DNA.

Regarding the RNA extracted for qRT-PCR, I first dissected the adult male mosquitoes in two halves, separating the abdomens (containing the testes) from the carcasses (thorax and head). To avoid RNA degradation, I placed the dissected parts in clean 1.5 mL tubes with 300 µL of DNA/RNA protection buffer on ice. I then proceeded with the RNA extraction as described above. The RNA quantities were measured via DeNovix DS-11 spectrophotometer (DeNovix) and normalised before the retro-transcription. The RNA retro transcription was performed using the Invitrogen Superscript IV kit (Invitrogen, #18091050) following the manufacturer's protocol. I combined 1 µL of 50 µM Oligo d(T)₂₀, 1 µL of 10 mM dNTP mix, and 50 ng of total mRNA in a 13 µL solution. I gently vortexed and incubated it at 65°C for 5 minutes and then on ice for 1 min. I prepared a second solution (RT reaction mix) as follows: 4 µL of 5x SSIV Buffer, 1 µL of 100 mM DTT, 1 µL of RNase out and 1 µL of Superscript Reverse Transcriptase. I combined the two solutions and incubated them at 55°C for 10 mins and then inactivated the reaction with incubation at 80°C for 10 minutes. To remove eventual residual

RNA, I added 1 μ L of RNase H and incubated again for 20 minutes at 37°C. For the qRT-PCR, I employed the SYBR Green real-time PCR master mix (ThermoFisher, #A46110). In addition, the cDNA Ct (Cycles to Threshold) was normalised with the housekeeping gene RPL19 using the following primers (RPL19 FW 5'-ccaactcgcgacaaaacatt-3', RPL19 RV 5'-accggcttcttgatgatcaga-3'). The primers employed to target the I-PpoI were DM182 (5'-atggcgtcaccatgctcaaattc-3') and DM236 (5'-gttccttgagaccgccacc-3') that produced an amplicon of 113 bp. The PCR reaction was performed on a StepONEplus Real-Time PCR System thermocycler (ThermoFisher, #4376600). Three biological replicates were analysed for each transgenic line.

The raw data obtained from the q-PCR were analysed using the Double Delta Ct method ($\Delta\Delta Ct$). Once I obtained the average of the Ct values for each target and housekeeping gene triplicate, I calculated the ΔCt values by subtracting the Ct average of the housekeeping gene from the Ct average of the target gene in each sample for each experimental condition. Then, I calculated the difference between the ΔCt of the control condition (the Y-I PpoI line abdomen samples) and each other experimental condition. Finally, I calculated the $2^{-\Delta\Delta Ct}$ value for each $\Delta\Delta Ct$ to get the expression fold change of each experimental condition compared to the control.

4.7 Statistical analysis

All the statistical analyses were carried out using GraphPad Prism (www.graphpad.com). The samples were first tested for normality and then tested again with the most appropriate statistical test. Since most of the samples had a not-normal distribution, the main statistical test used in these studies was the non-parametric Kruskal-Wallis test. The significance threshold (α) was set at 0.05. Hence only the P values < 0.05 were considered significant. When only two samples were compared (and then the Kruskal-Wallis test was not appropriate any more), like in the case of the females from the gypsy I-PpoI line compared with wild type females, the statistical test employed was the Wilcoxon test. The error bars in the figures represent the Standard Error of Mean (SEM) calculated as follows:

$$SEM = \frac{SD}{\sqrt{n}}$$

Where SD is the Standard Deviation and “n” is the number of elements in the sample.

4.8 Bioinformatics tools

The genomic databases used in this project were Vectorbase (www.vectorbase.org) to retrieve *A. gambiae* sequences and Flybase (www.flybase.org) to compare and look for orthologs in *D.melanogaster*. Benchling (www.benchling.com) was employed to design DNA constructs and PCR primers, set DNA digestions, and look for gRNAs target sites. HMMER (www.hmmer.org) was used to find homologies between protein domains from *A. gambiae* and *D.melanogaster*. GraphPad Prism v9.0 (www.graphpad.com) was used to perform statistical analyses and plot data.

4.9 The assembly of the Gal-UAS plasmids

To produce the plasmid containing the Gal 4 flanked by different promoters and terminator sequences and the plasmid for the responder line containing the UAS and the I-Ppo I under the control of the *beta-2* promoter, I used the primers listed in *Table 4-1*. All the PCRs described here were performed with Phusion High Fidelity Polymerase (NEB, #M0531S) according to the manufacturer's specifications described above. The Gal4 was amplified from the plasmid pSL*Cbxy-gal4-GF (kindly provided to our lab by Dr Gareth Lycett), producing an amplicon of 592 bp. The *beta-2* promoter was amplified from the plasmid pBac\ (b2-GFP-124LPpoI)(3xp3-DsRFP)-[124L] (available in our lab) producing an amplicon of 542 bp while the *beta-2* terminator produced an amplicon of 524 bp. The *vasa-2* promoter was amplified from [C63] 3xP3-attB-GFP_VASA2prom-RyAcrIIA4-VASA3utr and produced a fragment of 2338 bp, while its terminator sequence was 1048 bp in length. Finally, the zpg promoter and terminator were amplified from [C108] (attB-DsRed_zpg-hCas9-F2A-eGFP_(wt)b2prom-NLS-RyAcrIIA4), producing two amplicons of 1072 bp and 1074 bp respectively. The backbone was obtained via digestion of 4 µg of the plasmid [C18] (attBCFP_b2(5')-hCas9-F2A-eGFP-b2(3')) (available in our lab) with the restriction enzymes BshTI and SgsI (Thermofisher, #FD1464, #FD1894). The digestion was incubated for 1h at 37°C and then inactivated for 20 mins at 80°C. I isolated the DNA band corresponding to 3556 bp, which included the plasmid backbone, a DsRed marker, and an attB site for site-specific integration. After gel purification using a Monarch DNA Gel Extraction Kit from NEB, I combined the

appropriate fragments using the Gibson Assembly cloning kit (NEB, #T1020L) with 50-100 ng of vector and a 1:1 molar ratio. To confirm successful cloning, I performed colony PCR on the colonies obtained, followed by re-inoculation of the successful ones in 5 mL of LB solution with ampicillin and growth for 12h in a shaker incubator at 37°C. Next, I purified the resulting plasmids using a QIAprep Spin Miniprep Kit from (Qiagen, #27104) and sequenced them to ensure proper assembly. Once I verified the sequence correctness, I used the ZymoPURE II Plasmid Maxiprep kit (Zymo Research, # D4202) to Maxi-prep the successful plasmids, resulting in the following plasmids: attB_DsRed_b2_Gal4, attB_DsRed_Vasa2_Gal4, and attB_DsRed_zpg_Gal4.

To assemble the plasmid containing the UAS repeats for the reporter line, I used the same backbone as the Gal4 plasmids described above. I amplified the fragment containing a *gypsy* insulator sequence followed by 14 UAS repeats (790 bp) from the plasmid pBAC\ECFP\UAS14i-LUC-YFPnlsii-gypsy, generously provided to our lab by Dr Gareth Lycett. Additionally, I amplified a fragment containing the *beta-2* promoter-GFP-PpoI-*beta-2* terminator (2316 bp) from the plasmid pBac(b2-GFP-124LPpoI)(3xp3-DsRFP)-[124L], as well as a second *gypsy* sequence (428 bp). I used the Gibson Assembly cloning kit (NEB, #T1020L) to combine these fragments, following the same steps described before.

Element	Primer Name	Sequence (5'-> 3')
Gal 4 bd	DM 7	<u>ctgaagcaagcctcctgaag</u> ATGAAGCTACTGTCTTCTATCGAAC
	DM21	TTAGTTACCCGGGAGGATGTC
beta-2 tub promoter	DM5	<u>cgaccggtcaccggtg</u> CTAGCGTTCATAATTGATATAGTTTTG
	DM6	<u>caggaggcttgcctcaagcttgatgc</u> CTTGATATCTTTGAAACTGTGA
beta-2 terminator	DM22	<u>gacatcctcccggtactaa</u> CTAAAGCTAAATTGAACACCCTAAAT
	DM10	<u>ggcaccgcagcgcg</u> CGATTTAAGGACCGATTCCAA
vasa- 2 promoter	DM15	<u>cgaccggtcaccggtg</u> ATGTAGAACGCGAGCAAATT
	DM16	<u>cttcaggaggcttgcctcaagcttgatgc</u> GTTTCCTTTCTTTATTACCGG
vasa- 2 terminator	DM23	<u>gacatcctcccggtactaa</u> AGCGGCCGCTTGGGG
	DM18	<u>ggcaccgcagcgcg</u> CGCGAAAATGTGGCCATTAAACAG
zpg promoter	zpgFW	<u>ggatcccccaccggtcaccggtg</u> TCTTCCTGCGGCACATCCCTC
	zpgRV	<u>caggaggcttgcctcaagcttgatgc</u> CTCGATGCTGTATTGTGTTG
Zpg terminator	DM24	<u>gacatcctcccggtactaa</u> GAGGACGGCGAGAAGTAATC
	DM14	<u>ggcaccgcagcgcg</u> TCGCATAATGAACGAACCAAAG

Gypsy-UAS	DM1	<u>cgactctagaggatccccga</u> TCACGTAATAAGTGTGCGTTG
	DM19	GCTAGAGTCTCCGCTCGAC
B2-GFP- PpoI-B2ter	DM39	<u>gtcgagcggagactctage</u> CTAGCGTTCATAATTGATATAGTTTTG
	DM33	CTAGTTAAGATACATTGATGAGTTTG
Gypsy rev	DM42	<u>caaactcatcaatgtatcttaactag</u> TCACGTAATAAGTGTGCGTTG
	DM4	<u>gccacatgtggcgc</u> GAATTGATCGGCTAAATGGTATGGC

*Table 4-1 Primers employed for assembling the GAL-UAS plasmids.
The underlined sequences represent the Gibson overhangs.*

4.10 The assembly of the dCas9-mCherry plasmids

To disable the nuclease activity of the Cas9 enzyme, I used the plasmid [C18] (attBCFP_b2(5')-hCas9-F2A-eGFP-b2(3')) that was already present in our lab. This plasmid contained a Cas9 under the control of the *beta*-2 promoter and was less than 10kb in length. To introduce the two mutations needed to disable the nuclease activity of the Cas9, I re-amplified the entire plasmid in two different fragments using primer pairs DM52/DM53 and DM54/DM55, as described in *Table 4-2*. These primer pairs contained the necessary single and two nucleotide mutations (A>C and CA->GC, respectively) that produced two aminoacidic substitutions (D10A and H840A) in the RuvC and HNH catalytic domains, respectively. The two fragments obtained were 2511 bp and 7228 bp in length and were re-assembled using a two-fragment Gibson assembly. Briefly, I mixed 50 to 100 ng of the larger fragment in 1:1 and 1:3 molar ratios with the smaller fragment and incubated them for one hour at 50°C with an equal volume (10 µL) of Gibson Assembly master mix, as described by the manufacturer (NEB, #T1020L). I then transformed NEB 5-alpha competent cells, and plated them on LB-agar plates with ampicillin. I performed colony PCR on the colonies obtained and re-inoculated the successful colonies in 6 mL of LB medium with ampicillin to perform a miniprep. The minipreps were then digested to verify the correctness of the digestion profile, and the successful ones were sequenced. I finally obtained the plasmid [C18d], which contained a dCas9 instead of an active Cas9.

To assemble the plasmid D1 and D0 containing the dCas9 expression cassette under the control of the *vasa*-2 promoter, I used the plasmid p165 described in Hammond et al. 2016. I digested this plasmid with the restriction enzymes PvuI, PacI, and MluI (ThermoFisher, #FD0624, #FD2204, #FD0564) and gel-extracted the band corresponding to 4219 bp, which contained the plasmid backbone. I then amplified the dCas9 with the F2A epitope (4371 bp) from the plasmid [C18d], the *vasa*-2 promoter (2329 bp), a cassette expressing the GFP marker under the control of the 3xP3 promoter (1327 bp) from p165, and the mCherry fluorescent protein (746 bp) from the plasmid pB(B2 mCherry) using the oligos described in *Table 4-2*. Finally, I amplified a fragment containing the *vasa*-2 terminator and the cassette containing the U6 promoter, a Golden Gate cloning site for the gRNA spacer, and the gRNA backbone (from p165). Following the previously described steps, I performed two separate Gibson assemblies to obtain the plasmid D1 (PB-*vasa*-2-dCas9F2AmCherry-U6gRNA-[3xP3GFP]). First, I assembled an intermediate plasmid containing the backbone, the *vasa*-2 promoter, the dCas9

and the mCherry fluorescent protein. Then, to obtain the D1 plasmid, I linearised the intermediate plasmid with the restriction enzyme XhoI (ThermoFisher, #FD0695) and then performed a second Gibson assembly to knock in the *vasa-2* terminator with the gRNA expression cassette and the 3xP3 GFP cassette.

To generate the plasmid D0 (PB-Vasa2dCas9F2AmCherry[3xP3 GFP]), the newly obtained D1 plasmid was re-digested with SgsI (ThermoFisher, #FD1894) to eliminate the U6 promoter, the Golden Gate cloning site for the gRNA spacer, and the gRNA backbone. The 13870 bp band was then gel-purified and re-ligated using a T4 ligase (NEB) as per the manufacturer's instructions. The reaction mixture consisted of 50 ng of the linearised plasmid, 2 µL of T4 DNA Ligase Buffer (10X), 1 µL of T4 DNA ligase (Invitrogen, #15224017), and water up to 20 µL, which was incubated at room temperature for 2 hours. 50 µL of OneShot TOP10 competent cells (ThermoFisher, #C404010) were transformed using 2 µL of the reaction mixture. Twenty colonies were randomly selected and re-inoculated in 5 mL of LB medium with Chloramphenicol. Ligation was confirmed through diagnostic digestion with the restriction enzymes SgsI and Eco32I (ThermoFisher, #FD1894, #FD0303). The successful excision of the U6-gRNA expression cassette was confirmed by the absence of the relative 256 bp band produced by SgsI.

To introduce different gRNA spacers, Golden Gate cloning was employed. The oligos containing the desired spacer sequences (described in *Table 4-3*) were annealed in a 50 µL reaction with a final concentration of 100 uM of each primer and 2.5 µL of NaCl 1M. The annealed spacers were cloned into the proper plasmid containing the Golden Gate cloning site as described in (Engler, Kandzia and Marillonnet, 2008).

Element	Primer Name	Sequence (5' -> 3')
dCas9 1st fragment	DM 52	ATCGGCCTGGCCATCGGCACC
	DM53	AGGCACGATA ^{GC} GTCCACATC
dCas9 2nd fragment	DM54	<u>ggtgccgatgccaggccgat</u> GCTGTACTTCTTGTC
	DM55	<u>gatctggacgctatcgtgcct</u> CAGAGCTTTC
dCas9 F2A	RG568	ATGGACTATAAGGACCACGACGG
	RG286	<u>ctgcccttgctcaccata</u> GGGCCCCGGTTCTGA
mCherry	RG288	ATGGTGAGCAAGGGCGAGG
	DM185	<u>cataacaacccccccaaggcgccgct</u> CTCGAGTTACTTGTACAGCTCGTCCATGCC
vasa- 2 Promoter	DM184	<u>ctccacctcaccatgggacccacgcgcata</u> gtATGTAGAACGCGAGCAAATTCTTTTCC
	DM62	<u>gtcgtggtccttagtccat</u> GTTTCCTTTCTTTATTACCGG
vasa- 2 terminator	DM186	CTCGAGAGCGCCGCCTTGGGGTGGG
	DM187	<u>ggcgagctcgaattaaccattgtg</u> ACCGGTGGCGCGCCAAAAAAGCACCGACTCGGTGCCACTTTTTTC
3xP3 GFP	RG314	CCACAATGGTTAATTCGAGCTC
	DM107	<u>gtactccacctcaccatgcgat</u> CGCTCCGGAAGATACATTGATGAG

Table 4-2 Primers employed for the mutagenesis of the Cas9 and the cloning of the plasmids with the dCas9 fused with the mCherry fluorescent protein.

The underlined sequences represent the Gibson overhangs. The letters in red represent the mutations introduced to deactivate the nuclease domains of Cas9.

gRNA Spacer	Primer Name	Sequence (5' -> 3')
B2-32	B2 -32 FW	<u>tgctg</u> AAAATCCGCTAAATATCAAA
	B2 -32 RV	<u>aaac</u> TTTGATATTTAGCGGATTTTC
B2-62	B2 -62 FW	<u>tgctg</u> AGCGTTGCTCTCTAAGCGCT
	B2 -62 RV	<u>aaac</u> AGCGCTTAGAGAGCAACGCTC
B2-110	B2 -110 FW	<u>aaac</u> AGTAGGAATCGCGTATTCGGC
	B2 -110 RV	<u>tgctg</u> CCGAATACGCGATTCTACT
B2-511	B2 -511 FW	<u>tgctg</u> AAAGCTTATCGATACGCGTA
	B2 -511 RV	<u>aaac</u> TACGCGTATCGATAAGCTTTC

Table 4-3 Oligonucleotides employed for cloning the gRNA spacers into the dCas9 plasmids.

The underlined sequences represent the overhangs for Golden Gate cloning.

4.11 Amplification of the Pcaf and fusion with the dCas9

The Pcaf histone acetyltransferase was identified starting from *Drosophila*'s GCN5 and tracking down orthologs in *A. gambiae*. The orthology between these two HATs was identified using Vectorbase (Giraldo-Calderón et al., 2015). The “core” region of the Pcaf and its domains were identified using Vectorbase and HMMER. Once the Pcaf peptides sequence was retrieved from Vectorbase, the sequence was uploaded on HMMER to analyse the domain structures of the protein. Both the software showed the same domain structures and located the three main domains of the protein in the same regions. The Pcaf ORF was then amplified from *A. gambiae* genomic DNA using primers targeting the full ORF of the Pcaf gene (AGAP004434) (*Table 4-4*). A small aliquot of the PCR (5 µL) was used to check the band size (2280 bp) on an agarose gel, and a second PCR was performed using 1 µL of the previous PCR product to add Gibson overhangs, a nuclear localisation signal (NLS), and a stop codon at the end of the sequence. Another PCR was performed using 1 µL of the first PCR product to amplify the "core" region of the HAT and provide Gibson overhangs simultaneously, producing an amplicon of 1162 bp.

To assemble the dCas9-Pcaf and dCas9-Pcaf^{CORE} plasmids, the D1 plasmid was used as a starting template. The plasmid was digested with NdeI and XhoI (ThermoFisher, #FD0585, #FD0695) and gel-extracted, yielding a 6765 bp band containing the plasmid backbone, the GFP marker, and the gRNA expression cassette. The *vasa*-2dCas9 sequence was re-amplified with Gibson overhangs, and a sequence containing a 3XGGGGS motif was added to the end of the dCas9 to provide a flexible linker between the dCas9 and the Pcaf (*Table 4-4*). This resulted in an amplicon of 6639 bp. Two 3-fragment Gibson ligations were performed, one for the Pcaf wild type and one for the Pcaf^{CORE}, with a molar ratio of 1:1:1. The Gibson assembly was followed by transformation, colony PCR, mini and maxiprep as explained above. The plasmid sequence was verified through diagnostic digestion and Sanger sequencing.

Element	Primer Name	Sequence (5' -> 3')
Pcaf WT	DM 203	<u>ggaggaggaggatcaggaggaggaggatca</u> ATGAGTGAGAATCCAAAAAATCAGAACGGCG
	DM156	<u>tcactttttcttttgcctggccggccttttcgtggccggccgctttt</u> TTTATCCCACAGACCGATTTCTTCATTTTGGTTTGAAAATACC
Pcaf CORE	DM205	<u>ggaggaggaggatcaggaggaggaggatca</u> TCGCTCACCAAACCGGTAACG
	DM210	<u>cgcaacacataacaacccccacccaaggcggccgct</u> CTCGAGTCACTTTTTCTTTTTTGCCTGGCCGGCC
dCas9 3XGGGGS	DM202	<u>tgatcctcctcctcctgatcctcctcctcc</u> GTCGCCTCCCAGCTGAGACAGG

Table 4-4 Oligos used to amplify the Pcaf histone acetyltransferase and its truncate version (CORE). The underlined sequences represent the Gibson overhangs.

4.12 DNA microinjections in mosquito embryos

The *A. gambiae* mosquitoes used in the study were maintained in a controlled environment with a relative humidity of 70-80% and a temperature of 26-28 °C. Freshly-laid eggs were collected and aligned for microinjections using an Olympus SZX7 optical microscope. Donor plasmids designed for random integration were co-injected into the embryos with a helper plasmid expressing the transposase under the control of the *vasa-2* promoter, each at a concentration of 200 ng/μL in PBS (ThermoFisher, #10010023).

To establish the D0 line, which was the first line created using the PiggyBac system, more than 3000 eggs were injected, resulting in an average hatching rate of around 10%. Roughly 48 hours after the injections, hatched L1 larvae were screened for the transient expression of the eGFP marker in their tails using a Nikon Eclipse TE200 fluorescent microscope. The larvae that tested positive for the screening (around 5%) were isolated, grown to adulthood, and mated with the wild type counterparts. The cages containing the transient adults were repeatedly blood-fed, and every progeny obtained was screened for transgenic larvae expressing the eGFP marker in the eyes and the spinal cord. The transgenic individuals recovered were placed in separated cages, mated with wild type individuals and their progenies were screened again for the presence of the fluorescent marker. To confirm the presence of only a single transgene insertion in the genome, I checked that the ratio between positive and negative individuals was always close to 50%. This ratio was significantly above this threshold in none of the lines generated in this study. Among these G1 individuals, I selected the top three who gave progenies with the stronger and more consistent expression of the eGFP marker. Finally, I selected one G1 male as the founder of the transgenic line and used its progeny to start the line. For each other line generated in this study with the piggyBac system, I injected from 1000 to 3000 eggs, with hatching rates ranging from 15 to 30% and a transient rate among the survivals between 1% and 5%. The average egg-to-transgenic rate with the PiggyBac system ranged between 0.5% and 1%.

To create the UAS-gypsy-B2-I-PpoI line, embryos from the previously generated Y-attP docking line were injected with a solution containing both the UAS donor plasmid and a *vasa-2* φ31 Integrase helper plasmid at a concentration of 200 ng/μL. After injecting more than 4000 eggs and generating 17 male transients individuals (the transient females were discarded because the transgene insertion was aimed at the Y chromosome), I retrieved one single transgenic larva expressing the 3xP3-eGFP fluorescent marker. The larva was grown to

adulthood and mated with eight wild type females. As previously described, successful integrations of the transgenes were identified among the G1 progeny as eGFP-positive larvae, expressing the marker in the eyes and spinal cord. Finally, I verified the insertion in the Y chromosome, checking that males only inherited the transgene.

The injections attempted to generate the Gal4 transgenic lines employed the same attP/attB integration system described above. I performed these injections in a previously generated autosomal attP docking line. Unfortunately, after many attempts and around 6000 eggs injected (with an average hatching rate of around 5%), I retrieved only six transients, and none of them gave any transgenics.

4.13 Molecular confirmation of transgene integration

The integration of all the transgenes cassettes into *A. gambiae* wild type and Y-3xP3-attP-RFP docking lines were molecularly confirmed by PCR using the Qiagen Fast Cycling PCR kit (Qiagen, #203743) and genomic DNA extracted with the Wizard Genomic DNA purification kit (Promega, #A9282) as described above. In the lines containing the dCas9, a 392 bp portion of the protein was amplified using the primers RG360 (5'-ccagccactatgagaagc-3') and RG338 (5'-ttgtcgctcccagctga-3') to confirm the integration of the construct. The line containing the Y-linked I-Ppo I flanked by *gypsy* insulators was verified amplifying the full 489 bp of the I-Ppo I coding sequence using the primers DM182 (5'-atggcgctcaccaatgctcaaac-3') and RG174 (5'-taccacaaagtgactgccccttgttgaggac-3'). The upstream *gypsy* sequence using the primers RG134 (5'-tttatactccggcggtcgagg-3') and DM19 (5'-gctagagtctccgctcgac-3') and finally the downstream *gypsy* sequence using the primers DM177 (5'-cccgaacaaattgcgattcccagg-3') and DM176 (5'-gggacccacgcgcatagctag-3').

4.14 Mosquitoes' microdissections

The testes and spermathecae were dissected using PBS buffer, fine-tipped forceps, and a scalpel. The dissected organs were carefully isolated and placed under a glass coverslip on a microscopic slide to ensure they were levelled and ready for visualisation using an EVOS XL Core Cell imaging system (ThermoFisher).

For qRT-PCR analysis, the abdomens were dissected by separating the fourth distal segment from the rest of the carcass. This ensured that the testes were included in the specimen. The

dissection was performed using fine-tipped forceps and a scalpel, and the dissected samples were collected in a microcentrifuge tube for subsequent RNA extraction and analysis.

4.15 Phenotype assays

Progenies from intercrossed GFP-bearing mosquitoes (dCas9 and dCas9-Pcaf lines) and RFP-bearing mosquitoes (Y-linked B2 I-Ppo I and QFS1 lines) were screened for the presence of both markers and, when necessary, separated according to the sexes. Fluorescent-positive pupae males and females from each line were selected using a Nikon Eclipse TE200 fluorescent microscope and placed in a cage with a 5% glucose solution. The mosquitoes were left to mate for five days before they were blood-fed with cow blood using a Hemotek membrane feeding system. Three days after the blood meal, single females were placed in single cups for deposition. At least 30 blood-fed females from each cross were allowed to lay individually into cups filled for 1/4 with water and fitted with filter paper (Whatman) on the sides to ensure the correct hydration of the eggs. Each cup was covered with mesh with a small cotton swab on top, imbued with 5% glucose solution. Eggs and larvae were counted and screened for each female. All group statistical differences were calculated using GraphPad and reported using the Kruskal-Wallis nonparametric tests.

4.16 Amplicon sequencing of females from gypsy Y-Ppo I line.

DNA from ten single eGFP-negative females from the *gypsy* Y-PpoI line inbred for three generations was extracted with the Wizard Genomic DNA Purification kit (Promega, #A9282) according to the manufacturer protocol (described above). About 50 ng of these newly extracted DNA samples were used as a template to amplify 517 bp of the 28S ribosomal subunit multicopy gene from each sample. The Phusion HiFi MasterMix (NEB) was employed for the amplifications (as described before) using the primers described in (Galizi et al., 2014) RG23 (5'-gtaagggaagtcggaactggatccg-3') and RG24 (5'- gaggaagccctcccacctatgctgc-3'). After the amplification, aliquots of 5 µL of the PCR products were loaded on a 1% agarose gel at 100V for 1h for visualisation. The remaining PCR products were purified using a Monarch PCR & DNA Cleanup Kit and then digested for 2h using the commercially available Ppo I restriction

enzyme (Promega) according to the manufacturer's protocol. The digestions were loaded on a 1% agarose gel and run for 1h at 100V. The bands were then visualised on a Gel Visualization System (Azure Biosystem).

The amplicon sequencing was performed using the previously extracted DNA as templates for PCR amplification. This time, it was amplified a smaller portion of the 28S ribosomal subunit spanning 260 bp. The PCR was carried out as described before using the following primers DM234 (5'-gtgatggccccgggtggg-3') and DM235 (5'-ctcaacagggtcttcttccccgc-3'). The PCR product was purified and sent for amplicon sequencing analysis. Genewiz performed the amplicon sequencing, while the data analysis was performed via CRISPResso2 (Clement et al., 2019)

5 Bibliography

- Abudayyeh, O.O. et al. (2017) 'RNA targeting with CRISPR-Cas13'. Available at: <https://doi.org/10.1038/nature24049>.
- Adams, K.L. et al. (2021) 'Wolbachia cifB induces cytoplasmic incompatibility in the malaria mosquito vector', *Nature Microbiology*, 6(12), pp. 1575–1582. Available at: <https://doi.org/10.1038/s41564-021-00998-6>.
- Adugna, T., Yewhelew, D. and Getu, E. (2021) 'Bloodmeal sources and feeding behavior of anopheline mosquitoes in Bure district, northwestern Ethiopia', *Parasites and Vectors*, 14(1), pp. 1–12. Available at: <https://doi.org/10.1186/s13071-021-04669-7>.
- Ahmad, K. and Henikoff, S. (2001) 'Modulation of a transcription factor counteracts heterochromatic gene silencing in *Drosophila*', *Cell*, 104(6), pp. 839–847. Available at: [https://doi.org/10.1016/S0092-8674\(01\)00281-1](https://doi.org/10.1016/S0092-8674(01)00281-1).
- Ali, T. et al. (2017) 'Chromatin binding of Gcn5 in *Drosophila* is largely mediated by CP190', *Nucleic Acids Research*, 45(5), pp. 2384–2395. Available at: <https://doi.org/10.1093/nar/gkw1178>.
- Allshire, R.C. and Madhani, H.D. (2018) 'Ten principles of heterochromatin formation and function', *Nature Reviews Molecular Cell Biology*, pp. 229–244. Available at: <https://doi.org/10.1038/nrm.2017.119>.
- Alout, H. et al. (2014) 'Evaluation of ivermectin mass drug administration for malaria transmission control across different West African environments', *Malaria Journal*, 13(1). Available at: <https://doi.org/10.1186/1475-2875-13-417>.
- Alphey, L. et al. (2010) 'Sterile-Insect Methods for Control of Mosquito-Borne Diseases: An Analysis', *VECTOR-BORNE AND ZOONOTIC DISEASES*, 10(3). Available at: <https://doi.org/10.1089=vbz.2009.0014>.
- Alphey, L. et al. (2013) 'Genetic control of *Aedes* mosquitoes', *Pathogens and Global Health*, 107(4), pp. 170–179. Available at: <https://doi.org/10.1179/2047773213Y.0000000095>.
- Arrow, K.J., Panosian, C. and Gelband, H. (2004) *Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance*, Institute of Medicine (US) Committee on the Economics of Antimalarial Drugs. Available at: <https://doi.org/10.1093/ISBN-10: 0-309-09218-3>.
- Bannister, A.J. and Kouzarides, T. (2011) 'Regulation of chromatin by histone modifications', *Cell Research*, 21(3), pp. 381–395. Available at: <https://doi.org/10.1038/cr.2011.22>.
- Barckmann, B. et al. (2013) 'Three levels of regulation lead to protamine and Mst77F expression in *Drosophila*', *Developmental Biology*, 377(1), pp. 33–45. Available at: <https://doi.org/10.1016/j.ydbio.2013.02.018>.

Barkal, A.A. et al. (2016) ‘Cas9 functionally opens chromatin’, *PLoS ONE*, 11(3). Available at: <https://doi.org/10.1371/journal.pone.0152683>.

Beaghton, P.J. and Burt, A. (2022) ‘Gene drives and population persistence vs elimination: The impact of spatial structure and inbreeding at low density’, *Theoretical Population Biology*, 145, pp. 109–125. Available at: <https://doi.org/10.1016/j.tpb.2022.02.002>.

Berger, S.L. (2007) ‘The complex language of chromatin regulation during transcription’, *Nature*. 2007/05/25, 447(7143), pp. 407–412. Available at: <https://doi.org/10.1038/nature05915>.

Bernardini, F. et al. (2014) ‘Site-specific genetic engineering of the *Anopheles gambiae* y chromosome’, *Proceedings of the National Academy of Sciences of the United States of America*, 111(21), pp. 7600–7605. Available at: <https://doi.org/10.1073/pnas.1404996111>.

Bi, X. and Broach, J.R. (1999) ‘UAS(rpg) can function as a heterochromatin boundary element in yeast’, *Genes & Development*, 13(9), pp. 1089–1101. Available at: <https://doi.org/DOI 10.1101/gad.13.9.1089>.

Bolotin, A. et al. (2005) ‘Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin’, *Microbiology*, 151(8), pp. 2551–2561. Available at: <https://doi.org/10.1099/mic.0.28048-0>.

Bowman, G.D. and Poirier, M.G. (2015) ‘Post-translational modifications of histones that influence nucleosome dynamics’, *Chemical Reviews*, pp. 2274–2295. Available at: <https://doi.org/10.1021/cr500350x>.

Boyle, E.A. et al. (2017) ‘High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding’, *Proceedings of the National Academy of Sciences of the United States of America*. 2017/05/13, 114(21), pp. 5461–5466. Available at: <https://doi.org/10.1073/pnas.1700557114>.

Brady, O.J. et al. (2016) ‘Vectorial capacity and vector control: Reconsidering sensitivity to parameters for malaria elimination’, *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 110(2), pp. 107–117. Available at: <https://doi.org/10.1093/trstmh/trv113>.

Braun, S.M.G. et al. (2017) ‘Rapid and reversible epigenome editing by endogenous chromatin regulators’, *Nature Communications*, 8(1). Available at: <https://doi.org/10.1038/s41467-017-00644-y>.

Brocken, D.J.W., Tark-Dame, M. and Dame, R.T. (2017) ‘dCas9: A Versatile Tool for Epigenome Editing’, in *The CRISPR/Cas System: Emerging Technology and Application*. Available at: <https://doi.org/10.21775/9781910190630.02>.

Burt, A. (2003) ‘Site-specific selfish genes as tools for the control and genetic engineering of natural populations’, *Proceedings of the Royal Society B: Biological Sciences*, 270(1518), pp. 921–928. Available at: <https://doi.org/10.1098/rspb.2002.2319>.

Burt, A. and Deredec, A. (2018) ‘Self-limiting population genetic control with sex-linked genome editors’, *Proceedings of the Royal Society B: Biological Sciences*, 285(1883). Available at: <https://doi.org/10.1098/rspb.2018.0776>.

- Byrd, K. and Corces, V.G. (2003) ‘Visualization of chromatin domains created by the gypsy insulator of *Drosophila*’, *J Cell Biol.* 2003/08/20, 162(4), pp. 565–574. Available at: <https://doi.org/10.1083/jcb.200305013>.
- Carballar-Lejarazú, R., Brennock, P. and James, A.A. (2016) ‘Suppressor of hairy-wing, modifier of mdg4 and centrosomal protein of 190 gene orthologues of the gypsy insulator complex in the malaria mosquito, *Anopheles stephensi*’, *Insect Molecular Biology*, 25(4), pp. 460–469. Available at: <https://doi.org/10.1111/imb.12233>.
- Carballar-Lejarazú, R., Jasinskiene, N. and James, A.A. (2013) ‘Exogenous gypsy insulator sequences modulate transgene expression in the malaria vector mosquito, *Anopheles stephensi*’, *Proceedings of the National Academy of Sciences of the United States of America*, 110(18), pp. 7176–7181. Available at: <https://doi.org/10.1073/pnas.1304722110>.
- Catteruccia, F., Benton, J.P. and Crisanti, A. (2005) ‘An *Anopheles* transgenic sexing strain for vector control’, *Nature Biotechnology*, 23(11), pp. 1414–1417. Available at: <https://doi.org/10.1038/NBT1152>.
- Chavez, A. et al. (2016) ‘Comparison of Cas9 activators in multiple species’, *Nat Methods*. 2016/05/24, 13(7), pp. 563–567. Available at: <https://doi.org/10.1038/nmeth.3871>.
- Chen, S. and Corces, V.G. (2001) ‘The gypsy insulator of *drosophila* affects chromatin structure in a directional manner’, *Genetics*, 159(4), pp. 1649–1658. Available at: <https://doi.org/10.1093/genetics/159.4.1649>.
- Chen, X., Zaro, J.L. and Shen, W.C. (2013) ‘Fusion protein linkers: Property, design and functionality’, *Advanced Drug Delivery Reviews*, 65(10), pp. 1357–1369. Available at: <https://doi.org/10.1016/j.addr.2012.09.039>.
- Chiarella, A.M. et al. (2019) ‘Dose-dependent activation of gene expression is achieved using CRISPR and small molecules that recruit endogenous chromatin machinery’, *Nat Biotechnol.* 2019/11/13, 38(1), pp. 50–55. Available at: <https://doi.org/10.1038/s41587-019-0296-7>.
- Chiarella, A.M., Lu, D. and Hathaway, N.A. (2020) ‘Epigenetic control of a local chromatin landscape’, *International Journal of Molecular Sciences*, 21(3), pp. 6–8. Available at: <https://doi.org/10.3390/ijms21030943>.
- Choudhury, S.R. et al. (2016) ‘CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter’, *Oncotarget*, 7(29), pp. 46545–46556. Available at: <https://doi.org/10.18632/oncotarget.10234>.
- Clement, K. et al. (2019) ‘CRISPResso2 provides accurate and rapid genome editing sequence analysis’, *Nature Biotechnology*, 37(3), pp. 224–226. Available at: <https://doi.org/10.1038/s41587-019-0032-3>.
- Consortium, T.P.W.– S. and Ching, N.L. (2021) ‘Wolbachia-mediated sterility suppresses *Aedes aegypti* populations in the urban tropics’, *medRxiv*, p. 2021.06.16.21257922. Available at: <https://doi.org/https://doi.org/10.1101/2021.06.16.21257922>;
- Dattoo, M.S. et al. (2021) ‘High Efficacy of a Low Dose Candidate Malaria Vaccine, R21 in 1 Adjuvant Matrix-M™, with Seasonal Administration to Children in Burkina Faso’, *SSRN Electronic Journal* [Preprint]. Available at: <https://doi.org/10.2139/ssrn.3830681>.

- Dean A Baker, Tony Nolan, Bettina Fischer, Alex Pinder, A.C. and S.R. (2005) ‘A comprehensive gene expression atlas of sex- and tissue-specificity in the malaria vector, *Anopheles gambiae*’, *Current Biology*, 15(6). Available at: <https://doi.org/10.1016/j.cub.2005.03.005>.
- Deng, W. et al. (2015) ‘CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells’, *Proc Natl Acad Sci U S A*. 2015/09/02, 112(38), pp. 11870–11875. Available at: <https://doi.org/10.1073/pnas.1515692112>.
- Dhillon, N. et al. (2009) ‘DNA polymerase epsilon, acetylases and remodellers cooperate to form a specialized chromatin structure at a tRNA insulator’, *EMBO J*. 2009/07/25, 28(17), pp. 2583–2600. Available at: <https://doi.org/10.1038/emboj.2009.198>.
- Dixon, J.R. et al. (2012) ‘Topological domains in mammalian genomes identified by analysis of chromatin interactions’, *Nature*, 485(7398), pp. 376–380. Available at: <https://doi.org/10.1038/nature11082>.
- Dugar, G. et al. (2018) ‘CRISPR RNA-Dependent Binding and Cleavage of Endogenous RNAs by the *Campylobacter jejuni* Cas9’, *Molecular Cell*, 69(5), pp. 893–905.e7. Available at: <https://doi.org/10.1016/j.molcel.2018.01.032>.
- Dunham, I. et al. (2012) ‘An integrated encyclopedia of DNA elements in the human genome’, *Nature*, 489(7414), pp. 57–74. Available at: <https://doi.org/10.1038/nature11247>.
- Edgell, D.R. (2009) ‘Selfish DNA: Homing Endonucleases Find a Home’, *Current Biology*, 19(3), pp. 115–117. Available at: <https://doi.org/10.1016/j.cub.2008.12.019>.
- Engler, C., Kandzia, R. and Marillonnet, S. (2008) ‘A one pot, one step, precision cloning method with high throughput capability’, *PLoS ONE*, 3(11). Available at: <https://doi.org/10.1371/journal.pone.0003647>.
- Ewen-Campen, B. et al. (2017) ‘Optimized strategy for in vivo Cas9-activation in *Drosophila*’, *Proceedings of the National Academy of Sciences of the United States of America*, 114(35), pp. 9409–9414. Available at: <https://doi.org/10.1073/pnas.1707635114>.
- Fu, G. et al. (2010) ‘Female-specific flightless phenotype for mosquito control’, *Proceedings of the National Academy of Sciences of the United States of America*, 107(10), pp. 4550–4554. Available at: <https://doi.org/10.1073/pnas.1000251107>.
- Galizi, R. et al. (2014) ‘A synthetic sex ratio distortion system for the control of the human malaria mosquito’, *Nature Communications*, 5, pp. 1–8. Available at: <https://doi.org/10.1038/ncomms4977>.
- Galizi, R. et al. (2016) ‘A CRISPR-Cas9 sex-ratio distortion system for genetic control’, *Scientific Reports*, 6(April), pp. 2–6. Available at: <https://doi.org/10.1038/srep31139>.
- Galizi, R. and Jaramillo, A. (2019) ‘Engineering CRISPR guide RNA riboswitches for in vivo applications’, *Current Opinion in Biotechnology*, 55, pp. 103–113. Available at: <https://doi.org/10.1016/j.copbio.2018.08.007>.

- Gaszner, M. and Felsenfeld, G. (2006) 'Insulators: Exploiting transcriptional and epigenetic mechanisms', *Nature Reviews Genetics*, 7(9), pp. 703–713. Available at: <https://doi.org/10.1038/nrg1925>.
- Geci, R., Willis, K. and Burt, A. (2022) 'Gene drive designs for efficient and localisable population suppression using Y-linked editors', *PLoS genetics*, 18(12), p. e1010550. Available at: <https://doi.org/10.1371/journal.pgen.1010550>.
- Gerasimova, T.I., Byrd, K. and Corces, V.G. (2000) 'A chromatin insulator determines the nuclear localization of DNA', *Molecular Cell*, 6(5), pp. 1025–1035. Available at: [https://doi.org/10.1016/S1097-2765\(00\)00101-5](https://doi.org/10.1016/S1097-2765(00)00101-5).
- Gesto, J.S.M. et al. (2021) 'Reduced competence to arboviruses following the sustainable invasion of Wolbachia into native *Aedes aegypti* from Southeastern Brazil', *Scientific Reports*, 11(1), p. 10039. Available at: <https://doi.org/10.1038/s41598-021-89409-8>.
- Ghirlando, R. et al. (2012) 'Chromatin domains, insulators, and the regulation of gene expression', *Biochim Biophys Acta*. 2012/02/14, 1819(7), pp. 644–651. Available at: <https://doi.org/10.1016/j.bbagr.2012.01.016>.
- Gilbert, L.A. et al. (2013) 'CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes', *Cell*. 2013/07/16, 154(2), pp. 442–451. Available at: <https://doi.org/10.1016/j.cell.2013.06.044>.
- Giraldo-Calderón, G.I. et al. (2015) 'VectorBase: An updated Bioinformatics Resource for invertebrate vectors and other organisms related with human diseases', *Nucleic Acids Research*, 43(D1), pp. D707–D713. Available at: <https://doi.org/10.1093/nar/gku1117>.
- Glastad, K.M., Hunt, B.G. and Goodisman, M.A.D.D. (2018) 'Epigenetics in Insects: Genome Regulation and the Generation of Phenotypic Diversity', *Annu Rev Entomol*. 2018/10/05, 64(1), pp. 185–203. Available at: <https://doi.org/10.1146/annurev-ento-011118-111914>.
- Gomes, F.M. and Barillas-Mury, C. (2018) 'Infection of anopheline mosquitoes with Wolbachia: Implications for malaria control', *PLoS Pathogens*, 14(11). Available at: <https://doi.org/10.1371/journal.ppat.1007333>.
- Grosse, S.D. et al. (2011) 'Sickle Cell Disease in Africa A Neglected Cause of Early Childhood Mortality'. Available at: <https://doi.org/10.1016/j.amepre.2011.09.013>.
- Guo, T. et al. (2014) 'Functional organization and dynamics of the cell nucleus'. Available at: <https://doi.org/10.3389/fpls.2014.00378>.
- Hall, A.B. et al. (2016) 'Radical remodeling of the Y chromosome in a recent radiation of malaria mosquitoes', *Proceedings of the National Academy of Sciences of the United States of America*, 113(15), pp. E2114–E2123. Available at: <https://doi.org/10.1073/pnas.1525164113>.
- Hammond, Andrew, Galizi, R., et al. (2016) 'A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*', *Nature Biotechnology*. 2015/12/08, 34(1), pp. 78–83. Available at: <https://doi.org/10.1038/nbt.3439>.

Hammond, A et al. (2016) 'A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*', *Nat Biotechnol.* 2015/12/08, 34(1), pp. 78–83. Available at: <https://doi.org/10.1038/nbt.3439>.

Hammond, Andrew, Kyrou, K., et al. (2016) 'Gene drives and malaria control Driving endonuclease genes Previous limitations of CRISPR-based gene drives', *Nature Biotechnology*, 34(1), pp. 78–83. Available at: https://www.nature.com/articles/nbt.3439?__hstc=13887208.a07978b93986f5153ad75ef9157da47e.1471615194540.1473447065996.1473690478309.12&__hssc=13887208.3.1473690478309&__hsfp=1816377754.

Hammond, A.M. et al. (2017) 'The creation and selection of mutations resistant to a gene drive over multiple generations in the malaria mosquito', *PLoS Genetics*, 13(10), pp. 1–16. Available at: <https://doi.org/10.1371/journal.pgen.1007039>.

Hansen, I.A. et al. (2014) 'Four-way regulation of mosquito yolk protein precursor genes by juvenile hormone-, ecdysone-, nutrient-, and insulin-like peptide signaling pathways', *Frontiers in Physiology*, 5 MAR(March), pp. 1–8. Available at: <https://doi.org/10.3389/fphys.2014.00103>.

Helena Mangs, A. and Morris, B. (2007) 'The Human Pseudoautosomal Region (PAR): Origin, Function and Future', *Current Genomics*, 8(2), pp. 129–136. Available at: <https://doi.org/10.2174/138920207780368141>.

Hemingway, J. et al. (2004) 'The molecular basis of insecticide resistance in mosquitoes', *Insect Biochemistry and Molecular Biology*, 34(7), pp. 653–665. Available at: <https://doi.org/10.1016/j.ibmb.2004.03.018>.

HILDRETH, P.E. (1965) 'Doublesex, Recessive Gene That Transforms Both Males and Females of *Drosophila* Into Intersexes.', *Genetics*, 51(April), pp. 659–678. Available at: <https://doi.org/10.1093/genetics/51.4.659>.

Hilton, I.B. et al. (2015) 'Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers', *Nat Biotechnol.* 2015/04/08, 33(5), pp. 510–517. Available at: <https://doi.org/10.1038/nbt.3199>.

Hong WEI, G., Pei LIU, D. and Chuan LIANG, C. (2005) 'Chromatin domain boundaries: insulators and beyond', *Cell Research*, 15(4), pp. 292–300.

Horlbeck, M.A. et al. (2016) 'Nucleosomes impede cas9 access to DNA in vivo and in vitro', *eLife*, 5(MARCH2016), pp. 1–21. Available at: <https://doi.org/10.7554/eLife.12677>.

Isaac, R.S. et al. (2016) 'Nucleosome breathing and remodeling constrain CRISPR-Cas9 function', *eLife*, 5(APRIL2016), pp. 1–14. Available at: <https://doi.org/10.7554/eLife.13450>.

Ishino, Y. et al. (1987) 'Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isoenzyme conversion in *Escherichia coli*, and identification of the gene product', *Journal of Bacteriology*, 169(12), pp. 5429–5433. Available at: <https://doi.org/10.1128/jb.169.12.5429-5433.1987>.

- Jansen, R. et al. (2002) 'Identification of genes that are associated with DNA repeats in prokaryotes', *Molecular Microbiology*, 43(6), pp. 1565–1575. Available at: <https://doi.org/10.1046/j.1365-2958.2002.02839.x>.
- Janssens, D.H. et al. (2018) 'Automated in situ chromatin profiling efficiently resolves cell types and gene regulatory programs', *Epigenetics and Chromatin*, 11(1), pp. 1–14. Available at: <https://doi.org/10.1186/s13072-018-0243-8>.
- Jinek, M. et al. (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science*, 337(6096), pp. 816–821. Available at: <https://doi.org/10.1126/science.1225829>.
- Kahney, E.W., Snedeker, J.C. and Chen, X. (2019) 'Regulation of Drosophila germline stem cells', *Current Opinion in Cell Biology*, 60, pp. 27–35. Available at: <https://doi.org/10.1016/j.ceb.2019.03.008>.
- Kandul, N.P. et al. (2019) 'Transforming insect population control with precision guided sterile males with demonstration in flies', *Nature Communications*, 10(1). Available at: <https://doi.org/10.1038/s41467-018-07964-7>.
- Kauppi, L. et al. (2011) 'Distinct properties of the XY pseudoautosomal region crucial for male meiosis', *Science*. 2011/02/19, 331(6019), pp. 916–920. Available at: <https://doi.org/10.1126/science.1195774>.
- Kaya-Okur, H.S. et al. (2019) 'CUT&Tag for efficient epigenomic profiling of small samples and single cells', *Nature Communications*, 10(1), pp. 1–10. Available at: <https://doi.org/10.1038/s41467-019-09982-5>.
- Kelly, S.A., Panhuis, T.M. and Stoehr, A.M. (2012) 'Phenotypic plasticity: Molecular mechanisms and adaptive significance', *Comprehensive Physiology*, 2(2), pp. 1417–1439. Available at: <https://doi.org/10.1002/cphy.c110008>.
- Klein, S.L. et al. (2016) 'RTS,S malaria vaccine and increased mortality in girls', *mBio*, 7(2), pp. 1–2. Available at: <https://doi.org/10.1128/mBio.00514-16>.
- Klowden, M.J. (2013) 'Reproductive Systems', *Physiological Systems in Insects*, pp. 197–254. Available at: <https://doi.org/10.1016/B978-0-12-415819-1.00004-0>.
- Koutelou, E., Farria, A.T. and Dent, S.Y.R. (2021) 'Complex functions of Gcn5 and Pcaf in development and disease', *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1864(2), p. 194609. Available at: <https://doi.org/10.1016/j.bbagr.2020.194609>.
- Kouzarides, T. (2007) 'Chromatin Modifications and Their Function', *Cell*, 128(4), pp. 693–705. Available at: <https://doi.org/10.1016/j.cell.2007.02.005>.
- Krzywinska, E. et al. (2016) 'A maleness gene in the malaria mosquito *Anopheles gambiae*', *Science*, 353(6294), pp. 67–69. Available at: <https://doi.org/10.1126/science.aaf5605>.
- Krzywinska, E., Ferretti, L. and Krzywinski, J. (2022) 'Establishment and a comparative transcriptomic analysis of a male-specific cell line from the African malaria mosquito *Anopheles gambiae*', *Scientific Reports*, 12(1). Available at: <https://doi.org/10.1038/s41598-022-10686-y>.

- Kuhl, L.M. et al. (2020) ‘A dCas9-based system identifies a central role for Ctf19 in kinetochore-derived suppression of meiotic recombination’, *Genetics*, 216(2), pp. 395–408. Available at: <https://doi.org/10.1534/GENETICS.120.303384>.
- Kyrou, K. et al. (2018) ‘A CRISPR–Cas9 gene drive targeting doublesex causes complete population suppression in caged *Anopheles gambiae* mosquitoes’, *Nature Biotechnology*, 36(11), pp. 1062–1071. Available at: <https://doi.org/10.1038/nbt.4245>.
- Labbé, G.M.C. et al. (2012) ‘Female-specific flightless (fsRIDL) phenotype for control of *Aedes albopictus*’, *PLoS Neglected Tropical Diseases*, 6(7). Available at: <https://doi.org/10.1371/journal.pntd.0001724>.
- Landeem, E.L. et al. (2016) ‘Sex Chromosome-wide Transcriptional Suppression and Compensatory Cis-Regulatory Evolution Mediate Gene Expression in the *Drosophila* Male Germline’, *PLoS Biology*, 14(7), pp. 1–29. Available at: <https://doi.org/10.1371/journal.pbio.1002499>.
- Larson, M.H. et al. (2013) ‘CRISPR interference (CRISPRi) for sequence-specific control of gene expression’, *Nature Protocols*, 8(11), pp. 2180–2196. Available at: <https://doi.org/10.1038/nprot.2013.132>.
- Li, M. et al. (2017) ‘Germline Cas9 expression yields highly efficient genome engineering in a major worldwide disease vector, *Aedes aegypti*’, *Proceedings of the National Academy of Sciences of the United States of America*, 114(49), pp. E10540–E10549. Available at: <https://doi.org/10.1073/pnas.1711538114>.
- Li, M. et al. (2020) ‘Development of a confinable gene drive system in the human disease vector *aedes aegypti*’, *eLife*, 9, pp. 1–22. Available at: <https://doi.org/10.7554/eLife.51701>.
- Liang, F. et al. (1998) ‘Homology-directed repair is a major double-strand break repair pathway in mammalian cells’, *Proceedings of the National Academy of Sciences of the United States of America*, 95(9), pp. 5172–5177. Available at: <https://doi.org/10.1073/pnas.95.9.5172>.
- Lin, J.L. et al. (2019) ‘CRISPR-PIN: Modifying gene position in the nucleus via dCas9-mediated tethering’, *Synthetic and Systems Biotechnology*, 4(2), pp. 73–78. Available at: <https://doi.org/10.1016/j.synbio.2019.02.001>.
- Lin, S. et al. (2015) ‘In vivo transcriptional activation using CRISPR/Cas9 in *Drosophila*’, *Genetics*, 201(2), pp. 433–442. Available at: <https://doi.org/10.1534/genetics.115.181065>.
- Liu, K. et al. (2011) ‘Aquaporin water channel AgAQP1 in the malaria vector mosquito *Anopheles gambiae* during blood feeding and humidity adaptation’, *Proceedings of the National Academy of Sciences of the United States of America*, 108(15), pp. 6062–6066. Available at: <https://doi.org/10.1073/pnas.1102629108>.
- Liu, Z. et al. (2017) ‘Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector’, *Scientific Reports*, 7(1), pp. 1–9. Available at: <https://doi.org/10.1038/s41598-017-02460-2>.

- Luger, K., Dechassa, M.L. and Tremethick, D.J. (2012) ‘New insights into nucleosome and chromatin structure: An ordered state or a disordered affair?’, *Nature Reviews Molecular Cell Biology*, 13(7), pp. 436–447. Available at: <https://doi.org/10.1038/nrm3382>.
- Lynd, A. and Lycett, G.J. (2012) ‘Development of the bi-partite Gal4-UAS system in the African malaria mosquito, *Anopheles gambiae*’, *PLoS One*. 2012/02/22, 7(2), p. e31552. Available at: <https://doi.org/10.1371/journal.pone.0031552>.
- Maeder, M.L. et al. (2013) ‘CRISPR RNA-guided activation of endogenous human genes’, *Nature Methods*, 10(10), pp. 977–979. Available at: <https://doi.org/10.1038/nmeth.2598>.
- Mahmoudi, S. and Keshavarz, H. (2017) ‘Human Vaccines & Immunotherapeutics Efficacy of phase 3 trial of RTS, S/AS01 malaria vaccine: The need for an alternative development plan’. Available at: <https://doi.org/10.1080/21645515.2017.1295906>.
- Majori, G. (2012) ‘Short history of malaria and its eradication in Italy with short notes on the fight against the infection in the Mediterranean basin’, *Mediterranean Journal of Hematology and Infectious Diseases*, p. 201. Available at: <http://www.mjhid.org/article/view/9990>.
- Manning, L.R. and Manning, J.M. (2018) ‘Contributions to nucleosome dynamics in chromatin from interactive propagation of phosphorylation/acetylation and inducible histone lysine basicities’, *Protein Science*, 27(3), pp. 662–671. Available at: <https://doi.org/10.1002/pro.3359>.
- Marino, N.D. et al. (2020) ‘Anti-CRISPR protein applications: natural brakes for CRISPR-Cas technologies’, *Nature Methods*, 17(5), pp. 471–479. Available at: <https://doi.org/10.1038/s41592-020-0771-6>.
- Marinotti, O. et al. (2013) ‘The Genome of *Anopheles darlingi*, the main neotropical malaria vector’, *Nucleic Acids Research*, 41(15), pp. 7387–7400. Available at: <https://doi.org/10.1093/nar/gkt484>.
- Marraffini, L.A. (2015) ‘CRISPR-Cas immunity in prokaryotes’, *Nature*, 526(7571), pp. 55–61. Available at: <https://doi.org/10.1038/nature15386>.
- Melnikova, L. et al. (2019) ‘The same domain of Su(Hw) is required for enhancer blocking and direct promoter repression’, *Scientific Reports*, 9(1), pp. 1–12. Available at: <https://doi.org/10.1038/s41598-019-41761-6>.
- Messerschmidt, D.M., Knowles, B.B. and Solter, D. (2014) ‘DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos’, *Genes and Development*, 28(8), pp. 812–828. Available at: <https://doi.org/10.1101/gad.234294.113>.
- Michiels, F. et al. (1989) ‘A 14 bp promoter element directs the testis specificity of the *Drosophila* $\beta 2$ tubulin gene’, *EMBO Journal*, 8(5), pp. 1559–1565. Available at: <https://doi.org/10.1002/j.1460-2075.1989.tb03540.x>.
- Mojica, F.J.M. et al. (2000) ‘Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria’, *Molecular Microbiology*, 36(1), pp. 244–246. Available at: <https://doi.org/10.1046/j.1365-2958.2000.01838.x>.

- Mojica, F.J.M. et al. (2005) 'Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements', *Journal of Molecular Evolution*, 60(2), pp. 174–182. Available at: <https://doi.org/10.1007/s00239-004-0046-3>.
- Mojica, F.J.M. et al. (2009) 'Short motif sequences determine the targets of the prokaryotic CRISPR defence system', *Microbiology*, 155(3), pp. 733–740. Available at: <https://doi.org/10.1099/mic.0.023960-0>.
- Mojica, F.J.M. and Montoliu, L. (2016) 'On the Origin of CRISPR-Cas Technology: From Prokaryotes to Mammals', *Trends in Microbiology*, 24(10), pp. 811–820. Available at: <https://doi.org/10.1016/j.tim.2016.06.005>.
- Mojica, F.J.M. and Rodriguez-Valera, F. (2016) 'The discovery of CRISPR in archaea and bacteria', *FEBS Journal*, 283, pp. 3162–3169. Available at: <https://doi.org/10.1111/febs.13766>.
- Moore, L.D., Le, T. and Fan, G. (2013) 'DNA methylation and its basic function', *Neuropsychopharmacology*, 38(1), pp. 23–38. Available at: <https://doi.org/10.1038/npp.2012.112>.
- Müller, G.C. et al. (2010) *Field experiments of Anopheles gambiae attraction to local fruits/seedpods and flowering plants in Mali to optimize strategies for malaria vector control in Africa using attractive toxic sugar bait methods*. Available at: <http://www.malariajournal.com/content/9/1/262>.
- Naeem, M. et al. (2020) 'Latest Developed Strategies to Minimize the Off-Target Effects in CRISPR-Cas-Mediated Genome Editing', *Cells*, 9(7), pp. 1–23. Available at: <https://doi.org/10.3390/cells9071608>.
- Namekawa, S.H. et al. (2006) 'Postmeiotic Sex Chromatin in the Male Germline of Mice', *Current Biology*, 16(7), pp. 660–667. Available at: <https://doi.org/10.1016/j.cub.2006.01.066>.
- O'Geen, H. et al. (2017) 'DCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression', *Nucleic Acids Research*, 45(17), pp. 9901–9916. Available at: <https://doi.org/10.1093/nar/gkx578>.
- Oki, M. et al. (2004) 'Barrier Proteins Remodel and Modify Chromatin To Restrict Silenced Domains', *Molecular and Cellular Biology*, 24(5), pp. 1956–1967. Available at: <https://doi.org/10.1128/mcb.24.5.1956-1967.2004>.
- Otto, S.P. et al. (2011) 'About PAR: the distinct evolutionary dynamics of the pseudoautosomal region', *Trends Genet.* 2011/10/04, 27(9), pp. 358–367. Available at: <https://doi.org/10.1016/j.tig.2011.05.001>.
- Papathanos, P.A. et al. (2009) 'The vasa regulatory region mediates germline expression and maternal transmission of proteins in the malaria mosquito *Anopheles gambiae*: A versatile tool for genetic control strategies', *BMC Molecular Biology*. 2009/07/04, 10, pp. 1–13. Available at: <https://doi.org/10.1186/1471-2199-10-65>.
- Pereira, T.N. et al. (2018) 'Wolbachia significantly impacts the vector competence of *Aedes aegypti* for Mayaro virus', *Scientific Reports*, 8(1). Available at: <https://doi.org/10.1038/s41598-018-25236-8>.

- Piel, F.B. et al. (2010) 'Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis'. Available at: <https://doi.org/10.1038/ncomms1104>.
- Pikaart, M.I., Recillas-Targa, F. and Felsenfeld, G. (1998) 'Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators', *Genes & Development*, 12(18), pp. 2852–2862. Available at: <https://doi.org/DOI.10.1101/gad.12.18.2852>.
- Pourcel, C., Salvignol, G. and Vergnaud, G. (2005) 'CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies', *Microbiology*, 151(3), pp. 653–663. Available at: <https://doi.org/10.1099/mic.0.27437-0>.
- Poux, A.N. and Marmorstein, R. (2003) 'Molecular Basis for Gcn5/PCAF Histone Acetyltransferase Selectivity for Histone and Nonhistone Substrates', *Biochemistry*, 42(49), pp. 14366–14374. Available at: <https://doi.org/10.1021/bi035632n>.
- Pranckeviciene, E. et al. (2020) 'Nucleosome positioning sequence patterns as packing or regulatory', *PLoS Computational Biology*, 16(1), pp. 1–17. Available at: <https://doi.org/10.1371/journal.pcbi.1007365>.
- Provataris, P. et al. (2018) 'Signatures of DNA Methylation across Insects Suggest Reduced DNA Methylation Levels in Holometabola', *Genome Biol Evol.* 2018/04/27, 10(4), pp. 1185–1197. Available at: <https://doi.org/10.1093/gbe/evy066>.
- Pulecio, J. et al. (2017) 'CRISPR/Cas9-Based Engineering of the Epigenome', *Cell Stem Cell*. 2017/10/07, 21(4), pp. 431–447. Available at: <https://doi.org/10.1016/j.stem.2017.09.006>.
- Rhie, A. et al. (2021) 'Towards complete and error-free genome assemblies of all vertebrate species', *Nature*, 592(7856), pp. 737–746. Available at: <https://doi.org/10.1038/s41586-021-03451-0>.
- Roca Paixão, J.F. et al. (2019) 'Improved drought stress tolerance in Arabidopsis by CRISPR/dCas9 fusion with a Histone AcetylTransferase', *Scientific Reports*, 9(1). Available at: <https://doi.org/10.1038/s41598-019-44571-y>.
- Santel, A. et al. (2000) *The initiator element of the Drosophila $\beta 2$ tubulin gene core promoter contributes to gene expression in vivo but is not required for male germ-cell specific expression*, *Nucleic Acids Research*.
- Schoborg, T. et al. (2013) 'Chromatin insulator bodies are nuclear structures that form in response to osmotic stress and cell death', *Journal of Cell Biology*, 202(2), pp. 261–276. Available at: <https://doi.org/10.1083/jcb.201304181>.
- Smith, R.C. et al. (2007) 'Testis-specific expression of the $\beta 2$ tubulin promoter of *Aedes aegypti* and its application as a genetic sex-separation marker', *Insect Molecular Biology*, 16(1), pp. 61–71. Available at: <https://doi.org/10.1111/j.1365-2583.2006.00701.x>.
- Sougoufara, S., Ottih, E.C. and Tripet, F. (2020) 'The need for new vector control approaches targeting outdoor biting Anopheline malaria vector communities', *Parasites and Vectors*. BioMed Central. Available at: <https://doi.org/10.1186/s13071-020-04170-7>.

Strutt, S.C. et al. (2018) 'RNA-dependent RNA targeting by CRISPR-Cas9', *eLife*, 7, pp. 1–17. Available at: <https://doi.org/10.7554/eLife.32724>.

Tan, M. et al. (2011) 'Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification', *Cell*, 2011/09/20, 146(6), pp. 1016–1028. Available at: <https://doi.org/10.1016/j.cell.2011.08.008>.

Tanenbaum, M.E. et al. (2014) 'A protein-tagging system for signal amplification in gene expression and fluorescence imaging', *Cell*, 2014/10/14, 159(3), pp. 635–646. Available at: <https://doi.org/10.1016/j.cell.2014.09.039>.

Tauxe, G.M. et al. (2013) 'XTargeting a dual detector of skin and CO₂ to modify mosquito host seeking', *Cell*, 155(6), pp. 1365–1379. Available at: <https://doi.org/10.1016/j.cell.2013.11.013>.

Taxiarchi, C. et al. (2019) 'High-resolution transcriptional profiling of *Anopheles gambiae* spermatogenesis reveals mechanisms of sex chromosome regulation', *Scientific Reports*, 9(1). Available at: <https://doi.org/10.1038/s41598-019-51181-1>.

Taxiarchi, C. et al. (2021a) 'A genetically encoded anti-CRISPR protein constrains gene drive spread and prevents population suppression', *Nature Communications*, 12(1). Available at: <https://doi.org/10.1038/s41467-021-24214-5>.

Taxiarchi, C. et al. (2021b) 'A genetically encoded anti-CRISPR protein constrains gene drive spread and prevents population suppression', *Nature Communications*, 12(1). Available at: <https://doi.org/10.1038/s41467-021-24214-5>.

Terns, M.P. (2018) 'CRISPR-Based Technologies: Impact of RNA-Targeting Systems', *Molecular Cell*, 72(3), pp. 404–412. Available at: <https://doi.org/10.1016/j.molcel.2018.09.018>.

Terradas, G. et al. (2021) 'Inherently confinable split-drive systems in *Drosophila*', *Nature Communications*, 12(1). Available at: <https://doi.org/10.1038/s41467-021-21771-7>.

Terradas, G. et al. (2022) 'High-resolution in situ analysis of Cas9 germline transcript distributions in gene-drive *Anopheles* mosquitoes', *G3: Genes, Genomes, Genetics*, 12(1). Available at: <https://doi.org/10.1093/G3JOURNAL/JKAB369>.

Truong, L.N. et al. (2013) 'Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells', *Proceedings of the National Academy of Sciences of the United States of America*, 110(19), pp. 7720–7725. Available at: <https://doi.org/10.1073/pnas.1213431110>.

Tubio, J.M.C. et al. (2011) 'Evolutionary dynamics of the Ty3/Gypsy LTR retrotransposons in the genome of *Anopheles gambiae*', *PLoS ONE*, 6(1). Available at: <https://doi.org/10.1371/journal.pone.0016328>.

Turner, J.M. et al. (2005) 'Silencing of unsynapsed meiotic chromosomes in the mouse', *Nat Genet.* 2004/12/08, 37(1), pp. 41–47. Available at: <https://doi.org/10.1038/ng1484>.

Turner, J.M.A. et al. (2005) 'Silencing of unsynapsed meiotic chromosomes in the mouse', *Nature Genetics*, 37(1), pp. 41–47. Available at: <https://doi.org/10.1038/ng1484>.

- Turner, J.M.A. et al. (2006) ‘Pachytene asynapsis drives meiotic sex chromosome inactivation and leads to substantial postmeiotic repression in spermatids’, *Developmental Cell*, 10(4), pp. 521–529. Available at: <https://doi.org/10.1016/j.devcel.2006.02.009>.
- Turner, J.M.A. (2007) ‘Meiotic sex chromosome inactivation’, *Development*. 2007/03/03, 134(10), pp. 1823–1831. Available at: <https://doi.org/10.1242/dev.000018>.
- Waddington, C.H. (2012) ‘The epigenotype. 1942.’, *International journal of epidemiology*, 41(1), pp. 10–13. Available at: <https://doi.org/10.1093/ije/dyr184>.
- Wang, G.H. et al. (2021) ‘Combating mosquito-borne diseases using genetic control technologies’, *Nature Communications*. Nature Research. Available at: <https://doi.org/10.1038/s41467-021-24654-z>.
- Wang, H. et al. (2018) ‘CRISPR-Mediated Programmable 3D Genome Positioning and Nuclear Organization’, *Cell*, 175(5), pp. 1405–1417.e14. Available at: <https://doi.org/10.1016/j.cell.2018.09.013>.
- Waters, A.J. et al. (2018) ‘Rationally-engineered reproductive barriers using CRISPR & CRISPRa: an evaluation of the synthetic species concept in *Drosophila melanogaster*’, *Scientific Reports*, 8(1), pp. 1–15. Available at: <https://doi.org/10.1038/s41598-018-31433-2>.
- West, A.G., Gaszner, M. and Felsenfeld, G. (2002) ‘Insulators: Many functions, many mechanisms’, *Genes and Development*. 2002/02/05, 16(3), pp. 271–288. Available at: <https://doi.org/10.1101/gad.954702>.
- White, N.J. (2018) ‘Anaemia and malaria 11 Medical and Health Sciences 1108 Medical Microbiology 11 Medical and Health Sciences 1103 Clinical Sciences’, *Malaria Journal*, 17(1), pp. 1–17. Available at: <https://doi.org/10.1186/s12936-018-2509-9>.
- WHO (2021) *World Malaria Report 2021*, *World Malaria report Geneva: World Health Organization*. (2021). Licence: CC.
- Wise De Valdez, M.R. et al. (2011) ‘Genetic elimination of dengue vector mosquitoes’, *Proceedings of the National Academy of Sciences of the United States of America*, 108(12), pp. 4772–4775. Available at: <https://doi.org/10.1073/pnas.1019295108>.
- Wong, M.L. et al. (2020) ‘Natural *Wolbachia* infection in field-collected *Anopheles* and other mosquito species from Malaysia’, *Parasites and Vectors*, 13(1), pp. 1–15. Available at: <https://doi.org/10.1186/s13071-020-04277-x>.
- World Health Organization (2019) *World Malaria Report 2019*. Geneva. Available at: <https://www.who.int/publications-detail/world-malaria-report-2019>.
- Wu, B., Luo, L. and Gao, X.J. (2016) ‘Cas9-triggered chain ablation of cas9 as a gene drive brake’, *Nature Biotechnology*, pp. 137–138. Available at: <https://doi.org/10.1038/nbt.3444>.
- Xu, X. and Qi, L.S. (2019) ‘A CRISPR–dCas Toolbox for Genetic Engineering and Synthetic Biology’, *Journal of Molecular Biology*, 431(1), pp. 34–47. Available at: <https://doi.org/10.1016/j.jmb.2018.06.037>.
- Xue, B.K. and Leibler, S. (2018) ‘Benefits of phenotypic plasticity for population growth in varying environments’, *Proceedings of the National Academy of Sciences of the United States*

of America, 115(50), pp. 12745–12750. Available at:
<https://doi.org/10.1073/pnas.1813447115>.

Yang, J. and Corces, V.G. (2011) ‘Chromatin insulators: a role in nuclear organization and gene expression’, *Adv Cancer Res.* 2011/06/28, 110, pp. 43–76. Available at:
<https://doi.org/10.1016/B978-0-12-386469-7.00003-7>.

Zamyatin, A. et al. (2021) *Chromosome-level genome assemblies of the malaria vectors Anopheles coluzzii and Anopheles arabiensis*, *GigaScience*. Available at:
<https://doi.org/10.1093/gigascience/giab017>.

Zhang, X.H. et al. (2015) ‘Off-target effects in CRISPR/Cas9-mediated genome engineering’, *Molecular Therapy - Nucleic Acids*, 4(11), p. e264. Available at:
<https://doi.org/10.1038/mtna.2015.37>.

Zhao, Y. and Garcia, B.A. (2015) ‘Comprehensive catalog of currently documented histone modifications’, *Cold Spring Harbor Perspectives in Biology*, 7(9), pp. 1–20. Available at:
<https://doi.org/10.1101/cshperspect.a025064>.