

Institut für Insektenbiotechnologie
Professur für Insektenbiotechnologie im Pflanzenschutz
Justus-Liebig-Universität Gießen

**Development of CRISPR/Cas9-based tools and strains to improve
the Sterile Insect Technique in the Mediterranean fruit fly,
*Ceratitis capitata***



INAUGURAL-DISSERTATION
zur Erlangung des Doktorgrades (Dr. agr.)

im Fachbereich Agrarwissenschaften, Ökotoxologie und Umweltmanagement der
Justus-Liebig-Universität Gießen

vorgelegt von

Roswitha A. Aumann

geboren in Roding

Gießen, 2021

Mit Genehmigung des Fachbereichs Agrarwissenschaften, Ökotoxikologie und
Umweltmanagement der Justus-Liebig-Universität Gießen

1. Gutachter: Prof. Dr. Marc F. Schetelig

Justus-Liebig-Universität Gießen
Fachbereich Agrarwissenschaften, Ökotoxikologie und Umweltmanagement
Institut für Insektenbiotechnologie
Professur für Insektenbiotechnologie im Pflanzenschutz

2. Gutachter: Dr. Martin Rühl

Justus-Liebig-Universität Gießen
Fachbereich Biologie und Chemie
Institut für Lebensmittelchemie und Lebensmittelbiotechnologie
Gruppenleiter der Arbeitsgruppe für biochemische und molekularbiologische
Lebensmittelanalytik und Biotechnologie

3. Gutachter: Prof. Dr. Nikola Michael Pripic-Schäper

Justus-Liebig-Universität Gießen
Fachbereich Biologie und Chemie
Institut für Allgemeine Zoologie und Entwicklungsbiologie
Arbeitsgruppe Molekulare Entwicklungsbiologie

Prüfer: Prof. Dr. Annette Becker

Justus-Liebig-Universität Gießen
Fachbereich Biologie und Chemie
Institut für Botanik

Prüfer: Dr. Kostas Bourtzis

Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture
Insect Pest Control Laboratory, Seibersdorf (Österreich)

Vorsitz: Prof. Dr. Gunter P. Eckert

Justus-Liebig-Universität Gießen

Tag der Disputation: 12.07.2021

1	SUMMARY	1
2	INTRODUCTION	3
2.1	Tephritid fruit flies and their most devastating representative, <i>Ceratitidis capitata</i>	3
2.2	Insect pest control strategies: the Sterile Insect Technique	4
2.3	Classical genetic and transgenic approaches to improve the SIT	6
2.4	CRISPR/Cas gene editing and its potential to improve the SIT	12
2.5	Regulatory status of CRISPR-edited organisms	15
2.6	Research objectives	17
3	RESULTS	18
3.1	Expanding the toolbox: CRISPR/Cas9 HDR gene editing in <i>Ceratitidis capitata</i>	19
3.2	The Transformer Project: female-to-male sex conversion in <i>Ceratitidis capitata</i>	34
3.3	Solving a genetic puzzle in insect pest control: the <i>white pupae</i> gene	61
4	DISCUSSION	95
4.1	First CRISPR HDR gene editing in Tephritids	95
4.2	Improving SIT through CRISPR gene editing	97
4.3	Field application and regulatory status of CRISPR-edited organisms	104
4.4	A perspective for CRISPR-edited insects in SIT programs	108
5	REFERENCES	110
6	APPENDIX	121
7	CURRICULUM VITAE	132

1 Summary

The mass release of male sterilized insects, which leads to infertile matings and, consequently, to reduce an insect population, is the basis of the ‘Sterile Insect Technique’ (SIT) for pest control. Although the SIT has been used successfully since the 1950s, some key aspects need to be further improved and still established in most pest insects. For example, the cost-effective separation of males and females for generating a 100% male population, the so-called *sexing*, which is a key technology for the success of the SIT. New targeted methods such as CRISPR/Cas facilitate the development of genetically modified insects that enable the application or optimization of SIT programs for economically important pest species.

In this work, I present three studies that have significantly expanded the CRISPR/Cas genome editing and its applications in one of the most important pests, the Mediterranean fruit fly *C. capitata*. In the **first** study, I established sequence-specific, homology-directed repair (HDR) using CRISPR/Cas9 with an efficiency of up to 90%. This process was the world's first HDR process in Tephritids, and the high efficiency was essential for the other projects in this work. In the **second** study, all-male populations should be generated by manipulating sexual development. For this purpose, a temperature-sensitive variant of the *Drosophila melanogaster* sex determination gene *transformer-2* was targeted in the homologous gene in *C. capitata* by CRISPR/Cas9 HDR. The resulting progeny were exclusively male. Due to the permissive temperature range of this mutation in *C. capitata*, the conditional control of the gender conversion was not possible. In the **third** study, the molecular puzzle around the *white pupae* gene of medfly was solved. The phenotype has been used in Genetic Sexing Strains (GSS) for decades, but without the knowledge on the genotype. Through extensive genome sequencing and functional CRISPR/Cas9 knock-outs in the Mediterranean fruit fly, we were able to verify a metabolite transporter encoding gene as causal for the white pupal shell. This now allows the *white pupae* phenotype to be constructed in other species.

The results of this work are essential steps towards establishing a more universal approach to the development of GSS and optimizing SIT programs with CRISPR/Cas9 targeted mutagenesis without the use of transgenes – an approach in which the emerging strains are not classified as genetically modified organisms in many countries.

1 Zusammenfassung

Die Massenfreesetzung männlicher, sterilisierter Insekten, die zu unfruchtbaren Paarungen und damit zur Reduzierung einer Insektenpopulation führt, ist die Grundlage der "Sterilen Insekten Technik" (SIT) zur Schädlingsbekämpfung. Obwohl die SIT seit den 1950er Jahren erfolgreich eingesetzt wird, müssen einige Schlüsselaspekte bei den meisten Schadinsekten weiter verbessert oder erst noch etabliert werden. Beispielsweise ist die kostengünstige Trennung von Männchen und Weibchen zur Erzeugung einer 100%igen männlichen Population, dem sogenannten *Sexing*, eine Schlüsseltechnologie für den Erfolg des SIT. Neue Zielmethoden wie CRISPR/Cas ermöglichen die Entwicklung gentechnisch veränderter Insekten, die die Anwendung oder Optimierung von SIT-Programmen für wirtschaftlich wichtige Schädlingsarten ermöglichen.

In dieser Arbeit stelle ich drei Studien vor, die die CRISPR/Cas-Genomeditierung und ihre Anwendungen in einem der wichtigsten Schädlinge, der Mittelmeerfruchtfliege *C. capitata*, erheblich erweitert haben. In der **ersten** Studie habe ich sequenzspezifische, Homologie-gerichtete Reparatur (HDR) mittels CRISPR/Cas9 mit einem Wirkungsgrad von bis zu 90% etabliert. Dieser Prozess war der weltweit erste HDR-Prozess in Tephritiden, und die hohe Effizienz war für die weiteren Projekte in dieser Arbeit unerlässlich. In der **zweiten** Studie sollten 100% männliche Nachkommen durch Manipulation der sexuellen Entwicklung erzeugt werden. Zu diesem Zweck wurde eine temperaturempfindliche Variante des *Drosophila melanogaster* Geschlechtsbestimmungsgens *transformer-2* im homologen Gen in *C. capitata* mittels CRISPR/Cas9 HDR punktmutiert. Die daraus resultierenden Nachkommen waren ausschließlich männlich. Aufgrund des permissiven Temperaturbereichs dieser Mutation in *C. capitata* war eine konditionale Kontrolle der Geschlechtsumwandlung nicht möglich. In der **dritten** Studie konnte ich das molekulare Rätsel um das *white pupae* Gen in *C. capitata* lösen. Der Phänotyp wird seit Jahrzehnten in Genetischen Sexing Stämmen (GSS) verwendet, jedoch ohne den Genotyp zu kennen. Durch umfangreiche Genomsequenzierung und funktionelle CRISPR/Cas9 Knockouts in *C. capitata* konnten wir ein Metabolit-Transporter-kodierendes Gen als kausal für die weiße Puppenhülle verifizieren. Dadurch kann nun der weiße Puppen-Phänotyp auch in anderen Spezies erzeugt werden.

Die Ergebnisse dieser Arbeit sind wesentliche Schritte zur Etablierung eines universelleren Ansatzes für die Entwicklung von GSS und zur Optimierung von SIT-Programmen mittels zielgerichteter CRISPR/Cas9-Mutagenese ohne die Verwendung von Transgenen – ein Vorgehen, bei dem die entstehenden Stämme in vielen Ländern nicht als genetisch veränderte Organismen eingestuft werden.

2 Introduction

Despite the application of around three million metric tons of pesticide per year, pests destroy more than 40% of the global crop production (Pimentel, 2007). However, it was predicted that farmers will have to increase their cereal yields by at least 40-50% to be able to meet worldwide food supply needs in the future (Alexandratos, 1999; Maxmen, 2013), as the world population grows by 75-80 million people each year, resulting in a prospected world population of roughly nine billion people in 2050 (Alexandratos, 1999), and the number of undernourished people is rising (FAO et al., 2019).

While increasing crop yields is highly demanding *per se*, climate change, manifested in extreme weather events, rising temperatures, and changing rainfall patterns, may additionally stress the natural resources needed for crop production and substantially reduce harvests (Chakraborty and Newton, 2011; Wheeler and von Braun, 2013). Elevated temperatures may also positively influence the growth rate and expand the habitats of specific pest insect populations, and pests could become a year-round problem in several regions (Sultana et al., 2017; Deutsch et al., 2018). In parallel, agricultural intensification has led to increased pesticide applications, causing higher genetic selection pressure and growing resistance to these chemicals among pest insects, as well as a severe decline of beneficial insects (Whalon et al., 2008; Habel et al., 2019). So far, insect pests are already accounted for one-third of the world crop production losses (Pimentel, 2007; Maxmen, 2013). Considering the factors mentioned above, their impact might increase enormously in the near future. Therefore, sustainable insect pest control will play a crucial role in ensuring food security.

2.1 Tephritid fruit flies and their most devastating representative, *Ceratitis capitata*

The family of Tephritidae or ‘true fruit flies’, with nearly 5,000 species in 500 different genera, represents some of the most destructive, damaging, and economically important pest insects for the horticultural industry worldwide (White and Elson-Harris, 1992). About one-third of the known species feed on fruits; others infest and feed on stems, roots, shoots, flowers, buds, seeds, or leaves (Allwood and Leblanc, 1997). The genera *Anastrepha* (Schiner), *Bactrocera* (Macquart), *Dacus* (Fabricius), *Rhagoletis* (Loew), and *Ceratitis* (Macleay) pose the greatest threat to fruit and vegetable production, and several species among these genera are classified as quarantine pests (White and Elson-Harris, 1992; Jiang et al., 2018). Females lay their eggs under the skin of a broad range of fruits and vegetables, and larvae subsequently feed on the fruit flesh, causing the destruction of the host fruits and severe harvest losses (Fig. 1A). Furthermore, efforts to control these pest insects and associated provisions like quarantines, regulatory inspections, and trade restrictions cause losses and expenses for farmers, and are major cost factors of fruit production (Allwood and Leblanc, 1997; Gregory et al., 2009). The range of actual economic impact of fruit fly species differs depending on factors like host plants, mating behavior, generation time, lifespan,

temperature tolerance, survival during critical periods, dispersal characteristics and invasion potential (Allwood and Leblanc, 1997; Malacrida et al., 2007).

A highly successful global invader, and probably the most devastating among the Tephritids, is the Mediterranean fruit fly *Ceratitidis capitata* (Wiedemann) – also known as medfly – a highly polyphagous species with more than 250 different host plants (Fig. 1B). Medfly causes 20-25% loss of citrus, 91% of peaches, 55% of apricots, and 15% of plums in Jordan (Allwood and Leblanc, 1997), and 20-30% of mango and citrus in Africa (Badii et al., 2015). Average annual losses (1980-1989) attributed to medfly on citrus, pome, and stone fruits in Maghreb countries were 101.5, 41.9, and 71.3 thousand tons of fruit, respectively, causing a total loss of approximately US\$ 60 million – in addition to US\$ 9.6 million spent on pesticides to control medfly (IAEA, 1995). Costs for eradicating a medfly outbreak in Florida's Tampa Bay region in 1997 were about US\$ 25 million. Yet, the money was well spent, as controlling a potential establishment of medfly in the State of California was estimated to cost between US\$ 493-875 million and could have resulted in trade embargos causing revenue losses of about US\$ 564 million (Szyniszewska and Tatem, 2014). Initially, the habitat of *Ceratitidis* was Afrotropical. However, due to global fruit trade and tourist industry, *C. capitata* is now distributed in most tropical and temperate regions of the world and continuously threatens to invade or re-invade new areas (Malacrida et al., 2007). Until now, the distribution of medfly was limited to the South of Europe (Vera et al., 2002), however, it has already been shown that medfly can adapt and cope well with different climate conditions (Ricalde et al., 2012), and increasing temperatures will most probably expand the range of medfly northwards (Gutierrez and Ponti, 2011).

2.2 Insect pest control strategies: the Sterile Insect Technique

Several strategies can be applied to fight mobile insect pests like *C. capitata*. Measures have to be implemented area-wide and should be adjusted to the target population density, invasiveness, and infested area (rural or urban).

Chemical insecticides, acting as a stomach or contact poison, can act very quickly and effectively and are heavily used to fight pest insects (Devine and Furlong, 2007). However, their use is highly controversial due to several severe societal and ecological costs: many insecticides are not species-specific, i.e., they harm not just the intended species, but also non-target species. This potentially causes a shift in the agroecosystem, due to a negative impact on the environment, the biodiversity, beneficial insects and natural enemies of the pests (Epstein et al., 2000; Devine and Furlong, 2007). The lack of natural enemies can lead to pest resurgence, resulting in an additional and repeated need for insecticide applications (Wearing, 1982), which, in turn, promotes the rapid development of resistance and thereby reduces efficiency. These reasons, along with the effects on human health through water and soil pollution, residual effects on crops, and direct pesticide exposure, led to high public rejection (Kahn et al., 1990), and a steady demand for more environment-friendly pest control strategies (Hendrichs et al., 2007).

The concept of ‘integrated pest management’ (IPM), defined as ‘a decision support system for the selection and use of pest control tactics, singly or harmoniously coordinated into a management strategy, based on cost/benefits analyses that take into account the interests of and impacts on producers, society, and the environment’ (Kogan, 1998), puts one of its main emphases on decreasing chemical inputs by combining several strategies, such as cultural, physical, biological, and genetic control. Cultural control targets to disrupt a pests' reproductive cycle via farm sanitation and crop hygiene measures (Badii et al., 2015), physical or mechanical protection includes netting trees, bagging individual fruits, and trapping with sticky or pheromone traps. The use of parasitoids, predators, or pathogens is called biological control. Genetic control, also known as birth or autocidal control, is based on reducing the pest population's reproductive potential through induced sterility. A successful and widely used genetic control strategy, designed for area-wide (AW) IPM programs, is the Sterile Insect Technique (SIT).

The idea of controlling and suppressing pest insect populations by releasing sterile conspecifics into wild populations was developed independently during the 1930s and 1940s by three researchers: A. S. Serebrowskii, F. L. Vanderplant, and E. F. Knipling. While Serebrowskii and Vanderplant focused on hybridization-mediated sterility, Knipling suggested using ionizing radiation to induce dominant lethal mutations in the germline and thus achieve sterility (Klassen and Curtis, 2005). In 1954, Knipling proved his concept on the island of Curaçao and managed to eradicate the New World screwworm from 435 km² within 14 weeks (Baumhover et al., 1955; Klassen and Curtis, 2005). Subsequently, Knipling published on the ‘possibilities of insect control or eradication through the use of sexually sterile males’ (Knipling, 1955) and is, therefore, considered the SIT founder. In 1992, Knipling and his colleague R. C. Bushland were awarded the World Food Prize, acknowledging their environmentally friendly and effective measures to control insects that threaten crops and livestock production.

Today, the great potential of SIT is broadly acknowledged. Programs target several insect pests of importance – agricultural pests like the West Indian, the Mexican, the Queensland, the Oriental, the Mediterranean, and the Melon fruit fly, the pink bollworm or the codling moth, and pests of veterinary and medical importance, like the New World screwworm species, Tsetse flies, the yellow fever mosquito *Aedes aegypti*, or the malaria vector *Anopheles albimanus* (Klassen and Curtis, 2005).

The necessary steps during an SIT program are: 1) mass rearing of the insect, 2) sexing, i.e., separation of males and females, 3) marking of insects to enable monitoring of released flies in the field, 4) sterilization (classically by irradiation), 5) releases of large numbers of sterilized (male) insects to promote infertile mating with wild females, and 6) field control and data analysis to assess the success of the program (Fig. 1C). If a wild female mates with a sexually sterile male, the pest's reproductive cycle is interrupted. Females may still lay eggs; however, they are not viable, and no offspring will eclose (Knipling, 1955; Klassen, 2005; Wimmer, 2005).

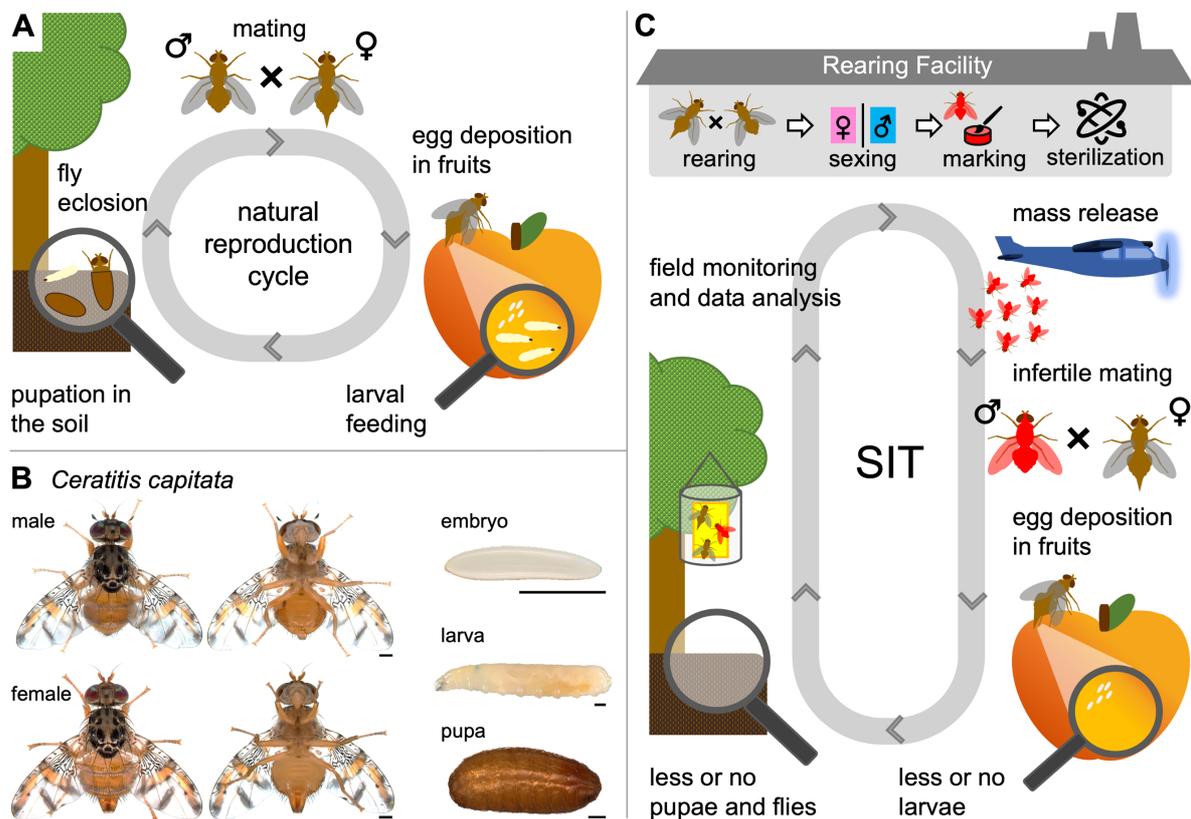


Fig. 1. The Mediterranean fruit fly (Diptera: Tephritidae): life cycle, characteristics, and genetic population control. **A)** In the wild population, male and female mate; females deposit eggs under the skin of fruits. Larvae subsequently feed on the fruit flesh and drop to the ground for pupation and eclosion of a new generation of flies. **B)** *Ceratitis capitata* male and female adult fly, dorsal and ventral view, embryo, 3rd instar larva, and pupa. Adult medflies have two light-colored stripes on the abdomen, a yellow-brownish band pattern on the wings, and a black thorax with irregular, white patches. Males have bristles with enlarged spatula-shaped tips on the head. Scale bar = 0.5 mm **C)** The Sterile Insect Technique (SIT) aims to reduce the reproductive potential of a pest population through induced sterility. Therefore, the target insect is mass-reared, sexed, marked with fluorescent dyes, sterilized by irradiation, and subsequently released over the infested area. Overflooding the area with sterile males promotes infertile mating of wild females with mass-reared males, whereby the pest population size gets reduced over time. The ratio of factory-produced and wild type flies is monitored by trapping to adjust the release numbers if necessary.

2.3 Classical genetic and transgenic approaches to improve the SIT

Key requirements for an SIT program are the ability to rear, sterilize, and release a sufficient amount of male insects to reach the necessary sterile to wild insect ratio in the field (Lance and McInnis, 2005), and the production of males with high fitness and mating competitiveness, able to successfully compete with their wild counterparts (Lance and McInnis, 2005). Thus, the correct setup of the mass rearing and release procedures is critical, as many aspects, e.g., the artificial diets, the temperature, the stock density, as well as the chilling and transport before the release, can negatively affect the insect quality and limit the success of an application (Lance and McInnis, 2005; Diallo et al., 2019). Other critical procedures are sexing, sterilization, and tracing of released flies. Those can be addressed and potentially improved by transgenic and/or classical genetic approaches.

Reproductive sterility

The currently used exposure to gamma rays from isotopic sources (cobalt-60 or cesium-137 (Robinson, 2002)), or high-energy electrons and X-rays to induce reproductive sterility is a safe and highly reliable method that has been established for over 300 arthropod species (Bakri et al., 2005). However, radiation can have a significant negative impact on the insects' biological quality, and thus the success of an SIT program (Alphey, 2007).

Transgenic genetic sterilization approaches based on the inheritance of dominant lethal genes were developed to address this issue. To facilitate rearing of such strains, conditionality, achieved, for example, via a tetracycline-controlled expression system (Tet-off), is an essential aspect. Under permissive conditions (+ tetracycline), normal rearing of the insect colony is possible, while under repressive conditions (- tetracycline), the lethal gene is activated, and no offspring is produced. Notably, the necessity of large amounts of antibiotics for strain maintenance in the factories is a general downside of tetracycline-inducible expression systems. It can decrease fly fecundity, is costly, and critical for the waste disposal (Schetelig and Handler, 2012a; Rashid et al., 2018). Using transgenic approaches, embryonic lethality systems were engineered for *D. melanogaster*, *C. capitata*, *Anastrepha suspensa*, and *Anastrepha ludens* (Horn and Wimmer, 2003; Schetelig et al., 2009; Schetelig and Handler, 2012a; Schetelig et al., 2016). Additionally, a late-acting system known as RIDL (Release of Insects carrying a Dominant Lethal) has been developed (Thomas et al., 2000; Gong et al., 2005). However, a recent paper demonstrated the risk of a genetic breakdown in such conditional lethality systems due to primary-site mutations and second-site suppressor effects (Zhao et al., 2020). Thereby, the need for a backup system, i.e. a second, functionally unrelated killing mechanism, became highly evident (Handler, 2016). Furthermore, it should be noted that these systems cannot be equated with SIT, as the release of fertile transgenic organisms is regulated differently by law (Black et al., 2011; Lutrat et al., 2019).

Monitoring/Marking

An SIT program's essential requirement is to distinguish WT and factory-reared flies for field-monitoring, detecting accidental contamination in the factory or determining if non-irradiated flies were released. Therefore, flies should optimally carry visible markers (Franz, 2005). Widely used are external fluorescent dyes, which are applied as fine dust. Yet, there are economic and biological costs, like the additional processing step and the effect on the insect quality, respectively, and the fine dust can pose a health hazard to facility workers (Parker, 2005). Furthermore, their reliability is limited, as the dye can be transferred to wild flies upon mating, and there is no possibility to distinguish the flies at the genomic level.

Morphological markers, such as a third white stripe on the abdomen of the medfly caused by a mutation in the *Sergeant-2* (Sr^2) gene, can be used to distinguish GSS and WT flies (Niyazi et al., 2005; Rempoulakis et al., 2016). Advantageously, such a marker does not require an additional processing step and is reliable. However, its use is limited to medfly, as the genetic basis of Sr^2

could not be resolved yet. Therefore, its detection at the DNA level is not possible, which would be important to track accidentally released non-irradiated flies and their offspring.

These limitations can be addressed by transgenic approaches (Alphey, 2002; Robinson and Hendrichs, 2005): stably integrated fluorescent proteins can serve as externally visible marker, are detectable at the DNA level, and are observable as well as provable even in dead flies (Nirmala et al., 2011; Simmons et al., 2011; Walters et al., 2012; Rempoulakis et al., 2016). Combined with specific promoters, they can either be expressed ubiquitously or just in specific tissues, e.g. testes (Scolari et al., 2008; Nirmala et al., 2011).

Sexing

Sexing is a very critical, yet an often unsolved issue (Franz et al., 2021). Released sterile females can still sting fruits or, in case of vector insects, transmit diseases, and furthermore reduce the efficiency of an SIT program by mating with the co-released sterile males (Hendrichs et al., 1995; Rendon et al., 2004). Additionally, sexing in an early developmental stage reduces rearing costs and increases the production rate (Lutrat et al., 2019). However, if no sexing strain is available and approved, bisexual releases have to be performed. In Panama, about 15 million mass-reared sterile male and female screwworm flies are released per week to maintain the New World screwworm free barrier zone between Panama and Colombia (Scott et al., 2017). Removing females at the pupal stage is predicted to save over US\$ 1 million per year, as the costs for marking, irradiation, transport, and release would be reduced by half (Concha et al., 2016; Lutrat et al., 2019). Sexing at the embryonic stage would additionally save space and cost during the larval rearing process. Thus, it is important to produce and release a male-only population to improve the economics of production and biological efficiency and safety in the field (Klassen, 2005). In principle, there are two options to achieve this: i) female elimination, which can be accomplished by mechanical sorting based on natural (sexual) dimorphisms like differences in size, development rate or phenotype (pupal color), or genetically, by introducing conditional female-specific lethality (Lutrat et al., 2019); and ii) female-to-male conversion. Here, female embryos (considering an XX karyotype in an XX/XY sex-determination system) are transformed into phenotypic male adults. This could be achieved via the introduction of mutations in genes involved in the sex-determination pathway. Such an approach would be highly beneficial for SIT mass rearing, because it can double the number of male offspring per parental egg capacity.

Important aspects of sexing methods are the male recovery rate and the female contamination rate, the time needed for sorting, the initial investment, the treatment costs, and the developmental sorting stage (embryonic, larval, pupal, adult stage) (Lutrat et al., 2019). For elimination of females as well as for sex conversion, conditionality is also crucial.

To improve several of those aspects, transgenic methods have been used to develop multiple transgenic sexing strategies (TSS) in several pest species. One was the sex-specific expression of a fluorescent marker gene that enabled mechanical sex-sorting (Catteruccia et al., 2005). Others

depended on the tetracycline-controlled conditional female-specific overexpression of a lethal or pro-apoptotic gene that mediated sex-specific lethality, e.g. in female-specific RIDL systems (fsRIDL) (Fu et al., 2007; Ant et al., 2012; Li et al., 2014), or in transgenic early-larval or early-embryonic sexing systems (TESS; Schetelig and Handler, 2012b; Ogaugwu et al., 2013; Yan and Scott, 2015; Concha et al., 2016; Schetelig et al., 2016; Yan et al., 2017; Concha et al., 2020; Yan et al., 2020), respectively. A recently developed strategy is the so-called subtractive transgene sex sorting (STSS; Das et al., 2020). Here, two transgenic strains, one with an conditional lethal construct on the X-chromosome, one with an conditional lethal construct on the Y-chromosome, are used to produce non-transgenic males in an two-step mating scheme (Das et al., 2020): rearing the Y-linked strain on restrictive diet (- Tet) results in only non-transgenic female offspring in the next generation. These females are then crossed to the X-linked strain (both sexes) on restrictive diet. Mating between non-transgenic females and males with X-linked lethality produces only transgene-free male offspring. Female offspring, and all offspring produced by mating of the X-linked strain *inter se*, will not survive on the restrictive diet (Das et al., 2020). However, this system has only been engineered in *D. melanogaster* so far.

Efforts have also been made to engineer (conditional) female-to-male sex conversion. First proof-of-principle studies were successfully conducted by targeting homologs of the *transformer* or *transformer-2* genes, which are essential for the female development, in *C. capitata*, *A. suspensa*, and *D. suzukii* (Saccone et al., 2007; Schetelig et al., 2012; Li and Handler, 2017), or by overexpressing the maleness factor in *C. capitata* (*Maleness-on-the-Y, MoY*) and *A. aegypti* (*Nix*) (Meccariello et al., 2019; Aryan et al., 2020). However, sex conversion systems developed so far are either based on transient knock-down via RNA interference (RNAi) and were neither conditional nor stable (Saccone et al., 2007; Schetelig et al., 2012), or not usable due unwanted side effects (Li and Handler, 2017; Aryan et al., 2020).

The most successful sexing systems so far are the so-called genetic sexing strains (GSS), which rely on classical genetics and have been developed for the Mediterranean fruit fly (GSS VIENNA 7 and VIENNA 8; Rendon et al., 2004; Augustinos et al., 2017; Franz et al., 2021), and the Mexican fruit fly, *A. ludens* (GSS Tapachula 7; Orozco et al., 2013; Zepeda-Cisneros et al., 2014). It should be mentioned that GSS have been developed for several species, but, so far, only GSS of these two species are developed enough to be used for automatic sexing on a mass-rearing scale over extended periods of time (Franz et al., 2021). The construction of a GSS requires two essential components: (1) a recessive mutation, which can act as a selectable marker for sex separation, e.g., a clearly visible phenotypic mutation or a conditionally lethal mutation which is not harmful under heterozygous and untriggered conditions, and (2) the rescue of the wild-type (WT) allele in males via linkage to the maleness factor or the Y-chromosome (Franz et al., 2021). This results in homozygous females showing the mutant phenotype, and heterozygous males with the wild-type phenotype (Franz et al., 2021).

The unknown molecular basis of the medfly GSS

Originally, medfly GSS were solely built on *white pupae* (*wp*; Rössler, 1979), a recessive mutation with white pupal case phenotype, and a radiation-induced translocation. The white pupae phenotype first spontaneously appeared in the 5th generation of an inbred line originating from a cross between an irradiated male and a non-treated female in 1977 (Rössler, 1979). It was then combined with a Y/autosome translocation T(Y;5), to place the WT allele on the Y-chromosome and rescue the WT phenotype (brown puparium) in males. Thus, the first-generation GSS (T:Y(*wp*⁺)101) was produced (Robinson and Van Heemert, 1982; Franz et al., 2021). Later, a recessive lethal mutation, *temperature-sensitive lethal* (*tsl*), which is heat-inducible and highly effective during the embryonic stage, was discovered during an ethyl methanesulphonate screening (Busch-Petersen, 1990). This led to the construction of the second-generation GSS, carrying the recessive mutations *tsl* and *wp*, as well as a translocation to achieve sex specificity (Rössler and Rosenthal, 1992; Augustinos et al., 2017; Franz et al., 2021). As the chromosomal breakpoint position of the translocation is critical for the system's stability and the vitality of the flies, it was necessary to induce and test several translocations; only two have been approved for mass rearing and are used in the GSS strains VIENNA 7 (T(Y;5)3-129; position 58B on the trichogen cells polytene chromosome map) and VIENNA 8 (T(Y;5)101; position 52B on the trichogen cells polytene chromosome map). To reduce recombination between the marker genes, medfly GSS may additionally carry the radiation-induced pericentric inversion 'D53', spanning a large region of chromosome 5 (Franz et al., 2021) (Fig. 2A). In the final product, males emerge from brown pupae and are resistant to high temperatures, while females emerge from white pupae and are sensitive to high temperatures (Franz et al., 2021) (Fig. 2B). A simple heat shock (34-35°C, 24 h) in the embryonic stage kills all female embryos of these GSS, while males are retained. This system improved mass-rearing tremendously and enabled a maximum weekly production of 3.5 billion sterile male flies in different facilities worldwide (Franz et al., 2021).

However, the strains still need improvement. GSS males are semi-sterile due to the segregation behavior of the Y/autosome translocation during male meiosis (Franz et al., 2021). Half of the offspring are genetically imbalanced and not viable, as they do not receive a normal chromosomal complement (Laven, 1969; Franz et al., 2021). Also, as the radiation-induced chromosomal breakpoints are at random positions, tedious genetic and cytogenetic analyses are necessary to select the most suitable translocation (Franz et al., 2021). Furthermore, the molecular genetic basis of *wp* and *tsl* has not been resolved, despite a more than 20 years lasting search. The inheritance pattern suggested that *wp* is monogenic, recessive, and autosomal (Rössler, 1979), and classical genetics and cytogenetic studies showed that *wp* and *tsl* are tightly linked and localized on the right arm of chromosome 5 (Kerremans and Franz, 1994; Zacharopoulou et al., 2017).

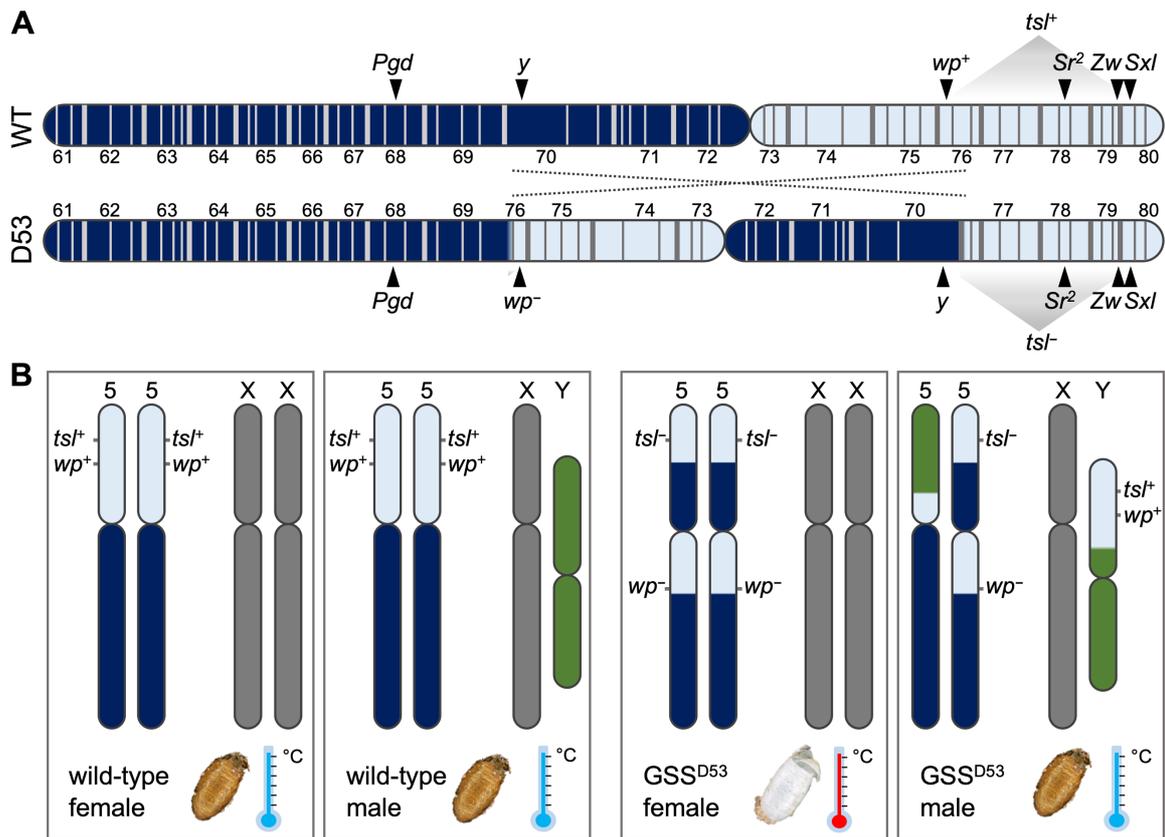


Fig. 2. The genetic basis of medfly VIENNA GSS. **A)** Schematics of a medfly chromosome 5 salivary gland polytene chromosome map without (WT) and with the pericentric inversion ‘D53’. To illustrate the inversion, the right arm of WT chromosome 5 is colored in light blue, the left arm in dark blue. Positions of the genes 6-phosphogluconate dehydrogenase (*Pgd*), yellow (*y*), white pupae (*wp*), Sergeant-2 (*Sr²*), glucose-6-phosphate 1-dehydrogenase (*Zw*), and sex lethal (*Sxl*), determined by deletion mapping or *in situ* hybridization, are indicated (Kerremans and Franz, 1994; Niyazi et al., 2005; Papanicolaou et al., 2016; Franz et al., 2021). Transposition mapping was used to determine the location of the *temperature-sensitive lethal* (*tsl*) locus on the trichogen cell chromosome map (59C-61C; Kerremans and Franz, 1994). The respective region is marked on the depicted salivary gland chromosome map. **B)** Chromosomes 5 (light blue/dark blue), X (grey) and Y (green) schematics for medfly WT and GSS females and males, under the assumption that *tsl* is located outside of the D53 inversion. In GSS carrying the D53 inversion (GSS^{D53}), females are homozygous for the *wp* and *tsl* mutations (*wp⁻*, *tsl⁻*) and the inversion, whereas males are heterozygous for all traits. A translocation (Y;5) (schematically shown) rescues the WT phenotype (*wp⁺*, *tsl⁺*) in GSS males. The phenotypes for *wp* and *tsl* are depicted below the chromosomes: *wp⁺⁺*, *wp⁺* = brown puparium; *wp⁻* = white puparium; *tsl⁺⁺*, *tsl⁺* = temperature resistant (blue thermometer); *tsl⁻* = temperature sensitive (red thermometer).

Establishing a generic sexing system for field use

Medfly GSS show significant success in SIT programs worldwide (Augustinos et al., 2017). Unfortunately, a targeted transfer of the GSS phenotypes to other species is not possible, because the genes and underlying mutations causing the *wp* and *tsl* phenotypes, as well as the translocation and inversion breakpoints, are unknown. White puparium color mutants do occur spontaneously in field or laboratory stocks and were also found in *Bactrocera dorsalis* (McCombs and Saul, 1992) and *Zeugodacus cucurbitae* (McInnis et al., 2004). However, for other Tephritids, such as *Bactrocera tryoni*, *Bactrocera oleae*, or *A. ludens*, no white pupae color mutant could be detected, despite extensive screening efforts (Ward et al., 2021).

Transgenic systems promise significant improvements of strains for SIT programs (Benedict and Robinson, 2003; Marec et al., 2005; Morrison et al., 2010; Häcker and Schetelig, 2018; Häcker et al., 2021) and also have the potential to be transferred to other target species. However, regulations concerning the release of transgenic or genetically modified (GM) insects often prohibit the use of such strains in the field (Reeves et al., 2012). Also, public acceptance, which can be critical for the adoption or rejection new technologies, might be low (Alphey et al., 2002; Panjwani and Wilson, 2016).

Therefore, the most promising approach to create applicable GSS for more pest species seems to unravel the genetic basis of the medfly GSS traits *w^p* and *ts^l* and find a (non-transgenic) way to transfer these to other insects. Knowledge about the genetic basis of these traits would furthermore allow the differentiation and a molecular tracing of WT and released GSS flies on a genomic level, which is an important safeguarding aspect for SIT programs. The discovery of maleness factors in different species (Hall et al., 2015; Krzywinska et al., 2016; Sharma et al., 2017) and recently in medfly (*MoY*; Meccariello et al., 2019) could furthermore eliminate the need of Y/autosome translocations and help to pave the way to construct GSS by linking the rescue alleles to the maleness determining region using gene editing technologies. The genome editing technology CRISPR/Cas HDR (clustered regulatory interspaced short palindromic repeats/CRISPR associated protein, homology-directed repair) (Doudna and Charpentier, 2014) may enable such transfers and facilitate a generic approach.

2.4 CRISPR/Cas gene editing and its potential to improve the SIT

CRISPR/Cas is an adaptive prokaryotic immune system that provides acquired immunity against foreign genetic elements such as phages and plasmids, based on homology-mediated detection and subsequent degradation (Barrangou et al., 2007; Horvath and Barrangou, 2010). Depending on the function of the involved Cas proteins, the organization of the *cas* operons and the signature of *cas* genes, CRISPR/Cas immune systems have been grouped into three types (I-III) and numerous subtypes (Makarova et al., 2011; Koonin and Makarova, 2019). Type I and III are the most abundant systems but are often not fully characterized due to the complex subunit structures of the targeting complexes (Brouns et al., 2008; Barrangou, 2015). The targeting complex of the well-studied Type II systems, however, requires only two components: Cas9, a single polypeptide DNA endonuclease that generates double-stranded breaks (DSB) (Sapranauskas et al., 2011), and a dual guide RNA (Jinek et al., 2012).

Genomes of most Bacteria and Archaea contain so-called clustered regularly interspaced short palindromic repeats (CRISPR), i.e., arrays of short, conserved repeat sequences, interspaced by unique DNA sequences of similar size (spacers) (Sapranauskas et al., 2011). Upon viral challenge, new spacers can be derived from the invaders' genome and be integrated into the CRISPR array (acquisition). However, new spacers are only selected and acquired if they are next to a protospacer adjacent motif (PAM), a short sequence motif proximal to the target sequence (5'-NGG-3' for

Streptococcus pyogenes Cas9). Together with *cas* genes, these CRISPR arrays subsequently provide protection against the foreign DNA, based on sequence similarity between the spacer and the invading DNA (Barrangou et al., 2007). The PAM site is crucial to discriminate between self- and non-self-sequences, as the targets found in the foreign DNA contain a PAM, whereas matching targets in the CRISPR locus itself do not contain a PAM and are therefore not targeted (Yosef et al., 2012). In case of a re-exposure to the invader, dual guide RNAs are produced: Precursor CRISPR RNA (pre-crRNA) molecules are transcribed from the CRISPR locus, and separately transcribed trans-activating CRISPR RNAs (tracrRNA), complementary to the repeat sequences in the pre-crRNAs, trigger their processing into crRNAs by RNase III (Jinek et al., 2012; Jiang and Doudna, 2017). Finally, a dual guide RNA is formed from the crRNA, which is responsible for recognizing the target sequence via RNA-DNA base pairing, and the tracrRNA, which forms a double-stranded stem to facilitate Cas9 recruitment (Deltcheva et al., 2011). Cas proteins bind the dual guide RNAs to form ribonucleoprotein (RNP) complexes, each complex containing a single ‘guide’ spacer sequence (Sternberg and Doudna, 2015). Upon specific RNA-DNA base pairing and in the presence of the PAM site, the Cas9 nuclease domains HNH and RuvC-like create a site-specific double-stranded break three nucleotides upstream of the PAM site by cleaving the complementary and non-complementary DNA strand, respectively (Jinek et al., 2012) (Fig. 3). In consequence, the invaders' genome is destroyed, and immunity is ensured.

Today, the acronym CRISPR/Cas is commonly associated with genome engineering: Cas9 and a designed and engineered guide RNA, combining the features of crRNA and tracrRNA in one single guide RNA (sgRNA), are introduced into cells, bacteria, or embryos to form RNP complexes, and are used to precisely edit genes via the induction of DSBs and their subsequent repair (Sternberg and Doudna, 2015). Cas9 can be delivered to the organism either as DNA (expression plasmid), mRNA, or heterologous expressed protein, the sgRNA(s) as *in vitro* transcribed RNA, or expressed through plasmids (Wilbie et al., 2019). The relative simplicity of the CRISPR/Cas9 system in terms of composition, design, and target sequence requirements, the ease of customization, its high efficiency and specificity based on RNA, as well as its comparably low price, has led to an extraordinarily fast adaptation of this technique by many researchers and for many different species and enabled remarkable innovations (Barrangou and Doudna, 2016).

Editing a genome is possible by exploiting the cell's endogenous DNA repair machinery, activated upon DNA cleavage by Cas9 (Fig. 3). Two repair pathways are of interest for gene editing: the non-homologous end-joining (NHEJ) and the homolog-directed repair (HDR) pathway (Ran et al., 2013). If DSBs are re-ligated through the NHEJ process, random insertion or deletion mutations might occur at the break site. NHEJ is therefore known as an ‘error-prone’ pathway, mostly utilized to mediate gene knock-outs, as indels can lead to frameshift mutations or premature stop codons, causing truncated or non-functional proteins. The introduction of such alterations is

categorized as a site-directed nuclease (SDN) type 1 application (Eckerstorfer et al., 2019). Furthermore, NHEJ can be used to create targeted genomic rearrangements (Schmidt et al., 2020a).

In the presence of a natural or a designed and exogenously introduced repair template, i.e., the sister chromatid or a double- or single-stranded DNA with homology arms flanking the target sequence, respectively, the HDR pathway might be activated and offers the option to generate precise, defined modifications at the target locus. If HDR is exploited for genome editing, it is referred to as SDN-2 if the repair template differs by one or a few nucleotides, and SDN-3 if the repair template differs by several kilobase pairs or if it is of foreign origin (Eckerstorfer et al., 2019). Due to its precision, HDR is often the preferred approach to gene editing. However, it is more challenging to achieve than NHEJ, because it is only active in dividing cells and therefore occurs at a lower frequency. Furthermore, its efficiency depends on the cell type and the genomic locus (Saleh-Gohari and Helleday, 2004), the length of the insert and the homology arms, and the position of the DSB relative to the editing site (Paquet et al., 2016; Paix et al., 2017).

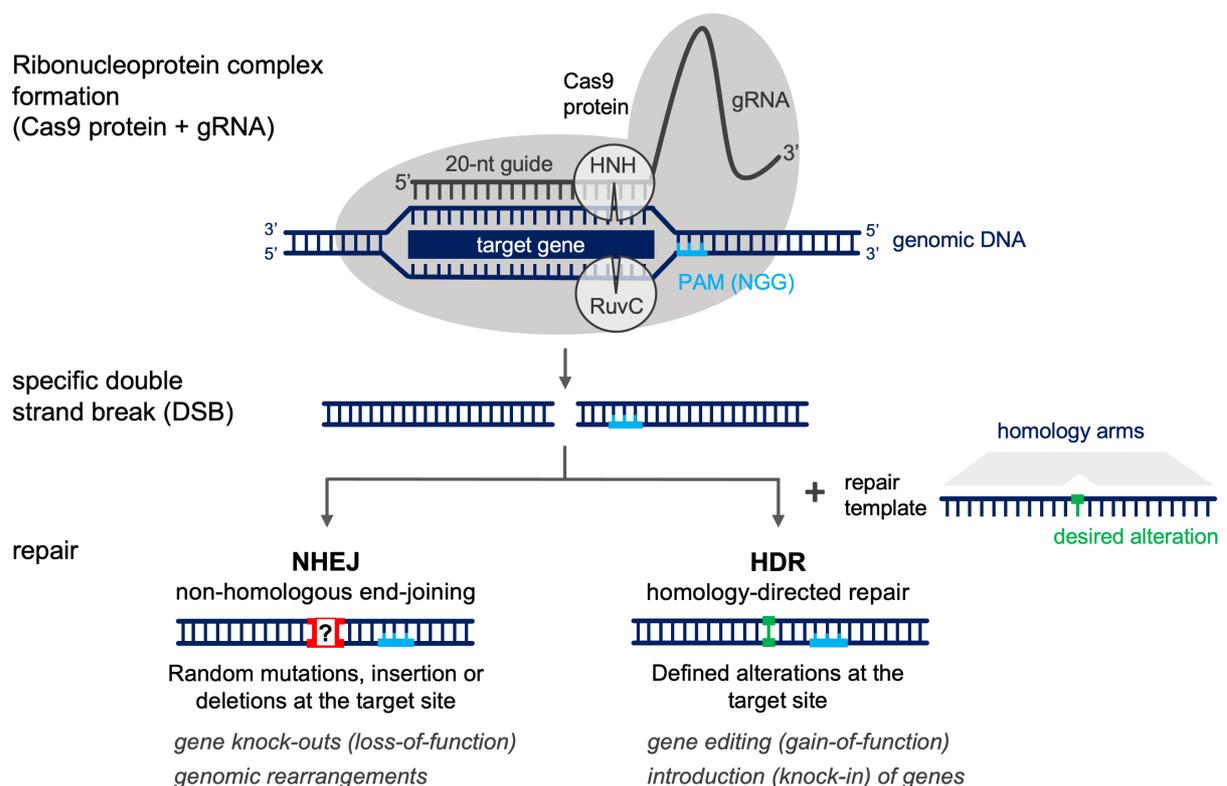


Fig. 3. Schematic illustration of CRISPR/Cas9 gene editing. A ribonucleoprotein (RNP) complex is formed by the specifically designed single gRNA and the Cas9 protein. RNA-DNA base pairing and the presence of the PAM (light blue) ensure site-specific cleavage of the genomic DNA three nucleotides upstream of the PAM site, mediated by the nuclease domains HNH and RuvC-like. The resulting double-strand break is then repaired by an intrinsic cellular repair pathway, most commonly NHEJ or HDR. Both pathways can be exploited for genome editing: the NHEJ pathway is used for gene knock-outs, due to random mutations, insertions or deletions at the target site, or targeted genomic rearrangements. The HDR pathway enables specific gene editing, such as introducing single nucleotide substitutions or the insertion/deletion of specific nucleotides.

CRISPR/Cas9 gene editing has also already been used in at least 26 insect species and holds great promise to improve critical steps like sexing or marking in SIT programs (Sun et al., 2017; O'Brochta et al., 2020). It could be used to alter the genome without inserting exogenous DNA and enable a 'scar-less' targeted transfer of known mutations from insect model organisms to homologous genes of other species via the HDR pathway or precise knock-outs through the NHEJ pathway. Furthermore, transgenic elements could be inserted at specific positions without the need of transposon-derived sequences flanking the construct. Thus, the risk of re-mobilization, a highly discussed issue in transposon-mediated transgenesis, could be eliminated (Dafa'alla et al., 2006). Moreover, CRISPR could be used for genome editing to achieve targeted chromosomal rearrangements such as inversions or translocations (Schmidt et al., 2020a), which could be used to reduce recombination events between traits or achieve sex specificity, respectively.

The first step in using CRISPR to engineer SIT relevant traits was to establish the technique in pest insects of interest, for example members of the Tephritid family. Therefore, genes whose functional knock-out would create a visible phenotype, like eye pigmentation alterations (*white*) or deformations (*paired, multiple edematous wings, transformer*), were targeted and knocked out using the NHEJ pathway in *A. suspensa* (Li and Handler, 2019), *A. ludens* (Sim et al., 2019), *B. dorsalis* (Bai et al., 2019; Zheng et al., 2019), *B. tryoni* (Choo et al., 2018), *B. oleae* (Meccariello et al., 2020), or *C. capitata* (Meccariello et al., 2017). However, HDR-mediated gene editing has not been established in the family of Tephritids so far.

2.5 Regulatory status of CRISPR-edited organisms

A significant advantage of CRISPR-based gene editing is its specificity. Contrary to classical mutation methods, which result in random alterations scattered over the whole genome, or transgenic approaches, which in insects mostly rely on the random insertion of the construct by a transposase (O'Brochta and Atkinson, 1996; Handler and O'Brochta, 2011), gene editing enables the targeted modification of a single gene, without necessarily integrating exogenous DNA. Such modifications are not distinguishable from natural mutations or mutations induced by classical methods, e.g., radiation or ethyl methanesulfonate (EMS) mutagenesis. Nevertheless, depending on the country's regulatory framework, CRISPR-edited organisms may be regulated as genetically modified organisms (GMO), whereas classical mutagenesis is commonly not. Many governments attempt to integrate novel techniques like CRISPR/Cas within their existing regulatory frame for GMOs. However, these frameworks were designed to distinguish between conventional breeding techniques, involving hybridization and classical mutagenesis, and recombinant DNA technology, involving the transfer of DNA from sexually non-compatible species (transgenesis) (Sprink et al., 2016; Fladung, 2017), and are therefore not suited for a technique like CRISPR. Furthermore, the evaluation of options and risks posed by gene editing can be conducted from different perspectives, whereby regulations worldwide strongly differ and can be contradictory (Panjwani and Wilson, 2016). In principle, there are two main types of regulatory approaches. The *process-based*

approach, focusing on the techniques used to produce the GMO that currently applies in the EU, and the *product-based* regulations, focusing on the risks of new products and novel traits, rather than the method of production (e.g. used in the US) (Camacho et al., 2014; Sprink et al., 2016; Wolt and Wolf, 2018). Additionally, there are many ‘middle ground’ options (e.g. in Australia, see below), and case-by-case decisions (Eckerstorfer et al., 2019; Mallapaty, 2019).

For Europe, the Court of Justice of the European Union (ECJ) ruled in July 2018 that CRISPR gene editing poses risks similar to older GM methods, and that edited crops will be assessed as GMO and be subject to the same stringent regulation according to the 2001 directive (Directive 2001/18/EC, Council Directive 90/220/EEC). The ruling contradicted an earlier opinion of the advocate-general of the Court of Justice, Michael Bobek. He had suggested an exemption for edited organisms that do not contain artificial or foreign DNA, based on the fact that GMO created via classical mutagenesis techniques are also not included in the directive, according to the ‘mutagenesis exemption’ (Advocate General’s Opinion in Case C-528/16, press release No 04/18). Furthermore, the ECJ’s decision raised grave concerns among scientists (Gene editing in legal limbo in Europe, 2017; Gene-edited plants cross European event horizon, 2018; Urnov et al., 2018). Notably, European GMO regulations do not apply for sterile organisms, as those are not considered organisms defined as ‘any biological entity capable of replication or of transferring genetic material’ (HCB Scientific Committee, 2017).

In the US, regulations depend on the responsible federal agency. The U.S. Department of Agriculture (USDA) decided not to regulate genome-edited crops without recombinant DNA, plant pest activity, or novel food safety attributes in the final product (Wolt and Wolf, 2018), and the first CRISPR-edited crops were cultivated and sold without USDA oversight in 2016 (Waltz, 2016a; Waltz, 2016b). In March 2018, genome-edited plants were excluded from regulatory oversight entirely. Regulation of GM animals, however, is under the purview of the U.S. Food and Drug Administration (FDA), which decided to apply the ‘premarket new animal drug’ regulatory evaluation for any kind of intentional genomic alteration in food animals, irrespective of the resulting product (Van Eenennaam, 2018; Van Eenennaam et al., 2019). Which federal agencies might oversee the regulation of GM insects depends on the species and the application; Different departments might be involved in the regulatory process, possibly creating complex regulatory issues due to split responsibilities (Eckerstorfer et al., 2019).

In Australia, all CRISPR gene editing technologies were governed by the same rules as conventional genetic modifications and required approval from a biosafety committee accredited by the Office of the Gene Technology Regulator (OGTR). In October 2019, Australian regulators ruled that SDN-1 edits (made without a repair template) do not differ from naturally occurring changes and are not an additional risk to the environment and human health (Mallapaty, 2019). Thus, CRISPR/Cas9 mediated knock-outs in plants, animals, and human cell lines will not be considered GMOs in Australia. However, alterations relying on a template or mediating foreign genetic material’s insertion are still regulated as GMOs.

2.6 Research objectives

My **first objective** was to establish CRISPR/Cas9 HDR-mediated gene editing in *C. capitata*, because only the NHEJ pathway had been established in Tephritids so far. A strategy for highly efficient targeted mutagenesis without the introduction of exogenous DNA was needed to facilitate the possible field use of the resulting strains in SIT programs (see 3.1). The resulting toolset should then be used to develop CRISPR/Cas9-based strains to improve SIT programs and enable the transfer of SIT-systems between species.

A conditional female-to-male conversion system could surpass currently used medfly GSS in terms of efficiency and rearing costs. Thus, my **second objective** was to use CRISPR/Cas9 HDR to generate specific mutations in the sex determination gene *transformer-2* to achieve conditional (temperature-dependent) sex conversion in medfly (3.2). The known mutations should be generated by CRISPR/Cas9 HDR without using any transgenes.

Next, CRISPR should be used for a major endeavor in SIT-related research, namely the identification of the molecular genetic basis of the medfly GSS, that could pave the way for an ‘generic approach’ to construct new or improved GSS in several novel species. Therefore, my **third objective** was to find the *white pupae* gene in medfly and to use CRISPR/Cas9 to verify its role by functional knock-outs and complementation of the original mutation, to subsequently establish new white pupae strains (3.3). Simultaneously, the knowledge of the causal mutation and the genetic structure was used to develop safeguarding assays to distinguish GSS and WT flies in the field. Such a safeguard tool is missing for medfly GSS and would allow tracing of possible contaminations or fertile escapers on a genomic level.

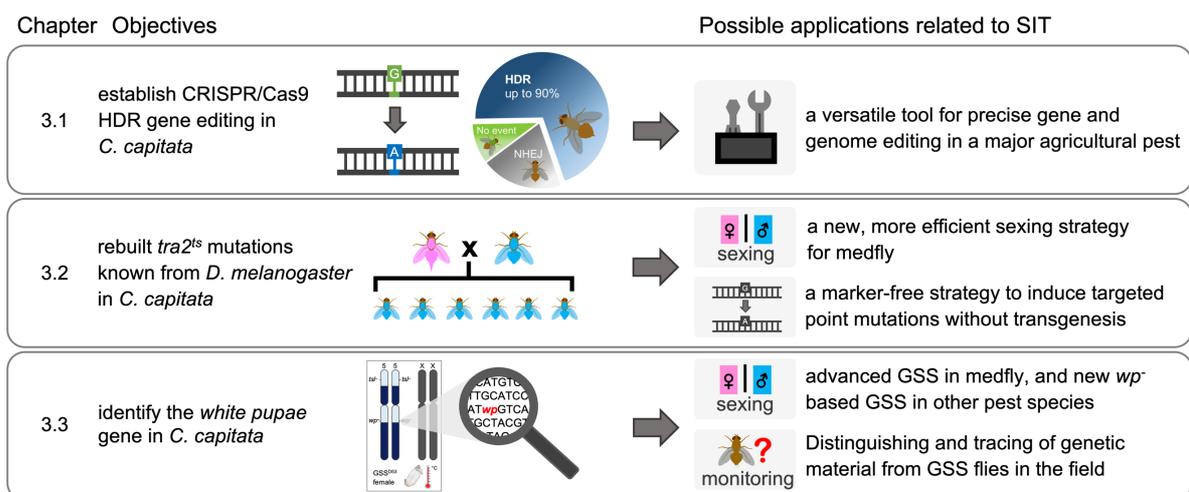


Fig. 4. Objectives of this thesis and their relevance to SIT

3 Results

This section includes three published manuscripts and provides information on:

- the main aim, results of each manuscript, and how it is embedded in the thesis
- the title, authors (* = co-first), journal and author contributions
- related presentations / talks / scientific outreach connected to each project

- 3.1 **Aumann RA**, Schetelig MF, Häcker I (2018) Highly efficient genome editing by homology-directed repair using Cas9 protein in *Ceratitis capitata*. *Insect Biochemistry and Molecular Biology* 101, 85-93.
<https://doi.org/10.1016/j.ibmb.2018.08.004>
- 3.2 **Aumann RA**, Häcker I, Schetelig MF (2020) Female-to-male sex conversion in *Ceratitis capitata* by CRISPR/Cas9 HDR-induced point mutations in the sex determination gene *transformer-2*. *Scientific Reports* 10, 18611.
<https://doi.org/10.1038/s41598-020-75572-x>
- 3.3 Ward CM*, **Aumann RA***, Whitehead MA, Nikolouli K, Leveque G, Gouvi G, Fung E, Reiling SJ, Djambazian H, Hughes MA, Whiteford S, Caceres-Barrios C, Nguyen TNM, Choo A, Crisp P, Sim SB, Geib SM, Marec F, Häcker I, Ragoussis J, Darby AC, Bourtzis K, Baxter SW, Schetelig MF (2021) White pupae phenotype of tephritids is caused by parallel mutations of a MFS transporter. *Nature Communications* 12, 491.
<https://doi.org/10.1038/s41467-020-20680-5>

3.1 Expanding the toolbox: CRISPR/Cas9 HDR gene editing in *Ceratitis capitata*

The first aim was to establish CRISPR/Cas9 HDR gene editing in medfly and find the most efficient system for targeted nucleic acid exchanges. The conversion of the enhanced green fluorescent protein eGFP into a blue fluorescent protein (BFP) was used as a proof of principle experiment. Here, I achieved CRISPR/Cas9 HDR gene editing in *C. capitata* and the family of Tephritids for the first time, which additionally resulted in an excellent efficiency of up to 90% HDR events and gained important insights in the design strategy of gRNAs and short repair templates. These findings proved valuable for the subsequent projects relying on efficient CRISPR/Cas9 (HDR) technology.

Title: Highly efficient genome editing by homology-directed repair using Cas9 protein in *Ceratitis capitata*

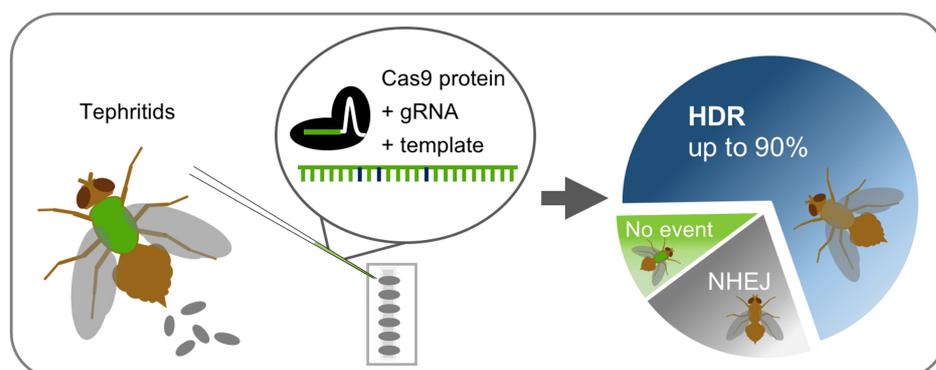
Authors: Aumann RA, Schetelig MF, Häcker I

Status: published in *Insect Biochemistry and Molecular Biology* (2018)

Contributions: all experiments were performed by Aumann RA

Presentations: This work was presented at the 3rd Research Coordination Meeting on ‘Comparing Rearing Efficiency and Competitiveness of Insect Sterile Male Strains Produced by Genetic, Transgenic or Symbiont-based Technologies’ of the FAO/IAEA in Bangkok, Thailand (2018, talk), the INSECTA conference in Gießen, Germany (2018, talk), and at the 11th GGL conference on Life Science in Gießen, Germany (2018, poster, awarded as best presentation).

Furthermore, I was teaching a course in the Master program Insect Biotechnology & Bioresources on gene editing tools (MP149, JLU Gießen), and discussed those novel technologies with a 9th grade school class on the scientific background of gene drives (in the framework of ‘Genomchirurgie im Diskurs – Wissenschaft im Dialog’; Georg-Christoph-Lichtenberg school Ober-Ramstadt, Germany).





Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmbHighly efficient genome editing by homology-directed repair using Cas9 protein in *Ceratitis capitata*Roswitha A. Aumann^a, Marc F. Schetelig^{a,b,*}, Irina Häcker^{a,b}^a Justus-Liebig-University Gießen, Institute for Insect Biotechnology, Department of Insect Biotechnology in Plant Protection, Winchesterstr. 2, 35394 Gießen, Germany^b Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Division of Bioresources, Department of Insect Pest and Vector Control, 35394 Gießen, Germany

ARTICLE INFO

Keywords:

ssODN-mediated homology-directed repair
CRISPR-Cas9 ribonucleoprotein complexes
Mediterranean fruit fly
Tephritids
Sterile insect technique

ABSTRACT

The Mediterranean fruit fly *Ceratitis capitata* is a highly polyphagous and invasive insect pest, causing enormous economic damage in horticultural systems. A successful and environment-friendly control strategy is the sterile insect technique (SIT) that reduces pest populations through infertile matings with mass-released, sterilized insects. However, the SIT is not readily applicable to each pest species. While transgenic approaches hold great promise to improve critical aspects of the SIT to transfer it to new species, they are suspect to strict or even prohibitive legislation regarding the release of genetically modified (GM) organisms. In contrast, specific mutations created via CRISPR-Cas genome editing are not regulated as GM in the US, and might thus allow creating optimal strains for SIT. Here, we describe highly efficient homology-directed repair genome editing in *C. capitata* by injecting pre-assembled CRISPR-Cas9 ribonucleoprotein complexes using different guide RNAs and a short single-stranded oligodeoxynucleotide donor to convert an enhanced green fluorescent protein in *C. capitata* into a blue fluorescent protein. Six out of seven fertile and individually backcrossed G₀ individuals generated 57–90% knock-in rate within their total offspring and 70–96% knock-in rate within their phenotypically mutant offspring. Based on the achieved efficiency, this approach could also be used to introduce mutations which do not produce a screenable phenotype and identify positive mutants with a reasonable workload. Furthermore, CRISPR-Cas HDR would allow to recreate mutations formerly identified in classical mutagenesis screens and to transfer them to related species to establish new (SIT-like) pest control systems. Considering the potential that CRISPR-induced alterations in organisms could be classified as non-GM in additional countries, such new strains could potentially be used for pest control applications without the need to struggle with GMO directives.

1. Introduction

Due to its ability to infest fruits of more than 250 plant species including some vegetables and nuts and to adapt to a wide range of temperate and tropical habitats, the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann; Diptera: Tephritidae) (Medfly) has become one of the most successful and economically important invasive insect pests worldwide. A successful control strategy for the Medfly as part of integrated pest management programs (IPM) is the sterile insect technique (SIT) (Knipling, 1955). The SIT is an area-wide environmental-friendly and species-specific approach. It is based on the mass release of sterilized male insects into the wild-type (WT) population, leading to

infertile matings and thereby to a decrease of the progeny. Repeated releases thus allow for the suppression of a pest population to an economically uncritical size or to prevent the infestation of new areas by preventative releases.

Although the SIT has been applied very successfully for some species (Gilles et al., 2014; Hendrichs et al., 1995; Vreysen, 2001), there are still several key aspects that could be improved to optimize existing SIT programs or that need to be developed to apply the SIT to new species. One of them is the generation of male-only populations for release, a process that is also called *sexing*. Male-only releases are more effective than bisexual releases (Rendon et al., 2004) as they prevent the mating of the sterile males with the co-released sterile females, thus

Abbreviations: BFP, blue fluorescent protein; bp, base pair(s); Cas, CRISPR associated; CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats; DSB, double-strand break; eGFP, enhanced green fluorescent protein; GMO, genetically modified organism; gRNA, guide RNA; HDR, homology directed repair; IPM, integrated pest management; NHEJ, non-homologous end-joining; PAM, protospacer adjacent motif; SIT, sterile insect technique; ssODN, single-stranded oligodeoxynucleotide; YFP, yellow fluorescent protein

* Corresponding author. Justus-Liebig-University Gießen, Institute for Insect Biotechnology, Department of Insect Biotechnology in Plant Protection, Winchesterstr. 2, 35394 Gießen, Germany.

E-mail address: marc.schetelig@agrar.uni-giessen.de (M.F. Schetelig).

<https://doi.org/10.1016/j.ibmb.2018.08.004>

Received 3 June 2018; Received in revised form 21 August 2018; Accepted 24 August 2018

Available online 26 August 2018

0965-1748/© 2018 Elsevier Ltd. All rights reserved.

increasing the chance of matings with the WT females. Moreover, the release of sterile females could still result in damage to the fruits and crops by oviposition, even if the eggs will not develop due to the sterilization step. Current Medfly SIT programs enable the mass production of billions of male flies per week by using an automated sexing system during the mass rearing process (Augustinos et al., 2017). A few such sexing systems were created by chemical or irradiation-based classical mutagenesis in the past 40 years, for example in *Anopheles albimanus* (Wiedemann) (Kaiser et al., 1978), *Ceratitidis capitata* (Franz and McInnis, 1995), *Zeugodacus cucurbitae* (Coquillett) (McInnis et al., 2004), *An. arabiensis* (Patton) (Yamada et al., 2012), *Bactrocera dorsalis* (Hendel) (Isasawin et al., 2012), *Anastrepha ludens* (Loew) (Orozco et al., 2013), and *Bactrocera carambolae* (Drew and Hancock) (Isasawin et al., 2014). Their transfer to new pest insects, however, is difficult and time-consuming, since classical mutagenesis results in random mutation products (Bose, 2016). Therefore, the mutations responsible for the effect are often not identified (Busch-Petersen, 1990; Rössler, 1979). As an alternative, transgenic sexing systems were developed based on the expression of lethal genes in female embryos (Schetelig and Handler, 2012) or the accumulation of a toxic protein later in development (Fu et al., 2007). These transgenic sexing systems have been quickly adapted to a number of pest species within just a few years (Ant et al., 2012; Concha et al., 2016; Li et al., 2014; Ogaugwu et al., 2013; Schetelig and Handler, 2012; Schetelig et al., 2016; Yan and Scott, 2015). However, to date, the release of transgenic organisms is highly regulated or even prohibited in many countries, while the use of strains created by classical mutagenesis is not restricted. Therefore, a tool that can create efficient and safe sexing systems that are acceptable for release similar to classical mutagenesis would be highly beneficial. CRISPR-Cas has the potential to become such a tool, as small sequence changes (InDels or SNPs) created by CRISPR-Cas that could also occur naturally are already classified as non-GM in the US and might be regulated similarly in other countries.

CRISPR-Cas enables the editing of genes in two ways following the introduction of a double-strand break (DSB) (Doudna and Charpentier, 2014; Ran et al., 2013; Zetsche et al., 2015), either by the non-homologous end-joining (NHEJ) or the homology-directed repair (HDR) pathway. While the NHEJ pathway, in simplified terms, is a somewhat ‘error-prone’ pathway, causing random insertions or deletions of nucleotides at the target site, the HDR pathway can be exploited to precisely manipulate the target sequence by providing a suitable DNA repair template including the desired alteration (Kim and Kim, 2014). This allows the introduction of specific sequence changes without leaving exogenous DNA sequences in the genome. Therefore, once established in a new pest species, CRISPR-Cas HDR could be the long-awaited tool to overcome the disadvantages of existing methods: classical mutagenesis with its low yield and its potential for additional unknown changes in the genome on one side, and transgenic methods with the introduction of exogenous genetic material and the regulatory restrictions on the other side.

The NHEJ and HDR DNA repair pathways are competing events within a cell. The balance between the two pathways differs widely among species, between different cell types within a species, as well as during different cell cycle phases of a single cell (Shrivastav et al., 2008). Specific genome alterations via CRISPR-Cas HDR can be achieved in a highly efficient manner by shifting the equilibrium towards the less efficient HDR pathway (Shrivastav et al., 2008). Improving the efficiency of HDR was explored by the inhibition of key enzymes of the NHEJ pathway like the DNA ligase IV using the inhibitor Scr7 (Hu et al., 2018; Maruyama et al., 2015) or by the controlled timing of Cas9 delivery according to cell-cycle dynamics (Gutschner et al., 2016; Lin et al., 2014). Other important aspects for a precise HDR event are the prevention of re-editing of the already modified locus. For example by introducing mutations in the protospacer adjacent motif (PAM) sequence or the guide RNA (gRNA) target site of the repair template, as well as considering the influence of the distance between

the DSB site and the desired mutation position on the mutagenesis efficiency (Kwart et al., 2017; Paquet et al., 2016).

To determine the efficiency of such HDR-improving methods, it is helpful to simultaneously quantify HDR and NHEJ events. This can be done for example by targeting an enhanced green fluorescent marker protein (eGFP) and converting it into the blue fluorescent protein (BFP) (Glaser et al., 2016) by introducing two single base substitutions in the chromophore of eGFP (Glaser et al., 2016; Heim et al., 1994). In this experimental setup, green fluorescence shows the absence of a CRISPR mutation event, blue fluorescence indicates an HDR event, and the loss of fluorescence represents unspecific mutation events caused by NHEJ repair. So far, in Medfly, mutant phenotypes could only be generated by the NHEJ pathway after CRISPR-Cas9-based gene disruption (Meccariello et al., 2017).

Here, we used the eGFP-to-BFP conversion approach to establish highly efficient CRISPR-Cas HDR genome editing in *C. capitata*, knocking-in a short single-stranded oligodeoxynucleotide (ssODN) repair template by injecting *in vitro* preassembled Cas9-gRNA ribonucleoprotein complexes and a single-stranded oligo donor into transgenic *C. capitata* embryos carrying an eGFP marker.

2. Material and methods

2.1. Fly rearing

The *Ceratitidis capitata* transgenic target line *TREhs43^{Ala5}_F1m2* flies (Schetelig et al., 2009) and wild-type *Egypt-II* flies (*EgII*, obtained from the FAO/IAEA Agriculture and Biotechnology Laboratory, Seibersdorf, Austria) were maintained in a controlled environment (26 °C, 48% RH, and a 14:10 light/dark cycle) and fed with a 3:1 (v/v) mixture of sugar and yeast extract, and water. Larvae were reared on a gel diet, containing carrot powder (120 g/l), agar (3 g/l), yeast extract (42 g/l), benzoic acid (4 g/l), HCl (25%, 5.75 ml/l) and Ethyl-4-hydroxybenzoate (2.86 g/l). Larvae and flies from injected embryos were reared under the same conditions. *TREhs43^{Ala5}_F1m2* flies used for CRISPR gene editing carry an eGFP marker under the control of the *D. melanogaster polyubiquitin* promoter (Schetelig et al., 2009). The eGFP marker gene is expressed in the head, thorax, and legs of the adult fly. Flies were anesthetized with CO₂ for screening, sexing, and the setup of backcrosses.

2.2. CRISPR-Cas9 reagents

Purified Cas9 protein was obtained from PNA Bio Inc (catalog number CP01). The lyophilized protein pellet was reconstituted to a stock concentration of 1 µg/µl in 20 mM Hepes, 150 mM KCl, 2% sucrose and 1 mM DTT (pH 7.5) by adding 25 µl nuclease free H₂O and stored at –80 °C until use.

Linear double-stranded DNA templates for the gRNAs were produced by a template-free PCR reaction with two partially overlapping oligos, containing 20 µl 5 × Q5 reaction buffer, 10 µl dNTP Mix (2 mM each), 5 µl of each primer (10 µM) and 1 µl Q5 HF polymerase (2U) (New England Biolabs, NEB) in a total volume of 100 µl. PCR reactions were run in a Bio-Rad C1000 Touch thermal cycler [98 °C, 30 s; 35 cycles of (98 °C, 10 s; 58 °C, 20 s; 72 °C, 20 s); 72 °C, 2 min] (Kalajdzic and Schetelig, 2017). For the synthesis of the guide RNA eGFP_gRNA2b primers P_986 (GAAATTAATACGACTCACTATAGGCTCGTGACCACCC TGACCTAGTTTTAGAGCTAGAAATAGC) and P_369 (GCACCGACTCGG TGCCACTTTTTCAAGTTGATAACGGACTAGCCTATTTTAACTTGCTATTCTAGCTCTAAAAC) were used, for the synthesis of eGFP_gRNA2b primers P_1172 (GAAATTAATACGACTCACTATAGGCTGAAGCACTGC ACGCCGTGTTTTAGAGCTAGAAATAGC) and P_369 were used. The forward primers (P_986, P_1172) encode the T7 polymerase-binding site followed by the specific gRNA target sequence and ending with the 20 nt complementary sequence that allows forward and reverse primers to anneal. Reverse primer (P_369) is a universal oligonucleotide that

can be used for all gRNAs encoding the Cas9 interacting portion of the gRNA sequence (Kalajdzic and Schetelig, 2017). The specific gRNA target sequence of gRNAs eGFP_gRNA2 and eGFP_gRNA2b was previously described (Glaser et al., 2016). Size verification was carried out using 2 µl of the reaction while the remaining 98 µl were purified using a PCR purification kit (DNA Clean & Concentrator™-25; Zymo Research, Irvine, CA, USA) and eluted in 30 µl elution buffer. Purity and concentration of the gRNA templates were measured with a spectrophotometer (BioTek Epoch2 microplate reader). gRNA *in vitro* transcription was performed with the HiScribe™ T7 High Yield RNA Synthesis Kit (NEB), using 500 ng purified DNA template for 16 h (overnight) at 37 °C according to the manufacturer's instructions. RNA samples were treated with TURBODNase (Ambion, Oberursel, Germany) to remove possible DNA contamination, and purified using the MEGAclean purification kit (Ambion) as described by (Kalajdzic and Schetelig, 2017). Purified gRNAs were aliquoted and stored at –80 °C until use.

The 140 bp single-stranded HDR template 'ssODN_BFP' (single-stranded oligodeoxynucleotide blue fluorescent protein; P_1000_G/BFP_ssODN_Glaser) to convert eGFP into BFP was described previously (Glaser et al., 2016) and synthesized by Eurofins Genomics (EXTREMER oligo, purified salt-free, quality control by CGE). It differs from the eGFP sequence by 3 bases (194C > G, 196T > C, 201C > G), whereby the first change (194C > G; Thr65 > Ser65) causes a reversion of eGFP back to wild-type GFP, the second (196T > C; Tyr66 > His66) converts GFP to BFP. The third SNP (201C > G) is a silent mutation, to further reduce the target sequence similarity after HDR (Glaser et al., 2016). The sequence of ssODN_BFP was verified by sequencing (Macrogen Europe, Amsterdam), after performing PCR using Platinum Taq polymerase (Invitrogen), using primers P_1160 (GGCATGGCGGACTTG) and P_1001 (CCTGAAGTTCATCTGCACCACC) in a Bio-Rad C1000 Touch thermal cycler [95 °C, 2 min; 35 cycles of (95 °C, 30 s; 50.5 °C, 30 s; 72 °C, 20 s); 72 °C, 2 min]. The PCR reaction contained 10 µl 10× Platinum PCR Buffer (-Mg), 1 µl MgSO₄ 50 mM, 1 µl dNTP Mix (2 mM each), 1 µl of each primer (10 µM), 0.2 µl Platinum Taq polymerase and 440 ng DNA template in a total volume of 20 µl.

2.3. Preparation of CRISPR injection mix

Injection mixes for microinjection of embryos contained 360 ng/µl Cas9 protein, 200 ng/µl gRNA_eGFP2 or gRNA_eGFP2b and 200 ng/µl ssODN_BFP in a 10 µl volume containing an end-concentration of 300 mM KCl, according to previous studies (Burger et al., 2016; Kistler et al., 2015; Meccariello et al., 2017). To inhibit NHEJ, 1 mM Scr7 (Xcess biosciences Inc., catalog number M60082-2, CAS 1533426-72-0) was added to the injection mix in one experiment. All mixes were freshly prepared on ice followed by an incubation step for 10 min at 37 °C to allow pre-assembly of gRNA-Cas9 ribonucleoprotein complexes and stored on ice prior to injections. To determine the stability of the gRNAs in the injection mixes after pre-assembly, we verified the presence of the gRNAs via gel electrophoresis (Suppl. Fig. 1).

2.4. Microinjection of embryos

For microinjection of homozygous *C. capitata* TREhs43hid^{Ala5}_F1m2 embryos eggs were collected over a 45–90 min period. Eggs were prepared for injection as previously described (Handler and Anthony, 2000) using chemical dechorionization (sodium hypochlorite, 3 min). In brief, embryos were fixed using double-sided sticky tape (Scotch 3M Double Sided Tape 665) onto a microscope slide and covered with halocarbon oil 700 (Sigma-Aldrich, Munich, Germany). Injections were performed using borosilicate needles (GB100F-10 with filaments; Science Products, Hofheim, Germany), drawn out on a Sutter P-2000 laser-based micropipette puller. The injection station consisted of a manual micromanipulator (MN-151, Narishige), an Eppendorf femtoJet 4i

microinjector, and an Olympus SZX2-TTR microscope (SDF PLAPO 1×PF objective). The microscope slide with the injected embryos was placed in a Petri dish containing moist tissue paper in an oxygen chamber (max. 2 psi) and stored at 25 °C, 60% RH for 72 h to allow larval hatching. Hatched first instar larvae were transferred from the oil to larval food.

2.5. Crossing and screening

Each G₀ adult survivor was individually crossed to three *EgII* WT males or female virgins. Eggs were collected three times, with an interval of one to two days. Both G₀ and G₁ flies were screened for eGFP fluorescence phenotype to detect CRISPR mutagenesis events, G₁ flies additionally were screened for BFP fluorescence.

2.6. Genomic DNA extraction, PCR and sequencing

Genomic DNA was extracted from single G₁ flies according to standard protocols. The DNA was used as a template to amplify the region surrounding the gRNA target sites. PCR was performed in a 50 µl reaction volume using DreamTaq polymerase (Life Technologies) according to the manufacturer's protocol, the primers P_145 (ACTTAATCGCCTTGCAGCACATCC) and P_55 (TGTGATCGCGCTTCTCGTT), or P_145 and P_176 (AGGCCACCTATTCGTCTTCC), and 150–250 ng template in a Bio-Rad C1000 Touch thermal cycler [95 °C, 3 min; 35 cycles of (95 °C, 30 s; 58 °C, 30 s; 72 °C, 1 min); 72 °C, 5 min]. Primer combination P_145/P_55 is spanning the first 630 bp of the 720 bp eGFP coding sequence, primer combination P_145/P_176 is spanning the whole eGFP sequence (P_145 is located in the transgene construct backbone between the P_UB promoter and eGFP, P_176 is located in the 3' pBac sequence of the transgene construct). The size of the PCR product was verified by running an aliquot of the reaction on an agarose gel. The remaining PCR product was purified using a PCR purification kit (DNA Clean & Concentrator™-25; Zymo Research). All PCR products were verified by sequencing (Macrogen Europe, Amsterdam; with Primer P_145).

2.7. gRNA off-target assessment and verification of CRISPR-HDR mutations

Assessment of potential off-target effects of eGFP_gRNA2 and eGFP_gRNA2b was performed using the "Find CRISPR sites" tool of the Geneious Software Package 10.2.2 (Kearse et al., 2012), using the *C. ceratitis* genome (Papanicolaou et al., 2016) assembly version Ccap_1.1 (GCF_000347755.2_Ccap_1.1_genomic.fna.gz) from NCBI as the off-target database. Verification of CRISPR-induced mutations in the sequencing results was performed by mapping the sequencing results of G₁ individuals to the eGFP reference sequence (Schetelig et al., 2009) using Geneious.

2.8. Equipment and settings for screening and image acquisition

Screening of transgenic flies was performed using a Leica M165 FC stereo microscope with the PLAN 0.8× LWD objective and the following epifluorescence filters: GFP-LP (Excitation 425/60 nm; barrier 480 LP nm), YFP (excitation 510/20 nm; barrier 560/40 nm) or ET DAPI BP (excitation 395/25 nm; barrier 460/50 nm). For bright field and fluorescent image acquisition of living flies, flies were anesthetized with CO₂ and placed on a 4 °C cooler. Images were taken with a fully automated Leica M205FC stereo microscope with a PLANAPO 1.0× objective and a 1× Leica DFC7000 T camera using the Leica LAS X software. In order to enhance screen and print display of the pictures the image processing software Adobe Photoshop CS5.1 was used to apply moderate changes to image brightness and contrast. Changes were applied equally across the entire image and for all images.

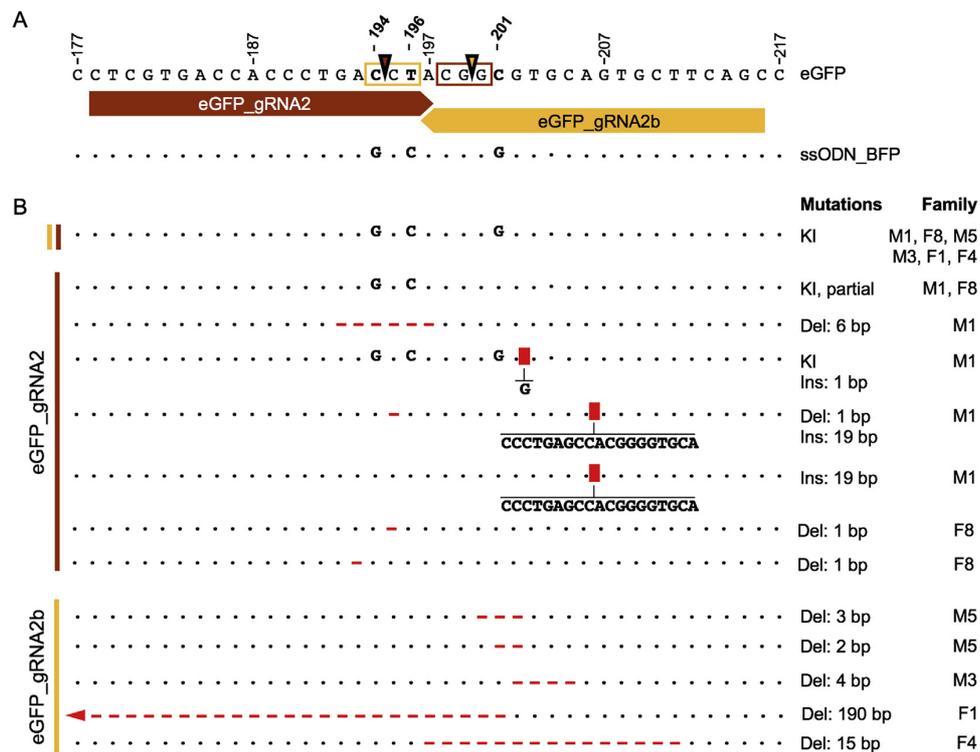


Fig. 1. Position of gRNAs, protospacer adjacent motifs (PAM), double strand breaks (DSB) and single nucleotide polymorphisms (SNPs) within the eGFP target sequence. **A)** Relative to the eGFP sequence the eGFP_gRNA2 (red) is sense- and the eGFP_gRNA2b (yellow) is anti-sense-oriented. PAM sequences are highlighted within the eGFP sequence, DSB sites are indicated by triangles. Related gRNA, PAM, and DSB site match in color. The ssODN_BFP sequence differs from the eGFP sequence in three positions shown as uppercase letters (194C > G, 196T > C, 201C > G), the consensus is shown as dots. **B)** Sequences of mutant eGFP alleles identified in G₁ individuals compared to the eGFP reference sequence. Consensus is shown as dots, knock-in (KI) mutant sites in uppercase letters, deletions (Del) as red lines, insertion sites (Ins) as red rectangles. Families that carried the respective mutation(s) are indicated on the right. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3. Results

3.1. Selection of gRNAs for eGFP mutagenesis in *C. capitata* and off-target analysis

Two previously evaluated guide RNAs against eGFP (Glaser et al., 2016), which were named eGFP_gRNA2 and eGFP_gRNA2b, were used to direct the Cas9 endonuclease activity towards the chromophore region of the eGFP fluorescent marker gene in the transgenic *C. capitata* strain *TREhs43hid^{Ala5}_F1m2* (Schetelig et al., 2009). The gRNAs target the same region, therefore one HDR repair template (single-stranded oligodeoxynucleotide blue fluorescent protein, ssODN_BFP (Glaser et al., 2016)) was used for both. They differ, however, in their orientation and cleavage site, as well as in the number of mismatches to their target sequence resulting from successful HDR (Fig. 1A) and their off-target activity.

In silico target site analysis in Geneious predicted an on-target activity score of 0.272 and zero off-target sites in the medfly genome (100% off-target score) for the eGFP_gRNA2. On-target activity scores are reported on a scale from 0 (no on-target activity) to 1 (highest on-target activity) (Doench et al., 2014). eGFP_gRNA2b has an on-target activity score of 0.329 but two off-target sites (98.94% off-target score: #1 score 4.23%; location NW_004524467.1 4,259,338 < 4,259,360; #2 score 1.13%; location NW_004523691.1 10, 017, 309 < 10, 017, 331; Ccap 1.1). Both off-target sites of eGFP_gRNA2b show four mismatches to the reference genome sequence. Importantly, none of the off-target sites is located in a coding sequence of *C. capitata* genome.

The repair template, ssODN_BFP, differs from the eGFP sequence by

three bases (194C > G, 196T > C, 201C > G; Fig. 1A), whereby the first change (194C > G; Thr65 > Ser65) causes a reversion of eGFP back to wild-type GFP and the second (196T > C; Tyr66 > His66) converts GFP to BFP (Glaser et al., 2016). The third SNP (201C > G) is a silent mutation to further reduce the gRNA-target sequence similarity after HDR and thus prevent re-editing of the target sequence (Glaser et al., 2016) (Fig. 1A).

3.2. CRISPR-Cas9 HDR mutagenesis in *Medfly*

Three different injections were conducted to establish CRISPR-Cas9 HDR in *Medfly*. Each G₀ adult survivor was screened for eGFP fluorescence to confirm the presence of the CRISPR target site, and individually backcrossed. Their offspring (G₁) were screened for eGFP and BFP fluorescence.

In the first experiment, the eGFP_gRNA2 was injected with recombinant Cas9 protein and the ssODN_BFP donor template into 243 embryos of the strain *TREhs43hid^{Ala5}_F1m2*, homozygous for eGFP. 16 reached the larval stage of which eight survived to adulthood (Table 1). These were individually backcrossed to *EgII* wild-type virgin females and males, respectively. Eggs of these crosses were collected three times, at an interval of one to two days. Three crosses (M1, F2, F8) were fertile and two out of these three families produced phenotypically WT offspring missing the eGFP marker fluorescence (Fig. 2). This effect was observed in 98 out of 116 flies (84%) in family M1, and 34 out of 42 flies (81%) in family F8 (Fig. 3A, D, and Suppl. Table 1). The loss of the eGFP fluorescence was interpreted as a positive CRISPR event (insertion/deletion or knock-in event at the target site). Blue fluorescence

Table 1

Overview of injections and resulting gene editing events. Shown is the number of injected eggs, number of surviving larvae, number of total and fertile G_0 adults, number of fertile adults which produced offspring phenotypically missing the eGFP marker, number of total and mutant (knock out-phenotype (KO)) G_1 , the average percentage of the KO phenotype within the total offspring, and the average percentage of the knock-in genotype (two or three base pairs altered) within the phenotypically mutant offspring. *For screening results of individual families see Suppl. Tables 1, 2, and 3.

	eGFP_gRNA2	eGFP_gRNA2 + Scr7	eGFP_gRNA2b
injected eggs	243	323	371
surviving larvae	16	79	19
total G_0 adults (male/female)	8 (3♂/5♀)	31 (17♂/14♀)	9 (5♂/4♀)
fertile G_0 adults (male/female)	3 (1♂/2♀)	27 (15♂/12♀)	4 (2♂/2♀)
G_0 crosses producing mutant offspring (family)	2 (M1, F8)	0	4 (M3, M5, F1, F4)
# G_1 total	219	1967	208
# G_1 mutant*	132	0	174
KO phenotype in G_1	60%	0%	84%
# knock-in genotype in G_1	94	0	138
knock-in genotype in G_1 (within eGFP-KO phenotype)	71%	0%	79%

was not observed in any of the G_1 flies.

In a second, independent experiment, 323 embryos of the target line were injected with eGFP_gRNA2 and 1 mM Scr7 additionally added to the injection mix. Scr7 is an inhibitor of Ligase IV, a key enzyme in the NHEJ repair pathway (Srivastava et al., 2012). In previous studies employing zinc finger nucleases in *D. melanogaster*, deficiency of Ligase IV resulted in increased HDR rates (Beumer et al., 2008). 79 of the injected embryos reached the larval stage with 31 surviving to adulthood (Table 1). These were backcrossed individually and eggs collected from 27 fertile crosses as described previously. In total, 1967 G_1 offspring were screened for the loss of eGFP fluorescence. However, none of the families produced offspring phenotypically missing the eGFP fluorescence (Suppl. Table 2). To exclude the possibility of gRNA degradation in injection mix containing Scr7, the stability of eGFP-gRNA2 in this mix was verified via gel electrophoresis (Suppl. Fig. 1).

In the third experiment, the injection of eGFP_gRNA2b-Cas9

complexes together with ssODN_BFP donor template yielded nine adult flies from 371 injected embryos (Table 1). Four of the nine individual crosses (M3, M5, F1, F4) were fertile, and all four produced mainly offspring phenotypically lacking the eGFP fluorescence (79%–100%), indicating a CRISPR-induced mutation (Fig. 3G, J, M, P, Suppl. Table 1). Again, none of the G_1 flies showed blue fluorescence. To determine if BFP is visible in flies homozygous for the mutation, G_1 individuals with knock-out (KO) phenotype from the eGFP_gRNA2b injection were inbred, and G_2 was screened for BFP fluorescence. From eight separate G_1 inbreeding crosses a total of 1982 G_2 offspring were screened. None of them showed BFP fluorescence, although two of the crosses (M3 (T3) and M5 (T3)) contained only individuals with the knock-in genotype, and thus should produce 25% offspring homozygous for the knock-in mutation (Suppl. Table 3).

3.3. Molecular verification of HDR or other mutagenic events

The genotype of phenotypically WT G_1 flies was analyzed via eGFP-specific PCR with primers P_145 and P_55 spanning the first 630 bp of the 720 bp eGFP sequence, and subsequent sequencing of the PCR product. All G_1 flies were analyzed except for four individuals in family M1 and two in family F8, as no PCR product and consequently, no sequence information could be obtained for these flies, probably due to genomic DNA degradation in the dead flies.

Sequencing of DNA amplicons from individuals of eGFP_gRNA2 injection revealed that 18 out of 94 phenotypically WT M1 offspring (19%) carried the complete knock-in genotype (three base pairs exchanged) and 50 (53%) carried a shorter version of the knock-in with only two of the three base pairs altered (194C > G, 196T > C; Figs. 1B and 3B, Suppl. Table 1). These two base changes are sufficient for a loss of the eGFP fluorescence, as the third base change, 201C > G, is only a silent mutation. Insertions or deletions caused the knock-out phenotype of the remaining 26 flies (28%) in the target region (four different mutation events; Figs. 1B and 3B, Suppl. Table 1). In case of family F8, sequencing showed that 17 out of 32 phenotypically WT flies carried the complete knock-in genotype (three bp HDR) (53%) and nine (28%) carried the shorter version of the knock-in. Moreover, two different deletion events caused by NHEJ repair were observed in six flies (19%) (Figs. 1B and 3E, and Suppl. Table 1). No other (point) mutations beyond the DSB region could be identified within the sequenced 630 bp in any of the G_1 individuals. For two M1 individuals (M1m5, M1f7) and four F8 individuals (F8f14, F8f15, F8f16, F8m18) additionally the

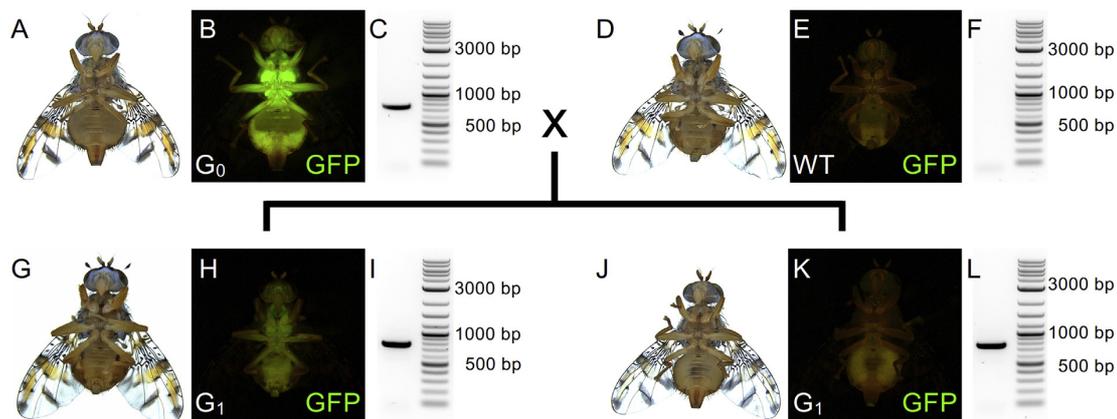


Fig. 2. Crossing scheme of G_0 individuals and molecular analysis of G_0 and G_1 flies. Shown are fly images in bright field (A, D, G, J) and corresponding eGFP fluorescence (B, E, H, K) as well as the respective PCR validating the presence or absence of the eGFP marker gene (C, F, I, L). The *TREhs43hid^{Alu5}F1m2* G_0 individuals, homozygous for the eGFP marker gene, injected with Cas9 and eGFP_gRNA2 or -2b, were individually crossed to WT *EgII* flies. G_1 offspring was either heterozygous for the eGFP marker (H) and positive in eGFP-specific PCR (I), or phenotypically missing the eGFP fluorescence (K), but still carrying the eGFP marker gene (L), which indicates a CRISPR-induced mutagenesis. DNA ladder used for agarose gels is the NEB 2log DNA-ladder; bp = base pair.

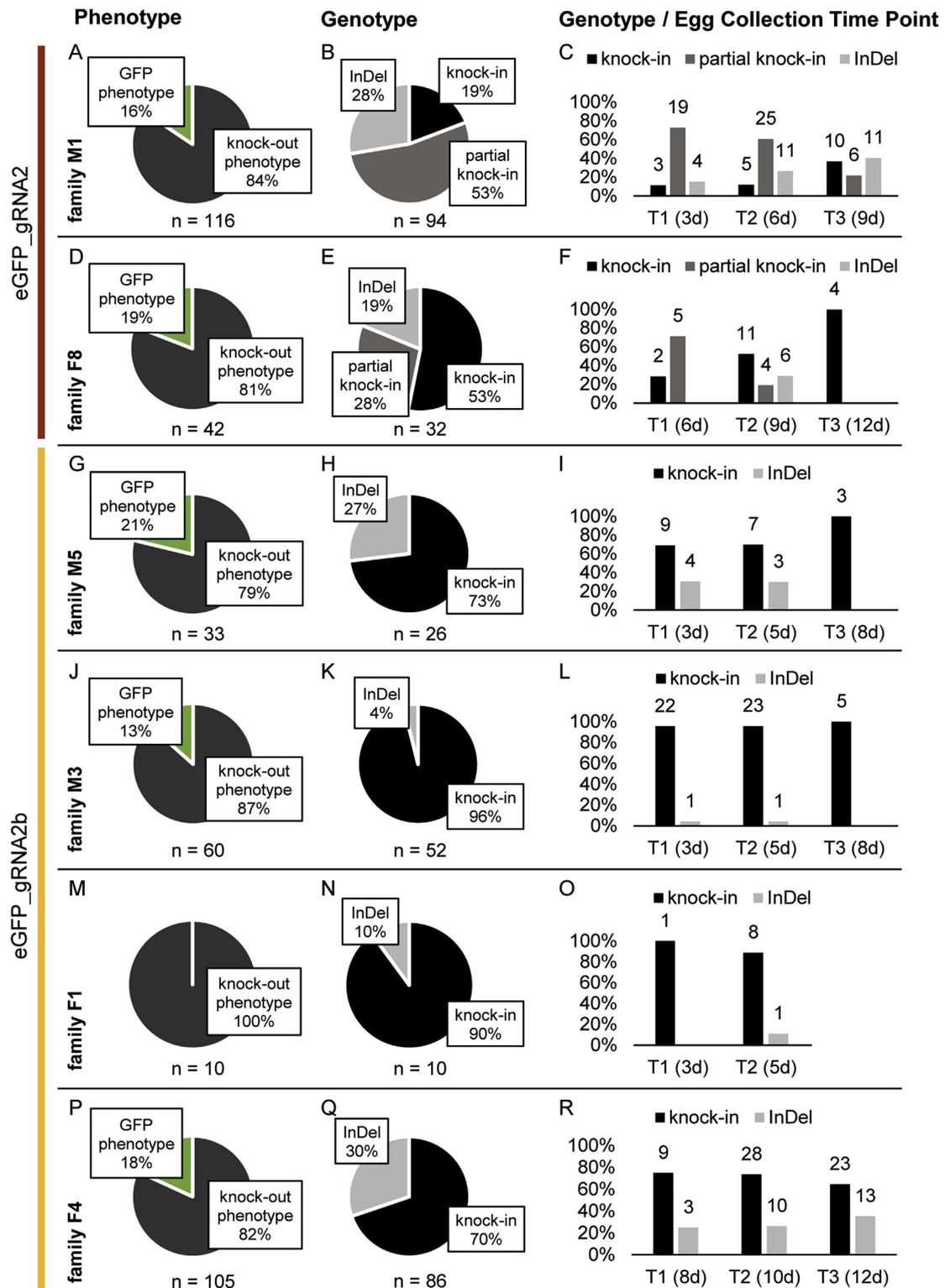


Fig. 3. Frequency of CRISPR-Cas-induced G₁ phenotypes and genotypes. Families M1 and F8 were injected with eGFP_gRNA2 (A-F), families M5, M3, F1 and F4 with eGFP_gRNA2b (G-R). In the first column, the absolute number of offspring per family and the percentage of phenotypes “GFP” (heterozygous) and “knock-out” (eGFP phenotypically missing) are shown (A, D, G, J, M, P). The second column shows the number of sequenced knock-out individuals and the percentage of different mutation types (knock-in, partial knock-in or insertion/deletion (InDel); B, E, H, K, N, Q). The third column shows the occurrence of the different mutation types per egg collection (T1, T2, T3) (C, F, I, L, O, R). Numbers above bars indicate absolute number of individuals per mutation per time point.

whole eGFP gene was sequenced (primers P_145/P_176) to exclude a mutation in the last 90 bp of the eGFP coding sequence that could result in a non-functional eGFP protein. Again, no sequence changes except for the three (or two) intended point mutations of the knock-in were found in any of the six sequenced individuals (data not shown), proving that the WT phenotype is indeed due to the knock-in genotype.

Interestingly, the two different HDR events in G₁ (complete three bp HDR versus two bp HDR) were not evenly distributed over the three egg collection time points (T1, T2, T3). In both families, the percentage of the complete HDR increased over time, whereas the rate of the partial HDR decreased. In the M1 family, 73% of the offspring from the first egg collection (T1) carried the partial knock-in, whereas only 22% of the offspring from the last egg collection (T3) carried this genotype. The complete knock-in was observed in 11.5% of the T1-offspring and 37% of the T3-offspring of M1. In the F8 family, the partial knock-in decreased from 71.4% in T1 to none in T3. In contrast, the complete knock-in increased from 28.5% (T1) to 100% (T3) (Fig. 3C, F). Due to the low sample numbers, however, no statistical analysis was performed to determine the significance of this finding.

In the second injection using eGFP_gRNA2 plus Scr7, no phenotypically wild-type individuals were found during the screening (Suppl. Table 2), and consequently, no PCR and sequence analysis was performed.

Sequence information for all G₁ individuals generated by the third injection with eGFP_gRNA2b was obtained with the primer combination P_145/P_55, and for at least two individuals per family also with primer combination P_145/P_176 (M5m17, M5f8, M5f9, M3m1, M3f1, F1m1, F1f2, F4m1, F4f1). Analyzing the amplicons confirmed efficient HDR in all four families, with 70%–96% complete knock-in genotype within the phenotypically WT offspring (Fig. 3H, K, N, Q). NHEJ caused one to two different mutation events per family, explaining the knock-out phenotype of the remaining flies (Fig. 1B, Suppl. Table 1). None carried any mutation in the eGFP gene other than the ones close to the DSB site (data not shown). As observed with the first injection, the relative occurrence of HDR events increased from the first to the third egg collection time point in families M5 and M3 (Fig. 3I, L). Family F1 produced only ten G₁ progeny in total from two egg collections, of which nine showed complete knock-in genotypes (Fig. 3O, Suppl. Table 1). In family F4, the percentage of knock-in events per egg collection slightly decreased over time (Fig. 3R). None of the analyzed individuals originating from the eGFP_gRNA2b injections carried the incomplete knock-in with only two bp changed instead of three that was observed with eGFP_gRNA2.

4. Discussion

Genome editing via CRISPR-Cas HDR in Medfly was successfully developed and evaluated, using a short ssODN repair template to introduce point mutations in the eGFP marker gene of the transgenic line *TREhs43hid^{Ala5}_F1m2*. We used two different gRNAs to target eGFP and one single-stranded repair template (ssODN_BFP) to achieve the conversion. After successful HDR, two mismatches were introduced to the target sequence of eGFP_gRNA2 (194C > G; 196T > C), while its PAM sequence remained intact. Regarding eGFP_gRNA2b, an HDR event introduced one mismatch to the target sequence (201C > G) and two to the PAM sequence (194C > G; 196T > C), whereby the PAM was eliminated (Fig. 1A).

The use of an end-concentration of 300 mM KCl in the injection mix, as reported previously for experiments using Cas9 protein (Burger et al., 2016; Meccariello et al., 2017), seemed to help solubilizing the pre-assembled Cas9-gRNA RNPs, as there were no issues regarding clogging of needles while injecting such high concentrations of protein, DNA, and RNA (360 ng/μl, 200 ng/μl, and 200 ng/μl, respectively).

While only 50% of the injection survivors were fertile, we observed high efficiency of CRISPR-induced mutations, not only in the frequency of CRISPR-positive families (six out of seven fertile G₀) but also in the

inheritance within the families. Between 79 and 100% of G₁ individuals within a family showed the phenotypic loss of eGFP fluorescence, indicating a mutation event and efficient targeting of the germ line in the G₀ individuals. Sequence analysis confirmed these events and moreover revealed a knock-in rate of up to 96% (Fig. 3, Suppl. Table 1). However, blue fluorescence was not observed even if the flies were homozygous for the knock-in. This would have been the phenotypic confirmation of a positive knock-in event. A reason for this could be the melanization of the medfly thorax in combination with the low quantum yield of the BFP. BFP with its excitation maximum at 382 nm and an emission peak at 448 nm is different from the commonly used enhanced cyan fluorescent protein (eCFP; excitation 439 nm, emission 476 nm). BFP exhibits only about 15–20% of the brightness of wild-type GFP (Day and Davidson, 2009), which again has a weaker fluorescence intensity than the eGFP used in our target strain (relative fluorescence intensities). It is also weaker in its fluorescence intensity than eCFP, which can be difficult to detect even in the eyes of a white eye phenotype of *Aedes aegypti*, where no pigmentation can interfere with the fluorescence (Häcker, unpublished). Therefore, we conclude that very weak BFP fluorescence is too dim to be visible over the melanization of the thorax, whereas the much brighter eGFP fluorescence can penetrate the pigmentation.

Besides the complete three base pair knock-in event, we also detected a ‘partial knock-in’ with only two out of three base pairs changed when we used eGFP_gRNA2, but not with eGFP_gRNA2b. It was reported earlier that during HDR often only the part of the repair template overlapping with the deletion caused by the DSB is utilized (Paquet et al., 2016; Yang et al., 2013). As small deletions are more common than large deletions, the probability for a mutation to be incorporated during the HDR event decreases with the increasing distance from the cleavage site. This finding could explain the missing third SNP in the first experiment (201C > G, ‘partial knock-in’), as that SNP is the one most distal to the DSB side of eGFP_gRNA2. However, we did not observe anything similar for eGFP_gRNA2b, although the distance between the cleavage site and the most distal SNP is similar (six bp for eGFP_gRNA2b, versus seven bp for eGFP_gRNA2). Alternatively, the occurrence of the partial knock-in could be the result of re-editing of the already modified locus (Kwart et al., 2017), as the PAM of eGFP_gRNA2 remains intact after HDR whereas the PAM of eGFP_gRNA2b becomes eliminated. To ensure precise modification of the target site, it is therefore essential to include PAM-site mutations (silent) into the repair template (Paquet et al., 2016). We speculate that re-editing might also be the reason for the relative increase of the complete knock-in rate compared to the partial knock-in in later egg collection time points observed with eGFP_gRNA2 (Fig. 3). In that case, Cas9 would have to be stable and active for prolonged times to re-edit the germline in older flies. Thereby the chance of larger deletions to be formed would increase, thus increasing the chance for the most distal SNP to be incorporated upon HDR. Such a phenomenon paired with high efficiency achieved with CRISPR-Cas HDR would offer a possibility to save time and resources in mutagenic screens. Further experiments are needed, however, to corroborate these findings.

Interestingly, the use of Scr7 in the injection mix increased the hatch rate compared to the two injections without Scr7 (24.5% versus 6.6% and 5.1%, respectively), but it did not yield any phenotypic CRISPR events in Medfly. We excluded a possibly decreased stability of the gRNA in the presence of Scr7 by gel electrophoresis (Suppl. Fig. 1). The batch of Scr7 used in the experiment had been purchased some time ago. Therefore, we cannot exclude a reduced activity of the compound. Why a reduced or complete lack of Scr7 activity would abolish any DSB repair is not clear, however. Scr7 inhibits DNA ligase IV (Srivastava et al., 2012), a key enzyme in the NHEJ pathway and has been shown to enhance the HDR rate in human cell cultures or mouse embryos (Maruyama et al., 2015). There is also a contradictory report on the inhibitory function of Scr7, however. Greco et al. reported a stronger inhibition of human ligases I and III compared to IV,

questioning the specificity of Scr7 for ligase IV (Greco et al., 2016). In insects, increased HDR rates have been shown only in a ligase IV-deficient background of *Drosophila melanogaster*, by injecting zinc finger nuclease together with circular donor DNA into embryos (Beumer et al., 2008, 2013). In contrast, *Ae. aegypti* ligase IV-deficient strains are not viable. The strains were lost after a few generations. Instead, successful improvement of single-strand annealing, a form of HDR, was achieved by RNAi-mediated knock-down of *lig4* or *ku70* (Basu et al., 2015). Knock-down of *lig4*, however, at the same time increased the rate of NHEJ. Scr7, to our knowledge, so far has not been tested to enhance HDR in insects. Therefore, further experiments with Scr7 at different concentrations will be interesting to investigate if it has any effect in insects.

CRISPR-Cas allows a wide variety of genome editing strategies from small InDels at defined positions in the genome via NHEJ to the targeted introduction of point mutations (SNPs) via HDR all the way to the knock-out or knock-in of complete genes. While gene knock-in most probably will be classified as GMO, ‘non-traceable’ CRISPR-induced mutations like InDels and SNPs potentially might be regarded as non-GMO. In the US, the United States Department of Agriculture, Animal and Plant Health Inspection Service (USDA APHIS) recently classified CRISPR-edited organisms as ‘not regulated’ under the code of federal regulations 7 CFR part 340 (“Introduction of organisms and products altered or produced through genetic engineering which are plant pests or which there is reason to believe are plant pests”). One example is the modified white button mushroom (*Agaricus bisporus*) with an anti-browning phenotype which is achieved by the introduction of small deletions (1–14 bp) in a specific polyphenol oxidase gene via CRISPR gene editing (Waltz, 2016). The classification of specific CRISPR-induced alterations as non-GMO would allow the application of this highly efficient and versatile technique for the development of new or improved strains for Medfly SIT programs and possibly for other related Tephritid fruit flies. It would also finally allow transferring successful (sexing or lethality) systems based on small mutations to new pest species. The release of these strains could be discussed in line with other solutions regarding public acceptance, which is vital to the establishment and success of novel and safe pest control systems.

In conclusion, we demonstrate that genome editing via CRISPR-Cas HDR using a short, single-stranded DNA repair template is highly efficient in the fruit pest *C. capitata*. This high efficiency would offer a possibility to save time and resources in mutagenic screens, especially when targeting genes that do not alter the phenotype. Further experiments will be needed, however, to corroborate these findings. Moreover, it remains to be seen if this high efficiency can be matched with larger repair templates. This is the first report of successful CRISPR-Cas HDR genome editing in the family of Tephritidae, which contains many economically important fruit pest species. The establishment of CRISPR-Cas genome editing in Medfly, therefore, is an essential step towards the application of this technique to other Tephritid fruit pests like *Bactrocera dorsalis*, *B. oleae*, *Anastrepha ludens*, and *A. suspensa* and will be crucial for the development of new strategies to fight these pest insects.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

All data generated or analyzed are included in this article.

Declaration of interest

None.

Funding

This work was supported by the LOEWE Center for Insect Biotechnology & Bioresources of the Hessian Ministry for Science and the Arts (MFS), the Fraunhofer Attract program of the Fraunhofer Society (MFS), and the Emmy Noether program of the German Research Foundation SCHE 1833/1-1 (MFS).

The funding sources were not involved in any of the following: study design, collection, analysis and interpretation of data, writing of the report, and the decision to submit the article for publication.

Authors' contributions

RA performed research; MFS, RA and IH designed research; RA, MFS and IH analyzed data; and RA, IH and MFS wrote the paper. MFS and IH were group leaders for the project. All authors have read and approved the final version of the manuscript.

Acknowledgements

We wish to thank Tanja Rehling and Julia Hehner for technical assistance, and Kate Sutton for proofreading of the manuscript.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.ibmb.2018.08.004>.

References

- Ant, T., Koukidou, M., Rempoulakis, P., Gong, H.F., Economopoulos, A., Vontas, J., Alphey, L., 2012. Control of the olive fruit fly using genetics-enhanced sterile insect technique. *BMC Biol.* 10, 51.
- Augustinos, A.A., Targovska, A., Cancio-Martinez, E., Schorn, E., Franz, G., Cáceres, C., Zacharopoulou, A., Bourtzis, K., 2017. *Ceratitis capitata* genetic sexing strains: laboratory evaluation of strains from mass-rearing facilities worldwide. *Entomol. Exp. Appl.* 164, 305–317.
- Basu, S., Aryan, A., Overcash, J.M., Samuel, G.H., Anderson, M.A., Dahlem, T.J., Myles, K.M., Adelman, Z.N., 2015. Silencing of end-joining repair for efficient site-specific gene insertion after TALEN/CRISPR mutagenesis in *Aedes aegypti*. *Proc. Natl. Acad. Sci. U. S. A.* 112, 4038–4043.
- Beumer, K.J., Trautman, J.K., Bozas, A., Liu, J.L., Rutter, J., Gall, J.G., Carroll, D., 2008. Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19821–19826.
- Beumer, K.J., Trautman, J.K., Mukherjee, K., Carroll, D., 2013. Donor DNA utilization during gene targeting with Zinc-Finger nucleases. *G3* 3, 657–664.
- Bose, J.L., 2016. Chemical and UV mutagenesis. *Meth. Mol. Biol.* 1373, 111–115.
- Burger, A., Lindsay, H., Felker, A., Hess, C., Anders, C., Chiavacci, E., Zaugg, J., Weber, L.M., Catena, R., Jinek, M., Robinson, M.D., Mosimann, C., 2016. Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* 143, 2025–2037.
- Busch-Petersen, E., 1990. Temperature Sensitive Lethal Factors and Puparial Colour Sex Separation Mechanisms in the Mediterranean Fruit Fly, *Ceratitis capitata* (Wied). International Atomic Energy Agency (IAEA), Vienna.
- Concha, C., Palavesam, A., Guerrero, F.D., Sagel, A., Li, F., Osborne, J.A., Hernandez, Y., Pardo, T., Quintero, G., Vasquez, M., Keller, G.P., Phillips, P.L., Welch, J.B., McMillan, W.O., Skoda, S.R., Scott, M.J., 2016. A transgenic male-only strain of the New World screwworm for an improved control program using the sterile insect technique. *BMC Biol.* 14, 72.
- Day, R.N., Davidson, M.W., 2009. The fluorescent protein palette: tools for cellular imaging. *Chem. Soc. Rev.* 38, 2887–2921.
- Doench, J.G., Hartenian, E., Graham, D.B., Thohova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., Root, D.E., 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* 32, 1262–1267.
- Doudna, J.A., Charpentier, E., 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science (Wash. D C)* 346, 1258096.
- Franz, G., McInnis, D.O., 1995. A promising new twist - a genetic sexing strain based on a temperature sensitive lethal (tsl) mutation. In: Morse, J.G. (Ed.), *The Mediterranean Fruit Fly in California: Defining Critical Research*. University of California, Riverside, USA, pp. 187–196.
- Fu, G., Condon, K.C., Epton, M.J., Gong, P., Jin, L., Condon, G.C., 2007. Female-specific

- insect lethality engineered using alternative splicing. *Nat. Biotechnol.* 25.
- Gilles, J.R., Schetelig, M.F., Scolari, F., Marec, F., Capurro, M.L., Franz, G., Bourtzis, K., 2014. Towards mosquito sterile insect technique programmes: exploring genetic, molecular, mechanical and behavioural methods of sex separation in mosquitoes. *Acta Trop.* 132 (Suppl. 1), S178–S187.
- Glaser, A., McColl, B., Vadolos, J., 2016. GFP to BFP Conversion: a versatile assay for the quantification of CRISPR/Cas9-mediated genome editing. *Mol. Ther. Nucleic Acids* 5, e334.
- Greco, G.E., Matsumoto, Y., Brooks, R.C., Lu, Z., Lieber, M.R., Tomkinson, A.E., 2016. SCR7 is neither a selective nor a potent inhibitor of human DNA ligase IV. *DNA Repair* 43, 18–23.
- Gutschner, T., Haemmerle, M., Genovese, G., Draetta, G.F., Chin, L., 2016. Post-translational regulation of Cas9 during G1 enhances homology-directed repair. *Cell Rep.* 14, 1555–1566.
- Handler, A.M.J., Anthony, A., 2000. In: *Insect Transgenesis: Methods and Applications*, first ed. CRC Press, Boca Raton.
- Heim, R., Prasher, D.C., Tsien, R.Y., 1994. Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc. Natl. Acad. Sci. U.S.A.* 91, 12501–12504.
- Hendrichs, J., Franz, G., Rendon, P., 1995. Increased effectiveness and applicability of the sterile insect technique through male-only releases for control of Mediterranean fruit flies during fruiting seasons. *J. Appl. Entomol.* 119, 371–377.
- Hu, Z., Shi, Z., Guo, X., Jiang, B., Wang, G., Luo, D., Chen, Y., Zhu, Y.S., 2018. Ligase IV inhibitor SCR7 enhances gene editing directed by CRISPR-Cas9 and ssODN in human cancer cells. *Cell Biosci.* 8, 12.
- Isasawin, S., Aketarawong, N., Lertsiri, S., Thanaphum, S., 2014. Development of a genetic sexing strain in *Bactrocera carambolae* (Diptera: Tephritidae) by introgression of sex sorting components from *B. dorsalis* Salaya1 strain. *BMC Genet.* 15 (Suppl. 2), S2.
- Isasawin, S., Aketarawong, N., Thanaphum, S., 2012. Characterization and evaluation of microsatellite markers in a strain of the oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae), with a genetic sexing character used in sterile insect population control. *Eur. J. Entomol.* 109, 331–338.
- Kaiser, P.E., Seawright, J.A., Dame, D.A., Joslyn, D.J., 1978. Development of a genetic sexing system for *Anopheles albimanus*. *J. Econ. Entomol.* 71, 766–771.
- Kalajdzic, P., Schetelig, M.F., 2017. CRISPR/Cas-mediated gene editing using purified protein in *Drosophila suzukii*. *Entomol. Exp. Appl.* 164, 350–362.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649.
- Kim, H., Kim, J.-S., 2014. A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.* 15, 321–334.
- Kistler, K.E., Voshall, L.B., Matthews, B.J., 2015. Genome engineering with CRISPR-Cas9 in the mosquito *Aedes aegypti*. *Cell Rep.* 11, 51–60.
- Knipling, E.F., 1955. Possibilities of insect control or eradication through the use of sexually sterile males. *J. Econ. Entomol.* 48, 459–462.
- Kwart, D., Paquet, D., Teo, S., Tessier-Lavigne, M., 2017. Precise and efficient scarless genome editing in stem cells using CORRECT. *Nat. Protoc.* 12, 329–354.
- Li, F., Wantuch, H.A., Linger, R.J., Belikoff, E.J., Scott, M.J., 2014. Transgenic sexing system for genetic control of the Australian sheep blow fly *Lucilia cuprina*. *Insect Biochem. Mol. Biol.* 51, 80–88.
- Lin, S., Staahl, B.T., Alla, R.K., DouDNA, J.A., 2014. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife* 3, e04766.
- Maruyama, T., Dougan, S.K., Truttmann, M., Bilate, A.M., Ingram, J.R., Ploegh, H.L., 2015. Inhibition of non-homologous end joining increases the efficiency of CRISPR/Cas9-mediated precise genome editing. *Nat. Biotechnol.* 33, 538–542.
- McInnis, D.O., Tam, S., Lim, R., Komatsu, J., Kurashima, R., Albrecht, C., 2004. Development of a pupal color-based genetic sexing strain of the melon fly, *Bactrocera cucurbitae* (Coquillett) (Diptera: Tephritidae). *Ann. Entomol. Soc. Am.* 97, 1026–1033.
- Meccariello, A., Monti, S.M., Romanelli, A., Colonna, R., Primo, P., Inghilterra, M.G., Del Corsano, G., Ramaglia, A., Iazzetti, G., Chiarore, A., Patti, F., Heinze, S.D., Salvemini, M., Lindsay, H., Chiavacci, E., Burger, A., Robinson, M.D., Mosimann, C., Bopp, D., Saccone, G., 2017. Highly efficient DNA-free gene disruption in the agricultural pest *Ceratitis capitata* by CRISPR-Cas9 ribonucleoprotein complexes. *Sci. Rep.* 7, 10061.
- Ogaugwu, C.E., Schetelig, M.F., Wimmer, E.A., 2013. Transgenic sexing system for *Ceratitis capitata* (Diptera: Tephritidae) based on female-specific embryonic lethality. *Insect Biochem. Mol. Biol.* 43, 1–8.
- Orozco, D., Meza, J.S., Zepeda, S., Solis, E., Quintero-Fong, J.L., 2013. Tapachula-7, a new genetic sexing strain of the Mexican fruit fly (Diptera: Tephritidae): sexual compatibility and competitiveness. *J. Econ. Entomol.* 106, 735–741.
- Papanicolaou, A., Schetelig, M.F., Arensburger, P., Atkinson, P.W., Benoit, J.B., Bourtzis, K., Castañera, P., Cavanaugh, J.P., Chao, H., Childers, C., Curril, I., Dinh, H., Doddapaneni, H., Dolan, A., Dugan, S., Friedrich, M., Gasperi, G., Geib, S., Georgakilas, G., Gibbs, R.A., Giers, S.D., Gomulski, L.M., González-Guzmán, M., Guillem-Amat, A., Han, Y., Hatzigeorgiou, A.G., Hernández-Crespo, P., Hughes, D.S.T., Jones, J.W., Karagkouni, D., Koskinioti, P., Lee, S.L., Malacrida, A.R., Manni, M., Mathiopoulos, K., Meccariello, A., Murali, S.C., Murphy, T.D., Muzny, D.M., Oberhofer, G., Ortego, F., Paraskevopoulou, M.D., Poelchau, M., Qu, J., Reczko, M., Robertson, H.M., Rosendale, A.J., Rosselot, A.E., Saccone, G., Salvemini, M., Savini, G., Schreiner, P., Scolari, F., Siciliano, P., Sim, S.B., Tsiamis, G., Ureña, E., Vlachos, I.S., Werren, J.H., Wimmer, E.A., Worley, K.C., Zacharopoulou, A., Richards, S., Handler, A.M., 2016. The whole genome sequence of the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann), reveals insights into the biology and adaptive evolution of a highly invasive pest species. *Genome Biol.* 17, 192.
- Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K.M., Gregg, A., Noggle, S., Tessier-Lavigne, M., 2016. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* 533, 125–129.
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308.
- Rendon, P., McInnis, D., Lance, D., Stewart, J., 2004. Medfly (Diptera: Tephritidae) genetic sexing: large-scale field comparison of males-only and bisexual sterile fly releases in Guatemala. *J. Econ. Entomol.* 97, 1547–1553.
- Rössler, Y., 1979. The genetics of the mediterranean fruit fly: a “white pupae” mutant. *Ann. Entomol. Soc. Am.* 72, 583–585.
- Schetelig, M.F., Caceres, C., Zacharopoulou, A., Franz, G., Wimmer, E.A., 2009. Conditional embryonic lethality to improve the sterile insect technique in *Ceratitis capitata* (Diptera: Tephritidae). *BMC Biol.* 7, 4.
- Schetelig, M.F., Handler, A.M., 2012. A transgenic embryonic sexing system for *Anastrepha suspensa* (Diptera: Tephritidae). *Insect Biochem. Mol. Biol.* 42.
- Schetelig, M.F., Targovska, A., Meza, J.S., Bourtzis, K., Handler, A.M., 2016. Tetracycline-suppressible female lethality and sterility in the Mexican fruit fly, *Anastrepha ludens*. *Insect Mol. Biol.* 25, 500–508.
- Shrivastav, M., De Haro, L.P., Nickoloff, J.A., 2008. Regulation of DNA double-strand break repair pathway choice. *Cell Res.* 18, 134–147.
- Srivastava, M., Nambiar, M., Sharma, S., Karki, Subhas S., Goldsmith, G., Hegde, M., Kumar, S., Pandey, M., Singh, Ram K., Ray, P., Natarajan, R., Kelkar, M., De, A., Choudhary, B., Raghavan, Sathees C., 2012. An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell* 151, 1474–1487.
- Vreysen, M.J., 2001. Principles of area-wide integrated tsetse fly control using the sterile insect technique. *Med. Trop. : revue du Corps de sante colonial* 61, 397–411.
- Waltz, E., 2016. Gene-edited CRISPR mushroom escapes US regulation. *Nature* 532, 293.
- Yamada, H., Benedict, M.Q., Malcolm, C.A., Oliva, C.F., Soliban, S.M., Gilles, J.R., 2012. Genetic sex separation of the malaria vector, *Anopheles arabiensis*, by exposing eggs to dieldrin. *Malar. J.* 11, 208.
- Yan, Y., Scott, M.J., 2015. A transgenic embryonic sexing system for the Australian sheep blow fly *Lucilia cuprina*. *Sci. Rep.* 5, 16090.
- Yang, L., Guell, M., Byrne, S., Yang, J.L., De Los Angeles, A., Mali, P., Aach, J., Kim-Kiselak, C., Briggs, A.W., Rios, X., Huang, P.Y., Daley, G., Church, G., 2013. Optimization of scarless human stem cell genome editing. *Nucleic Acids Res.* 41, 9049–9061.
- Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E.V., Zhang, F., 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* 163, 759–771.

Supplementary Information for

Highly efficient genome editing by homology-directed repair using Cas9 protein in *Ceratitis capitata*

Roswitha A. Aumann^a, Marc F. Schetelig^{a,b,*}, Irina Häcker^{a,b}

Justus-Liebig-University Gießen, Institute for Insect Biotechnology, Department of Insect Biotechnology in Plant Protection, Winchesterstr. 2, 35394 Gießen, Germany

Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Division of Bioresources, Department of Insect Pest and Vector Control, 35394 Gießen, Germany

*Corresponding author: marc.schetelig@agrار.uni-giessen.de

This PDF file includes:

- Supplementary Tables 1-3
- Supplementary Figure 1

	G ₀ family	G ₁ phenotype	Mutation type	Sequence changes	number of G ₁ flies
eGFP_gRNA2	M1	WT	knock-in	194C>G, 196T>C, 201C>G	18
		WT	partial knock-in	194C>G, 196T>C	50
		WT	6 bp deletion	192-197 GACCTA	18
		WT	knock-in, 1 bp insertion	194C>G, 196T>C, 201C>G, 202_G_203	3
		WT	1 bp deletion, 19 bp insertion	195 C, 206_CCCTGAGCCACGGGGTGCA_207	2
		WT	19 bp insertion	206_CCCTGAGCCACGGGGTGCA_207	3
		WT	no sequence information*		4
		GFP			18
				Σ 116	
	F8	WT	knock-in	194C>G, 196T>C, 201C>G	17
WT		partial knock-in	194C>G, 196T>C	9	
WT		1 bp deletion	195 C	3	
WT		1 bp deletion	193 A	3	
WT		no sequence information*		2	
GFP				8	
			Σ 42		
F2	GFP			61	
			Σ 61		
eGFP_gRNA2b	M5	WT	knock-in	194C>G, 196T>C, 201C>G	19
		WT	3 bp deletion	200-202 GCG	5
		WT	2 bp deletion	201-202 CG	2
		GFP			7
				Σ 33	
	M3	WT	knock-in	194C>G, 196T>C, 201C>G	50
		WT	4 bp deletion	202-205 GTGC	2
		GFP			8
				Σ 60	
	F1	WT	knock-in	194C>G, 196T>C, 201C>G	9
WT		190 bp deletion	11-201	1	
GFP				0	
			Σ 10		
F4	WT	knock-in	194C>G, 196T>C, 201C>G	60	
	WT	15 bp deletion	197-211 ACGGCGTGCAAGTCT	26	
	GFP			19	
			Σ 105		

Supplementary Table 1. Summary of all CRISPR-Cas events identified in G₁ individuals. Shown are the gRNAs used, the fertile G₀ families (injection survivors individually backcrossed to wild-type flies), the phenotype of the G₁ flies (WT = phenotypically missing eGFP fluorescence, GFP = heterozygous eGFP fluorescence), the mutation types, the sequence changes including the mutation positions in the eGFP sequence, and the number of flies carrying the mutation. SNPs are shown as 'position native base > mutated base', deletions as 'position deleted bases', insertions as 'position_inserted bases_position'. Five G₀ survivors of eGFP_gRNA2 injection and five G₀ survivors of eGFP_gRNA2b injection did not produce offspring and are not shown here. * For these flies no PCR product was obtained, probably due to genomic DNA degradation after death of the flies. Therefore, the eGFP locus could not be sequenced.

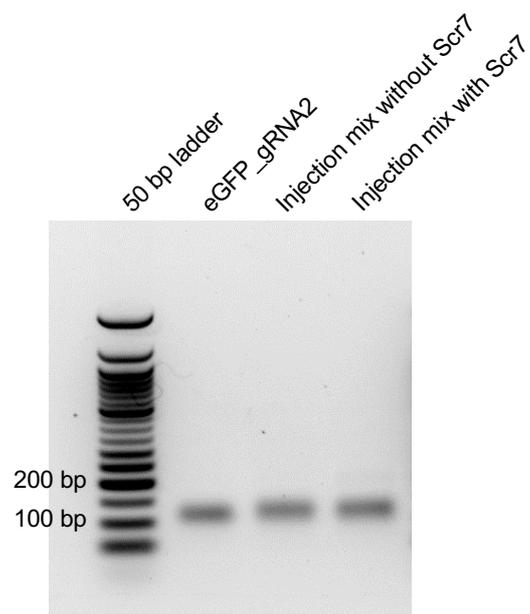
eGFP_gRNA2 + 1mM Scr7		
G ₀ family	G ₁ pheno-type	number of G ₁ flies
F1	GFP	3
F3	GFP	59
F4	GFP	27
F5	GFP	49
F6	GFP	21
F8	GFP	13
F9	GFP	13
F10	GFP	39
F11	GFP	56
F12	GFP	65
F13	GFP	27
F14	GFP	47
		Σ 419

eGFP_gRNA2 + 1mM Scr7		
G ₀ family	G ₁ pheno-type	number of G ₁ flies
M1	GFP	129
M2	GFP	88
M3	GFP	115
M4	GFP	77
M5	GFP	7
M6	GFP	146
M7	GFP	138
M9	GFP	101
M10	GFP	27
M11	GFP	84
M13	GFP	240
M14	GFP	113
M15	GFP	61
M16	GFP	158
M17	GFP	64
		Σ 1548

Supplementary Table 2. Fertile crosses of the eGFP_gRNA2 injection with Scr7. Shown are the fertile G₀ families (injection survivors individually backcrossed to wild-type flies), phenotype of the G₁ flies (GFP = heterozygous eGFP fluorescence), and the number of G₁ flies screened for GFP fluorescence. Four G₀ survivors of an eGFP_gRNA2+Scr7 injection did not produce offspring and are not shown here.

	G₁ cross	number of G₁ individuals	number of G₁ with knock-in genotype	number of G₁ with InDel genotypes	G₂ phenotype	number of G₂ flies
eGFP_gRNA2b	M3 (T1)	23 (9♂/14♀)	22	1	WT	111
	M3 (T2)	24 (9♂/15♀)	23	1	WT	320
	M3 (T3)	5 (1♂/4♀)	5	0	WT	132
	M5 (T3)	3 (1♂/2♀)	3	0	WT	104
	F1 (T2)	9 (4♂/5♀)	8	1	WT	171
	F4 (T1)	12 (6♂/6♀)	9	3	WT	126
	F4 (T2)	38 (18♂/20♀)	28	10	WT	811
	F4 (T3)	36 (20♂/16♀)	23	13	WT	207
						Σ 1982

Supplementary Table 3. G₁ sibling crosses and screening of G₂. Shown are the sibling crosses of G₁ knock-out phenotype individuals from different egg collection time points (T1, T2, T3). Described are the total number of G₁ individuals per cross (male/female), the number of knock-in and InDel genotypes within the cross, the phenotype of G₂ offspring, and the number of screened G₂. G₂ individuals were screened for eGFP and BFP fluorescence.



Supplementary Figure 1. Stability of eGFP_gRNA2 in injection mixes with and without Scr7. Shown is the eGFP_gRNA2 (200 ng/ μ l) and the injection mix (360 ng/ μ l Cas9 protein, 200 ng/ μ l eGFP_gRNA2, 300 mM KCl endconcentration, incubated 10 min at 37°C) without and with 1 mM Scr7. All samples show bands of similar intensity at the expected size of 120 bp. The DNA ladder used for agarose gel is the 50 bp DNA-ladder (NEB), bp = base pairs.

3.2 The Transformer Project: female-to-male sex conversion in *Ceratitis capitata*

To achieve temperature-dependent female-to-male sex conversion in *C. capitata*, CRISPR/Cas9 HDR from 3.1. was used to induce point mutations in the sex determination gene *transformer-2* (*tra2*). Although sexing of medflies in mass-rearing conditions is already possible by using genetic sexing strains, a temperature-dependent female-to-male conversion would still be beneficial for SIT programs, because double the number of males could be produced per parental egg capacity. Therefore, I attempted to reconstruct the original chemically induced temperature-sensitive *tra2^{ts}* mutations of *Drosophila melanogaster tra2* in *C. capitata tra2*. High HDR efficiency and the usage of non-lethal genotyping was critical for this experiment, because I decided to avoid the introduction of exogenous DNA and the use of a phenotypic marker. I succeeded to generate homozygous *tra2^{ts2}* mutants and 100% female-to-male conversion with this non-transgenic technique. Subsequently, I studied the restrictive temperature range of the *Cctra2^{ts2}* strain and the temperature-dependent fertility of medfly. It was not possible to identify a permissive temperature range at which normal rearing and maintenance of the *tra2^{ts2}* mutants would have been possible. Therefore, *tra2^{ts2}* could not be used to establish conditional sex conversion for an SIT application in medfly. However, the study provides a straightforward strategy to introduce point mutations without the use of transgenes.

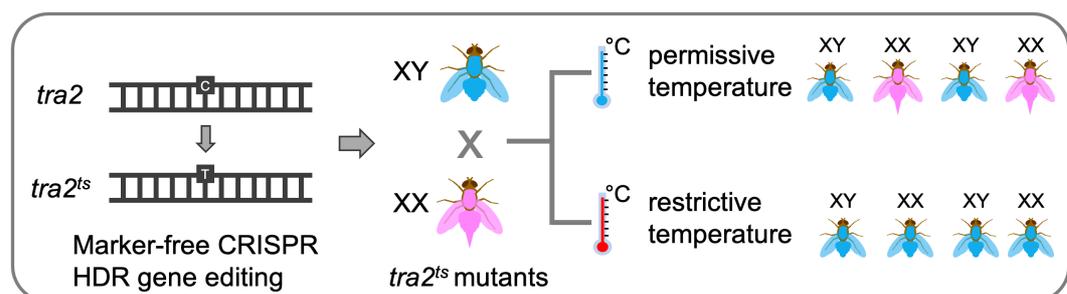
Title: Female-to-male sex conversion in *Ceratitis capitata* by a CRISPR/Cas9 HDR-induced point mutations in the sex determination gene *transformer-2*

Authors: Aumann RA, Häcker I, Schetelig MF

Status: published in *Scientific Reports* (2020)

Contributions: all experiments were performed by Aumann RA

Presentations: This work was presented at the 4th Research Coordination Meeting on ‘Comparing Rearing Efficiency and Competitiveness of Insect Sterile Male Strains Produced by Genetic, Transgenic or Symbiont-based Technologies’ of the FAO/IAEA in Adelaide, Australia (2019, talk), and the 12th GGL conference on Life Science in Gießen, Germany (2019, talk, awarded as best oral presentation).





OPEN

Female-to-male sex conversion in *Ceratitis capitata* by CRISPR/Cas9 HDR-induced point mutations in the sex determination gene *transformer-2*

Roswitha A. Aumann, Irina Häcker & Marc F. Schetelig

The Sterile Insect Technique (SIT) is based on the mass release of sterilized male insects to reduce the pest population size via infertile mating. Critical for all SIT programs is a conditional sexing strain to enable the cost-effective production of male-only populations. Compared to current female-elimination strategies based on killing or sex sorting, generating male-only offspring via sex conversion would be economically beneficial by doubling the male output. Temperature-sensitive mutations known from the *D. melanogaster transformer-2* gene (*tra2^{ts}*) induce sex conversion at restrictive temperatures, while regular breeding of mutant strains is possible at permissive temperatures. Since *tra2* is a conserved sex determination gene in many Diptera, including the major agricultural pest *Ceratitis capitata*, it is a promising candidate for the creation of a conditional sex conversion strategy in this Tephritid. Here, CRISPR/Cas9 homology-directed repair was used to induce the *D. melanogaster*-specific *tra2^{ts}* SNPs in *Ctra2*. 100% female to male conversion was successfully achieved in flies homozygous for the *tra2^{ts2}* mutation. However, it was not possible, to identify a permissive temperature for the mutation allowing the rearing of a *tra2^{ts2}* homozygous line, as lowering the temperature below 18.5 °C interferes with regular breeding of the flies.

The production of large populations of only male pest insects is a key factor for the Sterile Insect Technique (SIT), a highly successful, environment-friendly, and species-specific method of pest control. Proposed in 1955 by Knipling¹, the SIT is based on the sustained mass-release of sterile males into the existing pest population to reduce population size by infertile mating, and has been successfully applied to several pest species^{2–5}. The release of pure male populations is important because male-only releases are more effective than bisexual ones⁶ by preventing the mating of sterile males with the co-released sterile females. In addition, the release of sterile females could still result in crop damage due to oviposition or, in case of vector insects, in disease transmission. Sexing, classified as the removal of females from a mass-reared insect population, can be achieved by physical sorting, female-specific lethality, or by converting females into males⁷. Such solutions have been developed for multiple pest species using naturally occurring or classically induced mutations^{8–11} or transgenesis^{12–17}. Most of them, however, are not ready for mass-rearing yet. To allow efficient rearing of sexing strains and cost-effective operation of the program, important characteristics of sexing systems are the conditionality and early developmental time-point of the sexing, respectively. Currently, two conditional embryonically active systems exist for the devastating agricultural pest *Ceratitis capitata* (Wiedemann; Diptera: Tephritidae) (Mediterranean fruit fly, medfly). Medflies pose a vast economic threat to agriculture worldwide, as they feed on >260 plants (fruits, vegetables, nuts) and are highly invasive: Native to the Afrotropical region, medfly can now be found in most tropical and temperate regions^{18,19}.

In the successful medfly genetic sexing strains (GSS), VIENNA 7 and 8, an unknown recessive autosomal temperature-sensitive lethal (*tsl*) mutation eliminates all female embryos upon heat shock¹¹. The GSS males, however, are semi-sterile due to chromosomal rearrangements necessary to rescue the WT phenotype, resulting in 50% genetically imbalanced gametes and thus non-viable zygotes. In a conditional transgenic embryonic sexing system (TESS) medfly female embryos are killed by overexpression of a pro-apoptotic gene²⁰. The TESS

Department of Insect Biotechnology in Plant Protection, Institute for Insect Biotechnology, Justus-Liebig-University Gießen, Winchesterstr. 2, 35394 Giessen, Germany. email: marc.schetelig@agr.uni-giessen.de

can be switched off for strain maintenance by adding the antibiotic tetracycline to the fly food (Tet-off system). Compared to these systems, a sexing system based on temperature-inducible female-to-male conversion would have two advantages: (1) doubling or, compared to semi-sterile GSS, even quadrupling the number of males for the release and (2) abolishing the use of antibiotics. Both factors would considerably reduce costs and increase the efficiency of a medfly SIT program. However, population maintenance would presumably need to be done at reduced temperatures, which could decrease the productivity of the mass-rearing due to prolonged development times²¹. Currently, the production of one million sterile medfly pupae of the classical GSS is estimated at US\$ 250–500, depending on the production level and the location of the rearing facility²².

In search of genetic elements suitable to construct sexing or sex-conversion systems, insect sex determination pathways have been studied to identify essential genes and to understand their function. The *transformer-2* gene (*tra2*) is involved in the sex determination pathway of different insects, including *C. capitata*^{23,24}. In medfly, *transformer-2* is an auxiliary factor, necessary to establish and sustain the autoregulation of *transformer*, a gene known to be crucial for the sexual fate^{23,25,26}. As illustrated and described in detail elsewhere^{23,26}, maternal *Cctra* and *Cctra2* initiate a positive feedback loop in XX fertilized eggs and control the female-specific splicing of the downstream targets *doublesex* and *fruitless*^{23,26}. Switching off either *Cctra* or *Cctra2* leads to male development²⁶ and the transient knock-down of *Cctra2* during embryogenesis via RNA interference (RNAi) resulted in full sex-reversal of XX-karyotype flies into phenotypic males²³. In contrast to *Anastrepha suspensa*, where embryonic injection of dsRNA against *Astra2* resulted in sex-reversed XX males, which were infertile despite testes full of sperm bundles²⁷, medfly XX-karyotype males were fertile^{23,25}, indicating that male-fertility is not Y-dependent in *C. capitata*. Sex-reversion via RNAi-mediated knock-down of *tra2* was also shown in *Bactrocera tau* (Walker)²⁸ and *B. dorsalis* (Hendel)²⁹. However, to make use of the *tra2*-mediated sex-conversion for male-only production, it needs to be conditionally inducible and stable. In *Drosophila melanogaster*, two *tra2* temperature-sensitive mutations (*tra2^{ts1}*, *tra2^{ts2}*) are known, supposedly causing conformational changes in the protein structure at elevated (restrictive) temperatures (29 °C). These result in a loss of protein function and therefore in sex-conversion of XX embryos (male-only offspring). At permissive temperatures (e.g. 16 °C), a functional TRA2 protein allows healthy female development and rearing of the population^{30,31}. Due to the high conservation of TRA2 among different species^{23,32–36}, gene editing techniques such as CRISPR/Cas³⁷ can be used to exactly recreate temperature-sensitive *tra2* mutations known from *D. melanogaster* in homologous genes of pest insects. This has been shown by Li and Handler³⁸, who introduced the *D. melanogaster tra2^{ts2}* mutation together with a fluorescent marker into the *D. suzukii tra2* gene. 16 °C and 20 °C were permissive temperatures for *D. suzukii tra2^{ts2}* mutants³⁸, resulting in fertile and normally developed males and females. At 26 °C, all XX embryos developed as sterile intersex with sex combs and male-like genitalia, and all XY embryos showed dysmorphic testes and were sterile. However, the survival rate for both, wild-type and mutant flies was very low (5–10%) at this temperature and even lower at more elevated temperatures. While this temperature-sensitivity of *D. suzukii* would be problematic if the *tra2^{ts2}* mutation were to be used for sexing in an SIT application, this should not be an issue for medfly, which can be reared at 26 °C. Based on this fact and the promising results from the previous transient knock-down of *tra2* in *C. capitata*²³, *Cctra2* is a good candidate for the construction of a temperature-based sex-conversion system in medfly.

Hence, we used our previously established protocol for markerless CRISPR/Cas9-HDR in medfly yielding high-efficiency mutagenesis³⁹ to integrate the *D. melanogaster tra2^{ts1}* and *tra2^{ts2}* mutations into the *Cctra2* homolog. Omitting the use of a fluorescent marker gene should facilitate the use of non-transgenic strains in SIT programs, as CRISPR/Cas9-induced single nucleotide polymorphisms (SNP) are even considered non-GMO in certain countries⁴⁰.

Results

***Cctra2* mutagenesis: gRNA and repair template design.** CRISPR/Cas9 HDR gene editing was used to separately re-create the two temperature-sensitive *D. melanogaster tra2* mutations (*ts1*, *ts2*) in the *C. capitata* homolog *Cctra2* (NCBI Gene ID: 101452698). Positions of the mutations were determined by comparing amino acid sequence identity for *D. melanogaster* and medfly TRA2. The mutated Alanine151 in the *Dmel tra2^{ts1}* (Ala-151Val)³⁰ corresponds to *Ccap* Ala158, the Proline181 of the *Dmel tra2^{ts2}* mutation (Pro181Ser) to *Ccap* Pro188. The *ts2* mutation is located in a 19 aa linker region, which is a unique feature of TRA2 and highly conserved among species^{23,32–36} (Fig. 1a).

For both mutations, a single guide RNA (gRNA) and a 140 nt single-stranded oligodeoxynucleotide (ssODN) repair template were designed to introduce the amino acid exchanges corresponding to the *Dmel ts1* or *ts2* mutations (*ts1*: 158 Ala > Val, *ts2*: 188 Pro > Ser), to create temperature-sensitive versions of the CcTRA2 protein. The repair template ssODN_tra2_ts1 differs from the wild-type *tra2* ORF sequence by two bases, a C > T transition at position 473 of the CDS to introduce the *ts1* SNP and the silent mutation 477 G > A that removes the PAM sequence to prevent re-editing. ssODN_tra2_ts2 differs by one base introducing the *ts2* SNP (CDS: 562 C > T) (Fig. 1b).

Preliminary gRNA tests to confirm editing capability of *tra2^{ts}* positions. To assess the functionality of the *tra2^{ts1}* and *tra2^{ts2}* gRNAs, each was injected complexed with Cas9 protein and either without (non-homologous end joining, NHEJ, knock-out) or with repair template (homology-directed repair, HDR, knock-in). G₀ survivors of these injections were reared at 26 °C. 327 *Egypt II* wild-type (*EgII* WT) embryos were injected for *tra2_ts1* knock-out. Ten reached adult stage (six males, four females) (Table 1a). One male was fertile. The *ts1* injection with repair template (290 *EgII* embryos) yielded four viable but infertile adults (two males, two females), and three adults got stuck in the puparium while eclosing and died (two males, one female) (Table 1a). None of the *ts1* G₀ adults showed external phenotypic abnormalities. To check for editing activity of

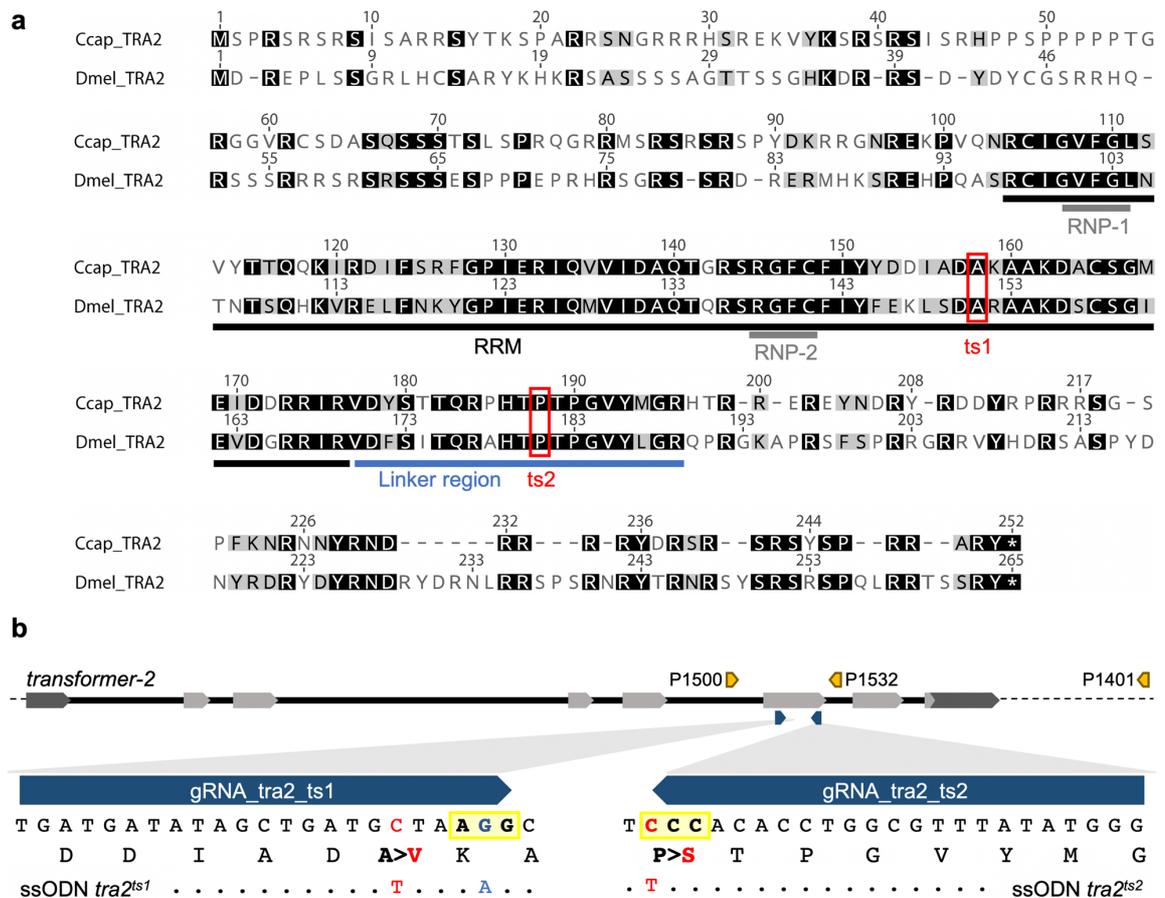


Figure 1. Strategy to re-create *D. melanogaster tra2^{ts}* alleles in *C. capitata tra2*. **(a)** Amino acid alignment of *C. capitata* and *D. melanogaster* TRA2. Shown are the RNA recognition motif (RRM, black), two ribonucleoprotein identifier sequences (RNP motifs, grey), the linker region (blue), and the position of the *tra2^{ts1}* and *tra2^{ts2}* mutations (red). Consensus is shown in black, amino acids with similar characteristics in grey. **(b)** Overview of *Cctra2* gene structure (*tra2* exon structure, light grey: CDS, dark grey: UTR), primers used for genotyping (P1500/P1401) or for genomic positive control PCRs (P1500/P1532 and P1500/P1401), position of single guide RNAs (blue arrows) and mutations mediated by the HDR repair templates (ssODN). PAM sequences are marked in light yellow, position of SNPs introduced by HDR are shown and marked either in blue (silent mutation) or in red (functional mutations). Resulting amino acid exchanges are indicated.

gRNA_tra2_ts1, the *tra2* genotype of four randomly chosen G₀ flies (two from each injection) was analysed by subcloning the *tra2*-specific PCR products. One of two knock-out injected G₀ flies showed a 1 bp deletion in one of five sequenced clones. One of the knock-in injected G₀ flies showed two independent events within five sequenced clones, the *tra2^{ts1}* HDR genotype or a 6 bp deletion (Supplementary Fig. S1a).

The *tra2^{ts2}* gRNA knock-out injection yielded six adult males from 367 injected *EgII* embryos (Table 1b), three of them were fertile. Additionally, three G₀ flies stuck in the puparium did not survive (two males, one intersex IS1-KO). The *tra2^{ts2}* knock-in mix was injected into 244 *EgII* embryos (Table 1b). Eight developed to adults (six males, two intersex: IS1, IS2), and four died during eclosing (one male, three intersex: IS3-6). Intersex flies showed varying degrees of phenotypically male and female characteristics (genital terminalia apparatus and bristles) (Fig. 2a), and were sterile. In contrast, all six G₀ males were fertile. The genotype of six *ts2* G₀ flies from the knock-out (males M5, M6, and intersex IS1-KO) and knock-in injection (IS1, IS4, IS5) was analysed. All showed NHEJ events ranging from 33 bp deletions to 4 bp insertions (Supplementary Fig. S1b). G₁ offspring from both injections was not analysed.

These experiments confirmed the editing activity of the *ts1* and *ts2* gRNAs. The lack of fertile G₀ in the *ts1* injections and the complete lack of females and appearance of intersexes in the *ts2* injections, however, indicated that 26 °C is a restrictive temperature for the *Cctra2^{ts}* mutations.

Evaluation of medfly rearing at low temperatures. To evaluate if *D. melanogaster* and *D. suzukii tra2^{ts2}* permissive temperatures, 16 °C or 16 to 20 °C, respectively^{30,31,38}, are applicable to medfly, newly eclosed

	26 °C, KO	26 °C, KI	19 °C, KI
(a) Injections for target <i>tra2^{ts1}</i>			
Injected embryos	327	290	169
Larvae (% hatching)	32 (9.8%)	28 (9.6%)	55 (32.5%)
Pupae	16	8	42
G ₀ adults viable; not viable (% eclosion)	10;0 (3.0%)	4;3 (2.4%)	33;0 (19.5%)
G ₀ viable males (fertile)	6 (1)	2 (0)	11 (>4)
G ₀ viable females (fertile)	4 (0)	2 (0)	22 (>2)
G ₀ viable intersex (fertile)	0	0	0
(b) Injections for target <i>tra2^{ts2}</i>			
Injected embryos	367	244	181
Larvae (% hatching)	29 (7.9%)	30 (12.3%)	52 (28.7%)
Pupae	12	18	17
G ₀ adults viable; not viable (% eclosion)	6;3 (2.4%)	8;4 (4.9%)	11;2 (7.1%)
G ₀ viable males (fertile)	6 (3)	6 (6)	5 (2)
G ₀ viable females (fertile)	0	0	0
G ₀ viable intersex (fertile)	0	2 (0)	6 (0)

Table 1. Summary of injections for targeted *Cctra2* knock-out or knock-in mutations. Shown is the mutation target *tra2^{ts1}* (a) and *tra2^{ts2}* (b), the strategy (knock-out (KO) or knock-in (KI)), the rearing temperature, the number of injected embryos and surviving G₀ larvae, pupae and adults, the larval and adult hatch rate, and the number, phenotypic sex and fertility of viable G₀ adults. Number of fertile flies for the *tra2^{ts1}* KI injection at 19 °C could not be exactly assessed, as only twelve flies were backcrossed individually and remaining 21 flies were backcrossed in three groups.

WT *EgII* (60–160 adults per experiment) were transferred from 26 °C to 16, 18, or 19.5 °C and eggs of these crosses were collected for seven days (for temperature profiles and egg collection timepoints see Supplementary Table S1 and Fig. S2a–c). At 19.5 °C, the number of adult offspring was reduced to about 40%, compared to 26 °C, and at 18 °C to about 1%. At 16 °C, no larvae hatched from more than 2,000 collected eggs. Hence, 19 °C was chosen as a possible rearing and potential permissive temperature for the subsequent *Cctra2^{ts}* injections.

CRISPR/Cas9-HDR injections at 19 °C do not produce stable *Cctra2^{ts1}* lines. Rearing of *ts1*-injected G₀ at 19 °C increased the number of adult G₀ survivors to 19.5% compared to 3% and 2.4% for the *ts1* injections at 26 °C (Table 1a). None of them showed external phenotypic abnormalities. Twelve G₀ flies were backcrossed individually to *EgII* (M5–M10 and F4–F9), remaining flies were backcrossed in three groups (M-group I, F-group I, II). After allowing sufficient time for mating and egg laying, all individually crossed G₀ flies were dissected to examine their reproductive organs. Phenotypes included females without ovaries (F4), only one ovary (F7), or normal ovaries (F6, F8, F9). Males showed normal reproductive organs, except for M7, which had no testes (Supplementary Fig. S3a, b). F5 and M10 died and could not be dissected. Overall, 47 G₁ flies eclosed from eight fertile families (F6, F8, M6, M8, M9, M10, group M_I and F_II). Since no phenotypic marker was inserted to track successful mutagenesis in G₁, non-lethal genotyping was used to analyse G₁ offspring reared at 19 °C for the presence of the *tra2^{ts1}* mutation. DNA was extracted from a single leg, and the *ts1* target site region was PCR-amplified and sequenced. 38 of 47 G₁ flies provided sufficient quality sequence information. All showed WT genotype.

CRISPR/Cas9-HDR successfully creates inheritable *Cctra2^{ts2}* mutation at 19 °C. Rearing of *ts2* HDR-injected G₀ at 19 °C yielded lower survival numbers than the *ts1* HDR injection, but still about twice as high as the experiments at 26 °C (7.1% compared to 2.4% and 4.9%; Table 1b). Injection of 181 *EgII* embryos resulted in five viable males and six intersex. Additionally, two intersex flies (IS9, IS10) died during eclosing. Males and intersex were individually backcrossed to WT virgin females. Eggs were collected every second day for 10 consecutive days (for temperature profile during egg collection see Supplementary Table S1, Fig. S2b). Two of the eleven crosses (M8, M11) produced G₁ offspring (Supplementary Table S2). After mating, all alive G₀ were dissected. Males M8, M9, and M11 showed normal testes, while M10 did not have testes (Supplementary Fig. S3c). Flies with intersex phenotype showed apparently normal ovaries but no spermathecae (IS6), hypertrophic testes (IS8), miniaturized testes (IS13), or no identifiable reproductive organs (IS7, IS11, IS12; Fig. 2b). To assess the karyotype of all 13 G₀ flies, PCR on Y-chromosome-specific repetitive elements was performed, whereby absence of a PCR signal implies a XX-karyotype. None of the intersex phenotype G₀ flies was positive for the Y-chromosome-specific PCR (Fig. 2c), indicating that all XX (female) karyotype G₀ embryos were transformed to intersex flies. The absence of phenotypically female G₀ in all three *tra2^{ts2}* injections indicates a high efficiency of gRNA_ *tra2^{ts2}* and the importance of the targeted position for proper TRA2 function in female sex development.

The *tra2* genotype of the G₁ flies was analysed via non-lethal genotyping. For family M8, ten of twelve analysed G₁ (83%) were heterozygous for the knock-in genotype (*tra2^{ts2}*), and two (17%) carried NHEJ events.

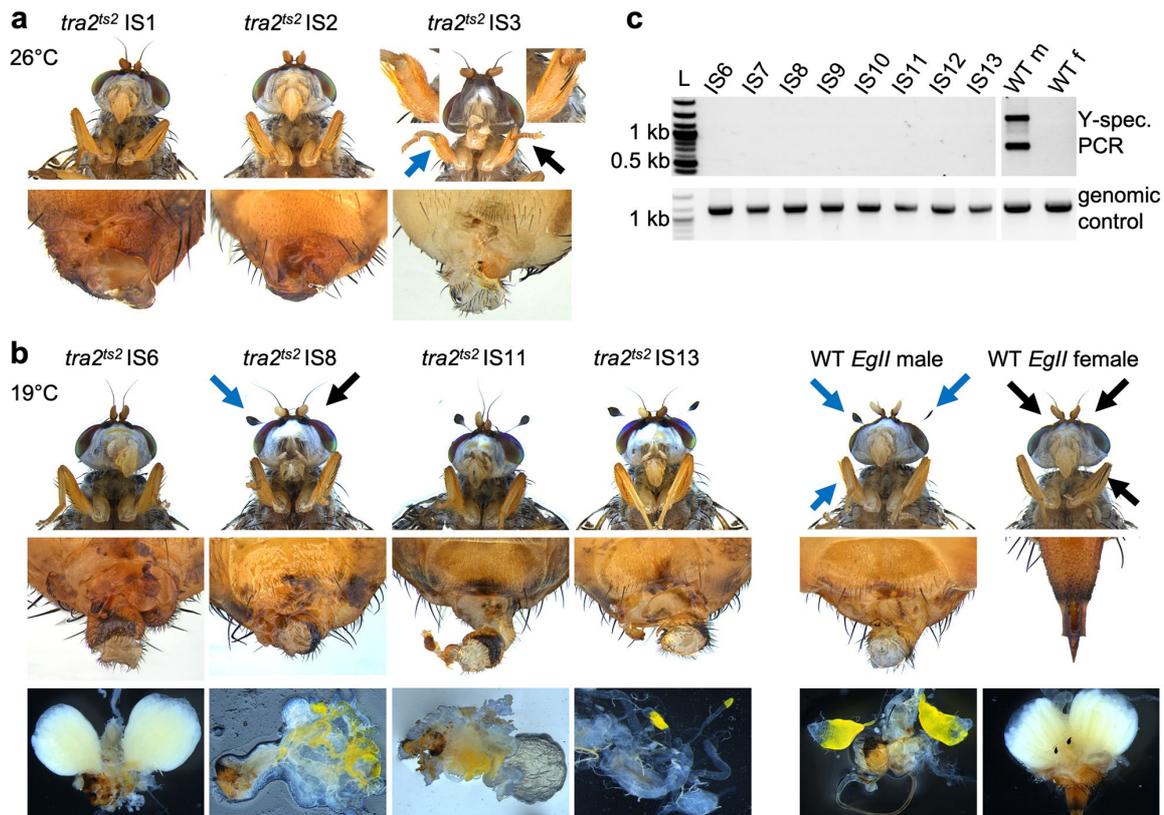


Figure 2. Somatic modification of *tra2^{ts2}* causes intersexuality with external and internal phenotypic abnormalities in G_0 . G_0 survivors of *tra2^{ts2}* KI injections reared at 26 °C (a) and 19 °C (b) show intersex phenotypes with malformed external and internal reproductive organs and mixed male- and female-specific characteristics. Phenotypes included deformed ovipositors (IS1, IS2), a mixture of male- and female-specific bristles on the femur (IS3) or the head (IS8), absent genitalia (IS3), and various degrees of deformed male genitalia, combined with ovaries without spermatheca (IS6), testes-like structures (IS8, IS13) or no identifiable reproductive organs (IS11). For comparison, wild-type males have two spatulated bristles on the head, non-pigmented bristles on the femur, and a prominent ovipositor. Male characteristics are highlighted by blue arrows, female characteristics by black arrows. (c) Karyotyping via Y-chromosome-specific PCR of intersex phenotype *tra2^{ts2}* KI individuals (19 °C) revealed XX-karyotype for all intersex individuals. Shown is the Y-chromosome-specific PCR (primers P1504/1505) on genomic DNA extracted from a single fly and a genomic positive control on *tra2* with primers P1401/1500 using the same DNA samples. Wild-type male (WT m) and female (WT f) are shown as control. Displayed are cropped parts from different gels. Uncropped versions of the gels are provided in the supplement (Supplementary Fig. S4). L = DNA ladder; kb = kilo base pairs.

The remaining individuals were not analysed due to low DNA quality. From the G_1 offspring of family M11, 60 flies were randomly chosen for genotyping. The heterozygous *tra2^{ts2}* genotype was found in 45 flies (75%). This percentage was similar in males (26 of 33) and females (19 of 27). Nine flies (15%) were WT and six flies could not be analysed (low DNA quality).

Inbreeding of the *ts2* mutation at 19 °C does not produce phenotypic females homozygous for *tra2^{ts2}*. Heterozygous *tra2^{ts2}* mutant G_1 flies were either inbred or backcrossed to *EgII* to ensure the propagation of the line if inbreeding should turn out to be sterile. Details on crosses, egg collection numbers, and temperature profiles are shown in Supplementary Tables S1, S3, and Fig. S2. Inbreeding of heterozygous M8 offspring produced 121 G_2 flies with a 1:2 female to male ratio (Supplementary Table S3). 27 of 78 phenotypic G_2 males were homozygous for the *ts2* mutation (*tra2^{ts2/ts2}*), 38 were heterozygous (*tra2^{ts2/WT}*), and 13 were WT (*tra2^{WT/WT}*, Fig. 3a). In contrast, none of the 38 phenotypic females were homozygous for *tra2^{ts2}*, 24 were heterozygous, and 14 had two WT *tra2* alleles (Fig. 3a). *Inter se* crosses of M11 offspring resulted in a similar phenotypic female to male ratio as M8 inbreeding (26 and 42, respectively). Non-lethal genotyping showed that also M11 inbreeding produced phenotypic *tra2^{ts2}*-homozygous males (21%), but no phenotypic females with two *tra2^{ts2}* alleles (Fig. 3a). Backcross of *tra2^{ts2}* heterozygous M11 offspring produced a 1:1 phenotypic sex ratio (Supplementary Table S3), which was not further analysed molecularly.

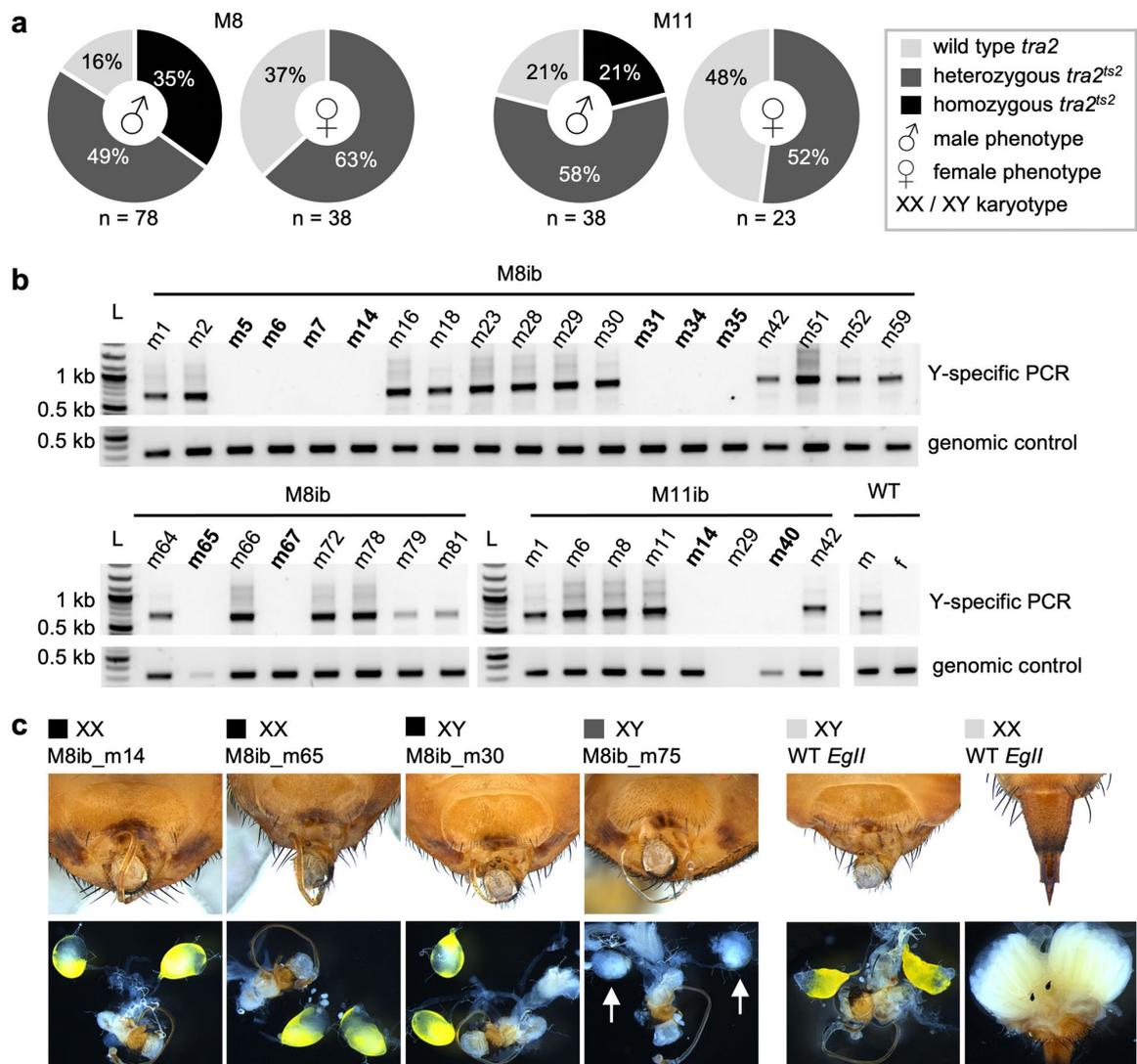


Figure 3. Analysis of *tra2^{ts2}* genotypes and phenotypes in G_2 . **(a)** Shown are frequencies of *tra2* genotypes (homozygous for the WT or the *tra2^{ts2}* allele, or heterozygous *tra2^{ts2}* mutants) within the number of analyzed individuals (n), found in phenotypic male or female G_2 offspring of family M8 and M11 inbreeding (ib). Both families are lacking homozygous *tra2^{ts2}* mutants with a female phenotype. **(b)** Karyotyping of phenotypic G_2 males via Y-chromosome specific PCR (primers P1504/P1505) on genomic DNA extracted from a single leg of family M8 and M11 offspring. A positive control PCR was performed on *tra2* with primers P1532/P1500 using the same DNA samples as in the Y-specific PCR, to exclude lack of PCR product due to DNA quality. Individuals lacking a signal in the Y-chromosome-specific PCRs but not in the genomic control PCR are marked in bold letters to indicate the XX-karyotype. M11ib_m29 was excluded from the analysis, due to low DNA quality. One phenotypic male (WT m) and female (WT f) from family M8 with WT *tra2* genotype are shown as controls. Displayed are cropped parts from different gels. Uncropped versions of the gels are provided in the supplement (Supplementary Fig. S5a and b). L = DNA ladder; kb = kilo base pairs. **(c)** Phenotypic male flies carrying the *tra2^{ts2}* mutation were dissected and compared to WT *EgII* flies to assess external and internal sexual organ formation. Shown are representative *tra2^{ts2}* homozygous XX (M8ib_m14, M8ib_m65) or XY (M8ib_m30) individuals as well as one XY male heterozygous for *tra2^{ts2}* (M8ib_m75). Black, dark and light grey boxes indicate the *tra2* genotype, with colors following the legend in (a). Mutants were not able to coil and store their distiphallus. Testes were normal or decolorized (M8ib_m75).

***tra2^{ts2}* homozygous XX embryos are transformed into phenotypic males at 19 °C.** The absence of phenotypic females homozygous for the *tra2^{ts2}* mutation in G₂ implied that XX embryos homozygous for *tra2^{ts2}* are either not viable or transformed into phenotypic males at 19 °C. Y-specific primers were used to assess the karyotype of 35 G₂ *tra2^{ts2}*-homozygous and 60 heterozygous male G₂ flies by PCR. In family M8, nine of 27 phenotypic males homozygous for *tra2^{ts2}* showed a signal in the control genomic PCR but not in the Y-chromosome-specific PCR, confirming the transformation of *tra2^{ts2}*-homozygous XX flies into phenotypic males. This also applied to two out of eight phenotypic males in family M11 (Fig. 3b). For one M11 offspring, M11ib_m29, no statement can be made as the control PCR failed to produce a signal. In contrast, all *tra2^{ts2}*-heterozygous males were positive for the Y-chromosome-specific PCR (Supplementary Fig. 6), excluding sex conversion as reason for the male-biased sex ratio in the G₂ heterozygotes.

Dissection of six XX- and four XY-karyotype males homozygous for *tra2^{ts2}*, and two XY *tra2^{ts2}*-heterozygous males (all G₂) showed that all *tra2^{ts2}*-homozygous males (XX and XY) had apparently normal or slightly decolorized testes. The two *tra2^{ts2}*-heterozygous males, in contrast, showed severely decolorized testes (Fig. 3c). In addition, across the G₂ offspring of both families, M8 and M11, 81.8% of the *tra2^{ts2}* homozygous XX males, 4.3% of the *tra2^{ts2}* homozygous XY males, and 16.6% of the *tra2^{ts2}* heterozygous XY males were not capable to coil and store their distiphallus (Fig. 3c). This phenotype was also observed in random samples of WT flies of different ages; while its penetrance in WT is higher at 19 °C (24.8%, n = 161) than at 26 °C (6.9%, n = 174), it is still markedly lower than observed in the *tra2^{ts2}* homozygous XX males (81.8%, n = 11) and might, therefore, be also an effect of the *ts2* mutation.

Rearing at lower temperature leads to low fertility rates. Based on the karyotyping experiments, 19 °C still is a restrictive temperature for the *ts2* mutation in *Ctra2*, contrary to *D. suzukii tra2^{ts2}* where 20 °C was permissive³⁸. Data from *D. melanogaster* suggests 16 °C as permissive temperature^{30,31}. However, medflies do not breed at such low temperatures, as the small-scale fertility tests at 16 °C had shown. To attain a permissive temperature for the medfly *tra2^{ts2}* mutation that does not affect breeding, the temperature was lowered to 18.5 °C, the mating threshold temperature⁴¹, for G₂ crossing and egg laying (Supplementary Table S1, Fig. S2c). *ts2*-homozygous XX and XY G₂ males were backcrossed to *EgII* females individually (13 crosses) or in groups (two crosses). *ts2*-heterozygous males and females were inbred (three crosses) or backcrossed (one group). Overall, during 13 days and 81 egg collections, more than 8,000 eggs were collected from these 19 crosses (Supplementary Table S4). A total of five larvae hatched from two egg collections of homozygous *tra2^{ts2}* XY male group-backcrosses, and only one survived to adulthood (M11ib_m1-het, Supplementary Table S4). Noteworthy, due to technical restrictions the temperature could not be kept constantly at 18.5 °C during the experiment, and these larvae hatched from a late egg collection (383 h; Supplementary Fig. S2c), prior to which the temperature had been above 18.5 °C for about two days. The male (G₃) was crossed to 40 *EgII* females but did not reproduce. Therefore, maintaining the *ts2* mutant strain by lowering the temperature to a permissive range was not possible.

Discussion

CRISPR/Cas9-HDR gene editing was used to create temperature-sensitive mutations in the *C. capitata* sex-determination gene *transformer-2*, equivalent to the two chemically induced point mutations in *D. melanogaster*^{30,31}. The *D. melanogaster tra2^{ts}* temperature-dependent sex-conversion phenotype promises great advantages for creating male-only populations needed for SIT programs, as it doubles the amount of male offspring per parental egg capacity, and only heat is needed for induction. Some countries do not regulate the use of organisms carrying CRISPR-induced SNPs as they could have also occurred by natural mutagenesis and selective breeding^{40,42}. Hence, only the *tra^{ts}* SNPs, but no exogenous DNA was inserted, to help facilitate a potential field release of *Ctra2^{ts}* strains. This was possible due to the high mutagenesis rate achieved with our previously published CRISPR/Cas9-HDR protocol³⁹, which we now successfully applied for the first time without using a visible phenotype.

The injections aiming at creating the *tra2^{ts1}* allele did not result in any mutant G₁ offspring at 19 °C, despite promising prerequisites; *ts1* gRNA and ssODN were functional in the preliminary tests at 26 °C, and the high number of G₀ adult survivors in the 19 °C injection increased the chance to obtain mutant offspring. Moreover, G₀ flies showed deformities of internal reproductive organs (Supplementary Fig. 3b). It can't be excluded, however, that these are the result of physical damage to the embryo caused by the injection. Possible reasons for the poor efficiency of the *ts1* knock-in could be the low on-target activity score of the *ts1* gRNA (0.045), or a stronger phenotypic impact of the *tra2^{ts1}* mutation compared to *tra2^{ts2}* as observed in *D. melanogaster*³⁰, which could reduce the chance to obtain viable *ts1* mutant flies. Testing of other *ts1* gRNAs could shed more light on possible reasons for the failure to create a stable *ts1* line; but considering the decreased viability in *D. melanogaster* and the permissive temperature issues in medfly, these experiments have little prospect for success.

In contrast, the *tra2^{ts2}* mutation could be introduced with high efficiency, detectable already from the absence of phenotypic females and the appearance of intersexes in G₀, in the frequency of HDR-positive fertile G₀ (100% at 19 °C), as well as in the high penetrance of the mutant genotype within their G₁ offspring (83% and 75% knock-in for family M8 and M11, respectively). This matches the higher on-target activity score of the *ts2* gRNA (0.140).

The observed overall higher survival rate of injected G₀ at 19 °C compared to 26 °C might be the result of a lower Cas9 editing activity⁴³ and a potentially associated off-target rate, but could also be connected to the reduced speed of embryonic development allowing more time for repair mechanisms to fix injection-induced damage to the embryo⁴⁴, which is unrelated to Cas9 editing. Extensive comparative injections would be needed to answer this question.

The lack of phenotypic females homozygous for *tra2^{ts2}* and the conversion of XX embryos into phenotypic males at 19 °C suggests that this is still a restrictive temperature for the *Ctra2^{ts2}* mutation, which does not allow correct protein folding, and indicates the importance of this position in the highly conserved TRA2 linker region

for correct protein conformation. This observation is in line with the results obtained for the *D. melanogaster tra2^{ts2}* mutation, where the temperature had to be lowered to 16 °C to generate fertile males and females, while 18 °C produced sterile males and females, and 29 °C resulted in sterile males and pseudomale-like intersexes³⁰. A further reduction of the temperature to an average of 18.3 °C, however, resulted in a loss of the strain due to mainly unviable eggs deposited by the G₂ generation. This was not unexpected since our small-scale tests with WT at 18 °C and 16 °C produced very little or no viable offspring, respectively. Furthermore, some males were not capable of coiling and storing their distiphallus. While this phenotype was also observed in WT males at low temperatures, it seems to be enhanced by the *tra2* mutant allele. However, the numbers are too small for a robust statement. Fertility and mating behaviour of this phenotype have not been assessed.

While the mean survivorship of medfly egg and larval stages at 15, 20, 25 and 30 °C has been reported to not differ significantly²¹, and the described threshold for ovarian maturation with 8.1 °C to 16.6 °C^{21,41,45} is also below the tested *Ctra2^{ts2}* permissive temperature of 18.3 °C, Prokopy and Hendrichs⁴¹ showed that 18.5 °C is the temperature threshold for mating in medfly. During the cross of *tra2^{ts2}* G₂ flies, temperatures were above the threshold mainly during the first days (1–72 h) and last days (337–517 h) of the crossing (Supplementary Table S1, Fig. S2c). As ovarian maturation takes up to 10 days at this temperature, and crosses have been set up with 3–5 d old flies, no successful mating could have been achieved during the first period above 18.5 °C. During the main egg collection period (72–336 h), temperature was mainly below 18.5 °C (Supplementary Fig. S2c). The successful mating appeared within the second period of exceeded temperature. A possible explanation for the loss of the *tra2^{ts2}* strain therefore is that the low temperatures prevented mating and eggs have not been fertilized until temperatures had exceeded 18.5 °C for at least 2 d. On the other hand, control crosses of *EgII* flies managed to produce a small amount of offspring at temperatures mainly lower than or equal to 18.5 °C (2,796 collected eggs, 16 larvae, 8 adults; Supplementary Fig. S2c), showing that low mating activity is taking place at or below the threshold. Therefore, it is possible that the *ts2* mutation, even in the heterozygous state, affects the fertility of the flies at temperatures lower than 18.5 °C. However, as numbers are very small, no robust statement is possible. Overall, using the *EgII* background for the *tra2^{ts}* experiments, it could not be determined if the permissive temperature for the medfly *tra2^{ts2}* mutation is lower than 18.5 °C or if the *ts2* mutant phenotype in medfly is not temperature-dependent at all.

As strains with different genetic backgrounds can have markedly different sensitivities for elevated or low temperatures due to adaptation mechanisms, using another medfly WT background might allow to investigate lower permissive temperatures for *ts2*. It might also be possible to induce cold acclimation in a WT strain by successively reducing the rearing temperature over several generations before generating the *tra2^{ts2}* mutation. This strategy would fail, however, if there is no acclimation with respect to the mating threshold, as shown for *B. tryoni*⁴⁶.

Moreover, with regard to the use of the *tra2^{ts2}* mutation for medfly sexing in a mass-rearing facility, the presumably low (< 18.5 °C) permissive temperature of the medfly *ts2* mutation would be problematic, as temperature and development time show a linear relation. At 19 °C, for example, the development from egg to adults takes about 32.7 d plus 9 d for ovarian maturation, compared to 17.4 d plus 5.3 d at 26 °C²¹. The even longer development times at < 18 °C would thus be problematic for the production scale and the cost-effectiveness of a mass-rearing program and investigations into lower temperatures would thus certainly not be relevant for insect pest control applications in medfly.

In conclusion, we demonstrated the successful creation of the *D. melanogaster tra2^{ts2}* point mutation in *C. capitata* via markerless CRISPR/Cas9-HDR gene editing and the importance of the respective amino acid for the correct function of TRA2 in the female sex-determination. The previously shown high HDR efficiency in medfly using a ssODN repair template to convert the marker gene eGFP (enhanced green fluorescent protein) into BFP (blue), could be confirmed in this study, where we achieved 100% knock-in efficiency (2 out of 2 fertile G₀) compared to 86% (6 out of 7 fertile G₀) in the previous study³⁹. Also, the high penetrance of mutant offspring within the G₁ with 75–83% in this study is similar compared to 90% in the previous one. It was not possible, however, to identify a permissive temperature at which the *tra2^{ts2}* mutation does not affect female development, as it would be located below the mating threshold of medfly. Therefore, it could not be determined if we hadn't reached the permissive temperature yet, or if the *tra2^{ts2}* phenotype in medfly, in contrast to *Drosophila*, is not temperature dependent. Based on the data presented here, a medfly sexing strain built solely on *tra2^{ts2}* would be unsuitable for an SIT program and mass-rearing, either because the rearing would be too slow to be productive on a large scale, or because the sex conversion could not be switched off for strain maintenance. Other possibilities to create a sex-conversion system in medfly could be to target other sex-determination genes, like *transformer*^{27,29,47}, or to force (over)expression of the maleness factor *MoY*, which induces masculinization in XX embryos⁴⁸. However, conditionality would need to be engineered for both options.

Material and methods

Rearing conditions. *Ceratitidis capitata* wild-type *Egypt-II* (*EgII*) flies were received from the FAO/IAEA Agriculture and Biotechnology Laboratory, Austria, and kept at 26 °C, 48% RH and 14/10 h light/dark cycle. For fertility tests, freshly eclosed *EgII* adult flies were transferred from 26 °C to 19.5 °C, 60% RH, 24 h light or 16 or 18 °C, 46–48% RH, 24 h light, where egg collections and subsequent rearing took place. *tra2^{ts}* mutants were kept at 19 °C or 18.5 °C, 60% RH, and 24 h light. Temperature and humidity were measured every five minutes of the experiment using an EL-USB-2 data logger (Lascar electronics, measurement precision for temperature ± 1 °C, for humidity ± 3%). Readout of the data logger showed that during the rearing of the *tra2^{ts2}* mutants, short-term variations of the temperature (+3 °C/–1 °C) occurred (see Supplementary Table S1 and Fig. S2). These could not be avoided due to technical restrictions of the experimental setup. Furthermore, the targeted temperature (19 °C) was once exceeded for 3.5 h up to a temperature of 25 °C during an outage of the air conditioning sys-

tem. This occurred during the late larval or pupal stage of *tra2^{ts2}* G₁. Feeding and screening conditions were as described in Aumann et al.³⁹.

CRISPR/Cas9 gene editing. Design of gRNAs targeting *tra2* (gRNA_tra2_ts1 and gRNA_tra2_ts2) and assessment of potential off-target effects was performed using the *C. capitata* genome version Ccap 2.1 (GCF_000347755.3, NCBI)⁴⁹ and the Software Package Geneious Prime⁵⁰. On-target activity score was 0.045 for gRNA_tra2_ts1, and 0.140 for gRNA_tra2_ts2 (scores are between 0 and 1; 1 = highest expected activity⁵⁰). Both gRNAs showed zero off-targets sites in the medfly genome. gRNA synthesis, in vitro transcription and purification was performed as described before³⁹, using primers P_1439 (GAAATTAATACGACTCACTATAGGTGA TGATATAGCTGATGCTAGT TTTAGAGCTAGAAATAGC) and P_369 (GCACCGACTCGGTGCCACTTT TCAAGTTGATAACGGACTAGCCTTATTTAACTTGCTATTTCTAGCTCTAAAAC) for gRNA_tra2_ts1 and primers P_1440 (GAAATTAATACGACTCACTATAGGCCATATAAACGCCAGGTGTGTTT TAGAGCTAGAAATAGC) and P_369 for gRNA_tra2_ts2. The sequences of the 140 bp single-stranded HDR templates 'ssODN_tra2_ts1' and 'ssODN_tra2_ts2' (EXTREMer oligo, Eurofins Genomics) were: ssODN_tra2_ts1 (sense): TGAGTAATCTACGCGTATGCGTCGATCATCGATTCCATGCCGGAACATGCGTCCTTGGCTGCTTTA ACATCAGCTATATCATCATAATAGATAAAGCAAAAGCCACGAGATCGGCCAGTCTGAAAAAAGA AAAAAATAG; ssODN_tra2_ts2 (antisense): AAACGATTTAAATCACATGCACATGCCAAGTATACCTTG TGTGTCGTCATATAAACGCCAGGTGTGGAAGTGTGTGGTCTCTGTGTAGTTAGTAATCTACGCG TATGCGTCGATCATCGATTCCATGCCGGAACAT; Base changes introducing the *ts1* or *ts2* mutation are shown in bold. Purified Cas9 protein (PNA Bio Inc.) was reconstituted to 1 µg/µl in 20 mM Hepes, 150 mM KCl, 2% sucrose and 1 mM DTT (pH 7.5).

Microinjection of embryos: 10 µl injection mix for knock-out experiments contained 360 ng/µl Cas9 protein and 200 ng/µl gRNA_tra2_ts1 or gRNA_tra2_ts2 in 300 mM KCl^{39,51}. For knock-in experiments, 200 ng/µl ssODN_tra2_ts1 or ssODN_tra2_ts2 were added to the mix. The mixes were freshly prepared on ice, incubated at 37 °C for 10 min to allow pre-assembly of gRNA-Cas9 ribonucleoprotein complexes and stored on ice prior to injections. For microinjection of WT *C. capitata* embryos, eggs were collected over a 30–50 min period, prepared for injection and handled afterwards as previously described³⁹. Injections were performed using siliconized quartz glass needles (Q100-70–7.5; LOT171381; Science Products, Hofheim, Germany), drawn out on a Sutter P-2000 laser-based micropipette puller. Injection equipment consisted of a manual micromanipulator (MN-151, Narishige), an Eppendorf FemtoJet 4i microinjector, and an Olympus SZX12-TTR microscope (SDF PLAPO 1xPF objective). Injection survivors were numbered successively across ts1 injections and ts2 injections, respectively.

Crossing strategies and dissection of internal reproductive organs. *Crossing of G₀:* Each G₀ adult injection survivor was individually crossed to three *EgII* WT males or virgin females, except for the 19 °C *ts1* knock-in injection. Here, six males and six females were individually backcrossed, the remaining flies were group-backcrossed (five G₀ males to 15 females, ten G₀ females to ten WT males, and six G₀ females to nine WT males). Eggs were collected three to five times, with an interval of one to two days. For the 19 °C knock-in experiments, G₁ and G₂ flies (if applicable) were kept individually until their genotype was assessed via non-lethal genotyping.

Crossing of tra2^{ts2} G₁: males and females heterozygous for *tra2^{ts2}* were inbred. Additionally, heterozygous males were backcrossed (Supplementary Table S3). Eggs were collected six times, with an interval of one to two days.

Crossing of tra2^{ts2} G₂: phenotypic males and females heterozygous for *tra2^{ts2}* were inbred (Supplementary Table S4). Additionally, four *tra2^{ts2}* heterozygous XY males, not capable of coiling and storing their distiphallus, were group backcrossed. *tra2^{ts2}* homozygous XY males were either backcrossed or crossed with heterozygous *tra2^{ts2}* females (Supplementary Table S4). Nine males homozygous for *tra2^{ts2}* with XX-karyotype, all not able to coil and store their distiphallus, were individually backcrossed to four females each. Eggs of the G₂ crosses were collected four to seven times over seven to 13 days (Supplementary Table S4).

Dissections: G₀ flies and single crossed G₂ flies were allowed to mate for 5–10 days (G₀) or 7–13 days (G₂) days. If still alive, they were then dissected to examine their internal reproductive organs.

Molecular analyses of G₀ mosaics. To analyse the mosaic genotype of G₀ flies, DNA was extracted from single flies according to a standard protocol. The target region encompassing the *ts1* and *ts2* mutant sites (1213 bp) was amplified using the *tra2*-specific primers P1401 (TGCTTGGTGGTCCGCAAATA) and P1500 (TGTGCATATACTAAAGGCTCTCCC), 50–100 ng DNA, and the Q5 High-fidelity DNA polymerase (New England Biolabs) according to the manufacturer's protocol in a Bio-Rad C1000 Touch thermal cycler [98 °C, 1 min; 35 cycles of (98 °C, 15 s; 56 °C, 30 s; 72 °C, 45 s); 72 °C, 2 min]. PCR fragments were purified using the Zymo Research DNA Clean & Concentrator⁵ kit and subcloned into the pCR4-blunt TOPO vector (Invitrogen) for sequencing. Three to five clones were sequenced using primer mfs13 (TGTA AACGACGCCAGT) (Macrogen Europe, Amsterdam) for each analysed fly. Verification of CRISPR-induced mutations from the sequencing results was performed using the Software Package Geneious Prime⁵⁰ by mapping the sequencing results to the *tra2* reference sequence (Gene ID: 101,452,698).

Non-lethal genotyping of G₁ and G₂ flies. To identify the *tra2* genotype of G₁ and G₂ flies, non-lethal genotyping was performed using an adapted version of the protocol established by Carvalho et al.⁵². A single leg of an anesthetized fly was cut at the proximal femur using scissors and homogenized in 50 µl buffer (10 mM Tris-Cl pH 8.2, 1 mM EDTA, 25 mM NaCl) for 15 s (6 m/s) using ceramic beads and a FastPrep-24 5G homogenizer (M.P. Biomedicals). 28.3 µl buffer mixed with 1.7 µl proteinase-K (2.5 U/mg) were added and incubated for 1 h at 37 °C. The reaction was stopped 4 min at 98 °C. The solution was cooled down on ice and directly

used as PCR template to amplify the region surrounding the *tra2* target site. A 25 μ l PCR reaction contained DreamTaq polymerase and buffer (Life Technologies), dNTPs and the *tra2*-specific primers P1401 and P1500 according to the manufacturer's instructions, and 3.75 μ l template solution in a Bio-Rad C1000 Touch thermal cycler [95 °C, 3 min; 35 cycles of (95 °C, 30 s; 56 °C, 30 s; 72 °C, 1 min); 72 °C, 5 min]. The size of the PCR product (1,213 bp) was verified on an agarose gel. PCR products were purified using the Zymo Research DNA Clean & Concentrator⁵ kit, sequenced using primer P1500, and subsequently analysed using Software Package Geneious Prime⁵⁰.

Molecular karyotyping-Y chromosome specific PCR. Y-specific repetitive elements were amplified from genomic DNA extracted either from a single fly (G_0) or a single leg (G_2) using the published Y-specific oligonucleotides P1504_Y-spec1 (TACGCTACGAATAACGAATTGG) and P1505_Y-spec2 (GCGTTTAAATATACAAATGTGTG)⁵³. 10 μ l PCR reactions contained either 50 ng DNA (single fly) or 3.75 μ l single-leg DNA template solution, and the Y-specific primers and DreamTaq PCR components as described above. PCR cycling conditions (Bio-Rad C1000 Touch) were [95 °C, 3 min; 35 cycles of (95 °C, 30 s; 58 °C, 30 s; 72 °C, 1 min); 72 °C, 5 min]. Absence of a PCR product was interpreted as absence of the Y chromosome (XX-karyotype). The same PCR conditions with primers P1532 (AGTGAAAACGATTAAATCACATGCAC) and P1500 for genomic DNA extracted from a single-leg, or P1401 and P1500 for DNA extracted from a single fly were used to amplify 328 bp or 1,213 bp fragments, respectively of *tra2* as a positive control PCR to confirm sufficient quality of extracted genomic DNA.

Equipment and settings for image acquisition. For bright field image acquisition of flies (either dead or anesthetized with CO₂ and placed on a 4 °C cooler) was carried out using a fully automated Leica M205FC stereo microscope with a PLANAPO 1.0 \times objective, a Leica DFC7000 T camera and the Leica LAS X 3.4.2.18368 software. To enhance screen and print display of the pictures the image processing software Fiji ImageJ Version 2.0.0⁵⁴ was used to apply moderate changes to image brightness and contrast. Changes were applied equally throughout the entire image and across all images.

Data availability

All data generated or analysed is included in this article or the supplement.

Received: 9 June 2020; Accepted: 13 October 2020

Published online: 29 October 2020

References

1. Knipling, E. F. Possibilities of insect control or eradication through the use of sexually sterile males. *J. Econ. Entomol.* **48**, 459–462 (1955).
2. Hendrichs, J., Franz, G. & Rendon, P. Increased effectiveness and applicability of the sterile insect technique through male-only releases for control of Mediterranean fruit-flies during fruiting seasons. *J. Appl. Entomol.* **119**, 371–377 (1995).
3. Vreysen, M. J. Principles of area-wide integrated tsetse fly control using the sterile insect technique. *Med. Trop. (Mars)* **61**, 397–411 (2001).
4. Ito, Y., Kakinohana, H., Yamagishi, M. & Kohama, T. Eradication of the Melon Fly, *Bactrocera cucurbitae*, from Okinawa, Japan, by means of the sterile insect technique, with special emphasis on the role of basic studies. *J. Asia-Pacif. Entomol.* **6**, 119–129 (2003).
5. Gilles, J. R. *et al.* Towards mosquito sterile insect technique programmes: exploring genetic, molecular, mechanical and behavioural methods of sex separation in mosquitoes. *Acta Trop.* **132**(Suppl), S178–187 (2014).
6. Rendon, P., McInnis, D., Lance, D. & Stewart, J. Medfly (Diptera: Tephritidae) genetic sexing: large-scale field comparison of males-only and bisexual sterile fly releases in Guatemala. *J. Econ. Entomol.* **97**, 1547–1553 (2004).
7. Lutrat, C. *et al.* Sex sorting for pest control: It's raining men!. *Trends Parasitol.* **35**, 649–662 (2019).
8. Yamada, H. *et al.* Genetic sex separation of the malaria vector, *Anopheles arabiensis*, by exposing eggs to dielidrin. *Malar. J.* **11**, 208 (2012).
9. Orozco, D., Meza, J. S., Zepeda, S., Solis, E. & Quintero-Fong, J. L. Tapachula-7, a new genetic sexing strain of the Mexican fruit fly (Diptera: Tephritidae): sexual compatibility and competitiveness. *J. Econ. Entomol.* **106**, 735–741 (2013).
10. McInnis, D. O. *et al.* Development of a pupal color-based genetic sexing strain of the Melon fly, *Bactrocera cucurbitae* (Coquillett) (Diptera: Tephritidae). *Ann. Entomol. Soc. Am.* **97**, 1026–1033 (2004).
11. Franz, G. Genetic sexing strains in Mediterranean fruit fly, an example for other species amenable to large-scale rearing for the Sterile Insect Technique. In *Sterile Insect Technique—Principles And Practice In Area-Wide Integrated Pest Management* (eds Dyck, V. A. *et al.*) 427–451 (Springer, Berlin, 2005).
12. Schetelig, M. F. & Handler, A. M. A transgenic embryonic sexing system for *Anastrepha suspensa* (Diptera: Tephritidae). *Insect. Biochem. Mol. Biol.* **42**, 790–795 (2012).
13. Fu, G. *et al.* Female-specific insect lethality engineered using alternative splicing. *Nat. Biotechnol.* **25**, 353–357 (2007).
14. Concha, C. *et al.* A transgenic male-only strain of the New World screwworm for an improved control program using the sterile insect technique. *BMC Biol.* **14**, 72 (2016).
15. Schetelig, M. F., Targovska, A., Meza, J. S., Bourtzis, K. & Handler, A. M. Tetracycline-suppressible female lethality and sterility in the Mexican fruit fly, *Anastrepha ludens*. *Insect. Mol. Biol.* **25**, 500–508 (2016).
16. Ant, T. *et al.* Control of the olive fruit fly using genetics-enhanced sterile insect technique. *BMC Biol.* **10**, 51 (2012).
17. Jin, L. *et al.* Engineered female-specific lethality for control of pest Lepidoptera. *ACS Synth. Biol.* **2**, 160–166 (2013).
18. Malacrida, A. R. *et al.* Globalization and fruitfly invasion and expansion: the medfly paradigm. *Genetica* **131**, 1–9 (2007).
19. Arias, M. B., Elfekih, S. & Vogler, A. P. Population genetics and migration pathways of the Mediterranean fruit fly *Ceratitis capitata* inferred with coalescent methods. *PeerJ* **6**, e5340–e5340 (2018).
20. Ogaugwu, C. E., Schetelig, M. F. & Wimmer, E. A. Transgenic sexing system for *Ceratitis capitata* (Diptera: Tephritidae) based on female-specific embryonic lethality. *Insect. Biochem. Mol. Biol.* **43**, 1–8 (2013).
21. Duyck, P. F. & Quilici, S. Survival and development of different life stages of three *Ceratitis spp.* (Diptera: Tephritidae) reared at five constant temperatures. *Bull. Entomol. Res.* **92**, 461–469 (2007).
22. Mumford, J. D. Application of benefit/cost analysis to insect pest control using the Sterile Insect Technique. In *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management* (eds Dyck, V. A. *et al.*) 481–498 (Springer, Berlin, 2005).

23. Salvemini, M. *et al.* *Ceratitis capitata transformer-2* gene is required to establish and maintain the autoregulation of *Cctra*, the master gene for female sex determination. *Int. J. Dev. Biol.* **53**, 109–120 (2009).
24. Gomulski, L. M. *et al.* Gene discovery in an invasive tephritid model pest species, the Mediterranean fruit fly *Ceratitis capitata*. *BMC Genom.* **9**, 243 (2008).
25. Pane, A., Salvemini, M., Delli Bovi, P., Polito, C. & Saccone, G. The *transformer* gene in *Ceratitis capitata* provides a genetic basis for selecting and remembering the sexual fate. *Development* **129**, 3715–3725 (2002).
26. Saccone, G., Salvemini, M. & Polito, L. C. The *transformer* gene of *Ceratitis capitata*: a paradigm for a conserved epigenetic master regulator of sex determination in insects. *Genetica* **139**, 99–111 (2011).
27. Schetelig, M. F., Milano, A., Saccone, G. & Handler, A. M. Male only progeny in *Anastrepha suspensa* by RNAi-induced sex reversion of chromosomal females. *Insect. Biochem. Mol. Biol.* **42**, 51–57 (2012).
28. Thongsaklaing, T., Nipitwattanaphon, M. & Ngernsiri, L. The *transformer2* gene of the Pumpkin fruit fly, *Bactrocera tau* (Walker), functions in sex determination, male fertility and testis development. *Insect. Mol. Biol.* **27**, 766–779 (2018).
29. Liu, G., Wu, Q., Li, J., Zhang, G. & Wan, F. RNAi-mediated knock-down of *transformer* and *transformer 2* to generate male-only progeny in the Oriental fruit fly, *Bactrocera dorsalis* (Hendel). *PLoS ONE* **10**, e0128892 (2015).
30. Amrein, H., Maniatis, T. & Nöthiger, R. Alternatively spliced transcripts of the sex-determining gene *tra-2* of *Drosophila* encode functional proteins of different size. *EMBO J.* **9**, 3619–3629 (1990).
31. Belote, J. M. & Baker, B. S. Sex determination in *Drosophila melanogaster*: analysis of *transformer-2*, a sex-transforming locus. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1568–1572 (1982).
32. Chandler, D. *et al.* Evolutionary conservation of regulatory strategies for the sex determination factor *transformer-2*. *Mol. Cell. Biol.* **17**, 2908–2919 (1997).
33. Wang, Y. *et al.* Molecular cloning, expression pattern analysis, and in situ hybridization of a *transformer-2* gene in the oriental freshwater prawn, *Macrobrachium nipponense* (de Haan, 1849). *Biotech* **9**, 205 (2019).
34. Martin, I., Ruiz, M. F. & Sanchez, L. The gene *transformer-2* of *Sciara* (Diptera, Nematocera) and its effect on *Drosophila* sexual development. *BMC Dev. Biol.* **11**, 19 (2011).
35. Li, X. *et al.* Two of the three *transformer-2* genes are required for ovarian development in *Aedes albopictus*. *Insect. Biochem. Mol. Biol.* **109**, 92–105 (2019).
36. Dauwalder, B., Amaya-Manzanares, F. & Mattox, W. A human homologue of the *Drosophila* sex determination factor *transformer-2* has conserved splicing regulatory functions. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 9004–9009 (1996).
37. Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096 (2014).
38. Li, J. & Handler, A. M. Temperature-dependent sex-reversal by a *transformer-2* gene-edited mutation in the spotted wing drosophila, *Drosophila suzukii*. *Sci. Rep.* **7**, 12363 (2017).
39. Aumann, R. A., Schetelig, M. F. & Häcker, I. Highly efficient genome editing by homology-directed repair using Cas9 protein in *Ceratitis capitata*. *Insect. Biochem. Mol. Biol.* **101**, 85–93 (2018).
40. Waltz, E. CRISPR-edited crops free to enter market, skip regulation. *Nat. Biotechnol.* **34**, 582 (2016).
41. Prokopy, R. J. & Hendrichs, J. Mating behavior of *Ceratitis capitata* on a field-caged host tree. *Ann. Entomol. Soc. Am.* **72**, 642–648 (1979).
42. Hamburger, D. Comparative analysis: the regulation of plants derived from Genome Editing in Argentina, Australia, Canada, the European Union, Japan and the United States. In *Regulation of Genome Editing in Plant Biotechnology: A Comparative Analysis of Regulatory Frameworks of Selected Countries and the EU* (eds Dederer, H.-G. & Hamburger, D.) 313–363 (Springer, Berlin, 2019).
43. Xiang, G., Zhang, X., An, C., Cheng, C. & Wang, H. Temperature effect on CRISPR-Cas9 mediated genome editing. *J. Genet. Genom.* **44**, 199–205 (2017).
44. Handler, A. M. & O'Brochta, D. A. Ch. 10: Transposable elements for insect transformation. In *Insect Development: Morphogenesis, Molting and Metamorphosis* (ed. Gilbert, L. I.) 459–496 (Elsevier BV, Amsterdam, 2009).
45. Tassan, R. L. *et al.* Mediterranean fruit fly life cycle estimations for the California eradication program. In *Fruit Flies of Economic Importance* (ed. Cavalloro, R.) 564–570 (Balkema, Amsterdam, 1983).
46. Meats, A. & Fay, H. A. C. The effect of acclimation on mating frequency and mating competitiveness in the Queensland fruit fly, *Dacus tryoni*, in optimal and cool mating regimes. *Physiol. Entomol.* **1**, 207–212 (1976).
47. Shukla, J. N. & Palli, S. R. Sex determination in beetles: Production of all male progeny by parental RNAi knockdown of *transformer*. *Sci. Rep.* **2**, 602 (2012).
48. Meccariello, A. *et al.* *Maleness-on-the-Y (MoY)* orchestrates male sex determination in major agricultural fruit fly pests. *Science* **365**, 1457–1460 (2019).
49. Papanicolaou, A. *et al.* The whole genome sequence of the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann), reveals insights into the biology and adaptive evolution of a highly invasive pest species. *Genome Biol.* **17**, 192 (2016).
50. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
51. Burger, A. *et al.* Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* **143**, 2025–2037 (2016).
52. Carvalho, G. B., Ja, W. W. & Benzer, S. Non-lethal genotyping of single *Drosophila*. *Biotechniques* **46**, 312–314 (2009).
53. Anleitner, J. E. & Haymer, D. S. Y enriched and Y specific DNA sequences from the genome of the Mediterranean fruit fly *Ceratitis capitata*. *Chromosoma* **101**, 271–278 (1992).
54. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

Acknowledgements

We wish to thank Jakob Martin, Tanja Rehling, Johanna Rühl, and Julia Hehner for technical assistance and help with insect rearing.

Author contributions

R.A.A. performed research; R.A.A., M.F.S. and I.H. designed research; R.A.A., M.F.S. and I.H. analysed data; and R.A.A., I.H. and M.F.S. wrote the paper. All authors have read and approved the final version of the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This study benefitted from discussions at meetings for the Coordinated Research Project, “Comparing Rearing Efficiency and Competitiveness of Sterile Male Strains Produced by Genetic, Transgenic or Symbiont-based Technologies”, funded by the International Atomic Energy Agency (IAEA). This work was supported by projects of the LOEWE Centre for Insect Biotechnology & Bioresources and the LOEWE Centre DRUID of the Hessian Ministry of Science and Arts (MFS), and has also been funded by the Hort Frontiers Fruit Fly Fund part of the Hort Frontiers strategic partnership initiative

developed by Hort Innovation, with co-investment from Macquarie University, USDA, and JLU Gießen and contributions from the Australian Government (FF17000 to MFS). The funding sources were not involved in any of the following: study design, collection, analysis and interpretation of data, writing of the report, and the decision to submit the article for publication.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-75572-x>.

Correspondence and requests for materials should be addressed to M.F.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

Supplementary Information for

Female-to-male sex conversion in *Ceratitis capitata* by CRISPR/Cas9 HDR-induced point mutations in the sex determination gene *transformer-2*

Roswitha A. Aumann, Irina Häcker, Marc F. Schetelig*

Justus-Liebig-University Gießen, Institute for Insect Biotechnology, Department of Insect Biotechnology in Plant Protection, Winchesterstr. 2, 35394 Gießen, Germany

*Corresponding author: marc.schetelig@agrar.uni-giessen.de

This PDF file includes:

- Supplementary Methods
- Supplementary Figures S1-S5
- Supplementary Tables S1-S4
- Supplementary References

Supplementary Methods

PCR reactions shown in Supplementary Figure S5 were done in a 10 µl reaction volume containing 3.75 µl single-leg DNA template solution, the *tra2*- or Y-chromosome-specific primers, and DreamTaq PCR components according to the manufacturer's protocol. PCR cycling conditions (Bio-Rad C1000 Touch) were [95°C, 3 min; 35 cycles of (95°C, 30 s; 58°C, 30 s; 72°C, 1 min); 72°C, 5 min]. Gel pictures were taken with a VersaDoc MP Molecular Imager (BioRad).

PCR reactions shown in Supplementary Figure S6 were done in a 10 µl reaction volume containing 2.5 µl single-leg DNA template solution, the Y-specific primers and DreamTaq PCR components according to the manufacturer's protocol. PCR cycling conditions (Bio-Rad C1000 Touch) were [95°C, 3 min; 40 cycles of (95°C, 30 s; 58°C, 30 s; 72°C, 1 min); 72°C, 5 min]. Gel pictures were taken with a Gel iX imager (Intas, Göttingen).

a	gRNA_tra2_ts1				
	T G A T G A T A T A G C T G A T G C T A A G G C	Individual	Injection	Mutation	
-.....	F1, G ₀	KO, 26°C	1 bp deletion	
-.....	F2, G ₀	KI, 26°C	6 bp deletion	
 T A	F2, G ₀	KI, 26°C	<i>tra2^{ts1}</i> KI	
b	gRNA_tra2_ts2				
	T C C C A C A C C T G G C G T T T A T A T G G G	Individual	Injection	Mutation	
-.....	IS1-KO, G ₀	KO, 26°C	33 bp deletion	
-.....	M5, G ₀	KO, 26°C	9 bp deletion	
-.....	M6, G ₀	KO, 26°C	4 bp insertion	
					
-.....	IS1, G ₀	KI, 26°C	4 bp deletion	
- G	IS4, G ₀	KI, 26°C	7 bp deletion, SNP	
-.....	IS5, G ₀	KI, 26°C	4 bp deletion	
 T	M8m3, G ₁	KI, 19°C	<i>tra2^{ts2}</i> KI	

Fig. S1. HDR and NHEJ events confirm ts1 and ts2 gRNA functionality. Sequences of mutant *tra2^{ts1}* (a) and *tra2^{ts2}* (b) alleles identified in G₀ or G₁ individuals compared to the *tra2* reference sequence. The consensus is shown as dots, knock-in (KI) mutant sites in red uppercase letters, NHEJ induced SNP as uppercase letters, deletions as dashes, insertions as a triangle. The identity of the analysed fly (F = female, M = male, IS = intersex), the injection type (knock-in = KI, knock-out = KO) and rearing temperature, as well as the mutation event are indicated on the right.

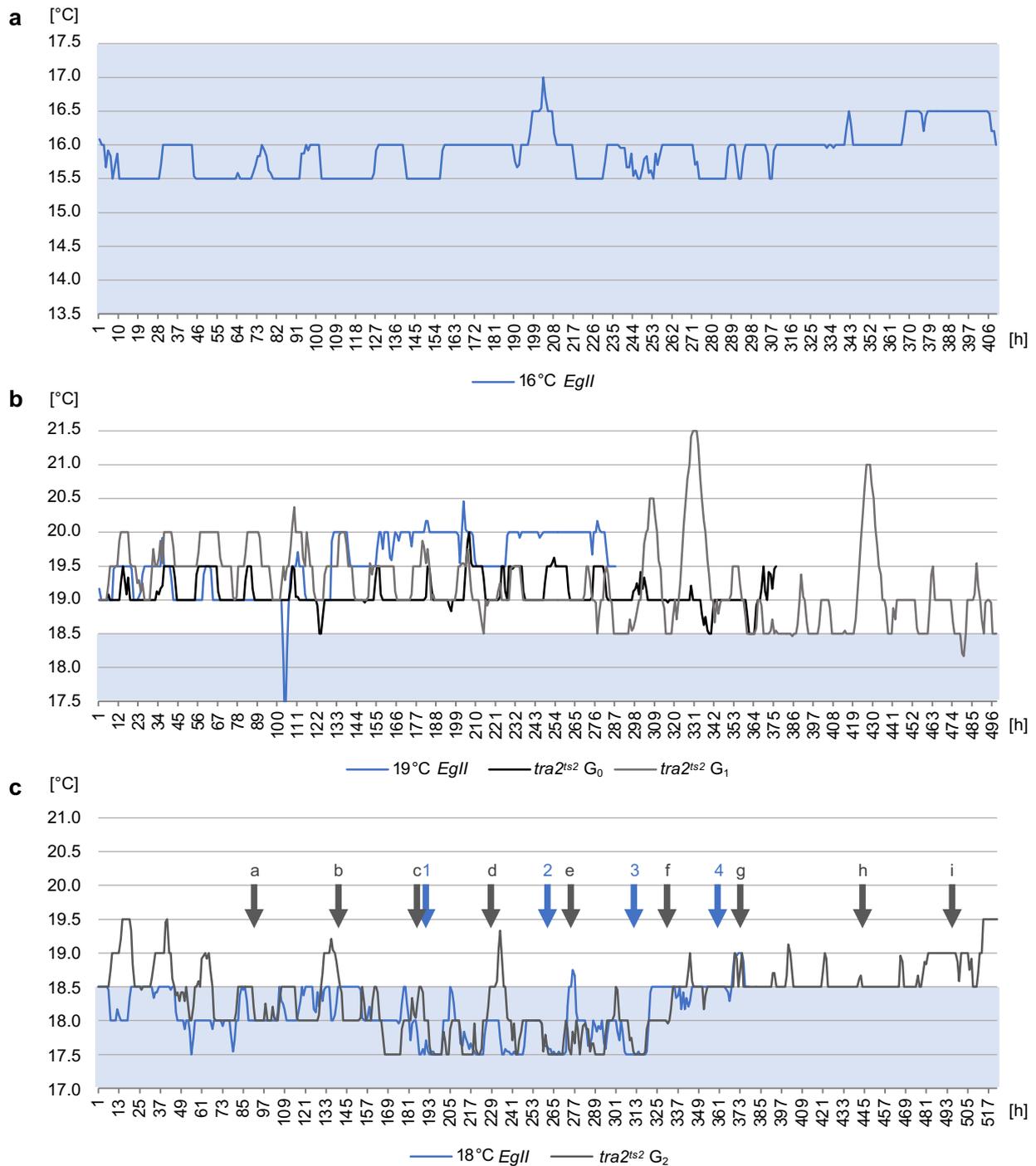


Fig. S2. Temperature profiles during *EglI* control and *tra2^{ts2}* crosses. Shown are the temperature profiles [°C] for the duration of the cross (from set up until the day of the last egg collection, in hours = h), for the wild type control cross at 16°C (**a**), the cross of *tra2^{ts2}* G₀ and G₁ flies as well as a *EglI* control at 19°C (**b**) and the cross of *tra2^{ts2}* G₂ flies and a control at 18°C (**c**). Crosses have been set up either with newly eclosed flies (all controls, G₀ cross), or with 3-5 d old flies, if flies needed to be genotyped first (G₁ and G₂ crosses). Temperature was recorded every 5 min. For a clearer visualization of the recordings, the average temperature per hour (average of 12 recording timepoints) is displayed here. Egg collection timepoints of *tra2^{ts2}* G₂ and control crosses are indicated by dark grey (a-i) or blue (1-4) arrows, respectively. For *tra2^{ts2}* G₂, larvae hatched only once, on egg collection timepoint 'g', from two crosses, the group crosses No. 17 and 18 (see Supplementary table S4). For the control, 16 larvae hatched from egg collection timepoints 2, 3 and 4 from overall 2,796 collected eggs. Temperatures below the published medfly mating-threshold (18.5°C; Prokopy and Hendrichs, 1979) are shaded in blue.

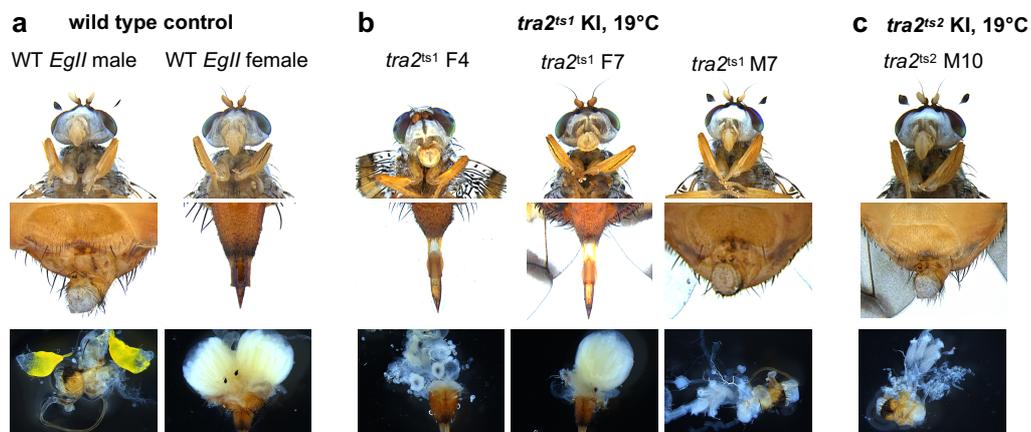


Fig. S3. CRISPR/Cas9 HDR injections at 19°C targeting *tra2^{ts1}* or *tra2^{ts2}* create G₀ flies with abnormal internal sexual organs. Shown are the head, with or without male-specific bristles (top row), the abdomen (middle row), and the internal reproductive organs (bottom row). **(a)** wild type control male and female fly. **(b)** *tra2^{ts1}* KI injection survivors did not show any external abnormalities, however, internal reproductive organs of three out of twelve dissected individuals were distorted; for example, F4 showed no ovaries but spermatheca, F7 had one developed ovary, M7 had no testes (F = female, M = male). **(c)** One out of four dissected male survivors of the *tra2^{ts2}* KI injection (19°C) showed internal abnormalities of sexual organs (no testes).

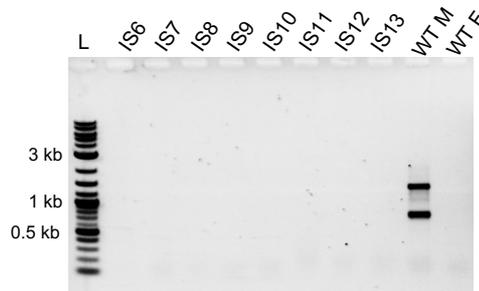
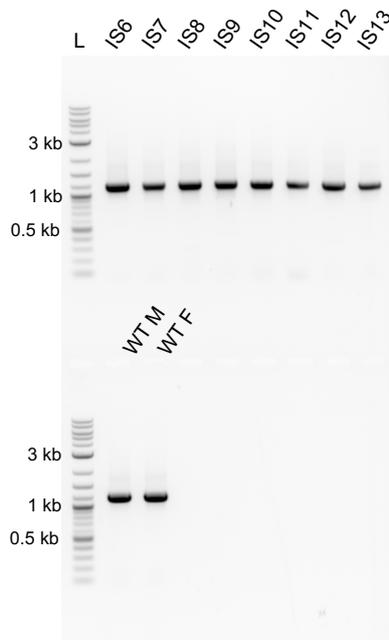
a Y-chromosome specific PCRs**b** genomic control PCRs

Fig. S4. Karyotyping of *tra2^{ts2}* injected G_0 intersex flies reveals partial masculinization of XX embryos. Uncropped version of the gels shown in Fig. 2c. **(a)** Y-chromosome specific PCR (primers P1504/1505) on genomic DNA extracted from single flies with intersex (IS) phenotype. A wild type *Egll* male (WT M) and female (WT F) served as positive and negative controls, respectively. **(b)** Positive control PCR on the same genomic DNA samples as in a) using primers P1401/P1500 to amplify a 1,213 bp fragment of *tra2*. The DNA ladder (L) used for agarose gels is the NEB 2-log DNA-ladder; kb = kilo basepairs.

a Y-chromosome specific PCRs

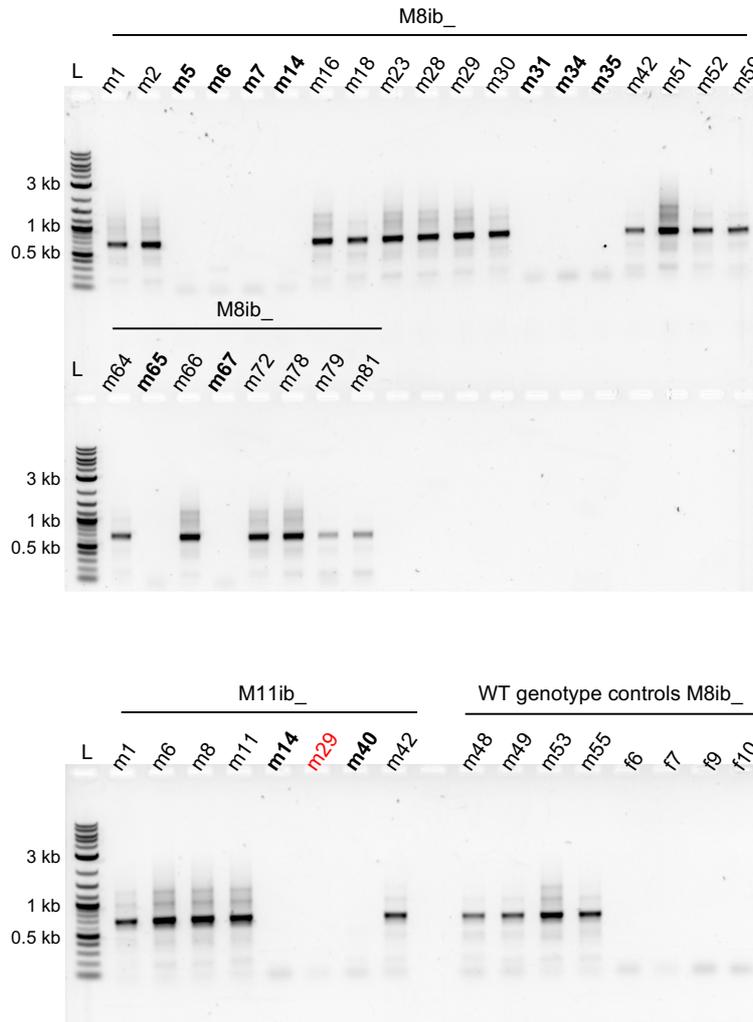


Fig. S5a. Karyotyping of phenotypic G_2 males homozygous for the $tra2^{ts2}$ mutation reveals sex transformation of XX embryos. Uncropped version of the gels shown in Fig. 3b. Y-chromosome specific PCR (primers P1504/1505) on genomic DNA extracted from a single leg of family M8 and M11 offspring. Out of 35 analysed $tra2^{ts2}$ males, eleven did not show a signal in the Y-chromosome specific PCRs and are marked in bold letters. PCRs were done at least twice to verify the results. M11ib_m29 was excluded from the analysis due to low DNA quality (see Fig. S5 b). Four phenotypic males and females from family M8 with WT $tra2$ genotype are shown as controls. The DNA ladder (L) used for agarose gels is the NEB 2-log DNA-ladder; kb = kilo basepairs. A positive control PCR was performed and is shown in Fig. S5b.

b genomic control PCRs

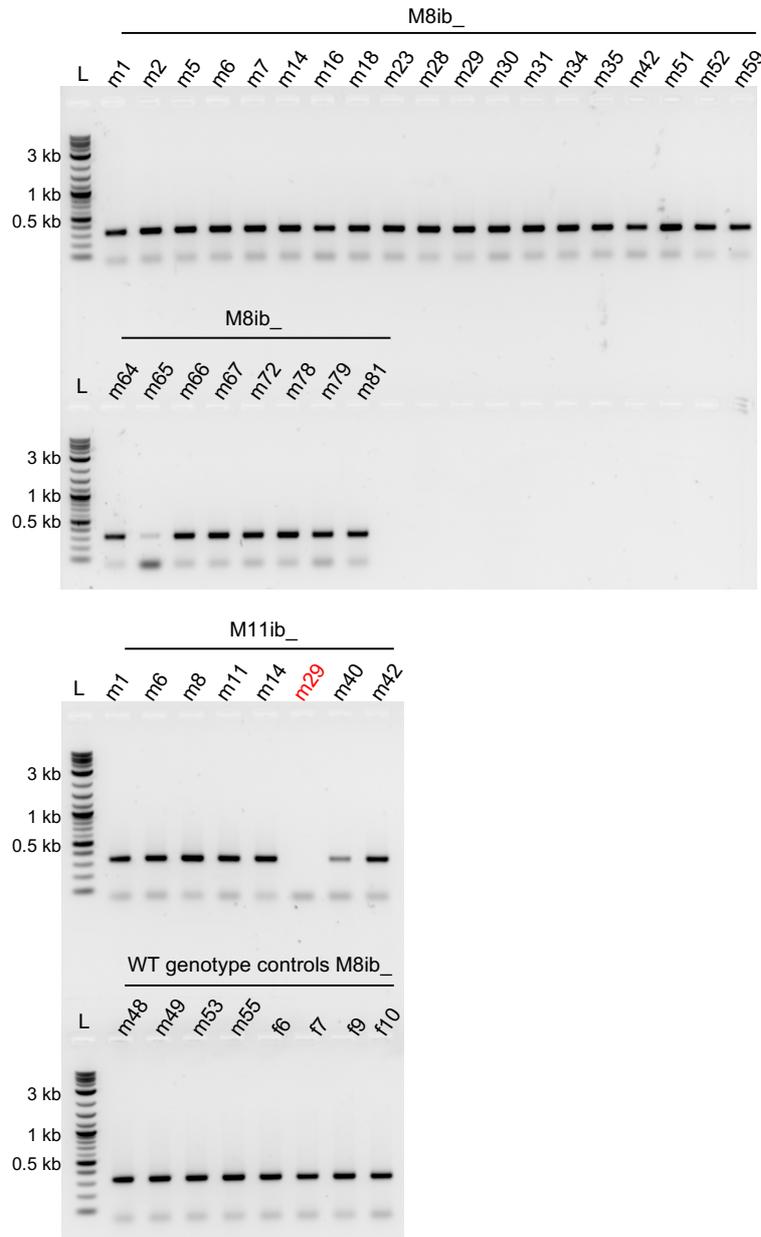


Fig. S5b. Karyotyping of phenotypic G_2 males homozygous for the $tra2^{ts2}$ mutation reveals sex transformation of XX embryos. Uncropped version of the gels shown in Fig. 3b. A positive control PCR was performed on *tra2* with primers P1532/P1500 using the same DNA samples as in the Y-specific PCR, to exclude lack of PCR product due to DNA quality. Individuals lacking a signal in the Y-chromosome-specific PCRs but not in the genomic control PCR are marked in bold letters to indicate the XX karyotype. M11ib_m29 was excluded from the analysis due to low DNA quality. Four phenotypic males and females from family M8 with WT *tra2* genotype are shown as controls. The DNA ladder (L) used for agarose gels is the NEB 2-log DNA-ladder; kb = kilo basepairs.

Y-chromosome specific PCRs

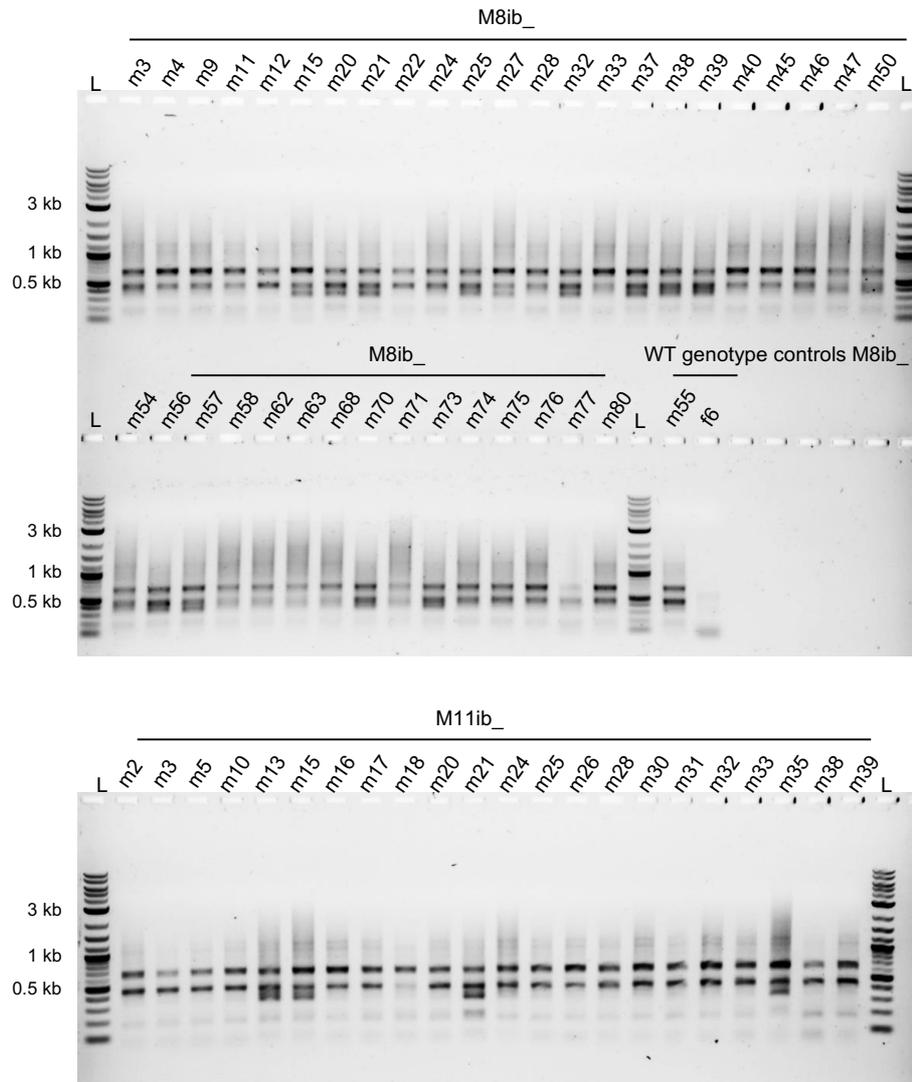


Fig. S6. Karyotyping of phenotypic G_3 males heterozygous for the $tra2^{ts2}$ mutation verifies Y-chromosome presence in all individuals. Y-chromosome-specific PCR (primers P1504/1505) on genomic DNA extracted from single legs of $tra2^{ts2}$ -heterozygous males, originating from family M8 and M11 inbreeding. One phenotypic male and female from family M8 with WT $tra2$ genotype are shown as controls. The DNA ladder (L) used for agarose gels is the NEB 2-log DNA-ladder; kb = kilo basepairs.

Table S1. Temperature profiles during *EgII* control and *tra2^{ts2}* crosses. Shown is the cross, the duration of the cross (from setup until last egg collection, in hours = h), the average temperature during this timeframe [°C], the lowest and highest temperature during this timeframe and their durations, and how long the temperature was above the mating-threshold (18.5°C), absolute [h], and in relation to the overall duration [%]. Durations of temperatures are based on the whole timeframe, and their continuity is shown in Supplementary Fig. S2.

cross	duration	average temp.	min. temp. (duration)	max. temp. (duration)	temp. ≥ 18.5°C
16°C <i>EgII</i> control	410 h	15.88°C	15.5°C (154.92 h)	17°C (1.75 h)	0 h (0%)
19.5°C <i>EgII</i> control	288 h	19.50°C	16.5°C (0.66 h)	20.5°C (2.17 h)	286 h (99.3%)
G ₀ cross <i>tra2^{ts2}</i>	377 h	19.10°C	18.5°C (12.67 h)	20°C (2.17 h)	376.67 h (99.9%)
G ₁ cross <i>tra2^{ts2}</i>	498 h	19.20°C	18.0°C (1.33 h)	21.5°C (4.42 h)	496.75 h (99.7%)
18°C <i>EgII</i> control	378 h	18.09°C	17.5°C (76.83 h)	19.5°C (0.08 h)	146.67 h (38.8%)
G ₂ cross <i>tra2^{ts2}</i>	522 h	18.30°C	17.5°C (71.42 h)	19.5°C (17.92 h)	305.83 h (58.6%)

Table S2. Fertile crosses of the *tra2^{ts2}* CRISPR-HDR injection reared at 19°C. Shown are the two fertile G₀ male flies, the number and genotype of females used for backcrossing, the number of G₁ pupae and adult male and female offspring.

fertile G₀ survivor	females: number, genotype	G₁ pupae	G₁ adults phenotype
M8	3x WT <i>EgII</i>	22	5 females 13 males
M11	3x WT <i>EgII</i>	96	37 females 45 males

Table S3. Crosses of the heterozygous *tra2^{ts2}* G₁ mutants of families M8 and M11 at 19°C. Shown is the number, family, and *tra2* genotype of males and females used to set up the inbreeding and backcrossing cages, the number of eggs collected per cage over all egg collection time points, and the number of G₂ pupae and male and female offspring. Inbreeding crosses are marked in bold.

males: number, family, <i>tra2</i> genotype	females: number, family, <i>tra2</i> genotype	eggs collected	G₂ pupae	G₂ adults phenotype
16, M11m, <i>tra2^{ts2/WT}</i>	17, M11f, <i>tra2^{ts2/WT}</i>	~ 280	71	26 females 42 males
10, M11m, <i>tra2^{ts2/WT}</i>	22, WT	~ 1.180	53	26 females 27 males
4, M8m, <i>tra2^{ts2/WT}</i>	2, M8f, <i>tra2^{ts2/WT}</i>	~ 300	126	38 females 82 males

Table S4. Crossing of G₂ *tra2^{ts2}* mutants at 18.5°C. Shown is: in column 1 the numbering of cages (No.), in column 2 the family identifier (M8ib/M11ib), specific name (single cross) or number of flies per cage (group cross), *tra2* genotype, karyotype and ability to coil the phallus of males, in column 3 the number and *tra2* genotype of females, in column 4 the number of eggs collected over all egg collection time points (7-13 days), in column 5 the number of hatched larvae and in column 6 the number of eclosed adults. Inbreeding crosses are marked in bold.

	No.	males: name (single cross) or number and family (group cross), <i>tra2</i> genotype, karyotype (XX or XY), unable to coil phallus (P)	females: number, <i>tra2</i> genotype	eggs collected	G₃ larvae	G₃ adults
Single cross (backcross)	1	M8ib_m5, <i>tra2^{ts2/ts2}</i> , XX, P	4, WT	222	0	0
	2	M8ib_m6, <i>tra2^{ts2/ts2}</i> , XX, P	4, WT	184	0	0
	3	M8ib_m7, <i>tra2^{ts2/ts2}</i> , XX, P	4, WT	111	0	0
	4	M8ib_m14, <i>tra2^{ts2/ts2}</i> , XX, P	4, WT	373	0	0
	5	M8ib_m30, <i>tra2^{ts2/ts2}</i> , XY, P	4, WT	308	0	0
	6	M8ib_m31 <i>tra2^{ts2/ts2}</i> , XX, P	4, WT	265	0	0
	7	M8ib_m34, <i>tra2^{ts2/ts2}</i> , XX, P	4, WT	319	0	0
	8	M8ib_m35, <i>tra2^{ts2/ts2}</i> , XX, P	4, WT	406	0	0
	9	M8ib_m42, <i>tra2^{ts2/ts2}</i> , XY	4, WT	359	0	0
	10	M8ib_m64, <i>tra2^{ts2/ts2}</i> , XY	4, WT	343	0	0
	11	M8ib_m65 <i>tra2^{ts2/ts2}</i> , XX, P	4, WT	478	0	0
	12	M8ib_m67 <i>tra2^{ts2/ts2}</i> , XX, P	4, WT	399	0	0
	13	M8ib_m81, <i>tra2^{ts2/ts2}</i> , XY	4, WT	303	0	0
Group cross (backcross or inbreeding)	14	4, M8ib_m, <i>tra2^{ts2/WT}</i> , XY, P	8, WT	500	0	0
	15	29, M8ib_m, <i>tra2^{ts2/WT}</i>	22, M8ibf, <i>tra2^{ts2/WT}</i>	1.743	0	0
	16	13, M11ib_m, <i>tra2^{ts2/WT}</i>	10, M11ibf, <i>tra2^{ts2/WT}</i>	511	0	0
	17	6, M8ib_m, <i>tra2^{ts2/ts2}</i> , XY	12, WT	751	1	0
	18	4, M11ib_m, <i>tra2^{ts2/ts2}</i> , XY	8, WT	535	4	1
	19	2, M8ib_m, <i>tra2^{ts2/ts2}</i>, XY	2, M8ibf, <i>tra2^{ts2/WT}</i>	216	0	0

Supplementary references

Prokopy, R. J. & Hendrichs, J. Mating behavior of *Ceratitis capitata* on a field-caged host tree. *Ann. Entomol. Soc. Am.* **72**, 642-648 (1979).

3.3 Solving a genetic puzzle in insect pest control: the *white pupae* gene

The *white pupae* (*wp*) mutation causes a vibrant white puparium that enables automated mechanical sorting and sex separation in genetic sexing strains (GSS). Despite its successful use in medfly SIT programs for over 30 years, its genetic basis and the underlying mechanism remained unknown. Unraveling this trait's nature on a molecular level could help to initiate SIT programs for many more insects of medical and agricultural importance, allowing comparable mutations to be created in the target species by minimally invasive technologies such as CRISPR/Cas9. To identify *wp*, cytogenetics, comparative genomics, and transcriptomic analysis of wild type and mutant strains were conducted, including introgression crossing, evaluation of putative targets, and the reconstruction of mutations. Therefore, I was part of a large consortium of researchers formed at the coordinated research project ‘Generic approach for the development of genetic sexing strains for SIT applications’ (CRP D44003), organized by the International Atomic Energy Agency (IAEA). My position was thankfully funded by the ‘Horticulture Innovation Australia’ (HORT). My contribution aimed to identify putative *wp* candidate genes in medfly and confirm their linkage to the naturally occurring medfly white pupae phenotype. In principle, several modifications of other genes along the pigmentation pathway could lead to a white pupae phenotype. However, it was crucial to identify ‘THE’ *wp* gene responsible for the phenotype in the established medfly GSS, because this gene is tightly linked to a second important GSS-trait, the *temperature-sensitive lethal* (*tsl*).

From cytogenetic analysis, it is known that *wp* is inside the chromosomal inversion ‘D53’, close to its right breakpoint on chromosome 5. Therefore, the first step in narrowing down the genomic location of the *wp* locus was to locate the inversion breakpoints, and we identified them by bioinformatic analysis (A. Darby and J. Ragoussis' labs). However, the left breakpoint was located within a scaffold gap in the new medfly genome assembly Ccap_3.2.1. I developed PCRs for both breakpoints based on my analysis of the wild type and inversion strains' sequence information. This way, I succeeded in verifying the right and identifying the exact sequences of the left breakpoint.

Second, *wp* candidate genes were chosen and subsequently mutated via CRISPR/Cas9 gene editing. I analyzed the candidate genes in different medfly genome assemblies based on the genomic and transcriptomic analysis of the laboratories of A. Darby, S. Geib, and J. Ragoussis' lab. The most promising candidates were localized on chromosomes by *in situ* hybridizations in K. Bourtzis' lab, and I edited the genes via knock-in and knock-out approaches. This resulted in identifying the causal *wp* gene: a metabolite transport protein containing a Major Facilitator-like superfamily domain (MFS_1). This gene was then indeed verified as ‘THE’ *wp* by crossing CRISPR/Cas9-mediated knock-out flies in a complementation cross to a strain carrying the naturally occurring *wp* mutation, which I identified to be a transposon-derived, 8 kb insertion. Novel *wp*^(CRISPR) lines were established and are currently evaluated in quality control tests at the

Insect Pest Control Laboratory of the Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture in Seibersdorf. Furthermore, I developed PCR assays to assess the inversion status and distinguish between the naturally occurring *wp* mutant allele and the wild type allele. The latter one was used in 2020 by Agencies in Australia to uncover the origin of multiple medfly outbreaks. Besides, it allows tracing of released GSS flies, and even eventually surviving progeny could be detected – a safeguard aspect which was missing so far in medfly SIT programs.

The identified *white pupae* gene appears to be well conserved among numerous insect species. Therefore, this study provides the groundwork to build *wp*-based GSS and establish SIT programs in multiple new pest species.

Title: **White pupae phenotype of tephritids is caused by parallel mutation of a MFS transporter**

Authors: Ward CM*, **Aumann RA***, Whitehead MA, Nikolouli K, Leveque G, Gouvi G, (* = co-first) Fung E, Reiling SJ, Djambazian H, Hughes MA, Whiteford S, Caceres-Barrios C, Nguyen TNM, Choo A, Crisp P, Sim SB, Geib SM, Marec F, Häcker I, Ragoussis J, Darby AC, Bourtzis K, Baxter SW, Schetelig MF

Status: **published** in *Nature Communications* (2021)

Contributions: Organization and compilation of all data from all authors, implementation of changes during corrections and revisions from reviewers and editors: **Aumann RA**

C. capitata / *Z. cucurbitae*

- CRISPR gene editing and molecular work on *wp* candidate genes and the D53 inversion: **Aumann RA**
- Bioinformatic analysis of chromosome 5: **Aumann RA**
- Sequencing efforts and bioinformatic analysis: Darby AC, Ragoussis J, **Aumann RA**, Nikolouli K, Whitehead M, Leveque G, Reiling SJ, Djambazian H, Hughes MA, Whiteford S, Marec F, Häcker I, Bourtzis K, Schetelig MF

B. dorsalis / *B. tryoni*

- CRISPR gene editing and molecular work Fung E, Nguyen TNM, Choo A, Crisp P
- Bioinformatics: Ward CM, Baxter SW
- Development of Introgressed Line: Caceres-Barrios C

C. capitata / *Z. cucurbitae* / *B. dorsalis*

- *in situ* hybridization: Gouvi G, Bourtzis K

Presentations: This work was presented at the Consultancy Meeting on ‘Isolation of genetic markers towards the development of generic methods for the construction of GSS for SIT applications’ of the FAO/IAEA in Vienna, Austria (2018, talk), the 1st Research Coordination Meeting on ‘Generic approach for the development of GSS for SIT applications’ of the FAO/IAEA in Vienna, Austria (2019, talk), and the HORT gene discovery meeting in Adelaide/Hahndorf, Australia (2019, talk).





ARTICLE

<https://doi.org/10.1038/s41467-020-20680-5>

OPEN

White pupae phenotype of tephritids is caused by parallel mutations of a MFS transporter

Christopher M. Ward ^{1,12}, Roswitha A. Aumann^{2,12}, Mark A. Whitehead³, Katerina Nikolouli ⁴, Gary Leveque^{5,6}, Georgia Gouvi ^{4,7}, Elisabeth Fung¹, Sarah J. Reiling⁵, Haig Djambazian⁵, Margaret A. Hughes³, Sam Whiteford³, Carlos Caceres-Barrios⁴, Thu N. M. Nguyen ^{1,8}, Amanda Choo ¹, Peter Crisp^{1,9}, Sheina B. Sim ¹⁰, Scott M. Geib¹⁰, František Marec ¹¹, Irina Häcker ², Jiannis Ragoussis ⁵, Alistair C. Darby³, Kostas Bourtzis⁴ , Simon W. Baxter ⁸  & Marc F. Schetelig ² 

Mass releases of sterilized male insects, in the frame of sterile insect technique programs, have helped suppress insect pest populations since the 1950s. In the major horticultural pests *Bactrocera dorsalis*, *Ceratitis capitata*, and *Zeugodacus cucurbitae*, a key phenotype white pupae (wp) has been used for decades to selectively remove females before releases, yet the gene responsible remained unknown. Here, we use classical and modern genetic approaches to identify and functionally characterize causal wp^- mutations in these distantly related fruit fly species. We find that the wp phenotype is produced by parallel mutations in a single, conserved gene. CRISPR/Cas9-mediated knockout of the wp gene leads to the rapid generation of white pupae strains in *C. capitata* and *B. tryoni*. The conserved phenotype and independent nature of wp^- mutations suggest this technique can provide a generic approach to produce sexing strains in other major medical and agricultural insect pests.

¹School of Biological Sciences, University of Adelaide, 5005 Adelaide, Australia. ²Department of Insect Biotechnology in Plant Protection, Justus-Liebig-University Gießen, Institute for Insect Biotechnology, Winchesterstr. 2, 35394 Gießen, Germany. ³Centre for Genomic Research, Institute of Integrative Biology, The Biosciences Building, Crown Street, L69 7ZB Liverpool, United Kingdom. ⁴Insect Pest Control Laboratory, Joint FAO/IAEA Programme of Nuclear Techniques in Food and Agriculture, Seibersdorf, 1400 Vienna, Austria. ⁵McGill University Genome Centre, McGill University, Montreal, QC, Canada. ⁶Canadian Centre for Computational Genomics (C3G), McGill University, Montreal, QC, Canada. ⁷Department of Environmental Engineering, University of Patras, 2 Seferi str., 30100 Agrinio, Greece. ⁸Bio21 Molecular Science and Biotechnology Institute, School of BioSciences, University of Melbourne, Melbourne 3010, Australia. ⁹South Australian Research and Development Institute, Waite Road, Urrbrae 5064, South Australia. ¹⁰USDA-ARS Daniel K. Inouye US Pacific Basin Agricultural Research Center, 64 Nowelo Street, Hilo, HI 96720, USA. ¹¹Biology Centre, Czech Academy of Sciences, Institute of Entomology, Branišovská 31, 370 05 České Budějovice, Czech Republic. ¹²Authors contributed equally to the study: Christopher M. Ward, Roswitha A. Aumann. ✉email: k.bourtzis@iaea.org; simon.baxter@unimelb.edu.au; marc.schetelig@agr.uni-giessen.de

ARTICLE

Tephritid species, including the Mediterranean fruit fly (medfly) *Ceratitidis capitata*, the oriental fruit fly *Bactrocera dorsalis*, the melon fly *Zeugodacus cucurbitae*, and the Queensland fruit fly *Bactrocera tryoni*, are major agricultural pests worldwide¹. The sterile insect technique (SIT) is a species-specific and environment-friendly approach to control their populations, which has been successfully applied as a component of area-wide integrated pest management programs^{2–4}. The efficacy and cost-effectiveness of these large-scale operational SIT applications has been significantly enhanced by the development and use of genetic sexing strains (GSS) for medfly, *B. dorsalis* and *Z. cucurbitae*^{5,6}.

A GSS requires two principal components: a selectable marker, which could be phenotypic or conditionally lethal, and the linkage of the wild-type allele of this marker to the male sex, ideally as close as possible to the male determining region. In a GSS, males are heterozygous and phenotypically wild type, whilst females are homozygous for the mutant allele thus facilitating sex separation^{6–8}. Puparium color was one of the first phenotypic traits exploited as a selectable marker for the construction of GSS. In all three species, brown is the typical puparium color. However, naturally occurring color mutants such as white pupae (wp)⁹ and dark pupae (dp)¹⁰ have occurred in the field or laboratory stocks. The wp locus was successfully used as a selectable marker to develop GSS for *C. capitata*, *B. dorsalis*, and *Z. cucurbitae*^{6,11,12}; however, its genetic basis has never been resolved.

Biochemical studies provided evidence that the white pupae phenotype in medfly is due to a defect in the mechanism responsible for the transfer of catecholamines from the hemolymph to the pupal cuticle¹³. In addition, classical genetic studies showed that the wp phenotype is due to a recessive mutation in an autosomal gene located on chromosome 5 of the medfly genome^{9,14}. The development of translocation lines combined with deletion and transposition mapping and advanced cytogenetic studies allowed the localization of the gene responsible for the wp phenotype on the right arm of chromosome 5, at position 59B of the trichogen polytene chromosome map¹⁵. In the same series of experiments, the wp locus was shown to be tightly linked to a *temperature-sensitive lethal* (*tsl*) gene (position 59B–61C), which is the second selectable marker of the VIENNA 7 and VIENNA 8 GSS currently used in all medfly SIT operational programs worldwide^{7,15}.

The genetic stability of a GSS is a major challenge, mainly due to recombination phenomena taking place between the selectable marker and the translocation breakpoint. To address this risk, a chromosomal inversion called D53 was induced and integrated into the medfly VIENNA 8 GSS (VIENNA 8^{D53+})^{6,8}. Cytogenetic analysis indicated that the D53 inversion spans a large region of chromosome 5 (50B–59C on trichogen polytene chromosome map) with the wp locus being inside the inversion, close to its right breakpoint⁶.

Extensive genetic and cytogenetic studies facilitated the development of a physical map of the medfly genome^{8,16}. The annotated gene set provided opportunities for the identification of genes or loci-associated mutant phenotypes, such as the wp and *tsl*, used for the construction of GSS^{16,17}. Salivary gland polytene chromosome maps developed for *C. capitata*, *B. dorsalis*, *Z. cucurbitae*, and *B. tryoni* show that their homologous chromosomes exhibit similar banding patterns. In addition, in situ hybridization analysis of several genes confirmed that there is extensive shared synteny, including the right arm of chromosome 5 where the *C. capitata* wp gene is localized⁸. Interestingly, two recent studies identified SNPs associated with the wp phenotype in *C. capitata* and *Z. cucurbitae* that were also on chromosome 5^{18,19}.

In this work, we employ different strategies involving genetics, cytogenetics, genomics, transcriptomics, gene editing, and

bioinformatics to identify independent natural mutations in a gene responsible for puparium coloration in three tephritid species of major agricultural importance, *C. capitata*, *B. dorsalis*, and *Z. cucurbitae*. We then functionally characterize causal mutations within this gene in *C. capitata* and *B. tryoni* resulting in development of new white pupae strains. Due to its conserved nature²⁰ and widespread occurrence in many insect species of agricultural and medical importance, we also discuss the potential use of this gene as a generic selectable marker for the construction of GSS for SIT applications.

Results

Resolving the *B. dorsalis* wp locus by introgression experiments. The *B. dorsalis* white pupae phenotype was introgressed into *B. tryoni* to generate a strain referred to as the *Bactrocera* introgressed line (*BIL*, Supplementary Fig. 1). To determine the proportion of *B. dorsalis* genome introgressed into *BIL*, whole-genome sequence data from male and female *B. dorsalis*, *B. tryoni*, and *BIL* individuals were analyzed. Paired-end Illumina short read data from single *B. oleae* males (SRR826808) and females (SRR826807) were used as an outgroup. Single copy orthologs across the genome ($n = 1,846$) were used to reconstruct the species topology revealing a species-specific monophyly (Fig. 1a) consistent with published phylogenies^{21,22}. Reconstruction also showed monophyly between *B. tryoni* and *BIL* across 99.2% of gene trees suggesting the majority of loci originally introgressed from *B. dorsalis* have been removed during backcrosses.

Genomes were partitioned into 100 kb windows and pairwise absolute genetic distance (d_{XY}) calculated between each species and *BIL* to estimate admixture. *Bactrocera dorsalis* was found to be highly similar to a small proportion of the *BIL* genome (Fig. 1b; purple), as indicated by d_{XY} values approaching the median value of *B. dorsalis* vs *B. tryoni* (Fig. 1b; yellow).

Two formal tests for introgression were also carried out, the f_d estimator (Fig. 1c) and topology weighting (Fig. 1d). Three distinct local evolutionary histories (Fig. 1d) were tested using d_{XY} and topology weighting across the *B. dorsalis* wp Quantitative Trait Locus (QTL) i) *BIL* is closest to *B. tryoni* (Fig. 1d; purple, expected across most of the genome), ii) *BIL* is closest to *B. dorsalis* (Fig. 1d; orange, expected at the wp⁻ locus), and iii) *BIL* is closest to *B. oleae* (Fig. 1d; green, a negative control). Across the nuclear genome the species topology was supported in 98.82% of windows. Both f_d and topology weighting confirmed a lack of widespread introgression from *B. dorsalis* into *BIL* with few ($n = 42$) discordant outlier windows. Genomic windows discordant across all three tests were considered candidate regions for the wp mutation. Four scaffolds accounting for 1.18% of the *B. dorsalis* genome met these criteria and only two, NW_011876372.1 and NW_011876398.1, showed homozygous introgression consistent with a recessive white pupae phenotype (Supplementary Fig. 2).

To resolve breakpoints within the *B. dorsalis* wp⁻ QTL, a windowed analysis across NW_011876398.1 and NW_011876372.1 was performed using d_{XY} (Fig. 1e), topology weighting (Fig. 1f) and f_d (Fig. 1g). The maximum range of the introgressed locus was 4.49 Mb (NW_011876398.1 was 2.9–5.94 Mb and NW_011876372.1 was 0–1.55 Mb) (Fig. 1e–g). The wp⁻ QTL was further reduced to a 2.71 Mb region containing 113 annotated protein coding genes through analyzing nucleotide diversity (π) among eight pooled *BIL* genomes (3.8 Mb on NW_011876398.1 to 0.73 Mb on scaffold NW_011876372.1, Supplementary Fig. 2).

Resolving the *C. capitata* wp by genome sequencing and in situ hybridization. Cytogenetic studies have determined the gene responsible for the white pupae phenotype to be localized on the right arm of chromosome 5, at position 59B of the trichogen polytene

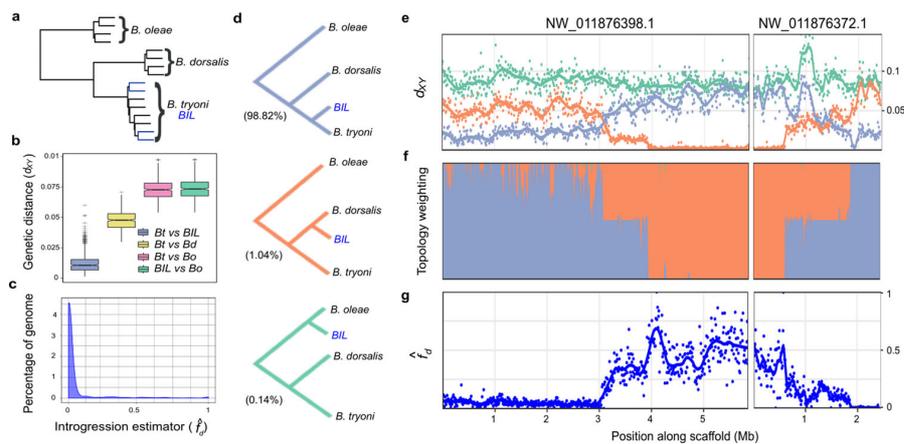


Fig. 1 Characterization of total introgression from *B. dorsalis* into the *Bactrocera* introgressed line and identification of the white pupae locus. **a** Species tree constructed from 1846 single copy ortholog gene trees for four haplotypes of *B. oleae*, *B. dorsalis*, *B. tryoni*, and *BIL*. Branches corresponding to *BIL* individuals are shown in blue. All nodes were well supported with posterior probabilities >0.97. **b** Nei's absolute genetic distance (d_{xy}) calculated for tiled 100 kb windows across the genome between *B. tryoni* vs *BIL* (*Bt* vs *BIL*); *B. tryoni* vs *B. dorsalis* (*Bt* vs *Bd*); *B. tryoni* vs *B. oleae* (*Bt* vs *Bo*); and *BIL* vs *B. oleae* (*BIL* vs *Bo*). Box and whisker graphs (including outliers) represent a summary of 2294 genomic windows. Boxes show the first and third inter quartile range (IQR) while whiskers extend to a maximum of $1.5 * IQR$. All values outside $1.5 * IQR$ are shown as plus signs. **c** The introgression estimator (f_d) calculated across tiled 100 kb windows to identify regions of disproportionately shared alleles between *BIL* and *B. dorsalis*, f_d (*Bt*, *BIL*, *Bd*; *Bo*). **d** The three evolutionary hypothesis/topologies of interest to identify introgressed regions and their representation across the genome: species (purple, 98.82%), introgression (orange, 1.04%) and a negative control tree (green, 0.14%). **e** Nei's absolute genetic distance (d_{xy}) calculated for tiled 10 kb windows across the candidate *wp* locus for *B. tryoni* vs *BIL* (purple), *B. dorsalis* vs *BIL* (orange), *B. oleae* vs *BIL* (green). **f** Topology weighting for each topology shown in **d**, calculated for 1 kb tiled local trees across the candidate *wp* locus. **g** The introgression estimator (f_d) calculated across tiled 10 kb windows for the comparison f_d (*Bt*, *BIL*, *Bd*; *Bo*) to identify the start and end of the introgressed locus. Source data are provided in a Source Data file.

chromosome map¹⁵. The equivalent of position 59B is position 76B of the salivary gland polytene chromosome map, inside but close to the right breakpoint of the D53 inversion (69C–76B on the salivary gland polytene chromosome map). Long read sequencing data were generated of the wild-type strain Egypt II (EgII, WT), the inversion line D53 and the genetic sexing strain VIENNA 8 (without the inversion; VIENNA 8^{D53-|}) (Supplementary Table 1) to enable a comparison of the genomes and locate the breakpoints of the D53 inversion, to subsequently narrow down the target region, and to identify *wp* candidate genes.

Chromosome 5-specific markers¹⁶ were used to identify the EgII_Ccap3.2.1 scaffold_5 as complete chromosome 5. Candidate D53 breakpoints in EgII scaffold_5 were identified using the alignment of three genome datasets EgII, VIENNA 8^{D53-|}, and D53 (see material and methods). The position of the D53 inversion breakpoints was located between 25,455,334 and 25,455,433 within a scaffold gap (left breakpoint), and at 61,880,224 bp in a scaffolded contig (right breakpoint) on EgII chromosome 5 (Ccap3.2.1; accession GCA_905071925) (Fig. 2a). The region containing the causal *wp* gene was known to be just next to the right breakpoint of the D53 inversion. Cytogenetic analysis and in situ hybridization using the WT EgII strain and the D53 inversion line confirmed the overall structure of the inversion, covering the area of 69C–76B on the salivary gland polytene chromosomes (Fig. 2), as well as the relative position of markers residing inside and outside the breakpoints (Fig. 2 and Supplementary Fig. 3). PCRs using two primer pairs flanking the predicted breakpoints (Supplementary Fig. 4) and subsequent sequencing confirmed the exact sequence of the breakpoints. Thereby, the wild-type status was confirmed for EgII flies and VIENNA 7^{D53+|} GSS males, which are heterozygous for the inversion. Correspondingly, these amplicons were not present in D53 males and females or in VIENNA 7^{D53+|} GSS females (all homozygous for the inversion)

(Supplementary Fig. 4). Positive signals for the inversion were detected in D53 and VIENNA 7^{D53+|} GSS males and females, but not in WT flies using an inversion-specific primer pair (Supplementary Fig. 4).

Genome and transcriptome sequencing reveal a single candidate *wp* gene. Orthologs within the QTL of *B. dorsalis*, *C. capitata*, and scaffolds known to segregate with the *wp* phenotype in *Z. cucurbitae* (NW_011863770.1 and NW_011863674.1)¹⁸ were investigated for null mutations under the assumption that errors within a conserved gene result in white pupae. A single ortholog containing fixed indels absent from wild-type strains was identified in each species. White pupae *B. dorsalis* and *BIL* strains showed a 37 bp frame-shift deletion in the first coding exon of LOC105232189 introducing a premature stop codon 210 bp from the transcription start site (Fig. 3a). Presence of the deletion was confirmed in silico using whole genome resequencing from the *wp* and wildtype mapped to the reference, and by de novo assembly of Illumina RNAseq data transcripts (Fig. 3a).

In *C. capitata*, a D53 Nanopore read alignment on EgII showed an independent approximate 8150 bp insertion into the third exon of LOC101451947 disrupting proper gene transcription 822 bp from the transcription start site (Fig. 3b). The insertion sequence is flanked by identical repeats, suggesting that it may originate from a transposable element insertion. The *C. capitata* mutation was confirmed in silico, as in *B. dorsalis*, using whole genome sequencing and RNAseq data (Fig. 3b).

Transcriptome data from the white pupae-based genetic sexing strain of *Z. cucurbitae* revealed a 13 bp deletion in the third exon of LOC105216239 on scaffold NW_011863770.1 introducing a premature stop codon (Fig. 3c).

The candidate white pupae gene in all three species had a reciprocal best BLAST hit to the putative metabolite transport

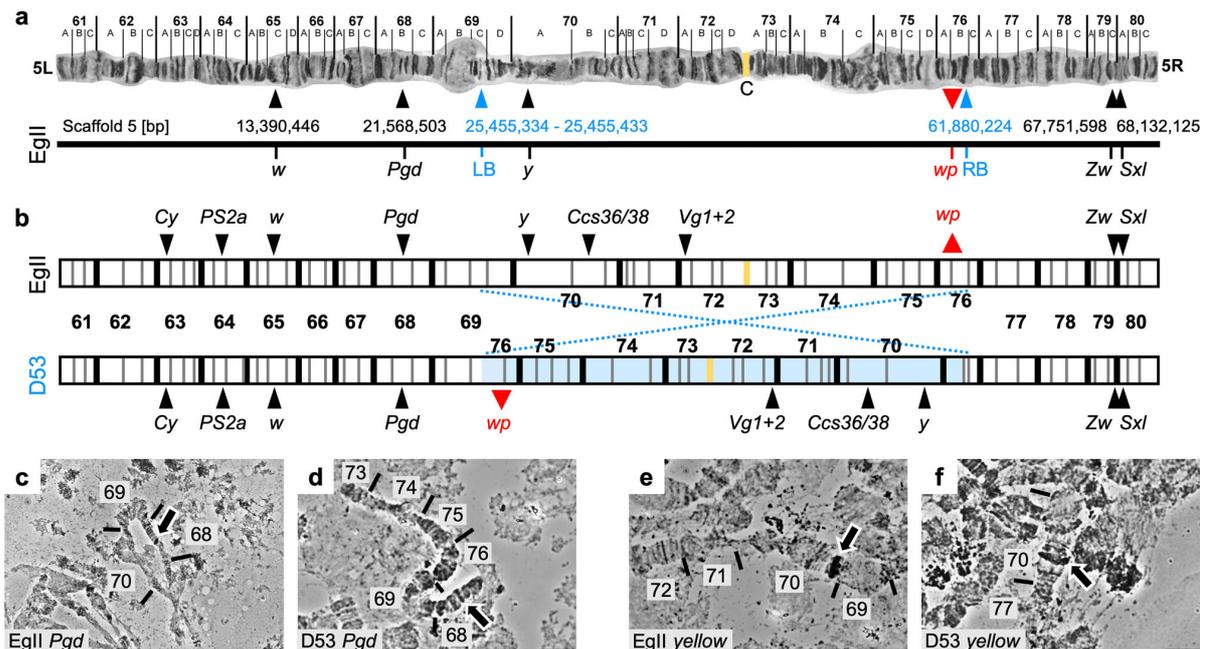


Fig. 2 Genomic positioning of the D53 inversion on chromosome 5 of *C. capitata*. **a** Chromosome scale assembly of *C. capitata* EgII chromosome 5. Shown are the positions of in situ mapped genes *white* (*w*), *6-phosphogluconate dehydrogenase* (*Pgd*), *glucose-6-phosphate 1-dehydrogenase* (*Zw*), and *sex lethal* (*Sxl*), the position of the D53 inversion breakpoints (blue; LB = left breakpoint, RB = right breakpoint), and the relative position of *white pupae* (*wp*) on the polytene chromosome map of chromosome 5⁷¹ (left (L) and right (R) chromosome arm, linked at the centromeric region (C)) and the PacBio-Hi-C EgII scaffold_5 (bp = base pairs), representing the complete chromosome 5 (Ccap3.2.1, accession GCA_905071925). The position of the *yellow* gene (*y*, LOC101455502) was confirmed on chromosome 5 70A by in situ hybridization, despite its sequence not been found in the scaffold assembly. **b** Schematic illustration of chromosome 5 without (EgII, WT) and with (D53) D53 inversion, with additional marker genes *Curly* (*Cy*), *integrin- α PS2* (*PS2a*), *white* (*w*), *chorion S36/38* (*Ccs36/38*), *vitellogenin-1/2-like* (*Vg1+2*). The inverted part of chromosome 5 is shown in light blue, the centromere in yellow. Two probes, one inside (*y*, 70A) and one outside (*Pgd*, 68B) of the left inversion breakpoint were used to verify the D53 inversion breakpoints by in situ hybridization. WT EgII is shown in **c** and **e**, D53 in **d** and **f**. Chromosomal segments are numbered, arrows in micrographs indicate in situ hybridization signal. In situ hybridizations were done at least in duplicates and at least ten nuclei were analyzed per sample, scale bar = 10 μ m. All replicates led to similar results. The source data underlying Fig. 2c–f are provided as a Source Data file.

protein CG14439 in *Drosophila melanogaster* and contains a Major Facilitator-like superfamily domain (MFS_1, pfam07690), suggesting a general function as a metabolite transport protein. In situ hybridization on polytene chromosomes of *B. dorsalis*, *C. capitata* and *Z. cucurbitae* was used to confirm the presence of the *wp* locus in the same syntenic position on the right arm of chromosome 5 (Fig. 3d–f). Therefore, all three species show a mutation in the same positional orthologous gene likely to be responsible for the phenotype in all three genera.

Knockout of the MFS gene causes white pupae phenotypes. An analogous *B. dorsalis wp*⁻ mutation was developed in *B. tryoni* by functional knockouts of the putative *Bt_wp* using the CRISPR/Cas9 system. A total of 591 embryos from the Ourimbah laboratory strain were injected using two guides with recognition sites in the first coding exon of this gene (Fig. 4a). Injected embryos surviving to adulthood ($n = 19$, 3.2%) developed with either wild-type brown ($n = 12$) or somatically mosaic white-brown puparia ($n = 7$, Supplementary Fig. 5). Surviving G₀ adults were individually backcrossed into the Ourimbah strain, resulting in potentially *wp*^{+|-(CRISPR)} heterozygous brown pupae (Fig. 4c). Five independent G₀ crosses were fertile (three mosaic white-brown and two brown pupae phenotypes). G₁ offspring were sibling mated and visual inspection of G₂ progeny revealed that three families contained white pupae individuals. Four distinct

frameshift mutations were observed in screened G₂ progeny (Fig. 4a) suggesting functional KO of putative *Bt_wp* is sufficient to produce the white pupae phenotype in *B. tryoni*. Capillary sequencing of cloned *Bt_MFS* amplicons revealed deletions ranging from a total of 4–155 bp, summed across the two guide recognition sites, introducing premature stop codons.

In *C. capitata*, CRISPR/Cas9 gene editing was used to knockout the orthologous gene and putative *Cc_wp*, LOC101451947, to confirm that it causes a white puparium phenotype. A mix of recombinant Cas9 protein and the gRNA_MFS, targeting the third exon and thereby the MFS domain of the presumed *Cc_wp* CDS (Fig. 4b), was injected into 588 EgII WT embryos of which 96 developed to larvae and 67 pupated. All injected G₀ pupae showed brown pupal color. In total, 29 G₀ males and 34 females survived to adulthood (9.3%) and were backcrossed individually or in groups (see material and methods) to a strain carrying the naturally occurring *white pupae* mutation (*wp*^{-(nat)}; strain #1402_22m1B)²³ (Fig. 4d). As *white pupae* is known to be monogenic and recessive in *C. capitata*, this complementation assay was used to reveal whether the targeted gene is responsible for the naturally occurring white pupae phenotype or if the mutation is located in a different gene. G₁ offspring would only show white pupae phenotypes if *Cc_wp* was indeed the *white pupae* gene, knocked-out by the CRISPR approach, and complemented by the natural mutation through the backcross (*wp*^{-(nat)|-(CRISPR)}). In the case that the *Cc_wp* is not the gene

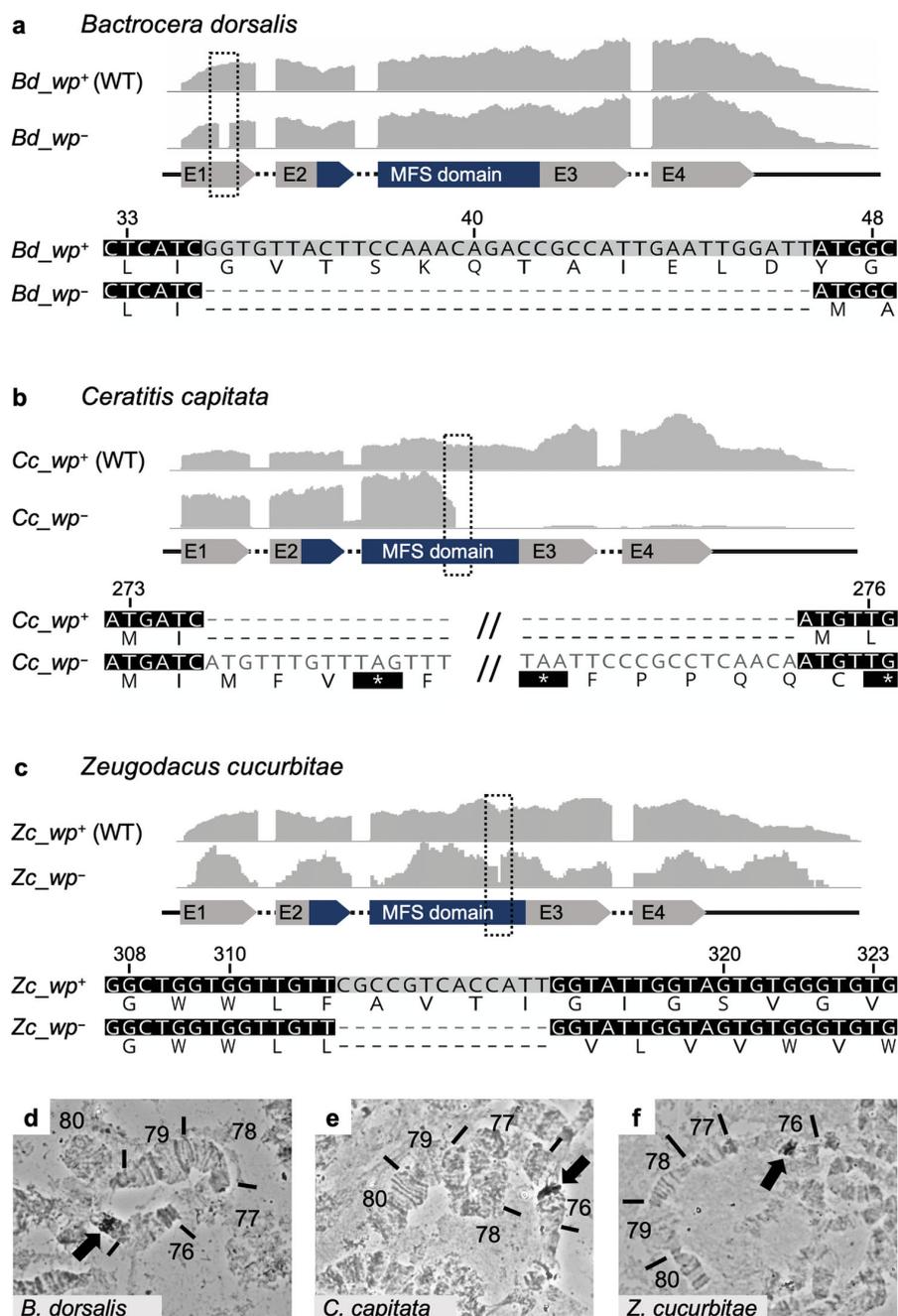
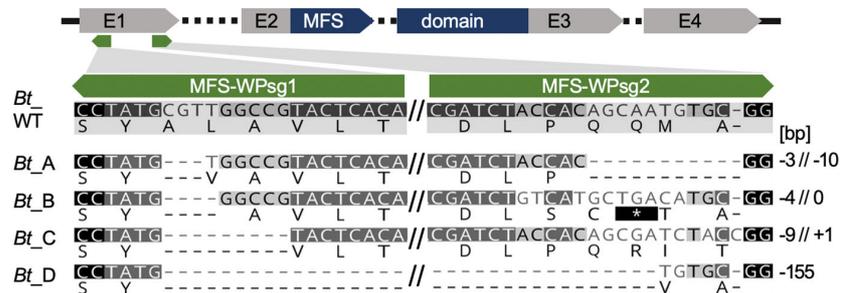
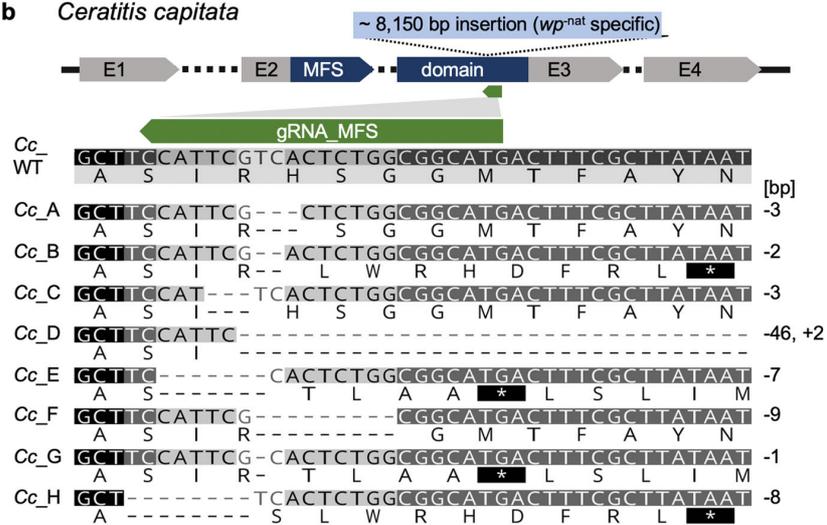


Fig. 3 Identification of the *wp* mutation in the transcriptomes of *B. dorsalis*, *C. capitata*, and *Z. cucurbitae*. The gray graphs show expression profiles from the candidate *wp* loci in WT (*wp*⁺) and mutant (*wp*⁻) flies at the immobile pupae stages of **a** *B. dorsalis*, **b** *C. capitata*, and **c** *Z. cucurbitae*. The gene structure (not drawn to scale) is indicated below as exons (arrows labeled E1–E4) and introns (dashed lines), the Major Facilitator Superfamily (MFS) domain is shown in blue. The positions of independent *wp* mutations (*Bd*: 37 bp deletion, *Cc*: approximate 8150 bp insertion, *Zc*: 13 bp deletion) are marked with black dashed boxes in the expression profiles and are shown in detail below the gene models based on de novo assembly of RNAseq data from WT and white pupae phenotype individuals (nucleotide and amino acid sequences). Deletions are shown as dashes, alterations on protein level leading to premature stop codons are depicted as asterisks highlighted in black. In situ hybridization on polytene chromosomes for **d** *B. dorsalis*, **e** *C. capitata*, and **f** *Z. cucurbitae* confirmed the presence of the *wp* locus on the right arm of chromosome 5 in all three species (arrows in micrographs). In situ hybridizations were done at least in duplicates and at least ten nuclei were analyzed per sample, scale bar = 10 μ m. The source data underlying Fig. 3d–f are provided as a Source Data file.

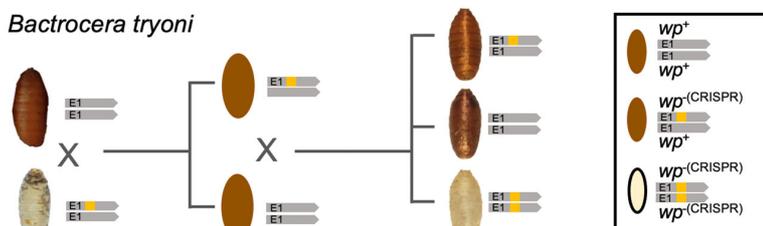
a *Bactrocera tryoni*



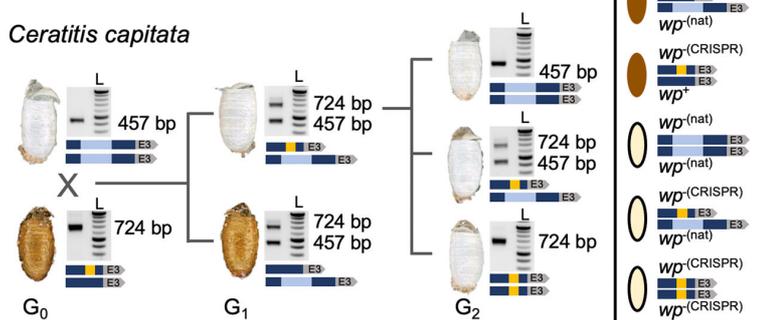
b *Ceratitis capitata*



c *Bactrocera tryoni*



d *Ceratitis capitata*



carrying the natural wp^- mutation, a brown phenotype would be observed for all offspring. Here, five out of 13 crosses, namely M1, M3, F2, F3, and F4, produced white pupae phenotype offspring. The crosses generated 221, 159, 70, 40, and 52 G_1 pupae, of which 10, 30, 16, 1, and 1 pupa respectively, were white. Fifty-seven flies

emerged from white puparia were analyzed via non-lethal genotyping, and all of them showed mutation events within the target region. Overall, eight different mutation events were seen, including deletions ranging from 1 to 9 bp and a 46 bp deletion combined with a 2 bp insertion (Fig. 4b). Five mutation events

Fig. 4 CRISPR/Cas9-based generation of homozygous $wp^{-(CRISPR)}$ lines in *B. tryoni* and *C. capitata*. A schematic structure of the *wp* CDS exons (E1, E2, E3, E4) including the MFS domain in *B. tryoni* (a) and *C. capitata* (b) are shown. Positions of gRNAs targeting the first and third exon in *B. tryoni* and *C. capitata*, respectively, are indicated by green arrows. Nucleotide and amino acid sequences of mutant *wp* alleles identified in G_1 individuals are compared to the WT reference sequence in *B. tryoni* (a) and *C. capitata* (b). Deletions are shown as dashes, alterations on protein level leading to premature stop codons are depicted as asterisks highlighted in black. Numbers on the right side represent InDel sizes (bp = base pairs). Crossing schemes to generate homozygous $wp^{-(CRISPR)}$ lines in *B. tryoni* (c) and *C. capitata* (d) show different strategies to generate *wp* strains. Bright-field images of empty puparia are depicted for both species. Genotype schematics and corresponding PCR analysis (for *C. capitata*) validating the presence of CRISPR-induced (orange) and natural (blue, for *C. capitata*) *wp* mutations are shown next to the images of the puparia. c Injected G_0 *B. tryoni* were backcrossed to the Ourimbah laboratory strain resulting in uniformly brown G_1 offspring (depicted as illustration because no images were acquired during G_1). G_1 inbreeding led to G_2 individuals homozygous for the white pupae phenotype. d Injected WT G_0 flies were crossed to flies homozygous for the naturally occurring wp^- allele ($wp^{-(nat)}$). $wp^{-(nat)}$ (457 bp amplicon) and $wp^{-(CRISPR)}$ or WT (724 bp amplicon) alleles were identified by multiplex PCR (left lane; L = NEB 2 log ladder). White pupae phenotypes in G_1 indicated positive CRISPR events. G_2 flies with a white pupae phenotype that were homozygous for the $wp^{-(CRISPR)}$ allele were used to establish lines. PCR was done once for each individual, $wp^{-(CRISPR)}$ alleles were verified and further analyzed via sequencing. The source data underlying Fig. 4d are provided as a Source Data file.

(B, D, E, G, H) caused frameshifts and premature stop codons. The remaining three (A, C, F), however, produced deletions of only one to three amino acids. Mutants were either inbred (mutation C) (Fig. 4d) or outcrossed to WT EgII (mutation A–H), both in groups according to their genotype. This demonstrated that *Cc_wp* is the gene carrying the $wp^{-(nat)}$, and that even the loss of a single amino acid without a frameshift at this position can cause the white pupae phenotype. Offspring from outcrosses of mutation A, D, and H, as well as offspring of the inbreeding (mutation C), were genotyped via PCR, and $wp^{+/- (CRISPR)}$ and $wp^{-(CRISPR)/-(CRISPR)}$ positive flies were inbred to establish homozygous $wp^{-(CRISPR)}$ lines.

Discussion

White pupae (*wp*) was first identified in *C. capitata* as a spontaneous mutation and was subsequently adopted as a phenotypic marker of fundamental importance for the construction of GSS for SIT^{6,9}. Full penetrance expressivity and recessive inheritance rendered *wp* the marker of choice for GSS construction in two additional tephritid species, *B. dorsalis* and *Z. cucurbitae*^{11,12}, allowing automated sex sorting based on pupal color. This was only possible because spontaneous *wp* mutations occur at relatively high rates either in the field or in mass rearing facilities and can easily be detected^{6,9}. Despite the easy detection and establishment of *wp* mutants in these three species, similar mutations have not been detected in other closely or distantly related species such as *B. tryoni*, *B. oleae*, or *Anastrepha ludens*, despite large screens being conducted. In addition to being a visible GSS marker used to separate males and females, the *wp* phenotype is also important for detecting and removing recombinants in cases where sex separation is based on a conditional lethal gene such as the *tsl* gene in the medfly VIENNA 7 or VIENNA 8 GSS^{6,7}. However, it took more than 20 years from the discovery and establishment of the *wp* mutants to the large-scale operational use of the medfly VIENNA 8 GSS for SIT applications^{6,9} and the genetic nature of the *wp* mutation remained unknown. The discovery of the underlying *wp* mutations and the availability of CRISPR/Cas genome editing would allow the fast recreation of such phenotypes and sexing strains in other insect pests. Isolation of the *wp* gene would also facilitate future efforts towards the identification of the closely linked *tsl* gene.

Using an integrated approach consisting of genetics, cytogenetics, genomics, transcriptomics, and bioinformatics, we identified the white pupae genetic locus in three major tephritid agricultural pest species, *B. dorsalis*, *C. capitata*, and *Z. cucurbitae*. Our study clearly shows the power of employing different strategies for gene discovery, one of which was species hybridization. In *Drosophila*, hybridization of different species has played a

catalytic role in the deep understanding of species boundaries and the speciation processes, including the evolution of mating behavior and gene regulation^{24–28}. In our study, we took advantage of two closely related species, *B. dorsalis* and *B. tryoni*, which can produce fertile hybrids and be backcrossed for consecutive generations. This allowed the introgression of the *wp* mutant locus of *B. dorsalis* into *B. tryoni*, resulting in the identification of the introgressed region, including the causal *wp* mutation via whole-genome resequencing and advanced bioinformatic analysis.

In *C. capitata*, we exploited two essential pieces of evidence originating from previous genetic and cytogenetic studies: the localization of *wp* to region 59B and 76B on chromosome 5 in the trichogen cells and salivary gland polytene chromosome map, respectively^{15,29}, and its position close to the right breakpoint of the large inversion D53⁶. This data prompted us to undertake a comparative genomic approach to identify the exact position of the right breakpoint of the D53 inversion, which would bring us in the vicinity of the *wp* gene. Coupled with comparative transcriptomic analysis, this strategy ensured that the analysis indeed tracked the specific *wp* locus on the right arm of chromosome 5, instead of any mutation in another, random locus which may participate in the pigmentation pathway and therefore result in the same phenotype. Functional characterization via CRISPR/Cas9-mediated knockout resulted in the establishment of new white pupae strains in *C. capitata* and *B. tryoni* and confirmed that this gene is responsible for the puparium's coloration in these tephritid species. Interestingly, the *wp* phenotype is based on three independent and very different natural mutations of this gene, a rather large and transposon-like insertion in *C. capitata*, but only small deletions in the two other tephritids, *B. dorsalis* and *Z. cucurbitae*. In medfly, however, CRISPR-induced in-frame deletions of one or three amino acids in the MFS domain were sufficient to induce the *wp* phenotype, underlining the importance of this domain for correct coloration of the puparium.

It is worth noting that in the first stages of this study, we employed two additional approaches, which did not allow us to successfully narrow down the *wp* genomic region to the desired level. The first was based on Illumina sequencing of libraries produced from laser micro-dissected (Y;5) mitotic chromosomes that carry the wild-type allele of the *wp* gene through a translocation from the fifth chromosome to the Y. This dataset from the medfly VIENNA 7 GSS was comparatively analyzed to wild-type (Egypt II) Y and X chromosomes, and the complete genomes of Egypt II, VIENNA 7^{D53-} GSS, and a D53 inversion line in an attempt to identify the chromosomal breakpoints of the translocation and/or inversion, which are close to the *wp* locus (Supplementary Table 2). However, this effort was not successful due to the short Illumina reads and the lack of a high-quality reference genome. The second approach was based on individual scale whole-genome

ARTICLE

resequencing/genotyping, and identifying fixed loci associated with pupal color phenotypes, which complemented the QTL analysis¹⁹. Seven loci associated with SNPs and larger deletions linked to the white pupae phenotype were analyzed based on their respective mutations and literature searches for their potential involvement in pigmentation pathways. However, we could not identify a clear link to the pupal coloration as shown by *in silico*, molecular, and *in situ* hybridization analysis (Supplementary Figs. 6 and 7, Supplementary Table 3).

The *wp* gene is a member of a Major Facilitator Superfamily (MFS). Orthologs of *white pupae* are present in 146 of 148 insect species aggregated in OrthoDB²⁰ v9 and are single copy in 133 species. Furthermore, *wp* is included in the benchmarking universal single copy ortholog (BUSCO) gene set for Insecta and according to OrthoDB³⁰ v10 has a below average evolutionary rate (0.87, OrthoDB group 42284at50557) suggesting an important and evolutionarily conserved function (Supplementary Fig. 8). Its ortholog in *Bombyx mori*, *muck*, was shown to participate in the pigmentation at the larval stage³¹ whereas in *D. melanogaster* peak expression is during the prepupal stage after the larva has committed to pupation³², which is the stage where pupal cuticle sclerotization and melanization occurs. It is known that the insect cuticle consists of chitin, proteins, lipids, and catecholamines, which act as cross-linking agents thus contributing to polymerization and the formation of the integument³³. Interestingly, the sclerotization and melanization pathways are connected and this explains the different mechanical properties observed in different medfly pupal color strains with the dark color cuticles being harder than the brown ones and the latter harder than the white color ones³⁴. The fact that the white pupae mutants are unable to transfer catecholamines from the hemolymph to the cuticle is perhaps an explanation for the lack of the brown pigmentation¹³.

The discovery of the long-sought *wp* gene in this study and the recent discovery of the *Maleness-on-the-Y* (*MoY*) gene, which determines the male sex in several tephritids³⁵, opens the way for the development of a generic approach for the construction of GSS for other species. Using CRISPR/Cas-based genome editing approaches, we can: (a) induce mutations in the *wp* orthologues of SIT target species and establish lines with *wp* phenotype and (b) link the rescue alleles as closely as possible to the *MoY* region. Given that the *wp* gene is present in diverse insect species including agricultural insect pests and mosquito disease vectors, this approach would allow more rapid development of GSS in SIT target species, members of diverse families, such as the agricultural pest species *A. ludens*, *A. fraterculus*, *B. dorsalis*, *B. corecta*, *B. oleae*, *Drosophila suzukii*, *Cydia pomonella*, *Pectinophora gossypiella*, *Lobesia botrana*; the livestock pests *Glossina morsitans*, *G. pallidipes*, *G. palpalis gambiensis*, *G. austeni*; and the mosquito disease vectors *Aedes aegypti*, *Aedes albopictus*, and *Anopheles arabiensis*. However, the biological quality of any new strain which is considered for SIT application should be first thoroughly tested in respect to their fitness and male mating competitiveness. In principle, these GSS will have higher fertility compared to the semi-sterile translocation lines⁶. In addition, these new generation GSS will be more stable since the rescue allele will be tightly linked to the male determining region thus eliminating recombination which can jeopardize the genetic integrity of any GSS. The concept of the generic approach can also be applied in species which lack a typical Y chromosome such as *Ae. aegypti* and *Ae. albopictus*. In these species, the rescue allele should be transferred close to the male determining gene (*Nix*) and the M locus^{36,37}. It is hence important for this generic approach to identify regions close enough to the male determining loci to ensure the genetic stability of the GSS and to allow the proper expression of the rescue alleles. In the present study,

we have already shown that CRISPR/Cas9-induced mutations resulting in the white pupae phenotype can be developed in SIT target species and the resulting strains provide already new opportunities for GSS based on visible markers.

Methods

Insect rearing. *Ceratitis capitata*, *B. dorsalis*, and *Z. cucurbitae* fly strains were maintained at 25 ± 1 °C, 48% RH and 14/10 h light/dark cycle, and fed with a mixture of sugar and yeast extract (3 v:1 v) and water. Larvae were reared on a gel diet, containing carrot powder (120 g/L), agar (3 g/L), yeast extract (42 g/L), benzoic acid (4 g/L), HCl (25%, 5.75 mL/L), and ethyl-4-hydroxybenzoate (2.86 g/L). Flies were anesthetized with N₂ or CO₂ for screening, sexing, and the setup of crosses. To slow down the development during the non-lethal genotyping process (*C. capitata*), adult flies were kept at 19 °C, 60% RH, and 24 h light for this period (1–4 days).

Bactrocera tryoni flies were obtained from New South Wales Department of Primary Industries (NSW DPI), Ourimbah, Australia and reared at 25 ± 2 °C, 65 ± 10% RH and 14/10 h light/dark cycle. Flies were fed with sugar, Brewer's yeast and water and larvae were reared on a gel diet, containing Brewer's yeast (204 g/L), sugar (121 g/L), methyl p-hydroxy benzoate (2 g/L), citric acid (23 g/L), wheat germ oil (2 g/L), sodium benzoate (2 g/L), and agar (10 g/L).

Introgression and identification of *wp* in *B. dorsalis*. Interspecific crosses between male *B. tryoni* (*wp*^{+/+}) and female *B. dorsalis* (*wp*^{-/-}) were carried out. The F₁ *wp*^{+/-} hybrids developed with brown puparia and were mass crossed. F₂ *wp*^{-/-} females were backcrossed into *B. tryoni wp*^{+/+} males. Backcrossing was then repeated five additional times to produce the white pupae *Bactrocera* introgressed line (*BIL*, Supplementary Fig. 1).

Genome sequencing using Illumina NovaSeq (2 × 150 bp, Deakin University) was performed on a single male and female from the *B. dorsalis wp* strain, *B. tryoni*, and the *BIL* (~ 26X) and two pools of five *BIL* individuals (~ 32X). Quality control of each sequenced library was carried out using FastQC v0.11.6 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and aggregated using ngsReports³⁸ v1.3. Adapter trimming was carried out using Trimmomatic³⁹ v0.38 and paired reads were mapped to the *B. dorsalis* reference genome (GCF_000789215.1) using NextGenMap⁴⁰ v0.5.5 under default settings. Mapped data were sorted and indexed using SAMtools, and deduplication was carried out using Picard MarkDuplicates v2.2.4 (<https://github.com/broadinstitute/picard>). Genotypes were called on single and pooled libraries separately with ploidy set to two and ten, respectively, using FreeBayes⁴¹ v1.0.2. Each strain was set as a different population in FreeBayes. Genotypes with less than five genotype depth were set to missing and sites with greater than 20% missing genotypes or indels filtered out using BCFtools⁴² v1.9. Conversion to the genomic data structure (GDS) format was carried out using SeqArray⁴³ v1.26.2 and imported into the R package gear⁴⁴ v0.1 for population genetic analysis.

Single copy orthologs were identified in the *B. dorsalis* reference annotated proteins (NCBI *Bactrocera dorsalis* Annotation Release 100) with BUSCO^{45,46} v3 using the dipteran gene set⁴⁵. Nucleotide alignments of each complete single copy ortholog were extracted from the called genotype set using gear v0.1 and gene trees built using RAXML⁴⁷ v8.2.10 with a GTR + G model. Gene trees were then imported into Astral III⁴⁸ v5.1.1 for species tree estimation. Genome scans of absolute genetic divergence (d_{XY}), nucleotide diversity (π), and the f_d estimator f_d were carried out using gear v0.1. Two levels of analysis were carried out: i) genome wide scans of non-overlapping 100 kb windows and ii) locus scans of 10 kb tiled windows. Local phylogenies were built for nucleotide alignments of non-overlapping 1 kb windows using RAXML v8.2.10 with a GTR + G model and topology weighting was calculated using TWISST⁴⁹.

Introgressed regions (i.e., candidate *wp* loci) were identified by extracting windows in the genome wide scan with topology weighting and f_d greater than 0.75 and visually inspecting the 'locus scan' data set for d_{XY} , f_d , and topology weighting patterns indicative of introgression. Nucleotide alignments of all genes within candidate *B. dorsalis* introgressed regions were extracted from the GDS using gear and visually inspected for fixed mutations in *B. dorsalis wp*, *BIL* individuals, and the two *BIL* pools. Candidate genes were then searched by tBLASTn against the *D. melanogaster* annotated protein set to identify putative functions and functional domains were annotated using HMMer⁵⁰. Mapped read depth was calculated around candidate regions using SAMtools⁵¹ depth v1.9 and each sample's read depth was normalized to the sample maximum to inspect putative deletions. Called genotypes were confirmed by de novo genome assembly of the *B. dorsalis wp* genome using MaSuRCA⁵² v3.3 under default settings. The de novo scaffold containing LOC105232189 was identified using the BLASTn algorithm. *In silico* exon-intron boundaries were then manually annotated in Geneious⁵³ v11.

Identification of the D53 inversion and *wp* in *C. capitata*. Multiple *C. capitata* strains were used for this study. Egypt II (EgII) is a wild-type laboratory strain. D53 is a homozygous strain with an irradiation-induced inversion covering the area 69C–76B on the salivary gland polytene chromosome map (50B–59C on the trichogen cells polytene chromosome map). VIENNA 7 and VIENNA 8 are two GSS in which two (Y;5) translocations, in the region 58B and 52B of the trichogen cells polytene chromosome map, respectively, have resulted in the linkage of the wild-type allele of the *wp* and *tsl* genes to the male determining region of the

Y chromosome. Thus, VIENNA 7 and VIENNA 8 males are heterozygous in the *wp* and *tsl* loci but phenotypically wild type while VIENNA 7 and VIENNA 8 females are homozygous for the mutant alleles and phenotypically white pupae, and they die when exposed to elevated temperatures. The VIENNA 7 and VIENNA 8 GSS can be constructed with and without the D53 inversion (VIENNA 7/8^{D53+} or ^{D53-}). When the GSS have the inversion, females are homozygous (^{D53++}) for D53 while males are heterozygous (^{D53+-})^{6,8,16}.

To perform whole genome sequencing of *C. capitata* strains, high-molecular-weight (HMW) DNA was extracted from *C. capitata* lines (males and females of the WT EgII strain, the VIENNA 7^{D53-/-} and VIENNA 8^{D53-/-} GSS and the inversion line D53) and sequenced. Freshly emerged, virgin and unfed males and females were collected from all strains. For 10X Genomics linked read and Nanopore sequencing, the HMW was prepared as follows: twenty individuals of each sex and strain were pooled, ground in liquid nitrogen, and HMW DNA was extracted using the QIAGEN Genomic tip 100/G kit (Qiagen, Germany). For PacBio Sequel an EgII line was created with single pair crossing and subsequent sibling-mating for six generations. In all generations adult and larval diet contained 100 µg/mL tetracycline. HMW DNA from G₆ individuals was prepared as follows: five males from this EgII line were pooled and ground in liquid nitrogen, and HMW DNA was extracted using the phenol/chloroform Phase Lock Gel tubes (QuantaBio)⁵⁴. DNA for Illumina applications was extracted from individual flies (Supplementary Table 1). PacBio de novo sequencing: samples were purified with AMPure beads (Beckman Coulter, UK) (0.6 volumes) and QC checked for concentration, size, integrity, and purity using Qubit (Qiagen, UK), Fragment Analyser (Agilent Technologies) and Nanodrop (Thermo Fisher) machines. The samples were then processed without shearing using the PacBio Express kit 1 for library construction and an input of 4 µg DNA following the manufacturer's protocol. The final library was size-selected using the Sage Blue Pippin (Sage Sciences) 0.75% cassette U1 marker in the range of 25–80 kb. The final library size and concentrations were obtained on the Fragment Analyser before being sequenced using the Sequel 1.2.1 chemistry with V4 primers at a loading on plate concentration of 6 pM and 10 h movie times. For Nanopore sequencing, the ligation sequencing kits SQK-LSK109 or SQK-RAD004 were used as recommended by the manufacturer (Oxford Nanopore Technologies, Oxford, United Kingdom). Starting material for the ligation library preparation were 1–1.5 µg HMW gDNA for the ligation libraries and 400 ng for the rapid libraries. The prepared libraries were loaded onto FLO-PRO002 (R9.4) flow cells. Data collection was carried out using a PromethION Beta with live high accuracy base calling for up to 72 h and with mux scan intervals of 1.5 h. Each sample was sequenced at least twice. Data generated were 7.7 Gb for EgII male, 31.09 Gb for D53 male, 26.72 Gb for VIENNA 7^{D53-/-} male, and 24.83 Gb for VIENNA 8^{D53-/-} male. Run metrics are shown in Supplementary Table 4. The PacBio data were assembled using CANU⁵⁵ v1.8 with two parameter settings: the first to avoid haplotype collapsing (genomeSize = 500 m corOutCoverage=200 *batOptions = -dg 3 -db 3 -dr 1 -ca 500 -cp 50) and the second to merge haplotypes together (genomeSize = 500 m corOutCoverage=200 correctedErrorRate=0.15). The genome completeness was assessed with BUSCO^{45,46} v3 using the dipteran gene set⁴⁵. The two assemblies were found to be duplicated due to alternative haplotypes. To improve the contiguity and reduce duplication, haploMerger2 v20161205 was used⁵⁶ and the assembly was assessed with BUSCO v3. Phase Genomics Hi-C libraries were made by Phase genomics from males (*n* = 2) of the same family used for PacBio sequencing. Initial scaffolding was completed by Phase Genomics, but edited using the Salsa⁵⁷ v2.2 and 3D-DNA (3D de novo assembly pipeline v180419; <https://github.com/theaidenlab/3d-dna>) software. The resulting scaffolds were allocated a chromosome number using chromosome specific markers¹⁶. Specific attention was made to the assembly and scaffolding of chromosome 5. Two contig misassemblies were detected by the Hi-C data and fitted manually. The new assembly (EgII_Ccap3.2.1) was then validated using the Hi-C data. Genes were called using the Funannotate v1.6.0-24f346 software making use of the Illumina RNAseq data generated by this project; mRNA mapping to the genome is described below.

To identify possible breakpoint positions, the Nanopore D53 fly assembly contig_531 was mapped onto the EgII_scaffold_5 (from the EgII_CCAP3.2_CANU_Hi-C_scaffolds.fasta assembly) using MashMap v2.0 (<https://github.com/marbl/MashMap>). This helped to visualize the local alignment boundaries (Supplementary Fig. 10). MashMap parameters were set to kmer size = 16; window size = 100; segment length = 500; alphabet = DNA; percentage identity threshold = 95%; filter mode = one-to-one. Subsequent to this, and to help confirm the exact location of the identified breakpoints, minimap2 (v2.17, <https://github.com/lh3/minimap2>) was used to align D53 as well as VIENNA 8^{D53-/-} and VIENNA 7^{D53-/-} Nanopore reads onto the EgII_scaffold_5 reference (Supplementary Fig. 10). Minimap2 parameters for Nanopore reads were: minimap2 -x map-ont -A 1 -a --MD -L -t 40. Samtools (v1.9, <https://github.com/samtools/samtools>) was used to convert the alignment.sam to bam and prepare the alignment file to be viewed in the Integrative Genomics Viewer (IGV, <http://software.broadinstitute.org/software/igv/>). The expectation was to see a leftmost breakpoint in D53 read set alignments but not in VIENNA 8^{D53-/-} and VIENNA 7^{D53-/-} when compared to the EgII reference (Supplementary Fig. 9). Due to an assembly gap in the EgII_scaffold_5 sequence, the exact location of the leftmost inversion breakpoint was not conclusive using this approach. A complementary approach was then used to facilitate detection of the leftmost inversion breakpoint in the D53 inversion line. Minimap2 was again used, but here D53 contig_531 was used as reference for the mapping of EgII male PacBio

reads as well as VIENNA 8^{D53-/-} male and VIENNA 7^{D53-/-} male Nanopore reads (Supplementary Fig. 10). Minimap2 parameters for PacBio reads were: minimap2 -x map-pb -A 1 -a --MD -L -t 40. Minimap2 parameters for Nanopore reads were: minimap2 -x map-ont -A 1 -a --MD -L -t 40. Samtools (v1.9, <https://github.com/samtools/samtools>) was used to convert the alignment.sam to bam and prepare the alignment file to be viewed in the Integrative Genomics Viewer (IGV, <http://software.broadinstitute.org/software/igv/>). The expectation was to see a common breakpoint for all three of the above read set alignments when compared to the D53 genome in the area of the inversion. Position ~3,055,294 was identified in the D53 contig_531 as the most likely leftmost breakpoint. To determine the rightmost breakpoint, D53, VIENNA 8^{D53-/-} and VIENNA 7^{D53-/-} male nanopore reads were aligned on the EgII_scaffold_5 sequence. The expectation was to see a breakpoint in D53 read set alignments but not in VIENNA 7^{D53-/-} and VIENNA 8^{D53-/-}. This is the case here, since read alignments coming from both sides of the inversion are truncated at one position (Supplementary Fig. 9). Findings from genome version EgII_Ccap3.2 were extrapolated to the manually revised genome version EgII_Ccap3.2.1.

Predicted D53 inversion breakpoints were verified via PCRs on EgII, D53, and VIENNA 7^{D53+} GSS male and female genomic DNA, using PhusionFlash Polymerase in a 10 µL reaction volume [98 °C, 10 s; 30 cycles of (98 °C, 1 s; 56 °C, 5 s; 72 °C, 35 s); 72 °C, 1 min] (Supplementary Fig. 4). Sequences of all oligonucleotides used in this study are listed in Supplementary Table 5. The primer pair for the right breakpoint was designed based on EgII sequence information, primers for the left breakpoint were designed based on D53 sequence information. The wild-type status of chromosome 5 (EgII male and female, VIENNA 7^{D53+/-} male) was amplified using primer pairs P_1794 and P_1798 (1950 bp) and P_1795 and P_1777 (690 bp). Chromosome 5 with the inversion (D53 male and female, VIENNA 7^{D53+/-} male and VIENNA 7^{D53+/-} female) was verified using primer pairs P_1777 and P_1798 (1188 bp) and P_1794 and P_1795 (1152 bp) and amplicon sequencing (Macrogen Europe, Amsterdam).

Transcriptomic analysis of *C. capitata*, *B. dorsalis*, and *Z. cucurbitae* species were then conducted for RNA samples from 3rd instar larval and pre-pupal stages (Supplementary Table 1). Total RNA was extracted by homogenizing three larvae of *C. capitata* and *B. dorsalis* and a single larvae of *Z. cucurbitae* in liquid nitrogen, and then using the RNeasy Mini kit (Qiagen). Three replicates per strain and time point were performed. mRNA was isolated using the NEBNext polyA selection and the Ultra II directional RNA library preparation protocols from NEB and sequenced on the Illumina NovaSeq 6000 using dual indexes as 150 bp paired end reads (library insert 500 bp). Individual libraries were sequenced to provide >1 million paired end reads per sample. Each replicate was then assembled separately using Trinity⁵⁸ v2.8.5. The assembled transcripts from Trinity were mapped to the Ccap3.2 genome using minimap⁵⁹ (parameters -ax splice:hq -uf). The Illumina reads were mapped with STAR⁶⁰ v2.5.2.a. IGV⁶¹ v2.6 was used to view all data at a genomic and gene level. Given that the white pupae GSS^{12,62} was used to collect samples for RNA extraction from single larvae of *Z. cucurbitae*, larval sex was confirmed by a maleness-specific PCR on the *MoY* gene of *Z. cucurbitae*³⁵ using cDNA synthesized with the OneStep RT-PCR Kit (Qiagen) and the primer pair ZcMoY1F and ZcMoY1R amplifying a 214 bp fragment. Conditions for a 25 µL PCR reaction using the 1× Taq PCR Master Mix kit (Qiagen) were: [95 °C, 5 min; 30 cycles of (95 °C, 1 min; 51 °C, 1 min; 72 °C, 1 min); 72 °C, 10 min]. Presence of a PCR product indicated a male sample. Each, male and female sample was a pool of three individuals. Three replicates per strain and time point were collected.

Cytogenetic verification of D53 inversion and *wp* genes. Polytene chromosomes for in situ hybridization were prepared from third-instar larvae salivary glands⁶³. In brief, the glands were dissected in 45% acetic acid and placed on a coverslip in a drop of 3:2:1 solution (3 parts glacial acetic acid: 2 parts water: 1-part lactic acid) until been transparent (approximately 5 min). The coverslip was picked up with a clean slide. After squashing, the quality of the preparation was checked by phase contrast microscope. Satisfactory preparations were left to flatten overnight at -20 °C and dipped into liquid nitrogen until the bubbling stopped. The coverslip was immediately removed with razor blade and the slides were dehydrated in absolute ethanol, air dried, and kept at room temperature.

Probes were prepared by PCR. Single adult flies were used to extract DNA with the Extract me kit (Blirt SA), following the manufacturer's protocol. NanoDrop spectrometer was used to assess the quantity and quality of the extracted DNA which was then stored at -20 °C until used. Primers (P_1790/P_1791, P_1821/P_1822, Pgd_probe_F/R, vgl1_probe_F/R, Sxl_probe_F/R, y_probe_F/R, zw_probe_F/R, P_1633/P_1634, Zc_F/R, Bd_F/R, P_1395/P_1396, P_1415/P_1416) were designed for each targeted gene using the Geneious Prime software. PCR was performed in a 25 µL reaction volume using 12.5 µL PCR Master mix 2x Kit (Thermo Fisher Scientific), 60–80 ng DNA, and the following PCR settings [94 °C, 5 min; 35 cycles of (94 °C, 45 s; 56 °C, 30 s; 72 °C, 90 s); 72 °C, 1 min].

Probe labeling was carried out according to the Dig DNA Labelling Kit manual (Roche). Prior to in situ hybridization⁶⁴, stored chromosome preparations were hydrated by placing them for 2 min at each of the following ethanol solutions: 70%, 50%, and 30%. Then they were placed in 2× SSC at room temperature for 2 min. The stabilization of the chromosomes was done by placing them in 2× SSC at 65 °C for 30 min, denaturing in 0.07 M NaOH 2 min, washing in 2× SSC for 30 s, dehydrating (2 min in 30%, 50%, 70%, and 95% ethanol), and air drying. Hybridization was performed on the same day by adding 15 µL of denatured probe

ARTICLE

(boiled for 10 min and ice-chilled). Slides were covered with a siliconized coverslip, sealed with rubber cement, and incubated at 45 °C overnight in a humid box. At the end of incubation, the coverslip was floated off in 2× SSC and the slide washed in 2× SSC for 3 × 20 min at 53 °C.

After 5 min wash in Buffer 1 (100 mM Tris-HCl pH 7.5/ 1.5 M NaCl), the preparations were in Blocking solution (Blocking reagent 0.5% in Buffer 1) for 30 min, and then washed for 1 min in Buffer 1. The antibody mix was added to each slide and a coverslip was added. Then the slides were incubated in a humid box for 45 min at room temperature, following 2 × 15 min washes in Buffer 1, and a 2 min wash in detection buffer (100 mM Tris-HCl pH 9.5/ 100 mM NaCl). The color was developed with 1 mL of NBT/BCIP solution during a 40 min incubation in the dark at room temperature. The removal of the NBT/BCIP solution was done by rinsing in water twice. Hybridization sites were identified using 40× or 100× oil objectives (phase or bright field) and a Leica DM 2000 LED microscope, with reference to the salivary gland chromosome maps⁶⁵. Well-spread nuclei or isolated chromosomes were photographed using a digital camera (Leica DMC 5400) and the LAS X software 3.7.0. All in situ hybridizations were performed at least in duplicates and at least ten nuclei were analyzed per sample.

Gene editing and generation of homozygous *w⁺* strains. For CRISPR/Cas9 gene editing in *B. tryoni*, purified Cas9 protein (Alt-R S.p. Cas9 Nuclease V3, #1081058, 10 µg/µL) and guide RNAs (customized Alt-R CRISPR/Cas9 crRNA, 2 nmol and Alt-R CRISPR/Cas9 tracrRNA, #1072532, 5 nmol) were obtained from Integrated DNA Technologies (IDT). The guide RNAs were individually resuspended to a 100 µM stock solution with nuclease-free duplex buffer before use. The two customized 20 bp crRNA sequences (Bt_MFS-1 and Bt_MFS-2) were designed using CRISPOR⁶⁶. Injection mixes for microinjection of *B. tryoni* embryos comprise of 300 ng/µL Cas9 protein, 59 ng/µL of each individual crRNA, 222 ng/µL tracrRNA, and 1× injection buffer (0.1 mM sodium phosphate buffer pH 6.8, 5 mM KCl) in a final volume of 10 µL. The guide RNA complex containing the two crRNAs and tracrRNA was prepared by heating at 95 °C for 5 min before cooling to room temperature. The Cas9 enzyme along with the other injection mix components were then added to the guide RNA complex and incubated at room temperature for 5 min to assemble the ribonucleoprotein (RNP) complexes. Microinjections were performed in *B. tryoni* Ourimbah laboratory strain embryos that were collected over a 1 h time period. Injections were performed under paraffin oil using borosilicate capillary needles (#30-0038, Harvard Apparatus) drawn out on a Sutter P-87 flaming/brown micropipette puller and connected to an air-filled 20 mL syringe, a manual MM-3 micromanipulator (Narishige) and a CKX31-inverted microscope (Olympus). Microscope slides with the injected embryos were placed on agar in a Petri dish inside a vented container containing moist paper towels at 25 °C (± 2 °C). Hatched first instar larvae were removed from the oil and transferred to larval food. Individual G₀ flies were crossed to six virgin flies from the Ourimbah laboratory strain and eggs were collected overnight for two consecutive weeks. G₁ flies were then allowed to mate inter se and eggs were collected in the same manner. G₂ pupae were then analyzed phenotypically and separated according to color of pupae (brown, mosaic, or white).

For *C. capitata* CRISPR/Cas9 gene editing, a guide RNA (gRNA_MFS), targeting the third CDS exon of *CcMFS* was designed and tested for potential off target effects using Geneious Prime⁵³ and the *C. capitata* genome annotation Ccap2.1¹⁶. In silico target site analysis predicted an on-target activity score of 0.615 (scores are between 0 and 1; higher score corresponds to higher expected activity⁶⁷) and zero off-targets sites in the medfly genome. gRNA_MFS was synthesized by in vitro transcription of linear double-stranded DNA template. Therefore, a linear DNA template was amplified in a 100 µL PCR reaction using primers P_1753 and P_369 and Q5 HF polymerase (NEB) according to the manufacturers protocol (Bio-Rad C1000 Touch thermal cycler [98 °C, 30 s; 35 cycles of (98 °C, 10 s; 58 °C, 20 s; 72 °C, 20 s); 72 °C, 2 min]). The PCR reaction was purified using the Clean and Concentrator-25 kit. Subsequently, 500 ng were transcribed using the HiScribe T7 High Yield RNA Synthesis kit (NEB), followed by a DNase treatment (Invitrogen) and a final purification of the RNA using the Mega Clear Kit (Invitrogen). Injection mix for microinjection of embryos contained 360 ng/µL Cas9 protein (1 µg/µL, dissolved in its formulation buffer (PNA Bio Inc, CP01)), 200 ng/µL gRNA_MFS, and an end-concentration of 300 mM KCl^{68,69}. The mix was freshly prepared on ice followed by an incubation step for 10 min at 37 °C to allow pre-assembly of gRNA-Cas9 RNP complexes and stored on ice until use. Microinjections were conducted in WT EgII *C. capitata* embryos, collected over a 30–40 min period, chemically dechorionized (sodium hypochlorite, 3 min), fixed on double-sided sticky tape (Scotch 3 M), and covered with halocarbon oil 700 (Sigma-Aldrich). For injections, siliconized quartz glass needles (Q100-70-7.5; LOT171381; Science Products, Germany), drawn out on a laser-based micropipette puller (Sutter P-2000), were used with a manual micromanipulator (MN-151, Narishige), an Eppendorf FemtoJet 4i microinjector, and an Olympus SZX16 microscope (SDF PLAPO 1xPF objective). Injected embryos were placed in an oxygen chamber (max. 2 psi), first instar larvae were transferred from the oil to larval food.

As complementation assay, reciprocal crosses between surviving G₀ adults and virgin adults of the white pupae strain #1402_22m1B (pBac_fa_attP-TREhs43-Ctra-1-hid^ΔSV40_a_PUB-nls-EGFP-SV40) (*w⁻*)²³ were set up either single paired (six cages) or in groups of seven to ten flies (seven cages). Eggs were

collected three times every 1–2 days. Progeny (G₁) exhibiting the white pupae phenotype (*w⁻*(nat)-(CRISPR)) were assayed via non-lethal genotyping and sorted according to mutation genotype (see Fig. 4). Genotypes 'A-H' were group-backcrossed to WT EgII (*w⁺*), genotype 'C' siblings mass-crossed. Eggs were collected four times every 1–2 days. Generation G₂ flies were analyzed via multiplex PCR using three primers, specific for *w⁺* and *w⁻*(CRISPR) or *w⁻*(nat) allele size, respectively (see molecular analyses of *w⁺* mutants and mosaics, *C. capitata* non-lethal genotyping). Offspring of outcross cages showed brown pupae phenotype and either *w⁺*(nat) or *w⁺*(-CRISPR) genotype. In order to make mutations A, D, and H homozygous, 40 flies (25 females, 15 males) were genotyped each, and *w⁺*(-CRISPR) positive flies were inbred (mutation A: 15 females, 7 males, mutation D: 12 females, 7 males, mutation H: 11 females, 8 males). G₃ showing white pupae phenotype was homozygous for *w⁻*(CRISPR) mutations A, D, or H, respectively, and was used to establish lines. Inbreeding of mutation C *w⁻*(nat)-(CRISPR) flies produced only white pupae offspring, based on either the *w⁻*(nat)-(nat), *w⁻*(nat)-(CRISPR), or *w⁻*(CRISPR)-(CRISPR) genotype. 94 flies (46 females, 48 males) were genotyped, homozygous *w⁻*(CRISPR) were inbred to establish a line (13 females, 8 males).

Molecular analyses of *w⁺* mutants and mosaics. In *B. tryoni*, genomic DNA was isolated for genotyping from G₂ pupae using the DNeasy Blood and Tissue Kit (Qiagen). PCR amplicons spanning both BtMFS guide recognition sites were generated using Q5 polymerase (NEB) with primers BtMFS_5primeF and BtMFS_exon2R. Products were purified using MinElute PCR Purification Kit (Qiagen), ligated into pGEM-t-easy vector (Promega) and transformed into DH5α cells. Plasmids were purified with Wizard Plus SV Minipreps (Promega) and sequenced.

In *C. capitata*, non-lethal genotyping was performed to identify parental genotypes before setting up crosses. Therefore, genomic DNA was extracted from single legs of G₁ and G₂ flies following an adapted version of an established protocol⁷⁰. Single legs of anesthetized flies were cut at the proximal femur, placed in vials containing ceramic beads and 50 µL buffer (10 mM Tris-Cl, pH 8.2, 1 mM EDTA, 25 mM NaCl), and homogenized for 15 s (6 m/s) using a FastPrep-24 5 G homogenizer. Then, 28.3 µL buffer and 1.7 µL proteinase-K (2.5 U/mg) were added. The reaction mix was incubated for 1 h at 37 °C, and subsequently cooled down on ice and used for PCR. For G₁ flies, PCR on *w⁺* was performed in a 25 µL reaction volume using the DreamTaq polymerase, primers P_1643 and P_1644, and 3.75 µL reaction mix, whereby different amplicon sizes were expected for different alleles (*w⁺* and *w⁻*(CRISPR): 724 bp, *w⁻*(nat): 8872 bp). The *w⁻*(nat) amplicon was excluded via PCR settings (95 °C, 3 min; 35 cycles of (95 °C, 30 s; 56 °C, 30 s; 72 °C, 1 min); 72 °C, 5 min). The 724 bp PCR product was verified by gel electrophoresis and purified from the PCR reaction using the DNA Clean & Concentrator-5 kit. PCR products were sequenced (P_1644) and analyzed using Geneious Prime⁵³. In generation G₂, flies were analyzed using multiplex PCR with primers P_1657, P_1643, and P_1644, to distinguish between the *w⁻*(nat) (457 bp; P_1643/P_1657), and *w⁻*(CRISPR) alleles (724 bp; P_1643/P_1644) using the above-described PCR protocol.

Image acquisition. Images of *B. tryoni* pupae were taken with an Olympus SZX16 microscope, Olympus DP74 camera, and Olympus LF-PS2 light source using the Olympus stream basic 2.3.3 software. Images of *C. capitata* pupae were taken with a Keyence digital microscope VHX-5000. Image processing was conducted with Adobe Photoshop CS5.1 software to apply moderate changes to image brightness and contrast. Changes were applied across the entire image.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this article is available as a Supplementary Information file. The datasets and insect strains generated and analyzed during the current study are available from the corresponding authors upon request. All sequences generated in this study from *B. dorsalis*, *B. tryoni*, *Bactrocera* introgressed line (BIL), *C. capitata*, and *Z. cucurbitae* samples are publicly available on NCBI within the ENA BioProject PRJEB36344 (for Ccap genome assembly EgII-3.2.1, WGS, PacBio, chromosome dissections, Illumina MiSeq, Illumina HiSeq 4000, RNAseq, Illumina NovaSeq 6000, Hi-C, and Nanopore data; see Supplementary Table 1 for detailed sample designation), BioProject PRJNA629430 (for WGS and Illumina DNAseq 2 × 250 PE data; see Supplementary Fig. 6 for detailed sample designation), and BioProject PRJNA682907 (for WGS and Illumina NovaSeq 6000 data; see Supplementary Table 1 for detailed sample designation). The source data underlying Figs. 1, 2c–f, 3d–f, and 4d, as well as Supplementary Figs. 3a–b, 4a–b, 4d–e, and 7 are provided as a Source Data file. Source data are provided with this paper.

Received: 16 June 2020; Accepted: 14 December 2020;

Published online: 21 January 2021

References

- Robinson, A. S. & Hooper, G. *Fruit Flies: Their Biology, Natural Enemies, and Control* Vol. 1 (Elsevier, 1989).
- Suckling, D. M. et al. Eradication of tephritid fruit fly pest populations: outcomes and prospects. *Pest Manag. Sci.* **72**, 456–465 (2016).
- Dyck, V. A. et al. *Sterile Insect Technique – Principles and Practice in Area-Wide Integrated Pest Management* (eds Dyck, V. A., Hendrichs, J. & Robinson, A. S.) (Springer, 2005).
- Vreysen, M., Robinson, A. S. & Hendrichs, J. *Area-Wide Control of Insect Pests: from Research to Field Implementation* (Springer, 2007).
- Rendon, P., McInnis, D., Lance, D. & Stewart, J. Medfly (Diptera: Tephritidae) genetic sexing: large-scale field comparison of males-only and bisexual sterile fly releases in Guatemala. *J. Econ. Entomol.* **97**, 1547–1553 (2004).
- Franz, G. *Sterile Insect Technique – Principles and Practice in Area-Wide Integrated Pest Management* (eds Dyck, V. A., Hendrichs, J. & Robinson, A. S.) (Springer, 2005).
- Augustinos, A. A. et al. *Ceratitis capitata* genetic sexing strains: laboratory evaluation of strains from mass-rearing facilities worldwide. *Entomol. Exp. Appl.* **164**, 305–317 (2017).
- Zacharopoulou, A. et al. A review of more than 30 years of cytogenetic studies of Tephritidae in support of sterile insect technique and global trade. *Entomol. Exp. Appl.* **164**, 204–225 (2017).
- Rössler, Y. The genetics of the Mediterranean fruit fly: a “white pupae” mutant. *Ann. Entomol. Soc. Am.* **72**, 583–585 (1979).
- Rössler, Y. & Koltin, Y. The genetics of the Mediterranean fruit fly, *Ceratitis capitata*: three morphological mutations. *Ann. Entomol. Soc. Am.* **69**, 604–608 (1976).
- McCombs, S. D. & Saul, S. H. Linkage analysis of five new genetic markers of the oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae). *J. Hered.* **83**, 199–203 (1992).
- McInnis, D. O. et al. Development of a pupal color-based genetic sexing strain of the melon fly, *Bactrocera cucurbitae* (Coquillett) (Diptera: Tephritidae). *Ann. Entomol. Soc. Am.* **97**, 1026–1033 (2004).
- Wappner, P. et al. *White pupa*: a *Ceratitis capitata* mutant lacking catecholamines for tanning the puparium. *Insect Biochem. Molec. Biol.* **25**, 365–373 (1995).
- Rössler, Y. & Rosenthal, H. Genetics of the mediterranean fruit fly (Diptera: Tephritidae): morphological mutants on chromosome five. *Ann. Entomol. Soc. Am.* **85**, 525–531 (1992).
- Kerremans, P. & Franz, G. Cytogenetic analysis of chromosome 5 from the Mediterranean fruit fly, *Ceratitis capitata*. *Chromosoma* **103**, 142–146 (1994).
- Papanicolaou, A. et al. The whole genome sequence of the Mediterranean fruit fly, *Ceratitis capitata* (Wiedemann), reveals insights into the biology and adaptive evolution of a highly invasive pest species. *Genome Biol.* **17**, 192 (2016).
- Papathanos, P. A. et al. A perspective on the need and current status of efficient sex separation methods for mosquito genetic control. *Parasites Vectors* **11**, 654 (2018).
- Sim, S. B. & Geib, S. M. A chromosome-scale assembly of the *Bactrocera cucurbitae* genome provides insight to the genetic basis of white pupae. *G3* **7**, 1927–1940 (2017).
- Sim, S. B., Ruiz-Arce, R., Barr, N. B. & Geib, S. M. A new diagnostic resource for *Ceratitis capitata* strain identification based on QTL mapping. *G3* **7**, 3637–3647 (2017).
- Zdobnov, E. M. et al. OrthoDB v9.1: cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. *Nucleic Acids Res.* **45**, D744–D749 (2017).
- San Jose, M. et al. Incongruence between molecules and morphology: a seven-gene phylogeny of *Dacini* fruit flies paves the way for reclassification (Diptera: Tephritidae). *Mol. Phylog. Evol.* **121**, 139–149 (2018).
- Choo, A., Crisp, P., Saint, R., O’Keefe, L. V. & Baxter, S. W. CRISPR/Cas9-mediated mutagenesis of the white gene in the tephritid pest *Bactrocera tryoni*. *J. Appl. Entomol.* **142**, 52–58 (2018).
- Ogaugwu, C. E., Schetelig, M. F. & Wimmer, E. A. Transgenic sexing system for *Ceratitis capitata* (Diptera: Tephritidae) based on female-specific embryonic lethality. *Insect Biochem. Mol. Biol.* **43**, 1–8 (2013).
- Davis, A. W. et al. Rescue of hybrid sterility in crosses between *D. melanogaster* and *D. simulans*. *Nature* **380**, 157–159 (1996).
- Araripe, L. O., Montenegro, H., Lemos, B. & Hartl, D. L. Fine-scale genetic mapping of a hybrid sterility factor between *Drosophila simulans* and *D. mauritiana*: the varied and elusive functions of “speciation genes”. *BMC Evol. Biol.* **10**, 385 (2010).
- Brideau, N. J. & Barbash, D. A. Functional conservation of the *Drosophila* hybrid incompatibility gene *Lhr*. *BMC Evol. Biol.* **11**, 57 (2011).
- Kotov, A. A. et al. piRNA silencing contributes to interspecies hybrid sterility and reproductive isolation in *Drosophila melanogaster*. *Nucleic Acids Res.* **47**, 4255–4271 (2019).
- Barbash, D. A. Ninety years of *Drosophila melanogaster* hybrids. *Genetics* **186**, 1–8 (2010).
- Bedo, D. G. & Zacharopoulou, A. Inter-tissue variability of polytene chromosome banding patterns. *Trends Genet.* **4**, 90–91 (1988).
- Kriventseva, E. V. et al. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Res.* **47**, D807–d811 (2019).
- Zhao, Y. et al. A major facilitator superfamily protein participates in the reddish brown pigmentation in *Bombyx mori*. *J. Insect Physiol.* **58**, 1397–1405 (2012).
- The modEncode Consortium et al. Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science* **330**, 1787–1797 (2010).
- Wright, T. R. The genetics of biogenic amine metabolism, sclerotization, and melanization in *Drosophila melanogaster*. *Adv. Genet.* **24**, 127–222 (1987).
- Bourtzis, K., Psachoulia, C. & Marmaras, V. J. Evidence that different integumental phosphatases exist during development in the Mediterranean fruit fly *Ceratitis capitata*: possible involvement in pupariation. *Comp. Biochem. Physiol. Part B* **98**, 411–416 (1991).
- Meccariello, A. et al. *Maleness-on-the-Y (MoY)* orchestrates male sex determination in major agricultural fruit fly pests. *Science* **365**, 1457–1460 (2019).
- Hall, A. B. et al. Sex determination. A male-determining factor in the mosquito *Aedes aegypti*. *Science* **348**, 1268–1270 (2015).
- Liu, P. et al. *Nix* is a male-determining factor in the Asian tiger mosquito *Aedes albopictus*. *Insect Biochem. Mol. Biol.* **118**, 103311 (2019).
- Ward, C. M. & Pederson, T. H. S. M. ngsReports: a Bioconductor package for managing FastQC reports and other NGS related log files. *Bioinformatics* **36**, 2587–2588 (2020).
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- Sedlazeck, F. J., Rescheneder, P. & von Haeseler, A. NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics* **29**, 2790–2791 (2013).
- Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv:1207.3907* (2012).
- Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
- Zheng, X. et al. SeqArray-a storage-efficient high-performance data format for WGS variant calls. *Bioinformatics* **33**, 2251–2257 (2017).
- Ward, C. M., Ludington, A. J., Breen, J. & Baxter, S. W. Genomic evolutionary analysis in R with gear. <https://doi.org/10.1101/2020.08.06.240754> (2020).
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
- Waterhouse, R. M. et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* **35**, 543–548 (2018).
- Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
- Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinforma.* **19**, 153 (2018).
- Martin, S. H. & Van Belleghem, S. M. Exploring evolutionary relationships across the genome using topology weighting. *Genetics* **206**, 429–438 (2017).
- Finn, R. D., Clements, J. & Eddy, S. R. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* **39**, W29–W37 (2011).
- Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Zimin, A. V. et al. The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669–2677 (2013).
- Kearse, M. et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).
- Green, M. R. & Sambrook, J. Isolation of High-Molecular-Weight DNA using organic solvents. *Cold Spring Harb. Protoc.* **2017**, pdb.prot093450 (2017).
- Koren, S. et al. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
- Huang, S. et al. HaploMerger: reconstructing allelic relationships for polymorphic diploid genome assemblies. *Genome Res.* **22**, 1581–1588 (2012).
- Ghurye, J., Pop, M., Koren, S., Bickhart, D. & Chin, C. S. Scaffolding of long read assemblies using long range contact information. *BMC Genomics* **18**, 527 (2017).
- Haas, B. J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).
- Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
- Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

ARTICLE

NATURE COMMUNICATIONS | <https://doi.org/10.1038/s41467-020-20680-5>

61. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
62. Zacharopoulou, A. & Franz, G. Genetic and cytogenetic characterization of genetic sexing strains of *Bactrocera dorsalis* and *Bactrocera cucurbitae* (Diptera: Tephritidae). *J. Econ. Entomol.* **106**, 995–1003 (2013).
63. Zacharopoulou, A. et al. The genome of the Mediterranean fruit fly *Ceratitis capitata*: localization of molecular markers by in situ hybridization to salivary gland polytene chromosomes. *Chromosoma* **101**, 448–455 (1992).
64. Mavragani-Tsipidou, P. et al. *Protocols for Cytogenetic Mapping of Arthropod Genomes* (ed Sakharov, I.) (CRC Press, Taylor and Francis Group, LLC, 2014).
65. Zacharopoulou, A. Polytene chromosome maps in the medfly *Ceratitis capitata*. *Genome* **33**, 184–197 (1990).
66. Concordet, J. P. & Haussler, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res.* **46**, W242–W245 (2018).
67. Doench, J. G. et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* **32**, 1262–1267 (2014).
68. Aumann, R. A., Schetelig, M. F. & Häcker, I. Highly efficient genome editing by homology-directed repair using Cas9 protein in *Ceratitis capitata*. *Insect Biochem. Mol. Biol.* **101**, 85–93 (2018).
69. Burger, A. et al. Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* **143**, 2025–2037 (2016).
70. Carvalho, G. B., Ja, W. W. & Benzer, S. Non-lethal PCR genotyping of single *Drosophila*. *Biotechniques* **46**, 312–314 (2009).
71. Gariou-Papalexiou, A. et al. Polytene chromosomes as tools in the genetic analysis of the Mediterranean fruit fly, *Ceratitis capitata*. *Genetica* **116**, 59–71 (2002).

Acknowledgements

This study was financially supported by the Joint FAO/IAEA Insect Pest Control Sub-programme of the Joint FAO/IAEA Programme of Nuclear Techniques in Food and Agriculture, the Emmy Noether program of the German Research Foundation (SHE 1833/1-1; to MFS), the LOEWE Center for Insect Biotechnology and Bioresources of the Hessen State Ministry for Higher Education, Research and the Arts (HMWK; to MFS), and the SITplus collaborative fruit fly program funded by the Hort Frontiers Fruit Fly Fund, part of the Hort Frontiers strategic partnership initiative developed by Hort Innovation, with co-investment from Macquarie University, USDA, and JLU Gießen and contributions from the Australian Government (FF17000 to MFS). The project was further supported by Hort Innovation, using the Apple & Pear, Strawberry, Citrus, Cherry, Summerfruit, Table Grape and Vegetable research and development levies (MT13059 to PC, AC, EF), with co-investment from South Australian Research and Development Institute (SARDI) and Primary Industries and Regions South Australia (PIRSA) and contributions from the Australian Government. Hort Innovation is the grower-owned, not-for-profit research and development corporation for Australian horticulture. SWB was supported by the Australian Research Council (FT140101303) and the Hermon Slade Foundation grant HSF 18/6. Furthermore, this work was supported by the Canadian Foundation for Innovation (33408, to JR) and Genome Canada Genome Technology Platform awards (JR), as well as the International Atomic Energy Agency research contracts no. 23358 (JR) and no. 23379 (FM) as part of the Coordinated Research Project “Generic approach for the development of genetic sexing strains for SIT applications”. The project benefitted from discussions at this CRP. Furthermore, resources were provided by the SCINet project of the USDA-ARS, the Detection, Control, and Area-wide Management of Fruit Flies and Other Quarantine Pests of Tropical/Subtropical Crops, and a USDA-NIFA grant (0500-00093-001-00-D, 2040-22430-026-00-D, and 2017-67012-26087 to SBS and SMG). The USDA-ARS is an equal opportunity/affirmative action employer, and all agency services are available without discrimination. Mention of commercial products and organizations in this manuscript is

solely to provide specific information. It does not constitute endorsement by USDA-ARS over other products and organizations not mentioned. The authors also wish to thank Tanja Rehling, Jakob Martin, and Johanna Rühl for technical assistance and Germano Sollazzo for input on injections and primers design (Justus-Liebig University Gießen and Insect Pest Control Laboratory); Elena Isabel Cancio Martinez, Thilakasiri Dammalage, Sohel Ahmad, and Gülizar Pillwax for insect rearing (Insect Pest Control Laboratory), Shu-Huang Chen (McGill University) for technical assistance with Nanopore library preparations; and Arjen van ’t Hof (University of Liverpool) for constructing libraries from micro-dissected chromosomes.

Author contributions

R.A.A., C.M.W., C.C., P.C., S.B.S., S.M.G., I.H., J.R., A.C.D., K.B., S.W.B., and M.F.S. designed the research; C.M.W., R.A.A., M.A.W., K.N., G.G., E.F., S.J.R., M.A.H., C.C., T.N.M.N., A.C., S.B.S., S.M.G., A.C.D., K.B., S.W.B., and M.F.S. performed the research; R.A.A., C.M.W., H.D., G.L., F.M., J.R., K.B., S.W.B., and M.F.S. contributed new reagents/analytic tools; C.M.W., R.A.A., M.A.W., K.N., G.L., G.G., H.D., S.W., T.N.M.N., A.C., S.B.S., S.M.G., I.H., J.R., A.C.D., K.B., S.W.B., and M.F.S. analyzed the data; R.A.A., C.M.W., K.N., G.L., G.G., S.J.R., S.W., A.C., S.B.S., S.M.G., I.H., J.R., A.C.D., K.B., S.W.B., and M.F.S. wrote the paper.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-20680-5>.

Correspondence and requests for materials should be addressed to K.B., S.W.B. or M.F.S.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

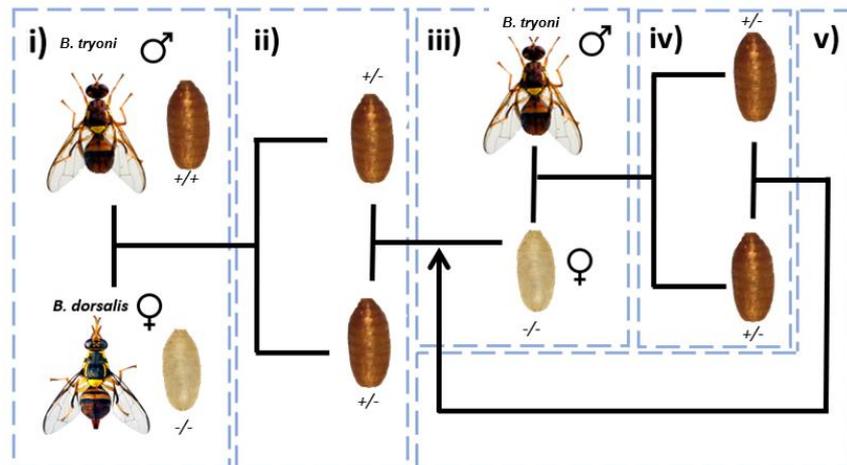


Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

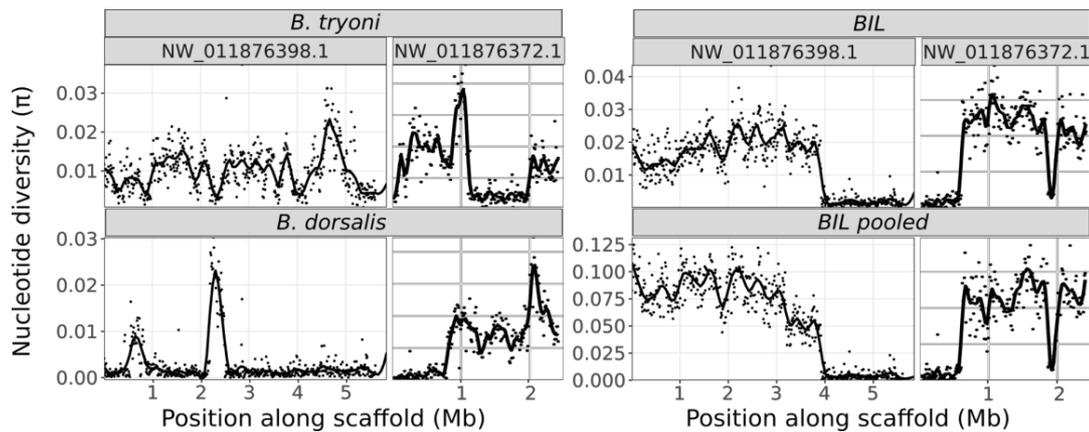
© The Author(s) 2021

White pupae phenotype of tephritids is caused by parallel mutations of a MFS transporter

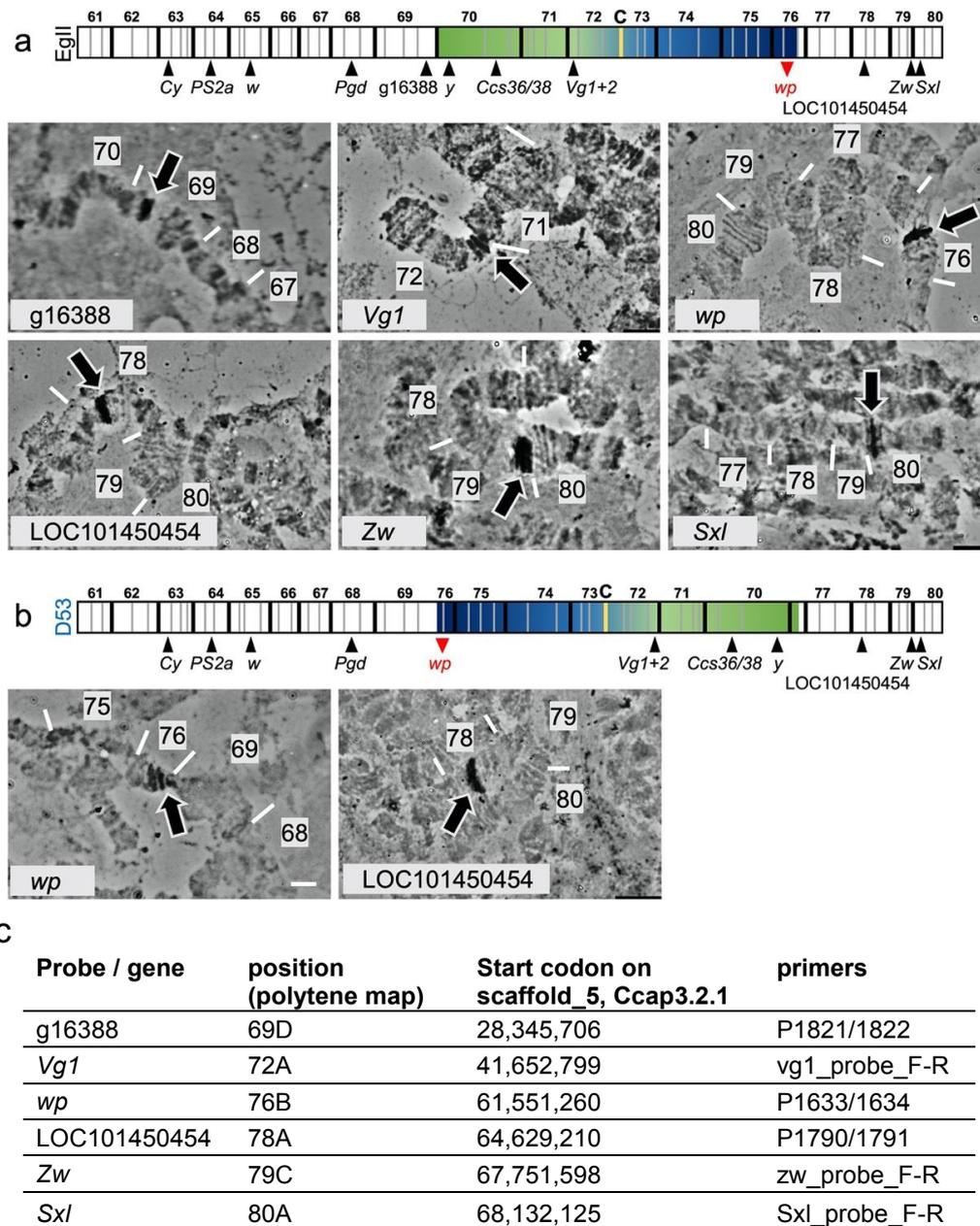
Ward and Aumann *et al.*



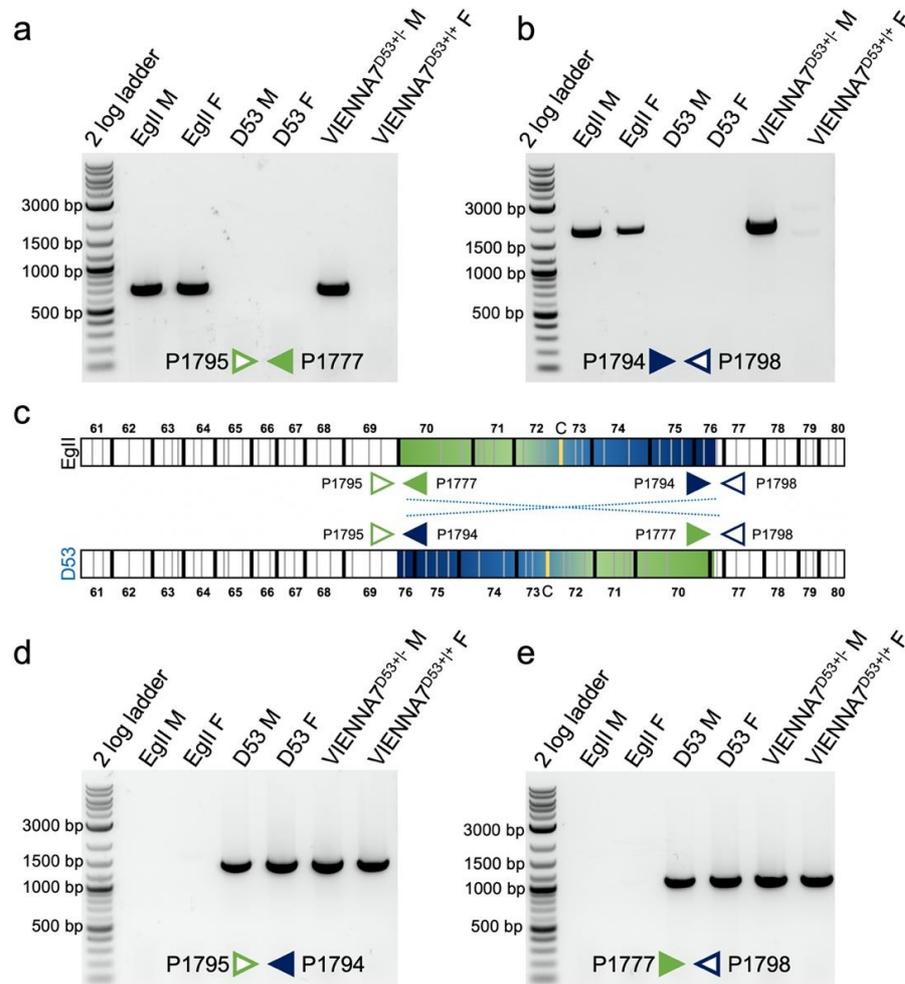
Supplementary Fig. 1. Introgression of the *B. dorsalis* white pupae locus and phenotype into *B. tryoni*. Stepwise methods are shown: i) Mass crosses between *B. tryoni* males and white pupae *B. dorsalis* females, ii) mass intercross of the F₁ progeny, iii) F₂ brown pupae were discarded and white pupae females were selectively backcrossed to *B. tryoni* males, iv) mass intercross progeny and v) reselect white pupae females and repeat backcrossing from step (iii) five additional times to produce the *Bactrocera Introgression Line* (BIL) with a white pupae phenotype.



Supplementary Fig. 2. Nucleotide diversity across two genomic scaffolds belonging to chromosome 5 (NW_011876398.1; 5.8 Mb and NW_011876372.1; 2.8 Mb) containing the *Bactrocera dorsalis* white pupae QTL. Calculations were carried out on resequenced genomes mapped to the *B. dorsalis* reference genome (GCF_000789215.1) treating each strain/species as a population. Upper left: *Bactrocera tryoni* Ourimbah strain individual pupa (wild type brown pupae, $n = 2$). Lower left: *Bactrocera dorsalis* Salaya1 strain (white pupae, $n = 2$). Upper right: *Bactrocera Introgression Line* (*BIL*, white pupae, $n = 2$). Lower right: pooled *BIL* individuals with two replicates of five samples per pool (white pupae). White pupae populations, *B. dorsalis*, *BIL* and *BIL pooled* show low levels of diversity between 3.8 - 5.8 Mb on NW_011876398.1 and 0 - 0.73 Mb, indicating the causal mutation is within this region.



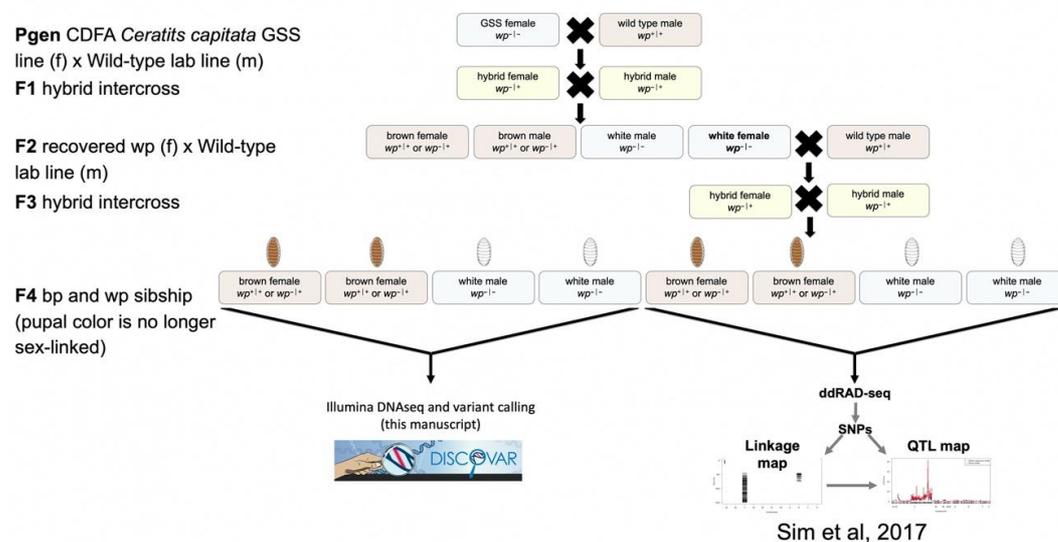
Supplementary Fig. 3. *In situ* hybridization of chromosome 5 marker genes in *Ceratitis capitata*. Known marker genes (*vitellogenin-1-like* (*Vg1*), *white pupae* (*wp*), *glucose-6-phosphate 1-dehydrogenase* (*Zw*), and *sex lethal* (*Sxl*)), and genes inside and outside the predicted inversion breakpoints (*g16388* and *LOC101450454*) were used to confirm the overall structure of scaffold 5 (*Ccap3.2.1*; accession [GCA_905071925](https://www.ncbi.nlm.nih.gov/nuclot/GCA905071925)) and the D53 inversion in WT *EgII* (a) and D53 (b). Chromosome 5 is shown as schematic illustration, including the additional marker genes *Curly* (*Cy*), *integrin- α PS2* (*PS2a*), *white* (*w*), *6-phosphogluconate dehydrogenase* (*Pgd*), *yellow* (*y*), and *chorion S36/38* (*Ccs36/38*). The inverted part (D53) is shown in a green-to-blue color gradient and the centromere is marked as 'C'. Exact position of the genes on the polytene map, the position of the start codons on *EgII* scaffold_5 in *Ccap3.2.1* and primers used to amplify the probes for *in situ* hybridization are named in (c), primer sequences are shown in Supplementary Table 5. *In situ* hybridizations were done at least in duplicates and at least ten nuclei were analyzed per sample, scale bar = 10 μ m. All replicates led to similar results. The source data underlying Supplementary Figure 3a and 3b are provided as a Source Data file.



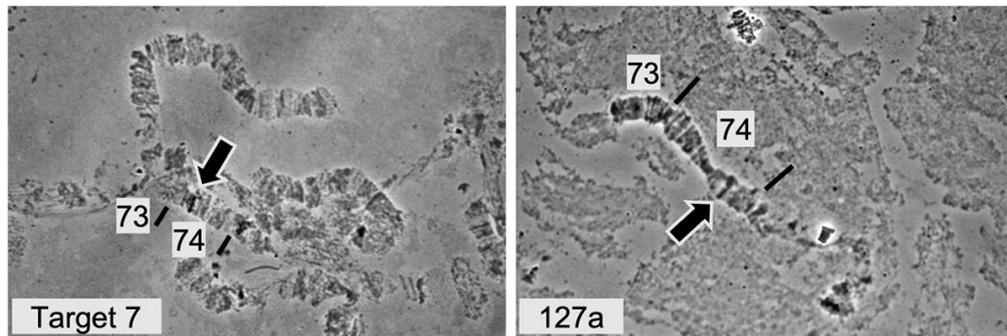
Supplementary Fig. 4. Verification of the D53 inversion breakpoints via PCRs. Two primer pairs were used to amplify the sequences spanning the inversion breakpoints in EgII (WT), D53 (homozygous for the inversion), and VIENNA7^{D53+} males (heterozygous for the inversion) and females (homozygous for the inversion). Schematic primer positions are shown in (c) (indicated by triangles), together with schematics of the chromosome 5 without (top) and with (bottom) the D53 inversion, which is shown as a reversed blue-to-green color gradient. The centromere is marked as 'C'. The left breakpoint in strains without inversion was amplified with primers P_1795 and P_1777 (a), the right breakpoint with the primer pair P_1794/P_1798 (b). For the amplification of the breakpoints in strains with inversion, the internal primers were used vice versa: (d) P_1795/P_1794, (e) P_1777/P_1798. Primers P_1794 and P_1798 were designed based on EgII chromosome 5, flanking the predicted breakpoint position (61,880,224 bp; Ccap3.2.1, accession [GCA_905071925](https://www.ncbi.nlm.nih.gov/nuccore/GCA_905071925)). Primers P_1795 and P_1777 were designed based on D53 nanopore data (D53 assembly contig_531), as the left breakpoint is within a scaffold gap on EgII chromosome 5. The ladder used for agarose gels is the NEB 2-log DNA-ladder, bp = base pairs; M = male; F = female. PCRs were done at least twice to confirm the result; primer sequences are shown in Supplementary Table 5. The source data underlying Supplementary Figure 4a, 4b, 4d, and 4e are provided as a Source Data file.



Supplementary Fig. 5. *Bactrocera tryoni* pupae exhibiting somatic mosaicism after embryonic injection of CRISPR/Cas9 guide RNA targeting exon 1 of the *white pupae* gene.



Supplementary Fig. 6. Relationship of the whole genome sequenced *C. capitata* individuals to the mapping population described in Sim *et al*¹: utilizing pangenomic assemblies to characterize region and mutations associated with white pupae in *Ceratitis capitata*. The putative region of the medfly genome thought to contain the *white pupae* mutation was previously identified as being on NW_004523946.1 of the AOHK00000000.1 WGS accession for *C. capitata*, determined by genotyping a mapping population of a wild type strain with the Vienna genetic sexing strain¹. However, genotyping was performed using restriction site associated genotype by sequencing, and the resolution of the map only gave resolution within several million base pairs of the actual causative mutation. To better explore putative mutations associated with the white pupae phenotype, single individual *de novo* assemblies were generated from stable white and stable brown individuals from the mapping population. Three white males (wp^{-1-}) (from population W8_3_2_2) and three brown females (wp^{+1+} or wp^{-1+}) (from population B8_3_2_1) from a common Pgen and parents were sequenced using PCR free library preparation, and each was sequenced to ~50x coverage using Illumina 2 x 250 paired-end sequencing. These libraries are made public on NCBI under SRA accessions SRR11649127-SRR11649132 associated with BioProject [PRJNA629430](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA629430). The libraries were sized selected to allow overlap of the read pairs through double sided Ampure size selection. The subsequent data was assembled in tandem using the DISCOVAR *de-novo* assembler, using the AOHK00000000.1 WGS assembly available on NCBI to anchor all of the individual-scale assemblies to each other and call variants relative to each other². This approach creates a single assembly graph, with all the individual assemblies being represented within this graph which was visualized using the NhoodInfo tool. A secondary method, using GATK 3.3.0 and the Unified Genotyper pipeline³ was used to create a .vcf containing variants between the genomes. To target for mutations associated with the white pupae phenotype, this was screened across the scaffolds within 20 Mb of the QTL loci and filtered for assigned genotype probability >0.995. Final variants from this pipeline were filtered using vcfutils⁴ to subset to those that were consistent between phenotypes, identifying loci that are homozygous for one allele in white pupae individuals and homozygous for another allele or heterozygous in brown pupae individuals. SNPs between the two variant calling approaches, in graph space, and mapping base VCF were merged, and the SNPEff tool⁵ was used to filter for SNPs causing high impact in protein coding regions based off of the annotation set accompanying the genome in NCBI. This resulted in a subset of variants on NW_004523946.1 that were putatively identified as causative, or directly linked to the causative mutation (Supplementary Table 3).



Supplementary Fig. 7. *In situ* hybridization on target 7 (see Supplementary Table 3), the region surrounding the locus with the highest LOD score and perfect LD with the white pupae phenotype, described in Sim *et al.*¹ (59,869,348 - 59,871,012 on scaffold_5, Ccap3.2.1), and 127a, a randomly chosen gene on the same scaffold in Ccap2.1 (LOC101458400; 59,396,172 - 59,397,378 on scaffold_5, Ccap3.2.1). *In situ* hybridizations were done at least in duplicates and at least ten nuclei were analyzed per sample. Pictures were taken without a scale bar and cannot be scaled. All replicates led to similar results. Source data are provided as a Source Data file.

1. *Bactrocera dorsalis*_XP_011212119
2. *Ceratitis capitata*_XP_004530515
3. *Zeugodacus cucurbitae*_XP_028898713
4. *Musca domestica*_XP_005176653
5. *Lucilia cuprina*_XP_023297119
6. *Drosophila melanogaster*_NP_001188550
7. *Aedes albopictus*_XP_029727765
8. *Anopheles gambiae*_XP_321808
9. *Tribolium castaneum*_NP_001308590
10. *Bombyx mori*_NP_001243989.1

```

10 20 30 40 50 60 70
MVPVKKVSSYVNSVAIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
MLPVLKVVSGFYNAYVIGVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
MVPVKKVSSYVNSVAIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
MIPVSSKLSGYSVAVIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
MIPVSSKLSGYSVAVIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
MISLVDMKRFLLKPVVIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
MMKRLSFLKPKPVVIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
MLPFLKLSFLKPKPVVIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
MFRSGLLGKLGDFYKVMVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
    
```

1. *Bactrocera dorsalis*_XP_011212119
2. *Ceratitis capitata*_XP_004530515
3. *Zeugodacus cucurbitae*_XP_028898713
4. *Musca domestica*_XP_005176653
5. *Lucilia cuprina*_XP_023297119
6. *Drosophila melanogaster*_NP_001188550
7. *Aedes albopictus*_XP_029727765
8. *Anopheles gambiae*_XP_321808
9. *Tribolium castaneum*_NP_001308590
10. *Bombyx mori*_NP_001243989.1

```

80 90 100 110 120 130 140
AAVMEEESCVAELSLNDIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
SLVKERDSCVAELSLNDIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
SMVKERDSCVAELSLNDIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
SAVONRTSCVAELSLNDIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
ADIKNRTSCVAELSLNDIAVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
AVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
AEELHLEHGHGQHVINGVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
AEFLIHEHGHGQHVINGVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
DTANSSEICLGFNININGVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
EKVNSSEICLGFNININGVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
    
```

1. *Bactrocera dorsalis*_XP_011212119
2. *Ceratitis capitata*_XP_004530515
3. *Zeugodacus cucurbitae*_XP_028898713
4. *Musca domestica*_XP_005176653
5. *Lucilia cuprina*_XP_023297119
6. *Drosophila melanogaster*_NP_001188550
7. *Aedes albopictus*_XP_029727765
8. *Anopheles gambiae*_XP_321808
9. *Tribolium castaneum*_NP_001308590
10. *Bombyx mori*_NP_001243989.1

```

150 160 170 180 190 200 210
AVAALIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
AVAALIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
AVAALIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
AVAALIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
CIPAMVIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
AVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
AVAALIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
AVAVLIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
VVAALIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
    
```

1. *Bactrocera dorsalis*_XP_011212119
2. *Ceratitis capitata*_XP_004530515
3. *Zeugodacus cucurbitae*_XP_028898713
4. *Musca domestica*_XP_005176653
5. *Lucilia cuprina*_XP_023297119
6. *Drosophila melanogaster*_NP_001188550
7. *Aedes albopictus*_XP_029727765
8. *Anopheles gambiae*_XP_321808
9. *Tribolium castaneum*_NP_001308590
10. *Bombyx mori*_NP_001243989.1

```

220 230 240 250 260 270 280
VYTRSNFENLWVRMAYLGLVAVLVAALIGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
    
```

1. *Bactrocera dorsalis*_XP_011212119
2. *Ceratitis capitata*_XP_004530515
3. *Zeugodacus cucurbitae*_XP_028898713
4. *Musca domestica*_XP_005176653
5. *Lucilia cuprina*_XP_023297119
6. *Drosophila melanogaster*_NP_001188550
7. *Aedes albopictus*_XP_029727765
8. *Anopheles gambiae*_XP_321808
9. *Tribolium castaneum*_NP_001308590
10. *Bombyx mori*_NP_001243989.1

```

290 300 310 320 330 340 350
INPAMIMIAASTIRHSGGMIFAYNADLYVNYVPPVDLGGWLLFAVLTIGLGSVGVVVGAVSDFKIVAKMG
ANPAMIMIAASTIRHSGGMIFAYNADLYVNYVPPVDLGGWLLFAVLTIGLGSVGVVVGAVSDFKIVAKMG
VNPAMIMIAASTIRHSGGMIFAYNADLYVNYVPPVDLGGWLLFAVLTIGLGSVGVVVGAVSDFKIVAKMG
KNPAMIMIAASTIRHSGGMIFAYNADLYVNYVPPVDLGGWLLFAVLTIGLGSVGVVVGAVSDFKIVAKMG
KNPAMIMIAASTIRHSGGMIFAYNADLYVNYVPPVDLGGWLLFAVLTIGLGSVGVVVGAVSDFKIVAKMG
MOPRVLLMIAASTIRHSGGMIFAYNADLYVNYVPPVDLGGWLLFAVLTIGLGSVGVVVGAVSDFKIVAKMG
LQPRVLLMIAASTIRHSGGMIFAYNADLYVNYVPPVDLGGWLLFAVLTIGLGSVGVVVGAVSDFKIVAKMG
LQPRVLLMIAASTIRHSGGMIFAYNADLYVNYVPPVDLGGWLLFAVLTIGLGSVGVVVGAVSDFKIVAKMG
SOBRLMIAASTIRHSGGMIFAYNADLYVNYVPPVDLGGWLLFAVLTIGLGSVGVVVGAVSDFKIVAKMG
    
```

1. *Bactrocera dorsalis*_XP_011212119
2. *Ceratitis capitata*_XP_004530515
3. *Zeugodacus cucurbitae*_XP_028898713
4. *Musca domestica*_XP_005176653
5. *Lucilia cuprina*_XP_023297119
6. *Drosophila melanogaster*_NP_001188550
7. *Aedes albopictus*_XP_029727765
8. *Anopheles gambiae*_XP_321808
9. *Tribolium castaneum*_NP_001308590
10. *Bombyx mori*_NP_001243989.1

```

360 370 380 390 400 410 420
IRSRALVIAVLSQIAIHPAFGGSVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
    
```

1. *Bactrocera dorsalis*_XP_011212119
2. *Ceratitis capitata*_XP_004530515
3. *Zeugodacus cucurbitae*_XP_028898713
4. *Musca domestica*_XP_005176653
5. *Lucilia cuprina*_XP_023297119
6. *Drosophila melanogaster*_NP_001188550
7. *Aedes albopictus*_XP_029727765
8. *Anopheles gambiae*_XP_321808
9. *Tribolium castaneum*_NP_001308590
10. *Bombyx mori*_NP_001243989.1

```

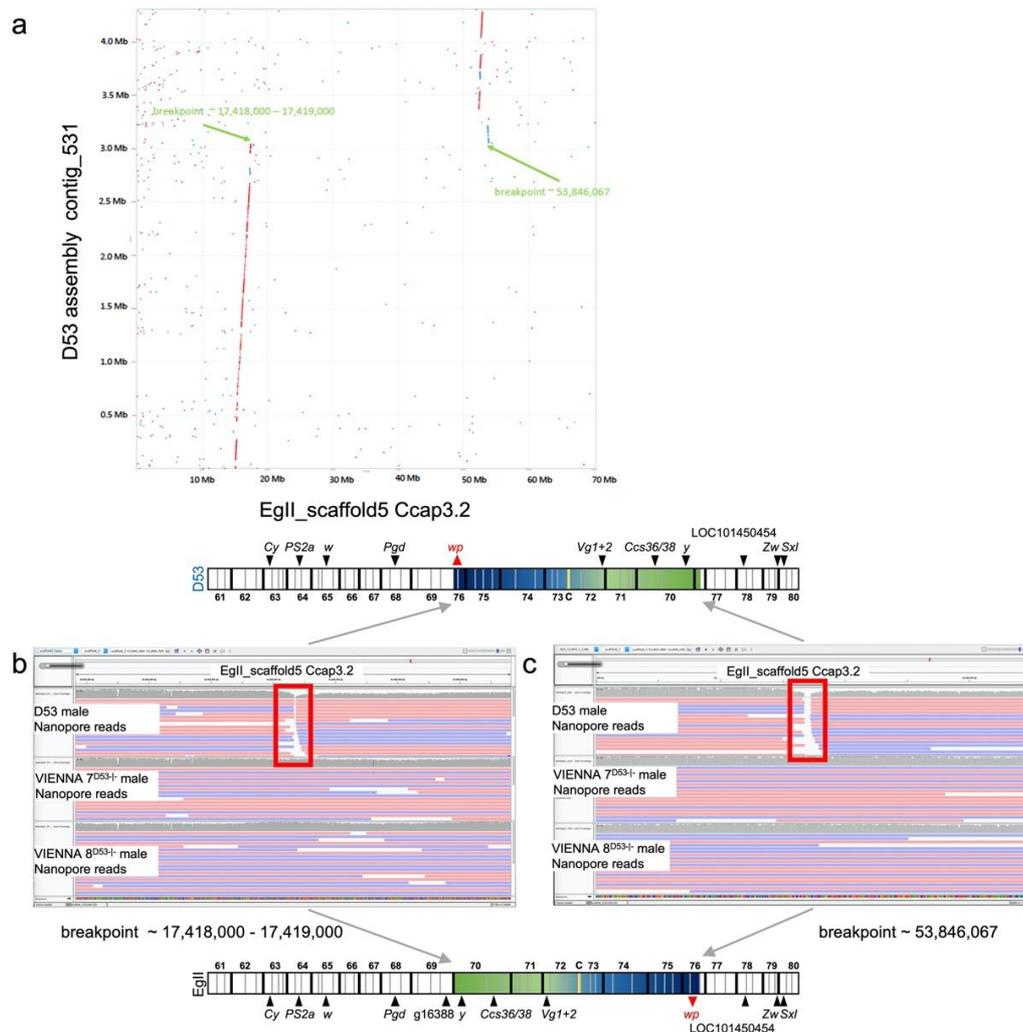
430 440 450 460 470 480 490
VMNNVGGNLPVLDVPAKALIGYRGSIMHLYAGVYVIGELGHVLLGVTSSKQIAIFELYGDACACGNINMIFTRHDLRQOC
    
```

1. *Bactrocera dorsalis*_XP_011212119
2. *Ceratitis capitata*_XP_004530515
3. *Zeugodacus cucurbitae*_XP_028898713
4. *Musca domestica*_XP_005176653
5. *Lucilia cuprina*_XP_023297119
6. *Drosophila melanogaster*_NP_001188550
7. *Aedes albopictus*_XP_029727765
8. *Anopheles gambiae*_XP_321808
9. *Tribolium castaneum*_NP_001308590
10. *Bombyx mori*_NP_001243989.1

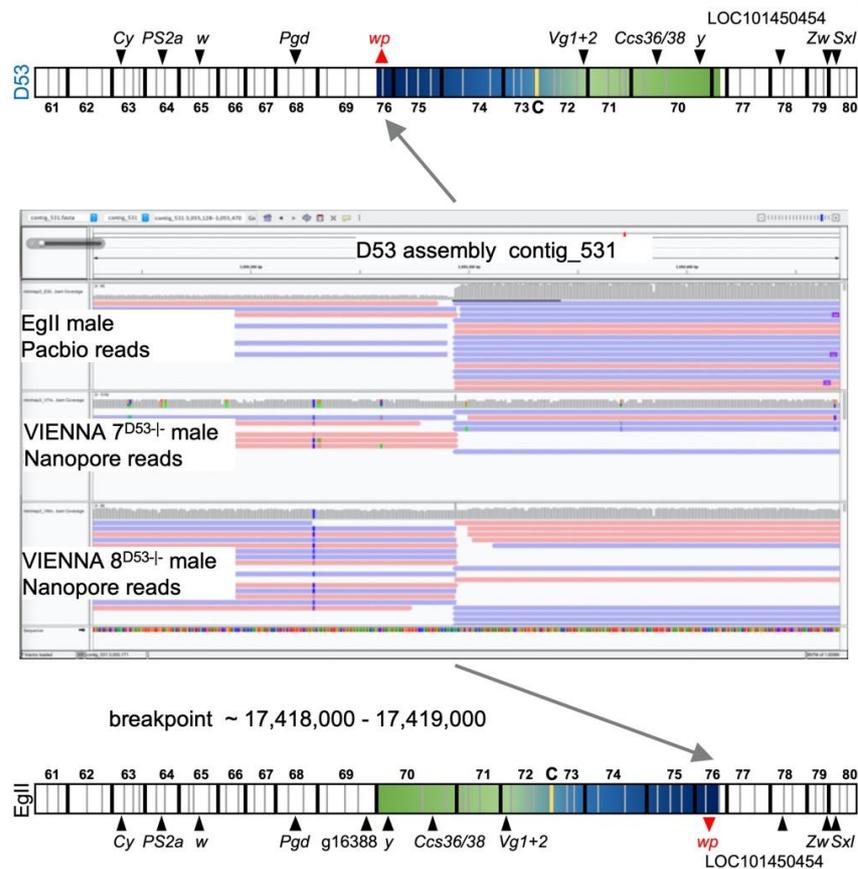
```

500 510 520 530 540 550 553
EAHTRQMR-GHNSVSVVIEVLPNHG--RPAVLPQHLQLSSNGGDK--YTPATR--EISRL
EAHTRQMR-GHNSVSVVIEVLPNHG--RPAVLPQHLQLSSNGGDK--YTPATR--EISRL
EAHTRQMR-GHNSVSVVIEVLPNHG--RPAVLPQHLQLSSNGGDK--YTPATR--EISRL
EVPARNIQ-GHNSVSVVIEVLPNHG--RPAVLPQHLQLSSNGGDK--YTPATR--EISRL
VLAARHMH-GHNSVSVVIEVLPNHG--RPAVLPQHLQLSSNGGDK--YTPATR--EISRL
DEHRNQMEGHNSVSVVIEVLPNHG--RPAVLPQHLQLSSNGGDK--YTPATR--EISRL
TSHANHMEGHNSVSVVIEVLPNHG--RPAVLPQHLQLSSNGGDK--YTPATR--EISRL
--GHNSVSVVIEVLPNHG--RPAVLPQHLQLSSNGGDK--YTPATR--EISRL
--GHNSVSVVIEVLPNHG--RPAVLPQHLQLSSNGGDK--YTPATR--EISRL
    
```

Supplementary Fig. 8. Amino acid sequence alignment of insect *white pupae* homologs using MUSCLE. Identical residues are shaded in black, conserved residues in shades of grey and the MFS domain is underlined.



Supplementary Fig. 9. Identification of the D53 inversion breakpoints. (a) The Egll (wild type; Egll) assembly was plotted vs the D53 Nanopore assembly. The Egll Ccap 3.2 assembly scaffold_5 is shown on the x-axis compared to the D53 assembly contig 531 using the MashMap long read mapping tool, providing a visual overview of the local alignment boundaries. The D53 assembly breaks its homology with the Egll assembly at the 17.4 Mbp position of scaffold 5 and continues further at position 53.84 Mbp, providing identification of candidate breakpoints. (b)-(c): Zoomed in view using Integrative Genome Viewer (IGV, <http://software.broadinstitute.org/software/igv/>) of D53 male (top) VIENNA 7^{D53-/-} male (middle) and VIENNA 8^{D53-/-} male (bottom) Nanopore reads aligned against the Egll Ccap 3.2 genome reference scaffold 5. Reads aligned in the sense direction are shown in red, reads aligning in antisense in blue. The alignments of D53 Nanopore reads (top) break in the positions of 17,418,000 for the left breakpoint on chromosome 5 (b) and at position 53,846,067 corresponding to the right breakpoint (c). However, VIENNA 7^{D53-/-} male and VIENNA 8^{D53-/-} male Nanopore reads align perfectly to the Egll reference without breaks. Annotated ideograms of the D53 inversion chromosome and the Egll wild type chromosome are shown at the top and bottom, respectively. Marker genes are described in Supplementary Fig. 3.



Supplementary Fig. 10. Example of D53 inversion breakpoint confirmation. Zoomed in view, using IGV, of the EgII male PacBio reads (top) VIENNA 7^{D53-/-} male (middle) and VIENNA 8^{D53-/-} male (bottom) Nanopore reads aligned against the D53 genome assembly (Nanopore reads) contig 531. Reads aligned in the sense direction are shown in red, reads aligning in antisense in blue. The alignments of EgII PacBio reads (top) break in the position of 17,418,000 for the left breakpoint on chromosome 5, along with VIENNA 7^{D53-/-} male and VIENNA 8^{D53-/-} (middle and bottom) Nanopore reads consistent with the presence of inversion only in the D53 genome. Annotated ideograms of the D53 inversion chromosome and the EgII wild type chromosome are shown at the top and bottom, respectively. Marker genes are described in Supplementary Fig. 3.

Supplementary Table 1. Strains used for genomic and transcriptomic analysis.

No.	Species	Strain	Sex	Genotype	Inversion	Trans-location	Pupa color	Illumina MiSeq Chromosome dissection	PacBio WGS	Hi-C WGS	Nanopore WGS	Illumina WGS	Illumina NovaSeq 6000 RNAseq
1	<i>C. capitata</i>	EgyptII (EgII)	Male	WT	NO	NO	brown	Y (ERS4426864), X (ERS4426857-ERS4426863)	Yes (ERS4426857-ERS4426863)	Yes (ERS45195-ERS45195)	Yes (ERS45475-91)	Yes (ERS4426868) HiSeq 4000	
2	<i>C. capitata</i>	EgII	Female	WT	NO	NO	brown					Yes (ERS4426869) HiSeq 4000	
3	<i>C. capitata</i>	Benakeion	Larvae / prepupae	WT	NO	NO	brown						Yes (ERS4426994 - ERS4426999)
4	<i>C. capitata</i>	<i>wp/ts/</i> (in EgII)	Larvae / prepupae	<i>wp/ts/</i> homo	NO	NO	white						Yes (ERS4427000 - ERS4427005)
5	<i>C. capitata</i>	D53 (in <i>wp/ts/</i> EgII)	Male	<i>wp/ts/</i> homo	YES	NO	white				Yes (ERS45475-90)	Yes (ERS4426872) HiSeq 4000	
6	<i>C. capitata</i>	D53 (in <i>wp/ts/</i> EgII)	Female	<i>wp/ts/</i> homo	YES	NO	white					Yes (ERS4426873) HiSeq 4000	
7	<i>C. capitata</i>	Vienna 7 ^{D63}	Male	<i>wp/ts/</i> hetero	NO	YES (Vienna 7)	brown	T(Y;5) (ERS4426866), X (ERS4426867)			Yes (ERS45475-92)	Yes (ERS4426870) HiSeq 4000	
8	<i>C. capitata</i>	Vienna 7 ^{D63}	Female	<i>wp/ts/</i> homo	NO	NO	white					Yes (ERS4426871) HiSeq 4000	
9	<i>C. capitata</i>	Vienna 8 ^{D63}	Male	<i>wp/ts/</i> hetero	NO	YES (Vienna 8)	brown				Yes (ERS45475-93)		
10	<i>B. dorsalis</i>	Saramburi	Larvae / prepupae	WT	NO	NO	brown						Yes (ERS4427006 - ERS4427011)
11	<i>B. dorsalis</i>	white pupae	Larvae / prepupae	<i>wp</i> homo	NO	NO	white						Yes (ERS4427012 - ERS4427017)

12	<i>B. dorsalis</i>	white pupae	Male	wp homo	NO	NO	white					Yes (SRR13206143) NovaSeq 6000	
13	<i>B. dorsalis</i>	white pupae	Female	wp homo	NO	NO	white					Yes (SRR13206144) NovaSeq 6000	
14	<i>B. tryoni</i>	Ourimbah	Male	WT	NO	NO	brown					Yes (SRR13206145) NovaSeq 6000	
15	<i>B. tryoni</i>	Ourimbah	Female	WT	NO	NO	brown					Yes (SRR13206146) NovaSeq 6000	
16	Hybrid	Bactrocera Introgressed Line	Male	wp homo	NO	NO	white					Yes (SRR13206141) NovaSeq 6000	
17	Hybrid	Bactrocera Introgressed Line	Male	wp homo	NO	NO	white					Yes (SRR13206142) NovaSeq 6000	
18	Hybrid	Bactrocera Introgressed Line	Pooled males	wp homo	NO	NO	white					Yes (SRR13206139) NovaSeq 6000	
19	Hybrid	Bactrocera Introgressed Line	Pooled Females	wp homo	NO	NO	white					Yes (SRR13206140) NovaSeq 6000	
20	<i>Z. cucurbitae</i>	wp GSS	Larvae / prepupae	Mixed. Individual preps. Sex, and by extrapolation the pupal color, was confirmed by MoY PCR	NO	YES	brown/ white					Yes (brown: ERS4427018 - ERS4427023; white: ERS4427024 - ERS4427029)	

Note: Accession numbers ENA BioProject [PRJEB36344/ERP119522](https://www.ncbi.nlm.nih.gov/bioproject/PRJEB36344/ERP119522), and BioProject [PRJNA682907](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA682907).

Supplementary Table 2. Illumina pair-end sequencing of *Ceratitis capitata* strains and micro-dissected chromosomes.

Strain	Chromosome (N) / Sex	Coverage	Accession number
Egypt II	Y (20)	~7 x	ERS4426864
Egypt II	X (20)	~6 x	ERS4426865
VIENNA 7 ^{D53} -	T(Y;5) (25)	~6 x	ERS4426866
VIENNA 7 ^{D53} -	X (20)	~8 x	ERS4426867
Egypt II	Male	~40 x	ERS4426868
Egypt II	Female	~45 x	ERS4426869
VIENNA 7 ^{D53} -	Male	~35 x	ERS4426870
VIENNA 7 ^{D53} -	Female	~30 x	ERS4426871
D53	Male	~32 x	ERS4426872
D53	Female	~27 x	ERS4426873

Note: Third instar larvae mitotic chromosomes were laser micro-dissected^{6,7} and used to construct libraries using the PicoPLEX WGA kit. These libraries were then multiplexed into a single run of Illumina Miseq platform for 250 bp paired-end sequencing. For the whole genome sequencing of Egypt II, VIENNA 7^{D53}- and D53 strains, fresh pupae (1 day after pupation) were used as the insect material for DNA extraction using the phenol/chloroform approach (see methods). This DNA was analyzed for size and quality. Single insects (males or females) were used from each of the three insect lines to make 6 indexed Illumina TruSeq PCR-free libraries (550 bp inserts) from the extracted genomic DNA. These libraries were then multiplexed into one lane of the Illumina HiSeq 4000 150 bp platform for paired-end sequencing. Sequence data are available at: ENA BioProject accession number [PRJEB36344](https://ena.ebi.ac.uk/ena/browser/view/PRJEB36344)/ERP119522.

Supplementary Table 3. Position of high impact variants identified using the union of the DISCOVAR and GATK genotyping methods performed on WGS data from phenotypic white and brown F4 *Ceratitis capitata* individuals from the mapping population described in Sim *et al*¹ (see Supplementary Fig. 6). No. 7 is described as the SNP locus displaying the strongest linkage to the wp phenotype.

No.	Scaffold	Position	Ref	Alt
1	NW_004523946.1	800831	C	T
2	NW_004523946.1	837972	C	C
3	NW_004523946.1	1576424	G	A
4	NW_004523946.1	2259830	AAC	A
5	NW_004523946.1	2262779	C	A
6	NW_004523946.1	2410888	A	ACAACAGGCATGCCAGCAAGTTGT GGCCGTCTTCCAACAACATGCTGCT ACAACACTACAACAGCCAAATGACGA GCCC GCCGTTGCAGCCTCAGCACCA GCCAAGGCTACATATGCAACACTGC GACATGCGATGGTTGTAGAGGCGC AAGC
7	NW_004523946.1	1353742	G	T

Supplementary Table 4. Nanopore sequencing run metrics.

Sample	Flow cell ID	Pores	Library kit	Bases (Gb)	N50 (kb)	MinKNOW PromethION Release	Guppy
Egypt II male	PAD68071	2827	SQK-LSK109	7.64	6.82	19.06.9	3.0.3
Egypt II male	PAD68071	159	SQK-RAD004	0.06	0.87	19.06.9	3.0.3
D53 male	PAD66870	6288	SQK-LSK109	19.88	14.04	19.06.9	3.0.3
D53 male	PAD66870	2991	SQK-RAD004	6.42	6.73	19.06.9	3.0.3
D53 male	PAD75166	2081	SQK-LSK109	4.79	0.91	19.10.2	3.2.6
VIENNA 7 ^{D53} -male	PAD60029	25	SQK-LSK109	0.21	4.86	19.06.9	3.0.3
VIENNA 7 ^{D53} -male	PAD73592	3270	SQK-LSK109	7.47	5.37	19.10.2	3.2.6
VIENNA 7 ^{D53} -male	PAE49665	5025	SQK-LSK109	9.34	5.59	19.12.5	3.2.8
VIENNA 7 ^{D53} -male	PAE47816	3749	SQK-LSK109	5.73	6.42	19.12.5	3.2.8
VIENNA 7 ^{D53} -male	PAE47455	1503	SQK-RAD004	3.97	3.88	19.12.5	3.2.8
VIENNA 8 ^{D53} -male	PAE03818	3557	SQK-LSK109	7.68	6.22	19.06.9	3.0.3
VIENNA 8 ^{D53} -male	PAD88883	3209	SQK-LSK109	7.02	6.19	19.10.2	3.2.6
VIENNA 8 ^{D53} -male	PAE16370	2473	SQK-LSK109	6.90	5.73	19.10.2	3.2.6
VIENNA 8 ^{D53} -male	PAE47104	3143	SQK-RAD004	3.23	3.41	19.12.5	3.2.8

Supplementary Table 5. List of primers used in this study.

Name	Sequence	Purpose
P_1753	GAAATTAATACGACTCACTATAGGCAT GCCGCCAGAGTGACGAAGTTTTAGAG CTAGAAATAGC	gRNA_MFS (<i>C. capitata</i>) <i>in vitro</i> synthetization
P_369	GCACCGACTCGGTGCCACTTTTTCAA GTTGATAACGGACTAGCCTTATTTTAA CTTGCTATTTCTAGCTCTAAAAC	gRNA_MFS (<i>C. capitata</i>) <i>in vitro</i> synthetization
P_1643	TTGAAGAGCGCACTTGCAAC	<i>Cc_wp</i> non-lethal genotyping G ₁
P_1644	TTCCCAACAGTGAATCCGG	<i>Cc_wp</i> non-lethal genotyping G ₁
P_1657	AAACGCTCTACAGATTGTGGA	multiplex PCR non-lethal genotyping G ₂
P_1794	ATCTACCAAATGAGAGAGAGAGCG	D53 inversion verification
P_1795	TTTTTCAAACCACTTGAACAACGC	D53 inversion verification
P_1777	TCCAGTGTTCTCTACTATGTTGCC	D53 inversion verification
P_1798	TCAGCTAACAGAACATGAATTCCG	D53 inversion verification
BtMFS_5primeF	TTTTTGCTTATCCCACTTCTGATT	PCR amplicons spanning both <i>BtMFS</i> guide recognition sites
BtMFS_exon2R	ACACCAGCAATTGTAAAGACCA	PCR amplicons spanning both <i>BtMFS</i> guide recognition sites
ZcMoY1F	AAGCCAGATCACGCAATCC	maleness-specific PCR on the <i>MoY</i> gene of <i>Z. cucurbitae</i>
ZcMoY1R	AGGACATCGTTATCTCCCCTG	maleness-specific PCR on the <i>MoY</i> gene of <i>Z. cucurbitae</i>
Bt_MFS-1	TGTGAGTACGGCCAACGCAT	customized 20 bp crRNA sequence
Bt_MFS-2	CGATCTACCACAGCAATGTG	customized 20 bp crRNA sequence
P_1790	AATCAAGTAAAGACAAAGCGGACG	<i>in situ</i> probe , <i>Cc</i> _LOC101450454 ‘
P_1791	TCATACGAACAGTTTGCCATAACG	<i>in situ</i> probe , <i>Cc</i> _LOC101450454 ‘
P_1821	TAAATGATTTGTCCGCTGAAGCC	<i>in situ</i> probe , <i>Cc</i> _g16388‘
P_1822	GAACACTATCCATGCTCTTGCC	<i>in situ</i> probe , <i>Cc</i> _g16388‘
P_1395	TCTCTGGGCAGCTCAAAGTG	<i>in situ</i> probe , <i>Cc</i> _target 7‘
P_1396	AAACCAACATTGCGGGCTC	<i>in situ</i> probe , <i>Cc</i> _target 7‘
P_1415	TCTCCCACACACTCAGGTCA	<i>in situ</i> probe , <i>Cc</i> _127a‘
P_1416	ACTCTCGTTGTCTGCTTGCA	<i>in situ</i> probe , <i>Cc</i> _127a‘
P_1633	TCCAGTGCAGTTCGGCTTAA	<i>in situ</i> probe <i>Cc_wp</i>
P_1634	CGGCTTTTACAACGCTTATGTTG	<i>in situ</i> probe <i>Cc_wp</i>
Zc_F	GTCATGACCACGCATTTGACG	<i>in situ</i> probe <i>Zc_wp</i>
Zc_R	GTTTATGGGTTTCGCCGCTG	<i>in situ</i> probe <i>Zc_wp</i>

Name	Sequence	Purpose
Bd_F	GCTGTTGCTGTTGCGTATGG	<i>in situ</i> probe <i>Bd_wp</i>
Bd_R	GTGGCGGGCGTATATTTGTC	<i>in situ</i> probe <i>Bd_wp</i>
Pgd_probe_F	TTGCTTTCTCTCCTTCTGCTT	<i>in situ</i> probe <i>Cc_Pgd</i>
Pgd_probe_R	TTCAAGTACTCACAAACGCTTGA	<i>in situ</i> probe <i>Cc_Pgd</i>
vg1_probe_F	ATCTCCGATAATCTCACAGGAAAT	<i>in situ</i> probe <i>Cc_Vg1</i>
vg1_probe_R	TCAGAGCGGGTCCATCGAAT	<i>in situ</i> probe <i>Cc_Vg1</i>
Sxl_probe_F	GCCAAAGAGATTGTTGTGTCAC	<i>in situ</i> probe <i>Cc_Sxl</i>
Sxl_probe_R	AAATTCCTGCACTGCGCTGT	<i>in situ</i> probe <i>Cc_Sxl</i>
y_probe_F	CCGTCGCCACTGTTGCTATT	<i>in situ</i> probe <i>Cc_y</i> (LOC101455502)
y_probe_R	AGTCGGGGTTGGTTGTTGTT	<i>in situ</i> probe <i>Cc_y</i> (LOC101455502)
zw_probe_F	TGCACACTGCTGCCATAGAT	<i>in situ</i> probe <i>Cc_Zw</i>
zw_probe_R	AGCCTCCTTAGCGGTTACAC	<i>in situ</i> probe <i>Cc_Zw</i>

Supplementary References

- 1 Sim, S. B., Ruiz-Arce, R., Barr, N. B. & Geib, S. M. A new diagnostic resource for *Ceratitis capitata* strain identification based on QTL mapping. *G3 (Bethesda)* **7**, 3637-3647 (2017).
- 2 Love, R. R., Weisenfeld, N. I., Jaffe, D. B., Besansky, N. J. & Neafsey, D. E. Evaluation of DISCOVAR *de novo* using a mosquito sample for cost-effective short-read genome assembly. *BMC Genomics* **17**, 187 (2016).
- 3 Poplin, R. *et al.* Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*, 201178 (2018).
- 4 Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156-2158 (2011).
- 5 Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80-92 (2012).
- 6 Fukova, I. *et al.* Probing the W chromosome of the codling moth, *Cydia pomonella*, with sequences from microdissected sex chromatin. *Chromosoma* **116**, 135-145 (2007).
- 7 Kubickova, S., Cernohorska, H., Musilova, P. & Rubes, J. The use of laser microdissection for the preparation of chromosome-specific painting probes in farm animals. *Chromosome Res* **10**, 571-577 (2002).

4 Discussion

The Sterile Insect Technique is an environmentally friendly method to suppress, contain, prevent the (re)introduction or even locally eradicate populations of pest insects and avoids the extensive use of insecticides. Over the years, the Mediterranean fruit fly has become a model organism for SIT research, because it is the only species with successful and well-established genetic sexing strains, combining a phenotypic marker (*wp*), a conditional lethal mutation (*tsl*), a recombination-reducing chromosomal inversion (D53), and a chromosome translocation to achieve stable sex specificity of the traits (Franz et al., 2021). However, the causal genes (*wp*, *tsl*) as well as the exact position of the chromosomal breakpoints could never be identified. Therefore, the targeted transfer to other species was not possible, and the application limited to medfly SIT programs. Moreover, even in medfly, the cost efficacy could be further improved via the conditional conversion of karyotypic females to males. Several transgenic solutions have been developed to enable and enhance sexing for multiple species (Lutrat et al., 2019). However, legal restrictions and lack of public acceptance often prevent their use. The site-specific editing tool CRISPR/Cas offers new possibilities for functional gene characterization and gene or genome editing. It could help to identify causal genes and transfer mutations to homologous genes in different species. Furthermore, certain products of techniques like CRISPR are not considered ‘genetically modified’ in several countries (Schmidt et al., 2020b), which could also help to facilitate field use of engineered insects. Therefore, this thesis aimed to establish the CRISPR/Cas9 HDR pathway in medfly and use minimal invasive CRISPR tools to investigate and improve SIT without the use of transgenes.

4.1 First CRISPR HDR gene editing in Tephritids

First, CRISPR/Cas9 HDR was established in medfly using an eGFP-to-BFP conversion approach (3.1). Previously, only the less demanding and imprecise NHEJ-mediated repair was achieved in medfly and other Tephritids (Meccariello et al., 2017; Li and Handler, 2019; Sim et al., 2019; Meccariello et al., 2020). By HDR, we introduced specific mutations, aiming to engineer distinct phenotypes like temperature-sensitivity or to repair undesired alterations. Adding CRISPR HDR to the ‘toolbox’ of medfly genome editing was an essential step towards genetic improvements relevant to future non-transgenic SIT programs. The novel assay which combines a single-strand repair template, a single guideRNA and Cas9 as a protein proved successful for the precise gene editing in the Tephritid family and was the basis for the subsequent CRISPR projects. The repair template and gRNAs were previously evaluated in eGFP-expressing human cell lines (Glaser et al., 2016), allowing to focus on adjusting parameters for medfly, rather than on designing gRNAs and assessing their efficiency and specificity. As transgenic strains designed for pest control typically carry fluorescent marker genes, this assay also could be useful to establish CRISPR/Cas HDR protocols in other pest species.

High HDR efficiency was achieved in the G₀ and G₁ generations with six out of seven G₀ families producing positive offspring, up to 90% HDR in all medfly G₁ offspring, and up to 96% within the phenotypically mutant offspring (Aumann et al., 2018). For comparison, the originally published GFP-to-BFP conversion assay, using a gRNA/Cas9 expressing plasmid, achieved HDR efficiencies of only 23.3% in K562-50 cells, and 5.8% in HEK293T-EGFP cells (Glaser et al., 2016). Noteworthy, we were able to reproduce our high HDR efficiency when introducing the *ts2* mutation in *Cctra2* (Aumann et al., 2020). So far, only one more study has conducted ssODN-mediated HDR in another Tephritid (*B. tryoni*). Here, the injection of RNP complexes and a 151 nt single-stranded donor template containing a one base substitution (*sh^tts1*) produced one G₀ family with HDR-positive offspring (out of seven fertile families, 14%) and 15% of its G₁ offspring carried the desired point mutation (Choo et al., 2020). A main difference compared to our studies in medfly using 200 ng/μl *in vitro* transcribed single guide RNA (Aumann et al., 2018; Aumann et al., 2020) was the use of a crRNA-tracrRNA complex (120 ng/μl crRNA + 220 ng/μl tracrRNA), which mimics the bacterial CRISPR system (Choo et al., 2020). Interestingly, the use of a single gRNA did not yield any *sh^tts1* HDR positive *B. tryoni* mutants in multiple injections using different concentration of Cas9 protein, sgRNA and ssODN (Choo et al., 2020).

The eGFP-to-BFP conversion assay further revealed frequent partial knock-in events. Their occurrence depended on the combination of gRNA and repair template. When using the ssODN to introduce a PAM site modification, no partial knock-ins were observed. Therefore, the partial knock-ins were probably due to re-editing events (Okamoto et al., 2019). Interestingly, no partial knock-in events were reported when the assay was done in cell culture (Glaser et al., 2016), but it might have remained unnoticed, as only a representative batch of five individual clones per phenotype was sequenced, while we analyzed all phenotypic mutant offspring. Incomplete HDR events were also observed when SNPs with different distance to the cutting site were introduced in porcine fetal fibroblasts (Wang et al., 2016). Notably, the ratio of partial and full knock-ins in the G₁ generation in medfly seemed to be reversely dependent on the age of the injected G₀ at the timepoint of egg laying: while the percentage of partial knock-ins in the offspring dropped with advanced parental age, the desired full HDR events increased. A similar trend (more HDR events in late G₁ collections) was also noticed in *C. elegans* (Paix et al., 2014). However, neither their nor our observations were statistically significant. Further investigation of this phenomenon could be worthwhile, because it would save labor and resources in mutagenic screens.

Knock-in of larger cargo sequences via CRISPR-HDR

The next important step in completing the medfly editing-toolbox would be to test and optimize the use of longer repair templates for HDR, as this might also be of interest to SIT related research, e.g., for creating male linkage of (marker) genes to enable sexing, or for inserting transgenes at defined genomic positions. It has been shown that the knock-in of a long cargo sequence is harder to achieve than ssODN-mediated repair and choosing the ‘correct’ length of the homology arms

in relation to the cargo size can be critical (Paix et al., 2017). Li and Handler achieved a germline transmission rate of 7.3% in *D. sukukii*, using Cas9 protein, *in vitro* transcribed gRNA and a plasmid repair template to introduce ~2 kb DNA, flanked by roughly 1 kb homology arms (Li and Handler, 2017). Gilles et al. reported 0-10% germline transmission and 2-34% positive G₁ offspring in *Tribolium castaneum*, using Cas9 and gRNA expressing plasmids, and a plasmid repair template with 1.2 kb cargo and 0.7-1 kb homology arms (Gilles et al., 2015). For the medfly, it has to be evaluated which ratio of homology arm to cargo sequence length produces optimal editing results, and whether using a plasmid, a linear single-, or double-stranded repair template is optimal for high transformation efficiency.

To develop high efficiency protocols for large cargo CRISPR HDR, strategies enabling the quick and direct readout of HDR and undesired NHEJ events would be highly beneficial. This could be achieved by integrating an exogenous marker gene fused to a promoter (e.g., P_{U6}-eGFP) into an endogenous gene producing a visible knock-out phenotype (e.g., *white*, *white pupae*, *yellow*). Thereby, NHEJ events (morphological marker) could be distinguished from HDR events (morphological and fluorescent markers) without the need of sequence analysis. Another option is to target and exchange fluorescent markers in transgenic strains. If a long repair template is used to introduce several distant SNPs, silent mutations should be included in the repair template sequence between the intended edits and the DSB to avoid creating partial knock-ins due to premature removal of the donor, and the PAM site should be recoded to prevent re-editing (Paix et al., 2017).

4.2 Improving SIT through CRISPR gene editing

The highly efficient CRISPR/Cas9 HDR protocol developed in the first part of this thesis was used to target two genes with the potential to improve SIT programs for medfly and other species: *transformer-2* and *white pupae*. In both projects, I omitted using extraneous modifications like marker genes to meet the criteria necessary for field use without GMO restrictions in several countries (Schmidt et al., 2020b). This strategy was only possible due to our HDR protocols' high mutation frequency, as identifying mutagenesis events without a phenotypic marker can otherwise be highly laborious and time-consuming.

White pupae – a key-element to build new genetic sexing strains

The discovery and verification of the molecular basis of 'the' medfly *white pupae* gene (3.3) was a long-awaited achievement in SIT-related research and holds great promise for the construction of GSS in other related and unrelated species, as shown for *B. tryoni* (Ward et al., 2021). Medfly GSS were solely built on *wp* until the discovery of *tsl* enabled developing the 'second generation' GSS, based on *wp* and *tsl*. Now, sexing could also be achieved based on the *tsl* phenotype alone – however, the integration of a genetically-linked visible marker provided a beneficial tracing option of the *tsl* mutation and allowed implementing a filter rearing system to eliminate any recombinant

insects (Fisher and Caceres, 1998; Morrison et al., 2010; Franz et al., 2021). Therefore, *wp* remains an essential part of the medfly GSS and is sufficient to build new GSS in other species.

The NCBI database describes the gene identified as causative for the white puparium phenotype in our medfly project as putative metabolite transport protein HI_1104, also known as *Spns1* (LOC101451947, XP_020715742.1; Protein Spinster Homolog 1; putative sphingolipid transporter). It contains a conserved protein domain of the Major Facilitator Superfamily (pfam07690; *MFS_1*; <https://www.ncbi.nlm.nih.gov/gene/101451947>), which is one of the two largest families of membrane transporters (Pao et al., 1998). A homologous gene can, e.g., be found on the *D. melanogaster* X-chromosome (CG14439, 80.26% identity to XP_020715742.1), with a peak expression in the white prepupal stage (modENCODE Development RNA-Seq). Interestingly, no pupal phenotype is described for CG14439 mutants available at the Bloomington Drosophila Stock Center (<https://bdsc.indiana.edu>). In medfly, *Spns1* (*wp*) is probably involved in transporting hemolymph catecholamines to the cuticle, because the mutant *wp* puparium comprises low levels of the catecholamines *N*- β -alanyldopamine (NBAD), *N*-acetyldopamine (NADA) and dopamine (DA) (Wappner et al., 1995). However, its hemolymph contains about ten times more NBAD, NADA, and DA than the WT hemolymph (Wappner et al., 1995). Lack of these substrates in the puparium cuticle prevents sclerotization and pigmentation, explaining the puparium color and the distinctive mechanical properties of different pupal color strains (Bourtzis et al., 1991; Wappner et al., 1995; Wappner et al., 1996).

During the project, four homozygous *wp*^(CRISPR) medfly strains based on different *wp* mutation events were established. So far, all strains behave similarly to the EgII WT strain in terms of fitness and fecundity. Quality control tests are currently performed at the Insect Pest Control Laboratory of the Joint FAO/IAEA Centre of Nuclear Techniques in Food and Agriculture in Seibersdorf to obtain statistically significant results (personal communication with Dr. Kostas Bourtzis, FAO/IAEA, September 2020). Interestingly, the deletion of a single amino acid in the *MFS_1* domain in the *wp* gene led to the same phenotype as the natural mutation - a roughly 8.150 bp large insertion of a transposable element. Analysis of this natural insertion indicated that it might be transposon-derived because identical repeats are flanking it. As expected from a putative transposon, the inserted sequence can be found multiple times in the medfly genome. Because the unambiguous assembly of short sequencing reads to such repetitive elements in the genome is complicated, identifying the natural *wp* mutation was only possible based on the combination of multiple short and long-read sequencing data paired with expression data. RNAseq and Illumina genome sequencing first showed an apparent inconsistency between WT and *wp* strains at this locus. A combination of Nanopore sequencing data spanning the region and supporting PCR results then confirmed the presence of the transposon-based insert. The phylogeny and lineage of the putative transposon have not been resolved yet, and it remains unknown if it is endogenous or was gained via horizontal gene transfer (Peccoud et al., 2017). It has been shown that the medfly genome contains several active endogenous transposable elements (Zhou and Haymer, 1997;

Gomulski et al., 2004; Papanicolaou et al., 2016). The transposable element disrupting the *white pupae* gene, however, seems to be stably integrated, as the phenotype was never lost during years of large-scale mass rearing.

Another important achievement based on our findings is the newly developed genetic tracking tool to safeguard the field application of the established medfly GSS VIENNA 7 and 8. Revealing the breakpoints of the D53 inversion and the genetic basis of the white pupae phenotype on a sequence level enabled me to develop two PCR assays to distinguish between GSS and WT flies on a genomic level. Until now, no such tool has been published, and it has not been possible to genetically identify released GSS flies for monitoring the frequency of potentially surviving progeny (hazard characterization). The first assay is based on the mutation of the *wp* gene. Three primers are used for the *wp*-assay, producing amplicons of different sizes for the WT and the mutant allele. This leads to either one amplicon of a distinctive size for WT or homozygous mutants, or two amplicons in case of heterozygosity. Thus, each PCR reaction has to produce a positive signal and decisions are not based on negative 'no signal' PCR results. The second assay uses the presence or absence of the D53 inversion. Two primer pairs, flanking the right and the left breakpoint of the inversion, are necessary for the inversion-based assay. WT or inversion status of the chromosome 5 (homo- or heterozygous) can be assessed using the respective primer combinations (see 6.2.4 for detailed protocols). The tracking system's importance has recently been demonstrated in Australia when researchers were asked to prove that a medfly outbreak in an area, where sterile releases of a GSS strain were previously conducted, was not due to insufficiently sterilized GSS males. Our *wp*-based assay was used to test outbreak and factory reared flies and prove their genetic lineage. The assay will now be regularly applied to analyze outbreaks (personal communication with Dr. Amanda Choo, University of Adelaide, and Dr. Peter Crisp, South Australian Research and Development Institute (SARDI). January - June 2020).

Furthermore, we released a novel, high-resolution, chromosome-scale genome assembly for medfly (accession number GCA_905071925) and several genomic and transcriptomic sequencing data sets of different WT and mutant strains (with and without D53 inversion, *wp*⁻, *tsl*⁻) as part of this study (Ward et al., 2021). These data will help to gain a better understanding of the molecular basis of medfly GSS, investigate target genes relevant to genetic control, and be highly relevant to a broad research community studying medfly and related Tephritids' biology and evolution.

tra2^{ts} – Learning from model species and re-building mutations via CRISPR HDR

Another strategy to improve SIT programs is to double the number of males via the conditional conversion of female embryos to males. Such a sex conversion can be achieved in Tephritids by targeting genes involved in sexual development such as *transformer* or *transformer-2* via RNAi (Pane et al., 2002; Saccone et al., 2007; Salvemini et al., 2009; Schetelig et al., 2012). However, the RNAi-based knock-down of genes lacks the conditionality that is necessary to rear the strains for an SIT application. Learning from the model organism *D. melanogaster*, temperature-

sensitivity, and therefore conditionality of certain phenotypes, can be achieved by introducing specific DNA mutations. Temperature-sensitive variants of *D. melanogaster transformer-2 (tra2)* were engineered in the *D. sukukii* and in *C. capitata* homologs (Li and Handler, 2017; Aumann et al., 2020; see. 3.2). In medfly, I could establish a homozygous *tra2^{ts2}* population within three generations without using a marker gene. This success was based on two factors; i) the high HDR editing efficiency – detectable by the absence of phenotypic females and the occurrence of intersexes in the G₀ generation, the frequency of HDR-positive, fertile G₀, and the high penetrance of the mutant genotype within their G₁ offspring, and ii) the development of a non-lethal genotyping assay based on DNA extraction from a single leg, which has not been published for medfly before (see 6.2.3 for detailed protocol). Thus, we reached the ‘ultimate goal of genetic engineering’, consisting of precisely rewriting the genome without adding unnecessary alterations (Paix et al., 2014). Interestingly, the phenotype of the *tra2^{ts}* neither in *D. sukukii* nor in medfly did match the one known from *D. melanogaster*. While *D. sukukii* was not able to cope with the high restrictive temperatures (> 26°C; Li and Handler, 2017), medfly was not able to breed at the supposedly low permissive temperatures (< 18.5°C; Aumann et al., 2020). Nevertheless, in both species, the conversion effect was demonstrated, which supports the general idea of transferring known mutations from model organisms to pest insects via CRISPR/Cas9 HDR.

Other examples for *D. melanogaster* genes with mutations causing putatively interesting phenotypes for the development of SIT strains (such as temperature-sensitive paralysis or lethality) are *paralysed*, *cacophony*, *no action potential* (Tarasoff and Suzuki, 1970; Suzuki et al., 1971; Wu and Ganetzky, 1980; Kawasaki et al., 2000), or *shibire* (Grigliatti et al., 1973; Grant et al., 1998). Several dominantly and recessively acting variants are, for example, known in *D. melanogaster shibire*, most causing lethality or temperature-sensitive paralysis. An initial screening of this candidate in medfly showed homology between *Dmel_shibire* and *Ccap_LOC101463232* (dynamin), indicating that the known mutations could be re-built in medfly and potentially exploited as a sexing system for SIT programs. *In situ* hybridization analysis of LOC101463232 also verified its location on the tip of the right arm of chromosome 5 (80A, salivary gland chromosomal map; Georgia Gouvi, FAO/IAEA; R.A. Aumann; Marc F. Schetelig, JLU Gießen; Kostas Bourtzis, FAO/IAEA; unpublished), which would be beneficial for a translocation approach in combination with *wp*.

However, the identification and utilization of genes and mutations known and proven successful in existing GSS seem to be most promising for implementing such phenotypes in other pest species. Genes of interest include *bent wings* (temperature-sensitive lethality and crippled wings in *B. tryoni*; Meats et al., 2002), *slow larvae* (slow larval development and light pigmentation in pupal and adult stage in *A. ludens*; Meza et al., 2019), *slow development* (slow larval development in *C. capitata*; Porras et al., 2020), *sergeant-2* (third abdominal stripe in *C. capitata*; Niyazi et al., 2005; Rempoulakis et al., 2016) and, most notably, *tsl*, causing temperature-sensitive lethality in the SIT model species *C. capitata*. The *tsl* system in medfly is currently in

use in VIENNA 7 and VIENNA 8 GSS (Franz et al., 1994; Franz et al., 2021). Thanks to the recent identification of *wp* in medfly, the chromosome-scale genome assembly (Ward et al., 2021), and previous cytogenetic analyses indicating that *tsl* is located in the interval 59B-61C (Kerremans and Franz, 1994) or 60B-61B (Niyazi et al., 2005; Franz et al., 2021) on the trichogen polytene chromosome map, the candidate region for *tsl* could be narrowed down to about 6-9 Mbp. It harbors about 600 genes, and the identification of *tsl* seems now feasible. Multiple criteria can be established to identify the most promising candidates among the genes in the target region. Based on the experience in *Drosophila* and several microorganisms, temperature-sensitive (*ts*) mutations are typically missense mutations, i.e., variants changing an amino acid sequence, which causes a change in the proteins' 3-dimensional structure at restrictive temperatures (Suzuki, 1970; Ben-Aroya et al., 2010). Therefore, one criterion to select *tsl* candidate genes can be the presence of a point mutation with an impact on the protein level. Our transcriptomic data of WT and mutant heat-treated and untreated embryos could furthermore be used to identify differential expression among the candidate genes, which could be another selection criterion (personal communication with Kostas Bourtzis, FAO/IAEA; Alistair Darby, Centre for Genomic Research, Liverpool; Marc F. Schetelig, JLU Gießen). However, it is unknown if *tsl* is differentially expressed as a consequence of the mutation or a heat treatment. Furthermore, genes with a matching phenotype in *D. melanogaster* could be chosen as candidate genes, especially if the homologous genes are located on the *D. melanogaster* X-chromosome, which shows conserved synteny to the medfly chromosome 5 (Papanicolaou et al., 2016; Zacharopoulou et al., 2017). It should also be taken into account that it is unknown whether the *tsl* phenotype is caused by one mutation (affecting a single gene), or if multiple factors cause it. Once candidate genes have been selected, CRISPR/Cas9 gene editing can be used to either rescue the WT phenotype in the *tsl* strain (HDR approach) or knock-out the WT allele (NHEJ or HDR), to check if the *tsl* phenotype can be triggered. Knock-out could be achieved either via the NHEJ pathway or via the insertion of an exogenous marker gene within the candidate's coding sequence to simplify the screening process. However, it should be considered that a homozygous knock-out of the gene might not produce the same conditional phenotype as the original mutation and might be lethal at any condition. Thus, 'hemizygous' individuals, created via backcross to the original *tsl* strain, should instead be used for temperature-sensitivity screens. Once the *tsl*, or another conditional lethal phenotype is identified in the medfly, this trait could enable the rapid development of new or advanced GSS in other species by editing the homologous genes as it was shown for *wp* (Ward et al., 2021).

Engineering sexing strains via CRISPR genome editing

CRISPR/Cas9-modified strains like *wp*^(CRISPR) could be used to establish a new generation of GSS with improved fertility. A CRISPR/Cas9-mediated transfer of *wp* and *tsl*, once its genetic basis is resolved, close to the maleness-determining factor (*MoY*; Meccariello et al., 2019) could engineer male-linkage without a translocation and the associated semi-sterility. However, it needs to be

evaluated if the WT allele transfer to the Y-chromosome is sufficient to rescue the WT phenotype in the presence of mutated copies on each chromosome 5, or if three copies of the genes (on both autosomes and on the Y-chromosome) are potentially harmful to the insect. Furthermore, the Y-chromosome transfer would require the insertion of a large sequence, which might be highly challenging due to the Y-chromosome's heterochromatic, repeat-rich nature (Buchman and Akbari, 2019). It should also be noted that even the (untraceable) HDR-mediated transfer of endogenous sequences, which basically mimics classical genetics and mutagenesis using irradiation, would imply GMO regulations in certain countries like Australia (Thygesen, 2019). Here, NHEJ-mediated mutations (SDN-1) are principally considered non-GMO, whereas HDR edited organisms (SDN-2 and SDN-3) are considered and regulated as GMO (Grohmann et al., 2019; Mallapaty, 2019).

Sex-linkage could also be achieved via a deliberate, CRISPR-induced inter-chromosomal rearrangement ('genome engineering'). The introduction of several systematic DSBs in the genome would allow to modify chromosomes directly and to create novel combinations relying solely on the NHEJ pathway, thus preventing the need for applying GMO regulations (Jiang et al., 2016; Brunet and Jasin, 2018; Schmidt et al., 2020a). Like current medfly GSS, these strains would be semi-sterile due to the translocation. Nevertheless, the targeted approach offers many advantages over random induction through irradiation. Irradiation-induced random translocations require tedious genetic and cytogenetic analyses. To assess which translocation might be suitable for GSS, the involved autosomes, the position of the translocation breakpoint(s), the position of the selectable markers, and the translocation's genetic consequences (e.g., sterility) must be known (Franz et al., 2021). When engineering a translocation by CRISPR, it is only necessary to know the target sites on an autosome and the Y-chromosome to design specific gRNAs for the induction of DSBs. The recent delineation of the autosomal T(Y;5) breakpoint in medfly VIENNA 8 on a genomic level will help assess such suitable positions in medfly (Jiannis Ragoussis, McGill University; Marc F. Schetelig, JLU Gießen; Kostas Bourtzis FAO/IAEA, unpublished). Stabilizing intra-chromosomal inversions like the D53 inversion could also be engineered by CRISPR (Schmidt et al., 2020a). Naturally occurring inversions and translocations have already been re-built in human cell lines with rearrangement rates of up to approximately 8% (Choi and Meyerson, 2014; Torres et al., 2014), and heritable chromosomal translocations in the Mbp range have also been engineered in the plant model organism *Arabidopsis thaliana* with frequencies up to 2.5% (Beying et al., 2020). The efficiency in insects needs to be evaluated. However, even lower efficiency would probably still be acceptable, considering that irradiation-induced translocations produced about 7% F₁ crosses with male-linked translocations in medfly, of which most are not suited to construct a GSS (Franz et al., 2021). Such difficulties could be increased in other insects, where the translocation frequencies are negatively affected by the size of the genome and chromosomes (Franz et al., 2021).

Das et al. (2020) suggested that CRISPR could be used to engineer subtractive transgene sex sorting (STSS) systems in pest insects. For STSS, two transgenic strains with a conditional lethal construct on the Y- or X-chromosome, respectively, need to be created (Das et al., 2020). The advantage of this system, compared to other TSS, is that males produced for release are transgene-free and would therefore not trigger GMO regulations in several countries (Schmidt et al., 2020b). As described above, linkage to the Y-chromosome – although probably highly challenging to achieve – would avoid translocation-associated semi-sterility known from current medfly GSS. However, STSS systems first need to be evaluated under mass-rearing conditions to test the feasibility of integrating them in an SIT program - in terms of economics, as the system requires the maintenance of two separate transgenic strains on (tetracycline) supplemented diet in the rearing facilities, and in terms of genetic stability, as mutations accumulated during mass-rearing could cause a genetic breakdown of the system (Das et al., 2020; Zhao et al., 2020). Furthermore, the public acceptance of such an approach needs to be assessed.

Gene drives – CRISPR-engineered pest control systems beyond SIT

Another option to improve pest control programs using CRISPR gene editing are gene drives (modification or suppression drive systems) (Gantz and Bier, 2015). The gene drive's basic idea is to 'overwrite' Mendelian inheritance by converting heterozygous mutations to homozygosity, a strategy referred to as 'active genetics' (Gantz and Bier, 2016). To achieve gene drive by CRISPR/Cas HDR, a gene drive cassette has to be inserted into the genome, consisting of the Cas9 gene and a gRNA sequence, surrounded by homology arms matching the sequence adjacent to the targeted cutting site (Gantz and Bier, 2015). Expression of this cassette in somatic cells and the germline will lead to a so-called mutagenic chain reaction: Cas9 and the gRNA will induce site-specific cleavage at the target site, followed by the insertion of the Cas9-gRNA cassette into that locus via the HDR pathway due to the provided homology arms. This is continued until no unedited allele remains (Gantz and Bier, 2015; Gantz et al., 2015; Gantz and Bier, 2016). The desired trait will then be spread throughout the population – until an NHEJ event destroys the target site, and the target locus can no longer be recognized by the gRNA, leading to resistance against the gene drive (Unckless et al., 2017; KaramiNejadRanjbar et al., 2018). Several types of gene drive systems (e.g., Homing-based drive, X-shredder, *Medea*, Toxin-antidote) and possible applications exist in insects, namely the removal of invasive species, the reversal of resistance alleles, or the spread of beneficial traits in endangered species. The main research focus, however, currently is on reducing vector-borne diseases, by either suppressing the vector insect or reducing transmission by making the insect 'resistant' to the pathogen (Champer et al., 2016). Gene drive systems have been engineered and tested in laboratories; however, several ethical and technical questions, e.g., about the removability, reversibility, and quickly evolving resistances, as well as the legal regulation of such a system, need to be discussed and answered (Marshall, 2010; Marshall et al., 2017; Jones et al., 2019; Zhao et al., 2020).

4.3 Field application and regulatory status of CRISPR-edited organisms

Before CRISPR-engineered strains could be used in a field release, they need to be thoroughly evaluated in terms of genetic stability and environmental impacts (benefits and risks) to fulfill current legal regulations and gain public acceptance. Several countries started to adapt their GM regulations to new breeding techniques (NBT) like CRISPR/Cas gene editing. However, the regulations differ greatly. The first country to apply regulations specifically for NBT was Argentina (Whelan and Lema, 2015). Here, product- and process-based criteria are used to determine the regulation status. Products that are free of transgenes after modifications are not regulated as GMO, even if a transgene was temporarily used during the development process (Whelan and Lema, 2015). This regulation applies now also in Brazil, Chile, Colombia, Israel and Japan (Schmidt et al., 2020b). Bangladesh, India, Kenya, Nigeria, Norway, Paraguay, and Uruguay are likely to follow soon (Schmidt et al., 2020b). In Australia, GM regulations apply if a repair template was used (Schmidt et al., 2020b). In the US, it depends on the organism, its use, and the responsible federal agency if gene editing without transgenesis is subject to additional regulations (Van Eenennaam et al., 2019). The European Union and New Zealand have the most stringent regulations. Here, any mutagenesis technique that did not exist before 2001, or the use of *in vitro* techniques to modify genetic material respectively, triggers GMO regulations (Schmidt et al., 2020b). Efforts are underway to develop specific guidance on GM insects' release (Brown et al., 2018). Until widely accepted rules are available, the regulatory status of CRISPR gene-edited insects will remain uncertain in several countries (Mumford, 2012). Yet, an *issue-specific global* regulation would be of great importance (Schwindenhammer, 2020). Issue-specific, because regulations must be tailored to each kind of transgenic construct or editing procedure, the species and its application, to take into account the manifold possibilities of genetic engineering and potential impacts upon release (Reeves et al., 2012). Global, because (GM) insects will cross national boundaries naturally and human-assisted through trade and travel (Häcker and Schetelig, 2021). One of the most important global regulatory frameworks for GM insect technology is the Cartagena Protocol on Biosafety (173 Parties) to the Convention on Biological Diversity (CBD, 196 Parties). It provides guidelines for the safe transfer, handling and use of living modified organisms resulting from modern biotechnology and provides an outline for a risk assessment methodology for activities (Annex III) (CBD, 2000). However, it should be noted, that some countries with substantial involvement in GM technology, e.g., the US and Australia, are not a party to the Cartagena Protocol.

EU regulation and its impact on GMO handling in Europe

In the European Union, the Directive 2001/18/EC regulates the deliberate release, labeling, traceability, and market launch of genetically modified organisms (GMOs). New breeding techniques, such as CRISPR, are regulated by this Directive and are treated like 'classic' transgenic

GMOs. In contrast, ‘first-generation technologies’ breeding techniques, based on the use of ionizing radiation or DNA-damaging chemicals (Urnov et al., 2018) to produce high frequencies of new mutations, are exempted by the so-called mutagenesis-exemption of the Directive 2001/18/EC, because, at the time the Directive was published, these techniques have been used for several years, showed a long safety record, and did not involve recombinant DNA. Therefore, products generated by first-generation technologies can be released without passing through the GMO risk assessment process and do not need to be labeled as GMO – even though first-generation technologies induce multiple random mutations that might be beneficial, deleterious, or lethal, and are scattered throughout the genome. Breeding-based selection can be used to isolate mutations that confer desirable traits, yet it is unknown how many additional mutations, known as ‘unintended effects’, are carried along with them. Mutations induced by CRISPR are targeted and highly specific, and unintended effects occur less frequently. Depending on the mutation type, CRISPR-induced alterations may even be indistinguishable from naturally occurring mutations or mutations resulting from random mutagenesis (e.g., SDN-1 and SDN-2). Thus, two identical products, one edited via classical mutagenesis, the other via genome editing, are differently governed by law. Consequently, it will be impossible to prove how the mutation was introduced, posing a significant problem from a regulatory perspective (Sprink et al., 2016; Grohmann et al., 2019). For this reason, the Council of the European Union has asked the European Commission to re-evaluate the EU policy framework (Rechtssache C-528/16) and to propose options to update it until the end of April 2021 (Rat der Europäischen Union, 2019). For the time being, the deliberate release of CRISPR-edited plants, animals, and insects will be subject to an environmental risk assessment according to the Directive 2001/18/EC, and all applications need to be evaluated case-by-case. Reasons for this decision seem to be concerns about the speed with which new varieties can be produced, as targeted mutagenesis is much more efficient than random mutagenesis (Group of Chief Scientific Advisors, 2018), and the limited existing knowledge about potential effects of the edited organism (Eckerstorfer et al., 2019).

Scientists criticize the current regulation as ‘being made without reference to scientific evidence’ (Group of Chief Scientific Advisors, 2018; Metje-Sprink et al., 2018; Urnov et al., 2018). Several renowned German scientific societies criticized the current legislation and asked for a different evaluation of NBT. Based on the facts that the speed with which new strains or varieties can be developed is not automatically linked to additional risks, that there is no scientific evidence for novel risks associated with gene editing, and that mutations induced by a new breeding technique are often indistinguishable from mutations resulting from first-generation technologies, they argue that new products should be evaluated based on the potential risks associated with the product – not the process. Legislation should take the kind of alteration (SDN-1/-2/-3) into account, and whether the mutation could as well occur naturally or result from conventional breeding techniques (Zentrale Kommission für die biologische Sicherheit, 2018;

Bioökonomierat, 2019; Max-Planck-Gesellschaft, 2019; Nationale Akademie der Wissenschaften Leopoldina et al., 2019; Andersen and Schreiber, 2020).

A recent publication evokes some hope concerning the potential use of SDN-1 and SDN-2 in Europe. The authors conclude that ‘impacts of the Judgment on the NBT might have been slightly overvalued’ and that the EU might consider the exemption of specific NBT applications as SDN-1/-2, once such techniques have been ‘used in a number of applications’ and ‘have a long safety record’ - if the clause ‘does not involve the use of recombinant nucleic acid molecules’ is not interpreted narrowly (Vives-Vallés and Collonnier, 2020). On the other hand, the risk assessment, cultivation and labelling of a GMO obtained via traditional mutagenesis can be regulated by EU Member States nationally (Vives-Vallés and Collonnier, 2020). Consequently, the Conseil d’État, France’s administrative court recently used this option to implement a stricter interpretation of the EU directive and ruled in February 2020 that all organisms derived via *in vitro* mutagenesis should be treated as GMO, including those that are subject to the EU mutagenesis exemption (Conseil d’État, 2020). As this ruling would negatively affect the free movement of goods within the EU, the EU commission requested a scientific statement on risk-assessment of *in vivo* and *in vitro* mutagenesis from EFSA (Mandat Nr. M-2020-016, EFSA Q-2020-00445).

It remains to be seen which decision the EU will arrive at in a re-evaluation and how the EU will re-harmonize legislation to avoid negative consequences and ensure a technically enforceable law. Currently, only the EU and New Zealand regulate all genome-edited organisms as GMOs, while many other countries already have adapted regulatory frameworks (Schmidt et al., 2020b).

Risk assessment for GM animals in Europe

A working group convened by the European Food Safety Authority (EFSA, 2010-2013) (Benedict et al., 2010) issued a so-called guidance document describing the six steps for the environmental risk assessment (ERA) of genetically modified animals according to the Directive 2001/18/EC: (1) problem formulation including hazard and exposure identification; (2) hazard characterization; (3) exposure characterization; (4) risk characterization; (5) risk management strategies; and (6) overall risk evaluation (EFSA, 2013; Mestdagh et al., 2014). Each of these six steps have to be applied to seven risk areas. For GM insects, these are (I) the persistence and invasiveness, including vertical gene transfer; (II) horizontal gene transfer; (III) interactions with target organisms and (IV) non-target organisms; (V) the environmental impacts of the specific techniques used for the management of the GM insect, (VI) impacts of the GM animal on biogeochemical processes and (VII) on human and animal health. Furthermore, the choice of comparators, the use of non-GM surrogates, long-term effects and uncertainty analyses need to be considered during the risk assessment process (EFSA, 2013). As this is highly expensive and time-consuming, mostly large companies can afford such processes, which triggers monopolies as well as a focus on economically important species and traits and prevents small- or medium-sized

enterprises or public research facilities from investing in research and product development (Nationale Akademie der Wissenschaften Leopoldina et al., 2019; Zaidi et al., 2019).

The first environmental risk assessment of a genetically modified insect in Europe was requested for a netted field release of the transgenic, self-limiting olive fruit fly strain OX3097D-Bol in Spain in 2013 by Oxitec Limited, the most important business entrepreneur in the field (Turner et al., 2018; Schwindenhammer, 2020). This would have been the first outdoor trial of a GM insect in the EU (Turner et al., 2018). OX3097D-Bol flies carry a fluorescent marker (DsRed2) and a conditionally expressed female-specific self-limiting trait (Tet-OFF), which is transmitted to their offspring (Turner et al., 2018). However, the application was withdrawn by Oxitec roughly three years after the first ERA had been handed in as it became apparent that the feasibility of anything other than a trial release was unlikely due to regulations and public rejection (Turner et al., 2018). In contrast, Oxitec's transgenic mosquitos have been released in the Cayman Islands (2009), Malaysia (2010) and Brazil (2011) (Antonelli et al., 2015), and open field trials of transgenic diamondback moth (OX4319L) and medfly (OX3864A) were permitted in New York State (Waltz, 2017) and Morocco (Panjwani and Wilson, 2016), respectively. All involved strains are based on the same principle as the self-limiting olive fruit fly. Overall, the imponderability of the application procedures and the possibility of public rejection in case of a successful application diminish the prospects for the use of transgenic insects in Europe.

Public acceptance as a decisive factor

The examples above highlight the worldwide differences in the GMO risk assessment processes. Scientists, the public, and stakeholders criticize these processes for lack of transparency, scientific quality of mostly generic regulatory processes, and incomprehensible ruling (Panjwani and Wilson, 2016). Furthermore, the examples highlight public acceptance as a decisive factor (Lehane and Aksoy, 2012). However, the extent to which the public was adequately informed (Enserink, 2010), and whether experts examined potential hazards thoroughly before the releases was often unknown (Reeves et al., 2012). This 'transparency challenge' involves regulators as well as business actors and scientists, and the relevance of the provided information might be considered differently by different stakeholders (Schwindenhammer, 2020; Häcker and Schetelig, 2021). Businesses like Oxitec need to protect their proprietary data, which could hinder a transparent and independent scientific evaluation (Schwindenhammer, 2020). As a solitary entrepreneur in the field, Oxitec claims technical and moral authority and is involved in promoting and distributing GM insects and developing the Cartagena Protocol and the WHO Guidance Framework on GM insects (Schwindenhammer, 2020). Such potential conflicts of interests, insufficient or badly communicated information, and missing transparency of all processes involving evaluation and release of GMOs can cause mistrust and rejection. These shortcomings can give rise to conspiracy theories, for example, that the Zika virus was caused by genetically modified mosquitos (Klofstad et al., 2019), or that the US and the WHO used the SIT as a 'smoke screen' to prepare for biological

warfare in India (Oh, New Delhi; oh, Geneva, 1975). As the current ruling demands that GMO trial locations need to be published in location registries, field trials have been targets of vandalism and attacks by GMO opponents. While trials could be conducted in other regions, the important data with respect to local strains and potential local effects cannot be produced (Stewart et al., 2000; Kuntz, 2012; Nationale Akademie der Wissenschaften Leopoldina et al., 2019). This leads to a vicious circle: more information and more thorough tests are requested, but the field evaluations necessary to generate the data are prevented or destroyed. In any case, large-scale assessments are indispensable for proper risk assessment, as recently shown for the widely established Tet-OFF embryonic lethality system in the model organism *D. melanogaster* (Zhao et al., 2020).

It will be critical for the success of field trials and the release of CRISPR-edited organism to gain public approval prior to any action. Main channels of information that can be used for public engagement are the TV, radio and internet (Kittayapong et al., 2019). However, successfully conducted field trials have also taught us how important door-to-door community campaigns and community education are to gain public support and acceptance (Liew et al., 2021). The methods advantages, as well as its disadvantages and limits, need to be communicated and explained understandably, but still accurately based on scientific data to avoid rejection based on inaccurate knowledge (Byrne et al., 2002; Wunderlich and Gatto, 2015; Yuan et al., 2019). The impact of the final product, e.g., an insect strain for improved pest control, should be put in relation to existing products or methods, such as the use of insecticides (Adalja et al., 2016; Panjwani and Wilson, 2016; Turner et al., 2018). New techniques and first-generation breeding techniques need to be precisely explained with regard to the underlying mutagenesis practices and definitions to prevent e.g. confusion of CRISPR gene editing with gene drives, or ‘transgenic’ with ‘genetically modified’. A recent survey shows that consumers currently would most probably not differentiate between CRISPR-edited and ‘traditional’ transgenic approaches – at least with respect to food (Shew et al., 2018).

4.4 A perspective for CRISPR-edited insects in SIT programs

Classical SIT programs do not draw negative attention as the insects are sterilized via radiation before their release. According to EU Directive 2001/18/EC (HCB Scientific Committee, 2017), such insects are not considered as organisms and are therefore not GMOs. They are ‘living modified organisms’ as defined by the Cartagena Protocol on Biosafety to the Convention on Biological Diversity but exempted from GMO regulations, because radiation is considered a mutagenesis technique with a long safety record.

An approach combining sterilization via radiation with traits introduced via the herewith established (scar-less) CRISPR editing, and sex-linkage via CRISPR-engineered Y/autosome translocation, could eventually be accepted by the public and be considered ‘non-GMO’ even in the EU (Vives-Vallés and Collonnier, 2020). The engineered alteration(s) would be known and

traceable on a DNA level, contrary to currently used traits generated by radiation or chemicals, whose molecular origin is unknown. Furthermore, SDN-1/-2 CRISPR-edited insects would be genetically modified, but not transgenic, because no exogenous DNA is incorporated in the genome. Combined with irradiation to induce sterility, insects will be incapable of reproducing and could not transfer traits to wild insects. This is in stark contrast to exceedingly discussed gene drive systems or the already used RIDL system, for which fertile survivors are expected after field releases (Gong et al., 2005; Harris et al., 2011). Therefore, such CRISPR-edited insects for SIT programs should not require regulation. Several countries (e.g. Argentina, Australia, Brazil, Canada, Chile, Colombia, Israel) already categorize SDN-1 alterations as non-GMO (Schmidt et al., 2020b), which facilitates field trials and the application of edited insects without the need for a GMO risk assessment in these countries.

Overall, CRISPR gene and genome editing show great potential to improve SIT relevant traits and enable a ‘generic approach’ to build or improve genetic sexing strains in several agricultural and medical important pest insect species. The findings in this thesis, namely the establishment of an efficient CRISPR/Cas9 HDR strategy using a short, single-stranded repair template and its application for pest control development, the implementation of a safeguard tool, and the identification of the *white pupae* gene, represent essential steps towards a versatile and safe use of CRISPR technologies for insect pest control.

5 References

- Adalja, A., Sell, T.K., McGinty, M., Boddie, C., 2016. Genetically Modified (GM) Mosquito use to reduce Mosquito-transmitted disease in the US: A community opinion survey. *PLoS currents* **8**.
- Alexandratos, N., 1999. World food and agriculture: Outlook for the medium and longer term. *Proc. Natl. Acad. Sci.* **96**, 5908-5914.
- Allwood, A.J., Leblanc, L., 1997. Losses caused by fruit flies (Diptera: Tephritidae) in seven pacific island countries, in: Allwood, A.J., Drew, R.A.I. (Eds.), *Management of Fruit Flies in the Pacific. A Regional Symposium*. ACIAR Proceedings Nadi, Fiji pp. 208-211.
- Alphey, L., 2002. Re-engineering the sterile insect technique. *Insect Biochem. Mol. Biol.* **32**, 1243-1247.
- Alphey, L., Beard, C.B., Billingsley, P., Coetzee, M., Crisanti, A., Curtis, C., Eggleston, P., Godfray, C., Hemingway, J., Jacobs-Lorena, M., James, A.A., Kafatos, F.C., Mukwaya, L.G., Paton, M., Powell, J.R., Schneider, W., Scott, T.W., Sina, B., Sinden, R., Sinkins, S., Spielman, A., Toure, Y., Collins, F.H., 2002. Malaria control with genetically manipulated insect vectors. *Science* **298**, 119-121.
- Alphey, L.S., 2007. Engineering insects for the Sterile Insect Technique, in: Vreysen, M.J.B., Robinson, A.S., Hendrichs, J. (Eds.), *Area-Wide Control of Insect Pests*. Springer, Dordrecht, pp. 51-60.
- Andersen, E., Schreiber, K., 2020. "Genome Editing" vor dem EuGH und seine Folgen. *Natur und Recht* **42**, 99-106.
- Ant, T., Koukidou, M., Rempoulakis, P., Gong, H.-F., Economopoulos, A., Vontas, J., Alphey, L., 2012. Control of the olive fruit fly using genetics-enhanced sterile insect technique. *BMC Biol.* **10**, 51.
- Antonelli, T., Clayton, A., Hartzog, M., Webster, S., Zilnik, G., 2015. Transgenic pests and human health: a short overview of social, cultural and scientific considerations, in: Adelman, Z.N. (Ed.), *Genetic Control of Malaria and Dengue*. Academic Press (Elsevier), pp. 1-27.
- Aryan, A., Anderson, M.A.E., Biedler, J.K., Qi, Y., Overcash, J.M., Naumenko, A.N., Sharakhova, M.V., Mao, C., Adelman, Z.N., Tu, Z., 2020. *Nix* alone is sufficient to convert female *Aedes aegypti* into fertile males and *myo-sex* is needed for male flight. *Proc. Natl. Acad. Sci.*, 202001132.
- Augustinos, A.A., Targovska, A., Cancio-Martinez, E., Schorn, E., Franz, G., Cáceres, C., Zacharopoulou, A., Bourtzis, K., 2017. *Ceratitidis capitata* genetic sexing strains: laboratory evaluation of strains from mass-rearing facilities worldwide. *Entomol. Exp. Appl.* **164**, 305-317.
- Aumann, R.A., Häcker, I., Schetelig, M.F., 2020. Female-to-male sex conversion in *Ceratitidis capitata* by CRISPR/Cas9 HDR-induced point mutations in the sex determination gene *transformer-2*. *Sci. Rep.* **10**, 18611.
- Aumann, R.A., Schetelig, M.F., Häcker, I., 2018. Highly efficient genome editing by homology-directed repair using Cas9 protein in *Ceratitidis capitata*. *Insect Biochem. Mol. Biol.* **101**, 85-93.
- Badii, K.B., Billah, M.K., Afreh Nuamah, K., Obeng Ofori, D., Nyarko, G., 2015. Review of the pest status, economic impact and management of fruit-infesting flies (Diptera: Tephritidae) in Africa. *African Journal of Agricultural Research* **10**, 1488-1498.
- Bai, X., Zeng, T., Ni, X.Y., Su, H.A., Huang, J., Ye, G.Y., Lu, Y.Y., Qi, Y.X., 2019. CRISPR/Cas9-mediated knockout of the eye pigmentation gene *white* leads to alterations in colour of head spots in the oriental fruit fly, *Bactrocera dorsalis*. *Insect Mol. Biol.* **28**, 837-849.
- Bakri, A., Mehta, K., Lance, D.R., 2005. Sterilizing insects with ionizing radiation, in: Dyck, V.A., Hendrichs, J., Robinson, A.S. (Eds.), *Sterile Insect Technique: Principles and practice in area-wide integrated pest management*. Springer Netherlands, Dordrecht, pp. 233-268.
- Barrangou, R., 2015. Diversity of CRISPR-Cas immune systems and molecular machines. *Genome Biol.* **16**, 247.
- Barrangou, R., Doudna, J.A., 2016. Applications of CRISPR technologies in research and beyond. *Nat. Biotechnol.* **34**, 933-941.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709-1712.
- Baumhover, A.H., Graham, A.J., Bitter, B.A., Hopkins, D.E., New, W.D., Dudley, F.H., Bushland, R.C., 1955. Screwworm control through release of sterilized flies. *J. Econ. Entomol.* **48**, 462-466.
- Ben-Aroya, S., Pan, X., Boeke, J.D., Hieter, P., 2010. Making temperature-sensitive mutants. *Methods Enzymol.* **470**, 181-204.
- Benedict, M., Eckerstorfer, M., Franz, G., Gaugitsch, H., Greiter, A., Heissenberger, A., Knols, B., Kumschick, S., Nentwig, W., Rabitsch, W., 2010. Defining environment risk assessment criteria for genetically modified insects to be placed on the EU market. *EFSA Supporting Publications* **7**, 71E.
- Benedict, M.Q., Robinson, A.S., 2003. The first releases of transgenic mosquitoes: an argument for the sterile insect technique. *Trends Parasitol.* **19**, 349-355.

- Beying, N., Schmidt, C., Pacher, M., Houben, A., Puchta, H., 2020. CRISPR-Cas9-mediated induction of heritable chromosomal translocations in *Arabidopsis*. *Nature Plants* **6**, 638-645.
- Bioökonomierat, 2019. Genome Editing: Europa benötigt ein neues Gentechnikrecht, Börmemo 07, Berlin.
- Black, W.C., Alphey, L., James, A.A., 2011. Why RIDL is not SIT. *Trends Parasitol.* **27**, 362-370.
- Bourtzis, K., Psachoulia, C., Marmaras, V.J., 1991. Evidence that different integumental phosphatases exist during development in the Mediterranean fruit fly *Ceratitis capitata*: possible involvement in pupariation. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* **98**, 411-416.
- Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., van der Oost, J., 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960-964.
- Brown, Z.S., Carter, L., Gould, F., 2018. An introduction to the proceedings of the *Environmental Release of Engineered Pests: Building an International Governance Framework*. *BMC Proceedings* **12**, 10.
- Brunet, E., Jasin, M., 2018. Induction of chromosomal translocations with CRISPR-Cas9 and other nucleases: Understanding the repair mechanisms that give rise to translocations. *Adv. Exp. Med. Biol.* **1044**, 15-25.
- Buchman, A., Akbari, O.S., 2019. Site-specific transgenesis of the *Drosophila melanogaster* Y-chromosome using CRISPR/Cas9. *Insect Mol. Biol.* **28**, 65-73.
- Busch-Petersen, E., 1990. Temperature sensitive lethal factors and puparial colour sex separation mechanisms in the Mediterranean fruit fly, *Ceratitis capitata* (Wied). International Atomic Energy Agency, Vienna.
- Byrne, P.F., Namuth, D.M., Harrington, J., Ward, S.M., Lee, D.J., Hain, P., 2002. Increasing public understanding of transgenic crops through the World Wide Web. *Public Understanding of Science* **11**, 293-304.
- Camacho, A., Van Deynze, A., Chi-Ham, C., Bennett, A.B., 2014. Genetically engineered crops that fly under the US regulatory radar. *Nat. Biotechnol.* **32**, 1087-1091.
- Carvalho, G.B., Ja, W.W., Benzer, S., 2009. Non-lethal genotyping of single *Drosophila*. *BioTechniques* **46**, 312-314.
- Catteruccia, F., Benton, J.P., Crisanti, A., 2005. An *Anopheles* transgenic sexing strain for vector control. *Nat. Biotechnol.* **23**, 1414-1417.
- CBD, 2000. Cartagena protocol on biosafety to the convention on biological diversity: Text and annexes. CBD, Montreal, Canada.
- Chakraborty, S., Newton, A.C., 2011. Climate change, plant diseases and food security: an overview. *Plant Pathol.* **60**, 2-14.
- Champer, J., Buchman, A., Akbari, O.S., 2016. Cheating evolution: engineering gene drives to manipulate the fate of wild populations. *Nat. Rev. Genet.* **17**, 146-159.
- Choi, P.S., Meyerson, M., 2014. Targeted genomic rearrangements using CRISPR/Cas technology. *Nat. Commun.* **5**, 3728.
- Choo, A., Crisp, P., Saint, R., O'Keefe, L.V., Baxter, S.W., 2018. CRISPR/Cas9-mediated mutagenesis of the *white* gene in the tephritid pest *Bactrocera tryoni*. *J. Appl. Entomol.* **142**, 52-58.
- Choo, A., Fung, E., Chen, I.Y., Saint, R., Crisp, P., Baxter, S.W., 2020. Precise single base substitution in the *shibire* gene by CRISPR/Cas9-mediated homology directed repair in *Bactrocera tryoni*. *BMC Genet.* **21**, 127.
- Concha, C., Palavesam, A., Guerrero, F.D., Sagel, A., Li, F., Osborne, J.A., Hernandez, Y., Pardo, T., Quintero, G., Vasquez, M., Keller, G.P., Phillips, P.L., Welch, J.B., McMillan, W.O., Skoda, S.R., Scott, M.J., 2016. A transgenic male-only strain of the New World screwworm for an improved control program using the sterile insect technique. *BMC Biol.* **14**, 72.
- Concha, C., Yan, Y., Arp, A., Quilique, E., Sagel, A., de León, A.P., McMillan, W.O., Skoda, S., Scott, M.J., 2020. An early female lethal system of the New World screwworm, *Cochliomyia hominivorax*, for biotechnology-enhanced SIT. *BMC Genet.* **21**, 143.
- Conseil d'État, 2020. Ruling on litigation No. 388649. <https://www.conseil-etat.fr/ressources/decisions-contentieuses/dernieres-decisions-importantes/conseil-d-etat-7-fevrier-2020-organismes-obtenus-par-mutagenese>
- Dafa'alla, T.H., Condon, G.C., Condon, K.C., Phillips, C.E., Morrison, N.I., Jin, L., Epton, M.J., Fu, G., Alphey, L., 2006. Transposon-free insertions for insect genetic engineering. *Nat. Biotechnol.* **24**, 820-821.
- Das, S.R., Maselko, M., Upadhyay, A., Smanski, M.J., 2020. Genetic engineering of sex chromosomes for batch cultivation of non-transgenic, sex-sorted males. *PLoS Genet.* **16**, e1009180.

- Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., Charpentier, E., 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**, 602-607.
- Deutsch, C.A., Tewksbury, J.J., Tigchelaar, M., Battisti, D.S., Merrill, S.C., Huey, R.B., Naylor, R.L., 2018. Increase in crop losses to insect pests in a warming climate. *Science* **361**, 916-919.
- Devine, G.J., Furlong, M.J., 2007. Insecticide use: Contexts and ecological consequences. *Agriculture and Human Values* **24**, 281-306.
- Diallo, S., Seck, M.T., Rayaissé, J.B., Fall, A.G., Bassene, M.D., Sall, B., Sanon, A., Vreysen, M.J.B., Takac, P., Parker, A.G., Gimonneau, G., Bouyer, J., 2019. Chilling, irradiation and transport of male *Glossina palpalis gambiensis* pupae: Effect on the emergence, flight ability and survival. *PLoS One* **14**, e0216802-e0216802.
- Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., Root, D.E., 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* **32**, 1262-1267.
- Doudna, J.A., Charpentier, E., 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096.
- Eckerstorfer, M.F., Engelhard, M., Heissenberger, A., Simon, S., Teichmann, H., 2019. Plants developed by new genetic modification techniques - Comparison of existing regulatory frameworks in the EU and non-EU countries. *Front. Bioeng. Biotechnol.* **7**.
- EFSA, 2013. Guidance on the environmental risk assessment of genetically modified animals. Panel on Genetically Modified Organisms. *EFSA Journal* **11**, 3200.
- Enserink, M., 2010. GM mosquito trial alarms opponents, strains ties in Gates-funded project. *Science* **330**, 1030-1031.
- Epstein, D.L., Zack, R.S., Brunner, J.F., Gut, L., Brown, J.J., 2000. Effects of broad-spectrum insecticides on epigeal Arthropod biodiversity in Pacific Northwest Apple Orchards. *Environ. Entomol.* **29**, 340-348.
- FAO, IFAD, UNICEF, WFP, WHO, 2019. The state of food security and nutrition in the world 2019. Safeguarding against economic slowdowns and downturns, Rome.
- Fisher, K., Caceres, C., 1998. A filter rearing system for mass reared genetic sexing strains of Mediterranean fruit fly (Diptera: Tephritidae), *Fifth International Symposium on Fruit Flies of Economic Importance: Area-wide control of fruit flies and other insect pests*, Penang, Malaysia, Penerbit Universiti Sains Malaysia, pp. 543-550.
- Fladung, M., 2017. Gene editing: Debate is failing Europe's geneticists. *Nature* **544**, 35.
- Franz, G., 2005. Genetic sexing strains in Mediterranean Fruit Fly, an example for other species amenable to large-scale rearing for the sterile insect technique, in: Dyck, V.A., Hendrichs, J., Robinson, A.S. (Eds.), *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer Netherlands, Dordrecht, pp. 427-451.
- Franz, G., Bourtzis, K., Caceres-Barrios, C., 2021. Practical and operational genetic sexing systems based on classical genetic approaches in fruit flies, an example for other species amenable to large-scale rearing for the sterile insect technique, in: Dyck, V.A., Hendrichs, J., Robinson, A.S. (Eds.), *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management (2nd ed.)*. IAEA. CRC Press, New York, pp. 575-604.
- Franz, G., Gencheva, E., Kerremans, P., 1994. Improved stability of genetic sex-separation strains for the Mediterranean fruit fly, *Ceratitidis capitata*. *Genome* **37**, 72-82.
- Fu, G., Condon, K.C., Epton, M.J., Gong, P., Jin, L., Condon, G.C., Morrison, N.I., Dafa'alla, T.H., Alphey, L., 2007. Female-specific insect lethality engineered using alternative splicing. *Nat. Biotechnol.* **25**, 353-357.
- Gantz, V.M., Bier, E., 2015. Genome editing. The mutagenic chain reaction: a method for converting heterozygous to homozygous mutations. *Science* **348**, 442-444.
- Gantz, V.M., Bier, E., 2016. The dawn of active genetics. *Bioessays* **38**, 50-63.
- Gantz, V.M., Jasinskiene, N., Tatarenkova, O., Fazekas, A., Macias, V.M., Bier, E., James, A.A., 2015. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. *Proc. Natl. Acad. Sci. USA* **112**, E6736-6743.
- Gene editing in legal limbo in Europe, 2017. *Nature* **542**, 392.
- Gene-edited plants cross European event horizon, 2018. *Nat. Biotechnol.* **36**, 776-776.
- Gilles, A.F., Schinko, J.B., Averof, M., 2015. Efficient CRISPR-mediated gene targeting and transgene replacement in the beetle *Tribolium castaneum*. *Development* **142**, 2832-2839.
- Glaser, A., McColl, B., Vadolas, J., 2016. GFP to BFP conversion: A versatile assay for the quantification of CRISPR/Cas9-mediated genome editing. *Molecular therapy. Nucleic acids* **5**, e334-e334.

- Gomulski, L.M., Torti, C., Murelli, V., Bonizzoni, M., Gasperi, G., Malacrida, A.R., 2004. Medfly transposable elements: diversity, evolution, genomic impact and possible applications. *Insect Biochem. Mol. Biol.* **34**, 139-148.
- Gong, P., Epton, M.J., Fu, G., Scaife, S., Hiscox, A., Condon, K.C., Condon, G.C., Morrison, N.I., Kelly, D.W., Dafa'alla, T., Coleman, P.G., Alphey, L., 2005. A dominant lethal genetic system for autocidal control of the Mediterranean fruit fly. *Nat. Biotechnol.* **23**, 453-456.
- Grant, D., Unadkat, S., Katzen, A., Krishnan, K.S., Ramaswami, M., 1998. Probable mechanisms underlying interallelic complementation and temperature-sensitivity of mutations at the *shibire* locus of *Drosophila melanogaster*. *Genetics* **149**, 1019-1030.
- Gregory, P.J., Johnson, S.N., Newton, A.C., Ingram, J.S., 2009. Integrating pests and pathogens into the climate change/food security debate. *J. Exp. Bot.* **60**, 2827-2838.
- Grigliatti, T.A., Hall, L., Rosenbluth, R., Suzuki, D.T., 1973. Temperature-sensitive mutations in *Drosophila melanogaster*. *Mol. Gen. Genet.* **120**, 107-114.
- Grohmann, L., Keilwagen, J., Duensing, N., Dagand, E., Hartung, F., Wilhelm, R., Bendiek, J., Sprink, T., 2019. Detection and identification of genome editing in plants: Challenges and opportunities. *Front. Plant Sci.* **10**.
- Group of Chief Scientific Advisors, 2018. A scientific perspective on the regulatory status of products derived from gene editing and the implications for the GMO directive: Statement by the group of chief scientific advisors. Publications Office of the European Union.
- Gutierrez, A.P., Ponti, L., 2011. Assessing the invasive potential of the Mediterranean fruit fly in California and Italy. *Biol. Invasions* **13**, 2661-2676.
- Habel, J.C., Samways, M.J., Schmitt, T., 2019. Mitigating the precipitous decline of terrestrial European insects: Requirements for a new strategy. *Biodivers. Conserv.* **28**, 1343-1360.
- Häcker, I., Bourtzis, K., Schetelig, M.F., 2021. Applying modern molecular technologies in support of the sterile insect technique, in: Dyck, V.A., Hendrichs, J., Robinson, A.S. (Eds.), *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management (2nd ed.)*. IAEA. CRC Press, New York, pp. 657-702.
- Häcker, I., Schetelig, M.F., 2018. Molecular tools to create new strains for mosquito sexing and vector control. *Parasites Vectors* **11**, 645.
- Häcker, I., Schetelig, M.F., 2021. Genome editing and its applications for insect pest control: Curse or blessing?, in: Hendrichs, J., Pereira, R., Vreysen, M. (Eds.), *Area-wide Integrated Pest Management: Development and Field Application*. CRC Press, Boca Raton, Florida, USA., pp. 809-842.
- Hall, A.B., Basu, S., Jiang, X., Qi, Y., Timoshevskiy, V.A., Biedler, J.K., Sharakhova, M.V., Elahi, R., Anderson, M.A.E., Chen, X.-G., Sharakhov, I.V., Adelman, Z.N., Tu, Z., 2015. A male-determining factor in the mosquito *Aedes aegypti*. *Science* **348**, 1268-1270.
- Handler, A.M., 2016. Enhancing the stability and ecological safety of mass-reared transgenic strains for field release by redundant conditional lethality systems. *Insect Sci.* **23**, 225-234.
- Handler, A.M., O'Brochta, D.A., 2011. Transposable elements for insect transformation, in: Gilbert, L.I. (Ed.), *Insect Biochem. Mol. Biol.* Academic Press, San Diego, pp. 90-133.
- Harris, A.F., Nimmo, D., McKemey, A.R., Kelly, N., Scaife, S., Donnelly, C.A., Beech, C., Petrie, W.D., Alphey, L., 2011. Field performance of engineered male mosquitoes. *Nat. Biotechnol.* **29**, 1034-1037.
- HCB Scientific Committee, 2017. Scientific Opinion of the High Council for Biotechnology concerning use of genetically modified mosquitoes for vector control in response to the referral of 12 October 2015 (Ref. HCB-2017.06.07), Paris, p. 142 pp.
- Hendrichs, J., Franz, G., Rendon, P., 1995. Increased effectiveness and applicability of the Sterile Insect Technique through male-only releases for control of Mediterranean fruit-flies during fruiting seasons. *J. Appl. Entomol.* **119**, 371-377.
- Hendrichs, J., Kenmore, P., Robinson, A.S., Vreysen, M.J.B., 2007. Area-wide integrated pest management (AW-IPM): Principles, practice and prospects, in: Vreysen, M.J.B., Robinson, A.S., Hendrichs, J. (Eds.), *Area-wide Control of Insect Pests*. Springer Netherlands, Dordrecht, pp. 3-33.
- Horn, C., Wimmer, E.A., 2003. A transgene-based, embryo-specific lethality system for insect pest management. *Nat. Biotechnol.* **21**, 64-70.
- Horvath, P., Barrangou, R., 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167-170.
- Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., Cradick, T.J., Marraffini, L.A., Bao, G., Zhang, F., 2013. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827-832.
- IAEA, 1995. Economic evaluation of damage caused by, and methods of control of, the Mediterranean fruit fly in the Maghreb, Vienna.

- Jiang, F., Doudna, J.A., 2017. CRISPR–Cas9 structures and mechanisms. *Annual Review of Biophysics* **46**, 505-529.
- Jiang, F., Liang, L., Li, Z., Yu, Y., Wang, J., Wu, Y., Zhu, S., 2018. A conserved motif within *cox 2* allows broad detection of economically important fruit flies (Diptera: Tephritidae). *Sci. Rep.* **8**, 2077.
- Jiang, J., Zhang, L., Zhou, X., Chen, X., Huang, G., Li, F., Wang, R., Wu, N., Yan, Y., Tong, C., Srivastava, S., Wang, Y., Liu, H., Ying, Q.-L., 2016. Induction of site-specific chromosomal translocations in embryonic stem cells by CRISPR/Cas9. *Sci. Rep.* **6**, 21918.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821.
- Jones, M.S., Delborne, J.A., Elsensohn, J., Mitchell, P.D., Brown, Z.S., 2019. Does the U.S. public support using gene drives in agriculture? And what do they want to know? *Science advances* **5**, eaau8462-eaau8462.
- Kahn, E., Jackson, R.J., Lyman, D.O., Stratton, J.W., 1990. A crisis of community anxiety and mistrust: the medfly eradication project in Santa Clara County, California, 1981-82. *Am. J. Public Health* **80**, 1301-1304.
- KaramiNejadRanjbar, M., Eckermann, K.N., Ahmed, H.M.M., Sánchez, C.H., Dippel, S., Marshall, J.M., Wimmer, E.A., 2018. Consequences of resistance evolution in a Cas9-based sex conversion-suppression gene drive for insect pest management. *Proc. Natl. Acad. Sci. USA* **115**, 6189-6194.
- Kawasaki, F., Felling, R., Ordway, R.W., 2000. A temperature-sensitive *paralytic* mutant defines a primary synaptic calcium channel in *Drosophila*. *The Journal of Neuroscience* **20**, 4885-4889.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647-1649.
- Kerremans, P., Franz, G., 1994. Cytogenetic analysis of chromosome 5 from the Mediterranean fruit fly, *Ceratitidis capitata*. *Chromosoma* **103**, 142-146.
- Kistler, Kathryn E., Voss hall, Leslie B., Matthews, Benjamin J., 2015. Genome Engineering with CRISPR-Cas9 in the Mosquito *Aedes aegypti*. *Cell Reports* **11**, 51-60.
- Kittayapong, P., Ninphanomchai, S., Limohpasmanee, W., Chansang, C., Chansang, U., Mongkalagoon, P., 2019. Combined sterile insect technique and incompatible insect technique: The first proof-of-concept to suppress *Aedes aegypti* vector populations in semi-rural settings in Thailand. *PLOS Neglected Tropical Diseases* **13**, e0007771.
- Klassen, W., 2005. Area-wide integrated pest management and the sterile insect technique, in: Dyck, V.A., Hendrichs, J., Robinson, A.S. (Eds.), *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer Netherlands, Dordrecht, pp. 39-68.
- Klassen, W., Curtis, C.F., 2005. History of the sterile insect technique, in: Dyck, V.A., Hendrichs, J., Robinson, A.S. (Eds.), *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer Netherlands, Dordrecht, pp. 3-36.
- Klofstad, C.A., Uscinski, J.E., Connolly, J.M., West, J.P., 2019. What drives people to believe in Zika conspiracy theories? *Palgrave Communications* **5**, 36.
- Knipling, E.F., 1955. Possibilities of insect control or eradication through the use of sexually sterile males. *J. Econ. Entomol.* **48**, 459-462.
- Kogan, M., 1998. Integrated pest management: Historical perspectives and contemporary developments. *Annu. Rev. Entomol.* **43**, 243-270.
- Koonin, E.V., Makarova, K.S., 2019. Origins and evolution of CRISPR-Cas systems. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **374**, 20180087-20180087.
- Krzywinska, E., Dennison, N.J., Lycett, G.J., Krzywinski, J., 2016. A maleness gene in the malaria mosquito *Anopheles gambiae*. *Science* **353**, 67-69.
- Kuntz, M., 2012. Destruction of public and governmental experiments of GMO in Europe. *GM Crops & Food* **3**, 258-264.
- Kwart, D., Paquet, D., Teo, S., Tessier-Lavigne, M., 2017. Precise and efficient scarless genome editing in stem cells using CORRECT. *Nat. Protoc.* **12**, 329-354.
- Lance, D.R., McInnis, D.O., 2005. Biological basis of the sterile insect technique, in: Dyck, V.A., Hendrichs, J., Robinson, A.S. (Eds.), *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer Netherlands, Dordrecht, pp. 69-94.
- Laven, H., 1969. Eradicating mosquitoes using translocations. *Nature* **221**, 958-959.
- Lehane, M.J., Aksoy, S., 2012. Control using genetically modified insects poses problems for regulators. *PLoS Negl Trop Dis* **6**, e1495.

- Li, F., Wantuch, H.A., Linger, R.J., Belikoff, E.J., Scott, M.J., 2014. Transgenic sexing system for genetic control of the Australian sheep blow fly *Lucilia cuprina*. *Insect Biochem. Mol. Biol.* **51**, 80-88.
- Li, J., Handler, A.M., 2017. Temperature-dependent sex-reversal by a *transformer-2* gene-edited mutation in the spotted wing drosophila, *Drosophila suzukii*. *Sci. Rep.* **7**, 12363.
- Li, J., Handler, A.M., 2019. CRISPR/Cas9-mediated gene editing in an exogenous transgene and an endogenous sex determination gene in the Caribbean fruit fly, *Anastrepha suspensa*. *Gene* **691**, 160-166.
- Liew, C., Soh, L.T., Chen, I., Li, X., Sim, S., Ng, L.C., 2021. Community engagement for *Wolbachia*-based *Aedes aegypti* population suppression for Dengue control: the Singapore experience, in: Hendrichs, J., Pereira, R., Vreysen, M. (Eds.), *Area-wide Integrated Pest Management: Development and Field Application*. IAEA. CRC Press, Boca Raton, Florida, USA., pp. 747-761.
- Lutrat, C., Giesbrecht, D., Marois, E., Whyard, S., Baldet, T., Bouyer, J., 2019. Sex sorting for pest control: It's raining men! *Trends Parasitol.* **35**, 649-662.
- Makarova, K.S., Haft, D.H., Barrangou, R., Brouns, S.J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F.J., Wolf, Y.I., Yakunin, A.F., van der Oost, J., Koonin, E.V., 2011. Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **9**, 467-477.
- Malacrida, A.R., Gomulski, L.M., Bonizzoni, M., Bertin, S., Gasperi, G., Guglielmino, C.R., 2007. Globalization and fruitfly invasion and expansion: the medfly paradigm. *Genetica* **131**, 1-9.
- Mallapaty, S., 2019. Australian gene-editing rules adopt 'middle ground'. *Nature*.
- Marec, F., Neven, L.G., Robinson, A.S., Vreysen, M., Goldsmith, M.R., Nagaraju, J., Franz, G., 2005. Development of genetic sexing strains in Lepidoptera: from traditional to transgenic approaches. *J. Econ. Entomol.* **98**, 248-259.
- Marshall, J.M., 2010. The Cartagena Protocol and genetically modified mosquitoes. *Nat. Biotechnol.* **28**, 896-897.
- Marshall, J.M., Buchman, A., Sánchez C, H.M., Akbari, O.S., 2017. Overcoming evolved resistance to population-suppressing homing-based gene drives. *Sci. Rep.* **7**, 3776.
- Max-Planck-Gesellschaft, 2019. Stellungnahme zu den wissenschaftlichen und translationalen Auswirkungen der Genom-Editierung und daraus resultierenden ethischen, rechtlichen und gesellschaftlichen Fragen.
- Maxmen, A., 2013. Crop pests: Under attack. *Nature* **501**, 15-17.
- McCombs, S.D., Saul, S.H., 1992. Linkage analysis of five new genetic markers of the oriental fruit fly, *Bactrocera dorsalis* (Diptera: Tephritidae). *J. Hered.* **83**, 199-203.
- McInnis, D.O., Tam, S., Lim, R., Komatsu, J., Kurashima, R., Albrecht, C., 2004. Development of a pupal color-based genetic sexing strain of the melon fly, *Bactrocera cucurbitae* (Coquillett) (Diptera: Tephritidae). *Ann. Entomol. Soc. Am.* **97**, 1026-1033.
- Meats, A., Maheswaran, P., Frommer, M., Sved, J., 2002. Towards a male-only release system for SIT with the Queensland fruit fly, *Bactrocera tryoni*, using a Genetic Sexing Strain with a temperature-sensitive lethal mutation. *Genetica* **116**, 97-106.
- Meccariello, A., Monti, S.M., Romanelli, A., Colonna, R., Primo, P., Inghilterra, M.G., Del Corsano, G., Ramaglia, A., Iazzetti, G., Chiarore, A., Patti, F., Heinze, S.D., Salvemini, M., Lindsay, H., Chiavacci, E., Burger, A., Robinson, M.D., Mosimann, C., Bopp, D., Saccone, G., 2017. Highly efficient DNA-free gene disruption in the agricultural pest *Ceratitidis capitata* by CRISPR-Cas9 ribonucleoprotein complexes. *Sci. Rep.* **7**, 10061.
- Meccariello, A., Salvemini, M., Primo, P., Hall, B., Koskinioti, P., Dalikova, M., Gravina, A., Gucciardino, M.A., Forlenza, F., Gregoriou, M.E., Ippolito, D., Monti, S.M., Petrella, V., Perrotta, M.M., Schmeing, S., Ruggiero, A., Scolari, F., Giordano, E., Tsoumani, K.T., Marec, F., Windbichler, N., Arunkumar, K.P., Bourtzis, K., Mathiopoulos, K.D., Ragoussis, J., Vitagliano, L., Tu, Z., Papathanos, P.A., Robinson, M.D., Saccone, G., 2019. *Maleness-on-the-Y (MoY)* orchestrates male sex determination in major agricultural fruit fly pests. *Science* **365**, 1457-1460.
- Meccariello, A., Tsoumani, K.T., Gravina, A., Primo, P., Buonanno, M., Mathiopoulos, K.D., Saccone, G., 2020. Targeted somatic mutagenesis through CRISPR/Cas9 ribonucleoprotein complexes in the olive fruit fly, *Bactrocera oleae*. *Arch. Insect Biochem. Physiol.*, e21667.
- Mestdagh, S., Devos, Y., Ehlert, C., Liu, Y., Podevin, N., Rodighiero, S., Waigmann, E., Kiss, J., Perry, J.N., Sweet, J.B., 2014. EFSA Guidelines on the environmental risk assessment of genetically modified animals in the EU: the process and risk assessment considerations. *Journal für Verbraucherschutz und Lebensmittelsicherheit* **9**, 85-91.
- Metje-Sprink, J., Menz, J., Modrzejewski, D., Sprink, T., 2018. DNA-Free genome editing: Past, present and future. *Front Plant Sci* **9**, 1957.

- Meza, J.S., Caceres, C., Bourtzis, K., 2019. Slow larvae mutant and its potential to improve the pupal color-based genetic sexing system in Mexican Fruit Fly, (Diptera: Tephritidae). *J. Econ. Entomol.* **112**, 1604-1610.
- Morrison, N., Franz, G., Koukidou, M., Miller, T., Saccone, G., Alphey, L., Beech, C., Nagaraju, J., Simmons, G., Polito, L., 2010. Genetic improvements to the sterile insect technique for agricultural pests. *AsPac J Mol Biol Biotechnol* **18**, 275-295.
- Mumford, J.D., 2012. Science, regulation, and precedent for genetically modified insects. *PLoS Negl Trop Dis* **6**, e1504.
- Nationale Akademie der Wissenschaften Leopoldina, Deutsche Forschungsgemeinschaft, Union der deutschen Akademien der Wissenschaften, 2019. Wege zu einer wissenschaftlich begründeten, differenzierten Regulierung genomeditierter Pflanzen in der EU, Halle (Saale).
- Nirmala, X., Olson, S.R., Holler, T.C., Cho, K.H., Handler, A.M., 2011. A DsRed fluorescent protein marker under polyubiquitin promoter regulation allows visual and amplified gene detection of transgenic Caribbean fruit flies in field traps. *BioControl* **56**, 333-340.
- Niyazi, N., Caceres, C., Delprat, A., Wornoayporn, V., Santos, E.R., Franz, G., Robinson, A.S., 2005. Genetics and mating competitiveness of *Ceratitidis capitata* (Diptera: Tephritidae) strains carrying the marker *Sergeant*, Sr2. *Ann. Entomol. Soc. Am.* **98**, 119-125, 117.
- O'Brochta, D., Tonui, W., Dass, B., James, S., 2020. A cross-sectional survey of biosafety professionals regarding genetically modified insects. *Applied Biosafety* **25**, 19-27.
- O'Brochta, D.A., Atkinson, P.W., 1996. Transposable elements and gene transformation in non-drosophilid insects. *Insect Biochem. Mol. Biol.* **26**, 739-753.
- Ogaugwu, C.E., Schetelig, M.F., Wimmer, E.A., 2013. Transgenic sexing system for *Ceratitidis capitata* (Diptera: Tephritidae) based on female-specific embryonic lethality. *Insect Biochem. Mol. Biol.* **43**, 1-8.
- Oh, New Delhi; oh, Geneva, 1975. *Nature* **256**, 355-357.
- Okamoto, S., Amaishi, Y., Maki, I., Enoki, T., Mineno, J., 2019. Highly efficient genome editing for single-base substitutions using optimized ssODNs with Cas9-RNPs. *Sci. Rep.* **9**, 4811.
- Orozco, D., Meza, J.S., Zepeda, S., Solis, E., Quintero-Fong, J.L., 2013. Tapachula-7, a new genetic sexing strain of the Mexican fruit fly (Diptera: Tephritidae): sexual compatibility and competitiveness. *J. Econ. Entomol.* **106**, 735-741.
- Paix, A., Folkmann, A., Goldman, D.H., Kulaga, H., Grzelak, M.J., Rasoloson, D., Paidemarry, S., Green, R., Reed, R.R., Seydoux, G., 2017. Precision genome editing using synthesis-dependent repair of Cas9-induced DNA breaks. *Proc. Natl. Acad. Sci.* **114**, E10745-E10754.
- Paix, A., Wang, Y., Smith, H.E., Lee, C.Y., Calidas, D., Lu, T., Smith, J., Schmidt, H., Krause, M.W., Seydoux, G., 2014. Scalable and versatile genome editing using linear DNAs with microhomology to Cas9 Sites in *Caenorhabditis elegans*. *Genetics* **198**, 1347-1356.
- Pane, A., Salvemini, M., Delli Bovi, P., Polito, C., Saccone, G., 2002. The *transformer* gene in *Ceratitidis capitata* provides a genetic basis for selecting and remembering the sexual fate. *Development* **129**, 3715-3725.
- Panjwani, A., Wilson, A., 2016. What is stopping the use of genetically modified insects for disease control? *PLoS Path.* **12**, e1005830.
- Pao, S.S., Paulsen, I.T., Saier, M.H., Jr., 1998. Major facilitator superfamily. *Microbiology and molecular biology reviews: MMBR* **62**, 1-34.
- Papanicolaou, A., Schetelig, M.F., Arensburger, P., Atkinson, P.W., Benoit, J.B., Bourtzis, K., Castañera, P., Cavanaugh, J.P., Chao, H., Childers, C., Curral, I., Dinh, H., Doddapaneni, H., Dolan, A., Dugan, S., Friedrich, M., Gasperi, G., Geib, S., Georgakilas, G., Gibbs, R.A., Giers, S.D., Gomulski, L.M., González-Guzmán, M., Guillem-Amat, A., Han, Y., Hatzigeorgiou, A.G., Hernández-Crespo, P., Hughes, D.S.T., Jones, J.W., Karagkouni, D., Koskinioti, P., Lee, S.L., Malacrida, A.R., Manni, M., Mathiopoulos, K., Meccariello, A., Murali, S.C., Murphy, T.D., Muzny, D.M., Oberhofer, G., Ortego, F., Paraskevopoulou, M.D., Poelchau, M., Qu, J., Reczko, M., Robertson, H.M., Rosendale, A.J., Rosselot, A.E., Saccone, G., Salvemini, M., Savini, G., Schreiner, P., Scolari, F., Siciliano, P., Sim, S.B., Tsiamis, G., Ureña, E., Vlachos, I.S., Werren, J.H., Wimmer, E.A., Worley, K.C., Zacharopoulou, A., Richards, S., Handler, A.M., 2016. The whole genome sequence of the Mediterranean fruit fly, *Ceratitidis capitata* (Wiedemann), reveals insights into the biology and adaptive evolution of a highly invasive pest species. *Genome Biol.* **17**, 192.
- Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K.M., Gregg, A., Noggle, S., Tessier-Lavigne, M., 2016. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature* **533**, 125-129.

- Parker, A.G., 2005. Mass-rearing for sterile insect release, in: Dyck, V.A., Hendrichs, J., Robinson, A.S. (Eds.), *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer Netherlands, Dordrecht, pp. 209-232.
- Peccoud, J., Loiseau, V., Cordaux, R., Gilbert, C., 2017. Massive horizontal transfer of transposable elements in insects. *Proc. Natl. Acad. Sci.*, 201621178.
- Pimentel, D., 2007. Area-wide pest management: Environmental, economic and food issues, in: Vreysen, M.J.B., Robinson, A.S., Hendrichs, J. (Eds.), *Area-wide Control of Insect Pests*. Springer Netherlands, Dordrecht, pp. 35-47.
- Porras, M.F., Meza, J.S., Rajotte, E.G., Bourtzis, K., Cáceres, C., 2020. Improving the phenotypic properties of the *Ceratitidis capitata* (Diptera: Tephritidae) temperature-sensitive lethal Genetic Sexing Strain in support of Sterile Insect Technique applications. *J. Econ. Entomol.*
- Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281-2308.
- Rashid, M.A., Andongma, A.A., Dong, Y.-C., Ren, X.-M., Niu, C.-Y., 2018. Effect of gut bacteria on fitness of the Chinese citrus fly, *Bactrocera minax* (Diptera: Tephritidae). *Symbiosis* **76**, 63-69.
- Rat der Europäischen Union, 2019. Beschluss (EU) 2019/1904 des Rates vom 8. November 2019 mit dem Ersuchen an die Kommission, eine Untersuchung im Lichte des Urteils des Gerichtshofs in der Rechtssache C-528/16 zu dem Status neuartiger genomischer Verfahren im Rahmen des Unionsrechts sowie — falls angesichts der Ergebnisse der Untersuchung angemessen — einen Vorschlag zu unterbreiten, in: Union, R.d.E. (Ed.). Amtsblatt der Europäischen Union, Brüssel.
- Reeves, R.G., Denton, J.A., Santucci, F., Bryk, J., Reed, F.A., 2012. Scientific standards and the regulation of genetically modified insects. *PLoS Neglected Tropical Diseases* **6**, e1502.
- Rempoulakis, P., Taret, G., Haq, I.U., Wornayporn, V., Ahmad, S., Sto Tomas, U., Dammalage, T., Gembinsky, K., Franz, G., Cáceres, C., Vreysen, M.J.B., 2016. Evaluation of quality production parameters and mating behavior of novel genetic sexing strains of the Mediterranean fruit fly *Ceratitidis capitata* (Wiedemann) (Diptera: Tephritidae). *PLoS One* **11**, e0157679-e0157679.
- Rendon, P., McInnis, D., Lance, D., Stewart, J., 2004. Medfly (Diptera: Tephritidae) genetic sexing: large-scale field comparison of males-only and bisexual sterile fly releases in Guatemala. *J. Econ. Entomol.* **97**, 1547-1553.
- Ricalde, M.P., Nava, D.E., Loeck, A.E., Donatti, M.G., 2012. Temperature-dependent development and survival of Brazilian populations of the Mediterranean fruit fly, *Ceratitidis capitata*, from tropical, subtropical and temperate regions. *Journal of insect science (Online)* **12**, 33-33.
- Robinson, A.S., 2002. Mutations and their use in insect control. *Mutation Research/Reviews in Mutation Research* **511**, 113-132.
- Robinson, A.S., Hendrichs, J., 2005. Prospects for the future development and application of the sterile insect technique, in: Dyck, V.A., Hendrichs, J., Robinson, A.S. (Eds.), *Sterile Insect Technique: Principles and Practice in Area-Wide Integrated Pest Management*. Springer Netherlands, Dordrecht, pp. 727-760.
- Robinson, A.S., Van Heemert, C., 1982. *Ceratitidis capitata* - a suitable case for genetic sexing. *Genetica* **58**, 229-237.
- Rössler, Y., 1979. The genetics of the Mediterranean fruit fly: a "white pupae" mutant. *Ann. Entomol. Soc. Am.* **72**, 583-585.
- Rössler, Y., Rosenthal, H., 1992. Genetics of the Mediterranean fruit fly (Diptera: Tephritidae): morphological mutants on chromosome five. *Ann. Entomol. Soc. Am.* **85**, 525-531.
- Saccone, G., Pane, A., De Simone, A., Salvemini, M., Milano, A., Annunziata, L., Mauro, U., Polito, L.C., 2007. New sexing strains for Mediterranean fruit fly *Ceratitidis capitata*: Transforming females into males, in: Vreysen, M.J.B., Robinson, A.S., Hendrichs, J. (Eds.), *Area-wide Control of Insect Pests*. Springer Netherlands, Dordrecht, pp. 95-102.
- Saleh-Gohari, N., Helleday, T., 2004. Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Res.* **32**, 3683-3688.
- Salvemini, M., Robertson, M., Aronson, B., Atkinson, P., Polito, L.C., Saccone, G., 2009. *Ceratitidis capitata transformer-2* gene is required to establish and maintain the autoregulation of *Cetra*, the master gene for female sex determination. *Int. J. Dev. Biol.* **53**, 109-120.
- Sapranaukas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., Siksnys, V., 2011. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* **39**, 9275-9282.
- Schetelig, M.F., Cáceres, C., Zacharopoulou, A., Franz, G., Wimmer, E.A., 2009. Conditional embryonic lethality to improve the sterile insect technique in *Ceratitidis capitata* (Diptera: Tephritidae). *BMC Biol.* **7**, 4.

- Schetelig, M.F., Handler, A.M., 2012a. Strategy for enhanced transgenic strain development for embryonic conditional lethality in *Anastrepha suspensa*. *Proc. Natl. Acad. Sci.* **109**, 9348-9353.
- Schetelig, M.F., Handler, A.M., 2012b. A transgenic embryonic sexing system for *Anastrepha suspensa* (Diptera: Tephritidae). *Insect Biochem. Mol. Biol.* **42**.
- Schetelig, M.F., Milano, A., Saccone, G., Handler, A.M., 2012. Male only progeny in *Anastrepha suspensa* by RNAi-induced sex reversion of chromosomal females. *Insect Biochem. Mol. Biol.* **42**, 51-57.
- Schetelig, M.F., Targovska, A., Meza, J.S., Bourtzis, K., Handler, A.M., 2016. Tetracycline-suppressible female lethality and sterility in the Mexican fruit fly, *Anastrepha ludens*. *Insect Mol. Biol.* **25**, 500-508.
- Schmidt, C., Schindele, P., Puchta, H., 2020a. From gene editing to genome engineering: restructuring plant chromosomes via CRISPR/Cas. *aBIOTECH* **1**, 21-31.
- Schmidt, S.M., Belisle, M., Frommer, W.B., 2020b. The evolving landscape around genome editing in agriculture. *EMBO reports* **21**, e50680.
- Schwindenhammer, S., 2020. The rise, regulation and risks of genetically modified insect technology in global agriculture. *Science, Technology and Society* **25**, 124-141.
- Scolari, F., Schetelig, M.F., Bertin, S., Malacrida, A.R., Gasperi, G., Wimmer, E.A., 2008. Fluorescent sperm marking to improve the fight against the pest insect *Ceratitis capitata* (Wiedemann; Diptera: Tephritidae). *New biotechnology* **25**, 76-84.
- Scott, M.J., Concha, C., Welch, J.B., Phillips, P.L., Skoda, S.R., 2017. Review of research advances in the screwworm eradication program over the past 25 years. *Entomol. Exp. Appl.* **164**, 226-236.
- Sharma, A., Heinze, S.D., Wu, Y., Kohlbrenner, T., Morilla, I., Brunner, C., Wimmer, E.A., van de Zande, L., Robinson, M.D., Beukeboom, L.W., Bopp, D., 2017. Male sex in houseflies is determined by *Mdmd*, a paralog of the generic splice factor gene *CWC22*. *Science* **356**, 642-645.
- Shew, A.M., Nalley, L.L., Snell, H.A., Nayga, R.M., Dixon, B.L., 2018. CRISPR versus GMOs: Public acceptance and valuation. *Global Food Security* **19**, 71-80.
- Sim, S.B., Kauwe, A.N., Ruano, R.E.Y., Rendon, P., Geib, S.M., 2019. The ABCs of CRISPR in Tephritidae: developing methods for inducing heritable mutations in the genera *Anastrepha*, *Bactrocera* and *Ceratitis*. *Insect Mol. Biol.* **28**, 277-289.
- Simmons, G.S., McKemey, A.R., Morrison, N.I., O'Connell, S., Tabashnik, B.E., Claus, J., Fu, G., Tang, G., Sledge, M., Walker, A.S., Phillips, C.E., Miller, E.D., Rose, R.I., Staten, R.T., Donnelly, C.A., Alphey, L., 2011. Field performance of a genetically engineered strain of Pink Bollworm. *PLoS One* **6**, e24110.
- Sprink, T., Eriksson, D., Schiemann, J., Hartung, F., 2016. Regulatory hurdles for genome editing: process- vs. product-based approaches in different regulatory contexts. *Plant Cell Rep.* **35**, 1493-1506.
- Sternberg, S.H., Doudna, J.A., 2015. Expanding the biologist's toolkit with CRISPR-Cas9. *Mol. Cell* **58**, 568-574.
- Stewart, C.N., Richards, H.A., Halfhill, M.D., 2000. Transgenic plants and biosafety: Science, misconceptions and public perceptions. *BioTechniques* **29**, 832-843.
- Sultana, S., Baumgartner, J.B., Dominiak, B.C., Royer, J.E., Beaumont, L.J., 2017. Potential impacts of climate change on habitat suitability for the Queensland fruit fly. *Sci. Rep.* **7**, 13025.
- Sun, D., Guo, Z., Liu, Y., Zhang, Y., 2017. Progress and prospects of CRISPR/Cas systems in insects and other Arthropods. *Frontiers in Physiology* **8**.
- Suzuki, D.T., 1970. Temperature-sensitive mutations in *Drosophila melanogaster*. *Science* **170**, 695-706.
- Suzuki, D.T., Grigliatti, T., Williamson, R., 1971. Temperature-sensitive mutations in *Drosophila melanogaster*. VII. A mutation (*para-ts*) causing reversible adult paralysis. *Proc. Natl. Acad. Sci. USA* **68**, 890-893.
- Szyniszewska, A.M., Tatem, A.J., 2014. Global assessment of seasonal potential distribution of Mediterranean fruit fly, *Ceratitis capitata* (Diptera: Tephritidae). *PLoS One* **9**, e111582-e111582.
- Tarasoff, M., Suzuki, D.T., 1970. Temperature-sensitive mutations in *Drosophila melanogaster*; VI. Temperature effects on development of sex-linked recessive lethals. *Dev. Biol.* **23**, 492-509.
- Thomas, D.D., Donnelly, C.A., Wood, R.J., Alphey, L.S., 2000. Insect population control using a dominant, repressible, lethal genetic system. *Science* **287**, 2474-2476.
- Thygesen, P., 2019. Clarifying the regulation of genome editing in Australia: situation for genetically modified organisms. *Transgenic Res.* **28**, 151-159.
- Torres, R., Martin, M.C., Garcia, A., Cigudosa, J.C., Ramirez, J.C., Rodriguez-Perales, S., 2014. Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system. *Nat. Commun.* **5**, 3964.
- Turner, G., Beech, C., Roda, L., 2018. Means and ends of effective global risk assessments for genetic pest management. *BMC proceedings* **12**, 13-13.

- Unckless, R.L., Clark, A.G., Messer, P.W., 2017. Evolution of resistance against CRISPR/Cas9 Gene Drive. *Genetics* **205**, 827-841.
- Urnov, F.D., Ronald, P.C., Carroll, D., 2018. A call for science-based review of the European court's decision on gene-edited crops. *Nat. Biotechnol.* **36**, 800-802.
- Van Eenennaam, A.L., 2018. The importance of a novel product risk-based trigger for gene-editing regulation in food animal species. *Crispr Journal* **1**, 101-106.
- Van Eenennaam, A.L., Wells, K.D., Murray, J.D., 2019. Proposed U.S. regulation of gene-edited food animals is not fit for purpose. *NPJ Sci Food* **3**, 3.
- Vera, M.T., Rodriguez, R., Segura, D.F., Cladera, J.L., Sutherst, R.W., 2002. Potential geographical distribution of the Mediterranean fruit fly, *Ceratitidis capitata* (Diptera: Tephritidae), with emphasis on Argentina and Australia. *Environ. Entomol.* **31**, 1009-1022.
- Vives-Vallés, J.A., Collonnier, C., 2020. The judgment of the CJEU of 25 July 2018 on mutagenesis: Interpretation and interim legislative proposal. *Front. Plant Sci.* **10**.
- Walters, M., Morrison, N.I., Claus, J., Tang, G., Phillips, C.E., Young, R., Zink, R.T., Alphey, L., 2012. Field longevity of a fluorescent protein marker in an engineered strain of the Pink Bollworm, *Pectinophora gossypiella* (Saunders). *PLoS One* **7**, e38547-e38547.
- Waltz, E., 2016a. CRISPR-edited crops free to enter market, skip regulation. *Nat. Biotechnol.* **34**, 582-582.
- Waltz, E., 2016b. Gene-edited CRISPR mushroom escapes US regulation. *Nature* **532**, 293.
- Waltz, E., 2017. GM moths with autocidal gene tested outdoors in New York state. *Nat. Biotechnol.* **35**, 896-896.
- Wang, K., Tang, X., Liu, Y., Xie, Z., Zou, X., Li, M., Yuan, H., Ouyang, H., Jiao, H., Pang, D., 2016. Efficient generation of orthologous point mutations in pigs via CRISPR-assisted ssODN-mediated homology-directed repair. *Molecular Therapy - Nucleic Acids* **5**, e396.
- Wappner, P., Hopkins, T.L., Kramer, K.J., Cladera, J.L., Manso, F., Quesada-Allué, L.A., 1996. Role of catecholamines and β -alanine in puparial color of wild-type and melanic mutants of the Mediterranean fruit fly (*Ceratitidis capitata*). *J. Insect Physiol.* **42**, 455-461.
- Wappner, P., Kramer, K., Hopkins, T., Merritt, M., Schaefer, J., Quesada-Allue, L., 1995. *White pupa*: A *Ceratitidis capitata* mutant lacking catecholamines for tanning the puparium. *Insect Biochem. Molec. Biol.* **25**, 365-373.
- Ward, C.M., Aumann, R.A., Whitehead, M.A., Nikolouli, K., Leveque, G., Gouvi, G., Fung, E., Reiling, S.J., Djambazian, H., Hughes, M.A., Whiteford, S., Caceres-Barrios, C., Nguyen, T.N.M., Choo, A., Crisp, P., Sim, S.B., Geib, S.M., Marec, F., Häcker, I., Ragoussis, J., Darby, A.C., Bourtzis, K., Baxter, S.W., Schetelig, M.F., 2021. White pupae phenotype of tephritids is caused by parallel mutations of a MFS transporter. *Nat. Commun.* **12**, 491.
- Wearing, C.H., 1982. Integrated pest management - progress and prospects, with special reference to horticulture. *New Zealand Journal of Experimental Agriculture* **10**, 87-94.
- Whalon, M.E., Mota-Sanchez, D., Hollingworth, R.M., 2008. Analysis of global pesticide resistance in arthropods, in: Whalon, M.E., Mota-Sanchez, D., Hollingworth, R.M. (Eds.), *Global Pesticide Resistance in Arthropods*. Cromwell Press, Trowbridge, pp. 5-30.
- Wheeler, T., von Braun, J., 2013. Climate change impacts on global food security. *Science* **341**, 508-513.
- Whelan, A.I., Lema, M.A., 2015. Regulatory framework for gene editing and other new breeding techniques (NBTs) in Argentina. *GM Crops Food* **6**, 253-265.
- White, I.M., Elson-Harris, M.M., 1992. *Fruit flies of economic significance: Their identification and bionomics*, Wallingford UK.
- Wilbie, D., Walther, J., Mastrobattista, E., 2019. Delivery aspects of CRISPR/Cas for *in vivo* genome editing. *Accounts of chemical research* **52**, 1555-1564.
- Wimmer, E.A., 2005. Eco-friendly insect management. *Nat. Biotechnol.* **23**, 432-433.
- Wolt, J.D., Wolf, C., 2018. Policy and governance perspectives for regulation of genome edited crops in the United States. *Front. Plant Sci.* **9**, 1606-1606.
- Wu, C.-F., Ganetzky, B., 1980. Genetic alteration of nerve membrane excitability in temperature-sensitive *paralytic* mutants of *Drosophila melanogaster*. *Nature* **286**, 814-816.
- Wunderlich, S., Gatto, K.A., 2015. Consumer perception of genetically modified organisms and sources of information. *Adv Nutr* **6**, 842-851.
- Yan, Y., Linger, R.J., Scott, M.J., 2017. Building early-larval sexing systems for genetic control of the Australian sheep blow fly *Lucilia cuprina* using two constitutive promoters. *Sci. Rep.* **7**, 2538.
- Yan, Y., Scott, M.J., 2015. A transgenic embryonic sexing system for the Australian sheep blow fly *Lucilia cuprina*. *Sci. Rep.* **5**, 16090.

- Yan, Y., Williamson, M.E., Davis, R.J., Andere, A.A., Picard, C.J., Scott, M.J., 2020. Improved transgenic sexing strains for genetic control of the Australian sheep blow fly *Lucilia cuprina* using embryo-specific gene promoters. *Mol. Genet. Genomics* **295**, 287-298.
- Yosef, I., Goren, M.G., Qimron, U., 2012. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* **40**, 5569-5576.
- Yuan, S., Ma, W., Besley, J.C., 2019. Should scientists talk about GMOs nicely? Exploring the effects of communication styles, source expertise, and preexisting attitude. *Sci. Commun.* **41**, 267-290.
- Zacharopoulou, A., Augustinos, A.A., Drosopoulou, E., Tsoumani, K.T., Gariou-Papalexiou, A., Franz, G., Mathiopoulos, K.D., Bourtzis, K., Mavragani-Tsipidou, P., 2017. A review of more than 30 years of cytogenetic studies of Tephritidae in support of sterile insect technique and global trade. *Entomol. Exp. Appl.* **164**, 204-225.
- Zaidi, S.S.-e.-A., Vanderschuren, H., Qaim, M., Mahfouz, M.M., Kohli, A., Mansoor, S., Tester, M., 2019. New plant breeding technologies for food security. *Science* **363**, 1390-1391.
- Zentrale Kommission für die biologische Sicherheit, 2018. Genome Editing - Auswirkungen des EuGH-Urteils auf die Pflanzenzüchtung.
- Zepeda-Cisneros, C.S., Meza Hernández, J.S., García-Martínez, V., Ibañez-Palacios, J., Zacharopoulou, A., Franz, G., 2014. Development, genetic and cytogenetic analyses of genetic sexing strains of the Mexican fruit fly, *Anastrepha ludens* Loew (Diptera: Tephritidae). *BMC Genet.* **15 Suppl 2**, S1-S1.
- Zhao, Y., Schetelig, M.F., Handler, A.M., 2020. Genetic breakdown of a Tet-off conditional lethality system for insect population control. *Nat. Commun.* **11**, 3095.
- Zheng, W., Li, Q., Sun, H., Ali, M.W., Zhang, H., 2019. Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated 9-mediated mutagenesis of the *multiple edematous wings* gene induces muscle weakness and flightlessness in *Bactrocera dorsalis* (Diptera: Tephritidae). *Insect Mol. Biol.* **28**, 222-234.
- Zhou, Q., Haymer, D.S., 1997. Molecular structure of *yoyo*, a gypsy-like retrotransposon from the Mediterranean fruit fly, *Ceratitidis capitata*. *Genetica* **101**, 167-178.

6 Appendix

6.1 MATERIAL AND DATASETS USED FOR THESE STUDIES.....	122
6.1.1 <i>C. capitata</i> strains	122
6.1.2 <i>C. capitata</i> genome assembly versions.....	122
6.1.3 Oligonucleotides and HDR templates	123
6.2 METHODS USED OR ESTABLISHED DURING THESE STUDIES.....	125
6.2.1 CRISPR/Cas9 gene editing.....	125
6.2.2 gRNA synthesis (Protocol: Predrag Kalajdzic)	127
6.2.3 Non-lethal genotyping of <i>C. capitata</i> using single-leg DNA (Protocol: R. A. Aumann)	129
6.2.4 PCR assays to distinguish VIENNA ^(D53) GSS and WT medflies	130
6.2.4.1 <i>white pupae</i> gene based multiplex PCR (Protocol: R. A. Aumann).....	130
6.2.4.2 D53 inversion detection assay (Protocol: R. A. Aumann).....	131

6.1 Material and datasets used for these studies

6.1.1 *C. capitata* strains

Strain	Sex	Pupa color/ marker gene	<i>wp/tsl</i> genotype	D53 inversion	(Y;5) transloc.	Used for project
Egypt II (EgII) ¹	male, female	brown	<i>wp</i> ⁺ / <i>tsl</i> ⁺	no	no	3.1, 3.2, 3.3
D53 ¹	male, female	white	<i>wp</i> ⁻ / <i>tsl</i> ⁻	yes	no	3.2
VIENNA 7 ¹	male	brown	<i>wp</i> ^{+/-} / <i>tsl</i> ^{+/-}	yes, het	yes	3.2
VIENNA 7 ¹	female	white	<i>wp</i> ⁻ / <i>tsl</i> ⁻	yes	no	3.2
<i>wp/tsl</i> (in EgII) ¹	male, female	white	<i>wp</i> ⁻ / <i>tsl</i> ⁻	no	no	3.2
1402_22m1B ²	male, female	white, eGFP	<i>wp</i> ⁻ / <i>tsl</i> ⁺	no	no	3.2
1247_F1m2 ³	male, female	brown, eGFP	<i>wp</i> ⁺ / <i>tsl</i> ⁺	no	no	3.1

¹ received from the IPCL Seibersdorf (FAO/IAEA)

² (Ogaugwu et al., 2013)

³ (Schetelig et al., 2009)

6.1.2 *C. capitata* genome assembly versions

Assembly version	Accession number	BioProjects	Sequencing technology
Ccap_1.0	GCA_000347755.1	PRJNA168120	Illumina HiSeq
Ccap_1.1	GCA_000347755.2	PRJNA168120	Illumina HiSeq
Ccap_2.1	GCA_000347755.4	PRJNA201381, PRJNA168120	Illumina HiSeq
Ccap_3.1	Not published	--	PacBio, Hi-C, Nanopore, Illumina HiSeq
Ccap_3.2	Not published	--	
Ccap_3.2.1	GCA_905071925	PRJEB36344, ERP119522	

6.1.3 Oligonucleotides and HDR templates

Name	Sequence	Purpose
P_986	GAAATTAATACGACTCACTATAGGCTCGTGACCAC CCTGACCTAGTTTTAGAGCTAGAAATAGC	gRNA_eGFP_2 <i>in vitro</i> synthesis
P_1172	GAAATTAATACGACTCACTATAGGCTGAAGCACTG CACGCCGTGTTTTAGAGCTAGAAATAGC	gRNA_eGFP_2b <i>in vitro</i> synthesis
P_1439	GAAATTAATACGACTCACTATAGGTGATGATATAG CTGATGCTAGTTTTAGAGCTAGAAATAGC	gRNA_Cc_tra2 ^{ts1} <i>in vitro</i> synthesis
P_1440	GAAATTAATACGACTCACTATAGGCCCATATAAAC GCCAGGTGTGTTTTAGAGCTAGAAATAGC	gRNA_Cc_tra2 ^{ts2} <i>in vitro</i> synthesis
P_1753	GAAATTAATACGACTCACTATAGGCAT GCCGCCAGAGTGACGAAGTTTTAGAG CTAGAAATAGC	gRNA_Cc_wp <i>in vitro</i> synthesis
P_369	GCACCGACTCGGTGCCACTTTTTCAA GTTGATAACGGACTAGCCTTATTTAA CTTGCTATTTCTAGCTCTAAAAC	gRNA <i>in vitro</i> synthesis (common reverse primer for all gRNAs)
P_1000	ACCCTGAAGTTCATCTGCACCACCGCAAGCTGCC CGTGCCCTGGCCACCCTCGTGACCACCCTGAGCCA CGGGGTGCAGTGCTTCAGCCGCTACCCCGACCACA TGAAGCAGCAGCACTTCTTCAAGTCCGCCATGCC	repair template for eGFP to BFP conversion
P_1441	TGAGTAATCTACGCGTATGCGTCGATCATCGATTTC CATGCCGGAACATGCGTCCTTGGCTGCTTTAACATC AGCTATATCATCATAATAGATAAAGCAAAGCCAC GAGATCGGCCAGTCTGAAAAAAGAAAAAATAG	repair template for tra2 ^{ts1} implementation
P_1442	AAACGATTTAAATCACATGCACATGCGAAGTATAC CTTGTGTGTCGTCCATATAAACGCCAGGTGTGGAA GTGTGTGGTCTCTGTGTAGTTGAGTAATCTACGCGT ATGCGTCGATCATCGATTTCATGCCGGAACAT	repair template for tra2 ^{ts2} implementation
P_55	TGTGATCGCGCTTCTCGTT	
P_145	ACTTAATCGCCTTGCAGCACATCC	eGFP gene editing genotyping
P_176	AGGCCACCTATTCGTCTTCC	
P_1001	CCTGAAGTTCATCTGCACCACC	Amplification and sequencing of eGFP to BFP repair template
P_1160	GGCATGGCGGACTTG	
P_1401	TGCTTGGTGGTCCGCAAATA	Cc_tra2 ^{ts} genotyping PCR
P_1500	TGTGCATATACTAAAGGCTCTCCC	
P_1504	TACGCTACGAATAACGAATTGG	Ccap Y-specific repetitive elements PCR (karyotyping)
P_1505	GCGTTTAAATATACAAATGTGTG	
P_1532	AGTGAAAACGATTTAAATCACATGCAC	Cc_tra2 ^{ts} genotyping PCR
P_1633	TCCAGTGCAGTTCGGCTTAA	
P_1634	CGGCTTTTACAACGCTTATGTTC	<i>in situ</i> probe Cc_wp

P_1643	TTGAAGAGCGCACTTGCAAC	<i>Cc_wp</i> genotyping multiplex PCR (natural mutation vs WT allele)
P_1644	TTCCCAACAGTGAATCCGG	
P_1657	AAACGCTCTACAGATTGTGGA	
P_1777	TCCAGTGTCTCTACTATGTTGCC	D53 inversion status examination
P_1794	ATCTACCAAATGAGAGAGAGAGCG	
P_1795	TTTTTGAAACCACTTGAACAACGC	
P_1798	TCAGCTAACAGAACATGAATTCCG	

6.2 Methods used or established during these studies

6.2.1 CRISPR/Cas9 gene editing

Here, the most important considerations for CRISPR/Cas9 gene editing experiments in *C. capitata* are summarized. The ‘Practical guide on genome-engineering with CRISPR-Cas9 in the mosquito *Aedes aegypti*’ by Kistler et al. (2015) was used as a starting point for this and is highly recommendable for more detailed information.

Cas9 protein

All experiments were done with recombinant Cas9 protein from *Streptococcus pyogenes*, obtained from PNA Bio Inc (catalog number CP01, Lot number PC16912). The protein was shipped in lyophilized form and subsequently reconstituted to a stock concentration of 1 µg/ µl in 20 mM Hepes, 150 mM KCl, 2% sucrose, and 1 mM DTT (pH 7.5), by adding nuclease-free ddH₂O to the protein pellet. Cas9 protein lyophilizate and ddH₂O were incubated at room temperature for 10 min and gently mixed by tapping to ensure complete dissolution while avoiding foam formation. The dissolved protein was stored at -80°C in single-use aliquots (3.6 µl). A concentration of 360 ng/µl Cas9 protein was used for microinjections of *C. capitata* embryos.

gRNA design and usage

gRNAs were designed using the Geneious software package (Kearse et al., 2012) and subsequently produced by a template-free PCR reaction with Q5 HF polymerase (NEB) followed by *in vitro* transcription (see 6.2.2 for detailed protocol). gRNAs for a gene knock-out via NHEJ were positioned either in an early exon or a conserved domain to ensure an editing event causes a non-functional protein. For ssODN knock-in (HDR), gRNAs were designed to bind as close as possible to the target site, as the likelihood of incorporation decreases rapidly with increasing distance to the target site (0 bp cut-to-mutation distance equals ~ 90-100% mutation incorporation, 10 bp cut-to-mutation distance equals ~ 35-50%; (Paquet et al., 2016; Kwart et al., 2017)). The Geneious tool ‘find CRISPR sites’ was used with the following prerequisites for the search: Target: N(20); PAM: NGG; Activity scoring: Doench 2014; Specificity scoring against an off-target database. For specificity scoring and off-target analysis, the most recent genome version and an algorithm proposed by the Zhang lab (Hsu et al., 2013) were used to score CRISPR sites. Denoted scores are between 0 and 100, with 100 being the highest specificity and lowest off-target activity. If possible, gRNAs without an off-target were chosen for experiments. The on-target activity was scored using the method from Doench et al. (2014). Scores are between 0 and 1, with a higher score representing higher expected activity (Doench et al., 2014). gRNAs with an on-target activity of > 0.1 showed good editing efficiency in my experiments. Produced gRNAs were stored at -80°C and thawed not more than 1-2 times. A concentration of 200 ng/µl was used for injections.

ssODN repair template

Single-stranded DNA was used as a repair template. Short templates (< 200 bp, ssODN) were obtained from Eurofins (‘EXTREmers’), longer templates were produced using the guide-it long ssDNA production kit from Takara (catalog number 632666), according to the manufacturer's protocol.

Microinjections and post-injection treatment

Embryos for microinjections were collected for 30-60 min, dechorionated by immersion in a 1:1 solution of Chlorix (DanKlorix Hygiene Reiniger mit Aktiv Chlor) and demineralized H₂O (3 min, freshly prepared), and afterwards washed with demineralized H₂O. Cas9 protein (360 ng/μl), gRNA (200 ng/μl), KCl (final concentration: 300 mM; include the 150 mM KCl in the Cas9 protein stock solution in the calculation), ddH₂O, and, for HDR experiments, ssODN repair template (200 ng/μl) were mixed (final volume: 10 μl) and incubated for 10 min at 37°C to enable RNP formation. The injection mix was prepared freshly for each injection day and stored on ice until use. 'HDR-injection mixes' tend to clog the needle and are more difficult to inject. Notably, not the protein itself but the mixture with the repair template seems to be the bigger issue, as injection mixes for knock-out experiments (without repair template) are less viscous than mixes for knock-in experiments. It proved to be helpful to use siliconized quartz needles (Q100-70-7.5; O.D.: 1.0 mm, I.d.:0.70; 7.5 cm length; Science Products, Hofheim) with a relatively short tip, instead of the frequently used borosilicated needles (GB100F-10 with filaments, 0.58x1.00x100 mm; Science Products, Hofheim). Needles were drawn out on a Sutter P-2000 laser-based micropipette puller with the following conditions (Heat = heat, Filament = Fil, Velocity = Vel, Delay = Del, Pull = Pull):

Quartz (Q100-70-7.5): Heat 730, Fil 4, Vel 40, Del 125, Pull 130 (program 60)

Borosilicate (GB100F-10): Heat 345, Fil 4, Vel 45, Del 180, Pull 160 (program 66)

The rearing of injected EgII G₀ at 19°C seems to increase the number of adult G₀ survivors (see 3.2). This assumption, however, was not systematically tested yet and would need to be verified in a large-scale study, e.g., the comparison of survival rates of CRISPR-HDR injections targeting a non-essential gene like eGFP at 19°C and 26°C. Hatched larvae were subsequently transferred to 25-26°C.

6.2.2 gRNA synthesis (Protocol: Predrag Kalajdzic)

1. Template-free PCR reaction using Q5 high fidelity polymerase for DNA-template synthesis

The DNA template needed for the subsequent *in vitro* transcription of the gRNA comprises three sequence parts: the T7 promotor, the target-site specific portion, and the Cas9 interacting sequence. The Cas9 interacting part is universal for all gRNAs. Therefore, a ‘common reverse oligonucleotide’ is used to synthesize it (P369, see 6.1.3). The target-specific sequence needs to be specifically designed and fused to the T7 promotor for subsequent transcription. Therefore, this part is ordered as ‘unique forward oligonucleotide’ (e.g., P_986, P_1172, P_1439; see 6.1.3). Oligonucleotides were ordered HPLC purified. Note that the PAM sequence is not part of the oligonucleotide and that the gRNA target sequence ideally should start with ‘GG’, as these are the first two obligate nucleotides transcribed by the T7 RNA polymerase (Kistler et al., 2015). If the target sequence does not start with GG, it should be added to the 5’ end of the 20 bp recognition sequence (Kistler et al., 2015).

PCR reaction setup

5x reaction buffer Q5	20 μ l
10 mM dNTPs (2 mM each)	10 μ l
10 μ M CRISPR_F primer	5 μ l
10 μ M CRISPR_R primer	5 μ l
Q5 polymerase (NEB)	1 μ l
Nuclease free ddH ₂ O	59 μ l
Total	100 μ l

PCR program: 98°C, 30 s; 35x (98°C, 10 s; 58°C, 20 s; 72°C, 20 s); 72°C, 2 min

- Run 2 μ l of the reaction on a 2% agarose gel to verify the size (single band, ~120 bp)
- Purify the remaining 98 μ l according to the manufacturers protocol (use e.g. Zymogen DNA Clean and Concentrator25 and elute the purified PCR product in 30 μ l TE buffer (60°C)).
- Check purity and concentration of the DNA. Typically, I got 150-200 ng/ μ l.

2. *in-vitro* transcription of gRNA using HiScribe T7 High Yield RNA Synthesis Kit (NEB)

IVT reaction setup (0.2 ml reaction tube)

10x reaction buffer	1.5 μ l (0.75X)
T7 RNA polymerase mix	1.5 μ l
NTPs	1.5 μ l each
Template DNA	500 ng
Nuclease free ddH ₂ O	X μ l
Total	20 μ l

incubation: in the thermocycler at 37°C (lid: 105°C), 16 hours

- Subsequently, transfer the RNA to a 1.5 ml reaction tube, add 2 μ l 10xTURBO DNase buffer and 1 μ l TURBO DNase (Invitrogen). Mix gently, incubate 20 min at 37°C
- Add 2 μ l resuspended DNase inactivation reagent, incubate 5 min at room temperature and mix occasionally.
- Centrifuge 1.5 min at 10,000 x g. Transfer RNA to a fresh tube

3. RNA purification via the MEGAclean Kit (Invitrogen, Ambion)

Follow the manufacturers protocol until step 5 (wash with 2 x 500 μ l wash solution). Then follow the procedure below instead of step 6:

- Pre-heat 110 μ l elution solution (per sample) to 70°C
- Place filter cartridge into a new collection tube
- Apply 50 μ l of pre-heated elution solution to the center of the filter. Close the cap and incubate at 70°C for 5 min
- Recover eluted RNA by centrifugation (1 min, 12,000 x g, room temperature)
- Transfer eluted RNA into a new reaction tube, store on ice.
- Repeat the elution procedure
- Store elution separately (Elution 1 and Elution 2) and measure its concentration. Typically, I got an RNA concentration of 1,000-2,000 ng/ μ l in Elution 1, and 200-500 ng/ μ l in Elution 2.
- Run ~200 ng RNA on a 2% agarose gel to verify its size and quality (single band, ~100 bp)
- Store small aliquots at -80°C. Do not thaw the RNA more than 1-2 times.

6.2.3 Non-lethal genotyping of *C. capitata* using single-leg DNA (Protocol: R. A. Aumann)

This protocol is based on a protocol designed for *Drosophila* (Carvalho et al., 2009), and was adapted and modified to facilitate genotyping of *C. capitata* based on genomic DNA extraction of a single leg (non-lethal).

Material

- Buffer A (for 10 ml: 100 μ l TRIS 1 M pH 8.0; 20 μ l EDTA 0.5 M pH 8.0; 50 μ l NaCl 5 M; add ddH₂O up to 10 ml)
- Proteinase K (Roche, 03115887001)
- 2 ml reaction tubes with screw caps, filled with ceramic beads (Lysing matrix D bulk, MP Biomedicals, Cat No. 6540434; beads are filled in the tubes by hand, using a spoon with 90-100 μ l volume)
- Heat block (e.g. Eppendorf ThermoStat)
- Micro scissor (e.g. Hammacher (Solingen), article nr. HSB 530-08, KG115)
- Spring steel tweezers (blunt)
- Vials for keeping single flies (e.g. *Drosophila* breeding container or container for plant tissue culture, 175 ml, clear; closed with a ceaprene stopper, 50 mm), equipped with feed and water (3:1 sugar-yeast mixture, filled in a screwcap of a 15 ml Greiner tube attached to the bottom of the vial, and a 2 ml reaction tube attached to the side, filled with water and a sponge cloth)
- Genotyping must be performed on virgin flies (either freshly eclosed or sexed)

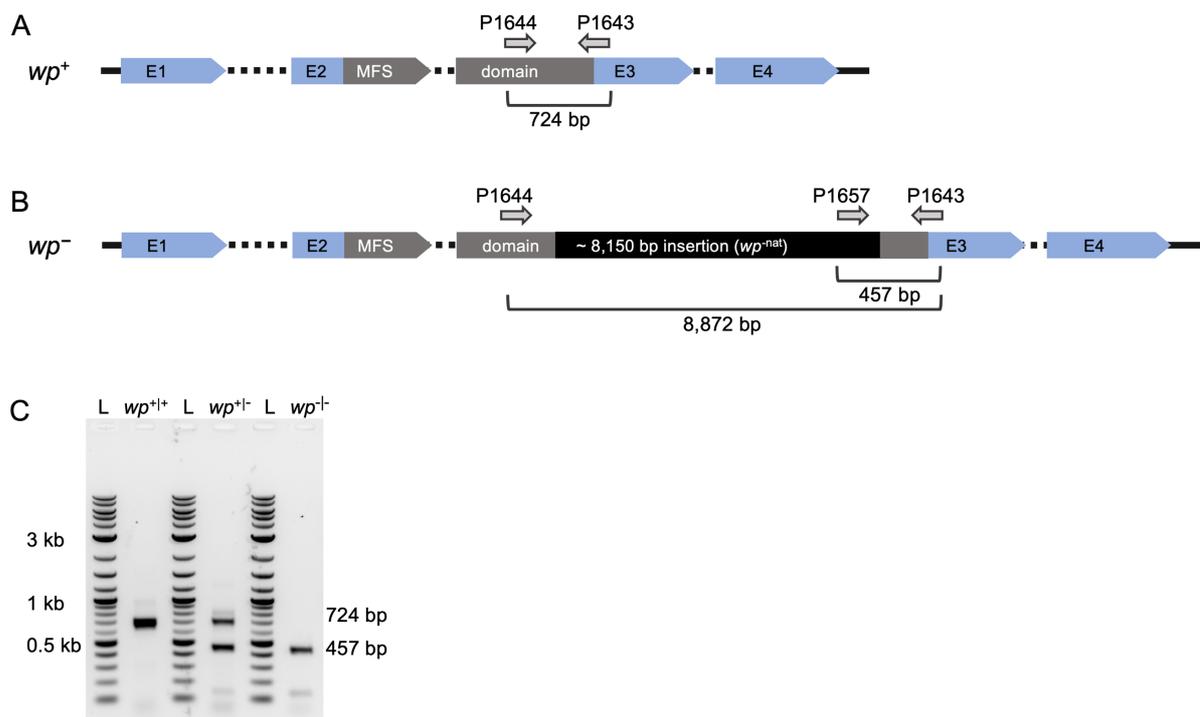
Procedure

- Add 50 μ l buffer A to each 2 ml reaction tube containing ceramic beads and store on ice.
- Carefully cut off a single middle leg of an CO₂-anesthetized fly using micro scissors while holding the leg with spring steel tweezers and place the leg in the reaction tube filled with beads and buffer. Make sure the leg is on the bottom and not attached to the side of the reaction tube. Store it on ice.
- Clean the equipment and proceed until all flies are sampled.
- Sampled flies need to be kept individual, e.g. in equipped *Drosophila* breeding container, until their genotype is assessed. Keeping the flies at cooler temperatures (19°C) proved to be beneficial, as the biological aging is slowed down.
- Homogenize the single legs for 15 sec (6 m/s) using a tissue homogenizer (e.g. Precellys 24), spin down and store on ice.
- Add 28.3 μ l buffer A and 1.7 μ l Protease K to each sample and carefully mix it. Spin down if necessary.
- Incubate at 37°C for 1 h, and subsequently at 98°C for 4 min to stop the reaction.
- Let the samples cool down to room-temperature.
- The solution can be used as template for PCR reaction. Use approximately 3.5 μ l for 25 μ l PCR reaction volume.

6.2.4 PCR assays to distinguish VIENNA^(D53) GSS and WT medflies

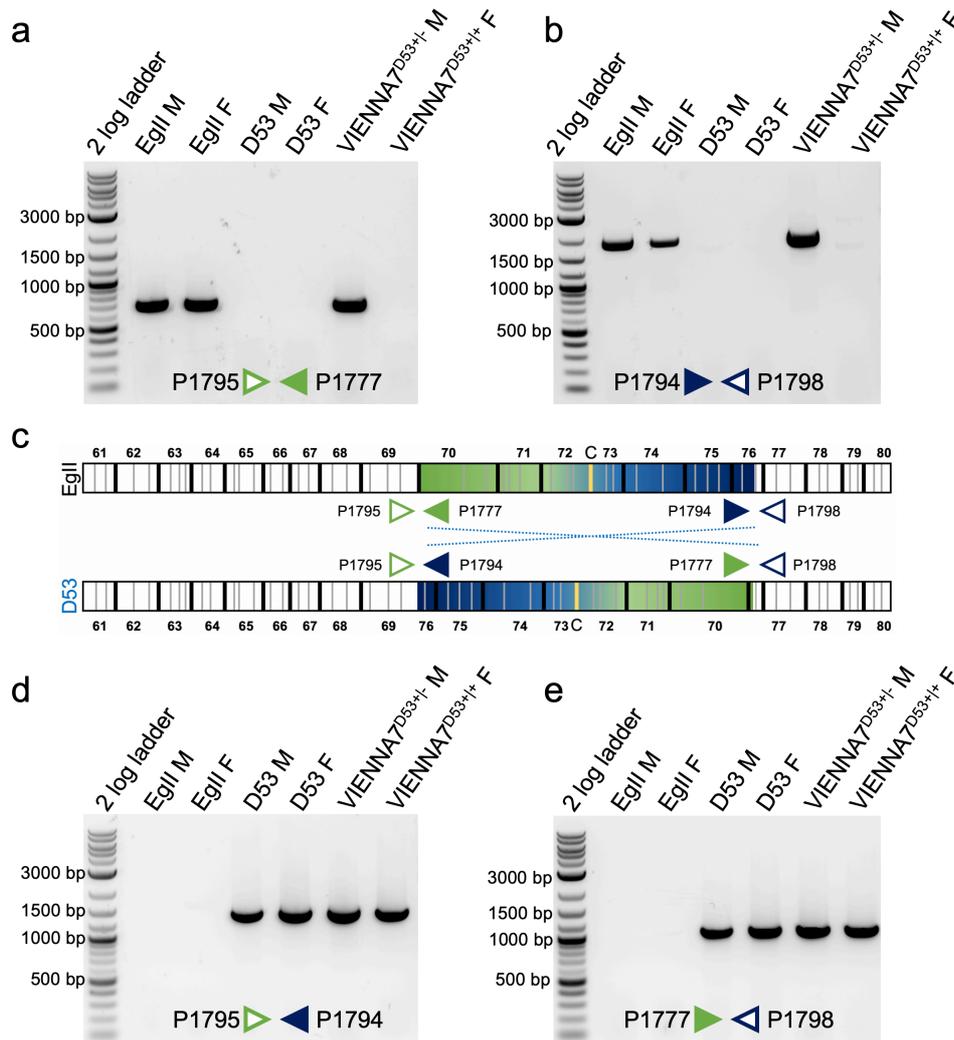
PCR assays were developed to be able to distinguish factory reared GSS flies and WT flies on genomic level, and to characterize eventually surviving progeny thereof, in the case that fertile GSS flies were released by accident (hazard characterization). The first assay is based on the *wp* mutation used in all VIENNA 7 and VIENNA 8 GSS; the second assay is based on the D53 inversion (VIENNA 7/8^{D53+}), which is used in most rearing facilities.

6.2.4.1 *white pupae* gene based multiplex PCR (Protocol: R. A. Aumann)



Presence or absence of the *wp*^{-(nat)} mutation can be detected via multiplex PCR. The *wp* gene structure without (A) and with (B) the naturally occurring mutation, as well as the positions of the three primers (P1644, P1643, P1657) are schematically shown (Exons 1 - 4 including the MFS domain, introns are shown as dashed lines). For the primer pair P1644/P1643, different amplicons are expected for brown (*wp*⁺) and white pupae (*wp*^{-(nat)}), 724 bp and 8,872 bp in size, respectively. The large *wp*^{-(nat)} amplicon can be excluded via PCR settings. A third primer (P1657), specific for the 8,150 bp insertion, can pair with P1643 to amplify 457 bp, but only in the presence of the *wp*^{-(nat)} mutation. Therefore, *wp*^{+/+} (WT) flies show one amplicon at 724 bp, *wp*^{+/-}(*nat*) flies, e.g. GSS males, or WT paired with GSS, show two signals (724 bp and 457 bp), and *wp*^{-(*nat*)/-(*nat*)} (GSS females) one at 457 bp, in a PCR using all three primers (C). PCRs were done using DreamTaq polymerase according to the manufacturer's protocol [95°C, 3 min; 35 cycles of (95°C, 30 s; 56°C, 30 s; 72°C, 1 min); 72°C, 5 min]. All primers were used equimolarly; oligonucleotide sequences are described in 6.1.3.

6.2.4.2 D53 inversion detection assay (Protocol: R. A. Aumann)



Ward et al., 2021

The presence or absence of the D53 inversion can be verified via PCRs spanning the breakpoints. The position of the inversion on chromosome 5 and the primers are schematically shown in (C). EgII is a WT strain without inversion, D53 is homozygous for the inversion. VIENNA 7^{D53+} (V7^{D53+}) GSS males are heterozygous; females are homozygous for the inversion. Primer pairs P1795/P1777 (A) and P1794/P1798 (B) can be used to amplify 690 bp and 1,950 bp amplicons, respectively, if the inversion is not present or heterozygous (wild type status of chromosome 5, e.g., in WT, V7^{D53+/-} male, or WT paired with a V7^{D53+/+} female). In the homozygous presence of the D53 inversion, these primers do not give a signal. Primer pairs P1794/P1795 (D) and P1777/P1798 (E) generate amplicons of 1,152 bp and 1,188 bp, respectively, only in the presence of the D53 inversion (D53, V7^{D53+/-} male, V7^{D53+/+} female, or WT paired with a V7^{D53+/+} female). PCRs were performed using PhusionFlash Polymerase according to the manufacturer's protocol [98°C, 10 s; 30 cycles of (98°C, 1 s; 56°C, 5 s; 72°C, 35 s); 72°C, 1 min]. Primer sequences are described in 6.1.3.

**Der Lebenslauf wurde aus der elektronischen
Version der Arbeit entfernt.**

**The curriculum vitae was removed from the
electronic version of the paper.**

Erklärung gemäß der Promotionsordnung des Fachbereichs 09 vom 07. Juli 2004 § 17 (2)

„Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe.

Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht.

Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.“

Gießen, den 08.03.2021

Roswitha A. Aumann