Insect Biochemistry and Molecular Biology 110 (2019) 34–44

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

An integrated protocol for targeted mutagenesis with CRISPR-Cas9 system in the pea aphid

G. Le Trionnaire[a,*], S. Tanguy[a], S. Hudaverdian[a], F. Gleonnec[a], G. Richard[a], B. Cayrol[b], B. Monsion[b], E. Pichon[b], M. Deshoux[b], C. Webster[b], M. Uzest[b], A. Herpin[c], D. Tagu[b]

[a] UMR 1349 IGEPP (INRA, University Rennes 1, Agrocampus Ouest), Rennes, France
[b] UMR BGPI (University Montpellier, CIRAD, INRA, Montpellier SupAgro), Montpellier, France
[c] INRA LPGP, Rennes, France

A B S T R A C T

CRISPR-Cas9 technology is a very efficient functional analysis tool and has been developed in several insects to edit their genome through injection of eggs with guide RNAs targeting coding sequences of genes of interest. However, its implementation in aphids is more challenging. Aphids are major pests of crops worldwide that alternate during their life cycle between clonality and sexual reproduction. The production of eggs after mating of sexual individuals is a single yearly event and is necessarily triggered by a photoperiod decrease. Fertilized eggs then experience an obligate 3-month diapause period before hatching as new clonal colonies. Taking into consideration these particularities, we developed in the pea aphid *Acyrthosiphon pisum* a step-by-step protocol of targeted mutagenesis based on the microinjection within fertilized eggs of CRISPR-Cas9 components designed for the editing of a cuticular protein gene (*stylin-01*). This protocol includes the following steps: **i**) the photoperiod-triggered induction of sexual morphs (2 months), **ii**) the mating and egg collection step (2 weeks), **iii**) egg microinjection and melanization, **iv**) the 3-month obligate diapause, **v**) the hatching of new lineages from injected eggs (2 weeks) and **vi**) the maintenance of stable lineages (2 weeks). Overall, this 7-month long procedure was applied to three different crosses in order to estimate the impact of the choice of the genetic combination on egg production dynamics by females as well as hatching rates after diapause. Mutation rates within eggs before diapause were estimated at 70–80%. The hatching rate of injected eggs following diapause ranged from 1 to 11% depending on the cross and finally a total of 17 stable lineages were obtained and maintained clonally. Out of these, 6 lineages were mutated at the defined sgRNAs target sites within *stylin-01* coding sequence, either at the two alleles (2 lineages) or at one allele (4 lineages). The final germline transmission rate of the mutations was thus around 35%. Our protocol of an efficient targeted mutagenesis opens the avenue for functional studies through genome editing in aphids.

1. Introduction

The CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9 system has emerged in the last few years as a very powerful genome editing tool allowing the generation of precise and targeted mutational events within the DNA sequence of various metazoans. This technology is derived from a bacterial “adaptive” immune system dedicated to the recognition and the cleavage of invasive viral or plasmid DNA. The unique properties of this mechanism have been used to engineer the “CRISPR-Cas9” system for mediating targeted mutagenesis. This system relies on the use of a single-guide RNA (sgRNA) and a Cas9 nuclease (Doudna and Charpentier, 2014). In short, the sgRNA contains a sequence of 20 nucleotides (nt) corresponding to the genomic region to be targeted, and a universal sequence of 120 nucleotides. The folding capacity of the latter allows the formation of the sgRNA-Cas9 ribonucleoprotein complex. Once delivered to the cells of the organism of interest, the Cas9-sgRNA complex first scans the genome until the targeted sequence is recognized by the sgRNA specific 20-nt motif through RNA-DNA complementarity. This then promotes the nuclease activity of the Cas9 protein at that site, which also depends on the presence on the genomic DNA of a 3-nt Protospacer Adjacent Motif (PAM) at the end of the 20-nt recognition site. After this recognition event, the endonucleolytic activity of the Cas9 generates a double-strand break (DSB) of the DNA that is then repaired by endogenous cellular systems such as the Non Homologous End Joining (NHEJ) or the Homology-Directed Repair (HDR) system. NHEJ is considered to be error-prone since the reparation sometimes generates mistakes such as insertions or deletions (InDels) of one to several nucleotides, thus causing genomic DNA mutations. These InDels can potentially generate frame shift mutations resulting in premature stop codons or altered proteins. CRISPR-Cas9 technology is commonly used in a wide range of organisms including animal and plant models. In insects, such a system has proved to be efficient in models such as *Drosophila* melanogaster (Bassett et al., 2013), *Aedes aegypti* (Kistler et al., 2015), *Bombyx mori* (Wang et al., 2013) or *Tribolium castaneum* (Gilles et al., 2015). Recently, it was also successfully implemented in a wide range of non-model insects ranging from Diptera, Lepidoptera, Orthoptera, Coleoptera, Hymenoptera and other arthropods (for

* Corresponding author.
E-mail address: gael.le-trionnaire@inra.fr (G. Le Trionnaire).

https://doi.org/10.1016/j.ibmb.2019.04.016
Received 8 February 2019; Received in revised form 16 April 2019; Accepted 17 April 2019
Available online 20 April 2019
0965-1748/ © 2019 Elsevier Ltd. All rights reserved.

detailed reviews, see Sun et al., 2017, Cui et al., 2017 and Taning et al., 2017). The main delivery method for this technology in insects (and in most animal models) now relies on the microinjection of a ribonucleoprotein (RNP) complex - containing the sgRNAs and purified Cas9 protein - within single cells or early embryos.

Implementation in non-model insect organisms remains challenging, therefore we aimed to optimize the CRISPR-Cas9 technology in aphids, which are hemipteran phloem-sap feeding insects known as major pests of most crops worldwide. Their life cycle is quite singular since they develop by clonal reproduction (viviparous parthenogenesis) during spring and summer. When summer ends, aphids can perceive photoperiod shortening and integrate this signal towards their developing embryos that in turn switch to a sexual reproductive mode, resulting in the production of males and sexual oviparous females. Sexual individuals then mate and females lay fertilized eggs on plants during the autumnal period. Then these eggs must experience an obligate diapause of three months over winter, before they hatch in the next spring as new parthenogenetic colonies (Le Trionnaire et al., 2013). Aphids display other life history traits (host plant adaptation mechanisms, endosymbiosis, phenotypic plasticity) that allow them to efficiently cope with biotic and abiotic stresses or increase damages on crops (especially as vectors of plant viruses). Deciphering the molecular mechanisms controlling these traits and especially identifying master regulator genes could potentially provide new molecular targets for specific aphid control strategies. Understanding and validating the function of these genes is thus a necessary pre-requisite. So far the peculiarities of the aphid model have made the development of functional analysis tools particularly complex. RNA interference (RNAi) is used to transiently silence the expression of aphid transcripts (Jaubert-Possamai et al., 2007). However, silencing lasts only for a few days in most cases and its efficiency depends on the targeted tissue and rarely exceeds 50% (Mulot et al., 2016). EMS mutagenesis is another alternative to generate stable mutants (Tagu et al., 2014). The availability of an efficient genome editing tool in aphids would thus be a major breakthrough in the aim of understanding gene functions in such an important crop pest.

Taking into consideration all the constraints inherent to the model, a step-by-step protocol for targeted mutagenesis with the CRISPR-Cas9 system was optimized in the pea aphid (*Acyrthosiphon pisum*), one of the model species used commonly in the aphid research community, and for which genomic resources are available (IAGC, 2010 and Legeai et al., 2010). In this study, the candidate gene for developing and validating the efficiency of the method was *stylin-01*, a gene coding for a cuticular protein from the CPR family, which is a prime candidate receptor of noncirculative plant viruses. This protein has recently been identified in the acrostyle, an organ in aphid maxillary stylets containing plant virus receptors across its surface (Webster et al., 2018). In the same study, the authors showed that the viral helper protein P2 of *Cauliflower mosaic virus* (CaMV) and an antibody targeting the surface-exposed peptide of Stylin-01 compete *in vitro* for binding to the acrostyle. Moreover, reducing Stylin-01 expression in the aphid vector *Myzus persicae* impacted its ability to transmit the CaMV. However, transcript levels were only partially reduced, despite numerous attempts made by multiple methods. In addition, a direct P2-Stylin-01 interaction was not observed, due to intrinsic properties of the protein, preventing Stylin-01 from being definitely identified as the receptor of CaMV. *A stylin-01* mutant would undoubtedly help define the precise role of Stylin-01 in noncirculative virus transmission. Here, we will describe in detail the following steps of the protocol: **i)** male and sexual female synchronized photoperiod-triggered induction, **ii)** mating and fertilized eggs synchronization, **iii)** sgRNAs and Cas9 protein microinjection, **iv)** diapause experienced by melanized eggs, **v)** hatching and maintenance of new colonies, **vi)** the mutation detection protocol, and **vii)** the effective germline transmission of mutations within the offspring of hatching colonies. Each critical step in the protocol will then be discussed in the light of our experience. This study represents the first successful genome editing protocol ever developed in aphids to create stable mutant lineages.

2. Material and methods

2.1. Single guide RNAs preparation

2.1.1. In-silico single guide RNAs design

The CRISPOR algorithm (http://crispor.tefor.net/, Concordet and Haeussler, 2018) was used to find targets for single guide RNA synthesis within the *stylin-01* coding sequence. Prior to this computational step, the genomic *stylin-01* DNA sequence from the *A. pisum* clones used in this study were checked to identify putative polymorphic sites. CRISPOR software specifically searches for target sequences of a N^{16-20}-NGG format, N^{16-20} corresponding to a succession of 16–20 nucleotides (N = A, T, G or C) and NGG to the PAM motif necessary for *Streptococcus pyogenes* Cas9 activity. This algorithm can also search within *A. pisum* genome for putative off-target sites (within exonic, intronic or intergenic regions) sharing respectively 0, 1, 2, 3 or 4 mismatches with the initial target sequence. A specificity score is then calculated for all target sequences and a measurement of their uniqueness within the genome is provided. Target sequences (up to 4, if possible) with the maximum specificity score (100 being the highest) and the lowest number of off-target sites were then selected. In our case, the selected sequences also need to begin with GG, since these two nucleotides are necessary for *in vitro* transcription using T7 polymerase (see below). Considering this additional constraint, target sites selection within *stylin-01* coding sequence were restricted to GG-$N^{(14-18)}$-NGG motifs.

2.1.2. In vitro single guide RNA synthesis

For single guide RNA synthesis (Fig. S1), the DR274 plasmid (Plasmid 42250, Addgene) was used as a template. To linearize the plasmid, a digestion with *Bsa1* restriction enzyme (New England Biolabs) was performed at 37 °C for 24 h in a total volume of 100 µL, using 80 U of enzyme and 8 µg of plasmid. The digestion quality was checked by electrophoresis on a 1.5% agarose gel containing SYBR safe (Life Technologies) and quantified with a Quantus Fluorometer (Promega). The linearized plasmid was then used for single guide RNA template preparation with specific primers. The Forward primer of 65 nt consists in a T7 promoter sequence at the 5′-end (20 nt) followed with the target site sequence defined by CRISPOR software (16–20 nt) and the beginning of the universal sequence of the guide RNA at the 3′-end (25 nt). The Reverse primer corresponds to a 23 nt sequence covering the end of the universal guide RNA sequence (see primers Table S1). PCR was subsequently performed using *Taq* polymerase (Promega) with the following conditions: one initial denaturation step of 95 °C for 5 min, 35 cycles of denaturation step at 94 °C for 30 s, one annealing step at 58 °C for 30 s and one elongation step at 72 °C for 20 s, and one final elongation step of 72 °C for 5 min. The primers used for the four RNA guides of *stylin-01* gene are listed in Table S1. PCR products were then purified using the GenElute PCR Clean Up kit (Sigma Aldrich) according to the manufacturer's instructions. Purified PCR products were quality checked on an agarose gel, quantified with Quantus Fluorometer (Promega) and concentrated by vacuum drying, before being used as a template for *in vitro* transcription of single guide RNA. 4 µg of purified PCR product were *in vitro* transcribed with the MAXIscript T7 transcription kit (Ambion) following manufacturer's instructions with an additional DNase digestion step. Single guide RNAs were purified once with acid phenol-chloroform (1:1) and once with chloroform (1:1), before isopropanol precipitation (1:1) at −20 °C. The RNA pellet was then dried using a vacuum desiccator (or dyer) and resuspended in 15 µL of RNase-free water. RNA quality was checked by electrophoresis under RNase-free conditions and the concentration was estimated with a Quantus fluorometer (Promega).

2.1.3. In vitro single guide RNA cleavage efficiency validation

The genomic DNA sequence of *stylin-01* was initially amplified by PCR using specific primers (Table S1) and *Taq* polymerase (Promega) starting with 2 µL of genomic DNA in a 50 µL reaction mix with the following steps: 95 °C for 5 min followed by 45 cycles of 94 °C for 30 s, 58 °C for 30 s and 72 °C for 20 s and a final step at 72 °C for 5 min. The amplification product was checked by electrophoresis on an agarose gel and used as template for an *in vitro* cleavage assay. Briefly, 100 ng of PCR target fragments were mixed with 300 ng of Cas9 protein (CP01, PNA Bio Inc) and 250 ng of each sgRNA individually in NEB buffer 3 containing BSA. The mix was incubated at 37 °C for 1 h. The Cas9 protein was then inactivated at 65 °C for 10 min. A negative control without sgRNA was also performed. The cleavage reaction efficiency was analyzed by electrophoresis on an agarose gel.

2.1.4. Injection mix preparation

A total volume of 40 µL of injection mix was prepared with 1600 ng of each sgRNA corresponding to a final concentration of 40 ng/µL for each sgRNA, 13.3 µg of recombinant Cas9 protein (CP01, PNA Bio Inc) at a final concentration of 333 ng/µL in RNase free water, with aliquots of 4 µL stored at −80 °C until microinjection of fertilized eggs.

2.2. Egg production and microinjection

2.2.1. Sexual morphs production

The ultimate purpose of this part of the protocol is to generate a sufficient number of fertilized eggs for the microinjection steps (see below). To this end, males and females have to be perfectly synchronized in order to optimize their mating once they reach adulthood. The five clones used in this study were initially sampled in eastern France on *Medicago sativa* for L9Ms10, L9Ms13 and L9Ms14, *Trifolium pratens* for L7Tp23 (Nouhaud et al., 2014) and *Genista tinctoria* for Gt2.01 (Peccoud et al., 2015). These clones are maintained in a controlled environment as clonal colonies on *Vicia fabae* plants. A Long Day (LD) photoperiodic regime (16 h of light and 8 h of darkness at 18 °C) allows multiplication by viviparous parthenogenesis. In order to induce the production of sexual morphs, a Short Day (SD) photoperiodic regime (12 h of light and 12 h of darkness at 18 °C) was applied to the different clones used in this study. Two distinct induction protocols were used: i) a so-called “partial induction protocol” (aphids are moved to SD conditions for a limited number of days and then moved back to LD) that allows production of males, and ii) a so-called “complete induction protocol” (aphids are moved to constant SD conditions) that allows production of sexual oviparous females. The partial induction starts exactly one week before the complete induction, allowing males and sexual females to reach adulthood simultaneously and ensures that they are both mature for mating. Females start to lay fertilized eggs a few days after mating (Fig. 1). The following cross combinations were produced: L9Ms10 (Female) x L7Tp23 (Male), L9Ms14 (F) x L9Ms10 (M) and Gt2.01 (F) x L9Ms13 (M).

2.2.1.1. Partial induction for male production. This protocol was applied to L7Tp23, L9Ms10 and L9Ms13 clones (Fig. 1). Initially reared under LD conditions, 20 synchronized third stage larvae (L3) aphids (Generation 1 or G1) of each clone were moved to SD conditions on *V. fabae* plants (2 aphids/plant) for exactly 9 days. In these conditions, adulthood was reached within 5 days and mature adults started to produce offspring 8 days later. For each adult individual, 2–3 L1 (Generation 2 or G2) were kept in order to obtain 40 L1 larvae of the same age. These sexuparae - which are the male-producers - were isolated on plants (1 aphid/plant). After 9 days under SD conditions these G2 individuals were moved back to LD conditions in which they reached adulthood 10 days later. Production of offspring started after 3 days. Each adult was kept on the plant until it produced around 20 individuals (Generation 3 or G3, batch 1) before being moved to another plant to produce around 20 individuals (G3, batch 2). Because of the specificities of this partial induction protocol, batch 1 contained only parthenogenetic females, while batch 2 was mainly composed of males. These newborn G3 males needed another 10 days to reach adulthood. In total, 800 males (40*20) were produced and were ready for the mating step (see below).

2.2.1.2. Complete induction for sexual females production. This protocol was applied to the second clone of the cross and started exactly 7 days after the partial induction of males (Fig. 1). Initially reared under LD conditions, 20 synchronized L3 aphids (Generation 1 or G1) from the L9Ms10, L9Ms14 or Gt2.01 clone were moved to SD conditions on *V. fabae* plants (2 aphids/plant). Since this induction is complete, aphids are never moved back to LD conditions. For each individual, around 4–5 L1 were kept (Generation 2 or G2) to obtain 70 L1 larvae of the same age. These sexuparae individuals are the female-producers. They were isolated (1 aphid/plant) until they reach adulthood after 10 days and started the production of offspring 3 days later. Each adult was kept on the plant until it produced around 20 individuals (Generation 3 or G3, Batch 1) and adults were discarded afterwards. Because of the specificities of this complete induction protocol all individuals were sexual females. These new born individuals needed another 10 days to reach adulthood. In total, 1400 females (70*20) were produced and ready for the mating step (see below).

2.2.2. Cross preparation, mating and egg collection

Thirty adult females and 15 adult males were placed on individual plants for mating, and an average of 6 plants were used each day. The mixing of sexual individuals started on the Wednesday before the week of injection (Fig. 2a). Usually four batches were set up: one on the Wednesday, one on the Thursday, one on the Friday and one on the Saturday. Once put on plants, the first eggs laid by sexual females were observable three days later and the peak of eggs production occurred after 5 days. As a consequence, the 4 successive batches of crosses corresponded to four rounds of injection the next week (from Monday to Thursday). On each day during the week of injections, aphids were placed on a new plant at precisely 9.00 a.m. to allow sexual females to lay eggs for exactly 4 h. Sexual females usually laid their eggs on the plant, on the plastic bag or on the plastic pot (Fig. 2b, c and 2d). Eggs were collected with a fine paintbrush under a magnifying glass and placed on a microscope slide with wet filter paper (Fig. 2e). With this protocol, eggs were collected within a 4 h window so that they were 2 h old on average. After egg collection, aphids were placed back on the plants before being manipulated exactly the same way on the following day.

2.2.3. Microinjection

Twenty eggs were aligned and orientated on a wet filter paper deposited on a microscope slide (Fig. 2f). The filter paper was cut diagonally and the “straight” or ventral side of the egg (opposite from the “curved” or dorsal side) was orientated to be in contact with the paper. The anterior pole of the egg was also distinguishable from the posterior pole under visible light using a dissecting microscope, as the anterior pole appears clearer and the posterior pole contains the bacteriocyte dark-green structure. Once aligned, eggs were injected using the FemtoJet Microinjector device (Eppendorf), and Eppendorf™ Femtotips™ Microinjection Capillary Tips (inside diameter of 0.5 µm and outside diameter of 0.7 µm). Capillaries were filled with the injection mix containing the Cas9 protein and the guide RNAs (see above) using dedicated Eppendorf Microloaders™ (Eppendorf). Eggs were injected at the anterior pole by a single pressure on the foot pedal for 0.3 s. An injection pressure of 500 hPa was applied and a compensation pressure of 50 hPa was set up to avoid liquid or cytoplasm reflux. A micromanipulator was used to quickly shift the capillary from one egg to another. After injection, eggs were kept on the microscope slide used for injection, then directly transferred to a Petri dish containing wet filter paper - to prevent eggs from desiccation - and placed in a climatic

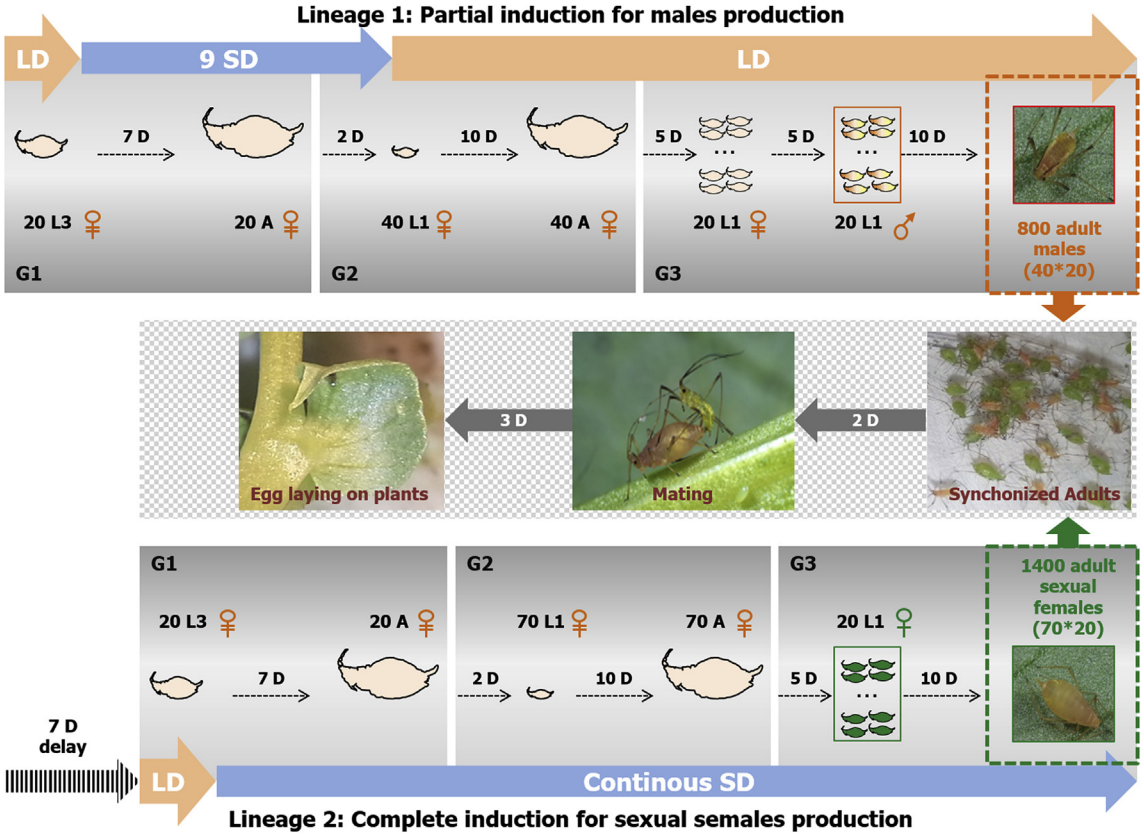

Fig. 1. – Sexual morph induction protocol. Lineage 1 undergoes a partial induction protocol where 20 L3 parthenogenetic aphids, initially maintained under long days photoperiod conditions (LD: 16 h of light, 8 h of night), are moved to short days (SD: 12 h of light, 12 h of night) for exactly 9 days before being moved back to LD conditions. Once the adults reproduce, the first 2 L1 offspring per adult are kept until they reach adulthood. Each of these 40 so-called "sexuparae" then produces a first batch of approximately 20 parthenogenetic females which are discarded, followed by a second batch of around 20 males that are kept for a total of 800 males. Seven days after lineage 1 induction has started, lineage 2 undergoes a complete induction protocol where 20 L3 parthenogenetic aphids under LD are moved to constant SD. These are kept until they are adults, and the first 3–4 L1 produced by each adult are then kept until they reach adulthood. Each of these 70 adult sexuparae produces a first batch of 20 sexual females that are kept for a total of 1400 females. These two delayed protocols allow a synchronized production of males and sexual females that all become adults at the same time, after which they are placed together (15 males for 30 females per plant) for mating. Around 5 days later, the first eggs are laid by females and the microinjection procedure can start. Abbreviations: LX: larval stage X, A: adult, GX: Generation X, D: Day.

chamber at 15 °C (Fig. 2g).

2.2.4. Egg melanization, diapause and hatching

Melanization occurs at the beginning of embryos development where they slowly acquire a black serosal cuticle necessary for cold resistance. Thus complete melanization indicates a non-disruptive injection. Eggs that did not reach this stage were considered damaged by the injection procedure and discarded (Fig. 3a). Melanization of injected eggs was complete within 3–4 days at 15 °C. Melanized eggs were then collected from the microscope slide (Fig. 3b). At this time, eggs were either directly transferred to plants, or treated before transferring them to prevent them from developing fungal or bacterial contamination during the obligate diapause period. This treatment consisted of an initial 5 min wash step in 5% bleach solution followed by three 5 min washes in ultrapure water. Eggs were then dried out on filter paper and carefully deposited with a fine paintbrush on the upper surface of *V. fabae* leaves (Fig. 3c and d). Plants were then kept in a climatic chamber at 4 °C (LD conditions) for exactly 85 days, a period corresponding to the duration of the obligate diapause they experience before hatching. After this period, plants were moved back at 18 °C (LD conditions) to allow diapause to be broken. During two weeks, plants were carefully observed every day under a magnifying glass and hatching aphids were collected. These individuals are the so-called "foundress" and correspond to viviparous parthenogenetic females (Fig. 3e). Each "foundress" was then individualized on a new single plant until they produced offspring by parthenogenesis. These new clonal lineages were then maintained by isolating two individuals (L2 or L3) every generation to a new plant.

2.3. Mutations detection

2.3.1. Mutations detection in injected eggs

Once injections were completed, sub-samples of 20–40 melanized eggs were collected and frozen at −80 °C prior to DNA extraction, which was carried out using the 'salting out' protocol previously described by Sunnucks et al. (1996). DNA from a single egg was extracted and resuspended in 20 µL of ultra-pure sterilized water. The *stylin-01* genomic region (containing the target sites of the four sgRNAs) was PCR amplified starting with genomic DNA (see above). The sizes of the amplicons were compared to the size of the wild-type genomic fragment by electrophoresis on a 1.5% agarose gel to evaluate the presence of insertions or deletions.

2.3.2. Mutations detection in hatched colonies

Once hatched, "foundress" individuals were isolated on a new plant until they produced offspring. As clones, newborns are all genetically identical. For each newly established lineage, a pool of 5 individuals at the 4th instar stage (L4) was frozen in liquid nitrogen. Genomic DNA was extracted using PureLink™ Genomic Plant DNA Purification kit (K183001, ThermoFisher Scientific). A PCR fragment of 1388 bp flanking the targeted region was amplified using Phusion High-Fidelity DNA Polymerase (F530S, ThermoFisher Scientific) with primers BC57F

Fig. 2. – Egg collection and injection procedure. On each day, during the week of injections, males and sexual females - initially put together for mating a few days before - are placed on a new plant at exactly 9.00 a.m. (**a** and **b**). On each plant, 30 sexual females and 15 males are placed together during 4 h so that females can lay eggs on the plant, the plastic bag, the pot or the tutor (**c** and **d**). Precisely 4 h later, eggs are collected with a fine paintbrush and aligned (**e**) on wet filter paper deposited a microscope slide (20 per slide). The ventral face (V) of the egg is placed in contact with the paper and the injection is performed on the dorsal face (D) at the anterior pole (A), which is opposite of the posterior pole (P) defined as the pole where the bacteriocyte is observable (**f**). Once injection is completed slides are placed into petri dishes containing wet filter paper in climatic chambers at 15 °C (**g**).

5′-AGACGCATCCACAATCACAA-3′ and BC57R 5′-CGGATGATCATTTC GAGGAG-3'. PCR cycling conditions included a denaturation step at 98 °C for 30 s, 35 cycles of 98 °C for 10 s, 64 °C for 30s, 72 °C for 45s followed by a final extension at 72 °C for 5 min. To analyze the genomic DNA sequence from the 2 alleles of the *stylin-01* gene, blunt-ended PCR products were cloned in pGGz plasmid (derived from pDonR-Zeo) and transformed in *E. coli* DH5 strain. Four to 20 colonies were selected by colony PCR (Sandhu et al., 1989) using GoTaq G2 polymerase (M7841, Promega, Corporation, Madison, WI, USA) with primers 13F 5′-AACA GGGGCTGGTGAAATG-3′ and 14R 5′-GGCCATATCGGTGGTCATC-3'. PCR amplifications included a denaturation step at 95 °C for 2 min, 35 cycles of 95 °C for 30 s, 60 °C for 30s, 72 °C for 100 s followed by a final extension at 72 °C for 5 min. Selected colonies were cultivated overnight, their plasmid DNA extracted using the Wizard Plus SV Minipreps DNA purification System (A1460, Promega, Corporation, Madison, WI, USA) and sequenced using the same primers (GENEWIZ, South

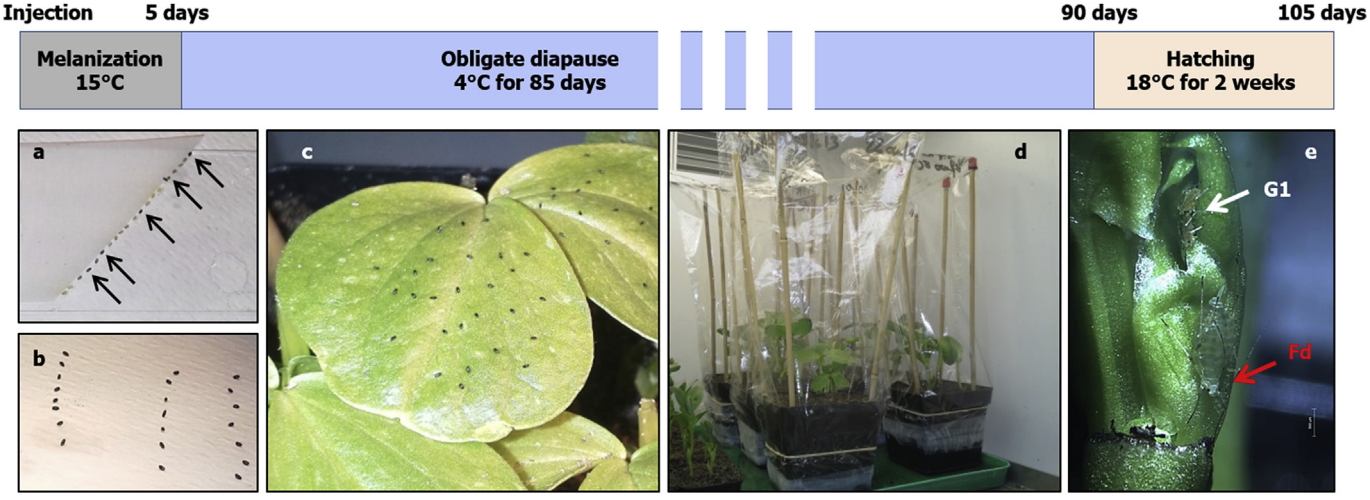

Fig. 3. – Diapause and hatching steps. Eggs that were not damaged by the injection usually melanized within 5 days (**a**). Melanized eggs are then collected directly from the microscope slide and placed on the upper surface of the leaves of a new plant (**b** and **c**). Plants are then covered with a plastic bag and placed in a climatic room for 85 days at 4 °C for the obligate diapause period (**d**). Once diapause is completed plants are moved back to 18 °C to allow egg hatching. The individual that hatches is called the "Foundress" (Fd) and each of those is placed individually on a new plant until it becomes an adult and start producing an offspring (G1) by parthenogenesis (**e**).

Table 1
Comparison of the egg production level and hatching rate of 3 different crosses.

Female	Male	Egg production	Egg treatment	Hatching rate
L9Ms10	L7Tp23	High (at least 10 eggs/ female)	Control	36%
L9Ms10	L7Tp23		H2O + Bleach	9%
L9Ms14	L9Ms10	Low (less than 5 eggs/ female)	Control	35%
L9Ms14	L9Ms10		H2O + Bleach	28%
Gt2.01	L9Ms13	High (at least 10 eggs/ female)	Control	77%
Gt2.01	L9Ms13		H2O + Bleach	33%

Plainfield, NJ, USA). Sequence alignment steps were performed using the Geneious software (http://www.geneious.com).

3. Results

3.1. Evaluation of egg production dynamics and hatching rate in three genetic backgrounds

Before applying CRISPR-Cas9 mutagenesis to the pea aphid, it was necessary to select lineages that are able to respond to the photoperiod shortening signal to produce sexual individuals. Three combinations of crosses were tested in this study: L9Ms10 (Female) x L7Tp23 (Male), L9Ms14 (F) x L9Ms10 (M) and L9Ms13 (F) x Gt2.01 (M). Two key parameters were considered for these crosses: **i)** the number of fertilized eggs laid by sexual females after mating and **ii)** egg hatching success after the 85 days of obligate diapause. Females usually start to lay eggs 3 days after mating (d.a.m) for a total period of 6 days with an optimum for eggs production at 6 d.a.m. Although the number of eggs laid by single females was not precisely recorded (Table 1), the females from crosses L9Ms10xL7Tp23 and L9Ms13xGt2.01 produced higher numbers of fertilized eggs compared to the females from the L9Ms14xL9Ms10 cross (at least 10 per female for the two first crosses, and less than 5 per female for the third one). Thirty six and 35% of the untreated eggs hatched and produced “foundress” lineages for L9Ms10xL7Tp23 and L9Ms14xL9Ms10 crosses, respectively, whereas 77% of eggs hatched in the case of Gt2.01xL9Ms13 cross (Table 1). Hatching rates were reduced for bleach-treated eggs, ranging from 9% (L9Ms10xL7Tp23), 28% (L9Ms14xL9Ms10) to 33% (Gt2.01xL9Ms13), suggesting that this treatment affects embryos viability and should be removed from the protocol (Table 1). Taken together, our results show that the Gt2.01xL9Ms13 cross is the most efficient for producing sufficient numbers of eggs for the microinjection step and post-diapause lineage selection.

3.2. Assessment of single guide RNAs in vitro cleavage efficiency

CRISPOR algorithm found four target sites for single guide RNAs synthesis fulfilling the GG-N^{14-18}-NGG motif constraint within the *stylin-01* coding sequence. The four selected sequences do not share perfect identity with other homologous regions within the pea aphid genome: the single guide RNAs synthesized to target these regions (namely sg1, sg2, sg3 and sg4) were therefore unlikely to generate off-target cleavages. Briefly, *stylin-01* genomic DNA is composed of three exons. Sg1 target site is located within the second exon and the three others are distributed within the last exon (Fig. 4). These four sgRNAs were tested for their ability to cleave the *stylin-01* genomic DNA sequence *in vitro* (Fig. 5). Electrophoresis profiles showed two bands corresponding to the expected cleavage fragments when sg1 and sg3 were present in the reaction, suggesting a high *in vitro* binding efficiency for these guides. Expected cleavage fragments were also observed for sg2 and sg4. However, a band corresponding to uncleaved PCR fragment remained, suggesting an effective but reduced efficiency of these guides *in vitro*. Altogether, this experiment ensured that all the designed sgRNAs were functional *in vitro*.

3.3. Effect of the microinjection on egg survival

In this study, three independent series of microinjections were performed on eggs originating from three different crosses. For each experiment, eggs were injected with two different mixtures: the first one comprising the 4 guide RNAs (sg1, sg2, sg3 and sg4) and the Cas9 protein; and the second one comprising only one guide RNA, sg4, and the Cas9 protein. Overall, 1698 eggs were injected in 5 days for the first cross (L9Ms10xL7Tp23), 412 eggs in 4 days for the second cross (L9Ms14xL9Ms10) and 527 eggs in 2 days for the third cross (Gt2.01xL9Ms13) (Table 2). The impact of microinjection on egg survival was estimated with the proportion of melanized eggs reported to the total number of injected eggs, considering melanization as a proxy measurement of survival. Egg survival rate was around 20%, 31% and 44.4% for the first, second, and third cross, respectively. These differences in egg survival might either reflect different embryo fitness for the three crosses, or an experimenter effect due to manual microinjection. The survival rate of microinjected eggs could thus exceed 45% under optimal conditions.

3.4. Estimation of the in ovo mutation rate

To rapidly estimate CRISPR-Cas9 efficiency after microinjection, up to 40 melanized eggs injected with the 4 sgRNAs/Cas9 mixture were collected within the pool of eggs kept for diapause. For L9Ms10xL7Tp23 cross, the PCR amplification was efficient for 36 eggs out of the 40 selected (Fig. 6). A single band corresponding to the wild type amplicon was observed for 8 eggs. Although this result may reflect the rate of “non-mutated” eggs, we cannot totally exclude the presence of small insertion/deletion events at the four target sites than cannot be detected with the PCR protocol used. For the 28 remaining eggs, the electrophoresis profiles revealed 1, 2 or even 3 bands with a size lower than the one expected for the wild type PCR fragment, either found alone or together with the wild type amplicon. These short PCR fragments probably arose from the combined action of at least 2 guide RNAs triggering large deletions in *stylin-01* genomic DNA sequence. In addition, small amplicons of various sizes were observed in the eggs, indicating that large re-arrangements occurred between different combinations of guide RNAs. These 28 eggs were thus assigned as “mutated”, bringing the mutation rate in the first experiment to at least 77.8%. Eighteen eggs were examined as part of the third experiment (Gt2.01xL9Ms13 cross), out of which 12 gave a PCR result. Four samples were assigned as “non-mutated” and 8 samples as “mutated”, indicating a mutation rate of at least 66.7%. These *in ovo* analyzes confirmed the efficiency of the CRISPR-Cas9 system in mutating aphid eggs at early stages post microinjection. They also showed that the injection of several guide RNAs could generate various combinations of large deletions within the genomic sequence of *sylin-01* without knowing - at this stage - if the observed mutations arose from somatic or germline cells.

3.5. Hatching and stable lineages establishment rate from injected eggs

After the obligate diapause period, embryos have completed their development and hatch as the so-called “foundress” individuals that initiate new clonal colonies. For the three cross combinations tested in this study hatching rates of injected eggs were very low (Table 2) and varied from 1.1% for L9Ms10xL7Tp23 cross (3 hatched eggs/274 eggs injected with sg 4 or with the four sgRNAs), to 4.9% for L9Ms14xL9Ms10 cross (5/101), and 11.7% for Gt2.01xL9Ms13 cross (23/196). These hatching rates were much lower than the ones observed for non-injected eggs in the three genetic backgrounds (36, 35 and 77% respectively, see Table 1). This suggests a potential damage occurring during the injection procedure or a toxic effect of the

>stylin-01 (ACYPI009006)

ATGCAGgtaagagtctctcgaagtttcgagaatggattgttattcgttataccggcgaggtaaccgtgctcacacgtttcacgcttttgcacag**GTCACTTTTGCCGTATCGTCGTTGCT**
GTTAGCTGTCGTCGCCGTCAGCGCGTACCCCGCATCACTGAACCCGGAATCCAGAGCCGCCATCT**TGG**TCCAAGATTCAGCACCCAACGCCGATGGATCATTCAAGAACAAgtatgcaca
attatttttcttaaaaaactttgaggtgaatttttttaagctcgctagtctcatgacgtttatgtacctacgcatttatcgatctcatcaagtatagcctcatcccccgacgtgaataatgt
tgaacgaaatctttgcgcacacttctctcggtcgggtcgacattatacttgttgcctgtagttatacgttatactattaacgcgtttcggtgctaataatagacgatttttcgtccagta
tctatattattattattatcattattttgaaaacgaatgcggcagcaataactacgacaatagttgcaatgttttttttttttactattttacgaaattattgttttctagtaaaatatt
gtaatattataacacgtaacatttcaaagttaaaaaactgaaaacctcaatttaaccgggtggcgttggtgtataaaatatcactgaattacgctcacggtagaccgcatgttgagcggt
ctacttataggttgaagaattcatcacttagtttaaatatttagagtcattagccaaccatggaccatctcaatagacggtttttcgttttagcttctatttttactataattatagttt
taataattaaaaaaagctcactcctatcggcttattcctaaccacattatattattctcgttgcag**TTTCCAAACCGAAAACGGAATCAAACAAGAATCAGTCGGATACTTGAAGGCTGG**
CCCAGAAGGACCCGTAGCTGTGTTCCAGGGAGCTTCTGCGTACGTCGCCCCAGACGGCCAGA**CCA**TCCAAATCGGATACATCGCC**GACGAGAACGGTTACCAGCCGTACGGCGCTCATTT**
GCCCACTCCACCACCAATCCCAGCTGAGATCCAAGAGTCGCTCAGATACCTCGCCTCTCTGCCCAGCACCC**CCGAACCAAAATACCAGTAA**

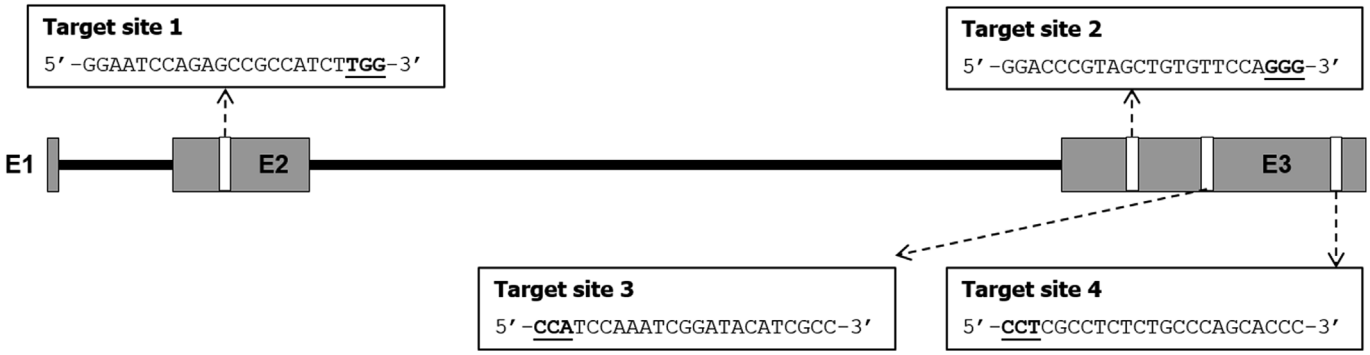

Fig. 4. – *Stylin-01* gene model and guide RNAs target sites. *Stylin-01* gene comprises 3 exons (E1, E2 and E3) and 2 introns for a genomic DNA length of 1.35 kb. CRISPOR software was used to identify putative target sites (in grey) comprising NGG motifs (underlined) at the 3′ end to allow potential DNA cleavage by sgRNA-Cas9 complex. One target site was predicted in exon 2, and 3 target sites in exon 3.

sgRNAs/Cas9 mix impacting embryos survival and hatching. These results also mirror the differences in hatching rates for eggs from the three genetic backgrounds (Table 1), the Gt2.01xL9Ms13 cross showing the highest rate compared with the two other crosses. Altogether, 31 "foundress" individuals out of the 571 injected and melanized eggs were obtained in the three experiments. These 31 "foundress" were installed individually on plants to allow their complete developmental to adulthood. Seventeen "foundress" aphids gave offspring, producing by parthenogenesis genetically identical individuals, leading to the establishment of stable lineages subsequently maintained as clonal colonies. Finally, 55% of the hatched "foundresses" ended up as stable linages (Table 2). A potential effect of mosaicism in the "foundress" might in part explain why 14 individuals could not produce offspring. The final rate of stable lineages setting is therefore closed to 3% (17/571) in the three combined genetic backgrounds.

3.6. Germline transmission rate estimation in stable lineages

All the individuals within the progeny of "foundress" aphids are genetically identical since they arise from diploid oocytes produced clonally. They can potentially originate from one or several germline nuclei within the syncytial early embryo. If only one germline nucleus has been targeted by the CRISPR-Cas9 system within the early-fertilized egg then all the progeny will share the same mutations (one or two different mutated alleles). If there are several germline nuclei at the time of injection, it is possible to detect within the progeny individuals displaying different mutation patterns (more than two mutated alleles). To detect mutated lineages, DNA was extracted from pooled individuals for the 17 stable lineages to amplify the *stylin-01* genomic DNA sequence. PCR fragments were then cloned in a plasmid and sequenced to discriminate between the two parental alleles. A total of 4 lineages (L2, L4, L7 and L15) originating from eggs injected with the four guide RNAs (sg1, sg2, sg3 and sg4) and 2 lineages (L14 and L16) originating from eggs injected with a single guide RNA (sg4) were mutated in the *stylin-01* gene (Table 2). In more details (Fig. 7), L2 lineage displayed multiple re-arrangements in the two mutated alleles and at the four sites targeted by sgRNAs: allele 1 showed a 10 nt deletion at target site 1, a 3 nt deletion at target site 2, a 3 nt insertion at target site 3 and a 13 nt insertion/6 nt deletion at target site 4, while allele 2 showed a large 380 nt deletion at target site 1, a 10 nt insertion/2 nt deletion at target site 2, a 2 nt insertion at target site 3, and a 7 nt deletion at target site 4. In L4, L7 and L15 lineages, insertions or deletions of a few nucleotides were observed at the different target sites for 1 of the 2 alleles. Unexpectedly for these three lineages, only one parental allele was

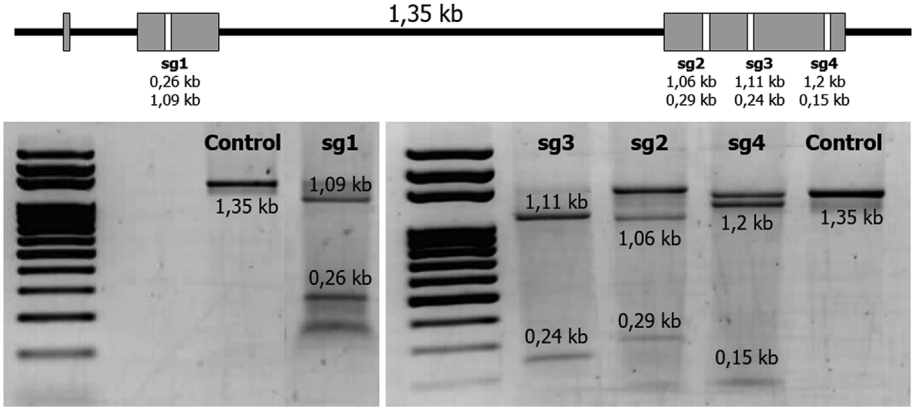

Fig. 5. sgRNAs *in vitro* cleavage assay. In order to test the ability of the 4 synthesized sgRNAs to efficiently cleave their predicted target sequence *in vitro* a PCR product from *stylin-01* genomic DNA (1.35 kb) was mixed with each sgRNA individually and Cas9 protein in its buffer. If the sgRNA actually targets its predicted sequence and slices the DNA fragment, two cleavage fragments will be observed on the agarose gel. The size of expected cleaved fragments are indicated on top of the gel for each sgRNA. The size of the bands that can be observed at the end of the reaction either for the control or for the 4 tested sgRNA are indicated on each lane of the gel.

Table 2
CRISPR-Cas9 mutagenesis of *stylin 1* gene in three different genetic backgrounds.

Cross	sgRNAs combination	*Foundress generation*			*Foundress offspring*	
		Injected eggs	Melanized eggs (%)	Hatched eggs (%)	Stable lineages (%)	Mutated lineages (%)
L9Ms10 (F) x L7Tp23 (M)	sg1-sg2-sg3-sg4	1059	199/1059 (18.8)	3/159 (1.9)	1/3 (33.3)	0/1 (0)
	sg4	639	140/639 (21.9)	0/115 (0)	0 (0)	0 (0)
L9Ms14 (F) x L9Ms10 (M)	sg1-sg2-sg3-sg4	257	89/257 (34.6)	4/74 (5.4)	1/4 (25)	1/1 (100)
	sg4	155	39/155 (25.2)	1/27 (3.7)	1/1 (100)	1/1 (100)
Gt2.01 (F) x L9Ms13 (M)	sg1-sg2-sg3-sg4	286	119/286 (41.6)	17/101 (16.8)	9/17 (52.9)	3/9 (33.3%)
	sg4	241	115/241 (47.7)	6/95 (6.3)	5/6 (83.3)	1/5 (25)

detected by cloning, none of the 20 sequences analyzed for each lineage allowed the detection of the second parental allele. For lineages originating from eggs injected with a single guide RNA sg4, L14 lineage showed re-arrangements on both alleles at the target site: a 3 nt insertion/2 nt deletion on allele 1 and a 6 nt deletion on allele 2. L16 lineage displayed a 7 nt deletion at target site 4 on allele 1, the other parental allele could not be detected in the set of sequenced clones. Since lineages were maintained clonally, the same protocol was applied several generations later to sequence the 6 lineages and perfectly identical mutational patterns as those previously described were observed (data not shown). Since 6 out the 17 established lineages were stably mutated, the germline transmission rate of the mutations was close to 35% (Table 2). These results thus demonstrate that the CRISPR-Cas9 protocol described here is fully functional in the pea aphid and efficient to generate stable mutant lineages on the two targeted alleles of *stylin-01* gene.

4. Discussion

4.1. Difficulties encountered in the aphid model to develop CRISPR-Cas9 mutagenesis

Alternation of reproductive mode as well as the obligate diapause fertilized eggs must experience are two specificities of the aphid life cycle that have caused difficulties in developing CRISPR-Cas9 mutagenesis in aphids. Direct injection within viviparous parthenogenetic females could be an alternative to egg injection. However, in a recent study, Jamison et al. (2018) tried to deliver sgRNA-Cas9 RNPs complexes designed to target the *tor* gene - a cuticle pigmentation gene - in viviparous aphid ovaries by electroporation. The analysis did not indicate any mutation (neither in the soma nor in the germline) of the targeted gene in the offspring of the treated individuals. It is indeed very unlikely that both components could efficiently reach germline cells that are embedded between several layers of tissues within the ovaries. Thus, fertilized eggs seem to be the most rational starting point and therefore the production and the collection of a significant number of eggs just after being laid by sexual females is the first critical point of the protocol. This implies the three main parameters we have first optimized in this protocol: **i)** the induction of sexual morphs, **ii)** the synchrony of males and females production and **iii)** a large number (1000s) of manipulated individuals to produce a sufficient number of eggs. Because diapause cannot be avoided, the *in ovo* PCR test is a proxy to estimate the success of the mutagenesis before eggs experience the 3 months of slow development. The second critical point of the protocol corresponds to the *a priori* choice of the cross that conditions genetic compatibility and by extension hatching rates and fitness of stable lineages. The main challenge of the aphid model in the view of developing targeted mutagenesis thus relies on the complexity of the biology of this insect. We discuss in the following paragraphs all the parameters that have been optimized to allow the production of stable mutant lineages in the pea aphid.

4.2. Lineages choice for sexual morphs induction, mating and hatching success

In order to produce sexual morphs, the two different lineages chosen for the cross must be able to fully respond to photoperiod shortening and alternate their reproductive mode. The implementation of two distinct induction protocols for the two lineages - one favoring a late production of males and the second an early production of oviparous females - ensures that sexual individuals become adult at the same period and are synchronously mature for mating. Self crosses must be

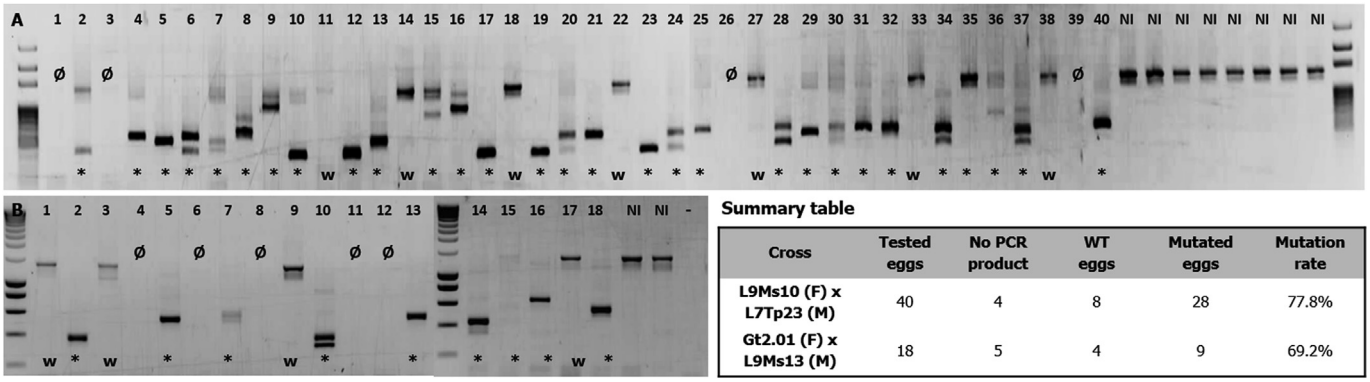

Fig. 6. – *In ovo* mutation rates estimation. For two different crosses (**A**: L9Ms10 x L7Tp23 and **B**: Gt2.01 x L9Ms13) eggs were injected with a sg1-sg2-sg3-sg4 and Cas9 protein mix. Respectively 40 and 18 melanized eggs from these two experiments were then collected for DNA extraction and for subsequent *stylin-01* genomic DNA region amplification. PCR products were loaded on a gel in order to identify deletion events between the different sgRNAs target sites, revealed by the presence of amplicons having a different size than the wild-type expected size. Eggs are classified into three different classes: wild-type eggs (**w**), mutated eggs showing deletion events (*) and eggs for which PCR did not work (**Ø**). In the summary table the ratio [mutated eggs/PCR positive eggs] gives an estimation of the *in ovo* mutation rate.

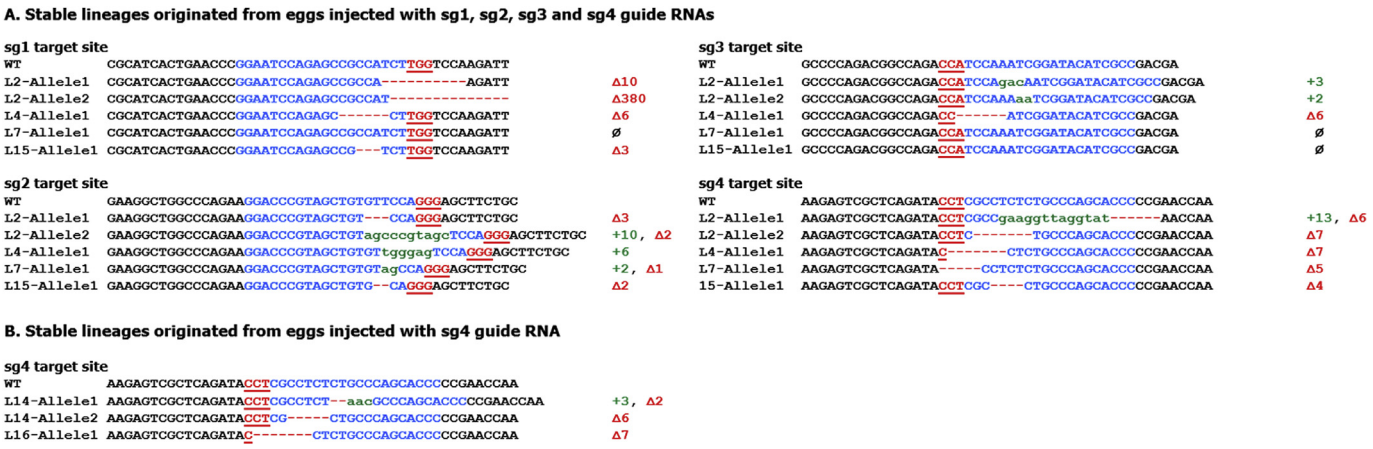

Fig. 7. – Sequences of the mutations detected in lineages edited with CRISPR-Cas9. Both *stylin-01* alleles from all the lineages originated from eggs injected with a combination of sg1, sg2, sg3 and sg4 (A) or only sg4 (B) were sequenced. For the first sgRNAs combination (A), four lineages (L2, L4, L7 and L15) were stably mutated whereas for the second sgRNAs combination (B) two lineages were stably mutated. For each sgRNA the target site change relative to the wild type sequence for each allele is indicated: insertion (+) or deletion (**Δ**). The 3 nucleotides corresponding to the PAM motif are underlined.

avoided since most of the time they are unfit and usually result in very low hatching rates after diapause. In the current study, the three different cross combinations tested resulted in variable egg production rates thus revealing different genetic compatibility levels between selected lineages. Nevertheless, the production of numerous synchronized sexual females (1400) and males (800) is a way to overcome this issue, and to produce the hundreds of eggs needed for further microinjection. Once melanized, eggs must experience a 3-month period of diapause at low temperature. The process of diapause in the pea aphid – and in many other insects - is a slowing but not a cessation of development (Shingleton et al., 2003). The early stages of embryogenesis progress at a temperature-independent rate whereas later stages of embryogenesis progress at a temperature-dependent rate. However embryos maintained at very high temperatures in the very early stages show severe developmental abnormalities. Diapause is thus an incompressible event that cannot be artificially overcome. Hatching is promoted by placing aphids at 18 °C during a 2-weeks period. The embryos originating from the three cross combinations tested in this study displayed various hatching rates from 35 to nearly 80%. The hatching rate, in our hands, corresponds to the most robust indicator of the genetic compatibility between two lineages selected for a cross. It is thus important, when starting to set up a targeted mutagenesis protocol, to optimize the choice of the two lineages that will allow reaching the necessary scale of the experiment.

4.3. Egg injection procedure: timing and effect on melanization and post-diapause hatching

Once the mating of adult individuals is completed, sexual females start to deposit the first fertilized eggs on plants three days later. Females then continue to lay eggs for approximately one week with a peak in production three days after the first laid ones. Every day of the week of microinjection it is mandatory to repeat the same routine: individuals are placed every morning on a new plant and eggs are collected exactly 4 h later. The age of the eggs can thus be estimated between 0 and 4 h after egg laying (hAEL). Lin et al. (2014) performed a detailed cytological observation of the 10 first cycles of nuclei division of pea aphid early embryos originating from fertilized eggs. The authors estimated that these 10 cycles are completed within 16 h and that the 5th cycle of division is reached after approximately 6 h. At that time, the embryo contains 16 nuclei in a syncytial structure. The authors also performed in situ hybridization of ApVas1 mRNA - a germline marker - within embryos. Interestingly, they could not observe the germ plasm after the 3rd nuclei division, which means that none of the eight observable nuclei are predetermined to be a germline nucleus. After the 4th division, the germ plasm is detectable which means that some nuclei will give rise to primordium germ cells. Whether the germline was of a single or multiple origin was not assessed in this study. It is known in drosophila that germ cells originate from three founding nuclei in the syncytial blastoderm embryo and continue amplification before moving into polar buds that give rise to the primordium germ cells (Campos-Ortega and Hartenstein, 2013 and Lindsley et al., 2016). Therefore in the case of the pea aphid, injecting eggs 4 h maximum after laying guarantees that they contain between 8 and 16 nuclei in a syncytial structure, but we do not know how many nuclei display a germline origin. CRISPR-Cas9 system could thus potentially generate more than two different mutated alleles within the offspring of the “foundress” individual. Nevertheless, we could only observed at most two mutated alleles within the offspring of the six mutated “foundress” suggesting that in these specific cases only one germline nucleus had been targeted. In a protocol aiming at performing germline transformation of various species of mosquitoes (including *A. gambiae* and *A. aegypti*) embryos are collected right after oviposition for 15 min then placed for 30 min under mosquito rearing specific conditions to allow their maturation/melanization. At the time of injection, embryos are thus 30–45 min old (Lobo et al., 2006). In *D. melanogaster* injections are also performed around 30 min after embryos collection which corresponds to the syncytial blastoderm stage (Kiehart et al., 2007). In both cases injections take place at the very early stages of embryogenesis to reduce mosaicism levels. As stated above, the early steps of aphid embryogenesis are slower so that a 0–4 h post laying window is comparable to a 30–45 min post-oviposition window in drosophila or mosquitoes. Another feature of aphid eggs is that bacteriocyte is observable at the early steps of embryogenesis at the posterior pole of the egg and is situated close to the future germline nuclei. To avoid any bacteriocyte damage the capillary was inserted at the opposite pole to allow a diffusion of the CRISPR-Cas9 components within the entire syncytium. Nevertheless, injecting those reagents at the “bacteriocyte pole” could be a way to preferentially reach germline nuclei. It has to be noted that given the large scale of these experiment it was not possible to test different sgRNAs concentrations or compare the impact of the injection of a Cas9 mRNA versus the Cas9 protein. These parameters had already been tested in the mosquito injection protocol (Kistler et al., 2015) so that we decided to directly inject the recombinant Cas9 protein with several sgRNAs at optimized concentrations. After the microinjection step, melanization is a first estimation of egg survival. Only eggs that had completed this step were conserved and placed on plants at low temperature. In our study, melanization rate ranged from 20 to 45%

depending on the experiment. In mosquitoes, hatching occurs between 3 and 5 days after microinjection (Lobo et al., 2006) whereas it happens between 2 and 3 days in drosophila (Kiehart et al., 2007). In the case of aphids, due to the obligate diapause, hatching cannot be observed before 85 days post-injection. For further comparisons hatching rates were estimated starting from melanized rather than injected eggs. As mentioned above, hatching rate is highly dependent on the genetic compatibility between the two lineages that are crossed. Nevertheless, the comparison of hatching rate of non-injected versus injected eggs mirrors the deleterious effect – including cytoplasm leakage, membrane deterioration and probably zygotic maternal mRNAs gradient perturbation - that the injection procedure can have on embryos fitness. Embryos survival rate after injection is estimated at 15–20% in *D. melanogaster* (Bassett et al., 2013) and approaches 10% in *A. aegypti* (Kistler et al., 2015). When considering the best cross combination (Gt2.01xL9Ms13), it appears that a survival rate of 11.7% (23 hatched lineages out of 195 melanized eggs) is quite similar to what has been observed in mosquitoes or drosophila. The individuals that hatch from eggs (also called foundress) then produce genetically identical offspring by clonal reproduction. Considering the peculiarity of clonal reproduction in aphids, if the germline nucleus was reached by the CRISPR-Cas9 system at the time of microinjection, the mutation would be fixed and all offspring will be identically mutated (see below). On the contrary, if only somatic cells were affected, then the offspring will not be mutated and somatic mutations would potentially affect the fitness of foundress individuals. This is illustrated by the fact that not all the hatched lineages gave birth to viable offspring. Indeed for the last experiment 14 out of the 23 hatched lineages gave offspring: these are referred as to stable lineages. The final stable lineages rate is thus of 7.1% (14 out of 195 melanized eggs).

4.4. *In vitro and in ovo guide efficiency validation: two necessary check points before diapause*

As mentioned above, the whole experiment requires 7 months for completion: 8 weeks for the induction of sexual morphs, 2 weeks for microinjection/melanization, 12 weeks of diapause, 2 weeks for hatching and another 4 weeks for the set up and maintenance of clonal lineages. Considering the complexity of the protocol, it is therefore crucial to verify *a priori* the cleavage efficiency of the sgRNA/Cas9 mixture injected within the egg. Indeed sgRNAs design and prediction mainly rely on the likeliness of the RNA sequence to anneal its target DNA sequence. In the recent years, many different bioinformatics tools were developed for such prediction purposes. In this study, 4 sgRNAs predicted to target the *stylin-01* coding sequence were designed and all of them appeared to efficiently cleave DNA *in vitro* when combined with the recombinant Cas9 protein. It thus appears that CRISPOR software (Concordet and Haeussler, 2018) offers a good chance to generate functional sgRNA molecules. Injecting several sgRNAs is a means to generate large deletions within the genomic DNA sequence of a given gene, and the observation of amplicons smaller than the wild-type is thus an easy way to confirm sgRNA efficiencies. As a side effect, the simultaneous injection of several guide RNAs might also increase the risk of off-target effects. In the future, methods such as the Digenome-seq (Kim et al., 2015), could be used to estimate the rate of off-target cleavages caused by specific sgRNAs, but were not investigated in this study. The approximate size of amplicons also appears to be sufficient to predict which combinations of sgRNAs are responsible for the large deletion events observed. In that regard the various patterns observed confirmed that the 4 sgRNAs were efficient to trigger DNA cleavage in early embryos and validated the functionality of the CRISPR-Cas9 system in the pea aphid. Finally, *in ovo* mutation rates could be estimated at 70–78% in pea aphid eggs, a rate close to what can be observed in drosophila (50–86%, Bassett et al., 2013) but higher than the rates in mosquitoes estimated at 24.6% (Kistler et al., 2015). Sacrificing several eggs in the early steps of the experiment to check the editing efficiency is thus an absolute pre-requisite to include before going through the long and complex steps of diapause and hatching.

4.5. *Germline transmission rate within established mutant lineages*

Combining the three successive microinjection experiments performed in this study, 17 stable lineages were eventually generated and maintained clonally. Depending on the efficiency of the system, it was possible to observe either one or the two alleles of *stylin-01* gene mutated and in the latter case to observe different mutations for the two alleles. Amplification, cloning and sequencing of the *stylin-01* genomic DNA region within the offspring of the 17 stable lineages revealed that 6 lineages (4 arising from eggs injected with the 4sgRNAs and two from eggs injected with only one sgRNA) contained an insertion/deletion of a few nucleotides within the target sequence of at least one sgRNAs. In all cases, open reading frames were modified with insertion/deletion events ranging from −10 nt to +13 nt. Large deletions (−380 nt or −75 nt) outside predicted target sites have also been observed in two cases. For two lineages L2 and L14, mutations in both *stylin-01* alleles were detected. In L2 lineage, multiple re-arrangements with insertions and deletions were observed at the four target sites for the two alleles. In the lineage originating from eggs injected with one sgRNA, L14, mutations were also observed on both alleles at sg4 target site. These two lineages correspond therefore to double heterozygous *stylin-01* mutants. For the other four mutated lineages, only one of the two parental alleles could be detected and its sequence analyzed. Surprisingly, the other parental allele could not be detected in these 4 lineages. It is possible that sequencing only 20 clones per PCR product would not guarantee the identification of both parental alleles. A possible cloning bias could also result in the fact that only one of the two alleles was preferentially introduced within the plasmid. In addition, the exact same mutations could be detected in later parthenogenetic generations suggesting the stable character of the mutations produced here (data not shown). In the aphid model, clonality thus allows germline mutations to be directly transmitted and fixed within their offspring. Finally, the germline transmission rate in this study reached 35%. In model organisms, germline transmission rate estimation requires a subsequent crossing generation. Those rates can reach 34.5% in Drosophila (Bassett et al., 2013) and 18.9% in mosquito (Kistler et al., 2015). Despite the low number of stable colonies generated in this study, the germline transmission rate of the mutations in the pea aphid is very close to what is observed in those models, which fully validates the current CRISPR-Cas9 mutagenesis protocol. In the case of heterozygous mutant lineages, it is possible to obtain homozygous mutant lineages by performing a self-cross. Mutated lineages must be submitted to a decreasing photoperiod treatment to allow the production of sexual morphs and their subsequent mating. After hatching, 25% of the lineages should thus be homozygous for the mutations, 50% heterozygous mutant and the remaining 25% would be of a wild-type genotype.

In conclusion, the current study provides a detailed step-by-step and reproducible framework of CRISPR-Cas9 mutagenesis to create and maintain stably edited aphid lineages. *Stylin-01* edited stable lineages have been successfully generated and colonies are now well established. Analyzing their phenotypes will require further investigations. To our knowledge this work corresponds to the first example of an efficient targeted mutagenesis in the aphid model. The availability of such a protocol thus represents a major advance for the aphid scientific community to decipher and validate gene function for various life history traits associated with aphids' peculiar biology and strongly responsible for the serious damages they cause within agro-ecosystems.

Acknowledgements

This work was funded by an INRA-SPE grant (2016–2017) lead by Gaël Le Trionnaire and Véronique Brault (INRA Colmar, UMR SVQV)

and the ANR SexAphid (2018–2020) project (ANR-17-CE20-0010) lead by Denis Tagu. The authors would like to thank Jean-Christophe Simon, Gaëtan Denis and Julie Jaquiéry (INRA Rennes, UMR IGEPP) for their advices on aphid crosses and rearing. Eric Marois and Stéphanie Blandin (CNRS Strabsourg, IBMC Unit) are greatly acknowledged for their advices on mosquito eggs microinjection protocol.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ibmb.2019.04.016.

References

Bassett, A.R., Tibbit, C., Ponting, C.P., Liu, J.-L., 2013. Highly efficient targeted mutagenesis of Drosophila with the CRISPR/Cas9 system. Cell Rep. 4, 220–228.
Campos-Ortega, J.A., Hartenstein, V., 2013. The Embryonic Development of Drosophila melanogaster. Springer Science & Business Media.
Concordet, J.-P., Haeussler, M., 2018. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res. 46, W242–W245.
The International Aphid Genomics Consortium (IAGC), 2010. Genome sequence of the pea aphid Acyrthosiphon pisum. PLoS Biol. 8, e1000313.
Cui, Y., Sun, J.L., Yu, L., 2017. Application of the CRISPR gene-editing technique in insect functional genome studies–a review. Entomol. Exp. Appl. 162, 124–132.
Doudna, J.A., Charpentier, E., 2014. The new frontier of genome engineering with CRISPR-Cas9. Science 346, 1258096.
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.
Jamison, B.V., Thairu, M.W., Hansen, A.K., 2018. Efficacy of in vivo electroporation on the delivery of molecular agents into aphid (Hemiptera: aphididae) ovarioles. J. Insect Sci. 18, 49.
Jaubert-Possamai, S., Le Trionnaire, G., Bonhomme, J., Christophides, G.K., Rispe, C., Tagu, D., 2007. Gene knockdown by RNAi in the pea aphid Acyrthosiphon pisum. BMC Biotechnol. 7, 63.
Kiehart, D.P., Crawford, J.M., Montague, R.A., 2007. Quantitative microinjection of Drosophila embryos. Cold Spring Harb. Protoc. 4 pdb. prot4718.
Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H.R., Hwang, J., Kim, J.-I., Kim, J.-S., 2015. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nat. Methods 12, 237.
Kistler, K.E., Vosshall, L.B., Matthews, B.J., 2015. Genome engineering with CRISPR-Cas9 in the mosquito Aedes aegypti. Cell Rep. 11, 51–60.
Le Trionnaire, G., Wucher, V., Tagu, D., 2013. Genome expression control during the photoperiodic response of aphids. Physiol. Entomol. 38, 117–125.
Legeai, F., Shigenobu, S., Gauthier, J.P., Colbourne, J., Rispe, C., Collin, O., Richards, S., Wilson, A.C., Murphy, T., Tagu, D., 2010. AphidBase: a centralized bioinformatic resource for annotation of the pea aphid genome. Insect Mol. Biol. 19, 5–12.
Lin, G.-w., Cook, C.E., Miura, T., Chang, C.-c., 2014. Posterior localization of ApVas1 positions the preformed germ plasm in the sexual oviparous pea aphid Acyrthosiphon pisum. EvoDevo 5, 18.
Lindsley, D.L., Hardy, R.W., Ripoll, P., Lindsley, D., 2016. Gonadal mosaicism induced by chemical treatment of sperm in Drosophila melanogaster. Genetics 202, 157–174.
Lobo, N.F., Clayton, J.R., Fraser, M.J., Kafatos, F.C., Collins, F.H., 2006. High efficiency germ-line transformation of mosquitoes. Nat. Protoc. 1, 1312.
Mulot, M., Boissinot, S., Monsion, B., Rastegar, M., Clavijo, G., Halter, D., Bochet, N., Erdinger, M., Brault, V., 2016. Comparative analysis of RNAi-based methods to down-regulate expression of two genes expressed at different levels in Myzus persicae. Viruses 8, 316.
Nouhaud, P., Peccoud, J., Mahéo, F., Mieuzet, L., Jaquiéry, J., Simon, J.C., 2014. Genomic regions repeatedly involved in divergence among plant-specialized pea aphid biotypes. J. Evol. Biol. 27, 2013–2020.
Peccoud, J., Mahéo, F., De La Huerta, M., Laurence, C., Simon, J.C., 2015. Genetic characterisation of new host-specialised biotypes and novel associations with bacterial symbionts in the pea aphid complex. Insect Conserv. Divers. 8, 484–492.
Sandhu, G.S., Precup, J.W., Kline, B.C., 1989. Rapid one-step characterization of recombinant vectors by direct analysis of transformed Escherichia coli colonies. Biotechniques 7, 689–690.
Shingleton, A.W., Sisk, G.C., Stern, D.L., 2003. Diapause in the pea aphid (Acyrthosiphon pisum) is a slowing but not a cessation of development. BMC Dev. Biol. 3, 7.
Sun, D., Guo, Z., Liu, Y., Zhang, Y., 2017. Progress and prospects of CRISPR/Cas systems in insects and other arthropods. Front. Physiol. 8, 608.
Sunnucks, P., England, P.R., Taylor, A.C., Hales, D.F., 1996. Microsatellite and chromosome evolution of parthenogenetic Sitobion aphids in Australia. Genetics 144, 747–756.
Tagu, D., Le Trionnaire, G., Tanguy, S., Gauthier, J.-P., Huynh, J.-R., 2014. EMS mutagenesis in the pea aphid Acyrthosiphon pisum. G3: Genes, Genomes, Genetics, g3 113, 009639.
Taning, C.N.T., Van Eynde, B., Yu, N., Ma, S., Smagghe, G., 2017. CRISPR/Cas9 in insects: applications, best practices and biosafety concerns. J. Insect Physiol. 98, 245–257.
Wang, Y., Li, Z., Xu, J., Zeng, B., Ling, L., You, L., Chen, Y., Huang, Y., Tan, A., 2013. The CRISPR/Cas system mediates efficient genome engineering in Bombyx mori. Cell Res. 23, 1414.
Webster, C.G., Pichon, E., van Munster, M., Monsion, B., Deshoux, M., Gargani, D., Calevro, F., Jimenez, J., Moreno, A., Krenz, B., 2018. Identification of plant virus receptor candidates in the stylets of their aphid vectors. J. Virol. 92 e00432-18.