Gene editing technologies and applications for insects
Valentino M Gantz¹,² and Omar S Akbari¹,²

Initially discovered in bacteria, CRISPR-based genome editing endonucleases have proven remarkably amenable for adaptation to insects. To date, these endonucleases have been utilized in a plethora of both model and non-model insects including diverse flies, bees, beetles, butterflies, moths, and grasshoppers, to name a few, thereby revolutionizing functional genomics of insects. In addition to basic genome editing, they have also been invaluable for advanced genome engineering and synthetic biology applications. Here we explore the recent genome editing advancements in insects for generating site-specific genomic mutations, insertions, deletions, as well as more advanced applications such as Homology Assisted Genome Knock-in (HACK), potential to utilize DNA base editing, generating predictable reciprocal chromosomal translocations, and development gene drives to control the fate of wild populations.

Addresses
¹ Division of Biological Sciences, Section of Cell and Developmental Biology, University of California, San Diego, La Jolla, CA 92092, USA
² Tata Institute for Genetics and Society, University of California, San Diego, La Jolla, CA 92093, USA

Corresponding author: Akbari, Omar S (oakbari@ucsd.edu)

Introduction
The increasingly refined ability to introduce altered traits into insects via genome editing offers scientists exciting opportunities for tackling public health and environmental issues in novel sustainable ways. Intentional spread of engineered traits through wild insect populations could be used to address numerous problems as varied as biocide resistance, ill effects associated with invasive species, and insect-borne diseases. For example, the replacement of herbicide — or pesticide — resistant alleles with sensitive ones may restore vulnerability to herbicides or pesticides, allowing for their continued use. Moreover, the introduction of genes that make an organism sensitive to a previously innocuous molecule may allow for that molecule to be utilized as a novel biocide. Additionally, the spread of certain genetic elements, for example ones that impede vector competence, or that cause deleterious recessive mutations, or that bias the sex ratio of a population, may be used to block the spread of vector disease, or suppress invasive organisms in a species-specific eco-friendly manner, respectively [1].

Recent genome editing and engineering approaches in insects have evolved from the early random chemical and radiation mutagenesis strategies for genome modification [2], to transgenesis methods based on transposable elements, and more recently to site-directed mutagenesis and transgenesis approaches utilizing sequence-specific nucleases enabling the manipulation of the genome with surgical accuracy. These techniques have permitted researchers to generate random, or planned modifications, within the genomes of insects to investigate the function of genes and their regulatory sequences, and to engineer synthetic genetic elements with novel functions. Editing strategies in recent years have been developed to exploit different nucleases beginning from sequence-specific zinc finger nucleases [3], to modular TALENs (transcription activator-like effectors nucleases) [4], and now RNA-guided nucleases adapted from bacterial adaptive immune systems, known as CRISPR/Cas (clustered regularly interspaced palindromic repeats/CRISPR associated systems). Together these tools have given birth to a new era of gene editing and genome engineering. In general, each of these technologies uses sequence-specific nucleases to generate double-stranded DNA breaks (or nicks) in regions of interest, this allows for targeted DNA modifications by taking advantage of endogenous mechanisms to repair broken DNA. Since cells are unable to divide further when harboring broken chromosomes, the nuclease generated cuts must be rapidly repaired by the cell to ensure survival. Two DNA repair pathways are usually employed for this purpose: (1) non-homologous end-joining, which can lead to small insertion and deletions (indels) at the break site; (2) homology-directed repair, which is designed to use the information on the intact chromosome to accurately repair the broken one. The latter can be turned to the researcher’s advantage as the cell can be tricked into using a synthetic construct as a template, and therefore leading to user-defined insertions or deletions (reviewed in [5]).

The CRISPR/Cas system has been particularly amenable to be used in insects and, thus far, it has been used in multiple species with minimal optimization steps (reviewed in [6]). These applications reduce the bacterial immunity complex to a simplified version composed of
two components: (1) the Cas9 endonuclease, which performs DNA cleavage, and (2) a synthetic guide RNA (gRNA) which pilots the nuclease to the target genomic location, programmed within its RNA sequence [7]. In insects, these two elements can be delivered as RNA, plasmid DNA, or encoded in the genome to increase efficiency (reviewed in [8]). When combined they can lead to whooping rates of mutagenesis at the target site efficiently disrupting the function of the target DNA sequence. Alternatively, when combined with an exogenous DNA source (single-stranded DNA or double-stranded plasmid DNA) harboring homology to the genomic target sequence on each side of the intended cleavage site, the intervening DNA ‘cargo sequence’ can be reliably and efficiently inserted at the cut site with efficiencies comparable to traditional transgenesis [9]. Here we discuss recent developments and focus on two broad aspects of utilizing gene editing in insects including functional genomics to elucidate gene function, and utilizing advanced gene editing for insect control.

**Practical uses of the CRISPR/Cas9 to investigate the genome’s function**

Many applications of CRISPR/Cas9 for genome editing have been developed to modify the DNA sequence of an insect’s genome. To utilize CRISPR, the Cas9 endonuclease, combined with one or more gRNAs, is delivered to the nucleus where the Cas9/gRNA complex is directed by base-complementarity of the guide RNA(s) with their user-specified genomic target to induce DNA cleavage (Figure 1a and b). Once a cut is generated on a chromosome, the cell senses the break and responds to it by via one of two main DNA-repair pathways: nonhomologous end joining (NHEJ) and homology-directed repair (HDR). While other repair pathways could be employed by the cell to repair the break (reviewed in [10,11]) NHEJ and HDR have been thus far co-opted successfully for efficient genome editing in several insects (reviewed in [12]), with extensive pioneer work and optimization in the fruit fly [13,14].

**Targeted mutations and knock-outs (KOs) with NHEJ**

NHEJ is an error-prone process that can be exploited to generate localized mutations, usually small indels, neighboring the cut site determined by the gRNA target sequence (Figure 1c). Additionally, if two (or more) gRNAs are used in concert, NHEJ can be used to obtain deletions of large DNA fragments by occasional loss of the sequence in between the two gRNAs (Figure 1d). This approach is extremely useful to investigate gene function by generating small mutations in the coding region of a gene, or its regulatory region when using one gRNA, and the generation of functional knockouts (KOs) by deleting entire genes or gene complexes using two gRNAs. For example, Gratzer et al. achieved genomic deletions of up to 14.2kb in *Drosophila melanogaster* using a comparable strategy [13].

**Targeted knock-ins (KIs) and knock-outs with HDR**

By utilizing CRISPR to first generate a double-stranded DNA break, foreign DNA sequences can be integrated with high efficiency into the genome at this break site by exploiting the HDR pathway. In general, this pathway utilizes the intact sister chromatid as a template to correct dsDNA breaks. However, when a DNA template (single or double stranded) with homology to the target DNA sequence abutting the CRISPR target cut site is provided, the HDR machinery can be exploited to ‘knock-in’ the DNA in between the homologous sequences. Single-stranded oligodeoxynucleotides (ssODN’s) of up to 200 nt can be readily ordered from commercial vendors and, by allowing 50 nt of homology to each side of the dsDNA cut, these molecules can be used to efficiently insert small sequences up to ~100 nt (e.g., attP docking site) into the target location at high frequency [15] (Figure 1e and f). Double-stranded DNA (i.e., plasmid DNA) can also be used as a template, in this case, it has been shown that 500–1000 bp homology arms (HAs) to each side of the dsDNA break seem to promote efficient insertion of the cargo DNA contained in between the HAs [13]. Depending on the gRNAs used in the strategy (Figure 1g–i), researchers have used HDR to knock-in a 17.3 kb cargo DNA in mosquitoes as part of a gene drive strategy (using one gRNA) [16] or a concerted knock-in/knock-out strategy in *D. melanogaster* [13], or a trans-species regulatory sequence replacement to generate chimeric animals [9].

**Complex genome editing & synthetic genomics**

Combining the use of CRISPR-based techniques with other genetic technologies has resulted in more powerful applications permitting a synthetic control of the genome. For example, a recent report by the Potter group used traditional transgenesis to insert a donor HDR-template, termed Homology Assisted CRISPR Knock-in (HACK). Once the donor HDR-template was inserted in the genome, the authors then used a gRNA to cut within genome encoded Gal4-expressing transgenes and convert them into a QF2 transactivator using homology sequence to the Gal4 transgene from the HACK element (general strategy outlined in Figure 2a). In this study, the researchers were able to generate HACKing lines that would essentially convert any *D. melanogaster* Gal4-expressing line into a comparable QF2 expressing one in only two genetic crosses [17]. In a separate report, nuclease-dead versions of Cas9 (termed dCas9) have been fused to DNA-base-modifier domains and employed as localized mutagens that are independent of any DNA cleavage. This system, first characterized in mammalian cell cultures [19] has been already used in diverse systems and in the future could be readily adapted to insects to generate precise edits in the genome limiting the risks associated with off-target effects (reviewed in [20]) (Figure 2b). Lastly, Buchman et al. used randomly-inserted fluorescently marked transgenes, carrying mutual homology
sequences, to generate predictable reciprocal translocations in the presence of a sequence-specific nuclease [18**]. This technique, which is used by the authors to generate high-threshold gene drive systems, could also be used in other insect species to generate sequential chromosomal translocations, and in the future could be adapted to engineer chromosomal inversions and generate balancer chromosomes to boost genetics in emerging model systems (Figure 2c).

Advanced applications for insect control
The ability to combat insect-borne diseases through engineering populations has immense potential, as insects act as vectors for a number of important diseases affecting humans, animals, and plants [21]. The impact of these diseases has been greatly aggravated by increased global movement of commodities, people, and animals, which is leading to the invasive spread of disease vectors and pathogens into new environments [22–24]. While
Complex genome editing examples. In panel (a) and active genetic element with homology arms and a gRNA to a given genomic location is inserted randomly using a transposable element (represented by the black arrows) (a'); further crossing of such transgenic animal with one carrying a source of Cas9, results in the element copying onto the predetermined location by homology-directed repair (a'') [17]. (b) describes the use of an activation-induced cytidine deaminase (AID), fused to a dead form of Cas9, to target mutagenic DNA base-editing in the neighborhood of the gRNA target site (b'); the endogenous DNA-repair mechanisms repair these mismatches creating occasional mutations in the targeted locus (b'') [19]. (c) Describes a strategy that takes advantage of inverted homologous sequences, separated by a gRNA target site, inserted into two separate transgenes (c'); in presence of Cas9 and the appropriate gRNA (in panel (c) they are supplied as transgenes), the gRNA cleaves both locations resulting in four chromosomal fragments with homologous sequences at their ends (c'') which trigger recombination and result in a predictable reciprocal translocation between the darker and the lighter chromosomes carrying the transgenes (c'''') [18]. Lastly, panel (d) describes the use of a transgene containing both a Cas9 and a gRNA genes inserted at the location targeted by the gRNA (d'); this arrangement generates an active genetic element capable of cutting the opposing (wild-type) chromosome (d'') and converting it to the same condition (d'''); when this process happens in the germline of an animal it generated a strong gene drive effect which dramatically modifies the inheritance pattern of the genetic construct from Mendelian (d''') to Super-mendelian (d'''') and can be taken advantage of for population suppression or modification applications.

vector control is an important component of disease prevention, it is often expensive, with the degree of protection provided being dependent, on a continuous basis, to the effort and money allocated for control. Additionally, specific methods of vector control, such as environmental modification, or use of chemical insecticides, may be impractical or have undesirable side effects in certain contexts. A complementary strategy for disease prevention, first articulated many decades ago [25], involves replacement of wild, disease-transmitting insects.
with individuals that are engineered to be refractory to disease transmission, but that are still subject to vector control [26]. A central appeal of this strategy is that in contrast to vector suppression via insecticides alone, population replacement is species-specific and potentially self-perpetuating. Notwithstanding, a major obstacle to implementation of population replacement approaches is that engineered traits are unlikely to confer an overall fitness benefit on organisms that harbor them, but instead typically reduce fitness and therefore are rapidly lost from populations through natural selection [27,28,29,30]. Therefore, an essential component of any wild population replacement strategy is the presence of a ‘gene drive’ mechanism that will ensure the spread of engineered transgenes to genotype or allele fixation in a modest number of generations following a release. This capability to catalytically spread engineered traits through wild populations would enable novel methods of addressing a plethora of significant worldwide issues, including the spread of insect-borne diseases, ill effects of invasive species, and biocide resistance.

Several naturally occurring selfish genetic elements, including transposons, meiotic drive, B-chromosomes, homing endonuclease genes (HEGs), and Medea elements, have been proposed as potential gene drive mechanisms (reviewed in [31,32**,33,34,35]), along with approaches relying on linking genes of interest to engineered chromosomes, such as translocations or compound chromosomes [25,36], or engineered underdominance [37,38] (Figure 2c). Some of these strategies, including Medea [39,40,41**], engineered underdominance [42], and HEGs [43] have been shown to have some capacity to mediate gene introgression in laboratory populations of insects. However, these systems have been difficult to engineer in diverse species and therefore much of the recent excitement has been redirected toward using CRISPR systems to generate gene drives (Figure 2d). In fact, significant progress has already been made and CRISPR homing based drive systems have recently been demonstrated to bias mendelian inheritance in insects with incredible efficiency [16**,44,45*]. Given these exciting results, future efforts are now aimed toward utilizing gene editing approaches to further optimize drive systems, transfer working drive technologies to other insects that transmit pathogens to humans, and to engineer new drive systems inspired from naturally existing systems (reviewed in [32**]).

Given the rapid scientific progress aimed at the development of gene drives articulated above, the discussion has now turned to the ethics and regulation of such systems. This is particularly complicated due to the fact that many of these systems are invasive and can spread beyond borders requiring international agreements to be made before any planned release. For example, Medea, HEGs, transposable elements, meiotic drive, and RNA-guided CRISPR-based homing gene drives, are predicted to be invasive drive mechanisms with low release thresholds, capable of spreading to high frequency even when a small number of individuals are introduced into a population [46,47]. Invasive gene drive mechanisms are ideal when the goal is to spread genes over a large area, particularly when migration rates between the release site and surrounding areas of interest are low. However, because such systems have low release thresholds, once introduced, the pre-transgenic state cannot easily be restored by diluting the replaced population with wild-type alleles such that the frequency of the gene drive alleles falls below the threshold frequency required for their spread. Therefore, given their potency and difficulty of removal, establishing international agreements and developing countermeasures and safeguards, prior to any planned release, is a high priority issue for the entire gene drive field [30,32**,48,49,50].

Conclusions
In summary, while still relatively new, CRISPR based gene editing technologies have already revolutionized functional genomics of insects. With CRISPR we now have the ability to rapidly modify, delete, and insert DNA nearly anywhere we desire in virtually any insect species. Additionally, with more advanced genome editing based technologies rapidly being developed such as HACK systems, DNA base-editors, site-specific chromosomal translocations, and gene drives, we may soon have the power to eradicate disease-transmitting insects in the wild. Given this exciting potential, it is important to continue optimizing gene editing technologies, to understand and overcome potential limitations and risks, and to engage with regulators, stakeholders, and the general public to ensure the safe and timely application of these promising technologies to address the problems they are intended to solve.

Conflict of interest
VMG is a founder of Synbal, Inc. and Agragene, Inc.

Acknowledgements
This work was supported in part by NIH grants 5K22AI113060, 1R21AI23937 and a Defense Advanced Research Project Agency (DARPA) Safe Genes Program Grant HR0011-17-2-0047 to OSA; and NIH grant DP5 OD023098 awarded to VMG. We thank Ethan Bier (UCSD) and Anthony A James (UCI) for their useful comments and edits.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as
● of special interest
** of outstanding interest


Development of a CRISPR-homing-based population replacement gene drive system inAnopheles stephensi.


This paper demonstrates that site-specific endonucleases can be exploited to engineer user defined reciprocal chromosomal translocations that can efficiently spread into a population if released above a critical threshold frequency.


A comprehensive review of the gene drive field outlining how CRISPR can be used to engineer various classes of gene drives.


42. Akbari OS, Matzen KD, Marshall JM, Huang H, Ward CM, Hay BA: A synthetic gene drive system for local, reversible


