Methods 69 (2014) 128–136

Contents lists available at ScienceDirect

Methods

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

CRISPR/Cas9 mediated genome engineering in *Drosophila*

Andrew Bassett *, Ji-Long Liu *

Medical Research Council Functional Genomics Unit, Department of Physiology, Anatomy and Genetics, University of Oxford, South Parks Road, Oxford OX1 3PT, United Kingdom

ARTICLE INFO

Article history:
Received 27 November 2013
Revised 28 January 2014
Accepted 7 February 2014
Available online 24 February 2014

Keywords:
Drosophila melanogaster
CRISPR
Cas9
Genome engineering
Targeted mutagenesis
RNA injection

ABSTRACT

Genome engineering has revolutionised genetic analysis in many organisms. Here we describe a simple and efficient technique to generate and detect novel mutations in desired target genes in *Drosophila melanogaster*. We target double strand breaks to specific sites within the genome by injecting mRNA encoding the Cas9 endonuclease and in vitro transcribed synthetic guide RNA into *Drosophila* embryos. The small insertion and deletion mutations that result from inefficient non-homologous end joining at this site are detected by high resolution melt analysis of whole flies and individual wings, allowing stable lines to be made within 1 month.

1. Introduction

Our ability to design DNA binding factors with exquisite specificity for desired target sequences has heralded a new wave of genome engineering techniques that allow targeted modifications of the genome to be achieved in many organisms [1–14]. This new genome engineering technology will enable more directed and elegant experiments to be performed to analyse structural and functional aspects of the genome.

The CRISPR/Cas9 system was discovered as a bacterial defence system against invading viral pathogens, which uses fragments of RNA from the virus to target cleavage of the viral DNA through complementary base pairing [15–20]. This system has recently been shown to be active in other systems, including mammals [1–3], insects [6–12] and plants [13], and can be easily modified to target double strand breaks (DSB) at any desired target sequence by supplying it with a short guide RNA that is complementary to the target site within the DNA. The endogenous system involves three components. The Cas9 protein is an endonuclease that binds to a structure within a trans-acting CRISPR RNA (tracrRNA). The tracrRNA base pairs with a CRISPR RNA (crRNA), the first 20 nt of which determine the specificity of the Cas9 endonuclease. A simplified two component system has been described that fuses the tracrRNA and crRNA into a single synthetic guide RNA (sgRNA), making delivery of the components easier [1,2,18].

The DSBs produced can be repaired by non-homologous end joining (NHEJ) or homologous recombination (HR), and both can be useful to introduce mutations into the underlying DNA [21]. NHEJ repair is error prone, and often results in small insertions or deletions (indels) at the cut site, that can be mutagenic. Targeting two DSBs can also result in the deletion of intervening sequences, to generate longer deficiencies [6]. Induction of a DSB also enhances rates of HR repair, which can be used to enhance gene targeting efficiencies by several orders of magnitude [22–24].

This system has been developed for use in many organisms, including *Drosophila*, where multiple methods of introducing the Cas9 and sgRNA components have been developed [6–12] (Table 1). The Cas9 protein can be introduced by injection of mRNA or an expression vector into the early embryo [6–8], or by using a transgenic strain that produces the Cas9 protein under a germline-specific or ubiquitous promoter [9–11]. The sgRNA itself can be produced by in vitro transcription [7,8], or expressed from a pol III promoter derived from the U6 snRNA gene [6,9–11]. The use of a pol III promoter avoids capping and polyadenylation of the transcript, which may inhibit its activity. Again, in vitro transcribed sgRNA or an expression plasmid can be injected into *Drosophila* embryos, or transgenic strains can be produced that express the sgRNA ubiquitously. These techniques can be used in different combinations, and each has advantages in certain circumstances, or for specific experiments (Table 1). For instance, the highest reproducibility and efficiency of mutagenesis can be achieved by crossing two transgenic lines, but it relies on generating a transgenic

Abbreviations: CRISPR, clustered regularly interspaced palindromic repeats; Cas, CRISPR associated; DSB, double strand break; NHEJ, non-homologous end joining; HR, homologous recombination.

* Fax: +44 1865 282849 (A. Bassett), +44 1865 285862 (J.-L. Liu).
E-mail addresses: andrew.bassett@dpag.ox.ac.uk (A. Bassett), jilong.liu@dpag.ox.ac.uk (J.-L. Liu).

http://dx.doi.org/10.1016/j.ymeth.2014.02.019
1046-2023/

Table 1
Comparison of CRISPR/Cas9 techniques in *Drosophila*.

Reference	Gratz et al. [6]	Bassett et al. [7]	Yu et al. [8]	Kondo and Ueda [9]	Sebo et al. [11]	Ren et al. [10]
Cas9 promoter	*hsp*70	T7	Sp6	*nos*	*vasa*	*nos*
Cas9 delivery	DNA injection	mRNA injection	mRNA injection	Transgenic	Transgenic	Transgenic
sgRNA promoter	U6	T7	T7	U6	U6	U6a, U6b, nos-mini
sgRNA delivery	DNA injection	sgRNA injection	sgRNA injection	Transgenic	DNA injection	DNA injection
Target genes	*Yellow*	*Yellow*, *white*	*Yellow*, *K*81, *CG*3708, *CG*9652, *kl*-3, *light*, *RpL*15	*White*, neuropeptide genes (*Ast*, *capa*, *Ccap*, *Crz*, *Eh*, *Mip*, *npf*), *mir*-219, *mir*-315	*EGFP*, *mRFP*	*White*
Mosaic G_0 (%)[a]	6–66	4–88	35.7–80	N/A	N/A	N/A
Germline mutants (among fertile flies) (%)[b]	5.9–20.7	0–79	35.7–100	0–100	35–71	0–100
F_1 mutant overall (%)[c]	0.25–1.37	0–34.5	2.1–98.9	0–99.4	7.7–24.7	0–74.2
Overall Timescale[d]	~1 month	~1 month	~1 month	~2–3 months	~1 month	~1 month
Applicable to all genetic backgrounds[e]	Yes	Yes	Yes	No	No	No

This table is modified from Table 1 in Bassett and Liu [37].
[a] Percentage of flies that exhibit mosaic expression in the injected generation, either visibly in males or detected using HRMA (high resolution melt analysis).
[b] Proportion of fertile flies giving rise to at least one mutant offspring.
[c] Total number of mutant G_1 offspring as a percentage of the total offspring.
[d] Approximate overall timescale including the time spent generating transgenic fly stocks (if applicable).
[e] All of the techniques involving transgenic delivery of Cas9 rely on injecting into specific fly lines, limiting their ability to compound mutations with existing lines, or generating mutations in other genetic backgrounds or *Drosophila* strains. N/A, not applicable to this technique, since Cas9 is germline restricted.

line expressing each desired sgRNA, which is relatively time consuming. Although giving good mutagenesis efficiency, all of the techniques involving transgenic Cas9 expression rely on injection or crossing to the transgenic fly lines, making it difficult to compound mutations with pre-existing alleles, or inject into different genetic backgrounds. The described technique has the advantage that it can be performed in essentially any genetic background, and there is no possibility of integration of DNA constructs into the genome, but does require care in the production and handling of the injected RNA.

Here we describe a detailed methodology to produce and inject mRNA encoding the Cas9 protein, and in vitro transcribed sgRNAs that can result in high efficiencies of mutagenesis of desired target genes by inefficient NHEJ. Up to 88% of flies have mosaic mutations in the target gene, which can be transmitted to up to 34.5% of total F_1 offspring [7] (Table 1). We also describe the application of high resolution melt analysis (HRMA) to provide a simple and effective system of detection of the resulting indel mutations to enable generation of stable mutant lines [7]. This technique utilises the fact that indel mutations change the melting temperature of PCR products spanning the target site to rapidly and accurately detect mosaic and heterozygous mutant flies.

2. Materials and methods

2.1. Overview

sgRNAs are designed to target the gene of interest that minimise potential off target effects and maximise mutagenic efficiency, and templates for their transcription are generated by a simple PCR. The sgRNA and mRNA encoding the Cas9 protein are generated by in vitro transcription, purified and coinjected into *Drosophila* embryos of essentially any genotype. Mosaic flies are identified by HRMA, and heterozygous mutant offspring from these flies are selected by analysis of PCR products from single wings by HRMA and sequencing. These flies are used to make stable stocks that can be used for further analysis. An overview of the process with approximate timings is shown in Fig. 1, and reagents required are listed in Supplementary Table 1.

2.2. sgRNA design

2.2.1. Target site choice

Cas9 is guided to 20 nt target sequences in the genome that are complementary to the 5′ end of the sgRNA, and these sequences must be followed by an NGG protospacer adjacent motif (PAM) sequence (Fig. 2A). The PAM sequence does not appear in the sgRNA, but is nevertheless required by the Cas9 protein for efficient endonucleolytic cleavage of the DNA. These sequences can be on either strand of the DNA, making the expected frequency of such sites approximately every 8 nt, although within certain genomic regions, this can be considerably less often. Some reports suggest that the NGG PAM sequence can be replaced by NAG [25], but the relative efficiencies of cleavage have not been directly tested.

Since sgRNAs therefore only have a 20 nt (target) + 2 nt (PAM) specificity determinant, and recent studies have shown that mismatches can be tolerated within the target sequence [25–30], targets must be carefully chosen to minimise the potential for off target effects. Ideally, sgRNA sequences should be chosen whose closest off target site differs by at least 4 nt, but this requirement can be relaxed if the mutations cluster towards the 3′ end of the sgRNA, closest to the PAM. Mutations of only 1–3 nt within the final 10 nt of the target sequence often prevent cleavage, especially if they are next to each other. Several websites have recently become available to enable simple design of sgRNA target sequences that minimise potential off targeting (for example http://crispr.mit.edu/ [29], http://www.flyrnai.org/crispr/ [10], http://tools.flycrispr.molbio.wisc.edu/targetFinder/ [6], http://www.e-crisp.org/E-CRISP/).

For mutation of protein coding genes, it is often desirable to choose target sequences that will result in failure to produce a functional protein. Target sites should be chosen that are within the coding sequence of the gene to induce frameshifts, at the translational start codon or at the splice acceptor or donor sites of a common exon. This is because only 2 of 3 indels will result in a frameshift, whereas removal of a splice site or start codon will prevent a functional protein being produced. The indels produced can also be used to remove other functional sites within the genome such as transcription factor binding sites, miRNA target sites, splice sites and transcriptional start sites as well as mutating protein coding genes.