Insect Biochemistry and Molecular Biology 72 (2016) 31–40

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

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

Expansion of CRISPR targeting sites in *Bombyx mori*

Baosheng Zeng [a, b], Shuai Zhan [a], Yueqiang Wang [a], Yuping Huang [a], Jun Xu [a, b], Qun Liu [a], Zhiqian Li [a], Yongping Huang [a, **], Anjiang Tan [a, *]

[a] Key Laboratory of Insect Developmental and Evolutionary Biology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:
Received 20 January 2016
Received in revised form
16 March 2016
Accepted 18 March 2016
Available online 24 March 2016

Keywords:
CRISPR/Cas9
U6 promoter
sgRNA
Transgene
Bombyx mori

ABSTRACT

The CRISPR/Cas9 system has been proven as a revolutionary genome engineering tool. In most cases, single guide RNA (sgRNA) targeting sites have been designed as GN19NGG or GGN18NGG, because of restriction of the initiation nucleotide for RNA Pol III promoters. Here, we demonstrate that the U6 promoter from a lepidopteran model insect, *Bombyx mori,* effectively expressed the sgRNA initiated with any nucleotide bases (adenine, thymine, guanine or cytosine), which further expands the CRISPR targeting space. A detailed expansion index in the genome was analysed when N20NGG was set as the CRISPR targeting site instead of GN19NGG, and revealed a significant increase of suitable targets, with the highest increase occurring on the Z sex chromosome. Transfection of different types of N20NGG sgRNAs targeting the enhanced green fluorescent protein (EGFP) combined with Cas9, significantly reduced EGFP expression in the BmN cells. An endogenous gene, *BmBLOS2*, was also disrupted by using various types of N20NGG sgRNAs, and the cleavage efficiency of N20NGG sgRNAs with different initial nucleotides and GC contents was evaluated *in vitro*. Furthermore, transgenic silkworms expressing Cas9 and sgRNAs targeting the *BmBLOS2* gene were generated with many types of mutagenesis. The typical transparent skin phenotype in knock-out silkworms was stable and inheritable, suggesting that N20NGG sgRNAs function sufficiently *in vivo*. Our findings represent a renewal of CRISPR/Cas9 target design and will greatly facilitate insect functional genetics research.

1. Introduction

The newly developed type II prokaryotic CRISPR/Cas9 system has been used for genome engineering in both model and non-model organisms (Cong et al., 2013; Feng et al., 2013; Friedland et al., 2013; Gratz et al., 2013; Hwang et al., 2013; Mali et al., 2013b; Wang et al., 2013a). The mechanism of single-guide RNA (sgRNA) guiding the Cas9 nuclease to cleave specific sites in the genome is well understood (Hsu et al., 2014; Sander and Joung, 2014). Cas9 targeting requires a 20-nucleotide sgRNA sequence that is complementary to the target and a protospacer-adjacent motif (PAM). Theoretically, any N20-PAM sequence in the genome could be set as a potential target. The expression of sgRNAs with a desired sequence is an essential step for CRISPR technology, and the RNA polymerase type III promoter U6 is commonly used for this purpose. In mammals, the U6 promoter drives the expression of sgRNAs with the requirement of a guanine nucleotide for transcription initiation, thus constraining the sgRNA target site to GN19NGG. Ran et al. provided an alternative way by adding one G at +1, thus making GN20NGG the target sequence can be efficiently utilized in mammals (Ran et al., 2013c). However, there are no further reports to study the potential side effects of the added 5′G, which artificially induces one mismatch, or whether it could be extended to other classes. Other commonly used promoters for short RNAs *in vitro* transcription, such as SP6, T7 and T3, also require one or two starting guanine nucleotide(s) (Adhya et al., 1981; Ma et al., 2014a; Melton et al., 1984; Pleiss et al., 1998). Recently, utilizing the ribozyme self-processing capacity, the RNA polymerase II promoter was used to express sgRNAs through a *ribozyme-gRNA-ribozyme*(RGR) gene system to produce mature sgRNAs (Gao and Zhao, 2014). Additionally, the H1 pol III promoter is an alternative option for sgRNA transcription, which could initiate both guanine and adenine (Ranganathan et al., 2014). In *Drosophila*, the Pol II *nanos* promoter has been used to express

* Corresponding author.
** Corresponding author.
E-mail addresses: yphuang@sibs.ac.cn (Y. Huang), ajtan01@sibs.ac.cn (A. Tan).

http://dx.doi.org/10.1016/j.ibmb.2016.03.006
0965-1748/

sgRNAs by adding the 5′ and 3′ untranslated region (UTR) of *nanos* gene to the upstream and downstream of sgRNA sequence (Ren et al., 2013).

The domestic silkworm *Bombyx mori* is a lepidopteran model insect for genetic research (Tan et al., 2013). In recent years, numerous genetic manipulation techniques have been established in *B. mori*, including transposon-based transgenesis, RNAi, zinc finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) (Sajwan et al., 2013; Takasu et al., 2010; Tamura et al., 2000). The CRISPR/Cas9 system has also been established in *B. mori*, both in cell lines and *in vivo* (Liu et al., 2014; Ma et al., 2014b; Wang et al., 2013b). The *in vitro* transcription of sgRNAs by the T7 promoter require the target sequence of GGN19NGG, and only GN19NGG sgRNAs expressed with the U6 promoter were applied, which substantially limits the extended usage of the CRISPR/Cas9 system in short coding genes or non-coding RNAs loss of function researches (Ho et al., 2015; Zhao et al., 2014).

In the present study, we demonstrate that the ability of *B. mori* U6 promoter (Tanaka et al., 2009) to express sgRNAs beginning with the other three nucleotides (T, A and C). The potential expansion of CRISPR sites in the *B. mori* genome was calculated and analysed. The specifically designed N20NGG sgRNAs under the control of the U6 promoter cooperated with the ubiquitously expressed Cas9 protein effectively knocked out both exogenous and endogenous genes *in vitro* and *in vivo*. The transparent skin phenotype induced through disruption of *BmBLOS2* by N20NGG type sgRNAs (combined with Cas9) was observed and inheritable. Finally, the cleavage efficiency of sgRNAs with different initial nucleotides with different GC contents was also evaluated to provide possible principles for future sgRNAs target design.

2. Materials and methods

2.1. Bioinformatics analysis

To determine the potential sgRNA target sites in the *B. mori* genome, the occurrence of AN19NGG, TN19NGG, GN19NGG, CN19NGG and N20NGG in each chromosome was calculated and plotted with Circos (v0.52, http://circos.ca/). Then, we evaluated the mean and median distance values of two adjacent GN19NGG and N20NGG sites. The frequencies of the GN19NGG and N20NGG sites in *B. mori* genes were also analysed and compared. We also examined the density of GN19NGG and N20NGG in all identified *B. mori* mRNA and microRNAs. Finally, the genomes of several other representative insect species, including the fruit fly, beetle, honeybee and pea aphid, were also similarly evaluated. The genome and annotation of the silkworm were downloaded from SilkDB. The genomes of the additional insects were obtained from FlyBase (for *Drosophila melanogaster*), BeetleBase (for *Tribolium castaneum*), HGD (for *Apis mellifera*) and AphidBase (for *Acyrthosiphon pisum*). We used custom Perl scripts to search all of the potential CRISPR sites in the genome and calculated the frequency and distance. The CRISPR sites at genes or miRNAs were determined by BEDtools (intersectBed, v2.17.0).

2.2. Silkworms and embryo injections

The *B. mori* strain Nistari, which is multivoltine and nondiapausing, was used in all of the experiments. The larvae were reared with fresh mulberry leaves or an artificial diet under standard conditions (Tan et al., 2013). The Cas9 and sgRNA expressing vectors were extracted using a QIAGEN Plasmid Midi Kit (QIAGEN, Hilden, Germany) and purified by phenol chloroform. The mixture of Cas9-sgRNA plasmids and helper plasmids (plasmids which express the *piggyBac* transposase, final concentration of each vector was 300 ng/μl) was microinjected into preblastoderm G0 embryos within 8 h after oviposition and microinjection was performed according to our previous report (Ling et al., 2014; Tan et al., 2013). For each independent group, 640 preblastoderm embryos were injected, then the injected embryos were incubated at 25 °C in a humidified chamber for 10–11 days until larval hatching (Tan et al., 2013).

2.3. Plasmid construction

To generate plasmids for the single-strand annealing (SSA) assay, we constructed seven pGL (firefly luciferase) plasmids within two 870-bp repeat sequences combined with different artificial sgRNA targets. The CMV promoter in the original pGL plasmid (Promega, Madison, WI) was replaced by IE1 promoter (Li et al., 2015a). Then the pGL plasmids containing 870 bp repeat sequence were constructed by two steps of cloning: 1) insert the pGL-part-1(including the first section of GL and first 870 bp repeat sequence) between NheI and BglII; 2) ligate the second part (sgRNA target, the other 870 bp repeat sequence and the rest part of GL ORF sequence) with BglII and BamHI (or XbaI). All the primers used were listed in Table s4. The U6:sgRNA scaffold:pol III terminator sequence was obtained using the U6 promoter as a template (U6 promoter was synthesized based on the sequence reported by Tanaka et al.) (Tanaka et al., 2009) and was amplified with forward primer located at the U6 promoter and reverse primer (sgRNA targeting sequence included) (Table s4), then sub-cloned into the pJET Blunt vector and verified by sequencing. The Cas9 ORF sequence was cloned from the PTD1-Cas9 plasmid(ViewSolid Biotech, Beijing, China) and sub-cloned into *pBac[IE1-DsRed2-IE1-SV40]*(Li et al., 2015a) to achieve *pBac[IE1-DsRed2-IE1-Cas9-SV40]* with AatII and ApaI. The vectors used for transgenic transformation were generated by inserting a U6-sgRNA expression cassette into *pBac[IE1-DsRed2-IE1-Cas9-SV40]* with KpnI and obtained *pBac[IE1-DsRed2-U6-sgRNA-IE1-Cas9-SV40]*. The *pBac[IE1-EGFP-SV40]* was constructed based on *pBac[IE1-DsRed2-SV40]* by using primers of BglII-HR5-IE1-F/NotI-HR5-IE1-R (HR5 as the enhancer) and NotI-EGFP-F/BamHI-EGFP-R. The related primers were presented in Table s4 (Li et al., 2015a).

2.4. BmN cell culture and transfection

A stable BmN cell line was maintained at 27 °C with the TC100 insect cell medium (PAN-Biotech, Aidenbach, Germany) supplemented with 10% foetal bovine serum (FBS) (Thermo Fisher Scientific, Waltham, MA). The plasmids (including plasmids used for *in vitro* knock out and SSA assay experiments) used for the transfection in the cells were extracted and purified with phenol chloroform. Transfection was performed using an Effectene Transfection Reagent kit following the manufacturer's recommended instructions (QIAGEN, Hilden, Germany). In total, 400 ng of DNA were transfected per well of a 24-well plate. Twenty-four hours post-transfection, the medium containing transfection reagents was replaced with the new growth medium. Cells were assessed by fluorescence microscopy or harvested for genomic DNA extraction 72 h after transfection.

2.5. SSA dual luciferase reporter assay

For each well of a 24-well plate, 100 ng of pGL reporter plasmid (containing the structure of 870 bp repeat sequence, sgRNA targets and 870 bp repeat sequence), 100 ng of pRL control plasmid(Promega, Madison, WI), 100 ng of Cas9 expression plasmid(*pBac[IE1-dsRed-IE1-Cas9-SV40]*) and 100 ng of sgRNA vector (U6-sgRNA expression plasmids) were transfected (Cradick et al., 2014).