Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice

Graphical Abstract

Highlights

- CRISVar allows the efficient generation of structural variations in ESCs

- ESCs carry duplications, inversions, deletions, or different combinations thereof

- ESC clones with specific SVs generate germline transmitting chimeras

Authors

Katerina Kraft, Sinje Geuer, ..., Darío G. Lupiáñez, Guillaume Andrey

Correspondence

lupianez@molgen.mpg.de (D.G.L.), andrey@molgen.mpg.de (G.A.)

In Brief

Kraft et al. now present a 10-week-long protocol to engineer deletions, inversions, and duplications over a megabase in mice. The authors generate ESCs carrying SVs and show that mice with deletion of Laf4 recapitulate a human malformation.

Kraft et al., 2015, Cell Reports 10, 833–839
February 10, 2015 ©2015 The Authors
http://dx.doi.org/10.1016/j.celrep.2015.01.016
Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice

Katerina Kraft, Sinje Geuer, Anja J. Will, Wing Lee Chan, Christina Paliou, Marina Borschiewer, Izabela Harabula, Lars Wittler, Martin Franke, Daniel M. Ibrahim, Bjort K. Kragesteen, Malte Spielmann, Stefan Mundlos, Dario G. Lupiánnez, and Guillaume Andrey

INTRODUCTION

Genomic structural variations (SVs) are large-scale structural differences in the genomic DNA ranging in size from a few kilobases to entire chromosomes. SVs may be unbalanced as in deletions, duplications, and insertions or balanced as in inversions and translocations or a combination thereof. SVs contribute to a large extent to the variability of our genome and are often associated with disease (Stankiewicz and Lupski, 2010). When occurring within the coding sequence of genes, they can affect the protein sequence and thereby protein function or stability. When encompassing one or several coding units, deletions or duplications lead to changes in gene dosage. Furthermore, it was shown that SVs could interfere with gene regulation by disrupting genomic architecture necessary for proper enhancer-promoter interactions. Such rearrangements can result in loss of WT interactions and/or ectopic enhancer-promoter interactions, thereby resulting in gene mis-expression (Montavon et al., 2012; Spielmann et al., 2012; Spielmann and Mundlos, 2013). To discriminate between these multiple effects and to study their complex molecular pathology in vivo modeling of SVs is required.

SVs can be induced by radiation via the induction of double-strand breaks, but this is a random process that cannot be targeted to specific genomic regions. So far, allelic series involving the remodeling of large DNA segments have been obtained from the cis and trans recombination of targeted loxP sites in mouse chromosomes (Hérault et al., 1998; Ruf et al., 2011; Spitz et al., 2005). However, these approaches are time consuming and laborious, involving the targeting of loxP sequences and subsequent mouse crossings with Cre driver animals, a procedure taking at least 12 months. More recently, Zinc Finger nucleases (ZNF) or transcription activator-like effector nucleases (TALENs) have been shown to induce targeted SVs of several hundred kilobases in mammalian cells. In vivo inversions and deletions up to 1 Mb could be obtained in zebrafish (Bauer et al., 2013; Gupta et al., 2013; Lee et al., 2012; Xiao et al., 2013). However, to our knowledge, these methods were not applied to generate large structural variants in mice. This might be due to the process of generating specific ZNF nucleases or TALEN that is rather slow and requires expert knowledge.

The recent development of the CRISPR/Cas technology has led to a wider use of genome editing, opening new possibilities to engineer SVs in various model systems (Hsu et al., 2014; Peng et al., 2014; Wang et al., 2013). Indeed, two synthetic guide RNAs (sgRNAs), targeted at two different positions of a chromosome, are able to induce inversions and deletions through non-homologous end joining (NHEJ). For instance, Xiao et al. (2013) were able to induce 1.5 kb deletions in zebrafish embryos by injecting sgRNAs and Cas9 directly into zygotes. Similarly, larger structural rearrangements involving deletions, inversions, and translocations were obtained in HEK293 or in murine erythroleukemia (MEL) cell lines (Canver et al., 2014; Choi and Meyer, 2014).

SUMMARY

Structural variations (SVs) contribute to the variability of our genome and are often associated with disease. Their study in model systems was hampered until now by labor-intensive genetic targeting procedures and multiple mouse crossing steps. Here we present the use of CRISPR/Cas for the fast (10 weeks) and efficient generation of SVs in mice. We specifically produced deletions, inversions, and also duplications at six different genomic loci ranging from 1.1 kb to 1.6 Mb with efficiencies up to 42%. After PCR-based selection, clones were successfully used to create mice via aggregation. To test the practicability of the method, we reproduced a human 500 kb disease-associated deletion and were able to recapitulate the human phenotype in mice. Furthermore, we evaluated the regulatory potential of a large genomic interval by deleting a 1.5 Mb fragment. The method presented permits rapid in vivo modeling of genomic rearrangements.
In this study, we applied the CRISPR/Cas technology in mouse embryonic stem cells (ESCs) and developed a 10-week protocol that we named CRISVar (CRISPR/Cas-induced structural variants) to efficiently produce deletions, inversions, and duplications in mice. We were able to rearrange targeted genomic intervals ranging from 1 kb to 1.6 Mb using the CRISPR/Cas system in ESCs (Figure 1A). Moreover, we show that ESCs harboring these mutations can be used to create chimeric animals. Finally, we show at the *Laf4* and *Epha4* loci the ability of CRISVar to model human pathogenic SVs and to evaluate the regulatory influence of a large gene desert, respectively.

**RESULTS**

**CRISPR/Cas for the Induction of SVs in Mouse ESCs**

To test the ability of CRISPR/Cas to generate structural variants in mice, we first targeted the pleiotropic and developmentally associated locus *Pitx1*. We designed two sgRNAs in the gene desert adjacent and telomeric to *Pitx1* promoter. The two-targeted sites are separated by 232 kb of intermediate gene desert. After transfection of ESCs with the CRISPR constructs, we screened 288 clones for inversions, deletions, and duplications using a PCR-based approach (Figures 1B and S1). We detected
Table 1. Synthesis of Targeted SVs Efficiencies

<table>
<thead>
<tr>
<th>Name of the Locus</th>
<th>H2afy</th>
<th>Bmp2</th>
<th>Ihh</th>
<th>Ptx1</th>
<th>Laf4</th>
<th>Epha4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size of the rearranged region</td>
<td>1,189 bp</td>
<td>3.7 kb</td>
<td>12.6 kb</td>
<td>32 kb</td>
<td>353 kb</td>
<td>1.672 Mb</td>
</tr>
<tr>
<td>Deleted</td>
<td>11 (3.8%)</td>
<td>12 (6.3%)</td>
<td>121 (42%)</td>
<td>9 (3.1%)</td>
<td>38 (13.2%)</td>
<td>4 (2.1%)</td>
</tr>
<tr>
<td>Inverted (two breakpoints mapped)</td>
<td>2 (0.7%)</td>
<td>3 (1.6%)</td>
<td>7 (2.4%)</td>
<td>3 (1%)</td>
<td>12 (4.2%)</td>
<td>3 (1.6%)</td>
</tr>
<tr>
<td>Inverted (one breakpoint mapped)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>9 (3.1%)</td>
<td>8 (2.8%)</td>
<td>20 (6.9%)</td>
<td>1 (0.5%)</td>
</tr>
<tr>
<td>Duplicated</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2 (0.7%)</td>
<td>81 (28.1%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Deleted/inverted (one breakpoint mapped)</td>
<td>1 (0.3%)</td>
<td>1 (0.5%)</td>
<td>3 (1.0%)</td>
<td>2 (0.7%)</td>
<td>0 (0%)</td>
<td>11 (3.8%)</td>
</tr>
<tr>
<td>Deleted/inverted (two breakpoints mapped)</td>
<td>0 (0%)</td>
<td>1 (0.5%)</td>
<td>0 (0%)</td>
<td>9 (3.1%)</td>
<td>1 (0.3%)</td>
<td>9 (3.1%)</td>
</tr>
<tr>
<td>Inverted/duplicated (two breakpoints mapped)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2 (0.7%)</td>
<td>17 (5.9%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Inverted/duplicated (one breakpoint mapped)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (0.3%)</td>
<td>4 (1.4%)</td>
<td>28 (10%)</td>
</tr>
<tr>
<td>Number of screened clones</td>
<td>288</td>
<td>192</td>
<td>288</td>
<td>288</td>
<td>288</td>
<td>192</td>
</tr>
<tr>
<td>Successfully aggregated clones</td>
<td>one deletion - - one inversion two deletions one inversion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Numbers and percentages are indicated for each type of rearrangements at every locus. Deletions and inversions were found in all cases. Duplications originated from translocations between homologous chromosomes; thereby one of the chromosomes loses a copy of the rearranged DNA while the other gains it. In our experiments, duplications segregated with a deletion in only half of the cases. In the others, we found duplications segregating with an inverted or a WT homologous region. A similar observation was described using ZNF nucleases in mammalian cells (Lee et al., 2012).

After successfully implementing CRISVar at the Ptx1 locus, we targeted other regions of the genome to rule out the possibility that the observed effects were locus specific. We aimed at regions near the genes H2afy, Bmp2, Ihh, Laf4 and Epha4 (Figure 2) with rearrangements of 1.1 kb, 3.7 kb, 12.6 kb, 353 kb, and 1.6 Mb, respectively (Table 1). Interestingly, the frequency at which SVs occurred was different from one locus to another without evident relation to the genomic distance separating the two CRISPR sites. We observed deletions and inversions at all sites, although with a variation in their frequency of around 10-fold between experiments. Duplications were observed at three of these five rearranged loci.

Next, we wanted to test whether the here-produced clones could be used to create viable animals. We aggregated ESCs from eight selected clones to produce animals carrying deletions, duplication, and inversions at the above-mentioned loci. We were able to produce highly chimeric animals for all rearrangement types and loci (Table 1). In our hands, all aggregated ESC clones performed to produce chimeric animals.

A 353 kb Intragenic Deletion of Laf4 Recapitulates a Human Malformation Syndrome

In order to confirm the effectiveness of this genetic tool to generate mouse models, we aimed at reproducing a human disease-associated SV of unknown pathogenicity (Steichen-Gersdorf et al., 2008). The SV had occurred de novo, but its functional relevance remained unclear. The patient suffered from multiple malformations, including shortening of the femur, aplasia of the fibula, a triangular tibia, and three toes (Figure 3B). Because of the tibia malformation, the diagnosis of a Nievergelt-like syndrome was made. To evaluate the function of LAF4, we first produced a conventional knockout by targeted recombination. We could not observe a phenotype in homozygous mice, showing that inactivation of LAF4 cannot reproduce the human phenotype in the mouse (data not shown). We therefore re-evaluated the published data and found that the deletion was smaller than...
originally thought, encompassing only nine exons of the LAF4 gene and not the entire genomic region. The breakpoints of this 500 kb deletion were mapped in introns 3 and 12, thus eliminating exons 4 to 12 of LAF4 without introducing a frameshift. The predicted product is a truncated 850 amino acids protein lacking a domain predicted to be involved in transcription activation (Ma and Staudt, 1996). To specifically model this SV, we designed sgRNA at the homologous regions in the mouse genome corresponding to the human breakpoints (Figures 2 and 3A). The screening of ESCs revealed clones with heterozygous and homozygous deletions of the 353 kb homologous region. In chimeric animals generated from heterozygous as well as homozygous clones, we observed a short zeugopod in the upper limb (Figure 3C). The lower-limb abnormalities recapitulated part of the human phenotype showing a small, triangular ossification center of the tibia and severe hypoplasia of the fibula (Figure 3C). In contrast to the patient, the mice had polydactyly of the feet with incomplete penetrance. Thus, the specific intragenic deletion in LAF4 generated by CRISPR/Cas recapitulated the patient’s phenotype, demonstrating the pathogenicity of this SV. It is likely that the truncated protein exerts a dominant-negative effect, thereby leading to the observed abnormalities of bone formation.

A 1.5 Mb Deletion of a Gene Desert Encompassing Epha4 Results in Hindlimb Hopping Gait

Large gene deserts are thought to harbor regulatory elements and often surround developmentally active genes. We aimed at using the here-described method to challenge the integrity of such a locus. At the Epha4/Pax3 extended locus, two gene deserts centromeric and telomeric to Epha4 might regulate either one or both genes. Epha4 has been shown to control neuronal guidance in hindlimbs. Pax3 is a transcription factor with important function in the migration of muscle progenitors in the limbs and neural crest migration. Mice with mutations in Pax3 show pigmentation defects, early lethality due to heart defects, spina bifida, and exencephaly. We induced a deletion extending 1 Mb centromeric and 350 kb telomeric to the Epha4 transcription unit (Figure 4A). Heterozygous clones were used to produce mice with a 1.5 Mb deletion, which were subsequently bred to homozygosity. We observed a neurological phenotype in these animals, consisting of a hopping gait, as previously described for Epha4 loss of function (Figure 4B; Movie S1) (Dottori et al., 1998). Other abnormalities were not observed in these mice. In particular, we did not detect any of the Pax3-associated phenotypes such as pigmentation defects and abnormalities of the spine or the brain. Furthermore, in situ hybridization of Pax3 in deletion embryos revealed a normal pattern of expression (data not shown). Our finding that the 1.5 Mb deletion results in a full recapitulation of the Epha4 knockout without additional abnormalities indicates that the region does not contain elements essential for Pax3 regulation.

DISCUSSION

CRISPR/Cas is a rapid and efficient method to specifically edit genomes (Cho et al., 2013; Wang et al., 2013). Here we show that by targeting two sgRNAs at two distal genomic sites we can induce the rearrangement of the intermediate DNA fragment up to 1.6 Mb in ESCs. We show that duplications, deletions, and inversions can be obtained in different clones, but they may also occur in various combinations together. We observed extensive variability in the targeting efficiency from one locus to another. Although the size of the re-arranged DNA region was previously shown to play an important role in targeting efficiency, we were not able to confirm this relationship (Canver et al., 2014). One of the possibilities is that the chromatin “openness” at the CRISPR target site affects the cutting efficiency of Cas9 and thereby increases the probability of NHEJ between the two distal breakpoints. In fact, it has been shown that Cas9 binds better at open chromatin sites, where it has higher off-target effects (Kuscu et al., 2014). In this view, the epigenetic status of the targeted region might influence the overall efficiency of the rearrangement.

Furthermore, we frequently found sequence variations at the SV breakpoints. The randomness of NHEJ after a double-strand break is likely to be accountable for this diversity. This might also explain why several inverted clones could be mapped at a single breakpoint only: the other one being too extensively modified for a proper PCR reaction to occur. This variability in the NHEJ process also pinpoints to the limitations of PCR-oriented allele detection; indeed, several false-negative clones with deletions, duplications, or inversions may have been missed because of

Figure 2. Rearrangements at Other Loci

Summary of the other alleles in which SVs were induced. The alleles are ordered from the smallest generated SV (top: H2afy) to the largest (bottom: Epha4). Sizes of the rearranged regions are indicated next to the locus name.
extensive rearrangements at their breakpoints, inhibiting proper PCR reactions to occur. Because of this uncertainty, these clones were not processed further and discarded. Given the high efficiency of CRISVar, this was not considered a problem. We also observed that WT alleles might harbor indels induced by CRISPR/Cas. Thus, the WT allele should also be sequenced at the target site.

The limitations of CRISVar can be efficiently bypassed by a proper screening of ESC clones and subsequent mouse crossing. Such extensive screening might be a limiting step if the method would be applied directly to mouse zygotes. However, it is unclear how efficient microinjections of sgRNAs and Cas9 directly into embryos would be. Further testing is thus required to compare the efficiency of both approaches. The advantages of producing SVs in ESCs are on the one hand the possibility of retargeting to introduce additional mutations and on the other hand the direct use of these cells in culture to investigate the effects of SVs.

Thus, with CRISVar, large and small SVs of various types can be produced with high efficiency and in a short time period to study the effect of genomic rearrangements and their pathogenic effect. As shown at the Laf4 locus, human SVs may exert their effects in a specific manner, necessitating an exact reproduction in the mouse genome to evaluate their pathogenicity. CRISPR/Cas-mediated genome editing can be used to efficiently create such rearrangements. Finally, as shown in the Eph4 example, the method can be used to challenge the integrity of genomic regulatory units, as they have recently been proposed (Marinić et al., 2013).

**EXPERIMENTAL PROCEDURES**

**CRISPR sgRNAs Selection and Cloning**

SgRNAs were designed flanking the regions to rearrange. We used the [http://crispr.mit.edu/](http://crispr.mit.edu/) platform to obtain candidate sgRNA sequences with little off-target specificity. Complementary strands were annealed, phosphorylated, and cloned into the BbsI site of pX459 or pX330 CRISPR/Cas vector (Addgene; see Table 1).

**ES Cell Culture and Transfection**

G4 ESCs (300,000) (George et al., 2007) were seeded on CD1 feeders and transfected with 8 μg of each CRISPR construct using FuGENE technology (Promega). When the construct originated from the pX330 vector, cells were cotransfected with a puromycine-resistant plasmid. PX459, in contrast, already contains a puromycine-resistance cassette. After 24 hr, cells were split and transferred onto DR4 puro-resistant feeders and selected with puromycin for 2 days. Clones were then grown for 5 to 6 more days, picked, and transferred into 96-well plates on CD-1 feeders. After 2 days of culture, plates were split in triplicates, two for freezing and one for growth and DNA harvesting.

Positive clones were thawed and grown on CD-1 feeders until they reach an average of four million cells. Three vials were frozen, and DNA was harvested from the rest of the cells to confirm genotyping.

**PCR-Based Genotyping**

Primers were designed on both sides of sgRNAs targets, at a distance of 100–300 bp from the cutting site. Each allele has thus a set of four primers: T1(fwd)/T2(rev) amplifying one targeted site and T3(fwd)/T4(rev) amplifying the other (Figure S1; Table S1). Deletions were mapped using T1 and T4 primers while inversions were detected with T2/T4 for the proximal breakpoints and T1/T3 for the distal one. Finally, duplications were detected with T2/T3 primers. Positive PCR bands were Sanger sequenced using one of the amplification primers.

---

**Figure 3. A 350 kb Deletion at the Laf4 Locus Recapitulates Nievergelt Syndrome**

(A) Schematic of the extended Laf4 genomic region. CRISPR/Cas binding sites are depicted in yellow. The induced deletion (350 kb) is shown below.

(B) Radiograph of patient lower leg showing triangular tibia, missing fibula, and oligodactyly.

(C) Skeletal preparations of WT and mutant embryos at E16.5. Cartilage is shown in blue and bone in red. Note the short ulna and radius with small ossification centers in forelimb and triangular tibia (magnification shown on right) and severely hypoplastic fibula in hindlimb.

See also Table S2.
Screening for *Laf4* Homozygote Deletion Clones

Quantification of copy-number variation of the deleted interval in comparison to other genomic positions was performed with a set of qPCR primers (see Table S2).

Mouse Aggregation

A frozen ESC vial was seeded on CD-1 feeders, and cells were grown for 2 days. Mice were generated by diploid or tetraploid aggregation (Artus and Hadjantonakis, 2011). All animal procedures were in accordance with institutional, state, and government regulations (LAGeSo).

Leg Footprints

Adult mice were encouraged to run over a white paper sheet after their hindfeet were soaked in blue ink.

Skeletal Preparation

E16.5 animals were processed and stained as described previously (Mundlos, 2000).

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Supplemental Experimental Procedures, one figure, two tables, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.01.016.

**AUTHOR CONTRIBUTIONS**


**ACKNOWLEDGMENTS**

We thank Phillip Grote and Tracie Pennimpede for discussions, advice, and sharing reagents. We thank Asita Steige for technical assistance. G.A. is supported by an Early Postdoc.Mobility grant from the Swiss National Science Foundation. D.G.L. is supported by the Fundación Alfonso Martín Escudero. This research was supported by a grant from the Deutsche Forschungsgemeinschaft to S.M.

**REFERENCES**


