



OPEN

Broad spectrum of CRISPR-induced edits in an embryonic lethal gene

Kayla T. B. Fuselier¹, J. Michael Salbaum² & Claudia Kappen^{1,3}✉

Mendelian genetics poses practical limitations on the number of mutant genes that can be investigated simultaneously for their roles in embryonic development in the mouse. While CRISPR-based gene editing of multiple genes at once offers an attractive alternative strategy, subsequent breeding or establishment of permanent mouse lines will rapidly segregate the different mutant loci again. Direct phenotypic analysis of genomic edits in an embryonic lethal gene in F0 generation mice, or F0 mouse embryos, circumvents the need for breeding or establishment of mutant mouse lines. In the course of genotyping a large cohort of F0 CRISPRants, where the embryonic lethal gene *T/brachyury* was targeted, we noted the presence of multiple CRISPR-induced modifications in individual embryos. Using long-read single-molecule Nanopore sequencing, we identified a wide variety of deletions, ranging up to 3 kb, that would not have been detected or scored as wildtype with commonly used genotyping methods that rely on subcloning and short-read or Sanger sequencing. Long-read sequencing results were crucial for accurate genotype–phenotype correlation in our F0 CRISPRants. We thus demonstrate feasibility of screening manipulated F0 embryos for mid-gestation phenotypic consequences of CRISPR-induced mutations without requiring derivation of permanent mouse lines.

Neural tube defects (NTDs) are developmental abnormalities of neural tube closure, and are thought to have multifactorial origins in gene–gene and gene–environment interactions¹. The neural tube is formed during the morphogenetic process known as gastrulation, and becomes closed at the end of neurulation. Disruptions in both of these processes are likely to cause embryonic defects, which—in severe cases—are lethal. In order to assess the contribution of candidate genes to the pathogenesis of NTDs in mouse embryos during gastrulation and neurulation, we employed a transient transgenic approach to mutagenize genes using the CRISPR/Cas9 system². The approach is called transient, as the phenotype resulting from the genetic manipulation is analyzed directly in the F0 generation, without establishment of permanent mouse lines^{3,4}. The advantage of such a strategy is that it allows for the screening of genes whose bi-allelic disruption may cause embryonic lethality, precluding the derivation of permanent mouse lines⁵. It has been estimated that 30% of murine null (knock-out) alleles are embryonic lethal in homozygous configuration⁶, the incidence of lethality in heterozygous configuration (haplolethal) is currently unknown.

Because each F0 gene-edited individual represents a unique genetic event, multiple individuals need to be generated and analyzed for morphogenetic defects to yield results that can be generalized over multiple distinct mutations by which the individuals might differ⁷, depending on accuracy of the CRISPR methodology at the given gene locus. While this can feasibly be achieved with transgenic mouse technology in two or three micro-injection sessions per individual gene, the incidence and severity of phenotypes in the resulting genetically manipulated F0 progeny also allows estimates about the phenotypic penetrance of mutations at a chosen locus. In addition, a high-throughput transgenic approach enables the simultaneous targeting of multiple genetic loci⁸, thereby facilitating studies of gene–gene interactions during neural tube closure.

To establish feasibility of this methodology, we chose to mutagenize a gene whose function in mouse development has been well-characterized⁹, and whose mutant phenotypes can be visually ascertained at mid-gestation: the *brachyury* locus, which encodes the T-Box transcription factor T. Mice with a homozygous deletion of the *T/brachyury* gene are embryonic lethal by embryonic day 10, with defects in mesodermal derivatives, such as the notochord, presomitic and somitic mesoderm¹⁰. Reductions in T expression are associated with defective posterior mesoderm development, ranging from complete absence of posterior mesoderm¹¹ to absent tail, abnormal somite development, short tail or curly tail, abnormal neural tube closure, kinked neural tube and spina bifida^{12,13}.

¹Department of Developmental Biology, Pennington Biomedical Research Center/Louisiana State University System, 6400 Perkins Road, Baton Rouge, LA 70808, USA. ²Department of Regulation of Gene Expression, Pennington Biomedical Research Center/Louisiana State University System, 6400 Perkins Road, Baton Rouge, LA 70808, USA. ³Peggy M. Pennington Cole Chair in Developmental Biology, Pennington Biomedical Research Center/Louisiana State University System, 6400 Perkins Road, Baton Rouge, LA 70808, USA. ✉email: claudia.kappen@pbrc.edu

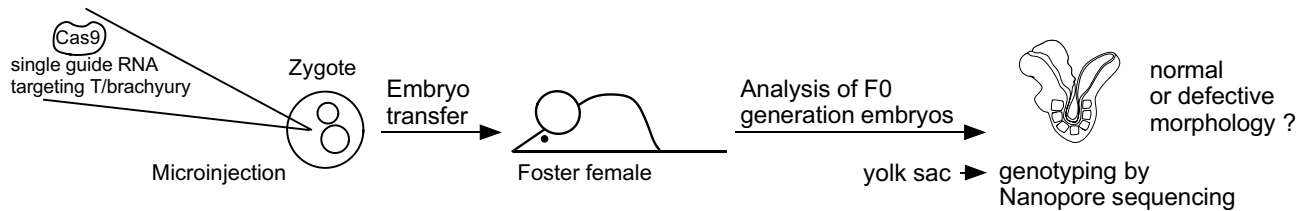


Figure 1. Transient transgenic strategy for CRISPR-induced mutagenesis. Schematic of the experimental approach: After microinjection of the Cas9/single-guideRNA ribonuclease complex into fertilized oocytes, the zygotes are transplanted into foster females. On gestational day E9.5, the morphology of individual F0 generation embryos is assessed, and the corresponding yolk sac tissue is collected for genotyping. This transient approach allows rapid screening of single, and potentially multiple, genes for developmental function without the need for breeding.

In vitro test of guide RNA for the T gene

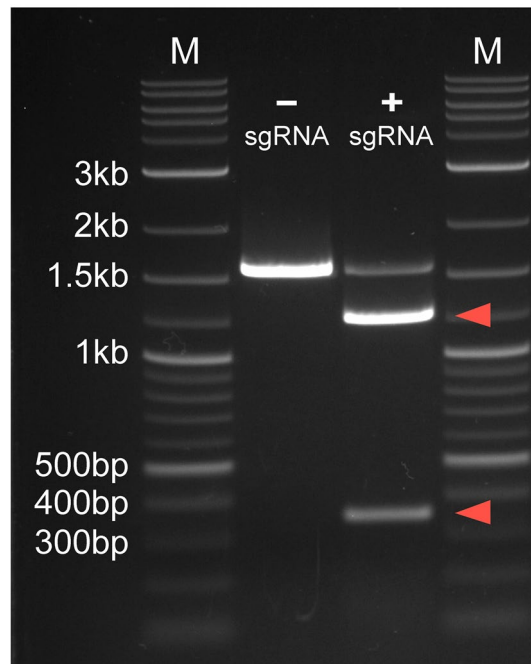


Figure 2. *In vitro* test of guide RNA for editing of the T gene. To detect CRISPR/Cas9-mediated cleavage at the intended targeted site *in vitro*, a 1592 bp region of the mouse T gene was PCR-amplified from genomic DNA and incubated with Cas9 protein and the single guide RNA (+ sgRNA). The control reaction contained the same substrate amplicon, and Cas9 protein, but no sgRNA (-sgRNA); the uncut PCR amplicon is close to 1.6 kb in size. Cleavage of the target amplicon by the sgRNA/Cas9 complex is detected by virtue of fragments of ~ 1250 bp and ~ 340 bp, respectively, in the sgRNA-containing sample (red arrowheads). M: Quick-Load® Purple 1 kb DNA Ladder (New England Biolabs, Ipswich, Massachusetts).

We here report that similar morphological anomalies were observed in mouse embryos where the T/brachyury gene was edited by CRISPR mutagenesis, carried out in a transient transgenic strategy that involves phenotyping and genotyping of individuals directly in the F0 generation. It was only by employing single-molecule long-read sequencing, as implemented in the Oxford Nanopore sequencing technology¹⁴, that we were able to detect the multiple CRISPR-edited mutant T/brachyury alleles in single F0 individuals, and to identify larger on-target deletions than have been reported from previous CRISPR mutagenesis efforts.

Results

The principal experimental approach of this study is depicted in Fig. 1. The gene chosen for targeting encodes the T-box transcription factor T, mutations of which are responsible for the brachyury phenotype in mice^{15,16}. The murine T gene consists of 8 exons and produces 3 known splice variants that are protein coding. We designed a guide RNA to target exon 3, as each of the known splice variants contain this highly conserved region. This exon codes for amino acid residues 158 through 202 of the T/brachyury protein, and encompasses a portion of the DNA binding-domain (amino acid residues 1–229). Thus, disruption of the third exon is expected to abolish transcription factor function. The efficacy of the specific guide RNA was ascertained *in vitro*, as shown in Fig. 2.

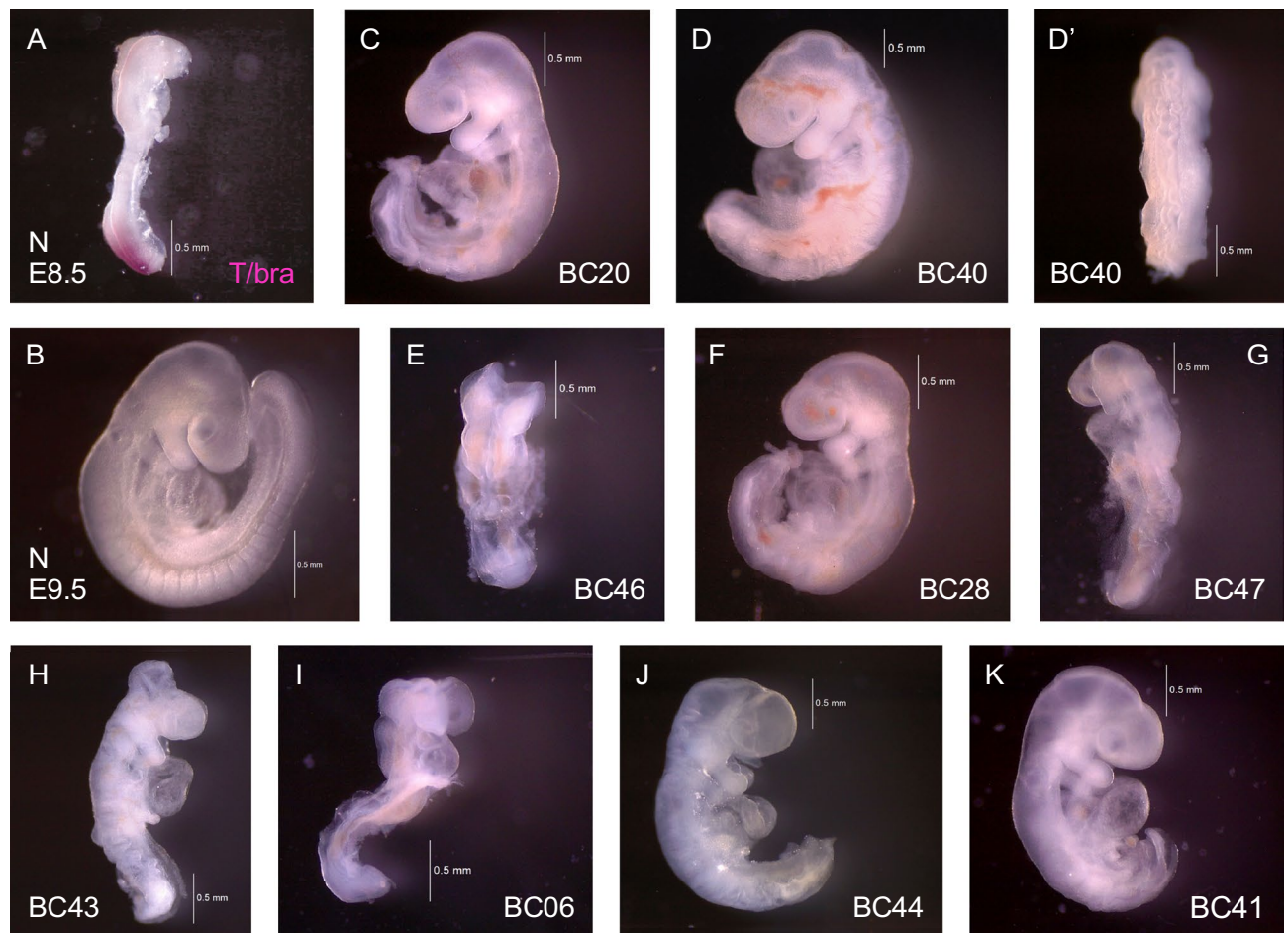


Figure 3. Morphological abnormalities of individual CRISPRants. Panels A and B show normal mouse embryos at E8.5 (A) and E9.5 (B). Based upon the earlier expression of T/brachyury in the primitive streak, notochord, and the posterior region (as shown by in situ hybridization in (A)), disruption of the coding region by targeting the third exon for CRISPR editing is expected to cause morphological abnormalities in somite formation, neural tube morphogenesis and reduced growth of the posterior. Panel B shows a normal mouse embryo with well-formed somite boundaries, and long posterior region. CRISPRants are designated according to their barcode numbers (see Table 1), and they are ordered in this figure by ascending content of mutant alleles. BC20 and BC40 (C, D, and D') carry only the $\Delta 11@-1 \Theta$ allele, likely in homozygous configuration, and they exhibit disrupted somite formation and an abnormal tail region, accompanied in BC40 by wavy neural tube (D'). (E) BC46 carries the $\Delta 11@-5 \Theta$ allele, and displays a disordered neural tube, as well as a somite-less and shortened posterior region. (F) BC28 contains both $\Delta 11@-1 \Theta$ and $\Delta 11@-5 \Theta$ deletions; while the neural tube has closed in the anterior region, somites are not discernable in the trunk and shortened posterior region, which also reveals a wavy neural tube. (G) BC47 harbors the longest 2728 bp and both 11 bp deletion alleles; the neural tube is disordered and wavy in the trunk, heart development is compromised and somites are absent in the posterior. (H) BC43 carries 3 deletions ranging from 719 to 998 bps, a 10 bp deletion and the one nucleotide insertion alleles; morphogenesis of somites and neural tube is abnormal, with wavy neural tube and shortened tail region. (I) BC06 contains 3 deletions over 1000 bps in size, and 691 bp and 100 bp deletion alleles; the entire trunk region is malformed and shorter, somites are absent, and heart development is also abnormal. (J) BC44 harbors 6 deletion alleles, and exhibits lack of somites, wavy neural tube from the hindbrain along the entire trunk, and a severely truncated tail region. (K) BC41 carries 6 different deletion alleles, and displays complete absence of somites and neural tube below the heart and in the posterior trunk.

We injected fertilized oocytes with Cas9 enzyme and a single-guide RNA targeted to the T/brachyury gene; these manipulated zygotes were transplanted into pseudopregnant foster females and allowed to develop until gestational day E9.5. Embryos were harvested, and the yolk sac associated with each individual was used for genotyping of the targeted region. Morphological analysis of embryos revealed a spectrum of phenotypes (representative examples are shown in Fig. 3), encompassing defects that have been described in the literature for various heterozygous and homozygous configurations of T-alleles in mutant mouse strains: wavy and kinked neural tubes, defective neural tube closure, malformed somites, short tails, absent tails, malformed posterior and trunk regions^{10,13,17,18}. All recovered embryos displayed one or more of these morphogenetic abnormalities. Notably, we found considerable phenotypic variability between individual F0 embryos, suggesting that the underlying gene editing events may also vary between gene-edited individuals, which we refer to as CRISPRants.

PCR Detection of larger deletions in CRISPRants

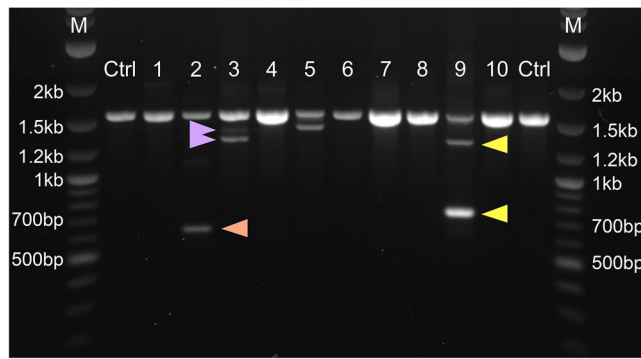


Figure 4. PCR amplification of genomic DNA reveals larger deletions. A 1592 bp region spanning the third exon of the T gene was amplified from genomic DNA of 10 individual CRISPRants (numbered 1 through 10) and genomic DNA from unmanipulated mice (Ctrl: control). The orange triangle identifies an amplicon of approximately 650 bp length in sample 2, indicative of a large deletion with an estimated length of 950 bp. Purple arrows point to bands in sample 3 of ~1.4 and ~1.3 kb size, respectively, indicating putative deletions of ~200 and ~300 bp size. Yellow arrows mark amplicons in sample 9 of ~1.3 kb and ~800 bp size, predicting the presence of deletions of ~300 bp and ~800 bp, respectively. M: Lane of marker DNA fragments of indicated sizes.

In order to genotype the putative edited target region in the T/brachyury gene CRISPRants, we initially planned to employ a widely used combination of PCR-based amplification, followed by subcloning and DNA sequencing. PCR amplification of the targeted region was expected to produce amplicons of ~1.6 kb from genomic DNA (see Fig. 4, Ctrl samples). Unexpectedly, however, visualization of the PCR products after gel electrophoresis revealed the presence of much smaller than expected amplicon sizes in several of the samples, and often multiple smaller bands within a single sample (see Fig. 4 for an example/s). We interpreted the smaller products to reflect larger-than-expected deletions, and the multitude of bands to reflect the presence of multiple events within a single CRISPRant. For example, 3 bands of approximate sizes of 0.8 kb, 1.3 kb and 1.6 kb are detected in sample #9 (Fig. 4), suggesting that deletions of potentially 0.8 kb and 0.3 kb could be present. In order to circumvent the laborious steps of subcloning for separation of distinct molecular species in the amplicon mixture, and subsequent sequencing of numerous clones for each sample/individual, we employed single-molecule long-read sequencing of amplicons from the target region, as afforded by the Nanopore sequencing approach.

Again, PCR was employed to amplify the region of interest, this time increasing the amplicon length to 6.6 kb around the targeted cut site Θ (theta), and used to prepare Nanopore sequencing libraries. Sequences were aligned to wildtype T locus control samples run alongside the experimental samples, and to the reference genome sequence. The software pipeline CRISPResso2 was used to visualize the sequence alignments and insertion or deletion (INDEL) events. Position of INDEL events were labelled in accordance to their distance from Θ , with negative values on the 5' left side, and positive values designating the 3' right side. A complete listing of the 140 editing events observed amongst 40 CRISPRants for T/brachyury is presented in Table 1.

These 140 editing events represent predominantly deletions (Δ), ranging from 1 to 2728 bps in size. Most commonly observed were 4 editing events per T CRISPRant (in 16 of 40 embryos), with some CRISPRants harboring fewer, or more - up to 6-distinct mutations (Fig. 5, and Table 1). The only insertion detected was a 1 bp insertion at position +1 relative to Θ (Fig. 6), which occurred in 10 (25%) of the CRISPRants. All other events were deletions.

Of the 130 deletion events, 80 unique alleles could be recognized; of those, the 6 deletion events that occurred in multiple embryos are shown in Fig. 6. The two most common deletions were 11 bps in size, and started at either -5 relative to Θ (in 21 of 40 CRISPRants; 52.5%) or at -1 Θ (also in 21 of 40 CRISPRants; 52.5%), respectively. Eleven of our T CRISPRants had only one of these 11 bp deletions, respectively, and 10 T CRISPRants harbored both (32 embryos with one or both of the recurrent 11 base deletion at -5 Θ or -1 Θ). Additional deletions observed in more than one embryo were 23 bp or 10 bp in size (each found in 5% of CRISPRants), 6 bp (in 17.5% of CRISPRants), and one base pair deletions in 4 CRISPRants (7.5%).

Consistent with the results of the initial PCR analysis (Fig. 4), we also detected much larger deletions, ranging up to 2728 bps in size (see Table 1). A schematic of all deletions longer than 10 bps is shown in Fig. 7. Notably, although the longer deletions all spanned or involved the targeted cut site Θ , many were located asymmetrically, with 25 deletions located to a much greater extent towards the 5' side of Θ compared to the 33 deletions located to a much greater extent towards the 3' side. Thus, deletions predominantly "spread" in either of the two directions from the cut site, and "spreading" evenly in both directions is observed mostly to involve fewer than 10 nucleotides on either side. There were 46 unique "start sites" for deletions, including θ itself (in 7 distinct events); 56 distinct "end sites" were observed.

The consequences of the CRISPR-induced mutations render the T/brachyury protein dysfunctional. The targeted Exon 3 encodes 72 amino acids of the N-terminal DNA-binding domain, contributing residues 159–230 to the protein, which contains 436 amino acids overall. Multiple protein changes would be predicted from the CRISPR-mediated double strand breaks and repair. For examples, the 1 bp insertion at Θ creates a frame shift,

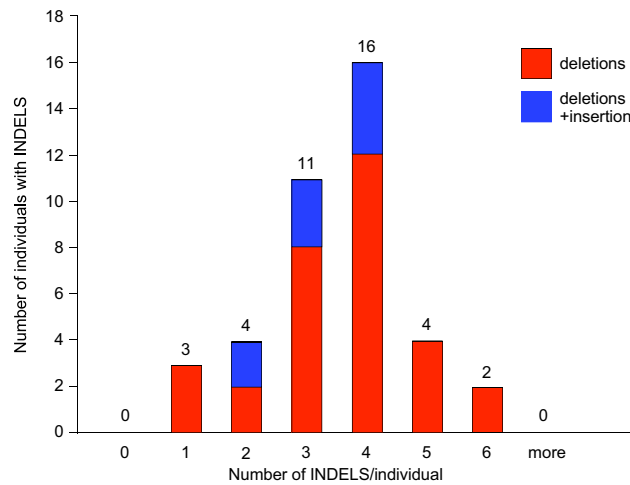


Figure 5. Occurrence of INDELS in individual CRISPs. The number of mutant alleles per individual was plotted. Red: deletions, Blue: insertions. 33 CRISPs contain three or more mutant alleles of the T/brachyury locus.

T/brachyury nucleotide sequence from -20 to Θ to +20 (5'-3')

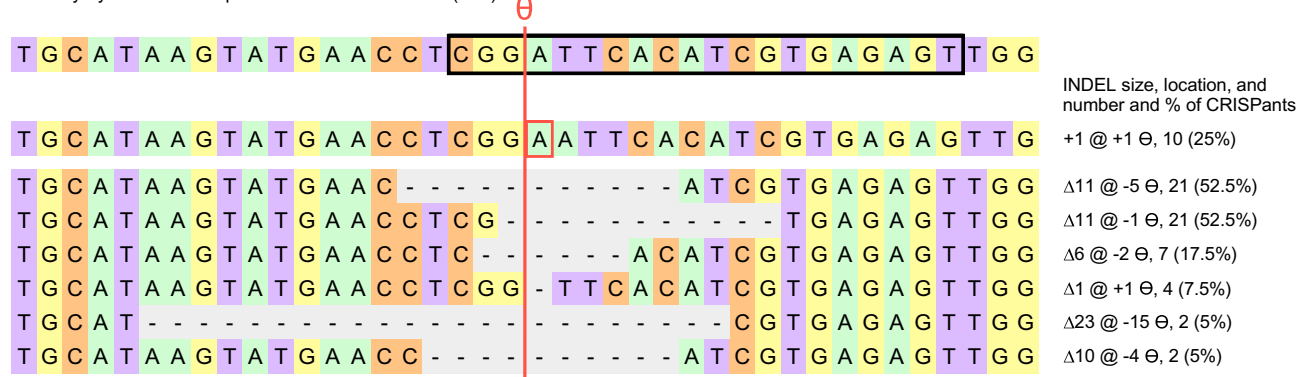


Figure 6. INDELS with occurrence in more than one individual. The sequence around Θ is shown, for 20 bp on either side. The black box frames the sequence complementary to the targeting portion of the single-guide RNA. The position of the cut site Θ is indicated by the vertical red line. Only one insertion was detected, the addition of an A, which was present in 10 independent CRISPs. Deletions that were found in 2 or more individuals are listed in descending order of frequency; 10 individual CRISPs had both 11 bp deletions.

changing 31 amino acids before a premature stop codon at position 201. The 11 bp deletion at -5 relative to Θ also shifts the frame, changing 29 amino acids before a premature stop codon at position 197. The 11 bp deletion at -1 Θ changes 27 amino acids before a premature stop codon. The 1 bp deletion at +1 Θ introduces a frame shift, changing 3 amino acids before a premature stop codon at position 172. While the 6 bp deletion at -2 Θ causes an in-frame loss of just 2 amino acids, those are located within the highly conserved DNA binding site. Obviously, the larger deletions will remove larger parts of the protein, critically altering protein structure and affecting DNA-binding function. In addition, deletions that extend more than 40 bp in the 5' direction of Θ will also remove the splice acceptor site at the 5' end of Exon 3, and deletions larger than 210 bp to the 3' of Θ will result in loss of the Exon 3 downstream splice donor site; both events, alone and in combination in the same allele, will alter the structure and length of any potential transcript, and of the protein produced, if any. As reduced expression as well as reduced activity of the T transcription factor is known to cause developmental defects in established genetic models, the CRISPR-induced mutations we detected in the F0 generation can explain the phenotypic alterations in all of the CRISPs embryos we recovered.

Intuitively, different CRISPR-edited alleles within the same individual were found to be represented in different fractions of all sequences for the individual, ranging from as much as 96.6% for the predominant allele in BC40 to as little as 2.32% for the minor allele in BC43, and wildtype sequence representation was found to range from 0.03% in BC21 to 32.8% in BC46 (Table 1). The proportions of alleles and their number in each individual sample are depicted in Fig. 8 as segments of pies. It is visually obvious that all CRISPs have unique genotype configurations. Minor fractions of alleles could conceivably be due to low cell numbers with the particular allele

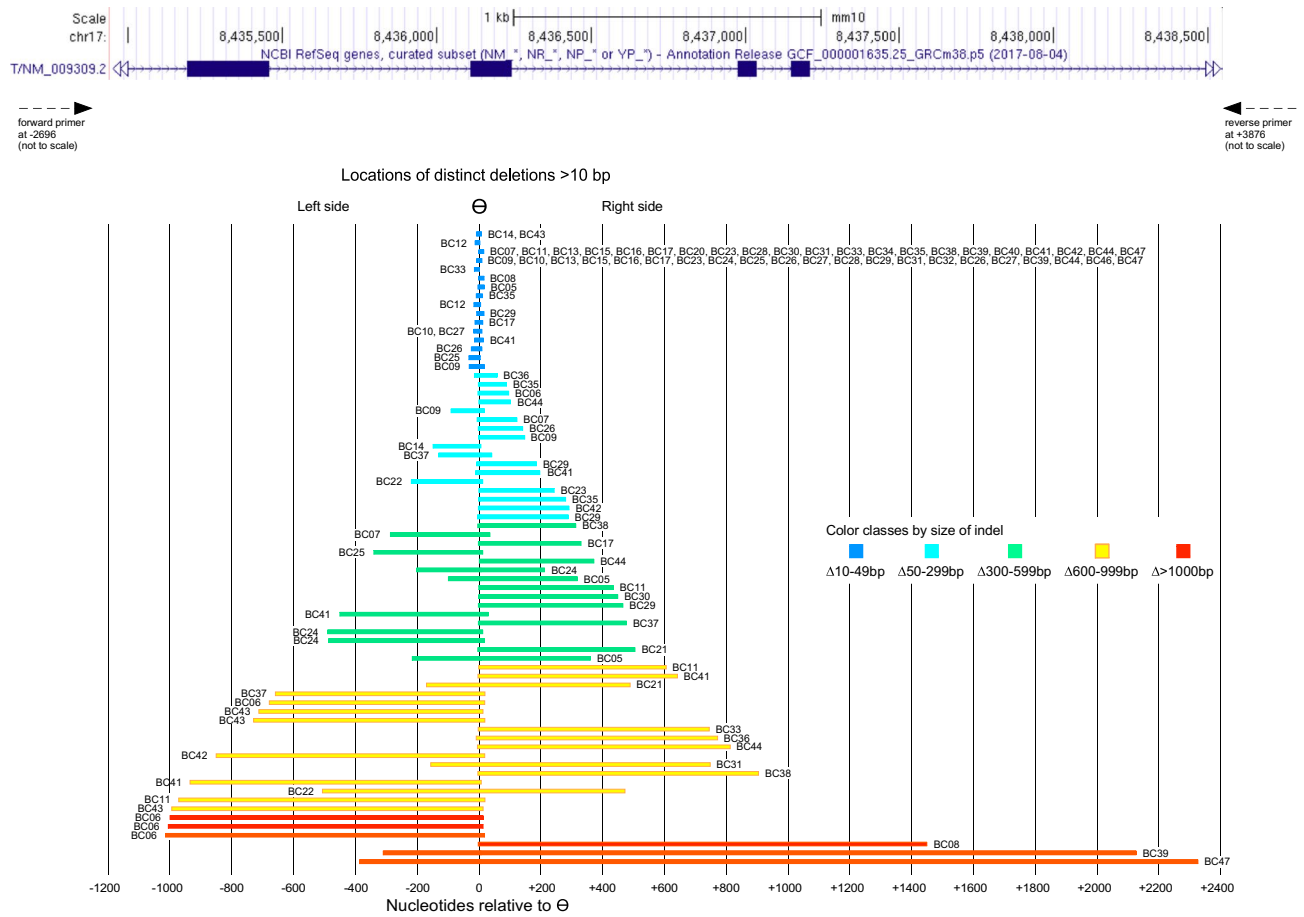


Figure 7. Locations of individual larger deletions in the T locus of CRISPRants. Each deletion of greater than 10 bp is shown as a single line, ordered by overall size from the bottom up. All deletions end at, originate at or span across Θ . Of note, a large fraction of the deletions are asymmetric relative to the cut site. Deletions are colored according to size, and the individual samples in which they appear are listed next to the respective line.

configuration within the genotyped tissue (here yolk sac), while more prominently represented alleles likely reflect (a) more predominant cell lineage(s). Due to the complexity of genetic configurations within any given individual and the low number of individuals with the smallest or largest numbers of mutant alleles, we were unable to establish an unequivocal correlation between number of mutant alleles and phenotype severity, or particular mutant alleles and phenotype severity. However, all recovered CRISPRants exhibited morphological anomalies consistent with on-target editing of the T/brachyury gene, demonstrating high phenotypic penetrance of the CRISPR-edited mutant alleles.

Discussion

Our goal in this work was to test the feasibility of a transient approach to identify new genes that—alone or in combination—contribute to morphogenesis and closure of the neural tube. Because neural tube defects originate early in post-implantation development, during the processes of gastrulation and neurulation, and because they can cause postnatal lethality, any genetic manipulation that targets this process may impact postnatal survival and preclude establishment of a live mutant strain. Therefore, we needed to screen for morphological defects in gene-edited embryos directly in the F0 generation, prior to birth. We chose to target a gene for which morphological phenotypes can be detected with ease, and at high penetrance under conditions of homo- as well as hetero-zygosity at the targeted locus, the gene encoding the T/brachyury transcription factor. Our results demonstrate that CRISPR-induced edits occurred in every individual recovered, and that they caused phenotypic anomalies that resemble those observed in conventional genetic mutations at the T/brachyury locus^{11–13,17,18}, including posterior growth defects, somite defects, wavy neural tube, open neural tube.

Historically, the transgenic mouse field operated under the convention that a minimum of three independent genetic events yielding the same outcome would be required to establish that the observed phenotypic outcome was caused by the genetic manipulation¹⁹, as opposed to some random parameter, which would not be expected to occur consistently among multiple individuals. However, in cases of incomplete penetrance or phenotype variability (also termed variable expressivity), larger sample numbers are necessary to document the phenotypic spectrum with convincing robustness. In the absence of prior evidence from, e.g. targeted mutant mouse lines, the optimal sample number can only be determined post-hoc. Phenotypic variability for T/brachyury mutant

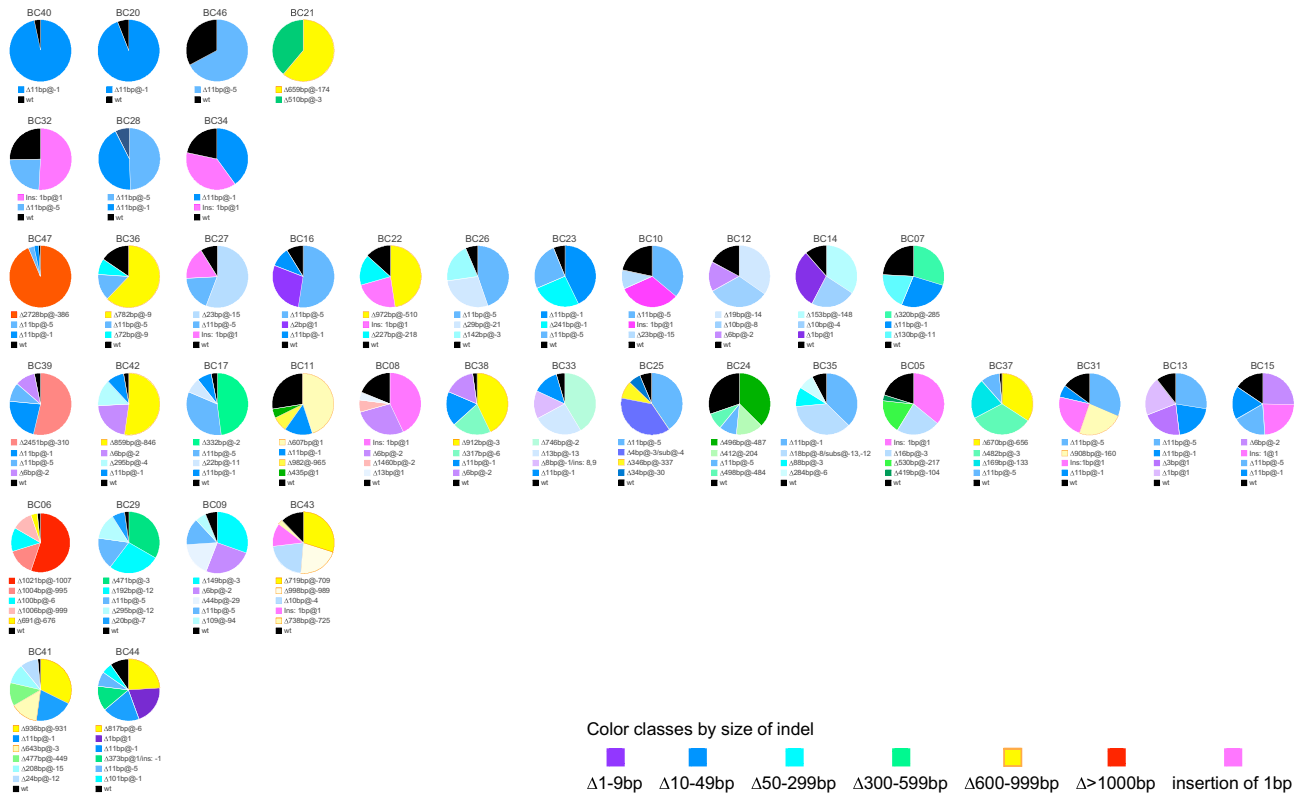


Figure 8. Representation of T/brachyury allele sequences within each individual. Colored segments depict the fraction of each edited allele among all sequences for the respective sample; wildtype sequences are indicated by black color. The color scheme corresponds to the same deletion size classes as defined in Fig. 7. Size of each INDEL (as listed in Table 1) is given in the legend below each sample. Samples are ordered by the total number of alleles detected (edited and wildtype), with individuals harboring 2 alleles at the top, and those with the most alleles at the bottom of the figure. Within each allele class, samples are ordered from left to right according to descending proportion of the most prominent edited allele.

mouse lines has been mentioned in the literature (although not quantified), but those reports used strains with genetic backgrounds different from FVB/N. In our CRISPR-mediated editing of the T/brachyury gene, 39 embryos from multiple pregnancies and different microinjection sessions were recovered, and they all exhibited malformations previously reported from multiple T-mutant strains^{10,13,17,18}. Thus, in recapitulating developmental defects with 100% penetrance within the known phenotypic spectrum, our work provides proof-of-concept for CRISPR-induced gene editing as a feasible strategy to assess gene function in embryonic development without the need for the establishment of permanent mouse lines.

By employing single-molecule long-read sequencing, we detected numerous distinct CRISPR-induced deletion events that involved longer regions of the targeted gene than initially expected. A large fraction of deletions was asymmetric relative to the cut site, a phenomenon that was also observed by others²⁰. Notably, the largest deletion we found was close to 3 kb in size. We would not have detected the larger deletions if we had followed the genotyping approaches that were standard in the field at the beginning of our project: Published protocols recommend sequencing of considerably shorter amplicons, such as for example in the 100–200 bp size range²¹. Under those criteria, (assuming sequencing spanned across Θ by 100 bp on either side), only 20 individuals out of our 40 CRISPs would have genotyped as CRISPR edited, and 54 of the 80 unique deletion events would have been missed. Similarly, following recommendations for amplification and sequence analysis included in the widely used Webtool CHOPCHOP (150–290 bp amplicons^{22,23}) or the CRISPR Gene Editing guide (recommending sequencing of 500–1500 bp, starting 200 bp upstream of Θ ²⁴), we would have detected only 30 out of the 81 unique INDEL events, and missed 23 unique deletions, respectively, scoring as successfully edited only 12 or 26 of our 40 CRISPs, respectively. Such low genotyping ascertainment would have been difficult to reconcile with the frequency of morphological defects seen in all of our CRISPs. Thus, the choice of long-read sequencing technology not only enabled us to identify genome edits in every recovered embryo, but was critical to our success in attributing abnormal phenotypes to underlying mutations.

Another advantage of the Nanopore MinION sequencer is the large number of reads that can be acquired, mitigating the inherently higher error rate of the technique by creating reliable consensus sequences²⁵. This is particularly useful for detection of CRISPR-mutagenesis in mouse strains or animal species for which no reference genome sequence exists^{26,27}. The largest deletion we observed encompasses 2728 base pairs, and we would have been able to recognize larger events, as long as they fell within the 6.6 kb size of our PCR amplicons. However, we cannot exclude the theoretical possibility that rare events extending beyond the ends of the amplified

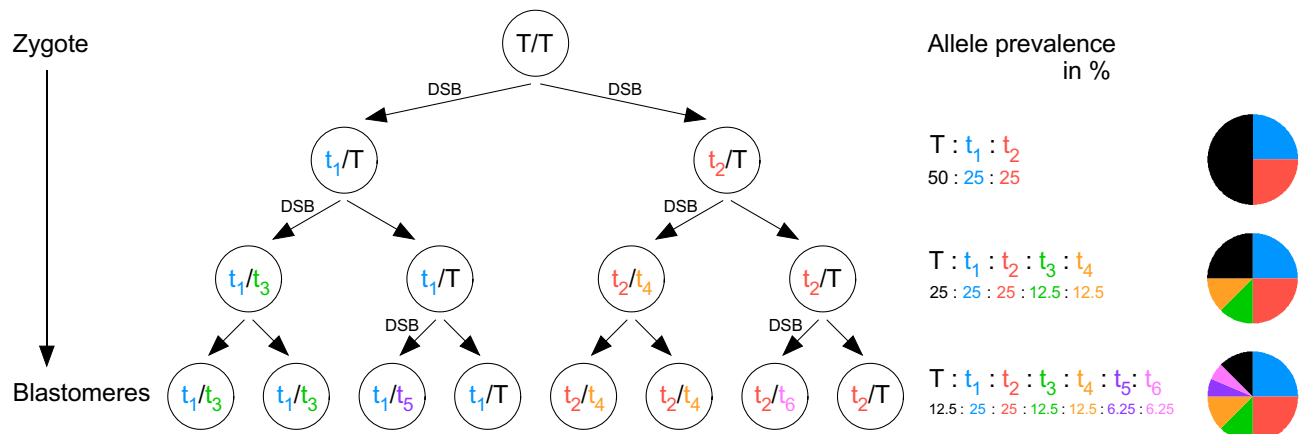


Figure 9. Successive CRISPR-induced mutations during the first embryonic cell divisions. Depicted is the most parsimonious theoretically possible scenario by which 6 mutant alleles in the same individual can be induced. T: wildtype allele, t₁₋₆: different mutant alleles (separate colors). As cells divide, DNA strand breaks (DSB) can occur on either sister chromatid, with retention of wildtype allele on the intact strand. The different alleles will be represented at different prevalence among all cells in successive generations, as visualized in the pie diagrams (colors correspond to the respective mutant alleles t₁₋₆).

product could be present in our CRISPs. For CRISPR transgenic approaches where larger amounts of genomic DNA are available than can be prepared than from yolk sac, such as for example when genotyping delivered F0 or F1 generation progeny, direct Nanopore sequencing of genomic DNA could potentially allow identification of even larger deletions, or structural alterations²⁸, and therefore be an attractive alternative to amplification- or cloning-based genotyping.

Our genotyping results demonstrate that we efficiently generated genome edits in every embryo recovered after microinjection and implantation. While we recovered 2 CRISPs with a single mutation and 4 CRISPs with just two mutations, it was the availability of a large collection of single-read sequences that enabled unequivocal identification of multiple mutant alleles in single individuals. The majority of CRISPs in our study harbored three or more mutant alleles, indicating that these latter individuals each were a mosaic of at least two (or more) genetically distinguishable cell populations. The extent to which cells with different genetic configurations are apportioned to trophoblast layer or the inner cell mass²⁹, which generates the embryo proper, is likely variable between CRISPs, and may account for the differential prevalence of particular alleles within an individual embryo. If live progeny were derived and bred to subsequent generations, only alleles that are present in germline cells would be transmitted to the next generation⁷. In addition, unequal representation of alleles among the sequences for a given CRISP could also be a reflection of the mixed cell composition of the yolk sac tissue that was the source of genomic DNA for genotyping. The yolk sac consists of two layers: a visceral endoderm layer derived from the primitive endoderm and an extra-embryonic mesoderm layer derived from epiblast cells that also generate tissues of the embryo proper³⁰. However, we cannot exclude the possibility that the embryo proper could harbor additional mutant alleles that are not represented in yolk sac cells.

In a given individual, up to 6 distinct mutant alleles were detected (see Table 1 and Fig. 8), in addition to low or appreciable fractions of wildtype sequence. Very minor fractions of wildtype sequence could be attributable to potential contamination during dissection with wildtype cells from the surrogate mother, such as blood or uterine decidual cells. But greater proportions of wildtype sequences suggest that a fraction of cells in an individual could be heterozygous for a CRISPR-induced mutation, due to incomplete editing at the single-cell stage. Yet, the large number of mutant alleles detected in 40 individual samples, with the majority of individuals harboring over 3 mutant alleles, and the fact that not a single animal with wholly unmodified genome was recovered, taken together indicate that CRISPR-mediated editing can proceed from the zygote through subsequent cell divisions. The most parsimonious explanation for the preponderance of more than 3 mutant alleles in our CRISPs is that editing occurred—at the least—during the cell divisions that produced the 2-cell and 4-cell stages; for the rarer occurrence of 5 and 6 mutant alleles in the same individual, editing would have to have continued at least through the 8-cell stage (Fig. 9).

If wildtype alleles are retained through one or more cell divisions, the multiple alleles in individuals are predicted to be present in unequal proportions even after the first three cell divisions (Fig. 9), which is consistent with our results (Fig. 8). In principle, differential representation of alleles could also be due to reduced efficiency of longer amplicon production in the PCR, but then one would expect that longer deletions be preferentially overrepresented, owing to shorter amplicon size. Instead, we find CRISPs with either of those scenarios: both large or small deletions can predominate in a given individual. For examples, in CRISPs BC06 and BC47 large deletions predominate, in CRISPs BC08 and BC31 small deletions are the most prominent allele, respectively (Table 1 and Fig. 8), arguing against technical artifact. Therefore, we attribute the differential representation of INDELS in individuals with multiple alleles to sequential editing events through successive cell divisions. Given that the majority (33 of 40) of our CRISPs harbors more than 3 distinct mutant alleles, the most plausible scenario is that most of their cells would be compound heterozygous for two different INDEL events.

In summary, employing single-molecule sequencing enabled us to identify numerous mutant alleles of the T/brachyury locus, a large fraction of which would have been missed when genotyping with the more commonly used shorter-read sequencing approaches. Furthermore, our study in T/brachyury-targeted CRISPRs provides proof-of-concept for the feasibility of screening manipulated F0 embryos for phenotypic consequences of CRISPR-induced mutations in genes where disruption of one or both alleles might cause embryonic lethality. The particular promise of this strategy is that it now allows simultaneous targeting of combinations of two or more genes with suspected functions in early embryonic development, and especially in morphogenesis of the neural tube.

Methods

Design of single guide RNA and in vitro testing for activity at the T/brachyury locus. The CRISPR single-guide RNA (sgRNA) was designed to target an essential coding region³¹ contained within all known isoforms of the T/brachyury gene mRNA, and its suitability was assessed with the online tool at crispr.mit.edu (<https://zlab.bio/guide-design-resources>). The targeting RNA sequence was selected with as few predicted off-target events as possible, and yielded a score of 88 (on a scale of 1–100 where higher is better) in the MIT Specificity test³² (ACTCTCACGATGTGAATCCG; NCBI location #17:8,436,143–8,436,163). In addition, the whole single-guide RNA used in this study contained the canonical PAM, CRISPR RNA [crRNA] and tracrRNA sequences for the Cas9 complex; it was purchased from Invitrogen (<https://www.thermofisher.com/us/en/home/life-science/genome-editing/genart-crispr/crispr-libraries/trueguide-grnas.html>). Re-analysis with the currently available tools CRISPOR³³, CRISPick^{34,35}, and CHOPCHOP v3²³ yielded the highest ranking for our guide RNA in all three algorithms.

DNA fragments PRC-amplified from FVB mouse genomic DNA were used to test the sgRNAs for activity in vitro. A 1592 bp segment of the T/brachyury gene was amplified by employing the following primers: T-Forward TGAGGAGGAAGGAGGCATTC (ENSEMBL Gene ID: ENSMUSG0000062327, Chromosome 17:8,434,905–8,434,925, GRCm38.p6) and T-Reverse GCTGGCGTTATGACTCACAG (ENSEMBL Gene ID: ENSMUSG0000062327, Chromosome 17: 8,436,477–8,436,497). PCR conditions consisted of 50 μ l reactions, with 50 ng of template, 200 nM primers, 250 μ M each dNTP and 1.25 U Taq DNA Polymerase (New England BioLabs). Amplification conditions employed Hotstart, with 2 min at 95 °C, followed by 30 cycles of: 30 s melting at 95 °C, 30 s annealing at 59 °C, 3 min elongation at 68 °C, and a final step of 10 min at 68 °C, after which the samples were kept at 4 °C.

The amplicon produced contained the intended target cut site Theta (Θ) in asymmetrical position, with offsets of approximately 1250 bp and 340 bp, respectively. The sgRNA was then tested on its ability to cut the double-stranded DNA amplicons in the presence of Cas9 Nuclease according to New England BioLab's protocol "In vitro digestion of DNA with Cas9 Nuclease, *S. pyogenes* (M0386)." The negative control did not contain the sgRNA for T/brachyury. The reactions consisted of: 30 μ l nuclease-free water, 4.5 μ l NEBuffer r3.1, 1.5 μ l 1 μ M Cas9 Nuclease, *S. pyogenes* (M0386), and 4.5 μ l 300 nM T sgRNA or 4.5 μ l nuclease-free water, respectively, and were incubated for 10 min at 25 °C, after which 4.5 μ l of 30 nM T amplicons were added for a total reaction volume of 45 μ l and further incubation for 30 min. The reactions were stopped by heating the samples to 65 °C for 5 min, then the DNA was ethanol precipitated, pellets were resuspended in TE buffer, mixed with Gel Loading Dye, Purple (6X) (NEB, B7024S), and resolved on a 1.1% agarose gel for separation of any fragments produced by the ribonucleoprotein complex (RNP). GelRed (Biotium) staining served to visualize DNA, and images were acquired with a BioRad Transluminator (Universal Hood II).

Mice. All mice utilized for this study were housed and maintained at Pennington Biomedical Research Center under specific pathogen-free (SPF) conditions, in individually vented cages.

Females of the inbred FVB/N mouse strain (Charles River Laboratories International, Inc.) were used as zygote donors, chosen for their breeding success, high oocyte yield and excellent microscopic visibility of fertilized oocyte pronuclei³⁶. These mice had a light cycle from 7 am to 7 pm and were superovulated before being mated. For superovulation, 5 IU of Pregnant Mare Serum Gonadotropin (PMSG, ProSpec-Tany TechnoGene Ltd) were injected intraperitoneally one hour past the middle of the light cycle two days before breeding, and 5 IU Human Chorionic Gonadotropin (hCG, ProSpec-Tany TechnoGene Ltd) were injected at 11am, approximately 1 h before mating was set up with a FVB/N male. The day of observation of a copulation plug was counted as gestational day E0.5. Females of the outbred CD-1 mouse strain (CrI:CD1(ICR), Charles River Laboratories International, Inc.) were used as surrogates, after mating with a vasectomized CD-1 male to induce pseudopregnancy. The female foster dams were used at > than 6 weeks of age, they weighed between 25 and 32 g, and they experienced the same light cycle as the donor mice.

Microinjection and transfer of zygotes. Single-cell zygotes were harvested from plugged FVB/N donors on E0.5 and were cultured in EmbryoMax® Advanced KSOM Embryo Medium (EMD Millipore) at 37 °C in an atmosphere containing 5% CO₂ before and after microinjection. Reagents to be microinjected were mixed fresh 30 min before injection and consisted of 25 ng/ μ l TrueGuide™ sgRNA for the gene of interest (T/brachyury) and 50 ng/ μ l TrueCut™ Cas9 Protein v2 (Invitrogen) diluted in EmbryoMax® Electroporation Buffer (EMD Millipore). The injection solution was stored on ice until use. Microinjections and transfer surgeries were performed as described by Behringer et al. in "Manipulating the Mouse Embryo"³⁷. A FemtoJet 4i Microinjector (Eppendorf) was employed for injections. Injected embryos were transplanted to the oviduct of pseudopregnant CD-1 females. A maximum of 18 injected zygotes were transferred into the infundibulum of a single uterine horn. Noon of the day of the embryo transfer transplantation surgery was designated day E0.5 of gestation. Surrogate females were allowed to carry until E9.5, when they were euthanized for embryo harvest.

Dissection, imaging, and fixation of embryos. Dissections of individual decidua were performed each in their own fresh dish of OptiMEM, and then the yolk sac and any extraembryonic tissues destined for genotyping were rinsed in separate wells of PBS in a 24-well plate. Embryos were imaged fresh in a plate of 5% FBS Opti-MEM® I Reduced Serum Medium (Life Technologies) under a Leica S9i stereo microscope with V4.12.0 Leica Application Suite software (Leica Microsystems Co.). Thereafter, the embryos were rinsed in ice-cold sterile PBS and fixed in 4% paraformaldehyde + 0.25% glutaraldehyde at 4 °C for 3 h. Fixed embryos were then dehydrated in methanol washes and stored at – 20 °C.

Isolation of genomic DNA for genotyping. For genotyping of genome editing events at the T/brachyury locus, visceral yolk sacs/visceral endoderm, chorion, and allantois (where available) were harvested from each embryo. These tissues were rinsed in sterile PBS buffer, stored individually in 50 µl PBS in microtubes, and frozen at – 20 °C until genomic DNA (gDNA) isolation by the Tail buffer/Proteinase K digestion protocol³⁸.

PCR amplification for sequencing. A fragment of 6572 bp size was PCR amplified from genomic DNA with approximately 3 kb on either side flanking the targeted cut site Θ . Amplifications were performed in 50 µl volume, with 50 ng of template, 200 nM primers, 250 µM each dNTP and 1.0 U Phusion High-Fidelity DNA Polymerase (New England BioLabs). Primer pairs for T CRISPs were T-Forward TCACCGAGAGGCAAT AAACC and T-Reverse GCTGGCGTTATGACTCACAG (Chromosome 17: 8,433,450–8,433,470, and Chromosome 17: 8,440,002–8,440,022, respectively; ENSEMBL Gene ID: ENSMUSG00000062327 GRCm38.p6). Amplification conditions were: Hotstart, with 1.5 min at 98 °C, followed by 28 cycles of 10 s melting at 98 °C, 30 s annealing at 67 °C, 6 min extension at 72 °C, with a final step of 10 min at 72 °C, and storage at 4 °C thereafter. When necessary, the reaction was performed twice to yield sufficient amounts of amplicon. Electrophoresis was employed to resolve the amplified products on 1% agarose gels with the 1 kb Plus DNA Ladder (Invitrogen) as size marker. GelRed (Biotium) staining was used to visualize DNA, and images were acquired with a BioRad Transluminator (Universal Hood II). After confirmation of a PCR product, the amplicons were purified with the QIAquick PCR Purification Kit (Qiagen) according to the manufacturer's protocol, and subsequently prepared for sequencing via Oxford Nanopore Technologies' MinION Portable Sequencer.

Preparation of libraries for the MinION portable sequencer. Purified PCR amplicons were bar-coded using LongAmp Taq 2 × Master Mix (New England BioLabs) and Oxford Nanopore Technologies (ONT) PCR Barcoding Expansion Pack 1–96 (EXP-PBC096) according to ONT's PCR barcoding (96) amplicons protocol. Barcoding primers were 24 bp in length, added to both ends. Sequencing adaptors were added using the 1D-Ligation Sequencing Kit (SQK-LSK109) instructions (ONT). At all steps, DNA was purified using AMPure XP beads (Agencourt) with a 1 × beads to sample ratio and quantified with a Qubit fluorometer. Up to 50 fmol of sequencing libraries were loaded on primed FLO-MIN106 flow cells (R9.4.1) (ONT). Runs were performed employing the ONT MinION sequencer and MinKNOW GUI at Kit SQK-LSK109, Exp-PBC096 barcoding, and High-accuracy basecalling settings for up to 24 h (ONT). The MinKNOW GUI and EPI2ME software demultiplexed and trimmed barcodes for each sample, placing fast5 and fastq reads in barcode-specific labelled folders. On average, close to 5000 sequence reads per sample were obtained; for samples yielding under 2000 reads, a new sequencing run was conducted.

Analysis of ONT sequencing data. De-multiplexed fastq reads were filtered via SeqKit³⁹ with the stipulation that all reads to be analyzed had to contain at least 250 bases from each end of the amplicon, simultaneously, collecting only complete reads for further analysis. After high accuracy base-calling, the CRISPResso2 software pipeline⁴⁰ was used to analyze genome editing of the amplicon sequences by aligning sequences wildtype T locus control samples run alongside the experimental samples, and to the reference genome sequence for the T/Brachyury gene. This pipeline quantifies insertions, mutations, and deletions to determine whether a read is modified or unmodified and then summarizes genome editing results in plots and datasets⁴¹. First, input reads were filtered according to quality score (phred33), by which potentially false-positive indels were removed. Amplicon minimum alignment score to the target amplicon was set at 60. For samples with deletions larger than 1 kb, this score was lowered to 45 (denoted in the command "-amas 45") and the Gap open option for Needleman-Wunsch alignment was changed from the default of – 20 to – 30 (denoted in the command "-needleman_wunsch_gap_open-30") to allow the pipeline to assess reads containing these large deletions.

Research involving laboratory animals. All animal experimentation protocols in this study were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) at Pennington Biomedical Research Center/Louisiana State University System. All procedures involving mice were compliant with the "Guide for Care and Use of Animals in Laboratory Research" (<https://pubmed.ncbi.nlm.nih.gov/21595115/>; <https://doi.org/10.17226/12910>), as required by the US National Institutes of Health. The currently active protocol #1044 was approved by the Pennington Biomedical Research Center IACUC for the period from December 18, 2018 through December 17, 2021. Experimentation and analysis were performed following ARRIVE (Animal Research: Reporting In Vivo Experiments) guidelines (<https://arriveguidelines.org/arrive-guidelines>), which were also applied in the writing of this manuscript.

Research involving recombinant DNA. All experiments involving the production and purification of recombinant DNA were reviewed and approved by the Institutional Biological Safety Committee at Pennington Biomedical Research Center/Louisiana State University System.

Data availability

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request, in compliance with institutional policies of Pennington Biomedical Research Center/Louisiana State University System.

Received: 27 May 2021; Accepted: 15 November 2021

Published online: 09 December 2021

References

1. Wallingford, J. B., Niswander, L. A., Shaw, G. M. & Finnell, R. H. The continuing challenge of understanding, preventing, and treating neural tube defects. *Science* **339**, 1222002 (2013).
2. Doudna, J. A. & Charpentier, E. Genome editing: The new frontier of genome engineering with crispr-cas9. *Science* **346**, 1258096 (2014).
3. Guimier, A. *et al.* Mmp21 is mutated in human heterotaxy and is required for normal left-right asymmetry in vertebrates. *Nat. Genet.* **47**, 1260–1263 (2015).
4. Sunagawa, G. A. *et al.* Mammalian reverse genetics without crossing reveals nr3a as a short-sleeper gene. *Cell Rep.* **14**, 662–677 (2016).
5. Papaioannou, V. E. & Behringer, R. R. Early embryonic lethality in genetically engineered mice: Diagnosis and phenotypic analysis. *Vet. Pathol.* **49**, 64–70 (2011).
6. Dickinson, M. E. *et al.* High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508–514 (2016).
7. Singh, P., Schimenti, J. C. & Bolcun-Filas, E. A mouse geneticist's practical guide to crispr applications. *Genetics* **199**, 1–15 (2015).
8. Wang, H. *et al.* One-step generation of mice carrying mutations in multiple genes by crispr/cas-mediated genome engineering. *Cell* **153**, 910–918 (2013).
9. Gluecksohn-Waelsch, S. & Erickson, R. P. The t-locus of the mouse: Implications for mechanisms of development. *Curr. Top. Dev. Biol.* **5**, 281–316 (1970).
10. Rashbass, P., Cooke, L. A., Herrmann, B. G. & Beddington, R. S. A cell autonomous function of brachyury in t/t embryonic stem cell chimaeras. *Nature* **353**, 348–351 (1991).
11. Beisaw, A. *et al.* Brachyury directs histone acetylation to target loci during mesoderm development. *EMBO Rep.* **19**, 118–134 (2018).
12. Vidigal, J. A. *et al.* An inducible rna interference system for the functional dissection of mouse embryogenesis. *Nucleic Acids Res.* **38**, e122 (2010).
13. Pennimpede, T. *et al.* In vivo knockdown of brachyury results in skeletal defects and urorectal malformations resembling caudal regression syndrome. *Dev. Biol.* **372**, 55–67 (2012).
14. Lu, H., Giordano, F. & Ning, Z. Oxford nanopore minion sequencing and genome assembly. *Genom. Proteom. Bioinform.* **14**, 265–279 (2016).
15. Dobrovolskaia-Zavadskaja, N. On spontaneous tail mortification in newborn mice and on the existence of a 'non-viable' hereditary trait (factor). *Soc. Biol.* **1**, 116–118 (1927).
16. Herrmann, B. G. *et al.* Cloning of the t gene required in mesoderm formation in the mouse. *Nature* **343**, 617–622 (1990).
17. Wilson, V. & Beddington, R. Expression of t protein in the primitive streak is necessary and sufficient for posterior mesoderm movement and somite differentiation. *Dev. Biol.* **192**, 45–58 (1997).
18. Wilson, V., Manson, L., Skarnes, W. C. & Beddington, R. S. The t gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. *Development* **121**, 877–886 (1995).
19. Saunders, T. L. The history of transgenesis. *Methods Mol. Biol.* **2066**, 1–26 (2020).
20. Shin, H. Y. *et al.* Crispr/cas9 targeting events cause complex deletions and insertions at 17 sites in the mouse genome. *Nat. Commun.* **8**, 15464 (2017).
21. Ran, F. A. *et al.* Genome engineering using the crispr-cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
22. Montague, T. G. *et al.* Chopchop: A crispr/cas9 and talen web tool for genome editing. *Nucleic Acids Res.* **42**, W401–W407 (2014).
23. Labun, K. *et al.* Chopchop v3: Expanding the crispr web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171–W174 (2019).
24. Luo, Y. *Crispr Gene Editing: Methods and Protocols* (Humana Press, 2019).
25. Jain, M., Olsen, H. E., Paten, B. & Akeson, M. The oxford nanopore minion: Delivery of nanopore sequencing to the genomics community. *Genome Biol.* **17**, 239 (2016).
26. Payne, A. *et al.* Readfish enables targeted nanopore sequencing of gigabase-sized genomes. *Nat. Biotechnol.* **39**, 442–450 (2021).
27. Loose, M. Finding the needle: Targeted nanopore sequencing and crispr-cas9. *CRISPR J.* **1**, 265–267 (2018).
28. Miyamoto, S. *et al.* Nanopore sequencing reveals a structural alteration of mirror-image duplicated genes in a genome-editing mouse line. *Congenit. Anom.* **60**, 120–125 (2020).
29. Schrode, N. *et al.* Anatomy of a blastocyst: Cell behaviors driving cell fate choice and morphogenesis in the early mouse embryo. *Genesis* **51**, 219–233 (2013).
30. Kaufman, M. H. & Bard, J. B. L. *The Anatomical Basis of Mouse Development* (Academic Press, 1999).
31. Meier, J. A., Zhang, F. & Sanjana, N. E. Guides: Sgrna design for loss-of-function screens. *Nat. Methods* **14**, 831–832 (2017).
32. Hsu, P. D. *et al.* DNA targeting specificity of rna-guided cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
33. Concordet, J. P. & Haeussler, M. Crispor: Intuitive guide selection for crispr/cas9 genome editing experiments and screens. *Nucleic Acids Res.* **46**, W242–W245 (2018).
34. Doench, J. G. *et al.* Optimized sgrna design to maximize activity and minimize off-target effects of crispr-cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
35. Sanson, K. R. *et al.* Optimized libraries for crispr-cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416 (2018).
36. Taketo, M. *et al.* Fvb/n: An inbred mouse strain preferable for transgenic analyses. *Proc. Natl. Acad. Sci. USA* **88**, 2065–2069 (1991).
37. Nagy, A., Gertsenstein, M., Vintersten, K. & Behringer, R. *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory Press, 2003).
38. Wang, Z. & Storm, D. R. Extraction of DNA from mouse tails. *Biotechniques* **41**, 410–412 (2006).
39. Shen, W., Le, S., Li, Y. & Hu, F. Seqkit: A cross-platform and ultrafast toolkit for fasta/q file manipulation. *PLoS ONE* **11**, e0163962 (2016).
40. Clement, K. *et al.* Crispresso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).
41. Pinello, L. *et al.* Analyzing crispr genome-editing experiments with crispresso. *Nat. Biotechnol.* **34**, 695–697 (2016).

Acknowledgements

We thank Ms. Bethani Bell for help with animal husbandry, and Mss. Sydney Jones, Jennifer Bailey and Jennifer Martin for technical assistance on dissection days. This project was funded by the Eunice Kennedy Shriver

National Institute of Child Health and Human Development Grant 5R01-HD078283 to C.K. and the Peggy M. Pennington Cole Endowed Chair in Maternal Biology & the Risk of Obesity. The generation of CRISPs was performed in the Pennington Biomedical Research Center Transgenic Core Facility, which received support from the PBRC/Louisiana Nutrition and Obesity Research Center, funded by the National Institute for Diabetes and Digestive and Kidney Diseases through NORC Grant 1P30-DK072476, and from the PBRC Center for Biomedical Research Excellence, funded by the National Institute of General Medical Sciences COBRE Grant 1P30-GM118430.

Author contributions

K.T.B.F. conducted experiments, analyzed the results, and drafted the manuscript; J.M.S. provided training on the transgenic methods and helpful suggestions on sequencing strategy; C.K. conceived the study, acquired the funding, composed the figures, and edited the final version of the manuscript. All authors reviewed and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021