

OPEN

# MultiFRAGing: Rapid and Simultaneous Genotyping of Multiple Alleles in a Single Reaction

Cassidy Petree & Gaurav K. Varshney \*

Powerful and simple, RNA-guided CRISPR/Cas9 technology is a versatile genome editing tool that has revolutionized targeted mutagenesis. CRISPR-based genome editing has enabled large-scale functional genetic studies through the generation of gene knockouts in a variety of model organisms including zebrafish, and can be used to target multiple genes simultaneously. One of the challenges associated with the large scale application of this technique to zebrafish is the lack of a cost-effective method by which to identify mutants. To address this, we optimized the high-throughput, high-resolution fluorescent PCR-based fragment analysis method to develop MultiFRAGing - a robust and cost-effective method to genotype multiple targets in a single reaction. Our approach can identify indels in up to four targets from a single reaction, which represents a four-fold increase in genotyping throughput. This method can be used by any laboratory with access to capillary electrophoresis-based sequencing equipment.

Following completion of the human genome sequencing project, identification of candidate disease genes has been the focus of much genetic research. With the development of less expensive sequencing technologies, such genes are being discovered at a rapid rate, but functional validation remains slow. Most of the knowledge of gene function has been generated using gene knockout technology in model organisms<sup>1</sup>.

The zebrafish (*Danio rerio*) has become a popular model organism for many reasons including high fecundity, optically transparent embryos and larvae, external development and the ease with which various types of genetic manipulation can be performed<sup>2</sup>. A large number of zebrafish mutants have been generated using a variety of random mutagenesis approaches involving either the chemical mutagen N-ethyl-N-nitrosourea (ENU) or insertional mutagens (Retroviruses, Transposons)<sup>2</sup>. Recent progress in the transformative, targeted, and simple RNA-guided CRISPR/Cas9-based genome editing method has expedited genetic manipulation in many systems, including zebrafish<sup>3-5</sup>. In addition to CRISPR/Cas9, Transcription Activator-like Effector Nuclease (TALENs), and Zinc Finger Nucleases (ZFNs) are other targeted mutagenesis methods being used to generate knockouts for the purposes of developing disease models and understanding disease pathology in zebrafish<sup>2,6</sup>. Recently, the structure-guided endonuclease (SGN) - a DNA-guided genome editing tool that uses flap endonuclease I (FEN-1) fused to the FokI endonuclease - was also used to generate large deletions in the zebrafish genome<sup>7,8</sup>. While CRISPR/Cas9 is a RNA-guided endonuclease and TALENs and ZFNs are engineered proteins, the SGN nuclease functions in a similar manner: it induces a double stranded break (DSB) at a DNA target site which is then repaired by an error-prone, non-homologous end-joining method that often leaves insertions and/or deletions during the repair.

Because of its simplicity, low-cost and ability to target multiple sites simultaneously, CRISPR/Cas9 is becoming most popular genome editing tool, and many workflows are available for generating large numbers of gene knockouts<sup>9-13</sup>. While the generation of knockouts is relatively straight forward, the identification of mutations in a high-throughput, affordable manner remains a challenge. More than 70% of the indels generated by CRISPR/Cas9 (or TALENs/ZFNs) are less than 20 bp, making genotyping challenging<sup>9</sup>. The most sensitive method used to identify indels involves amplification of the target region followed by cloning and sequencing - a labor intensive and time-consuming process not amenable to high-throughput technology. A number of other methods including high-resolution melt analysis (HRMA), melting curve analysis, restriction fragment length polymorphism (RFLP) analysis, PAGE-based screening, and the Surveyor assay are suitable to identify mutations on a small scale<sup>14-18</sup>. We and others previously adopted a high-throughput, high-resolution fragment analysis-based method

Genes & Human Disease Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, 73104, USA. \*email: [Gaurav-varshney@omrf.org](mailto:Gaurav-varshney@omrf.org)

to genotype CRISPR-induced alleles<sup>9,19,20</sup>, and later demonstrated that it can also be used to determine guide RNA activities *in vivo*<sup>21</sup>. Fragment analysis involves generation of double-stranded fluorescently labeled fragments using PCR, and subsequent separation by size using capillary electrophoresis; software determines the relative size of each fluorescently labeled fragment by comparison with a size standard to generate the genotype of each amplicon<sup>22</sup>.

Over the past 10 years, zebrafish has emerged as a preferred model organism to study various human diseases, and the use of CRISPR/Cas9 is fueling this growth<sup>6</sup>. As the throughput of CRISPR/Cas9-mediated mutagenesis increases, researchers are able to target multiple genes simultaneously, and the development of a multiplex genotyping method to reduce both cost and labor is needed. Here, we detail the development of MultiFRAGing, a multiplexing fragment analysis pipeline that could genotype up to four targets in a single reaction, increasing the throughput up to 4-fold while significantly reducing cost.

## Materials and Methods

**Ethics statement, and zebrafish care.** The zebrafish experiments were carried out in compliance with the National Institutes of Health guidelines for animal handling and research under Oklahoma Medical Research Foundation (OMRF) Institutional Animal Care and Use Committee (IACUC) approved protocol 17-01. Zebrafish were housed in an AAALAC (Association for Assessment and Accreditation of Laboratory Animal Care) accredited facility. All zebrafish handling, embryo care, and microinjections were performed according to procedures described in *The Zebrafish Book*<sup>23</sup>. Wildtype (WT) zebrafish strain TAB-5 was used for all experiments. Zebrafish embryos were maintained in E3 embryo medium with 0.00002% methylene blue and raised at 28 °C.

**Generation of mutant lines using CRISPR/Cas9 in zebrafish.** The guide RNAs (sgRNAs), Cas9 mRNA synthesis, and microinjections were carried out as described earlier<sup>9,10</sup>. Injected eggs were raised to the adulthood to generate founder fish. Six to eight founder fish were outcrossed with the wild type fish to generate heterozygous progenies (F<sub>1</sub>). Progenies from founders carrying mutations were raised to adulthood to generate the F<sub>1</sub> generation, adults of which were genotyped using fragment analysis as previously described<sup>9,10,22</sup>.

To generate mutant lines with multiple alleles for multiplex genotyping, homozygous *dfnb31a*−/−; *dfnb31b*−/− double mutants were crossed with *grhl2a*−/−; *grhl2b*−/− double mutants. The progenies were raised to adulthood, and fish carrying five different alleles (*dfnb31a* T1, *dfnb31b* T1, *grhl2a* T1, *grhl2a* T2, and *grhl2b* T1) in four genes were used to establish this method.

**Primer design for fragment analysis.** Primers were designed to amplify amplicons 180–300 bp in length, usually keeping the target site in the middle of the amplified fragments. In principle, fluorescently labeled amplicons can be generated by labelling one primer with a fluorophore, but this approach is not cost effective as designing fluorescent labeled primers for each target is expensive. To make it more affordable, we previously modified the method to include an adapter sequence (M13Fwd or T3 or SP6) to tail the 5′-end of gene-specific forward primers<sup>9</sup>. A third fluorescently labeled primer (M13Fwd-FAM, T3-TAMRA or SP6-HEX) was designed for use with gene-specific primers. This strategy avoids the cost associated with the fluorescent labeling of individual primers. The list of primers is listed in Table 1.

In order to avoid stutter peaks in genotyping, we added a 7-nucleotide tag (PIGTAIL) at the end of the gene-specific reverse primer. Taq DNA polymerases often catalyzes the non-templated addition of nucleotides to the 3′ end of PCR amplicons in a primer-specific activity that can introduce errors in accurate genotyping. Brownstein *et al.* demonstrated that the inclusion of a GTGTCTT sequence (PIGTAIL) in the reverse primer could suppress this activity<sup>24</sup>. The PIGTAIL was later adopted by Sood *et al.* in genotyping indels in zebrafish<sup>22</sup>. The sizes of fluorescent amplicons are calculated as follows: size of amplicon amplified by gene-specific primers + (size of tailed sequence attached to forward primer + 7 bp PIGTAIL sequence).

Multiplexing can be done in following ways:

- (1) Amplicons separated by two/three different sizes amplified in a single reaction.
- (2) Amplicons separated by two/three different colors amplified in a single reaction.
- (3) Amplicons separated by three/four different sizes amplified separately, and pooled together for capillary electrophoresis.
- (4) Amplicons separated by three different colors and two sizes amplified separately, and pooled together for capillary electrophoresis.

**Genomic DNA extraction.** Genomic DNA was extracted using a previously described method<sup>10</sup>. Fins were clipped from heterozygous adults, tail pieces were dissolved in 30 μl 50 mM NaOH, heated to 95 °C for 20 minutes to completely dissolve all tissue, vortexed and centrifuged. 30 μl 100 mM Tris-HCl (pH 8.0) was added as a neutralization solution. DNA was diluted 10X with nuclease-free water. Extracted DNA can be stored at −20 °C for up to six months. For 2–5 dpf embryos, we recommend using 10 μl of NaOH and 10 μl Tris-HCl and DNA should be diluted to 50X.

**Multiplex PCR setup.** First, all three primers were mixed together as follows:

- 5 μL 100 μM Fluorescent Primer
- 3 μL 100 μM Gene-specific forward Primer
- 5 μL 100 μM Gene-specific reverse Primer
- 487 μL water

Primer Name	Sequence
M13Fwd	TGTA AACGACGGCCAGT
SP6	ATTTAGGTGACACTATAG
T3	ATTAACCTCACTAAAGG
M13Fwd-FAM	/56-FAM/TGTA AACGACGGCCAGT
SP6-HEX	/5HEX/ATTTAGGTGACACTATAG
T3-TAMRA	/56-TAMN/ATTAACCTCACTAAAGG
M- <i>grhl2a</i> T2-F-FAM	tgtaaacgacggccagtTCCGAACACCACCATCACTA
M- <i>grhl2a</i> T2-F-HEX	atttaggtgacctatagTCCGAACACCACCATCACTA
M- <i>grhl2a</i> T2-F-TAMRA	attaaccctactaaaggTCCGAACACCACCATCACTA
M- <i>grhl2a</i> T2-R	gtgtcttATTGAAGCAAGCCGTTCTGT
M- <i>grhl2b</i> T1-F-FAM	tgtaaacgacggccagtGAAACAGCAGCCAAATGGAG
M- <i>grhl2b</i> T1-F-HEX	atttaggtgacctatagGAAACAGCAGCCAAATGGAG
M- <i>grhl2b</i> T1-F-TAMRA	attaaccctactaaaggGAAACAGCAGCCAAATGGAG
M- <i>grhl2b</i> T1-R	gtgtcttGTCTGTAGTGTCCCCCTGA
M- <i>dfnb31a</i> T1-F-FAM	tgtaaacgacggccagtGTGCTGATGCTGTCAGGAGA
M- <i>dfnb31a</i> T1-F-HEX	atttaggtgacctatagGTGCTGATGCTGTCAGGAGA
M- <i>dfnb31a</i> T1-F-TAMRA	attaaccctactaaaggGTGCTGATGCTGTCAGGAGA
M- <i>dfnb31a</i> T1-R	gtgtcttGCTCGGATCAGCTTCTGT
M- <i>dfnb31b</i> T1-F-FAM	tgtaaacgacggccagtCACCTTGACTGCCTCTCCAT
M- <i>dfnb31b</i> T1-F-HEX	atttaggtgacctatagCACCTTGACTGCCTCTCCAT
M- <i>dfnb31b</i> T1-F-TAMRA	attaaccctactaaaggCACCTTGACTGCCTCTCCAT
M- <i>dfnb31b</i> T1-R	gtgtcttGGCTTCTGTTTCAGCACCT

**Table 1.** Primer sequences used in this study.

Any polymerase can be used for PCR, we tested multiple polymerases with variable results, reactions and conditions for two different polymerases are as follows:

(a) PCR reactions using **Platinum Taq** Polymerase were set-up in 20  $\mu$ l final volume as follows:

- 2  $\mu$ L 10x Buffer
- 0.6  $\mu$ L MgCl<sub>2</sub>
- 0.4  $\mu$ l dNTP Mix (10 m M, New England Biolabs, Cat # N0447L)
- 1.2  $\mu$ L Primer mix (for each)
- 0.16  $\mu$ l Platinum Taq Polymerase (ThermoFisher Scientific, Cat # 100021275)
- 2  $\mu$ L DNA Template
- To 20  $\mu$ L with Water

PCR was performed using the following conditions:

5 min denaturation at 94 °C followed by 34 cycles of: 94 °C for 30 sec, 57 °C for 30 sec, and 72 °C for 30 sec; and 5 min final extension at 72 °C.

(b) PCR reactions using Qiagen **HotStarTaq** Polymerase were set-up in 20  $\mu$ l final volume as follows:

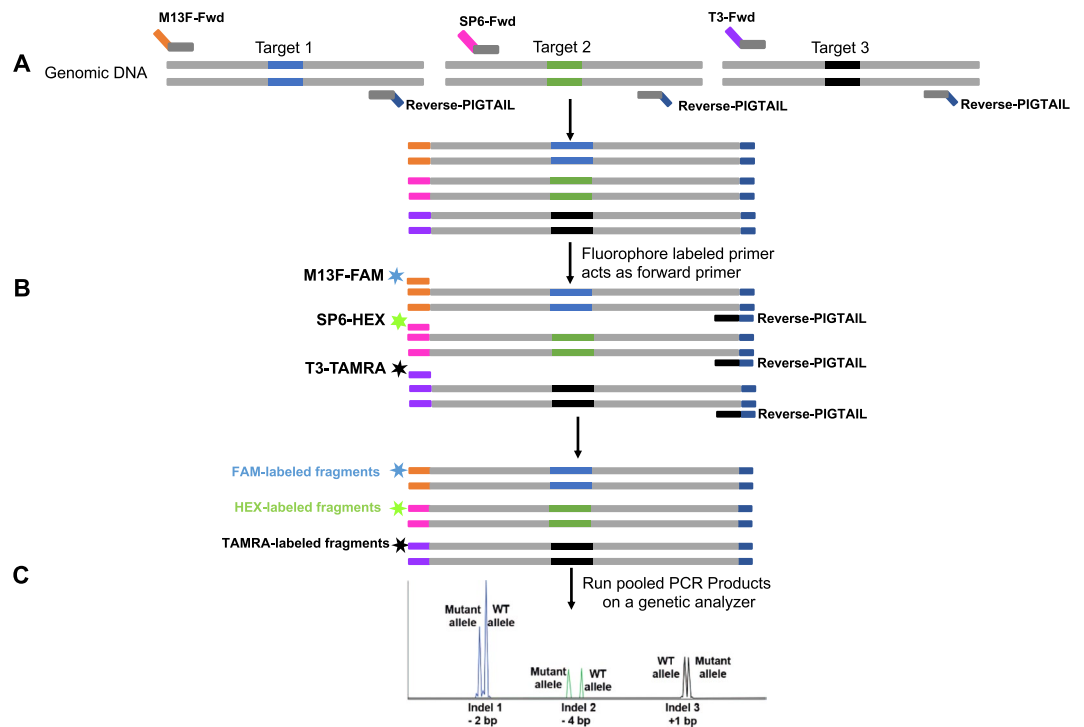
- 2  $\mu$ L 10x Buffer
- 0.4  $\mu$ l dNTP Mix (10 m M, New England Biolabs, Cat # N0447L)
- 2  $\mu$ L Primer mix (for each)
- 0.1  $\mu$ l Qiagen HotStar Taq Plus (Qiagen LLC, Cat # 203609)
- 2  $\mu$ L DNA Template
- To 20  $\mu$ L with Water

PCR was performed using the following conditions:

15 min denaturation at 95 °C, 30 seconds denaturation at 94 °C followed by 34 cycles of: 94 °C for 30 sec, 55 °C for 30 sec, and 72 °C for 30 sec; and 5 min final extension at 72 °C.

Successful PCR amplification was confirmed by electrophoresis on 2% agarose gels.

**Fragment analysis by capillary electrophoresis.** Here, fluorescent PCR fragments were separated by capillary electrophoresis on a Genetic Analyzer (ABI 3500-XL), although fragment analysis can be performed on any ABI Genetic Analyzer platform. A detailed protocol for fragment analysis has been described earlier<sup>10</sup>. Briefly,



**Figure 1.** Overview of MultiFRAGing method. (A) Primer design strategy for multiple targets. Gene-specific primers are designed to generate 180–300 bp fragments. The gene-specific forward primers are attached with an adapter sequence (M13Fwd, SP6, and T3), and the reverse primer contains a short PIGTAIL sequence to suppress the stutter peaks. Tailed gene-specific primers add M13Fwd or SP6 or T3 adapter sequences in the first few PCR cycles. (B) A third primer with same adapter sequence attached to a fluorescent dye (FAM, HEX and TAMRA) is used to generate fluorescently labeled fragments. After few PCR cycles, fluorophore-labeled primers act as forward primer and bind to the respective adapter sequences. In the subsequent PCR cycles, most fragments will incorporate fluorophore thus generating double-stranded fluorescent fragments. Multiple fragments are generated in a single PCR reaction. These fragments can either be tagged with same dye and generate products of different size or tagged with different dyes. (C) Pooled PCR products are mixed with a size standard to run on a genetic analyzer. Fragments sizes are plotted, and indel size can be measured based on the expected size of the wild type fragment. Wild-type samples will have one size (allele), while heterozygous samples will show two sizes (alleles).

double-stranded, fluorescently-labelled PCR fragments were mixed with a size standard. We used GenScan 400 HD ROX size standard (ThermoFisher Scientific, Cat # 402985) as an internal size marker: ROX was diluted 1:100 in Hi-Di formamide (ThermoFisher Scientific, Cat #4311320) and 9  $\mu$ l of diluted mix was added to 1  $\mu$ l of PCR product. Samples were mixed and then denatured at 95 °C for 5 minutes before separation on the Genetic Analyzer. For multiplexing where amplicons were generated using individual PCR reactions, 2  $\mu$ l PCR products from each reaction were pooled, mixed and then 2  $\mu$ l of the mixture was added to Hi-Di Formamide: ROX mix. The following conditions were used for fragment analysis: Application type: Fragment; Capillary Length: 50 cm; Polymer: POP7; Dye Set: DS30; Filter Set: D; and injection time 15 seconds. Allele sizes were determined using GeneMapper software (ThermoFisher Scientific, Cat # A38888).

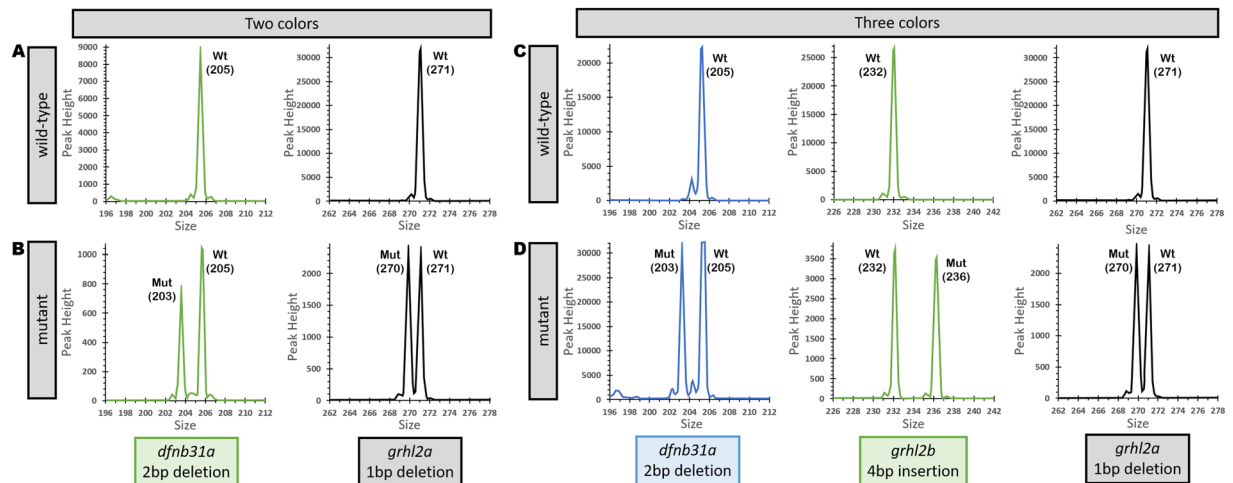
**Cloning and sequencing of PCR products.** To verify the size of indels, PCR products from the fragment analysis reaction were directly sub-cloned into a pCR4-TOPO vector (ThermoFisher Scientific, Cat # K457502). Plasmid was extracted using a Zymo Plasmid Miniprep kit (Zymo Research, Cat # D4054), and 100 ng DNA from individual clones sequenced using a BigDye Terminator Cycle Sequencing kit (ThermoFisher Scientific, Cat #4337456). The resulting DNA fragments were purified and sequenced using ABI Genetic Analyzer 3500Xl (Applied Biosystems), and aligned with wild type (reference) sequences using SnapGene software (GSL Biotech LLC).

## Results

Our aim was to establish a reliable multiplex method to identify indels from multiple targets in a single PCR reaction that would save time, cost, and increase genotyping efficiency (Fig. 1). The fragment analysis workflow presented here involves labeling fragments with fluorescent dyes to allow multiple colors of fluorescent dyes to be detected in a single sample. ABI genetic analyzers can accommodate at least five different fluorescent dyes (one of which is reserved for a size standard). We used the DS-30 dye set with 6-FAM (blue), Hex (green), NED (yellow), and Rox (Red) (ThermoFisher Scientific, Cat # 4345827). Rox was used as the labeled size standard, leaving the three other colors available. (If more colors are needed, the DS-33 dye set - which contains 6-FAM (Blue), VIC (Green), NED (Yellow), PET (Red) and Liz (Orange) - can be used.) As described in the methods section,

Gene	Wild type allele size	Mutant Allele size	Indel size
<i>dfnb31a</i>	205.0	203.25	2 bp del
<i>dfnb31b</i>	300.46	287.39	13 bp del
<i>grhl2a</i>	268.58	267.77	1 bp del
<i>grhl2a</i>	268.74	258.61	10 bp del
<i>grhl2b</i>	232.14	236.19	4 bp ins

**Table 2.** Summary of wild type and mutant alleles, and predicted indel sizes.



**Figure 2.** Fragment Analysis PCR plots from wild-type samples and a pool of samples amplified together. (A) Wild-type controls for size comparison with two targets separated by different dyes. (B) Two targets are separated based on different dyes, showing a 2 bp or 1 bp deletion when compared to wild-type sizes. (C) Wild-type controls for size comparison with three targets separated by different dyes. (D) Three targets are separated by dye color, and show a 2 bp deletion, 4 bp insertion, and 1 bp deletion compared to wild-type sizes.

gene-specific forward primers were tailed with an adapter sequence (in this case M13Fwd, T3, and SP6 sequences, though any sequence that lacks similarity to the genome could be used). The same sequences were used for dye-labeling. We replaced the NED dye with TAMRA because it uses same color (yellow), is readily available and inexpensive to synthesize.

To establish this method, we used a mutant line carrying mutations at five distinct sites in four different genes (*dfnb31a T1*, *dfnb31b T1*, *grhl2a T1*, *grhl2a T2*, and *grhl2b T1*). Summary of mutant alleles is listed in Table 2, and detailed data is available in Supplementary Table 1. We tested following combination of fragments based on two parameters: fragments labelled with two or three colors, fragments of two or three sizes, and fragments with different sizes and colors together. First, we genotyped these alleles separately and confirmed the five independent alleles listed in Table 2. These alleles were then used to develop the multiplex method retrospectively. The fluorescent PCR products were combined with a size standard, and run on capillary electrophoresis to identify indels.

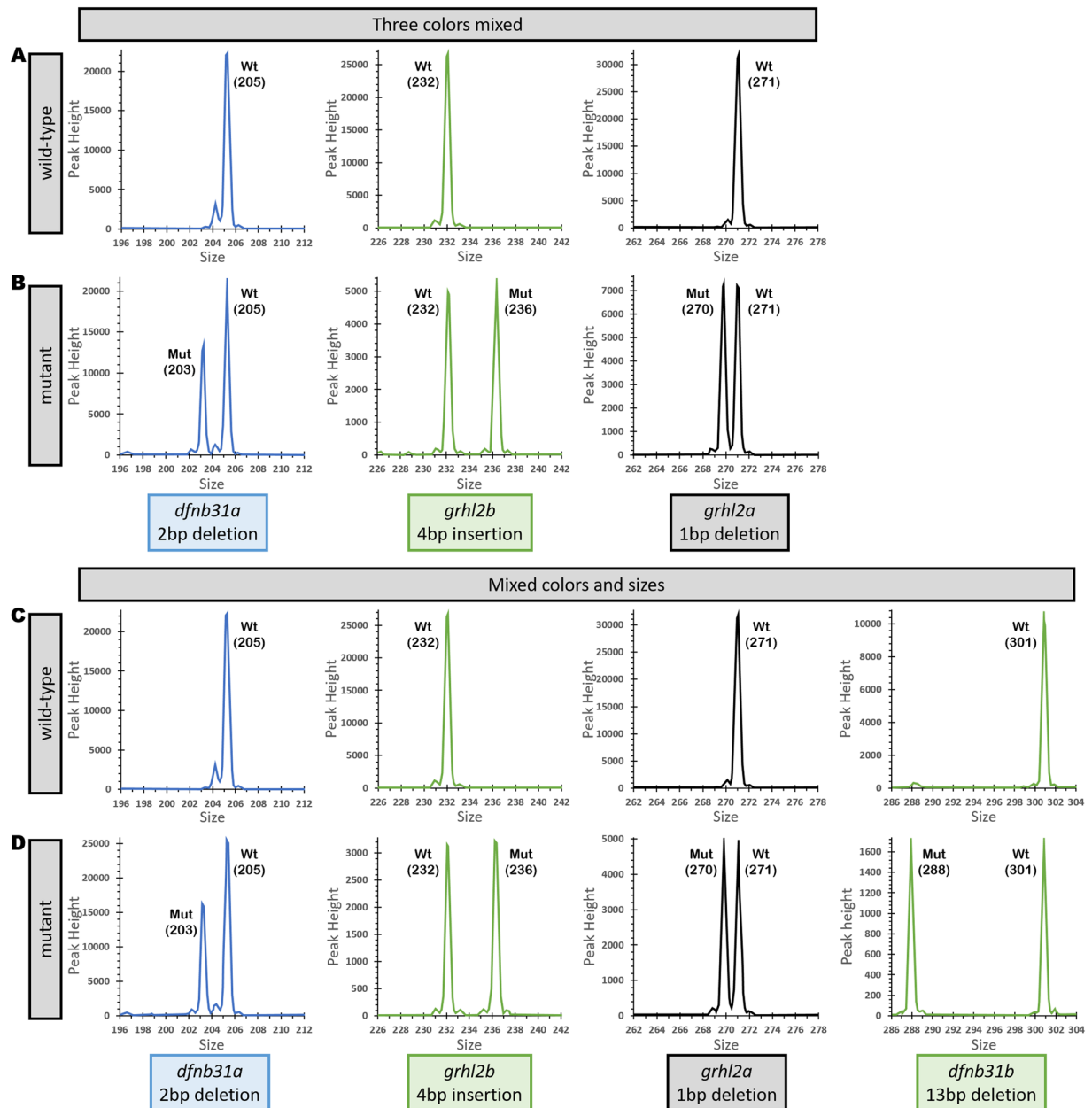
#### Amplification of multiple fragments separated by color, and multiplex genotyping in a single reaction.

We tested two different strategies to multiplex: amplicons separated by size and/or colors. It has been shown that the majority of indels induced by CRISPR/Cas9 are less than 20 bp in size, which makes it possible to design specific primers to generate fragments of different sizes (within a 180–300 bp range), thereby allowing us to vary both fragment length and dye color (Table 1). To test this approach, we first amplified two fragments using two gene-specific primer sets, together with unique third primers linked to different fluorophores (TAMRA and HEX dyes) simultaneously. The resulting amplicons were subjected to fragment analysis in a single reaction. Genotyping showed that control samples (wild-type) generated single peaks (corresponding to the size of the amplicon) for each allele (Fig. 2A), and heterozygous samples generated two peaks (mutant and wild-type). In both cases the second peak was smaller than the wild-type control indicating deletions (2 bp and 1 bp) (Table 2). After testing two fragments by simultaneous amplification and genotyping, we tested three different fragments separated by three unique fluorophores- FAM (Blue), HEX (Green) and TAMRA (Yellow), and genotyped them simultaneously in a single reaction. As expected, three different alleles [corresponding to a 2 bp deletion (*dfnb31a*), 4 bp insertion (*grhl2b*), and 1 bp deletion (*grhl2a*)] were identified successfully (Fig. 2D); the size of the mutant allele was determined by size comparison with control peaks (Fig. 2C).

#### Amplification of multiple fragments separated by size, and multiplex genotyping in a single reaction.

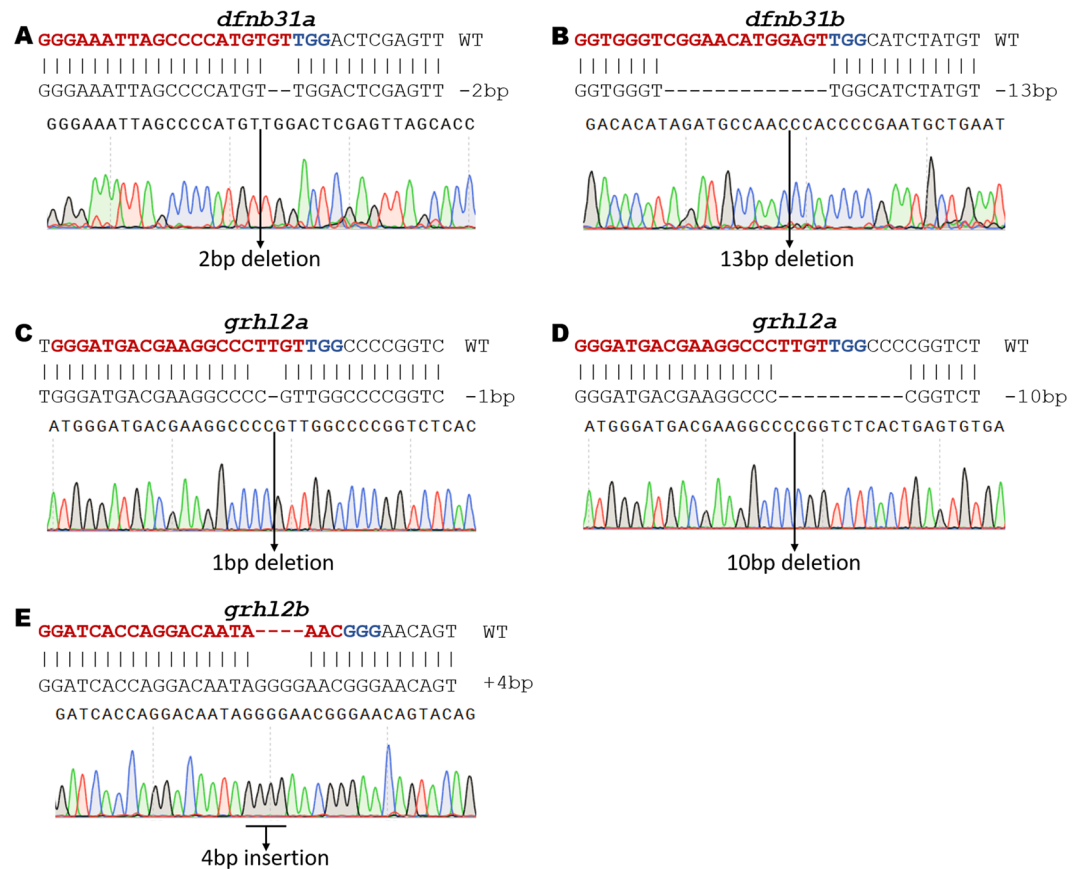
We hypothesized that given most indels are less than 20 bp, multiple targets can be separated and identified





**Figure 3.** Fragment Analysis plots from wild-type controls and mutants from fragments separated by color, sizes and amplified individually, then pooled together for genetic analyzer. (A) Wild-type controls for size comparison with three targets separated by different dyes. (B) Three targets separated by different dyes showing indels of 2 bp deletion, 4 bp insertion, and 1 bp deletion when compared to wild-type sizes. (C) Wild-type controls for size comparison with targets separated by size and dye color. (D) Four targets separated by a combination of different sizes and dye colors showing indels of 2 bp deletion, 4 bp insertion, 1 bp deletion, and 13 bp deletion when compared to wild-type sizes.

by size. The choice of size is dependent on the fragment analysis size standard; we use 400 HD ROX, which is designed for sizing DNA fragments in the 50–400 nucleotide range. We normally design primers to generate fragments between 180–300 bp. For other size standards (e.g. 500 HD ROX, or 500 LIZ), amplicons of up to 500 bp can be generated. We tested this strategy by amplifying two, three, and four targets of different sizes simultaneously in a single reaction. Two and three targets amplified successfully, however, four target amplification was not successful. We tested different DNA polymerases (Platinum Taq, AmpliTaq Gold, and QIAGEN HotStarTaq), including some specialized for multiplex PCR (NEB Multiplex PCR 5X Master Mix, and Phusion Multiplex PCR Master Mix). Surprisingly, standard polymerases were more effective than multiplex PCR master mixes: Qiagen HotStarTaq, and Platinum Taq performed slightly better than AmpliTaq Gold. We successfully amplified three targets of 205 bp, 232 bp, and 269 bp simultaneously (Supplementary Fig. 1). When the collection of amplified fragments was genotyped, each of the two and three mutant alleles were successfully identified (Supplementary



**Figure 4.** Validation of indels by Sanger Sequencing. Each indel that was identified by fragment analysis was sequenced using the Sanger method to establish the correlation between fragment analysis and Sanger sequencing Data. All indels from fragment analysis showed similar indel size in Sanger sequencing. (A) Alignment and chromatogram confirming a 2 bp deletion in *dfnb31a*. (B) Alignment and chromatogram confirming a 13 bp deletion in *dfnb31b*. (C) Alignment and chromatogram confirming a 1 bp deletion in *grhl2a*. (D) Alignment and chromatogram confirming a 10 bp deletion in *grhl2a*. (E) Alignment and chromatogram confirming a 4 bp insertion in *grhl2b*.

Fig. 2B,D respectively). The control sizes are shown in Supplementary Fig. 2A,C. Thus, this strategy of amplifying three targets simultaneously can increase genotyping throughput by three-fold.

**Genotyping of individually amplified fragments in a pool.** In many cases only one sgRNA is used to target a single gene, and these targets are amplified individually. We tested whether individually amplified targets with different colors and/or sizes could be pooled together and still allow indels to be identified in a single reaction. We distinguished these targets in multiple ways: by size, color, or a mix of colors and sizes. To test this approach, 3 or 4 different targets were amplified individually in separate reactions. We combined 2  $\mu$ l of three or four different PCR products, and 2  $\mu$ l of pooled product was run on ABI genetic analyzer for genotyping in a single well. First, we mixed fragments amplified from three targets separated by three colors. In a second pool, a combination of sizes (203, 232, 271, and 301 bp) and colors (fragments of 232 and 301 bp have the HEX fluorophore, while 203 uses FAM, and 271 uses TAMRA fluorophore) was tested. In both cases, all alleles were identified successfully and accurately as shown in Fig. 3C,D. Furthermore, pooling three and four fragments of different sizes and genotyping them in a single reaction yielded the same results: all mixed alleles were identified successfully (Supplementary Fig. 3B,D). Wild-type sizes are shown in Supplementary Fig. 3A,B. These results demonstrate that it is possible to pool multiple PCR products to increase the genotyping throughput up to four-fold, and save the cost of consumables time, and labor.

**Verification of mutant alleles by sanger sequencing.** While this fragment analysis-based genotyping provides the size of the indels, it does not provide the exact sequences of mutations. In order to do this, amplicons must be sequenced the first time mutants are identified (which we normally do to validate the fragment analysis data). Here, we confirmed all alleles tested to demonstrate the feasibility of multiplex method. We genotyped five different alleles using fluorescent PCR: 2 bp deletion in *dfnb31a*; 13 bp deletion *dfnb31b*; 1 bp insertion and 10 bp deletion in *grhl2a*; 4 bp deletion in *grhl2b* (Fig. 4A–E). Amplicons from heterozygous samples were sub-cloned in a TOPO vector, and DNA from individual clones was sequenced. Sanger sequencing confirmed all five alleles

as identified from fluorescent PCR. Heterozygous samples can also be sequenced directly, and sequencing data can be analyzed by Poly Peak Parser (<http://yosttools.genetics.utah.edu/PolyPeakParser/>), a tool to parse double peaks from Sanger sequencing chromatograms into wild-type and alternative alleles (Supplementary Fig. 4A,B)<sup>25</sup>.

## Discussion

Emerging genome editing technologies such as CRISPR/Cas9, and TALENs have revolutionized targeted mutagenesis in zebrafish. While they allow gene knockouts to be generated in a high-throughput manner, the identification of mutated alleles at a similar rate and in a cost-effective manner has been lacking. Several strategies are being used such as DNA sequencing, fluorescent PCR based fragment analysis, T7 endonuclease assay, PAGE-based genotyping, high-resolution melting analysis (HRMA), and melting curve analysis<sup>18</sup>. Each of these techniques has their own strengths and weaknesses: the most sensitive method of identifying allele is Sanger sequencing but it is expensive, time-consuming, and low throughput; HRM, melting curve analysis, and fluorescent PCR methods are high-throughput; the HRM and melting curve analysis methods are low cost, but they lack sensitivity due to the high number of polymorphisms in zebrafish. The fluorescent PCR based fragment analysis method is high-throughput and highly sensitive, but also high cost. The fragment analysis method was first used to identify indels generated by ZFNs in zebrafish<sup>20</sup>, and later modified to include a third fluorescent primer (to reduce cost) and a PIGTAIL (to suppress stutter peaks<sup>22</sup>). This method was used to identify indels generated by CRISPR/Cas9 in a high-throughput manner<sup>9,10</sup>, and modified to develop CRISPR-STAT (CRISPR-Somatic Tissue Activity Test), an assay to determine the sgRNA activity *in vivo*<sup>21</sup>. In order to further reduce cost and increase throughput, we adapted this method to identify multiple alleles in a single reaction, showing that fragment analysis can be used to genotype multiple alleles simultaneously. The result is a more robust and cost-effective genotyping method. The cost of this method is approximately \$2–3 per reaction in our lab. All alleles identified by fragment analysis were confirmed by sequencing, and the sequencing data was co-related with the fragment analysis data, which confirms the sensitivity of this method.

Multiplexing can be done using two different approaches: amplification of up to three targets in a single PCR reaction, followed by genotyping of the products in a single reaction; another approach is to amplify targets individually and pool the products for genotyping in a single reaction, which allows us to genotype up to four targets. Another way to genotype four targets simultaneously is to amplify two targets simultaneously in a single reaction, and then pool fragments from two reactions. While we have demonstrated the use of multiplexing in zebrafish, this method can be easily used in other systems.

Received: 31 October 2019; Accepted: 4 February 2020;

Published online: 21 February 2020

## References

- Housden, B. E. *et al.* Loss-of-function genetic tools for animal models: cross-species and cross-platform differences. *Nat. Rev. Genet.* **18**, 24–40, <https://doi.org/10.1038/nrg.2016.118> (2017).
- Varshney, G. K., Sood, R. & Burgess, S. M. Understanding and Editing the Zebrafish Genome. *Adv. Genet.* **92**, 1–52, <https://doi.org/10.1016/bs.adgen.2015.09.002> (2015).
- Jao, L. E., Wente, S. R. & Chen, W. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc. Natl Acad. Sci. USA* **110**, 13904–13909, <https://doi.org/10.1073/pnas.1308335110> (2013).
- Hwang, W. Y. *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **31**, 227–229, <https://doi.org/10.1038/nbt.2501> (2013).
- Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278, <https://doi.org/10.1016/j.cell.2014.05.010> (2014).
- Varga, M. *et al.* Zebrafish Models of Rare Hereditary Pediatric Diseases. *Diseases* **6**, <https://doi.org/10.3390/diseases6020043> (2018).
- Xu, S. *et al.* An alternative novel tool for DNA editing without target sequence limitation: the structure-guided nuclease. *Genome Biol.* **17**, 186, <https://doi.org/10.1186/s13059-016-1038-5> (2016).
- Varshney, G. K. & Burgess, S. M. DNA-guided genome editing using structure-guided endonucleases. *Genome Biol.* **17**, 187, <https://doi.org/10.1186/s13059-016-1055-4> (2016).
- Varshney, G. K. *et al.* High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. *Genome Res.* **25**, 1030–1042, <https://doi.org/10.1101/gr.186379.114> (2015).
- Varshney, G. K. *et al.* A high-throughput functional genomics workflow based on CRISPR/Cas9-mediated targeted mutagenesis in zebrafish. *Nat. Protoc.* **11**, 2357–2375, <https://doi.org/10.1038/nprot.2016.141> (2016).
- Gagnon, J. A. *et al.* Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* **9**, e98186, <https://doi.org/10.1371/journal.pone.0098186> (2014).
- Burger, A. *et al.* Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Dev.* **143**, 2025–2037, <https://doi.org/10.1242/dev.134809> (2016).
- Moreno-Mateos, M. A. *et al.* CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting *in vivo*. *Nat. Methods* **12**, 982–988, <https://doi.org/10.1038/nmeth.3543> (2015).
- VanLeuven, A. J., Park, S., Menke, D. B. & Lauderdale, J. D. A PAGE screening approach for identifying CRISPR-Cas9-induced mutations in zebrafish. *Biotechniques* **64**, 275–278, <https://doi.org/10.2144/btn-2018-0012> (2018).
- Dahlem, T. J. *et al.* Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet.* **8**, e1002861, <https://doi.org/10.1371/journal.pgen.1002861> (2012).
- Yu, C., Zhang, Y., Yao, S. & Wei, Y. A PCR based protocol for detecting indel mutations induced by TALENs and CRISPR/Cas9 in zebrafish. *PLoS One* **9**, e98282, <https://doi.org/10.1371/journal.pone.0098282> (2014).
- Thomas, H. R., Percival, S. M., Yoder, B. K. & Parant, J. M. High-throughput genome editing and phenotyping facilitated by high resolution melting curve analysis. *PLoS One* **9**, e114632, <https://doi.org/10.1371/journal.pone.0114632> (2014).
- D'Agostino, Y. *et al.* A Rapid and Cheap Methodology for CRISPR/Cas9 Zebrafish Mutant Screening. *Mol. Biotechnol.* **58**, 73–78, <https://doi.org/10.1007/s12033-015-9905-y> (2016).
- Yang, Z. *et al.* Fast and sensitive detection of indels induced by precise gene targeting. *Nucleic Acids Res.* **43**, e59, <https://doi.org/10.1093/nar/gkv126> (2015).
- Foley, J. E. *et al.* Targeted mutagenesis in zebrafish using customized zinc-finger nucleases. *Nat. Protoc.* **4**, 1855–1867, <https://doi.org/10.1038/nprot.2009.209> (2009).



21. Carrington, B., Varshney, G. K., Burgess, S. M. & Sood, R. CRISPR-STAT: an easy and reliable PCR-based method to evaluate target-specific sgRNA activity. *Nucleic Acids Res.* **43**, e157, <https://doi.org/10.1093/nar/gkv802> (2015).
22. Sood, R. *et al.* Efficient methods for targeted mutagenesis in zebrafish using zinc-finger nucleases: data from targeting of nine genes using CompoZr or CoDA ZFNs. *PLoS one* **8**, e57239, <https://doi.org/10.1371/journal.pone.0057239> (2013).
23. Westerfield M. The zebrafish book: A guide to the laboratory use of the zebrafish (*Danio Rerio*). 4th ed., *Univ. of Oregon Press, Eugene*. 1–354 (2000).
24. Brownstein, M. J., Carpten, J. D. & Smith, J. R. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques* **20**(1004–1006), 1008–1010, <https://doi.org/10.2144/96206st01> (1996).
25. Hill, J. T. *et al.* Poly peak parser: Method and software for identification of unknown indels using sanger sequencing of polymerase chain reaction products. *Dev. Dyn.* **243**, 1632–1636, <https://doi.org/10.1002/dvdy.24183> (2014).

## Acknowledgements

We thank Rachel Smith for help with figures, Duane Goins and the staff of the comparative medicine department for animal husbandry work. This research is supported by a grant from NIH/COBRE GM103636 (Project 3), and the Presbyterian Health Foundation (PHF) Grant to G.K.V.

## Author contributions

G.K.V. designed the experiments, C.P. performed the experiments, C.P. and G.K.V. analyzed data, and G.K.V. wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-59986-1>.

**Correspondence** and requests for materials should be addressed to G.K.V.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020