Molecular Therapy
Nucleic Acids
Original Article

# Allele-specific CRISPR-Cas9 editing inactivates a single nucleotide variant associated with collagen VI muscular dystrophy

Véronique Bolduc,[1,4] Katherine Sizov,[1,4] Astrid Brull,[1] Eric Esposito,[1] Grace S. Chen,[1] Prech Uapinyoying,[1,2] Apurva Sarathy,[1] Kory R. Johnson,[3] and Carsten G. Bönnemann[1]

[1]Neurogenetics and Neuromuscular Disorders of Childhood Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA; [2]Center for Genetic Medicine Research, Children's National Research and Innovation Campus, Children's National Hospital, Washington, DC 20012, USA; [3]Bioinformatics Core, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA

**The application of allele-specific gene editing tools can expand the therapeutic options for dominant genetic conditions, either via gene correction or via allelic gene inactivation in situations where haploinsufficiency is tolerated. Here, we used allele-targeted CRISPR-Cas9 guide RNAs (gRNAs) to introduce inactivating frameshifting indels at an SNV in the *COL6A1* gene (c.868G>A; G290R), a variant that acts as dominant negative and that is associated with a severe form of congenital muscular dystrophy. We expressed SpCas9 along with allele-targeted gRNAs, without providing a repair template, in primary fibroblasts derived from four patients and one control subject. Amplicon deep sequencing for two gRNAs tested showed that single-nucleotide deletions accounted for the majority of indels introduced. While activity of the two gRNAs was greater at the G290R allele, both gRNAs were also active at the wild-type allele. To enhance allele selectivity, we introduced deliberate additional mismatches to one gRNA. One of these optimized gRNAs showed minimal activity at the WT allele, while generating productive edits and improving collagen VI matrix in cultured patient fibroblasts. This study strengthens the potential of gene editing to treat dominant-negative disorders, but also underscores the challenges in achieving allele selectivity with gRNAs.**

## INTRODUCTION

Collagen VI is a secreted, multimeric protein abundant in the connective tissues of skeletal muscle.[1,2] The main collagen type VI genes, *COL6A1*, *COL6A2*, and *COL6A3*, are expressed by interstitial muscle fibroblasts and not by myofibers.[3,4] They encode three proteins, called, respectively, collagen α1, α2, and α3(VI) chains, each structured with a collagenous central triple helical domain composed of Gly-X-Y repeats.[5] Owing to the small glycine in these repeats, the three alpha chains can tightly intertwine to form a monomer, composed of one of each of the chains, that further self-assemble into dimers and tetramers.[6,7] Tetramers are secreted and polymerize end-to-end into a microfibrillar network that constitutes the collagen VI matrix[6] and that contacts with myofibers.[3]

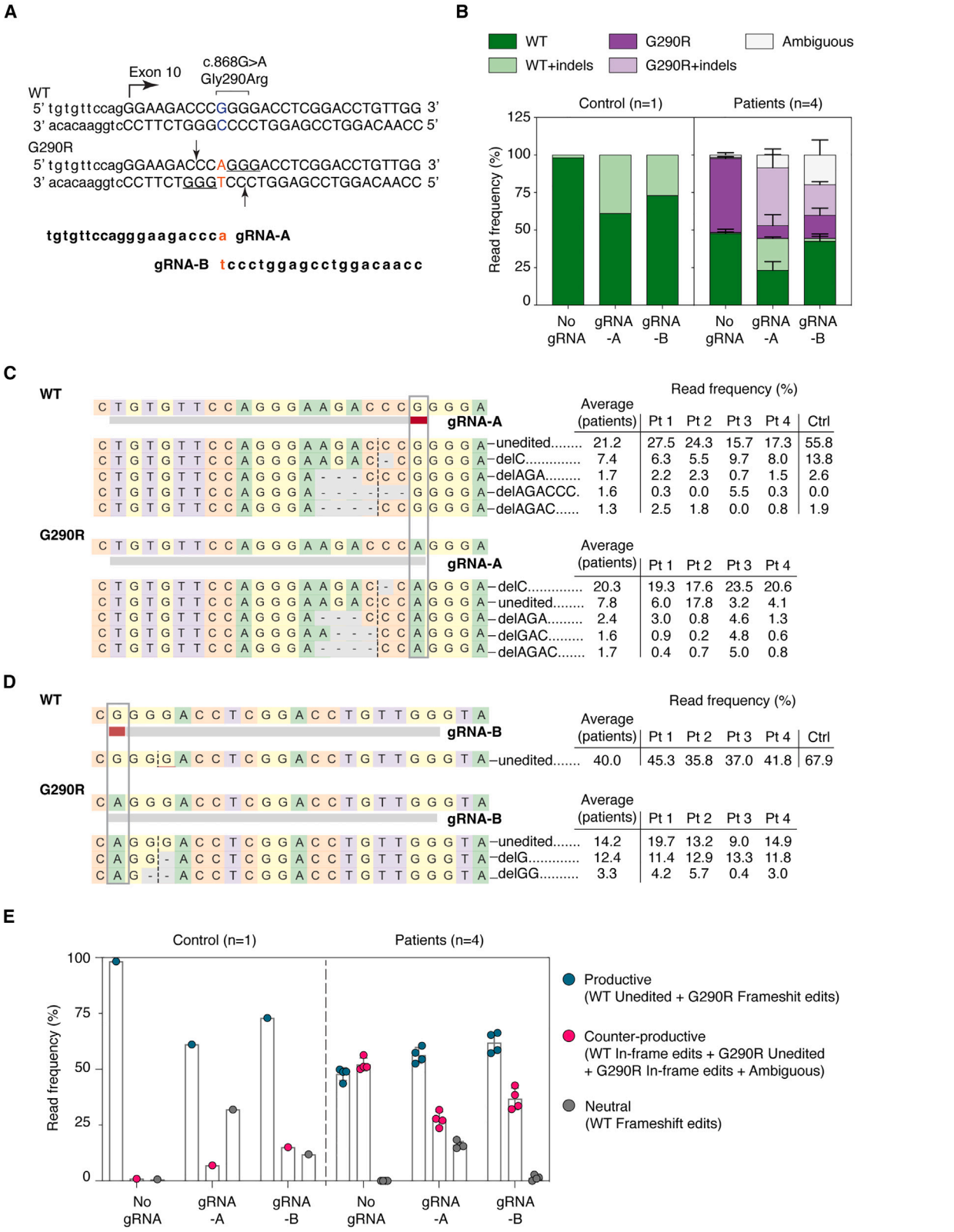
Integrity of the triple helical domains of collagen α1, α2, and α3(VI) chains is thus essential for proper collagen VI production. Dominantly acting variants account for over 65% of all *COL6* pathogenic variants that cause collagen VI-related muscular dystrophies (COL6-RD),[8] and they cluster at the amino termini of the triple helical domains.[2] Initial coiling of the alpha chains, to form monomers, proceeds from the carboxy toward the amino termini of the triple helical domains. As a result, pathogenic variants at the amino termini allow the alpha chains harboring them to be incorporated, rather than excluded, into monomers, and to further propagate to dimer and tetramer stages.[9] For this reason, these "assembly-competent" variants exert strong dominant-negative effects. One type of variant that is frequently reported is glycine substitutions of the first position of the Gly-X-Y repeat.[2,10] Typically, such glycine substitutions are not tolerated and cause structural changes to tetramers that appear "kinked" on electron microscopy and that exhibit a reduced ability to polymerize in the matrix.[9,11] By mechanisms not yet fully elucidated, dysfunctional collagen VI production leads to skeletal muscle weakness, progressive joint contractures, and respiratory failure, which can be fatal.[1,12]

There are currently no treatments for COL6-RD. Given that *COL6A1*, *COL6A2*, and *COL6A3* genes are haplosufficient (i.e., losing or inactivating one copy of either of these genes does not result in a clinical phenotype),[12,13] we and others have postulated that allele-specific silencing would be a promising therapeutic strategy for dominant-negative variants causing COL6-RD. This can be achieved with antisense technologies, such as small interfering RNAs or gapmer antisense oligonucleotides,[14] adapted to be allele specific. COL6-RD primary cultures treated with these molecules provided evidence

Received 20 November 2023; accepted 12 July 2024;
https://doi.org/10.1016/j.omtn.2024.102269.

[4]These authors contributed equally

**Correspondence:** Carsten G. Bönnemann, Neurogenetics and Neuromuscular Disorders of Childhood Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA.
**E-mail:** carsten.bonnemann@nih.gov




*(legend on next page)*

that specifically knocking down the pathogenic transcripts attenuates the dominant-negative effect and improves collagen VI matrix deposition[15–18]; however, complete and sustained mutant allele knockdown with stringent allele selectivity has yet to be demonstrated in preclinical animal models of COL6-RD.

Gene editing via the CRISPR-Cas9 system is also adaptable to achieve allele-specific gene inactivation.[19] Engineered CRISPR-Cas9 is a two-component system that includes expression of a nuclease (e.g., *Streptococcus pyogenes* Cas9 [SpCas9]) concomitantly with a single guide RNA (gRNA) to induce a double-stranded break at desired genomic locations.[20] Using a gRNA sequence complementary to a single allele enables allele-specific targeting, even for single nucleotide changes.[19] The creation of double-stranded breaks then activates the cellular DNA repair pathways, which, in the absence of template, are conducted primarily by the canonical non-homologous end joining (NHEJ) pathway, or by the microhomology-mediated end joining (MMEJ) pathway (also called alternative end joining) (reviewed in references[21,22]). Both NHEJ and MMEJ are error-prone and readily introduce indels during the repair process following a Cas9-induced double-stranded break. The DNA modifications, if frameshifting, can permanently inactivate expression of the gene in which they were introduced.

Allele-specific CRISPR-Cas9-induced gene inactivation has been tested in various cellular and animal models of dominant diseases, with variable degrees of Cas9 on-target activity and allele discrimination.[23–29] Here, we applied this precision medicine to a recurrent glycine substitution in *COL6A1* (G290R).[9,10] We show that the indel profiles favor gene inactivation and that the gRNA design can be optimized for enhanced allele selectivity with the addition of base mismatches.

## RESULTS

### CRISPR-Cas9 gRNAs selected for the *COL6A1* SNV (c.868G>A; G290R) preferentially introduce frameshifting edits at the variant allele

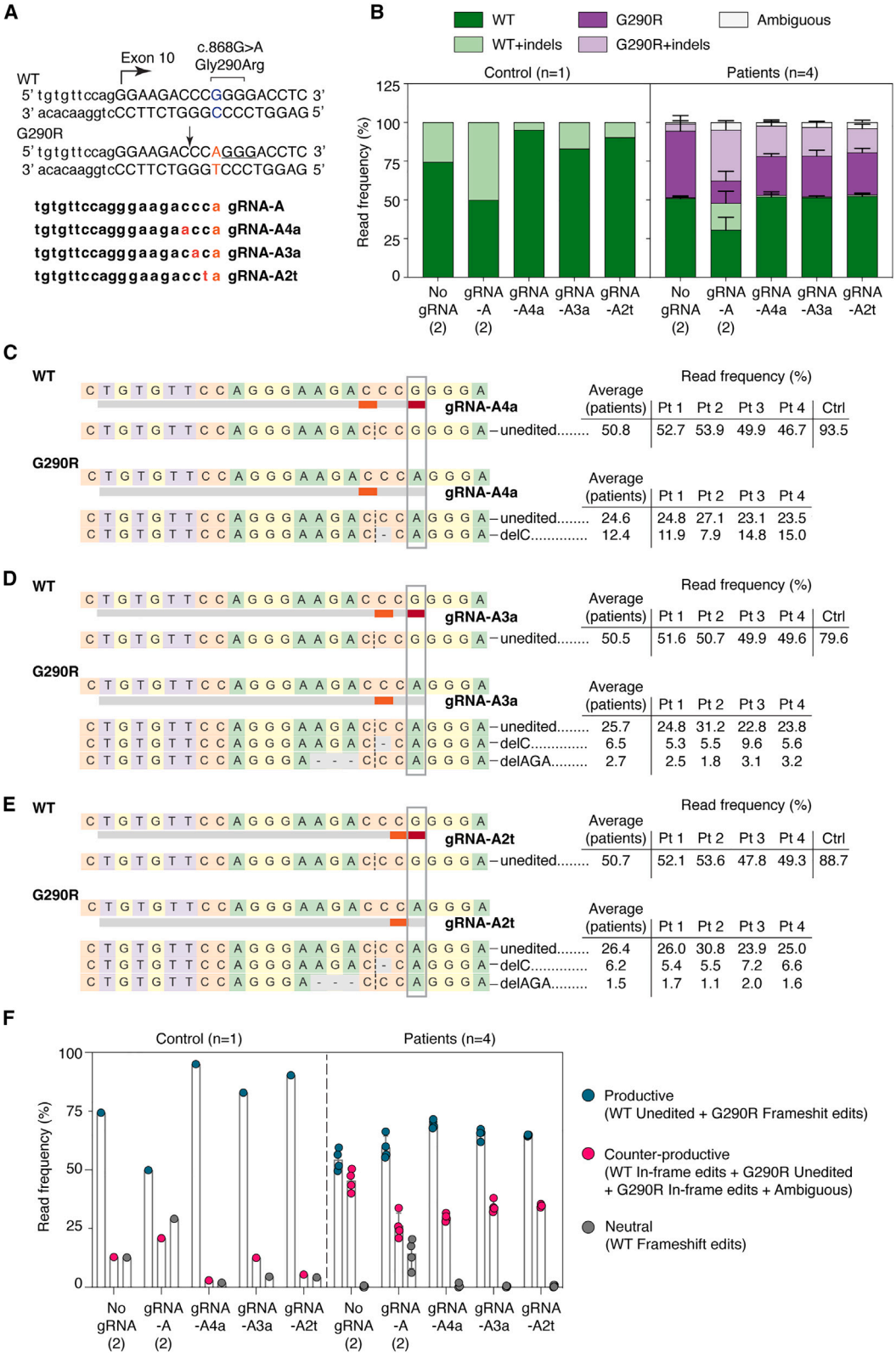
At the *COL6A1* c.868G>A locus, we selected two gRNAs (gRNA-A and gRNA-B; Figures 1A, S1A, and S1B) in which the c.868 nucleotide was located at position 1 of the protospacer (i.e., at a 1-nt distance relative to the NGG protospacer adjacent motif [PAM]). The gRNAs had no significant predicted off-target sites, including toward the other *COL6* genes. We prepared plasmid constructs, each expressing an SpCas9-GFP fusion protein concomitantly with either gRNA-A or gRNA-B, or without gRNA (No gRNA; Figure S1C). To test our gene editing strategy, we nucleofected the constructs into primary dermal fibroblast cells obtained from four independent patients carrying the *COL6A1* c.868G>A (G290R) variant (Pt1–Pt4), and from one non-neuromuscular disease patient as a control (Ctrl). Dermal-derived fibroblasts express collagen VI at high levels and are therefore a good surrogate for the muscle interstitial fibroblasts.[30] GFP-enriched cell populations, collected 48 h after nucleofection, were used for all further analyses (Figure S1C).

We performed targeted deep sequencing of the locus of interest to measure the precise editing frequencies (Figure S1C). The total number of reads was tallied for each sample (average of 188,607 reads per sample; Figure S2A) and served as denominators to calculate read frequencies. Both gRNAs efficiently introduced indels at the c.868G>A (G290R) allele in patient cells (38.4% and 20.3% of total reads were G290R allele modified with indels, for gRNA-A and gRNA-B, respectively, vs. 0.8% for No gRNA; Figure 1B). While they were active at the G290R allele, both gRNAs also modified the native (wild-type [WT]) allele in patient cells (21.2% and 1.9% of total reads, for gRNA-A and gRNA-B, respectively) and in the control sample (38.9% and 27.0% of total reads, for gRNA-A and gRNA-B, respectively). With gRNA-B, deletions frequently encompassed the c.868 nucleotide site, such that a larger proportion of reads were ambiguous (19.8% of total reads; Figure 1B). It is likely that a fraction of these represent modified WT reads and that the 1.9% read frequency observed in this sample is underestimated. The 27.0% of modified reads in the control sample suggests that gRNA-B does not effectively discriminate between the two alleles. Overall, both gRNAs were more active at the G290R allele compared to the WT allele, but they did not fully discriminate the G290R from the WT alleles.

We next examined the indel edits generated by either gRNA. For gRNA-A, in patient cells, deletions were observed far more frequently than insertions (54.7% of total reads, G290R and WT alleles combined, vs. 4.0% of total reads, respectively; Figure S2B). Of these, 1-bp deletions were the most frequent (33.0% of total reads, G290R and WT alleles combined; Figure S2C). A single edit, deletion of a cytosine at the cleavage site (c.865delC), accounted for most of the modified reads (27.7% of total reads, for G290R and WT alleles combined; Figures 1C and S2D). Looking at the G290R allele only, this edit represented nearly half of the reads (43.4% of G290R reads, $n = 4$ patient cells). Preference for this modification was observed consistently in the four patient cell lines tested (Figures 1C and S2D), as well as in the control cell line (13.8% of total reads; Figures 1C and S2D). Similarly, for gRNA-B, the most frequent edits were deletions (19.9% of total

**Figure 1. Two selected gRNAs introduce frameshifting edits preferentially, but not specifically, at the *COL6A1* c.868G>A (G290R) variant allele**

(A) Two gRNAs (gRNA-A and gRNA-B) were designed to specifically anneal to the *COL6A1* c.868G>A; G290R (marked in red) variant. The PAM sequences utilized (underlined) and the Cas9 cleavage sites (arrows) are shown. (B) Histogram plots distribution of the targeted re-sequencing read frequencies analyzed with CRISPResso2, expressed as a percentage of total reads for each sample. For patient samples, bars represent the average of four individuals ± SD. (C and D) The most frequent (≥1.5%) reads were aligned to their respective allele (WT, top; G290R, bottom) for gRNA-A (C) or gRNA-B (D). The read frequency, as well as the average (for the four patient samples) is reported. (E) Reads were compiled according to the functional outcome, whether they were (1) productive (unedited WT reads or G290R reads containing a frameshifting edit), (2) damaging (that likely produces a dominant-negative product, such as unedited G290R reads, WT, or G290R reads containing in-frame edits, or ambiguous reads that encompass the splice site), or (3) neutral (WT reads that include frameshifting edits). $n = 4$ patient lines ± SD.




*(legend on next page)*

reads, for G290R and WT alleles combined), and more precisely, 1-bp deletions (13.6% of total reads, for G290R and WT alleles combined; Figures S2B and S2C). The ambiguous reads were also modified with deletions, but they were not tallied in these analyses since they could not be attributed to either allele. For gRNA-B, the single most frequent edit was c.871delG (12.4% of total reads for G290R and WT alleles combined), and this was also observed as the preferential indel in the four patient cell lines tested and in the control cell line (Figures 1D and S2D).

While the preferential edits (c.865delC for gRNA-A, c.871delG for gRNA-B) induce frameshifting and premature terminations in the coding sequence—and consequently silence the allele in which they are introduced—other motifs can retain the reading frame (e.g., deletions or insertions of multiples of 3 bp, splice site indels that would cause in-frame exon 10 skipping). Edits that preserve the reading frame likely produce dominant-negative chains. To account for the functional outcome of the indel edits introduced by Cas9/gRNA-A or -B, we sorted the resulting edits as productive (i.e., desired), neutral, or counter-productive (i.e., undesired). For example, frameshifting indels on the G290R allele were productive, since they silence the dominant-negative allele, whereas frame-preserving indels on either the WT or the G290R allele were counter-productive, since they likely create dominant-negative products. Frameshifting indels on the WT allele were classified as neutral, as they silence this allele but do not cause a dominant-negative product. In patient cells, for gRNA-A, 49.7% of reads were productive and 25.2% were counter-productive (Figure 1E). In comparison, for gRNA-B, these values were 55.9% and 29.7%, respectively (Figure 1E). Overall, indels introduced by the cleavage of Cas9 in the presence of either gRNA were more frequently productive.

### Addition of intentional mismatches to gRNA-A sequence increases allele specificity

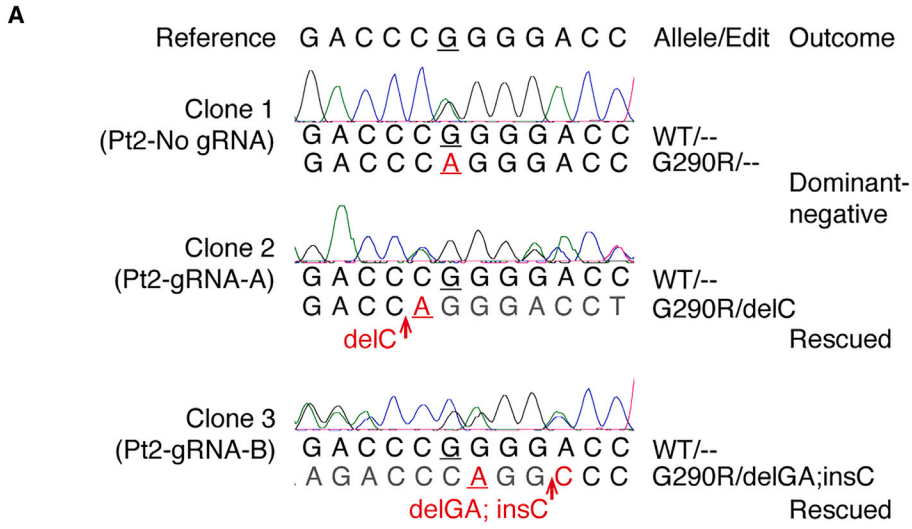
gRNA-A was efficient at inactivating the G290R allele but lacked specificity. To further increase allelic specificity, we deliberately introduced additional base mismatches in the gRNA-A sequence. Each of the re-designed gRNAs now present with two mismatches compared to the WT allele, destabilizing binding and thus editing of the WT allele; meanwhile, the single mismatch compared to the G290R allele should still allow for binding and editing of this allele. We tested three different re-designed gRNAs, in which the deliberate mismatches were introduced in position 2, 3, or 4 of the protospacer (Figure 2A). These gRNAs were nucleofected into the patient and control cell lines, and GFP-enriched cells were analyzed, as described before (Figure S1C). All three re-designed gRNAs were less active than the parental guide at the WT allele—in particular, gRNA-A4a, for which only 5.0% of reads included indels in the control sample (Figure 2B). As a trade-off, the attenuated gRNAs were less active at the G290R allele as well (19.7%, 18.5%, and 15.5% of reads were G290R + indels for gRNA-A4a, gRNA-A3a, and gRNA-A2t, respectively, compared to 32.8% of G290R + indels reads for gRNA-A; Figure 2B). Of note, the control cell line treated with the No gRNA condition is a replicate from the previous experiment (Figure 1). Unexpectedly, we found in this replicate that the sequences did not consist of 100% WT reads, possibly because of PCR contamination. Introduction of the mismatches did not change the editing patterns: deletions, in particular 1-bp deletions, were still the most common edits (Figure S3). Again, the preferential edit was c.865delC (Figures 2C–2E). The rate of productive editing was similar for each of the three re-designed gRNAs (69.5%, 65.2%, and 64.7% of reads were productive for gRNA-A4a, gRNA-A3a, and gRNA-A2t, respectively) and was higher than for the parental gRNA-A (59.5%; Figure 2F).

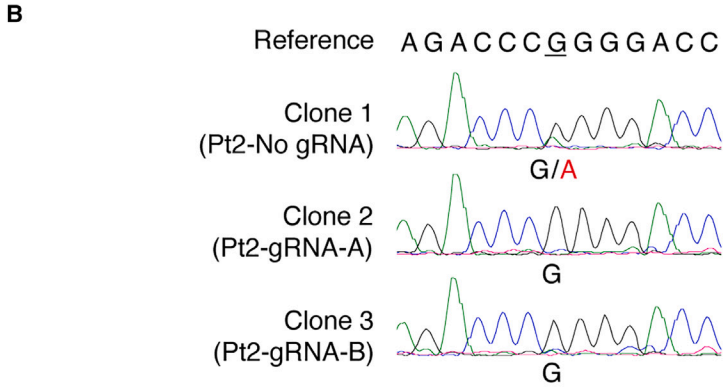
### Frameshifting edits effectively inactivate the dominant-negative G290R variant

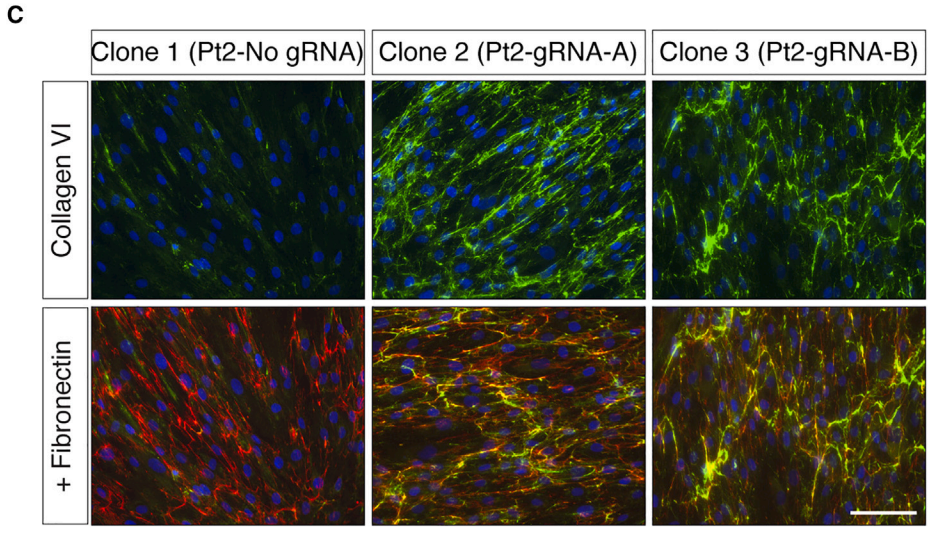
As a first step to determining the effect of suppressing the G290R allele on collagen VI matrix production, we prepared clonal cell lines following CRISPR-Cas9 editing (Figure S1C). Three clonal cell lines were successfully expanded and analyzed by Sanger sequencing. Clonal cell lines Pt2-gRNA-A and Pt2-gRNA-B acquired frameshifting edits (delC, or delGA; insC) exclusively on the c.868G>A (G290R) allele (Figure 3A). These frameshifting edits are predicted to trigger nonsense-mediated decay, thereby silencing the targeted allele. We confirmed by cDNA sequencing that only the c.868G allele was detected (Figure 3B), consistent with the absence of expression of the c.868G>A allele. As a result of the G290R allele inactivation, we expected the dominant-negative effect exerted by the mutant collagen VI alpha chain to be completely abolished in the corrected clonal cell lines. To test this, we immunostained the collagen VI deposition in the matrix using a commercial antibody that detects collagen α3(VI). Since collagen VI is produced from the three basic chains (α1– α3) that undergo a complex multistep assembly process, independently of which chain harbors the pathogenic variant, collagen VI matrix will be affected, and consequently, a collagen VI antibody specific for any of the three chains can be used to assess collagen VI secretion and deposition (Figure S4A). With this staining, we found that a collagen VI matrix was re-established in the Pt2-gRNA-A and Pt2-gRNA-B clonal cells (Figure 3C).

**Figure 2. Addition of mismatches to the gRNA sequences increases allele specificity and preserves the repair outcome**

(A) The gRNA sequence was modified to include an intentional mismatch either at position 4, 3, or 2 of the protospacer. The sequences of the three new attenuated gRNAs, as well as the PAM sequence utilized (underlined) and the Cas9 cleavage site (arrow), are shown. (B) Targeted re-sequencing (Illumina MiSeq) of the *COL6A1* c.868G>A locus in four patient and one control primary cells after gene editing with gRNA-A, gRNA-A4a, gRNA-A3a, gRNA-A2t, or No gRNA. In this experiment, gRNA-A and No gRNA were replicates from the previous experiment (Figure 1). Sequencing reads were analyzed by CRISPResso2 and plotted as in Figure 1. For the patient samples, bars represent the average of four individuals ± standard deviation. (C–E) Alignment of the most frequent reads (reaching a read frequency of at least 1.5% in either the patient or the control samples), for gRNA-A4a (C), gRNA-A3a (D), and gRNA-A2t (E). delC remains the most frequent edit at the G290R allele for all three new re-designed gRNAs. (F) Reads were compiled according to the functional outcome, as described in Figure 1. $n = 4$ patient lines ± SD.




*(legend on next page)*

### Partial inactivation of the G290R variant improves collagen VI matrix deposition

In a scenario where this CRISPR-Cas9 therapy would be administered systemically, edits could vary from cell to cell. While some cells would have complete abolition of *COL6A1* expression (if both alleles are modified with frameshifting edits) and some would produce a dominant-negative α1 chain (if an in-frame edit is introduced or if there is no editing), a proportion of cells, would, as desired, have a selective disruption of the c.868G>A allele by the introduction of inactivating frameshifts. In the latter case, the dominant-negative effect exerted by the mutant α1 chain would be completely abolished, allowing for the normal α1 chains to form tetramers that secrete, diffuse, and polymerize in the extracellular space, as opposed to the unedited scenario in which every cell carries the dominant-negative acting allele that interferes with collagen VI matrix formation.

To survey the collagen VI matrices produced by mixed populations of edited G290R cells, we treated cells with either of our two initial gRNAs (gRNA-A and gRNA-B) or with one of the re-designed gRNAs (gRNA-A4a). We cultured cells in a single dish after GFP enrichment (Figure S1C). Using immunofluorescence and confocal microscopy, we examined the overall collagen VI matrix appearance (Figure 4A), including morphology of the collagen VI microfibrils (Figure 4B). Cultures from patients carrying collagen VI glycine substitutions typically display matrices that are reduced in abundance and that are speckled compared to controls,[9,10] likely due to the increased retention of mutant tetramers and/or reduced adhesion of the secreted tetramers that get washed off during the staining procedure. We quantified the abundance of collagen VI in the cultures (Figure 4C) and the percentage of collagen VI overlapped with fibronectin, a known interactor of collagen VI in the extracellular space[31,32] (Figures S4B and 4D). Here, collagen VI abundance in both Pt1 and Pt2 cultures was increased with gRNA-A4a and gRNA-B as opposed to No gRNA, reaching statistical significance for gRNA-A4a in Pt2 only (Figure 4C). Furthermore, gRNA-A4a significantly increased collagen VI overlap with fibronectin consistently in both patients' cultures (Figure 4D). Importantly, gRNA-A4a showed no detrimental effect on the collagen VI matrix abundance and overlap with fibronectin in the Ctrl cell line (Figures 4C and 4D). Normal matrices show long and continuous microfibrillar structures of collagen VI, but mutant glycine matrices usually appear dotty and lack microfibrillar structures (Figures 4A and 4B). We used this feature to quantify the circularity ("roundness") of microfibrils detected in images and hypothesized that our editing approach would reduce their circularity. We found that in both patients' cultures, rounder microfibrils (>0.5) occupied a greater surface area and that this effect was mitigated with all three gRNAs, in particular with gRNA-A4a, which showed the greater effect (Figure 4E). This suggests that reducing the expression of dominant-negative products even in a subpopulation of cells improves overall collagen VI assembly.

Taken together, these data show that random frameshifting edits preferentially introduced at the *COL6A1* c.868G>A (G290R) allele by the action of CRISPR-Cas9 and selective gRNAs can rescue the production of collagen VI.

## DISCUSSION

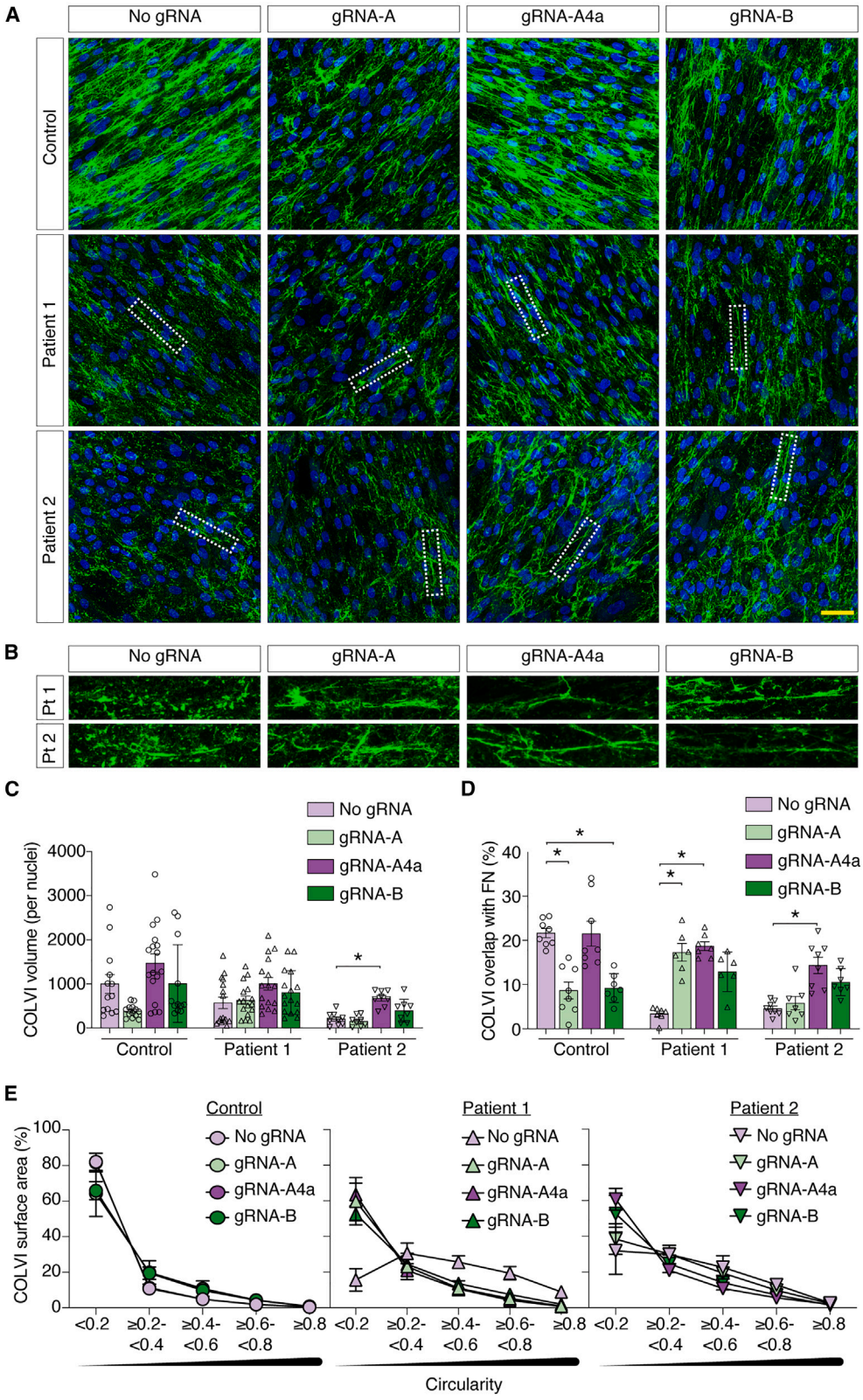
In this study, we used CRISPR-Cas9 gene editing to introduce frameshifting edits to a recurrent dominant-negative *COL6A1* variant (G290R) to achieve allele-specific gene disruption. Given that variants acting as dominant negative are a frequent cause of COL6-RD and given that harboring a single copy of either of the three *COL6* genes is not associated with a clinical phenotype,[12,13] allele-specific gene disruption is a suitable approach to treat COL6-RDs. There also is the advantage that introducing changes at the DNA level, as opposed to the RNA level, can provide permanent inactivation of the dominant allele, allowing edited alleles to be passed on to the progeny of proliferating cells, supporting long-term rescue.

Gene disruption by CRISPR-Cas9 is determined by the resulting DNA repair processes and relies on those to introduce frameshifts into the coding sequence of the targeted gene. Recent findings suggest that end joining repair profiles following Cas9 cleavage are not random, but for any given gRNA have rather predictable outcomes.[33,34] In agreement with this notion, our deep-sequencing data showed that the repair profiles were highly similar in four independent primary fibroblast cell lines from patients. Most important, the majority of edits observed, both for gRNA-A (and its derivatives) and gRNA-B, were productive (i.e., generated the desired frameshifts at the G290R allele). Two single 1-bp deletion edits (c.865delC for gRNA-A, gRNA-A4a, gRNA-A3a, gRNA-A2t, and c.871delG for gRNA-B) were predominant and contributed most to the repair profiles. The gRNAs tested in this study effectively introduced frameshifts to the coding sequence and support the use of CRISPR-Cas9-induced gene disruption to inactivate the *COL6A1* G290R variant.

While obtaining productive edits by engaging the cellular DNA repair machinery was successful, achieving allele selectivity was more challenging. In our study, we used a guide-specific approach, in which the variant was in the protospacer rather than in the PAM sequence. The targeting of PAM-creating variants has been shown in several contexts to be the ideal scenario for high allele discrimination, since

**Figure 3. Frameshifting edits at the G290R allele effectively abolish the dominant-negative effect**

Pt2 cells were clonally expanded after gene editing with either gRNA-A, gRNA-B, or No gRNA, and three clones were analyzed. (A) Genomic DNA sequencing traces obtained from the three clones (Pt 2-No gRNA, Pt 2-gRNA-A, and Pt 2-gRNA-B), aligned to the reference sequence. Clone 2 (Pt2-gRNA-A) includes a delC edit only at the G290R allele, while clone 3 (Pt2-gRNA-B) includes another frameshifting edit (delGA; insC) only at the G290R allele. (B) cDNA sequencing from the clones in (A) shows that clones with G290R allele-specific frameshifting edits exclusively express the WT allele. (C) Matrix immunostaining for collagen VI (green) and fibronectin (red) in clonal populations in which the G290R allele is inactivated. Nuclei were identified with DAPI (blue). Scale bar, 100 μm.




*(legend on next page)*

Cas9 has no detectable activity on the non-targeted allele due to the absence of a PAM sequence.[23,35,36] However, a pathogenic variant may not necessarily create a PAM site for Cas9, as this was the case for our G290R variant. While in some cases Cas nucleases with different PAM requirements can be considered instead,[26,37] certain positions of the protospacer are thought to also provide sequence specificity, such as the PAM-proximal seed sequence positions 1–8 (reviewed in reference[38]). Positioning a variant within this region should thus provide allele selectivity. In particular, it was previously reported that an imperfect base-pairing rA:dG mismatch at position 1 (like our gRNA-A:protospacer WT sequence) nearly abolished cleavage activity.[39] This was in contrast to our observations, given the suboptimal allele specificity achieved with gRNA-A and gRNA-B in which the variant was in position 1. Additional gRNA design strategies are available to enhance allelic discrimination, such as the utilization of truncated gRNAs[24,40] or the introduction of an additional, deliberate mismatch in the gRNA sequence,[38,41] which is the strategy we selected here. In our experimental design, introduction of a mismatch at a 3-nt distance of the variant (gRNA-A4A) destabilized sufficiently hybridization to the WT allele to almost abrogate cutting activity completely in a control cell line. This simple design strategy adds to the design options for precision CRISPR-Cas therapeutics and could be combined with the use of engineered Cas9 with greater fidelity,[42–44] or with the use of alternate Cas nucleases.[45,46]

A recent study applied this allele-specific gene disruption strategy to a different recurrent *COL6A1* glycine substitution (G293R) and reported high allele specificity and productive introduction of frame-shifting edits.[47] By incorporating the intentional mismatches in the gRNA design, we achieved allele specificity to levels comparable to the López-Márquez study for our target as well. One factor that could have contributed to the higher specificity observed by López-Márquez et al. using an unaltered gRNA is the position of the variants within the protospacer, which were located at either position 2 or 8.[47] Factors other than the primary sequence, however, may also account for the different specificity levels we and López-Márquez et al. observed, such as the cellular concentration of Cas9/gRNA and the time of exposure to Cas9 nuclease.[39] We used plasmid-based expression of Cas9 and gRNAs for 48 h, which "mimics" a viral delivery; López-Márquez et al. instead delivered CRISPR components as ribonucleoprotein complexes that best reflect a nanoparticle-based delivery.[47] In these current study designs, however, it is not possible to determine the contribution of each of these factors to the levels of allele specificity observed. Similarly to the López-Márquez study, we observed productive NHEJ repair outcomes and improvement of collagen VI matrix production, further supporting this CRISPR-Cas9-induced gene disruption strategy. The two studies together demonstrate that CRISPR-Cas9 approaches can be tailored to different recurrent dominant disease-causing glycine variants in the *COL6A1* gene, and thus likely also to other such disease-causing variants in the three *COL6* genes.

Dominant variants on either of the three *COL6* genes, such as glycine substitutions located at the N termini of the triple helical domains, frequently have profound dominant-negative effects on collagen VI assembly when they are "assembly competent" (i.e., when they can be carried forward into the final secreted product, the multimers of a1(VI), a2(VI) and a3(VI) protein chains that constitute the collagen VI extracellular matrix). Of note, patients who carry such variants in somatic mosaicism show considerably milder phenotypes.[48,49] Hence, even a subpopulation of cells capable of depositing a normal collagen VI matrix, as an *in vivo* CRISPR-Cas9 therapy would likely produce, has the potential to ameliorate the disease. Additionally, as noted before, every corrected cell is expected to generate an expanding progeny of corrected cells. Here, we show that clonally expanded correctly edited cells show a normalized matrix deposition, but also realize that the potential effect *in vivo* is best modeled in our experiment, where we kept primary cells as a mixed culture after editing had been applied. Testing this experimental scenario, we found that the overall extracellular microfibrillar structure of collagen VI in the cultures was also improved. While there are no established assays to verify functional recovery of this matrix in culture, the qualitative and quantitative improvement we report is encouraging and consistent with the observation of the milder phenotype observed in the individuals somatically mosaic for dominant-negative *COL6* variants. Importantly, the approach we are testing here edits and inactivates the pathogenic allele (avoiding the WT), such that the collagen α1(VI) produced originates from the WT copy of the gene, with the assumption that it assembles normally in the matrix. Additional experimentation is needed to determine what threshold of corrected cells is necessary to attain a restored and functional matrix, but these data provide an excellent rationale for allele-specific gene disruption of collagen VI genes.

Collagen VI in muscle is mainly expressed by muscle interstitial fibroblasts.[3,4] As a cellular model, we used here primary fibroblastic cells derived from unaffected or patient skin. Skin-derived fibroblasts are

**Figure 4. Collagen VI matrix production is improved after gene editing**

(A–E) Confocal microscopy of matrix immunostaining for bulk cultures after gene editing with either gRNA-A, gRNA-A4a, gRNA-B, or No gRNA. (A) Representative merge stacks images showing the abundance of collagen VI (green) secretion in each culture (nuclei, blue). Scale bar, 50 μm. (B) Enlarged image insets, identified by the dotted rectangles in (A), showing that the structure of the collagen VI microfibrils is dotty in patients' cultures, but that it is improved after gene editing. (C) Quantification of the collagen VI abundance, normalized to the number of nuclei. Each data point represents one quantified field. Four fields were surveyed for each experiment, and experiments were replicated three times. Bars represent average ± SEM. *$p < 0.05$, Kruskal-Wallis analysis of variance, followed by Dunn's multiple comparisons test. (D) Percentage of the collagen VI surface area that overlaps with fibronectin (FN). Each data point represents one quantified field. Four fields were surveyed for two replicate experiments. Bars represent average ± SEM. *$p < 0.05$, Kruskal-Wallis analysis of variance, followed by Dunn's multiple comparisons test. (E) Collagen VI surface area plotted against the degree of circularity (roundness) of detected fibrils in acquired fluorescence images. Data points represent the average of eight fields from two replicate experiments ± SEM.

capable of producing and depositing collagen VI efficiently and are well established to reflect the abnormal collagen VI deposition in a wide variety of COL6-RD patients.[30] While dermal primary fibroblasts certainly differ from muscle fibroblasts *in vivo*, they can serve as a model to test the on-target activity of the system. However, given that the genomic and epigenetic contexts influence Cas9 activity, working with a relevant cell type is essential, in particular for the survey of genomic off-targets, which we did not evaluate systematically in the present study. To translate our approach *in vivo*, an appropriate systemic delivery system is required, whether it is based on lipid nanoparticles or viral vectors such as adeno-associated viruses (AAVs). The target cell type for a COL6-RD therapy is the muscle interstitial fibroblasts, as opposed to the myofiber and/or the satellite cell that is for other types of muscular dystrophies. To date, no such delivery system has been established for the muscle fibroadipogenic progenitors (FAPs), and none of the common AAV serotypes has been identified to efficiently transduce FAPs. The development of AAV capsids or nanoparticles directed at the muscle fibroblasts is an area of interest for the COL6-RD field.

CRISPR-Cas9 is a powerful gene editing tool that can mediate targeted gene disruption. Two recent phase 1 clinical studies made use of this experimental design, in which Cas9 and gRNAs were administered *ex vivo* to transfusion-dependent β-thalassemia and sickle cell disease patients,[50] and *in vivo* to transthyretin amyloidosis patients.[51] In these trials, expression levels of the targeted genes were strongly reduced; however, the gRNAs were designed to target both alleles. Modifying the system to achieve allele-specific gene disruption would expand the application to additional conditions, including those caused by dominant and dominant-negative variants[19] that occur in haplosufficient genes. Our study focused on a variant-specific approach; however, in combination with the López-Márquez et al. study, it provided a strong proof of principle for the application of this approach to other dominant *COL6* variants. Our study also supports mismatch gRNA design as an additional tool to the precision engineering of allele-specific gRNAs. To make this approach into a more universal tool for the COL6-RD patient community, an interesting strategy would be to develop a repertoire of allele-specific gRNAs for each allele of common SNPs, and to target those polymorphisms as proxy for pathogenic variants located in *cis* to the polymorphisms.[52] Combined with an appropriate delivery system, the approach could be applied to a larger portion of the patient population.

## MATERIALS AND METHODS

### gRNA design

Possible gRNA designs were analyzed *in silico* using the Broad Institute's sgRNA Designer (http://www.broadinstitute.org/rnai/public/analysis-tools/sgrna-design), now called CRISPick (https://portals.broadinstitute.org/gppx/crispick/public),[53,54] and the Massachusetts Institute of Technology's former CRISPR Design (http://www.crispr.mit.edu)[39] tools. gRNAs overlapping the pathogenic variant site were listed and compared for their on-target score, number of off-targets, and the distance of the mutation site to the PAM sequence (Figures S1A and S1B). The number of off-target matches reported includes coding regions and non-coding regions of coding genes as well as cutting frequency determination scores between 0.2 and 1.0, inclusively. Given the desire for allele discrimination, the two guide sequences chosen were those for which the mutation site was in the seed sequence proximal to the PAM sequence and that scored highest for each design tool.

### Cloning

Plasmid pSpCas9(BB)-2A-GFP (PX458) was obtained from Addgene (catalog no. 48138; (Watertown, MA), courtesy of Dr. Feng Zhang, and selected guide sequences were cloned according to the protocol described in Ran et al.[55] Following cloning, plasmid DNA was isolated using an endotoxin-free maxiprep kit (Qiagen, Germantown, MD). Plasmids were verified by Sanger sequencing (GENEWIZ, South Plainfield, NJ).

### Subjects

Skin biospies were obtained based on standard operating procedures from subjects recruited at the NIH through the Pediatric Neuromuscular Clinic (12-N-0095, approved by the institutional review board of the National Institute of Neurological Disorders and Stroke [NINDS]). The four patients were two males and two females, aged between 6 and 30 years old at the time of skin biopsy (average = 18.8 years old). The site of the skin biopsy could only be traced back for one of the patients as the flank. The control was an unidentified patient (not neuromuscular-related disease) of a maximum of 4 years of age at the time of skin biopsy.

### Cell culture and nucleofection of Cas9/gRNA plasmids

Normal control and patient dermal fibroblast cell lines were established from skin biopsies. Cells were maintained in DMEM (Gibco/Thermo Fisher Scientific, Waltham, MA) with 10% fetal bovine serum (FBS) (Corning/Thermo Fisher Scientific) and 1% penicillin/streptomycin (P/S) (Gibco/Thermo Fisher Scientific) in 5% $CO_2$ at 37°C. Prior to conducting experiments, cells were tested for mycoplasma contamination (MycoAlert mycoplasma detection kit, Lonza, Basel, Switzerland) and treated if necessary (MycoZap Reagent, Lonza). At the time of nucleofection, 2–4 × $10^5$ cells were resuspended in 100 μL of P2 medium (Lonza) containing 10 μg Cas9 or Cas9/guide plasmidic DNA and transferred to a cuvette (Lonza). Cells were nucleofected with the DT-130 program using the 4D-Nucleofector system with Core and X unit (Lonza). Nucleofected cells were incubated for 10 min at room temperature before being transferred to a 6-well plate containing pre-warmed and equilibrated medium (DMEM + 10% FBS). After 48 h, cells were trypsinized, washed twice with PBS, and resuspended in 500 μL cell sorting medium (145 mM NaCl, 5 mM KCl, 1.8 mM $CaCl_2$, 0.8 mM $MgCl_2$, 10 mM HEPES, and 10 mM glucose) containing 1 μg/mL DAPI (Sigma-Aldrich, St. Louis, MO).

### Cell sorting and Cell culture

Cells were sorted at the NINDS Flow and Imaging Cytometry Core Facility on the MoFlo Astrios Cell Sorter (Beckman Coulter Life Sciences, Indianapolis, IN). The strategy for gating was to first exclude

dead cells using 1 μg/mL DAPI staining and then sort on viable $DAPI^-$ cells that also expressed GFP. The threshold for GFP expression was established using untransfected cells as controls. $GFP^+$ cells were either sorted in complete culture medium for cell culture and DNA extractions (DMEM + 10% FBS + 1% P/S + 1% amphotericin B) or directly in Trizol LS (Invitrogen/Thermo Scientific) for RNA extractions.

For cell culture, cells were pelleted and washed once in culture medium, and then were either resuspended in complete medium and plated in a single dish or resuspended in conditioned medium and plated by serial dilution in 96-well plates. Conditioned medium was obtained by collecting medium from WT primary fibroblasts that were cultured for at least 2 days to confluency and by filtering the conditioned medium with 0.2 μm filters. In the 96-well plates, wells in which a single cell had been seeded were flagged for observation. Once cells reached confluency in the flagged wells, they were successively expanded in larger wells until the population size was large enough to perform assessments.

#### DNA/RNA isolation, amplification, and Sanger sequencing

DNA was isolated from fibroblast cells using the Puregene Kit (Qiagen) according to instructions. DNA was amplified using the 2× KAPA Taq ReadyMix PCR Kit (Roche Sequencing, Indianapolis, IN) for 35 cycles using an annealing temperature of 64°C. Primers for this reaction were as follows: forward, 5′-cacactgcctgttccttgtg-3′, and reverse, 5′-gtcgagcctcactcaccttc-3′. PCR products were sent for purification and Sanger sequencing (GENEWIZ), and sequencing reads were aligned to the reference sequence using DNASTAR Lasergene (DNASTAR, Madison, WI). Sequence chromatograms were read manually to analyze the composition of edited alleles at the c.868G>A location.

Total RNA was isolated from fibroblast cells with TRIzol (Invitrogen/Thermo Fisher Scientific) according to the manufacturer's instructions. RNA, 500 ng, was treated with recombinant DNaseI using the DNA-free DNA Removal Kit (Ambion/Thermo Fisher). The DNA-free RNA sample was then used for reverse transcription, using the SuperScript III Reverse Transcriptase (Invitrogen/Thermo Fisher) and random primers. Amplification was performed using the Advantage 2 Polymerase Mix (Takara Bio USA, Mountain View, CA) and the following primers: forward, 5′-ccatcgtggacatgatcaaa-3′, and reverse 5′-ccctcgtctccagatggtc-3′. Sanger sequencing was performed and analyzed as described above.

#### Targeted re-sequencing

Following cell sorting, cells were pelleted by centrifugation at 10,000 × *g* for 1 min and washed once with PBS. DNA was isolated as described above. DNA library for targeted re-sequencing was prepared using the following procedure. A first PCR reaction was prepared from genomic DNA using the KAPA HiFi HotStart PCR Kit (Roche Sequencing) in a total reaction volume of 20 μL, with an annealing temperature of 68°C for 15 cycles (except for the three Pt3 samples, for which 25 cycles were used at this step because of low DNA yield). The forward primers for this reaction (5′-ctccttggcccaaatcctat-3′) were each extended in 5′ with a nucleotide stagger, a barcode unique to each sample, and the Illumina forward extension (5′-acactctttccctacacgacgctcttccgatct-3′), while the reverse primers (5′-agagaccagctccgaggtc-3′) were also extended with a nucleotide stagger, a barcode unique to each sample, and the Illumina reverse extension (5′- gtgactggagttcagacgtgtgctcttccgatct-3′). A second round of amplification was performed, in duplicate, using 8 μL PCR products from the first round, using the same conditions as above, except for the number of cycles, which was increased to 25. The primers used for the second round were the Illumina adapter primers forward (5′-aatgatacggcgaccaccgagatctacactctttccctacacgac-3′) and reverse (5′-caagcagaagacggcatacgagatgtgactggagttcagacgtgt-3′). Reaction duplicates were pooled and purified with the QIAquick PCR Purification Kit (Qiagen). Products were sent for targeted re-sequencing on the Illumina MiSeq platform at the NIH Intramural Sequencing Center (https://nisc.nih.gov/index.htm). For the second series of targeted re-sequencing experiments (Figure 2), the same protocol was applied, with the exception of the gene-specific reverse primer in the first round of amplification that was modified (5′-tagtgctgtgcaaggctgag-3′) to generate a shorter product.

Paired-end sequencing reads were analyzed with CRISPResso2 (http://crispresso.pinellolab.partners.org)[56] using the default settings, with the exception of the quantification window size, which was set to 4 bp. Reads were classified as unedited or containing indels. Reads that could not be called either WT or G290R, because the indel encompassed the c.868 nucleotide site, were classified as ambiguous.

#### Immunostaining

Fibroblasts ($0.8 \times 10^4$) were seeded into 8-chamber tissue culture slides (Corning/Thermo Fisher Scientific). The culture medium was replaced the following day to supplement with 50 μg/mL L-ascorbic acid (Wako Chemicals USA, Richmond, VA). After 2 days, the medium was once again replaced to maintain a sufficient supply of L-ascorbic acid. After a total of 3–4 days of L-ascorbic acid treatment, cells were washed 1× in PBS and then fixed for 10 min at room temperature in 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, PA). Matrix immunostaining was performed using the same procedure and antibodies as described previously.[57] Antibody used for fibronectin was F3648 (Sigma) at a dilution of 1:800.

#### Microscopy and image analysis

Epifluorescence images were acquired on an inverted Nikon Eclipse Ti microscope (Nikon Instruments, Melville, NY) equipped with an sCMOS pco.edge 4.2 LT camera (Excelitas PCO GmbH, Kelheim, Germany). Confocal images were acquired using a TCS SP5 II system (Leica Microsystems, Buffalo Grove, IL), with 40× or 63× objectives. A thickness of 0.508 μm was obtained, and z stacks were acquired using 0.5-μm-sized steps. Collagen VI matrix deposition was quantified according to methods set forth in Bolduc et al.[17] For collagen VI/fibronectin imaging and analysis, confocal images

were acquired using a Stellaris 8 system (Leica), using the 40× objective, for a total of 3.12-µm-stacked images. Stacks were merged and analyzed using Aivia Go version 13.1 (Leica). Collagen VI and fibronectin signals first were detected using the pixel classifier trained on at least eight separate images and then threshold using cell count recipes included in the package. Using the Object Relation Tool, collagen VI "objects" with ≥25% fibronectin overlap were compiled, and their surface area summed. The overlap area was divided by the total collagen VI area to determine the percentage of collagen VI overlap. Object circularity, a measure of how circular an object is and measured with the formula below, was analyzed for collagen VI.

$$\text{Object circularity} = \frac{4\pi \times (\text{object area})}{(\text{object perimeter})^2}$$

With this calculation, a value of 1 indicates a perfect circle. For high-resolution image acquisition, the Lightning basic detection package of the Stellaris 8 system (Leica) was used.

## DATA AND CODE AVAILABILITY

Targeted re-sequencing data from this study have been submitted to the NCBI Sequence Read Archive (BioProject ID PRJNA1023208).

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2024.102269.

## ACKNOWLEDGMENTS

We thank Dragan Maric at the NINDS Flow and Imaging Cytometry Core Facility, as well as Alice Young and the NIH Intramural Sequencing Center staff. We thank Fady Guirguis and Janelle Hauserman for their input. We thank Feng Zhang for providing the Cas9 plasmid (Addgene catalog no. 48138). We thank Mon-Li Chu for providing the collagen α1(VI) antibody. This work was supported by the Division of Intramural Research of the NIH, NINDS (1ZIANS003129). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

## AUTHOR CONTRIBUTIONS

Conceptualization, V.B., K.S., and C.G.B. Methodology, K.S., V.B., and A.S. Investigation, V.B., K.S., E.E., A.B., G.S.C., and P.U. Data curation, K.R.J., E.E., and P.U. Software, E.E. and P.U. Writing – original draft, V.B. Writing – review & editing, V.B., K.S., E.E., A.B., G.S.C., A.S., P.U., and C.G.B. Visualization, V.B., K.S., E.E., and A.B. Supervision, V.B. and C.G.B. Funding acquisition, C.G.B.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

1. Bönnemann, C.G. (2011). The collagen VI-related myopathies: muscle meets its matrix. Nat. Rev. Neurol. *7*, 379–390. https://doi.org/10.1038/nrneurol.2011.81.
2. Lamande, S.R., and Bateman, J.F. (2018). Collagen VI disorders: Insights on form and function in the extracellular matrix and beyond. Matrix Biol. *71–72*, 348–367. https://doi.org/10.1016/j.matbio.2017.12.008.
3. Zou, Y., Zhang, R.Z., Sabatelli, P., Chu, M.L., and Bonnemann, C.G. (2008). Muscle interstitial fibroblasts are the main source of collagen VI synthesis in skeletal muscle: implications for congenital muscular dystrophy types Ullrich and Bethlem. J. Neuropathol. Exp. Neurol. *67*, 144–154. https://doi.org/10.1097/nen.0b013e3181634ef7.
4. Braghetta, P., Ferrari, A., Fabbro, C., Bizzotto, D., Volpin, D., Bonaldo, P., and Bressan, G.M. (2008). An enhancer required for transcription of the Col6a1 gene in muscle connective tissue is induced by signals released from muscle cells. Exp. Cell Res. *314*, 3508–3518. https://doi.org/10.1016/j.yexcr.2008.08.006.
5. Chu, M.L., Conway, D., Pan, T.C., Baldwin, C., Mann, K., Deutzmann, R., and Timpl, R. (1988). Amino acid sequence of the triple-helical domain of human collagen type VI. J. Biol. Chem. *263*, 18601–18606.
6. Furthmayr, H., Wiedemann, H., Timpl, R., Odermatt, E., and Engel, J. (1983). Electron-microscopical approach to a structural model of intima collagen. Biochem. J. *211*, 303–311.
7. Engvall, E., Hessle, H., and Klier, G. (1986). Molecular assembly, secretion, and matrix deposition of type VI collagen. J. Cell Biol. *102*, 703–710.
8. Allamand, V., Brinas, L., Richard, P., Stojkovic, T., Quijano-Roy, S., and Bonne, G. (2011). ColVI myopathies: where do we stand, where do we go? Skeletal Muscle *1*, 30. https://doi.org/10.1186/2044-5040-1-30.
9. Pace, R.A., Peat, R.A., Baker, N.L., Zamurs, L., Morgelin, M., Irving, M., Adams, N.E., Bateman, J.F., Mowat, D., Smith, N.J., et al. (2008). Collagen VI glycine mutations: perturbed assembly and a spectrum of clinical severity. Ann. Neurol. *64*, 294–303. https://doi.org/10.1002/ana.21439.
10. Butterfield, R.J., Foley, A.R., Dastgir, J., Asman, S., Dunn, D.M., Zou, Y., Hu, Y., Donkervoort, S., Flanigan, K.M., Swoboda, K.J., et al. (2013). Position of Glycine Substitutions in the Triple Helix of COL6A1, COL6A2, and COL6A3 is Correlated with Severity and Mode of Inheritance in Collagen VI Myopathies. Hum. Mutat. *34*, 1558–1567. https://doi.org/10.1002/humu.22429.
11. Lamandé, S.R., Morgelin, M., Selan, C., Jobsis, G.J., Baas, F., and Bateman, J.F. (2002). Kinked collagen VI tetramers and reduced microfibril formation as a result of Bethlem myopathy and introduced triple helical glycine mutations. J. Biol. Chem. *277*, 1949–1956. https://doi.org/10.1074/jbc.M109932200.
12. Mohassel, P., Foley, A.R., and Bonnemann, C.G. (2018). Extracellular matrix-driven congenital muscular dystrophies. Matrix Biol. *71–72*, 188–204. https://doi.org/10.1016/j.matbio.2018.06.005.
13. Foley, A.R., Hu, Y., Zou, Y., Yang, M., Medne, L., Leach, M., Conlin, L.K., Spinner, N., Shaikh, T.H., Falk, M., et al. (2011). Large genomic deletions: a novel cause of Ullrich congenital muscular dystrophy. Ann. Neurol. *69*, 206–211. https://doi.org/10.1002/ana.22283.
14. Levin, A.A. (2019). Treating Disease at the RNA Level with Oligonucleotides. N. Engl. J. Med. *380*, 57–70. https://doi.org/10.1056/NEJMra1705346.
15. Gualandi, F., Manzati, E., Sabatelli, P., Passarelli, C., Bovolenta, M., Pellegrini, C., Perrone, D., Squarzoni, S., Pegoraro, E., Bonaldo, P., and Ferlini, A. (2012). Antisense-induced messenger depletion corrects a COL6A2 dominant mutation in Ullrich myopathy. Hum. Gene Ther. *23*, 1313–1318. https://doi.org/10.1089/hum.2012.109.
16. Noguchi, S., Ogawa, M., Kawahara, G., Malicdan, M.C., and Nishino, I. (2014). Allele-specific Gene Silencing of Mutant mRNA Restores Cellular Function in Ullrich Congenital Muscular Dystrophy Fibroblasts. Nucleic Acids *3*, e171. https://doi.org/10.1038/mtna.2014.22.
17. Bolduc, V., Zou, Y., Ko, D., and Bonnemann, C.G. (2014). siRNA-mediated Allele-specific Silencing of a COL6A3 Mutation in a Cellular Model of Dominant Ullrich Muscular Dystrophy. Mol. Ther. Nucleic Acids *3*, e147. https://doi.org/10.1038/mtna.2013.74.
18. Marrosu, E., Ala, P., Muntoni, F., and Zhou, H. (2017). Gapmer Antisense Oligonucleotides Suppress the Mutant Allele of COL6A3 and Restore Functional Protein in Ullrich Muscular Dystrophy. Mol. Ther. Nucleic Acids *8*, 416–427. https://doi.org/10.1016/j.omtn.2017.07.006.

19. Wu, J., Tang, B., and Tang, Y. (2020). Allele-specific genome targeting in the development of precision medicine. Theranostics *10*, 3118–3137. https://doi.org/10.7150/thno.43298.

20. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821. https://doi.org/10.1126/science.1225829.

21. Chang, H.H.Y., Pannunzio, N.R., Adachi, N., and Lieber, M.R. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nat. Rev. Mol. Cell Biol. *18*, 495–506. https://doi.org/10.1038/nrm.2017.48.

22. Xue, C., and Greene, E.C. (2021). DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. Trends Genet. *37*, 639–656. https://doi.org/10.1016/j.tig.2021.02.008.

23. Christie, K.A., Courtney, D.G., DeDionisio, L.A., Shern, C.C., De Majumdar, S., Mairs, L.C., Nesbit, M.A., and Moore, C.B.T. (2017). Towards personalised allele-specific CRISPR gene editing to treat autosomal dominant disorders. Sci. Rep. *7*, 16174. https://doi.org/10.1038/s41598-017-16279-4.

24. Li, P., Kleinstiver, B.P., Leon, M.Y., Prew, M.S., Navarro-Gomez, D., Greenwald, S.H., Pierce, E.A., Joung, J.K., and Liu, Q. (2018). Allele-Specific CRISPR-Cas9 Genome Editing of the Single-Base P23H Mutation for Rhodopsin-Associated Dominant Retinitis Pigmentosa. CRISPR J. *1*, 55–64. https://doi.org/10.1089/crispr.2017.0009.

25. Gyorgy, B., Loov, C., Zaborowski, M.P., Takeda, S., Kleinstiver, B.P., Commins, C., Kastanenka, K., Mu, D., Volak, A., Giedraitis, V., et al. (2018). CRISPR/Cas9 Mediated Disruption of the Swedish APP Allele as a Therapeutic Approach for Early-Onset Alzheimer's Disease. Molecular therapy. Nucleic Acids *11*, 429–440. https://doi.org/10.1016/j.omtn.2018.03.007.

26. György, B., Nist-Lund, C., Pan, B., Asai, Y., Karavitaki, K.D., Kleinstiver, B.P., Garcia, S.P., Zaborowski, M.P., Solanes, P., Spataro, S., et al. (2019). Allele-specific gene editing prevents deafness in a model of dominant progressive hearing loss. Nat. Med. *25*, 1123–1130. https://doi.org/10.1038/s41591-019-0500-9.

27. Rabai, A., Reisser, L., Reina-San-Martin, B., Mamchaoui, K., Cowling, B.S., Nicot, A.S., and Laporte, J. (2019). Allele-Specific CRISPR/Cas9 Correction of a Heterozygous DNM2 Mutation Rescues Centronuclear Myopathy Cell Phenotypes. Mol. Ther. Nucleic Acids *16*, 246–256. https://doi.org/10.1016/j.omtn.2019.02.019.

28. Feliciano, C.M., Wu, K., Watry, H.L., Marley, C.B.E., Ramadoss, G.N., Ghanim, H.Y., Liu, A.Z., Zholudeva, L.V., McDevitt, T.C., Saporta, M.A., et al. (2021). Allele-Specific Gene Editing Rescues Pathology in a Human Model of Charcot-Marie-Tooth Disease Type 2E. Front. Cell Dev. Biol. *9*, 723023. https://doi.org/10.3389/fcell.2021.723023.

29. Shin, J.W., Hong, E.P., Park, S.S., Choi, D.E., Seong, I.S., Whittaker, M.N., Kleinstiver, B.P., Chen, R.Z., and Lee, J.M. (2022). Allele-specific silencing of the gain-of-function mutation in Huntington's disease using CRISPR/Cas9. JCI Insight *7*, e141042. https://doi.org/10.1172/jci.insight.141042.

30. Jimenez-Mallebrera, C., Maioli, M.A., Kim, J., Brown, S.C., Feng, L., Lampe, A.K., Bushby, K., Hicks, D., Flanigan, K.M., Bonnemann, C., et al. (2006). A comparative analysis of collagen VI production in muscle, skin and fibroblasts from 14 Ullrich congenital muscular dystrophy patients with dominant and recessive COL6A mutations. Neuromuscul. Disord. *16*, 571–582. https://doi.org/10.1016/j.nmd.2006.07.015.

31. Sabatelli, P., Bonaldo, P., Lattanzi, G., Braghetta, P., Bergamin, N., Capanni, C., Mattioli, E., Columbaro, M., Ognibene, A., Pepe, G., et al. (2001). Collagen VI deficiency affects the organization of fibronectin in the extracellular matrix of cultured fibroblasts. Matrix Biol. *20*, 475–486.

32. Brull, A., Sarathy, A., Bolduc, V., Chen, G.S., McCarty, R.M., and Bonnemann, C.G. (2024). Optimized allele-specific silencing of the dominant-negative *COL6A1* G293R substitution causing collagen VI-related dystrophy. Nucleic Acids *35*, 102178. https://doi.org/10.1016/j.omtn.2024.102178.

33. Shen, M.W., Arbab, M., Hsu, J.Y., Worstell, D., Culbertson, S.J., Krabbe, O., Cassa, C.A., Liu, D.R., Gifford, D.K., and Sherwood, R.I. (2018). Predictable and precise template-free CRISPR editing of pathogenic variants. Nature *563*, 646–651. https://doi.org/10.1038/s41586-018-0686-x.

34. Allen, F., Crepaldi, L., Alsinet, C., Strong, A.J., Kleshchevnikov, V., De Angeli, P., Palenikova, P., Khodak, A., Kiselev, V., Kosicki, M., et al. (2018). Predicting the mutations generated by repair of Cas9-induced double-strand breaks. Nat. Biotechnol. *37*, 64–72. https://doi.org/10.1038/nbt.4317.

35. Monteys, A.M., Ebanks, S.A., Keiser, M.S., and Davidson, B.L. (2017). CRISPR/Cas9 Editing of the Mutant Huntingtin Allele In Vitro and In Vivo. Mol. Ther. *25*, 12–23. https://doi.org/10.1016/j.ymthe.2016.11.010.

36. Bchetnia, M., Dionne Gagne, R., Powell, J., Morin, C., McCuaig, C., Duperee, A., Germain, L., Tremblay, J.P., and Laprise, C. (2022). Allele-Specific Inactivation of an Autosomal Dominant Epidermolysis Bullosa Simplex Mutation Using CRISPR-Cas9. CRISPR J. *5*, 586–597. https://doi.org/10.1089/crispr.2021.0132.

37. Keough, K.C., Lyalina, S., Olvera, M.P., Whalen, S., Conklin, B.R., and Pollard, K.S. (2019). AlleleAnalyzer: a tool for personalized and allele-specific sgRNA design. Genome Biol. *20*, 167. https://doi.org/10.1186/s13059-019-1783-3.

38. Rabinowitz, R., and Offen, D. (2021). Single-Base Resolution: Increasing the Specificity of the CRISPR-Cas System in Gene Editing. Mol. Ther. *29*, 937–948. https://doi.org/10.1016/j.ymthe.2020.11.009.

39. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832. https://doi.org/10.1038/nbt.2647.

40. Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat. Biotechnol. *32*, 279–284. https://doi.org/10.1038/nbt.2808.

41. Zheng, T., Hou, Y., Zhang, P., Zhang, Z., Xu, Y., Zhang, L., Niu, L., Yang, Y., Liang, D., Yi, F., et al. (2017). Profiling single-guide RNA specificity reveals a mismatch sensitive core sequence. Sci. Rep. *7*, 40638. https://doi.org/10.1038/srep40638.

42. Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature *529*, 490–495. https://doi.org/10.1038/nature16526.

43. Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. Science *351*, 84–88. https://doi.org/10.1126/science.aad5227.

44. Ikeda, A., Fujii, W., Sugiura, K., and Naito, K. (2019). High-fidelity endonuclease variant HypaCas9 facilitates accurate allele-specific gene modification in mouse zygotes. Commun. Biol. *2*, 371. https://doi.org/10.1038/s42003-019-0627-8.

45. Maule, G., Casini, A., Montagna, C., Ramalho, A.S., De Boeck, K., Debyser, Z., Carlon, M.S., Petris, G., and Cereseto, A. (2019). Allele specific repair of splicing mutations in cystic fibrosis through AsCas12a genome editing. Nat. Commun. *10*, 3556. https://doi.org/10.1038/s41467-019-11454-9.

46. Wang, Y., Qi, T., Liu, J., Yang, Y., Wang, Z., Wang, Y., Wang, T., Li, M., Li, M., Lu, D., et al. (2023). A highly specific CRISPR-Cas12j nuclease enables allele-specific genome editing. Sci. Adv. *9*, eabo6405. https://doi.org/10.1126/sciadv.abo6405.

47. López-Márquez, A., Morin, M., Fernandez-Penalver, S., Badosa, C., Hernandez-Delgado, A., Natera-de Benito, D., Ortez, C., Nascimento, A., Grinberg, D., Balcells, S., et al. (2022). CRISPR/Cas9-Mediated Allele-Specific Disruption of a Dominant COL6A1 Pathogenic Variant Improves Collagen VI Network in Patient Fibroblasts. Int. J. Mol. Sci. *23*, 4410. https://doi.org/10.3390/ijms23084410.

48. Donkervoort, S., Hu, Y., Stojkovic, T., Voermans, N.C., Foley, A.R., Leach, M.E., Dastgir, J., Bolduc, V., Cullup, T., de Becdelievre, A., et al. (2015). Mosaicism for dominant collagen 6 mutations as a cause for intrafamilial phenotypic variability. Hum. Mutat. *36*, 48–56. https://doi.org/10.1002/humu.22691.

49. D'Amico, A., Fattori, F., Tasca, G., Petrini, S., Gualandi, F., Bruselles, A., D'Oria, V., Verardo, M., Carrozzo, R., Niceta, M., et al. (2017). Somatic mosaicism represents an underestimated event underlying collagen 6-related disorders. Eur. J. Paediatr. Neurol. *21*, 873–883. https://doi.org/10.1016/j.ejpn.2017.07.009.

50. Frangoul, H., Altshuler, D., Cappellini, M.D., Chen, Y.S., Domm, J., Eustace, B.K., Foell, J., de la Fuente, J., Grupp, S., Handgretinger, R., et al. (2021). CRISPR-Cas9 Gene Editing for Sickle Cell Disease and beta-Thalassemia. N. Engl. J. Med. *384*, 252–260. https://doi.org/10.1056/NEJMoa2031054.

51. Gillmore, J.D., Gane, E., Taubel, J., Kao, J., Fontana, M., Maitland, M.L., Seitzer, J., O'Connell, D., Walsh, K.R., Wood, K., et al. (2021). CRISPR-Cas9 In Vivo Gene Editing for Transthyretin Amyloidosis. N. Engl. J. Med. *385*, 493–502. https://doi.org/10.1056/NEJMoa2107454.

52. Christie, K.A., Robertson, L.J., Conway, C., Blighe, K., DeDionisio, L.A., Chao-Shern, C., Kowalczyk, A.M., Marshall, J., Turnbull, D., Nesbit, M.A., and Moore, C.B.T. (2020). Mutation-Independent Allele-Specific Editing by CRISPR-Cas9, a Novel Approach to Treat Autosomal Dominant Disease. Mol. Ther. *28*, 1846–1857. https://doi.org/10.1016/j.ymthe.2020.05.002.

53. Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol. *34*, 184–191. https://doi.org/10.1038/nbt.3437.

54. Sanson, K.R., Hanna, R.E., Hegde, M., Donovan, K.F., Strand, C., Sullender, M.E., Vaimberg, E.W., Goodale, A., Root, D.E., Piccioni, F., and Doench, J.G. (2018). Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. Nat. Commun. *9*, 5416. https://doi.org/10.1038/s41467-018-07901-8.

55. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308. https://doi.org/10.1038/nprot.2013.143.

56. Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, J.K., Bauer, D.E., and Pinello, L. (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat. Biotechnol. *37*, 224–226. https://doi.org/10.1038/s41587-019-0032-3.

57. Bolduc, V., Foley, A.R., Solomon-Degefa, H., Sarathy, A., Donkervoort, S., Hu, Y., Chen, G.S., Sizov, K., Nalls, M., Zhou, H., et al. (2019). A recurrent COL6A1 pseudoexon insertion causes muscular dystrophy and is effectively targeted by splice-correction therapies. JCI Insight *4*, e124403. https://doi.org/10.1172/jci.insight.124403.

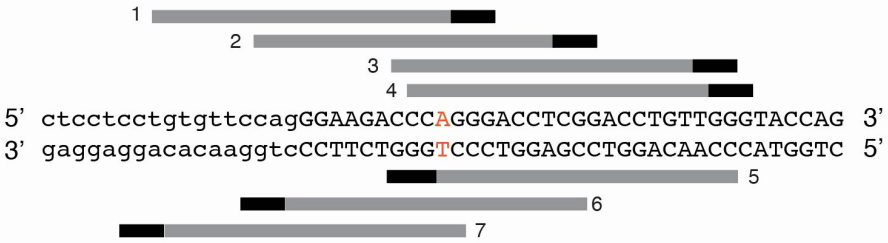
**OMTN, Volume *35***

## Supplemental information

# Allele-specific CRISPR-Cas9 editing inactivates a single nucleotide variant associated with collagen VI muscular dystrophy

**Véronique Bolduc, Katherine Sizov, Astrid Brull, Eric Esposito, Grace S. Chen, Prech Uapinyoying, Apurva Sarathy, Kory R. Johnson, and Carsten G. Bönnemann**

**A**




**B**

| ID | Target / PAM | Strand | Broad efficacy score | Broad combined rank | MIT specificity score (0-100) | Number Off-targets (0-1-2 mismatches) | Location of variant vs PAM |
|---|---|---|---|---|---|---|---|
| 5* | CCAACAGGTCCGAGGTCCCT/GGG | (-) | -0.03 | 3 | 63 | 0-1-0 | proximal |
| 4 | CCAGGGACCTCGGACCTGTT/GGG | (+) | -0.26 | 7 | 47 | 0-1-0 | distal |
| 3 | CCCAGGGACCTCGGACCTGT/TGG | (+) | -0.21 | 4 | 44 | 0-1-1 | distal |
| 1** | tgtgttccagGGAAGACCCA/GGG | (+) | 0.86 | 1 | 43 | 0-2-6 | proximal |
| 6 | CGAGGTCCCTGGGTCTTCCc/tgg | (-) | -0.03 | 5 | 39 | 0-2-0 | distal |
| 7 | CTGGGTCTTCCctggaacac/agg | (-) | 0.11 | 6 | 35 | 0-1-10 | distal |
| 2 | cagGGAAGACCCAGGGACCT/CGG | (+) | 0.13 | 2 | 32 | 0-2-6 | proximal |

*gRNA-B **gRNA-A

**C**

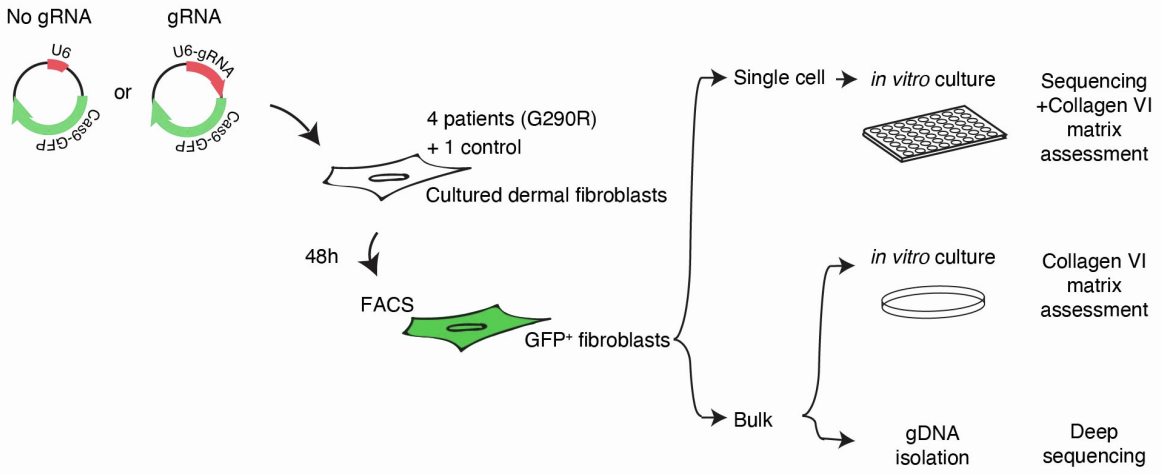



**Figure S1. Strategy for allele-specific gene editing of the *COL6A1* c.868G>A (G290R) variant.**

**(A)** The localization of potential allele-specific gRNAs for the c.868G>A variant is illustrated along the *COL6A1* genomic sequence. gRNAs hybridizing to the negative (-) strand are shown below, while gRNAs hybridizing to the positive (+) strand are shown above the sequence. **(B)** The sequences of all potential gRNAs are listed, together with the sequences of the adjacent PAM sites.

The presence of the mutation within a 10-nt distance of the PAM site was considered proximal. ID 1 was chosen as gRNA-A and ID 5 was chosen as gRNA-B. **(C)** Schematics of the experimental procedures followed in the study. Note that all experiments were conducted on $GFP^{+}$-sorted cells (enriching for cells transfected with the gene editing constructs).

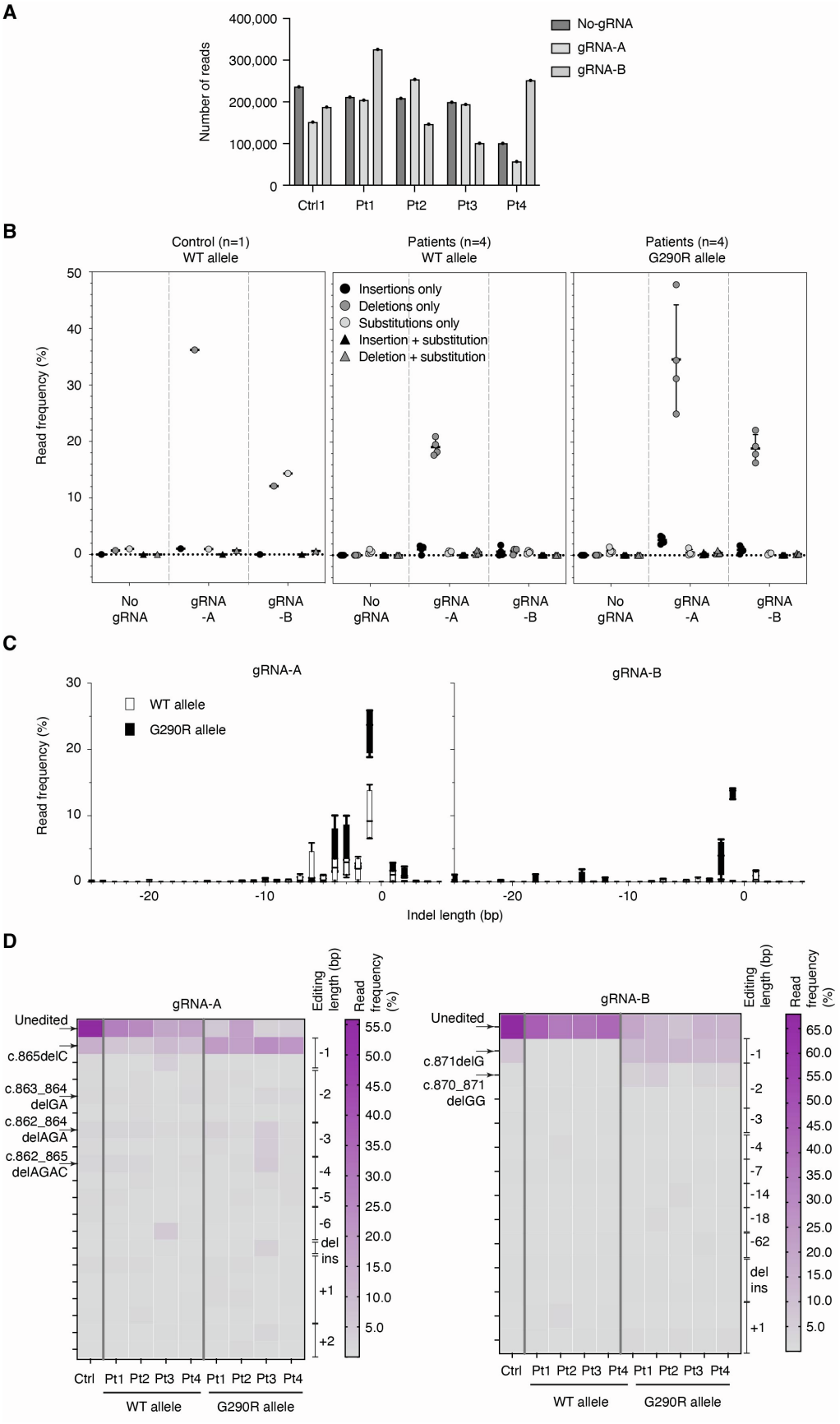

A
Number of reads
400,000
300,000
200,000
100,000
0
No-gRNA
gRNA-A
gRNA-B
Ctrl1
Pt1
Pt2
Pt3
Pt4
B
Control (n=1)
WT allele
Patients (n=4)
WT allele
Patients (n=4)
G290R allele
Read frequency (%)
Insertions only
Deletions only
Substitutions only
Insertion + substitution
Deletion + substitution
No gRNA
gRNA -A
gRNA -B
C
gRNA-A
gRNA-B
WT allele
G290R allele
Read frequency (%)
Indel length (bp)
D
gRNA-A
gRNA-B
Editing length (bp)
Read frequency (%)
Unedited
c.865delC
c.863_864 delGA
c.862_864 delAGA
c.862_865 delAGAC
c.871delG
c.870_871 delGG
del
ins
Ctrl
Pt1 Pt2 Pt3 Pt4
WT allele
G290R allele

**Figure S2. Targeted re-sequencing analysis of the repair outcomes after gene editing with gRNA-A or gRNA-B.**
Analysis of Illumina Mi-Seq sequencing reads for four patient and one control primary cells co-transfected with Cas9-GFP and with either gRNA-A, gRNA-B, or without gRNA (No gRNA). **(A)** Total number of sequencing reads, per sample. The total number of reads was subsequently used as the denominator to calculate the read frequency in each sample. **(B)** Total read frequencies per indel/edit type (deletions, insertions, substitutions, or combination thereof) and per allele type (WT or G290R). Bars represent average ± standard deviation. **(C)** Total read frequencies per indel length, and per allele type, reported as box and whisker plots, for gRNA-A (left) and gRNA-B (right). On the x axes, the negative scale represents length of deletions, while the positive site represents length of insertions. **(D)** Heatmaps of the top 20 motifs for their total read frequencies, identified at the G290R allele in patient samples, and displayed by individual.

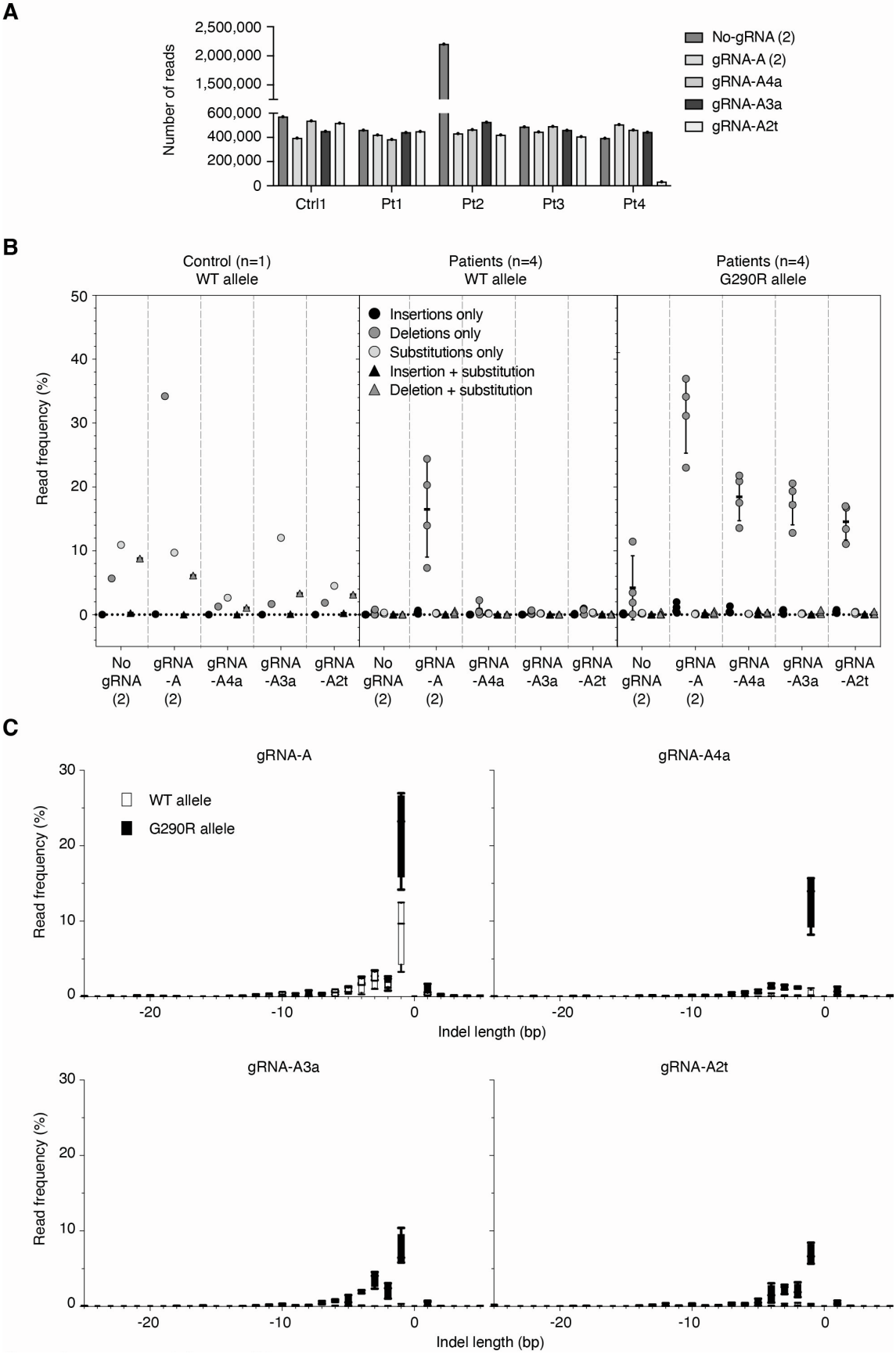

A
Number of reads
2,500,000
2,000,000
1,500,000
600,000
400,000
200,000
0
Ctrl1
Pt1
Pt2
Pt3
Pt4
No-gRNA (2)
gRNA-A (2)
gRNA-A4a
gRNA-A3a
gRNA-A2t
B
Control (n=1)
WT allele
Patients (n=4)
WT allele
Patients (n=4)
G290R allele
Read frequency (%)
Insertions only
Deletions only
Substitutions only
Insertion + substitution
Deletion + substitution
No gRNA (2)
gRNA -A (2)
gRNA -A4a
gRNA -A3a
gRNA -A2t
C
gRNA-A
gRNA-A4a
gRNA-A3a
gRNA-A2t
WT allele
G290R allele
Read frequency (%)
Indel length (bp)

**Figure S3. Targeted re-sequencing analysis of the repair outcomes after gene editing with gRNA-A4A, gRNA-A3A or gRNA-A2T.**
Analysis of Illumina Mi-Seq sequencing reads for four patient and one control primary cells co-transfected with Cas9-GFP and with either gRNA-A, gRNA-A4a, gRNA-A3a, gRNA-A2t, or without gRNA (No gRNA). Note that gRNA-A and No gRNA in this experiment are replicates from the previous experiment. **(A)** Total number of sequencing reads, per sample. The total number of reads was subsequently used as the denominator to calculate the read frequency in each sample. **(B)** Total read frequencies per indel/edit type (deletions, insertions, substitutions, or combination thereof) and per allele type (WT or G290R). Bars represent average ± standard deviation. **(C)** Total read frequencies per indel length, and per allele type, reported as box and whisker plots, for each of the gRNA. On the x axes, the negative scale represents length of deletions, while the positive site represents length of insertions. **(D)** Heatmaps of the top 20 motifs for their total read frequencies, identified at the G290R allele in patient samples, and displayed by individual.

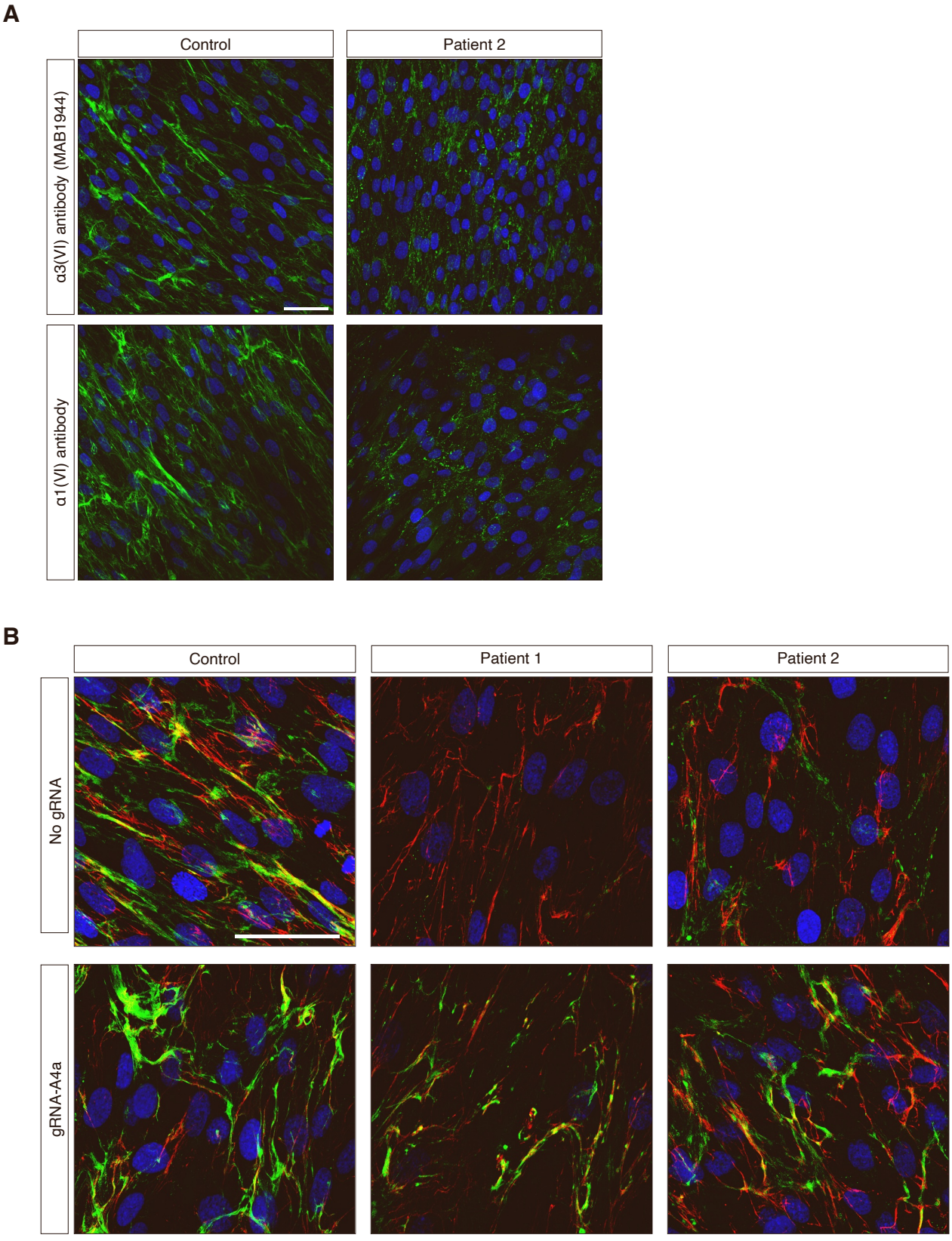

A
Control
Patient 2
α3(VI) antibody (MAB1944)
α1(VI) antibody
B
Control
Patient 1
Patient 2
No gRNA
gRNA-A4a

**Figure S4. Assessment of the collagen VI matrix in the extracellular space.**
Immunostaining of the collagen VI matrix secreted by patient-derived cultured fibroblasts. **(A)** Comparision of a Collagen α1(VI) antibody (gift of M.L. Chu) and a Collagen α3(VI) antibody (MAB1944) to detect the collagen VI matrix. Green signal is collagen VI, blue signal represents the nuclei. Bar = 50 µm. **(B)** High-resolution confocal microscopy of the collagen VI matrix (green) interaction with fibronectin (red). Nuclei are labeled in blue. Bar = 50 µm.