Citation: Molecular Therapy—Nucleic Acids (2016) 5, e307; doi:10.1038/mtna.2016.19
Official journal of the American Society of Gene & Cell Therapy

www.nature.com/mtna

Gene Editing for the Efficient Correction of a Recurrent *COL7A1* Mutation in Recessive Dystrophic Epidermolysis Bullosa Keratinocytes

Cristina Chamorro[1,2], Angeles Mencía[2,3], David Almarza[1,4], Blanca Duarte[1,2,5], Hildegard Büning[6,7], Jessica Sallach[6], Ingrid Hausser[8], Marcela Del Río[2,3,5], Fernando Larcher[1,2,3,5] and Rodolfo Murillas[1,2,5]

Clonal gene therapy protocols based on the precise manipulation of epidermal stem cells require highly efficient gene-editing molecular tools. We have combined adeno-associated virus (AAV)-mediated delivery of donor template DNA with transcription activator-like nucleases (TALE) expressed by adenoviral vectors to address the correction of the c.6527insC mutation in the *COL7A1* gene, causing recessive dystrophic epidermolysis bullosa in a high percentage of Spanish patients. After transduction with these viral vectors, high frequencies of homology-directed repair were found in clones of keratinocytes derived from a recessive dystrophic epidermolysis bullosa (RDEB) patient homozygous for the c.6527insC mutation. Gene-edited clones recovered the expression of the *COL7A1* transcript and collagen VII protein at physiological levels. In addition, treatment of patient keratinocytes with TALE nucleases in the absence of a donor template DNA resulted in nonhomologous end joining (NHEJ)-mediated indel generation in the vicinity of the c.6527insC mutation site in a large proportion of keratinocyte clones. A subset of these indels restored the reading frame of *COL7A1* and resulted in abundant, supraphysiological expression levels of mutant or truncated collagen VII protein. Keratinocyte clones corrected both by homology-directed repair (HDR) or NHEJ were used to regenerate skin displaying collagen VII in the dermo-epidermal junction.

Molecular Therapy—Nucleic Acids (2016) **5**, e307; doi:10.1038/mtna.2016.19; published online 5 April 2016
Subject Category: Gene Insertion, Deletion & Modification

Introduction

Precise modification of the genome of cultured human cells to revert disease-causing mutations has been facilitated by recent advances in molecular tools, and specifically by the development of customized programmable genomic nucleases[1] and highly efficient gene and cell targeting protocols based on recombinant adeno-associated (rAAV) vectors.[2,3] Customizable nucleases include zinc-finger nucleases, transcription activator-like nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPRs) nucleases. All of them are designed to bind specific DNA sequences and create site-specific double-strand DNA breaks (DSBs). These are repaired mostly by the nonhomologous end joining DNA repair pathway (NHEJ), which is prone to introducing small insertions and deletions (indels) at the DSB sites. However, in the presence of donor DNA molecules homology-directed repair (HDR) may occur by engaging the cellular homologous recombination DNA repair pathway, albeit at low frequency. *Ex-vivo* preclinical[4,5] and clinical[6] protocols based on TALEN and Zinc-finger nucleases show the therapeutic potential of gene editing. Adeno-associated virus (AAV) vectors are a fundamental tool in current gene replacement-based gene therapy protocols, and have also shown great promise in many preclinical and clinical trial settings.[7] In addition, AAV vectors can deliver gene-targeting constructs as single-stranded linear DNA molecules that serve as donor templates for homologous recombination, allowing high targeting frequencies[2] and have the capability to deliver DNA efficiently and without significant toxicity into hard to transduce primary cell cultures, overcoming a critical hurdle for *ex vivo* gene therapy in patient cells.[3,8–10] The use of target-specific nucleases to generate DSBs in the target locus greatly increases AAV vector-mediated gene-targeting frequencies.[11–13]

Recessive dystrophic epidermolysis bullosa (RDEB) is a hereditary skin disease caused by loss of function mutations in *COL7A1*, a gene expressed by skin keratinocytes and fibroblasts that encodes the pro-α1 (VII) chain of type VII collagen (C7). Three alpha chains fold into a triple helical conformation to give rise to the procollagen molecules which after being secreted undergo additional processing and arrange into type VII collagen fibers, the main component of the anchoring fibrils that connect the epidermal basement membrane to the dermal tissue. C7 deficiency causes loss of dermo-epidermal adhesion resulting in blister formation and

The first two authors contributed equally to this work.

[1]Epithelial Biomedicine Division, Centro de Investigaciones Energéticas Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain; [2]Instituto de Investigación Sanitaria de la Fundación Jiménez Díaz, Madrid, Spain; [3]Department of Biomedical Engineering, Carlos III University (UC3M), Madrid, Spain; [4]Present address: Institute of Cellular Medicine, University of Newcastle, Newcastle Upon Tyne, UK; [5]Centro de Investigación Biomédica en Red en Enfermedades Raras (CIBERER) U714, Madrid, Spain; [6]Center for Molecular Medicine Cologne (CMMC), University of Cologne, Köln, Germany; [7]Institute of Experimental Hematology, Hannover Medical School, Hannover, Germany; [8]Institute of Pathology, Universitätsklinikum Heidelberg, Heidelberg, Germany. Correspondence: Rodolfo Murillas, Epithelial Biomedicine Division, Cutaneous Disease Modelling Unit, CIEMAT-CIBERER (Centre for Biomedical Research on Rare Diseases) U714, Avenida Complutense 40, Madrid 28040, Spain. E-mail: rodolfo.murillas@ciemat.es Or Fernando Larcher, Epithelial Biomedicine Division, Cutaneous Disease Modelling Unit, CIEMAT-CIBERER (Centre for Biomedical Research on Rare Diseases) U714, Avenida Complutense 40, Madrid 28040, Spain. E-mail: fernando.larcher@ciemat.es

Keywords: AAV vectors; epidermolysis bullosa; gene editing; indels, homologous recombination; TALEN
Received 4 February 2016; accepted 13 February 2016; published online 5 April 2016. doi:10.1038/mtna.2016.19

scarring.[14] We have recently described a recurrent frame-shift mutation (c.6527insC) in the exon 80 of *COL7A1* gene that produces a premature termination codon (PTC). The high prevalence of this mutation (46% of RDEB alleles in Spanish patients)[15] justifies the development of targeted therapies. Although HDR allows the most precise genetic correction, that is, the restoration of the complete coding sequence of the mutated gene, strategies based on NHEJ-mediated indel generation can be conceived. The generation of indels in the vicinity of the pathogenic mutation can lead to restoration of the reading frame, although usually involving limited changes in the aminoacid sequence. When the introduction of indels results in alteration or elimination of intron splice sites, exon skipping can produce a transcript encoding a truncated form of the protein. The feasibility of frame restoration by indel generation with nucleases has previously been demonstrated in cells from Duchenne muscular dystrophy patients, achieving reading frame restoration of the dystrophin transcript and dystrophin protein expression by introducing indels with TALENs[16] and exon skipping of a PTC-containing exon by using zinc-finger nucleases.[17] The *COL7A1* gene is also suitable for exon-skipping approaches, since all exons in the *COL7A1* triple helix-coding region are very short, encode Gly-X-Y aminoacid repeats beginning with intact codons for the glycine residues and are arranged in frame.[18] Since NHEJ-mediated repair occurs much more frequently than HDR, gene editing by designer nucleases-generated indels should be much easier to implement than HDR-based gene targeting, which usually requires the introduction of antibiotic resistance cassettes to select for homologous recombination events.

Here, we address the *ex-vivo* correction of the highly recurrent c.6527insC mutation using gene-editing approaches based on site-specific nucleases and gene-targeting constructs delivered by nonintegrating viral vectors.

Results

Generation of pairs of TALE nucleases designed to cut in the proximity of the c.6527insC mutation

Pairs of TALENs were designed with TALEffector-Nucleotide Targeter design Software[19] and constructed using the Golden Gate method.[20] A total of 17 TALENs were constructed (**Supplementary Figure S1a**). Since carboxy terminal deletions of TALEN nucleases have significantly increased nuclease activity,[21] we constructed TALEN proteins retaining 30 aminoacids of the carboxy terminal region, and included a sequence encoding a nine aminoacids HA tag between the TALE binding and FokI nuclease domains (+30+HA design) to allow immunodetection with anti HA antibodies. The N-terminal domain was left intact. Expression in these constructs is driven by the phosphoglycerate kinase (PGK) promoter. We also constructed TALENs with the same DNA binding repeats but based on a design retaining 63 aminoacids at the carboxy terminus and having a 152 aminoacids N-terminal deletion, using the pCAG-T7-TALEN as destination vector (+63 design) (**Supplementary Figure S1b**).

TALEN constructs with the +30+HA design were individually transfected in 293T cells to analyze their correct expression by western blot using anti HA antibodies (**Supplementary Figure S2a**). For functionality assays, pairs of TALENs were transfected into 293T cells and tested for their ability to generate indels at the nuclease target region by performing the Cel I/ Surveyor nuclease assay.[22] All TALEN pairs tested with this protocol showed functionality to some extent, and comparable results were found for our +30+HA design and the +63 design (**Supplementary Figure S2b**). The +30+HA design was used in subsequent experimentation *since it allowed efficient immunodetection of TALEN proteins*.

TALENs delivery to patient keratinocytes by adenoviral vectors

To equip keratinocytes with TALENs, we decided for viral vector-mediated gene delivery. Specifically, we constructed bicistronic adenoviral vectors for coexpression of TALENs and a marker gene (GFP) under the control of the PGK promoter (Ad-TALENs). Of the candidates shown to be active in the transient transfection assay in 293T cells, three pairs of TALENs were chosen (**Figure 1a**).

The ability of Ad-TALENs to transduce human keratinocytes was assessed in patient-derived RDEB-E67A6 immortalized keratinocytes, carrying the highly recurrent

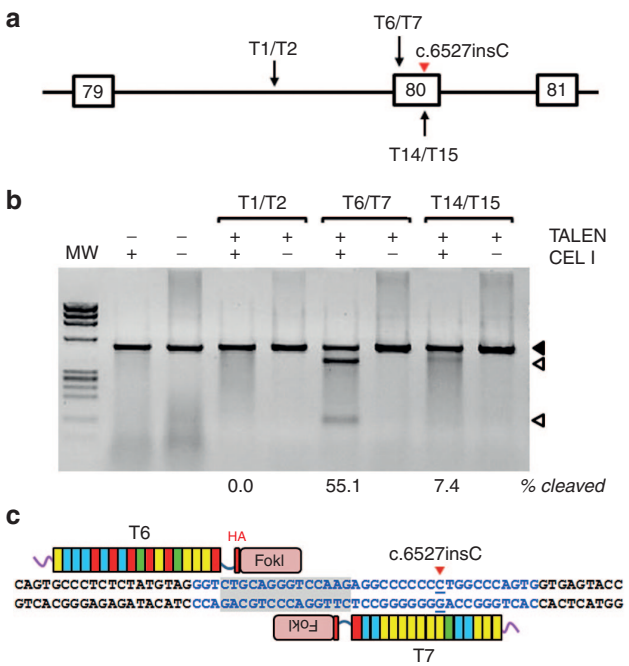

Figure 1 Design and activity of transcription activator-like nucleases (TALENs) **targeted to exon 80 and intron 79.** Three pairs of TALENs with targets across intron 79 and exon 80 (**a**) were expressed with adenoviral vectors in RDEB-E67A6 keratinocytes. A polymerase chain reaction (PCR) product spanning the TALEN target sites was generated with primers F1/R and analyzed with the Cel I (Surveyor) mutation detection assay, showing efficient indel generation activity for the T6/T7 pair, only weak activity for T14/T15 and no detectable activity for T1/T2. Solid arrowhead indicates uncleaved DNA, arrowheads indicate cleavage fragments. Percentage of cleavage for each PCR product is shown at the bottom. IX molecular weight marker (MW) (**b**). Schematic drawing showing the position of T6 and T7 TALENs on a genomic DNA sequence fragment containing *COL7A1* exon 80 (blue) and the c.6527insC mutation site (red arrowhead). TALENs spacer sequence is shaded (**c**). RDEB, recessive dystrophic epidermolysis bullosa.

c.6527insC mutation in both alleles of the *COL7A1* gene.[23] Infection with a multiplicity of infection (MOI) of 1,000 vector genomes (vg)/cell resulted in strong and generalized GFP expression for all six vectors. Expression of TALEN proteins was analyzed by western blot (**Supplementary Figure S3a**) and immunofluorescence (**Supplementary Figure S3b**) with anti-HA tag antibodies, showing highly efficient expression at comparable levels for all TALEN proteins.

Indel generation was assessed for each TALEN pair by using the Cel I mutation detection assay in RDEB-E67A6 keratinocytes transduced with the three pairs of Ad-TALENs vectors at a MOI of 1,000 vg/cell for each construct. Efficient indel generation activity was found for the T6/T7 pair, while only weak activity was detected for T14/T15 and no detectable activity could be found for T1/T2 (**Figure 1b**). The T6/T7 pair was therefore chosen for subsequent gene-editing experiments (**Figure 1c**).

c.6527insC mutation correction by HDR

To address the correction of the c.6527insC mutation in exon 80 of *COL7A1* gene by homologous recombination, we constructed AAV vector-based targeting constructs. The vector genomes contained genomic DNA sequences homologous to the *COL7A1* target region, including a wild-type exon 80 (having a run of six cytosines versus seven cytosines in the c.6527insC allele) in the right donor arm, and neomycin resistance (neo)-expressing cassettes flanked by AAV-packaging signals (**Figure 2a**). We used two different strategies for

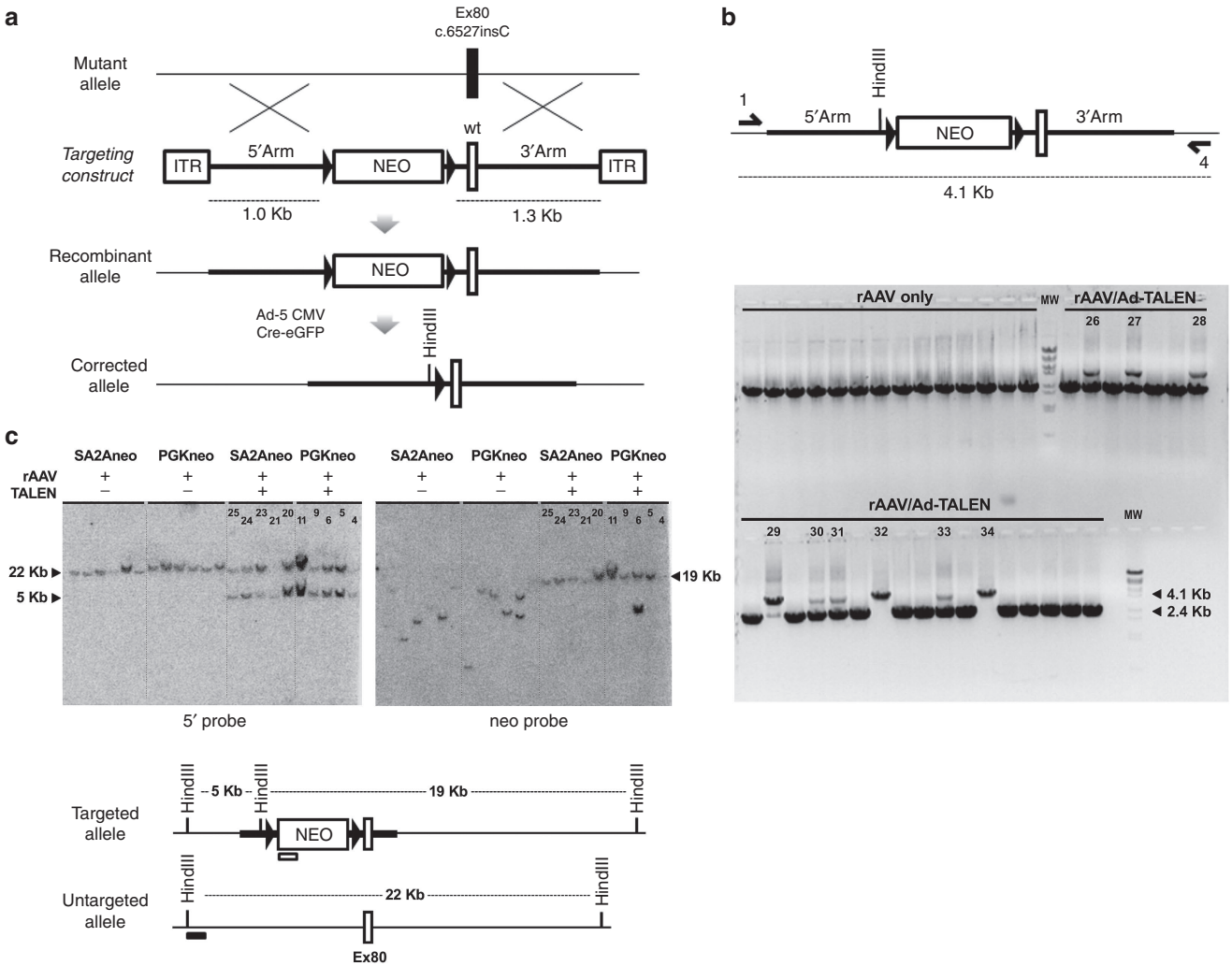

Figure 2 Gene targeting by homologous recombination strategy and genotyping of G418-resistant RDEB-E67A6 clones. AAV vector-delivered targeting constructs consist of left (1.0 kb) and right (1.3 kb) homology arms flanking a floxed neo selection cassette. The right arm contains the wild type sequence for exon 80. Ad-Cre transduction allows selection cassette removal to generate the corrected allele (**a**). Homologous recombination was analyzed by polymerase chain reaction using primer pair 1 and 4 external to the targeting construct that amplify a 2.4 kb in the untargeted allele and 4.1/4.2 kb (for SA-2A-neo /PGK-neo cassettes) in the targeted allele. AAV-SA-2A-neo-targeted clones are shown (**b**). Southern blot analysis. Genomic DNA from recombinant clones digested with Hind III, single cutter in the donor cassette, hybridized to a neo probe (right panel) showed a single band of the expected size (19 kb), indicating integration of a single copy of the targeting vector, in all but one of the recombinant clones. In nonrecombinant controls, a single band with different mobility for every clone suggests random integrations of a single copy of the targeting vector in all but one clone. A 5' external probe (left panel) showed bands corresponding to the wild-type (WT) (22 kb) and recombinant (5 kb) alleles in recombinant clones and just the WT band in control clones (**c**). RDEB, recessive dystrophic epidermolysis bullosa.

neo expression, a PGK-neo cassette for constitutive expression of neo driven by the PGK promoter, and a gene trap cassette (SA-2A-neo) in which a promoterless neo sequence is fused to a ribosomal-skipping 2A peptide sequence and is preceded by a splice acceptor (SA) sequence. This gene-trap strategy should result in neo expression only upon integration of the targeting construct into an intron of an expressed gene, and would therefore result in selection of fewer off-target integration events. Selection cassettes were flanked by loxP sites to allow for Cre recombinase-mediated cassette removal after recombination. These constructs were designated as AAV-PGK-neo and AAV-SA-2A-neo (**Supplementary Figure S4**). These vector genome constructs were packaged into the recently developed AAV2-derived engineered capsid variants with high keratinocyte transduction ability designated as AAV-Kera1 and AAV-Kera2. These are AAV2 peptide insertion variants displaying different 7-mer peptides at capsid position 587.[3] We tested AAV preparations with both capsid variants to assess the possibility that they might yield different rates of recombination.

We cotransduced RDEB-E67A6 cells with AAV vectors carrying the targeting constructs and adenoviral vectors expressing the T6/T7 TALEN pair. After treating the cells with G418, resistant clones were isolated and genotyped by PCR to detect homologous recombination using primers complementary to the genomic sequence flanking the targeting construct. Specifically, a DNA fragment of 2.4 kb will be amplified from the untargeted allele, while a fragment of 4.1/4.2 kb (depending on which neo cassette was used) indicates the targeted allele (primers 1 and 4) (**Figure 2b**). The presence of the recombinant allele was further verified by polymerase chain reaction (PCR) using primer pair 1 and 2 (specific for PGK-neo cassette) or 1 and 2B (specific for SA-2A-neo cassette), amplifying through the left arm of the construct, and primer pair 3 and 4, amplifying through the right arm (**Supplementary Figure S5**).

We performed five independent targeting experiments, using PGK-neo and SA-2A-neo AAV vectors. Results for each experiment are shown in **Table 1**. In total, targeting efficiency was 11% (19 of 166 clones) for the PGK-neo AAV vectors and 39% (15 of 38) for the SA-2A-neo vectors, as assessed by PCR genotyping of G-418 resistant clones. AAV vectors with capsid variants Kera1 and Kera2 yielded comparable recombination efficiencies (**Table 1**). Four biallelically targeted clones (4 of 15) were obtained when using the gene trap design, while only monoallelic correction was achieved with the PGK-neo constitutive expression design. Thus, the use of the more stringent, gene trap-based selection strategy resulted in a more than threefold increase in targeting efficiency over the constitutive neo expression strategy. The presence of the *COL7A1*/2A-neo fusion transcript in recombinant clones obtained by transduction with the SA-2A-neo-AAV vector was verified by reverse transcription-polymerase chain reaction (RT-PCR) using primers specific to exon 73 of *COL7A1* and to the neo resistance sequence. Sequencing of the RT-PCR product showed an in-frame fusion transcript comprised of *COL7A1*, 2A and neo sequences (**Supplementary Figure S6**). The correction of the c.6527insC mutation in exon 80 was verified by PCR amplification of the targeted allele using primer pair 3 (in the neo selection cassette) and 4 (in genomic sequence external to the targeting construct) (**Supplementary Figure S5**) and sequencing the PCR product. Thirty-two of 34 (94%) recombinant clones carried the corrected allele. To unambiguously analyze the mutation status of the targeted clones, we also performed PCR with primers Col7KI-B2-F (in the genomic sequence at the base of the right homology arm) and 4, amplifying a fragment spanning the exon 80 both in targeted and untargeted alleles (**Supplementary Figure S7**). For every targeting experiment, a control with AAV targeting vector only (without Ad-TALENs) was done. PCR genotyping showed that out of a total of 98 G418-resistant control clones (65 for the PGKneo vectors and 33 for the gene trap vectors) none of them had undergone homologous recombination (**Figure 2b**).

A subset of 10 recombinant and 12 control clones were further analyzed by Southern blot with a neo probe to assess whether multiple targeting vector integrations had taken place. An external 5' probe produced two bands corresponding to the wild type (22 kb) and targeted allele (5 kb) for all the recombinant clones analyzed, and a single wild type allele band for controls, confirming PCR genotyping results (**Figure 2c**, left panel). All but one of the targeted clones analyzed with the neo probe (9 of 10) showed a single band of the expected size (19 kb), indicating integration of a single copy of the targeting vector. One clone showed two

Table 1 Efficiency of c.6527insC correction by homologous recombination

Targeting construct	AAV capsid	TALENs T6-T7	Clones analyzed	Recombinant clones	Corrected clones	Homozygous insertion
PGK-neo	H1	(+)	53	6	5	0
		(–)	20	0	-	0
	H3	(+)	30	4	4	0
		(+)	83	9	9	0
		(–)	45	0	-	0
SA-2A-neo	H1	(+)	14	6	6	1
		(–)	19	0	-	0
	H3	(+)	24	9	8	3
		(–)	14	0	-	0

RDEB-E67A6 keratinocytes were transduced with rAAV gene targeting vectors containing a PGK-neo or SA-2A-neo selection cassettes. Vectors packaged with two different recombinant capsids, H1 or H3, were tested, with or without TALENs-expressing adenoviral vectors Ad-T6 and Ad-T7. Targeting efficiencies are shown for each experiment.

bands, indicating an additional integration. Eleven out of 12 non-recombinant control clones (transduced with targeting vector in the absence of TALENs) probed with neo showed a single band with different mobility for every clone, suggesting random integration in different genomic locations of a single copy of the targeting vector. Only one clone showed two bands, indicating two integrations (**Figure 2c**, right panel).

To avoid silencing of *COL7A1* gene expression by the neo selection cassette inserted in intron 79, as had been shown in a hypomorphic mouse model,[24] the cassette was removed by infection of recombinant clones with a Cre-expressing adenoviral vector (Ad-Cre). Twenty correctly targeted clones were infected and removal of the cassette was analyzed by PCR amplification using primers external to the targeting construct that detect different mobility bands for the selection cassette-containing allele (4 kb) and untargeted or corrected allele (2.4 kb). Disappearance of the 4 kb band indicated efficient cassette elimination in every clone. Hind III digestion of the PCR product allowed to discriminate the recombinant corrected vs the untargeted allele after Cre expression (**Supplementary Figure S8**).

RT-PCR expression analysis showed that *COL7A1* was expressed in 19 different recombinant clones that had the corrected sequence (6Cs) at the c.6527insC mutation site, before and after Cre expression. Keratinocytes from a healthy donor (WT HK) and from a c.6527insC heterozygous donor (HT HK), as well as RDEB-E67A6 parental cells, were used as controls. All corrected clones expressed *COL7A1* at levels higher than unedited RDEB-E67A6 parental cells (**Figure 3a**). An uncorrected control clone (clone 17), targeted by homologous recombination but maintaining the c.6527insC mutation, displayed *COL7A1* expression levels

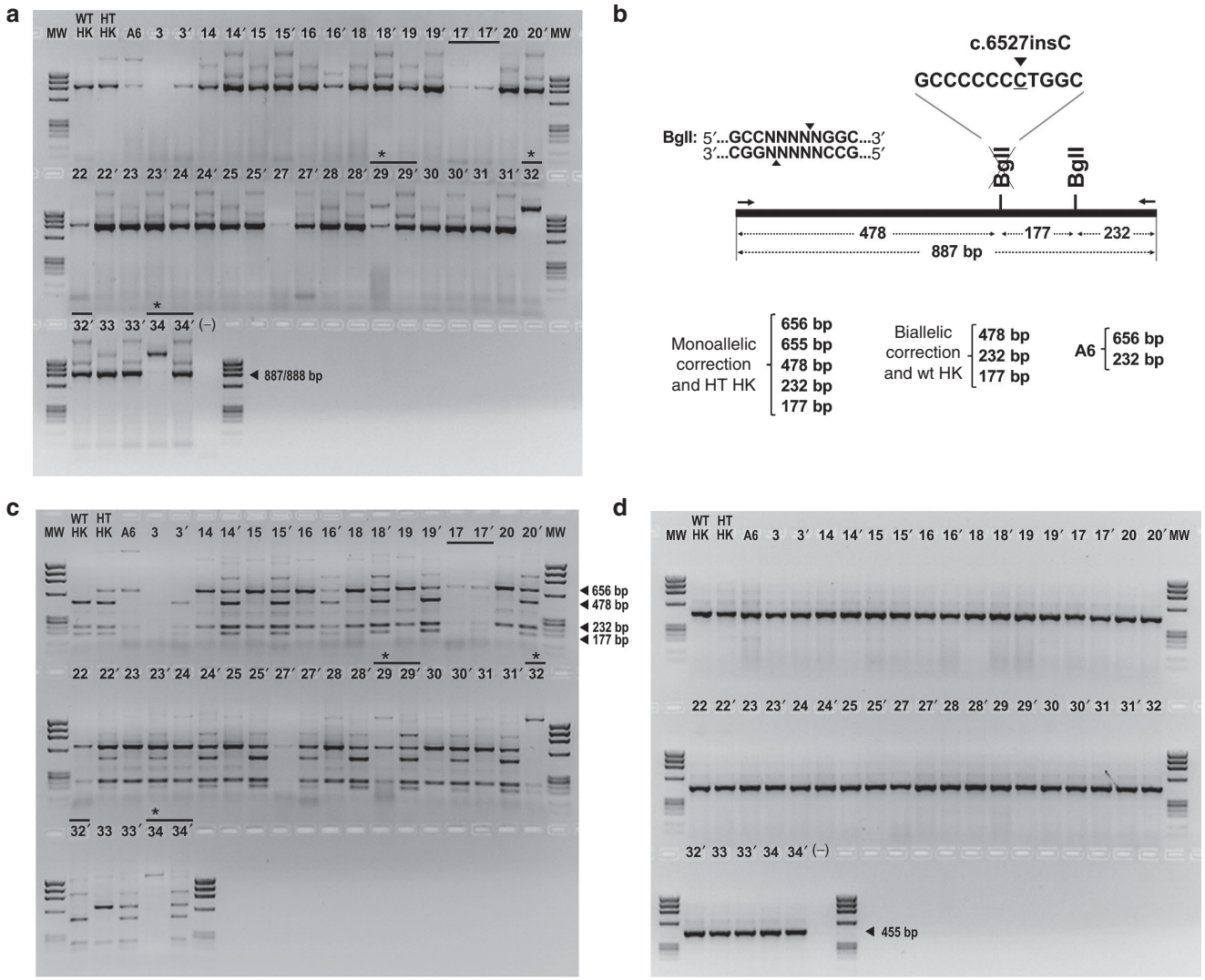

Figure 3 Reverse transcription-polymerase chain reaction (RT-PCR) analysis of *COL7A1* expression in homology-directed repair-corrected RDEB-E67A6 keratinocyte clones. Nineteen recombinant clones that had the corrected sequence (6Cs) at the c.6527insC mutation site were analyzed before and after transduction with Ad-Cre. Clone numbers are indicated above the lanes before (n) and after (n') transduction with Ad-Cre. Clone 17 (uncorrected control) is underlined. Biallelically targeted clones (29, 32, 34) are labeled with asterisks. Normal human keratinocytes from an unaffected donor (HK) and a heterozygous carrier of the c.6527insC allele (HT) were used as controls of *COL7A1* physiological expression levels. (**a**). Bgl I restriction patterns of RT-PCR bands corresponding to transcripts from WT (corrected) and c.6527insC alleles (b) allow to compare transcription from each allele (**c**). Glyceraldehyde-3-phosphate dehydrogenase expression was analyzed as a loading control (**d**). RDEB, recessive dystrophic epidermolysis bullosa.

comparable to RDEB-E67A6 cells, before or after Cre-mediated deletion of the selection cassette (**Figure 3a**, underlined lanes). Sequencing of the PCR products showed that, prior to cassette removal, only the uncorrected allele (containing the c.6527insC mutation, 7Cs at the mutation site) was expressed, while after cassette removal both the corrected (6Cs at the mutation site) and uncorrected alleles (7Cs) were expressed. In clones targeted at both *COL7A1* alleles with the gene trap vector (clones 29, 32, and 34) (**Figure 3a**, labeled with asterisks) a band of the expected size was found after Ad-Cre transduction (same as in clones having a single corrected allele). However, the RT-PCR band for these biallelically targeted clones prior to Cre expression was of higher molecular weight than expected for *COL7A1* cDNA amplification. Sequencing revealed that this band corresponded to a fusion between the neo resistance cassette sequence and *COL7A1* sequence. Clone 29 displayed both the *COL7A1* band and the fusion band before Ad-Cre, probably because it is a mixed population.

Transcription from the wt (6Cs) and c.6527insC (7Cs) alleles can be differentiated by digestion of RT-PCR products with BglI restriction enzyme, since the c.6527 mutation abolishes this site (**Figure 3b**). BglI restriction patterns confirmed that, prior to cassette removal, the RT-PCR band represented transcription from the c.6527 mutant allele only in all monoallelically targeted and corrected clones, and after Cre expression comparable amounts of corrected and uncorrected alleles were expressed in these clones (**Figure 3c**). The restriction pattern for the band from clone 17 (homologous recombinant without correction of the c.6527insC mutation) was consistent with expression of the uncorrected allele before and after Ad-Cre. To further quantify the expression of the corrected and uncorrected alleles before and after Cre-mediated deletion of the selection cassette, we performed T-A cloning in pGEM-T plasmid vector and sequencing of the RT-PCR products for clones PGKneo-HDR 18 (one allele targeted with the AAV-PGKneo vector), gene trap HDR-33 (one allele targeted with the AAV-SA-2A-neo gene trap vector) and gene trap HDR-32 (both alleles targeted with the AAV-SA-2A-neo gene trap vector). Approximately 10 colonies were analyzed for each clone. Before Cre expression, only transcripts corresponding to the untargeted allele were found for the PGKneo vector-targeted clone 18. Transcripts corresponding to the untargeted allele and fusion transcripts between neo and *COL7A1* were found in clone 33, monoallelically targeted with the gene trap vector and only fusion transcripts were found in clone 32, biallelically targeted with the gene trap vector. After excision of the neo cassette, monoallelically targeted clones 18 and 33 expressed both the untargeted (7Cs) and corrected (6Cs) alleles at similar ratios, while biallelically targeted clone 32 expressed only the corrected allele (**Supplementary Figure S9**).

COL7A1 frame restoration by NHEJ: isolation of keratinocyte clones with indels in COL7A1 exon 80

To obtain keratinocyte clones carrying insertions and deletions (indels) in the exon 80 of *COL7A1* gene produced by NHEJ-mediated repair of DSBs, we transduced RDEB-E67A6 cells with adenoviral vectors for expression of T6 and T7 TALENs (MOI 1,000 vg/cell) and obtained clones by limiting dilution. The presence of indels was analyzed for each clone by sequencing a PCR fragment spanning the nuclease target site. When overlapping chromatogram traces were found (**Supplementary Figure S10**) the sequence of both alleles was sorted out for each clone by using the Poly peak parser software.[25] Out of 113 clones analyzed, 30 (26.5 %) contained indels in this region. Although a total of nineteen different indel alleles were found, some of them occurred with higher frequency. Eleven clones had a 1 bp deletion, four clones presented an 18 bp deletion and two had a 19 bp deletion. These recurrent mutations were found as heterozygous as well as homozygous and compound heterozygous mutant alleles (**Figure 4**).

COL7A1 gene expression was analyzed by RT-PCR in one clone for each mutation or combination of mutations. Remarkably, high *COL7A1* transcript expression levels were found by RT-PCR in most indel-carrying clones, both in clones with frame-shifting indels causing frame restoration and in clones with indels not leading to frame recovery, compared to the RDEB-E67A6 parental cell line (**Figure 5a**). Human keratinocytes from an unaffected donor (WT) and a heterozygous carrier of the c.6527insC allele were used as controls of *COL7A1* physiological expression levels. RT-PCR products were sequenced to analyze transcript sequences, revealing overlapping chromatogram traces corresponding to edited and unedited alleles for the indel region in clones with one edited allele, and to differently edited alleles in clones with edition of both alleles. To sort out both alleles and quantify their relative expression, we performed T-A cloning and sequencing of the RT-PCR products. This analysis showed that both the re-framed and the unedited alleles were expressed at a similar ratio, regardless of whether frame restoration was caused by intra-exonic or exon skipping indel mutations (**Supplementary Figure S11**). Intra-exonic deletions Δ1, Δ7, Δ16, and Δ19, compatible with frame shifting-mediated recovery of *COL7A1* expression, resulted in transcripts encoding limited changes in aminoacid sequence in the exon 80 region. The most common indel (Δ1) resulted in four aminoacids changed within the exon 80-encoded sequence. Transcripts from clones lacking intron processing signal at the 5' (Δ20, Δ34, Δ43) or 3'(Δ40, Δ42) sequence of exon 80 skipped this exon altogether, resulting in a truncated collagen transcript lacking the sequence corresponding to the 12 aminoacids encoded by exon 80. Out of frame transcript sequences were found in RT-PCR products from clones with intra-exonic deletions that did not restore the reading frame (Δ6 ins3, Δ6, Δ8, Δ2 and Δ5, **Figure 4**). However, a clone with an 18 bp biallelic deletion within the exon 80 sequence (Δ18, **Figure 4**) yielded in-frame transcripts skipping exon 80 and exons 80–81 (**Figure 5b**). In total, 22 of the 113 clones analyzed (19.5%) presented indels that result in expression of transcripts encoding mutated or truncated C7.

C7 protein expression in NHEJ-corrected and HDR-corrected clones

We analyzed C7 expression by western blot in clones corrected by HDR, before and after Ad-Cre-mediated removal of the selection cassette, and NHEJ. For HDR-corrected clones we analyzed protein extracts from five clones: 4, 6, and 51, corrected with the PGK-neo vector, and 32 (bi-allelic correction) and 33, corrected with the SA-2Aneo gene trap vector. We found that,

Figure 4 Indels in RDEB-E67A6 keratinocytes transduced with Ad-T6/T7. Thirty out of 113 clones analyzed had indels in the target site for the transcription activator-like nucleases pair T6-T7. Nineteen different indel mutations were found. The number of clones having each indel and the presence of indels in one or both alleles of *COL7A1* are shown on the right. RDEB, recessive dystrophic epidermolysis bullosa.

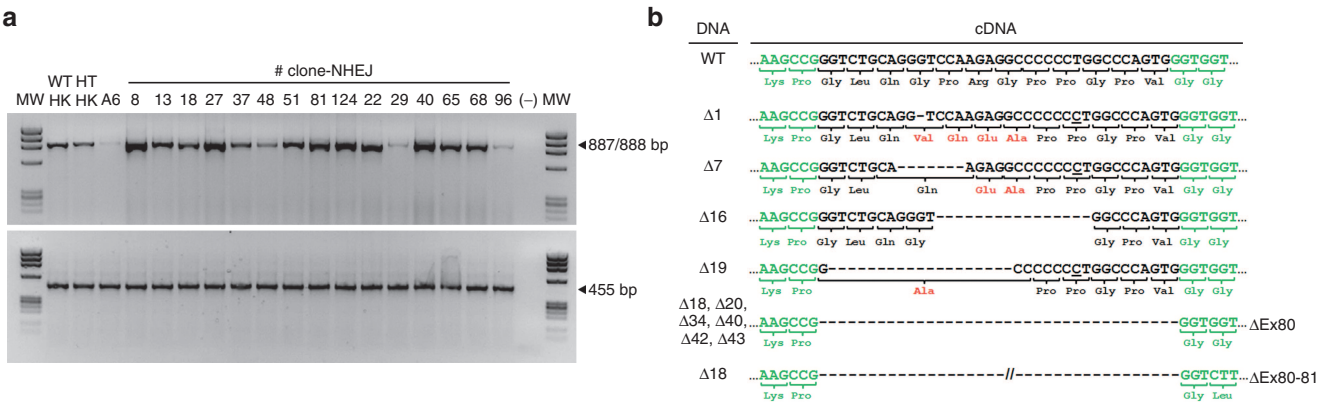

Figure 5 Reverse transcription-polymerase chain reaction (RT-PCR) analysis of *COL7A1* expression in NHEJ-corrected RDEB-E67A6 keratinocyte clones. *COL7A1* gene expression was analyzed by RT-PCR in one clone for each mutation or combination of mutations. High *COL7A1* transcript expression levels in relation to RDEB-E67A6 parental cell line were found, both in clones with frame-shifting indels causing frame restoration (clones 8, 13, 18, 27,37, 48, 51, 124, and 22) and in clones with indels not leading to frame recovery (29, 40, 65, 68, and 96), compared to the RDEB-E67A6 parental cell line (A6). Normal human keratinocytes from an unaffected donor (HK) and a heterozygous carrier of the c.6527insC allele (HT) were used as controls of *COL7A1* physiological expression levels. Glyceraldehyde-3-phosphate dehydrogenase (455 bp) was used as a loading control (**a**). RT-PCR bands were cloned in plasmids and sequenced. Sequence of the transcripts corresponding to each indel in the region flanking the nuclease target site, and the ensuing changes in C7 aminoacid sequence, are shown (**b**). RDEB, recessive dystrophic epidermolysis bullosa.

after Ad-Cre infection, C7 was expressed in every corrected clone analyzed at levels comparable to control human keratinocytes. No expression was detected in RDEB-E67A6 parental cells or in clone 17, homologous recombinant without correction of the c.6527insC mutation included as a negative control. Consistent with the RT-PCR results, showing lack of expression of the corrected allele before neo selection cassette removal, no C7 protein expression was found in corrected clones before Ad-Cre infection (**Figure 6a**, lower panels).

For NHEJ-mediated correction we analyzed by WB all clones carrying indel alleles, alone or in combination, that had been found to result in expression of re-framed *COL7A1* transcripts by frame shifting (Δ1, Δ7, Δ16 and Δ19) or exon skipping (Δ18, Δ20, Δ34, Δ43, Δ40, Δ42). We also analyzed clones with indels

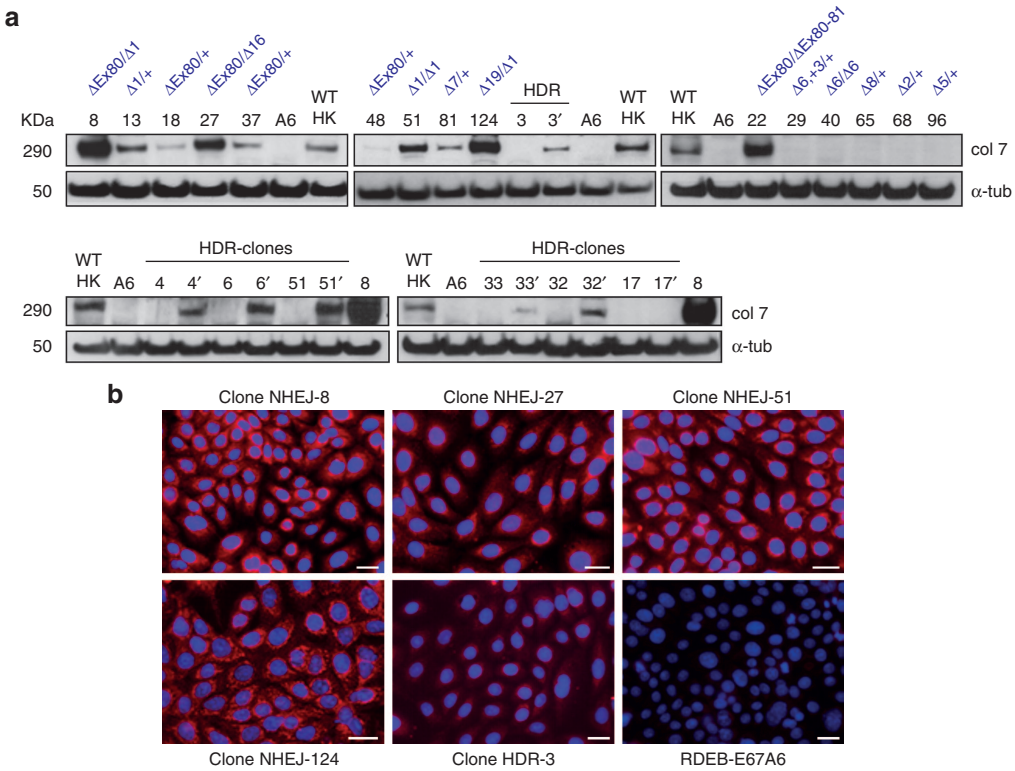

Figure 6 C7 protein expression in HDR and NHEJ-corrected RDEB-E67A6 keratinocyte clones. NHEJ (upper panels) and HDR (lower panels) corrected keratinocyte clone lysates were analyzed by Western blot to detect C7 expression. Clone numbers are shown above lanes. Indel genotypes for each NHEJ-corrected clone are shown above every clone number. For every HDR-corrected clone, cell lysates before (n) and after (n') neo cassette excision were analyzed (**a**). Immunofluorescence staining to detect C7 in cells from four NHEJ-corrected clones (NHEJ-8, 27, 51, and 124) and one HDR-corrected clone (HDR-3). RDEB-E67A6 parental cells are shown as negative control (**b**). Scale bar= 25 μm. HDR, homology-directed repair; NHEJ, nonhomologous end joining; RDEB, recessive dystrophic epidermolysis bullosa.

not leading to *COL7A1* reading frame restoration (Δ6 ins3, Δ6 homozygous, Δ8, Δ2 and Δ5). C7 protein expression was found in every clone expressing restored frame transcripts. Although the different indel-restored transcripts encode different truncated or mutated forms of C7, the estimated size of these proteins is very similar to full-length wild type C7 and a single band of apparently the same mobility was found for every clone (**Figure 6a**, upper panels). A remarkable overexpression of C7 protein was found in several clones, as compared to physiological expression levels in control human keratinocytes or HDR-corrected RDEB-E67A6 keratinocytes (Clone NHEJ-8 sample was included in the HDR-corrected clones gels for comparison of C7 amounts in HDR versus NHEJ correction). C7 overexpression was most pronounced in clones having both alleles restored, either in compound heterozygosity (Δ19/Δ1, ΔEx80/Δ1, ΔEx80/Δ16) or homozygosity (Δ1/Δ1, Δ18/Δ18) (**Figure 6a**).

Immunofluorescence was used to detect C7 in cells from four NHEJ-corrected clones and one HDR-corrected clone. Consistent with the western blot results, prominent cytoplasmic staining was detected in NHEJ-corrected cells as compared to HDR-corrected cells (**Figure 6b**).

In vivo analysis of C7 incorporation into the DEJ of gene-edited keratinocyte grafts

To determine whether production of C7 by *COL7A1* gene-edited cells results in C7 deposition in the dermo-epidermal junction (DEJ) of skin, corrected clones of RDEB-E67A6 keratinocytes were used to generate skin equivalents, as described previously.[23] NHEJ-corrected clones 8 (indels ΔEx80/Δ1) and 37 (ΔEx80) that had shown by WB analysis high and low C7 expression levels respectively, and HDR-corrected clone 3/Cre, with expression levels comparable to human keratinocytes control, were used. For engraftment follow up, the cells were transduced with a GFP-expressing retroviral vector.[23] Eight weeks after transplantation onto immunodeficient mice, GFP-positive regenerated skins were excised for analysis.

Immunofluorescence analysis to detect the human C7 protein in graft sections using a human-specific anti C7 monoclonal antibody[26] showed a very strong signal in the DEJ of skin regenerated from clone NHEJ-8 keratinocytes. A weaker signal, comparable to control skin generated from control human keratinocytes, was observed in grafts generated from clones NHEJ-37 and HDR-3/Cre (**Figure 7a**).

Discussion

Cutaneous *ex vivo* gene therapy based on the specific correction of disease-causing mutations in epidermal stem cell clones can now be pursued by using new gene-editing methods. Targeting the epidermal stem cell will allow the generation of persistent corrected skin grafts with therapeutic utility for the treatment of monogenic genodermatoses such

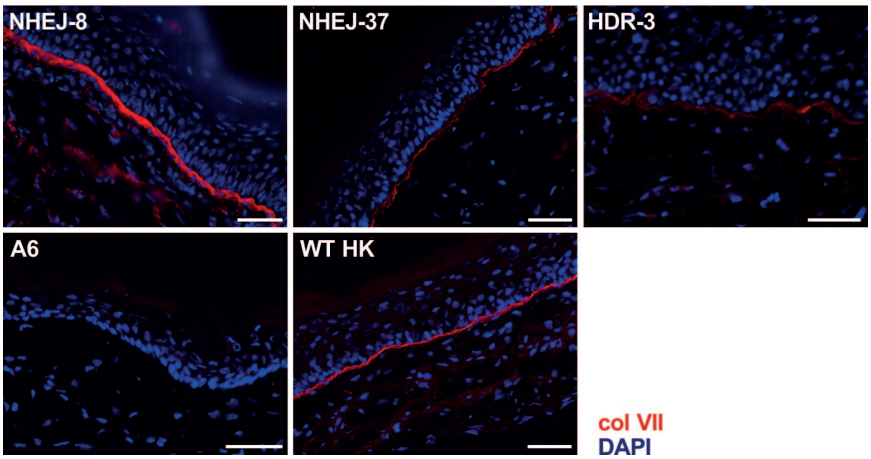

Figure 7 ***In vivo*** **analysis of C7 incorporation into the dermo-epidermal junction of gene-edited RDEB-E67A6 keratinocyte grafts**. Immunofluorescence staining for the detection of human C7 expression in sections of grafts from NHEJ-8, NHEJ-37, and HDR-3/Cre clones. Grafts from RDEB-E67A6 parental cells and wild-type human keratinocytes are shown as negative and positive controls respectively. Scale bar= 50 µm. HDR, homology-directed repair; NHEJ, nonhomologous end joining; RDEB, recessive dystrophic epidermolysis bullosa.

as RDEB. We have previously achieved long-term engraftment of retroviral vector transduced-single human epidermal holoclones,[27] and this approach has been used by others for the correction of RDEB patient-derived human keratinocytes by gene replacement.[28] However, the potential genotoxicity risk and the very large size and repetitive nature of *COL7A1* cDNA are important hurdles for efficient transduction and expression of C7 protein with retroviral vectors. Implementation of precise and efficient gene-editing protocols will therefore be necessary to advance towards clinically meaningful DEB gene therapy.

Targeted gene addition in human epithelial stem cells by nuclease-mediated homologous recombination has been previously reported by our group. We combined IDLV-mediated delivery of donor DNA with adenoviral vector-mediated expression of zinc finger nucleases to execute the targeted integration of a GFP expression vector into the AAVS1 "safe harbor" locus in a small fraction of primary keratinocyte cells that, upon grafting, persisted as small foci in skin regenerated in immunodeficient mice.[29] We then achieved long-term skin regeneration from a single AAVS1 gene-targeted human epidermal stem cell clone.[30] Several groups have also addressed the correction of DEB by derivation of human keratinocytes from patient-derived iPS cells. Osborn *et al.*[31] used TALEN-mediated targeted correction of patient fibroblasts and reprogrammed them to derive iPS cells that showed some potential as a source of keratinocytes in teratoma formation assays. Sebastiano et al performed AAV vector-mediated HDR correction of iPS cells derived from DEB patients and subsequently differentiated them to human keratinocytes.[32] In a similar fashion, iPS cells generated from naturally revertant keratinocytes from DEB patients were differentiated to keratinocytes.[33] However, the long term skin regeneration capacity and safety profile of human keratinocytes derived from iPS cells is still unknown. Achieving the correction of RDEB-causing mutations in human keratinocyte clones, with proven long-term skin regeneration ability and clinical relevance, would be highly preferable for translation to the clinical practice.

We have now developed gene-editing tools for the correction of a recurrent mutation accounting for 46% of DEB alleles in the Spanish DEB population. We have achieved a high percentage of gene correction by homologous recombination-mediated gene targeting in patient-derived immortalized keratinocytes by combining donor DNA delivery using AAV vectors with site-specific TALEN nucleases delivered with adenoviral vectors.

We explored the capability of two novel AAV2-derived vectors, Kera1 and Kera2, with engineered capsids specifically developed to efficiently transduce human keratinocytes[3] to carry our targeting constructs and we found that both of them drove comparable, high rates of homologous recombination.

Although homologous recombination mediated by AAV vectors without the aid of site-specific nucleases has been reported in human keratinocytes[10] we did not find any recombinants in the absence of nucleases (0 recombinant clones out of 98 rAAV-only controls analyzed). Importantly, in our experiments, genetic correction of the c.6527insC mutation was achieved in the majority of recombinant clones (32 of 34) and all corrected clones analyzed (a total of 19) expressed the *COL7A1* gene. These results show the potential of these new vectors for the development of gene targeting-based therapies of heritable skin diseases.

Randomly integrated AAV vectors proviruses are located at sites of spontaneous ds break formation[34] and this property of rAAVs has been used to identify nucleases off-target sites.[35] We have shown here that expression of TALE nucleases together with rAAV targeting constructs does not seem to result in increased off-target integrations, since most G418 resistant RDEB-E67A6 clones analyzed by Southern blotting with a neo probe had a single transgene integration, regardless of whether they had been transduced with AAV only or AAV plus TALENs.

In this study we have demonstrated that production of truncated or mutated forms of C7 in keratinocytes from c.6527 patients can be achieved efficiently and without antibiotic selection by generating indels in the *COL7A1* exon 80 sequence. After transduction with our adenoviral

vector-delivered TALENs, a high percentage (22 of 113, 19.5%) of RDEB-E67A6 keratinocyte clones carried indels compatible with frame restoration in at least one allele. All of the frame-restoring indel mutations resulted in *COL7A1* mRNA and C7 protein expression. Some of the indels were recurrent, with a single G deletion at the nuclease target site accounting for 30% of the mutations. Limited diversity of NHEJ-generated mutant alleles has been described before.[36] WB analysis showed remarkable C7 protein overexpression in several NHEJ-reframed clones, as compared to normal control keratinocytes. However, C7 protein expression in clones corrected by HDR was in the same range as in human control keratinocytes. These data suggest that NHEJ-mediated frame restoration, but not HDR-mediated correction, might result in upregulation of COL7A1 expression, as indicated by increased COL7A1 transcript levels in NHEJ-edited clones (**Figure 5a**). In addition, transcripts from NHEJ-edited alleles might be translated at a higher rate.

Although the relation between genotype and phenotype in RDEB is not always clear, a general rule can be drawn that the most severe forms of the disease are caused by PTC-causing mutations that result in the complete absence of C7, while mild or moderate forms of the disease are associated with missense or splice site mutations that result in the production of mutated or truncated C7 protein, even if the patient is a compound heterozygote carrying a PTC in the other allele.[15] A mini C7 lacking part of its central collagenous domain retained the functions and characteristics of a full-length C7 α chain[37] and antisense oligoribonucleotide-mediated skipping of the exon 70 of *COL7A1*, containing a recurrent PTC mutation, achieved C7 re-expression in patient cells and resulted in the formation of anchoring fibrils.[38] We have shown here that human skin regenerated from a keratinocyte clone (NHEJ-8) carrying indels in the *COL7A1* gene that result in the production of two C7 protein variants, one with a substitution of four aminoacids and other with a truncation of 12 aminoacids within the C7 collagenous domain, displays C7 in the dermo-epidermal junction. Therefore, NHEJ-mediated correction of the c.6527insC mutation, resulting in the expression of truncated or altered forms of C7, might have the potential to correct the skin fragility phenotype, or at least shift the pathological phenotype from severe to mild.

We have shown the feasibility of efficiently correcting a recurrent RDEB causing mutation in human keratinocytes by gene-editing approaches with the aid of TALEN nucleases and non-integrating viral vectors. These molecular tools will allow the *ex vivo* manipulation of epidermal stem cells from a large cohort of RDEB patients to generate transplantable human skin for therapeutic purposes. Although achieving HDR-mediated correction of RDEB keratinocyte holoclones remains a challenging task, we have shown that NHEJ-mediated gene edition is a highly efficient alternative for the development of clonal gene therapy protocols.

Materials and methods

TALEN construction. TALE nucleases were constructed using the Golden Gate Talen kit (Addgene, Cambridge, MA) as described in Cermak *et al.*[20] Destination vector pCAG-T7-TALEN (Addgene) was used to generate TALEN proteins with N and C terminal truncations (+63 design). A destination vector designed to express a C-terminal deletion of HA-tagged TALE proteins fused to the FokI domain was generated by ligating an adapter into the AatII/EcoRV sites of pTAL3. The adapter was prepared by annealing oligos 5′-tacccatacgacgtcccagactacgcgat-3′ and 5′-atcgctgagtctgggacgtcgtaggtgtaacgt-3′. A BglII/AflII fragment from the resulting plasmid was cloned into BamHI/AflII sites of a plasmid containing the PGK promoter and polyadenylation sequences, to yield the final PGK-driven destination vector (+30+HA design, **Supplementary Figure S1**).

CelI (Surveyor) analysis of TALEN pairs activity. PCR fragments spanning the TALENs' target sites were generated with primers F1/R, F2/R and 80F/82R. F1: 5′-gtgagtggtggctgaagcac-3′; F2: 5′-tctgtgtgtggttgtatgtgga-3′; R: 5′-accccaccaaggaaactga-3′; 80F: 5′-caagtgaggcccagattgag-3′; 82R: 5′-ggcatggacacagcttgaag-3′. PCR products were subjected to melting and reannealing to form heteroduplexes and homoduplexes, digested with Surveyor nuclease (Transgenomic, Omaha, NE) at 42 °C for 1 hour, resolved in 1.5% agarose gels, and visualized by ethidium bromide staining. Molecular weight marker was IX (Roche Diagnostics, Mannheim, Germany). Percentage of cleaved DNA was calculated by using the following formula: % cleaved= $((b + c)/(a + b + c)) \times 100$, where a is the integrated intensity of the undigested PCR product and b and c are the integrated intensities of each cleavage product.

Adenoviral vectors. Recombinant serotype 5 adenoviral vectors were produced by site-specific recombination between Ad5 genome-containing pBHGloxΔE1, 3Cre plasmid, and shuttle plasmids pDC315-TALEN-iresGFP after their cotransfection into 293 cells.[39] Shuttle vectors were constructed by cloning a NheI /EcoRI fragment, comprising the PGK promoter and TALEN coding sequences, from the TALEN expression vectors with the +30+HA design into XbaI/EcoRI sites of pDC315iGFP plasmid. After calcium phosphate-mediated transfection of both plasmids, 293 cells were embedded in agarose medium (0.5% agarose in Dulbecco's modified Eagle medium (DMEM) 10% fetal bovine serum) and observed for plaque appearance. When plaques reached 2–3 mm in diameter, they were isolated by punching out agar plugs with a Pasteur pipette and eluted in 1 ml of 1× phosphate-buffered saline (PBS) $^{+/+}$ with 10% glycerol for −80 °C storage. For virus expansion, a subconfluent 60-mm dish of 293 cells was adsorbed with 400 microliters of plaque picks for 1 hour. Following adsorption, 5 ml of MEM+5% horse serum was added and incubated until complete cytopathic effect was observed. Cells were then harvested by gentle pipetting and genomic DNA was extracted from the cellular pellet and digested with HindIII for restriction digestion analysis of the viral genome. Supernatants were frozen at −80 °C after adding glycerol to 10%, and used for later preparation of large scale stocks once the DNA structure of the recombinant virus was verified. 150-mm dishes of subconfluent 293 cells were infected with the supernatants. Once cytopathic effect was observed, cells were harvested by gentle pipetting; cellular pellets were resuspended in DMEM and lysed by subjecting them to three cycles of freezing and thawing. The lysate was added on top

of a cesium chloride gradient in SW40 tubes and centrifuged at 35,000 rpm at 10 °C for 3 hours. The band containing the viruses was collected by piercing the side of the tube with a needle and syringe. The viruses were then purified in PD10 columns (Amersham, Burlington, MA) and concentration of viral particles based on DNA content was determined as follows: $OD_{260} \times$ dilution factor $\times 1.1 \times 10^{12}$ = viral particles/ml.

Adeno-associated virus vector construction. DNA fragments containing *COL7A1* homology arms were PCR amplified using BAC RP11-148G20 as a template with primers Col7KI-B1-F: 5′-gcggccgcggactggatgggaaac-3′ and Col7KI-B1-R: 5′-gctagcaagcttagcacacacgggcctgcac-3- for left arm, and Col7KI-B2-F: 5-ggatcctctatgagaggcagtccatgg-3′ and AAV-Br2-R: 5′-actagtgaggaactcagtgcctc-3′ for right arm, and cloned in the pGEM-T vector (Promega, Madison, WI) for sequence verification. To facilitate genotyping, a KpnI site was introduced into the right arm by using Quikchange site directed mutagenesis kit (Agilent, Santa Clara, CA) with primers F: 5′-gtgccctgccccaggtaccagtactgccttcag-3′ and complementary reverse primer. A BamHI/EcoRI fragment containing the right arm, and a NotI/HpaI fragment for the left arm, were sequentially cloned in pLoxP plasmid, containing the floxed PGK-neo cassette. A SpeI/SpeI fragment from the resulting plasmid was cloned in the XbaI sites of pSub201 plasmid,[40] containing the AAV ITRs, to produce the AAV-PGKneo construct. The gene trap cassette (SA-2A-neo) was constructed by overlapping PCR of fragments containing a splice acceptor sequence from the first intron of mouse N4bp1 gene, a 2A peptide sequence and a neo resistance sequence. Overlapping fragments were constructed with the following primers: a splice acceptor-containing fragment (SA) of 248 bp was amplified with SA-F:5′-gggcccactctgtaaattatacaaag-3′ and SA-R:5′-aggaccgggggttttcttcc-3′, using a previous construct containing a SA-2A fusion as a template; 2A peptide-ATG less neo fragment was amplified with 2A-F: 5′-gaagaaaaccccggtcctttgattgaac-3′ and 2A-R: 5′-ctcaccataggaccgggggttttcttc-3′, using a previous construct containing 2A-neo as a template. The PCR-mediated fusion of both fragments was performed using primers SA-F and 2A-R. The PCR product was cloned in pGEM-T and verified by sequencing. An XhoI/RsrII fragment containing the gene trap neo cassette was cloned into XhoI/RsrII sites of AAV-PGKneo to generate the AAV-SA-2A-neo construct.

Production of AAV targeting vectors. HEK293 cells were cotransfected with a total of 37.5 micrograms of pRC-Kera1 or pRC-Kera2, AAV-PGKneo or AAV-SA-2A-neo construct, and pXX6.[41] Cells were harvested 48 hours post-transduction, lysed and purified by iodixanol density step gradient centrifugation.[42] Particle titer (vector genomes per microliter) was determined by qPCR (LightCycler System, Roche Diagnostics, Mannheim, Germany) from vector genomes isolated by DNeasy Tissue kit (Qiagen, Hilden, Germany) from vector preparations using neo-specific primers.

Genotyping for detection of homologous recombinants. Genomic DNA was isolated by isopropanol precipitation of keratinocytes lysates (lysis buffer was Tris pH8 100 mmol/l, ethylene diamine tetra acetic acid 5 mmol/l, sodium dodecyl sulfate 0.2%, NaCl 200 mmol/l, 1 mg/ml proteinase K (Roche Diagnostics, Mannheim, Germany)) and resuspended in Tris-EDTA buffer. Approximately 100 ng of genomic DNA were used for PCR genotyping. Targeted integration was assessed by PCR (Expand Long Template PCR system, Roche Diagnostics) using primers specific for human genomic sequence flanking the homology region present in the targeting construct 1:5′-cctcctggactccctggaac-3′and 4: 5′-tctcctttaggtccgacaggg-3′. PCR program was: 94 °C for 1 minute, 10 cycles of 94 °C for 10 seconds, 55 °C for 30 seconds, 68 °C for 1 minute and 50 seconds, followed by 25 cycles of 94 °C for 10 seconds, 55 °C for 30 seconds, 68 °C for 1 minute and 50 seconds with 15 seconds/cycle increments, then 68 °C for 7 minutes. Targeted integration 5′ junction was assessed with primers 1 and 2A: 5′-tacccggtagaattgacctgc-3′ for the PGKneo-based targeting construct and 1 and 2B: 5′-tgtgtgtgaggctatttgtctgtt-3′ for the SA-2A gene trap-based targeting construct, and 3′ junction was assessed with primers 3: 5′-aatgtgtcagtttcatagcctg-3 and 4. PCR products were resolved on 1% agarose gels and visualized by ethidium bromide staining.

For Southern blot analysis, approximately 10 µg of genomic DNA were digested with Hind III, separated on a 0.8% agarose gel, transferred to a nylon membrane (Amersham XL, Amersham, Burlington, MA) by alkaline transfer and probed with P^{32}-labeled probes. Probes were generated by PCR, cloned in pGEM-T (Promega) easy and verified by sequencing. The 5′ probe was made with primers 5′probeF: 5′-cccagcaggttctaccttg-3′ and 5′probeR: 5′-aagcttttctcctgctgggcctc-3′ using human genomic DNA as a template and the neo probe with primers 2AF: 5′-gaagaaaaccccggtcctttgattgaac-3′ and Neo R: 5′-gcacttcgcccaatagca-3′ using the AAV-SA-2A-neo targeting construct as a template.

RT-PCR analysis. Total RNA was extracted from keratinocytes with miRNeasy Mini Kit (Qiagen), and complementary DNA (cDNA) was synthesized using the SuperScript III First-Strand Synthesis System (Invitrogen, Carlsbad, CA). The following primers: F: 5′-ctggtctggcccttggggagag-3′ and R: 5′-cttctctccctttgctccaggg-3′, were used to amplify a 887-bp fragment spanning exons 73–88 of *COL7A1*. The PCR conditions were: step 1: 94 °C 3 minutes; step 2: 94 °C 45 seconds, 68 °C 30 seconds, 72 °C 1 minute, 10 cycles with 1 °C decrease of annealing temperature in each cycle; step 3: 94 °C 45 seconds 63 °C 30 seconds 72 °C 1 minute, 25 cycles; step 4: 72 °C 10 minutes. The human glyceraldehyde-3-phosphate dehydrogenase gene was analyzed as a loading control with glyceraldehyde-3-phosphate dehydrogenase-specific primers: F: 5′-accacagtccatgccatcac-3′ and R: 5′-tccaccaccctgttgctgt-3′.

Western blot analysis. Keratinocytes were grown to confluence and lysed in protein extraction buffer (50 mmol/l Tris-HCl, pH 7.5, 100 mmol/l NaCl, 1% Nonidet P-40, 4 mmol/l ethylene diamine tetra acetic acid) containing proteinase inhibitors cocktail (Complete Mini, ethylene diamine tetra acetic acid-free; Roche Diagnostics). Lysates were incubated for 30 minutes on ice and centrifuged at 15,000×g for 30 minutes at 4 °C. Supernatants were collected and protein concentrations were measured using the Bradford assay (BioRad, Hercules,

CA). For each sample, 50 µg of total protein was resolved on NuPAGE Novex 3–8% Tris-Acetate gel electrophoresis (Invitrogen) and electrotransferred to nitrocellulose membranes (Invitrogen). For type VII collagen analysis, blots were probed with anti-type VII collagen polyclonal antibody (Calbiochem, San Diego, CA) diluted at 1:50. Antibody against α-tubulin (Sigma Aldrich, St Louis, MO) was used as a loading control. For detection of HA-tagged TALENs transfected in 293T cells, a 1/1000 dilution of anti HA tag 16B12 mouse monoclonal antibody (Biolegend, San Diego, CA) was used. Visualization was performed by incubating with horse radish peroxidase-conjugated anti-rabbit antibody (Amersham) and West Pico Chemiluminescent Substrate (Pierce, Rockford, IL).

Sequencing. Plasmids and PCR products were sequenced using Big Dye Terminator V.1.1 Cycle Sequencing kit and examined on a 3730 DNA Analyser (Life Technologies, Carlsbad, CA). Sequences were analyzed using Sequencher (Gene Codes, Ann Harbor, MI) and Poly Peak Parser software packages.[25]

Cell culture, transfections, transduction with viral vectors, and cloning. RDEB-E67A6 keratinocytes were cultured in Kci medium (2:1 DMEM) (GIBCO-BRL, Barcelona, Spain) /HAM'S F12 (GIBCO-BRL) containing 10% Hyclone fetal bovine calf serum (GE Healthcare, Logan, UT), 0.1 nmol/l choleric toxin (Sigma Aldrich) 2 nmol/l T3 (Sigma Aldrich), 5 microg/ml insulin (Sigma Aldrich), 0.2 µg/ml hydrocortisone(Sigma Aldrich), 12 microg/ml adenine(Sigma Aldrich), 0.2% Primocin (Sigma Aldrich) as previously described.[23] For viral vector transduction, cells were trypsinized and infected in suspension with AAV targeting vector constructs (MOI: 30,000) and TALEN-expressing adenoviral vectors (MOI: 1,000) in the presence of polybrene (8 µg/ml) (Sigma Aldrich). 1×10^6 transduced cells were plated in a 35-mm dish without feeder layer and cultured for 48–72 hours. Cells were then trypsinized and plated in ten 100-mm plates with 2×10^6 lethally irradiated 3T3 feeder cells per plate. When cell clones were visible, G418 (100 µg/ml) (GIBCO-BRL) was added to the medium. After 10 days of antibiotic selection, resistant clones were collected using cloning cylinders (Sigma Aldrich) and expanded for cryopreservation and genomic DNA extraction. HEK293T cells were cultured in DMEM 10% FCS and transfected with 5 µg of TALEN-expressing constructs using the calcium phosphate coprecipitation method.

Immunofluorescence staining. For immunofluorescence detection of human C7 in gene-edited RDEB-E67A6 keratinocytes or in skin grafts, cells grown on glass cover slips or 7 µm frozen sections of grafted skin tissue were fixed in methanol/acetone (1:1) for 10 minutes at –20 °C. After washing three times in PBS and once in PBS with 3% bovine serum albumin (Sigma Aldrich) for 30 minutes, cells or cryosections were incubated with LH7.2 monoclonal antibody (Sigma Aldrich) for 1 hour in a humidified chamber at room temperature. Secondary antibody (AlexaFluor594, Invitrogen) was used at 1/1,000 dilution. After the final washing step in PBS, preparations were mounted using Mowiol (Hoechst, Somerville, NJ) mounting medium and 46-diamidino-2-phenyl indole 20 µgml^{-1} (Sigma Aldrich) for nuclei visualization. For detection of TALENs transfected in 293T cells, cells grown on glass cover slips and fixed with 4% paraformaldehyde for 10 minutes were incubated with a 1/1,000 dilution of anti HA tag mouse monoclonal antibody 16B12 (Biolegend) following the same protocol as described above.

Generation of skin equivalents and grafting onto immunodeficient mice. RDEB-E67A6 keratinocytes (2.5×10^5 cells) were seeded on RDEB fibroblasts-containing fibrin dermal equivalents prepared as previously described.[43] Bioengineered skin equivalents were grafted onto the back of immunodeficient nu/nu mice ($n = 3$) according to Del Rio *et al.*[44] including the step of retrovirus-mediated enhanced green fluorescent protein cell labelling. Engraftment was monitored by enhanced green fluorescent protein fluorescence visualization. Eight weeks after grafting, mice were sacrificed and grafts harvested for skin immunofluorescence analysis. Animal studies were approved by our institutional animal care and use committee according to all legal regulations.

Supplementary material

Figure S1. TALEN pairs targeting the *COL7A1* exon 80 region.
Figure S2. Expression of TALEN proteins and functionality of TALEN pairs transfected in 293T cells.
Figure S3. Expression of TALEN proteins in RDEB-E67A6 keratinocytes transduced with Ad-TALENs vectors.
Figure S4. AAV targeting vectors.
Figure S5. PCR genotyping by amplification of selection cassette-genome junctions.
Figure S6. *COL7A1*/2A-neo fusion transcript in recombinant clones obtained by transduction with the SA-2A-neo-AAV vector.
Figure S7. Genotyping c.6527insC mutation correction.
Figure S8. Neo selection cassette excision.
Figure S9. *COL7A1* alleles expression in HDR-targeted RDEB-E67A6 clones analyzed by T-A cloning and sequencing.
Figure S10. Sequence analysis of the nuclease target region in clones of RDEB-E67A6 keratinocytes transduced with T6/T7 TALENs.
Figure S11. *COL7A1* alleles expression in NHEJ-mutated RDEB-E67A6 clones analyzed by T-A cloning and sequencing.

Acknowledgments This work was supported in part by grants SAF2014-54885-R from MINECO to MDR; S2010/ BMD-2420 and S2010/ BMD-2359 from Dirección General de Investigación de la Comunidad de Madrid to MDR and FL respectively; PI14/00931 from Instituto de Salud Carlos III to FL and by funding from Center for Molecular Medicine Cologne (CMMC) of the University of Cologne to H.B. There are no conflicts of interest.

1. Cox, DB, Platt, RJ and Zhang, F (2015). Therapeutic genome editing: prospects and challenges. *Nat Med* **21**: 121–131.
2. Khan, IF, Hirata, RK and Russell, DW (2011). AAV-mediated gene targeting methods for human cells. *Nat Protoc* **6**: 482–501.
3. Sallach, J, Di Pasquale, G, Larcher, F, Niehoff, N, Rübsam, M, Huber, A *et al.* (2014). Tropism-modified AAV vectors overcome barriers to successful cutaneous therapy. *Mol Ther* **22**: 929–939.

4. Rio, P, Baños, R, Lombardo, A, Quintana-Bustamante, O, Alvarez, L, Garate, Z *et al.* (2014). Targeted gene therapy and cell reprogramming in Fanconi anemia. *EMBO Mol Med* **6**: 835–848.
5. Biffi, A (2015). Clinical translation of TALENS: Treating SCID-X1 by gene editing in iPSCs. *Cell Stem Cell* **16**: 348–349.
6. Tebas, P, Stein, D, Tang, WW, Frank, I, Wang, SQ, Lee, G *et al.* (2014). Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N Engl J Med* **370**: 901–910.
7. Mingozzi, F and High, KA (2011). Therapeutic *in vivo* gene transfer for genetic disease using AAV: progress and challenges. *Nat Rev Genet* **12**: 341–355.
8. Chamberlain, JR, Deyle, DR, Schwarze, U, Wang, P, Hirata, RK, Li, Y *et al.* (2008). Gene targeting of mutant COL1A2 alleles in mesenchymal stem cells from individuals with osteogenesis imperfecta. *Mol Ther* **16**: 187–193.
9. Petek, LM, Fleckman, P and Miller, DG (2010). Efficient KRT14 targeting and functional characterization of transplanted human keratinocytes for the treatment of epidermolysis bullosa simplex. *Mol Ther* **18**: 1624–1632.
10. Melo, SP, Lisowski, L, Bashkirova, E, Zhen, HH, Chu, K, Keene, DR *et al.* (2014). Somatic correction of junctional epidermolysis bullosa by a highly recombinogenic AAV variant. *Mol Ther* **22**: 725–733.
11. Händel, EM, Gellhaus, K, Khan, K, Bednarski, C, Cornu, TI, Müller-Lerch, F *et al.* (2012). Versatile and efficient genome editing in human cells by combining zinc-finger nucleases with adeno-associated viral vectors. *Hum Gene Ther* **23**: 321–329.
12. Miller, DG, Petek, LM and Russell, DW (2003). Human gene targeting by adeno-associated virus vectors is enhanced by DNA double-strand breaks. *Mol Cell Biol* **23**: 3550–3557.
13. Porteus, MH, Cathomen, T, Weitzman, MD and Baltimore, D (2003). Efficient gene targeting mediated by adeno-associated virus and DNA double-strand breaks. *Mol Cell Biol* **23**: 3558–3565.
14. Tolar, J and Wagner, JE (2015). A biologic Velcro patch. *N Engl J Med* **372**: 382–384.
15. Escámez, MJ, García, M, Cuadrado-Corrales, N, Llames, SG, Charlesworth, A, De Luca, N *et al.* (2010). The first COL7A1 mutation survey in a large Spanish dystrophic epidermolysis bullosa cohort: c.6527insC disclosed as an unusually recurrent mutation. *Br J Dermatol* **163**: 155–161.
16. Ousterout, DG, Perez-Pinera, P, Thakore, PI, Kabadi, AM, Brown, MT, Qin, X *et al.* (2013). Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol Ther* **21**: 1718–1726.
17. Ousterout, DG, Kabadi, AM, Thakore, PI, Perez-Pinera, P, Brown, MT, Majoros, WH *et al.* (2015). Correction of dystrophin expression in cells from Duchenne muscular dystrophy patients through genomic excision of exon 51 by zinc finger nucleases. *Mol Ther* **23**: 523–532.
18. Christiano, AM, Hoffman, GG, Chung-Honet, LC, Lee, S, Cheng, W, Uitto, J *et al.* (1994). Structural organization of the human type VII collagen gene (COL7A1), composed of more exons than any previously characterized gene. *Genomics* **21**: 169–179.
19. Doyle, EL, Booher, NJ, Standage, DS, Voytas, DF, Brendel, VP, Vandyk, JK *et al.* (2012). TAL effector-nucleotide targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucleic Acids Res* **40**(Web Server issue): W117–W122.
20. Cermak, T, Doyle, EL, Christian, M, Wang, L, Zhang, Y, Schmidt, C *et al.* (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* **39**: e82.
21. Miller, JC, Tan, S, Qiao, G, Barlow, KA, Wang, J, Xia, DF *et al.* (2011). A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* **29**: 143–148.
22. Qiu, P, Shandilya, H, D'Alessio, JM, O'Connor, K, Durocher, J and Gerard, GF (2004). Mutation detection using Surveyor nuclease. *Biotechniques* **36**: 702–707.
23. Chamorro, C, Almarza, D, Duarte, B, Llames, SG, Murillas, R, García, M *et al.* (2013). Keratinocyte cell lines derived from severe generalized recessive epidermolysis bullosa patients carrying a highly recurrent COL7A1 homozygous mutation: models to assess cell and gene therapies *in vitro* and in vivo. *Exp Dermatol* **22**: 601–603.
24. Fritsch, A, Loeckermann, S, Kern, JS, Braun, A, Bosl, MR, Bley, TA, *et al.* (2008). A hypomorphic mouse model of dystrophic epidermolysis bullosa reveals mechanisms of disease and response to fibroblast therapy. *J Clin Invest* **118**: 1669–1679.
25. Hill, JT, Demarest, BL, Bisgrove, BW, Su, YC, Smith, M and Yost, HJ (2014). Poly peak parser: Method and software for identification of unknown indels using sanger sequencing of polymerase chain reaction products. *Dev Dyn* **243**: 1632–1636.
26. Leigh, IM, Eady, RA, Heagerty, AH, Purkis, PE, Whitehead, PA and Burgeson, RE (1988). Type VII collagen is a normal component of epidermal basement membrane, which shows altered expression in recessive dystrophic epidermolysis bullosa. *J Invest Dermatol* **90**: 639–642.
27. Larcher, F, Dellambra, E, Rico, L, Bondanza, S, Murillas, R, Cattoglio, C *et al.* (2007). Long-term engraftment of single genetically modified human epidermal holoclones enables safety pre-assessment of cutaneous gene therapy. *Mol Ther* **15**: 1670–1676.
28. Droz-Georget Lathion, S, Rochat, A, Knott, G, Recchia, A, Martinet, D, Benmohammed, S *et al.* (2015). A single epidermal stem cell strategy for safe ex vivo gene therapy. *EMBO Mol Med* **7**: 380–393.
29. Coluccio, A, Miselli, F, Lombardo, A, Marconi, A, Malagoli Tagliazucchi, G, Gonçalves, MA *et al.* (2013). Targeted gene addition in human epithelial stem cells by zinc-finger nuclease-mediated homologous recombination. *Mol Ther* **21**: 1695–1704.
30. Duarte, B, Miselli, F, Murillas, R, Espinosa-Hevia, L, Cigudosa, JC, Recchia, A *et al.* (2014). Long-term skin regeneration from a gene-targeted human epidermal stem cell clone. *Mol Ther* **22**: 1878–1880.
31. Osborn, MJ, Starker, CG, McElroy, AN, Webber, BR, Riddle, MJ, Xia, L *et al.* (2013). TALEN-based gene correction for epidermolysis bullosa. *Mol Ther* **21**: 1151–1159.
32. Sebastiano, V, Zhen, HH, Haddad, B, Derafshi, BH, Bashkirova, E, Melo, SP *et al.* (2014). Human COL7A1-corrected induced pluripotent stem cells for the treatment of recessive dystrophic epidermolysis bullosa. *Sci Transl Med* **6**: 264ra163.
33. Umegaki-Arao, N, Pasmooij, AM, Itoh, M, Cerise, JE, Guo, Z, Levy, B *et al.* (2014). Induced pluripotent stem cells from human revertant keratinocytes for the treatment of epidermolysis bullosa. *Sci Transl Med* **6**: 264ra164.
34. Miller, DG, Petek, LM and Russell, DW (2004). Adeno-associated virus vectors integrate at chromosome breakage sites. *Nat Genet* **36**: 767–773.
35. Petek, LM, Russell, DW and Miller, DG (2010). Frequent endonuclease cleavage at off-target locations in vivo. *Mol Ther* **18**: 983–986.
36. Gagnon, JA, Valen, E, Thyme, SB, Huang, P, Akhmetova, L, Ahkmetova, L *et al.* (2014). Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. *PLoS One* **9**: e98186.
37. Chen, M, O'Toole, EA, Muellenhoff, M, Medina, E, Kasahara, N and Woodley, DT (2000). Development and characterization of a recombinant truncated type VII collagen "minigene". Implication for gene therapy of dystrophic epidermolysis bullosa. *J Biol Chem* **275**: 24429–24435.
38. Goto, M, Sawamura, D, Nishie, W, Sakai, K, McMillan, JR, Akiyama, M *et al.* (2006). Targeted skipping of a single exon harboring a premature termination codon mutation: implications and potential for gene correction therapy for selective dystrophic epidermolysis bullosa patients. *J Invest Dermatol* **126**: 2614–2620.
39. Ng, P and Graham, FL (2002). Construction of first-generation adenoviral vectors. *Methods Mol Med* **69**: 389–414.
40. Samulski, RJ, Chang, LS and Shenk, T (1987). A recombinant plasmid from which an infectious adeno-associated virus genome can be excised *in vitro* and its use to study viral replication. *J Virol* **61**: 3096–3101.
41. Xiao, X, Li, J and Samulski, RJ (1998). Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J Virol* **72**: 2224–2232.
42. Hacker, UT, Wingenfeld, L, Kofler, DM, Schuhmann, NK, Lutz, S, Herold, T *et al.* (2005). Adeno-associated virus serotypes 1 to 5 mediated tumor cell directed gene transfer and improvement of transduction efficiency. *J Gene Med* **7**: 1429–1438.
43. Llames, SG, Del Rio, M, Larcher, F, García, E, García, M, Escamez, MJ *et al.* (2004). Human plasma as a dermal scaffold for the generation of a completely autologous bioengineered skin. *Transplantation* **77**: 350–355.
44. Del Rio, M, Larcher, F, Serrano, F, Meana, A, Muñoz, M, Garcia, M *et al.* (2002). A preclinical model for the analysis of genetically modified human skin in vivo. *Hum Gene Ther* **13**: 959–968.

Supplementary Information accompanies this paper on the Molecular Therapy–Nucleic Acids website (http://www.nature.com/mtna)