AWARD NUMBER: **W81XWH-18-1-0706**

TITLE: **Translating a Stem Cell-Based Therapy for Epidermolysis Bullosa into the Clinic**

PRINCIPAL INVESTIGATOR: **Dennis R. Roop**

CONTRACTING ORGANIZATION: **University of Colorado Denver, Aurora, CO**

REPORT DATE: **October 2021**

TYPE OF REPORT: **Annual report**

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
October 2021	Annual report	15Sep2020-14Sep2021

4. TITLE AND SUBTITLE
Translating a Stem Cell-Based Therapy for Epidermolysis Bullosa into the Clinic

5a. CONTRACT NUMBER
W81XWH-18-1-0706

5b. GRANT NUMBER
PR171428

6. AUTHOR(S)
Dennis R. Roop

E-Mail: Dennis.Roop@cuanschutz.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Colorado Denver

12801 E 17th Ave, Aurora, CO 80045

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The proposal develops a stem-cell based therapy for recessive dystrophic epidermolysis bullosa (RDEB), which is one of the most severe forms of epidermolysis bullosa (EB), a group of rare inherited skin blistering diseases. To accomplish this goal, we are utilizing genetic correction of patient-specific induced pluripotent stem cells (iPSCs) followed by the differentiation of these corrected iPSCs into epidermal cells and fibroblasts for the generation of composite full thickness skin grafts for transplantation. During this funding period, we continued working on characterizing genetically corrected iPSC-derived skin cells and on developing an *in vivo* bioluminescent assay to determine potential tumorigenicity of iPSC-derived cells. In addition, we have continued working with our cGMP-compliant facility to advance our protocols toward clinical manufacturing. Specifically, we have identified the conditions to perform the purification of keratinocytes differentiated from iPSCs using the CliniMACS Prodigy. This step is important to ensure that keratinocytes that are being transplanted back to the patient are authentic and are not contaminated with partly differentiated cells. As indicated in the previous report, the initiation of many *in vivo* studies was delayed due to the COVID-19 pandemic and the closure of research laboratories on the University of Colorado Anschutz Medical Campus. Therefore, we have requested a no cost extension until September 2022, which will allow us to complete the goals of the original application. This request has been recently approved by the DOD.

15. SUBJECT TERMS
Epidermolysis Bullosa (EB), Recessive Dystrophic EB (RDEB), Induced Pluripotent Stem Cells (iPSC)

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	39	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

1. INTRODUCTION:

This application proposes to advance our previously developed induced pluripotent stem cell (iPSC)-based therapy for recessive dystrophic epidermolysis bullosa (RDEB) into the clinic by adapting the production of genetically corrected patient-specific iPSC-derived epidermal sheets and composite full-thickness skin grafts to current Good Manufacturing Practice (cGMP) standards and by generating a set of preclinical data for submission of an Investigational New Drug (IND) application. Therefore, this application directly addresses the FY17 PRMRP Topic Area "**Epidermolysis Bullosa**". This application is also relevant to one of the Areas of Encouragement identified by the Department of Defense (DoD), the Department of Veterans Affairs (VA), and other relevant stakeholders (as indicated in Appendix 2 of the Program Announcement) since, if successful, the study will result in the approval of a Phase I Clinical trial for a product that may enhance wound healing in inherited epidermolysis bullosa (EB). There are significant procedural differences when a cell therapy product is manufactured under general laboratory settings vs when the same product is manufactured under cGMP-compliant conditions in a cGMP facility. In this proposal, we will transfer the technologies that we have developed using our previous awards from the federal government and private foundations to a product development laboratory at a cGMP-compliant facility. We will also perform a pilot small-scale cGMP production run of genetically corrected RDEB iPSCs and epidermal progenitors derived from genetically corrected RDEB iPSCs. In addition, we will develop a composite skin graft consisting of both genetically corrected iPSC-derived keratinocytes and fibroblasts as an alternative to genetically corrected iPSC-derived epidermal sheets for the treatment of RDEB. If successful and proven to be safe in a clinical trial for EB, the iPSC-based therapy could then be easily expanded to monogenic diseases affecting internal organs, where the difficulty in monitoring adverse effects of an iPSC-based therapy would make them unlikely first targets. The iPSC-based therapy may potentially be applied to military personnel who develop severe blistering following exposure to vesicants, or who suffer from burns over a large portion of their body. In addition, stem-cell based therapies could also be used to accelerate wound repair in military personnel who experience acute injuries, or in older veterans with chronic wounds.

2. KEYWORDS:

Epidermolysis Bullosa (EB); Recessive Dystrophic EB (RDEB), Induced Pluripotent Stem Cells (iPSC); stem cell-based therapy; current Good Manufacturing Practice (cGMP) standards.

3. ACCOMPLISHMENTS:

What were the major goals of the project?

Aim 1: To perform a pilot cGMP-compliant production of genetically corrected RDEB iPSCs.

Task 1.1 To verify the reproducibility of simultaneous reprogramming and gene editing on cells isolated from three independent patients sharing the *COL7A1*$^{c.7485+G>A}$ mutation (3 human subjects)	**100% completed**
Task 1.2 To validate the absence of off-target events post-correction	**90% completed**
Task 1.3 To validate custom-made antibodies to quantify the level of WT Col 7 post-*COL7A1*$^{c.7485+G>A}$ correction	**100% completed**
Task 1.4 To examine the functionality of corrected RDEB iPSCs	**100% completed**
Task 1.5 To transfer the protocols to a product development laboratory at a cGMP-compliant facility, perform optimizations and prepare documentation/batch record	**50% completed**
Task 1.6 To implement a pilot small-scale cGMP-compliant run of the protocol for the generation of a MCB of genetically corrected iPSCs at a cGMP-compliant facility	**10% completed**
Milestone(s) Achieved: Successful pilot cGMP production of genetically corrected RDEB iPSCs	In progress
Local IRB/IACUC Approval	**Completed**
Milestone Achieved: HRPO/ACURO Approval	**Completed**

Aim 2: To perform a pilot cGMP-compliant production of epidermal progenitors from genetically corrected iPSCs..

Task 2.1 To confirm reproducibility of keratinocyte differentiation protocol on iPSCs with the corrected *COL7A1*$^{c.7485+G>A}$ mutation (3 human subjects)	**100% completed**
Task 2.2 To transfer the protocol for the generation of epidermal progenitors to a product development laboratory at a cGMP-compliant facility, perform optimizations and prepare documentation/batch record	**70% completed**
Task 2.3 To implement a pilot small-scale cGMP-compliant run of the protocol for the generation and characterization of genetically corrected iPSC-derived epidermal progenitors at a cGMP-compliant facility	**10% completed**
Milestone(s) Achieved: Successful pilot cGMP production of genetically corrected iPSC-derived epidermal progenitors	In progress

Aim 3: To develop a cGMP-compatible protocol for the generation of a composite skin graft and to generate IND-enabling safety and efficacy data for the FDA

Task 3.1 To optimize a protocol for the differentiation of iPSCs into a fibroblast lineage	**100% completed**
Task 3.2 To examine the functionality of iPSC-derived fibroblasts (12 mice for xenotransplantation)	**70% completed**
Task 3.3 To develop a cGMP-compatible protocol for the differentiation of *COL7A1*$^{c.7485+G>A}$ corrected iPSCs into fibroblasts	**100% completed**
Task 3.4 To generate a composite graft using genetically corrected iPSC-derived keratinocytes and fibroblasts in organotypic cultures and verify type VII collagen deposition	**70% completed**
Task 3.5 To assess wound closure by iPSC-derived composite grafts in immunocompromised mice	**30% completed**
Task 3.6 To assess tumorigenecity and the presence of residual iPSCs in the composite grafts	**50% completed**
Milestone(s) Achieved: Generation of Pre-IND-enabling safety and efficacy data for composite grafts using cGMP-compatible protocols	In progress

What was accomplished under these goals?

For this reporting period describe: 1) major activities; 2) specific objectives; 3) significant results or key outcomes, including major findings, developments, or conclusions (both positive and negative); and/or 4) other achievements. Include a discussion of stated goals not met. Description shall include pertinent data and graphs in sufficient detail to explain any significant results achieved. A succinct description of the methodology used shall be provided. As the project progresses to completion, the emphasis in reporting in this section should shift from reporting activities to reporting accomplishments.

Task 1.1 To verify the reproducibility of simultaneous reprogramming and gene editing on cells isolated from three independent patients sharing the COL7A1$^{c.7485+G>A}$ *mutation (3 human subjects).*

This task has been accomplished in prior funding periods. Please, see previous annual and semi-annual reports.

Task 1.2 To validate the absence of off-target events post-correction.

As indicated in previous progress reports, we did not identified any off-target modifications in our three corrected RDEB iPSC clones (CO1-60, CO2-6 and CO4-194). We are currently collecting specimens for whole-genome sequencing to assess the rate of mutation accumulation in iPSC-derived keratinocytes and other potential off-target events promoted by our Cas9-mediated gene correction strategy. Once the results of the sequencing are obtained, we will analyze the level of mutations in 130 selected genes that are associated with solid tumors. These 130 genes will be taken from the CLIA-certified STAMPT assay performed at Stanford University: https://www.stanfordlab.com/esoteric/test-stanford-actionable-mutation-panel-for-solid-tumors.html. Given that RDEB patients are particularly prone to SCC, we will also assess a set of genes implicated in SCC: BHLHE41, KRT17, MAL, ERCC2, HRAS, KRAS, etc. This analysis will ensure that we have not expanded a clone with a preexisting cancer transformation event.

Since we have multiple patients carrying the Colorado founder mutation, we analyzed the origin of this mutation. This study was important to identify the baseline of mutations in these patients in preparation to our whole genome sequencing analysis. In addition, this study would facilitate the identification of additional patients suitable for our clinical trial based on ancestry. This study included several patients with the Colorado

mutation, as well as multiple RDEB patients with other recurrent mutations (this part of the study was funded by other sources in collaboration with other researchers). We found that the Colorado mutation originated on Native American haplotype with a substantial amount of Sephardic ancestry. These results are summarized in our recent publication (see the publication section of the report), and funding from the Department of Defense (DOD) (W81XWH-18-1-0706) was acknowledged (see Appendix).

Task 1.3 To validate custom-made antibodies to quantify the level of WT Col 7 post-COL7A1$^{c.7485+G>A}$ correction.

This task has been accomplished in prior funding periods. Please, see previous annual and semi-annual reports.

Task 1.4 To examine the functionality of corrected Col7 in RDEB iPSCs (18 mice for teratoma and 48 mice for xenotransplantation)

This task has been accomplished in prior funding periods. All generated iPSC clones express pluripotency markers, show a normal karyotype and can be differentiated into all three germ layers via a teratoma formation assay. All clones have been successfully differentiated into keratinocytes and show the restoration of Col7 expression.

Task 1.5 To transfer the protocols to a product development laboratory at a cGMP-compliant facility, perform optimizations and prepare documentation/batch record.

We continue working with the Gates Biomanufacturing Facility (GBF), to adapt our reprogramming and gene editing protocols to cGMP standards. During the reporting period, we have trained GBF personnel to produce modified mRNA and maintain iPSCs. We are currently training them to perform our reprogramming and gene editing protocols using RDEB primary fibroblasts.

Task 1.6 To implement a pilot small-scale cGMP-compliant run of the protocol for the generation of a MCB of genetically corrected iPSCs at a cGMP-compliant facility.

The initiation of this aim was delayed due to the Covid-19 pandemic. GBF personnel have reviewed our material list and are currently testing cGMP materials in preparation for the pilot cGMP-compliant run.

Task 2.1 To confirm reproducibility of keratinocyte differentiation protocol on iPSCs with the corrected COL7A1$^{c.7485+G>A}$ mutation (3 human subjects).

This task has been accomplished in prior funding periods. Please, see previous annual and semi-annual reports.

Task 2.2 To transfer the protocol for the generation of epidermal progenitors to a product development laboratory at a cGMP-compliant facility, perform optimizations and prepare documentation/batch records.

We are working with the GBF to develop a protocol for the purification of authentic iPSC-derived keratinocytes based on CD104 expression using the CliniMACS Prodigy. This aim was also delayed due to the pandemic. During this funding period, we have optimized conditions for the purification of keratinocytes derived from genetically corrected iPSCs using the CliniMACS Prodigy. This step is important to ensure that keratinocytes that are being transplanted back to the patients are authentic and are not contaminated with partly differentiated cells.The CliniMACS Prodigy® is a closed platform with the capability of providing fully automated and scalable cell manufacturing procedures of various cell types on a single device and within a single process setup. Our preliminary results (see previous reports) indicate that the enrichment for the CD104 marker produces the most authentic keratinocytes. Before initiating the Prodigy run, we screened other antibodies to identify the best combination of markers to autheniticate iPSC-derived keratinocytes post-enrichment. Based on

our results, the iPSC differentiation culture contains keratinocytes, as well as fibroblast-like cells expressing CD90. These CD90 cells can be eliminated by a regular MACS enrichment procedure using CD104 beads (**Fig. 1**). Therefore, we screened antibodies specific to fibroblast markers (CD29, CD73, CD90 and CD105), keratinocyte markers (CD49f, D54 and CD104) and iPSC markers (SSEA4 and TRA-1-60). Table 1 below shows the result of our flow cytomtry analysis using keratinocytes (HEK), fibroblasts (HDF) and iPSCs.As shown in Table 1, iPSCs express a substantial amount of the keratinocyte-specific CD49f marker. Therefore, this marker will not be used for validation purposes. However, the CD104 marker shows a high specificity toward keratinocytes. Similarly, CD90 is more specific to fibroblasts and iPSCs and therefore can be used as a negative marker for enriched iPSC-derived keratinocytes.

Table 1: Keratinocyte, fibroblast and iPSC marker afinity by flow cytometry

MFI	**HEK 1**	**HEK 2**	**HEK 3**	**HDF 1**	**HDF 2**	**iPSC 1**
a-CD49f (hybridoma)	76.17	68.58	49.92	10.41	11.01	14.37
a-CD49f (REA)	6.05	6.66	3.78	2.05	2.09	6.69
a-CD54	3.63	5.43	6.38	2.43	1.78	1.74
a-CD104	**95.49**	**84.31**	**31.65**	**1.5**	**1.46**	**1.5**
a-CD29	636.84	579.54	347.34	91.78	112.84	30.6
a-CD73	476.55	441.37	257.83	80.32	136.17	1.2
a-CD90	**7.11**	**5.43**	**3.9**	**17.45**	**12.22**	**34.38**
a-CD105	1.31	1.23	1.3	2.35	3.4	0.47
a-SSEA4	1.27	1.42	1.44	2.04	1.66	8.54
a-TRA-1-60	3.06	2.89	3.99	2.62	1.07	2.56

Since the Prodigy run requires cells to be in suspension for a prolong time, we evaluated the survival of keratinocytes in suspension at different intervals in the presence of two different buffers suitable for the Prodigy enrichment: PBS and CnT-PR medium. We have noticed a drastic drop in cell survival during the first hour of cells being in suspension. However, the cells survived better in the presence of CnT-PR (**Fig.2**).

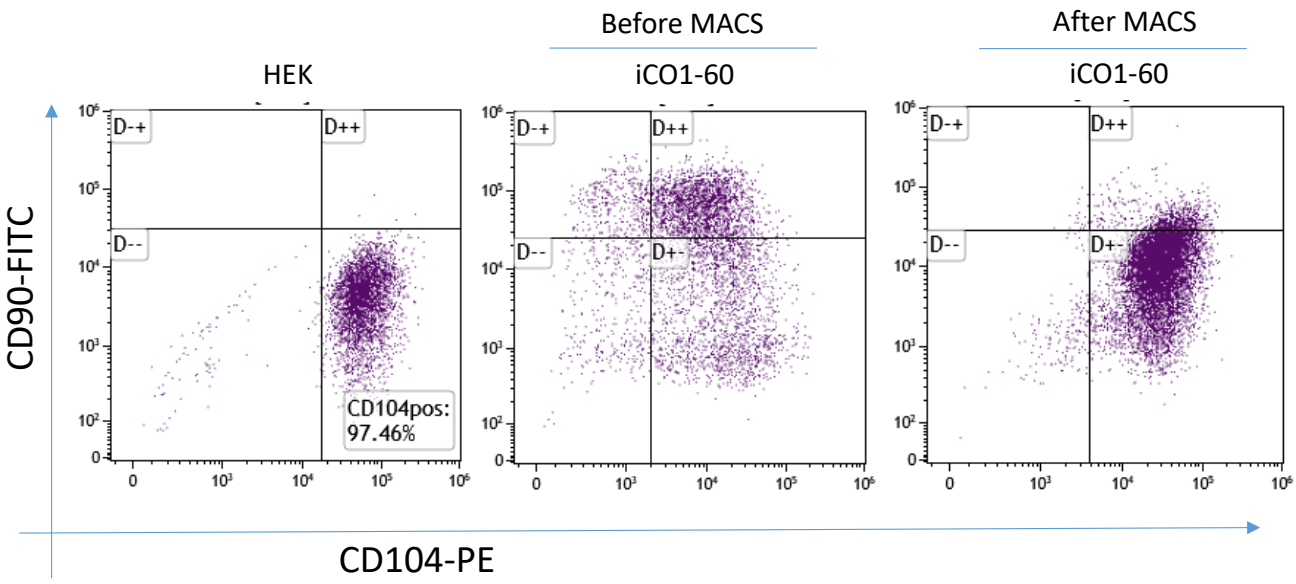


Fig. 1. Enrichment for iPSC-derived keratinocytes using α-CD104 magentic beads. Genetically corrected RDEB iCO1-60 IPSCs were differentiated into keratinocytes (iKCs) using 1µM retinoic acid and 5 ng/µL BMP4. At day 45, iKCs were harvested (middle panel) and enriched for CD104 using MACS (right panel). The marker expression profile in the resulting culture was similar to that of healthy keratinocytes (HEK, left panel). The cells expressed high CD104 and low CD-90 levels.

When we performed a MACS enrichment protocol (**Fig. 3**) at a manufacturing scale using approx 4 x 10^7 keratinocytes, the procedure took approximately 6 hours, which is similar to the Prodigy run. We sucessfully recovered approx. 30% of live cells, which indicates that for the cGMP prodcution run we will need to recalculate the starting cell culture number taking into account a 30% survival rate of enriched keratinocytes post-enrichment. Importantly, when MACs enrichment was performed with PBS vs CnT-PR using a smaller number of cells (1 x10^6), the cell recovery rate was similar between the conditions. However, after plating the cells purified in the presence of CnT-PR survived substantially better (**Fig. 4**). Therefore, the Prodigy run will be performed using CnT-PR as the buffer. We are currently preparing for the actual CliniMACS Prodigy run where 25% of keratinocytes will be mixed with 75% of CD90+ MSCs (to mimic cell populations in iPSC diffferentiation cultures), and the efficiency of keratinocyte enrichment will be assessed. The eluted keratinocytes will also be tested on a funtional level to ensure that the procedure does not compromise cell characterisitcs. We are also training GBF personnel to perform iPSC differentiation into keratinocytes. Once this is accomplished, the Prodigy enrichment run on iPSC-derived keratinocyte cultures will be performed.

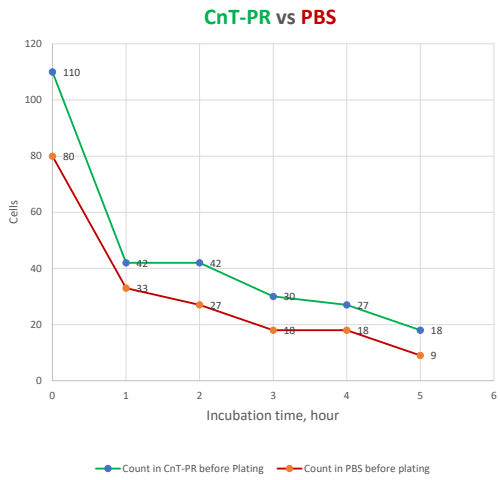


Fig. 2. Survival rate of keratinocytes in suspension in the presence of PBS vs CnT-PR. Healthy keratinocytes were detached and resuspended in either PBS (red) or CnT-PR(green). The cell suspension was kept at +4^{o}C for the indicated time on the X axis. Live cells were calculated manually using Trypan blue. The use of CnT-PR improves cell survival in suspension.

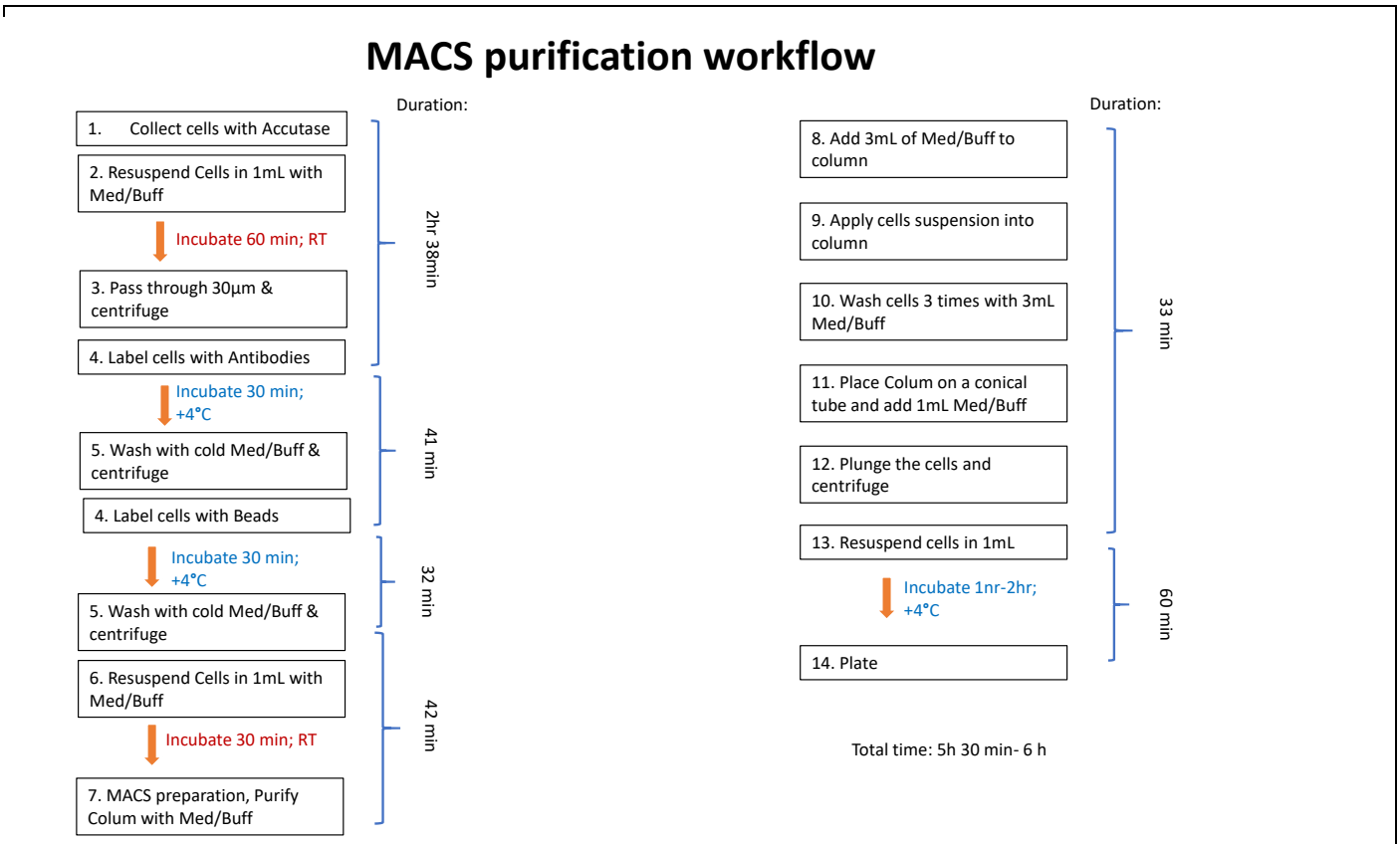


Fig. 3. A schematic depicting the MACS enrichment protocol for iPSC-derived keratinocytes.

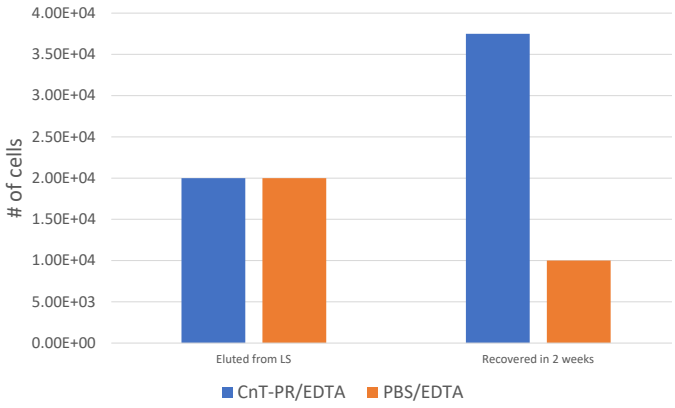


Fig. 4 Keratinocyte recovery after the MACS enrichment protocol performed with PBS vs CnT-PR. Healthy keratinocytes were subjected to the MACS enrichment protocol as depicted in Fig. 3 above using either PBS or CnT-PR as a buffer. The number of live cells was calculated immediately after elution from the column (Blue and orange columns on the left: Eluted from LS). The resulting numbers were similar between PBS and CnT-PR. Eluted cells were also plated and expanded over the two-week period. The cells processed using PBS as a buffer did not proliferate, while the cells processed with CnT-PR proliferated normally (Blue and orange columns on the left: Recovered in 2 weeks).

<u>Task 2.3</u> To implement a pilot small-scale cGMP-compliant run of the protocol for the generation and characterization of genetically corrected iPSC-derived epidermal progenitors at a cGMP-compliant facility

GBF personnel have reviewed our material list and are currently testing cGMP materials in preparation for the pilot cGMP-compliant run. We have provided two clones of genetically corrected RDEB iPSCs to the GBF. These clones were tested for mycoplasma, bacterial contamination and blood borne pathogens and found to be negative. They have already been expanded under cGMP conditions in preparation for the cGMP-compliant manufacturing of iPSC-derived keratinocytes.

Task 3.1 To optimize a protocol for the differentiation of iPSCs into a fibroblast lineage.

We have optimized the protocol for the differentiation of iPSCs into fibroblasts (see our previous reports).

Task 3.2 To examine the functionality of iPSC-derived fibroblasts.

Gene corrected RDEB iPSCs have been successfully differentiated into fibroblasts using the protocol that we described in the previous report. These cells are now being grown as 3D skin equivalents and if successful the grafting will be performed in the next few months to validate the functionality of these cells.

Task 3.3 To develop a cGMP-compatible protocol for the differentiation of COL7A1$^{c.7485+G>A}$ corrected iPSCs into fibroblasts.

We successfully modified our fibroblast differentiation protocol to eliminate the embryoid body step. We have also successfully replaced FBS with pharmaceutical grade human serum, making our protocol cGMP-compatible. This new protocol has been used to differentiate gene corrected RDEB iPSCs into fibroblasts in Task 3.2. The resulting cells look identical to normal fibroblasts and express vimentin, CD90 and Collagen I as markers of fibroblast differentiation.

Task 3.4 To generate a composite graft using genetically corrected iPSC-derived keratinocytes and fibroblasts in organotypic cultures and verify type VII collagen deposition.

As described in the previous report, we have modified the procedure to produce composite grafts for future transplantation. Instead of using devitalized dermis, we focused on using either Collagen I or fibrin gel as a substrate for dermal fibroblasts in 3 D skin equivalents. Fibrin is already approved for clinical applications as a wound bandage. Therefore, from regulatory perspectives, the production of fibrin-based 3D skin equivalents will be easier to adapt to cGMP manufacturing. Since the procedure for 3D skin equivalents was initially adapted to Collagen I, we tested if the fibrin gel can be used instead as a more clinically relevant matrix. As expected, both Collagen I- and fibrin-based approaches produce good quality 3D skin equivalents suitable for transplantation (**Fig. 5**). We are currently producing multiple 3D skin equivalents using keratinocytes derived from genetically corrected RDEB iPSCs.

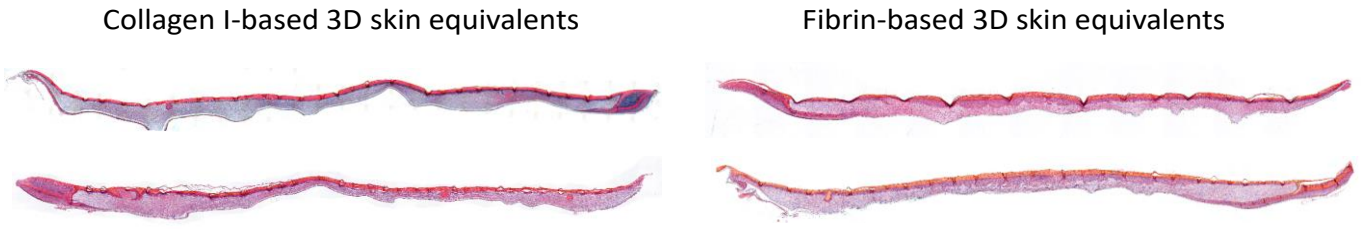


Fig. 5. Comparison between 3D skin equivalents produced using collagen I vs fibrin. The images show H&E images of the successful generation of 3D skin equivalents using either collagen I (left) or fibrin (right). Both protocols result in the formation of a well-developed epidermis and dermis.

Task 3.5 *To assess wound closure by iPSC-derived composite grafts in immunocompromised mice.*

We have grafted several 3D skin equivalents using genetically corrected RDEB iPSC-derived cells and are expecting the first results shortly. To design the safety studies correctly, we need to know how long human cells can persist on an immunocompromised mouse. The FDA would require us to monitor the mice for as

long as the human cells persist. In order to address this question, we employed the IVIS spectrum system (www.perkinelmer.com) that is capable of detecting bioluminescence in living cells and whole organisms. We transduced healthy human keratinocytes with a bicistronic lentivirus encoding dT Tomato and AKALuciferase. AKALuciferase is a mutant version of the firefly luciferase that produces near infrared bioluminescence from its optimized D-Luciferin analogue AKAlumine or "TokeOni" (its trademark name). The advantage of near infrared bioluminescence is that it can be detected deep within an organism and thus requiring fewer number of cells to produce a robust signal. We grafted transduced cells using two different grafting chamber assays. One is our standard grafting assay where 5 $x10^6$ keratinocytes are mixed with 5 $x10^6$ fibroblasts and delivered as a cell suspension into a silicone chamber implanted on the back of an immunocompromised mouse. In another assay, the cells are transplanted as a stepwise procedure using a low-density fibrin gel formulation. Specifically, 1 $x10^6$ fibroblasts are resuspended in the fibrin gel and delivered into the chamber first to recreate a human dermis. This is followed by delivering 2 $x10^6$ keratinocytes resuspended in the fibrin gel. Fibrin prevents cell leakage from the chamber and facilitates the formation of the human skin in the wound. After transplanting transduced keratinocytes, we tested the ability of the IVIS system to detect *in vivo* AKALuciferase expressing keratinocytes at 5, 6 and 7 weeks post-transplantation. Visualization was done in anesthetized mice injected with 100 µl of saturating levels of TokeOni (15mM). Both grafting systems showed successful engraftment of AKALuciferase positive keratinocytes at all three time points (**Fig. 6**). However, by week 7, the signal from AKALuciferase starts diminishing, indicating that human keratinocytes do not persist for long. These results may give us an opportunity to discuss the shortening of our *in vivo* toxicology studies with the FDA during our pre-IND meeting. This experiment is still ongoing, and more similar grafting experiments are being initiated for statistical analysis. In parallel, we have initiated the differentiation of healthy iPSCs with an AAVS locus-specific knock in of the dT Tomato-p2a- AKALuciferase construct into keratinocytes. Once differentiation is accomplished, these cells will be grafted using similar grafting assays to determine how long iPSC-derived keratinocytes persist on immunodeficient mice. This experiment will define the length of our safety studies for the IND application.

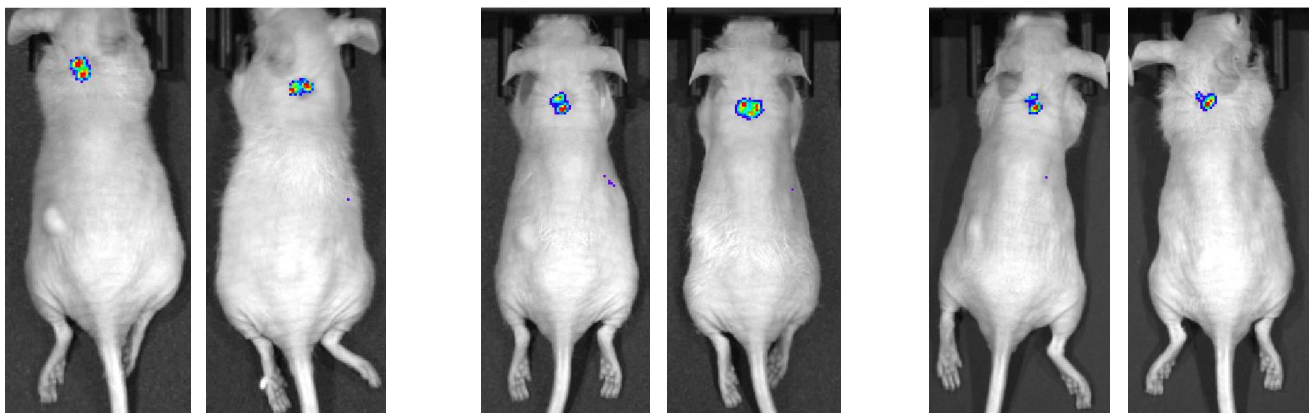


Fig. 6. Persistence of human keratinocytes expressing AKALuciferase on a nude mouse. Human keratinocytes were transduced with a lentivirus expressing AKALuciferase and then grafted using two types of grafting assays. In one assay, the fibrin gel was used to transplant fibroblasts and keratinocytes (the mouse on the left). In the other assay, fibroblasts and keratinocytes were grafted in a suspension using a classical grafting procedure. The grafts were analyzed using IVIS at 5, 6 and 7 weeks post-grafting. By week 7, the signal from the grafts starts to diminish.

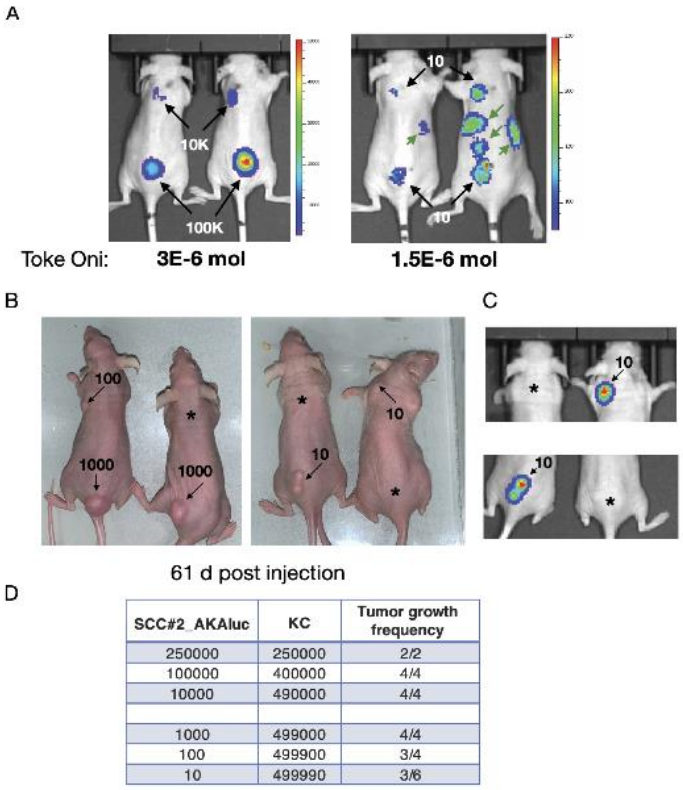
Task 3.6 *To assess tumorigenecity and the presence of residual iPSCs in the composite grafts.*

I. Summary of the previous semi-annual progress report: development of bioluminescent skin cancer growth models in immunocompromised mice

The overall goal of these studies was to determine the potential tumorigenicity of human induced Pluripotency Stem Cells (iPSC) or iPSC-derived keratinocytes. In the last progress report, we established the *in vivo* growth of skin cancer cells in immunocompromised mice as a positive control while developing *in vivo* bioluminescent imaging approaches to enhance our ability to detect rare oncogenic events in living animals. We employed the IVIS spectrum system (www.perkinelmer.com) with cells co-expressing the AKAluciferase bioluminescence producing enzyme. AKAluciferase produces near infrared bioluminescence in the presence of its optimized analogue, TokeOni (Iwano et al. Science 2018 and Kuchimaru et al. Nature 2016). The advantages of this system are 3-fold: (1) near infrared bioluminescence can be detected deep within an organism and thus requires fewer cells to produce a robust signal; (2) Bioluminescence is highly quantifiable, and (3) it allows repeated imaging for the tracking of bioluminescent cells in live animals over an extended period of time without the use of invasive procedures.

As we described in the previous progress report, we transduced two independent skin Squamous Cell Carcinoma Cell (SCC) cell lines, namely RDEBSCC_2(SCC#2) and IC1, with a dual expressing dT-Tomato/Akaluciferase lentivirus and bulk sorted these cells to use in our bioluminescent system. Transduced cells were then injected subcutaneously (sq) in Matrigel or grafted in a slurry with human keratinocytes (KCs) and fibroblasts (FBs) using a silicone chamber system inserted into Nu/J athymic nude mice. Using these *in vivo* models, we optimized the imaging kinetics of bioluminescence production and established preliminary growth curves from SCC cells grown in a sq or silicone graft chamber environment. These experiments indicated that SCC cells grown in a sq environment showed faster growth kinetics compared to SCC cells grown in the chamber environment. The slower growth of SCC cells in the graft chamber was likely due to the more dynamic environment of the chamber, which mimics the events found in skin wounding and healing. It takes several weeks for KCs and FBs in the graft to arrange themselves into functional units of skin (e.g., dermal vs epidermal structures). Last, we also began dilution studies to determine the minimum number of SCC cells that could be detected after injecting cells into the sq environment. In initial experiments we were able to detect down to 100 cells, 1 d after injection. However, we were unable to detect lower than 100 cells due to background bioluminescence originated from the liver metabolism of the TokeOni (See Nakayama et al, Int J. Mol Sci, 2020).

II. **Overcoming background bioluminescence of TokeOni.** To overcome the background signal when imaging less than 1000 cells, we titrated the concentrations of TokeOni down to a ½ saturating dose (100 ml of a 15 mM stock solution or 1.5E-6 mol). This allowed the detection of 10 SCC cells 1 day post injection. We used SCC#2_AKAluc cells in these experiments because this cell line produced more robust bioluminescence than the other SCC cell lines we tested (See Section III). Lowering the dose of TokeOni was a successful strategy to minimize the background because the affinity of AKAluciferase for Toke Oni is much higher than that of the endogenous mouse liver enzymes (**Fig. 7A**). The competition for TokeOni was more evident when there was a greater number of tumor cells engrafted in the mice (See Section III and **Fig. 10**). Nevertheless, we were able to observe palpable tumor growth from 10 SCC cells after 4 weeks (n=6 injection sites on 4 mice) in two independent experiments. We did observe that half of the tumors began to regress by 2 weeks as measured by a decrease in bioluminescent signal (See **Fig. 7B-C** and data not shown). We also observed tumor regression from injection sites that received 100 SCC cells, but this was less frequent. On the other hand, we did not observe tumor regression with ≥ 1000 SCC cells (**Fig. 7D**) when using young recipient mice. In summary, our results indicate that cancer initiating cells are common in the RDEBSCC_2 cell line, allowing us to detect a little as 10 bioluminescent cells using the IVIS Spectrum system.



SCC#2_AKAluc	KC	Tumor growth frequency
250000	250000	2/2
100000	400000	4/4
10000	490000	4/4
1000	499000	4/4
100	499900	3/4
10	499990	3/6

Fig. 7. Summary of tumor growth from limiting number of cancer cells. (A) SCC#2_AKAluc cells were resuspended in Matrigel with human keratinocytes (KC) to a total of 500,000 cells and injected sq on the lower back and between the shoulder blades (black arrows) of nude mice. At 24h post injection, mice were imaged on the Ivis Spectrum system with either saturating levels of TokeOni for mice with ≥ 1000 cancer cells or a 1/2 dose for mice with ≤ 1000 cells. Imaging was done weekly, and mice were monitored for tumor growth. Note the presence of background bioluminescence mid dorsally (right panel, green arrows). Luminescent scale in counts is shown on right of images. (B) Regression of tumor growth was seen with limiting numbers of SCC cells. Large tumor nodules were evident at sites of injection with ≥ 1000 cancer cells. However, the frequency of growth decreased at sites with ≤ 100 cells (right panel, asterisks). (C) Regression observed in B was confirmed by imaging. (D) Summary of two independent experiments showing the frequency of tumors at sites of injection with limiting number of SCC#2_AKAluc cells.

III. **Characterization of bioluminescent metastatic skin cancers cells from the sq environment.** As described in our first progress report, we tested several human skin cancer cell lines recommended by our collaborators or that were reported to form sq tumors quickly (See Watt et al. Oncogene 2011; Atasanova et al. Clin Cancer Res 2019; Welch et al. Int J Cancer 1991). **Fig. 8A** shows the growth profiles for the three SCC cell lines (SCC#2, IC1, and 106K) and one melanoma cell line (C8161) we tested after injecting 1E6 cells (n=3 mice per line) in a classical Matrigel sq flank tumor assay. Both the IC1 and SCC#2 cell lines reached maximum bioluminescence by 25-28 d post injection, whereas the 106K cell line formed tumors at a much slower rate, never reaching the volumes nor bioluminescence of the two other cell lines. In contrast to SCC cells, the C8161 cell line grew rapidly and after 4 weeks (**Fig. 8A-B**) forced the termination of the experiment due to a heavy tumor burden incurred by the recipient mice.

Next, we determined if skin tumor cells left the primary site of injection and infiltrated internal tissues. First, we looked at SCC cells. Analysis of the internal organs of SCC#2_AKAluc bearing mice showed a

background TokeOni signal in the liver of the recipient mice, but negligeable bioluminescence in other organs (**Fig. 8C**). In contrast, C8161_AKAluc tumor bearing mice showed detectible bioluminescent signal from

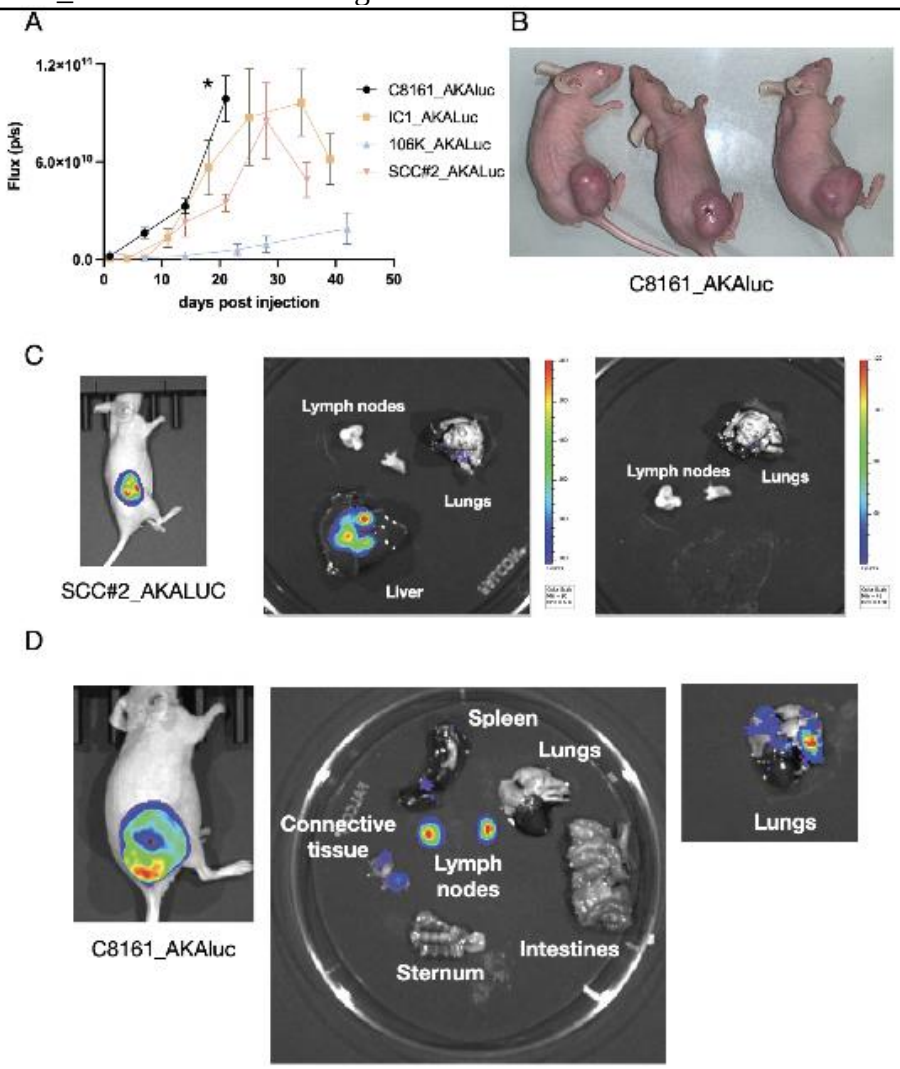


Fig. 8. Detection of bioluminescent metastasis from sq tumors. (A) AKAluciferase expressing cells were suspended in Matrigel and injected sq into the flank of nude mice (n=3, 1E6 cells/mouse). Mice were then imaged on the IVIS Spectrum system the following day to establish a base line reading and every 7 d thereafter. The graph shows the growth profiles of three SCC (IC1,106K, and SCC#2) and one melanoma (C8161) cell lines. Saturating doses of TokeOne (3E-6 mol) were used to image tumors. Mean +/-SEM is shown. *-mice sacrificed due to excessive tumor volume. (B) Left panel shows mice with large sq tumors 4 weeks after injecting C8161 cells. (C) Mice bearing a flank SCC#2 tumor (57 d post injection) and its internal organs were imaged 10 min after TokeOni administration. Note the background bioluminescence in the liver (middle panel) and minimal signal in other organs in its absence (right panel) (D) Mice bearing a flank C8161 tumor were analyze as in (C). Note the bioluminescence in the lymph nodes, connective tissues and lung likely indicating the presence of metastatic C8161_AKALuc cells (middle and right panels). Representative images are shown.

several organs, notably, the lungs (**Fig. 8D**). This indicated that unlike the SCC cell lines, the C8161 cell line can form flank tumors that frequently metastasize, consistent with what has been reported for this cell line in the literature (See Welch et al. Int J Cancer 1991 for example).

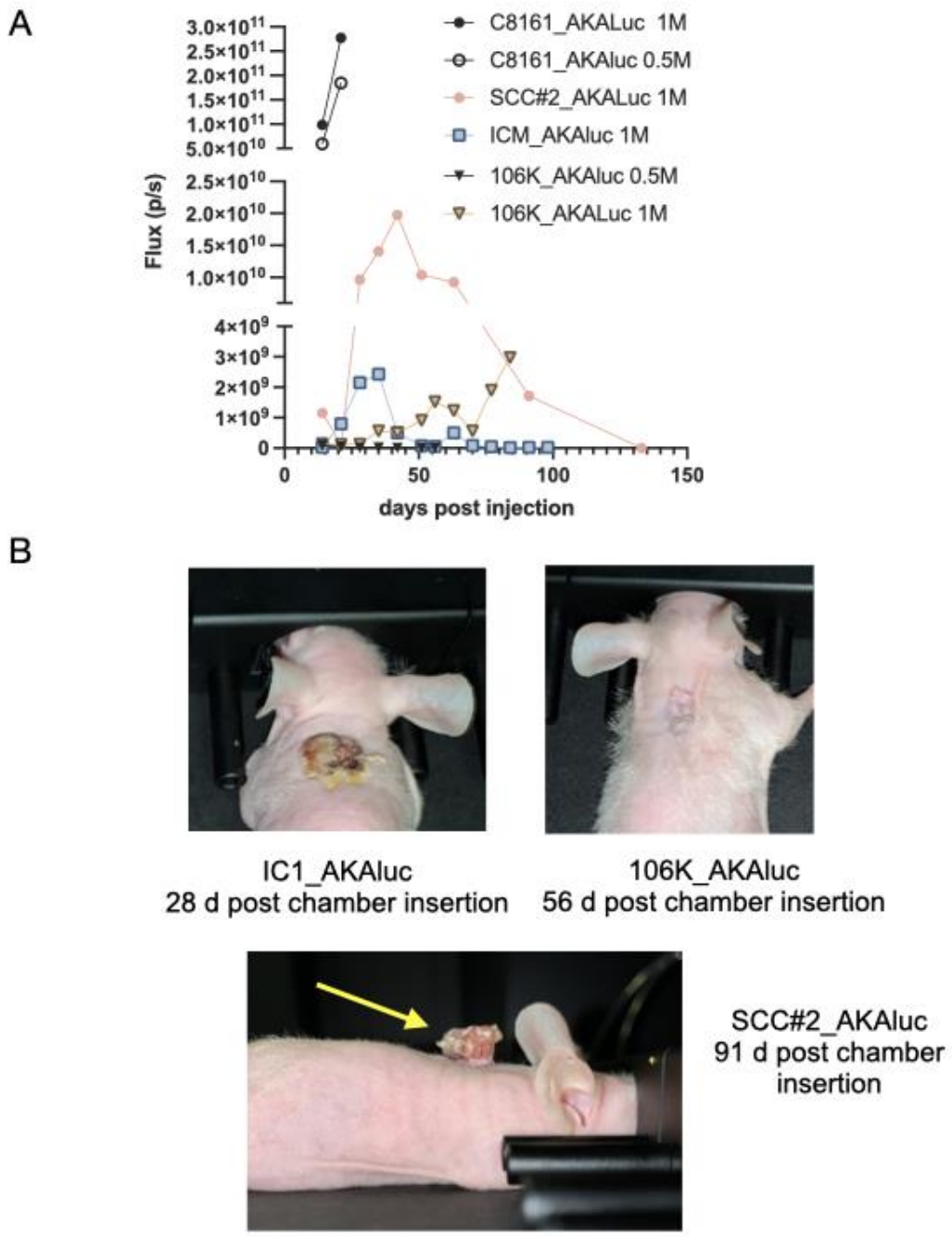


Fig. 9. Growth profiles of skin cancer cells in the graft chamber environment. (A) AKAluciferase expressing cells (0.5-1E6) were mixed with human 5E6 KCs and 5E6 Fibroblasts (FB) and layered into a silicon chamber that was inserted into the back skin of nude mice. The graphs show the growth profiles from three SCC (IC1,106K, and SCC#2) and one melanoma (C8161) cell line in the graft chamber system. C8161 containing chambers developed large subdermal tumors by 28 d that resulted in the mice being removed from the study due excessive tumor burden. (B) Images show the superficial growth of SCC cells after removal of the graft chamber. In the three SCC cell lines tested, tumor growth was associated with the scab that forms at the insertion site of the graft, where KC and FB reform skin. Lower panel shows a mouse with an SCC#2_AKAluc tumor. The superficial tumor fell off by 133 d post chamber insertion resulting in loss of bioluminescence (See graph in (A)).

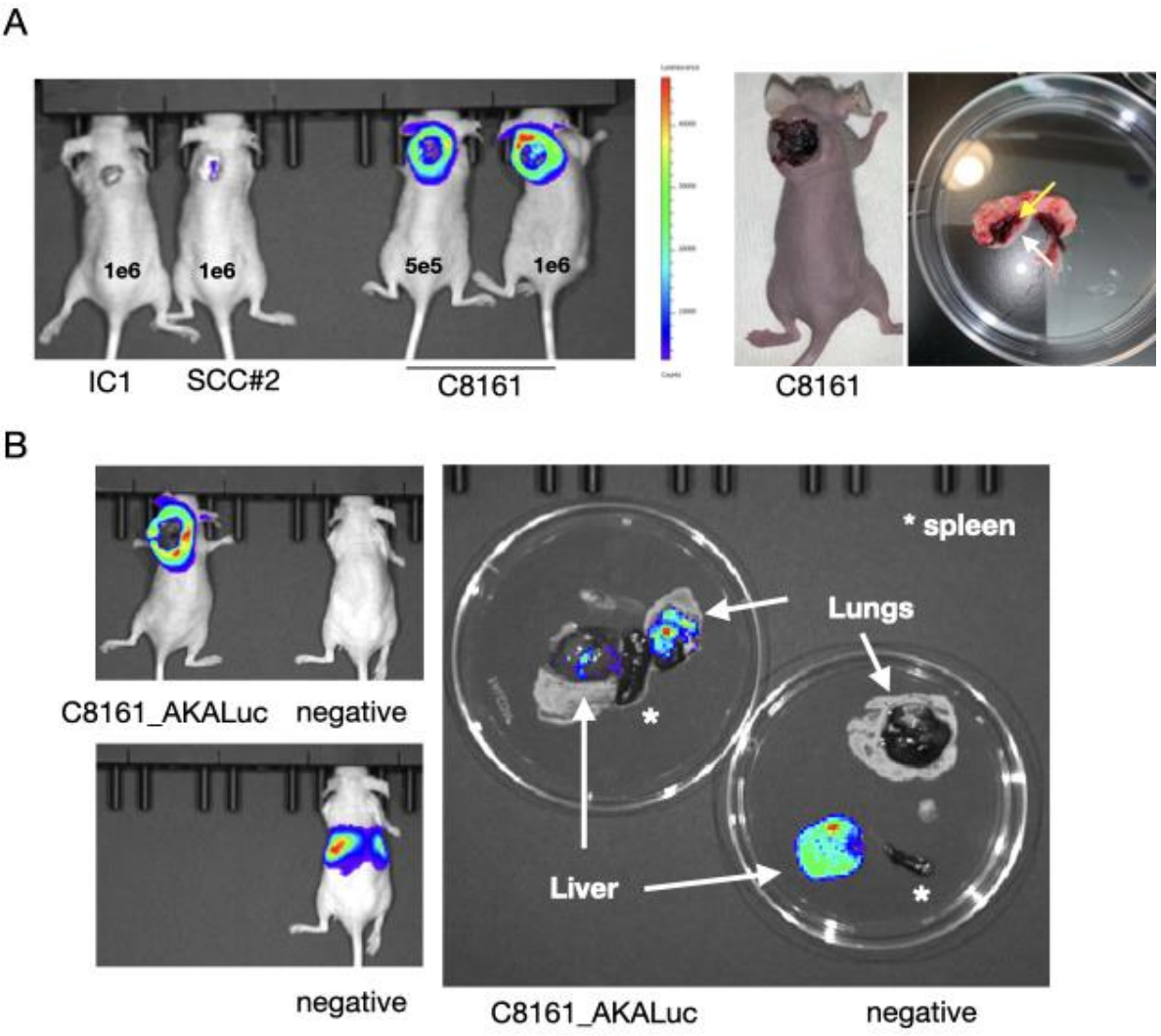


Fig. 10. Bioluminescent metastasis from the graft chamber environment using C8161 cells. (A) In the left panel, the IVIS Spectrum system was used to compare the bioluminescence production from graft chambers containing IC1_AKAluc (35 d post grafting), SCC#2_AKAluc SCC (63 d post grafting), or C8161 (28 d post grafting) cells. The number of tumor cells used in each graft is depicted in the image. Note the apparent low bioluminescence in IC1 and SCC#2 tumors compared to the overwhelming signal from the C8161 tumors. The right panel shows the necrotic center of C8161 tumors at the site of the graft and after the tumor was dissected. The yellow arrow points to necrotic tissue while the white arrow shows non-necrotic tumor growth. (B) Control (negative), C8161 tumor-bearing mice, and their internal organs were imaged 10 min after TokeOni administration. Note the bioluminescent background in the liver of the intact negative mouse and after its dissection. This background was diminished in the liver of C8161 tumor bearing mice due to the higher metabolism of TokeOni by AKAluciferase in tumor cells. Also note the enlarged spleen in C861 tumor bearing mice.

IV. **Bioluminescent metastasis from the skin graft environment.** Cell-based therapies targeting skin will require the introduction of iPSC-derived cells into patients while minimizing the potential for oncogenic events. For this reason, we wanted to understand more deeply the kinetics of tumor cell growth and the potential metastatic spread in a more relevant environment, namely the silicone graft chamber (skin reconstitution) system. In these experiments 0.5-1E6 skin cancer cells were mixed with hKC and hFB. A slurry of cells was then deposited in a silicon chamber embedded into the skin of athymic nude mice. After 7 d the chambers were removed to allow for the hKC and hFB to reconstitute skin. In the grafts that received skin cancer cells, we waited an additional 7 d after the removal of the chambers before starting weekly imaging. The graph in **Fig. 9A** shows the growth profiles of the three SCC cell lines (IC1, SCC#2, and 106K) and the one melanoma cell line (C8161) we tested. Maximal bioluminescence occurred between 35-42 d post chamber insertion in the IC1 and SCC#2 cell lines, which then declined gradually for SCC#2 cells and more sharply for IC1 cells. SCC#2 cells also produced the largest increases in bioluminescence of all the SCC cell lines we tested. Another important feature in the graft chamber environment was the outward growth of SCC tumors (See **Fig. 9B**). Regarding this point, histological analysis showed considerable keratinization of the tumors (data not shown). In time the superficial growths dropped off the mouse with a concomitant drop in bioluminescence (See **Fig. 9A**). This was the case for the mouse in **Fig. 9B** with the yellow arrow, which lost its superficial growth 91 d after reaching its maximal bioluminescence production (**Fig. 9B).** Necropsy of mice at the end of the experiments did not reveal obvious indications of metastatic spread. This was similar to what we observed when SCC cells were injected into the sq environment (See **Fig. 8**).

We also tested either 0.5E6 or 1E6 106K cells in the graft environment, where this cell line grew more slowly than either IC1 or the SCC#2 cell lines. The bioluminescence from the 0.5E6 cells decreased continually until the signal disappeared after 56 d post chamber insertion. On the other hand, the chamber that received 1E6 cells initially followed the growth pattern of the IC1 and SCC#2 cell lines, namely peaking and then showed a decrease in bioluminescence. The decrease in signal coincided with the loss of superficial material but these tumors began to regrow (See **Fig. 9A**).

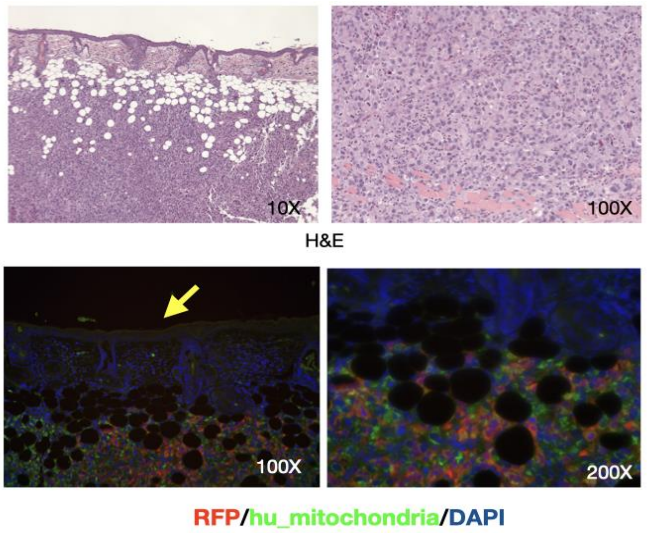


Fig. 11. Detection of C8161_AKAluc tumors cells by immunostaining. Biopsies from C8161 graft chamber tumors were process for histology using standard techniques. Top panels show H&E stained cross sections. Note the presence of tumor cells underneath the skin adjacent to the graft chamber. Bottom panels show co-staining of tumor sections with an RFP antibody (Rockland) that cross reacts with the dT-tomato protein (Red color in images) and with anti-human (hu) mitochondrial antibody (Abcam) that distinguishes human cells from mouse cells (Green color in images). Note that the human mitochondria are stained in a punctate pattern. The yellow arrow points to the epidermis. Magnification is stated on the bottom right corner of each panel.

In sharp contrast to the behavior of SCC cells, mice bearing the C8161 cell line developed subdermal tumors rapidly and had to be euthanized shortly after 28 d post chamber insertion. This was due to the excessive tumor burden and presence of ulcerated/necrotic skin (**Fig. 9A** and **Fig. 10A**). C8161 cells quickly invaded deeply into the surrounding tissue of the chamber graft (**Fig. 11**). We analyzed tumor biopsies using immunohistochemical approaches. We were able to confirm the presence of C861 and SCC tumor cells (not shown) using a red fluorescent protein (RFP) antibody that also recognizes the Td-tomato protein. We also co-stained for a human (hu) mitochondrial antigen that helped to distinguish the exogenous human cells from the recipient's mouse cells. However, the detection of Td-Tomato expression (RFP) was uneven and variable in comparison to the detection of the hu mitochondrial antigen. This likely reflects the use of bulk sorted cancer cells that results in a heterogenous population of cancer cells expressing different levels of both Td-Tomato and AKAluciferase. The internal organs of C861 bearing mice also showed strong bioluminescent light production (See example in Fig. 10B), indicating metastatic spread from the graft chamber environment. Lastly, C8161 bearing mice also had enlarged spleens, showing the activation of the environment. Lastly, C8161 bearing mice also had enlarged spleens, showing the activation of the immune system. Indeed, numerous immune cells were evident in the primary tumors by H&E staining (**Fig. 11).**

Growth profiles of iPS cell (iPSC) derived teratomas by bioluminescence detection. Having established the growth of skin cancer cells in our hands, we also began experiments to characterize the growth of iPSC-derived teratomas *in vivo*. Towards this goal, we created a human iPSC line that stably co-expressed the dT-Tomato and AKAluciferase proteins. We functionally tested this line *in vitro* by incubating colonies with TokeOni for bioluminescent production (**Fig. 12A**). Teratoma formation of iPSCs is considered a hallmark test of pluripotency as tumors arising from iPSCs contain features of all three germ layers (Takahashi and Yamanaka Cell 2006; Takahashi et al., Cell 2007; Nelakanti et al., Curr Protoc Stem Cell Biol 2016; and Kogut et al. Nat Comm 2018). However, the literature varies considerably on the procedures and background strain of immunocompromised mice to use in teratoma formation. In the following experiments, we use athymic nude mice as hosts due to the absence of hair that facilitates the visual and bioluminescent monitoring of tumor formation. To achieve a near single cell suspension, AKAluc_iPSC colonies were digested with Accutase until colonies were in small clumps and floating. The clumps were then broken down further until a near single cell suspension was achieved. Over digestion or over trituration of iPSCs reduced their viability. In the first phase of our studies, we used 1-6E6 AKAluc_iPSCs suspended in high concentration Matrigel and delivered cells sq to the flank of mice. **Fig. 12B** shows the mice injected with 1E6 iPSCs. However, even with 1E6 cells, the bioluminescence in AKALuc_iPSCs was much lower than what we observed in AKAluciferase expressing cancer cells. This was evident by the detection of the TokeOni background bioluminescence in the mid dorsal/ventral areas of the mice (**Fig. 12B**). Lower expression of Akaluciferase may impact our ability to detect rare iPSCs. In ongoing studies, we have begun to accumulate data on the teratoma growth profiles from our AKALuc_iPSC line (**Fig. 12B**).

In summary, we have characterized in greater detail the growth profiles of skin SCC cells both in sq and graft chamber environments. We documented the superficial growth of SCC cells in the graft chamber, which indicates that common human SCC cell lines may have a low propensity for metastasis in the nude mouse background. We have also established a bioluminescent model of skin metastasis using the C8161 cell line. Last, we began to characterize the teratoma growth profiles of human iPSCs.

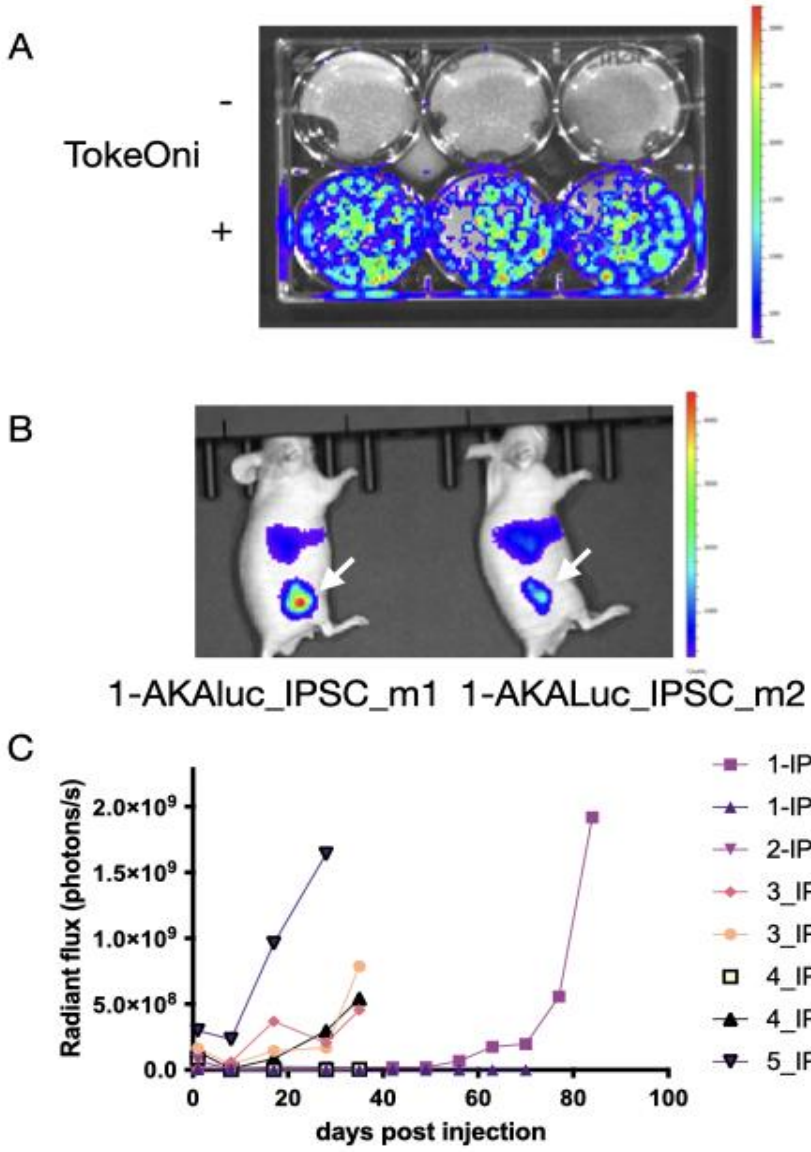


Fig. 12. Bioluminescent detection of iPSC-teratoma growth in nude mice. iPS cells were made to co-express stably the dT-Tomato and AKAluciferase proteins. A clonal colony was selected and expanded to be used in these experiments. (A) AKAluc_iPS cell colonies were imaged in 6 well plates using the IVIS Spectrum system in the presence or absence of TokeOni. (B) A mix of Matrigel and 1-6E6 AKAluc_iPS cells was injected sq into the flank of athymic nude mice. Mice were then imaged weekly. Mice depicted in the image received ~1E6 iPS cells/mouse. Note the detection of background TokeOni bioluminescence mid-dorsal/ventral just above the signal from AKAluc_iPS cells (arrows). (C) Graph shows the summary of ongoing experiments measuring the bioluminescence from developing iPSC teratomas in nude mice.

What opportunities for training and professional development has the project provided?

Nothing to Report

How were the results disseminated to communities of interest?

Nothing to Report

What do you plan to do during the next reporting period to accomplish the goals?

We will continue characterizing keratinocytes derived from genetically corrected RDEB iPSCs, including assessing our new *in vivo* grafting system. We will also differentiate our AKALuciferase positive iPSCs into keratinocytes to determine the persistence of iPSC-derived keratinocytes *in vivo* and their potential tumorigenicity using our new bioluminescent assay as described above. We will also continue working with our cGMP-compliant facility to perform pilot cGMP-compliant manufacturing of genetically corrected iPSCs and iPSC-derived keratinocytes as proposed in the original application.

4. **IMPACT:**

What was the impact on the development of the principal discipline(s) of the project?

We anticipate that our project will develop a stem-cell based therapy for the treatment of epidermolysis bullosa (EB), a group of rare inherited skin blistering diseases. EB derives from genetic mutations in structural proteins of the skin and sentences those afflicted to a life of severe pain and disability due to constant blistering and scarring. The development of a stem cell-based therapy is a complex process that needs to be reproducible and performed under clinically relevant standards. During this funding period, we continued to optimize a procedure for the generation of full thickness skin equivalents and developed a novel *in vivo* grafting assay to assess wound closure of genetically corrected RDEB iPSC-derived cells. We have also assessed the sensitivity of our *in vivo* bioluminescent assay using a mouse model in preparation for our iPSC-tumorigenicity studies. These accomplishments will bring us a step closer toward approval for a clinical trial to treat EB.

What was the impact on other disciplines?

Stem cell-based strategies similar to the one proposed in this application, whereby patient cells are genetically corrected and reprogrammed into immature induced pluripotent stem cells (iPSCs) that can be subsequently differentitated into target cell types for transplantation, can be applied to virtially any other currently incurable monogenic disease, including cystic fibrosis, Fanconi anemia, beta thalassemia, etc. However, unlike other monogenic diseases, EB, and especially RDEB, may represent an ideal platform to initially test an iPSC-based therapy due to the orphan nature of EB and its severity. The experiments using our *in vivo* bioluminescent tumor assay and the procedure for grafting genetically corrected RDEB iPSC-derived skin cells will provide important safety data that will be applicable for the use of iPSC-derived cells to regenerate other tissues and organs.

What was the impact on technology transfer?

Nothing to Report

What was the impact on society beyond science and technology?

Nothing to Report

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change

Nothing to Report

Actual or anticipated problems or delays and actions or plans to resolve them

We have obtained approval for a no cost extension from the DOD, which will allow us to mitigate all delays caused by the COVID-19 pandemic discussed in the previous reports.

Changes that had a significant impact on expenditures

Nothing to Report

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Significant changes in use or care of human subjects

Nothing to Report

Significant changes in use or care of vertebrate animals

Nothing to Report

Significant changes in use of biohazards and/or select agents

Nothing to Report

6. PRODUCTS:

- **Publications, conference papers, and presentations**

Report only the major publication(s) resulting from the work under this award.

Journal publications.

The following publication has been published with the acknowledgement of funding support from the Department of Defense (DOD) (W81XWH-18-1-0706) (see Appendix for the actual reprint):

Warshauer EM, Brown A, Fuentes I, Shortt J, Gignoux C, Montinaro F, Metspalu M, Youssefian L, Vahidnezhad H, Jacków J, Christiano AM, Uitto J, Fajardo-Ramírez ÓR, Salas-Alanis JC, McGrath JA, Consuegra L, Rivera C, Maier PA, Runfeldt G, Behar DM, Skorecki K, Sprecher E, Palisson F, Norris DA, Bruckner AL, Kogut I, Bilousova G, Roop DR (2021). Ancestral patterns of recessive dystrophic epidermolysis bullosa mutations in Hispanic populations suggest sephardic ancestry. *Am J Med Genet A.* 2021 Aug 26;. doi: 10.1002/ajmg.a.62456. [Epub ahead of print] PubMed PMID: 34435747.

Books or other non-periodical, one-time publications.

Nothing to Report

Other publications, conference papers and presentations.

Nothing to Report

- **Website(s) or other Internet site(s)**

Nothing to Report

- **Technologies or techniques**

Nothing to Report

- **Inventions, patent applications, and/or licenses**

Nothing to Report

- **Other Products**

Nothing to Report

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name:	*Dennis Roop*
Project Role:	*PI*
Nearest person month worked:	*0.6*
Contribution to Project:	*Dr. Roop oversees the project as a PI.*
Funding Support:	*National Institute of Health, EB Charities; institutional support, DEBRA International, Avita Medical, DOD*

Name: Ganna Bilousova
Project Role: Co-Investigator
Nearest person month worked: 3
Contribution to Project: Dr. Bilousova prepares regulatory compliance documents and oversees the work related to iPSC generation and differentiation.
Funding Support: National Institute of Health, EB Charities, institutional support, DEBRA International, DOD, Avita Medical

Name: Igor Kogut
Project Role: Co-Investigator
Nearest person month worked: 2.4
Contribution to Project: Dr. Kogut oversees the transfer of our technologies into a cGMP compliant facility.
Funding Support: National Institute of Health, EB Charities, institutional support, DEBRA International, DOD

Name: Andrii Rozhok
Project Role: Instructor
Nearest person month worked: 0.8
Contribution to Project: Mr. Rozhok develops a platform to analyze whole-genome sequencing data.
Funding Support: Institutional support, National Institute of Health, DOD

Name: Josiah Fernandes
Project Role: PRA
Nearest person month worked: 1.7
Contribution to Project: Mr. Fernandez assists Dr. Bilousova in the generation of 3D skin equivalents and differentiation of iPSCs. He also performs molecular analysis of iPSC-derived cells.
Funding Support: National Institute of Health, institutional support, DOD

Name: Chann Makara Han
Project Role: PRA
Nearest person month worked: 1.2
Contribution to Project: Mr. Han assists Drs. Kogut and Bilousova in analyzing iPSCs and detecting off-target events in the Cas9-mediated gene correction strategy. He also characterizes generated iPSCs.
Funding Support: National Institute of Health, institutional support, DEBRA International, DOD

Name: Maryna Pavlova
Project Role: RA
Nearest person month worked: 6
Contribution to Project: Dr. Pavlova works on the differentiation of iPSCs into keratinocytes. She also assists Dr. Bilousova in the generation of 3D skin equivalents and in vivo grafting.
Funding Support: DEBRA International, DOD

Name: Christopher Taylor
Project Role: PRA
Nearest person month worked: 4
Contribution to Project: Mr. Taylor assists Dr. Bilousova in mouse grafting experiments and maintenance of the mouse colony.
Funding Support: Institutional support, DOD

Name: *Velmurugan Balaiya*
Project Role: *RA*
Nearest person month worked: *1.6*
Contribution to Project: *Dr. Balaiya performs in vivo xenografting experiments.*
Funding Support: *Institutional support, Avita Medical, DOD*

Name: *Jocelyn Castillo Flores*
Project Role: *PRA*
Nearest person month worked: *12*
Contribution to Project: *Ms. Castillo Flores assist Dr. Pavlova in iPSC-differentiation.*
Funding Support: *DOD*

Name: *Sean Vieau*
Project Role: *PRA*
Nearest person month worked: *6*
Contribution to Project: *Mr. Vieau assists Drs. Pavlova and Balaiya in in vivo experiments.*
Funding Support: *DEBRA International, DOD*

Name: *Shennea McGarvey*
Project Role: *PRA*
Nearest person month worked: *0.6*
Contribution to Project: *Ms McGarvey differentiates iPSCs into fibroblasts and develops a cGMP-compliant protocol for the derivation of fibroblasts.*
Funding Support: *Institutional support, National Institute of Health, DOD*

Name: *Enrique Torchia*
Project Role: *RA*
Nearest person month worked: *10*
Contribution to Project: *Dr. Torchia develops in vivo tumorigenicity assays to determine the safety of iPSC-derived cells.*
Funding Support: *DOD*

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Dr. Kogut has received an NIH R01 award from NIAMS entitled "Developing a platform for human somatic cell rejuvenation, expansion and genetic engineering using synthetic RNA molecules". Dr. Bilousova is a Co-Investigator on this award. The major goal is to develop novel technology platforms for somatic cell expansion and genetic engineering using a rejuvenating RNA cocktail. There is no scientific overlap with the current DOD award.

Aims:

1. To characterize the effect of the rejuvenating RNA cocktail on the expansion level and molecular characteristics of somatic cells.
2. To assess the functionality of cells treated with the RNA cocktail.
3. To develop a platform for the generation of genetically modified human somatic cell lines using the rejuvenating cocktail.

Start and end date: 05/01/2021 - 04/30/2026
Level of effort: Dr. Kogut -30%; Dr. Bilousova – 15%
Contact: Alexey Belkin; alexey.belkin@nih.gov

The key personnel were able to accommodate this new grant without affecting the percent effort for the current award by modifying their other institution support.

What other organizations were involved as partners?

No changes.

8. **SPECIAL REPORTING REQUIREMENTS**

COLLABORATIVE AWARDS:

QUAD CHARTS:

Included with the report. The chart was modified to include the approved no cost extension.

9. **APPENDICES:**

A copy of the recent publication that acknowledges the current award is included.

Received: 24 May 2021 | Revised: 11 July 2021 | Accepted: 14 July 2021

DOI: 10.1002/ajmg.a.62456

ORIGINAL ARTICLE

AMERICAN JOURNAL OF medical genetics PART A WILEY

Ancestral patterns of recessive dystrophic epidermolysis bullosa mutations in Hispanic populations suggest sephardic ancestry

Emily Mira Warshauer[1,2] | Adam Brown[3] | Ignacia Fuentes[4,5] | Jonathan Shortt[6] | Chris Gignoux[6] | Francesco Montinaro[7,8] | Mait Metspalu[7] | Leila Youssefian[9] | Hassan Vahidnezhad[9] | Joanna Jacków[10,11] | Angela M. Christiano[10,12] | Jouni Uitto[9] | Óscar R. Fajardo-Ramírez[13,14] | Julio C. Salas-Alanis[13,15] | John A. McGrath[11] | Liliana Consuegra[16] | Carolina Rivera[16,17] | Paul A. Maier[18] | Goran Runfeldt[18] | Doron M. Behar[7,18] | Karl Skorecki[19] | Eli Sprecher[20,21] | Francis Palisson[5,22] | David A. Norris[1,2] | Anna L. Bruckner[1] | Igor Kogut[1,2] | Ganna Bilousova[1,2] | Dennis R. Roop[1,2]

[1]Department of Dermatology, University of Colorado School of Medicine, Anschutz Medical Campus, Aurora, Colorado, USA

[2]Charles C. Gates Center for Regenerative Medicine, University of Colorado School of Medicine, Anschutz Medical Campus, Aurora, Colorado, USA

[3]Avotaynu Research Partnership LLC, Englewood, New Jersey, USA

[4]Centro de Genética y Genómica, Facultad de Medicina Clínica Alemana, Universidad del Desarrollo, Santiago, Chile

[5]Fundación DEBRA Chile, Santiago, Chile

[6]Colorado Center for Personalized Medicine, University of Colorado Anschutz Medical Campus, Aurora, Colorado, USA

[7]Estonian Biocentre, Institute of Genomics, University of Tartu, Tartu, Estonia

[8]Department of Biology and Genetics, University of Bari, Bari, Italy

[9]Department of Dermatology and Cutaneous Biology, Sidney Kimmel Medical College and Jefferson Institute of Molecular Medicine, Thomas Jefferson University, Philadelphia, Pennsylvania, USA

[10]Department of Dermatology, Columbia University, New York, New York, USA

[11]St. John's Institute of Dermatology, King's College London (Guy's Campus), London, UK

[12]Department of Genetics and Development, Columbia University, New York, New York, USA

[13]DEBRA Mexico, Azteca Guadalupe, Mexico

[14]Tecnologico de Monterrey, Escuela de Medicina y Ciencias de la Salud, Monterrey, Mexico

[15]Instituto Dermatologico de Jalisco, Zapopan, Mexico

[16]Fundación DEBRA Colombia, Bogotá, Colombia

[17]Department of Medical Genetics, Pediatric Hospital, Fundacion Cardioinfantil-Universidad del Rosario, Bogotá, Colombia

[18]Gene by Gene, Genomic Research Center, Houston, Texas, USA

[19]Azrieli Faculty of Medicine of the Galilee, Bar-Ilan University, Safed, Israel

[20]Department of Dermatology, Tel-Aviv Sourasky Medical Center, Tel Aviv, Israel

[21]Department of Human Molecular Genetics, Sackler Faculty of Medicine, Tel-Aviv University, Tel Aviv, Israel

[22]Facultad de Medicina Clínica Alemana Universidad del Desarrollo, Santiago, Chile

Correspondence
Emily Mira Warshauer, Department of Dermatology, Center for Regenerative Medicine, University of Colorado School of Medicine, Anschutz Medical Campus, Aurora,

Abstract

Recessive dystrophic epidermolysis bullosa (RDEB) is a rare genodermatosis caused by mutations in the gene coding for type VII collagen (*COL7A1*). More than

CO, USA.
Email: emily.warshauer@ucdenver.edu

Funding information
Avotaynu Foundation; Cure EB; Dystrophic Epidermolysis Bullosa Research Assocation International; Epidermolysis Bullosa Medical Research Foundation; Epidermolysis Bullosa Research Partnership; Gates Frontiers Fund; National Institute of Arthritis and Musculoskeletal and Skin Diseases, Grant/Award Numbers: R01AR059947, U01AR075932; U.S. Department of Defense, Grant/Award Number: W81XWH-18-1-0706

800 different pathogenic mutations in *COL7A1* have been described to date; however, the ancestral origins of many of these mutations have not been precisely identified. In this study, 32 RDEB patient samples from the Southwestern United States, Mexico, Chile, and Colombia carrying common mutations in the *COL7A1* gene were investigated to determine the origins of these mutations and the extent to which shared ancestry contributes to disease prevalence. The results demonstrate both shared European and American origins of RDEB mutations in distinct populations in the Americas and suggest the influence of Sephardic ancestry in at least some RDEB mutations of European origins. Knowledge of ancestry and relatedness among RDEB patient populations will be crucial for the development of future clinical trials and the advancement of novel therapeutics.

KEYWORDS
epidermolysis bullosa, genetics, genodermatoses

1 | INTRODUCTION

Recessive dystrophic epidermolysis bullosa (RDEB) is a rare genodermatosis characterized by severe skin fragility and blistering resulting in chronic wounds with progressive fibrosis (Mittapalli et al., 2019), caused by mutations in the gene coding for type VII collagen (*COL7A1*) (Hovnanian et al., 1994). More than 800 different pathogenic mutations in *COL7A1* have been described to date (Stenson et al., 2017); however, the ancestral origins of these mutations have not been precisely identified. Furthermore, little knowledge exists of the descendants of isolated Sephardic communities on the Iberian Peninsula, populations that characteristically propagate recessive diseases such as RDEB (Nogueiro et al., 2015a; Nogueiro et al., 2015b). The successful identification of common origins in distinct RDEB populations will provide an important resource as new treatments involving gene editing approaches for RDEB and other severe genetic skin diseases become viable for early clinical trials. In this study, 32 Hispanic RDEB patient samples carrying common *COL7A1* gene mutations from the Southwestern United States, Mexico, Chile, and Colombia were genotyped to investigate the ancestral origins for *COL7A1* mutations and determine the extent to which shared ancestry contributes to RDEB in these populations.

2 | RESULTS

The 32 Hispanic RDEB patients included in this study from the Southwestern United States, Mexico, Chile, and Colombia (see Table 1) carry pathogenic *COL7A1* gene mutations that were genotyped with different sequencing techniques and confirmed by Sanger Sequencing.

To further investigate ancestral patterns in the Hispanic RDEB population, a principal component analysis (PCA) and admixture analysis were performed (Alexander et al., 2009) on genomewide genotyping data alongside a reference panel consisting of 500 individuals from African, Middle Eastern, and European ancestry (Chang et al., 2015). In the PCA, the majority of the individuals are scattered between European and Native American populations (Figure 1). For Admixture, we identified three different European components, evident in Figures 2 and S1, modal in Western Europe (blue), Middle East and the Caucasus (red), and the Arabian peninsula (green), supporting the presence of Sephardic ancestry. For the American component, we identified at $K = 8$, a high variation in ancestral components, with an almost complete overlap between ancestry and population, possibly reflecting isolation and inbreeding dynamics, well-known in Jewish populations. Furthermore, Admixture analysis shows similar European and Native American ancestral components in the majority of the RDEB patients, with a minor, but consistent contribution from African groups (pink in Figure 2). A non-negative least squares (NNLS) haplotype-based method also supported overriding European and Native American origins, yet the majority (84%) of individuals harbor a small component (>2%) of North African and Middle-Eastern ancestry (Figure 3, shown in green), which may result from interactions between North Africa or Jewish populations and Iberian populations. The evidence of North African and Middle-Eastern lineage substantiates, at least in part, the influence of Sephardic Jews in the Hispanic RDEB population.

To further explore the origin of these mutations, we estimated local ancestry across the *COL7A1* locus in all 32 RDEB patients using a reference panel consisting of African, European, and Native American populations (see Methods and Materials). Pathogenic mutations at the locus were determined separately from the local ancestry and are thus not phased with the local ancestry estimates, however most patients (25/32) had just one ancestry at both copies of the locus, allowing us to gain insights into the ancestral source of the 20 mutations present on these haplotypes (Table 2). Seventeen of the mutations appear in patients with a single ancestry at both copies of *COL7A1* (Table 2). Two additional mutations (c.7708delG and c.5532 + 1G > T) are present only in individuals with either a single ancestry at both copies of the locus or mixed ancestry at the locus, and may

TABLE 1 *COL7A1* mutations in Hispanic RDEB patients

FTDNA kit no.	Mutation 1	Mutation 2	Sephardic	Test site
MK40784	c.6527_6528insC	c.8329C > T	21%	Chile
730,451	c.2470insG	c.2470insG	18%	Mexico
MK40790	c.6527_6528insC	c.5856 + 1G > T	15%	Chile
MK40779	c.7485 + 1G > A	c.4635 + 5G > A	10%	Colombia
MK42684	c.2470insG	c.2470insG	9%	Mexico
730,456	c.2470insG	c.2470insG	7%	Mexico
662,380	c.7485 + 5G > A	c.7485 + 5G > A	8%	Colorado
MK40771	c.8200G > A	c.4965C > T	7%	Colombia
MK40789	c.7708delG	c.7876-1G > A	5%	Chile
662,381	c.7485 + 5G > A	c.7485 + 5G > A	4%	Colorado
MK40776	c.7651G > C	c.4012G > A	4%	Colombia
MK40782	c.6781C > T	c.2044C > T	<2%	Colombia
MK40777	c.1584G > T	c.1584G > T	—	Colombia
MK40787	c.2005C > T	c.4342-2A > G	—	Chile
MK42706	c.5108G > A	c.IVS23-1G > A	—	Mexico
730,454	c.2470insG	c.2470insG	—	Mexico
MK40794	c.2992 + 2 T > G	c.6527_6528insC	—	Chile
MK40795	c.3264_5293del	c.5532 + 1G > T	—	Chile
MK40785	c.3759 + 2 T > G	c.3759 + 2 T > G	—	Chile
MK40772	c.425A > G	c.8833 T > C	—	Colombia
MK40775	c.4510G > T	c.4510G > T	—	Colombia
MK40778	c.4678G > A	c.4678G > A	—	Colombia
MK40793	c.5532 + 1G > T	c.8245G > A	—	Chile
MK40792	c.5932C > T	c.5932C > T	—	Chile
MK40780	c.6091G > A	c.6091G > A	—	Colombia
MK40783	c.6527_6528insC	c.6527_6528insC	—	Chile
MK40788	c.6527_6528insC	c.8528-1G > A	—	Chile
MK40773	c.6527_6528insC	c.5604 + 1G > A	—	Colombia
MK40791	c.7708delG	c.7708delG	—	Chile
MK40786	c.7708delG	c.8393 T > G	—	Chile
MK40774	c.8046 + 6G > A	c.5047C > T	—	Colombia
MK42683	c.8709del11	c.G2899del11	—	Mexico

thus be attributed to a single ancestral population in our data. Notably, a single mutation, c.2470insG, appears to have originated independently in both ancestral European and Native American populations. In total, 10 of the mutations appear to have arisen on Native American haplotypes, 8 on European haplotypes, and 4 are observed only in individuals with mixed ancestry and cannot be determined.

In order to evaluate the temporal origins of *COL7A1* mutations in our dataset, we looked for shared haplotypes between patients and identified an enrichment of identity by descent (IBD) in the region surrounding the *COL7A1* gene on chromosome 3 (Figure 4), supporting recent shared ancestry among at least some of the patients, even though none are known to be closely related. From among the 32 RDEB patients, we identified three different IBD clusters, each representing a single shared haplotype, indicating recent shared ancestry between these individuals (Table 3).

The largest cluster is comprised of 4 individuals (2 from Chile, 1 from Colombia, and 1 from Mexico) (Table 3, Cluster A). The observation that ¾ of the patients carry at least one copy of the pathogenic RDEB mutation c.6527_6528insC mutation, with no other *COL7A1* mutation being shared by all individuals in the cluster, suggests that the haplotype shared by individuals in the cluster carries the c.6527_6528insC mutation. It is possible that the pathogenic mutation of the patient who does not carry the c.6527_6528insC mutation was incompletely characterized, or that the shared haplotype does not encompass the mutation. Local ancestry indicates that the shared haplotype is of European origin.

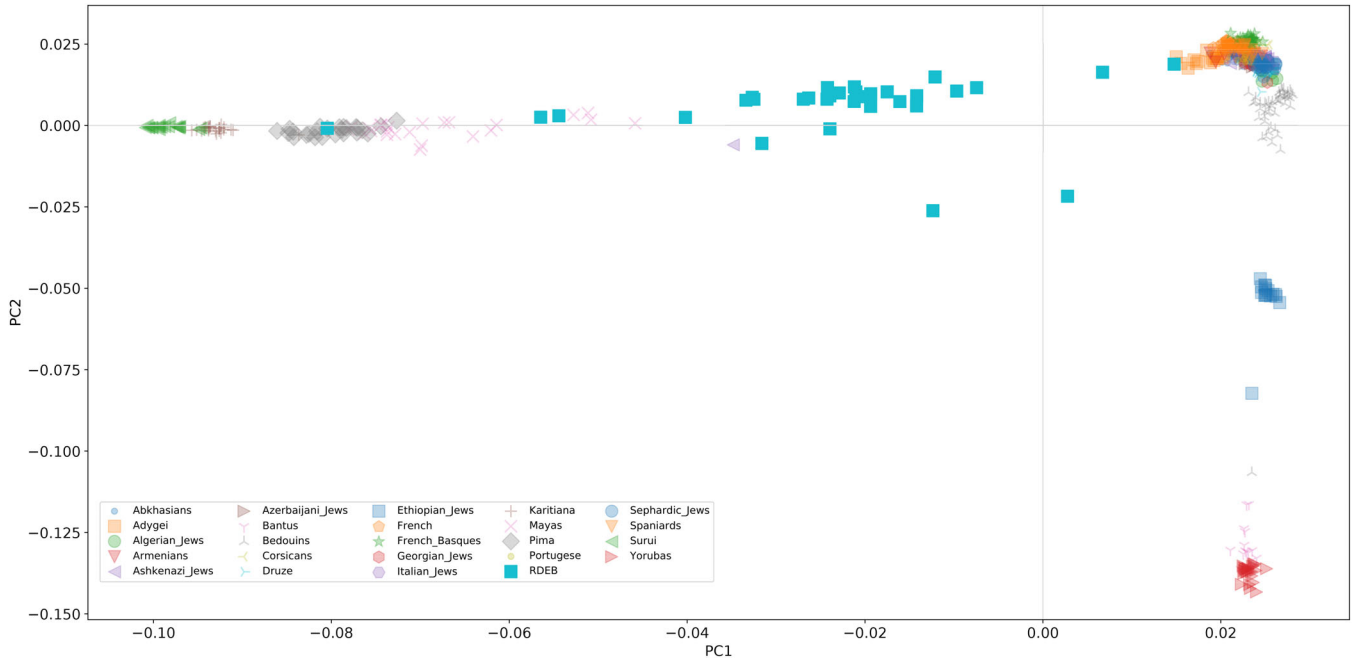


FIGURE 1 A Principal Component Analysis (PCA) was carried out on a dataset composed of 516 individuals in total, including 32 RDEB samples. The RDEB samples are marked with a teal blue square "■" and represent substantial admixture with Native American and European individuals, including Sephardic Jews represented with light blue circles in the right upper corner "●". The legend in the left lower corner denotes the color markers exhibited in the PCA analysis. The *X* and *Y* axis refer to Principal Component (PC) 1 and 2, respectively

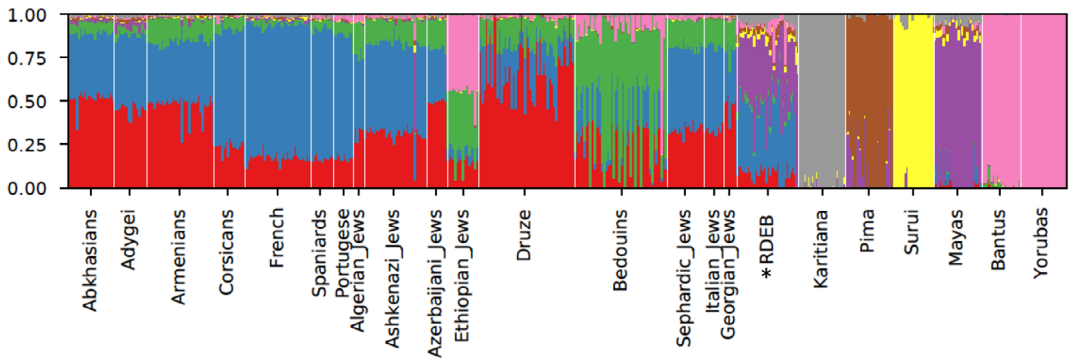


FIGURE 2 Admixture analysis. Population structure of 32 RDEB samples inferred by Admixture analysis. Each individual is represented by a vertical (100%) stacked column of genetic components proportions shown in color for K8 (*K* = number of subpopulations). The RDEB population is marked with an asterisk

A second IBD haplotype is shared by two RDEB patients from Mexico (Table 3, Cluster B). Both patients are homozygous for the c.2470insG mutation, and both are homozygous for the shared haplotype. Local ancestry across the *COL7A1* locus on the shared haplotype indicates that it is likely of European origin.

The third IBD haplotype is shared by two RDEB patients from Chile carrying the c.7708delG mutation (Table 3, Cluster C), one of whom is homozygous for the shared haplotype and the mutation. Native American local ancestry across the *COL7A1* locus in all patients suggests that this mutation arose in the Americas.

Multiple other mutations are observed multiple times in the RDEB patients without being part of an IBD cluster. We do not rule out the possibility that at least some of these mutations may have been inherited from a common ancestor, but recombination in intervening generations has interrupted the inherited haplotype to a size below detection thresholds. On the other hand, it is also reasonable to consider that at least some of these mutations may have arisen independently.

3 | DISCUSSION

Admixed populations present challenges in understanding the history of disease, particularly in regions where each population can have a

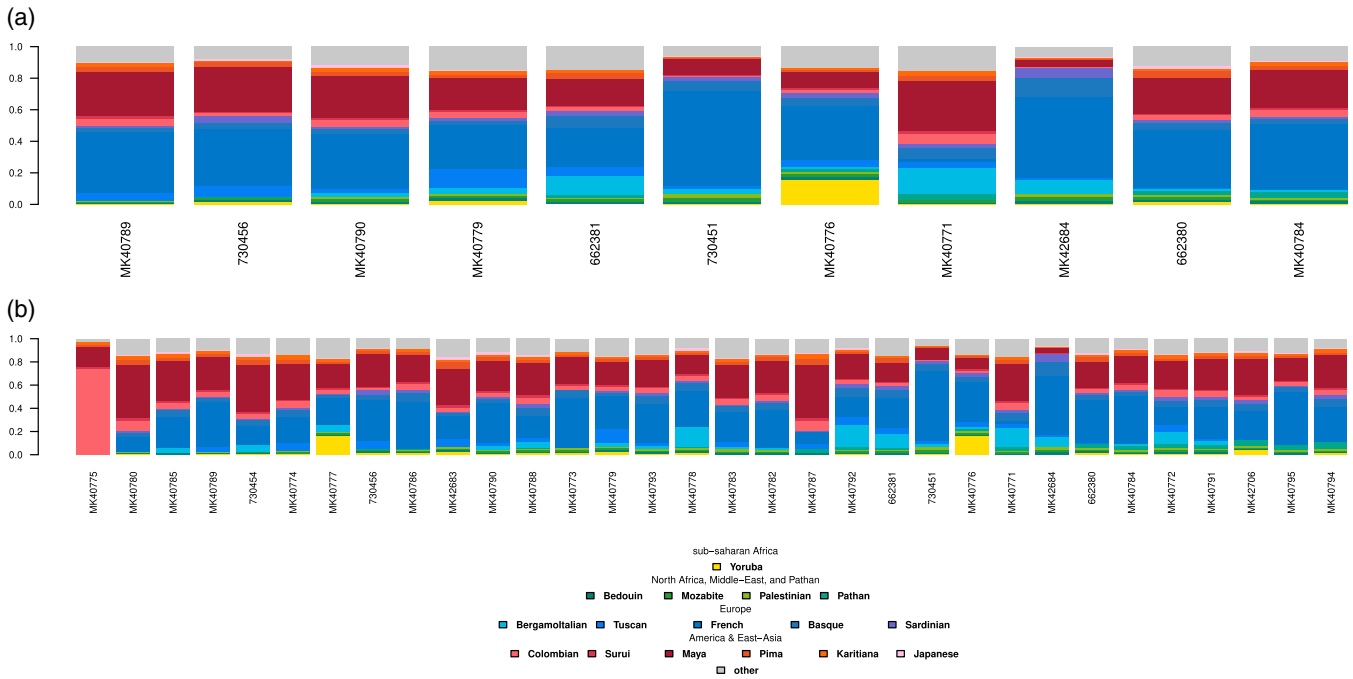


FIGURE 3 NNLS haplotype method. The NNLS Haplotype method exhibits a bar plot of the ancestral composition for a set of surrogate groups, as demonstrated in the legend. Implemented in CP and merged with HGDP sequences, each individual (recipient) was painted as a combination of genomic fragments inherited by all the HGDP panel (donors). Thirty-two Hispanic RDEB patient samples carrying common mutations in the *COL7A1* gene sorted according to Middle Eastern ancestry from the Southwestern United States, Mexico, Chile, and Colombia (a). Asterisks in (a) highlight the samples shown in (b). The analysis of RDEB samples with affinity to Sephardic origins sorted according to Middle Eastern ancestry demonstrate North African and Middle Eastern origins, highlighted in green, effectively supporting hidden genetic ancestry in heavily admixed individuals (b)

different admixture history. The high frequency of shared haplotypes harboring pathogenic RDEB mutations among unrelated patients in our data suggests that founder or other demographic events may have increased the prevalence of these pathogenic haplotypes in Hispanic populations in the Americas. We also present evidence that RDEB mutations in multiple Hispanic populations have ancestral origins from both European and Native American ancestral populations, with at least one mutation (c.2470insG) having origins from both ancestral populations. We note that because the number of observations of each mutation in our dataset is limited by our patient sampling, it is possible that other mutations presented here also have other origins that simply are not sampled in our patients. Thus, we cannot exclude the possibility that these other mutations have independent origins in other ancestral populations.

Eleven of the RDEB patients presented here had substantial amounts (>4%) of Sephardic ancestry using FamilyTreeDNA ancestry testing, a finding confirmed by a significantly higher proportion of European and American ancestry in these patients compared to the remainder of the RDEB patients (Figure 5). The finding is not entirely unexpected as it is known that Sephardic individuals arrived in the Americas in the late 15th century, and Sephardic ancestry has been previously detected in some Hispanic populations, with some Sephardic mutations being observed in former Spanish colonies such as the Southwestern US, Mexico, and El Salvador (Ellis et al., 1998; Ostrer, 2016; Shahrabani-Gargir et al., 1998; Struewing et al., 1995; Velez et al., 2012). Notably, significant admixture also occurred between North African non-Jewish populations and Sephardic Jews following their expulsion from Spain during the Inquisition (Aksentijevich et al., 1993; Campbell et al., 2012), which may explain the ancestral

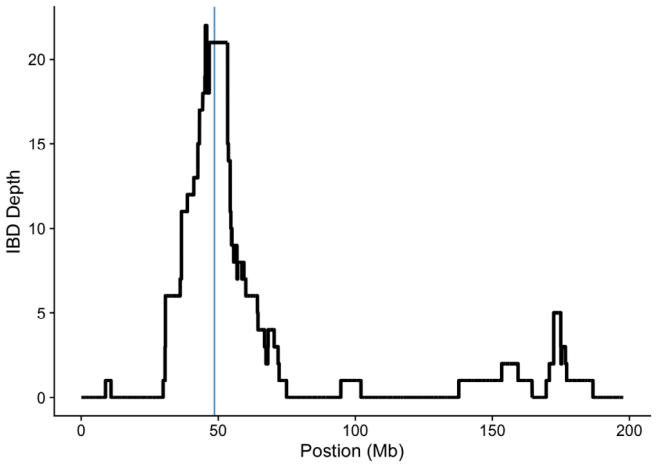


FIGURE 4 Identity by descent (IBD). Cumulative depth of identical by descent (IBD) segments from all pairwise sample comparisons across chromosome 3 (blue vertical line indicates the location of *COL7A1*)

TABLE 2 Counts of local ancestry at COL7A in RDEB patients for each mutation

	Counts of local ancestry at COL7A in RDEB patients for each mutation		
Mutation	**AMR**	**EUR**	**Mixed**
c.7485 + 5G > A	4	0	0
c.7708delG	3	0	1
c.2470insG	2	6	0
c.4510G > T	2	0	0
c.1584G > T	2	0	0
c.6091G > A	2	0	0
c.5108G > A	1	0	0
c.8200G > A	1	0	0
c.2005C > T	1	0	0
c.IVS23-1G > A	1	0	0
c.4965C > T	1	0	0
c.6527_6528insC	0	6	1
c.4678G > A	0	2	0
c.3759 + 2 T > G	0	2	0
c.5932C >	0	2	0
c.7651G > C	0	1	0
c.2992 + 2 T > G	0	1	0
c.3264_5293del	0	1	0
c.5532 + 1G > T	0	1	1
c.8709del11	0	1	0
c.425A > G	0	0	1
c.8046 + 6G > A	0	0	1
c.7485 + 1G > A	0	0	1
c.6781C > T	0	0	1

components in the NNLS related to African and Middle Eastern populations (Figures 2 and 3).

The recognition that the European haplotype shared by individuals carrying the c.6527_6528insC mutation represents patients from three separate populations could suggest that the arrival of the mutation on the haplotype predates European settlement. This knowledge would be consistent with estimates of the first occurrence of the c.6527_6528insC mutation more than 3000 years ago when pre-Roman communities settled in the Iberian Peninsula (Sanchez-Jimeno et al., 2013). Interestingly, this region was also home to a closed endogamous community of Sephardic Jews during a time period more than a millennium ago, coinciding with the estimated origin of the c.6527_6528insC founder mutation (Adams et al., 2008). The c.6527_6528insC mutation remains frequent on the Iberian Peninsula of Spain and it is possible that at least some Hispanic populations inherited this mutation through Hispanic or Sephardic migration. Data from the original source of Sephardic Jews on the Iberian Peninsula is limited and the question of Sephardic origins is challenging as Sephardic and Spanish ancestry signals are likely to be largely overlapping due to centuries of cohabitation (Álvarez-Álvarez et al., 2018; Nogueiro et al., 2015b).

From the analyses presented here, the observation that at least some RDEB mutations appear to have European origins lends some support to the hypothesis that some of the pathogenic RDEB mutations presented here are of Sephardic origin. Population frequencies represented in gnomAD indicate slight enrichment with European populations in eight Hispanic RDEB mutations (Table S1), providing rationale for further investigation of Sephardic ancestry in these variants (Collins et al., 2020; Karczewski et al., 2020; Landrum et al., 2020). Interestingly, there are a few rare variants demonstrating enrichment with African populations, which may represent the Sephardic populations of North Africa (Gonçalves et al., 2005). Establishing more confident and specific inferences of the ancestral origins of the mutations present in the Hispanic RDEB patients will benefit from additional sampling from ancient and modern-day Sephardic populations as well as from the populations from which the RDEB samples were collected.

Gene-editing treatments for RDEB patients using guide RNAs specific to pathogenic mutations are anticipated in the near future (Bonafont et al., 2019; Mencía et al., 2018). As regulatory agencies may consider each guide RNA as a separate drug, separate clinical trials for each RNA would be required. Patient populations who share the same founder mutations therefore represent an increasingly important resource to facilitate early clinical trials and advance novel treatments.

TABLE 3 Counts of RDEB mutations and local ancestry populations within identity by descent COL7A1 haplotypes

IBD cluster ID	Count of patients in cluster	Count of patient haplotype zygosity for cluster (#homozygous/#heterozygous)	Count of patient mutation Zygosity (#homozygous/#heterozygous)							Local ancestry counts at *COL7A1* in cluster			
			c.6527_6528insC	c.5604 + 1G > A	c.8528-1G > A	c.8709del11	c. G2899del11	c.2470insG	c.7708delG	c.7876-1G > A	AMR	EUR	Mixed
A	4	0/4	1/2	0/1	0/1	0/1	0/1	0/0	0/0	0/0	0	3	1
B	2	2/0	0/0	0/0	0/0	0/0	0/0	2/0	0/0	0/0	0	2	0
C	2	1/1	0/0	0/0	0/0	0/0	0/0	0/0	1/1	0/1	2	0	0

4 | MATERIALS AND METHODS

Thirty-two RDEB homozygous and compound heterozygous patient samples from the Southwestern United States, Mexico, Chile, and Colombia with common mutations in *COL7A1* included in this study were previously identified using different sequencing technologies and subsequently confirmed by Sanger sequencing (Table 1). Informed written consent was obtained from all patients in concordance with Institutional Review Board approval from the USA: Colorado Multiple Institutional Review Board (COMIRB no: 09-0192), Mexico: Universidad de Monterrey (132012-CE), Chile: Comité Ético Científico, Facultad de Medicina, Clinica Alemana—Universidad del Desarrollo (Project number 2013-145), and Colombia: Universidad del Rosario (CIE-UR DVO005 1149-CV1192).

4.1 | Illumina bead chip array and FTDNA family finder analysis

DNA samples were genotyped by GeneByGene, Inc. utilizing the Illumina Human OmniExpress BeadChip array and analyzed with the FTDNA Family Finder autosomal DNA test. The Family Finder test returns results for about 690,000 pairs of single nucleotide polymorphisms (SNPs) on the 22 pairs of autosomal chromosomes. Autosomal SNPs are clustered into sets about 50–100 SNPs long that are predefined based on the reliability, variability, average centiMorgans (cM), and density of the SNPs. The Family Finder software then evaluates SNP sets for matching as half identical or a non-match based on an autosomal DNA algorithm demonstrating shared DNA segments. Adjacent SNP sets are also analyzed to see if they qualify as identical by descent (IBD) segments. A segment is considered a candidate autosomal match if it contains at least 500 SNPs, and it is at least 1 cM long. Then, proprietary rules based on total shared cM and longest segment cM are applied to infer whether the match is valid.

The RDEB DNA samples were then subset to approximately 245,000 unlinked SNPs and run through ADMIXTURE using FamilyTreeDNA's proprietary global reference panel called myOrigins2. The 24 global population clusters represent modern human genetic variation and results reflect admixture between historical gene pools. They include Sephardic, Ashkenazi, North and Central America, South America, British Isles, Scandinavia, Finland, West and Central Europe, Southeast Europe, East Europe, Iberia, West Middle East, East Middle East, Asia Minor, North Africa, East Central Africa, South Central Africa, West Africa, Central Asia, South Central Asia, Siberia, Northeast Asia, Southeast Asia, and Oceania. Affinity to Sephardic origins is determined significant by at least a 4% match to the Sephardic cluster (Chacón-Duque et al., 2018).

4.2 | Identity by descent analysis

We merged the genotyping data from RDEB patients with genotypes from a combined reference dataset consisting of individuals from the

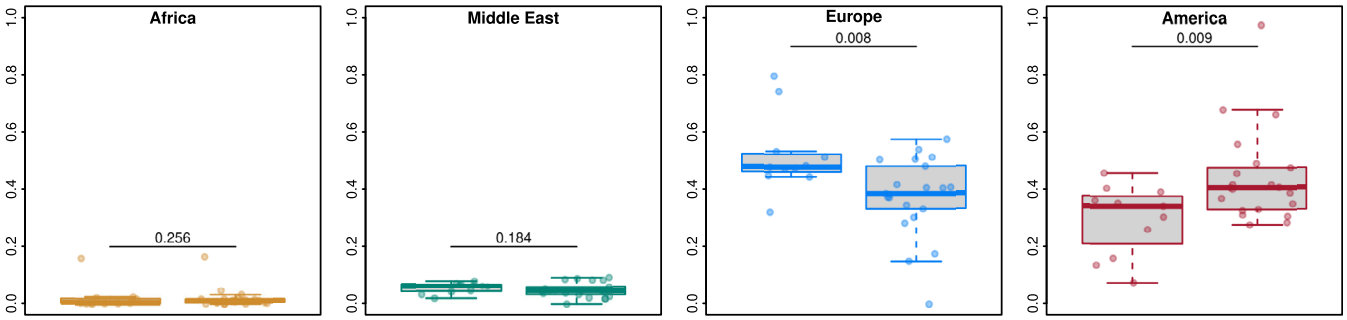


FIGURE 5 Ancestries proportions. This figure represents the ancestry proportions for Africa, the Middle East, Europe and America. Within each panel, the left hand side bar graph represents RDEB samples with Sephardic ancestry and the right hand side bar graph represents RDEB samples without Sephardic ancestry. The 11 RDEB samples with Sephardic ancestry clearly demonstrate significantly greater European ancestry and less American ancestry than RDEB samples without Sephardic ancestry. There are no significant differences noted among the African or Middle Eastern ancestry

Human Genome Diversity Project (HGDP) (Cann et al., 2002) and 1000 Genomes (Genomes Project et al., 2010; Genomes Project et al., 2012; Genomes Project et al., 2015), and phased the entire merged dataset using EAGLE2 (Loh et al., 2016) with default parameters. Phased haplotypes were then fed through iLASH to detect identical by descent (IBD) regions greater than or equal to 3 cM between pairs of RDEB patients. Cumulative IBD depth at each marker was calculated using a custom python script that counts the total number of times each locus is part of an IBD segment shared between two individuals. We used DASH (Gusev et al., 2011) with a minimum cluster size of 3 and a minimum haplotype length of 3 cM to identify clusters of patients that share the same IBD haplotypes across the entirety of *COL7A1*.

4.3 | Local ancestry analysis

We used phased haplotypes (see "Identity by Descent Analysis") from RDEB patients in RFMix (Maples et al., 2013) (version 1.5.4) with a reference panel consisting of European and nonadmixed African populations (ESN, GWD, MSL, LWK, YRI) from 1000 Genomes and 107 indigenous American individuals from HGDP to train the random forests. The estimates for local ancestry presented in the article are generated from the RFMix Viterbi output.

4.4 | PCA and admixture analysis

In order to explore the ancestry of the analyzed individuals, an exploratative analysis was carried out using a PCA and Admixture approach (Alexander et al., 2009). We merged the genotype data with publicly available genome-wide datasets (Behar et al., 2010; Behar et al., 2013; Kushniarevich et al., 2015; Ongaro et al., 2019; Tambets et al., 2018; Tamm et al., 2019; Yunusbayev et al., 2012; Yunusbayev et al., 2015) using PLINK 1.9, a widely used program for research in population genetics (Chang et al., 2015). After merging, SNPs and individuals characterized by less than 3% and 5% of missing data were retained for a total of 87,181 SNPs, and 516 individuals. A PCA was carried out using the PCA flag in PLINK1.9 (Chang et al., 2015).

The Admixture analysis was carried out on the dataset with random seed and cross-validation based on data resampling. Although the most supported configuration is $K = 9$, we are showing $K = 8$, for which European and the Middle East ancestries show a reasonable degree of differentiation (Figure S1).

4.5 | CP and NNLS

The ancestry composition of the analyzed individuals was explored utilizing haplotype method implemented in ChromoPainter (CP, Lawson et al., 2012). The analyzed samples were merged with the HGDP sequences in order to increase the density of the data (Bergström et al., 2019). First, the sequence data SNPs coordinates were lifted from Ghch38 to hg19 using Picard Tools (Picard, 2020). The merged dataset was subsequently phased using shapeit ver 2. using default settings (Delaneau et al., 2011). Second, each individual (recipient) was painted as a combination of genomic fragments inherited by all the HGDP panel (donors), an effective way to unlock the hidden genetic ancestry in heavily admixed individuals (Montinaro et al., 2015; Ongaro et al., 2019). The nuisance parameters were set as $M = 0.0018$ and $n = 409$, estimated based on expectation maximization run in a subset of five individuals from each population.

The resulting copying vectors were subsequently normalized and each RDEB patient individual ancestry was reconstructed as a combination of copying vectors from all the populations in the HGDP dataset, using a modified version of the NNLS function of R software implemented in GlobeTrotter (Hellenthal et al., 2014; Leslie et al., 2015). The results for the RDEB patients are summarized in Figure 4. Populations contributing on average less than 1% were labeled as others.

ACKNOWLEDGMENTS

We thank all the patients and their families for helping us to carry out this study. This study was supported by the Avotaynu Foundation,

Epidermolysis Bullosa Research Partnership, Epidermolysis Bullosa Medical Research Foundation, Cure EB, National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) of the National Institute of Health (NIH) (R01AR059947 and U01AR075932), the Department of Defense (DOD) (W81XWH-18-1-0706), Dystrophic Epidermolysis Bullosa Research Association (DEBRA) International, and the Gates Frontiers Fund. We also want to give special thanks to Stephen Berman, MD, the Founding Director of the Epidermolysis Bullosa Center of Excellence at Children's Hospital Colorado, who has cared for Hispanic RDEB patients in Colorado for over three decades and originally suggested to us that these patients may be of Converso ancestry.

CONFLICT OF INTEREST

The authors state no conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization: Emily Mira Warshauer, Adam Brown, Ignacia Fuentes, Karl Skorecki, Eli Sprecher, Dennis R. Roop. Formal Analysis: Emily Mira Warshauer, Adam Brown, Ignacia Fuentes, Jonathan Shortt, Chris Gignoux, Francesco Montinaro, Mait Metspalu, Leila Youssefian, Hassan Vahidnezhad, Joanna Jacków, Angela M. Christiano, Jouni Uitto, Óscar R. Fajardo-Ramírez, Julio C. Salas-Alanis, John A. McGrath, Liliana Consuegra, Carolina Rivera, Paul A. Maier, Goran Runfeldt, Doron M. Behar, Karl Skorecki, Eli Sprecher, Francis Palisson, David A. Norris, Adam Brown, Igor Kogut, Ganna Bilousova, Dennis R. Roop. Writing—original draft: Emily Mira Warshauer, Ignacia Fuentes, Jonathan Shortt, Chris Gignoux, Francesco Montinaro, Dennis R. Roop. Writing—review and editing: Emily Mira Warshauer, Adam Brown, Ignacia Fuentes, Jonathan Shortt, Chris Gignoux, Francesco Montinaro, Paul A. Maier, Goran Runfeldt, Karl Skorecki, Dennis R. Roop.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Emily Mira Warshauer https://orcid.org/0000-0002-4055-5196
Hassan Vahidnezhad https://orcid.org/0000-0003-4298-9147
Jouni Uitto https://orcid.org/0000-0003-4639-807X

REFERENCES

Adams, S. M., Bosch, E., Balaresque, P. L., Ballereau, S. J., Lee, A. C., Arroyo, E., López-Parra, A. M., Aler, M., Gisbert Grifo, M. S., Brion, M., Carracedo, A., Lavinha, J., Martínez-Jarreta, B., Quintana-Murci, L., Picornell, A., Ramon, M., Skorecki, K., Behar, D. M., Calafell, F., & Jobling, M. A. (2008). The genetic legacy of religious diversity and intolerance: Paternal lineages of Christians, Jews, and Muslims in the Iberian Peninsula. *American Journal of Human Genetics*, *83*(6), 725–736. https://doi.org/10.1016/j.ajhg.2008.11.007

Aksentijevich, I., Pras, E., Gruberg, L., Shen, Y., Holman, K., Helling, S., Prosen, L., Sutherland, G. R., Richards, R. I., & Dean, M. (1993). Familial Mediterranean fever (FMF) in Moroccan Jews: Demonstration of a founder effect by extended haplotype analysis. *American Journal of Human Genetics*, *53*(3), 644–651.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Álvarez-Álvarez, M. M., Risch, N., Gignoux, C. R., Huntsman, S., Ziv, E., Fejerman, L., Esteban, M. E., Gayà-Vidal, M., Sobrino, B., Brisighelli, F., Harich, N., Cruciani, F., Chaabani, H., Carracedo, Á., Moral, P., Burchard, E. G., Via, M., & Athanasiadis, G. (2018). Genetic analysis of Sephardic ancestry in the Iberian Peninsula. *bioRxiv*, 325779. https://doi.org/10.1101/325779

Behar, D. M., Metspalu, M., Baran, Y., Kopelman, N. M., Yunusbayev, B., Gladstein, A., Tzur, S., Sahakyan, H., Bahmanimehr, A., Yepiskoposyan, L., Tambets, K., Khusnutdinova, E. K., Kusniarevich, A., Balanovsky, O., Balanovsky, E., Kovacevic, L., Marjanovic, D., Mihailov, E., Kouvatsi, A., … Rosenberg, N. A. (2013). No evidence from genome-wide data of a Khazar origin for the Ashkenazi Jews. *Human Biology*, *85*(6), 859–900. https://doi.org/10.3378/027.085.0604

Behar, D. M., Yunusbayev, B., Metspalu, M., Metspalu, E., Rosset, S., Parik, J., Rootsi, S., Chaubey, G., Kutuev, I., Yudkovsky, G., Khusnutdinova, E. K., Balanovsky, O., Semino, O., Pereira, L., Comas, D., Gurwitz, D., Bonne-Tamir, B., Parfitt, T., Hammer, M. F., … Villems, R. (2010). The genome-wide structure of the Jewish people. *Nature*, *466*(7303), 238–242. https://doi.org/10.1038/nature09103

Bergström, A., McCarthy, S. A., Hui, R., Almarri, M. A., Ayub, Q., Danecek, P., Chen, Y., Felkel, S., Hallast, P., Kamm, J., Blanché, H., Deleuze, J.-F., Cann, H., Mallick, S., Reich, D., Sandhu, M. S., Skoglund, P., Scally, A., Xue, Y., … Tyler-Smith, C. (2019). Insights into human genetic variation and population history from 929 diverse genomes. *bioRxiv*, 674986. https://doi.org/10.1101/674986

Bonafont, J., Mencía, Á., García, M., Torres, R., Rodríguez, S., Carretero, M., Chacón-Solano, E., Modamio-Høybjør, S., Marinas, L., León, C., Escamez, M. J., Hausser, I., del Río, M., Murillas, R., & Larcher, F. (2019). Clinically relevant correction of recessive dystrophic epidermolysis bullosa by dual sgRNA CRISPR/Cas9-mediated gene editing. *Molecular Therapy*, *27*(5), 986–998. https://doi.org/10.1016/j.ymthe.2019.03.007

Campbell, C. L., Palamara, P. F., Dubrovsky, M., Botigué, L. R., Fellous, M., Atzmon, G., Oddoux, C., Pearlman, A., Hao, L., Henn, B. M., Burns, E., Bustamante, C. D., Comas, D., Friedman, E., Pe'er, I., & Ostrer, H. (2012). North African Jewish and non-Jewish populations form distinctive, orthogonal clusters. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(34), 13865–13870. https://doi.org/10.1073/pnas.1204840109

Cann, H. M., de Toma, C., Cazes, L., Legrand, M. F., Morel, V., Piouffre, L., Cavalli-Sforza, L. L., Bodmer, W. F., Bonne-Tamir, B., Cambon-Thomsen, A., Chen, Z., Chu, J., Carcassi, C., Contu, L., du, R., Excoffier, L., Ferrara, G. B., Friedlaender, J. S., Groot, H., … Cavalli-Sforza, L. L. (2002). A human genome diversity cell line panel. *Science*, *296*(5566), 261–262. https://doi.org/10.1126/science.296.5566.261b

Chacón-Duque, J. C., Adhikari, K., Fuentes-Guajardo, M., Mendoza-Revilla, J., Acuña-Alonzo, V., Barquera, R., Quinto-Sánchez, M., Gómez-Valdés, J., Everardo Martínez, P., Villamil-Ramírez, H., Hünemeier, T., Ramallo, V., Silva de Cerqueira, C. C., Hurtado, M., Villegas, V., Granja, V., Villena, M., Vásquez, R., Llop, E., … Ruiz-Linares, A. (2018). Latin Americans show wide-spread Converso ancestry and imprint of local native ancestry on physical appearance. *Nature Communications*, *9*(1), 5388. https://doi.org/10.1038/s41467-018-07748-z

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience*, *4*, 7. https://doi.org/10.1186/s13742-015-0047-8

Collins, R. L., Brand, H., Karczewski, K. J., Zhao, X., Alföldi, J., Francioli, L. C., Khera, A. V., Lowther, C., Gauthier, L. D., Wang, H., Watts, N. A., Solomonson, M., O'Donnell-Luria, A., Baumann, A., Munshi, R., Walker, M., Whelan, C. W., Huang, Y., Brookings, T., … Talkowski, M. E. (2020). A structural variation reference for medical and population genetics. *Nature*, *581*(7809), 444–451. https://doi.org/10.1038/s41586-020-2287-8

Delaneau, O., Marchini, J., & Zagury, J. F. (2011). A linear complexity phasing method for thousands of genomes. *Nature Methods*, *9*(2), 179–181. https://doi.org/10.1038/nmeth.1785

Ellis, N. A., Ciocci, S., Proytcheva, M., Lennon, D., Groden, J., & German, J. (1998). The Ashkenazic Jewish bloom syndrome mutation blmAsh is present in non-Jewish Americans of Spanish ancestry. *American Journal of Human Genetics*, *63*(6), 1685–1693. https://doi.org/10.1086/302167

1000 Genomes Project Consortium, Abecasis, G. R., Altshuler, D., Auton, A., Brooks, L. D., Durbin, R. M., Gibbs, R. A., Hurles, M. E., & McVean, G. A. (2010). A map of human genome variation from population-scale sequencing. *Nature*, *467*(7319), 1061–1073. https://doi.org/10.1038/nature09534

1000 Genomes Project Consortium, Abecasis, G. R., Auton, A., Brooks, L. D., DePristo, M. A., Durbin, R. M., Handsaker, R. E., Kang, H. M., Marth, G. T., & McVean, G. A. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature*, *491*(7422), 56–65. https://doi.org/10.1038/nature11632

1000 Genomes Project Consortium, Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G., & Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393

Gonçalves, R., Freitas, A., Branco, M., Rosa, A., Fernandes, A. T., Zhivotovsky, L. A., Underhill, P. A., Kivisild, T., & Brehm, A. (2005). Y-chromosome lineages from Portugal, Madeira and Açores record elements of Sephardim and Berber ancestry. *Annals of Human Genetics*, *69*(Pt 4), 443–454. https://doi.org/10.1111/j.1529-8817.2005.00161.x

Gusev, A., Kenny, E. E., Lowe, J. K., Salit, J., Saxena, R., Kathiresan, S., Altshuler, D. M., Friedman, J. M., Breslow, J. L., & Pe'er, I. (2011). DASH: A method for identical-by-descent haplotype mapping uncovers association with recent variation. *American Journal of Human Genetics*, *88*(6), 706–717. https://doi.org/10.1016/j.ajhg.2011.04.023

Hellenthal, G., Busby, G. B. J., Band, G., Wilson, J. F., Capelli, C., Falush, D., & Myers, S. (2014). A genetic atlas of human admixture history. *Science*, *343*(6172), 747–751. https://doi.org/10.1126/science.1243518

Hovnanian, A., Hilal, L., Blanchet-Bardon, C., de Prost, Y., Christiano, A. M., Uitto, J., & Goossens, M. (1994). Recurrent nonsense mutations within the type VII collagen gene in patients with severe recessive dystrophic epidermolysis bullosa. *American Journal of Human Genetics*, *55*(2), 289–296. Retrieved from. https://www.ncbi.nlm.nih.gov/pubmed/8037207

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., MacArthur, D., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., … Daly, M. J. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, *581*(7809), 434–443. https://doi.org/10.1038/s41586-020-2308-7

Kushniarevich, A., Utevska, O., Chuhryaeva, M., Agdzhoyan, A., Dibirova, K., Uktveryte, I., Möls, M., Mulahasanovic, L., Pshenichnov, A., Frolova, S., Shanko, A., Metspalu, E., Reidla, M., Tambets, K., Tamm, E., Koshel, S., Zaporozhchenko, V., Atramentova, L., Kučinskas, V., … Balanovsky, O. (2015). Genetic heritage of the Balto-Slavic speaking populations: A synthesis of autosomal, mitochondrial and Y-chromosomal data. *PLoS One*, *10*(9), e0135820. https://doi.org/10.1371/journal.pone.0135820

Landrum, M. J., Chitipiralla, S., Brown, G. R., Chen, C., Gu, B., Hart, J., Hoffman, D., Jang, W., Kaur, K., Liu, C., Lyoshin, V., Maddipatla, Z., Maiti, R., Mitchell, J., O'Leary, N., Riley, G. R., Shi, W., Zhou, G., Schneider, V., … Kattman, B. L. (2020). ClinVar: Improvements to accessing data. *Nucleic Acids Research*, *48*(D1), D835–d844. https://doi.org/10.1093/nar/gkz972

Lawson, D. J., Hellenthal, G., Myers, S., & Falush, D. (2012). Inference of population structure using dense haplotype data. *PLoS Genetics*, *8*(1), e1002453. https://doi.org/10.1371/journal.pgen.1002453

Leslie, S., Winney, B., Hellenthal, G., Davison, D., Boumertit, A., Day, T., Hutnik, K., Royrvik, E. C., Cunliffe, B., Wellcome Trust Case Control Consortium 2, International Multiple Sclerosis Genetics Consortium, Lawson, D. J., Falush, D., Freeman, C., Pirinen, M., Myers, S., Robinson, M., Donnelly, P., & Bodmer, W. (2015). The fine-scale genetic structure of the British population. *Nature*, *519*(7543), 309–314. https://doi.org/10.1038/nature14230

Loh, P. R., Danecek, P., Palamara, P. F., Fuchsberger, C., A Reshef, Y., K Finucane, H., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G. R., Durbin, R., & L Price, A. (2016). Reference-based phasing using the haplotype reference consortium panel. *Nature Genetics*, *48*(11), 1443–1448. https://doi.org/10.1038/ng.3679

Maples, B. K., Gravel, S., Kenny, E. E., & Bustamante, C. D. (2013). RFMix: A discriminative modeling approach for rapid and robust local-ancestry inference. *American Journal of Human Genetics*, *93*(2), 278–288. https://doi.org/10.1016/j.ajhg.2013.06.020

Mencía, Á., Chamorro, C., Bonafont, J., Duarte, B., Holguin, A., Illera, N., Llames, S. G., Escámez, M. J., Hausser, I., del Río, M., Larcher, F., & Murillas, R. (2018). Deletion of a pathogenic mutation-containing exon of COL7A1 allows clonal gene editing correction of RDEB patient epidermal stem cells. *Molecular Therapy—Nucleic Acids*, *11*, 68–78. https://doi.org/10.1016/j.omtn.2018.01.009

Mittapalli, V. R., Kühl, T., Kuzet, S. E., Gretzmeier, C., Kiritsi, D., Gaggioli, C., Bruckner-Tuderman, L., & Nystrom, A. (2019). STAT3 targeting in dystrophic epidermolysis bullosa. *British Journal of Dermatology*. https://doi.org/10.1111/bjd.18639

Montinaro, F., Busby, G. B., Pascali, V. L., Myers, S., Hellenthal, G., & Capelli, C. (2015). Unravelling the hidden ancestry of American admixed populations. *Nature Communications*, *6*, 6596. https://doi.org/10.1038/ncomms7596

Nogueiro, I., Teixeira, J., Amorim, A., Gusmao, L., & Alvarez, L. (2015a). Echoes from Sepharad: Signatures on the maternal gene pool of crypto-Jewish descendants. *European Journal of Human Genetics*, *23*(5), 693–699. https://doi.org/10.1038/ejhg.2014.140

Nogueiro, I., Teixeira, J. C., Amorim, A., Gusmao, L., & Alvarez, L. (2015b). Portuguese crypto-Jews: The genetic heritage of a complex history. *Frontiers in Genetics*, *6*, 12. https://doi.org/10.3389/fgene.2015.00012

Ongaro, L., Scliar, M. O., Flores, R., Raveane, A., Marnetto, D., Sarno, S., Gnecchi-Ruscone, G. A., Alarcón-Riquelme, M. E., Patin, E., Wangkumhang, P., Hellenthal, G., Gonzalez-Santos, M., King, R. J., Kouvatsi, A., Balanovsky, O., Balanovska, E., Atramentova, L., Turdikulova, S., Mastana, S., … Montinaro, F. (2019). The genomic impact of European colonization of the Americas. *Current Biology*, *29*(23), 3974–3986.e4. https://doi.org/10.1016/j.cub.2019.09.076

Ostrer, H. (2016). The origin of the p.E180 growth hormone receptor gene mutation. *Growth Hormone and IGF Research*, *28*, 51–52. https://doi.org/10.1016/j.ghir.2015.08.003

Picard. (2020). Retrieved from http://broadinstitute.github.io/picard/

Sanchez-Jimeno, C., Cuadrado-Corrales, N., Aller, E., Garcia, M., Escamez, M. J., Illera, N., Trujillo-Tiebas, M. J., Ayuso, C., Millán, J. M., & Del Rio, M. (2013). Recessive dystrophic epidermolysis bullosa: The origin of the c.6527insC mutation in the Spanish population. *British*

Journal of Dermatology, *168*(1), 226–229. https://doi.org/10.1111/j.1365-2133.2012.11128.x

Shahrabani-Gargir, L., Shomrat, R., Yaron, Y., Orr-Urtreger, A., Groden, J., & Legum, C. (1998). High frequency of a common bloom syndrome Ashkenazi mutation among Jews of polish origin. *Genetic Testing*, *2*(4), 293–296. https://doi.org/10.1089/gte.1998.2.293

Stenson, P. D., Mort, M., Ball, E. V., Evans, K., Hayden, M., Heywood, S., Hussain, M., Phillips, A. D., & Cooper, D. N. (2017). The human gene mutation database: Towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Human Genetics*, *136*(6), 665–677. https://doi.org/10.1007/s00439-017-1779-6

Struewing, J. P., Abeliovich, D., Peretz, T., Avishai, N., Kaback, M. M., Collins, F. S., & Brody, L. C. (1995). The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. *Nature Genetics*, *11*(2), 198–200. https://doi.org/10.1038/ng1095-198

Tambets, K., Yunusbayev, B., Hudjashov, G., Ilumäe, A. M., Rootsi, S., Honkola, T., Vesakoski, O., Atkinson, Q., Skoglund, P., Kushniarevich, A., Litvinov, S., Reidla, M., Metspalu, E., Saag, L., Rantanen, T., Karmin, M., Parik, J., Zhadanov, S. I., Gubina, M., … Metspalu, M. (2018). Genes reveal traces of common recent demographic history for most of the Uralic-speaking populations. *Genome Biology*, *19*(1), 139. https://doi.org/10.1186/s13059-018-1522-1

Tamm, E., di Cristofaro, J., Mazières, S., Pennarun, E., Kushniarevich, A., Raveane, A., Semino, O., Chiaroni, J., Pereira, L., Metspalu, M., & Montinaro, F. (2019). Publisher correction: Genome-wide analysis of Corsican population reveals a close affinity with northern and Central Italy. *Scientific Reports*, *9*(1), 18827. https://doi.org/10.1038/s41598-019-55185-9

Velez, C., Palamara, P. F., Guevara-Aguirre, J., Hao, L., Karafet, T., Guevara-Aguirre, M., Pearlman, A., Oddoux, C., Hammer, M., Burns, E., Pe'er, I., Atzmon, G., & Ostrer, H. (2012). The impact of Converso Jews on the genomes of modern Latin Americans. *Human Genetics*, *131*(2), 251–263. https://doi.org/10.1007/s00439-011-1072-z

Yunusbayev, B., Metspalu, M., Järve, M., Kutuev, I., Rootsi, S., Metspalu, E., Behar, D. M., Varendi, K., Sahakyan, H., Khusainova, R., Yepiskoposyan, L., Khusnutdinova, E. K., Underhill, P. A., Kivisild, T., & Villems, R. (2012). The Caucasus as an asymmetric semipermeable barrier to ancient human migrations. *Molecular Biology and Evolution*, *29*(1), 359–365. https://doi.org/10.1093/molbev/msr221

Yunusbayev, B., Metspalu, M., Metspalu, E., Valeev, A., Litvinov, S., Valiev, R., Akhmetova, V., Balanovska, E., Balanovsky, O., Turdikulova, S., Dalimova, D., Nymadawa, P., Bahmanimehr, A., Sahakyan, H., Tambets, K., Fedorova, S., Barashkov, N., Khidiyatova, I., Mihailov, E., … Villems, R. (2015). The genetic legacy of the expansion of Turkic-speaking nomads across Eurasia. *PLoS Genetics*, *11*(4), e1005068. https://doi.org/10.1371/journal.pgen.1005068

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Warshauer, E. M., Brown, A., Fuentes, I., Shortt, J., Gignoux, C., Montinaro, F., Metspalu, M., Youssefian, L., Vahidnezhad, H., Jacków, J., Christiano, A. M., Uitto, J., Fajardo-Ramírez, Ó. R., Salas-Alanis, J. C., McGrath, J. A., Consuegra, L., Rivera, C., Maier, P. A., Runfeldt, G., … Roop, D. R. (2021). Ancestral patterns of recessive dystrophic epidermolysis bullosa mutations in Hispanic populations suggest sephardic ancestry. *American Journal of Medical Genetics Part A*, 1–11. https://doi.org/10.1002/ajmg.a.62456