HOKKAIDO UNIVERSITY

Title	Cas9-guided haplotyping of three truncation variants in autosomal recessive disease
Author(s)	Natsuga, Ken; Furuta, Yoshikazu; Takashima, Shota et al.
Citation	Human mutation, 43(7), 877-881 https://doi.org/10.1002/humu.24385
Issue Date	2022-07
Doc URL	https://hdl.handle.net/2115/90104
Rights	This is the peer reviewed version of the following article: Natsuga, K., Furuta, Y., Takashima, S., Nohara, T., Huang, H.-Y., Shinkuma, S., Nakamura, H., Katsuda, Y., Higashi, H., Hsu, C.-K., Fukushima, S., & Ujiie, H. (2022). Cas9-guided haplotyping of three truncation variants in autosomal recessive disease. Human Mutation, 43, 877- 881, which has been published in final form at 10.1002/humu.24385. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions. This article may not be enhanced, enriched or otherwise transformed into a derivative work, without express permission from Wiley or by statutory rights under applicable legislation. Copyright notices must not be removed, obscured or modified. The article must be linked to Wiley's version of record on Wiley Online Library and any embedding, framing or otherwise making available the article or pages thereof by third parties from platforms, services and websites other than Wiley Online Library must be prohibited.
Type	journal article
File Information	Hum Mutat 43 877-881.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Cas9-guided haplotyping of three truncation variants in autosomal recessive disease

Ken Natsuga[1*], Yoshikazu Furuta[2], Shota Takashima[1], Takuma Nohara[1], Hsin-Yu Huang[3], Satoru Shinkuma[4], Hideki Nakamura[1], Yousuke Katsuda[5], Hideaki Higashi[2], Chao-Kai Hsu[3,6,7], Satoshi Fukushima[8], Hideyuki Ujiie[1]

[1]Department of Dermatology, Faculty of Medicine and Graduate School of Medicine, Hokkaido University, Sapporo, Japan

[2]Division of Infection and Immunity, International Institute for Zoonosis Control, Hokkaido University, Sapporo, Japan

[3]Institute of Clinical Medicine, College of Medicine, National Cheng Kung University, Tainan, Taiwan

[4]Department of Dermatology, Nara Medical University School of Medicine, Kashihara, Japan

[5]Division of Materials Science and Chemistry, Faculty of Advanced Science and Technology, Kumamoto University, Kumamoto, Japan

[6]Department of Dermatology, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan

[7]International Center for Wound Repair and Regeneration (iWRR), National Cheng Kung University, Tainan, Taiwan

[8]Department of Dermatology and Plastic Surgery, Faculty of Life Sciences, Kumamoto University, Kumamoto, Japan

***Correspondence**:

Ken Natsuga, M.D., Ph.D.

Department of Dermatology, Faculty of Medicine and Graduate School of Medicine, Hokkaido University, N15 W7, Sapporo 0608638, Japan

Tel: +81-117067387

E-mail: natsuga@med.hokudai.ac.jp

ABSTRACT

An autosomal recessive disease is caused by biallelic loss-of-function mutations. However, when more than two disease-causing variants are found in a patient's gene, it is challenging to determine which two of the variants are responsible for the disease phenotype. Here, to decipher the pathogenic variants by precise haplotyping, we applied nanopore Cas9-targeted sequencing (nCATS) to three truncation *COL7A1* variants detected in a patient with recessive dystrophic epidermolysis bullosa (EB). The distance between the most 5' and 3' variants was approximately 19 kb at the level of genomic DNA. nCATS successfully demonstrated that the most 5' and 3' variants were located in one allele while the variant in between was located in the other allele. Interestingly, the proband's mother, who was phenotypically intact, was heterozygous for the allele that harbored the two truncation variants, which could otherwise be misinterpreted as those of typical recessive dystrophic EB. Our study highlights the usefulness of nCATS as a tool to determine haplotypes of complicated genetic cases. Haplotyping of multiple variants in a gene can determine which variant should be therapeutically targeted when nucleotide-specific gene therapy is applied.

Keywords

Haplotyping, nanopore sequencing, epidermolysis bullosa, CRISPR/Cas9

Biallelic mutations are required for the development of autosomal recessive diseases. Generally, each disease-causing variant is present in each allele. In contrast, it is exceedingly rare to detect three or more variants in a causative gene in a patient. In such a case, determining which two variants are disease-causing has proved challenging because of the inability to discern haplotypes using conventional sequencing technologies.

Single-molecule third-generation sequencing systems (e.g., the nanopore sequencers of Oxford Nanopore Technologies and the PacBio sequencers) can be pivotal in this field. These sequencers typically generate long to ultra-long reads. A combination of CRISPR/Cas9 technology and nanopore sequencing has been further shown to enable enrichment and sequencing of specific regions in genomic DNA (gDNA) (Gilpatrick et al., 2020) without PCR procedures or reverse transcription procedures, which can induce artificial homologous recombination (Meyerhans et al., 1990; Natsuga et al., 2022; Negroni et al., 1995). This method is called nanopore Cas9-targeted sequencing (nCATS) and has been utilized to detect somatic mutations and chromosomal rearrangements (Gilpatrick et al., 2020; Natsuga et al., 2022; Stangl et al., 2020; Watson et al., 2020; Wongsurawat et al., 2020).

Here, by taking advantage of three truncation *COL7A1* variants found in a recessive dystrophic epidermolysis bullosa (EB) patient, we show that nCATS can delineate the haplotypes of multiple variants, thus deciphering the disease-causing ones.

The proband was born to non-consanguineous parents and suffered from EB, a congenital skin fragility disorder (Has et al., 2020), as indicated by skin erosions and blisters at birth (**Figure 1a**). His EB subtype was classified as recessive dystrophic EB (Has et al., 2020), based on the following factors: (1) the absence of anchoring fibrils on electron microscopy (**Figure 1b**); (2) the absence of type VII collagen (COL7), which is the main component of anchoring fibrils (Watanabe et al., 2018), at the dermo-epidermal junction; and (3) skin-split at the level beneath laminin 332 (L332) and type IV collagen (COL4), which are the components of the basement membrane (lamina densa) (**Figure 1c**).

Whole-exome sequencing and subsequent Sanger sequencing of the proband's gDNA revealed three heterozygous truncation variants in *COL7A1* (NM000094.4), which encodes COL7 (**Figure 1d, Supp. Figure S1**). All variants were expected to lead to premature termination codons (PTCs). c.1474_1505del (p.Glu492TrpfsTer46) and c.2778_2779del

(p.Ala927AspfsTer26) were novel *COL7A1* variants, while c.6781C>T (p.Arg2261Ter; dbSNP, rs772381373) was previously described in RDEB cases (Kern et al., 2009; Kern et al., 2006; Sato-Matsumura et al., 2003).

The mutational profile of the proband indicated the three possible diplotypes, in which two of the three variants were present in one allele and the remaining variant was present in the other allele (diplotype-1, diplotype-2, and diplotype-3 in **Figure 2a**). We reasoned that nCATS (Gilpatrick et al., 2020) on the proband's gDNA would precisely identify the haplotypes allocated by the variants. We designed three nCATS experiments (nCATS-1, nCATS-2, and nCATS-3 in **Figure 2a**), each of which spanned either two of the three or all the variants. The length of the reads obtained by each nCATS experiment was 4 kb, 15 kb, and 19 kb, respectively. The coverage of the pooled samples was 87k (0.9 Gb). nCATS-1 primarily detected two kinds of reads, each harboring c.1475_1505del or c.2778_2779del (**Figure 2b**, **Supp. Table S1**). Most of the reads of nCATS-2, covering c.2778_2779del and c.6781C>T, also had one of the two variants, but not both (**Figure 2c**, **Supp. Table S2**). In line with these results, nCATS-3 identified the reads with both c.1474_1505del and c.6781C>T. The other major reads harbored only c.2778_2779del (**Figure 2d**,

Supp. Table S3). These results demonstrated that the most 5' and 3' variants were present in one allele, and the variant in between was present in the other allele (Proband, **Figure 2e**).

To determine whether the variants of the proband were inherited from the parents, who were phenotypically intact, or developed *de novo*, we performed Sanger sequencing on the parents' gDNA. The father was heterozygous for c.2778_2779del while the mother was heterozygous for c.1474_1505del and c.6781C>T (**Figure 1d, Supp. Figure S1**). We further confirmed that the mother's c.1474_1505del and c.6781C>T were in the same allele by nCATS-3 (coverage: 56k (0.5 Gb)) (**Figure 2d**, **Supp. Table S4**). These data revealed that, of the three truncation variants of the proband, one was inherited from the father while the other two came from the mother (**Figure 2e**).

Our data suggest that c.1474_1505del and c.2778_2779del are mainly pathogenic for recessive dystrophic EB in the proband because c.1474_1505del is located upstream of c.6781C>T in the same allele. This diplotype information also dictates that, when nucleotide-specific gene therapy (e.g., gene editing) is applied, c.2778_2779del should be targeted. By contrast, correction of one of c.1474_1505del or c.6781C>T would be insufficient because the other variant

would still be present in the same allele. Notably, the mother harbored two heterozygous truncation mutations. Moreover, her Sanger sequencing results could have been interpreted as those of typical recessive dystrophic EB if haplotyping by nCATS had not been performed.

Although regular genetic testing might not always need nCATS, this method can be an option for haplotyping somatic mutations (Natsuga et al., 2022), haplotyping three or more variants in one patient like our study, and identifying repeat expansions (Mizuguchi et al., 2021; Sone et al., 2019). One nCATS experiment costs hundreds of USD, mostly for a MinION flow cell. Flongle, a smaller flow cell, might reduce the cost if its use is verified for nCATS in the future.

In our study, only nCATS-3 was ultimately sufficient to determine the proband's diplotype. However, as the reads of nCATS-3 were much fewer than those of nCATS-1 and nCATS-2 (**Supp. Table S1-S4**), nCATS-1 and nCATS-2 were confirmatory for this analysis. Furthermore, in scenarios where the most 5' and 3' variants are far from each other (e.g., > 50 kb), the combination of nCATS experiments (like nCATS-1 and nCATS-2 in our study) will be needed.

In summary, our study has shown that nCATS is a promising tool for deciphering pathogenic mutations among multiple variants by precise haplotyping. Thus, the haplotypes identified by nCATS can help clinicians design the optimal gene therapy for patients.

Data Availability

All sequence data produced by MinION sequencers and haplotype count data were deposited to Gene Expression Omnibus (GEO) with accession no. GSE196673. During the review process, the data at GEO is accessible with the following private token: sxwtmseizxidben.

Conflict of Interests

None to declare.

Acknowledgements

We thank Ms. Mika Tanabe for her technical assistance. This work was supported by the World-leading Innovative and Smart Education (WISE) Program (1801) from the Ministry of Education, Culture, Sports, Science, and Technology, Japan, and AMED (ID: 21ak0101168h0001).

References

Gilpatrick, T., Lee, I., Graham, J. E., Raimondeau, E., Bowen, R., Heron, A., Downs, B., Sukumar, S., Sedlazeck, F. J., & Timp, W. (2020). Targeted nanopore sequencing with Cas9-guided adapter ligation. *Nat Biotechnol*, *38*(4), 433-438. https://doi.org/10.1038/s41587-020-0407-5

Has, C., Bauer, J. W., Bodemer, C., Bolling, M. C., Bruckner-Tuderman, L., Diem, A., Fine, J. D., Heagerty, A., Hovnanian, A., Marinkovich, M. P., Martinez, A. E., McGrath, J. A., Moss, C., Murrell, D. F., Palisson, F., Schwieger-Briel, A., Sprecher, E., Tamai, K., Uitto, J., . . . Mellerio, J. E. (2020). Consensus reclassification of inherited epidermolysis bullosa and other disorders with skin fragility. *Br J Dermatol*, *183*(4), 614-627. https://doi.org/10.1111/bjd.18921

Kern, J. S., Gruninger, G., Imsak, R., Muller, M. L., Schumann, H., Kiritsi, D., Emmert, S., Borozdin, W., Kohlhase, J., Bruckner-Tuderman, L., & Has, C. (2009). Forty-two novel COL7A1 mutations and the role of a frequent single nucleotide polymorphism in the MMP1 promoter in modulation of disease severity in a large European dystrophic epidermolysis bullosa cohort. *Br J Dermatol*, *161*(5), 1089-1097. https://doi.org/10.1111/j.1365-

2133.2009.09333.x

Kern, J. S., Kohlhase, J., Bruckner-Tuderman, L., & Has, C. (2006). Expanding the COL7A1 mutation database: novel and recurrent mutations and unusual genotype-phenotype constellations in 41 patients with dystrophic epidermolysis bullosa. *J Invest Dermatol*, *126*(5), 1006-1012. https://doi.org/10.1038/sj.jid.5700219

Meyerhans, A., Vartanian, J. P., & Wain-Hobson, S. (1990). DNA recombination during PCR. *Nucleic Acids Res*, *18*(7), 1687-1691. https://doi.org/10.1093/nar/18.7.1687

Mizuguchi, T., Toyota, T., Miyatake, S., Mitsuhashi, S., Doi, H., Kudo, Y., Kishida, H., Hayashi, N., Tsuburaya, R. S., Kinoshita, M., Fukuyama, T., Fukuda, H., Koshimizu, E., Tsuchida, N., Uchiyama, Y., Fujita, A., Takata, A., Miyake, N., Kato, M., . . . Matsumoto, N. (2021). Complete sequencing of expanded SAMD12 repeats by long-read sequencing and Cas9-mediated enrichment. *Brain*, *144*(4), 1103-1117. https://doi.org/10.1093/brain/awab021

Natsuga, K., Furuta, Y., Takashima, S., Nohara, T., Kosumi, H., Mai, Y., Higashi, H., & Ujiie, H. (2022). Detection of revertant mosaicism in epidermolysis

bullosa through Cas9-targeted long-read sequencing. *Hum Mutat*, *43*(4), 529-536. https://doi.org/10.1002/humu.24331

Negroni, M., Ricchetti, M., Nouvel, P., & Buc, H. (1995). Homologous recombination promoted by reverse transcriptase during copying of two distinct RNA templates. *Proc Natl Acad Sci U S A*, *92*(15), 6971-6975. https://doi.org/10.1073/pnas.92.15.6971

Sato-Matsumura, K. C., Sawamura, D., Goto, M., Goto, M., Nakamura, H., & Shimizu, H. (2003). A novel insertion mutation in COL7A1 identified in Hallopeau-Siemens recessive dystrophic epidermolysis bullosa. *Acta Derm Venereol*, *83*(2), 137-138. https://doi.org/10.1080/00015550310007526

Sone, J., Mitsuhashi, S., Fujita, A., Mizuguchi, T., Hamanaka, K., Mori, K., Koike, H., Hashiguchi, A., Takashima, H., Sugiyama, H., Kohno, Y., Takiyama, Y., Maeda, K., Doi, H., Koyano, S., Takeuchi, H., Kawamoto, M., Kohara, N., Ando, T., . . . Sobue, G. (2019). Long-read sequencing identifies GGC repeat expansions in NOTCH2NLC associated with neuronal intranuclear inclusion disease. *Nat Genet*, *51*(8), 1215-1221. https://doi.org/10.1038/s41588-019-0459-y

Stangl, C., de Blank, S., Renkens, I., Westera, L., Verbeek, T., Valle-Inclan, J. E., González, R. C., Henssen, A. G., van Roosmalen, M. J., Stam, R. W., Voest, E. E., Kloosterman, W. P., van Haaften, G., & Monroe, G. R. (2020). Partner independent fusion gene detection by multiplexed CRISPR-Cas9 enrichment and long read nanopore sequencing. *Nature Communications*, *11*(1), 2861. https://doi.org/10.1038/s41467-020-16641-7

Watanabe, M., Natsuga, K., Shinkuma, S., & Shimizu, H. (2018). Epidermal aspects of type VII collagen: Implications for dystrophic epidermolysis bullosa and epidermolysis bullosa acquisita. *J Dermatol*, *45*(5), 515-521. https://doi.org/10.1111/1346-8138.14222

Watson, C. M., Crinnion, L. A., Hewitt, S., Bates, J., Robinson, R., Carr, I. M., Sheridan, E., Adlard, J., & Bonthron, D. T. (2020). Cas9-based enrichment and single-molecule sequencing for precise characterization of genomic duplications. *Laboratory Investigation*, *100*(1), 135-146. https://doi.org/10.1038/s41374-019-0283-0

Wongsurawat, T., Jenjaroenpun, P., De Loose, A., Alkam, D., Ussery, D. W., Nookaew, I., Leung, Y.-K., Ho, S.-M., Day, J. D., & Rodriguez, A. (2020). A novel Cas9-targeted long-read assay for simultaneous detection of IDH1/2

mutations and clinically relevant MGMT methylation in fresh biopsies of diffuse glioma. *Acta Neuropathologica Communications*, *8*(1), 87. https://doi.org/10.1186/s40478-020-00963-0

Figure 1. Proband with three truncation variants in *COL7A1*.

(a) Clinical manifestations of the proband. Erosions are indicated by arrowheads. (b) Electron micrographs of the proband's skin, showing the absence of anchoring fibrils (dotted arrows). Scale bar: 500 nm. (c) Immunofluorescence studies of the proband's skin specimens. Skin-splits are indicated by asterisks. The absence of type VII collagen is indicated by dotted arrows. Scale bar: 50 μm. (d) The variant profiles of the proband and parents. The variant nomenclature is based on NM_000094.4.

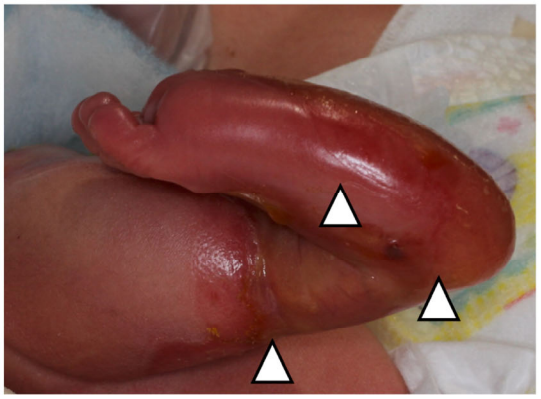
Figure 2. Haplotyping through nCATS.

(a) Diagram of the nCATS strategy. Three possible diplotypes of the proband are depicted (diplotype 1–3). (b–d) Frequency of the reads corresponding to the haplotypes. Each pie slice represents a haplotype as indicated in the figures. The results of nCATS-1 (b), nCATS-2 (c), and nCATS-3 (d) are shown. The raw data are summarized in **Supp. Tables S1–S4**. (e) Family tree with uncovered diplotypes of the proband and parents.

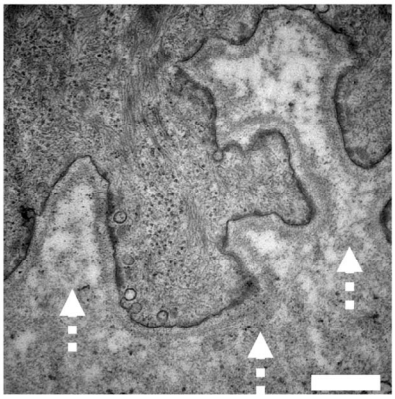
Figure 1. Proband with three truncation variants in *COL7A1*

a

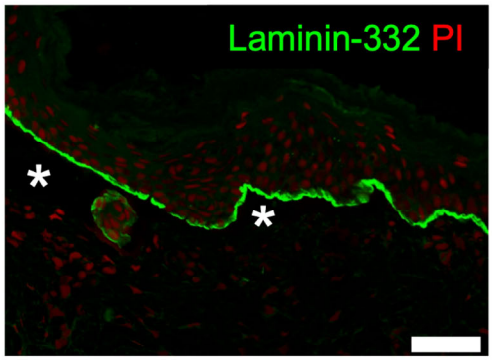


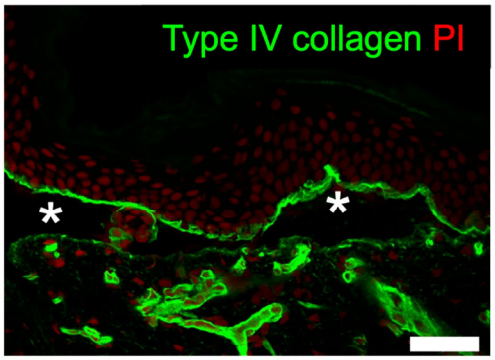


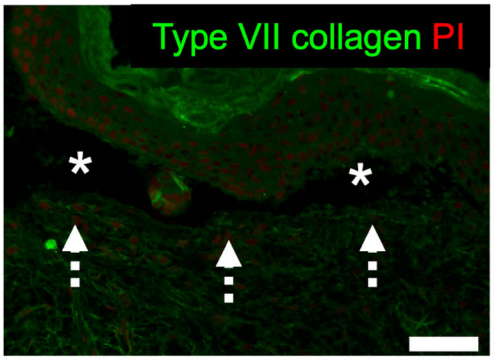
b



c







d

COL7A1	c.1474_1505del (p.Glu492TrpfsTer46)	c.2778_2779del (p.Ala927AspfsTer26)	c.6781C>T (p.Arg2261Ter)
Proband	**Heterozygous**	**Heterozygous**	**Heterozygous**
Father	Wild type	**Heterozygous**	Wild type
Mother	**Heterozygous**	Wild type	**Heterozygous**

Figure 2. Haplotyping through nCATS

