Title: Exome sequencing in a consanguineous Pakistani family identifies a mutational hotspot in the *COL7A1* gene, causing recessive dystrophic epidermolysis bullosa

Authors:

Atta Ur Rehman[1] MPhil, Virginie G. Peter[1] MSc, Mathieu Quinodoz[1] MSc, Muhammad Dawood[2] MBBS, Carlo Rivolta[1,3,4,5] PhD

1. Department of Computational Biology, Unit of Medical Genetics, University of Lausanne, Switzerland
2. Department of Dermatology, District Headquarter Hospital Khar, Bajaur, Khyber Pakhtunkhwa, Pakistan
3. Department of Genetics and Genome Biology, University of Leicester, Leicester, United Kingdom
4. Clinical Research Center, Institute of Molecular and Clinical Ophthalmology Basel (IOB)
5. Department of Ophthalmology, University Hospital Basel, Basel, Switzerland

Corresponding Author's Email: carlo.rivolta@unil.ch

Conflict of interest statement: The authors have no conflict of interest to declare.

Exome sequencing in a consanguineous Pakistani family identifies a mutational hotspot in the *COL7A1* gene, causing recessive dystrophic epidermolysis bullosa

Dear Editor,

Epidermolysis bullosa (EB) is a rare inherited blistering condition of the skin that affects approximately 1 in 120,000 people [1]. EB is heterogeneous both clinically and genetically. To date, there are more than 30 known clinical variants of EB, with casual mutations reported in as many as 18 genes [2]. With a prevalence of 1 in ~350,000 individuals, dystrophic epidermolysis bullosa (DEB) is one of the major forms of EB [1]. DEB may be inherited as a recessive (RDEB, MIM 226600) or dominant (DDEB, MIM 131750) disease, each further classified into different clinical subtypes [3].

All major forms of DEB, whether dominant or recessive, are attributable to mutations in the *COL7A1* gene, which encodes type VII collagen [4]. This gene is relatively large (~32 kb, 118 exons) and, according to the *COL7A1* gene variants database, it has been found to carry more than 800 individual mutations (http://www.col7a1-database.info/ Accessed 2018-12-20). While the majority of DEB-associated mutations are family-specific, some common mutations have also been identified [5].

Here, we present a consanguineous Pakistani family with four affected individuals suffering from RDEB. The patients were 3 to 12 years old, all born to consanguineous and healthy parents (Fig. 1a). The proband is the first child of double first cousins. She presented with severe skin fragility and multiple skin injuries at birth, followed by mechanically-induced generalized blistering of the skin. Blisters with atrophic scarring were, however, more frequent on hands, feet, knees, elbows, and flexural areas (Fig. 2c and d). Occasionally, the proband experienced chronic wounds at the blistering

sites that lasted for more than 6 months. Anonychia, mild dental decay and milia were also observed and tissue granulation was restricted to the abdomen. At the time of examination (12 years of age), the proband was suffering from anemia and growth retardation. Blistering of the oral mucosa was a common feature that caused severe swallowing difficulties, especially in her early childhood (Fig. 2a). Progressive contractures of hands and feet were noticed. Excessive blistering, contractures, and pseudosyndactylism of toes resulted in a 'mitten-feet' (Fig. 2d). Generally, more than 40% of her skin was affected by blisters and associated symptoms. In contrast, the proband's two affected brothers and one first cousin showed less severe symptoms. For instance, although generalized blistering of the skin and involvement of the oral mucosa was a common feature (Fig. 2b), these patients lacked symptoms like milia, atrophic scarring, chronic wounds, anonychia, tissue granulation, growth retardation, and pseudosyndactylism of toes. The proportion of affected skin in these patients was also significantly lower. Taken together, the clinical presentation of the proband was clearly indicative of dystrophic epidermolysis bullosa. We further characterized her clinical subtype as "RDEB-generalized intermediate" by using the electronic version of a recently developed clinical diagnostic matrix for EB [6].

Because of the marked genetic heterogeneity displayed by EB, we performed exome sequencing on the proband's DNA. We first performed homozygosity mapping on the sequencing data, which enabled us to identify numerous runs of homozygosity (ROHs), globally accounting for 248.57 Mb of the autosomal exome (Fig. 1b). Notably, chromosomes 1, 2, and 3 alone contributed to more than 72% of the total homozygosity (Fig. 1b). Following a stringent computational analysis of exome data, we retained 37 high-quality and rare (MAF<1%) homozygous variants that had an impact on the protein level (not synonymous nor intronic). Among them, a homozygous missense

(NM_000094.3:c.8038G>A:p.Gly2680Ser, genome build hg19) in exon 108 of the *COL7A1* gene was considered as the potential candidate for the disease etiology. Upon Sanger validation, this variant perfectly co-segregated with the phenotype in the family (Fig. 1a and c). *In silico* analysis predicted this Glycine-to-Serine substitution to be highly pathogenic. The functional significance of this variant could also be recognized by the fact that codon 2680 is highly conserved across vertebrate species, both at nucleotide and amino acid levels (Fig. 1d). Finally, c.8038G>A is an extremely rare allele (allele frequency = 0.00001992 in gnomAD, absent in South Asians and with highest population-specific allele frequency of 0.000123 in Africans) with no homozygotes identified [7].

The clinical outcome and the inheritance pattern of DEB largely depends on the nature and position of mutations in the *COL7A1* gene, especially when glycine substitutions within the triple helix of collagen VII are involved, notably in DDEB. However, a clear genotype-phenotype correlation is still missing in the literature [3, 8]. In particular, missense variants at codon 2680 of *COL7A1* gene have been described in association with different clinical subtypes of EB. For example, c.8039G>A:p.Gly2680Asp was reported in a proband from the UK to be associated with DDEB, subtype pruriginosa [9]. The same amino acid, changed to serine, was detected as a *de novo* mutation in a Polish patient with severe generalized RDEB, in a compound heterozygous state with another missense variant (c.425A>G:p.Lys142Arg) [5]. Lastly, this variant was reported in conjunction with a novel frameshift mutation (c.4871delC:p.Pro1624GlnfsTer86) in a Chinese family with Bart's Syndrome, a mild form of RDEB [10]. This highlights the fact that the absence of genotype-phenotype correlation in EB is not uncommon.

In summary, we describe here the occurrence of a missense variant c.8038G>A:p.Gly2680Ser in the *COL7A1* gene, detected homozygously in a consanguineous Pakistani family with RDEB. Although this variant was previously reported to be associated with EB, to the best of our knowledge, it was never identified in a homozygous state. Our findings suggest that this mutation can be pathogenic per se, and possibly associated with generalized intermediate RDEB. Furthermore, the emergence of this particular glycine substitution in patients from diverse ethnic backgrounds such as China, UK, Poland, and Pakistan indicates that this variant most likely constitutes a recurrent mutational hotspot in the *COL7A1* gene, rather than a germline mutation present at low levels in the general population.

Acknowledgments

We are highly indebted to the patients and the healthy members of the family for their participation in this study. We are also thankful to the Swiss Confederation for the award of ‘Swiss Government Excellence Fellowship for PhD’ to the first author of this article.

References

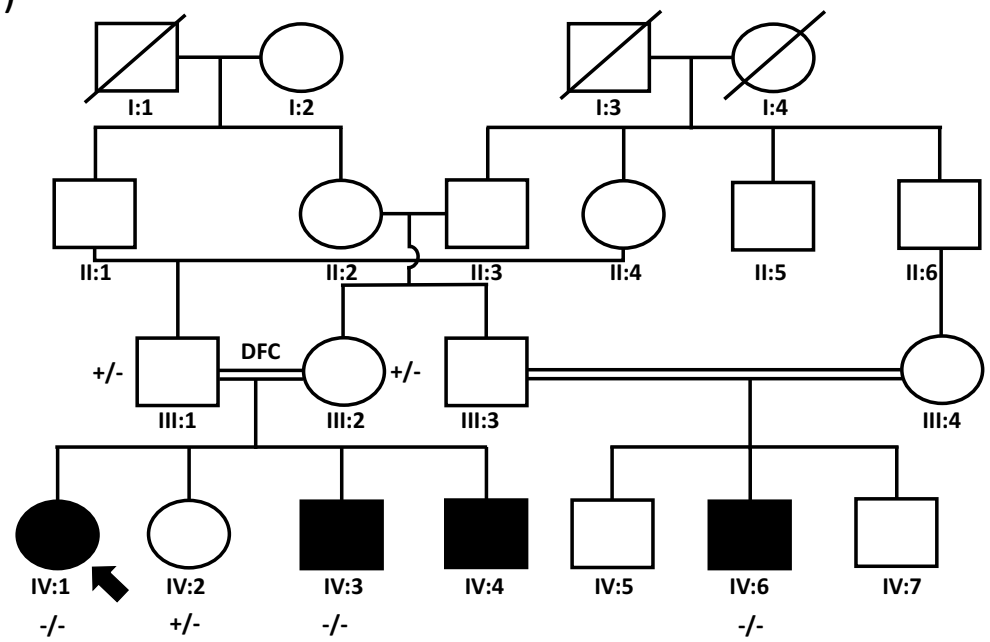
1. Fine, J.D., Epidemiology of Inherited Epidermolysis Bullosa Based on Incidence and Prevalence Estimates From the National Epidermolysis Bullosa Registry, Jama Dermatol 152 (2016) 1231-1238. https://doi.org/10.1001/jamadermatol.2016.2473
2. McGrath, J.A., Recently Identified Forms of Epidermolysis Bullosa, Ann Dermatol 27 (2015) 658-666. https://doi.org/10.5021/ad.2015.27.6.658
3. Pfendner, E.G. Lucky, A.W., *Dystrophic Epidermolysis Bullosa*, in *GeneReviews((R))*, M.P. Adam, et al., Editors. 2018: Seattle (WA).
4. Fine, J.D., Eady, R.A.J., Bauer, E.A., Bauer, J.W., Bruckner-Tuderman, L., Heagerty, A., Hintner, H., Hovnanian, A., Jonkman, M.E., Leigh, I., McGrath, J.A., Mellerio, J.E., Murrell, D.E., Shimizu, H., Uitto, J., Vahlquist, A., Woodley, D. Zambruno, G., The classification of inherited epidermolysis bullosa (EB): Report of the Third International Consensus Meeting on Diagnosis and Classification of EB, J Am Acad Dermatol 58 (2008) 931-950. https://doi.org/10.1016/j.jaad.2008.02.004
5. Wertheim-Tysarowska, K., Sobczynska-Tomaszewska, A., Kowalewski, C., Kutkowska-Kazmierczak, A., Wozniak, K., Niepokoj, K., Klausegger, A., Sypniewska-Jutkiewicz, J., Stepien, A. Bal, J., Novel and recurrent COL7A1 mutation in a Polish population, Eur J Dermatol 22 (2012) 23-8. https://doi.org/10.1684/ejd.2011.1583
6. Yenamandra, V.K., Moss, C., Sreenivas, V., Khan, M., Sivasubbu, S., Sharma, V.K. Sethuraman, G., Development of a clinical diagnostic matrix for characterizing inherited epidermolysis bullosa, Brit J Dermatol 176 (2017) 1624-1632. https://doi.org/10.1111/bjd.15221
7. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T., Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., Cooper, D.N., Deflaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M., Poplin, R., Rivas, M.A., Ruano-Rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K., Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-Luna, M.T., Weisburd, B., Won, H.H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman, C.M., Kathiresan, S., Laakso, M., McCarroll, S., McCarthy, M.I., McGovern, D., McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M., Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G., Daly, M.J., MacArthur, D.G. Exome Aggregation, C., Analysis of protein-coding genetic variation in 60,706 humans, Nature 536 (2016) 285-91. https://doi.org/10.1038/nature19057
8. Knopfel, N., Noguera-Morel, L., Hernandez-Martin, A., Garcia-Martin, A., Garcia, M., Mencia, A., Pedrero, R.M., de Lucas, R., Escamez, M.J. Torrelo, A., Identical COL71A1 heterozygous mutations resulting in different dystrophic epidermolysis bullosa phenotypes, Int Congr Ser 35 (2018) E94-E98. https://doi.org/10.1111/pde.13367
9. Almaani, N., Liu, L., Dopping-Hepenstal, P.J.C., Lai-Cheong, J.E., Wong, A., Nanda, A., Moss, C., Martinez, A.E., Mellerio, J.E. McGrath, J.A., Identical Glycine Substitution Mutations in Type VII Collagen May Underlie Both Dominant and Recessive Forms of Dystrophic Epidermolysis Bullosa, Acta Derm-Venereol 91 (2011) 262-266. https://doi.org/10.2340/00015555-1053
10. Yan, Y.S., Meng, Z.Y., Hao, S.J., Wang, F., Jin, X.H., Sun, D.G., Gao, H.F. Ma, X., Five Novel COL7A1 Gene Mutations in Three Chinese Patients with Recessive Dystrophic Epidermolysis Bullosa, Ann Clin Lab Sci 48 (2018) 100-105.

Legends:

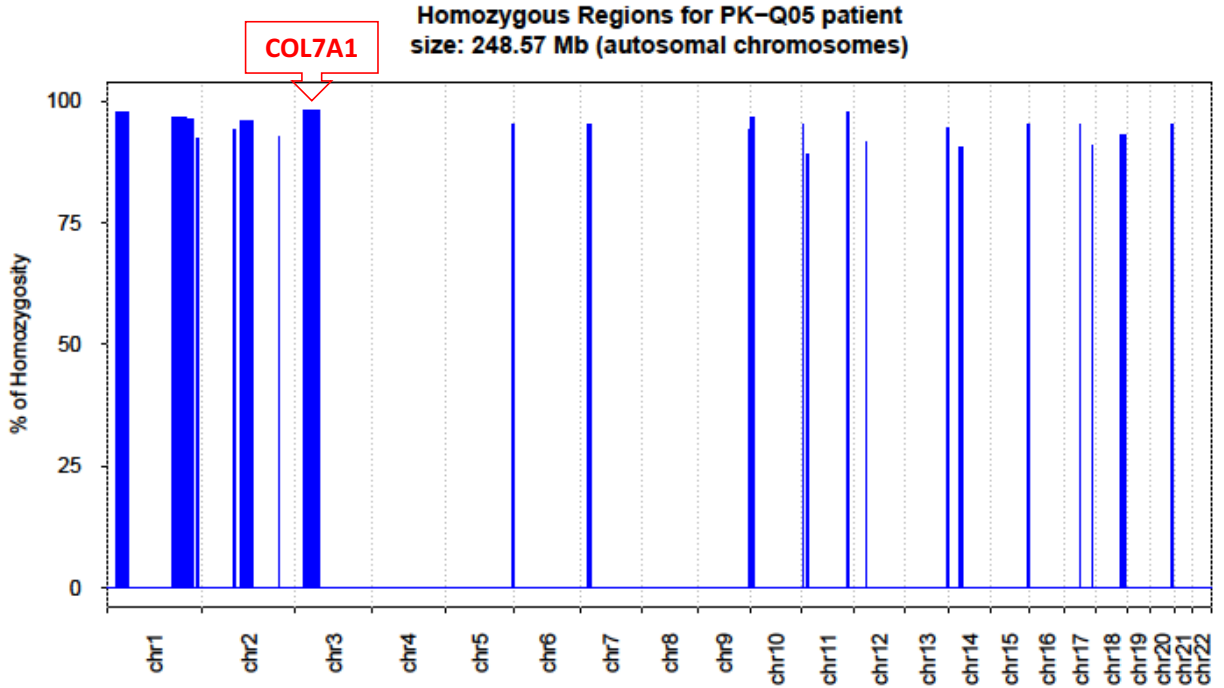
Figure 1: Molecular diagnosis of patients with RDEB. (a) Pedigree of the family. Filled symbols represent affected individuals, whereas open symbols indicate healthy subjects. The proband is designated by a black arrow. DFC = Double First Cousin. Double lines between symbols indicate consanguineous marriages. (+): wild type allele, (-): mutant allele, p.Gly2680Ser. (b) Homozygosity map displaying runs of homozygous (ROH) across the genome as vertical peaks. The position of *COL7A1* is shown on the top of a large 43-Mb ROH on chromosome 3. (c) Sanger sequencing results for all available family members. The missense variant (c.8038G>A:p.Gly2680Ser) is indicated by a red arrow. While both parents and a healthy sister of the proband are heterozygous, all patients are homozygous for this mutation. (d) Output from the UCSC genome browser, showing conservation of the missense variant c.8038G>A:p.Gly2680Ser (highlighted in red box) across vertebrates.

Figure 2: Clinical presentation of patients with RDEB. (a) Proband showing oral blisters, mild teeth decay and facial milia. (b) Proband's brother showing oral blisters. (c) Proband's hands showing excessive skin eruption as a consequence of frequent blisters. (d) Proband's feet, displaying progressive contractures coupled with pseudosyndactylism of toes, resulting in 'mitten-feet'.

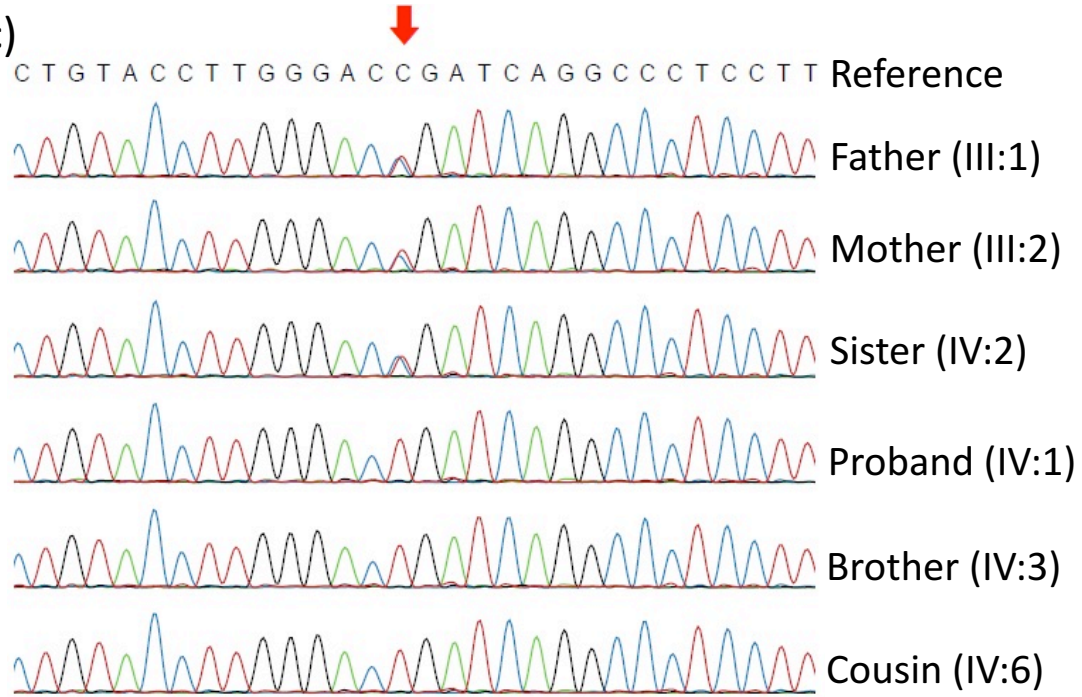
(a)


I:1
I:2
I:3
I:4
II:1
II:2
II:3
II:4
II:5
II:6
+/-
DFC
+/-
III:1
III:2
III:3
III:4
IV:1
-/-
IV:2
+/-
IV:3
-/-
IV:4
IV:5
IV:6
-/-
IV:7

(b)

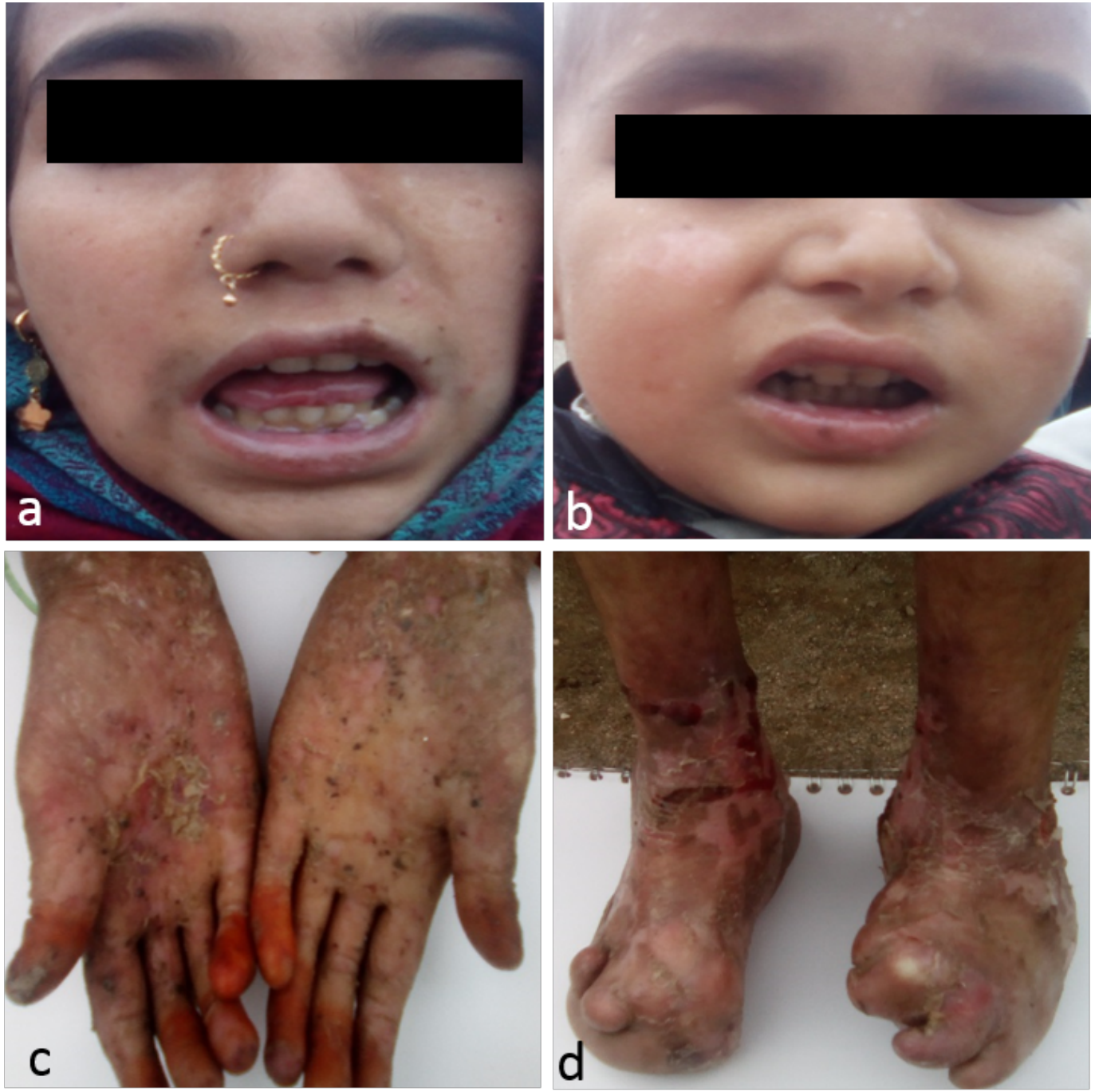

Homozygous Regions for PK−Q05 patient
size: 248.57 Mb (autosomal chromosomes)
COL7A1
% of Homozygosity
100
75
50
25
0
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chr20
chr21
chr22

(c)


C T G T A C C T T G G G A C C G A T C A G G C C C T C C T T
Reference
Father (III:1)
Mother (III:2)
Sister (IV:2)
Proband (IV:1)
Brother (IV:3)
Cousin (IV:6)

(d)


COL7A1
K 2682
P 2681
G 2680
I 2679
L 2678
G 2677
Multiz Alignments
Gaps
Human
Chimp
Gorilla
Rhesus
Mouse
Pig
Dog
Elephant
Chicken
Lizard
Zebrafish


a
b
c
d