Genotype–Phenotype Correlation in Recessive Dystrophic Epidermolysis Bullosa: When Missense Doesn't Make Sense

To the Editor:

All forms of dystrophic epidermolysis bullosa (DEB) result from mutations in the gene encoding type VII collagen, *COL7A1*, the major component of anchoring fibrils at the dermal–epidermal junction (Christiano *et al*, 1994). Approximately 200 different pathogenic mutations have been reported and, to some extent, a paradigm for genotype–phenotype correlation has emerged. Typically, severe generalized forms of recessive DEB result from premature termination codons on both *COL7A1* alleles (Uitto and Christiano, 1994), whereas autosomal dominant cases are caused by heterozygous glycine substitutions within the collagenous triple helix (Christiano *et al*, 1994). It is clear, however, that the clinical diversity seen in both recessive and dominant forms of DEB is mirrored by a spectrum of mutations in *COL7A1*, with clinical heterogeneity partially being explained through an accumulating database of other missense, splice site, or insertion/deletion frameshift mutations (Whittock *et al*, 1999).

In recessive DEB, the repertoire of mutations includes: nonsense mutations, splice site mutations, internal deletions or insertions, "silent" glycine substitutions within the triple helix, and non-glycine missense mutations within the triple helix or non-collagenous NC-2 domain (Whittock *et al*, 1999). We now report an example of a pathogenic missense mutation within the NC-1 domain. The nucleotide substitution results in cryptic splicing and an in-frame deletion of 29 amino acids from the part of the protein with cartilage matrix protein homology, thus providing clinicopathological evidence for a new type of mutation in recessive DEB and demonstrating the importance of this part of type VII collagen in epidermal–dermal adhesion.

The proband is an 8-y-old Korean male with a history of blisters and erosions at birth but only limited trauma-induced blistering during infancy. But at the age of 7 y, there was a worsening of blisters with development of milia and scarring. There was never any history of skin itching. Physical examination revealed bullae (some of which were hemorrhagic), erosions, scars, and milia on his lower legs, hands, feet, and pinnae. He also had dystrophic nails with some loss of toenails (Fig 1*A*). Notably, the skin on his shins was the most affected site (Fig 1*B*). No other family member was affected and both parents are clinically normal. The clinical features were suggestive of a mild form of recessive DEB or possibly *de novo* dominant disease, including the pretibial variant. Indirect immunofluorescence microscopy of the affected individual's skin using an anti-type VII collagen monoclonal antibody (LH7.2; Fig 1*C*) (Heagerty *et al*, 1986) showed bright labeling at the dermal–epidermal junction that was of an intensity very similar to normal control skin (Fig 1*D*). This finding is unhelpful in establishing the precise diagnosis and in distinguishing between mild recessive and *de novo* dominant DEB.

Following ethical approval, informed consent, and in compliance with the Helsinki guidelines, DNA extraction from peripheral blood samples from the patient and his parents was performed using QIAamp DNA Blood Midi Kits (Qiagen, Crawley, West Sussex, UK). PCR amplification of the *COL7A1* gene was performed, and heteroduplex analysis was carried out as described previously (Christiano *et al*, 1997b). Any fragment with a different mobility to amplified control DNA was sequenced using dRhodamine dye terminator technology on an ABI 310 capillary system (PE Biosystems, Warrington, UK).

Sequencing of genomic DNA showed that the patient is heterozygous for two nucleotide substitutions: 341G>T in exon 3 and 5797C>T in exon 70, the former inherited from the mother and the latter from the father (GenBank No. L02870). The mutation 341G>T converts a glycine (GGG) residue to valine (GTG) residue, and is designated G114V (Fig 2*A*). The 5797C>T mutation changes arginine (CGA) to a premature stop codon (TGA) and is designated R1933X (Fig 2*B*). Thus, the affected individual is a compound heterozygote for the mutations G114V/R1933X in *COL7A1*. The mutation R1933X has been reported once previously in an individual with Hallopeau–Siemens recessive DEB (compound heterozygote for R578X/R1933X) (Whittock *et al*, 1999) but G114V is a new finding. The possibility of this missense change being a non-pathogenic polymorphism was excluded by direct sequencing of 100 ethnically matched (Korean/Chinese/Japanese) control chromosomes.

To assess the potential pathogenicity of the mutation G114V, the genomic sequence was examined using the Delila software package (Schneider, 1997a), which scans genomic sequences with weight matrices for sites with positive R_i (individual information content) values, and the results were displayed using the Sequence Walker program (Schneider, 1997b). The 341G>T nucleotide substitution, located 86 nucleotides upstream from the exon 3 natural donor, shows high semblance to the donor splice site consensus (Zhang, 1998) and activates a new strong cryptic site (R_i value increases from 0.0 to 8.2 bits) at this position (Fig 2*C* and *D*), whereas the natural donor site remains

Abbreviation: DEB, dystrophic epidermolysis bullosa

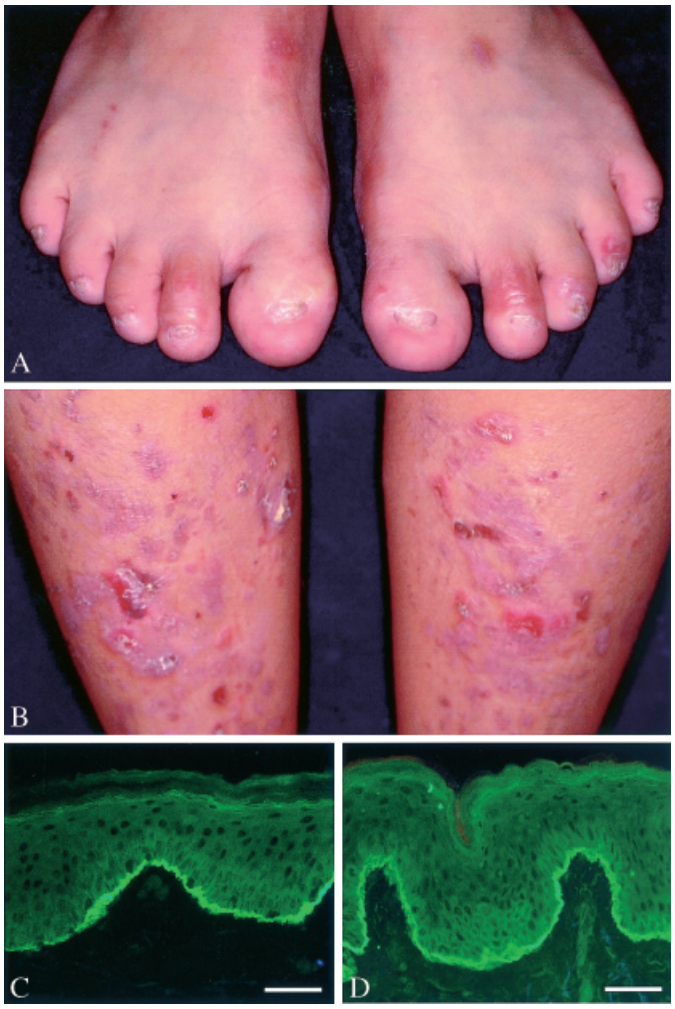

Figure 1
Clinicopathological assessment is consistent with either localized recessive dystrophic epidermolysis bullosa or *de novo* dominant disease. The clinical features include (*A*) dystrophic or rudimentary toenails and (*B*) blisters, erosions, and inflammatory papules on the shins. Indirect immunofluorescent staining with LH7.2 monoclonal antibody (NC-1 domain of type VII collagen) shows bright linear labeling at the dermal–epidermal junction in the patient's skin (*C*) of intensity and distribution similar to normal control skin (*D*). *Scale bar* = 40 μm.

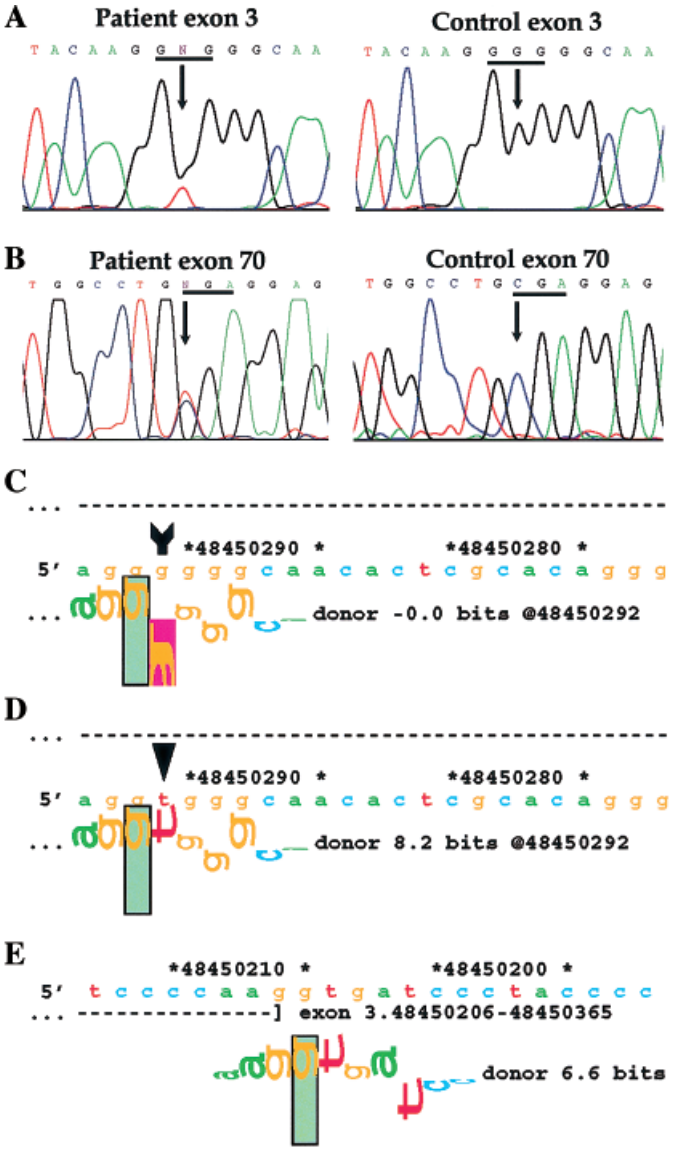

Figure 2
Nucleotide sequencing of the patient's genomic DNA reveals compound heterozygosity for two single nucleotide substitutions in *COL7A1*. (*A*) There is a heterozygous 341G>T transition in exon 3 (maternally inherited), which is predicted to change a glycine residue to valine (G114V). (*B*) There is also a heterozygous 5797C>T transition in exon 70 (paternally inherited), which is predicted to change an arginine residue to a stop codon (R1933X). The mutation 341G>T, however, is also predicted to activate a new strong cryptic site, with the R_i value increasing from 0.0 bits (*C*) to 8.2 bits (*D*) at this position, whereas the natural donor site remains unchanged at 6.6 bits (*E*). This suggests cryptic splicing induced by the substitution 341G>T. No such changes were evident in assessing 5797C>T.

unchanged at 6.6 bits (Fig 2*E*). This change is predicted to result in aberrant splicing with removal of the last 87 nucleotides of exon 3, as well as some residual splicing at the natural donor splice site.

Following informed consent, extraction of total RNA from skin biopsies of the proband, mother, and control was performed using Qiagen RNeasy Mini Kits (Qiagen) with subsequent reverse transcription using ABgene Reverse-iT 1st Strand Synthesis Kit (ABgene House, Epsom, Surrey, UK). The following sets of exonic primers were used to amplify and sequence *COL7A1* complementary DNA (cDNA) in the region of the mutation G114V: sense primer 5′-ATGACCCACGGACAGAGTTC-3′ and anti-sense primer 5′-AACTCGCTTCAGCTCCTCAG-3′ (PCR product 302 base pairs; nucleotide positions 257–558). The annealing temperature used was 55°C. Agarose gel electrophoresis of the amplified PCR products showed two bands in the proband's and mother's samples compared with a single band in the control (Fig 3*A*). Direct sequencing of the smaller band (proband and mother) confirmed the anticipated 87-bp deletion in exon 3, corresponding to nucleotides 340–426 (340del87) (Fig 3*B*). This is predicted to result in an in-frame deletion of 29 amino acids from the NC-1 domain of type VII collagen. This deletion occurs within the region of cartilage matrix protein homology and is upstream from the nine fibronectin III-like repeats, within which the epitope for the LH7.2 antibody occurs (Tanaka *et al*, 1994) (Fig 3*C*). This finding helps explain the "normal" intensity anti-type VII collagen immunostaining, as well as demonstrating, in a

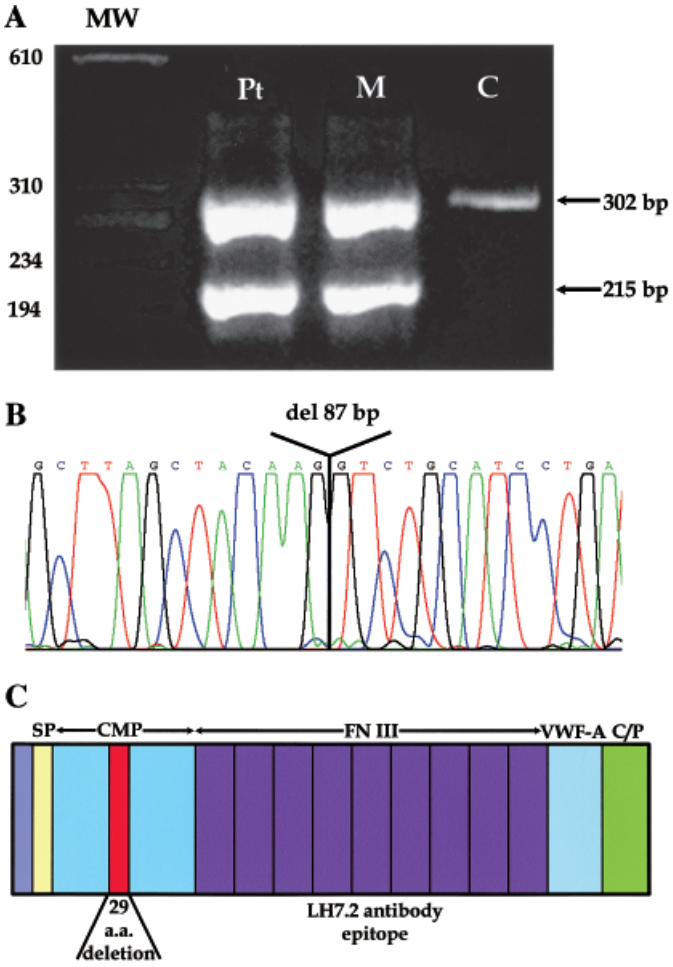

Figure 3
Direct sequencing of complementary DNA (cDNA) from the patient's and mother's skin confirms cryptic splicing in exon 3. (*A*) PCR of cDNA between nucleotide positions 257 and 558 reveals a single 302 bp band in a control (*C*) and two bands in the proband (Pt) and mother (M), one of which is 302 bp and the other 215 bp, i.e., 87 bp shorter. (*B*) Direct sequencing of the lower band reveals an 87 bp deletion between nucleotides 340 and 426. (*C*) Schematic diagram of the NC-1 domain of *COL7A1* messenger RNA showing the position of the 29 amino acid deletion within the CMP domain, upstream from the epitope for the LH7.2 antibody in the middle of the FN III-like repeats. SP, signal peptide; CMP, domain with homology to cartilage matrix protein; FN III, fibronectin-type III-like motifs; VWF-A, von Willebrand factor-type A-like domain; C/P, cysteine- and proline-rich domain (Christiano *et al*, 1992).

patient with DEB, the functional role of the cartilage matrix protein homology region in dermal–epidermal adhesion.

The possible effects of the paternal nonsense mutation on aberrant splicing were also examined using the same approach. The nucleotide substitution 5797C>T, which is located 25 nucleotides upstream from the exon 70 natural donor, shows an increase in R_i value from −1.7 to 5.8 bits at this position, whereas the natural donor site remains unchanged at 8.7 bits. Although the cryptic site created by this substitution is a legitimate site, it is 2.9 bits weaker than the natural site, which would correspond to an 8-fold ($2^{\Delta Ri} = 2^{2.9}$) difference in splicing, assuming minimal conversion of information to energy dissipated (Rogan *et al*, 1998). Even though this site is presumably scanned first, the thermodynamic difference is so marked that it cannot compete for binding to the U2 spliceosome. Indeed, RT-PCR amplification across the site of the mutation R1933X using the proband's cDNA did not disclose any evidence of cryptic splicing (data not shown).

Our study has identified the pathogenic relevance of a particular nucleotide substitution in *COL7A1*, 341C>T, in inducing cryptic splicing within the NC-1 domain. It is likely that this previously unreported phenomenon in type VII collagen may also account for the molecular pathology in a number of other patients with recessive DEB. Indeed, a retrospective analysis of our non-Hallopeau Siemens generalized RDEB patient mutation database has identified some similar sequence variants in the NC-1 domain that may also potentially lead to cryptic splicing. These include 245T>G (V82G) in exon 2 and 3359G>A (R1120K) in exon 25. But there is no evidence that such a phenomenon may also occur at other sites.

Specifically, from the collated *COL7A1* mutation databases, a small number of potential pathogenic (non-glycine) missense mutations have been reported: these include R1772W (exon 61); Q1924P (exon 70); R2002C, R2008C, and R2008G (all exon 73); R2063G and R2063W (exon 74); and R2791W, M2798K (both exon 113). Of these mutations, R2008G and R2063G result in new glycine residues within the triple helix and may therefore disrupt the structure of the collagenous triple helix. Q1924P occurs within a consensus splice site, but it is predicted to have only a limited effect on splicing (R_i value reduces from 10.9 to 8.9 bits). Of the remainder, analysis with the Delila software package predicts no cryptic splicing.

Mutations resulting in aberrant splicing account for at least 15% of all DEB mutations, and occur in both recessive and dominant disease (Whittock *et al*, 1999). Over 30 single nucleotide substitutions within consensus sites (mostly at the +1 or −1 position) have been described, although RNA analysis is usually lacking. It has mostly been assumed that such mutations result in exon skipping but, even in the absence of available RNA, it is now evident that the splicing software can add to genotype–phenotype understanding. For example, the mutation IVS35-1G>C has been reported in a three-generation family with transient bullous disease of the newborn (Christiano *et al*, 1997a), an autosomal dominant condition that ameliorates with increasing age. Analysis using the Delila software reveals that this mutation results in a "leaky" acceptor splice site (R_i value reduced from 11.8 to 4.5), thus predicting both in-frame skipping of exon 36 (78 bp) and normal splicing. It is plausible, therefore, that relative differences in these two possible splice choices at different ages (neonate versus adult) account for the clinical pattern of this family's DEB. To add to the molecular complexity of splicing anomalies in other forms of dominant and recessive DEB, it has been shown that certain exonic sequences outside consensus splice sites also may influence splicing and a few examples of internal deletions (up to 28 bp) in exons 73 or 87 of *COL7A1* have been reported (Cserhalmi-Friedman *et al*, 1998; Sakuntabhai *et al*, 1998).

Overall, this study further expands the spectrum of different types of pathogenic mutations in *COL7A1* in individuals with recessive DEB, as well as highlighting the subtle but consequential effects of cryptic splicing in this inherited skin disorder. The study also illustrates the utility of com-

putational tools that can accurately predict cryptic splicing, such as Automated Splice Site Analyses (https://splice.cmh.edu/cgi-bin/protected/menu.cgi?menu.html) or Splice Site Prediction by Neural Network (http://www.fruitfly.org/seq_tools/splice.html).

Vesarat Wessagowit,* Soo-Chan Kim,† Se Woong Oh,† and John A. McGrath*

*Genetic Skin Disease Group, St John's Institute of Dermatology, St Thomas' Hospital, London, UK; †Department of Dermatology and Cutaneous Biology Research Institute, Yongdong Severance Hospital, Yonsei University College of Medicine, Dogok-dong, Kangnam-gu, 135-720, Seoul, Korea

DOI: 10.1111/j.0022-202X.2005.23650.x

Manuscript received October 16, 2004; revised October 28, 2004; accepted for publication October 29, 2004

Address correspondence to: Vesarat Wessagowit, Genetic Skin Disease Group, St John's Institute of Dermatology, St Thomas' Hospital, Lambeth Palace Road, London SE1 7EH, UK. Email: vesarat.wessagowit@kcl.ac.uk

References

Christiano AM, Fine JD, Uitto J: Genetic basis of dominantly inherited transient bullous dermolysis of the newborn: A splice site mutation in the type VII collagen gene. J Invest Dermatol 109:811–814, 1997a

Christiano AM, Hoffman GG, Zhang X, Xu Y, Tamai Y, Greenspan DS, Uitto J: Strategy for identification of sequence variants in COL7A1 and a novel 2-bp deletion mutation in recessive dystrophic epidermolysis bullosa. Hum Mutat 10:408–414, 1997b

Christiano AM, Rosenbaum LM, Chung-Honet LC, *et al*: The large non-collagenous domain (NC-1) of type VII collagen is amino-terminal and chimeric. Homology to cartilage matrix protein, the type III domains of fibronectin and the A domains of von Willebrand factor. Hum Mol Genet 1:475–481, 1992

Christiano AM, Ryynanen M, Uitto J: Dominant dystrophic epidermolysis bullosa: Identification of a Gly−>Ser substitution in the triple-helical domain of type VII collagen. Proc Natl Acad Sci USA 91:3549–3553, 1994

Cserhalmi-Friedman PB, McGrath JA, Mellerio JE, *et al*: Restoration of open reading frame resulting from skipping of an exon with an internal deletion in the COL7A1 gene. Lab Invest 78:1483–1492, 1998

Heagerty AH, Kennedy AR, Gunner DB, Eady RA: Rapid prenatal diagnosis and exclusion of epidermolysis bullosa using novel antibody probes. J Invest Dermatol 86:603–605, 1986

Rogan PK, Faux BM, Schneider TD: Information analysis of human splice site mutations. Hum Mutat 12:153–171, 1998

Sakuntabhai A, Hammami-Hauasli N, Bodemer C, *et al*: Deletions within COL7A1 exons distant from consensus splice sites alter splicing and produce shortened polypeptides in dominant dystrophic epidermolysis bullosa. Am J Hum Genet 63:737–748, 1998

Schneider TD: Information content of individual genetic sequences. J Theor Biol 189:427–441, 1997a

Schneider TD: Sequence walkers: A graphical method to display how binding proteins interact with DNA or RNA sequences. Nucleic Acids Res 25: 4408–4415, 1997b

Tanaka T, Takahashi K, Furukawa F, Imamura S: The epitope for anti-type VII collagen monoclonal antibody (LH7:2) locates at the central region of the N-terminal non-collagenous domain of type VII collagen. Br J Dermatol 131:472–476, 1994

Uitto J, Christiano AM: Molecular basis for the dystrophic forms of epidermolysis bullosa: Mutations in the type VII collagen gene. Arch Dermatol Res 287:16–22, 1994

Whittock NV, Ashton GH, Mohammedi R, *et al*: Comparative mutation detection screening of the type VII collagen gene (COL7A1) using the protein truncation test, fluorescent chemical cleavage of mismatch, and conformation sensitive gel electrophoresis. J Invest Dermatol 113:673–686, 1999

Zhang MQ: Statistical features of human exons and their flanking regions. Hum Mol Genet 7:919–932, 1998