

**Identified six single nucleotide variants in the *COL17A1* gene that
altered RNA splicing by minigene splicing assays**

Yingfei Shao^{1,2}, Ran Zhang¹

¹Department of Nephrology, The Affiliated Qingdao Municipal Hospital of Qingdao
University, Qingdao, China

²Wenzhou Medical University Renji College, Wenzhou, China.

Correspondence:

Ran Zhang, Department of Nephrology, The Affiliated Qingdao Municipal Hospital of
Qingdao University, Qingdao, China. Email:18562788709@163.com

Table S1 Primer sequences for PCR amplification of exons in *COL17A1* from this study

Exon-Primer	Sequences (5'-3')	Product size (bp)
pSPL3- <i>COL17A1</i> -Exon4-F	CCGCTCGAGACCAACGAAACTGGAGCAT	520
pSPL3- <i>COL17A1</i> -Exon4-R	CTAGCTAGCAGTAGGCTTCAAATGCTG	
pSPL3- <i>COL17A1</i> -Exon14-F	CCGCTCGAGCATAAACCAGCCATCCAG	532
pSPL3- <i>COL17A1</i> -Exon14-R	CTAGCTAGCGGGATGGCTATGGTTTCA	
pSPL3- <i>COL17A1</i> -Exon15-F	CCG CTCGAGGAAACCATAGCCATCCCA	619
pSPL3- <i>COL17A1</i> -Exon15-R	CTA GCTAGCCAGGCTTTTGCTTCAGTA	
pSPL3- <i>COL17A1</i> -Exon18-F	CCG CTCGAGTAAGGCTCTGCTGTGATG	699
pSPL3- <i>COL17A1</i> -Exon18-R	CTA GCTAGCGGTATGGGAAGAAGAAAC	
pSPL3- <i>COL17A1</i> -Exon22-F	CCGCTCGAGTTTGGCAACCCTACCTGA	655
pSPL3- <i>COL17A1</i> -Exon22-R	CTAGCTAGCTGCTTGGCTCCTTCTTGC	
pSPL3- <i>COL17A1</i> -Exon31-F	CCGCTCGAGCCCCCTTCTCCTTTCCACC	481
pSPL3- <i>COL17A1</i> -Exon31-R	CTAGCTAGCGCCAGACTTAGTTACCCAGA	
pSPL3- <i>COL17A1</i> -Exon46-F	CCGCTCGAGCTCTGCCTCAGTTTCTCCAT	625
pSPL3- <i>COL17A1</i> -Exon46-R	CTAGCTAGCCTAATGCCTGCCAGATGC	

Table S2 Mutagenesis primers sequences for selected variants in *COL17A1* from this study

Variant-Primer	Sequences (5'-3')
pSPL3- <i>COL17A1</i> -Intron4-c.202+6T>G-F	ACTCAAGTGAGGGCCCACTA
pSPL3- <i>COL17A1</i> -Intron4-c.202+6T>G-R	TAGTGGGCCCTCACTTGAGT
pSPL3- <i>COL17A1</i> -Exon14-c.1139C>T-F	ATCGCTGTAAGTGCCTCAGAT
pSPL3- <i>COL17A1</i> -Exon14-c.1139C>T-R	ATCTGACGCACTTACAGCGAT
pSPL3- <i>COL17A1</i> -Intron15-c.1222+4A>G-F	CCGAAGCTAATGGTAGGAAC
pSPL3- <i>COL17A1</i> -Intron15-c.1222+4A>G-R	GTTCCCTACCATTAGCTTCGG
pSPL3- <i>COL17A1</i> -Exon18-c.1467G>A-F	TTTGTTGCAGCAGAGGAGGT
pSPL3- <i>COL17A1</i> -Exon18-c.1467G>A-R	ACCTCCTCTGCTGCAACAAA
pSPL3- <i>COL17A1</i> -Exon22-c.1834G>A-F	TCAAAAAGACAGCGTGGGTA
pSPL3- <i>COL17A1</i> -Exon22-c.1834G>A-R	TACCCACGCTGTCTTTTGA
pSPL3- <i>COL17A1</i> -Intron31-c.2335+3A>G-F	AAAGCCAGGTGACAGAGCAG
pSPL3- <i>COL17A1</i> -Intron31-c.2335+3A>G-R	CTGCTCTGTCACCTGGCTTT
pSPL3- <i>COL17A1</i> -Intron45-c.3071-5G>A-F	TTGGTCAGCAGGTGACAGTATT
pSPL3- <i>COL17A1</i> -Intron45-c.3071-5G>A-R	AATACTGTCACCTGCTGACCAA
pSPL3- <i>COL17A1</i> -Exon46-c.3198C>T-F	ACGTTGTGAGTTACTTACGG
pSPL3- <i>COL17A1</i> -Exon46-c.3198C>T-R	CCGTAAGTAACTCACAACTG

Table S3 Sequences for qPCR primer

Primer	Sequences (5'-3')
c.1834G>A	
SD6F	TCTGAGTCACCTGGACAACC
R1	TCCAGGTTACGCTGCCTTT
R2	ATCTCCAGGTGCTGCCTTTT
c.3071-5G>A	
F1	TCGACCCAGCAGTGACAGTAT
F2	ACCCAGCACAGGTGACAGTAT
SA2R	ATCTCAGTGGTATTTGTGAGC
R3	CTCCAGGTTGCTGGGTCG
GAPDH	
F	ACACCCACTCCTCCACCTTTG
R	TCCACCACCCTGTTGCTGTAG

Table S4 The sequences of wild-type and variants in *COL17A1* from this study

[illegible]

GGATAGAATAGAAAAGGACCGCCTCCAGGGCATGGCACCCGCGGCGGGAGCAGACCTGGACAAAAT
TGGGCTGCACAGTGACAGCCAGGAGGAGCTCTGGATGTTCTGTGAGGAAGAAGCTAATGATGGAACA
GGAAAATGgtgagagcatcgccgggcacaggccagctcccctgagtccttctcttctgcatctatgtggtatgtgcatgtgctgtgagcaccatcatcgg
gccactctcgccctctctcccctgatcctctcaatagcgctgagaggtaggtgggaccaagactgtgtcccactttactctggaanaactgagttctggaag
aataattgctcaagatcactggactagccaaattagaatctgatctctagtcttctgctcagtttcttctccatacc

COL17A1-Exon18
with c.1467G>A

taaggctctgctgtgatgctgagcagtaatcctgctggtgccacggcagattatctctgcatgagtattctagcatgaagaagccaccacaggcccaatccctgtg
gtcaggctctccactaggagcgagtgtatggaagcggaaggaccggccctcagtgacgcagtggtgtctgttctgtgcagCAGAGGAGGTGAGGA
AGCTGAAGGCGCGTGTGGATGAGCTGGAGAGGATCAGGAGGAGCATACTGCCCTATGGGGACAGCAT
GGATAGAATAGAAAAGGACCGCCTCCAGGGCATGGCACCCGCGGCGGGAGCAGACCTGGACAAAAT
TGGGCTGCACAGTGACAGCCAGGAGGAGCTCTGGATGTTCTGTGAGGAAGAAGCTAATGATGGAACA
GGAAAATGgtgagagcatcgccgggcacaggccagctcccctgagtccttctcttctgcatctatgtggtatgtgcatgtgctgtgagcaccatcatcgg
gccactctcgccctctctcccctgatcctctcaatagcgctgagaggtaggtgggaccaagactgtgtcccactttactctggaanaactgagttctggaag
aataattgctcaagatcactggactagccaaattagaatctgatctctagtcttctgctcagtttcttctccatacc

COL17A1-Exon21/22

tttggcaacctacctgaatcccactagtactgctttgggaattgggcacagctctctggtgaccactctctcttctgcttcttctcagGAGATCGAGGG
TTCCCTGGGACTCCAGgtatgtactctctgatgacctccaggcctgccagtggtgcagctgctcaatcctagctgatggcttctgggaggaaaacaggg
gctgagatatacttctgtgttcacagGTATCCCTGGGCCCCCTTGGGCCACCCAGGTCCACAAGGACCAAAGGGTCAAA
AAGGCAGCGTGGtaagaacacattttcttattcatttattctcattcttctcatttcaataaataacattcattaagtgcctgctgtatgtgaggccttgttttagg
ctctggagagacaccagtgaaccaaacagagaacacctgctgtgtgagctgagatttgggttgggagaagacagtaactgtgagaatagtgaagtctatagc
agtttacgggtgatgcacgacatgggaaggagagcctggaggatgttgggaatgagagctggggagagtaattttacacagggaagccagggaaggcctccct
gagaggccaccatttgagcaaagattcgaagggaaggaaggccaagca

COL17A1-Exon21/22
with c.1834G>A

tttggcaacctacctgaatcccactagtactgctttgggaattgggcacagctctctggtgaccactctctcttctgcttcttctcagGAGATCGAGGG
TTCCCTGGGACTCCAGgtatgtactctctgatgacctccaggcctgccagtggtgcagctgctcaatcctagctgatggcttctgggaggaaaacaggg
gctgagatatacttctgtgttcacagGTATCCCTGGGCCCCCTTGGGCCACCCAGGTCCACAAGGACCAAAGGGTCAAA
AAGGCAGCGTGAgtaagaacacattttcttattcatttattctcattcttctcatttcaataaataacattcattaagtgcctgctgtatgtgaggccttgttttagg
ctctggagagacaccagtgaaccaaacagagaacacctgctgtgtgagctgagatttgggttgggagaagacagtaactgtgagaatagtgaagtctatagc
agtttacgggtgatgcacgacatgggaaggagagcctggaggatgttgggaatgagagctggggagagtaattttacacagggaagccagggaaggcctccct
gagaggccaccatttgagcaaagattcgaagggaaggaaggccaagca

COL17A1-Exon31

ccccctctccttccacctgccagcaagtcaattaggaaaacctcagaaatcatcagttcagatgtgataaccagtttagcttctcctaaatgtaaacacttccacgttggc
cattctctctctagGTGAACAAGGTCTTACTGGGATGCCTGGAATCCGTGGCCCCACCAGGACCTTCTGGAGAC
CCAGGAAAGCCAGgtaacagagcaggaagaggagccatggccctgcttagtgtgtgtgcttctgtgtgtggtatgaggcatgtgtgtgcgtgtgcat
gtgcagggtgtgcagggtgtggggaagggtgtatatgtgcatgggtgtgtgtgtcatgcagggtttacaatgcacccatgtgtgccttatcccgtttctccagca
gctcagcagggaactgccatagtgttaggttgttgcgtgcagtggtgttaggccctggcacctctgggtaactaagtctggc

COL17A1-Exon31
with c.2335+3A>G

ccccctctccttccacctgccagcaagtcaattaggaaaacctcagaaatcatcagttcagatgtgataaccagtttagcttctcctaaatgtaaacacttccacgttggc
cattctctctctagGTGAACAAGGTCTTACTGGGATGCCTGGAATCCGTGGCCCCACCAGGACCTTCTGGAGAC
CCAGGAAAGCCAGgtgacagagcaggaagaggagccatggccctgcttagtgtgtgtgcttctgtgtgtggtatgaggcatgtgtgtgcgtgtgca
tgtgcagggtgtgcagggtgtggggaagggtgtatatgtgcatgggtgtgtgtgtcatgcagggtttacaatgcacccatgtgtgccttatcccgtttctccagca
gctcagcagggaactgccatagtgttaggttgttgcgtgcagtggtgttaggccctggcacctctgggtaactaagtctggc

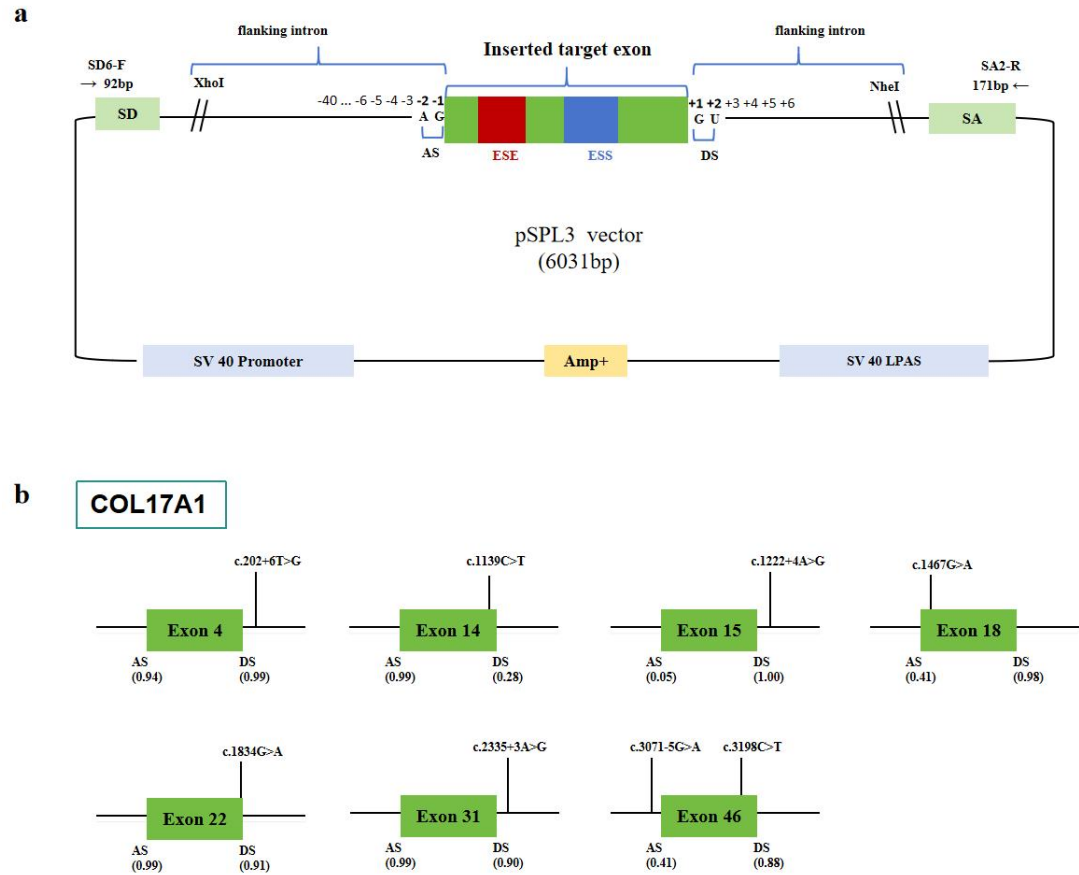
COL17A1-Exon46

ctctgcctcagtttctccatctgtaatttaggggaaggacagtgaggggaagtctaggctccctcctggcccagccttgggcccattgtgtgggcggggcaggcatgt
tatcatgtgcttcaggtcacctccgtggggaagaatgtctcttggacagcgttgactttgcttcttcttggctggcagGTGACAGTATTAGATCTT
ACCTATCCGGAGTTCAGGGTCCCCCAGGCCACCTGGTCCCCCAGGACCTGTACCAACCATCACAGG
CGAGACTTTCGACTACTCAGAGCTGGCAAGCCACGTTGTGAGCTACTTACGGAgtgaagccctccccacgcctgc
acctggcaccttctcctgtgggtctgcttcttctccagagcctcatctctcgttctgttcttcttccctctctggaacatgagtttaattccgaattcataataatcg
ggccactgggatacagtgtaacaagggtctgtccagattcaggcgccacgctggagagacaccagtgagccttccggcaattgactaatgtgttcctaattata
ggttcttccacctgtgcatgaagggtggcagcatctggcaggcattag

COL17A1-Exon46 with c.3071-5G>A	ctctgcctcagtttctccatctgtaatttaggggaaggacagtgaggggaagtctaggctccctcctggcccagccttgggcccatgtgtggggcggggcaggcatgt tatcatgtgcttcaggtcacctccgtggggagaacatgtcctcttgggacaggcttgactttgctttgttccttggtcagcagGTGACAGTATTAGATCTT ACCTATCCGGAGTTCAGGGTCCCCCAGGCCACCTGGTCCCCCAGGACCTGTCACCAACCATCACAGG CGAGACTTTCGACTACTCAGAGCTGGCAAGCCACGTTGTGAGCTACTTACGGAgtaaagccctccccacgcctgc acctggcaccttctcctgtgggtctgcttttgcctccagagcctcatctcctcgtttgctgtttgctttgcccttctctggaacatgagtttaattccgaattcataataatcg ggccactgggatacagtgtaacaagggtctgtccagattcaggcgccacgctggagagacaccgagtgaacctccggcaattgactaatgtgttcctaattata ggttcttcccacctgtgcatgaagggtggcagcatctggcaggcattag
COL17A1-Exon46 with c.3198C>T	ctctgcctcagtttctccatctgtaatttaggggaaggacagtgaggggaagtctaggctccctcctggcccagccttgggcccatgtgtggggcggggcaggcatgt tatcatgtgcttcaggtcacctccgtggggagaacatgtcctcttgggacaggcttgactttgctttgttccttggtggcagGTGACAGTATTAGATCTT ACCTATCCGGAGTTCAGGGTCCCCCAGGCCACCTGGTCCCCCAGGACCTGTCACCAACCATCACAGG CGAGACTTTCGACTACTCAGAGCTGGCAAGCCACGTTGTGAGTTACTTACGGAgtaaagccctccccacgcctgc acctggcaccttctcctgtgggtctgcttttgcctccagagcctcatctcctcgtttgctgtttgctttgcccttctctggaacatgagtttaattccgaattcataataatcg ggccactgggatacagtgtaacaagggtctgtccagattcaggcgccacgctggagagacaccgagtgaacctccggcaattgactaatgtgttcctaattata ggttcttcccacctgtgcatgaagggtggcagcatctggcaggcattag

Lowercase letter: sequences of flanking introns; Capital letter: sequences of target exons; Red letter: mutation sites; “ag”: sequences of acceptor splice sites; “gt”: sequences of donor splicing sites; shading “AG/GT”: cryptic splice sites.

Figure S1 The schematic diagram of the minigene splicing assay constructed by pSPL3 exon trapping vector and position of 8 variants selected in *COL17A1*.



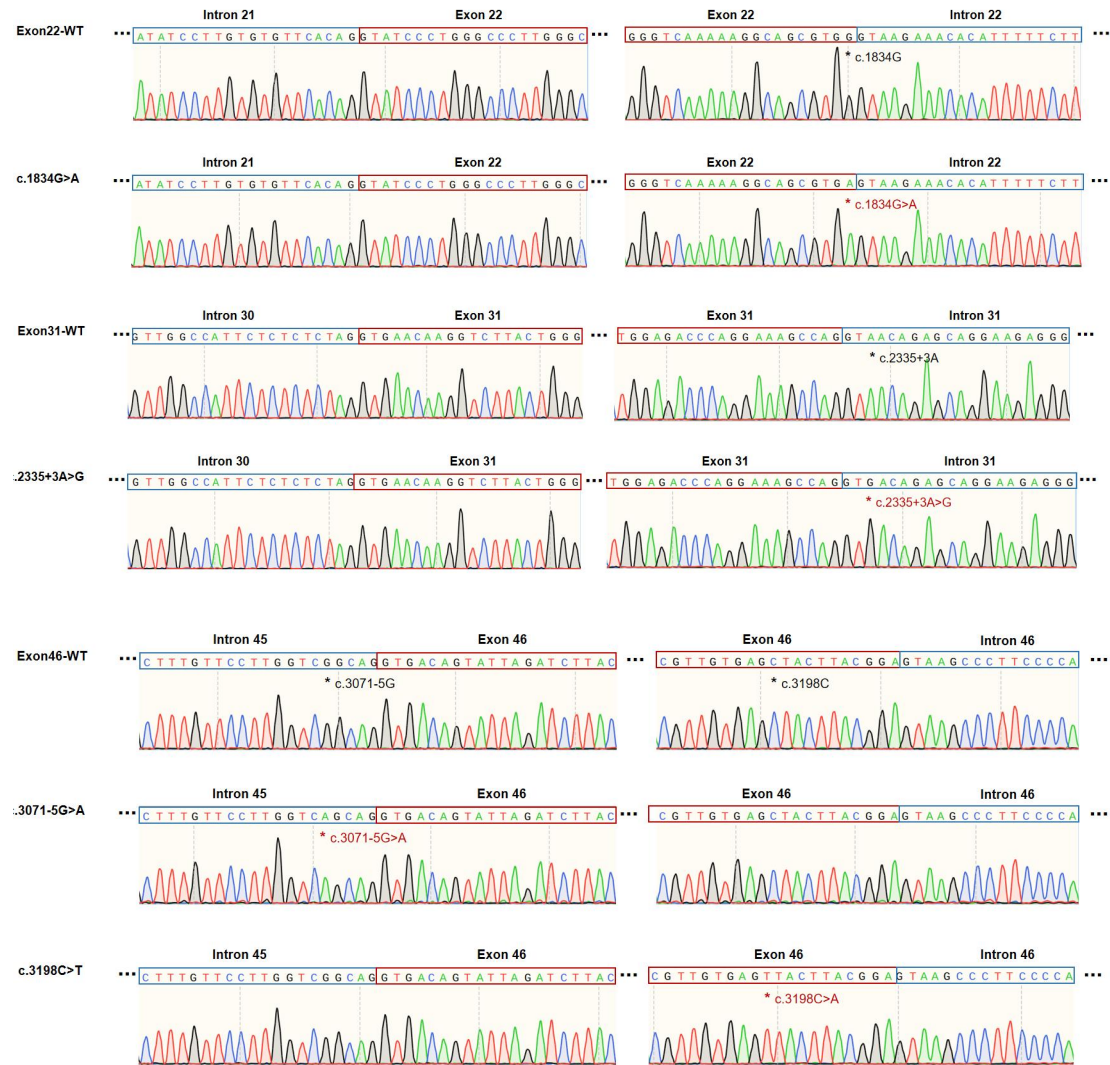
(a). The pSPL3 vector includes two exons (SD and SA) and a functional intron. Transcription starts at the SV40 promoter and ends at the LPAS. The target exon containing partial flanking introns was inserted into pSPL3 vector via XhoI and NheI cloning sites to form the wild-type or mutant plasmid. Splicing regulatory elements located on exons, such as ESEs and ESSs. The positions of intronic variants selected: 3 to 40 nt upstream of a canonical splice acceptor site or 3 to 6 nt downstream of a canonical splice donor site.

(b). Position of 8 variants in *COL17A1*. The green boxes and black horizontal lines represent the coding exons and flanking introns sequences, respectively. The BDGP scores of donor and acceptor splice sites are represented in decimal, as shown at the bottom.

The sizes of boxes and lines are out of proportion. LPAS, late poly(A) signal; AS, acceptor splice site; DS, donor splice site; ESEs, exonic splicing enhancers; ESSs, exonic splicing silencers.

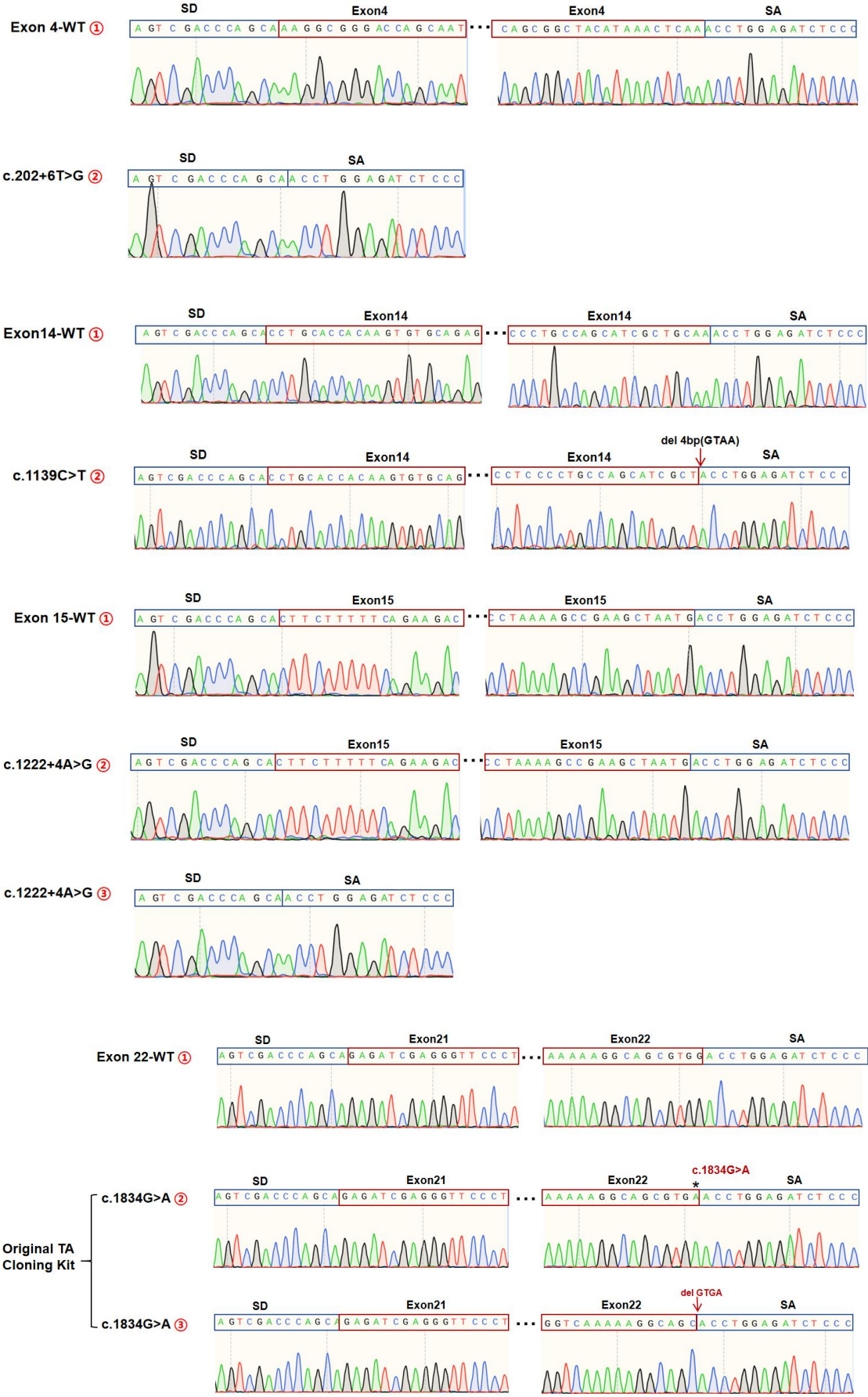
Figure S2 Sequencing diagrams of variants in *COL17A1*.

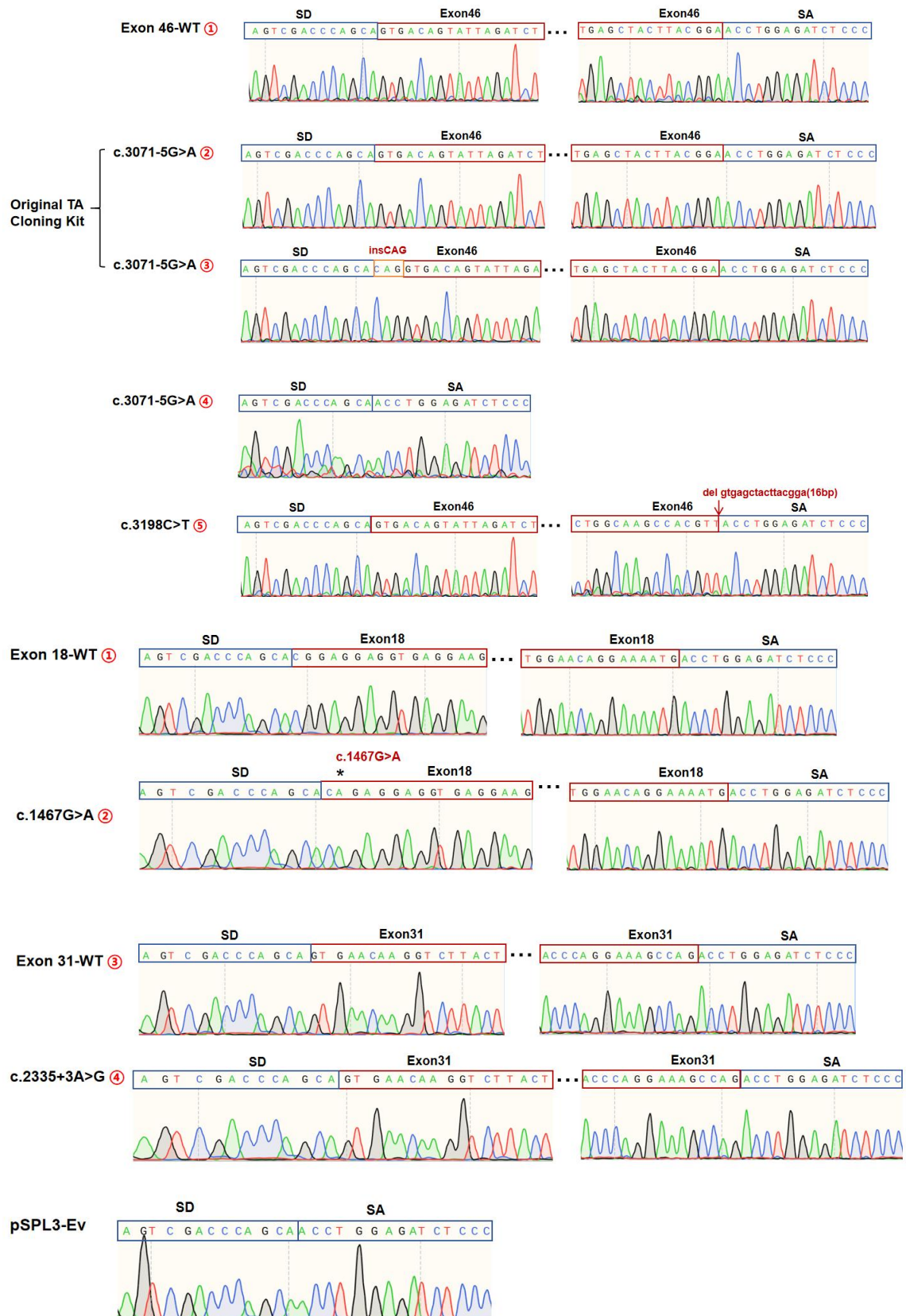




The red boxes and blue boxes are respectively partial sequences of target exons and flanking introns; * indicates the variant site.

Figure S3 Sanger sequencing analysis figure of RT-PCR products.

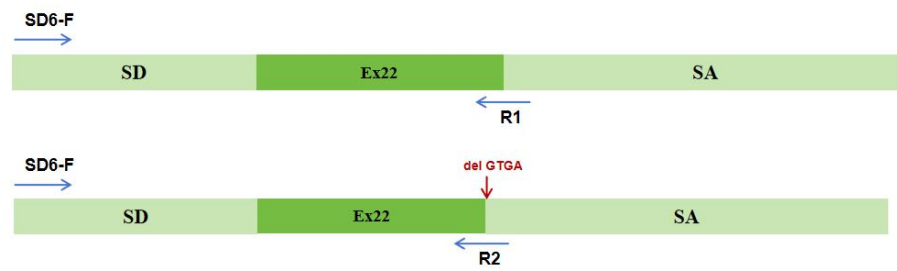




WT, wild type; EV, SD and SA exons of the pSPL3 vector; * indicates the variant site.

Figure S4 Location diagram of qPCR primers

c.1834G>A

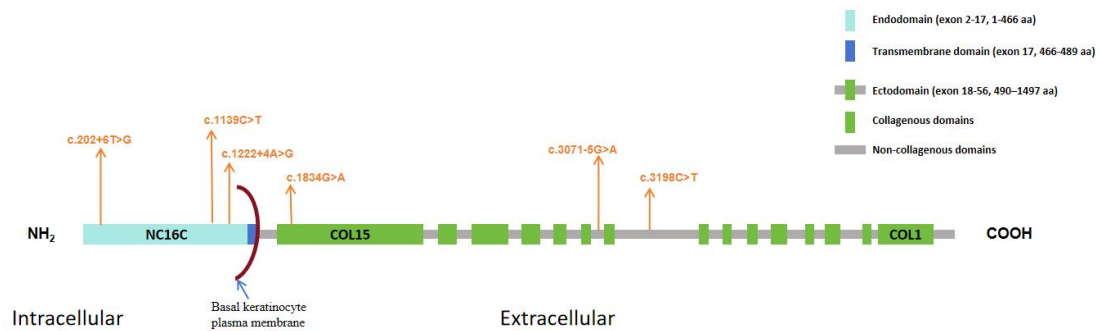


c.3071-5G>A



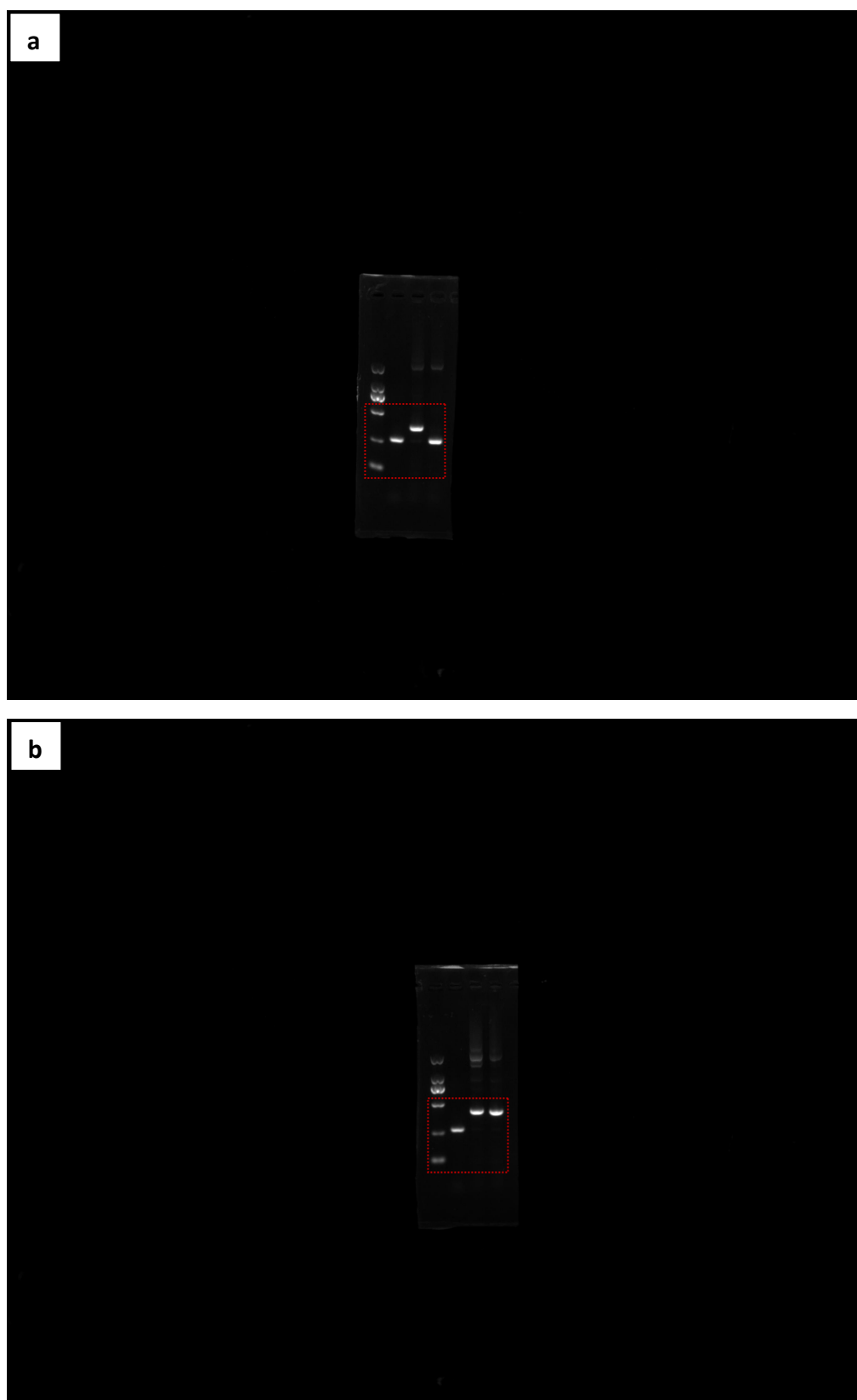
The boxes represent the coding exons, and their sizes are out of proportion. → indicates the primers site

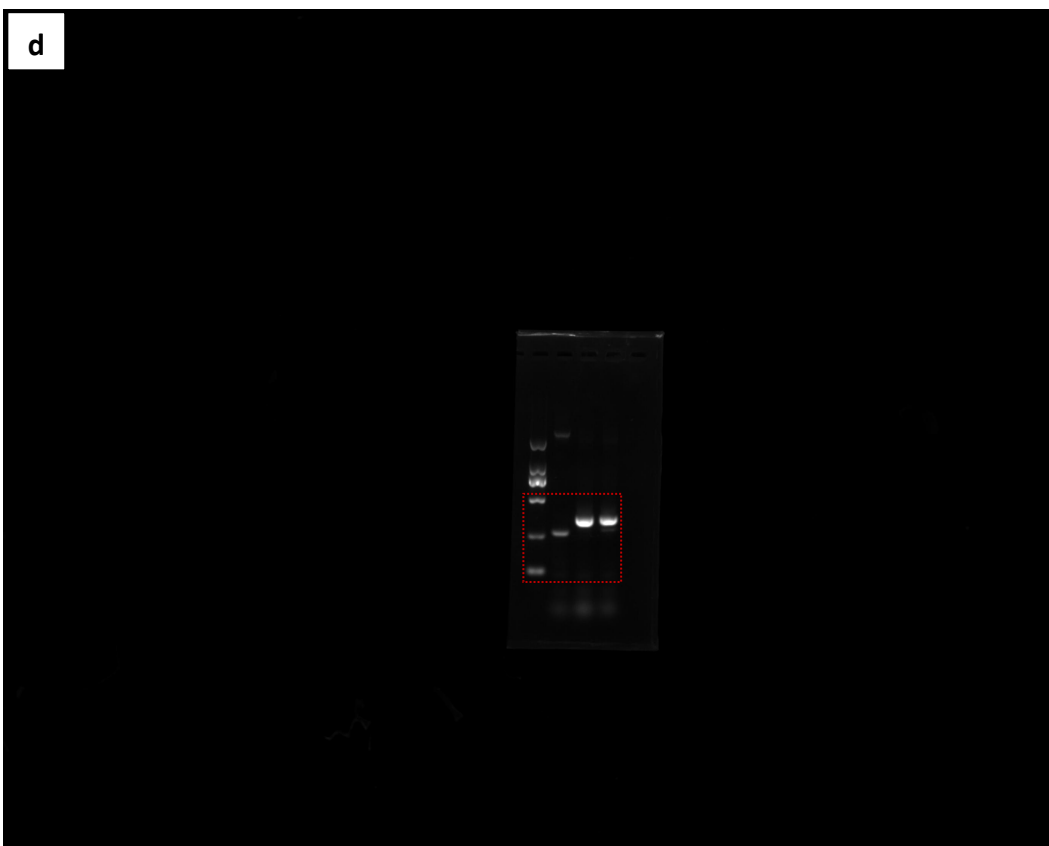
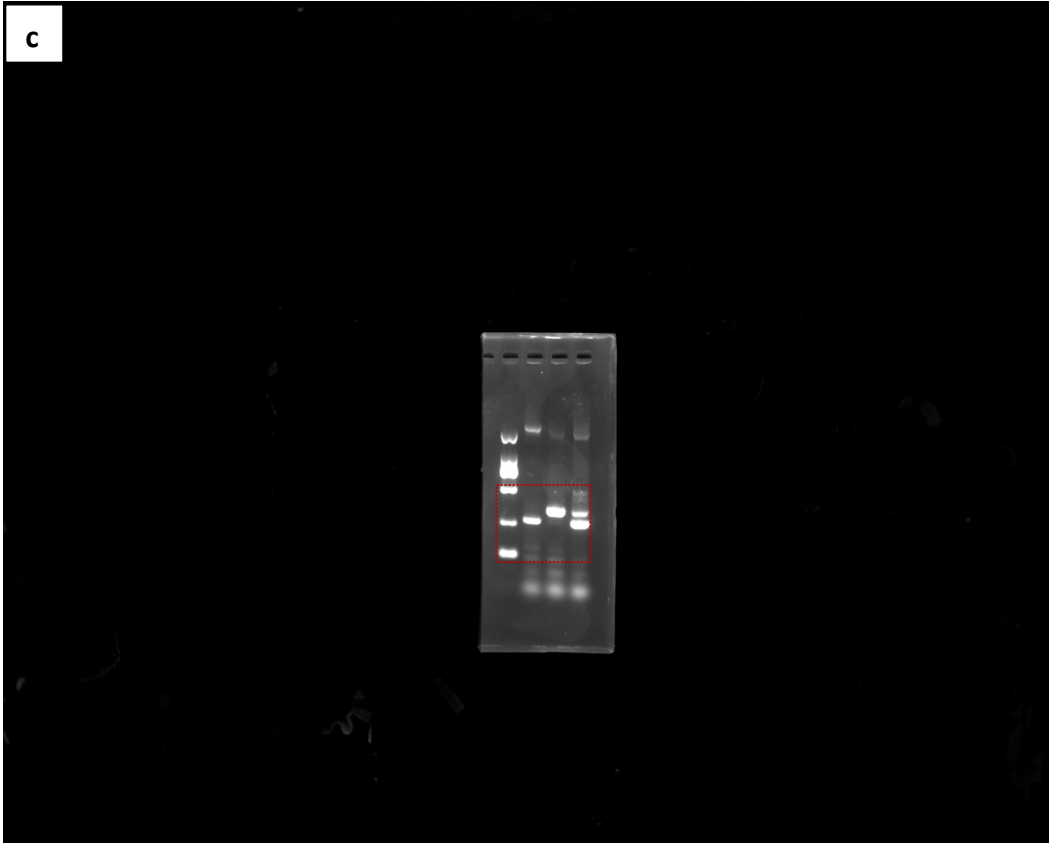
Figure S5 Diagrammatic representation of the domain organization of the collagen XVII protein

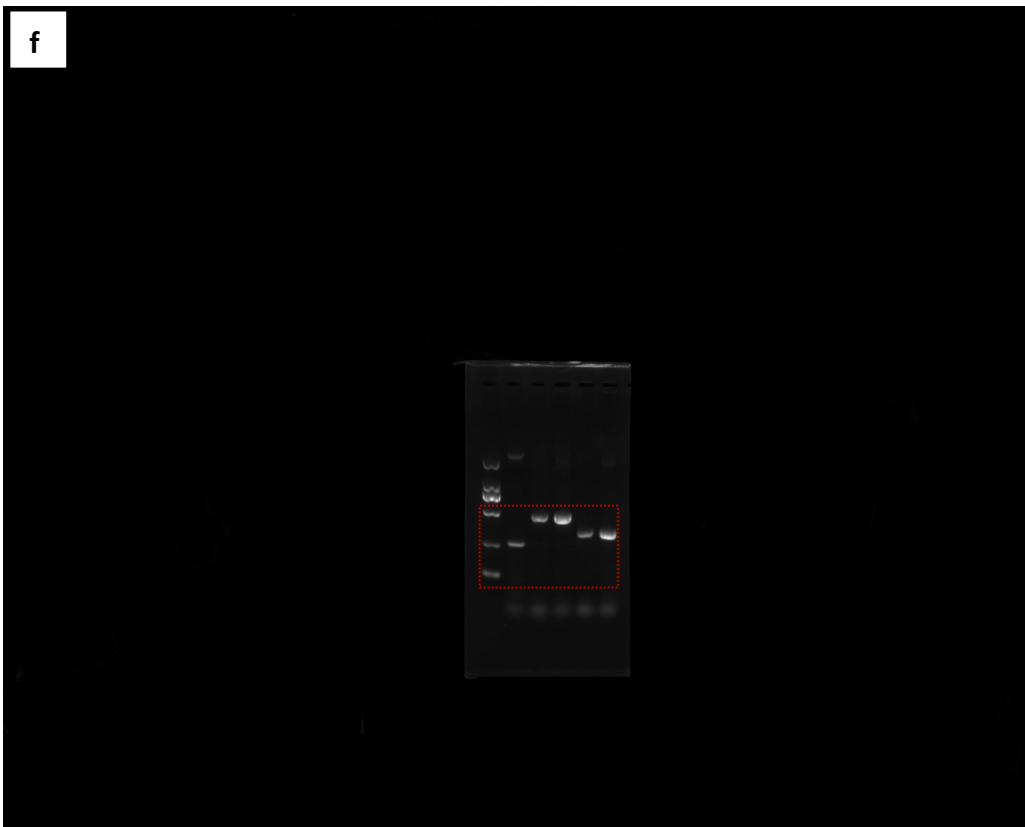


The extracellular domain consists of 15 collagenous (COL1–COL15, green vertical boxes) and flanked by stretches of non-collagen sequence (NC1–NC16a, gray horizontal lines). The non-collagenous domain, NC16, extends from the extracellular matrix to the cytoplasm and includes a transmembrane segment, NC16b, flanked by NC16a at the C-terminal end and NC16c at the N-terminal end. The *COL17A1* variants in this study are denoted in orange text.

Figure S6 Uncropped gel images for Figure 1.







Red dashed box: cropped gels.