Supporting Online Material for

Mitotic Recombination in Patients with Ichthyosis Causes Reversion of Dominant Mutations in *KRT10*

Keith A. Choate, Yin Lu, Jing Zhou, Murim Choi, Peter M. Elias, Anita Farhi, Carol Nelson-Williams, Debra Crumrine, Mary L. Williams, Amy J. Nopper, Alanna Bree, Leonard M. Milstone, and Richard P. Lifton‡

‡ To whom correspondence should be addressed. E-mail: richard.lifton@yale.edu

This PDF file includes:

Materials and Methods:

Human Subjects. Patients were referred for evaluation of ichthyosis with confetti. Patients and participating family members provided consent to the study protocol which was approved by the Yale Human Investigation Committee. Affected subjects provided a venous blood sample and biopsies of affected and revertant skin. Unaffected relatives provided a blood sample.

Keratinocyte culture. Skin biopsy samples were obtained and placed in DMEM (Invitrogen) with 2X penicillin/streptomycin (Invitrogen) for transport. Samples were incubated overnight in 1X dispase. Keratinocytes were isolated from the resulting epidermal sheet using 0.05% trypsin - 2% EDTA and were plated onto a feeder layer of mitomycin C - treated 3T3 cells in DMEM/F12 medium. When individual colonies of keratinocytes were well-established, cells were split and propagated in EpiLife (Invitrogen) (1, 2).

Genomic DNA preparation and analysis. Genomic DNA was isolated from venous blood and cultured disease and revertant keratinocyte cultures via standard methods. SNP genotyping was performed on the Illumina platform using CNV370-Duo chips following manufacturer's protocols. Results from SNP genotyping enabled calculation of the maximum likelihood interval for the position of the disease gene. We specified that sites of mitotic recombination are marked by the onset of LOH on 17q, that they occur proximal to the location of the disease locus, and that their distribution is stochastic, as supported by their individually distinct positions between the centromere and 34.5 Mb on 17q. From these considerations one can calculate a confidence interval for distal boundary of the disease locus from the equation: $L_x = 1 - (\theta_{\text{breakpoint}} / \theta_{\text{locus}})^n$ where L_x is the likelihood that the gene lies within the interval between the location of the most distal observed mitotic recombination site and a more distal site at position x, $\theta_{\textbf{breakpoint}}$ is the distance between the centromere and the most distal site of mitotic recombination seen among revertants, $\theta_{\textbf{locus}}$ is the distance between the centromere and position x, and n represents the number of LOH events observed. This calculation can assume either that the mitotic recombination map is proportional to the meiotic recombination map (in which case x is measured in cM), or that the mitotic map is proportional to physical distance on 17q (in which case x is measured in Mb). Assuming the latter, the 99.9% confidence interval for the location of the ichthyosis with confetti gene spans the 3.2 Mb interval from 34.5 to 37.7 Mb on the

physical map of chromosome 17q. Results using the meiotic recombination map yield a disease gene location that is virtually identical.

Overlapping primer sets tiling across the chromosome 17 interval from 34.5 to 37.7 MB were generated using an hg18 human genome reference masked with SNPmasker and GenomeMasker (3) and the web interface for Primer3 (4) was employed for primer design to achieve 8500 bp amplicons with 200 bp overlap. Long range PCR was performed using KAPA LongRange polymerase with 1M betaine (Sigma). Primer sets were optimized on genomic control DNA using 50^{o}, 55^{o}, 60^{o}, and 65^{o} C annealing temperatures, and any sets failing these conditions were successively re-designed by division into smaller amplicons with halved target size until a complete tiling path across the interval was established. 225 primer sets were ultimately used to produce overlapping amplicons across the interval from the unaffected parents and affected offspring of kindred 107. PCR products were purified using Ampure beads (Beckman), quantitated using PicoGreen (Invitrogen) and pooled in equimolar amounts. 5 μg of pooled DNA was used as input material for library preparation and 75 base single end sequencing on the Illumina GAIIx instrument. The Maq program was used to map Illumina reads to the hg18 reference genome (5). Reads that mapped to multiple locations and base calls with raw quality scores <45 were discarded. Reads mapping to the critical interval were retained for analysis and Perl scripts were used to generate statistics on coverage. Genotype calls and quality scores were annotated with SAMtools (6). All bases deviating from the reference sequence were noted. Using an automated pipeline, variants were annotated for novelty using dbSNP (build 130) and data from four personal genomes (7). A Perl script was then used to compare variants between the affected subject and both parents and variants confirmed to be present only in the affected offspring were further examined.

Sanger sequencing. The Primer3 web interface was used to generate primers for amplification of the entire *KRT10* coding region by PCR. DNA from affected subjects, family members, and control samples was used as a template. Products were analyzed using Sanger sequencing of both strands. Resulting traces were analyzed using Sequencher 9.0 (Genecodes).

cDNA preparation and analysis. Total cellular RNA was isolated from keratinocytes using RNeasy (Qiagen). Reverse transcription was performed using Superscript II (Invitrogen) with oligo-dT priming per standard protocols. *KRT10*-specific products were produced by PCR using KAPA HiFi polymerase and specific primers in the 5' and 3' UTRs and products were TA cloned

into pCR2.1TOPO vector (Invitrogen). Plasmid DNA was isolated from the resulting clones and subjected to Sanger sequencing. At least 4 independent mutant cDNAs were identified for each subject. Full-length wild-type and mutant cDNA were cloned into pCDNA3.1(+) for expression studies in PLC cells.

Western analysis. Wild-type keratinocytes from a healthy 22 year-old male donor and mutant and revertant keratinocytes from affected subjects were grown on mitomycin-C treated 3T3 cell feeder layers in the presence of 1.2 mM calcium for 10 days to induce differentiation. Cells were washed twice with PBS and lysed by sonication in 50mM Tris HCl PH 7.5/1% SDS/1X protease inhibitor. Insoluble material was pelleted by centrifugation. Protein concentration was determined by Bio-Rad DC protein assay (500-0114, Bio-Rad). 8 μg of normal control and revertant keratinocyte, and 20 ug of diseased keratinocyte protein extract were loaded per lane of a 12% Bis-Tris gel (Invitrogen NuPAGE NP0341). Protein was transferred onto PVDF membrane and blocked with 5% nonfat milk overnight at 40^{o}C. Blots were probed with 1:100 mouse anti-human keratin 10 whose epitope has been mapped to the amino terminal half of the protein ([8]) (LH2, Santa Cruz Biotechnology, sc-53252) and 1:2000 mouse anti-ß actin (Sigma A-2066) primary antibodies, washed, and then probed with peroxidase conjugated sheep anti-mouse IgG (GE Healthcare RPN4201V) at 1:10,000 dilution. The blot was washed and developed with ECL reagents (GE Healthcare RPN2106V) (2).

Transfections. PLC cells (CRL-9204, ATCC) were seeded at a density of 2 x 10^5 cells per well of 4-well culture slides in serum-free medium. Cells were transfected using Lipofectamine 2000 (Invitrogen) per standard protocols and fixed for immunofluorescence studies at 48 hours post-transfection.

Immunofluorescence microscopy. Tissue samples were embedded in OCT medium and 5 micron sections were cut. Tissue sections or transfected cells grown on 4-well chamber slides (Nuncalon) were fixed with 2:1 methanol/acetone at -20˚C for 7 minutes. Blocking was performed with 10% donkey serum/1% BSA for 1 hour at room temperature. Slides were incubated with primary antibody for 1 hour, washed 4 times with 1X PBS, incubated 1 hour with secondary antibody, and again washed 3 times with 1X PBS before mounting with Mowiol/1% n-propyl gallate (Polysciences). Primary antibodies used include: 1:100 mouse anti-human KRT10 (LH2, Santa Cruz Biochemicals), 1:100 rabbit anti-human fibrillarin (AB5821, Abcam), mouse anti-human KRT1 (34ß34, Covance). Secondary antibodies used include CY3 donkey

anti-rabbit IgG and CY2 donkey anti-mouse IgG (Jackson Immunoresearch). DAPI was used as a nuclear counterstain (Santa Cruz Biochemicals).

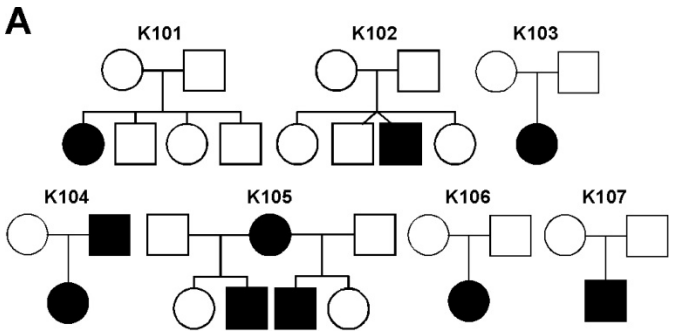

Fig. S1. Ichthyosis with confetti kindreds. The structure of 7 kindreds studied is shown. In 5 there is a single affected subject who is the offspring of unaffected parents, while in two affected subjects are the offspring of one affected parent.

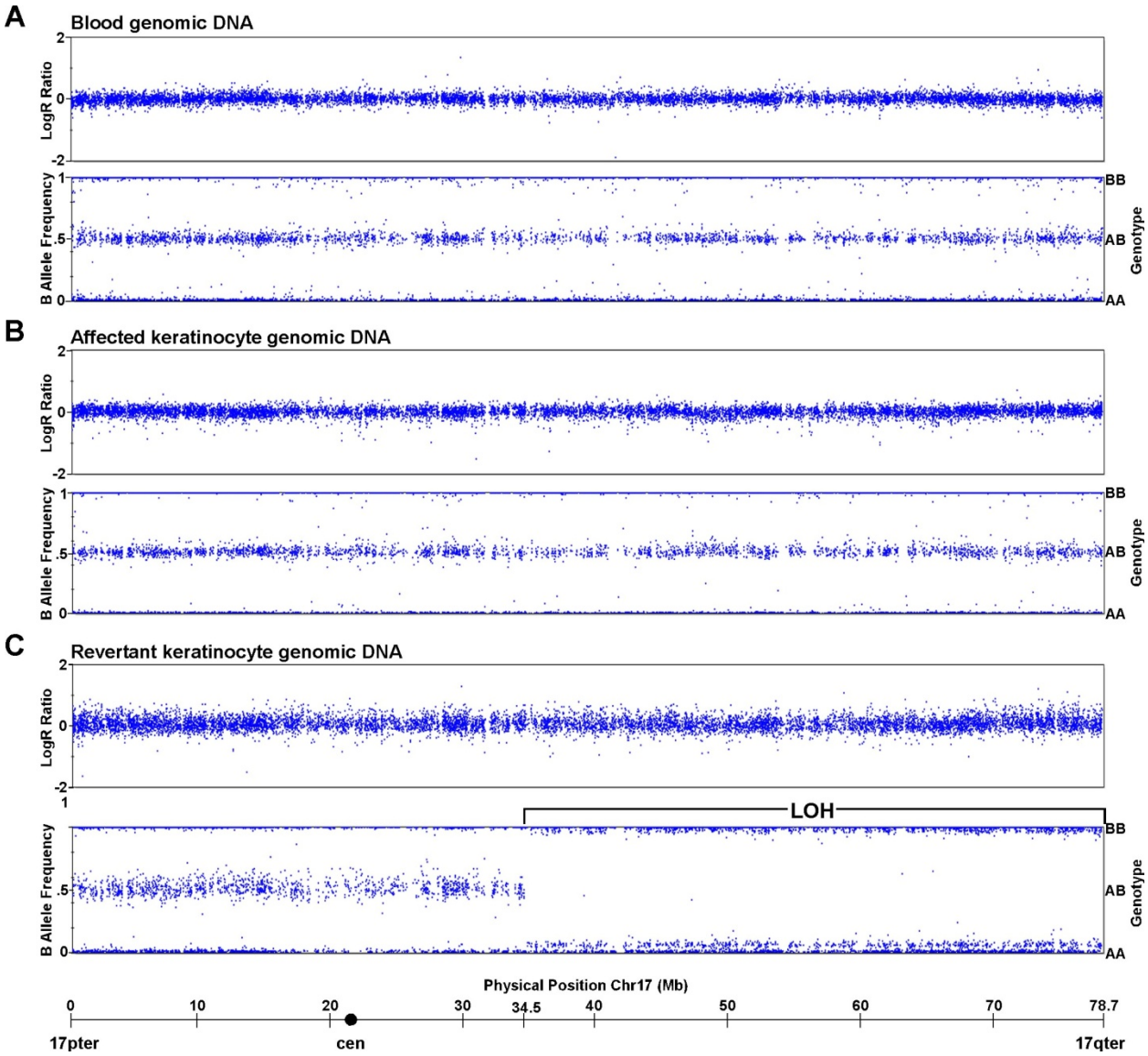

Fig. S2. Revertant spots show loss of heterozygosity on 17q. In each panel the log R ratio of the SNP signal intensity and the genotypes of SNPs along chromosome 17 in subject 106-1 are shown using Illumina 370K genotyping. (**A**) Results from blood; (**B)** Results from diseased keratinocytes; (**C**) Results from revertant keratinocytes. The revertant keratinocytes show LOH across 43 Mb, but no change in copy number. Similar results were obtained for 31 additional blood-revertant keratinocyte pairs. These long segments of copy-neutral LOH are consistent with an origin via mitotic recombination.

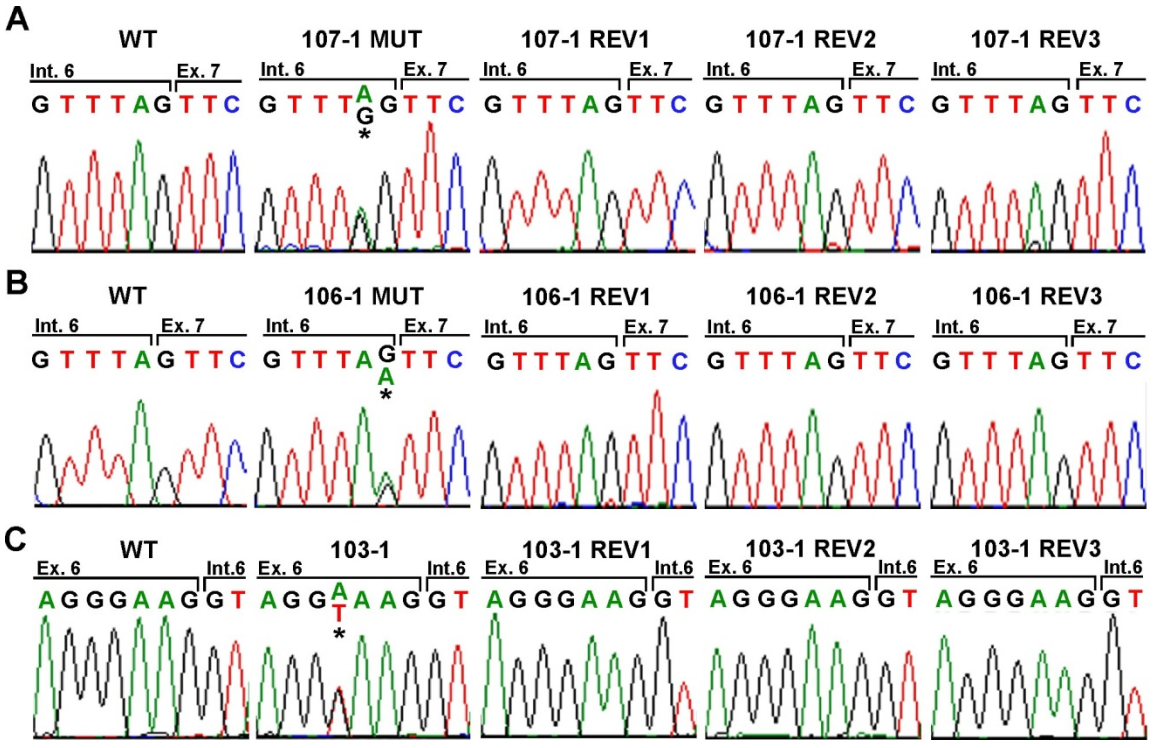

Fig. S3. Revertant skin shows loss of the mutant allele. Keratinocytes were cultured from revertant skin. In each subject, the disease-causing mutation identified in *KRT10* from blood DNA by Sanger sequencing is shown, along with absence of the mutation in keratinocytes from independent revertants. (**A**) Sequence from an unrelated wild-type sample, 107-1 blood DNA, and three revertant keratinocyte clones shows loss of the mutant G allele in each of the revertant clones. (**B, C**) show loss of mutant allele in revertants from subjects 106-1 and 103-1.

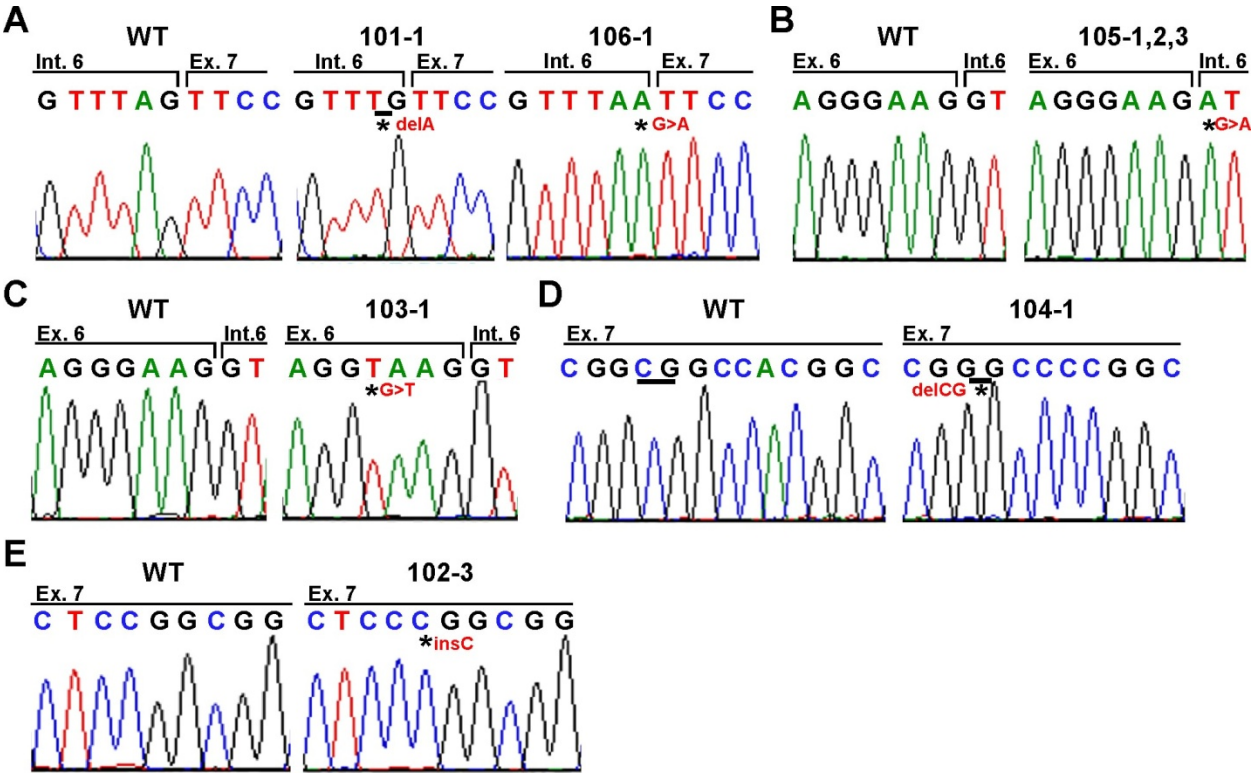

Fig. S4. Mutations in *KRT10* cause ichthyosis with confetti. Indicated segments of *KRT10* from patients with IWC were amplified by PCR, and mutations detected by direct sequencing of the PCR product. Because many of these heterozygous mutations are insertion-deletions, the mutant sequences are shown as TA-cloned products for clarity. In kindreds 101, 106, 103 and 102, both parents and the affected offspring were sequenced, revealing the mutations in affected subjects to be *de novo*. (**A**) Sanger sequencing of kindreds 101 and 106 shows *de novo* mutations in the canonical intron 6 splice acceptor site. (**B**) Sequencing of affected individuals in kindred 105 shows mutation in the canonical intron 6 splice donor site. (**C**) A *de novo* G→T substitution in exon 6 of 103-1 generates a novel splice donor site and that leads to a 5bp deletion in the encoded cDNA. (**D**) In kindred 104, a two base (CG) deletion within exon 7 arises on the background of a private A→C polymorphism. (**E**) In kindred 102, a *de novo* single base (C) insertion within exon 7 is present. All mutations lead to frameshifts into the

same open reading frame encoding an arginine-rich carboxy terminal frameshift peptide (Fig. 3D).

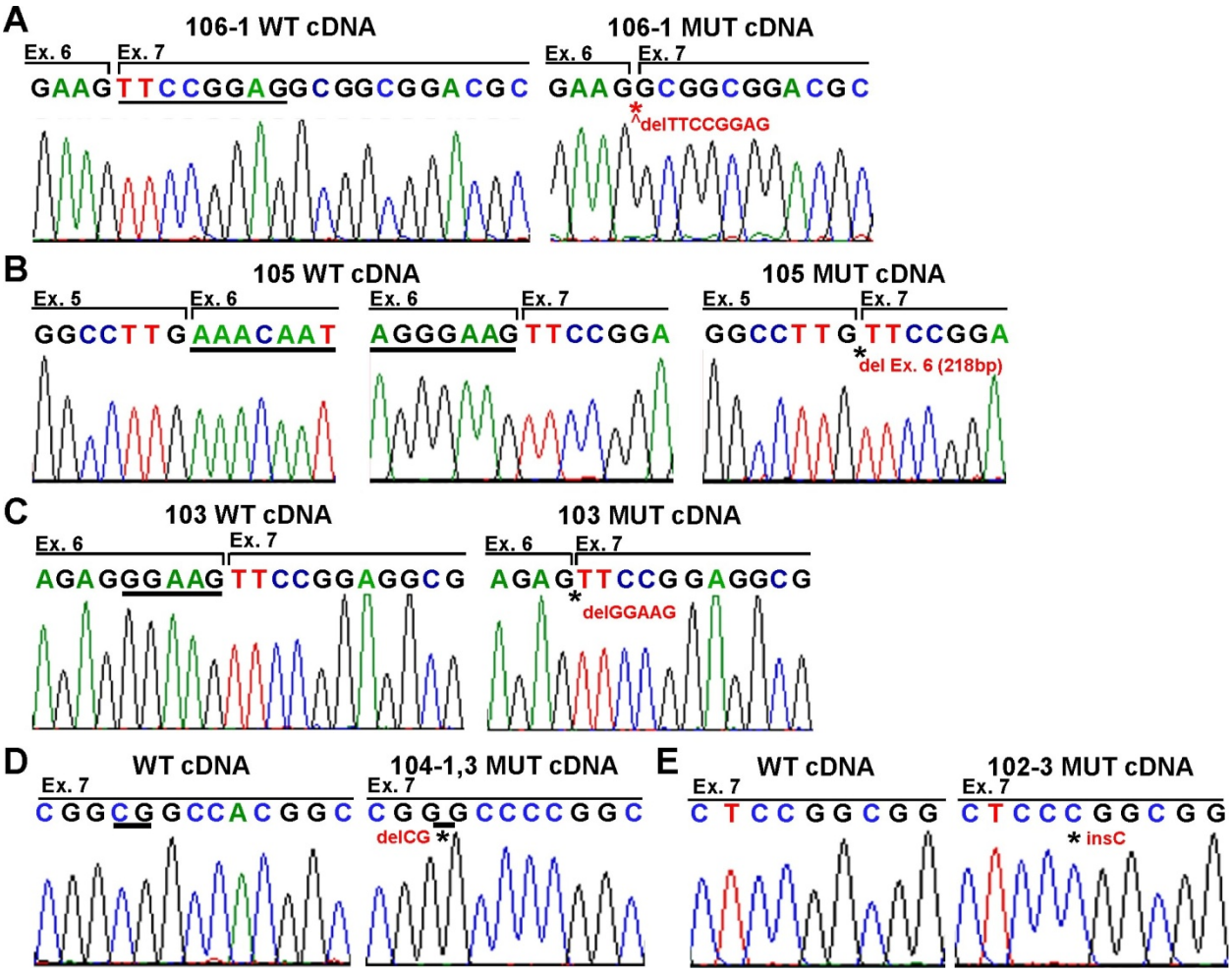

Fig. S5. Mutations in *KRT10* result in frameshift mutations. cDNA from diseased keratinocytes was prepared, PCR was performed using *KRT10*-specific primers, and the resulting product was TA cloned. Unlike the products of normal subjects or revertant keratinocytes, sequencing of independent clones from diseased keratinocytes showed 2 splice forms. (**A**) In 106-1, mutation of the canonical splice acceptor site of intron 6 (TAG to TGG), results in use of an alternative AG splice acceptor site, producing an 8 base deletion in the cDNA and leading to a frameshift mutation. (**B**) In kindred 105, the GT→AT mutation leads to skipping of exon 6 with a junction resulting between exons 5 and 7. (**C**) In 103-1, the G>T substitution at c1369 results in a novel exon 6 splice donor, use of which leads to deletion of 5 bp. (**D**) In kindred 104, a CG

deletion within exon 7 arises on the background of a private A→C polymorphism. (**E**) In kindred 102, a C insertion within exon 7 is present. All mutations result in frameshift into the same alternative reading frame, producing a very arginine-rich C-terminal peptide.

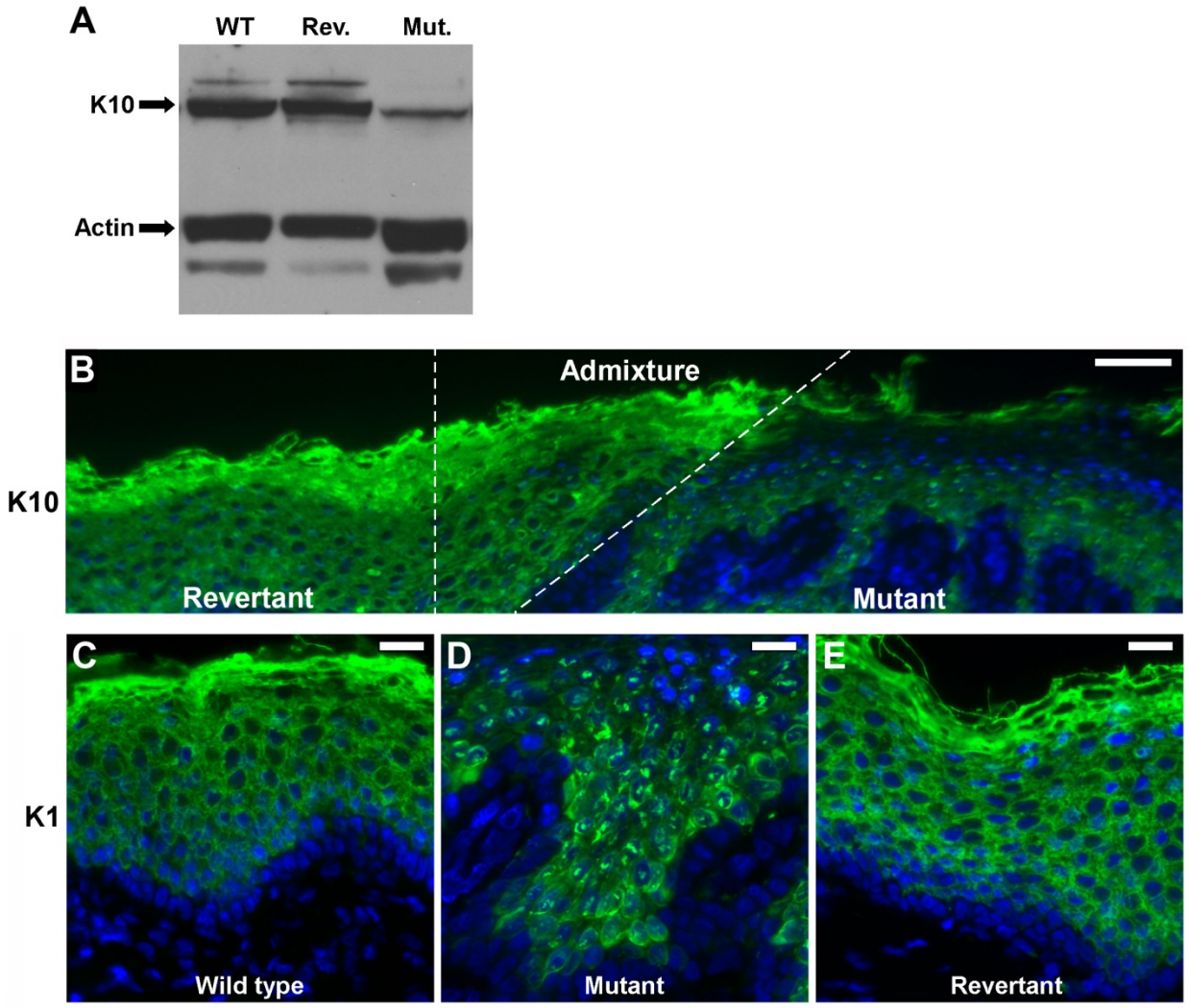

Fig S6. Reduced expression and mislocalization of keratins in IWC. (**A)** Mutant, but not revertant, keratinocytes express reduced levels of keratin 10. Mutant and revertant cultured keratinocytes from subject 106-1 and an age-matched normal control were grown in the presence of 1.2 mM calcium on a 3T3 cell feeder layer, harvested, and subjected to SDS-PAGE electrophoresis. The resulting blot was probed with antibodies to keratin 10 and ß-actin was used as a protein loading control. While wild-type and revertant cells express virtually identical levels of keratin 10, mutant cells express significantly lower levels. (**B**) Keratin 10 staining of skin from a biopsy across the junction of mutant and revertant skin from kindred 102. Keratin 10 was stained with monoclonal antibody LH2, nuclei with DAPI and the results visualized with

immunofluorescence microscopy. In the revertant skin, keratin 10 is expressed in suprabasalar layers; no keratin 10 is detected in the nucleus. In mutant skin, keratin 10 levels are lower overall and discrete staining within the nucleus is detected. Scale bar = 100 μm. (**C-E**) Mislocalization of keratin 1 in IWC. Higher power images of normal, mutant and revertant skin stained with DAPI and antibodies to keratin 1 reveal discrete nuclear foci of keratin 1 in mutant skin which are absent in normal and revertant skin. Scale bars = 10 μm.

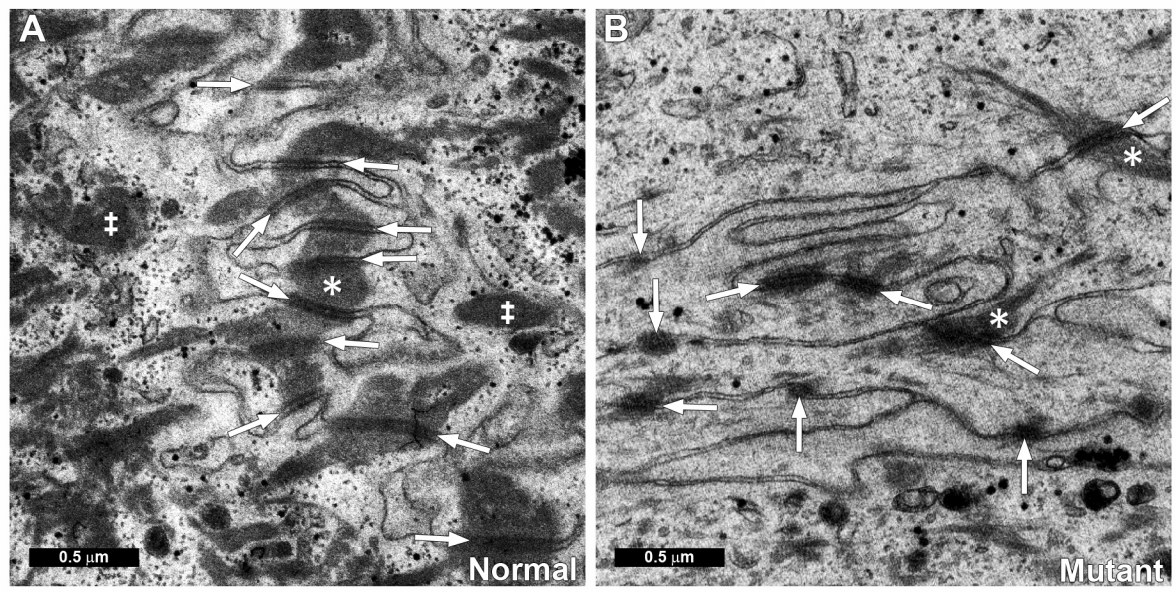

Fig. S7. Electron microscopy of wild-type and IWC spinous layer skin. (**A**) In normal skin, desmosomes (white arrows) from neighboring cells are invested with dense bundles of keratin filaments which are seen in longitudinal sections (*) and *en face* (‡). (**B**) In diseased IWC skin, the cytosol is relatively devoid of keratin filaments, desmosomes are poorly formed and filament bundles attached to desmosomes are sparse and tapered. Many desmosomes lack intermediate filaments entirely. Scale bar = 0.5 μm.

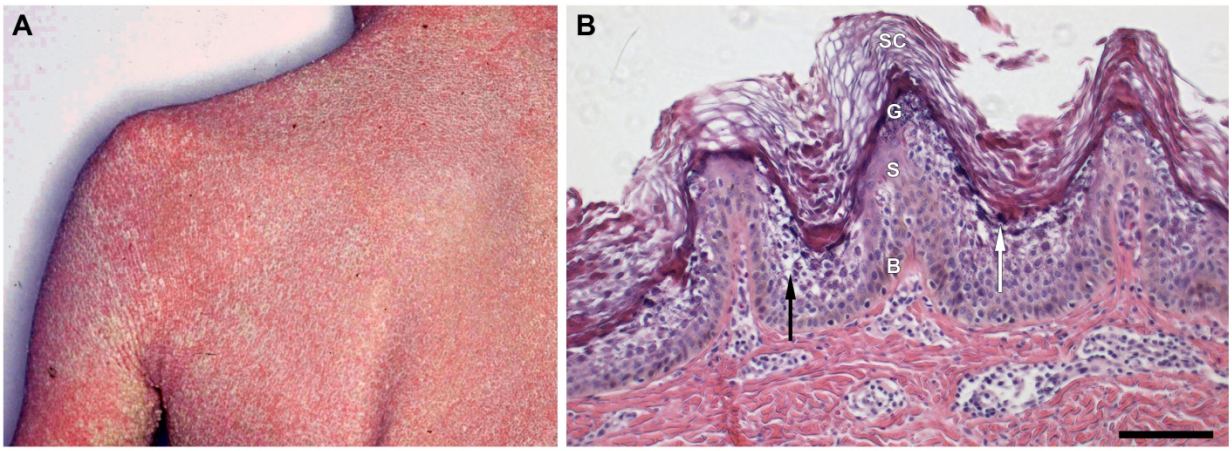

Fig. S8. Epidermolytic ichthyosis (EI) has a clinical appearance and histology distinct from ichthyosis with confetti. (**A**) EI is characterized by skin fragility and blistering with hyperkeratosis, variable palmoplantar keratoderma, and background erythema without development of patches of normal-appearing skin. Shoulder of a young male with EI showing erythema and thick hyperkeratosis. (**B**) The histology of EI is characterized by acanthosis, hyperkeratosis, hypergranulosis with clumping of keratohyalin granules (white arrow), and epidermolysis with separation between cells developing in the upper spinous and granular layer (black arrow). This is in contrast to IWC which shows no epidermal fragility, absence of a granular layer, and parakeratosis (retention of nuclei in the stratum corneum). Basal layer (labeled 'B'), stratum spinosum ('S'), granular layer ('G') and stratum corneum ('SC') are denoted. Scale bar = 50 μm.

Tables:

Table S1. *De novo* mutation in IWC subject 107-1. Illumina sequencing reads from PCR amplicons spanning the IWC critical interval were screened for mutations present in the affected subject that were absent in both parents. A single heterozygous *de novo* mutation (T to C mutation) was detected, at base 36,228,941 on chromosome 17 (build Hg18). The high Samtools quality score of this mutation provided high likelihood that the sequence of the affected patient differs from reference. This base lies in the splice acceptor sequence of intron 6 of *keratin 10*, which has 5' to 3' orientation directed toward 17p. In the 5' to 3' orientation of *KRT10*, this mutation changes the wild-type 'A' of the 'TAG' splice acceptor site, to 'G' (TGG). There was no evidence of this mutation in either parent. This mutation was confirmed by Sanger sequencing (Fig. 3A).

	Inferred	# of reads with each base				Samtools
	genotype	T (WT)	C (Mutant)	A	G	quality score
Father	T/T	43	0	0	0	0
Mother	T/T	38	0	0	0	0
Affected	C/T	13	19	0	0	121

Table S2. Mutations in *KRT10* in IWC kindreds. For each affected individual, mutation, expected cDNA consequence, and predicted protein change is shown. All mutations affect exon 7, leading to frameshift into the same arginine-rich reading frame. * = *de novo* mutation

Affected	Mutation	Consequence	Predicted Protein Change
107-1*	intron 6 splice acceptor TAG→TGG	New splice acceptor → 8 base deletion in cDNA	del S459→G460, S458R fs x 120
101-1*	intron 6 splice acceptor TAG→TG	New splice acceptor → 8 base deletion in cDNA	del S459→G460, S458R fs x 120
106-1*	intron 6 splice acceptor TAG→TAA	New splice acceptor → 8 base deletion in cDNA	del S459→G460, S458R fs x 120
105-1,2,3	intron 6 splice donor GT→AT	Exon 6 skipped in cDNA, frameshift at codon 458	del K386→S458, G459F fs x 122
102-3*	exon 7 c.1450insC	Frameshift at codon 484	G484R fs x 97
104-1,3	exon7 c.1560delCG	Frameshift at codon 521	G521P fs x 59
103-1*	exon 6 c.1369 G→T	New splice donor → 5 base deletion in cDNA	del S458, G457F fs x 122

Supporting Online Material References and Notes:

S1. K.A. Choate, D.A. Medalie, J.R. Morgan, P.A. Khavari, Corrective gene transfer in the human skin disorder lamellar ichthyosis. *Nature Med.* **2**, 1259 (1996).

S2. J. Zhou, J. G. Haggerty, L. M. Milstone, Growth and differentiation regulate CD44 expression on human keratinocytes. *In Vitro Cell. Dev. Biol.* **35**: 228 (1999).

S3. R. Andreson, T. Puurand, M. Remm, SNPmasker: automatic masking of SNPs and repeats across eukaryotic genomes. *Nucleic Acids Res*. **34**, W651 (2006).

S4. S. Rozen and H. J. Skaletsky, Primer3 on the WWW for general users and for biologist programmers. In: S. Krawetz, S. Misener (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* (Humana Press, Totowa, NJ, 2000) pp 365-386.

S5. H. Li, J. Ruan, R. Durbin, Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* **18**, 1851 (2008).

S6. H. Li, *et al.*, The sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*. **25**, 2078 (2009).

S7. M. Choi, *et al.*, Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl. Acad. Sci. USA*. **106**, 19096 (2009).

S8. J. Reichelt, *et al.*, Out of balance: consequences of a partial keratin 10 knockout. *J. Cell Sci.* **110**, 2175 (1997).