Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 13069–13074, November 1997
Genetics

The Tabby phenotype is caused by mutation in a mouse homologue of the *EDA* gene that reveals novel mouse and human exons and encodes a protein (ectodysplasin-A) with collagenous domains

ANAND K. SRIVASTAVA[*†], JOHANNA PISPA[‡], ANDREW J. HARTUNG[*], YANGZHU DU[*], SINI EZER[§], TED JENKS[*], TOKIHIKO SHIMADA[¶], MAIJA PEKKANEN[‡], MARJA L. MIKKOLA[‡], MINORU S. H. KO[¶], IRMA THESLEFF[‡], JUHA KERE[§], AND DAVID SCHLESSINGER[‖]

[*]J. C. Self Research Institute of Human Genetics, Greenwood Genetic Center, Greenwood, SC 29646; [‡]Institute of Biotechnology and [§]Department of Medical Genetics, Haartman Institute, 00014 University of Helsinki, Helsinki, Finland; [¶]Center for Molecular Medicine and Genetics, Wayne State University, Detroit, MI 48202; and [‖]Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110

Communicated by Albert de la Chapelle, University of Helsinki, Helsinki, Finland, October 3, 1997 (received for review August 13, 1997)

ABSTRACT **Mouse Tabby (*Ta*) and X chromosome-linked human *EDA* share the features of hypoplastic hair, teeth, and eccrine sweat glands. We have cloned the *Ta* gene and find it to be homologous to the *EDA* gene. The gene is altered in two *Ta* alleles with a point mutation or a deletion. The gene is expressed in developing teeth and epidermis; no expression is seen in corresponding tissues from *Ta* mice. *Ta* and *EDA* genes both encode alternatively spliced forms; novel exons now extend the 3′ end of the *EDA* gene. All transcripts recovered have the same 5′ exon. The longest *Ta* cDNA encodes a 391-residue transmembrane protein, ectodysplasin-A, containing 19 Gly-Xaa-Yaa repeats. The isoforms of ectodysplasin-A may correlate with differential roles during embryonic development.**

Ta was the first inherited mouse epidermal deficiency to be identified and the first determined by "a totally sex-linked gene" (1). The characteristic hair defects, tooth abnormalities, and eccrine sweat gland morphology in *Ta* mice (1–3) have striking counterparts in a human X chromosome-linked disease, anhidrotic ectodermal dysplasia (EDA; MIM 305100; ref. 4). The syntenically equivalent positions of the *EDA* and *Ta* genes (see Fig. 1*A*) supported their likely identity (5–9). More direct analysis of *Ta* and *EDA* became more feasible recently with the isolation of the human *EDA* gene (10, 11). The gene is normally expressed in keratinocytes and glandular structures in skin and is mutated in various patients by nine different point mutations as well as translocations and deletions (11).

Using a human EDA cDNA as a hybridization probe, we have now recovered corresponding mouse cDNAs and adjoining genomic sequence for the *Ta* gene. It is confirmed to be the homologue of *EDA*. It is comparably expressed in ectodermal tissues early in mouse embryonic development and loses expression by mutation in each of two different *Ta* strains. Three alternatively spliced *Ta* gene transcripts, encoding ectodysplasin-A, were recovered from wild-type mice. Several exons are included that have now been found in EDA transcripts as well, although they had not been seen in earlier experiments (11). The longest cDNA species encodes collagen-like repeats and a cysteine-rich region in a 3′ exon. This is consistent with other recent findings that the EDA protein may be involved in cell–matrix interactions (12) critical for the formation of skin.

MATERIALS AND METHODS

Tabby Alleles and Cell Lines. DNAs for tabby alleles were from The Jackson Laboratory. These are from Tabby strain B6CBACa-A⟨w-j⟩/A-Ta jp/++ (genotype *Ta jp/Y*) and the respective control male mice B6CBACa-A⟨w-j⟩/A-*Ta jp*/++ and from Tabby ⟨6J⟩ strain C57BL/6J-A⟨W-J⟩-*Ta*⟨*6J*⟩ (genotype *Ta*⟨*6J*⟩/*Y*) and the respective control male C57BL/6J-A⟨W-J⟩ *Ta*⟨*6J*⟩ +/Y. Other control mouse DNAs used in this study were BALB/cBy, 3T6, and strain C3H/HeJ +/Y.

Mouse cDNA and Genomic Clones. A mouse 17-day embryo 5′ stretch plus cDNA library (CLONTECH) was screened by hybridization with a human EDA cDNA clone (11). The inserts of five recovered cDNAs were amplified with λgt11 vector primers, subcloned into the pcr2.1 vector (Invitrogen), and sequenced. Sequencing was done manually by using the *fmol* DNA cycle sequencing system (Promega). An additional cDNA (clone D85) was obtained from Research Genetics (Huntsville, AL) and sequenced. Ta A (5,004 bp) includes two overlapping cDNAs, 7–19-0 (3,995 bp) and D85 (1,372 bp); Ta B (1,879 bp) includes two identical clones, 7–19-3 and 7–25-3; and Ta C (1,562 bp) includes clone 7–25-2.

By using a 121-bp mouse PCR product amplified with the primers 5′-CGGTTAAAAGTTGTTTCAACGG and 5′-CGGTGCTTGTTTTTTCGTGTC, corresponding to nucleotides 358–479 of the published *EDA* genomic sequence (11), as a hybridization probe, genomic λ clones containing exon 1 were isolated from a 129Sv mouse genomic library (Stratagene). Plasmid subclones of the λ DNA insert were sequenced on an Automated Pharmacia LKB A.L.F. DNA Sequencer. Genomic and cDNA sequences were analyzed by BLAST and FASTA against GenBank sequences.

DNA and RNA Analysis. PCR assays used *Taq* polymerase (Boheringer Mannheim) and initial denaturation at 95°C for 150 s followed by 35 cycles of 95°C for 30 s, 65°C (57°C for exon 3) for 30 s, and 72°C for 30–60 s. Genomic DNA (25–50 ng/20 μl of reaction mixture) was used as template. Rapid amplification of 3′ cDNA ends was done by nested PCR on poly(A)-primed cDNA by using a human placenta marathon-ready cDNA kit (CLONTECH) with manufacturer's instructions. Nested cDNA specific primers were designed from the 3′ end of human EDA exon 1 (11). The primer in the first reaction was 5′-CCTAAGCAGCAGCCATTGGAACCGG and in the second reaction was 5′-AGCCGCACTCCACTCTGACTCCCAG. A 510-bp product was cloned into the pGEM-T vector (Promega) and sequenced.

Exons 1 and 3 of the Ta gene were screened for mutations by using restriction endonuclease fingerprinting analysis (13). Approximately 100 ng of 589-bp PCR product for exon 1

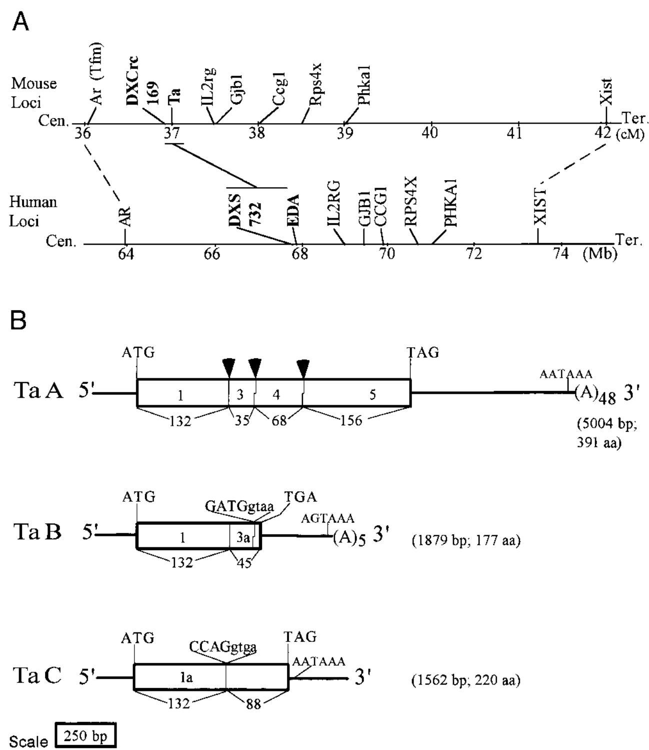
The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1997 by The National Academy of Sciences 0027-8424/97/9413069-6$2.00/0
PNAS is available online at http://www.pnas.org.

Abbreviations: d.p.c., days post coitum; EGF, epidermal growth factor.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF016627–AF016632 and Y13438).

[†]To whom reprint requests should be addressed at: J. C. Self Research Institute of Human Genetics, Greenwood Genetic Center, 1 Gregor Mendel Circle, Greenwood, SC 29646. e-mail: anand@ggc.org.



D

```
Mouse  MGYPEVERRE PLPAAAPRER GSQGCGCRGA PARAGEGNSC 40
Human  .......... L......... .......G.. .......... 40

Mouse  RLFLGFFGLS LALHLLTLCC YLELRSELRR ERGTESRLGG 80
Human  L......... .......... .......... ...A...... 80

Mouse  PGAPGTSGTL SSPGSLDPVG PITRHLGQPS FQQQPLEPGE 120
Human  S.T....... ..L.G...DS ...S...... PK........ 120

Mouse  DPLPPDSQDR HQMALLNFFF PDEKAYSEEE SRRVRRNKRS 160
Human  AA.HS....G .......... ....P..... .......... 160

Mouse  KSGEGADGKS TQVIFFP                          177
Human  ..N....... ..L.LYHF                         178
```

C

Ta A cDNA

```
-175      TGGAGCTGCACATGCGGCTGCTCCCTGTTCTGTCCCGCCCAGCCACCGTCGCTCA
-120 GGAACGGGTCCCTGCAGCCCCCAGCCGATGGCAGGACAGTAGTCGCCTGTCAGGGGTCGT
-60  GAAGGACTGAGGCAGAGGCAGAGGCTCCCGGGCCTCAGATAGTGGTTGTCTCTGGAGGCC
1    ATGGGCTACCCAGAGGTAGAGCGCAGGGAACCCCTGCCTGCGGCAGCGCCAAGGGAGCGG
1     M  G  Y  P  E  V  E  R  R  E  P  L  P  A  A  A  P  R  E  R

61   GGCAGCCAGGGCTGCGGCTGTCGCGGGGCCCCTGCTCGGGCGGGCGAAGGGAACAGCTGC
21    G  S  Q  G  C  G  C  R  G  A  P  A  R  A  G  E  G  N  S  C

121  CGGCTCTTCCTGGGTTTCTTTGGCCTCTCGCTGGCCCTCCACCTGCTGACGCTGTGCTGC
41    R  L  F  L  G  F  F  G  L  S  L  A  L  H  L  L  T  L  C  C

181  TACCTAGAGTTGCGGTCCGAATTGCGGCGGGAACGGGGAACCGAGTCCCGCCTCGGTGGC
61    Y  L  E  L  R  S  E  L  R  R  E  R  G  T  E  S  R  L  G  G

241  CCGGGTGCTCCTGGCACCTCTGGCACCCTAAGCAGCCCTGGGAGCCTCGACCCGGTGGGT
81    P  G  A  P  G  T  S  G  T  L  S  S  P  G  S  L  D  P  V  G

301  CCCATCACCCGCCACCTGGGGCAGCCGTCCTTTCAACAGCAGCCTTTGGAGCCGGGAGAA
101   P  I  T  R  H  L  G  Q  P  S  F  Q  Q  Q  P  L  E  P  G  E

361  GATCCACTCCCCCCTGACTCCCAGGACCGGCACCAGATGGCCCTCCTGAATTTCTTCTTT
121   D  P  L  P  P  D  S  Q  D  R  H  Q  M  A  L  L  N  F  F  F

421  CCTGATGAAAAGGCATATTCTGAAGAGGAAAGTAGGCGTGTTCGCCGCAATAAGAGAAGC
141   P  D  E  K  A  Y  S  E  E  E  S  R  R  V  R  R  N  K  R  S

481  AAAAGTGGTGAAGGAGCAGATGGTCCTGTTAAAAACAAGAAAAAGGGAAAGAAGGCAGGG
161   K  S  G  E  G  A  D  G  P  V  K  N  K  K  K  G  K  K  A  G

541  CCACCTGGGCCCAACGGCCCCCCAGGACCTCCAGGACCTCCGGGACCCCAGGGACCTCCA
181   P  P  G  P  N  G  P  P  G  P  P  G  P  P  G  P  Q  G  P  P

601  GGGATTCCAGGAATTCCTGGGATTCCAGGAACAACTGTTATGGGACCACCTGGCCCACCT
201   G  I  P  G  I  P  G  I  P  G  T  T  V  M  G  P  P  G  P  P

661  GGCCCTCCTGGTCCTCAAGGACCCCCTGGCCTCCAAGGACCTTCTGGTGCTGCTGATAAA
221   G  P  P  G  P  Q  G  P  P  G  L  Q  G  P  S  G  A  A  D  K

721  ACTGGAACTCGGGAAAATCAGCCAGCTGTGGTGCATCTGCAGGGCCAAGGGTCAGCAATT
241   T  G  T  R  E  N  Q  P  A  V  V  H  L  Q  G  Q  G  S  A  I

781  CAAGTCAAAAATGATCTTTCAGGTGGAGTGCTCAATGACTGGTCTCGCATCACTATGAAC
261   Q  V  K  N  D  L  S  G  G  V  L  N  D  W  S  R  I  T  M  N

841  CCTAAGGTGTTTAAACTACATCCCCGCAGCGGGGAGCTGGAGGTACTGGTGGACGGCACC
281   P  K  V  F  K  L  H  P  R  S  G  E  L  E  V  L  V  D  G  T

901  TACTTCATCTATAGTCAGGTAGAAGTCTACTACATCAACTTCACTGACTTTGCCAGCTAC
301   Y  F  I  Y  S  Q  V  E  V  Y  Y  I  N  F  T  D  F  A  S  Y

961  GAGGTGGTGGTGGATGAGAAGCCCTTCCTGCAGTGCACCCGCAGCATTGAGACAGGGAAG
321   E  V  V  V  D  E  K  P  F  L  Q  C  T  R  S  I  E  T  G  K

1021 ACCAACTACAACACTTGCTATACTGCAGGCGTGTGCCTCCTCAAGGCCAGGCAGAAAATC
341   T  N  Y  N  T  C  Y  T  A  G  V  C  L  L  K  A  R  Q  K  I

1081 GCCGTGAAGATGGTGCACGCTGACATCTCTATCAATATGAGCAAGCACACCACCTTCTTC
361   A  V  K  M  V  H  A  D  I  S  I  N  M  S  K  H  T  T  F  F

1141 GGGGCCATCAGGCTGGGCGAAGCCCCTGCATCCTAGATTCTCCCATTCCATCCTGGCCCA
381   G  A  I  R  L  G  E  A  P  A  S  *
```

Ta B protein

```
1    MGYPEVERREPLPAAAPRERGSQGCGCRGAPARAGEGNSCRLFLGFFGLS
51   LALHLLTLCCYLELRSELRRERGTESRLGGPGAPGTSGTLSSPGSLDPVG
101  PITRHLGQPSFQQQPLEPGEDPLPPDSQDRHQMALLNFFFPDEKAYSEEE
151  SRRVRRNKRSKSGEGADgkstqviffp.
```

Ta C protein

```
1    MGYPEVERREPLPAAAPRERGSQGCGCRGAPARAGEGNSCRLFLGFFGLS
51   LALHLLTLCCYLELRSELRRERGTESRLGGPGAPGTSGTLSSPGSLDPVG
101  PITRHLGQPSFQQQPLEPGEDPLPPDSQDRHQvshlggaaaleapsparl
151  ggglglraqgtlplrakfqgrswewagvlgrgcpgqvvlgsclgssrpsp
201  vpwswkaqparaapgevwaa.
```

FIG. 1. (*A*) The location of the Tabby (*Ta*) locus on the mouse X chromosome genetic map (17) and EDA on a syntenic physical map of the human X chromosome (18). (*B*) Partial genomic organization of the *Ta* gene and corresponding transcript isoforms (Ta A, Ta B, and Ta C). Transcript size and protein length for each transcript are shown in brackets. Translated regions for exons (boxes) are drawn to scale and numbered. Note that no exon 2 corresponding to human *EDA* exon 2 has been detected in mouse. Numbers below each box represent amino acids encoded by the respective exon. Arrowheads show the locations of the exon–intron junctions in mouse genomic DNA. Split codons (see text) are shown by extensions within a box. Alternative exons (1a in Ta C and 3a in Ta B) are indicated. Potential splice site sequences are shown at 3′ ends of exon 1 and exon 3 in Ta C and Ta B transcripts. Sequence at the exon 4–intron junction (–TCTGgtgagt—) is not shown. Polyadenylylation site AATAAA or AGTAAA and poly(A) tail are indicated. (*C*) Nucleotide sequence (first 1,375 bp of 5,004-bp sequence, GenBank accession no. AF016628), predicted amino acid sequence of Ta A, and predicted peptide sequence of Ta B and Ta C. The predicted start of translation at +1 nucleotide with the in-frame stop codon (TAG) at nucleotides 1,174–1,176 yields an ORF of 1,176 bp that encodes a predicted protein of 391 amino acids in Ta A. A putative transmembrane domain is underlined, and the Gly-Xaa-Yaa collagenous repeats are in underlined italics. The locations of identified introns are indicated by arrowheads. Cysteine residues are indicated in underlined boldface. Amino acids in lowercase type are translated by in-frame adjacent intron sequences in alternatively spliced cDNAs Ta B and Ta C. (*D*) Amino acid sequence comparison of the mouse Ta B protein with its human *EDA* homologue. Identities are shown as dots.

(primers 5′-CAGATAGTGGTTGTCTCTGGAG and 5′-AACAACCTGACCTGGACAACCTCT) was digested with *Bfa*I and *Mbo*II. A 339-bp product for exon 3 (primers 5′-ATGGCCCTCCTGAATTTCTTC and 5′-CTTGTT-GAGAGTAACCTTGA) was digested with *Acc*I and *Ssp*I restriction endonucleases and analyzed essentially as in ref. 14.

PCR products with abnormally migrating bands were sequenced manually as described above.

Other primers used were as follows: for exon 3, 5′-ATGGCCCTCCTGAATTTCTTC and 5′-ACTTACCATCTGCTCCTTCAC (112-bp PCR product); for exon 5, 5′-TGACTTTGCCAGCTACGAGGTG and 5′-TCTTCACGGCGATTTTCTGC (146-bp PCR product); for exon 4–exon 5 junction, 5′-CCCACCTGGCCCTCCTGGTCCTC and 5′-CAGATGCACCACAGCTGGCT; for Ex1–Ex3, 5′-AACAGCAGCCTTTGGAACCGG and 5′-ACTTACCATCTGCTCCTTCAC (174-bp PCR product).

Northern blots (CLONTECH) of multiple mouse adult tissues and mouse embryos were hybridized with a 589-bp exon 1 fragment of Ta A (see above) or with a 613-bp noncoding 3′ terminal probe from clone 7–19-0. Filters were washed with low and high stringencies according to manufacturer's (CLONTECH) recommendations. RNA isolation from mouse embryos and reverse transcription-coupled PCR were essentially as in ref. 15.

***In Vitro* Transcription and Translation.** For *in vitro* transcription a 1.84-kb *Nco*I–*Sph*I fragment of the mouse cDNA was cloned into a modified pGEM3 vector. T7 RNA polymerase was used to transcribe the linearized template according to manufacturer's instruction (Promega). Approximately 1 μg of RNA was used for *in vitro* translation in a rabbit reticulocyte lysate by following the supplier's instructions (Promega) with [^{35}S]methionine (Amersham).

***In Situ* Hybridization.** Paraffin-embedded tissue sections were from NMRI × CBAT6T6 hybrid embryos or B6CBACa-Aw-J/A-Ta embryos. The stock JR 0314 was obtained from The Jackson Laboratories. A 398-bp cDNA fragment (nucleotides 84–481 of Ta A) was cloned as a blunt fragment into the *Sma*I sites of pGEM3 (Promega). An antisense probe was made by *in vitro* RNA transcription using T7 RNA polymerase (Promega). Hybridizations were performed as in ref. 16. Because expression was weak, [^{33}P]UTP (Amersham) was used; the exposure time was 15 days. Images were digitized with ImagePro software and edited with Adobe PHOTOSHOP and MICROGRAFX DESIGNER software.

RESULTS

Ta cDNAs and Alternate Splicing. Five cDNA clones were isolated from a mouse embryo cDNA library screened by hybridization with a human cDNA probe (clone 27G4 insert; ref. 11) containing exons 1 and 2 of the *EDA* gene (11). Insert sizes of recovered clones ranged from 0.7 kb to 4.0 kb. One additional clone (1.37 kb) was identified with the 3′ end sequence of a mouse cDNA by sequence comparison in the GenBank database. Sequence comparisons suggested that three different 3′ exons exist. The three different forms were inferred to represent alternative splicing and are termed Ta A, Ta B, and Ta C (Fig. 1*B*).

We assembled a consensus sequence of Ta A, Ta B, and Ta C by the analysis of sequences from the six mouse cDNA clones. We identified three mRNAs, 5,004 bp (Ta A), 1,879 bp (Ta B), and 1,562 bp (Ta C), encoding proteins of 391, 177, and 220 amino acids, respectively (Fig. 1 *B* and *C*) in an ORF that has an optimal initiation sequence (19). All three transcripts begin with a common 5′ untranslated region (160–175 bp) that is 92% identical to the corresponding segment of the human EDA transcript, followed by a 396-bp segment 87% identical to the coding region of human *EDA* exon 1 (11). This segment encodes 132 amino acid residues (Fig. 1*B*) and is 88% identical to the first 132 residues of the EDA protein (11), confirming the identity of the *Ta* gene (Fig. 1*D* and see below).

Novel Mouse *Ta* and Human *EDA* Exons and Gene Structure. Southern blot analysis of mouse genomic DNAs digested with *Taq*I restriction endonuclease gave six strongly hybridizing mouse X-chromosome-specific fragments with a 3.60-kb Ta A cDNA probe that contains two *Taq*I restriction sites, suggesting the presence of at least four *Ta* exons.

A complete alignment between the three transcript isoforms (Ta A, Ta B, and Ta C) and the published *EDA* gene sequence (11) identified a common exon 1 region and revealed a tentative genomic organization of the *Ta* gene (Fig. 1*B*).

Interestingly, the original 3′ exon (exon 2) observed in *EDA*—a 218-bp sequence encoding only three amino acids—has not been detected in mouse, but to maintain parallel nomenclature of human and mouse exons, the immediate next coding region of 106 bp, common in Ta A and Ta B, is called exon 3. Two additional exons were identified within an additional 3′ coding region of 674 bp in Ta A and were called exon 4 and exon 5 (Fig. 1 *B* and *C*). To determine the exon 4–exon 5 junction, a PCR product of approximately 2.2 kb was amplified from the corresponding sequence in mouse genomic DNA and sequenced with exon 4 primer. Sequence analysis of genomic and cDNA sequences confirmed the presence of an intron (Fig. 1 *B* and *C*).

A stop codon and a 3′ untranslated region are present in each of the alternatively spliced forms Ta A, B, and C. In Ta C, the region homologous to mouse exon 1 in alternate exon 1a continues with 1,005 bp of partially translated sequence similar to adjacent intron 1 sequence in the human *EDA* gene (11). We isolated corresponding genomic DNA and confirmed that this sequence is identical to the sequence adjacent to exon 1 in mouse genomic DNA (GenBank accession no. AF016632) and contains a splice junction (Fig. 1*B* and data not shown).

In Ta B, at the position where it deviates from exon 3 sequence in Ta A, the alternate exon 3a contains a splice junction and continues with a partially translated longer novel sequence (Fig. 1*B*). Comparisons of Ta B and Ta A cDNAs and respective encoded protein sequences reveal that exon 4 in Ta A begins with a split codon.

Reciprocal comparative analysis of *EDA* and *Ta* sequences has led to the extension of the known span of the human gene. No sequence equivalent to *EDA* exon 2 has been observed in the mouse, but the homologous *Ta* exons 3 and 5 in human have now been isolated in EDA transcripts. Corresponding alternative transcripts in human have also been recovered in cDNAs. For example, for exon 3, nested primers from the 3′ end of EDA exon 1 were used for rapid amplification of 3′ cDNA ends from placental cDNA. A 517-bp product was obtained and sequenced to show that the new human exon was indeed homologous to Ta exon 3. The first 143 bp of the human sequence shows 89% homology to the corresponding mouse sequence. A translation stop codon is found at base 139 of the new exon. The encoded protein species would contain 178 amino acids, with 89% homology to the Ta B protein (Fig. 1*D*). Partial sequencing of another human cDNA clone confirmed the presence of an exon homologous to Ta exon 5. Furthermore, the sequence-tagged sites (STSs) for Ta exons 3 and 5 amplified specific products from human genomic DNA (Fig. 2*A*, lane 9). By STS content in yeast artificial chromosomes, the human sequences map about 80 kb distal to the *EDA* exon 2 in a human yeast artificial chromosome contig (refs. 10 and 11 and data not shown).

Molecular Lesions in Tabby Alleles. Studies of genomic structure confirm that the *Ta* gene is mutated in two independently arising Tabby strains, consistent with its identification as the affected gene. When each exon was assessed by PCR assays in each strain, one *Ta* allele (Fig. 2*A*, lane 2) was specifically deleted for an STS specific for exon 1. Deletion was further verified with additional primers spanning exon 1 and by Southern blot analysis with a Ta A cDNA probe hybridized to mouse genomic DNAs digested with *Taq*I (data not shown). This *Ta* strain would lack the first 132 amino acids encoded by exon 1.

The analysis of exon 1 and exon 3 PCR products by restriction endonuclease fingerprinting (13, 14) detected a mobility shift in exon 1-specific fragments for a second allele, Tabby ⟨*6J*⟩. Sequencing results are shown in Fig. 2*B*. Tabby ⟨*6J*⟩ has a single base deletion at position +550 of the coding region in exon 1. This results in a frame-shift mutation that would truncate Ta protein

prematurely. Corresponding fragments from eight different control mouse DNAs all had the wild-type sequence.

***Ta* Encodes Alternative Forms of the Transmembrane Protein Ectodysplasin-A.** Ta transcripts have different and longer 3′ exons (Fig. 1*B*) and encode correspondingly alternate forms of mouse Ta protein, called ectodysplasin A (Fig. 1 *B* and *C*).

Analogous to human EDA protein, the hydrophobicity profiles of the mouse Ta proteins are consistent with a class II-type transmembrane protein (20). The EDA protein (135 amino acids) originally reported contains a series of four

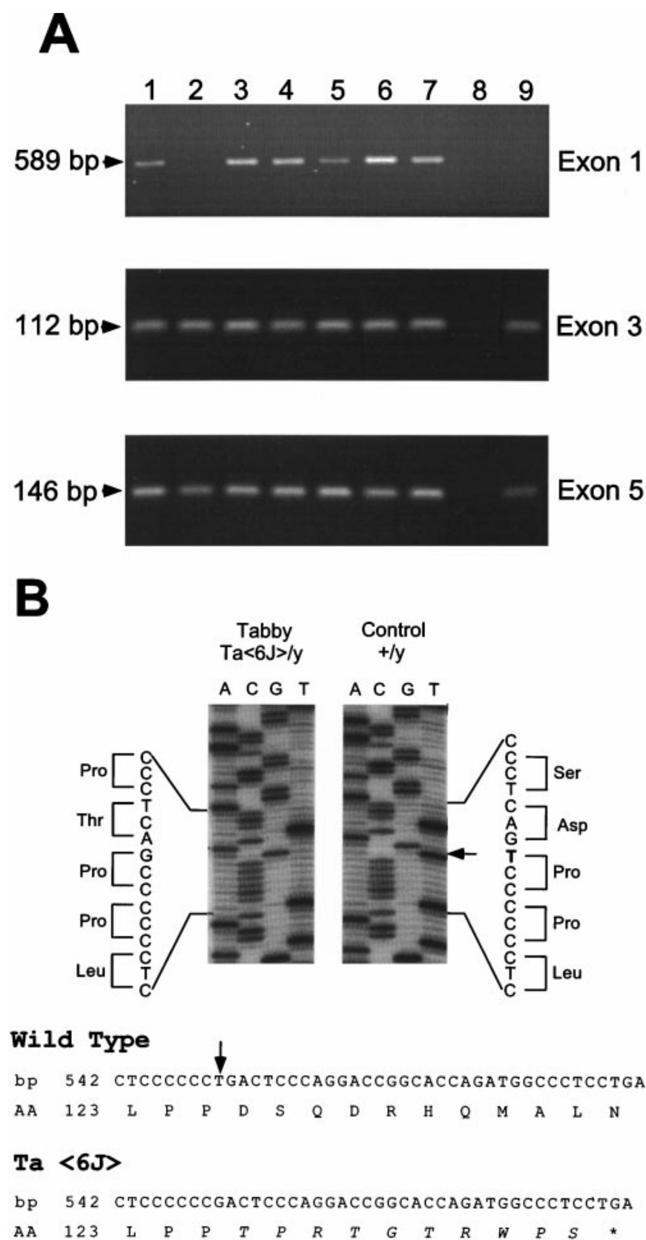


FIG. 2. Analysis of rearrangement in *Ta* gene in two Tabby alleles. (*A*) PCR-based detection of exons deleted in Tabby alleles. Products for exons 1, 3 ,and 5 were amplified using oligonucleotide primer pairs and visualized on EtBr-stained agarose gels. Deletion is evident for exon 1 in one *Ta* allele (lane 2). Lanes: 1 and 3, control mouse DNAs, respectively, for *Ta* (lane 2) and Tabby ⟨6J⟩ (lane 4); 5–7, other mouse DNAs; 8, water; 9, human DNA. (*B*) Exon 1 sequence analysis of mutated Tabby ⟨6J⟩ allele and respective mouse control. Point mutation 550delT is indicated by an arrow. Sequence reads in the sense direction from bottom to top (*Upper*). (*Lower*) Nucleotide sequence and translated amino acids of normal control (wild type) and mutant *Ta* ⟨*6J*⟩ allele. Amino acids shown in italics are due to a frame-shift mutation and results in premature stop codon (*).

collagen-like Gly-Xaa-Yaa repeats in the C-terminal hydrophilic domain of 74 amino acids (11). The Ta protein in its longest predicted form (391 amino acids in Ta A) contains an additional run of 19 Gly-Xaa-Yaa amino acid triplets with only one interruption (—GTTVMGPP—). The putative region of triple helix has a high proline content in the Xaa and Yaa positions. The Ta A protein also contains a cysteine-rich noncollagenous C-terminal region (cysteine residues at +332, +346, and +352). Altogether, five distinct domains can be recognized in the largest deduced Ta protein species: I, a 39-residue N-terminal cytoplasmic domain; II, a 22-residue single transmembrane domain; III, a 118-residue noncollagenous domain; IV, a 59-residue collagenous domain; and V, a 153-residue C-terminal cysteine-rich domain (Fig. 1*C*). Sequence comparison with the nonredundant sequence database suggests that ectodysplasin-A is a novel type of membrane-associated protein with short collagenous motifs.

The predicted molecular weight of the longest Ta form is 41,602 Da that is in good agreement with the results obtained *in vitro* from translation of the corresponding template (band migrating at 42–43 kDa; data not shown).

Expression Pattern and Promoter Region of Ta. As seen for *EDA* (11), Northern blot analysis using adult tissue blots with mouse cDNA as a probe detected mRNA species ranging in size from 6.0 kb to 1.35 kb (Fig. 3 *Right*). This pattern of expression closely resembles that obtained for human mRNA with an EDA cDNA (10, 11).

Northern blot analysis using exon 1 as a probe on RNA preparations from mouse embryos revealed an apparently biphasic expression during fetal development (Fig. 3 *Left*). A 4.5-kb species expressed at low levels predominated in RNAs isolated from 7 days post coitum (d.p.c.) embryos (detected after low-stringency wash), whereas an approximately 5.5- to 6.0-kb mRNA was more apparent in RNAs from 11-, 15-, and 17-day embryos. After high-stringency washes only the larger species of 5.5–6.0 kb was detected. Other minor transcripts of approximately 1.8 kb and 2.0–2.4 kb were also detected at a very low level (Fig. 3 *Left*). An identical expression pattern was seen with a 3′ terminal probe from Ta A (data not shown). The timing of Ta mRNA expression during early fetal development was also monitored by reverse transcription-coupled PCR using RNAs isolated from phased mouse embryos and a pair of primers, one from exon 1 and one from 3 (Ex1–Ex3). Expression was not detected in RNAs from the pre- (0–4 d.p.c.) and periimplantation (5–6 d.p.c.) periods. Expression of the gene started around gastrulation (7–8 d.p.c.) and continued strongly through early (9 and 10–11 d.p.c.) and late (12–13 and 14–15 d.p.c.) organogenesis, and early fetal growth and development (16–17 d.p.c.) period. Expression was sharply reduced and continued low until birth (18 and 19 d.p.c.; data not shown).

The spatial expression pattern of *Ta* was studied in preparations of normal mouse embryonic teeth and skin, two tissues affected in adult Tabby animals (1, 2). The probe, a 398-bp cDNA fragment from exon 1, detected expression at embryonic stage E15 (late cap stage) in the outer enamel epithelium and the dental lamina of the teeth (Fig. 4 *a* and *b*). At embryonic stage E14, expression was detected in the epidermis (Fig. 4 *e* and *f*). In agreement with its identification as the *Ta* gene, no expression was seen in corresponding sections from Tabby teeth (Fig. 4 *c* and *d*) or skin (Fig. 4 *g* and *h*).

The putative promoter region of the gene has properties consistent with wide-spread expression of the gene and with possibly enhanced transcription in ectodermal tissues. A 3.5-kb tract was sequenced from a λ clone recovered by probing with exon 1 DNA. The sequence (GenBank accession no. Y13438) contains an upstream region that includes presumptive promoter elements (nucleotides 1–2,374), followed by exon 1 (571 bp; nucleotides 2,375–2,945), and 618 bp of adjacent intronic sequence (nucleotides 2,946–3,564). The same "intronic" sequence was present in Ta C cDNA (Fig. 1*B*).

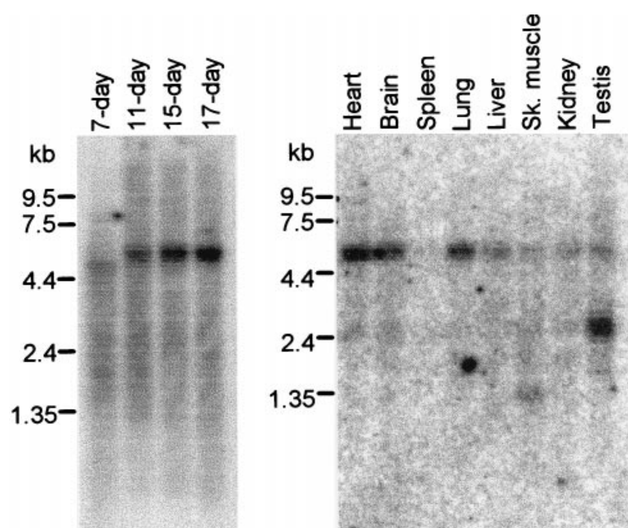


FIG. 3. Northern blot analysis of the expression of *Ta* gene. A 589-bp PCR product specific for exon 1 region was hybridized to RNA samples derived from staged mouse embryos (*Left*) and mouse adult tissues (*Right*).

As in the *EDA* gene, the overall G+C content of the putative promoter region was high, 66% G+C over 1.5 kb spanning exon 1 and a segment 5′ of it. A possible binding site for the transcription factor LEF-1, thought to be involved in ectodermal development (21) and also present in the *EDA* genomic sequence (11), is in the mouse sequence at nucleotides 1,886–1,894. The 2 kb directly preceding exon 1 are 82% identical to the *EDA* promoter region, with higher homology around the LEF-1 binding sequence (96% for nucleotides 1,769–2,018).

DISCUSSION

***Ta* and *EDA* Genes and Their Transcripts.** Many sequence tracts, including segments of the putative promoter and all mouse exons, are conserved between *Ta* and *EDA*. Both are inferred to encode isoforms of a protein that we have named ectodysplasin-A.

Direct analysis of cDNAs indicates that both human and mouse encode protein isoform Ta B, and very likely Ta A and Ta C. The mRNA species originally described for *EDA*, however, included exon 2, resulting in a unique protein species of 135 amino acids. Possibly the gene has evolved an additional exon/isoform or perhaps an equivalent of the exon 2 will yet be found in mouse. Such a possibility is suggested in Fig. 1*B*.

Northern blot analyses detected Ta/EDA transcripts as large as 6 kb–larger than any yet recovered from cDNA libraries. However, all the mRNA and protein species seem to begin with the same exon 1. One speculative possibility is that the common exon 1 is a generic tag for an "EDA group protein," or "ectodysplasin-A," and that the variant 3′ ends would then tailor gene action for distinct functional roles.

The additional telomeric exons now known for the *EDA* gene increase its known extent from 220 kb (11) to more than 300 kb. This rationalizes puzzling earlier results in which DNA from a number of patients showed deletions and mutations in exon 1 or deletion in the earlier described exon 2, but DNA from one male patient DNA contained both exons. Instead, that DNA was deleted for a segment starting 9 kb distal to exon 2 and extending 80–100 kb further (11, 22). The disorder in that patient might involve the deletion of the new *EDA* exon 3.

***Ta* Expression and Regulatory Pathways.** The organs affected in *EDA* and Tabby are epithelial appendages arising as a result of interactions between the mesenchyme and epithelium during fetal development. It seems likely that the gene product is involved early in organogenesis, perhaps functioning in the regulation of tissue interactions. Recombination grafting of epidermal and dermal layers from wild-type and *Ta* mice has suggested a correspondingly complex function for the *Ta* gene in the development of affected tissues (23). In particular, the function of *Ta* in hair follicle initiation in the tail region could not be assigned to either epidermis or dermis, whereas body hair morphology was influenced solely by the type of epidermis (23). Thus, the Tabby gene may have more than one role, perhaps in both epithelial and mesenchymal cell types.

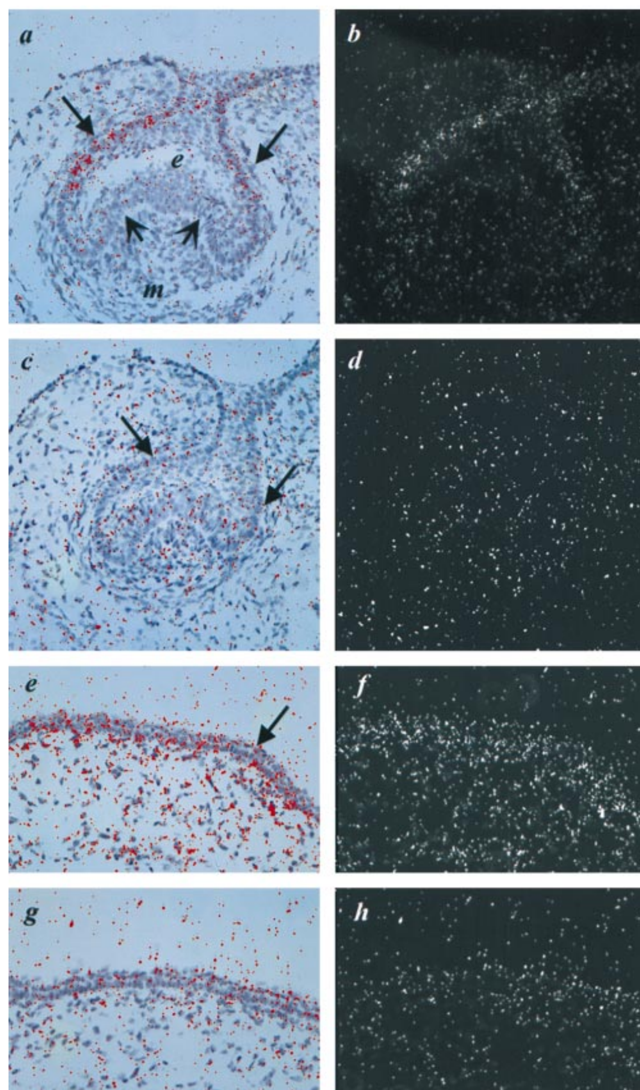


FIG. 4. Expression of *Ta* mRNA in embryonic tooth and skin. *In situ* hybridization was performed with a ^{33}P-labeled antisense RNA probe. (*a* and *b*) A wild-type late cap stage E15 tooth. *Ta* mRNA is expressed in the outer enamel epithelium (arrow) and the dental lamina. No expression in the inner enamel epithelium (arrowhead) and the mesenchyme. (*c* and *d*) In corresponding Tabby teeth there is no expression (arrow indicate outer enamel epithelium). (*e* and *f*) Neck skin from a wild-type E14 embryo. Tabby expression is concentrated in the epidermis (arrow). (*g* and *h*) In Tabby neck skin there is no expression. (*a, c, e,* and *g*) Digital bright-field images overlaid with expression in red. (*b, d, f,* and *h*) Digital dark-field images. e, epithelium; m, mesenchyme.

EDA is expressed in several fetal and adult tissues (11), including hair follicles, neuroectoderm, thymus, bone, and epidermis, and particularly in the body epithelium of 9- to 13-week human embryos (24). In agreement, *Ta* is expressed in embryonic epidermis in skin sections at E14, and in E15 teeth, which are at cap stage, expression is seen in the outer enamel epithelium and the dental lamina.

Further hints about *Ta* function come from the structure of the gene. Of particular interest is the potential LEF-1 binding

sequence in the promoter region of *Ta* and *EDA* genes. LEF-1, a DNA-binding protein that is necessary for tooth, hair, and mammary gland morphogenesis (21, 25), may be involved in regulating *Ta* gene expression. LEF-1 has been shown to be involved in mediating the effects of *wnt* signaling. It is also associated with E-cadherin signaling in which epidermal growth factor (EGF) is also involved, possibly through interaction with the EGF receptor (26–29). Interestingly, EGF can rescue dermal ridge and sweat gland development in newborn *Ta* mice (30), and EGF receptor expression is reduced in *Ta* and *EDA* fibroblasts (31). It is suggestive that Tabby expression overlaps with EGF receptor expression in E15 teeth (32).

Predicted structural features of ectodysplasin-A suggest possible interactions with membrane-associated proteins as well as with extracellular matrix, which have been independently implicated in ectodermal differentiation. The membrane association of the EDA protein has recently been established (12), and both the C-terminal collagen-like repeats and the cysteine-rich noncollagenous region are characteristics of several membrane-associated collagenous proteins (33).

Some evidence consistent with a direct role for ectodysplasin A in the regulation of cell interactions comes from transfection experiments in which the EDA product caused MCF-7 cells to round up and detach (12). The *Ta* gene and protein now provide reagents to extend functional analysis to interacting genes, possibly including mouse genes *dl*, *cr*, and *Slk* (34–36), as well as some of the 150 genetically distinct human ectodermal dysplasias (37).

Note. After submission of this manuscript, recent independent work (38) reported a 1.5-kb cDNA PCR product with the same mutations in the same *Ta* alleles and described a protein similar to the Ta A protein reported herein but lacking 14 amino acids (amino acids 296–390). That protein may represent yet another isoform of Ta.

We thank Ed Michaud, Suvi Taira, and Kari Alitalo for the mouse genomic library; Peter Lalley for somatic cell hybrid DNAs; Outi Montonen for discussions; Angela Brown, Kaija Kettunen, Merja Mäkinen, and Riikka Santalahti for excellent technical assistance; and Lars Paulin and the DNA synthesis and sequencing laboratory for the genomic sequencing. The work has been supported by institutional fund of the J. C. Self Research Institute to A.K.S., the Sigrid Juselius Foundation and the Academy of Finland to J.K., and the Academy of Finland and the Human Frontier Science Program to I.T.

1. Falconer, D. S. (1952) *Nature (London)* **169,** 664–665.
2. Blecher, S. R. (1986) *J. Invest. Dermatol.* **87,** 720–722.
3. Sundberg, J. P. (1994) in *Handbook of Mouse Mutations with Skin and Hair Abnormalities* (CRC, Boca Raton, FL), pp. 455–462.
4. McKusick, V. A. (1994) in *Mendelian Inheritance in Man*, (Johns Hopkins Univ. Press, Baltimore and London), 11th Ed.
5. Cattanach, B. M., Rasberry, C., Evans, E. P., Dandolo, L., Simmler, M. C. & Avner, P. (1991) *Cytogenet. Cell Genet.* **56,** 137–143.
6. Zonana, J., Gault, J., Davies, K. J. P., Jones, M., Browne, D., Litt, M., Brockdorff, N., Rastan, S., Clarke, A. & Thomas, N. S. T. (1993) *Am. J. Hum. Genet.* **52,** 78–84.
7. Brockdorff, N., Kay, G., Cattanach, B. M. & Rastan, S. (1991) *Mammal. Genome* **1,** 152–157.
8. Brockdorff, N., Kay, G., Smith, S., Kerr, J. T., Hamvas, R. M. J., Brown, SDM. & Rastan, S. (1991) *Genomics* **10,** 17–22.
9. Zonana, J., Jones, M., Browne, D., Litt, M., Kramer, P., Becker, H. W., Brockdorff, N., Rastan, S., Daveis, K. P., Clarke, A. & Thomas, N. S. T. (1992) *Am. J. Hum. Genet.* **51,** 1036–1046.
10. Srivastava, A. K., Montonen, O., Saarialho-Kere, U., Chen, E., Baybayan, P., Pispa, J., Limon, J., Schlessinger, D. & Kere, J. (1996) *Am. J. Hum. Genet.* **58,** 126–132.
11. Kere, J., Srivastava, A. K., Montonen, O., Zonana, J., Thomas, N., Ferguson, B., Munoz, F., Morgan, D., Clarke, A., Baybayan, P., Chen, E. Y., Ezer, S., Saarialho-Kere, U., de la Chapelle, A. & Schlessinger, D. (1996) *Nat. Genet.* **13,** 409–416.
12. Ezer, S., Schlessinger, D., Srivastava, A. K. & Kere, J. (1997) *Hum. Mol. Genet.* **6,** 1581–1587.
13. Liu, Q. & Sommer, S. S. (1995) *BioTechniques* **18,** 470–477.
14. Du, Y.-Z., Srivastava, A. K. & Schwartz, C. E. (1997) *Hum. Mutat.*, in press.
15. Takahashi, N. & Ko, M. S. H. (1994) *Genomics* **23,** 202–210.
16. Wilkinson, D. & Green, J. (1990) in *Postimplantation Mammalian Embryos*, eds. Copp, A. J. & Cokroft, D. L., (Oxford Univ. Press, London), pp. 155–171.
17. *Human/Mouse Homology Maps*, The National Center for Biotechnology Information and The Genome Database.
18. Nagaraja, R., MacMillan, S., Kere, J., Jones, C., Cox, S., Schmatz, M., Terrell, J., Shomaker, M., Jermak, C., Hott, C., Masisi, M., Mumm, S., Srivastava, A. K., Pilia, G., Feartherstone, T., Mazzarella, R., Kesterson, S., McCauley, B., Railey, B., Burough, F., Nowotny, V., States, D., Brownstein, B. & Schlessinger, D. (1997) *Genome Res.* **7,** 210–222.
19. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8148.
20. Singer, S. J. (1990) *Annu. Rev. Cell. Biol.* **6,** 247–296.
21. Travis, A., Amsterdam, A., Belanger, C. & Grosschedl, R. (1991) *Genes Dev.* **5,** 880–894.
22. Zonana, J., Jones, M., Clarke, A., Gault, J., Muller, B. & Thomas, N. S. T. (1994) *J. Med. Genet.* **31,** 287–292.
23. Pennycuik, P. R. & Raphael, K. A. (1984) *Genet. Res.* **43,** 51–63.
24. Montonen, O., Ezer, S., Saarialho-Kere, U., Herva, R., Karjalainen-Lindsberg, M.-L., Kaitila, I., Schlessinger, D., Srivastava, A. K., Thesleff, I. & Kere, J. (1997) *J. Histochem. Cytochem.*, in press.
25. van Genderen, C., Okamura, R. M., Farinas, I., Quo, R. G., Parslow, T. G., Bruhn, L. & Grosschedl, R. (1994) *Genes Dev.* **15,** 2691–2703.
26. Hoschuetzky, H., Aberle, H. & Kemler, R. (1994) *J. Cell Biol.* **127,** 1375–1380.
27. Miller, J. R. & Moon, R. T. (1996) *Genes Dev.* **10,** 2527–2539.
28. Behrens, J., von Kries, J. P., Kühl, M., Bruhn, L., Wedlich, D., Grosschedl, R. & Birchmeier, W. (1996) *Nature (London)* **382,** 638–642.
29. Huber, O., Korn, R., McLaughlin, J., Ohsugi, M., Herrmann, B. G. & Kemler, R. (1996) *Mech. Dev.* **59,** 3–10.
30. Blecher, S. R., Kapalanga, J. & Lalonde, D. (1990) *Nature (London)* **345,** 542–544.
31. Vargas, G. A., Fantino, E., George-Nascimento, C., Gargus, J. J. & Haigler, H. T. (1996) *J. Clin. Invest.* **97,** 2426–2432.
32. Partanen, A.-M. & Thesleff, I. (1987) *Dev. Biol.* **120,** 186–197.
33. Pihlajaniemi, T. & Rehn, M. (1995) *Prog. Nucleic Acid Res.* **50,** 225–262.
34. Falconer, D. S., Fraser, A. S. & King, J. W. B. (1951) *J. Genet.* **50,** 324–344.
35. Philips, R. J. S. (1975) *Mouse News Lett.* **23,** 30.
36. Cattanach, B. (1975) *Mouse News Lett.* **53,** 29.
37. Pinheiro, M. & Freire-Maia, N. (1994) *Am. J. Med. Genet.* **53,** 153–162.
38. Ferguson, B. M., Brockdorff, N., Formstone, E., Ngyuen, T., Kronmiller, J. E. & Zonana, J. (1997) *Hum. Mol. Genet.* **6,** 1589–1594.