Genomic Structure and Identification of Novel Mutations in Usherin, the Gene Responsible for Usher Syndrome Type IIa

M. D. Weston,^{1,*} J. D. Eudy,^{2,*} S. Fujita,³ S.-F. Yao,² S. Usami,⁴ C. Cremers,⁵ J. Greenburg,⁶ R. Ramesar,⁶ A. Martini,⁷ C. Moller,⁸ R. J. Smith,⁹ J. Sumegi,² and William J. Kimberling¹

¹Department of Genetics, Boys Town National Research Hospital, and ²Center for Molecular Genetics, Monroe-Meyer Institute for Genetics and Rehabilitation, University of Nebraska Medical Center, Omaha; ³Department of Otorhinolaryngology, Hirosaki University School of Medicine, Hirosaki, and ⁴Department of Otorhinolaryngology, Shinshu University School of Medicine, Matsumoto, Japan; ⁵Katholieke Universiteit Nijmegen, Nijmegen, The Netherlands; ⁶University of Cape Town Medical School, Department of Genetics, Cape Town, South Africa; ⁷ENT Department, Universita di Ferrara, Ferrara, Italy; ⁸Department of Audiology, Sahlgrenska University Hospital, Gotebörg, Sweden; ⁹Molecular Otolaryngology Research Laboratories, Department of Otolaryngology-Head and Neck Surgery, University of Iowa, Iowa City

Usher syndrome type IIa (USHIIa) is an autosomal recessive disorder characterized by moderate to severe sensorineural hearing loss and progressive retinitis pigmentosa. This disorder maps to human chromosome 1q41. Recently, mutations in USHIIa patients were identified in a novel gene isolated from this chromosomal region. The USH2A gene encodes a protein with a predicted molecular weight of 171.5 kD and possesses laminin epidermal growth factor as well as fibronectin type III domains. These domains are observed in other protein components of the basal lamina and extracellular matrixes; they may also be observed in cell-adhesion molecules. The intron/exon organization of the gene whose protein we name "Usherin" was determined by direct sequencing of PCR products and cloned genomic DNA with cDNA-specific primers. The gene is encoded by 21 exons and spans a minimum of 105 kb. A mutation search of 57 independent USHIIa probands was performed with a combination of direct sequencing and heteroduplex analysis of PCR-amplified exons. Fifteen new mutations were found. Of 114 independent USH2A alleles, 58 harbored probable pathologic mutations. Ten cases of USHIIa were true homozygotes and 10 were compound heterozygotes; 18 heterozygotes with only one identifiable mutation were observed. Sixtyfive percent (38/58) of cases had at least one mutation, and 51% (58/114) of the total number of possible mutations were identified. The allele 2299delG (previously reported as 2314delG) was the most frequent mutant allele observed (16%; 31/192). Three new missense mutations (C319Y, N346H, and C419F) were discovered; all were restricted to the previously unreported laminin domain VI region of Usherin. The possible significance of this domain, known to be necessary for laminin network assembly, is discussed in the context of domain VI mutations from other proteins.

Introduction

Usher syndrome (MIM 276900), an autosomal recessive disorder characterized by moderate to profound sensorineural hearing loss and retinitis pigmentosa (RP), is the most frequent cause of combined deafness and blindness in the industrialized world. The incidence of Usher syndrome in the United States is estimated at 4.4/100,000 (Boughman et al. 1983). The syndrome is both clinically and genetically heterogeneous and has been divided into three clinical subtypes, types I, II (MIM 276901), and III (MIM 276902; Smith et al. 1994). The subtypes differ in the degree of hearing loss and in the presence or ab-

* These two authors contributed equally to this work.

sence of vestibular abnormalities. Usher type I patients are profoundly deaf and have absent vestibular responses, whereas Usher type II patients suffer from moderate to severe hearing loss and have normal vestibular responses. Usher type III patients exhibit progressive hearing loss and variable vestibular abnormalities. In all forms of Usher syndrome, the hearing loss is accompanied by progressive RP.

At present, 10 distinct genetic loci have been identified for the three clinical Usher phenotypes: 6 loci correspond to the Usher type I phenotype, 3 for the Usher type II phenotype, and 1 for Usher type III (Kimberling et al. 1990; Kaplan et al. 1992; Kimberling and Smith 1992; Smith et al. 1992; Keats et al. 1994; Sankila et al. 1995; Wayne et al. 1996, 1997; Chaib et al. 1997; Hmani et al. 1999; Pieke-Dahl et al., in press). Although it was originally believed that the Usher syndrome type I phenotype was the most common form of Usher syndrome, our experience is that the Usher type II phenotype accounts for more than half of Usher cases. Most

Received June 18, 1999; accepted January 14, 2000; electronically published March 22, 2000.

Address for correspondence and reprints: Dr. William J. Kimberling, Boys Town National Research Hospital, 555 North 30th Street, Omaha, NE 68131. E-mail: kimberling@boystown.org

 $^{^{\}odot}$ 2000 by The American Society of Human Genetics. All rights reserved. 0002-9297/2000/6604-00052.00

Usher type II families show linkage to the type IIa locus on 1q41, and the gene responsible for Usher syndrome type IIa (USH2A) has recently been identified (Eudy et al. 1998). Although the function of the protein (named "Usherin") is unknown, Usherin has a predicted molecular weight of 171.5 kD and has significant sequence similarity to a number of extracellular matrix proteins and receptors containing laminin-like epidermal growth factor and fibronectin type III repeats.

A comprehensive mutation-detection strategy for differential diagnosis of patients with Usher syndrome requires knowledge of the genomic organization of the USH2a gene, USH2A. In this report we describe the intron/exon genomic organization and sequence of the USH2A gene, from which a set of 21 primer pairs suitable for mutation testing was developed. Heteroduplex and direct sequencing analyses of the DNA from 57 independent USHIIa patients were performed to determine the frequency and distribution of USH2A mutations.

Methods

Genomic Structure

To identify and isolate bacterial artificial chromosomes (BACs) containing the USH2A gene, BAC libraries from Genome Systems and Research Genetics were screened with PCR by use of primer pairs from the critical region of the USH2A gene; these were listed in the database maintained by the Whitehead Institute for Biomedical Research/MIT Center for Genome Research. We obtained additional sequence-tagged sites by sequencing the ends of the BACs obtained from the initial BAC library screening.

PCR Conditions

Oligonucleotide primers were designed from the consensus sequence of the USH2A cDNA sequence and used for PCR and cycle-sequencing reactions. BAC clones were grown in 100 ml of Luria-Bertani broth supplemented with 12.5 mg/ml chloramphenicol overnight and purified with Qiagen tip 100 columns. Human genomic DNA was isolated from whole blood with a Puregene kit (Gentra Systems).

Primer combinations for PCR reactions with *Taq* Gold (Perkin Elmer) or Elongase (Gibco-BRL) were tested with human genomic DNA and BAC clones as templates. PCR reactions were as follows: 10 ng BAC or 200 ng human genomic DNA, 1 × reaction buffer, 0.2 mM each dNTP, 2.0 mM MgCl₂, 0.2 mM each primer; 95°C for 10 min followed by 37 cycles of 95°C for 30 s, 55°C for 30 s, and 72°C for 2 min. Elongase enzyme reactions were as follows: 10 ng BAC or 200 ng human genomic DNA, 1 × reaction buffer, 0.2 mM each dNTP, 1.5 mM or 2.0 mM MgCl₂, 0.2 mM each primer for

genomic templates, and 0.4 mM each primer for BAC templates; 94°C for 30 s followed by 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 10 min.

DNA Sequencing

PCR products were purified and concentrated prior to sequencing by Microcon 100 filters (Ambion). ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction kits (Perkin Elmer) were used to generate sequence from both purified PCR products and BAC DNA templates. PCR products were sequenced with 4 pmol Usherin-specific primers and 4 ml ready-reaction mix in a total volume of 10 ml for 25 cycles of 95°C for 10 s, 50°C for 5 s, and 60°C for 4 min. BAC DNA (500 ng-2 mg) was sequenced in 20-ml reaction volumes by 30 pmol of SP6, T7 pBELOBAC11 vector primers or USH2A primers, 8 ml ready-reaction mix in a total volume of 20 ml for 95°C for 5 min followed by 99 cycles of 95°C for 20 s, 55°C for 30 s, and 60°C for 4 min. Sequencing reactions were precipitated and resuspended in a 5:1 ratio of formamide/dye loading buffer (6 ml for PCR products and 2 ml for BAC DNA). Two microliters were loaded onto 4.8% (36-cM plates) or 5.25% (48cM plates) PAGE-PLUS® (Amresco) sequencing gels and run on the ABI Prism model 377 DNA sequencer according to ABI protocols. Sample files were generated from gel files by ABI Sequencing Analysis 3.0. Fractura 2.0 (ABI) was used to remove vector and ambiguous sequences. Contig assembly was performed by the AU-TOASSEMBLER program (ABI).

Mutation Screening

Table 1 lists the intronic primers used in this study and the amplification sizes of their PCR products surveyed for mutations. A set of 95 anonymous control DNA samples was also screened to assess the frequency of USH2A mutations in a population unaffected by USHIIA.

Heteroduplex Analysis

PCR products amplified from samples taken from USHIIa probands and cloned control DNA were mixed, heated at 95°C for 3 min, and cooled to 25°C over a 45-min period. The reannealed reaction products were then electrophoresed through 35 cm × 43 cm × 1 mm $1 \times \text{MDE}^{\text{(3)}}$ (FMC) at 850 V for 18–24 h, stained with 1 mg/ml ethidium bromide, and visualized under ultraviolet light. Mixing with control PCR DNA prior to heteroduplex formation was done to allow detection of both homozygous and heterozygous mutations.

Direct Link Sequencing

Five-microliter PCR products were treated with 10 U of shrimp alkaline phosphatase (SAP; Amersham) for 30

Table 1

USH2A Primers Used in Mutation Analysis of 57 USHIIa Patients

		Exon	PCR
Primer		Ampli-	Size
Set	Primer Sequence	fied	(bp)
Exon1U	AATGGATTGAGGTGCATGAG	1	525
Exon1L	ACCCAACACCATCTGTCTGT		
Exon2U	GCCTGGGATGAGCTTCAG	2	840
Exon2L	GGTTTGGAATTCAGGCTGA		
Exon3U	CACCACTGTAACTGCACAATACC	3	344
Exon3L	CTGCTGCAGATTTTGTGAGTAGA		
Exon4U	GTCTTCCCAGCTGAACAAAGTA	4	375
Exon4L	GTGGTAATTTGTTCAGTAGCCCTAG		
Exon5U	GTCAGGTATTGCTTGGTAAACAG	5	173
Exon5L	CAGCATTTATCCTTTCGGTTC		
Exon6U	TGACATTCATTTGTAACGACTCC	6	442
Exon6L	AAGTTTGTGGGGCATTTGTTG		
Exon7U	CCATGGTTTGATATATACTGATGG	7	336
Exon7L	CACCAGCCTAGAGAGCTAGC		
Exon8U	CAACATTTTGATTTCTGTTTTGC	8	370
Exon8L	TGCTCTGACATCTTAATGTGCT		
Exon9U	CACACAATGCATATAGTCCTAGG	9	267
Exon9L	TGTTAGGCCAAGATTAAGTTCAT		
Exon10U	TGATATGTGCTTTACTTCTGGTG	10	356
Exon10L	GCATTGTAGATAGAAGCACACAG		
Exon11U	TGGCAGGTAGAGATGAAAGG	11	371
Exon11L	GCAAATGCAGTCTTCAATTCTAC		
Exon12U	CCCTGTCTTGTACCTAATGAGC	12	323
Exon12L	TTCCAGATGGTAATAGAGATGTGA		
Exon13U	GCAGTAGCATTGTTTGTGTCTC	13	816
Exon13L	GTAGAAGCCACAAACCAGAAAC		
Exon14U	GGGAATTAGTGCCTTGGTAGAG	14	379
Exon14L	GAAGTTATTGCTTTGCAACTGC		
Exon15U	AAGCCGTCTTACTCTACAATGCT	15	360
Exon15L	TTCTGATGGGTTCTAAATGGAG		
Exon16U	GCAATCCTGAATCAGAAAGACC	16	354
Exon16L	CCACAACAGCATTTATCCTCAA		
Exon17U	GAGAGGAAAGCAGTTAGCAATG	17	626
Exon17L	GATTCTCATTCATGTCTTGACCA		
Exon18U	AAGTAACCCCTTTGTCTGATGAGT	18	378
Exon18L	GGAAACATTTGCATTCAGAGG		
Exon19U	TCAGAAACTAAATGAATGTGTGA	19	379
Exon19L	TGCCCTGTTTAATCAATATAGAG		
Exon20U	TGGTGGTTGGCAATAATTCC	20	381
Exon20L	GAGTTAGTGAGGGAGGAGAAGACA		
Exon21U	AGCCATACAGATACTTGAAACC	21	501
Exon21L	GCTATCAAAGGGCTGAATTAG		

min at 37°C, and the enzyme was inactivated at 80°C for 15 min. SAP was used to dephosphorylate unincorporated dNTPs prior to cycle sequencing. Sequencing reactions contained 2.8 ml of the BigDye Terminator Cycle Sequencing Ready Reaction mix (Perkin Elmer), 2.5 ml SAP-treated PCR product, and 20 pmol primer. Cycling conditions were 30 cycles of 95°C for 10 s, 55°C for 5 s, and 60°C for 4 min. Sequencing reactions were precipitated in 60% ethyl alcohol (final), dried, and resuspended in 4 ml of a 5:1 ratio of formamide/loading dye. One microliter was loaded onto 4.8% PAGE-PLUS[®] (Amresco) 36-cm sequencing gels and run on the ABI Prism model 377 DNA sequencer according to

ABI protocols. Sample files were generated from gel files by ABI Sequencing Analysis 3.3. The Fractura 2.2 program was used to remove ambiguous sequences and identify heterozygous base positions. Sequence data were aligned and analyzed for mutations by Sequence Navigator 1.0.1.

Comparative Analysis

A total of 32 proteins bearing laminin domain VI motifs were used to evaluate the functional significance of the three missense mutations found within domain VI of Usherin (see fig. 4 below). The first 300 amino acids from each of these protein sequences were compared with amino acids 250–550 of the Usherin protein by a progressive pairwise alignment strategy by the MULTI-ALIGN program (Pôle Bio-Informatique Lyonnais; Corpet 1988).

Results

Genomic Structure

Figure 1 presents an ~325-kb contig assembly across the region that contains the human USH2A gene. Amplification products produced by Elongase by primers specific for both USH2A cDNA and BAC DNA corresponded to seven fragments containing the following exons with their intervening introns: 1-3, 5-8, 10-11, 12-13, 17-18, 18-19, and 19-20. The intron sizes were estimated by agarose gel electrophoresis or were sequenced (IVS1, -6, -7, and -17). Attempts to amplify the following introns were unsuccessful: 3, 4, 8, 9, 11, 13, 14, 15, 16, and 20. Intron/exon boundaries were determined initially by manual inspection of sequence data, then confirmed by direct sequencing of PCR products from BAC DNA by USH2A cDNA primers. The initial determination of intron/exon boundaries was by manual inspection of sequence data. All boundaries were found to be consistent with the invariant intronic AG and GT features that flank the beginning and end of the majority of vertebrate exons, respectively. The contig assembly program AUTOASSEMBLER (ABI) was used on all sequence data generated. Seventeen separate sequence contigs from an input of 174 sequence files were generated and used to confirm the genomic organization. The total length of the 17 contigs was 33.6 kb.

The Usherin gene is encoded by 21 exons (fig. 1). Table 2 lists the sizes and starting positions of each of the 21 exons, the surrounding intronic sequence, intron sizes (where known), and GenBank accession numbers from the National Center for Biotechnology Information. Reliable genomic sequence extends 667 nt 5' from the start of the cDNA sequence. The last genomic contig extends 605 bp 3' from the published end of USH2A.

A significant sequence difference between the published cDNA and genomic DNA concerns the ATG



Figure 1 BAC/PAC contig of the USH2A critical region on chromosome 1q41 developed from the use of sequence tagged sites and expressed sequence tags from the region (see Results section). The location and orientation of the 21 USH2A exons within the contig is shown. Those portions of the gene that are not shaded black are untranslated, as predicted from the cDNA sequence.

translation start site. A stretch of five thymidines was previously observed to start at cDNA position +5 relative to the ATG start site; this was found to be six thymidines on all other genomic-based templates. The additional thymidine removes five amino acids from the beginning of the cDNA's longest open-reading frame (fig. 2A). A reexamination of the USH2A cDNA primary sequence data for this region revealed that the cDNA does contain six thymidines, in complete agreement with the genomic sequence. The poly-A tail of the cDNA sequence corresponds to a 20-nt imperfect poly-A stretch interrupted by 2 Gs in the genomic sequence (fig. 2B). It is likely that the true 3' end occurs downstream of this imperfect poly-A feature. A potential polyadenylation signal (aataaa) was found 340 nucleotides 3' from the end of the poly-A feature. The SignalP program (SignalP V1.1 World Wide Web Prediction Server, Center for Biological Sequence Analysis) predicted that the first 31 aa of Usherin would be a eukaryotic signal peptide (Nielsen et al. 1997). Therefore, the change in primary structure of Usherin from that previously reported (Eudy et al. 1998) does not alter the predicted role of this motif as a signal peptide for potential transport into the extracellular space. The sequence difference reduces the total length of the Usherin protein from 1551 to 1546 amino acids and results in a change of nomenclature for reported mutations; thus, the 2314delG (E772FS) mutation that we previously reported has become 2299delG (E767FS).

Mutations

Table 1 provides the primers, designed from intronic sequence, that were used in the mutation testing we describe here. Exons 1, 2, 5, 13, 17, and 21 were surveyed for mutations by direct sequencing. Exons 3, 14, 16, and 19 were screened by heteroduplex analysis, and heteroduplex-positive samples were sequenced to identify the mutation. To compare the sensitivity of the two methods for the detection of USH2A mutations, exons 4, 6–12, 15, 18, and 20 were screened by both methods.

Seven sequence changes— $373A \rightarrow G$, $802G \rightarrow A$,

Table 2

Intron/Exon Boundaries for the 21 Exons of the Usherin Gene

Exon	Exon Length	cDNA Start	5' Acceptor Slice Site	3' Donor Splice Site	GenBank Accession Number	Intron Length	IVS
1	182	1	Not applicable	aagaaatggtaactcagcttattt	AF091873	673 bp	1
2	688	183	cgtttttgattcccagATACCAGC	GGTGTAATgtaagtagtagtgtag	AF091873	~3 kb	2
3	166	871	ccggttgcttcctaagGTGTGTTA	GTGTGCAGgtgagtaaaataaaaa	AF091874		3
4	133	1037	tggtaatttcttacagGTGCATCA	TTTAAATGgtaagtttctgacttc	AF091875		4
5	64	1170	atgtgtgtatttgcagGTTTAGAG	ACAAACAGgtaagcagatgttgca	AF091876	~2.1 kb	5
6	295	1234	ctcttgactcccacagAGAGATTC	AGTATCAGgtaatgagaaacgata	AF091877	952 bp	6
7	185	1529	aactttccattttcagGTGTTTTA	CTTTCCAAgtatgaatattttaa	AF091877	472 bp	7
8	222	1714	tttctttcattaacagTTTTACTC	AGTGGGAGgtattattatgaattt	AF091877		8
9	94	1936	atgctcatgtttgcagATGTCAGT	GACTTCATgtgagttctattttat	AF091878		9
10	196	2030	atattttttgttctagTGTGATCG	CACTACAGgtaagtagcaaataaa	AF091879	~2.8 kb	10
11	131	2226	cctttccctgatgcagGAAGGAAC	GTGATCAGgtgagtttttctgagt	AF091880		11
12	196	2357	gttgtatcttctctagATTGGAGG	CGTTATTGgtacgtgtaacgcaaa	AF091881	~3 kb	12
13	642	2553	tatattttatctttagGGCTTAGG	TCAACCAGgtaagaaagaaatgta	AF091882		13
14	184	3195	atttctttatttctagGTTTTTAT	ACTGGAAGgtacatctttataagc	AF091883		14
15	164	3379	tttgccttgttttcagATGTCAGC	CAGCAAAAgtaagtgaactctggt	AF091884		15
16	159	3543	cttcattgtaatttagCTCCATTC	CCCATACAgtgagtttaagggttt	AF091885		16
17	495	3702	tttcttttcatttcagGTATTCAA	ACTAAATGgtaagaatccaaggct	AF091886	1,045 bp	17
18	270	4197	tatttaatgtctctagGAATAATT	AGAATCAGgtaaagatcaatgttt	AF091886	~2 kb	18
19	170	4467	ttgttctcattgatagCACCTGTA	TTTCACAGgtattgttattttaa	AF091887	~6 kb	19
20	145	4637	cttttttgccttgaagCTGTTGCA	AGCAGCAGgtaagcaagttttaat	AF091888		20
21	1531 +	4782	gttttctaatatttagCACCAGCA	Not applicable	AF091889		

 $779T \rightarrow G$, $3883C \rightarrow T$, $IVS15+35A \rightarrow G$, 4510-4511 insA, and $4457A \rightarrow G$ —were found by sequence analysis. Twenty-four sequence changes were observed first by heteroduplex analysis. In instances where PCR DNA was subjected to both heteroduplex analysis and sequencing, the detection rate for heteroduplex analysis was 92% of that of direct sequencing. Heteroduplex analysis did not detect any mutations not observed by direct sequencing. Thus, direct sequencing is more sensitive than the heteroduplex method at picking up mutations.

Table 3 lists 31 different nucleotide changes that were observed in this sample and gives their frequencies in the patients with Usher II and unaffected control patients; table 3 also identifies the restriction enzymes that we used to verify the base changes. The frequency of USH2A mutations detected is 58 of 114 USH2A alleles (51%). The number of USHIIa cases homozygous for mutations is 10, the number of compound heterozygotes is 10, and the number of cases with only one identified mutant allele is 18. USH2A mutations were not detected in 19 patients with Usher II. These proportions lead us to estimate that we have detected 69% of the true USH2A mutations and that 4 of the 19 individuals with no observed mutation harbor a change in USH2A that remains to be identified. This further leads us to conclude that ~26% (15/57) of the families with USH2A in this sample are not linked to USH2A. This value corresponds well with an estimate, obtained by heterogeneity analysis of linkage marker data (Pieke-Dahl et al. 1997) 18% of cases are not linked to 1q41.

The locations of the 17 pathologic mutations on a motif-based map of the Usherin protein are summarized in figure 3. The protein has one laminin type VI domain, 10 laminin-like epidermal growth factor domains, and four fibronectin type III domains. It is of interest that the three probable pathologic missense mutations are limited to the domain VI region. All the other pathologic mutations we observed produced nonsense codons, frame shifts of mRNA translation, or splice mutations.

Nonsense Mutations

Seven stop mutations distributed throughout the Usherin gene were observed. None were observed in the series of 95 controls. All individuals had a typical type II phenotype.

The 779T \rightarrow G (L260X) mutation was observed in a Swedish type II family as a compound heterozygote with the 2299delG mutation discussed below. A G \rightarrow A transition was seen at bp 1227 in exon 7 (W409X) in two apparently unrelated Dutch families, one homozygous and the other a compound heterozygote with 1256G \rightarrow T (C419F). In exon 10, a 1696C \rightarrow T (Q566X) mutation was observed in one Swedish family as a heterozygote. A 1876C \rightarrow T (R626X) mutation in exon 11 was observed in an American family of mixed northern European ancestry that also carried a 2299del mutation. The 2541C \rightarrow A (C847X) mutation in exon 13 was observed with the 2299delG allele as a compound heterozygote in an American family with African heritage. There were two stop mutations in exon 18; one is an unusual double

```
A.
   CDNA
         (AF055580)
                  5'~TTTATCAGGAGGAGAATGCTTTTT~GTAAACATGAATTGCCCAGTTCTTTCA~3'
                              M L F
                                    VNMNCPVLS~
   Genomic (AF091873)
                 5'~TTTATCAGGAGGAGAATGCTTTTTTGTAAACATGAATTGCCCAGTTCTTTCA~3
                                          М
                                           N
                                              CPV
                                                    L
                                                      S~
B.
   CDNA
         Genomic (AF091889) 5'~CTTTATATTCTTAAAATGGCAAAAAAGAAAAGAAAAAAGGTAAG~236bp~TAATAAA3'
```

Figure 2 Significant sequence differences between the previously reported cDNA (GenBank accession number AF055580) and genomic DNA (GenBank accession numbers AF091873 and AF091889) at the 5' and 3' ends of the Usherin gene. *A*, The genomic structure predicts a start methionine five codons downstream from the ATG predicted by the published cDNA sequence because of an insertion of a T residue. A reexamination of the primary cDNA sequence data agrees with the genomic DNA data. This predicts a 1546-aa protein. *B*, At the 3' end of the gene, a potential polyadenylation signal (aataaa) was found 340 nt 3' from the end of a 20-nt imperfect poly-A region found in the genomic sequence that otherwise lines up with the end of the USH2A cDNA.

transversion 3840-41GA \rightarrow CT (MR1280IX), and the other is 3883C \rightarrow T (R1295X). The double transversion occurred in an American family with European ancestry. R1295X was seen in the heterozygous state in one sibship of a large Swedish family with three involved sibships; for both families, the companion mutation was not detected.

Frameshift Mutations

Five mutations cause frameshifts with premature termination. The first, a four-base insertion in exon 6, 921-922insCAGC (H308FS), was observed in two compound heterozygotes. The families were reportedly unrelated, but both were of English-Scottish ancestry. One family segregated the 2299delG mutation on the other allele; the other inherited a novel $2052A \rightarrow G$ (Q684Q) maternal allele that could create a possible cryptic splice acceptor site 81 nt 3' of the normal splice acceptor site inside exon 12. One additional $2052A \rightarrow G$ allele from an unrelated USHIIa patient was observed; this mutation was not observed in any of the 95 control samples by heteroduplex analysis. A 1679delC (P560FS) mutation occurred in combination with 2299delG in an individual who has mixed northern European ancestry. One consanguineous individual from Venezuela was homozygous for 2898delG (T967FS), causing a frameshift starting at codon 967. A deletion of 2 bp, 4338-9delCT (Q1447FS), was observed in a French Acadian family with a well-documented Usher II phenotype. Although the parents of the affected children are believed to be distantly related, the children are predicted to be compound heterozygote, with the companion mutation still undetected.

The most common Usherin mutation was 2299delG (E767FS). We have observed this mutation in 31 of 192 alleles. The increase in the total number of patients tested for this mutation as well as portions of exons 12, 13, 14, and 20 is due to our previous work (Eudy et al. 1998). This mutation occurred among USHIIa patients

with a variety of ancestries, but all were European. The single exception was the 2299delG/C840X compound heterozygote with African heritage, discussed above.

Splicing Mutations

To evaluate the 2052A \rightarrow G (Q684Q) mutation, we used the FITCONSENSUS program from the GCG package to calculate quality scores for the authentic exon 12 acceptor site and the potential cryptic site created by the mutation with the eukaryotic 3' splice acceptor consensus table (Mount 1982). The normal splice site quality score was 42.89; the Q684Q mutation score was 42.28. This mutation was absent from the 95 control samples; however, its potential to produce aberrant splice products has not been tested by reverse transcriptase-PCR methods. The other nonpathologic alterations from table 3, namely 2109T \rightarrow C (D703D) and 4371G \rightarrow A (S1457S), were also checked for potential splicing effects by FITCONSENSUS. No significant changes in the quality scores for the natural splice sites in these exons was observed. The 4371G→A does creates a new AG site in exon 20, but the quality score calculated from FIT-CONSENSUS at this potential cryptic acceptor site was 37.67, well below that of the natural site (44.28). In addition, restriction digests of PCR products from the 95 normal controls indicated that $2109T \rightarrow C$ and 4371G \rightarrow A are both present (table 3).

Missense Mutations

Seven different amino acid changes were observed in the sample groups. Three of these, $373A\rightarrow G$ (T125A), $1931A\rightarrow T$ (D644V), and $4457A\rightarrow G$ (K1486R), are clearly polymorphic, as indicated on the basis of their frequency in the USHIIa population or the normal controls (table 3). In addition, both alleles of the $373A\rightarrow G$ polymorphism were observed from BAC clone sequences of exon 2. The $802G\rightarrow A$ (G268R) change was observed in a Swedish family in which inheritance of the gene was maternal. We were not able to test for the presence of

Table 3

DNA Sequence Changes in USH2A

Exon	Base Change	Codon Change	Frequency in USH2A Patients	Family Origin	Frequency in Controls ^a	Predicted Pathology	Restriction Enzyme ^b
DNA sequence changes in Usherin believed	1						
not to be pathologic:							
2	c.373A→G	T125A	27/94		^c	None	
4	IVS3-80T→C		9/94		14/190	?	MseI = N
5	c.802G→A	G268R	1/94			?	
7	IVS7+36-39del gatt		11/94 ^d		14/190	?	BsrI = Y
8	c.1419C→T	T473T	22/94			None	
9	IVS9+34C→A		34/94			None	
11	c.1931A→T	D644V	11/94 ^d		18/190	None	MaeIII = Y
12	c.2109T→C	D703D	1/94		3/188	?	FokI = N
14	c.2880T→C	N960N	1/192 ^e		0/190	?	
15	IVS15+35A→G		38/94			None	
18	IVS17-8T→G		25/94		^c	None	
18	c.3945C→T	N1315N	5/94		2/190	?	NsiI = N
20	c.4371G→A	\$1457S	4/94		3/188	?	Eco57I = Y
21	c.4457A→G	K1486R	35/94		76/188	None	A f l I I = N
Probable pathologic mutations							,
in Usherin:							
4	c.779T→G	L260X	1/94	Swedish		Nonsense	
6	c.921-922insCAGC	H308FS	2/94	American		Frameshift	
6	c.956G→A	C319Y	2/94	American	0/190	Missense	BsmI = N
6	c.1042A→C	N346H	2/94	Swedish, American	0/190	Missense	MlnI = Y
7	c.1227G→A	W409X	3/94	Dutch		Nonsense	
7	c.1256G→T	C419F	5/94 ^d	Dutch	0/190	Missense	ApoI = Y
10	c.1679delC	P560FS	1/94	American		Frameshift	
10	c.1696C→T	Q566X	1/94	Swedish		Nonsense	MseI = Y
11	c.1876C→T	R626X	1/94	American		Nonsense	
12	c.2052A→G	Q684Q	2/192°	American	0/190	Cryptic splicing	
13	c.2299delG	E767FS	31/192°	Many	0/190	Frameshift	
13	c.2541C→A	C847X	1/94	American		Nonsense	
14	c.2898delG	T967FS	2/192°	Spanish	0/190	Frameshift	
18	c.3840-41GA→CT	MR1280IX	1/94	American		Nonsense	
18	c.3883C→T	R1295X	1/94	Swedish		Nonsense	
20	c.4338-9delCT	C1447FS	1/192°	American	0/190	Frameshift	
21	c.4510-4511insA	R1504FS	1/94	American		Frameshift	

NOTE.—An ellipsis (...) = not applicable; a question mark (?) = unknown.

^a A plate of 95 normal controls was screened for the presence of the mutation.

^b If the restriction enzyme equals Y, the enzyme will cleave the rarer variant. If N, the enzyme will cleave the common allele.

^c These polymorphisms were identified from the sequences of various BAC clones.

^d These changes were observed in *cis* (disequilibrium) within the USHIIa group.

* Portions of exons 12, 13, 14, and 20 were originally screened by heteroduplex analysis in 96 USHIIa and 95 control samples (Eudy et al. 1998).



Figure 3 Linear structure of Usherin showing the significant domains. The protein has one laminin type VI domain, 10 laminin-like epidermal growth factor domains, and four fibronectin type III domains. Numbers below the figure are amino acid positions for each motif. The relative locations of 17 pathologic mutations from an analysis of 57 individuals with USHIIa are shown above the figure. Three missense mutations, all within domain VI, were discovered. Q684Q creates a punitive cryptic 5' splice acceptor site. The other mutations produce frame shifts of mRNA translation or encode nonsense codons.

G268R in the control samples, because this change did not produce a heteroduplex pattern and no restriction sites were altered by it. However, exon 5 DNA sequence from amplified gorilla DNA was homozygous for the rare human variant (unpublished observations), and so this base change is not considered likely to be pathologic.

The three missense mutations in USH2A, $956G \rightarrow A$ (C319Y), $1042A \rightarrow C$ (N346H), and $1256G \rightarrow T$ (C419F) (fig. 3), all are located within domain VI and are absent from 95 control samples (fig. 3). The first, C319Y, was observed to be homozygous in the proband of a Hispanic American family with no family history of consanguinity. N346H was observed in the same large Swedish family described above that segregated 1295X. The N346H mutation was also observed in an American family as a compound heterozygote with 2299delG. The C419F mutation was observed exclusively in five Dutch families. Three of five Dutch families were simple heterozygotes, two were compound heterozygotes, and one carried 2299delG and the other W409X.

To evaluate the potential impact of these substitutions on the Usherin protein, a multiple sequence alignment was generated by 32 protein sequences with type VI domains, to identify conserved residues. A portion of the aligned set of laminin type VI globular domains from representative proteins and the locations of the three Usherin missense mutations is shown in figure 4. The C319Y and C419F mutations involve cysteine residues absolutely conserved in the type VI domains of these proteins. The N346H mutation occurs at a nonconserved residue, but the alignment identity of the asparagine residue with human laminin γ 1, *Caenorhabditis elegans* laminin-like protein-1, human netrin-1, and *C. elegans* UNC6 suggests that this residue is also functionally conserved.

Disequilibrium

A 1931A→T polymorphism causes a nonconservative amino acid change, D644V, in the third laminin-like epidermal growth factor (L-EGF) domain of Usherin. The polymorphism was observed in the USH2a population as two D644V homozygotes and seven D644V heterozygotes, for a gene frequency of .096. In the control population, one homozygote and 16 heterozygotes were found, resulting in a gene frequency of .094. The C419F mutation appears to be in disequilibrium with D644V, because all five Dutch C419F alleles carried in cis the rarer D644V allele. None of the 95 control samples carried the C419F mutation, as determined by digestion of exon 7 PCR products with ApoI. Another sequence variant, IVS7+36-39delGATT, was observed to be in complete disequilibrium with all the USH2a patients who have the D644V allele. Unfortunately, we were unable to find USH2A mutations segregating with the 6 IVS7+36-39delGATT/D644V alleles in USH2a patients who did not have the C419F mutation. Concerning the IVS7+36-39delGATT allele among the control group with the D644V allele, BsrI digests of exon 7 PCR products showed that 4 of 16 D644V heterozygotes do not carry an IVS7+36-39delGATT allele, and the one normal control D644V homozygote was heterozygous for IVS7+36-39delGATT.

Discussion

Knowledge of the gene structure of USH2A was a prerequisite before a thorough screening of USHIIa patients for mutations in that gene could be undertaken. A full description of the genomic organization of the Usherin gene will facilitate the screening for additional patho36.

METS MOUSE

NE TA_DROME NE TA_DROME NE TA_DROME UNICG_CAREL NETI_FROG NETI_HUKAN LNGI_DROME LNGI_HUKAN LNB2_HUKAN LNB2_HUKAN LNB2_HUKAN CELUOSF8 LNA3B_HUMAN	30: 24: 22: 12: 34: QRPFINSAGE 1: MGOPOPAELY 34: Q	FSQQA SAAGI AYFSQFSHRA HELRGTTFMP PSVDPNHFGA	PPEDPCYNKA SSNDPCYFEG PDHDPCHDHT ALE-CYDPYG DGNPCYDRAT AMDECTDEGG QQA QAPAPDVPG QEPEFSY-G TVGIP	LEPRS CIEDF K-PKK CIEDF GRPVROVPEF PRRCIEDF PRRCIEDF R-POKCLPEF R-POKCLPEF CSRGSCYPAT CAEGSCYPAT CAEGSCYPAT GPRAQLHPTY	VIAAYDAPVVASSICG VIAAYDAPVVASSICG VIAAYCHPVIASDICG- VISAFCZVIASDICG VISAFCZVIASDICG VISAFCZVIASSICG VISAFNLCUVTINCG COLVCATRILASSICG COLVCATRILASSICG COLVCATRILASSICG CHLICACKLSVISICG CHLICACKLSVISICG CHLICACKLSVISICG		RCEY-ODH RCEY-ODH RCELLRDGN RCVVTEKGE RCVVTEKGE RCVVSERGE RCVVSERGE RCVVOSGHTG RCVVOSGHTG RCVVOSGHTG RCVOSGHTG RCVVSHLQD PCIVSHLQD RCIVSHLQE RCIVSHLQE RCIVSHLEE RCUVSHL	ERS GIIREQ GR ERL-RS HKN QRSV VTKS QMK EKK QTK QTK		9 AAD 2680: 2 024567 7 024568 4 934710 9 AAB 87983 5 U75586 1 915215 1 018823 4 911047 1 013751 8 955268 1 907942 8 AF039041 1 AAC 51867
LMAS_MOUSE LMA_CAEEL LMA2_MOUSE LMA1_HUMAN Usherin_HUMAN Consensus Mutations	43: 23: 18: 286:	E Q R	SR RR0SQA	ERGLFUNI HQQRGLFDAV QRGLFDAV LEVFSGDL	FULATINSL-ITATION FULATINSL-ITATION UMASNAL-ITTNAICG- LAWASNAHISTNAIGA- LAWASNAHISTNAIGA- LAWASNAH-STRAIG STAR CONSTRUCTION	aa deletion	LYCKLVGGPV EYCKLVEHVL MYCKLVEHVL MYCKLVEHV- RYCIPNDAGD AC 091E,LMA21 C319Y (U	AGGD PN-QTI LRKTTNT PGQPV PGRPV Kouse, dy ²³): sherin)	QGQYQDICHA: 3 QGQYQDICHA: 3 QSPQDDICHA: 100 RNPQDRINQ: 8 RNPQRINQ: 7 TAON: 330 TAON: 330	0 061001 0 061001 0 AF074902 4 060675 1 P25391 0 AAC23748
NET3 HOUSE NETA_DROME NETB_DROME UNC6_CAEL NETI_FROG NETI_FROG NETI_FROG LIG_DROME LIG_DROME LIG_HUMAN LIM3_HUMAN LIM3_HUMAN LIM3_HUMAN LIM3_HUMAN LIM3_BONNE LIMA_SCAEL LIMA_ROMSE LIMA_ROMSE LIMA_HUMAN USherin_HUMAN Usherin_HUMAN	70: SDPQRAHSAD 93: TDPLRSFPAR 87: QRYGPA 95: RNHFOSHPAS 61: SDAKRAHP PS 112: NDHP3 82: RHEGFSHPAK 95: GOPHLOHGAA 72: ROPHNY 94: RRPFSARDUHGAA 72: ROPHNY 94: RRPFSARDUHGAA 72: ROP	LLTSAPGTAS SLTD LNNSNN ALTD LNNSNN LLTD LNNSNN FLTD LNNFNN FLTD LNNFHN FLTD LNNFHN FLTD LNNFHN FLTD YNGON YSHRUENY -SHRUENY V-SHRUENY V-SHRUENY V-TNADGS P-ENALDGT I-TNADGS PLSFVNDNDY H34 GH(IC6_CAEEL, eX		QAP EPVTC3G YPHDPDSAP TP5L3P 	FNOTLINY BER-ABELVING DISTLIN'S GREATELINY DISTLIN'S GREATELINY DISTLIN'S GREATELINY DISTLIN'S GREATELINY DISTLIN'S GREATELINY ON THIS GREATELINY DISTLIN'S GREATELINY DISTLIN'S GREATELINY DISTLIN'S GREATELINY DISTLIN'S GREATELINY DISTLIN'S GREATELINY DISTLIN'S GREATELINY DISTLINY D	LRUCS APP LSDCPRSPER LSDCPRSPER LSDCS-PRP LDTCS-PRP LDTCS-PRP LDTCS-PRP LKPHT-SSP RF RFDG-FNP HTPTS-PRP IKSAN-SPRP IKSAN-SPRP IKSAN-SPRP IKAN-SPRP IKAN-SPRP IKAN-SPRP IKAN-SPRP IKAN-SPRP IKAN-SPRP IKAN-SPRP IKAN-SPRP	TSVALLESOD DSTAVIYSTD DSTAVIYSTD DSTALYSAD ESHAUFXSAD ESHAUFXSAD ESTIYSTS STIYSTS STIYST AGHLIPSSD AAHLIPSSD AAHLIPSSD AAHLIPSSD ELVVL285D GLVL285D	HORSIVALGE HELSIOFFOR FORMUTERS FORM	Solution Solution Solution Solution<	 AAD28603 Q24557 Q24567 Q24568 P34710 AAB87983 U75586 P15215 Q18623 P1047 Q13751 P5268 P07942 AAC51867 Q00174 Q61001 AF074902 Q60575 P25391 AAC23748 sherin)
NET3 NOUSE NETA DROME NETA DROME UNCG CAEEL NETI FROG NETI HUMAN LINGI DROME LINGI DROME LINGI HUMAN LINSI HUMAN LINSI HUMAN CELWO3F6 LINASH HUMAN LINSI HUMAN LINASHOUSE LINASH UNSE LINASH UNSE LINASH UNSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE LINASHOUSE	157:D WGR LPAPAD 162:L GR PARAGT 160:T MGR PDRAKI 181:I GGR PDRAKI 181:I GGR PDVSI 184:H MWR PHRAPI 198:T GSL PDSRAI 172:T GSL SDRAPI 155:T GSK PMR GGR 155:T GSK PMR GGR 155:T GST PMR GGR 173:S PGI PEG PP 161:E GGFR AM 171:Y GGFR TG 171:Y GGFR TG 199:O GGI PATTGV 175:L WHI YFR GGP 162:R WHITPRR GD 421:X GGKRNNG L	GPSGFGFEAL GRIN-EHEAR SKYN-EJEAR TKSN-EJEAR TKSN-EJEAR TKON-EJEAT RKGEGEAHAL JFON-EATAO RKGEGEAHAL JFON-EATAO RKGEGEAHAL JFON-EATAO SWODII KKHTDII KKHTDII KKHTDII KKHTDII KKHTDII KKHT-DII FRYGKDEGH PSYAKDDEJI PSYAKDDEJI PSYAKDEJI FYPADDEJI	PPA-POAUS CINSCHOTGC CINSCHOTGC CINSCHOTGC CINSCHOTGC CINSCHOTCS CINSCHOT CINSC	D-GGELLARS RIA DS ANGR TAIN GGNRTAIN 	V 00	HODWYTNDE HODWYTNDE HODWYTNDE HODWYTNDE HODWYTNDE HODWYTNDE HODWYTNDE HODWYTNDE HORWYTNSAL HOEWYTNDE HORWYTNSAL HOEWYTNDE HOEWTRUDE HOEW	TV LTRP TV FARLEP TV FARLEP	POLLKVKNAN		AAD28603 024567 024568 P34710 AAB87983 U75586 P15215 018823 P11047 013751 P55268 P07942 AF039041 AAC51867 000174 061071 AF074902 060675 P25391 AAC23748

Brider unt Briene

TTT Dec 25

Figure 4 A portion of an aligned set of laminin domain VI regions from representative proteins for the evaluation of three Usherin missense mutations. The MULTIALIGN program generated a multiple alignment by use of 32 different domain VI-containing sequences with a progressive pairwise alignment strategy (Corpet 1988). Twenty sequences, representing the three classes of domain VI-bearing laminin subunits (α , β , and γ) and netrins were chosen for display in this figure. Protein name and amino acid numbering are to the left of the lineup, and GenBank or SwissProt accession numbers are to the right. Included at the bottom of the protein lineup are the locations of relevant mutations (see Discussion section) within domain VI of UNC-6 (ev437, rh46, and ev436hs), LAMB3 (E210K), LAMA1 (dy²ⁱ), and Usherin (C319Y, N346H, and C419F, in bold). The locations of netrin-specific cysteines are shown by arrows (Wadsworth et al. 1996). Dark-shaded residues are >90% conserved in all 32 aligned sequences. Light-shaded residues are conserved 50%–90% in all 32 aligned sequences. Both the C319Y and C419F mutations involve cysteine residues absolutely conserved in the type VI domains of these proteins. The N346H mutation occurs at a nonconserved residue. Note, however, the alignment identity of the asparagine residue with LMG1_HUMAN, LML1_CAEEL, and all examples of netrin proteins with the exception of NET3_MOUSE.

logic mutations in patients and families that have USHIIa. The distribution of mutations across the gene might suggest certain alternatives for more-expedient clinical testing. To date, 17 mutations have been identified (fig. 3). Combined 2299delG homozygotes and heterozygotes account for $\sim 20\%$ of all USHIIa patients and may be the most frequent mutation causing RP in the human population. Thus, a reasonable first pass for USHIIa molecular diagnosis can be made by testing for the

presence of the 2299delG mutation. The occurrence of apparent disequilibrium with rare polymorphisms (D644V, IVS7+36-39delgatt) suggests that other as-yet-undetected mutations may also be relatively common.

Usher syndrome is highly heterogeneous, with three different clinical types and at least 10 different genes. Usher type II has three subtypes: USHIIa on chromosome 1q is due to mutations in the Usherin gene, USHIIb on chromosome 3p (Hmani et al. 1999), USHIIc on chromosome 5q (Pieke-Dahl et al., in press), and a fourth group of Usher II families, affected members of which have a gene that remains to be localized (data not shown). The non-USHIIa families are in a minority, probably accounting for <30% of all cases of Usher II. This presents a problem for the positional cloning of the USH2B and USH2C genes. However, screening of Usher II patients for mutations in Usherin can help in differentiating those attributable to USH2A from those attributable to USH2B or USH2C, thus reducing the impact of heterogeneity on the search for those two genes. However, it is important to know how thoroughly USH2A mutations are detected. Our results indicate that ~69% of all the true USH2A mutations have been detected by the methods used in this study.

The interfamilial variation of both the severity of hearing loss and the onset of RP in individuals with documented USHIIa may turn out to be due to particular mutations. Although no clear phenotypic variation can be assigned to the particular mutations presented here, a recent study of the prevalence of the 2299delG mutation included 10 atypical Usher cases and revealed a surprisingly high 40% carrier frequency, which indicates that environmental or genetic factors influence the effect of Usherin mutations (Liu et al. 1999).

The delineation of mutations involved in diseases provide an invaluable resource for identification of regions of the protein critical to its proper function. This strategy is especially useful with a novel protein such as Usherin, the function of which can only be speculated on at this stage of our knowledge. The discovery of three putative missense mutations (C319Y, N346H, and C419F) in the laminin type VI domain of this protein mark this region for a potentially significant functional role in the cochlea and retina. Of particular interest are the other classes of proteins containing domain VI motifs: laminin subunits $\alpha 1$, $\alpha 2$, $\alpha 5$, $\beta 1-\beta 3$, $\gamma 1$, $\gamma 3$, and netrins. Laminins (reviewed in Timpl 1996) are the major noncollagenous components of basement membranes that mediate cell adhesion, growth, migration, and differentiation. They function as heterotrimers composed of α , β , and γ subunits that combine to form 12 known heterotrimeric assembly forms, laminin-1 through -12 (Miner et al. 1997; Koch et al. 1999). Domain VI of the laminin subunits is crucial to a threearm interaction model in which individual laminin heterotrimers polymerize via calcium ion-dependent domain VI interactions into quasihexagonal networks (Yurchenco and Cheng 1993). Netrins are small diffusable proteins that control guidance of central nervous system commissural axons at the midline and peripheral motor axons (reviewed in Tessier-Lavigne and Goodman 1996). A homolog of netrin in *C. elegans*, UNC-6, is one of the cues in the extracellular matrix that guides dorsoventral migrations of pioneer axons and migrating cells along the body wall on the epidermis (Ishii et al. 1992). It has been proposed that netrins may copolymerize within the laminin network in the basement membrane via its domain VI (Wadsworth et al. 1996).

Several mutations in the domain VI region of the genes encoding laminin subunits β 3 (E210K), α 2 (dy^{2J}), and the C. elegans netrin UNC-6 (ev437 and ev436hs) have been shown (fig. 4). These mutations appear to permit at least some protein translation, with the encoded proteins producing less-severe phenotypes than null mutations. The mutation E210K, in the domain VI of laminin β 3, is predictive of a less severe, nonlethal form of epidermolysis bullosa (MIM 226650; Mellerio et al. 1998), a genetically heterogeneous group of diseases characterized by easy blistering and fragility of the skin, where mutations in the genes encoding various components of the hemidesmosomal-anchoring filament complex and epidermal basement membrane have been identified (Pulkkinen and Uitto 1999). Similarly, mutations in laminin $\alpha 2$ and components of the dystrophin-associated glycoprotein complex cause various forms of muscular dystrophy (Culligan et al. 1998). The muscular dystrophy mouse dy^{2J} produces a less severe phenotype because of a 57-aa deletion within the domain VI of laminin $\alpha 2$ (Sunada et al. 1995). The altered laminin-2 heterotrimer localizes to the basement membrane but is defective in laminin polymerization (Colognato et al. 1999). The UNC-6 mutations ev436hs and ev437 cause the selective loss of some, but not all, cell and axonal migrations directed by UNC-6 cues (Wadsworth et al. 1996).

The less severe phenotypes associated with these domain VI mutants appear to be in contrast to those of Usherin, in particular the C319Y homozygote. A review of clinical records indicates that this individual is phenotypically indistinguishable from several 2299delG homozygotes. In this regard, the domain VI UNC-6 mutation rh46 (A136P; fig. 4) produces a defective cell and axonal migration phenotype indistinguishable from the UNC-6 ev400 (Q78X) stop-codon mutant (Wadsworth et al. 1996).

Although netrin and laminin subunits share domain VI in common with Usherin, one difference is the motif's location in the protein. The domain VI in laminin subunits and netrin molecules is located at the NH₂ terminal Weston et al.: Mutations in the Usher IIa Gene

end of these proteins. In the case of Usherin, its domain VI is preceded by ~ 300 aa of novel sequence (fig. 3). The potential functional role of Usherin in the development or maintenance of the sensory systems of the ear and eye based solely on domain structure similarities with laminins and netrins would be speculative. Recently, however, mutations in the human homolog of the Drosophila melanogaster protein crumbs (CRB1) have been identified in patients with a severe and specific form of autosomal recessive RP (RP12 [MIM 600105]) on chromosome 1q31-32.1. The CRB1 protein, which contains 19 EGF-like domains, three laminin A G-like domains, and a C-type lectin domain, is implicated in neuronal development of the retina; this fact is based on domain similarities with other EGF-like domain-containing proteins that are essential for neuronal development (den Hollander et al. 1999). It is quite likely that the mutated cysteine residues we have identified in patients with USHIIa disrupt intra- or intermolecular disulfide bonds. Whether these changes lead to an unstable protein more sensitive to proteolytic degradation or to an alteration in the structure and thus function of the protein is unknown. The identification of Usherin protein-binding partners, its subcellular localization, and its spatial and temporal pattern of expression will provide the evidence for a model of the function of this molecule in the cochlea and retina.

Acknowledgments

This work was supported by NIH-NIDCD grant 5P01DC01813 (to W.J.K.), NIDCD grant R01DC02842 (to R.J.H.S.), and by The Foundation Fighting Blindness, Hunt Valley, MD (to W.J.K. and J.G.). In Holland, the study was supported by grants from the Dutch Foundation for the Hand-icapped Child, the Dutch Foundation "De Drie Lichten," The Heinsius Houbolt Foundation, and De Landelijke Stichting voor Blinden en Slechtzienden. In South Africa, the study was supported in part by the Retinal Preservation Foundation of Southern Africa (to J.G.). We especially want to thank all the patients and their families for their participation in this study.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

- National Center for Biotechnology Information (NCBI), http: //www.ncbi.nlm.nih.gov/ (for GenBank accession nos.)
- Online Mendelian Inheritance in Man (OMIM), http://www .ncbi.nim.nih.gov/omim (for Usher syndrome [MIM 276900], subtypes I, II [MIM 276901], and III [MIM 276902], epidermolysis bullosa [MIM 226650], and RP12 [MIM 600105]
- Pôle Bio-Informatique Lyonnais, http://ferrari.ibcp.fr/NPSA/ npsa_multalin.html (for MULTIALIGN program)
- SignalP V1.1 World Wide Web Prediction Server, Center for

Biological Sequence analysis (CBS), http://genome.cbs.dtu .dk/services/SignalP/

Whitehead Institute for Biomedical Research/MIT Center for Genome Research, http://www.genome.wi.mit.edu

References

- Boughman JA, Vernon M, Shaver KA (1983) Usher syndrome: definition and estimate of prevalence from two high risk populations. J Chronic Dis 36:595–603
- Chaib H, Kaplan J, Gerber S, Vincent C, Munnich A, Weissenbach J, Petit C (1997) A newly identified locus for Usher syndrome type I, USH1E, maps to chromosome 21q21. Hum Mol Genet 6:27–31
- Colognato H, Winkelmann DA, Yurchenco PD (1999) Laminin polymerization induces a receptor-cytoskeleton network. J Cell Biol 145:619–631
- Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res 16:10881–10890
- Culligan KG, Mackey AJ, Finn DM, Maguire PB, Ohlendieck K (1998) Role of dystrophin isoforms and associated proteins in muscular dystrophy. Int J Mol Med 2:639–648
- den Hollander AI, ten Brink JB, de Kok YJ, van Soest S, van Den Born LI, van Driel MA, van De Pol DJ, et al (1999) Mutations in a human homologue of *Drosophila* crumbs cause retinitis pigmentosa (RP12). Nat Genet 23:217–221
- Eudy JD, Weston MD, Yao S, Hoover DM, Rehm HL, Ma-Edmonds M, Yan D, et al (1998) Mutation of a gene encoding a protein with extracellular matrix motifs in Usher syndrome type IIa. Science 280:1753–1757
- Hmani M, Ghorbel A, Boulila-Elgaied A, Ben Zina Z, Kammoun W, Drira M, Chaabouni M, et al (1999) A novel locus for Usher syndrome type II, USH2B, maps to chromosome 3 at p23-24.2. Eur J Hum Genet 7:363–367
- Ishii N, Wadsworth WG, Stern BD, Culotti JG, Hedgecock EM (1992) UNC-6, a laminin-related protein, guides cell and pioneer axon migrations in *C. elegans*. Neuron 9:873–881
- Kaplan J, Gerber S, Bonneau D, Rozet JM, Delrieu O, Briard ML, Dollfus H, et al (1992) A gene for Usher syndrome type I (USH1A) maps to chromosome 14q. Genomics 14: 979–987
- Keats BJ, Nouri N, Pelias MZ, Deininger PL, Litt M (1994) Tightly linked flanking microsatellite markers for the Usher syndrome type I locus on the short arm of chromosome 11. Am J Hum Genet 54:681–686
- Kimberling W, Smith RJ (1992) Gene mapping of the Usher syndromes. Otolaryngol Clin North Am 25:923–934
- Kimberling WJ, Weston MD, Moller C, Davenport SL, Shugart YY, Priluck IA, Martini A, et al (1990) Localization of Usher syndrome type II to chromosome 1q. Genomics 7:245–249
- Koch M, Olson PF, Albus A, Jin W, Hunter DD, Brunken WJ, Burgeson RE, et al (1999) Characterization and expression of the laminin γ 3 chain: a novel, non-basement membraneassociated, laminin chain. J Cell Biol 145:605–618
- Liu XZ, Hope C, Liang CY, Zou JM, Xu LR, Cole T, Mueller RF, et al (1999) A mutation (2314delG) in the Usher syndrome type IIa gene: high prevalence and phenotypic variation. Am J Hum Genet 64:1221–1225
- Mellerio JE, Eady RA, Atherton DJ, Lake BD, McGrath JA

(1998) E210K mutation in the gene encoding the β 3 chain of laminin-5 (LAMB3) is predictive of a phenotype of generalized atrophic benign epidermolysis bullosa. Br J Dermatol 139:325–331

- Miner JH, Patton BL, Lentz SI, Gilbert DJ, Snider WD, Jenkins NA, Copeland NG, et al (1997) The laminin α chains: expression, developmental transitions, and chromosomal locations of α 1–5, identification of heterotrimeric laminins 8–11, and cloning of a novel α 3 isoform. J Cell Biol 137: 685–701
- Mount SM (1982) A catalogue of splice junction sequences. Nucleic Acids Res 10:459–472
- Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10:1–6
- Pieke-Dahl SA, Moller CG, Kelley PM, Astuto LM, Cremers CWRJ, Gorin MB, Kimberling WJ. Genetic heterogeneity of Usher syndrome type II: localization to chromosome 5q. Am J Med Genet (in press)
- Pieke-Dahl SA, Weston MD, Kimberling WJ (1997) Genetic heterogeneity of the Usher syndromes. Assoc Res Otolaryngol 20:A874
- Pulkkinen L, Uitto J (1999) Mutation analysis and molecular genetics of epidermolysis bullosa. Matrix Biol 18:29–42
- Sankila EM, Pakarinen L, Kaariainen H, Aittomaki K, Karjalainen S, Sistonen P, de la Chapelle A (1995) Assignment of an Usher syndrome type III (USH3) gene to chromosome 3q. Hum Mol Genet 4:93–98
- Smith RJ, Berlin CI, Hejtmancik JF, Keats BJ, Kimberling WJ,

Lewis RA, Moller CG, et al (1994) Clinical diagnosis of the Usher syndromes. Usher Syndrome Consortium. Am J Med Genet 50:32–38

- Smith RJ, Steel KP, Barkway C, Soucek S, Michaels L (1992) A histologic study of nonmorphogenetic forms of hereditary hearing impairment. Arch Otolaryngol Head Neck Surg 118:1085–1094
- Sunada Y, Bernier SM, Utani A, Yamada Y, Campbell KP (1995) Identification of a novel mutant transcript of laminin α 2 chain gene responsible for muscular dystrophy and dysmyelination in dy^{2J} mice. Hum Mol Genet 4:1055–1061
- Tessier-Lavigne M, Goodman CS (1996) The molecular biology of axon guidance. Science 274:1123–1133
- Timpl R (1996) Macromolecular organization of basement membranes. Curr Opin Cell Biol 8:618–624
- Wadsworth WG, Bhatt H, Hedgecock EM (1996) Neuroglia and pioneer neurons express UNC-6 to provide global and local netrin cues for guiding migrations in *C. elegans*. Neuron 16:35–46
- Wayne S, Der Kaloustian V, Schloss M, Polomeno R, Scott DA, Hejtmancik JF, Sheffield VC, et al (1996) Localization of the Usher syndrome type ID gene (Ush1D) to chromosome 10. Hum Mol Genet 5:1689–1692
- Wayne S, Lowry RB, McLeod DR, Knaus R, Farr C, Smith RJH (1997) Localization of the Usher syndrome type IF to chromosome 10. Am J Hum Genet 61:A300
- Yurchenco PD, Cheng YS (1993) Self-assembly and calciumbinding sites in laminin: a three-arm interaction model. J Biol Chem 268:17286–17299