ARTICLE

Conversion and Compensatory Evolution of the γ-Crystallin Genes and Identification of a Cataractogenic Mutation That Reverses the Sequence of the Human *CRYGD* Gene to an Ancestral State

Olga V. Plotnikova,* Fyodor A. Kondrashov,* Peter K. Vlasov, Anastasia P. Grigorenko, Evgeny K. Ginter, and Evgeny I. Rogaev

We identified a mutation in the *CRYGD* gene (P23S) of the γ-crystallin gene cluster that is associated with a polymorphic congenital cataract that occurs with frequency of ~0.3% in a human population. To gain insight into the molecular mechanism of the pathogenesis of γ-crystallin isoforms, we undertook an evolutionary analysis of the available mammalian and newly obtained primate sequences of the γ-crystallin genes. The cataract-associated serine at site 23 corresponds to the ancestral state, since it was found in *CRYGD* of a lower primate and all the surveyed nonprimate mammals. Crystallin proteins include two structurally similar domains, and substitutions in mammalian CRYGD protein at site 23 of the first domain were always associated with substitutions in the structurally reciprocal sites 109 and 136 of the second domain. These data suggest that the cataractogenic effect of serine at site 23 in the N-terminal domain of CRYGD may be compensated indirectly by amino acid changes in a distal domain. We also found that gene conversion was a factor in the evolution of the γ-crystallin gene cluster throughout different mammalian clades. The high rate of gene conversion observed between the functional *CRYGD* gene and two primate γ-crystallin pseudogenes (*CRYGEP1* and *CRYGFP1*) coupled with a surprising finding of apparent negative selection in primate pseudogenes suggest a deleterious impact of recently derived pseudogenes involved in gene conversion in the γ-crystallin gene cluster.

Cataracts are characterized by opaqueness of all or part of the eye crystallin lens[1] and are the most common cause of blindness in the world, with congenital cataracts frequently resulting in blindness or visual impairment in children.[2] The estimated prevalence is 2.2–2.49 cases per 10,000 live births,[3] and ~50% of all infantile cataract cases are genetic.[2] Most cases occur as isolated pathologies, but some forms are associated with other abnormalities.[4] Although congenital cataracts can be transmitted as a recessive or an X-linked trait, autosomal dominant inheritance occurs most frequently and exhibits both clinical variability and genetic heterogeneity.[3]

Clinical and molecular genetics studies have led to the identification of multiple candidate disease loci for congenital cataracts. Mutations in genes encoding four specific types of proteins have been described in association with the phenotype of nonsyndromic inherited cataracts. These include members of the α-, β-, and γ-crystallin families[5,6] (MIM +123580, +123590, *123610, *123620, *123630, *123631, +600929, +123680, +123690, and *123730), three transcription factors (MAF[7] [MIM *177075], PITX3[8] [MIM +602669], and HSF4[9] [MIM *602438]), cytoskeletal protein BFSP2[10] (MIM *603212), and membrane-transport proteins MIP[1] (MIM +154050), GJA3 (CX46)[11] (MIM *121015), and GJA8 (CX50)[12] (MIM *600897). Approximately half of all mutations associated with congenital cataracts are located in crystallin genes.[13]

Crystallins are the major water-soluble structural proteins expressed in the mammalian eye lens and consist of three major families—the α-, β-, and γ- crystallins[14]—with the γ-crystallin composing up to 40% of the soluble proteins expressed in the lens.[15] In humans, the γ-crystallin gene cluster is located on chromosome 2q33-q35 and consists of genes *CRYGA* (MIM *123660; GenBank accession numbers M17315 and M17316), *CRYGB* (MIM *123670; GenBank accession number M19364), *CRYGC* (MIM +123680; GenBank accession numbers K03003 and K03004), and *CRYGD* (MIM +123690; GenBank accession numbers K03005 and K03006)[16] (encoding γ*A*-, γ*B*-, γ*C*-, and γ*D*-crystallins, respectively), with cataract-associated mutations in two of these genes (*CRYGD* and *CRYGC*) that code for the most abundant γ-crystallin proteins in the lens.[17] Two other γ-crystallin genes—*CRYGEP1* (GenBank accession numbers K03007 and K03008) (encoding γ*E*-crystallin) and *CRYGFP1* (GenBank accession numbers K03009 and K03010) (encoding γ*F*-crystallin) (both MIM

From the Laboratory of Molecular Brain Genetics, Research Center of Mental Health (O.V.P.; A.P.G.; E.I.R.), Engelhardt Institute of Molecular Biology (P.K.V.), Research Center of Medical Genetics (E.K.G.), Vavilov Institute of General Genetics (E.I.R.), Russian Academy of Sciences, and Lomonosov Moscow State University (E.I.R.), Moscow; Department of Psychiatry, Brudnick Neuropsychiatric Research Institute, University of Massachusetts Medical School, Worcester (O.V.P.; A.P.G.; E.I.R.); and Section on Ecology, Behavior and Evolution, Division of Biological Sciences, University of California at San Diego, La Jolla (F.A.K.)

Received January 3, 2007; accepted for publication March 30, 2007; electronically published May 16, 2007.

Address for correspondence and reprints: Dr. Evgeny I. Rogaev, Department of Psychiatry, Brudnick Neuropsychiatric Research Institute, University of Massachusetts Medical School, 303 Belmont Street, Worcester, MA 01604. E-mail: Evgeny.Rogaev@umassmed.edu

* These two authors contributed equally to this work.

Am. J. Hum. Genet. 2007;81:32–43. 0002-9297/2007/8101-0004$15.00
DOI: 10.1086/518616

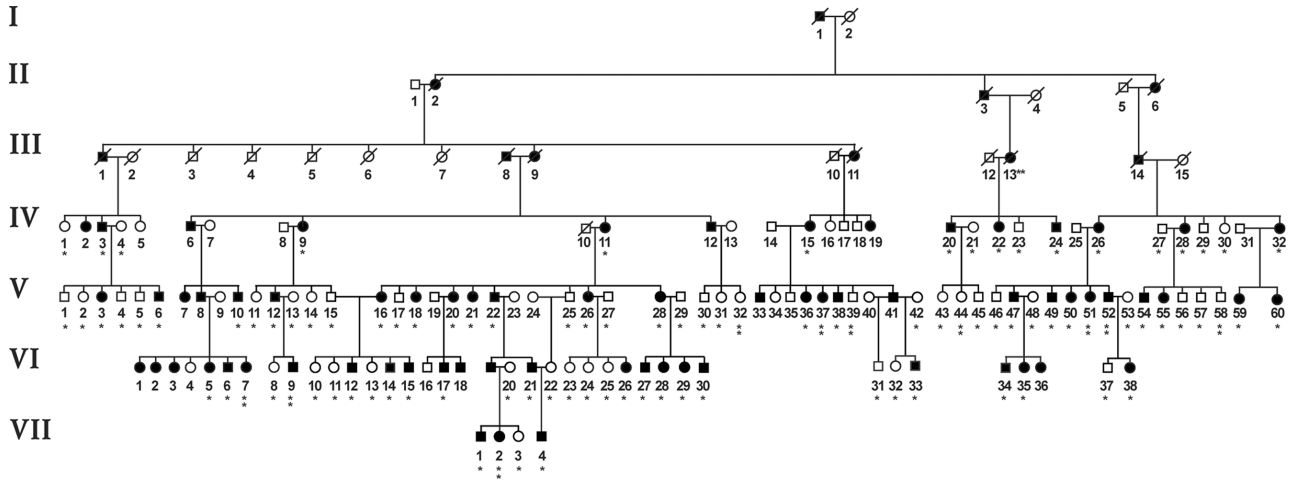

Figure 1. Abridged PCC-affected pedigree selected from the genetic isolate. The mutation is transmitted as an autosomal dominant trait. The affected individuals are represented by blackened squares (males) and circles (females), and the unaffected individuals are represented by unblackened symbols. Family members participating in this study are indicated by an asterisk (*). One asterisk indicates subjects genotyped by restriction enzyme–digestion analysis, and two asterisks indicate individuals genotyped by direct-sequencing analysis.

*123660)—are also located in the same cluster. However, in humans, they harbor a stop codon and are considered pseudogenes, whereas, in other nonprimate mammals, these genes appear functional.[18] Cataractogenesis is anticipated to be a strong factor in selection processes of genes for lens proteins; however, very little is yet known about the evolution of different members of γ-crystallin genes, especially in human and primate lineages. We have previously linked nonnuclear polymorphic congenital cataract (PCC [MIM %601286]) to the γ-crystallin gene cluster (*CRYG*) on the human chromosome 3q33-35 in a large pedigree from a Central Asian population.[19] Here, we screened the PCC-affected pedigree for mutations in the *CRYGA–CRYGD* genes and performed an evolutionary and structural analysis of the mutation and the γ-crystallin gene family.

Material and Methods

Mutation Analysis

The collection of DNA samples from subjects with PCC was described elsewhere.[19] The genomic sequence of human *CRYGA–CRYGD* genes was obtained from the GenBank database.[20] To search for mutations, the protein-coding regions of these genes were amplified by PCR, by use of genomic DNA from probands of the PCC-affected pedigree. Pairs of oligonucleotide primers flanking the exons of human *CRYGA–CRYGD* genes were designed manually or by Primer3 and were used for PCR amplification and sequencing of the PCR products (primer oligonucleotide sequences are available from the authors on request). PCR was performed for 32 cycles at 94°C for 3 min, with an annealing temperature of 56°C–58°C for 30 s, and at 72°C for 4 min. Each PCR was performed in a volume of 25 μl that contained 10–20 pmol of each primer, 1× reaction buffer, 50 ng DNA, 200 μM dNTP, 2.5–3 mM $MgCl_2$, and 0.2 U *Taq* polymerase. The PCR products were purified with electrophoresis in a 1% agarose gel, 1× TBE buffer, and the QIAEX II Kit gel extraction kit (QIAGEN). The purified PCR products were sequenced directly with use of an ABI Prism 310 Automated Sequencer with the ABI Prism BigDye Terminator cycle sequencing kits (Applied Biosystems).

The C70T mutation in the *CRYGD* gene was initially found in selected probands by direct sequencing. The presence or absence of the mutation was elucidated further by restriction enzyme–digestion assay in genomic DNA samples from all affected and unaffected family members of the pedigree. To distinguish the genotypes of unaffected and heterozygous individuals for this particular mutation, we designed nucleotide substitutions in one of the primers (reverse int) to create a new site for *Bpm*I-restriction endonuclease in the mutant C70T allele. Exon 2 of the *CRYGD* gene was amplified by two rounds of PCR with the primers direct ext (5′-GCAGCCCCACCCGCTCA-3′) and reverse ext (5′-GGGTAATACTTTGCTTATGTGGGG-3′) and then with internal primers direct int (5′-AGCCATGGGGAAGGTGAG-3′) and reverse int (5′-AGTAGGGCTGCAGGCTGG-3′). The PCR products were digested for 3–4 h at 37°C with *Bpm*I, and resulting DNA restriction fragments were analyzed on a 7% polyacrylamide gel and were visualized using ethidium bromide staining. In total, we analyzed 54 individuals with cataract and 46 unaffected individuals from the Middle Asian PCC pedigree. In addition, families with obesity from the same genetic isolate (22 individuals) were genotyped. We also tested 512 control chromosomes from white (206 chromosomes from Russians) and mixed white and Mongolian (224 chromosomes from Tatars and 82 chromosomes from Bashkirs) populations. The cataract-associated mutation (C70T) was detected in affected individuals from the PCC-affected pedigree only.

Sequencing of Primate Genes

To determine nucleotide sequences for ORFs of functional γ-crystallin genes (*CRYGA–CRYGD*) in primates, we used the PCR oligonucleotide primers based on human sequences or redundant oligonucleotide primers based on macaque, chimpanzee, and hu-

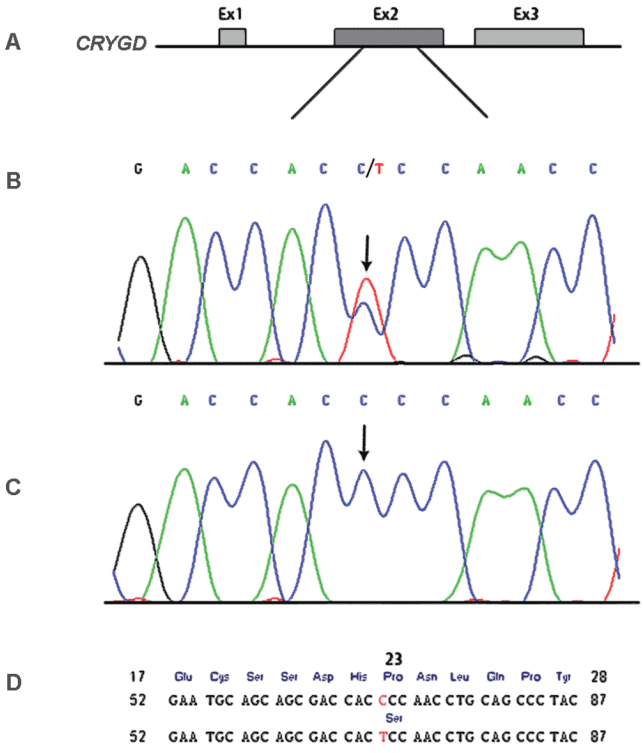

Figure 2. Identification of the mutation in individuals with PCC. *A,* The three exons (Ex) that comprise *CRYGD*. *B* and *C,* Sequence chromatogram from a heterozygous patient carrying the P23S mutation (*B*) and a homozygous wild-type sequence from an unaffected individual (*C*). *D,* Exon sequence of *CRYGD* bearing the cataract-associated PS23S mutation.

man intronic sequences, such that these primers flanked exons 1, 2, and 3. The PCRs and sequencing were performed as described above. The γ-crystallin gene sequences were determined in species of the following primate families: *Hominidae* (*Pan paniscus* [pygmy chimpanzee, GenBank accession numbers EF467187, EF467196, EF467205, and EF467214]; *Pan troglodytes* [chimpanzee, GenBank accession numbers EF467190, EF467199, EF467208, and EF467217]; *Gorilla gorilla* [gorilla, GenBank accession numbers EF467183, EF467192, EF467201, and EF467210]; *Pongo pigmaeus* [orangutan, GenBank accession numbers EF467188, EF467197, EF467206, and EF467215]); *Hylobatidae* (*Hylobates lar* [gibbon, GenBank accession numbers EF467184, EF467193, EF467202, and EF467211); *Cecropithecidae* (*Macaca mulatta* [rhesus monkey, GenBank accession numbers EF467186, EF467195, EF467204, and EF467213]); *Cebidae* (*Lagothrix lagotricha* [common woolly monkey, GenBank accession numbers EF467185, EF467194, EF467203, and EF467212]); *Ateles geoffroyi* [black-handed spider monkey, GenBank accession numbers EF467182, EF467191, EF467200, and EF467209]); and *Callitrichidae* (*Saguinus labiatus* [red-chested mustached tamarin, GenBank accession numbers EF467189, EF467198, EF467207, and EF467216]). In addition, we determined gene sequence for the putative pseudogene *CRYGFP1* in *Pan paniscus* (GenBank accession number EF492219), *G. gorilla* (GenBank accession number EF492217), *Pongo pigmaeus* (GenBank accession number EF492220), and *H. lar* (GenBank accession number EF492218). PCR primer oligonucleotide sequences are available from the authors on request. Genomic DNA samples were obtained from Coriell Cell Repositories and from our own collection of primate DNAs.

Bioinformatic and Structural Analysis

In addition to the sequenced primate genes, we used the sequence information of the γ-crystallin genes from GenBank.[20] We also used the information on gene order from completely sequenced mammalian genomes of *Monodelphis domestica, Canis familiaris, Mus musculus,* and *Bos taurus,* using the UCSC Genome Browser.[21] The rat genome (*Rattus norvegicus*) was excluded from the synteny analysis because of a likely error of assembly of the γ-crystallin cluster. A multiple alignment of all sequences was made using the MUSCLE alignment program.[22] We reconstructed the phylogeny of the γ-crystallin genes, using a Bayesean approach as implemented in MrBayes, with 1 million iterations (mcmc ngen = 1,000,000 in MrBayes) with the General Time Reversible model.[23] Sequence divergence between genes was estimated using the codeml program in the PAML package.[24] Exon 1, which encodes 3 aa, was omitted from the phylogenetic analysis.

Results

A Novel Pathogenic Mutation in the GRYGD *Gene*

Elsewhere, we established a link between PCC and the cluster of γ-crystallin genes (*CRYG*) at chromosome 2q33-35 in a large, unique pedigree of a family from a Central Asian population.[19] This population of mixed white and Mongolian origin is characterized by tribe ancestry, high endogamy, and complex ethnic genesis. Two inherited monogenic diseases were accumulated in these populations: an autosomal dominant cataract (PCC), with frequency of the mutant gene of ~0.26%, and autosomal recessive obesity, with frequency of the mutant gene of ~2.47%.[19,25,26] The large 7-generation PCC-affected pedigree characterized by high endogamy and a high coefficient of inbreeding (>3%) was selected for the molecular-genetic study. Of the 157 pedigree individuals, 105 had PCC, and DNA samples from 100 members, including 54 affected individuals, were available for mutation analysis of the *CRYG* cluster (fig. 1).

We amplified and sequenced the coding regions of the *CRYGA–CRYGD* genes and identified a novel nonsynonymous C70T (P23S) mutation in the coding region of exon 2 of the *CRYGD* gene, which was found to cosegregate with the disease (fig. 2). The presence or absence of the mutation was confirmed by sequence analysis of four affected and four unaffected individuals from the PCC-affected pedigree and then by restriction enzyme–digestion assay (as described in the "Material and Methods" section) of DNA samples from all available members from the PCC-

The figure is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Figure 3. Sequence alignment of γ-crystallins. The legend is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Species	Sites and residues													
Primates	22	23	24	...	48	49	50	...	108	109	110	...	136	137	138
γD-H. sap	H	P	N		P	N	Y		C	S	C		S	N	Y
γD-P.trog	H	P	N		P	N	Y		C	S	C		S	N	Y
γD-P.panis	H	P	N		P	N	Y		C	S	C		S	N	Y
γD-P.pyg	H	P	N		P	N	Y		C	S	C		S	N	Y
γD-G.gor	H	P	N		P	N	Y		C	S	C		S	N	Y
γD-H.lar	H	P	N		P	N	Y		C	S	C		S	N	Y
γD-M.mul	H	P	N		P	N	Y		C	S	S		P	N	Y
γD-A.geof	C	P	N		P	N	Y		C	T	S		P	N	Y
γD-L.lagot	C	P	N		P	N	Y		C	T	S		P	N	Y
γD-S.lab	H	S	N		P	N	Y		C	T	S		P	N	Y
Others															
γD-B.taur	H	S	N		P	N	Y		C	S	S		P	N	Y
γD-C.fam	H	S	N		P	N	Y		C	S	S		P	N	Y
γD-R.nor	H	S	N		P	N	F		C	P	S		T	N	Y
γD-M.mus	H	S	N		P	N	F		C	P	S		T	N	Y
Opossum	H	S	N		P	N	Y		C	S	S		P	N	Y

Figure 4. Sequence patterns in the interacting fragments of two γ-crystallin domains. Correlated serine and proline residues in CRYGD are shown in red and green, respectively. *H. sap* = *H. sapiens*; *P. trog* = *Pan troglodytes*; *P. panis* = *Pan paniscus*; *P. pyg* = *Pongo pygmaeus*; *G. gor* = *G. gorilla*; *M. mul* = *Macaca mulatta*; *A. geof* = *A. geoffroyi*; *L. lagot* = *L. lagotricha*; *S. lab* = *S. labiatus*; *B. taur* = *B. taurus*; *C. fam* = *C. familiaris*; *R. nor* = *R. norvegicus*; *M. mus* = *Mus musculus*.

affected pedigree. All affected individuals—but none of the related unaffected individuals from the PCC-affected pedigree or other unaffected, unrelated control individuals—were found to be heterozygous for this mutation. Three common synonymous SNPs in *CRYGB* (C192T) and *CRYGD* (T51C and T392C) were also detected. These polymorphic changes, however, showed no cosegregation with PCC in this pedigree. We found no *CRYGD* C70T cataract-associated mutation in unaffected individuals from the same genetic isolate or in 512 control chromosomes from populations of white or mixed white and Mongolian origin (see the "Material and Methods" section). The data strongly demonstrated that the nonsynonymous C70T (P23S) substitution in *CRYGD* is the only mutation in the γ-crystallin gene cluster that segregates with PCC.

Compensatory Evolution Reveals Structural Characteristics of the γ-Crystallins

Mutations that have a pathogenic effect when they are harbored in a human gene can be benign in other, sometimes closely related, organisms.[27,28] Such cases are known as "compensated pathogenic deviations" (CPDs), since it is thought that the deleterious effect of such mutations is neutralized by another, compensatory mutation. Unlike mutations in γ-crystallins described elsewhere, the mutation we describe here also appears to be a CPD, such that the disease-causing variant is found in the normal *CRYGD* sequence of several wild-type organisms, including one primate (fig. 3).

The basis of compensations of CPDs is usually the maintenance of structural stability within a single molecule,[27–30] although, in a few cases, the compensatory mutation and the CPD may be located on two different interacting proteins.[27] To investigate the molecular nature of the compensation of the P23S substitution, we assembled sequences of the γ-crystallin genes from the available mammalian genomes (human, chimpanzee, macaque, dog, mouse, rat, cow, and opossum) and the genew sequences for several primates that we determined in this study (*A. geoffroyi, S. labiatus, L. lagotricha, Macaca mulatta, H. lar, Pongo pygmaeus, G. gorilla, Pan paniscus,* and *Pan troglodytes*). The resulting multiple alignment was analyzed on the basis of the available crystal structure of the human CRYGD protein[31] (Protein Data Bank ID 1h4a).

The β- and γ-crystallin polypeptides fold into Greek key motifs that form two structurally similar domains.[31–33] On the γD-crystallin structure, site 23 interacts with position 49, and, since many compensatory substitutions for CPDs have been found in interacting sites, we surveyed the amino acid at site 49 in mammalian γD-crystallin genes. However, we found that site 49 and the neighboring sites are generally conserved throughout evolution and show no evidence of compensatory evolution with site 23 (fig. 4).

Thus, we undertook a correlation analysis in search of compensatory substitutions in the entire γD-crystallin protein. We searched for the compensatory site on the basis of the expected pattern of compensatory evolution[27]; that is, all species harboring serine at site 23 must have a single predicted compensatory amino acid at another site. No single site conformed to this prediction. However, sites 109 and 136 conformed in conjunction, such that the pathogenic state S23 was not observed together with the human state in either site 109 or 136 (fig. 4). Thus, in the course of evolution at site 23, in the common ancestor of

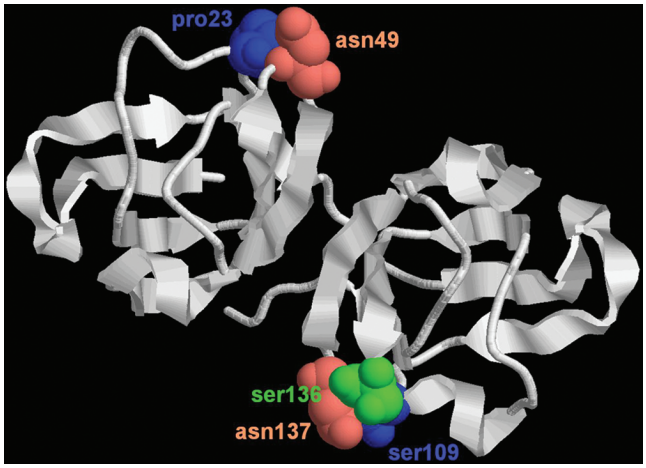

Figure 5. Contact regions in two symmetrical crystallin domains and four similar Greek key motifs form β-strands in two joined protein domains. The sample is based on the structure of the human γD-crystallin protein.

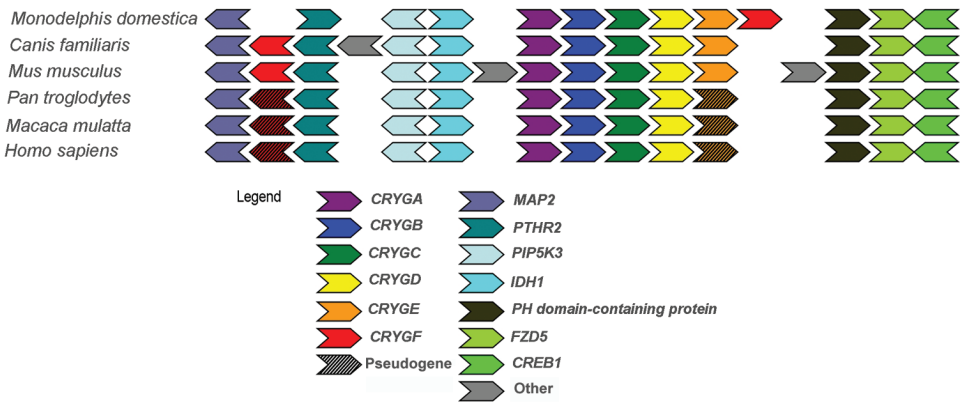

Figure 6. Gene order of paralogous γ-crystallin genes in mammalian genomes

primates, the ancestral serine changed to proline. In *S. labiatus,* the proline reversed to the ancestral serine, apparently without deleterious effects, whereas the same reversal in humans results in cataract formation. The deleterious effect of P23S substitution in humans is most likely related to the P136S substitution that also occurred in the primate, apparently the *Hominidae-Hylobatidae* common ancestor. Interestingly, the amino acid is also reverted to the ancestral state (S136P) in *S. labiatus* that have 23S in the wild-type allele. In general, the presence of serine at site 23 appears to be tolerated under the condition that either site 109 or site 136 is occupied by a proline. Remarkably, site 109 corresponds to the same position in the second domain as site 23 in the first γD-crystallin domain, whereas the corresponding site of the interaction site 49 in the second domain is site 137 (fig. 5).

Apparent Selection and Gene Conversion in the γ-Crystallin Gene Family

There is a large diversity of crystallin genes in higher animals; the γ-crystallin family is mammalian specific.[34] The γ-crystallin family is located in tandem with one of the genes (*CRYGFP1*) slightly removed from the rest of the cluster (fig. 6). The high sequence similarity of some of the γ-crystallin genes in humans[14] and a phylogenetic analysis of the gene family in rats[35] indicated that the γ-crystallin genes undergo gene conversion, which is apparently restricted to exon 2.[35] In addition, it is thought that two of the six γ-crystallin genes (*CRYGEP1* and *CRYGFP1*) have turned into a pseudogene in humans and chimpanzees, as evidenced by the presence of stop codons in the beginning of the second codon.[14,18] Several diseases are caused by gene conversion from a degenerate pseudogene into a functional gene,[36] including cataractogenesis by gene conversion of the β-crystallin pseudogene to one of the β-crystallin functional genes.[37] The presence of closely related pseudogenes in the γ-crystallins coupled with the reported gene-conversion events opens up the possibility that the same mechanism is the cause of some fraction of familial cataracts.

Gene conversion can make it particularly difficult to establish orthology[38]; therefore, we relied on synteny to resolve orthologous relationships within the γ-crystallin gene family. Indeed, the syntenic structure is well preserved within the mammalian clade, with only one of the genes (*CRYGFP1*) separated from an otherwise tandem arrangement of the γ-crystallin genes in the common ancestor of placental mammals (fig. 6).

In rats, gene conversion appears to preferentially affect the second exon of the γ-crystallins.[35] Thus, we constructed separate phylogenies for exons 2 and 3 for all six of the γ-crystallin genes for all available mammals, which, for the first time, included primates (fig. 7). On a phylogeny, gene-conversion events appear at a point of common ancestry of paralagous sequences.[38] In exon 2, gene conversion was found across all of the genes in the γ-crystallin family and in all surveyed taxa, but it was particularly common in nonprimate mammals and in *CRYGD, CRYGEP1,* and *CRYGFP1* (fig. 7*A*).

Since the divergence of the macaque and human lineages, there have been at least two fixed gene-conversion events in the macaque genome (one between *CRYGEP1* and *CRYGFP1* and one between *CRYGD* and the preconverted *CRYGE* or *CRYGFP1* sequence). In the human-chimpanzee lineage, there have been a gene-conversion event between *CRYGE* and *CRYGFP1* after the macaque split and a probable gene-conversion event around the time of divergence from macaques (fig. 7*A*). The rate of gene conversion in primates appears to be higher in *CRYGEP1, CRYGFP1,* and *CRYGD,* since only one gene-conversion event involving other genes (between *CRYGB* and *CRYGC*) was fixed around the time of divergence of the surveyed primate species (fig. 7*A*). The rate of gene conversion in exon 3 was not as high as in exon 2, but some conversion events were still observed (fig. 7*A*). Exon 3 of *CRYGEP1* and *CRYGFP1* was involved in at least five gene-conversion

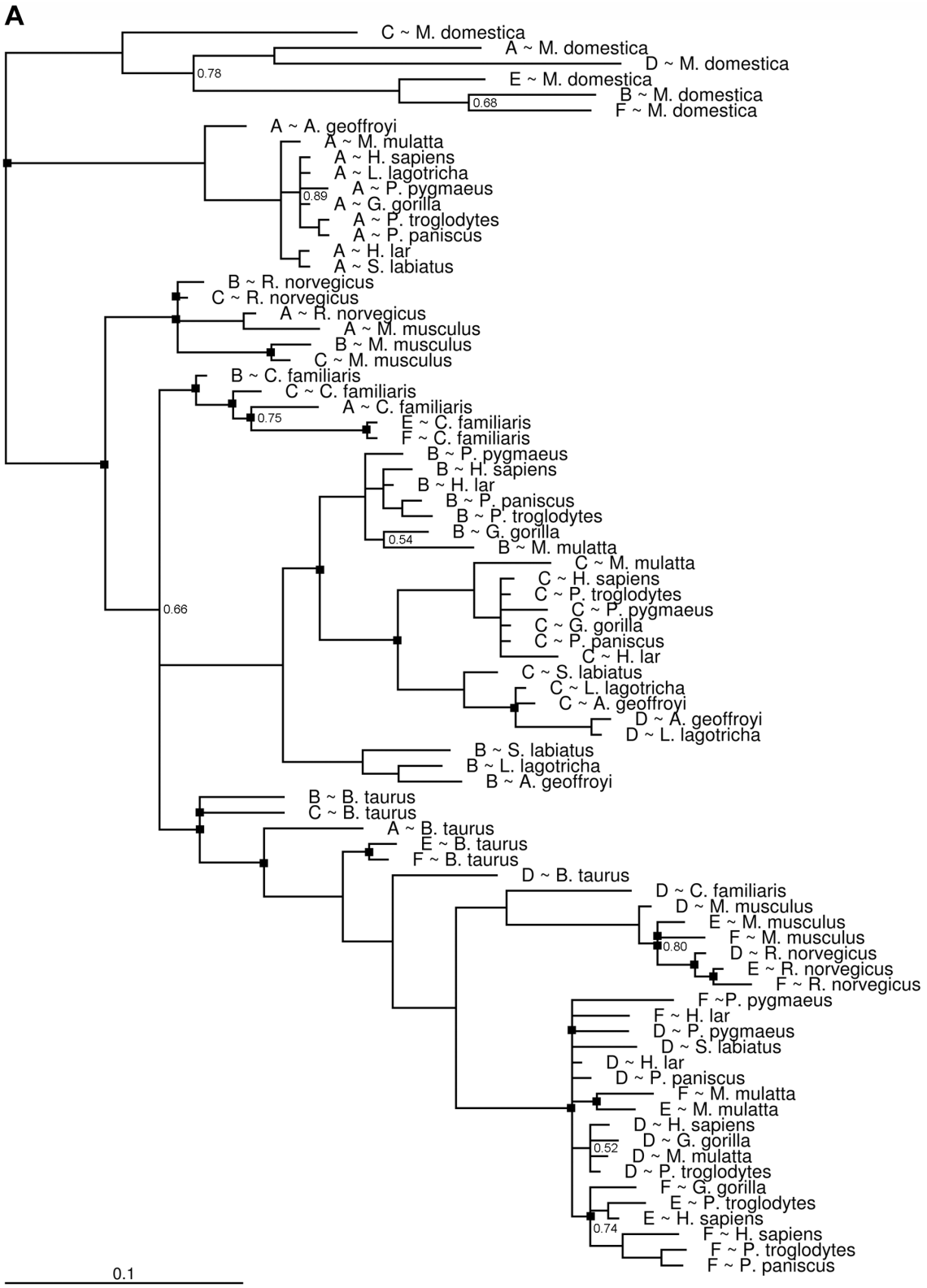

Figure 7. Bayesian phylogenies of exon 2 (*A*) and exon 3 (*B*) of the γ-crystallins. Gene-conversion events are shown on the phylogeny with a blackened square, and only posterior probabilities <0.90 are indicated. A, B, C, D, E, and F correspond to the *CRYGA, CRYGB, CRYGC, CRYGD, CRYGE*, and *CRYGF* genes, respectively.

B

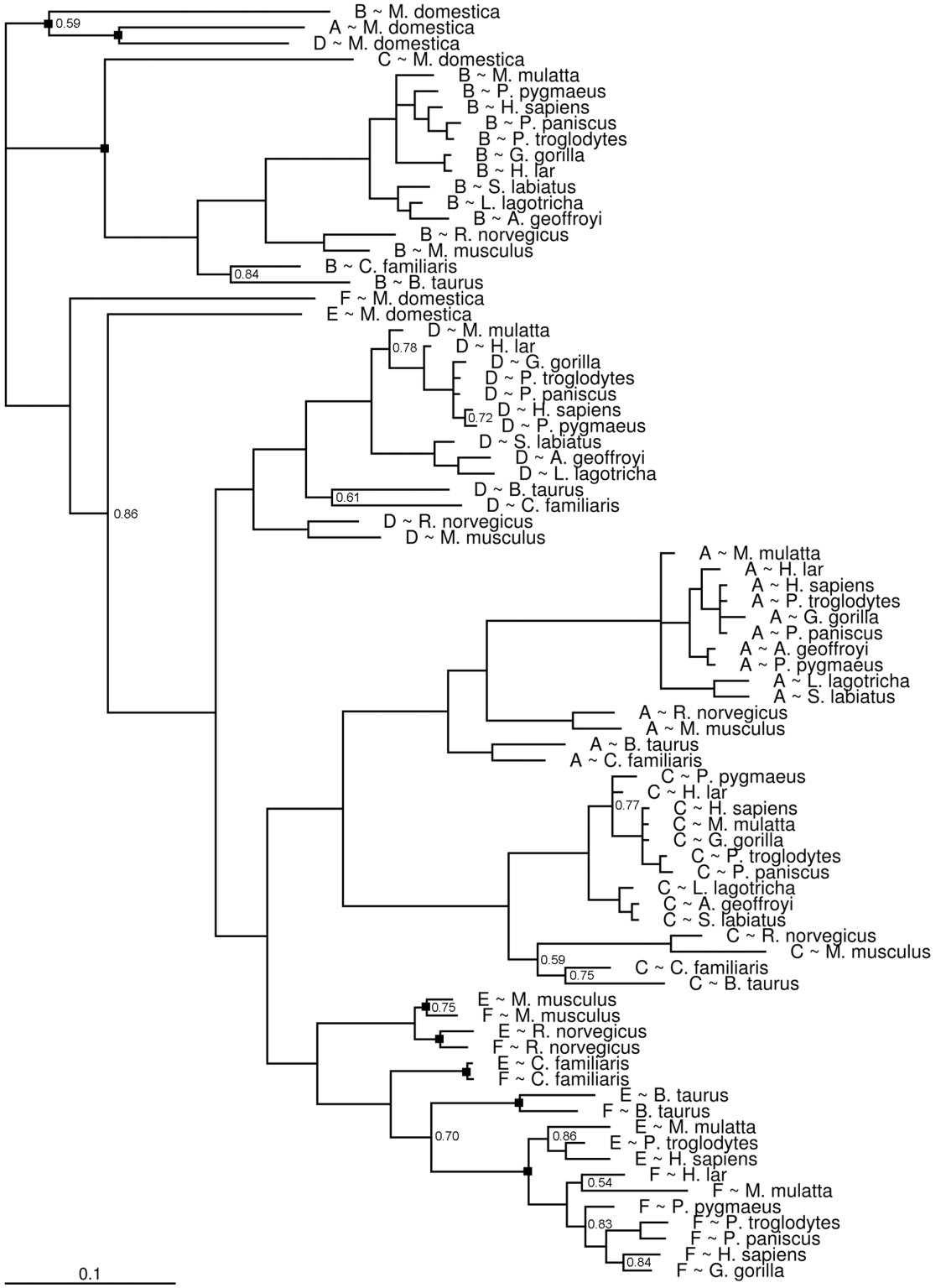
B ~ M. domestica
0.59
A ~ M. domestica
D ~ M. domestica
C ~ M. domestica
B ~ M. mulatta
B ~ P. pygmaeus
B ~ H. sapiens
B ~ P. paniscus
B ~ P. troglodytes
B ~ G. gorilla
B ~ H. lar
B ~ S. labiatus
B ~ L. lagotricha
B ~ A. geoffroyi
B ~ R. norvegicus
B ~ M. musculus
B ~ C. familiaris
0.84
B ~ B. taurus
F ~ M. domestica
E ~ M. domestica
D ~ M. mulatta
0.78
D ~ H. lar
D ~ G. gorilla
D ~ P. troglodytes
D ~ P. paniscus
0.72
D ~ H. sapiens
D ~ P. pygmaeus
D ~ S. labiatus
D ~ A. geoffroyi
D ~ L. lagotricha
D ~ B. taurus
0.61
D ~ C. familiaris
0.86
D ~ R. norvegicus
D ~ M. musculus
A ~ M. mulatta
A ~ H. lar
A ~ H. sapiens
A ~ P. troglodytes
A ~ G. gorilla
A ~ P. paniscus
A ~ A. geoffroyi
A ~ P. pygmaeus
A ~ L. lagotricha
A ~ S. labiatus
A ~ R. norvegicus
A ~ M. musculus
A ~ B. taurus
A ~ C. familiaris
C ~ P. pygmaeus
C ~ H. lar
0.77
C ~ H. sapiens
C ~ M. mulatta
C ~ G. gorilla
C ~ P. troglodytes
C ~ P. paniscus
C ~ L. lagotricha
C ~ A. geoffroyi
C ~ S. labiatus
C ~ R. norvegicus
C ~ M. musculus
0.59
C ~ C. familiaris
0.75
C ~ B. taurus
E ~ M. musculus
0.75
F ~ M. musculus
E ~ R. norvegicus
F ~ R. norvegicus
E ~ C. familiaris
F ~ C. familiaris
E ~ B. taurus
F ~ B. taurus
E ~ M. mulatta
0.70
0.86
E ~ P. troglodytes
E ~ H. sapiens
F ~ H. lar
0.54
F ~ M. mulatta
F ~ P. pygmaeus
0.83
F ~ P. troglodytes
F ~ P. paniscus
F ~ H. sapiens
0.84
F ~ G. gorilla
0.1

Table 1. Sequence Divergence between Human and Macaque γ-Crystallin Genes

Gene	No. of Substitutions		dn:ds[a]	P[b]
	Nonsynonymous	Synonymous		
CRYGA	4.12	6.25	.1338	<.05
CRYGB	7.22	14.93	.1001	<.001
CRYGC	5.16	5.08	.1953	NS
CRYGD	9.32	4.31	.3322	NS
CRYGEP1	11.46	15.12	.1041	<.001
CRYGFP1	21.22	17.06	.2236	<.01

[a] The rate of synonymous (ds) and nonsynonymous (dn) evolution.
[b] The *P* value of a χ^2 test of the number of synonymous and nonsynonymous substitutions observed compared with the number expected under neutrality. NS = not significant.

events since the radiation of placental mammals: two in rodents, one each in dog and cow, and one in primates before the divergence of humans and macaques.

The relatively high rate of gene conversion in primates that involves two of the genes that have been pseudogenized in humans and a functional gene reveals the possibility that some of the degenerate mutations that have accumulated in the pseudogene may be transferred to the functional gene, with deleterious or pathogenic consequences. Such frequent gene conversion between a pseudogene and a functional gene should lead to a pattern of negative selection in the pseudogene, because selection would act against degenerate mutations in the pseudogene if gene conversion events frequently transfer such mutations to a functional gene.[39,40]

Although *CRYGEP1* and *CRYGFP1* are thought to be pseudogenes in humans and chimpanzees,[18] it is not known whether they are functional in other primates. We retrieved the macaque *CRYGEP1* and *CRYGFP1* sequences from orthologous (determined by synteny) sections of the macaque genome, and we compared the obtained sequences with those of other γ-crystallin genes from different species. The macaque *CRYGFP1* sequence contained three frameshift mutations, strongly suggesting that it is a pseudogene. The functional status of *CRYGEP1* was not as apparent. Although the macaque *CRYGEP1* gene showed seven amino acid changes in evolutionarily conserved sites in all γ-crystallins, no inframe stop codons or frameshift mutations were observed.

A common way to test the strength of selection acting on protein-coding genes is to compare the rate of synonymous (ds) and nonsynonymous (dn) evolution.[38] A neutrally evolving sequence, such as a pseudogene, is expected to show the same rate of evolution in these two types of sites (dn:ds = 1). We compared the rates of evolution between human and macaque sequences for all six genes in the γ-crystalline cluster (table 1). We found no evidence to support a neutral level of evolution in *CRYGEP1* (dn:ds = 0.1041) that would suggest that *CRYGEP1* is an active gene in the macaque; however, we also found that the dn:ds ratio significantly differs from 1 even for the *CRYGFP1* pseudogene (dn:ds = 0.2236). Such apparent selection is commonly observed on the phenotypic level[41]; however, to our knowledge, it has been observed only once on a molecular level, in a mouse c-ubiquitin cluster by an uneven crossing-over event.[42] We determined the *CRYGFP* sequence in *Pan paniscus, Pongo pygmaeous,* and *H. lar* and found signs of pseudogenization of the gene in different primate species (table 2) (see the "Discussion" section). To our knowledge, this is the first study that observes apparent selection in nonfunctional sequence due to gene conversion with a functional gene.

Discussion

To date, six different mutations in the *CRYGD* gene have been found in patients with cataracts: R14C,[43] R36S,[44,45] R58H,[46] W156X,[47] P23T,[47] and E107A[48] (MIM +123690.0001–123690.0006). The seventh mutation identified here (P23S) appears to be the cause of the nonnuclear PCC. This mutation is accumulated in an isolated population (with frequency of ~0.26%) along with autosomal recessive obesity (with frequency of the mutant-gene allele of ~2.5%). We found no connection or cosegregation of these pathologies, which were probably inherited from different ancestral founders. The PCC type of cataract is characterized by a nonprogressive phenotype and partial opacity of the lens, which has a variable location on the periphery between the fetal nucleus of the lens and the equator. The opacities are irregular and look similar to a bunch of grapes or a lump of cotton balls and may be present simultaneously in different lens layers.[19] Another amino acid substitution at the same site has been described elsewhere, a proline→threonine substitution (P23T) in a family with congenital cataracts.[47,49–53] Generally, the clinical manifestation of the PCC-affected family was different from that of patients with other forms of familial cataracts caused by the P23T mutation of the *CRYGD* gene. The P23T mutation has been shown to cosegregate with variable phenotypes, such as the lamellar cataract in an Indian family and the fasciculiform, coralliform, and cerulean cataracts (MIM 608983) in a Moroccan family.[47,49–53] The segregation frequency of PCC strongly corresponds to an autosomal dominant form of inheritance (frequency [±SD] of 0.56 ± 0.04) with >90%–97% penetrance.[19,25] Thus, the variability of clinical phenotype may be caused by epigenetic factors during embryonic development or by gene modifiers.

In vitro experimental data about the nature of the P23T substitution suggested that this change does not significantly alter the stability structure of the protein but, rather, affects the protein solubility,[32,33] resulting in clusters of the P23T-mutant protein. Because of a high similarity of the side chains of thrionine and serine, Evans et al.[32] also considered the impact of a nonnatural P23S substitution and found that P23S also affects the solubility of γD-crystallin, although not as profoundly. This effect may be caused by a change in the hydrogen-binding characteristics of the protein-water interface. A substitution of a

Table 2. Signs of Pseudogenization in the Primate *CRYGFP1* Pseudogene

Species	Sign(s) of Pseudogenization
H. sapiens	56 Y→STOP and 1-nt deletion at site 422
Pan troglodytes	56 Y→STOP, 1-nt deletion at site 422 and deletion of the 5′ end of exon 3 sites 252–267
Pan paniscus	56 Y→STOP, 1-nt deletion at site 422, deletion of the 5′ end of exon 3 sites 252–267, and A insertion at site 455
G. gorilla	56 Y→STOP
Pongo pygmaeus	42 C→STOP
H. lar	Second intron splice sites AT...GG
Macaca mulatta	1 M→K, G deletion at site 30, 28-nt insertion at site 425, and 4-nt insertion at site 511

proline, since it is an imino acid that does not have a hydrogen bond–forming NH group, is particularly capable of affecting solubility in water. Thus, it is likely that the P23S substitution is compensated for in another part of the γD-crystallin protein.

This observation is in agreement with our observation that the probable compensatory substitution is a reversal to a proline in structurally reciprocal sites in a different domain. The distribution patterns of the proline/serine residues that we describe here may play an important role in the protein-water–system stabilization. In particular, these sites can be the part of the crystallin interaction interface with other lens proteins where, possibly, γD-crystallins are connected with each other. Substitutions in these sites may lead to protein aggregation in solution that dramatically changes the lens crystal transparency. The observation that two sites, 109 and 136, may compensate for the same CPD (P23S) is not completely unexpected, since examples of compensatory interactions involving more than two interacting sites have been described.[28,29]

We find evidence to support active gene conversion between the pseudogene *CRYGFP1* and the functional *CRYGD* gene copy in recent evolution (since the divergence of human and macaque lineages). In addition, we find that a conversion event between these two genes also occurred sometime in the primate lineage before the divergence of humans and macaques (fig. 7). The phylogeny can reveal only fixed gene-conversion events, such that, if gene conversion events still occur but are not fixed, they will not be observed. In particular, the probability of fixation of a gene-conversion event will be much lower if it brings a deleterious substitution into a functional gene.[39,40] Thus, gene conversion between the pseudogenes and the functional gene copies may have occurred with a much higher frequency on a mutational level than is apparent from the phylogeny that reveals only fixed events on an evolutionary scale.

The observation of apparent negative selection in pseudogenes depends on the assumption that these genes became pseudogenes before the divergence of humans and macaques. There are no shared stop codons or frameshift mutations between the human and macaque pseudogenes; however, frequent gene conversion, which is seen between these pseudogenes and their functional copies, will erase such shared stop codons. To demonstrate this point, we sequenced the *CRYGFP1* pseudogene from *Pan paniscus, G. gorilla, Pongo pygmaeus,* and *H. lar,* in addition to an already available sequence from *H. sapiens, P. troglodytes,* and *Macaca mulatta*. We found that the *CRYGFP1* sequence is a pseudogene in each of these species; however, whereas the higher ape species share a stop codon, other species show different signs of pseudogenization (table 2). The possibility of very recent and independent pseudogenizations of the same gene in four separate lineages is extremely remote and is unlikely to explain our observations.

There are two nonexclusive ways in which gene conversion between a functional gene and a pseudogene can lead to apparent negative selection. The first is gene conversion of the functional sequence over that of the pseudogene. A comparison of pseudogene sequences from two species would reveal apparent negative selection ($dn < ds$) if such gene-conversion events were fixed after the divergence of the two species. Alternatively, selection may act on mutations in the pseudogene if such mutations are deleterious when they are converted to the functional gene. Whereas the first model is simple and requires only one fixed gene conversion of a functional copy to the pseudogene, the latter model encompasses many parameters, such as mutation rate and gene-conversion rate, on a population level. To delineate the exact conditions under which gene conversion from a pseudogene to a functional gene can lead to apparent selection in a pseudogene, we would require an extensive population genetics model, which is beyond the scope of our work here. However, it is clear that, for apparent selection to show such a strong pattern of selection (dn:ds for the pseudogene γF-crystallin was very similar to dn:ds for other functional *CRYG* genes and substantially deviates from a neutral expectation of dn:ds = 1; see table 1), the rate of gene conversion should be at least on the order of the rate of emergence of the potentially deleterious substitutions in the pseudogene.

Since we observe apparent selection in the pseudogene, it is reasonable to hypothesize that some mutations that may cause cataracts in humans may originate in functional genes from gene-conversion events from either the *CRYGEP1* or the *CRYGFP1* pseudogene. Indeed, the activation of the *CRYGEP1* pseudogene may lead to Coppock-like cataracts,[16] and gene conversion leading to genetic disorders has been observed for several diseases,[36] including cataract formation through gene conversion in the β-crystallins.[37] The observation that the *CRYGD* gene has the most cataract-causing mutations described in humans (including P23T) of all functional *CRYG* genes[3] may be

explained either by a higher level of expression of the *CRYGD* gene compared with the other *CRYG* genes[18] or by a higher rate of gene conversion of *CRYGD* with pseudogenes, which may harbor such mutations.

We checked the sequence of mutations known to cause cataracts in humans against those of the two pseudogenes (fig. 3). Of the seven surveyed mutations, we found one that corresponded to a state found in a pseudogene—P23T in *CRYGD,* which is the most common and likely independently derived mutation underlying clinical heterogeneity for different forms of cataracts. It is unclear whether this or other mutations also originate through gene conversion. Nevertheless, our evolutionary analysis suggested that negative selection of γ-crystallin pseudogenes is likely driven by gene conversion of the pseudogenes with functional genes that may result in cataractogenic γ-crystallin alleles.

We describe a human polymorphic congenital cataract caused by a mutation that reversed an amino acid in the *CRYGD* gene to an ancestral state found in nonprimate mammals. This cataract-associated mutation may be compensated for by indirect mechanisms related to the overall protein solubility, through substitutions in a symmetric protein domain. In addition, we found gene-conversion events in the γ-crystallin gene cluster in several mammalian species that involve the interaction of pseudogenes and functional genes in the primate lineage. The observed negative selection in the pseudogene in the course of human-macaque divergence is likely to be the result of apparent selection due to frequent gene conversions between the pseudogenes and the functional genes. The data suggest that some cataractogenic mutations might appear in functional γ-crystallin genes from pseudogenes through gene-conversion events contributing to conservation of the pseudogene sequence.

Acknowledgments

This study was supported by the Biodiversity and Dynamics of Gene Pools program of the Presidium of the Russian Academy of Sciences (support to E.I.R.). E.I.R. is also supported in part by the National Institute of Diabetes and Digestive and Kidney Diseases and National Institute of Neurological Disorders and Stroke (National Institutes of Health), and F.A.K. is supported by a National Science Foundation graduate research fellowship.

Web Resources

Accession numbers and URLs for data presented herein are as follows:

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for sequence information on *CRYGA* [accession numbers M17315 and M17316], *CRYGB* [accession number M19364], *CRYGC* [accession numbers K03003 and K03004], *CRYGD* [accession numbers K03005 and K03006], *CRYGEP1* [accession numbers K03007 and K03008], and *CRYGFP1* [accession numbers K03009 and K03010]; *Pan paniscus CRYGA, CRYGB, CRYGC, CRYGD,* and *CRYGFP1* [accession numbers EF467187, EF467196, EF467205, EF467214, and EF492219, respectively]; *Pan troglodytes CRYGA, CRYGB, CRYGC,* and *CRYGD* [accession numbers EF467190, EF467199, EF467208, and EF467217, respectively], *G. gorilla CRYGA, CRYGB, CRYGC, CRYGD,* and *CRYGFP1* [accession numbers EF467183, EF467192, EF467201, EF467210, and EF492217, respectively]; *Pongo pygmaeus CRYGA, CRYGB, CRYGC, CRYGD,* and *CRYGFP1* [accession numbers EF467188, EF467197, EF467206, EF467215, and EF492220, respectively]; *H. lar CRYGA, CRYGB, CRYGC, CRYGD,* and *CRYGFP1* [accession numbers EF467184, EF467193, EF467202, EF467211, and EF492218, respectively]; *Macaca mulatta CRYGA, CRYGB, CRYGC,* and *CRYGD* [accession numbers EF467186, EF467195, EF467204, and EF467213, respectively]; *L. lagotricha CRYGA, CRYGB, CRYGC,* and *CRYGD* [accession numbers EF467185, EF467194, EF467203, and EF467212, respectively]; *A. geoffroyi CRYGA, CRYGB, CRYGC,* and *CRYGD* [accession numbers EF467182, EF467191, EF467200, and EF467209, respectively]; and *S. labiatus CRYGA, CRYGB, CRYGC,* and *CRYGD* [accession numbers EF467189, EF467198, EF467207, and EF467216, respectively])

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for PCC and other genetic forms of cataracts)

Primer3, http://frodo.wi.mit.edu/

Protein Data Bank, http://www.pdb.org/ (for the human CRYGD protein [ID 1h4a])

University of California Santa Cruz (UCSC) Genome Browser, http://genome.ucsc.edu/

References

1. Francis PJ, Berry V, Moore AT, Bhattacharya S (1999) Lens biology: development and human cataractogenesis. Trends Genet 15:191–196
2. Francis PJ, Berry V, Bhattacharya SS, Moore AT (2000) The genetics of childhood cataract. J Med Genet 37:481–488
3. Reddy MA, Francis PJ, Berry V, Bhattacharya SS, Moore AT (2004) Molecular genetic basis of inherited cataract and associated phenotypes. Surv Ophthalmol 49:300–315
4. SanGiovanni JP, Chew EY, Reed GF, Remaley NA, Bateman JB, Sugimoto TA, Klebanoff MA (2002) Infantile cataract in the collaborative perinatal project: prevalence and risk factors. Arch Ophthalmol 120:1559–1565
5. Blundell T, Lindley P, Miller L, Moss D, Slingsby C, Tickle I, Turnell B, Wistow G (1981) The molecular structure and stability of the eye lens: x-ray analysis of γ-crystallin II. Nature 289:771–777
6. Graw J (1997) The crystallins: genes, proteins and diseases. Biol Chem 378:1331–1348
7. Jamieson RV, Perveen R, Kerr B, Carette M, Yardley J, Heon E, Wirth MG, van Heyningen V, Donnai D, Munier F, et al (2002) Domain disruption and mutation of the bZIP transcription factor, MAF, associated with cataract, ocular anterior segment dysgenesis and coloboma. Hum Mol Genet 11:33–42
8. Semina EV, Ferrell RE, Mintz-Hittner HA, Bitoun P, Alward WL, Reiter RS, Funkhauser C, ack-Hirsch S, Murray JC (1998) A novel homeobox gene *PITX3* is mutated in families with autosomal-dominant cataracts and ASMD. Nat Genet 19:167–170
9. Bu L, Jin Y, Shi Y, Chu R, Ban A, Eiberg H, Andres L, Jiang H,

Zheng G, Qian M, et al (2002) Mutant DNA-binding domain of HSF4 is associated with autosomal dominant lamellar and Marner cataract. Nat Genet 31:276–278
10. Conley YP, Erturk D, Keverline A, Mah TS, Keravala A, Barnes LR, Bruchis A, Hess JF, FitzGerald PG, Weeks DE, et al (2000) A juvenile-onset, progressive cataract locus on chromosome 3q21-q22 is associated with a missense mutation in the beaded filament structural protein–2. Am J Hum Genet 66: 1426–1431
11. Bassnett S, Missey H, Vucemilo I (1999) Molecular architecture of the lens fiber cell basal membrane complex. J Cell Sci 112:2155–2165
12. Berry V, Francis P, Kaushal S, Moore A, Bhattacharya S (2000) Missense mutations in MIP underlie autosomal dominant "polymorphic" and lamellar cataracts linked to 12q. Nat Genet 25:15–17
13. Hejtmancik JF, Smaoui N (2003) Molecular genetics of cataract. Dev Ophthalmol 37:67–82
14. Meakin SO, Breitman ML, Tsui LC (1985) Structural and evolutionary relationships among five members of the human γ-crystallin gene family. Mol Cell Biol 5:1408–1414
15. Meakin SO, Du RP, Tsui LC, Breitman ML (1987) γ-Crystallins of the human eye lens: expression analysis of five members of the gene family. Mol Cell Biol 7:2671–2679
16. Brakenhoff RH, Henskens HA, van Rossum MW, Lubsen NH, Schoenmakers JG (1994) Activation of the γ E-crystallin pseudogene in the human hereditary Coppock-like cataract. Hum Mol Genet 3:279–283
17. Pande A, Pande J, Asherie N, Lomakin A, Ogun O, King JA, Lubsen NH, Walton D, Benedek GB (2000) Molecular basis of a progressive juvenile-onset hereditary cataract. Proc Natl Acad Sci USA 97:1993–1998
18. Brakenhoff RH, Aarts HJ, Reek FH, Lubsen NH, Schoenmakers JG (1990) Human γ-crystallin genes: a gene family on its way to extinction. J Mol Biol 216:519–532
19. Rogaev EI, Rogaeva EA, Korovaitseva GI, Farrer LA, Petrin AN, Keryanov SA, Turaeva S, Chumakov I, St George-Hyslop P, Ginter EK (1996) Linkage of polymorphic congenital cataract to the γ-crystallin gene locus on human chromosome 2q33–35. Hum Mol Genet 5:699–703
20. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2006) GenBank. Nucleic Acids Res 34:D16–D20
21. Kuhn RM, Karolchik D, Zweig AS, Trumbower H, Thomas DJ, Thakkapallayil A, Sugnet CW, Stanke M, Smith KE, Siepel A, et al (2006) The UCSC Genome Browser database: update 2007. Nucleic Acids Res 34:D590–D598
22. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32: 1792–1797
23. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574
24. Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci 13: 555–556
25. Ginter EK, Turaeva S, Revazov AA, Panteleeva OA, Artykov OA (1983) Medical genetic study of the population of Turkmenia. III. Hereditary pathology in Turkmen Nokhurlis. Genetika 19:1344–1352
26. Ginter EK, Petrin AN, Spitsyn VA, Rogaev EI (1991) An attempt to locate the gene for congenital cataracts using linkage analysis. Genetika 27:1840–1849
27. Kondrashov AS, Sunyaev S, Kondrashov FA (2002) Dobzhansky-Muller incompatibilities in protein evolution. Proc Natl Acad Sci USA 99:14878–14883
28. Kern AD, Kondrashov FA (2004) Mechanisms and convergence of compensatory evolution in mammalian mitochondrial tRNAs. Nat Genet 36:1207–1212
29. Fukami-Kobayashi K, Schreiber DR, Benner SA (2002) Detecting compensatory covariation signals in protein evolution using reconstructed ancestral sequences. J Mol Biol 319: 729–743
30. Poon A, Chao L (2005) The rate of compensatory mutation in the DNA bacteriophage phiX174. Genetics 170:989–999
31. Basak A, Bateman O, Slingsby C, Pande A, Asherie N, Ogun O, Benedek GB, Pande J (2003) High-resolution X-ray crystal structures of human γD crystallin (1.25 A) and the R58H mutant (1.15 A) associated with aculeiform cataract. J Mol Biol 328:1137–1147
32. Evans P, Wyatt K, Wistow GJ, Bateman OA, Wallace BA, Slingsby C (2004) The P23T cataract mutation causes loss of solubility of folded γD-crystallin. J Mol Biol 343:435–444
33. Pande A, Annunziata O, Asherie N, Ogun O, Benedek GB, Pande J (2005) Decrease in protein solubility and cataract formation caused by the Pro23 to Thr mutation in human γ D-crystallin. Biochemistry 44:2491–2500
34. Aarts HJ, den Dunnen JT, Leunissen J, Lubsen NH, Schoenmakers JG (1988) The γ-crystallin gene families: sequence and evolutionary patterns. J Mol Evol 27:163–172
35. den Dunnen JT, Moormann RJ, Lubsen NH, Schoenmakers JG (1986) Concerted and divergent evolution within the rat γ-crystallin gene family. J Mol Biol 189:37–46
36. Bischof JM, Chiang AP, Scheetz TE, Stone EM, Casavant TL, Sheffield VC, Braun TA (2006) Genome-wide identification of pseudogenes capable of disease-causing gene conversion. Hum Mutat 27:545–552
37. Vanita, Sarhadi V, Reis A, Jung M, Singh D, Sperling K, Singh IR, Burger I (2001) A unique form of autosomal dominant cataract explained by gene conversion between β-crystallin B2 and its pseudogene. J Med Genet 38:392–396
38. Li WH (1997) Molecular evolution. Sinauer, Sunderland, MA
39. Teshima KM, Innan H (2004) The effect of gene conversion on the divergence between duplicated genes. Genetics 166: 1553–1560
40. Kondrashov FA, Gurbich TA, Vlasov PV (2007) Selection for functional uniformity of *tuf* duplicates in γ-proteobacteria. Trends Genet 23:215–218
41. Barton NH (1990) Pleiotropic models of quantitative variation. Genetics 124:773–782
42. Perelygin AA, Kondrashov FA, Rogozin IB, Brinton MA (2002) Evolution of the mouse polyubiquitin-C gene. J Mol Evol 55: 202–210
43. Stephan DA, Gillanders E, Vanderveen D, Freas-Lutz D, Wistow G, Baxevanis AD, Robbins CM, VanAuken A, Quesenberry MI, Bailey-Wilson J, et al (1999) Progressive juvenile-onset punctate cataracts caused by mutation of the γD-crystallin gene. Proc Natl Acad Sci USA 96:1008–1012
44. Kmoch S, Brynda J, Asfaw B, Bezouska K, Novak P, Rezacova P, Ondrova L, Filipec M, Sedlacek J, Elleder M (2000) Link between a novel human γD-crystallin allele and a unique cataract phenotype explained by protein crystallography. Hum Mol Genet 9:1779–1786
45. Gu J, Qi Y, Wang L, Wang J, Shi L, Lin H, Li X, Su H, Huang

S (2005) A new congenital nuclear cataract caused by a missense mutation in the γD-crystallin gene (*CRYGD*) in a Chinese family. Mol Vis 11:971–976

46. Héon E, Priston M, Schorderet DF, Billingsley GD, Girard PO, Lubsen N, Munier FL (1999) The γ-crystallins and human cataracts: a puzzle made clearer. Am J Hum Genet 65:1261–1267
47. Santhiya ST, Shyam MM, Rawlley D, Vijayalakshmi P, Namperumalsamy P, Gopinath PM, Loster J, Graw J (2002) Novel mutations in the γ-crystallin genes cause autosomal dominant congenital cataracts. J Med Genet 39:352–358
48. Messina-Baas OM, Gonzalez-Huerta LM, Cuevas-Covarrubias SA (2006) Two affected siblings with nuclear cataract associated with a novel missense mutation in the *CRYGD* gene. Mol Vis 12:995–1000
49. Mackay DS, Andley UP, Shiels A (2004) A missense mutation in the γD crystallin gene (*CRYGD*) associated with autosomal dominant "coral-like" cataract linked to chromosome 2q. Mol Vis 10:155–162
50. Xu WZ, Zheng S, Xu SJ, Huang W, Yao K, Zhang SZ (2004) Autosomal dominant coralliform cataract related to a missense mutation of the γD-crystallin gene. Chin Med J 117: 727–732
51. Nandrot E, Slingsby C, Basak A, Cherif-Chefchaouni M, Benazzouz B, Hajaji Y, Boutayeb S, Gribouval O, Arbogast L, Berraho A, et al (2003) Gamma-D crystallin gene (CRYGD) mutation causes autosomal dominant congenital cerulean cataracts. J Med Genet 40:262–267
52. Shentu X, Yao K, Xu W, Zheng S, Hu S, Gong X (2004) Special fasciculiform cataract caused by a mutation in the γD-crystallin gene. Mol Vis 10:233–239
53. Burdon KP, Wirth MG, Mackey DA, Russell-Eggitt IM, Craig JE, Elder JE, Dickinson JL, Sale MM (2004) Investigation of crystallin genes in familial cataract, and report of two disease associated mutations. Br J Ophthalmol 88:79–83