



HAL
open science

Mutation spectrum of EYS in Spanish patients with autosomal recessive retinitis pigmentosa

Isabel Barragan, Salud Borrego, Juan Ignacio Pieras, María González-del Pozo, Javier Santoyo, Carmen Ayuso, Montserrat Baiget, José María Millán, Marcela Daniela Mena, Mai Abd El-Aziz, et al.

► **To cite this version:**

Isabel Barragan, Salud Borrego, Juan Ignacio Pieras, María González-del Pozo, Javier Santoyo, et al.. Mutation spectrum of EYS in Spanish patients with autosomal recessive retinitis pigmentosa. *Human Mutation*, 2010, 31 (11), 10.1002/humu.21334 . hal-00573065

HAL Id: hal-00573065

<https://hal.science/hal-00573065>

Submitted on 3 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Mutation spectrum of *EYS* in Spanish patients with autosomal recessive retinitis pigmentosa

Journal:	<i>Human Mutation</i>
Manuscript ID:	humu-2010-0146.R1
Wiley - Manuscript type:	Mutation in Brief
Date Submitted by the Author:	11-Jun-2010
Complete List of Authors:	<p>Barragan, Isabel; Hospital Universitario Virgen del Rocío/CSIC/ Universidad de Sevilla, Unidad de Gestión Clínica de Genética, Reproducción y Medicina Fetal, Instituto de Biomedicina de Sevilla (IBIS); Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p> <p>Borrego, Salud; Hospital Universitario Virgen del Rocío/CSIC/ Universidad de Sevilla, Unidad de Gestión Clínica de Genética, Reproducción y Medicina Fetal, Instituto de Biomedicina de Sevilla (IBIS); Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p> <p>Pieras, Juan Ignacio; Hospital Universitario Virgen del Rocío/CSIC/ Universidad de Sevilla, Unidad de Gestión Clínica de Genética, Reproducción y Medicina Fetal, Instituto de Biomedicina de Sevilla (IBIS); Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p> <p>González-del Pozo, María; Hospital Universitario Virgen del Rocío/CSIC/ Universidad de Sevilla, Unidad de Gestión Clínica de Genética, Reproducción y Medicina Fetal, Instituto de Biomedicina de Sevilla (IBIS); Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p> <p>Santoyo, Javier; Centro de Investigación Príncipe Felipe, Departamento de Bioinformática y Genómica; Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p> <p>Ayuso, Carmen; Instituto de Investigación Sanitaria Fundación Jiménez Díaz-CIBERER, Departamento de Genética; Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p> <p>Baiget, Montserrat; Hospital de la Santa Creu i Sant Pau, Servei de Genética; Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p> <p>Millán, José; Hospital Universitari La Fe, Unidad de Genética; Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p> <p>Mena, Marcela; Hospital Universitario Virgen del Rocío/CSIC/ Universidad de Sevilla, Unidad de Gestión Clínica de Genética,</p>

	<p>Reproducción y Medicina Fetal, Instituto de Biomedicina de Sevilla (IBIS); Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p> <p>ABD EL-AZIZ, Mai; UCL-Institute of Ophthalmology</p> <p>Audo, Isabelle; INSERM, U968; CNRS, UMR_7210; UPMC Univ Paris 06, UMR_S 968, Institut de la Vision; Centre Hospitalier National d'Ophthalmologie des Quinze-Vingts, INSERM-DHOS CIC 503; Institute of Ophthalmology, Department of Molecular Genetics</p> <p>Zeitz, Christina; INSERM, U968; CNRS, UMR_7210; UPMC Univ Paris 06, UMR_S 968, Institut de la Vision</p> <p>Littink, Karin; Radboud University Nijmegen Medical Centre, Department of Human Genetics; The Rotterdam Eye Hospital</p> <p>Dopazo, Joaquín; Centro de Investigación Príncipe Felipe, Departamento de Bioinformática y Genómica; Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p> <p>Bhattacharya, Shomi; CABIMER, Cellular Therapy and Regenerative Medicine</p> <p>Antañolo, Guillermo; Hospital Universitario Virgen del Rocío/CSIC/ Universidad de Sevilla, Unidad de Gestión Clínica de Genética, Reproducción y Medicina Fetal, Instituto de Biomedicina de Sevilla (IBIS); Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)</p>
Key Words:	EYS, Retinitis Pigmentosa, Spanish population, mutation, functional domain, recurrent mutation

 scholarONE™
Manuscript Central

Review

Mutation spectrum of *EYS* in Spanish patients with autosomal recessive retinitis pigmentosa

BARRAGÁN I.^{1,2}, BORREGO S.^{1,2}, PIERAS J.I.^{1,2}, GONZÁLEZ-DEL POZO M.^{1,2}, SANTOYO J.^{2,3}, AYUSO C.^{2,4}, BAIGET M.^{2,5}, MILLÁN J.M.^{2,6}, MENA M.^{1,2}, ABD EL-AZIZ M. M.⁷, AUDO I.^{8,9,10,11,12}, ZEITZ C.^{8,9,10}, LITTINK K.W.^{13,14}, DOPAZO J.^{2,3}, BHATTACHARYA S. S.¹⁵ and ANTIÑOLO G.^{1,2}

¹ *Unidad de Gestión Clínica de Genética, Reproducción y Medicina Fetal, Instituto de Biomedicina de Sevilla (IBIS), Hospital Universitario Virgen del Rocío/CSIC/ Universidad de Sevilla, Sevilla, Spain*

² *Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Valencia, Spain*

³ *Departamento de Bioinformática y Genómica, Centro de Investigación Príncipe Felipe, Valencia, Spain.*

⁴ *Departamento de Genética, Instituto de Investigación Sanitaria Fundación Jiménez Díaz, Madrid, Spain*

⁵ *Servei de Genética, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain*

⁶ *Unidad de Genética, Hospital Universitari La Fe, Valencia, Spain*

⁷ *UCL-Institute of Ophthalmology, Bath Street, London, EC1V 9EL. United Kingdom*

⁸ *INSERM, U968, Paris, F-75012, France*

⁹ *CNRS, UMR_7210. Paris, F-75012, France*

¹⁰ *UPMC Univ Paris 06, UMR_S 968, Institut de la Vision, Paris, F-75012, France*

¹¹ *Centre Hospitalier National d'Ophthalmologie des Quinze-Vingts, INSERM-DHOS CIC 503, Paris, F-75012, France*

¹² *Department of Molecular Genetics, Institute of Ophthalmology, London, UK*

¹³ *Department of Human Genetics, Radboud University Nijmegen Medical Centre, P.O. Box 9101, 6500 HB Nijmegen, The Netherlands.*

¹⁴ *The Rotterdam Eye Hospital, P.O. Box 70030, 3000 LM Rotterdam, The Netherlands.*

¹⁵ *Department of Cellular Therapy and Regenerative Medicine, Andalusian Molecular Biology and Regenerative Medicine Centre (CABIMER), Sevilla, Spain*

*Correspondence to Guillermo Antiñolo, MD, PhD

Unidad de Gestión Clínica de Genética, Reproducción y Medicina Fetal, Instituto de Biomedicina de Sevilla (IBIS)

Hospital Universitario Virgen del Rocío

Avenida Manuel Siurot s/n, 41013 Sevilla, Spain

Phone +34 955 012 772 E-mail: guillermo.antinolo.ssapa@juntadeandalucia.es

or

Professor Shomi S Bhattacharya,

Received <date>; accepted revised manuscript <date>.

2 <Barragán et al.>

Department of Celular Therapy and Regenerative Medicine
Andalusian Molecular Biology and Regenerative Medicine Centre (CABIMER)
Avda. Americo Vespucio s/n. Edif. CABIMER
Parque Científico y Tecnológico Cartuja 93,1092 - Sevilla (SPAIN)
Phone +34 954 468 004 E-mail: shomi.bhattacharya@cabimer.es

Contract grant sponsor: Fondo de Investigación Sanitaria, Spain; Consejería de Salud; Consejería de Innovación, Ciencia y Empresa; Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER).
; Contract grant number: PI081131, PI-0334/2007, PI08-CTS-03687

Short Title: Mutation spectrum of *EYS*.

Communicated by <Please don't enter>

ABSTRACT: Retinitis pigmentosa (RP) is a heterogeneous group of inherited retinal dystrophies characterised ultimately by the loss of photoreceptor cells. We have recently identified a new gene (*EYS*) encoding an ortholog of *Drosophila* spacemaker (spam) as a commonly mutated gene in autosomal recessive RP. In the present study, we report the identification of 73 sequence variations in *EYS*, of which 28 are novel. Of these, 42.9% (12/28) are very likely pathogenic, 17.9% (5/28) are possibly pathogenic, whereas 39.3% (11/28) are SNPs. In addition, we have detected 3 pathogenic changes previously reported in other populations. We are also presenting the characterisation of *EYS* homologues in different species, and a detailed analysis of the *EYS* domains, with the identification of an interesting novel feature: a putative coiled-coil domain. Majority of the mutations in the arRP patients have been found within the domain structures of *EYS*. The minimum observed prevalence of distinct *EYS* mutations in our group of patients is of 15.9% (15/94), confirming a major involvement of *EYS* in the pathogenesis of arRP in the Spanish population. Along with the detection of three recurrent mutations in Caucasian population, our hypothesis of *EYS* being the first prevalent gene in arRP has been reinforced in the present study.

©2010 Wiley-Liss, Inc.

KEY WORDS: *EYS*, Retinitis Pigmentosa, Spanish population, mutation, functional domain, recurrent mutation

INTRODUCTION

Retinitis pigmentosa (RP) is a heterogeneous group of inherited retinal dystrophies featured by the loss of photoreceptor cells and clinically characterized by pigmentary deposits at mid periphery of the retina that are visible on fundus examination. Patients present with night blindness as the initial symptom, which is followed by the constriction of the visual field and progressive loss of visual acuity, leading to complete blindness after several decades [Hamel, 2006]. Prevalence of nonsyndromic RP is approximately 1 in 4000. The condition may segregate as an autosomal dominant, autosomal recessive, or an X-linked recessive trait [Humphries et al., 1990]. All genes identified to date are believed to account for roughly 50% of all retinal dystrophy cases [Pomares et al., 2007]. The autosomal recessive form of RP is the commonest worldwide, accounting for approximately 39% of cases in Spain [Ayuso et al., 1995]. To date, 29 loci have been reported being responsible for arRP, of which 25 genes have been identified (<http://www.sph.uth.tm.edu/Retnet/>). However, all together the reported loci are responsible for only ~35-45% of the recessive RP cases, although none of them independently account for a substantial proportion of arRP (more than 10%) [Daiger et al., 2007; Hartong et al., 2006]. In contrast, the *RP25* locus, identified by our group in 1998 [Ruiz et al., 1998], was estimated to be linked to 27.7% of Spanish arRP families [Barragán et al., 2008]. Recently, we have identified a new gene encoding an ortholog of *Drosophila* spacemaker (*spam*) corresponding to *RP25* as a commonly mutated gene in arRP. The identification of six independent mutations, together with the presence of linked families from different ancestral origins, supports *EYS* (*Eyes Shut Homologue*, (MIM# 612424) as one of the first major genes reported for arRP [Abd El-Aziz et al., 2008]. Spanning over 2 Mb within the *RP25* locus (6p12.1-6q15), *EYS* is the largest gene identified to be expressed in the human eye so far, and the fifth largest gene overall in the human genome. The longest isoform of *EYS* encodes a protein of 3165 amino acids whose function remains to be elucidated. Considering the evolutionary data and the known function of the only characterised homologue, *EYS* is likely to have a role in the modelling of retinal architecture [Zelhof et al., 2006]. The identification of the gene for *RP25* reveals what might be the genetic basis for a significant proportion of arRP cases and thus paves the way for genetic counselling, prenatal detection, and treatment. However, further characterisation of the novel *EYS* protein as well as an extended mutation spectrum of *EYS*-related arRP would be valuable to undertake.

In the present study based on 94 families, we report the identification of 73 sequence variations in *EYS*, of which 28 are novel. Of these novel changes, 42.9% (12/28) are very likely pathogenic, 17.9% (5/28) are possibly pathogenic, whereas 39.3% (11/28) are SNPs. In addition, we have detected 3 pathogenic changes previously reported in other populations. The estimated prevalence of distinct *EYS* mutations in our group of patients is of 15.9%, confirming the significant involvement of *EYS* in the pathogenesis of the arRP in the Spanish population. Besides, we present a detailed bioinformatic characterisation of *EYS* and its homologues, which would aid in the determination of the pathogenic nature of newly identified variations in *EYS*.

MATERIALS AND METHODS

Subjects and Clinical Data: Our current cohort of study comprises 94 unrelated Spanish families affected by arRP, all derived from the Ophthalmology Service of different Hospitals throughout Spain. The participating families conform to the phenotypic and inheritance patterns of arRP. A group of matching control individuals was also recruited. Informed consent was obtained from all participants in the study, in accordance with the tenets of the Declaration of Helsinki (Edinburgh, 2000). Clinical diagnosis was based on visual acuity, fundus photography, computerized testing of central and peripheral visual fields and electroretinography (ERG) findings. Clinical features of RP include initial hemeralopy, restriction of visual field, gradual increased bone spicule pigmentation and decrease of visual acuity, attenuation of retinal vessels, and waxy disc pallor.

Bioinformatic characterisation of *EYS*: Firstly, ExPASy ProtParam tool was employed to determine the physical and chemical parameters of *EYS*. Secondly, InterProScan program was used to search for known domains and functional sites within *EYS*. For a further characterisation of *EYS*, we used Coils and Secpred to analyse the secondary structure. SignalP 3.0 was utilised to predict the presence and localisation of signal peptide cleavage sites. The characterisation of *EYS* homologues was performed in different steps. Firstly, Blast analyses of the human *EYS* cDNA and encoded protein were run to detect annotated homologous proteins. However, only human and *Drosophila* were found to be completely annotated in the databases. Therefore, we employed the BLAT tool at the UCSC Genome Bioinformatics Site to identify and map sequences with high identity to the target sequences. Human and *Drosophila* *EYS* protein were used as the basis of this search. An *in-silico* splicing site

4 <Barragán et al.>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

characterisation of the positive genomic region ensued to build the homologous cDNA which was then translated into protein. Also, we identified those sequencing gaps within the genomic region of *EYS* homologues in different species. Comparison of protein homologues were performed using *bl2seq* (NCBI) and *EMBOSS Pairwise Alignment Algorithms: Needle and Water* (EBI) alignments. The fully characterised proteins were aligned using *MUSCLE (MULTiple Sequence Comparison by Log-Expectation)* program at EBI.

PCR-based direct genomic sequencing of EYS: Peripheral blood samples were collected from all subjects for genomic DNA purification using an automated DNA extractor (*MagNA Pure LC Instrument*, Roche Diagnostics, Switzerland). Forty-eight pairs of primers were designed using the *Primer 3 Output program* (<http://frodo.wi.mit.edu/primer3/>) in order to screen the forty coding exons, the three non coding exons, the intronic flanking sequences and the regulatory factor binding sites of *EYS* (**Genbank Reference Sequence and Version FJ416331; GI: 212675237; Transcript Reference Sequence: NM_001142800.1**). PCR conditions and primer sequences employed are available upon request. The amplified products were subsequently purified using an enzymatic procedure, according to manufacture's recommendations (*EXOSAP-IT[®]*, USB Corporation) and sequenced with a ready reaction kit (*BigDye Terminator Cycle FS Ready Reaction Kit*; PE-Applied Biosystems, Foster City, CA). The fragments obtained were purified using fine columns (*Sephadex G-501*, Sigma-Aldrich Co.) and resolved on an automated sequencer (*3730 DNA Analyzer*, Applied Biosystems, USA). Finally, the data was analysed using *Lasergene DNASTAR[®]* software (DNASTAR, Inc). **Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence, according to journal guidelines (www.hgvs.org/mutnomen). The initiation codon is codon 1.** In order to evaluate the pathogenicity of the novel variants, we employed various softwares which analyse the potential role of a given variant on the function or structure of the encoded protein based on conservation and homology, physical properties of the amino acids, prediction of the protein disorder, or binding to transcription factor binding sites (TFBS) (*Conseq*: <http://conseq.tau.ac.il/>; *PolyPhen* (prediction of functional effect of human nSNPs): <http://coot.embl.de/PolyPhen/>; *SIFT* (Sorting Intolerant From Tolerant): <http://blocks.fhrc.org/sift/SIFT.html>; *Disopred*: <http://bioinf.cs.ucl.ac.uk/disopred/disopred.html> [Ramensky et al., 2002]. Besides, the tool *DiANNA* was employed for disulfide connectivity prediction when the variation affected a Cys residue, and *NetPhos2.1* (*NetPhos 2.0 Server*: <http://www.cbs.dtu.dk/services/NetPhos/>) together with *Diphos* (*Disorder-Enhanced Phosphorylation Sites Predictor*: <http://core.ist.temple.edu/pred/pred.html>), to predict the alteration of phosphorylation [Blom et al., 1999]. In addition, intronic variants were evaluated for affecting any regulatory process at the transcriptional or splicing levels (*TESS Transcription Element Search System*: <http://www.cbil.upenn.edu/cgi-bin/tess/tess>; <http://www.fruitfly.org/seqtools/splice.html>; *Splice SignalAnalysis*: <http://www.ebi.ac.uk/asd-srv/wb.cgi>; *Alternative Splicing DataBase*: <http://hazleton.lbl.gov/~teplitski/alt/>; http://www.fruitfly.org/cgi-bin/seq_tools/splice.pl; *Splicing Element Annotation*: <http://genes.mit.edu/acescan2/index.html>; *ESEfinder*: <http://rulai.cshl.edu/tools/ESE2/>; *Rescue-ESE*: <http://genes.mit.edu/burgelab/rescue-ese/>; *ACESCAN2 Web Server*: <http://genes.mit.edu/acescan2/index.html>; *NetGene2 Server* (<http://www.cbs.dtu.dk/services/NetGene2/>; http://www.ensembl.org/Homo_sapiens/generegulationview; <http://www.cisred.org/content/software>; <http://regrna.mbc.nctu.edu.tw/html/about.html>. [Yeo et al., 2004; Matlin et al., 2005; Wang & Marin 2006; Fairbrother et al., 2002; Brunak et al., 1991].

GeneChip 6.0 data analysis (see Affymetrix Genotyping Console 2.1 user manual for details): When sample requirements were fulfilled, patients underwent Copy Number Variations (CNVs) analysis with *Genechip 6.0 Affymetrix array*. The CEL intensity files were loaded into *Genotyping Console v2.1* (Affymetrix Inc.) for analysis. All samples passed the initial contrast QC metric (>0.4) that measures the ability of the intensity files to resolve SNPs into three genotyping clusters. Copy number data were generated by comparing intensities for both SNP and copy number probes *in silico* to the HapMap control provided by Affymetrix. The resulting log₂ ratios were then analysed using a Hidden Markov Model (HMM) to generate copy number calls for each probe. The quality of the log₂ data was assessed by the degree of variation, determined by the MAPD metric. MAPD is defined as the Median of the Absolute value of all Pairwise Differences between log₂ ratios for a given chip. High MAPD >0.4 (using the HapMap control) is considered to be the cut-off at which copy numbers can no longer be accurately called. None of the samples included in this study had a MAPD >0.4. Using the copy number calls provided by *Genotyping Console v2.1* as a guide, a more detailed analysis was performed by interrogation of the plots of log₂ ratios paying particular attention to CNV regions called by the HMM. A minimum log₂ ratio cut-off of +/-0.3 was used for autosomal CNVs. This excluded any false positive calls made by the HMM algorithm. Identified CNVs in the region of *EYS* were checked against the Database of Genomic Variants (DGV) in

Mutation spectrum of *EYS* in Spanish patients with autosomal recessive retinitis pigmentosa 5

Genotyping Console v2.1. CNVs were excluded from further analysis if they matched a known CNV exactly, or if occurred in a region that did not include any of the coding exons of *EYS* [McCarroll et al., 2006; de Smith et al., 2007; Sharp et al., 2005]. The allelic difference and loss of heterozygosity plots generated from the difference in allele intensity for each SNP were analysed to investigate parental consanguinity and as an aid to the interpretation of potential mosaicism. Genotyping was performed using the Birdseed v.2 algorithm. All samples had call rates >97.5%.

RESULTS AND DISCUSSION

Molecular diagnosis of RP is a challenging task given the important genetic heterogeneity of these groups of diseases. For most genes, many different mutations with similar consequences are known, yet other mutations in the same gene may cause different diseases. Particularly, for recessive RP, which is the most prevalent form of the disease, estimated to comprise from 50 to 60% of all RP cases, 29 loci have been described as pathogenic. In aggregate, the known mutations in arRP genes cause about ~35-45% of all cases of this form of the disease [Daiger et al., 2007; Hartong et al., 2006]. Some authors suggest that 50-60% of all arRP associated causal loci have already been identified. However, mutations in individual genes do not account for a significant proportion of arRP cases.

EYS is the largest gene identified to be expressed in the human eye so far, and appears to be a frequent cause of arRP. Thus, *EYS* encoded protein EYS emerges as a relevant player in arRP pathogenesis. In previous studies, a total of 8 mutations had been identified [Abd El-Aziz et al., 2008; Collin et al., 2008] by a combination of different screening methods such as direct genomic sequencing, MLPA or CGH arrays. The domain structure of EYS has been predicted from the characterised sequence of *EYS* [Abd El-Aziz et al., 2008; Collin et al., 2008] as having several EGF-like and Laminin G domains.

Here we report on the molecular screening of *EYS* in a Spanish cohort of patients with arRP. Besides, we are presenting the characterisation of EYS homologues in different species, and a detailed analysis of the EYS domains, with the identification of an interesting novel feature.

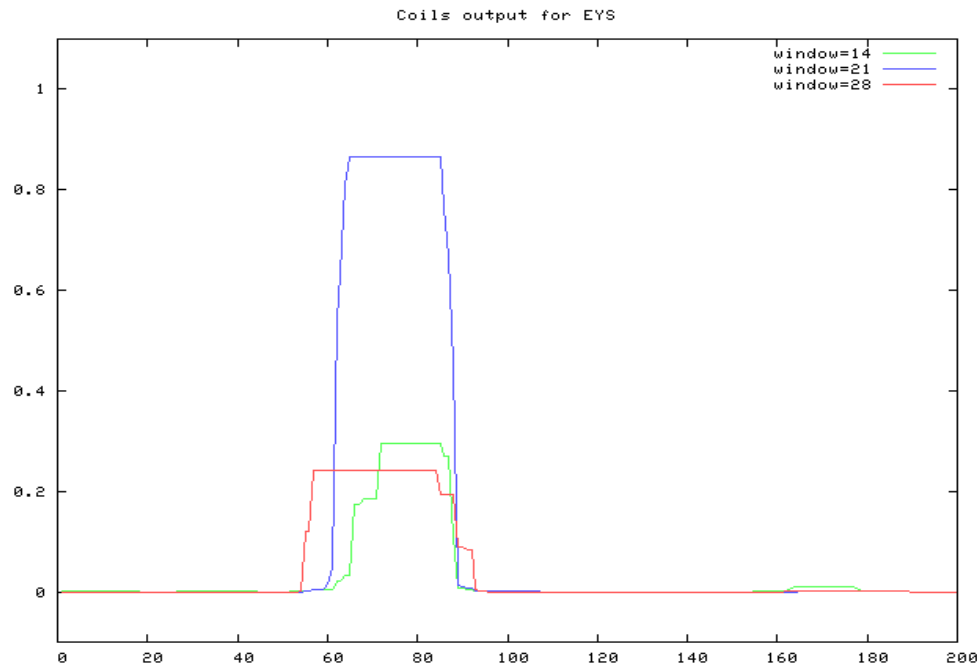
Protein domains structure

We present here the identification of a putative coiled-coil structure, which is an interesting novel feature, in the central portion of the protein coincident with a region of Alpha helix overrepresentation (Fig. 1, [Supplementary Figure S1](#)). The insight into the functional repercussion of both the already known signal peptide domain and the novel coiled-coil domain reported here supports a structural role for this new protein, which would be secreted and polymerize into a scaffolding that would contribute to the human retinal architecture. This is consistent with the function of *Eys* in *Drosophila*, where it is secreted by photoreceptor cells [Husain et al., 2006] and is essential for the formation of the matrix-filled intertrabecular space. The signal peptide and its cleavage site consensus sequence located in the N-terminal region of EYS (Fig. 1, Fig. 2) may confer a secretory nature to the protein or result in an intracellular or cytoplasmic location of the mature protein [Jarjanazi et al., 2008]. Remarkably, we have identified the consensus sequence for this feature in *EYS* homologues in a number of species such as orangutan, dog, horse, marmoset, monkey and chimpanzee (data available on request). Accordingly, *Drosophila* Spacemaker and other proteins which share several of EYS domains have been found to be secreted and to have a structural function, such as SCUBE [Yang et al., 2002] or CMG-2 [Bell et al., 2001]. The members of SCUBE gene family contain both a signal peptide domain and multiple EGF-like repeats. Interestingly, both EYS and SCUBE1 share homology with the same protein families, such as members of the fibrillin and Notch families among others. SCUBE1 and 2 are known to form oligomers and manifest a stable association with the cell surface in vascular endothelial cells [Yang et al., 2002]. CMG-2, containing a potential signal peptide, targets to the endoplasmic reticulum and shows affinity for the basement membrane matrix proteins, collagen type IV and laminin. Similarly, CMG-1, which encodes a protein with coiled-coil domains, was observed to target to an intracellular vesicular compartment and may play as well a structural role since it has been postulated that this gene may be implicated in the regulation of capillary formation in an *in-vitro* model of endothelial cell morphogenesis [Bell et al., 2001].

Furthermore, the fact that we have identified a putative coiled-coil domain within EYS reinforces the idea of EYS being a key player in the organisation of human retina. Coiled coils are important structural motifs involved in a

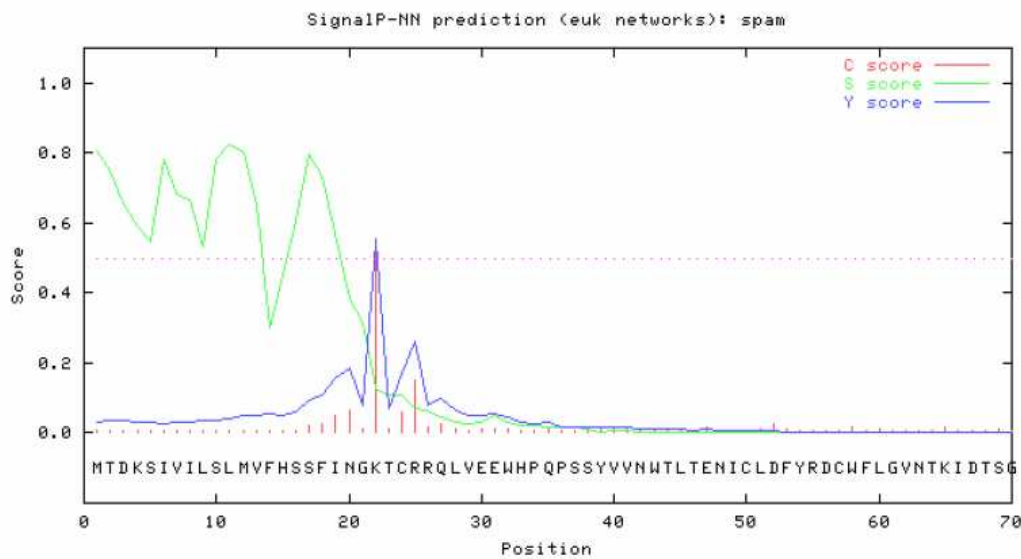
6 <Barragán et al.>

1
2
3 variety of important interactions [Mason & Arndt, 2004]. Dystrophin, which resembles *EYS* in several aspects such
4 as presenting a point mutation/deletion pattern of mutations in human disease and belonging to the group of five
5 longest genes in the human genome, is known to play structural roles among others. Dystrophin also contains
6 coiled-coil domains, which are responsible for the assembly of heterodimers of the so called Dystrophin
7 glycoprotein complex [Sadoulet et al., 1997; Böhm et al., 2008]. In this regard, it is important to mention that
8 *Drosophila* Spacemaker interacts with Prominin and the cell adhesion molecule Chaoptin to choreograph the
9 partitioning of rhabdomeres into an open system, critically affecting retinal morphogenesis [Zelhof et al., 2006].
10



36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

Figure 1. Coiled coil domain prediction in human EYS protein.



Mutation spectrum of *EYS* in Spanish patients with autosomal recessive retinitis pigmentosa 7

Annotation of new homologues

With the aim of evaluating evolutionary conservation, we have performed the bioinformatic characterisation of *EYS* homologues in several species. Apart from *Drosophila Eyes Shut*, which was the first *Eys* gene to be annotated, and human *EYS*, only fragmented information was available of other *EYS* homologues. Here we report in detail the structures of Zebrafish, Chicken, Platypus, Opossum, Horse, Dog, Marmoset and Orang-utan homologues and pairwise comparisons with human *EYS/EYS* (Table 1, Supplementary Figure S1). However, *EYS* seems to be absent in some species such as those insects with a close rhabdom system or in mammals in mouse, rat and guinea pig, which represent two of the three major rodent clades [Abd El-Aziz et al., 2008].

It is noteworthy that the previously reported and newly identified functional domains described here are conserved throughout evolution. Consistently, as shown in Table 1, the more distant a species is from human, the lower the percentages of identity and similarity. Concerning the signal peptide in *Drosophila*, some studies report that spacemaker secretion would be upon interaction of *Eys* with a receptor, which could promote its spreading from the stalk to the rhabdomere to fill the interrhabdomeral space (IRS) [Husain et al., 2006]. Thus, there would be no need of a signal peptide for secretion from the photoreceptor cells.

Table 1. *EYS* homologues characterisation.

Specie	Chromosomal interval (bp)	Genomic Length (bp)	Genomic Identity%	Protein Identity%	Protein Similarity%
Drosophila	chr2L:2323799-2357874	34,076	NA	23.9	37.9
Zebrafish	chr13:37105545-37227793	122,249	NA	44.0	60.3
Chicken	chr3:87827859-88369092	541,234	71.9	33.1	40.3
Platypus	chr1:33486985-34313969	1,756,349	71.7	31.8	36.6
Opossum	chr2:312522958-314990462	2,467,505	71.9	35.1	40.5
Horse	chr20:56733300-58137441	4,247,273	80.5	61.4	67.1
Dog	chr12:30185993-31709408	1,701,147	81.9	62.8	68.4
Marmoset	Several contigs*	973,997	91.4	88.8	92.2
Orangutan	chr6:63945057-65776065	1,831,009	97.3	96.8	97.5
Human	chr6:64488454-66262024	1,773,571	100.0	100.0	100.0

* Available on request

NA=Not Available

Pathogenic nature of the identified changes.

In this study, 12 novel very likely pathogenic changes have been identified in 10 families. Of these 10 families, 5 present mutations in both alleles, whereas the remaining 5 have mutations in just one allele. The clearly pathogenic variants consisted of 6 truncating mutations, 1 in frame deletion of 300 nucleotides leading to a protein truncation of 100 aminoacids, 1 splice site mutation and 4 missense changes. Out of the 28 novel variations, we have also identified 5 possible pathogenic changes in 5 separate families. In addition, we have detected 3 pathogenic variations previously published in ours and other populations [Abd El-Aziz et al., 2008; Collin et al., 2008] (Table 2, Table 3, Figs. 3 and 4). As mentioned in the Methods section, the sequence variants were designated in accordance with the Human Genome Variation Society recommendations (<http://www.hgvs.org/mutnomen/>). All the patients with mutations had received a defined clinical diagnosis of RP with a recessive mode of inheritance and were Spanish. The variations were regarded as pathogenic changes as long as they met the criteria of pathogenicity, i.e absence in 200 control individuals and the segregation with the disease phenotype within the family (Fig. 4). Particularly, missense mutations were considered pathogenic according to their effect on functional *EYS* domains that they target, their evolutionary conservation and/or to the fact that they are found together with a second variant, especially if this is truncating.

8 <Barragán et al.>

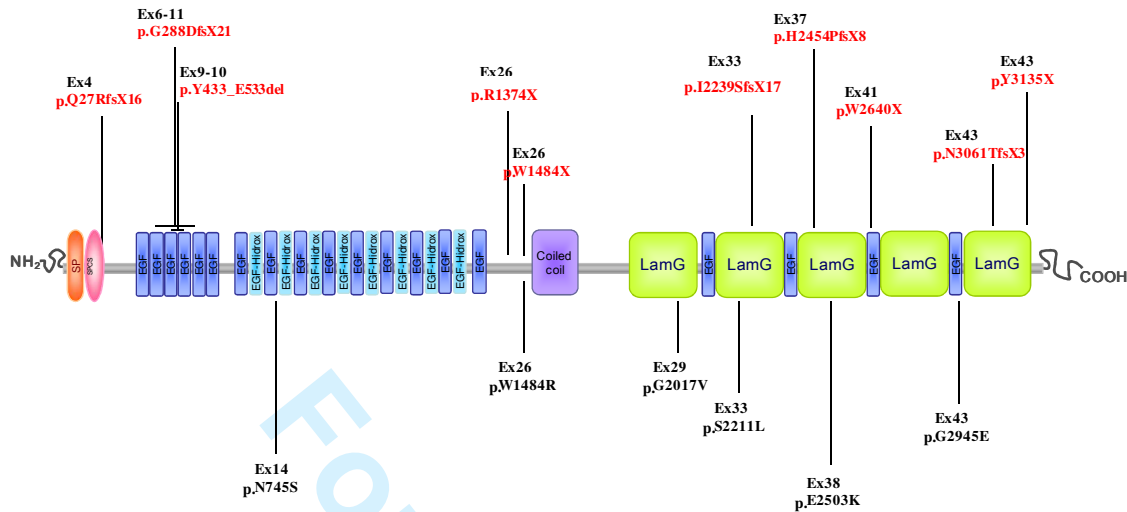


Figure 3. *EYS* mutation distribution along the domain structure of *EYS/EYS*. 5' UTR and splice site variations are not included in this depiction.

Table 2. Mutation spectrum of *EYS* gene in Spanish families.

Family ID	Nucleotide change	Predicted effect	Domains	Location in gene	Type of change	Reference of the variation
Families with novel very likely pathogenic changes and both alleles affected						
RP1052M	c.[1300-17039_1599+22208del]	p.Y433_E533del	EGF-like	Exons 9 and 10	Heterozygous	This study
	+ [9178_9181delA TAA]	p.N3061TfsX3	LamininG	Exon 43	Heterozygous	This study
RP1237M	c.[2234A>G]	p.N745S	EGF	Exon 14	Heterozygous	This study
	+ [7919G>A]	p.W2640X	EGF	Exon 41	Heterozygous	This study and Abd El-Aziz et al., 2008
RP84B	c.6050G>T	p.G2017V	LamininG	Exon 29	Homozygous	This study
RP136B	c.[6632C>T]	p.S2211L	LamininG	Exon 33	Heterozygous	This study
	+ [8834G>A]	+ p.G2945E	EGF	Exon 43	Heterozygous	This study
RP367B	c.[4120C>T]	p.R1374X	Close to EGF	Exon 26	Heterozygous	This study
	+ [6424+1G>T]	Splice mutation		Intron 31	Heterozygous	This study

Mutation spectrum of *EYS* in Spanish patients with autosomal recessive retinitis pigmentosa 9

Families with single novel very likely pathogenic changes						
RP358B	c.7361delA	p.H2454PfsX8	LamininG	Exon 37	Heterozygous	This study
RP60	c.78_79dupGC	p.Q27RfsX16	Signal peptide cleavage site	Exon 4	Heterozygous	This study
RP180M	c.862-10671_1766+10020del	p.G288DfsX21	EGF	Exons 6-11	Heterozygous	This study
RP33	c.4451G>A	p.W1484X	Close to coiled-coil	Exon 26	Heterozygous	This study
RP81	c.2234A>G	p.N745S	EGF	Exon 14	Heterozygous	This study
Families with novel possible pathogenic changes						
VRP8	c.4450T>C	p.W1484R	Close to coiled-coil	Exon 26	Heterozygous	This study
RP383	c.7507G>A	p.E2503K	LamininG	Exon 38	Heterozygous	This study
RP228B	c.-462G>C	-	-	Exon 1 5'UTR	Homozygous	This study
RP107B	c.-204G>A	-	-	Exon 3 5'UTR	Heterozygous	This study
RP509M	c.-349G>T	-	-	Exon 2 5'UTR	Heterozygous	This study
Families with very likely pathogenic changes reported in other populations						
RP155B	c.6714delT	p.I2239SfsX17	LamininG	Exon 33	Homozygous	This study Collin et al., 2008.
RP194B	c.9405T>A	p.Y3135X	LamininG	Exon 43	Homozygous	This study Collin et al., 2008.

Genbank Reference Sequence and Version FJ416331; GI: 212675237; Transcript Reference Sequence: NM_001142800.1

Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence, according to journal guidelines (www.hgvs.org/mutnomen). The initiation codon is codon 1.

10 <Barragán et al.>

Table 3. Novel variations identified by direct genomic sequencing of *EYS*

Gene_exon	Nucleotide Change	Predicted effect	Patients	Frequency - Controls frequency
<i>EYS</i> _1	c.-462G>C	-	2/188	0/400
<i>EYS</i> _2	c.-349G>T	-	1/188	0/400
<i>EYS</i> _3	c.-204G>A	-	2/188	0/400
<i>EYS</i> _4	c.748+209delA	-	1/188	-
	c.78_79dupGC	p.Q27RfsX16	1/188	0/400
<i>EYS</i> _6-11	c.862-10671_1766+10020del	p.G288DfsX21	2/12	0/400
<i>EYS</i> _9-10	c.1300-17039_1599+22208del	p.Y433_E533del	2/12	0/400
<i>EYS</i> _11	c.1766+61A>G	-	7/188	-
<i>EYS</i> _13	c.2024-14_-13insT	-	13/188	-
<i>EYS</i> _14	c.2234A>G	p.N745S	2/188	0/400
<i>EYS</i> _17+18	c.2733T>C	p.N911N	1/188	-
<i>EYS</i> _24	c.3684+61T>A	-	2/188	-
<i>EYS</i> _25	c.3877+18_22delAGATA	-	10/188	68/400
	c.4450T>C	p.W1484R	1/188	0/400
	c.4451G>A	p.W1484X	1/188	0/400
	c.4120C>T	p.R1374X	1/188	0/400
<i>EYS</i> _29	c.6050G>T	p.G2017V	1/188	0/400
	c.5959A>C	T1987P	1/188	2/400
<i>EYS</i> _30	c.6119T>A	p.V2040D	1/188	2/400
<i>EYS</i> _31	c.6424+1G>T	-	1/188	0/400
<i>EYS</i> _33	c.6632C>T	p.S2211L	1/188	0/400
<i>EYS</i> _37	c.7361delA	p.H2454PfsX8	1/188	0/400
<i>EYS</i> _38	c.7507G>A	p.E2503K	1/188	0/400
	c.7578+18C>T	-	2/188	6/400
<i>EYS</i> _39	c.7666A>T	p.S2556C	2/188	20/400
	c.7723+64T>A	-	3/188	34/400
<i>EYS</i> _43	c.9178_9181delATAA	p.N3061TfsX3	1/188	0/400
	c.8834G>A	p.G2945E	1/188	0/400

Genbank Reference Sequence and Version FJ416331; GI: 212675237; Transcript Reference Sequence: NM_001142800.1

Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence, according to journal guidelines (www.hgvs.org/mutnomen). The initiation codon is codon 1.

Families with novel very likely pathogenic changes and both alleles affected

As previously mentioned, 5 out of the 10 families bearing novel very likely pathogenic changes have both alleles affected. In 4 of them, they occurred as compound heterozygotes and hence it is sufficient to explain the recessive phenotype in their corresponding patients (Table 2, Fig. 4). In the case of family RP1052M, 1 frameshift deletion of 4 nucleotides involving the loss of the last residues of the Laminin G domain in the C-terminal region

Mutation spectrum of *EYS* in Spanish patients with autosomal recessive retinitis pigmentosa 11

of *EYS* together with an in frame deletion of 300 nucleotides leading to a protein truncation of 100 amino acid residues that disrupts one EGF-like domain were identified in the same patient. In family RP1237M, the truncating p.W2640X change and the missense p.N745S variation, also occurring in heterozygosity in RP81 family, are showing a recessive segregation through generations I and II. The same applies to p.S2211L and p.G2945E double heterozygotes in family RP136B (Fig. 4). These protein substitutions p.S2211L and p.G2945E are each transmitted by one of the progenitors. The altered residues are part of Laminin G and EGF domains, respectively. Turning to evolutionary conservation, both Serine and Glycine are present in these positions in *EYS* homologues characterised in this study. The new residue at amino acid position 2211 is of different polarity than Serine, and the substitution of Glycine to Glutamate in position 2945 introduces an acidic polarity in a previously hydrophobic position. Whereas the latter is not tolerated according to computational predictions, the former implies the loss of one phosphorylation site (NetPhos2.1). Finally, all affected members of family RP367B were compound heterozygous for a splice site and a nonsense mutation (c.6424+1G>T, p.R1374X) (Fig. 4). The c.6424+1G>T variation is predicted to lead to an abolishment of the donor splice site located at this position. It is known that splice sites mutations may disrupt protein function by diverse mechanisms such as exon skipping or the use of cryptic acceptor sites, presenting even multiple splice outcomes for a mutation in a given splice site [Takahara et al., 2002].

The fifth family of this group, RP84B, presents a homozygous coding variation, p.G2017V, which alters a residue that lies within a Laminin G domain of the protein and it is predicted not to be tolerated by SIFT. Family segregation of this variation shows a transmission pattern compatible with the recessive trait of the disease (Fig. 4).

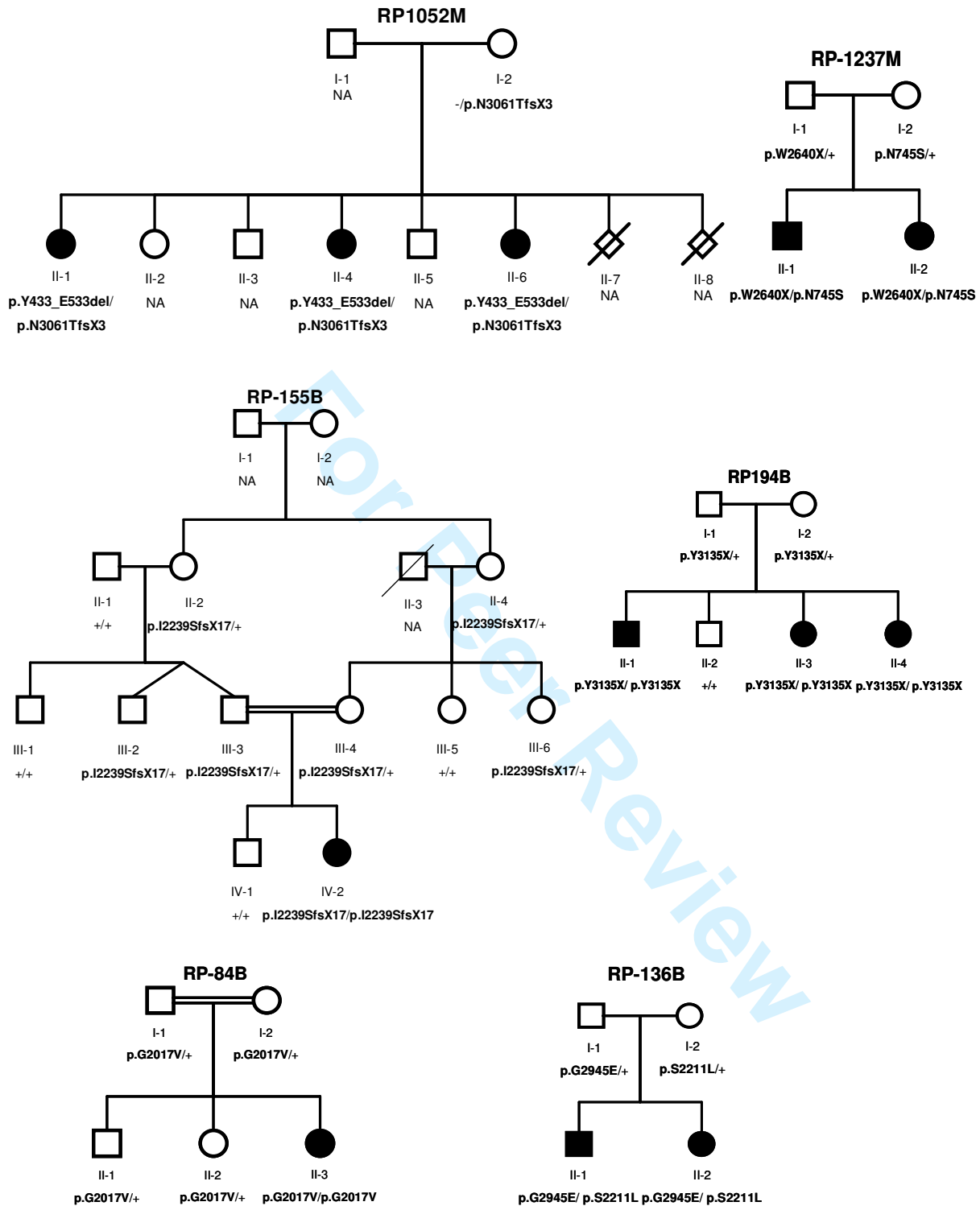
Families with single novel very likely pathogenic changes

The rest of the families comprising the group with novel very likely pathogenic changes present only single mutations. This is the case in families RP358B, RP60, RP180M, RP33 and RP81, with p.H2454PfsX8, p.Q27RfsX16, p.G288DfsX21, p.W1484X and p.N745S mutations, respectively (Table 2, Fig. 4).

Families with novel possible pathogenic changes

The group of families with possibly pathogenic variations is composed of 2 families bearing missense variations which did not appear in 200 control individuals and affect important domains of the protein (VRP8 and RP383, Table 2 and Fig. 4), and 3 additional families with variations in the 5' UTR segment of *EYS* (RP228B, RP107B and RP509M, Table 2 and Fig. 4). Of these, transversion c.-462G>C deserves special interest as it has been identified in a homozygous state in the proband of the consanguineous family RP228B. The hypothesized pathogenic potential for this change would ensue from its position in regulatory sequences important for protein translation [Scheper et al., 2007]. Moreover, family segregation is compatible with disease.

12 <Barragán et al.>



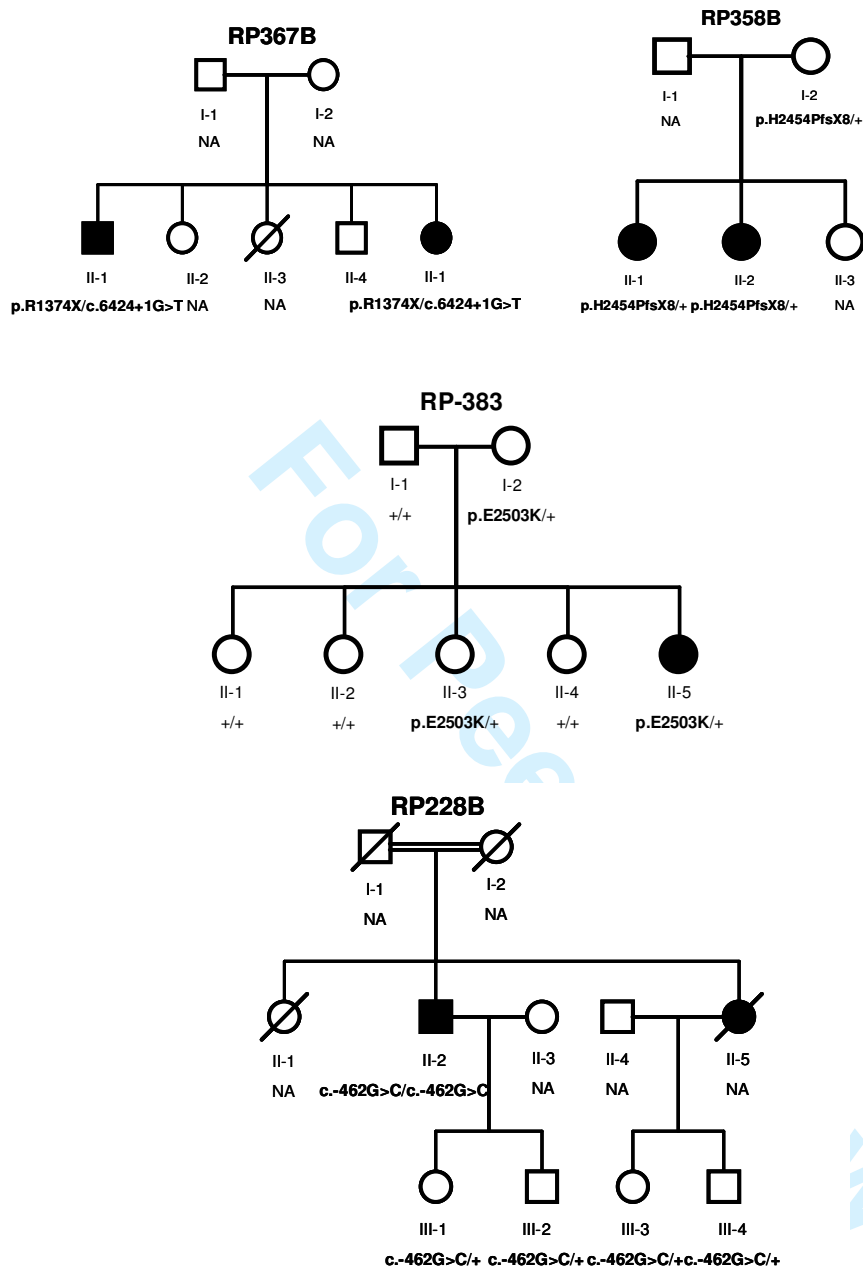
Mutation spectrum of *EYS* in Spanish patients with autosomal recessive retinitis pigmentosa 13

Figure 4. Family segregation of variations identified in the studied arRP families. Below the individuals, genotypes are presented for each change detected to segregate with the RP. For example, p.G2017V/p.G2017V represents homozygous mutants; p.G2017V/+ indicates heterozygous carriers, +/+ indicates individuals carrying two wild-type alleles, whereas p.R1374X/c.6424+1G>T represents individuals presenting both mutations as compound heterozygous. NA means non available DNA sample.

Considering only the most likely pathogenic variations (truncating, stop, and frameshift), a prevalence estimate of 9.6% of distinct *EYS* variants in the Spanish arRP population can be drawn. Additionally, if the very likely pathogenic changes are included in the prevalence estimation, the figure could rise up to 15.9%. In an additional

14 <Barragán et al.>

1
2
3 study performed in a separate population from United Kingdom, we have recently published eleven other novel
4 mutations within *EYS* with probable allele frequency of 11% [Abd El-Aziz et al., 2010].

5 Furthermore, it is worth mentioning that many of the domains that feature *EYS* protein are targeted by
6 pathogenic variations. Yet, different pathogenic mechanisms are postulated depending on the nature of the
7 variation. Probably, in the case of variations generating a premature stop codon, most of these altered mRNA
8 transcripts will be lost through Nonsense Mediated Decay (NMD) [Frischmeyer et al., 2002].

9
10 With regard to the coding mutations, there is not a clear clustering of both the previously reported mutations
11 and those identified in this study to indicate a common target in the primary sequence of the protein. Although the
12 distribution of these mutations implicated in arRP reveals 15 affected exons, a trend for alterations in residues of
13 the C-terminal region containing alternating EGF-Laminin G domains of the protein is observed. Particularly, 4
14 out of the 6 missense variations are found in the Laminin G and EGF domains of the second half of the protein
15 (Fig. 3). This is consistent with the hypothesis presented by other groups that the C-terminal region would be
16 crucial for the function of the protein [Collin et al., 2008], and with the homology results presented in the current
17 study, which outline the C-terminal region as one of the highly conserved intervals. Accordingly, the homology
18 analysis presented here also reveals the high degree of evolutionary conservation of all the domains presenting
19 altered residues by the mutations identified in arRP patients (Supplementary Figure S1).

20 Interestingly, mutations p.I2239SfsX17, p.Y3135X and p.W2640X had been previously reported as disease
21 causing in 2 Dutch and 1 Spanish family respectively [Abd El-Aziz et al., 2008; Collin et al., 2008]. We have
22 performed an extensive haplotype analysis and based on all available marker data we get different genotype
23 information associated with the mutant alleles. Therefore, it is very likely that they are recurring mutations.
24 Identifying recurrent mutations in Caucasian and especially specific populations such as the Spanish one provides
25 an essential source for the molecular and clinical diagnosis of such a heterogeneous disease. Furthermore, this fact
26 reinforces our hypothesis that *EYS* is the first prevalent gene in arRP [Abd El-Aziz et al., 2008].

27 The identification of 6 missense variations within the disease related changes in *EYS* differs from the mainly
28 deletion/truncation mutations reported in previous studies ([Abd El-Aziz et al., 2008; Collin et al., 2008].
29 Nonetheless, the identification of missense mutations in arRP patients have already been reported [Sun et al., 1997;
30 Molday et al., 2000], suggesting a refinement of the model based on the observation that some missense alleles
31 might behave as true null allele at the functional level and may be responsible for severe impairment of protein
32 function. Furthermore, majority of missense mutations reported here are located in functional conserved *EYS*
33 domains and are more prevalent in the domains of the second half of the protein, thereby indicating a pathologic
34 role for such variants (Fig. 3, Supplementary Figure S1).

35 It is noteworthy that a significant proportion of families in our study with an *EYS* mutation had only one
36 identified mutation. Interestingly, structural variations have been found in 2 out of the 6 assessed patients. Besides,
37 in the previous paper we had identified 2 large deletions which are not detectable by direct genomic sequencing
38 [Abd El-Aziz et al., 2008] and this may explain why in a proportion of patients the second mutation remained
39 unidentified as reported in the current work. Additional experiments consisting of Copy Number Variations
40 (CNVs) or MLPA (Multiplex Ligation-dependent Probe Amplification) analysis would be useful to rule out long
41 heterozygous deletions.

42 The identification of distinct mutations in *EYS* reveals a probable mutation frequency of 15.9% in the Spanish
43 arRP population. Along with the detection of three recurrent mutations in Caucasian population, our hypothesis of
44 *EYS* being the first prevalent gene in arRP has been reinforced in the present study.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Mutation spectrum of *EYS* in Spanish patients with autosomal recessive retinitis pigmentosa 15

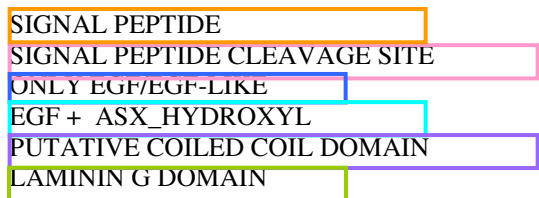
REFERENCES

- 1
2
3
4
5
6 - Abd El-Aziz MM, Barragan I, O'Driscoll CA, Goodstadt L, Prigmore E, Borrego S, Mena M, Pieras JI, El-Ashry MF, Safieh
7 LA, Shah A, Cheetham ME, Carter NP, Chakarova C, Ponting CP, Bhattacharya SS, Antinolo G. 2008. *EYS*, encoding an
8 ortholog of *Drosophila* spacemaker, is mutated in autosomal recessive retinitis pigmentosa. *Nat Genet* 40:1285-1287.
- 9 - Abd El-Aziz MM, O'Driscoll CA, Kaye RS, Barragan I, El-Ashry MF, Borrego S, Antiñolo G, Pang CP, Webster A,
10 Bhattacharya SS. 2010. Identification of Novel Mutations in the ortholog of *Drosophila* eyes shut Gene (*EYS*) Causing
11 Autosomal Recessive Retinitis Pigmentosa. *Invest Ophthalmol Vis Sci*, *in press*.
- 12 - Ayuso C, Garcia-Sandoval B, Najera C, Valverde D, Carballo M, Antinolo G. 1995. Retinitis pigmentosa in Spain. The
13 Spanish multicentric and multidisciplinary group for research into retinitis pigmentosa. *Clin Genet* 48:120-122.
- 14 - Barragán I, Abd El-Aziz MM, Borrego S, El-Ashry MF, O'Driscoll C, Bhattacharya SS, Antiñolo G. Linkage validation of
15 RP25 Using the 10K genechip array and further refinement of the locus by new linked families. 2008. *Ann Hum Genet*
16 72:454-462.
- 17
18 - Bell SE, Mavila A, Salazar R, Bayless KJ, Kanagala S, Maxwell SA, Davis GE. 2001. Differential gene expression during
19 capillary morphogenesis in 3D collagen matrices: regulated expression of genes involved in basement membrane matrix
20 assembly, cell cycle progression, cellular differentiation and G-protein signaling. *J Cell Sci* 114:2755-2773.
- 21
22 - Blom N, Gammeltoft S, Brunak S. 1999. Sequence- and structure-based prediction of eukaryotic protein phosphorylation
23 sites. *J Mol Biol* 294:1351-1362.
- 24
25 - Böhm S, Jin H, Hughes SM, Roberts RG, Hinitz Y. 2008. Dystrobrevin and dystrophin family gene expression in zebrafish.
26 *Gene Expr Patterns* 8:71-78.
- 27
28 - Brunak S, Engelbrecht J, Knudsen S. 1991. Prediction of Human mRNA Donor and Acceptor Sites from the DNA Sequence.
29 *J Mol Biol* 220:49-65.
- 30
31 - Collin RW, Littink KW, Klevering BJ, van den Born LI, Koenekoop RK, Zonneveld MN, Blokland EA, Strom TM, Hoyng
32 CB, den Hollander AI, Cremers FP. 2008. Identification of a 2 Mb human ortholog of *Drosophila* eyes shut/spacemaker
33 that is mutated in patients with retinitis pigmentosa. *Am J Hum Genet* 83:594-603.
- 34
35 - Daiger SP, Bowne SJ, Sullivan LS. 2007. Perspective on Genes and Mutations Causing Retinitis Pigmentosa. *Arch*
36 *Ophthalmol* 125:151-158.
- 37
38 - Fairbrother WG, Yeh RF, Sharp PA, Burge CB. 2002. Predictive identification of exonic splicing enhancers in human genes.
39 *Science* 297:1007-1013.
- 40
41 - Frischmeyer PA, van Hoof A, O'Donnell K, Guerrero AL, Parker R, Dietz HC. 2002. An mRNA surveillance mechanism
42 that eliminates transcripts lacking termination codons. *Science* 22:2258-2261.
- 43
44 - Hamel C. Retinitis pigmentosa. 2006. *Orphanet J Rare Dis* 1:40.
- 45
46 - Hartong DT, Berson EL, Dryja TP. 2006. Retinitis pigmentosa. *Lancet* 368:1795-1809.
- 47
48 - Humphries P, Farrar GJ, Kenna P, McWilliam P. 1990. Retinitis pigmentosa: genetic mapping in X-linked and autosomal
49 forms of the disease. *Clin Genet* 38:1-13.
- 50
51 - Husain N, Pellikka M, Hong H, Klimentova T, Choe KM, Clandinin TR, Tepass U. 2006. The agrin/perlecan-related protein
52 eyes shut is essential for epithelial lumen formation in the *Drosophila* retina. *Dev Cell* 11:483-93.
- 53
54 - Jarjanazi H, Savas S, Pabalan N, Dennis JW, Ozcelik H. 2008. Biological implications of SNPs in signal peptide domains of
55 human proteins. *Proteins* 70:394-403.
- 56
57 - Mason JM, Arndt KM. 2004. Coiled coil domains: stability, specificity, and biological implications. *ChemBiochem* 5:170-
58 176.
- 59
60 - Matlin AJ, Clark F, Smith CW. 2005. Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol*
6:386-398.

16 <Barragán et al.>

- 1
2
3 - Molday LL, Rabin AR, Molday RS. 2000. ABCR expression in foveal cone photoreceptors and its role in Stargardt macular
4 dystrophy. *Am J Ophthalmol* 130:689.
- 5
6 - Pomares E, Marfany G, Brión MJ, Carracedo A, Gonzalez-Duarte R. 2007. Novel high-throughput SNP genotyping
7 cosegregation analysis for genetic diagnosis of autosomal recessive retinitis pigmentosa and Leber congenital amaurosis.
8 *Hum Mutat* 28:511–516.
- 9
10 - Ramensky V, Bork P, Sunyaev S. 2002. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res* 30:3894–3900.
- 11
12 - Ruiz A, Borrego S, Marcos I, Antiñolo G. 1998. A major locus for autosomal recessive retinitis pigmentosa on 6q, determined
13 by homozygosity mapping of chromosomal regions that contain gamma-aminobutyric acid-receptor clusters. *Am J Hum
14 Genet* 62:1452–1459.
- 15
16 - Sadoulet-Puccio HM, Rajala M, Kunkel LM. 1997. Dystrobrevin and dystrophin: an interaction through coiled-coil motifs.
17 *Proc Natl Acad Sci U S A* 94:12413–12418.
- 18
19 - Scheper GC, van der Knaap MS, Proud CG. 2007. Translation matters: protein synthesis defects in inherited disease. *Nat Rev
20 Genet* 8:711–723.
- 21
22 - Sun H, Nathans J. 1997. Stargardt's ABCR is localized to the disc membrane of retinal rod outer segments. *Nat Genet* 17:15–
23 16.
- 24
25 - Takahara K, Schwarze U, Imamura Y, Hoffman GG, Toriello H, Smith LT, Byers PH, Greenspan DS. 2002. Order of intron
26 removal influences multiple splice outcomes, including a two-exon skip, in a COL5A1 acceptor-site mutation that results in
27 abnormal pro-alpha1(V) N-propeptides and Ehlers-Danlos syndrome type I. *Am J Hum Genet* 71:451–465.
- 28
29 - Wang M, Marin A. 2006. Characterization and prediction of alternative splice sites. *Gene* 366:219–227.
- 30
31 - Yang RB, Ng CK, Wasserman SM, Colman SD, Shenoy S, Mehraban F, Komuves LG, Tomlinson JE, Topper JN. 2002.
32 Identification of a novel family of cell-surface proteins expressed in human vascular endothelium. *J Biol Chem* 277:46364–
33 46373.
- 34
35 - Yeo G, Hoon S, Venkatesh B, Burge CB. 2004. Variation in sequence and organization of splicing regulatory elements in
36 vertebrate genes. *Proc Natl Acad Sci U S A* 101:15700–15705.
- 37
38 - Zelfhof AC, Hardy RW, Becker A, Zuker CS. 2006. *Nature* 443:696–699.
- 39
40 - McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, Dallaire S, Gabriel SB, Lee C, Daly MJ, Altshuler
41 DM, International HapMap Consortium. Common deletion polymorphisms in the human genome. 2006. *Nat Genet* 38:86–
42 92.
- 43
44 - Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Seagraves R, Oseroff
45 VV, Albertson DG, Pinkel D, Eichler EE. 2005. Segmental duplications and copy-number variation in the human genome.
46 *Am J Hum Genet* 77:78–88.
- 47
48 - de Smith AJ, Tsalenko A, Sampas N, Scheffer A, Yamada NA, Tsang P, Ben-Dor A, Yakhini Z, Ellis RJ, Bruhn L, Laderman
49 S, Froguel P, Blakemore AI. 2007. Array CGH analysis of copy number variation identifies 1284 new genes variant in
50 healthy white males: implications for association studies of complex diseases. *Hum Mol Genet* 16:2783–2794.
- 51
52
53
54
55
56
57
58
59
60

Mutation spectrum of EYS in Spanish patients with autosomal recessive retinitis pigmentosa 17



p.Q27RfsX16



13	Orangutan	MTDKSIIILSLMVFHSSSFINGKT-CRRQSVVEEWHQPSSYVNVNWTLTENICLDFYRDCWF	59
14	Human	MTDKSIVILSLMVFHSSSFINGKT-CRRQLVEEWHQPSSYVNVNWTLTENICLDFYRDCWF	59
15	Marmoset	MTDKSIIILSLVVFHSSSFINGKT-CRRQLVEEWHQPSSYVNVNWTLTENICMDFYRDCWI	59
16	Horse	MANKSVMLLSLVVLFHSSVINGKMTCKWQLVEEWHQTPSSYVNVNWTLTENICTDFYGDWCWF	60
17	Dog	MTNKSIMLLSLVVLHSSVTFINGKTTCKWQLVEEWHQTPSSYVNVNWTLTENICTDFYGDWCWF	60
18	Opposum	-----	
19	Platypus	-----	
20	Chicken	-----	
21	Zebrafish	-----	
22	Drosophila	-----MAESPQIRRDMGRLCAT	17
25	Orangutan	LGVNTKIDTSGNQAVPQICPLQIQLDILVISSEPSLQFPEINLMNVSETSFVGCVQNTT	119
26	Human	LGVNTKIDTSGNQAVPQICPLQIQLDILVISSEPSLQFPEINLMNVSETSFVGCVQNTT	119
27	Marmoset	LGVNTKIDTSGNQVVPQICPLQIQLDILVISSEPSLQFPEINLMNVSETSFIDCVQNTT	119
28	Horse	GDVNTKMSTLGNQVVPQICPLQIQLDILVISSEPSLQSPENLMNVSEASFIDCMQAT	120
29	Dog	GDVNTKMSTLGNQVVPQICPLQIQLDILVISSEPSLQSPENLMNVSEASFIDCMQAT	120
30	Opposum	-----PQICPLQIQYGDKLIISSELSHFGLMNLNMVSKDTFINCLQND	45
31	Platypus	-----	
32	Chicken	-----	
33	Zebrafish	-----	
34	Drosophila	WPSKDSBEDGAGTALRAATPLTANGATTGLSVTLAPKDMQRNHLKMPATIEKPTITAT	77
37	Orangutan	TEDQLLFGCRLKGMHTVNSKWLSVGTHYFITVMASQPSPCPLGLRLNVTVKQQFCQESLS	179
38	Human	TEDQLLFGCRLKGMHTVNSKWLSVGTHYFITVMASQPSPCPLGLRLNVTVKQQFCQESLS	179
39	Marmoset	TEDQLLFGCRLKGMHTVNSKWLSVGTHYFITVMASQPSLCSLGLRLNVTVKQQFCQESLS	179
40	Horse	TEDQLLFGCRLKGMHTVNSQWLSVGTHYFITVMANGPSLCQLGLRLNVTVKQQFCQEPLH	180
41	Dog	TEDQLLFGCRLKGMHTVNSQWLSVGTHYFITVMANGPSLCQLGLRLNVTVKQQFCQEPLH	180
42	Opposum	AQDELLEFGCRLKGTHEVNPQWLSIGTHYLLTVVERSPFLCKLGLRLNVTVKQQFCQASQN	105
43	Platypus	-----	
44	Chicken	-----	
45	Zebrafish	-----	
46	Drosophila	IASSSSTSTSTTRKSVTATRSKLNPNILLPTLRILARGLLLPALILAILVGSQAGFAC	137
49	Orangutan	SEFCSGHGKCLSEAVWSKIYSCHCQPPFSGKYCQ-ELDACSFKPKKNGGSCINKRENWDGQ	238
50	Human	SEFCSGHGKCLSEAVWSKIYSCHCQPPFSGKYCQ-ELDACSFKPKKNGGSCINKRENWDEQ	238
51	Marmoset	SEICSGHGKCLSEVWSKIYSCHCQPPFSGKYCQ-EHDACSSKPKKNGGSCINKRGKWDKQ	238
52	Horse	SEFCSGHGKCLSEVWSKVYSCHCQPPFSGKYCQ-EVDACSHKPCENNGGSCINKRGKWNKQ	239
53	Dog	SEFCSGHGKCLSEVWSKVYSCHCQPPFSGKYCQ-EVDACSHKPCENNGGSCINKRGKWNKQ	239
54	Opposum	APFCSGHGRCLSEVWNKTYVCHCEPPYSGEFCQ-EVDECFQNPCHNNGGICINKRE----	159
55	Platypus	-----	
56	Chicken	-----FKGTRCETDIDECASSPKNGATCIDQPGN----	30
57	Zebrafish	-----	
58	Drosophila	LSNPCVFGVCIDGLN-SSYSYCYCIDGYTGIQCQTNWDECWSSPCQNGGTCVDGVAYYN--	194

18 <Barragán et al.>

1			
2			
3			
4	Orangutan	GYECVCHPPFTGKNCSEIIGQCQPHVCFHGNCNSITNSFICECDEQFSGPFCEMSTKPC	298
5	Human	AYECVCHPPFTGKNCSEIIGQCQPHVCFHGNCNSITNSFICECDEQFSGPFCEVSAKPC	298
6	Marmoset	GYECVCHPPFTGKNCSEIIGQCQPHICFHGNCNSITNSFICECDERFKGPFCVSTKPC	298
7	Horse	GYECICHPPFTGKNCSEIIDQCQPYVCFHGNYSNITNSFICECDEPFSG-----	289
8	Dog	GYECICHPPFTGKNCSEIIDQCQPYVCFHGNYSNITNSFICECDEPFSG-----	289
9	Opposum	-----FTGRNCEEVIDYCR-----	173
10	Platypus	-----	
11	Chicken	-----YFCQCMAPFKFVN-----	43
12	Zebrafish	-----	
13	Drosophila	---CTCPEGFSGSNCEENVDECMSNPCQNGGLCRDRTINGYICTCQPGYLG-----	241
14			
15			
16	Orangutan	VSLLCWKRIGICPNSSSAYTYECPKGSSSQNGETDVSECSLVP CQNGTDCIQISNDVMCIC	358
17	Human	VSLLFWKRIGICPNSSSAYTYECPKGSSSQNGETDVSEFSLVPCQNGTDCIKISNDVMCIC	358
18	Marmoset	VSLLCWKRIGICPNSSSAYTCECPKRSPSQNDEIDVNDCSLIPCQNGTDCIKISNDVMCIC	358
19	Horse	-----	
20	Dog	-----	
21	Opposum	-----	
22	Platypus	-----	
23	Chicken	-----	
24	Zebrafish	-----	
25	Drosophila	-----	
26			
27			
28	Orangutan	SPIFTDLLCKSIQTSCEFSFLRNNATCKKWEKDYHCSCISGFTEKNCEKAIDHCRLLSIN	418
29	Human	SPIFTDLLCKSIQTSCEFSFLRNNATCKKCEKDYPCSCISGFTEKNCEKAIDHCKLLSIN	418
30	Marmoset	SPIFTDILCRSIQTSYESFPLKNTTICKKCEKEYHC-----KNCEKVIDHCRLLSIN	410
31	Horse	-----KNCEKVIDHCRLLCVN	305
32	Dog	-----R-CSYYLGRIDREFCIL	304
33	Opposum	-----LLSIN	178
34	Platypus	-----	
35	Chicken	-----	
36	Zebrafish	-----	
37	Drosophila	-----SHCELDVAV	250
38			
39	Orangutan	CLNEEWCFNIIGRF-YVCIPGCPKNPCWFLKNVYLIHQHL CYCGVTFHGICQDKGPAQFE	477
40	Human	CLNEEWCFNIIGRFKYVCIPGCTKNPCWFLKNVYLIHQHL CYCGVTFHGICQDKGPAQFE	478
41	Marmoset	CLNEEWCFNIIGRF-YVCIAGCKINPCWLKKNVYSIHQHPCYYGVTCRGICQDKGPAQFQ	469
42	Horse	CQNEGWC FN IIGRF-----	319
43	Dog	DVED-----	308
44	Opposum	CLNEGLCFNIIGGF-----	192
45	Platypus	-----	
46	Chicken	-----	
47	Zebrafish	-----	
48	Drosophila	CETGTGAR-----	258
49			
50			
51	Orangutan	YVWQLGFAGSEGEK CQGV IDAYFFFLAANCTEDAICVNDPEDYNSSCRFPREGTKEICANG	537
52	Human	YVWQLGFAGSEGEK CQGV IDAYFFFLAANCTEDATYVNDPEDNNSCWFPHEGTKEICANG	538
53	Marmoset	CVWQLGFTGSEGEK CQRV IDVYFFFLAANCTEDAIYVNMPEDINNSCWFPCEGTKEICANG	529
54	Horse	-----	
55	Dog	-----	
56	Opposum	-----	
57	Platypus	-----	
58			
59			
60			

Mutation spectrum of *EYS* in Spanish patients with autosomal recessive retinitis pigmentosa 19

1			
2			
3	Chicken	-----	
4	Zebrafish	-----	
5	Drosophila	-----	
6			
7			
8	Orangutan	CSCLSEEDSQEYLYLCFLRWAGNMYLENTTDDQENECQHEAICKDEINRPRCSCSLSYIG	597
9	Human	CSCLSEEDSQEYRYLCFLRWAGNMYLENTTDDQENECQHEAVCKDEINRPRCSCSLSYIG	598
10	Marmoset	CSSFVKEDSQEYQCLCFLRWADKLYLENTTDDQENEYQHEAICKDEINRPRCSCSLSCSG	589
11	Horse	-----	
12	Dog	-----	
13	Opposum	-----	
14	Platypus	-----	
15	Chicken	-----	
16	Zebrafish	-----	
17	Drosophila	-----	
18			
19			
20	Orangutan	RLCVVNVDYCLGNQSI SVHGLCLALSHKNCISLQRYERNICEIDTEDCKSVCKNGTTS	657
21	Human	RLCVVNVDYCLGNHSISVHGLCLALSHNCNCSGLQRYERNICEIDTEDCKSASCKNGTTS	658
22	Marmoset	RLCVVNVDYCLGNQSI SVHGLCLALSHKNCISLQRYERNICEIDTEDCKSVCKNGTTS	649
23	Horse	-----RVCLGNESISVHGLCLVRLHNCNCSLQRYERNICEIETEDCKSVCKNGTTS	372
24	Dog	-----CLGNESISVHGLCLVRLHNCNCSLQRYERNICEIETEDCKSVCKNGTTS	359
25	Opposum	-----	
26	Platypus	-----	
27	Chicken	-----	
28	Zebrafish	-----	
29	Drosophila	-----	
30			
31	Orangutan	IHLRGYFFCKCVPGFK---GTQREIDIDECASHPCKNGAATCIDQPGNYFCQVPPFK--V	712
32	Human	THLRGYFFRKCVPGFK---GTQCEIDIDECASHPCKNGAATCIDQPGNYFCQVPPFK--V	713
33	Marmoset	IHLRGYVFCCKCVPGFKGFEGTRCKIDVDECASHPCKNGAATCIDQPGNYLCQVPPFKFPV	709
34	Horse	IHSSGYFFCKCVPGFK---GTQCEIDIDECASHPCKNGAATCIDQPGNYFCQVPPFK--V	392
35	Dog	IHSSGYFFCKCVPGFKGT---RSETDTERASHPCKNGAATCADQPGNYFCQVAPFS--V	414
36	Opposum	-----	193
37	Platypus	-----	
38	Chicken	-----	
39	Zebrafish	-----	
40	Drosophila	-----CQFGGECIEGP	269
41			
42			
43			
44			
45	Orangutan	VDGFSCCLGNPGYVGRICEQDIDDCILNACEHNSICKDLHLSYQCVCCLSGWEGNFCEQESN	772
46	Human	VDGFSCCLNPGYVGRICEQDIDDCILNACEHNSICKDLHLSYQCVCCLSDWEGNFCEQESN	773
47	Marmoset	VDGFSCCLNPGYVGRICEQDIDDCILNACEHNSICKELHLSYQCVCCLSGWEGNFCEQESN	769
48	Horse	AVGFSCCLSAACVGLRICEQDIDDCILNACEHNSICKDLHLSYQCVCCLSGWEGNFCEQESN	432
49	Dog	AVGFSCCLSAACVGLRICEQDIDDCILNACEHNSICKDLHLSYQCVCCLSGWEGNFCEQESN	454
50	Opposum	VDGFSCCLNPGYAGLRICEQDIDDCIHTACENNSICKDLHLSYQCVCCLSGWEGNFCEQESN	234
51	Platypus	-----	
52	Chicken	--GFYCLNPGYAGLRICEQDIDDCIHTACENNSICKDLHLSYQCVCCLSGWEGNFCEQESN	82
53	Zebrafish	-----MCEDTSENGCLFIISVR	17
54	Drosophila	GLEFTCDCPAGWHGRICQEEINECASSPCQNGGVCVDKLAAY	311
55			
56			
57			
58			
59			
60			

p.N745S



20 <Barragán et al.>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

ECKMNPCKNNSTCIDLYKSYRCECTSGWTGQNCSEEINECDSDPCMNGGLCHESTIPGQF	832
ECKMNPCKNNSTCIDLYKSYRCECTSGWTGQNCSEEINECDSDPCMNGGLCHESTIPGQF	833
ECEMNPCKNNSTCIDLYKSYRCECTSGWTGQNCSEEINECDSDPCMNGGLCHESTIPGQF	829
-----CECTSGWTGQNCSEEINECDSDPCLNGALCHESTIPGQF	471
-----CECTSGWTGQNCSEEINECDSDPCLNGALCHESTIPGQF	493
-----CGCTPGWTGQNCSEEINECDASPCKNGAICQESTVPGQF	273
-----RCICKPEWTGQNCSEELNKCDSNPCMGATCFNSAVPGKV	40
-----CLCTAGWTGPDCESEDINECDSEPCLNATCYESVKQGF	121
-----SAKSPASQSQWHHEPPTSSTHSPTL---LITVETPPGEW	53
-----ACACPMGYTGINCEEILICADNPQNNALCLMEEG--VP	349
	* * * * *

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

VCLCPPLYTGRFCHQRYNPCDLLHNPGRNNSTC--LALVDGNQHICICREEFEGKNCIDV	890
VCLCPPLYTGRFCHQRYNCDLLHNPGRNNSTC--LALVDANQHICICREEFEGKNCIDV	891
VCLCPPLYTGRFCHQRYNPCCELLNPNCRNNSTC--LALVDGNQHICICR-EFEGKHCEIDT	886
VCLCPFFFYGKFC-----LNDPCRNNATC--LTLVDGQRYCVCR-----	508
VCLCPFFFYGKFCQYRSSCDPLNDPCRNNATC--LTLVDGQRYCVCR-----	539
VCLCPFFFYGFCCQIYNPCDMAYNLCINNSTC--LITVDGNSNCVCR-----	319
ECLCPPLYTGAFCEHHYNSCETFHKPCINNSTC--LTRADGNTEFICF-----	86
VCICPPFYTGDFCHQRFSPCELPYNPCINNSTC--LAQADGNPMCICK-----	167
VRVLSPTQPAPVCPQG-----ICLNGGTCPVSLPSGASSFFCD-----	92
TCYCVDPDYHGEKCEFQYDECQLG-PRCMNGGVC-----IDGVDTFSCS-----	391
	* * * * *

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

KECLFLSCQDYGDCEDMVNNFRICIRPGFSGSLCEIEINECSSEPCKNNGTQVDLTNRFF	950
KDCLFLSCQDYGDCEDMVNNFRICIRPGFSGSLCEIEINECSSEPCKNNGTQVDLTNRFF	951
NECLFLPCQGYGDCEAIVNNFRICIRPGFSGSLCEIEINECSSEPCKNNGTQVDLTNRFF	946
-----ERCVRPGFSGPLCEIETNECSSKPKNNGTQVDLTN---	545
-----ERCVRPGFSGPLCEIETNECSSKPKNNGTQVDLTN---	576
-----CTGSPGFSGSQCEIEINECYSTPCKNNGTQVDLNR---	355
-----IGARCETDIDECDSPCKNRRANCIDQPG---	114
-----TGASVSMAS---	176
-----CPPLLTMGLCECLMVGEESLDCNYTAPATQSPPR---	425

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

CNCFPGYHGPFCELDVKCKISPCLDDEENCVYRTDGYNCLCAPGYTGINCEINLDECLSE	1010
CNCFPEYHGPFCELDVKCKISPCLDDEENCVYRTDGYNCLCAPGYTGINCEINLDECLSE	1011
CNCFPGYHGPLCELDINECKTSPCLDEENCVYRADGYNCLCAPGYTGINCEINLDECLSE	1006
-----RCLSK	550
-----RCLSK	581
-----CEVNIDECLSE	366
-----RTTTTSTMA	434

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken

PCLHDGYCIDGINHYTCDCKSGFFGTHCETNANDCLSNPCLHGRCTELINEYPCSCDADG	1070
PCLHDGYCIDGINHYTCDCKSGFFGTHCETNANDCLSNPCLHGRCTELINEYPCSCDADG	1071
PCLHDGYCIDGINHYTCDCKSGFFGTHCETNANDCLSNPCLHG-----	1049
PCLHDGAWTDGVNHYTCDCKQSGFLGTHCETNANDCLSNPCLHG-----	593
PCLHDGAWTDGVNHYTCDCKQSGFLGTHCETNANDCLSNPCLHGR-----	625
PCLNDGICADGISYYTCYCKDGFIGTHCETNADACLSDPCLHG-----	409
-----NYFCQCVAPFK-----VVDGFYCLCN-----	135

Mutation spectrum of EYS in Spanish patients with autosomal recessive retinitis pigmentosa 21

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Zebrafish
Drosophila

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

```
-----  
PPTVRPVTPPETTVSPSRASEEVEIIVVTTSAPAEVVTSVLSPS-----
```

478

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

```
TSTQCKIKINDCTSIPCMNEGFCQKSAHGFTICIPRGYTGAYCEKSIDNCAEPEFNSVIC  
TSTQCKIKINDCTSIPCMNEGFCQKSAHGFTICIPRGYTGAYCEKSIDNCAEPELNSVIC  
-----NLGLGPSIPCMNEGFCQKSANGFTICQHGVTGAYCEERIDNCAKRELNSTLC  
-----IDCTSVSCLNEGICQKSVHGVTICIPGGYTGIVYCEMHVDGSAEPEPNLVLC  
-----  
-----  
-----  
-----  
-----
```

1130
1131
1102
676

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

```
LNGGICVDGPGHTFDCR-CLPGFSGQFCEININECSSSPCLHGADCEDHINGYVCKCQPG  
LNGGICVDGPGHTFDCR-CLPGFSGQFCEININECSSSPCLHGADCEDHINGYVCKCQPG  
LNGGICVDGPGDTFECRRCLPGFSGQFCDININECSSSPCLNGANCEDHINGYVCKCQPG  
-----R-CLPGFSGQFCEININECSSSPCLNGANCEDHINGYICKCQRG  
LNGGICVDGAGRTLYCR-CLPGFSGQFCEININECSSSPCLNGANCEDHINGYICKCQRG  
-----RRCPEGFSGNFCEVININECSSSPCLHGADCEDLVNRFNCKCQQG  
-----PGYAGLRCDQDIDDCIINTCDHNSTCKDLH-----  
-----LGHYVRWRLMAYQDPAKTMG-----  
-----CPLHFTGRLCEQDITVFSPR-----  
-----SSSSSSEEGVSVEIKTPTVAPPEESGSHSISVEQTTAVPAQPE
```

1189
1190
1162
636
735
453
165
197
112
520

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

```
WSGHHCNEKELECIIPNSCVHELCMENEPGSTCLCTPGFMTCSIGLLCGDEIRRIITCLTPIF  
WSGHHCNEKELECIIPNSCVHELCMENEPGSTCLCTPGFMTCSIGLLCGDEIRRIITCLTPIF  
WSGRRCNEKEVECIIPNSCVHELCMENEPGSTCLCIPGFMTCSIGLLCGDEIRRIISCLPPVF  
-----TCSIGLLCGDERRRITCLSPIS  
-----TCSIGLLCGDERRRITCLSPIS  
-----  
-----  
-----  
-----PESEQEPESKHPPESES
```

1249
1250
1222
658
757

537

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

```
QRTDPISTQTYTVPPSETLVSSFFPSIKATRIPAIMDTYPVDQGPKQTGIVKHDILPTTGL  
QRTDPISTQTYTIPPSETLVSSFFPSIKATRIPAIMDTYPVDQGPKQTGIVKHDILPTTGL  
QRTDAIVTQIYAVPPSETLVSSFFPSVKATRISTRMDTYPVDQGPKQTGIVKHDILPTTDW  
ARTDTISTQTHTPAPATSVHNFPRIGAPRLWTMDTYPVDQGPKQTDIFKHVDVLPPTGL  
ARTDTISTQTHTPAPATSVHNF-----MDTYPVDQGPKQTDIFKHVDVLPPTGL  
-----  
-----  
-----  
-----ASESETETEEEIIPGTTARPPTSRSSSSSEESPSIFTTLPPLPGKPQTSASSESSGEVVT
```

1309
1310
1282
718
807

597

Orangutan
Human
Marmoset
Horse

```
AALRISTPLESYLLEELIVTRELSAKHGLLSSADVSSSRFLNFGIHDPAQIVQDKTSVSH  
ATLRISTPLESYLLOELIVTRELSAKHSLSSADVSSSRFLNFGIRDPAQIVQDKTSVSH  
ATLRISTLSLEYLLKELIVTRELSAKHSLSSADVSSSRFLNFGIHDPAQIDRGKTSVSN  
AALGTGISFERYLLKHVIAAKELLLAKHSLSPSSADVSSSRFLNFGVPGPAQVWVKTSVPH
```

1369
1370
1342
778

22 <Barragán et al.>

1			
2			
3	Dog	AALGTGISFERYLLKHVIAAKELLAKHSLPSSTDVSSSRFLNFGVPGPAQVWVGKTSVPH	867
4	Opposum	-----	
5	Platypus	-----	
6	Chicken	-----	
7	Zebrafish	-----	
8	Drosophila	SEETYTVPHFEVSGSKSESGSEEVTTVRPTAAPSITISVDITSSGSSSSSSSESVEVFTTP	657
9			
10		p.R1374X	
11		↓	
12			
13	Orangutan	MPIRTSAAATLGFFFPDRRARTPFIMSSVMSDFIFPTQSLLFENYQTVASSATPTTSVIRS	1429
14	Human	MPIRTSAAATLGFFFPDRRARTPFIMSSLMSDFIFPTQSLLFENCQTVALSATPTTSVIRS	1430
15	Marmoset	MPIQTSAAATLGFFLSDRRARTPFIMSSLMTDFISPTQSLLFENYQTVASSATIMTSVIRS	1402
16	Horse	LPIQASAAATPRFFFLDRGERT-----SSMTDFIFPTQSLLFESDRSVASSATTMSSVISG	833
17	Dog	LPIQASAAATPRFFFLDRGERT-----SSMTDFIFPTQSLLFESDRSVASSATTMSSVISG	922
18	Opposum	-----IFPTQVSSLKNHQITISLSATDMSSAISS	481
19	Platypus	-----LPVKAIILKQVTTSSSSVIHKGSTQAP	192
20	Chicken	-----LVWTSYRFVLLDFLDSFVKNKLMSSVIHHL	226
21	Zebrafish	-----FDGNSFLELPSLTSLSLQSDTYFPS----	136
22	Drosophila	APVFVQRVTTIETSISIDYVTPPLPETTPRVVVPVRPTFAPEPPLDVVETTASTHHLW	717
23			
24			
25			
26			
27			
28	Orangutan	IPGADIELNRQSLLSRGFLLTAASISATPVVSRGAQEDIEEYSADSLISRREHWRLLSPS	1489
29	Human	IPGADIELNRQSLLSRGFLLIAASISATPVVSRGAQEDIEEYSADSLISRREHWRLLSPS	1490
30	Marmoset	IPGADIKLNRHSLLSRGFLLTAASISATPVVSRRAQEDIEEYSAVSLISRKEHWRLLSSS	1462
31	Horse	ILGADVVELNRHSLLSHGFLLKTASTGAPPVVMGAQEGIEEYSAVSLISRREYWRLLSSS	893
32	Dog	ILGADVVELNRHSLLSHGFLLKTASTGAPPVVMGAQEGIEEYSAVSLISRREYWRLLSSS	982
33	Opposum	IPGDEIELNSHSFLSHRFL-----QEDTREYSAFSLSS-----	515
34	Platypus	LLS-----EYSIMAVAS-----	204
35	Chicken	VYMQP-----	232
36	Zebrafish	-----	
37	Drosophila	TEVPTTAAPFFTEYPAEVLITTHRTSAGRFTTVQPPAGVTTTSPTEDSSVELPTPHTPQI	777
38			
39	Orangutan	MSPIFPAKIIISKQVTILNSSALHREFGKAFNPESEYQAITTEASSNQRLTNIKSQAADSLR	1549
40	Human	MSPIFPAKVIIISKQVTILNSSALHREFSTKAFNPESEYQAITTEASSNQRLTNIKSQAADSLR	1550
41	Marmoset	MSPIFPAKIIISKQVTILNSSAVHREFDTKAFIPSEYQVITEASSNQRLTNIKSQAADSLR	1522
42	Horse	MPPISPAKVIIISKQVAIVNSSSLHREFTTQDSIPSEYQVITEASSNQRLTNIKSQSADSLS	953
43	Dog	MPPISPAKVIIISKQVAIVNSSSLHREFTTQDSIPSEYQVITEASSNQRLTNIKSQSADSLS	1042
44	Opposum	-----QPADSLS	522
45	Platypus	-----SKGRLTNIKSQSADSLS	221
46	Chicken	-----VKIMSMG	239
47	Zebrafish	-----RSSDKR	143
48	Drosophila	VVTILDSEVIPSLLITTTGSPTHHHHHHHHPHEAEGETTLQPLEEDEHHHHHHHDEF TTP	837
49			
50			
51	Orangutan	ELSQTCAATCSMTEIKSSREFSDQVLHSHKQSHFYETFWMNSAILASWCALMGAQTITSGHS	1609
52	Human	ELSQTCAATCSMTEIKSSREFSDQVLHSHKQSHFYETFWMNSAILASWYALMGAQTITSGHS	1610
53	Marmoset	ELSQICTTCSMTEIKSSHEFSQVSHSKQSHFYETFWMNSAILASWYALMGAQTITSGHS	1582
54	Horse	ELSQTCAATCSMTEIKSSHEFSQVLHSHKQSHFYETFWMNSEILASWYALMRTQTITSGHS	1013
55	Dog	ELSQTCAATCSMTEIKSSHEFSQVLHSHKQSHFYETFWMNSEILASWYALMRTQTITSGHS	1102
56	Opposum	ELSQTCVTCMTKMKSSDECSVQALHSHKQSQFYEPFWMNSAILSSWYTLTGATVITSGHS	582
57	Platypus	ELIQTCAATCSMTEIKPSDEFHQVLHSHKQYQFYETFWMNSAILTSWYTLMRTTITSGHS	281
58			
59			
60			

Mutation spectrum of EYS in Spanish patients with autosomal recessive retinitis pigmentosa 23

1			
2			
3	Chicken	TLVSVKKSNIFFGKLFDFVFPDQVLHKSQSPFYEAFFWMNSAILNSWYALMGATAITSAYL	299
4	Zebrafish	ILYLTMKSG-----RTPHGSLLYCREQDLGERFLHVFLQNARAVARLGCGAAHILTA	194
5	Drosophila	QPVEITTGHPQLQTEDLIGVQEPVVVTTESPFAPAETTVVPVVVPATIAPLGTAAPPATPA	897
6			
7			
8	Orangutan	FSSATEITPSVAFTEVPSLFPSKKS\$AKRTILSSSLEESITLSSNLDVNLCLDKTCLSIVP	1669
9	Human	FSSATEITPSVAFTEVPSLFPSKKS\$AKRTILSSSLEESITLSSNLDVNLCLDKTCLSIVP	1670
10	Marmoset	FSSATEITPSVAFTEVPSLFPSKMSAKRTILSASLEESITLSSNLDVNLCSSEMTCLSIVP	1642
11	Horse	FSSATEIMPVSAFMEVSSSFPSKKS\$TKRRI\$TPSVEDSIALSTNLDANLCLDKTRLSIVP	1073
12	Dog	FSSATEIMPVSAFMEVSSSFPSKKS\$TKRRI\$TPSVEDSIALSTNLDANLCLDKTRLSIVP	1162
13	Opposum	FSSVTEITPSVEFTELSPPFSFKK-----	606
14	Platypus	FSPATEITSSVEFTELSSSFLPKMS-----	306
15	Chicken	FSSSRITSSVEFTEIHPHPLQK-----	323
16	Zebrafish	VAAQNIRIDSLVAITVRYALPSQN-----	218
17	Drosophila	PVPPATTTPPPSPPLATEPTLPPVTLPPVTQPPPTIPPSTQSAQLPPPTS	957
18			
19			
20	Orangutan	SQTISSDLMNSDLTSKMTTDELSV\$ANILKLLKIRQYGITMGPT\$EVLNQD\$SLLDMEKSKG	1729
21	Human	SQTISSDLMNSDLTSKMTTDELSV\$ENILKLLKIRQYGITMGPT\$EVLNQESLLDMEKSKG	1730
22	Marmoset	SQTM\$LDL\$MNSDLTSQPTNDQ\$LSV\$ENILKLLQIRQYGITMGPT\$EELNQD\$SLLDMEKCKG	1702
23	Horse	SQTVSSDLLNSDLTSELT-EDLSV\$ENILKLLKIGQYGITMGPT\$EVLNQD\$NLLAVHESKG	1132
24	Dog	SQTVSSDLLNSDLTSELT-EDLSV\$ENILKLLKIGQYGITMGPT\$EVLNQD\$NLLAVHESKG	1221
25	Opposum	-----	
26	Platypus	-----	
27	Chicken	-----	
28	Zebrafish	-----	
29	Drosophila	AINVYTTDPGPPTASQTKPSVTESS\$EEVEGTNTVSTGGRGSGGVPEEKAGDVDCIKLGCY	1017
30			
31	Orangutan	SHTL\$FKLHPSDSSLD\$FELNLQIY\$PDVTLKTYSEITHAN\$DFKNNL\$PPLTGSV\$PDFSEVTTN	1789
32	Human	SHTL\$FKLHPSDSSLD\$FELNLQIY\$PDVTLKTYSEITHAN\$DFKNNL\$PPLTGSV\$PDFSEVTTN	1790
33	Marmoset	SHTL\$FKLHPCDSSLD\$FQNLQ\$SHPDVTLR\$TYSEI\$IHAN\$DLKNNL\$PPLTGSIP\$DFSEVSTN	1762
34	Horse	SHKQLKLHTSDRSLDFELNLPSH-----PLKNNLPPYMD\$RSDLSEVTSN	1177
35	Dog	SHKQLKLHTSDRSLDFELNLLKQ-----EVR-----	1247
36	Opposum	-----	
37	Platypus	-----	
38	Chicken	-----	
39	Zebrafish	-----	
40	Drosophila	NGGTCVTTSEGR\$CVCRFDRQGP-----	1040
41			
42			
43	Orangutan	VAFYTVSATPALS\$IQTSS\$MSVIRPDW\$PYFTDYMT\$SLKKEV\$KTSSEW\$SKWELQPSVQYQE	1849
44	Human	VAFYTVSATPALS\$IQTSS\$MSVIRPDW\$PYFTDYMT\$SLKKEV\$KTSSEW\$SKWELQPSVQYQE	1850
45	Marmoset	VAFYTVSATPALS\$IQTSS-MSVTRPEW\$PDFTDYV\$NALKK\$DIKTSSEW\$SKWELQPSVQYQE	1821
46	Horse	VAFYTVSAT-----QSLPVQTST\$SSEW\$SKWELQPSVHDWE	1212
47	Dog	-----TSSEW\$SKWELQPSVHDWE	1265
48	Opposum	-----	
49	Platypus	-----	
50	Chicken	-----	
51	Zebrafish	-----	
52	Drosophila	-----	
53			
54			
55	Orangutan	FPTASWHLPFTRSLTLSSLESILAPQQLTISDFSCVRYYGDSYLEFQNVVLLNPQNNISLE	1909
56	Human	FPTASRHLPFTRSLTLSSLESILAPQRLMISDFSCVRYYGDSYLEFQNVALLNPQNNISLE	1910
57	Marmoset	FPTASRHLLFTRSF\$TLSSLESILAPQQLMISDFSCVRYYGDSYLEFQNVLLNPQNNIYLE	1881

24 <Barragán et al.>

1				
2				
3	Horse	SPAASQTPAITRSLTLP	SDFTVCVYGD	1272
4	Dog	SPAASQTPAITRSLTLP	SDFTVCVYGD	1325
5	Opposum	-----	SDFSYMQYFGDSYLK	636
6	Platypus	-----	-----	
7	Chicken	-----	IHTLIGLSNFGEQFLH	350
8	Zebrafish	-----	-----	
9	Drosophila	-----LCELPIIIRNAAF	SGDSYVSHRIYKDIGG	1083
10				
11				
12	Orangutan	FQTFSSYG	-----PGEAKFKSINTT	1964
13	Human	FQTFSSYG	-----PGEAKFKSINTT	1970
14	Marmoset	FQTFSSYG	-----PGEAKLKSINTT	1941
15	Horse	FQTFSSYG	-----PGEAKLKNINTT	1332
16	Dog	FQTFSSYG	-----	1360
17	Opposum	FQTFNSHGL	-----	671
18	Platypus	-----NFGEQFLH	-----	323
19	Chicken	FSCGNSQN	-----	381
20	Zebrafish	-----	-----	
21	Drosophila	KVRTRATNGL	-----	1124
22				
23				
24				
25				
26	Orangutan	YTLLIR-QELYPCKAEL	-----	2023
27	Human	YTLLIR-QELDPCNAEL	-----	2029
28	Marmoset	YTLLIR-QELDPCNAEL	-----	2000
29	Horse	YALLIR-QELDPCNAEL	-----	1386
30	Dog	-----RQELDPCNAEL	-----	1408
31	Opposum	-----	-----	
32	Platypus	-----	-----	
33	Chicken	-----	-----	
34	Zebrafish	-----	-----	
35	Drosophila	-----	-----	
36				
37	Orangutan	PVKNFTGCIEVIEINN	-----	2083
38	Human	PVKNFTGCIEVIEINN	-----	2089
39	Marmoset	PVKNFTGCIEVIEINN	-----	2060
40	Horse	PVENFTGCIEVIEINN	-----	1446
41	Dog	PVENFTGCIEVIEINN	-----	1468
42	Opposum	-----	-----	
43	Platypus	-----	-----	
44	Chicken	-----	-----	
45	Zebrafish	-----	-----	
46	Drosophila	-----	-----LQTMLLSEL	1133
47				
48				
49	Orangutan	WTSVSPSVAAP	-----	2143
50	Human	WTSVSPSVAAP	-----	2149
51	Marmoset	WTSITPSAAAP	-----	2120
52	Horse	WTSLSAPPAAP	-----	1505
53	Dog	WTSLSAPPAAP	-----	1527
54	Opposum	-----SVCQDSL	-----	711
55	Platypus	-----VKLGCD	-----	352
56	Chicken	-----VSSLEGY	-----	412
57	Zebrafish	-----NGQLCF	-----	246
58				
59				
60				

p.G2017V



Mutation spectrum of EYS in Spanish patients with autosomal recessive retinitis pigmentosa 25

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Drosophila ETPVNTGHEITIRAELEDFSRNYTHCNASLLVNDTLAMSGDQPTWLKLLPPR----- 1184

Orangutan FNGNSYLELPFLKFVLEKEHNRVTIYLTIKTNSLNGTILYSNGNNGFKQFLHLFLVEGR 2203
Human FNGNSYLELPFLKFVLEKEHNRVTIYLTIKTNSLNGTILYSNGNNGCGKQFLHLFLVEGR 2209
 Marmoset FNGNSYLELPFLNSVLEKEHNRVTIYLTIKTNSLNGTVLYSNGNDFGKQFLHLFLVEGR 2180
 Horse FNGNSYLELPFLKSVLEKEHNRIVTIYLTIKTNTLNGTILYS----FGQQFLHLFLLEGR 1561
 Dog FNGNSYLELPFLKSVLEKEHNRIVTIYLTIKTNTLNGTILYSN---FGQQFLHLFLLEGR 1584
 Opossum -----
 Platypus -----
 Chicken -----
 Zebrafish -----
 Drosophila -----

p.S2211L p.I2239SfsX17

Orangutan PSVKYGCNSQNILTVSANYSINTNAFTPIITIRYTPVGS PGVVCMIEMTADGKPPVQKK 2263
Human PSVKYGCNSQNILTVSANYSINTNAFTPIITIRYTPVGS PGVVCMIEMTADGKPPVQKK 2269
 Marmoset PSVKYGCNSQNILTVSANYSINTNAFIPITIRYTMPVGS PGVICMIEMTADGKSPAQKK 2240
 Horse PTVKYGCNSQNILTVSANYSINTNVFIPITIR----- 1594
 Dog PTVKYGCNSQNILTVSANYSINTNVFIPITIRYTI PVGSPGVACMIEMTADGKPPIQKK 1644
 Opossum -----
 Platypus -----
 Chicken -----
 Zebrafish -----
 Drosophila -----

Orangutan DTEISQASQAYFESMFLGHIPANVQ----IHKKSGPVYGFRCILDQVNNKEFFIIDEA 2319
Human DTEISHASQAYFESMFLGHIPANVQ----IHKKAGPVYGFRCILDQVNNKEFFIIDEA 2325
 Marmoset DTKISHASQAYFESMFLGHIPANVQ----IHKKAGPVYGFRCILDQVNNKEFFIIDEA 2296
 Horse -----QAYFESMFLGHVPTNVK----IHKKAGPIYGFRCIRELQVNDKEFFIIDEA 1642
 Dog DTETPHASQAYFESMFLGHVPTNVK----IHKKAGPIYGFRCIRELQVNDKEFFIIDEA 1700
 Opossum -----KFGSIFLGHIPAKVK----VHENMGQTYGYRGCIREFQVNNKELFIIDEA 757
 Platypus -----IRSQANLGSMLGNIPAN-----GFRGCIREMQVNNKELFIIDEA 392
 Chicken -----QITFGSTFLGNVPVHKE----VPECAGQIRGYKGCIRDFQVNNKELFIIDDA 460
 Zebrafish -----EVVFGPTFLGGFPVLE----LHNSGNVSGFIGCIRELQMGSKELYVVGEA 294
 Drosophila ---LHTPEAILNTWLHLGGAPQAPIGLIIE LPPAQSGSFTGCLHTLRINGQAREIFGDA 1241
 . ** * * : ** : : : : : : : . . . *

Orangutan RHGKNIENCHVPWCAHHLCRNNGTCIS-----DNENLFCECPRLYS 2360
Human RHGKNIENCHVPWCAHHLCRNNGTCIS-----DNENLFCECPRLYS 2366
 Marmoset RRGKNIENCHVPWCAHHPCHNNGTCIS-----DSENLFCCEPRLYS 2337
 Horse LRGRNIENCHVPWCAHHLCHNNGTCIS-----DSENFCECPRLSS 1683
 Dog LRGRNIENCHVPWCAHHLCHNNGTCIS-----DSENFCECPRLSS 1741
 Opossum LSGKNIENCHVPVCNYHPCRNGGTCIS-----DTENWLCCECLQLYS 798
 Platypus LSGRNIENCHIPVCDYHPCRNGGTCIS-----NTENWFCECPGLYS 433
 Chicken LGGRNVENCHVPICDYHPCRNGGTCIRS-----DAENWFCECPKLYS 502
 Zebrafish IRGQNIQNCDAAVCQHQPCCRNGGTCISLNPPVPLPNLCFLRKLHQSDAESWFCACPSLYS 354
 Drosophila LDGFGITEGSLACLSSPCRNGAACIKIETN-----DLDENGEKAEKWKCKCPTGYM 1293
 * . : * * * : * . * *

Orangutan GKLCQFASCENNPCGNGATCVPKSGTDIVCLCPYGRSGPLCTD-----AINITQPRFSGT 2415
Human GKLCQFASCENNPCGNGATCVPKSGTDIVCLCPYGRSGPLCTD-----AINITQPRFSGT 2421
 Marmoset GKLCQFASCENNPCGNGATCVPKSGTDIVCLCPYGRSGPLCTD-----AINIIQPRFSGT 2392

26 <Barragán et al.>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

GKLCQFATCENNPCGNGATCVPKSGTEIVCLCPYGRSGVLCTD	-----	AINITQPSFSGT	1738
GKLCQFATCENNPCGNGATCVPKSGTEIVCLCPYGRSGVLCTD	-----	AINITQPSFSGT	1796
GKLCQFATCENNPCGNGATCFPKSNRDVCLCPYGRSGVLCTD	-----	AINITKPSFSGT	853
GKLCQFSSCEKNPCGYGATCFPKSNQDAVCLCPYGRGILCND	-----	AITITRPSFSSST	488
GRLCQFMTCEDESPCGNGATCFPKSRQDVVCLCPYGRSGILCND	-----	VVNISQPSFSGT	557
GKLCQFTACERNPCARGATCVPQTQLEAACLCPYGRQGLLCDEGRSRQ	AINITRPKFSGL		414
GPTCEISVCEDNPCQYGGTVCVQFPGSGYLCLCPLGKHGHYCEHN	-----	LEVALPSFSG-	1347

p.H2454PfsX8

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

DAFGYTSFLAYSRI SDISFH YEFHLKFQLANNHSALQNNLIFFTGQK	GHGLNGDDFLAVG	2475
DAFGYTSFLAYSRI SDISFH YEFHLKFQLANNHSALQNNLIFFTGQK	GHGLNGDDFLAVG	2481
DAFGYTSFLAYSRI SDISFH YEFHLKFQLANNHSALQNNLIFFTGQK	GHGLNGDDFLAVG	2452
DAFGYTSFLAYS RVPD IGF D YEFH VTFQLANNHSALQNNLIFFTGQK	GHGRNGDDFLAVG	1798
DAFGYTSFLAYS RVPD IGF D YEFH VTFQLANNHSALQNNLIFFTGQK	GHGRNGDDFLAVG	1856
DSFGFTSFLAYSRI PDISF YEFHLKFQLANNNSALQDNLIFFTGQK	GQGLNGDDFLAVG	913
DAFGYTSFLAYSRI PDISS YEFRLKFQLAANISALQNNLIFFTGQK	KGRG-----	538
DVFGYTSFLAYSTIPDI TFY YEFHLKFQLLNHHSALQDNLIFFTGQK	GQGLNGDDFLVLG	617
DEFGYSSVAYPSIPSTGH F YEFHLKLT FANNASALRNLI LFSGQK	GQGLSGDDFFALG	474
SVNGLSSFVAYTVP--IPLEYSLELSFKILP-QTMSQISLLAFFGQS	GYHDEKSDHLAVS	1404

p.E2503K

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

LLNGSVVYSYNLGSGLASIRSE P-----LNL SLGVHTVHLGKFFQ---	LKVDDHK NKSII	2527
LLNGSVVYSYNLGSGLASIRSE P-----LNL SLGVHTVHLGKFFQ	EGWLKVDDHK NKSII	2536
LLNGSVVYSYNLGSGLASIRSE P-----LDLSLGVHTVHLGKFFQ	EGWLKVDDHK NKSII	2507
LRDGRVVYSYNLGSGLASVSSDP-----LDRSLGIHAVRLGRFLQ	MGWLKVDDHK NKSIV	1853
LRDGRVVYSYNLGSGLASVSSDP-----LDRSLGIHAVRLGRFLQ	MGWLKVDDHK NKSIV	1911
LRNGCLVYSYNLGSGLTANLHSDP-----LNL SLRVHV VHLGRSFQT	GWLVKVDH QNKSIT	968
-----	VDNQKNKSIT	548
LCDGRVVYSYNLGSGLATII SKP-----LDLTLNIHVIHLGRYLQ	KGWLVKVDQKNKT TTT	672
VRNGRIVHKYNLGSGLATII SDR-----LNPRINIHTVHFGRYLK	TGWLVKVNQKRRRTGT	529
FIQGYIMLTWNLGAGPRRI FTQKPIDFRLDAPRVPYEIKVGRIGRQ	AWLSVDGKFNITGR	1464

Orangutan
Human
Marmoset
Horse
Dog
Opposum
Platypus
Chicken
Zebrafish
Drosophila

APGRLAGLNVFSQFYVGGYSEYTPDLLPNGADFKNGFQGCIFTLQ	VRTEKDG HFRGLGNP	2587
APGRLVGLNVFSQFYVGGYSEYTPDLLPNGADFKNGFQGCIFTLQ	VRTEKDG HFRGLGNP	2596
APGRLVGLNVFSQFYVGGYSEYTPDLLPNGADFKNGFQGCIFTLQ	VRTEKDG HFRDLGNP	2567
APGRLVGLNVFSQFYVGGYSEYTPPELLPNGSEFKNGFQG-----	-----	1892
APGRLVGLNVFSQFYVGGYSEYTPPELLPNGSEFKNGFQG-----	-----	1950
SPGRLVGLNVFSQFYVGGYNE YTPPELLPNESKFNQGFQG-----	-----	1007
SPGRLVGLNVFSQFYVGGYSE YTPPELLPNSDFKNGFQG-----	-----	587
SPGRLVGLNVFSQFYLGGYHE YTPPELLPKGSRFKNGFQGCIFD	VQVRNMQEFKSPGTP	732
SPGRLMGLNFTFSQLYIGGYEE YTPPELLPPGSRFQNSFQGCIF	DMLFRTRQDGK FHALGGP	589
SPGSGSRMDVLPILYLGGH EIANFNTLPHDLPLHSGFQGCIDV	QLKAG-----QVTVPLQ	1520

p.W2640X

Orangutan
Human
Marmoset
Horse
Dog

EGHPNAGRSVGQCHASPCSLMKCGNGGTCIESGTSV-CNCTTGW	KGAFCTETVSTCDPEH	2646
EGHPNAGRSVGQCHASPCSLMKCGNGGTCIESGTSVYCNC	TGWKGSFCTETVSTCDPEH	2656
EGHPNAGRSVGQCHTSPCSLMKCGNGGTCIES-----GCNCTTG	WKGAFCTETVSTCDPEH	2623
-----	CDCPSGWKGAFC TEMVSTCDPEH	1915
-----	CDCPSGWKGAFC TEMVSTCDPEH	1973

Mutation spectrum of EYS in Spanish patients with autosomal recessive retinitis pigmentosa 27

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Opposum	-----CIFDVQVRTRKDHFRSLGDPEG	1030
Platypus	-----CDCVLGWKGMFCTETVSI CDQEH	610
Chicken	EGHPNSGRSVGQC KASPCSLIKCRNGGKCMESG--STVCNCLAGWKGAFC TEMVTVC DPEH	791
Zebrafish	DIRPLSGRNVGQC GVNPCSLVFC HNGGTCVDSGSSVYCQVFGWK GALCSEKVSFCDAEH	649
Drosophila	ETRGVVRGRGVGQC GTR ECHR HACQHDGACLQHGATFTICQEGWYGPLCAQPTN PCD SFN	1580

Orangutan	DPPHHC SRGATCISLPHGYTCFCPLGTTGIYCEQ ALSISDPSFRSNELSWMS-----	2698
Human	DPPHHC SRGATCISLPHGYTCFCPLGTTGIYCEQ ALSISDPSFRSNELSWMS-----	2708
Marmoset	DPPHNCSRGATCISLPHGYTCYCP LGTTGIYCE--ALSISDASFRSNELSWMS-----	2674
Horse	DPPHNCSK GATCVPLPHGYTCRCPLGTTGIYCE--ALSVSDPSFRSHEL SWMS-----	1966
Dog	DPPHNCSK GATCVPLPHGYTCRCPLGTTGIYCE--ALSVSDPSFRSHEL SWMS-----	2024
Opposum	HPNAGRSVG-QCEVFPCSLIKCQNGGTC-----IALSISDPSFRSHDSSWMS-----	1076
Platypus	DPPHQCRPGATCVPLTIGYTCHCPLGTTGIHCE--ALSISDPSFRRNES SWMS-----	661
Chicken	DPPHLCKQGGTCVPLP NGYMCHCPLGTSGYTCEQDISISDPSFRSNKSSWMS-----	843
Zebrafish	IPPPFCARGSTCVPLSDGYTCQCP LGSAGLHCQQAITISDPPFFSGNQSSWMS-----	701
Drosophila	N---KCYEDATCVPLVNGYECDCPVGR TGKNCEEVIRSLSDVSLTGRRSYLAVRWPYLYD	1637

Orangutan	-----F	2699
Human	----- F	2709
Marmoset	-----F	2675
Horse	-----F	1967
Dog	-----F	2025
Opposum	-----F	1077
Platypus	-----F	662
Chicken	-----F	844
Zebrafish	-----F	702
Drosophila	GGDKLGAKRSQMVSYRNF TKKLMPPKPIITPSSHVMKLLNEVEKQRSFSPVPLMGSKSF	1697

Orangutan	ASFHVRRKTHIQ LQFQPLAADGILFYAAQHLKAQSG--DFLCISLVNSSVQLRYNLGDRT	2757
Human	ASFHVRRKTHIQ LQFQPLAADGILFYAAQHLKAQSG--DFLCISLVNSSVQLRYNLGDRT	2767
Marmoset	ASFVRRKTHIQ LQFQPLAADGILFYAAQHLKAQSG--DFLCISLANGSVQLRYNLGDRT	2733
Horse	SSFRIKRKTHIQ LQFRPLSADGILFYVAQNLKAQSG--DFLCISLVNGSVQLRYNLGDRT	2025
Dog	SSFRIKRKTHIQ LQFRPLSADGILFYVAQNLKAQSG--DFLCISLVNGSVQLRYNLGDRT	2083
Opposum	APFYIRQKMHIQ LQFQALSTDGILFYTAQHLHSRSG--DFLCLSLTRGSVQLRYNLGDRT	1135
Platypus	APFHIRHKTHIQ LQFQPLSADGILFYTAQHLSTHSG--DFLCLSLANGYVQLRYNLGDRT	720
Chicken	APFYIRHKTHIKLQ FQPLSPDGILFYTAQRLGTQSG--DFLCISLVNGFIQLRYNLGDKT	902
Zebrafish	PPINIRHRTHVQLQ FQTLSP EGILFYTAQHLSTHSG--DFLSISLSAGFLQLRYNLGNQT	760
Drosophila	EEHHRVQFFFIEFQLRPLSERGLLLYFGTLN NNQDKKIGFVSLSLQGGVVEFRISGPSNH	1757

Orangutan	IILETLQKVTINGSTWHI IKAGRVGAEGYLDLDGINVTEKASTKMSSLDTNTDFYIGGVS	2817
Human	IILETLQKVTINGSTWHI IKAGRVGAEGYLDLDGINVTEKASTKMSSLDTNTDFYIGGVS	2827
Marmoset	IILETLQKVTINGSTWHI IKAGRVGAEGYLDLDGINVTEKASTKMSSLDTNTDFYIGGVS	2793
Horse	IILETLQKVN MNGSTWHVIKAGRVGAEGYLDLDGKT VTEKAKAEMNSLDTNTDFYIGGVS	2085
Dog	IILETLQKVN MNGSTWHVIKAGRVGAEGYLDLDGKT VTEKAKAEMNSLDTNTDFYIGGVS	2143
Opposum	VILQSLQK-----WHTIKAGRVGNEGYLELDGINVTKGATAGMRALDTSTDFYIGGVS	1188
Platypus	VILQSFQKVDTS GDAWHLIKAGR HGTEGYLTDGNTVTKATGRMTVLDTNTDFYVGGVS	780
Chicken	IVLQAVQKVHADGQ TWHVLKVRGVGNEGYVDLDGINITHASAGMNVLDTH TDFYVGGVS	962
Zebrafish	IVLQSPKELDVTGVRWHTVKAGRE GNSGFLIVDGESVTRNSSEGSTTLDVGANIFIGGIS	820
Drosophila	VTVRSVRMLAIG--EWHKIKMAQRGRWLT LWVEGSASSALAPSAEVLVEPDSL LYIGGLK	1816

28 <Barragán et al.>

1
2
3
4 Orangutan
5 **Human**
6 Marmoset
7 Horse
8 Dog
9 Opposum
10 Platypus
11 Chicken
12 Zebrafish
13 Drosophila

SLNLVNPMAIENEPEVGFQGCIRQVIIINNQELQLTEFGAKGGSNVGDCDGTACGYNTCRNG	2877
SLNLVNPMAIENEPEVGFQGCIRQVIIINNQELQLTEFGAKGGSNVGDCDGTACGYNTCRNG	2887
SLNLINPMAIENEPEIGFQGCIREVIIINNQELQLTESGAKGGSNVGDCDGTACGYNICRNG	2853
SLNLVNPMAIANEPEVGFQGCIREVIIINNQELQLTELGAKGGSNVGDCDGTACGYNVCNR	2145
SLNLVNPMAIANEPEVGFQGCIREVIIINNQELQLTELGAKGGSNVGDCDGTACGYNVCNR	2203
SLSLVNPMAIENEPEVGFNGCVREILINGRELKLTEAGAKRGSNVGDCDGTPCGYKVCENK	1248
SLDSVNSMAVENDPVGFDDGCVREIFINNRELKLTDKGAKDGLNIGDCDGTACGYTVCKNK	840
SLNLVNSMATENEPTGFGSIREIVINDKELKLTVDPKGGANIGDCDGTVCYSVCKNN	1022
SLNTVSIDAVEKELVGFTEGGIREVVVNGQELLELTETGALDGANVGDWDGTACGYKVCNK	880
DVSKLPHNAISGFPIPFRCVVRGLVVSCTRIVLNETNIVESRNIRDGDGTACGGDSCESE	1876
. . . : * * * : * * : * * . . . : * * * * * * * *	

p.G2945E

18 Orangutan
19 **Human**
20 Marmoset
21 Horse
22 Dog
23 Opposum
24 Platypus
25 Chicken
26 Zebrafish
27 Drosophila

TECTVNGTTF-SCRCLPDWAGNTCNQSVYCLNNLCLHQSLCIPN-QSFSYSCLCTLGWVG	2935
TECTVNGTTF-SCRCLPDWAGNTCNQSVYCLNNLCLHQSLCIPN-QSFSYSCLCTLGWVG	2945
GKCTVNGTTF-SCRCLPDWAGNTCNQSAAYCLNNLCLHQSLCIPD-QSFSYSCLCTLGWVG	2911
GECVNGTTF-SCQCSPPWAGNTCEQSAAYCLNNLCLHQSLCVPD-QSSSYRCLCTLGWEG	2203
GECVNGTTF-SCQCSPPWAGNTCEQSAAYCLNNLCLHQSLCVPD-QSSSYRCLCTLGWEG	2261
GQCRAGQSKF-SCKCLQPWIGKRCSESANCRNNLCLHHSRCIPV-QPAAYICLPLGWVG	1306
GECILHHTNF-SCKCTPGWAGNTCEQSMNCLNNKCCQHSLCIPD-NTFSYSCACPLGWVG	898
GTCQVESSGF-SCSCPPQGWIGNTCESVHCLHNRCSRQALCIPQPALLSYTCVCPGLWSG	1081
GHCHPSAG-----	888
GHCWLEDEKLQPHICPEYAKGDRCEYSETCKLIPCKNNGRCLRS-----GRCSCPNGWGG	1931
* *	

29 Orangutan
30 **Human**
31 Marmoset
32 Horse
33 Dog
34 Opposum
35 Platypus
36 Chicken
37 Zebrafish
38 Drosophila

RYCENKTSF-STAKFMGNSYIKYIDPNYRMRNLQ-----FTTISLNFSTTKTEG	2983
RYCENKTSF-STAKFMGNSYIKYIDPNYRMRNLQ-----FTTISLNFSTTKTEG	2993
RYCENKTSF-STAKFMGNSYIKYIDPNYRMRNLH-----FTTISLNFSTTKTEG	2959
RYCENKISF-STAKFMGNSYIKYIDPDYRMRNHH-----FTTVSLNFSTTEG	2251
RYCENKISF-STAKFMGNSYIKYIDPDYRMRNHH-----FTTVSLNFSTTEG	2309
RYCDNETSF-ITAKFVGNYSYIKYTDPNYKRDLR-----FTAVSLNFSTTITEG	1354
GYCETEILF-LIAKFQGNYSYIKHTDPNYGKRNH-----FTTVSLNFSTTEG	946
KHCDSKISF-FTAKFVGNYSYIKYIDPLYGKRDLO-----YSRISLNFSTTQIEG	1129
-----DG	890
FYCEIAMS KPTTPSFRGNSYLILPPPRIPMKDKRRGPSLYVRPREAIQVSLNFSTIEPDG	1991
: *	

42 Orangutan
43 **Human**
44 Marmoset
45 Horse
46 Dog
47 Opposum
48 Platypus
49 Chicken
50 Zebrafish
51 Drosophila

LIVWMGTAQNEENDFLAIGLHNQTLKIAVNLGERISVPMSYNGTFCCNK-WHHVVIQN	3042
LIVWMGIAQNEENDFLAIGLHNQTLKIAVNLGERISVPMSYNGTFCCNK-WHHVVIQN	3052
LIVWMGIAQNEENDFLAIGLHNQTLKIAVNLGERISVPMSYNGTFCCNK-WHHVVIQN	3018
LIVWIGKAQNEENDFLAIGLHNQSLKIAVNLGESISVPVIYSNGTFCCNK-WHHVIVSQN	2310
LIVWIGKAQNEENDFLAIGLHNQSLKIAVNLGESISVPVIYSNGTFCCNK-WHHVIVSQN	2368
LILWMGKAEHEENDFLAIGVHNRTLKVMVNLGERISVP-----WHHVIVSQN	1401
LILWMGKAEHEEDNFLAIGLSNGLKVVVNLGNLSVP-----WHYITVAQN	993
LMVWLKAEDEEDNFLAIGLANGRLKVVINLGERISVPMIHSKDSICTDERWHFVTIQN	1189
LIFWMGKAESSEDDHLAVGLQDGYLKISVNLGERITALPLVYQN-SFCCNY-WNYLSITHN	948
LLLWS---EHERSKFLGLGLEAGHLKLAASNLGSTNDTVRAPASGFIADGAWHWTSVLLD	2048
* : *	

p.N3061TfsX3

55 Orangutan
56 **Human**

QTLIKAYVNNLILSEDIDP-----HKNFVALNYDGI CYLGGFEY	3082
QTLIKAYINNLLILSEDIDP-----HKNFVALNYDGI CYLGGFEY	3092

Mutation spectrum of *EYS* in Spanish patients with autosomal recessive retinitis pigmentosa 29

1			
2			
3	Marmoset	QTLIKAYLNNSLILSEDIDP-----HKNFVALNYDGCYLGGFY	3058
4	Horse	QTLIKAYLDDNLILSEDIDP-----HKKFVALNYDGISYLGGFY	2350
5	Dog	QTLIKAYLDDNLILSEDIDP-----HKKFVALNYDGISYLGGFY	2408
6	Opposum	QTYFKVYLDGDLILFEDIDP-----TKNYIALNYGGVSYLGGFEF	1441
7	Platypus	QTIKQVYLDSDILILFEDIDP-----QRKHIALNYGGVSYLGGFEL	1033
8	Chicken	QTCIKVYLDDEELIIFEDIDP-----HRKYTALNYGGICYFGGFEL	1229
9	Zebrafish	RTLTIQVYVNEERVIFEDIDP-----FEQYVAVNYGGVIYLGGFEL	988
10	Drosophila	RSRLELQLDGEVIFTERLPEGGRSLGSTTPRSTLAGRRKNSSKEPTISYEDVFYLGGFNP	2108
11		:: :: :: . :: * :	
12			
13			

p.Y3135X



14			
15	Orangutan	GRKVSIVTQEIFKTNFV-----FFQDP-KKIELIK-LEGYNVYDGDEQNEVT---	3127
16	Human	GRKVNIVTQEIFKTNFVGVKIKDVVFFQEP-KKIELIK-LEGYNVYDGDEQNEVT---	3144
17	Marmoset	GRKVNIIITQEIFKTNFVGVKIKDVVFFQDP-KKIELIK-SEGYNVYDGDEQNEVT---	3110
18	Horse	GRKVNIVTQEIFKRNDFVGVKIKD-VFFQDS-KKIELIK-SEGYNVYNGDEQNVT----	2400
19	Dog	GRKVNIVTQEIFKRNDFVGVKIKD-VFF-----	2433
20	Opposum	GRGVMVITQGLFNRFVIGKIKDVVFFQDQ-KKIELIK-SEGYNIYNGDE-----	1488
21	Platypus	GHNVSIVTQGLYSQHFVGVKIKDVVFFQES-TKIQLIE-LEGYNVYSGDE-----	1080
22	Chicken	GRKVHTATAGLFQKEFIGKIKDIALFQDS-KKIQLMK-GEGYNVHNGDHRN-----	1278
23	Zebrafish	NRDVASVTSGVFTKGFEGSIKDVFLYQDT-KQLQFLQTCEGFNVYQGE-----	1036
24	Drosophila	SDSVSRRTKGRFFDPPFKGCLQDIQFGAEPITAIISDFSTYQGENIGSCDLHGDEPLTV	2165
25		. * * : * :	

Supplementary Figure S1. *EYS* homologues protein alignment and domain distribution with identified variations. 5' UTR and splice site variations are not included in this figure.

ACKNOWLEDGEMENTS

We would like to thank the families who participated in the study, as well as Almudena Ávila, Teresa Jaijo and Sara Bernal for their help on DNA collection. This study was funded by PN de I+D+I 2008- 2011, Instituto de Salud Carlos III (ISCIII) –Subdirección General de Evaluación y Fomento de la Investigación, Fondo de Investigación Sanitaria (PI081131), Spain; Consejería de Salud (PI-0334/2007), Consejería de Innovación, Ciencia y Empresa (PI08-CTS-03687), Junta de Andalucía, Spain; El Centro de Investigación Biomédica en Red de Enfermedades Raras is an initiative of the Instituto de Salud Carlos III. J.I. P. was supported by Fondo de Investigación Sanitaria and M. G-dP by Consejería de Innovación, Ciencia y Empresa, Junta de Andalucía, Spain. The research of K.W.L. was funded by the Stichting Wetenschappelijk Onderzoek Oogziekenhuis Prof. Dr. H.J. Flieringa, Rotterdam.