DEPARTAMENTO DE CIÊNCIAS DA VIDA

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Autosomal Dominant Optic Atrophy Type 1: Study of 4b and 5b exons of OPA1 gene contribution

Gonçalo Ferreira

2014

DEPARTAMENTO DE CIÊNCIAS DA VIDA

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Autosomal Dominant Optic Atrophy Type 1: Study of 4b and 5b exons of OPA1 gene contribution

Tese realizada no âmbito do Mestrado em Evolução e Biologia Humanas, com a orientação científica do Doutor Filipe Silva (COCV-IBILI, FMUC) e a orientação interna da Professora Doutora Eugénia Cunha (DCV, FCTUC).

Gonçalo Ferreira

2014

Índice

Lista de figuras

Resumo

A atrofia óptica autossómica dominante (ADOA) é uma neuropatia óptica hereditária que foi descrita pela primeira vez no final do século XIX. A ADOA é uma doença rara, apresentando uma prevalência que varia consoante a localização geográfica. Na Dinamarca, é possível encontrar uma prevalência superior ao resto do mundo, sendo de 1:10.000. Caracteriza-se pela perda insidiosa da visão, desde moderada a grave, com início na infância. A forma sindrómica desta doença, apesar de ser rara, pode estar associada a consequências mais graves, como por exemplo ataxia, surdez, miopatia, neuropatia, entre outras.

O gene *OPA*1 está localizado no braço longo do cromossoma 3 (3q28-29), tendo sido identificadas mais de 200 mutações em doentes com ADOA, das quais 27% são *missense*, 27% são uma forma alternativa de *splicing*, 23% são *frameshift*, 16,5% são *nonsense* e 6% devem-se a rearranjos pontuais ou cromossómicos. Este gene codifica para 8 isoformas proteicas diferentes, devido a *splicing* alternativo, existindo variações na sua expressão nos diferentes tecidos.

Muitos investigadores consideram que os exões alternativos 4b e 5b não codificam para a proteína responsável pelo fenótipo de ADOA. No entanto, foi possível determinar através de estudos funcionais, que estes dois exões parecem ser muito importantes para a função da proteína Opa1 na manutenção da mitocôndria, para além de definirem as suas 8 isoformas. Esta proteína caracteriza-se por ter dez domínios: um domínio GTPase, um domínio intermédio, um domínio GTPase *effector* (GED), um domínio de importação para a mitocôndria (MIS) na região N-terminal, três domínios transmembranares (TM1, TM2a e TM2b) e três domínios *coiled-coil* (CC-0, CC-I e CC-II). A proteína Opa1 tem um local de clivagem, em que o domínio MIS é removido após importação para a mitocôndria, e pode ter dois locais de clivagem extra localizados nos exões 5 e 5b, podendo originar uma forma longa ou duas formas curtas. Assim, o processo de *splicing* pode dar origem a várias isoformas e através de vários processos proteolíticos, pode originar várias formas longas e/ou várias formas curtas, podendo estas interagir de forma diferente na mitocôndria.

O estudo genético do presente trabalho foi realizado em 12 amostras de DNA de doentes com suspeita clínica de ADOA. O procedimento laboratorial incluiu PCR com *primers* específicos para os exões alternativos 4b e 5b, seguido de electroforese em gel

de agarose, purificação com ExoSAP, sequenciação pelo método de Sanger dos fragmentos purificados, purificação dos produtos de sequenciação por precipitação com etanol, electroforese capilar dos produtos de sequenciação purificados e análise dos resultados da sequenciação utilizando os programas Sequencing Analysis v5.4 e SeqScape® Software v2.5.

Após a análise dos exões alternativos 4b e 5b, e das respectivas regiões adjacentes, não foi identificada nenhuma variação de sequência. Deste modo, confirmou-se que estes exões apresentam raras variações de sequência, sendo extremamente conservados entre os primatas, assim como toda a proteína.

De modo a determinar a conservação evolutiva dos exões 4b e 5b do gene *OPA1* e da proteína Opa1, procedeu-se ao alinhamento das sequências de nucleótidos dos exões alternativos 4b e 5b, e dos aminoácidos constituintes da proteína Opa1 utilizando as sequências de várias espécies de primatas, vertebrados e invertebrados. Assim, tendo em conta os dados obtidos, foi possível confirmar que os exões em estudo deverão ter uma função primordial uma vez que após o estudo de 12 amostras de indivíduos com ADOA não foi encontrada uma única variação de sequência. A importância dessa função nos primatas é reforçada pela elevada homologia entre espécies. Apesar de não ser possível refutar a hipótese de estes exões alternativos não contribuírem para o fenótipo ADOA, a sua análise não deve ser descurada e deve ser parte integrante do estudo do gene *OPA1*.

Palavras-chave: ADOA, Gene OPA1, *Splicing*, Variação de sequência, Mitocôndria.

Abstract

Autosomal Dominant Optic Atrophy Kjer type (ADOA) is a hereditary optic neuropathy described for the first time in the end of the XIX century. ADOA is a rare disorder with a variable prevalence according to the geographical localization. In Denmark, it is possible to find a higher prevalence (1:10,000) than in the rest of the world. It is characterized by insidious vision loss, from moderate to severe, initiating in childhood. Syndromic ADOA may be associated with more severe features, such as ataxia, deafness, myopathy, neuropathy, among other.

OPA1 gene is located in the long arm of chromosome 3 (3q28-29), with more than 200 mutations identified in ADOA patients, 27% of them are missense, 27% are a form of alternative splicing, 23% are frameshift, 16.5% are nonsense and 6% are due to nucleotide or chromosomal rearrangements. This gene codes for 8 different protein isoforms due to alternative splicing, existing variation of its expression in the several tissues.

Several researchers consider that alternative exons 4b and 5b are not translated in ADOA phenotype responsible protein. However, it was possible to determine by functional studies that these two exons seem to be very important for Opa1 protein in mitochondria maintenance, beyond defining its 8 isoforms. This protein has 10 different domains: a Mitochondrial Import Sequence (MIS), three transmembrane (TM1, TM2a and TM2b), three coiled-coil (CC-0, CC-I and CC-II), a GTPase, a Middle and a GTPase effector domain (GED). The Opa1 protein has a cleavage point, where the MIS domain is removed after mitochondria importation, and possibly two other extra cleavage points located in exons 5 and 5b, which may originate a long or two short isoforms. Accordingly, the splicing process may give origin to several isoforms and through several proteolytic processes, may originate several long and/or short isoforms, which may present different ways for interacting in mitochondria.

This project genetic study was performed in 12 DNA samples from patients with ADOA clinical suspicion. The laboratorial procedure included PCR with specific primers for alternative exons 4b and 5b, followed with agarose gel electrophoresis, ExoSAP purification, Sanger sequencing of the purified fragments, purification of the sequencing products by ethanol precipitation, capillary electrophoresis of the purified

sequencing products and analysis of the sequencing results using the Sequencing Analysis v5.4 and the SeqScape® Software v2.5 programs.

After sequencing analysis of exons 4b and 5b of *OPA1* gene, it was not possible to find any variant in the coding exons neither in their surrounding regions. Accordingly, it was confirmed that these exons present rare sequence variations, being extremely conserved among primates, as well as the whole protein.

Aiming to determine the evolutionary conservation of *OPA1* gene exons 4b and 5b, and of Opa1 protein, alignment of alternative exons 4b and 5b nucleotide sequences, and complete Opa1 protein amino acids was performed using several primates, vertebrate and invertebrate species sequences. As so, taking into consideration the obtain data, it was possible to confirm that the studied exons must have an essential role since not a single sequence variation was found after the study of the 12 samples from ADOA patients. The significance of such function in primates is reinforced by the high homology between species. Despite it is not possible to disprove the hypothesis that these alternative exons do not contribute for the ADOA phenotype, its analysis should not be neglected and should be included in the study of *OPA1* gene.

Key words: .ADOA, *OPA1* Gene, Splicing, Sequence variation, Mitochondria

Agradecimentos: Quero agradecer ao Doutor Filipe Silva por me ter proposto este projecto, o qual aceitei, facultando-me todo o material necessário para a sua realização e ajudando-me com aspectos tanto práticos como teóricos, tal como a Professora Doutora Manuela Grazina. Para finalizar, também gostaria de agradecer à Professora Doutora Eugénia Cunha por ter aceite ser orientadora interna.

Este trabalho foi preparado para submissão no formato de artigo científico e de acordo com as Instruções para Autores da revista *Molecular Vision* (http://www.molvis.org/molvis/instructions.html) (Anexo A).

Autosomal Dominant Optic Atrophy Type 1: Study of 4b and 5b exons of OPA1 gene contribution

G. Ferreira[1,2,3], C. Ribeiro[3], A. Reis[4], J. Pratas[3], M. Castelo-Branco[4,5], E. Silva[5,6], M. Grazina[3,5], F. Silva[1,5]

1 Centre of Ophthalmology and Vision Sciences, IBILI, Faculty of Medicine, University of Coimbra, Coimbra, Portugal
2 Department of Life Sciences, Faculty of Sciences and Technology, University of Coimbra, Coimbra, Portugal
3 CNC- Centre for Neurosciences and Cell Biology- Biochemical Genetics Laboratory, Coimbra, Portugal
4 Visual Neuroscience Laboratory, Institute of Biomedical Research on Light and Image (IBILI), Faculty of medicine, University of Coimbra, Coimbra, Portugal
5 Faculty of Medicine, University of Coimbra, Coimbra, Portugal
6 Center for Hereditary Eye Diseases, Centro Hospitalar e Universitário de Coimbra E.P.E., Coimbra, Portugal

Corresponding author:
Filipe Silva
Centre of Ophthalmology and Vision Sciences, IBILI
Faculty of Medicine, University of Coimbra
Azinhaga Santa Comba, Celas
3000-548 Coimbra
Portugal
Email: fmpereiras2003@yahoo.com.br

Key words: .ADOA, *OPA1* Gene, Splicing, Sequence variation, Mitochondria

Abstract

Purpose: The purpose of this work was to analyze the exons 4b and 5b of *OPA1* gene, absent from the isoform NM_0155060.2, in a group of 12 Autosomal Dominant Optic Atrophy Kjer type (ADOA) suspected patients without any pathogenic mutation in the other 28 *OPA1* gene exons, in order to identify possible new mutations that may be responsible for the phenotype. Additionally, it was aimed to determine the degree of evolutionary conservation of the two alternative exons in a group of primates, and of the Opa1 protein in a group of primates and in a group of non-primate species.

Methods: Analysis of 12 DNA samples was conducted with the purpose of studying alternative exons 4b and 5b of the *OPA1* gene. The PCR fragments were amplified and separated on a 1% agarose gel and purified with ExoSAP-IT to remove the excess of ddNTP`s and primers. Sequencing reactions followed by ethanol precipitation and capillary electrophoresis allowed determining the nucleotide sequence of the PCR products that included exons 4b and 5b of *OPA1* gene. Analysis of the sequencing results was performed with programs Sequencing Analysis v5.4 and SeqScape® Software v2.5. Finally, in order to determine the evolutionary conservation of the exons 4b and 5b, and of the entire Opa1 protein, alignment of the nucleotide sequences from a group of primates, and Opa1 amino acids sequence from a group of primates and a group of non-primates, was manually performed.

Results: After sequencing analysis of exons 4b and 5b of *OPA1* gene, it was not possible to find any variant neither in the coding exons nor in their surrounding regions. It was also verified through evolutionary conservation studies that the exons 4b and 5b of *OPA1* gene, and the entire Opa1 protein has almost the same nucleotide and amino acid sequences among primates. However, the Opa1 protein is very different among vertebrate and invertebrate species, with the exception of a few regions of the protein where a perfect alignment could be identified. Finally, if it is considered only the alignment among mammals, it is possible to identify a high similarity between the amino acid sequences.

Conclusions: Surprisingly, no variants were identified in the exons 4b and 5b of *OPA1* gene, neither in their surrounding regions. This may allow concluding that a variant at these exons would influence the function or alter the structure of Opa1 protein very significantly and with a severe impact in the phenotype. Upon analysis of the evolutionary conservation of the alternative exons, it was possible to conclude that such exons are highly conserved among primates. The same conclusion was obtained for the evolutionary conservation analysis of the Opa1 protein among primates. However, the Opa1 protein it is not conserved among vertebrate and invertebrate species. Nevertheless, it is noteworthy that the studied mammals showed a considerable conservation of the Opa1 protein. These results suggest that *OPA1* gene and the corresponding protein have suffered major changes over the time until the evolutionary separation of the mammals, with further few changes until the separation of the primates, showing an extremely high similarity in sequence, and probably function, among the latest species.

Introduction

Hereditary optic atrophy refers to a heterogeneous group of genetic diseases; the most common disease of this group is the autosomal dominant optic atrophy (ADOA) [1]. It was described for the first time in the late nineteenth century [2,3]. The ADOA is a rare disease, with a prevalence that varies by geographic location. In Denmark, the prevalence is about 1:10000 [4], while in the rest of the world it may range between 1:35.000 [5] and 1:50.000 [6].

ADOA results from the degeneration of the optic nerve fibers, which causes decreased visual acuity, lack of color vision, centrocecal scotoma and pallor of the optic nerve [7,8]. The ADOA usually manifests itself as a non-syndromic disease, presenting a slow progression and it is rarely asymmetric [4,9]. In a non-syndromic form, the optic disc shows some bilateral and symmetrical pallor on the temporal side; in some cases, a kind of "excavation" similar to glaucoma may be observed, which may indicate the loss of retinal nerve fibers[9-11]. Another common feature is the presence of a centrocecal scotoma. Paracentral scotoma can happen but it is less frequent; in that situation, the central visual field is affected, while the peripheral field remains normal. In syndromic cases, there is a more severe visual deficit, hearing loss, myopathy, and peripheral neuropathy, among other features [8-11].

So far, *OPA1* is the only identified gene associated to ADOA, contributing to the phenotype in over 75% of the cases. The other genes and *loci* associated with Optic Atrophy have a minimal contribution of less than 1% [9].

The *OPA1* gene is localized in the long arm of the chromosome 3 (3q28-29), more specifically from base pair 193,310,932 until base pair 193,415,591 comprising 30 coding exons, 3 of which are alternative exons, and 1 non-coding exon [5,12,13].

The *OPA1* protein is expressed ubiquitously in the human tissues [14] and presents 10 different domains: GTPase, a Middle domain, a GTPase effector domain (GED), a Mitochondrial Import Sequence (MIS) domain in the N-terminal region, three transmembrane domains (TM1, TM2a and TM2b) and three coiled-coil domains (CC-0, CC-I and CC-II) [12].

The exons 4, 4b and 5b are alternatively involved in the process of splicing, and combinations of these exons give origin to 8 different mRNA transcripts [5,12,13]. Splicing is an important process to include or exclude these alternative exons,

originating different isoforms. This is a very complex process that involves more than 300 proteins, some of them binding to some *cis*-acting elements, the SRE (Splicing Regulatory Elements) [15-20], enabling recruiting other proteins of the spliceossome: U1, U2, U4/6 and U5 for the inclusion of the RNA exons and, subsequently, the formation of specific proteins [18,20-23]. The various isoforms are expressed in different tissues. The isoforms 1 (NM_015560.2) and 4 (NM_130833.2) have higher expression in the retina, fetal brain and heart while isoform 7 (NM_130836.2) has high expression in liver, kidney and colon. In all tissues, the isoforms 3 (NM_130832.2) and 2 (NM_130831.2) are weakly expressed and isoforms 5 (NM_130834.2) and 6 (NM_130835.2) have an intermediate expression [5]. The transcript precursors of the 8 mRNAs are directed to the mitochondria by the MIS domain, which is removed by the Mitochondrial processing peptidase (MPP) after importation, originating the long isoforms (l-*OPA1*). The Opa1 protein has two other points of cleavage, named S1 (exon 5) and S2 (exon 5b), which enable the proteases to also act in these specific points, giving origin to one or two shorter isoforms (s-*OPA1*). Both isoforms are bond to the mitochondria, however, the long isoforms are located in the inner mitochondrial membrane (IMM) while the shorter isoforms are in the periphery of the IMM, being able to diffuse to the inter-membrane space (IMS) and bind to the outer mitochondrial membrane (OMM) [24].

Some studies revealed that protein Opa1 plays an important role for the fusion of mitochondria [25-29]. Accordingly, the combination of a long and a short *OPA1* isoforms is important for mitochondrial fusion activity. Additionally, the isoforms 1 (*NM_015560.2*), 2 (*NM_130831.2*), 4 (*NM_130833.2*) and 7 (*NM_130836.2*), which lack the alternative exon 4b, promote a robust tubulation of mitochondria, while the isoforms 3 (*NM_130832.2*), 5 (*NM_130834.2*), 6 (*NM_130835.2*), and 8 (*NM_130837.2*), that express the alternative exon 4b, show less tubulation, which is correlated with mitochondrial fission activity [24]. Thus, the several Opa1 protein isoforms may have an important role in the maintenance of mitochondria and cell morphology, with a potential contribution for causing several diseases. Although many researchers do not pay much attention to the alternative exons of the *OPA1* gene, since they claim that they do not generate the ADOA phenotype [13].

Some studies show that the alternative splicing is more frequent in higher eukaryotes than in lower eukaryotes and the percentage of exons that undergo alternative splicing is higher in vertebrates than in invertebrates [30]. There are three

proposed mechanisms that suggest the emergence of new exons. Exon shuffling is the first mechanism. It occurs in 10% of the genes and arises when a constitutive exon is duplicated, being involved in the alternative splicing. The second explanatory mechanism consists in the transformation of the constitutive exons into alternative exons by cumulative mutation in the exons and introns, close to splice sites. Therefore, there is a "relaxation" of the splice site and may be no exon inclusion. The third explanatory mechanism is related to the exonization of some introns through Alu sequences. This mechanism may explain the appearance of 5% of alternative exons; however, it may have had major importance in human evolution. More than half of the human genome derives from transposable elements (TE), more specifically Alu sequences. These are short sequences with about 300 nucleotides, abundant in the introns that have several sites similar to the splice sites. Some mutations allow the spliceosome proteins to recognize these pseudo sites as an exon [30-33]. When these sequences within exons suffer some kind of mutation, they may form a new exon (exonization) which, in most cases, can be disadvantageous. Moreover, it can cause various diseases, but it can also give some benefits, for example, being implicated in the variability of transcripts and in the regulation of the gene [30,33,34]. However, Alu sequences can also be found in exons (intronization), and it may also be responsible for some exon modifications. Additionally, the Alu sequences can originate the non-inclusion of certain exons in the RNA, form a new constitutive exon, or “transform” a constitutive exon into an alternative exon, 5´ UTR and/or 3´ UTR. But sometimes a sequence can bind itself to an exon, rending it bigger or even separating an exon in two [33].

The first objective for the present work was to analyze the alternative exons 4b and 5b of OPA1 gene, absent from the isoform NM_0155060.2, in a group of 12 ADOA patients without any pathogenic mutation identified in the other 28 *OPA1* gene exons, in order to find possible new mutations that may be responsible for the phenotype. Another objective was to determine the degree of evolutionary conservation of the two alternative exons in a group of primates, and of the Opa1 protein in a group of primates and in a group of non-primate species.

Methods

The study followed the tenets of the Declaration of Helsinki and was approved by local Ethics Committee. Prior to the inclusion in the study, informed consent was obtained from all subjects after a full given explanation.

Twelve affected individuals from 12 families with ADOA were tested. The diagnosis was primarily based on classical clinical criteria. Average age of the patient's group was 36 ± 19 years, ranging from 11 to 73 years. Participants were submitted to a complete ophthalmological examination, including the best-corrected visual acuity (BCVA) slit-lamp examination of anterior chamber, IOP measurement (Goldmann applanation tonometer), angle and fundus examination (non-contact lens), cataract grading by the Lens Opacities Classification System II (LOCS), and the assessment of subjective visual complaints. No cases of high intraocular pressure or cataract were found. Exclusion criteria included pseudophakic and aphakic eyes, media opacities, retinal diseases, and high ammetropies (sphere > ± 4D; cylinder > ± 2D). Patients were explicitly checked for the presence of other neuro-ophthalmologic pathologies (besides ADOA).

OPA1 exon 4b and exon 5b genotyping was performed according to a method previously published [35]. Analysis of the obtained sequencing results was performed with programs Sequencing Analysis v5.4 (Applied Biosystems®) and SeqScape v2.5 (Applied Biosystems®) allowing their comparison with a reference sequence. Finally, using the nucleotide sequence data available in the websites Ensemble [36], NCBI [37] and UCSC [38], in order to determine the evolutionary conservation of the exons 4b and 5b, and of the entire Opa1 protein, alignment of the nucleotide sequences from a group of primates, and Opa1 amino acids sequence from a group of primates and a group of non-primates, was manually performed.

Results

After sequencing analysis of exons 4b and 5b of *OPA1* gene (Figure 1), it was not possible to find any variant neither in the coding exons nor in their surrounding regions.

It was also verified through evolutionary conservation studies that the exons 4b (Figure 2) and 5b (Figure 3) of *OPA1* gene, and the entire Opa1 protein present almost the same nucleotide and amino acid sequences among primates (Figure 4). Nevertheless, it is interesting to notice that in all the primates but Orangutan exists a single constitutive exon that codes for a specific amino acid sequence, while in the Orangutan the same amino acid sequence is coded by three different exons separated by introns (Figure 4). However, the function and the structure of the Opa1 protein in the Orangutan should not be significantly different of the other primates Opa1 proteins.

Moreover, the Opa1 protein is very different among vertebrate and invertebrate species (Figure 5), with the exception of a few regions of the protein where a perfect alignment could be identified. Such regions are mainly located at the C-terminus of the protein and give the functional similarity between the orthologous Opa1 proteins (Figure 5). On the opposite, the N-terminus of the protein shows a very low resemblance (Figure 5) indicating that each vertebrate and invertebrate species has specificity for Opa1 protein transport to mitochondria and isoform expression. Finally, if it is considered only the alignment among mammals, it is possible to identify a considerable similarity between the amino acid sequences, except for the N-terminus (Figure 5).

Discussion

The purpose of this work was to analyze the exons 4b and 5b of *OPA1* gene in a group of 12 Autosomal Dominant Optic Atrophy Kjer type patients in order to identify possible new mutations that may be responsible for the phenotype. After sequencing analysis of exons 4b and 5b of *OPA1* gene, it was not possible to find any variant neither in the coding exons nor in their surrounding regions. According to Delettre and collaborators [13], the truncating mutations occur between exon 17 and exon 28, 26% occur in the GTPase domain, 16% occur in the first seven exons and most often they occur in the three last exons of the *OPA1* gene. The missense mutations typically occur in the GTPase domain, the deletions and insertions originate a frameshift with a premature stop codon, the nonsense mutations occur throughout the gene and mutations due to splicing usually occur in the GTPase domain. None were found in the alternatively spliced exons 4, 4b and 5b [13]. Thus, mutations in the spliced exons 4, 4b and 5b may not generate a phenotype, since other isoforms can compensate for the mutated one [5,13]. However, special attention should be taken to exons 4b and 5b since they are very important for several reasons. The first reason refers to the splicing process, through which several isoforms that can be important in the cells maintenance are originated [5,12,24]; the second refers to the mtDNA cellular content that may be variable depending on the *OPA1* exons 4b and 5b expression [29]; and the third refers to the mitochondrial fusion and fission that may influence cells fate [25-28].

It is clear that the splicing process can be a very complex but useful mechanism that can originate several similar proteins. But, does the inclusion or exclusion of alternative exons originate proteins that can provide the cell with higher or lower stability? Three of the *OPA1* gene exons are alternatives, so its origin can be questioned according to the theories previously mentioned [30,31]. Focusing only in the two alternative exons that were sequenced in the 12 patients with ADOA, and considering that exon 4b has 54 bp while the exon 5b has 111 bp [5], the hypothesis of its origin happening through duplications of other exons can be excluded, because these are the smallest exons in the gene. Another hypothesis refers to the Alu sequences, by questioning if the origin of the 4b and 5b exons occurred due to the insertion of Alu sequences in the two introns that later suffered mutations. Although it is not possible to exclude that this might have happened, it is unlikely since the Alu sequences usually

have 300 nucleotides, as previously mentioned [30]. Recently, Gallus and collaborators studied an Italian family with ADOA and found an insertion of an Alu sequence with 327 bp in the intron 7, localized 27 pb from the exon 8, resulting in the non-inclusion of exon 8 in the *OPA1* mRNA and, consequently, in the Opa1 protein [39]. One might think that the non-inclusion of an exon would result in a phenotype more severe; however, it did not diverge a lot from an ADOA phenotype. The last hypothesis for the origin of the alternative exons of the gene, and the one that is more likely to occur, is that the exons 4b and 5b might have been constitutive exons that, after the ”relaxation” of the splice sites due to the cumulative mutations, have changed the expression from constitutive to alternative, with implications in the cellular homeostasis [31].

Maintenance and distribution of mtDNA are essential for energy metabolism, mitochondrial lineage in primordial germ cells, and to prevent mtDNA instability, which may lead to many debilitating diseases. Elachouri and colleagues [29], using siRNAs specific to *OPA1* alternative exons, evidenced that silencing of the *OPA1* variants leads to mtDNA depletion, secondary to inhibition of the mtDNA replication and to marked alteration of mtDNA distribution in nucleoid and nucleoid distribution throughout the mitochondrial network. Additionally, evidence that the abundance of exon 4b in *OPA1* transcripts correlates with mtDNA copy number, suggesting a common regulatory mechanism for both processes was also provided [29]. The role played by *OPA1* exon 4b interfacing the mitochondrial genome organization and its replication machinery with the membrane dynamics may also provide an explanatory basis for including the alternative exons in the *OPA1* mutation screening of ADOA patients.

Several studies indicate that Opa1 protein, along with other proteins, might be responsible for the fission/fusion of the mitochondria, possibly participating in the events that lead to apoptosis [24,28]. As previously stated, the expression of alternative exon 4b isoforms is highly correlated with fragmentation of mitochondria due to fission while the isoforms without exon 4b are strongly associated with mitochondria tubulation [24]. Previously, increased mitochondria fission has been associated to apoptosis while increased mitochondria fusion has been associated to cell survival [26,28]. Considering that exon 4b expression plays a relevant role in fission/fusion, this suggests that such exon is a key element for apoptosis/cell survival mechanisms, giving one more evidence for the significant role of Opa1 protein and *OPA1* alternative exons in the cellular homeostasis.

This work also aimed to determine the degree of evolutionary conservation of the two alternative exons in a group of primates, and of the Opa1 protein in a group of primates and in a group of non-primate species. As so, it was verified through evolutionary conservation studies that the exons 4b and 5b of *OPA1* gene, and the entire Opa1 protein present almost the same nucleotide and amino acid sequences among primates, allowing to conclude that such exons and protein are highly conserved among primates. It is very interesting that the only difference in the exons 4b and 5b nucleotide sequence was identified in *Pan traglodytes* exon 5b, where a G was present instead of an A at position 726 of the coding sequence (Figure 3). Consequently, in this species protein, an amino acid Arginine is present instead of Histamine at position 242 (Figure 4). Despite this amino acid difference, it does not seem to influence significantly the protein structure since both are polar amino acids and have similar biochemical characteristics. For the same reason, it seems that this amino acid difference has no influence in the TM2b domain, neither in the S2 proteolytic cleavage site (amino acids 217-223), two very important features for the Opa1 protein post-translational processing that require the presence of alternative exon 5b coded amino acids sequence.

Furthermore, it is also interesting that a single human constitutive exon that encodes a specific sequence of amino acids is split in three exons, with exactly the same amino acid sequence but one, in orangutan. This suggests that, in the process of orangutan differentiation from the other primates, the constitutive exon could have been divided in three or maybe two small introns could have been inserted.

The evolutionary conservation analysis of the Opa1 protein in a group of vertebrates and invertebrates species allowed concluding that this protein is not conserved among these species (Figure 2). This is more obvious if it is taken into consideration the *Drosophila melanogaster* and the *Caenorhabditis elegans* species in which the homology with the human Opa1 protein is very low (Figure 5). However, this is not surprising since it refers to two species evolutionarily very distant from humans, even considering that they share a considerable homology in terms of proteome. Furthermore, a special attention must be given to the organ that is mainly affected in ADOA, the eye. The eye is one of the most complex structures in nature, and its design presents a huge variation across the animal kingdom. The *Drosophila melanogaster* compound eye, made of 800 linked mini-eyes, looks nothing like the single-lensed human eye, while the tiny eyeless *Caenorhabditis elegans* is only able to detect flashes of light and responds to them by quickly wriggling away. Considering this, it is relevant

to question which has been the contribution of Opa1 protein and, more specifically, of the two alternative exons 4b and 5b for the development of the eye in the vertebrate species that is significantly different of the organs with similar function in the invertebrates. Certainly, this is a possible research interest for the future. Nevertheless, it is noteworthy that the studied mammals showed a considerable conservation of Opa1 protein, suggesting that *OPA1* gene and protein have suffered major changes over the time until the evolutionary separation of the mammals, with further few changes until the separation of the primates, showing an extremely similar sequence and function among the latest species.

At this point it is necessary to mention that the alignment of the *OPA1* gene alternative exons 4b and 5b was also attempted for the group of vertebrates and invertebrates species. However, it was not possible to be performed because such sequences were not found in the Ensembl [36], NCBI [37] and UCSC [38] databases. This may be because exons 4b and 5b are not present in the genome of those species or because the transcripts and the isoforms with those exons were not investigated and/or identified yet, which does not mean that they do not exist.

In conclusion, even considering that it was not possible to refute the hypothesis that *OPA1* alternative exons 4b and 5b do not contribute for ADOA phenotype, it is very clear that the analysis of such exons should not be neglected and it should be mandatory to include in every mutation screening of *OPA1* gene.

Acknowledgments

We are grateful to the patients that kindly agreed to participate in this study. FS is a Post-Doc grant holder (SFRH_BPD_71016_2010) from FCT. This work was partially supported by FCT project PEST-C/SAU/LA0001/2011.

References

1. Ayrignac X, Liauzun C, Lenaers G, Renard D, Amati-Bonneau P, de Sèze J, Dollfus H, Hamel C, Bonneau D, Labauge P. OPA3-related autosomal dominant optic atrophy and cataract with ataxia and areflexia. *Eur Neurol* 2012; 68: 108-10.

2. Batten B. A family suffering from hereditary optic atrophy. *Trans Ophthalmol Soc UK* 1896; 1: 125.

3. Snell S. Disease of the optic nerve I. hereditary or congenital optic atrophy and allied cases. *Trans Ophthalmol Soc UK* 1897; 17: 66-81.

4. Kjer P. Infantile optic atrophy with dominant mode of inheritance: a clinical and genetic study of 19 Danish families. *Acta Ophthalmol Suppl* 1959; 164: 1-147.

5. Delettre C, Griffoin JM, Kaplan J, Dollfus H, Lorenz B, Faivre L, Lenaers G, Belenguer P, Hamel CP. Mutation spectrum and splicing variants in the OPA1 gene. *Hum Genet* 2001; 109: 584–91.

6. Yu-Wai-Man P, Griffiths PG, Burke A, Sellar PW, Clarke MP, Gnanaraj L, Ah-Kine D, Hudson G, Czermin B, Taylor RW, Horvath R, Chinnery PF. The prevalence and natural history of dominant optic atrophy due to OPA1 mutations. *Ophthalmology* 2010; 117: 1538-46.

7. Chevrollier A, Guillet V, Loiseau D, Gueguen N, de Crescenzo MA, Verny C, Ferre M, Dollfus H, Odent S, Milea D, Goizet C, Amati-Bonneau P, Procaccio V, Bonneau D, Reynier P. Hereditary optic neuropathies share a common mitochondrial coupling defect. *Annals of Neurology* 2008; 63: 794-8.

8. Votruba M, Moore AT, Bhattacharya SS. Demonstration of a founder effect and fine mapping of dominant optic atrophy locus on 3q28-qter by linkage disequilibrium method: a study of 38 British Isles pedigrees. *Hum Genet* 1998; 102: 79-86.

9. Lenaers G, Hamel C, Delettre C, Amati-Bonneau P, Procaccio V, Bonneau D, Reynier P, Milea D. Dominant optic atrophy. *Orphanet Journal of Rare Diseases* 2012; 7: 46.

10. Cohn AC, Toomes C, Potter C, Towns KV, Hewitt AW, Inglehearn CF, Craig JE, Mackey DA. Autosomal dominant optic atrophy: penetrance and expressivity in patients with OPA1 mutations. *Am J Ophthalmol* 2007; 143: 656-62.

11. Ferre M, Bonneau D, Milea D, Chevrollier A, Verny C, Dollfus H, Ayus C, Defoort S, Vignal C, Zanlonghi X. Molecular screening of 980 cases of suspected hereditary optic neuropathy with a report on 77 Novel OPA1 mutations. *Human Mutation* 2009; 30: 692-705.

12. Landes T, Leroy I, Bertholet A, Diot A, Khosrobakhsh F, Daloyau M, Davezac N, Miquel MC, Courilleau D, Guillou E, Olichon A, Lenaers G, Arnauné-Pelloquin L, Emorine LJ, Belenguer P. OPA1 (dys)functions. *Semin Cell Dev Biol* 2010; 21: 593-8.

13. Delettre C, Lenaers G, Pelloquin L, Belenguer P, Hamel CP. OPA1 (Kjer type) dominant optic atrophy: a novel mitochondrial disease. *Mol Genet Metab* 2002; 75: 97-107.

14. Delettre C, Lenaers G, Griffoin JM, Gigarel N, Lorenzo C, Belenguer P, Pelloquin L, Grosgeorge J, Turc-Carel C, Perret E, Astarie-Dequeker C, Lasquellec L, Arnaud B, Ducommun B, Kaplan J, Hamel CP. Nuclear gene OPA1, encoding a mitochondrial dynamin-related protein, is mutated in dominant optic atrophy. *Nat Genet* 2000; 26: 207-10.

15. Liu W, Zhou Y, Hu Z, Sun T, Denise A, Fu XD, Zhang Y. Regulation of splicing enhancer activities by RNA secondary structures. *FEBS Lett* 2010; 584: 4401-7.

16. Cyphert TJ, Suchanek AL, Griffith BN, Salati LM. Starvation actively inhibits splicing of glucose-6-phosphate dehydrogenase mRNA via a bifunctional ESE/ESS element bound by hnRNP K. *Biochim BiophysActa* 2013; 1829: 905-15.

17. Blencowe BJ. Exonic splicing enhancers: mechanism of action, diversity and role 643 in human genetic diseases. *Trends Biochem Sci* 2000; 25: 106-10.

18. Shepard PJ, Hertel KJ. The SR protein family. *Genome Biol* 2009; 10: 242.1-9

19. Long JC, Caceres JF. The SR protein family of splicing factors: master regulators of gene expression. *Biochem J* 2009; 417: 15–27.

20. Graveley BR. Sorting out the complexity of SR protein functions. *Rna* 2000; 6: 1197–211.

21. Mayeda A, Screaton GR, Chandler SD, Fu XD, Krainer AR. Substrate specificities of SR proteins in constitutive splicing are determined by their RNA recognition motifs and composite pre-mRNA exonic elements. *Molecular and Cellular Biology* 1999; 19: 1853–63.

22. McCullough AJ, Berget SM. An intronic splicing enhancer binds U1 snRNPs to enhance splicing and select 5' splice sites. *Mol Cell Biol* 2000; 20: 9225-35.

23. Caceres JF, Kornblihtt AR. Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet* 2002; 18:186–193.

24. Song Z, Chen H, Fiket M, Alexander C, Chan DC. OPA1 processing controls mitochondrial fusion and is regulated by mRNA splicing, membrane potential, and Yme1L. *J Cell Biol* 2007, 178: 749-55.

25. Luo LF, Hou CC, Yang WX. Nuclear factors: roles related to mitochondrial deafness. *Gene* 2013; 520: 79-89.

26. Chan DC. Mitochondrial fusion and fission in Mammals. *Annu Rev Cell Dev Biol* 2006; 22: 79-99.

27. Chan DC. Fusion and fission interlinked processes critical for mitochondrial health. *Annu Rev Genet* 2012; 46: 265-87.

28. Chen H, Chan DC. Emerging functions of mammalian mitochondrial fusion and fission. *Hum Mol Genet* 2005; 14: 1-7.

29. Elachouri G, Vidoni S, Zanna C, Pattyn A, Boukhaddaoui H, Gaget K, Yu-Wai-Man P, Gasparre G, Sarzi E, Delettre C, Olichon A, Loiseau D, Reynier P, Chinnery PF, Rotig A, Carelli V, Hamel CP, Rugolo M, Lenaers G. OPA1 links human mitochondrial genome maintenance to mtDNA replication and distribution. *Genome Res* 2011; 21: 12-20.

30. Keren H, Lev-Maor G, Ast G. Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet* 2010; 11: 345-55.

31. Kim E, Goren A, Ast G. Alternative splicing: current perspectives. *BioEssays* 2008; 30: 38-47.

32. Sorek R, Ast G, Graur D. Alu-containing exons are alternatively spliced. *Genome Res* 2002; 12: 1060-7.

33. Sela N, Mersch B, Gal-Mark N, Lev-Maor G, Hotz-Wagenblatt A, Ast G. Comparative analysis of transposed element insertion within human and mouse genes reveals Alu's unique role in shaping the human transcripome. *Genome Biol* 2007; 8: 1-15.

34. Silke SS, Männel DN, Hehlgans T, Brosius J, Schmitz J. From "junk" to gene: curriculum vitae of a primate receptor isoform gene. *J Mol Biol* 2004; 4: 883-6.

35. Sousa T. Genomic Cross-Talk Research in Leber's Hereditary Optic Neuropathy. Thesis for Cellular and Molecular Biology Master degree in the Department of Life Sciences, Faculty of Sciences and Technology, University of Coimbra, 2012.

36. Flicek P, Ahmed I, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, Gil L, García-Girón C, Gordon L, Hourlier T, Hunt S, Juettemann T, Kähäri AK, Keenan S, Komorowska M, Kulesha E, Londen I, Maurel T, McLaren WM, Muffato M, Nag R, Overduin B, Pignatelli M, Pritchard B, Pritchard E, Riat HS, Ritchie GRS, Ruffier M, Schuster M, Sheppard D, Sobral D, Taylor K, Thormann A, Trevanion S, White S, Wilder SP, Aken BL, Birney E, Cuningham F, Dunham I, Harrow J, Herrero J, Hubbard TJP, Johnson N, Kinsella R, Parker A, Spudich G, Yates A, Zadissa A and Searle SMJ. Ensembl 2013. *Nucleic Acids Res* 2013; 41: D48-55.

37. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J and Savers EW. GenBank. *Nucleic Acids Res* 2013; 41: D36-42.

38. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. The human genome browser at UCSC. *Genome Res* 2002; 12: 996-1006.

39. Gallus GN, Cardaioli E, Rufa A, Da Pozzo P, Bianchi S, D'Eramo C, Collura M, Tumino M, Pavone L, Federico A. Alu-element insertion in an OPA1 intron sequence associated with autosomal dominant optic atrophy. *Mol Vis* 2010; 16: 178-83.

Legends of figures

Figure 1 - Electropherogram from sequencing analysis of *OPA1* exon 5b. Note that all the nucleotide sequence matches the normal reference sequence.

Figure 2 - Evolutionary conservation of *OPA1* exon 4b among primates. Note that *OPA1* exon 4b is highly conserved among primates.

Figure 3 - Evolutionary conservation of *OPA1* exon 5b among primates. Note that *OPA1* exon 5b is very conserved among primates.

Figure 4 - Evolutionary conservation of Opa1 protein among primates. Exons are represented in alternating black and blue letters. Red letters represent amino acids coded by two consecutive exons. Amino acids with yellow shading are non-conserved. Note that Opa1 protein is very conserved among primates.

Figure 5 - Evolutionary conservation of Opa1 protein among vertebrates and invertebrates species. Exons are represented in alternating black and blue letters. Red letters represent amino acids coded by two consecutive exons. Note that only specific amino acids sequences are conserved among vertebrates and invertebrates species.

Figures

Figure 1

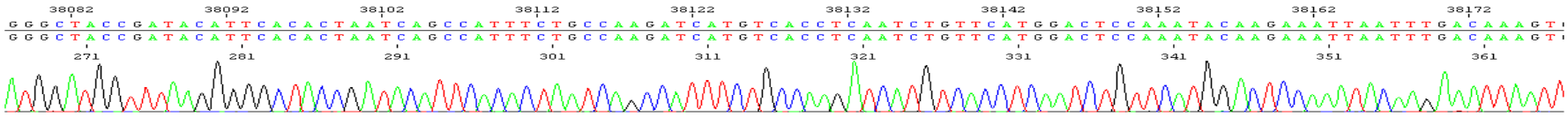


Figure 2

```
H. sapiens          GTCACAAATTGGTTAGTGAAGTCATAGGAGCTTCTGACCTACTTCTCTTGTTAG
G. gorilla          GTCACAAATTGGTTAGTGAAGTCATAGGAGCTTCTGACCTACTTCTCTTGTTAG
P. troglodytes      GTCACAAATTGGTTAGTGAAGTCATAGGAGCTTCTGACCTACTTCTCTTGTTAG
P. abelii           GTCACAAATTGGTTAGTGAAGTCATAGGAGCTTCTGACCTACTTCTCTTGTTAG
M. mulatta          GTCACAAATTGGTTAGTGAAGTCATAGGAGCTTCTGACCTACTTCTCTTGTTAG
```

Figure 3

```
H. sapiens          GGTCTGCTTGGTGAGCTCATTCTCTTACAACAACAAATTCAAGAGCATGAAGAGGAAGCGCGCAGAGCCGCTGGCCAATATAGCACGAGCTATGCCCAACAGA
G. gorilla          GGTCTGCTTGGTGAGCTCATTCTCTTACAACAACAAATTCAAGAGCATGAAGAGGAAGCGCGCAGAGCCGCTGGCCAATATAGCACGAGCTATGCCCAACAGA
P. troglodytes      GGTCTGCTTGGTGAGCTCATTCTCTTACAACAACAAATTCAAGAGCGTGAAGAGGAAGCGCGCAGAGCCGCTGGCCAATATAGCACGAGCTATGCCCAACAGA
P. abelii           GGTCTGCTTGGTGAGCTCATTCTCTTACAACAACAAATTCAAGAGCATGAAGAGGAAGCGCGCAGAGCCGCTGGCCAATATAGCACGAGCTATGCCCAACAGA
M. mulatta          GGTCTGCTTGGTGAGCTCATTCTCTTACAACAACAAATTCAAGAGCATGAAGAGGAAGCGCGCAGAGCCGCTGGCCAATATAGCACGAGCTATGCCCAACAGA


H. sapiens          AGCGCAAG
G. gorilla          AGCGCAAG
P. troglodytes      AGCGCAAG
P. abelii           AGCGCAAG
M. mulatta          AGCGCAAG
```

Figure 4


H. sapiens MWRLRRAAVACEVCQSLVKHSSGIKGSLPLQKLHLVSRSIYHSHHPTLKLQRPQLRTSFQQFSSLTNLPLRKLKFSPIKYGYQPRRNFWPARLATRLLKLRYLIL
G. gorilla MWRLRRAAVACEVCQSLVKHSSGIKGSLPLQKLHLVSRSIYHSHHPTLKLQRPQLRTSFQQFSSLTNLPLRKLKVSPIKYGYQPRRNFWPARLATRLLKLRYLIL
P. troglodytes MWRLRRAAVACEVCQSLVKHSSGIKGSLPLQKLHLVSRSIYHSHHPTLKLQRPQLRTSFQQFSSLTNLPLRKLKFSPIKYGYQPRRNFWPARLATRLLKLRYLIL
P. abelii MWRLRRAAVACEVCQSLVKHSSGIKGSLPLQKLHLVSRSIYHSHYPTLKLQRPQLRTSFQQFSSLTNLPLRKLKFSPIKYGYQPRRNFWPARLATRLLKLRYLIL
M. mulatta MWRLRRAAVACEVCQSLVKHSSGIKGSLPLQKLHLVSRSIYHSHHPTLKLQRPQLRTSFQQFSSLTNLPLRKLKFSPIKYGYQPRRNFWPARLATRLLKLRYLIL
H. sapiens GSAVGGGYTAKKTFDQWKDMIPDLSEYKWIVPDIVWEIDEYIDFEKIRKALPSSEDLVKLAPDFDKIVESLSLLKDFFTSGHKLVSEVIGASDLLLLLGSPEETA
G. gorilla GSAVGGGYTAKKTFDQWKDMIPDLSEYKWIVPDIVWEIDEYIDFEKIRKALPNSEDLVKLAPDFDKIVESLSLLKDFFTSGHKLVSEVIGASDLLLLLGSPGETA
P. troglodytes GSAVGGGYTAKKTFDQWKDMIPDLSEYKWIVPDIVWEIDEYIDFEKIRKALPNSEDLVKLAPDFDKIVESLSLLKDFFTSGHKLVSEVIGASDLLLLLGSPGETA
P. abelii GSAVGGGYTAKKTFDQWKDMIPDLSEYKWIVPDIVWEIDEYIDFEKIRKALPNSEDLVKLAPDFDKIVESLSLLKDFFTSGHKLVSEVIGASDLLLLLGSPGETA
M. mulatta GSAVGGGYTAKKTFDQWKDMIPDLSEYKWIVPDIVWEIDEYIDFEKIRKALPNSEDLVKLAPDFDKIVESLNLLKDFFTSGHKLVSEVIGASDLLLLLGSPGETA
H. sapiens FRATDRGSESDKHFRKGLLGELILLQQQIQEHEEEARRAAGQYSTSYAQQKRKVSDKEKIDQLQEELLHTQLKYQRILERLEKENKELRKLVLQKDDKGIHHRKL
G. gorilla FRATDHGSESDKHFRKGLLGELILLQQQIQEHEEEARRAAGQYSTSYAQQKRKVSDKEKIDQLQEELLHTQLKYQRILERLEKENKELRKLVLQKDDKGIHHRKL
P. troglodytes FRATDHGSESDKHFRKGLLGELILLQQQIQEREEEARRAAGQYSTSYAQQKRKVSDKEKIDQLQEELLHTQLKYQRILERLEKENKELRKLVLQKDDKGIHHRKL
P. abelii FRATDHGSESDKHFRKGLLGELILLQQQIQEHEEEARRAAGQYSTSYAQQKRKVSDKEKIDQLQEELLHTQLKYQRILERLEKENKELRKLVLQKDDKGIHHRKL
M. mulatta FRATDHGSESDKHFRKGLLGELILLQQQIQEHEEEARRAAGQYSTSYAQQKRKVSDKEKIDQLQEELLHTQLKYQRILERLEKENKELRKLVLQKDDKGIHHRKL
H. sapiens KKSLIDMYSEVLDVLSDYDASYNTQDHLPRVVVVGDQSAGKTSVLEMIAQARIFPRGSGEMMTRSPVKVTLSEGPHHVALFKDSSREFDLTKEEDLAALRHEIEL
G. gorilla KKSLIDMYSEVLDVLSDYDASYNTQDHLPRVVVVGDQSAGKTSVLEMIAQARIFPRGSGEMMTRSPVKVTLSEGPHHVALFKDSSREFDLTKEEDLAALRHEIEL
P. troglodytes KKSLIDMYSEVLDVLSDYDASYNTQDHLPRVVVVGDQSAGKTSVLEMIAQARIFPRGSGEMMTRSPVKVTLSEGPHHVALFKDSSREFDLTKEEDLAALRHEIEL
P. abelii KKSLIDMYSEVLDVLSDYDASYNTQDHLPRVVVVGDQSAGKTSVLEMIAQARIFPRGSGEMMTRSPVKVTLSEGPHHVALFKDSSREFDLTKEEDLAALRHEIEL
M. mulatta KKSLIDMYSEVLDVLSDYDASYNTQDHLPRVVVVGDQSAGKTSVLEMIAQARIFPRGSGEMMTRSPVKVTLSEGPHHVALFKDSSREFDLTKEEDLAALRHEIEL

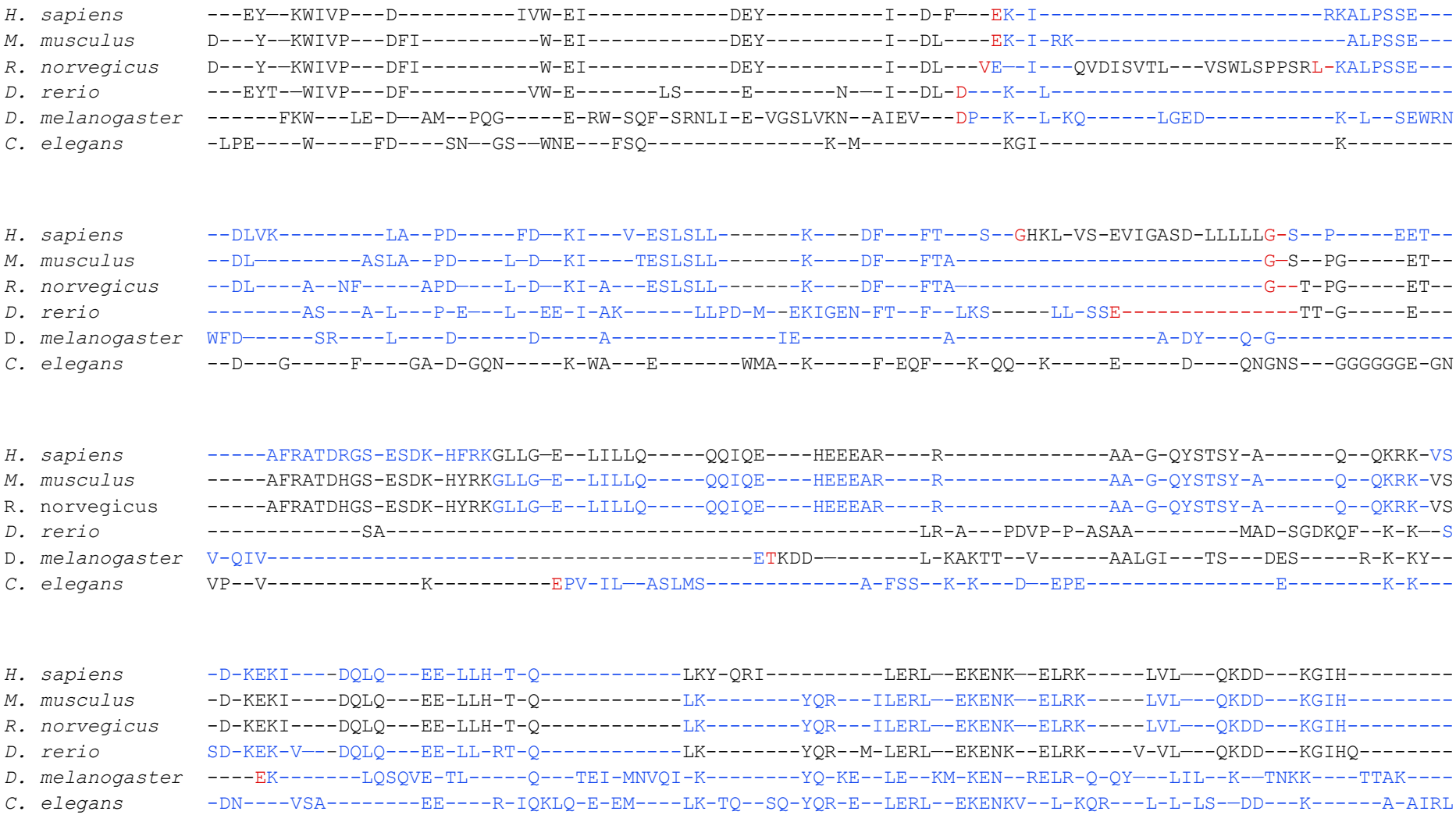
```
H. sapiens        RMRKNVKEGCTVSPETISLNVKGPGLQRMVLVDLPGVINTVTSGMAPDTKETIFSISKAYMQNPNAIILCIQDGSVDAERSIVTDLVSQMDPHGRRTIFVLTKVD
G. gorilla        RMRKNVKEGCTVSPETISLNVKGPGLQRMVLVDLPGVINTVTSGMAPDTKETIFSISKAYMQNPNAIILCIQDGSVDAERSIVTDLVSQMDPHGRRTIFVLTKVD
P. troglodytes    RMRKNVKEGCTVSPETISLNVKGPGLQRMVLVDLPGVINTVTSGMAPDTKETIFSISKAYMQNPNAIILCIQDGSVDAERSIVTDLVSQMDPHGRRTIFVLTKVD
P. abelii         RMRKNVKEGCTVSPETISLNVKGPGLQRMVLVDLPGVINTVTSGMAPDTKETIFSISKAYMQNPNAIILCIQDGSVDAERSIVTDLVSQMDPHGRRTIFVLTKVD
M. mulatta        RMRKNVKEGCTVSPETISLNVKGPGLQRMVLVDLPGVINTVTSGMAPDTKETIFSISKAYMQNPNAIILCIQDGSVDAERSIVTDLVSQMDPHGRRTIFVLTKVD

H sapiens         LAEKNVASPSRIQQIIEGKLFPMKALGYFAVVTGKGNSSESIEAIREYEEEFFQNSKLLKTSMLKAHQVTTRNLSLAVSDCFWKMVRESVEQQADSFKATRFNLE
G. gorilla        LAEKNVASPSRIQQIIEGKLFPMKALGYFAVVTGKGNSSESIEAIREYEEEFFQNSKLLKTSMLKAHQVTTRNLSLAVSDCFWKMVRESVEQQADSFKATRFNLE
P. troglodytes    LAEKNVASPSRIQQIIEGKLFPMKALGYFAVVTGKGNSSESIEAIREYEEEFFQNSKLLKTSMLKAHQVTTRNLSLAVSDCFWKMVRESVEQQADSFKATRFNLE
P. abelii         LAEKNVASPSRIQQIIEGKLFPMKALGYFAVVTGKGNSSESIEAIREYEEEFFQNSKLLKTSMLKAHQVTTRNLSLAVSDCFWKMVRESVEQQADSFKATRFNLE
M. mulatta        LAEKNVASPSRIQQIIEGKLFPMKALGYFAVVTGKGNSSESIEAIREYEEEFFQNSKLLKTSMLKAHQVTTRNLSLAVSDCFWKMVRESVEQQADSFKATRFNLE

H. sapiens        TEWKNNYPRLRELDRNELFEKAKNEILDEVISLSQVTPKHWEEILQQSLWERVSTHVIENIYLPAAQTMNSGTFNTTVDIKLKQWTDKQLPNKAVEVAWETLQEE
G. gorilla        TEWKNNYPRLRELDRNELFEKAKNEILDEVISLSQVTPKHWEEILQQSLWERVSTHVIENIYLPAAQTMNSGTFNTTVDIKLKQWTDKQLPNKAVEVAWETLQEE
P. troglodytes    TEWKNNYPRLRELDRNELFEKAKNEILDEVISLSQVTPKHWEEILQQSLWERVSTHVIENIYLPAAQTMNSGTFNTTVDIKLKQWTDKQLPNKAVEVAWETLQEE
P. abelii         TEWKNNYPRLRELDRNELFEKAKNEILDEVISLSQVTPKHWEEILQQSLWERVSTHVIENIYLPAAQTMNSGTFNTTVDIKLKQWTDKQLPNKAVEVAWETLQEE
M. mulatta        TEWKNNYPRLRELDRNELFEKAKNEILDEVISLSQVTPKHWEEILQQSLWERVSTHVIENIYLPAAQTMNSGTFNTTVDIKLKQWTDKQLPNKAVEVAWETLQEE

H. sapiens        FSRFMTEPKGKEHDDIFDKLKEAVKEESIKRHKWNDFAEDSLRVIQHNALEDRSISDKQQWDAAIYFMEEALQARLKDTENAIENMVGPDWKKRWLYWKNRTQEQ
G. gorilla        FSRFMTEPKGKEHDDIFDKLKEAVKEESIKRHKWNDFAEDSLRVIQHNALEDRSISDKQQWDAAIYFMEEALQARLKDTENAIENMVGPDWKKRWLYWKNRTQEQ
P. troglodytes    FSRFMTEPKGKEHDDIFDKLKEAVKEESIKRHKWNDFAEDSLRVIQHNALEDRSISDKQQWDAAIYFMEEALQARLKDTENAIENMVGPDWKKRWIYWKNRTQEQ
P. abelii         FSRFMTEPKGKEHDDIFDKLKEAVKEESIKRHKWNDFAEDSLRVIQHNALEDRSISDKQQWDAAIYFMEEALQARLKDTENAIENMVGPDWKKRWLYWKNRTQEQ
M. mulatta        FSRFMTEPKGKEHDDIFDKLKEAVKEESIKRHKWNDFAEDSLRVIQHNALEDRSISDKQQWDAAIYFMEEALQARLKDTENAIENMVGPDWKKRWLYWKNRTQEQ
```

```
H sapiens         CVHNETKNELEKMLKCNEEHPAYLASDEITTVRKNLESRGVEVDPSLIKDTWHQVYRRHFLKTA----LNHCNLCRRGFYYYQRHFVDSELECNDVVLFWRIQRM
G. gorilla        CVHNETKNELEKMLKCNEEHPAYLASDEITTVRKNLESRGVEVDPSLIKDTWHQVYRRHFLKTA----LNHCNLCRRGFYYYQRHFVDSELECNDVVLFWRIQRM
P. troglodytes    CVHNETKNELEKMLKCNEEHPAYLASDEITTVRKNLESRGVEVDPSLIKDTWHQVYRRHFLKTA----LNHCNLCRRGFYYYQRHFVDSELECNDVVLFWRIQRM
P. abelii         CVHNETKNELEKMLKCNEEHPAYLASDEITTVRKNLESRGVEVDPSLIKDTWHQVYRRHF-KTALNHSLNHCNLCRRGFY-YQRHFVDSELECNDVVLFWRIQRM
M. mulatta        CVHNETKNELEKMLKCNEEHPAYLASDEITTVRKNLESRGVEVDPSLIKDTWHQVYRRHFLKTA----LNHCNLCRRGFYYYQRHFVDSELECNDVVLFWRIQRM
```

```
H. sapiens        LAITANTLRQQLTNTEVRRLEKNVKEVLEDFAEDGEKKIKLLTGKRVQLAEDLKKVREIQEKLDAFIEALHQEK
G. gorilla        LAITANTLRQQLTNTEVRRLEKNVKEVLEDFAEDGEKKIKLLTGKRVQLAEDLKKVREIQEKLDAFIEALHQEK
P. troglodytes    LAITANTLRQQLTNTEVRRLEKNVKEVLEDFAEDGEKKIKLLTGKRVQLAEDLKKVREIQEKLDAFIEALHQEK
P. abelii         LAITANTLRQQLTNTEVRRLEKNVKEVLEDFAEDGEKKIKLLTGKRVQLAEDLKKVREIQEKLDAFIEALHQEK
M. mulatta        LAITANTLRQQLTNTEVRRLEKNVKEVLEDFAEDSEKKIKLLTGKRVQLAEDLKKVREIQEKLDAFIEALHQEK
```

Figure 5


H. sapiens
M. musculus
R. norvegicus
D. rerio
D. melanogaster
C. elegans


H. sapiens ---EY--KWIVP---D----------IVW-EI----------DEY----------I--D-F---EK-I----------RKALPSSE---
M. musculus D---Y--KWIVP---DFI----------W-EI----------DEY----------I--DL----EK-I-RK----------ALPSSE---
R. norvegicus D---Y--KWIVP---DFI----------W-EI----------DEY----------I--DL---VE--I---QVDISVTL---VSWLSPPSRL-KALPSSE---
D. rerio ---EYT--WIVP---DF----------VW-E-------LS-----E-------N---I--DL-D---K--L----------
D. melanogaster ------FKW---LE-D--AM--PQG-----E-RW-SQF-SRNLI-E-VGSLVKN--AIEV---DP--K--L-KQ------LGED----------K-L--SEWRN
C. elegans -LPE----W-----FD----SN--GS--WNE---FSQ----------K-M----------KGI----------K---------
H. sapiens --DLVK---------LA--PD-----FD--KI---V-ESLSLL-------K----DF---FT---S--GHKL-VS-EVIGASD-LLLLLG-S--P-----EET--
M. musculus --DL---------ASLA--PD----L-D--KI----TESLSLL-------K----DF---FTA----------G--S--PG-----ET--
R. norvegicus --DL----A--NF-----APD----L-D--KI-A---ESLSLL-------K----DF---FTA----------G--T-PG-----ET--
D. rerio --------AS---A-L---P-E---L--EE-I-AK------LLPD-M--EKIGEN-FT--F--LKS-----LL-SSE----------TT-G-----E---
D. melanogaster WFD------SR----L----D------D-----A----------IE----------A----------A-DY---Q-G----------
C. elegans --D---G-----F----GA-D-GQN-----K-WA---E-------WMA--K-----F-EQF---K-QQ--K-----E-----D----QNGNS---GGGGGGE-GN
H. sapiens -----AFRATDRGS-ESDK-HFRKGLLG--E--LILLQ-----QQIQE----HEEEAR----R----------AA-G-QYSTSY-A------Q--QKRK-VS
M. musculus -----AFRATDHGS-ESDK-HYRKGLLG--E--LILLQ-----QQIQE----HEEEAR----R----------AA-G-QYSTSY-A------Q--QKRK-VS
R. norvegicus -----AFRATDHGS-ESDK-HYRKGLLG--E--LILLQ-----QQIQE----HEEEAR----R----------AA-G-QYSTSY-A------Q--QKRK-VS
D. rerio ----------SA----------LR-A---PDVP-P-ASAA---------MAD-SGDKQF--K-K--S
D. melanogaster V-QIV----------ETKDD---------L-KAKTT--V------AALGI---TS---DES-----R-K-KY--
C. elegans VP--V----------K----------EPV-IL--ASLMS----------A-FSS--K-K---D--EPE----------E--------K-K---
H. sapiens -D-KEKI----DQLQ---EE-LLH-T-Q----------LKY-QRI----------LERL--EKENK--ELRK-----LVL---QKDD---KGIH---------
M. musculus -D-KEKI----DQLQ---EE-LLH-T-Q----------LK--------YQR---ILERL--EKENK--ELRK-----LVL---QKDD---KGIH---------
R. norvegicus -D-KEKI----DQLQ---EE-LLH-T-Q----------LK--------YQR---ILERL--EKENK--ELRK-----LVL---QKDD---KGIH---------
D. rerio SD-KEK-V---DQLQ---EE-LL-RT-Q----------LK--------YQR--M-LERL--EKENK--ELRK----V-VL---QKDD---KGIHQ--------
D. melanogaster ----EK-------LQSQVE-TL-----Q---TEI-MNVQI-K--------YQ-KE--LE--KM-KEN--RELR-Q-QY---LIL--K--TNKK----TTAK----
C. elegans -DN----VSA--------EE----R-IQKLQ-E-EM----LK-TQ--SQ-YQR-E--LERL--EKENKV--L-KQR---L-L-LS--DD---K------A-AIRL

```
H. sapiens        HRK--L-KK--SLIDMYSEVLD-V-LSD----YDAS----Y---NT-QD-HLPRVVVVGDQSA-GKTSVLEM-I-AQ-ARIFPRGSGEMMT-RS-PVKVTL-SEG
M. musculus       HRK--L-KK--SLIDMYSEVLD-V-LSD----YDAS----Y---NT-QD-HLPRVVVVGDQSA-GKTSVLEM-I-AQ-ARIFPRGSGEMMT-RS-PVKVTL-SEG
R. norvegicus     HRK--L-KK--SLIDMYSEVLD-V-LSD----YDAS----Y---NT-QD-HLPRVVVVGDQSA-GKTSVLEM-I-AQ-ARIFPRGSGEMM-TRS-PVKVTL-SEG
D. rerio          -RK---VKK--SLIDMYSEVLD--ILSD----YD----SNY---NT-QD-HLPRVVVVGDQSA-GKTSVLEM-I-AQ-ARIFPRGSGEMMT-RS-PVKVTL-SEG
D. melanogaster   --KI---KK--SLIDMYSEVLDE--LS----GYD--TG--YTMA----D-HLPRVVVVGDQS-SGKTSVLE-SI-A-KARIFPRGSGEMMT-R-APVKVTLA-EG
C. elegans        --K-RL--K-RSLIDMYSEVLD---L--LNE-YD-S--S-Y---NTS-DN-LPRVVVVGDQSA-GKTSVLEM--VAQ-ARIFPRGSGEMMT-R-APVKVTL-SEG

H. sapiens        P-HHVAL--FK-DSS-REF-DLT-KEE--DL-AA--LR----HE--IEL-RMRKN--V-KEG-C---TVSP-ET-ISLN------VKGP-GL-QRMVLVDLPGVI
M. musculus       P-HHVAL--FK-DSS-REF-DLT-KEE--DL-AA--LR----HE--IEL-RMRKN--V-KEG-C---TVSP-ET-ISLN------VKGP-GL-QRMVLVDLPGVI
R. norvegicus     P-HHVAL--FK-DSS-REF-DLT-KEE--DL-AA--LR----HE--IEL-RMRKN--V-KEG-C---TVSP-ET-ISLN------VKGP-GL-QRMVLVDLPGVI
D. rerio          P-HHVA-M-FK-DSS-REF-DL-GKEE--DL-AA--LR----HE--IEL-RMRK--SV-KEGQ----TVSP-ET-I------SLSVKGP-GI-QRMVLVDLPGV-
D. melanogaster   PY-HVA--QF-RDS-DRE-YDLT-KE--SDL---QDLR-RDV-EF-----RM-K-ASVR--G---GKTVS-NE-VI---AMT---VKGP-GL-QRMVLVDLPG-I
C. elegans        PY-HVA--QF-RDSS-REF-DLT-KE-T-DLQ--Q-LRN----E-T-E-VRMR-N-SVR-----DGKTVS-NE-VISL---T---VKGPN-LP-RMVLVDLPGV-

H. sapiens        --NTV-T-----SGMAP---DTKET----I------FSISKA-YM--QNPNAIILCIQDGSVDAERSI-VTDLVS-QM---DP-H---G-RRTIF-VLTKVD-LA
M. musculus       --NTV-T-----SGMAP---DTKET----I------FSISKA-YM--QNPNAIILCIQDGSVDAERSI-VTDLVS-QM---DP-H---G-RRTIF-VLTKVD-LA
R. norvegicus     --NTV-T-----SGMAP---DTKET----I------FSISKA-YM--QNPNAIILCIQDGSVDAERSI-VTDLVS-QM---DP-H---G-RRTIF-VLTKVD-LA
D. rerio          IS-TV-TT-----GMA--A-DTKET----I------FSISKA-YM--QNPNAIILCIQDGSVDAERSI-VTDLVS-QM---DP----QGKR-TIF-VLTKVD-LA
D. melanogaster   IS-T-MT-V-D---MA---SDTK----DSI--HQMT----K-HYMS--NPNAIILCIQDGSVDAERS-NVTDLV-MQ---CDP--L--G-RRTIF-VLTKVD-LA
C. elegans        IS-TV-T--AD---MA-R-----ETKDD-IIR--M--S--KAH-M-E-NPNAIILCIQDGSVDAERS-NVTDLVS---SI-DP---S-GKR-TI-LVLTKVDM-A

H. Sapiens        EKNV--AS--PS-RI--QQ-I---IEGKLFPMKALGY-FA-VVTGK-GNSS--ESIEA-----I-------R-EYEE--EFFQ-NS---KLL---K------TSM
M. musculus       EKNV--AS--PS-RI--QQ-I---IEGKLFPMKALGY-FA-VVTGK-GNSS--ESIEA-----I-------R-EYEE--EFFQ-NS---KLL---K------TSM
R. norvegicus     EKNV--AS--PS-RI--QQ-I---IEGKLFPMKALGY-FA-VVTGK-GNSS--ESIEA-----I-------R-EYEE--EFFQ-NS---KLL---K------TSM
D. rerio          EKN--LAS--PS-RI--QQ-I--V-EGKLFPMKALGY-FA-VVTGK-G--SPNESI-----DSIKD--------YEED--FFQ-NSR---LL---K--DG----M
D. melanogaster   E---ELA--DP-DRIRK---ILS---GKLFPMKALGYY-A-VVTG-RG----------RKDDSI-DAIRQ----YEED--FF-KNS---KL-FH--RR-GVI--M
C. elegans        EKN--LA-N-P-DRI-K--KIL---EGKLFPMKALGY-F--GVVTG-RGNSS----------DSI-DEIR-K---YEE-N-FF---STSQ-LL----R-DGV----
```

```
H. sapiens        -LKA-H-QV-TT-R-NL-SLAVSDC-FW-KMVRE-SV--EQ-Q-ADS-FK-AT-RFN-----LET-EWKNN-YPRLREL------DRNELF--EKAK-N-EIL-D
M. musculus       -LKA-H-QV-TT-R-NL-SLAVSDC-FW-KMVRE-SV--EQ-Q-ADS-FK-AT-RFN-----LET-EWKNN-YPRLREL------DRNELF--EKAK-N-EIL-D
R. norvegicus     -LKA-H-QV-TT-R-NL-SLAVSDC-FW-KMVRE-SV--EQ-Q-ADS-FK-AT-RFN-----LET-EWKNN-YPRLREL------DRNELF--EKAK-N-EIL-D
D. rerio          -LKA-H-QV-TT--KNL-SLAVSDC-FW-KMVRE-SV--EQ-Q-AD-AFK-A-SRFN-----LET-EWKNN-YPRLREL------DRNEL-------------YD
D. melanogaster   P----H-QV--TSR-NL-SLAVSD-RFW-KMVRE---TIEQ-Q-AD-AFK-AT-RFN-----LET-EWKNNF-PRLRE--SG--RD--ELF-D-KAKG--EIL-D
C. elegans        -LK-P-SQ-MTT-R-N-MSLAVSDC-FWR-MVR-DS--IE-SQT-D-AF-----R--AAKFNLE-AEWKNHF---LR--IRQLNRD--EL-YD-KARG--EIL-D

H. sapiens        E-V--IS-LSQ----V--TPKHWEE-ILQQS---LWE------RVSTH---------V-I-E-------NIYL-PAAQTMN-----SG-------TFNTT--VDI
M. musculus       E-V--IS-LSQ----V--TPKHWEE-ILQQS---LWE------RVSTH---------V-I-E-------NIYL-PAAQTMN-----SG-------TFNTT--VDI
R. norvegicus     E-V--IS-LSQ----V--TPKHWEE-ILQQS---LWE------RVSTH---------V-I-E-------NIYL-PAAQTMN-----SG-------TFNTT--VDI
D. rerio          E-V--IS-LSQ----V--TPKHWE-SILQ---KKLWE------RVSTH---------V-I-E-------NIYL-PAAQTMN-----SG--------TFNTT-VDI
D. melanogaster   E-VVT---LSQ--I----------------SAKK-W---DDAL--ST-KLWEKLSNYV--FES-------IYL-PAAQ--------SGSQ-NS----FNT-MVDI
C. elegans        EI-V---NL--SLIGVEE----WE---------KL--LQD--------KLW---S---GI--SSHVFDQ-I-LMPA-----YASSSSGS--------FNTT-VDI

H. sapiens        KLK---QW--TDKQ-LP-N--KA-V-EVA--WE-T-LQ---EE-F---SRF-MT------EP----KGKEHD---D-----IFDK--LKE-AV-K--E-E-SIKR
M. musculus       KLK---QW--TDKQ-LP-N--KA-V-EVA--WE-T-LQ---EE-F---SRF-MT------EP----KGKEHD---D-----IFDK--LKE-AV-K--E-E-SIKR
R. norvegicus     KLK---QW--TDKQ-LP-N--KA-V-EVA--WE-T-LQ-D-E--F---SRF-MT------EP----KGKEHD---D-----IFDK--LKE-AV-K--E-E-SIKR
D. rerio          KLK---QW--TDKQ-LPH---KA--LEVA--WE-T-LQ---EE-FA---RF-M------AE----YKGK--D-Q-D-D---IFDK--LKE-AV-KD---E-SIKR
D. melanogaster   KL---RQWAE---QALP--A-K-SV-E-A-GWE--ALQQ---E-F--IS--LM-------E-R-S-K-K---AQ-DHDG--IFD--QLK-SAVV----DEA-I-R
C. elegans        KLKHF---A--DKQ-L---AQK-SI--ET-GW-DT-L---KE-VF-F--R----QINQDA--RT--K-K------DHD-PV-FD-A-LKE-AV---IE-EA----

H. sapiens        ----HK--WN--D-FAE-----D--SLRVIQH-NA-L-EDR------SISDK-QQ---WD-AAI---YFMEEA-LQ---A---RL---KDT--E-----NA----
M. musculus       ----HK--WN--D-FAE-----D--SLRVIQH-NA--LEDR------SISDK-QQ---WD-AAI---YFMEEA-LQ------GRL---KDT--E-----NA----
R. norvegicus     ----HK--WN--D-FAE-----D--SLRVIQH-NA-L-EDR------SISDK-QQ---WD-AAI---YFMEEA-LQ------GRL---KDT--E-----NA----
D. rerio          ----HK--WN------ER-A-MD--SLRVIQH-NA-L-EDR------SITDKPQ----WD-AAIQ---FMEE-TLQ----S--RL---KDT--ES-VI--AD---
D melanogaster    R---H--SW--EDK-----AI-DM--LRVIQ-LN-TL-EDRFVH------DK--Q---EWDS-A----------------------VK--FLESSV--NA-KLV
C. elegans        -MLTH-G-W-D-DK-A-----MD-Y-LRVIQ-LNA--MED----AVQ---DK---RS-WD-AA--CN-F-----L-HKAASE-RLAAVK------------K--
```

```
H. sapiens          ------------IENMV------GP--------DWK------E----KM--LK--------------C-N-EE--H-PAY-LA---SDE--ITTV-R-KNL-ES-
M. musculus         ------------IENM-I-----GP--------DWK------E----KM--LK---------------VNDE---H-PAY-LA---SDE--ITTV-R-KNL-ES-
R. norvegicus       -----------------IENM-IGP--------DWK------E----KM--LK---------------VNDE---H-PAY-LA---SDE--ITTV-R-KNL-ES-
D. rerio            ---------------MV------GP--------DWK------ER------LLK-----------LH-----E--DHT-AY-LA--N-DEV--TTV-R-KNL-E--
D. melanogaster     QTEETLAQ------------MF-GPGQMRRITH-W-----D------K-IL-KN--D--T-K---H-------L---P--TL-TH--DE-L-TTV-R-KN-----
C. elegans          Q----L--TDDR-----------GPG--------W-ASIQDE--LYS--IL---GADPE-HKQAL-T---D---D------------D---I-TVIRR-N-IE-T

H. sapiens          ---RGV-EVDP------SLI-KDT-WHQ--V---YRRHFL-K-TA-L-----N-----------HCNL----CRRGF--YY-Y-Q---RHFV-------DSEL-
M. musculus         ---RGV-EVDP------SLI-KDT-WHQ--V---YRRHFL-K-TA-L-----N-----------HCNL----CRRGF--YY-Y-Q---RHF-------IDSEL-
R. norvegicus       ---RGV-EVDP------SLI-KDT-WHQ--V---YRRHFL-K-TA-L-----N-----------HCNL----CRRGF--YY-Y-Q---RHF-------IDSEL-
D. rerio            -A-RGV-EVDP-V-----LI-KDT-WHQ-LF-----RRHFLQK--A-LL----------------HCNL----CRRGF--YY-Y-Q---RHFV-------DSEL-
D. melanogaster     --QR-----D-NVD----------------V----------KH--FL--------QQALQRAKD-C--R--------KAYYLYTQQ-------GAEC-I-S---
C. elegans          K---GVLEV-P---TE-S-IRK--QW--KL-VFKK---HFL--------ERIINSS-------KD-C-L-SM---------Y----QMYR---QG---MI-----

H. Sapeins          EC----NDVVLFWRI-QRMLA--------------ITAN-T----LRQQL---TNT-EV--RRLE-KN--VK-EVLEDFA-EDG-----------EKK-IK---L
M. musculus         EC----NDVVLFWRI-QRMLA--------------ITAN-T----LRQQL---TNT-EV--RRLE-KN--VK-EVLEDFA-EDG-----------EKKV-K---L
R. norvegicus       EC----NDVVLFWRI-QRMLA--------------ITAN-T----LRQQL---TNT-EV--RRLE-KN--VK-EVLEDFA-EDG-----------EKKV-K---L
D. rerio            EC----NDVVLFWRI-QRML--------------GITAN-T----LRQQL---TNT-EV--RRLE-KN--VK-EVLEDF-GED-----------NEKKV----QL
D. melanogaster     -CS----DVVLFW---Q------V---IRIQK---I----T-GNALRQQ--VI-N-RE--ARRL-DK-EI-KA-VL-D---E---FS---DDE--EKK---GY--
C. elegans          E-GQDI-D-------CQ----TIVLFY-RIQKMV----NLTC-NALRQQ---ITNT-E-H-RRLE-K-EI-K-EVL-D----D-W-SQEP-D-I--KK--K-Y--

H. Sapeins          L--TGKR-VQ-L-A-EDL---KKVRE--IQEKL--DAF-IEA---L-HQ--EK-
M. musculus         L--TGKR-VQ-L-A-EDL---KKVRE--IQEKL--DAF-IEA---L-HQ--EK-
R. norvegicus       L--TGKR-VQ-L-A-EDL---KKVRE--IQEKL--DAF-IEA---L-HQ--EK-
D. rerio            --ITG-RRVQ-L-A-EDL---KKVRE--IQEKLE--AF-IEA---L-H-K-EK-
D. melanogaster     LL-TGKR-V--LLAEE-LI---KVR--QIQEKLEE--F-I---NSLN-Q--EK-
C. elegans          -L-TG-RRV-DL-AEE--IQQ--VR-R-IQEKLEE--FM--AQ--L--Q-REKI
```

Anexo A

Instruções para Autores da revista *Molecular Vision*

Molecular Vision
Biology and Genetics in Vision Research

Molecular Vision Instructions to Authors

IMPORTANT NOTES:

We editorially reject many papers: Do not let this happen to you.

Absolute "musts" before you submit:

1. Carefully edit your manuscript in American English. Use a copy editor. Write concisely and clearly. Shorter is usually better.

2. Present good data. Insufficient data, small sample groups, or underpowered studies lead to immediate rejection.

3. Conduct sound statistical analyses.

4. Present your data in clearly-labeled, high-resolution figures.

5. Tables must be carefully structured and short. Do not include columns or rows of information irrelevant to the manuscript.

6. Make a significant contribution to vision sciences.

Molecular Vision uses an online manuscript submission system. All **new manuscript submissions should use the submission system**. Instructions for manuscript preparation and the journal guidelines are detailed below and must be followed.

A major benefit of submitting a manuscript to Molecular Vision is the potential for rapid publication. For this to happen, the following instructions must be carefully followed. Departures from these instructions create more work for the editorial staff and will result in delays. Substantial departures will result in the manuscript being returned to the authors without consideration.

Synopsis of Guidelines

For questions not covered in these instructions, please email the Editors.

Types of Articles:

Molecular Vision encourages the submission of manuscripts on the molecular biology, the cell biology, or the genetics of the visual system (ocular and cortical). Manuscripts should present original, unpublished material not being considered for publication elsewhere and written to be accessible to vision scientists. If accepted, the material (data and text) shall not be published elsewhere without the consent of the Editors and Publisher of Molecular Vision. While there is no limit on the length of a manuscript, manuscripts are expected to be concise.

Following are descriptions of the types of manuscripts Molecular Vision accepts with an outline of the required elements for each manuscript type. Detailed descriptions of the elements of a manuscript are given in the section on "Preparation of manuscripts."

Research Article:

A detailed description of original, unpublished work covering a positive or negative result of significance. Reports of negative results must be investigations that other investigators would be likely to pursue in the absence of the report.

Molecular Vision does not publish simple reports of sequence data or identification of polymorphisms. Reports of new sequence (nucleotide or protein) may accompany related biological data. Simple sequencing data should be submitted to GenBank. New SNPs should be reported to appropriate sites such as http://www.ncbi.nlm.nih.gov/SNP/.

A research article must include the following sections: Title Page, Structured Abstract, Introduction, Methods, Results, Discussion, Acknowledgments, and References. Figures, Tables, Appendices, and Supplements should be included as appropriate.

Technical Brief:

This is a short account detailing a novel method or unique use of current technology that of itself, regardless of the experimental question studied, represents a significant addition to scientific enquiry.

A technical brief must include the following sections: Title Page, Descriptive Abstract, Introduction, Methods, Results, Discussion, Acknowledgments, and References. Figures, Tables, Appendices, and Supplements should be included as appropriate.

Review:

The Editors solicit suggestions for comprehensive articles reviewing the current status of a particular field or topic. The Editors should be contacted in advance with topics so that appropriateness for publication is confirmed. Appropriateness will be determined by the importance of the topic, lack of existing reviews on topic, and timeliness of request. Reviews of appropriate topics are still subject to peer review; a determination that a topic is acceptable does not guarantee acceptance of a manuscript.

A review must include the following sections: Title Page, Descriptive Abstract, Introduction, Discussion, Acknowledgments, and References. If Results are to be presented, the manuscript should include both a Methods and a Results section. Figures, Tables, Appendices, and Supplements should be included as appropriate.

Preparation of manuscripts:

The sections of the manuscript (Abstract, Introduction, Methods, etc.) should be clearly labelled with the name of the section on a line by itself. Figure captions and Tables should follow the References section. Each paragraph should be preceded and followed by a blank line. The Editors recommend that authors read the Molecular Vision Style Guide before preparing a manuscript for submission.

If any figures, graphs, tables, or data previously published are used, written permission of the publisher of the previous work must be presented. This includes works by the authors of the current submission. Obtaining this permission is the sole responsibility of the author.

Manuscript information

All manuscripts should begin with the title of the manuscript, the names of the authors, the authors' institutional affiliations, and contact information for the corresponding author. Author names should be in order of first name, middle initial, and last name. After each author name, an affiliation needs to be cited in superscript numbers. Each affiliation should

start with superscript number and each department at the same University and Institute needs to have a separate affiliation number

We allow only one corresponding author. If the authors wish to have a co-corresponding author, we acknowledge the co-corresponding author in Acknowledgments section and provide the email addresses of both corresponding authors.

Abstract

The abstract should concisely (in fewer than 4000 characters) summarize the work presented. Manuscripts submitted as a Research Article must have a structured abstract consisting of four subsections: Purpose, Methods, Results, and Conclusions. Other types of manuscripts must include a descriptive abstract that details the topics covered in the manuscript.

Introduction

A succinct introduction without subheadings should describe the purpose or goals that led to the production of the manuscript. This should include a concise review of relevant literature.

Methods

This section should detail everything that would be required to replicate the work presented. Non-proprietary names should be used whenever possible. Where relevant, the name of the supplier of items used in the investigation should be identified. Suppliers should be identified by their full company name and location (city, state/country); if the company maintains an online presence, a URL may also be given. The supplier name and location (city, state/country) should be cited in full on first mention and after that, it shall be cited only by company name.

Studies involving animals should include a statement that animal care guidelines comparable those published by the Institute for Laboratory Animal Research (Guide for the Care and Use of Laboratory Animals) or the US Public Health Service (Public Health Service Policy on Humane Care and Use of Laboratory Animals) were followed.

Studies that involve human subjects should indicate that an appropriate institutional review board approved the project. The authors must document that informed consent was obtained. The authors should verify that they adhered to the tenets of the Declaration of Helsinki (JAMA 1997; 277:925-926).

Results

The findings of the investigation should be presented without interpretation or discussion. In short manuscripts where the interpretation of the data is relatively simple, the data may be discussed in a combined Results and Discussion section.

Statistics

Provide sample sizes (N) of each sample group in each experiment. For data presented with statistical analyses, the type of statistical test(s) must be named. The level of significance must be provided. It is the responsibility of the authors to ensure that the assumptions of a given test are met by their data. If this information is not provided, the manuscript will be returned or rejected.

Discussion

An interpretation and commentary on the data (research article) or technique (technical brief) should be presented without speculating beyond the scope of the investigation. For reviews, the discussion is the bulk of the manuscript. It may be sectioned to fit the topic being reviewed and each section may be titled by the author.

Acknowledgments

Authors may briefly mention individuals making significant non-authorship contributions to the manuscript.

Funding support for the work presented should be detailed. Specific grant numbers should be provided.

Authors shall disclose any commercial interests in the subject of the manuscript or in entities discussed in the manuscript.

All prior presentation of the manuscript's data at meetings should be indicated; such presentations should not appear among the manuscripts references.

A co-corresponding author can be acknowledged here, and both corresponding authors email addresses can be provided here.

Citations and References

Citations in the text are noted in appropriate places by numbers in brackets (e.g., [3,5,8-12]).

References must be numbered in the order of appearance in the text of the manuscript. All references cited in the text should be listed in the References with corresponding numbers. References should be numbered sequentially and provided as an Arabic number followed by a period. Please do not use parentheses or brackets after or enclosing the reference number. Here is the correct format:

1. Watson JD, Crick FHC. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. Nature 1953; 171:737-38.

Below are incorrect formats of references:

(1) Watson JD, ...

[1] Watson JD, ...

1) Watson JD, ...

1 Watson JD, ...

References should follow the Vancouver style as described by the International Committee of Medical Journal Editors (Ann Intern Med 1997; 126:36-47) except for part 33 which is outdated (see Example 4 below). This style orders elements of the source of journal articles from least to most specific and has been adopted by the National Library of Medicine.

References to unpublished work should be made parenthetically in the body of the text and not listed as a citation in the References section. If the data comes from some or all of the authors of this work, you may simply list it as "[data not shown]." If the data was provided by some other party (e.g., "John Smith"), the party that conveyed the information should be named as "(Personal communication, Dr. John Smith, Department, Institution, City, State or province, Country)." If you want to provide a more detailed acknowledgement of a person's contribution, you may add that to the Acknowledgements section. Citations to submitted manuscripts are not allowed because there is no guarantee that the source material will ever exist in the form in which it was cited.

Please list all author names in each reference. Do not use “et al”.

Molecular Vision has compiled a page of citation tools with links to tools that are useful in preparing the References section of a manuscript. Additionally, the page includes excerpts of the most commonly used citation types in the Vancouver Style. Remember that incomplete or inaccurate references are not useful to the reader. Following are a few referencing examples that cover what most Molecular Vision articles require:

1. Watson JD, Crick FHC. Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. Nature 1953; 171:737-38.

2. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor (NY): Cold Spring Harbor Press; 1989.

3. Watson JD. How we did it. In: Pauling LC, Kant E, Nietzsche FW, editors. We may have been wrong, progress in thinking. Vol 29. New York: Putnam; 1995. p.

278-99.

4. Wistow G. Peptide sequences for β-crystallins of a teleost fish. Mol Vis 1995; 1:1 <http://www.molvis.org/molvis/v1/a1/>.

Figures

Figures should be numbered with Arabic numerals according to their sequence of appearance in the text, where they are cited as "Figure 1", "Figure 2", etc. Figure captions should follow the References section in the body of the manuscript. Each figure caption should have a title and body. The text of a caption should be sufficient to explain the figure without referring to the body of the manuscript. Captions for figures presenting data must describe the result presented. Labels and abbreviations must also be explained in the caption. Authors should feel free to use color figures in any way that better communicates their message.

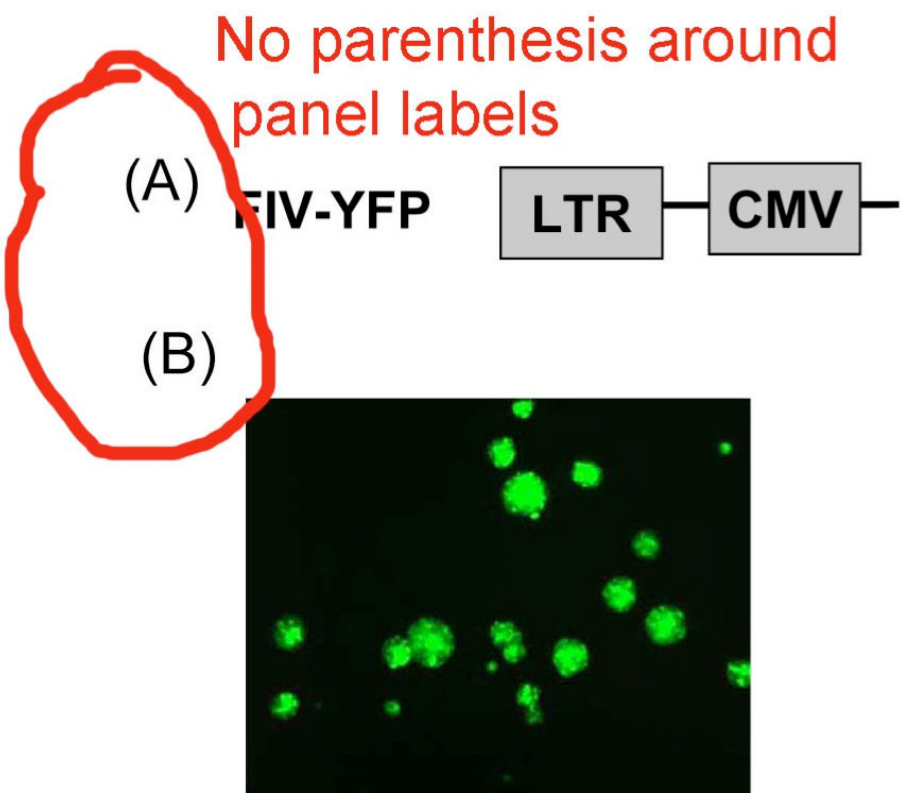
Figures that are composites of multiple images (e.g., gels, micrographs) should contain readily discernable white space between the different images. This also applies to composite images of non-adjacent lanes of the same gel.

Figures by number, title, caption, and image file (tiff or jpg) are submitted separately from main text at the submission site. The main text should contain figure captions as well.

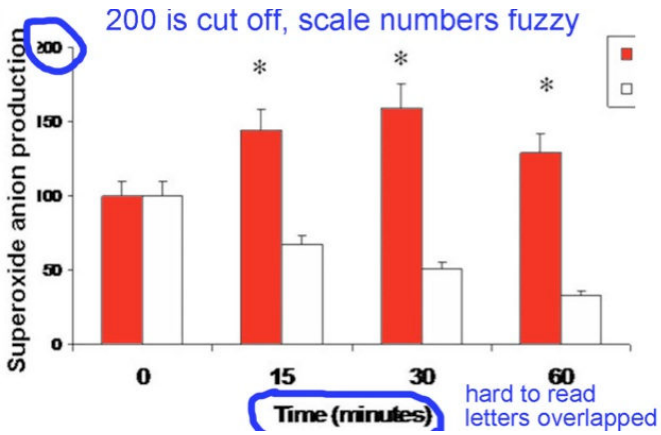
Figures must be high resolution, at least 300 dpi, in a full page size image, or on final size image. All text on an image must be legible (at least 8 points) when reduced to final publication size in the PDF. IE, Consider that a single column is only 3.375 inches wide and make sure your text is clearly legible when reduced to that width.

Please see the following examples of bad images:

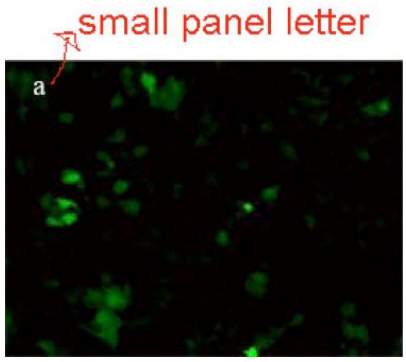
1. Please do not use parentheses in panel labels. Use an uppercase alphabetic character for the panel tag (cf., A, B, C, ….). Do not use numerals. Do not use A1, A2, etc.



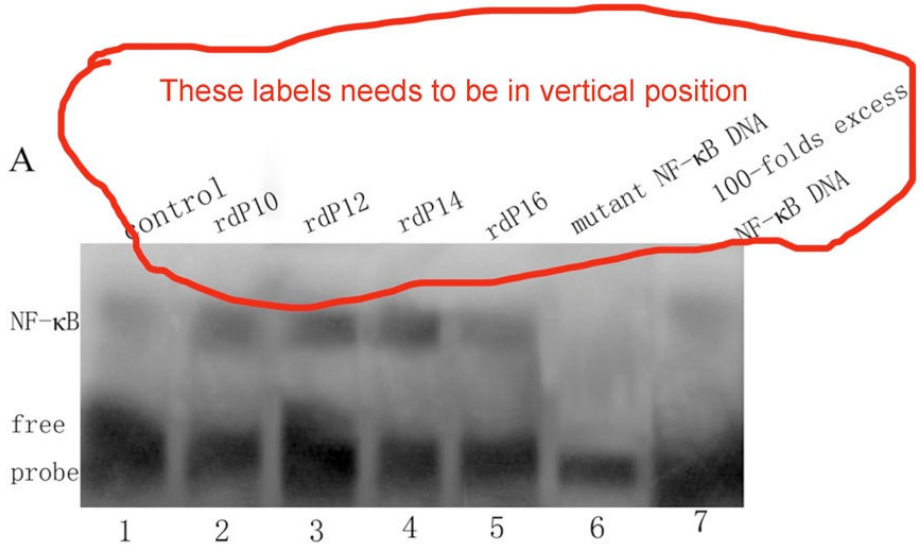
2. Please make sure labels are not cut off, and leave some space between letters. Below, “200” is cut off. The “minutes” label is hard to read.



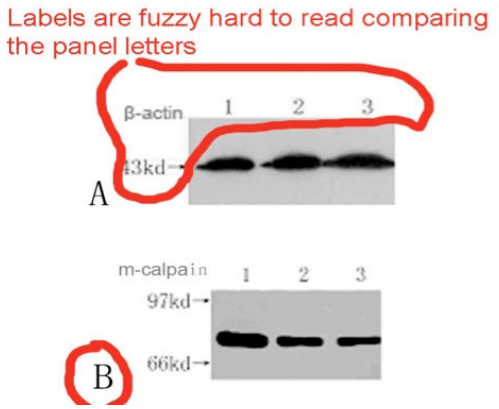
3. Please do not use lowercase panel letters; always use an uppercase letter for each panel tag.



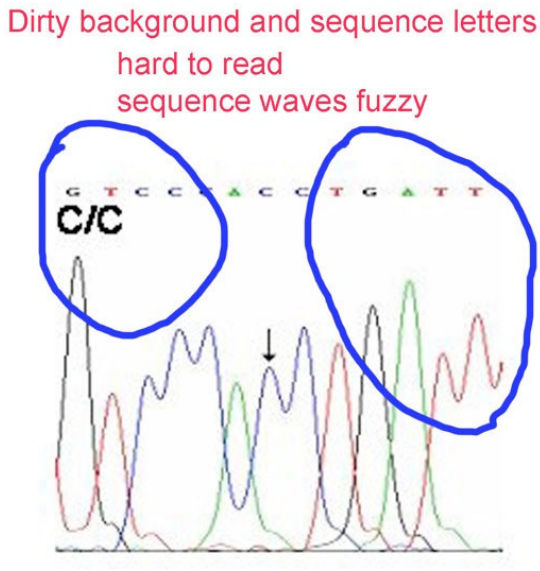
4. In the next panel below, the labeling is legible, but it needs to be in a vertical orientation.



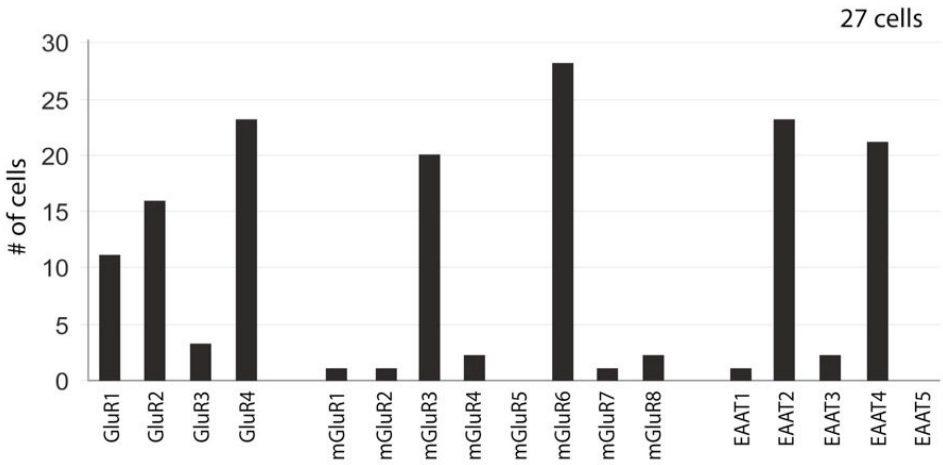
5. In the next panel, panel tags A and B are easy to read but β-actin, 43kd, m-calpain, 97kd, 67kd and 1, 2, 3 are faint and fuzzy. All the labels need to be sharp. Panel tags, A and B, must be in the upper left corner of the image. Note that "kd" is incorrect and it should be "kDa".



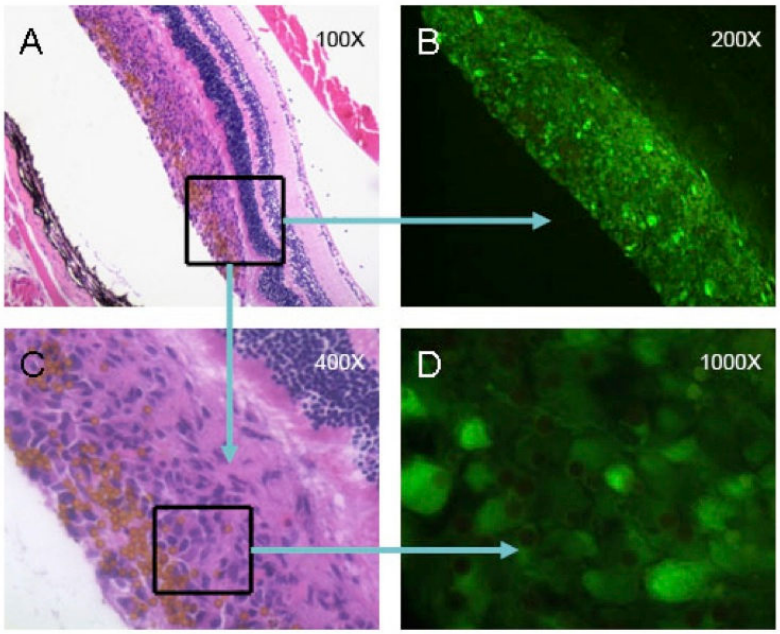
6. In the panel below, especially in images of sequencing chromatograms, uneven background must be corrected. Sequence lettering ought to be replaced with crisp lettering.



7. In the next panel, the image labels are all sharp. This is the correct display of text on the vertical axis. However, Error Bars are apparently missing in this example.

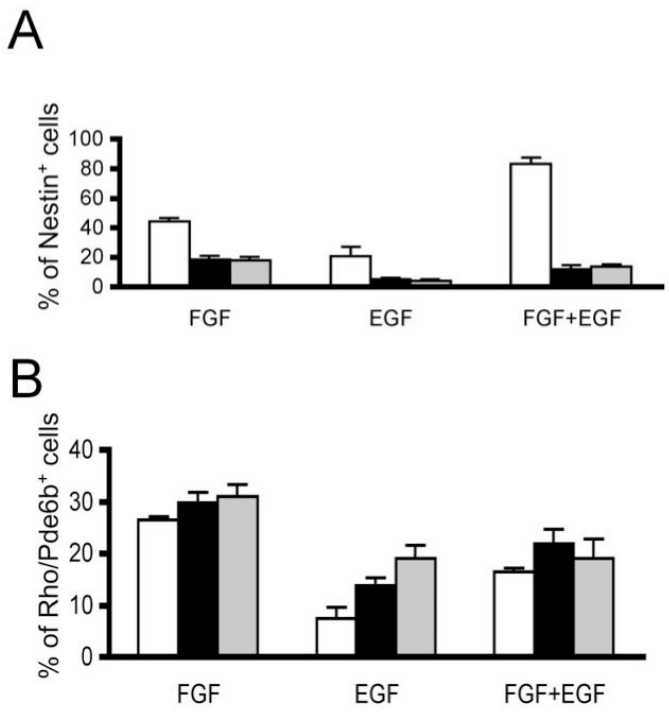


8. In the next image, a figure with multiple panels, the caption should explain each and every panels. All micrographs should contain a scale bar. Please position the scale bar in the lower right corner of the last panel. It is not sufficient to provide the magnification in the caption or on the image.



Examples of good figure images:

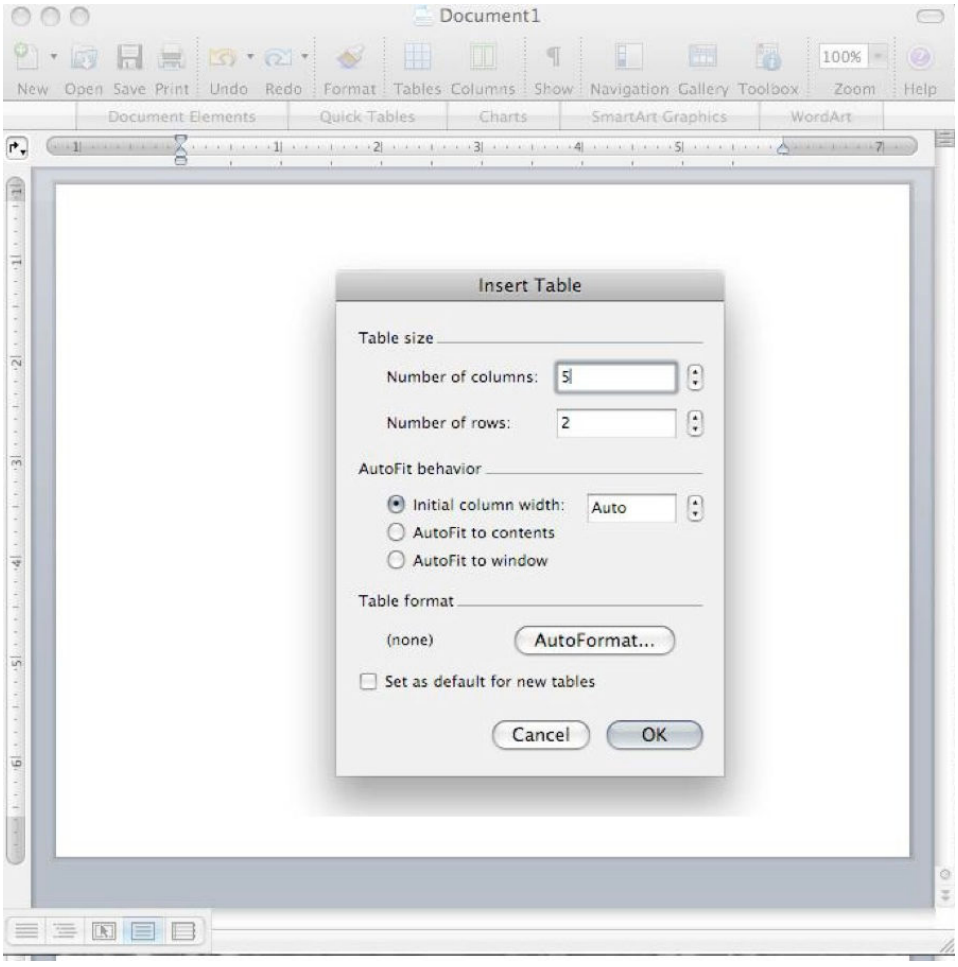
In the next panels, lines and objects are sharp, and the labels are in the upper left corner. All the other labels are easy to read. Note that Error Bars are shown. These should be defined as either standard deviations or standard error of the mean in the caption, and sample sizes per group should be provided in the caption.



Tables

Tables must be submitted as a Table-formatted Word document. Excel file format is also allowed. Tables should be numbered with Arabic numerals according to their sequence of appearance in the text, where they are cited as "Table 1", "Table 2", etc. All tables should be cited in the text body. Each table should have a title and caption. The text of a caption should be sufficient to explain the table without referring to the body of the manuscript.

Tables that are images are not permitted, and the manuscript will be rejected without review.



Tables must be: 1 page or smaller in a Word or excel file with 1 inch margins on all sides, 12 point font size, Times new roman font style, portrait or landscape and no more than 8 columns wide. Tables longer than one page will automatically be converted to supplementary files and will be shown as appendices.

Appendices

All large tables (over one page in size) and datasets should be provided as appendices and will be regarded as supplementary materials of the publication. All Appendices should be cited in the text body. Please prepare appendices as Word doc tables or as excel files. Other file formats are acceptable but must be approved by the Editors in Chief. These file

formats must be compatible with standards set by the National Library of Medicine for publisher submission to PubMed and PubMed Central.

Release of Data

Every manuscript should contain all necessary data for the reader to reach the conclusions reported by the authors. This is much easier to accomplish today because of the public databases available for sequence and structure data. Depositing data in these databases makes it easily available by reference to an identification code that may be cited throughout the literature. Further, it allows for a systematic organization and searching of a class of data.

Authors submitting manuscripts containing newly reported nucleotide or protein sequences must deposit those sequences in the GenBank database. Submissions reporting new gene expression data must submit those data to the Gene Expression Omnibus (GEO). Submissions reporting new three-dimensional structures must submit those structures to the Protein Data Bank. Such submissions should include all structural data supporting the manuscript's conclusions, including any derived atomic coordinates. **The accession or identification numbers for these deposits must be provided to Molecular Vision prior to acceptance and must be released (available to the public) prior to publication.** It is the authors responsibility to arrange for the release of this information with the relevant database. **Any delay in receipt of functioning accession numbers will delay publication. Failure to submit the full datasets will result in rejection of the manuscript.** Authors are also required to include accession numbers for any sequences or structures relevant to their manuscript.

Molecular Vision reserves the right at any time to require the submission of other large datasets to major public databases as the editors deem appropriate. These data sets will include metabolomics and protein data sets.

Genome-wide association studies

Molecular Vision encourages authors of genetic association studies to follow prevailing standards (Little et al. PLoS Med 2009; 6: 151-62) for reporting results. Replication of prior studies, whether positive or negative are of value to the community and will receive consideration until the association is clearly established or refuted. Replication studies will receive priority if they extend the work of previous studies and provide a better understanding of the nature of the association. Novel findings in genetic association studies will be given priority if they are replicated in an independent group of subjects. Such replication is important for initial observations and novel findings in replication studies, whether they are derived from genome-wide or candidate gene approaches. Although present standards include guidelines of $p < 1.0 \times 10^{-5}$, and sample sizes of at least 250 members per group, these likely will become more strict in the near future. Clearly, sample size and significance depend on the questions being asked and the nuances of the hypotheses, which may require larger sample sizes and stricter levels of significance.

Meta-analyses

A meta-analysis typically is an attempt to identify relationships across several study results, assessing contrast or agreement, and possibly generating new hypotheses whose testing will move the field forward.

For purposes of publishing in Molecular Vision, a meta-analysis must generate new, testable hypotheses.

To quality as a full Research Article a meta-analysis should communicate at least a modest amount of new results (previously unpublished) and should examine aggregated data (e.g., p-value) for each marker (e.g., SNP) for each of numerous SNPs from several prior GWAS studies. The meta-analysis should include clinical criteria for selecting cohorts, ethnicity, gender, genotyping methods, and error rates from the groups that did the initial studies.

A meta-analysis that presents and uses an analytical approach that is new and distinct from previous approaches may be considered a Technical Brief.

A meta-analysis that lacks novel results or findings or analytical approach is a Review Article.

Genetic Analysis

For pedigree analysis and linkage studies, relatively large pedigrees should provide lod scores (>+3.0 or <-2.0) sufficient to accept or exclude linkage to a new locus unless also supported by identification of a high probability mutation in a gene in the included region or other evidence of biological significance. Studies confirming previously published results will be considered as long as they provide additional insight into the existing association or linkage data. This includes identification or refinement of the biological import, the importance of previously identified linkage or associations in specific populations, and the genetic mechanism for population differences.

Polymorphism studies

Polymorphism reports are not generally acceptable unless there are a large number of new polymorphisms reported such that the resource needs explanation in the literature, or the manuscript provides insight into a biological process of general interest. In general, polymorphisms should be deposited in a suitable public database such as dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) prior to publication.

New Mutation Studies

The standards for publication of new mutation studies in "Molecular Vision" have become stricter.

We will still consider manuscripts that describe a gene lesion that causes disease in a new gene, provided that the manuscript data are convincing. Subsequently, for new mutations found in the same gene with the same disease, we expect a larger number of new mutations in each article, due to the increasing ease of detection of causative mutations and a glut of articles. For example, many gene lesions have been identified in rhodopsin that cause retinitis pigmentosa. We will not consider a manuscript on just a single new mutation in rhodospin. Starting now, we would consider two new novel mutations in rhodopsin. We expect that in the near future we will likely expect several new causative mutations per paper. The new standards are: Two or more new mutations per gene per article in 2012, three or more new mutations in 2013, and so on.

Manuscripts that include functional biological data or experimental data regarding a new mutation will be given significantly higher priority over those that lack it. Such manuscripts are not subject to the above rule on the number of new mutations. To be more competitive for the limited number of articles that Molecular Vision can publish, we strongly urge each article to contain supporting experimental data of this nature.

Use of immortalized cell lines

Manuscripts reporting experiments in which immortalized cell lines are used must include data and documentation that demonstrate that the actual cells used in the experiments reported in the manuscript exhibit the correct phenotype and genotype. Authors should read this article by the editors on the authentication of cell lines (The Editors. On authentication of cell lines. Mol. Vis. 2013; 19:1848-1851). The following instructions not withstanding, authors will be held to the strictures of that editorial.

In brief, authors must demonstrate that the cells actually used are of the correct species of origin, the correct sex and genotype, and express genes and gene products that are specific to the pertinent cell type. It is expected that the cell line expresses genes necessary for hypothesis testing. This may require specific differentiation. Regardless, correct expression must be documented in the manuscript. Cells used in experiments should be within a few passages of authentication (typically five passages). These standards hold even if the cell lines are considered "established" and were obtained from reputable sources. A statement of cell handling protocol that includes passage information and authentication data, and certification documentation must be provided in Methods sections of submitted manuscripts. These will be part of the freely-available article. Submissions that use primary cell cultures or animals to verify hypotheses tested first in immortalized cell lines will be viewed in general more favorably than work that relies solely on cell lines. Again, authors will be held to those standards in the editorial (The Editors. On authentication of cell lines. Mol. Vis. 2013; 19:1848-1851).

Use of the RGC-5 cell line to model retinal ganglion cells

The RGC-5 cell line was originally reported to be derived from postnatal day 1 rat retina cells, expressed markers specific to retinal cells, and was sensitive to trophic factor withdrawal and glutamate toxicity following treatment with differentiation factors (Krishnamoorthy et al., Brain Res Mol Brain Res. 2001;86(1-2):1-12). Several research groups report that even cells obtained from the originating laboratories are not of rat origin and do not express genes and proteins specific to retina or retinal ganglion cells (e.g., Van Bergen et al., IOVS 2009 50:4267-4272, Krishnamoorthy et al., Invest Ophthal Vis Sci. 2013;54:5712-5719).

New manuscripts containing data derived from RGC-5 cells will be editorially rejected without review.

Negative Data

Studies of a negative nature are accepted only if the logic for carrying out the experiment is compelling, it is highly likely that other groups in the field would inevitably perform the same futile studies repeatedly, and such a report would warn them away from wasting time and money. Studies showing an inability to confirm previously reported association or linkage data will be considered depending on comparability to the original study in terms of disease phenotype, population of interest, methodology, and the power of the repeat study to exclude association or linkage (e.g., 95% probability limits for the odds ratio).

Submission of manuscripts:

New manuscript submissions should use our online submission system.

Instructions for warranting author responsibilities and contributions are handled within the online submission system. Each and every author must personally answer these queries. Please assure that all author names and email addresses are correct and current when submitting your manuscript. Each author must have his/her own unique email address. The corresponding author is not authorized to answer on behalf of the other authors. US government employees must answer the same set of questions online as other authors. Errors at this step will greatly delay the completion of the submission process and block the review process.

Author Responsibilities, Licensing, and Copyrights:

Each author shall sign an online statement warranting that he/she authored the work, actively participated in the reported experiments, and he/she approves and agrees with representation of the data in the article. Each author shall warrant that the article is original, is not being considered for publication by another journal, and has not been previously published other than as an abstract.

Each author shall sign an online statement that, if the article is accepted for publication, the article and all supplementary materials will be made freely available online immediately with a Creative Commons Attribution-NonCommercial-NoDerivatives License 3.0, or CC BY-NC-ND 3.0 (see http://creativecommons.org/licenses/by-nc-nd/3.0/ for license terms). The authors retain copyright and grant Molecular Vision an irrevocable, royalty-free, perpetual license to publish and distribute the article, in all formats now known or later developed, and to identify Molecular Vision as the original publisher.

Competing interests

The authors shall declare any competing, commercial, or conflicts of interests, regarding this publication. Each author warrants this in the online submission system.

Resubmission of revised Manuscripts

Revisions to active manuscripts shall be resubmitted through the online manuscript system.

Do not submit a revised manuscript as a new manuscript.

Notices about republication

For noncommercial use, republished material must include a citation to the original work in Molecular Vision and must not be changed from the original. See http://creativecommons.org/licenses/by-nc-nd/3.0/ for license terms.

Articles authored by a student may be reproduced for inclusion in the student's thesis or dissertation, whether in an online repository at the student's institution or through a commercial service used by the student's institution. Such republications must include a citation to the original work in Molecular Vision and must not be changed from the original. They do not require further permissions from Molecular Vision for republication in a thesis or dissertation submitted in partial fulfillment of requirements for an advanced scholarly degree. CC BY-NC-ND 3.0 (see http://creativecommons.org/licenses/by-nc-nd/3.0/ for license terms) applies to derivative works other than the thesis or dissertation.

Commercial publishers shall ask permission from Molecular Vision to republish parts of an article. Specific directions are online at Commercial Permissions. The publisher also shall ask permission of The Authors directly.

Contact the editors at <molvis@emory.edu>.

ISSN 1090-0535