# Identifying the ARR3 gene responsible for a cone dystrophy using genetic browsers

Hayder Muhsin Khalfa , Ameer Ali Shakir , Noor Hassan Khalfa

**Abstract**—This study aims to identify the gene responsible for a cone dystrophy within a family. Prioritizing several possible genes that can be associated with cone dystrophies this task investigates the gene with cone associated function and highest expression in the retina. ARR3 gene has been found to have the highest expression within the retina compared to other genes found on the locus of the X chromosome. Using different website and electronic genetic databases the section on the x chromosome has to be identified and examined. Identification of the genes in the region on the X chromosome is essential since all genes have to be analyzed. Each gene found need to be assessed in terms of function and region of expression within the human body. Genes expressed in the retina posing a photoreceptor function are prioritized to narrow down the focus on a particular gene. Top priortised genes are also assessed on how many EST (expressed sequence tags) they have. Top priortised genes are also assessed on how many EST (expressed sequence tags) they have. After the top gene has been selected thorough analysis is required to identify the function of the gene, this is achieved using literature search to determine the association of the gene and photoreceptors. The selected gene contains exons which are the coding regions of the gene, primers have to be designed to amplify the sequence of each exon present in the gene. ARR3-201 transcript of ARR3 gene found in the region between the two different microsatellite markers provided has been analyzed and primers have been designed to amplify its exons. Evidence has been provided in the literature that this gene is localized in cone photoreceptors and its associated with different functions in the cones photoreceptors furthermore the expression rate was significantly high in the retina which provides more evidence that ARR3 gene is possibly associated with cone dystrophy identified within the family.

**Index Terms**— ARR3, Cone Dystrophy, Gene, X chromosome, Genetic Browsers, Sequence Tags, EST.

## 1 Introduction

This study aims to identify the gene responsible for a cone dystrophy within a family. Prioritizing several possible genes that can be associated with cone dystrophies this task investigates the gene with cone associated function and highest expression in the retina. ARR3 gene has been found to have the highest expression within the retina compared to other genes found on the locus of the X chromosome. To aid the process of genetic identification of possible mutations that occurred within the ARR3 gene a pair of primers have been designed for each exon present in the ARR3 gene.

## 2 Strategy to identify the target gene

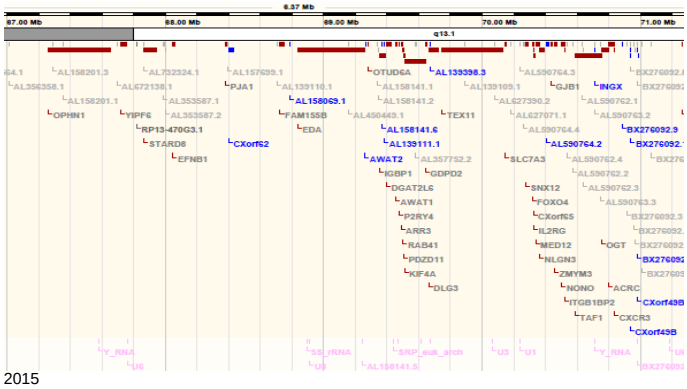
This task follows a particular strategy to identify the gene responsible for the cone dystrophy. As suggested by the information provided the cone dystrophy is X linked and responsible genes fall in between two possible microsatellite markers. Using different website and electronic genetic databases the section on the x chromosome has to be identified and examined. Identification of the genes in the region on the X chromosome is essential since all genes have to be analyzed. Each gene found need to be assessed in terms of function and region of expression within the human body. Genes expressed in the retina posing a photoreceptor function are prioritized to narrow down the focus on a particular gene. Top priortised genes are also assessed on how many EST (expressed sequence tags) they have. After the top gene has been selected thorough analysis is required to identify the function of the gene, this is achieved using literature search to determine the association of the gene and photoreceptors. The selected gene contains exons which are the coding regions of the gene, primers have to be designed to amplify the sequence of each exon present in the gene.

## 3 Location of the X chromosome linked with the disease

The region identified by the microsatellite markers provided and has been located on the X chromosome between 67,251,011-70,436,328 and segregating region Xq12-21. Below is the identified section on the X chromosome which contains the suspected gene responsible for the cone dystrophy.


## 4 Gene expression in humans

Below is a table showing the genes within the identified region on the X chromosome. The table contains the formal name of the gene and its expression in the human body. The expression of each gene is associated with the prioritizing of the top genes responsible for the cone dystrophy in the family (3).



| Gene | Name | Locus tag | Expression | Map position |
|---|---|---|---|---|
| OPHN1 | Oligophrenin 1 | RP6-201G10.1 | Highly expressed in adipose tissues, small expression in the eye. | Xq12 |
| YIPF6 | Yip1 domain family, member 6 | RP13-253K9.1 | Highly expressed in pituatry glands, small expression in retina | Xq12 |
| STARD8 | StAR-related lipid transfer (START) domain containing 8 | | Highly expressed in nerves with very small expression in the eye. | Xq13.1 |
| EFNB1 | Ephrin-B1 | RP3-496G1.1 | Highly expressed in ascites with a degree of expression in the eye. | Xq12 |
| PJA1 | Praja ring finger 1 | chrX:68,297,429-68,302,065 | Highly expressed in the umbilical cord with small expression in the eye. | Xq13.1 |
| FAM155B | Family with sequence similarity 155, member B | chrX:68,641,803-68,669,076 | Highly expressed in the thyroid with small expression in the eye. | Xq13.1 |
| EDA | Ectodysplasin A | | Highly expressed in the spleen with no expression in the retina. | Xq12-q13.1 |
| OTUD6A | OTU domain containing 6A | | | Xq13.1 |
| IGBP1 | Immunoglobulin (CD79A) binding protein 1 | | Highly expressed in the cervix with relative expression in the eye. | Xq13.1-q13.3 |
| DGAT2L6 | Diacylglycerol O-acyltransferase 2-like 6 | RP13-26D14.8 | Highly expressed in skin cells with no expression in the eye. | Xq13.1 |
| RAB41 | RAB41, member RAS oncogene family | RP13-26D14.4 | Highly expressed in the eye compared with other areas. | Xq13.1 |
| AWAT1 | Acyl-CoA:wax alcohol acyltransferase 1 | RP13-26D14.7 | Highly expressed in heart tissues with no expression in the eye | Xq13.1 |
| P2RY4 | Pyrimidinergic receptor P2Y, G-protein coupled, 4 | RP13-26D14.5 | No Results found | Xq13 |
| ARR3 | Arrestin 3, retinal (X-arrestin) | RP13-26D14.6 | Highly expressed in the eye. | Xcen-q21 |
| PDZD11 | PDZ domain containing 11 | RP13-26D14.3 | Highly expressed in bladder cells with relative expression in the eye. | Xq13.1 |
| KIF4A | Kinesin family member 4A | | Highly expressed in the cervix with relative expression in the eye | Xq13.1 |
| GDPD2 | Glycerophosphodiester phosphodiesterase domain containing 2 | | Highly expressed in the spleen with low expression in the eye. | Xq13.1 |
| DLG3 | DLG3, discs, large homolog 3 | RP11-291O7.3 | Highly expressed in the pancreas with some expression in the eye | Xq13.1 |
| TEX11 | Testis expressed 11 | RP11-291O7.2 | Highly expressed in the testis with no expression in the eye. | Xq13.1 |
| SLC7A3 | Solute carrier family 7 (cationic amino acid transporter, y+ system), member 3 | | Highly expressed in the thymus with no expression in the eye. | Xq13.1 |
| SNX12 | Sorting nexin 12 | | Highly expressed in the cervix with reasonable expression in the eye. | Xq13.1 |
| CXORF65 | Chromosome X open reading frame 65 | | Highly expressed in the bone marrow with no expression in the eye. | Xq13.1 |
| IL2RG | Interleukin 2 receptor, gamma (severe combined immunodeficiency) | | Highly expressed in the blood with very small expression in the eye. | Xq13.1 |
| MED12 | Mediator complex subunit 12 | | Highly expressed in the bone marrow with some small expression in the eye. | Xq13 |
| NLGN3 | Neuroligin 3 | | Highly expressed in the brain with some small expression in the eye. | Xq13.1 |
| GJB1 | Gap junction protein, beta 1, 32kDa | | Highly expressed in skin with no expression in the eye. | Xq13.1 |
| ZMYM3 | Zinc finger, MYM-type 3 | DXS6673E | Highly expressed in the ovary with good expression in the eye | Xq13.1 |
| NONO | Non-POU domain containing, octamer-binding | | Highly expressed in the oesophagus with high expression in the eye | Xq13.1 |
| ITGB1BP2 | Integrin beta 1 binding protein (melusin) 2 | | Highly expressed in the umbilical cord with no expression in the eye | Xq12-q13.1 |
| TAF1 | TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa | | Highly expressed in the larynx with small expression in the eye | Xq13.1 |

## 5 TOP 9 GENES

Below is a table showing the top nine genes that have been selected after a thorough analysis of their expression and function using different genetic web browsers. Each gene in the table was investigated using the eye browse from the NEIBANK (3) browser to determine the number of EST (expressed sequence tags) each gene had. The number of EST has been counted manually for each gene.


- *Hayder Muhsin Khalfa: Assistant lecturer in Biotechnology, Department of Biology, Faculty of science, Kufa university, Iraq E-mail: HaydarHussain@gmail.com*
- *Ameer Ali Shakir: Department of Biology, Faculty of science, Kufa university*
- *Noor Hassan Khalfa: Department of Computer Science, College of science for Girls, Babylon University, Iraq*


| | | | | |
|---|---|---|---|---|
| | | phosphorylation by casein kinase II (1) association of arrestin-3 with the FSHR is dependent on receptor phosphorylation, however phosphorylation of the third intracellular loop residues is not needed for arrestin-3 association. Autoimmunity to this antigen and retinal vasculitis is not altered in patients with Behcet's disease (1). | | |
| KIF4A | Kinesin family member 4A | Motor protein that translocates PRC1 to the plus ends of interdigitating spindle microtubules during the metaphase to anaphase transition (1), an essential step for the formation of an organized central spindle midzone and midbody and for successful cytokinesis. May play a role in mitotic chromosomal positioning and bipolar spindle stabilization (1). | 1 | Highly expressed in the cervix with relative expression in the eye |
| DLG3 | DLG3, discs, large homolog 3 | This gene encodes a member of the membrane-associated guanylate kinase protein family. The encoded protein may play a role in clustering of NMDA receptors at excitatory synapses (1). It may also negatively regulate cell proliferation through interaction with the C-terminal region of the adenomatosis polyposis coli (1). tumor suppressor protein. Mutations in this gene have been associated with X-linked mental retardation. Alternatively spliced transcript variants have been described (1). | 0 | Highly expressed in the pancreas with some expression in the eye. |
| ZMYM3 | Zinc finger, MYM-type 3 | Identified a family of multiprotein corepressor complexes that function through modifying chromatin structure to keep genes silent (1). The polypeptide composition of these complexes includes a common core of 2 subunits. Other subunits of these complexes include ZNF261, GTF2I (601679), and polypeptides associated with cancer-causing chromosomal translocations (1). | 4 | Highly expressed in the ovary with good expression in the eye |
| NONO | Non-POU domain containing, octamer-binding | This gene encodes an RNA-binding protein which plays various roles in the nucleus, including transcriptional regulation and RNA splicing (1). A rearrangement between this gene and the transcription factor E3 gene has been observed in papillary renal cell carcinoma. Alternatively spliced transcript variants have been described. Pseudogenes exist on Chromosomes 2 and 16 (1). | 8 | Highly expressed in the esophagus with high expression in the eye |

| Gene | Name | Known function | EST no | Expression |
|---|---|---|---|---|
| OPHN1 | Oligophrenin 1 | Rho-GTPase-activating protein that promotes GTP hydrolysis of Rho subfamily members. Mutations in this gene are responsible for OPHN1-related X-linked mental retardation with cerebellar hypoplasia and distinctive facial dysmorhphism (1). | 1 | Highly expressed in adipose tissues, small expression in the eye. |
| STARD8 | StAR-related lipid transfer (START) domain containing 8 | START; STeroidogenic Acute Regulatory (STAR) related lipid Transfer Domain. These domains are 200-210 amino acid in length and occur in proteins involved in lipid transport (phosphatidylcholine) and metabolism, signal transduction, and transcriptional (1). | 0 | Highly expressed in nerves with very small expression in the eye. |
| IGBP1 | Immunoglobulin (CD79A) binding protein 1 | The proliferation and differentiation of B cells is dependent upon a B-cell antigen receptor (BCR) complex. Binding of antigens to specific B-cell receptors results in a tyrosine phosphorylation reaction through the BCR complex and leads to multiple signal transduction pathways (1). | 1 | Highly expressed in the cervix with relative expression in the eye. |
| RAB41 | RAB41, member RAS oncogene family | RAB; Rab subfamily of small GTPases, Rab6; Rab6 subfamily. Rab6 is involved in microtubule-dependent transport pathways through the Golgi and from endosomes to the Golgi. Rab6A of mammals is implicated in retrograde transport through the Golgi stack (1). | 2 | Small expression but Mainly expressed in the eye compared with other areas. |
| ARR3 | Arrestin 3, retinal (X-arrestin) | The molecular interactions of arrestin2 and arrestin3 and their individual domains with the components of the two MAPK cascades (1). The light-dependent translocation of cone arrestin suggests a role in light-dark adaptation of cones, Arrestin in all conformations binds JNK3 comparably (1), whereas Mdm2 preferentially binds cone arrestin 'frozen' in the basal state. Regulation of arrestin-3 | 20 | Highly expressed in the eye. |

## 6 ORDER OF PRIORITIZED 5 GENES

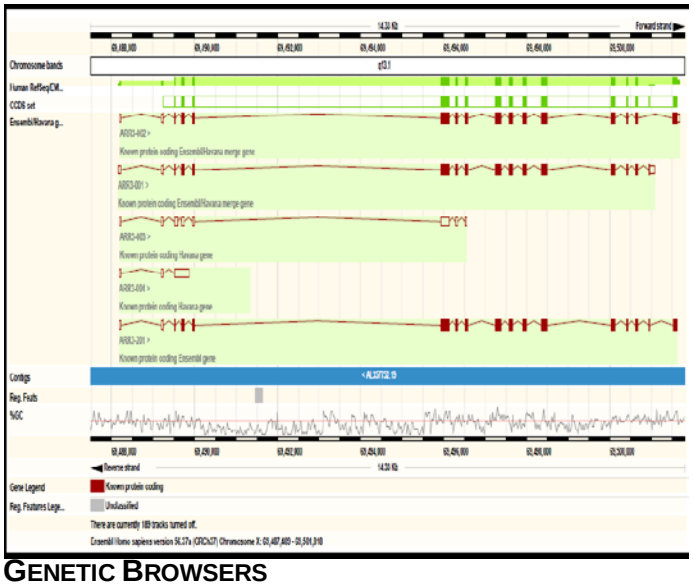
Below is a table showing the order of the prioritized top 9 genes and the known function of the gene as well as the processes involved in. The order of genes has been determined by the number of EST each gene has and the function it preformed in the eye. After an intense investigation I have found the above a meaningful prioritization since most of the genes selected have a high expression and high EST compared with other genes found on the section of the X chromosome.

| Order | Gene | No of EST's | Function | Process |
|---|---|---|---|---|
| 1 | ARR3 | 20 | G-protein-coupled receptor binding<br>opsin binding<br>phosphoprotein binding<br>protein domain specific binding (1 ) | regulation of protein amino acid phosphorylation<br>response to stimulus<br>sensory perception<br>signal transduction<br>visual perception<br>endocytosis (1) |
| 2 | NONO | 8 | DNA binding<br>RNA binding<br>identical protein binding<br>nucleotide | DNA recombination<br>DNA repair<br>RNA splicing<br>mRNA pro- |

| | | | | |
|---|---|---|---|---|
| | | | binding (1) | cessing regulation of transcription response to DNA damage stimulus (1) |
| 3 | ZMYM3 | 4 | DNA binding metal ion binding zinc ion binding (1) | multicellular organismal development (1) |
| 4 | RAB41 | 2 | GTP binding nucleotide binding (1) | protein transport small GTPase mediated signal transduction (1) |
| 5 | OPHN1 | 1 | GTPase activator activity Rho GTPase activator activity actin binding (1) | actin cytoskeleton organization axon guidance endocytosis nervous system development signal transduction substrate-bound cell migration, cell extension (1) |

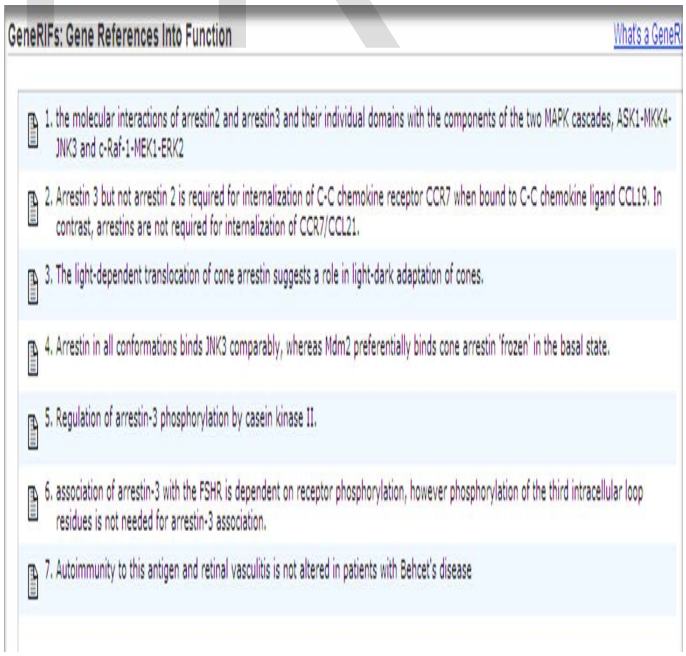
## 7 GENE ANALYSIS

The gene selected as a top candidate gene is ARR3 gene and this is because the above mentioned function of the arrestin gene identify it to be a suitable candidate gene for the primer design furthermore its function is associated with sensory perception, signal transduction and visual perception which all are the characteristics of the cones photoreceptors. Below is a diagram showing the different transcripts of the ARR3 gene. The ARR3 gene has been selected as top candidate gene to investigate in order to determine the gene responsible for the cone dystrophy in the family. The selection of the ARR3 gene has been determined by the number of EST it has as well as the expression and functionality. The ARR3 gene has 20 EST and was highly expressed in the retina furthermore literature search exposed its association with cone photoreceptors. The ARR3 gene falls in the following region Chromosome X: 69,488,155-69,501,690(2).




## 8 GENETIC BROWSERS

Below is a series of screenshots taken from the NCBI browser showing further details on the ARR3 gene. The ARR3 gene is associated with cone photoreceptors and its biological process is associated with visual perception, signal transduction and responses to a stimulus. There are several articles which provide evidence to prove that the ARR3 gene is associated with the cones photoreceptors as shown in the above screenshot. Below is a screen shot showing the expression location of the arrestin gene and as shown it's mainly and highly expressed in the eye.



**Names and origin**

| | |
|---|---|
| Protein names | *Recommended name:* **Arrestin-C** *Alternative name(s):* Cone arrestin; Short name=C-arrestin; Short name=cArr; Retinal cone arrestin-3; X-arrestin |
| Gene names | Name: **ARR3** Synonyms: ARRX, CAR |
| Organism | **Homo sapiens (Human)** |

| | |
|---|---|
| Biological process | Sensory transduction<br>Vision |
| Coding sequence diversity | Alternative splicing<br>Polymorphism |
| Technical term | 3D-structure<br>Complete proteome |

Gene Ontology (GO)

| | |
|---|---|
| Biological process | response to stimulus<br>Inferred from electronic annotation. Source: UniProtKB-KW<br>signal transduction Ref.1<br>Traceable author statement. Source: ProtInc<br>visual perception Ref.1<br>Traceable author statement. Source: ProtInc |
| Cellular component | cytoplasm Ref.1<br>Traceable author statement. Source: ProtInc<br>soluble fraction Ref.1<br>Traceable author statement. Source: ProtInc |

| | | |
|---|---|---|
| connective tissue | 0 | 0 / 149583 |
| ear | 0 | 0 / 16341 |
| embryonic tissue | 0 | 0 / 215827 |
| esophagus | 0 | 0 / 20211 |
| eye | 184 | 39 / 211508 |
| heart | 0 | 0 / 90297 |
| intestine | 0 | 0 / 235594 |
| kidney | 0 | 0 / 212554 |
| larynx | 0 | 0 / 24480 |

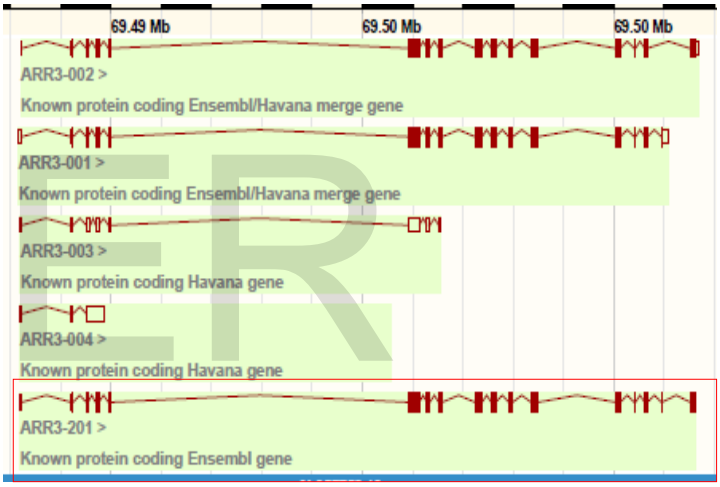
9

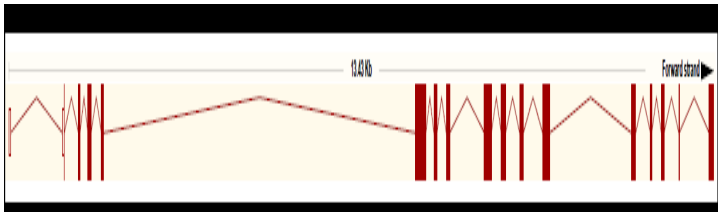
## TRANSCRIPTS WITHIN THE ARR3 GENE

Below is a table showing the number of transcripts within the ARR3 gene. Highlighted in red is the transcript selected to be analyzed in order to design primers for each exon it contains. The ARR3-201 transcript have been selected since it contains the highest number of exons furthermore it's a protein coding transcript and it covers most of the different exons found on other transcript in the ARR3 gene. The ARR3-201 transcript contains 17 different exons which require a forward and backward primer in order to amplify.

Below is an image of the ARR3-201 transcript selected for analysis and primer design.

| Name | Transcript ID | Protein ID | Description |
|---|---|---|---|
| ARR3-001 | ENST00000374495 | ENSP00000363619 | protein_coding |
| ARR3-002 | ENST00000374480 | ENSP00000363604 | protein_coding |
| ARR3-003 | ENST00000480877 | No protein product | processed_transcript |
| ARR3-004 | ENST00000477379 | No protein product | processed_transcript |
| ARR3-201 | ENST00000307959 | ENSP00000311538 | protein_coding |




## 10 PRIMERS FOR ARR3 GENE

Below is table of left and right primers designed to amplify the different exons on the ARR3-201 transcript. The primers have been designed using the geneomic sequence and CDNA sequence of ARR3-201 transcript.

| Exon No | Primer length (bp) | Product size | Left Primer | Right Primer |
|---|---|---|---|---|
| 1 | 38 | 171 | CCCCACCAATCAAAACACTC | AGGGATGCTAAGAGCAAGGG |
| 2-3 | 337 | 460 | GTATGCTGGAGTGGGGTGAG | TACCCACCTATCCCTTTTGC |
| 4-5 | 302 | 442 | GTCTCAGGGGTCTGGGTTG | CACTCATCGGGTTACCTGTC |
| 6-7 | 411 | 539 | TGACCCAGTTAAGATTCCCC | TGTCCCACTGTTCCTCTTCC |
| 8 | 68 | 192 | TTATAGGCCCATGAGGAAGG | GATCTGTTTCCTGGCAGAGG |
| 9-10 | 407 | 527 | GGGAGAGGTCTGTGGGTCTG | TTTTGGGAAAAGACCTGTGG |
| 11 | 73 | 200 | TGGGGTCTCAAGGATGACTG | ACTGCCTCCTGGACCTCTTC |
| 12 | 138 | 265 | CTCTTGGGAGTAAAAGTATGCTTAG | AGCATGTGGAATGGTTGAGG |
| 13-14 | 388 | 526 | AATGCATGAGAAGAAACTGGG | TCCTTATCCCCTTGTCCTCC |
| 15-16 | 343 | 481 | CTAAGGGAGGGAGGTTCAGG | TCGTCATAACATGGAGGGTG |
| 17 | 91 | 226 | AGGACACCCAGAAAGGGG | GCTAGGCACATCTGAACAAAC |

## 11 Conclusion

ARR3-201 transcript of ARR3 gene found in the region between the two different microsatellite markers provided has been analyzed and primers have been designed to amplify its exons. The ARR3 has been selected as a top candidate gene after an intense investigation using different genetic browsers and published articles. Evidence has been provided in the literature that this gene is localized in cone photoreceptors and its associated with different functions in the cones photoreceptors furthermore the expression rate was significantly high in the retina which provides more evidence that ARR3 gene is possibly associated with cone dystrophy identified within the family.



## References

[1] http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?taxid=9606&CHR=X&BEG=67074048&END=70613291&build=37
[2] http://www.ensembl.org/Homo_sapiens/Location/Overview?r=X:67251011-70436328
[3] http://eyebrowse.cit.nih.gov/cgi-bin/hgTracks?clade=vertebrate&org=Human&db=hg19&position=chrX%3A67251011-70436328&pix=620&hgsid=183&hgt.dummyEnterButton.x=0&hgt.dummyEnterButton.y=0
[4] http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene&cmd=retrieve&dopt=full_report&list_uids=407
[5] http://www.ensembl.org/Homo_sapiens/Transcript/Summary?db=core;g=ENSG00000120500;r=X:67251011-70436328;t=ENST00000307959
[6] http://www.uniprot.org/uniprot/P36575
[7] http://ihg2.helmholtz-muenchen.de/ihg/ExonPrimer.html
[8] http://www.ncbi.nlm.nih.gov/UniGene/ESTProfileViewer.cgi?uglist=Hs.308