UCLA

UCLA Previously Published Works

Title

Refinement of the Physical Location and the Genomic Characterization of the CRSP2 (EXLM1) Gene on Xp11.4

Permalink

https://escholarship.org/uc/item/2t83v1rr

Authors

Gorin, MB

Demirci, FY

Ramser, J

et al.

Publication Date

2023-12-11

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Taylor & Francis
Taylor & Francis Group

DNA Sequence

ISSN: 1042-5179 (Print) (Online) Journal homepage: http://www.tandfonline.com/loi/imdn19

Refinement of the Physical Location and the Genomic Characterization of the CRSP2 (EXLM1) Gene on Xp11.4

F. Yesim K. Demirci, Juliane Ramser, Nicola J. White, Brian W. Rigatti, Alfons Meindl, Karen F. Lewis, Gaiping Wen & Michael B. Gorin

To cite this article: F. Yesim K. Demirci, Juliane Ramser, Nicola J. White, Brian W. Rigatti, Alfons Meindl, Karen F. Lewis, Gaiping Wen & Michael B. Gorin (2003) Refinement of the Physical Location and the Genomic Characterization of the CRSP2 (EXLM1) Gene on Xp11.4, DNA Sequence, 14:2, 123-127, DOI: 10.1080/1042517021000056952

To link to this article: http://dx.doi.org/10.1080/1042517021000056952

Published online: 11 Jul 2009.

Submit your article to this journal

Article views: 4

View related articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=imdn19

Download by: [UCLA Library] **Date:** 27 April 2017, At: 16:35

DNA Sequence, 2003 Vol. 14 (2), pp. 123–127

Taylor & Francis
healthsciences

Short Communication

Refinement of the Physical Location and the Genomic Characterization of the CRSP2 (EXLM1) Gene on Xp11.4

F. YESIM K. DEMIRCI[a,b], JULIANE RAMSER[c], NICOLA J. WHITE[d], BRIAN W. RIGATTI[a], ALFONS MEINDL[c], KAREN F. LEWIS[d], GAIPING WEN[e] and MICHAEL B. GORIN[a,b,*]

[a]*Department of Ophthalmology, School of Medicine, University of Pittsburgh, EEINS Building, 203 Lothrop Street, Rm 1027 Pittsburgh, PA 15213, USA;* [b]*Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15213, USA;* [c]*Department of Medical Genetics, Ludwig-Maximilians-University, Munich, Germany;* [d]*Discovery Genetics, GlaxoSmithKline Research and Development, Stevenage, UK;* [e]*Department of Genome Analysis, Institute of Molecular Biotechnology, Jena, Germany*

(Received 27 August 2002)

In the course of our search for the gene responsible for X-linked cone-rod dystrophy (COD1), we constructed a physical map and contig (encompassing the region between DXS556 and DXS228), and identified sequences showing homologies to the expressed sequence tags (ESTs) that matched CRSP2 (EXLM1) transcript. We confirmed the expression of the CRSP2 gene in the retina and refined its exact genomic location between DXS1368 and DXS993. We demonstrated that the entire transcript is encoded within 31 exons. Primers were designed for mutation analysis of the exons by direct sequencing of PCR products from genomic DNA, and revealed no mutations in COD1 families. We subsequently excluded CRSP2 as a candidate for COD1 by demonstrating the causative mutations in the RPGR. However, due to its expression in different tissues and its contribution to transcriptional regulation, CRSP2 may be a candidate for other diseases that map to this region of the X chromosome.

Keywords: CRSP2; EXLM1; Xp11.4; Cone dystrophy; RPGR; Transcriptional regulation

X-linked cone-rod dystrophy (COD1; MIM 304020) is a rare, progressive visual disease primarily affecting the cone photoreceptors (Pinckers and Timmerman, 1981; Pinckers and Deutman, 1987; Jacobson *et al.*, 1989). In the course of our search for the COD1 gene (Hong *et al.*, 1994; Seymour *et al.*, 1998; Demirci *et al.*, 2001), we constructed a physical contig encompassing the region between DXS556 and DXS228 (Xp11.4), comprised of P1-derived artificial chromosomes (PACs), bacterial artificial chromosomes (BACs) and yeast artificial chromosomes (YACs). This complete physical contig was used for high throughput sequencing and BLAST searches to identify homologies with expressed sequence tags (ESTs) and cDNAs in GenBank. This enabled us to identify several EST hits (originating from different tissues including retina) that corresponded to CRSP2 transcript (NM_004229, 7984 nt) (MIM 300182, the cofactor required for Sp1 transcriptional activation subunit 2). This transcript was originally designated as EXLM1 (AB006651) when it was characterized and mapped to Xp11.2-p11.4 (Yoshikawa *et al.*, 1998).

Previous studies indicate that CRSP2 (EXLM1) is conserved through evolution in mammals and escapes X-chromosome inactivation (Yoshikawa *et al.*, 1998). The vitamin D receptor-interacting protein complex component DRIP150 and the thyroid hormone receptor-associated protein TRAP170 mRNAs are also identical to CRSP2 (EXLM1) mRNA, and were identified in multi-protein complexes involved in transcriptional regulation (Hittelman *et al.*, 1999; Rachez *et al.*, 1999; Ryu *et al.*, 1999). Therefore, as a potential candidate for COD1, we characterized CRSP2 (EXLM1) genomic structure and designed primers for mutation analysis of the exons in COD1 families.

We performed PCR with gene specific primers (forward: 5′-CCAATCTGGAGATTCCACATCAA-3′, reverse: 5′-TTAATAAGCGAATTGCATGG-3′) to confirm the expression of the CRSP2 (EXLM1) gene

*Corresponding author. Tel.: +1-412-647-2205. Fax: +1-412-647-5880. E-mail: gorinmb@msx.upmc.edu

ISSN 1042-5179 print/ISSN 1029-2365 online © 2003 Taylor & Francis Ltd
DOI: 10.1080/1042517021000056952

in human retina using retina cDNA libraries, and to determine its exact genomic location on Xp11.4 using the genomic clones comprising our physical contig. The PCR conditions for genomic DNA amplification were: initial denaturation at 95°C for 5 min, followed by 35 cycles of 45 s of denaturation at 95°C, 10 s of annealing at 55°C and 60 s of extension at 72°C, and a final extension at 72°C for 10 min. PCR was carried out in a 20 ul reaction volume containing 1 × PCR buffer II, 1.5 mM $MgCl_2$, 200 uM each dATP, dCTP, dGTP, dTTP, 0.5 uM each forward and reverse primer and 0.04 U/ul AmpliTaq Gold (Applied Biosystems, Foster City, CA). The retinal expression of the CRSP2 (EXLM1) gene was confirmed in two independent retina cDNA libraries (Fig.1). The exact location of the gene was refined to be between DXS1368 and DXS993 on Xp11.4 (the order is tel-DXS556-DXS8042-DXS6821-DXS1368-CRSP2 (EXLM1)-DXS993-DXS228-cen). The physical map and contig spanning the region between the markers DXS6821 and DXS993 is shown in Fig. 2.

The cDNA transcript was compared with the fragmented, high throughput genomic sequencing data from our contig to specify intron–exon junctions in silico. Two overlapping genomic clones [PAC16186 (AC093029) and RP11-169L17 (AC092474)] that were identified to harbor the entire CRSP2 (EXLM1) gene were further sequenced in order to obtain a complete genomic information. Thirty-one exons were determined and these exons were amplified and screened for mutations using the primers listed in Table I. CRSP2 (EXLM1) has a 4365-nt coding sequence (CDS: 120-4484) encoding a protein of 1453 amino acids (when the start and the stop codons are not considered). The coding region includes exons 2 through 30 and parts of the exons 1 and 31, indicating that all exons contribute to the translated protein. Although 5′ UTR appears to be short, exon 31 contains a very long 3′ UTR (3500 bp).

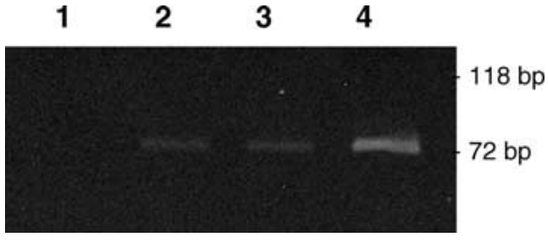

FIGURE 1 Confirmation of the expression of CRSP2 (EXLM1) gene in the retina. Aliquots from two different retina cDNA libraries were subjected to 35 cycles of PCR with gene specific primers (expected PCR fragment size: 72 bp, location within the cDNA: exon 16). The cDNA libraries were also tested with the intronic primers flanking exon 16 to demonstrate that there was no genomic DNA contamination. Amplification was observed in only genomic DNA but not in the libraries (Data not shown). Lanes marked by numbers indicate the template, 1 = distilled water (negative control), 2 = Lambda gt10 retina cDNA library (from Dr Jeremy Nathans), 3 = Stratagene retina cDNA library, 4 = Genomic DNA (positive control).

PCR primers flanking the exons were designed in order to generate PCR fragments that would also include the intron–exon junctions for detecting the splice site mutations (Table I). The exons were amplified from leucocyte genomic DNA from six COD1-affected males (2 affected males from each of the 3 families) and two unaffected males, and screened for mutations by direct sequencing of PCR products in both strands with either ABI377 or ABI3700 automated sequencer. The PCR conditions for genomic DNA amplification were as follows: initial denaturation at 95°C for 5 min, followed by 35 cycles of 45 s of denaturation at 95°C, 10 s of annealing at 53–58°C (Table I) and 60 s of extension at 72°C, and completed with a final extension at 72°C for 10 min. PCR was done in a standard 50 ul reaction volume containing 1 × PCR buffer II, 1.5 mM $MgCl_2$, 200 uM each dATP, dCTP, dGTP, dTTP, 0.5 uM each forward and reverse primer, and 0.04 U/ul AmpliTaq Gold (Applied Biosystems). After gel checking, PCR products were purified using QIAquick PCR purification kit (Qiagen, Valencia, CA) and used as template for sequencing reactions. No causative mutations for COD1 were identified by the sequence analysis of the exons with the flanking intronic sequences. The sequence variants observed during our analyses were as follows: 644T>C (Asp175Asp), 815C>A (Ala232Ala), 3912T>G (Leu1265Val), 4471T>G (Val1451Gly) and 4769T>C (3′UTR). Four of these variants (644T>C, 3912T>G, 4471T>G, and 4769T>C) were detected both in affected and unaffected males. The search in dbSNP revealed two SNP cluster ID's for 4769T>C (rs3185561 and rs1128513). The allele frequencies for 644T>C and 815C>A in thirty X chromosomes from healthy female subjects were 0.76 and 0.03, respectively. 3912T>G and 4471T>G were observed in all chromosomes we screened (patients and healthy female subjects), indicating that the "T" allele that is present at these positions in some of the GenBank clones (but not detected in our samples) is either a rare allele in the North American population or a sequencing artifact.

There are currently eight mRNA sequences related to CRSP2 (EXLM1) in GenBank (AB006651.1, NM_004229.1, AF304448.1, AF135802.1, AF104256.1, AK023368.1, AF070563.1, AJ012077.1). AJ012077 represents an alternatively spliced form with the partial retention of intron 30 as the 3′ end of the mRNA and is also confirmed with the identification of the overlapping ESTs in GenBank. BLAST search in the human EST database of GenBank using the full-length cDNA sequence (7984 nt) reveals over 300 EST hits from several different tissues including brain, eye, skin, lung, kidney, muscle (skeletal), prostate, ovary, colon, small intestine and bone marrow. Although our analyses of the exons and

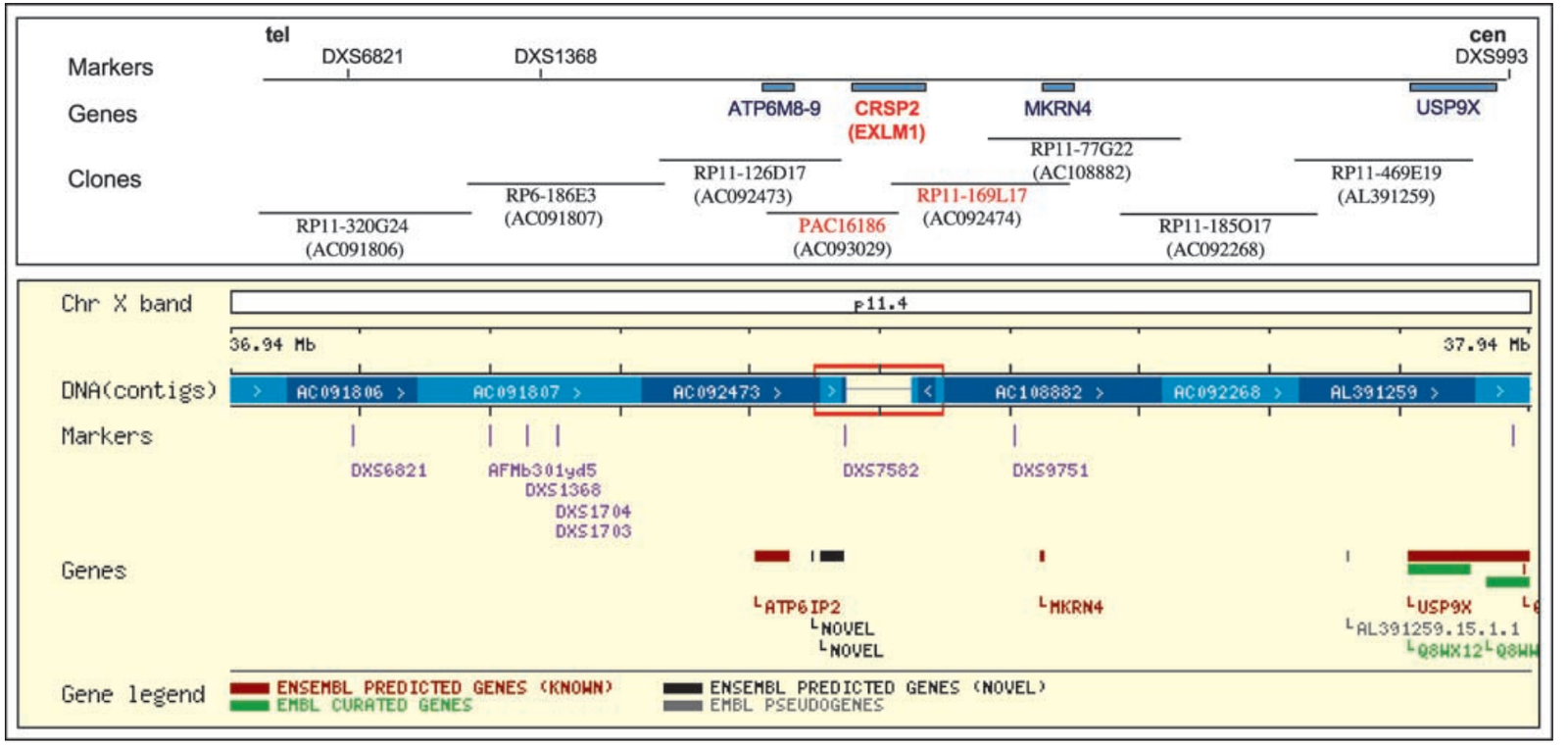

FIGURE 2 Comparison of our physical map and contig with the Ensembl view (http://www.ensembl.org/) of the same region on Xp11.4. Note the location of the CRSP2 (EXLM1) gene and the genomic clones which were used for high throughput sequencing. The contig represents a subset of clones isolated from two different centers and which has been submitted to GenBank. PAC16186 was isolated in the UK (N.J.W. and K.F.L.) and the other clones except AL391259 were isolated and sequenced in Germany (J.R., A.M. and G.W.).

TABLE I Position and nucleotide sequence of intron/exon boundaries and PCR primer pairs for EXLM1 exons

Exon	Position in cDNA (bp)	Acceptor Splice Site	Donor Splice Site	Forward Primer (5′ > 3′)	Position relative to Exon	Reverse Primer (5′ > 3′)	Position relative to Exon	PCR fragment length (bp)	Annealing temperature (°C)
1	1–334		GACCT/gtaagtg	GGCGGACTTCTCGCTACAT	−63	AGGGAGGTGCAAACGTCTC	+75	472	53
2	335–361	ttcttag/ACTGC	GAAAG/gtgagta	GCTGTTTACCCAGTTTTTCTGA	−118	CAGCCCTACACCACCCATAC	+86	231	53
3	362–467	tttatag/GAAAA	GTGCG/gtgagtt	TGGTTGTAAGCAAAGGCTGA	−112	TCACTCAAATCAGCAGGGATA	+105	323	55
4	468–641	tttaaag/ATGAT	TTAGG/gtaagtt	CCCCTTGAAGTTTACTTTGGA	−96	AATCCAATCCAATCCAGTGC	+66	336	55
5	642–771	tttgcag/GATAA	AGTTG/gtatgta	CGCAACACTGTATAAAGACAGGA	−86	TCAGTCTCCCAAAGTGCTGA	+85	301	55
6	772–900	ctaaaag/CAAAT	AGGAG/gtaaatc	TGAAAATTCTAGGTCAGGAGGA	−79	CTTGGCCTCCCAAAGTAGTG	+126	334	55
7	901–1008	acggaag/ATGGG	CCTAC/gtatcct	CCGTGGAAGTTTCAGTGTTTC	−164	GGCAAGGAATGTAGGCTGCT	+107	379	55
8	1009–1141	ctatcag/ATTCT	TGGAA/gtaagta	TTACGATGAGGCATTGGAGA	−155	GCCCTATACGCAGGTTTCAC	+178	466	55
9	1142–1292	tctttag/TCAAC	TGAAG/gtaaaag	TTCCAGGCTGTGGCTAAACT	−101	GGTGCTCTAATGCCAGACCT	+262	514	56
10	1293–1404	tatgcag/ATCGA	AAACT/gtaggtt	TCAGTGTGATTCCAAAAGGA	−110	TGCAGAAATACCAGCTTCCA	+127	349	53
11	1405–1530	tgttcag/CTTCC	ACTTG/gtaagtt	GGATCCTTCAATAGTTGGTTTG	−101	ATACCCACTTTCAACACTGCAT	+122	349	56
12	1531–1609	tctgtag/ACCAG	CTTAA/gtagctt	GCCAGTGTACATTTGCCATT	−242	TCTAAGGTACCAAACGGCACT	+142	463	56
13	1610–1769	aattcag/GTTTT	ACATT/gtaagta	TCATTATGGCCATGTGTCCT	−145	CCCCTGGGTTTCACTATACC	+129	434	53
14	1770–1964	tttttag/GTTGT	GCAAG/gtatgat	ATTCTGAGCCTTACTTGATGATG	−86	CCTTATTGCTGCCTCGTCTC	+122	403	56
15	1965–2099	attctag/TTGTC	TGGAG/gtaagtt	TTGTCCACCATTTTCCATGA	−213	ACCCTACCACTCCTCCCAAA	+137	485	56
16	2100–2176	tttctag/TTGTC	TTAAA/gtaagcc	AGGTGCCACAAGAATAGAGAC	−91	AGAAGCAGCTTTCAGAGTGTG	+65	233	56
17	2177–2337	tcaccag/AATTC	GCAAG/gtgagat	GCTGTAATGTCATACCTTGTCTG	−86			565	56
18	2338–2484	cctgtag/GACCA	ACCAG/gtacgtg			CTGTGCATGTTTCACTTCAGTT	+85		
19	2485–2576	tgtttag/ACATA	GCTCA/gtaagtc	TTATCTGTTTGTGCTGTTGCTT	−157	TCCCTAACTTTTACTTGAGCTCTTA	+112	361	54
20	2577–2729	tatttag/ATTAG	TACAG/gtaatta	AACAGCACATCTGGAGTTCG	−166	AGCACGGCTGGCATACTAAG	+53	372	56
21	2730–2999	cttatag/GTACT	TAAAG/gtaacag	CTCTTCAAGGGGAAGTTTTATG	−132	TTTGTTGCTTTTCCCCATTA	+92	494	54
22	3000–3140	tccgtag/ACGTT	AACAG/gtaatat	TTCCCGCTTCAAAATTAACA	−160	CAGAAATGTGATACCACTTGGA	+40	341	53
23	3141–3252	ttttcag/CCATT	TCCAG/gtaagat	AAGACTTGTGCTGGTGATGTG	−174	TTCTCAATGTGGGGTGTGTG	+74	360	56
24	3253–3384	ttattag/GAAAT	ACATG/gtgagtc	GATTTTTAACTTGGGGTCCTTG	−78	CAATATTTTTGCTACTTCCTCAA	+120	330	54
25	3385–3567	tttatag/GAACT	AACAA/gtaagta	CTGTGTAAAGATTTGAATGCTGTC	−84	GAGGATGCCAACAAGATAGG	+70	337	56
26	3568–3803	ccaaaag/GTTCT	AAACG/gtatggg	ATGATGAAAGCAAAACAGGA	−75	AATGTCACGTTAGTTCAGATCC	+91	402	58
27	3804–3983	tttgtag/CTGCA	CAAGA/gtatgtg	AACTTAGTGGTTTTCCTTTTGG	−92			569	55
28	3984–4091	cctgtag/GTTGC	AGCTG/gtaagtc			CTACATAGATGGGTGAAAACCA	+94		
29	4092–4217	aaaatag/TTCCC	TTTTT/gtaagta	TTCTAACCTCTCAGACTTTCCTACC	−83	AGGAATCCTGGCCTCTATTTG	+123	332	56
30	4218–4410	tatacag/CTTCA	ACAAG/gtacatg	GACAGATGTTTTATGATGGGTTT	−53	AGGAGAAAGGGGTGTCTCTCA	+73	319	53
31	4411–7984	gttgcag/GTGAA		CCACAGCTTAATCCCCAATA	−159	TTGCCAGTTTAGAATATTTAGCTC		363	56
				TTTTTAAACCAGGAAGGCTGAC		TCCACCAGGTTAAAGATGAAGG		600	56
				CACCCCAAGGAAGCAATATTTA		ATTCTTCACAATGAGGGGAATG		632	56
				TGCATGAAGTTTCTGTGAGCTT		GTGGGAGATGAAGGCAGAAATA		607	56
				AGCCTTCTTTGCCTTTCTTCTT		GATCGAACCTGGGTAACAAAAG		632	56
				TGCTTGGTACTTGGAATTGA		CAGTCCAAAGAATGGCTGTTAG		603	56
				CAGTAATCTGTTCAGCCAAAGG		TGCCTAGCACACAGTAAGCTCT		599	56
				ACTTCCAAGTGGCCAGTGATT		AGCACCTAGCAAAGGTTAGACG	+153	600	56

Sizes of the exons for EXLM1 and the forward and reverse primers (5′- > 3′) generating PCR fragments including exons and flanking intron–exon junctions. Exons 17&18 and 27&28 were amplified together with incorporation of a small intron between the exon pairs. Exon 31 was amplified and sequenced using 8 primer pairs, producing 8 overlapping fragments. The locations for the primers (for the first 5′-nt of the primer) are also indicated (the first nt preceding the 5′-end of the exon being −1 and the first nt following the 3′-end of the exon being +1).

flanking intronic sequences did not reveal any disease-causing mutations, there remained the possibility of the mutations outside of these regions that could affect the splicing or stability of the RNA transcript. Nevertheless, we subsequently excluded CRSP2 (EXLM1) as a candidate, by remapping COD1 to Xp11.4-p21.1 and identifying the causative mutations in the RPGR exon ORF15 (Demirci *et al.*, 2002). However, the expression of CRSP2 (EXLM1) in retina and other tissues and its contribution to the transcriptional regulation, may be relevant for other diseases that map to this region of the X chromosome.

Acknowledgements

We are grateful to the families who have participated in this study. This research was supported by the NIH Grant EY13130 (MBG); the NIH Core Grant for Vision Research EY08098; the Eye & Ear Foundation of Pittsburgh, Pittsburgh, PA (MBG); Fight For Sight, Research Division of Prevent Blindness America (FYKD); Research to Prevent Blindness, NY, NY; and the German Ministry for Research and Education KW9974 (AM).

References

Demirci, F.Y., White, N.J., Rigatti, B.W., Lewis, K.F. and Gorin, M.B. (2001) "Identification, genomic structure, and screening of the vacuolar proton-ATPase membrane sector-associated protein M8-9 gene within the COD1 critical region (Xp11.4)", *Molecular Vision* **7**, 234–239.

Demirci, F.Y., Rigatti, B.W., Wen, G., Radak, A.L., Mah, T.S., Baic, C.L., Traboulsi, E.I., Alitalo, T., Ramser, J. and Gorin, M.B. (2002) "X-linked cone-rod dystrophy (locus COD1): Identification of mutations in RPGR exon ORF15", *American Journal of Human Genetics* **70**, 1049–1053.

Hittelman, A.B., Burakov, D., Iniguez-Lluhi, J.A., Freedman, L.P. and Garabedian, M.J. (1999) "Differential regulation of glucocorticoid receptor transcriptional activation via AF-1-associated proteins", *EMBO Journal* **18**, 5380–5388.

Hong, H-K., Ferrell, R.E. and Gorin, M.B. (1994) "Clinical diversity and chromosomal localization of X-linked cone dystrophy (COD1)", *American Journal of Human Genetics* **55**, 1173–1181.

Jacobson, D.M., Thompson, H.S. and Bartley, J.A. (1989) "X-linked progressive cone dystrophy, Clinical characteristics of affected males and female carriers", *Ophthalmology* **96**, 885–895.

Pinckers, A. and Deutman, A.F. (1987) "X-linked cone dystrophy, An overlooked diagnosis?", *International Ophthalmology* **10**, 241–243.

Pinckers, A. and Timmerman, G.J. (1981) "Sex-difference in progressive cone dystrophy. I.", *Ophthalmology and Paediatric Genetics* **1**, 17–24.

Rachez, C., Lemon, B.D., Suldan, Z., Bromleigh, V., Gamble, M., Naar, A.M., Erdjument-Bromage, H., Tempst, P. and Freedman, L.P. (1999) "Ligand-dependent transcription activation by nuclear receptors requires the DRIP complex", *Nature* **398**, 824–828.

Ryu, S., Zhou, S., Ladurner, A.G. and Tjian, R. (1999) "The transcriptional cofactor complex CRSP is required for activity of the enhancer-binding protein Sp1", *Nature* **397**, 446–450.

Seymour, A.B., Dash-Modi, A., O'Connell, J.R., Shaffer-Gordon, M., Mah, T.S., Stefko, S.T., Nagaraja, R., Brown, J., Kimura, A.E., Ferrell, R.E. and Gorin, M.B. (1998) "Linkage analysis of X-linked cone-rod dystrophy: localization to Xp11.4 and definition of a locus distinct from RP2 and RP3", *American Journal of Human Genetics* **62**, 122–129.

Yoshikawa, H., Fujiyama, A., Nakai, K., Inazawa, J. and Matsubara, K. (1998) "Detection and isolation of a novel human gene located on Xp11.2-p11.4 that escapes X-inactivation using a two-dimensional DNA mapping method", *Genomics* **49**, 237–246.