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
2017-05-16

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


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
Short Communication

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

F. YESIM K. DEMIRCIa,b, JULIANE RAMSERc, NICOLA J. WHITEd, BRIAN W. RIGATTIa, ALFONS MEINDLc, KAREN F. LEWIsd, GAIPING WENE and MICHAEL B. GORINa,b,*

aDepartment of Ophthalmology, School of Medicine, University of Pittsburgh, EEINS Building, 203 Lothrop Street, Rm 1027 Pittsburgh, PA 15213, USA; bDepartment of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15213, USA; cDepartment of Medical Genetics, Ludwig-Maximilians-University, Munich, Germany; dDiscovery Genetics, GlaxoSmithKline Research and Development, Stevenage, UK; eDepartment 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
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 53–58°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 20μl reaction volume containing 1 × PCR
buffer II, 1.5 mM MgCl₂, 200 uM each dATP, dCTP,
dGTP, dTTP, 0.5 uM each forward and reverse primer
and 0.04 U/μl AmpliTaq Gold (Applied Biosystems,
Foster City, CA). The retinal expression of the CRSP2
(EXLM1) gene was confirmed in two indepen-
dent 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 (AC093209) 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
coding region 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).

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 μl
reaction volume containing 1 × PCR buffer II, 1.5 mM MgCl₂, 200 uM each dATP, dCTP, dGTP,
dTTP, 0.5 uM each forward and reverse primer, and
0.04 U/μl 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 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).
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>
<thead>
<tr>
<th>Exon</th>
<th>Position in cDNA (bp)</th>
<th>Acceptor Splice Site</th>
<th>Donor Splice Site</th>
<th>Forward Primer (5'-3')</th>
<th>Position relative to Exon</th>
<th>Reverse Primer (5'-3')</th>
<th>Position relative to Exon</th>
<th>PCR fragment length (bp)</th>
<th>Annealing temperature (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1–334</td>
<td>tac/tgag/ACTGC</td>
<td>GACCT/gtaagtg</td>
<td>GGGGAGTCTTCGCTACAT</td>
<td>–63</td>
<td>AGGGGAAGTTCACCAACACCCT</td>
<td>+75</td>
<td>472</td>
<td>53</td>
</tr>
<tr>
<td>2</td>
<td>335–361</td>
<td>GAAAG/gtaagaag</td>
<td>GCTGTTACCTACGATTTCGA</td>
<td>–118</td>
<td>CACGCTGCTACCAACACCCT</td>
<td>+86</td>
<td>253</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>362–467</td>
<td>TGCC/gtaagaag</td>
<td>GCTGTTACCTACGATTTCGA</td>
<td>–112</td>
<td>TACCTGACCAATCAGGAGTTTA</td>
<td>+103</td>
<td>325</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>468–641</td>
<td>TTAGG/gtaaggt8</td>
<td>CCCCCGGATAGTTACCTTTGGA</td>
<td>–96</td>
<td>AATCCATCTACCCATCTCAG</td>
<td>+66</td>
<td>336</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>642–771</td>
<td>GGTTAG/GATGtaata</td>
<td>CCCGACCTCAACTTAAAGCTAGA</td>
<td>–86</td>
<td>TACGCTCTGACCAATCAGGAGTTTA</td>
<td>+85</td>
<td>301</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>772–900</td>
<td>AAAAG/gtaataat</td>
<td>TGAATAATCTTACGAGGGAGA</td>
<td>–79</td>
<td>TTGGGAGTCACCAATCAGGAGTTTA</td>
<td>+126</td>
<td>354</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>901–1008</td>
<td>AAGGAG/AATGCC</td>
<td>CCGTGCTTATGCTTTGTTT</td>
<td>–164</td>
<td>GCAGAAAGTATGAGCCTCAG</td>
<td>+107</td>
<td>379</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1009–1141</td>
<td>ATACAG/ATCTT</td>
<td>TTGAGAGGTACCTTTGGAAGA</td>
<td>–155</td>
<td>GCCCTTATACCGAGCTTCAG</td>
<td>+178</td>
<td>466</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1142–1292</td>
<td>TCAATG/GAAGagaag</td>
<td>TTTGAGGAGGTTTCTAGCTTCA</td>
<td>–101</td>
<td>TCTAATGCTAATAGCAGG</td>
<td>+262</td>
<td>514</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1293–1404</td>
<td>AACCT/gtaatta</td>
<td>TACGCTGCTACCAATCAGGAGTTTA</td>
<td>–130</td>
<td>TCTAATGCTAATAGCAGG</td>
<td>+127</td>
<td>349</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>1405–1530</td>
<td>TGGCA/GAAGagaag</td>
<td>GCTGTTACCTACGATTTCGA</td>
<td>–101</td>
<td>TACCTGACCAATCAGGAGTTTA</td>
<td>+129</td>
<td>349</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1531–1609</td>
<td>ACCAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–242</td>
<td>TCTAATGCTAATAGCAGG</td>
<td>+142</td>
<td>463</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>1610–1769</td>
<td>ACAGT/GTTACTT</td>
<td>TCTGTTACCTACGATTTCGA</td>
<td>–145</td>
<td>CTTGGGAGTCACCAATCAGGAGTTTA</td>
<td>+129</td>
<td>434</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1770–1964</td>
<td>TTTGAG/GAAGagaag</td>
<td>TTTGAGGAGGTTTCTAGCTTCA</td>
<td>–101</td>
<td>TCTAATGCTAATAGCAGG</td>
<td>+127</td>
<td>349</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1965–2099</td>
<td>ATACAG/ATCTT</td>
<td>TTGAGAGGTACCTTTGGAAGA</td>
<td>–155</td>
<td>GCCCTTATACCGAGCTTCAG</td>
<td>+178</td>
<td>466</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>2100–2176</td>
<td>TTTTAC/GTGTTCT</td>
<td>TTTGAGGAGGTTTCTAGCTTCA</td>
<td>–101</td>
<td>TCTAATGCTAATAGCAGG</td>
<td>+142</td>
<td>463</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>2177–2338</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2338–2484</td>
<td>ACCAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>2485–2576</td>
<td>ACCAG/TTGTAACATACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2577–2729</td>
<td>AAGCT/GTTACTTCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>2720–2999</td>
<td>AAGCT/GTTACTTCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>2992–3140</td>
<td>AAGCT/GTTACTTCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>3141–3252</td>
<td>ATACAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>3253–3384</td>
<td>ATACAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>3385–3567</td>
<td>ATACAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>3568–3703</td>
<td>ATACAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>3704–3803</td>
<td>ATACAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>3804–3983</td>
<td>ATACAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>3984–4091</td>
<td>ATACAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>4092–4217</td>
<td>ATACAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>4218–4410</td>
<td>ATACAG/ATCTCCTT</td>
<td>CAGCTGAGCTTACCTTTGGAAGA</td>
<td>–86</td>
<td>TTCTAGCTGCTACCAATCAGGAGTTTA</td>
<td>+122</td>
<td>403</td>
<td>56</td>
<td></td>
</tr>
</tbody>
</table>

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, 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


