J. Appl. Genet. 42(1), 2001, pp. 89-102

Large-scale analysis of sequence tags in Xp11.4-11.3 and evaluation of candidate genes for X-linked ocular diseases

Carsten M. PUSCH

Molecular Genetics Laboratory, University Eye Hospital II, Tübingen, Germany

Abstract. The gene-rich region of Xp11.4-Xp11.3 was characterized by increasing the physical marker density. Sequence tags (STSs) were generated by IRS- and DOP-PCR techniques, subsequent cloning, sequencing, and creation of primer pairs for single-copy sites. A total of 224 novel STSs were collected, providing an average marker density of 18 kb in the Xp11.4-Xp11.3 region which is assumed to be ~4 Mb in size. Sequence analysis of generated and established STSs via data base searches identified a novel gene highly homologous with the protein phosphatase 1 inhibitor 2 (*IPP-2*) and two pseudogenes; all of which map to the ~1.5 Mb proximal region of the critical region for X-linked congenital stationary night blindness type I (CSNB1) between markers DXS993 and DXS228. Using well-defined DNA panels, 69 STSs were fine-mapped to this ~1.5 Mb region, providing a marker coverage of one marker per 22 kb. No allelic loss was observed when the total STS content was applied to patient DNAs by PCR-mediated amplification. However, given the association of this region with a number of inherited ocular diseases, the data presented here provide valuable tools for genetic linkage and large-scale association studies.

Key words: genetic diseases, human X-chromosome, IPP-2, night vision, physical map, polymerase chain reaction, pseudogene, sequence tags

Introduction

Although the relative density of X-chromosomal DNA markers has steadily increased over the past years, prominent gaps remain to be closed. This specially holds true for regions on the short arm of the human X chromosome. One of them is the gene-rich region Xp11 where a number of ocular disorders displaying severe

Received: September 5, 2000. Accepted: December 5, 2000.
Correspondence: C.M. PUSCH, Molekulargenetisches Labor, Universitäts-Augenklinik, Auf der Morgenstelle 15, 72076 Tübingen, Germany, e-mail: carsten.pusch@uni-tuebingen.de.

phenotypes have been mapped genetically (Figure 1). However, intense mapping efforts are required to obtain sequence data that may be associated with the respective causative genes. For example, X-linked congenital stationary night blindness (XLCSNB) represents a distinct subgroup of non-progressive night vision disorders. The characteristic night vision defect from birth XLCSNB is frequently associated with other clinical problems such as high myopia, nystagmus, decreased visual acuity and occasionally strabismus and optic disc abnormalities. Various linkage studies have narrowed the critical interval for the complete form of CSNB (i.e. CSNB1) to a consensus interval between markers DXS556 and DXS228 (i.e. 3-5 cM) (BERGEN et al. 1995, THISELTON et al. 1995, HARDCASTLE et al. 1997, BOYCOTT et al. 1998, ROZZO et al. 1999). Moreover, a number of other ocular or retinal disorders map also to the same region or into close vicinity (Figure 1). Extensive positional cloning efforts, however, have not yet led to the identification of the causative gene for CSNB1 or of any other predicted disease-related genes, largely owing to a lack of sequence information and useful markers within human Xp11.4-Xp11.3.

For a better characterization of the interval in Xp11.4-11.3, sequence tagged sites (STSs) were generated by interspersed repeat sequence (IRS-) and degenerate oligonucleotide-primed (DOP-) polymerase chain reaction (PCR) techniques. The novel physical markers were assigned to the candidate region and their chromosomal Xp-sublocalization was determined by remapping them to a set of specific and precisely typed templates. The STSs were further used as a source for homology searches at public data bases in order to attain additional sequence information for the identification of potential candidate genes. In addition to the newly defined markers a set of established markers and cDNAs was used for deletion screening in CSNB1 patients by PCR and Southern hybridization.

Methods

Generation and characterization of STSs

Thirteen primer sequences for Inter-*Alu*-, Inter-*Kpn*-, and DOP-PCR were used:

5-GTCCTGCGGCCGCCTCCCAAAGTGCTGGGATTACAG-3,
5-AGGCTGGATCCTGAGCYRWGATYRYRCCAYTGC-3,
5-TTGCAGTGAGCCGAGATCGCCN-3,
5-GGCGATCTCGGCTCACTGGAAN-3,
5-CTCCCAAAGTGCTGGGATTACAGG-3,
5-GAGGTTGCAGTGAGCYRAGAT-3,
5-GAGGCTGCAGTGAGCYRTGAT-3,
5-GAGYRAGACTCYRTCTCAAAAAAA-3,
5-GATGTGCTGCTGGATTCAG-3,
5-GGACATGAACAGACAGTATN-3,

5-CATCACGTTGTACACCTTAAATAT-3,
5-CAATTGTCTGCTGTGAGTKTAYAATT-3;and
5-CCGACTCGANNNNNNATGTGG-3.

Two-step IRS-PCRs were carried out using 80 ng DNA and 20 pmol of each primer in various combinations in a standard 50 µl reaction. The temperature profile used for amplification on a GeneAmp 9600 PCR cycler (ABI Perkin-Elmer, Weiterstadt, Germany) was 5 minutes at 96°C for 1 min, 10 cycles at 94°C for 30 sec, 72°C for 3 min, 15 cycles at 94°C for 20 sec, 68°C for 2 min, and a final step at 72°C for 10 min. For DOP PCRs (exclusively applied onto PFGE-separated and gel-extracted YACs) the annealing temperatures used were 50°C and 40°C.

Cloning of PCR products into the pCR®2.1 vector was performed as recommended by the manufacturer (Invitrogen, Groningen, The Netherlands). The recombinants were sequenced using M13 forward/reverse primers and ABI BigDye terminator cycle sequencing chemistry on an ABI 377 automated sequencer. Sequence analysis was performed with the ABI analysis software and the Lasergene package (DNAStar Inc., Madison, WI, USA). The software program REPEATMASKER was employed to identify single copy regions, while PRIMER (GCG) was used to generate unique primer pairs for amplification (Table 1). STSs were remapped by PCR to (i) a whole genome panel, (ii) a subregional panel of the X chromosome (both Coriell Cell Repositories, Camden, NJ, USA), (iii) a rodent/human radiation-reduced hybrid panel covering the Xp11.4-Xp11.3 region (BERGER et al. 1992), (iv) clones from two yeast artificial chromosome (YAC) libraries (Research Genetics, USA; RZPD, Germany), and (v) a collection of 15 previously typed YACs. Prior to that, libraries (ii) to (v) had been precisely typed with 19 published markers specific to the Xp11.3-Xp11.4 location and flanking regions: tel-DXS989, DXS1058, DXS556, DXS361, DXS8042, DXS1015E, DXS1368, DXS574, DXS993, DXS8012, DXS1207, DXS1201, DXS8085, DXS228, DXS7, MAOA/B, DXS8080, DXS8083, and DXS1003-cen. This was necessary to allow a precise remapping of the newly generated STSs to these templates. All PCR mixes and cycle-sequencing reactions were prepared using the pipetting workstation model BioRobot 9600 (Qiagen, Hilden, Germany).

Sequence analysis and data base searches

Novel STSs were characterized by submitting their single copy portion to homology searches (e.g. BLAST, FASTA, HOMOLOGY, GAP; GCG) at NCBI, EMBL and hgmp data bases. Matches displaying >90% identity on nucleic acid, and >80% similarity on amino-acid level were analysed by sequence comparison and multiple sequence alignments (MALIGN, CLUSTAL, PileUp; GCG) in THC (TIGR) and UNIGENE clusters, using various software packages. The NIX anal-

ysis package at hgmp was applied for identification and characterization of larger DNA segments.

Additionally, data base sequence information of established markers was retrieved from GenBank in NCBI and subsequently submitted to hgmp, EMBL, and the TIGR Human Gene Index Data base for tentative human consensus sequence analysis. Primer pairs were designed for a collection of cDNA fragments, STSs, expressed sequence tags (ESTs) and full length cDNAs, putatively localized in Xp between the markers DXS556 and DXS228.

Patients and mutation screening

Nine male CSNB1 patients selected from the RetDis repository of the University of Tübingen were analysed in this study. Patients were diagnosed as having the CSNB1 phenotype based on clinical examination as well as electroretinographical and psychophysical testing. Defective rod scotopic vision was demonstrated in all patients through electroretinographical recordings (ERGs) generated in response to a low luminance flash light in dark adaptation. All patients showed a negative ERG, i.e. a normal 'a' wave and a diminished 'b' wave typical of the type II or Schubert-Bornschein type of ERG, absent oscillatory-potential wavelets and a relatively normal 30Hz flicker response. Dark-adaptation curves were monophasic with an elevated final threshold indicating a defect in the rod system.

DNA was extracted from blood lymphocytes by standard procedures (MILLER et al. 1988). DNA was also purified from three non-affected male control persons. All investigations followed the tenets of the Declaration of Helsinki, and informed consent and full institutional review board approval were obtained.

Previously, patients had been typed and mapped in family linkage studies to the Xp11.4-Xp11.3 region, and a minimal CSNB1 region of a ~1.5 Mb array was defined between the markers DXS993 and DXS228 (PUSCH et al., unpublished data) by haplotype analyses. Male patients of different haplotypes were selected in order to exclude genetic heterogeneity as a source of error.

Three approaches were chosen to detect mutations in CSNB1 patients.

(i) Deletion screening in patients and controls, performed by PCR, using a number of different protocols specific for the STS primer pair of interest.

(ii) Deletion screening by Southern hybridization, which proceeded as follows. Restricted genomic DNA was separated on 0.7% NuSieve agarose gels (Biozym, Hessisch Oldendorf, Germany) in 1× TBE buffer and subsequently transferred onto Qiabrane plus membranes (Qiagen, Hilden, Germany) by capillary transfer in 0.4 M NaOH, 0.025 M KOH, and 1 M NaCl. The membranes were soaked in 2 × SSC and pre-hybridized overnight in 10 ml of HybrisolXR™ (Oncor, Heidelberg, Germany) at 55°C. One hundred µg/ml unlabelled sonified salmon sperm DNA, 100 µg/ml heparin and 10% dextran-sulphate were added. Probe DNA was labelled using each 1.48 MBq ^{32}P α-dCTP, ^{32}P α-dATP (Amersham Pharmacia

Biotech Europe) and the Random Primed DNA Labeling Kit (Boehringer Mannheim, Mannheim, Germany). The probe was then purified by two passages through polymer cotton/sephadex G-50 columns at room temperature. The hybridization was carried out at 60°C for 20 h. Following hybridization, the filters were washed twice for 30 min in 4 × SSC and 0.15% SDS at 68°C, and twice for 10 min in 0.1 × SSC, 0.1% SDS at 60°C. The filters were exposed to X-ray films at –80°C for 15 h.

(iii) A touchdown protocol involving PCR amplification of putative candidate genes in a standard 50 µl reaction on a GeneAmp 9600 PCR cycler: an initial denaturation of 5 minutes at 94°C for 1 min, 5 cycles at 94°C for 15 sec, 60°C for 30 sec, and 72°C for 30 sec, 5 cycles at 94°C for 15 sec, 57°C for 30 sec, and 25 cycles at 94°C for 15 sec, 52°C for 15 sec, 72°C for 30 sec, and a final step at 72°C for 5 min. The PCR products obtained were purified using Centricon 100 (Millipore, Eschborn, Germany) one-way concentrators and sequenced with the respective forward and reverse PCR primers as described.

Results

Generation and mapping of STSs

STSs are short tracts of operationally unique DNA sequences that can be detected by PCR. STSs were generated by IRS- and DOP-PCR, using degenerate oligonucleotides that amplified a population of different fragments originating from 15 YACs, 4 rodent/human radiation-reduced hybrids, and flow-sorted X-chromosomes covering the Xp11.4-Xp11.3 region. A total of 62 amplifications were performed in triplicate, using varying amounts of template DNA and applying diverse annealing temperatures. The resulting products were cloned into plasmid vectors, and 788 clones were selected for sequencing. Doublets were excluded from further processing by multiple sequence alignments. Unique oligonucleotide pairs for subsequent STS amplification were generated by excluding repetitive regions with the REPEATMASKER software (GCG). Assignment of 224 novel/unique STSs – including 31 markers available from data bases – to the CSNB1 interval was confirmed by remapping them to (i) a whole genome panel, (ii) a subregional panel of the X chromosome, (iii) a rodent/human radiation-reduced hybrid panel covering the Xp11.4-Xp11.3 region, and (iv) two collections of YAC clones. Assuming approximately 4 cM (i.e. ~4 Mb) for this region, the marker content provides an average of 18 kb sequence tag density in the Xp11.4 region from DXS556 to DXS228. From the total of 224 sequence tags, 69 were fine-mapped to the ~1.5 Mb array of the minimal CSNB1 region between markers DXS993 and DXS228 (Table 1, Figure 1). The evaluated marker coverage of one marker per 22 kb for this sublocation is in agreement with the one previously observed for the whole region. Fine-mapping of the STSs allowed for

Table 1. Collection of newly generated sequence tags. Order of primer pairs corresponds to the boxed lists shown in Figure 1 (i.e. from distal to proximal)

Name	Sequence (5' → 3' orientation)	Temp. [°C]	Size [bp]/feature
1	2	3	4
STR1F	GAACAGCAAACCAAATCCAAA	57	(GA)n(TG)n
STR1R	TCTTGGCTTTACTCCCTCCC		
STR2F	TGTGGTGGAGTGTCTCCAAA	58	(TG)n
STR2R	AGGCTGCAGCTGGAGAATC		
W14-9F	AGATCATGCCACTGCACTC	54	129
W14-9R	TTGCTGTATGTTGCCCAGG		
W2-56F	CCCAAAGTGCTGGGATACAGG	62	265
W2-56R	AGCCTCTAGATCTGCACTGTCC		
W9-22F	CACCAAGAGTAAACCCTAACG	58	109
W9-22R	TATCAATATCCCCCACCAGAG		
W11-8F	GGCAACACACACCCTATTC	53	107
W11-8R	TGTACCAAAGCCTGGACTG		
W15-27F	ATCACCCAGCAGAGAACAC	53	114
W15-27R	GGGCACATGGTATGTTGAG		
W23-44F	TCCCTCAACATACCATGTGCC	59	107
W23-44R	TTCTGACCCACCAACTACTGC		
W17-1F	ATGCTGAAATGATCTCGTCTG	55	100
W17-1R	CAAAGAGCACACAAGTGCTG		
W11-9F	GATTACAGGTGTGAGCCACC	58	223
W11-9R	TCTGACAGCTCTTTCTTTTGTG		
W14-10F	CTGATTGACTTGTAGTCTTGGA	57	145
W14-10R	AGCCTATATGATGGATGACTTG		
W3-12F	ATGATGAACATCCCAGCA	59	87
W3-12R	GTACAGCACTGGCTGTACTTC		
W1-1F	GAACCGTGATCATGGCTCAG	58	252
W1-1R	GACTAATAATTGTCAAAGCAGGG		
STR17F	AGGTTGCAGTGAGCCAAGAT	60	(CT)n(CA)n
STR17R	CTAGAGTGCAGTGGTGCGAC		
STR21F	TTTAAGTGCCTTTTTCTTGGAA	54	(TG)n
STR21R	TTCTTCATCATAGGTACATCC		
STR22F	ATCGATGCCGTGTGTCCATA	55	(CA)n
STR22R	TCAACGTCCATTCAGGTTCC		
W21-19F	TAGCACTTTCACTTTTTGATCTCC	60	543
W21-19R	TGTTTGGTTGTTATTGCCTTCC		
W24-32F	TCTCACACTGGCCTGAAGC	55	147
W24-32R	TGGCAAGTCTGCCTCATTAG		

1	2	3	4
W20-32F W20-32R	GCTATACTTCTCTTTGCCCAGC TGTTGCTCAGGCTGAGAAAC	57	81
W20-21F W20-21R	CTTGGCATGTTTCTACGTGC AAAAGACATTTCAGGCAGAAGG	56	136
W3-21F W3-21R	CACAAAATCATTTTATAAGCCGG CTCCTGAGCTCAGGTGATCC	58	102
W20-9F W20-9R	AACAAGAGCTGCTTCTGCCC CATTCTCCATCTGACCACCC	57	108
W21-22F W21-22R	CAGGTTAAGAAGCTGGTGATCC AGGAAATAATTGAGCCGAAGG	56	179
W22-35F W22-35R	AACAGTACCACAAACAAAGCCC GCCTATGTCCCCATCTTTCC	58	188
W21-34F W21-34R	TCAACCTTTTTGTATCCTTGGG AAAAACATGGAATAGGTGGTGC	57	195
W21-28F W21-28R	GCATCCTCCTTATCTCCTGGG CGTAGCAGCTATAGAAGCAATC	61	149
W20-27F W20-27R	CGGAAATTTGCTTGATTGAGC GCAAAGCAAAGACTGGAACC	55	195
W21-17F W21-17R	AATAAACCCCAGCATACTGGTC AATCTCCAGGGGAGAAAAAGAC	59	100
W21-24F W21-24R	TCATGGTTGGCTGTATATCGAG GAAAACATAAGGCCATGTTCTAG	59	105
W21-31F W21-31R	GTATCAAAGGGGACTTAGTG GGAAGATTCTGAAGGTAAGC	53	162
W22-36F W22-36R	TGCTGCTTCTAAGTAGCCAAC GCTCAGGTTATGGGTTATCTG	57	138
W21-35F W21-35R	ACTAGCCATAGATTGACCCTTG TCAACTGTGAGTGCTTCTGTTC	59	118
W21-39F W21-39R	AAAAATGTATTGGATGGAACGG AATCCCAGGCATGTAACTATTC	56	158
W13-43F W13-43R	CTCCTGAGCTCAGGTGATCC GTAATCCCAGCACTTTGGGAG	59	325
STR3F STR3R	GGACCTGACAACTAATGGGC ATTCATTTTCCCCACAACCC	60	(TG)n
STR5F STR5R	TCAAAGGTACACCTGCACTCA TGTTGCTGCAAATGACATGA	55	(TA)n
STR18F STR18R	TTCAAATAAACCAATCACATTGC TAAAGATGCCCATACCCTGG	58	(TG)n
W20-19F W20-19R	TGTACACCCATGTAATCATCACC CAGTTCAATAATGCCTCAACAGTG	61	130
W21-26F W21-26R	GAGTGTTGAGGATTTCCAGG CCTACCATGTTCCCATTTCTT	55	133

1	2	3	4
W23-23F W23-23R	CCAAACCACTGTGGATTCTAGC GAAGAAAAACCTGGGCAGACTG	61	120
W23-45F W23-45R	GGACCAAACAGGGACAGAAC TGTGAACTCCACCTACCACC	57	108
W15-21F W15-21R	GGATTGTTAAGGCTACTCGGG GGAAGTGCAAGGCTAATAAATG	59	250
STR9F STR9R	CTGCAGTCTCAAGCACTCCAC TGCCTAGGTGTTAATACACACCA	56	(CA)n
W20-28F W20-28R	TCACTAGAGACCCTCATGCC GACTACATGCACCAAGTCCA	57	197
W20-29F W20-29R	CAAACAATGAAGTCCATTGAGG AAGGACAGATGTCTCTCCAAG	57	102
W21-37F W21-37R	CAAGAGGAATACGCAGACAC GCCATTGAAGTACCACAACC	55	135
W19-13F W19-13R	GGCAGAAGTCTGGAAACTGG GTGATGATTACATGGGTGTACA	57	185
W20-13F W20-13R	TGTGTCTCCCAGTAGAATGCG CCCAATTTCTCTGGTAGGCTC	59	174
W21-32F W21-32R	AGGAAGCAACAAGAACATCC TCTAGCTCTTCCAACTTTCC	55	162
W19-4F W19-4R	CATACATTCACGTAACATCCAC AGTGATAGGAAATAGGTCAGTG	57	105
W3-81F W3-81R	CCCTCAAATGACTATAGCCCTG TTTTTAGTCTTCCAGCAGAGGC	60	369
W21-36F W21-36R	GTGGTTCTCAAAGTATGGTCCC TGGTCCTCTAACATCAACATGC	60	176
SNP1F SNP1R	TTATATCAGCTGGGAATCTGCG ATTCTACAAGCCCAAGTTTAGGC	61	C-A
SNP2F SNP2R	CAGATTTCAGTTCAATCATCATTGAC ATTCTACAAGCCCAAGTTTAGGC	65	G-T
SNP3F SNP3R	GAGGAACAAAGGACAGAGGTTG TCTCTCGCAGCCCTTCTCTC	61	A-Δ
STR13F STR13R	TCAGACTGAAGATGACAGGTGA GACATCCTGCCTAGTTCCCA	60	(CA)n
STR15F STR15R	AGTACGAAGATCGTCCCAGG CCTCATCTGGGGTTGGAGTA	61	(CA)n
W2-60F W2-60R	GGCCATGAGTCATTGTGTTG GATTCTGATCCAACCTCTACGG	60	179
W2-62F W2-62R	TGTAGTGATGTGACAGTGCAGC ATTTTCCTATATTTGTGGCTGAACT	61	110

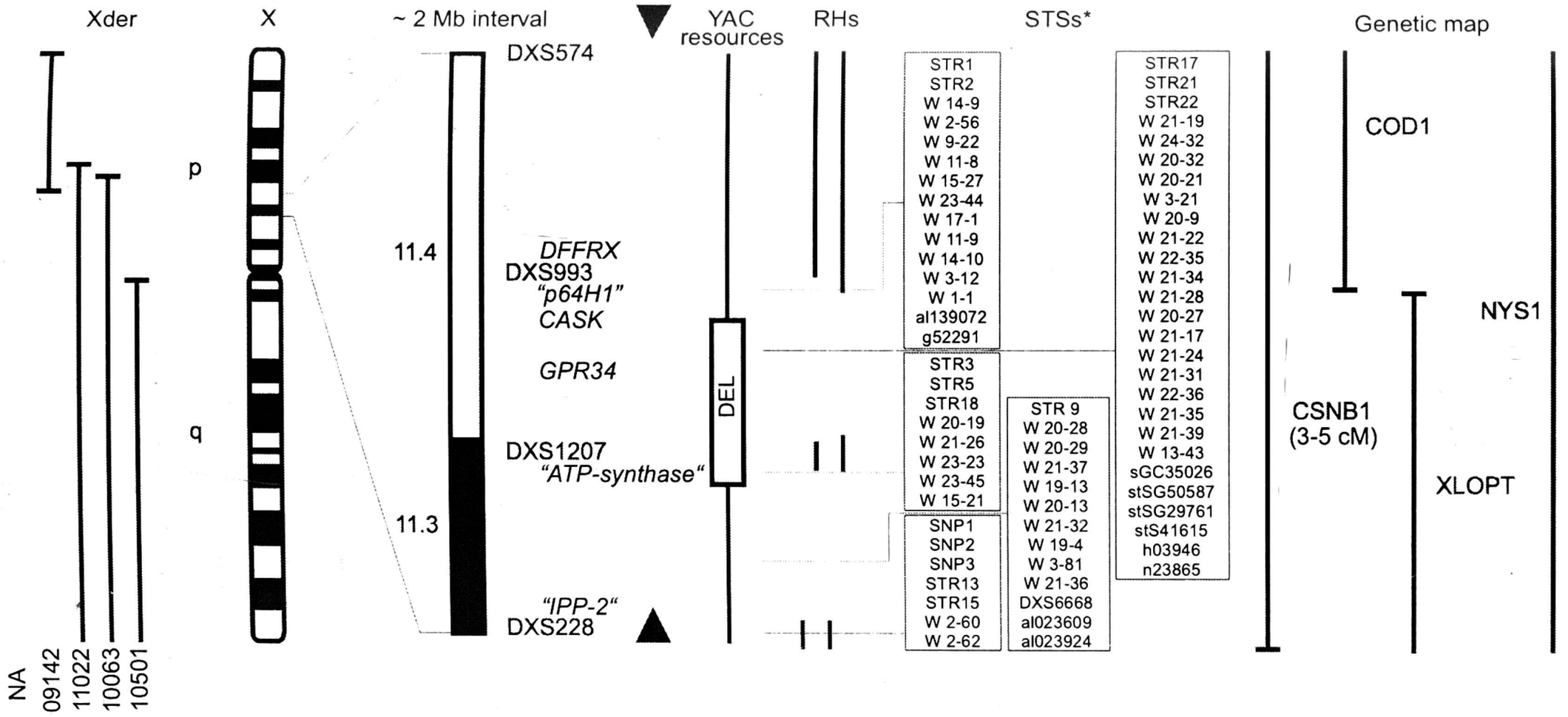

Figure 1: X-chromosome and proximal region of the CSNB1 critical interval between markers DXS574 and DXS228.

Polymorphic STSs are denoted as STR. Abbreviations used: Xder = NIGMS HGMCR human/rodent somatic cell hybrid, RH = rodent/human radiation-reduced hybrid, YAC = yeast artificial chromosome, SNP = single nucleotide polymorphism, DEL = consensus of YAC deletion regions, NYS1 = idiopathic congenital nystagmus, COD1 = X-linked progressive cone dystrophy, XLOPT = X-linked optic atrophy. Filled arrowheads mark the direction of mapping/sequencing efforts prior to starting the analyses presented here. Top: Genome Sequencing Centre Jena (Germany). Bottom: Sanger Centre Hinxton (UK). The region in between was not available from both sequencing groups.

assigning the 69 markers (i.e. 12 from data bases, 57 new sequence tags) to 5 distinct sublocations (Figure 1). Furthermore, out of the 57 new sequence tags identified, 11 display a highly polymorphic character, as each was confirmed in 35 control subjects (PUSCH et al., unpublished data). These short tandem repeat (STR) markers may therefore prove useful for refining linkage or haplotype studies.

Assessing microdeletions in CSNB1 patients

The 224 markers, together with the short tandem repeat-based sequence tags, ESTs and cDNA fragments already mapped and published elsewhere, were used in PCR screening for possible microdeletions in DNA from control subjects and 9 CSNB1 patients. In none of the cases investigated was a loss of allelic signals detected, indicating either that the deletions were non-existent in these cases or that they were too small to be detected by the coverage chosen. Three genes, namely *CASK* (DIMITRATOS et al. 1998), *DFFRX* (JONES et al. 1996), and *GPR34* (MARCHESE et al. 1999), which have been recently localized to the Xp11.4 region, were tested for rearrangements. Southern hybridization techniques showed the banding patterns of patients and healthy controls to be identical.

Data base searches, candidate gene assembly and analysis

The novel marker sequences were used for constructing an STS map of the CSNB1 region, and were therefore characterized by homology searches in a number of data bases. With markers w17-1, w20-13, and w2-60/62 (Table 1) hits were obtained with three clones (hCIT200L4, dJ50A13 and dJ154K9 respectively) in different phases of high throughput genomic sequencing (HTGS). It was possible to assign them to regions adjacent to the markers DXS993, DXS1207, and DXS228, respectively. Sequence analysis using the NIX software revealed sequence similarity to three genes highly homologous with the subunit 'c' of mitochondrial ATP synthase, a membrane protein of the intracellular chloride channel protein p64H1 and the protein phosphatase 1 inhibitor-2 (*IPP-2*).

Expression of the subunit 'c' of mitochondrial ATP synthase has been reported in sheep and a possible association with ovine ceroid lipofuscinosis has been investigated (MEDD et al. 1993). A match with a 136 aa rat ATP synthase protein (Q06645) was obtained, exhibiting a similarity of 72% to the sequence in dJ50A13 (position 126,008-126,412). However, the presence of several stop codons and frame shifts rendered this sequence a pseudogene.

Intracellular chloride channels, the probable seat of the second candidate gene, are located in vesicular organelles and may function as regulators of organelle volume and intracellular electrolyte balance. The p64H1 protein has recently been described as a member of a family that shows marked expression in the rat brain and retina (CHUANG et al. 1999). Data base analysis revealed that this channel on hCIT200L4 displays 93.1% identity in a 1.015 nt overlap to *Homo sapiens* p64H1

(AF109196; 1633 bp) and the human H1 chloride channel (AF097330; 999 bp). However, two stop codons in the putative coding region raised the possibility of HTGS sequencing errors. A total of 1256 bp (position 156,345-157,601) was sequenced in 9 patients and control persons to confirm the presence of the two stop codons and thereby the classification of the p64H1 homologue as a pseudogene.

The third candidate gene was found on a 136,144 bp PAC clone dJ154K9 and is highly homologous with the inhibitor 2 of protein phosphatase 1 (*IPP-2* or *PPP1R2*; AJ133812, X78873). Protein phosphatase inhibitors are a group of proteins that have been assigned a fundamental role in the coordination of regulation processes through inhibition of phosphatases. The large number of IPPs and the existence of PP-1 isoforms account for the diverse functions ascribed to IPP both in a tissue-specific and developmental, stage-dependent manner (SAKAGAMI et al. 1995, OLIVER, SHENOLIKAR 1998). The congenital nature and occasional association of CSNB1 with other ocular abnormalities raise the possibility of a developmental defect in CSNB1 and render the newly identified IPP a tentative candidate gene for CSNB1. The IPP coding region is located at position 113,299-112,691 of the PAC sequence, encoding a 204 aa protein (Z94277.sw_release). An amino acid sequence comparison using the algorithm GAP matches *IPP-2* (P41236) vs. Z94277 with a similarity of 52%, and an identity of 43%. FASTX analysis of Z94277.owl using the SwissPir data bases determined up to 46% identity in 200 aa overlaps of diverse IPP matches. In contrast, using GAP for comparison of the corresponding nucleic acid sequences (X78873 vs. dJ154K9) reveals a coincidence of coding sequences as high as 63.3%, and 65% when the 3'-UTR of 647 bp is additionally included in the comparison. Using Z94277.owl for a BLASTP analysis in SwissPir, protein phosphatases of several species were identified, amongst them up to 68% positive matches of IPPs in humans, rats, and rabbits.

To determine whether this protein phosphatase inhibitor is the disease-causative gene in CSNB1, direct sequencing was carried out on genomic DNA from 9 CSNB individuals from different families and three healthy males. A total of 3,190 bp including the entire IPP coding region (Z94277.sw_release) as well as 5'- and 3'-UTRs were amplified and sequenced in all cases. No mutations were found to segregate with the disease phenotype in any of the patients examined, ruling out a causative role of this IPP in the etiology of CSNB1. I was able to identify a sequencing error at nucleotide position 112,331 of the reference sequence dJ154K9, consistent with an adenine to cytosine change. In addition, three single nucleotide polymorphisms (SNPs) were identified at position 114,173 (C-A), 114,541 (G-T) upstream of the IPP coding region, and at position 112,064 (A-Δ) downstream of IPP (Figure 1, Table 1).

Discussion

Genetic changes on the human X-chromosome have been linked to a variety of inherited and sporadic disorders. Some sets of disorders have been traced to specific genes, while others have been merely linked to chromosomal regions. Although the relative density of anonymous markers on the X-chromosome has steadily increased over the past years, significant gaps still remain on this map, particularly in some regions on the short arm of chromosome X. One of them is the region Xp11.4-Xp11.3, and further mapping efforts are required to elucidate the genomic sequences associated with various disease states there. One reason is that these regions are refractory to YAC cloning, as demonstrated by the significant number of chimeric and/or partly deleted YAC clones. This is most probably due to the presence of gene-rich regions and high GC-contents. Thus, the new markers obtained from this region will help to establish a physical map based on sequence-ready templates, such as BACs and PACs, that will provide a basis for large-scale association genetic studies and for the identification and characterization of further genes.

I have contributed to the characterization of this region that is so poorly represented with useful markers by mapping of a protein phosphatase inhibitor and the two pseudogenes of the intracellular chloride channel p64H1 protein and the subunit 'c' of mitochondrial ATP synthase. It was possible to identify and sublocalize 224 STSs within that region, thus improving its current mapping status. Moreover, the identification of 3 SNPs and 11 polymorphic STR markers located in Xp11.4 will prove useful in subsequent genetic mapping studies.

The chromosomal region Xp11.4-Xp11.3 marks a gene-rich region to which several other eye diseases of currently unknown etiology, such as idiopathic congenital nystagmus (CABOT et al. 1999), X-linked progressive cone dystrophy (HONG et al. 1994), Åland island eye disease (GLASS et al. 1993) and X-linked optic atrophy (ASSINK et al. 1997) have been mapped (Figure 1). The extension of the physical mapping of the Xp11 region should facilitate future investigations into the molecular basis of the above-mentioned ocular disorders and presents the IPP-2 homologue as a candidate gene for the ocular disorders of idiopathic congenital nystagmus (NYS1) and X-linked optic atrophy (XLOPT).

Acknowledgements. This work was supported by a grant from the Federal Ministry of Education, Science, Research and Technology (Fö. 01KS9602) and the Interdisciplinary Center of Clinical Research (IZKF) Tübingen. The author has no financial or proprietary interest in the subject matter mentioned in this article. I am grateful to A. MEINDL (Munich) and W. BERGER (Berlin) for helpful discussions, for providing the rodent/human radiation-reduced hybrid panel and a collection of YACs.

REFERENCES

ASSINK J.J.M., TIJMES N.T., TEN BRINK J.B., OOSTRA R.J., RIEMSLAG F.C., DE JONG P.T., BERGEN A.A. (1997). A gene for X-linked optic atrophy is closely linked to the Xp11.4-Xp11.2 region of the X-chromosome. Am. J. Hum. Genet. 61: 934-939.

BERGEN A.A.B., TEN BRINK J.B., RIEMSLAG F., SCHUURMAN E.J., TIJMES N. (1995). Localization of a novel X-linked congenital stationary night blindness locus: close linkage to the RP3 type retinitis pigmentosa gene region. Hum. Mol. Genet. 4: 931-935.

BERGER W., MEINDL A., DE LEEUW B., DE ROOS A., VAN DE POL T.J., MEITINGER T., VAN DER VELDE-VISSER S.D., ACHATZ H., GEURTS VAN KESSEL A., CREMERS F.P., ROPERS H.H. (1992). Generation and characterization of radiation reduced hybrids and isolation of probes from the proximal short arm of the human X chromosome. Hum. Genet. 90: 243-246.

BOYCOTT K.M., PEARCE W.G., MUSARELLA M.A., WELEBER R.G., MAYBAUM T.A., BIRCH D.G., MIYAKE Y., YOUNG R.S., BECH-HANSEN N.T. (1998). Evidence for genetic heterogeneity in X-linked congenital stationary night blindness. Am. J. Hum. Genet. 62: 865-875.

CABOT A., ROZET J.M., GERBER S., PERRAULT I., DUCROQ D., SMAHI A., SOUIED E., MUNNICH A., KAPLAN J. (1999). A gene for X-linked idiopathic congenital nystagmus (NYS1) maps to chromosome Xp11.4-p11.3. Am. J. Hum. Genet. 64: 1141-1146.

CHUANG J.Z., MILNER T.A., ZHU M., SUNG C.H. (1999). A 29 kDa intracellular chloride channel p64H1 is associated with large dense-core vesicles in rat hippocampal neurons. J. Neurosci. 15: 2919-2928.

DIMITRATOS S.D., STATHAKIS D.G., NELSON C.A., WOODS D.F., BRYANT P.J. (1998). The location of human CASK at Xp11.4 identifies this gene as a candidate for X-linked optic atrophy. Genomics 51: 308-309.

GLASS I.A., GOOD P., COLEMAN M.P., FULLWOOD P., GILES M.G., LINDSAY S., NEMETH A.H., DAVIES K.E., WILLSHAW H.A., FIELDER A. (1993). Genetic mapping of a cone and rod dysfunction (Aland Island eye disease) to the proximal short arm of the human X chromosome. J. Med. Genet. 30: 1044-50.

HARDCASTLE A.J., DAVID-GRAY Z.K., JAY M., BIRD A.C., BHATTACHARYA S.S. (1997). Localization of CSNBX (CSNB4) between the retinitis pigmentosa loci RP2 and PR3 on proximal Xp. Invest. Ophthalmol. Vis. Sci. 38: 2650-2755.

HONG H.K., FERRELL R.E., GORIN M.B. (1994). Clinical diversity and chromosomal localization of X-linked cone dystrophy (COD1). Am. J. Hum. Genet. 55: 1173-1181.

JONES M.H., FURLONG R.A., BURKIN H., CHALMERS I.J., BROWN G.M., KHWAJA O., AFFARA N.A. (1996). The *Drosophila* developmental gene fat facets has a human homologue in Xp11.4 which escapes X-inactivation and has related sequences on Yq11.2. Hum. Mol. Genet. 5: 1695-1701.

MARCHESE A., SAWZDARGO M., NGUYEN T., CHENG R., HONG H.H., NOWAK T., IM D.S., LYNCH K.R., GEORGE S.R., O'DOWD B.F. (1999). Discovery of three novel orphan G-protein-coupled receptors. Genomics 56: 12-21.

MEDD S.M., WALKER J.E., JOLLY R.D. (1993). Characterization of the expressed genes for subunit c of mitochondrial ATP synthase in sheep with ceroid lipofuscinosis. Biochem. J. 293: 65-73.

MILLER S.A., DYKES D.D., POLESKY H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 16: 1215.

OLIVER C.J., SHENOLIKAR S. (1998). Physiologic importance of protein phosphatase inhibitors. Frontiers in Biosciences 3: 961-972.

ROZZO C., FOSSARELLO M., GALLERI G., MIANO M.G., CICCODICOLA A., SOLE G., PIRASTU M. (1999). Complete congenital stationary night blindness maps on Xp11.4 in a Sardinian family. Eur. J. Hum. Genet. 7: 574-578.

SAKAGAMI H., KONDO H. (1995). Molecular cloning of the cDNA for rat phosophatase inhibitor-2 and its wide gene expression in the central nervous system. J. Chem. Neuroanatom. 8: 259-266.

THISELTON D.L., LINDSAY S., KAMAKARI S., HARDCASTLE A.J., ROUSTAN P., BHATTACHARYA S.S. (1995). Genetic and physical mapping of five novel microsatellite markers on human Xp21.1-p11.22. Genomics 25: 279-281.