



Transcriptional analysis of the candidate region for incontinentia pigmenti (IP2) in Xq28.

U C Rogner, N S Heiss, P Kioschis, et al.

Genome Res. 1996 6: 922-934

Access the most recent version at doi:[10.1101/gr.6.10.922](https://doi.org/10.1101/gr.6.10.922)

References This article cites 57 articles, 9 of which can be accessed free at:
<http://genome.cshlp.org/content/6/10/922.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Copyright © Cold Spring Harbor Laboratory Press

RESEARCH

Transcriptional Analysis of the Candidate Region for Incontinentia Pigmenti (IP2) in Xq28

Ute C. Rogner,¹ Nina S. Heiss,¹ Petra Kioschis, Stefan Wiemann, Bernhard Korn, and Annemarie Poustka²

Deutsches Krebsforschungszentrum, Abteilung Molekulare Genomanalyse, D-69120 Heidelberg, Germany

The hereditary form of incontinentia pigmenti (IP2) is a rare disorder characterized by abnormalities of the tissues and organs derived from the ectoderm and neuroectoderm and has been linked to Xq28 distal to the factor VIII gene (F8C). Four YAC clones covering the 1.1-Mb candidate region at the telomere of Xq28 were subjected to direct cDNA selection and Alu long-range PCR. The products of both methods were subsequently used to isolate 154 cosmid clones that were assembled into five cosmid contigs. This first-generation cosmid map covered the region almost entirely and was used as a basis for constructing a transcript map that was in turn integrated with the physical YAC and cosmid maps. To isolate specifically coding sequences, exon trapping and cDNA selection methods were combined. Exon trapping was carried out on YAC Alu-PCR products, YAC Alu long-range PCR products, and on pools of cosmids. The region-specific enriched cDNA library was then screened by using the exon trap products as complex probes. To ensure a more complete analysis, the products from cDNA selection experiments were also used to screen conventional oligo(dT) primed cDNA libraries. Twenty overlapping cDNA contigs were assembled and computer analyses were performed to identify EST hits, open reading frames, protein motifs, and protein sequence homologies. Five of the cDNA contigs corresponded to known sequences such as the factor VIII, c6.1A, and c6.1B genes, and both distal copies of the factor VIII intron 22 repeat sequence. Expression patterns of the 15 new cDNA contigs were analyzed by Northern blot and RT-PCR studies and these data were integrated with expression data obtained from known EST sequences. Although a more detailed analysis of this 1.1-Mb region with respect to the structure and function of the genes will only ultimately be possible by a global sequencing approach, an analysis of all novel transcripts as candidate genes for incontinentia pigmenti is already in progress.

Incontinentia pigmenti, also known as Bloch-Sulzberger syndrome, is a rare genodermatosis that usually appears at birth or shortly thereafter. The disease is characterized by swirled patterns of hyperpigmentation that are also evident in biopsies of the streaks that show pigment within dermal macrophages ("incontinent pigment"). Associated findings include seizures, mental retardation, strabismus, cataracts, and delayed or impaired dentition (Carney 1976; Isselbacher et al. 1994). The pigmentary lesions disappear with aging and this can be interpreted as a selection against cells in which the mutation-bearing X

chromosome is active (Migeon et al. 1989, Coleman et al. 1993). Several sporadic cases exhibiting an X-autosomal translocation with the breakpoint located at Xp11, indicated that the gene for incontinentia pigmenti (*IP1*; McKusick 308300) may lie in Xp11.12 (Gorski et al. 1991; Reed et al. 1994). Linkage analyses of 12 families assigned the gene for the familial form of incontinentia pigmenti (*IP2*; McKusick 308310) to Xq28 (Sefiani et al. 1991). This was further confirmed by Smahi et al. (1994) who presented a study with 21 additional families and were able to reduce the candidate region to an area between F8C and the telomere.

The human chromosomal region Xq28 is in the focus of medical and scientific interest because of the large number of disease genes that have been mapped here. At least 10 of the more

¹These authors contributed equally to this work.

²Corresponding author.

E-MAIL a.poustka@DKFZ-Heidelberg.de; FAX 49-6221-423454.

NOVEL TRANSCRIPTS IN CANDIDATE REGION FOR IP2

than 20 disease genes remain still uncloned. In the course of a systematic analysis, a physical map covering the 12 Mb of the terminus of the long arm of the human X chromosome was constructed (Poustka et al. 1991; Dietrich et al. 1992). YAC contigs spanning 7.5 Mb between the iduronate-2-sulphatase gene (IDS) and the telomere were assembled and compared with the genetic and physical data (Palmieri et al. 1994; Rogner et al. 1994). So far, transcriptional maps were described in the regions around the *LICAM* and *G6PD* genes (Korn et al. 1992; Sedlacek et al. 1993), the candidate region for the *MTM* gene (Kioschis et al. 1996), and the region around the polymorphic DXS52 loci (Heiss et al. 1996).

The composition of the telomeric region of Xq28 is in itself most interesting. A second pseudoautosomal region was described at the Xq and Yq telomeres (Freije et al. 1992), which encompasses ~400 kb proximal to the telomere and was obviously a result of an ectopic recombination mediated by long interspersed element (LINE) sequences (Kvaloy et al. 1994). Furthermore, the largest intron of the factor VIII gene (Gitschier et al. 1984; Wood et al. 1984), intron 22, harbors two smaller factor VIII-associated genes, F8A (Levinson et al. 1990) and F8B (Levinson et al. 1992). One of the two genes, F8A, is part of the 9.5-kb intron 22 homologous sequences (int22h) that has two more copies located close to the telomere (Naylor et al. 1993, 1995). Intrachromosomal inversions between the intragenic and either one of the two extragenic copies are responsible for almost half of the cases of severe hemophilia (Lakish et al. 1993; Naylor et al. 1993). To date, only two other genes are mapped distal to factor VIII, c6.1A and c6.1B (Kenrick et al. 1992). Both genes are associated with T-cell prolymphocytic leukemia in patients with ataxia telangiectasia (Fisch et al. 1993; Thick et al. 1994).

As a prerequisite to the positional cloning of the gene responsible for the familial form of incontinentia pigmenti (*IP2*), we constructed yeast artificial chromosome (YAC) (Rogner et al. 1994) and cosmid contigs covering the 1.1-Mb region between the factor VIII gene and the telomere of Xq28. Direct cDNA selection (Lovett et al. 1991; Parimoo et al. 1991; Korn et al. 1992) and exon amplification methods (Buckler et al. 1991; Church et al. 1994) were combined with screening of oligo(dT)-primed cDNA libraries to identify novel transcripts in the region. The resulting contigs of overlapping cDNAs were mapped to the cosmids and YACs, leading to the first inte-

grated transcriptional map within the candidate region for incontinentia pigmenti (*IP2*).

RESULTS

Construction of Cosmid Contigs

Cosmid clones were isolated from an Xq28-specific cosmid library (Kioschis et al. 1991), from flow-sorted X chromosome libraries (Nizetic et al. 1991), and from the Lawrence Livermore Laboratory library by using Alu-PCR, Alu long-range PCR, and direct cDNA selection products from YACs I011, 1792, C869, and I234 (Rogner et al. 1994). This facilitated the identification of a total of 154 cosmid clones corresponding to a sixfold coverage in the 1.1-Mb region. The cosmids were first assigned to seven intervals of overlap on the YAC map. Five DNA markers from the region (St35.549, DXS115, St35.955, St35.617, and DXYS64; Palmieri et al. 1994; Rogner et al. 1994) and five YAC end probes (CA2-91, 869-89, 138-91, 869-91, and 497-91; Rogner et al. 1994) were then used to assemble one-third of the cosmids into five bins. The probes F85' (Gitschier et al. 1984; Wood et al. 1984) and TelBam3.4 (Brown et al. 1990) did not detect any cosmids, but mapped to the YACs ACA2 and I011, respectively (Fig. 2, below).

Deleted cosmid clones were eliminated at this stage of the analysis and the 78 remaining unassigned cosmids were hybridized individually to gridded filters carrying all region-specific cosmids. Seventy cosmids were analyzed further by *EcoRI* and *HindIII* double digests to estimate their extent of overlap. In addition to comparing digestion patterns of the cosmids and to confirm the extent of the overlaps, corresponding Southern blots were hybridized with CA-repeat (polydA-dC, Pharmacia), Alu-repeat (Perlino et al. 1985), and L1-repeat (Kpn 1.2, 1.5, 1.8 kb; Shafit-Zagardo et al. 1982) probes (see Fig. 1). Hybridizations for all three repeat probes were different and overlaps were detectable in all cases, especially for the Alu probe (Fig. 1). Hybridization of the complex cDNA selection and exon trap products onto the cosmid digests provided further confirmation of cosmid overlaps at a high resolution (data not shown). The overlaps of the cosmid contigs were also verified by hybridizing 22 individual newly isolated cDNA clones and four PCR-amplified sequence tagged sites (STSs) (DXS1108, DXYS154, DXYS225, and DXYS227; Kvaloy et al. 1994; Li and Hamer 1995). The STS

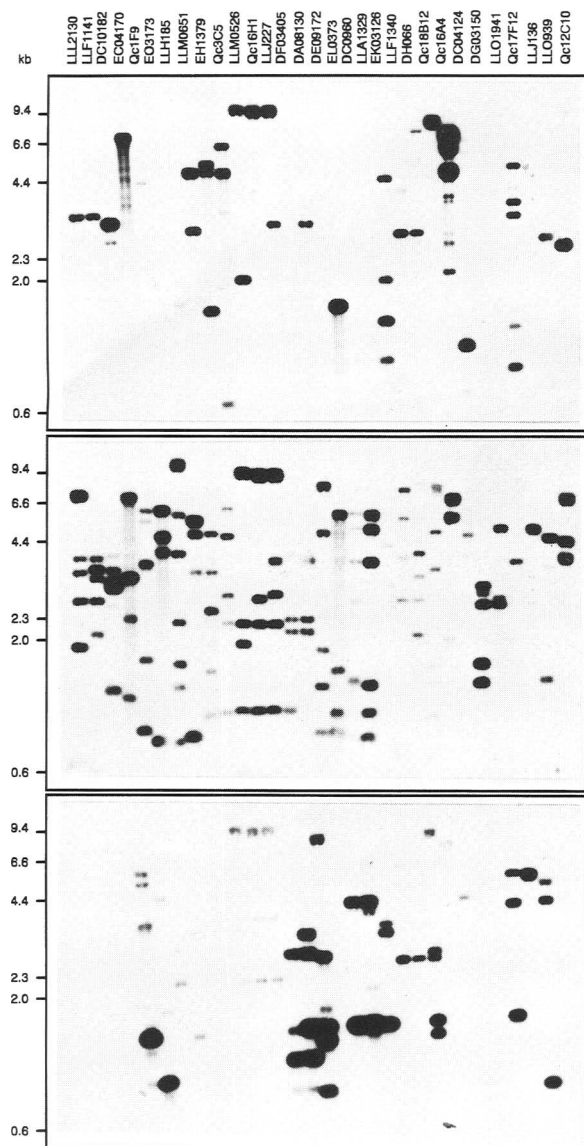


Figure 1 Blots carrying *Eco*RI and *Hind*III double-digested DNA of cosmids from the region between the YAC end CA2-91 and DXYS154 (see Fig. 2). The same blot was hybridized to a polyCA probe (*top* panel), an Alu-repeat probe (*middle* panel) and a L1-repeat probe (*bottom* panel) to confirm cosmid overlaps.

JXYQ mapped to the YACs I011, I792, and C869. However, because this marker was not amplified from any cosmid, it was positioned in the gap distal to DXS1108 according to results described previously (Kvaloy et al. 1994; Li and Hamer 1995). Finally, a first-generation cosmid map as depicted in Figure 2 was constructed. Because attempts to isolate cosmids within the gaps failed,

the exact size of the gaps could not be determined but were estimated to be relatively small according to the previously established pulsed-field gel electrophoresis (PFGE) and YAC maps (Fig 2.; Poustka et al. 1991; Palmieri et al. 1994; Rogner et al. 1994).

Construction and Screening of cDNA Libraries

To construct a region-specific cDNA library we used gel-purified inserts from YACs I011, I792, C869, and I234 (Rogner et al. 1994) for direct cDNA selection (Korn et al. 1992) from human fetal brain, human fetal liver, and human adult skeletal muscle tissues. In total, 6144 clones were picked into microtiter dishes and gridded onto nylon membranes. To minimize unnecessary analysis of unwanted clones, the filters were first hybridized with probes containing Alu-, L1-, and CA-repeat probes, total human, and total yeast DNA. Only 410 clones were found to contain repetitive sequences.

As part of a combined approach for the isolation of transcripts, we carried out exon trapping on three different substrates. Exon trapping (Buckler et al. 1991; Church et al. 1994) was performed on Alu-PCR and Alu long-range PCR products (Wilgenbus et al. 1995; Heiss et al. 1996a) of the four YACs and on pools of cosmids. All region-specific cosmids were used for exon trapping and were pooled according to the intervals of overlaps of the YACs. Finally, each pool covered between 100 and 200 kb and contained a minimum of 8 and a maximum of 57 cosmids. In light of the fact that the sizes of the cosmid pools varied and the efficiency with which exons are trapped is reduced considerably with increased DNA complexity (Church et al. 1994), we might have expected a relative reduction in the recovery of exons from the larger cosmid pools. However, because of the redundancy of cosmids in the larger pools, we anticipated that this effect would be compensated. The use of exon trap products as probes on enriched cDNA libraries is specific and efficient (Heiss et al. 1996a) and hence we considered it appropriate to use products of the three exon trap experiments as complex probes to screen the enriched cDNA library.

The hybridization of YAC Alu- and Alu long-range PCR-derived exon trap products identified 17 and 29 different cDNA clones, respectively. The majority of clones (159) however, were isolated by screening with cosmid-derived exon trap products. Only 10 of these clones were identical

NOVEL TRANSCRIPTS IN CANDIDATE REGION FOR IP2

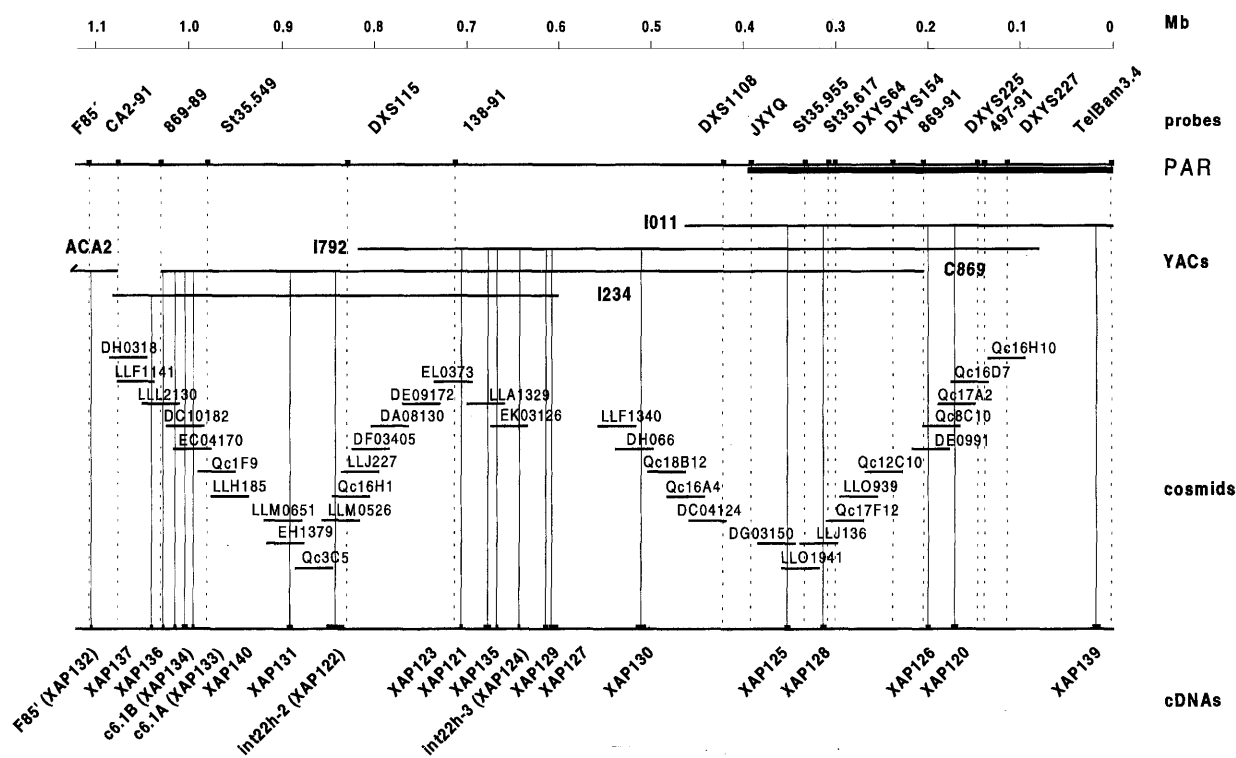


Figure 2 Transcriptional map of the IP2 candidate region. Probes: YAC-ends are abbreviated by –89 for left and –91 for right end probes (Rogner et al. 1994). The STS JXYQ (Li and Hamer 1995) defines the proximal end of the pseudoautosomal region (PAR). YACs: Full names are given in Rogner et al. (1994). Cosmids: A minimal overlap of 35 cosmids is shown. Libraries are abbreviated as follows: Xq28 specific library (Qc) (Kioschis et al. 1991); ICRFc104 (D); ICRFc100 (E) (Nizetic et al. 1991); and Lawrence Livermore library (LL).

to clones also isolated by the Alu-PCR approach. No identical clones were found on comparing the Alu-PCR with the Alu long-range PCR experiment or the Alu long-range PCR to the cosmid exon trap experiment. In summary, a total of 195 cDNA clones were found to be positive using exon trap products as complex probes on the region-specific cDNA library. To obtain larger cDNA clones and to generate longer sequences we additionally used the cDNA selection products as complex hybridization probes on conventional cDNA libraries. The cDNA elution products of the four YACs were hybridized onto gridded oligo(dT)-primed cDNA libraries constructed from fetal brain and fetal thymus mRNA (ICRFp507 and ICRFp508). Twenty conventional cDNA clones were isolated and subsequently hybridized back to the region-specific enriched cDNA library. In this way, a further 189 enriched cDNA clones were identified of which, however, one-third were identical to clones already isolated by the cosmid-derived exon trap products.

Altogether, 736 clones of the enriched cDNA

library were found to be positive in at least one of the experiments described above. These clones, plus the 410 clones that contained repetitive elements, represented 18% of the entire enriched cDNA library. The remaining clones that were not positive after the combined hybridization approach may, in part, represent contaminating yeast clones, genomic clones and/or unspliced cDNA clones. Furthermore, the library is likely to contain cDNA clones that were missed because of their absence in the corresponding exon trap probes.

Assembly of cDNA Contigs

Our initial aim was to isolate a maximum of sequence per cDNA and because the isolation of individual exon trap products does not contribute to an increase in the length of the sequence, we only sequenced cDNA clones. A subset of 289 cDNA clones were sequenced from both ends. These included 17 clones from the YAC-Alu PCR exon trap experiments, 29 clones from the YAC

Alu long-range PCR exon trap experiments, and 149 from the cosmid exon trap experiments. Twenty clones were derived from conventional cDNA libraries and 74 were obtained by hybridizing the conventional clones back onto the enriched cDNA libraries. Complete sequences were generated for insert sizes up to 800 bp, which was also the average size of inserts in the enriched region-specific cDNA library. Sequences were assembled using the XGAP program from the Staden package (Dear and Staden 1991) and this resulted in the assembly of 20 nonoverlapping cDNA contigs (Table 1). To ensure that the cDNA contigs exhibited no further overlap, dot-plot analyses were performed with the Genskipper program (version 1.2; EMBL, Heidelberg).

Three unique contigs (XAP120, XAP121, and XAP122) resulted from the hybridization of Alu-PCR exon trap products to the enriched cDNA library. Six of the contigs (XAP123–XAP128) were found using Alu long-range PCR exon trap products as probes on the region-specific library. Screening of the same library with cosmid-derived exon trap products led to the identification of eight additional cDNA contigs (XAP 129–136) of which one, XAP131, also overlapped with cDNA clones that had been isolated from conventional oligo(dT)-primed libraries. Three contigs (XAP137, XAP139, and XAP140) were assembled exclusively from conventional cDNA clones. At least one clone from each contig was chosen for hybridization and mapping onto cosmids and YACs (Fig. 2).

Data-base Comparisons

To identify homologies to known sequences we employed the BLASTN, BLASTX, and BLASTP programs (Altschul et al. 1990) of the Heidelberg Unix Sequence Analysis Resources (HUSAR) (Table 1). Known genes in the region were comprised of the 5' sequence of the factor VIII gene (XAP132 in Table 1; Gitschier et al. 1984; Wood et al. 1984), and the c6.1A and the c6.1B cDNAs (XAP133 and XAP134 in Table 1; Kenwick et al. 1992). Other known sequences that were re-isolated included XAP122 and XAP124 (Table 1). XAP122 exhibited identity to a partial human cDNA derived from the microdissection clone GF4 (Yokoi et al. 1994), which is also part of all three int22h sequences (Naylor et al. 1995), whereas XAP124 overlapped the int22h sequence itself (Naylor et al. 1993, 1995). We failed, how-

ever, to isolate the IL-9 receptor gene that was mapped recently to the Xq pseudoautosomal region (Kermouni et al. 1995) and is known to lie at the telomere (D'Esposito et al. 1996) where we had failed to isolate cosmids.

BLASTN and BLASTX analyses revealed that 2 of the 15 novel cDNA contigs exhibited high homologies to known cDNAs and proteins. XAP120 was interesting because it showed a 94% and 99% homology to two human expressed sequence tags (ESTs). One of these, X92396, has already been mapped to Xq28 (D'Esposito et al. 1996). Furthermore, 84% and 67% homologies to synaptobrevin and cellubrevin DNA sequences from *Rattus* sp. (McMahon et al. 1993) and *Arabidopsis thaliana* were observed in 269 and 110 bp of the 1371-bp cDNA contig, respectively (Table 1). BLASTX comparisons showed that XAP120 exhibited 34%–43% identical and 53%–68% similar amino acids over a length of 206 amino acids to a *Homo sapiens* synaptobrevin-related protein (SWISS-PROT accession no. P47192). Furthermore, this sequence demonstrated 35% identity and 60% similarity to 89 amino acids of a *Saccharomyces cerevisiae* synaptobrevin homolog (SWISS-PROT accession no. P31109). The second cDNA contig that demonstrated homology to known sequences was XAP121, which exhibited a 96%–98% homology to three human ESTs and also showed a 66% homology in 462 bp to a cDNA of the p64 bovine chloride channel protein (Table 1; Landry et al. 1993). Moreover, XAP121 showed 60% and 62% homology across 153 and 339 bp, respectively, to two *Rattus* cDNAs encoding proteins similar to the bovine chloride channel protein (Table 1). At the protein level, XAP121 also demonstrated 66%–72% similarity and 80%–91% identity in 50 amino acids to the p64 bovine chloride channel protein (SWISS-PROT accession nos. P35526, pironly: A47104, Pir). BLASTN analyses revealed that although seven of the remaining cDNA contigs exhibited between 91% and 100% homology to large numbers of mostly redundant human ESTs, to this point, none of these had been mapped previously to Xq28. Besides showing identity to many ESTs, none of these cDNA contigs were homologous to known genes or proteins (Table 1). Nonetheless, by using the Genskipper program to analyze the predicted open reading frames (ORFs) for possible protein motifs, we found that XAP127 and XAP130 contained possible cell attachment sites (RGD). In addition, a histone H4 signature (GAKRH) was found in the ORF of XAP127 and a putative ty-

NOVEL TRANSCRIPTS IN CANDIDATE REGION FOR IP2

Table 1. cDNA Contigs Distal to the Factor VII Gene

cDNA contig	no. of clones	length bp	BLASTN homologies	ORF bp	BLASTX homologies	protein sites	Expression pattern	Genbank accession number
XAP120	18	1371	human ESTs: R27644 (94%); X92396 (99%); rat EST: H33932 (84%); <i>Arabidopsis thaliana</i> EST: M90418 (67%)	663	synapto-brevin	none	ubi 3.2 kb	X97503
XAP121	13/3*	1072/317	human ESTs: T52280 (98%); T73274 (96%); R06362 (96%); F08553 (64%); R121143 (64%); rat ESTs: H95101 (62%); H32187 (60%); bovine EST: L16547 (66%)	882/129	bovine chloride channel protein	none	RT-PCR: skm, ili	X97509
XAP122	20	3777	human ESTs: D16474 (99%); X86012 (99%); R55174 (99%); Z19829 (97%)	540	int22h-2	none	RT-PCR: br (1)	X97508
XAP123	7	579	none	273	none	none	ubi 4.2, 2.8, 1.7 kb	X97507
XAP124	2	518	human EST: L23870 (99%)	216	int22h-3	none	hea, skm, kid > 10 kb	X97506
XAP125	6	1196	human ESTs: T6756 (99%); H24223 (100%)	183	none	none	hea, skm, kid > 10 kb (2)	X97504
XAP126	2	863	none	213	amidation	cell attachment	ubi 2.3 kb	X97505
XAP127	2	1809	none	375	none	histone	hea, skm > 10 kb, 2.3 kb	X97510
XAP128	5	1136	none	594	none	none	ubi 2.3 kb	X97511
XAP129	11	1142	none	570	none	none	pla smear	X97512
XAP130	20/8*	1191/833	human ESTs: Z45989 (98%); T32618 (96%); R18299 (98%); T65328 (99%); T81256 (100%); R42275 (94%)	1035/708	none	cell attachment	hea, skm 4.4, 1.7 kb	X97513
XAP131	25/2*	756/418	human ESTs: D53331 (97%); D60306 (95%); R68783 (99%); R51425 (99%); R51425 (99%); Z45872 (98%); T35659 (99%); T35595 (99%); T90241 (100%); R34624 (99%); T31469 (100%); R30678 (100%); R58767 (95%); D60406 (98%); T08770 (99%); Z41853 (100%)	597/201	none	Tyr phosphorylation	hea, bra, liv, skm, pan, kid 6 kb	X97514
XAP132	2	510	FVII	n.d.	factor VIII	no data	ubi 9 kb (3)	in database
XAP133	18	1190	c6.1A	n.d.	c6.1A	no data	ubi 2.6 kb (4)	in database
XAP134	2	590	c6.1B	n.d.	c6.1B	no data	ubi 5kb (4)	in database
XAP135	2	470	human ESTs: T64415 (98%); R54772 (77%); R63298 (100%)	159	none	none	ubi 5, 3, 2.4, 2, > 1 kb	X97515
XAP136	2	448	human ESTs: H21199 (98%); H39163 (97%); H49123 (99%); R94014 (88%); T78543 (91%); H66519 (97%); H43518 (98%); H78051 (94%); R59643 (92%); R93970 (94%); H42518 (92%); H38957 (95%); H49058 (95%); H24548 (93%)	210	none	none	ubi 6 kb, skm, kid 7 kb	X97516
XAP137	3	482	human ESTs: H09739 (100%); T17155 (100%); F02764 (99%); R40707 (100%); F04398 (99%)	141	none	none	hea, skm 1.9 kb	X97517
XAP139	7/4*	746/855	human ESTs: Z40757 (100%); T47278 (96%); T34907 (97%); F11048 (97%); R89629 (99%); F02886 (94%); T11665 (100%); T32439 (97%); F10620 (97%); T51217 (97%); R43475 (92%); R41279 (93%); R49647 (93%); T05363 (100%); T99608 (96%); H27200 (97%); H04565 (89%); T95981 (95%)	279/348	none	none	ubi 2 kb; hea, skm 6.7 kb	X97518
XAP140	3	594/551	none	177/282	polyCA in 3' end	none	ubi 2.1 kb; pla 6.5 kb	X97519

References: Kenwick et al. 1992; Levinson et al. 1992; Naylor et al. 1993; Yokoi et al. 1994.

*Two different halves of one contig of which the complete sequence still needs to be determined. RT-PCR primers used for the amplification of a 236-bp product for XAP121 were 5'-GAGCTTTTGGTGGCTGTA-3' and 5'-CAAGCTACAATCAGCCAGTGA-3'. (Tyr) Tyrosine; (ubi) ubiquitous; (hea) heart; (skm) skeletal muscle; (pla) placenta; (liv) liver; (pan) pancreas; (kid) kidney; (fbr) fetal brain; (fli) fetal liver.

ROGNER ET AL.

rosine phosphorylation site (KKLDEQY) in XAP131 (Table 1).

Expression Studies

Northern blot hybridizations were performed for all novel cDNAs in order to determine transcript sizes and tissue expression. cDNA clones that overlapped into the ORFs of the 15 novel sequences (as listed in Table 1, except for the known genes XAP122, XAP124, XAP132, XAP133, and XAP134) were chosen for hybridizations on Northern blots carrying mRNA from different adult tissues (Clontech; Table 1; Fig. 3). Nine cDNAs identified ubiquitously expressed transcripts, whereas three were only expressed in muscle and heart (Table 1). XAP129 was found to be expressed only in placenta, and XAP131 was expressed in all tissues except placenta and lung (Table 1). It became evident that as many as eight of the cDNA contigs exhibited complex transcription patterns detecting more than one transcript in all or only a few tissues present on the Northern blot (XAP123, XAP126, XAP127, XAP130, XAP135, XAP136, XAP139, and

XAP140) and XAP129 hybridized as a smear. Because the data indicated that these genes probably undergo complex alternative splicing but was also suggestive of the presence of multiple gene families possibly not mapping to Xq28, it was necessary to confirm and extend these findings. To this end, we hybridized cDNAs covering the length of each of these contigs onto panels bearing *Eco*RI-digested total genomic DNA from human female, human male, the Q1Z cell line carrying Xq27.3-tel on a hamster background (Warren et al. 1990), and a cell line carrying Xp-Xq27.3 (Warren et al. 1990). A lane with hamster DNA only was included as a control. For XAP123, XAP127, XAP130, XAP136, and XAP140 the hybridizations revealed identical bands in all lanes for each cDNA contig except in those lanes carrying DNA from the Xp-Xq27.3 cell line and hamster control DNA. This showed clearly that these cDNAs map to one locus in Xq28 (data not shown). cDNAs from XAP135, however, hybridized as two bands in lanes containing human female and male DNA, but only as one band in the Xq28-containing Q1Z cell line, strongly indicating that XAP135 is part of a gene family. The

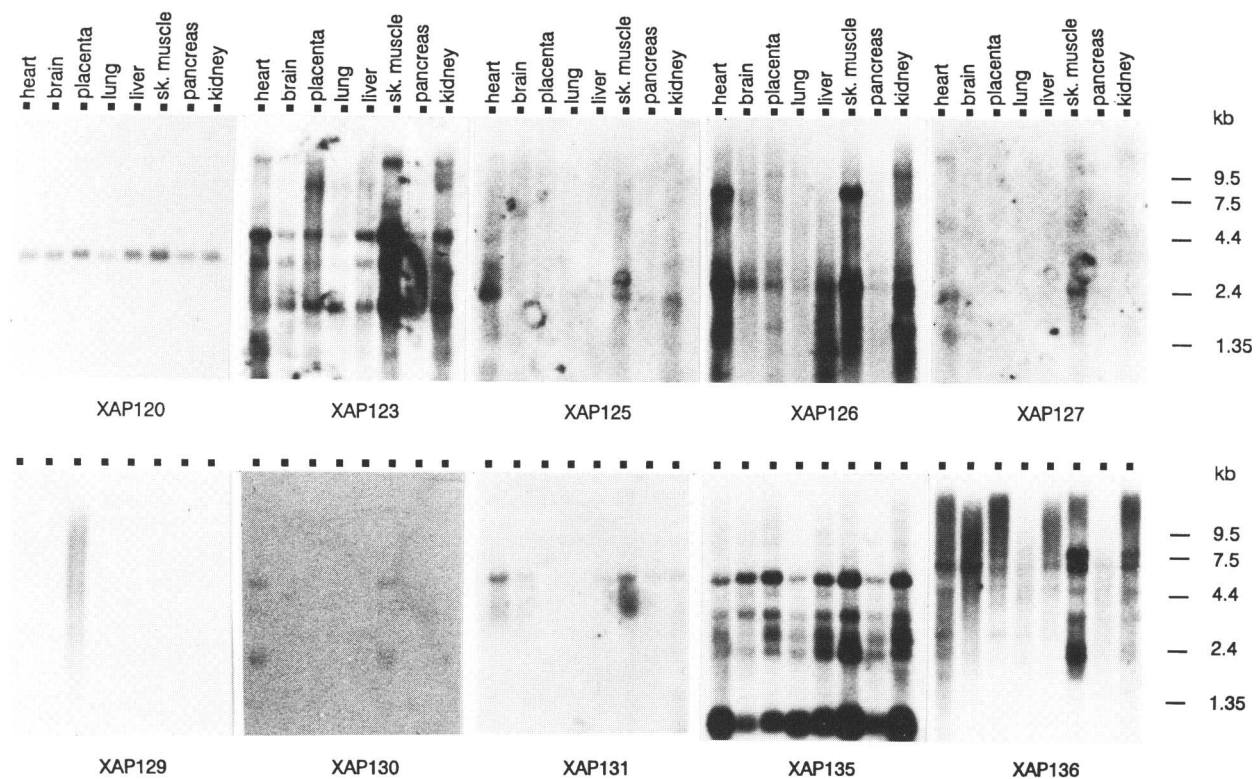


Figure 3 Northern blot analysis of 10 cDNA contigs distal of the factor VIII gene. Note that XAP125 identified the same transcript as XAP128, and XAP126 identified the same transcripts as XAP139.

NOVEL TRANSCRIPTS IN CANDIDATE REGION FOR IP2

hybridization results for XAP126 and XAP139 were ambiguous in that multiple bands were observed in all lanes.

Besides providing data on tissue expression, Northern hybridizations also served to verify whether each of the cDNAs represented a unique gene or different parts of the same gene. Some of the cDNA contigs showed identical size and expression of the transcripts. The two contigs XAP125 and XAP128, for example, detected a ubiquitously expressed transcript of 2.3 kb. XAP126 and XAP139 both identified a ubiquitous transcript of 2 kb and an additional transcript of 6.7 kb in muscle and heart (Table 1; Fig. 3). In conjunction with the mapping data (Fig. 2), this suggested that each of these pairs of cDNA contigs represented different parts of the same gene.

Because of the importance of relating tissue expression of the novel cDNAs to the tissues affected by the incontinentia pigmenti disease, it was of interest to correlate the tissue origin of the ESTs with the Northern blot data, especially for those cDNA contigs that exhibited high homologies to human ESTs but showed no evidence of ubiquitous expression on Northern blots. From Northern blots, XAP130 appeared to be expressed only in heart and skeletal muscle (Table 1), whereas its homologous ESTs were derived from diverse tissues such as infant brain, fetal liver, spleen, and gall bladder. Similarly, XAP131, which from Northern data was found to be expressed in heart, brain, liver, skeletal muscle, pancreas, and kidney, exhibited homologies to ESTs that were derived from heart and brain but also from lung. From the Northern hybridizations, XAP137 was expressed only in heart and muscle, whereas the EST expression information showed that it was also expressed in brain tissue. It was moreover informative to find ESTs derived from tissues that were not covered by the multiple tissue Northern blots. The ubiquitously expressed XAP125 cDNA contig for example, contained an EST that was derived from a human breast cDNA library. Likewise, the ubiquitously expressed XAP136 cDNA contig was found to be expressed also in pineal gland tissue and XAP139 was found to be expressed additionally in spleen, breast, and ovary.

Although XAP121 did not hybridize on a Northern blot, it was likely that this transcript is expressed at very low levels. Therefore, we checked its expression pattern by reverse transcriptase (RT)-PCR analyses. Primers were chosen

within the putative ORF of the cDNA contig using the Oligo 4.1 program (Medprobe, Oslo, Norway). The primers were tested for sequence specificity on dilutions of the same cDNA templates that had been used for cDNA enrichment. On performing RT-PCR on mRNA from human fetal brain, fetal liver, and adult skeletal muscle (Clontech), XAP121 was found to be expressed in skeletal muscle and fetal liver (Table 1). By comparing these data with the tissue expression of the homologous human ESTs, its expression in liver was confirmed and was found additionally to be expressed in brain and spleen.

DISCUSSION

To date, transcriptional analysis has been focusing on the GC-rich regions within human Xq28 (Korn et al. 1992; Bione et al. 1993; Sedlacek et al. 1993) and this has led to the identification of a large number of genes. Little is known about the content and distribution of genes in GC-poor regions, however. Rare cutter restriction sites representing putative CpG islands are almost absent within 1 Mb from the telomere (Poustka et al. 1991; Maestrini et al. 1992; Palmieri et al. 1994; Rogner et al. 1994), leading to the assumption that the organization or the nature of transcriptional units could also be different in this region (Duret et al. 1995). To study the 1.1-Mb region distal to the factor VIII gene in more detail, we constructed cosmid contigs that covered the region almost entirely. The construction of a transcriptional map necessitated the assembly of a cosmid map that simultaneously provided a basis for refined mapping and genomic sequencing. Cosmids were selected by Alu long-range PCR and direct cDNA selection products by using YACs as substrates. This permitted the construction of cosmid and transcript maps in parallel and demonstrated that the approach can be efficiently applied to regions already covered by YACs but not by cosmids.

To identify putative genes, we employed a strategy that combines direct cDNA selection from YACs with exon trapping that was carried out on Alu-PCR and Alu long-range PCR products, as well as on cosmids. A combination of cDNA selection and exon trapping was chosen to overcome the different drawbacks inherent to these two methods. Direct cDNA selection is very sensitive but frequently leads to the enrichment of unspliced or genomic sequences (Rogner et al. 1995), whereas the use of exon trap libraries is

ROGNER ET AL.

limited by the accumulation of cryptic splice products and a high redundancy factor making the isolation of authentic and unique exon trap products laborious (Brennan and Hochgeschwender 1995). A combination of both methods has a synergistic effect and artifacts are eliminated almost completely (Heiss et al. 1996a). However the isolation of cDNAs is also restricted when using the combined approach because only those cDNAs that were enriched by both methods can be isolated. To maximize the isolation of novel transcripts, we carried out exon trapping on three different templates and hybridized the cDNA selection products onto a conventional cDNA library. In this way, five of six known genes and 15 novel cDNA contigs were identified and mapped to the region. Despite using a combined approach when isolating cDNAs, we cannot exclude the possibility that not all transcripts were identified in this 1.1-Mb region.

Because most of the novel cDNAs showed no homologies to known genes, proteins or protein motifs, a considerable amount of characterization will be required to understand the structure, nature, and function of these genes. XAP120, however, demonstrated significant homologies to several synaptobrevin-related sequences from different species (Table 1) and because the homology did not exceed 84% at the DNA level and 68% at the protein level, this suggested that XAP120 may represent a novel member of this gene family. Like the rat cellubrevin, XAP120 was expressed in all tissues tested, whereas other synaptobrevins are reported to be more brain-specific (McMahon et al. 1993). XAP121 exhibited significant homology to the p64 bovine chloride channel protein (Table 1) and possibly represents the human homolog of this gene. The XAP121 transcript was not detectable on Northern blots indicative of its low expression in the tissues tested. This was in agreement with the observation made by Landry et al. (1993) who likewise found low levels of expression of the bovine protein in all cells.

To verify that the identified transcripts belonged to the telomeric region, each cDNA contig was mapped to the YACs and cosmids distal to the factor VIII gene (Fig. 2). Northern blot expression studies, however, revealed that several cDNAs detected multiple transcripts raising the possibility that these could have arisen owing to complex alternative splicing or alternatively, owing to the presence of multiple gene families. In the latter case, cross-hybridizations may have

contributed to a localization in Xq28. The chance of having isolated members of different gene families that do not map to Xq28 was rather low because of the combined approaches that were used in the construction of the transcript map. Furthermore, the many redundant sequences that comprised an assembled cDNA contig were identical and not similar. Nonetheless, hybridization analyses on total genomic DNA were carried out in order to confirm and extend the cosmid and YAC mapping data. A single locus in Xq28 was demonstrated for all cDNA contigs except for XAP135, which exhibited one locus in Xq28 as well an additional locus elsewhere in the genome.

The mapping data showed that the cDNA contigs were not evenly distributed along the region. Fourteen transcripts mapped within the proximal 550 kb, corresponding to one transcript every 40 kb on average. This result is in good agreement to the findings in a 700-kb region around DXS52 (Heiss et al. 1996b) and a 600-kb region between DXS304 and DXS497 (Kioschis et al. 1996) in Xq28. A 170-kb region proximal of XAP125 was completely devoid of cDNA contigs and only five cDNA contigs mapped to the most distal part (Fig. 2). Considering that four contigs belong to only two genes, the average distribution of genes would be one in every 130 kb for the pseudoautosomal region. The distance between the two cDNA contigs XAP126 and XAP139 that exhibited an identical expression pattern was ~180 kb suggestive of a gene that spans a distance comparable with that of the factor VIII gene (Gitschier et al. 1984). Our results for the region distal to the factor VIII gene are in contrast to reports for the GC-rich regions around the L1CAM and G6PD genes (Korn et al. 1992; Bione et al. 1993; Sedlacek et al. 1993), where gene densities of up to one gene in 15 kb have been described. Although additional novel genes may ultimately fill the gaps, taking the present literature into consideration, the present map represents the most complete transcript map of the candidate region for IP2. Ultimately however, completing analyses of the enriched cDNA libraries in conjunction with a global genomic sequencing approach, will facilitate identification and integration of all the genes in this region.

Because incontinentia pigmenti (IP2) is a disease that affects the development of tissues derived from the ectoderm and neuroectoderm, all brain-expressed transcripts should putatively be

NOVEL TRANSCRIPTS IN CANDIDATE REGION FOR IP2

considered as candidate genes for this disease. Testing for the expression of the novel cDNAs in tissues that are symptomatically altered in patients affected with IP2 will be required. Because full-length cDNA sequences are necessary as a first step in determining the function of the genes, the present cDNA sequences are currently being extended by further screening of cDNA libraries and rapid amplification of sequence ends (RACE). Here again, sequencing of this region will be required to gain better insight into the structure and distribution of the genes in this region. In an attempt to identify the gene underlying the familial form of incontinentia pigmenti, a systematic analysis of the sequences for mutations on patient DNA has been initiated. In light of other linkage analyses that have shown that besides IP2, three other diseases, Waismann syndrome (Gregg et al. 1991), Goeminne syndrome (TKC) (Zuffardi and Fraccaro 1982), and periventricular heterotopia (Eksioglu et al. 1996) are linked to distal Xq28, the novel transcripts should also be considered as putative candidate genes for these disorders.

METHODS

Cosmid and cDNA Libraries

Cosmids indicated by Qc were isolated from an Xq28-specific cosmid library constructed from a somatic cell hybrid carrying human Xq28 on a hamster background (Warren et al. 1990; Kioschis et al. 1991). Cosmids from the Imperial Cancer Research Fund (ICRF) flow-sorted X chromosome libraries (ICRFc104 and ICRFc100) are indicated by D and E, respectively (Nizetic et al. 1991). Clones from the Lawrence Livermore National Laboratory chromosome X library (LLOXNCO1 and LLNLc110) are indicated by LL. Oligo(dT)-primed cDNA libraries and were constructed at the ICRF by S. Meier-Ewert and M.-L. Yaspo from digest of DNA from fetal brain (ICRFp507) and fetal thymus (ICRFp508) RNA and ligated into pSPORT.

cDNA Selection

Direct cDNA selection was essentially performed as described by Korn et al. (1992). YAC inserts were excised from low-melting point agarose pulsed-field gels and phenol-chloroform purified. One microgram of DNA was subjected to biotinylation by nick translation using biotin-16-dUTP. Competition of 450-ng labeled DNA was carried out at 65°C for 2 hr using 50 μ g of Cot-1 DNA, 5 μ g of pYAC4 vector and 50 μ g of total yeast DNA in 750 mM NaCl, 50 mM sodium phosphate (pH 7.2), and 5 mM EDTA. The cDNA (100 ng) was subsequently hybridized to the biotinylated DNA at 65°C for 16 hr. The biotinylated DNA was then bound to streptavidin-coated magnetic beads for 1 hr at room temperature. The beads were washed six times with

0.2 \times SSC at 65°C. The cDNAs were eluted in TE buffer, PCR amplified, and size-selected on Chromaspin 400 columns (Clontech). A second round of enrichment was followed by amplification of the cDNAs using SK primers tailed with dUMP residues. The cDNA was re-size-selected, and aliquots were cloned into pAMP10 and transformed into DH5 α cells.

Alu-PCR

Alu-PCR reactions were carried out using the primers Alu-5', Alu-3' (Tagle and Collins 1992), 517, TC-65, ALE1, ALE3, PDF66, and PDJ34 in separate reactions and primers Alu-S and ALU-J (Parrish and Nelson 1994) mixed in a ratio of 5:1 (Wilgenbus et al. 1995). The sequence CGGAATTC-GGATCCGAGG providing an *EcoRI* and *BamHI* site was added to the primers. Alu-PCR conditions were as described previously (Rogner et al. 1994). Alu long-range PCR (Wilgenbus et al. 1995; Heiss et al. 1996a) was carried out with an initial 2 min denaturation step at 94°C, followed by 35 cycles consisting of 20 sec at 96°C, 1 min at 55°C, 5 min at 72°C, and a final extension of 10 min at 72°C. PCR products deriving from the same YAC were pooled and purified over Qiaquick spin columns (Qiagen).

Exon Trapping

Exon trapping on Alu-PCR products was performed as described previously (Heiss et al. 1996a), using the *EcoRI* and *BamHI* restriction sites in the Alu primers for cloning into pSpl3b vector (Integrated Genetics). One hundred fifty-four cosmids were pooled according to seven regions of overlap of the YACs I011, I792, C869, and I234 (Rogner et al. 1994), digested with *EcoRI* and *HindIII*, and cloned into pSpl3b. After electroporation into DH5 α cells, DNA was isolated from 28 libraries consisting of 8,000–10,000 clones each. Two to 3 μ g of phenol-purified DNA was transfected into COS-7 cells by lipofection in serum-free medium using Lipofectamine (Gibco). Total RNA was Trizol-extracted (Gibco). First-strand cDNA synthesis was carried out using the SA2 (ATC TCA GTG GTA TTT GTG AGC) primer. Primary PCR was performed with an initial denaturation for 2 min at 94°C, followed by 10 cycles with 1 min at 94°C, 1 min at 60°C, 1 min and 3 min at 72°C, and a final 10 min extension at 72°C using the SA2 and SD6 (TCTGAGTCACCTGGACAACC) primers. PCR products were digested with 50 units of *BstXI* at 55°C for 16 hr. Secondary PCR was carried out using the USD2 (CUACU-ACUACUAGTGAAGTGCACCTGTGACAAGCTGC) and USA4 (CUACUACUACUACACCTGAGGAGTGAAT-TGGTC) primers under the same conditions except that the samples were cycled 25 times. Aliquots of secondary PCR products were used as complex hybridization probes on gridded cDNA filters, cosmid digest filters, or cloned into pAmp10 vector (Gibco).

Sequencing

Plasmid DNA was extracted using the Bio Robot 9600 and Qiawell Ultra System (Qiagen). Sequencing reactions were performed with the *Taq* Dye Deoxy Terminator Cycle Se-

ROGNER ET AL.

quencing Kit (Perkin Elmer, ABI) and run on an ABI 373A automated DNA sequencer.

Southern Blotting and Hybridization

Southern blotting, probe labeling, competition, and hybridization were performed as described in Poustka (1990).

RT-PCR

First strand cDNA synthesis was performed using 1 μ g of polyA⁺ mRNA and the Marathon Amplification Kit as recommended by the supplier (Clontech). The cDNAs were diluted 1:20 to 1:200 for RT-PCR amplification. RT-PCR was carried out with a 3 min denaturation step at 94°C, followed by 30 cycles with 1 min denaturation at 94°C, 1 min annealing at 60°C, and 1 min extension at 72°C. The final extension was 10 min at 72°C. The same conditions were used for STS amplification with the exception that the annealing was done at 50°C.

ACKNOWLEDGMENTS

We thank Petra Wilgenbus and Alexandra Ochs for extensive sequencing work, Sabine Henze for excellent technical assistance, and Gregor Haberhauer and Renate Gaul for helping with the computer analyses. We are indebted to Dr. Assunta Albanese for the picking and spotting of cDNA clones. We thank Dr. Günther Zehetner, Nadja Pohl, and Iris Dorbandt for providing gridded library filters and clones. We also thank Prof. Jean-Louis Mandel for the St35-probes. We gratefully acknowledge Dr. Zdenek Sedlacek for critical reading of the manuscript. The X chromosome-specific cosmid library LLNLc110 was constructed at the Human Genome Center, Lawrence Livermore National Laboratory under auspices of the National Laboratory Gene Library Project sponsored by U.S. Department of Energy. This work was supported by grants from the Deutsche Forschungsgemeinschaft and the European Community Genome Analysis Program. The sequence data described in this paper have been submitted to the EMBL data library under accession nos. X97503–X97519.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Altschul, S.F., W. Gish, W. Miller, E.W. Meyers, and D.J. Lipman. 1990. Basic alignment search tool. *J. Mol. Biol.* **215**: 403–410.

Bione, S., F. Tamanini, E. Maestrini, C. Tribioli, A. Poustka, G. Torri, S. Rivella, and D. Toniolo. 1993. Transcriptional organization of a 450-kb region of the human X chromosome in Xq28. *Proc. Natl. Acad. Sci.* **90**: 10977–10981.

Brennan, M.B. and U. Hochgeschwender. 1995.

Commentary: So many needles, and so much hay. *Hum. Mol. Genet.* **4**: 153–156.

Brown, W.R.A., P.J. Mackinnon, A. Villasante, N. Spurr, V.J. Buckle, and M.J. Dobson. 1990. Structure and polymorphism of human telomere-associated DNA. *Cell* **63**: 119–132.

Buckler, A., D. Chang, S. Graw, D. Brook, D. Haber, P. Sharp, and D. Housman. 1991. Exon amplification: A strategy to isolate mammalian genes based on RNA splicing. *Proc. Natl. Acad. Sci.* **88**: 4005–4009.

Carney, R.G. 1976. Incontinentia pigmenti: A world statistical analysis. *Arch. Dermatol.* **112**: 535–542.

Church, D.M., C.J. Stotler, J.L. Rutter, J.R. Murrell, J.A. Troffatter, and A.J. Buckler. 1994. Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. *Nature Genet.* **6**: 98–105.

Coleman, R., S.A. Genet, J.I. Harper, and A.O. Wilkie. 1993. Interaction of incontinentia pigmenti and factor VIII mutations in a female with biased X inactivation, resulting in hemophilia. *J. Med. Genet.* **30**: 497–500.

Dear, S. and R. Staden. 1991. A sequence assembly and editing program for efficient management of large projects. *Nucleic Acids Res.* **19**: 3907–3911.

D'Esposito, M., A. Ciccodicola, F. Gianfrancesco, T. Esposito, L. Flagiello, and M. D'Urso. 1996. A novel gene in the Xq28 PAR. *Nature Genet.* **13**: 227–229.

Dietrich, A., B. Korn, and A. Poustka. 1992. Completion of the physical map of Xq28: The location of the gene for L1CAM on the human X chromosome. *Mamm. Genome* **3**: 168–172.

Duret, L., D. Mouchiroud, and C. Gautier. 1995. Statistical analysis of vertebrate sequences reveals that long genes are scarce in GC-rich isochores. *J. Mol. Evol.* **40**: 308–317.

Eksioglu, Y.Z., I.E. Scheffer, P. Caedenas, J. Knoll, F. DiMario, G. Ramsby, M. Berg, K. Kamuro, S.F. Berkovic, G.M. Duyk, J. Parisi, P.R. Huttenlocher, and C.A. Walsh. 1996. Periventricular heterotopia: An X-linked dominant epilepsy locus causing aberrant cerebral cortical development. *Neuron* **16**: 77–87.

Fisch, P., A. Forster, P.D. Sherrington, M.J. Dyer, and T.H. Rabbitts. 1993. The chromosomal translocation t(X;14)(q28;q11) in T-cell pro-lymphocytic leukemia breaks with one gene and activates another. *Oncogene* **8**: 3271–3276.

Freije, D. and D. Schlessinger. 1992. A 1.6-Mb contig of Yeast Artificial Chromosomes around human factor VIII gene reveals three regions homologous to probes DXS115 and two for the DXYS64 locus. *Am. J. Hum. Genet.* **51**: 66–80.

Gitschier, J., T.M. Goralka, K.L. Wion, E.Y. Chen, D.H.

NOVEL TRANSCRIPTS IN CANDIDATE REGION FOR IP2

- Eaton, G.A. Vehar, D.J. Capon, and R.M. Lawn. 1984. Characterization of the human factor VIII gene. *Nature* **312**: 326–330.
- Gorski, J.L., E.N. Burchright, C.E. Harden, C.K. Stein, T.W. Glover, and E.L. Reyner. 1991. Localization of DNA sequences to a region within Xp11.21 between incontinentia pigmenti (IP1) X-chromosomal translocation breakpoints. *Am. J. Hum. Genet.* **48**: 63–64.
- Gregg, R.G., A.B. Metzberg, K. Hogan, G. Sekhon, and R. Laxova. 1991. Waisman syndrome, a human X-linked recessive basal ganglia disorder with mental retardation: Localization to Xq27.3-qter. *Genomics* **9**: 701–706.
- Heiss, N.S., B. Korn, U.C. Rogner, and A. Poustka. 1996a. Generation of specific exon trap probes from YACs by using Alu long-range PCR products. *Methods Mol. Cell Biol.* (in press).
- Heiss, N.S., U.C. Rogner, P. Kioschis, B. Korn, and A. Poustka. 1996b. Transcriptional mapping in a 700 kb region around the DXS2 locus in Xq28: Isolation of six novel transcripts and a novel ATPase isoform (hPMCA5). *Genome Res.* **6**: 478–491.
- Isselbacher, K.J., E. Braunwals, J.D. Wilson, J.B. Martin, A.S. Fauci, and D.L. Kasper. 1994. *Harrison's principles of internal medicine*, 13th edition. McGraw-Hill Inc., New York, NY.
- Kenwrick, S., B. Levinson, S. Taylor, A. Shapiro, and J. Gitschier. 1992. Isolation and sequence of two genes associated with a CpG island 5' of the factor VIII gene. *Hum. Mol. Genet.* **1**: 179–186.
- Kermouni, A., E. Van Roost, K.C. Arden, J.R. Vermeesch, S. Weiss, D. Godelaine, J. Flint, C. Lurquin, J.-P. Szikora, D.R. Higgs, P. Marynen, and J.-C. Renaud. 1995. The IL-9 receptor gene (IL9R): Genomic structure, chromosomal localization in the pseudoautosomal region of the long arm of the sex chromosomes, and identification of IL9R pseudogenes at 9qter, 10pter, 16pter, and 18pter. *Genomics* **29**: 371–382.
- Kioschis, P., B. Gross, D. Nizetic, G. Zehetner, H. Lehrach, and A. Poustka. 1991. Molecular analysis of the Xq28 region. *Cytogenet. Cell. Genet.* **58**: 2070.
- Kioschis, P., U.C. Rogner, E. Pick, S.M. Klauk, N. Heiss, R. Siebenhaar, B. Korn, J.F. Coy, J. Laporte, S. Liechti-Gallati, and A. Poustka. 1996. A 900 kb contig and ten new transcripts within the candidate region for myotubular myopathy (MTM1). *Genomics* **33**: 365–373.
- Korn, B., Z. Sedlacek, A. Manca, P. Kioschis, D. Konecki, H. Lehrach, and A. Poustka. 1992. A strategy for the selection of transcribed sequences in Xq28 region. *Hum. Mol. Genet.* **4**: 235–242.
- Kvaloy, K., F. Galvagni, and W.R.A. Brown. 1994. The sequence organization of the long arm pseudoautosomal region of the human sex chromosomes. *Hum. Mol. Genet.* **3**: 771–778.
- Lakich, D., H.H. Kazazian, S.E. Antonarakis, and J. Gitschier. 1993. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nature Genetics* **5**: 236–241.
- Landry, D., S. Sullivan, M. Nicolaidis, C. Redhead, A. Edelman, M. Field, Q. Al-Awqati, and J. Edwards. 1993. Molecular cloning and characterization of p64, a chloride channel protein from kidney microsomes. *J. Biol. Chem.* **268**: 14989–14955.
- Levinson, B., S. Kenwrick, D. Lakish, G. Hammonds, and J. Gitschier. 1990. A transcribed gene in an intron of the human factor VIII gene. *Genomics* **7**: 1–11.
- Levinson, B., S. Kenwrick, P. Gamel, K. Fisher, and J. Gitschier. 1992. Evidence for a third transcript from the human factor VIII gene. *Genomics* **14**: 585–589.
- Li, L. and D. Hamer. 1995. Recombination and allelic association in the Xq/Yq homology region. *Hum. Mol. Genet.* **4**: 2013–2016.
- Lovett, M., J. Kere, and L.M. Hinton. 1991. Direct selection: A method for the isolation of cDNAs encoded by large genomic regions. *Proc. Natl. Acad. Sci.* **88**: 9628–9632.
- Maestrini, E., F. Tamanini, P. Kioschis, E. Gimbo, P. Marinelli, C. Tribioli, M. d'Urso, G. Palmieri, A. Poustka, and D. Toniolo. 1992. An archipelago of CpG islands in Xq28: Identification and fine mapping of 20 new CpG islands of the human X chromosome. *Hum. Mol. Genet.* **1**: 275–280.
- McMahon, H.T., Y.A. Ushkaryov, L. Edelmann, E. Link, T. Binz, H. Niemann, R. Jahn, and T.C. Südhof. 1993. Cellubrevin is a ubiquitous tetanus-toxin substrate homologous to a putative synaptic vesicle fusion protein. *Nature* **364**: 346–349.
- Migeon, B.R., J. Axelman, S. de Beur, D. Valle, G.A. Mitchell, and K.N. Rosenbaum. 1989. Selection against lethal alleles in females heterozygous for incontinentia pigmenti. *Am. J. Hum. Genet.* **44**: 100–106.
- Naylor, J., A. Brinke, S. Hassock, P.M. Green, and F. Giannelli. 1993. Characteristic mRNA abnormality found in half the patients with severe haemophilia A is due to large DNA inversions. *Hum. Mol. Genet.* **2**: 1773–1778.
- Naylor, J.A., D. Buck, P. Green, H. Williamson, D. Bentley, and F. Giannelli. 1995. Investigation of the factor VIII intron 22 repeated region (int22h) and the associated inversion junctions. *Hum. Mol. Genet.* **4**: 1217–1224.
- Nizetic, D., G. Zehetner, A.P. Monaco, L. Gellen, B.D. Young, and H. Lehrach. 1991. Construction, arraying, and high density screening of large insert libraries of human chromosomes X and 21; Their potential use as reference libraries. *Proc. Natl. Acad. Sci.* **88**: 3233–3237.
- Palmieri, G., G. Romano, A. Ciccodicola, A.

ROGNER ET AL.

- Casamassimi, C. Campanile, T. Esposito, V. Cappa, A. Lania, S. Johnson, R. Reinbold, A. Poustka, D. Schlessinger, and M. d'Urso. 1994. YAC contig organization and CpG island analysis in Xq28. *Genomics* **24**: 149–158.
- Parimoo, S., S.R. Patanjali, H. Shukla, D.D. Chaplin, and S.M. Weissman. 1991. cDNA selection: Efficient PCR approach for the selection of cDNAs encoded in large chromosomal DNA fragments. *Proc. Acad. Natl. Sci.* **88**: 9623–9627.
- Parrish, J. and D. Nelson. 1994. Practical aspects of fingerprinting human DNA using Alu polymerase chain reaction. *Methods Mol. Cell Biol.* **5**: 74–77.
- Perlino, E., G. Paonessa, and G. Ciliberto. 1985. Alu sequences transcription in *X. laevis* oocytes: Nuclear-cytoplasmic partitioning and evidence for 3' end processing reactions. *Nucleic Acids Res.* **13**: 8359–8377.
- Poustka, A. 1990. Physical mapping by PFGE. *Methods: A Companion to Methods in Enzymology* **1**: 204–211.
- Poustka, A., A. Dietrich, G. Langenstein, D. Toniolo, S.T. Warren, and H. Lehrach. 1991. Physical map of human Xq27-qter: Localizing the region of the fragile X mutation. *Proc. Natl. Acad. Sci.* **88**: 8302–8306.
- Reed, V., S. Rider, G.L. Maslen, E. Hatchwell, H.J. Blair, I.C. Uwechue, I.W. Craig, S.H. Laval, A.P. Monaco, and Y. Boyd. 1994. A 2-Mb YAC contig encompassing three loci (DXF34, DXS14, and DXS390) that lie between Xp11.2 translocation breakpoints associated with incontinentia pigmenti type 1. *Genomics* **20**: 341–346.
- Rogner, U.C., P. Kioschis, K. Wilke, W. Gong, E. Pick, A. Dietrich, U. Zechner, H. Hameister, A. Pragliola, G.E. Herman, J.R.W. Yates, H. Lehrach, and A. Poustka. 1994. A YAC clone map spanning 7.5 megabases of human chromosome band Xq28. *Hum. Mol. Genet.* **3**: 2137–2146.
- Rogner, U.C., K. Wilke, E. Steck, B. Korn, and A. Poustka. 1995. The melanoma antigen (MAGE) family is clustered in the chromosomal band Xq28. *Genomics* **29**: 725–731.
- Sedlacek, Z., B. Korn, D.S. Konecki, R. Siebenhaar, J.F. Coy, P. Kioschis, and A. Poustka. 1993. Construction of a transcription map of a 300 kb region around the human G6PD locus by direct cDNA selection. *Hum. Mol. Genet.* **11**: 1865–1869.
- Sefiani, A., R. M'rad, L. Simard, A. Vincent, C. Julier, L. Holvoet-Vermaut, S. Heuertz, N. Dahl, J.F. Stadler, M.O. Peter, C. Moraine, J. Maleville, J. Boyer, I. Oberlé, D. Labuda, and M.C. Hors-Cayla. 1991. Linkage relationship between incontinentia pigmenti (IP2) and nine terminal X long arm markers. *Hum. Genet.* **86**: 297–299.
- Shafit-Zagardo, B., J.J. Maio, and F.L. Brown. 1982. KpnI families of long interspersed repetitive DNAs in human and other primate genomes. *Nucleic Acids Res.* **10**: 3175–3193.
- Smahi, A., C. Hyden-Granskog, B. Peterlin, P. Vabres, S. Heuertz, M.C. Fulchignoni-Lataud, N. Dahl, P. Labrune, B. Le Marec, C. Piussan, A. Taieb, H. von Koskull, and M.C. Hors-Cayla. 1994. The gene for the familial form of incontinentia pigmenti (IP2) maps to the distal part of Xq28. *Hum. Mol. Genet.* **3**: 273–278.
- Tagle, D.A. and F.S. Collins. 1992. An optimized Alu-PCR primer pair for human-specific amplification of YACs and somatic hybrids. *Hum. Mol. Genet.* **1**: 121–122.
- Thick, J., Y.F. Mak, J. Metcalfe, D. Beatty, and A.M. Taylor. 1994. A gene on chromosome Xq28 associated with T-cell prolymphocytic leukemia in two patients with ataxia telangiectasia. *Leukemia* **8**: 564–573.
- Warren, S.T., S.L. Knight, J.F. Peters, C.L. Stayton, G.G. Consalez, and F. Zhang. 1990. Isolation of the human chromosomal band Xq28 within somatic cell hybrids by fragile site breakage. *Proc. Natl. Acad. Sci.* **87**: 3856–3860.
- Wilgenbus, K.K., A. Mincheva, B. Korn, P. Lichter, and A. Poustka. 1995. IRS-long range (LR) PCR: A simple method for efficient amplification of human genomic DNA from complex sources. *Methods Mol. Cell Biol.* **5**: 214–221.
- Wood, W.I., D.J. Capon, C.C. Simonsen, D.L. Eaton, J. Gitschier, B. Keyt, P.H. Seeburg, D.H. Smith, P. Hollingshead, K.L. Wion, E. Delwart, E.G.D. Tuddenham, G.A., Vehar, and R.M. Lawn. 1984. Expression of active human factor VIII from recombinant clones. *Nature* **312**: 330–337.
- Yokoi, H., S. Hadana, H. Kogi, X. Kang, K. Wakasa, and J.-E. Ikeda. 1994. Isolation of expressed sequences encoded by the human Xq terminal portion using microclone probes generated by laser microdissection. *Genomics* **20**: 404–411.
- Zuffardi, O. and M. Fraccaro. 1982. Gene mapping and serendipity. The locus for Torticollis, Keloids, Cryptorchidism and Renal Dysplasia (31430, McKusick) is at Xq28, distal to the G6PD locus. *Hum. Genet.* **62**: 280–281.

Received February 7, 1996; accepted in revised form June 27, 1996.