European Journal of Human Genetics (2001) 9, 121–129


www.nature.com/ejhg

ARTICLE

A high-resolution integrated map spanning the *SDHD* gene at 11q23: a 1.1-Mb BAC contig, a partial transcript map and 15 new repeat polymorphisms in a tumour-suppressor region

Bora E Baysal*[,1], Joan E Willett-Brozick[1], Peter EM Taschner[2], JG Dauwerse[2], Peter Devilee[2] and B Devlin[1]

[1]*Department of Psychiatry, The University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA;*
[2]*Department of Human Genetics, The Leiden University Medical Center, Leiden, The Netherlands*

Chromosomal region 11q22-q23 is a frequent target for deletion during the development of many solid tumour types, including breast, ovary, cervix, stomach, bladder carcinomas and melanoma. One of the most commonly deleted subregions contains the *SDHD* gene, which encodes the small subunit of cytochrome *b* (cybS) in mitochondrial complex II (succinate-ubiquinone oxidoreductase). Germline mutations in *SDHD* cause hereditary paraganglioma type 1 (*PGL1*), and suggest a tumour suppressor role for cybS. We present a high-resolution physical map spanning *SDHD*, covered by 19 YACs and 20 BACs. An approximate 1.1-Mb gene-rich region around *SDHD* is spanned by a complete BAC contig. Twenty-six new STSs are developed from the BAC clone ends. In addition to the discovery and characterisation of 15 new simple tandem repeat polymorphisms, we provide integrated positional information for 33 ESTs and known genes, including KIAA1391, *POU2AF1* (*OBF1*), *PPP2R1B*, *CRYAB*, *HSPB2*, *DLAT*, *IL-18*, *PTPS*, KIAA0781 and KAIA4591, which is mapped by *Not*I site cloning. We describe full-length transcript sequence for *PPP2R1B*, encoding the protein phosphatase 2A regulatory subunit A beta isoform. We also discover a processed pseudogene for *USA-CYP*, a cyclophilin associated with U4/U6 snRPNs, and a novel gene, *DDP2*, encoding a mitochondrial protein similar to the X-linked deafness-dystonia protein, which is juxtaposed 5′-to-5′ to *SDHD*. This map will help assess this gene-rich region in PGL and in other common tumours. *European Journal of Human Genetics* (2001) 9, 121–129.

Keywords: chromosomal band 11q23; hereditary paraganglioma; LOH; physical map; imprinting

Introduction

Somatic deletions in the chromosomal region 11q22-q23 are frequent for many tumour types, including breast, ovary, cervix, stomach, bladder carcinomas and melanoma.[1] These deletions map to several non-overlapping sub-regions, thus revealing potential sites harbouring tumour suppressor activity. Because a 6 Mb interval between D11S2000 and D11S897 is one of the most frequently deleted sub-regions in these common tumours, it may harbour either several distinct tumour suppressor genes or a single gene of fundamental importance to the process of neoplasia.

In fact, several candidate tumour suppressor genes are located into this region. Germ-line mutations in *ATM* cause ataxia-telangiectasia (AT).[2] AT patients have a predisposition to malignancies, such as B-cell lymphomas and T-cell leukaemias; however, the role of *ATM* as a tumour suppressor gene in solid tumours is controversial.[3] Somatic mutations involving *PPP2R1B*, which encodes a protein phosphatase 2A subunit, have been identified in a small fraction of common

*Correspondence: Bora E Baysal, 3811 O'Hara Street, Western Psychiatric Institute and Clinic, R1445, Pittsburgh, PA, 15213-2593, USA.
Tel: +1 412 624 1093; Fax: +1 412 624 1754;
E-mail: baysalbe@msx.upmc.edu
Received 24 July 2000; revised 21 September 2000; accepted 26 September 2000

tumours that show regional loss-of-heterozygosity (LOH).[4,5] Whether *PPP2R1B* serves a tumour suppressor role at the cellular level remains to be established.[6] Finally, *SDHD* encodes the small subunit of the cytochrome *b* (cybS) in mitochondrial complex II (succinate-ubiquinone oxidoreductase).[7,8] Germline mutations in *SDHD* cause hereditary paraganglioma type 1 (*PGL1*), a rare disorder characterised by the development of mostly benign, highly vascular tumours in the head and neck.[9] PGL tumours develop in an autosomal dominant pattern after paternal transmission of the mutant gene, whereas no phenotype occurs after maternal transmission, suggesting that *SDHD* is subject to genomic imprinting.[10]

The germline loss-of-function mutations in the paternal alleles and the subsequent somatic loss of normal maternal alleles in PGL tumours define a tumour suppressor role for *SDHD*. Based on the phenotypic similarity between the PGL tumours and the CB's response to chronic hypoxic exposure, we have hypothesised that loss of *SDHD* compromises oxygen sensing and leads to tumour development in the paraganglionic tissues. Because it has been postulated that hypoxic environment provides a selective pressure towards more aggressive phenotypes in solid tumours,[11] it is conceivable that loss of *SDHD* could play a wider role in common tumorigenesis.

We present a high-resolution physical map around *SDHD* to help refine minimal critical regions of LOH at 11q23 in common tumours and to facilitate assessment of candidate tumour suppressor genes. This map contains an overlapping set of small-insert-YAC and BAC clones, as well as large-insert CEPH mega-YACs across a 3 Mb region, and provides integrated physical information for 33 ESTs/known genes and 15 new (CA)n simple tandem repeat polymorphisms (STRPs).

Materials and methods

EST/STS content mapping, BAC library screening, insert analysis, and identification of vector end sequences

EST/STS content mapping was performed using standard PCR amplification as described elsewhere.[12] We initially selected positional candidate ESTs using the radiation hybrid data at Whitehead Institute for Biomedical Research/MIT Center for Genome Research (http://carbon.wi.mit.edu:8000/cgi-bin/contig/phys_map). BAC library screening, clone purification, and DNA isolation were performed according to the instructions provided by Research Genetics, with slight modifications. BAC insert sizes were identified by pulsed field gel electrophoresis (PFGE)[12] following *Not*I restriction enzyme digestion (New England Biolabs). Individual CEPH mega-YAC and CITB Cal-Tech human BAC clones, PCR primers for ESTs/genes and other STSs were purchased from Research Genetics (Huntsville, AL, USA). Small insert YAC clones were obtained from a chromosome 11-specific YAC library constructed at the Roswell Park Cancer Institute (RPCI).[13] BAC clone ends were isolated by creating *Alu*I, *Rsa*I and *Hae*III bubble-vectorette libraries of clones,[14] and by PCR amplification with a vector end specific primer and a bubble primer. The primers specific for BAC vector pBeloBAC11 ends were BAC-T7: 5′-CGGGGATCCTCTAGAGTCGA-3′, BAC-SP6: 5′-ATGACCATGATTACGCCAAGCT-3′. To isolate ends of small insert RPCI YAC yRP4-f-7 from bubble-vectorette libraries, we used pYAC4 primers in the vicinity of its cloning site.

Fluorescent *in situ* hybridisation (FISH) analysis

Regionally localised BACs were labelled by nick-translation with biotin-14-dUTP (Gibco BRL) or digoxigenin-16-dUTP (Roche Diagnostics) as described by Kievits *et al*.[15] *In situ* hybridisation using the labelled BACs as probes was performed according to Dauwerse *et al*.[16] Slides containing Epstein Barr virus-transformed cells were processed for Fiber-FISH as described by Datson *et al*.[17] For Fiber-FISH, pairs of BACs labelled in red and green were used as probes.

Identification and characterisation of novel STRP and gene-specific SNP

To identify novel CA repeats, we used regionally-localised BACs; 15 (CA)10-NN primers, excluding (CA)11, were individually coupled with bubble primer to PCR amplify regional BAC bubble-vectorette libraries. Annealing temperature was 58°C. After PCR amplification of 30–35 cycles, products were analysed by agarose gel electrophoresis. Selected bands were isolated from agarose gels using commercial protocol (Quiaex II gel purification kit, Qiagen), and sequenced from both directions with a linker specific primer and the relevant (CA)n-NN primer. Based on the new sequence identified from one side of the (CA)n, a unique reverse primer was designed and bubble-vectorette libraries amplified through the (CA)n repeat to obtain sequence from the other side of the repeat. Heterozygosities for the new repeats were assessed by polyacrylamide gel (6%) electrophoresis, using PGEM3zf/-40 sequencing reaction products as a molecular weight marker. SNPs were identified by single strand conformational polymorphism (SSCP) analysis and sequencing. All DNA sequencing reactions were performed by ABI 373A version 1.2.0 fluorescent sequencer, as described by the manufacturer.

Cloning *Not*I sites

To identify *Not*I sites, regional BAC clones were isolated and subjected to Pulsed Field Gel Electrophoresis (PFGE) after *Not*I digestion. A panel of 4 bp-frequent cutter enzymes, including *Rsa*I, *Hae*III, *Alu*I, *Sau3A*I, were used to digest the BAC clones containing *Not*I sites and the products were ligated to plasmid vector pBlueScript SK-(Stratagene), which had been digested with *Sma*I (for blunt end fragments) and *Bam*HI (for *Sau*3AI). The ligated molecules were subcloned into XL-1 Blue bacterial host by electroporation (GenePulser, BioRad). Then the transformants were double-digested with *Not*I/

*Eco*RI and analysed for inserts using standard methodology. The plasmids containing *Not*I inserts were partially sequenced and subjected to BLAST analysis to identify nucleotide matches.

PPP2R1B 3′-UTR characterisation

A regional EST WI-19325, which is contained in the Whitehead whole-genome EST radiation hybrid mapping database (http://www-genome.wi.mit.edu), mapped close to *PPP2R1B*. In addition, a cDNA contig THC297422, identified in the Institute for Genomic Research (TIGR) database (http://www.tigr.org), matched WI-19325. Several large-insert cDNA clones including #396996, 396962, 524344, 400072, 520238 mapping inside this cDNA contig were identified, purchased from ATCC, and partially sequenced by an ABI 373A Version 1.2.0 fluorescence sequencer (Applied Biosystems). The available cDNA sequence was further extended by sequential BLAST analysis[18] using the GenBank EST database (http://www.ncbi.nlm.nih.gov/BLAST). To detect and analyse matching human EST sequences, the complete *PPP2R1B* 3′-UTR sequence, stripped of repetitive sequences using RepeatMasker program (http://Dot.imgen.bcm.tmc.edu:9331/seq-util/seq-util.html), was used in BLAST analysis (search performed July 1999).

Results

Contig construction and STS mapping

To nucleate a physical map in the region between D11S1986 and D11S1347, the last published borders bracketing *PGL1* prior to its cloning, we selected a set of STSs, ESTs, CEPH mega- and RPCI small-insert YACs obtained from either

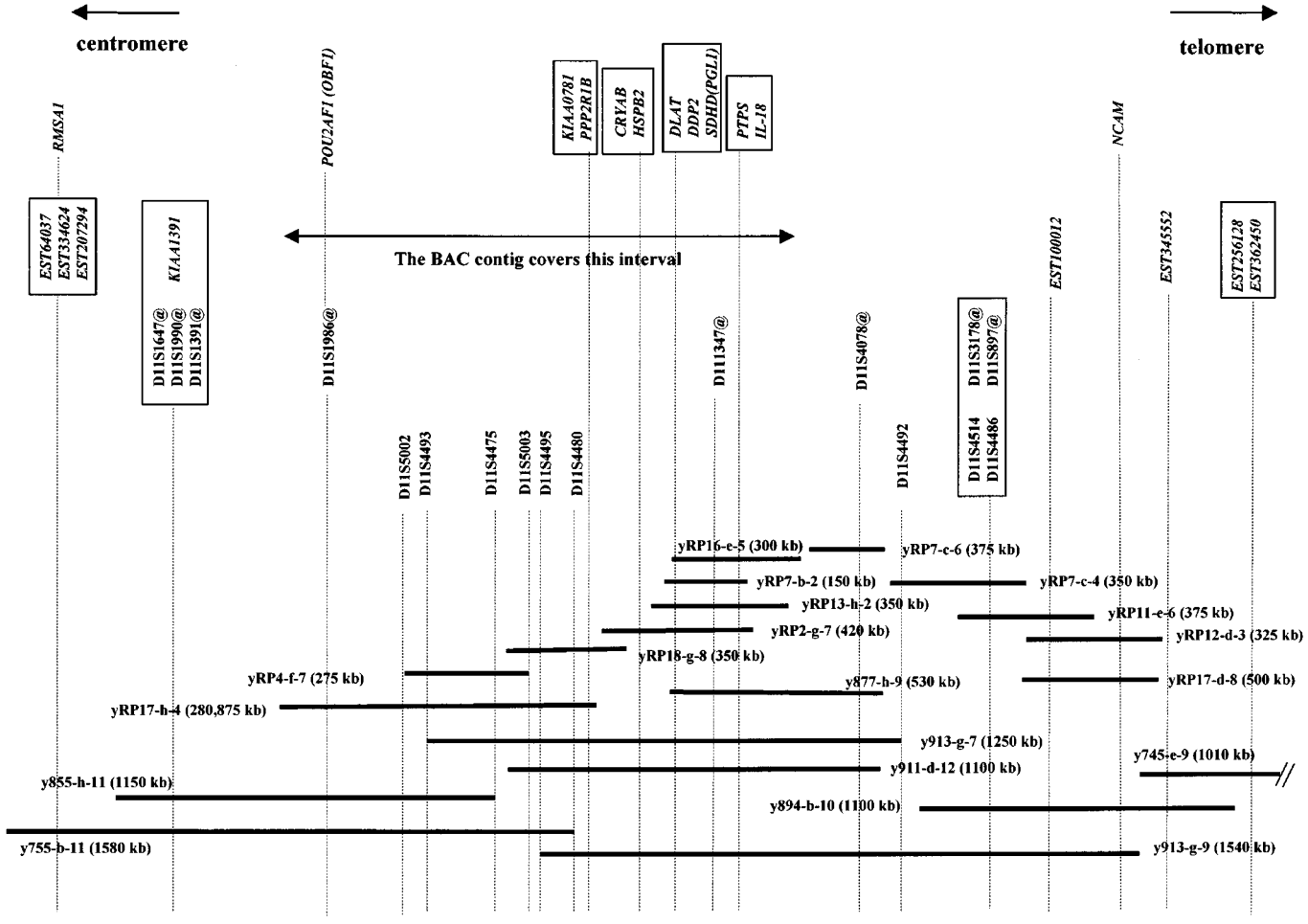

Figure 1 Contig of large-insert CEPH mega-YACs, and small-insert RPCI YACs. The contig spans more than 3-Mb. The insert sizes of the clones are depicted in parenthesis. Genes, ESTs, STRPs (shown with @ extension), and non-polymorphic STSs are localised to distinct intervals defined by clone overlaps. The following genes/ESTs from 11q23 do not map to this genomic interval: PC8, HSPA8, SGC33878, SGC30085, SGC34343, WI-11320, D11SS2294E, D11SS2285E, WI-17649, D11S2284E, WI-19664, WI-20930, SGC34431, WI-19587, WI-18163, WI-19362, D11S2293E, D11S2226E, D11S2251E, D11S2320E.

genome-wide or previous regional mapping data. This initial map provided full mega- and small-insert YAC-contigs spanning the *PGL1* critical region (Figure 1). This map located D11S1986 approximately 1 Mb centromeric to D11S1347, less than originally estimated based on other regional maps.[19] Our analysis showed D11S1986 is an expressed STRP derived from the 3′-untranslated region of *POU2AF1* (*OBF1*), which encodes the B-cell octamer binding factor.[20]

Using these initial YAC contig data, we selected several ESTs and STSs for PCR-screening of BAC library pools. New STSs were created from BAC insert ends and used to confirm their regional localisation and to determine clone overlap patterns. After many rounds of BAC end sequencing and library screening, a BAC contig composed of 20 BACs, including three BACs detected in the TIGR database, were constructed (Figure 2). The BAC contig spanned approximately 1.1 Mb. The insert sizes and *Not*I restriction sites of each BAC were analysed by PFGE, uncovering the presence of three *Not*I sites in the region, consistent with a previous regional *Not*I map.[21]

To create new STSs, the insert-ends of 15 selected clones were sequenced. Of the 27 STSs created, 26 mapped back to the clones from which they were derived and to other regional clones. The STS derived from the SP6-end of B103-G-1 did not map back to other regional clones, thus implying potential chimerism in this BAC. Some insert ends contained expressed sequences as detected by BLAST analysis. These expressed sequences included *POU2AF1* (*OBF1*),[22] *IL-18*, encoding interleukin 18,[23] and a novel processed pseudogene

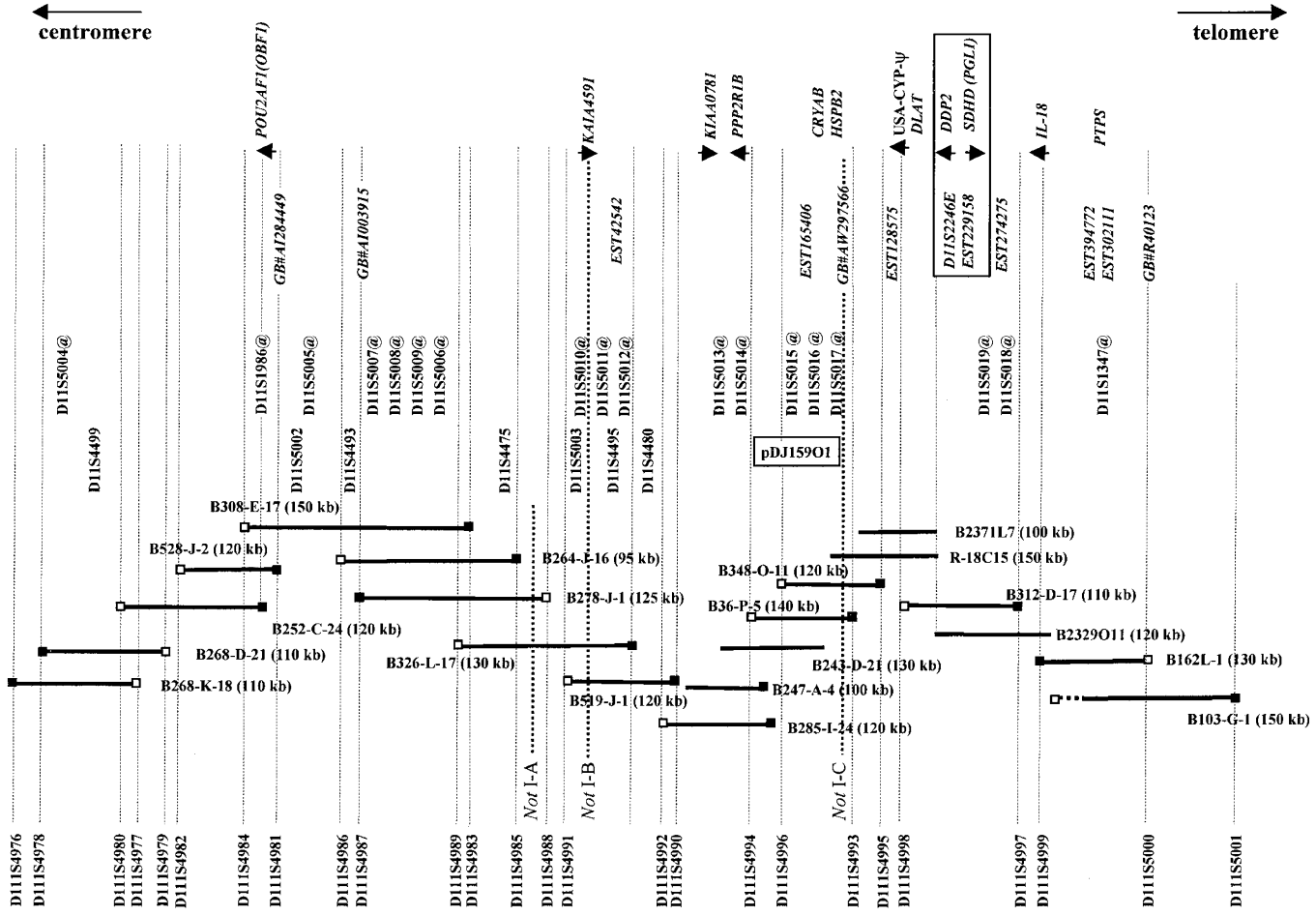

Figure 2 1.1-Mb BAC contig around the *SDHD* gene. The contig is composed of 20 BACs with their sizes indicated in parenthesis. Genes and ESTs are in italics. STRPs are labelled with an '@' extension. New STSs derived from the BAC ends are depicted at the bottom. Open and closed boxes at the end of BACs denote the SP6- and T7-insert-ends, respectively. BACs B2371L7, R-18C15 and B2329O11 are identified in the TIGR database by searching with a genomic sequence containing *SDHD*. The pDJ159O1 box shows the location of a previously characterized PAC clone[34] and the five STRPs used in the repositioning of the *PGL1* critical region.[19] Two of the STRPs, D11S5011 and D11S5019, have been previously described.[9] The arrows underlying some genes depict their 5′-to-3′ orientation. dbEST numbers are used to label ESTS.

for *USA-CYP*, encoding a cyclophilin associated with U4/U6 snRPNs.[24,25] Complete sequencing of the *USA-CYP* pseudo-gene revealed that it is 92% identical to the *USA-CYP* mRNA sequence, lacks introns, contains in-frame premature stop codons, and is followed by a poly(A) sequence. These features strongly suggest this pseudo-gene has arisen by retrotransposition of *USA-CYP* mature RNA.

Fluorescent *in situ* hybridisation (FISH) analysis

To confirm BAC overlap patterns, seven clones, B308-E-17, B278-J-1, B519-J-1, B285-I-24, B36-P-5, B348-O-11, and B312-D-17, were selected from the BAC contig. None of these clones was found to be chimaeric in metaphase FISH (data not shown). The order and orientation of these clones in the BAC contig, which had been determined by STS-based methods, was confirmed by Fiber-FISH (Figure 3). Predicted overlap patterns for each clone were observed. The BAC pairs B278-J-16/B519-J-1 and B348-O-11/B312-D-17 did not reveal any overlap consistent with the STS content analysis. No gaps were observed between the clones in each pair. Since the maximum resolution of fiber-FISH is about 5 kb, we estimated that the BAC ends are less than 5 kb apart. The distance from the centromeric end of B308-E-17 to the telomeric end of B312-D-17 is estimated to be about 700 kb. The continuity of the BAC signals also demonstrated the integrity of the clones.

Identification of novel (CA)n STRP and SNP

To search for novel (CA)n repeats, several BACs were selected. Using a PCR-based approach, three to four (CA)n sequences

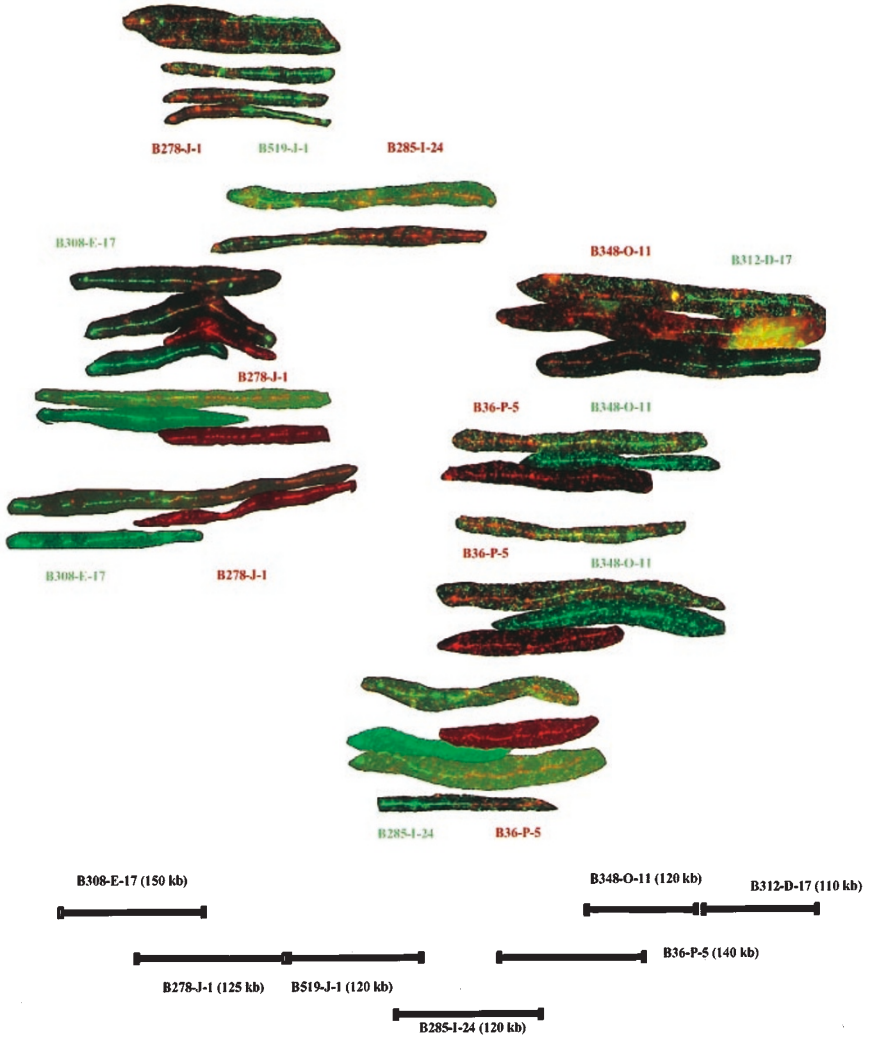

Figure 3 Fiber-FISH mapping of seven 11q23 BACs. Two or three examples of fluorescent signals from each pair of BACs, which have been labelled in red and green, confirm their order as shown in the contig underneath. The overlap between the BAC pairs B308-E-17 and B278-J-1, B285-I-24 and B36-P-5 and B348-O-11 can clearly be observed as a yellow signal. In case of overlap, the signal from each BAC is also shown separately.

per BAC were identified. Sequences were first assessed for the number uninterrupted CA repeats, on the assumption that the degree of polymorphism would be proportional to the repeat number. Loci having less than 8 CA repeats were not evaluated further. On one occasion, however, captured sequences revealed a (TA)n motif, which proves to be a highly polymorphic locus (D11S5004). Thirteen of the new STRPs had heterozygosities larger than 0.5, as determined by analysing a number of unrelated Caucasian chromosomes (Table 1). SSCP analysis of regional genes revealed four new coding region SNPs; a synonymous and a non-synonymous SNP in *CRYAB*, a synonymous SNP in *HSP2*, and a 3′-UTR SNP in KIAA0781.

Transcript mapping

Gene/EST mapping showed a high transcript density in the region, and provided relative locations of KIAA1391, *POU2AF1* (*OBF1*), *PPP2R1B*, *CRYAB*, *HSPB2*, *DLAT*, *IL-18*, *PTPS*, KIAA0781 and KAIA4591. While a number of ESTs clustered to the centromeric third of y755-b-11 (Figure 1), their order could not be established in the absence of small-insert clones covering that region. KIAA1391, a novel gene that mapped in a region frequently deleted in B-cell chronic lymphocytic leukaemia,[26] shows strong homology to a chromosome X-linked rho-GTPase-activating protein predominantly expressed in haematopoietic cells.[27]

Several ESTs in databases, including EST222225 (D11S966E), derived from KIAA0781, a large transcript encoding a novel serine-threonine kinase.[28] By mapping KIAA0781-derived amplicons in the contig, we found this gene to be oriented 5′-centromeric to 3′-telomeric, with its 3′-end located only 10 kb from the 3′-end of *PPP2R1B* (data not shown) (Figure 2).

In addition to *SDHD* and *DLAT*, EST274275, EST198329, EST229158, and D11S2246E mapped to BAC B312-D-17. By BLAST analysis, all three ESTs clustered to an approximate 29 kb genomic sequence (GenBank accession number AB026906) containing the exons for *SDHD*: while EST198329 and EST229158 map immediately 5′ to *SDHD*, D11S2246E, which is mostly composed of repeat sequences, is derived from *SDHD* intron 3. Further database analysis revealed that EST198329 is derived from *DDP2*, encoding a protein similar to the mitochondrial protein DDP of X-linked deafness dystonia syndrome and that EST229158 was derived

Table 1 New simple tandem repeat polymorphisms in the vicinity of *SDHD*

DNA seg. number	*Locus*	*Derived from BAC clone/ Repeat motif*	*Primer sequences*	*Range of alleles No. of alleles/heterozygosity/ No. of chromosomes tested*
D11S5004	B268D21-TA	B268D21	5′-GTGAATCCTCCATAGTCTGGT-3′	134–182 bp
		Complex (TA)n	5′-AAGCCAATTGCAATGGGAAGT-3′	8/0.82/35
		Largest *n*=12		
D11S5005	B308E17CA-C	B308E17/$(CA)_{14}$	5′-CCACTACACTATCACTACACT-3′	99–113 bp
			5′-GTTGCCAGTGTGGTACATGT-3′	7/0.76/77
D11S5006	B308E17CA-F	B308E17	5′-GAGCCAAGATCGAGTCACT-3′	214–226 bp
		$(CA)_{19}..(AAC)_8$	5′-GGGTCCTTTTGACATACCTCA-3′	6/0.77/36
D11S5007	B264J16CA-CC	B264J16	5′-CCTCTTAGGACTCTCCTGTT-3′	234–238 bp
		Complex (CA)n	5′-ACAATGTTCCATGAACTGAGTGT-3′	3/0.51/26
		Largest *n*=8		
D11S5008	B264J16CA-TC	B264J16	5′-AATAATTAAGATCATGACACC-3′	93–103 bp
		$(CA)_{18}$	5′-GGAAGGAGAAGAGGAAATGCT-3′	5/0.62/36
D11S5009	B264J16CA-GT2	B264J16	5′-ATCATGCACCAAGAGAAGTAC-3′	126–140 bp
		$(CA)_{16}$	5′-CTTTAGTAGTCTATCGTCTATCT-3′	3/0.52/31
D11S5010	B519J1CA-A[a]	B519J1	5′-AGCTGGGAGTACAGGTGTG-3′	212–220 bp
		$(CA)_{14}$	5′-CTCAGTTTTCTAATACTCTATTGA-3′	5/0.76/65
D11S5012	B519J1CA-C	B519J1	5′-GCACTCCTATGTTCATTACAG-3′	185–197 bp
		$(CA)_{18}$	5′-GTTTCATATGGCAGGAGTCC-3′	6/0.73/34
D11S5013	B247A4CA-A	B247A4	5′-AAACGCAGCAAGACCCTGTT-3′	140–158 bp
		$(CA)_{10}$	5′-AGGGGTTCATGACCATCCTA-3′	3/0.11/45
D11S5014	B247A4CA-F	B247A4	5′-GTCTCTCCAAGTTGGGTAGTA-3′	138–162 bp
		$(CA)_{12}$	5′-ATAGGCATGTGCCACCACAT-3′	6/0.79/33
D11S5015	B348O11CA-A	B348O11	5′-GCAACTCAAGAGCCCATCAA-3′	147–172 bp
		$(CA)_{20}$	5′-CCATGTTGTAGCATGTGACA-3′	8/0.79/23
D11S5016	B348O11CA-C	B348O11	5′-GTTCCTGCTACTGTTCTCCA-3′	144–154 bp
		$(CA)_{20}$	5′-TTATGGAGATGAGACCAGACTA-3′	6/0.66/45
D11S5017	B348O11CA-D	B348O11	5′-GCGTATGCATGCGCTTGCA-3′	136–148 bp
		$(CA)_{17}$	5′-GGCCGTTTGTGGCCTTAAGT-3′	7/0.61/42
D11S5018	B312D17CA-AC	B312D17	5′-CTTTCTGGCTCGTAGCCGTA-3′	240–244 bp
		$(CA)_{12}$	5′-TGAATGAGGAAACGATCATAGA-3′	2/0.26/35
D11S5020	*PPP2R1B*	*PPP2R1B*	5′-TCCCACTGTTGCTGAGACAT-3′	134–146 bp
			5′->CATCCTTGGCAGCTTACATCA-3′	7/0.78/45

[a]A null allele is observed for D11S5010.

from a distinct gene with no detectable homology (GenBank accession number AK001588). The analysis of *DDP2* (GenBank accession numbers AF150087, AF165967) and *SDHD* (GenBank accession number NM_003002) transcripts in Genbank database revealed that the two genes are oppositely oriented and that their transcription start sites were separated by less than 150 bp.

*Not*I site cloning

Because most *Not*I sites in the human genome mark CpG islands that are associated with expressed sequences,[29] we attempted to clone the *Not*I sites in this region of 11q23. We cloned and sequenced two of the three sites, *Not*I-C and *Not*I-B (Figure 2), identified in the BAC contig. Sequences derived from both *Not*I sites mapped back to the regional clones by hybridisation and PCR amplification.

BLAST analysis of the sequences around the *Not*I sites identified unmapped EST sequences. The sequence of **5′-CTCATTGAAGCCCTCCTGCAGGACGTCCAG-TAAGTTCCCTCGGGTCAGGCAGGCTAG-CATTGTTCCTGCTGGGGGCTCCCCGGGCGGGGTCTC-CAGCCCG**GCGGCCGC**-3′** flanked the *Not*I-C site (underlined) and identified a complete match (in bold) to several ESTs similar to alpha actinin (for example GenBank accession number AW297566). The sequence of **5′-ACTTCGCTGTGGTGAAGCTGGGGCGGCACCGGATCAC-CAAGACGGAGGTG**CGGCCCGGGGCTCGGCGG-GAGCGTCCGAGGCGAGGGTTCGGGAGAGGAGCTGCT-TACCGAGAGGGGCGGCCGC-3′ flanked the *Not*I-B site (underlined) and identified a complete match (in bold) to cDNA clone KAIA4591 (GenBank accession number AK000396), which also showed significant sequence overlap with the 5′-region of KIAA0781 (above). These data suggest that 5′-end of KIAA0781 spans the *Not*I-B site which is located at least 100-kb centromeric from its 3′-end (Figure 2).

Full characterisation of PPP2R1B transcript

PP2R1B has two major transcript isoforms on Northern analysis, 2.5 and 5.0 kb.[14] cDNA sequences compatible with the 2.5 kb transcript have been previously identified. By EST mapping, cDNA clone sequencing and database search, we obtained a 5.6 kb cDNA sequence and showed that the larger transcript arises by alternative polyadenylation signal usage. The extended 3′-UTR sequence of *PPP2R1B* also revealed a novel (CA)n STRP, D11S5020, which is highly polymorphic (Table 1). By RT – PCR experiments, we confirmed the newly-identified sequence derives from the 3′-UTR of *PPP2R1B*, it is not interrupted by an intron at the genomic level and that D11S5020 is an expressed STRP (data not shown). The complete 5.6 kb *PPP2R1B* sequence has been deposited to GenBank under accession number AF163473.

Finally, to test whether positions of most *PPP2R1B* EST 3′-ends would be compatible with either 2.5 kb or 5.0 kb sizes observed in Northern analysis, we mapped the ends of 3′-

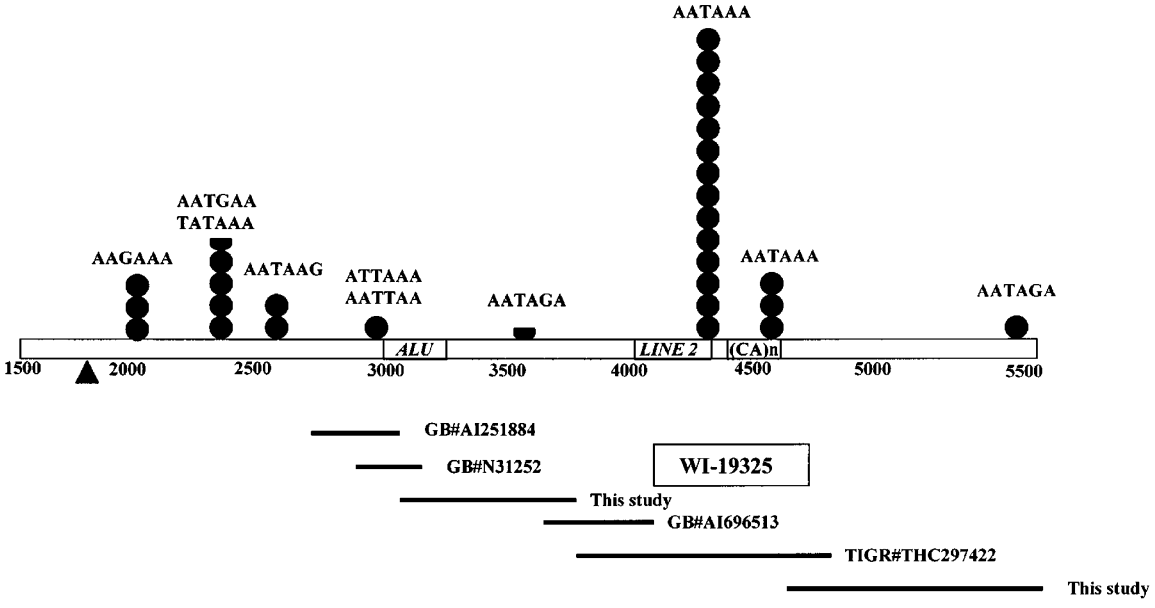

Figure 4 Summary of the characteristics of *PPP2R1B* 3′-UTR. The vertical arrowhead denotes the stop codon at nt 1873.[14] The horizontal bars below the cDNA scale denote sources of sequences initially used to assemble the 3′-UTR sequence and filled circles denote 3′-ends of matching EST sequences in GenBank. Each filled semi-circle corresponds to a single EST sequence. Poly(A) signal motifs preceding the transcript ends are shown on top. Most of 3′-ends are located after the first consensus Poly(A) signal positioned at the end of a LINE repeat sequence. Alu repeat and (CA)n sequences are also shown. The numbering of the cDNA shown on the scale follows the one used in Baysal *et al*.[14]

PPP2R1B-derived EST sequences in GenBank EST database onto the 5.6 kb transcript sequence. The 3′-ends of EST clones cluster in two regions (Figure 4): 33% between nt 2050 – nt 2650, and 59% between nt 4290 – nt 2630. Thus, with added Poly(A) tails, these transcript sizes approximated the observed major band sizes in the Northern analysis. All mapped EST clones had either the consensus polyadenylation signal, 5′-AATAAA-3′, or one of its single nucleotide neighbours within 75 bp of termination.[30] Almost half of the EST ends (48%) mapped following the first consensus polyadenylation signal at nt 4272. Interestingly, this major consensus signal occurs within a long interspersed repeat element, LINE2, and adds *PPP2R1B* to the growing number of genes that utilise repetitive elements for transcript polyadenylation.[31]

Discussion

First draft genomic sequences are now available for many chromosomal regions through the multi-nationally sponsored Human Genome Project. However, until ordered, gap-free and fully-annotated genomic sequences are available for each chromosome, high-resolution, informative physical maps, especially in regions of biological interest, should continue to be useful by complementing raw sequence data. Three mega-YAC-based physical maps spanning this region of chromosome 11q23 have been previously reported,[21,32,33] however they do not provide detailed coverage for the interval presented in this study. The abundance of transcripts and the presence of three *Not*I sites within an approximate 1.1-Mb interval indicate a gene-rich genomic domain around the *SDHD* gene. Partial sequence data from a PAC clone pDJ159O1 that maps within this region also confirms high transcript density.[34] The abundance of transcripts is consistent with the presence of distinct tumour suppressor genes for different tumour types, although a single tumour suppressor gene fundamentally important to neoplastic transformation cannot be ruled out.

The mapped transcripts already suggest several candidate tumour suppressor genes. For example, KIAA1391 and *POU2AF1* (*OBF1*) map on mega-YAC y755-b-11, which is frequently deleted in B-cell chronic lymphocytic leukaemia.[26] KIAA0781, a large novel serine-threonine kinase, is localised immediately centromeric to *PPP2R1B*. This map also reveals that *PPP2R1B* and *SDHD* are separated by only 250 kb. Thus large somatic deletions spanning this region cannot be informative regarding targeted gene(s); instead, identification of intragenic deletions and/or mutations will be necessary to evaluate each gene. The high density of genes within the BAC contig is also reflected by the noteworthy proximity. For example, *CRYAB* and *HSPB2* are oriented 5′-to-5′ within less than a 1-kb intergenic interval, and they potentially share some regulatory elements.[35] A similar mechanism of coordinated gene expression may also operate for *SDHD* and *DDP2*, which show immediate 5′-to-5′ organisation.

Our discovery of 15 new STRPs brings their total number to 24 within the 1.1-Mb interval spanned by the BAC contig, including two STRPs within 400-kb of the *SDHD* gene[9] and five STRPs derived from regional PAC clone pDJ159O1.[19] The high intensity of STRPs provides a powerful tool for the evaluation of new PGL families for linkage and haplotype analyses, as well as for the assessment of PGL and other tumour types for somatic deletions. If certain alleles of the newly characterised, expressed dinucleotide repeat at *PPP2R1B*, D11S5020, affect mRNA stability, they may be associated with higher risk of tumorigenesis. Finally, although the inheritance of the disease phenotype in PGL1 is determined strictly by the sex of the transmitting parent, previous work on *SDHD* showed no evidence for imprinting.[9] The resources developed herein may also be useful to determine the precise mechanism of the parent-specific transmission in PGL1.

Acknowledgements

We thank Tom Shows for the gift of RPCI chromosome 11 small-insert YAC library, Lauren Richey and Jamie Antis for their technical help. This work is supported by a research project grant (VM-180) from American Cancer Society to B Devlin. The EMBL/GENBANK accession numbers for the nucleotide data are AF163473, AQ936561 and AG936560.

References

1 Koreth J, Bakkenist CJ, McGee JO: Chromosome 11q and cancer: a review. *J Pathol* 1999; **187**: 28 – 38.
2 Savitsky K, Bar-Shira A, Gilad S *et al*: A single ataxia telangiectasia gene with a product similar to PI-3 kianse. *Science* 1995; **268**: 1749 – 1753.
3 Yuille MA, Coignet LJ: The ataxia telangiectasia gene in familial and sporadic cancer. *Recent Results Cancer Res* 1998; **154**: 156 – 173.
4 Wang SS, Esplin ED, Li JL *et al*: Alterations of the *PPP2R1B* gene in human lung and colon cancer. *Science* 1998; **282**: 284 – 287.
5 Calin GA, di Iasio MG, Caprini E *et al*: Low frequency of alterations of the alpha (*PPP2R1A*) and beta (*PPP2R1B*) isoforms of the subunit A of the serine-threonine phosphatase 2A in human neoplasms. *Oncogene* 2000; **19**: 1191 – 1195.
6 Campbell IG, Manolitsas T: Absence of *PPP2R1B* gene alterations in primary ovarian cancers. *Oncogene* 1999; **18**: 6367 – 6369.
7 Hirawake H, Taniwaki M, Tamura A, Kojima S, Kita K: Cytochrome b in human complex II (succinate-ubiquinone oxidoreductase): cDNA cloning of the components in liver mitochondria and chromosome assignment of the genes for the large (*SDHC*) and small (*SDHD*) subunits to 1q21 and 11q23. *Cytogenet Cell Genet* 1997; **79**:; 132 – 138.
8 Hirawake H, Taniwaki M, Tamura A, Amino H, Tomitsuka E, Kita K: Characterization of the human *SDHD* gene encoding the small subunit of cytochrome b (cybS) in mitochondrial succinate-ubiquinone oxidoreductase. *Biochim Biophys Acta* 1999; **1412**: 295 – 300.
9 Baysal BE, Ferrell RE, Wilett-Brozick JE *et al*: Mutations in *SDHD*, a mitochondrial complex II gene, in hereditary paraganglioma. *Science* 2000; **287**: 848 – 851.
10 Van der Mey AG, Maaswinkel-Mooy PD, Cornelisse CJ, Schmidt PH, Van de Kamp JJ: Genomic imprinting in hereditary glomus tumours: evidence for new genetic theory. *Lancet* 1989; **2**: 1291 – 1294.

11 Graeber TG, Osmanian C, Jacks T *et al*: Hypoxia-mediated selection of cells with diminished apoptotic potential in solid tumours. *Nature* 1996; **379**: 88 – 91.
12 Baysal BE, Van Schothorst EM, Farr JE, James MR, Devilee P, Richard III CW: A high-resolution STS, EST, and gene-based physical map of the hereditary paraganglioma region on chromosome 11q23. *Genomics* 1997; **44**: 214 – 221.
13 Qin SZ, Nowak NJ, Zhang JL *et al*: A high-resolution physical map of human chromosome 11. *Proc Natl Acad Sci USA* 1996; **93**: 3149 – 3154.
14 Baysal BE, Farr J, Goss J, Devlin B, Richard III CW: Genomic organization and precise physical location of protein phosphatase 2A regulatory subunit A beta isoform gene on chromosome band 11q23. *Gene* 1998; **217**: 107 – 116.
15 Kievits T, Dauwerse JG, Wiegart J *et al*. Rapid subchromosomal localization of cosmids by nonradioactive in situ hybridization. *Cytogenet Cell Genet* 1990; **53**: 134 – 136.
16 Dauwerse JG, Jumelet EA, Wessels JW *et al*: Extensive cross-homology between the long and the short arm of chromosome 16 may explain leukemic inversions and translocations. *Blood* 1992; **79**: 1299 – 1304.
17 Datson NA, Semina E, Van Staalduinen AA *et al*: Closing in on the Rieger syndrome gene on 4q25: mapping translocation breakpoints within a 50-kb region. *Am J Hum Genet* 1996; **59**: 1297 – 1305.
18 Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. *J Mol Biol* 1990; **215**: 403 – 410.
19 Baysal BE, Van Schothorst EM, Farr JE *et al*: Repositioning the hereditary paraganglioma critical region on chromosome band 11q23. *Hum Genet* 1999; **104**: 219 – 225.
20 Strubin M, Newell JW, Matthias P: OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. *Cell* 1995; **80**: 497 – 506.
21 Arai Y, Hosoda F, Nakayama K, Ohki M: A yeast artificial chromosome contig and *Not*I restriction map that spans the tumour suppressor gene(s) locus, 11q22.2-q23.3. *Genomics* 1996; **35**: 196 – 206.
22 Junker S, Brondum-Nielsen K, Newell JW, Matthias P, Tommerup N: Assignment of the human gene for Oct-binding factor-1 (OBF1), a B-cell-specific coactivator of octamer-binding transcription factors 1 and 2, to 11q23.1 by somatic cell hybridization and in situ hybridization. *Genomics* 1996; **33**: 143 – 145.
23 Nolan KF, Greaves DR, Waldmann H: The human interleukin 18 gene IL18 maps to 11q22.2-q22.3, closely linked to the DRD2 gene locus and distinct from mapped IDDM loci. *Genomics* 1998; **51**: 161 – 163.
24 Teigelkamp S, Achsel T, Mundt C *et al*: The 20kD protein of human [U4/U6.U5] tri-snRNPs is a novel cyclophilin that forms a complex with the U4/U6-specific 60kD and 90kD proteins. *RNA* 1998; **4**: 127 – 141.
25 Horowitz DS, Kobayashi R, Krainer AR: A new cyclophilin and the human homologues of yeast Prp3 and Prp4 form a complex associated with U4/U6 snRNPs. *RNA* 1997; **3**: 1374 – 1387.
26 Stilgenbauer S, Liebisch P, James MR *et al*: Molecular cytogenetic delineation of a novel critical genomic region in chromosome bands 11q22.3-q23.1 in lymphoproliferative disorders. *Proc Natl Acad Sci USA* 1996; **93**: 11837 – 11841.
27 Tribioli C, Droetto S, Bione S *et al*: An X chromosome-linked gene encoding a protein with characteristics of a rhoGAP predominantly expressed in hematopoietic cells. *Proc Natl Acad Sci USA* 1996; **93**: 695 – 699.
28 Nagase T, Ishikawa K, Suyama M *et al*: Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. *DNA Res* 1998; **5**: 277 – 286.
29 Lindsay S, Bird AP: Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature* 1987; **327**: 336 – 338.
30 Sheets MD, Ogg SC, Wickens MP: Point mutations in AAUAAA and the poly (A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. *Nucleic Acids Res* 1990; **18**: 5799 – 5805.
31 Mager DL, Hunter DG, Schertzer M, Freeman JD: Endogenous retrovirus provide the primary polyadenylation signal for the two new human genes (HHLA2 and HHLA3). *Genomics* 1999; **59**: 255 – 263.
32 di Iasio MG, Calin G, Tibiletti MG *et al*: Refinement of the LOH region 1 at 11q23.1 deleted in human breast carcinomas and sublocalization of 11 expressed sequence tags within the refined region. *Oncogene* 1999; **18**: 1635 – 1638.
33 Tunnacliffe A, Jones C, Le Paslier D *et al*: Localization of Jacobsen syndrome breakpoints on a 40-Mb physical map of distal chromosome 11q. *Genome Res* 1999; **9**: 44 – 52.
34 Burbee DG, Baysal BE, Korcz J *et al*: Cloning and DNA sequencing of a t(9;11) translocation region associated with bipolar disorder. *Am J Hum Genet* 1996; **59**: A1445.
35 Iwaki A, Nagano T, Nakagawa M, Iwaki T, Fukumaki Y: Identification and characterization of the gene encoding a new member of the alpha-crystallin/small hsp family, closely linked to the alphaB-crystallin gene in a head-to-head manner. *Genomics* 1997; **45**: 386 – 394.