PERGAMON

European Journal of Cancer 38 (2002) 1126–1132

European Journal of Cancer

www.ejconline.com

Novel expressed sequences obtained by means of a suppression subtractive hybridisation analysis from the 6q21 region that is frequently deleted in gastric cancer

B. Carvalho[a], R. Seruca[a], C.H.C.M. Buys[b], K. Kok[b,*]

[a]*Instituto de Patologia e Imunologia Molecular da Universidade do Porto, IPATIMUP, Oporto, Portugal*
[b]*Department of Medical Genetics, University of Groningen, Antonius Deusinglaan 4, 9713 AW Groningen, The Netherlands*

Received 30 July 2001; received in revised form 31 October 2001; accepted 21 November 2001

Abstract

In our search for genomic regions that are involved in the development of gastric cancer, we recently identified a 2-cM minimal region of overlapping heterozygous deletions in 6q16.3-q23.1. Here, we describe an application of the suppression subtraction method (SSH) to search for genes in this small region of the genome, taking advantage of the fact that many human genes present on yeast artificial chromosomes (YACs) are expressed in yeast. Subtraction was performed with two virtually contiguous YACs that cover a region of approximately 2.5 Mb. Combined forward and reversed subtractions resulted in the identification of 12 clones of human origin, all of which could be confirmed by sequence analysis as originating from the 6q21 region. Expression in human tissues could be confirmed by Northern analysis for two of the clones, one of them showing a high level of expression in stomach tissue.

Keywords: Subtraction technique; Stomach neoplasms; Suppressor genes; Chromosomes human pair 6; Chromosomes yeast artificial

1. Introduction

In our search for genomic regions that are involved in the development of gastric cancer, we recently identified a 2-cM minimal region of overlap of heterozygous deletions in 6q16.3-q23.1 [1]. As long as the DNA sequence in many parts of the human genome has not been completed and annotated, one still has to use experimental methods to search for genes in such a limited genomic region. These methods include direct screening of cDNA libraries, methods of PCR-based cDNA selection, cross-species sequence homology searches, screening for CpG islands, exon trapping strategies and screening for splice sites (reviewed in Ref. [2]). Although all of these methods can be useful in the identification of genes, some are very laborious and time-consuming. There is also a variety of methods allowing the identification of differentially expressed genes, such as cDNA subtractive hybridisation (SH) (reviewed in Ref. [3]), differential display (DD) [4], representative difference analysis (RDA) [5], serial analysis of gene expression (SAGE) [6] and suppression subtractive hybridisation (SSH) [7]. These latter methods, however, analyse complete genomes instead of restricted regions. In this study, we describe the application of a subtraction method to analyse a small region of the genome. Our strategy took advantage of the fact that many human genes present on yeast artificial chromosomes (YACs) are expressed in yeast [8]. Thus, in a cDNA subtraction protocol using two yeast clones that contain different recombinant YACs, the resulting pool of cDNAs will be enriched for genes localised on the human inserts of the YACs. Among the various subtraction methods available, we chose the SSH strategy [7,9] to analyse the gene contents of two YACs from the 2-cM smallest region of overlap of heterozygous deletions in gastric cancer.

* Corresponding author. Tel.: +31-50-363-2925; fax: +31-50-363-2947.
E-mail address: k.kok@medgen.azg.nl (K. Kok).

0959-8049/02/$ - see front matter
PII: S0959-8049(02)00014-X

2. Materials and methods

2.1. Isolation of RNA and DNA from yeast

Yeast cells containing different CEPH YACs were grown for 48 h at 30 °C. Total RNA was isolated according to the RNAzol™ B procedure (Campro Scientific, Veenendaal, The Netherlands). Poly A^+ RNA was obtained using a mRNA purification kit (Amersham Pharmacia Biotech, Uppsala, Sweden). DNA was isolated by standard salt-chloroform extraction [10].

2.2. Suppression subtractive hybridisation

First-strand cDNA was synthesised from 2 μg of poly A^+ RNA using AMV reverse transcriptase from Clontech's PCR-Select™ cDNA subtraction kit. Subtraction was performed according to the instructions of Clontech's PCR-Select™ cDNA Subtraction kit, (Clontech, Palo Alto, USA), with the following modification: diluted cDNAs (subtracted and unsubtracted samples) were further diluted 50× before the primary polymerase chain reaction (PCR). PCR products were subcloned into pCR™ II using a TA Cloning kit (Invitrogen, Carlsbad, USA). The transformed bacteria were plated on selective plates and white colonies were selected randomly, resulting in two cDNA libraries, specific for the forward and reverse subtractions, respectively.

2.3. Subtraction efficiency

Primers for the yeast housekeeping gene *ZWF1* (encoding glucose-6-phosphate dehydrogenase, accession no. M34709) were designed in order to check the depletion of non-differentially expressed sequences. Primer sequences were: forward 5′-GGATTCCAGAG GCTTACGAG-3′, reverse 5′-GGGTGCTTTTCGGG CATAAC-3′, resulting in an amplicon of 232 bp. Amplification was for 30 cycles with denaturation at 94 °C for 0.5 min, annealing at 58 °C for 0.5 min, and extension at 72 °C for 2 min. The initial denaturation step at 94 °C and the final extension at 72 °C lasted for 3 min.

2.4. Differential screening

Inserts of the selected colonies were amplified by touch-PCR. The PCR was carried out in a 25-μl reaction volume containing 10 mM Tris–HCl (pH 8.0), 50 mM KCl, 1.5 mM $MgCl_2$, 200 mM of each deoxynucleotide triphosphate, 0.4 mM of each primer (M13 forward 5′-GTAAAACGACGGCCAG-3′; M13 reverse 5′-CAGGAAACAGCTATGAC-3′), and 0.5 U of Taq DNA polymerase (Amersham Pharmacia Biotech). Amplification was for 30 cycles with denaturation at 94 °C for 0.5 min, annealing at 50 °C for 0.5 min, and extension at 72 °C for 1.5 min. The initial denaturation step at 94 °C and the final extension step at 72 °C lasted for 3 min. Part of the PCR product (7 μl) was analysed on a 2% agarose gel. PCR products were denatured by the addition of an equal volume of 0.5M NaOH, and 3-μl aliquots were dot-blotted on nylon membranes (Hybond-N+; Amersham Pharmacia Biotech). Of each filter, four identical copies were made. DNA was cross-linked to the membranes by incubation at 80 °C for 2 h. Membranes were hybridised with both non-cloned subtracted cDNA pools. cDNA clones that hybridised predominantly with the cDNA pool from which they originated, were considered specific for that cDNA pool.

2.5. Southern and northern analysis

Inserts from the selected clones were hybridised to Southern membranes containing *EcoRI* digests of human DNA, DNA from the two YACs used in the subtraction (Y776A5, Y785D2) and from YACs overlapping with them (Y798G12, Y911F5). DNA from YAC 915E4, mapping to distal 6q was used as negative control. The inserts were labelled with 30 μCi [α-^{32}P] deoxycytidine triphosphate (dCTP) by random primer labelling. Probes and Southern membranes were prehybridised with 10 and 15 μg human Cot-1 DNA (GibcoBRL, Life Technologies), respectively. Hybridisation was overnight at 65 °C in a buffer containing 0.5M NaHPO4 (pH 7.0), 7% SDS, and 1 mM ethylene diamine tetraacetic acid (EDTA). Four washings, with solutions of decreasing stringencies (2×SSC (150 mM NaCl, 15 mM Na_3 citrate, pH 7.0)/0.1% sodium dodecyl sulphate (SDS); 1×SSC/0.1% SDS; 0.3×SSC/0.1% SDS; 0.1×SSC/0.1% SDS, respectively), were used at 65 °C. The cDNA clones were hybridised to commercially available multiple-tissue Northern blots that contain mRNA samples from stomach, thyroid, spinal cord, lymph node, trachea, adrenal gland and bone marrow (Clontech), following the supplier's instructions.

2.6. Sequencing analysis

Selected clones were sequenced using an ABI Prism™ 377 DNA Sequencer (Applied Biosystems, Foster City, USA) and M13 primers from the pCR™ vector. The cDNA sequences were compared with the available databases using the basic local alignment search tool (BLAST) (http://www.ncbi.nlm.nih.gov) and the local alignment tool at the UCSC (University of California at Santa Cruz) website (http://genome.ucsc.edu) searches.

3. Results

From the Whitehead Institute database, 4 YACs were selected that according to their sequence tagged site

(STS) marker content mapped within the 2-cM smallest region of overlap (SRO) of heterozygous deletions [1]. YAC 776A5 (990 kb) contains marker D6S278 flanking the SRO on the proximal side and markers D6S1594 and WI1240 that map within the SRO. YAC 785D2 (1370 kb) contains the markers WI3490 and D6S1647 that also map within the SRO. The STS content of each YAC was confirmed by PCR (data not shown). YACs were tested by fluorescent *in situ* hybridisation (FISH) to be non-chimeric. The relative position of the YACs as depicted in Fig. 1 is based on their STS content.

3.1. cDNA subtraction and differential screening

The two virtually contiguous, but non-overlapping, YACs 776A5 and 785D2 were used for subtraction. Forward and reverse subtractions were performed in parallel, with Y776A5 as the tester in the forward subtraction and Y785D2 as the tester in the reverse subtraction. The efficiency of the subtraction was assessed by determining the abundance of the yeast *ZWF1* gene, a housekeeping gene, in the cDNA pools. When the unsubtracted and the subtracted cDNA pools were subjected to PCR using *ZWF1* gene-specific primers, a PCR product with the expected length was obtained only for the unsubtracted cDNA pool (Fig. 2a). This indicates that the subtraction step had indeed resulted in the depletion of the yeast *ZWF1* gene. After subcloning of the subtracted cDNA populations, two libraries of 192 cDNA clones each were generated, corresponding to the forward and the reverse subtraction, respectively. The inserts of the clones, as determined by touch-PCR, varied in length between 100 and 800 bp. Seven out of the 384 clones had no insert.

In the differential screening, the membranes that carried the dot-blotted inserts of the cloned cDNAs were subsequently hybridised with the two cDNA pools resulting from the forward and reversed subtractions. Clones were considered specific for the cDNA pool from which they were derived, when a strong hybridisation signal with that cDNA pool was observed in combination with absence of a signal or a very faint hybridisation signal with the other cDNA pool. Inserts from the clones selected by the differential screening were hybridised to the dot-blotted cDNA libraries to identify identical clones. The 30 clones that were selected for further analysis appeared to represent 21 independent clones (Tables 1 and 2).

3.2. Analysis of the selected clones

All selected clones were hybridised to *EcoR1*-digested human placenta DNA, to *EcoR1*-digested DNA from the YACs that were used in the subtraction, and to *EcoR1*-digested DNA from two other YACs that were overlapping with them. An example of such an analysis is shown in Fig. 2b. Hybridisation of clone BKR4, selected from the forward subtraction, to human DNA resulted in a single hybridising band of approximately 9 kb. When the same clone was hybridised to the two

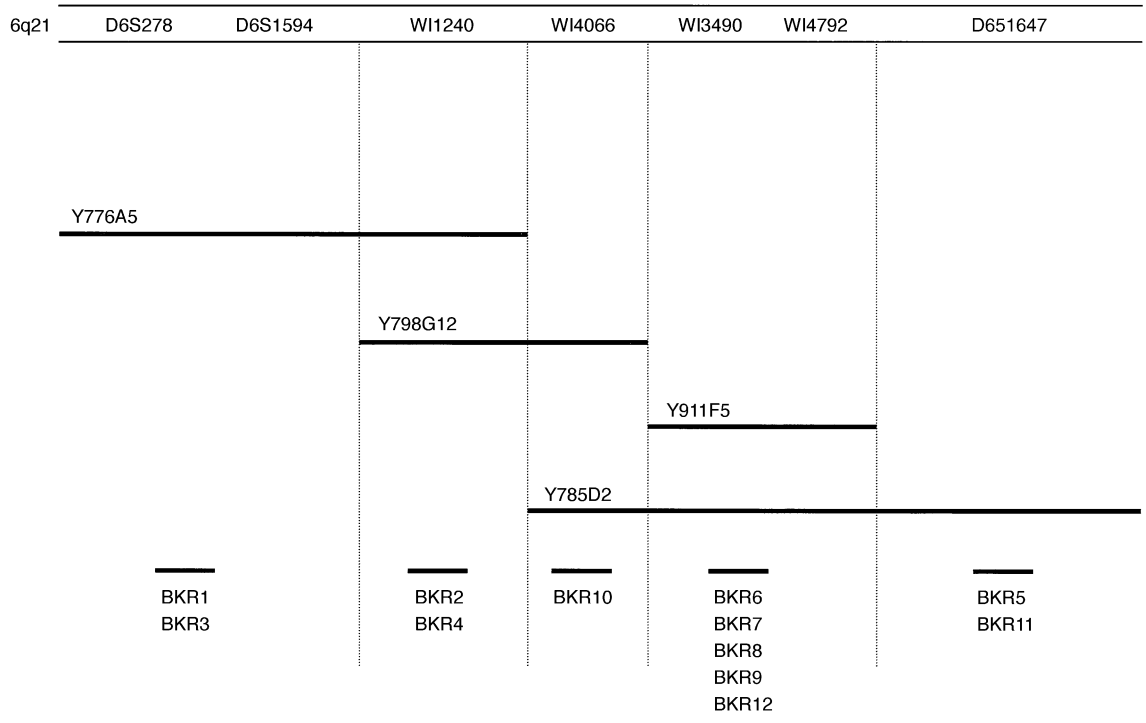

Fig. 1. Schematic representation of the yeast artificial chromosome (YAC) contig in the 6q region and location of the human clones relative to the YACs.

YACs used in the subtraction, a single band of the same size as for human DNA was detected in the tester YAC (Y776A5), but not in the driver YAC (Y785D2). A band with the same size was also observed in Y798G12 that partly overlaps with Y776A5. Data from the Southern analysis of the selected clones are summarised in Table 1 for the forward subtraction and in Table 2 for the reverse subtraction. In total, four out of seven clones of the forward subtraction gave rise to a unique band when hybridised to human DNA and to a band of similar size when hybridised with the tester YAC (Y776A5). The remaining three clones did not give a signal with human DNA, but gave specific and identical bands with all YACs.

In the reverse subtraction, eight out of 14 clones gave rise to a unique band when hybridised to human DNA. All these eight clones gave a band of similar size when hybridised with the tester YAC (Y785D2). For the remaining six clones, a specific signal was seen in the hybridisation to all YACs, but not in the hybridisation to human DNA.

Hybridisation of the clones to the set of four YACs gave us some information about the position of these clones within the YAC contig. A schematic representation of the physical location of the clones is depicted in Fig. 1.

Comparison of the sequences of the 12 human clones (BKR1–BKR12) to the BLAST and BLAT databases did not reveal any homology to known genes. However, all clones did show an almost complete identity with sequences from 6q21 and mapped within a 2.5-Mb segment. As a first approach to verify that the selected clones indeed represented sequences that are transcribed in (a subset of) human tissues, we hybridised their inserts to multiple-tissue Northern blots. For seven clones no signal was observed. Clone BKR8 from the reversed subtraction was not subjected to Northern analysis since the comparison of its sequence with the database showed a high degree of homology with LTR7-HERVH elements. For clone BKR1 of the forward subtraction, a signal was repeatedly found only with trachea mRNA (Fig. 3). The length of the mRNA recognised by clone BKR1 is approximately 2 Kb. Clone BKR12 from the reversed subtraction detected a mRNA of approximately 1.2 Kb in all samples present on the membrane (Fig. 3). The strongest signal was observed for stomach and adrenal gland.

4. Discussion

Our previous study on allelic imbalances in 60 gastric carcinomas with 19 microsatellite markers predominantly from 6q16-q23, resulted in the identification of a smallest

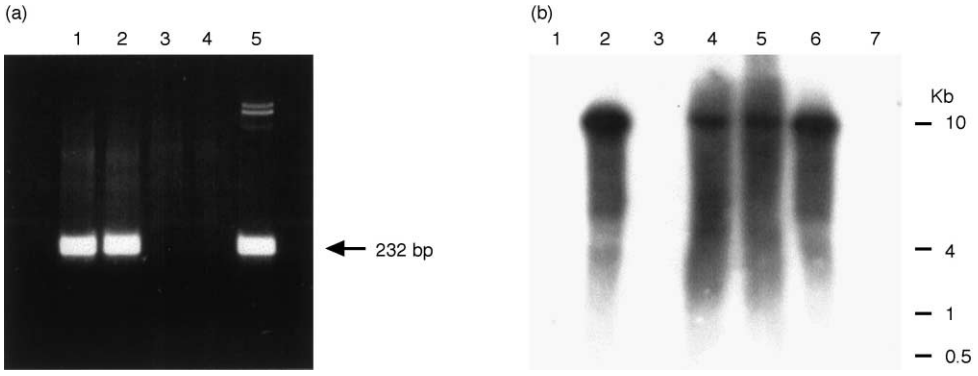

Fig. 2. (a) *ZWF1* depletion after subtraction of yeast artificial chromosomes (YACs) 776A5 and 785D2. Lanes 1 and 2, amplification with *ZWF1*-specific primers of unsubtracted cDNA from 776A5 (tester) and 785D2 (tester), respectively; lanes 3 and 4, amplification with *ZWF1*-specific primers of subtracted cDNA from 776A5 (tester) and 785D2 (tester), respectively; lane 5, genomic DNA from YAC 785D2; (b) Southern analysis of clone BKR4 from Y776A5. Lane 1, Y915E4 (6q27); lane 2, Y776A5; lane 3, Y785D2; lanes 4 and 5, total human DNA; lane 6, Y798G12; lane 7, Y911F5 (all samples digested with *EcoR1*).

Table 1
Southern analysis data of selected clones from the forward subtraction (tester Y776A5)

Clones	GenBank accession no.	Size (bp)	Human placenta	Size (kb)	Y776A5	Y785D2	Y798G12	Y911F5	Y915E4	Similar clones
BKR1	AF439961	364	+	5.0	+	–	–	–	–	$n=5$
26	–	–	–	–	+	+	+	+	+	–
52	–	–	–	–	+	+	+	+	+	–
63	–	–	–	–	+	+	+	+	+	–
BKR2	AF439962	584	+	12.0	+	–	+	–	–	$n=1$
BKR3	AF439963	282	+	7.0	+	–	–	–	–	$n=0$
BKR4	AF439964	480	+	8.5	+	–	+	–	–	$n=6$

region of overlap of heterozygous deletions of approx. 2 cM [1]. Instead of trying to narrow down the region by increasing the number of tumours or by including more markers, we decided to assemble a YAC contig for this region as a starting point for a gene search. The reason to search for genes on these YACs is the fact that the published human sequence still contains large gaps, also in the D6S278–D6S1647 region. Moreover, as our results also appear to indicate, it is likely that not all the human genes have been identified yet, or accurately annotated to the human genome working draft sequence. Thus, there is still a need for experimental approaches to search for unknown human genes. Several methods have been developed to identify genes in large segments of genomic DNA. Direct screening of cDNA libraries with whole YACs can be successful, but is dependent on the expression of the gene within the cDNA library used [11–13]. Exon trapping has also been used successfully, but is a time-consuming procedure known to often generate many false-positives. In general, cosmid-size fragments are used as starting material. That would have necessitated the construction of a cosmid contig of the region of interest.

Reports that human and mouse genes encoded on YACs are transcribed in yeast [8,14,15] suggest that these genes could in principle be identified by methods that analyse differentially expressed genes. By using two yeast clones, containing different YACs in a subtraction protocol, it should be possible to specifically isolate the transcribed sequences that are encoded by these YACs. Thus, a subtraction method could be used to identify unknown genes in a limited region of the genome.

Table 2
Southern analysis data of selected clones from the reverse subtraction (tester Y785D2)

Clones	GenBank accession no.	Size (bp)	Human placenta	Size (kb)	Y776A5	Y785D2	Y798G12	Y911F5	Y915E4	Similar clones
BKR5	AF439965	642	+	3.5	–	+	–	–	–	$n=1$
BKR6	AF439966	422	+	4.0	–	+	–	+	–	$n=0$
11	–	–	–	–	+	+	+	+	+	–
BKR7	AF439967	562	+	9.0	–	+	–	+	–	$n=0$
21	–	–	–	–	+	+	+	+	+	–
BKR8	AF439968	810	+	5.5	—	+	—	+	—	$n=2$
BKR9	AF439969	164	+	5.5	–	+	–	+	–	$n=0$
68	–	–	–	–	+	+	+	+	+	–
BKR10	AF439970	583	+	6.0	–	+	+	–	–	$n=0$
78	–	–	–	–	+	+	+	+	+	–
119	–	–	–	–	+	+	+	+	+	–
BKR11	AF439971	584	+	10.0	–	+	–	–	–	$n=0$
BKR12	AF439972	389	+	7.0	–	+	–	+	–	$n=1$
146	–	–	–	–	+	+	+	+	+	–

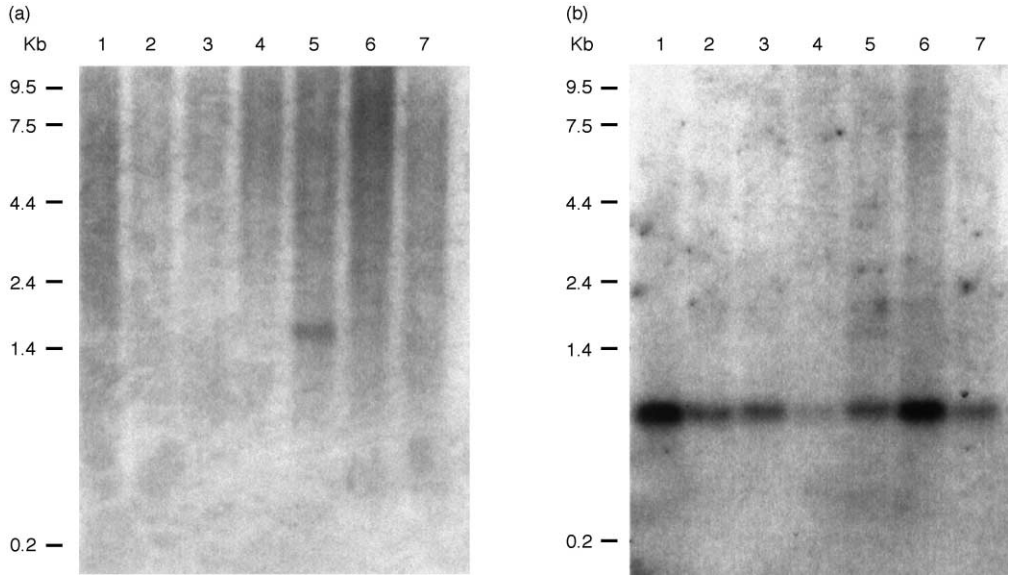

Fig. 3. Northern analysis of clone BKR1 (panel a) and clone BKR12 (panel b). Lane 1, stomach; lane 2, thyroid; lane 3, spinal cord; lane 4, lymph node; lane 5, trachea; lane 6, adrenal gland; lane 7, bone marrow.

Diatchenko and colleagues [7] have developed a PCR-based cDNA subtraction technique (SSH) that has some advantages compared with other subtraction methods. (a) It equalises transcript abundance among the target cDNA population, which allows the detection of rare transcripts. (b) Only one round of subtraction is needed, and (c) due to the suppression effect during the PCR, physical separation of the unhybridized cDNAs (differentially expressed transcripts) is not necessary. Diatchenko and colleagues [7] used this method to subtract human testis cDNAs against a mixture of cDNA derived from 10 different human tissues and achieved a high level of enrichment of testis-specific cDNAs. They used the subtracted cDNA mixture to screen a cosmid library constructed from flow-sorted human Y chromosomes. By a subsequent Northern analysis with 37 selected cosmids, five different chromosome Y-specific genes expressed in testis were identified. The application of this method [16] also resulted in the identification of 27 genes differentially expressed in an oestrogen receptor-positive breast carcinoma cell line and an oestrogen receptor-negative breast carcinoma cell line.

Here, we have applied this subtraction approach to two non-overlapping YACs that covered a region of approx. 2.3 Mb. The combined forward and reversed subtractions resulted in the identification of 21 unique clones. Of these, 12 appeared to be of human origin, as they hybridised to human DNA and to the tester YAC. The remaining nine clones hybridised to all yeast DNA samples present on the membrane, but not to human DNA. Thus, these clones presumably represent yeast cDNAs. This conclusion was confirmed by sequence analysis of some of these clones (data not shown), that indicated a high degree of homology with yeast sequences. The detection of yeast cDNAs may be caused by the differential expression of some yeast genes between the two cultures or may be intrinsic to the method that only guarantees enrichment of differentially expressed genes. The efficiency of the subtraction was indicated by the fact that transcripts of the yeast housekeeping gene *ZWF1* could no longer be detected in the subtracted cDNA pool.

The strategy used in this work resulted in the identification of 12 clones that map to the human DNA segment covered by the YACs 776A5 and 785D2. The fact that the whole procedure, from mRNA isolation until differential screening of the dot-blots, can be done in only 12 days, clearly demonstrates the strength of this method to identify novel transcripts in Mb-size sections of the human genome.

All 12 clones could be mapped to an approximately 2.5 Mb segment of 6q21 in the April 2001 release of the UCSC Human Genome Project Working Draft; six of them at very short distances from, or even from within the genomic region of, spliced expressed sequence tagged site (ESTs) and/or human mRNAs. Our failure to detect expression for seven clones in our preliminary Northern analysis could be caused by an overall low expression in human tissues in combination with the small size of the clones. Alternatively, some of the clones may consist of intronic sequences that were not spliced out during the mRNA maturation process. Splicing of foreign transcripts in yeast may not be efficient [17], as human sequences do not contain some of the consensus sequences needed by the splicing machinery in yeast to properly excise the introns [14,18]. Clone BKR3 of the forward subtraction appears to map within an intron of the nuclear receptor subfamily 2, group E, member 1 gene (*NR2E1*). Novel transcripts were detected for two of the clones. One of them is highly expressed in stomach and is being further analysed with respect to its putative role in gastric cancer development.

Acknowledgements

We thank Dr Anke van den Berg for carefully reading the manuscript. This work was supported by PRAXIS XXI (Project PRAXIS/PSAU/C/SAU/115/96).

References

1. Carvalho B, Seruca R, Carneiro F, Buys CHCM, Kok K. Substantial reduction of the gastric carcinoma critical region at 6q16.3-q23.1. *Genes Chromosomes Cancer* 1999, **26**, 29–34.
2. Parrish JE, Nelson DL. Methods for finding genes. A major rate-limiting step in positional cloning. *Genet Anal Tech Appl* 1993, **10**, 29–41.
3. Sagerström CG, Sun BI, Sive HL. Subtractive cloning: past, present, and future. *Annu Rev Biochem* 1997, **66**, 751–783.
4. Liang P, Pardee AB. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 1992, **257**, 967–971.
5. Lisitsyn N, Lisitsyn N, Wigler M. Cloning the differences between two complex genomes. *Science* 1993, **259**, 946–951.
6. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW. *Serial analysis of gene expression* 1995, **270**, 484–487.
7. Diatchenko L, Lau Y-FC, Campbell AP, *et al*. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc Natl Acad Sci USA* 1996, **93**, 6025–6030.
8. Still IH, Vince P, Cowell JK. Direct isolation of human transcribed sequences from yeast artificial chromosomes through the application of RNA fingerprinting. *Proc Natl Acad Sci USA* 1997, **94**, 10373–10378.
9. Diatchenko L, Lukyanov S, Lau Y-FC, Siebert PD. Suppression subtractive hybridization: a versatile method for identifying differentially expressed genes. *Methods Enzymol* 1999, **303**, 349–380.
10. Müllenbach R, Lagoda PJ, Welter C. An efficient salt-chloroform extraction of DNA from blood and tissues. *Trends Genet* 1989, **5**, 391.
11. Elvin P, Slynn G, Black D, *et al*. Isolation of cDNA clones using yeast artificial chromosome probes. *Nucleic Acids Res* 1990, **18**, 3913–3917.
12. Van der Steege G, Draaijers TG, Grootscholten PM, *et al*. A provisional transcript map of the spinal muscular atrophy (SMA) critical region. *Eur J Hum Genet* 1995, **3**, 87–95.
13. Timmer T, Terpstra P, van den Berg A, *et al*. A comparison of genomic structures and expression patterns of two closely related

flanking genes in a critical lung cancer region at 3p21.3. *Eur J Hum Genet* 1999, **7**, 478–486.
14. Trachtulec Z, Forejt J. Transcription and RNA processing of mammalian genes in Saccharomyces cerevisiae. *Nucleic Acids Res* 1999, **27**, 526–531.
15. Kunze B, Hellwig-Bürgel T, Weichenhan D, Traut W. Transcription and proper splicing of a mammalian gene in yeast. *Gene* 2000, **246**, 93–102.
16. Kuang WW, Thompson DA, Hoch RV, Weigel RJ. Differential screening and suppression subtractive hybridization identified genes differentially expressed in an estrogen receptor-positive breast carcinoma cell line. *Nucleic Acids Res* 1998, **26**, 1116–1123.
17. Langford CJ, Nellen W, Niessing J, Gallwitz D. Yeast is unable to excise foreign intervening sequences from hybrid gene transcripts. *Proc Natl Acad Sci USA* 1983, **80**, 1496–1500.
18. Langford CJ, Gallwitz D. Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. *Cell* 1983, **33**, 519–527.