Rapid communication

Identification of a Crohn's disease specific transcript with potential as a diagnostic marker

D A Lafontaine, S Mercure, J-P Perreault

Abstract

***Background*—A long time goal of the medical research community has been the identification of a reliable and valid marker for Crohn's disease.**

***Aim*—To identify differences in the genetic expression patterns of healthy and diseased tissues.**

***Method*—The RNA arbitrarily primed polymerase chain reaction (RAP-PCR) procedure was modified to improve its potential to identify clinical markers in heterogeneous RNA populations.**

***Results*—With this procedure, a 1065 bp PCR product associated with the inflammation that occurs in Crohn's disease was identified, cloned and sequenced. Northern blot hybridisations showed that this novel sequence originates from a unique RNA species of 3.1 kb. Dot blot hybridisations clearly showed that this RNA species was specific to Crohn's disease. Moreover, its abundance seemed to correlate with the severity of inflammation. Finally, this RNA species was also detected in macroscopically normal areas from Crohn's disease specimens, suggesting that it appears either early during the disease or at least before severe manifestations.**

***Conclusion*—This finding of a 3.1 kb RNA species permits the discrimination of Crohn's disease manifestations. Although further clinical work is required, this transcript appears to have definite potential as a diagnostic marker.**

(*Gut* 1998;**42**:878–882)

Keywords: Crohn's disease; inflammatory bowel disease; molecular marker; RNA arbitrarily primed PCR[1]

Département de Biochimie, Université de Sherbrooke, Sherbrooke, Québec J1H 5N4, Canada
D A Lafontaine
S Mercure
J-P Perreault

Correspondence to:
Dr Perreault.

Accepted for publication
20 February 1998

Crohn's disease is a chronic inflammatory bowel disease (IBD) that can develop anywhere along the gastrointestinal tract. Its causes are unknown; however, it is believed to be due to a combination of factors involving diet, genetic background, immunological responses, and the environment.[1–3] For example, it has been proposed that an immunological event, triggered by a conventionally presented antigen, in individuals with a particular genetic background may be involved in the lack of correct repression of the inflammation.[4 5]

In tissues affected by IBDs such as Crohn's disease and ulcerative colitis, several genes have been shown to be differentially expressed—for example: (*a*) the expression of mRNA for interleukin (IL)-1 and IL-1 receptor antagonist differs in colonic biopsy specimens from patients with IBD and inflammatory controls from patients with acute colitis[6]; (*b*) IL-12 is expressed and actively released by intestinal lamina propria mononuclear cells from Crohn's disease specimens in contrast with ulcerative colitis specimens[7]; (*c*) tumour necrosis factor α, IL-1β, and IL-6 have all been shown to be overexpressed in the inflamed areas of Crohn's disease specimens as compared with both normal areas and other controls.[8] These genes are all related to the immunological response leading to inflammation. However, neither the presence of these mRNAs nor that of their respective proteins appear to be useful as specific markers for unequivocally distinguishing Crohn's disease affected tissues from other intestinal diseases.

An urgent requirement for Crohn's disease diagnosis is the development of a clinical molecular marker that positively discriminates this pathology from others affecting the bowel. Within this context, the primary aim has been to identify differences in the genetic expression patterns of healthy and diseased tissues. To date, RNA fingerprinting—for example, differential display and RNA arbitrarily primed-polymerase chain reaction (RAP-PCR)—using reverse transcriptase coupled to the PCR appears to be the most promising approach for the identification of such molecular markers.[9 10] We have refined the RAP-PCR procedure to help to identify clinical markers in heterogeneous RNA populations—for example, surgically resected specimens from either different individuals or different pathologies. Using this procedure, we report the finding of a cellular transcript that positively identifies Crohn's disease affected tissues; the procedure therefore has potential as a diagnostic tool for Crohn's disease.

Materials and methods

PATIENTS, TISSUES, AND RNA EXTRACTION

This study was approved by the ethical review board of the Université de Sherbrooke, Centre

The sequence will be deposited in Genebank.

de Recherche Clinique. Forty one tissue specimens from intestinal surgical resections on 37 patients with active Crohn's disease, ulcerative colitis, diverticulitis, familial polyposis, lymphoma, or adenocarcinoma were obtained. These samples were washed with saline solution, and then small sections from severely inflamed, less inflamed, or nearby macroscopically uninflamed areas were isolated and rapidly snap frozen in liquid nitrogen before storage at −70°C. If multiple specimens were taken from one patient, they were taken from areas with different degrees of inflammation. Sections from the same areas were fixed for both histopathological analysis and confirmation of the clinical diagnosis. Total RNA was extracted from the frozen tissues using guanidinium isothiocyanate (Trisol Reagent; Life Technologies, Gibco-BRL, Gaithersburg, Maryland, USA)[11], quantified by UV spectroscopy, and its quality verified by 1% agarose gel electrophoresis. The RNA samples were normalised by equating the amounts of rRNA present. Degraded RNA samples, probably the result of extensive tissue necrosis, were rejected.

RAP-PCR

RAP-PCR was performed basically as described previously[9 10 12 13] on RNA populations isolated from different areas of surgically resected specimens. These specimens were obtained from patients with Crohn's disease (two specimens from inflamed, one from less inflamed, and one from non-inflamed areas). Similarly, samples from inflamed areas of one patient with ulcerative colitis, one with familial polyposis and two with diverticulitis were chosen as controls of inflammation, and two samples from altered areas of adenocarcinoma were chosen as controls of proliferation. Before the reverse transcriptase reaction, 1 μg RNA was incubated in a volume of 10 μl for 10 minutes at room temperature with DNase I as recommended by the manufacturer (amplification grade; Life Technologies). From the resulting RNA samples, 0.250 μg was used as templates for reverse transcriptase reactions using random hexamer primers and Moloney murine leukaemia virus reverse transcriptase enzyme following the manufacturer's recommended protocol (Superscript II RNase H^-; Life Technologies). The random hexamer primers allowed all (or most) of the RNA molecules to be represented in the resulting cDNA pool, and offered the advantage that a single cDNA pool was sufficient to perform all subsequent steps regardless of the arbitrary primers used in the PCR assays. The primers (10 μl; 6.6 μM) were annealed to the RNA by incubation at 70°C for 10 minutes followed by five minutes on ice. The quantities of the various cDNA pools were normalised by specific amplification of a 497 bp fragment from glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA using 3 μl of the cDNA samples as template and 1 μM both sense (dGTGAAGGTCGGTGTCAACGG corresponding to positions 1462–1481 and 70–89 of the gene and cDNA respectively—exon 1) and antisense (dGTCATCCATGACAACTTTGG corresponding to positions 3974–3993 and 547–566 of the gene and cDNA respectively— exon 6) primers (data not shown). Preliminary experiments allowed us to establish the linear portion of the amplification curve under the conditions used (PCR mixtures contained 10 mM Tris/HCl, pH 8.3, 2.0 mM $MgCl_2$, 50 mM KCl, 0.2 mM dGTP, 0.2 mM dCTP, 0.2 mM dTTP, 0.2 mM dATP (or 1 μM dATP and 20 μCi [α-^{35}S]dATP) and 1.5 units *Taq* DNA polymerase (Pharmacia Biotech, Uppsala, Sweden) in a final volume of 100 μl. We performed 32 cycles of one minute at 94°C, one minute at 55°C, and one minute at 72°C, with an extension time of five seconds/cycle). After this amplification, the volume of the various cDNA pools required to give equivalent amounts of the 497 bp fragment were determined and used in all subsequent amplifications. We tested more than 20 arbitrarily chosen oligonucleotides, which served as both sense and antisense primers for PCR. A normalised aliquot from the cDNA pool was added to the PCR mixture in a final volume of 100 μl containing 10 mM Tris/HCl, pH 8.3, 2.0 mM $MgCl_2$, 50 mM KCl, 0.2 mM dGTP, 0.2 mM dCTP, 0.2 mM dTTP, 1 μM dATP, 1 μM primer, and 1.5 units *Taq* DNA polymerase. Two initial cycles of selection (successive incubations of one minute at 94°C, two minutes at 37°C, slow increase of 2°C/minute up to 72°C, two minutes at 72°C) were performed. Subsequently, 20 μCi [α-^{35}S]dATP was added to the mixture and 28 amplification cycles (one minute at 94°C, two minutes at 60°C, and two minutes at 72°C) were performed. After amplification, the mixtures were extracted with phenol/chloroform, precipitated with propan-2-ol, washed with ethanol, and dried. The resulting pellets were resuspended in 10 μl loading buffer (0.3% each of bromophenol blue and xylene cyanol, 10 mM EDTA, pH 7.5, 97.5% deionised formamide), and a 3 μl aliquot was then heat denatured at 90°C for one minute, snap cooled on ice for five minutes, and electrophoresed on a 5% Long Ranger (JT Baker, Phillipsburg, New Jersey, USA) sequencing gel (4 mm × 20 cm × 60 cm). The dried gels were quantified using a PhosphorImager (Molecular Dynamics). Alternatively, the RAP-PCR products were analysed by 1% agarose gel electrophoresis.

PCR PRODUCT CLONING AND SEQUENCING

One PCR product (1065 bp) appeared to be specific to the inflamed Crohn's disease specimens (see the Results section). The gel slice containing this band was isolated and the DNA extracted, precipitated with ethanol, washed with ethanol, and finally dried. This DNA fragment was cloned by taking advantage of the property of *Taq* DNA polymerase to add an adenosine at the 3' ends of the PCR products. Consequently, the PCR amplified fragment was ligated in a "sticky end" fashion to a linearised vector pCRII possessing an extra thymidine residue at each 5' end, as recommended by the manufacturer (TA cloning kit; Invitrogen, Carlsbad, California, USA). The resulting clone (pCD1) was sequenced in both

directions by the dideoxyribonucleotide chain termination method using the T7 sequencing kit (Pharmacia Biotech).

NORTHERN AND DOT BLOT HYBRIDISATIONS

Northern and dot blot hybridisations were performed as described previously.[14] All RNA samples were extracted, treated, and normalised as described previously. To eliminate RNA loading discrepancies, all membranes were stained with methylene blue (Molecular Research Center, Cincinnati, Ohio, USA), which permits visualisation of the rRNA. Northern and dot blot hybridisation probes were prepared from the pCD1 clone. For the northern blot hybridisations, both sense and antisense riboprobes were prepared by in vitro transcription of linearised vector using either Sp6 or T7 RNA polymerase in the presence of [α-^{32}P]UTP (Amersham Life Science, Arlington Heights, Illinois, USA). RNA samples (20 μg) isolated from different resected tissues were denatured in the presence of glyoxal/dimethyl sulphoxide, fractionated by 1% agarose gel electrophoresis, and transferred overnight to nylon filters (Hybond N$^+$; Amersham Life Science). The filters were prehybridised at 65°C for two hours and hybridised at 65°C overnight (about 16 hours) as described previously.[14] After hybridisation, the filters were successively washed twice in 2 × SSC (20 × SSC is 3 M NaCl/0.3 M sodium citrate/NaOH, pH 7.0) for five minutes at room temperature, once in 2 × SSC/1% sodium dodecyl sulphate (SDS) for 30 minutes at 65°C, once in 0.1 × SSC/0.5% SDS for 30 minutes at 65°C, and finally in 0.1 × SSC for 30 minutes at room temperature. The filters were then analysed by either autoradiography or PhosphorImager. For the dot blot hybridisations, the probe consisted of the entire 1065 bp insert which could be isolated from pCD1 by taking advantage of the two *Eco*RI sites flanking the cloning site in pCRII. The gel-purified DNA insert was then labelled using the multiprime DNA labelling system in the presence of [α-^{32}P]dCTP (Amersham Life Science). The hybridisation procedure was similar to the northern blot analysis, except that RNA samples (10 μg) were applied to the filter under vacuum. Several hybridisation membranes were probed either independently or simultaneously with a GAPDH probe consisting of the 497 bp PCR product (see earlier) labelled using the multiprime DNA labelling system in the presence of [α-^{32}P]dCTP. This procedure confirmed that all RNA samples were equivalent with respect to amount.

Results

RNA fingerprinting methods—for example, differential display and RAP-PCR—are semiquantitative and can be used to analyse RNA populations for differentially regulated genes[9 10] Both of these methods have been used successfully on many occasions, primarily for RNA isolated from relatively homogeneous populations such as bacterial or cell cultures.[12 13] However, several groups working on the characterisation of gene expression associated with clinically important pathologies have been unable to identify molecular markers using current RNA fingerprint methods. Two reasons for these failures are the high heterogeneity of the RNA samples obtained when working with tissues, and the bias introduced by using a minimum number of specimens—for example, only one diseased and one healthy sample. To circumvent the problem of RNA heterogeneity, we added a cDNA pool normalisation step as well as several control points to ensure both the quality and quantity of the samples. Ten RNA samples from patients with ulcerative colitis, diverticulitis, familial polyposis, adenocarcinoma, or Crohn's disease were selected in order to remove any marker associated with inflammation, proliferation, or a particular individual (see under specimens and methods). On assay of more than 20 primers, an 18-mer oligonucleotide (sequence: dGCTGTTTCCTTCCCCGTC; melting temperature 51.7°C) was observed to produce patterns that included a PCR product of 1065 bp that was specific to severely as well as less inflamed Crohn's disease specimens. This product was analysed in greater detail.

The 1065 bp PCR product was gel extracted and cloned into the pCRII vector. The resulting construct was named pCD1, and its 1065 bp insert DNA was sequenced in both directions (fig 1A). A BLAST search[15] of nucleic acid databases did not reveal any significant homologies. The sequence includes several potential open reading frames. However, because it does not correspond to the complete RNA molecule (see below), the identification of a putative protein would be premature. Northern blot hybridisation using either sense or antisense radiolabelled riboprobes synthesised from pCD1 was performed with RNA isolated from different patients from those used for RAP-PCR. Only one of the probes gave a signal, thereby permitting the identification of the polarity of the natural RNA species. A single band corresponding to an mRNA of about 3.1 kb was detected exclusively in lanes containing RNA populations isolated from Crohn's disease inflamed areas (fig 1B, lanes 2 and 3). No hybridisation signal was detected in samples from either a normal area of a Crohn's disease specimen or an inflamed area of an adenocarcinoma (fig 1B, lanes 1 and 4), even though the amounts of both the rRNA and GAPDH mRNA were shown to be virtually the same by methylene blue staining and GAPDH probing of the membrane respectively (data not shown). These results support the conclusion that the RAP-PCR 1065 bp product is unique to inflamed tissues from patients with Crohn's disease.

The detection of a single RNA species using this 1065 bp fragment as a probe suggested the possibility of performing dot blot hybridisations. This method allows the screening of a greater number of specimens,[13] and therefore should provide an excellent preliminary evaluation of the discriminatory potential of this RNA species for Crohn's disease. An example of a dot blot is presented in fig 1C. Clearly, a

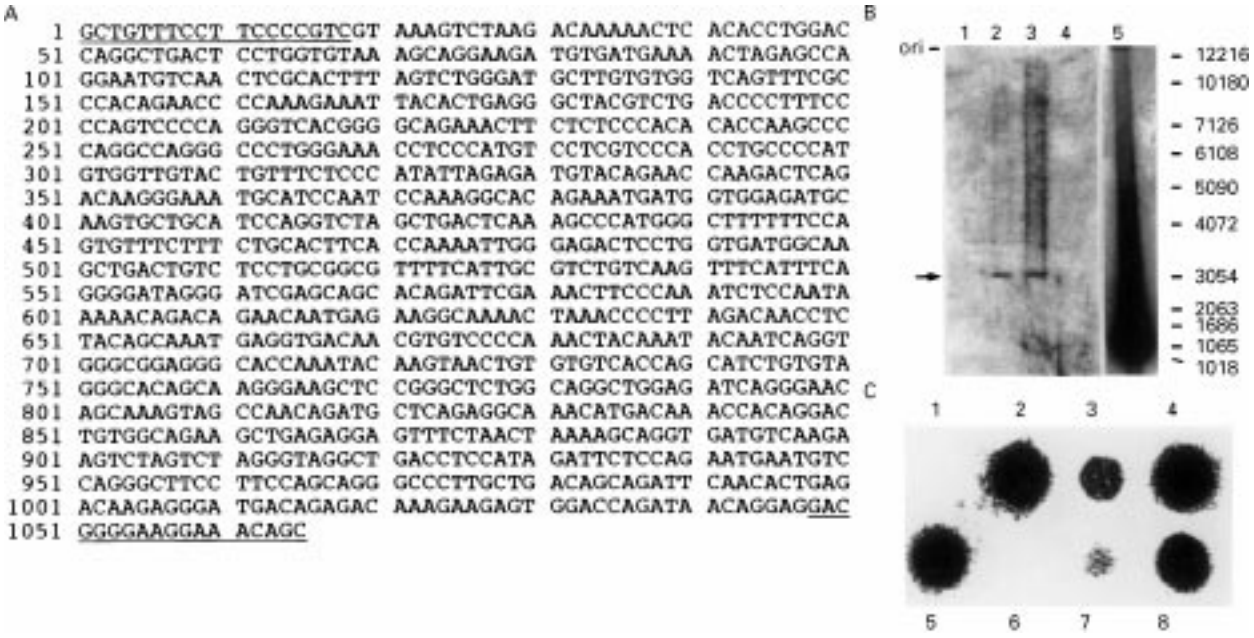

Figure 1 Characterisation of the 1065 bp fragment. (A) Complete nucleotide sequence of the 1065 bp PCR product. The underlined nucleotides at both ends correspond to the PCR primer used. (B) Northern blot hybridisation. Lane 1 is a sample from a non-inflamed area of a specimen from a patient with Crohn's disease, lanes 2 and 3 are Crohn's disease samples from inflamed areas from two different patients, lane 4 is a sample from an adenocarcinoma, and lane 5 is 0.1 ng of a pCD1 insert. The arrow indicates the 3.1 kb RNA species. The positions of single stranded DNA molecular mass (Da) markers are indicated on the right. (C) Example of a dot blot hybridisation assay. Sample 1 is an inflamed ulcerative colitis specimen; samples 2–4 are Crohn's disease specimens, sample 5 is the non-radioactive probe (control), sample 6 is from an individual with an adenocarcinoma, sample 7 is from an inflamed area of a patient with Crohn's disease, but showing signs of remission, and sample 8 is from an inflamed area from an individual with an indeterminate IBD.

positive signal is detected only with samples from the inflamed areas of Crohn's disease tissues (samples 2–4), and not in control samples (samples 1 and 6). Furthermore, it seems that the abundance of the RNA species in question correlates with the severity of the inflammation; the more severe the inflammation, the stronger the signal. For example, samples 2 and 4 were from severely inflamed areas of patients with Crohn's disease, while sample 3, which has a less intense signal, was from a less severely affected patient. In addition, a Crohn's disease case in which remission was observed produced a barely visible signal (sample 7). These results suggest that the 3.1 kb mRNA species is closely related to the severity of the inflammation in Crohn's disease. Finally, a clinically and pathologically indeterminate IBD sample was analysed (sample 8). The 3.1 kb RNA was detected, suggesting that the patient is suffering from Crohn's disease. This result shows the potential of the 3.1 kb RNA species as a diagnostic marker for Crohn's disease.

Table 1 Detection of the 3.1 kb RNA species in various specimens

Diagnosis	*Macroscopic aspect of the specimens*	*No tested*	*No in which 3.1 kb RNA detected*
Crohn's disease	++	9	9
	+	5	5
	−	4	1
Ulcerative colitis	++	4	0
Diverticulitis	++	3	0
Familial polyposis	++	2	0
Adenocarcinoma	++	4	0
	−	1	0
Lymphoma	++	4	0
	−	1	0
Indeterminate IBD	++	4	1

Macroscopic aspect of the specimens: ++, severely altered; +, altered; −, not altered.

Results of the dot blot hybridisation, as well as of other similar experiments to detect the 3.1 kb RNA species, are summarised in table 1. Briefly, all 13 Crohn's disease specimens were positive, while all 17 control samples were negative, for the presence of this RNA molecule. Four samples isolated from the macroscopically normal areas of Crohn's disease tissues were analysed. In one of these cases, the 3.1 kb transcript was detected in the normal area. Finally, one of four undetermined cases of IBD was shown to be positive for the presence of this RNA species. Thus, from the 41 samples representing different pathological states, 15 samples compatible with a Crohn's disease diagnosis were shown to be positive for the presence of the 3.1 kb RNA species. This RNA species was not detected in samples isolated from several parts of the gastrointestinal tract, as well as other parts of the human body, from patients suffering from either IBD or other unrelated diseases (data not shown). The fact that the RNA species was not detected in these samples suggests that either it is not present in non-Crohn's disease affected tissues or that it is present at a level too low to be detected by hybridisation techniques.

Discussion

A long time goal of the medical research community has been the development of a clinical molecular marker for Crohn's disease. The primary focus of this study was to identify differences between genetic expression patterns of Crohn's disease and other intestinal diseases.

Using a refined RAP-PCR procedure useful for studying resected samples—for example, inflamed tissues—we detected a 1065 bp PCR product associated with the Crohn's disease manifestations present in surgical specimens. The DNA sequence of this PCR product showed no significant homology with any of those present in the various nucleic acid databases. However, preliminary PCR amplifications of DNA isolated from different individuals indicate that this 1065 bp fragment is present in genomic DNA (data not shown). In contrast, northern blot hybridisations show that this PCR product originates from a unique RNA species of 3.1 kb detected only in those patients suffering from Crohn's disease.

Dot blot hybridisation showed that the 3.1 kb RNA species permits the discrimination of Crohn's disease alterations from those of other intestinal diseases, since this RNA species was detectable in all Crohn's disease specimens from inflamed areas but absent in all controls. Moreover, it was also detected in macroscopically normal areas from a Crohn's disease specimen, although in lower amount. This latter result suggests that this RNA molecule appears either early during the disease or at least before severe manifestations. This 3.1 kb species could be an interesting molecular marker, since it seems to appear relatively early in Crohn's disease development. However, it should be kept in mind that the number of specimens examined is relatively small and insufficient for significant statistical analysis. Furthermore, samples from patients living within a defined geographic area were used. Therefore additional worldwide trials are required to validate the use of this RNA species as a molecular marker of Crohn's disease. Furthermore, it would be interesting to verify the expression of the corresponding gene when Crohn's disease occurs in other parts of the gastrointestinal tract.

In summary, we have identified a Crohn's disease specific transcript which possesses potential as a diagnostic marker. To allow further development of this potential diagnostic tool, the gene as well as the transcript and any putative expression product must be characterised. Further clinical work is required to validate the worldwide use of this potential marker, which may also provide a new avenue for research into Crohn's disease.

We thank Dr F Beuvon, Dr D B Ménard, and Dr J Poisson for their suggestions and for providing the surgical specimens used in this study. This research was supported by a grant from MRC (Canada) to J-P P. D L was sponsored by a graduate fellowship from the FRSQ (Québec). J-P P is an MRC scholar.

1 Podolsky DK. Inflammatory bowel disease. *N Engl J Med* 1991;**325**:928–37.
2 Thayer WR. The aetiology of inflammatory bowel disease: the role of infectious agents. In: Allan RN, Keighley MRB, Hawkins C, eds. *Inflammatory bowel diseases.* New York: Churchill Livingstone, 1990:147–63.
3 Mishina D, Katsel P, Brown ST, *et al.* On the etiology of Crohn's disease. *Proc Natl Acad Sci USA* 1996;**93**:9816–20.
4 Fewell DP, Snook FA. Immunology of ulcerative colitis and Crohn's disease. In: Allan RN, Keighley MRB, Hawkins C, eds. *Inflammatory bowel diseases.* New York: Churchill Livingstone, 1990:127–46.
5 Strober W, Ehrhardt RO. Chronic intestinal inflammation: an unexpected outcome in cytokine or T cell receptor mutant mice. *Cell* 1993;**75**:203–5.
6 Isaacs KL, Sartor RB, Haskill S. Cytokine messenger, RNA profiles in inflammatory bowel disease mucosa detected by polymerase chain reaction amplification. *Gastroenterology* 1992;**103**:1587–95.
7 Monteleone G, Biancone L, Marasco R, *et al.* Interleukin 12 is expressed and actively released by Crohn's disease intestinal lamina propria mononuclear cells. *Gastroenterology* 1997;**112**:1169–78.
8 Reimund J-M, Wittersheim C, Dumont S, *et al.* Increased production of tumour necrosis factor-α, interleukin-1β, and interleukin-6, by morphologically normal intestinal biopsies from patients with Crohn's disease. *Gut* 1996;**39**: 684–9.
9 Liang P, Pardee AB. Differential display of eukaryotic messager RNA by means of the polymerase chain reaction. *Science* 1992;**257**:967–71.
10 McClelland M, Mathieu-Daude F, Welsh J. RNA fingerprinting and differential display using arbitrarily primed PCR. *Trends Genet* 1995;**11**:242–6.
11 Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidium-isothiocyanate-phenol-chloroform extraction. *Anal Biochem* 1987;**162**:156–9.
12 Dalal SS, Welsh J, Tkachenko A, *et al.* Rapid isolation of tissue-specific and developmentally regulated brain cDNAs using RNA arbitrarily primed PCR (RAP-PCR). *J Mol Neurosci* 1994;**5**:93–104.
13 Sunday ME. Differential display RT-PCR for identifying novel gene expression in the lung. *Am J Physiol* 1995;**269**: L273–84.
14 Sambrook J, Fritsch EF, Maniatis T. Analysis of RNA. In: *Molecular cloning: A laboratory manual.* 2nd edn. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 1989:7.37–7.52.
15 Altschul SF, Gish W, Miller W, *et al.* Basic local alignment search tool. *J Mol Biol* 1990;**215**:403–10.