Pergamon

European Journal of Cancer Vol. 30A, No. 9 pp. 1391–1394, 1994

0959–8049/94 $7.00+0.00

0959-8049(94)00277-0

News

European Pancreatic Cancer Reference Library System, EPCRLS

T.M. Gress, F. Müller-Pillasch, G. Adler, G. Zehetner and H. Lehrach

INTRODUCTION

THE AIM of the European Pancreatic Cancer Reference Library System (EPCRLS) is to provide resources offering the possibility to correlate and coordinate research on molecular biology of pancreatic cancer. Cancerogenesis, in general, is thought to be a complex multistep process in which genetic alterations accumulate to bring about the neoplastic phenotype. Early studies using mRNA/cDNA saturation–hybridisation techniques showed that the complexity of the mRNA of wild type and transformed cells may differ by up to 10% [1, 2]. Thus, studies of gene expression in cancer appear to be a suitable tool for a global evaluation of genetic alterations leading to the malignant phenotype of cancer cells. Conventional studies of gene expression are based on techniques allowing the assessment of only one or a few genes involved. In pancreatic cancer, genetic alterations of growth factors, growth factor receptors and oncogenes have been described and reviewed elsewhere [3, 4]. A much larger number of genes can be expected to be involved in primary and secondary processes leading to the malignant phenotype. The EPCRLS was designed to accelerate the understanding of the molecular biology of pancreatic cancer, by offering the resources and the technical means required to combine data obtained by use of different experimental approaches and by different groups involved in the research of the molecular biology of pancreatic cancer and of other tumours. The EPCRLS makes use of automated molecular techniques, originally devised in the context of the genome project [5] based on the use of cDNA libraries arrayed in uniform formats on filters suitable for hybridisation.

cDNA LIBRARIES

cDNA libraries were prepared from the pancreatic cancer cell lines PATU 8988s and t [6], from two pancreatic cancer tissues (reference no. 535 = ductual carcinoma G3, reference no. 603 = ductual carcinoma G2) and from two pancreatic control tissues (reference no. 503 and 343/1). Both pancreatic cancer cell lines were derived from the same tumour, the clone 8988t (G1 in nude mice xenografts) is highly differentiated whereas the clone 8988s is poorly differentiated (G2–3 in nude mice xenografts). Pancreatic tissues were provided by the Department of Surgery of the University of Ulm, with written consent of the patients, and after approval by the local Ethics Committee. Cancer tissues were obtained from patients undergoing surgery for pancreatic cancer, and control tissues were obtained from healthy individuals through an organ donor programme. RNA from cell lines and shock frozen human pancreatic tissue was extracted by use of a standard guanidinium thiocyanate extraction, followed by centrifugation in a caesium chloride gradient (Pharmacia LKB, Freiburg, Germany). cDNA libraries were synthesised using Oligo-dT primers and reverse transcriptase according to standard protocols [7]. cDNA was cloned into the plasmid vector pSPORT^R (Gibco BRL, Paisley, U.K.). Care was taken to achieve average insert sizes of >1.5 kb and initial complexities of $>1 \times 10^6$ clones. Libraries were neither amplified nor enriched to obtain a collection of cDNA clones representative for the mRNA of the cell line or tissue of interest.

AUTOMATED MOLECULAR TECHNIQUES

For an automated analysis, cDNA libraries have to be presented in a regular arrangement allowing computerised data analysis and processing. Techniques used for this purpose were devised in the context of the human genome project and have been described elsewhere in detail [8–11]. In brief, 20 736 clones of each cDNA library are randomly plated on agar plates and picked into individual wells of 386-well microtitre plates with a "picking robot". After an overnight incubation at 37°C, 20 736 cDNA clones of each library are spotted from microtitre plates on to 20 × 20 cm Nylon membranes with the help of a "spotting robot". These automated procedures allow the production of an unlimited number of identical *in-situ* membranes for each cDNA library. After an overnight incubation on agar plates, an *in situ* colony forms on the membrane for each spotted cDNA clone. After additional processing procedures (see [11] for details) *in situ* library membranes are ready to be used for hybridisations (Figure 1).

USES

Arrayed library filters are suitable for hybridisations with a wide range of different probes (some examples are given in Table 1). While filters made from amplified randomly plated libraries currently retain the advantage that clone densities are higher than for standard one-clone/one-spot arrays, all types of arrayed library filters eliminate the need for multiple rounds of clone purification and can, therefore, be used for standard gene

Correspondence to T.M. Gress.
T.M. Gress, F. Müller-Pillasch and G. Adler are at the Medizinische Klinik, Abt.I, Robert Koch Str. 8, 89081 Ulm, Germany; and G. Zehetner and H. Lehrach are at the Genome Analysis Laboratory, Imperial Cancer Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, U.K.
Received 13 May 1994; accepted 9 June 1994.

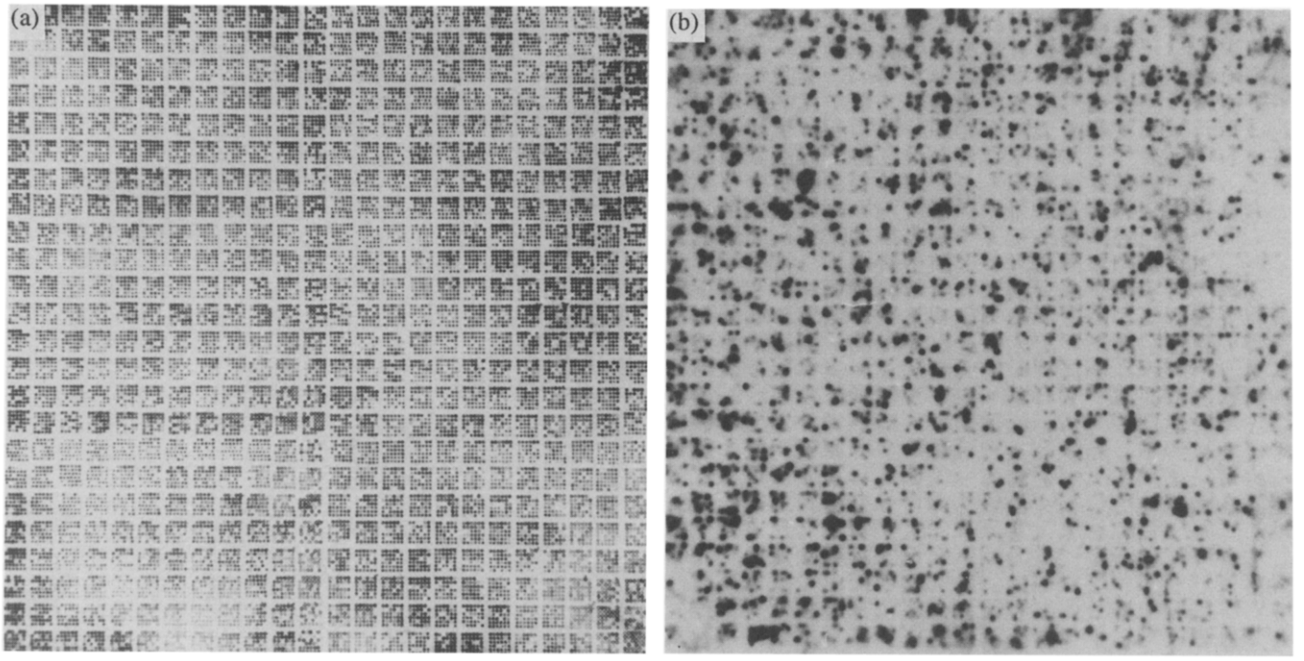

Figure 1. (a) A formatted cDNA library array of pancreatic cancer cell line hybridised with a vector probe is shown. This hybridisation allows the detection of all clones on the library and is required for all further normalisation procedures and as reference for the automated image analysis. (b) Same filter as shown in (a) hybridised with a cDNA pool probe prepared from human pancreatic tissue, allowing identification of clones expressed at middle to high abundance in the human pancreas.

Table 1. Uses of EPCRLS cDNA library arrays

Probes	Uses
Simple	
Single copy fragments	Isolation of homologous genes, determination of gene abundance
Unique oligonucleotides	Gene isolation, e.g. with amino acid consensus sequences
Complex	
Tissue cDNA pools	Identification of cDNAs with tissue-specific gene expression
Clone pools (cosmid, exon-trapp)	Identification of chromosome-specific cDNAs
Random oligonucleotides	Partial sequencing by hybridisation
Repetitive DNA	Identification of repetitive sequences to avoid spurious results
Sequence motifs	Identification of cDNAs sharing the same amino acid domain

isolation and identification protocols. Currently, the EPCRLS offers a total of six cDNA library filters from pancreatic tissues or cell lines, amounting to a total of 124 416 cDNA clones, which should suffice for most gene isolation purposes. For isolation of very rare genes, it may still be necessary to use conventional cDNA library screening techniques. Regular arrays of cDNA libraries display their main advantage when using complex probes hybridising with hundreds or thousands of library clones [12]. These uses take advantage of the ability to correlate the quantitative information generated by a series of different hybridsation probes for each cDNA clone in the library (see Table 1, for examples). In the interest of brevity, we will describe only a few hybridisations performed in our laboratories which will form part of the data available within the EPCRLS. To avoid spurious results, all cDNA library filters are hybridised with a number of control probes as, e.g. Poly-A homopolymers (identification of large Poly-A tails), repetitive DNA (identification of repetitive sequence elements) or vector DNA (normalisation for amount of DNA per spot). Hybridisations with labelled probes of DNA, complementary to the complete mRNA of a cell or tissue (=cDNA pool probe), allow the identification of clones containing sequences abundantly expressed in the tissue used to generate the probe [13]. Presently, we are using cDNA pool probes of pancreatic tissues (control, chronic pancreatitis, carcinoma) and cell lines. We are planning to include hybridisations with cDNA pools from other tissues. These experiments will produce data that can usually only be obtained by northern blot hybridisations. Hybridisations with cloned chromosome-specific DNA (e.g. from cosmids [14] or exon-trap probes [15]) will allow the isolation of cDNA sequences belonging to a specific chromosomal localisation.

RECOMMENDATIONS FOR THE USE OF cDNA LIBRARY ARRAYS

High-density cDNA library arrays are basically suitable for any hybridisation technique. In our hands, hybridisations in six × SSC, five × Denhardts solution, 0.5% SDS, 100 μg yeast t-RNA/ml, and 50% formamide at 42°C gave the best results. The hybridisation solution can be supplemented with 50 μg of sonicated total human placental DNA (Sigma, Poole, U.K.)/ml and/or 10 μg of polyU homopolymer/ml (Pharmacia), if required. For an optimal clone localisation and for automated data analysis, it is necessary to add ^{35}S-labelled vector probe to

each hybridisation with a ^{32}P-labelled probe to generate a vector grid as shown in Figure 1a. The ^{35}S and ^{32}P signal can easily be distinguished on the same autoradiograph by positioning, e.g. a sheet of polypropylen between filter and X-ray film during exposure. To obtain their clone of interest, the users should either send us the autoradiographies (with a clearly visible vector grid) or provide us with the coordinates of the clone (*x*/*y* coordinates starting from the top left corner of the library filter).

AUTOMATED DATA ANALYSIS AND PROCESSING

Automated techniques for data analysis and processing are required for an efficient evaluation of hybridisation results obtained with complex probes yielding thousands of positive clones (see Figure 1b). The regular array of cDNA clones on the library membranes allows the use of a computerised image analysis system, based on a CCD video camera mounted on an XY-robotic arm (developed in collaboration with Professor Frey, Dept of Industrial Image Analysis/Fachhochschule Ulm/ Germany), a frame grabber and a computer. Grey values obtained for each hybridising clone are initially stored in a local standard database, allowing relational queries and data processing (e.g. procedures to normalise for different amounts of DNA and differences in colony growth or identification of repetitive sequence elements, large Poly-A tails). Files containing processed data for each library clone will be transferred to the centralised computer database at the Imperial Cancer Research Fund (ICRF) in London.

CENTRALISED COMPUTER DATABASE

The data will form part of the newly developed Reference Library Database 2 (RLDB2), which is an improved version of the original RLDB, used for several years for the administration and data storage of the Reference Library System at the ICRF [5, 16]. RLDB2 has a strongly object-oriented scheme and is set up under the relational ORACLE Database Management system. An experimental server has been installed which allows on-line access to the RLDB2 over the internet by using the World Wide Web (WWW) browser Mosaic (a freely available programme developed by NCSA) as client programme. The RLDB2 information server can be accessed by pointing the WWW browser to the Universal Resource Locator (URL) http://gea.lif.icnet.uk/. This allows anybody who has access to a UNIX workstation, Macintosh or DOS/Windows computer, which is connected to the internet, to retrieve information either from the public part of the database or after identification with user number and password from the users own private data. For scientists without direct internet access, dumps of the relevant information from the database will be made available in text files via anonymous ftp or e-mail. Initially, the database will mainly contain basic data produced in our laboratories (e.g. control data, expression patterns). The amount of biological data will increase significantly over time after acceptance by a large scientific community.

CONCLUSIONS

The use of formatted cDNA library filters allows the pooling of all data produced by a variety of different techniques and laboratories, and to keep track of all information known on each cDNA clone. By making cDNA library resources from pancreatic cancer tissues and cell lines publicly available, with financial support from the European Community, all users will have the possibility to work with the same set of cDNA clones. The EPCRLS will make it easier to correlate data obtained by use of different approaches and by different groups. Additional data to each library will be obtained by the participation of groups from Eastern and Central Europe (G. Varga, Hungarian Academy of Sciences, Budapest, Hungary; T. Lushinkova and M. Kelve, Estonian Academy of Sciences, Tallinn, Estonia; L. Djondjurov, Bulgarian Academy of Sciences, Sofia, Bulgaria). These groups are supported by the European Community to contribute to the characterisation of cDNA clones (nucleic acid sequence, chromosomal localisation ...). The principle of the EPCRLS is summarised in Figure 2. The participation of a large scientific community and, in particular, the feedback of experimental data will be essential to ensure the largest possible success of the system. Other tumours may be included in the cDNA reference library system, if a wide acceptance is obtained.

1. Groudine M, Weintraub H. Activation of cellular genes by avian RNA tumor viruses. *Proc Natl Acad Sci USA* 1980, **77**, 5351–5354.
2. Moyzis R, Grady D, Li D, Mirvis S, Tso's P. Extensive homology of nuclear ribonuclic acid and polysomal poly(adenylic) messenger ribonucleic acid between normal and neoplastically transformed cells. *Biochemistry* 1980, **19**, 821–832.
3. Gillespie J, Poston GJ. The molecular biology of pancreatic adenocarcinoma. In Johnson CD, Imrie CW, eds. *Pancreatic Disease. Progress and Prospects*. London, Springer Verlag, 1991, 47–64.
4. Lemoine NR, Hall PA. Growth factors and oncogenes in pancreatic cancer. *Baillieres Clin Gastroenterol* 1990, **4**, 815–832.
5. Lehrach H, Drmanac R, Hoheisel J, *et al.* Hybridization fingerprinting in genome mapping and sequencing. In Davies K, Tilghman SM, eds. *Genome Analysis*, Vol. I: *Genetic and Physical Mapping*. Cold Spring Harbor, New York, Cold Spring Harbor, Laboratory Press, 1990, 39–81.

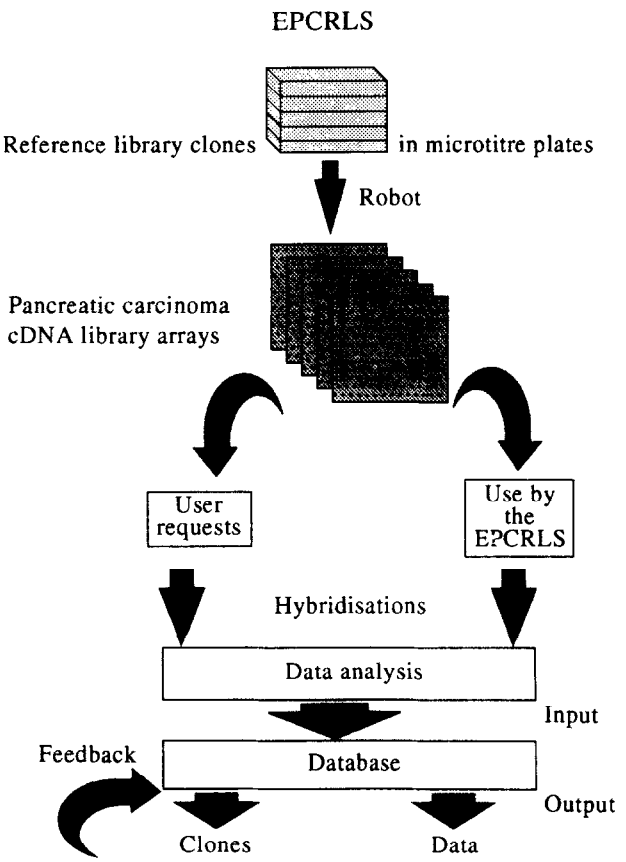

Figure 2. Flow chart outlining the basic principle of the EPCRLS.

6. Elsässer HP, Lehr U, Agricola B, Kern HF. Establishment and characterisation of two cell lines with different grade of differentiation derived from one primary human pancreatic adenocarcinoma. *Virchows Arch B Cell Pathol* 1992, **61**, 295–306.
7. Gubler U, Chua AO. Establishment of cDNA libraries in lambda-gt 10. In Brown TA, ed. *Essential Molecular Biology. A Practical Approach*, Vol. II. Oxford, IRL Press, 1991, 39–56.
8. Hoheisel JD, Lennon GG, Zehetner G, Lehrach H. Use of high-coverage reference libraries of Drosophila melanogaster for relational data analysis. A step towards mapping and sequencing the genome. *J Mol Biol* 1991, **220**, 903–914.
9. Hoheisel JD, Maier E, Mott R, *et al.* High resolution cosmid and P1 maps spanning the 14 Mb genome of the fission yeast *S. pombe*. *Cell* 1993, **73**, 109–120.
10. Maier E, Hoheisel JD, McCarthy L, *et al.* Yeast artificial chromosome clones completely spanning the genome of *Schizosaccharomyces pombe*. *Nature Genet* 1992, **1**, 273–277.
11. Nizetic D, Zehetner G, Monaco A, Gellen L, Young B, Lehrach H. 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 USA* 1991, **88**, 3233–3237.
12. Lennon GG, Lehrach H. Hybridization analyses of arrayed cDNA libraries. *TIG* 1991, **7**, 314–317.
13. Gress TM, Hoheisel JD, Lennon GG, Zehetner G, Lehrach H. Hybridization fingerprinting of high density cDNA-library arrays with cDNA pools derived from whole tissues. *Mammalian Genome* 1992, **3**, 609–619.
14. Riordan JR, Rommens JM, Kerem BS, *et al.* Identification of the cystic fibrosis gene: cloning and characterisation of complementary DNA. *Science* 1989, **245**, 1066–1073.
15. North MA, Sanseau P, Buckler AJ, *et al.* Efficiency and specificity of gene isolation by exon amplification. *Mammalian Genome* 1993, **4**, 466–474.
16. Zehetner G, Lehrach H. The reference library system–sharing biological material and experimental data. *Nature* 1994, **367**, 489–491.

Acknowledgements—We would like to thank U. Lacher, K. Monn and C. Gartmann-Gutjahr for excellent technical assistance and Prof. Frey (Fachrochschule Ulm/Germany) for his support concerning the image analysis of hybridsation results. This study was supported by the European Community BIOMED1-Programme Grant No. ERBBMH1-CT-92-0401, European Community PECO-Programme Grant No. ERBCIPDCT930109 and Deutsche Forschungsgemeinschaft (DFG) Grant No. Gr1010/3-1.