

cDNA Microarrays for Pharmacogenomic Analysis of Cystic Fibrosis

Meera Srivastava, Ofer Eidelman, and Harvey B. Pollard

1. Introduction

Cystic fibrosis (CF) is a single-gene disorder with a complex phenotype, in which multiple organs are affected. The pulmonary complications of CF, including mucous plugging and chronic bacterial infection of the lung, represent the major cause of morbidity and mortality (*1*). It has long been suggested that mutations in CFTR, localized to the apical membrane of airway epithelia, lead to abnormalities in the fluid lining the airway surface. Accumulating evidence suggests that the inflammatory response in the CF lung may be excessive. However, it is not clear to what extent this reflects unusual persistence of stimulation of this response by bacteria or other pathogens, vs to what extent defects in CFTR might by themselves stimulate or cause dysregulation of inflammatory responses. Changes in the multigene patterns of expression can provide clues about regulatory mechanisms and broader cellular functions and biochemical pathways. In the context of cystic fibrosis disease and treatment, the knowledge gained from these types of measurements can help determine the causes and consequences of disease, how drugs and drug candidates work in cells and organisms, and what gene products might have therapeutic uses themselves or may be appropriate targets for therapeutic intervention.

For our own studies, we have concentrated on HEK-293 cells expressing recombinant wild-type or $\Delta F508$ mutant CFTR (*2*), and the naturally occurring CFTR mutant IB3 epithelial cells, repaired with wild-type CFTR using an AAV vector (*3*). We describe here how the cells and cellular processes are studied using DNA arrays that allow complex mixtures of RNA and DNA to be interrogated in a parallel and quantitative fashion. Unlike the recombinant HEK-

293 system, IB3 cells mirror the high IL8 secretion phenotype typical of the CF airway. Therefore, we have tended in recent times to concentrate in the latter system.

Our own approach has been to appreciate that not every site in the airway is identical physiologically to others, and to take these differences into account when devising and interpreting experiments. For example, CF epithelial cells high in the airway, such as nasal epithelia, secrete much less pro-inflammatory IL8 than those from lower in the airway. Furthermore, there are instances in which immortalized CF airway cells have lost their CF dependent pro-inflammatory phenotype of high IL8 secretion. For example, there are lines of IB3 cells that have lost their high IL8 secretion phenotype. In this case, one may want either to avoid these cells or use them in a knowledgeable manner.

The basic strategy is to grow cells of interest, treat them as dictated by the experiment, isolate messenger RNA, and analyze by cDNA or oligonucleotide array methods. Other approaches are available according to the taste and pocketbook of the investigator. For the academic investigator, cDNA arrays are less expensive than oligonucleotide arrays. These arrays also have the advantage of 25+ years of experience with dot-blot Northern assays.

2. Materials

2.1. Growth of IB3 Cells

1. Serum-free LHC-8 medium.

2.2. For DNase 1 Treatment of Total RNA

1. RNase-free DNase I (Boehringer-Mannheim #776-785, 10 U/ μ L): dilute to 1 U/ μ L in 1 \times DNase I buffer before to use.
2. 10 \times DNase I buffer: 400 mM Tris-HCl, pH 7.5, 100 mM NaCl, 60 mM MgCl₂.
3. Phenol/chloroform/isoamyl alcohol (25/24/1), equilibrated with 0.1 M sodium citrate (pH 4.5), 1 mM EDTA.
4. 95% ethanol.
5. 2 M NaOAc (pH 4.5).
6. 10 \times termination mix: 0.1 M EDTA, pH 8.0, 1 mg/mL glycogen (Sigma #G1508 or Boehringer-Mannheim #901-393).

2.3. cDNA Probe Synthesis

1. 10 \times dNTP mix (5 mM each of dCTP, dGTP, dTTP).
2. CDS primer mix.
3. 5 \times reaction buffer.
4. MMLV reverse transcriptase.
5. DTT (100 mM).
6. 10 \times termination mix: 0.1 M EDTA, pH 8.0, 1 mg/mL glycogen.
7. C₀t-1 DNA (1 mg/mL).
8. Deionized water.
9. (α -³²P)dATP (10 μ Ci/ μ L; 3000 Ci/mmol; Amersham #PB10204).

2.4. Purification of Probe

1. Nucleospin extraction spin columns (Clontech).
2. 2-mL collection tubes.
3. Buffer NT2 (Clontech).
4. Buffer NT3 (Clontech) (add 15 mL of 95% ethanol before use).
5. Buffer NE (Clontech).

2.5. Choice of cDNA Microarrays

We have routinely purchased arrays from Clontech and used radiolabeled probes for quantitation. Other companies supply such arrays, and we do not represent any to be better than others.

2.6. Hybridization and Washing

1. Sheared salmon testes DNA (10 mg/mL; Sigma #D7656).
2. 10× denaturing solution: 1 M NaOH, 10 mM EDTA.
3. 2× neutralizing solution: 1 M NaH₂PO₄, pH 7.0.
4. 20× SSC (1× SSC=0.15M NaCl and 0.015M sodium citrate).
5. 20% sodium dodecyl sulfate (SDS).
6. Wash solution 1: 2× SSC, 1% SDS.
7. Wash solution 2: 0.1× SSC, 0.5% SDS.

2.7. Analysis of Arrays

Phosphorimager. We happen to use an instrument from Molecular Dynamics. However, other instruments are available that to our knowledge are just as functional.

3. Methods

3.1. Origin of the IB3 Cells

IB3 cells are SV40 transformed cells derived from tracheal epithelial cells of a CF patient who received a lung transplant at Johns Hopkins University (4). The IB3 cells are compound heterozygous mutants for the CFTR locus. One allele is the frequent mutation, $\Delta F508$; the other is the less common mutation W1282X. The only expressed CFTR allele in IB3 cells is the $\Delta F508$ (P. Zeitlin, personal communication). Two kinds of AAV-mediated repairs have been performed on this cell. One repair ("C38") was with a truncated CFTR missing the first 119 residues. The second repair was with a full-length CFTR ("S9").

3.2. Cell Culture

IB3 cells and repaired cell lines were obtained from Dr. Pam Zeitlin, Johns Hopkins University School of Medicine, and cultured in serum-free LHC-8 medium (Gibco, BRL). The cells are grown to confluence, washed with fresh medium, and then replaced in the incubator. Samples of supernatant medium

were removed at the times shown and placed at -80°C until assayed for secreted materials, such as IL8. The cells themselves are taken for RNA isolation (*see Subheading 3.3.*).

3.3. Treatment of Cells with Bacteria or Drugs

If bacteria are to be included in the experiment, the LHC-8 medium must be modified by omitting the proprietary gentamycin. This allows the bacteria to survive during exposure to the cells. Cells are seeded at 2×10^5 /well, and incubated overnight with drugs or carriers. At time zero, cells are washed with fresh medium and mixed with bacteria. In the case of *Pseudomonas aeruginosa* we use a ratio of bacteria/cell of 15/1.

3.4. Preparation of Total RNA

Total RNA is prepared by the method of Chirgwin et al (5).

3.5. Treatment of Total RNA with DNase 1

The quality of the RNA used to make probes is the most important factor influencing the sensitivity and reproducibility of the hybridization pattern. A poor-quality RNA preparation leads to a high background on the membrane and an inaccurate hybridization pattern. These problems are typically caused by residual RNase and genomic DNA contamination. We use the Atlas Pure method to produce high yields of quality total RNA that is virtually free of genomic DNA, nucleases, and other impurities.

1. Combine the following reagents in a 1.5-mL microcentrifuge tube for each sample: 500 μL total RNA (1 mg/mL), 100 μL 10 \times DNase I buffer, 50 μL DNase I (1 U/ μL), 350 μL deionized water, 1.0 mL total volume
2. Incubate the reactions at 37°C for 30 min.
3. Add 100 μL of 10 \times termination mix.
4. Split each reaction into two 1.5-mL microcentrifuge tubes (550 μL per tube).
5. Add 550 μL of phenol/chloroform/isoamyl alcohol (25/24/1; pH 4.5) to each tube and vortex thoroughly.
6. Spin in a microcentrifuge at 14,000 rpm for 10 min at 4°C to separate phases.
7. Carefully transfer the top aqueous layer to a new 1.5-mL microcentrifuge tube. Avoid pipetting any material from the interface or lower phase.
8. Repeat **steps 5–7** with 550 μL of chloroform
9. Add 1/10 vol (50 μL) of 2 M NaOAc (pH 4.5) and 2.5 vol (1.5 μL) of 95% ethanol.
10. After incubating in ice for 10 min, spin in a microcentrifuge at 14,000 rpm for 15 min at 4°C .
11. Wash the pellet with 500 μL of 80% ethanol.
12. Dissolve the precipitate in 250 μL of RNase-free H_2O and check the quality of total RNA by electrophoresing 0.5–2 μg on a denaturing formaldehyde/agarose/ethyl bromide gel.

Total RNA should have the two bright 28S and 18S rRNA bands at approximately 4.5 and 1.9 kb, respectively. The ratio of intensities of these bands should be 1.5–2.5:1.

3.6. Probe Synthesis from Total RNA

1. Prepare a master mix for all labeling reactions plus one extra reaction to ensure that you have sufficient volume. Combine the following reagents in a 0.5-mL microcentrifuge tube at room temperature:

	per rxn	4 rxns
5× reaction buffer	2 μL	8 μL
10× dNTP mix (for dATP label)	1 μL	4 μL
[α - ^{32}P]dATP (3000 Ci/mmol, 10 $\mu\text{Ci}/\mu\text{L}$)	3.5 μL	14 μL
DTT (100 mM)	0.5 μL	2 μL
Total volume	7 μL	28 μL

(NB: ^{32}P or ^{33}P can be used interchangeably)

2. Preheat a polymerase chain reaction (PCR) thermal cycler to 70°C.
3. For each reaction, combine the following in a labeled 0.5-mL PCR tube:

For experimental RNA samples	For control RNA samples
1–2 μL RNA (2–5 mg)	1 μL RNA
1 μL CDS primer mix	1 μL CDS primer mix

To each tube, add deionized H₂O to a final volume of 3 μL (if necessary).

4. Incubate tubes in a preheated PCR thermal cycler at 70°C for 2 min.
5. Reduce the temperature of the thermal cycler to 50°C (or 48°C if you are using an unregulated heating block or water bath) and incubate tubes for 2 min. During this incubation, add 1 μL MMLV reverse transcriptase per reaction to the master mix (add 4 μL MMLV RT for the 4-reaction master mix). Mix by pipetting, and keep the master mix at room temperature.
6. After completion of the 2-min incubation at 50°C, add 8 μL of master mix to each reaction tube.
7. Mix the contents of the tubes by pipetting and immediately return them to the thermal cycler.
8. Incubate tubes in the PCR thermal cycler at 50°C (or 48°C) for 25 min.
9. Stop the reaction by adding 1 μL of 10× termination mix.
10. Proceed with the column chromatography under **Subheading 3.6**. If necessary, store your probe on ice or at 4°C for a few hours.

3.7. Purification of Labeled cDNA from Unincorporated ^{32}P - or ^{33}P -Labeled Nucleotides by Column Chromatography

1. Dilute probe synthesis reactions to 200 μL total volume with buffer NT2.
2. Place a NucleoSpin extraction spin column into a 2-mL collection tube, and pipet the sample into the column. Centrifuge at 14,000 rpm for 1 min. Discard collection tube and flowthrough into the appropriate container for radioactive waste.

3. Insert the NucleoSpin column into a fresh 2-mL collection tube. Add 400 μL of buffer NT3 to the column. Centrifuge at 14,000 rpm for 1 min. Discard collection tube and flowthrough.
4. Repeat **step 3** twice.
5. Transfer the NucleoSpin column to a clean 1.5-mL microcentrifuge tube. Add 100 μL of buffer NE, and allow column to soak for 2 min.
6. Centrifuge at 14,000 rpm for 1 min to elute purified probe.
7. Check the radioactivity of the probe by scintillation counting:
 - a. Add 2 μL of each purified probe to 5 mL of scintillation fluid in separate scintillation-counter vials.
 - b. Count ^{32}P -labeled samples on the ^{32}P channel, and calculate the total number of counts in each sample.

3.8. Hybridizing cDNA Probes to the cDNA Microarray

1. Prepare a solution of ExpressHyb and sheared salmon testes DNA:
 - a. Prewarm 5 mL of ExpressHyb at 68°C.
 - b. Heat 0.5 mg of the sheared salmon testes DNA at 95–100°C for 5 min, then chill quickly on ice.
 - c. Mix heat-denatured sheared salmon testes DNA with prewarmed ExpressHyb. Keep at 68°C until use.
2. Wet the Atlas array by placing it in a dish of deionized H_2O , and then place the membrane into the hybridization bottle. Pour off all the water from the hybridization bottle. Add 5 mL of the solution prepared in **step 1**. Ensure that the solution is evenly distributed over the membrane. Perform this step quickly to prevent the array membrane from drying (prehybridize for 30 min with continuous agitation at 68°C).
3. Prepare probe for hybridization by following the appropriate steps below.
 - a. Mix together:
 - Labeled probe (entire pool; $0.5\text{--}20 \times 10^6$ cpm): $\sim 200 \mu\text{L}$
 - 10 \times denaturing solution (1 M NaOH, 10 mM EDTA): $\sim 22 \mu\text{L}$
 - Total volume: $\sim 222 \mu\text{L}$
 - b. Incubate at 68°C for 20 min.
 - c. Add the following to your denatured probe:
 - $\text{C}_0\text{t-1}$ DNA (carrier DNA): 5 μL
 - 2 \times neutralizing solution (1 M NaH_2PO_4 , pH 7.0): $\sim 225 \mu\text{L}$
 - Total volume: $\sim 450 \mu\text{L}$
 - d. Continue incubating at 68°C for 10 min.
4. Add the mixture prepared in **step 3** directly to the prehybridization solution.
5. Hybridize overnight with continuous agitation at 68°C.
6. The next day, prewarm wash solution 1 (2 \times SSC, 1 % SDS) and wash solution 2 (0.1 \times SSC, 0.5% SDS) at 68°C.
7. Carefully remove the hybridization solution and replace with 200 mL of prewarmed wash solution 1. Wash the Atlas array for 30 min with continuous agitation at 68°C. Repeat this step three more times.
8. Perform one 30-min wash in 200 mL of prewarmed wash solution 2 with continuous agitation at 68°C.

9. Perform one final 5-min wash in 200 mL of 2× SSC with agitation at room temperature.
10. Using forceps, remove the Atlas array from the container and shake off excess wash solution.
11. Immediately wrap the damp membrane in plastic wrap and expose the Atlas array to a phosphorimaging screen at room temperature.

3.9. Imaging and Quantitation of the cDNA Microarray

1. The imaged data from the Storm PhosphorImager are subjected to graphical organization with the IMP program, or an equivalent. The exact identification of points is verified by the PSCAN program, or an equivalent. With a limited number of points (up to 1000) the data can be conveniently downloaded to a Microsoft Excel spreadsheet. Expression of each gene, in duplicate, are ratioed to the ubiquitin standard or to a basket of genes, chosen on the basis of no change under experimental conditions. This is a matter of taste and skill in design.
2. In the glass slide “chip” format with fluorescent labeled cDNAs, a ratio of messages from two conditions is obtained directly. Companies supplying these “chips” have proprietary software for analysis.

3.10. Data Mining Algorithms from cDNA Arrays

3.10.1. Hierarchical Clustering Algorithms

1. Several versions are available for this algorithm. We have had good experience with the Scanalyze software developed at Stanford University.
2. This algorithm can be used to find similarities between cells under different treatment conditions. A discussion of the caveats important in this approach are given by Pollard et al. (6).

3.10.2. The GRASP Algorithm

1. The analytic strategy we have employed routinely is embodied in the GRASP algorithm. In this algorithm, changes in gene expression due to a given set of experimental conditions are quantitated in terms of the number of standard deviations (SDs) from the mean of all the genes in the array (*see* Srivastava et al., 1999). This technique vastly increases the statistical power of the analysis. Following ratioing to the ubiquitin standard, as mentioned above, each gene on each blot is analyzed in duplicate. In the specific case of the 588 Clontech array, the duplicate positions have an error of ca. 1.3%. Duplicate samples on duplicate blots in this system (or other) were found to vary from each other by only ca. 12%. An expression data for a given gene in a treated cell are then ratioed to expression in the parental cell line. The log of this ratio is then graphed against the log of an equivalent ratio of gene expressions for a different treatment to that of the parental cell line. The diagonal of this orthogonal plot is the distribution of equivalently expressing genes. The relationship of different genes to this diagonal can then be expressed as being inside or outside the ± 1 SD region around the diagonal. Knowledge of relative positions on the orthogonal plot, outside the ± 1

SD region, then allow one to state that the expression of the given gene is or is not related to the mutation or experimental condition. Genes are then assessed as being a given multiple of SDs from the mean of the entire distribution. Mutation on condition-specific genes are then available by inspection.

2. The advantages of the GRASP orthogonal plot is that the specific effects of drugs or other variables can be assessed directly. For example, the place on the plot may vary from the “control” if a drug is added to the cells. The two positions on the graph constitute a “vector,” with magnitude and direction. Movements “toward” or “away” from the diagonal can be interpreted in terms of possible medicinal value. We have specifically developed these genomic vectors for the study of the CFTR-mimetic drug CPX in recombinant HEK-293 cells (2) and refer the reader to that paper for more details.

3.10.3. GENESAVER Algorithm

The GENESAVER algorithm (gene space vector) allows one to use physiological variables to define genes involved in processes related to the variable. In one conceptual formulation of the fundamental problem in cystic fibrosis, it is thought that the epithelial cells in the lung express an intrinsic inflammatory phenotype. This is manifest by elevated IL8 secretion from the epithelial cells. So, one can use the GENESAVER algorithm to determine which genes are expressed in proportion to IL8 secretion under various conditions (3). More generally, any physiological parameter can be used to evaluate the genomic basis of a biological process or disease.

The GENESAVER algorithm is applied as follows.

3. The first step is to identify a physiological parameter that is causatively related to or is symptomatically indicative of the disease. In the case of CF, one can use IL8 or other pro-inflammatory cytokine in the lung. For diabetes, it might be blood sugar. The choice is limited by the disease of interest. This relevant parameter can also be the expression level of a given gene. The relevant physiological parameter is then measured under the different conditions and the results are used to create a multidimensional mathematical vector. Finally, at the time of the physiological variable, one collects RNA from affected cells or tissues, or surrogates that one hopes might reflect the affected tissue.
4. The next analytic step is to take each gene on the array, to calculate a similar multidimensional vector based on the expression levels of that gene in the various experimental conditions. The components of these vectors should be the logarithm of the ratio between the expression level in each experimental condition and in the control for that given gene. We use one of the conditions (i.e., “control”) as a benchmark, and create an $(n-1)$ -dimensional vector (where n is the number of experimental conditions). This allows us to compare between different genes with disparate expression levels, and also to eliminate possible effects from dissimilar binding efficiencies for the cDNAs of different genes.
5. The final step in the analysis is to calculate the angle in multidimensional space between the vector for each gene and the vector for the physiological parameter.

Small angles between vectors indicate that there is a correlation between the changes in expression levels of specific genes and the changes in the physiological parameter. Angles close to 180° indicate anticorrelation between the gene and the parameter. This means that in those experimental conditions where the parameter increases, the gene expression is reduced, and vice versa.

6. The advantage of the GENESAVER algorithm is that the approach allows one to identify genes that exhibit a pattern of expression that is similar to a physiological parameter relevant to the studied disease. The advantages of this approach include the ability to test a hypothesis regarding signaling pathway genes involved in the physiological or disease process of interest. In the case of our CF experiments, IL8 secretion is highly correlated with selected genes from the TNF α R/NF κ B pathway (3). However, since the algorithm requires no a priori assumptions, it is useful for investigations that are “discovery driven.” For those interested in the details of this algorithm, and its application to specific conditions, please contact the authors.

References

1. Welsh, M. J., Tsui, L.-C., Bost, T. F., and Beaudet, A. L. (1995) Cystic fibrosis, in *The Methods and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds.), McGraw-Hill, New York, pp. 3799–3876.
2. Srivastava, M., Eidelman, O., and Pollard, H. B. (2000) Gene expression microarray pharmacogenomics of wt and mutant CFTR, and of the CF drug CPX. *Mol. Med.* **5**, 753–767.
3. Eidelman, O., Srivastava, M., Zhang, J., Murtie, J., Jacobson, K., Metcalf, E., Weinstein, D., and Pollard, H. (2001) Control of the proinflammatory state in cystic fibrosis lung epithelial cells by genes from the TNF α R/NF κ B pathway. *Mol. Med.* **7**, 523–534.
4. Zeitlin, P. L., Lu, L., Hwang, T.-C., Rhim, J., Cutting, G. R., Keiffer, K. A., Craig, R., and Guggino, W. B. (1991) A cystic fibrosis bronchial epithelial cell line: immortalization by adeno12-SV40 infection. *Am. J. Respir. Cell Mol. Biol.* **4**, 313–319.
5. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294–5299.
6. Pollard, H. B., Eidelman, O., Jacobson, K. A., and Srivastava, M. (2001) Pharmacogenomics of cystic fibrosis. *Mol. Intervent.* **1**, 54–63.



<http://www.springer.com/978-0-89603-897-4>

Cystic Fibrosis Methods and Protocols

Skach, W.R. (Ed.)

2002, XVI, 615 p., Hardcover

ISBN: 978-0-89603-897-4

A product of Humana Press