

## Investigation of Acquired Resistance to EGFR-Targeted Therapies in Lung Cancer Using cDNA Microarrays

Kian Kani, Raffaella Sordella, and Parag Mallick

### Abstract

Clinical tools to accurately describe, evaluate, and predict an individual's response to cancer therapy are a field-wide priority; in many advanced cancers, only 10–20% of individuals will have a clinical benefit from therapy, yet we treat the entire population. Furthermore, many therapies are initially effective, but lose effectiveness over time. Here we describe methods to derive in vitro models of resistance to EGFR tyrosine kinase inhibitors. We additionally describe approaches to characterize possible mechanisms of resistance by genomic and transcriptomic approaches.

**Key words:** Gefitinib, Erlotinib, Non-small-cell lung cancer, Therapeutic response, Proteomics, Resistance

---

### 1. Introduction

Mechanistically, changes in cancer cell growth patterns and rates may originate with alterations in growth signaling networks. The concept of “oncogene addiction” has been demonstrated in a variety of mouse tumorigenesis models in which continued expression of the transforming oncogene is required for tumor maintenance. Numerous targeted therapeutics attempt to halt the growth of a tumor via specific disruptions in these networks. For example, the efficacy of drugs such as imatinib illustrate the importance of BCR-ABL in chronic myelogenous leukemia and the dramatic responses of epidermal growth factor receptor (EGFR)-mutant lung tumors to EGFR tyrosine kinase inhibitors (TKIs), such as erlotinib and gefitinib, suggest that even complex epithelial cancers may exhibit dependency on a single activated kinase for survival. Specifically

in the case of lung cancer, much excitement has recently been generated by the finding that a group of non-small-cell lung cancer (NSCLC) patients benefit highly from treatment with selective EGFR inhibitors (1). The clinical success of tyrosine kinase inhibitors (TKIs) has dramatically influenced research and medical oncology over the past decade. Unfortunately, these therapies are often only effective in a small percentage of patients. Furthermore, even when initially effective, therapies can lose effectiveness over time, presumably through the tumor's molecular evolution of resistance to the therapeutic. Despite extensive study, *de novo* and acquired resistance to targeted therapeutic agents remains a major obstacle to improving remission rates and achieving prolonged disease-free survival.

Through the use of derived models of resistance numerous mechanisms of response and resistance have been uncovered. NSCLC patients with activation mutations in the EGFR gene exhibit dramatic sensitivity to targeted TKIs (2). These mutations remove some or all of the negative allosteric regulatory mechanisms in the EGFR kinase domain (3). The most frequent activating mutations are the in-frame deletions of leucine-747 to glutamic acid-749 (DLRE) and the leucine to arginine substitution at codon 858 (L858R), which account for 44% and 41% of all the mutations in EGFR cases in NSLC, respectively (4, 5). Subsequent studies also identified a mutation of threonine to methionine at 790 that confers resistance. In addition to mutations in the EGFR, amplification of *c*-MET has recently been identified as another potential acquired resistance mechanism. In a preliminary study, 22% (4/18) of cases of NSCLC with acquired resistance to gefitinib/erlotinib have been shown to harbor *c*-MET amplification (6). Interestingly, some of these tumors contained both the T790M mutation and *c*-MET amplification. Although the role of *c*-MET amplification in EGFR TKI-acquired resistance has yet to be definitively proven, current studies suggest that the presence of T790M and *c*-MET amplification could account for approximately 50% of all cases (6). Thus, despite considerable progress in identifying acquired resistance mechanisms, new determinants of EGFR TKI have yet to be uncovered in the majority of tumors that have relapsed.

The accumulation of genetic/epigenetic aberrations that ultimately lead to the acquisition of resistance to a particular TKI treatment in fact usually results in dramatic changes in the expression pattern of multiple genes. In this regard, a strong correlation between gene expression and copy number variations has been observed. As such, gene expression profiling or microarray analysis, by enabling the measurement of thousands of genes in a single RNA sample, provides a powerful research tool for elucidating genetic interactions and response to drug treatments.

Simplistically, while DNA sequence analysis instructs us on what a cell could possibly do as a consequence of the accumulations

of genetic aberrations, the expression profile data conveys us what it is actually happening in a given cell in a particular cell state. The basic idea behind all the different variety of microarray platforms that have been developed is simple: a glass slide or membrane is spotted or “arrayed” with DNA fragments or oligonucleotides that represent specific gene coding regions. Purified RNA is then fluorescently or radioactively labeled and hybridized to the slide/membrane. After thorough washing, the raw data is obtained by laser scanning or autoradiographic imaging. At this point, the data may then be entered into a database and analyzed by a number of statistical methods. The up-front design of a study is critical to a research study’s ability to draw sound conclusions in the particular case of identifying signaling pathways or genes whose dysregulation could contribute to the acquisition of resistance to a given drug treatment. Here, we introduce several techniques for deriving models of resistance. We also describe genomic and transcriptomic techniques for characterizing the mechanism of resistance of those models.

---

## 2. Materials

### ***2.1. In Vitro Selection of Resistant Populations***

1. NSCLC cell lines: e.g., HCC827 cells.
2. Type II classification biological hood and safety cabinet.
3. Cell growth media compatible for cells under investigation: e.g., Dulbecco’s Modified Eagle’s Medium (DMEM) containing 1% of dialyzed fetal bovine serum.
4. Gefitinib: 99% purità, prepare 10 mM stock solution in DMSO.
5. Cell counter, microscope.
6. 96-well plates: flat bottom for cell culture.
7. Cell viability assay kit: e.g., the MTS or MTT assays are commercially available from many sources.
8. Ethyl methane sulfonate (EMS): 600 µg/ml in cell culture medium.

### ***2.2. DNA Sequencing***

DNA can be prepared for sequencing by use of several commercially available kits.

1. DNeasy Blood and Tissue Kit.
2. Pipettes and pipette tips.
3. Vortexer.
4. Microcentrifuge tubes (1.5 ml or 2 ml).
5. Microcentrifuge with rotor for 1.5 ml and 2 ml tubes.
6. Ethanol: 96–100% purity.

7. Phosphate-based saline (PBS).
8. Lysis buffer: Autoclaved PBS supplemented with 1 mM EDTA, 0.5% NP40, 25 mM DTT, 5 mM MgCl<sub>2</sub>, and 1× RNase Inhibitor.

### **2.3. RNA Isolation**

1. Trizol reagent.
2. Chloroform.
3. Ethanol: 200 proof quality.
4. Dulbecco's phosphate-buffered saline (DPBS).
5. Sodium acetate: 3 M, pH 5.2.
6. RNase-free water: 1 ml of 0.1% (v/v) diethylpyrocarbonate (DEPC) in 999 ml of water. Let stand overnight. Autoclave prior to use.
7. Cell scraper.
8. Centrifuge tubes: 15 ml, round bottom, polypropylene.
9. Centrifuge tubes: 50 ml, conical, polypropylene.
10. Plastic tubes: 1.5 ml.
11. PCR tubes: 0.2 ml, thin wall.
12. Centrifugal filter: MicroCon 100.
13. High-speed centrifuge.

### **2.4. Quantitative PCR**

1. PCR tubes: 0.2 ml, thin wall (or plate).
2. Plastic tubes: 1.5 ml, RNase free.
3. Primers for genes of interest.
4. DNA polymerase: JumpStart Taq kit.
5. Reverse transcription: ImProm-II Rt system.

### **2.5. Transcriptomic Profiling**

1. Ethanol: 200 proof quality.
2. Dulbecco's phosphate-buffered saline (DPBS).
3. Sodium acetate: 3 M, pH 5.2.
4. Nucleotides: dATP, dCTP, dGTP, dTTP, 100 mM each, store frozen at -20°C.
5. Primers: 1 mg/ml pd(T)<sub>12-18</sub> in water, store frozen at -20°C.
6. Oligo primer: anchored, 5'-TTT TTT TTT TTT TTT TTT TTV N-3', resuspend at 2 mg/ml, store frozen at -20°C.
7. Fluorescent dyes: 1 mM CyTM3-dUTP and 1 mM CyTM5-dUTP, store at -20°C, protect from light.
8. Rnase inhibitor: RNasina, store at -20°C.
9. SUPERSRIPT™ II Rnase H<sup>-</sup> Reverse Transcriptase Kit: store at -20°C.
10. Human C0t-1 DNA: 10 mg/ml, store frozen at -20°C.

11. Ethylenediaminetetraacetic acid (EDTA): 0.5 M, pH 8.0.
12. NaOH: 1 N.
13. Tris(hydroxymethyl)aminomethane (Tris)-HCl: 1 M, pH 7.5.
14. TE buffer (1×): 10 mM Tris, 1 mM EDTA pH 7.4.
15. Diethylpyrocarbonate (DEPC)-treated water: 0.1% v/v diethylpyrocarbonate in water.
16. Tris-Acetate electrophoresis buffer (TAE, 50×): 2 M Tris acetate pH 8.0, 55 mM EDTA in DEPC-treated water.
17. Centrifuge tubes: 15 ml, round bottom, polypropylene.
18. Centrifuge tubes: 50 ml, conical bottom, polypropylene.
19. Plastic tubes: 1.5 ml.
20. PCR tubes: 0.2 ml, thin wall.
21. Centrifugal filter: MicroCon 100.
22. High-speed centrifuge for 15-ml tubes.
23. Nucleotides: 100 mM dGTP, dATP, dCTP, and dTTP.
24. First-Strand Buffer: 5×, provided with Superscript II kit.
25. Poly d(A): 8 mg/ml in water.
26. Yeast tRNA: 4 mg/ml in water.
27. SSC buffer (20×): 3 M sodium chloride, 300 mM sodium citrate, pH 7.0.
28. Denhardt's blocking solution (50×): 1% Ficoll type 400, 1% polyvinylpyrrolidone, and 1% bovine serum albumin.
29. For our studies, we used the Affymetrix platform including the Affymetrix U95Av2 array that contains approximately 12,600 human genes and the Affymetrix GeneChip Scanner 3000. Other alternative systems may also be used.

---

## 3. Methods

### **3.1. Generation of Gefitinib Resistant Cell Populations in Cell Culture**

#### *3.1.1. Titration Method*

1. Maintain cells for several passages to ensure normal growth (see Note 1).
2. Plate the cells at 3,500 cells per well in quadruplicate in 96-well plates (flat bottom, cell culture plated).
3. 18 h after cell plating, wash the cells with fresh medium and then add medium supplemented with different concentrations (e.g., threefold dilution series) of gefitinib (see Note 2). Incubate the cells for 2–3 days (see Note 3).

4. Measure cell viability, for example, using the MTS assay with UV detection at 490 nm in a plate reader. Determine the drug concentration at which 50% of all cells are viable (IC<sub>50</sub> value). This drug concentration is used as a starting point for subsequent cell dosing.
5. Start dosing the cells at one order of magnitude below the determined IC<sub>50</sub> value (e.g., 100 nM).
6. Plate the cells in a 10-cm cell-culture plate.
7. Incubate the cells with 10 nM gefitinib (see Note 4) and replenish drug containing media every 24 h.
8. Monitor cell viability and increase gefitinib concentration as the cells regain normal growth kinetics (see Note 5).
9. Expand the cells to desired quantities for subsequent experiments (see Notes 6 and 7).

### **3.1.2. Rapid Selection Method**

In addition to the above described titration method, a rapid selection protocol with or without exposure to mutagen can be employed. A mutagen such as EMS can be used to increase genetic aberrations. In this case, cells are treated prior to selection with EMS at a concentration of 600 µg/ml and allowed to recover for 72 h.

1. Maintain the cells for several passages to ensure normal growth.
2. Measure IC<sub>50</sub> by performing a MTS viability assay with different concentrations of gefitinib as described in Subheading 3.1.1.
3. Plate the cells in a 10 cm cell culture plate at a 30% confluency (i.e.,  $6 \times 10^4$  cells per 10 cm<sup>2</sup>).
4. Incubate the cells with three times the IC<sub>50</sub> in the presence of 5% serum (selection media). In the case of the HCC827 cells, this corresponds to 1 µM gefitinib.
5. Replenish every 24 h with new selection media (RPMI, 5% FBS, Strep/pen Invitrogen).
6. Expand the cells to the desired quantities for subsequent experiments.

### **3.2. Characterization of Resistant Cell Populations by DNA Sequencing**

1. Grow cells to 65% confluence in 10-cm plates and wash three times with PBS.
2. Scrape the cells off the plate with 400 µl of lysis buffer.
3. Add 40 µl proteinase K. Check pH. If required, adjust pH using hydrochloric acid or glacial acetic acid.
4. Add 400 µl Buffer AL provided in the DNeasy kit (without added ethanol) to the sample and mix thoroughly by vortexing.
5. Add 400 µl ethanol (96–100%) and mix again thoroughly by vortexing. It is important that the sample and the ethanol are mixed thoroughly to yield a homogeneous solution.

6. Pipette the mixture from step 5 (including any precipitate) into the DNeasy Mini spin column placed in a 2-ml collection tube (provided). Centrifuge at  $\geq 6,000 \times g$  for 1 min. Discard the flow-through and collection tube.
7. Place the DNeasy Mini spin column in a new 2-ml collection tube (provided), add 500  $\mu$ l Buffer AW1 (provided in the DNeasy kit), and centrifuge for 1 min at  $\geq 6,000 \times g$ . Discard the flow-through and collection tube.
8. Place the DNeasy Mini spin column in a new 2-ml collection tube (provided), add 500  $\mu$ l Buffer AW2 (provided in the DNeasy kit), and centrifuge for 3 min at  $20,000 \times g$  to dry the DNeasy membrane.
9. Place the DNeasy Mini spin column in a clean 1.5 ml or 2 ml microcentrifuge tube (not provided) and pipette 400  $\mu$ l Buffer AE (provided in the DNeasy kit) directly onto the DNeasy membrane. Incubate at room temperature for 1 min and then centrifuge for 1 min at  $\geq 6,000 \times g$  to elute the sample.
10. Measure the DNA quality and purity by use of OD 260/280 and send out for commercial sequencing (see Note 8).

### **3.3. Interrogation of Markers of Epithelial to Mesenchymal Transition by Quantitative PCR**

Here, we provide a reliable protocol to assess the expression of most commonly used epithelial to mesenchymal transition (EMT) markers by using QPCR-based techniques (see Note 9).

#### **3.3.1. RNA Isolation**

Many protocols and commercially available kits are currently available for the generation of RNA. We routinely employ the following method based on a combination of phase extraction and chromatography.

1. Wash subconfluent cells (i.e., approximately  $5 \times 10^6$  cells) twice with cold PBS on ice. Scrape the cells with 1 ml of cold PBS and transfer them into sterile RNase-free 1.5-ml microcentrifuge tubes. Pellet by centrifugation at  $13,000 \times g$  for 10 s and remove the supernatant by gently aspiration.
2. Lyse cells by repetitive pipetting in 1 ml of Trizol.
3. Add two-tenth volume of chloroform (i.e., 200  $\mu$ l) and shake vigorously for 15 s.
4. Let mixture stand for 3 min at room temperature. Centrifuge at  $13,000 \times g$  for 15 min at  $4^\circ\text{C}$ .
5. Take off the aqueous phase (i.e., upper phase) and transfer it to a polypropylene tube.

6. Add 0.53 volumes of ethanol (i.e., 500  $\mu$ l) slowly while vortexing (see Note 10).
7. Let mixture stand for 10 min at room temperature and centrifuge at  $13,000\times g$  for 10 min at 4°C.
8. The RNA forms a pellet on the side or bottom of the tube. Discard the supernatant.
9. Wash the pellet two times by adding 1 ml of 75% ethanol and centrifuging at  $10,000\times g$  for 5 min at 4°C.
10. Dry the pellet for 5–10 min and resuspend RNA in 80  $\mu$ l of DEPC-treated water at 65°C.
11. Resuspend RNA at approximately 1 mg/ml in DEPC-treated water.
12. Concentrate the sample by centrifugation ( $500\times g$ ) in a MicroCon 100 filter unit (see Note 11).
13. Determine the concentration of RNA by spectrophotometry. Store at -80°C (see Note 12). To measure the RNA concentration, take 2–5  $\mu$ l RNA sample and dilute with RNase-free water to a final volume of 1 ml in a 1.5-ml microcentrifuge tube (i.e., 200–500 dilution of the RNA sample). Transfer 1 ml of RNase-free water to a clean cuvette and read absorbance as blank. Pipette the diluted RNA sample in to a clean cuvette and read absorbance at 260 nm and 280 nm. To determine RNA concentration of the original sample, use the following formula:  

$$[\text{RNA } \mu\text{g}/\mu\text{l}] = A_{260} \times 33 \times \text{dilution factor}/1,000.$$

### 3.3.2. Reverse Transcriptase Reaction

Many commercial kits are available for the reverse transcriptase reaction. The following protocol has been optimized for the ImProm-II Rt system.

1. Preparation of a reaction master mix is highly recommended to give best reproducibility. Mix all reagents but template in a common mix, using ~10% more than needed. A 20  $\mu$ l reaction contains 4  $\mu$ l reaction buffer, 3 mM  $\text{MgCl}_2$ , 0.5 mM of each dNTP, 1  $\mu$ / $\mu$ l ribonuclease inhibitor (we suggest recombinant RNAasin), and 1  $\mu$ l of reverse transcriptase (we suggest the Improm-II system, see Note 13).
2. Anneal RNA and oligos for 5 min at 25°C.
3. For each reaction, combine 15  $\mu$ l of reverse transcriptase mix with RNA and oligos for a final volume of 20  $\mu$ l.
4. The first-strand cDNA is synthesized by incubating the reaction for 60 min at 42°C (i.e., the extension reaction) and terminated by incubating at 70°C for 15 min.

### 3.3.3. Quantitative PCR

Quantitative PCR uses the linearity of DNA amplification to determine relative amounts of a known sequence in a sample by using a fluorescent reporter in the reaction. When the DNA is in

**Table 1**  
**Pipetting scheme for QPCR**

Volume*	Reagent	Final concentration
25 $\mu$ l	2 $\times$ JumpStart Taq ReadyMix	1.5 units Taq DNA polymerase, 10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl <sub>2</sub> , 0.001% gelatin, 0.2 mM dNTP, stabilizers
– $\mu$ l	Forward primer	50–1,000 nM (see Notes 15 and 16)
– $\mu$ l	Reverse primer	50–1,000 nM
– $\mu$ l	Template DNA	10–100 ng
– to 50 $\mu$ l	Water	
50 $\mu$ l	Total volume	

**Table 2**  
**PCR cycle conditions**

	Temperature	Time
Initial denaturation	94°C	2 min*
Denaturation	94 °C	15 s
Annealing/extension	60°C or 5°C below lowest primer $T_M$	1 min**
Cycle	40	
(Optional) Hold	4°C – only if products will be run out on a gel	

the log linear phase of amplification, the amount of fluorescence increases above the background. The point at which the fluorescence becomes measurable is called the Threshold Cycle ( $C_T$ ) or crossing point.

1. A reaction master mix (see Note 14) of all reagents but template is assembled as a common mix according to Table 1.
2. Aliquot the master mix into a 200  $\mu$ l PCR tube or plate.
3. Add 1  $\mu$ l of RT reaction.
4. Mix gently by vortexing and briefly centrifuge to collect all components at the bottom of the tube.
5. Perform thermal cycling according to Table 2 (see Note 17).

### 3.4. Gene Expression Profiling

#### 3.4.1. RNA Labeling

The first step in gene expression profiling analysis is the RNA isolation. We refer to Subheading 3.3.1 for a detailed protocol.

1. Anneal the oligo dT(12–18) primer to the RNA in the reaction mixture defined in Table 3 using a 0.2-ml thin-wall PCR

**Table 3**  
**Pipetting scheme for RNA annealing**

Component	Addition for Cy5 labeling	Addition for Cy3 labeling
Total RNA (>7 mg/ml)	150–200 µg	50–80 µg
dT(12–18) primer (1 µg/µl)	1 µl	1 µl
DEPC H <sub>2</sub> O	to 17 µl	to 17 µl

**Table 4**  
**Pipetting scheme for RNA labeling**

Component	Volume [µl]
5× first-strand buffer	8
10× low T dNTPs mix	4
Cy5 or Cy3 dUTP (1 mM)	4
0.1 M DTT	4
Rnasin (30 µ/µl)	1
Superscript II (200 µ/µl)	2
Total volume	23

tube so that incubations can be carried out in a PCR cycler (see Note 18).

- Heat to 65°C for 10 min and transfer on ice for 2 min.
- Add 23 µl of the reaction mixture defined in Table 4 containing either Cy5-dUTP or Cy3-dUTP nucleotides, mix well by pipetting, and use a brief centrifuge spin to concentrate the liquid in the bottom of the tube.
- Incubate at 42°C for 30 min. Then add 2 µl Superscript II. Make sure that the enzyme is well mixed in the reaction volume and incubate at 42°C for 30–60 min (see Note 19).
- Stop the reaction by adding 5 µl of 0.5 M EDTA (see Note 20).
- Add 10 µl of 1 N NaOH and incubate at 65°C for 30 min to hydrolyze residual RNA. Cool to room temperature (see Note 21).
- Neutralize the reaction by adding 25 µl of 1 M Tris-HCl, pH 7.5.
- Desalt the labeled cDNA by adding the following neutralized reaction: 400 µl of TE pH 7.5 and 20 µg of human C0t-1 DNA. Transfer the solution to a MicroCon 100 cartridge. Pipette to mix and spin for 10 min at 500×g.

9. Wash again by adding 200  $\mu\text{l}$  TE pH 7.5 and concentrating to about 20–30  $\mu\text{l}$  by centrifuging at  $500\times g$  for 8–10 min.
10. Recover the sample by inverting the concentrator over a clean collection tube and spinning for 3 min at  $500\times g$ .
11. Take a 2–3  $\mu\text{l}$  aliquot of the Cy5-labeled cDNA for analysis, leaving the remaining 18–28  $\mu\text{l}$  for hybridization.
12. Run this probe on a 2% agarose gel (6 cm wide  $\times$  8.5 cm long, 2 mm wide teeth) in Tris Acetate Electrophoresis Buffer (TAE, see Note 22).
13. Scan the gel on a fluorescence scanner (setting: red fluorescence, 200  $\mu\text{m}$  resolution, see Note 23).

### 3.4.2. Hybridization

1. As an initial step, it is necessary to determine the volume of the hybridization solution required (see Note 24). Usually, 0.033  $\mu\text{l}$  of hybridization solution is used for each  $\text{mm}^2$  of slide surface area covered by the coverslip for each array is used (i.e., an array covered by a 24 mm by 50 mm coverslip will require 40  $\mu\text{l}$  of hybridization solution).
2. For a 40  $\mu\text{l}$  hybridization, pool the Cy3- and Cy5-labeled cDNAs into a single 0.2-ml thin-wall PCR tube and adjust the volume to 30  $\mu\text{l}$  by either adding DEPC-treated water, or by removing water in a vacuum concentrator. If using a vacuum device to remove water, do not use high heat or heat lamps to accelerate evaporation as the fluorescent dyes might be degraded.
3. For a 40- $\mu\text{l}$  hybridization, combine the components specified in Table 5.
4. Mix the components well by pipetting, heat at  $98^\circ\text{C}$  for 2 min in a PCR cycler, cool quickly to  $25^\circ\text{C}$ , and add 0.6  $\mu\text{l}$  of 10% SDS.
5. Centrifuge for 5 min at  $14,000\times g$ . The labeled cDNAs have a tendency to form small, very fluorescent, aggregates

**Table 5**  
**Pipetting scheme for array hybridization**

Component	High sample blocking	High array blocking
Cy5 + Cy3 probe	30 $\mu\text{l}$	28 $\mu\text{l}$
Poly d(A) (8 mg/ml)	1 $\mu\text{l}$	2 $\mu\text{l}$
Yeast tRNA (4 mg/ml)	1 $\mu\text{l}$	2 $\mu\text{l}$
Human C0t-1 DNA (10 mg/ml)	1 $\mu\text{l}$	0 $\mu\text{l}$
20 $\times$ SSC	6 $\mu\text{l}$	6 $\mu\text{l}$
50 $\times$ Denhardt's blocking solution	1 $\mu\text{l}$ (optional)	2 $\mu\text{l}$
Total volume	40 $\mu\text{l}$	40 $\mu\text{l}$

which result in bright, punctate background on the array slide. Hard centrifugation will pellet these aggregates, avoiding their introduction to the array.

6. Apply the labeled cDNA to a 24 mm × 50 mm glass coverslip and then touch with the inverted microarray (see Note 25).
7. Place the slide in a microarray hybridization chamber. Add 5  $\mu$ l of 3 $\times$  SSC to the reservoir and seal the chamber. Submerge the chamber in a 65°C water bath and allow the slide to hybridize for 16–20 h (see Note 26).
8. Wash off unbound fluorescent cDNA by incubating in prewarmed 2 $\times$  SSC, 0.2% SDS washing solution at 65°C for 5 min. Then, wash the array three times in 2 $\times$  SSC, 0.2% SDS washing solution at RT for 2 min. Subsequently, air-dry the array.
9. Remove the hybridization chamber from the water bath, cool and carefully dry off. Unseal the chamber and remove the slide (see Note 27).
10. Place the slide, with the coverslip still affixed, into a Coplin jar filled with 0.5 $\times$  SSC/0.01% SDS wash buffer. Allow the coverslip to fall from the slide and then remove the coverslip from the jar with a forceps. Allow the slide to wash for 2–5 min.
11. Transfer the slide to a fresh Coplin jar filled with 0.06 $\times$  SSC. Allow the slide to wash for 2–5 min (see Note 28).
12. Transfer the slide to a slide rack and centrifuge at low speed for 3 min in a clinical centrifuge equipped with a horizontal rotor for microtiter plates (see Note 29).
13. Obtain image of array. Images are typically obtained using microarray scanners. A large variety of microarrays scanners are available with different features such as sensitivity, resolution, scan area, rapidity of scan, excitation and emission filters, and autofocus capability. For our studies, we used the Affymetrix GeneChip Scanner 3000.

### **3.5. Data Analysis**

#### *3.5.1. Array Segmentation, Background Intensity Extraction, and Target Detection*

1. The first step in any analysis of microarray data is the extraction of quantitative information from the images resulting from the readout of hybridizations. As a preliminary analysis, the data from each experiment is assessed for overall quality and to determine the extent of difference between the experimental sample and the reference sample. A segmentation of the array is initially performed to decode the array information. Since each element of an array is printed automatically to a predefined location and because specific orientation markers are given, the initial positioning of the grid can be automatically aligned.
2. The background of the microarray image is usually not uniform over the entire array. Therefore, it is necessary to extract local background intensities. Owing to many technical reasons,

the changes of fluorescent background across an array are usually gradual and smooth. In case more abrupt changes are observed, we recommend caution in the interpretations of the data. Conventionally, pixels near the bonding box edge are taken to be the background pixels, and the average gray level of these pixels are used for the estimation of the local background intensity.

3. A fixed thresholding method is used in image analysis to determine the signal corresponding to a specific target over the background. A threshold value ( $T$ ) corresponding to changes of over three times the value of standard deviation ( $s$ ) over the local background mean ( $m$ ) is usually used (e.g.,  $T = m + 3s$ ). However, this method based on a simple fixed thresholding fails quite often due to variability of the background and the signal, particularly when the signal is weak (a frequent finding in cDNA array experiments). To avoid these problems, a more sophisticated thresholding method such as the Mann–Whitney method may be implemented (7).
4. The local background value is subtracted from the reported sample intensities from the red channel (R) and the green channel (G), and then the ratio (R/G) is calculated. Clearly, the ratio measurement is the ratio of two average intensity measurements. We usually use Affymetrix Expression Console QC metrics to pass the image data, although other software packages are available and can be used.

### 3.5.2. Data Visualization

A second general step in the analysis of gene expression profile data is the collection of experimental data into a database that supports both further mathematical analysis and connection to the available knowledge about the structure and function of the individual genes (Database Design & Development). Usually, this step is automatically generated by current software packages.

### 3.5.3. Gene Annotation

1. Gene annotation provides functional information on a given gene. In addition, depending on the tool utilized this analysis can also provide other useful information such as for example the location of each gene within a particular chromosome. Having identified some set of regulated genes, the next step in expression profiling analysis involves looking for patterns within the regulated set. The way in which particular genes' expression patterns vary across the samples can be analyzed in a variety of ways. In expression clustering analysis, a query asks whether the similarities of patterns between genes suggest involvement in common processes, or whether differences in the individual gene expression patterns can be aligned with differences in the sample types. In Discriminative Gene List analysis, genes having the most differential behavior between

samples can be used to try to identify the particular cellular activities that differ between classes of samples. By using these two types of analysis, one can start to make biological sense of expression profiling data. Many analytical tools are currently available for data mining and to determine whether, for example, proteins made from genes with similar patterns of expression perform related functions, whether they are chemically alike or have similar subcellular localization.

2. As part of available tools, Gene ontology analysis provides a standard way to define these relationships. Gene ontologies start with very broad categories, e.g., “metabolic process,” and break them down into smaller categories, e.g., “carbohydrate metabolic process,” and finally into quite restrictive categories such as “inositol and derivative phosphorylation.” The Molecular Signatures Database, the Comparative Toxicogenomics Database and the Ingenuity Gene Network Diagram are examples of resources to further categorize genes.

#### 3.5.4. Validation by Statistical Analysis

1. Data analysis of microarrays has become an area of intense research. Simply stating that a group of genes are regulated by at least twofold lacks a solid statistical footing. The usual typical three replicates in each group are in fact not sufficient from a statistical perspective and can easily create a deceptive difference greater than twofold. Rather than identifying differentially expressed genes using a fold change cutoff, one can use a variety of statistical tests such as ANOVA, all of which estimate how often we would observe the data by chance alone. This type of analysis also allows the identification of genes that, despite subtle variation in expression, could have important biological functions.
2. In addition to treatment (i.e., drug) or comparison of multiple cell lines, it is also helpful to include information such as hybridization, operator, scanner, etc. as variable control items.
3. The output of an ANOVA test includes two values: a  $p$ -value and a mean square for each factor tested. The mean of the mean square for each factor is used to quantify the contribution of each factor to the experimental variation. The  $p$ -values indicate instead the significance of an observed difference in the expression. The lower the  $p$ -value, the more confidence one has that the difference detected are not due to random chance. A  $p$ -value of 0.05 is typically defined to indicate significance, since it estimates a 5% probability of observing the data by chance. A  $p$  value alone of course is not sufficient since for example analysis of 10,000 genes on a microarray would result in the identification of 500 genes at  $p < 0.05$  even if there were no difference between the experimental groups. A variety of analysis packages from bioinformatics

companies are available to account for this so-called multiple hypothesis testing. Among these tools, we use the Significance Analysis of Microarrays (SAM) system. In addition, we suggest following the MAQC Project recommendations for microarray analysis.

#### *3.5.5. Validation by QPCR and Western Blot Analysis*

Ultimately, changes in the expression of certain genes of interest should be further validated by more direct methods such as quantitative PCR and western blot analysis. While high throughput DNA microarrays lack the quantitative accuracy of QPCR, it takes about the same time to measure the gene expression of a few dozen genes via QPCR as it would to measure an entire genome using DNA microarrays. Thus, we suggest performing semiquantitative DNA microarray analysis experiments to identify candidate genes and then perform QPCR on some of the most interesting candidate genes to validate the microarray results. Western blot analysis of some of the protein products of differentially expressed genes is also recommended, since the mRNA levels do not necessarily correlate to the amount of expressed protein. We refer to the previous section for a detailed protocol for QPCR analysis.

---

## 4. Notes

1. Cell growth should be carefully monitored, as some cells take longer to obtain a resistant phenotype (see Subheading 3.1.1).
2. Drug resistance can be obtained by incremental titration of a particular drug over a several month time course. For example, the IC<sub>50</sub> of the HCC827 cell line treated with gefitinib is in the nanomolar range (10–100 nM). Once the IC<sub>50</sub> has been determined for a particular cell line and drug, the initial dosage can be set. For example, if the IC<sub>50</sub> for the HCC827 cell line with gefitinib is ~100 nM, then one could begin dosing with 10 nM of gefitinib for 1 month. The exact concentration of gefitinib has to be established by a case-by-case basis, but initial dosing at concentrations above the IC<sub>50</sub> can result in apoptosis and the release of cell debris that can prevent proper cell attachment and growth (see Subheading 3.1.1).
3. The rate of the development of an acquired resistance cell line also depends on the frequency of drug replenishment. Gefitinib is an anilinoquinazoline and thus will undergo latent hydrolysis in aqueous and is insoluble in alkaline pH. Therefore, the replenishment of the gefitinib containing media will determine the effective inhibition time of EGFR. The half-life of the therapeutic in vitro will determine the rate and possibly influence the mechanism of resistance. This parameter should

be addressed prior to dosing and documented in any manuscript (see Subheading 3.1.1).

4. Acquired resistance can develop over a course of several months. For the HCC827 example, if the initial dosing started at 10 nM, then one can incrementally increase gefitinib threefold until a particular stopping point is achieved. Cell viability should be monitored regularly and drug concentration increased as cells regain their normal growth kinetics at each dosing level. In our hands, HCC827 gefitinib-resistant cells demonstrated normal growth kinetics after 4 months and ultimately handled gefitinib up to 10  $\mu$ M. However, as with any therapeutic in high concentration, the number of off-target hits increases leading to overlapping inhibition of multiple kinases. It would be reasonable to generate a line that is resistant to gefitinib at concentrations above 10  $\mu$ M as long as the final resistant pool was maintained in 1  $\mu$ M to avoid overlapping inhibition (see Subheading 3.1.1).
5. After approximately 2 weeks, resistant clones start to be visible at a frequency of 0.01% (i.e., 100 clones for  $1 \times 10^6$  selected cells). Each clone can then be expanded from a 96-well plate to a 10 cm dish in the presence of selecting medium (see Subheading 3.1.1).
6. At this point, usually the resistant derived cell lines have become stably resistant to the drug treatment and can be grown in normal media without drug (see Subheading 3.1.1).
7. The in vitro development of acquired resistance starts with a homogenous cell type, which has been a priori deemed sensitive to a particular therapy. In the case of NSCLC tissue, several cell lines with activation mutations in the EGFR kinase domain have been isolated and cataloged at various bioresource centers (ATCC in the USA, JCRB in Japan). Once the appropriate cell line has been obtained, one should confirm the sensitivity to EGFR targeted tyrosine kinase inhibitors (TKIs) by performing a cell viability assay to demonstrate the half maximal inhibitory concentration (IC<sub>50</sub>) of the particular TKI (see Subheading 3.1.1).

Limitations: The described protocols make use of in vitro-based cell culture systems. Although these systems are easy to implement and are based on a limited number of variables, they do not represent the high complexity of actual tumors. New experimental evidence is now indicating that also the tumor microenvironment could play an important role in modifying the response of cells to drug treatment. As such new protocols based on in vivo selections are currently under development. In addition, one should also be aware that the cell lines currently in use are usually highly genetic unstable and therefore genomic abnormalities observed in this culture

system could not reflect actual mechanism of resistance. The possibility of utilizing early passages tumors cells could overcome these limitations.

8. In order to characterize modes of acquired TKI resistance, it is imperative to sequence exon 18–20 of EGFR. This is particularly important because a low percentage of T790M is sufficient to confer resistance (8, 9) even when introduced in cis (10). As a result, traditional sequencing might not be sufficiently stringent enough to identify a small proportion of T790M mutant bearing EGFR genes. One such method relies on an exonuclease derived from celery (CEL I) (11). This enzyme demonstrates preference for cutting at 3' to the mismatched nucleotides on both strands of DNA. CEL I recognizes deletions and insertion mutations irrespective of context within the DNA. This has been adapted to fully automated and commercially available sequencing such as the SURVEYOR (12). DNA should be prepared from any tissue using aseptic technique and purified highly. Polymerase chain amplification of the region of interest is then performed on heterozygote alleles followed by heat denaturation and slow reannealing generates a population of dsDNA heteroduplexes. Enzymatic digestion of the PCR product by SURVEYOR yields cleaved fragments. These fragments can be separated by HPLC column to further increase signal-to-noise prior to sequencing. This technique has been applied to EGFR mutation screens with success in the past (13) (see Subheading 3.2).
9. Resistant populations obtained either by rapid selection protocol or by progressive titration of drug usually display a high degree of heterogeneity in their morphological features. While the majority display characteristic of epithelial cells (i.e., cobble stone morphology, growth in patches) other are characterized by more mesenchymal like appearance (i.e., spindle shape). These morphological differences are usually associated with changes in their growth properties, motility, and invasive capabilities. At the molecular level the switch to a more mesenchymal phenotype (i.e., EMT) has been associated with the activation of complex network of signaling molecules and transcription factors (e.g., TGF $\beta$ , WNT, Notch, NF-KB, STAT3 signaling axis and transcriptional regulators such as Snail, Slug, ZEB1, ZEB2, TWIST, and FOXM1). Interestingly, the cell rewiring that lead to the acquisition of mesenchymal features has also been shown to lead to resistance to certain drug treatment. Decreased sensitivity to erlotinib in NSCLC derived cell lines has for example been recently associated with EMT-like properties (see Subheading 3.3).
10. The ethanol should be added drop by drop and allowed to mix completely with the supernatant before more ethanol is added.

If a high local concentration of ethanol is produced, the RNA will precipitate (see Subheading 3.3.1).

11. This step removes many residual, small to medium-sized molecules that inhibit the reverse transcription reaction in the presence of fluorescently derivatized nucleotides (see Subheading 3.3.1).
12. A ratio between 1.7 and 2 of absorbance at 260 over 280 represent good RNA (see Subheading 3.3.1).
13. The reverse transcriptase reaction is usually set up in 20  $\mu$ l reaction and utilize up to 1  $\mu$ g of RNA allowing the detection of all the target here described. We recommend using a mixture of random priers and oligo (dT) with a final concentration of 0.5  $\mu$ g. It is important to use sterile, nuclease-free reaction tubes and to perform all the reactions on ice (see Subheading 3.3.2).
14. We usually use a volume for 50  $\mu$ l reaction; however, volumes may be scaled to give the desired reaction volumes (see Subheading 3.3.3).
15. Primers are usually kept as a 100  $\mu$ M stock (see Table 1).
16. Table 6 lists the primers that we successfully used to detect the expression of selective EMT markers. The following Web site provide a useful source of validated QPCR primers: <http://web.ncifcrf.gov/rtp/gel/primerdb/default.asp> (see Table 1)
17. Optimal cycling parameters vary with probe design and thermal cyler. We found that an initial denaturation step of greater than 2 min is unnecessary and it is actually not recommended (see Subheading 3.3.3).
18. Because the incorporation rate for Cy5-dUTP is less than that of Cy3-dUTP, more RNA is labeled in the former case to achieve equivalent signal (see Subheading 3.4.1).

**Table 6**  
**Validated primers for detecting the expression of EMT markers**

Gene	Forward primer	Reverse primer
TGFB1	CAACAATTCTGGCGATACCT	GCTAAGGCGAAAGCCCTCAAT
TGFB2	CCCCGAGGTGATTTCCATC	CAGACAGTTTCGGAGGGGA
Ecad	CGAGAGCTACACGTTACGG	GGCCTTTTGACTGTAATCACACC
Vim	AGAACTTTGCCGTTGAAGCTG	CCAGAGGGAGTGAATCCAGATTA
SNAIL	AATCGGAAGCCTAACTACAGCG	GTCCCAGATGAGCATTGGCA
ZEB2	GCTCCGAAGCTGGCAAGAA	GGGACTTGTCATATGCAGTTT

19. The superscript polymerase is very sensitive to denaturation at air–liquid interfaces, so be very careful to suppress foaming in all handling of this reaction (see Subheading 3.4.1).
20. It is important to stop the reaction first with EDTA before adding NaOH, since nucleic acids precipitate in alkaline magnesium solutions (see Subheading 3.4.1).
21. The purity of the sodium hydroxide solution used in this step is crucial. Slight contamination or long storage in a glass vessel can produce a solution that will degrade the Cy5 dye molecule (see Subheading 3.4.1).
22. For maximal sensitivity, do not add ethidium bromide to the gel or running buffer (see Subheading 3.4.1).
23. Platforms: With the continue advances in microarray technologies multiple platforms are currently available both form commercial and for in-house generation. While in principle cDNA and oligos platforms can be easily generated, depending on the number of planned experiments, it can be more prudent to seek the service of an academic microarray core facility or a commercial service. The estimated cost for purchasing a clone set and/or an oligos library, robot, printing pins, and the reagents could vary significantly, but one can expect to need at least \$100,000 to establish such a platform. Using commercial platforms it also eliminates the manufacturing expertise required and ensures more robust and reproducible results. In order to decide which commercial system to use one should consider the total content of the array, the availability of gene annotation, standardize protocol and reagents as well as platform performance, cost, and general acceptance. For our studies, we use an Affimetrix U95Av2 (Sata Clara,CA) that contains approximately 12,600 human genes (see Subheading 3.4.1).
24. The volume of the hybridization solution is critical. If not enough solution is used, the coverslip will bow toward the slide in the center and the hybridization will be nonuniform and likely air bubbles over some portion of the arrayed ESTs will be also introduced. On the contrary, if too much solution is added the coverslip will move too easily compromising the accuracy of the experiment (see Subheading 3.4.2).
25. The hybridization solution is added to the coverslip first, since some aggregates of fluor remain in the solution and will bind to the first surface they touch (see Subheading 3.4.2).
26. There are a wide variety of commercial hybridization chambers. It is worthwhile to prepare a mock hybridization with a blank slide, load it in the chamber, and incubate it to test for leaks, or drying of the hybridization fluid, either of which cause severe fluorescent noise on the array (see Subheading 3.4.2).

27. As there may be negative pressure in the chamber after cooling, it is necessary to remove water from around the seals so that it is not pulled into the chamber and onto the slide when the seals are loosened (see Subheading 3.4.2).
28. The sequence of washes may need to be adjusted to allow for more aggressive noise removal, depending on the source of the sample RNA. Useful variations are to add a first wash which is 0.5× SSC/0.1% SDS or to repeat the normal first wash twice (see Subheading 3.4.2).
29. If the slide is simply air dried, it frequently acquires a fluorescent haze. Centrifuging off the liquids results in a lower fluorescent background. As the rate of drying can be quite rapid, it is suggested that the slide be placed in the centrifuge immediately upon removal from the Coplin jar (see Subheading 3.4.2).

Limitations: Numerous gene expression profiling studies have identified statistically significant changes in expression that occur upon acquisition of resistance. Although informative, this type of analysis bears several limitations. First, because of financial constraints, usually expression profiling experiments are limited to a small number of observations (i.e., three biological replicas) reducing the statistical power of the experiment, making it impossible for the experiment and as such preventing the identification of subtle yet biologically important changes. In addition, cells in fact use many other mechanisms to regulate proteins in addition to altering the amount of mRNA, so these genes may stay consistently expressed even when protein concentrations are rising and falling.

## References

1. Scaltriti, M., and Baselga, J. (2006) The epidermal growth factor receptor pathway: a model for targeted therapy. *Clin. Cancer Res.* **12**, 5268–5272.
2. Lynch, T. J., Bell, D. W., Sordella, R., Gurubhagavatula, S., Okimoto, R. A., Brannigan, B. W., Harris, P. L., Haserlat, S. M., Supko, J. G., Haluska, F. G., Louis, D. N., Christiani, D. C., Settleman, J., and Haber, D. A. (2004) Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **350**, 2129–2139.
3. Choi, S. H., Mendrola, J. M., and Lemmon, M. A. (2007) EGF-independent activation of cell-surface EGF receptors harboring mutations found in gefitinib-sensitive lung cancer. *Oncogene* **26**, 1567–1576.
4. Kumar, A., Petri, E. T., Halmos, B., and Boggon, T. J. (2008) Structure and clinical relevance of the epidermal growth factor receptor in human cancer. *J. Clin. Oncol.* **26**, 1742–1751.
5. Shigematsu, H., and Gazdar, A. F. (2006) Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. *Int. J. Cancer* **118**, 257–262.
6. Engelman, J. A., Zejnullahu, K., Mitsudomi, T., Song, Y., Hyland, C., Park, J. O., Lindeman, N., Gale, C. M., Zhao, X., Christensen, J., Kosaka, T., Holmes, A. J., Rogers, A. M., Cappuzzo, F., Mok, T., Lee, C., Johnson, B. E., Cantley, L. C., and Janne, P. A. (2007) MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* **316**, 1039–1043.
7. Jiang, N., Leach, L. J., Hu, X., Potokina, E., Jia, T., Druka, A., Waugh, R., Kearsey, M. J., and Luo, Z. W. (2008) Methods for evaluating gene expression from Affymetrix microarray datasets. *BMC Bioinformatics* **9**, 284.

8. Kwak, E. L., Sordella, R., Bell, D. W., Godin-Heymann, N., Okimoto, R. A., Brannigan, B. W., Harris, P. L., Driscoll, D. R., Fidias, P., Lynch, T. J., Rabindran, S. K., McGinnis, J. P., Wissner, A., Sharma, S. V., Isselbacher, K. J., Settleman, J., and Haber, D. A. (2005) Irreversible inhibitors of the EGF receptor may circumvent acquired resistance to gefitinib. *Proc. Natl. Acad. Sci. USA* **102**, 7665–7670.
9. Pao, W., Miller, V. A., Politi, K. A., Riely, G. J., Somwar, R., Zakowski, M. F., Kris, M. G., and Varmus, H. (2005) Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Med.* **2**, e73.
10. Engelman, J. A., Mukohara, T., Zejnullahu, K., Lifshits, E., Borrás, A. M., Gale, C. M., Naumov, G. N., Yeap, B. Y., Jarrell, E., Sun, J., Tracy, S., Zhao, X., Heymach, J. V., Johnson, B. E., Cantley, L. C., and Janne, P. A. (2006) Allelic dilution obscures detection of a biologically significant resistance mutation in EGFR-amplified lung cancer. *J. Clin. Invest.* **116**, 2695–2706.
11. Oleykowski, C. A., Bronson Mullins, C. R., Godwin, A. K., and Yeung, A. T. (1998) Mutation detection using a novel plant endonuclease. *Nucleic Acids Res.* **26**, 4597–4602.
12. Qiu, P., Shandilya, H., D'Alessio, J. M., O'Connor, K., Durocher, J., and Gerard, G. F. (2004) Mutation detection using Surveyor nuclease. *Biotechniques* **36**, 702–707.
13. Janne, P. A., Borrás, A. M., Kuang, Y., Rogers, A. M., Joshi, V. A., Liyanage, H., Lindeman, N., Lee, J. C., Halmos, B., Maher, E. A., Distel, R. J., Meyerson, M., and Johnson, B. E. (2006) A rapid and sensitive enzymatic method for epidermal growth factor receptor mutation screening. *Clin. Cancer Res.* **12**, 751–758.