Genomic Approaches to Lung Cancer

Roman K. Thomas,1,3 Barbara Weir,1,3 and Matthew Meyerson1,2,3

Abstract

The last decade has seen remarkable success in clinical targeting of activated oncogenes in cancer. We have applied gene expression profiling, high-density single-nucleotide polymorphism arrays, and systematic resequencing of genes to identify novel oncogenes and signaling pathways in lung cancer that might represent therapeutic targets. Systematic resequencing of tyrosine kinase genes has led to the discovery of somatic mutations in the epidermal growth factor receptor (EGFR) gene in lung adenocarcinomas. These mutations range in frequency from 10% in Caucasian to 40% in East Asian patients. Lung cancer–derived EGFR mutations are oncogenic and are tightly associated with clinical response to the EGFR kinase inhibitors erlotinib and gefitinib. Furthermore, gene expression profiling has been shown to classify patients according to their clinical outcome, indicating that the application of this technique may help in guiding patient selection for therapy in the future. Finally, genome-wide analyses of copy number gains and losses were successfully applied to detect gene amplifications and deletions. Taken together, the application of genomics technologies has led to important discoveries with clinical implications in lung cancer that might help to improve clinical care for patients suffering from this highly fatal tumor.

The complete sequencing of the human genome and the concurrent development of technologies that allow for high-throughput generation of genomic data have opened avenues for a systematic approach to understanding the complex biology of lung cancer (1–3). These technologies mainly comprise gene expression arrays, single-nucleotide polymorphism (SNP) arrays, and high-throughput capillary sequencing. The efforts to determine novel cancer-causing alterations in the lung cancer genome were at least in part inspired by the enormous clinical success of small-molecule inhibitors of tyrosine kinases. The fact that almost all patients with chronic myeloid leukemia can achieve a complete remission by administration of imatinib, an inhibitor that is active against BCR-ABL kinase (the fusion protein of the translocation causing chronic myeloid leukemia), has created a paradigm that is the basis for many cancer genomics projects: find a genetic lesion and find the agent that is active against the lesion (4–6). Genetic changes continued to be detected that are strongly associated with clinical response to inhibitors targeting the respective lesion (mostly kinases). Examples are the clinical success of imatinib in the presence of mutations in c-KIT, platelet-derived growth factor receptor-α, and platelet-derived growth factor receptor-β in gastrointestinal stromal tumor (7–10) and hypereosinophilic syndrome (11), and in acute myeloid leukemia harboring FLT3 mutations (12–14), and, finally, of epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKI) in EGFR-mutant lung cancer (15–17).

The initial application of high-throughput genomic technologies has led to important insights into the biology of lung cancer. For example, recent studies have reported that different types of lung cancer can be subcategorized based on their gene expression profiles (18–20). Additionally, by making use of high-density SNP arrays, it was found that this technology is capable of detecting loss of heterozygosity (LOH) as well as copy number changes and homozygous deletions (21, 22). Finally, using high-throughput sequencing, novel somatic mutations were found in lung cancer, among them mutations in the EGFR gene (16, 23–27). Here, we will give an overview of recent results in the field of lung cancer genomics, with their effect on patient care, and discuss the potential future developments of this field.

Global Gene Expression Analysis Can Subcategorize Lung Cancer and Predict Survival

Gene expression arrays offer the possibility of simultaneously analyzing the transcription of several thousands of genes in a semiquantitative manner. Expression arrays can be broadly subdivided into cDNA arrays and oligonucleotide arrays. On a cDNA array, clones corresponding to the transcripts to be analyzed are spotted onto a matrix. On most oligonucleotide arrays, oligonucleotides corresponding to the transcript and mismatched control oligonucleotides are synthesized on a matrix by photolithography. After hybridizing a given sample on the array, signal intensities can be determined by scanning.
the array that are directly correlated to the amount of a given transcription within the sample.

In lung cancer, several studies have been done using both cDNA and oligonucleotide gene expression arrays (18–20). Sample sizes in these studies ranged from 60 to 186, and the number of interrogated transcripts varied between 5,000 and 20,000. Using unsupervised clustering of the expression array data, all samples in the initial studies were reliably distributed into groups that correspond to the known histologic subtypes (18, 19). Thus, the histopathologic phenotypes of lung cancer correspond to particular gene expression patterns. Furthermore, novel subtypes were identified within the histologic subtype of adenocarcinoma. This is particularly important because this category comprises patients with markedly differing outcomes, for example, the bronchioloalveolar carcinoma (BAC) subtype, which is histopathologically difficult to determine but is characterized by a more favorable prognosis. Among the clusters identified was one (cluster C4) showing features of type II pneumocytes (Fig. 1; ref. 18). Interestingly, 10 of 14 tumors in this category were BACs, compared with a total of 15 BAC cases among the remaining 113 adenocarcinomas, suggesting that cluster C4 is characteristic of BACs. Other clusters discovered in this analysis were cluster C1, with a predominant expression of proliferation-associated genes; cluster C2, with expression of genes reminiscent of neuroendocrine tumors; and cluster C3, which shares one group of genes with cluster C2 and one with cluster C4, including some of the type II pneumocyte-associated genes (Fig. 1). Importantly, cluster C2 was associated with inferior patient outcomes compared with non-C2 patients. In contrast, patients in cluster C4 had a more favorable outcome, corroborating that this cluster in fact recapitulates the histologic subtype of BACs. Thus, novel subtypes can be identified in lung adenocarcinomas by gene expression array analyses that are predictive of patient survival.

In another study investigating expression profiles of lung adenocarcinomas, predominantly early-stage disease, hierarchical clustering revealed three subgroups, one of which was highly enriched for BACs (20). Although the results in this third study also showed the same degree of heterogeneity in gene expression data among lung adenocarcinoma samples, they underline the notion that BACs show a particular and characteristic gene expression pattern. In the study by Garber et al. (19), hierarchical clustering of expression array data from lung cancer samples also recapitulated the histopathologic subtype. Interestingly, in this study, too, the adenocarcinoma samples formed subclusters, whereas squamous cell carcinomas and small-cell lung cancer samples clustered tightly together within their respective groups. Some degree of overlap in these subcategories with those identified by Bhattacharjee underline the reproducibility of the results across the two different experimental platforms used, namely, oligonucleotide arrays in the study by Bhattacharjee and cDNA arrays in the study by Garber.

A large collaborative study is now under way in which large new data sets on lung adenocarcinoma samples will be collected by the participating centers. These data sets will then be pooled and analyzed by a central group of biostatisticians. First, analyses show little variance between expression array data collected on a given sample by all participating centers (28). Additionally, hierarchical clustering done on these data reproducibly clustered the samples into identical groups, demonstrating that such an intergroup project is in fact feasible. The results from that endeavor will hopefully help to differentiate the different subtypes of adenocarcinoma, as determined by their gene expression signatures, on a more robust level and to identify transcriptional patterns reflecting particular mechanisms of oncogenic transformation in these subgroups.

Beer et al. (20) have identified a gene expression signature predictive of patients at high risk for poor overall survival. They built a 50-gene predictor using leave-one-out cross-validation to predict the survival of patients in a training set of patients from their data set ($n = 86$). When applying this predictor to a data set on 84 patients from an independent study (18), they were able to predict survival of lung adenocarcinoma patients with a high degree of statistical significance. Importantly, the predictor was able to separate patients with early-stage adenocarcinoma according to their risk. This latter result implies that, using gene expression data, patients at high risk for recurrence who might benefit from adjuvant therapy may be identified before treatment. Notably, in this study, some of the gene expression results were confirmed by Northern blot and immunohistochemistry, validating the results by microarray analysis and representing a first step to the development of a clinically feasible test for predicting survival. Hopefully, in the near future, the results of all gene expression studies in lung cancer will translate into a clinically widely applicable test (e.g., immunohistochemistry) to allow for early identification of patients at high risk.

One approach to modeling molecular circuitries in cancer cells is to identify the effect of gene mutations on gene expression patterns. This approach has the potential to increase the understanding of the molecular consequences of a given mutation, thereby identifying key mediators of oncogenic signaling. These molecules would then represent attractive therapeutic targets. Additionally, such an approach may lead to the detection of genes specifically up-regulated or down-regulated in the mutant cases. The evaluation of antibodies against the proteins encoded by those genes for use in immunohistochemistry may finally result in broadly applicable predictive tests. In the case of lung cancer, such a characteristic gene expression profile was recently identified in lung cancers harboring mutations in the KRAS proto-oncogene (29). Using expression array data from a KRAS-mutant mouse model of lung adenocarcinoma, the authors were able to infer a KRAS mutation signature by using the tool of gene set enrichment analysis. Transferring that signature to human adenocarcinomas correctly identified KRAS-mutant lung tumors. These results were confirmed by real-time PCR and, most importantly, by experimentally silencing the central genes involved in this signature using small interfering RNA technology. Given the current lack of a potent clinically available KRAS inhibitor, one intriguing possibility would be to find inhibitors against key mediators of mutant KRAS function. Additionally, it would be interesting to determine the gene expression signatures characteristic of the different EGFR gene mutations (discussed later in the text).

Taken together, gene expression array analyses have led to important insights into lung cancer, such as the existence of molecularly defined subclasses of lung adenocarcinomas. In an ongoing collaborative study, these subgroups will hopefully
Fig. 1. Gene expression clusters and histologic differentiation within lung adenocarcinoma subclases. Genes expressed at high levels in specific subsets of adenocarcinomas. A, colon metastases. B, proliferation-related gene expression (C1). C, neuroendocrine gene expression (C2). D, ornithine decarboxylase 1 and surfactant gene expression (C3 and C2). E, type II pneumocyte gene expression (C4, C3, and normal lung). F, histopathologic degree of differentiation (red, poor; yellow, moderate; green, well; white, not available or irrelevant). G, estimated nucleated tumor content (white, not determined or irrelevant; gray, 30-40%; blue, 40-70%; black, >70%). Reproduced with permission from Bhattacharjee et al. (18).
Genomic Approaches to Lung Cancer

Large-scale screens analyzing copy number changes and regions of LOH have made an immense effect on the study of lung cancer genomics. These studies give insight into patterns of genetic alteration, relative frequencies of certain genetic changes, and the possibility of identifying potentially “drug-gable” targets and diagnostic markers.

Cancer-specific copy number alterations and LOH represent important changes found in cancer cells. Amplified regions of the genome may include oncogenes, whereas deletions and regions of LOH may harbor tumor suppressor genes. Screens for copy number alteration and LOH in cancer cells have undergone large gains in resolution over the last few years. Work in cytogenetics gave some of the first insight into the role of large genetic changes in cancer cells. For example, the cytogenetic technique of comparative genomic hybridization (CGH) allows large copy number alterations to be reliably detected and quantified throughout the genome using metaphase spreads (30). A high-throughput, array-based version of this technique is array-CGH. These arrays use bacterial artificial chromosomes or cloned cDNAs to interrogate the changes in copy number across the genome at a higher resolution than conventional CGH (31, 32). Cancer-specific LOH has typically been identified using RFLPs or microsatellite analysis. These methods, which many consider the gold standard for LOH analysis, are difficult to use in large-scale studies. Although genome-wide microsatellite and RFLP analyses have been done, most studies have been limited by the low number of markers used (33–36).

Millions of SNP loci have now been identified by the completion of the human genome sequence, making them good markers for studying cancer genetics. Arrays interrogating SNP loci are commercially available in 10,000 and 100,000 SNP loci format (37, 38), with an array of 500,000 SNP loci being in development. These arrays contain 40 probes for each SNP locus, representing both mismatch and perfect match probes. SNP microarrays are very suitable for large-scale screens like array-CGH methods. Copy number has been inferred from the fluorescent intensity of hybridization of tumor DNA to these arrays (22, 39, 40). SNP arrays have successfully been used in genome-wide screens for detecting LOH in lung (21, 41) and other cancers (42–44). Importantly, the regions of LOH found by SNP array analysis also correlate with previous studies using microsatellite markers, thus demonstrating their accuracy (36, 41). The power of SNP arrays mainly lies in the fact that, unlike array-CGH, they allow for the simultaneous detection of copy number alterations, genotype, and when compared with a matched normal sample, LOH (22). This provides a distinction between copy-neutral LOH events and those LOH events with copy number alterations (22).

The high resolution of genomic maps produced by SNP array analysis provides an opportunity to detect small regions of copy number alterations. In lung cancer, the use of SNP array analyses has led to the detection of previously unknown regions of copy number change, as well as known regions of both amplification (e.g., Myc family members) and homozygous deletion (such as CDKN2A and PTEN; ref. 40). Novel amplifications of chromosomal regions 12p11 and 22q11 in non–small cell lung cancer (NSCLC) were identified as undergoing amplification using arrays interrogating 100,000 SNP loci (40). An amplified region on chromosome segment 8q12-13 in small-cell lung cancer was previously identified using 10,000 SNP arrays (22), and the borders of the region were narrowed through a study with 100,000 SNP arrays (40). Novel regions of recurrent homozygous deletion were found using SNP arrays at chromosomes 3q25 and 9p23 (40). These two regions contain very few genes; the 3q25 region contains arylacetamide deacetylase (AADAC) and succinate receptor 1 (SUCLCR1), whereas the 9p23 region contains protein tyrosine phosphatase, receptor type D (PTPRD) and an uncharacterized spliced transcript (BC028038) that shares some exons with PTPRD (40). The size of these regions reflects the capacity of SNP arrays to detect small regions of copy number alteration.

Although the amplifications found have often been of considerable size and are unlikely to benefit greatly from the increased resolution offered by the 500,000 SNP arrays, these novel arrays will be useful in identifying even smaller regions of heterozygous and homozygous deletion. This will enable the identification of regions down to the single gene or intron level. Changes to the current algorithm for obtaining copy number from signal intensity have been made to uncover allele-specific copy number for each SNP. This novel statistical approach has the ability to uncover false calls of LOH due to the high amplification of one allele over the other.

SNP array analyses also have the potential to reveal pathways disrupted in tumorigenesis; for instance, in lung adenocarcinoma, distinct members of the CDK4/CDK6 pathway may be subjected to copy number alterations in different tumors (40). When one or more additional genomic data sets are added, such as gene expression or mutational status, the potential to reveal tumorigenic pathways could be greatly increased.

Somatic Oncogene Mutations in Lung Cancer Detected by High-Throughput Resequencing

The availability of a reference version of the human genome in combination with the development of high-throughput capillary sequencers have made it possible to comparatively analyze a patient’s cancer and “normal” genomes to catalog the somatic genetic alterations in entire gene families of interest. Given that oncogenes typically activate specific signaling pathways that might be amenable to therapeutic interdiction, systematic resequencing approaches are frequently focused on this class of cancer genes. In a first systematic resequencing project, mutations in the BRAF proto-oncogene were identified in roughly 60% of malignant melanomas and some other cancers, including papillary thyroid cancers, colorectal cancers, and lung cancer (24, 45). Extending these analyses, this gene was found to be mutated in ~2% of primary adenocarcinomas

4 LaFramboise et al., unpublished results.
of the lung (26, 27). Importantly, BRAF mutations were exclusively found in tumors that did not harbor mutations in KRAS, a proto-oncogene that is mutated in 15% to 20% of NSCLC patients (20-30% of adenocarcinoma patients; ref. 46). Because the proteins encoded by these genes are located in the same pathway, one might speculate that a particular selection pressure on activating this pathway is active during lung tumorigenesis. Studies investigating the efficacy of inhibitors specifically targeting the mutationally activated pathways are under way.

In a subsequent study that aimed at systematically analyzing the family of phosphatidylinositol 3-kinase kinases, mutations in the PIK3CA gene encoding the p110α subunit of the phosphatidylinositol kinase were identified in a large fraction of colorectal adenocarcinomas but also in 1 of 24 NSCLC cases (23). The mutation found in the lung cancer sample, E545K, affects a highly conserved region in the helical domain of the protein. Further analysis of the most frequent PIK3CA mutants, among them the E545K, revealed that these mutations are in fact oncogenic (47). Because phosphatidylinositol kinase activates the protein kinase B/Akt pathway (48), it is conceivable that these mutations contribute to activation of this oncogenic pathway in NSCLC. Interestingly, activation of the Akt pathway has been shown in a substantial portion of primary NSCLC samples by immunohistochemistry (49, 50). Given the lack of mutations in the PTEN tumor suppressor gene, which negatively regulates phosphatidylinositol 3-kinase activity, in NSCLC (51), it can be speculated that in a subset of cases, PIK3CA mutations contribute to activation of the Akt pathway in NSCLC. It would be interesting to determine the actual prevalence of PIK3CA mutations in a larger panel of primary lung cancer samples and, also, if targeting the Akt pathway proves beneficial clinically in NSCLC.

**Somatic EGFR Mutations in NSCLC and their Therapeutic Implications**

In 2004, heterozygous mutations in the EGFR gene encoding the EGFR were detected in NSCLC by our group and others (15–17). These mutations were clustered around the kinase domain of the EGFR kinase and, most importantly, were strongly associated with clinical response to gefitinib and erlotinib, small-molecule kinase inhibitors targeting the EGFR kinase. Mutations mainly fell into three groups: the heterozygous substitution mutation G719S located in the glycine-rich nucleotide triphosphate-binding P loop, heterozygous small in-frame deletion mutations in exon 19 deleting the ELREA residues, and the heterozygous substitution mutation L858R that is located directly adjacent to the DFG motif of the activation loop. The residues affected by substitution mutations are all highly conserved within the protein kinases, pointing toward a role in malignant transformation. Numerous other reports have added to these initial findings and the respective approximate frequencies are as follows: exon 19 deletion mutations, 45%; L858R, 40%; in-frame duplications/insertions, 10%; other (including G719S), 5% (52–55). Mutations were more frequent in adenocarcinomas, in women, in never-smokers, and in patients with East Asian ethnicity (56). Functional experiments showed that the H3255 adenocarcinoma cell line harboring the L858R mutation was extremely sensitive to treatment with gefitinib, with an IC₅₀ of only 40 nmol/L, compared with wild-type cell lines that were growth inhibited only at concentrations that were 100-fold higher (16). Additionally, levels of phosphorylated EGFR, extracellular signal-regulated kinase 1/2, and Akt were strongly reduced by gefitinib treatment in H3255 cells but not in wild-type cells, implying a role for these proteins in mutant EGFR-dependent malignant transformation. Similarly, cells genetically engineered to express one of the deletion mutations or the L858R mutation showed enhanced EGFR phosphorylation upon addition of EGF as well as enhanced killing by gefitinib compared with cells expressing wild-type EGFR (15). In a subsequent analysis, Sordella et al. (57) showed that lung cancer–derived EGFR mutations activate the antiapoptotic Akt and signal transducers and activators of transcription pathways, providing a link between these mutations and malignant transformation. Also, by specifically silencing the mutant EGFR allele using small interfering RNA technology, they were able to show induction of apoptosis in EGFR-mutant but not wild-type cells and, thus, a dependency of EGFR mutants on this oncogene.

It is still an unresolved issue whether the mechanism by which the mutations confer response to EGFR TKIs is by changing the biochemical properties of the kinase, by targeting an activated kinase on which the cell has become strictly dependent, or both. Experiments from our own laboratory now formally show that the lung cancer–derived EGFR mutations are in fact inducing oncogenic transformation: Retroviral introduction of different substitution mutations, including G719S and L858R as well as the L747_E749del A750P and the D770_N771insNPG, into NIH-3T3 cells led to anchorage-independent growth, formation of multiple foci, and growth in a xenograft model (58). Furthermore, infection of human tracheobronchial epithelial cells with retroviruses expressing these mutations also led to oncogenic transformation. The mutants were found to induce constitutive, ligand-independent, autophosphorylation of EGFR and to lead to activation of Shc and signal transducers and activators of transcription 3. Finally, the different mutants were found to differentially confer sensitivity to EGFR TKIs, except for the insertion mutation, which is highly resistant. Thus, lung cancer–derived EGFR mutations induce oncogenic transformation by leading to constitutive kinase activity of EGFR and activation of downstream signaling pathways, and these mutations confer markedly different degrees of sensitivity to treatment with EGFR TKIs, with the substitution mutations and the deletions leading to increased sensitivity, and the insertion leading to resistance.

Although the frequency of EGFR gene mutations in NSCLC seems to be roughly 10% in Caucasian patient populations, in East Asian patients, the proportion is ~30% (56). Interestingly, these features match with the ones of patients who responded to EGFR TKI treatment in initial phase II trials (59–63). Sequencing the EGFR kinase domain in 31 responders to either gefitinib or erlotinib revealed mutations in 25 (81%), but in none of 29 nonresponders (9%; refs. 15–17). Although the patients analyzed in these studies were highly selected (i.e., patients showing progressive disease were analyzed in the control group but no patients experiencing disease stabilization), results from several additional studies support these preliminary findings. Most recently, it was found that EGFR
copy number changes, as determined by fluorescence in situ hybridization, and even EGFR protein expression may be better predictors of survival (64, 65). When looking at these data, one has to take several factors into considerations: A survival advantage may also exist for the patients experiencing stable disease, whereas EGFR mutations were found in patients experiencing major, occasionally, dramatic responses. Additionally, it was found that EGFR copy number changes are correlated with the presence of mutations (66). Because it may well be the wild-type allele that is amplified, one may easily imagine that the mutant allele might have escaped detection due to the limited sensitivity of Sanger dideoxy sequencing (67). Finally, in most cases, sequencing is being done on paraffin-embedded archival specimens. Because in most studies EGFR TKIs have been used at a very late clinical stage, the question remains whether these analyses actually represent the tumor cell clone that is being treated. Future prospective studies will help to determine which is the molecular alteration that is most stringently associated with response to EGFR TKIs.

Unfortunately, all NSCLC patients treated with the EGFR TKI gefitinib or erlotinib will eventually relapse and succumb to their tumor. In chronic myeloid leukemia, acquired resistance to imatinib was found to be mainly caused by the emergence of secondary resistance mutations in the BCR-ABL kinase domain. In some cases, these mutations were existent before treatment, indicating that they contribute to malignant growth and implying that continuous molecular monitoring of a patient’s tumor cell clone should help guiding the optimal therapy. Interestingly, one of the residues frequently mutated in chronic myeloid leukemia, T315, is conserved in EGFR (T790) and was shown crystallographically to bind erlotinib via a bridging water molecule (68). When the T790M mutation, analogous to the T315I mutation in chronic myeloid leukemia, was introduced into CHO-K1 cells, they became resistant to EGFR TKI treatment (69). More recently, the T790M mutation was detected in an NSCLC tumor specimen that was obtained at time of relapse following successful treatment with gefitinib, but not in the pretreatment specimen that contained an exon 19 deletion mutation of EGFR, indicating that T790M is in fact a clinically meaningful resistance mutation (70). Similar results were also reported by another group who found the T790M mutation in three of six patients with acquired or primary resistance to gefitinib or erlotinib (71). Irreversible inhibitors of the EGFR kinase could effectively kill cells carrying the T790M resistance mutation in concentrations that might be achievable in patients (70, 72, 73). One of these inhibitors is currently in clinical trials.

In our laboratory, we have now used ultradeep pyrosequencing in a blinded fashion to detect mutations in the EGFR kinase domain in patients with known mutational status. We were able to detect all mutations that had been previously detected by Sanger sequencing. However, we were also able to detect mutations in two samples that had been classified as unmutated by Sanger sequencing. These results encouraged us to sequence the EGFR kinase domain in DNA that had been extracted from the pleural effusion of a patient with fulminant relapse to EGFR TKI. The tumor content in such a clinical sample is very low. However, using ultradeep pyrosequencing, we were able to detect the T790M mutation in addition to an exon 19 deletion mutation. These results show that cancer gene mutation detection for both clinical and research purposes may now be feasible without the need for sample purification.

Taken together, the discovery of EGFR mutations causing clinical responses to EGFR TKIs, second-site mutations such as T790M causing relapse, and, finally, irreversible kinase inhibitors overcoming resistance are an impressive example of how high-throughput technologies can be successfully applied in collaborative projects involving clinicians, genome biologists, and basic scientists, finally leading to improvement in lung cancer patient care.

**Conclusions**

Great success has been achieved in the last decade in the field of lung cancer research that is intimately connected with the completion of the human genome project and the development of high-throughput genomic technologies. Examples like the recently discovered EGFR gene mutations in NSCLC show how these findings can directly translate into the clinic. With the development and use of additional technologies and by
combining analyses of different genomic data sets, more such discoveries can be made that may ultimately help in treating the patient suffering from lung cancer.

Open Discussion

The questions and discussion below follow from the oral presentation given at the Third Conference on Novel Agents in the Treatment of Lung Cancer and do not correspond directly to the written manuscript which is a more general review.

Dr. Daniel Haber: If you were to play the devil’s advocate in regard to this highly sensitive mutation detection analysis, you could say that this is very important if you are looking for mutation in mixed or heterogenous tumor samples. If, on the other hand, there are problems with either rare PCR artifacts from paraffin sections or even rare biological variation, like a rare cell that does have an EGFR mutation but isn’t driving the tumor, then you will be picking up these mutations with your analysis. It will look reproducible, but the biological significance will be unclear. So, are you getting into a situation where your ability to pick up such rare mutations may be difficult to interpret, and how do you deal with that?

Dr. Meyerson: The PCR artifact, in principle, is a stochastically rare artifact, because every PCR reaction is independent starting from a single DNA molecule. Unless you get the same artifact reproducibly many times, it shouldn’t be detected. The biological artifact would be the same. However, if we find in 5% of the cells that there is a EGFR mutation but it is completely nonrelevent, how do we sort that out? It’s a good question.

Dr. Haber: It is an interesting theme because there have been rare EGFR mutations in other tumors that don’t seem to be responsive to these drugs. You wonder, in some cases, is the mutation an early genetic event driving the cancer in some types of cancer, whereas in others is it a late event and less relevant?

Dr. Meyerson: I do think that other cancers have EGFR mutations that are not going to be clinically significant, so I think that some other mutation types may be found. If these other mutations are significant, maybe they will also be early genetic events.

Dr. Glennwood Goss: Are there subpopulations of wild-type EGFR that have increased sensitivity or increased resistance? The extrapolation from that is the clinical question. With the drugs that are currently available, if we escalated the doses, would we increase the response rate in the population that is sensitive to the drugs?

Dr. Meyerson: That is a really good question to present. Does the level of EGFR expression or the coexpression of other ErbB family members affect the sensitivity? We don’t have any really good data to address that.

Dr. Thomas Lynch: You really couldn’t do it in model system, could you?

Dr. Meyerson: I don’t think so. We can introduce other ErbB members, and that is something we are interested in doing. Potentially, we could try to select for cells with different expression levels. We have BA/F3 cells that are transformed with wild-type EGFR. Thus far, those show a lower IC50 for erlotinib and gefitinib than do the NIH-3T3 cells. We see an ~ 80 nmol/L IC50 in those cells. We are now trying to figure out if those results are genuine or not.

Dr. Paul Bunn: But in resistant cells that have wild-type receptor, many of them have lots of receptor, you are going to get toxicity in normal tissue before you get a concentration that can affect those cells. They are not resistant because of something with the receptor itself; it is because of the associated proteins.

Dr. Goss: That would be certainly true of the very resistant cells, but there may be a spectrum of sensitivity within the wild-type group.

Dr. Bunn: Right, there are some cells with wild-type receptors that are sensitive, but that is because of the expression of these related proteins, not because of the receptor itself. There is nothing different about the receptors. It is not a quantitative difference in amount of receptor either.

Dr. Alice Shaw: Do you have data on exon 20 insertion mutations and their effects on tumor behavior?

Dr. Meyerson: The transformed cell lines are much more resistant to erlotinib than gefitinib, but otherwise behave similarly as far as we can tell. There are differences in the phosphorylation pattern on the EGFR. In the deletion mutation, tyrosine 1045 phosphorylation, or tyrosine 1069 phosphorylation using the new numbering, is much higher than in the insertion mutant. That is the CBL binding site. And in fact, we find that the expression level of that receptor is significantly lower. We have not yet seen any other signaling differences, but we are trying to understand if there is a difference in signaling mechanisms. Finally, we have not seen any cancer-derived cell lines yet that have that mutation.

Dr. Alan Sandler: There was another pan irreversible inhibitor of the erbB family, CI 1033. Great for causing rashes, easy on the tumor. It was studied only in a small number of patients, but there were no responses. How would you explain the potential disparity with that agent? Has anybody looked and found that compound versus any of the Wyeth compounds?

Dr. Meyerson: Not yet, so I don’t know.

References

11. Cools J, DeAngelo DJ, Gotlib J, et al. A tyrosine kinase created by fusion of the PDGFRa and FIP1L1


Genomic Approaches to Lung Cancer

Roman K. Thomas, Barbara Weir and Matthew Meyerson

*Clin Cancer Res* 2006;12:4384s-4391s.