

# Identifying Cancer Genes—A Review

a report by

**Lynette M Sholl, MD** and **Lucian R Chirieac, MD**

*Department of Pathology, Brigham and Women's Hospital, Harvard Medical School*

DOI: 10.17925/OHR.2007.00.2.8

More than 350 cancer genes have been identified to date (see [www.sanger.ac.uk/genetics/CGP](http://www.sanger.ac.uk/genetics/CGP)) using a combination of epidemiological, genetic, and laboratory techniques. Decades of research into cancer genes have established some basic principles of tumorigenesis, including the roles of tumor suppressor genes and oncogenes, the loss and gain of which, whether through copy number changes or through inactivating or activating point mutations, can initiate or promote the evolution of cancer. Building on these principles, technologies that permit rapid sequencing or simultaneous analysis of copy number changes in hundreds of thousands of genes, coupled with advanced statistical analysis, are leading to rapid identification of new candidate cancer genes in a large spectrum of human tumors.

Many cancer genes have been identified through the use of positional cloning to characterize a recurrent chromosomal anomaly such as a translocation (t) or a chromosome gain or loss.<sup>1</sup> Other responsible genes, principally those associated with inherited cancer syndromes, have been uncovered through a combination of epidemiological studies and linkage analysis. Finally, a subset of cancer genes has been identified in the laboratory based on homology to other known genes and/or biological characteristics. High-throughput sequencing<sup>2</sup> and microarray<sup>3</sup> platforms that exploit the genetic changes (such as mutations and copy number loss or gain) involved in tumorigenesis promise to accelerate advances in this field.



Lynette M Sholl, MD, is a Senior Anatomic Pathology Resident at Brigham and Women's Hospital, Boston. Her interests include the application of molecular techniques to neoplastic and non-neoplastic pulmonary diagnostics and research. She completed a fellowship in molecular genetics pathology at the Harvard Medical School and an internship in internal medicine at the Hospital of the University of Pennsylvania. Dr Sholl received her MD from Stanford University School of Medicine in 2003.



Lucian R Chirieac, MD, is an Assistant Professor in the Department of Pathology at Brigham and Women's Hospital (BWH) and Harvard Medical School, Boston. He provides laboratory expertise for the Pulmonary Medical Oncology and General Thoracic Surgery Departments at BWH and Dana-Farber/Harvard Cancer Center. He is involved in establishing national guidelines for the diagnosis of lung cancer and assumes a senior role in the histopathological and immuno-phenotypic analysis of lung cancer. Dr Chirieac's clinical duties include acting as the reference pathologist for one of the largest academic pulmonary oncology services in the US.

E: [lchirieac@partners.org](mailto:lchirieac@partners.org)

## Chromosome Translocations and Tumorigenesis

Historically, cytogenetic observations have enabled cancer gene discovery. In 1956, the human karyotype was defined as containing 46 chromosomes, which were classified according to size and centromeric position.<sup>4</sup> Several years later, Peter Nowell and David Hungerford identified an abnormal chromosome in patients with chronic myeloid leukemia (CML),<sup>5</sup> which became known as the Philadelphia chromosome. With advances in chromosomal banding techniques, the Philadelphia chromosome was identified as chromosome 22 with a deletion on the long arm linked with chromosome 9 as its t partner. Genetic-mapping techniques identified the breakpoint cluster region (BCR) and Abelson murine leukemia virus homolog (ABL) as the fusion partners in the t(9;22) product.<sup>6</sup> The chimeric bcr-abl protein drives abl kinase activity with the functional consequence of increased myeloid proliferation. In this proliferative state, additional pathogenic mutations develop, eventually leading to transformation into an acute leukemia. In the absence of treatment, the disease inevitably progresses. However, the development of tyrosine kinase inhibitors, such as imatinib mesylate, has led to remarkable improvements in survival.<sup>7</sup>

Currently, CML is defined by the presence of the t(9;22), and the t has been shown to be both necessary and sufficient for disease.<sup>7</sup> Similar t events drive numerous hematological disorders (acute leukemias and lymphomas), sarcomas,<sup>8</sup> and few carcinomas.<sup>9</sup> Furthermore, it is increasingly recognized that the t defines the neoplasm, predicts the biological behavior and therapeutic response, and, as in the case of CML, even directs the type of therapy.<sup>10</sup>

In most cases, a cytogenetically visible chromosomal abnormality provides the first clue regarding a tumor's underlying genetic abnormality. However, cancer gene discovery by this method is limited by the resolution of chromosomal banding techniques, as only large structural rearrangements or large gains and losses of chromosomal material will be detected. Cryptic translocations (those involving the telomeres or t between chromosome material with similar banding patterns) or sub-megabase changes require alternative detection methods, including comparative genomic hybridization (CGH) and fluorescence *in situ* hybridization (FISH).<sup>4</sup>

## Inherited Cancer Syndromes

Many of the most common cancers, including breast and colon carcinomas, are clinically and pathologically heterogeneous and appear to be driven by a complex set of environmental and genetic interactions. Decades of clinical observation and epidemiological studies have identified a number of inherited cancer syndromes, most famously breast

cancer (BRCA)-related breast and ovarian cancer syndromes, hereditary non-polyposis colorectal carcinoma (HNPCC), and familial adenomatous polyposis (FAP). While these syndromes represent only a small fraction of cancer cases in the population, recognition of the causative mutations has provided significant and general insight into mechanisms of tumorigenesis and DNA repair.

Documentation of the hereditary nature of HNPCC came in the 19th century, when Dr Aldred Warthin described a family with multiple gastrointestinal cancers. Over 100 years later, epidemiological studies

Many of the most common cancers, including breast and colon carcinomas, are clinically and pathologically heterogeneous and appear to be driven by a complex set of environmental and genetic interactions.

confirmed that HNPCC was inherited in an autosomal dominant manner.<sup>11</sup> Genetic studies using microsatellite markers, which are short, usually non-coding DNA sequences containing tandem repeats of two to four nucleotides that are variable within the population and can serve as markers of inheritance, showed tight linkage of the disease with markers on the short arm of chromosome 2 (2p).<sup>12</sup> Researchers attempted to identify allelic loss on 2p (loss of heterozygosity (LOH)) as an explanation for tumorigenicity. While LOH was not seen, an interesting observation was made: the tumor cells contained multiple new microsatellite loci, which suggested impaired repair of repeated sequences. Subsequently, this process was termed microsatellite instability.

Concurrent investigations in yeast found mutated mismatch repair genes leading to microsatellite instability. These findings prompted a search for human homologs of the yeast mismatch repair genes, one of which was found on the short arm of 2p and is now known as MSH2. It is one of a handful of mismatch repair genes that can be mutated in HNPCC.<sup>13</sup> The familial nature of FAP was also recognized for decades before the first clues into the genetic basis of the disease were identified. In 1986, an interstitial deletion on the long arm of chromosome 5 (5q) was seen in a patient with FAP. Fine mapping at this locus ultimately led to the identification of the adenomatous polyposis coli (APC) gene. APC is now recognized as a pivotal tumor suppressor gene that is mutated early in the pathogenesis of most sporadic colorectal carcinomas.<sup>14</sup>

#### Identification of Cancer Genes *In Vitro*

Ras-Raf-MEK-ERK is one of the most commonly mutated pathways in human carcinomas. The first clues to this pathway's role in tumorigenesis came from *in vitro* biological assays in the early 1980s, when a genomic fragment derived from the DNA of a lung carcinoma cell line with homology to the Kirsten rat sarcoma virus was shown to transform fibroblasts.<sup>15</sup> Activating mutations in the KRAS gene have been found in up to 25% of human tumors and are instrumental in both tumor initiation and progression.<sup>16</sup> In addition, other members of the Ras-Raf-MEK-ERK pathway,

including NRAS, HRAS, and BRAF, have been implicated in tumorigenesis via their role in mediating proliferative and survival responses to growth factors. Germline mutations in members of this signaling pathway are associated with several congenital syndromes, including Noonan syndrome (PTPN11, KRAS) and cardio-facial-cutaneous syndrome (KRAS, BRAF, MEK). Studies of these syndromes are shedding light on the developmental function of these genes, with implications for our understanding of how disordered signaling through this pathway influences tumorigenesis.<sup>17</sup>

#### Oncogene Copy Number Gain in Tumorigenesis

Amplification of receptor tyrosine kinase genes is another known mechanism contributing to tumorigenesis. Human epidermal growth factor receptor 2 (HER2)/neu represents the best described and most commonly clinically exploited amplified cancer gene. Originally cloned in the mid-1980s in tumor cell lines, the gene was named for its sequence homology to the HER2.<sup>18</sup> Subsequently, it has been shown to be amplified in a subset of gastric, lung, and salivary carcinomas, but is most commonly amplified in breast carcinoma. In up to 30% of breast cancer patients, HER2/neu amplification is present and is associated with more aggressive disease. A recombinant monoclonal antibody to the HER2 protein product trastuzumab has been shown to improve outcomes in patients with HER2/neu amplification, which can be detected via immunodetection of the overexpressed protein product or via *in situ* hybridization for the amplified gene.<sup>19</sup>

This model has been extended to other tumors as well. Mutations in the kinase domain and/or amplifications of the epidermal growth factor receptor (EGFR) in non-small-cell lung cancer predict response to tyrosine kinase inhibitors (gefitinib, erlotinib), which have proven efficacy in a subgroup of patients, typically non-smoking women of Asian descent.<sup>20</sup> Histopathological classification is of somewhat limited utility in identifying potential responders; rather, sequencing of the EGFR kinase domain and *in situ* hybridization, via either a fluorescence or chromogenic approach, promise to serve as important ancillary studies in lung cancer diagnosis.<sup>21,22</sup>

#### High-throughput Approaches to Cancer Gene Discovery

Currently, systematic analysis of cancer genomes by high-throughput methods is driving the discovery of important actors in tumorigenesis.

Human epidermal growth factor receptor 2/neu represents the best described and most commonly clinically exploited amplified cancer gene.

Typical approaches include comparative genomic hybridization with single nucleotide polymorphism (SNP) or bacterial artificial chromosome (BAC) microarrays. These methods permit detection of deletions or duplications ('copy number changes') in tumors compared with normal DNA at hundreds of thousands of loci. For instance, a recent study of nearly 250 lung adenocarcinoma samples examined copy number changes at over 238,000 SNP loci. Statistical analysis revealed over 50 recurrent copy

number changes, both large-scale (gain or loss of over half of a chromosome arm) and focal changes (several hundred kilobases to several megabases). Further analysis of these sites confirmed previous observations of loss of tumor suppressor genes such as cyclin-dependent kinase inhibitors (CDKN2A/CDKN2B), and gain of proto-oncogenes (KRAS, EGFR, and ERBB2) in lung adenocarcinoma. More importantly, this large-scale analysis identified new regions of frequent genomic loss and gain, with one promising, recurrent amplification occurring at 14q13.3, a locus that contains a gene encoding transcription factor critical to type II pneumocyte formation.<sup>22</sup>

Systematic sequencing of the 'cancer genome' can identify recurrent mutations from which new cancer genes can be inferred. While cytogenetic and comparative genomic analysis can effectively identify changes on the scale of hundreds of kilobases or larger, pathogenic point mutations cannot be detected by these methods. High-throughput sequencing is starting to elucidate the mutational signature of various human cancers. In a study that sequenced over 500 protein kinase genes encompassing 270 megabases of DNA from 210 different human cancers, numerous recurrent mutations were identified. Mutations that

lead to a change in the amino acid sequence and, in particular, those that are located in the kinase domain are predicted to be the most likely to lead to gene activation and thus drive abnormal protein kinase activity. Using mathematical algorithms to predict pathogenic mutations, the authors were able to detect well-established cancer-related protein kinases, as well as new candidate cancer genes that will require biological validation but may be promising new targets for protein kinase inhibitors in cancer therapy.<sup>23</sup>

## Summary

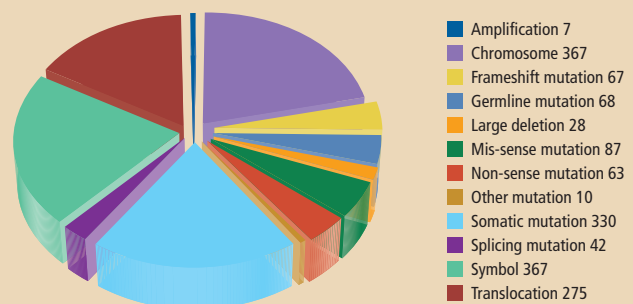
Since its inception in the mid-20th century, the field of cancer genetics has identified numerous oncogenes and tumor suppressor genes that have been implicated in tumorigenesis in many different body systems. Traditionally, the process of cancer gene identification has generally been laborious, frequently requiring decades of collaboration between clinicians, pathologists, geneticists, epidemiologists, and basic scientists. In light of improvements in cytogenetic, molecular, and computational technologies, investigators in the field today appear increasingly poised to more quickly identify candidate cancer genes and to translate these findings into promising new therapeutic targets. ■

1. K Fair, Le Beau MM, Characterization of translocations in human cancer. In: *Methods in molecular biology*, Totowa: Humana Press, 2003;237.
2. Greenman C, Stephens P, Smith R, et al., Patterns of somatic mutation in human cancer genomes, *Nature*, 2007;446:153–8.
3. Pollack JR, Perou CM, Alizadeh AA, et al., Genome-wide analysis of DNA copy-number changes using cDNA microarrays, *Nat Genet*, 1999;23:41–6.
4. Dal Cin P, Morton CC, Cytogenetics for the Hematologist. In: Handin RI, Lux SE, Stossel TP (eds), *Blood: Principles and Practice of Hematology*, Philadelphia: Lippincott, 2003;96.
5. Nowell PC, Hungerford DA, Chromosome studies on normal and leukemic human leukocytes, *J Natl Cancer Inst*, 1960;25:85–109.
6. Tefferi A, The history of myeloproliferative disorders: before and after Dameshek, *Leukemia: official journal of the Leukemia Society of America*, Leukemia Research Fund, UK, 2007.
7. Jaffe ES, Harris NL, Stein H, Vardiman JW, *Tumours of Haematopoietic and Lymphoid Tissues*, Lyon: IARC Press, 2001.
8. van de Rijn M, Fletcher JA, Genetics of soft tissue tumors, *Annu Rev Pathol*, 2006;1:435–66.
9. Argani P, Ladanyi M, Translocation carcinomas of the kidney, *Clin Lab Med*, 2005;25:363–78.
10. Druker BJ, Talpaz M, Resta DJ, et al., Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia, *N Engl J Med*, 2001;344:1031–7.
11. Lynch HT, de la Chapelle A, Hereditary colorectal cancer, *N Engl J Med*, 2003;348:919–32.
12. Lindblom A, Tannergard P, Werelius B, et al., Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer, *Nat Genet*, 1993;5:279–82.
13. Kinzler KW, Vogelstein B, Lessons from hereditary colorectal cancer, *Cell*, 1996;87:159–70.
14. Aoki K, Taketo MM, Adenomatous polyposis coli (APC): a multifunctional tumor suppressor gene, *J Cell Sci*, 2007;120:3327–35.
15. Der CJ, Krontiris TG, Cooper GM, Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and Kirsten sarcoma viruses, *Proc Natl Acad Sci U S A*, 1982;79:3637–40.
16. Kranenburg O, The KRAS oncogene: past, present, and future, *Biochim Biophys Acta*, 2005;1756:81–2.
17. Schubert S, Shannon K, Bollag G, Hyperactive Ras in developmental disorders and cancer, *Nature reviews, Cancer*, 2007;7:295–308.
18. Coussens L, Yang-Feng TL, Liao YC, et al., Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene, *Science*, 1985;230:1132–9.
19. Slamon DJ, Leyland-Jones B, Shak S, et al., Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2, *N Engl J Med*, 2001;344:783–92.
20. Sharma SV, Bell DW, Settleman J, et al., Epidermal growth factor receptor mutations in lung cancer, *Nature reviews, Cancer*, 2007;7:169–81.
21. Cappuzzo F, Hirsch FR, Rossi E, et al., Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer, *J Natl Cancer Inst*, 2005;97:643–55.
22. Sholl LM, John lafrate A, Chou YP, et al., Validation of chromogenic *in situ* hybridization for detection of EGFR copy number amplification in nonsmall cell lung carcinoma, *Mod Pathol*, 2007;20:1028–35.
23. Weir BA, Woo MS, Getz G, et al., Characterizing the cancer genome in lung adenocarcinoma, *Nature*, 2007;450(7171):893–8.

## Cancer Genetics—At a Glance

- Currently, more than 1% of all human genes are implicated via mutation in cancer.
- Of every 100 women who carry the faulty version of the BRCA2 breast cancer gene, between 80 and 85% will develop breast cancer at some point in their life.
- By removing the gene SATB1 from cancer cells, researchers at Lawrence Berkeley National Laboratory in California have been able to halt the progress of cancer cells and even return the cell to a non-cancerous state.
- A mutation of the K-ras gene is present in approximately 40% of cancer patients.


### Cancer Gene Census—The Causes of Gene Mutation




Sources: *Cancer Gene Census*, Wellcome Trust Sanger Institute, Cancer Research UK, Daily Telegraph (London), March 12, 2008, *Laboratory Talk*.

The American Society for


# Clinical Laboratory Science



ASCLS serves as the voice of all clinical laboratory professionals, creating a vision for the advancement of the clinical laboratory practice field, and advocating the value and the role of the profession ensuring safe, effective, efficient, equitable, and patient centered health care.



## Benefits of Membership include:

- Monthly newsletter
  - Reduced rates for ASCLS and constituent society continuing education events
  - Member savings at the ASCLS Online Store
  - Local, state and national network of colleagues
  - Scientific Assembly listserv
  - Stay informed on professional issues
- 



For more information:  
Visit the ASCLS website at  
[www.ascls.org](http://www.ascls.org)

or call the ASCLS office at  
**301-657-2768**