Somatic alterations in the human cancer genome

Barbara Weir, Xiaojun Zhao, and Matthew Meyerson*

Department of Medical Oncology
Dana-Farber Cancer Institute
Department of Pathology
Harvard Medical School
44 Binney Street
Boston, Massachusetts 02115
*Correspondence: matthew_meyerson@dfci.harvard.edu

Most human malignancies are caused by somatic alterations within the cancer genome, leading to oncogene activation or tumor suppressor gene inactivation. The sequence of the human genome has enabled systematic approaches to identify cancer genome alterations, including point mutations, copy number increases and decreases, loss of allelic heterozygosity, and chromosome translocations. Systematic cancer genome analysis has recently led to the discovery of somatic mutations in the *BRAF*, *PIK3CA*, and *EGFR* genes, among others. With further development of targeted cancer therapies and improvement in genome analysis technology, genome-wide surveys of cancer will likely become tools for diagnosis as well as discovery.

Introduction

Cancer is caused by alterations in the control and activity of genes that in turn regulate cell growth and differentiation, leading to abnormal cell proliferation. These "cancer-related genes" fall into two major classes that have opposite effects on normal cell proliferation and opposite modes of alteration in cancer cells. Tumor suppressor genes normally repress cell growth and are inactivated in cancer, while oncogenes, which normally stimulate cell growth, become hyperactivated in cancer (Weinberg, 1996).

Specific changes in the cancer genome sequence lead to disregulation of cancer-related genes. Among these changes are germline variations that lead to hereditary cancer predispositions, the acquisition of transforming DNA or RNA sequences from cancer viruses, and somatic changes in the cancer genome. Each of these mechanisms may lead to the activation of oncogenes or the inactivation of tumor suppressor genes. Epigenetic mechanisms, such as DNA methylation or histone modification, have also been postulated to promote oncogenesis by modifying the activity of cancer-related genes, but the evidence for this mechanism, although substantial, is less definitive to date.

Among these mechanisms, somatic genomic alteration appears to be the major causative factor in most human cancers. Early in the last century, Boveri first proposed that somatic changes in chromosomes could lead to cancer (Knudson, 2000, 2002). The first physical reification of this concept was the demonstration that chronic myelogenous leukemia (CML) is consistently associated with a recurrent chromosomal translocation, the Philadelphia chromosome, as shown by Nowell and Hungerford in 1960 (Nowell and Hungerford, 1960).

Remarkably, research building on the discovery of the Philadelphia chromosome, spanning the past four decades, has led to the development of one of the first effective targeted therapies for cancer: the use of the tyrosine kinase inhibitor imatinib or Gleevec to treat CML. The major milestones in this work were the findings that the Philadelphia chromosome represents a translocation between chromosomes 9 and 22 (Rowley, 1973a, 1973b), that this translocation leads to a fusion between the Bcr gene and the Abl tyrosine kinase gene (reviewed in Daley and Ben-Neriah, 1991), that imatinib can block the growth of CML cells in model systems by inhibiting Abl kinase activity (Druker et al., 1996), and finally that imatinib treatment of CML leads to remission in the overwhelming majority of patients (Druker et al., 2001). The path from the Philadelphia chromosome to imatinib treatment, as well as the similar path from the discovery of the *neu* oncogene to the development of therapy with trastuzumab or Herceptin (Akiyama et al., 1986; Bargmann et al., 1986; Coussens et al., 1985; King et al., 1985; Slamon et al., 2001), provides a new paradigm for cancer treatment by targeting the somatic molecular alterations that cause cancer.

A recent review has thoroughly summarized the range of known genomic alterations leading to cancer gene activation and inactivation (Futreal et al., 2004). This review will focus on the opportunities for cancer gene discovery based on the recent sequencing of the human genome and discuss some of the recent findings based on whole-genome analysis.

Cancer gene discovery by genome-wide screens prior to the completion of the human genome sequence

Before the recent determination of the complete human genome sequence, a variety of systematic "genome-wide" approaches led to the discovery of most of the important known cancer genes. These include cancer cell cytogenetics, oncogene transfection, mapping of genes that cause familial cancer syndromes, and genome-wide searches for allelic loss of heterozygosity and homozygous deletions.

The first evidence of cancer causation by genomic changes was the discovery of chromosome translocations in cancer, beginning with the Philadelphia chromosome that is pathogenic for CML. Using the techniques of cytogenetics, most notably chromosome banding, probably the first method for genome-wide exploration of cancer and other diseases, numerous oncogenes activated by chromosome translocation have been identified, as described in recent reviews (Mitelman, 2000; Mitelman et al., 1997; Rabbitts, 1994). Another early fruitful approach for oncogene discovery in cancer was the identification of genes isolated from cancer cells that could transform tissue culture cells in vitro. While oncogene transformation assays are not based on knowledge of the genome sequence, they

are nevertheless unbiased, genome-wide screens. These approaches have led to the identification of many cellular oncogenes, including the human homologs of *myc*, *ras*, *abl*, *neu*, their relatives, and others. The identification of *Her2-Neu* or *ERBB2* (Bargmann et al., 1986) by this approach is particularly important as it has helped lead to the development of the targeted breast cancer therapy trastuzumab (Herceptin) (Slamon et al., 2001). While the pace of oncogene discovery using transfection approaches has slowed dramatically, it is unclear whether new ways of library construction and gene delivery may be useful for this purpose.

The discovery of tumor suppressor genes has been mainly achieved by the study of familial cancer syndromes and by mapping of regions of cancer-specific allelic loss of heterozygosity. Here, one of the key developments was the production of genome maps, in particular by the use of restriction fragment length polymorphisms (RFLPs), making it possible to map genetic disorders systematically. This was first applied to a variety of X-linked disorders such as chronic granulomatous disease and Duchenne's muscular dystrophy. In the cancer field, the recognition of tumor suppressor genes underlying hereditary disorders began with the cloning of the retinoblastoma tumor suppressor gene, accomplished by a combination of genetic linkage studies using RFLPs and mapping of homozygous deletions (Cavenee et al., 1983; Friend et al., 1986). Another example was the cloning of the adenomatosis polyposis coli (*APC*) tumor suppressor gene through a combination of linkage analysis, deletion mapping, and mutation discovery in individuals with familial adenomatous polyposis, a disease which often leads to colorectal cancer (reviewed in Fearnhead et al., 2001).

The discovery of the first tumor suppressor gene, *RB1* (Friend et al., 1986), confirmed the Knudson "two-hit" hypothesis (Knudson, 1971), which states that tumor suppressor genes are inactivated by a recessive mutation in one allele followed by the loss of the other wild-type allele, a phenomenon termed loss of heterozygosity (LOH). By using initially RFLPs, then microsatellite markers, and then single-nucleotide polymorphisms, it has been possible to generate genome-wide maps of cancer LOH. These LOH maps have proven critical in localizing regions that are commonly deleted in cancer, such as chromosome 9p and chromosome 18q. These regions have subsequently been shown to harbor tumor suppressor genes, *p16INK4a* on chromosome 9 (Kamb et al., 1994; Nobori et al., 1994) and *SMAD4*/*DPC4* on chromosome 18 (Hahn et al., 1996). Comparative genomic hybridization studies have also identified many common regions of loss in cancers, pointing to potential tumor suppressor gene loci (reviewed in Knuutila et al., 1999). Cytogenetic and LOH studies also identified a potential tumor suppressor gene locus at 10q23 in some cancers, including prostate and glioblastoma (reviewed in Dahia, 2000). Representational difference analysis (RDA) and deletion mapping identified a candidate tumor suppressor gene, *PTEN*, from the areas of common deletion at this locus (reviewed in Dahia, 2000). RDA is a method that subtracts genomic representations of tumor from normal DNA to identify regions of gene deletion using multiple enrichment rounds of hybridization and PCR; the inverse subtraction of normal from tumor DNA representations is used to identify tumor viruses and high-level amplifications (Lisitsyn et al., 1995).

The amplification of chromosomal regions is a frequent way of increasing gene expression in cancer cells. Commonly amplified regions have been found in many cancer types through cytogenetic methods such as comparative genomic hybridization (reviewed in Knuutila et al., 1998). These regions often point to the location of potential oncogenes.

Genomic instability is often a driving force behind changes in the cancer genome. Mutations in nonhereditary polyposis colon cancer as well as xeroderma pigmotensum, Fanconi anemia, and ataxia telangiectasia, among others, lead to defects in DNA repair (D'Andrea and Grompe, 2003; Hoeijmakers, 2001; Lengauer et al., 1998). These in turn may cause mutations in oncogenes or tumor suppressor genes; for example, cancer cell lines with high rates of microsatellite instability have been found to contain presumed inactivating mutations in the gene for TGF β receptor type II (Markowitz et al., 1995). Whether similar mechanisms underlie the development of mutations in sporadic cancers remains unclear.

Cancer molecular pathogenesis after the human genome sequence

The draft sequencing of the complete human genome (Lander et al., 2001; Venter et al., 2001) now makes it possible to query the cancer genome systematically in ways that were hitherto impossible. The two major advances are the opportunity to look throughout the genome at the full range of copy number and allelic changes in cancer compared to the germline and the ability to survey particular gene families (e.g., kinases, phosphatases, G protein-coupled receptors, etc.) in a complete manner for somatic mutations.

Systematic copy number analysis

Oncogene activation in cancer is often a consequence of chromosomal copy number amplification, while tumor suppressor gene inactivation is often caused by either hemizygous deletion associated with mutation or by homozygous deletion. Thus, the identification of copy number alterations is a powerful tool for cancer gene discovery. With the mapping of the genome and the identification of large numbers of expressed genes, it became possible to explore the full range of genome copy number alterations by hybridization to arrays of bacterial artificial chromosomes (BACs) or to cloned cDNAs (reviewed in Albertson and Pinkel, 2003).

Higher-density tools for genome copy number analysis have become available with the completion of the human genome sequence. These include genome tiling BAC arrays (Ishkanian et al., 2004), oligonucleotide arrays for CGH (Brennan et al., 2004; Lucito et al., 2003), copy number analysis using SNP microarrays (Bignell et al., 2004; Zhao et al., 2004), and digital karyotyping (Wang et al., 2002). The resulting quantitative measurements of DNA copy number provide precise mapping information for amplicons and regions harboring homozygous deletions throughout the genome.

Digital karyotyping, a genomic version of the serial analysis of gene expression (SAGE) technique, uses the quantification of short, unique tags along the genome to assess copy number changes through sequencing (Wang et al., 2002). Recently, a new version of digital karyotyping has been developed using tags released from type IIB restriction enzymes, which is promising for the analysis of paraffin-embedded clinical samples (Tengs et al., 2004).

Finally, a new sequence-based method allows the simultaneous detection of complex structural changes including translocations at the same time as copy number analysis, by sequencing BAC ends and aligning them along the known

genomic sequence (Volik et al., 2003). The applications of this promising approach are yet to be developed.

Systematic LOH analysis

Traditionally, polymorphic markers, such as RFLPs and microsatellites, have been used to detect LOH through allelotyping the DNA from a cancer sample and a corresponding normal sample (Vogelstein et al., 1989). However, these assays are difficult to automate and are not readily scalable. Although unbiased genome-wide analyses have also been performed, most studies have used only a limited number of markers. The completion of the human genome has allowed for the identification of millions of SNP loci, making them ideal markers for genetic analysis. Due to their frequency in the genome and ease to automate using an array platform (Wang et al., 1998), they provide an attractive method to analyze LOH at a genome-wide level (Lindblad-Toh et al., 2000; Mei et al., 2000). LOH patterns generated by SNP array have a high degree of concordance with previous microsatellite analyses of the same cancer samples (Lindblad-Toh et al., 2000). In addition, shared regions of LOH from SNP arrays are able to cluster lung cancer samples into subtypes (Janne et al., 2004), and distinct patterns of LOH are found to associate with clinical features in primary breast, bladder, and prostate tumors (Hoque et al., 2003; Lieberfarb et al., 2003; Wang et al., 2004b). The hybridization intensity of DNA from tumor samples to high-density SNP arrays can also be used to infer copy number with a high degree of accuracy (Bignell et al., 2004; Zhao et al., 2004). Using SNP arrays, copy number analysis can be performed in combination with LOH analysis to distinguish copy number gains, copy neutral loss of heterozygosity, and copy number losses, providing a comprehensive map of genetic alterations within a cancer cell (Zhao et al., 2004).

Expression analysis and gene target discovery

Recent gene expression profiling revealed that lymphoblastic leukemias with MLL translocation had a unique expression signature and can be separated from conventional acute lymphoblastic and acute myelogenous leukemias. Among the features of the MLL group was high expression of the receptor tyrosine kinase gene *FLT3* (Armstrong et al., 2002). Subsequent studies have shown that *FLT3* is frequently mutated in MLL, and evidence suggests that the inhibitor PKC412 is active against MLL cells in xenografts (Armstrong et al., 2003).

Expression analysis has also been combined with copy number analyses, such as fluorescence in situ hybridization (FISH), array CGH, and other methods, in an effort to elucidate the critical target genes within DNA amplicons. Several array CGH studies have shown a genome-wide correlation of gene expression with copy number alterations and have proved useful in individual amplicon refinement (Pollack et al., 2002; Wolf et al., 2004). For example, through tissue microarray FISH and RT-PCR, a minimally amplified region around *ERBB2* was identified in a large number of breast tumors; in addition, gene amplification was found to be correlated with increased gene expression in a subset of those genes (Kauraniemi et al., 2003).

Cancer genome sequencing by gene family: kinases, phosphatases, PI3-kinases

The completion of the human genome sequence allows for the direct identification of mutations involved in cancer, by resequencing DNA from tumor tissue. Mutations in cancer-causing genes are increasingly being identified, and these searches can be focused on specific protein families. In all cases, functional analyses are needed to distinguish noncausative passenger mutations from those with true roles in tumorigenesis. Several stories are emerging from the identification of these mutations and, in some cases, their correlation with clinical features. Such studies have led to new insights into cancer pathogenesis with immediate and long-term clinical implications.

By using denaturing capillary gel electrophoresis to detect mutations, activating mutations in the *BRAF* kinase gene were discovered in a variety of malignancies (Davies et al., 2002). Notably, the *BRAF* mutations occur in over 60% of melanomas, suggesting that *BRAF* could be a therapeutic target for the treatment of melanoma and other cancers (Davies et al., 2002; Tuveson et al., 2003). Activating *BRAF* mutations have also been found in other cancers, including papillary thyroid cancers, colorectal cancers, and primary lung adenocarcinomas (Tuveson et al., 2003). *BRAF* mutations are generally nonoverlapping with mutations in *KRAS*, a member of the same pathway frequently activated in human cancers (Brose et al., 2002; Naoki et al., 2002; Rajagopalan et al., 2002). This suggests that only one mutational event in this pathway may be needed for tumorigenesis.

Advances in sequencing technology now make it possible to perform exon resequencing for gene families involved in cellular signaling pathways, such as tyrosine kinases, tyrosine phosphatases, and phosphatidylinositol 3-kinases (Bardelli et al., 2003; Lynch et al., 2004; Paez et al., 2004; Samuels et al., 2004; Stephens et al., 2004; Wang et al., 2004a). These studies have identified cancer-specific somatic mutations in several tyrosine kinase and tyrosine phosphatase genes as well as in one PI3 kinase gene (*PIK3CA*), which encodes the p110α catalytic subunit of phosphatidylinositol 3-kinase.

Somatic mutations within the *PIK3CA* gene were located in regions where mutations are predicted to increase its catalytic activity and were found in a significant fraction of colorectal cancers and glioblastomas (Samuels et al., 2004) as well as breast carcinomas (Bachman et al., 2004). The region harboring *PIK3CA* is often amplified in many different types of cancers and is both amplified and overexpressed in cervical and ovarian cancers, consistent with a likely role as an oncogene (Ma et al., 2000; Shayesteh et al., 1999).

Systematic sequencing of protein kinase genes in colorectal cancers led to the identification of somatic mutations in the *NTRK3*, *FES*, *KDR*, *EPHA3*, *NTRK2 MLK4*, and *GUCY2F* genes, with a high prevalence of mutations found in the kinase domain (Bardelli et al., 2003). Followup studies of these kinase genes in colorectal and other cancers will determine whether they play a causative role in tumorigenesis.

In non-small cell lung carcinoma, the most common cause of cancer death in the United States and worldwide, systematic exon resequencing of tyrosine kinase genes identified mutations within the epidermal growth factor receptor (*EGFR*) tyrosine kinase gene (Paez et al., 2004). The frequencies of these mutations within lung adenocarcinomas vary between samples from Japan (30%) and the United States (15%). The reason for this difference between distinct populations is unknown but is consistent with the concept that molecular mechanisms of oncogenesis may vary among such groups.

The finding of mutations in *EGFR* in lung adenocarcinoma is of significant clinical interest, because the compound gefitinib (Iressa), an EGFR kinase inhibitor, has shown significant activity in the treatment of lung adenocarcinoma patients, most notably in patients from Japan, nonsmokers, and women (Miller et al., 2004). Three groups, using different approaches, found

that mutations in the *EGFR* kinase domain in non-small cell carcinoma specimens correlated closely with patient responses to gefitinib (Lynch et al., 2004; Paez et al., 2004; Pao et al., 2004). Two studies approached *EGFR* mutation in a hypothesis-driven manner (Pao et al., 2004; Lynch et al., 2004), while another study came from a genome-scale effort to sequence all tyrosine kinase genes in cancer, as described above (Paez et al., 2004). Clinical testing for *EGFR* mutation in lung carcinoma is now available and may aid in the selection of patients for therapy with gefitinib and with related agents such as erlotinib (Tarceva), for which clinical response is also related to *EGFR* mutation (Pao et al., 2004).

Another recent study found kinase domain mutations of *ERBB2* in 4% of primary lung carcinomas and 10% of lung adenocarcinomas (Stephens et al., 2004). The relationship between these mutations and susceptibility to kinase inhibitors has not yet been described.

Finally, predicted inactivating mutations within six protein tyrosine phosphatases genes, *PTPRF*, *PTPRG*, *PTPRT*, *PTPN3*, *PTPN13*, and *PTPN14*, have been identified by systematic exon resequencing of these genes from colorectal cancers (Wang et al., 2004a). Biochemical studies of the most commonly altered gene, *PTPRT*, suggest that it is likely to function as a tumor suppressor gene (Wang et al., 2004a). In addition to colorectal cancer, some of these mutations have been found in other cancer types, and these mutations cluster in evolutionarily conserved domains, giving more evidence that they play a functional role in tumorigenesis (Wang et al., 2004a).

In summary, the approach of systematic exon resequencing has begun to identify a wide variety of somatic mutations in cancers, many of which have significant therapeutic implications. It is likely that many more cancer-causing mutations will be identified via targeted genome resequencing, in addition to those somatic alterations identified using complementary approaches such as systematic copy number, LOH, and expression analyses. As the power of sequencing technologies continues to increase with the development of novel high-throughput approaches (Brenner et al., 2000; Mitra et al., 2003; Shendure et al., 2004), the scale of cancer genome sequencing should increase proportionally. One can envision a future in which the sequencing of cancer-specific mutations will take its place alongside microscopic examination of cancer histology as the mainstay of cancer diagnosis.

References

Akiyama, T., Sudo, C., Ogawara, H., Toyoshima, K., and Yamamoto, T. (1986). The product of the human c-erbB-2 gene: a 185-kilodalton glycoprotein with tyrosine kinase activity. Science *232*, 1644–1646.

Albertson, D.G., and Pinkel, D. (2003). Genomic microarrays in human genetic disease and cancer. Hum. Mol. Genet. *12*, R145–R152.

Armstrong, S.A., Staunton, J.E., Silverman, L.B., Pieters, R., den Boer, M.L., Minden, M.D., Sallan, S.E., Lander, E.S., Golub, T.R., and Korsmeyer, S.J. (2002). MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat. Genet. *30*, 41–47.

Armstrong, S.A., Kung, A.L., Mabon, M.E., Silverman, L.B., Stam, R.W., Den Boer, M.L., Pieters, R., Kersey, J.H., Sallan, S.E., Fletcher, J.A., et al. (2003). Inhibition of FLT3 in MLL. Validation of a therapeutic target identified by gene expression based classification. Cancer Cell *3*, 173–183.

Bachman, K.E., Argani, P., Samuels, Y., Silliman, N., Ptak, J., Szabo, S., Konishi, H., Karakas, B., Blair, B.G., Lin, C., et al. (2004). The PIK3CA gene is mutated with high frequency in human breast cancers. Cancer Biol. Ther. *3*, 772–775.

Bardelli, A., Parsons, D.W., Silliman, N., Ptak, J., Szabo, S., Saha, S., Markowitz, S., Willson, J.K., Parmigiani, G., Kinzler, K.W., et al. (2003). Mutational analysis of the tyrosine kinome in colorectal cancers. Science *300*, 949.

Bargmann, C.I., Hung, M.C., and Weinberg, R.A. (1986). The neu oncogene encodes an epidermal growth factor receptor-related protein. Nature *319*, 226–230.

Bignell, G.R., Huang, J., Greshock, J., Watt, S., Butler, A., West, S., Grigorova, M., Jones, K.W., Wei, W., Stratton, M.R., et al. (2004). High-resolution analysis of DNA copy number using oligonucleotide microarrays. Genome Res. *14*, 287–295.

Brennan, C., Zhang, Y., Leo, C., Feng, B., Cauwels, C., Aguirre, A.J., Kim, M., Protopopov, A., and Chin, L. (2004). High-resolution global profiling of genomic alterations with long oligonucleotide microarray. Cancer Res. *64*, 4744–4748.

Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D.H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M., et al. (2000). Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat. Biotechnol. *18*, 630–634.

Brose, M.S., Volpe, P., Feldman, M., Kumar, M., Rishi, I., Gerrero, R., Einhorn, E., Herlyn, M., Minna, J., Nicholson, A., et al. (2002). BRAF and RAS mutations in human lung cancer and melanoma. Cancer Res. *62*, 6997–7000.

Cavenee, W.K., Dryja, T.P., Phillips, R.A., Benedict, W.F., Godbout, R., Gallie, B.L., Murphree, A.L., Strong, L.C., and White, R.L. (1983). Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. Nature *305*, 779–784.

Coussens, L., Yang-Feng, T.L., Liao, Y.C., Chen, E., Gray, A., McGrath, J., Seeburg, P.H., Libermann, T.A., Schlessinger, J., and Francke, U. (1985). Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene. Science *230*, 1132–1139.

Dahia, P.L. (2000). PTEN, a unique tumor suppressor gene. Endocr. Relat. Cancer *7*, 115–129.

Daley, G.Q., and Ben-Neriah, Y. (1991). Implicating the bcr/abl gene in the pathogenesis of Philadelphia chromosome-positive human leukemia. Adv. Cancer Res. *57*, 151–184.

D'Andrea, A.D., and Grompe, M. (2003). The Fanconi anaemia/BRCA pathway. Nat. Rev. Cancer *3*, 23–34.

Davies, H., Bignell, G.R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M.J., Bottomley, W., et al. (2002). Mutations of the BRAF gene in human cancer. Nature *417*, 949–954.

Druker, B.J., Tamura, S., Buchdunger, E., Ohno, S., Segal, G.M., Fanning, S., Zimmermann, J., and Lydon, N.B. (1996). Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. Nat. Med. *2*, 561–566.

Druker, B.J., Talpaz, M., Resta, D.J., Peng, B., Buchdunger, E., Ford, J.M., Lydon, N.B., Kantarjian, H., Capdeville, R., Ohno-Jones, S., and Sawyers, C.L. (2001). Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N. Engl. J. Med. *344*, 1031–1037.

Fearnhead, N.S., Britton, M.P., and Bodmer, W.F. (2001). The ABC of APC. Hum. Mol. Genet. *10*, 721–733.

Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.A., Rapaport, J.M., Albert, D.M., and Dryja, T.P. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature *323*, 643–646.

Futreal, P.A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N., and Stratton, M.R. (2004). A census of human cancer genes. Nat. Rev. Cancer *4*, 177–183.

Hahn, S.A., Schutte, M., Hoque, A.T., Moskaluk, C.A., da Costa, L.T., Rozenblum, E., Weinstein, C.L., Fischer, A., Yeo, C.J., Hruban, R.H., and Kern, S.E. (1996). DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science *271*, 350–353.

Hoeijmakers, J.H. (2001). Genome maintenance mechanisms for preventing cancer. Nature *411*, 366–374.

Hoque, M.O., Lee, C.C., Cairns, P., Schoenberg, M., and Sidransky, D. (2003). Genome-wide genetic characterization of bladder cancer: a comparison of high-density single-nucleotide polymorphism arrays and PCR-based microsatellite analysis. Cancer Res. *63*, 2216–2222.

Ishkanian, A.S., Malloff, C.A., Watson, S.K., DeLeeuw, R.J., Chi, B., Coe, B.P., Snijders, A., Albertson, D.G., Pinkel, D., Marra, M.A., et al. (2004). A tiling resolution DNA microarray with complete coverage of the human genome. Nat. Genet. *36*, 299–303.

Janne, P.A., Li, C., Zhao, X., Girard, L., Chen, T.H., Minna, J., Christiani, D.C., Johnson, B.E., and Meyerson, M. (2004). High-resolution single-nucleotide polymorphism array and clustering analysis of loss of heterozygosity in human lung cancer cell lines. Oncogene *23*, 2716–2726.

Kamb, A., Gruis, N.A., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavtigian, S.V., Stockert, E., Day, R.S., Johnson, B.E., and Skolnick, M.H. (1994). A cell cycle regulator potentially involved in genesis of many tumor types. Science *264*, 436–440.

Kauraniemi, P., Kuukasjarvi, T., Sauter, G., and Kallioniemi, A. (2003). Amplification of a 280-kilobase core region at the ERBB2 locus leads to activation of two hypothetical proteins in breast cancer. Am. J. Pathol. *163*, 1979–1984.

King, C.R., Kraus, M.H., and Aaronson, S.A. (1985). Amplification of a novel v-erbB-related gene in a human mammary carcinoma. Science *229*, 974–976.

Knudson, A.G., Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. Proc. Natl. Acad. Sci. USA *68*, 820–823.

Knudson, A.G. (2000). Chasing the cancer demon. Annu. Rev. Genet. *34*, 1–19.

Knudson, A.G. (2002). Cancer genetics. Am. J. Med. Genet. *111*, 96–102.

Knuutila, S., Bjorkqvist, A.M., Autio, K., Tarkkanen, M., Wolf, M., Monni, O., Szymanska, J., Larramendy, M.L., Tapper, J., Pere, H., et al. (1998). DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies. Am. J. Pathol. *152*, 1107–1123.

Knuutila, S., Aalto, Y., Autio, K., Bjorkqvist, A.M., El-Rifai, W., Hemmer, S., Huhta, T., Kettunen, E., Kiuru-Kuhlefelt, S., Larramendy, M.L., et al. (1999). DNA copy number losses in human neoplasms. Am. J. Pathol. *155*, 683–694.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Lengauer, C., Kinzler, K.W., and Vogelstein, B. (1998). Genetic instabilities in human cancers. Nature *396*, 643–649.

Lieberfarb, M.E., Lin, M., Lechpammer, M., Li, C., Tanenbaum, D.M., Febbo, P.G., Wright, R.L., Shim, J., Kantoff, P.W., Loda, M., et al. (2003). Genome-wide loss of heterozygosity analysis from laser capture microdissected prostate cancer using single nucleotide polymorphic allele (SNP) arrays and a novel bioinformatics platform dChipSNP. Cancer Res. *63*, 4781–4785.

Lindblad-Toh, K., Tanenbaum, D.M., Daly, M.J., Winchester, E., Lui, W.O., Villapakkam, A., Stanton, S.E., Larsson, C., Hudson, T.J., Johnson, B.E., et al. (2000). Loss-of-heterozygosity analysis of small-cell lung carcinomas using single-nucleotide polymorphism arrays. Nat. Biotechnol. *18*, 1001–1005.

Lisitsyn, N.A., Lisitsina, N.M., Dalbagni, G., Barker, P., Sanchez, C.A., Gnarra, J., Linehan, W.M., Reid, B.J., and Wigler, M.H. (1995). Comparative genomic analysis of tumors: detection of DNA losses and amplification. Proc. Natl. Acad. Sci. USA *92*, 151–155.

Lucito, R., Healy, J., Alexander, J., Reiner, A., Esposito, D., Chi, M., Rodgers, L., Brady, A., Sebat, J., Troge, J., et al. (2003). Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. Genome Res. *13*, 2291–2305.

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., et al. (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.

Ma, Y.Y., Wei, S.J., Lin, Y.C., Lung, J.C., Chang, T.C., Whang-Peng, J., Liu, J.M., Yang, D.M., Yang, W.K., and Shen, C.Y. (2000). PIK3CA as an oncogene in cervical cancer. Oncogene *19*, 2739–2744.

Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R.S., Zborowska, E., Kinzler, K.W., and Vogelstein, B. (1995). Inactivation of the type II TGF-β receptor in colon cancer cells with microsatellite instability. Science *268*, 1336–1338.

Mei, R., Galipeau, P.C., Prass, C., Berno, A., Ghandour, G., Patil, N., Wolff, R.K., Chee, M.S., Reid, B.J., and Lockhart, D.J. (2000). Genome-wide detection of allelic imbalance using human SNPs and high-density DNA arrays. Genome Res. *10*, 1126–1137.

Miller, V.A., Kris, M.G., Shah, N., Patel, J., Azzoli, C., Gomez, J., Krug, L.M., Pao, W., Rizvi, N., Pizzo, B., et al. (2004). Bronchioloalveolar pathologic subtype and smoking history predict sensitivity to gefitinib in advanced non-small-cell lung cancer. J. Clin. Oncol. *22*, 1103–1109.

Mitelman, F. (2000). Recurrent chromosome aberrations in cancer. Mutat. Res. *462*, 247–253.

Mitelman, F., Mertens, F., and Johansson, B. (1997). A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. Nat. Genet. *15*, 417–474.

Mitra, R.D., Shendure, J., Olejnik, J., Edyta-Krzymanska-Olejnik, and Church, G.M. (2003). Fluorescent in situ sequencing on polymerase colonies. Anal. Biochem. *320*, 55–65.

Naoki, K., Chen, T.H., Richards, W.G., Sugarbaker, D.J., and Meyerson, M. (2002). Missense mutations of the BRAF gene in human lung adenocarcinoma. Cancer Res. *62*, 7001–7003.

Nobori, T., Miura, K., Wu, D.J., Lois, A., Takabayashi, K., and Carson, D.A. (1994). Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers. Nature *368*, 753–756.

Nowell, P.C., and Hungerford, D.A. (1960). Chromosome studies on normal and leukemic human leukocytes. J. Natl. Cancer Inst. *25*, 85–109.

Paez, J.G., Janne, P.A., Lee, J.C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F.J., Lindeman, N., Boggon, T.J., et al. (2004). EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science *304*, 1497–1500.

Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L., et al. (2004). EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc. Natl. Acad. Sci. USA *101*, 13306–13311.

Pollack, J.R., Sorlie, T., Perou, C.M., Rees, C.A., Jeffrey, S.S., Lonning, P.E., Tibshirani, R., Botstein, D., Borresen-Dale, A.L., and Brown, P.O. (2002). Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. Proc. Natl. Acad. Sci. USA *99*, 12963–12968.

Rabbitts, T.H. (1994). Chromosomal translocations in human cancer. Nature *372*, 143–149.

Rajagopalan, H., Bardelli, A., Lengauer, C., Kinzler, K.W., Vogelstein, B., and Velculescu, V.E. (2002). Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status. Nature *418*, 934.

Rowley, J.D. (1973a). Chromosomal patterns in myelocytic leukemia. N. Engl. J. Med. *289*, 220–221.

Rowley, J.D. (1973b). Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. Nature *243*, 290–293.

Samuels, Y., Wang, Z., Bardelli, A., Silliman, N., Ptak, J., Szabo, S., Yan, H., Gazdar, A., Powell, S.M., Riggins, G.J., et al. (2004). High frequency of mutations of the PIK3CA gene in human cancers. Science *304*, 554.

Shayesteh, L., Lu, Y., Kuo, W.L., Baldocchi, R., Godfrey, T., Collins, C., Pinkel, D., Powell, B., Mills, G.B., and Gray, J.W. (1999). PIK3CA is implicated as an oncogene in ovarian cancer. Nat. Genet. *21*, 99–102.

Shendure, J., Mitra, R.D., Varma, C., and Church, G.M. (2004). Advanced sequencing technologies: methods and goals. Nat. Rev. Genet. *5*, 335–344.

Slamon, D.J., Leyland-Jones, B., Shak, S., Fuchs, H., Paton, V., Bajamonde, A., Fleming, T., Eiermann, W., Wolter, J., Pegram, M., et al. (2001). Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic

breast cancer that overexpresses HER2. N. Engl. J. Med. *344*, 783–792.

Stephens, P., Hunter, C., Bignell, G., Edkins, S., Davies, H., Teague, J., Stevens, C., O'Meara, S., Smith, R., Parker, A., et al. (2004). Lung cancer: intragenic ERBB2 kinase mutations in tumours. Nature *431*, 525–526.

Tengs, T., LaFramboise, T., Den, R.B., Hayes, D.N., Zhang, J., DebRoy, S., Gentleman, R.C., O'Neill, K., Birren, B., and Meyerson, M. (2004). Genomic representations using concatenates of Type IIB restriction endonuclease digestion fragments. Nucleic Acids Res. *32*, e121.

Tuveson, D.A., Weber, B.L., and Herlyn, M. (2003). BRAF as a potential therapeutic target in melanoma and other malignancies. Cancer Cell *4*, 95–98.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., et al. (2001). The sequence of the human genome. Science *291*, 1304–1351.

Vogelstein, B., Fearon, E.R., Kern, S.E., Hamilton, S.R., Preisinger, A.C., Nakamura, Y., and White, R. (1989). Allelotype of colorectal carcinomas. Science *244*, 207–211.

Volik, S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q., Kowbel, D., Huang, G., Lapuk, A., Kuo, W.L., et al. (2003). End-sequence profiling: sequence-based analysis of aberrant genomes. Proc. Natl. Acad. Sci. USA *100*, 7696–7701.

Wang, D.G., Fan, J.B., Siao, C.J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., et al. (1998). Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science *280*, 1077–1082.

Wang, T.L., Maierhofer, C., Speicher, M.R., Lengauer, C., Vogelstein, B., Kinzler, K.W., and Velculescu, V.E. (2002). Digital karyotyping. Proc. Natl. Acad. Sci. USA *99*, 16156–16161.

Wang, Z., Shen, D., Parsons, D.W., Bardelli, A., Sager, J., Szabo, S., Ptak, J., Silliman, N., Peters, B.A., van der Heijden, M.S., et al. (2004a). Mutational analysis of the tyrosine phosphatome in colorectal cancers. Science *304*, 1164–1166.

Wang, Z.C., Lin, M., Wei, L.J., Li, C., Miron, A., Lodeiro, G., Harris, L., Ramaswamy, S., Tanenbaum, D.M., Meyerson, M., et al. (2004b). Loss of heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancers. Cancer Res. *64*, 64–71.

Weinberg, R.A. (1996). How cancer arises. Sci. Am. *275*, 62–70.

Wolf, M., Mousses, S., Hautaniemi, S., Karhu, R., Huusko, P., Allinen, M., Elkahloun, A., Monni, O., Chen, Y., Kallioniemi, A., and Kallioniemi, O.P. (2004). High-resolution analysis of gene copy number alterations in human prostate cancer using CGH on cDNA microarrays: impact of copy number on gene expression. Neoplasia *6*, 240–247.

Zhao, X., Li, C., Paez, J.G., Chin, K., Janne, P.A., Chen, T.H., Girard, L., Minna, J., Christiani, D., Leo, C., et al. (2004). An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. Cancer Res. *64*, 3060–3071.