3 Impact of Protein Interaction Technologies on Cancer Biology and Pharmacogenetics

Rebecca Blanchard Raftogianis[1] and Andrew K. Godwin[2]

Departments of [1]Pharmacology and [2]Medical Oncology, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Protein–Protein Interactions: A Molecular Cloning *Manual*, © 2002 by Cold Spring Harbor Laboratory Press, Chapter 3.

SEQUENCE VARIATIONS: SUSCEPTIBILITY TO CANCER AND TREATMENT BASED ON GENOTYPE-SPECIFIC PROTEIN INTERACTIONS

New genomic and proteomic technologies and the completion of the human genome sequence promise a revolution in our understanding of human disease. These developments are expected to transform dramatically the study of diseases such as cancer and to allow the molecular classification of such diseases. Furthermore, the identification and functional characterization of genetic variation in the human genome will be a major scientific effort in the post-genomics era. Identification of genetic variation is the current focus of several genome projects and is proceeding at a rapid pace (Service 2000). Many diverse sources of data have shown that any two individuals are more than 99.9% identical in sequence, which means that the differences among individuals in our own species that can be attributed to genes falls in a mere 0.1% (or ~2,900,000 nucleotides) of the genome (Venter et al. 2001). Characterization of the functional significance of those variants will be laborious and will likely continue to be a scientific focus for many years. Two areas of particular interest are genetic variants that predispose individuals to disease and variants that affect individual response to drugs (pharmacogenetics/pharmacogenomics). Genetic variation can impart a functional phenotype via a myriad of mechanisms, including the introduction or interruption of protein–protein interactions or protein–ligand interactions. The focus of this chapter is to discuss the application of the techniques described in this book to the study of functional significance of human genetic variation in the context of human disease susceptibility and pharmacogenetics.

CANCER AND THE GENOME

Cancer is an important public health concern in the United States and around the world. After heart disease, it is the second leading cause of death, accounting for 23.3% of deaths in the United States (Greenlee et al. 2000). In the year 2000, about 1,220,100 new cases of invasive cancer were expected to be diagnosed in the United States. In addition, ~1.3 million cases of basal and squamous cell skin cancer are diagnosed annually (Greenlee et al. 2000). An estimated 552,000 Americans were predicted to have died of cancer in 2000, which translates into more than 1500 people a day. Among men, the most common cancers are cancers of the prostate (29%), lung and bronchus (14%), and colon and rectum (10%). Among women, the three most commonly diagnosed cancers are cancers of the breast (30%), lung and bronchus (12%), and colon and rectum (11%) (Greenlee et al. 2000). In 1987, lung cancer surpassed breast cancer as the leading cause of cancer death in women and is expected to account for at least 25% of all female cancer deaths in the coming years. The good news is that following more than 70 years of increases, the recorded number of total cancer deaths among men in the United States has declined for the first time, from a peak of 281,898 in 1996 to 281,110 in 1997. In contrast, among women, the recorded number of total cancer deaths continues to increase, although the rate of increase has diminished in recent years. These somewhat encouraging trends are primarily associated with improved screening techniques and the subsequent increase in diagnosis at an early stage when most, but not all, cancers are more successfully treated. Unfortunately, most current cancer therapies have limited efficacy in curing late-stage disease. Therefore, there continues to be a great and immediate need to develop new approaches to (1) diagnose cancer early in its clinical course, (2) more effectively treat advanced stage disease, (3) better predict a tumor's response to therapy prior to the actual treatment, and (4) ultimately prevent disease from arising through the use of chemopreventive strategies. These goals can only be accomplished through a better understanding of how certain genes and their encoded proteins contribute to disease onset and tumor progression and how they influence the response of patients to drug therapies. Innovations in genetic, biological, and biochemical approaches are necessary to realize these goals.

Genetic Risk Factors

The risk of developing high-incidence cancers (e.g., lung, breast/ovarian, colon, prostate) is not uniformly distributed throughout the population. In part, this lack of uniformity may be explained by different environmental, occupational, and recreational exposure histories (e.g., ultraviolet light from the sun, inhaled cigarette smoke, incompletely defined dietary factors). These carcinogens can affect one or multiple stages of carcinogenesis through both genetic and epigenetic mechanisms (Shields and Harris 2000; Rothman et al. 2001). However, only a small fraction of exposed individuals ultimately develop cancer. Moreover, the risk among individuals with similar exposures is unevenly distributed. Individuals in certain families have been observed to have greater risk and earlier onset of cancer. Families ascertained through a cancer proband (i.e., from the Latin propositus [male] or proposita [female], the individual through whom the family is ascertained) also show significant excess of other cancers in their family members. This excess in cancer incidence persists even when rare Mendelian forms segregating in families are removed. These observations suggest that common cancer susceptibility may, in part, be determined by host genetic factors in addition to Mendelian factors identified to date.

At least two classes of genes have been identified that may determine the risk of developing human cancer. The first class is composed of a restricted number of genes that are directly involved in tumorigenesis, such as proto-oncogenes, tumor suppressor genes, and DNA mismatch repair genes. These can be thought of as the cancer-causing genes (Fearon 1997; Godwin et al. 1997). The risk associated with germ-line mutations in this class of genes, e.g., tumor suppressor genes (including *RB1, TP53, CDKN2, WT1, NF1, NF2, TSC1, TSC2, VHL, PTCH, PTEN, LKB1, SMAD4, APC, MEN1, CDH1, TGFBR2, EXT1, EXT2*), oncogenes (*RET, MET, KIT, CDK4*), DNA mismatch repair genes (*MSH2, MLH1, PMS1, PMS2, MSH6*), and two genes related to DNA repair and tumor suppression (*BRCA1, BRCA2*) is high, but the risk alleles are rare in the population (Fig. 1). For example, the risk of developing a cancer in a RB mutation carrier is extremely high (>1000-fold), whereas the "at risk" allele is infrequent (~1 in 20,000 live births) in the general population. These genes are also frequent targets for mutations in sporadic forms of the disease. These cancer susceptibility genes encode proteins that perform diverse cellular functions, including transcription, cell cycle control, DNA repair, and apoptosis. The proteins encoded by these genes often function through protein–protein interactions, and thus their characterization is particularly amenable to the use of techniques described in this book. The second class is broader and not completely defined (Rothman et al. 2001). The most frequently studied group in this case is composed of genes involved in metabolic detoxification pathways and steroid and amino acid metabolism (Evans and Relling 1999). The risk associated with the second class is significantly lower (maybe two- to tenfold) but is likely to be of greater public health significance as a consequence of the higher frequency (as great as 50% in certain populations) of the risk-modifying factors (Fig. 1). For example, the cytochromes P450 (CYP) evolved to catalyze the metabolism of numerous structurally diverse exogenous and endogenous molecules. Approximately 55 different CYP genes are present in the human genome and are classified into different families and subfamilies on the basis of sequence homology. Members of the CYP3A subfamily, for example, catalyze the oxidative, peroxidative, and reductive metabolism of structurally diverse endobiotics, drugs, and protoxic or procarcinogenic molecules (Rendic and Di Carlo 1997). Moreover, because CYP3A metabolizes estrogens to 2-hydroxyestrone, 4-hydroxyestrone, and 6α-hydroxylated estrogens, all of which have been implicated in estrogen-mediated carcinogenicity, variation in CYP3A may influence the circulating levels of these estrogens and the risk of breast cancer. CYP3A4, a member of the CYP3A family, has been shown to be associated with oxidative deactivation of testosterone. Rebbeck and colleagues demonstrated that a single base change in the *CYP3A4* gene was significantly associated with higher clinical stage and grade in men with prostate tumors (Rebbeck et al. 1998), indicating that mutations in *CYP3A4* may influence prostate carcinogenesis. Kuehl and colleagues have recently shown that a single-nucleotide polymorphism (SNP) in *CYP3A5* leads to alternative splicing and protein truncation and results in the absence of CYP3A5

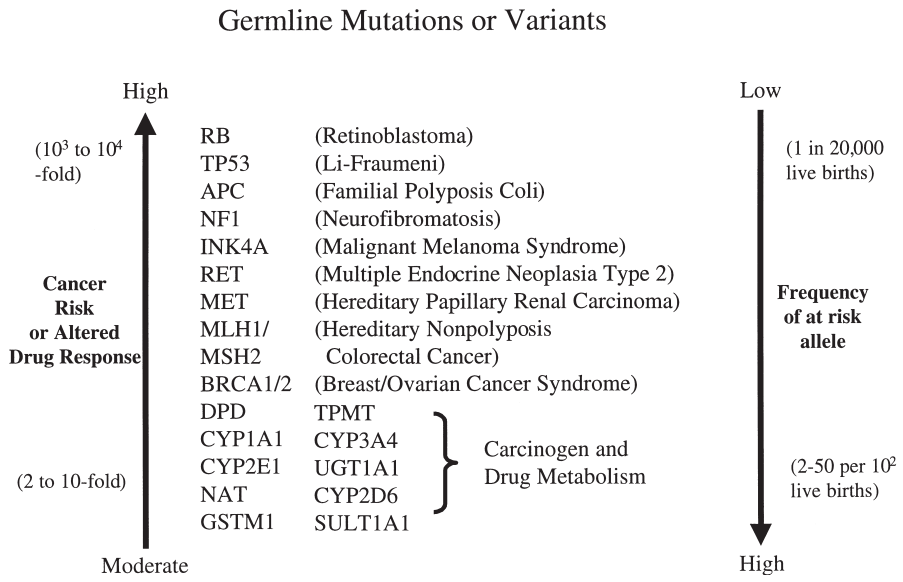

FIGURE 1. Example of the inverse correlation between allelic frequency and cancer risk associated with inherited cancer susceptibility genes versus the cancer risk or change in drug response associated with variations in drug-metabolizing genes. The cancer syndrome associated with germ-line mutations in each of the cancer susceptibility genes listed is shown in parentheses. (DPD) Dihydropyrimidine dehydrogenase; (CYP1A1) cytochrome P450 1A1; (CYP2E1) cytochrome P450 2E1; (NAT) *N*-acetyltransferase; (GSTM1) glutathione-*S*-transferase, Mu 1; (TPMT) thiopurine methyltransferase; (CYP3A4) cytochrome P450 3A4; (UGT1A1) UDP-glucuronosyltransferase 1A1; (CYP2D6) cytochrome P450 2D6; (SULT1A1) sulfotransferase 1A1.

from tissues in some homozygous individuals (Kuehl et al. 2001). Because CYP3A5 represents at least 50% of the total hepatic CYP3A content in people polymorphically expressing CYP3A5, *CYP3A5* may be the most important genetic contributor to interindividual differences in CYP3A-dependent drug clearance and in response to many medicines.

Molecular genetic methods are being combined with genomic science approaches to facilitate the identification and large-scale characterization of constitutional DNA variation in these kinds of candidate loci. The actual number of genes present in the human genome that contribute to cancer susceptibility as well as response to drugs is unknown and will require many years of additional research to uncover the normal function of the encoded proteins and how these carefully regulated activities are disrupted in cancer. New protein interaction technologies will be key to uncovering how sequence variations as small as a single-nucleotide substitute in any given gene can potentially alter the function of the protein through its interaction with other proteins and/or substrates. Regarding cancer, genes involved in cancer susceptibility, cancer progression, and response to drug therapy will be key. Each of these classes is discussed in some detail below.

CANCER GENES

Despite the prevalence of cancer, little is known about the molecular events that occur during its development. Cancer, both hereditary and nonhereditary, is a multistep process that involves alterations in many specific genes. The normal cell has multiple independent mechanisms that regulate its growth and differentiation, and several separate events are required to override these control mechanisms. The fundamental mechanisms underlying the genetic basis of cancer are

being defined and involve, as indicated above, alterations in genes that have been classified into three general categories: (1) proto-oncogenes, which are involved in growth promotion and whose defects leading to cancer are a gain of function; (2) tumor suppressor genes, which are negative regulators of growth and whose loss of function contributes to cancer; and (3) DNA repair genes, which maintain the integrity of the genome and whose loss of function causes increased accumulation of mutations in other critical cancer-causing genes. Progress is being made in isolating these genes and the proteins they encode, determining the normal cellular functions of the proteins, and investigating the mechanisms of tumorigenesis. Recently, it was predicted that ~1.3% (or roughly 351) of the genes present in the human genome are likely to be proto-oncogenes, whereas as many as 130 human DNA repair genes have already been identified (Lander et al. 2001; Wood et al. 2001). The number of classic tumor suppressor genes is relatively small, 20–30; however, loss of heterozygosity (LOH) studies suggest dozens of additional tumor suppressors have yet to be classified (discussed below). A combination of somatic mutations and chromosomal alterations affecting these genes and other genes is thought to be the driving force behind tumorigenesis. Much work remains to uncover how these proteins function normally to understand how they contribute to cancer. However, a number of these have been drawn into nets of interactions with other proteins and are beginning to provide hints to their activity.

Oncogenes

Proto-oncogenes/oncogenes are genes whose action promotes cell proliferation in a positive manner. The normal, nonmutant versions are properly called proto-oncogenes, and the mutant versions, the oncogenes, are excessively or inappropriately active. Mutations converting normal proto-oncogenes into oncogenes are gain-of-function mutations, and these mutations include point mutations; structural alterations such as insertions, deletions, inversions, and translocations; gene amplification; and hypomethylation. Oncogenes were originally identified on the basis of their similarity to retroviral sequences, which were known to be able to transform cells. Many cellular oncogenes have now been identified; however, their involvement in cancer has yet to be fully elucidated. The most commonly studied oncogenes are those of the SRC, RAS, MYC, and ERBB family of proteins (Godwin et al. 1997). Although the functions of these genes initially appeared to be unrelated, elucidation of protein interactions has in some cases shown commonalities of effect. Furthermore, studies of these proteins share some areas with studies of signal transduction (see Chapter 2).

Tumor Suppressor Genes

The concept that genes could suppress cell growth came from early studies of somatic cell fusion (Harris et al. 1969; Stanbridge 1976). In those experiments, fusion of tumorigenic cells with normal cells resulted in hybrids that could continue to grow in culture but were no longer tumorigenic in animals. Moreover, when some of the hybrid cells re-expressed the tumorigenic phenotype, that re-expression correlated with specific loss of chromosomes derived from the normal, nontumorigenic parental cell. In 1969, de Mars proposed that in certain familial cancers, gene carriers might be heterozygous for a recessive mutation and that the cancer appears because of subsequent somatic mutation causing an individual cell to become homozygous for the cancer-causing gene (DeMars 1969). Epidemiological analysis of retinoblastoma and Wilms' tumor provided important evidence for this theory (Knudson 1971; Knudson and Strong 1972) and led to what is now referred to as Knudson's "two-hit" hypothesis. Two forms of the disease were observed: patients with a family history of disease who often presented with multiple bilateral lesions and patients without family history who typically presented with single unilateral lesions and at a later age than those with familial disease. Knudson and Strong speculated that the familial form of dis-

ease represented the inheritance of a mutation predisposing to neoplasia and that only one additional rate-limiting genetic event was necessary for tumorigenesis. Sporadic tumors, on the other hand, required two independent genetic lesions and therefore would be slower to develop. Because two independent lesions occurring in the same cell would be rare, only unilateral tumors would be expected. This hypothesis was confirmed at the molecular level with the isolation of the *RB1* gene (Cavenee et al. 1983; Dunn et al. 1989). The development of retinoblastoma has been shown to involve loss of both alleles at a single locus, with individuals with the hereditary form of retinoblastoma having one mutated allele in the germ line (Marshall 1991). Retinoblastoma is one of several dozen prototypic tumor suppressor genes (e.g., *TP53, APC, WT1, NF1, PTEN*) that conform to the classic definition of Knudson's two-hit hypothesis.

Increasingly, different mechanisms of mutation are being implicated in the inactivation of tumor suppressor genes leading to neoplasia (both hereditary and nonhereditary forms). These include point mutations, deletions, insertions, hypermethylation, alterations in genomic imprinting, loss of genetic material leading to haploinsufficiency, dominant negative mutations, and homozygous deletions (Fearon 2000; Hanahan and Weinberg 2000; Prowse et al. 2001). Based on the idea of tumor suppressors as components of complexes, these genetic lesions predict multiple modes of action, including maintenance of genomic stability, programmed cell death (apoptosis), DNA repair, and cell cycle control (Fearon 1997; Godwin et al. 1997; Lundberg and Weinberg 1999; Hanahan and Weinberg 2000). Furthermore, as more and more individuals elect to undergo diagnostic testing at the DNA level for disease susceptibility, the range of known mutations in these genes will continue to grow. This scenario is already giving rise to a new set of problems, namely, delineating deleterious mutations from benign polymorphisms. Protein technology that will help to assess the functional consequences of a single residue change in a protein will be key. For example, at present, over 860 and 880 different sequence variants (deleterious mutations, naturally occurring polymorphisms, and unclassified variants) have been identified in *BRCA1* and *BRCA2*, respectively, and more than 50% have been reported only once (see Breast Cancer Information Core, www.nhgri.nih.gov/ Intramural_research/Lab_transfer/Bic/index.html). Although a few missense mutations within these genes have been identified as conferring a predisposition to breast cancer, most such mutations have not been associated with a known phenotype. Like nonsense mutations, missense mutations result from the substitution of a single nucleotide within a coding codon. Unlike nonsense mutations, the substitution results in a functional codon but encodes a different amino acid at that position. The problem with missense changes is that it is not always simple to determine whether the amino acid substitution will adversely affect the protein's function and thereby contribute to the disease phenotype. If the missense change is commonly found in control populations (ethnically matched disease-free individuals with no family history of breast and/or ovarian cancer), it is deemed to be a naturally occurring polymorphism. However, many of these variants are found in only a limited number of families and are referred to as variants of unknown significance. The frequent discovery of variants of unknown significance is a major problem from a clinical standpoint because many patients who undergo genetic testing are left to interpret these ambiguous results while trying to make important health-care decisions. Therefore, new protein technologies will be necessary to complement genetic studies and to help define the consequence of these single-base changes in regard to protein function and, thus, cancer risk.

Finally, LOH studies are the most frequently used method to define a region that may harbor a tumor suppressor gene. Whole-genome allelotyping studies of nearly every type of common cancers revealed that LOH has been observed on every chromosome arm. These studies suggest that a large number of tumor suppressor genes may exist, most of which have yet to be identified and their functions characterized. As the technology continues to improve, functional genomics and high-throughput screening methods will provide powerful means to identify the genetic components that influence human health and disease. Proteomic approaches will further expand the molecular profile of disease and define the functional pathways.

DNA Repair Genes

The human genome, like other genomes, encodes information to protect its own integrity (Lindahl and Wood 1999). DNA repair enzymes continuously monitor chromosomes to correct damaged nucleotide residues generated by exposure to carcinogens and cytotoxic compounds. Genomic instability caused by the great variety of DNA-damaging agents would be an overwhelming problem for cells and organisms if it were not for DNA repair. Proteins for base excision repair, nucleotide excision repair, and mismatch excision repair have all been identified, and defects in some of these proteins have been shown to be closely associated with cancer development. For example, proteins encoded by mismatch repair genes correct occasional errors of DNA replication as well as heterologies formed during recombination. The bacterial *mutS* and *mutL* genes encode proteins responsible for identifying mismatches, and there are numerous homologs of these genes in the human genome, of greater variety than those found in yeast and nematodes. Some of these proteins are specialized for locating distinct types of mismatches in DNA, some are specialized for meiotic and/or mitotic recombination, and some have functions yet to be determined. Of the 130-plus DNA repair genes, at least 11 are mismatch excision repair genes (Wood et al. 2001), and some of these genes (e.g., *hMSH2, hMLH1, hPMS1*, and *hPMS2*), when mutated, have been shown to predispose to hereditary nonpolyposis colorectal cancer (HNPCC) or Lynch syndrome (Lynch and Smyrk 1996), as well as sporadic forms of the disease. *hMLH1* and *hMSH2* are believed to account for most HNPCC families with an identifiable deleterious mutation (Peltomaki and Vasen 1997). Studies of these proteins from bacteria to mammals emphasize their assembly into recombination "machines." Furthermore, study of the locations of the mutations in the context of the assembled structure of components should both provide insight into basic mechanisms and suggest targets for therapeutic intervention as discussed below.

The study of protein–protein interactions has provided important insights into the functions of many of the known oncogenes, tumor suppressors, and DNA repair proteins. These studies have also demonstrated that members of each class of protein participate in overlapping pathways and, when mutated, each theoretically could contribute to cancer. A strong example of this is the human *TP53* gene. *TP53* codes for a protein product (referred to as p53) that has an important biological function as a cell cycle checkpoint. p53, originally detected by virtue of its ability to form a stable complex with the SV40 large-T antigen (Fig. 2), has been a constant source of fasci-

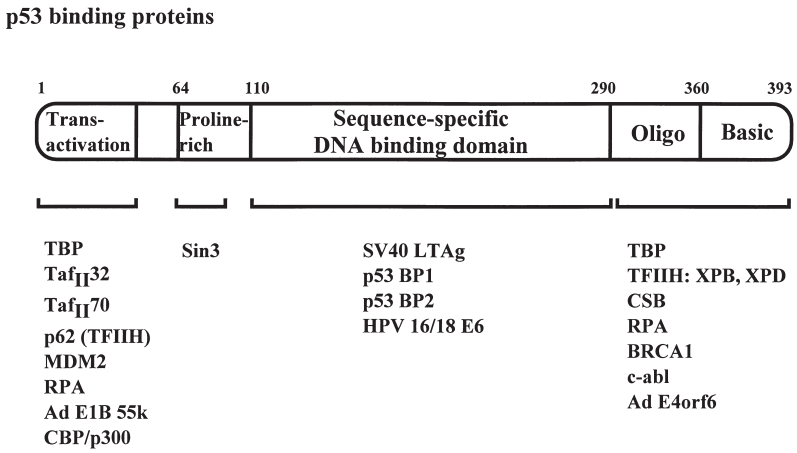

FIGURE 2. Schematic of p53 protein and its interacting proteins. The various protein domains and the proteins that interact with each of these regions are shown (Mansur et al. 1995; Dobner et al. 1996; Gottlieb and Oren 1996; Ko and Prives 1996; Levine 1997; Lill et al. 1997; Zhang et al. 1998; Nie et al. 2000; Zilfou et al. 2001). (This figure was provided courtesy of Dr. Maureen Murphy, Fox Chase Cancer Center, Philadelphia.)

nation since its discovery over a decade ago (for reviews, see Gannon and Lane 1990; Soussi et al. 1990). The gene encoding this 53-kD nuclear phosphoprotein was initially considered to be a cellular oncogene because introduction of expression vectors containing mutant *TP53* cDNA clones by transfection could transform recipient cells in concert with an activated *ras* gene. Subsequently, however, several convergent lines of research indicated that normal (wild-type) p53 actually functioned as a tumor suppressor.

The *TP53* gene is one of the most commonly altered genes identified in human tumors (e.g., sporadic osteosarcomas, soft-tissue sarcomas, brain tumors, leukemias, and carcinomas of the breast, colon, lung, and ovary) occurring in a large fraction (perhaps even half) of the total cancers in the United States (Hollstein et al. 1996; Levine 1997; Hainaut et al. 1998). Unlike many tumor suppressor genes, missense mutations represent a high proportion (>70%) of *TP53* mutations (http://www.iarc.fr/p53/homepage.html). Furthermore, in contrast to the retinoblastoma gene *RB*, where the hereditary syndrome served as the basis for identification of the causal gene, *TP53* was discovered and subsequently found to have a role in hereditary cancer. In 1990, Li and colleagues identified germ-line *TP53* mutations in a series of families with Li-Fraumeni syndrome (LFS), which features diverse childhood cancers as well as early-onset breast cancers (Malkin et al. 1990).

The p53 exists at low levels in virtually all normal cells. Wild-type p53 acts as a negative regulator of cell growth that is induced following DNA damage and mediates cell cycle arrest in late G_1. In some contexts, wild-type p53 can induce apoptosis (programmed cell death) and, in the absence of the wild-type protein, leads to resistance to ionizing radiation and chemotherapeutic agents. For example, in normal cells with DNA damaged by ultraviolet or γ irradiation, progression through the cell cycle is blocked at G_1, coincident with a sharp rise in the levels of p53. During the subsequent arrest of growth, repair of DNA is completed before the cells proceed into S phase. If, however, genomic damage is excessive, the cell undergoes apoptosis, which requires normally functioning p53. Cells can escape apoptosis in the absence of a functional p53 protein, thus allowing the cell to survive and replicate its damaged DNA, which in turn leads to the propagation of the mutation. Therefore, p53 has been described as the "guardian of the genome" because it prevents entry into S phase unless, or until, the genome has been cleared of potentially damaging mutations. In addition, because many chemotherapeutic drugs are believed to kill tumor cells by inducing apoptosis, loss of p53 function may also directly decrease the cells' sensitivity to such cytotoxic agents, enhancing the emergence of drug-resistant populations of cancer cells.

The biochemical mechanisms by which p53 acts in regulating cell proliferation are not fully understood; however, through multiple approaches, p53 has been shown to mediate growth suppression in part through its specific DNA-binding and transcriptional regulatory abilities (El-Deiry et al. 1994; Ko and Prives 1996). Many of these functional studies have relied on defining functional domains of p53 through their interaction with specific binding partners (Mansur et al. 1995; Dobner et al. 1996; Gottlieb and Oren 1996; Ko and Prives 1996; Levine 1997; Lill et al. 1997; Zhang et al. 1998; Nie et al. 2000). As shown in Figure 2, p53 interacts with a host of proteins, including the tumor suppressor, BRCA1, and the proto-oncogene, c-ABL. Recently, Zilfou and colleagues demonstrated through protein interaction studies that the corepressor SIN3 directly interacts with the proline-rich domain of p53 and protects p53 from proteosome-mediated degradation (Zilfou et al. 2001). This interaction is particularly noteworthy because deletion of the proline-rich domain of p53 renders this protein capable of functioning as an activator of transcription, but incapable of inducing programmed cell death, or apoptosis. Elucidation of the functional consequences of the p53–SIN3 interaction is likely to reveal new mechanisms of tumor suppression and apoptosis induction by p53 (Zilfou et al. 2001).

Many oncogenes, tumor suppressors, and DNA repair proteins appear to be part of much larger complexes. For example, Wang and colleagues recently reported that a set of proteins associates with BRCA1 to form a large megadalton protein complex, referred to as BASC (*B*RCA1-*a*ssociated genome *s*urveillance *c*omplex). This complex includes tumor suppressors; the DNA damage repair proteins MSH2, MSH6, MLH1, ATM, and BLM; and the RAD50–MRE11–NBS1 protein

complex (Wang et al. 2000). In addition, DNA replication factor C (RFC), a protein complex that facilitates the loading of PCNA onto DNA, is also part of the BASC. The association of BRCA1 with MSH2 and MSH6, both of which are required for transcription-coupled repair, provides a possible explanation for the role of BRCA1 (and BRCA2) in this pathway. These interactions are consistent with a role of both proteins in some aspects of transcription-coupled DNA repair and DNA recombination.

Studies of the BRCA1 protein following DNA damage (discussed below) have also uncovered a role for other tumor suppressors and oncogenes (i.e., ATM, ATR, CHK2, and CDK2) in regulating BRCA1 activity. BRCA1 exists in nuclear foci but is hyperphosphorylated and disperses after DNA damage (Scully et al. 1997a; Thomas et al. 1997). This dose-dependent change in the state of BRCA1 phosphorylation is accompanied by a specific loss of the BRCA1-containing nuclear foci during S phase. After BRCA1 dispersal, BRCA1, BRCA2, BARD1, and RAD51 accumulate focally on $PCNA^+$ replication structures, implying an interaction of BRCA1/BRCA2/BARD1/RAD51-containing complexes with damaged, replicating DNA (Fig. 3). Lee and colleagues have reported that the human CHK2 kinase (CDS1/CHK2) regulates BRCA1 function after DNA damage by phosphorylating Ser-988 of BRCA1 (Lee et al. 2000). CHK2 and BRCA1 were shown to interact and colocalize within discrete nuclear foci but to separate after γ irradiation. Phosphorylation of BRCA1 at Ser-988 was shown to be required for the release of BRCA1 from CDK2. ATR, a mammalian homolog of yeast S-phase checkpoint gene products, in part controls BRCA1 phosphorylation following hydroxyurea treatment (Tibbetts et al. 2000). Recently, ATR was found to colocalize with BRCA1 in somatic cells, both before and after replication arrest (Tibbetts et al. 2000). BRCA1 has also been shown to be phosphorylated by CDK2 at Ser-1497, concordant with the G_1/S-specific increase in BRCA1 phosphorylation, independent of DNA damage (Ruffner et al. 1999). Phosphorylation of BRCA1 in response to DNA damage has been shown to be dependent on ATM (Cortez et al. 1999). Falck and colleagues recently demonstrated, through various protein studies, a functional link between ATM, CHK2, the phosphatase CDC25A, and CDK2. These proteins have been implicated in radioresistant DNA synthesis, and defects in some have been shown to predispose to, or promote, tumorigenesis (Galaktionov et al. 1995; Bell et al. 1999; Rotman and Shiloh 1999; Falck et al. 2001). Together, these studies emphasize how various approaches to study protein function are uncovering the ways in which key pathways converge and how defects in these components can contribute to cancer.

PHARMACOGENETICS: CURRENT STATUS AND FUTURE GOALS

Whereas the description of cancer genetics above details genetic changes that lead to onset of disease, pharmacogenetics deals with individual response to the environment, whether in susceptibility to mutagenic insult or in ability to respond to therapeutic drugs. Pharmacogenetics has been studied historically in the context of variable drug metabolism (Weber 1997). Several functionally significant genetic polymorphisms within genes that encode drug-metabolizing enzymes have been identified. Classically, those discoveries have been made after observation of variable human response to a drug, often with the variant phenotype manifesting as a toxic drug reaction. For example, ~1 in 300 Caucasian individuals are deficient in the ability to metabolically inactivate the chemotherapeutic agent 6-mercaptopurine (Weinshilboum and Sladek 1980). Individuals with such a deficiency are at risk for profound myelosuppression and ultimately death if they are prescribed standard doses of 6-mercaptopurine (Lennard 1997; Krynetski and Evans 1998). The deficient phenotype is encoded by one of several genetic polymorphisms in the human thiopurine methyltransferase (TPMT) gene (Krynetski et al. 1996; Otterness et al. 1997). Clinicians are now able to phenotype or genotype individuals for the TPMT polymorphism(s) prior to initiation of 6-mercaptopurine therapy (Fig. 4). Functionally significant polymorphisms

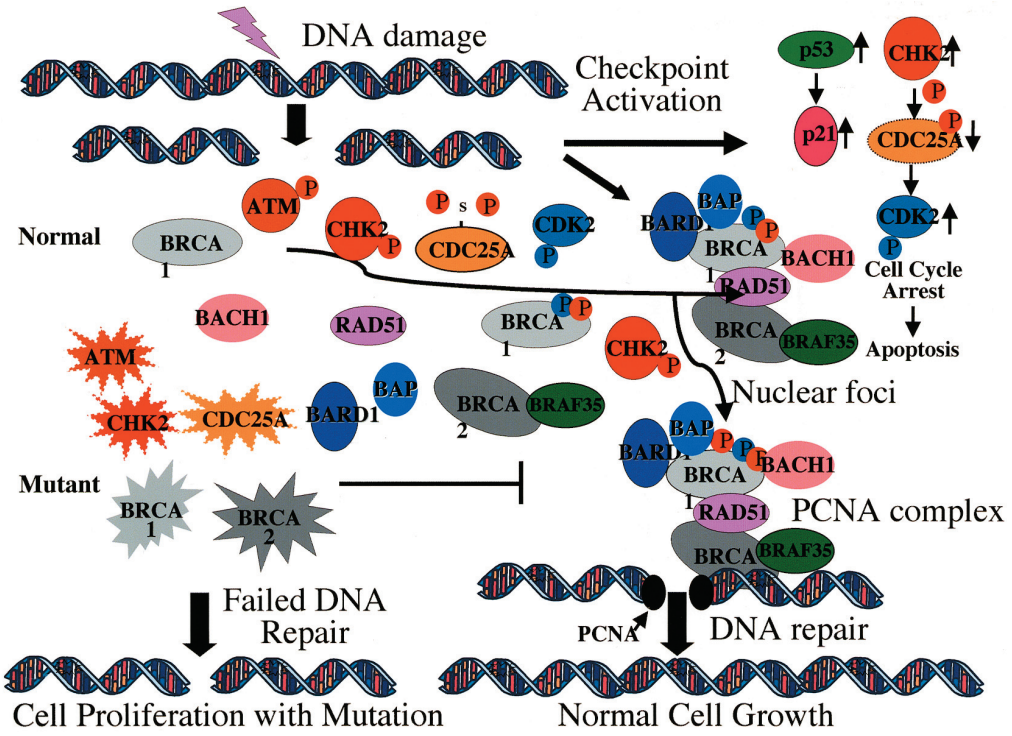

FIGURE 3. Role of BRCA1/2 in DNA replication and repair. Following DNA damage, cell cycle checkpoints are activated, including *p53*. p53 induces *p21* and *CHK2* expression, and signals cell cycle arrest until the damaged DNA can be repaired. The CDC25A phosphatase activates the cyclin-dependent kinase 2 (CDK2), which is also needed for DNA synthesis, but becomes degraded in response to DNA damage. Ionizing radiation-induced destruction of CDC25A requires both ATM and the CHK2-mediated phosphorylation of CDC25A on Ser-123. Failure to repair leads to apoptosis, or programmed cell death (as depicted in the upper right-hand corner). BRCA1 and BRCA2 participate in DNA repair by an unknown mechanism, which is thought to involve transcription-coupled repair. BRCA1 is known to be phosphorylated by ATM and CDK2 in a damage-free cell. BRCA1 is phosphorylated by CHK2 and ATR after damage; phosphorylated BRCA1 associates with each of the proteins BARD1, RAD51, BRCA2, and PCNA during the processes of repair and replication (*Normal*). DNA damage is successfully repaired and normal cell growth proceeds. Mutations in BRCA1, BRCA2, ATM, CHK2, or CDC25A proteins (*Mutant*) may cause dysfunction in this repair/replication complex/system due to their failure to interact with other protein. In this case, DNA repair fails and proliferation continues with damaged DNA. The role of other proteins such as BAP1 (BRCA1-associated protein-1), BRAF35 (BRCA2-associated factor 35), and BACH1 (BRCA1-associated C-terminal helicase) in repair is not known and will require additional studies.

have also been identified in several other drug-metabolizing genes, some of which are associated with variable clinical response to drugs (Ingelman-Sundberg et al. 1999; Iyer et al. 1999; Rettie et al. 1999; Wormhoudt et al. 1999).

In addition to detoxifying and eliminating drugs and metabolites, drug-metabolizing enzymes are often required for activation of prodrugs. For example, many opioid analgesics are activated by CYP2D6 (Poulsen et al. 1996), rendering the 2–10% of the population who are homozygous for nonfunctional CYP2D6 mutant alleles relatively resistant to opioid analgesic effects. Thus, it is not surprising that there is remarkable interindividual variability in the adequacy of pain relief when uniform doses of codeine are widely prescribed.

More recently, pharmacogenetic research has expanded to include the study of drug transporters, drug receptors, and drug targets. As the Human Genome Project progresses, scientists will have a map of hundreds of thousands of common genetic polymorphisms. As this scope expands, genetic profiles, defined by multiple gene loci, will be identified that are associated with particu-

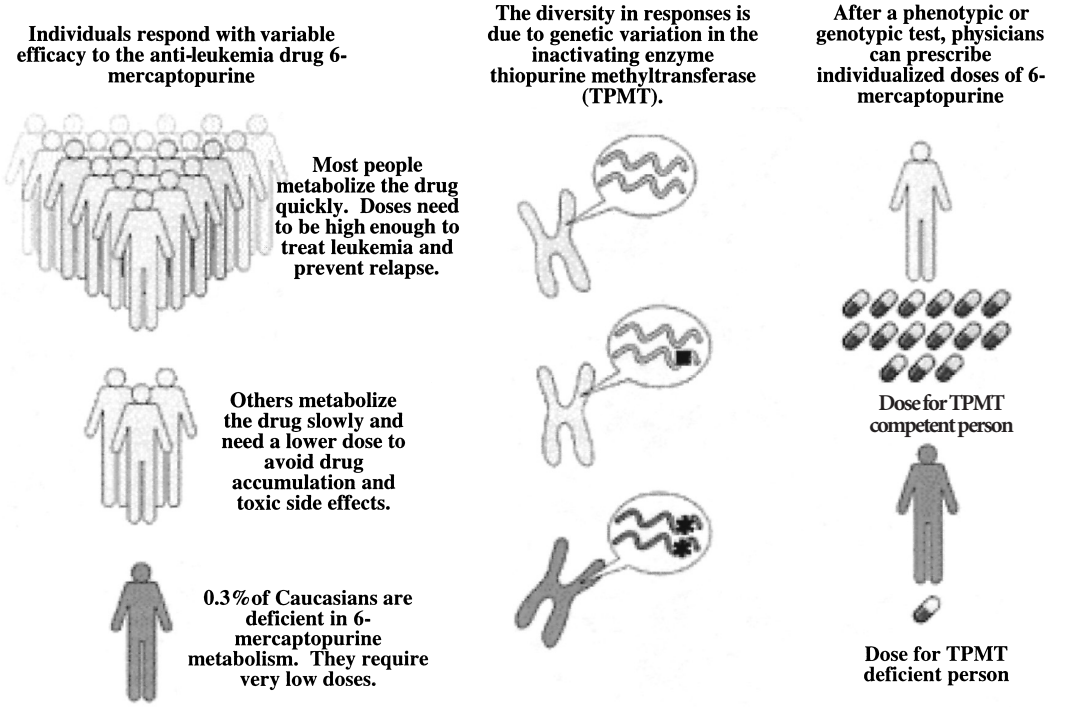

FIGURE 4. Thiopurine methyltransferase (TPMT) deficiency and pharmacogenetics of thiopurine therapy. The TPMT paradigm is used to illustrate a classic pharmacogenetic case in which variation in drug response is caused by a single gene defect. The first panel illustrates phenotypic differences in thiopurine metabolism and response. The second panel indicates the associated genetic variation (a black square marks a heterozygous deficient allele and the asterisks mark deficient alleles expressed homozygously). The third panel illustrates modifications in thiopurine dosing based on TPMT genetics.

larly favorable or adverse responses to drugs. It is hoped that this information will allow clinicians to use genotype information to individualize the prescribing of drugs. As more genetic variants are associated with clinical syndromes and responses, increased efforts will be necessary to understand the molecular mechanisms underlying these phenotypic changes, some of which undoubtedly involve protein–protein interactions.

Many functionally significant polymorphisms alter encoded amino acids that are not involved in the catalytic regions of enzymes. The functional significance of these changes suggests effects with regard to the interaction of these proteins with other species, including enzymes, regulators, and drugs. In fact, a common mechanism by which polymorphisms cause phenotypic change is via altered protein stability. Typically, the amino acid change results in increased degradation of the variant protein, resulting in a lower steady-state level of protein within the cell. For example, two of the TPMT allozymes associated with TPMT deficiency, TPMT*2 and TPMT*3A, are degraded an order of magnitude more rapidly than the wild-type protein (Tai et al. 1997). The specific protein interactions contributing to this regulatory change are not known, however, and application of the techniques such as those described in this book would elucidate this mechanism further.

As described above, study of protein interactions has contributed greatly to the general understanding of mechanisms of cancer through a convergence of fundamental knowledge about signal transduction and DNA repair. The study of protein interactions within the context of classic pharmacogenetics has thus far been relatively limited because most pharmacogenetic investigation has focused on the interaction of proteins with small molecules. However, as the focus of pharmacogenetic research expands to include classes of proteins that function as protein com-

plexes, the tools described in this book will become integrated into pharmacogenetic research. For example, tumor resistance to antiestrogenic therapy of breast cancer has long been recognized but not well understood at the mechanistic level. The past 2 years have given rise to significant advances in our understanding of estrogen receptor (ER)-mediated cell proliferation and mechanisms of antiestrogenic response. Those studies have revealed great complexity in ER signaling and transcriptional activation. It is clear that many coregulators are capable of forming protein complexes with ligand-bound ER (Clarke et al. 2001). The specific coregulators that comprise a given complex determine whether the complex activates or represses transcription of target genes. It has also been recognized that genetic variation in ERs or in proteins that interact with ERs might contribute to antiestrogen resistance (MacGregor and Jordan 1998; Clarke et al. 2001). This level of understanding sets the stage for elucidation of the mechanistic underpinnings of resistance to antiestrogen therapy, and those studies will undoubtedly involve the characterization of protein interactions.

Response to Anticancer Therapies

As indicated above, new clinical applications relating to DNA repair and drug-metabolizing genes are certain to emerge. Tumor cells often acquire resistance to radiation or therapeutic drugs. Genomics approaches, such as array technology, will be used to define any DNA repair genes that may be overexpressed in this context. Furthermore, it will be important to find ways to manipulate DNA repair and drug metabolism specifically or transport selectively to avoid precancerous lesions while promoting the activity of chemotherapeutic drugs. Genetic polymorphisms in relevant genes will be identified and efforts will be made to correlate them with effects on activity of the respective proteins, with response to particular therapies and with clinical outcomes. Although a number of polymorphisms in DNA repair and drug-metabolizing genes are being reported, there is little functional information on the consequence of the attendant amino acid changes. It will be important to find out which polymorphisms actually affect protein functions and then concentrate on these epidemiological and clinical studies. For example, homozygosity for a specific polymorphism in the DNA ligase subunit XRCC1 is associated with higher sister chromatid exchange frequencies in smokers, suggesting an association of this allele with a higher risk for tobacco- and age-related damage (Duell et al. 2000). Furthermore, with the use of gene and protein array techniques, it should be possible to compare expression profiles of DNA repair and drug-metabolizing genes in normal and tumor cells—information that could eventually lead to individually tailored therapies with chemicals and radiation. For example, tumors with low levels of nucleotide excision repair should be more susceptible to treatment with cisplatin (Claij and te Riele 1999). In comparison, mismatch-repair-deficient cells are highly tolerant to alkylating chemotherapeutic drugs (Koberle et al. 1999). The rapidly expanding knowledge of the human genome, coupled with automated methods for detecting gene polymorphisms, provides the tools to elucidate these polygenic determinants of drug effects, thus fueling the burgeoning field of pharmacogenomics. However, methods to evaluate the functional relevance of many of these polymorphisms on protein function will be paramount.

USE OF PROTEIN–PROTEIN INTERACTION STUDIES TO ANALYZE FUNCTIONAL CONSEQUENCES OF GENETIC VARIATION

As described above, study of protein interactors has contributed greatly to the general understanding of mechanisms of cancer genetics by helping to elucidate the function of many cancer-associated genes. Substantial investments are now being made within the pharmaceutical and biotechnology industries to use genomic and proteomic strategies for the discovery of therapeu-

tic targets. It is anticipated that, over the next decade, DNA and protein array technologies, high-throughput screening systems, enhanced mass spectrometry methods, and advanced bioinformatics will be merged to permit rapid elucidation of complex genetic components of human disease. The techniques described in this book are particularly important for correlating genotypic changes with specific biochemical or biophysical phenotypic changes associated with the variant proteins (genotype/phenotype correlation). Some techniques are amenable to the study of DNA interactions and have been applied as genotyping methodologies. The next pages provide examples of how these techniques have been applied or may be applied to the study of functional aspects of genetic variation.

Tagged Fusion Proteins

GST-fusion proteins, or similarly "tagged" proteins, are used extensively in pull-down assays, affinity purification, and far-western analyses (Chapter 4). These classic techniques have been used extensively in the study of protein–protein interactions, including proteins that represent drug targets and are genetically polymorphic. For example, the p53 tumor suppressor gene is one of the most frequently mutated genes in human tumors (Levine 1997). A common Pro72Arg polymorphism has been identified in the polyproline domain of the human p53 protein (Harris et al. 1986). Using GST pull-down assays, Thomas et al. (1999) have shown that the Pro-72 allozyme binds the transcriptional activators TAFII32 and TAFII70 with significantly greater affinity than does the Arg-72 allozyme. These authors further demonstrated that the Pro-72 allozyme is a better *trans*-activator of p53 *trans*-activated genes than the Arg-72 allozyme. These studies begin to determine a molecular mechanism by which the Pro72Arg polymorphism might confer a phenotypic change to the protein. These observations have allowed investigators to further hypothesize that this common polymorphism might influence the biology of the p53 protein.

Coimmunoprecipitation

Like the GST-fusion assays, coimmunoprecipitation (Chapter 5) has become a classic technique for the study of protein–protein interactions. This technique is a rigorous method of establishing physiologically relevant protein interactions. Coimmunoprecipitation and yeast two-hybrid approaches (as discussed below) have been the driving force behind the functional characterization of many of the so-called "orphan" proteins. As indicated above, the *BRCA1* and *BRCA2* breast/ovarian cancer susceptibility genes were identified in 1994 and 1995, respectively (Miki et al. 1994; Wooster et al. 1995; Tavitigian et al. 1996). The primary amino acid sequence of BRCA2 showed a weak similarity to BRCA1 over a restricted region (Wooster et al. 1995; Tavitigian et al. 1996), and only a low level of homology was seen among other known proteins. Not until Scully and colleagues demonstrated a direct interaction using coimmunoprecipitation between BRCA1 and RAD51 and BRCA2 and RAD51 were their potential roles in the cellular response to DNA damage realized (Scully et al. 1997b; Wong et al. 1997). Eukaryotic RAD51 proteins are homologs of bacterial RecA and are required for recombination during mitosis and meiosis and for recombinational repair of double-strand DNA breaks (Shinohara et al. 1992). Through yeast two-hybrid and biochemical assays, they demonstrated that the RAD51 protein interacts specifically with the eight evolutionarily conserved BRC motifs encoded in exon 11 of BRCA2 (Wong et al. 1997). Further coimmunoprecipitation studies using smaller portions of BRCA2 defined at least two additional RAD51-binding domains, residues 982–1066 and 1139–1266 (Katagira et al. 1998). These studies indicated that BRCA2 may interact with RAD51 through multiple sites of BRCA2 and that BRCA2 and RAD51 interact and colocalize in a BRCA1–BRCA2–RAD51 complex (Chen et al. 1997). BRCA2 has also been shown to form in vivo complexes with p53 (Marmorstein et al. 1998). Exogenous BRCA2 expression in cancer cells inhibits p53's transcriptional activity, and RAD51 coexpression enhances BRCA2's inhibitory effects. These findings, made possible using

protein interaction technologies, demonstrate that BRCA2 physically and functionally interacts with two key components of cell cycle control, one involving apoptosis and one involving DNA repair. These approaches have led to the identification of more than a dozen proteins that associate with specific domains of BRCA1 and/or BRCA2 (Fig. 5).

These studies are also helping to predict which missense changes may be deleterious mutations on the basis of their location in the protein and possible effects on protein–protein interactions. Understanding protein–protein interactions may also lead to a better understanding of cancer risks. For example, BRCA2 proteins have several significant repeated motifs that are not found in BRCA1 but are conserved in all mammalian BRCA2 proteins that have been sequenced. Eight internal repeats of 30–80 amino acids, known as BRC motifs, are encoded in exon 11 of the human BRCA2 gene (Fig. 5). Each repeat is variably conserved, suggesting that the core sequence was duplicated eight times during evolution, but that many of the repeats are now redundant (Chen et al. 1995). The BRC motifs were later demonstrated to be involved in BRCA2-mediated DNA repair through their interactions with RAD51 (Wong et al. 1997). The BRC repeats lie within a large region spanning exon 11 that has been deemed the ovarian cancer cluster region (OCCR), because mutations in this region have been associated with an increased frequency of ovarian cancer (Gayther et al. 1997; Neuhausen et al. 1998).

Although coimmunoprecipitation can be cumbersome for the detection of novel protein interactions, it is ideally suited for studying genetic variation causing loss or gain of function within known protein interactions. For example, Marin et al. (2000) have used this technique to demonstrate that a mutant form of p53 with Arg at codon 72 has a higher affinity for p73 (a p53 homolog and binding protein) than does the mutant Pro-72 allozyme. These observations again suggest biological significance of the common p53 polymorphism, interestingly, within the context of independent p53 mutations. Furthermore, the ability of mutant p53 to block the chemotherapeutic activity of etoposide correlates with the ability of the protein to bind p73 (Blandino et al. 1999), and p73 plays a role in cisplatin-induced apoptosis (Gong et al. 1999). Collectively, these observations suggest that the common Pro72Arg p53 polymorphism might be associated with chemosensitivity of selected tumors (Marin et al. 2000).

Yeast Two-hybrid

Although most often used to identify novel partners for a known protein, the yeast two-hybrid system (Chapter 7) can be adapted to accommodate the study of variant protein interactions. For example, Maier et al. (1998) have used the yeast two-hybrid approach to analyze the genotype–phenotype correlation among mouse aromatic hydrocarbon receptor (AHR) alleles. The AHR is a transcriptional activator of several drug-metabolizing genes. Translocation of the AHR to the nucleus depends on its interaction with a small-molecule ligand and the ARNT protein (Gu et al. 2000). Affinity of the AHR allozymes for the ARNT protein, in the presence of an AHR ligand such as dioxin, was quantified with a β-galactosidase reporter system (Maier et al. 1998). These authors reported a 15-fold difference in the ligand affinity of two mouse AHR alleles, suggesting biological significance of the polymorphisms. Subsequently, the functional significance of human AHR genetic variation has been demonstrated in studies associating AHR genotype with level of expression of AHR target genes (Smart and Daly 2000).

More recent innovations to the yeast two-hybrid system offer additional promise for pharmacogenetic research. For example, the three-hybrid system includes "hook," "bait," and "fish" components for detecting small ligand–protein receptor interactions (Licitra and Liu 1996). This represents a system in which the genotype–phenotype correlation might be analyzed for small molecule binding proteins such as the drug-metabolizing enzymes. In addition, the reverse two-hybrid system, which selects for interference of a specific protein–protein interaction, may facilitate the detection of variant proteins in which there is a loss of function (Leanna and Hannink 1996).

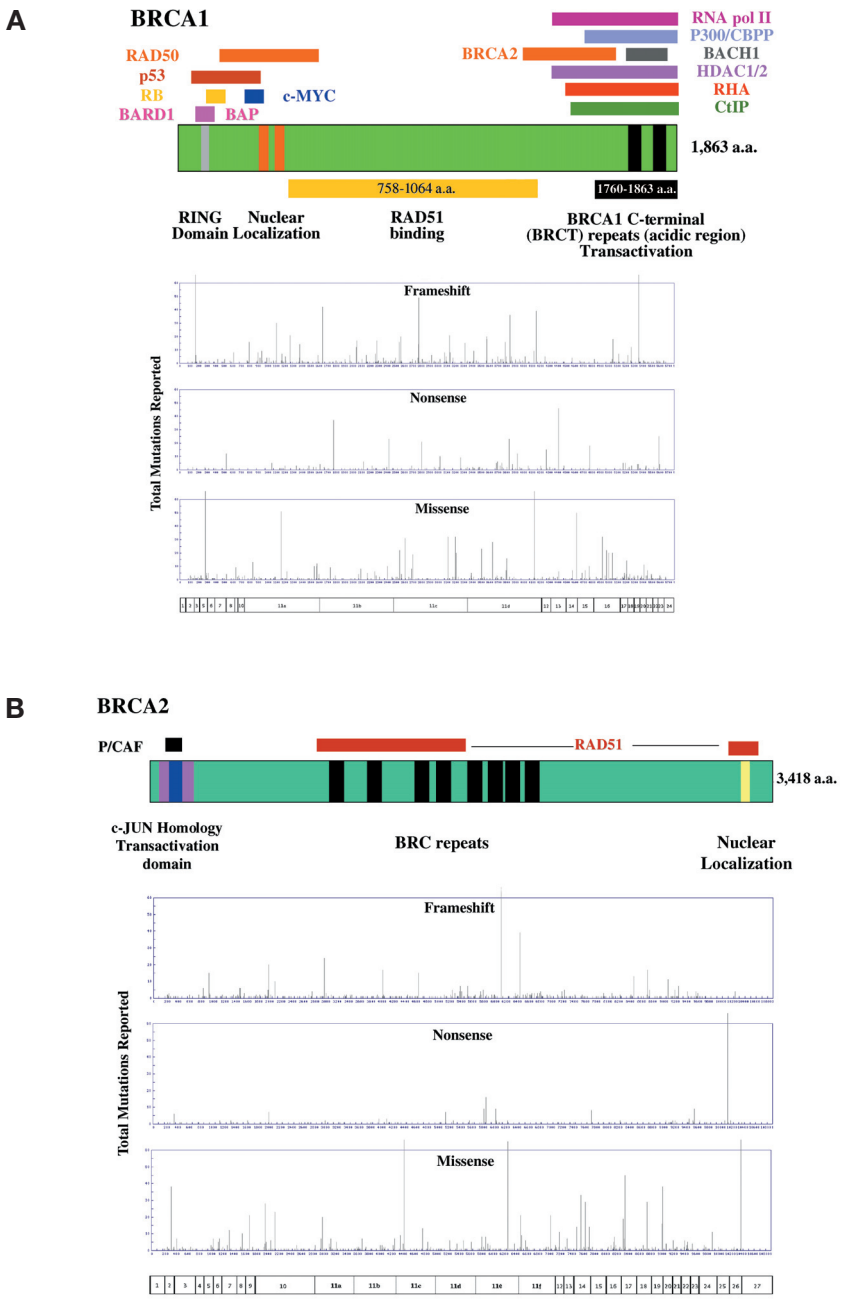

FIGURE 5. Schematic of BRCA1 (*A*) and BRCA2 (*B*) and their interacting proteins. Transcription-coupled repair appears to involve BRCA1 through its association with many proteins, including RNA pol II, acetyl transferases, and deacetylases (P300/CPB, P/CAF, ACRT/SCR-1). During DNA replication, BRCA1 is associated with the replication protein PCNA (see Fig. 3) and with Rad50 and Rad51 during homologous recombinational repair. (BARD) BRCA1-associated RING domain 1 protein; (BAP1) BRCA1-associated protein-1 (ubiquitin conjugating enzymes); (BRCT) BRCA1 C-terminal repeats; (HDAC) histone deacetylase 1 and 2; (RHA) RNA helicase A; (CtIP) CtBp-interacting protein; (BACH1) BRCA1-associated C-terminal helicase, a member of the DEAH helicase; (P/CAF) p300/CBP-associated factor (with histone acetylase activity). The lower panels represent the position and the number of times each of the disease-associated mutations (frameshift, nonsense, and missense) has been reported by BIC (see www.nhgri.nih.gov/Intramural_research/Lab_transfer/Bic/index.html).

FRET

Fluorescence resonance energy transfer (FRET) measures molecular proximity by detecting the excited state of acceptor and donor fluorophore-tagged molecules. This technique, as described in Chapter 10, has been applied to measure protein–protein interactions by alternatively labeling separate proteins of interest with donor or acceptor fluorophores and allows the physiological validation of proposed protein–protein interactions in real time. Intriguingly, the application of this technique to pharmacogenetic research is best exemplified by the dye-terminator incorporation genotyping assay, which detects oligonucleotide proximity rather than protein proximity (Chen and Kwok 1999). In this approach, an allele-specific, donor dye-labeled oligonucleotide is extended by DNA polymerase in PCR using acceptor dye-labeled dNTPs. Once the donor and acceptor dyes become part of a new molecule (via allele-specific PCR amplification), intramolecular FRET is detected and quantified. Another FRET-associated genotyping assay, the Invader assay, was recently developed (Lyamichev et al. 1999; Mein et al. 2000). The Invader assay relies on the specificity of recognition and cleavage, by a flap endonuclease (FEN), of overlapping "signal" and "invader" oligonucleotides that hybridize to target DNA containing a polymorphic site. In the presence of a match between the signal oligonucleotide and the DNA template, the signal oligonucleotide is cleaved to drive a secondary cleavage reaction with a FRET label. The signal is detected with a fluorescence plate reader. An advantage of this technology is the high-throughput flexibility of the assay as well as its sensitivity.

Biacore

Surface plasmon resonance (Chapter 14) enables the detection and quantification of bonding events between molecules by monitoring changes in the refractive index of molecular surfaces upon intermolecular interaction. This technology is useful in obtaining kinetic data regarding the interaction between proteins, lipids, nucleotides, and small molecules. Because of this versatility, there is great potential for the application of this technique to the study of phenotypic changes associated with genetic variation. For example, Baynes et al. (2000) have used Biacore technology to study the binding characteristics of wild-type and mutant human p85α proteins to lipid substrates. p85α is a PI3 kinase involved in insulin-stimulated glucose disposal. These investigators identified a common p85α polymorphism as well as one mutation in subjects with severe insulin resistance. Cells expressing the polymorphic variant of p85α maintained similar PI3 kinase activity toward lipid substrates as cells expressing the wild-type protein. However, cells expressing the mutant p85α exhibited significantly lower levels of PI3 kinase activity. These authors then used the Biacore technique to determine whether the mutant protein was defective in binding lipid substrates. Their data demonstrated that wild-type, polymorphic, and mutant p85α each have similar binding capacity, thus suggesting that the mechanism by which the mutant protein exerted a phenotypic change was not via altered substrate binding.

Atomic Force Microscopy

AFM (Chapter 13) maps the topography of a surface by detecting minute changes in bonding forces at the surface of a molecule. This technique is often used to study ligand–receptor interactions. Yip et al. (1998) have used AFM to demonstrate that a mutant form of the human insulin protein forms much weaker self-associated oligomeric structures than does wild-type insulin. Consequently, in solution, a higher percentage of mutant insulin adopted monomeric conformations as opposed to hexamers. These observations subsequently contributed to the development of insulin therapies with better bioavailability because the efficacy of commercial insulin preparations is dependent on the dissolution of insulin hexamers to monomers.

Similar to Biacore technologies, AFM has been applied as a technique for high-throughput genotyping. One of the most difficult aspects of genotyping is the assignment of haplotypes

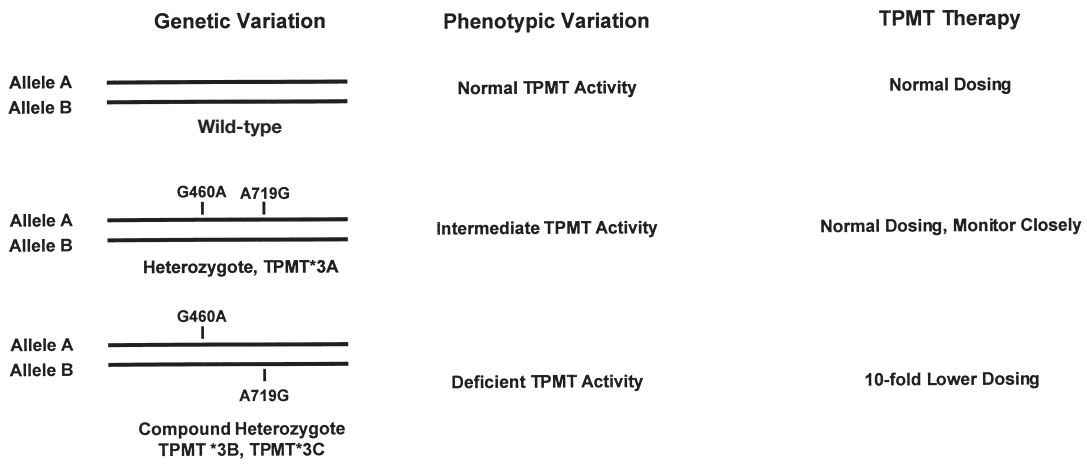

FIGURE 6. Haplotype–phenotype correlation with thiopurine methyltransferase. The figure illustrates the importance of being able to distinguish technically between individuals who carry one copy of the TPMT*3A allele (and hence would not need dose modification) and those who are compound heterozygotes expressing one copy of the *3B and one of the *3C allele (and hence would need very low doses of thiopurines). The *3A allele is defined by two SNPs, a G→A transition at nucleotide 460 and an A→G transition at nucleotide 719 in the cDNA. These two positions are separated by 13 kb of sequence on human chromosome 6. The *3B allele is defined only by the SNP at position 460, and the *3C allele by the SNP at position 719.

(Service 2000). The ability to assign a haplotype is sometimes paramount to clinical decisions. For example, the TPMT polymorphic phenotype described previously in this chapter is encoded by several *TPMT* alleles. One of those alleles, *3A, is defined by two separate SNPs at loci within the gene that are separated by more than 13 kb of DNA sequence. However, each of these SNPs also defines separate alleles (*3B and *3C, respectively) that each encode deficient TPMT enzyme. In the clinical setting, it would be critical to distinguish between an individual who is heterozygous for the *3A allele and one who is a compound heterozygote (*3B/*3C). That is true because in the former case, the individual would be predicted to have an "intermediate" TPMT phenotype and could be prescribed standard doses of mercaptopurine drugs. However, the individual who is a *3B/*3C compound heterozygote would be TPMT deficient and would likely experience profound myelosuppression with that same dose (Fig. 6). Woolley et al. (2000) have described modification of the AFM technique, using carbon nanotube probes, to determine genomic haplotypes. This approach involves the hybridization of fluorescently labeled, allele-specific oligonucleotide probes to target DNA fragments, followed by detection of the presence and spatial location of the labels by AFM. A major advantage of this methodology is the ability to codetect SNPs as far as several kilobases (perhaps 100 kb) apart and its amenability to high-throughput genotyping (Service 2000).

CONCLUSIONS

Susceptibility to cancer and the development of effective approaches to treat these diseases will depend greatly on our ability to decipher the function of hundreds, maybe thousands, of proteins. With the sequencing of the human genome and the identification of thousands of previously uncharacterized genes now a reality, the work to elucidate the function of the proteins they encode is just beginning. The number of proteins directly or indirectly involved in cancer susceptibility, cancer progression, and response to drug therapy is difficult, if not impossible, to predict.

If one were to consider just the oncogenes, tumor suppressor genes, and DNA repair genes as the sole genes associated with cancer,the number might approach more than 500. However, it is clear that not all of these genes have a role in human disease and that they are clearly not the only ones that contribute to the risk of disease. The class of genes involved in metabolic detoxification pathways and steroid and amino acid metabolism, for example, is quite large and is likely to be intimately involved in many facets of cancer. Furthermore, it is likely that a small but significant fraction of the several million SNPs present in any genome will result in multiple forms of the same protein with slightly modified biochemical and biological properties. Therefore, new genomic and proteomic technologies will be necessary to identify important genomic risk factors and translate these findings into a functional reality. In this book, the authors have presented new and exciting advances in the area of protein interaction technologies that will help to move forward this area of research. Specific protein–protein and protein–ligand interactions are central to most biological processes and are the focus of many avenues of research to develop small molecule-based therapies that will disrupt these essential interactions. We are only beginning the post-genomics era. Therefore, we can look forward to an equally exciting decade of great discoveries and challenges that will be necessary to define the function of thousands of new proteins as well as their potentially important variant forms and allozymes.

REFERENCES

Baynes K.C., Beeton C.A., Panayotou G., Stein R., Soos M., Hansen T., Simpson H., O'Rahilly S., Shepherd P.R., and Whitehead J.P. 2000. Natural variants of human p85 alpha phosphoinositide 3-kinase in severe insulin resistance: A novel variant with impaired insulin-stimulated lipid kinase activity. *Diabetologia* **43:** 321–331.

Bell D.W., Varley J.M., Szydlo T.E., Kang D.H., Wahrer D.C., Shannon K.E., Lubratovich M., Verselis S.J., Isselbacher K.J., Fraumeni J.F., Birch J.M., Li F.P., Garber J.E., and Haber D.A. 1999. Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome. *Science* **286:** 2528–2531.

Blandino G., Levine A.J., and Oren M. 1999. Mutant p53 gain of function: Differential effects of different p53 mutants on resistance of cultured cells to chemotherapy. *Oncogene* **18:** 477–485.

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.

Chen H., Lin R., Schiltz R., Chakravarti D., Nash A., Nagy L., Privalsky M., Nakatani Y., and Evans R. 1997. Nuclear receptor coactivaor ACTR is a novel histone acetyltransferase and forms a multimeric activation complex with P/CAF and CPB/p300. *Cell* **90:** 569–580.

Chen X. and Kwok P.Y. 1999. Homogeneous genotyping assays for single nucleotide polymorphisms with fluorescence resonance energy transfer detection. *Genet. Anal.* **14:** 157–163.

Chen Y., Chen C.-F., Riley D.J., Allred D.C., Chen P.-L., von Hoff D., Osborne C.K., and Lee W.-H. 1995. Aberrant subcellular localization of BRCA1 in breast cancer. *Science* **270:** 789–791.

Claij N. and te Riele H. 1999. Microsatellite instability in human cancer: A prognostic marker for chemotherapy? *Exp. Cell Res.* **246:** 1–10.

Clarke R., Leonessa F., Welch J.N., and Skaar T.C. 2001. Cellular and molecular pharmacology of antiestrogen action and resistance. *Pharmacol. Rev.* **53:** 25–71.

Cortez D., Wang Y., Qin J., and Elledge S. 1999. Requirement of ATM-dependent phosphorylation of brca1 in the DNA damage response to double-strand breaks. *Science* **286:** 1162–1166.

DeMars R. 1969. Genetics concepts and neoplasia; a collection of papers. In *23rd Annual Symp. Fundamental Cancer Research* (ed. W.A.Wilkins), pp. 105–106, Baltimore.

Dobner T., Horikoshi N., Rubenwolf S., and Shenk T. 1996. Blockage by adenovirus E4orf6 of transcriptional activation by the p53 tumor suppressor. *Science* **272:** 1470–1473.

Duell E.J., Wiencke J.K., Cheng T.J., Varkonyi A., Zuo Z.F., Ashok T.D., Mark E.J., Wain J.C., Christiani D.C., and Kelsey K.T. 2000. Polymorphisms in the DNA repair genes XRCC1 and ERCC2 and biomarkers of DNA damage in human blood mononuclear cells. *Carcinogenesis* **21:** 965–971.

Dunn J.M., Phillips R.A., Zhu X., Becker A., and Gallie B.L. 1989. Mutations in the RB1 gene and their effects on transcription. *Mol. Cell. Biol.* **9:** 4596–4604.

El-Deiry W.S., Harper J.W., O'Conner P.M., Velculescu V.E., Canman C.E., Jackman J., Peintenpol J.A.,

Burrell M., Hill D.E., Wang Y., Wilman K.G., Mercer W.E., Kastan M.B., Konh K.W., Elledge S.J., Kinzler K.W., and Vogelstein B. 1994. WAF1/CIP1 is induced in p53-mediated G1 arrest and apoptosis. *Cancer Res.* **54:** 1169–1174.

Evans W.E. and Relling M.V. 1999. Pharmacogenomics: Translating functional genomics into rational therapeutics. *Science* **286:** 487–491.

Falck J., Mailand N., Syljuasen R.G., Bartek J., and Lukas J. 2001. The ATM-Chk2-Cdc25A checkpoint pathway guards against radioresistant DNA synthesis. *Nature* **410:** 842–847.

Fearon E.R. 1997. Human cancer syndromes: Clues to the origin and nature of cancer. *Science* **278:** 1043–1050.

———. 2000. BRCA1 and E-cadherin promoter hypermethylation and gene inactivation in cancer-association or mechanism? *J. Natl. Cancer Inst.* **92:** 515–517.

Galaktionov K., Lee A.K., Eckstein J., Draetta G., Meckler J., Loda M., and Beach D. 1995. CDC25 phosphatases as potential human oncogenes. *Science* **269:** 1575–1577.

Gannon J. and Lane D. 1990. Interactions between SV40 T antigen and DNA polymerase alpha. *New Biol.* **2:** 84–92.

Gayther S., Mangion J., Russell P., Seal S., Barfoot R., Ponder B., Stratton M., and Easton D. 1997. Variation of risks of breast and ovarian cancer associated with different germline mutations of the BRCA2 gene. *Nat. Genet.* **15:** 103–105.

Godwin A.K., Schultz D.C., Hamilton T.C., and Knudson A.G. 1997. Oncogenes and tumor suppressor genes. In *Gynecologic oncology: Principles and practice* (ed. W.J. Hoskins et al.), pp. 107–148. Lippencott, Philadelphia.

Gong J.G., Costanzo A., Yang H.Q., Melino G., Kaelin W.G., Jr., Levrero M., and Wang J.Y. 1999. The tyrosine kinase c-Abl regulates p73 in apoptotic response to cisplatin-induced DNA damage (see comments). *Nature* **399:** 806–809.

Gottlieb T.M. and Oren M. 1996. p53 in growth control and neoplasia. *Biochim. Biophys. Acta* **1287:** 77–102.

Greenlee R.T., Murray T., Bolden S., and Wingo P.A. 2000. Cancer statistics, 2000. *CA Cancer J. Clin.* **50:** 7–33.

Gu Y.Z., Hogenesch J.B., and Bradfield C.A. 2000. The PAS superfamily: Sensors of environmental and developmental signals. *Annu. Rev. Pharmacol. Toxicol.* **40:** 519–561.

Hainaut P., Hernandez T., Robinson A., Rodriguez-Tome P., Flores T., Hollstein M., Harris C.C., and Montesano R. 1998. IARC Database of p53 gene mutations in human tumors and cell lines: Updated compilation, revised formats and new visualisation tools. *Nucleic Acids Res.* **26:** 205–213.

Hanahan D. and Weinberg R.A. 2000. The hallmarks of cancer. *Cell* **100:** 57–70.

Harris H., Miller O.J., Klein G., Worst P., and Tachibana T. 1969. Suppression of malignancy by cell fusion. *Nature* **223:** 363–368.

Harris N., Brill E., Shohat O., Prokocimer M., Wolf D., Arai N., and Rotter V. 1986. Molecular basis for heterogeneity of the human p53 protein. *Mol. Cell. Biol.* **6:** 4650–4656.

Hollstein M., Shomer B., Greenblatt M., Soussi T., Hovig E., Montesano R., and Harris C.C. 1996. Somatic point mutations in the p53 gene of human tumors and cell lines: Updated compilation. *Nucleic Acids Res.* **24:** 141–146.

Ingelman-Sundberg M., Oscarson M., and McLellan R.A. 1999. Polymorphic human cytochrome P450 enzymes: An opportunity for individualized drug treatment. *Trends Pharmacol. Sci.* **20:** 342–349.

Iyer L., Hall D., Das S., Mortell M.A., Ramirez J., Kim S., Di Rienzo A., and Ratain M.J. 1999. Phenotype-genotype correlation of in vitro SN-38 (active metabolite of irinotecan) and bilirubin glucuronidation in human liver tissue with UGT1A1 promoter polymorphism. *Clin. Pharmacol. Ther.* **65:** 576–582.

Katagira T., Saito H., Shinohara A., Ogawa H., Kamada N., Nakamura Y., and Miki Y. 1998. Multiple possible sites of BRCA2 interacting with DNA repair protein RAD51. *Genes Chromosomes Cancer* **21:** 217–222.

Knudson A.G. 1971. Mutation and cancer: Statistical study of retinoblastoma. *Proc. Natl. Acad. Sci.* **68:** 820–823.

Knudson A.G. and Strong L.C. 1972. Mutation and cancer: A model for Wilms' tumor of the kidney. *J. Natl. Cancer Inst.* **48:** 313–324.

Ko L.J. and Prives C. 1996. p53: Puzzle and paradigm. *Genes Dev.* **10:** 1054–1072.

Koberle B., Masters J.R., Hartley J.A., and Wood R.D. 1999. Defective repair of cisplatin-induced DNA damage caused by reduced XPA protein in testicular germ cell tumours. *Curr. Biol.* **9:** 273–276.

Krynetski E.Y. and Evans W.E. 1998. Pharmacogenetics of cancer therapy: Getting personal. *Am. J. Hum. Genet.* **63:** 11–16.

Krynetski E.Y., Tai H.L., Yates C.R., Fessing M.Y., Loennechen T., Schuetz J.D., Relling M.V., and Evans W.E. 1996. Genetic polymorphism of thiopurine S-methyltransferase: Clinical importance and molecular mechanisms. *Pharmacogenetics* **6:** 279–290.

Kuehl P., Zhang J., Lin Y., Lamba J., Assem M., Schuetz J., Watkins P.B., Daly A., Wrighton S.A., Hall S.D.,

Maurel P., Relling M., Brimer C., Yasuda K., Venkataramanan R., Strom S., Thummel K., Boguski M.S., and Schuetz E. 2001. Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nat. Genet.* **27:** 383–391.

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.

Leanna C.A. and Hannink M. 1996. The reverse two-hybrid system: A genetic scheme for selection against specific protein/protein interactions. *Nucleic Acids Res.* **24:** 3341–3347.

Lee J.-S., Collins K.M., Brown A.L., Lee C.-H., and Chung J.H. 2000. hCds1-mediated phosphorylation of BRCA1 regulates the DNA damage response. *Nature* **404:** 201–204.

Lennard L. 1997. Methyl transferases. In *Comprehensive toxicology* (ed. F.P. Guengerich), pp. 437–454, Pergamon, London.

Levine A.J. 1997. p53, The cellular gatekeeper for growth and division. *Cell* **88:** 323–331.

Licitra E.J. and Liu J.O. 1996. A three-hybrid system for detecting small ligand-protein receptor interactions. *Proc. Natl. Acad. Sci.* **93:** 12817–12821.

Lill N.L., Grossman S.R., Ginsberg D., DeCaprio J., and Livingston D.M. 1997. Binding and modulation of p53 by p300/CBP coactivators. *Nature* **387:** 823–827.

Lindahl T. and Wood R.D. 1999. Quality control by DNA repair. *Science* **286:** 1897–1905.

Lundberg A.S. and Weinberg R.A. 1999. Control of the cell cycle and apoptosis. *Eur. J. Cancer* **35:** 1886–1894.

Lyamichev V., Mast A.L., Hall J.G., Prudent J.R., Kaiser M.W., Takova T., Kwiatkowski R.W., Sander T.J., de Arruda M., Arco D.A., Neri B.P., and Brow M.A. 1999. Polymorphism identification and quantitative detection of genomic DNA by invasive cleavage of oligonucleotide probes. *Nat. Biotechnol.* **17:** 292–296.

Lynch H. and Smyrk T. 1996. Hereditary nonpolyposis colorectal cancer (Lynch syndrome): An updated review. *Cancer* **78:** 1149–1167.

MacGregor J.I. and Jordan V.C. 1998. Basic guide to the mechanisms of antiestrogen action. *Pharmacol. Rev.* **50:** 151–196.

Maier A., Micka J., Miller K., Denko T., Chang C.Y., Nebert D.W., and Alvaro P. 1998. Aromatic hydrocarbon receptor polymorphism: Development of new methods to correlate genotype with phenotype. *Environ. Health Perspect.* **106:** 421–426.

Malkin D., Li F., Strong L., Fraumeni J., Nelson C., Kim D., Kassel J., Gryka M., Bisehoff F., Tainsky M., and Friend S. 1990. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* **250:** 1233–1238.

Mansur C.P., Marcus B., Dalal S., and Androphy E.J. 1995. The domain of p53 required for binding HPV 16 E6 is separable from the degradation domain. *Oncogene* **10:** 457–465.

Marin M.C., Jost C.A., Brooks L.A., Irwin M.S., O'Nions J., Tidy J.A., James N., McGregor J.M., Harwood C.A., Yulug I.G., Vousden K.H., Allday M.J., Gusterson B., Ikawa S., Hinds P.W., Crook T., and Kaelin W.G., Jr. 2000. A common polymorphism acts as an intragenic modifier of mutant p53 behaviour. *Nat. Genet.* **25:** 47–54.

Marmorstein L.Y., Ouchi T., and Aaronson S.A. 1998. The *BRCA2* gene product functionally interacts with p53 and RAD51. *Proc. Natl. Acad. Sci.* **95:** 13869–13874.

Marshall C.J. 1991. Tumor suppressor genes. *Cell* **64:** 313–326.

Mein C.A., Barratt B.J., Dunn M.G., Siegmund T., Smith A.N., Esposito L., Nutland S., Stevens H.E., Wilson A.J., Phillips M.S., Jarvis N., Law S., de Arruda M., and Todd J.A. 2000. Evaluation of single nucleotide polymorphism typing with invader on PCR amplicons and its automation. *Genome Res.* **10:** 330–343.

Miki Y., Swensen J., Shattuck-Eidens D., Futreal P., Harshman K., Tavtigian S., Liu Q., Cochran C., Bennett L., Ding W., et al. 1994. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* **266:** 66–71.

Neuhausen S.L., Godwin A.K., Gershoni-Baruch R., Schubert E., Garber J., Stoppa-Lyonnet D., Olah E., Csokay B., Serova O., Lalloo F., et al. 1998. Haplotype and phenotype analysis of nine recurrent BRCA2 mutations in 111 families: Results of an international study. *Am. J. Hum. Genet.* **62:** 1381–1388.

Nie Y., Li H.H., Bula C.M., and Liu X. 2000. Stimulation of p53 DNA binding by c-Abl requires the p53 C terminus and tetramerization. *Mol. Cell. Biol.* **20:** 741–748.

Otterness D., Szumlanski C., Lennard L., Klemetsdal B., Aarbakke J., Park-Hah J.O., Iven H., Schmiegelow K., Branum E., O'Brien J., and Weinshilboum R. 1997. Human thiopurine methyltransferase pharmacogenetics: Gene sequence polymorphisms. *Clin. Pharmacol. Ther.* **62:** 60–73.

Peltomaki P. and Vasen H. 1997. Mutations predisposing to hereditary nonpylyposis colorectal cancer: Database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. *Gastroenterology* **113:** 1146–1158.

Poulsen L., Brosen K., Arendt-Nielsen L., Gram L.F., Elbaek K., and Sindrup S.H. 1996. Codeine and morphine in extensive and poor metabolizers of sparteine: pharmacokinetics, analgesic effect and side effects. *Eur. J. Clin. Pharmacol.* **51:** 289–295.

Prowse A., Frolov A., and Godwin A.K. 2001. The genetics of ovarian cancer. In *American Cancer Society*

Atlas of clinical oncology. B.C. Decker, Hamilton, Ontario. (In press.)
Rebbeck T.R., Jaffe J.M., Walker A.H., Wein A.J., and Malkowicz S.B. 1998. Modification of clinical presentation of prostate tumors by a novel genetic variant in CYP3A4. *J. Natl. Cancer Inst.* **90:** 1225–1229.
Rendic S. and Di Carlo F.J. 1997. Human cytochrome P450 enzymes: A status report summarizing their reactions, substrates, inducers, and inhibitors. *Drug Metab. Rev.* **29:** 413–580.
Rettie A.E., Haining R.L., Bajpai M., and Levy R.H. 1999. A common genetic basis for idiosyncratic toxicity of warfarin and phenytoin. *Epilepsy Res.* **35:** 253–255.
Rothman N., Wacholder S., Caporaso N.E., Garcia-Closas M., Buetow K., and Fraumeni J.F. 2001. The use of common genetic polymorphisms to enhance the epidemiologic study of environmental carcinogens. *Biochim. Biophys. Acta* **1471:** C1–10.
Rotman G. and Shiloh Y. 1999. ATM: A mediator of multiple responses to genotoxic stress. *Oncogene* **18:** 6135–6144.
Ruffner H., Jiang H., Craig A., Hunter T., and Verman I. 1999. BRCA1 is phosphorylated at serine 1497 in vivo at a cyclin-dependent kinase 2 phosphorylation site. *Mol. Cell. Biol.* **19:** 4843–4854.
Scully R., Chen J., Ochs R.L., Keegan K., Hoekstra M., Feunteun J., and Livingston D.M. 1997a. Dynamic changes of BRCA1 subnuclear location and phosphorylation state are initiated by DNA damage. *Cell* **90:** 425–435.
Scully R., Chen J., Plug A., Xiao Y., Weaver D., Feunteun J., Ashley T., and Livingston D.M. 1997b. Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* **88:** 265–275.
Service R.F. 2000. DNA imaging. Getting a feel for genetic variations (news). *Science* **289:** 27–28.
Shields P.G. and Harris C.C. 2000. Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. *J. Clin. Oncol.* **18:** 2309–2315.
Shinohara A., Ogawa H., and Ogawa T. 1992. Rad51 protein involved in repair and recombination in *Saccharomyces cerevisiae* is a recA-like protein. *Cell* **69:** 457–470.
Smart J., and Daly A.K. 2000. Variation in induced CYP1A1 levels: Relationship to CYP1A1, Ah receptor and GSTM1 polymorphisms. *Pharmacogenetics* **10:** 11–24.
Soussi T., Caron de Fromentel C., and May P. 1990. Structural aspects of the p53 protein in relation to gene evolution. *Oncogene* **5:** 945–952.
Stanbridge E.J. 1976. Suppression of malignancy in human cells. *Nature* **260:** 17–20.
Tai H.L., Krynetski E.Y., Schuetz E.G., Yanishevski Y., and Evans W.E. 1997. Enhanced proteolysis of thiopurine S-methyltransferase (TPMT) encoded by mutant alleles in humans (TPMT*3A, TPMT*2): Mechanisms for the genetic polymorphism of TPMT activity. *Proc. Natl. Acad. Sci.* **94:** 6444–6449.
Tavitigian S., Rommens J., Couch F., Shattuck-Eidens D., Neuhausen S., Merajver S., Thorlacius S., Offit K., Stoppa-Lyonnet D., Belanger C., et al. 1996. The complete *BRCA2* gene and mutations in chromosomes 13q-linked kindreds. *Nat. Genet.* **12:** 333–337.
Thomas J.E., Smith M., Tonkinson J., Rubinfeld B., and Polakis P. 1997. Induction of phosphorylation on *BRCA1* during the cell cycle and after DNA damage. *Cell Growth Differ.* **8:** 801–809.
Thomas M., Kalita A., Labrecque S., Pim D., Banks L., and Matlashewski G. 1999. Two polymorphic variants of wild-type p53 differ biochemically and biologically. *Mol. Cell. Biol.* **19:** 1092–1100.
Tibbetts R., Cortez D., Brumbaugh K.M., Scully R., Livingston D., Elledge S.J., Abraham R.T. 2000. Functional interactions between BRCA1 and the checkpoint kinase ATR during genotoxic stress. *Genes Dev.* **14:** 2989–3002.
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.
Wang Y., Cortez D., Yazdi P., Neff N., Elledge S.J., and Qui J. 2000. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev.* **14:** 927–939.
Weber W.W. 1997. *Pharmacogenetics*. Oxford Press, New York.
Weinshilboum R.M. and Sladek S.L. 1980. Mercaptopurine pharmacogenetics: Monogenic inheritence of erythrocyte thiopurine methyltransferase activity. *Am. J. Hum. Genet.* **32:** 651–662.
Wong A., Pero R., Ormonde P., Tavigian S., and Bartel P. 1997. Rad51 interacts with the evolutionarily conserved BRC motifs in the human breast cancer susceptibility gene *BRCA2*. *J. Biol. Chem.* **272:** 31941–31944.
Wood R.D., Mitchell M., Sgouros J., and Lindahl T. 2001. Human DNA repair genes. *Science* **291:** 1284–1289.
Woolley A.T., Guillemette C., Li Cheung C., Housman D.E., and Lieber C.M. 2000. Direct haplotyping of kilobase-size DNA using carbon nanotube probes (see comments). *Nat. Biotechnol.* **18:** 760–763.
Wooster R., Bignell G., and Lancaster J. 1995. Identification of the breast cancer susceptibility gene BRCA2. *Nature* **378:** 789–792.
Wormhoudt L.W., Commandeur J.N., and Vermeulen N.P. 1999. Genetic polymorphisms of human N-acetyltransferase, cytochrome P450, glutathione-S-transferase, and epoxide hydrolase enzymes: Relevance to xenobiotic metabolism and toxicity. *Crit. Rev. Toxicol.* **29:** 59–124.

Yip C.M., Brader M.L., DeFelippis M.R., and Ward M.D. 1998. Atomic force microscopy of crystalline insulins: The influence of sequence variation on crystallization and interfacial structure. *Biophys. J.* **74:** 2199–2209.

Zhang H., Somasundaram K., Peng Y., Tian H., Zhang H., Bi D., Weber B.L., and El-Deiry W.S. 1998. BRCA1 physically associates with p53 and stimulates its transcriptional activity. *Oncogene* **16:** 1713–1721.

Zilfou J.T., Hoffman W.H., Sank M., George D.L., and Murphy M. 2001. The co-repressor Sin3 directly interacts with the proline-rich domain of p53 and protects p53 from proteosome-mediated degradation. *Mol. Cell. Biol.* **21:** 3974–3985.