Molecular and Genetic Epidemiology of Cancer in Low- and Medium-Income Countries

Jyoti Malhotra, MD, MPH

ABSTRACT

Background: Genetic and molecular factors can play an important role in an individual's cancer susceptibility and response to carcinogen exposure. Cancer susceptibility and response to carcinogen exposure can be either through inheritance of high penetrance but rare germline mutations that constitute heritable cancer syndromes, or it can be inherited as common genetic variations or polymorphisms that are associated with low to moderate risk for development of cancer. These polymorphisms can interact with environmental exposures and can influence an individual's cancer risk through multiple pathways, including affecting the rate of metabolism of carcinogens or the immune response to these toxins. Thus, these genetic polymorphisms can account for some of the geographical differences seen in cancer prevalence between different populations.

Objectives: This review explores the role of molecular epidemiology in the field of cancer prevention and control in low- and medium-income countries.

Findings: Using data from Human Genome Project and HapMap Project, genome-wide association studies have been able to identify multiple susceptibility loci for different cancers. The field of genetic and molecular epidemiology has been further revolutionized by the discovery of newer, faster, and more efficient DNA-sequencing technologies including next-generation sequencing.

Conclusions: The new DNA-sequencing technologies can play an important role in planning and implementation of cancer prevention and screening strategies. More research is needed in this area, especially in investigating new biomarkers and measuring gene-environment interactions.

Key Words: cancer, genetics, low-income countries, molecular epidemiology

© 2014 Icahn School of Medicine at Mount Sinai. Annals of Global Health 2014;80:418-425

INTRODUCTION

Genetic and inherited factors play a significant role in predicting an individual's susceptibility to cancer.[1] The risk for nasopharyngeal cancer is 4- to 10-fold among individuals with a first-degree relative with the cancer.[2] Similarly, first-degree relatives of breast cancer patients have a 2-fold higher risk for developing breast cancer[3,4] and, a family history of pancreatic cancer is associated with a 9-fold increased risk compared with the general population.[5] Carcinogenesis is considered to be a multistep, evolving process that involves interaction between genetic and environmental factors.[6,7]

Molecular or genetic epidemiology is the discipline of epidemiological study that focuses on the contribution of genetic and environmental risk factors to cancer risk and the interplay between them at the cellular and molecular levels.[8,9] A major breakthrough in this field was the sequencing of complete human DNA by The Human Genome Project.[10] This was followed by the creation of genome-wide databases of common genetic variations in humans. This information, in addition to recent advances in genotyping technology, has provided the foundation to more comprehensively examine genetic variations in relation to cancer risk using genome-wide association studies (GWAS; Fig. 1). GWAS have identified loci for multiple new cancer-susceptibility genes[11] and use the design of observational epidemiology studies (case-control and cohort studies) to determine association of low penetrance alleles with cancer risk by calculating odds ratio or relative risk. Human genome is comprised of discrete regions, of linkage disequilibrium (LD), where nearby single nucleotide polymorphisms (SNPs) show strong correlation with one another.[12] The International HapMap Project was initiated in 2002 to study LD patterns across the genome in multiple populations. The HapMap Project genotyped more than 3 million SNPs in 269 samples from 4 populations (Northern and Western European ancestry, Chinese,

Japanese, and Yorubans from Nigeria).[13] The completion of these 2 projects and the advances in genome-sequencing technologies including next-generation sequencing and high-volume SNP genotyping platforms such as the Illumina BeadArray and the Affymetrix GeneChip array have revolutionized the practice of molecular epidemiology. Commons terms used in the field of molecular and genetic epidemiology and their definitions are listed in (Table 1).

GENETIC AND MOLECULAR MECHANISMS OF CARCINOGENESIS

A human gene is composed of coding (exons) and noncoding (introns) regions, and regulatory DNA sequences.[14] A DNA mutation occurring in germline cells may be passed on to progeny and thereby lead to so-called less common heritable cancer syndromes (Table 2). The degree of cancer risk associated with inheriting a mutation depends on gene penetrance, which is the probability that an individual carrying a mutation will develop disease and may be modified by a combination of genetic and environmental factors. Gene mutations associated with inherited cancer syndromes usually involve oncogenes or tumor suppressor genes. Several hundred proto-oncogenes have been identified that have the potential to be converted into oncogenes and lead to cancer. They usually are involved in processes such as cell division, apoptosis, signal transduction, and cell differentiation.

Additionally, cancer susceptibility also can be inherited through more common nucleotide sequence variations. These variations include SNPs, 1 base-pair substitutions, sequence insertions and deletions, and highly variable repeating nucleotide segments. SNPs are the most common form of variation and exist approximately every 300 bp on average across the genome.[13] An SNP is a DNA sequence variation occurring when a single nucleotide (A, T, C, or G) in the genome differs between members of a species. Between ethnically different populations, there are 7 million differences at the level of single-base pairs occurring at a frequency of at least 5% of the population.[15] Most SNPs have no functional consequence if they occur in a noncoding sequence, or might have modest effects that may interact with environmental factors to increase cancer susceptibility.

Epigenetic mechanisms of carcinogenesis are increasingly being identified as key steps in pathways from exposure to cancer. These mechanisms do not depend on structural changes in DNA but on functional regulation such as DNA methylation and histone modification. These processes are important determinants of gene expression as they determine the process by which the instructions in genes are converted to mRNA, directing protein synthesis.[16] Genomic instability also plays an important role in carcinogenesis
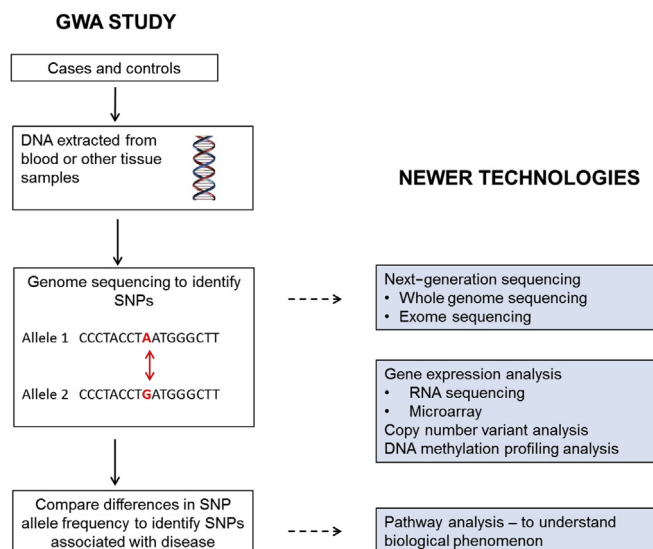

Figure 1. Methods used to study genetic variations in molecular and genetic epidemiology. Abbreviations: GWA, genome-wide association; SNP, single nucleotide polymorphism.

both in inherited and sporadic cancers and is broadly classified into microsatellite instability and chromosome instability. In inherited cancers, genomic instability is mostly caused by mutations in DNA-mismatch repair.[17]

CANCER SUSCEPTIBILITY FROM HERITABLE CANCER SYNDROMES

Once a cancer is ascertained to have a genetic component through familial aggregation studies, the next step is to identify the pattern of Mendelian inheritance (autosomal dominant or recessive) using segregation analysis. Then linkage analysis is used to localize the chromosomal region containing the cancer-causing gene. This approach has been used to identify multiple inherited cancer syndromes. For instance, linkage analysis was conducted as a genome-wide search among 23 families with multiple individuals affected with early-onset breast cancer to identify associated genetic loci.[18] They linked a region on chromosome 17q21 and positional cloning was used to localize the BRCA1 gene in this region.[19] Two years later BRCA2, located on chromosome 13q12−13, was identified.[20] Both BRCA1 and BRCA2 are tumor suppressor genes. Mutations in BRCA1 and BRCA2 also are associated with other malignancies, such as ovarian, prostate, and pancreatic cancers; malignant melanoma; gallbladder and bile duct cancers; and stomach cancer.[21,22]

In general, women with BRCA1 have a lifetime risk for breast cancer of 60% to 85% and prevalence of BRCA1 or BRCA2 mutations in the general population range from 1 in 200 to 1 in 1000.[23] Factors that modify penetrance appear to be reproductive factors and exogenous hormones.[24] Cancer risk from BRCA1 and BRCA2 is reported to be highest in Ashkenazi Jews.[25] It is estimated that 5% to 10% of all breast cancers can be

Table 1. Common Terms in Genetic and Molecular Epidemiology

Term	Definition
Allele	One of ≥2 versions of a gene. An individual inherits 1 allele from each parent. If the 2 alleles are the same, the individual is homozygous for that gene. If the alleles are different, the individual is heterozygous
Copy number variant	Number of copies of a particular gene that vary between individuals
DNA adducts	Part of DNA that is covalently bound to a chemical or carcinogen and can lead to carcinogenesis
DNA methylation	Process involving addition of a methyl group to the cytosine or adenine DNA nucleotides that may then affect the expression of genes in cells
DNA sequencing	Technique used to determine the exact sequence of bases (A, C, G, T) in a DNA molecule
Epigenetics	Study of functional changes to the genome that do not involve a change in the nucleotide sequence. Examples of mechanisms that produce such changes are DNA methylation and histone modification
Exome	Part of the genome formed by exons or coding regions, the sequences that are transcribed into mature RNA
Gene expression analysis	Gene expression is the process by which information from a gene is used in the synthesis of a functional gene product like a protein or functional RNA. This involves multiple processes including transcription, RNA splicing, translation, and post-translational modification of a protein. Methods used for gene expression analysis include microarray technologies and RNA sequencing
Locus	Physical location of a gene or DNA sequence on a chromosome
Microarray technology	Technology used to study the expression of genes. A DNA or RNA sample is placed in contact with gene sequences on a glass slide. Complementary base pairs between the sample and the gene sequences produce light that is measured and these represent genes that are expressed in the sample.
Microsatellite instability	Microsatellites are repetitive DNA sequences variable from person to person. Microsatellite instability results from defect in DNA-mismatch repair
Minor allele frequency	Frequency at which the least common allele occurs in a given population
Next-generation sequencing	Recent sequencing technology also known as high-throughput sequencing that enables rapid sequencing of large stretches of DNA spanning even entire genomes in a small amount of time
Oncogene	Gene that if mutated can contribute to the development of a cancer. In their normal, unmutated state, oncogenes are called proto-oncogenes, and they play a role in the regulation of cell division. For example, *myc* gene is a proto-oncogene that, when activated, can cause hematologic malignancies
Pathway analysis	Method to determine the biological pathways including signaling and metabolic pathways using the list of genes that are determined to be implicated in carcinogenesis using GWAS or gene expression analysis
Tumor suppressor gene	Gene that directs the production of a protein that plays a role in keeping cell division in check. When mutated, it may contribute to the carcinogenesis. For example, *TP53* gene codes for p53 tumor suppressor protein

GWAS, genome-wide association studies.

attributed to highly penetrant germline mutations and mutations in *BRCA1* and *BRCA2* contribute to 90% of these.[26] Other heritable syndromes associated with breast cancer include Li-Fraumeni familial cancer syndrome (mutations in *P53*)[23] and Cowden's syndrome (mutation in *PTEN*).[26]

Linkage analysis also has been used to identify a heritable cancer syndrome in melanoma. As 5% to 10% of melanomas occur among individuals with a strong family history of the disease, linkage analyses of melanoma kindreds in the 1990s was performed. The analyses identified a familial melanoma susceptibility locus on the short arm of chromosome 9.[27,28] This led to the identification, of the *CDKN2A* gene, which is the most common cause of inherited susceptibility to melanoma. *CDKN2A* mutation penetrance has been estimated to be 30% by age 50 years and 67% by age 80 and varies by geographical location.[29]

The 2 heritable syndromes associated with increased risk for colorectal cancer are familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC or Lynch syndrome). FAP is caused by an inherited mutation in the *APC* gene, a tumor suppressor gene that regulates formation of a protein names β-catenin involved in activation of transcription of various oncogenes.[30,31] This is inherited as an autosomal dominant disorder and is manifested as numerous adenomatous polyps in the epithelium of the large intestine, which almost always progress to cancer. HNPCC is also an autosomal dominant disorder that is associated with an increased risk for colorectal cancer and various other cancers, including endometrial and ovarian cancers.[32,33] HNPCC accounts for up to 3% of all colorectal carcinomas[34] and 5% to 10% of all ovarian cancers.[32] This syndrome is associated with high-penetrance germline mutations in DNA-mismatch repair genes: *MSH2, MLH1, PMS1,* and *PMS*[35,36] and leads to microsatellite instability. This defect in DNA repair leads to

Table 2. Common Heritable Cancer Syndromes

	Gene	Syndrome	Pattern of Inheritance	Cancer Type
Oncogenes	*RET*	Multiple endocrine neoplasia 2, familial medullary thyroid cancer	Dominant	Pheomchromocytoma, medullary thyroid
	MET	Familial papillary renal carcinoma syndrome	Dominant	Renal
Tumor suppressor genes	*APC*	Familial adenomatous polyposis	Dominant	Colon
	BRCA1	Hereditary breast-ovarian syndrome	Dominant	Breast, ovarian
	BRCA2	Hereditary breast-ovarian syndrome	Dominant	Breast, ovarian, prostate, pancreas, melanoma
	CDKN2A	Familial melanoma	Dominant	Melanoma
	MEN1	Multiple endocrine neoplasia 1	Dominant	Pituitary, parathyroid, pancreas
	NF2	Neurofibromatosis 2	Dominant	Schwannoma
	p53	Li-Fraumeni syndrome	Dominant	Breast, brain, adrenal, sarcoma, leukemia
	PTEN	Cowden syndrome	Dominant	Breast, thyroid, GI, skin
	Rb	Hereditary retinoblastoma	Dominant	Retina
	VHL	von Hippel-Lindau syndrome	Dominant	Renal
DNA repair genes	*MSH2, MLH1, PMS1, PMS*	Hereditary nonpolyposis colon cancer	Dominant	Colon, endometrium, ovary, other GI, brain, skin
Chromosome instability syndromes	*ATM*	Ataxia-telangiectasia	Recessive	Lymphoma, leukemia
	BLM	Bloom syndrome	Recessive	Lymphoma, leukemia
	FANC(A-P)	Fanconi anemia	Recessive	Leukemia
	XP(A-G)	Xeroderma pigmentosum	Recessive	Skin

GI, gastrointestinal.

spontaneous genetic mutations to accumulate in colonic mucosa, which predisposes to the dysplasia, and eventually development of invasive cancers.[37] Degree of microsatellite instability can be measured by polymerase chain reaction and is used as a marker of DNA repair capacity.[38] Patients with heritable syndromes develop cancers at young ages. For instance, women with HNPCC tend to develop endometrial cancer 15 years earlier than the general population.[39]

Multiple other inherited cancer syndromes have been identified. Multiple endocrine neoplasia (MEN) types 2A and 2B account for approximately 20% of medullary thyroid cancers and are attributed to germline mutations in the *ret* proto-oncogene on chromosome 10.[40,41] Fanconi anemia, ataxia telangiectasia, and Bloom syndrome are inherited disorders associated with inherent chromosome instability and are autosomal recessive disorders. They are associated with increased risk for many cancers, including skin cancer and acute leukemia.[42,43]

CANCER SUSCEPTIBILITY FROM COMMON GENETIC POLYMORPHISMS OR SNPS

The polygenic model suggests that multiple genetic variations or polymorphisms may confer a small amount of cancer risk individually, yet in combination, may result in modest susceptibility to cancer.[44] GWAS have been used to identify multiple genetic polymorphisms and cancer susceptibility loci and are more powerful than linkage analysis in identifying genes with modest effects.[45] GWAS of high-risk prostate cancer families in the United States and Sweden showed the first prostate cancer susceptibility locus in 1996 on the long arm of chromosome 1 (HPC1) at 1q24-25.[46] Another polymorphism that has been identified in prostate cancer is in 8q24 that is associated with a population attributable risk of approximately 8% in whites and 16% in blacks. This may be contributing to higher incidence of prostate cancer observed in black men compared with men of European ancestry.[47] GWAS also have identified susceptibility loci for multiple other cancers including lung cancer with susceptibility loci identified at 15q25, 5p15, and 6p21. The loci at 15q25 correspond to nicotine acetylcholine receptor unit and may be contributing to lung carcinogenesis in response to smoking.[48,49] Polymorphisms also influence cancer risk by influencing the response to environmental factors like environmental toxins, diet, alcohol, or tobacco (Fig. 2).

POLYMORPHISMS RELATED TO RESPONSE TO METABOLIZING CARCINOGENS

Many exogenous carcinogens require activation by metabolizing enzymes to their activated forms. On the other
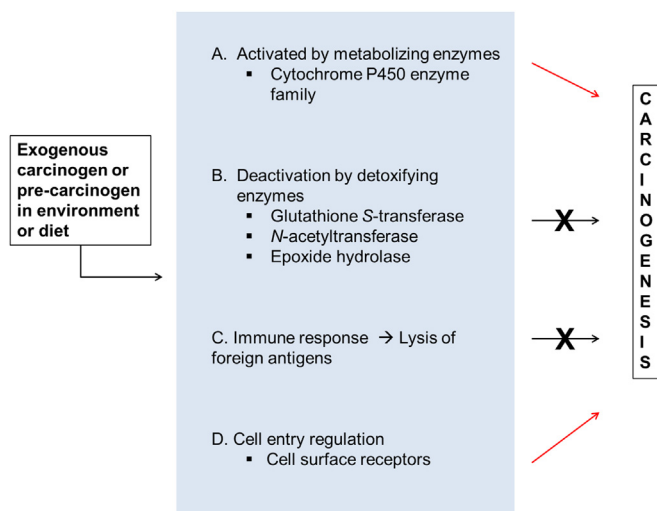
Figure 2. Role of genetic variations and interactions with environmental factors in carcinogenesis.

hand, detoxification enzymes deactivate carcinogens or their intermediate metabolites. Inherited polymorphisms in these enzymes may alter their rate of activation or detoxification, thus increasing or decreasing the carcinogenic potential of the environmental exposures. Metabolizing enzymes belong to the cytochrome P450 (CYP) enzyme family while detoxifying enzymes include glutathione S-transferase (GST), N-acetyltransferase (NAT2), and epoxide hydrolase.[50,51]

CYP Enzymes

Polymorphisms on *CYP1A1* gene are associated with increased risk for lung cancer among smokers in Japanese as well as white populations.[52,53] CYP polymorphisms also have been shown to be associated with increased gastric cancer in high-risk individuals in China in a case-cohort analysis. Heterozygous or homozygous carriage of the variant *CYP1A1*2A* allele was associated with an almost halved risk for gastric cancer, adjusted relative risk (RR) 0.47.[54] Similarly, a Taiwanese study found an almost 3-fold increased risk for stomach cancer among carriers of the c2/c2 genotype for the *CYP2E1* gene, an important activator of nitrosamines.[55]

NAT2

Cooking of animal proteins at high temperature creates carcinogens such as heterocyclic amines and polycyclic aromatic hydrocarbons (PAHs).[56] Heterocyclic amines are metabolized by a number of enzymes including NAT2 and polymorphisms in the gene for these enzymes divide the population into slow acetylators (about 59% of the white population) and rapid acetylators.[57] NATs also play a role in detoxifying aromatic amines found in tobacco smoke.[58] In a large case-control study and meta-analysis in whites, NAT2 slow acetylators had a 40% increased risk for developing bladder cancer; this association was more pronounced in cigarette smokers.[59]

GST

Chemical carcinogens like intermediates of PAHs found in tobacco smoke are detoxified by enzymes including GST and polymorphisms in genes coding for these have been associated with a substantially greater risk for lung cancer in nonsmokers from environmental tobacco smoke when compared with individuals who were heterozygous or homozygous carriers of the wild-type GSTM1 allele.[60] The M1 locus in the *GST* gene is entirely absent in approximately 50% of whites, and this deletion has been associated quite consistently with increased risk for lung cancer in this population and accounts for 17% of lung cancer cases.[61]

POLYMORPHISMS RELATED TO RESPONSE TO DIET AND ALCOHOL

Studies have identified potential interactions between dietary intake and genetic variants and some gene-diet associations have been established. For instance, in breast cancer pathogenesis, studies have demonstrated the interactive effect between folate intake and the 1-carbon metabolism-related genes. SNPs related to the folate-metabolizing enzyme methylenetetrahydrofolate reductase (*MTHFR*) gene modify associations between folate intake and breast cancer. In a nested case-control study, a positive association between dietary folate intake and breast cancer was observed in participants with *MTHFR 677TT* genotype.[62,63]

Alcohol or ethanol is primarily oxidized to acetaldehyde (its carcinogenic form) by enzyme alcohol dehydrogenase (ADH).[64] Although most ethanol metabolism occurs in the liver, ADH is also expressed in the oral cavity and upper aerodigestive tract.[65,66] Several genes including *ADH1B* and *ADH1C* are involved in alcohol metabolism and polymorphisms at these gene loci can influence the risk for cancer. For instance, several studies have reported that *ADH1B*1* allele, which is the fast metabolizing allele found in majority of Asians, is associated with an increased risk for head and neck cancer.[67]

POLYMORPHISMS RELATED TO RESPONSE TO INFECTIONS

The human leukocyte antigen (*HLA*) genes encode proteins required for the presentation of foreign antigens, including viral peptides, to the immune system for targeted lysis. Because virtually all nasopharyngeal cancers (NPC) contain Epstein-Barr virus (EBV), individuals who inherit *HLA* alleles with a reduced ability to present EBV antigens may have an increased risk for developing NPC, whereas individuals with *HLA* alleles that present EBV efficiently may have a lower risk.[68] In southern Chinese and other Asian populations, *HLA-A2-B46* and *B17* are generally associated with a 2- to 3-fold increase in NPC risk.[69,70] In contrast, lower risk for NPC was found

in association with *HLA-A11*.[70] In Thailand[71] and in China,[72] polymorphisms in the polymeric immunoglobulin receptor (*PIGR*), a cell surface receptor proposed to mediate EBV entry into the nasal epithelium are associated with 2- to 3-fold increased risk for NPC. These polymorphisms may partly account for higher prevalence of NPC seen in South-East Asia.

In Egypt and the Middle East where *Schistosoma hematobium* is endemic, squamous cell carcinomas of the bladder predominate compared with the more common transitional-cell carcinoma seen in America and Europe and these differ at the molecular level.[73,74] A study reported high frequency of 9p loss of heterozygosity (LOH) at *CDKN2* gene in *Schistosoma*-associated bladder cancers at 65% compared with 39% in non—*Schistosoma*-associated bladder cancers. This suggests that *CDKN2* gene on 9p may contribute to the development of the majority of schistosomiasis-associated bladder tumors.[74]

Other Polymorphisms Related to Cancer Risk

Epidermal growth factor (EGF) and transforming growth factor (TGF)-β, through interaction with cell surface receptors, induce growth signals and are important for tumor growth and progression. A case-control study in Japan showed that individuals with the *EGFR A/A* or *A/G* genotype showed a significantly lower risk for gastric cancer than those with the *G/G* genotype (adjusted odds ratio [OR], 0.56). This polymorphism is associated with reduced production of EGF.[75] Another study in China also showed that EGF promoter polymorphisms were associated with a significantly decreased risk for gastric cancer.[76] Similarly, in a case-control study of 675 gastric cancer cases and 704 healthy controls in a Chinese population, variant alleles of the promoter polymorphisms, *TGFB1 C-509T* and *TGFBR2 G-875A*, were associated with a significantly decreased risk for gastric cancer (OR, 0.65 for -509CT/TT and 0.67 for -875GA/AA).[77] Although many factors like diet and *Helicobacter pylori* infection have been implicated in increased prevalence of gastric cancer in China and Japan, these genetic variants also may play a role.

Polymorphisms in *XRCC1*, which is involved in DNA repair activity, are associated with 20% tp 50% lower risk for NPC in Taiwan[78] and southern China.[79] In a large population-based case-control study in breast cancer, a positive association for *XRCC1* codon 399 Arg/Gln or Gln/Gln genotypes compared with Arg/Arg was found among blacks (OR 1.7; 95% CI, 1.1-2.4) but not whites.[80]

CLINICAL IMPLICATIONS

The role of genetic and molecular factors in cancer susceptibility can be used clinically in cancer prevention strategies by identifying high-risk individuals in screening programs based on genetic susceptibility loci. However,

this is an evolving area as risk prediction models based on SNPs to predict cancer risk have shown only modest benefit. Adding 7 SNPs identified from GWA analyses to the original Gail model Breast Cancer Risk Assessment Tool yielded only a modest improvement in area under the curve (AUC) statistic from 0.607 to 0.632. As per an estimate, at least 280 more SNPs are needed to improve the performance of the test.[81] Although the SNP- based risk score had only a moderate discriminatory accuracy, it still improves the predictive ability of the prediction model. Future studies incorporating both genetic polymorphisms and established cancer risk factors in risk prediction models are needed.[82]

Molecular and genetic epidemiology have been used in molecular biomarkers to determine exposure to a potential carcinogen to identify at-risk populations or to measure the efficacy of cancer prevention or control programs. Tobacco smoke contains inhaled carcinogens including the PAHs and *N*-nitrosamines, which are metabolized by the P450 enzyme system and exert carcinogenic effects through the formation of DNA adducts.[83] Lung cancer patients have significantly higher levels of aromatic/PAH-DNA adducts[84,85] even after adjustment for potential confounders. This can be used to identify individuals who would benefit most from early intervention. A study of 400 smokers showed how DNA adducts can be used to monitor efficacy of a smoking-cessation program. Blood samples were drawn from the participants before they began the program and then at multiple time points after smoking cessation. Levels of PAH-DNA and other adducts reflected cessation; within 8 weeks of quitting, their concentrations were significantly reduced.[86]

Other carcinogens are also known to form DNA adducts that can be used as a biomarker. AFB1 is a fungal metabolite present in grains and cereals as a result of improper storage and induces DNA damage in humans.[87] A prospective study of 18,244 men in Shanghai (22 incident cases of liver cancer) used assays for aflatoxin exposure including DNA adducts to assess the relation between aflatoxin exposure and liver cancer. Individuals with liver cancer were more likely to have detectable concentrations of these aflatoxin-related DNA adducts (relative risk [RR], 2.4; 95% CI, 1.0-5.9).[88] A subsequent study in Taiwan showed similar results.[89] This can be used as a useful resource in planning liver cancer prevention and screening programs in high-prevalence areas like Asia.

Limitations

One major limitation in conducting GWAS is the need for a large sample size as they evaluate a large number of SNPs increasing the chance of type I error due to multiple comparisons being performed simultaneously. For instance, analysis of 5356 invasive breast cancer cases and more than 7000 controls in the National Cancer Institute Breast and Prostate Cancer Cohort[90] found an

association between genetic variations at the *CYP19A1* locus with a 10% to 20% increase in endogenous estrogen levels, but not with breast cancer risk. This was thought to be due to the lack of statistical power. Relatively few main effects of low-penetrance alleles have been found to be reproducible,[91] and even fewer gene-environment interactions are confirmed. One method to address this is by pooling of results of several GWAS to attain statistical power. But this is often hindered by a lack of consistency in quantifying cancer risk. Another source of bias in GWAS is population stratification. This occurs when cases and controls are selected from populations with different ethnic backgrounds leading to false-positive results.[92]

CONCLUSIONS

Genetic and molecular factors can play an important role in an individual's cancer susceptibility and response to carcinogen exposure. Therefore, this is an important area to identify potential areas for cancer prevention and screening strategies. More research is needed in this area, especially in investigating new biomarkers and measuring gene-environment interactions. Large collaborative studies and consortia to coordinate research efforts and enhance statistical power by increasing sample size are needed.

References

1. Ponder BA. Cancer genetics. Nature 2001;411:336—41.
2. Zeng YX, Jia WH. Familial nasopharyngeal carcinoma. Semin Cancer Biol 2002;12:443—50.
3. Rebora P, Czene K, Reilly M. Timing of familial breast cancer in sisters. J Natl Cancer Inst 2008;100:721—7.
4. Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA. Family history and the risk of breast cancer: a systematic review and meta-analysis. Int J Cancer 1997;71:800—9.
5. Klein AP, Brune KA, Petersen GM, et al. Prospective risk of pancreatic cancer in familial pancreatic cancer kindreds. Cancer Res 2004;64:2634—8.
6. Shields PG. Inherited factors and environmental exposures in cancer risk. J Occup Med 1993;35:34—41.
7. Barrett JC, Shelby MD. Mechanisms of human carcinogens. Prog Clin Biol Res 1992;374:415—34.
8. Perera FP. Molecular epidemiology: insights into cancer susceptibility, risk assessment, and prevention. J Natl Cancer Inst 1996;88:496—509.
9. Shields PG, Harris CC. Molecular epidemiology and the genetics of environmental cancer. JAMA 1991;266:681—7.
10. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860—921.
11. Manolio TA, Collins FS. The HapMap and genome-wide association studies in diagnosis and therapy. Annu Rev Med 2009;60:443—56.
12. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in the human genome. Science 2002;296:2225—9.
13. International HapMap C. A haplotype map of the human genome. Nature 2005;437:1299—320.
14. Pearson H. Genetics: what is a gene? Nature 2006;441:398—401.
15. Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. Nat Rev Genet 2009;10:241—51.
16. Kim YI. Folate and DNA methylation: a mechanistic link between folate deficiency and colorectal cancer? Cancer Epidemiol Biomarkers Prev 2004;13:511—9.
17. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability—an evolving hallmark of cancer. Nat Rev Mol Cell Biol 2010;11:220—8.
18. Hall JM, Lee MK, Newman B, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. Science 1990;250:1684—9.
19. Miki Y, Swensen J, Shattuck-Eidens D, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 1994;266:66—71.
20. Schutte M, Rozenblum E, Moskaluk CA, et al. An integrated high-resolution physical map of the DPC/BRCA2 region at chromosome 13q12. Cancer Res 1995;55:4570—4.
21. Breast Cancer Linkage C. Cancer risks in BRCA2 mutation carriers. J Natl Cancer Inst 1999;91:1310—6.
22. Thompson D, Easton DF. Breast Cancer Linkage C. Cancer Incidence in BRCA1 mutation carriers. J Natl Cancer Inst 2002;94:1358—65.
23. Ellisen LW, Haber DA. Hereditary breast cancer. Annu Rev Med 1998;49:425—36.
24. Narod SA. Modifiers of risk of hereditary breast cancer. Oncogene 2006;25:5832—6.
25. Struewing JP, Hartge P, Wacholder S, et al. The risk of cancer associated with specific mutations of BRCA1 and BRCA2 among Ashkenazi Jews. N Engl J Med 1997;336:1401—8.
26. Bennett IC, Gattas M, Teh BT. The genetic basis of breast cancer and its clinical implications. Aust N Z J Surg 1999;69:95—105.
27. Cannon-Albright LA, Goldgar DE, Meyer LJ, et al. Assignment of a locus for familial melanoma, MLM, to chromosome 9p13-p22. Science 1992;258:1148—52.
28. Nancarrow DJ, Mann GJ, Holland EA, et al. Confirmation of chromosome 9p linkage in familial melanoma. Am J Hum Genet 1993;53:936—42.
29. Bishop DT, Demenais F, Goldstein AM, et al. Geographical variation in the penetrance of CDKN2A mutations for melanoma. J Natl Cancer Inst 2002;94:894—903.
30. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. Cell 1996;87:159—70.
31. Meera Khan P, Tops CM, van de Broek M, et al. Close linkage of a highly polymorphic marker (D5S37) to familial adenomatous polyposis (FAP) and confirmation of FAP localization on chromosome 5q21-q22. Hum Genet 1988;79:183—5.
32. Lynch HT, Casey MJ, Lynch J, White TE, Godwin AK. Genetics and ovarian carcinoma. Semin Oncol 1998;25:265—80.
33. Aaltonen LA, Salovaara R, Kristo P, et al. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. N Engl J Med 1998;338:1481—7.
34. Samowitz WS, Curtin K, Lin HH, et al. The colon cancer burden of genetically defined hereditary nonpolyposis colon cancer. Gastroenterology 2001;121:830—8.
35. Papadopoulos N, Nicolaides NC, Wei YF, et al. Mutation of a mutL homolog in hereditary colon cancer. Science 1994;263:1625—9.
36. Nicolaides NC, Papadopoulos N, Liu B, et al. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 1994;371:75—80.
37. De Jong AE, Morreau H, Van Puijenbroek M, et al. The role of mismatch repair gene defects in the development of adenomas in patients with HNPCC. Gastroenterology 2004;126:42—8.
38. Kinoshita M, Nakamura J, Kusaka H, et al. Automated and simultaneous identification of microsatellite instability by fluorescence-based polymerase chain reaction (PCR) in four loci. Clin Chim Acta 1999;279:15—23.
39. Vasen HF, Watson P, Mecklin JP, et al. The epidemiology of endometrial cancer in hereditary nonpolyposis colorectal cancer. Anticancer Res 1994;14:1675—8.
40. Niccoli-Sire P, Murat A, Baudin E, et al. Early or prophylactic thyroidectomy in MEN 2/FMTC gene carriers: results in 71 thyroidectomized patients. The French Calcitonin Tumours Study Group (GETC). Eur J Endocrinol 1999;141:468—74.
41. Eng C, Clayton D, Schuffenecker I, et al. The relationship between specific RET proto-oncogene mutations and disease phenotype in multiple endocrine neoplasia type 2. International RET mutation consortium analysis. JAMA 1996;276:1575—9.
42. Popp HD, Bohlander SK. Genetic instability in inherited and sporadic leukemias. Genes Chromosomes Cancer 2010;49:1071—81.
43. Taylor AM. Chromosome instability syndromes. Best Pract Res Clin Haematol 2001;14:631—44.

44. Lange K. An approximate model of polygenic inheritance. Genetics 1997;147:1423—30.

45. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273:1516—7.

46. Smith JR, Freije D, Carpten JD, et al. Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. Science 1996;274:1371—4.

47. Amundadottir LT, Sulem P, Gudmundsson J, et al. A common variant associated with prostate cancer in European and African populations. Nat Genet 2006;38:652—8.

48. Landi MT, Chatterjee N, Yu K, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet 2009;85:679—91.

49. Truong T, Hung RJ, Amos CI, et al. Replication of lung cancer susceptibility loci at chromosomes 15q25, 5p15, and 6p21: a pooled analysis from the International Lung Cancer Consortium. J Natl Cancer Inst 2010;102:959—71.

50. Lang M, Pelkonen O. Metabolism of xenobiotics and chemical carcinogenesis. IARC Sci Publ 1999:13—22.

51. Taningher M, Malacarne D, Izzotti A, Ugolini D, Parodi S. Drug metabolism polymorphisms as modulators of cancer susceptibility. Mutat Res 1999;436:227—61.

52. Kawajiri K, Fujii-Kuriyama Y. P450 and human cancer. Jpn J Cancer Res 1991;82:1325—35.

53. Xu X, Kelsey KT, Wiencke JK, Wain JC, Christiani DC. Cytochrome P450 CYP1A1 MspI polymorphism and lung cancer susceptibility. Cancer Epidemiol Biomarkers Prev 1996;5:687—92.

54. Roth MJ, Abnet CC, Johnson LL, et al. Polymorphic variation of Cyp1A1 is associated with the risk of gastric cardia cancer: a prospective case-cohort study of cytochrome P-450 1A1 and GST enzymes. Cancer Causes Control 2004;15:1077—83.

55. Wu MS, Chen CJ, Lin MT, et al. Genetic polymorphisms of cytochrome p450 2E1, glutathione S-transferase M1 and T1, and susceptibility to gastric carcinoma in Taiwan. Int J Colorectal Dis 2002;17:338—43.

56. Knize MG, Salmon CP, Pais P, Felton JS. Food heating and the formation of heterocyclic aromatic amine and polycyclic aromatic hydrocarbon mutagens/carcinogens. Adv Exp Med Biol 1999;459:179—93.

57. Cascorbi I, Drakoulis N, Brockmoller J, Maurer A, Sperling K, Roots I. Arylamine N-acetyltransferase (NAT2) mutations and their allelic linkage in unrelated Caucasian individuals: correlation with phenotypic activity. Am J Hum Genet 1995;57:581—92.

58. Ambrosone CB, Freudenheim JL, Graham S, et al. Cigarette smoking, N-acetyltransferase 2 genetic polymorphisms, and breast cancer risk. JAMA 1996;276:1494—501.

59. Garcia-Closas M, Malats N, Silverman D, et al. NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. Lancet 2005;366:649—59.

60. Bennett WP, Alavanja MC, Blomeke B, et al. Environmental tobacco smoke, genetic susceptibility, and risk of lung cancer in never-smoking women. J Natl Cancer Inst 1999;91:2009—14.

61. McWilliams JE, Sanderson BJ, Harris EL, Richert-Boe KE, Henner WD. Glutathione S-transferase M1 (GSTM1) deficiency and lung cancer risk. Cancer Epidemiol Biomarkers Prev 1995;4:589—94.

62. Ericson U, Sonestedt E, Ivarsson MI, et al. Folate intake, methylenetetrahydrofolate reductase polymorphisms, and breast cancer risk in women from the Malmo Diet and Cancer cohort. Cancer Epidemiol Biomarkers Prev 2009;18:1101—10.

63. Suzuki T, Matsuo K, Hirose K, et al. One-carbon metabolism-related gene polymorphisms and risk of breast cancer. Carcinogenesis 2008;29:356—62.

64. Bosron WF, Li TK. Genetic polymorphism of human liver alcohol and aldehyde dehydrogenases, and their relationship to alcohol metabolism and alcoholism. Hepatology 1986;6:502—10.

65. Muto M, Hitomi Y, Ohtsu A, et al. Acetaldehyde production by non-pathogenic Neisseria in human oral microflora: implications for carcinogenesis in upper aerodigestive tract. Int J Cancer 2000;88:342—50.

66. Wight AJ, Ogden GR. Possible mechanisms by which alcohol may influence the development of oral cancer—a review. Oral Oncol 1998;34:441—7.

67. Brennan P, Lewis S, Hashibe M, et al. Pooled analysis of alcohol dehydrogenase genotypes and head and neck cancer: a HuGE review. Am J Epidemiol 2004;159:1—16.

68. Simons MJ, Wee GB, Chan SH, Shanmugaratnam K, Day NE, de-The G. Immunogenetic aspects of nasopharyngeal carcinoma (NPC)

III. HL-a type as a genetic marker of NPC predisposition to test the hypothesis that Epstein-Barr virus is an etiological factor in NPC. IARC Sci Publ 1975:249—58.

69. Simons MJ, Day NE, Wee GB, et al. Nasopharyngeal carcinoma V: immunogenetic studies of Southeast Asian ethnic groups with high and low risk for the tumor. Cancer Res 1974;34:1192—5.

70. Simons MJ, Chan SH, Wee GB, et al. Nasopharyngeal carcinoma and histocompatibility antigens. IARC Sci Publ 1978:271—82.

71. Hirunsatit R, Kongruttanachok N, Shotelersuk K, et al. Polymeric immunoglobulin receptor polymorphisms and risk of nasopharyngeal cancer. BMC Genet 2003;4:3.

72. Fan Q, Jia WH, Zhang RH, et al. [Correlation of polymeric immunoglobulin receptor gene polymorphisms to susceptibility of nasopharyngeal carcinoma]. Ai Zheng 2005;24:915—8.

73. Ramchurren N, Cooper K, Summerhayes IC. Molecular events underlying schistosomiasis-related bladder cancer. Int J Cancer 1995;62:237—44.

74. Shaw ME, Elder PA, Abbas A, Knowles MA. Partial allelotype of schistosomiasis-associated bladder cancer. Int J Cancer 1999;80:656—61.

75. Hamai Y, Matsumura S, Matsusaki K, et al. A single nucleotide polymorphism in the 5' untranslated region of the EGF gene is associated with occurrence and malignant progression of gastric cancer. Pathobiology 2005;72:133—8.

76. Jin G, Miao R, Deng Y, et al. Variant genotypes and haplotypes of the epidermal growth factor gene promoter are associated with a decreased risk of gastric cancer in a high-risk Chinese population. Cancer Sci 2007;98:864—8.

77. Jin G, Wang L, Chen W, et al. Variant alleles of TGFB1 and TGFBR2 are associated with a decreased risk of gastric cancer in a Chinese population. Int J Cancer 2007;120:1330—5.

78. Cho EY, Hildesheim A, Chen CJ, et al. Nasopharyngeal carcinoma and genetic polymorphisms of DNA repair enzymes XRCC1 and hOGG1. Cancer Epidemiol Biomarkers Prev 2003;12:1100—4.

79. Cao Y, Miao XP, Huang MY, et al. Polymorphisms of XRCC1 genes and risk of nasopharyngeal carcinoma in the Cantonese population. BMC Cancer 2006;6:167.

80. Duell EJ, Millikan RC, Pittman GS, et al. Polymorphisms in the DNA repair gene XRCC1 and breast cancer. Cancer Epidemiol Biomarkers Prev 2001;10:217—22.

81. Gail MH. Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk. J Natl Cancer Inst 2008;100:1037—41.

82. Li H, Yang L, Zhao X, et al. Prediction of lung cancer risk in a Chinese population using a multifactorial genetic model. BMC Med Genet 2012;13:118.

83. Alberg AJ, Samet JM. Epidemiology of lung cancer. Chest 2003;123(Suppl):21S—49S.

84. Veglia F, Matullo G, Vineis P. Bulky DNA adducts and risk of cancer: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2003;12:157—60.

85. Tang D, Santella RM, Blackwood AM, et al. A molecular epidemiological case-control study of lung cancer. Cancer Epidemiol Biomarkers Prev 1995;4:341—6.

86. Mooney LA, Santella RM, Covey L, et al. Decline of DNA damage and other biomarkers in peripheral blood following smoking cessation. Cancer Epidemiol Biomarkers Prev 1995;4:627—34.

87. Wild CP, Hasegawa R, Barraud L, et al. Aflatoxin-albumin adducts: a basis for comparative carcinogenesis between animals and humans. Cancer Epidemiol Biomarkers Prev 1996;5:179—89.

88. Ross RK, Yuan JM, Yu MC, et al. Urinary aflatoxin biomarkers and risk of hepatocellular carcinoma. Lancet 1992;339:943—6.

89. Yu MW, Lien JP, Chiu YH, Santella RM, Liaw YF, Chen CJ. Effect of aflatoxin metabolism and DNA adduct formation on hepatocellular carcinoma among chronic hepatitis B carriers in Taiwan. J Hepatol 1997;27:320—30.

90. Haiman CA, Dossus L, Setiawan VW, et al. Genetic variation at the CYP19A1 locus predicts circulating estrogen levels but not breast cancer risk in postmenopausal women. Cancer Res 2007;67:1893—7.

91. Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. Nat Genet 2001;29:306—9.

92. Knowler WC, Williams RC, Pettitt DJ, Steinberg AG. Gm3;5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. Am J Hum Genet 1988;43:520—6.