STATE-OF-THE-ART REVIEW

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

 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,

2214-9996/

From the Division of Hematology and Oncology, and Institute of Translational Epidemiology, Icahn School of Medicine at Mount Sinai, New York, NY. Address correspondence to J.Y.; e-mail: jyoti.malhotra@mountsinai.org

The author has no conflicts of interest to declare.

http://dx.doi.org/10.1016/j.aogh.2014.09.011

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

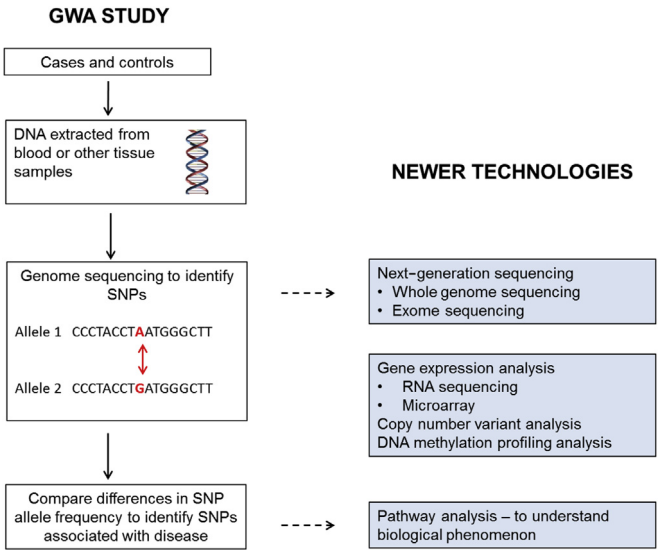

Figure 1. Methods used to study genetic variations in molecular and genetic epidemiology. Abbreviations: GWA, genome-wide association; SNP, single nucleotide polymorphism.

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