The contribution of molecular epidemiology to the identification of human carcinogens: current status and future perspectives

P. Boffetta1,2* & F. Islami1,3

1The Tisch Cancer Institute and Institute for Translational Epidemiology, Mount Sinai School of Medicine, New York, USA; 2International Prevention Research Institute, Lyon, France; 3Digestive Disease Research Center, Tehran University of Medical Sciences, Tehran, Iran

Received 6 August 2012; revised 14 September 2012; accepted 18 September 2012

Background: The use of biological-based markers of exposure, intermediate effect, outcome, and susceptibility has become standard practice in cancer epidemiology, which has contributed to identification of several carcinogenic agents. Nevertheless, with the exception of biological agents, this contribution, in terms of providing sufficiently strong evidence as required by the International Agency for Research on Cancer (IARC) monographs, has been modest.

Materials and methods: We discuss the overall contribution of molecular epidemiology to identification of carcinogens, with focus on IARC monographs.

Results: For many carcinogens, valid biological markers of exposure and mechanisms of actions are not available. Molecular markers are usually assessed in single biological samples, which may not represent the actual exposure or biological events related to carcinogens. The contribution of molecular epidemiology to identification of carcinogens has mainly been limited to the carcinogens acting through a genotoxic mechanism, i.e. when carcinogens induce DNA damage. A number of factors, including certain hormones and overweight/obesity, may show carcinogenic effects through nongenotoxic pathways, for which mechanisms of carcinogenicity are not well identified and their biomarkers are sparse.

Conclusion: Longitudinal assessment of biomarkers may provide more informative data in molecular epidemiology studies. For many carcinogens and mechanistic pathways, in particular nongenotoxic carcinogenicity, valid biological markers still need to be identified.

Key words: cancer, carcinogenicity, epidemiology, genotoxicity, molecular epidemiology

introduction

The field of molecular cancer epidemiology has been established some 40 years ago, with the early applications of molecular markers to population-based cancer studies [1]. Since then, the use of increasingly sophisticated biological-based markers of exposure, intermediate effect, outcome, and susceptibility has become standard practice in cancer epidemiology. The maturity of the field is reflected by the increasing number of textbooks, journals, academic programs, and professional societies [2–5].

One of the main domains of cancer epidemiology has been etiology research, which is the study of causes of human cancer, with the ultimate goal of identifying strategies for cancer prevention. The causes of cancer can be classified as inherited (genetic and epigenetic variants) and acquired (sometimes defined as ‘environmental’ in a broad sense) [6]. Despite strong evidence from family-based studies of a strong inherited component in human cancer etiology, the elucidation of the genes and variants responsible for an increased (or decreased) risk of cancer has been elusive: some 100 genetic factors have been identified through family-based studies [7], and a comparable number of low risk variants have been identified in recent years through genome-wide association studies [8]. Collectively, these factors explain a small proportion of cancers in the population although some may confer a high risk to carriers, and some special populations comprise a relatively large proportion of carriers of risk variants. However, our limited current knowledge of the ways that genetic and epigenetic factors influence quantitative traits and disease risk may be a reason for why hereditary factors have not been recognized as strong risk factors for many cancers at the general population level.
Compared with hereditary factors, the associations between environmental factors and cancers are better known. Modern cancer epidemiology was started in the early 1950s with the conduct of carefully designed cohort and case–control studies which rapidly lead to the identification of important, nongenetic causes of cancer, primarily tobacco smoking, alcohol drinking, and several occupational agents [9–13]. Since the 1980s, however, it became apparent that the yield of cancer epidemiology research, and the number of newly discovered carcinogenic agents, was decreasing (Figure 1). The rising of the new field of molecular cancer epidemiology was seen as an opportunity to elucidate the remaining etiologic questions.

Reviews of the causes of human cancer confront the following challenges: (i) evaluating the overall strength of evidence for various etiologic hypotheses and (ii) identifying the role of different components of the evidence (for example, by study design). These problems are best exemplified by the inconsistencies between the two reviews of the evidence of a particular mechanism is evaluated and classified as group 1. Therefore, strong mechanistic evidence may have an important role in evaluation of carcinogenicity of agents when the evidence from human and animal carcinogens identiﬁed in similar categories, but with different criteria adapted to the nature of animal studies. Mechanistic and other relevant evidence may include data on preneoplastic lesions, tumor pathology, genetic and related effects, structure–activity relationships, metabolism and toxicokinetics, physicochemical parameters, and analogous biological agents. The strength of the evidence that any carcinogenic effect observed is due to a particular mechanism is evaluated and classified as weak, moderate, or strong. Finally, the body of evidence is considered as a whole using criteria developed by IARC, and agents are classified as group 1, ‘carcinogenic to humans’; group 2A, ‘probably carcinogenic to humans’; group 2B, ‘possibly carcinogenic to humans’; group 3, ‘not classiﬁable as to its carcinogenicity to humans’; and group 4, ‘probably not carcinogenic to humans’ [18].

Group 1 category is used when there is sufﬁcient evidence of carcinogenicity in humans. However, when evidence of carcinogenicity in humans is less than ‘sufﬁcient’ but there is sufﬁcient evidence of carcinogenicity in experimental animals and strong evidence in exposed humans that the agent acts through a relevant mechanism of carcinogenicity, the agent may be classiﬁed as group 1. Therefore, strong mechanistic data may have an important role in evaluation of carcinogenicity of agents when the evidence from human cancer studies is less than sufﬁcient.

**molecular epidemiology and identification of human carcinogens**

**the IARC monographs program**

The evaluations in each of the IARC monographs on the evaluation of carcinogenic risks to humans are collectively made in case-by-case basis by a group of experts, who critically review the pertinent peer-reviewed scientiﬁc literature. The working group members have signiﬁcant knowledge and experience in research related to the carcinogenicity of the agents being reviewed and do not have real or apparent conﬂicts of interests. The ﬁrst monograph was published in 1972 [17]. By 2008, 99 monographs were conducted, in each of them, the data on carcinogenicity of one or several agents were evaluated. In 2009, IARC conducted ﬁve sessions of the monographs, collectively called as volume 100, to re-evaluate the human carcinogens identiﬁed in the ﬁrst 99 volumes. In 2011 and 2012 (by August), ﬁve additional groups met after the completion of the volume 100 project. In this article, we only consider the ﬁrst 100 monographs.

The evidence for carcinogenicity in epidemiological studies is classiﬁed by the monographs as follows: (i) ‘sufﬁcient evidence of carcinogenicity’, which indicates that a causal association is credible, and chance, bias, and confounding could be ruled out with reasonable conﬁdence; (ii) ‘limited evidence of carcinogenicity’, which shows a causal interpretation is credible, but chance, bias, or confounding could not be ruled out with reasonable conﬁdence; (iii) ‘inadequate evidence of carcinogenicity’, which indicates no data on cancer in humans are available or the available data do not permit a conclusion about the presence or absence of a causal association; and (iv) ‘evidence suggesting lack of carcinogenicity’. The evidence for carcinogenicity in animals is classiﬁed in similar categories, but with different criteria adapted to the nature of animal studies.

**Figure 1.** Decade of publication of key results on occupational carcinogens.
mechanistic studies in animals. Direct evidence from molecular epidemiology and other relevant data. Total 107

*11 biological agents.

For an additional 12 agents, the evaluation of sufficient evidence in humans was derived by the results of epidemiologic studies, which used molecular markers to measure the relevant exposures (and in some cases, potential confounders). Eleven of these carcinogens are biological agents (Table 2); in the most informative studies, assessment of infection was based on laboratory methods, which justifies their classification as molecular epidemiologic studies. In some instances, a relatively simple laboratory test had sufficient sensitivity and specificity to allow an accurate estimate of the association between the agent and cancer risk; a notable example here is the use of HBV surface antigen as an indicator of HBV infection in earlier biomarker studies of hepatocellular carcinoma [23]. In other instances, however, early laboratory methods were not sufficiently sensitive or specific to provide conclusive evidence of a causal association. This was the case of human papilloma virus infection, whose etiologic role in cervical carcinogenesis became clear when polymerase chain reaction (PCR)-based amplification was applied to large-scale epidemiologic studies (Table 3) [24–26].

The only nonbiologic agent for which molecular epidemiology played a critical role in exposure assessment and its evaluation as an IARC group 1 carcinogen is aflatoxin. In this case, the development of urinary markers of exposure, in particular the aflatoxin B1-DNA adduct at N7 of guanine, and its application in several prospective studies, in which samples were collected at enrollment and stored during follow-up, reduced exposure misclassification, and revealed a strong and consistent association between aflatoxin and hepatocellular carcinoma risk [27].

For the remaining 13 group 1 agents, the evidence of carcinogenicity in humans was less than sufficient (Table 4). Evidence for carcinogenicity in animals for one agent (etoposide) was ‘inadequate’ and for the other 12 agents was ‘sufficient’. The classification of these 13 agents as established human carcinogens was supported by strong mechanistic evidence. This can be considered as a contribution to molecular epidemiology: strong evidence of a relevant mechanism of carcinogenicity has been shown in humans exposed to 11 of these agents. For two dioxin-like compounds, 3,4,5,3'-pentachlorobiphenyl and 2,3,4,7,8-pentachlorodibenzofuran, evidence for similarity of their carcinogenic action with that of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), an established carcinogen, was provided by mechanistic studies in animals.

Table 1. Agents classified in group 1 (IARC monographs volumes 1–100), by type of evidence

<table>
<thead>
<tr>
<th>Type of evidence</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct evidence from traditional epidemiology</td>
<td>82</td>
</tr>
<tr>
<td>Direct evidence from molecular epidemiology</td>
<td>12</td>
</tr>
<tr>
<td>Mechanistic evidence from molecular epidemiology and other relevant data</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>107</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Biological agent</th>
<th>Established mechanistic events</th>
<th>A key study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epstein–Barr virus</td>
<td>Cell proliferation, inhibition of apoptosis, genomic instability, and cell migration</td>
<td>[80]</td>
</tr>
<tr>
<td>Hepatitis B virus</td>
<td>Inflammation, liver cirrhosis, and chronic hepatitis</td>
<td>[81]</td>
</tr>
<tr>
<td>Hepatitis C virus</td>
<td>Inflammation, liver cirrhosis, and liver fibrosis</td>
<td>[82]</td>
</tr>
<tr>
<td>Kaposi’s sarcoma herpes virus</td>
<td>Cell proliferation, inhibition of apoptosis, genomic instability, and cell migration</td>
<td>[83]</td>
</tr>
<tr>
<td>Human immunodeficiency virus, type 1</td>
<td>Immunosuppression (indirect action)</td>
<td>[84]</td>
</tr>
<tr>
<td>Human papillomavirus</td>
<td>Immortalization, genomic instability, inhibition of DNA damage response, and antiapoptotic activity</td>
<td>[24]</td>
</tr>
<tr>
<td>Human T-cell lymphotropic virus, type 1</td>
<td>Immortalization and transformation of T cells</td>
<td>[85]</td>
</tr>
<tr>
<td>Helicobacter pylori</td>
<td>Inflammation, oxidative stress, altered cellular turnover and gene expression, methylation, and mutation</td>
<td>[86]</td>
</tr>
<tr>
<td>Clonorchis sinensis</td>
<td>–</td>
<td>[87]</td>
</tr>
<tr>
<td>Opisthorchis viverrini</td>
<td>Inflammation, oxidative stress, and cell proliferation</td>
<td>[88]</td>
</tr>
<tr>
<td>Schistosoma haematobium</td>
<td>Inflammation and oxidative stress</td>
<td>[89]</td>
</tr>
</tbody>
</table>

Table 3. Results of early case–control studies of cervical cancer using different HPV assays

<table>
<thead>
<tr>
<th>Assay</th>
<th>Odds ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serological [25]</td>
<td>5.7 (2.1, 15.4)</td>
</tr>
<tr>
<td>DNA, no PCR [26]</td>
<td>3.7 (3.0, 4.5)</td>
</tr>
<tr>
<td>DNA with PCR [24]</td>
<td>46.2 (18.5, 115)</td>
</tr>
</tbody>
</table>

main classes of mechanisms of action of carcinogens

Carcinogens induce their deleterious effects through various mechanisms. Here, we only discuss the major mechanisms that contributed to identification of carcinogens in the monograph volume 100 series, as examples of such mechanisms of action. We do not discuss the mechanistic evidence for the agents that were classified as group 1 carcinogens because of sufficient evidence from human studies. Also, our intention was not to agree or disagree on how mechanistic data have been used in the evaluations, but just to show them as examples of how molecular epidemiology has been used to reach the evaluations.

exposure and interaction with cellular targets: DNA and protein adducts and ah receptor

A wide range of genotoxic carcinogens can covalently bind with DNA and make DNA adducts. These adducts could
indicate the biologically effective dose of carcinogens [28]. In addition to exposure to external factors, several endogenous factors, including oxidative stress, can lead to DNA adduct formation [29]. The frequency of DNA adducts is also influenced by DNA repair ability [28]. In animal models, DNA adduct formation has been shown to be a required step (but might not a sufficient step) in cancer development [30]. In humans, DNA adducts have been also suggested as biomarkers of cancer risk [27, 31]. DNA adducts and A>T>T:A transversions in the TP53 gene similar to those in experimental animals exposed to aristolochic acid have been observed in urothelial tumors from aristolochic acid nephropathy patients and Balkan endemic nephropathy patients; these two types of nephropathy have very similar clinical and morphological characteristics. This evidence contributed to classification of aristolochic acid as an IARC group 1 carcinogen [32]. The ability for formation of DNA adducts, in combination with other mechanistic pathways, have also contributed to the

Table 4. Group 1 agents with less than sufficient evidence in humans but with strong mechanistic evidence in humans

<table>
<thead>
<tr>
<th>Agents</th>
<th>Evidence from cancer studies in humans</th>
<th>Evidence from cancer studies in animals</th>
<th>Mechanistic evidence in humans*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaldehyde associated with consumption of alcoholic beverages</td>
<td>Inadequate</td>
<td>Sufficient</td>
<td>Very higher risk of certain upper aerodigestive tract cancers in aldehyde dehydrogenase deficient populations in genetic epidemiology studies</td>
</tr>
<tr>
<td>Areca nut</td>
<td>Inadequate</td>
<td>Sufficient</td>
<td>Genotoxicity, including sister chromatid exchange (SCE) and micronuclei (MN) formation; induction of oral preneoplastic disorders with high propensity to progress to malignancy; areca nut is the primary ingredient in betel quid, a carcinogen in humans</td>
</tr>
<tr>
<td>Aristolochic acid</td>
<td>Limited</td>
<td>Sufficient</td>
<td>Genotoxicity, including DNA adducts and A&gt;T&gt;T:A transversions in TP53 in patients with severe renal nephropathy or urothelial tumors</td>
</tr>
<tr>
<td>Benzo[a]pyrene</td>
<td>Inadequate</td>
<td>Sufficient</td>
<td>Genotoxicity, including SCE, MN formation, DNA damage and 8-oxo-deoxyguanosine, and specific diolepoxide-induced DNA adducts which cause TP53 and K-RAS mutations</td>
</tr>
<tr>
<td>Benzidine, dyes metabolized to Ethylen oxide</td>
<td>Inadequate</td>
<td>Sufficient</td>
<td>Strong mechanistic evidence indicating that these dyes are converted to benzidine, a carcinogen to humans, and produce DNA adducts and other genotoxic effects similar to those of benzidine</td>
</tr>
<tr>
<td>Neutron radiation</td>
<td>Inadequate</td>
<td>Sufficient</td>
<td>Genotoxicity, including SCE, aneuploidy, mixed lineage leukemia (MLL) gene translocations</td>
</tr>
<tr>
<td>4,4'-Methlenbis (2-chloroaniline) (MOCA)</td>
<td>Inadequate</td>
<td>Sufficient</td>
<td>Genotoxicity similar to other aromatic amines that are known to cause cancer of the urinary bladder in humans, including DNA adducts, protein adducts, SCE, MN formation in urothelial cells of exposed workers</td>
</tr>
<tr>
<td>4-(N-Nitrosomethylamino)-1-(3-pyridyl)-1-butane (NNK), N-Nitrosonornicotine (NNN)</td>
<td>Inadequate</td>
<td>Sufficient</td>
<td>Genotoxicity, including DNA and hemoglobin adducts. The uptake and metabolic activation of NNK and NNN have been clearly documented in smokeless tobacco users</td>
</tr>
<tr>
<td>Neutron radiation</td>
<td>Inadequate</td>
<td>Sufficient</td>
<td>Similar but more severe damage than from gamma rays, including structural and numerical chromosomal aberrations (including rings, dicentrics, and acentric fragments)</td>
</tr>
<tr>
<td>3,4,5,3’,4’-Pentachlorobiphenyl (PCB-126)</td>
<td>Inadequate</td>
<td>Sufficient</td>
<td>Strong evidence of action through the same Ah receptor-mediated mechanism (based on animal studies) as 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), a carcinogen in humans</td>
</tr>
<tr>
<td>2,3,4,7,8-Pentachlorodibenzofuran (2,3,4,7,8-PCDF)</td>
<td>Inadequate</td>
<td>Sufficient</td>
<td>Strong evidence of action through the same Ah receptor-mediated mechanism (based on animal studies) as TCDD, a carcinogen in humans</td>
</tr>
<tr>
<td>Ultraviolet radiation</td>
<td>Inadequate*</td>
<td>Sufficient</td>
<td>Specific C&gt;T transition in TP53 in squamous cell carcinoma and solar keratosis at DNA bases where known photoproducts could be formed</td>
</tr>
</tbody>
</table>

*aMechanistic evidence from animal studies is not included in this table (except for PBC-126 and 2,3,4,7,8-PCDF).

*bThere is sufficient evidence in humans for the carcinogenicity of solar radiation, the use of ultraviolet-emitting tanning devices, and welding (another source of ultraviolet radiation).
classification of benzo[a]pyrene (B[a]P), 4,4′-methlenbis (2-chloroaniline) (MOCA), 4-(N-nitrosomethylamino)-1-(3-pyridyl)-1-butaneone (NNK), and N-[4]-nitrosonornicotine (NNN) as IARC group 1 carcinogens [19, 22]. NNK and NNN are the most abundant strong carcinogens in smokeless tobacco; their uptake and metabolic activation has been clearly documented in smokeless tobacco users.

Similar to DNA adducts, various carcinogens can lead to protein adduct formation. Protein adducts are generally stable, and unlike most DNA adducts, they are not enzymatically repaired [33]. Protein adducts are suitable for being used as biomarkers of biologically effective dose of exposures. For this purpose, globin and albumin are usually used, because they have relatively long lifetime and are obtainable from blood samples. However, other biological samples, such as hair and tissue, and other proteins, including histones and collagen, may also be used. Although protein adducts are not mechanistically involved in carcinogenesis, they can be good surrogates for DNA adducts and have been linked to cancer risk [34, 35]. Therefore, they have also been used for estimating cancer risk [34, 36]. A dose-related increase in frequency of ethylene oxide-derived hemoglobin adducts in humans and rodents exposed to ethylene oxide and a dose-related increase in the frequency of ethylene oxide-derived DNA adducts in exposed rodents have been shown [19]. Ethylene oxide also has mutagenic and elastogenic effects and can induce sister chromatid exchange (SCE), chromosomal aberrations, and micronucleus formation in the lymphocytes of exposed workers, as well as heritable translocations in the germ cells of exposed rodents [19]. The ability of ethylene oxide in producing protein adducts was one of the mechanisms for which this agent was classified as an IARC group 1 carcinogen [19]. Protein adduct formation was also one of the several mechanisms by which B[a]P was classified in the same category [19].

The Aryl hydrocarbon receptor (Ahr) is a ligand-dependent cytosolic transcription factor that is normally inactive. Following ligand binding to certain chemicals, including polychlorinated biphenyls (PCBs), B[a]P, and dioxin-like compounds, Ahr translocate into the nucleus and dimerize with AhR nuclear translocator, leading to changes in gene transcription, gene expression, cell replication, and apoptosis [37]. The complex of Ahr/AhR nuclear translocator also induces activity of several drug-metabolizing enzymes, notably xenobiotic phase I metabolizing enzymes, such as cytochrome P450 (CYP) A1, CYP1A2, and CYP1B, and phase II enzymes, such as glutathione S-transferase [38]. Ahr is involved in tumor initiation and plays a role in tumor promotion and perhaps in tumor progression [39]. There is strong evidence that TCDD has carcinogenic effects in humans through receptor-mediated mechanisms: the primary mechanism is the modification of cell replication and apoptosis and subsequently promotion of tumor development; the secondary mechanism relates to increases of oxidative stress which leads to DNA damage [19]. Based on this mechanistic evidence, TCDD was upgraded as an IARC Group I carcinogen in one of the earlier monographs [40]; in the evaluation of this compound in monograph 100, the working group also concluded that there was sufficient evidence for carcinogenicity of TCDD in humans based on epidemiological studies [19]. There is strong evidence supporting similar Ahr-mediated mechanisms for carcinogenicity of 3,4,5,3′,4′-pentachlorobiphenyl and 2,3,4,7,8-pentachlorodibenzo-furan in humans based on their mechanisms of carcinogenicity in experimental animals [19], which provided strong mechanistic evidence for classifying these two compounds as IARC group 1 carcinogens [19].

Specific genetic alterations: TP53 and k-RAS mutations
TP53 is a tumor suppressor gene that encodes a transcription factor responding to several forms of cellular stress. The p53 protein is involved in the response pathways against genotoxic DNA damage [41] and regulation of the activity of several genes, including some of the genes involved in the cellular cycle control (e.g. p21 gene). The mutated TP53 found in some neoplasms is not able to activate p21, allowing for an uncontrolled cell growth [42–44]. Somatic TP53 gene alterations have been reported in many cancer types. The frequency of alterations (5%–80%) depends on the type, stage, and etiology of tumors [45]. The majority (75%) of TP53 mutations are missense substitutions; the frequency of other main alterations is as following: frameshift insertions and deletions (9%), nonsense mutations (7%), and silent mutations (5%) [46]. In addition to endogenous factors, several environmental factors, including ionizing radiation, ultraviolet rays, and oxidizing and cytotoxic agents, can induce TP53 mutations [47]. Exposure to ultraviolet radiation induce mutations in several genes in human cell model systems, and mutations have been detected in several genes in human tumors at DNA bases where known photoproducts could be formed, including the TP53 gene in squamous cell carcinoma of the skin and solar keratosis [48]. This supports a major role of these mutations in carcinogenicity of ultraviolet radiation and has contributed to classification of ultraviolet radiation as an IARC group 1 carcinogen. Also, the ability in inducing TP53 mutations was one of the mechanistic pathways that contributed to aristolochic acid and B[a]P being upgraded to IARC group 1 carcinogens [32].

RAS genes encode Ras proteins, which are among Guanosine-5′-triphosphate-binding proteins [49]. Normal Ras proteins are involved in the control of several cellular functions, including proliferation, differentiation, and survival [50]. In ~30% of human cancers, activating mutations in RAS have been reported [51]. This activating transformation is implicated in various cancers. The major categories of RAS genes, including K-, H-, and N-RAS, seem to be primarily associated with different types of cancers, among which K-RAS are mainly associated with lung, colon, pancreas, and biliary tract cancers [52]. K-RAS mutation can be induced by various environmental exposures, including polycyclic aromatic hydrocarbons (including B[a]P) and aristolochic acid [51, 53]. K-Ras mutations, as well as some other mechanistic pathways, contributed to classification of B[a]P and aristolochic acid as IARC group 1 carcinogens [19, 32].

Structural genomic alterations
Several structural alternations in the genome have been used in cancer studies. Some of these alternations, including SCE and micronuclei (MN) formation, are unspecific, whereas some
others, including mixed lineage leukemia (MLL) translocations, are specific to certain cancer types. SCE may occur during DNA replication. Two sister chromatids break and rejoin with one another, which results in symmetrical exchange of DNA segments between the sister chromatids [54]. As the exchanges are symmetrical, SCEs themselves do not necessarily cause adverse health outcomes, and they are not considered mutations [55]. Although mechanisms of SCE formation are not clear, SCEs can be induced by UV light and a large number of genotoxic chemicals, including tobacco use [56]. Genotoxic factors can induce high SCE levels in vitro and in animal models, and quantification of these levels is generally easy [54]. Therefore, SCE test has been used to investigate the DNA damage in blood lymphocytes as an indicator of exposure to genotoxic factors. However, high SCE frequencies have not been shown as a predictor of cancer [57]. This may be attributable to high interindividual variability in exposure to genotoxic factors and DNA repair capacity, the two main factors that determine SCE levels [57].

MNs are small cytoplasmic bodies having a portion of acentric chromosome or the whole chromosome, which were not carried during the anaphase to the opposite poles, due to defect in the repair of DNA breaks [58]. This results in missing of part or whole chromosomes in the daughter cell. MN formation is influenced by age, gender, diet, several lifestyle factors, including exercise, alcohol drinking, and smoking, and some genetic defects and hereditary disorders [59]. MN test have shown a high reliability in identifying genome damage in genotoxicity studies [60]. Association between MN frequency and risk of several types of cancer has been reported in epidemiological studies, including prospective studies [60, 61].

Induction of SCE and MN by an agent per se is not generally considered as strong evidence for carcinogenicity of the agent, but it may support carcinogenicity of agents in the presence of other mechanistic evidence. SCE and MN formation have been reported with exposure to many carcinogenic agents; here, we present two examples. Areca nut chewing, alone or in combination with other substances, is a practice in Southern Asia [22]. Areca nut is the primary ingredient in betel quid, which is a known carcinogen to humans [22]. Continuous local irritation of buccal epithelial cells by areca nut can generate chronic inflammation, oxidative stress, and cytokine production, and subsequently genotoxic effects, including SCE and MN, in exposed oral keratinocytes with uncontrolled proliferation and hyperplastic/dysplastic lesions [22]. With long-term use, these preneoplastic lesions may progress to malignancy. Based on this evidence, areca nut chewing was upgraded from IARC carcinogenicity group 2A to group 1 [22]. The mechanisms for which MOCA was upgraded to a group 1 carcinogen were genotoxic activities similar to those of other aromatic amines that are known to cause bladder cancer in humans, including DNA and protein adducts, SCE, and MN formation [19].

MLL protein is a histone methyltransferase that seems to be involved in the epigenetic maintenance of transcriptional memory [62]. Translocations in MLL gene (in 11q23) are involved in human leukemia [63]. Overall, MLL translocations are found in ~10% of human acute leukemia cases, but there is a large variation in this proportion depending on the type of leukemia and the age of patients (reported in >70% of infant leukemia and in <10% of leukemia in older children) [64, 65]. There is sufficient evidence from cancer studies in human that etoposide, a topoisomerase-II-inhibiting chemotherapy agent, in combination with cisplatin and bleomycin can cause acute myeloid leukemia [32]. Each of these chemotherapy agents are genotoxic, with evidence of induction of DNA damage, chromosomal aberrations, and aneuploidy. However, etoposide alone has been classified as IARC group 1 human carcinogens because, unlike the other two drugs, it can induce chromosomal translocations affecting the MLL gene in cultured mouse and human cells [32]. Among treatment-related 11q23 translocations in leukemia patients, ~85% have been treated with topoisomerase-II-inhibiting drugs, primarily etoposide [66].

challenges and future perspectives

Molecular epidemiology has contributed to identification of several carcinogens. However, with the exception of biological agents, its contribution in terms of providing sufficiently strong evidence as required by the IARC monographs has been modest. Molecular epidemiology may provide evidence for exposure to and biological effects of carcinogens. This can be particularly helpful with mixed compounds, as identification of the effect of individual exposures in traditional epidemiologic studies could be challenging. As an example, molecular epidemiology has provided information on exposures or has been used for validation of exposure assessment methods in occupational studies, in which individuals are usually exposed to several compounds simultaneously, and these exposures differ by job type, location, and time [6]. Also, as discussed above, the carcinogenic effects of several such compounds have been shown in molecular epidemiological studies.

Nevertheless, biological markers of exposure and mechanisms of actions for many other carcinogens need to be identified. Some cases of cancer may be associated with long-term, low-dose exposures to carcinogens [67, 68]. Owing to temporal variability of exposures, assessment of biomarkers in a single biological sample may not provide accurate information about lifelong exposure to the respective carcinogens. Even in cases of cancers associated with high exposures in a relatively short time, samples may not be collected in the most appropriate time for assessment of the exposure. Therefore, attempts should be made to assess exposures in repeated samples in different periods to have a more comprehensive picture about patterns of exposure. The concept of ‘exposome’, assessment of exposure from perinatal period until adulthood [69, 70], has been introduced to address the variability of exposure in epidemiological studies, which also applies to molecular epidemiology.

Another challenge in molecular epidemiology is that its contribution has mainly been limited to the identification of carcinogens acting through a genotoxic mechanism, i.e. when carcinogens induce DNA damage. There is increasing understanding of the role of nongenotoxic carcinogens. Some examples of such carcinogens are certain hormones (such as 17β-estradiol and androstenedione) [71, 72], immunosuppressants (such as ciclosporine) [73, 74], metals (such as arsenic) [75, 76], and overweight/obesity [77, 78].
These carcinogens may act through different mechanisms, including receptor modification, effects on immune system, and epigenetic changes, but these mechanisms are not clear in the majority of cases. The potential role of nongenotoxic mechanisms in carcinogenicity of IARC group 1 carcinogens is reviewed elsewhere [79]. Overall, as nongenotoxic mechanisms of carcinogenicity are not well identified and their biomarkers are sparse, the contribution of these mechanisms to identification of human carcinogens has been limited. This area of molecular epidemiology deserves priority through the development of biological markers of nongenotoxic mechanisms and alterations that are specific to carcinogens.

**Disclosure**

The authors have declared no conflicts of interest.

**References**
