OPINION

Impact of race/ethnicity on molecular pathways in human cancer

John K. Wiencke

Understanding the molecular circuitry of the cancer cell is within the grasp of the basic scientist; however, harnessing this knowledge to predict cancer risk requires integration of molecular and population sciences. But, what role, if any, does race/ethnicity have in cancer research and, more specifically, in the nature of genetic and epigenetic alterations that programme the malignant behaviour of the cancer cell?

Race, as it is used in common discourse, is a subdivision of a species formed by a group of individuals that share common biological characteristics that distinguish them from other groups[1]. The concept of ethnicity emphasizes cultural, socioeconomic, religious and political qualities of human groups, including language, diet, dress, kinship relation systems and historical or territorial identity[2]. The United States Census and biomedical researchers collect both types of data to categorize populations. There is abundant epidemiological evidence that self-identified race/ethnicity is associated with differences in cancer incidence and mortality. For example, over the 5-year period ending in the year 2000, national cancer statistics from the United States show an average annual prostate cancer incidence of 277 per 100,000 for African-American men compared with 168 per 100,000 among Caucasians. Racial differences in death rates for the disease were even more evident. An average of 73 prostate cancer deaths per 100,000 for African-American men compared with 30 per 100,000 for Caucasians were recorded[3].

Another example is early-onset breast cancer, which is more common among African-American compared with Caucasian women, and breast cancer mortality is higher among African-American women in all age groups[4–8]. By contrast, certain minority populations have reduced risks of developing some types of cancer. Primary brain tumours are more common in Caucasians, compared with minority non-whites[9]. African Americans were reported to have lower survival rates after diagnosis of primary brain tumour compared with Caucasians[10], whereas another study reported a higher incidence of survival among African Americans[11] . To address the complex issues regarding cancer risk, race and ethnicity, data are commonly collected by health researchers. This information can be used to obtain information about social class, possible environmental exposures and genotype.

There is far from a consensus on the value of racial information in cancer research. It has been argued that racial categories are no longer useful in aetiological research because they are too vague and imprecise[12]. Others point to the use of such classification schemes for epidemiological and clinical investigations[13,14]. Moreover, the political ramifications of collecting racial data continue to be intensely debated. In California's special recall election that was held on 7 October 2003, voters rejected the Racial Privacy Initiative (Proposition 54), which sought to ban the state from collecting racial data in all but a few exempted cases. Sixty-four percent of voters voted against the proposal, reflecting the concern that limiting the collection of racial information would slow the progress of cancer research.

Ancestry and racial categories

To help answer the question of whether there is a valid biological meaning to racial categories and whether these categories might help to explain the molecular features and aetiological heterogeneity of cancer in different populations, we can turn to the work of evolutionary biologists and population geneticists. Studies that use molecular-marker analysis show that human populations worldwide can be subdivided into groups that are consistent with race, based on ancestry within one of five continents[15]. These groups include African, Caucasian (European and Middle Eastern), Asian, Pacific Islander and Native American. DNA markers, including short tandem repeats (minisatellites) and single-nucleotide polymorphisms, have been used to determine relatedness and lineage within human populations (FIG. 1).

An example of such a marker is the Duffy-blood-group antigen, a glycosylated protein that was first recognized as the erythrocyte receptor for the human malaria parasite *Plasmodium knowlesi*[16]. A point mutation within the gene locus for Duffy (*FY*), which is located at 1q21-1q22, leads to lack of expression of the Duffy antigen in red blood cells. This mutation is very rare in most racial groups, but is present in 100% of

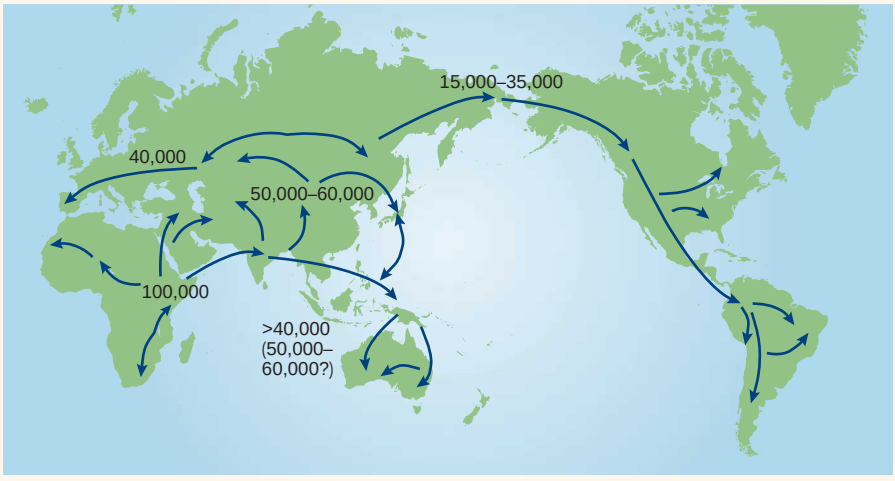

Figure 1 | **Classification of major racial/ethnic groups, based on the migration of modern *Homo sapiens*.** Genetic differentiation of humans according to their migration patterns and establishment of genetically isolated populations over time provides the basis for racial categories according to continental ancestry. The scheme that is outlined above begins with a radiation from east Africa to the rest of Africa about 100,000 years ago, and is followed by an expansion from the same area to Asia — probably by two routes, southern and northern — between 60,000 and 40,000 years ago. Oceania, Europe and America were settled from Asia in that order. Genetic divergence is brought about by relative isolation of groups in different environments and through the actions of genetic drift and differential natural selection. Figure adapted from REF. 15 © (2003) Nature Publishing Group.

Table 1 | Examples of potential differences in cancer type by race/ethnicity

Cancer	Population	Number of subjects in study	Molecular marker	Patient/tumour characteristics	References
Gastric	Texan Hispanics, Caucasians, African Americans	107	*CDKN2A* methylation	EBV infection more common in Texan Hispanics; no difference between Caucasians and African Americans	46
Glioma	Japanese	63	*CDKN2A* deletion	*CDKN2A* deletion less common in tumours from Japanese compared with Caucasians	47
Glioma	Northern Californian Caucasians, non-Caucasians	172	*TP53* mutation	Higher prevalence of *TP53*-mutation-positive gliomas in non-Caucasians	51
Breast	Detroit African Americans, midwest US rural Caucasians, Scottish Caucasians	75	*TP53* mutation	Higher frequency of all transition-type mutations in African Americans	53
Breast	African Americans, Native Americans, Asian/Pacific Islanders, Hispanic whites, non-Hispanic whites	93,317	ER/PR receptor	Greater risk for ER/PR-negative breast cancer and different histological profiles in ethnic minorities	56
Pancreatic cancer	Detroit African Americans, Caucasians	410	*KRAS* mutation and mutational spectra	Similar frequency; different *KRAS* mutational spectra	58
Lung cancer	Louisiana African Americans, Caucasians	111	*KRAS* mutation and mutational spectra	Increased prevalence of mutant *KRAS* in African Americans, but same mutational spectra	59
Colorectal cancer	African Americans	22	Microsatellite instability	Threefold higher prevalence of high-grade microsatellite instability in tumours from African Americans	61
Paediatric acute leukaemia	Caucasians, African Americans, Hispanics, Asians, mixed & others	8,447	Cytogenetic profile, immunophenotype	Higher risk for T-cell phenotypes, lower risk for hyperdiploid karyotype, shorter EFS and poor outcomes among standard risk categories for African-American children	62–65

EBV, Epstein–Barr virus; EFS, event-free survival; ER/PR, oestrogen receptor/progesterone receptor; US, United States.

native Africans and about 70% of African Americans. The mutation has been shown to occur 46 base pairs upstream of the transcription initiation site, in the protein's consensus binding site for the transcription factor GATA1, leading to loss of Duffy expression[17]. There are many such loci that display large differences in allele frequencies among ancestral populations[18,19]. Genetic determinants of cancer risk could be linked to these ancestral associations, and further study of these racial categories could help to identify new susceptibility loci.

There are several ways in which race/ethnicity could affect the results and interpretation of cancer studies. Certain ancestral populations carry mutations or polymorphisms in genes that encode proteins thought to be directly involved in carcinogenesis. Several notable examples are found within Ashkenazi-Jewish populations. The first involves founder mutations in the *BRCA1* and *BRCA2* genes, which are associated with breast and ovarian cancer. Founder mutations are those that occur in a specific population and that were introduced to the group by an ancestor in whom the original mutation occurred. Two mutations in *BRCA1* (185delAG and 5382insC) and one mutation in *BRCA2* (6174delT) are common in the Ashkenazi Jewish population[20–22] — the *BRCA1* 185delAG mutation has an approximately 1% prevalence[23,24]. Investigators have yet to agree on whether the clinical and pathological characteristics of early-onset breast or ovarian cancer in carriers of these founder mutations are different from those in non-carriers[25–28].

A second example that has arisen in Ashkenazi-Jewish populations involves a common genetic variation in the adenomatous polyposis coli (*APC*) gene, which might cause a predisposition to colorectal cancer. A transversion from T to A at codon 1307 (I1307K) in the *APC* gene converts the wild-type sequence to a homopolymer tract that is thought to be unstable and prone to mutation[29]. *APC* I1307K is found in about 6% of Ashkenazi-Jewish individuals[30–32]. Whereas early studies indicated a modestly increased risk for colorectal cancer and unique molecular features of the tumours among *APC* I1307K carriers compared with non-carriers[29,31,33], more recent studies have led to questions about the importance of this polymorphism as a risk factor for colorectal cancer[34]. One potentially fruitful approach to clarifying the molecular and epidemiological features of cancer that are associated with either the *BRCA1*/*BRCA2* founder mutations or the *APC* I1307K polymorphism is to explore the possible interactions of these variants with other genes and environmental influences.

Certain race/ethnicity data have also been associated with exposure to specific cancer-causing agents. Socioeconomic factors (for example, income or education) are often linked to environmental exposures that are important in modifying cancer risk. For

Table 2 | Cancer risk factors associated with methylation of *CDKN2A*

Exposure/cancer risk factor	Association with *CDKN2A* methylation	Cancer	References
Cigarette smoking	Increased frequency of methylation	Non-small-cell lung cancer	71–75
Cigarette smoking	Increased frequency of methylation	Squamous-cell head and neck cancer	76
Human papillomavirus infection	Decreased frequency of methylation	Squamous-cell head and neck cancer	77
Epstein–Barr virus infection	Increased frequency of methylation	Gastric cancer	46,78
Epstein–Barr virus infection	Increased frequency of methylation	Nasopharyngeal cancer	79
Hepatitis B and C infection	Increased frequency of methylation	Hepatocellular carcinoma	78,80,81
HIV-1/KSHV infection	Increased frequency of methylation	Kaposi sarcoma, HIV-1-associated lymphoma	82,83

HIV, human immunodeficiency virus; KSHV, Kaposi's sarcoma herpesvirus.

example, tobacco smoking is more prevalent among some ethnic groups compared with others[35] (for example, African Americans versus Latinos), and tobacco-related cancer risks are greater for smokers compared with non-smokers. Clearly, direct measurements of carcinogen exposure would be a more valid approach to risk analysis than collection of race/ethnicity information. But the specific environmental factors that cause cancers are often ill-defined, unknown or inaccessible.

The unique genetic features of racial groups, in combination with environmental factors, can also influence carcinogenic mechanisms and lead to biologically important differences in the molecular profile of a tumour. These racial/ethnic differences therefore determine not only cancer risk, but also potential responses to preventative measures and treatment. Recent epidemiological and clinical studies using molecular markers indicate that racial differences in cancer types do exist (TABLE 1), so a systematic evaluation of these issues is appropriate.

Cancer cells to human populations

At the heart of approaches to studying cancer at the population level is the idea of a 'web of causality' that underlies any complex disease. Applying this concept to carcinogenesis, it is likely that cellular control pathways are subject to disruptions through distinct mechanisms that are triggered by different combinations of environmental and genetic factors. For example, the main risk factor for liver cancer is cirrhosis, which is often a result of hepatitis B (HBV) or hepatitis C (HCV) infection. Other conditions that modify liver cancer risk and form the 'web' of causal factors include dietary exposure to fungal toxins (for example, aflatoxin), age, sex, duration and severity of liver disease, concurrent alcohol consumption, and genetic conditions that lead to iron accumulation in the liver (haemachromatosis)[36,37]. Several oncogenic pathways have been implicated in malignant transformation of liver cells. Mutations and allelic deletions in *TP53*, which are found in about 30% of liver cancer cases, have been associated predominantly with exposure to aflatoxin B1 and HBV infection. Conversely, mutations in the gene that encodes β-catenin occur in about 22% of liver cancer cases, but are rare in HBV-associated tumours[38].

Another body of evidence to show that specific genomic alterations are associated with environmental risk factors can be found in looking at the prevalence in different cancers of epigenetic inactivation of the *CDKN2A* gene — which encodes a cyclin-dependent kinase inhibitor, INK4A (also known as p16), that regulates the retinoblastoma (Rb) cell-cycle control pathway[39]. Differential inactivation of *CDKN2A* provides one example of a gene that is altered by environmental factors. This gene has been found to be inactivated through single-base mutation, chromosomal deletion and an epigenetic mechanism that involves aberrant methylation within its promoter region, leading to cancer[40]. This aberrant promoter methylation inhibits *CDKN2A* expression and leads to defects in cell-cycle control, a common event in transformation. The epigenetic mechanism for disruption of *CDKN2A* expression has been linked to different aetiological agents, including viral exposure and cigarette smoking, in several tumour types (TABLE 2).

Therefore, if environment can influence cancer type at the molecular level, it follows that differences in exposure patterns among racial/ethnic subgroups might lead to differences in cancer susceptibility, irrespectively of any intrinsic genetic differences between groups. For example, levels of exposure to several viral agents (TABLE 2) show regional and ethnic variations[41,36,42]. A mixture of aetiological factors could therefore be important in determining cancer risk. Interactions between environmental and genetic factors should also be considered in determining cancer susceptibility. However, many of the environmental triggers that underlie common cancers remain unknown, or can only be defined imprecisely. So, can race/ethnicity data add any significant information, above and beyond the known exposure-risk categories, to help identify different causal pathways and risk groups?

Variations in cellular control pathways

To bring into sharper focus the contributions of race/ethnicity to cancer, it is useful to look at the problem from the perspective of variations in the cellular control pathways that are commonly linked to cancer.

Cell cycle. The tumour-suppressor gene *CDKN2A* is methylated and therefore inactivated in some virus-associated tumours (TABLE 2), but other mechanisms for disrupting the INK4A–Rb pathway exist in different tumour types. About 10% of gastric adenocarcinomas have a distinct histology, proximal anatomic location, male predominance and are associated with Epstein–Barr-virus (EBV) infection[43,44]. Inactivation of *CDKN2A* through aberrant promoter methylation is more common in EBV-related gastric cancers than in non-EBV-associated tumours[45]. A multi-ethnic study by Vo *et al.*[46] showed that the presence of EBV and silencing of *CDKN2A* by methylation was significantly more common in gastric tumour samples that were taken from Texan Hispanics compared with those from non-Hispanic whites or African-Americans, and was also more common among men (TABLE 1). These findings indicate that there are ethnic differences in tumour virology and in gastric cancer pathogenesis, although future studies are needed to determine whether EBV exposure alone or in combination with other host factors underlies these associations.

In a second study, inactivation of *CDKN2A* by chromosomal deletion was less common in malignant glioma samples from Japanese patients compared with those from Caucasians[47]. The clinical significance of this difference is unknown, but given the central role of the Rb cell-cycle

Box 1 | The IARC *TP53* Mutation Database

In 1991, a database of somatic *TP53* mutations in human cancers and cell lines was initiated by Monica Hollstein and Curtis Harris. Since 1994, this database has been maintained at the International Agency for Research on Cancer (IARC) in Lyon, France, and is made freely available as a service to the scientific community. The IARC *TP53* Mutation Database can be used for the following purposes:

- To perform regular reviews of the TP53 mutation literature.
- To develop electronic formats for compiling, sorting and retrieving mutation data.
- To perform research on *TP53* mutation patterns.

The current version of the database is 'R8', which was released in June 2003. The R8 dataset includes 18,585 somatic mutations that were reported in 1,680 original publications and 225 germline mutations that were reported in 98 publications (published between 1989 and June 2002). Functional information on more than 200 p53 mutant proteins is now available. The database can be accessed at http://www.iarc.fr/p53/.

checkpoint in gliomagenesis, further studies might reveal defects at other points in the Rb pathway in tumours from Japanese patients. With respect to aetiology, the causes of glioma in adults are obscure. Factors such as increasing age (up to age 80), male gender and Caucasian non-Hispanic race/ethnicity are all associated with increased risk. In fact, race is one of the few factors that is consistently associated with risk for this devastating cancer.

Apoptosis. The control of apoptosis by p53 is another pathway that varies depending on race/ethnicity. The p53 protein mediates the cellular response to DNA damage and proliferative signals, and selectively activates different subsets of target genes that can modulate apoptosis, growth arrest, DNA repair or differentiation[48,49]. The presence and nature of *TP53* mutations has been proposed for use as a tool to identify carcinogen exposure, and could also be used to determine the influence of race/ethnicity on human carcinogenesis. Mutations in *TP53* are among the most common events in human cancer, as some 18,585 acquired mutations have been catalogued in the International Agency for Research on Cancer (IARC) TP53 mutation database[50] (BOX 1). It is important to note, however, that less than 10% of these can be linked with exposure to specific environmental factors, so further molecular epidemiological studies are required.

To address this issue, we recently carried out a study in the San Francisco Bay area to identify associations between characteristics such as race and ethnicity with the presence and type of *TP53* mutation in a population-based sample of adult gliomas. Surprisingly, tumours from non-whites were five times more likely to have mutations in exons 5–8 of *TP53*, and there were also subtle differences in the mutational spectra that were observed in different ethnic groups[51]. Gliomas that contained *TP53* mutations and that were found more commonly in non-whites could arise from lower-grade malignancies that recurred later as high-grade aggressive glioblastoma multiforme (GBM). Clinical outcome for patients with GBM are dismal irrespective of p53 status; only 2–5% of patients who are originally diagnosed with GBM will survive for more than 3 years[52]. Potential environmental exposures, as well as germline genetic differences that are associated with *TP53* mutations in patients with glioma, are under investigation.

Breast cancer provides another example of the association between race/ethnicity and *TP53* mutations, cancer risk and prognosis. African-American women with breast cancer have a worse prognosis compared with other groups in the United States. One comparison of the mutational spectra within breast tumours from women of different ethnic backgrounds from the United States reported significantly higher proportions of transition-type mutations in *TP53* in tumours from African-American women, compared with Caucasians[53]. The issue of racial differences in breast cancer incidence and prognosis has been extensively examined, although age of onset, as well as other potential biases in reporting, could account for some of the reported differences[54,55]. In this regard, it is important to note the Surveillance Epidemiology and End Results (SEER) study of 95,523 patients with breast cancer, all of whom were more than 50 years old. This study found that women from ethnic minorities have a greater risk of oestrogen-receptor/progesterone-receptor-negative breast cancer and that their tumours showed different histological profiles, compared with those of non-Hispanic white women[56]. These findings could partly explain the reported poorer survival among these populations.

The most widely cited example of a link between *TP53* and environmental carcinogen exposure occurs in tobacco-related cancers. In terms of the association between smoking and *TP53* mutational spectra, race has not been adequately addressed. Although the IARC Mutation Database (BOX 1) has been updated over time to include classification of smoking status, many of the data lacks annotations on race. A recent analysis has indicated that racial differences could be very important in determining tobacco-related cancer risk, and that the widely cited association between G→T transversions and lung cancer might be an artefact of the unequal distribution of racial groups that have been assigned to smoker and non-smoker categories[57].

Proliferation. Genomic alterations that affect proliferative signals are reported to vary by race/ethnicity. In a study of 410 patients (166 African Americans and 244 Caucasians) with a histological diagnosis of pancreatic ductal adenocarcinoma, patients from the two races/ethnicities were compared according to the clinicopathological characteristics of their tumours, including the presence and types of *KRAS* mutations at codon 12. Codon 12 contains the most common activating mutation in human cancer. African Americans had more frequent *KRAS* mutations that resulted in glycine→valine amino-acid substitutions than Caucasians[58]. These studies could be relevant to the observation that African Americans have a higher incidence of pancreatic adenocarcinoma than do Caucasians.

Lung cancer risks and mortality rates are higher among African-American men than any other group in the United States. Analysis of lung tumours from African Americans in the Mississippi River corridor in Louisiana — a region with very high mortality rates from lung cancer — showed that 32/116 (27.6%) contained *KRAS* mutations in either codon 12 or 13. This frequency is comparable to that reported for Caucasians, although the mutation spectrum was strikingly different. Of the 32 mutations observed, an abnormally high proportion of cysteine and serine mutations was found in lung cancers from African Americans compared with lung cancers in Caucasians that have been reported in the literature[59].

Other mechanisms. Even though there is limited research on the subject, other cancer-related mechanisms have been studied in relation to race/ethnicity. Overall, levels of DNA methylation were reported to be lower

in squamous-cell lung cancers from African Americans compared with Caucasians[60], and high-grade (extensive) microsatellite instability — examined in a case series of colorectal cancers — was more common among tumours from African Americans compared with other groups[61]. Prognostic characteristics have also been found to vary between racial/ethnic groups of children with acute lymphoblastic leukaemia[62]. African-American children with acute leukaemia were more likely to present with high-risk features and to have poorer outcomes compared with Caucasian children[63–65]. The cytogenetic and molecular pathways that are involved in these racial differences have not been identified.

Race and human-genome science

Although scientists debate the value of racial information[12], it is likely to be counterproductive to continue to ignore race while searching for the molecular underpinnings of human cancer. Developments in evolutionary biology and genetics compel us to address the value of ethnic and racial categories to ensure that we do not pass up any opportunity to improve the prospects for cancer prevention and patient outcome, or to gain a more complete understanding of cancer pathogenesis[66,67]. This recommendation does not indicate a static categorization scheme for race/ethnicity, nor one of strict divisions between continental groups, because migration and interbreeding degrade the endogamy that is required to maintain genetically clustered groups. The implications of unique allelic combinations among racial groups for pharmacology have recently been outlined[68,69].

Future directions

It could be said that our approach to investigating cancer pathogenesis has been 'race blind' or 'race neutral'. Many of the time-honoured tools for dissecting the crucial control pathways in cancer are cancer cell lines from patients whose race/ethnicity is unknown. As we move forward, cancer biologists should collect data that are based on the basic demographic characteristics of the patients whose tumours they study. Even the substantial international database of acquired mutations in the *TP53* gene is poorly annotated for race — a fact that could threaten the validity of some of the conclusions that are drawn from this extensive database. Future submissions of specimens to the database should include race/ethnicity information. Clearly, we should encourage a closer collaboration between the population scientist, who views cancer pathogenesis as a multifactorial web of causal processes, and the cancer biologist.

It is clear that exposure to infectious and chemical agents affects the genetic and epigenetic profiles of tumours, and it is also known that these exposures vary according to race/ethnicity. Exposure variability by race/ethnicity is the simplest way in which differences in cancer type can arise. As the specific aetiological exposures that lead to most cancer types are incompletely understood, race/ethnicity information could be useful for understanding how differences among populations can affect carcinogenesis. Attention to racial differences might help in identifying new cancer-causing agents — if a specific cancer is prevalent among one racial/ethnic group, investigations into the lifestyle and environment of that group could uncover a previously unrecognized carcinogen. Clues to help explain race/ethnicity differences in cancer risk and prognosis will come about through combining race/ethnicity information with molecular profiles.

From population genetics, we know that race/ethnicity categories correspond to and help identify unique germline alleles and allelic combinations. These ancestral genetic groupings can modify both cancer risk and the molecular subtype of tumours. It is possible that there are complex interactions between ancestry-specific genes and race-associated environmental exposures. In addition to collecting race/ethnicity data, population-specific polymorphisms have been identified that can be used to directly estimate individuals' ancestry, and this approach could supplant or replace self-reported ancestry[19,70]. It is still too early to assess the full impact of race/ethnicity in human carcinogenesis and many questions have been raised by recent research. But, it is important to note that as therapies evolve that target specific pathway defects, information on differences in cancer pathways between patient groups will become increasingly more clinically relevant. Therefore, in the future, clinical scientists can expect more, not less, emphasis on learning about the racial makeup of the individuals who are involved in clinical trials.

John K. Wiencke is at the University of California San Francisco, Department of Epidemiology and Biostatistics, 500 Parnassus Avenue, San Francisco, California 94143, USA.
e-mail: Wiencke@itsa.ucsf.edu

doi: 10.1038/nrc1257

1. Cavalli-Sforza, L. L. & Bodmer, W. F. *The Genetics of Human Populations* (W. H. Freeman, San Francisco, 1971).
2. Lee, S. S., Mountain, J. & Koenig, B. A. The meanings of "race" in the new genomics: implications for health disparities research. *Yale J. Health Policy Law Ethics* **1**, 33–75 (2001).
3. Weir, H. K. *et al.* Annual report to the nation on the status of cancer, 1975-2000, featuring the uses of surveillance data for cancer prevention and control. *J. Natl Cancer Inst.* **95**, 1276-1299 (2003).
4. Howe, H. L. *et al.* Annual report to the nation on the status of cancer (1973 through 1998), featuring cancers with recent increasing trends. *J. Natl Cancer Inst.* **93**, 824–842 (2001).
5. Ragland, K. E., Selvin, S. & Merrill, D. W. Black–white differences in stage-specific cancer survival: analysis of seven selected sites. *Am. J. Epidemiol.* **133**, 672–682 (1991).
6. Carter, C. L., Allen, C. & Henson, D. E. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. *Cancer* **63**, 181–187 (1989).
7. Boyer-Chammard, A., Taylor, T. H. & Anton-Culver, H. Survival differences in breast cancer among racial/ethnic groups: a population-based study. *Cancer Detect. Prev.* **23**, 463–473 (1999).
8. O'Malley, C. D., Le, G. M., Glaser, S. L., Shema, S. J. & West, D. W. Socioeconomic status and breast carcinoma survival in four racial/ethnic groups: a population-based study. *Cancer* **97**, 1303–1311 (2003).
9. Wrensch, M., Minn, Y., Chew, T., Bondy, M. & Berger, M. S. Epidemiology of primary brain tumors: current concepts and review of the literature. *Neuro-oncol.* **4**, 278–299 (2002).
10. Barnholtz-Sloan, J. S., Sloan, A. E. & Schwartz, A. G. Racial differences in survival after diagnosis with primary malignant brain tumor. *Cancer* **98**, 603–609 (2003).
11. Shaw, E. G. *et al.* Ethnic differences in survival of glioblastoma (GBM): a secondary analysis of the radiation therapy oncology group (ROTC) recursive partioning analysis (RPA) database. *Neuro-oncol.* **5**, 296–297 (2003).
12. Schwartz, R. S. Racial profiling in medical research. *N. Engl. J. Med.* **344**, 1392–1393 (2001).
13. Risch, N., Burchard, E., Ziv, E. & Tang, H. Categorization of humans in biomedical research: genes, race and disease [comment]. *Genome Biol.* **3**, 2007 (2002).
14. Burchard, E. G. *et al.* The importance of race and ethnic background in biomedical research and clinical practice. *N. Engl. J. Med.* **348**, 1170–1175 (2003).
15. Cavalli-Sforza, L. L. & Feldman, M. W. The application of molecular genetic approaches to the study of human evolution. *Nature Genet.* **33** (Suppl.), S266–S275 (2003).
16. Hadley, T. J. & Peiper, S. C. From malaria to chemokine receptor: the emerging physiologic role of the Duffy blood group antigen. *Blood* **89**, 3077–3091 (1997).
17. Tournamille, C., Colin, Y., Cartron, J. P. & Le Van Kim, C. Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nature Genet.* **10**, 224–228 (1995).
18. Smith, M. W. *et al.* Markers for mapping by admixture linkage disequilibrium in African American and Hispanic populations. *Am. J. Hum. Genet.* **69**, 1080–1094 (2001).
19. Parra, E. J. *et al.* Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *Am. J. Phys. Anthropol.* **114**, 18–29 (2001).
20. Neuhausen, S. L. *et al.* Haplotype and phenotype analysis of six recurrent *BRCA1* mutations in 61 families: results of an international study. *Am. J. Hum. Genet.* **58**, 271–280 (1996).
21. Berman, D. B. *et al.* A common mutation in *BRCA2* that predisposes to a variety of cancers is found in both Jewish Ashkenazi and non-Jewish individuals. *Cancer Res.* **56**, 3409–3414 (1996).
22. Moslehi, R. *et al. BRCA1* and *BRCA2* mutation analysis of 208 Ashkenazi Jewish women with ovarian cancer. *Am. J. Hum. Genet.* **66**, 1259–1272 (2000).
23. Struewing, J. P. *et al.* The carrier frequency of the *BRCA1* 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. *Nature Genet.* **11**, 198–200 (1995).
24. Roa, B. B., Boyd, A. A., Volcik, K. & Richards, C. S. Ashkenazi Jewish population frequencies for common mutations in *BRCA1* and *BRCA2*. *Nature Genet.* **14**, 185–187 (1996).
25. Robson, M. *et al. BRCA*-associated breast cancer: absence of a characteristic immunophenotype. *Cancer Res.* **58**, 1839–1842 (1998).
26. Lee, J. S. *et al.* Survival after breast cancer in Ashkenazi Jewish *BRCA1* and *BRCA2* mutation carriers. *J. Natl Cancer Inst.* **91**, 259–263 (1999).
27. Ravid, A. *et al.* Immunohistochemical analyses of sporadic and familial (185delAG carriers) ovarian cancer in Israel. *Eur. J. Cancer* **36**, 1120–1124 (2000).
28. Yair, D. *et al.* p53 and WAF1 polymorphisms in Jewish-Israeli women with epithelial ovarian cancer and its association with *BRCA* mutations. *BJOG* **107**, 849–854 (2000).
29. Gryfe, R., Di Nicola, N., Gallinger, S. & Redston, M. Somatic instability of the *APC* I1307K allele in colorectal neoplasia. *Cancer Res.* **58**, 4040–4043 (1998).

30. Laken, S. J. *et al.* Familial colorectal cancer in Ashkenazim due to a hypermutable tract in *APC*. *Nature Genet.* **17**, 79–83 (1997).
31. Prior, T. W. *et al.* The I1307K polymorphism of the *APC* gene in colorectal cancer. *Gastroenterology* **116**, 58–63 (1999).
32. Rozen, P. *et al.* Prevalence of the I1307K *APC* gene variant in Israeli Jews of differing ethnic origin and risk for colorectal cancer. *Gastroenterology* **116**, 54–57 (1999).
33. Woodage, T. *et al.* The *APC* I1307K allele and cancer risk in a community-based study of Ashkenazi Jews. *Nature Genet.* **20**, 62–65 (1998).
34. Strul, H. *et al.* The I1307K adenomatous polyposis coli gene variant does not contribute in the assessment of the risk for colorectal cancer in Ashkenazi Jews. *Cancer Epidemiol. Biomarkers Prev.* **12**, 1012–1015 (2003).
35. Tobacco use among U. S. racial/ethnic minority groups: African Americans, American Indians and Alaska Natives, Asian Americans and Pacific Islanders, Hispanics. A Report of the Surgeon General. Executive summary. *MMWR Recomm. Rep.* **47**, 1–16 (1998).
36. Kensler, T. W., Qian, G. S., Chen, J. G. & Groopman, J. D. Translational strategies for cancer prevention in liver. *Nature Rev. Cancer* **3**, 321–329 (2003).
37. Ming, L. *et al.* Dominant role of hepatitis B virus and cofactor role of aflatoxin in hepatocarcinogenesis in Qidong, China. *Hepatology* **36**, 1214–1220 (2002).
38. Levy, L., Renard, C. A., Wei, Y. & Buendia, M. A. Genetic alterations and oncogenic pathways in hepatocellular carcinoma. *Ann. NY Acad. Sci.* **963**, 21–36 (2002).
39. Shapiro, G. I., Edwards, C. D. & Rollins, B. J. The physiology of $p16^{INK4A}$-mediated G1 proliferative arrest. *Cell Biochem. Biophys.* **33**, 189–197 (2000).
40. Liggett, W. H. Jr & Sidransky, D. Role of the *p16* tumor suppressor gene in cancer. *J. Clin. Oncol.* **16**, 1197–1206 (1998).
41. Armstrong, A. A. *et al.* Epstein–Barr virus and Hodgkin's disease: further evidence for the three disease hypothesis. *Leukemia* **12**, 1272–1276 (1998).
42. Munoz, N. *et al.* Risk factors for HPV DNA detection in middle-aged women. *Sex Transm. Dis.* **23**, 504–510 (1996).
43. Fukayama, M., Chong, J. M. & Uozaki, H. Pathology and molecular pathology of Epstein–Barr virus-associated gastric carcinoma. *Curr. Top. Microbiol. Immunol.* **258**, 91–102 (2001).
44. Young, L. S. & Murray, P. G. Epstein–Barr virus and oncogenesis: from latent genes to tumours. *Oncogene* **22**, 5108–5121 (2003).
45. Chong, J. M. *et al.* Global and non-random CpG-island methylation in gastric carcinoma associated with Epstein–Barr virus. *Cancer Sci.* **94**, 76–80 (2003).
46. Vo, Q. N. *et al.* Epstein–Barr virus in gastric adenocarcinomas: association with ethnicity and *CDKN2A* promoter methylation. *J. Clin. Pathol.* **55**, 669–675 (2002).
47. Mochizuki, S. *et al.* Homozygous deletion of the *p16*/*MTS-1*/*CDKN2* gene in malignant gliomas is infrequent among Japanese patients. *Int. J. Oncol.* **15**, 983–989 (1999).
48. Prives, C. & Hall, P. A. The p53 pathway. *J. Pathol.* **187**, 112–126 (1999).
49. Oren, M. *et al.* Regulation of p53: intricate loops and delicate balances. *Biochem. Pharmacol.* **64**, 865–871 (2002).
50. Olivier, M. *et al.* The IARC *TP53* database: new online mutation analysis and recommendations to users. *Hum. Mutat.* **19**, 607–614 (2002).
51. Chen, P. *et al.* Ethnicity delineates different genetic pathways in malignant glioma. *Cancer Res.* **61**, 3949–3954 (2001).
52. Burton, E. C. *et al.* Aberrant p53, mdm2, and proliferation differ in glioblastomas from long-term compared with typical survivors. *Clin. Cancer Res.* **8**, 180–187 (2002).
53. Hill, K. A. & Sommer, S. S. p53 as a mutagen test in breast cancer. *Environ. Mol. Mutagen.* **39**, 216–227 (2002).
54. Middleton, L. P., Chen, V., Perkins, G. H., Pinn, V. & Page, D. Histopathology of breast cancer among African-American women. *Cancer* **97**, 253–257 (2003).
55. Rose, D. P. & Royak-Schaler, R. Tumor biology and prognosis in black breast cancer patients: a review. *Cancer Detect. Prev.* **25**, 16–31 (2001).
56. Li, C. I., Malone, K. E. & Daling, J. R. Differences in breast cancer hormone receptor status and histology by race and ethnicity among women 50 years of age and older. *Cancer Epidemiol. Biomarkers Prev.* **11**, 601–607 (2002).
57. Paschke, T. Analysis of different versions of the IARC *p53* database with respect to G→T transversion mutation frequencies and mutation hotspots in lung cancer of smokers and non-smokers. *Mutagenesis* **15**, 457–458 (2000).
58. Pernick, N. L. *et al.* Clinicopathologic analysis of pancreatic adenocarcinoma in African Americans and Caucasians. *Pancreas* **26**, 28–32 (2003).
59. Hunt, J. D. *et al.* Differences in *KRAS* mutation spectrum in lung cancer cases between African Americans and Caucasians after occupational or environmental exposure to known carcinogens. *Cancer Epidemiol. Biomarkers Prev.* **11**, 1405–1412 (2002).
60. Piyathilake, C. J. *et al.* Race- and age-dependent alterations in global methylation of DNA in squamous cell carcinoma of the lung (United States). *Cancer Causes Control* **14**, 37–42 (2003).
61. Ashktorab, H. *et al.* High incidence of microsatellite instability in colorectal cancer from African Americans. *Clin. Cancer Res.* **9**, 1112–1117 (2003).
62. Carroll, W. L. Race and outcome in childhood acute lymphoblastic leukemia. *JAMA* **290**, 2061–2063 (2003).
63. Bhatia, S. *et al.* Racial and ethnic differences in survival of children with acute lymphoblastic leukemia. *Blood* **100**, 1957–1964 (2002).
64. Kadan-Lottick, N. S., Ness, K. K., Bhatia, S. & Gurney, J. G. Survival variability by race and ethnicity in childhood acute lymphoblastic leukemia. *JAMA* **290**, 2008–2014 (2003).
65. Pui, C. H. *et al.* Results of therapy for acute lymphoblastic leukemia in black and white children. *JAMA* **290**, 2001–2007 (2003).
66. Stephens, J. C. *et al.* Haplotype variation and linkage disequilibrium in 313 human genes. *Science* **293**, 489–493 (2001).
67. Lin, S. S. & Kelsey, J. L. Use of race and ethnicity in epidemiologic research: concepts, methodological issues, and suggestions for research. *Epidemiol. Rev.* **22**, 187–202 (2000).
68. Wilson, J. F. *et al.* Population genetic structure of variable drug response. *Nature Genet.* **29**, 265–269 (2001).
69. Desai, A. A., Innocenti, F. & Ratain, M. J. Pharmacogenomics: road to anticancer therapeutics nirvana? *Oncogene* **22**, 6621–6628 (2003).
70. Devlin, B., Roeder, K. & Wasserman, L. Genomic control, a new approach to genetic-based association studies. *Theor. Popul. Biol.* **60**, 155–166 (2001).
71. Kim, D. H. *et al.* $p16^{INK4a}$ and histology-specific methylation of CpG islands by exposure to tobacco smoke in non-small cell lung cancer. *Cancer Res.* **61**, 3419–3424 (2001).
72. Soria, J. C. *et al.* Aberrant promoter methylation of multiple genes in bronchial brush samples from former cigarette smokers. *Cancer Res.* **62**, 351–355 (2002).
73. Toyooka, S. *et al.* Smoke exposure, histologic type and geography-related differences in the methylation profiles of non-small cell lung cancer. *Int. J. Cancer* **103**, 153–160 (2003).
74. Yanagawa, N. *et al.* Frequent epigenetic silencing of the p16 gene in non-small cell lung cancers of tobacco smokers. *Jpn. J. Cancer Res.* **93**, 1107–1113 (2002).
75. Kersting, M. *et al.* Differential frequencies of $p16^{INK4a}$ promoter hypermethylation, *p53* mutation, and *K-ras* mutation in exfoliative material mark the development of lung cancer in symptomatic chronic smokers. *J. Clin. Oncol.* **18**, 3221–3229 (2000).
76. Hasegawa, M. *et al.* Patterns of gene promoter methylation in squamous cell cancer of the head and neck. *Oncogene* **21**, 4231–4236 (2002).
77. Gasco, M. *et al.* Epigenetic inactivation of 14-3-3 σ in oral carcinoma: association with $p16^{INK4a}$ silencing and human papillomavirus negativity. *Cancer Res.* **62**, 2072–2076 (2002).
78. Osawa, T. *et al.* Reduced expression and promoter methylation of *p16* gene in Epstein–Barr virus-associated gastric carcinoma. *Jpn. J. Cancer Res.* **93**, 1195–1200 (2002).
79. Tong, J. H. *et al.* Quantitative Epstein–Barr virus DNA analysis and detection of gene promoter hypermethylation in nasopharyngeal (NP) brushing samples from patients with NP carcinoma. *Clin. Cancer Res.* **8**, 2612–2619 (2002).
80. Yang, J. M. *et al.* Effect of HCV infection on expression of several cancer-associated gene products in HCC. *World J. Gastroenterol.* **5**, 25–27 (1999).
81. Shim, Y. H., Yoon, G. S., Choi, H. J., Chung, Y. H. & Yu, E. p16 hypermethylation in the early stage of hepatitis B virus-associated hepatocarcinogenesis. *Cancer Lett.* **190**, 213–219 (2003).
82. Platt, G., Carbone, A. & Mittnacht, S. p16INK4a loss and sensitivity in KSHV associated primary effusion lymphoma. *Oncogene* **21**, 1823–1831 (2002).
83. Fang, J. Y., Mikovits, J. A., Bagni, R., Petrow-Sadowski, C. L. & Ruscetti, F. W. Infection of lymphoid cells by integration-defective human immunodeficiency virus type 1 increases *de novo* methylation. *J. Virol.* **75**, 9753–9761 (2001).

Acknowledgements

J.K.W. is supported by the National Institutes of Environmental Health Sciences.

Competing interests statement

The author declares that he has no competing financial interests.

Online links

DATABASES

The following terms in this article are linked online to:

Cancer.gov: http://cancer.gov/

acute lymphoblastic leukaemia | brain cancer | breast cancer | colorectal cancer | gastric cancer | liver cancer | lung cancer | ovarian cancer | pancreatic cancer | prostate cancer

LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink/

APC | *BRCA1* | *BRCA2* | *CDKN2A* | *FY* | GATA1 | *KRAS* | Rb | *TP53* | β-catenin

FURTHER INFORMATION

International Agency for Research on Cancer *TP53* mutation database: http://www.iarc.fr/p53/

The Surveillance Epidemiology and End Results (SEER): http://seer.cancer.gov/

Access to this interactive links box is free online.