

## Introduction to the Post-Human Genome Project era, a target for interactions between polygenic and/or multiphenotypical components in cancer control in South America

Introducción a la post era del Proyecto Genoma Humano: la interrelación entre componentes multi-genéticos y multi-fenotípicos en el control del cáncer en América Latina como una meta

José Iscovich<sup>1,2</sup>

<sup>1</sup> International Institute of Fertility, 142/4 Achuzá Ra'anana, 43300, Israel.

<sup>2</sup> Division of Environmental and Occupational Health, Department of Life Sciences, Bar-Ilan University, Ramat Gan, Israel.

**Abstract** *Epidemiological studies have suggested that the propensity to develop malignancy involves a complex mix of genetic and environmental determinants, however both older and innovative techniques display unresolved questions regarding etiology. Current barriers to achieving the potential benefit from this understanding are: 1) incomplete background on the various environmental and genetic factors involved in the carcinogenesis mechanism; 2) difficulties in accurately differentiating specific molecular subtypes and measuring the effective cellular exposure dose; and 3) difficulties in determining the multifactorial interaction between genetic and environmental factors. To extrapolate Human Genome Project research findings to the Post-Human Genome Project era, South America provides a large population and large-pedigree families, thus including genetically heterogeneous and less heterogeneous groups. An initial strategy might be to trace high risk populations and the respective exposures to which they are susceptible, such as: 1) migration, identifying rural migrant populations; 2) inherent susceptibility, studying "long term homogeneous populations" or large families living in similar rural environments; and 3) dissection of gene-environmental interaction.*

**Key words** Neoplasms; Population Genetics; Human Genome

**Resumen** *Estudios epidemiológicos han demostrado que la susceptibilidad de la población a las enfermedades malignas está basada en interrelaciones genéticas hereditarias y no hereditarias. Las técnicas epidemiológicas tradicionales no han resuelto los problemas básicos de los mecanismos etiológicos. Las barreras existentes son: 1) el conocimiento incompleto de las etapas del mecanismo de la carcinogénesis accionada por factores genéticos y ambientales; 2) la dificultad en delimitar subtipos específicos de neoplasmas basados en mecanismos moleculares definidos, y las dosis efectivas de exposición celular; y 3) la capacidad en determinar la interrelación en el mecanismo genético-ambiental. Anticipándose al futuro, América Latina presenta la oportunidad para capitalizar los recientes avances en el conocimiento molecular, aplicando técnicas epidemiológicas y biológicas. La primera estrategia es la delimitación de poblaciones de alto riesgo o que vengán a representar alto riesgo en el futuro. Las estrategias deberían identificar: 1) poblaciones de migración reciente a áreas urbanas con antecedentes genéticos homogéneos; 2) susceptibilidad genética en familias numerosas residentes en zonas rurales con antecedentes genéticos homogéneos y menos homogéneos; y 3) estudios biológicos específicos de la interrelación genética ambiental en distintas situaciones multi-genéticas y multi-fenotípicas entre sus componentes.*

**Palabras clave** Neoplasias; Genética Poblacional; Genoma Humano

## One chapter, one challenge

The primary prevention of human cancer relies on the idea that reducing a population's exposure to a causal risk factor will result in decreased cancer incidence. Given that organisms are adversely affected by hazardous environments, susceptible individuals are by definition affected more severely or more quickly under hostile surroundings. Hence, primary prevention is reaching new areas in the interrelationship between polygenic (multiple genes) and multifactorial (genetic and environmental) events. Owing to the rapid development in the fields of molecular biology and molecular epidemiology since the late 1980s, scientists have expanded the capacity of epidemiological techniques to identify the biologically effective dose at tissue targets (for example, DNA), early biological effects (for example, DNA alterations), and variations in individual susceptibility. The goals are a chain of accurate steps from host-biological subtypes, determining exposure, and genetic-multifactorial interrelations (Table 1).

Available biologic evidence shows the strengths and limitations of these types of associations, particularly in susceptibility to chemically-induced cancers in the micro-environmental field (read specific occupational environments) and the extent to which genetic heterogeneity and other factors play a role in the process. A typical example from traditional epidemiological research on potential cancer-causing exposures with successful application in the occupational and para-occupational areas of preventive interventions is that of asbestos-related malignancies. However, not all associations are causal and not all exposure-cancer associations (with perhaps a thousand associations inferred) are biologically plausible.

More recent epidemiological studies suggest that the propensity to develop a malignancy involves a complex mix of genetic and environmental determinants and therefore shows patterns of inheritance that do not follow simple Mendelian transmission (Lander & Schork, 1994). However, both older and innovative techniques still face unresolved fundamental questions regarding etiology, susceptibility, and environmental carcinogenesis. As mentioned above, epidemiological studies have used measurement in hazardous environments, where individuals are affected more severely or more quickly under hostile surroundings. Two examples lead to different approaches to the dilemma of this gene-environmental interrelationship.

As a first example, case-control epidemiological studies of female breast cancer, describing an association between genetic deficiency in isoforms of the detoxification enzyme glutathione S-transferase (GST, mainly class *mu* in post-menopausal women) have conflicting results (Kelsey et al., 1997; Helzlsouer et al., 1998). Could differences in the environmental component of the putative gene-environmental interaction be responsible for dissimilar results such as geographic distribution, exposure, and/or ethnicity (Laden et al., 1997; Hartmann et al., 1997)? If so, these lines point to a diversity of mutational exposures in breast cancer, where environmental studies should be guided by a classification into genetically homogeneous populations to show sufficient evidence of biological plausibility.

The second example is from the study of the so-called "A-bomb survivors", focusing on the quantified leukemia relative risk of 75,991 Japanese survivors for whom radiation exposure levels were available (Shimizu et al., 1990). Two hundred and two survivors developed leukemia between 1950 and 1985, but the relative risk declined from 11.7 to 1.8 over time. The exposed Japanese population is genetically less heterogeneous than other populations in the cancer surveillance network, but even in this example, distinctive mutational patterns were able to differentiate the highly susceptible population with a short exposure-to-diagnosis interval. In terms of similar evidence of environmental-gene interaction, it is reasonable to cite the example of the differential risk of workers exposed to benzene in the development of leukemia and bone marrow aplasia (Aksoy, 1985; Linet et al., 1996). Far from being fully clarified, this genetically complex disease involving traditional exposure-related association (environmental or occupational) with cancer has produced novel approaches to define the role of genetic susceptibility in epidemiological studies of cancer etiology.

Individual genetic susceptibility is a major etiologic factor of cancer. While there has been impressive progress in understanding monogenic diseases prone to cancer risk, for example ataxia-teleangiectasia (Gilad et al., 1998), or which are associated with an inherent dominant gene such as bilateral retinoblastoma (Sipfel et al., 1998), such studies have had relatively little impact on prevention, implications in cancer control, and public health perspectives, due to the low incidence of most of these diseases. In contrast, as was pointed out, more important traits related to the propensity to develop neoplasms follow more complex poly-

Table 1

Goals in molecular epidemiology, exposure, and neoplasms.

The Host
1) Inheritable variations in carcinogenesis involving metabolic enzymes. 2) Determination of polygenic etiology in carcinogenesis mechanisms. 3) Germ-line mutations in tumor-associated gene(s). 4) Sib-pair design (concordant or discordant) to identify new environmental factor(s), controlling for genetic heterogeneity. 5) DNA repair mechanism.
The Environment
1) Measurement of exposure to mutagenic agents. 2) Intracellular effective exposure. 3) Identification, determination, and measurement of reliable biomarkers for a specific exposure. 4) DNA and protein adducts. 5) Dissection of multifactorial components (genetic and environmental).
The Neoplasm
1) Somatic mutations in disease progression. 2) Disturbance of architecture, invasion by malignant cells, nuclear anaplasia, multicentricity, and cell types. 3) Early biological effect. 4) Molecular subsets. 5) Association with other primary neoplasms.

genetic and multifactorial patterns (Weeks & Lathrop, 1995; Risch & Merikangas, 1996). Such complexity makes it difficult to characterize and identify each and every factor contributing to cancer etiology in a given individual or in the population at large.

The following are some of the current barriers to achieving the potential benefit from this understanding:

1) Incomplete understanding of the various environmental and genetic factors involved in the carcinogenesis mechanism and difficulties in their assessment.

2) Difficulties in accurately differentiating specific molecular mechanisms.

3) Limitations in measuring the effective cellular exposure dose and its interrelationship with the polygenetic etiology and the effective individual exposure.

4) Limited ability to determine the multiple interaction between genetic and environmental factors. This interaction gene/mutation specific creates "*the phenomenon of differential genetic susceptibility to certain environmental influences*" (Xu & Schork, 1997:521).

Another barrier to identifying environmental factor(s) for neoplasms is the paucity of appropriate epidemiological studies conducted in the field, for example, individual develop-

mental and differential phenotypical modes of transmission, genetic heterogeneity, incomplete genetic penetrance, and delayed age of onset. Of course, a potential effect of the use of molecular epidemiology is that it will increase the number of subsets in each traditionally defined neoplasm type, and as molecular assays become more sensitive, the probability of false-positive findings increases and diagnostic specificity is reduced.

By the years 2003-2005 it is expected that the 3 billion base-pairs human genome will have been completely explored and its 50,000-100,000 genes deciphered (Boguski, 1995; Schuler et al., 1996). As a result, it will become possible to identify mutant genes and polymorphic alleles associated with the development of most neoplasms. In addition, the ability to accurately measure molecular markers of exposure (for example, DNA protein adduct, see Table 1) will help in the analysis of the differential contribution of each environmental carcinogen to the development of malignancies (Perera, 1996, 1997). To date, the most common use of biomarkers in cancer research is in analyzing the complex interplay of various molecular determinants in the initiation of neoplasms.

Illustrating this situation of the conjunction of molecular epidemiology techniques and the

information to come from the Human Genome Project, some potential future issues will be the solution to small sample sizes in study populations, inappropriate selection of controls, subclassification of neoplasms, and difficulties in identifying multifactorial interactions between polygenic and/or multiphenotypical components. This situation is currently unsatisfactory using traditional epidemiological methods and will continue to pose problems if the Human Genome Project potential is not applied to develop an understanding of novel gene-environmental interaction in an area like Latin America with a broad socioeconomic transition.

These new aspects point to a diversity of challenges in cancer research and may be critical for future studies of genes related to environmental susceptibility to incorporate larger scale studies with more sophisticated methods of neoplasm subclassification into etiologically more homogeneous groups. The critical element, at least up to now, has been the potential public impact and potential benefits associated with interventions based on common genetic risks. A full understanding of the mechanism involved in gene-environmental mechanisms could lead to preventive measures, including the management of high-risk groups carrying inherent germ-line genes.

In the face of inconsistent population studies and elusive environmental risk factors related to the environmental-genetic complex which begins to emerge in specific cancer sites for example in ductal breast cancer, understanding the role of GSTs in the glutathione conjugation of estrogens and lipid metabolism in different germ-line mutation carrier populations (Zhu & Conney, 1998) one would ask what might be the appropriate critical approach to evaluating this threshold of evidence for accepting putative etiology associated with potential public health impact.

To extrapolate the Human Genome Project research findings to the Post-Human Genome Project era in order to the study cancer in human population, three important concepts could apply: a) the sample size must be large; b) large study should cover both, long-term isolated populations and those having migrated recently from remote to more urban and industrial areas; and c) and the markers must be informative (Ishibe & Kelsey, 1997).

An initial strategy could be to trace high-risk populations and the respective exposures to which they are susceptible. The following approach would facilitate this strategy:

a) Migration study by identifying rural migrant populations living in urban areas. By

identifying the rural areas from which the population came and comparing cancer burden between the urban (host) and rural (origin) areas, it is possible to estimate the attributable fraction of neoplasm associated with the urban environment without confounding from genetic factors. Rural migrants may be moving on a large scale from the same region, and they may be genetically homogeneous, thus facilitating gene-mapping studies (Khoury et al., 1998) from China. The following is an example from two parallel studies conducted in Los Angeles (Douer et al., 1996). Acute promyelocytic leukemia (or acute myelocytic leukemia FAB-M3) was found to be significantly higher in Hispanics than in non-Hispanics (proportions: 37.% and 6.5%, respectively,  $p < 0.00001$ ) in a hospital-based study on recent migrant groups. A similar trend, although to a lesser but still significant degree, was found in a population-based epidemiological study at age 30-69 years in the same area, for at least 15 years prior to the diagnosis of acute promyelocytic leukemia in the migrant Hispanic population as compared to the non-Hispanic population (proportion 24.3.% and 8.3%, respectively,  $p < 0.0075$ ). The difference in estimated risks of acute promyelocytic leukemia in the two studies, performed in the same study area, may reflect different environmental exposures, as the authors point out, due to the migration effect. The rate in the population-based study is much closer to the non-Hispanic population, but still significantly higher, probably as a result not only of differences in genetic susceptibility, but also of long-term residence in Los Angeles (at least 15 years), compared with the hospital series chapter of the study.

b) Inherent susceptibility by studying large families or "long-term homogeneous populations" living in similar rural environments and in the native population. In South America, there are still isolated populations that are ideal for researching linkage disequilibrium mapping (Jorge, 1995). Studying disequilibrium mapping in this population in large-pedigree families with high birth rates and a younger population could prove useful for valuable segregation for narrowing chromosomal regions. Searching large-pedigree families could facilitate the study, of discordant sib-pairs in gene mapping, useful to identify new environmental risk factors, controlling for genetic heterogeneity.

c) Dissection of gene-environmental interaction. Differential genetic susceptibility to specific environmental exposures is an essential part to elucidate the pathogenesis of neo-

Table 2

Cancer incidence in selected regional areas and cancer sites, 1988-1992<sup>1</sup>.

Cancer site (ICD-9)	Latin America			Canada & USA SEER-White			South Europe			Scandinavia			India			Australia		
	ASR	SIR	(95% CI)	ASR	SIR	(95% CI)	ASR	SIR	(95% CI)	ASR	SIR	(95% CI)	ASR	SIR	(95% CI)	ASR	SIR	(95% CI)
<b>Men</b>																		
Esophagus (150)	7.1	100	(94.7; 105.5)	4.2	87.4	(85.2; 89.6)	5.4	95.2	(92.1; 98.5)	3.6	76.9	(74.3; 79.5)	9.7	177.2	(170.4; 184.2)	4.8	99.2	(95.0; 103.5)
Stomach (151)	26.7	100	(97.2; 102.8)	9.3	55.2	(54.3; 56.1)	21.6	105.1	(103.3; 106.9)	12.2	78.5	(77.1; 79.9)	9.6	49.8	(47.9; 51.7)	10.9	63.0	(61.2; 64.8)
Colon (153)	10.8	100	(95.6; 104.5)	27.4	399.8	(395.9; 403.6)	18.5	221.2	(217.1; 225.2)	18.2	283.0	(278.9; 287.1)	2.8	37.4	(34.9; 40.1)	22.8	389.6	(382.8; 396.6)
Larynx (161)	5.7	100	(94.4; 105.9)	6.3	141.2	(138.2; 144.2)	12.6	255.1	(249.6; 260.7)	3.5	80.8	(78.0; 83.7)	6.3	131.3	(125.1; 137.7)	5.1	114.5	(109.7; 119.5)
Lung (162)	26.1	100	(97.3; 102.8)	63.7	341.5	(339.3; 343.7)	63.3	294.8	(292.0; 297.7)	38.0	210.8	(208.6; 213.0)	11.9	58.2	(56.2; 60.3)	46.7	249.5	(246.2; 253.0)
Breast (175)	0.3	100	(77.3; 127.3)	0.6	277.7	(260.3; 296.1)	0.6	210.0	(189.0; 232.7)	0.4	187.2	(169.2; 206.6)	0.4	140.5	(114.9; 170.1)	0.6	262.7	(232.3; 296.0)
Prostate (185)	36.4	100	(97.3; 102.8)	79.7	426.1	(423.8; 428.5)	24.1	97.2	(95.6; 98.8)	46.1	267.4	(265.1; 269.7)	5.7	23.7	(22.3; 25.2)	51.5	275.1	(271.7; 278.6)
Bladder (188)	9.0	100	(95.0; 105.2)	20.9	383.1	(378.9; 387.4)	24.4	369.9	(364.1; 375.8)	19.0	361.6	(356.5; 366.8)	3.7	57.3	(52.1; 62.9)	15.3	277.2	(270.7; 283.9)
NHL (200+202)	7.4	100	(95.4; 104.8)	14.4	255.6	(252.1; 259.1)	8.7	137.0	(133.2; 140.8)	10.4	196.5	(192.5; 200.6)	3.8	56.8	(53.7; 60.0)	13.2	233.7	(227.6; 239.9)
All sites, but skin	204.6	100	(99.0; 101.0)	342.3	253.2	(252.5; 253.9)	283.7	177.6	(176.8; 178.4)	252.6	197.5	(196.7; 198.2)	115.0	74.9	(74.1; 75.8)	315.2	228.3	(227.1; 229.5)
<b>Women</b>																		
Esophagus (150)	1.1	100	(90.7; 110.0)	1.3	109.9	(10.5; 114.3)	0.8	44.6	(41.0; 48.5)	1.2	116.1	(110.7; 121.7)	7.5	546.1	(521.3; 571.9)	2.2	194.9	(184.4; 205.8)
Stomach (151)	13.3	100	(96.4; 103.7)	3.9	55.6	(54.4; 56.8)	9.9	98.2	(96.0; 100.5)	6.4	96.6	(94.6; 98.6)	4.6	50.3	(47.5; 53.2)	7.6	64.3	(61.9; 66.8)
Colon (153)	10.0	100	(95.8; 104.4)	21.1	397.9	(394.2; 401.7)	13.7	181.2	(177.8; 184.7)	16.2	318.6	(314.4; 329.9)	2.2	31.3	(28.7; 34.1)	21.9	381.6	(374.8; 388.5)
Larynx (161)	0.8	100	(86.2; 115.4)	1.3	210.8	(201.6; 220.4)	0.6	91.9	(83.7; 100.6)	0.5	86.3	(78.9; 94.2)	0.9	141.9	(123.9; 161.8)	0.6	99.5	(87.9; 112.1)
Lung (162)	7.3	100	(95.3; 104.9)	30.5	603.5	(598.3; 608.7)	8.4	141.0	(137.6; 144.5)	13.7	269.5	(265.2; 273.9)	2.7	48.9	(45.2; 52.7)	15.6	319.5	(312.5; 326.6)
Breast (174)	44.4	100	(98.2; 101.8)	82.6	238.8	(237.6; 240.1)	56.2	137.5	(136.2; 138.9)	68.1	197.1	(195.6; 198.6)	24.3	63.4	(61.9; 64.9)	67.4	187.5	(185.4; 189.6)
Cervix Uteri (180)	24.5	100	(97.7; 102.3)	7.7	37.0	(36.3; 37.6)	9.5	41.5	(40.5; 42.6)	9.2	44.3	(43.3; 45.2)	25.9	106.9	(104.5; 109.4)	9.9	46.9	(45.5; 48.3)
Ovary (183)	6.4	100	(98.4; 104.7)	10.7	215.3	(212.1; 218.5)	7.6	132.9	(129.4; 136.5)	7.6	155.8	(152.3; 159.3)	5.7	99.7	(95.0; 104.6)	8.6	170.5	(165.3; 175.9)
Bladder (188)	2.5	100	(91.5; 109.1)	5.5	399.6	(392.1; 407.2)	3.9	208.7	(201.3; 216.3)	4.8	362.2	(353.3; 371.2)	1.0	54.8	(47.8; 62.5)	4.3	312.6	(300.4; 325.3)
NHL (200+202)	5.3	100	(94.8; 105.4)	9.5	264.0	(260.0; 268.1)	6.0	130.7	(126.8; 134.6)	7.1	207.4	(202.9; 212.0)	2.3	54.2	(50.2; 58.4)	8.8	239.4	(233.4; 246.6)
All sites except skin	181.3	100	(99.1; 100.9)	264.5	209.3	(208.6; 209.8)	191.4	123.7	(123.1; 124.4)	227.6	187.6	(186.9; 188.3)	119.5	79.0	(78.2; 79.9)	242.1	187.8	(186.7; 189.0)

<sup>1</sup> Age-adjusted incidence rates per 100,000 and standardized incidence ratios (SIR) were computed by Parkin et al. (1997). ICD-9 is International Classification of Diseases, Ninth Edition; SIR is age-standardized, ASR is age-adjusted standard rate to the world population and CI confidence interval. Reference category: for each neoplasm and cancer registry, to similar neoplasm in all of South America. Cancer registration in South America includes: Concordia, 1990-94 (Argentina); Belém 1989-91, Goiânia 1990-93, Porto Alegre 1990-92 (Brazil); Cali 1987-91 (Colombia); Costa Rica 1988-92; Quito 1988-92 (Ecuador); Lima 1990-91, Trujillo 1988-90 (Peru); Puerto Rico 1981-91; Montevideo 1990-92 (Uruguay). United State of America includes incidence in white population from the Surveillance, Epidemiology, and End Results (SEER) Program. Southern Europe includes Croatia (1988-92); Hérault, Isère, Tarn 1988-92 (France); Italy; Malta 1992-93; Granada, Mallorca, Murcia, Tarragona 1988-92 (Spain). Scandinavia includes Finland 1987-92, Denmark, Iceland, Norway and Sweden 1988-92.

Table 3

Cancer incidence by selected sites and cancer registration areas, 1987-1994<sup>1</sup>.

Cancer Registry	Esophagus		Colon		Lung		Breast		Prostate		Non-Hodgkin Lymphoma	
	SIR	95% CI	SIR	95% CI	SIR	95% CI	SIR	95% CI	SIR	95% CI	SIR	95% CI
<b>Men</b>												
Concordia, 1990-94	340.3	251.7; 449.8	232.2	171.8; 307.0	285.6	243.0; 333.7	141.3	4.2; 787.3	86.2	63.6; 114.3	60.2	30.0; 107.6
Belém, 1989-91	118.7	89.8; 153.8	55.5	39.7; 75.6	134.0	117.8; 151.8	288.1	115.6; 593.4	72.8	60.5; 87.1	67.4	50.5; 88.2
Goiânia, 1990-93	158.2	125.2; 197.2	60.4	44.1; 80.9	87.2	74.5; 101.5	78.2	9.4; 282.5	101.3	86.7; 117.6	73.8	56.1; 95.5
Porto Alegre, 1990-92	327.8	287.8; 371.9	167.1	143.6; 193.3	303.3	283.4; 324.2	136.5	44.2; 318.5	147.2	133.1; 162.4	121.4	100.5; 145.4
Cali, 1987-91	71.3	56.7; 88.5	84.2	71.0; 99.1	116.6	106.7; 127.1	86.3	27.9; 201.3	154.8	143.1; 167.3	114.0	98.3; 131.4
Costa Rica, 1988-92	55.3	46.2; 65.8	66.1	57.8; 75.2	64.2	59.1; 69.7	77.9	35.7; 147.9	92.2	85.9; 98.9	87.5	77.9; 97.9
Quito, 1988-92	61.6	45.7; 81.3	51.3	39.4; 65.6	60.8	44.6; 80.8	121.9	39.5; 284.5	125.9	113.6; 139.2	134.5	114.9; 156.3
Lima, 1990-91	35.0	27.4; 44.2	72.0	62.8; 82.2	79.9	73.8; 86.4	78.0	33.6; 153.6	98.1	91.2; 105.4	110.2	98.5; 123.0
Trujillo, 1988-90	69.6	38.0; 116.8	66.9	40.8; 103.2	60.8	44.6; 80.8	98.0	2.9; 545.9	118.0	94.6; 145.4	135.9	97.1; 185.1
Puerto Rico, 1988-91	160.1	147.4; 173.6	180.0	168.9; 191.6	87.6	82.8; 92.6	111.9	68.4; 172.8	222.6	215.3; 230.1	122.4	112.0; 133.5
Montevideo, 1990-92	245.0	218.5; 273.9	344.7	318.4; 372.5	360.7	344.1; 377.9	314.8	189.5; 491.5	175.5	164.4; 187.1	177.4	155.6; 201.4
<b>Women</b>												
											<b>Cervix uteri</b>	
Concordia, 1990-94	260.0	134.4; 454.2	248.3	187.5; 322.7	149.3	98.4; 217.2	166.9	144.9; 191.3	150.7	123.6; 182.0	77.5	38.7; 138.6
Belém, 1989-91	114.2	68.7; 178.3	57.0	42.2; 75.4	127.3	101.8; 157.3	74.2	66.9; 82.0	275.8	258.2; 294.4	49.0	33.0; 69.9
Goiânia, 1990-93	128.0	78.2; 197.7	102.5	81.7; 126.9	156.3	127.3; 190.0	97.3	89.0; 106.2	138.1	125.8; 151.4	73.6	53.3; 99.2
Porto Alegre, 1990-92	258.2	203.1; 323.8	175.7	154.7; 198.8	234.9	207.7; 264.7	144.7	136.5; 153.2	97.1	88.5; 106.3	135.9	112.8; 162.4
Cali, 1987-91	169.7	130.1; 217.6	93.8	80.5; 108.8	166.1	146.0; 188.3	95.5	89.9; 101.4	141.5	132.9; 150.6	127.9	109.1; 149.1
Costa Rica, 1988-92	90.1	68.6; 116.2	78.1	68.9; 88.2	86.7	75.9; 98.7	71.3	67.6; 75.1	104.6	99.0; 110.5	79.5	68.4; 91.8
Quito, 1988-92	97.8	63.9; 143.3	87.3	72.3; 104.5	73.3	57.9; 91.5	71.4	65.7; 77.4	133.9	124.1; 144.3	146.3	123.0; 172.8
Lima, 1990-91	44.9	29.3; 65.9	80.2	70.3; 91.1	120.2	106.6; 135.3	85.8	81.5; 90.2	115.6	109.4; 122.1	136.3	120.6; 153.4
Trujillo, 1988-90	46.5	9.6; 136.0	70.0	44.4; 105.1	85.2	53.4; 129.0	79.2	67.2; 92.7	232.7	206.4; 261.5	119.2	78.6; 173.5
Puerto Rico, 1988-91	137.1	116.5; 160.2	162.1	151.7; 173.0	116.2	106.4; 126.7	114.3	110.5; 118.2	43.0	40.0; 46.2	110.2	99.3; 122.1
Montevideo, 1990-92	310.9	263.4; 364.5	344.8	321.2; 369.6	149.0	131.9; 167.7	273.7	264.1; 283.4	86.3	79.1; 94.0	200.1	176.5; 225.9

<sup>1</sup> Reference category: for each neoplasm and cancer registry, to similar neoplasm in all South America (reference = 100) as in Parkin et al., 1997.

plasms. By identifying large non-hereditary neoplasms in high-risk populations and new disease-susceptibility genes (after the Human Genome Project), it will facilitate identification of the precise pathogenesis mechanisms. One barrier in traditional epidemiological research is the apparent lack of success in identifying the environmental risk factor(s) which can contribute to the paucity of controlling environmental exposure misclassification.

The potential advantage in these studies using gene-environmental dissection is that occupational and environmental exposures to risk factors are determined within the disease group and compared between those who carry a specific gene-susceptibility and those who do not (Dorman et al., 1988). Comparison analysis of the genotypical determinant of the disease and its associated absolute risk across the population (migrants, rural/urban, industrialized/primary economy) may provide clues as to specific mechanisms of pathogenesis and etiology, in addition to new therapeutic and specific preventive interventions. Few epidemiological studies have been conducted that analyze the interplay between environmental factors and DNA or cytogenetic changes; most of them were retrospective and focused on this association, without adjusting for genetic factors (see Table 1). An example was an epidemiological study on women working in agriculture, the textile industry, and housewives using a mailing survey through the Swedish Environmental Cancer registry, in which emphasis was given to ascertaining past exposure, showing increased risk in specific hematopoietic neoplasms (i.e., adjusted standardized incidence ratio of 3.6 for multiple myeloma in textile workers) (Linnet et al., 1994). Moreover, as the author mentioned, the study lacked detailed information on specific exposure types and duration of exposure, and occupational data were compared with compiled groups of neoplasms rather than with dissected gene-environmental components. Conduction of comprehensive data-bases that are informative (including incidence and an occupational-environmental matrix), representative (population-based, large-pedigree families, and isolated populations), and with reliable biological samples in South America should allow for testing innovative hypotheses in the post-Human Genome Project era.

Traditionally, an ecological study examines the relationship between the disease and the direct event, such as environmental exposure across various populations. In addition, the use of large numbers of cells across regions will

somehow ensure random cross-regional relation between the disease, the exposure, and their covariates (Cohen, 1990). Furthermore, it is sometimes assumed that for spatial correlation to be compared, region itself must be a confounder on the individual level after other factors are controlled. There are already a number of descriptive epidemiological reviews that have accurately illustrated cross-sectional estimates and trends in cancer incidence and mortality by region (Coleman et al., 1993; Parkin et al., 1997; Pisani et al., 1997), and it is not intended to replicate them here. Instead, focus will be on recent data and the potential of increasing informative sources for interaction between polygenic and/or multiphenotypic components. It is beyond the purpose of these comments to fully explore cancer registration in South America. Data presented here, comprise more or less 5-6% of the population living in the region, which apart from statistical variation, remain low and randomly representative of demographic and environmental components of the continent.

Given this limitation, Table 2 shows cancer burden from recently published incidence data by population-based registries across regions compared to South America as a population reference. There is a substantial variation in the incidence of these selected cancer sites. Briefly, all neoplasms in both sexes are lower compared with others regions. Moreover, the esophagus, stomach, and uterine cervix showed higher incidence rates. In a more detailed analysis based on the same cancer registration network (Table 3), a heterogeneity of incidence rates could shed light on the selected cancer sites. It is likely that the information processes among the various registries and diagnostic criteria tend to explain the heterogeneity of incidence rates, yet clear patterns emerge from the network information. Prostate and uterine cervical cancer varied inversely across South America. Incidence of esophagus, colon, and breast cancers in both gender differs significantly between the highest southern and lowest northern areas. These results from the two areas indicate clear patterns which may suggest different environmental factors and/or demographic components.

Returning to the enigma in the title, we still have to live with it in times to come. The philosophical drama represented in Greek theater is still upon our shoulders. However, in taking one step forward in a long march, the ties among experts in different fields must further enhance "the enigma in the title".

## References

- AKSOY, M., 1985. Malignancies due to occupational exposure to benzene. *American Journal of Industrial Medicine*, 7:395-402.
- BOGUSKI, M. S., 1995. Hunting for genes in computer data bases. *New England Journal of Medicine*, 333:645-647.
- COHEN, B. L., 1990. Ecological versus case-control studies for testing a linear-no threshold dose-response relationships. *International Journal of Epidemiology*, 19:680-684.
- COLEMAN, M. P.; ESTEVE, J.; DAMIECKI, P.; ARSLAN, A. & RENARD, H., 1993. *Trends in Cancer Incidence and Mortality*. IARC Scientific Publications 121. Lyon: International Agency for Research on Cancer.
- DORMAN, J. S.; TRUCCO, M.; LAPORTE, R. E. & KULLER, L. H., 1988. Family studies: the key to understanding the genetic and environmental etiology of chronic disease? *Genetic Epidemiology*, 5:305-310.
- DOUER, D.; PRESTON-MARTIN, S.; CHANG, E.; NICHOLS, P. W.; WATKINS, K. J. & LEVINE, A. M., 1996. High frequency of acute promyelocytic leukemia among Latinos with acute myeloid leukemia. *Blood*, 87:308-313.
- GILAD, S.; CHESSA, L.; KHOSRAVI, R.; RUSSELL, P.; GALANTY, Y.; PIANE, M.; GATTI, R. A.; JORGENSEN, T. J.; SHILOH, Y. & BAR-SHIRA, A., 1998. Genotype-phenotype relationships in ataxia-telangiectasia and variants. *American Journal of Human Genetics*, 62:551-561.
- HARTMANN, A.; BLASZYK, H.; KOVACH, J. S. & SOMMER, S. S., 1997. The molecular epidemiology of p53 gene mutations in human breast cancer. *Trends in Genetics*, 13:27-33.
- HELZLSOUER, K. J.; SELMIN, O.; HUANG, H. Y.; STRICKLAND, P. T.; HOFFMAN, S.; ALBERG, A. J.; WATSON, M.; COMSTOCK, G. W. & BELL, D., 1998. Association between glutathione S-transferase M1, P1, and T1 genetic polymorphisms and development of breast cancer. *Journal of the National Cancer Institute*, 90:512-518.
- ISHIBE, N. & KELSEY, K. T., 1997. Genetic susceptibility to environmental and occupational cancer. *Cancer Causes and Control*, 8:505-513.
- JORGE, L. B., 1995. Linkage disequilibrium as a gene-mapping tool. *American Journal of Human Genetics*, 56:11-14.
- KELSEY, K. T.; HANKINSON, S. E.; COLDITZ, G. A.; SPRINGER, K.; GARCIA-CLOSAS, M.; SPIEGELMAN, D.; MANSON, J. E.; GARLAND, M.; STAMPFER, M. J.; WILLET, W. C.; SPEIZER, F. E. & HUNTER, D. J., 1997. Glutathione S-transferase class mu deletion polymorphism and breast cancer: results from prevalent versus incident cases. *Cancer Epidemiology, Biomarkers and Prevention*, 6:511-515.
- KHOURY, M. J.; ADAMS, M. J. & FLANDERS, W. D., 1988. An epidemiological approach to ecogenetics. *American Journal of Human Genetics*, 42:89-95.
- LADEN, F.; SPIEGELMAN, D.; NEAS, L. M.; COLDITZ, G. A.; HANKINSON, S. E.; MANSON, J. E.; BYRNE, C.; ROSNER, B. A.; SPEIZER, F. E. & HUNTER, D. J., 1997. Geographic variation in breast cancer incidence rates in a cohort of US women. *Journal of the National Cancer Institute*, 89:1373-1378.
- LANDER, E. S. & SCHORK, N. J., 1994. Genetic dissection of complex traits. *Science*, 265:2037-2048.
- LINET, M. S.; MCLAUGHLIN, J. K.; MALKER, H. S.; CHOW, W. H.; WEINER, J. A.; STONE, B. J.; ERICSSON, J. L. & FRAUMENI Jr., J. F., 1994. Occupation and hematopoietic and lymphoproliferative malignancies among women: a linked registry study. *Journal of Occupational Medicine*, 36:1187-1198.
- LINET, M. S.; YIN, S. N.; TRAVIS, L. B.; LI, C. Y.; ZHANG, Z. N.; LI, D. G.; ROTHMAN, N.; LI, G. L.; CHOW, W. H.; DONALDSON, J.; DOSEMECI, M.; WACHOLDER, S.; BLOT, W. J. & HAYES, R. B., 1996. Clinical features of hematopoietic malignancies and related disorders among benzene-exposed workers in China. Benzene Study Group. *Environmental Health Perspectives*, 104(Sup. 6): 1353-1364.
- PARKIN, D. M.; WHELAN, S. L.; FERLAY, J.; RAYMOND, L. & YOUNG, J., 1997. *Cancer Incidence in Five Continents*. Vol. VII. IARC Scientific Publications, 143. Lyon: International Agency for Research on Cancer.
- PERERA, F. P., 1996. Molecular epidemiology: insights into cancer susceptibility, risk assessment, and prevention. *Journal of the National Cancer Institute*, 88:496-509.
- PERERA, F. P., 1997. Environment and cancer: who are susceptible? *Science*, 278:1068-1073.
- PISANI, P.; PARKIN, M. D.; MUÑOZ, N. & FERLAY, J., 1997. Cancer and infection: estimates of the attributable fraction in 1990. *Cancer Epidemiology, Biomarkers and Prevention*, 6:387-400.
- RISCH, N. & MERIKANGAS, K., 1996. The future of genetic studies of complex human diseases. *Science*, 273:1516-1517.
- SCHULER, G. D.; BOGUSKI, M. S.; STEWART, E. A.; STEIN, L. D.; GYAPAY, G.; RICE, K.; WHITE, R. E.; RODRIGUEZ-TOME, P.; AGGARWAL, A.; BAJOREK, E.; BENTOLILA, S.; BIRREN, B. B.; BUTLER, A.; CASTLE, A. B.; CHIANNILKULCHAI, N.; CHU, A.; CLEE, C.; COWLES, S.; DAY, P. J. R.; DIBLING, T.; DROUOT, N.; DUNHAM, I.; DUPRAT, S.; EAST, C.; EDWARDS, C.; FAN, J.-B.; FANG, N.; FIZAMES, C.; GARRETT, C.; GREEN, L.; HADLEY, D.; HARRIS, M.; HARRISON, P.; BRADY, S.; HICKS, A.; HOLLOWAY, E.; HUI, L.; HUSSAIN, S.; LOUISDIT-SULLY, C.; MA, J.; MACGILVERY, A.; MADER, C.; MARATUKULAM, A.; MATISE, T. C.; MCKUSICK, K. B.; MORISSETTE, J.; MUNGALL, A.; MUSELET, D.; NUSBAUM, H. C.; PAGE, D. C.; PECK, A.; PERKINS, S.; PIERCY, M.; QIN, F.; QUACKENBUSH, J.; RANBY, S.; REIF, T.; ROZEN, S.; SANDERS, C.; SHE, X.; SILVA, J.; SLONIM, D. K.; SODERLUND, C.; SUN, W.-L.; TABAR, P.; THANGARAJAH, T.; VEGA-CZARNY, N.; VOLLRATH, D.; VOYTICKY, S.; WILMER, T.; WU, X.; ADAMS, M. D.; AUFRAY, C.; WALTER, N. A. R.; BRANDON, R.; DEHEJIA, A.; GOODFELLOW, P. N.; HOULGATTE, R.; HUDSON Jr., J. R.; IDE, S. E.; IORIO, K. R.; LEE, W. Y.; SEKI, N.; NAGASE, T.; ISHIKAWA, K.; NOMURA, N.; PHILLIPS, C.;



- POLYMEROPOULOS, M. H.; SANDUSKY, M.; SCHMITT, K.; BERRY, R.; SWANSON, K.; TORRES, R.; VENTER, J. C.; SIKELA, J. M.; BECKMANN, J. S.; WEISSENBACH, J.; MYERS, R. M.; COX, D. R.; JAMES, M. R.; BENTLEY, D.; DELOUKAS, P.; LANDER, E. S. & HUDSON, T. J., 1996. A gene map of the human genome. *Science*, 274:540-546.
- SHIMIZU, Y.; KATO, H. & SCHULL, W. J., 1990. Studies of the mortality of A-bomb survivors. 9. Mortality, 1950-1985: Part 2. Cancer mortality based on the recently revised doses (DS86). *Radiation Research*, 121:120-141.
- SIPPEL, K. C.; FRAIOLI, R. E.; SMITH, G. D.; SCHALKOFF, M. E.; SUTHERLAND, J.; GALLIE, B. L. & DRYJA, T. P., 1998. Frequency of somatic and germ-line mosaicism in retinoblastoma: implications for genetic counseling. *American Journal of Human Genetics*, 62:610-619.
- WEEKS, D. E. & LATHROP, G. M., 1995. Polygenic diseases: methods for mapping complex diseases traits. *Trends in Genetics*, 11:513-519.
- XU, X. & SCHORK, N. J., 1997. Linking genes and environmental exposure: why China presents special opportunities. *Cancer Causes and Control*, 8:518-523.
- ZHU, B. T. & CONNEY, A. H., 1998. Functional role of estrogen metabolism in target cells: review and perspectives. *Carcinogenesis*, 19:1-27.