The Impact of the Human Genome Project on Risk Assessment

Katarzyna Doerffer , Paul Unrau

Atomic Energy of Canada, Radiation Biology and Health Physics, Chalk River, Ontario, Canada K0J 1J0

INTRODUCTION

The radiation protection approach to risk assessment assumes that cancer induction following radiation exposure is purely random. Present risk assessment methods derive risk from cancer incidence frequencies in exposed populations and associate disease outcomes totally with the level of exposure to ionizing radiation. Exposure defines a risk factor that affects the probability of the disease outcome. But cancer risk can be affected by other risk factors such as underlying genetic factors (predisposition) of the exposed organism. These genetic risk factors are now becoming available for incorporation into ionizing radiation risk assessment. Progress in the Human Genome Project (HGP) will lead to direct assays to measure the effects of genetic risk determinants in disease outcomes. When all genetic risk determinants are known and incorporated into risk assessment it will be possible to reevaluate the role of ionizing radiation in the causation of cancer.

The distribution of genetic risk determinants in time and space is governed by the biological and social processes involved in reproduction, such as the extent of inbreeding (endogamy) . We know already that endogamy causes the genetic risk determinants to be non-randomly distributed in populations [1]. These non-random distributions of genetic risk factors need to be considered in explaining cancer "hot-spots" observed by epidemiology.

Genetic determinants of cancer risk can be defined because the relevant phenotypes and their genotypes are being mapped and sequenced. Of 6456 genetic disorders identified in Online Mendelian Inheritance in Man (OMIM), 641 are associated directly, indirectly, or by implication with cancer. These cancer risk genes, when fully described, will define the list of genetic cancer determinants for use in risk assessment. Examples which impinge upon present risk assessments include cancer prone syndromes associated with ionizing radiation sensitivity.

ANALYSIS OF GENETIC DISORDERS ASSOCIATED WITH CANCER

Rapid developments in the HGP is leading to detailed knowledge about genes, providing information about genetic risk determinants. This information is stored in molecular biological databases which mirror the current knowledge of genetics and molecular biology. Information about those genes associated with cancer seems to be most relevant to risk assessment, as cancer risk is of most concern.

OMIM is the on-line version of V.A. McKusick's Mendelian Inheritance In Man [2]. OMIM describes all known genetic disorders and traits. The access interface uses Information Retrieval Experiment (IRX) software [3]. The Genome Data Base (GDB) is directly linked to OMIM. GDB collects genetic and physical chromosome maps, information about genes arranged by chromosome location, polymorphisms (normal and mutant gene variants), probes and associated information. GDB also provides accession numbers for links to nucleotide sequence repositories such as GenBank.

Because these databases can be interconnected, it is possible to link the phenotypic description of a particular disorder with molecular mutation at the DNA level and the detailed protein changes causing the disorder. This

allows us to define in molecular terms the genetic risk determinants we need to consider in ionizing radiation risk assessment. We built a set of computer tools to access these databases, analyze the data, and use it to evaluate hereditary contributions to radiation risk. Our goal was to define a set of all known phenotypes (physical manifestations of genotypes) and genotypes (underlying genetic blueprints) and all described biochemical events associated with cancer. Such a list is expected to contain the bulk of presently known or suspected genetic risk determinants; it is from this list that the targets of molecular epidemiology will be selected for analysis.

We operationally distinguished three major sorts of genetic effects as predisposing to cancer:
.1 direct associations,
.2 indirect associations and,
.3 molecular biological traits (biochemical events) associated with cancer.

Direct associations are those genetic disorders, in which a specific gene is involved, which when mutated manifests as cancer of a particular type. The primary outcome of mutation in such a gene is a cancer itself. Examples include Retinoblastoma (MIM# 180200) and Thyroid Carcinoma (MIM# 188550) which is also associated with ionizing radiation. This group includes phenotypes in which tumor suppressor genes and proto-oncogenes are involved.

Indirect associations include genetic disorders in which, for example, DNA repair is less accurate, leading to somatic mutation manifested as an increased incidence of cancer. Examples include Bloom syndrome (MIM# 210900), Neurofibromatosis (MIM# 162200), and Ataxia Telangiectasia (MIM# 208900) which is also associated with ionizing radiation sensitivity. This group includes phenotypes in which modulator genes are involved.

Molecular biological traits are a group of OMIM entries that describe all known molecular biological events leading to or associated with cancers. Examples include the Tumor Suppressor Gene, Hela Cell Type (MIM# 191181), Oncogene MYC (MIM# 164840), Transformation Gene: Oncogene AMV (MIM# 189990), Carcinoembryonic Antigen (MIM# 114890), and the DNA Damage Inducible Gene (MIM# 126335) which is induced by ionizing radiation.

We separated the 641 OMIM entries associated with cancer into direct, indirect, and trait subsets. The results are summarized in Table 1.

Table 1

type of association with cancer	# of disorders	# of known gene locations	# of detailed gene maps	# of sequenced genes
direct	94	53	27	15
indirect	107	52	24	20
traits	440	390	166	218

In Table 1 the number of disorders of any category were defined by reading the entries, checking for linkage to GDB, and determining the details of mapping and sequence data. The entries reflect progress in the HGP; mapping at increasing resolution is followed by sequencing the actual gene. Increasing numbers of genetic disorders associated with cancer are becoming known, and genes and map locations are being determined with increasing precision. The sequence information is what is ultimately required to develop molecular assays for the presence or absence of variants of this class of genes in individuals. All these disorders are in some way connected with cancer risk; the precise nature of the connection is one of the major outputs of the HGP.

Ultimately, all genetic disorders will be tied to mapped, sequenced genes. At this interim period, however, the 253 sequenced genes are primary targets of molecular epidemiology . Thus, we are primarily interested in those

mapped and sequenced genes whose phenotypes include sensitivity to ionizing radiation.

In risk assessment, the value of having a list of genes affecting cancer risk lies in the ability to combine molecular biological assays and epidemiological methods to define cancer risk sources in individuals and thus populations. Knowing the genes and their sequences gives us the potential for direct measurement of genetic risk determinants. By measuring doses and determining genetic risk factors, a better understanding of the ionizing radiation component of cancer risk will be obtained. Thus the presence or absence of genetic risk factors in individuals will affect the risk assessment process. Ultimately it will be possible to determine risk on the individual level and tie the risk to real causes.

Consider an example of a genetically determined, ionizing radiation sensitive disorder which is expected to affect risk assessment. Ataxia telangiectasia has long been a model human genetic disease for studies in radiation research, and ultimately in risk assessment. It belongs to the indirect set of cancer associated genes. AT is described in detail in a number of recent publications [4,5,6,7,8]. Cancer predisposition in heterozygotes, who carry one mutant copy of the AT gene, is estimated to be about 3 - 4 fold higher than the general population, and breast cancer risk in carrier women may be five-fold higher than non-carriers. It has also been hypothesized that breast cancer risks in irradiated carriers may be further increased [9].

While AT is a rare disease, up to 1.5% of the world population may be carriers of one defective copy of the gene. The relevance of AT to risk assessment is that heterozygotes, or carriers, may show both cancer predisposition and moderate radiation sensitivity. This means that up to 1.5% of individuals may be significantly more radiation sensitive than other members of the population, and at altered cancer risk, owing to carrying a mutant copy of the AT gene. Non-carriers will not have the risk determinant nor the attributable risk.

Remember that AT is only one of the 641 genetic risk determinants found in our search, and can account for only a proportion of genetically mediated or influenced cancer risk. Assays to measure the actual genetic risk determinants at the DNA sequence level are required to define the presence or absence of genetic risk determinants in individuals. Such direct assays for the presence or absence of gene variants linked with cancer predisposition, susceptibility, or expression will lead to a new understanding of the magnitude of ionizing radiation risk.

The distribution of genetic risk determinants in the individuals who make up populations is governed by biological processes. We have discussed two indicators of endogamy which are associated with non-random distribution of genetic risk determinants, namely ethnicity and consanguinity [1]. The genetic consequence of endogamy is that genetic risk factors located on DNA shared between group members is more likely to lead to observed genetic disorders. The AT variants described to date show that both ethnicity and consanguinity affect their temporal and spatial distribution, and their associated cancer risks in populations.

Acquiring the ability to assay genes directly, to determine both their relationship to disease and their distribution in populations will increase the certainty of the risk assessment process by identification and direct measurement of its deterministic genetic risk factors. At the same time it will reduce the uncertainty of risk assessment by decreasing the probabilistic components of the process.

REFERENCES

1. P. Unrau and K. Doerffer. *Perspectives in Biology and Medicine.* (Fall 1995.)
2. V.A. McKusick. *Johns Hopkins University Press,* Baltimore (1992).
3. GDB, OMIM User Guide. *Baltimore: Johns Hopkins University* (1993).
4. K. Savitsky et al. *Science,* 268, 1749-1753 (1995.)
5. V.A. Zakian. *Cell,* 82, (1995).
6. K.L. Hari, A. Santerre, et al. *Cell,* 82, (1995).
7. P.W. Greenwell, S.L. Kronmal, et al. *T Cell,* 82 (1995).
8. D.M. Morrow, D.A. Tagle, et al. *Cell,* 82, (1995).
9. M. Swift, et al. *New Eng. J. Med.* 325, 1831-1836 (1991).