

Gene-Environment Interrelations in Prostate Cancer

Richard B. Hayes

INTRODUCTION

With rapidly expanding knowledge about the human genome and its variations in populations, great opportunities are becoming available to incorporate genetic assessments into epidemiologic studies of prostate cancer. In particular, the study of gene-environment interrelations provides a potentially powerful approach for identifying the causes of this disease. As importantly, correlative knowledge about the biologic implications of genetic risk factors for this disease will enhance our understanding of underlying biologic mechanisms in prostate carcinogenesis. To attain this goal, however, will require redoubled efforts and new approaches to the way epidemiologic research is carried out. Here, we describe the current status of gene-environment investigations in prostate cancer epidemiology, illustrate barriers to progress, and attempt to provide some directions for future work.

In a most general sense, all diseases have genetic and environmental components. For some (e.g., highly infectious diseases) environmental factors predominate, while for others (e.g., inherited syndromes) the genetic component is most important. In 1981, Doll and Peto (1) attributed about 85 percent of cancer to environmental causes. Genetic determinants were not explicitly considered, although about 2 percent of cancers were attributed to congenital causes. With increased knowledge about rare mutations that show moderate to high penetrance (i.e., high relative risk) for cancer, as seen in high-risk families (e.g., *p53*, *BRCA1*, *FAP*), it can currently be estimated that about 10–15 percent of cancer has a predominantly inherited basis, due typically to single gene disorders. Beyond this relatively small fraction of cancers, it is also suspected that the large remainder of sporadic (environmental) cancers also have an important genetic component, due to complex interactions of one or more polymorphic genetic variants of low penetrance (i.e., low relative risk) which contribute singly or in combination to the overall cancer burden because of their high (cumulative) prevalence in the population (i.e., high attributable risk).

One or more high-penetrance genes appear responsible for a substantial proportion of prostate cancer in high-risk

families, although the mode of inheritance of familial cancer is unclear. Risks tend to be greater for men who report prostate cancer in their brothers than in their fathers, consistent with recessive or X-linked transmission (2), however, segregation analyses tend to show an autosomal dominant pattern (3, 4). Although the rare cancer genes that track with prostate cancer in high-risk families are certainly of importance for carriers, a substantially larger component of prostate cancer in the general population may be due to common low-penetrance polymorphisms (5).

An environmental role in prostate cancer is strongly suggested by the some 30-fold variation internationally in disease incidence and the observation that prostate cancer rates in immigrants tend to approach those of the host country. The challenge in prostate cancer epidemiology is to exploit the rapidly growing knowledge about the molecular basis of human genetic variation and statistical approaches to the analysis of gene-environment, gene-gene, and more complex relations in the etiology of this disease (6). With growing understanding of prostate cancer biology, population- and family-based epidemiologic studies are beginning to elucidate the impact of human genetic variation on this disease. As etiologic pathways become defined, it will be increasingly possible to integrate multifactorial genetic and environmental components into these studies.

MODELS OF GENE-EXPOSURE RELATIONS

Considering a genetic polymorphism and an environmental exposure as dichotomous factors, these relations can be summarized for disease incidence (I) and associated risks (R) as shown in table 1. This model can be generalized to continuous factors and could include biochemical parameters, even though the relative genetic and environmental contribution of these parameters may be uncertain. With comparisons to the referent ($R_{ref} = 1.0$), R_g estimates risk due to the gene (main effect of the allelic variant, in the absence of the environmental factor), R_e estimates risk due to the environmental factor (main effect of the environment, in the absence of the at-risk allelic variant), and R_{g-e} estimates the joint effect of both. Other comparisons of interest are the contingent risks: $R_{g^*e} = I_{11}/I_{10}$ and $R_{eg^*} = I_{11}/I_{01}$. For example, the impact of a genetic factor may be most effectively studied among a subgroup of individuals with a particular environmental exposure, that is R_{g^*e} .

The genetic and environmental effects on disease risk are considered independent, on a multiplicative scale, if the observed joint effect, R_{g-e} , is consistent with $R_g \times R_e$ or, on

Received for publication September 5, 2000, and accepted for publication February 27, 2001.

Abbreviation: PhIP, 2-amino-1-methyl-6-phenylimidazol[4,5-*b*]pyridine.

From the Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.

Correspondence to Dr. Richard B. Hayes, Occupational Studies Branch, Division of Cancer Epidemiology and Genetics, EPS 8114, Bethesda, MD 20892 (e-mail: hayesr@mail.nih.gov).

TABLE 1. Relations between a genetic polymorphism and an environmental exposure when both are considered as dichotomous variables

Exposure	Susceptibility genotype	Incidence (<i>I</i>)	Risk (<i>R</i>)
No	No	I_{00}	R_{ref}
No	Yes	I_{01}	R_g
Yes	No	I_{10}	R_e
Yes	Yes	I_{11}	R_{g-e}

an additive scale, with $R_g + R_e - 1$. If R_{g-e} differs from its expected value (based upon the respective multiplicative or additive model), a gene-environment interaction is said to exist. Interaction is tested by examination of the statistical significance of the deviation (θ) of R_{g-e} from its model-based expectation. Similarly, gene-gene (R_{g-g}) and environment-environment (R_{e-e}) effects are also of interest.

GENE-EXPOSURE RELATIONS IN PROSTATE CANCER

Cancer genes identified in high-risk families

At least three loci on chromosome 1 (7-11), one locus on chromosome X (12), and one locus on chromosome 20 (13) have been identified by linkage analysis in high-risk families as potential sites for high-penetrance prostate cancer genes; however, confirmation in other family series has been difficult, suggesting that familial prostate cancer is heterogeneous (9, 14, 15). Recently, linkage studies in Utah pedigrees identified the first putative prostate cancer susceptibility gene, *HPC2/ELAC2*, located on 17p (16). Case-control studies also point to relatively common polymorphic variants in *ELAC2* associated with low-penetrance risk for prostate cancer (two- to threefold risk) (16, 17).

There is evidence that environmental factors can influence the penetrance of high-risk cancer genes (18), however, the impact of environmental factors in high-risk prostate cancer families has not been explored. Given the apparent genetic heterogeneity of familial prostate cancer, gene-environment analyses in this setting will be enhanced by knowledge of the specific high-risk genes and their cancer-related mutations.

Metabolic polymorphisms and chemical exposures in prostate cancer

The model for gene-environment interactions derives historically from observations of idiosyncratic reactions to pharmacologic agents. For example, about 50 percent of subjects in Western populations have an inherited decreased ability to effectively *N*-acetylate certain drugs such as isoniazid. This phenotypic trait is due to certain polymorphic variants in *N*-acetyltransferase-2 (*NAT2*), the principal gene responsible for the *N*-acetylation of these compounds. Lower et al. (19) first extended this model in pharmacogenetics to cancer, and Cartwright et al. (20) related this poly-

morphic trait to risk of bladder cancer in workers exposed to selected aromatic amine dyes.

Etiologic leads in prostate cancer are being explored following this model of gene-environment interaction. Meat cooked at high temperatures produces heterocyclic amines including 2-amino-1-methyl-6-phenylimidazol[4,5-*b*]pyridine (PhIP), a compound that causes invasive prostate cancer in a rat model (21). Human prostate tissue is capable of metabolically activating cooked meat carcinogens (22), and in a small study, a polymorphic variant in *NAT1* was associated with increased risk for prostate cancer (23). In comparing progress on gene-environment interactions in prostate cancer with the bladder cancer model, the exposure (PhIP and other products of high-temperature cooking) has not yet been related directly to prostate cancer risk in humans. The genetic associations with *NAT1* (and other genes involved in heterocyclic amines metabolism) need to be substantiated, and suggested gene-environment interrelations need to be evaluated in studies of sufficient statistical power (see below).

Other metabolic polymorphisms have also been evaluated, however, environmental correlates are uncertain and have not yet been considered in any detail. Prostate cancer was weakly associated with the *CYP2D6*B* allele in one study (24) and, in another study, with the nondeleted (functional) genotype of *GSTT1* (25). No differences were found in a small case-control study of prostate cancer assessing the 609 C→T polymorphism in *NQO1* (the *NAD(P)H*: quinone oxidoreductase gene) (26).

Gene-gene relations in prostate cancer

Gene-gene effects of common polymorphisms in the androgen receptor gene (*AR*) have been studied for prostate cancer. The length of CAG repeat sequences tend to be shorter in African-American than white and Asian men (27). Several, but not all, studies show shorter CAG repeat sequences among prostate cancer cases, particularly for advanced disease (28-34).

Stanford et al. (31) found little effect of CAG and GGN repeat length polymorphisms in *AR* (when considered as main effects, i.e., in the absence of the other factor); however, a twofold joint effect (R_{g-g}) was found for the combination of short repeats in both CAG and GGN. Xue et al. (35) also found little effect for the CAG repeat polymorphism in *AR* and for a polymorphism in the prostate-specific antigen gene, however, a fivefold increase in risk for prostate cancer was found among subjects who had both at-risk genotypes.

ANALYTICAL CONSIDERATIONS

The bladder cancer model discussed above is something of an ideal case. First, employment in dye factories is associated with a high attributable risk (most bladder cancer cases can be presumed to be due to the exposure) and the putative chemical exposures are relatively specific. Second, the gene is functionally relevant in that it inactivates at least some of these compounds, and the polymorphic variants in this gene alter its metabolic capacity.

In prostate cancer, environmental risk factors are not well defined, particularly with regard to specific chemical agents; consequently, attributable risks for prostate cancer and associated metabolic pathways are uncertain. Expanded gene-environment studies, such as those for PhIP and *NAT1*, provide a new approach to better characterize environmental risk factors by identifying high risks for particular exposures in genetic sub-groups.

Statistical tests for interaction of genetic variants and environment require large sample sizes. For relatively common exposures (20 percent) and gene variants (20 percent) with main effect risks (R_g and R_e) of twofold, about 7,800 subjects (3,900 cases and 3,900 controls) are needed in a case-control study to evaluate a twofold or greater multiplicative interaction ($\theta = 2$, that is $R_{g-e} = 8$, versus an expected risk under the null of $R_{g-e} = 4$) (36). In planning studies to assess interaction, we have little guidance as to what a reasonable minimum detectable interaction is. For example, a recent meta-analysis showed a multiplicative effect of only 1.3 (37) for smoking and *NAT2* in relation to bladder cancer. At this level of interaction, about 40,000 subjects would be needed for our example. The underlying biologic basis for the choice of the interaction parameter (e.g., additive or multiplicative) is also unclear.

Considered more broadly, however, gene-environment and gene-gene effects that express as statistical interactions in epidemiologic studies may only account for a relatively small proportion of this disease. For example, if androgen-related genetic polymorphisms (e.g., CAG repeats in *AR*) influence risk by altering cell turnover, this would likely influence prostate cancer risk independent of factors (i.e., chemical carcinogens) that cause mutations in these cells. Similarly, within a broad range of environmental exposures, the kinetics of DNA repair associated with polymorphisms in repair genes is likely to be independent of the rate of genetic damage.

Where the test for statistical interaction is not the primary focus of study, sample sizes tend to become manageable. For example, to ask whether a common exposure (20 percent) is associated with a twofold or greater risk of disease in a biologically interesting study group (i.e., the contingent risk, $R_{e|g}$) requires only about 270 subjects (135 cases and 135 controls). (In practice, one is more likely to evaluate genetic effects in populations with interesting environmental exposures (i.e., $R_{g|e}$), as in the occupational bladder cancer example.) In contrast to the large sample sizes required for evaluation of interaction terms, assessment of joint effects can be carried out with substantially fewer subjects. A joint effect of $R_{g-e} = 2.0$ or greater for common exposures (20 percent) and gene variants (20 percent) would require only about 1,100 subjects (i.e., assuming independence of the risk factor distribution in controls, about 4 percent of controls would be positive for both factors). Further efficiencies could, of course, be achieved for interaction and joint main effect tests by over-sampling of the environmentally (or genetically) exposed group.

Interaction tests (θ) are attractive because a positive finding suggests a necessary biologic interrelation; however, their use in practice may be limited, for statistical and biologic rea-

sons. However, studies designed to estimate joint effects (R_{g-e}) efficiently provide information about combinations of factors, although the analysis does not distinguish factors on the same or on different disease pathways. Another efficient approach to the elaboration of gene-environment interrelations is the study of populations with unique exposure or genetic characteristics.

FUTURE DIRECTIONS

Although both environment and genetics appear to play important roles in human prostate cancer, few strong risk factors have consistently been identified. The epidemiologic study of genes and environment in combination may provide insights that are not possible by examining one factor at a time.

The large sample sizes required to address interrelations of genetic and environmental factors in population studies has been stressed in this presentation. New initiatives, including multi-investigator collaborations, will be needed to achieve this. Modern cancer epidemiology has been shifting from case-control to cohort investigations to better address temporal relations in cancer etiology. Infrastructure development and maintenance for these large-scale efforts will continue to expand. Family-based linkage studies will also require larger scale approaches to characterize high-penetrance prostate cancer genes.

Epidemiologists studying prostate cancer have examined a limited number of genetic variants in a few candidate genes as main effects or in association with other genes. Studies to assess gene-environment interrelations are only beginning. Epidemiologists will soon have available many thousands of variants in the human genome for study. Two basic approaches can be considered for use of these data for identifying disease genes: linkage analyses in cancer-prone pedigrees and (what geneticists refer to as) association studies, that is family- and population-based case-control studies.

For linkage analysis, genome-wide scans with a large number of random genetic markers (short tandem repeats and single-nucleotide polymorphisms) are being used to search for prostate cancer loci. Large collaborative investigations are being carried out to resolve discrepancies between studies. Linkage studies will be strengthened by the growing number of random markers available for the human genome, although the number of markers and the number of family pedigrees required for loci identification is still uncertain (38), particularly for genetic associations of low relative risk (e.g., < 2) (39). In this genetically-based approach, triage of potentially true biologic associations from the large number of false positives will require the ability to rapidly verify findings in independent studies. Associations that survive this culling process may provide leads for the identification of the specific genes involved, recognizing that finding the underlying specific genetic variants and their functional correlates can be a laborious process (16, 40).

The second basic approach, association studies, will consider genetic variants in candidate genes that are along putative causal pathways for prostate cancer, for example in

hormone and growth factor regulating genes and their biochemical antecedents. Nucleotide base variants of known functional significance or at least resulting in changes in amino acid sequence or level of protein expression will receive priority. Even with this more restricted biologically-based approach, the identification within these pathways of single and multiple interrelating genes of potential etiologic significance will be a great challenge. Exploiting knowledge about environmental exposures in population-based studies will be essential for refining this search.

Genetic stratification approaches are being developed to address threats of confounding by ethnicity (i.e., population stratification) (41) in population-based studies, even if such biases are minor (42). Case-sibling (or -cousin) studies, which by design control for confounding by ethnicity, may also play an expanded role in prostate cancer epidemiology (43, 44) as a bridge between classical linkage analyses and population-based studies.

Information management and technology for high-throughput genetic analysis are advancing at a rapid pace; however, careful validation of the gene variant databases and the methods used for genotyping will be essential for successful epidemiologic research. Most of the genetic variants in the human genome database are derived from just a few people, leaving many variants unidentified. Adequate checks for erroneous variants are also not systematized. Sequencing in a sufficient number of subjects (perhaps 100–200) will be needed to characterize genetic variation for specific candidate genes in populations of epidemiologic interest. For random single-nucleotide polymorphism investigations, sequencing is clearly impractical at this time; however, the single-nucleotide polymorphism variants identified from the human genome database should be verified in a similar manner.

Improvements in case definition are also crucial for advancement in this field. The distinction of clinically significant prostate cancer is increasingly difficult, particularly due to widespread prostate cancer screening with prostate-specific antigen. Limiting epidemiologic analyses to advanced stage cancer partially addresses this problem, however, the development of biologic markers predictive for cancer aggressiveness would allow for the study of the full spectrum of clinically relevant prostate cancer.

In summary, the integration of genetic and environmental factors in large-scale epidemiologic studies of prostate cancer are needed to increase our insight about the causes and etiologic mechanisms underlying this disease. Methodological developments in information management and statistical analysis will be needed to keep pace with the potential richness of the soon to be available genetic study material.

REFERENCES

- Doll R, Peto R. The causes of cancer. New York, NY: Oxford University Press, 1981.
- Monroe KR, Yu MC, Kolonel LN, et al. Evidence of an X-linked or recessive genetic component to prostate cancer risk. *Nat Med* 1995;1:827–9.
- Carter BS, Beaty TH, Steinberg GD, et al. Mendelian inheritance of familial prostate cancer. *Proc Natl Acad Sci U S A* 1992;89:3367–71.
- Schaid DJ, McDonnell SK, Blute MI, et al. Evidence for autosomal dominant inheritance of prostate cancer. *Am J Hum Genet* 1998;62:1425–38.
- Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer. *N Engl J Med* 2000;343:78–85.
- Seminara D, ed. Innovative study designs and analytic approaches to the genetic epidemiology of cancer. *J Natl Cancer Inst Monogr* 1999;26:1–105.
- Smith JR, Freije D, Carpten JD, et al. Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. *Science* 1996;274:1371–4.
- Berthon P, Valeri A, Cohen-Akenine A, et al. Predisposing gene for early-onset prostate cancer, localized on chromosome 1q42.2–43. *Am J Hum Genet* 1998;62:1416–24.
- Gibbs M, Chakrabarti L, Stanford JL, et al. Analysis of chromosome 1q42.2–43 in 152 families with high risk of prostate cancer. *Am J Hum Genet* 1999;64:1087–95.
- Gibbs M, Stanford JL, McIndoe RA, et al. Evidence for a rare prostate cancer-susceptibility locus at chromosome 1p36. *Am J Hum Genet* 1999;64:776–87.
- Gronberg H, Smith J, Emanuelsson M, et al. In Swedish families with hereditary prostate cancer, linkage to the *HPC1* locus on chromosome 1q24–25 is restricted to families with early-onset prostate cancer. *Am J Hum Genet* 1999;65:134–40.
- Xu JF, Meyers D, Freije D, et al. Evidence for a prostate cancer susceptibility locus on the X chromosome. *Nat Genet* 1998;20:175–9.
- Berry R, Schroeder JJ, French AJ, et al. Evidence for a prostate cancer-susceptibility locus on chromosome 20. *Am J Hum Genet* 2000;67:82–91.
- Whittemore A, Lin IG, Oakley-Girven I, et al. No evidence of linkage for chromosome 1q42.2–43 in prostate cancer. *Am J Hum Genet* 1999;65:254–6.
- Xu J. Combined analysis of hereditary prostate cancer linkage to 1q24–25: results from 772 hereditary prostate cancer families from the International Consortium for Prostate Cancer Genetics. *Am J Hum Genet* 2000;66:945–57.
- Tavtigian SV, Simard J, Teng DHF, et al. A candidate prostate cancer susceptibility gene at chromosome 17q. *Nat Genet* 2001;27:172–80.
- Rebbeck TR, Walker AH, Zeigler-Johnson C, et al. Association of *HPC2/ELAC2* genotypes and prostate cancer. *Am J Hum Genet* 2000;67:1014–19.
- Narod SA, Risch H, Moslehi R, et al. Oral contraceptives and the risk of hereditary ovarian cancer. *N Engl J Med* 1998;339:424–8.
- Lower GM Jr, Nilsson T, Nelson CE, et al. *N*-acetyltransferase phenotype and risk in urinary bladder cancer: approaches in molecular epidemiology. Preliminary results in Sweden and Denmark. *Environ Health Perspect* 1979;29:71–9.
- Cartwright RA, Rogers HJ, Barham-Hall D, et al. Role of *N*-acetyltransferase phenotypes in bladder carcinogenesis: a pharmacogenetic epidemiologic approach to bladder cancer. *Lancet* 1982;2:842–5.
- Shirai T, Cui L, Takahashi S, et al. Carcinogenicity of 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (PhIP) in the rat prostate and induction of invasive carcinomas by subsequent treatment with testosterone propionate. *Cancer Lett* 1999;143:217–21.
- Williams JA, Martin FL, Muir GH, et al. Metabolic activation of carcinogens and expression of various cytochrome P450 in human prostate tissue. *Carcinogenesis* 2000;21:1683–9.
- Fukutome K, Watanabe M, Shiraiishi T, et al. *N*-acetyltransferase 1 genetic polymorphism influences the risk of prostate cancer development. *Cancer Lett* 1999;136:83–7.
- Febbo PG, Kantoff PW, Giovannucci E, et al. Debrisoquine hydroxylase (*CYP2D6*) and prostate cancer. *Cancer Epidemiol Biomarkers Prev* 1998;7:1075–8.
- Rebbeck TR, Walker AH, Jaffe JM, et al. Glutathione *S*-transferase-mu (*GSTM1*) and -theta (*GSTT1*) genotypes in the

- etiology of prostate cancer. *Cancer Epidemiol Biomarkers Prev* 1999;7:357-62.
26. Steiner M, Hillenbrand M, Borkowski M, et al. 609 C→T polymorphism in *NAD(P)H*: quinone oxidoreductase gene in patients with prostatic adenocarcinoma or benign prostatic hyperplasia. *Cancer Lett* 1999;135:67-71.
 27. Irvine RA, Yu MC, Ross RK, et al. The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. *Cancer Res* 1995;55:1937-40.
 28. Ingles SA, Ross RK, Yu MC, et al. Association of prostate cancer risk with genetic polymorphisms in vitamin D receptor and androgen receptor. *J Natl Cancer Inst* 1997;89:166-70.
 29. Giovannucci E, Stampfer MJ, Krithivas K, et al. The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc Natl Acad Sci U S A* 1997;94:3320-3.
 30. Hakimi JM, Schoenberg MP, Rondinelli RH, et al. Androgen receptor variants with short glutamine or glycine repeats may identify unique subpopulations of men with prostate cancer. *Clin Cancer Res* 1997;3:1599-608.
 31. Stanford JL, Just JJ, Gibbs M, et al. Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer Res* 1997;57:1194-8.
 32. Correa-Cerro L, Wohr G, Haussler J, et al. (CAG)_nCAA and GGN repeats in the human androgen receptor gene are not associated with prostate cancer in a French-German population. *Eur J Hum Genet* 1999;7:357-62.
 33. Bratt O, Borg A, Kristofferson U, et al. CAG repeat length in the androgen receptor gene is related to age at diagnosis of prostate cancer and response to endocrine therapy, but not to prostate cancer risk. *Br J Cancer* 1999;81:672-6.
 34. Hsing AW, Gao YT, Wu G, et al. Polymorphic CAG and GGN repeat lengths in the androgen receptor gene and prostate cancer risk: a population-based case-control study in China. *Cancer Res* 2001;60:5111-16.
 35. Xue WM, Irvine RA, Yu MC, et al. Susceptibility to prostate cancer: interaction between genotypes at the androgen receptor and prostate-specific antigen loci. *Cancer Res* 2000;60:839-41.
 36. Lubin JH, Gail MH. On power and sample-size for studying features of the relative odds of disease. *Am J Epidemiol* 1990;131:552-66.
 37. Marcus PM, Hayes RB, Vineis P, et al. Cigarette smoking, *N*-acetyltransferase 2 acetylation status, and bladder cancer risk: a case-series meta-analysis of a gene-environment interaction. *Cancer Epidemiol Biomarkers Prev* 2000;9:461-7.
 38. Terwilliger JD, Goring HHH. Gene mapping in the 20th and 21st centuries: statistical methods, data analysis, and experimental design. *Hum Biol* 2000;72:63-132.
 39. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;273:1516-17.
 40. Althuser D, Daly M, Kruglyak L. Guilt by association. *Nat Genet* 2000;26:135-7.
 41. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 1999;65:220-8.
 42. Wacholder S, Rothman N, Caporaso N. Population stratification in epidemiologic studies of common genetic variants and cancer: quantification of bias. *J Natl Cancer Inst* 2000;92:1151-8.
 43. Weinberg CR, Umbach DM. Choosing a retrospective design to assess joint genetic and environmental contributions to risk. *Am J Epidemiol* 2000;152:197-203.
 44. Risch NJ. Searching for genetic determinants in the new millennium. *Nature* 2001;405:847-56.