

Gene-environment Interactions and the Complexity of Human Genetic Diseases

Christoph E. Nabholz, PhD; Jan von Overbeck, MD

In the assessment of mortality and morbidity risk, the ability of family history and genetic test results to predict the age of occurrence, severity, and long-term prognosis of 'genetic' diseases is important. An increasing number of gene-gene and gene-environment interactions have been demonstrated in a number of monogenic Mendelian diseases. These interactions can significantly modify the clinical presentation (disease phenotype) of diseases previously regarded purely as 'genetic.' As a result, 'genetic' diseases can be positioned in a continuum between classic Mendelian and complex disease where the extremes, pure genetic or solely non-genetic, do not exist. The position of any given disease in this continuum is defined by three components: the major gene(s) contributing to the phenotype, the variability added by modifier genes and the significance of environmental factors influencing the phenotype. As the predictive value of genetic test results can be significantly influenced by additional genetic and environmental risk factors, a better understanding of these factors may influence the quantification of mortality and morbidity risk.

Address: Swiss Reinsurance Company, Mythenquai 50/60, 8022 Zurich Switzerland; ph: 41 43 285 7278; fax: 41 43 282 7278; e-mail: Christoph_Nabholz@swissre.com.

Correspondent: Dr. Christoph Nabholz, Genetics Consultant.

Key words: Genetics, environment, genotype, phenotype, cystic fibrosis, breast cancer, mortality, morbidity, insurance.

The rapid identification of genes associated with human disease has revolutionized the field of medical genetics. Developing new genome analysis tools, and the sequencing and assembly of the human genome has given us fresh molecular insights into disease and new diagnostic and prognostic tools.^{1,2} This enormous effort, of developing new technologies and managing the large scale sequencing of the human genome, was the product of the Human Genome Project. This landmark project started in 1990 and ended—ahead of time and below budget—in April 2003.³

The immediate benefit from this effort was the facilitation of identifying mutations in human genes that lead to disease. So far, more

than 1500 disease genes, containing mutations or polymorphisms known to cause disease, have been mapped and annotated on the human genome. The identification of these alterations in the human genome allowed for the development of new diagnostic tools, such as genetic tests. A genetic test works by analyzing human DNA or RNA, the result of which is used to identify the presence or absence of alterations in the patient's genetic material. Deviations from the norm can show a predisposition towards, or actually confirm the presence of, a disease.

More than 1000 disease-causing DNA alterations in the disease gene loci can be detected with commercially available tests.⁴ This is 3 times as many as 5 years ago, and

the number is expected to double over the next 5 years. Most of these genetic tests have been developed to identify disease-conferring genes, thereby predicting a genetic predisposition towards disease in individuals before they start to show any symptoms (predictive genetic testing), or confirming disease in patients who show some signs of symptoms (diagnostic genetic testing).

Genetic tests are currently used in a range of different clinical medical disciplines:

- Carrier testing: Thalassaemia
- Pre-implantation or prenatal testing: Down syndrome
- Newborn screening: Phenylketonuria
- Predictive genetic testing: Huntington's disease
- Diagnostic testing: Hereditary hemochromatosis
- Pharmacogenetic testing: CYP2D6 (drug metabolism)

While there are some overlaps, genetic tests differ both in their clinical use and relevance.² Many of these diseases are considered to be inherited in a single gene, so-called "monogenic," fashion according to Mendel's law (often referred to as Mendelian diseases).

This paper focuses on two relevant monogenic Mendelian diseases: cystic fibrosis and familial breast cancer. During the last 20 years, disease-causing mutations in key genes have been identified for most Mendelian disorders. Initially it was hoped that identifying the disease-causing mutations in the disease gene loci would correlate with the age of onset and the severity of the disease. However, a lot more has since been learned about the etiology of many Mendelian diseases, and it has become questionable whether the genotype alone can, and should, be used to predict the future clinical outcome. It is recognized that phenotypic differences in all Mendelian disorders are caused by the combined action of a few major genes, by a variable number of modifier genes and by environmental factors.

Even for mutations associated with a very high probability of disease, the clinical pre-

sentation and prognosis are significantly influenced by environmental factors. Consequently, molecular genetic DNA test results are not deterministic and can only identify the level of susceptibility to disease. In fact patients may never get the disease, or at least not until they are very old. Hence, if the goal of using genetic test results as part of a successful personalized clinical strategy to prevent or cure disease is to be realized, more needs to be learned about the contribution of modifier genes and environmental risk factors before mortality and morbidity risk can be properly quantified.

CYSTIC FIBROSIS: A CLASSIC MENDELIAN DISEASE?

Cystic fibrosis (CF) is the most common life-limiting autosomal recessive disorder in Caucasians,⁵ with a frequency of about 1 in 2500. CF is considered to be a monogenic Mendelian recessive disorder caused by mutations in the cystic fibrosis trans-membrane conductance regulator (*CFTR*) gene. When *CFTR* was first cloned, it was hoped that the mutation analysis of the *CFTR* loci alone would be sufficient to predict the phenotypic manifestation in the patient.⁶ Today, we know that mutations in the *CFTR* gene loci almost always cause the CF phenotype. The combination of the genotype with the presence of one or more characteristic CF phenotypic features can be considered diagnostic. However, while the genotype correlates with the pancreatic status, it does not correlate with the presence or severity of pulmonary disease in CF.⁷⁻¹⁰

To simplify the genotype-phenotype analysis of the 1000-plus disease-causing mutations found in the *CFTR* loci, the identified mutations were grouped into six classes.¹¹ The defined *CFTR* classes are based on the predicted functional consequences for the *CFTR* protein (Figure 1). Out of the disease-causing mutations described, the ones that lead to a loss of function of the *CFTR* protein tend to be severe.^{7,9,10} Even for the most common genotype, such as the $\Delta F508/\Delta F508$, the

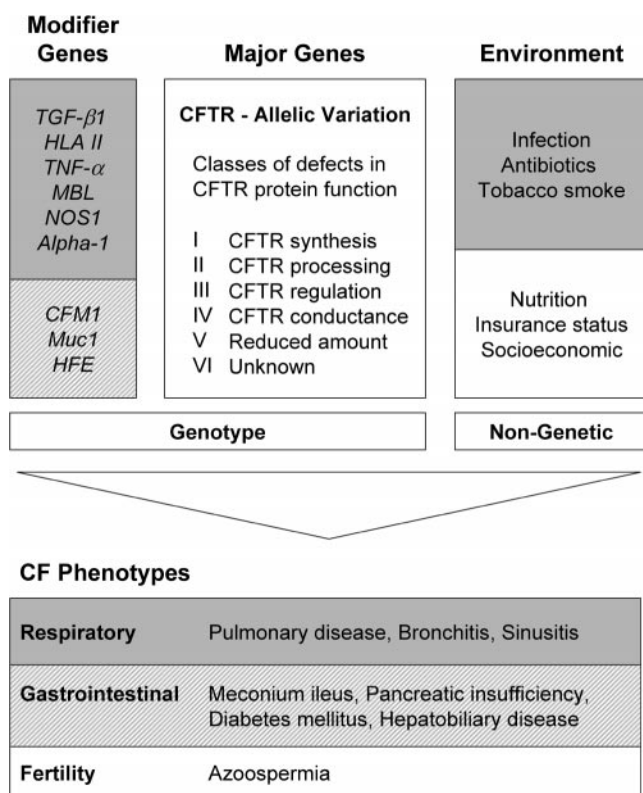


Figure 1. *Phenotype Modification in Cystic Fibrosis.* Gene-gene and gene-environment interactions modulate the phenotype of cystic fibrosis (CF). See text for details. (CF = cystic fibrosis; CFTR = cystic fibrosis transmembrane conductance regulator; TGF-β1 = transforming growth factor beta (1); HLA II = human leukocyte antigen class II; TNF-α = tumor necrosis factor alpha; MBL = mannose-binding lectin; NOS1 = nitric oxide synthase (1); Alpha-1 = α1-antichymotrypsin; CFM1 = cystic fibrosis modulator locus 1; Muc1 = mucin 1; HFE = hemochromatosis).

clinical course of pulmonary disease (the pulmonary phenotype—the main cause of mortality) varies dramatically.¹² This suggests that the pulmonary disease is strongly influenced by other genetic or environmental factors. In recent years, modifier genes and environmental risk factors have been identified that significantly contribute to the CF phenotype (Figure 1).⁹

For example, polymorphisms in the nitric oxide synthase 1 (*NOS1*) gene correlate with the variability in the frequency and severity of microbial infections, indicating an interaction between variants of *NOS1* and the environment.¹³ Similar gene-environment interactions were found for *CFTR* and other CF modifier genes.^{9,10} Both the list of genetic as

well as environmental CF modifying risk factors is constantly increasing (Figure 1, gray areas). So far, more than 9 genetic CF modifiers have been described.^{9,10,13–20}

The many environmental CF modifiers include specific risk factors such as exposure to tobacco smoke, but also other loosely defined risk factors such as health insurance or socioeconomic status. The latter may reflect nutritional status, access to antibiotics, or other underlying CF modifiers.^{21–23}

Therefore, this monogenic recessive Mendelian disease has turned out to be a rather complex disorder. Even though genotype-phenotype correlation is imprecise, neonatal screening programs have been introduced in many countries. As a consequence, 50% of all patients with CF in the United States are diagnosed by the time they are 6 months old and 90% by age 8.⁵ Modern screening is based on immunoreactive trypsinogen assays performed on blood spots. If elevated, the positive results are confirmed by molecular genetic DNA testing. To complicate it even further, a recent study has reported that some patients with a milder CF phenotype do not have mutations in the *CFTR* loci.²⁴ As a result, clinicians base their clinical strategy on the presence of CF phenotypes, genetic test results and on other genetic and environmental risk factors.⁵

THE CONTINUUM FROM MONOGENIC TO COMPLEX DISEASE

It is now understood that many inherited diseases are complex disorders, rather than diseases that follow simple monogenic Mendelian patterns, as once thought.^{8,25} The number of classic Mendelian disorders explained by DNA alteration in a single gene locus associated with a single disease phenotype is shrinking. It has been proposed that all the etiology of all diseases can be positioned in a continuum between classic Mendelian inheritance and complex disorders where neither extreme, diseases with pure genetic or solely environmental origin, exists. The position of any given disease in this continuum

is defined by three components: the major gene(s) contributing to the phenotype, the variability added by modifier genes, and the significance of environmental factors influencing the phenotype.

THE BRCA BIOLOGICAL PATHWAY IN BREAST AND OVARIAN CANCER

Conventional thinking had been that familial breast cancer likely followed conventional monogenic Mendelian patterns based on discovery of associated tumor suppressor genes *BRCA1* and *BRCA2*.²⁶ However, the "placement" of *BRCA* associated breast cancer in the continuum has shifted towards multifactorial disorders over time. Breast cancer is the most common malignancy among women in industrialized developed countries. About 10% of all breast cancer cases have a family history of breast cancer or ovarian cancer. Less than half of these familial cases are linked to mutations in the *BRCA* genes.²⁷

Women carrying disease-causing mutations in the *BRCA* genes are at a significantly increased risk of developing breast cancer and ovarian cancer. Male *BRCA* mutation carriers rarely develop breast cancer. The estimated breast cancer risk for *BRCA* mutation carriers at age 70 ranges from 40% to 80%, and the risk of developing ovarian cancer is 15% to 65%.²⁷ The probability of developing breast cancer by mutations in the *BRCA* loci, at a given age, is referred to as the penetrance of the mutation. In addition to the breast cancer risk, mutations in the two genes have been linked to various other cancers such as prostate and colon cancer in *BRCA1* carriers and prostate, pancreatic and other cancers in *BRCA2* carriers.^{28,29} The variability of the penetrance of the *BRCA* genes was believed to be due mainly to different mutations in the *BRCA* genes (allelic variation) and *BRCA* modifying genes.²⁶ Among the genetic modifiers of *BRCA* susceptibility associated with breast cancer are the DNA-repair gene *RAD51* and with ovarian cancer, the *HRAS1* oncogene.^{8,26}

The results of the recent New York Breast

Cancer Study (NYBCS) show that environmental and lifestyle factors can significantly modify the penetrance of *BRCA* mutations, as well.³⁰ This cohort study identified a number of important risk modifiers that might have implications for breast cancer molecular genetic screening programs of the general population and suggested new preventive strategies for *BRCA* mutation carriers.

The NYBCS cohort study specifically included Ashkenazi Jewish probands only, because this founder population harbors three ancient *BRCA* mutant alleles. Since this population only requires screening for a few mutations, rather than requiring the sequencing of entire genes, molecular genetic DNA testing was facilitated and more accurate. Probandes were selected with an incidence of primary, invasive breast cancer, regardless of family history of cancer. To improve the accuracy of risk estimates, relatives of *BRCA* carriers were genotyped. With these results, the authors were able to calculate the definitive cancer risk. The lifetime risk of breast cancer among mutation carriers was 82% for both genes. The lifetime risk of ovarian cancer was 54% for *BRCA1* and 23% for *BRCA2* mutations. That compares to 10% risk for breast cancer and less than 2% risk for ovarian cancer in the general population.

A major finding of the NYBCS was that non-genetic factors significantly alter the penetrance of identical *BRCA* mutations. The birth cohort analysis showed that by age 50 the observed penetrance was 24% for women born before 1940, but 67% for women born later. This could not be explained by changes in medical practice over the decades covered by the study. Importantly, the non-genetic effect seen in the birth cohort study was much larger than for any of the reported modifier genes identified so far. This is consistent with the results of an earlier paper which proposed that environmental risk factors can significantly modify the penetrance of *BRCA* mutations and that those environmental risk factors are of increasing prevalence.³¹

The increasing penetrance of *BRCA* mutations is likely due to the increasing prevalence

of environmental risk factor modification. This is reflected in the NYBCS result that half of the identified *BRCA* mutation carriers reported no family history of breast or ovarian cancer (among mothers, sisters, grandmothers or aunts). Even though there was no positive family history for breast cancer, their breast cancer risk was calculated to be as high as the risk of a carrier with the same *BRCA* mutation who had a high-incidence family history. This may have clinical consequences, since *BRCA* mutation carriers with no family history would have been identified before the onset of breast cancer if they had been screened for *BRCA* mutations. Family history is a risk factor that reflects the contribution of major gene(s) to the phenotype, with variability added by modifier genes and significant environmental factors. Hence, use of high-incidence breast and ovarian cancer family history as a surrogate for *BRCA* mutation screening in the general population will fail to detect a proportion of women who have high-risk *BRCA* mutations.

The NYBCS study identified two lifestyle risk factors (physical exercise and lack of obesity in adolescence) that significantly delay breast cancer onset. Other modifying environmental risk factors for *BRCA*-associated breast cancer include age at menarche/puberty, reproductive/hormonal factors and smoking.³²

GENETIC TESTING: TODAY AND TOMORROW

The use of genetic tests to identify a future disease in presymptomatic patients is the most promising future application in clinical medicine. Almost daily we are confronted with news about genes that are associated with a human disease. Inevitably, great hopes are raised that identification of disease genes will allow for an immediate breakthrough in treatment, screening, or even a preventive clinical strategy. However, identifying disease-associated genes is only the first step in a long process. Long-term follow-up studies, many new techniques, drugs and clinical

strategies have to be developed before this information can be effectively used in clinical practice to treat or prevent a disease.

Clinicians have been put in a position where susceptibility gene tests are available long before successful clinical strategies are developed. As with any medical test, consumers are naturally anxious about taking genetic tests. They fear the worst; they are concerned about their health and the financial consequences that might follow if a genetic disorder is identified, including any potential impact on their insurance premiums, their job and other aspects of day-to-day life. Moreover, because preventive or curative strategies may not be currently available, genetic tests are believed to be deterministic, and therefore, the use of the results is considered discriminatory. However, while a test could indicate the presence of a genetic disorder, it does not generally mean that the patient will actually develop the disease identified by the test.

Most human diseases have a significant inheritable component. Identification of these genetic determinants will promote a greater understanding of the molecular mechanism of disease. While this information will inevitably be the basis for research into new drugs and new treatments, it might take generations to transform this knowledge into clinical medicine. The greatest impact of medical genetic research in the last decade was its role in improving our understanding of the pathophysiology of disease, and the development of new diagnostic tools.

Clinical practice has always been limited by its inability to distinguish biochemical abnormalities that are an *outcome* of disease from those that actually *cause* the disease. Research in medical genetics has allowed geneticists to identify proteins and biochemical pathways underlying common complex diseases, such as asthma and schizophrenia. It is now possible to distinguish mechanistically distinct forms of these complex diseases. Ultimately, this will lead to the definition of new subtypes and change the way human diseases are classified. As this new taxonomy

comes into common usage, it will change both the theory and the practice of clinical medicine.³³ One current example is the use of gene-expression profiles for the more accurate prognosis of breast cancer, which should improve the selection of patients for adjuvant systematic therapy.³⁴⁻³⁷

It has been questioned whether the identified disease genes associated with familial disorders are relevant for sporadic cases of a disorder that shows the same phenotype. As each gene inevitably interacts with many other genes, it can only be part of a biological pathway rather than being the only determinant. Genes in the very same network can cause a similar molecular dysfunction when altered and can therefore show the same phenotype. Hence, identifying the inherited component of the familial disorder might be key to identifying the components of the sporadic occurrence of the disorder.

The power of trying to identify the biological pathway of a disease by searching for other components was demonstrated by Hughes-Davies et al in a recent paper.³⁶ They were able to identify a new protein, EMSY, in the BRCA breast cancer disease pathway, by its ability to bind to the BRCA2 protein. More importantly, they showed that the EMSY gene is frequently over-expressed, almost exclusively in sporadic breast and ovarian cancer. The protein-protein interaction of EMSY with BRCA2, together with the clinical phenotypic overlap between familial BRCA2 and EMSY-over-expressed sporadic breast and ovarian cancer, suggest that both genes are in the same BRCA2 pathway.

Finally, there was speculation that EMSY over-expression may mimic loss of function of BRCA2 in inherited breast cancer, and may contribute to tumorigenesis of a subset of sporadic cancers.³⁶ Even though the same disease phenotype is found, the genetic defect causing breast cancer is different. Yet, the underlying biological pathway seems to be the same. Importantly, this example shows that study of inherited breast cancer allows the identification of a key regulatory pathway for a subset of sporadic breast cancer. Even

though this strategy has proven to be successful, it has been argued that the lack of understanding of the environmental risk factors contributing to common complex diseases is part of the reason why scientists have been largely unsuccessful in identifying the underlying genetic components.³⁷

CONCLUSION

When assessing mortality and morbidity risk, the ability of family history and genetic test results to predict the age of occurrence, severity, and long-term prognosis of 'genetic' diseases is important. It has been shown by the NYBCS cohort and other studies that environmental risk factors of increasing prevalence can significantly alter the risk of high penetrance mutations.³⁰ The growing prevalence of these environmental risk factors is reflected by the finding that a low incidence of family history of breast or ovarian cancer does not mean low risk. In summary, if clinicians are to use genetic tests in clinical practice to predict a future phenotype, more needs to be learned about gene-gene and gene-environment interactions.^{25,38} Hence, a better understanding of these additional risk factors may improve the quantification of mortality and morbidity risk.

We thank Alan Hester, Alan Tyler, Beatrice Baldinger, Kevin Somerville and Tim Dickenson for helpful comments on the manuscript.

REFERENCES

1. Guttmacher A, Collins F. Genomic medicine—a primer. *N Engl J Med.* 2002;347:1512–1520.
2. Burke W. Genetic testing. *N Engl J Med.* 2002;347:1867–1875.
3. Collins F, Green E, Guttmacher A, Guyer M. A vision for the future of genomics research. *Nature.* 2003;422:835–847.
4. Children's Health System and University of Washington, Seattle. Available at: <http://www.geneclinics.org/>.
5. Ratjen F, Doring G. Cystic fibrosis. *Lancet.* 2003;361:681–689.
6. Tsui L, Buchwald M, Barker D, et al. Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker. *Science.* 1985;230:1054–1057.

7. McKone E, Emerson S, Edwards K, Aitken M. Effect of genotype on phenotype and mortality in cystic fibrosis: a retrospective cohort study. *Lancet*. 2003;361:1671–1676.
8. Badano J, Katsanis N. Beyond Mendel: an evolving view of human genetic disease transmission. *Nat Rev Genet*. 2002;3:779–789.
9. Burke W. Genomics as a probe for disease biology. *N Eng J Med*. 2003;349:969–974.
10. Zielenski J. Genotype and phenotype in cystic fibrosis. *Respiration*. 2000;67:117–133.
11. Tsui L. The spectrum of cystic fibrosis mutations. *Trends Genet*. 1992;8:392–398.
12. Burke W, Aitken M, Chen S, Scott C. Variable severity of pulmonary disease in adults with identical cystic fibrosis mutations. *Chest*. 1992;102:506–509.
13. Grasmann H, Knauer N, Buscher R, Hubner K, Drazen J, Ratjen F. Airway nitric oxide levels in cystic fibrosis patients are related to a polymorphism in the neuronal nitric oxide synthase gene. *Am J Respir Crit Care Med*. 2000;162:2172–2176.
14. Arkwright P, Laurie S, Super M, et al. TGF-beta(1) genotype and accelerated decline in lung function of patients with cystic fibrosis. *Thorax*. 2000;55:459–462.
15. Aron Y, Polla BS, Bienvenu T, Dall'ava J, Dusser D, Hubert D. HLA class II polymorphism in cystic fibrosis. A possible modifier of pulmonary phenotype. *Am J Respir Crit Care Med*. 1999;159:1464–1468.
16. Mahadeva R, Sharples L, Ross-Russell R, Webb A, Bilton D, Lomas D. Association of alpha(1)-antichymotrypsin deficiency with milder lung disease in patients with cystic fibrosis. *Thorax*. 2001;56:53–58.
17. Parmley R, Gendler S. Cystic fibrosis mice lacking Muc1 have reduced amounts of intestinal mucus. *J Clin Invest*. 1998;102:1798–1806.
18. Garred P, Pressler T, Madsen H, et al. Association of mannose-binding lectin gene heterogeneity with severity of lung disease and survival in cystic fibrosis. *J Clin Invest*. 1999;104:431–437.
19. Hull J, Thomson A. Contribution of genetic factors other than CFTR to disease severity in cystic fibrosis. *Thorax*. 1998;53:1018–1021.
20. Rohlf E, Shaheen N, Silverman L. Is the hemochromatosis gene a modifier locus for cystic fibrosis? *Genet Test*. 1998;2:85–88.
21. Britton J. Effects of social class, sex, and region of residence on age at death from cystic fibrosis. *BMJ*. 1989;298:483–487.
22. Campbell PW III, Parker R, Roberts B, Krishnamani M, Phillips J III. Association of poor clinical status and heavy exposure to tobacco smoke in patients with cystic fibrosis who are homozygous for the F508 deletion. *J Pediatr*. 1992;120:261–264.
23. Curtis J, Burke W, Kassner A, Aitken M. Absence of health insurance is associated with decreased life expectancy in patients with cystic fibrosis. *Am J Respir Crit Care Med*. 1997;155:1921–1924.
24. Groman J, Meyer M, Wilmott R, Zeitlin P, Cutting G. Variant cystic fibrosis phenotypes in the absence of CFTR mutations. *N Eng J Med*. 2002;347:401–407.
25. Merikangas K, Risch N. Genomic priorities and public health. *Science*. 2003;302:599–601.
26. Narod SA. Modifiers of risk of hereditary breast and ovarian cancer. *Nat Rev Cancer*. 2002;2:113–123.
27. Wooster R, Weber B. Breast and ovarian cancer. *N Engl J Med*. 2003;348:2339–2347.
28. Ford D, Easton D, Bishop D, Narod S, Goldgar D. Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. *Lancet*. 1994;343:692–695.
29. The Breast Cancer Linkage Consortium. Cancer risks in BRCA2 mutation carriers. *J Natl Cancer Inst*. 1999;91:1310–1316.
30. King M, Marks J, Mandell J. Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. *Science*. 2003;302:643–646.
31. Narod S. Hereditary breast-ovarian cancer: how can we use the new DNA markers to improve patient management? *Clin Invest Med*. 1993;16:314–317.
32. Levy-Lahad E, Plon S. Cancer. A risky business—assessing breast cancer risk. *Science*. 2003;302:574–575.
33. Bell J. The double helix in clinical practice. *Nature*. 2003;421:414–416.
34. van't Veer L, Dai H, van de Vijver M, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002;415:530–536.
35. van de Vijver M, He YD, van't Veer L, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med*. 2002;347:1999–2009.
36. Hughes-Davies L, Huntsman D, Ruas M, et al. EMSY Links the BRCA2 Pathway to Sporadic Breast and Ovarian Cancer. *Cell*. 2003;115:523–535.
37. Chakravarti A, Little P. Nature, nurture and human disease. *Nature*. 2003;421:412–414.
38. Perera F. Environment and cancer: who are susceptible? *Science*. 1997;278:1068–1073.