

## Commentary

# The contribution of inherited genotype to breast cancer

Timothy R Rebbeck

Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, and Cancer Center, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA

**Correspondence:** Timothy R Rebbeck, PhD, Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, 904 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, USA. Tel: +1 215 898 1793; fax: +1 215 573 2265; e-mail: [trebbeck@cceb.med.upenn.edu](mailto:trebbeck@cceb.med.upenn.edu)

Received: 9 January 2002

Revisions requested: 5 February 2002

Revisions received: 13 February 2002

Accepted: 26 February 2002

Published: 4 April 2002

*Breast Cancer Res* 2002, 4:85-89

© 2002 BioMed Central Ltd  
(Print ISSN 1465-5411; Online ISSN 1465-542X)

## Abstract

The etiology of breast cancer is complex, and is likely to involve the actions of genes at multiple levels along the multistage carcinogenesis process. These inherited genotypes include those that affect the propensity to be exposed to breast carcinogens, and those associated with breast tumorigenesis directly. In addition, inherited genotypes may influence response to breast cancer chemoprevention and treatment. Studies relating inherited genotypes with breast cancer incidence and mortality should consider a broader spectrum of genes and their potential roles in multistage carcinogenesis than have been typically evaluated to date. Understanding the role of inherited genotype at different stages of carcinogenesis could improve our understanding of cancer biology, may identify specific exposures or events that correlate with carcinogenesis, or target relevant biochemical pathways for the development of preventive or therapeutic interventions.

**Keywords:** breast cancer, etiology, genetics, prevention

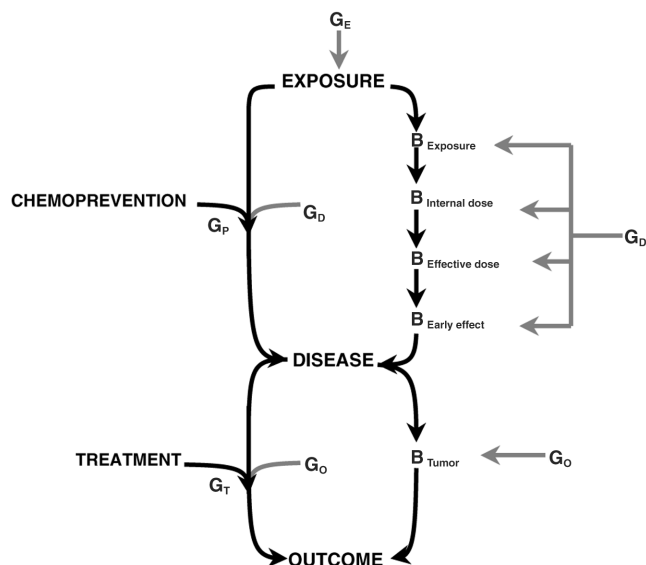
## Introduction

The multistage model of carcinogenesis describes the progression of normal cells to initiated and preneoplastic cells, and finally to malignant and metastatic disease. Inherited genotypes are known to play a prominent role in the initiation of breast tumors. For example, it is well known that germline mutations in *BRCA1* can initiate breast tumorigenesis [1]. However, inherited genotypes may also act at a variety of other steps in the multistage process of breast carcinogenesis. These genotypes may not only affect disease susceptibility, but also disease outcome.

Figure 1 presents a number of classes of genes acting in multistage carcinogenesis. In this presentation, G denotes genotypes at a single locus. Subscripts to G are used to distinguish different classes of genes that may act at different points in the development and progression of a

tumor. As described in detail below,  $G_E$  are involved in metabolic pathways that determine propensity to be exposed to breast carcinogens.  $G_P$  are involved in the metabolism of chemopreventive agents. These genes can determine the degree to which an individual's cancer risk may be modified, or the degree to which an individual will suffer adverse side effects from exposure to the agent.  $G_D$  include genes that are directly involved in the etiology of a tumor. These genes are typically involved in studies addressing disease susceptibility, and generally represent those genes considered in evaluating cancer risk.  $G_T$  are genes involved in the determination of drug dissemination and metabolism. Thus,  $G_T$  determine in part the efficacy or toxicity of a drug.  $G_O$  are determinants of a tumor's natural history, as it progresses from a more benign to a more advanced stage. Each of these classes of genes will be defined in this commentary, with examples that illustrate

Figure 1



Role of inherited susceptibility genotypes along the continuum of multistage carcinogenesis to determine exposure ( $G_E$ ), disease risk ( $G_D$ ), efficacy and response to chemoprevention ( $G_P$ ) or treatment ( $G_T$ ), and natural history or disease outcome ( $G_O$ ). B denotes biomarkers of exposure ( $B_{\text{exposure}}$ ), internal dose ( $B_{\text{internal dose}}$ ), effective dose ( $B_{\text{effective dose}}$ ), or early effect ( $B_{\text{early effect}}$ ) as defined by Schulte and Perera [12].

how inherited genotype information can be used to better understand the multistep carcinogenesis process and improve cancer detection, prevention, and treatment. It is important to note that the genes that fall into each of these classes are not mutually exclusive: genes involved in the metabolism of xenobiotic compounds could play the role of any of these multistep events, while other genes may play a greater role in one step than others. Implicit in this model is the role that inherited genotypes may play in cancer prevention by better defining modifiable aspects of carcinogenesis.

### Inherited genetic determinants of carcinogenic exposures: $G_E$

$G_E$  include genes involved in metabolic pathways that determine propensity to be exposed to endogenous or exogenous carcinogens. For example, a number of investigators have related  $G_E$  involved in neurotransmitter metabolism (e.g. *DRD4*, *SLCA3*, and *5-HTTLPR*) with propensity to smoke, smoking initiation at an early age, ability to stop smoking, and nicotine dependence [2–5]. Similarly, a number of investigators have related  $G_E$  of steroid hormone metabolism to interindividual variability in hormone-related risk factors. Among the most widely studied of these genes is *CYP17*, a member of the cytochrome P450 multigene family, which is responsible for conversion of progesterone

to androstenedione in the biosynthesis of estrone. A number of authors have evaluated the relationship of a 5'-untranslated region polymorphism at *CYP17* with age at menarche [6–9], use of hormone replacement therapy [6], or circulating hormone levels [9]. The relationship of *CYP17* genotype with age at menarche was not consistent in all studies, and there have not been independent confirmations of the reported associations of hormone replacement therapy use or circulating hormone levels. These examples point to the possibility that  $G_E$  may influence propensity to be exposed to 'traditional' risk factors, may elucidate biologically relevant prevention pathways, and may identify individuals who should be targeted for exposure modification or other preventive strategies.

### Inherited genetic determinants of chemoprevention response and toxicity: $G_P$

$G_P$  include genes that permit the prediction of an individual's response to a chemopreventive agent, and may define those individuals who could have increased toxicities when exposed to these agents. Few examples of  $G_P$  in cancer chemoprevention exist. Tamoxifen, an antiestrogenic drug used for the prevention of breast cancer, is metabolized to the antiestrogenic 4-hydroxytamoxifen (4-OHT) by *SULT1A1* and to N-desmethyl tamoxifen by *CYP3A4* [10,11]. The N-desmethyl form is quantitatively the major antiestrogenic metabolite, although the 4-OHT form is 100-fold more antiestrogenic than the N-desmethyl form. Therefore, the genes *SULT1A1* and *CYP3A4* are candidate  $G_P$ , because the variants that affect the metabolism of tamoxifen to 4-OHT or N-desmethyl tamoxifen indicate which individuals have either an optimal chemopreventive response or the potential for increased toxicity following tamoxifen administration. Although *SULT1A1* and *CYP3A4* encode key enzymes in the metabolism of tamoxifen, a number of other genes, including *CYP2D6* and *CYP2C9*, are also involved in tamoxifen metabolism. Furthermore, sulfate conjugation of 4-OHT by *SULT1A1* can reduce the antiestrogenic properties of tamoxifen metabolites. As with most complex metabolic systems, numerous  $G_P$  with multiple functions need to be considered to fully evaluate the pharmacogenetics of cancer chemoprevention. Knowledge of  $G_P$  may dictate that the protective or toxic effects of an agent occur only in a subset of individuals with a specific genotype.

### Inherited genetic determinants of cancer etiology: $G_D$

In the model presented in Figure 1,  $G_D$  link the relationship between exposure and disease. This relationship can be made independently of any other information about the complex pathways involved in the process of carcinogenesis. Alternatively, this relationship can be made considering other biomarkers, such as those of exposure, internal dose, (biologically) effective dose, or early effect, as described in detail by Schulte and Perera [12]. The

consideration of  $G_D$ , intermediate biomarkers, and exposures in association studies of disease has been referred to as a 'level crossing' model [13].

Two  $G_D$  subclasses have been the focus of much research in the genetic etiology of breast cancer, and are usually referred to as 'high penetrance' and 'low penetrance'. There are a few high penetrance  $G_D$  with allelic variants (mutations) that confer a high degree of risk to the individual, but occur infrequently in the general population. Because these alleles are rare, the proportion of breast cancer in the population that may be explained by these genotypes is low. Because of the large magnitude of effects these genotypes have on cancer risk, one hallmark of high penetrance  $G_D$  is the creation of a Mendelian (usually autosomal dominant) pattern of cancer in families. Only a small proportion (perhaps 5%) of breast cancer can be attributed to inherited single gene mutations in genes such as *BRCA1*, *BRCA2*, *TP53*, *PTEN*, and *STK11*. Even in women who have inherited these mutations, the penetrance of breast cancer may be modified by other genes or exposures [14,15].

Low penetrance  $G_D$  confer a small to moderate degree of breast cancer risk to the individual. It is expected that disease-associated alleles in these genes may be relatively common, and thus may explain a relatively larger proportion of cancer in the population. The search for low penetrance  $G_D$  generally relies on knowledge about biochemical or physiological pathways involved in breast carcinogenesis (i.e. 'candidate gene' studies). A number of study designs are available to evaluate the role of  $G_D$  in breast cancer etiology, including traditional epidemiologic methods of case-case studies, case-control studies, and cohort studies (reviewed by Thomas [16], Caporaso [17], and Langholz [18]), and family-based studies [19,20]. There are numerous examples of pathways that define candidate low penetrance  $G_D$  in breast cancer etiology. These include genes involved in the metabolism of environmental carcinogens (e.g. *CYP2D6*, *CYP2C19*, *GSTM1*, *GSTP1*, *GSTT1*, *NAT1*, *NAT2*), those involved in steroid hormone metabolism (e.g. *CYP17*, *CYP19*, *CYP1A1*, *CYP1A2*, *CYP3A4*, *COMT*), and others, including those involved in DNA damage recognition and repair. Despite the many potential candidate genes, few consistent associations between these genes and breast cancer risk have been identified. Many of these studies have been undertaken in small sample sets, and there have rarely been confirmatory analyses using independent study samples. Therefore, it is difficult to assess whether consistent and strong associations exist. While many of these genes may play a role in breast cancer etiology, there remains relatively little information about which low penetrance  $G_D$  are involved, and in which populations these effects of  $G_D$  on breast cancer risk will be observed.

### Inherited genetic determinants of treatment response and toxicity: $G_T$

$G_T$  are conceptually similar to  $G_P$  because they are involved in the dissemination and metabolism of a pharmacological agent, and thereby contribute to an agent's efficacy and toxicity. Therefore,  $G_T$  may influence the structure or amount of these agents, influence drug pharmacodynamics and metabolism, and permit the prediction of interindividual variability in the response to drug treatment, or its toxicity. A paradigm for the role of inherited genotype in dictating treatment regimen is the use of drugs that contain thiopurine, which are metabolized by thiopurine methyltransferase (TPMT) [21]. Approximately one in 300 US Caucasians carry a TPMT genotype that results in the inability to methylate thiopurine drugs, which can cause potentially fatal myelosuppression [22]. Therefore, acquiring knowledge of an individual's TPMT genotype and phenotype has become part of standard clinical practice with treatments involving thiopurines [22,23]. Other classes of enzymes, including dihydropyrimidine dehydrogenase, aldehyde dehydrogenases, glutathione S-transferases, uridine diphosphate glucuronosyl-transferases, and cytochromes P450, may also have pharmacogenetic significance in determining cancer chemotherapy regimens [24]. These include improved determination of type, timing, and dose of treatment tailored to an individual's  $G_T$ .

### Inherited genetic determinants of clinical outcome: $G_O$

The natural history of cancer progression may be influenced by  $G_O$ , which affect the tumor histopathology, including the stage or grade of disease, the rate of disease progression, or the propensity for metastasis. In contrast to the study of inherited genetic variants and clinical outcome, there is a large body of research that evaluates histopathological measures or somatic genetic mutations (e.g. loss or amplification of specific genes in tumors; estrogen or progesterone receptor positivity) with the natural history of cancer progression, and therefore clinical prognosis. A paradigm for this type of marker is that of HER2 amplification/overexpression in breast cancer prognosis [25].

Inherited (germline) mutations in candidate  $G_O$  may be associated with disease prognosis if they are involved in metabolic events that lead to tumorigenesis and progression. These events include regulation of somatic DNA damage or repair directly (via the metabolism of compounds that induce DNA damage) and by metabolism of steroid hormones that influence tumor growth. Therefore, some candidate  $G_D$  may also be candidate  $G_O$ . Furthermore, because some genes may affect both natural history (i.e.  $G_O$ ) and treatment response (i.e.  $G_T$ ), it is also important to evaluate these potentially separate or complementary effects.

## Implications for molecular epidemiology

The ability to identify consistent associations between inherited genotype and breast cancer requires that appropriate epidemiological and clinical studies be designed specifically to address these issues. Most studies that report the effects of genotype on tumor traits have not been specifically designed to study these relationships, but instead were reported in the context of case-control or cohort studies. Often, the criteria for determining the inclusion or exclusion of a case in a study are not adequately defined, or are sometimes inappropriately collected for the evaluation of natural history and prognosis. Studies are often not specifically designed to have adequate statistical power for the evaluation of these questions. In particular, prospective follow-up of a well-defined cohort of patients may be inadequate or incomplete. Evidence for the limitation of sample size of many studies is given by the many reports in which genotype was inferred to have no effect when the associated *P* values fell in the range 0.05–0.10, even though moderately large effect sizes were observed.

An important implication for chemoprevention trials is that sample size requirements will be dictated not only by the proposed reduction in cancer incidence in the treatment arm, but also by the proportion of individuals in the population who carry a particular variant genotype. These considerations could make the design and execution of chemopreventive studies considerably more difficult to achieve than if  $G_p$  were not considered.

Inconsistent inferences between studies are also common. A number of factors can be identified as contributing to these inconsistencies. A major concern is limited statistical power to identify or replicate associations in some studies. Differences in the definition and distribution of tumor stages and grades may affect study inferences. For example, differences in cancer screening practices at the time of case ascertainment may result in inconsistencies among studies. When this information is unavailable or not described, results appropriate for each population but that differ across populations may be interpreted as inconsistent findings. Studies that evaluate inherited genotype and tumor characteristics, natural history, or prognosis could benefit by reporting the distribution of screening practices and the stages or grades of tumors studied. Similarly, ethnic or exposure differences among study populations could induce apparent inconsistencies. These differences may reflect real differences across populations, and not a failure of the methodology.

Finally, multiple etiological factors should be considered simultaneously to determine whether inherited genotype provides additional information about natural history, prognosis, or treatment response that is independent of other traits including histopathological characteristics. Inherited genotypes will only have value in predicting outcome if

they provide readily accessible information beyond that routinely used and collected from these patients. Advances in these areas will improve our understanding of the multistep carcinogenesis process, and will potentially improve cancer detection, prevention, and treatment.

## References

1. Welch PL, King MC: **BRCA1 and BRCA2 and the genetics of breast and ovarian cancer.** *Hum Mol Genet* 2001, **10**:705-713.
2. Lerman C, Shields PG, Audrain J, Main D, Cobb B, Boyd NR, Caporaso N: **The role of the serotonin transporter gene in cigarette smoking.** *Cancer Epidemiol Biomarkers Prev* 1998, **7**:253-255.
3. Sabol SZ, Nelson ML, Fisher C, Gunzerath L, Brody CL, Hu S, Sirota LA, Marcus SE, Greenberg BD, Lucas FR 4th, Benjamin J, Murphy DL, Hamer DH: **A genetic association for cigarette smoking behavior.** *Health Psychol* 1999, **18**:7-13.
4. Lerman C, Caporaso NE, Audrain J, Main D, Bowman ED, Lockshin B, Boyd NR, Shields PG: **Evidence suggesting the role of specific genetic factors in cigarette smoking.** *Health Psychol* 1999, **18**:14-20.
5. Lerman C, Caporaso NE, Audrain J, Main D, Boyd NR, Shields PG: **Interacting effects of the serotonin transporter gene and neuroticism in smoking practices and nicotine dependence.** *Mol Psychiatry* 2000, **5**:189-192.
6. Feigelson HS, Coetzee GA, Kolonel LN, Ross RK, Henderson BE: **A polymorphism in the CYP17 gene increases the risk of breast cancer.** *Cancer Res* 1997, **57**:1063-1065.
7. Weston A, Pan CF, Bleiweiss IJ, Ksieski HB, Roy N, Maloney N, Wolff MS: **CYP17 genotype and breast cancer risk.** *Cancer Epidemiol Biomarkers Prev* 1998, **7**:941-944.
8. Helzlsouer KJ, Huang HY, Strickland PT, Hoffman S, Alberg AJ, Comstock GW, Bell DA: **Association between CYP17 polymorphisms and the development of breast cancer.** *Cancer Epidemiol Biomarkers Prev* 1998, **7**:945-949.
9. Haiman CA, Hankinson SE, Spiegelman D, Colditz GA, Willett WC, Speizer FE, Kelsey KT, Hunter DJ: **The relationship between a polymorphism in CYP17 with plasma hormone levels and breast cancer.** *Cancer Res* 1999, **59**:1015-1020.
10. Dehal SS, Kupfer D: **CYP2D6 catalyzes tamoxifen 4-hydroxylation in human liver.** *Cancer Res* 1997, **57**:3402-3406.
11. Jacolot F, Simon I, Dreano Y, Beaune P, Riche C and Berthou F: **Identification of the cytochrome P450 IIIA family as the enzymes involved in the N-demethylation of tamoxifen in human liver microsomes.** *Biochem Pharmacol* 1991, **41**:1911-1919.
12. Schulte PA, Perera FP (Ed): *Molecular Epidemiology: Principles and Practices.* San Diego: Academic Press; 1993.
13. Sing CF, Boerwinkle E, Turner ST: **Genetics of primary hypertension.** *Clin Exp Hypertens A* 1986, **8**:623-651.
14. Rebbeck TR, Kantoff PA, Krithivas K, Godwin AK, Daly MB, Narod SA, Garber JE, Weber BL, Brown M: **Modification of BRCA1-associated breast cancer penetrance by androgen receptor CAG repeat length variants.** *Am J Human Genet* 1999, **64**:1371-1377.
15. Rebbeck TR, Wang Y, Kantoff PW, Krithivas K, Neuhausen S, Godwin AK, Daly MB, Narod SA, Brunet J-S, Vesprini D, Garber JE, Lynch HT, Weber BL, Brown M: **Modification of BRCA1- and BRCA2-associated breast cancer risk by AIB1 genotype and reproductive history.** *Cancer Research* 2001, **61**:5420-5424.
16. Thomas DC: **Design of gene characterization studies: an overview.** *J Natl Cancer Inst Monogr* 1999, **26**:17-23.
17. Caporaso N, Rothman N, Wacholder S: **Case-control studies of common alleles and environmental factors.** *J Natl Cancer Inst Monogr* 1999, **26**:25-30.
18. Langholz B, Rothman N, Wacholder S, Thomas DC: **Cohort studies for characterizing measured genes.** *J Natl Cancer Inst Monogr* 1999, **26**:39-42.
19. Gauderman WJ, Witte JS, Thomas DC: **Family-based association studies.** *J Natl Cancer Inst Monogr* 1999, **26**:31-37.
20. Goldstein AM, Andrieu N: **Detection of interaction involving identified genes: available study designs.** *J Natl Cancer Inst Monogr* 1999, **26**:49-54.

21. Krynetski EV, Evans WE: **Genetic polymorphism of thiopurine S-methyltransferase: molecular mechanisms and clinical importance.** *Pharmacology* 2000, **61**:136-146.
22. Krynetski EY, Tai HL, Yates CR, Fessing MY, Loennechen T, Schuetz JD, Relling MV, Evans WE: **Genetic polymorphism of thiopurine S-methyltransferase: clinical importance and molecular mechanisms.** *Pharmacogenetics* 1996, **6**:279-290.
23. Otterness D, Szumlanski C, Lennard L, Klemetsdal B, Aarbakke J, Park-Hah JO, Iven H, Schmiegelow K, Branum E, O'Brien J, Weinsilbom R: **Human thiopurine methyltransferase pharmacogenetics: gene sequence polymorphisms.** *Clin Pharmacol Ther* 1997, **62**:60-73.
24. Iyer L, Ratain MJ: **Pharmacogenetics and cancer chemotherapy.** *Eur J Cancer* 1998, **34**:1493-1499.
25. Menard S, Fortis S, Castiglioni F, Agresti R, Balsari A: **HER2 as a prognostic factor in breast cancer.** *Oncology* 2001, **61**(suppl 2):67-72.