CHEKs and balances: accounting for breast cancer

Lawrence C. Brody

Genome Technology Branch, National Human Genome Research Institute, Bethesda, Maryland 20892, USA. e-mail: lbrody@helix.nih.gov

© 2002 Nature Publishing Group http://genetics.nature.com

Two highly penetrant loci have been linked to familial breast cancer (*BRCA1* and *BRCA2*), but do not explain most breast cancer familial aggregation. A new study suggests that a variant in a gene involved in DNA repair may account for some of this unexplained risk.

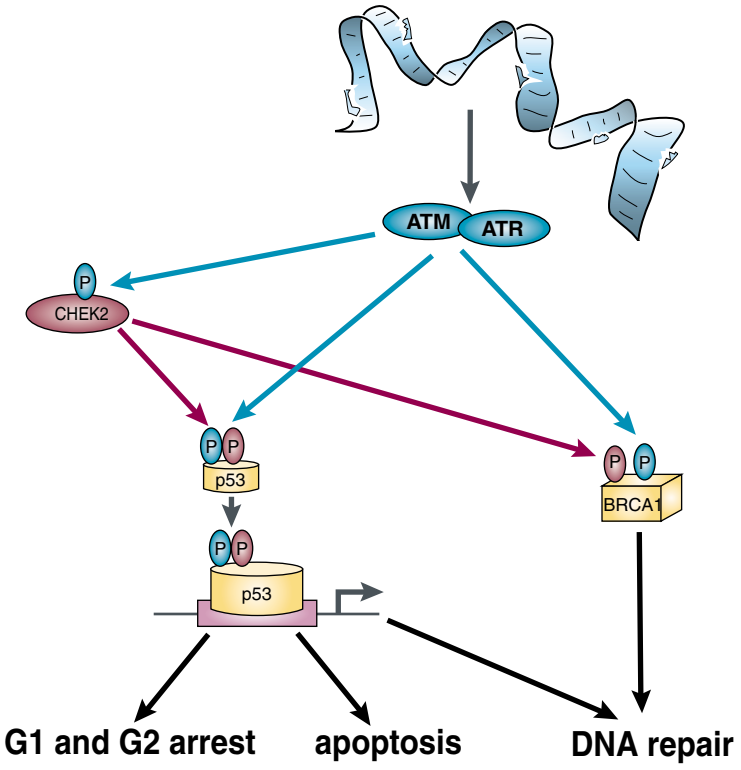
Despite differences in interpretation[1], the results from migration studies[2] (which measure the effects of environment) and twin studies[3] (which measure the effects of genes) roughly agree: genes and environment share the stage when it comes to several common cancers—colorectal, breast and prostate. This interplay has complicated the identification of genes responsible for cancer development. Whereas the active cataloging of human variation, especially single nucleotide polymorphisms (SNPs), will speed the discovery of the genetic bases for disease[4], significant hurdles need to be overcome before we can routinely screen all SNPs in the genome in the setting of a traditional case–control study. Until then, how should we proceed? On page 55 of this issue, Hanne Meijers-Heijboer and colleagues[5] provide an example of the pragmatist's approach to identifying a genetic variant that confers breast-cancer risk.

Risk variants are somewhat arbitrarily divided into low-penetrance and high-penetrance alleles. The hallmark of the latter is the segregation of cancer within families that approximates mendelian inheritance. Most mutations in *BRCA1* and *BRCA2* fall into this category and account for 3–8% of all breast cancer cases. Genes harboring high-penetrance alleles are best identified by genome-wide screens using family-based methods. In contrast, given a relatively short list of candidate loci, association studies are the method of choice to study highly prevalent, low-penetrance alleles. Meijers-Heijboer *et al.*[5] gathered families with the aim of identifying rare high-penetrance alleles linked to breast cancer risk but ended up uncovering a relatively common allele of low penetrance. By doing what was practical, rather than waiting the years that it would take to carry out the ideal study, the authors found a probable culprit.

The suspects

Linkage to *BRCA1* and *BRCA2* has been excluded in a significant fraction of 'breast cancer families'[6]. To identify additional high-penetrance loci, Meijers-Heijboer *et al.*[5] carried out genome-wide linkage studies in a collection of such families. The largest extended family in their set, EUR60 (17 cases of breast cancer), showed a hint of linkage (lod score of 1.2) to chromosome 22, in a region that contained two genes, *EP300* and *CHEK2*. Despite the inconclusive linkage result, the authors chose to screen these genes for mutations in family EUR60.

No mutations were found in *EP300*, a gene encoding a histone acetyltransferase. However, one previously described protein-truncating mutation in *CHEK2*, 1100delC[7,8], was found to segregate with breast cancer in a branch of family EUR60, although it was carried by only 8 of the 17 affected women in the family. Clearly, this mutation was not behaving as a high-penetrance allele. The authors had to gather additional evidence before concluding that *CHEK2**1100delC has a role in breast cancer.



The gene *CHEK2* is part of the DNA damage response pathway. DNA damage is sensed by the ATM and ATR kinases. The signal is transmitted via the phsophorylation of CHEK2, p53 and BRCA1. CHEK2 also directly phosphorylates p53 and BRCA1. Activated p53 induces the transcription of genes responsible for cell cycle arrest and DNA repair. In the case of extensive DNA damage, apoptotic pathways are induced. After phosphorylation by ATM and CHEK2, BRCA1 acts as a scaffold to organize various proteins involved in DNA repair. Germline mutations that increase breast cancer risk have been found in the genes that encode four of the five proteins depicted in this pathway. Mutations in these genes are also associated with the following cancer syndromes: *ATM*, ataxia-telangiectasia; *BRCA1*, inherited breast and ovarian cancer syndrome; *TP53*, Li-Fraumeni syndrome; and *CHEK2*, male and female breast cancer risk[5].

© 2002 Nature Publishing Group http://genetics.nature.com

The case

The circumstantial evidence that supports an involvement of *CHEK2* in cancer is quite convincing[9]. The cell-cycle checkpoint kinase CHEK2 is directly connected to the DNA damage response pathway (see figure), where it receives signals from the DNA damage sensors ATM/ATR (germline mutations in *ATM* have also been implicated in breast cancer[10]). Once activated, CHEK2 phosphorylates several target proteins that in turn lead to cell-cycle arrest and the activation of DNA repair pathways. In terms of inherited breast cancer, the most relevant targets of CHEK2 are p53 and BRCA1. The phosphorylation of BRCA1 on serine 988 by CHEK2 correlates with the dispersal of BRCA1 to sites of DNA repair[11]. Thus, *CHEK2* presents a biologically plausible candidate gene for breast cancer—providing one of the pillars of risk factor assessment.

But it does not stop here. To fulfill the criteria of a candidate disease gene, the variant under consideration should alter function in some way, the statistics connected to the finding should have suitably small *P* values and the finding should be replicated in an independent sample[12,13]. Do the data presented by Meijers-Heijboer *et al.*[5] provide these additional pillars? The *CHEK2* variant has a deletion of an important functional domain shown to adversely affect the activity of CHEK2 in experimental systems. Given that two genes were screened and the data analyzed by stratifying it several different ways, some modest correction to *P* values for multiple hypothesis testing would be appropriate; however, it is unlikely that one can 'correct away' the *P* values ($<10^{-6}$) reported in the paper. Pillars two and three are therefore in place.

What of pillar four—a replication set? Here is where the pragmatic approach was taken. It is unclear whether Meijers-Heijboer *et al.*[5] selected any of the six additional samples sets analyzed as *a priori* independent replication sets. Whereas a large number of samples were tested (total of 1,620 controls and 1,071 cases) and all seem to be of northern European origin, their makeup is inconsistent. Some of the sample sets are large, others small; some came with matched controls, others are matched to convenience controls; some were ascertained based on family history, others based on age of diagnosis. Given the mixed nature of the samples and ascertainment, it is reassuring to find that the carrier frequency for 1100delC was consistent between samples of the same type.

The verdict

*CHEK2**1100delC was not associated with breast cancer in cases selected without regard to family history or ones drawn from families with documented *BRCA1* and *BRCA2* mutations. The absence of an effect in *BRCA1* and *BRCA2*–positive families is consistent with biological data placing these genes in the same pathway as *CHEK2*.

An association between breast cancer and 1100delC was found only for *BRCA1/BRCA2*–negative families compared with controls. As the authors discuss, these results suggest that 1100delC may act as a low-penetrance allele only in the context of a positive family history, possibly in epistasis with variants in as-yet-unidentified genes. Only a small number (<10%) of controls were drawn from married-in individuals selected along with the cases. Therefore, an alternative explanation—that the selection of families has introduced some confounding variable—cannot be ruled out.

Confirmation of these findings in additional samples will address the concern of selection bias and help to determine the relative risk associated with carrying *CHEK2**1100delC. Likewise, additional studies will be required before the accuracy of the authors' estimate that 1% of female and 9% of male breast cancer is attributable to this allele. This study also highlights the challenge of translating the finding of a positive gene association to the clinical arena. Precisely because it is a low penetrance allele, presymptomatic testing for *CHEK2**1100delC will have little positive or negative predictive value. Unless there are populations that carry this variant at a high frequency, evaluating the role of *CHEK2**1100delC will remain difficult. Nevertheless, the work of Meijers-Heijboer *et al.*[5] serves as a positive example of how the application of genomics, genetic epidemiology and molecular pathogenesis can contribute to the understanding of human disease. □

1. Hoover, R.N. *N. Engl. J. Med.* **343**, 135–136 (2000).
2. Ziegler, R.G. *J. Natl Cancer Inst.* **85**, 1819–1827 (1993).
3. Lichtenstein, P. *et al. N. Engl. J. Med.* **343**, 78–85 (2000).
4. Collins, F.S. *et al. Science* **282**, 682–689 (1998).
5. Meijers-Heijboer, H. *et al. Nature Genet.* **31**, 55–59 (2002).
6. Easton, D.F. *Breast Cancer Res.* **1**, 14–17 (1999).
7. Wu, X. *et al. J. Biol. Chem.* **276**, 2971–2974 (2001).
8. Lee, SB. *et al. Cancer Res.* **61**, 8062-8067 (2001).
9. Bartek, J. *et al. Nature Rev. Mol. Cell. Biol.* **2**, 877–869 (2001).
10. Lee, J.S. *et al. Nature* **404**, 201–204 (2000).
11. Chenevix-Trench, G. *et al. J. Natl Cancer Inst.* **94**, 205–215 (2002).
12. Editorial, *Nature Genet.* **22**, 1–2 (1999).
13. Dahlman, I. *et al. Nature Genet.* **30**, 149–150 (2002).