

Commentary

Genes Involved in Breast Cancer Progression

Analysis of Global Changes in Gene Expression or Retroviral Tagging?

Emmett V. Schmidt

From the Massachusetts General Hospital Cancer Center, Charlestown, and the Pediatric Service, Massachusetts General Hospital, Boston, Massachusetts

The simple observation that breast cancer develops later in life has long suggested that more than one genetic event is involved in its development. The genetic simplicity of childhood tumors provides an interesting contrast. Loss of a single oncogene such as the retinoblastoma gene product (pRb) is sufficient to cause childhood eye tumors. In contrast, although pRb loss is a frequent contributor in sporadic breast cancer, pRb loss is apparently insufficient to completely induce breast cancer.¹ These simple observations have long been the basis for the idea that more than one cancer gene is involved in the final tumor's development. By inference, this idea implies that a preneoplastic stage of cancer will lack those genes that drive the final step in cancer development.

The field of molecular oncology is now graduating from an era of gene discovery into an era in which we will seek to determine how cancer genes interact. Several new methodologies promise to help identify interacting events. Array-based approaches show promise, but lack the power to simultaneously demonstrate the genetic function of the proposed interactions. Proviral tagging initiated the discovery of oncogenes so it should not be surprising that it may have the potential to contribute to the next analyses of breast cancer by simultaneously identifying interactors and providing genetic evidence that they interact.

Histological evidence suggests that several preneoplastic states exist that precede invasive breast tumors.² Such lesions provide additional support to the idea that more than one event contributes to the final cancers. These histological lesions include atypical ductular hyperplasia, atypical lobular hyperplasia, ductal carcinoma *in situ* (DCIS), and lobular carcinoma *in situ*. These lesions are thought to fall on a histological continuum between the terminal duct lobular units from which breast cancers

arise, and the final invasive breast cancer. They are thought to be preneoplastic lesions because they are less common in specimens from patients not suffering from breast cancer, and their presence identifies a higher risk for subsequent development of breast cancer. An alternative view might hold that these lesions simply represent surrogate markers that reflect cancer predisposition within a breast. In this alternative view, the lesions themselves are not on a direct continuum between normal and invasive cells.

A direct test of the relationship between preneoplastic lesions and the ultimate tumor awaits the development of experimental systems that can be genetically manipulated to induce a transition from a preneoplastic state into a neoplastic state by addition or subtraction of genetic information. Several models could be considered to explain the genetic differences between preneoplasia and neoplasia (Figure 1). Although stochastic addition of additional genetic events remains the dominant model in cancer biology, other situations may be found as experimental paradigms change. Clinically, identification of mechanisms by which preneoplastic lesions transform to invasive tumors holds the promise that we might treat lesions in a gene-specific manner to prevent development of the invasive phenotype at a point before the tumor has acquired its final genetic diversity and invasive qualities. Such a possibility is approaching reality considering the success trials of tamoxifen in patients with a high risk to develop breast cancer.³

Recent technological developments have rejuvenated the effort to clarify the relationship between preneoplasia and cancer. These experiments are based on the assumption that genes associated with the change from normal to preneoplasia will be found again in the neoplastic tissue, if the lesions are genetically related. The combinations of laser capture microdissection, compar-

Accepted for publication October 1, 2002.

Address reprint requests to Emmett V. Schmidt, M.D., Ph.D., Department of Pediatrics, Massachusetts General Hospital, 13th St., Building 149, 7th Fl., Charlestown, MA 02129. E-mail: schmidt@helix.mgh.harvard.edu.

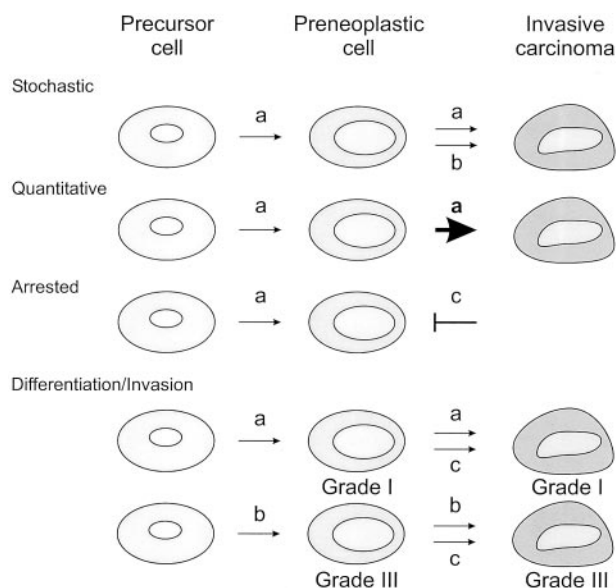


Figure 1. Model for genetic events that accompany the transition from precursor to invasive tumor. The genetic interrelationships between genes associated with progression to both preneoplasia and neoplasia are depicted (a, b, c). Several models can be contemplated including the current paradigm that stepwise stochastic events govern neoplastic progression.

ative genomic hybridization, and array analyses have started to reveal these genetic connections. Hints that a field of genetic loss might give rise to preneoplasia came from an analysis demonstrating loss of heterozygosity in normal tissue adjacent to breast carcinomas.⁴ In a later study, simple microdissection was used to look at loss of heterozygosity at 15 loci in 399 precursor lesions. This study revealed synchronous changes in 45% of atypical ductular hyperplasia, 77% of noncomedo DCIS, and 80% of comedo DCIS.⁵ The identification of synchronous changes would suggest that gene a in the stochastic model (Figure 1) indeed exists and persists through to the final tumor. Laser capture microdissection was also used to isolate precursor lesions in a subsequent study to evaluate genomic instability in precursor lesions.⁶ This study confirmed the existence of progressive chromosomal instability during tumor progression. However, loss of heterozygosity at some markers was observed only in the component of DCIS but not in the invasive component of the same tumors. Such a finding also adds the caveat that putative preneoplastic lesions might actually represent arrested tumor development in which the first hit is accompanied by a second hit that stops the tumor along the pathway to cancer, a possibility rarely considered in these analyses (the arrested preneoplasia model in Figure 1). In this model, not all genetic changes in a preneoplastic cell are necessarily causing the preneoplasia.

Comparative genomic hybridization has been similarly applied to the analysis of precursor lesions. However, the genetic interrelationships became significantly more complex when additional attention was paid to the stage of differentiation of the precursor lesions and the accompanying tumors.^{7,8} These studies suggested the existence of quite separable first events that were then accompanied by a shared event in the transition to cancer

(Figure 1, differentiation/invasion model). In such a model the shared event is not the first event. This possibility should be kept in mind when considering the stochastic model of cancer. The comparative genomic hybridization studies specifically revealed conserved chromosomal loss from precursor to invasive lesion, but only if tumors of similar differentiation status were compared. For example, well-differentiated DCIS shared genetic changes with tubular and tubulo-lobular carcinoma, whereas poorly differentiated DCIS shared changes with grade 3 ductal invasive carcinoma. Loss of 16q and a low average number of aberrations per case were seen in highly differentiated invasive breast cancers and their precursors. These losses of 16q were not seen in high-grade lesions or their precursors. This more complex data set suggests that the genetic lesions that drive the first step to preneoplasia might also drive the state of differentiation of the tumor and that the acquisition of an invasive phenotype might be caused by genes shared across the pathways leading to either poorly differentiated or highly differentiated phenotypes.

The plethora of recent studies describing array analyses of breast tumors suggests the potential power of transcriptome analysis.⁹⁻¹¹ These studies have confined their analyses to bulk tumors in which factors including tissue invasion by inflammatory cells often creeps into the analysis. Given the microscopic size of breast cancer precursor lesions, a true analysis of the transcriptome changes accompanying breast cancer progression awaits the application of microarrays to laser capture-microdissected lesions. A first proof of the potential power of this tool was recently published and further applications to larger numbers of matched normal, preneoplastic, and neoplastic samples from the same patient can be anticipated.¹²

With human samples, the investigator is caught by the problem that 'there are lies, damn lies, and then there are statistics.' What about animal models? The transgenic model that received the first patent for genetically engineered mice that develop cancer also provided an initial demonstration of the complexity of breast cancer in mice.¹³ Despite the provision of a single oncogene, the *myc* oncogene, in every cell in the developing mammary gland, the target tissue initially remained normal and the mice stochastically developed mammary adenocarcinomas as they aged. Classically, this was explained as owing to the need for a second collaborating event to complement the transgenic *myc* gene, or perhaps a third when *myc* mice and *ras* mice were first bred.¹⁴ Retroviral tagging had originally revealed the world of oncogenes to us, so it did not take long to use the mouse mammary tumor virus (MMTV) to infect transgenic mice that had already been seeded with the first hit. The goal was to determine whether another random hit by insertion mutagenesis could tell us about the nature of interacting genes in a model in which hyperplasia took quite some time to evolve to neoplasia.^{15,16} *Int-1* and *int-2* were thereby connected as collaborating events.

The technique of retroviral tagging has generally been more advanced in mouse models of leukemia, perhaps because the genes identified by MMTV mutagenesis

have been less clearly linked to breast cancer and because MMTV-induced tumors differ histologically from human tumors.¹⁷ Three recent studies have again asserted the power of retroviral tagging, in the leukemia world at least, in an era in which the mouse genome sequence is now available.¹⁸⁻²⁰ The most sophisticated study used mice that expressed a myc transgene, but had the usual genetic target of Moloney virus insertions, pim-1, knocked out. These mice were infected by Maloney virus to try to activate those genes that could substitute for Pim-1 and collaborate with *c-myc*.¹⁹ The insertion sites were then directly sequenced by a splinkerette-based polymerase chain reaction system. The key to the power of this study was the availability of the mouse genomic sequence. The sequences tagged by the retroviral insertions immediately revealed the site of insertion and a list of 52 candidate genes that could substitute for pim-1 collaboration in myc-induced lymphomas emerged. The list is not unlike many such lists. It contains enough known oncogenes to suggest that the unknown genes are worth further characterization.

Two additional studies confirmed the power of this approach. The second study simply applied the new sophisticated tools of inverse polymerase chain reaction, coupled with a comparison to the Celera genome sequence, to identify insertion sites in B cell tumors arising in AKXD and NFS.V mice.²⁰ The 149 targeted genes fell into the ras, Notch, Jak/Stat, and nuclear factor- κ b pathways. The third study started with mice null in the Cdkn2 locus that lacked both p16^{INK4a} and p19^{arf} and again infected them with Maloney murine leukemia virus. This approach again improved the genetic power of the analysis by starting with the deletion of an entire locus involved in both p53 and cyclin D/Cdk4-dependent signaling to ask what events might fall in collaborating pathways. Some old standby insertion sites were found again including myc, pim-1, and bmi-1. Insertion into ras and MAP kinase pathway members constituted the bulk of the remainder. At this point, we do not know to what extent the pathways activated by these retroviral insertions are indeed specific to the proposed collaborations and how much of their spectrum is because of any specificities of the retroviruses used. However, the accumulated list of oncogenes is likely to keep investigators busy for some time.

MMTV led the way into the world of oncogenes, but it has always been less tractable than the murine leukemia viruses. The cloning of an infectious MMTV that could be engineered *in vivo* provided a critical first step in using MMTV in mutagenesis.²¹ Dr. Greg Shackleford and colleagues,¹⁵ then working in the laboratory of Dr. Harold Varmus, used the MMTV infectious clone to infect Wnt-1 transgenic mice to demonstrate the expected collaboration with int-2/fgf-3 and hst-1. Further work on the novel insertion sites identified by insertional mutagenesis of the Wnt-1 transgenics revealed a new fgf, fgf-8.^{22,23} The identification of the sequences in which the retrovirus has inserted provided the main difficulty in all of these studies. To circumvent this problem, Dr. Shackleford's laboratory²⁴ therefore developed a retrovirus containing a bacterial supF to allow direct selection of insertional tar-

gets once they were cloned from tumor DNA. This virus is equipotent in tumor induction and indeed provided a significant advantage in isolating insertion sequences by this rapid cloning technique before the mouse genomic sequence became available. My laboratory's experience with this technique should, however, send a warning signal to those embarking on this kind of study. We originally showed that cyclin D1 is probably the significant breast cancer gene located at 11q13.²⁵ We infected MMTV-cyclin D1 mice with the supF-containing MMTV in hopes of identifying those genes that would interact with cyclin D1 in tumorigenesis. The MMTV-cyclin D1 mice have some significant advantages as a model to study neoplastic progression because they exhibit a prolonged hyperplastic state for several months before tumor development. However, cyclin D1 is also a key molecule that appears to be downstream of the function of a large number of genes including erbB2, and ras.²⁶ Infection of MMTV-cyclin D1 mice by the Shackleford MMTV produced no acceleration of tumorigenesis.²⁷ Although this might result from an inefficient infection of the mice, the presence of multiple new exogenous proviruses in the majority of tumors argued against it. More likely, this result reflects the intimate involvement of cyclin D1 in the pathways of most or all of the targets for MMTV insertion mutations including both Wnts and Fgfs.^{28,29}

In this issue of *The American Journal of Pathology*, Chatterjee and colleagues³⁰ have now applied insertion mutagenesis to identify genes collaborating with p53 in murine mammary adenocarcinomas. This work, from the laboratory of Dr. Archibald Perkins, started with their earlier demonstration that a dominant oncogenic version of p53 could collaborate with erbB2 in mammary tumorigenesis.³¹ In that initial publication they experienced the same frustration as my own laboratory in that MMTV-erbB2 did not collaborate with MMTV infection. Because erbB2 is upstream of cyclin D1, the similarity to our failure to demonstrate MMTV interactions with cyclin D1 may be for similar reasons. By this reasoning, erbB2 may also similarly function to regulate most of the genetic pathways that are targeted by MMTV, all of which function to regulate the cyclin D1/Cdk4/pRb pathway. In this "silk purse out of a sow's ear" interpretation of two negative results, the erbB2 > ras > MAP kinase pathway provides the same kind of oncogenic signal as Wnt-1 and the Fgf family that are usually targeted by MMTV insertions so MMTV infection adds nothing to the genetic abnormalities already present in the tumors. Moreover, it would suggest that p53 activation might be the second event needed for tumorigenesis by all of them. The logical next step was therefore to infect the MMTV-p53^{172H} mice with MMTV.

Perseverance in the face of unanticipated difficulty provides an additional satisfying element of the current study. Rather than using the infectious clone of MMTV, these authors simply crossed their mice to C3H/He, which will result in infection of the progeny by the milk-borne MMTV of the C3H/He mice. This approach immediately complicated matters because of strain differences in the mice. The results of the initial breeding revealed an F1 shortening of tumor latency compared with the parental FVB-based p53 result. However, tumor latency was

not different between the infected transgenics and the infected nontransgenics, a huge problem if one wants to identify collaborating oncogenes. The authors chose to forge ahead based on the consideration that a hidden suppressive effect of the FVB strain itself might be confounding their results. Indeed, one generation later the p53-expressing infected mice developed tumors more rapidly than the MMTV-infected strain alone. The existence of a mammary adenocarcinoma-suppressive effect in the FVB strain of mice is an important result in its own right, because this is the standard strain for most transgenic experiments. Nevertheless, the important result lies in the p53-interactors revealed by the MMTV insertions.

The infection scheme used by Chatterjee and colleagues³⁰ should identify genes that can complement p53 function. By their earlier failure to accelerate MMTV-erbB2 with MMTV infection, some of the insertions in the p53 strain should therefore still be the usual MMTV insertion sites. Satisfyingly, Wnt-1, Wnt3, Wnt10b, Fgf4, and Fgf8 indeed reappeared in this study. More satisfyingly, a new set of insertion sites appeared. A Blast search of the public databases gave them an intriguing list of their own. Their insertion sites included CD44, a novel member of the cadherin family, a number of new transcription factors, as well as members of both the Notch and nuclear factor- κ B signaling pathways. They discuss the implications of these genes in their own article.

Where will these studies go next? Although retroviral insertions provide an immediate test of function implied by their acceleration of tumorigenesis, their relevance to human cancer remains an open question. The Annapolis meeting of mouse and human pathologists¹⁷ recently concluded that features of mammary adenocarcinomas caused by MMTV-driven transgenes expressing the actual human genes involved in breast cancer were histologically close to the human diseases they sought to model. However, fundamental differences exist. For example, mice seem particularly susceptible to loss of the alternative reading frame of the Cdkn2 locus that encodes p19^{ARF} in mice and p14^{ARF} in humans.³² Human tumors far more frequently exhibit abnormalities of the p16^{INK4A} transcript in that same locus.³³ This may well represent a fundamental difference in the oncogenic wiring in the two species. Thus, as we move forward to identify the genetic interactions that govern the second or third steps in breast tumor progression, a method that continuously checks for conservation between mouse and human, but that uses the genetic power of the mouse, may help us to sharpen our focus on members of the long lists of potential oncogenic events emerging from both array analyses and insertional mutagenesis studies. The prize should be to identify those events that are significant in the human disease.

References

1. Lee EY, To H, Shew JY, Bookstein R, Scully P, Lee WH: Inactivation of the retinoblastoma susceptibility gene in human breast cancers. *Science* 1988, 241:218–221
2. Allred DC, Mohsin SK, Fuqua SA: Histological and biological evolution of human premalignant breast disease. *Endocr Relat Cancer* 2001, 8:47–61
3. Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, Vogel V, Robidoux A, Dimitrov N, Atkins J, Daly M, Wieand S, Tan-Chiu E, Ford L, Wolmark N: Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst* 1998, 90:1371–1388
4. Deng G, Lu Y, Zlotnikov G, Thor AD, Smith HS: Loss of heterozygosity in normal tissue adjacent to breast carcinomas. *Science* 1996, 274:2057–2059
5. O'Connell P, Pekkel V, Fuqua SA, Osborne CK, Clark GM, Allred DC: Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci. *J Natl Cancer Inst* 1998, 90:697–703
6. Shen CY, Yu JC, Lo YL, Kuo CH, Yue CT, Jou YS, Huang CS, Lung JC, Wu CW: Genome-wide search for loss of heterozygosity using laser capture microdissected tissue of breast carcinoma: an implication for mutator phenotype and breast cancer pathogenesis. *Cancer Res* 2000, 60:3884–3892
7. Buerger H, Otterbach F, Simon R, Poremba C, Diallo R, Decker T, Riethdorf L, Brinkschmidt C, Dockhorn-Dworniczak B, Boecker W: Comparative genomic hybridization of ductal carcinoma in situ of the breast—evidence of multiple genetic pathways. *J Pathol* 1999, 187:396–402
8. Buerger H, Otterbach F, Simon R, Schafer KL, Poremba C, Diallo R, Brinkschmidt C, Dockhorn-Dworniczak B, Boecker W: Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes. *J Pathol* 1999, 189:521–526
9. Eisen MB, Spellman PT, Brown PO, Botstein D: Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 1998, 95:14863–14868
10. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D: Molecular portraits of human breast tumours. *Nature* 2000, 406:747–752
11. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH: Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002, 415:530–536
12. Sgroi DC, Teng S, Robinson G, LeVangie R, Hudson Jr JR, Elkhoulou AG: In vivo gene expression profile analysis of human breast cancer progression. *Cancer Res* 1999, 59:5656–5661
13. Stewart TA, Pattengale PK, Leder P: Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. *Cell* 1984, 38:627–637
14. Sinn E, Muller W, Pattengale P, Tepler I, Wallace R, Leder P: Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo. *Cell* 1987, 49:465–475
15. Shackelford GM, MacArthur CA, Kwan HC, Varmus HE: Mouse mammary tumor virus infection accelerates mammary carcinogenesis in Wnt-1 transgenic mice by insertional activation of int-2/Fgf-3 and hst/Fgf-4. *Proc Natl Acad Sci USA* 1993, 90:740–744
16. Tsukamoto AS, Grosschedl R, Guzman RC, Parslow T, Varmus HE: Expression of the int-1 gene in transgenic mice is associated with mammary gland hyperplasia and adenocarcinomas in male and female mice. *Cell* 1988, 55:619–625
17. Cardiff RD, Anver MR, Gusterson BA, Hennighausen L, Jensen RA, Merino MJ, Rehm S, Russo J, Tavassoli FA, Wakefield LM, Ward JM, Green JE: The mammary pathology of genetically engineered mice: the consensus report and recommendations from the Annapolis meeting. *Oncogene* 2000, 19:968–988
18. Lund AH, Turner G, Trubetskoy A, Verhoeven E, Wientjens E, Hulsman D, Russell R, DePinho RA, Lenz J, Van Lohuizen M: Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. *Nat Genet* 2002, 32:160–165
19. Mikkers H, Allen J, Knipscheer P, Romeyn L, Hart A, Vink E, Berns A: High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. *Nat Genet* 2002, 32:153–159
20. Suzuki T, Shen H, Akagi K, Morse HC, Malley JD, Naiman DQ, Jenkins NA, Copeland NG: New genes involved in cancer identified by retroviral tagging. *Nat Genet* 2002, 32:166–174
21. Shackelford GM, Varmus HE: Construction of a clonable, infectious,

- and tumorigenic mouse mammary tumor virus provirus and a derivative genetic vector. *Proc Natl Acad Sci USA* 1988, 85:9655–9659
22. Kapoun AM, Shackleford GM: Preferential activation of Fgf8 by proviral insertion in mammary tumors of Wnt1 transgenic mice. *Oncogene* 1997, 14:2985–2989
 23. MacArthur CA, Shankar DB, Shackleford GM: Fgf-8, activated by proviral insertion, cooperates with the Wnt-1 transgene in murine mammary tumorigenesis. *J Virol* 1995, 69:2501–2507
 24. Jiang Z, Shackleford GM: Mouse mammary tumor virus carrying a bacterial supF gene has wild-type pathogenicity and enables rapid isolation of proviral integration sites. *J Virol* 1999, 73:9810–9815
 25. Wang TC, Cardiff RD, Zukerberg L, Lees E, Arnold A, Schmidt EV: Mammary hyperplasia and carcinoma in MMTV-cyclin D1 transgenic mice. *Nature* 1994, 369:669–671
 26. Lee RJ, Albanese C, Fu M, D'Amico M, Lin B, Watanabe G, Haines III GK, Siegel PM, Hung MC, Yarden Y, Horowitz JM, Muller WJ, Pestell RG: Cyclin D1 is required for transformation by activated Neu and is induced through an E2F-dependent signaling pathway. *Mol Cell Biol* 2000, 20:672–683
 27. Gadd M, Pisc C, Branda J, Ionescu-Tiba V, Nikolic Z, Yang C, Wang T, Shackleford GM, Cardiff RD, Schmidt EV: Regulation of cyclin D1 and p16(INK4A) is critical for growth arrest during mammary involution. *Cancer Res* 2001, 61:8811–8819
 28. Boudreau N, Andrews C, Srebrow A, Ravanpay A, Cheresh DA: Induction of the angiogenic phenotype by Hox D3. *J Cell Biol* 1997, 139:257–264
 29. Rimerman RA, Gellert-Randleman A, Diehl JA: Wnt1 and MEK1 cooperate to promote cyclin D1 accumulation and cellular transformation. *J Biol Chem* 2000, 275:14736–14742
 30. Chatterjee G, Rosner A, Han Y, Zelazny ET, Li B, Cardiff RD, Perkins AS: Acceleration of mouse mammary tumor induced-induced murine mammary tumorigenesis by a p53^{172H} transgene: influence of FVB background on tumor latency and identification of novel sites of proviral insertion. *Am J Pathol* 2002, 161:2241–2253
 31. Zelazny E, Li B, Anagnostopoulos AM, Coleman A, Perkins AS: Co-operating oncogenic events in murine mammary tumorigenesis: assessment of ErbB2, mutant p53, and mouse mammary tumor virus. *Exp Mol Pathol* 2001, 70:183–193
 32. Kamijo T, Zindy F, Roussel MF, Quelle DE, Downing JR, Ashmun RA, Grosveld G, Sherr CJ: Tumor suppression at the mouse INK4a locus mediated by the alternative reading frame product p19ARF. *Cell* 1997, 91:649–659
 33. Quelle DE, Cheng M, Ashmun RA, Sherr CJ: Cancer-associated mutations at the INK4a locus cancel cell cycle arrest by p16INK4a but not by the alternative reading frame protein p19ARF. *Proc Natl Acad Sci USA* 1997, 94:669–673