Biochemical and Biophysical Research Communications **270,** 171–175 (2000)
doi:10.1006/bbrc.2000.2392, available online at http://www.idealibrary.com on IDEAL®

Nuclear Localization Signals of the BRCA2 Protein

Ken-ichi Yano, Keiko Morotomi, Hiroko Saito, Masahiro Kato, Fumie Matsuo, and Yoshio Miki[1]
Department of Molecular Diagnosis, Cancer Institute, Japanese Foundation for Cancer Research, 1-37-1 Kami-ikebukuro, Toshima-ku, Tokyo 170-8455, Japan

Received March 1, 2000

***BRCA2* is a tumor suppressor gene whose germline mutations increase the lifetime risk of breast cancer. *BRCA2* encodes a large nuclear protein involved in DNA repair, but the location of its functional domain has been unclear. Here, we report nuclear localization signals (NLSs) of the BRCA2 protein. By expressing various portions of the BRCA2 protein tagged with enhanced green fluorescent protein in HeLa cells, we show that the C-terminal domain is necessary for nuclear localization. Two regions in the C-terminal domain were identified with functional NLSs by site-directed mutagenesis analyses. The NLSs locate between the germline mutation found in the most downstream position and the polymorphic stop codon, suggesting that defects in the proper nuclear transport of the BRCA2 protein are causative of carcinogenesis. Our data thus provide a possible explanation for the high frequency of frame-shift and nonsense mutations in *BRCA2* of hereditary breast cancer patients.**

BRCA2 is a breast cancer susceptibility gene whose germline mutations can lead to high lifetime risk of breast cancer in both females and males (1, 2). More than 100 independent germline mutations in *BRCA2* have been reported in families with a high incidence of breast cancer (2). Most of these mutations are frameshift or nonsense mutations that result in premature termination of translation (2). The *BRCA2* gene encodes a 3418-amino acid protein (1) containing eight short repetitive sequences, termed BRC repeats (3), that interact with the Rad51 protein encoded by a eukaryotic homolog of the *recA* gene of *Escherichia coli* (4). Upon DNA damage, the BRCA2 protein coexists with the Rad51 and BRCA1 proteins as nuclear foci (5). Targeted disruption of *BRCA2* in mice has been reported to result in hypersensitivity to ionizing radiation and frequent abnormalities of chromosomes (6). Taken together, these observations indicate that *BRCA2* is involved in the repair of DNA double-strand breaks and plays an important role in the maintenance of genomic integrity (7).

Although information on the structure-function relationship of the BRCA2 protein is important to understand the molecular basis of *BRCA2*-associated carcinogenesis, no distinct functional domains, except the BRC repeats, have yet been defined. One reason for this is that the BRCA2 protein has no significant sequence similarity to any other known proteins, and another is that the protein's unusually large size hampers the biochemical analysis of its biological function (1).

Recently, several groups have reported that the BRCA2 protein is localized in the nucleus, which location is consistent with its role in DNA repair (8, 9). The transport of proteins into organella is a strictly regulated process, and proteins destined for the nucleus carry specific amino acid sequences for nuclear import known as nuclear localization signals (NLSs) (10). In this study, we searched for NLSs of the BRCA2 protein. By expressing various portions of the BRCA2 protein tagged with enhanced green fluorescent protein (EGFP), we identified NLSs in the C-terminal domain of the BRCA2 protein. We then demonstrated that bacterial β-galactosidase fused to the NLS region of the BRCA2 protein could enter the nucleus. Finally, we proposed that disruption of proper subcellular distribution may be closely related to *BRCA2*-associated carcinogenesis.

MATERIALS AND METHODS

Plasmid constructions. Various portions of the *BRCA2* cDNA were introduced into the multicloning sites of the pEGFP vectors (Clonetech, California). Site-directed mutageneses of the plasmids were performed by using a QuickChange mutagenesis kit (Stratagene, California) according to the supplier's protocol. To construct the plasmid expressing a bacterial β-galactosidase fused to NLS, the region of *BRCA2* cDNA corresponding to amino acids 3258 to 3319 was amplified by PCR and inserted into pcDNA3.1HisB-lacZ (Invitrogen, Netherlands) at the BsiWI site. The integrities of all plasmids were confirmed by sequencing.

Cell culture and transient tranfections. HeLa cells were maintained in Dulbecco's modified Eagle's medium supplemented with

[1] To whom correspondence and reprint requests should be addressed. Fax: +81-3-5394-4035. E-mail: yosmiki@ims.u-tokyo.ac.jp.

0006-291X/00 $35.00

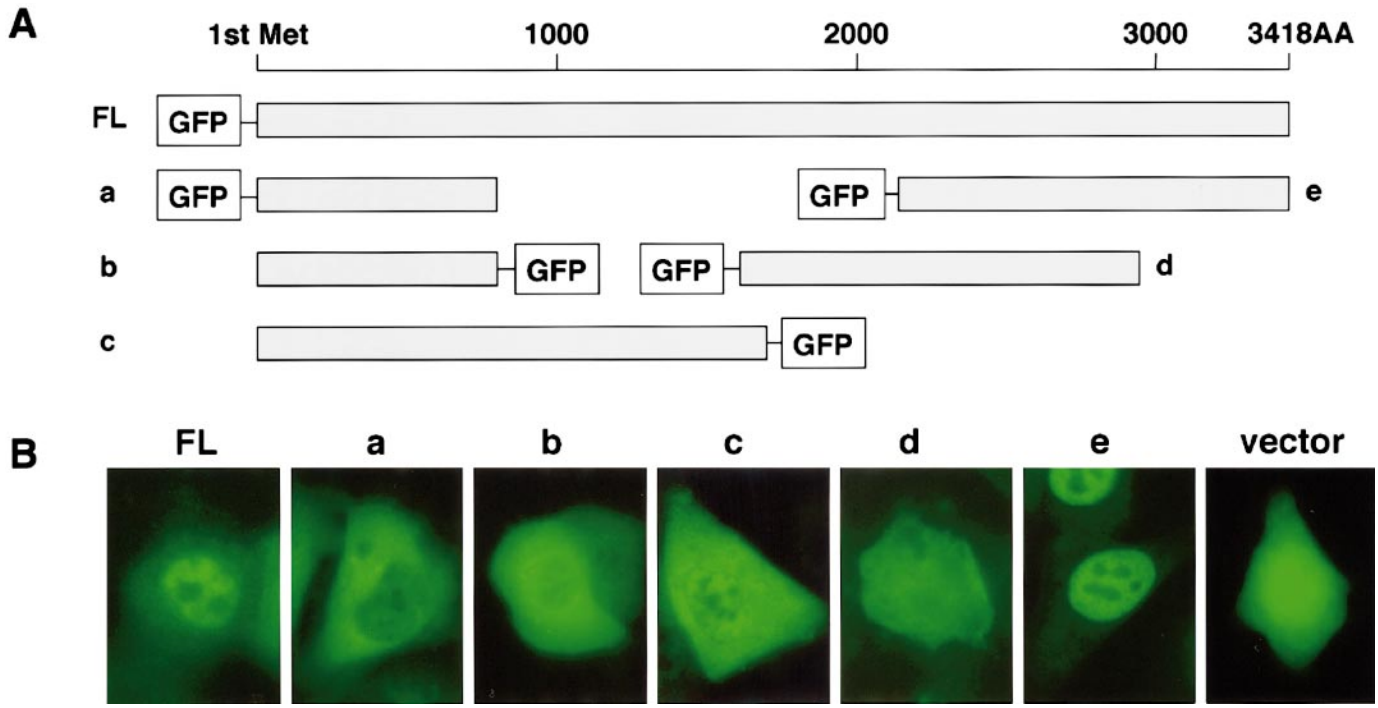

FIG. 1. Identification of a domain in the BRCA2 protein required for nuclear localization. (A) Schematic representation of the EGFP-BRCA2 fusion proteins. Various portions of the *BRCA2* cDNA were cloned into pEGFP vectors to express the EGFP-BRCA2 fusion proteins under the control of CMV promoter. Each fusion protein contains a fragment of the BRCA2 protein as follows: FL, amino acids 1-3418; a and b, 1-815; c, 1-1761; d, 1596-2940; e, 2126-3418. (B) Subcellular distribution of the EGFP-BRCA2 fusion proteins in the HeLa cell. The EGFP-BRCA2 fusion proteins were expressed transiently in the HeLa cell. Fluorescence of the fusion proteins was directly observed by an Axioplan2 microscope (Zeiss) equipped with a Quips Smart Capture System (Vysis). Representative micrographs are shown.

10% fetal bovine serum. For transfections, cells were grown on culture slides (Becton Dickinson, New Jersey) to subconfluency. Transfection was performed by using the TransIT lipofection reagent according to the manufacturer's instructions (Mirus, Wisconsin). After 2 hr, the cells were washed once with PBS, and cultured in fresh medium for 24 hr. The transfected cells were washed twice with PBS, fixed with 4% paraformaldehyde, and then subjected to fluorescent microscopy.

Observation of subcellular distributions of fusion proteins. Subcellular distributions of EGFP-BRCA2 fusion proteins were observed directly under an Axioplan-2 microscope (Carl Zeiss, Germany) equipped with a Quips Smart Capture System (Vysis, Illinois). Localization of β-galactosidase-NLS fusion protein was detected by immunofluorescent microscopy (11). Briefly, transfected cells were fixed as above and permealized with 0.2% Triton X in PBS for 5 min. The permealized cells were blocked with 2% bovine serum albumin in PBS for 30 min, and then reacted with anti-Xpress antibody (2 μg/ml; Invitrogen) for 1 hr. After washing three times with PBS containing 0.05% Tween 20, the cells were incubated with anti-mouse IgG coupled to fluorescein isothiocyanate (Santa Cruz, California) for 30 min. After a final five washings with PBS containing 0.05% Tween 20, fluorescence was observed as described above.

RESULTS

To identify the NLS in the BRCA2 protein, we constructed a series of plasmids that expressed the various portions of the BRCA2 protein fused to EGFP (Fig. 1A). These fusion proteins were transiently expressed in the HeLa cell, and their subcellular distribution was observed by fluorescent microscopy. As shown in Fig. 1B, the N-terminal and central domains of the BRCA2 protein fused to EGFP were dispersed throughout the cytoplasm and nucleus. On the other hand, the fusion protein of the C-terminal domain of the BRCA2 protein (Fig. 1B, panel e) was able to enter nucleus, demonstrating that this domain should contain NLSs.

In the amino acid sequence of this C-terminal region, we found short arrays of basic amino acids homologous to a classic basic NLS at four positions (Fig. 2A). To examine whether these sequences are required for nuclear localization, we carried out a site-directed mutagenesis of each basic sequence. As shown in Fig. 2A, two EGFP-BRCA2 fusion proteins carrying mutations of KK2946NQ and KR3383NG, respectively, were found to localize predominantly in the nucleus, indicating that these regions are not necessary for nuclear localization. In contrast, significant impairment of the nuclear localization was observed with EGFP-BRCA2 proteins carrying the mutations of KR3267NG and KK3314NQ, respectively (Fig. 2A, panels c and d). When both basic sequences are disrupted (KR3267NG and KK3314NQ), the cytoplasmic distribution of the mutant protein was significantly enhanced (Fig. 2A, panel f). This result indicated that the arrays of basic amino acids in these two proximal positions are indis-

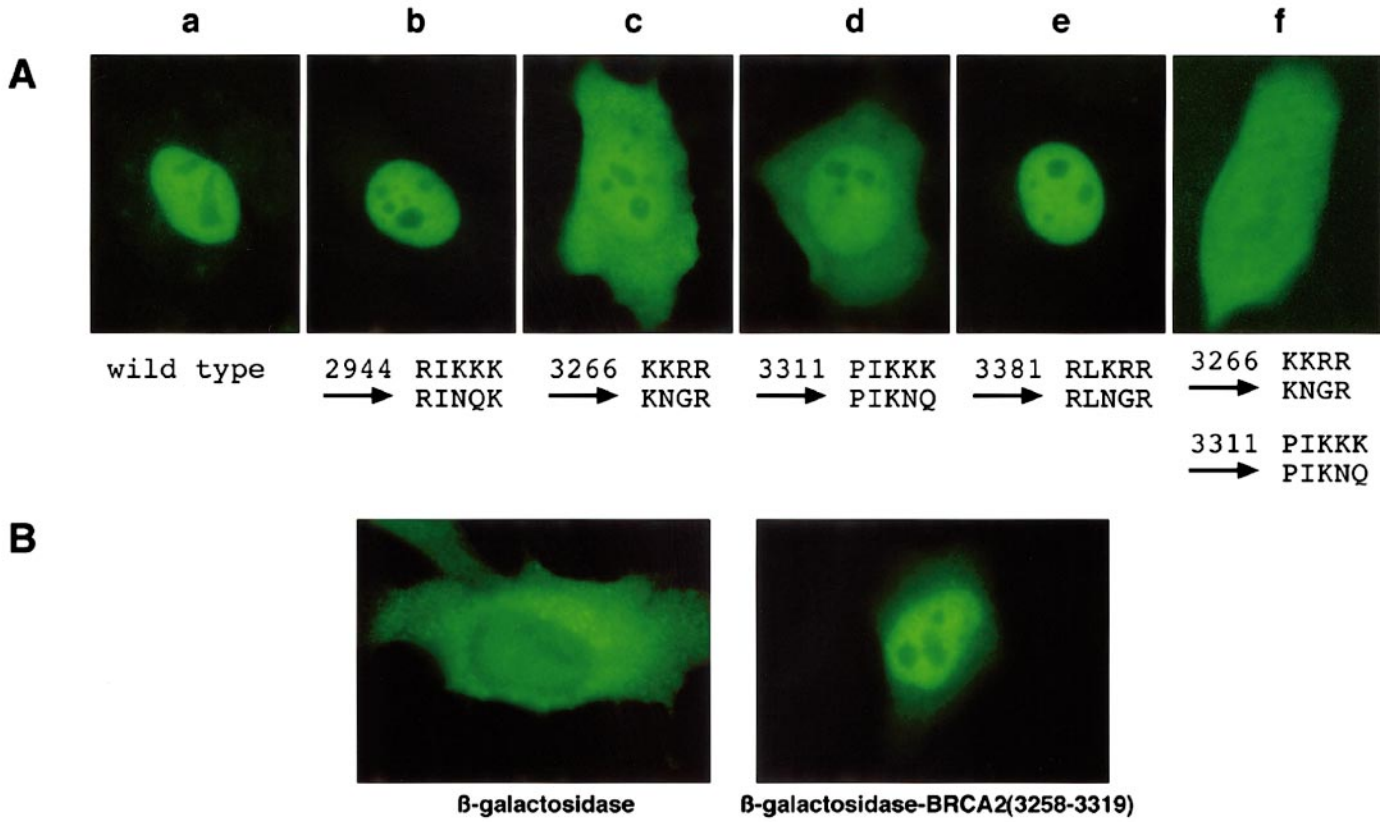

FIG. 2. Identification of the NLSs in the C-terminal domain of the BRCA2 protein. (A) Site-directed mutagenesis of basic amino acid arrays in the C-terminal domain of the BRCA2 protein. The plasmid e shown in Fig. 1 (A) (amino acids 2126–3418) was subjected to site-directed mutagenesis performed with a QuickChange mutagenesis kit (Stratagene). Subcellular localizations of the mutated proteins were observed as in Fig. 1. (B) The region containing two short arrays of basic amino acids is sufficient to transport the bacterial β-galactosidase into the nucleus. The Xpress-tagged β-galactosidase fused to the basic regions of the BRCA2 protein (right) and the Xpress-tagged β-galactosidase alone (left) were expressed in the HeLa cells. Subcellular distribution of the proteins was observed by immunofluorescent microscopy with anti-Xpress antibody.

pensable for the nuclear localization of the BRCA2 protein.

We then examined whether these basic sequences function as NLSs. A plasmid carrying *Escherichia coli lacZ* fused to the short region for the two basic arrays of the BRCA2 protein was transfected to the HeLa cells, and intracellular distribution of the fusion protein was observed by immunofluorescent microscopy. As shown in Fig. 2B, the β-galactosidase fused to the basic regions (amino acids 3258 to 3319) entered into the nucleus, while the β-galactosidase alone was distributed largely in cytoplasm. This result indicates that the basic amino acids in the C-terminal region of the BRCA2 protein constitute functional NLSs.

Many germline mutations have been reported throughout the coding region of the *BRCA2* gene in families with a high incidence of breast cancer. Interestingly, most of these mutations are insertion, deletion or nonsense mutations that invariably lead to truncation of the C-terminal region of the BRCA2 protein (2). The germline mutation mapped at the most downstream position in the *BRCA2* gene is a deletion at nucleotide 9808 that introduces a premature termination at codon 3195 (12). A polymorphic stop codon has also been reported at codon 3326 (13), implying that the domain from codons 3195 to 3326 carries a function critical for tumor suppression. It is noteworthy that the NLSs identified in this study locates between the germline mutation at the most downstream position and the polymorphic stop codon (Fig. 3).

DISCUSSION

BRCA2 is a tumor suppressor gene that encodes a large protein involved in the maintenance of genomic integrity through repair of DNA double strand breaks (2, 7). Dysfunction of *BRCA2* is believed to increase the lifetime risk of breast cancer, although the molecular mechanism underlying this relation remains unclear. In an attempt to define the functional domain, we here determined the region required for the nuclear localization of the BRCA2 protein. We demonstrated that two short arrays of basic amino acids were indispensable for the nuclear localization, and that the bacterial

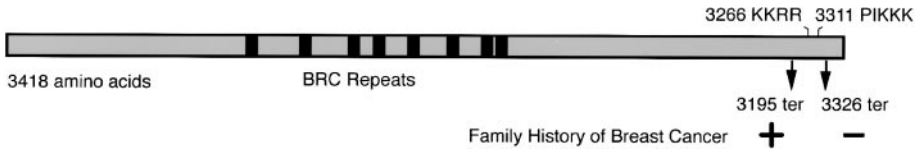

FIG. 3. The NLSs of the BRCA2 protein are mapped on the C-terminal region critical for tumor suppression. The BRCA2 protein is shown schematically. The germline mutation closest to the C-terminus and the polymorphic stop codon are shown, although many germline mutations in *BRCA2* have been reported throughout the coding sequence. The germline mutation closest to the C-terminus (at codon 3915) is associated with family history of breast cancer. The polymorphic termination at codon 3326, on the other hand, is not associated with family history of breast cancer. Positions of the NLSs are also indicated.

β-galactosidase fused to these basic regions entered the nucleus, indicating that the basic regions function as NLSs. Site-directed mutagenesis experiments revealed that disruption of one of the two basic regions reduced the nuclear localization of the EGFP-BRCA2 protein. Disruption of both the basic regions significantly enhanced the cytoplasmic distribution of the mutant protein. Many nuclear proteins contain a monopartite NLS that is a simple array of several basic amino acids, while others have a bipartite NLS consisting of two arrays of basic amino acids separated by spacer sequences (10). At present, more detailed experiments are required to examine whether two arrays of basic amino acids function as two independent monopartite NLSs or cooperatively as a single bipartite NLS.

In this study, we also found that the NLSs locate on the region critical for tumor suppression of the BRCA2 protein. Many germline mutations have been mapped through the coding region of the *BRCA2* gene (2). Most of these germline mutations are nonsense mutations or frameshifts that generate premature termination codons (2). The germline mutation at the most downstream position is a deletion at nucleotide 9808 that introduces premature termination at codon 3195 (12). At codon 3326, a polymorphic stop codon not associated with family history of breast cancer has been reported (13). We mapped the NLSs between the most downstream germline mutation in the family with high incidence of breast cancer and the polymorphic stop codon, and this mapping indicated that the products of most of the mutated *BRCA2* genes lack the NLS. In general, protein transport across the nuclear pore is a highly regulated process (10), and mislocalization is thought to impair the normal function of nuclear proteins (10, 14). We therefore hypothesize that the mislocalization of the mutated BRCA2 protein lacking the NLSs results in deficient maintenance of the genomic integrity, leading to the increased risk of breast cancer. Our data provides a possible explanation for the high frequency of frame-shift type and nonsense mutations in the *BRCA2* gene of hereditary breast cancer patients.

During the preparation of this manuscript, Spain *et al.* reported a similar study on the NLS of the BRCA2 protein, demonstrating that the C-terminal region is required for the nuclear transport (15). Spain *et al.* also identified two NLSs in the C-terminal domain of the BRCA2 protein. One NLS is present at codon 3266, consistent with our present findings. The other NLS, however, is at codon 3381, while the basic sequence at codon 3311, in our hand, is indispensable for efficient nuclear localization. Although at present we do not know the reason for this discrepancy, we imagine that it might reflect the difference in the experimental systems, e.g., the cell lines, studied. And again, despite this discrepancy, we point out that the first NLS (at codon 3266) is important for nuclear transport both in our experimental system and that of Spain *et al.* (15). Moreover, there is fairly general agreement that the NLS of the BRCA2 protein locates in the region downstream the most C-terminal germline mutation which is associated with family history of breast cancer.

In conclusion, we identified the NLSs that were mapped in the C-terminal region critical for tumor suppression, while further investigation will be required to elucidate the molecular mechanism of *BRCA2*-associated carcinogenesis.

ACKNOWLEDGMENTS

We thank Mayumi Soga for preparation of the manuscript. This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture in Japan.

REFERENCES

1. Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C., Micklem, G., Barfoot, R., Hamoudi, R., Patel, S., Rice, C., Biggs, P., Hashim, Y., Smith, A., Connor, F., Arason, A., Gudmundsson, J., Ficenec, D., Kelsell, D., Ford, D., Tonin, P., Bishop, D. T., Spurr, N. K., Ponder, B. A. J., Eeles, R., Peto, J., Devilee, P., Cornelisse, C., Lynch, H., Narad, S., Lenoir, G., Egilsson, V., Barkadottir, R. B., Easton, D. F., Bentley, D. R., Futreal, P. A., Ashworth, A., and Stratton, M. R. (1995) *Nature* **378,** 789–792.
2. Rahman, N., and Stratton, M. R. (1998) *Annu. Rev. Genet.* **32,** 95–121.
3. Bork, P., Blomberg, N., and Nilges, M. (1996) *Nat. Genet.* **13,** 22–23.
4. Wong, A. K. C., Pero, R., Ormonde, P. A., Tavtigian, S. V., and Bartel, P. L. (1997) *J. Biol. Chem.* **272,** 31941–31944.
5. Chen, J., Silver, D. P., Walpita, D., Cantor, S. B., Gazdar, A. F., Tomlinson, G., Couch, F. J., Weber, B. L., Ashley, T., Livingston, D. M., and Scully, R. (1998) *Mol. Cell* **2,** 317–328.

6. Patel, K. J., Yu, V. P. C. C., Lee, H., Corcoran, A., Thistlethwaite, F. C., Evans, M. J., Colledge, W. H., Friedman, L. S., Ponder, B. A. J., and Venkitaraman, A. R. (1998) *Mol. Cell* **1,** 347–357.
7. Chen, J. J., Silver, D., Cantor, S., Livingston, D. M., and Scully, R. (1999) *Cancer Res.* **59,** 1752s–1756s.
8. Chen, P. L., Chen, C. F., Chen, Y., Xiao, J., Sharp, Z. D., and Lee, W. H. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 5287–5292.
9. Marmorstein, L. Y., Ouchi, T., and Aaronson, S. A. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 13869–13874.
10. Nigg, E. A. (1997) *Nature* **386,** 779–787.
11. Harlow, E., and Lane, D. (1999) Using Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York.
12. Hakansson, S., Johannsson, O., Johansson, U., Sellberg, G., Loman, N., Gerdes, A. M., Holmberg, E., Dahl, N., Pandis, N., Kristoffersson, U., Olsson, H., and Borg, A. (1997) *Am. J. Hum. Genet.* **60,** 1068–1078.
13. Mazoyer, S., Dunning, A. M., Serova, O., Dearden, J., Puget, N., Healey, C. S., Gayther, S. A., Mangion, J., Stratton, M. R., Lynch, H. T., Goldgar, D. E., Ponder, B. A., and Lenoir, G. M. (1996) *Nat. Genet.* **14,** 253–254.
14. Matsumoto, T., Shimamoto, A., Goto, M., and Furuichi, Y. (1997) *Nat. Genet.* **16,** 335–336.
15. Spain, B. H., Larson, C. J., Shihabuddin, L. S., Gage, F. H., and Verma, I. M. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 13920–13925.