## Short Communication

# *NAT1\*10* and *NAT1\*11* Polymorphisms and Breast Cancer Risk[1]

**Robert C. Millikan**[2]

Department of Epidemiology, School of Public Health, Lineberger
Comprehensive Cancer Center, School of Medicine, University of North
Carolina, Chapel Hill, North Carolina 27599-7400

### Abstract

**Several recent epidemiological studies examined the
association of *N*-acetyltransferase (*NAT*) 1 and 2
genotypes and breast cancer risk. Taken together, these
studies do not support a strong role for the most common
*NAT* alleles in etiology of breast cancer. Only one study
estimated odds ratios (ORs) for the relatively rare
*NAT1\*11* allele: a strong positive association for the
*NAT1\*11* allele and breast cancer was reported, as well
as strong combined effects for *NAT1\*11*-containing
genotypes and two environmental factors, smoking and
red meat consumption. To further address the association
of *NAT1\*11* and breast cancer, an analysis was
performed using previously collected data from the
Carolina Breast Cancer Study, a population-based, case-
control study conducted in North Carolina. The OR for
*NAT1\*11*-containing genotypes and breast cancer was 0.5
(95% confidence interval, 0.2–1.3) among white women;
ORs were not calculated among African Americans
because only one participant exhibited the *NAT1\*11*
allele. There was no evidence for combined effects of
*NAT1\*11* and smoking. Unfortunately, the results of both
studies of *NAT1\*11* are imprecise and lack sufficient
statistical power to address fully the potential
contribution of *NAT1\*11* to breast cancer. These results
illustrate that the limitations imposed by sample size, as
well as incomplete knowledge of biological function, need
to be considered when planning and interpreting studies
of genetic polymorphisms and environmental exposures.**

### Introduction

A recent study by Zheng *et al.* (1) examined the role of *NAT*[3]
genetic polymorphisms and breast cancer risk. *NAT1* and *NAT*2
are involved in detoxication of aryl amines found in tobacco
smoke and in activation of heterocyclic amines found in cooked
meat. The relationship between *NAT* genotype and breast can-
cer risk has been examined in several recent studies (2–7).
Several of these studies reported interactions between *NAT*

genotype and environmental factors (smoking or diet), but these
interactions were observed among subgroups (*e.g.,* pre- or
postmenopausal women). The results are not consistent across
studies, and, taken together, they do not support a strong role
for *NAT1* or *NAT2* genotypes in risk of breast cancer.

Zheng *et al.* (1) reported a positive association between the
*NAT1\*11* allele and breast cancer risk, as well as strong combined
effects for *NAT1\*11* genotype and two environmental exposures,
cigarette smoking and red meat consumption. *NAT1* encodes a
variety of alleles, including *NAT1\*3*, *NAT1\*4, NAT1\*10,* and
*NAT1\*11*. Correlations between *NAT1* genotype and metabolic
phenotype are poorly understood (reviewed in Ref. 5). However,
recent evidence suggests that the enzyme encoded by the
*NAT1\*11* allele exhibits increased metabolic activation of *N*-hy-
droxy aromatic amines, relative to protein products of the *NAT1\*3*
and *NAT1\*4* alleles (reviewed in Ref. 1). In the study of Zheng *et
al.* (1), the OR for *NAT1\*11/*any genotype (presence of one or
more copy of the *NAT1\*11* allele) compared to *NAT1\*3-* or
*NAT1\*4* -containing genotypes (the more common alleles) was 3.9
(95% CI, 1.5–10.5). This OR is based on 11 cases and 7 controls
with *NAT1\*11/*any genotypes, out of a total of 154 cases and 328
controls. The OR for the combination of *NAT1\*11*/any genotype
and ever smoking (compared with the combination of *NAT1\*3-* or
*NAT1\*4*-containing genotypes and never smoking) was 13.2 (95%
CI, 1.5–116.0), based on five cases and one control with both
exposures. The OR for the combination of *NAT1\*11*/any genotype
and the highest tertile of red meat consumption (compared with the
combination of *NAT1\*3-* or *NAT1\*4*-containing genotypes and the
lowest tertile of meat consumption) was 6.1 (95% CI, 1.1–33.2),
based on five cases and two controls with both exposures. In
contrast, the OR for *NAT1\*10/*any genotype (compared with
*NAT1\*3-* or *NAT1\*4*-containing genotypes) was 1.3 (95% CI, 0.8
- 1.9). The OR for the combination of *NAT1\*10*/any genotype and
smoking (compared with the combination of *NAT1\*3-* or *NAT1\*4*-
containing genotypes and never smoking) was 1.4 (95% CI, 0.7–
2.9), and the OR for the combination of *NAT1\*10*/any genotype
and high levels of red meat consumption was 1.6 (95% CI, 0.6–
3.1). This study is the first to report a positive association between
the *NAT1\*11* allele and risk of cancer of any site and is also the
first to investigate interactions with environmental exposures.

To further examine the association of *NAT1\*11* and
*NAT1\*10* genotypes and breast cancer, previously collected
data from the Carolina Breast Cancer Study, a population-
based, case-control study of breast cancer, were used (5).

### Materials and Methods

The study design of the Carolina Breast Cancer Study and
methods for genotyping of *NAT1* and *NAT2* have been de-
scribed previously (5). ORs for breast cancer and 95% CI were
calculated using unconditional logistic regression models to
examine associations for *NAT1\*10* and *NAT1\*11* alleles.
PROC GENMOD of the software package SAS (version 6.11;
SAS Institute, Cary, NC) was used to incorporate offset terms
derived from sampling probabilities used to identify eligible
participants and to adjust for age (as an 11-level ordinal vari-

*Table 1*  ORs for *NAT1\*10* and *NAT1\*11* genotypes and breast cancer

| *NAT1* genotype | Cases | Controls | OR (95% CI)[a] |
|---|---|---|---|
| Whites | *n = 290*[b] | *n = 273*[b] | |
| *NAT1\*4/\*4, \*3/\*4, \*3/\*3* | 176 | 163 | Referent |
| *NAT1\*10/*any genotype | 108 | 103 | 1.0 (0.7–1.4) |
| *NAT1\*11/*any genotype | 6 | 11 | 0.5 (0.2–1.3) |
| African Americans | *n = 199*[b] | *n = 192*[b] | |
| *NAT1\*4/\*4, \*3/\*4, \*3/\*3* | 58 | 47 | Referent |
| *NAT1\*10/*any genotype | 141 | 145 | 0.8 (0.5–1.2) |
| *NAT1\*11/*any genotype | 1 | 0 | NA[c] |

[a] Adjusted for age and family history.
[b] Column totals exceed the number of cases and controls due to the overlap between *NAT1\*10/*any genotype and *NAT1\*11/*any genotype.
[c] NA, not applicable.

*Table 2*  ORs for the association of breast cancer with combinations of *NAT1\*10* and *NAT1\*11* genotypes and smoking

| | *NAT1* genotype | | | | | |
|---|---|---|---|---|---|---|
| | *NAT1\*4/\*4, \*3/\*4, \*3/\*3* | | *NAT1\*10/*any genotype | | *NAT1\*11/*any genotype | |
| | Cases/controls | OR (95% CI)[a] | Cases/controls | OR (95% CI)[a] | Cases/controls | OR (95% CI)[a] |
| Whites | | | | | | |
| Smoking status | | | | | | |
| Never | 79/78 | Referent | 52/51 | 1.0 (0.6–1.6) | 4/5 | 0.7 (0.2–2.7) |
| Ever | 97/85 | 1.1 (0.7–1.8) | 56/52 | 1.1 (0.6–1.7) | 2/6 | 0.3 (0.1–1.7) |
| African-Americans | | | | | | |
| Smoking status | | | | | | |
| Never | 32/30 | Referent | 79/87 | 0.7 (0.4–1.4) | 0/0 | NA[b] |
| Ever | 26/17 | 1.3 (0.6–3.0) | 62/58 | 1.0 (0.6–2.0) | 0/0 | NA |

[a] Adjusted for age and family history.
[b] NA, not applicable.

able reflecting 5-year age categories) and family history (defined as the presence of one or more first-degree relatives with breast cancer). Definitions of smoking and menopausal status were used as described previously, and race was classified according to self-report (5).

### Results

ORs for *NAT1\*10/*any and *NAT1\*11/*any genotypes and breast cancer, adjusted for age and family history of breast cancer using logistic regression, are presented in Table 1. There was no association for *NAT1\*10/*any genotype and breast cancer in African Americans or whites, and an inverse association for *NAT1\*11/*any genotype and breast cancer was seen in whites. The latter estimate was very imprecise due to the low frequency of the *NAT1\*11* allele. ORs for the joint effects of genotype and smoking are presented in Table 2. An inverse association with breast cancer for the combination of *NAT1\*11/*any genotype and smoking was observed among whites, although the estimate was imprecise. The role of red meat intake was not addressed because this information was not collected in our study.

### Discussion

Our results do not support a role for *NAT1\*10* or *NAT1\*11* alleles in breast cancer risk. Both our study and the study of Zheng *et al.* (1) observed ORs close to the null value for *NAT1\*10/*any genotype and for the combined effects of *NAT1\*10/*any genotype and ever smoking. However, in contrast to Zheng *et al.* (1), we did not observe strong main effects for *NAT1\*11/*any genotype or evidence of combined effects of smoking and *NAT1\*11/*any genotype.

There are several possible explanations for the differences

between our results and those of Zheng *et al.* (1). First, there may be differences in the populations studied. Our study population included African Americans (*n* = 391) and whites (*n* = 563). We categorized as "white" seven Native Americans, three Asian Americans, and three women who listed their race as "multiracial." ORs for whites did not differ after excluding these 13 individuals. The study population of Zheng *et al.* (1) was reported as "virtually all" Caucasian. In our study, frequencies of the *NAT1\*10* allele differed by race, but ORs for *NAT1\*10* genotype and breast cancer were similar in African Americans and whites (Table 1). The *NAT1\*11* allele was observed in only one African American. Our study population was approximately half premenopausal and half postmenopausal (5), whereas that of Zheng *et al.* (1) was entirely postmenopausal. We conducted additional analyses stratifying on the basis of menopausal status. ORs did not differ substantially in premenopausal *versus* postmenopausal women, although estimates were extremely imprecise. For example, in postmenopausal white women, the adjusted OR for *NAT1\*11/*any genotype (compared with *NAT1\*3*- or *NAT1\*4*-containing genotypes) was 0.6 (95% CI, 0.1–2.7), and the OR for the combination of *NAT1\*11/*any genotype and ever smoking (compared with the combination of *NAT1\*3*- or *NAT1\*4*-containing genotypes and never smoking) was 0.4 (95% CI, 0.04 - 4.41). Distributions of most traditional risk factors and associations with breast cancer were similar in the two studies (1, 5), and the time period for case ascertainment in the study of Zheng *et al.* (Ref. 1; 1992–1994) overlapped that of our study (1993–1996).

A second source of potential differences in results is that the methods for genotyping *NAT1* differed in the two studies. Both studies used PCR/RFLP-based methods, but the assay of

Zheng *et al.* (1) detected several alleles not detected by our study: (*a*) *NAT1*14*; (*b*) *NAT1*15*; (*c*) *NAT1*17*; and (*d*) *NAT1*22.* These four alleles were quite rare in the study of Zheng *et al.* (1), and failure to include them in our analysis would not affect ORs for *NAT1*10*/any genotype and *NAT1*11/*any genotype. Allele frequencies (*q*) for *NAT1*10* and *NAT1*11* did not differ significantly among controls across the two studies. For *NAT1*10*, $q = 0.17$ for controls in Zheng *et al.* (1) and 0.21 for white controls in our study ($P = 0.18$, $\chi^2$ test). For *NAT1*11*, $q = 0.01$ for controls in Zheng *et al.* (1) and 0.02 for white controls in our study ($P = 0.28$). The similarity in allele frequencies among controls suggests that the laboratory methods were comparable for detecting these alleles. Methods for classifying dose and duration of smoking differed in the two studies, but similar definitions for "ever" and "never" smoking were used. Odds in both studies were adjusted for age and family history (defined in both studies as "one or more first-degree relatives with breast cancer").

The most plausible explanation for the difference in results between our study and that of Zheng *et al.* (1) is random error. The difference in ORs for *NAT1*11/*any genotype is due to different estimates of allele frequency for *NAT1*11* among cases [$q = 0.04$ for cases in Zheng *et al.* (1), and $q = 0.01$ among white cases in our study ($P = 0.01$)]. Because *NAT1*11* alleles are rare, differences in allele frequencies among cases could have arisen due to chance. Assuming random error as an explanation for differences in results, homogeneity *P*s (8) were calculated comparing CIs derived from the study of Zheng *et al.* (1) and postmenopausal whites in our study. These tests provide strong evidence for heterogeneity: the *P* comparing CIs for the OR for *NAT1*11/*any genotype (compared with *NAT1*3-* or *NAT1*4-*containing genotypes) was 0.06; and the *P* comparing CIs for the OR for *NAT1*11/*any genotype and smoking (compared with *NAT1*3-* or *NAT1*4-*containing genotypes and never smoking) was 0.03. The sample sizes [308 cases and 656 controls for Zheng *et al.* (1) and 290 white cases and 273 white controls in our study] yield roughly 80% power to detect an OR of 3.0 or greater (0.33 or less) for genotypes with a frequency of 4% at a significance level of 0.05 (9). However, neither study had adequate power to estimate combined effects for *NAT1*11* genotype and environmental factors. Several methods for estimating sample size for case-control studies of gene-environment interaction have been developed (10–12). Even the most optimistic of these methods suggest that over 2000 cases and 2000 controls would be needed to investigate interactions between *NAT1*11/*any genotype and smoking or red meat consumption. In addition, Rothman *et al.* (13) showed that the presence of even small amounts of genotype or exposure misclassification can increase sample size requirements substantially. Smith and Day (14) remind us that when sample size is limiting, "very many of the reported 'significant' interaction effects are no more than chance observations." Thus, as concluded by Zheng *et al.* (1), studies of gene-environment interaction based on small numbers of participants should properly be regarded as "preliminary."

The need for large sample sizes is an important challenge for epidemiologists studying gene-environment interaction. The advent of high throughput techniques will allow investigators to conduct genotyping assays on large numbers of participants and increase power to estimate gene-environment interactions (15). Empirical Bayes' methods may prove useful for addressing data sparseness and associations arising due to chance (16, 17). However, for loci such as *NAT1*, where knowledge of biological function and toxokinetics is incomplete, studies of gene-environment interaction will continue to present significant challenges in interpretation. Repetition in multiple study populations may contribute little to causal inference. Thus, in addition to the problems of sample size and statistical power, the limitations imposed by incomplete knowledge of biological function need to be considered when planning and interpreting studies that estimate joint effects for genetic polymorphisms and environmental exposures.

## References

1. Zheng, W., Dietz, A., Campbell, D., Wen, W-Q., Cerhan, J., Sellers, T., Folsom, A., and Hein, D. *N*-Acetyltransferase 1 genetic polymorphism, cigarette smoking, well-done meat intake, and breast cancer risk. Cancer Epidemiol. Biomark. Prev., *8:* 233–239, 1999.

2. Ambrosone, C., Freudenheim, J., Graham, S., Marshall, J., Vena, J., Brasure, J., Michalek, A., Laughlin, R., Nemoto, T., Gillenwater, K., Harrington, A., and Shields, P. Cigarette smoking, *N*-acetyltransferase 2 genetic polymorphisms, and breast cancer risk. J. Am. Med. Assoc., *276:* 1494–1501, 1996.

3. Ambrosone, C., Freudenheim, J., Sinha, R., Graham, S., Marshall, J., Vena, J., Laughlin, R., Nemoto, T., and Shields, P. Breast cancer risk, meat consumption and *N*-acetyltransferase (NAT2) genetic polymorphisms. Int. J. Cancer, *75:* 825–830, 1998.

4. Hunter, D., Hankinson, S., Hough, H., Gertig, D., Garcia-Closas, M., Spiegelman, D., Manson, J., Colditz, G., Willett, W., Speizer, F., and Kelsey, K. A prospective study of NAT2 acetylation genotype, cigarette smoking, and risk of breast cancer. Carcinogenesis (Lond.), *18:* 2127–2132, 1997.

5. Millikan, R., Pittman, G., Newman, B., Tse, C-K., Selmin, O., Rockhill, B., Savitz, D., Moorman, P., and Bell, D. Cigarette smoking, *N*-acetyltransferases 1 and 2, and breast cancer risk. Cancer Epidemiol. Biomark. Prev., *7:* 371–378, 1998.

6. Morabia, A., Bernstein, M., Heritier, S., Bouchardy, I., and Morris, M. *NAT2*-smoking interaction with respect to breast cancer in menopausal women. Am. J. Epidemiol., *147:* S45, 1998.

7. Huang, C-S., Chern, H-D., Shen, C-Y., Hsu, S-M., and Chang, K-J. Association between *N*-acetyltransferase 2 (*NAT2*) genetic polymorphism and development of breast cancer in postmenopausal Chinese women in Taiwan, an area of great increase in breast cancer incidence. Int. J. Cancer, *82:* 175–179, 1999.

8. Poole, C., and Greenland, S. Random-effects meta-analyses are not always conservative. Am. J. Epidemiol., *150:* 469–475, 1999.

9. Schlesselman, J. Case-Control Studies: Design, Conduct, Analysis. New York: Oxford University Press, 1982.

10. Hwang, S-J., Beaty, T., Liang, K-Y., Coresh, J., and Khoury, M. Minimum sample size estimation to detect gene-environment interaction in case-control designs. Am. J. Epidemiol., *140:* 1029–1037, 1994.

11. Foppa, I., and Spiegelman, D. Power and sample size calculations for case-control studies of gene-environment interactions with a polytomous exposure variable. Am. J. Epidemiol., *146:* 596–604, 1997.

12. Garcia-Closas, M., and Lubin, J. Power and sample size calculations in case-control studies of gene-environment interactions: comments on different approaches. Am. J. Epidemiol., *149:* 689–692, 1999.

13. Rothman, N., Garcia-Closas, M., Stewart, W., and Lubin, J. The impact of misclassification in case-control studies of gene-environment interactions. *In:* P. Vineis, N. Malats, M. Lang, A. d'Errico, N. Caporaso, J. Cuzick, and P. Boffetta (eds.), Metabolic Polymorphisms and Susceptibility to Cancer, pp. 89–96. Lyon, France: IARC, 1999.

14. Smith, P., and Day, N. The design of case-control studies: the influence of confounding and interaction effects. Int. J. Epidemiol., *13:* 356–365, 1984.

15. Blomeke, B., and Shields, P. Laboratory methods for the determination of genetic polymorphisms in humans. *In:* P. Vineis, N. Malats, M. Lang, A. d'Errico, N. Caporaso, J. Cuzick, and P. Boffetta (eds.), Metabolic Polymorphisms and Susceptibility to Cancer, pp. 133–147. Lyon, France: IARC, 1999.

16. Greenland, S., and Robins, J. Empirical-Bayes adjustments for multiple comparisons are sometimes useful. Epidemiology, *2:* 244–251, 1991.

17. Aragaki, C., Greenland, S., Probst-Hensch, N., and Haile, R. Hierarchical modeling of gene-environment interactions: estimating *NAT2** genotype specific-specific dietary effects on adenomatous polyps. Cancer Epidemiol. Biomark. Prev., *6:* 307–314, 1997.

AACR American Association for Cancer Research

# *NAT1\*10* and *NAT1\*11* Polymorphisms and Breast Cancer Risk

Robert C. Millikan

| | |
|---|---|
| **Updated version** | Access the most recent version of this article at:<br>http://cebp.aacrjournals.org/content/9/2/217 |

| | |
|---|---|
| **Cited articles** | This article cites 14 articles, 3 of which you can access for free at:<br>http://cebp.aacrjournals.org/content/9/2/217.full#ref-list-1 |
| **Citing articles** | This article has been cited by 4 HighWire-hosted articles. Access the articles at:<br>http://cebp.aacrjournals.org/content/9/2/217.full#related-urls |

| | |
|---|---|
| **E-mail alerts** | Sign up to receive free email-alerts related to this article or journal. |
| **Reprints and Subscriptions** | To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org. |
| **Permissions** | To request permission to re-use all or part of this article, use this link<br>http://cebp.aacrjournals.org/content/9/2/217.<br>Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC)<br>Rightslink site. |