Inconsistent Association Between the STK15 F31I Genetic Polymorphism and Breast Cancer Risk

Olivia Fletcher, Nichola Johnson, Claire Palles, Isabel dos Santos Silva, Valerie McCormack, John Whittaker, Alan Ashworth, Julian Peto

STK15 may be a low-penetrance breast cancer susceptibility gene, and several reports suggest that women who are homozygous for the polymorphic variant F31I have an increased risk of breast cancer. To evaluate this potential breast cancer allele, we genotyped 507 patients with two primary breast cancers and 875 population-based control subjects for the STK15 F31I polymorphism. All statistical tests were two-sided. The Ile/Ile homozygous genotype was not associated with an increased risk in white women of British descent. The odds ratio for developing two primary breast cancers) in Ile/Ile homozygotes was 0.63 (95% confidence interval [CI] = 0.34 to 1.13), which corresponds to an odds ratio of 0.79 (95% CI = 0.58 to 1.06) for a first primary breast cancer. A meta-analysis of this study and other published studies showed statistically significant heterogeneity in the odds ratio estimates (*P*<.001). This heterogeneity could reflect either population-specific linkage disequilibrium with a functional variant or artifacts such as population stratification or publication bias. [J Natl Cancer Inst 2006; 98:1014–8]

The STK15 gene encodes a serine/threonine kinase that acts as a key regulator of mitotic chromosome segregation *(1,2)*. Studies in transgenic mice have suggested that STK15 is a candidate low-penetrance tumor susceptibility gene *(3)*, and analyses of archival tumor blocks *(4)* and experiments examining the role of the STK15 gene in chromosome segregation have provided compelling evidence that STK15 has a role in the etiology of cancer. A single-nucleotide polymorphism in the STK15 gene, T91A, results in the polymorphic substitution of isoleucine (Ile) for phenylalanine (Phe) at residue 31 (F31I) *(3)*. Several reports *(5–8)* suggest that women who are homozygous for the Ile/Ile allele have an increased risk of breast cancer.

To further examine the association between this polymorphism and the risk of breast cancer, we analyzed the prevalence of this variant in a series of 507 patients with two primary breast cancers and 875 healthy control subjects. We specifically selected patients with two breast cancers for inclusion in this study because association studies based on cancer patients with a family history of the disease or with multiple primary cancers have greater power to detect cancer susceptibility genes than studies of unselected patients, most of whom have a single primary breast cancer and no family history *(9,10)*.

Our case patients were women with two primary breast cancers who were identified through the English Cancer Registries, as previously described *(11)*. 448 (88.4%) of our case patients had bilateral disease and 59 (11.6%) had developed a later second primary cancer in the same breast. Control subjects were non-blood female relatives and friends of the case patients (n = 382) and women who were recruited from mammography

Affiliations of authors: The Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, London, UK (OF, NJ, CP, AA); Non-communicable Disease Epidemiology Unit, London School of Hygiene & Tropical Medicine, London, UK (IdSS, VM, JW, JP); Cancer Research UK Epidemiology & Genetics Unit, Institute of Cancer Research–Sutton, Surrey, UK (JP).

Correspondence to: Olivia Fletcher, PhD, The Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, U.K. (e-mail: olivia.fletcher@icr.ac.uk).

See "Notes" following "References."

DOI: 10.1093/jnci/djj268

clinics in England (n = 493). Written informed consent was obtained from all study subjects. This study was approved by the South East Multi-centre Research Ethics Committee. All women reported being white and of British descent. DNA from case patients and control subjects was extracted from peripheral blood lymphocytes using a QIAmp DNA Blood Mini kit (Qiagen, Sussex, United Kingdom). Case patients and control subjects were genotyped for the STK15 T91A single-nucleotide polymorphism (rs2273535) by a polymerase chain reaction-restriction fragment length polymorphism assay that used the restriction enzyme ApoI. This polymorphism has a minor allele frequency of 21% in European control populations, resulting in a rare homozygote frequency of 4.4% *(5,6)*. Primer sequences and reaction conditions used for genotyping SKT15 are available at http://jncicancerspectrum.oxfordjournals.org/jnci/content/vol98/issue14.

If a second breast cancer arises independently of the first primary, the odds ratio (OR) for having two primary tumors in carriers of a susceptibility genotype (ψ_B) is approximately the square of the odds ratio (ψ_U) for unselected first primary cancers *(9)*. However, if women who have developed a first breast cancer are at increased risk for nongenetic reasons (e.g., environmental risk factors, or, conceivably, a somatic effect of the first cancer), the square root is not the appropriate transformation from ψ_B to ψ_U. We have used the square root transformation and presented our results in terms of ψ_U to make our results compatible with those from other studies, but our statistical conclusions are virtually unaffected by this convention (see Appendix I). Odds ratios with 95% confidence intervals (CIs) were calculated using the STATA statistical package (version 8.0; Stata Corporation, College Station, TX). All statistical tests were two-sided.

We found that the Ile/Ile homozygous genotype did not confer an increased risk for developing two primary breast cancers (number of case patients by genotype: Phe/Phe, N = 335; Phe/Ile, N = 154; Ile/Ile, N = 18; number of control subjects by genotype: Phe/Phe, N = 547; Phe/Ile, N = 280, Ile/Ile, N = 48; OR [ψ_B] = 0.63, 95% CI = 0.34 to 1.13). Because there was no a priori reason to assume a recessive effect of STK15 F31I on breast cancer risk, we also estimated the multiplicative risk per allele by fitting a codominant model. The resulting estimates of the odds ratio for first primary cancers in Ile/Ile homozygotes were 0.79 (95% CI = 0.58 to 1.06) under a recessive model and 0.85 (95% CI = 0.70 to 1.03) under a codominant model. Both upper 95% confidence limits were close to unity. Thus, although we cannot rule out a positive association between STK15 F31I and the risk of breast cancer in other populations, we conclude that this association is negligible or negative in English women regardless of the mode of genetic transmission.

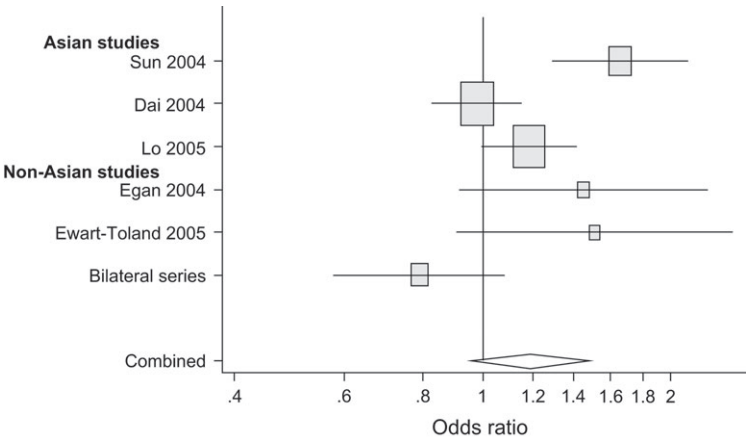

Fig. 1. Forest plot of odds ratios and 95% confidence intervals for STK15 Ile/Ile homozygotes versus Ile/Phe heterozygotes and Phe/Phe homozygotes combined. **Shaded squares** indicate the odds ratio, with the square size proportional to the inverse variance of the study-specific estimate. **Horizontal lines** represent 95% confidence intervals. The **diamond** represents the combined, random-effects estimate of the odds ratio and 95% confidence interval. The **vertical line** indicates the null effect (odds ratio = 1.0).

To combine our data with that of other published studies, we searched PubMed and EMBASE databases using the keywords "stk15," "aurora," "aurka," and "BTAK" alone or in combination with "breast cancer" or "case–control" or "cohort" for English language articles published between January 1, 1966, and October 31, 2005. Reference lists within all relevant articles and reviews were searched to identify publications not captured by the computerized searches. This search identified only five studies *(5–8,12)*, three of which were conducted among Asian populations. We calculated Cochran's Q statistic *(13)* to test for heterogeneity among the odds ratio estimates for the studies and the I^2 statistic *(14)* to quantify between-study variation. Because there was evidence of statistically significant heterogeneity between studies, we used a random-effects model to estimate the combined odds ratio *(15)*.

Fig. 1 and Table 1 show estimates under a recessive model of the odds ratio for a first primary breast cancer among homozygotes (Ile/Ile) from this study (ψ_U = 0.79, which we calculated as the square root of ψ_B = 0.63, our observed odds ratio in cases with two primary breast cancers, see Appendix I) and from the other published series of unselected cases *(5–8,12)*. However, the pooled odds ratio for all studies combined (OR = 1.19, 95% CI = 0.96 to 1.48) is difficult to interpret because the odds ratio estimates in the different studies were heterogeneous (Cochran's Q = 20.47, P<.001, I^2 = 75.6%). When we conducted separate meta-analyses for the three Asian studies *(7,8,12)* and for our data and the two non-Asian studies *(5,6)*, we still found evidence of between-study heterogeneity, particularly among the Asian studies (Asian studies: Cochran's Q = 12.2, P = .002, I^2 = 83.6%; non-Asian studies: Cochran's Q = 7.41, P = .03, I^2 = 73.0%). The Asian populations may not be genetically similar, because the Ile/Ile homozygote frequency was statistically significantly lower in Han Chinese control subjects *(7)* than in Shanghai Chinese control subjects *(8,12)* (36.9% versus 45.0%, difference = 8.1%, 95% CI = 3.1% to 13.1%; P = .002, Fisher's exact test). There may also be genetic differences among populations in the non-Asian studies. The case patients and control subjects in our study were white English residents, whereas the North American subjects studied by Egan et al. *(5)* included many of central European origin and the North American subjects studied by Ewart-Tolland et al. *(6)* included Asians, African Americans, and Hispanics.

Table 1. Characteristics and results of case–control studies of associations between STK15 F31I and breast cancer*

Study, year	Case patients	Control subjects	Power to detect OR of 1.5 in recessive model†	OR (95% CI) Recessive model *I*/*I* versus (*F*/*I*+*F*/*F*)	OR (95% CI) Codominant model *I*/*I* versus *F*/*F*	OR (95% CI) Codominant model *F*/*I* versus *F*/*F*
			Asian			
Sun et al., 2004	*N* = 520 Han Chinese *F*/*F n* = 50 (9.6%) *F*/*I n* = 214 (41.1%) *I*/*I n* = 256 (49.3%)	*N* = 520 Han Chinese *F*/*F n* = 66 (12.7%) *F*/*I n* = 262 (50.4%) *I*/*I n* = 192 (36.9%)	89% (α = 0.05) 73% (α = 0.01)	1.66 (1.28 to 2.14)	1.76 (1.14 to 2.72)	1.08 (0.70 to 1.67)
Dai et al., 2004	*N* = 1102 Chinese (Shanghai) *F*/*F n* = 121 (11.0%) *F*/*I n* = 491 (44.6%) *I*/*I n* = 490 (44.5%)	*N* = 1186 Chinese (Shanghai) *F*/*F n* = 149 (12.6%) *F*/*I n* = 503 (42.4%) *I*/*I n* = 534 (45.0%)	>99% (α = 0.05) 98% (α = 0.01)	0.98 (0.83 to 1.16)	1.13 (0.86 to 1.49)	1.20 (0.91 to 1.59)
Lo et al., 2005	*N* = 707 Taiwanese *F*/*F n* = 71 (10.0%) *F*/*I n* = 288 (40.7%) *I*/*I n* = 348 (49.3%)	*N* = 1969 Taiwanese *F*/*F n* = 196 (10.0%) *F*/*I n* =887 (45.0%) *I*/*I n* = 886 (45.0%)	>99% (α = 0.05) 97% (α = 0.01)	1.18 (0.99 to 1.41)	1.08 (0.80 to 1.48)	0.89 (0.66 to 1.23)
			Non-Asian			
Egan et al., 2004	*N* = 940 Caucasian (mostly central European ancestry) *F*/*F n* = 559 (59.5%) *F*/*I n* = 331 (35.2%) *I*/*I n* = 50 (5.3%)	*N* = 830 Caucasian (mostly central European ancestry) *F*/*F n* = 516 (62.2%) *F*/*I n* = 283 (34.1%) *I*/*I n* = 31 (3.7%)	43% (α = 0.05) 20% (α = 0.01)	1.45 (0.90 to 2.37)	1.49 (0.92 to 2.45)	1.08 (0.89 to 1.32)
Ewart-Toland et al., 2005	*N* = 898 89% Caucasian, 4% Asian, 6% African-American *F*/*F n* = 533 (59.4%) *F*/*I n* = 303 (33.7%) *I*/*I n* = 62 (6.9%)	*N* = 448 80% Caucasian, 12% Hispanic, 4% Asian, 1% African-American *F*/*F n* = 279 (62.3%) *F*/*I n* = 148 (33.0%) *I*/*I n* = 21 (4.7%)	35% (α = 0.05) 15% (α = 0.01)	1.51 (0.89 to 2.64)	1.55 (0.91 to 2.73)	1.07 (0.83 to 1.38)
This series‡	*N* = 507 Caucasian (English) *F*/*F n* = 335 (66.1%) *F*/*I n* = 154 (30.3%) *I*/*I n* = 18 (3.6%)	*N* = 875 Caucasian (English) *F*/*F n* = 547 (62.5%) *F*/*I n* = 280 (32.0%) *I*/*I n* = 48 (5.5%)	93% (α = 0.05) 81% (α = 0.01)	0.79 (0.58 to 1.06)	0.78 (0.57 to 1.04)	0.95 (0.84 to 1.07)

*OR = odds ratio; CI = confidence interval; *N* = total number of case patients or control subjects; α = two-sided statistical significance level; *F* = variant of STK15 with phenylalanine at codon 31, *I* = variant of STK15 with isoleucine at codon 31.

†Power calculations were for detecting an odds ratio of 1.5 for rare homozygotes (*I*/*I*) versus all other genotypes (*F*/*I* + *F*/*F*) at statistical significance levels of 5% and 1%. Calculations were based on a rare homozygote frequency of 40% in Asian populations and 4.4% in non-Asian populations (*5–8,12*).

‡For the data from our series, the ORs and 95% CIs shown are the square roots of the values for case patients with two primary cancers.

This heterogeneity among the published studies in our meta-analysis could also reflect population-specific linkage disequilibrium between the STK15 F31I polymorphism and another functional variant in STK15 or in another gene. Lo et al. *(8)* suggested that the F31I polymorphism in the Taiwanese population in their study was in linkage disequilibrium with a more extended haplotype that was associated with breast cancer risk. If this was the case, then the F31I polymorphism might be associated with increased breast cancer risk in some populations and with reduced risk in others. A less interesting, but perhaps more plausible, explanation for the heterogeneity among studies is an artifact such as population stratification or publication bias. The odds ratio estimate from the first published report of a genetic association is often greater than estimates reported in subsequent studies *(16,17)*. The first published study of STK15 F31I in breast cancer *(7)* reported a statistically significant association (OR = 1.66, 95% CI = 1.28 to 2.14). When we excluded this study from the meta-analysis, the combined odds ratio estimate for the five remaining studies was not statistically significantly greater than unity (OR = 1.09, 95% CI = 0.90 to 1.31) and the between-study heterogeneity was reduced, although still just statistically significant (Cochran's Q = 10.00, P = .04, I^2 = 60.0%).

An important advantage of association studies of women with two primary breast cancers patients is the gain in statistical efficiency. Table 2 shows that the reduction in sample size that can be achieved by studying bilateral breast cancers ranges from fourfold to more than fivefold depending on the relative risk, the allele frequency, and the genetic model. Under a polygenic model for breast cancer susceptibility, a large number of genes that individually confer low risks act in combination, resulting in a wide spectrum of risk in the population. To reliably identify individuals who are at high risk of breast cancer, it may be necessary to detect large numbers of relevant "polygenes" that confer odds ratios as low as 1.2. If the true odds ratio in homozygotes were 1.2, the number of bilateral case patients needed to detect a single nucleotide polymorphism such as STK15 T91A with a minor allele frequency of approximately 0.2 would

Table 2. Numbers of unselected breast cancer cases and cases with two primary breast cancers required to detect odds ratios of 1.2 to 2.5 associated with a susceptibility allele, assuming a case patient/control subject ratio of 1:1, 80% statistical power, and 1% level of statistical significance*

	OR (ψ_U) (homozygotes versus noncarriers)				
	1.2	1.5	1.7	2.0	2.5
Recessive model					
Allele frequency = 0.05					
Unselected	261 700	48 500	27 000	14 900	7800
Two primaries	60 900	10 400	5500	2900	1400
Unselected/two primaries	4.3	4.7	4.9	5.1	5.6
Allele frequency = 0.10					
Unselected	66 000	12 250	6800	3800	2000
Two primaries	15 400	2600	1400	735	360
Unselected/two primaries	4.3	4.7	4.9	5.2	5.6
Allele frequency = 0.20					
Unselected	17 100	3200	1800	1000	530
Two primaries	4000	700	380	200	100
Unselected/two primaries	4.3	4.6	4.7	5.0	5.3
Codominant model					
Allele frequency = 0.05					
Unselected	28 400	5500	3200	1800	970
Two primaries	6900	1300	710	390	210
Unselected/two primaries	4.1	4.2	4.5	4.6	4.6
Allele frequency = 0.10					
Unselected	15 100	3000	1700	950	530
Two primaries	3700	690	385	215	120
Unselected/two primaries	4.1	4.3	4.4	4.4	4.4
Allele frequency = 0.20					
Unselected	8600	1700	1000	560	315
Two primaries	2100	410	230	135	75
Unselected/two primaries	4.1	4.1	4.3	4.2	4.2

*Under the recessive model only rare homozygotes are at increased risk, whereas under the multiplicative codominant model the odds ratio (OR) in heterozygotes is the square root of that in rare homozygotes. The range of ORs chosen reflects those reported for low-penetrance breast cancer alleles (e.g., TP53 Arg72Pro: OR = 1.27, 95% confidence interval [CI] = 1.02 to 1.59; GSTP1: OR = 1.60, 95% CI = 1.08 to 2.59; CHEK2*1100delC: OR = 2.34, 95% CI = 1.72 to 3.20) *(18,21)*.

be 2100 for a codominant effect or 4000 for a recessive effect. The corresponding numbers of unselected case patients that would be required would be 8600 (codominant) and 17 100 (recessive). Moreover, an observed odds ratio of 1.2 in a large study of unselected case patients could be due to the combined effects of chance and the biases inherent in genetic epidemiology. A gene that conferred an odds ratio of 1.2 for a first primary breast cancer would give an odds ratio of more than 1.4 in breast cancer patients with two primaries, which because of its greater magnitude would be less likely to be the result of such biases. Population-based studies of familial breast cancer cases or cases with two primaries also have other advantages. For example, a trend in the prevalence of the risk allele with an increasing number of affected relatives, as was seen for the 1100delC allele of the CHEK2 gene *(18)*, provides independent evidence of a real genetic effect. Moreover, the cancer rate in relatives of breast cancer patients with two primaries may indicate whether a risk allele interacts multiplicatively or additively with other undiscovered genes *(11)*.

Appendix I

The high rate of breast cancer in the contralateral breast in breast cancer survivors is likely to be due mainly to genetic predisposition. Their rate (per breast) is the same as in their identical twins and is independent of the stage of their first cancer *(19,20)*. Ignoring elimination of susceptible women and assuming that the two primary breast cancers are independent events and that the prevalence of genetic susceptibility that confers a relative risk, r, is p (e.g., the prevalence of rare homozygotes under a recessive model), then the odds of being genetically susceptible are $p/(1-p)$ in the population, $pr/(1-p)$ for first primary cancers, and $pr^2/(1-p)$ in women with two primary cancers, so $\psi_U = r$ and $\psi_B = r^2$.

If women who have developed a first breast cancer are at increased risk for nongenetic reasons, the square root is not the appropriate transformation from ψ_B to ψ_U, but our statistical conclusions would be virtually the same under any model. The statistical significance of the difference of ψ_U from unity will be the same as for ψ_B for any monotonically increasing transformation that maps unity to itself, and the statistical significance of a test for heterogeneity between non-Asian studies based on ψ_B rather than ψ_U would be $P = .05$ instead of $P = .03$ ($Q = 6.03$, $P = .05$, $I^2 = 66.8\%$ compared with $Q = 7.41$, $P = .03$, $I^2 = 73.0\%$).

References

(1) Dutertre S, Descamps S, Prigent C. On the role of aurora-A in centrosome function. Oncogene 2002;21:6175–83.

(2) Hirota T, Kunitoku N, Sasayama T, Marumoto T, Zhang D, Nitta M, et al. Aurora-A and an interacting activator, the LIM protein Ajuba, are required for mitotic commitment in human cells. Cell 2003;114:585–98.

(3) Ewart-Toland A, Briassouli P, de Koning JP, Mao JH, Yuan J, Chan F, et al. Identification of Stk6/STK15 as a candidate low-penetrance tumor-susceptibility gene in mouse and human. Nat Genet 2003;34:403–12.

(4) Bischoff JR, Anderson L, Zhu Y, Mossie K, Ng L, Souza B, et al. A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. EMBO J 1998;17:3052–65.

(5) Egan KM, Newcomb PA, Ambrosone CB, Trentham-Dietz A, Titus-Ernstoff L, Hampton JM, et al. STK15 polymorphism and breast cancer risk in a population-based study. Carcinogenesis 2004;25:2149–53.

(6) Ewart-Toland A, Dai Q, Gao YT, Nagase H, Dunlop MG, Farrington SM, et al. Aurora-A/STK15 T+91A is a general low penetrance cancer susceptibility gene: a meta-analysis of multiple cancer types. Carcinogenesis 2005; 26:1368–73.

(7) Sun T, Miao X, Wang J, Tan W, Zhou Y, Yu C, et al. Functional Phe31Ile polymorphism in Aurora A and risk of breast carcinoma. Carcinogenesis 2004;25:2225–30.

(8) Lo YL, Yu JC, Chen ST, Yang HC, Fann CS, Mau YC, et al. Breast cancer risk associated with genotypic polymorphism of the mitosis-regulating gene Aurora-A/STK15/BTAK. Int J Cancer 2005;115:276–83.

(9) Antoniou AC, Easton DF. Polygenic inheritance of breast cancer: Implications for design of association studies. Genet Epidemiol 2003;25:190–202.

(10) Houlston RS, Peto J. The future of association studies of common cancers. Hum Genet 2003;112:434–5.

(11) Johnson N, Fletcher O, Naceur-Lombardelli C, dos Santos Silva I, Ashworth A, Peto J. Interaction between CHEK2*1100delC and other low-penetrance breast-cancer susceptibility genes: a familial study. Lancet 2005; 366:1554–7.

(12) Dai Q, Cai QY, Shu XO, Ewart-Toland A, Wen WQ, Balmain A, et al. Synergistic effects of STK15 gene polymorphisms and endogenous estrogen exposure in the risk of breast cancer. Cancer Epidemiol Biomarkers Prev 2004;13:2065–70.

(13) Cochrane WG. The combination of estimates from different experiments. Biometrics 1954;10:101–29.

(14) Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–60.
(15) Sharp S, Sterne J. Meta-analysis. Stata Tech Bull 1997;38:9–14.
(16) Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. Nat Genet 2001;29:306–9.
(17) Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 2003; 33:177–82.
(18) CHEK2 Breast Cancer Case-Control Consortium. CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10 860 breast cancer cases and 9,065 controls from 10 studies. Am J Hum Genet 2004;74:1175–82.
(19) Peto J, Mack TM. High constant incidence in twins and other relatives of women with breast cancer. Nat Genet 2000;26:411–4.
(20) Harvey EB, Brinton LA. Second cancer following cancer of the breast in Connecticut, 1935–82. Natl Cancer Inst Monogr 1985; 68:99–112.
(21) Dunning AM, Healey CS, Pharoah PD, Teare MD, Ponder BA, Easton DF. A systematic review of genetic polymorphisms and breast cancer risk. Cancer Epidemiol Biomarkers Prev 1999;8:843–54.

NOTES

This study was funded by Cancer Research UK and Breakthrough Breast Cancer.

The funding sources had no role in the study design or collection and analysis of data, nor in the preparation of the manuscript and the decision to submit the paper for publication.

Manuscript received December 2, 2005; revised May 5, 2006; accepted May 30, 2006.