Is *TGFBR1*6A* a Susceptibility Allele for Nonsyndromic Familial Colorectal Neoplasia?

Denise Daley,[1,8] Wendi Morgan,[2] Susan Lewis,[2] Joseph Willis,[4] Robert C. Elston,[1,5] Sanford D. Markowitz,[3,5,7] and Georgia L. Wiesner[2,3,6]

Departments of [1]Epidemiology and Biostatistics, [2]Genetics, [3]Pathology and [4]Medicine, [5]Case Comprehensive Cancer Center, [6]Center for Human Genetics, [7]University Hospitals of Cleveland/Case Western Reserve University and Howard Hughes Medical Institute, and Cleveland, [8]Ohio and the University of British Columbia, Vancouver, BC, Canada

Abstract

Our analysis definitely excludes the possibility of the *TGFBR1*6A* allele increasing the risk of colorectal neoplasia in our sample population. A recent study validating linkage of colorectal cancer to chromosome 9q also excluded the *TGFBR1*6A* allele as a disease-causing variant in that sample. We conclude that there remains an unidentified susceptibility locus in the region 9q22.2-31.2. (Cancer Epidemiol Biomarkers Prev 2007;16(5):892–4)

Introduction

Several case-control studies have reported an association between a common variant of the type I transforming growth factor-β receptor (*TGFBR1*6A*) and human colorectal cancer (1-5). This gene colocalizes to the same chromosomal region (9q22.2-31.2) identified by the Colon Neoplasia Sibling Study (6) and recently validated in studies from Sweden and United Kingdom as a putative susceptibility locus for colorectal neoplasia (7, 8). To determine whether the *TGFBR1*6A* allele is responsible for the observed linkage at 9q22.2-31.2, we genotyped 53 kindreds from the Colon Neoplasia Sibling Study with colon cancer or advanced colon adenomas for the presence of the *TGFBR1*6A* or *TGFBR1*9A* allele. We then examined the evidence for linkage and allelic association with the *TGFBR1*6A* allele.

Ascertainment and Phenotyping. Ascertainment of individuals and family members and exclusion criteria were as previously described (6). From this sample we selected the 53 original kindreds used to identify the linkage signal on chromosome 9 (6). We utilized all siblings within a sibship and classified the individuals as affected, unaffected, or unknown as previously described (6).

Statistical Methods. Genotypes from the *TGFBR1* locus (genotyping details available from the authors) along with the markers from the original genome scan (6) were used to obtain multipoint probabilities that sib-pairs shared zero, one, or two alleles identical by descent. These were used to determine whether study subjects were ''linked'' or ''unlinked'' to chromosome 9q22.2-31.2. In the absence of linkage, all sibling pairs are expected to share one allele identical by descent; in the presence of linkage, we expect increased allele sharing between concordant sib-pairs (i.e., >1) and decreased allele sharing between discordant sib-pairs (i.e., <1). We considered affected individuals as linked if they shared two alleles identical by descent with another affected sibling with probability of $\geq$0.85 or shared zero or one allele with an unaffected sibling with probability of 1 (n = 36). We classified affected individuals as unlinked, if they shared zero or one allele identical by descent with another affected sibling with probability of $\geq$0.95 or if they shared two alleles identical by descent with an unaffected sibling with a probability of $\geq$0.70 (n = 16). A final group comprised the individuals for whom identity by descent could not be so clearly distinguished, and these individuals were classified as unknown (n = 63).

We refined our Haseman-Elston regression analysis (9-11) by including the number of *TGFBR1*6A* alleles (0, 1, 2, 3, or 4) for each pair of sibs as a covariate in a multiple regression equation to determine if the *TGFBR1*6A* genotypes could account for the observed linkage signal. This provides a direct means to test for allelic association that, if there is linkage, is attributable to *TGFBR1*6A* or an allele in linkage disequilibrium with it. Reported P values were confirmed by comparison to a Monte Carlo sample of the permutation distribution created by permuting the allele sharing values relative to the pair labels (concordant or discordant). The number of permutations done, both across sibships of the same size and within sibships (always between 2,308 and 126,580), was sufficient to assure with 95% confidence that the estimated P value was within 5% of the true P value. Family-based tests of association (12-14) were also done.

Results and Discussion

The *TGFBR1*6A* allele was identified in 21% of affected siblings and in 28% of unaffected siblings (Table 1). The distribution of the *TGFBR1*6A* genotypes in the linked and unlinked affected individuals showed that proportionately more 9A/6A heterozygotes (4 of 16, 25%) were in the affected unlinked group than in the affected linked group (4 of 36, 11%; Table 1). This proportion changed little when individuals with small adenomas or late onset colon cancer were classified as affected (25% unlinked group versus 16% linked; see Table 1). Among all affected and unaffected individuals (parents and offspring) there is no significant difference in the percentage of subjects with one or more *TGFBR1*6A* alleles (Table 1). In fact, more siblings who had normal colon screening examinations had one or more *TGFBR1*6A* alleles than the siblings affected with colorectal cancer or advanced adenomas (28% versus 21%, respectively).

Inclusion of the *TGFBR1* genotypes into the Haseman-Elston regression analysis resulted in a reduction of the linkage signal at D17S1786, which is the most significant marker from the

Received 11/16/06; revised 2/15/07; accepted 2/21/07.

Grant support: USPHS grants U01 CA82901 (S.D. Markowitz), P30CA43703 (Case Western Reserve University/Ireland Comprehensive Cancer Center), K23 1K23CA81308 (G.L. Wiesner), P41 RR0365 and GM28356 (R.C. Elston), and DAMD17-03-1-0289 (D. Daley).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Denise Daley, 166 Burrard Building, 1081 Burrard Street, Vancouver, BC, Canada V6Z1Y6. Phone: 604-682-2344; Fax: 604-806-8351. E-mail: ddaley@mrl.ubc.ca

doi:10.1158/1055-9965.EPI-06-0965

Table 1. Distribution of genotypes among linked and unlinked affected sibs and among affected and unaffected individuals

Subgroup	9A/9A	9A/6A	6A/6A	Total
Genotypes in affected individuals				
Linked	31 (86%)	4 (11%)	1 (3%)	36
Unlinked	12 (75%)	4 (25%)	0	16
Unknown	53 (87%)	8 (13%)	0	61*
Total	96 (85%)	16 (14%)	1 (1%)	113
Distribution including individuals with small ademonas or late onset colon cancer as affected individuals				
Linked	42 (82%)	8 (16%)	1 (2%)	51
Unlinked	21 (75%)	7 (25%)	0	28
Unknown	33 (97%)	1 (3%)	0	34
Total	96 (85%)	16 (14%)	1 (1%)	113
Genotypes in affected and unaffected individuals				
All affected individuals	90	24	1	115 (27%) †
All unaffected individuals	36	12	0	48 (25%) †
Affected siblings	88	23	1	112 (21%) †
Unaffected siblings	26	10	0	36 (28%) †

*Two individuals in the unknown subgroup failed for genotyping at the *TGFBR1* gene.
† Indicates percentage of sample with one or more *6A* alleles.

original genome scan and is within 2cM of *TGFBR1* (Table 2a and b) rather than an increase, as would be expected for a susceptibility variant. To determine whether including individuals with small adenomas (n = 27) or colon cancer after the age of 65 (n = 2) would change our results, we repeated the analysis with them as affected. This resulted in further loss of signal at the *TGFBR1*6A* site (P = 0.16). The results of adding the number of *TGFBR1*6A* alleles as a pair-specific covariate in the Haseman-Elston regression analysis are shown in Table 2c. The analysis encapsulates the evidence for association/linkage disequilibrium being provided by the *TGFBR1* genotypes: when adding the pair-specific covariate, the β estimate for allele sharing reflects the remainder of linkage that is not accounted for by the covariate. If the *TGFBR1*6A* allele were the sole disease-causing variant in our sample, all the evidence for linkage would be captured by the covariate as association/linkage disequilibrium, resulting in a loss of linkage significance at D9S1786 but significance of the covariate *TGFBR1*. Rather than this scenario, the Haseman-Elston regression coefficient for the allele sharing at D9S1786 remained almost the same after inclusion of the covariate (Table 2c). Furthermore, relaxing the definition of unlinked (sharing zero or one allele identical by descent with another affected sibling with probability of $\geq$0.70 or sharing two alleles identical by descent with an unaffected sibling with a probability of $\geq$0.65) to increase their numbers made little difference. Thus, the evidence for linkage to this chromosomal region is provided by genotypes at a location other than that of the *TGFBR1*6A/9A* genotypes. All family-based tests of association were also negative (data not shown). One might question the power of the CNNS study to exclude the *TGFBR1*6A* allele, given that a recent metaanalysis case yielded an estimated odds ratio of 1.2 (4). From a pure association point of view (ignoring linkage), we have little power to exclude this locus. However, Blackwelder and Elston (15) showed that only 70 concordant sib-pairs are necessary for 95% power, given a dominant model with a population prevalence of 1% when $\theta = 0$ (i.e., directly at the disease locus, as is our hypothesis for the *TGFBR1*6A* allele). With 93 concordant sibling pairs we have 99% power to detect linkage at the disease locus. This assumes that *TGFBR1*6A* is the sole disease-causing variant in our sample, which is the hypothesis being tested. Although our analysis cannot exclude the possibility of the *TGFBR1*6A* allele increasing the risk of colorectal neoplasia in some populations, it can definitely be excluded as a susceptibility allele at the nearby linked locus we found on chromosome 9. A recent study validating linkage of colorectal cancer to chromosome 9q also excluded the *TGFBR1*6A* allele as a disease-causing variant in that sample (8). We conclude that there remains an unidentified susceptibility locus in the region 9q22.2-31.2.

Acknowledgments

We thank Christine Ticknor for the development of the assay used to genotype and distinguish the TGFBR1*6A and *9A alleles.

References

1. Pasche B, Luo Y, Rao PH, et al. Type I transforming growth factor β receptor maps to 9q22 and exhibits a polymorphism and a rare variant within a polyalanine tract. Cancer Res 1998;58:2727–32.
2. Pasche B, Kolachana P, Nafa K, et al. TβR-I (6A) is a candidate tumor susceptibility allele. Cancer Res 1999;59:5678–82.
3. Kaklamani VG, Hou N, Bian Y, et al. TGFBR1*6A and cancer risk: a meta-analysis of seven case-control studies. J Clin Oncol 2003;21:3236–43.
4. Pasche B, Kaklamani V, Hou N, et al. TGFBR1*6A and cancer: a meta-analysis of 12 case-control studies. J Clin Oncol 2004;22:756–8.
5. Bian Y, Caldes T, Wijnen J, et al. TGFBR1*6A may contribute to hereditary colorectal cancer. J Clin Oncol 2005;23:3074–8.
6. Wiesner GL, Daley D, Lewis S, et al. A subset of familial colorectal neoplasia

Table 2. Haseman-Elston regression linkage analysis of 9q22.2-31.2 in 53 kindreds with colon cancer and colon adenomatous polyps

Marker	(*a*) Original estimates*			(*b*) Including *TBFBR1* genotypes †			(*c*) Including *TGFBR1*6A* alleles (0, 1, 2, 3, or 4) as a pair-specific covariate at D9S1786					
	β ‡	SE§	P	β	SE	P	β for D9S1786	SE	P	β for the covariate ‡	SE	P
D9S283	0.1584	0.0583	0.0036	0.1584	0.0585	0.003	NA	NA	NA	NA	NA	NA
D9S1786	0.1796	0.0539	0.0005	0.1684	0.0552	0.001	0.1784	0.0559	0.0008	0.0032	0.0271	0.906
TGFBR1‖	0.1772	0.0554	0.0008	0.1297	0.0557	0.012	NA	NA	NA	NA	NA	NA

Haseman-Elston regression under the following conditions

NOTE: (*a*) original multipoint identical by descent estimates without the *TGFBR1*6A* and *TGFBR1*9A* genotypes. (*b*) multipoint identical by descent estimates with the *TGFBR1*6A* and *TGFBR1*9A* genotypes. (*c*) regression analysis including *TGFBR1*6A* alleles as a covariate.
Abbreviation: NA, not available.
*Analysis included the 17 polymorphic markers used in the original linkage study of this region to estimate the Haseman-Elston regression coefficient β.
†Analysis using the original 17 polymorphic and the *TBFBR1*6A* and *9A* genotypes to estimate β (7).
‡The Haseman-Elston regression coefficient β estimates the amount of trait variation due to a gene at this location. The covariate regression coefficient β in (*c*) estimates the effect of the number of *TBFBR1*6A* alleles from 0 to 4 that are carried by a pair of sibs on being affected.
§SE associated with the estimate of β.
‖The allele sharing for *TGFBR1* is interpolated based upon the allele sharing at the 17 polymorphic markers but does not actually use the *TGFBR1* genotypes in the original estimates.

kindreds linked to chromosome 9q22.2–31.2. Proc Natl Acad Sci U S A 2003; 100:12961–5.
7. Skoglund J, Djureinovic T, Zhou XL, et al. Linkage analysis in a large Swedish family supports the presence of a susceptibility locus for adenoma and colorectal cancer on chromosome 9q22.32-31.1. J Med Genet 2006;43:e7.
8. Kemp ZE, Carvajal-Carmona LG, Barclay E, et al. Evidence of linkage to chromosome 9q22.33 in colorectal cancer kindreds from the United Kingdom. Cancer Res 2006;66:5003–6.
9. Haseman JK, Elston RC. The investigation of linkage between a quantitative trait and a marker locus. Behav Genet 1972;2:3–19.
10. Cardon LR. A sib-pair regression model of linkage disequilibrium for quantitative traits. Hum Hered 2000;50:350–8.
11. Fulker DW, Cherny SS, Sham PC, Hewitt JK. Combined linkage and association sib-pair analysis for quantitative traits. Am J Hum Genet 1999;64: 259–67.
12. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 1993;52:506–16.
13. Spielman RS, Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium. test. Am J Hum Genet 1998;62:450–8.
14. Li H, Fan J. A general test of association for complex diseases with variable age of onset. Genet Epidemiol 2000;19 Suppl 1:S43–9.
15. Blackwelder WC, Elston RC. A comparison of sib-pair linkage tests for disease susceptibility loci. Genet Epidemiol 1985;2:85–97.