Family-based association analysis validates chromosome 3p21 as a putative nasopharyngeal carcinoma susceptibility locus

Zhaoyang Zeng, PhD[1,], Yanhong Zhou, PhD[1,*], Wenling Zhang, PhD[1,*], Xiaoling Li, PhD[1,*], Wei Xiong, PhD[1], Huaying Liu, PhD[1], Songqing Fan, PhD[1], Jun Qian, PhD[1], Lili Wang, PhD[1], Zheng Li, PhD[1], Shourong Shen, BS[2], and Guiyuan Li, PhD[1]*

Purpose: Nasopharyngeal carcinoma (NPC) poses one of the serious health problems in southern Chinese, with an incidence rate ranging from 15 to 50/100,000. In our previously linkage analysis, a locus on 3p21 was identified to link to NPC. In this study, family-based association analysis was performed to test the transmission disequilibrium of chromosome 3p in 18 high-risk nasopharyngeal carcinoma families of Hunan province in southern China. **Methods:** Single locus and multi-point of transmission disequilibrium test was performed by Genehunter program package with 15 microsatellite markers on chromosome 3p in 18 nasopharyngeal carcinoma pedigrees. **Results:** A major transmission disequilibrium peak was observed near D3S1568, which possessed 20 alleles or haplotypes of 6 loci, spanning a 12.4 cM region from D3S1298 to D3S1289 on chromosome 3p21.31-3p21.2, and 3 alleles or haplotypes reached high significantly difference ($P < 0.01$). **Conclusion:** These results reflected a link disequilibrium between this chromosome region and a nasopharyngeal carcinoma susceptibility locus, and provided further evidence that a novel nasopharyngeal carcinoma susceptibility gene may be located in this chromosome region. These alleles or haplotypes transmitting disequilibrium in nasopharyngeal carcinoma pedigrees may act as the highly risk molecular markers after verified in large population. ***Genet Med* 2006:8(3):156–160.**

Key Words: *nasopharyngeal carcinoma, chromosome 3p, transmission disequilibrium test*

Nasopharyngeal carcinoma (NPC), one of the most common malignant tumors in Southeast Asia and southern China, shows regional and familial clustering as other human cancers.[1] Epidemiological studies suggest that 5–10% of this familial aggregation derives from inherited susceptibility, which implies that genetic factors play an important role in the pathogenesis of NPC. The crucial etiologic factors involved in NPC include the Epstein–Barr virus, chemical carcinogens, radiation, structural or functional mutation of oncogenes and tumor suppressor genes and chromosomal aberration.[2–5] We have performed linkage analysis of high-heterozygosity microsatellite (STR) loci on Chromosome regions of 3p, 6q and 9p in high-risk families from Hunan Province of south China,[6–9] and a locus on 3p21 was identified as a putative NPC susceptibility locus.[1]

From the [1]Cancer Research Institute, Central South University, Changsha, 410078 Peoples' Republic of China; [2]The Third Xiangya Hospital, Central South University, Changsha, 410078 Peoples' Republic of China.

Wei Xiong and Guiyuan LI, Cancer Research Institute, Central South University, 110 Xiangya Road, Changsha, Hunan 410078, Peoples' Republic of China.

Submitted for publication August 26, 2005.

Accepted for publication October 31, 2005.

**These authors contributed equally to this study.*

DOI: 10.1097/01.gim.0000196821.87655.d0

To further verify the linage of chromosome 3p and NPC, in the present study, family-based association analysis was performed to test the transmission disequilibrium in 18 families for 15 markers on this chromosome region.

MATERIALS AND METHODS

Subjects

18 high-risk NPC families are recruited from Hunan Province in southern China. Most of these families were collected by Xiangya Hospital of Central South University and Hunan Tumor Hospital, Changsha, Hunan, China. All patients were diagnosed by pathologic examination, and the age at diagnosis of NPC was confirmed from medical records or other independent sources. Thirty-six affected and 93 unaffected individuals were used in this study (Table 1). Written informed consent was obtained from all studied participants. The study was approved by the ethical review committees of the appropriate institutions. Five to ten milliliter peripheral blood was obtained from each individual. Genomic DNA was extracted according to the routine phenol-chloroform procedure, and diluted to the final concentration of 20 ng/μl.

Genotyping analysis

Primers sequences of 15 loci used in this study were obtained either from ABI PRISM® Linkage Mapping Set v2.0, or from

Table 1
Characteristics of families with NPC genotyped in this study

Category	Result
Number of families analyzed	18
Average cases per family (range)	3.417 (2–7)
Average genotyped cases per family (range)	2.917 (2–5)
Total affected individuals genotyped (male/female)	36 (18/18)
Total unaffected individuals genotyped (male/female)	93 (43/50)
Mean ± SD age at diagnosis (range)	48.36 ± 15.27 (20–84)

GDB database (http://www.gdb.org), synthesized and labeled with different fluorescent dyes (FAM or HEX) by Shanghai Bioasia biological company.

PCR was performed in a volume of 5 μl, containing 10 mmol/L Tris-HCl, pH8.3, 50 mmol/L KCl, 1.5 mmol/L $MgCl_2$, 0.2 mmol/L each of dNTPs, 1 μmol/L of each primer, 20 ng genomic DNA and 0.25 unit Hotstar Taq (Qiagen), in a GeneAmp PCR system 9700 (PE Applied Biosystems). The thermocycling began with a first stage of 95°C for 15 minutes, followed by 10 cycles of 94°C for 30 seconds, 64°C for 60 seconds (declines 0.5°C after each cycle), and 72°C for 45 seconds, then 25 cycles of 94°C for 30 seconds, 58°C for 60 seconds and 72°C for 45 seconds, completed with a final stage of 72°C for 10 minutes. 1.5 PCR products, 2.5 μl deionized formamide and 0.5 μl GeneScan® 400HD[ROX] Size standard were added to 0.5 μl Gel Loading Dye. After mixing and denaturing on 95°C for 3 minutes, the mixture was moved onto ice. 0.8 μl of the mixture was loaded on 5% polyacrylamide gel, which contained 7.5 M urea, electrophoreses on ABI PRISM ™377 DNA Sequencer, at 2000V for two hours.

Data collection and analysis

Software ABI PRISM® 377 XL Data Collection 3.0 was used in data collection. Program ABI PRISM® GeneScan® analysis 3.7 was used for adjusting the traces of electrophoresis, correcting the molecular markers and measuring the size of amplification products. Genotyping was performed by using Genotyper® 2.1, and Transmission Disequilibrium Test (TDT) was carried out by using Genehunter 2.1.[10,11]

RESULTS

Genotyping

All microsatellite loci used in this study were dinucleotide repeat polymorphism. A serial of fragments with two different basepairs were found by genotyping for each locus. For convenience, we gave each allele a serial number as their genotype according to their length by increasing 2 bp. The heterozygosity (H) of these loci was from 0.906 to 0.210 and the polymorphism information content (PIC) of these lici was from 0.898 to 0.188. D3S3560 had only two alleles, Therefore the H and PIC of this locus were low, and the H and PIC of other 14 loci were all more than 0.6. Genotype and allele distribution in subjects of some loci were mentioned in Xiong et al.[6]

Transmission disequilibrium test

Transmission disequilibrium test of single locus or haplotypes, which were consisted by two to five adjacent loci, was performed by using programs TDT, TDT2, TDT3, TDT4 and TDT5 of software Genehunter respectively.

One hundred fourteen alleles in 15 loci were observed, after performing single locus TDT analysis by program TDT, 5 alleles have positive correlation and four alleles have negative correlation with NPC, reached significant different ($P < 0.05$). In these alleles, 2 alleles (Allele 8 of D3S1298 and Allele6 of D3S3624) reached high significantly different level ($P < 0.01$) (Table 2).

A Log Odds (lods) or an NPL score peak was obtained between D3S1298 and D3S1289 analyzed by two-point or multipoint, and parametric or non-parametric linkage analysis using Linkage or Genehunter software in our previously study.[1] In this study, single-locus TDT analysis showed that located in this chromosome region, D3S1568 (alleles 5 and 10), D3S1289 (alleles 7 and 2) and D3S3624 (alleles 4 and 6) possessed two alleles correlating to NPC, respectively, and one of alleles of each locus, D3S1298 and D3S3624, reached high significantly different level. These results indicated that the TDT analysis was matched our previous linkage analysis.

Sixteen haplotypes were found associated with NPC (Table 3) by two loci TDT analysis using program TDT2. As a matter of convenience, each locus was assigned a serial number according its genetic distance to 3p-ter (Table 4), and their serial numbers represented the haplotype in Table 3.

The results of program TDT2 showed that the eighth locus, D3S1568, and its adjacent locus comprised most of the haplo-

Table 2
Loci and alleles reached significantly different level in single locus TDT analysis

Loci	Allele	Trans[a]	Untrans[a]	χ^2	*P*-value	Significance[b]
D3S1298	8	10	1	7.36	0.0067	++
D3S3624	6	17	3	9.80	0.0017	++
D3S1297	2	18	7	4.84	0.0278	+
D3S1568	5	4	0	4.00	0.0455	+
D3S1289	7	18	7	4.84	0.0278	+
D3S3553	12	5	0	5.00	0.0253	+
D3S3681	5	4	0	4.00	0.0455	+
D3S3624	4	0	4	4.00	0.0455	−
D3S1300	7	3	13	6.25	0.0124	−
D3S1568	10	2	10	5.33	0.0209	−
D3S1289	2	0	4	4.00	0.0455	−

[a]Trans, transmitted alleles or haplotypes; Untrans, untransmitted alleles or haplotypes.
[b]+ Positive correlation between NPC and allele or haplotype; − negative correlation ($P < 0.05$); ++/−high correlation ($P < 0.01$).

Table 3
Haplotypes reached significantly different level in two loci TDT2 analysis

Loci	Haplotype	Trans	Untrans	χ^2	*P*-value	Significance
1–2	5–3	0	6	6.00	0.0143	−
2–3	5–5	4	0	4.00	0.0455	+
5–6	8–9	4	0	4.00	0.0455	+
6–7	3–6	6	0	6.00	0.0143	+
7–8	5–6	0	4	4.00	0.0455	−
7–8	5–10	0	4	4.00	0.0455	−
7–8	6–11	7	0	7.00	0.0082	++
8–9	5–2	4	0	4.00	0.0455	+
8–9	6–2	0	4	4.00	0.0455	−
9–10	2–2	0	4	4.00	0.0455	−
11–12	3–5	4	0	4.00	0.0455	+
11–12	4–12	5	0	5.00	0.0253	+
12–13	5–5	8	1	5.44	0.0196	+
12–13	8–6	4	0	4.00	0.0455	+
13–14	7–3	0	6	6.00	0.0143	−
14–15	2–5	4	0	4.00	0.0455	+

Trans, transmitted alleles or haplotypes; Untrans, untransmitted alleles or haplotypes.
+ Positive correlation between NPC and allele or haplotype; − negative correlation ($P < 0.05$); ++/−− high correlation ($P < 0.01$).

types, D3S3624-D3S1568 (7–8) and D3S1568-D3S3560, (8–9) associated with NPC (5 haplotypes), and haplotype D3S3624 (allele 6)—D3S1568 (allele 11) reached high significant different level (++, $P < 0.01$).

Haplotypes consisting of three adjacent loci were found by the TDT analysis with program TDT3, 11 haplotypes comprised of seven loci correlated to NPC ($P < 0.05$), but no haplotype was found to reach a highly significant different level ($P < 0.01$) (Table 5).

Similarly, D3S1568 and its adjacent loci possessed most NPC associated haplotypes including D3S3624-D3S1568-D3S3560 (7-8-9) and D3S1568-D3S3560-D3S1289 (8-9-10).

Three haplotypes which consisted of four adjacent loci reached significantly different levels, including haplotypes comprised of D3S1568 (D3S3624-D3S1568-D3S3560-D3S1289) (Table 6). But no haplotype, which comprised of five loci, reached significantly different levels (date not shown).

As a matter of convenience and obviousness, we summarized the above-mentioned results of TDT analyses of 15 loci in one table (Table 4), range them by their chromosome position. In this table, positive or negative correlation between allele or haplotype and NPC was not considered. If one allele or one haplotype correlated to NPC, an asterisk (*) was marked on the loci or the central position of the haplotype, if an allele or a haplotype reached high significant difference ($P < 0.01$), two asterisks were marked. If several alleles or haplotypes correlate to NPC, the asterisk would be added up.

Table 4
TDT analysis results of fifteen loci

Number	Loci	Chromosome	Genetic distance 3p-ter	TDT	TDT2	TDT3	TDT4
1	D3S1297	3p26.2	2.5	*			
					*		
2	D3S1489	3p25.3	21.4			*	
					*		*
3	D3S1266	3p22.3	46.9			*	
4	D3S3727	3p22.3	50.4			*	
5	D3S1298	3p21.31	56.7	**			
					*		
6	D3S3564	3p21.31	63.3				
					*		
7	D3S3624	3p21.31	65.1	***			
					****		
8	D3S1568	3p21.31	67.8	**		***	
					**		*
9	D3S3560	3p21.31	67.9			***	
					*		
10	D3S1289	3p21.2	69.1	**			
11	D3S1582	3p21.2	70.3				
					**		
12	D3S3553	3p21.1	75.8	*			
					**		
13	D3S1300	3p21.1	79.0	*		*	
					*		*
14	D3S1285	3p14.3	91.0			*	
					*		
15	D3S3681	3p12.1	111.0	*			

*One allele or one haplotype correlated to NPC.
**An allele or haplotye reached high significant difference.

As shown in Table 4, the distribution of the asterisks appeared as three peaks: A major TDT peak was observed near D3S1568, which possessed 20 alleles or haplotypes in six loci spanning 12.4 cm from D3S1298 to D3S1289. 3 of 20 alleles or haplotypes reached high significant difference. These 20 alleles or haplotypes reflected the linkage disequilibrium between this chromosome region and NPC, and these results provided further evidence that a novel NPC susceptibility gene may be located in this chromosome region. In addition, two minor TDT peaks were observed near D3S1489 and D3S1300, respectively, and it was consistent with our previous results of linkage analysis also.[1]

Table 5
Haplotypes reached significantly different level in three loci TDT3 analysis

Loci	Haplotype	Trans	Untrans	χ^2	*P*-value	Significance
1–2–3	5–3–6	0	4	4.00	0.045500	−
2–3–4	3–6–8	0	5	5.00	0.025347	−
3–4–5	2–2–10	0	5	5.00	0.025347	−
7–8–9	5–6–2	0	4	4.00	0.045500	−
7–8–9	5–10–2	0	4	4.00	0.045500	−
7–8–9	6–11–2	4	0	4.00	0.045500	+
8–9–10	6–2–7	0	4	4.00	0.045500	−
8–9–10	9–2–4	0	4	4.00	0.045500	−
8–9–10	11–2–5	4	0	4.00	0.045500	+
12–13–14	6–7–2	0	4	4.00	0.045500	−
13–14–15	7–2–4	0	4	4.00	0.045500	−

Trans, transmitted alleles or haplotypes; Untrans, untransmitted alleles or haplotypes.
+ Positive correlation between NPC and allele or haplotype; - negative correlation ($P < 0.05$).

Table 6
Haplotypes reached significant difference in four loci TDT4 analysis

Loci	Haplotype	Trans	Untrans	χ^2	*P*-value	Significance
1–2–3–4	5–3–6–8	0	4	4.00	0.0455	−
7–8–9–10	5–6–2–7	0	4	4.00	0.0455	−
12–13–14–15	6–7–2–4	0	4	4.00	0.0455	−

Trans, transmitted alleles or haplotypes; Untrans, untransmitted alleles or haplotypes. − Negative correlation ($P < 0.05$).

DISCUSSION

Because of its broad distribution on human genome, and its high heterozygosity as well as polymorphism information content, microsatellite is widely used in fine genetic linkage mapping, locating disease genes, individual distinguishing, parentage identification, etc. It also made great progress in locating and cloning the susceptibility genes of malignant tumors.[1,12–14]

In this study, we genotyped the 15 microsatellite loci on the short arm of chromosome 3 in 18 NPC families collected from Hunan Province, and TDT analysis was performed. The results of TDT analysis were consistent with our previous linkage analysis. Near D3S1568, spanning from D3S1298 to D3S1289, 20 NPC-associated alleles or haplotypes were observed, and the locus D3S1568 is exactly the region containing the lods score peak of our previously linkage analysis.[1] It reflected that one or more novel NPC-associated susceptibility genes, especially correlating with Hunan familial NPC, locates in the region near locus D3S1568, in which there was a 630 Kb homozygous deletion of cancer, and many candidate tumor suppressor genes were identified.[15]

We also found alleles or haplotypes in other loci, such as D3S1297-D3S1489-D3S1266 and D3S1300-D3S1285-D3S3681, were associated to NPC. In our previous linkage analysis, in addition to the major peek on D3S1568, two minors lods score peaks on D3S1489 and D3S1300 were observed. Though the lods or NPL scores of these loci did not reach a significant level, it implies that one or more genes contribute minor effects on the development of NPC that may be located in this region. Near D3S1489, NPC associated gene NAG7 was cloned. Preliminary function analysis of this gene shows that it may play a certain role in the development of NPC.[16–18]

Recently, a high lods score was obtained on chromosome 4p15.1-q12 by employing a Genome-wide scan in 20 Cantonese nasopharyngeal carcinoma families, which indicated that there exists a susceptibility gene in this region.[14] In addition, previous studies suggested that loci HLA-A2, HLA-B17, HLA-BW46,[19,20] D6S1581[7] and polymorphisms of CYP450,[21] NGX6,[22] UBAP1,[23] NOR1[24] genes were associated with NPC. In our opinion, considering the complicated pathogenesis resulted from heterogeneity and the environmental factors as well as multiple genes involved in NPC, it is acceptable that different results were obtained in the studies of different geography region and subjects.

ACKNOWLEDGMENTS

This work was supported by grants from the National Natural Science Foundation of China (30300201, 30470955, 30200312, 30300063, 30330560), the Program for New Century Excellent Talents in University (NCET-04-0761) the Special Fore Funds for Major State Basic Research of China (No. 2005 CCA 03200 and the Special Funds of Science & Technology Department of Hunan Province (05SK1001-1, 04SK2002). We are extremely grateful to the many patients, physicians, and volunteers in Southern China who willingly contributed to these studies.

References

1. Xiong W, Zeng ZY, Xia JH, Xia K, et al. A susceptibility locus at chromosome 3p31 linked to family nasopharyngeal carcinoma. *Cancer Research* 2004;64:1972–1974.
2. Zhan F, Jiang N, Cao L, Deng L, et al. Primary study of differentially expressed cDNA sequences in cell line HNE1 of human nasopharyngeal carcinoma by cDNA representational difference analysis. *Chin J Med Genet* 1998;15:341–344.
3. Jiang N, Zhan F, Xie Y, Zeng Z, et al. Establishment of partial gene expression map of 7q32 in nasopharyngeal carcinoma and primary culture normal nasopharyngeal epithelial cells. *Chin J Med Genet* 1998;15:267–270.
4. Tang X, Yang J, Deng L, Jiang N, et al. Detailed allelic loss mapping on 7q32 in nasopharyngeal carcinoma. *Chin J Med Genet* 2000;17:153–156.
5. Li Z, Wang L, Zhang X, Zhang L, et al. Chromosomal aberration analyzed by comparative genomic hybridization in nasopharyngeal carcinoma. *Chin J Med Genet* 2001;18:338–342.
6. Xiong W, Zeng ZY, Xiong F, Shen SR, et al. Genetic polymorphism of 8 STR loci on short arm of chromosome 3. *Chin J Med Genet* 2003;20:413–416.
7. Xiong W, Zeng ZY, Shen SR, Li XL, et al. Studies on the relationship between D6S1581, a high frequency allele imbalance locus, and genetic susceptibility to nasopharyngeal carcinoma. *Chin J Med Genet* 2003;20:311–314.
8. Xiong W, Zeng ZY, Shen SR, Li XL, et al. Polymorphism of two novel SNPs, which locate on chromosome 9p21-22, in Han Chinese of Hunan. *Chin J Med Genet* 2003; 20:203–206.
9. Zeng ZY, Xiong W, Xiong F, Shen SR, et al. Information Behavior of 7 STR Loci on Chromosome 9p in Gene Scanning. *Heredity (Beijing)* 2003;25:543–548.
10. Lathrop GM, Lalouel JM, Julier C, Ott J. Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. *Am J Hum Genet* 1985;37: 482–498.
11. Ott J, editor. Analysis of Human Genetic Linkage (revised edition). Baltimore: Johns Hopkins University Press, 1991.
12. Hall JM, Lee MK, Newman B, Morrow JE, et al. Linkage of early-onset familial breast cancer to chromosome 17q21. *Science* 1990;250:1684–1689.

13. Smith JR, Freije D, Carpten JD, Gronberg H, et al. Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. *Science* 1996; 274:1371–1374.
14. Feng BJ, Huang W, Shugart YY, Lee MK, et al. Genome-wide scan for familial nasopharyngeal carcinoma reveals evidence of linkage to chromosome 4. *Nat Genet* 2002;31:395–399.
15. Lerman MI, Minna JD. The 630-kb lung cancer homozygous deletion region on human chromosome 3p21.3: identification and evaluation of the resident candidate tumor suppressor genes. The International Lung Cancer Chromosome 3p21.3 Tumor Suppressor Gene Consortium. *Cancer Res* 2000;60:6116–6133.
16. Tan C, Li J, Wang J, Xiang Q, et al. Proteomic analysis of differential protein expression in human nasopharyngeal carcinoma cells induced by NAG7 transfection. *Proteomics* 2002;2:306–312.
17. Tan C, Peng C, Huang YC, Zhang QH, et al. Effects of NPC-associated gene NAG7 on cell cycle and apoptosis in nasopharyngeal carcinoma cells. *Chin J Cancer* 2002; 21:449–455.
18. Tan C, Li J, Peng C, Zhang QH, et al. The effects of NAG7 on gene expressional profile of HNE1 cells using cDNA microarray analysis. *Prog Biochem biophys* 2003;30:99–106.
19. Henderson BE, Louie E, SooHoo Jing J, Buell P, et al. Risk factors associated with nasopharyngeal carcinoma. *N Engl J Med* 1976;295:1101–1106.
20. Lu SJ, Day NE, Degos L, Lepage V, et al. Linkage of a nasopharyngeal carcinoma susceptibility locus to the HLA region. *Nature* 1990;346:470–471.
21. He ZM, Yuan JH, Wang SL, Lai JP, et al. Association between genetic polymorphism of human cytochrome P450 2E1 gene and susceptibility to nasopharyngeal carcinoma. *Chin J Cancer* 1999;18:517–519.
22. Xiong W, Zeng ZY, Li XL, Li WF, et al. Single-nucleotide polymorphisms in NGX6 gene and their correlation with Nasopharyngeal carcinoma. *Acta Bioch Bioph Sin* 2002;34:512–515.
23. Xiong W, Zeng ZY, Shen SR, Li XL, et al. Studies of Single-Nucleotide Polymorphisms in UBAP1 Gene and Association with Nasopharyngeal Carcinoma. *Prog Biochem and Biophys* 2002;29:766–770.
24. Xiong W, Zeng ZY, Xiao BY, Xiong F, et al. Studies of Association between Single-Nucleotide Polymorphisms in NOR1, a novel oxidored-nitro domain-containing protein Gene, and Nasopharyngeal Carcinoma. *Prog Biochem and Biophys* 2003;30: 401–405.