

ARTICLE

Complex segregation analysis of nasopharyngeal carcinoma in Guangdong, China: evidence for a multifactorial mode of inheritance (complex segregation analysis of NPC in China)

Wei-Hua Jia^{1,5}, Andrew Collins^{2,5}, Yi-Xin Zeng^{*1,5}, Bing-Jian Feng¹, Xing-Juan Yu¹, Li-Xi Huang¹, Qi-Sheng Feng¹, Ping Huang¹, Ming-Hong Yao³ and Yin Yao Shugart^{4,5}

¹Department of Experimental Research, Cancer Center, Sun Yat-sen University, GuangZhou, The Key Laboratory of Ministry of Education, China; ²Human Genetics, University of Southampton, Southampton General Hospital, UK; ³Department of Information, Cancer Center, Sun Yat-sen University, GuangZhou, The Key Laboratory of Ministry of Education, China; and ⁴Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, USA

The striking geographical and ethnic distribution of nasopharyngeal carcinoma (NPC) suggests the involvement of genetic and environmental factors in NPC development. The purpose of this study is to investigate the fit of single gene, polygenic and multifactorial models to the observed pattern of transmission of NPC in a hospital-based family history study conducted by the Cancer Center of Sun Yat-Sen University (CCSYU) in Guangzhou, China. Complex segregation analysis of a total of 1903 Cantonese pedigrees ascertained at CCSYU was conducted using a unified mixed model after the pedigrees were partitioned into 3737 nuclear families. The mixed model assumes that a phenotype is influenced by the additive and independent effect of a major gene, together with multifactorial components (genetic and environmental) and a random environmental effect. The current results do not provide evidence for a major gene and the observed data are best explained by a multifactorial mode of inheritance for NPC.

European Journal of Human Genetics (2005) 13, 248–252. doi:10.1038/sj.ejhg.5201305

Published online 13 October 2004

Keywords: nasopharyngeal carcinoma (NPC); complex segregation analysis; mode of inheritance; multifactorial model

Introduction

The incidence rates of nasopharyngeal carcinoma (NPC) are low throughout most of the world, typically less than 1 per 100 000.¹ However, NPC is common in Southeast Asia and Northern African countries. Moreover, NPC is among the leading causes of death in Cantonese and several other

southern Chinese populations.² In fact, the incidence rate per 100 000 for males is higher than 40 in some Cantonese-speaking populations in Guangdong province.

Familial clustering of NPC has been widely documented in the Chinese population.^{3–9} One rare documented example⁷ is one family living in Nanhai County of Guangdong Province having 15 NPC affected family members, two family members affected with hepatic cell carcinomas and one with breast cancer. In southern Chinese populations, several case–control studies have confirmed that the proportion of NPC cases with a positive first-degree family history is greater than 5%, with similar

*Correspondence: Dr Y-X Zeng, Cancer Center, Sun Yat-sen University, 651 DongFeng Road East, GuangZhou 510060, China.

Tel: +86 1380 9770 743; Fax: +86 20 8734 3295;

E-mail: yxzeng@gzsums.edu.cn

⁵These authors contributed equally to the manuscript.

Received 22 April 2004; revised 22 July 2004; accepted 18 August 2004

proportions found in high-risk areas such as Hong Kong (7.2%), Yulin (6.0%), and Guangzhou (5.9%).^{3,4} All of these data suggest an important role for familial patterns of inheritance in NPC pathogenesis. Complex segregation analysis has been successfully used to define the model of inheritance in esophageal and colorectal cancers.^{10,11} However, to date, there have been no large-scale complex segregation analyses to elucidate the genetic mechanisms underlying carcinogenesis in NPC patients.

This study aims to investigate the fit of single gene, polygenic and multifactorial models to the observed pattern of transmission of NPC in a hospital-based family study conducted by Cancer Center of Sun Yat-Sen University (CCSYSU). A total of 1903 Cantonese pedigrees ascertained through probands treated at the CCSYSU were enrolled, and complex segregation analysis was performed using the program POINTER.¹²

Subjects and methods

From January 1999, we interviewed Cantonese NPC patients receiving treatment in the Department of NPC at CCSYSU, which is the largest center for cancer prevention and treatment in southern China. After we received permission from their physicians, we contacted the patients regarding participation in the study and we then interviewed all consented patients. By June 2001, 1903 Cantonese cases were enrolled in the study.

Each patient's diagnosis was pathologically confirmed as NPC. Each consenting NPC patient was regarded as a proband, and all families were ascertained through a single proband, with the exception of two families for which two probands were independently ascertained. In these two cases, where more than one family member was being treated for NPC, a single questionnaire was administered to avoid duplication of family information.

A structured questionnaire was administered to the proband of each family, where patients were asked to provide verbal answers. Information was gathered on the probands themselves, as well as the family history of three successive generations (first- and second-degree relatives). Whenever probands were unsure of their answers, we interviewed additional family members to ensure the accuracy of the information. Questions included the proband's contact information, date of birth, sex, date of diagnosis, birthplace, birthplace of grandparents, his or her physician, dialect and other general characteristics. The family history included any incidence of cancer in the proband's relatives, including cause and date of death for deceased relatives, status of cancer occurrence, type of cancer, date of diagnosis, hospital and records of diagnosis. The age and gender information of the probands are summarized in Table 1.

Complex segregation analysis was performed using the unified mixed model.¹² The mixed model assumes that a

Table 1 Descriptive information for the NPC probands

Age group	Male	Proportion	Female	Proportion
<20	8	0.56	6	0.92
20–29	103	6.45	62	9.47
30–39	392	24.55	211	32.21
40–49	564	35.32	225	34.35
50–59	370	23.17	105	16.03
>60	159	9.96	46	7.02
Total	1597	100	655	100

phenotype is influenced by the additive and independent effect of a major gene, together with a multifactorial component (genetic and environmental) and a random environmental effect. Pedigrees were partitioned into nuclear families and analyzed using the program POINTER under joint likelihood.¹² Nuclear families, including all first- and second-degree relatives in the sample, were coded and totaled 3747 families.

The ascertainment probability, as used in POINTER, is ~ 0 if the probability of ascertaining a family increases in proportion to the number of affected offspring (single selection) and close to 1 if the probability of ascertaining a family is independent of the number of affected offspring (complete selection). Where appropriate, first-degree relatives of the proband were partitioned into nuclear families containing the proband as a parent (complete selection) and as a child (incomplete selection). Nuclear families of second-degree relatives containing affected individuals were coded for truncate selection (ascertainment probability = 1). For incomplete selection only two families in this large sample had more than one independently ascertained proband, suggesting a relatively low ascertainment probability, but not single selection. We therefore used a value of 0.1 but we also examined all models while varying the ascertainment probability over the range 0.001–0.2, and the results were found to be highly robust to changes in the specified ascertainment model (data not shown).

All individuals with known age (at the time of ascertainment) were assigned to liability classes according to age and gender. Liability classes and their frequencies were based on the observed cumulative risk of NPC over 10-year age intervals from a Hong Kong cancer database over the period 1990–1994.¹³ Such information is not available in the Cantonese population in Guangdong, and the Hong Kong Cantonese population is an appropriate representative population to Guangdong Cantonese.

The POINTER program permits the construction of four male and four female liability classes, which describe age-specific risks. Therefore, eight classes were formed according to the age ranges given in Table 2. Cumulative incidence figures, to the mid-point of each class, were calculated given the rates per 100 000 as described in

Lee *et al* (2003).¹⁴ As any gene causing specific mortality must decrease in frequency with age, we determined risks using the approach described by Iselius *et al*.¹⁴ Using the mortality figures from Lee *et al*,¹³ we calculated cumulative mortalities (Table 2) and risk, R , attributed to the j th liability class as: $R_j = (I_j - M_{j-1}) / (1 - M_{j-1})$, where I_j is the cumulative incidence to the mid-point and M_{j-1} is the cumulative specific mortality to the end of the preceding class. Although some data on age at the onset of disease are available, the data are incomplete and possibly subject to error because of the difficulty in defining onset. Further studies might usefully distinguish early and late onset as a route to discriminating between genetic and environmental etiology.

The parameters estimated in the analysis are major locus dominance (d , where $d=0$ refers to a recessive major gene, $d=0.5$ additive and $d=1$ dominant), difference on the liability scale between homozygotes (t), and gene frequency (q).¹² There are two parameters to test multifactorial heritability. The parameter H represents polygenic heritability in children, where $H = Ck/V$, in which Ck is the multifactorial component and V is the overall variance. The second parameter is Z for which HZ represents the multifactorial heritability in parents, where $Z = Ca/Ck$, the ratio of the multifactorial component in adults and children. Significant deviation of Z from 1 suggests a

Table 2 Gender-specific cumulative incidence, mortality and risk

Age range	Cumulative incidence ^a	Cumulative mortality	Risk
<i>Males</i>			
<30	0.0000115	0.0000050	0.0000115
30–49	0.0004275	0.0002640	0.0004225
50–69	0.0014895	0.0009470	0.0012258
70+	0.0023505	0.0012260	0.0014048
<i>Females</i>			
<30	0.0000080	0.0000020	0.0000080
30–49	0.0001795	0.0000800	0.0001775
50–69	0.0005840	0.0003040	0.0005040
70+	0.0008985	0.0004030	0.0005947

^aCumulative incidences in the table are calculated to the middle of each class.

Table 3 Results of POINTER segregation analysis of NPC

Model	d	t	q	H	Z	$-2 \ln L+C$	AIC
Sporadic	—	—	—	—	—	–2232.2	–2232.2
Polygenic	—	—	—	0.88	1.00	–4226.8	–4224.8
Multifactorial	—	—	—	0.95	0.61	–4262.2	–4258.2
Recessive	(0)	3.52	0.0694	—	—	–3833.9	–3829.9
Dominant	(1)	2.28	0.0011	—	—	–4144.6	–4140.6
Additive	(0.5)	4.50	0.0013	—	—	–4147.4	–4143.4
Generalized single locus	0.522	4.31	0.0012	—	—	–4147.5	–4141.5

d : where $d=0$ refers to a recessive major gene, $d=0.5$ additive and $d=1$ dominant; t : difference on the liability scale between homozygotes; q : gene frequency; H : the polygenic heritability; Z : the ratio of the multifactorial component in adults and children; and AIC: Akaike's Information Criterion.

generational difference in multifactorial heritability. To test competing hypotheses, minus twice log likelihood ($-2 \ln L + C$), calculated under a more general model, is subtracted from that for a reduced model. The difference in large samples is distributed as a χ^2 with a number of degrees of freedom equal to the number of free parameters. For non-nested models, Akaike's Information Criterion (AIC) is used. AIC is computed as $-2 \ln L + C$ plus twice the number of estimated parameters. The model with the lowest AIC has the most parsimonious fit to the data.

Results

The results in Table 2 show gender-specific cumulative incidence, mortality and risk of NPC in the study population. The results from POINTER (Table 3) provide overwhelming support for a model that does not include major gene inheritance but reflects a polygenic/multifactorial pattern. The difference in AIC between the best-fitting single gene model (additive) and the polygenic model is 81.4. Furthermore, support for a multifactorial model over a polygenic model has a χ^2 (1 df) of 35.4. The difference in AIC between the multifactorial model and the best-fitting single gene model (additive) is 114.8. Attempts to fit the full mixed model (major genes plus a polygenic component) always resulted in elimination, through bounding, of the major gene so that only the polygenic component remained.

The results of the present study strongly support a multifactorial mode of inheritance with no evidence for a substantial role of major genes in the inheritance of NPC. Our results are consistent with various epidemiological and genetic studies. Below, we will briefly review a body of evidence for both environmental and genetic factors for NPC based on previous findings.

Discussion

EBV infection

Epstein–Barr virus (EBV) is a ubiquitous herpes virus that infects more than 90% of the human population and establishes a lifelong viral persistence in the host. EBV has been consistently identified as an important risk factor for

NPC, with a dose–response relationship between EBV antibodies and NPC risk.^{15–18}

Cigarette smoking

A recent IARC monograph (2002)¹⁹ reviewed existing evidence and established cigarette smoking as a risk factor for NPC based on the consistent findings of increased risk in both high- and low-risk geographic regions for NPC, the presence of a dose–response relationship with duration or amount of smoking, and a reduction in risk after quitting.

Salted fish consumption

Cantonese-style salted fish, which contains high levels of nitrosamines, has long been thought to be a risk factor for NPC.^{8,20–26} Together, these studies have suggested earlier first exposure, longer duration of consumption, higher frequency of consumption and steaming the fish instead of frying or grilling increases the risk for NPC. More recently, Ward *et al*²⁷ confirmed that nitrosamine and nitrite intake during childhood may play a role in the development of NPC.

Fruits and vegetables

The association between NPC and fruit and vegetable intake depends strongly on whether the fruits and vegetables are fresh or preserved. As reviewed by Yu and Yuan,⁸ a number of studies have reported a protective effect of higher intake of fresh fruits and vegetables. However, the findings are not always consistent. On the contrary, increased risk of NPC was observed in association with frequent adult consumption of preserved fruits and vegetables.

Genetic studies

HLA Bw46 has been shown to be associated with increased risk of NPC.^{28,29} Hildesheim *et al*³⁰ reported an association of HLA class I and II alleles (HLA-A*0207 and HLA-B*4601) and extended haplotypes with NPC in a well-designed case–control study conducted in Taiwan. Further, a meta-analysis of 13 HLA association studies in southern China performed by Goldsmith *et al*³¹ reported evidence for positive associations between NPC and the HLA alleles A2, B14 and B46 and negative associations for the alleles A11, B13 and B22. Genetic linkage study has been used as one of the most common approaches for identification of cancer susceptibility genes.³² More recently, a genome-wide search was conducted on a series of high-risk NPC families from Cantonese-speaking populations in the Pearl River area, Guangdong Province. Both multipoint parametric and nonparametric linkage analysis supported evidence for linkage.³³ The inclusion of EBV antibody titer as a covariate in the nonparametric linkage analysis gave stronger linkage signals.

The results presented here concord with those reported by Chen *et al*.³⁴ They conducted a community-based case–

control study with 347 cases and controls in Taiwan and suggested that EBV infection is a risk factor for NPC and the effect of cigarette smoking on NPC is modified by age. They estimated the heritability of NPC as 0.60 based on a multifactorial inheritance model.

Results from this study suggest that NPC is a multifactorial malignancy with complex etiology involving both genetic and environmental factors. The penetrance of the susceptibility genes could be low, just like in breast cancer.³⁵ It has become apparent that, for a multifactorial malignancy, family-based genome-wide linkage studies may have their limitations, particularly in the absence of accurately measured relevant environmental factors. One of our future goals is to obtain more complete epidemiological data including EBV infection, smoking, consumption of salted fish, and preserved vegetables on NPC families in addition to the biological material that we have collected. We will also take advantage of newly developed genetic epidemiological tools such as a covariate-based linkage analysis method³⁶ to test for linkage between the NPC risk and genetic markers that are biologically plausible, allowing for potential interaction between these genes and environmental factors. It has been shown by Goddard *et al*³⁷, Olson *et al*³⁸ and Devlin *et al*³⁹ that power to detect linkage increases when a covariate that measures heterogeneity is included in the analysis model. We are in the process of setting up a long-term collaboration among clinicians, radiologists, epidemiologists, molecular biologists as well as statistical geneticists to achieve a collection of a large number of families, documented with confirmed diagnosis, age of diagnosis, history of EBV infection, cigarette smoking and dietary factors. We remain hopeful that dissecting the complex etiology for a multifactorial trait such as NPC is feasible.

Acknowledgements

This work is supported by National Natural Science Foundation (30000141), National Key Basic Research Projects (973 Projects, G19980510) and National Science and Technology Project 2002BA711A03 of China. Yin Yao Shugart is supported, in part, by an Innovation Award from the Bloomberg School of Public Health at Johns Hopkins University.

References

- 1 Parkin DM, Muir CS: *Cancer incidence in five continents*. Lyon: IARC Sci Publ, 1997, vol 7, pp 334–337.
- 2 Jeannel D, Bouvier G, Huber A: Nasopharyngeal carcinoma: an epidemiological approach to carcinogenesis. *Cancer Surv* 1999; **33**: 125–155.
- 3 Yu MC, Ho JH, Lai SH, Henderson BE: Cantonese style salted fishes as a cause of nasopharyngeal carcinoma: report of a case–control study in Hong Kong. *Cancer Res* 1986; **46**: 956–961.
- 4 Yu MC, Garabrant DH, Huang TB, Henderson BE: Occupational and other non-dietary risk factors for nasopharyngeal carcinoma in Guangzhou, China. *Int J Cancer* 1990; **45**: 1033–1039.
- 5 Chen DL, Huang TB: A case–control study of risk factors of nasopharyngeal carcinoma. *Cancer Lett* 1997; **117**: 17–22.

- 6 Ung A, Chen CJ, Levine PH *et al*: Familial and sporadic cases of nasopharyngeal carcinoma in Taiwan. *Anticancer Res* 1999; **19**: 661–665.
- 7 Zhang F, Zhang JM: Clinical hereditary characteristics in nasopharyngeal carcinoma through Ye-Liang's family cluster. *Chin Med J* 1999; **112**: 185–187.
- 8 Yuan JM, Wang XL, Xiang YB, Gao YT, Ross RK, Yu MC: Non-dietary risk factors for nasopharyngeal carcinoma in Shanghai, China. *Int J Cancer* 2000; **85**: 364–369.
- 9 Zeng YX, Jia WH: Familial nasopharyngeal carcinoma. *Semin Cancer Biol* 2002; **12**: 443–450.
- 10 Zhang W, Bailey-Wilson JE, Li W *et al*: Segregation analysis of esophageal cancer in a moderately high-incidence area of northern China. *Am J Hum Genet* 2000; **67**: 110–119.
- 11 Houlston RS, Collins A, Kee F, Collins BJ, Shields DC, Morton NE: Segregation analysis of colorectal cancer in Northern Ireland. *Hum Hered* 1995; **45**: 41–48.
- 12 Morton NE, Rao DC, Lalouel JM: *Methods in genetic epidemiology*. Basel: Karger, Chapter 5; 1983, pp 62–98.
- 13 Lee AW, Foo W, Mang O *et al*: Changing epidemiology of nasopharyngeal carcinoma in Hong Kong over a 20-year period (1980–99): an encouraging reduction in both incidence and mortality. *Int J Cancer* 2003; **103**: 680–685.
- 14 Iselius L, Slack J, Littler M, Morton NE: Genetic epidemiology of breast cancer in Britain. *Ann Hum Genet* 1991; **55**: 151–159.
- 15 Lin TM, Chen KP, Lin CC *et al*: Retrospective study on nasopharyngeal carcinoma. *J Natl Cancer Inst* 1973; **51**: 1403–1408.
- 16 De-The G, Dubouch P, Fontaine C *et al*: Natural antibodies to EBV-VCA antigens in common marmosets (*Callithrix jacchus*) and response after EBV inoculation. *Intervirology* 1980; **14**: 284–291.
- 17 Chen JY, Liu MY, Chen CJ *et al*: Antibody to Epstein–Barr virus-specific DNase as a marker for the early detection of nasopharyngeal carcinoma. *J Med Virol* 1985; **17**: 47–49.
- 18 Chan CK, Mueller N, Evans A *et al*: Epstein–Barr virus antibody patterns preceding the diagnosis of nasopharyngeal carcinoma. *Cancer Causes Control* 1991; **2**: 125–131.
- 19 IARC monograph: *Tobacco smoking and tobacco smoke (Group 1)*. Lyon, 2002.
- 20 Ho JH, Huang DP, Fong YY: Salted fish and nasopharyngeal carcinoma in southern Chinese. *Lancet* 1978; **2**: 626.
- 21 Yu MC, Huang TB, Henderson BE: Diet and nasopharyngeal carcinoma: a case–control study in Guangzhou, China. *Int J Cancer* 1989; **43**: 1077–1082.
- 22 Yu MC, Mo CC, Chong WX, Yeh FS, Henderson BE: Preserved foods and nasopharyngeal carcinoma: a case–control study in Guangxi, China. *Cancer Res* 1988; **48**: 1954–1959.
- 23 Ning JP, Yu MC, Wang QS, Henderson BE: Consumption of salted fish and other risk factors for nasopharyngeal carcinoma (NPC) in Tianjin, a low-risk region for NPC in the People's Republic of China. *J Natl Cancer Inst* 1990; **82**: 291–296.
- 24 Sriamporn S, Vatanasapt V, Pisani P, Yongchaiyudha S, Rungpitarangsi V: Environmental risk factors for nasopharyngeal carcinoma: a case–control study in north-eastern Thailand. *Cancer Epidemiol Biomarkers Prev* 1992; **1**: 345–348.
- 25 Lee HP, Gourley L, Duffy SW, Esteve J, Lee J, Day NE: Preserved foods and nasopharyngeal carcinoma: a case–control study among Singapore Chinese. *Int J Cancer* 1994; **59**: 585–590.
- 26 Armstrong RW, Imrey PB, Lye MS, Armstrong MJ, Yu MC, Sani S: Nasopharyngeal carcinoma in Malaysian Chinese: salted fish and other dietary exposures. *Int J Cancer* 1998; **77**: 228–235.
- 27 Ward MH, Pan WH, Cheng YJ *et al*: Dietary exposure to nitrite and nitrosamines and risk of nasopharyngeal carcinoma in Taiwan. *Int J Cancer* 2000; **86**: 603–609.
- 28 Lu CC, Chen JC, Jin YT, Yang HB, Chan SH, Tsai ST: Genetic susceptibility to nasopharyngeal carcinoma within the HLA-A locus in Taiwanese. *Int J Cancer* 2003; **103**: 745–751.
- 29 Ooi EE, Ren EC, Chan SH: Association between microsatellites within the human MHC and nasopharyngeal carcinoma. *Int J Cancer* 1997; **74**: 229–232.
- 30 Hildesheim A, Apple RJ, Chen CJ *et al*: Association of HLA class I and II alleles and extended haplotypes with nasopharyngeal carcinoma in Taiwan. *J Natl Cancer Inst* 2002; **94**: 1780–1789.
- 31 Goldsmith DB, West TM, Morton R: HLA associations with nasopharyngeal carcinoma in Southern Chinese: a meta-analysis. *Clin Otolaryngol* 2002; **27**: 61–67.
- 32 Brancolini V, Devoto M: Genetic linkage studies for the identification of cancer-related genes. *Ann Ist Super Sanit* 1996; **32**: 173–180.
- 33 Feng BJ, Huang W, Shugart YY *et al*: Genome-wide scan for familial nasopharyngeal carcinoma reveals evidence of linkage to chromosome 4. *Nat Genet* 2002; **31**: 395–399.
- 34 Chen CJ, Liang KY, Chang YS *et al*: Multiple risk factors of nasopharyngeal carcinoma: Epstein–Barr virus, malarial infection, cigarette smoking and familial tendency. *Anticancer Res* 1990; **10**: 547–553.
- 35 Weber BL, Nathanson KL: Low penetrance genes associated with increased risk for breast cancer. *Eur J Cancer* 2000; **36**: 1193–1199.
- 36 Olson JM: A general conditional-logistic model for affected-relative pair linkage studies. *Am J Hum Genet* 1999; **65**: 1760–1769.
- 37 Goddard KA, Witte JS, Suarez BK, Catalona WJ, Olson JM: Model-free linkage analysis with covariates confirms linkage of prostate cancer to chromosomes 1 and 4. *Am J Hum Genet* 2001; **68**: 1197–1206.
- 38 Olson JM, Goddard KAB, Dudek DM: The amyloid precursor protein locus and very-late-onset Alzheimer disease. *Am J Hum Genet* 2001; **69**: 895–899.
- 39 Devlin B, Jones BL, Bacanu S-A, Roeder K: Mixture models for linkage analysis of affected sibling pairs and covariates. *Genet Epidemiol* 2002; **22**: 52–65.