Evaluation of rare and common variants from suspected familial or sporadic nasopharyngeal carcinoma (NPC) susceptibility genes in sporadic NPC

Zhiwei Liu[1], **Alisa M. Goldstein**[1], **Wan-Lun Hsu**[2], **Kelly J. Yu**[1], **Yin-Chu Chien**[2], **Cheng-Yu Ko**[3], **James Jer-Min Jian**[4], **Yung-An Tsou**[5], **Yi-Shing Leu**[6], **Li-Jen Liao**[7], **Yen-Liang Chang**[8], **Cheng-Ping Wang**[3], **Jia-Shing Wu**[4], **Chun-Hung Hua**[5], **Jehn-Chuan Lee**[6], **Tsung-Lin Yang**[3], **Chu-Hsing Hsiao**[9], **Ming-Shiang Wu**[10], **Ming-Hsui Tsai**[5], **Kuei-Kang Huang**[11], **Kai Yu**[1], **Kristie Jones**[1], **Bin Zhu**[1], **Meredith Yeager**[1], **Guoqin Yu**[1], **Pei-Jen Lou**[3,†], **Chien-Jen Chen**[2,†], **Allan Hildesheim**[1,†] **GEV-NPC group**[*]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA

[2]Genomics Research Center, Academia Sinica, Taipei, Taiwan

[3]Department of Otolaryngology, National Taiwan University Hospital and National Taiwan University, College of Medicine, Taipei, Taiwan

[4]Department of Head and Neck Surgery, Koo Foundation Sun Yat-Sen Cancer Center, Taipei, Taiwan

[5]Department of Otorhinolaryngology, China Medical University Hospital, Taichung, Taiwan

[6]Department of Otolaryngology, MacKay Memorial Hospital, Taipei, Taiwan

[7]Department of Otolaryngology, Far Eastern Memorial Hospital, New Taipei City, Taiwan

[8]Department of Otolaryngology, Cathay General Hospital, Taipei, Taiwan

---

Correspondence to: Zhiwei Liu, PhD, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Rockville, MD 20892, Phone: +0012402762726; zhiwei.liu@nih.gov.
[†]These authors contributed equally to this work
[*]GEV-NPC Study Group include
Genomics Research Center, Academia Sinica, Taipei, Taiwan (Wan-Lun Hsu, Yin-Chu Chien, Chien-Jen Chen)
Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA (Allan Hildesheim, Zhiwei Liu, Kelly J. Yu)
Department of Otolaryngology, National Taiwan University Hospital and National Taiwan University, College of Medicine, Taipei, Taiwan (Cheng-Yu Ko, Cheng-Ping Wang, Tsung-Lin Yang, Pei-Jen Lou, Chun-Nan Chen, Tseng-Cheng Chen, Chih-Feng Lin)
Department of Radiation Oncology, Koo Foundation Sun Yat-Sen Cancer Center, Taipei, Taiwan (James Jer-Min Jian, Skye Hongiun Cheng, Yu-Chen Tsai, Yih-Lin Chung, Jia-Shing Wu, Ming-Jiung Liu, Kuei-Kang Huang)
Department of Pathology and Laboratory Medicine, Koo Foundation Sun Yat-Sen Cancer Center, Taipei, Taiwan (Mei-Hua Tsou)
Department of Hematology and Medical Oncology, Koo Foundation Sun Yat-Sen Cancer Center, Taipei, Taiwan (Hsin-Hsuan Chen)
Division of Otolaryngology Head and Neck Surgery (Ching-Yuan Lin, Shyuang-Der Terng, Fang-Yin Lin, Hsin-I Huang)
Department of Otorhinolaryngology, China Medical University Hospital, Taichung, Taiwan (Yung-An Tsou, Chun-Hung Hua, Ming-Hsui Tsai)
Department of Otolaryngology, MacKay Memorial Hospital, Taipei, Taiwan (Yi-Shing Leu, Jehn-Chuan Lee)
Department of Otolaryngology, Far Eastern Memorial Hospital, New Taipei City, Taiwan (Li-Jen Liao)
Department of Otolaryngology, Cathay General Hospital, Taipei, Taiwan (Yen-Liang Chang)
Graduate Institute of Epidemiology and Preventative Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan (Chu-Hsing Hsiao)
Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan (Ming-Shiang Wu)

[9]Graduate Institute of Epidemiology and Preventative Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan

[10]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

[11]Department of Radiation Oncology, Koo Foundation Sun Yat-Sen Cancer Center, Taipei, Taiwan

Abstract

Background: Genetic susceptibility is associated with nasopharyngeal carcinoma (NPC). We previously identified rare variants potentially involved in familial NPC and common variants significantly associated with sporadic NPC.

Methods: We conducted targeted gene sequencing of 20 genes (16 identified from the study of multiplex families, 3 identified from a pooled analysis of NPC genome wide association study [GWAS], and 1 identified from both studies) among 819 NPC cases and 938 controls from two case-control studies in Taiwan (independent from previous studies). A targeted, multiplex PCR primer panel was designed using the custom Ion AmpliSeq Designer v4.2 targeting the regions of the selected genes. Gene-based and single-variant tests were conducted.

Results: We found that NPC was associated with combined common and rare variants in *CDKN2A/2B* (P=1.3×10^{-4}), *BRD2* (P=1.6×10^{-3}), *TNFRSF19* (P=4.0×10^{-3}) and *CLPTM1L/TERT* (P=5.4×10^{-3}). Such associations were likely driven by common variants within these genes, based on gene-based analyses evaluating common variants and rare variants separately (e.g. for common variants of *CDKN2A/2B*, P=4.6×10^{-4}; for rare variants, P=0.04). We also observed a suggestive association with rare variants in *HNRNPU* (P=3.8×10^{-3}) for NPC risk. In addition, we validated four previously reported NPC-risk associated SNPs.

Conclusion: Our findings confirm previously reported associated variants and suggest that some common variants in genes previously linked to familial NPC are associated with the development of sporadic NPC.

Impact: NPC-associated genes, including *CLPTM1L/TERT*, *BRD2*, and *HNRNPU*, suggest a role for telomere length maintenance in NPC etiology.

Keywords

case-control; familial cases; nasopharyngeal carcinoma; sporadic cases; susceptibility gene; Taiwan; targeted sequencing

Introduction

While Epstein-Barr virus (EBV) infection is a necessary cause for virtually all undifferentiated nasopharyngeal carcinoma (NPC), only a small proportion of infected individuals develop NPC (1,2). Inherited genetic susceptibility is hypothesized to be an important risk factor for NPC among EBV infected individuals, and agnostic and targeted studies of sporadic and familial NPC support this notion (3,4).

We recently completed a whole-exome sequencing study among 251 individuals from 97 multiplex NPC families from Taiwan in which novel, rare variants were implicated in the

development of familial NPC (5). These variants are located in genes potentially involved in notch signaling (*NOTCH1*, *DLL3*, *LFNG*, *MAML1*, *MFNG*, *PSEN2*), magnesium transport (*NIPAL1*), EBV entry into epithelial cells (*ITGB6*), modulation of EBV (*BCL2L12*, *NEDD4L*, *LAMC2*), telomere biology (*CLPTM1L*, *BRD2*, *HNRNPU*), DNA repair (*PRKDC*, *MLH1*) or modulation of cAMP signaling (*RAPGEF3*). Separate findings from a meta-analysis of sporadic NPC noted significant associations for common polymorphisms within genes involved in telomere biology (*CLPTM1L/TERT*), apoptosis (*MECOM*, *TNFRSF19*), and cell cycle regulation (*CDKN2A/2B*) (6). Taken together, these studies of familial and sporadic NPC point to a set of candidate genes that might be important in NPC pathogenesis.

To follow-up on these findings, we performed a targeted gene sequencing of 20 genes/gene regions in an independent set of 1,757 NPC cases and controls from two studies in Taiwan. In this study, we aimed 1) to evaluate whether variability in genes involved in familial NPC might also be associated with sporadic NPC and 2) to determine whether additional polymorphisms in genes previously associated with sporadic NPC could be identified.

Materials and Methods

Study population

The two case-control studies in Taiwan have been described previously (7,8). In brief, the first case-control study, conducted between 1991 and 1994, recruited incident NPC cases from the Taipei metropolitan area and age, sex and residential area matched controls. Of 378 eligible cases and 372 controls, 369 cases (98%) and 320 controls (86%) agreed to participate. The second case-control study, conducted between 2010 and 2014, recruited prevalent and incident NPC cases diagnosed after January 1, 2007 from Northern and Central Taiwan and age and sex frequency matched controls from the same geographical regions. Of 1,850 eligible cases and 1,885 controls, 1,600 cases (86%) and 1,804 controls (96%) agreed to participate.

For the present study, we included all individuals with DNA available from the first case-control study (241 cases and 234 controls), and all individuals from the initial phase of the second case-control study (578 cases and 704 controls), as samples were not available for the remaining participants. For cases and controls, separately, we noted a similar distribution of age and sex for those who were included and excluded from each study (*P*>0.05). In total, we included 819 NPC cases and 938 controls. Characteristics of the NPC cases and controls are presented in Supplementary Table 1. Of 819 NPC cases, 256 were prevalent (i.e., recruited after treatment was initiated/completed) and 563 were incident (i.e., recruited at the time of diagnosis and prior to treatment initiation). Written informed consent was obtained from all participants and both studies were approved by human subject committees in Taiwan and the United States.

Targeted Sequencing Pipeline

DNA was extracted from the peripheral blood lymphocytes using the QIAamp DNA Blood Mini Kit (Qiagen). The quantity and quality of the genomic DNA were assessed by

Nanodrop 1000 (Thermo Scientific) and Qubit (Life Technologies), respectively. A targeted, multiplex PCR primer panel was designed using the custom Ion AmpliSeq Designer v4.2 (Life Technologies, Carlsbad, CA, USA) targeting the coding regions of the 20 selected genes/gene regions. Sample DNA (30ng) was amplified using this custom AmpliSeq primer pool, and libraries were prepared following the manufacturer's Ion AmpliSeq Library Preparation protocol (Life Technologies, Carlsbad, CA, USA).

Raw sequencing reads generated by the Ion Torrent sequencer were quality and adaptor trimmed by Ion Torrent Suite and then aligned to the hg19 reference sequence by TMAP (https://github.com/iontorrent/TS/tree/master/Analysis/TMAP) using default parameters. Resulting BAM files were aligned using GATK LeftAlignIndels module. Amplicon primers were trimmed from aligned reads. Variant calls and filtrations were made by Torrent Variant Caller 5.0 and Genome Analysis Toolkit (GATK, UnifiedGenotyper v3.1) (9). Variant annotation was done by snpEff, SnpSift (http://snpeff.sourceforge.net/) and Annovar (10). On average, we achieved a read length of 224 bp (with a minimum of 198 bp) across the samples, with an average of mean depth of 1113x (interquartile range: 849x-1296x). Insertions and deletions (INDELs) were not called due to the potential high false positive rate of INDELs calling for the data generated from the Ion Torrent sequencer. A total of 1966 variants were identified. Genotypes with read depth (DP) < 8 or genotype quality (GQ) < 20 were considered as ambiguous calls and therefore were considered as missing values. Variants missing in more than 5% of the samples in cases and controls were excluded (N=120). We excluded multiallelic variants (172 SNPs) and SNPs with Hardy–Weinberg equilibrium $P < 10^{-4}$ (N=2). Six candidate SNPs, which were significantly associated with NPC in a previous GWAS pooled analysis (6), were interrogated with the Ion Torrent panel; as compared with the variant calling data, no discrepancy was observed.

Statistical analysis

We first focused on genes previously identified from a study of NPC multiplex families (N=17), to test the hypothesis that genes involved in familial NPC are also associated with sporadic NPC. We then focused on genes associated with sporadic NPC from a previously reported NPC GWAS (N=4; 1 overlapped with the list from the family study) to both confirm previous findings and to determine whether there is support for an association with NPC for additional variants within those genes.

For gene-based analyses, we used the Sequence Kernel Association combined sum test (SKAT-C) to estimate the effect of common variants and combined effect of common (minor allele frequency - MAF -   1% among all study subjects) and rare (MAF <1% among all study subjects) variants with NPC risk (11). To evaluate the association between rare variants alone and NPC risk, we used the SKAT optimal adjusted (SKAT-O) test that encompasses the SKAT test and burden test. All SKAT analyses were performed using default weights, adjusting for sex and age groups (in 10-year category). For single-variant analyses (restricted to common variants), associations between SNPs and NPC were analyzed by logistic regression under a log additive model, adjusting for study, sex, and age groups (in 10-year category).

We considered a gene-based association significant if the nominal P value was $< 2.5 \times 10^{-3}$, corresponding to a Bonferroni correction for 20 genes evaluated. For the four SNPs with previously reported genome-wide significance ($P < 5 \times 10^{-8}$) (6), we considered P value of <0.05 as significant. For other SNPs, we considered a single-variant association significant if the nominal P value was $< 2.7 \times 10^{-4}$, corresponding to a Bonferroni correction for 187 tests (total number of common variants evaluated). Sensitivity analysis was conducted among incident cases and all controls.

Based on the current sample size (~800 cases and ~900 controls), after correction for multiple test (significant level = 2.7×10^{-4}), we had at least 80% power to detect an odds ratio (OR) as low as 3.4, 1.9, and 1.6, corresponding to MAF of 1%, 5%, and 10%, respectively. Due to a modest sample size, we have a limited power to identify disease association for rare variants.

Results

Number of variants:

From the selected 20 genes/gene regions, we identified a total of 1670 variants among 1,757 participants (819 NPC; 938 controls), of which 187 were common (MAF   1%).

Gene-based Analyses:

The association between the combined effect of rare and common variants and NPC risk is summarized in Table 1. For genes previously identified from a study of multiplex families, the most significant association was observed for *BRD2* located on Chromosome 6 ($P_{\text{SKAT-C}} = 0.0016$). A suggestive association was observed for *CLPTM1L/TERT* located on Chromosome 5 ($P_{\text{SKAT-C}} = 0.0054$). In analyses restricted to rare variants (MAF <1%), however, no evidence for significant associations for these two genes was noted, suggesting that the observed association was more likely explained by common variants within these gene regions. We observed a suggestive association with NPC for rare variants in *HNRNPU* located on Chromosome 1 ($P_{\text{SKAT-O}} = 0.0038$).

For genes previously identified from a meta-analysis of NPC GWAS, the most significant association was observed for *CDKN2A/2B* located on Chromosome 9 ($P_{\text{SKAT-C}} = 1.3 \times 10^{-4}$), followed by a suggestive association for *TNFRSF19* located on Chromosome 13 ($P_{\text{SKAT-C}} = 0.0040$). Again, no evidence for significant associations for these two genes was noted in analyses restricted to rare variants.

SNP-based Analyses: The association between individual common SNPs and NPC risk is summarized in Table 2 and Supplemental Tables 2 and 3. For SNPs from genes previously identified from a multiplex family study, suggestive evidence in support of an association with NPC was observed for rs78231671 ($P = 4.7 \times 10^{-4}$, Table 2), an intronic SNP in the *PRKDC* gene region, and for rs76146382 ($P = 5.8 \times 10^{-4}$) in the *BRD2* gene region.

We next evaluated the four SNPs that were previously reported to be significantly associated with NPC in a meta-analysis (6). All four SNPs, located within the *CLPTM1L/TERT, CDKN2A/2B, TNFRSF19,* and *MECOM* genes/gene regions, were significantly associated

with NPC, and the direction of the association was the same as previously reported ($P <$ 0.05, Table 2). Of an additional 29 SNPs evaluated within these four genes (Supplemental Table 2), 10 were in suggestive LD with the primary a-priori SNPs evaluated ($r^2 > 0.20$ with the primary a-priori SNP, Supplementary Figures 1A–D). Of the remaining 19 SNPs, we found suggestive evidence for an association with NPC for one SNP within *CLPTM1L/ TERT,* rs13167280 ($P = 0.012$, Supplementary Table 2), and one SNP within *MECOM* region, rs17466625 ($P = 0.011$).

The association between other common SNPs and NPC risk is summarized in Supplementary Table 3. Associations of similar magnitude were observed in analysis among incident cases and controls (Supplementary Tables 4 and 5).

Discussion

In this targeted gene sequencing study, we extended our findings from a study of NPC multiplex families aimed at identifying genes associated with familial NPC risk by showing that 2 genes, including *CLPTM1L/TERT* and *BRD2*, are also associated with sporadic NPC risk in Taiwan. A suggestive association was observed for *HNRNPU*. In addition, we confirmed results from a recent meta-GWAS of NPC by showing that 4 previously reported SNPs in genes *CDKN2A/2B*, *MECOM*, *TNFRSF19*, and *CLPTM1L/TERT* locus also conferred risk for sporadic NPC in this Taiwanese sample.

The biological implication for genes *CDKN2A/2B*, *TNFRSF19*, and *CLPTM1L/TERT* in NPC pathogenesis have been discussed previously (6,12). These genes are involved in cell growth control (*CDKN2A/2B* and *TNFRSF19*), and telomere length maintenance (*CLPTM1L/TERT*). Both *BRD2* and *HNRNPU* are also involved in telomere biology (13,14). BRD2 is one of the host chromatin factors interacting with the Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 that affects chromatin structure and maintenance of the telomere repeat (15). Of note, the most significantly associated SNP within *BRD2* was rs76146382. This SNP is in LD with an extended HLA haplotype known to be associated with NPC, HLA-A*3303~B*5801~DRB1*0301 ($r^2 > 0.6$) (7,16). It remains to be determined whether the association with *BRD2* observed in the present study is explained by or independent of this previously reported HLA association. The role of *HNRNPU* in NPC pathogenesis is not clear. Overexpression of *HNRNPU* has been associated with telomere shortening (13). In addition, *HNRNPU* shows a strong homology with sequence of EBV nuclear antigen 1 (EBNA1), which is important for NPC development (17). Whether such similarity would have biological implication is unclear. Taken together, although other mechanisms merit further investigation, our findings further support a role for telomere maintenance in NPC pathogenesis.

Variants in other genes previously linked to familial NPC were not associated with the risk of sporadic NPC. This finding could be due to the fact that genes involved in familial NPC pathogenesis differ from genes in sporadic NPC pathogenesis or to false positive findings in our previous familial NPC study. In addition, false negative findings cannot be ruled out given the modest sample size and power of our study to identify disease associations for rare

variants that occur with very low frequency in the population and account for a small proportion of NPC cases. More studies are needed to confirm our findings.

In summary, variants in several genes previously linked to familial NPC were associated with the risk of sporadic NPC and the risk was more likely to be driven by common rather than rare variants within these genes. We also confirmed previously reported associations from NPC GWAS for SNPs in four genes. Our findings highlighted the important role for telomere length maintenance in NPC pathogenesis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Abbreviations list:

DP	read depth
EBNA1	EBV nuclear antigen 1
EBV	Epstein-Barr virus
GQ	genotype quality
GWAS	genome wide association study
INDEL	Insertion and deletion
OR	odds ratio
MAF	minor allele frequency
NPC	nasopharyngeal carcinoma
SKAT-C	Sequence Kernel Association combined sum test
SKAT-O	Sequence Kernel Association test optimal adjusted

Reference

1. Pathmanathan R, Prasad U, Sadler R, Flynn K, Raab-Traub N. Clonal proliferations of cells infected with Epstein-Barr virus in preinvasive lesions related to nasopharyngeal carcinoma. The New England journal of medicine 1995;333(11):693–8 doi 10.1056/NEJM199509143331103. [PubMed: 7637746]

2. Niedobitek G Epstein-Barr virus infection in the pathogenesis of nasopharyngeal carcinoma. Molecular pathology : MP 2000;53(5):248–54. [PubMed: 11091848]

3. Bei JX, Jia WH, Zeng YX. Familial and large-scale case-control studies identify genes associated with nasopharyngeal carcinoma. Seminars in cancer biology 2012;22(2):96–106 doi 10.1016/j.semcancer.2012.01.012S1044-579X(12)00014-4 [pii]. [PubMed: 22313875]

4. Hildesheim A, Wang CP. Genetic predisposition factors and nasopharyngeal carcinoma risk: a review of epidemiological association studies, 2000–2011: Rosetta Stone for NPC: genetics, viral infection, and other environmental factors. Seminars in cancer biology 2012;22(2):107–16 doi 10.1016/j.semcancer.2012.01.007. [PubMed: 22300735]

5. Yu G, Hsu WL, Coghill AE, Yu KJ, Wang CP, Lou PJ, et al. Whole Exome Sequencing of Nasopharyngeal Carcinoma Families Reveals Novel Variants Potentially Involved in Nasopharyngeal Carcinoma Submitted 2018.

6. Bei JX, Su WH, Ng CC, Yu K, Chin YM, Lou PJ, et al. A GWAS Meta-analysis and Replication Study Identifies a Novel Locus within CLPTM1L/TERT Associated with Nasopharyngeal Carcinoma in Individuals of Chinese Ancestry. Cancer Epidemiol Biomarkers Prev 2016;25(1):188–92 doi 10.1158/1055-9965.EPI-15-0144. [PubMed: 26545403]

7. Hildesheim A, Apple RJ, Chen CJ, Wang SS, Cheng YJ, Klitz W, et al. Association of HLA class I and II alleles and extended haplotypes with nasopharyngeal carcinoma in Taiwan. Journal of the National Cancer Institute 2002;94(23):1780–9. [PubMed: 12464650]

8. Hsu WL, Chien YC, Yu KJ, Ko CY, Lin CY, Tsou YA, et al. Cigarette Smoking, Epstein-Barr Virus and Nasopharyngeal Carcinoma: Results from a Multicenter Case-Control Study in Taiwan Submitted 2018.

9. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43(5):491–8 doi 10.1038/ng.806. [PubMed: 21478889]

10. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research 2010;38(16):e164 doi 10.1093/nar/gkq603. [PubMed: 20601685]

11. Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. American journal of human genetics 2013;92(6): 841–53 doi 10.1016/j.ajhg.2013.04.015. [PubMed: 23684009]

12. Bei JX, Li Y, Jia WH, Feng BJ, Zhou G, Chen LZ, et al. A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. Nat Genet 2010;42(7):599–603 doi 10.1038/ng.601. [PubMed: 20512145]

13. Fu D, Collins K. Purification of human telomerase complexes identifies factors involved in telomerase biogenesis and telomere length regulation. Mol Cell 2007;28(5):773–85 doi 10.1016/j.molcel.2007.09.023. [PubMed: 18082603]

14. Deng Z, Wang Z, Lieberman PM. Telomeres and viruses: common themes of genome maintenance. Frontiers in oncology 2012;2:201 doi 10.3389/fonc.2012.00201. [PubMed: 23293769]

15. Ballestas ME, Kaye KM. The latency-associated nuclear antigen, a multifunctional protein central to Kaposi's sarcoma-associated herpesvirus latency. Future Microbiol 2011;6(12):1399–413 doi 10.2217/fmb.11.137. [PubMed: 22122438]

16. Gourraud PA, Khankhanian P, Cereb N, Yang SY, Feolo M, Maiers M, et al. HLA diversity in the 1000 genomes dataset. PloS one 2014;9(7):e97282 doi 10.1371/journal.pone.0097282. [PubMed: 24988075]

17. Fischer N, Voss MD, Mueller-Lantzsch N, Grasser FA. A potential NES of the Epstein-Barr virus nuclear antigen 1 (EBNA1) does not confer shuttling. FEBS Lett 1999;447(2–3):311–4. [PubMed: 10214968]

Table 1.

Genes associated with nasopharyngeal carcinoma (NPC) in Taiwan identified by SKAT gene-based tests among all cases and controls.

Chr.	Gene	Cytoband	Common and rare variants combined		Common variants (MAF ≥ 1%)		Rare variants (MAF <1%)	
			N. of variants	P_{SKAT-C} [a]	N. of variants	P_{SKAT-C} [a]	N. of variants	P_{SKAT-O} [a]
Gene previously identified from a study of multiplex families and a meta-analysis of NPC GWAS								
5	CLPTM1L/TERT	5p15.33	166	0.0054	14	0.0028	152	0.60
Gene previously identified from a study of multiplex families								
6	BRD2	6p21.32	79	0.0016	20	0.0010	59	0.25
1	HNRNPU	1q44	46	0.016	6	0.55	40	0.0038
8	PRKDC	8q11.21	260	0.14	17	0.16	243	0.41
4	NIPAL1	4p12	25	0.47	0	N/A	25	0.68
7	LFNG	7p22.3	83	0.34	7	0.23	76	0.13
19	BCL2L12	19q13.33	27	0.24	1	0.27	26	0.36
5	MAML1	5q35.3	32	0.36	3	0.50	29	0.67
12	RAPGEF3	12q13.11	103	0.27	17	0.33	86	0.45
3	MLH1	3p22.2	61	0.29	9	0.12	52	0.89
22	MFNG	22q13.1	33	0.41	2	0.65	31	0.34
19	DLL3	19q13.2	39	0.45	4	0.16	35	0.10
9	NOTCH1	9q34.3	266	0.61	20	0.55	246	0.53
2	ITGB6	2q24.2	65	0.64	7	1.00	58	0.44
1	PSEN2	1q42.13	58	0.66	6	0.32	52	0.73
1	LAMC2	1q25.3	100	0.68	26	0.71	74	0.62
18	NEDD4L	18q21.31	97	0.88	9	0.83	88	0.70
Gene previously identified from a meta-analysis of NPC GWAS								
9	CDKN2A/2B	9p21.3	32	0.00013	6	0.00046	26	0.04
13	TNFRSF19	13q12.12	33	0.0040	8	0.00061	25	0.88
3	MECOM	3q26.2	65	0.19	5	0.042	60	0.92

Abbreviation: Chr: Chromosome; GWAS: genome-wide association study; MAF: minor allele frequency; SKAT: Sequence Kernel Association test; SKAT-C: Sequence Kernel Association combined sum test; SKAT-O: Sequence Kernel Association test with optimal adjustment (burden test).

[a] All models were adjusted for study, sex and age group (in 10-year category).

Table 2.

Selected results from associations of common variants (MAF 1%) with nasopharyngeal carcinoma (NPC) risk in Taiwan among all cases and controls.

SNP	Gene	Chr	Position (bp)	REF	EFFECT	MAF	MAF		OR per allele [a] (95% CI)	P
							Control	Case		
Gene previously identified from a study of multiplex families [b]										
rs78231671	*PRKDC*	8	48695175	T	C	0.037	0.026	0.049	1.93 (1.34–2.82)	0.00047
rs76146382	*BRD2*	6	32943229	T	C	0.073	0.058	0.089	1.57 (1.22–2.03)	0.00058
Gene previously identified from a pooled analysis of NPC GWAS [c]										
rs31489	*CLPTM1L/TERT*	5	1342714	C	A	0.161	0.177	0.142	0.78 (0.65–0.94)	0.011
rs1412829	*CDKN2A/2B*	9	22043926	A	G	0.089	0.105	0.071	0.66 (0.52–0.83)	0.00052
rs9510787	*TNFRSF19*	13	24205195	A	G	0.326	0.304	0.351	1.24 (1.08–1.43)	0.0026
rs6774494	*MECOM*	3	169082633	G	A	0.385	0.403	0.366	0.84 (0.73–0.97)	0.017

Abbreviation: Chr: Chromosome; CI: confidence interval; EFFECT: effect allele; GWAS: genome-wide association study; MAF: minor allele frequency; NPC: nasopharyngeal carcinoma; OR: odds ratio; REF: reference allele

[a] All models were adjusted for study, sex, and age group (in 10-year category).

[b] Ref: #5.

[c] Selected based on results from a pooled analysis (ref: #6).