OXFORD

JNCI J Natl Cancer Inst (2022) 114(12): djac177

https://doi.org/10.1093/jnci/djac177
First published online September 6, 2022
Article

Whole-Exome Sequencing Study of Familial Nasopharyngeal Carcinoma and Its Implication for Identifying High-Risk Individuals

Tong-Min Wang, PhD,[1,‡‡] Yong-Qiao He, PhD,[1,‡] Wen-Qiong Xue, MS,[1] Jiang-Bo Zhang, PhD,[1] Yun-Fei Xia, MD,[1,2] Chang-Mi Deng, MD,[1] Wen-Li Zhang, MD,[1] Ruo-Wen Xiao, MD,[1] Ying Liao, PhD,[1] Da-Wei Yang, PhD,[3] Ting Zhou, MS,[1] Dan-Hua Li, MD,[1] Lu-Ting Luo, MS,[3] Xia-Ting Tong, MS,[3] Yan-Xia Wu, MD,[1] Xue-Yin Chen, BS,[1] Xi-Zhao Li, MS,[1] Pei-Fen Zhang, MS,[1] Xiao-Hui Zheng, PhD,[1] Shao-Dan Zhang, BS,[1] Ye-Zhu Hu, BS,[1] Fang Wang, MD,[1] Zi-Yi Wu, MD,[1] Mei-Qi Zheng, MS,[3] Jing-Wen Huang, MS,[3] Yi-Jing Jia, MS,[3] Lei-Lei Yuan, MS,[3] Rui You, MD,[1,4] Guan-Qun Zhou, MD,[1,2] Li-Xia Lu, MD,[1,2] Yu-Ying Liu, MD,[5] Ming-Yuan Chen, MD,[1,4] Lin Feng, PhD,[1] Wei Dai, PhD,[6] Ze-Fang Ren, MD,[3] Hai-Qiang Mai, MD,[1,4] Ying Sun, MD,[1,2] Jun Ma, MD,[1,2] Wei Zheng, MD,[7] Maria Li Lung, PhD,[6] Wei-Hua Jia, PhD[1,*]

[1]State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Sun Yat-sen University Cancer Center, Guangzhou, P. R. China; [2]Department of Radiation Oncology, Sun Yat-sen University Cancer Center, Guangzhou, P. R. China; [3]School of Public Health, Sun Yat-sen University, Guangzhou, P. R. China; [4]Department of Nasopharyngeal Carcinoma, Sun Yat-sen University Cancer Center, Guangzhou, P. R. China; [5]Department of Cancer Prevention, Sun Yat-sen University Cancer Center, Guangzhou, P. R. China; [6]Department of Clinical Oncology, University of Hong Kong, Hong Kong (Special Administrative Region), People's Republic of China; and [7]Division of Epidemiology, Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt Epidemiology Center, Vanderbilt University School of Medicine, Nashville, TN, USA

[‡]Authors contributed equally to this work.

[*]**Correspondence to:** Wei-Hua Jia, PhD, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Sun Yat-sen University Cancer Center, Guangzhou, P. R. China (e-mail: jiawh@sysucc.org.cn).

Abstract

Background: Nasopharyngeal carcinoma (NPC) is closely associated with genetic factors and Epstein-Barr virus infection, showing strong familial aggregation. Individuals with a family history suffer elevated NPC risk, requiring effective genetic counseling for risk stratification and individualized prevention. **Methods:** We performed whole-exome sequencing on 502 familial NPC patients and 404 unaffected relatives and controls. We systematically evaluated the established cancer predisposition genes and investigated novel NPC susceptibility genes, making comparisons with 21 other familial cancers in the UK biobank (N = 5218). **Results:** Rare pathogenic mutations in the established cancer predisposition genes were observed in familial NPC patients, including *ERCC2* (1.39%), *TP63* (1.00%), *MUTYH* (0.80%), and *BRCA1* (0.80%). Additionally, 6 novel susceptibility genes were identified. *RAD54L*, involved in the DNA repair pathway together with *ERCC2*, *MUTYH*, and *BRCA1*, showed the highest frequency (4.18%) in familial NPC. Enrichment analysis found mutations in *TP63* were enriched in familial NPC, and *RAD54L* and *EML2* were enriched in both NPC and other Epstein-Barr virus–associated cancers. Besides rare variants, common variants reported in the studies of sporadic NPC were also associated with familial NPC risk. Individuals in the top quantile of common variant-derived genetic risk score while carrying rare variants exhibited increased NPC risk (odds ratio = 13.47, 95% confidence interval = 6.33 to 28.68, $P = 1.48 \times 10^{-11}$); men in this risk group showed a cumulative lifetime risk of 24.19%, much higher than those in the bottom common variant-derived genetic risk score quantile and without rare variants (2.04%). **Conclusions:** This study expands the catalog of NPC susceptibility genes and provides the potential for risk stratification of individuals with an NPC family history.

Nasopharyngeal carcinoma (NPC), a malignancy that originates from the mucosal epithelium of the nasopharynx and is closely associated with Epstein-Barr virus (EBV) infection, shows distinctive geographic distribution around the world, with high incidence rates in Southern China, Southeast Asia, North Africa, and the Arctic (1-7). Approximately 10% of patients have an NPC

Received: April 1, 2022; **Revised:** July 28, 2022; **Accepted:** August 31, 2022

family history, and first-degree relatives of NPC patients showed more than fourfold elevated NPC risk compared with those without such a history (8,9). Extensive epidemiological observations of the unbalanced geographical distribution of disease incidence, familial clustering, and the estimated high heritability (61.3%) (10) suggested the importance of genetic factors on NPC development. Developing an efficient genetic screening tool to identify high-risk individuals with an NPC family history could provide the potential for individualized cancer screening and surveillance. However, no genetic testing panel is currently available for genetic consultation for individuals with an NPC family history.

Many cancer predisposition genes, such as *BRCA1*, *BRCA2*, and *TP53*, were identified and recommended to be tested by the National Comprehensive Cancer Network (NCCN) Clinical Practice Guidelines in Oncology for cancer risk assessment for individuals with a cancer family history (11,12). However, whether and how the established cancer predisposition genes could be applied in NPC screening for individuals with an NPC family history remain to be investigated. Recently, sequencing studies of familial NPC patients have identified rare variants with potentially deleterious effects in 14 susceptibility genes (13-15). In addition to the rare variants identified in familial NPC, genome-wide association studies (GWASs) conducted mostly in sporadic NPC cases and controls have identified associations in the human leukocyte antigen (HLA) region as well as other common variants in *CDKN2B-AS1*, *CLPTM1L/TERT*, etc (16-21). These findings indicate the complex nature of the NPC genetic etiology and raise questions of whether there are additional genes involved in NPC susceptibility.

To expand the catalog of NPC susceptibility genes and construct an efficient genetic risk prediction model, we performed whole-exome sequencing (WES) on individuals from high-risk NPC families recruited from the High-risk Nasopharyngeal Carcinoma Family Screening Program (22) as well as from the Sun Yat-sen University Cancer Center biobank. We evaluated rare pathogenic or likely pathogenic variants of the established cancer predisposition genes and searched for novel NPC susceptibility genes. We established a genetic risk score (GRS) using common and rare variants to assess the risk stratification of the individuals with an NPC family history.

ARTICLE

Methods

Study Samples

A total of 502 familial NPC cases, 76 unaffected relatives, and 328 healthy controls recruited from the High-risk Nasopharyngeal Carcinoma Family Screening Program (22) or the Sun Yat-sen University Cancer Center biobank were used to perform WES (Supplementary Methods; Supplementary Figure 1, available online). All the familial NPC patients have 2 or more NPC cases among the first- (89.04%), second-, or third-degree relatives in their family. The Institutional Review Board of Sun Yat-Sen University Cancer Center approved this study. Informed consent was obtained from all study participants.

Five publicly available NPC WES datasets were downloaded for estimating the prevalence of mutations in the cancer predisposition gene and the novel NPC susceptibility genes (13,23-26) (SRA291701, SRP035573, SRA288429, PRJEB12830, HRA000052, and HRA000053). After quality control, a total of 269 samples remained in the analysis.

To compare the mutation frequencies between familial NPC and other familial cancers, WES data of 200 619 samples (27) from the UK biobank were obtained. A total of 21 types of familial cancers with a sample size of more than 50 were investigated in this study (Supplementary Methods; Supplementary Table 1, available online), adding up to 5218 patients.

Germline Variant Calling and Annotation

We followed the recommendations of Genome Analysis Toolkit (GATK) Best Practices (28,29) to analyze the WES data. ANNOtate VARiation (ANNOVAR) (30) and Ensembl Variant Effect Predictor (VEP) (31) were used to annotate the qualified variants. Variants with a minor allele frequency less than 0.5% in the East Asian population were defined as rare variants and were further filtered as pathogenic or likely pathogenic variants according to the guidelines recommended by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (32). We curated a total of 162 autosome cancer predisposition genes from literatures review and the COSMIC database (release v92) (11,33-60) (Supplementary Methods; Supplementary Table 2, available online). In addition, rare variants in the 14 reported NPC susceptibility genes (13-15) were reviewed in our study samples.

Reported common single nucleotide polymorphisms (SNPs) associated with NPC risk were obtained from the GWAS catalog, and the proxy SNPs ($r^2 > 0.8$) available in our WES data were used for downstream analysis. HLA typing of 4-digit classical class-I alleles was performed by using Optitype (version 1.3.3) (61). Detailed information is shown in Supplementary Methods (available online).

Construction of GRSs and Cumulative Risk Projection

Among the total 906 samples, 787 were unrelated and randomly split into training (228 cases, 166 controls) and testing samples (227 cases, 166 controls). We generated 4 GRSs using common variants with *P* less than .05 in the training samples, which were pruned with different r^2 thresholds (0.01, 0.1, 0.5, and 0.7). The best variables for the construction of common variant-derived GRS (GRS_C) were selected by the highest area under the curve (AUC) and the strongest association with familial NPC risk in the fivefold cross-validation of the training samples. The GRS_C was constructed in the training samples with randomization and then validated in the testing samples. A GRS integrating common and rare variants (GRS_{CR}) was also constructed by multiple regression in the training samples and validated in the testing samples. We classified all the samples into 4 groups based on the GRS_C or GRS_{CR} quantiles of the healthy controls. The Individualized Coherent Absolute Risk Estimator tool (62,63) was used to estimate the individual cumulative NPC risks in future 10-year and cumulative lifetime for individuals at different quantiles of GRS_C and rare variant–carrying status (Supplementary Methods, available online).

Statistical Analysis

Gene-based association analysis was performed using the STAAR (64) method and including the genetic correlation matrix in the model for the control of sample relatedness. Sex, age, and the first 3 principal components were adjusted in the model. Genes with *P* less than .05 in the omnibus test in the STAAR framework (STAAR-O) and with the rare variants enriched in

familial NPC patients were identified as potential novel NPC susceptibility genes (Supplementary Methods, available online). A permutation test was performed by random selection (for 10 000 times) of the rare variants in the flanking regions of the potential susceptibility genes. A generalized linear mixed model association test was used to evaluate the rare variant–collapsing effects, the common variants effects, and the genetic risk quantiles effects on familial NPC risk, controlling sample relatedness by using genetic correlation matrix in the analysis (65). Enrichment analysis was performed to identify genes with enriched prevalence at each specific cancer type, and genes with Fisher exact test *P* less than .05 were identified as suggestive statistically significant enrichment in a specific familial cancer. The Benjamini-Hochberg method was used for multiple test corrections (66). All statistical analysis was conducted using R software (version 3.5.0).

Results

Association of Rare Pathogenic Variants in Established Cancer Predisposition Genes With Familial NPC Risk

We performed WES on 502 familial NPC patients and 404 unaffected relatives and controls from NPC endemic areas, with an average depth of 93.53× on target (Supplementary Table 3, available online). Among the 502 familial NPC patients, the mean (SD) age at diagnosis was 47.56 (10.63) years, and 361 patients (71.91%) were male (Supplementary Table 4, available online).

We assessed the rare pathogenic or likely pathogenic variants in 162 established cancer predisposition genes in the familial NPC patients, patients from publicly available NPC WES datasets, as well as other familial cancers in the UK biobank (21 cancer types, N = 5218). Overall, 15 of the 162 cancer predisposition genes were identified in familial NPC, and the rare pathogenic or likely pathogenic variants were carried by 36 (7.17%) familial NPC patients, showing a higher frequency compared with the controls (3.47%; Figure 1, A). Recurrent pathogenic or likely pathogenic variants were observed in 4 genes with frequencies of 0.80%-1.39% in familial NPC. *ERCC2* showed the highest mutation frequency. Seven familial NPC patients (1.39%) and 1 publicly available NPC case (0.37%) carried *ERCC2* mutations, all of which are in the DNA-binding helicase domains (Figure 1, B). *TP63* is the second-ranking gene. Five familial NPC patients (1.00%) and 3 publicly available NPC cases (1.12%) carried *TP63* mutations, and *TP63*:NM_003722: c.1807G>C: p.D603H is the mutation hotspot, classified as “likely pathogenic” in InterVar (Figure 1, C). Enrichment analysis incorporating WES data of 21 other familial cancers in the UK biobank found that the mutations in *TP63* showed suggestive enrichment in familial NPC, whereas other familial cancers did not show statistically significant enrichment for this gene ($P = 7.29 \times 10^{-3}$, $P_{adj} = .14$) (Supplementary Tables 5 and 6, available online). In addition, 4 familial NPC patients carried *BRCA1* mutations and 4 familial NPC patients carried *MUTYH* mutations (Figure 1, A, D-E). Overall, increased familial NPC risk was observed for the carriers of any identified rare pathogenic or likely pathogenic variants in the established cancer predisposition genes (odds ratio [OR] = 2.18, 95% confidence interval [CI] = 1.11 to 4.25, $P = .02$). Detailed information on the mutations is shown in Supplementary Table 7 (available online).

Identification of Rare Variants on the Six Novel NPC Susceptibility Genes

To identify additional NPC susceptibility genes, high-impact rare variants consisting of pathogenic or likely pathogenic variants and truncation variants were included for gene-based association analysis. Six potential novel NPC susceptibility genes were identified with $P_{STAAR\text{-}O}$ less than .05 (Figure 2, A; Supplementary Table 8, available online), among which *RAD54L* showed the highest mutation frequency in familial NPC. A total of 21 familial NPC patients (4.18%) and 1 publicly available NPC sample (0.37%) carried *RAD54L* high-impact variants (Figure 2, B). By incorporating WES data of another 21 familial cancers in the UK biobank, we found that *RAD54L* was statistically significantly enriched not only in familial NPC (4.18% in familial NPC vs 1.67% in other familial cancers, $P_{fisher} = 4.13 \times 10^{-4}$, $P_{adj} = .01$) but also in non-Hodgkin lymphoma, a cancer with some subtypes associated with EBV infection (4.98% in non-Hodgkin lymphoma vs 1.72% in other familial cancers, $P_{fisher} = 4.74 \times 10^{-4}$, $P_{adj} = .01$; Figure 2, H; Supplementary Table 9, available online). Additionally, recurrent high-impact variants were observed in *CAPN3* (1.59%), *CEP152* (1.39%), *EML2* (1.39%), *LRRC19* (1.20%), and *ZNF135* (1.20%) in familial NPC (Figure 2, C-G). Notably, *EML2* was enriched in both familial NPC ($P_{fisher} = 2.06 \times 10^{-5}$, $P_{adj} = 1.0 \times 10^{-3}$) and Hodgkin lymphoma, which is also an EBV-associated malignancy ($P_{fisher} = 6.04 \times 10^{-3}$, $P_{adj} = .12$), and the mutations in *LRRC19* ($P_{fisher} = 4.45 \times 10^{-7}$, $P_{adj} = 1.0 \times 10^{-3}$) and *CEP152* ($P_{fisher} = 1.33 \times 10^{-4}$, $P_{adj} = .01$) were specifically enriched in familial NPC (Figure 2, H; Supplementary Table 9, available online). Overall, increased familial NPC risk was observed for the carriers of any rare high-impact variants in the 6 genes (OR = 4.56, 95% CI = 2.23 to 9.34, $P = 3.28 \times 10^{-5}$). Detailed information on the mutations is shown in Supplementary Table 10 (available online).

We also evaluated the rare variants in 14 reported NPC susceptibility genes in our samples (13-15). Focusing on the rare high-impact variants on these genes as well as considering the reported variants in the original studies, we found that 2.59% of familial NPC patients carried *MST1R* mutations, 1.39% carried *BCL2L12* mutations, and 1.20% carried *MLH1* mutations (Supplementary Table 11, available online). Collapsing all the identified variants in the 14 genes, increased familial NPC risk for the variant carriers was observed (OR = 1.76, 95% CI = 0.94 to 3.29, $P = .079$).

Association of Common Variants Identified in NPC GWASs With Familial NPC Risk

Common SNPs with *P* less than 5×10^{-8} in the NPC GWASs (16-21) were evaluated in familial NPC. We found 4 of 6 proxy SNPs available in our WES data were replicated with *P* less than .05 (Figure 3; Supplementary Table 12, available online). The top signal was rs1136695 (OR = 0.46, 95% CI = 0.36 to 0.60, $P = 3.21 \times 10^{-9}$), a proxy in *HLA-A* locus, showing an elevated effect size on familial NPC risk compared with that observed in sporadic NPC studies (16,18). Given the top signal hits the HLA class I region, we performed HLA typing for 4-digit class-I alleles. Ten classical HLA class-I alleles were statistically significantly associated with familial NPC risk, including the reported protective alleles HLA-A*11:01 and HLA-C*12:02, and the reported risk alleles HLA-A*02:07, HLA-A*33:03, HLA-B*46:01, HLA-B*58:01, HLA-C*01:02, and HLA-C*03:02[20,67,68]. In addition, 2 alleles, HLA-B*15:02 and HLA-C*06:02, which have not

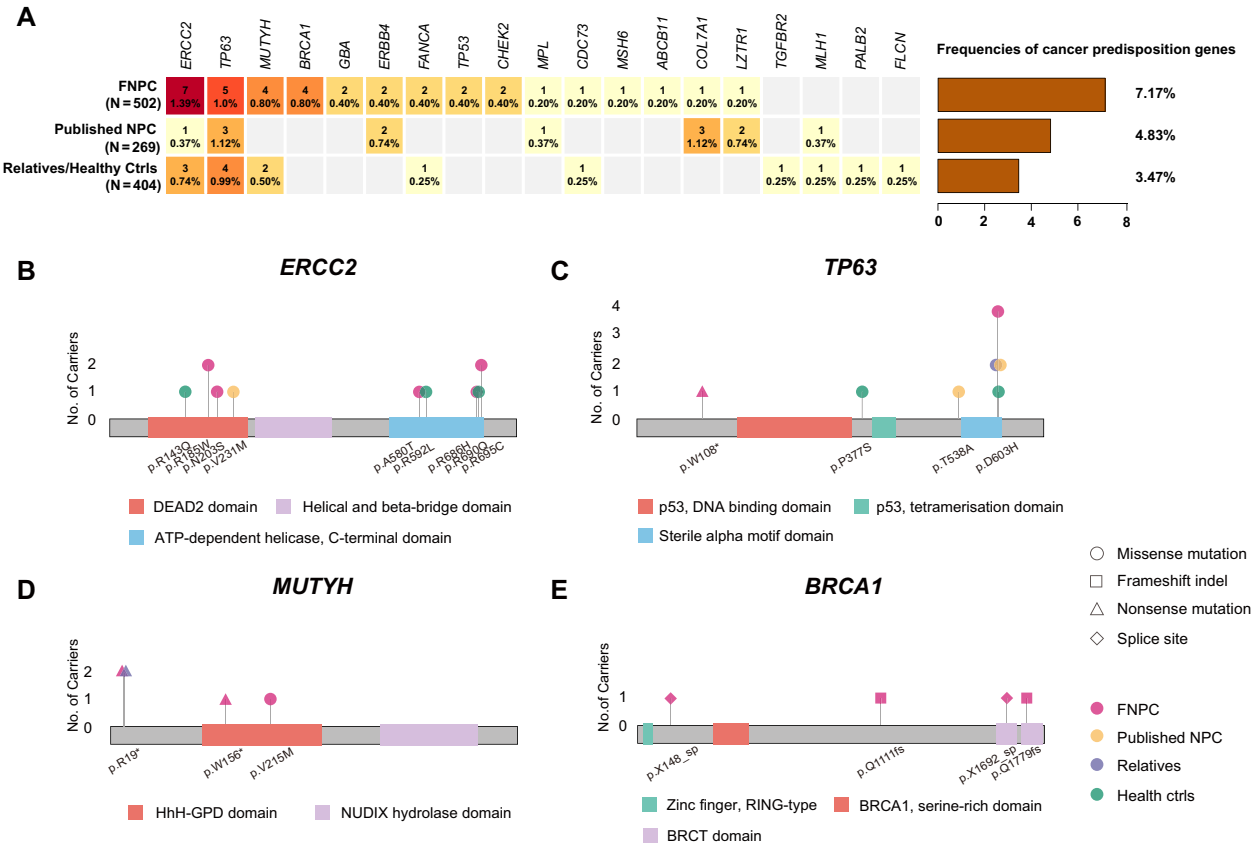

Figure 1. The rare pathogenic or likely pathogenic variants of the established cancer predisposition genes. **A)** The frequencies of rare pathogenic or likely pathogenic variants in 19 established cancer predisposition genes in the familial nasopharyngeal carcinoma (NPC) cases, the published NPC samples and the relatives or healthy controls (Ctrls) (**left**), and the mutations of 15 genes were identified in familial NPC cases. The overall frequencies of the cancer predisposition genes in the familial NPC cases, the published NPC samples, and the relatives/healthy controls are shown (**right**). **B-E)** The **lollipop plots** illustrate all the rare pathogenic or likely pathogenic variants in the 4 most frequently mutated genes (*ERCC2*, *TP63*, *MUTYH*, and *BRCA1*). The y-axis indicates the number of carriers for each mutation. The box colors indicate domains of each gene. The colors of the points indicate the individuals from the familial NPC cases (**red**), the published NPC samples (**yellow**), unaffected relatives (**purple**), and healthy controls (**green**); the shapes indicate the mutation types: **circle** for missense mutation, **square** for frameshift indel, **rectangle** for nonsense mutation, and **diamond** for splice site. FNPC = familial nasopharyngeal carcinoma.

been reported in NPC genetic studies, were associated with familial NPC risk (OR = 0.63, 95% CI = 0.42 to 0.94, $P = .03$ and OR = 0.39, 95% CI = 0.18 to 0.83, $P = .01$, respectively; Figure 3).

NPC Risk Stratification Using GRSs Derived by Identified Common and Rare Variants

We developed a GRS_C in the training samples (Supplementary Tables 13 and 14, available online) and then validated the GRS_C in the test samples (AUC = 0.71; Supplementary Table 15, available online). We found individuals in the top quantile of GRS_C had 6.72 times the NPC risk compared with those in the bottom quantile (95% CI = 4.19 to 10.76, $P = 2.29 \times 10^{-15}$; Figure 4, A). Given that common and rare variants are 2 important genetic components for NPC risks, we then performed stratification analysis by rare variant carrying status and found the NPC risk for individuals in different GRS_C quantiles could be stratified by rare variants. Carriers of any identified rare variants in the top quantile of GRS_C exhibited dramatically increased risk, with an OR of 13.47 (95% CI = 6.33 to 28.68, $P = 1.48 \times 10^{-11}$; Table 1) compared with the noncarriers of rare variants in the bottom quantile of GRS_C. Subsequently, we developed a GRS_{CR} in the training samples and validated the risk stratification ability in the test samples (AUC = 0.737; Supplementary Table 16, available online). We observed individuals in the top quantile of GRS_{CR} showed 9.21 times the risk compared with those in the bottom quantile (95% CI = 5.57 to 15.25, $P = 5.99 \times 10^{-18}$; Figure 4, A), implicating that inclusion of rare variants could improve the risk stratification performance of the model using common variants alone.

The cumulative lifetime and the future 10-year NPC risks for individuals with an NPC family history were calculated. The average lifetime NPC risks were 24.19% for men and 8.06% for women, who were in the top GRS_C quantile and in the meantime carried any identified rare variants, whereas the risks were 2.04% for men and 0.62% for women in the bottom GRS_C quantile and not carrying rare variants. Similar trends were observed for the future 10-year NPC risks (Figure 4, B and C; Supplementary Figure 2, available online).

Discussion

In this study, we for the first time—to our knowledge—evaluated the rare pathogenic or likely pathogenic variants in the established cancer predisposition genes for familial NPC and expanded the current genetic architecture of NPC by identifying 6 novel NPC

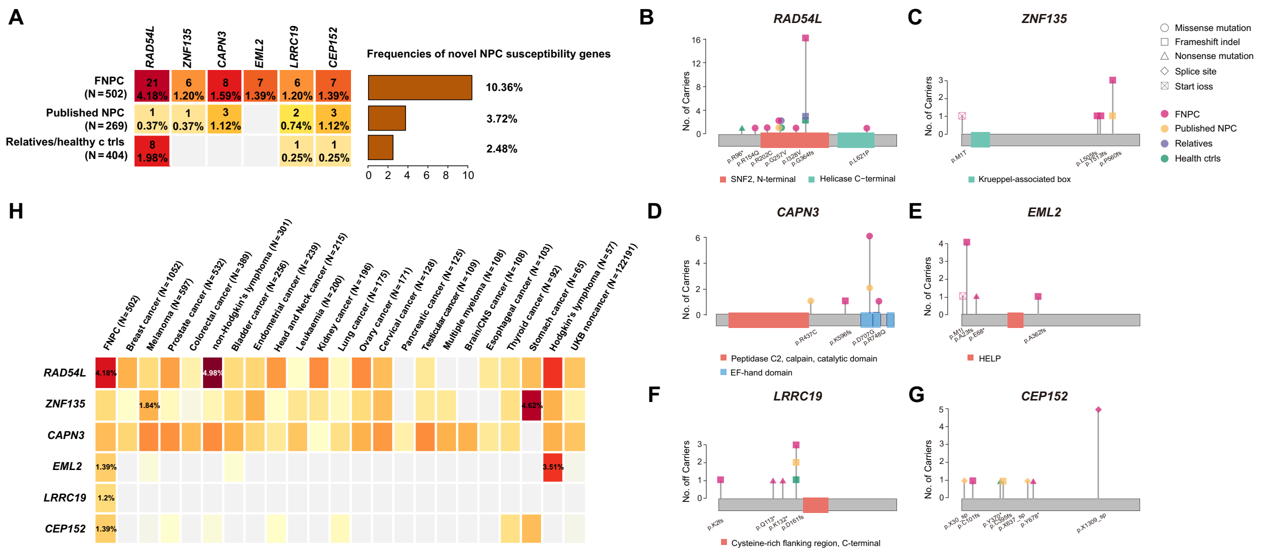

Figure 2. The rare high-impact variants of the nasopharyngeal carcinoma (NPC) novel susceptibility genes. **A)** The frequencies of rare high-impact variants in 6 novel NPC susceptibility genes in the familial NPC cases, the published NPC samples, and the relatives or healthy controls are shown (**left**). The overall frequencies in the familial NPC cases, the published NPC samples, and the relatives or healthy controls are shown (**right**). **B-G)** The **lollipop plots** illustrate all the high-impact variants identified in the 6 NPC susceptibility genes. The y-axis indicates the number of carriers for each mutation. The box colors indicate domains of each gene. The colors of the points indicate the individuals from the familial NPC cases (**red**), the published NPC samples (**yellow**), unaffected relatives (**purple**), and healthy controls (Ctrls) (**green**); the shapes indicate the mutation types: **circle** for missense mutation, **square** for frameshift indel, **rectangle** for nonsense mutation, **diamond** for splice site, and **square with a cross** for start loss mutation. **H)** The frequencies of rare high-impact variants of each NPC susceptibility gene identified in familial NPC as well as other familial cancers from the UK Biobank (UKB) are shown. The genes with statistically significant enrichments ($P < .05$) in a specific cancer type are highlighted by showing the frequencies of the genes in the **corresponding boxes**. FNPC = familial nasopharyngeal carcinoma.

susceptibility genes. By comparing the mutation spectrum of familial NPC with other family cancers, we identified genes enriched in familial NPC as well as genes enriched in both familial NPC and other EBV-associated cancers. Besides rare variants, common variants reported in the NPC GWASs, which were conducted mostly in sporadic cases and controls, were also associated with familial NPC risk. We developed a GRS integrating all the identified common and rare variants, which showed good performance in NPC risk stratification. This study improves the understanding of NPC genetic etiology and provides important evidence for NPC risk stratification and personalized prevention for the population with an NPC family history.

Family history is a well-established risk factor for many cancers (69-71), and individuals with a positive cancer family history need effective genetic consultation and risk assessment. Common and rare variants are 2 critical genetic components for cancer susceptibility (72) and could collectively contribute to a more precise risk prediction of cancers (73) than using family history alone (74-76). Moreover, using GRS could further stratify high-risk individuals with a positive family history and provide a stable and feasible supplement to cancer family history (77). In this study, we used the common and rare variants to construct the GRS for NPC risk prediction. We showed that individuals with an NPC family history could be stratified into different genetic risk groups by common and rare variants. The GRS derived by common variants alone showed 6.72 times the risk for the top quantile compared with the bottom; a largely improved risk stratification performance was achieved by adding rare variants, showing 13.47 times risk for the individuals in the top GRS_C quantile and meanwhile carrying rare variants. These findings suggest diverse risk management and health surveillance strategies may be applied for the populations with an NPC family history. More aggressive NPC screening may be suggested for risk reduction and early diagnosis for the high–genetic risk subgroup. Particularly, individuals carrying rare variants with a large effect size and more important clinical impact may be referred to clinical intervention for risk reduction and early detection.

The pathogenesis role of EBV infection on NPC development has long been suggested. Evidence showed that EBV infection and reactivation could promote host genomic instability by inducing DNA damage and inhibiting DNA repair, which potentially result in tumorigenesis (78-81). In this study, we identified several potential susceptibility genes related to DNA damage repair or/and EBV infection. For example, the identified gene *RAD54L* is involved in the homologous recombination and repair of DNA (82-84) and was proposed as a potential candidate gene for cancer susceptibility (85,86). The polymorphisms in *RAD54L* were associated with EBV seropositive status of IgA antibody against viral capsid antigen (VCA-IgA) (87), an important biomarker for NPC screening and early diagnosis. Rare *RAD54L* variants were enriched in both familial NPC and non-Hodgkin lymphoma, implicating the shared mechanism of *RAD54L* on NPC and other EBV-associated cancers (88,89). *ERCC2* is important in DNA excision repair, acting as an essential subunit of the general transcription factor IIH (TFIIH) complex (90). The polymorphisms of *ERCC2* were associated with NPC risk (13,91,92). In this study, 1.39% of familial NPC carried *ERCC2* rare pathogenic mutations in DNA-binding helicase domains, suggesting the inherited pathogenic mutations may affect the normal DNA-binding function. Moreover, *ERCC2* and the TFIIH complex could be targeted by *EBNA2*, an EBV-encoded transactivator, and this interaction may impair the DNA repair functions of TFIIH (93). *TP63*, a transcription factor of the p53 family and a key regulator of epithelial-cell differentiation (94), was an established cancer predisposition gene enriched in familial NPC. *TP63* may be involved in the development of NPC and many other cancers (95-98). Notably,

Variants	Cytoband	Minor allele	MAF (FNPC)	MAF (relatives/ctrls)	OR (95% CI)	*P*
SNPs						
rs402710	5p15.33	T	0.27	0.33	0.81 (0.64 to 1.01)	.06
rs465498	5p15.33	G	0.17	0.21	0.81 (0.62 to 1.08)	.14
rs2747421	6p22.1	C	0.42	0.33	1.47 (1.17 to 1.85)	9.76×10^{-4}
rs1136695	6p22.1	A	0.20	0.35	0.46 (0.36 to 0.60)	3.21×10^{-9}
rs146420652	6p21.32	T	0.05	0.09	0.54 (0.35 to 0.82)	4.16×10^{-3}
rs564398	9p21.3	C	0.06	0.10	0.63 (0.43 to 0.93)	.02
HLA alleles						
A*11:01	6p22.1	P	0.15	0.28	0.45 (0.34 to 0.60)	1.82×10^{-8}
A*02:07	6p22.1	P	0.25	0.11	2.97 (2.15 to 4.10)	3.46×10^{-11}
A*33:03	6p22.1	P	0.12	0.08	1.83 (1.19 to 2.80)	5.61×10^{-3}
B*46:01	6p22.1	P	0.26	0.17	1.97 (1.45 to 2.68)	1.34×10^{-5}
B*15:02	6p22.1	P	0.07	0.10	0.63 (0.42 to 0.94)	.03
B*58:01	6p22.1	P	0.09	0.06	1.98 (1.23 to 3.18)	4.94×10^{-3}
C*01:02	6p22.1	P	0.30	0.21	1.63 (1.24 to 2.16)	5.43×10^{-4}
C*03:02	6p22.1	P	0.09	0.05	2.30 (1.44 to 3.69)	5.39×10^{-4}
C*12:02	6p22.1	P	0.02	0.04	0.47 (0.23 to 0.97)	.04
C*06:02	6p22.1	P	0.01	0.03	0.39 (0.18 to 0.83)	.01

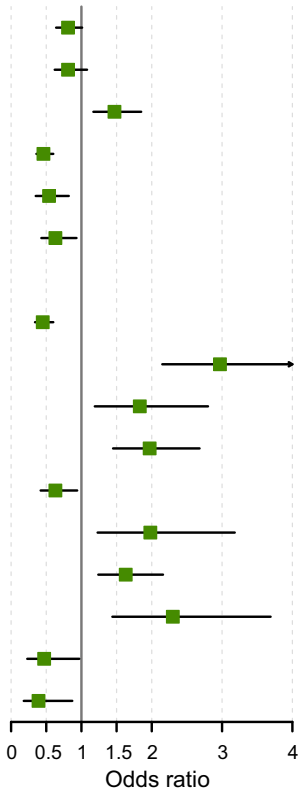

Figure 3. The associations of common variants with familial nasopharyngeal carcinoma (NPC) risk. The **forest plot** shows the odds ratios (ORs) and corresponding confidence intervals (CIs) of the proxied of genome-wide association studies–identified single nucleotide polymorphisms (SNPs) (**upper**) and the human leukocyte antigen (HLA) alleles (**lower**). The odds ratios and *P* values were calculated by using generalized linear mixed model adjusting for age, sex, and the first 3 principal components. The HLA minor allele "P" indicates present of the corresponding allele. Ctrls = controls; FNPC = familial nasopharyngeal carcinoma; MAF = minor allele frequency.

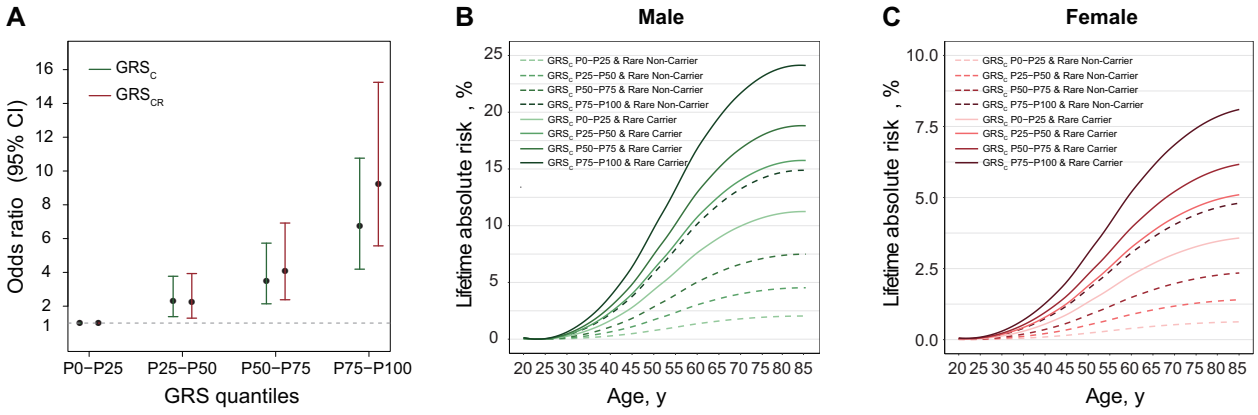

Figure 4. The performance of risk stratifications using genetic risk scores (GRS) derived by common variants and incorporated with rare variants. **A**) The effects of GRS derived by common variants (GRS_C, **green**) and GRS derived by common and rare variants (GRS_{CR}, **red**) on familial nasopharyngeal carcinoma (NPC) risk in the overall samples. The odds ratios and the corresponding 95% confidence intervals of each quantile are shown by **error bars**. The cumulative absolute risks of developing NPC between age 20 years and a specific age (*x*-axis) for male (**B**) and female (**C**) with NPC family history are shown. The **dashed lines** indicate the risks for individuals not carrying any identified rare variants, and the **solid lines** indicate individuals carrying any of the identified rare variants.

TP63 could transactivate many DNA damage repair genes (99,100). Additionally, *TP63* could interact with EBV-encoded oncoprotein *LMP2A* and affect the regulation of epithelial differentiation (101). Taken together, these findings suggest that abnormal host genetics related to DNA repair pathways or/and EBV infection may collaboratively contribute to NPC

Table 1. The effects of GRS_C quantiles and rare variant carrying status on familial NPC risk

	Noncarrier of rare variant[a]				Carrier of rare variant[b]			
GRS_C	FNPC, No. (%)	Relatives/ controls, No. (%)	OR (95% CI)[c]	P^c	FNPC, No. (%)	Relatives/ controls, No. (%)	OR (95% CI)[c]	P^c
P0-P25	24 (6.14)	87 (23.84)	1.0 (Referent)	—	8 (7.21)	5 (12.82)	5.80 (1.72 to 19.52)	4.53×10^{-3}
P25-P50	54 (13.81)	87 (23.84)	2.25 (1.28 to 3.96)	4.96×10^{-3}	23 (20.72)	10 (25.64)	8.34 (3.50 to 19.88)	1.73×10^{-6}
P50-P75	94 (24.04)	90 (24.66)	3.78 (2.19 to 6.54)	1.97×10^{-6}	28 (25.23)	10 (25.64)	10.15 (4.33 to 23.79)	9.69×10^{-8}
P75-P100	219 (56.01)	101 (27.67)	7.84 (4.67 to 13.16)	6.65×10^{-15}	52 (46.85)	14 (35.9)	13.47 (6.33 to 28.68)	1.48×10^{-11}

[a]Individuals not carrying any identified rare variants. CI = confidence interval; FNPC = familial nasopharyngeal carcinoma; GRS_C = common variant-derived genetic risk score; NPC = nasopharyngeal carcinoma; OR = odds ratio.

[b]Individuals carrying any identified rare variants.

[c]Odds ratios, 95% confidence intervals, and P values were calculated by using generalized linear mixed model.

carcinogenesis (80), and further studies are warranted to investigate whether the NPC susceptibility genes contribute to the disease risk by influencing EBV infections or whether there is an interplay between host genetics and EBV infection on NPC development.

Family-based studies have the advantage to obviate the problem of population stratification introduced in unrelated case-control studies. However, there are many difficulties in the recruitment of high-risk NPC families and in the accumulation of enough numbers of well-characterized families with both cases and first-degree relatives in successive generations. Therefore, we used a hybrid design by including samples from both NPC families with familial cases and unaffected relatives as well as familial case-unrelated control. Although we identified 6 potential susceptibility genes with $P_{STAAR\text{-}O}$ less than .05 and permutation P less than .05, they did not surpass multiple test correction. The power of this study remains to be improved by additional recruitment of NPC families with a more complete structure and inclusion of a larger number of familial cases and independent controls. In addition, this study only consists of individuals from the NPC endemic area, and additional validation of the identified NPC susceptibility genes and the constructed GRS in external study cohorts is warranted. Moreover, there is some progress on the identification of cancer predisposition genes and NPC-associated SNPs, and additional comprehensive study including novel cancer predisposition genes and NPC-associated SNPs could be performed to improve the performance of the model (63,102-104). Lastly, our current prediction model includes only the genetic factors. A more comprehensive model should be developed by combining genetic factors with environmental risk factors (105,106).

In summary, we systematically evaluated the established cancer predisposition genes in familial NPC samples and identified 6 novel NPC susceptibility genes. The rare variants confer statistically significantly increased risk of NPC and provide added value for risk stratification. This study expands the understanding of NPC genetic etiology and provides a potential tool for the management of the population with an NPC family history, which could benefit the precision risk prediction, prevention, and early diagnosis of the disease.

Funding

This study was supported by grants from the National Key Research and Development Program of China (2021YFC2500400), the Basic and Applied Basic Research Foundation of Guangdong Province, China (2021B1515420007), the Science and Technology Planning Project of Guangzhou, China (201804020094), Sino-Sweden Joint Research Programme (81861138006), the Special Support Program for High-level Professionals on Scientific and Technological Innovation of Guangdong Province, China (2014TX01R201), National Natural Science Foundation of China (82003520, 81973131, 81903395, 81803319), National Science Fund for Distinguished Young Scholars of China (81325018), the Science and Technology Planning Project of Guangdong Province, China (2019B030316031), the Key Area Research and Development Program of Guangdong Province, China (2019B110233004).

Notes

Role of the funder: The funders had no role in the study design, collection, and analysis; decision to publish; or preparation of the manuscript.

Disclosures: The authors have declared no conflicts of interest.

Author contributions: Tong-Min Wang: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Software; Visualization; Writing—original draft; Writing—review and editing. Yong-Qiao He: Conceptualization; Data curation; Formal analysis; Writing—original draft; Writing—review and editing. Wen-Qiong Xue, Jiang-Bo Zhang, Yun-Fei Xia, Ze-Fang Ren: Data curation; Resources. Chang-Mi Deng, Wen-Li Zhang: Data curation; Formal analysis. Ruo-Wen Xiao: Data curation; Writing—original draft. Ying Liao, Da-Wei Yang, Ting Zhou, Dan-Hua Li, Lu-Ting Luo, Xia-Ting Tong, Yan-Xia Wu, Xue-Yin Chen, Fang Wang, Zi-Yi Wu, Mei-Qi Zheng, Jing-Wen Huang, Yi-Jing Jia, Lei-Lei Yuan: Data curation. Xi-Zhao Li, Pei-Fen Zhang, Xiao-Hui Zheng, Shao-Dan Zhang, Ye-Zhu Hu, Rui You, Guan-Qun Zhou, Li-Xia Lu, Yu-Ying Liu, Ming-Yuan Chen, Lin Feng, Hai-Qiang Mai, Ying Sun, Jun Ma: Resources. Wei Dai, Wei Zheng, Maria Li Lung: Writing—review and editing. Wei-Hua Jia: Conceptualization; Data curation; Funding acquisition; Investigation; Methodology; Resources; Supervision; Writing—original draft; Writing—review and editing.

Acknowledgements: We thank all patients and relatives from the NPC high-risk families, who participated in this study. This research has been conducted using the UK Biobank Resource under Application Number 58450.

ARTICLE

Data Availability

The baseline sample information and genetic information data have been deposited in the Research Data Deposit public platform (www.researchdata.org.cn, accession number: RDDA2022918529). The genotyping data reported in this paper have been deposited in the National Genomics Data Center (NGDC, Nucleic Acids Res 2021), Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number PRJCA010171 (https://ngdc.cncb.ac.cn/gvm/). Other data underlying this article will be shared on reasonable request to the corresponding author.

References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424.
2. Tang LL, Chen WQ, Xue WQ, et al. Global trends in incidence and mortality of nasopharyngeal carcinoma. *Cancer Lett*. 2016;374(1):22-30.
3. Chan AT. Nasopharyngeal carcinoma. *Ann Oncol*. 2010;21(suppl 7):vii308-12.
4. Chen YP, Chan ATC, Le QT, et al. Nasopharyngeal carcinoma. *Lancet*. 2019; 394(10192):64-80.
5. Tsang CM, Tsao SW. The role of Epstein-Barr virus infection in the pathogenesis of nasopharyngeal carcinoma. *Virol Sin*. 2015;30(2):107-121.
6. Tsao SW, Tsang CM, Lo KW. Epstein-Barr virus infection and nasopharyngeal carcinoma. *Philos Trans R Soc Lond B Biol Sci*. 2017;372(1732):20160270.
7. Young LS, Dawson CW. Epstein-Barr virus and nasopharyngeal carcinoma. *Chin J Cancer*. 2014;33(12):581-590.
8. Jia WH, Feng BJ, Xu ZL, et al. Familial risk and clustering of nasopharyngeal carcinoma in Guangdong, China. *Cancer*. 2004;101(2):363-369.
9. Liu Z, Chang ET, Liu Q, et al. Quantification of familial risk of nasopharyngeal carcinoma in a high-incidence area. *Cancer*. 2017;123(14):2716-2725.
10. Huang SF, Hsiao JH, Young CK, et al. Familial aggregation of nasopharyngeal carcinoma in Taiwan. *Oral Oncol*. 2017;73:10-15.
11. Rahman N. Realizing the promise of cancer predisposition genes. *Nature*. 2014;505(7483):302-308.
12. Daly MB, Pal T, Berry MP, et al.; CGC. Genetic/familial high-risk assessment: breast, ovarian, and pancreatic, version 2.2021, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. 2021;19(1):77-102.
13. Dai W, Zheng H, Cheung AK, et al. Whole-exome sequencing identifies MST1R as a genetic susceptibility gene in nasopharyngeal carcinoma. *Proc Natl Acad Sci USA*. 2016;113(12):3317-3322.
14. Sasaki MM, Skol AD, Bao R, et al. Integrated genomic analysis suggests MLL3 is a novel candidate susceptibility gene for familial nasopharyngeal carcinoma. *Cancer Epidemiol Biomarkers Prev*. 2015;24(8):1222-1228.
15. Yu G, Hsu WL, Coghill AE, et al. Whole-exome sequencing of nasopharyngeal carcinoma families reveals novel variants potentially involved in nasopharyngeal carcinoma. *Sci Rep*. 2019;9(1):9916.
16. Bei JX, Li Y, Jia WH, et al. A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. *Nat Genet*. 2010;42(7): 599-603.
17. Chin YM, Mushiroda T, Takahashi A, et al.; Malaysian NPC Study Group. HLA-A SNPs and amino acid variants are associated with nasopharyngeal carcinoma in Malaysian Chinese. *Int J Cancer*. 2015;136(3):678-687.
18. Cui Q, Feng QS, Mo HY, et al. An extended genome-wide association study identifies novel susceptibility loci for nasopharyngeal carcinoma. *Hum Mol Genet*. 2016;25(16):3626-3634.
19. Ng CC, Yew PY, Puah SM, et al. A genome-wide association study identifies ITGA9 conferring risk of nasopharyngeal carcinoma. *J Hum Genet*. 2009;54(7): 392-397.
20. Tang M, Lautenberger JA, Gao X, et al. The principal genetic determinants for nasopharyngeal carcinoma in China involve the HLA class I antigen recognition groove. *PLoS Genet*. 2012;8(11):e1003103.
21. Tse KP, Su WH, Chang KP, et al. Genome-wide association study reveals multiple nasopharyngeal carcinoma-associated loci within the HLA region at chromosome 6p21.3. *Am J Hum Genet*. 2009;85(2):194-203.
22. Zheng MQ, Wang TM, Liao Y, et al. Nasopharyngeal Epstein-Barr virus DNA loads in high-risk nasopharyngeal carcinoma families: familial aggregation and host heritability. *J Med Virol*. 2020;92(12):3717-3725. doi: 10.1002/jmv.26198.
23. Lin DC, Meng X, Hazawa M, et al. The genomic landscape of nasopharyngeal carcinoma. *Nat Genet*. 2014;46(8):866-871.
24. Zheng H, Dai W, Cheung AK, et al. Whole-exome sequencing identifies multiple loss-of-function mutations of NF-kappaB pathway regulators in nasopharyngeal carcinoma. *Proc Natl Acad Sci USA*. 2016;113(40):11283-11288.
25. Li YY, Chung GT, Lui VW, et al. Exome and genome sequencing of nasopharynx cancer identifies NF-kappaB pathway activating mutations. *Nat Commun*. 2017;8:14121.
26. You R, Liu YP, Lin DC, et al. Clonal mutations activate the NF-kappaB pathway to promote recurrence of nasopharyngeal carcinoma. *Cancer Res*. 2019; 79(23):5930-5943.
27. Szustakowski JD, Balasubramanian S, Kvikstad E, et al.; UKB-ESC Research Team. Advancing human genetics research and drug discovery through exome sequencing of the UK Biobank. *Nat Genet*. 2021;53(7):942-948.
28. Van der Auwera GA, O'Connor BD. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. 1st ed. O'Reilly Media; 2020.
29. Poplin R, Ruano-Rubio V, DePristo MA, et al. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv. 2018. doi: 10.1101/201178:201178.
30. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010; 38(16):e164.
31. McLaren W, Gil L, Hunt SE, et al. The ensembl variant effect predictor. *Genome Biol*. 2016;17(1):122.
32. Richards S, Aziz N, Bale S, et al.; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424.
33. Zhang J, Walsh MF, Wu G, et al. Germline mutations in predisposition genes in pediatric cancer. *N Engl J Med*. 2015;373(24):2336-2346.
34. Lessel D, Vaz B, Halder S, et al. Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nat Genet*. 2014;46(11):1239-1244.
35. Wang L, Yamaguchi S, Burstein MD, et al. Novel somatic and germline mutations in intracranial germ cell tumours. *Nature*. 2014;511(7508):241-245.
36. Hanks S, Perdeaux ER, Seal S, et al. Germline mutations in the PAF1 complex gene CTR9 predispose to Wilms tumour. *Nat Commun*. 2014;5(1):4398.
37. Noetzli L, Lo RW, Lee-Sherick AB, et al. Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. *Nat Genet*. 2015;47(5):535-538.
38. Cybulski C, Carrot-Zhang J, Kluzniak W, et al. Germline RECQL mutations are associated with breast cancer susceptibility. *Nat Genet*. 2015;47(6): 643-646.
39. Kiiski JI, Pelttari LM, Khan S, et al. Exome sequencing identifies FANCM as a susceptibility gene for triple-negative breast cancer. *Proc Natl Acad Sci USA*. 2014;111(42):15172-15177.
40. Yokoyama S, Woods SL, Boyle GM, et al. A novel recurrent mutation in MITF predisposes to familial and sporadic melanoma. *Nature*. 2011;480(7375): 99-103.
41. Hussin J, Sinnett D, Casals F, et al. Rare allelic forms of PRDM9 associated with childhood leukemogenesis. *Genome Res*. 2013;23(3):419-430.
42. Robles-Espinoza CD, Harland M, Ramsay AJ, et al. POT1 loss-of-function variants predispose to familial melanoma. *Nat Genet*. 2014;46(5):478-481.
43. Whitworth J, Smith PS, Martin JE, et al.; NIHR BioResource Rare Diseases Consortium. Comprehensive cancer-predisposition gene testing in an adult multiple primary tumor series shows a broad range of deleterious variants and atypical tumor phenotypes. *Am J Hum Genet*. 2018;103(1):3-18.
44. Shah S, Schrader KA, Waanders E, et al. A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia. *Nat Genet*. 2013;45(10):1226-1231.
45. Altmuller F, Lissewski C, Bertola D, et al. Genotype and phenotype spectrum of NRAS germline variants. *Eur J Hum Genet*. 2017;25(7):823-831.
46. Lam WW, Hatada I, Ohishi S, et al. Analysis of germline CDKN1C (p57KIP2) mutations in familial and sporadic Beckwith-Wiedemann syndrome (BWS) provides a novel genotype-phenotype correlation. *J Med Genet*. 1999;36(7): 518-523.
47. Moriyama T, Metzger ML, Wu G, et al. Germline genetic variation in ETV6 and risk of childhood acute lymphoblastic leukaemia: a systematic genetic study. *Lancet Oncol*. 2015;16(16):1659-1666.
48. Topka S, Vijai J, Walsh MF, et al. Germline ETV6 mutations confer susceptibility to acute lymphoblastic leukemia and thrombocytopenia. *PLoS Genet*. 2015;11(6):e1005262.
49. Peacock JD, Dykema KJ, Toriello HV, et al. Oculoectodermal syndrome is a mosaic RASopathy associated with KRAS alterations. *Am J Med Genet A*. 2015;167(7):1429-1435.
50. Raskin L, Guo Y, Du L, et al.; Colon Cancer Family Registry (CCFR). Targeted sequencing of established and candidate colorectal cancer genes in the Colon Cancer Family Registry Cohort. *Oncotarget*. 2017;8(55):93450-93463.
51. Sung PL, Wen KC, Chen YJ, et al. The frequency of cancer predisposition gene mutations in hereditary breast and ovarian cancer patients in Taiwan: from BRCA1/2 to multi-gene panels. *PLoS One*. 2017;12(9):e0185615.
52. Savage SA, Giri N, Jessop L, et al. Sequence analysis of the shelterin telomere protection complex genes in dyskeratosiscongenita. *J Med Genet*. 2011;48(4): 285-288.
53. Nelson DS, van Halteren A, Quispel WT, et al. MAP2K1 and MAP3K1 mutations in Langerhans cell histiocytosis. *Genes Chromosomes Cancer*. 2015;54(6): 361-368.
54. Betti M, Casalone E, Ferrante D, et al. Germline mutations in DNA repair genes predispose asbestos-exposed patients to malignant pleural mesothelioma. *Cancer Lett*. 2017;405:38-45.

55. Chandrasekharappa SC, Chinn SB, Donovan FX, et al. Assessing the spectrum of germline variation in Fanconi anemia genes among patients with head and neck carcinoma before age 50. *Cancer*. 2017;123(20):3943-3954.
56. Lee J, Xiao YY, Sun YY, et al. Prevalence and characteristics of hereditary non-polyposis colorectal cancer (HNPCC) syndrome in immigrant Asian colorectal cancer patients. *BMC Cancer*. 2017;17(1):843.
57. Fostira F, Saloustros E, Apostolou P, et al. Germline deleterious mutations in genes other than BRCA2 are infrequent in male breast cancer. *Breast Cancer Res Treat*. 2018;169(1):105-113.
58. da Silva FM, Jorge AA, Malaquias A, et al. Nutritional aspects of Noonan syndrome and Noonan-related disorders. *Am J Med Genet A*. 2016;170(6):1525-1531.
59. Woodward EL, Olsson ML, Johansson B, et al. Allelic variants of PRDM9 associated with high hyperdiploid childhood acute lymphoblastic leukaemia. *Br J Haematol*. 2014;166(6):947-949.
60. Jafri M, Wake NC, Ascher DB, et al. Germline mutations in the CDKN2B tumor suppressor gene predispose to renal cell carcinoma. *Cancer Discov*. 2015;5(7):723-729.
61. Szolek A, Schubert B, Mohr C, et al. OptiType: precision HLA typing from next-generation sequencing data. *Bioinformatics*. 2014;30(23):3310-3316.
62. Pal Choudhury P, Maas P, Wilcox A, et al. iCARE: an R package to build, validate and apply absolute risk models. *PLoS One*. 2020;15(2):e0228198.
63. He YQ, Wang TM, Ji M, et al. A polygenic risk score for nasopharyngeal carcinoma shows potential for risk stratification and personalized screening. *Nat Commun*. 2022;13(1):1966.
64. Li X, Li Z, Zhou H, et al.; TOPMed Lipids Working Group. Dynamic incorporation of multiple in silico functional annotations empowers rare variant association analysis of large whole-genome sequencing studies at scale. *Nat Genet*. 2020;52(9):969-983.
65. Chen H, Wang C, Conomos MP, et al. Control for population structure and relatedness for binary traits in genetic association studies via logistic mixed models. *Am J Hum Genet*. 2016;98(4):653-666.
66. Yoav Benjamini YH. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B (Methodol)*. 1995;57(1):289-300.
67. Tian W, Zhu FM, Wang WY, et al. Sequence-based typing of HLA-A gene in 930 patients with nasopharyngeal carcinoma in Hunan province, southern China. *Tissue Antigens*. 2015;86(1):15-20.
68. Ning L, Ko JM, Yu VZ, et al. Nasopharyngeal carcinoma MHC region deep sequencing identifies HLA and novel non-HLA TRIM31 and TRIM39 loci. *Commun Biol*. 2020;3(1):759.
69. Yu H, Frank C, Sundquist J, et al. Common cancers share familial susceptibility: implications for cancer genetics and counselling. *J Med Genet*. 2017;54(4):248-253.
70. Albright F, Teerlink C, Werner TL, et al. Significant evidence for a heritable contribution to cancer predisposition: a review of cancer familiality by site. *BMC Cancer*. 2012;12:138.
71. Turati F, Negri E, La Vecchia C. Family history and the risk of cancer: genetic factors influencing multiple cancer sites. *Expert Rev Anticancer Ther*. 2014;14(1):1-4.
72. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet*. 2008;40(6):695-701.
73. Darst BF, Sheng X, Eeles RA, et al. Combined effect of a polygenic risk score and rare genetic variants on prostate cancer risk. *Eur Urol*. 2021;80(2):134-138.
74. Ruan HL, Qin HD, Shugart YY, et al. Developing genetic epidemiological models to predict risk for nasopharyngeal carcinoma in high-risk population of China. *PLoS One*. 2013;8(2):e56128.
75. Ho PJ, Ho WK, Khng AJ, et al. Overlap of high-risk individuals predicted by family history, and genetic and non-genetic breast cancer risk prediction models: implications for risk stratification. *BMC Med*. 2022;20(1):150.
76. Jia G, Lu Y, Wen W, et al. Evaluating the utility of polygenic risk scores in identifying high-risk individuals for eight common cancers. *JNCI Cancer Spectr*. 2020;4(3):pkaa021.
77. Mavaddat N, Michailidou K, Dennis J, et al.; NBCS Collaborators. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet*. 2019;104(1):21-34.
78. Gruhne B, Sompallae R, Masucci MG. Three Epstein-Barr virus latency proteins independently promote genomic instability by inducing DNA damage, inhibiting DNA repair and inactivating cell cycle checkpoints. *Oncogene*. 2009;28(45):3997-4008.
79. Hau PM, Tsao SW. Epstein-Barr virus hijacks DNA damage response transducers to orchestrate its life cycle. *Viruses*. 2017;9(11):341.
80. Lung ML, Cheung AK, Ko JM, et al. The interplay of host genetic factors and Epstein-Barr virus in the development of nasopharyngeal carcinoma. *Chin J Cancer*. 2014;33(11):556-568.
81. Yee Ko JM, Dai W, Wun Wong EH, et al. Multigene pathway-based analyses identify nasopharyngeal carcinoma risk associations for cumulative adverse effects of TERT-CLPTM1L and DNA double-strand breaks repair. *Int J Cancer*. 2014;135(7):1634-1645.
82. Swagemakers SM, Essers J, de Wit J, et al. The human RAD54 recombinational DNA repair protein is a double-stranded DNA-dependent ATPase. *J Biol Chem*. 1998;273(43):28292-28297.
83. Sigurdsson S, Van Komen S, Petukhova G, et al. Homologous DNA pairing by human recombination factors Rad51 and Rad54. *J Biol Chem*. 2002;277(45):42790-42794.
84. Zhou Q, Huang J, Zhang C, et al. The bromodomain containing protein BRD-9 orchestrates RAD51-RAD54 complex formation and regulates homologous recombination-mediated repair. *Nat Commun*. 2020;11(1):2639.
85. Doddato G, Valentino F, Giliberti A, et al. Exome sequencing in BRCA1-2 candidate familias: the contribution of other cancer susceptibility genes. *Front Oncol*. 2021;11:649435.
86. Earl J, Galindo-Pumarino C, Encinas J, et al. A comprehensive analysis of candidate genes in familial pancreatic cancer families reveals a high frequency of potentially pathogenic germline variants. *E Bio Medicine*. 2020;53:102675.
87. Shen GP, Pan QH, Hong MH, et al. Human genetic variants of homologous recombination repair genes first found to be associated with Epstein-Barr virus antibody titers in healthy Cantonese. *Int J Cancer*. 2011;129(6):1459-1466.
88. Shannon-Lowe C, Rickinson AB, Bell AI. Epstein-Barr virus-associated lymphomas. *Philos Trans R Soc Lond B Biol Sci*. 2017;372(1732):20160271.
89. Heslop HE. Biology and treatment of Epstein-Barr virus-associated non-Hodgkin lymphomas. *Hematology Am Soc Hematol Educ Program*. 2005;2005(1):260-266. 10.1182/asheducation-2005.1.260:260-6.
90. Drapkin R, Reardon JT, Ansari A, et al. Dual role of TFIIH in DNA excision repair and in transcription by RNA polymerase II. *Nature*. 1994;368(6473):769-772.
91. Yang ZH, Du B, Wei YS, et al. Genetic polymorphisms of the DNA repair gene and risk of nasopharyngeal carcinoma. *DNA Cell Biol*. 2007;26(7):491-496.
92. Lin H, Lin D, Zheng C. Association of XPD Lys751Gln polymorphism with head and neck cancer susceptibility: evidence from 11,443 subjects. *Diagn Pathol*. 2014;9(1):15.
93. Tong X, Drapkin R, Reinberg D, et al. The 62- and 80-kDa subunits of transcription factor IIH mediate the interaction with Epstein-Barr virus nuclear protein 2. *Proc Natl Acad Sci U S A*. 1995;92(8):3259-3263.
94. Romano RA, Smalley K, Magraw C, et al. ΔNp63 knockout mice reveal its indispensable role as a master regulator of epithelial development and differentiation. *Development*. 2012;139(4):772-782.
95. Kaufmann O, Fietze E, Mengs J, et al. Value of p63 and cytokeratin 5/6 as immunohistochemical markers for the differential diagnosis of poorly differentiated and undifferentiated carcinomas. *Am J Clin Pathol*. 2001;116(6):823-830.
96. Lakshmanachetty S, Balaiya V, High WA, et al. Loss of TP63 promotes the metastasis of head and neck squamous cell carcinoma by activating MAPK and STAT3 signaling. *Mol Cancer Res*. 2019;17(6):1279-1293.
97. Rocco JW, Leong CO, Kuperwasser N, et al. p63 mediates survival in squamous cell carcinoma by suppression of p73-dependent apoptosis. *Cancer Cell*. 2006;9(1):45-56.
98. Chiang CT, Chu WK, Chow SE, et al. Overexpression of delta Np63 in a human nasopharyngeal carcinoma cell line downregulates CKIs and enhances cell proliferation. *J Cell Physiol*. 2009;219(1):117-122.
99. Lin YL, Sengupta S, Gurdziel K, et al. p63 and p73 transcriptionally regulate genes involved in DNA repair. *PLoS Genet*. 2009;5(10):e1000680.
100. Bretz AC, Gittler MP, Charles JP, et al. DeltaNp63 activates the Fanconi anemia DNA repair pathway and limits the efficacy of cisplatin treatment in squamous cell carcinoma. *Nucleic Acids Res*. 2016;44(7):3204-3218.
101. Fotheringham JA, Mazzucca S, Raab-Traub N. Epstein-Barr virus latent membrane protein-2A-induced DeltaNp63alpha expression is associated with impaired epithelial-cell differentiation. *Oncogene*. 2010;29(30):4287-4296.
102. Capasso M, Montella A, Tirelli M, et al. Genetic predisposition to solid pediatric cancers. *Front Oncol*. 2020;10:590033.
103. Toss A, Quarello P, Mascarin M, et al. Cancer predisposition genes in adolescents and young adults (AYAs): a review paper from the Italian AYA working group. *Curr Oncol Rep*. 2022;24(7):843-860.
104. Taeubner J, Wieczorek D, Yasin L, et al. Penetrance and expressivity in inherited cancer predisposing syndromes. *Trends Cancer*. 2018;4(11):718-728.
105. Xue WQ, Qin HD, Ruan HL, et al. Quantitative association of tobacco smoking with the risk of nasopharyngeal carcinoma: a comprehensive meta-analysis of studies conducted between 1979 and 2011. *Am J Epidemiol*. 2013;178(3):325-338.
106. Jia WH, Luo XY, Feng BJ, et al. Traditional Cantonese diet and nasopharyngeal carcinoma risk: a large-scale case-control study in Guangdong, China. *BMC Cancer*. 2010;10:446.