Rare POLN mutations confer risk for familial nasopharyngeal carcinoma through weakened Epstein-Barr virus lytic replication

Ruo-Wen Xiao,a,1 Fang Wang,a,b,1 Tong-Min Wang,a,1 Jiang-Bo Zhang,a,1 Zi-Yi Wu,a Chang-Mi Deng,a Ying Liao,a Ting Zhou,a Da-Wei Yang,a Si-Qi Dong,a Wen-Qiong Xue,a Yong-Qiao He,a Xiao-Hui Zheng,a Xi-Zhao Li,a Pei-Fen Zhang,a Shao-Dan Zhang,a Ye-Zhu Hu,a Yu-Ying Liu,a Yun-Fei Xia,e Song Gao,a Jian-Bing Mu,f Lin Feng,a* and Wei-Hua Jiaa*

aState 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, Guangdong, China
bDepartment of Radiation Oncology, Affiliated Cancer Hospital & Institute of Guangzhou Medical University, Guangzhou, Guangdong, China
cGuangzhou Women and Children’s Medical Center, Guangzhou Medical University, Guangzhou, Guangdong, China
dScreening Center for Cancer Prevention, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, China
eDepartment of Radiation Oncology, Sun Yat-sen University Cancer Center, Guangzhou, Guangdong, China
fLaboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institute of Health, Rockville, Maryland, USA

Summary

Background Nasopharyngeal carcinoma (NPC) exhibits significant familial aggregation; however, its susceptibility genes are largely unknown. Thus, this study aimed to identify germline mutations that might contribute to the risk of familial NPC, and explore their biological functions.

Methods Whole-exome sequencing was performed in 13 NPC pedigrees with multiple cases. Mutations co-segregated with disease status were further validated in a cohort composed of 563 probands from independent families, 2,953 sporadic cases, and 3,175 healthy controls. Experimental studies were used to explore the functions of susceptibility genes and their disease-related mutations.

Findings The three rare missense mutations in POLN (DNA polymerase nu) gene, P577L, R303Q, and F545C, were associated with familial NPC risk (5/576 [0.87%] in cases vs. 2/3374 [0.059%] in healthy controls with an adjusted OR of 44.84 [95% CI: 1.91-514.34, p = 2.25 × 10−3]). POLN was involved in Epstein-Barr virus (EBV) lytic replication in NPC cells in vitro. POLN promoted viral DNA replication, immediate-early and late lytic gene expression, and progeny viral particle production, ultimately affecting the proliferation of host cells. The three mutations were located in two pivotal functional domains and were predicted to alter the protein stability of POLN in silico. Further assays demonstrated that POLN carrying any of the three mutations displayed reduced protein stability and decreased expression levels, thereby impairing its ability to promote complete EBV lytic replication and facilitate cell survival.

Interpretation We identified a susceptibility gene POLN for familial NPC and elucidated its function.

Funding This study was funded by the National Key Research and Development Program of China (2021YFC2500400); the National Key Research and Development Program of China (2020YFC1316902); the Basic and Applied Basic Research Foundation of Guangdong Province, China (2021B1515420007); the National Natural Science Foundation of China (81973131); the National Natural Science Foundation of China (82003520); the National Natural Science Foundation of China (81903395).

Copyright © 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

*Corresponding author at: State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, 651 Dongfeng East Road, BLDG 2, RM903, Guangzhou, Guangdong, PR China. 510060.
E-mail addresses: fengl@sysucc.org.cn (L. Feng), jiawh@sysucc.org.cn (W.-H. Jia).
1 Ruo-Wen Xiao, Fang Wang, Tong-Min Wang, and Jiang-Bo Zhang contributed equally to this work.
Introduction

Nasopharyngeal carcinoma (NPC) is a malignancy arising from the epithelium of the nasopharynx, characterized by an extraordinary ethnic and geographical distribution. It is rare in most parts of the world but is endemic in Southern China (particularly in Guangdong), Southeast Asia, Northern Africa, and Alaska. Familial clustering has been widely reported, and more than 10% of patients have a positive family history of NPC in endemic regions. Familial NPC is defined as a family with one or more NPC patients, in addition to the proband within three generations. The risk dramatically increased 4 to 10-fold among individuals with a first-degree relative with NPC compared with those without a family history. Familial NPC cases also have a significantly younger age of onset than sporadic cases. The geographical distribution and familial clustering of NPC suggest that genetic factors may play a critical role in its aetiology. Therefore, it is important to identify genetic susceptibility genes for NPC screening in high-risk populations, which would contribute to earlier diagnosis and better prognosis.

Linkage studies have demonstrated the contribution of some loci to the risk of familial NPC, including 6p22, 4p15.1–q12, 3p21.31–21.2, and 5p13.1, to the risk of familial NPC. To date, only a few studies have reported the susceptibility genes for familial NPC in the Southern Chinese population, including rare mutations in *MST1R* and 12 other genes.

**Research in context**

**Evidence before this study**

Previous linkage studies have demonstrated the contribution of certain chromosomal regions, such as 6p22, 4p15.1–q12, 3p21.31–21.2, and 5p13.1, to the risk of familial NPC. To date, only a few studies have reported the susceptibility genes for familial NPC in the Southern Chinese population, including rare mutations in *MST1R* and 12 other genes.

**Added value of this study**

This study identified a susceptibility gene *POLN* for familial NPC, supporting that germline mutations in DNA repair factors are associated with NPC risk. Functional studies have uncovered the substantial function of *POLN* in the course of EBV lytic replication and have provided experimental evidence for the involvement of germline mutations in EBV-related oncogenesis.

**Implications of all the available evidence**

Identification of susceptibility genes for NPC screening in high-risk populations could contribute to earlier diagnosis and improved prognosis, especially in those with a familial background. Clarifying the mechanism by which the susceptibility gene contributes to NPC development is helpful to deepen our understanding of the aetiology of NPC.

**Keywords:** Familial nasopharyngeal carcinoma; Whole-exome sequencing; *POLN* gene; Rare germline mutations; Epstein-Barr virus

We found that rare *POLN* mutations are associated with familial NPC risk. *POLN* encodes a DNA polymerase with translesion DNA synthesis activity, and participation in cross-linking repair and homologous recombination. We further demonstrated the functional
relevance of POLN in EBV lytic replication and the possible mechanism by which rare functional mutations influence familial NPC risk.

Methods

Study populations and samples collection
The NPC multiplex families were recruited from “High-risk Nasopharyngeal Carcinoma Family Screening Program” in Sun Yat-sen University Cancer Center (SYSUCC, Guangzhou, Guangdong, China). Patients and their unaffected relatives from 13 high-risk families were enrolled in this study. Each family had whole blood or saliva samples available from at least two NPC patients and at least one unaffected first-degree relative (Figure S8). WES was conducted on available samples, including 35 familial cases and 34 unaffected relatives (Table S1).

A total of 351 additional familial cases, defined as having at least one family member within third-degree relatives suffering NPC, were recruited from SYSUCC from 2008 to 2015 for Sanger sequencing. In addition to the public database, an in-house WES database, including 199 healthy individuals recruited from Guangdong, China, was also used for mutation filtering (Table S2).

For independent replication, 212 unrelated familial cases and 2,953 sporadic cases were recruited from SYSUCC from 2004 to 2017, and 3,175 healthy controls were recruited from Guangdong China from 2009 to 2016 (Table S2). All patients were Han Chinese and were histopathologically diagnosed.

Additionally, mRNA samples from 84 NPC tissues and 12 non-cancerous nasopharyngeal tissues were obtained from the SYSUCC biobank. Protein samples were obtained from 8 additional NPC tissues and 3 non-cancerous nasopharyngeal tissues. All NPC tissues were pathologically classified as undifferentiated non-keratinising tumours.

Sequencing and genotyping
Genomic DNA was extracted from peripheral blood and saliva using an automated workstation (Chemagic Star, Bonaduz, Switzerland) and quantified using the Picogreen reagent (Invitrogen, Carlsbad, CA, USA). 2.5 μg of DNA from each individual was fragmented using Covaris ultrasonicator (Woburn, MA, USA). Indexed paired-end libraries were prepared using the SureSelectXT Human All Exon Kit (V6) (Agilent, Cat#5190-8863, Santa Clara, CA, USA). paired-end 150-bp read-length sequencing was performed using Illumina HiSeq 2500 technology (San Diego, CA, USA).

The WES data were analysed according to the recommendations of GATK Best Practices.28 The sequence reads were aligned to the human reference genome (hg19/GRCh37) using the Burrows-Wheeler aligner and Picard tools (http://broadinstitute.github.io/picard/) were used to sort the bam files and mark duplicates. GATK was then used to conduct indel realignment, base quality score recalibration, variant calling, and variant quality score recalibration (VQSR).28 The sensitivity thresholds for single-nucleotide polymorphisms (SNPs) and indels were chosen at 99.5% and 99%, respectively. The quality of the mutations was checked by GATK and plink to remove abnormal samples or contamination. Genotypes with quality scores < 20 or read depths < 8 were treated as missing values. Heterozygotes with alternative allele frequencies > 0.75 or < 0.25 were also considered ambiguous calling. Mutations with missing rates of > 20% were excluded from further analyses. ANNOVAR30 was used to annotate the qualified mutations.

For Sanger sequencing of POLN gene, we used the Primer-BLAST online database (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) to design 26 pair primers (Table S7) for all exons and exon-intron boundaries of POLN. High fidelity DNA polymerase (Genestar, Beijing, CHN) was used for PCR. Amplified DNA fragments were sent for Sanger sequencing (Majorbio, Shanghai, CHN). All sequencing results were aligned to the POLN RefSeq sequence (NG_046934.1) using the DNAMAN software for mutation detection. We applied Sanger sequencing to validate the mutations identified by WES and TaqMan genotyping assays (Figure S1), and partial samples carried wild-type POLN identified by TaqMan genotyping assay (Figure S9).

TaqMan SNP Genotyping Assay (Applied Biosystems, ABI, Carlsbad, CA, USA) followed by the LightCycler Real-Time PCR 480 system (Roche, Basel, Switzerland) was used for genotyping, according to the manufacturer’s instructions. Blank and positive controls were used for each experiment. Probe sequences are listed in Table S8. The 5 μl reaction system included 0.25 μl primer plus probe, 1 μl DNA (25 ng), 2.5 μl TaqMan universal MMIX II with UNG (ABI, Cat#4440046) and 1.25 μl H₂O. The reaction conditions were as follows: incubation at 50 °C for 2 min, polymerase activation at 95 °C for 10 min, and 40 cycles of amplification at 95 °C for 15 s and 60 °C for 1 min.

Mutation selection
To identify candidate mutations that may play important roles on NPC development, we first conducted a co-segregation analysis on the 13 families to identify mutations segregating with NPC onset, that is, shared by all NPC cases but absent in the unaffected relatives in the family. We then filtered the non-synonymous rare mutations with minor allele frequency (MAF) < 0.1% in all public databases: 1000 Genomes (all population and East Asians), gnomAD (all population and East Asians with genome and exome data, respectively),
ESPD500, ExAC (all population and East Asians with and without TCGA data, respectively), the Kaviar database, and our in-house WES database with 199 healthy individuals. Among the resulting 145 rare non-synonymous co-segregating mutations, we used the dbNSFP database in ANNOVAR to conduct functional prediction. The mutations predicted as “damaging” in SIFT, polyphen2 HVAR, and FATHMM and located in the protein domain were further filtered. We performed a literature search for target gene(s) with functional relevance to cancer. A detailed mutation selection flowchart is shown in Figure 1a.

The mutations identified in 351 unrelated familial NPC cases were filtered by population frequency using the selection strategies aforementioned. Those located in the protein domain and predicted as “damaging” in SIFT, polyphen2 HVAR, and Mutation Taster were further selected for the functional study.

Cell culture
HK1 EBV+ and Tet-BZLF1/CNE2 EBV+ NPC cells were obtained from Professor Mu-Sheng Zeng (Sun Yat-sen University, China).31-32 Tet-BZLF1/CNE2 EBV+ cells conditionally expressed the exogenous protein Zta under the control of a tetracycline-inducible promoter (Figure S10). Both NPC cell lines were cultured in RPMI 1640 medium (Gibco, Cat# 11875093, Carlsbad, CA, USA) supplemented with 10% foetal calf serum (Gibco, Cat#16140071). Geneticin (700 μg/ml) (G418) (Sigma, Cat#G5013) was added to the culture medium to maintain the EBV-positive rate. Human embryonic kidney 293T (HEK-293T) cells used for virus packaging were purchased from ATCC (ATCC, CRL-3216). HEK-293T cells were cultured in Dulbecco’s Modified Eagle’s Medium (DMEM) (Gibco, Cat# 11965092) with 10% foetal calf serum. All cells were cultured at 37 °C in an incubator with a humidified atmosphere of 5% CO2 and tested negative for mycoplasmas (Figure S11).

Plasmids, mutagenesis
POLN full-length cDNA (NCBI: NM_181808.3) was obtained from Vigene Biosciences (Rockville, MD, USA) and cloned into the pCDH-CMV-MCS-EF1-Puro lentiviral vector using the ClonExpress Entry One Step Cloning Kit (Vazyme, Cat#C112, Nanjing, Jiangsu, CHN) for wild-type (WT) POLN expression plasmid construction. We performed polymerase chain reaction (PCR) to generate POLN P577L, F545C, and R303Q mutants by site-directed mutagenesis of the WT POLN expression plasmid with specific primers (Table S7) and used DpnI restriction enzyme (Thermo Scientific, Cat#ER1701, Waltham, MA, USA) to digest the WT template. HA-tags were fused to the C-terminus of the WT and mutant POLN plasmids. All plasmids were verified using DNA sequencing.

Lentivirus assembly and transduction
For lentiviral production, HEK-293T cells were co-transfected with PMD2G, psPAX2, and pCDH POLN WT or mutant plasmids using polyethyleneimine (PEI) (YEASEN, Cat#40816ES02, Shanghai, CHN). The supernatant medium was refreshed 12 h after transfection. Lentivirus particles were collected at 48 h and 72 h after transfection. NPC cells were seeded in 6-well plates and cultured until they reached approximately 70% confluency. Then, cells were transduced with the negative control, POLN WT, or mutant lentivirus particles at the same multiplicity in the presence of 8 μg/ml of polybrene (Solarbio, Cat# H8761, Beijing, CHN). Real-time quantitative polymerase chain reaction (RT-qPCR) and western blotting were used to examine POLN expression (Figure S12, S5, S3 and S7).

RNA interference of POLN
The siRNA duplexes used to knockdown POLN were purchased from RiboBio (Guangzhou, CHN). The commercial transfection reagent Lipofectamine 3000 (Cat # L1000075; Invitrogen) was used for cell transfection. The siRNA sequences used are listed in Table S9. siRNA efficacy was confirmed by RT-qPCR and western blotting (Figure S13 and S4).

RNA extraction, reverse transcription, and real-time quantitative polymerase chain reaction
Total RNA was extracted from cells and tissues using TRIzol reagent (Invitrogen, Cat#15596026) following the manufacturer’s protocol. Total RNA was reverse-transcribed using the PrimeScript RT reagent Kit with gDNA Eraser (TaKaRa, Cat#RR047A, Japan). RT-qPCR was performed using a LightCycler 480 System (Roche) with SYBR Green I Master (Roche, Cat#04887352001) on a LightCycler 480 System (Roche). The primers used for POLN, BZLF1, BRLF1, BKRF4, BOLF1, BLLF1, and GAPDH are listed in Table S7. GAPDH was used as an internal control to normalise the relative expression of the genes. The relative mRNA expression was calculated using the comparative Ct method 2−ΔΔCt.

Protein Extraction and Western Blot analysis
Proteins were extracted using a whole-cell lysis assay (KeyGEN BioTECH, Cat#KGP2100, Nanjing, Jiangsu, CHN). BCA Assay Reagent (CwbioTech, Cat#CW0014S, Beijing, CHN) was used to measure protein concentration. Protein lysates (20 μg) were electrophoresed on a 10% sodium dodecyl sulfate (SDS) (Aladdin, Cat#S108346, Shanghai, CHN) polyacrylamide gel and transferred onto PVDF membranes (Merck Millipore, Cat#IPVH00010, Billerica, MA, USA). Membranes were blocked with TBST (Solarbio, Cat#T1085) containing 5% non-fat dried milk (BIO-RAD, Cat#1706404, Hercules, CA, USA) and incubated overnight at 4 °C.
Figure 1. Identification of three POLN mutations associated with familial NPC. (a) Workflow for the identification and validation of the association between POLN and familial NPC. (b) Pedigree of FA NPC family. DNA samples of individuals I-1, I-2, II-1, II-2, and II-3 were extracted for WES. (c) The frequencies of the POLN P577L, R303Q, and F545C rare mutations in familial cases, sporadic cases, and controls, respectively. Logistic regression analysis was conducted to compare the frequencies of all three mutations in familial NPC cases with healthy controls after adjusting for age and sex.

NPC, Nasopharyngeal carcinoma; WES, whole-exome sequencing.
Cytofluorography

Cells were washed twice with phosphate-buffered saline (PBS) (Gibco, Cat#10010-044), fixed with 4% paraformaldehyde (G-CLONE, Cat# PN4204, Beijing, CHN) for 10 min, and blocked with 5% BSA (MP, Cat# 9048-46-8, Santa Ana, CA, USA) for 1 h, followed by incubation with POLN (1:250) (Thermo Scientific, Rabbit, Cat# PA5-65162, RRID: AB_2663861) and BMRF1 (1:250) (Abcam, mouse, Cat# ab30541, RRID: AB_732194, Cambs, UK) primary antibodies at 4 °C overnight. After three washes in TBST, the coveslips were incubated with Alexa Fluor 555 goat anti-mouse IgG (1:1,000) (CST, Cat# D3072) and Alexa Fluor Plus 647 donkey anti-rabbit IgG (1:1,000) (CST, Cat# D3072) for 1 h in a dark room. The coveslips were washed three times and stained with DAPI (Solarbio, Cat# DAPI, Wuhan, Hubei, CHN) for 5 min at RT. After drying and mounting with ProLong Gold Antifade Mountant (Invitrogen, Cat#P36930), the subcellular localisation of POLN and BMRF1 proteins was analysed using a confocal microscope (ZEISS, LSM880 with fast airyscan, Oberkochen, Germany). ZEN Blue software was used for the quantification of colocalisation between BMRF1 and POLN.

Flow Cytometry

The cells were collected and washed twice with PBS. The percentage of GFP-positive cells was evaluated by flow cytomtery (CytoFLEX 2; Beckman, Brea, CA, USA) using signal wavelength analysis.

EBV DNA extraction and copy number detection

Cell culture media used for determining extracellular genome copies were treated with DNase I (Vazyme, Cat# EN401/402) before DNA extraction to digest cell-free DNA which was not detected in the DNase-resistant viral particles. Total DNA was extracted from cells and culture media simultaneously using the Blood/Cell/ Tissue Genomic DNA Extraction Kit (TIANGEN, Cat# DP304, Beijing, CHN), following the manufacturer’s protocol. EBV DNA copy numbers were determined by RT-qPCR using primers targeting the EBV BALF5 and β-Globin genes of the human genome (Table S7). Standard sample ladders (10^3, 10^4, 10^5 copies/μl) were used to generate a standard curve. The relative EBV DNA copy number in cells was expressed as the ratio of the copy number of the EBV genome to that of the β-Globin DNA.

Cell Viability Assay

One thousand NPC cells were seeded in each well of a 96-well plate and cultured for one to five days. Ten microlitres of Cell Counting Kit-8 (CCK8) (APExBIO, Cat# C0008, Houston, TX, USA) solution was added to each well. After incubation for 2 h, the absorbance was measured at 450 nm using a spectrophotometer (BIO-Tek epoch, Winooski, VT, USA).

Statistical analysis

Statistical analyses were performed using R (version 3.5.0) and GraphPad Prism (version 9.0.1). Logistic regression adjusted for age and sex was used to compare the mutation frequencies of mutations in familial NPC cases or sporadic NPC cases with those in healthy controls. Student’s t-test was conducted to compare the age of onset between P577L carriers and non-carriers in familial NPC cases. Experimental data are presented as the mean ± standard deviation (SD) from three separate experiments. A two-sided Student’s t-test was used to compare data points between the two groups. Two-way analysis of variance (ANOVA) was used to test the statistical significance of the growth difference between the two groups. The results were considered statistically significant at p < 0.05.
POLN mutations identified in additional familial NPC cases

Sanger sequencing of all POLN exons and exon-intron boundaries was performed in 351 additional familial NPC cases. The clinical characteristics of these cases are described in Table S2. To investigate other mutations in POLN, we integrated the 351 familial NPC cases with the WES data of the 13 families, excluding the FA family, and found a total of 40 mutations in POLN among these 363 familial NPC cases (Table S5). Three rare mutations were predicted to be damaging. We detected one more POLN P577L mutation carrier (female, age at onset: 42 years). Additionally, two additional POLN rare mutations, R303Q (exon6: c.G908A) carried by a male patient diagnosed at 41 years of age and F545C (exon15: c.T1634G) carried by a male patient diagnosed at 42 years of age were found. Sanger sequencing was performed to validate these mutations (Figure S1).

POLN mutations are associated with an increased risk of familial NPC

Genotyping of three POLN mutations (P577L, F545C, and R303Q) was performed in an independent cohort of 212 familial NPC cases, 2,953 sporadic NPC cases, and 3,175 healthy controls (Table S2). The P577L mutation was detected in five individuals, including one patient with familial NPC (diagnosed at 43 years of age), three sporadic patients (diagnosed at 68, 42, and 40 years of age respectively), and one healthy control (60-year-old). The R303Q mutation was detected in one sporadic patient (diagnosed at 65 years of age) and one healthy control (59-year-old). No additional F545C mutation was detected in these samples. The genotypes of all mutation carriers were validated using Sanger sequencing (Figure S1).

Combining all the independent samples included in this study (576 independent familial cases [13 for WES, 563 for validation], 2,953 sporadic cases, and 3,374 healthy controls), the POLN rare damaging mutations showed a significantly higher frequency in familial NPC cases (5/376, 0.8%)) than in healthy controls (2/3374, 0.059%) with an adjusted OR of 44-21 (95% CI: 4.95-214.59; logistic regression, p = 2.25 × 10^-7) (Figure 1c). Additionally, these rare damaging mutations showed a relatively higher frequency in sporadic NPC cases (4/2953, 0.135%) than in healthy controls (2/3374, 0.059%) with an adjusted OR of 2.41 (95% CI: 1.04-13.21; logistic regression, p = 0.31). We also observed a younger age of onset in the familial patients carrying P577L (mean age: 39.67 years) than in those not carrying this mutation (mean age: 47.95 years, Student’s t-test, p = 0.096).

We further analysed WES data from publicly available NPC studies to investigate the mutations of POLN in NPC cases (Table S6). POLN P577L, R303Q, and F545C mutations were not detected in these
studies. However, we found 14 additional POLN rare mutations in 19 NPC cases, including one synonymous mutation, six missense mutations, one frameshift mutation, and six intronic mutations.

**POLN is involved in the completion of EBV lytic replication**

POLN was reported to interact with DNA repair factors Rad51 and PCNA, which have been shown to be involved in EBV DNA replication. To investigate the function of POLN in EBV DNA replication, we examined the subcellular localisation of POLN and BMRF1 (DNA processivity factor, completely coincided with the newly synthesized EBV DNA) in two EBV-positive NPC cell lines, HK1 EBV+ and Tet-BZLF1/CNE2 EBV+. After lytic replication induction, we found that POLN was colocalised with the BMRF1 protein (Manders' Colocalization Coefficients: HK1 EBV+: 0.91, Tet-BZLF1/CNE2 EBV+: 0.87) (Figure 2a and Figure S2a), suggesting that POLN may tend to be located in EBV replication compartments. Thus, we examined whether POLN expression affected EBV DNA replication. The two cell lines contained an exogenous recombinant EBV strain that expressed enhanced green fluorescent protein (GFP). We overexpressed POLN in both cell lines and observed that cells overexpressing POLN showed a higher percentage of GFP-positive cells than the control, with and without lytic replication induction (Figure 2b and Figure S2b). We further confirmed this result by quantifying the intracellular EBV DNA copy number using RT-qPCR. Overexpression of POLN increased the intracellular EBV genome copy number (Figure 2c and Figure S2c). Next, we used two siRNAs to knock down the expression of POLN in the two cell lines, and found that the knockdown of POLN decreased the percentage of GFP-positive cells compared to the control, with and without lytic replication induction (Figure 2d and Figure S2d). These results suggested that POLN plays a substantial role in EBV DNA replication.

Viral IE genes BZLF1 and BRLF1 play critical roles in the initiation of lytic replication of EBV, both of which were upregulated in POLN-overexpressing cells (Figure 2f and Figure S2f, S3f), and downregulated in POLN knockdown cells (Figure 2g and Figure S2g, S4g). The expression of EBV true late genes is strictly dependent on viral DNA replication. Our results showed that the expression of true late genes BZLF1, BRLF1, and BLLF1 was up-regulated at least 1.5-fold in both cell lines which overexpressed POLN with and without EBV reactivation, and the knockdown of POLN decreased their expression (Figure 2h and Figure S2h). POLN promotes EBV DNA replication and increases viral true late gene expression, which encouraged us to investigate whether POLN affects progeny viral particle production. The amount of DNase-resistant extracellular viral particles
Figure 2. POLN is involved in the completion of EBV lytic replication. POLN was overexpressed or knocked down in HK1 EBV+ cells, twenty-four hours after induction of EBV lytic replication by TPA/NaB, cells were harvested. (a) Immunofluorescence (left panel) was performed to show the location of BMRF1 (red) and POLN (green) in HK1 EBV+ cells, DAPI (blue). The picture on the right panel was performed to show the location of β-actin (green) and POLN (red) in HK1 EBV+ cells, DAPI (blue).
in cell culture supernatants was evaluated and the results showed that the extracellular EBV DNA copy number was increased when POLN was overexpressed and decreased when POLN was knocked down (Figure 2i and Figure S2i). These findings revealed that POLN was involved in the completion of EBV lytic replication.

POLN inhibits the proliferation of NPC cells and shows lower expression in NPC tissues

In epithelial cells, the latent infection and abortive/incomplete lytic infection of EBV in host cells have been shown to contribute to NPC carcinogenesis.\textsuperscript{24,45} In complete lytic cycle, productively infected cells eventually lyse, resulting in the inhibition of cell proliferation.\textsuperscript{24} Based on the finding that POLN plays a role in complete lytic reactivation and viral production, we subsequently evaluated the effect of POLN on NPC cell proliferation. CCK8 assay showed that overexpression of POLN significantly inhibited the proliferation of NPC cells (Figure 3a), whereas knockdown of POLN promoted the proliferation of NPC cells (Figure 3b).

Next, we examined POLN mRNA expression in 12 non-cancerous nasopharyngeal tissues and 84 NPC tissues using RT-qPCR and found lower POLN expression showed the scatterplot of colocalisation between BMRF1 and POLN. (b) The effect of POLN overexpression on the proportion of GFP-positive cells. (c) The effect of POLN overexpression on the intracellular EBV genome copy numbers (d) The effect of POLN knockdown on the proportion of GFP-positive cells. (e) The effect of POLN knockdown on the intracellular EBV genome copy numbers. (f) The effect of POLN overexpression on the mRNA expression level of EBV IE genes BZLF1 and BRLF1. (g) The effect of POLN knockdown on the mRNA expression level of EBV IE genes BZLF1 and BRLF1. (h) The effect of POLN overexpression or knockdown on the mRNA expression level of EBV true late genes BKRF4, BOLF1, and BLLF1. (i) The effect of POLN overexpression or knockdown on the extracellular EBV genome copy numbers. (b-i) Data were presented as mean ± SD (N = 3), ***p < 0.001, **p < 0.01, *p < 0.05, ns: not significant (Student’s t-test).

EBV, Epstein-Barr virus; SD, Standard deviation; WT, wild type.
in NPC tissues (Figure 3c). We also obtained gene expression data from GEO datasets (https://www.ncbi.nlm.nih.gov/geo/) (GSE12452). Consistently, POLN was lower in NPC tissues (n=10) than in normal nasopharyngeal tissues (n=31) (Figure 3d). To further confirm this finding, we examined POLN protein expression in five non-cancerous nasopharyngeal samples and eight clinical NPC samples by western blotting, and found lower POLN protein expression in NPC tissues (Figure 3e).

The three rare damaging mutations impaired the POLN protein stability and decreased the protein expression

POLN is a 100.3-kDa protein with 900 amino acids and two functional domains, DnaQ_like_exo and DNA_poly_A theta (Figure 4a). The three identified amino acids were located within the functional domains and showed high evolutionary conservation in five orthologous proteins (Figure 4b). We used INPS (http://inps.biocomp.unibo.it) to predict the thermodynamic free energy change for the three mutations and found decreased stability in the mutant protein. We further performed molecular modeling of the 194-859 amino acid region of POLN (PDB 4XVK) (https://www.rcsb.org/structure/4XVK) using the PyMOL software to evaluate the potential structural changes induced by the mutations (Figure 4c). Arg 303 (basic) was thought to be involved in a salt bridge interaction with Asp 335 (acidic) (distance: 2.5 Å), and the Arg to Gln (R303Q) mutation was predicted to impair this interaction (Figure 4d). Substitution of Pro with Leu (P577L) was predicted to induce the loss of conformational constraint ability and, therefore, may destabilise the protein structure. P577L and F545C are located near the DNA-binding site (residues 573-576 and 543-544, respectively) and were predicted to weaken the interaction with DNA (8.5 Å for Pro577 versus 9.3 Å for Leu577, 4.6 Å for Phe545 versus 6.9 Å for Cy545) (Figure 4e-4f). Phe545 is involved in the formation of π-π stacking interactions with DNA, and the substitution of Phe with Cys was predicted to impair this interaction. In summary, the three identified amino acid substitutions were predicted to impair POLN protein stability and slightly decrease its DNA-binding ability.

We performed a cycloheximide (CHX) chase experiment to analyse the protein stability of the WT and mutant POLN. Consistent with the results of the above in silico analysis, the three mutants showed shorter degradation half-lives than WT POLN (Figure 4g-4h). Moreover, we performed a western blot assay to assess the effect of mutations on protein expression and found a reduction in POLN protein expression levels in all three mutant types, compared to WT POLN (Figure 4i). Moreover, the R303Q and P577L mutated POLN protein showed lower expression levels than those of the F545C mutation. Both bioinformatics and experimental results suggested that the three mutations impaired POLN protein stability by shortening its degradation half-life.

The three rare damaging mutations impair the function of POLN in EBV lytic replication and NPC cells proliferation

We have observed that POLN is involved in EBV lytic cycle, including the process of EBV DNA replication, IE genes and true late genes expression, as well as progeny EBV particles production. As the three mutations have been shown to decrease the protein stability and expression of POLN, we further investigated whether these mutations impair its function. We overexpressed WT POLN and mutant POLN harbouring R303Q, P577L, and F545C in HK1 EBV+ and Tet-BZLF1/CNE2 EBV+ cell lines, using lentiviruses at the same multiplicity. The mRNA expression levels of the WT and mutant POLN were similar (Figure 5a). Twenty-four hours after lytic replication induction, we examined the percentage of GFP-positive cells and measured intracellular EBV DNA titrations together with DNase-resistant extracellular viral particles. Expression of each of the three mutant POLN resulted in a lower percentage of GFP-positive cells (Figure 5b and Figure S6a) and lower intracellular EBV DNA copy numbers (Figure 5c and Figure S6b) than those detected in WT POLN. Consistently, a decreased extracellular viral particle EBV genome copy number was detected in NPC cells overexpressing mutant POLN (Figure 5d and Figure S6c). Compared to WT POLN, mutant POLN decreased the expression of IE genes BZLF1 and BRLF1 (Figure 5d and Figure S6d, S7), and true late genes BKRF4, BOLF1 and BLLF1 (Figure 5e and Figure S6e). The results of the cell viability assay showed that the mutations weakened the inhibitory effect of POLN on the proliferation of NPC cells (Figure 5f and Figure S6f). The P577L and R303Q mutations showed a greater effect than the F545C mutation. These results suggest that the three rare mutations could impair the function of POLN in promoting EBV lytic replication and inhibiting cell proliferation, which indicates the important role of these mutations in NPC development.

Discussion

Using co-segregation and association analyses in multiplex NPC families and an independent validation cohort of familial cases, sporadic cases, and healthy controls (N=6,890), we identified three germline mutations, R303Q, P577L, and F545C of POLN, which confer susceptibility to familial NPC. By integrating bioinformatics analysis and functional assays, we showed that compared with WT POLN, the three mutations decreased the protein expression and stability of POLN,
Figure 4. The efficacy of mutations on POLN protein. (a) Top, schematic representation of POLN gene (reference sequences: NM_181808.4), which was composed of 26 exons, the solid black rectangles indicated exons and the gray horizontal bars introns.
impaired the function of POLN in EBV lytic replication, and promoted cell proliferation.

We and others have identified genetic mutations in some DNA repair genes that are associated with NPC susceptibility, such as those involved in base excision repair (XRCCL, XRC2, XRC3), homologous recombination repair (XRCC1, XRCC2, XRCC3), nucleotide excision repair (ERCC1, ERCC2, ERCC3), and DNA cross-link repair.60 The cellular condition of the S-phase cell response transducers to orchestrate its life cycle.64 POLN is also a DNA repair factor and is mainly responsible for translesion DNA synthesis.26 and DNA cross-link repair.65 The cellular condition of the S-phase cell cycle appears to favor viral lytic replication.66 The expression level of POLN is the highest in the S-phase cell cycle.27 In addition, EBV lytic replication-related DNA repair factors Rad51 and PCNA can interact with POLN, and this interaction only occurs in the S phase.27 These findings indicate that the DNA repair factor POLN may participate in EBV lytic replication.

Lytic or latent infection or abortive lytic infection that contributes to EBV-associated oncogenesis remains obscure. In EBV-associated lymphoma, lytic infection may contribute to oncogenesis, and deletions in the viral genome, such as the BART region, were found to promote EBV lytic reactivation.67 However, for EBV-associated epithelial cancers, deletions in the EBV genome are rare, and BART-encoding miRNAs are highly expressed, which promotes latent EBV infection. Traditionally, EBV in NPC was classified as latency II, in which EBNA1, LMP1, and LMP2A are expressed at the protein level, but antibody levels for lytic antigens (Zta, Rta, and VCA) increased in NPC sera, suggesting that lytic infection may contribute to NPC carcinogenesis. In addition to IHC which detects protein expression, mRNA detection by RT-PCR and recent single-cell RNA sequencing has revealed EBV antigen expression in NPC. In one study of NPC biopsies, all eight samples were positive for the 1E gene BZLF1, but only five of the eight samples tested positive for the early gene BMLF1 expression, and none tested positive for another early gene BBLF3 by RT-PCR assays.68 In NPC single-cell sequencing, both latent genes (LMPs and EBNA) and a few lytic genes (A73, BARF0, BFRF1, and BZLF1) were expressed, but no mRNAs of late lytic genes were detected.69 Thus, NPC tumours harbour cells undergoing an incomplete or abortive lytic phase, in which some early lytic genes are expressed but no viral particles are produced. Therefore, abortive/incomplete lytic infection of epithelial cells is highly likely to contribute to NPC carcinogenesis. The gene we identified, POLN, facilitated complete lytic reactivation and viral production, and its mutations failed to support EBV production, which is in line with the in vivo scenario. For individuals carrying these three mutations, EBV may also be prone to latent infection or abortive lysis in epithelial cells, facilitating the survival and malignant transformation of EBV-infected epithelial cells (Figure 6).

As a multifactorial disease, a complex interplay between host genetics, EBV infection, and environmental factors has been shown to be involved in NPC. Many lines of evidence indicate that host heredity interacts with EBV infection and plays a critical role in disease development. We previously reported that human genetic variants of homologous recombination repair genes are associated with EBV antibody titres in healthy Cantonese.70 Recently, we found that nasopharyngeal EBV loads are positively correlated between relative pairs from NPC multiplex families and have significantly high heritability.71 EBV infection status could modulate the role of genetic polymorphism in NPC.72 However, laboratory evidence of the interaction between the etiological factors of NPC is lacking. Our functional experiments validated that POLN participates in the completion of EBV lytic replication. In spontaneous lytic infection (low-level lytic activation) or induced lytic infection (high-level lytic activation), POLN function was impaired by the deleterious mutations identified in familial NPC.

We performed an integrated study combining bioinformatics analysis and functional experiments to explore germline mutations in NPC and demonstrated how rare mutations in the POLN exon region contribute

---

Bottom, schematic of the POLN protein (reference sequences: NP_861524.2) showing its different domains, DnaQ_like_exo (green) and DNA_pol_A_7ha (blue). Dashed lines indicated the location of mutations. (b) Amino acid conservation analysis. Multiple sequence alignment of POLN ortholog protein across six species (ClustalW, BioEdit software) showed evolutionary conservation at the site of the identified amino acid substitution (indicated by the red arrow). (c) 3D representation of the crystal structure of the 194-859 amino acid region of POLN (PDB 4XVK). The position of mutated residues was highlighted in pink ball, DNA in orange, PyMOL software (Version 1.7.0.0) was used for visualization. (d-f) The interaction and distance modification resulting from the POLN mutations. The WT and mutated residues were colored by gray and red, respectively. The red double arrow line showed the interaction distance. (g-i) HK1 EBV+ and Tet-BZLF1/CNE2 EBV+ cells were transduced with vector, WT POLN lentivirus or mutant POLN lentivirus harbouring R303Q, P577L, and F545C mutations. (g-h) Cells were treated with CHX for 0 h, 2 h, 4 h and 8 h, and the protein was collected. (g) The changes of HA (POLN) protein were analyzed by western blot. (h) Quantification of CHX chase assays. Image J was used to analyze the grayscale of each band. Dotted line indicated 50% of normalized HA (POLN) protein. (i) In western blot assay, endogenous and exogenous POLN expressions were detected using POLN or HA antibodies, respectively.
to NPC susceptibility. Our results provide experimental evidence of the interaction between host genetics and EBV infection in NPC carcinogenesis, highlighting the importance of DNA repair factors in EBV lytic replication and carcinogenesis. However, this study had several limitations. First, whether POLN mutations are linked to NPC in other populations remains unknown. Our analyses were restricted to participants of Han Chinese ancestry, which limits the generalisability of our findings to diverse populations. Second, we did not have direct evidence to support POLN as a bona fide polymerase for EBV, as in vitro viral DNA replication assays were lacking; alternatively, whether POLN participates in EBV lytic activation in an enzymatic activity-independent manner was unclear, as none of the mutation sites were located in DNA-binding or catalytic sites of POLN but greatly affected POLN protein stability.

Figure 5. Mutations impaired the function of POLN in EBV lytic replication and cell proliferation. (a-e) HK1 EBV+ cells were transduced with vector, WT POLN lentivirus and mutant POLN lentivirus harbouring R303Q, P577L, and F545C mutations, 24 h after induction of viral lytic replication by TPA/NaB, cells were harvested. (a) The effect of POLN mutations on the proportion of GFP-positive cells. (b) The effect of POLN mutations on the intracellular EBV genome copy numbers. (c) The effect of POLN mutations on the extracellular EBV genome copy numbers. (d) The effect of POLN mutations on the mRNA expression level of EBV IE genes BZLF1 and BRLF1. (e) The effect of POLN mutations on the mRNA expression level of EBV true late genes BKRF4, BOLF1, and BLLF1. (f) CCK8 assay was used to detect the proliferation of HK1 EBV+ cells expressing vector, WT POLN, and mutant POLN. Data were presented as mean ± SD (N = 4), ***p < 0.001, **p < 0.01, *p < 0.05, ns: not significant (Student’s t-test). Black asterisk: compared with vector, red asterisk: compared with WT POLN.
EBV-associated cancers, such as EBV-positive gastric cancer and EBV-related lymphomas like Burkitt’s lymphoma and NK/T cell lymphoma, were currently unknown. Finally, to clarify whether somatic mutations in POLN affect NPC development through the “two-hit” model, paired blood and NPC tissue samples are needed to analyse the impact of POLN somatic mutations on gene expression and tumour characteristics.

In summary, our study identified a susceptibility gene POLN for familial NPC, and highlighted its function in regulating EBV lytic replication. Our study supports the assumption that DNA repair factors play an important role in NPC development and that the interaction between the etiological factors of NPC is important for tumourigenesis.

Contributors
Jia W.H. conceived and supervised the study. Feng L. co-supervised, designed functional experiments and interpreted the results. Xiao R.W. and Wang F. performed key laboratory assays, including functional experiments, Sanger sequencing, TaqMan genotyping, and data analysis. Wang T.M. worked on sequencing data, bioinformatic and statistics analysis. Zhang J.B. coordinated part workflow. Zhang J.B., Wu Z.Y., Deng C.M., Liao Y. and Dong S.Q. constructed DNA libraries for WES, genotyping, Sanger sequencing. Wu Z.Y., Deng C.M., Liao Y., Zou T., Yang D.W., He Y. Q., Xue W.Q., Liu Y.Y. and Xia Y. F. recruited and collected blood samples of NPC family members, collected the clinical information of participants. Zheng X.H., Li X.Z., Zhang P.F., Zhang S.D. and Hu Y.Z. extracted DNA. Mu J.B. gave important consultation, critical revision and manuscript edition. Gao S. consulted for protein structure prediction in silico. Jia W.H. and Xiao R. W. wrote the manuscript. Jia W.H. and Feng L. revised the manuscript with the assist of Wang F., Xiao R.W. and Wang T.M. Jia W.H., Feng L., Xiao R.W. and Wang F. directly accessed and verified the underlying data reported in the manuscript. All authors read and approved the final version of the manuscript.

Data sharing statement
The datasets used and/or analyzed during the current study are available from the corresponding author (Prof. Wei-Hua Jia: jiawh@sysucc.org.cn) on reasonable request. The key raw data of this research have been deposited in the Research Data Deposit public platform (www.researchdata.org.cn, accession number: RDDB2022136179).

Declaration of interests
No conflicts of interest exist for any of the authors.

Acknowledgements
This work was supported by the National Key Research and Development Program of China.
54 Qin HD, Shugart YY, Bei JX, et al. Comprehensive pathway-based association study of DNA repair gene variants and the risk of nasopharyngeal carcinoma in a population from an endemic area in South China. *Cell Res* 2015;25(11):1273–1280.

55 Jin S, Li R, Chen MY, et al. Single-cell transcriptomic analysis defines the interplay between tumor cells, viral infection, and the microenvironment in nasopharyngeal carcinoma. *J Virol* 2016;90(21):10807–10819.

56 Chen JH, Deng Z, Norseen J, Lieberman PM. Regulation of the translesion polymerase eta recruitment to DNA damage sites. *J Virol* 2017;91(19):e00600-17.

57 Zhou J, Liu B, Zhang L, et al. The protective role of polymorphism of the xeroderma pigmentosum group C gene in nasopharyngeal carcinoma. *Nat Rev Microbiol* 2019;17(11):691–700.

58 Yang ZH, Li YB, Liu SY, et al. Genetic polymorphisms of the DNA repair gene XRCC1 and risk of nasopharyngeal carcinoma. *Cancer Lett* 2017;400:20–29.

59 Yang ZH, Du B, Wei YS, et al. Genetic polymorphisms of the DNA repair gene XRCC1 and risk of nasopharyngeal carcinoma. *Cancer Lett* 2017;400:20–29.

60 Ban EZ, Lye MS, Chong PP, Yap YY, Lim SYC, Abdul Rahman H. Contribution of double strand break repair gene XRCC3 genotypes to nasopharyngeal carcinoma. *PLoS One* 2014;9(11):e0112188.

61 Yang ZH, Du B, Wei YS, et al. Genetic polymorphisms of the DNA repair gene XRCC1 and risk of nasopharyngeal carcinoma. *Cancer Lett* 2017;400:20–29.

62 Yang ZH, Liang WB, Jia J, Wei YS, Zhou B, Zhang L. The xeroderma pigmentosum group C gene polymorphisms and genetic susceptibility of nasopharyngeal carcinoma. *Ann Oncol* 2008;19(5):779–784.

63 Yang ZH, Du B, Wei YS, et al. Genetic polymorphisms of the DNA repair gene XRCC1 and risk of nasopharyngeal carcinoma. *PLoS One* 2014;9(11):e0112188.

64 Yang ZH, Du B, Wei YS, et al. Genetic polymorphisms of the DNA repair gene XRCC1 and risk of nasopharyngeal carcinoma. *PLoS One* 2014;9(11):e0112188.

65 Yang ZH, Du B, Wei YS, et al. Genetic polymorphisms of the DNA repair gene XRCC1 and risk of nasopharyngeal carcinoma. *PLoS One* 2014;9(11):e0112188.

66 Yang ZH, Du B, Wei YS, et al. Genetic polymorphisms of the DNA repair gene XRCC1 and risk of nasopharyngeal carcinoma. *PLoS One* 2014;9(11):e0112188.