BRCA2 loss-of-function germline mutations are associated with esophageal squamous cell carcinoma risk in Chinese

Josephine Mun-Yee Ko1,2*, Lvwen Ning1,2*, Xue-Ke Zhao3, Annie Wai Yeeng Chai1, Lisa Chan Lei1, Sheyne Sta Ana Choi1, Lihua Tao1, Simon Law3, Ava Kwong3,4, Nikki Pui-Yue Lee1, Kin-Tak Chan5, Anthony Lo5, Xin Song2, Pei-Nan Chen2, Yun-Li Chang6, Li Dong Wang2 and Maria Li Lung2,3

1Department of Clinical Oncology, University of Hong Kong, Hong Kong, People’s Republic of China
2Henan Key Laboratory for Esophageal Cancer Research of the First Affiliated Hospital, State Key Laboratory for Esophageal Cancer Prevention & Treatment, Zhengzhou University, Zhengzhou, Henan, People’s Republic of China
3Department of Surgery, University of Hong Kong, Hong Kong, People’s Republic of China
4Hereditary Breast Cancer Family Registry Cancer Genetics Center, Hong Kong Sanatorium and Hospital, Hong Kong, People’s Republic of China
5Division of Anatomical Pathology, Queen Mary Hospital, Hong Kong, People’s Republic of China
6Department of Gastroenterology, Shanghai University of Medicine and Health Sciences, Affiliated Zhoused University, Shanghai, People’s Republic of China

Esophageal squamous cell carcinoma (ESCC) occurs with highest frequency in China with over 90% mortality, highlighting the need for early detection and improved treatment strategies. We aimed to identify ESCC cancer predisposition gene(s). Our study included 4,517 individuals. The discovery phase using whole-exome sequencing (WES) included 186 familial ESCC patients from high-risk China. Targeted gene sequencing validation of 598 genes included 3,289 Henan and 1,228 moderate-risk Hong Kong Chinese. A WES approach identified BRCA2 loss-of-function (LOF) mutations in 3.23% (6/186) familial ESCC patients compared to 0.21% (9/4300) in the ExAC East Asians (odds ratio [OR] = 15.89, p = 2.48 × 10−19). BRCA2 LOF mutation frequency in the combined Henan cohort has significantly higher prevalence (OR = 10.55, p = 0.0039). Results were independently validated in an ESCC Hong Kong cohort (OR = 10.64, p = 0.022). One Hong Kong pedigree was identified to carry a BRCA2 LOF mutation. BRCA2 inactivation in ESCC was via germline LOF mutations and wild-type somatic allelic loss via loss of heterozygosity. Gene-based association analysis, including LOF mutations and rare deleterious missense variants defined with combined annotation dependent depletion score ≥30, confirmed the genetic predisposition role of BRCA2 (OR = 9.50, p = 3.44 × 10−8), and provided new evidence for potential association of ESCC risk with DNA repair genes (POLQ and MSH2), inflammation (TTC39B) and angiogenesis (KDR). Our findings are the first to provide compelling evidence of the

*J.M.-Y.K., L.N. and X.-K.Z. are co-first authors

Additional Supporting Information may be found in the online version of this article.

Key words: esophageal susceptibility gene, synthetic lethality, Chinese, BRCA2, loss-of-function mutations

Abbreviations: CADD: combined annotation dependent depletion; CPG: cancer predisposition gene; EC: esophageal cancer; ESCC: esophageal squamous cell carcinoma; KDR: kinase insert domain receptor; LOF: loss-of-function; LOH: loss-of-heterozygosity; MSH2: MutS Homolog 2; OR: odds ratio; POLQ: polymerase Q; RDV: rare deleterious variant; SCC: squamous cell carcinoma; TTC39B: tetratricopeptide repeat domain 39B; WT: wild type

Conflict of interest: None of the authors have any potential conflict to disclose (financial, professional or personal) relevant to the manuscript.

Grant sponsor: Asian Fund for Cancer Research Limited; Grant sponsor: Collaborative Research Fund; Grant number: C7031-15G; Grant sponsor: Doctoral Team Foundation of the First Affiliated Hospital of Zhengzhou University; Grant number: 2016-BSTDJ-03; Grant sponsor: General Research Fund; Grant number: 17118517; Grant sponsor: National Key R&D Program "Precision Medicine" of China; Grant number: 2016YFC0901403; Grant sponsor: National Natural Science Foundation of China; Grant numbers: 81872032, U1804262; Grant sponsor: Seed Fund for Basic Research from the University of Hong Kong; Grant number: 201611159317

DOI: 10.1002/ijc.32619

History: Received 26 Apr 2019; Accepted 24 Jul 2019; Online 9 Aug 2019

Correspondence to: Maria Li Lung, Department of Clinical Oncology, University of Hong Kong, Room L6-43, 6/F, Laboratory Block, Li Ka Shing Faculty of Medicine Building, 21 Sassoon Road, Pokfulam, Hong Kong, People’s Republic of China, Tel.: (852)-3917-9783, Fax: (852)-2816-6279, E-mail: mlilung@hku.hk; or Li Dong Wang, Henan Key Laboratory for Esophageal Cancer Research of the First Affiliated Hospital, State Key Laboratory for Esophageal Cancer Prevention & Treatment, Zhengzhou University, Zhengzhou, Henan, 450052, People’s Republic of China, Tel.: 86-371-67967677, Fax: 86-371-66658335, E-mail: ldwang2007@126.com

Int. J. Cancer: 00, 00–00 (2019) © 2019 UICC
role of \textit{BRCA2} in ESCC genetic susceptibility in Chinese, suggesting defective homologous recombination is an underlying cause in ESCC pathogenesis, which is amenable to therapeutic options based on synthetic lethality approaches such as targeting \textit{BRCA2} with PARP1 inhibitors in ESCC.

What’s new?
Esophageal cancer occurs at unusually high rates in Henan, China, and prognosis is poor. These authors set out to find genes associated with esophageal squamous cell carcinoma (ESCC) in this population. Based on whole exome sequencing, the researchers showed that loss of function mutations in \textit{BRCA2} increase ESCC risk. They also found evidence suggesting an association between ESCC and two other DNA repair genes, \textit{POLQ} and \textit{MSH2}. This is the first study of familial ESCC in a genetically-enriched cohort of patients from Henan. Screening for \textit{BRCA2} mutations may improve cancer detection and prognosis among patients with a family history of ESCC.

Introduction
Esophageal cancer (EC) is prevalent worldwide with 456,000 new cases annually, with about half occurring in Mainland China. EC is ranked sixth for cancer mortality with 400,000 deaths in 2012.\textsuperscript{1} Preventive and screening measures for earlier detection with biomarkers are needed for high-risk areas. In Mainland China and Hong Kong, the major histological EC subtype is squamous cell carcinoma (SCC), while adenocarcinoma is more common in Western countries. These two histological subtypes have distinct etiologies and pathogenesis, reflected by their different somatic mutational landscapes.\textsuperscript{2} EC exhibits striking geographical variations with more than 21-fold difference between high-risk and low-risk regions. The high-risk region of Northern China has an incidence >100/100,000; esophageal SCC (ESCC) is a top cause of cancer mortality with distinct etiologic risk factors and host genetic susceptibilities.\textsuperscript{4–6} Synergistic interaction of alcohol and smoking consumption and two functional variants in \textit{ALDH2} and \textit{ADH1B} polymorphisms increase ESCC risk; genome-wide association study in Japanese suggests alcohol/acetaldehyde metabolism plays an important etiologic role in the ESCC pathogenesis in moderate-risk regions.\textsuperscript{5} ESCC genome-wide association study identify multiple low-penetration moderate-risk common genetic variants for its development in Chinese.\textsuperscript{5,6} The extent of gene and environmental interactions responsible for the underlying molecular pathogenesis during ESCC development varies in different regions.\textsuperscript{1,3–6}

\textit{BRCA2} was first identified as a cancer predisposition gene (CPG) in breast cancer and later in ovarian, prostate and pancreatic cancers.\textsuperscript{7–9} \textit{BRCA2} is involved in homologous recombination to repair DNA double-strand breaks and acts downstream of the Fanconi anemia-BRCA pathway to repair DNA interstrand cross-links. Bi-allelic \textit{BRCA2} inactivation causes Fanconi anemia.\textsuperscript{10} Fanconi anemia patients are characterized by early development of cancers including ESCC.\textsuperscript{11} Although previous studies in Chinese and Turkmen suggested \textit{BRCA2} may play a role in the genetic susceptibility for familial ESCC, evidence for the association of \textit{BRCA2} loss-of-function (LOF) mutations remains unclear.\textsuperscript{12–14}

Before the advent of next-generation sequencing technological advancements, to screen for \textit{BRCA2} mutations was a challenging task due to its huge size; hence, earlier studies were limited by sample size and the molecular tools. Large comprehensive studies to identify high-penetration ESCC susceptibility genes are lacking, despite evidence for familial aggregation.\textsuperscript{15} Historical migrations from north-central to southern China suggest that inherited genetic factors contribute to ESCC development in high-risk regions.\textsuperscript{15–17} This current study is the first to utilize a genetically enriched cohort of patients with familial ESCC from high-risk Henan, located near the Tai-Hang Mountain region in Northern China, to identify high penetration ESCC predisposition genes using a comprehensive unbiased exome-sequencing strategy.

Materials and Methods
ESCC cases were confirmed by histopathology. Henan samples were collected from high-risk Linxian and Anyang counties from Linzhou Center Hospital and Yaocun Esophageal Cancer Hospital (2001 to 2014). Approval for use of human blood and/or information was obtained from the Committee for Ethical Review of Research Involving Human Subjects at Zhengzhou University. The study was conducted according to the Declaration of Helsinki principles. Blood samples of ESCC cases from moderate-risk Hong Kong (Queen Mary Hospital) were used for validation. These studies were approved by the HKU Institutional Review Board. Informed written consent was obtained from all participants. Table S1 summarizes the clinical information of all participants.

Study of familial ESCC: discovery by exome sequencing and two-step validation by target capture sequencing of Henan and Hong Kong ESCC

To identify possible CPGs, exome sequencing was used to analyze 186 familial ESCC patients, defined as families involving two generations with at least one first-degree relative developing ESCC in addition to the proband, in the discovery phase from high-risk Henan. Results were independently validated in Henan cohort (1,935 cases and 1,168 controls) and Hong Kong cohort (338 cases).
and 890 controls). The samples in the validation phase were all sequenced using a 598 gene-based capture design, except 42 Hong Kong ESCC cases and 890 Hong Kong non-cancer controls, which are from our previous genomic exome studies.\textsuperscript{18,19}

Library preparation and sequencing details are described in the supplementary methods. Briefly, genomic DNA was fragmented by sonication before library preparation using the KAPA HTP Library Prep Kit. Exome sequencing and targeted 598-gene capture utilized NimbleGen SeqCap EZ capture kits (Roche, Basel, Switzerland). Table S2 contains the gene list for target capture, which consisted of genes involved in DNA repair pathways (as previously identified by whole-exome sequencing with BRCA2 being the top candidate CPG); shortlisted cancer-related genes from familial ESCC studies with LOF variants; and genes of interest based on our previous functional and genomic studies with microarray differential profiling for genes critical for ESCC carcinogenesis.\textsuperscript{20–24}

**Bioinformatics analysis**

Single nucleotide variants and small insertion and deletions were called from the exome and target capture sequencing, as previously described following the genome analysis toolkit guidelines.\textsuperscript{18,25} Sanger sequencing validated 99.5% (182/183) germline variants. Pathogenic variants are identified according to the recommendation from American College of Medical Genetics and Genomics.\textsuperscript{26} Combined annotation dependent depletion (CADD) score\textsuperscript{27} was used to assess damage effects of variants of unknown significance. The rare variants of unknown significance with CADD score of \( \geq 30 \) (version1.4) are considered deleterious variants. In order to evaluate the deleterious effect of the BRCA2 missense substitutions, Align-GVGD\textsuperscript{28} was utilized to classify all rare missense substitutions into C0, C15, C25, C35, C45, C55, C65 categories with “prior probabilities of pathogenicity.”\textsuperscript{29} Analysis details are provided in supplementary materials and methods.

For the gene-based association analysis of genes with rare deleterious variants (RDVs), the Chi-square or Fisher’s exact test (two-tailed) under the dominant mode was used to examine differences in RDV frequencies (minor allele frequency <1%) between cases and controls on the 2X2 contingency table constructed using the number of cases with and without RDVs versus those in the controls. An individual with multiple RDVs was only counted once in the calculation. The \( p \)-value adjusted for Bonferroni correction, considering the 598 genes, was the cutoff (\( p < 8.36 \times 10^{-5} \)); otherwise a \( p < 0.05 \) was considered statistically significant.

![Figure 1](image-url). Lollipop schematic diagrams of (a) BRCA2 germline LOF mutations and deleterious missense variants (CADD score \( \geq 30 \)) in Chinese ESCC from Henan and Hong Kong, and (b) TTC39B, POLQ, MSH2 and KDR mutational spectra from Henan.
**BRCA2 loss of heterozygosity (LOH)**

Formalin-fixed paraffin-embedded tissue sections were stained with hematoxylin and eosin to evaluate carcinoma content for ESCC. BRCA2 mutation-positive individuals. DNA extraction using ARCTURUS® PicoPure® DNA Extraction Kit (Thermo Fisher Scientific, Waltham, MA, USA) was performed after manual macrodissection of cancer cells by pathologists (A.L. and L.H.T.) according to manufacturer’s instructions. PCR was performed with 5 μl extracted DNAs. Sanger sequencing was performed to score the LOH at mutated sites and to confirm germline LOF mutations and recurrent RDVs identified in BRCA2.

**Data availability**

Sequence data were deposited at the European Genome-phenome Archive (EGA), accession number EGAS00001003423. The data support the findings of our study are available from the corresponding author upon reasonable request.

**Results**

**Identified BRCA2 as top candidate cancer predisposition gene for ESCC risk**

In the exome analysis discovery phase, BRCA2 was ranked first in the top candidate CPG list for shortlist prioritization for validation, based on number of LOF mutations identified. Germline BRCA2 LOF mutations were detected with significantly higher prevalence (odds ratio [OR] = 15.89, 2.48 × 10⁻⁵) in familial ESCC Henan patients (6/186, 3.23%) compared to East Asians in the non-TCGA Exome Aggregation Consortium (ExAC) database (9/4300, 0.21%).

In the validation phase, target capture sequencing of BRCA2 identified 13 additional BRCA2 LOF mutations in 1935 Henan ESCC patients (13/1935, 0.67%) vs. 1/1168, 0.086%, p = 0.023, OR = 7.89) (Fig. 1a and Table S3). Combining Henan cohorts showed BRCA2 LOF mutations (19/2121, 0.90%) associated with high ESCC risk (OR = 10.55, p = 0.0035, Table 1) compared to controls (1/1168, 0.086%). The great effect (OR = 10.64, p = 0.022, Table 1) was independently validated in a second cohort in Hong Kong ESCC (4/338, 1.18%) and controls with no known cancers (1/890, 0.11%). All BRCA2 LOF mutations were validated by Sanger sequencing. Table 1 provides risk comparisons to other public databases. With a total combined Henan and Hong Kong sample size of 4,517 Chinese, this current study comprehensively sequenced 598 genes and is the first to report the significant higher prevalence of BRCA2 LOF mutations in Chinese ESCC patients compared to controls (p = 9.71, p = 6.80 × 10⁻⁸). Table 2 summarizes the clinical information of patients with BRCA2 LOF mutations.

The majority of the BRCA2 LOF mutations were unique and occurred beyond the RAD51-binding domain localized to the helical and DNA-binding domains (Fig. 1a, Fig. S1 and Table S3). The protein truncating mutations occurring >50 bp upstream of the final splice junction (mRNA coordinate approximately <9,600) are expected to cause nonsense-mediated decay, leading to truncated BRCA protein dysfunction. In the exome analysis discovery phase, BRCA2 was ranked first in the top candidate CPG list for shortlist prioritization for validation, based on number of LOF mutations identified. Germline BRCA2 LOF mutations were detected with significantly higher prevalence (odds ratio [OR] = 15.89, 2.48 × 10⁻⁵) in familial ESCC Henan patients (6/186, 3.23%) compared to East Asians in the non-TCGA Exome Aggregation Consortium (ExAC) database (9/4300, 0.21%).

In the validation phase, target capture sequencing of BRCA2 identified 13 additional BRCA2 LOF mutations in 1935 Henan ESCC patients (13/1935, 0.67%) vs. 1/1168, 0.086%, p = 0.023, OR = 7.89) (Fig. 1a and Table S3). Combining Henan cohorts showed BRCA2 LOF mutations (19/2121, 0.90%) associated with high ESCC risk (OR = 10.55, p = 0.0035, Table 1) compared to controls (1/1168, 0.086%). The great effect (OR = 10.64, p = 0.022, Table 1) was independently validated in a second cohort in Hong Kong ESCC (4/338, 1.18%) and controls with no known cancers (1/890, 0.11%). All BRCA2 LOF mutations were validated by Sanger sequencing. Table 1 provides risk comparisons to other public databases. With a total combined Henan and Hong Kong sample size of 4,517 Chinese, this current study comprehensively sequenced 598 genes and is the first to report the significant higher prevalence of BRCA2 LOF mutations in Chinese ESCC patients compared to controls (p = 9.71, p = 6.80 × 10⁻⁸). Table 2 summarizes the clinical information of patients with BRCA2 LOF mutations.

The majority of the BRCA2 LOF mutations were unique and occurred beyond the RAD51-binding domain localized to the helical and DNA-binding domains (Fig. 1a, Fig. S1 and Table S3). The protein truncating mutations occurring >50 bp upstream of the final splice junction (mRNA coordinate approximately <9,600) are expected to cause nonsense-mediated decay, leading to truncated BRCA protein dysfunction. BRCA2 was ranked first in the top candidate CPG list for shortlist prioritization for validation, based on number of LOF mutations identified. Germline BRCA2 LOF mutations were detected with significantly higher prevalence (odds ratio [OR] = 15.89, 2.48 × 10⁻⁵) in familial ESCC Henan patients (6/186, 3.23%) compared to East Asians in the non-TCGA Exome Aggregation Consortium (ExAC) database (9/4300, 0.21%).

In the validation phase, target capture sequencing of BRCA2 identified 13 additional BRCA2 LOF mutations in 1935 Henan ESCC patients (13/1935, 0.67%) vs. 1/1168, 0.086%, p = 0.023, OR = 7.89) (Fig. 1a and Table S3). Combining Henan cohorts showed BRCA2 LOF mutations (19/2121, 0.90%) associated with high ESCC risk (OR = 10.55, p = 0.0035, Table 1) compared to controls (1/1168, 0.086%). The great effect (OR = 10.64, p = 0.022, Table 1) was independently validated in a second cohort in Hong Kong ESCC (4/338, 1.18%) and controls with no known cancers (1/890, 0.11%). All BRCA2 LOF mutations were validated by Sanger sequencing. Table 1 provides risk comparisons to other public databases. With a total combined Henan and Hong Kong sample size of 4,517 Chinese, this current study comprehensively sequenced 598 genes and is the first to report the significant higher prevalence of BRCA2 LOF mutations in Chinese ESCC patients compared to controls (p = 9.71, p = 6.80 × 10⁻⁸). Table 2 summarizes the clinical information of patients with BRCA2 LOF mutations.

The majority of the BRCA2 LOF mutations were unique and occurred beyond the RAD51-binding domain localized to the helical and DNA-binding domains (Fig. 1a, Fig. S1 and Table S3). The protein truncating mutations occurring >50 bp upstream of the final splice junction (mRNA coordinate approximately <9,600) are expected to cause nonsense-mediated decay, leading to truncated BRCA protein dysfunction.
## Table 2. Clinical information of 24 Henan and HK ESCC patients with BRCA2 LOF variants

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Location of BRCA2 LOF variants</th>
<th>Gender</th>
<th>Age</th>
<th>Family history</th>
<th>No. of ESCC cases in family</th>
<th>No. of generations</th>
<th>Relationship of other affected family members with the index case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Henan ESCC patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHeFH0015</td>
<td>exon11:c.G6547T:p.E2183X</td>
<td>Male</td>
<td>51</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
<td>Mother</td>
</tr>
<tr>
<td>AHeFH0089</td>
<td>exon11:c.C6656G:p.S2219X</td>
<td>Male</td>
<td>57</td>
<td>Yes</td>
<td>3</td>
<td>2</td>
<td>Father and mother</td>
</tr>
<tr>
<td>AHeFH0167</td>
<td>exon23:c.9090dupA:p.T3030fs</td>
<td>Male</td>
<td>57</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
<td>Mother</td>
</tr>
<tr>
<td>AHeFH0187</td>
<td>exon20:c.8854_8854del:p.K2849fs</td>
<td>Female</td>
<td>57</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
<td>Mother</td>
</tr>
<tr>
<td>AHeFH0193</td>
<td>exon11:c.T2790G:p.Y930X</td>
<td>Male</td>
<td>55</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
<td>Mother</td>
</tr>
<tr>
<td>AHeFH0068</td>
<td>exon24:c.9247dupA:p.V3082fs</td>
<td>Male</td>
<td>54</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
<td>Father</td>
</tr>
<tr>
<td>AHeFH0567</td>
<td>Splicing</td>
<td>Male</td>
<td>55</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
<td>Father</td>
</tr>
<tr>
<td>AHeFH0692</td>
<td>exon8:c.657_658del:p.T219fs</td>
<td>Male</td>
<td>40</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
<td>Father</td>
</tr>
<tr>
<td>AHeSP122</td>
<td>exon24:c.9247dupA:p.V3082fs</td>
<td>Male</td>
<td>56</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>AHeSP357</td>
<td>exon11:c.C6193T:p.Q2065X</td>
<td>Male</td>
<td>51</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>AHeSP429</td>
<td>exon23:c.9090dupA:p.T3030fs</td>
<td>Male</td>
<td>50</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>AHeSP567</td>
<td>exon23:c.9090dupA:p.T3030fs</td>
<td>Female</td>
<td>45</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>AHeSP829</td>
<td>exon15:c.C7558T:p.R2522X</td>
<td>Male</td>
<td>59</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>AHeSP902</td>
<td>exon11:c.5067dupA:p.A1689fs</td>
<td>Female</td>
<td>68</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>HeSP123</td>
<td>exon25:c.C9328T:p.R3128X</td>
<td>Female</td>
<td>67</td>
<td>No</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FH-B5</td>
<td>exon27:c.A9976T:p.K3326X</td>
<td>Female</td>
<td>62</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
<td>Mother</td>
</tr>
<tr>
<td>AHeFH0402</td>
<td>exon27:c.10255dupT:p.I3418fs</td>
<td>Male</td>
<td>52</td>
<td>Yes</td>
<td>5</td>
<td>1</td>
<td>Elder brothers</td>
</tr>
<tr>
<td>Average ± SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55.0 ± 6.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HK ESCC patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,005,087</td>
<td>exon18:c.8065_8066del:p.C2689fs</td>
<td>Female</td>
<td>69</td>
<td>Unknown</td>
<td>–</td>
<td>–</td>
<td>Right and left breast and hepatocellular carcinoma</td>
</tr>
<tr>
<td>2,007,002</td>
<td>exon11:c.6100_6104del:p.R2034fs</td>
<td>Male</td>
<td>54</td>
<td>Unknown</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>2,008,097</td>
<td>exon11:c.2170delA:p.K724fs</td>
<td>Female</td>
<td>59</td>
<td>Unknown</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>2015124</td>
<td>exon10:c.1794_1798del:p.T598fs</td>
<td>Male</td>
<td>75</td>
<td>Yes</td>
<td>3</td>
<td>2</td>
<td>Father and sister</td>
</tr>
<tr>
<td>2011008</td>
<td>exon10:c.1794_1798del:p.T598fs</td>
<td>Female</td>
<td>79</td>
<td>Yes</td>
<td>3</td>
<td>2</td>
<td>Father and brother</td>
</tr>
<tr>
<td>Average ± SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64.3 ± 9.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.59, 9.59, 9.510.5</td>
</tr>
</tbody>
</table>

1Same family.
2Average age ± SD not including 2,011,008; including 2,011,008 is 67.2 ± 10; 2,011,008 is genotyped using Sanger sequencing.
while those further downstream are predicted to encode a truncated BRCA2 protein resulting in cytoplasmic BRCA2 due to the loss of the nuclear localization signals. Two recurrent BRCA2 LOF mutations, c.9090dupA:p.T3030fs identified in three and c.9247dupA:p.V3082fs in two patients, constituted 26.3% (5/19) of all LOF mutations detected in Henan ESCC cases.

The target sequencing approach enables comprehensive coverage of both exonic BRCA2 and its regulatory regions including the promoter together with the 5'/3' untranslated regions. Two rare recurrent SNPs, one was exonic and the other was located at 3' untranslated region, showed significantly higher prevalence in Henan ESCC. The first recurrent variant c.G7522A:p.G2508S with OR of 7.75 (p = 0.026) was detected in 0.66% ESCC patients (14/2,121) compared to 0.086% controls (1/1,168). G2508S, mapped to the BRCA2 helical domain, is predicted to be highly deleterious (CADD score = 31). For a Hong Kong ESCC patient carrying the G2508S variant, the matched formalin-fixed and paraffin-embedded primary tumor demonstrated selective partial LOH of the wild-type (WT) BRCA2 allele (Fig. S2a). The second recurrent SNP, rs56003538 (chr13:32973297 A>G), residing in the BRCA2 3' untranslated region, associated with a significantly higher ESCC risk with OR of 7.20 (p = 0.025) in 0.61% cases (13/2,121) compared to 0.086% controls (1/1,168). In the Henan cohort; 22.06% (139/630) BRCA2 variants are rare missense/non-frameshift small insertion and deletions/LOF variants and 5.4% (34/630) are located at the 5'/3' untranslated regions (Table S4). In addition to CADD score,

![Image](wileyonlinelibrary.com)

Figure 2. (a) Pedigree with BRCA2 c.1794_1798del:p.T598fs deletion from Hong Kong. (b) Germline and tumor chromatopherograms by Sanger sequencing confirm the presence of BRCA2 LOF mutation and the selective loss of the wild-type allele by LOH in the tumor of patient 2,015,124. [Color figure can be viewed at wileyonlinelibrary.com]
Table 3. Potential CPGs identified by gene-based association analysis

<table>
<thead>
<tr>
<th>Genes</th>
<th>Case (WT)</th>
<th>Case (RDV)</th>
<th>Control (WT)</th>
<th>Control (RDV)</th>
<th>ORs</th>
<th>p values</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRCA2</td>
<td>2,087</td>
<td>34</td>
<td>1,166</td>
<td>2</td>
<td>9.50</td>
<td>3.44 × 10⁻⁵</td>
</tr>
<tr>
<td>POLQ</td>
<td>2,091</td>
<td>28</td>
<td>1,163</td>
<td>5</td>
<td>3.11</td>
<td>0.016</td>
</tr>
<tr>
<td>TTC39B</td>
<td>2,110</td>
<td>11</td>
<td>1,168</td>
<td>0</td>
<td>Undefined</td>
<td>0.010</td>
</tr>
<tr>
<td>KDR</td>
<td>2,112</td>
<td>9</td>
<td>1,168</td>
<td>0</td>
<td>Undefined</td>
<td>0.031</td>
</tr>
<tr>
<td>MSH2</td>
<td>2,109</td>
<td>12</td>
<td>1,167</td>
<td>1</td>
<td>6.64</td>
<td>0.041</td>
</tr>
<tr>
<td>PALB2²</td>
<td>2,113</td>
<td>8</td>
<td>1,168</td>
<td>0</td>
<td>Undefined</td>
<td>0.057</td>
</tr>
</tbody>
</table>

¹Based on rare deleterious variants (RDVs) defined by variants of unknown significance with CADD score ≥30 and pathogenic variants.
²PALB2 showed a trend of association but did not reach statistical significance.

we also assessed the damaging effect of BRCA2 missense variants utilizing Align-GVGD scores (Table S5).²⁸ All cases and controls from Henan are grouped according the LOF and Align-GVGD score; if one individual carries multiple variants, then this individual will be classified into the group with high score (LOF is assumed to have the highest score). Table S5 summarizes the number of individuals in each group between cases and controls and the association result was calculated by logistic regression, adjusted for gender and age.²⁹ The deleterious missense variants (groups C55 and C65) are significantly associated with ESCC risk in Henan (p = 0.014, OR = 6.10). There is a linear trend of the log ORs of sequential C-scores (p = 0.02, Table S5).²⁹

One Hong Kong familial ESCC patient (2015124) carrying a BRCA2 germline LOF mutation had two additional first-degree family members in two generations having EC (Table 2 and Fig. 2). His sister (2011008) developed ESCC and carried the identical BRCA2 germline variant. Selective LOH of WT BRCA2 alleles was observed in formalin-fixed and paraffin-embedded samples from both. BRCA2 is first inactivated by rare germline deleterious LOF mutations. Further clinical relevance of BRCA2, as the earliest causal factor in ESCC tumorigenesis, was demonstrated by selective LOH of the WT allele, as the second mechanism for inactivation of WT BRCA2 allele, supporting the classical 2-hit tumor suppressor gene inactivation model. The frequency of somatic WT allele inactivation by selective LOH in matched formalin-fixed and paraffin-embedded tissues was significantly higher in Hong Kong ESCC patients genotyped with BRCA2 LOF mutations, compared to those with missense mutations (4/4, 100% vs. 4/18, 22.2%, p = 0.0096) (Figs. 2b and Figs. S2a and S2b). We conclude that in a subset of ESCC patients, the germline BRCA2 LOF mutations may be important drivers for ESCC genetic predisposition with germline allele being functionally inactivated by the protein truncation and losing the remaining WT copy by genomic deletion.

Potential cancer predisposition genes for ESCC risk by gene-based association analysis

The secondary analysis considered RDVs, including the LOF mutations and missense variants with CADD score ≥30, using a case-control association approach; the strongest association effect was detected for BRCA2 (OR = 9.50, p = 3.44 × 10⁻⁵), which is the only identified gene remaining significant after Bonferroni correction. At the gene level, Henan ESCC patients, 1.6%, (34/2121) harbored more BRCA2 RDVs compared to controls, 0.17%, (2/1168). Table S6 and Table 3 show the details of deleterious RDVs of BRCA2, TTC39B, POLQ, KDR, MSH2 and PALB2 from the gene-based association analysis. Figure 1b shows the schematic lollipop of TTC39B, POLQ, KDR and MSH2 for RDVs detected in Henan ESCC cases. Suggestive associations for four potential candidate genes: Tetra tripeptide repeat domain 39B (TTC39B) (OR = undefined, p = 0.10), Polymerase Q (POLQ) (OR = 3.11, p = 0.16), Kinase insert domain receptor (KDR) (OR = undefined, p = 0.031) and Mut S homolog (MSH2) (OR = 6.62, p = 0.041) were enriched with RDVs with nominal association with risk of ESCC in Henan Chinese (p < 0.05).

Discussion

ESCC is a deadly cancer with over 90% mortality, highlighting the need for improved prevention or early detection and treatment strategies. Despite previous studies suggesting inherited genetic components may be involved in ESCC development, the underlying genetic factors remain unclear.⁵,¹⁵,¹⁷ We applied a next-generation sequencing approach for a large sample size of Chinese ESCC patients (n = 2,459), aiming to identify the CPG(s) for ESCC susceptibility. Logistic regression analysis for common SNPs in the current study detected and validated the modest effect (rs2274223 and rs3765524, OR = 1.6) of the PLCE1 locus identified from our previous genome-wide association study (Table S7).⁵ The current study comprehensively sequenced BRCA2 to provide the first report of germline BRCA2 LOF mutation association with ESCC risk in Chinese. BRCA2 LOF mutations were detected in 3.23% (1/31) of 186 Henan familial ESCC in the discovery phase and in 80.0% (1:125) and 1.16% (1:86) of the larger ESCC populations in high-risk Henan and moderate-risk Hong Kong validation cohorts. The fact that Fanconi anemia patients may have early development of ESCC⁰,¹¹ supports our novel findings for the genetic susceptibility role of BRCA2 with ESCC risk. Although previous studies suggested BRCA2 may play a role in the genetic susceptibility for familial ESCC, evidence for associations of LOF mutations of BRCA2...
remained unclear and was both contradictory in different ethnic groups and limited by small sample sizes (case sample size <750).12–14 We now provide the first evidence of a germline BRCA2 LOF mutation (c.1794_1798del:p.T598fs) observed in a Hong Kong EC pedigree of three family members from two generations developing EC (Fig. 2). It is classified as pathogenic according to ClinVar (Table S3) and was present in the Hong Kong proband, who developed ESCC and bladder carcinoma, and his sister. In this family, another sibling developed lung carcinoma and died young. Genotyping of the mutation was also not possible in the father due to death or lack of consent for surviving siblings. Further follow-up of the younger family members in the next generation will be needed to demonstrate c.1794_1798del:p.T598fs cosegregation with ESCC.

BRCA2 family lineages with the same mutation have great phenotypic heterogeneity, resulting in the development of multi-cancer phenotypes.31,32 One 69-year old female Hong Kong ESCC patient with the c.8065_8066del variant also developed multiple cancers: right breast carcinoma at 36 years and left breast and hepatocellular carcinoma at 68 and 71 years, respectively. The variant is located at the oligonucleotide fold (OB1) in exon 18 and is predicted to encode a truncating variant being degraded by nonsense-mediated decay.22 Interesting patterns of phenotypic heterogeneity are observed for Caucasian BRCA1/2 families, but were unknown for Chinese or Asian families.51 In Chinese, data from the Hong Kong Hereditary Breast Cancer Family Registry recorded 104 of 3,002 families (3.46%) with probands having either breast and/or ovarian cancers, who also had a family history of EC.33,34 Twelve of these carried BRCA2 mutations and 8 families had a confirmed linkage between BRCA2 mutation with EC (5/8 had BRCA1 mutations and 3/8 had BRCA2 mutations, unpublished data). None of the BRCA2 LOF mutations in the current study were present in the existing Hong Kong Hereditary Breast Cancer Family Registry database. One limitation of the current study is the lack of details for history of other cancers for the Henan ESCC patients carrying the BRCA2 LOF mutations. Further follow-up studies in the BRCA2 carriers in these Chinese ESCC pedigrees will improve the understanding of the cancer penetrance and expressivity, that is, the types of cancers, of BRCA2.

The majority of the unique BRCA2 LOF mutations (13/19, 68.4%) observed in ESCC did not overlap with the breast cancer clustered regions and ovarian cancer clustered regions.51 Most LOF mutations (12/19, 63.2%) distributed beyond the BRCA2 repeats affect the helical domain and DNA binding domains (Fig. 1a and Fig. S1). BRCA2 not only plays a role in homologous recombination, but also prevention of R-loops, which result in replication stress and genomic instability leading to cancer.55 Further functional characterization of the observed BRCA2 LOF mutations regarding R-loop formation is expected to improve mechanistic understanding of the molecular etiology of ESCC.

The prevalence and spectrum of BRCA2 mutations vary substantially in different geographical regions and ethnicities, which result in phenotypic variations.36 The pathogenic role and biological significance of K3326X in BRCA2 remain unclear and controversial in cancer. It is generally considered benign for breast and ovarian cancers; our study also suggests that the K3326X is infrequent in Henan and Hong Kong ESCC patients (0.04%, 1/2459 Chinese ESCC patients), in line with a previous smaller Chinese ESCC study.53 In contrast, a deleterious effect of BRCA2 K3326X (rs11571833) with OR ranging from two to five was observed in familial pancreatic cancer,57 upper aerodigestive tract SCC,58 lung SCC59 and ESCC in Turkmen,12,14 predominantly being detected in cancers of squamous cell in origin, including oral cavity, larynx, esophagus and lung. The deleterious effect of K3326X in ESCC varied among different ethnic groups was present in European, Latin American and Indian populations, but rare in Asians, including the Japanese and Chinese populations.12,13,38 Further studies are required to study the variable role of K3326X on ESCC risk among different ethnic groups. The landscape of germline BRCA2 mutations varies among ESCC patients from different ethnic groups. Interestingly, the current study is the first to report the association of high ESCC risk (OR = 7.75) in Henan Chinese with the BRCA2 G2508S rare recurrent deleterious variant and the novel rare recurrent 3' untranslated region BRCA2 SNP, rs56003538 (OR = 7.20). The missense variant G2508S with CADD score 31 is scored 55 by Align-GVGD. It accounts for the majority (15/24, 62.5%) of ESCC patients classified in the pooled deleterious missense variants group (C55 and C65), which is significantly associated with ESCC risk (OR = 6.1, 24 cases vs. 2 controls, p = 0.014, Table S5). Interestingly, these rare missense substitutions in BRCA2 are explaining a similar fraction of ESCC as the protein truncating variants. The association of G2508S with ESCC risk should be interpreted with caution, as a functional study showed G2508S reduced activity with the homology-directed repair assay, but had limited impact on its activity with the poly(ADP ribose) polymerase inhibitor assay and ssDNA binding activity.60 BRCA2 G2508S, interestingly, is not only associated with ESCC in Henan Chinese with high ORs, but also with moderate risk in breast cancer in the Asian population.60 Further studies are required to assess the specific association of G2508S and rs56003538 with ESCC risk in Chinese or Asian ESCC populations.
differ much from the ESCC patients for non-BRCA2 carrier (64.4%). The delayed age of onset in Hong Kong ESCC patients carrying BRCA2 LOF mutations may be partly due to modifiable environmental protective risk factors, such as chronic nutritional deficiencies of folate, interacting differentially for ESCC development in geographically distinct regions. It is interesting to note that ESCC and pancreatic cancers are both gastrointestinal tract cancers that occur more frequently in males, have a later age of onset and lack consistent evidence for a hormonal basis compared to breast and ovarian cancers. Larger sample size studies are needed to substantiate these observations.

The gene-based association analysis confirmed that BRCA2 has the strongest association after correction for multiple test adjustment. Suggestive evidence for the potential association of two other DNA repair genes, MSH2 and POLQ, further reinforces the importance for maintenance of genomic stability. POLQ is involved in DNA damage response needed for translesion synthesis and alternative non-homologous end-joining (alt-NHEJ) for DNA double-strand break repair. The demonstration of the potential genetic predisposition role of POLQ in cancer is novel. Since both BRCA2 and POLQ are involved in DNA double-strand break repair, the findings suggest that defective homologous recombination may be an essential underlying causal genetic factor in ESCC pathogenesis. PALB2, a partner and colocalizer to BRCA2, complexes with BRCA2 for stabilization and facilitation of DNA repair. Despite a trend of higher frequency of Henan ESCC patients carrying PALB2 LOF variants (5/2121 vs. 0/1168, p = 0.17), this did not reach statistical significance. The gene-based association analysis showed a trend of higher frequency of eight RDVs (including five LOF and three missense variants with CADD score ≥30) identified in Henan ESCC cases, while none were detected in the controls (8/2121, 0.38% vs. 0/1168, 0%, p = 0.057, Table 3). A larger study with improved methods to define variant of unknown significance is required to provide further evidence for the contribution of PALB2 in ESCC risk. MSH2 functions in mismatch repair. Microsatellite instability study suggested microsatellite instability-low (30.6%) predominates over microsatellite instability-high (8.1%) in ESCC and that microsatellite instability may play a role in ESCC carcinogenesis. The nominal association of KDR and TTC39B with ESCC risk may provide novel insights and clues to suggest the involvement of angiogenesis and inflammation in ESCC development. KDR encodes vascular endothelial growth factor receptor 2 (VEGFR2) and implicates the importance of VEGF signaling. TTC39B, which is a high-density lipoprotein promoting liver X receptor (LXR) degradation, has anti-atherogenic, cholesterol removal and anti-inflammatory activities. Further studies are needed with larger cohorts to support for the associations of PALB2, POLQ, MSH2, TTC39B and KDR with ESCC risk. Functional studies are needed to elucidate the molecular mechanisms of the RDVs in these potential CPGs in ESCC pathogenesis.

Despite BRCA2 conferring great effects for familial ESCC in Henan, a substantial fraction of familial clustering and high incidence of ESCC cases still remains unexplained. The missing heredity component may be located at non-coding regulatory regions or other susceptibility loci may act in a gene-environmental interaction or gene-lifestyle dependent manner. Further next-generation sequencing analysis to identify ESCC susceptibility loci may incorporate the environmental and other etiologic risk factors to uncover additional genetic loci underlying ESCC carcinogenesis.

Although the ESCC population-attributable fraction remains modest, our data have an important potential implication on therapeutic options based on synthetic lethality approaches. Poly(ADP ribose) polymerase 1 inhibitors or BRCA2 synthetic lethal targets including RAD52, FEN1, APEX2 may warrant further consideration for those ESCC patients with BRCA2 deficiency. We expect the current novel findings to have beneficial impact on clinical practice for identification of some high-risk familial ESCC patients for earlier cancer detection by next-generation sequencing screening with BRCA2 germline mutations and management, as a useful prognosis marker for treatment for improved personalized care for ESCC patients.

Acknowledgements

The authors acknowledge the Centre for Genomic Sciences at the University of Hong Kong for technical assistance in establishment of hybridization-based target capture platform with NimbleGen SeqCap probes. This work was supported by the Asian Fund for Cancer Research Limited and Collaborative Research Fund (grant number C7031.15G) to M.L.L.; the Seed Fund for Basic Research from the University of Hong Kong (grant number 201611159317) and General Research Fund (grant number 17118517) to J.M.Y.K.; and the National Key R&D Program "Precision Medicine" of China (grant number 2016YFC0901403), Doctoral Team Foundation of the First Affiliated Hospital of Zhengzhou University (grant number 2016-BSTDJF-03) and the National Natural Science Foundation of China (grant number 81872032, 18042622) to L.D.W.

References


