Long non-coding RNA CASC8 polymorphisms are associated with the risk of esophageal cancer in a Chinese population

Yonghua Sang1, Haiyong Gu2, Yongbing Chen1, Yijun Shi3, Chao Liu3, Lu Lv3, Yifeng Sun2 & Yongsheng Zhang4

1 Department of Thoracic and Cardiovascular Surgery, The Second Affiliated Hospital of Soochow University, Suzhou, China
2 Department of Thoracic Surgery, Shanghai Chest Hospital, Shanghai Jiaotong University, Shanghai, China
3 Department of Cardiothoracic Surgery, Affiliated People’s Hospital of Jiangsu University, Zhenjiang, China
4 Department of Pathology, The Second Affiliated Hospital of Soochow University, Suzhou, China

Keywords
CASC8; esophageal cancer; lncRNA; polymorphisms.

Correspondence
Yifeng Sun, Department of Thoracic Surgery, Shanghai Chest Hospital, Shanghai Jiaotong University, Shanghai 200030, China. Tel: 18017321353 Fax: 021-22200268 Email: sunyifeng1121@163.com

Yongsheng Zhang, Department of Pathology, The Second Affiliated Hospital of Soochow University, Suzhou, China. Tel: 18962116930 Fax: 0512-67783457 Email: shengyongzh@126.com

Yonghua Sang and Haiyong Gu contributed equally to this work.

Received: 14 May 2020; Accepted: 22 July 2020.

doi: 10.1111/1759-7714.13612

Thoracic Cancer 11 (2020) 2852–2857

Introduction

Esophageal cancer, an important threat to the normal function of human digestion, ranks sixth among causes of cancer death.1, 2 There are various subtypes of esophageal cancer such as adenocarcinoma and squamous cell carcinoma.3, 4 The incidence and mortality of esophageal cancer ranks eighth and sixth in all malignant tumors in the world; approximately 456 000 new cases of esophageal cancer are diagnosed each year, and this malignancy has a severe effect on people’s lives and health.5 Despite the use of multimodal therapy, the incidence of esophageal cancer continues to increase, which is mostly attributed to its complex pathogenic mechanism.6 Most researchers believe that esophageal cancer is a multifactorial tumor. Biological factors, lifestyle, habits, and environmental factors may be related to the occurrence of esophageal cancer.7–9

Although the human genome has 3 billion DNA base pairs, only 1.5% of the genome contains coding DNA.10, 11 However, the remaining non-coding regions play vital regulatory roles. Most disease-related single nucleotide
polymorphisms (SNPs) occur in non-coding regions, and many of these SNPs are associated with cancer. Polymorphisms in many genes including long non-coding RNAs (lncRNAs) are closely related to tumorigenesis and tumor development. LncRNA is a type of non-coding RNA that does not participate in protein coding and has a sequence length of >200 nucleotides. Many studies have investigated the correlation between lncRNAs and the pathogenesis of various diseases, including tumors, and the results suggest that SNPs in lncRNA are associated with the susceptibility to cancer.

Cancer susceptibility candidate 8 (CASC8) is an lncRNA with no protein-coding potential that is located in the 8q24 region. LncRNAs originating from the 8q24 region including CASC8 play a critical role in the regulation of MYC, which is important for the development of multiple tumors, and the expression of CASC8 is regulated by long-range interaction of the MYC enhancer with the CASC8 promoter. SNPs in the CASC8 gene, such as rs7837328, rs6983267, and rs7014346, are correlated with the risk of cancer, including prostate, breast, colorectal, and gastric cancers. However, the effect of CASC8 SNPs on esophageal squamous cell carcinoma (ESCC) remains unclear.

Therefore, we performed a hypothesis-driven study to assess the molecular mechanisms associated with functional CASC8 SNPs in ESCC.

### Table 1 Distribution of selected demographic variables and risk factors in ESCC cases and controls

| Variable                        | Cases (n = 949) | Controls (n = 1369) | P-value¹ |
|---------------------------------|----------------|---------------------|---------|
| Age (years) mean ± SD           | 62.56 (± 8.60) | 62.04 (± 9.09)      | 0.167   |
| Age (years)                     |                |                     |         |
| <63                             | 472            | 729                 | 0.096   |
| ≥63                             | 477            | 640                 |         |
| Sex                             |                |                     | 0.107   |
| Male                            | 659            | 907                 |         |
| Female                          | 290            | 462                 |         |
| Tobacco use                     |                |                     |         |
| Never                           | 547            | 947                 | <0.001  |
| Ever                            | 402            | 422                 |         |
| Alcohol use                     |                |                     | <0.001  |
| Never                           | 666            | 1073                |         |
| Ever                            | 283            | 296                 |         |

¹Two-sided χ² test and student’s t-test; Bold values are statistically significant (P < 0.05).

### Table 2 Primary information on CASC8 rs10505477 C > T and rs1562430 A > G polymorphisms

| Genotyped SNPs | CASC8 | CASC8 |
|----------------|-------|-------|
| Chromosome     | 8     | 8     |
| Gene official symbol | CASC8 | CASC8 |
| Function       | Intron variant | Intron variant |
| Chr Pos (GRCh38.p12) | 127 395 198 | 127 375 606 |
| Regulome DB score⁷ | 5 | 5 |
| TFBS¹          | —     | —     |
| Splicing (ESE or ESS) | — | — |
| miRNA (miRanda) | — | — |
| miRNA (Sanger)  | —     | —     |
| nsSNP          | —     | —     |
| MAF³ for East Asian in database (1000 Genomes) | 0.389 | 0.177 |
| MAF in our controls (n = 1369) | 0.410 | 0.163 |
| P-value for HWE⁴ | 0.265 | 0.019 |
| Test in our controls |         |         |
| Genotyping method | Hi-SNP | Hi-SNP |
| % Genotyping value | 97.37% | 97.80% |

¹TFBS, transcription factor binding site (http://snpinfo.niehs.nih.gov/snpinfo/snpfunc.html). ³MAF, minor allele frequency. ⁴HWE, Hardy-Weinberg equilibrium.

### Methods

#### Study subjects

The current study was confirmed and approved by the Review Board of Jiangsu University (Zhenjiang, China). Written informed consent was obtained from all participants. The study enrolled 949 patients from the Affiliated People’s Hospital of Jiangsu University and Affiliated Hospital of Jiangsu University (Zhenjiang, China) between October 2008 and June 2013. A total of 1369 normal controls were selected from the above hospitals during the study.
same time period and frequency-matched to the patients with respect to age (± 5 years) and sex. Information for each participant was collected through a questionnaire, including information on drinking, smoking, age, sex, and diet. Venous blood (2 mL) was collected from each participant for CASC8 genotyping.

**Polymorphism genotyping**

Genomic DNA samples were isolated using the QIAamp DNA Blood Mini Kit (Qiagen, Berlin, Germany) and the extracted DNA sample was amplified by polymerase chain reaction (PCR). Hi-SNP high-throughput genotyping methods were used to genotype the amplification products (Shanghai Biowing Applied Biotechnology CO. LTD, Shanghai, China).

**Statistical analysis**

Hardy-Weinberg equilibrium for each SNP in the control subjects was detected using the chi-squared test. Student’s t-tests and χ² tests were performed to detect differences in factors collected in the questionnaire and CASC8 rs10505477 C > T and rs1562430 A > G genotypes. The relationships between CASC8 rs10505477 C > T and rs1562430 A > G SNPs and risk of ESCC were examined by computing odds ratios (ORs) and 95% confidence intervals (CIs) using logistic regression analyses, including crude ORs and adjusted ORs, after adjusting for age, sex, smoking, and drinking status. All statistical analyses were performed using SAS 9.1.3 (SAS Institute, Cary, NC, USA).

**Results**

**Patient characteristics**

The distribution of the demographic characteristics of the 949 cases and 1369 normal controls is shown in Table 1. Statistical analysis showed no significant difference in age or sex between the two groups (P = 0.167 and P = 0.107). However, the cases group included a significantly higher number of smokers and drinkers (both P < 0.001), suggesting that smoking and drinking are related to the development of ESCC. Minor allele frequencies (MAFs) in the controls were similar to East Asian MAFs in the 1000 Genomes database for these SNPs (Table 2).

### Table 3 Logistic regression analyses of associations between CASC8 rs10505477 C > T and rs1562430 A > G polymorphisms and risk of ESCC

| Genotype                | Cases (n = 949) | Controls (n = 1369) | Crude OR (95% CI) | P-value | Adjusted OR† (95% CI) | P-value |
|-------------------------|----------------|---------------------|-------------------|---------|-----------------------|---------|
| CASC8 rs10505477 C > T  |                |                     |                   |         |                       |         |
| CC                      | 317 34.05      | 471 35.52           | 1.00              | 1.00    |                       |         |
| CT                      | 439 47.15      | 622 46.91           | 1.05 (0.87–1.27)  | 0.620   | 1.03 (0.85–1.24)       | 0.800   |
| TT                      | 175 18.80      | 233 17.57           | 1.12 (0.88–1.42)  | 0.375   | 1.10 (0.86–1.41)       | 0.437   |
| TT vs. CT vs. CC        |               |                     | 0.670             |         |                       |         |
| CC + CT                 | 756 81.20      | 1198 82.43          |                   |         |                       |         |
| TT                      | 175 18.80      | 233 17.57           | 1.06 (0.85–1.34)  | 0.597   | 1.07 (0.85–1.36)       | 0.550   |
| C allele                | 1234 81.20     | 1812 58.97          |                   |         |                       |         |
| T allele                | 636 18.80      | 846 41.03           |                   |         |                       |         |
| CASC8 rs1562430 A > G   |                |                     |                   |         |                       |         |
| AA                      | 645 69.65      | 927 69.12           | 1.00              | 1.00    |                       |         |
| AG                      | 250 26.99      | 390 29.08           | 0.92 (0.76–1.11)  | 0.393   | 0.91 (0.75–1.09)       | 0.304   |
| GG                      | 31 3.34        | 24 1.78             | 1.86 (1.08–3.19)  | 0.025   | 2.05 (1.18–3.55)       | 0.010   |
| GG vs. AG vs. AA        |               |                     | 0.042             |         |                       |         |
| AG + GG                 | 281 30.34      | 414 30.87           | 0.98 (0.81–1.17)  | 0.789   | 0.97 (0.81–1.16)       | 0.725   |
| AA+AG                   | 895 96.65      | 1317 98.21          |                   | 1.00    | 1.00                   |         |
| GG                      | 31 3.34        | 24 1.78             | 1.90 (1.11–3.26)  | 0.020   | 2.11 (1.22–3.64)       | 0.007   |
| A allele                | 1540 83.15     | 2244 83.66          |                   |         |                       |         |
| G allele                | 312 16.84      | 438 16.33           |                   |         |                       |         |

†Adjusted for age, sex, smoking status and alcohol consumption; bold values are statistically significant (P < 0.05).
Table 4 Stratified analyses between CASC8 rs1562430 A > G polymorphism and ESCC risk by sex, age, smoking status, and alcohol consumption

| Variable            | GASC8 rs1562430 A > G (case/control) | Adjusted OR (95% CI); P-value | Adjusted OR (95% CI); P-value | Adjusted OR (95% CI); P-value |
|---------------------|--------------------------------------|-------------------------------|-------------------------------|-------------------------------|
|                     | AA | AG | GG | AG + GG | AA | AG | GG | AG + GG | AA | AG | GG | AG + GG | GG vs. (AG + AA) |
| **Sex**             |    |    |    |         |    |    |    |         |    |    |    |         |                |
| Male                | 457| 161| 24 | 185/276 | 1.00| 0.80 (0.64–1.01); | 2.47 (1.27–4.81); | 0.88 (0.71–1.11); | 2.62 (1.35–5.10); | 0.005 |
| Female              | 188| 89 | 7 | 96/138  | 1.00| 1.16 (0.84–1.61); | 1.31 (0.48–3.58); | 1.17 (0.85–1.61); | 1.25 (0.46–3.41); | 0.660 |
| **Age**             |    |    |    |         |    |    |    |         |    |    |    |         |                |
| <63                 | 322| 117| 22 | 139/222 | 1.00| 0.84 (0.64–1.01); | 2.50 (1.26–4.95); | 0.95 (0.73–1.23); | 2.62 (1.33–5.17); | 0.005 |
| ≥63                 | 323| 133| 9 | 142/192 | 1.00| 0.97 (0.74–1.27); | 1.41 (0.55–3.61); | 0.99 (0.76–1.29); | 1.42 (0.56–3.63); | 0.460 |
| **Smoking status**  |    |    |    |         |    |    |    |         |    |    |    |         |                |
| Never               | 368| 143| 21 | 164/281 | 1.00| 0.94 (0.74–1.20); | 1.98 (1.05–3.73); | 1.01 (0.80–1.28); | 2.02 (1.07–3.78); | 0.029 |
| Ever                | 277| 107| 10 | 117/133 | 1.00| 0.85 (0.62–1.15); | 2.51 (0.78–8.10); | 0.90 (0.66–1.21); | 2.64 (0.82–8.45); | 0.104 |
| **Alcohol consumption** |    |    |    |         |    |    |    |         |    |    |    |         |                |
| Never               | 448| 179| 22 | 201/314 | 1.00| 0.99 (0.80–1.24); | 1.98 (1.07–3.66); | 1.05 (0.85–1.31); | 1.98 (1.07–3.66); | 0.029 |
| Ever                | 197| 71 | 9 | 80/100  | 1.00| 0.72 (0.50–1.03); | 2.73 (0.73–10.25); | 0.78 (0.54–1.11); | 3.01 (0.80–11.27); | 0.102 |

† The genotyping was successful in 926 (97.6%) ESCC cases, and 1341 (98.0%) controls for CASC8 rs1562430 A > G. ‡ Adjusted for age, sex, smoking status and alcohol consumption (besides stratified factors accordingly) in a logistic regression model.
subgroups (P = 0.042). In a recessive model using CASC8 rs1562430 AA/GG genotypes as the reference group, the GG homozygous genotype (GG vs. AA/GG: adjusted OR = 2.11, 95% CI: 1.22–3.64, P = 0.007) was significantly associated with increased risk of ESCC, whereas the AG/GG homozygous genotype (AG/GG vs. AA/GG: adjusted OR = 0.97, 95% CI: 0.81–1.16, P = 0.725) was not associated with ESCC risk. However, when the CASC8 rs1562430 AA homozygous genotype was used as the reference group, the GG genotype was significantly associated with increased risk of ESCC (GG vs. AA: adjusted OR = 2.05, 95% CI: 1.18–3.55, P = 0.010), whereas the AG genotype was not associated with ESCC risk (AG vs. AA: adjusted OR = 0.91, 95% CI: 0.75–1.09, P = 0.304). The CASC8 rs10505477 C>T SNP was not associated with ESCC risk (Table 3).

### Stratified analyses of associations between CASC8 polymorphisms and ESCC risk

Stratified analysis was performed to further assess the possible correlation between the CASC8 rs1562430 A>G SNP and ESCC risk in the recessive model (Table 4). The results showed that the CASC8 rs1562430 GG genotype was significantly associated with increased risk of ESCC among men (GG vs. AA: adjusted OR = 2.47, 95% CI: 1.27–4.81, P = 0.008), patients younger than 63 years (GG vs. AA: adjusted OR = 2.50, 95% CI: 1.26–4.95, P = 0.009), non-smokers (GG vs. AA: adjusted OR = 1.98, 95% CI: 1.05–3.73, P = 0.034), and nondrinkers (GG vs. AA: adjusted OR = 1.98, 95% CI: 1.07–3.66, P = 0.031) (Table 4). In addition, in a recessive model using CASC8 rs1562430 AA/GG genotypes as the reference group, the GG homozygous genotype was significantly associated with increased risk of ESCC among men (GG vs. AA/GG: adjusted OR = 2.62, 95% CI: 1.35–5.10, P = 0.005), patients younger than 63 years (GG vs. AA/GG: adjusted OR = 2.62, 95% CI: 1.33–5.17, P = 0.005), non-smokers (GG vs. AA/GG: adjusted OR = 2.02, 95% CI: 1.07–3.78, P = 0.029), and nondrinkers (GG vs. AA/GG: adjusted OR = 1.98, 95% CI: 1.07–3.66, P = 0.029) (Table 4).

### Discussion

There are many possible causes of ESCC including both environmental and genetic factors. In the present study, we investigated the association between SNPs in the IncRNA CASC8 gene and susceptibility to ESCC. We found that rs1562430 was significantly associated with increased risk of ESCC. In stratification analyses, we found that the increased ESCC risk was significantly associated with CASC8 rs1562430 GG genotype among subjects for males, never-drinkers, never-smokers and those age < 60.

The CASC8 gene is located at 8q24 and has no translation capabilities; however, it can affect the progression of the disease by regulating the function of the coding region. CASC8 gene polymorphisms play important roles in different cancers. Furthermore, the IncRNA CASC8 suppresses the proliferation of bladder cancer cells by downregulating glycolysis. The results of this study suggested that the CASC8 SNP rs1562430 could be a predictive biomarker for susceptibility to ESCC.

### Acknowledgment

This manuscript was supported by the following funds: (i) Science and Technology Plan Project of Suzhou (SYS201725). (ii) Youth program of National Natural Science Foundation of China (81802194). (iii) “National tutorial training” program of Key health young talents of Suzhou.

### Disclosure

All authors declared no conflict of interest.

### References

1. Haas JF, Schottenfeld D. Epidemiology of esophageal cancer. In: Gastrointestinal Tract Cancer. USA: Springer 1978.
2 Arnal MJ, Arenas AF, Arbeloa AL. Esophageal cancer: Risk factors, screening and endoscopic treatment in Western and eastern countries. World J Gastroenterol 2015; 26: 15–25.

3 Kuwano H, Kato H, Miyazaki T et al. Genetic alterations in esophageal cancer. Surg Today 2005; 35 (1): 7–18.

4 Matsushima K, Isomoto H, Kohno S, Nakao K. MicroRNAs and esophageal squamous cell carcinoma. Digestion 2010; 82 (3): 138–44.

5 Ferlay J, Soerjomataram I, Dikshit R et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015; 136 (5): E359–86.

6 Fan YJ, Song X, Li J et al. Esophageal and gastric cardiac cancers on 4238 Chinese patients residing in municipal and rural regions: A histopathological comparison during 24-year period. World J Surg 2008; 32 (9): 1980–8.

7 Afridi SP, Khan A, Waheed I et al. High risk factors in patients with carcinoma esophagus. J Coil Physicians Surg Pak 2000; 10: 68–70.

8 Brown LM, Devesa SS. Epidemiologic trends in esophageal and gastric cancer in the United States. Surg Oncol Clin N Am 2002; 11 (2): 235–56.

9 Yang CX, Wang HY, Wang ZM et al. Risk factors for esophageal cancer: A case-control study in South-Western China. Asian Pac J Cancer Prev 2005; 6 (1): 48–53.

10 Sarah D, Carrie AD, Angelika M et al. Landscape of transcription in human cells. Nature 2012; 489 (7414): 101–8.

11 Hagen T. An integrated encyclopedia of DNA elements in the human genome. Nature 2012; 489 (7414): 57–74.

12 Duggan D, Freedman ML, Monteiro ANA et al. Principles for the post-GWAS functional characterization of cancer risk loci. Nat Genet 2011; 43 (6): 513–8.

13 Harrow J, Frankish A, Gonzalez JM et al. GENCODE: The reference human genome annotation for the ENCODE project. Genome Res 2012; 22 (9): 1760–74.

14 Chen G, Qiu C, Zhang Q, Liu B, Cui Q. Genome-wide analysis of human SNPs at long intergenic noncoding RNAs. Hum Mutat 2013; 34 (2): 338–44.

15 Schmitz SU, Grote P, Herrmann BG. Mechanisms of long noncoding RNA function in development and disease. Cell Mol Life Sci 2016; 73 (13): 2491–509.

16 Li L, Feng T, Lian Y et al. Role of human noncoding RNAs in the control of tumorigenesis. Proc Natl Acad Sci USA 2009; 106 (31): 12956–61.

17 Huarte M. The emerging role of IncRNAs in cancer. Nat Med 2015; 21 (11): 1253–61.

18 Bang-Shun H, Hui-Ling S, Tao X et al. Association of genetic polymorphisms in the IncRNAs with gastric cancer risk in a Chinese population. J Cancer 2017; 8 (4): 531–6.

19 Du M, Wang W, Jin H et al. The association analysis of IncRNA HOTAIR genetic variants and gastric cancer risk in a Chinese population. Oncotarget 2015; 6 (31): 31255–62.

20 Li Y, Jinf F, Ding Y, He Q, Zhong Y, Fan C. Long noncoding RNA, CCAT1, polymorphisms are associated with the risk of colorectal cancer. Cancer Genet 2018; 222–223: 13–9.

21 Ma G, Gu D, Lv C et al. Genetic variant in 8q24 is associated with prognosis for gastric cancer in a Chinese population. J Gastroenterol Hepatol 2015; 30 (4): 689–95.

22 Xiang J, Yang L, Chen L et al. The long noncoding RNA regulation at the MYC locus. Curr Opin Genet Dev 2015; 33: 41–8.

23 Fearnhead P, Yu K, Chatterjee N et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet 2007; 39 (5): 645–9.

24 Zanke BW, Greenwood CM, Rangrej J et al. Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nat Genet 2007; 39 (8): 989–94.

25 He J, Wilkens LR, Strom DO et al. Generalizability and epidemiologic characterization of eleven colorectal cancer GWAS hits in multiple populations. Cancer Epidemiol Biomarkers Prev 2011; 20 (1): 70–81.

26 Tomlinson I, Webb E, Carvajal-Carmona L et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat Genet 2007; 39 (8): 984–8.

27 Gu H, Ding G, Zhang W et al. Replication study of PLCE1 and C20orf54 polymorphism and risk of esophageal cancer in a Chinese population. Mol Biol Rep 2012; 39 (9): 9105–11.

28 He W, Wang B, Mu K et al. Association of single-nucleotide polymorphisms in the IL27 gene with autoimmune thyroid diseases. Endocr Connect 2019; 8 (3): 173–81.

29 Konrad H, Pitt JJ, Wahlberg BM et al. The 8q24 Gene Desert: An oasis of non-coding transcriptional activity. Front Genet 2012; 3: 69.

30 Hu R, Zhong P, Xiong L et al. Long noncoding RNA cancer susceptibility candidate 8’r suppresses the proliferation of bladder cancer cells via regulating glycolysis. DNA Cell Biol 2017; 36: 767–74.

31 Li L, Lv L, Liang Y et al. Association of 8q23-24 region (8q23.3 loci and 8q24.21 loci) with susceptibility to colorectal cancer: A systematic and updated meta-analysis. Int J Clin Exp Med 2015; 8 (11): 21001.

32 Zhou CP, Pan HZ, Li FX, Hu NY, Li M, Yang XX. Association analysis of colorectal cancer susceptibility variants with gastric cancer in a Chinese Han population. Genet Mol Res 2014; 13 (2): 3673–80.

33 Ghousaini M, Song H, Koessler T et al. Multiple loci with different cancer specificities within the 8q24 gene desert. J Natl Cancer Inst 2008; 100: 962–6.

34 Xuemin Z, Qin C, Chunya H et al. Polymorphisms on 8q24 are associated with lung cancer risk and survival in Han Chinese. PLOS One 2012; 7 (7): e41930.