OncoTargets and Therapy

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Association of polymorphisms in MALAT1 with the risk of esophageal squamous cell carcinoma in a Chinese population

This article was published in the following Dove Medical Press journal:
OncoTargets and Therapy

Yan Qu[1]
Na Shao[2]
Wenjing Yang[1]
Jianbo Wang[1,*]
Yufeng Cheng[1,*]

[1]Department of Radiation Oncology, Qilu Hospital of Shandong University, Jinan, Shandong 250012, People's Republic of China; [2]Department of Oncology, Shandong Provincial Hospital Affiliated to Shandong University, Jinan, Shandong 250021, People's Republic of China

*These authors contributed equally to this work

Objective: The main aim of this study was to investigate the association of polymorphisms in long non-coding RNA metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) with the risk of esophageal squamous cell carcinoma (ESCC) in a Chinese population.

Methods: A total of 245 ESCC patients and 490 gender- and age-matched cancer-free controls were genotyped for four tag single nucleotide polymorphisms (SNPs) of MALAT1 (rs3200401 C > T, rs1122709 C > G, rs664589 C > G, and rs619586 A > G). Statistical analyses including chi-squared test and logistic regression were performed to identify the association between the tag SNPs and risk of ESCC, and false discovery rate (FDR) <25% was applied to adjust for multiple comparisons.

Results: We found that rs3200401 C > T polymorphism of MALAT1 was significantly associated with increased risk of ESCC (CT vs CC: adjusted OR =1.59, 95% CI =1.07–2.35, *P*=0.021; TT vs CC: adjusted OR =2.27, 95% CI =1.04–4.96, *P*=0.039; dominant model [CT+TT vs CC]: adjusted OR =1.68, 95% CI =1.16–2.43, *P*=0.006). In the stratified analysis, rs3200401 TT and CT/TT genotypes were associated with increased risk of ESCC compared with CC genotype in subgroup of never drinking (TT vs CC: adjusted OR =2.34, 95% CI =1.02–5.34, *P*=0.044; CT/TT vs CC: adjusted OR =1.52, 95% CI =1.02–2.26, *P*=0.041). However, compared with AA genotype, MALAT1 rs619586 GG was associated with decreased risk of ESCC in ever drinking subgroup (GG vs AA: adjusted OR =0.38, 95% CI =0.15–0.99, *P*=0.049). The results remained significant after FDR adjustment (FDR value <0.25) except for the comparison between rs619586 GG and AA genotype in ever drinking subgroup.

Conclusion: Taken together, our findings proposed that polymorphism rs3200401 C > T in MALAT1 gene is associated with increased risk of ESCC. Since the association between rs619586 A > G polymorphism and ESCC risk was not significant after FDR adjustment, there was a minor possibility that rs619586 A > G might be a protective factor for ESCC.

Keywords: MALAT1, polymorphism, esophageal squamous cell carcinoma, single nucleotide polymorphism, susceptibility

Introduction

Esophageal cancer (EC) is the sixth leading cause of cancer-related death worldwide.[1] According to the latest epidemiological data, the incident cases and death tolls in China due to EC were found to be 477,900 and 375,000 each year, which ranked third and fourth among all the malignant tumors, respectively.[2] Although great advances have been achieved in diagnosis and treatment during the past decades, the mortality due to EC remains dismal, owing to the lack of early detection and effective treatment strategies for metastatic disease. As esophageal squamous cell carcinoma (ESCC) accounts

Correspondence: Jianbo Wang; Yufeng Cheng
Department of Radiation Oncology, Qilu Hospital of Shandong University, 107 West Wenhua Road, Jinan, 250012, People's Republic of China
Tel +86 185 6008 1329; +86 187 6978 5399
Email qlwjb2008@163.com; qlchengyf@163.com

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/OTT.S191155

for 90% of EC cases in China, increased efforts are required to predict, diagnose, and treat ESCC more effectively.

Long non-coding RNAs (lncRNAs) are a type of ncRNAs that are more than 200 nucleotides (nts) in length and are unable to encode proteins.[3] They have been found to play important roles in the progress of several common diseases (eg, myocardial infarction and hyperglycemia).[4–6] In malignant tumors, increasing number of lncRNAs were found to have important functions in the process of genesis, progress, and treatment response of tumor cells.[7–9]

Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), also known as nuclear-enriched transcript 2 (NEAT2), located on chromosome 11q13, is a long intergenic non-coding RNA (lincRNA) with >8,000 nts.[10] It was reported that MALAT1 is expressed abnormally and participated in the regulation of proliferation, migration, invasion, and apoptosis of many cancers,[11] such as non-small-cell lung cancer,[12] ESCC,[13] gastric cancer,[14] and hepatocellular carcinoma.[15]

Single nucleotide polymorphisms (SNPs) were well-recognized to directly regulate lncRNAs expression and alter the function of their corresponding lncRNAs.[16–18] Polymorphisms of MALAT1 had been reported to be associated with susceptibility to pulmonary arterial hypertension,[19] coronary atherosclerotic heart disease,[20] onset of congenital heart disease,[21] and colorectal cancer.[22] However, whether the polymorphisms of MALAT1 play any role in the risk of ESCC remains unknown. Considering that upregulation of lncRNA MALAT1 contributed to proliferation and metastasis in ESCC,[13] we performed a case-control study in the Chinese population to identify the association between MALAT1 polymorphisms and risk of ESCC.

Patients and methods

Study patients

A total of 245 ESCC patients diagnosed by pathology were consecutively recruited from Qilu Hospital of Shandong University from June to December 2017. Subjects diagnosed with more than one primary tumor were excluded from the study. Four hundred and ninety controls were selected from the physical examination center and matched to the ESCC patients by age (±5 years) and gender during the same time period. All participants were genetically unrelated. Data on basic characteristics (eg, age and gender) and environmental exposure (eg, smoking status and alcohol use) were obtained using face-to-face questionnaires and from patient records. Each participant donated 2 mL of venous blood willingly and provided informed consent to participate in the study. The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Qilu Hospital.

SNP selection and genotyping

SNPs in the MALAT1 genes were selected on the basis of the 1000 Genomes database (http://browser.1000genomes.org/Homo_sapiens/UserData/Haploview) and the software Haploview 4.2. Criteria of minor allele frequency (MAF) >0.05 was set to obtain the tag SNPs in the Chinese population. Ultimately, four candidate SNPs were chosen: rs3200401 C > T, rs1122709 C > G, rs664589 C > G, and rs619586 A > G. Genomic DNA was extracted from 2 mL of peripheral blood using a DNA Blood Mini Kit (#51104 QIAamp DNA Blood Mini Kit; Qiagen, NV, Venlo, the Netherlands) under the guidance of the manufacturer's protocol. Genotyping for the SNPs was done by using the TaqMan allelic discrimination assay on LightCycler 480 II system (Hoffman-La Roche Ltd., Basel, Switzerland). The primers used for amplification are shown in Table S1, which were self-designed by Primer Express 3.0 (Thermo Fisher Scientific, Waltham, MA, USA). Genotyping was performed blindly and the results were analyzed by using Sequencher Software (version 5.4.5).

Statistical analysis

The goodness-of-fit chi-squared test was used to test Hardy–Weinberg equilibrium (HWE) of SNPs among the control subjects. To determine the differences in the distribution of basic characteristics between the patients and controls, the selected exposure variables were evaluated using chi-squared tests. The association between MALAT1 polymorphisms and risk of ESCC was examined by computing the ORs and 95% CIs using logistic regression analysis and adjusting for smoking status and alcohol use. Besides presenting *P*-values, the false discovery rate (FDR) was also analyzed using the Benjamini and Hochberg method.[23] Since the main design of our study was to find the potential association between MALAT1 polymorphisms and risk of ESCC and considering the number of samples and SNPs that were included in our research, we set the FDR cutoff for noteworthy SNPs at 25% level. Statistical analysis was performed with SPSS 17.0 (SPSS Inc., Chicago, IL, USA).

Results

Basic characteristics

Locus information regarding MALAT1 polymorphisms and MAF in controls is listed in Table S1. Results showed that

the distribution of MALAT1 rs3200401 C > T, rs1122709 C > G, rs664589 C > G, and rs619586 A > G genotypes in controls agreed with HWE (P>0.05). Table S2 shows the basic characteristics of the 245 ESCC patients and 490 controls. Age and gender were not significantly different between the case and control groups (P=0.999, P=0.999), which indicated that these groups matched well with respect to these parameters. There was also no significant difference in the smoking status (P=0.439). However, the proportion of drinking was significantly higher in the case group compared with controls (P=0.012).

Association of MALAT1 rs3200401 C > T, rs1122709 C > G, rs664589 C > G, and rs619586 A > G polymorphisms with ESCC risk

The genotype frequencies of MALAT1 rs3200401 C > T, rs1122709 C > G, rs664589 C > G, and rs619586 A > G polymorphisms are listed in Table 1. After adjusting for smoking status and alcohol use, logistic regression analysis indicated that rs3200401 C > T significantly increased the risk of ESCC (CT vs CC: adjusted OR =1.59, 95% CI =1.07–2.35, P=0.021; TT vs CC: adjusted OR =2.27, 95% CI =1.04–4.96, P=0.039; dominant model [CT+TT vs CC]: OR =1.68, 95% CI =1.16–2.43, P=0.006). The significance remained even after FDR correction (FDR =0.168, 0.208, 0.096, respectively). However, rs1122709 C > G, rs664589 C > G, and rs619586 A > G polymorphisms were not statistically associated with ESCC risk (P>0.05).

Association of MALAT1 rs3200401 C > T, rs1122709 C > G, rs664589 C > G, and rs619586 A > G polymorphisms with ESCC risk in different stratification groups

As is shown in Table 2, stratification between rs3200401 and risk of ESCC according to smoking status and alcohol use was further conducted. In never drinking subgroup, after adjusting for smoking status, we found that rs3200401 CT/TT increased the risk of ESCC compared with CC genotype (TT vs CC: adjusted OR =2.34, 95% CI =1.02–5.34, P=0.044; CT/TT vs CC: adjusted OR =1.52, 95% CI =1.02–2.26, P=0.041 [Table 2]), which passed the correction of FDR (FDR =0.208, 0.208, respectively).

The genotype frequencies of MALAT1 rs619586 A > G polymorphism in the subgroup analysis are shown in Table 3. In ever drinking subgroup, after adjusting for smoking status, rs619586 GG genotype was associated with a decreased risk of ESCC (GG vs AA: adjusted OR =0.38, 95% CI =0.15–0.99, P=0.049 [Table 3]). However, the analysis did not pass the FDR correction (FDR =0.432).

In addition, after logistic regression analysis, rs1122709 G > C and rs664589 G > C polymorphisms were not associated with the risk of ESCC in any subgroup (Tables S3 and S4).

Discussion

The pathogenesis of ESCC is complex, and multiple factors may contribute to the etiology of ESCC. Our study indicated that MALAT1 polymorphisms were associated with a risk of ESCC in a Chinese population. We identified that rs3200401, a tag SNP of MALAT1, was a pivotal genetic risk factor against ESCC. Stratified analysis showed that in never drinking subgroup, rs3200401 TT and CT/TT genotype increased the risk of ESCC compared with CC genotype after adjusting for smoking status. All the results remained statistically significant after correction for FDR. In addition, in ever drinking subgroup, rs619586 GG genotype was associated with decreased risk of ESCC compared with AA, while the significance disappeared after adjusting for FDR. Further investigation was required to explore whether T allele of rs3200401 and G allele of rs619586 regulated MALAT1 expression in vivo. These results indicated that rs3200401 T allele and rs619586 G allele may be the key factors to influence the risk of ESCC.

MALAT1 was found to be overexpressed in numerous types of cancers including ESCC,[13] and MALAT1 showed significant effects on proliferation, migration, invasion, and apoptosis of cancer cells.[24] In clinical application, the upregulation of MALAT1 in several types of cancer tissues and its association with biological behavior of tumor cells made it a potential diagnostic or predictive biomarker.[6,12,25–27]

MALAT1 co-localizes with SC35 splicing domains, indicating that it may function in RNA metabolism.[28] MALAT1 was found to participate in mRNA processing, splicing, and exporting.[29] Except for RNA spicing, MALAT1 functioned by regulating the expression of many pivotal genes, including melanoma inhibitory activity 2, roundabout 1, glypican 6, latrophilin 2, CUB domain containing protein 1 ATP-binding cassette, subfamily A, and member 1 in lung cancer,[30] and caspase-3, caspase-8, Bcl-2, and Bcl-2-associated X protein in cervical cancer cells,[31] thus promoting the development of cancer. In addition, MALAT1 interacted with many other genes and was involved in several signaling pathways (ATM-CHK2 pathway,[13] PI3K-AKT pathway,[31,32] and

Table 1 Logistic regression analysis for associations between selected SNPs and risk of ESCC

Genotype	Cases	Controls	*P*-value[a]	Crude OR[a] (95% CI)	*P*-value[b]	Adjusted OR[b] (95% CI)
rs3200401						
CC	148	338		Ref		Ref
CT	79	133	0.017*	1.60 (1.09–2.35)	**0.021***[,a]	1.59 (1.07–2.35)
TT	18	19	0.026*	2.39 (1.11–5.14)	**0.039***[,b]	2.27 (1.04–4.96)
Dominant						
CC	148	338		Ref		Ref
CT+TT	97	152	0.004*	1.70 (1.18–2.45)	**0.006***[,c]	1.68 (1.16–2.43)
Recessive						
CC+CT	227	471		Ref		Ref
TT	18	19	0.059	2.06 (0.97–4.36)	0.084	1.96 (0.91–4.20)
rs1122709						
CC	123	269		Ref		Ref
CG	105	198	0.558	1.11 (0.78–1.57)	0.574	1.11 (0.78–1.57)
GG	17	23	0.071	2.07 (0.94–4.57)	0.092	2.01 (0.89–4.52)
Dominant						
CC	123	269		Ref		Ref
CG+GG	122	221	0.344	1.18 (0.84–1.65)	0.370	1.17 (0.83–1.65)
Recessive						
CC+CG	228	467		Ref		Ref
GG	17	23	0.086	0.51 (0.23–1.10)	0.110	0.52 (0.24–1.16)
rs664589						
CC	106	205		Ref		Ref
CG	117	236	0.351	1.37 (0.71–2.62)	0.466	1.28 (0.66–2.50)
GG	22	49	0.656	1.15 (0.62–2.16)	0.832	1.07 (0.56–2.04)
Dominant						
CC	106	205		Ref		Ref
CG+GG	139	285	0.310	1.21 (0.84–1.74)	0.325	1.21 (0.83–1.76)
Recessive						
CC+CG	223	441		Ref		Ref
GG	22	49	0.510	0.81 (0.44–1.50)	0.664	0.87 (0.47–1.63)
rs619586						
AA	85	177		Ref		Ref
AG	132	248	0.619	0.90 (0.58–1.38)	0.691	0.92 (0.59–1.41)
GG	28	65	0.799	0.93 (0.51–1.68)	0.968	1.01 (0.55–1.87)
Dominant						
AA	85	177		Ref		Ref
AG+GG	160	313	0.626	0.90 (0.60–1.36)	0.755	0.94 (0.62–1.42)
Recessive						
AA+AG	217	425		Ref		Ref
GG	28	65	0. 964	0.99 (0.58–1.69)	0.814	1.07 (0.62–1.85)

Notes: *P*-value[a] and crude OR[a]: conditional logistic regression. *P*-value[b] and adjusted OR[b]: conditional logistic regression with adjustment for smoking status and alcohol use. *Significant *P*-values ($P<0.05$). Significant *P*-values (FDR <0.25) are given in bold after being adjusted by FDR; FDR = *[a]0.168, *[,b]0.208, *[,c]0.096.
Abbreviations: ESCC, esophageal squamous cell carcinoma; SNP, single nucleotide polymorphism; FDR, false discovery rate.

extracellular-signal-regulated kinase/mitogen-activated protein kinase pathway[33]), thereby playing an important role in the progression of cancer cells. However, studies reporting on the underlying mechanisms are very few. Although our study found that rs3200401 C > T polymorphism was associated with increased risk of ESCC, it was also reported to be associated with better survival in advanced lung adenocarcinomas.[34] The opposite role of rs32000401 polymorphism

Table 2 Stratified analysis for associations between variant genotype of rs3200401 and ESCC risk

Variables	rs3200401 C > T (ESCC/control)				Adjusted OR (95% CI); *P*-value				
	CC	CT	TT	CT+TT	CC	CT	TT	CT+TT	TT vs (CT+CC)
Smoking									
Ever	46/98	26/37	4/3	30/40	1.00	1.69 (0.88–3.26) 0.116	2.79 (0.54–14.38) 0.220	1.78 (0.95–3.35) 0.073	2.36 (0.47–11.95) 0.298
Never	102/240	53/96	14/16	67/112	1.00	0.49 (0.23–1.03) 0.060	0.63 (0.29–1.40) 0.255	1.41 (0.96–2.06) 0.078	1.90 (0.90–3.99) 0.091
Drinking									
Ever	59/100	30/38	6/6	36/44	1.00	0.52 (0.16–1.75) 0.293	0.73 (0.21–2.58) 0.629	0.68 (0.39–1.20) 0.180	1.72 (0.52–5.67) 0.372
Never	89/238	49/95	12/13	61/10	1.00	1.40 (0.92–2.14) 0.122	2.34 (1.02–5.34) **0.044***,a	1.52 (1.02–2.26) **0.041***,b	2.10 (0.93–4.74) 0.074

Notes: *P*-value and adjusted OR: binary logistic regression with adjustment for smoking status and alcohol use (besides accordingly stratified factors). *Significant values ($P<0.05$). Significant *P*-values (FDR <0.25) are given in bold after being adjusted by FDR; FDR = *[a]0.208, *[b]0.208.
Abbreviations: ESCC, esophageal squamous cell carcinoma; FDR, false discovery rate.

in different cancers remains to be illuminated for detailed mechanism. It was reported that the interaction between lncRNAs and other molecules was probably associated with their structure, which might be altered by the polymorphisms of lncRNAs.[35] The rs3200401 C > T variant is located at the region M of MALAT1 (6,008–7,011 nts), one of the binding sites to SRSF2.[36] In addition, SRSF2 and phosphorylated SRSF2 were reported to be correlated with aggressive features of lung adenocarcinoma.[37] In a recent report, lncRNA SNP database was used to predict potential functions of rs3200401, suggesting that the C > T variation of rs3200401 causes 1.62 kcal/mol minimal free energy (ΔG) change and that it may alter the structural features of MALAT1, resulting in weakened interaction between MALAT1 and its binding protein SRSF2.[34] Besides, MALAT1 was reported to be associated with phosphorylation of SRSF2, interaction with SR proteins as a "molecular sponge," and regulation of the alternative splicing of pre-mRNAs.[28,38] Taken together, it was biologically possible that SNP rs3200401 C > T may alter the expression levels of cancer-associated genes, thus participating in the carcinogenesis and cancer development. A previous report suggested that rs619586 A > G was associated with decreased cancer risk in a meta-analysis study,[39] which was in line with our findings. In pulmonary arterial hypertension, MALAT1 with rs619586 G allele was reported to function as a competing endogenous RNA which exhibits more affinity toward miR-214, consequently inhibiting the vascular endothelial cell proliferation and migration in vitro by shortening S–M phase transition.[19] Nevertheless, more studies are required to clarify the mechanism as to how SNPs of MALAT1 functioned in the development of cancers.

Several limitations existed in this study. First, because all the participants were selected from a local hospital, there was a potential possibility of selection and information

Table 3 Stratified analysis for associations between variant genotype of rs619586 and ESCC risk

Variables	rs619586 A > G (ESCC/control)				Adjusted OR (95% CI); *P*-value				
	AA	AG	GG	AG+GG	AA	AG	GG	AG + GG	GG vs (AG+AA)
Smoking									
Ever	25/39	43/74	8/25	51/99	1.00	0.68 (0.34–1.34) 0.261	0.41 (0.15–1.12) 0.081	0.61 (0.32–1.18) 0.142	0.53 (0.22–1.30) 0.165
Never	60/138	89/174	20/40	109/214	1.00	1.18 (0.79–1.75) 0.415	1.16 (0.62–2.14) 0.645	1.18 (0.80–1.72) 0.407	1.05 (0.59–1.86) 0.864
Drinking									
Ever	31/42	56/78	8/24	64/102	1.00	0.84 (0.46–1.52) 0.560	0.38 (0.15–0.99) 0.049*	0.73 (0.41–1.31) 0.289	2.32 (0.98–5.52) 0.057
Never	54/135	76/170	20/41	96/211	1.00	1.12 (0.74–1.71) 0.585	1.30 (0.69–2.43) 0.419	1.16 (0.78–1.72) 0.478	1.21 (0.68–2.16) 0.516

Notes: *P*-value and adjusted OR: binary logistic regression with adjustment for smoking status and alcohol use (besides accordingly stratified factors). *Significant values ($P<0.05$). No significant *P*-values (FDR <0.25) are found after being adjusted by FDR.
Abbreviations: ESCC, esophageal squamous cell carcinoma; FDR, false discovery rate.

bias. Therefore, we used stratified analysis as compensation. Second, the sample size of our study was relatively small, which might have impacted the statistical power. Furthermore, for the stratification analysis, a larger sample size including more areas is required to fully validate the results.

Conclusion

Our findings suggest that MALAT1 rs3200401 C > T polymorphism is associated with an increased risk of ESCC in a Chinese population, and that rs3200401 polymorphism could be a novel and effective predictive biomarker for ESCC. Also, there is a minor possibility that rs619586 A > G polymorphisms may be associated with decreased risk of ESCC. Further studies in diverse ethnicities and functional analysis are required to confirm our findings.

Ethics approval and informed consent

All participants were informed about the purpose of the study and all signed the written informed consent. The study was approved by the Ethics Committee of Qilu Hospital of Shandong University and the study was conducted in accordance with the Declaration of Helsinki.

Acknowledgment

We appreciate all subjects who participated in this study. This work was supported by the National Natural Science Foundation of China (Nos 81602007, 81773228, and 81572958), Science Foundation of Shandong Province (No ZR2016HB68), Nanshan Project of Yantai for Shandong University (No 2014QLKY31), and Key Research and Development Program of Shandong Province (No 2017GSF18153).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin*. 2011;61(2):69–90. doi:10.3322/caac.20107
2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65(2):87–108. doi:10.3322/caac.21262
3. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell*. 2009;136(4):629–641. doi:10.1016/j.cell.2009.02.006
4. Uchida S, Dimmeler S. Long noncoding RNAs in cardiovascular diseases. *Circ Res*. 2015;116(4):737–750. doi:10.1161/CIRCRESAHA.116.302521
5. Puthanveetil P, Chen S, Feng B, Gautam A, Chakrabarti S. Long noncoding RNA MALAT1 regulates hyperglycaemia induced inflammatory process in the endothelial cells. *J Cell Mol Med*. 2015;19(6):1418–1425. doi:10.1111/jcmm.12576
6. Liu JY, Yao J, Li XM, et al. Pathogenic role of lncRNA-MALAT1 in endothelial cell dysfunction in diabetes mellitus. *Cell Death Dis*. 2014;5:e1506. doi:10.1038/cddis.2014.466
7. Winkle M, Kluiver JL, Diepstra A, van den Berg A. Emerging roles for long noncoding RNAs in B-cell development and malignancy. *Crit Rev Oncol Hematol*. 2017;120:77–85. doi:10.1016/j.critrevonc.2017.08.011
8. Liu XH, Sun M, Nie FQ, et al. Lnc RNA HOTAIR functions as a competing endogenous RNA to regulate HER2 expression by sponging miR-331-3p in gastric cancer. *Mol Cancer*. 2014;13:92. doi:10.1186/1476-4598-13-92
9. Merry CR, Forrest ME, Sabers JN, et al. DNMT1-associated long noncoding RNAs regulate global gene expression and DNA methylation in colon cancer. *Hum Mol Genet*. 2015;24(21):6240–6253. doi:10.1093/hmg/ddv343
10. Ji P, Diederichs S, Wang W, et al. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*. 2003;22(39):8031–8041. doi:10.1038/sj.onc.1206928
11. Zhao M, Wang S, Li Q, Ji Q, Guo P, Liu X. MALAT1: a long noncoding RNA highly associated with human cancers. *Oncol Lett*. 2018;16(1):19–26. doi:10.3892/ol.2018.8613
12. Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. *Nature*. 2012;489(7414):101–108. doi:10.1038/nature11233
13. Hu L, Wu Y, Tan D, et al. Up-regulation of long noncoding RNA MALAT1 contributes to proliferation and metastasis in esophageal squamous cell carcinoma. *J Exp Clin Cancer Res*. 2015;34:7. doi:10.1186/s13046-015-0123-z
14. Okugawa Y, Toiyama Y, Hur K, et al. Metastasis-associated long noncoding RNA drives gastric cancer development and promotes peritoneal metastasis. *Carcinogenesis*. 2014;35(12):2731–2739. doi:10.1093/carcin/bgu200
15. Lai MC, Yang Z, Zhou L, et al. Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation. *Med Oncol*. 2012;29(3):1810–1816. doi:10.1007/s12032-011-0004-z
16. Gong J, Tian J, Lou J, et al. A functional polymorphism in lncLAMC2-1:1 confers risk of colorectal cancer by affecting miRNA binding. *Carcinogenesis*. 2016;37(5):443–451. doi:10.1093/carcin/bgw024
17. Tang X, Gao Y, Yu L, et al. Correlations between lncRNA-SOX2OT polymorphism and susceptibility to breast cancer in a Chinese population. *Biomark Med*. 2017;11(3):277–284. doi:10.2217/bmm-2016-0238
18. Du M, Wang W, Jin H, et al. The association analysis of lncRNA HOTAIR genetic variants and gastric cancer risk in a Chinese population. *Oncotarget*. 2015;6(31):31255–31262. doi:10.18632/oncotarget.5158
19. Zhuo Y, Zeng Q, Zhang P, Li G, Xie Q, Cheng Y. Functional polymorphism of lncRNA MALAT1 contributes to pulmonary arterial hypertension susceptibility in Chinese people. *Clin Chem Lab Med*. 2017;55(1):38–46. doi:10.1515/cclm-2016-0056
20. Wang G, Li Y, Peng Y, Tang J, Li H. Association of polymorphisms in MALAT1 with risk of coronary atherosclerotic heart disease in a Chinese population. *Lipids Health Dis*. 2018;17(1):75. doi:10.1186/s12944-018-0728-2
21. Li Q, Zhu W, Zhang B, et al. The MALAT1 gene polymorphism and its relationship with the onset of congenital heart disease in Chinese. *Biosci Rep*. 2018;38(3):BSR20171381
22. Li Y, Bao C, Gu S, et al. Associations between novel genetic variants in the promoter region of MALAT1 and risk of colorectal cancer. *Oncotarget*. 2017;8(54):92604–92614. doi:10.18632/oncotarget.21507
23. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995;57(1):289–300.
24. Gutschner T, Hammerle M, Diederichs S. MALAT1 – a paradigm for long noncoding RNA function in cancer. *J Mol Med*. 2013;91(7):791–801. doi:10.1007/s00109-013-1028-y

25. Zheng HT, Shi DB, Wang YW, et al. High expression of lncRNA MALAT1 suggests a biomarker of poor prognosis in colorectal cancer. *Int J Clin Exp Pathol*. 2014;7(6):3174–3181.
26. Pang EJ, Yang R, Fu XB, Liu YF. Overexpression of long non-coding RNA MALAT1 is correlated with clinical progression and unfavorable prognosis in pancreatic cancer. *Tumour Biol*. 2015;36(4):2403–2407. doi:10.1007/s13277-014-2850-8
27. Glover AR, Zhao JT, Ip JC, et al. Long noncoding RNA profiles of adrenocortical cancer can be used to predict recurrence. *Endocr Relat Cancer*. 2015;22(1):99–109. doi:10.1530/ERC-14-0457
28. Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics*. 2007;8:39. doi:10.1186/1471-2164-8-109
29. Spector DL, Lamond AI. Nuclear speckles. *Cold Spring Harb Perspect Biol*. 2011;3(2). doi:10.1101/cshperspect.a000646
30. Gutschner T, Hammerle M, Eissmann M, et al. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res*. 2013;73(3):1180–1189. doi:10.1158/0008-5472.CAN-12-2850
31. Xu S, Sui S, Zhang J, et al. Downregulation of long noncoding RNA MALAT1 induces epithelial-to-mesenchymal transition via the PI3K-AKT pathway in breast cancer. *Int J Clin Exp Pathol*. 2015;8(5):4881–4891.
32. Tian Y, Zhang X, Hao Y, Fang Z, He Y. Potential roles of abnormally expressed long noncoding RNA UCA1 and Malat-1 in metastasis of melanoma. *Melanoma Res*. 2014;24(4):335–341. doi:10.1097/CMR.0000000000000080
33. Wu XS, Wang XA, Wu WG, et al. MALAT1 promotes the proliferation and metastasis of gallbladder cancer cells by activating the ERK/MAPK pathway. *Cancer Biol Ther*. 2014;15(6):806–814. doi:10.4161/cbt.28584
34. Wang JZ, Xiang JJ, Wu LG, et al. A genetic variant in long non-coding RNA MALAT1 associated with survival outcome among patients with advanced lung adenocarcinoma: a survival cohort analysis. *BMC Cancer*. 2017;17(1):167. doi:10.1186/s12885-017-3151-6
35. Hrdlickova B, de Almeida RC, Borek Z, Withoff S. Genetic variation in the non-coding genome: involvement of micro-RNAs and long non-coding RNAs in disease. *Biochim Biophys Acta*. 2014;1842(10):1910–1922. doi:10.1016/j.bbadis.2014.03.011
36. Miyagawa R, Tano K, Mizuno R, et al. Identification of cis- and trans-acting factors involved in the localization of MALAT-1 noncoding RNA to nuclear speckles. *RNA*. 2012;18(4):738–751. doi:10.1261/rna.028639.111
37. Gout S, Brambilla E, Boudria A, et al. Abnormal expression of the pre-mRNA splicing regulators SRSF1, SRSF2, SRPK1 and SRPK2 in non small cell lung carcinoma. *PLoS One*. 2012;7(10):e46539. doi:10.1371/journal.pone.0046539
38. Tripathi V, Ellis JD, Shen Z, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*. 2010;39(6):925–938. doi:10.1016/j.molcel.2010.08.011
39. Huang X, Zhang W, Shao Z. Association between long non-coding RNA polymorphisms and cancer risk: a meta-analysis. *Biosci Rep*. 2018;38(4). doi:10.1042/BSR20180365

Supplementary materials

Table S1 Primary information of selected SNPs for lncRNA MALAT1

SNP	Primer	Primer sequence	Base change	Location in the chromosome	MAF in our controls	HWE
rs3200401	Forward	GGGTGCCTG TGGGGTTTT	C > T	65,504,361	0.174	0.201
	Reverse	CCACCACCAGA AATGAACAAAA				
rs11227209	Forward	TTTTAGCAAC GCAGAAGCCC	C > G	65,497,960	0.249	0.075
	Reverse	TGTTATGCCTGG TTAGGTATGAGC				
rs664589	Forward	CTGTTGGCAC GAACACCTTC	C > G	65,501,878	0.341	0.178
	Reverse	TGGTACACCCA GTGGCTCATA				
rs619586	Forward	TGGGAGCAA GTCGCAGGA	A > G	65,498,698	0.386	0.132
	Reverse	GAGAACAACT CGCATCACCG				

Abbreviations: HWE, Hardy–Weinberg equilibrium; MAF, minor allele frequency; SNP, single nucleotide polymorphism; lncRNA, long non-coding RNA.

Table S2 Selected characteristics in ESCC cases and controls

Variables	Cases (N=245, n [%])		Controls (N=490, n [%])		*P*-value[a]
Age (years)	61.50 (39–79)		62.00 (39–79)		0.999
≥60	140	56.91%	280	57.14%	
<60	105	43.09%	210	42.86%	
Gender					0.999
Male	185	75.61%	370	75.51%	
Female	60	24.39%	120	24.49%	
Smoking					0.439
Ever	76	30.89%	138	28.16%	
Never	169	69.11%	352	71.84%	
Drinking					0.012*
Ever	95	38.62%	144	29.39%	
Never	150	61.38%	346	70.61%	

Notes: [a]Two-sided chi-squared test. *Significant values ($P<0.05$).
Abbreviation: ESCC, esophageal squamous cell carcinoma.

Table S3 Stratified analysis for associations between variant genotype of rs1122709 and ESCC risk

Variables	rs1122709 C > G (ESCC/control)				Adjusted OR (95% CI); *P*-value				
	CC	CG	GG	CG+GG	CC	CG	GG	CG+GG	GG vs (CG+CC)
Smoking									
Ever	37/85	32/49	7/4	39/53	1.00	1.30 (0.69–2.42) 0.417	3.33 (0.86–12.97) 0.083	1.45 (0.80–2.64) 0.223	2.99 (0.79–11.36) 0.108
Never	86/184	73/149	10/19	83/168	1.00	1.04 (0.71–1.52) 0.837	1.12 (0.50–2.50) 0.792	0.95 (0.66–1.38) 0.798	1.10 (0.50–2.41) 0.822
Drinking									
Ever	49/182	40/55	6/7	46/62	1.00	1.14 (0.66–1.99) 0.640	1.25 (0.39–4.06) 0.708	1.15 (0.68–1.97) 0.599	1.18 (0.37–3.74) 0.776
Never	74/187	65/143	11/16	76/158	1.00	1.09 (0.73–1.63) 0.678	1.62 (0.72–3.68) 0.247	1.14 (0.78–1.69) 0.500	0.64 (0.29–1.42) 0.275

Notes: *P*-value and adjusted OR: binary logistic regression with adjustment for smoking status and alcohol use (besides accordingly stratified factors). *Significant values ($P<0.05$).
Abbreviation: ESCC, esophageal squamous cell carcinoma.

Table S4 Stratified analysis for associations between variant genotype of rs664589 and ESCC risk

Variables	rs664589 C > G (ESCC/control)				Adjusted OR (95% CI); *P*-value				
	CC	CG	GG	CG+GG	CC	CG	GG	CG+GG	GG vs (CG+CC)
Smoking									
Ever	37/67	31/62	8/9	39/71	1.00	0.95 (0.51–1.78) 0.880	1.43 (0.48–4.27) 0.523	1.02 (0.56–1.85) 0.948	1.46 (0.51–4.20) 0.482
Never	69/138	86/174	14/40	100/214	1.00	1.42 (0.72–2.79) 0.309	1.41 (0.73–2.74) 0.306	0.94 (0.65–1.37) 0.752	0.71 (0.37–1.34) 0.287
Drinking									
Ever	39/57	47/70	9/17	56/87	1.00	1.16 (0.65–2.05) 0.620	0.86 (0.34–2.19) 0. 757	1.10 (0.63–1.90) 0.743	0.80 (0.33–1.91) 0.611
Never	67/148	70/166	13/32	83/198	1.00	0.91 (0.61–1.37) 0.661	0.85 (0.42–1.72) 0.642	0.90 (0.61–1.33) 0.604	0.87 (0.45–1.75) 0.727

Notes: *P*-value and adjusted OR: binary logistic regression with adjustment for smoking status and alcohol use (besides accordingly stratified factors). *Significant values ($P<0.05$).
Abbreviation: ESCC, esophageal squamous cell carcinoma.