International Journal of General Medicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

The Association of EGFL7 Polymorphism and Expression with Cervical Cancer Susceptibility and Pathogenesis

Weipeng Liu[1,2,*], Zhixin Niu[1,*], Zhiling Yan[3], Lei Shi[1], Shao Zhang[3], Chao Hong[1], Shuying Dai[4], Li Shi[1], Yufeng Yao[1]

[1]Institute of Medical Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Kunming, People's Republic of China; [2]Department of Orthopedics, The First People's Hospital of Yunnan Province, Kunming University of Science and Technology Affliated Hospital; the Key Laboratory of Digital Orthopedics of Yunnan Province; the Clinical Medicine Center of Spinal and Spinal Cord Disorders of Yunnan Province, Kunming, People's Republic of China; [3]Department of Gynaecologic Oncology, No.3 Affiliated Hospital of Kunming Medical University, Kunming, People's Republic of China; [4]Department of Pathogen Biology and Immunology Faculty of Basic Medical Science, Kunming Medical University, Kunming, People's Republic of China

*These authors contributed equally to this work

Correspondence: Li Shi; Yufeng Yao, Institute of Medical Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Kunming, People's Republic of China, Email shili.imb@gmail.com; leoyyf@gmail.com

Purpose: Previous studies have highlighted diverse roles for EGFL7 in various cancers, but its specific function in cervical cancer remains unclear. This study aimed to elucidate the role of EGFL7 in cervical cancer susceptibility and pathogenesis.

Methods: We genotyped three EGFL7 SNPs in 694 healthy controls, 408 cervical intraepithelial neoplasia (CIN) patients, and 934 cervical cancer (CC) patients using TaqMan probe-based real-time PCR. EGFL7 expression was measured in tumor tissues and cell lines via qRT-PCR. RNA sequencing was used to explore the biological functions of EGFL7 in cervical cancer.

Results: The G allele frequency of rs9411260 was significantly elevated in both the cervical cancer (CC) ($P = 0.006$) and cervical intraepithelial neoplasia (CIN) ($P = 0.016$) groups compared to the control group. EGFL7 mRNA expression was markedly down-regulated in cervical cancer tissues and HeLa/C33A cells compared to normal tissues and ECT1/E6E7 cells. Furthermore, modulation of EGFL7 expression was found to alter pathways associated with cell adhesion, migration, and ECM-receptor interaction.

Conclusion: EGFL7 polymorphisms and expression are associated with cervical cancer susceptibility. Dysregulated EGFL7 appears to contribute to cervical cancer progression by affecting cell adhesion and migration, offering new insights into its pathogenesis and potential therapeutic targets.

Keywords: EGFL7, SNPs, cervical cancer, cell adhesion, migration

Introduction

Cervical cancer ranks as the fourth most prevalent malignancy affecting women's health worldwide, in the annals of 2020, there were more than 600,000 new cases of cervical cancer and about 340,000 deaths worldwide.[1] Persistent infection of high-risk types of human papillomavirus (HPV) is the main risk factor of cervical cancer.[2] The majority of women infected with HPV can clear the virus, while a small minority endure persistent infection, culminating in cervical cancer.[3] Thus, it emerges that the host's genetic factors may cast a profound shadow over the susceptibility to cervical cancer.[4] Research indicates cervical cancer has substantial heritability, largely explained by common single nucleotide polymorphisms (SNPs).[5–7] Consequently, the study of SNPs in oncogenes and tumor suppressor genes is growing,[8] enhancing our understanding of the disease and providing new avenues for diagnosis and treatment.

Epidermal growth factor-like domain 7 (EGFL7) is a 41-kDa secreted protein that can be expressed in various cell types. EGFL7 was initially reported to be involved in homeostasis regulation during embryogenesis.[9] Subsequently,

https://doi.org/10.2147/IJGM.S563592
Received: 27 August 2025
Accepted: 19 December 2025
Published: 8 January 2026

studies found that EGFL7 plays a pivotal role in angiogenesis and vascular development.[10–12] Recent investigations have unveiled EGFL7's role in metastasis,[13] proliferation[14] and invasion[15] of tumor cells, positioning it as a significant player in malignancies such as acute myeloid leukemia[16] breast cancer,[17] hepatocellular carcinoma,[18] lung cancer,[19] and malignant pleural mesothelioma.[20] The high expression of EGFL7 is a marker indicating poor prognosis for various cancers. It can regulate the Notch and PI3K/AKT signaling pathways, thereby causing cell migration, invasion and angiogenesis in different types of cancer.[21,22] Nevertheless, the role of EGFL7 in human cervical cancer remains enigmatic. A study showed that expression of epidermal growth factor-like domain 7 may be a predictive marker of the effect of neoadjuvant chemotherapy for locally advanced uterine cervical cancer.[23] Our prior research established a significant association between the common SNP rs4636297 in the EGFL7 gene region and the incidence of cervical cancer.[24] Given that single nucleotide polymorphisms (SNPs) within gene regulatory and coding regions could confer risk of cervical cancer by regulating the expression of specific genes, we speculate SNPs within EGFL7's transcriptional and coding regions may play a role in modulating EGFL7 expression and contribute to cervical cancer susceptibility.

In this article, we investigated the association between EGFL7 polymorphisms and cervical cancer susceptibility in the Han Chinese population. Concomitantly, we also examined EGFL7 expression in cervical cancer tissues and cells (HeLa and C33A). Moreover, we further explored the potential biological function of EGFL7 in cervical cancer tumorigenesis.

Material and Methods

Subjects and Tissue Samples

In this investigation, the case group comprised 408 patients with CIN and 934 patients with cervical cancer, while the control group consisted of 694 healthy individuals undergoing hospital-based physical examinations during the same period. Diagnoses of CIN and cervical cancer adhered to the "Diagnosis and Treatment: Obstetrics and Gynaecology" guidelines and the International Federation of Gynaecology and Obstetrics (FIGO 2009) criteria at the Third Affiliated Hospital of Kunming Medical University, covering the period from July 2021 to May 2023, patients with other malignancies, chronic diseases, or those previously exposed to chemotherapy or radiotherapy were excluded. This study was approved by the Ethics Committee of the Third Affiliated Hospital of Kunming Medical University (approval no. KYCS202195), and all participants provided written informed consent.

A collection of 63 matched sets of primary cervical cancer tumors and adjacent normal tissues was procured from patients at the Third Affiliated Hospital of Kunming Medical University. All tissues underwent pathological examination and were promptly frozen at −80 °C.

SNP Selection

In this study, 1670 bp upstream of the EGFL7 transcription start site was chosen as the promoter region according to previous studies.[25,26] JASPAR (http://jaspar.genereg.net/) was used to predict whether the SNPs in the promoter region of EGFL7 are located in the transcription factor binding site and disrupt the binding of specific transcription factors.[27] Missense variants with a minor allele frequency (MAF) greater than 0.05 were called and filtered using the Ensembl Variant Effect Predictor (http://www.ensembl.org/vep).[28] Consequently, rs1332793, rs9411260, and rs2297538 were deemed the candidate SNPs for our present study.

DNA Extraction and Sequencing

Venous blood samples, amounting to 10 mL, were meticulously collected from the study participants. The genomic DNA was subsequently extracted using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany). We employed the Multiskan Skyhigh full-wavelength enzyme plate (ND-2000, Thermo Fisher Scientific) to assess the DNA's concentration and purity. Post-extraction, the genomic DNA was preserved within a −20 °C refrigerator for future utilization.

Three SNPs rs1332793, rs9411260 and rs2297538 were genotyped by TaqMan probe real-time fluorescence quantitative polymerase chain reaction (RTFQ-PCR). The probes and primers were designed and produced by ThermoFisher Scientific Company (Waltham, MA, USA), and TaqMan Genotyping Master Mix was purchased from ABI. PCR

amplification was carried out in 384-well reaction plates with 2.5 μL Master Mix, 0.125 μL primer and probe (FAM and VIC) mix, 1.375 μL ddH2O and 1 μL genomic DNA in each well. Amplification was conducted in a QuantStudio 6 Flex Fast Real-Time PCR system as follows: 95 °C preheat denaturing for 10 min, 92°C for 10s and 60 °C for 1 min, all iterated over 40 cycles. As a validation measure, we conducted Sanger sequencing for 20 randomly selected individuals, assuring the absence of genotyping errors.

Quantitative RT–PCR

Total RNA was isolated from tissues or cells using TRIzol reagent. The PrimeScriptTM RT reagent Kit with gDNA Eraser (TaKaRa Bio Inc, Tokyo, Japan) was used to synthesize cDNA. We reverse transcribed 1 μg of total RNA, and then diluted cDNA at a final concentration of 20 μg/μL. Quantitative real-time PCR was carried out under the following conditions to detect the expression levels of EGFL7 in cervical cancer tissues, corresponding adjacent tissues and cervical cancer cells: denaturation at 95 °C for 10 min, followed by 40 cycles comprising denaturation at 95 °C for 15 seconds, annealing at 60 °C for 15 seconds, and extension at 72 °C for 15 seconds. GAPDH was used as the internal control of EGFL7. All the primers are listed in Table S1.

Western Blotting

Cells were washed with phosphate buffered solution (PBS). Then cells were lysed in RIPA and complete Protease Inhibitor Cocktail (Roche) mixed lysis buffer for 30 min. Subsequently, this mixture was concentrated under 12,000g at 4 °C for 10 minutes. Protein concentration was quantified by BCA Protein Assay Kits (Thermo Scientific). Total 50 μg of protein was loaded into each lane and samples were subjected to 12% SDS-PAGE, followed by a 1-hour transfer onto PVDF membranes. These membranes were then blocked with 5% non-fat milk for 2 hours at room temperature and subsequently incubated with the anti-EGFL7 antibody (Abcam, cat# ab256451) and anti-beta actin antibody (Servicebio, cat# GB15003) overnight, followed by secondary antibodies (Cell Signaling, cat# 7074) for 1 hour. Membranes were rinsed in TBST buffer and exposed to Western ECL Substrate (BIO-RAD) for 3 minutes before visualization.

Cell Culture

The cell lines utilized in this study were originally sourced from the American Type Culture Collection (ATCC). Ect1/E6E7, HeLa, and C33A cells were nurtured in Dulbecco's Modified Eagle medium (DMEM, Gibco), which was enriched with 10% fetal bovine serum (FBS, Gibco). All cell lines were maintained in an incubator set at 37 °C with 5% CO2.

Construction of Stable EGFL7 Overexpression HeLa Cell Line

Lentiviral constructs, pLV4ltr-EV and pLV4ltr-EGFL7, were created to infect HeLa cells for 72 hours. Following the infection, the cells underwent a 2-week selection period in a medium supplemented with 2 μg/mL purinomycin (Solarbio Life Sciences). After this selection period, the efficiency of infection was assessed using flow cytometry. RT-qPCR and Western blot analyses were conducted to further validate the expression levels of EGFL7.

RNA-Sequence Analysis

Total RNA was extracted from six samples (three control and three EGFL7-overexpression replicates). RNA quality was confirmed using an Agilent 2100 Bioanalyzer and gel electrophoresis. mRNA was enriched by rRNA depletion with the Ribo-Zero™ Magnetic Kit, fragmented, and reverse-transcribed into cDNA using the NEBNext Ultra RNA Library Prep Kit. The cDNA underwent end repair, A-tailing, and Illumina adapter ligation, followed by purification with AMPure XP Beads, size selection, and PCR amplification. Libraries were sequenced on an Illumina Novaseq6000 at Guangzhou Gene Denovo Biotechnology Co., Ltd. Raw reads were processed with fastp to remove adapters and low-quality bases. Clean reads were aligned to an rRNA database with Bowtie2 to discard ribosomal RNA. The remaining reads were mapped to the reference genome using HISAT2. Gene abundance was quantified by StringTie and RSEM, normalized via TPM. Differential expression analysis was performed with DESeq2, identifying significant DEGs using a threshold of $|\text{Log2 FC}| > 0.585$ and P-value < 0.05. Subsequently, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses (PEA) were conducted on the differentially expressed genes (DEGs) to understand

their biological functions and signaling pathways.[29] The transcriptome data analyses were performed through the utilization of an online platform (Omicsmart, Guangzhou, China).

Statistical Analysis

GraphPad Prism 8.3.0 software was used for statistical analysis. An analysis of Hardy-Weinberg equilibrium was conducted in the control group to assess population representativeness. Age differences among the CIN, CC, and control groups were assessed using one-way ANOVA, with the LSD test applied for multiple comparison correction. A post-hoc power analysis was performed using the G*Power software (v3.1.9.2)[30] for the primary significant findings related to rs9411260. With our sample size (934 cases, 694 controls), an alpha level of 0.05, and the observed effect size, the achieved statistical power reached 80.4%. Allele frequencies and genotype distribution disparities for rs1332793, rs9411260, and rs2297538 in the cervical cancer and control groups were assessed through the χ2 test and further validated using multivariate logistic regression analysis with adjustment for age. Strong linkage between SNPs was evaluated using SHEsis online software.[31] The SnpStats online software[32] was employed to determine the optimal genetic pattern based on AIC (akaike information criterion) and BIC (bayesian information criterion) values. Haploview 4.2 software[33] was used to calculate linkage disequilibrium (LD) among these SNPs. Haplotypes were constructed, and χ2 tests were utilized to assess the differing distributions of haplotypes among the CIN, CC, and control groups. Relative expression levels of EGFL7 were represented as the means of $2^{-\Delta\Delta C}$t,[34] prior to analysis, the amplification efficiencies of the target (EGFL7) and reference (GAPDH) genes were validated via standard curves, confirming that the efficiencies were approximately equal and thus satisfying the key condition for using the $2^{-\Delta\Delta Ct}$ method. The normality of data distribution and homogeneity of variances were verified using the Shapiro–Wilk test and F-test, respectively. For data that met the assumptions of parametric tests, a two-tailed *t*-test was used for inter-group comparisons. For data that violated the normality assumption, the non-parametric Mann–Whitney *U*-test was applied.

Results

Association of EGFL7 Gene SNPs with CIN and CC

A total of 2036 subjects, comprising 694 healthy individuals, 408 CIN patients, and 934 cervical cancer patients were enrolled in the current study. The general characteristics of the subjects are presented in Table 1. No significant difference in age was found s among the control, CIN, and CC groups.

All three SNPs (rs1332793, rs9411260, and rs2297538) conformed to Hardy-Weinberg equilibrium (HWE) in the control groups (P > 0.05). The allelic and genotypic distributions of these SNPs among the control, CIN, and CC groups are outlined in Table 2. Notably, rs2297538 exhibited no differences in allele and genotype distributions between the control and CC groups, or between the control and CIN groups (P > 0.017). However, the allele and genotype distributions of rs1332793 displayed significant disparities between the control and CIN groups (P<0.017). Specifically, the C allele of rs1332793 appeared to be associated with an increased risk of CIN (OR=1.322, 95% CI: 1.095~1.597). Furthermore, the frequencies of the G allele of rs9411260 were significantly higher in both the CC group (P = 0.006, OR= 1.272, 95% CI: 1.07~1.513, Table 2) and the CIN group (P = 0.016, OR= 1.294, 95% CI: 1.047~1.599, Table 2) compared to the control group. These findings suggest a potential association between EGFL7 polymorphism and cervical cancer progression.

Table 1 Characteristics of the Subjects Enrolled in the Current Study

	Control	CIN	CC	*F*	*P* value
N	694	408	934		
Age	46.48±7.40	45.71±7.43	46.01±9.86	1.148	0.3175

Table 2 The Allelic and Genotypic Distribution of SNPs in EGFL7 Genes Among Control, CIN and Cervical Cancer Groups

SNP	Control	CIN	CC	Control vs CC		Control vs CIN	
				P value	OR[95% CI]	*P* value	OR[95% CI]
rs1332793							
C	395(0.285)	280(0.343)	558(0.299)	0.375	1.072 [0.919~1.25]	0.003	1.322[1.095~1.597]
T	993(0.715)	536(0.657)	1310(0.701)				
C/C	56(0.081)	43(0.105)	81(0.087)	0.665		0.011	
C/T	283(0.408)	194(0.475)	396(0.424)				
T/T	355(0.512)	171(0.419)	457(0.489)				
rs9411260							
G	256(0.184)	186(0.228)	419(0.224)	0.006	1.272[1.07~1.513]	0.016	1.294[1.047~1.599]
A	1132(0.816)	630(0.772)	1449(0.776)				
A/A	461(0.664)	247(0.605)	566(0.606)	0.020		0.033	
A/G	210(0.303)	136(0.333)	317(0.339)				
G/G	23(0.033)	25(0.061)	51(0.055)				
rs2297538							
A	221(0.159)	120(0.147))	299(0.160)	0.985	1.002[0.829~1.211]	0.423	0.909 [0.72~1.148]
G	1167(0.841)	696(0.853)	1569(0.840)				
A/A	20(0.029)	18(0.044)	22(0.024)	0.713		0.063	
A/G	181(0.261)	84(0.206)	255(0.273)				
G/G	493(0.710)	306(0.750)	657(0.703)				

Inheritance Model Analysis in CIN, CC, and Control Groups

Inheritance model analysis was conducted to determine the most suitable inheritance model (including codominant, dominant, recessive, overdominant and log-additive) for each SNP within the control, CIN, and CC groups. The results showed that the C/T-C/C genotype of rs1332793, in the dominant model, posed a risk for CIN (P = 0.0031, OR=1.45, 95% CI: 1.14–1.85, Table 3). Additionally, the rs9411260 genotype exhibited a significant association with cervical cancer risk in a log-additive model (P = 0.006, OR=1.27, 95% CI: 1.08–1.52, Table 3).

Haplotype Analysis in CIN, CC, and Control Groups

Linkage disequilibrium coefficient D (D') among these 3 SNPs exceeded 0.7, signifying their linkage disequilibrium. Haplotypes for these three SNPs (rs1332793-rs9411260-rs2297538) were constructed, and the distributions of haplotypes with frequencies exceeding 3% were analyzed within the control, CIN, and CC groups. The findings revealed that, in comparison to the control group, the CGG frequency was notably higher (OR = 2.086, 95% CI: 1.528~2.848, P = 2.54e-6), while the TAG frequency significantly decreased in the CIN group (P = 0.003, OR= 0.754, 95% CI: 0.624~0.911, Table 4).

eQTL Analysis of EGFL7 Single Nucleotide Polymorphisms and Tissue-Specific Gene Expression Levels

To assess the impact of EGFL7 single-nucleotide polymorphisms on EGFL7 expression, we examined the correlations between the genotypes of rs9411260 and rs1332793 and EGFL7 expression using data from the GTEx project.[35] According to the GTEx database (Figure S1), the G/G genotype of rs9411260 correlated with decreased EGFL7 expression in adipose (P = 4.08e-13), skin (P = 2.51e-8), nerve tissues (P = 2.21e-7), and cultured fibroblasts (P = 1.29e-7), implying potential roles for rs9411260 in regulating EGFL7 expression. However, the eQTL analysis results for rs9411260 in the 63 cervical cancer tissues we collected did not reach conventional significance levels, likely due to the small sample size (data not shown).

https://doi.org/10.2147/IJGM.S563592

Table 3 Inheritance Model Analysis of SNPs in *EGFL7* Gene Among Control, CIN and CC Groups

SNPs	Models	Genotypes	Control	CIN	CC	Control vs CC				Control vs CIN			
						OR[95% CI]	P value	AIC	BIC	OR[95% CI]	P value	AIC	BIC
rs1332793	Codominant	T/T	355 (51.1%)	171 (41.9%)	457 (48.9%)	1	0.66	2227.4	2249	1	0.011	1448.8	1468.8
		C/T	283 (40.8%)	194 (47.5%)	396 (42.4%)	1.09 (0.88–1.33)				1.43 (1.10–1.85)			
		C/C	56 (8.1%)	43 (10.5%)	81 (8.7%)	1.12 (0.78–1.64)				1.59 (1.03–2.5)			
	Dominant	T/T	355 (51.1%)	171 (41.9%)	457 (48.9%)	1	0.38	2225.5	2241.7	1	0.0031	1447.1	1462.1
		C/T-C/C	339 (48.9%)	237 (58.1%)	477 (51.1%)	1.09(0.90–1.33)				1.45 (1.14–1.85)			
	Recessive	T/T-C/T	638 (91.9%)	365 (89.5%)	853 (91.3%)	1	0.64	2226	2242.2	1	0.17	1453.9	1468.9
		C/C	56 (8.1%)	43 (10.5%)	81 (8.7%)	1.09 (0.76–1.56)				1.35 (0.88–2.04)			
	Overdominant	T/T-C/C	411 (59.2%)	214 (52.5%)	538 (57.6%)	1	0.53	2225.9	2242	1	0.03	1451.1	1466.1
		C/T	283 (40.8%)	194 (47.5%)	396 (42.4%)	1.06 (0.87–1.30)				1.32 (1.03–1.69)			
	Log-additive	-	-	-	-	1.08(0.92–1.25)	0.37	2225.5	2241.7	1.32 (1.10–1.59)	0.0037	1447.4	1462.4
rs9411260	Codominant	A/A	461 (66.4%)	247 (60.5%)	566 (60.6%)	1	0.019	2220.4	2241.9	1	0.037	1451.2	1471.3
		G/A	210 (30.3%)	136 (33.3%)	317 (33.9%)	1.22 (0.99–1.51)				1.19 (0.92–1.56)			
		G/G	23 (3.3%)	25 (6.1%)	51 (5.5%)	1.82 (1.09–3.03)				2.04 (1.14–3.70)			
	Dominant	A/A	461 (66.4%)	247 (60.5%)	566 (60.6%)	1	0.017	2220.6	2236.8	1	0.057	1452.2	1467.2
		G/A-G/G	233 (33.6%)	161 (39.5%)	368 (39.4%)	1.28 (1.04–1.56)				1.28 (0.99–1.64)			
	Recessive	A/A-G/A	671 (96.7%)	383 (93.9%)	883 (94.5%)	1	0.035	2221.8	2238	1	0.028	1451	1466
		G/G	23 (3.3%)	25 (6.1%)	51 (5.5%)	1.69 (1.02–2.77)				1.92 (1.08–3.45)			
	Overdominant	A/A-G/G	484 (69.7%)	272 (66.7%)	617 (66.1%)	1	0.13	2223.9	2240.1	1	0.33	1454.9	1469.9
		G/A	210 (30.3%)	136 (33.3%)	317 (33.9%)	1.18 (0.95–1.45)				1.14 (0.88–1.49)			
	Log-additive	-	-	-	-	1.27 (1.08–1.52)	0.006	2218.7	2234.9	1.30 (1.05–1.59)	0.017	1450.2	1465.2
rs2297538	Codominant	G/G	493 (71%)	306 (75%)	657 (70.3%)	1	0.73	2227.6	2249.2	1	0.059	1452.2	1472.2
		A/G	181 (26.1%)	84 (20.6%)	255 (27.3%)	1.05 (0.84–1.32)				0.74 (0.55–1.00)			
		A/A	20 (2.9%)	18 (4.4%)	22 (2.4%)	0.83 (0.44–1.54)				1.43 (0.74–2.78)			
	Dominant	G/G	493 (71%)	306 (75%)	657 (70.3%)	1	0.8	2226.2	2242.4	1	0.14	1453.6	1468.6
		A/G-A/A	201 (29%)	102 (25%)	277 (29.7%)	1.03 (0.83–1.28)				0.81 (0.61–1.08)			
	Recessive	G/G-A/G	674 (97.1%)	390 (95.6%)	912 (97.6%)	1	0.51	2225.8	2242	1	0.2	1454.2	1469.2
		A/A	20 (2.9%)	18 (4.4%)	22 (2.4%)	0.81 (0.44–1.49)				1.54 (0.80–2.94)			
	Overdominant	G/G-A/A	513 (73.9%)	324 (79.4%)	679 (72.7%)	1	0.62	2226	2242.2	1	0.034	1451.3	1466.3
		A/G	181 (26.1%)	84 (20.6%)	255 (27.3%)	1.05 (0.85–1.32)				0.73 (0.54–0.98)			
	Log-additive	-	-	-	-	1.00 (0.83–1.20)	0.99	2226.3	2242.4	0.91 (0.72–1.15)	0.42	1455.2	1470.2

Note: The statistical significant threshold was set at $P < 0.017$ (0.05/n, n=3) after Bonferroni correction.

Table 4 The Distribution of Haplotype Among Control, CIN and CC Groups

Haplotypes[a]	Control	CIN	CC	Control VS CC		Control VS CIN	
	Counts (2n=1388)	Counts (2n=816)	Counts (2n=1868)	OR[95% CI]	*P* value	OR[95% CI]	*P* value
C A A	10(0.007)	6(0.008)	1(0.001)	/	/	/	/
C A G	133(0.096)	92(0.113)	162(0.087)	0.882 [0.693~1.122]	0.307	1.202 [0.907~1.593]	0.200
C G A	170(0.122)	88(0.108)	251(0.134)	1.095 [0.889~1.350]	0.394	0.867 [0.659~1.141]	0.308
C G G	81(0.059)	94(0.115)	144(0.077)	1.325 [1.000~1.755]	0.050	2.086 [1.528~2.848]	2.54e-6
T A A	41(0.029)	24(0.029)	27(0.015)	/	/	/	/
T A G	948(0.683)	508(0.623)	1259(0.674)	0.918 [0.786~1.071]	0.276	0.754 [0.624~0.911]	0.003
T G A	0(0.000)	2(0.003)	20(0.011)	/	/	/	/
T G G	5(0.003)	2(0.003)	4(0.002)	/	/	/	/

Notes: [a]haplotypes, constructed by rs1332793-rs9411260-rs2297538. The statistical significant threshold was set at P < 0.0125 (0.05/n, n = 4) after Bonferroni correction.

Abbreviations: CIN, Cervical intraepithelial neoplasia; CC, Cervical cancer; vs Versus; OR, Odds ratio; CI, Confidence Interval.

EGFL7 mRNA Expression in Cervical Cancer Tissues and Cells

EGFL7 mRNA expression was quantified using qRT-PCR. The findings revealed a notable reduction in EGFL7 expression within cervical cancer cells (HeLa and C33A) when compared to ECT1/E6E7 cells (Figure 1A). Subsequently, we assessed EGFL7 mRNA in 63 cervical cancer tissues alongside their corresponding noncancerous tissues. As depicted in Figure 1B, EGFL7 expression exhibited significant downregulation in cervical cancer tumor tissues when contrasted with normal cervical tissues. This observation was further corroborated by the Cancer Genome Atlas (TCGA) RNA-Seq database, which also demonstrated significantly lower EGFL7 expression in cervical cancer tissue than in normal controls or matched normal cervical tissues (Figure 1C). These collective results underscore the potential involvement of EGFL7's dysregulated expression in cervical cancer pathogenesis.

Establishing Stable EGFL7 Overexpression in HeLa Cells

To establish a stable EGFL7-overexpressing HeLa cell line, lentiviral transduction was employed. Infection efficiency was evaluated by flow cytometry based on GFP expression, which revealed more than 90% of cells were GFP-positive in

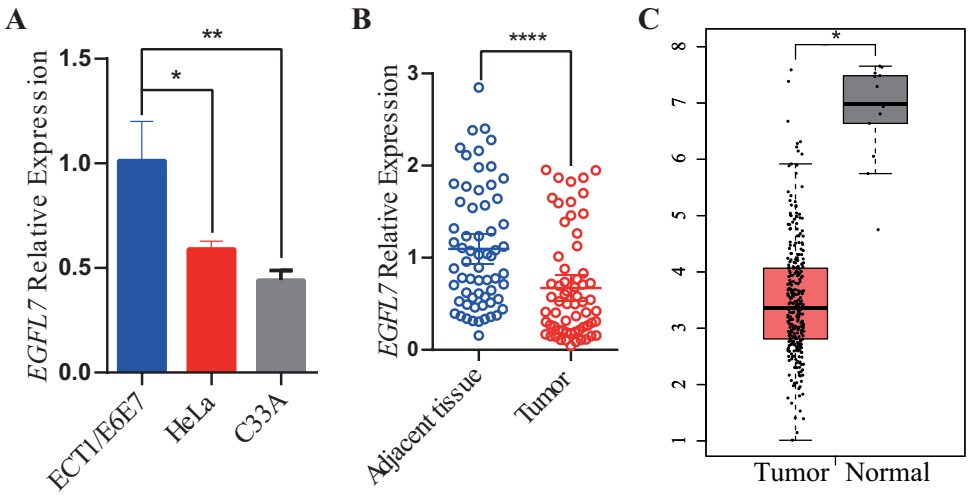

Figure 1 EGFL7 expression in cervical cancer cells and tissues. (**A**) EGFL7 mRNA levels were measured by qRT-PCR in ECT1/E6E7, HeLa, and C33A cells, with EGFL7 expression normalized to beta-actin levels. Data were compared using a two-tailed paired *t*-test. (**B**) EGFL7 mRNA expression was examined via qRT-PCR in 63 cervical cancer tissues and their matched noncancerous tissues, with EGFL7 expression normalized to beta-actin levels. Data were compared using a two-tailed paired *t*-test. (**C**) Relative EGFL7 expression in cervical cancer tissue compared to normal tissue from the GEPIA database. The symbol ns represents no statistical significance, P < 0.05 was considered statistically significant, indicated by * for P < 0.05, ** for P < 0.01and **** for P < 0.0001.

https://doi.org/10.2147/IJGM.S563592

both the pLV4ltr-EV and pLV4ltr-EGFL7 groups, confirming highly efficient transduction (Figure 2A). Subsequent RT-qPCR and Western blot analyses were performed to validate the upregulation of EGFL7. The results indicated a marked increase in EGFL7 expression in pLV-EGFL7-infected cells compared with the control group. Specifically, RT-qPCR demonstrated that the relative EGFL7 mRNA level was 339.6 ± 111.43 in EGFL7-overexpressing cells versus 1 ± 0.21 in control cells (Figure 2B). Consistent with this, Western blot analysis also confirmed substantial elevation of EGFL7 protein expression (Figure 2C and Figure S2).

Functional Annotations of EGFL7-Driven DEGs

To identify differentially expressed genes (DEGs) driven by EGFL7 in HeLa cells, we conducted RNA-seq analysis. Applying thresholds of |Log2 FC| > 0.585 and P-value < 0.05, we identified 246 upregulated and 364 downregulated genes in the HeLa group (Figure 3A). The RNA-seq analysis confirmed EGFL7 as the most significantly upregulated

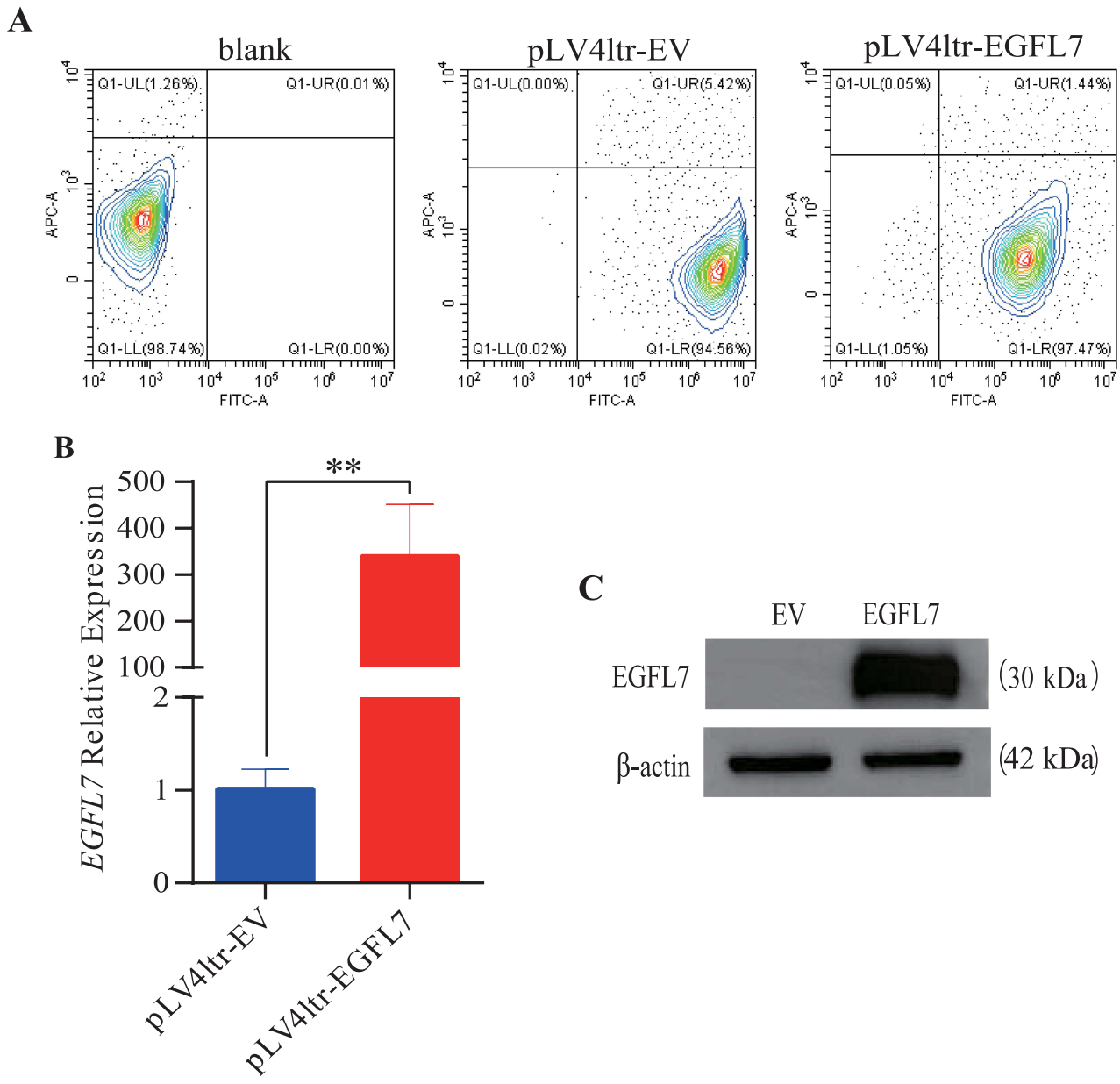

Figure 2 Establishment of stable EGFL7 overexpression HeLa cell lines.(**A**) Flow cytometric analysis of pLV4ltr-EV and pLV4ltr-EGFL7 infected cells. (**B**) EGFL7 expression levels were confirmed via RT-qPCR. (**C**) Expression levels of EGFL7 were validated through Western blotting and the cropped blots were displayed. The symbol ns represents no statistical significance, ** for P < 0.01.

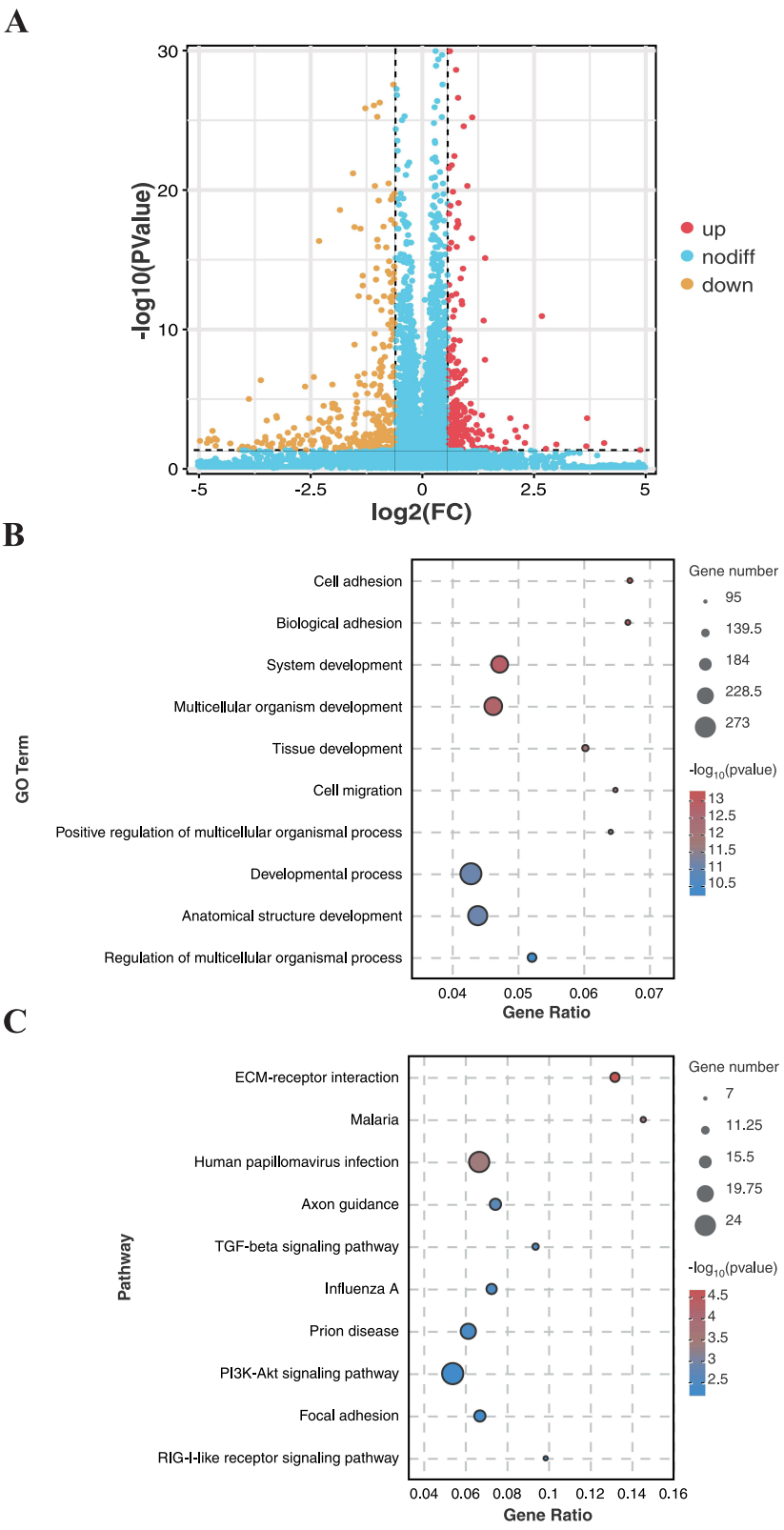

Figure 3 Comparison of gene expression profiles Induced by EGFL7 in cervical cancer. (**A**) Volcano plot showing significantly differentially expressed genes (DEGs) in HeLa cells. (**B**) Bubble plot illustrating the results of GO functional enrichment analysis. (**C**) Bubble plot displaying the outcomes of KEGG analysis.

https://doi.org/10.2147/IJGM.S563592

gene, with its expression (TPM) rising from 3.91 ± 0.41 in controls to 1305.27 ± 76.29 in EGFL7-overexpressing HeLa cells (P = 0.001). This result is consistent with the qRT-PCR validation data presented in Figure 2B. Our GO enrichment analysis unveiled the top ten significantly enriched biological process (BP) terms, encompassing cell adhesion, biological adhesion, system development, multicellular organism development, tissue development, cell migration, positive regulation of multicellular organismal processes, developmental processes, anatomical structure development, and regulation of multicellular organismal processes (Figure 3B). Furthermore, our KEGG results indicated the top ten significantly enriched pathways, including ECM-receptor interaction, Malaria, Human papillomavirus infection, Axon guidance, TGF-beta signaling pathway, Influenza A, Prion disease, PI3K-Akt signaling pathway, Focal adhesion, and RIG-I-like receptor signaling pathway (Figure 3C). These functional annotations imply that manipulating EGFL7 expression could induce alterations in cervical cancer cell adhesion and migration.

Discussion

Our study has established a clear link between EGFL7 polymorphisms and the risk of cervical cancer. It is noteworthy that a previous investigation we conducted identified another common SNP, rs4636297, situated within the EGFL7 gene region, as a factor associated with cervical cancer susceptibility.[24] Additionally, in 2017, a study by Hansen et al unveiled a significant correlation between the EGFL7 3'UTR variant rs1051851 and the overall survival of metastatic colorectal cancer patients.[36] These findings suggest that EGFL7 polymorphisms may have implications for various cancer types.

Numerous studies have highlighted EGFL7's atypical expression in a range of human cancers. It has been noted that EGFL7 exhibits high expression levels in many cancer types, including colorectal cancer,[22] gastric cancer,[37] hepatocellular carcinoma[38] and pancreatic cancer.[39] Conversely, a handful of studies have indicated that EGFL7 is underexpressed in cases of malignant pleural mesothelioma,[20] primary bladder and prostate cancer.[40] These diverse findings underscore the multifaceted roles EGFL7 may play in different cancer types. Our results showed that the EGFL7 expression was downregulated in cervical cancer tissues and cells, this is in accordance with findings in malignant pleural mesothelioma,[20] primary bladder and prostate cancer.[40] This alignment is further substantiated by data from the Cancer Genome Atlas (TCGA) RNA-Seq database.

Moreover, expression quantitative trait locus (eQTL) analyses revealed that the G/G genotype of rs9411260 is associated with reduced EGFL7 expression in multiple tissue types, including adipose, skin, nerve tissues, and cultured fibroblasts. Notably, JASPAR-based transcription factor binding site prediction indicated that rs9411260 resides within a putative binding site for NFKB1, and the G allele may impair NFKB1 binding affinity, potentially leading to diminished transcriptional activation of EGFL7. Given the significantly elevated frequency of the rs9411260 G allele in the cervical cancer cohort, we propose that this variant may contribute to cervical cancer susceptibility by disrupting NFKB1-mediated transcriptional regulation, thereby reducing EGFL7 expression. This downregulation of EGFL7 may in turn represent a novel risk mechanism in cervical carcinogenesis.

Functional annotations of EGFL7-driven DEGs suggest that modifying EGFL7 expression could impact ECM-receptor interaction, as well as the adhesion and migration of cervical cancer cells. Previous research has established EGFL7's involvement in intercellular and cell-matrix communication.[10,41] Li et al documented EGFL7's role in reducing endothelial cell adhesion molecule expression.[42] Furthermore, EGFL7 was initially recognized for its capacity to inhibit platelet-derived growth factor (PDGF-BB)-induced smooth muscle cell migration[43] Recent studies illuminated that EGFL7 could modulate cell migration by interacting with extracellular matrix (ECM) sensing integrins.[41,44] Our findings align with these prior studies, indicating that abnormal EGFL7 expression may disrupt the cell adhesion and migration capabilities of cervical cancer cells through cell-matrix communication. As a result, EGFL7's involvement in cervical cancer progression becomes evident.

Based on our findings, therapeutic strategies targeting EGFL7 could be explored. Given its downregulation in cervical cancer, restoring EGFL7 expression via gene therapy or small-molecule activators may counteract tumor progression. For patients with risk-associated SNPs like rs9411260, personalized approaches such as EGFL7 mimetics or epigenetic modulators could be beneficial. Alternatively, targeting downstream pathways like ECM-receptor interaction or PI3K-Akt signaling may offer additional therapeutic avenues. Further preclinical studies are needed to validate these strategies.

Conclusions

EGFL7 polymorphisms and expression are associated with cervical cancer susceptibility and progression. The rs9411260 G allele may serve as a genetic risk biomarker, while EGFL7 downregulation suggests its potential as a therapeutic target. Further in vivo and protein-level studies are warranted to validate these findings.

Abbreviations

EGFL7, epidermal growth factor-like domain 7; CC, cervical cancer; CIN, cervical intraepithelial neoplasia; HPV, human papillomaviruses; HWE, Hardy-Weinberg equilibrium; ORs, odds ratios; SNP, single-nucleotide polymorphism; DEGs, differentially expressed genes; RTFQ-PCR, real-time fluorescence quantitative polymerase chain reaction; AIC, akaike information criterion; BIC, bayesian information criterion; ECM, extracellular matrix; TCGA, the cancer genome atlas; GTEx, genotype-tissue expression.

Data Sharing Statement

The datasets generated and/or analyzed during the current study are available from the corresponding author, Dr. Yufeng Yao (leoyyf@gmail.com), for non-commercial research purposes upon reasonable request. Additionally, the raw sequence reads and expression level data have been deposited in the Gene Expression Omnibus under the accession number GSE248647.

Ethics Approval and Consent to Participate

This study received approval from the Ethics Committee of the Third Affiliated Hospital of Kunming Medical University (approval no. KYCS202195). All procedures were in accordance with the approved guidelines and principles expressed in the Helsinki Declaration and its later amendments or comparable ethical standards of 1975. And all participants provided written informed consent.

Acknowledgments

We are grateful to patients and healthy control individuals who participated in this study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by grants from the National Science Foundation for Young Scientists of China (82103190), the Yunnan Provincial Science and Technology Department (202501AT070049 and 202201AU070163), Yunnan Province Xingdian Talent Support Program (XDYC-QNRC-2023-0590 and XDYC-CYCX-2023-0074) and Yunnan Provincial High-Level Health Technology Talent Training Special Program (H-2024074). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a Cancer Journal for Clinicians*. 2021;71(3):209–249. doi:10.3322/caac.21660
2. Walboomers JM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of Pathology*. 1999;189(1):12–19. doi:10.1002/(sici)1096-9896(199909)189:1<12::Aid-path431>3.0.Co;2-f

3. Chen D, Gyllensten U. Lessons and implications from association studies and post-GWAS analyses of cervical cancer. *Trends in Genetics: TIG*. 2015;31(1):41–54. doi:10.1016/j.tig.2014.10.005
4. Bahrami A, Hasanzadeh M, Shahidsales S, et al. Genetic susceptibility in cervical cancer: from bench to bedside. *Journal of Cellular Physiology*. 2018;233(3):1929–1939. doi:10.1002/jcp.26019
5. Magnusson PK, Lichtenstein P, Gyllensten UB. Heritability of cervical tumours. *International Journal of Cancer*. 2000;88(5):698–701. doi:10.1002/1097-0215(20001201)88:5<698::aid-ijc3>3.0.co;2-j
6. Vink JM, van Kemenade FJ, Meijer CJ, Casparie MK, Meijer GA, Boomsma DI. Cervix smear abnormalities: linking pathology data in female twins, their mothers and sisters. *European Journal of Human Genetics*. 2011;19(1):108–111. doi:10.1038/ejhg.2010.139
7. Chen D, Cui T, Ek WE, Liu H, Wang H, Gyllensten U. Analysis of the genetic architecture of susceptibility to cervical cancer indicates that common SNPs explain a large proportion of the heritability. *Carcinogenesis*. 2015;36(9):992–998. doi:10.1093/carcin/bgv083
8. Zhang X, Zhang L, Tian C, Yang L, Wang Z. Genetic variants and risk of cervical cancer: epidemiological evidence, meta-analysis and research review. *BJOG*. 2014;121(6):664–674. doi:10.1111/1471-0528.12638
9. Fitch MJ, Campagnolo L, Kuhnert F, Stuhlmann H. Egfl7, a novel epidermal growth factor-domain gene expressed in endothelial cells. *Developmental Dynamics: an Official Publication of the American Association of Anatomists*. 2004;230(2):316–324. doi:10.1002/dvdy.20063
10. Chim SM, Kuek V, Chow ST, et al. EGFL7 is expressed in bone microenvironment and promotes angiogenesis via ERK, STAT3, and integrin signaling cascades. *Journal of Cellular Physiology*. 2015;230(1):82–94. doi:10.1002/jcp.24684
11. Lacko LA, Hurtado R, Hinds S, Poulos MG, Butler JM, Stuhlmann H. Altered feto-placental vascularization, feto-placental malperfusion and fetal growth restriction in mice with Egfl7 loss of function. *Development*. 2017;144(13):2469–2479. doi:10.1242/dev.147025
12. Nikolic I, Stankovic ND, Bicker F, et al. EGFL7 ligates αvβ3 integrin to enhance vessel formation. *Blood*. 2013;121(15):3041–3050. doi:10.1182/blood-2011-11-394882
13. Luo BH, Xiong F, Wang JP, et al. Epidermal growth factor-like domain-containing protein 7 (EGFL7) enhances EGF receptor-AKT signaling, epithelial-mesenchymal transition, and metastasis of gastric cancer cells. *Journal Pone*. 2014;9(6):e99922. doi:10.1371/journal.pone.0099922
14. Zhai W, Zhu R, Ma J, et al. A positive feed-forward loop between LncRNA-URRCC and EGFL7/P-AKT/FOXO3 signaling promotes proliferation and metastasis of clear cell renal cell carcinoma. *Molecular Cancer*. 2019;18(1):81. doi:10.1186/s12943-019-0998-y
15. Shen X, Han Y, Xue X, et al. Epidermal growth factor-like domain 7 promotes cell invasion and angiogenesis in pancreatic carcinoma. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*. 2016;77:167–175. doi:10.1016/j.biopha.2015.12.009
16. Papaioannou D, Shen C, Nicolet D, et al. Prognostic and biological significance of the proangiogenic factor EGFL7 in acute myeloid leukemia. *Proceedings of the National Academy of Sciences of the United States of America*. 2017;114(23):E4641–e47. doi:10.1073/pnas.1703142114
17. Philippin-Lauridant G, Baranzelli MC, Samson C, et al. Expression of Egfl7 correlates with low-grade invasive lesions in human breast cancer. *International Journal of Oncology*. 2013;42(4):1367–1375. doi:10.3892/ijo.2013.1820
18. Campagnolo L, Telesca C, Massimiani M, et al. Different expression of VEGF and EGFL7 in human hepatocellular carcinoma. *Digestive and Liver Disease*. 2016;48(1):76–80. doi:10.1016/j.dld.2015.09.019
19. Fan C, Yang LY, Wu F, et al. The expression of Egfl7 in human normal tissues and epithelial tumors. *Int J Biol Markers*. 2013;28(1):71–83. doi:10.5301/JBM.2013.10568
20. Andersen M, Trapani D, Ravn J, et al. Methylation-associated Silencing of microRNA-126 and its Host Gene EGFL7 in Malignant Pleural Mesothelioma. *Anticancer Research*. 2015;35(11):6223–6229.
21. de Oliveira C, Gonçalves PG, Bidinotto LT. Role of EGFL7 in human cancers: a review. *Journal of Cellular Physiology*. 2023;238(8):1756–1767. doi:10.1002/jcp.31084
22. Juan Z, Dake C, Tanaka K, Shuixiang H. EGFL7 as a novel therapeutic candidate regulates cell invasion and anoikis in colorectal cancer through PI3K/AKT signaling pathway. *International Journal of Clinical Oncology*. 2021;26(6):1099–1108. doi:10.1007/s10147-021-01888-x
23. Yamauchi M, Fukuda T, Wada T, et al. Expression of epidermal growth factor-like domain 7 may be a predictive marker of the effect of neoadjuvant chemotherapy for locally advanced uterine cervical cancer. *Oncology Letters*. 2016;12(6):5183–5189. doi:10.3892/ol.2016.5318
24. Yan Z, Zhou Z, Li C, et al. Polymorphisms in miRNA genes play roles in the initiation and development of cervical cancer. *Journal of Cancer*. 2019;10(20):4747–4753. doi:10.7150/jca.33486
25. Harris TA, Yamakuchi M, Kondo M, Oettgen P, Lowenstein CJ. Ets-1 and Ets-2 regulate the expression of microRNA-126 in endothelial cells. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 2010;30(10):1990–1997. doi:10.1161/atvbaha.110.211706
26. Liu W, Zhang Y, Huang F, et al. The Polymorphism and Expression of EGFL7 and miR-126 Are Associated With NSCLC Susceptibility. *Frontiers in Oncology*. 2022;12:772405. doi:10.3389/fonc.2022.772405
27. Fornes O, Castro-Mondragon JA, Khan A, et al. JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Research*. 2020;48(D1):D87–d92. doi:10.1093/nar/gkz1001
28. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biology*. 2016;17(1):122. doi:10.1186/s13059-016-0974-4
29. Ali A, Mohan J, Nadaf TAA, Ravishankar H, Deepa KRJSCS. Bioinformatics-Driven Discovery of Signaling Pathwaysand GenesInfluencing Cervical Cancer. SN Computer Sci. 2024;5(8):1–7.
30. Kang H. Sample size determination and power analysis using the G*Power software. *Journal of Educational Evaluation for Health Professions*. 2021;18:17. doi:10.3352/jeehp.2021.18.17
31. Shi YY, He L. SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Research*. 2005;15(4):97–98. doi:10.1038/sj.cr.7290272
32. Solé X, Guinó E, Valls J, Iniesta R, Moreno V. SNPStats: a web tool for the analysis of association studies. *Bioinformatics*. 2006;22(15):1928–1929. doi:10.1093/bioinformatics/btl268
33. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*. 2005;21(2):263–5. doi:10.1093/bioinformatics/bth457
34. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. 2001;25(4):402–408. doi:10.1006/meth.2001.1262
35. Battle A, Brown CD, Engelhardt BE, Montgomery SB. Genetic effects on gene expression across human tissues. *Nature*. 2017;550(7675):204. doi:10.1038/nature24277

36. Hansen TF, Andersen RF, Olsen DA, Sørensen FB, Jakobsen A. Prognostic importance of circulating epidermal growth factor-like domain 7 in patients with metastatic colorectal cancer treated with chemotherapy and bevacizumab. *Sci Rep*. 2018;7(1):2388. doi:10.1038/s41598-017-02538-x
37. Deng QJ, Xie LQ, Li H. Overexpressed MALAT1 promotes invasion and metastasis of gastric cancer cells via increasing EGFL7 expression. *Life Sciences*. 2016;157:38–44. doi:10.1016/j.lfs.2016.05.041
38. Yang MY, Wu F, Fang F, et al. Serum epidermal growth factor-like domain 7 serves as a novel diagnostic marker for early hepatocellular carcinoma. *Cancer*. 2021;21(4):772. doi:10.1186/s12885-021-08491-3
39. Zhou L, Li J, Zhao YP, et al. Prognostic significance of epidermal growth factor-like domain 7 in pancreatic cancer. *HBPD INT*. 2014;13 (5):523–528. doi:10.1016/s1499-3872(14)60272-1
40. Saito Y, Friedman JM, Chihara Y, Egger G, Chuang JC, Liang G. Epigenetic therapy upregulates the tumor suppressor microRNA-126 and its host gene EGFL7 in human cancer cells. *Biochemical and Biophysical Research Communications*. 2009;379(3):726–731. doi:10.1016/j.bbrc.2008.12.098
41. Heissig B, Salama Y, Takahashi S, Okumura K, Hattori K. The Multifaceted Roles of EGFL7 in Cancer and Drug Resistance. *Cancers*. 2021;13 (5):1014. doi:10.3390/cancers13051014
42. Chen S, Li Y, Zhu Y, et al. SERPINE1 Overexpression Promotes Malignant Progression and Poor Prognosis of Gastric Cancer. *Journal of Oncology*. 2022;2022:2647825. doi:10.1155/2022/2647825
43. Soncin F, Mattot V, Lionneton F, et al. VE-statin, an endothelial repressor of smooth muscle cell migration. *THE EMBO Journal*. 2003;22 (21):5700–5711. doi:10.1093/emboj/cdg549
44. Michael M, Parsons M. New perspectives on integrin-dependent adhesions. *Current Opinion in Cell Biology*. 2020;63:31–37. doi:10.1016/j.ceb.2019.12.008