Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Genetic Variant rs10251977 in Long Non-coding RNA EGFR-AS1 Creates a New Binding Site for miR-891b and Modulates the Expression of EGFR A/D Isoforms

Dhamodharan Shankar
Dr. ALM PG Institute of Basic Medical Sciences, University of Madras
Rose Mathew
Dr. ALM PG Institute of Basic Medical Sciences, University of Madras
Sundaram Reddy Chakkarappan
Dr. ALM PG Institute of Basic Medical Sciences, University of Madras
Uma Dharshini Karuppiah Vijayamuthuramalingam
Dr. ALM PG Institute of Basic Medical Sciences, University of Madras
Arulmurugan Ramalingam
Royapettah Government Hospital & Kilpauk Medical College
Subbiah Shanmugam
Royapettah Government Hospital & Kilpauk Medical College
Inoue Ituro
National Institute of Genetics
Munirajan Arasambattu (akmunirajan@gmail.com)
Dr. ALM PG Institute of Basic Medical Sciences, University of Madras

Research Article

Keywords: HNSCC, EGFR, rs10251977, EGFR-AS1, miR-891b

Posted Date: December 31st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-136285/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Version of Record: A version of this preprint was published at Scientific Reports on April 22nd, 2021. See the published version at https://doi.org/10.1038/s41598-021-88161-3.

A genetic variant *rs10251977* in long non-coding RNA EGFR-AS1 creates a new binding site for miR-891b and modulates the expression of EGFR A/D isoforms

Shankar Dhamodharan[1,+], Mathew Maria Rose[1,+], Reddy Chakkarappan Sundaram[1,4], Karuppiah Vijayamuthuramalingam Uma Dharshini[1], Ramalingam Arulmurugan[2], Shanmugam Subbiah[2], Ituro Inoue[3], Arasambattu Kannan Munirajan[1,*].

[1]Department of Genetics, Dr. ALM PG Institute of Basic Medical Sciences, University of Madras, Taramani Campus, Chennai-600 113, India.

[2]Center for Oncology, Royapettah Government Hospital & Kilpauk Medical College, Royapettah, Chennai-600 014, India.

[3]Division of Human genetics, National Institute of Genetics, Mishima- 411-8540, Japan

[4]Department of Health Research – Multi Disciplinary Research Unit (DHR-MRU), Dr. ALM PG Institute of Basic Medical Sciences, University of Madras, Taramani Campus, Chennai - 600 113,India.

[+] These authors contributed equally to this work

[*]**Correspondence to:**

Arasambattu Kannan Munirajan, PhD

Department of Genetics

Dr. ALM PG Institute of Basic Medical Sciences

University of Madras, Taramani Campus

Chennai – 600113, India.

Phone: +91-44-24547064, Fax: +91-44-24540709

E-mail: akmunirajan@gmail.com; akmunirajan@unom.ac.in.

Abstract:

Tyrosine kinase inhibitor (TKI) is one of the effective chemo-preventive approaches against tumors that deregulate EGFR pathway. About 80% of HNSCC patients overexpress EGFR, making TKI an effective treatment against this cancer. Recently a synonymous variant *rs10251977* in exon 20 of *EGFR* reported to act as a prognostic marker in HNSCC. Analysis of this germline variant in blood samples of oral cancer patients showed a similar frequency in cases and controls. Further, in-silico analysis showed that this polymorphism creates binding site for miR-891b in EGFR-AS1. The EGFR-AS1 expression modulates the EGFR A/D isoforms through alternative splicing. Our bioinformatic analysis showed enrichment of alternative splicing marks H3K36me3 and presence of a few intronic polyA sites spanning around exon 15a and 15b of EGFR facilitating the skipping of exon 15b and thereby promoting the splicing of EGFR-A isoform. In addition, the presence of PTBP1 binding site in EGFR and EGFR-AS1 enhances the expression of EGFR- A isoform by preventing the premature termination. Expression profiling of EGFR-AS1 along with miR-891b level and *rs10251977* polymorphism status in oral cancer patients may be useful for targeted therapy.

KEYWORDS: HNSCC; EGFR; rs10251977; EGFR-AS1; miR-891b

Introduction

Oral squamous cell carcinoma (OSCC) is one of the prevalent cancers worldwide. According to the GLOBOCAN 2018 report from India, cancer of lip, oral cavity is the top most cancer in men and fourth most in women[1] and often diagnosed in advanced stages, making it difficult for the therapeutic management. Tobacco chewing/smoking, alcohol consumption and infection with human papilloma virus (HPV) 16/18 are the major risk factors of OSCC[2]. In India, tobacco chewing with betel quid, slaked lime, and areca nut combined with smoking or drinking significantly increases the risk. Despite the advances in diagnosis and treatment the mortality rate of oral cancer patients has not markedly improved over the past three decades and the 5-year survival rate remains less than 50%. Even with the advances in drug discovery and treatment against cancer, chemoresistance and tumor recurrence remains an obstacle for development of effective therapeutic management in the patients with OSCC[3].

The etiology (58-90%) of HNSCC is attributed with the aberrant activity of epidermal growth factor receptor (EGFR). It belongs to the family of receptor tyrosine kinases ErbB and plays an important role in cell cycle regulation, proliferation, cell migration and other physiological processes[4]. Overexpression of EGFR is observed in the early stages of the oral tumorigenesis and linked with the advanced stages of the tumor, thus it could serve as the effective drug target[5]. Though various evidences proved EGFR signalling is associated with progression of HNSCC, the effective treatment outcome is not achieved upon targeting EGFR with monoclonal antibodies and/or TKIs[6]. Therefore, it is essential to identify robust biomarkers for the success of targeted therapy.

Alternative splicing is a common mechanism employed by eukaryotic cells to create diversified proteomics profile from single gene[7,8]. This mechanism is common in receptor kinase family genes. EGFR generates a functionally important four splice variants through alternative splicing where 10.5 kb and 5.8 kb belongs to variant 1 class encoding 170 kDa protein called isoform A. This variant will be generated by skipping of two exons 15a and 15b in *EGFR*, whereas 1.8 kb, 2.4 kb and 3.0 kb are three other isoforms formed from a read-through of a exon-intron boundary and incorporation of alternate exons 15a and 15b, which encodes 60, 80 and 110 kDa proteins known as isoform B, C and D, respectively[9,10].

Currently, long non-coding RNAs (lncRNAs), a class of heterogeneous RNA molecules (>200 nt) with no protein-coding potential, has been identified as the prognostic marker in various cancer. The lncRNAs were aberrantly expressed in various cancers and play regulatory roles in several cellular pathways promoting proliferation, stem cell pluripotency, cellular reprogramming, cellular transformation, and tumorigenesis[3]. A recent study established the prognostic potential of EGFR-AS1 in tyrosine kinase inhibitors (TKIs) treatment and knockdown of EGFR-AS1 induced regression of squamous cell carcinoma of head and neck and aids in mediating a ligand addiction[6].

Antisense non-coding transcript of *EGFR* locus (EGFR-AS1), a 2.8 kb transcript and was shown to promote the stability of its *cis* partner EGFR-A isoform[11]. However, the role of EGFR-AS1 in maintaining the stability of EGFR-A isoform is yet to be studied in detail. This study is focussed on understanding the molecular consequence of a germline polymorphism *rs10251977* (c.2361G>A) in EGFR-AS1 that creates a binding site for miR-891b and modulates the stability of the EGFR-AS1. Additionally, we explored the role of EGFR-AS1 in the alternative splicing of EGFR A/D isoforms.

Results

Screening of the germline variant *rs10251977* in oral cancer patients of South Indian origin

We screened for the variant *rs10251977* (c.2361G>A) in exon 20 of *EGFR* in 180 oral cancer patients with age and sex matched 184 cancer-free controls. The demographic details of both cases and controls were presented in Table 1. Oral cancer patients and control subjects' genotype frequencies were in agreement with Hardy–Weinberg equilibrium with p values 0.79 and 0.35, respectively (Table 2). The genotype frequencies were found to be 40% (72/180) of GG, 47.2% (85/180) of GA and 12.8% (23/180) of AA in oral cancer patients, 36.4% (67/184) of GG, 50.5% (93/184) of GA and 13.1% (24/184) of AA in control subjects and no significant difference was found between cases and controls. A similar result was observed in dominant model (GA+AA vs GG), recessive model (GG+GA vs AA) and allelic model (G vs A) with an OR of 0.86 (p= 0.55), 0.98 (p= 0.93) & 0.92 (p= 0.64), respectively.

EGFR-AS1 is overexpressed in OSCCs

The EGFR-AS1 expression was analyzed in 48 oral tumor samples and 8 independent normal tissues by RT-qPCR (Table 3). A significant upregulation of EGFR-AS1 was observed in tumor

tissues compared to that of normal tissues ($p<0.001$) (Figure 1a). The expression levels were correlated with clinic-pathological characteristics of the tumors, there is a trend in increased level of EGFR-AS1 was observed without any statistically significant association (Figure S1a). Interestingly, in tumors with GA and AA (N=25) genotypes, we noticed high level expression of EGFR-AS1 albeit with no statistical significance ($p=0.653$), and this could be due to the small sample size (Figure S1b).

EGFR-AS1 in maintaining the level of EGFR A and D isoforms

To know the functional relationship of EGFR-AS1 expression with EGFR and its isoforms, we analysed the relative expression levels of EGFR isoforms A and D and D/A ratio with reference to the expression level of EGFR-AS1. The tumor samples were stratified into two groups as EGFR-AS1 high and low expression groups based on the median values of EGFR-AS1 expression (Table 3). Tumors with expression level above median value is considered as high expression group and those with fold change below median value as low expression group. We observed a statistically significant difference in the expression level of EGFR-A form with reference to EGFR-AS1 level ($p <0.0001$) and no significant difference was found for EGFR-D form (Figure 1b). But we observed significant difference in the EGFR isoform A and D levels in both the groups (Figure 1c), and found a significantly high level of EGFR isoform D transcript in low EGFR-AS1 expressing group compared to the high expression group ($p=0.017$) and this could be due to the modulation of RNA splicing. To address this, we carried out in-silico analysis to identify the EGFR-AS1/EGFR binding partners that has functional association with alternative splicing.

First, we predicted the binding partners of EGFR-AS1 by using online database lncRNAtor[12]. EGFR-AS1 showed significant mechanistic association with three RNA binding proteins FIP1, hnRNPU, and PTBP1, of which PTBP1 showed a significant association ($p=8.43E^{-26}$) (Figure S2a). To evaluate the binding of this alternative splicing factors in *EGFR* locus we used RBPmap online tool[13] to predict the binding motif of the RNA binding proteins in both EGFR-AS1 and region between exon 15 and 16 of *EGFR* gene and identified several consensus motif of PTBP1 (Table S1 and S2), suggesting the PTBP1 role in alternative splicing of EGFR-A isoform. To further support this hypothesis, we checked the binding partners for PTBP1 using STITCH online tool, and found that most of its interacting partners HNRNPK, HNRNPU,

147 HNRNPA1, HNRNPA3, HNRNPD, HNRNPL and SNRPA were known to play a role in
148 alternative splicing (Figure 2a).

149 Previous studies have shown that H3K36me3 plays an important role in exon skipping by
150 recruiting alternative splicing factors[14]. We checked whether the H3K36me3 marks are present
151 in EGFR region by using UCSC genome browser. Interestingly we found that enhanced
152 H3K36me3 marks were present around the skipped region spanning the exon 15a and b region
153 (Figure 2b), and to our surprise they also carried a few polyA sites spanning around region exon
154 15a leading to the skipping of exon 15b and prevents premature termination thereby reducing
155 EGFR-D isoform. EGFR-AS1 may act as a scaffold to recruit major alternative splicing factors
156 in association with PTBP1 to bring exon skipping and promoting EGFR-A expression.

157 **EGFR variant *rs10251977* creates miR-891b binding site in EGFR-AS1**

158 Apart from acting as scaffold, lncRNAs may also act as miRNA sponges in the cytoplasm.
159 Tan *et al*., reported that the presence of minor allele in EGFR-AS1 decreased its steady state
160 level[6] and this could be due to the miRNA sponging mediated degradation of lncRNA.
161 We used a web-based tool called lncRNASNP2 to determine the functional impact of minor
162 allele in EGFR-AS1.

163 The minor allele generates a new binding site for miR-891b (Figure 3a and 3b). We chose this
164 miRNA and another miRNA, miR-138-5p (Figure 3c) which targets EGFR and remains
165 unaffected by the presence of either major or minor allele for expression study (Figure S2b).
166 Both the miRNAs targets EGFR-A isoform, as confirmed by TargetScan and miRwalk tools
167 (Figure 3d). Gene Set Enrichment Analysis (GSEA) for miR-891b and miR-138-5p revealed that
168 they are involved in pathways related to cancer and MAPK signaling pathway, respectively
169 (Table S3 and S4).

170 **Correlation of EGFR-AS1 level with the miRNA level**

171 The expression levels of both miR-891b and miR-138-5p, were analyzed in 48 oral tumor tissues
172 and 8 adjacent normal tissues and were found to be significantly downregulated in tumor tissues
173 compared to that of normal tissues with a p-value of 0.03 and 0.047, respectively (Figure 4a &
174 4b). Tumors that over expressed miR-891b and miR-138-5p were found to have a significantly
175 low level of EGFR-AS1 ($p= 0.04$ and 0.02, respectively) (Figure 4c & 4d). The EGFR D/A
176 isoforms ratio showed no statistically significant difference between high and low miRNA

expression groups, but the tumors with overexpressed miRNAs showed a trend of increased D/A ratio (Figure S3a and 3b). When correlated with the expression level of EGFR-A, the tumors with an increased level of miR-891b had significantly low level of EGFR-A (p-value- 0.02) (Figure 4e) suggesting the possibility of ceRNA network operating in these tumors. Our results suggest that the presence of minor allele generates the binding site for miR-891b thereby switching the mechanism of EGFR-AS1 from a scaffold to miRNA sponge modulating the expression level of EGFR-A isoform.

Discussions

EGFR is found to be frequently mutated and deregulated in multiple cancers, especially malignancies of epithelial origin in humans. Thus, tyrosine kinase inhibitors have been increased attention towards targeted therapy in cancers[15-17]. Various studies focused on EGFR variants conferring sensitivity towards TKI treatments. Recently Tan *et al*., identified a synonymous variant *rs10251977* (c.2361G>A) present in exon 20 of *EGFR* having greater prognostic implication towards TKI treatments particularly in squamous cell carcinoma of head and neck[6]. In the current study the prevalence of this polymorphism is known to be 40% (72/180) of GG genotype, 47.2% (85/180) of GA genotype and 12.8% (23/180) of AA genotype among oral cancer patients with a minor allele frequency of 0.36. This is consistent with the previous report where they found increased prevalence of the same polymorphism in head and neck cancer patients from South Indian origin with 61.24% and 13.95% of both GA and AA genotypes, respectively[5]. They also observed that the polymorphism had a significant risk association in cancer development. Several studies emphasized the prognostic significance of this EGFR variant in TKI response[18-20], whereas the wild type genotype predicted to have a good prognosis in metastatic colorectal cancer[21]. Interestingly, Koh *et al*., reported that this polymorphism is more prevalent in squamous cell carcinoma compared to that of adenocarcinoma of lung, and having a significant response for gefitinib treatment in patients who carries a wild type allele[22]. Previous study on this variant in advanced oesophageal squamous cell carcinoma recorded that patients with heterozygous genotype shows a poor response[23]. However, we did not observe any significant association of this polymorphism with cancer risk. The increased prevalence of this variant and its strong association with therapeutic response prompted us to explore its functional significance. The variant allele created a new binding site for miR-891b reducing the level of EGFR-AS1 and leading to increase in the EGFR D/A isoform ratio. Using the online STITCH database, we have found that lncRNA EGFR-AS1 has a positive interaction with PTBP1 and

hnRNPs, major players in alternative splicing and essential for the generation of EGFR-A isoform with both transmembrane and tyrosine kinase domain.

Since the last decade, non-coding RNAs have been one of the main focuses in the field of functional genomics. Emerging evidence has shown the altered expression signatures of a large number of lncRNAs in several human malignancies[24]. In recent years, many of lncRNAs have been proposed to have pivotal role in carcinogenesis. However, the functions and mechanisms of lncRNAs responsible for the development and progression of oral cancer need to be understood. Natural antisense transcripts (NATs) are a group of RNA transcripts that are suggested to play roles in alternative splicing, genomic imprinting, miRNA sponging, also X- chromosome inactivation and they regulate the gene expression both in *cis*- and *trans*- manner[25-27]. *cis*-NAT plays major role in regulation of its sense partner in various manners. A myriad of lncRNA comes under this class of RNA and various studies reported their *cis* role. Recently, we reported a *cis*- mediated regulation of OIP5-AS1 controlling expression of *OIP5* gene by acting as miRNA sponge[28]. Various pan cancer analysis showed the role of deregulated NATs in carcinogenesis and acting as both predictive marker as well as drug targets[29].

EGFR-AS1, a lncRNA transcribed from the antisense strand of EGFR, was found to be overexpressed in several cancers[30,11,31-33]. Our study observed a significant upregulation of EGFR-AS1 in oral cancer patients suggesting the oncogenic potential of this lncRNA, which is consistent with above findings. When the EGFR-AS1 expression profile was analysed with reference to genotypes we found a decreased expression of EGFR in GG compared to GA+AA genotype, but without any statistical significance. The altered expression pattern of EGFR-AS1 might be due to previously reported non-canonical RNA editing mechanism which maintains the allele specific lncRNA based on germline polymorphism[34].

In this study, there was a significant change in the EGFR D:A isoform ratio with the low level expression of EGFR-AS1. Alternative splicing of receptor tyrosine kinases (RTKs) has greater clinical implication, where truncated RTKs arise by alternative splicing of exons or activation of intronic polyA sites[35]. These truncated forms lack transmembrane and tyrosine kinase domain and thus can be used as a good prognostic marker[6]. EGFR has three different soluble truncated isoforms such as 110 kDa, 80 kDa and 60 kDa where EGFR D form (110 kDa) is highly expressed in human placenta[15] and found to be enriched in urine of patients with squamous cell

carcinoma[36]. Several studies reported the prognostic values of these soluble EGFR in various cancers[37-40].

Our bioinformatic analysis revealed that premature termination of EGFR transcript results in the production of EGFR-D isoform by recognition of intronic polyA (IPA) sites and activating premature cleavage and polyadenylation. Additionally, our biomining approach discovered the RNA binding proteins such as PTBP1 and hnRNPU, which were known to play an important role in alternate splicing by significantly interacting with EGFR-AS1. The binding sites for these proteins were present in the skipped exons 15a and 15b, which resides between exons 15 and 16 of *EGFR*. Previous studies have shown that PTBP1 overexpression was known to function as splicing silencer of exon 3 of *BIM* gene by binding to its intron 2[41] and was also known to act as a splicing reprogrammer of *PKM* gene in pancreatic cancer[42]. Additionally, PTBP1 and PTBP2 are larger family of non-conserved cryptic exons repressors[43].

Previous studies have established that *SNRNPA* (U1snRNP) prevents the formation of truncated RTK isoforms by negatively controlling the use of IPA sites[35]. Surprisingly, STITCH analysis showed that PTBP1 has a mechanical interaction with SNRNPA which indicates the role of PTBP1 in regulating RTKs isoforms. Besides these several studies emphasized the role of chromatin signatures in alternative splicing Luco *et al*., reported the role of H3K36me3 in recruiting PTBP1 in *FGFR* exon 3b and promotes alternative splicing of exon 3c in mesenchymal cells[44]. Our analysis for chromatin signatures using UCSC genome browsers showed that *EGFR* exon 15a region is enriched with H3K36me3 mark supporting their role in alternative splicing through PTBP1. Earlier study reported that HuR binding proteins promote alternative splicing of NF1 and FAS transcripts by inducing localized histone hyperacetylation[45]. Recent finding identified EGFR-AS1 and HuR interaction suggesting its role in controlling alternative splicing via chromatin modification[32]. These observations collectively suggest the role of EGFR-AS1 in controlling the expression of EGFR-A from suppressing the formation of EGFR-D.

Certain lncRNAs were shown to competitively sponge the miRNAs and protects its protein coding RNA partners from miRNA mediated repression and modulate their stability. This ceRNA hypothesis is one such phenomenon by which NATs may modulate *cis* regulation in controlling the expression of its sense partner[46]. Our biomining revealed that this germline variant *rs10251977* in EGFR-AS1 regulates the stability of the lncRNA by miRNA mediated

degradation. The minor allele generates a new binding site for miR-891b and resulting in diminished level of EGFR-AS1. This is evident from our results that the decreased level of EGFR-AS1 in tumors that expressed high level of miR-891b. Interestingly, we also found that miR-891b binding site was found in 3'UTR of EGFR-A isoform. We also analyzed another miRNA, miR-138-5p whose binding site is present in both EGFR-AS1 and EGFR-A isoform and its high expression level correlated with reduced level of EGFR-AS1 and EGFR-A. Our study showed both miRNAs were tumor suppressors and downregulated in tumor samples. In addition to support the ceRNA function, recent studies have reported miRNA sponging activity of EGFR-AS1[47,33] and miR-891b mediated gene regulation[48-50].

The above findings suggest that EGFR-AS1, promotes tumorigenesis by increasing the level of EGFR-A through the modulation of alternative splicing by preventing the premature termination through PTBP1. The germline polymorphism *rs10251977* (G>A) in EGFR-AS1 creates binding site for miR-891b which decreases the level of EGFR-AS1 and thereby increasing the EGFR-D/A ratio which confers the susceptibility in tumor by promoting the EGFR-AS1 to act as miRNA sponge. These findings provide an insight on functional importance of germline polymorphism *rs10251977* (G>A) in EGFR-AS1 and orchestrated interactions between non-coding RNAs and coding RNA in maintaining cellular homeostasis and its dysregulation may lead to development of cancer.

Methods and materials

Clinical specimens

The present study was conducted after the approval by the Institutional Ethics Committees (IEC) of Madras Medical College, Chennai (No.04092010) and Government Arignar Anna Memorial Cancer Hospital, Kancheepuram (No.101041/e1/2009-2) and also within the ethical framework of Dr. ALM PG Institute of Basic Medical Sciences, Chennai. The study includes a total of 180 participants diagnosed with OSCC and 184 healthy controls. Blood samples were collected to study the germline variant in *EGFR*. Forty-eight tumor tissue samples and eight normal tissues from individuals free from cancer were collected for expression studies. The patient's contextual and clinicopathological characteristics were documented with standard questionnaire following the IEC guidelines and written informed consent was obtained from each patient, after explaining

about the research study. The tissues were collected in RNAlater solution (Ambion, USA) and transported to the laboratory in cold-storage and stored at -80^{o} C.

DNA isolation

The blood samples were subjected for DNA isolation using Proteinase K digestion and PCI extraction method. The isolated DNA was quantified using NanoDrop2000 UV-Vis spectrophotometer (Thermo Scientific, USA) and its integrity was assessed by resolving in 1% agarose gel in Mupid gel electrophoresis (TaKaRa, Japan) and diluted to 100ng/µL stored at -20^{o}C which was later used for screening germline variants.

PCR amplification and sequencing of germline variant *rs10251977* in EGFR

The exon 20 of *EGFR* gene harbouring the SNP was amplified from 364 DNA samples (180 oral cancer patients and 184 healthy controls) using the sequence specific primers EGFR Fwd:5′- M13- CCTTCTGGCCACCATGCGAA-3′ and Rev:5′-CGCATGTGAGGATCC TGGCT -3′ (where M13 in the EGFR Fwd is the universal sequencing primer 5′ - tgtaaaacgacggccagt -3′). Polymerase chain reaction (PCR) was carried in 30µL volume containing 100ng of genomic DNA, 1X PCR buffer, 1.5mM $MgCl_2$, 100µM dNTP (Takara, Japan), 80nM primers (Sigma India) and 0.5U of Takara Taq polymerase (Takara, Japan). The PCR is performed using the following thermal conditions: 94ºC for 2 min for initial denaturation, followed by 40 cycles of 94ºC for 30 sec, 60ºC for 30 sec and 72ºC for 30 sec, and 7 min of final extension at 72ºC. The PCR products (298 bp) were electrophoresed in 2 % agarose gel. The purified PCR products were sequenced at Macrogen Inc, South Korea.

RNA isolation and quality control

The tumor samples soaked in RNA later solutions and frozen in -80^{o}C were thawed on ice, washed twice with 1x ice cold PBS, homogenized with zirconium beads using MicroSmash MS-100 automated homogenizer (Tomy Digital Biology, Japan). The total RNA was isolated using the RNAeasy mini kit (Qiagen, Germany) following the manufacturer's protocol. RNA was quantified using NanoDrop2000 UV-Vis spectrophotometer (Thermo Scientific, USA) and the RNA integrity was assessed by resolving in DEPC treated 1% agarose gel in Mupid gel electrophoresis (TaKaRa, Japan).

cDNA synthesis and quantitative Real-Time PCR

The cDNA conversion was carried out from total RNA using custom designed miRNA seed specific stem-loop primers for miRNAs (Table S5) and a random hexamer primer for coding/ non-coding RNAs, with 2µg of RNA for mRNA and lncRNAs and 10ng of RNA for miRNAs. The RNA samples were pre-incubated at 65˚C for 20 min followed by 55˚C for 90 min, 72˚C for 15 min and final hold at 4˚C. cDNA conversion was performed using a Reverse Transcription kit (Invitrogen; Thermo Fisher Scientific, Inc. USA) and cDNAs were further diluted 25-fold and stored at -20˚C.

Real-time RT-qPCR was performed in ABI Quantstudio 6 Flex (ABI Lifetechnology, USA). The reactions were performed in 384 well optical plates with 10µl total volume using 1µl of cDNAs as template, TaqMan® 2X Universal Master mix (No AmpErase® UNG; Thermo Fisher Scientific, Inc. USA), specific forward primers (Table S6), universal reverse primer 5'-TCGTATCCAGTGCGTCGA GT -3', and fluorescein amidite-labeled minor groove binder probes 5'-CAGAGCCACCTGGGCAATTTT-3' for miRNAs expression. The EGFR A/D isoforms and EGFR-AS1 expression levels were analyzed using SYBR-Green master mix (KAPA SYBR FAST qPCR Kits, USA) with the gene specific primers (Table S7), by following the thermal cycling conditions: 50˚C for 2 min and 95˚C for 10 min once, followed by 95˚C for 15 sec and 60˚C for 1 min for 40 cycles. GAPDH served as an endogenous control for coding and non-coding genes, and RNU44 as an endogenous control for miRNA. NTC reactions were used in all the experiments. All the reactions were carried out in triplicates, mean Ct was used for analysis and the expression level was calculated using $2^{-\Delta\Delta Ct}$ method.

In-silico approaches

To predict the proteins interacting with EGFR-AS1, lncRNAtor[12] an online database was used and consensus motifs were predicted using the RBPMap[13] to identify the RNA binding partners and motifs of interest such as PTBP1 in both EGFR-AS1 and skipped exons of *EGFR*. The motifs with highest stringency were selected for the study. The STITCH online tool was used to collect the molecular partners for PTBP1. UCSC genome browser was used to visualize the epigenetic signal that promotes alternative splicing and also polyA sites to predict premature termination. To assess the functional impact of the polymorphism *rs10251977* in EGFR-AS1, lncRNASNP2 tool was used to identify the loss or creation of binding site for miRNA in the lncRNA or the structural impact on lncRNA was analysed. Gene targets of the miRNA were

identified by Targetscan and MiRwalk online tools. Gene Set Enrichment Analysis (GSEA) done by MiRwalk web interface to understand the molecular pathways affected by the selected miRNAs.

Statistical analysis

The frequency of alleles and genotypes were compared between the patient's and control groups by chi-square test. Fisher exact test and odds ratio (OR) with 95% confidence interval (CI) was calculated to find the risk association. Differences between the means were presented as mean ± SEM and analysed using Student's *t*-test (Mann-Whitney) using Graph Pad Prism statistical software, v 6.01 (Graph Pad software Inc, USA). All tests were two-tailed and a p value of <0.05 was considered as statistically significant.

References:

1 Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians* **68**, 394-424 (2018).

2 Bose, P., Brockton, N. T. & Dort, J. C. Head and neck cancer: from anatomy to biology. *International journal of cancer* **133**, 2013-2023 (2013).

3 Arunkumar, G. *et al.* Expression profiling of long non-coding RNA identifies linc-RoR as a prognostic biomarker in oral cancer. *Tumor Biology* **39**, 1010428317698366 (2017).

4 Elferink, L. A. & Resto, V. A. Receptor-tyrosine-kinase-targeted therapies for head and neck cancer. *Journal of signal transduction* **2011** (2011).

5 Nagalakshmi, K., Jamil, K., Pingali, U., Reddy, M. V. & Attili, S. S. Epidermal growth factor receptor (EGFR) mutations as biomarker for head and neck squamous cell carcinomas (HNSCC). *Biomarkers* **19**, 198-206 (2014).

6 Tan, D. S. *et al.* Long noncoding RNA EGFR-AS1 mediates epidermal growth factor receptor addiction and modulates treatment response in squamous cell carcinoma. *Nature medicine* **23**, 1167 (2017).

7 Sawicka, K., Bushell, M., Spriggs, K. A. & Willis, A. E. Polypyrimidine-tract-binding protein: a multifunctional RNA-binding protein. *Biochemical Society Transactions* **36**, 641-647 (2008).

8 Lin, J.-C., Tsao, M.-F. & Lin, Y.-J. Differential impacts of alternative splicing networks on apoptosis. *International journal of molecular sciences* **17**, 2097 (2016).

9 Albitar, L. *et al.* EGFR isoforms and gene regulation in human endometrial cancer cells. *Molecular cancer* **9**, 1-13 (2010).

10 Weinholdt, C. *et al.* Prediction of regulatory targets of alternative isoforms of the epidermal growth factor receptor in a glioblastoma cell line. *BMC bioinformatics* **20**, 434 (2019).

11 Hu, J. *et al.* Long noncoding RNA EGFR-AS1 promotes cell proliferation by increasing EGFR mRNA stability in gastric cancer. *Cellular Physiology and Biochemistry* **49**, 322-334 (2018).

12 Park, C., Yu, N., Choi, I., Kim, W. & Lee, S. lncRNAtor: a comprehensive resource for functional investigation of long non-coding RNAs. *Bioinformatics* **30**, 2480-2485 (2014).

13 Paz, I., Kosti, I., Ares Jr, M., Cline, M. & Mandel-Gutfreund, Y. RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucleic acids research* **42**, W361-W367 (2014).

14 Brown, S. J., Stoilov, P. & Xing, Y. Chromatin and epigenetic regulation of pre-mRNA processing. *Human molecular genetics* **21**, R90-R96 (2012).

15 Abou-Fayçal, C., Hatat, A.-S., Gazzeri, S. & Eymin, B. Splice variants of the RTK family: their role in tumour progression and response to targeted therapy. *International journal of molecular sciences* **18**, 383 (2017).

16 Schlessinger, J. Receptor tyrosine kinases: legacy of the first two decades. *Cold Spring Harbor perspectives in biology* **6**, a008912 (2014).

17 Sigismund, S., Avanzato, D. & Lanzetti, L. Emerging functions of the EGFR in cancer. *Molecular oncology* **12**, 3-20 (2018).

18 Mu, X. L. *et al.* Gefitinib-sensitive mutations of the epidermal growth factor receptor tyrosine kinase domain in Chinese patients with non–small cell lung cancer. *Clinical cancer research* **11**, 4289-4294 (2005).

19 Sasaki, H. *et al.* EGFR polymorphism of the kinase domain in Japanese lung cancer. *Journal of Surgical Research* **148**, 260-263 (2008).

20 Leichsenring, J. *et al.* Synonymous EGFR variant p. Q787Q is neither prognostic nor predictive in patients with lung adenocarcinoma. *Genes, Chromosomes and Cancer* **56**, 214-220 (2017).

21 Bonin, S. *et al.* A synonymous EGFR polymorphism predicting responsiveness to anti-EGFR therapy in metastatic colorectal cancer patients. *Tumor Biology* **37**, 7295-7303 (2016).

22 Koh, Y. W. *et al.* Q787Q EGFR polymorphism as a prognostic factor for lung squamous cell carcinoma. *Oncology* **90**, 289-298 (2016).

23 Wang, W.-P., Wang, K.-N., Gao, Q. & Chen, L.-Q. Lack of EGFR mutations benefiting gefitinib treatment in adenocarcinoma of esophagogastric junction. *World Journal of Surgical Oncology* **10**, 14 (2012).

24 Bolha, L., Ravnik-Glavač, M. & Glavač, D. Long noncoding RNAs as biomarkers in cancer. *Disease markers* **2017** (2017).

25 He, Y., Vogelstein, B., Velculescu, V. E., Papadopoulos, N. & Kinzler, K. W. The antisense transcriptomes of human cells. *Science* **322**, 1855-1857 (2008).

26 Lindsay, M. A., Griffiths-Jones, S., Wight, M. & Werner, A. The functions of natural antisense transcripts. *Essays in biochemistry* **54**, 91-101 (2013).

27 Wenric, S. *et al.* Transcriptome-wide analysis of natural antisense transcripts shows their potential role in breast cancer. *Scientific reports* **7**, 1-12 (2017).

28 Arunkumar, G. *et al.* LncRNA OIP5-AS1 is overexpressed in undifferentiated oral tumors and integrated analysis identifies as a downstream effector of stemness-associated transcription factors. *Scientific reports* **8**, 1-13 (2018).

29 Chiu, H.-S. *et al.* Pan-cancer analysis of lncRNA regulation supports their targeting of cancer genes in each tumor context. *Cell reports* **23**, 297-312. e212 (2018).

30 Qi, H.-l. *et al.* The long noncoding RNA, EGFR-AS1, a target of GHR, increases the expression of EGFR in hepatocellular carcinoma. *Tumor Biology* **37**, 1079-1089 (2016).

31 Dong, Z.-Q., Guo, Z.-Y. & Xie, J. The lncRNA EGFR-AS1 is linked to migration, invasion and apoptosis in glioma cells by targeting miR-133b/RACK1. *Biomedicine & Pharmacotherapy* **118**, 109292 (2019).

32 Wang, A. *et al.* Long noncoding RNA EGFR-AS1 promotes cell growth and metastasis via affecting HuR mediated mRNA stability of EGFR in renal cancer. *Cell death & disease* **10**, 1-14 (2019).

33 Xu, Y.-H., Tu, J.-R., Zhao, T.-T., Xie, S.-G. & Tang, S.-B. Overexpression of lncRNA EGFR-AS1 is associated with a poor prognosis and promotes chemotherapy resistance in non-small cell lung cancer. *International journal of oncology* **54**, 295-305 (2019).

34 Shah, M. Y. *et al.* Cancer-associated rs6983267 SNP and its accompanying long noncoding RNA CCAT2 induce myeloid malignancies via unique SNP-specific RNA mutations. *Genome research* **28**, 432-447 (2018).

35 Vorlová, S. *et al.* Induction of antagonistic soluble decoy receptor tyrosine kinases by intronic polyA activation. *Molecular cell* **43**, 927-939 (2011).

36 Ilekis, J. V., Gariti, J., Niederberger, C. & Scoccia, B. Expression of a truncated epidermal growth factor receptor-like protein (TEGFR) in ovarian cancer. *Gynecologic oncology* **65**, 36-41 (1997).

37 Jantus-Lewintre, E. *et al.* Analysis of the Prognostic Value of Soluble Epidermal Growth Factor Receptor Plasma Concentration in Advanced Non–Small-Cell Lung Cancer Patients. *Clinical lung cancer* **12**, 320-327 (2011).

38 Guillaudeau, A. *et al.* EGFR soluble isoforms and their transcripts are expressed in meningiomas. *PLoS One* **7**, e37204 (2012).

39 Zampino, M. *et al.* Epidermal growth factor receptor serum (sEGFR) level may predict response in patients with EGFR positive advanced colorectal cancer treated with gefitinib? *Cancer chemotherapy and pharmacology* **63**, 139-148 (2008).

40 MÜLLER, V. *et al.* Prognostic and predictive impact of soluble epidermal growth factor receptor (sEGFR) protein in the serum of patients treated with chemotherapy for metastatic breast cancer. *Anticancer research* **26**, 1479-1487 (2006).

41 Juan, W. C., Roca, X. & Ong, S. T. Identification of cis-acting elements and splicing factors involved in the regulation of BIM Pre-mRNA splicing. *PLoS One* **9**, e95210 (2014).

42 Calabretta, S. *et al.* Modulation of PKM alternative splicing by PTBP1 promotes gemcitabine resistance in pancreatic cancer cells. *Oncogene* **35**, 2031-2039 (2016).

43 Ling, J. P. *et al.* PTBP1 and PTBP2 repress nonconserved cryptic exons. *Cell reports* **17**, 104-113 (2016).

44 Luco, R. F. *et al.* Regulation of alternative splicing by histone modifications. *Science* **327**, 996-1000 (2010).

45 Zhou, H.-L. *et al.* Hu proteins regulate alternative splicing by inducing localized histone hyperacetylation in an RNA-dependent manner. *Proceedings of the National Academy of Sciences* **108**, E627-E635 (2011).

46 Kulkarni, A., Anderson, A. G., Merullo, D. P. & Konopka, G. Beyond bulk: a review of single cell transcriptomics methodologies and applications. *Current opinion in biotechnology* **58**, 129-136 (2019).

47 Feng, Z. *et al.* LncRNA EGFR-AS1 upregulates ROCK1 by sponging miR-145 to promote Esophageal squamous cell carcinoma cell invasion and migration. *Cancer Biotherapy & Radiopharmaceuticals* **35**, 66-71 (2020).

48 Belleannée, C. *et al.* Role of microRNAs in controlling gene expression in different segments of the human epididymis. *PloS one* **7**, e34996 (2012).

49 Dong, Q. *et al.* MicroRNA-891b is an independent prognostic factor of pancreatic cancer by targeting Cbl-b to suppress the growth of pancreatic cancer cells. *Oncotarget* **7**, 82338 (2016).

50 Xu, S. *et al.* Down-regulation of PARP1 by miR-891b sensitizes human breast cancer cells to alkylating chemotherapeutic drugs. *Archives of gynecology and obstetrics* **296**, 543-549 (2017).

Acknowledgement

We were very grateful for the patients agreeing to provide clinical samples for this study. SD and MMR acknowledge the research fellowship from University Grant Commission (UGC), Govt. of India. The study was supported by grants DHR- starter grant (No. V.25011/536-HRD/2016-HR), DHR, Government of India and NIG-JOINT(1A2019), Japan to Dr. A. K. Munirajan. We also gratefully acknowledge the DST-FIST, UGC-SAP and DHR-MRU infrastructural facility.

Authors Contributions

AKM (corresponding author) designed the study, supervised the experiments, data analysis and reviewed the manuscript. SD and MMR participated in study design, performed all the experiments, analyzed the data and wrote the manuscript. RCS assisted in study design, analysed the data and reviewed the manuscript. KVUD assisted the real-time PCR experiments and bioinformatics analysis. RA & SS provided tumor samples and clinical data. II critically reviewed the manuscript. All authors read and approved the final manuscript.

Competing interests

The author declared that there are no competing interests.

Figure legends

Figure 1 Role of EGFR-AS1 in maintaining the level of EGFR A and D isoforms in OSCC
a. Relative expression level of lncRNA EGFR-AS1 in oral tumors compared with normal tissues. **b.** Graph showing the relative fold change of EGFR A and D isoforms in relation to EGFR-AS1 high and low expression groups. **c.** Graph shows the EGFR D/A ratio in relation to EGFR high and low expression groups. (Statistical significance represented as p-value **** for < 0.0001, *** for <0.001, * for < 0.01, ns for non-significant, two tailed Student's t-test).

Figure 2 *In-silico* characterization of EGFR-AS1 role in alternative splicing of EGFR
a. STITCH analysis showing the binding partners of PTBP1. **b.** UCSC genome browser data shows the H3K36me3 marks (green box) and polyA sites (orange box) prevalence between exon 15 and 16 skipped region (blue box) of EGFR.

Figure 3 *rs10251977* minor allele's effect on EGFR-AS1
a. Absence of minor allele in EGFR-AS1 (Blue arrow). **b.** Prediction of effect of minor allele in EGFR-AS1 generating binding site for miR-891b using lncRNASNP2 (Red arrow indicating presence of variant allele). **c.** Figure showing the binding site for another miRNA, miR-138-5p in EGFR-AS1. **d.** miRWalk database showing the targets of both miR-891b and miR-138-5p (EGFR encircled in both miRNAs).

Figure 4 miR-891b sponging by EGFR-AS1 promoting the expression of EGFR-A isoform
a. Graph showing the relative fold change of miR-891b. **b.** Graph showing the relative fold change of miR-138-5p. **c.** Graph shows the relative fold change of EGFR-AS1 in relation to miR-891b. **d.** Graph showing the relative fold change of EGFR-AS1 in relation to miR-138-5p levels. **e.** Graph shows the relative fold change of EGFR-AS1 in relation to miR-891b. (Statistical significance represented as * for P < 0.01, two tailed Student's t-test).

Tables

Table 1 Demographic and clinical data of Oral cancer patients and healthy controls

Table 2 Genotype and allele frequency in South Indian oral cancer patients and controls

Table 3 Relationship between EGFR-AS1 expression and clinicopathological characteristics in oral cancer patients

Supplementary material

Figure S1

a. Graph showing the relative fold change of EGFR-AS1 in relation to clinicopathological characteristics. **b.** Relative fold change of EGFR-AS1 in patients with GG and GA+ AA genotype of *rs10251977*.

Figure S2

a. Predicted EGFR-AS1 interacting proteins using online database lncRNAtor.

b. LncRNASNP2 showing the miRNA target sites, with miR-138-5p (within box) is targeted independent of the variant *rs10251977*.

Figure S3

a. Graph showing EGFR D/A ratio level in relation to miR-891b levels.

b. Graph showing EGFR D/A ratio level in relation to miR-138-5p levels.

Table S1 Table showing the prevalence of binding motif of PTBP1 in EGFR-AS1 using RBPmap online tool

Table S2 Table showing the prevalence of binding motif of PTBP1 in EGFR 15 and 16 exons using RBPmap online tool

Table S3 miR-891b miRWalk_GSEA_results (Excel file)

Table S4 miR-138-5p miRWalk_GSEA_results (Excel file)

Table S5 List of Universal reverse transcription primers used for cDNA synthesis

Table S6 List of gene specific forward primers used for real time PCR experiments

Table S7 List of primers used for SYBR Green gene expression assays

Figure.1

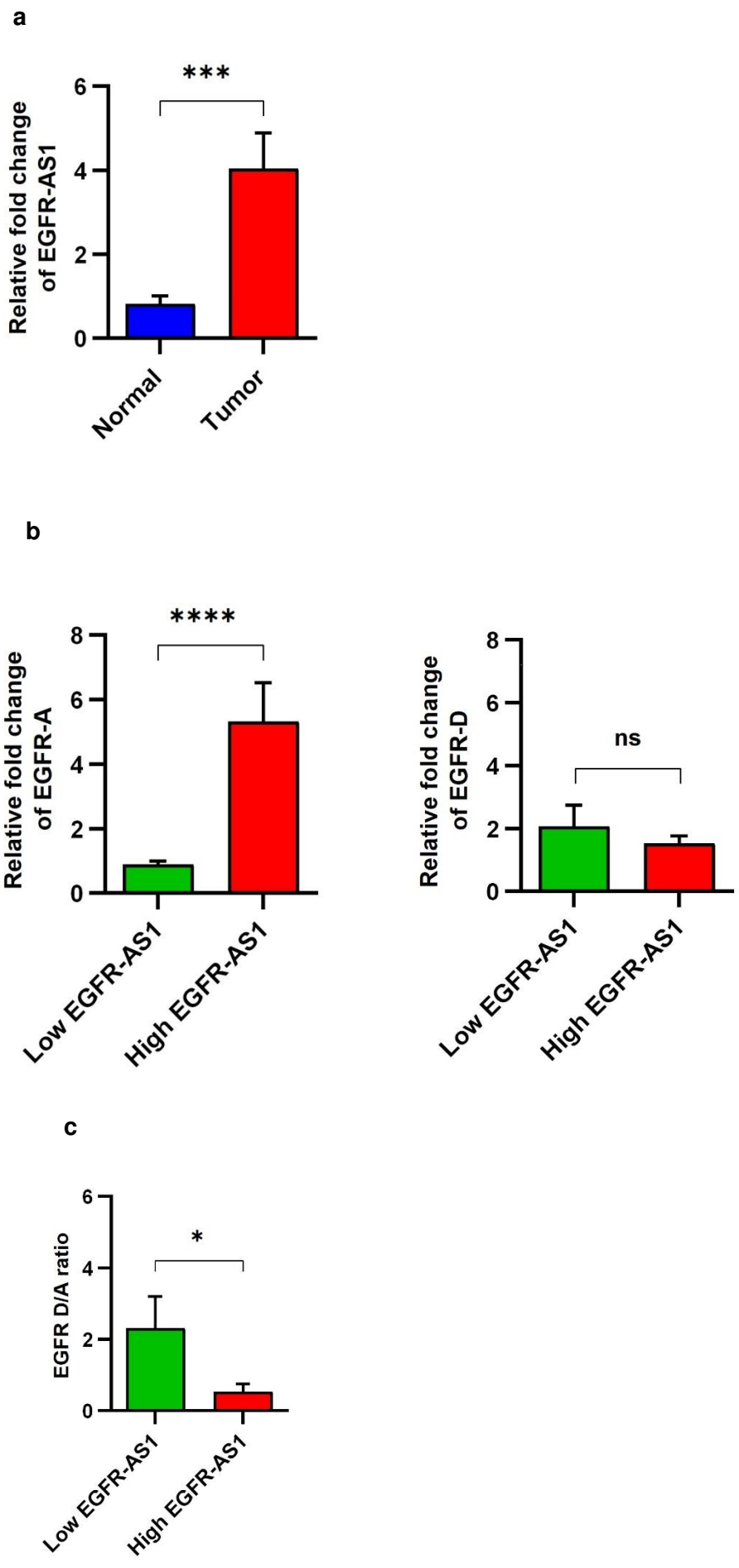
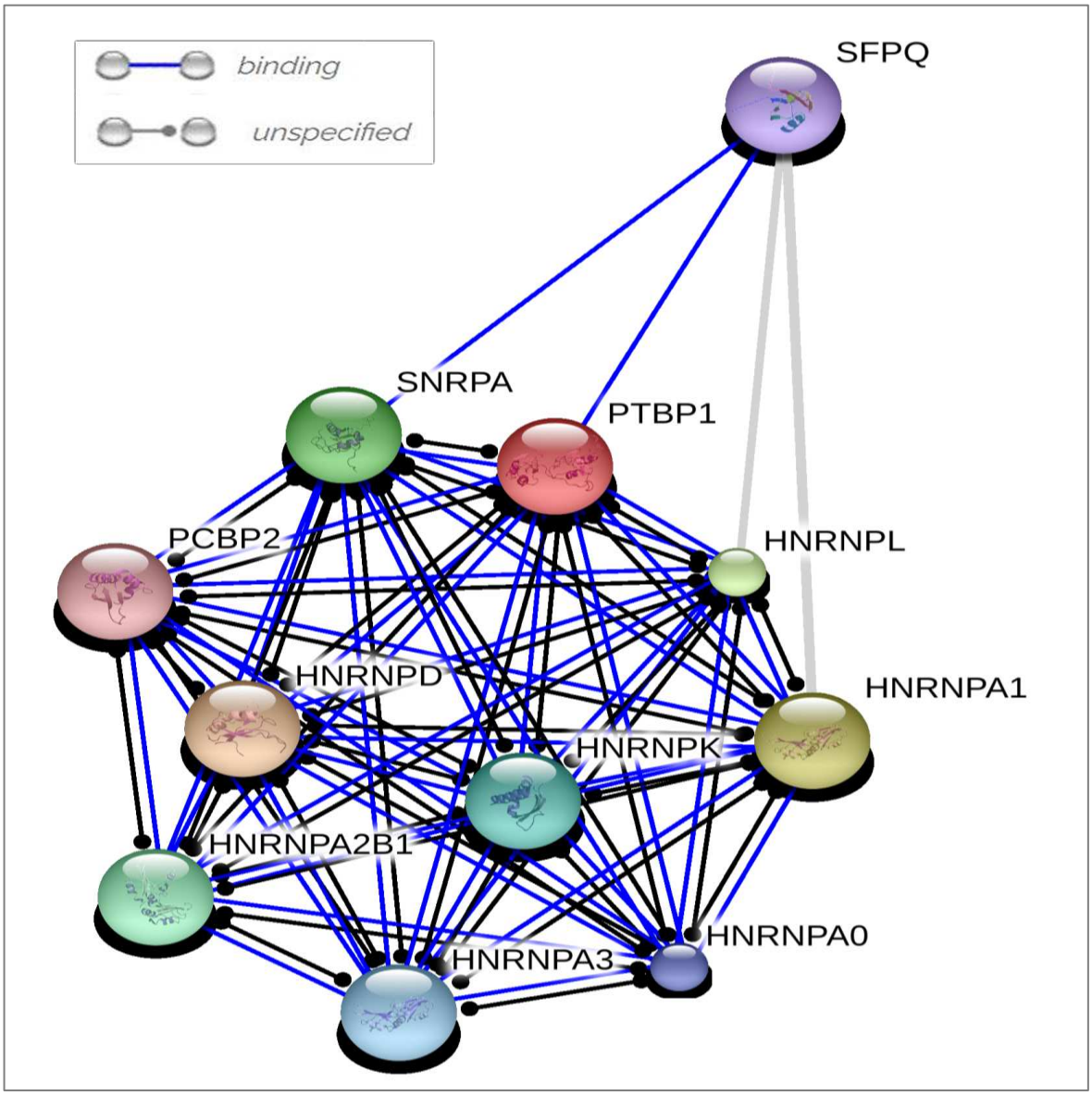

a
***
Relative fold change of EGFR-AS1
6
4
2
0
Normal
Tumor
b
****
Relative fold change of EGFR-A
8
6
4
2
0
Low EGFR-AS1
High EGFR-AS1
ns
Relative fold change of EGFR-D
8
6
4
2
0
Low EGFR-AS1
High EGFR-AS1
c
*
EGFR D/A ratio
6
4
2
0
Low EGFR-AS1
High EGFR-AS1

Figure.2

a


binding
unspecified
SFPQ
SNRPA
PTBP1
PCBP2
HNRNPL
HNRNPD
HNRNPA1
HNRNPK
HNRNPA2B1
HNRNPA0
HNRNPA3

b

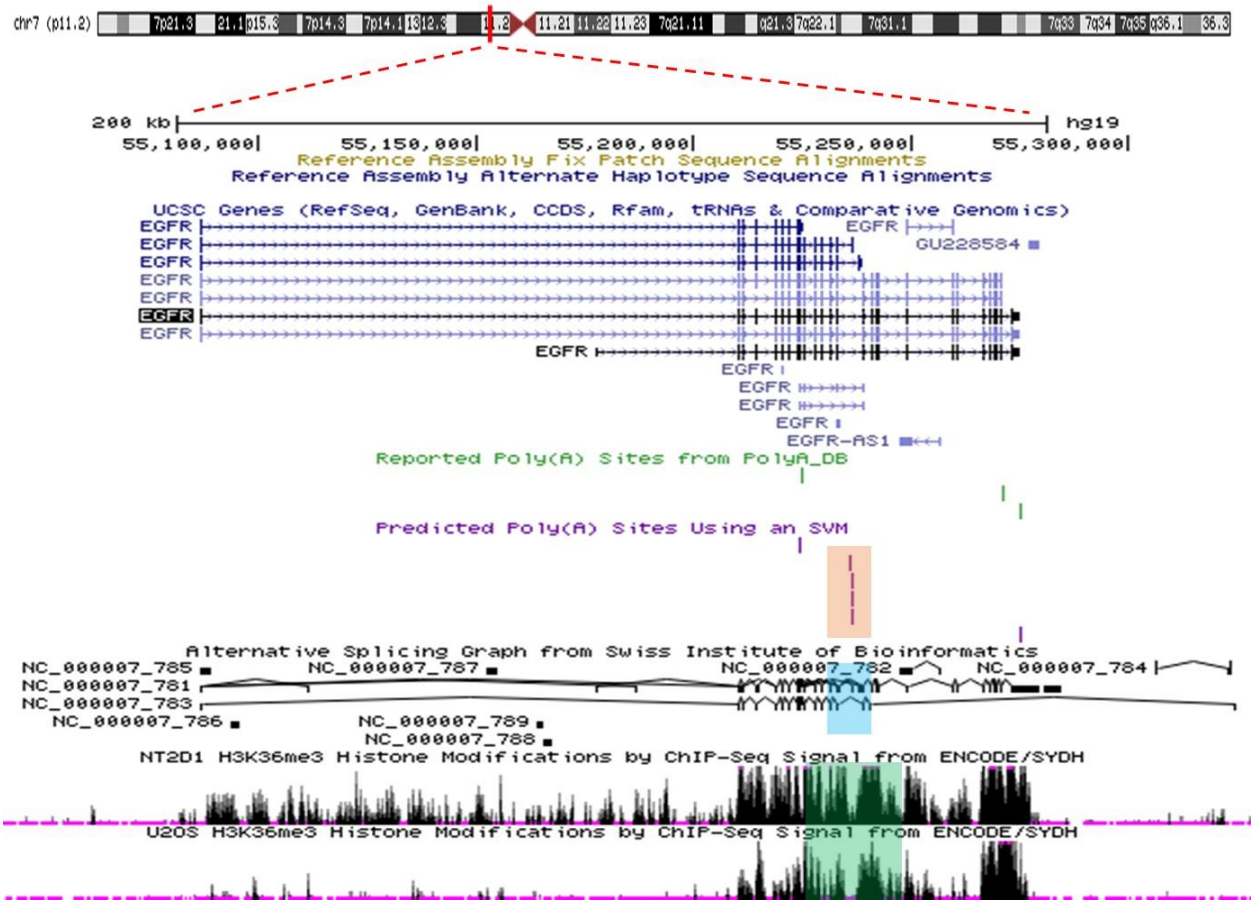

chr7 (p11.2)
200 kb
hg19
55,100,000
55,150,000
55,200,000
55,250,000
55,300,000
Reference Assembly Fix Patch Sequence Alignments
Reference Assembly Alternate Haplotype Sequence Alignments
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
EGFR
GU228584
EGFR-AS1
Reported Poly(A) Sites from PolyA_DB
Predicted Poly(A) Sites Using an SVM
Alternative Splicing Graph from Swiss Institute of Bioinformatics
NC_000007_785
NC_000007_787
NC_000007_782
NC_000007_784
NC_000007_781
NC_000007_783
NC_000007_786
NC_000007_789
NC_000007_788
NT2D1 H3K36me3 Histone Modifications by ChIP-Seq Signal from ENCODE/SYDH
U2OS H3K36me3 Histone Modifications by ChIP-Seq Signal from ENCODE/SYDH

Figure.3

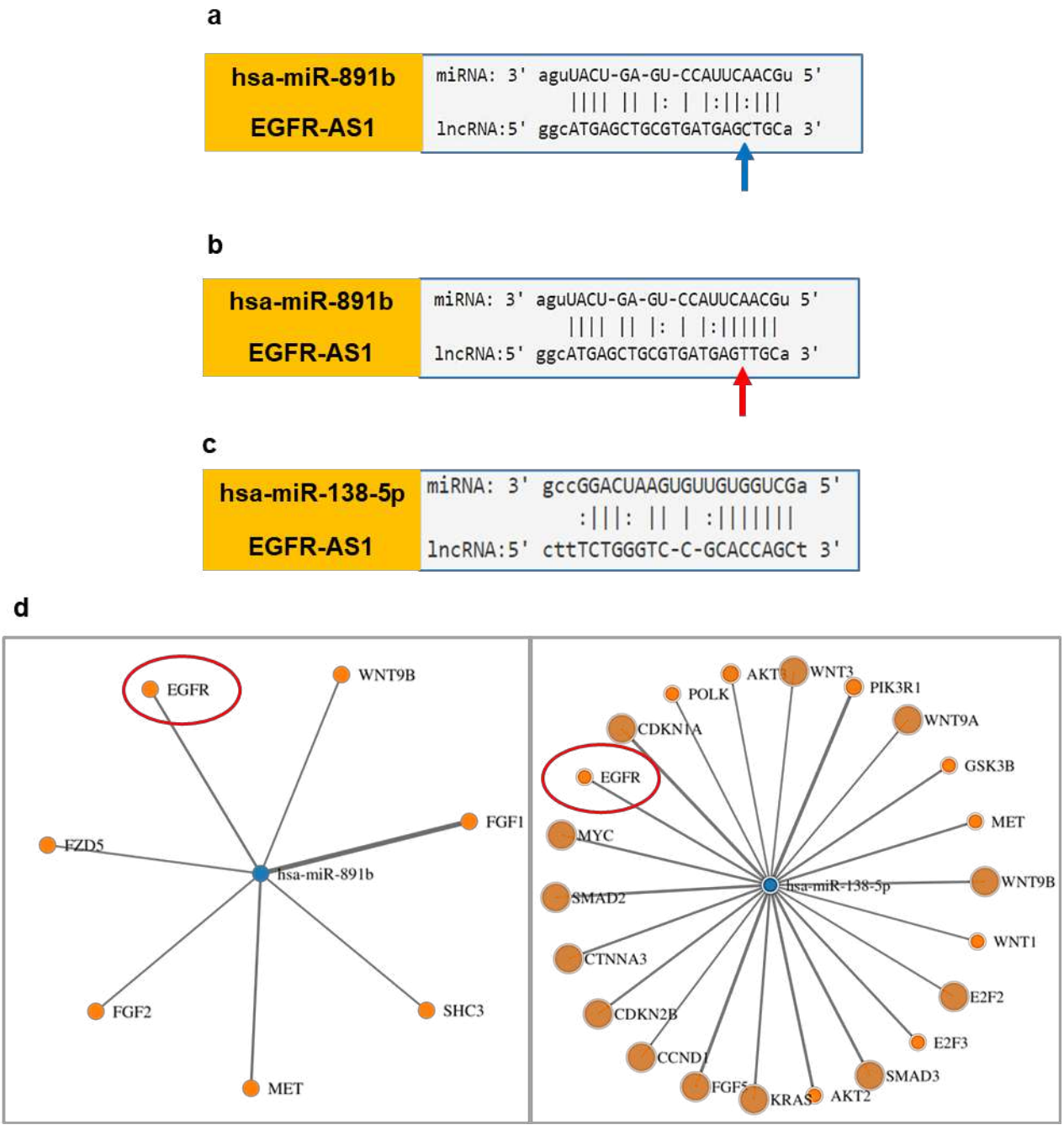
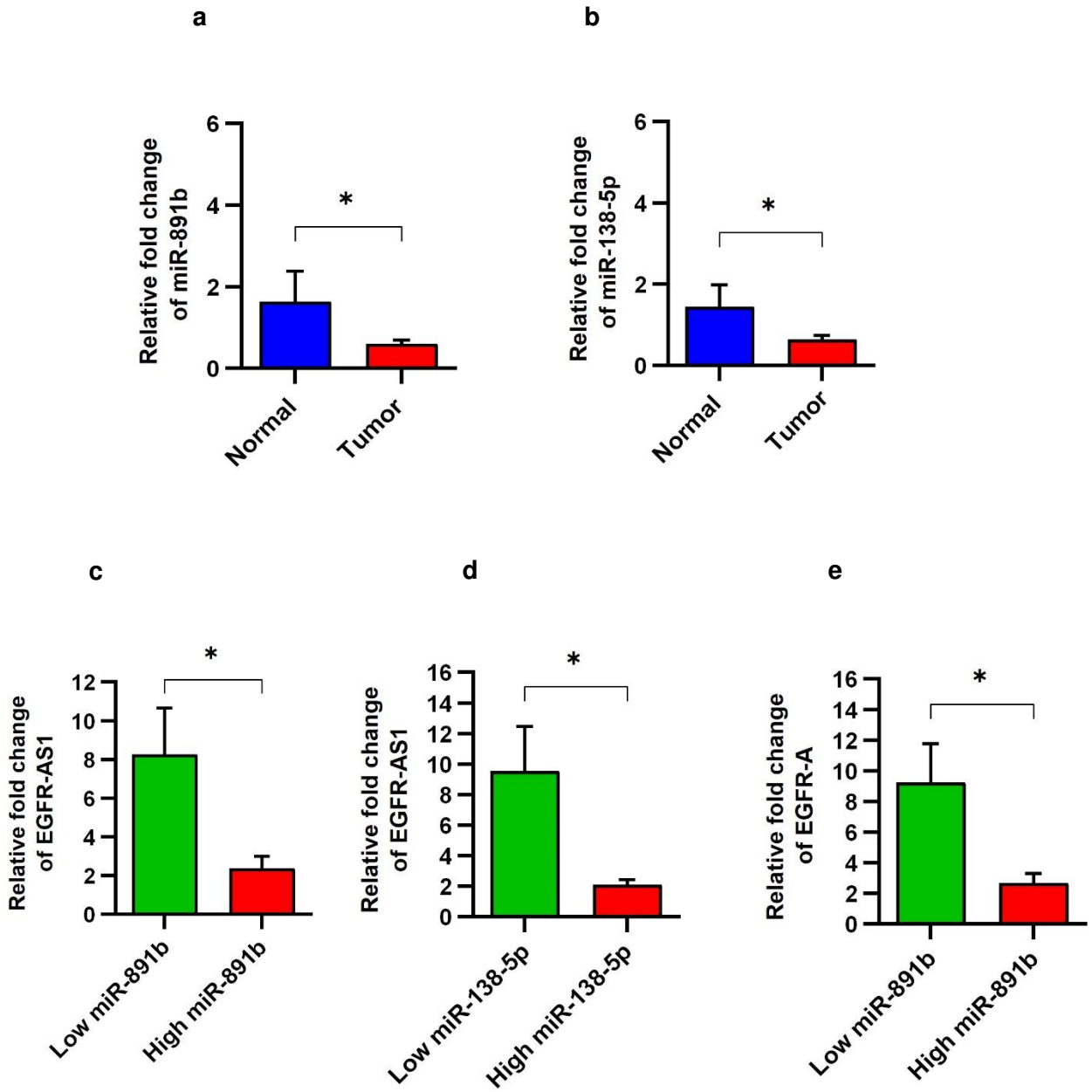

a
hsa-miR-891b
EGFR-AS1
miRNA: 3' aguUACU-GA-GU-CCAUUCAACGu 5'
lncRNA:5' ggcATGAGCTGCGTGATGAGCTGCa 3'
b
hsa-miR-891b
EGFR-AS1
miRNA: 3' aguUACU-GA-GU-CCAUUCAACGu 5'
lncRNA:5' ggcATGAGCTGCGTGATGAGTTGCa 3'
c
hsa-miR-138-5p
EGFR-AS1
miRNA: 3' gccGGACUAAGUGUUGUGGUCGa 5'
lncRNA:5' cttTCTGGGTC-C-GCACCAGCt 3'
d
EGFR
WNT9B
FGF1
FZD5
hsa-miR-891b
FGF2
SHC3
MET
AKT3
WNT3
PIK3R1
POLK
WNT9A
CDKN1A
GSK3B
EGFR
MET
MYC
hsa-miR-138-5p
WNT9B
SMAD2
WNT1
CTNNA3
E2F2
CDKN2B
E2F3
CCND1
SMAD3
FGF5
KRAS
AKT2

Figure.4


a
b
c
d
e
Relative fold change of miR-891b
Relative fold change of miR-138-5p
Relative fold change of EGFR-AS1
Relative fold change of EGFR-AS1
Relative fold change of EGFR-A
Normal
Tumor
Low miR-891b
High miR-891b
Low miR-138-5p
High miR-138-5p
Low miR-891b
High miR-891b
*

Tables

Table 1 Demographic and clinical data of Oral cancer patients and healthy controls

Clinical parameters	Oral cancer patients (n = 180)	Controls (n = 184)
Age in years, Mean ± SD	53.1 ± 11.0 (26 – 80)	50.4 ± 9.3 (25 – 82)
Sex (Male / Female)	124/56	125/59
Risk habits N (%)		
Smoking	96 (53.3)	9 (4.9)
Chewing	75 (41.6)	-
Alcoholics	46 (25.5)	4 (2.2)
Smoking + chewing	25 (13.8)	-
Smoking + alcoholics	16 (8.8)	55 (29.9)
Smoking + chewing +alcoholics	19 (10.5)	-
No risk habits	42 (23.3)	116 (63)
Histological differentiation N (%)		
Poor	26 (14.4)	-
Moderate	77 (42.8)	-
Well	77 (42.8)	-
Tumor Stage N (%)		
≤T2	34 (18.9)	-
>T2	146 (81.1)	-
Nodal invasion N (%)		
Positive	150 (83.3)	-
Negative	30 (16.7)	-

Table2 Genotype, Allele frequency in South Indian oral cancer patients and controls

rs10251977; c.2361G>A		Oral Cancer N (%)	Control N (%)	Odds ratio	95% CI	p-value
Genotype	GG	72 (40)	67 (36.4)	1 (Ref)		
	GA	85 (47.2)	93 (50.5)	1.17	0.75-1.83	0.47
	AA	23 (12.8)	24 (13.1)	1.12	0.57-2.17	0.73
Dominant model	GA+AA	108 (60)	117 (63.6)	0.86	0.56-1.31	0.55
Recessive model	GG+GA	157 (87.2)	160 (87)	0.98	0.53-1.80	0.93
Allelic model	G	229 (63.6)	227 (61.7)	1 (Ref)		
	A	131 (36.4))	141 (38.3)	0.92	0.68-1.24	0.64
HWE $\chi 2$		0.072	0.88			
HWE p-value		0.79	0.35			

Table 3 Relationship between EGFR-AS1 expression and clinicopathological characteristics in oral cancer patients. #Fisher's exact test with two tailed P value

Clinical characteristics	EGFR-AS1 expression level		p-value[#]	OR	95% CI
	Low (N=24)	High (N=24)			
Age (Years)					
≤51	11	14	0.5634	0.6044	0.1929-1.893
>51	13	10			
Sex					
Female	6	5	0.999	1.267	0.3281-4.891
Male	18	19			
Tobacco habits					
Smoking					
Yes	17	14	0.546	1.735	0.5238-5.745
No	7	10			
Chewing					
Yes	10	10	0.7697	1	0.3173-3.152
No	14	14			
Histological Differentiation					
Poor and moderate	17	14	0.546	1.735	0.5238-5.745
Well	7	10			
Tumor Stage					
≤T2	3	8	0.831	0.8571	0.1736-4.233
>T2	21	16			
Nodal invasion					
Positive	22	20	0.1905	2.2	0.3627-13.34
Negative	2	4			

Figures

Figure.1

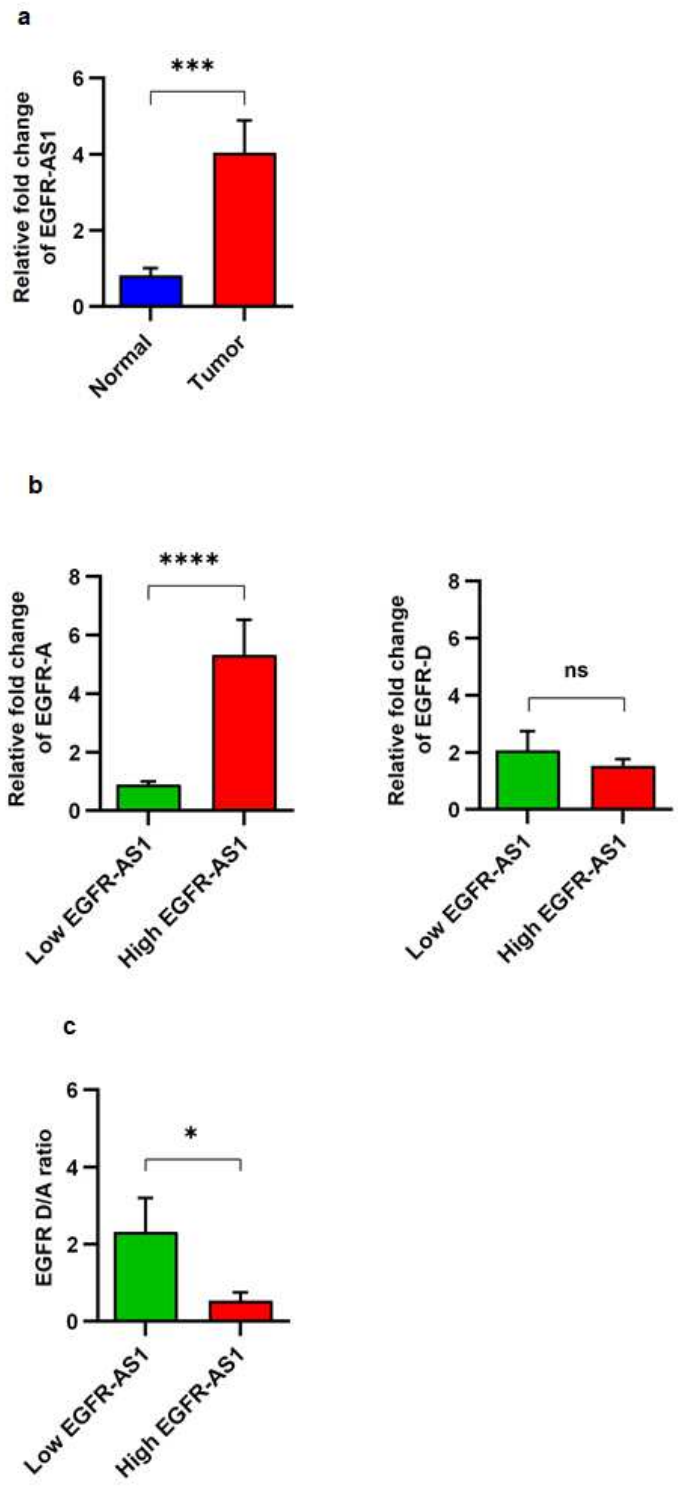


Figure 1

Role of EGFR-AS1 in maintaining the level of EGFR A and D isoforms in OSCC a. Relative expression level of lncRNA EGFR-AS1 in oral tumors compared with normal tissues. b. Graph showing the relative fold change of EGFR A and D isoforms in relation to EGFR-AS1 high and low expression groups. c. Graph

shows the EGFR D/A ratio in relation to EGFR high and low expression groups. (Statistical significance represented as p-value **** for < 0.0001, *** for <0.001, * for < 0.01, ns for non-significant, two tailed Student's t-test).

Figure.2

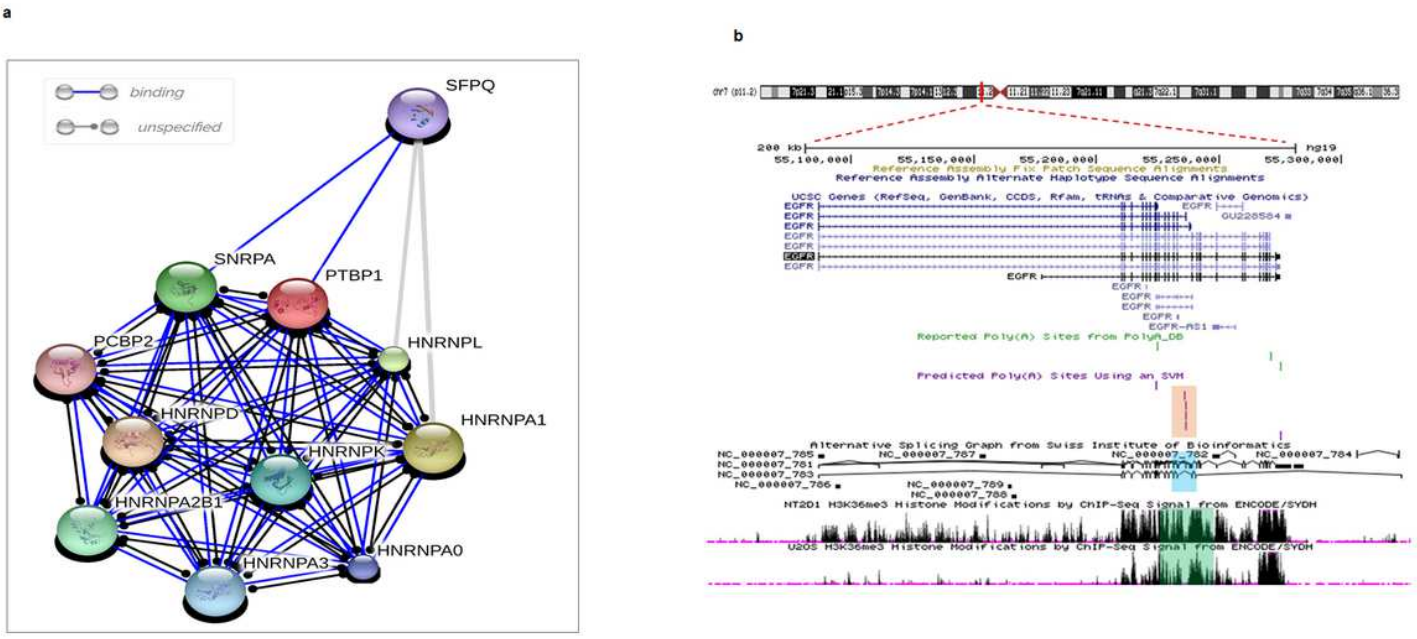


Figure 2

In-silico characterization of EGFR-AS1 role in alternative splicing of EGFR a. STITCH analysis showing the binding partners of PTBP1. b. UCSC genome browser data shows the H3K36me3 marks (green box) and polyA sites (orange box) prevalence between exon 15 and 16 skipped region (blue box) of EGFR.

Figure.3

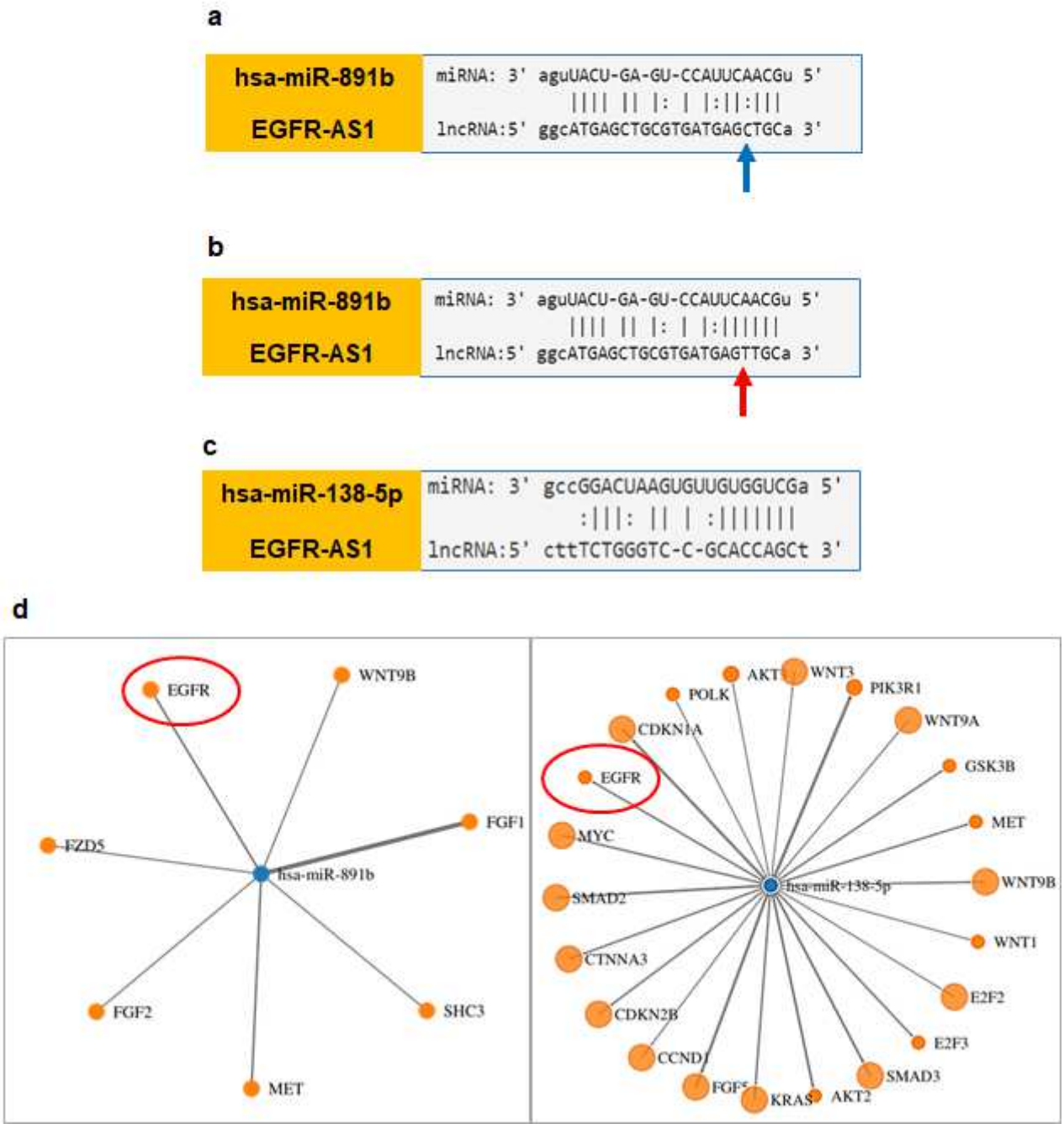


Figure 3

rs10251977 minor allele's effect on EGFR-AS1 529 a. Absence of minor allele in EGFR-AS1 (Blue arrow). b. Prediction of effect of minor allele in EGFR-AS1 generating binding site for miR-891b using lncRNASNP2 (Red arrow indicating presence of variant allele). c. Figure showing the binding site for another miRNA, miR-138-5p in EGFR-AS1. d. miRWalk database showing the targets of both miR-891b and miR-138-5p (EGFR encircled in both miRNAs).

Figure.4

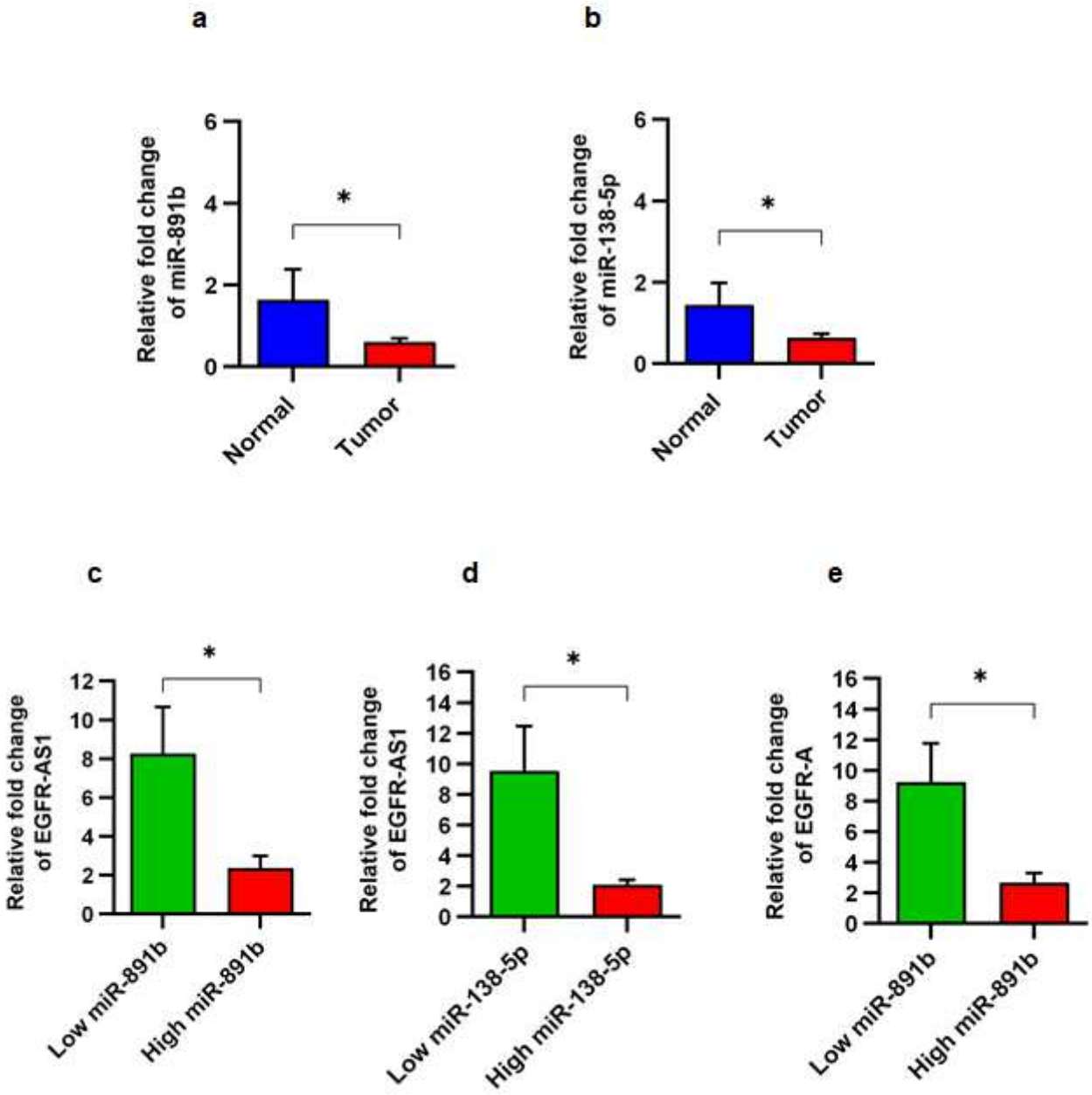


Figure 4

miR-891b sponging by EGFR-AS1 promoting the expression of EGFR-A isoform a. Graph showing the relative fold change of miR-891b. b. Graph showing the relative fold change of miR-138-5p. c. Graph shows the relative fold change of EGFR-AS1 in relation to miR-891b. d. Graph showing the relative fold change of EGFR-AS1 in relation to miR-138-5p levels. e. Graph shows the relative fold change of EGFR-AS1 in relation to miR-891b. (Statistical significance represented as * for $P < 0.01$, two tailed Student's t-test).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupptableS4miR1385pmiRWalkGSEAresults.xls
- SupptableS3miR891bmiRWalkGSEAresults.xls
- SupplementarymaterialsEGFRAS1manuscript.docx