Supplementary Material

Table1: Demographic information of healthy vs OSCC patients.

Table2: Fusion primers.

Table3: PCR conditions.

Figure S1: Bioanalyzer graph of cfDNA size of OSCC patients: The arrow mark in black shows peak of cfDNA size in bioanalyzer graph.

Figure S2: Fusion sequence: the purple sequence indicates sequence of TRMO, red sequence indicates junction and blue sequence indicates TRNT1.

Figure S3: Fusion validation in other OSCC samples: A. Sanger sequencing of the PCR products confirmed the fusion junction in our samples with results from the ChiTaRS fusion database. B. Agarose gel image of fusion (highlighted by the red box) identification in other patients' cfDNA after PCR.

Figure S4: Coding potential of the fusion sequence. A. The CPAT score suggests that fusion to be non-coding. B. The CNIT result was also used to validate the coding potential and likewise suggests that the fusion gene is a non-protein coding sequence.

Table1A: Demographic parameters of healthy vs OSCC patients.

Details	Healthy	OSCC Patients
Gender	Male: 10 (40%) Female: 15 (60%)	Male: 14 (47%) Female: 16 (53%)
Age (mean)	52.90526	58.25306
cfDNA concentration	Mean 1.04 Median 1.2	Mean 23.72 Median 22.83

Table1B: Details of OSCC patients.

OSCC Patients	**cfDNA>0.3 ng/□l** **(N=30)**
Smoking status	Current – 7 Past – 4 Never – 19
Alcohol abuse	NA
Diabetes Mellitus	Present: 9 (30%) Absent: 21 (70%)
IHD	Present: 5 (16%) Absent: 25 (84%)
S/P CVA	Present: 3 (10%) Absent: 27 (90%)
Other past malignancy	Present: 7 (23%) Absent: 23 (77%)
Primary site	Tongue – 20 (67%) Alveolar ridge – 8 (27%) FOM – 1 (3%) RMT – 1 (3%)
Ki67	Positive: 16 (53%)
P53	Positive: 14 (47%)
Keratinization	Present: 20 (67%)
P16	Positive: 3 (10%)

Table2: Fusion primers:

Forward	CTCCTGAGAACCACCAGACA
Reverse	GCTGCACCCCACTAATGTGT

Table 3: PCR conditions.

Initial denaturation	95 ^{o}C	2 mins	
Denaturation	95 ^{o}C	25 sec	28 cycles
Annealing	53^{o}C	35 sec	
Extension	72 ^{o}C	60 sec	
Final Extension	72 ^{o}C	5 mins	
Hold	4^{o}C	∞	

Figure S1.

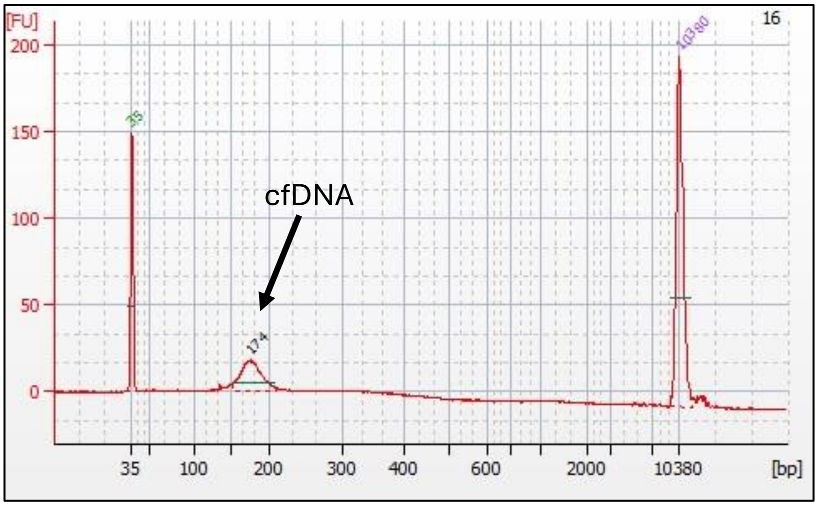


Figure S2.

GGGGCTTGGAGGAGCCGGGGCCTCGGCCTACAGCGACCCCGTGCGGCTGCGTGAAGCCGGCTCTGGAGACAG
GGAATCTTTTAACTGAGCCAGTCGGCTACTTGGAATCTTGTTTCTCGGCCAAGAATGGTACTCCAAGACAGCCAT
CCATTTGTAGCTATTCTCGAGCCTGTTTGAGGATTAGAAAGAGGATCTTTAATAATCCTGAACATTCCTTGATGG
GCCTAGAACAGTTTTCTCATGTTTGGATTTTGTTTGTTTTTCACAAAAACGGTCATTTGAGCTGTAAGGCAAAAGT
GCAGCCTCCTAGGCTGAATGGTGCAAAGACTGGAGTTTTTTCCACAAGGAGCCCTCATCGTCCCAATGCAATAG
GACTGACCCTGGCCAAGCTGGAAAAGGTAGAAGGTAACCCATTTCATTTCTACTTTTACCTTTTTCCTCACCCAAA
AAAGCCTTATAAAATAAACAGCATTCACCACTGGCTGTTTATTCAGACGAGTACCAATGATGTTAACTCCTGAGA
ACCACCAGACATTGAAGAAAATACAATGATTGGGGGGAGGGATAGTATTGGGAGATATACCTAATGCTAGATG
ACACATTAGTGGGGTGCAGCGCACCAGCATGGCACATGTATACCTATGGTAACTAACCCTGCACAATGGTGCAC
ATGTTACCCCTAAAACCTTAGAGGTATAATTAAAAAAAAACACAAAAAAGAAGT

Figure S3.

A.


ATCTCTCTCCTGAGAACCACCAGACATTGAAGAAAATACAATGATTGGGGGGAGGGATAGTATGGGA Sample
TGTTAACTCCTGAGAACCACCAGACATTGAAGAAAATACAATGATTGGGGGGAGGGATAGTATTGGG Parent sequence

B.

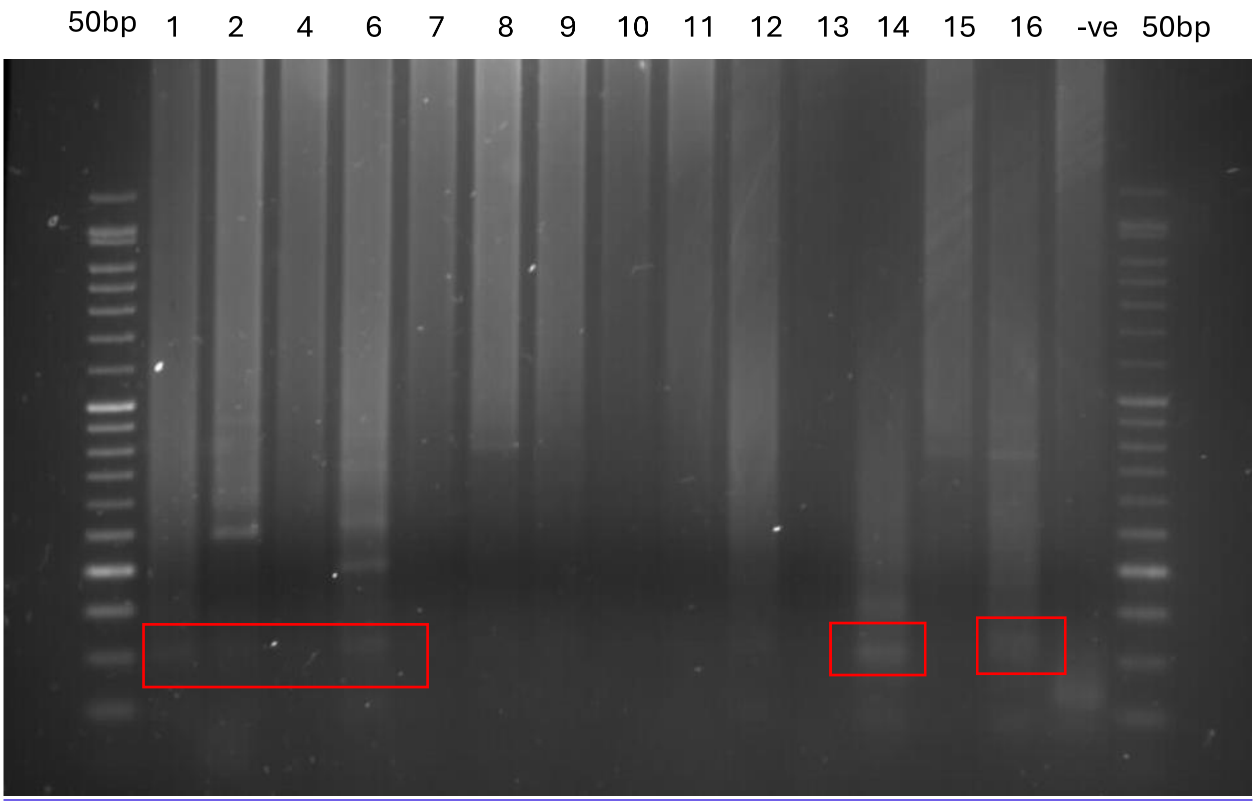
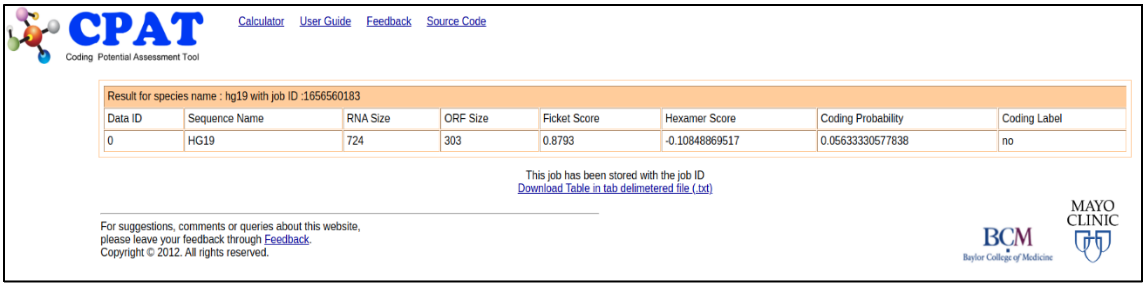


Figure S4.

A.



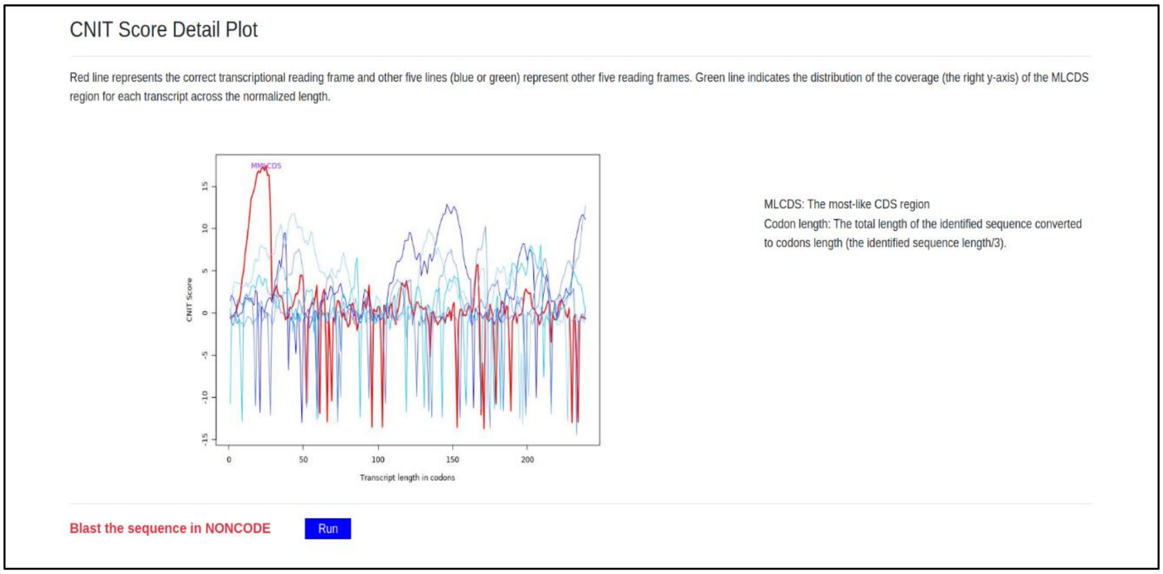
Result for species name : hg19 with job ID :1656560183

Data ID	Sequence Name	RNA Size	ORF Size	Ficket Score	Hexamer Score	Coding Probability	Coding Label
0	HG19	724	303	0.8793	-0.10848869517	0.05633330577838	no

This job has been stored with the job ID
Download Table in tab delimetered file (.txt)

For suggestions, comments or queries about this website,
please leave your feedback through Feedback.
Copyright © 2012. All rights reserved.

B.



CNIT Score Detail Plot

Red line represents the correct transcriptional reading frame and other five lines (blue or green) represent other five reading frames. Green line indicates the distribution of the coverage (the right y-axis) of the MLCDS region for each transcript across the normalized length.

MLCDS: The most-like CDS region
Codon length: The total length of the identified sequence converted to codons length (the identified sequence length/3).

Blast the sequence in NONCODE