Appendix Information

Wei-fa Yang, Na Qin, Xueyao Song, Canhua Jiang, et al. Genomic signature of mismatch repair deficiency in areca nut-related oral cancer.

Appendix Methods

Study subjects

Surgical resected tumors and matched normal and peripheral blood samples were obtained from 113 newly diagnosed OSCC patients without any treatment or neoadjuvant therapy prior to operation in the South China Oral Cancer Cohort (SCOCC). All patients had pathological diagnosis of OSCC from Xiangya Hospital of Central South University (98 OSCC patients) and Jiangsu Stomatological Hospital of Nanjing Medical University (15 OSCC patients) in China. Demographic information (age, gender, smoking, and AN chewing history) was collected from each patient. Among all 113 patients, 89 had a history of AN chewing, which was defined as the behavioral use of AN for at least three years. Frozen tumor and paired normal specimens were stained with hematoxylin and eosin and microscopically evaluated by two independent pathologists. Samples were frozen in liquid nitrogen and stored at -80°C. Only OSCC tissues with the malignant cell purity of over 70% were selected for DNA and/or RNA extraction and subsequent sequencing. The study was approved by the local ethics committee and written informed consent was obtained from each participant. Detailed demographic and clinical characteristics are provided in **Appendix Table 1**.

To further compare the genomic and transcriptome features of patients with different AN exposure, we integrated sequencing data of 325 OSCC patients from TCGA database and 50 OSCC patients from a published Taiwan study (43 patients with AN chewing history)(Chen et al., 2017) (**Appendix Table 1**). The 325 non-Asian and non-Indian TCGA OSCC patients were considered AN-negative since AN chewing is rare in Europe and the United States(Humans, 2012). Clinical and follow-up information for these patients was obtained from the Firehose Broad GDAC (http://gdac.broadinstitute.org/runs/stddata__2016_01_28/data/HNSC/20160128/).

DNA extraction and whole-exome sequencing

Genomic DNA was extracted from frozen oral tissues using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) and from blood samples using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. The quality and quantity of the extracted samples were assessed using the NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, DE, USA), Qubit 2.0 Fluorometer (Life Technologies, CA, USA) and 1% agarose

gel electrophoresis.

Whole-exome sequencing was performed on matched tumor-blood samples from 113 OSCC Chinese patients. Exome capture was performed on 1.5μg of high-quality genomic DNA per sample by SureSelect Human All Exon V6 (Agilent Technologies, Inc., Santa Clara, CA, USA) followed by the 150-bp paired-end sequencing on the HiSeq X Ten instrument (Illumina, San Diego, CA, USA) according to the manufacturer's protocols.

In 325 TCGA OSCC patients, 316 tumor-blood pairs were examined by whole-exome sequencing and 320 tumor samples were subjected to RNA sequencing. Additionally, 50 tumor-blood OSCC pairs from Taiwan were detected by whole-exome sequencing.

Total RNA extraction and RNA sequencing

Total RNA was extracted from 46 matched tumor-normal samples (40 AN-related and 6 AN-negative) using the RNeasy Mini Kit (Qiagen, Hilden, Germany). The quality and quantity of extracted RNA were assessed using the NanoDrop 2000 (Thermo Fisher Scientific, Wilmington, DE, USA), Qubit 2.0 Fluorometer (Life Technologies, CA, USA) and 1% agarose gel electrophoresis. RNA integrity was assessed using the RNA Nano 6000 Assay Kit (Agilent Technologies, CA, USA).

A total of 3μg of high-quality RNA per sample was used for ribosomal RNA removal by the Epicentre Ribo-zero rRNA Removal Kit (Epicentre, USA) and the sequencing library was prepared using the rRNA-depleted RNA by the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB, USA) following manufacturer's recommendations. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina, San Diego, CA, USA), followed by the 150-bp paired-end sequencing on the HiSeq X Ten instrument (Illumina, San Diego, CA, USA) according to the manufacturer's protocols.

Areca nut extract preparation

The method of preparing ANE has been described in a previous study(Xu et al., 2017). In brief, we first ground and extracted 30 g of tender AN from dried nuts at 4°C for 1 h with deionized water, and then centrifuged the squeezed extract at 12,000 g for 10 minutes. The supernatant was harvested, filtered and concentrated by lyophilization. The weighted dry powder was stored

at -20°C before dissolved in deionized water as ANE for research use.

RNA and protein expression detection

Total RNA was isolated from OSCC cell lines using Trizol reagent (Invitrogen). 1000 ng of RNA was reverse transcribed into cDNA using PrimeScript™ RT Master Mix (Takara) according to the manufacturer's instructions. Quantitative RT-PCR (RT-qPCR) was performed with the SYBR PCR Master Mix reagent kit (Takara) and QuantStudio 7 Flex real-time PCR System (Applied Biosystems) according to the manufacturer's instructions. The primers were described in **Appendix Table 7**.

Cells were lysed with the mammalian protein extraction reagent RIPA (Beyotime). Protein lysates were cleared by centrifugation and concentrations were determined using the BCA protein assay kit (Beyotime). The lysate was combined with 4×LDS Sample Buffer and 10×Sample Reducing Agent (Life technologies) followed by denaturation at 95°C for 5 min. For Western blotting, 40 μg of protein was separated by 10% SDS-PAGE and transferred to a 0.45-μm PVDF membrane (Millipore). Membranes were then blocked in 5% skim milk and incubated with rabbit anti-ATG2A (1:1000, ab226931, Abcam), anti-ATG9A (1:1000, ab108338, Abcam), anti-MLH1 (1:1000, ab92312, Abcam), anti-MSH2 (1:1000, ab227941, Abcam) and mouse anti-GAPDH (1:1000, AG019, Beyotime), respectively. Immunoreactive proteins were visualized using a molecular imager (Bio-Rad).

Cell viability assessment in vitro

Cell culture and Cell transfection

Oral and laryngeal cancer cells (CAL27 and HEP2) were maintained in DMEM supplemented with 10% FBS. Transfection of OSCC cells (1.0×10^5 cells) with specific siRNAs targeting ATG2A and WEE1 (Invitrogen U.S.A) constructs was carried out using lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. The sequences of siRNAs are described in **Appendix Table 7**.

Cell proliferation and Colony formation assays

Cell viability was measured with a Cell Counting Kit-8 (CCK8, Dojindo, Japan) according to the manufacturer's instructions. The transfected cells were incubated in 60 mm culture plates for 24 hours, and then 3,000 cells in 200 ul of culture medium were seeded in E-Plate 96. 96-

well plates were incubated overnight at 37°C with 5% CO2. Next, 10μl of CCK8 was added to each well with 100 μl of medium at the indicated time points, and the absorbance was measured at 450 nm after two hours using a microplate reader. For the colony formation assay, 800 transfected cells were seeded into 60 mm culture plates and incubated for 12 days at 37°C with 5% CO_2.

Transwell migration assays

Cell migration was measured in Costar Transwell plates (6.5mm diameter insert, 8.0 mm pore size, polycarbonate membrane, Corning Sparks, MD). The lower chamber was filled with 600 μl of medium containing 10% fetal bovine serum. Next, 3×10^4 cells in 300 μl of serum-free medium were added to the upper chamber of the Transwell system. After incubation at 37°C for 24 hours, non-migrated cells were removed from the upper surface of the membrane using a cotton swab.

Whole-exome sequencing data processing

Read mapping

The quality score and base-call distributions of raw sequencing reads were accessed with the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). The sequenced reads were mapped to the reference genome (GRCh37) using Burrows-Wheeler Aligner (BWA-MEM) algorithm (http://bio-bwa.sourceforge.net/) with default parameters(Li and Durbin, 2009). The duplicates were marked with Picard (v1.70, http://broadinstitute.github.io/picard/) and discarded for further analyses. We then performed local realignment and base quality score recalibration (BQSR) using the Genome Analysis Toolkit (GATK, version 3.7) with default settings(McKenna et al., 2010). The mapping rate and sequencing coverage for each sample are listed in **Appendix Table 8**.

Somatic variants calling

Somatic single-nucleotide variations (SNVs) and small insertions and deletions (indels) were detected using the Mutect2 mode in GATK following the best practice (https://software.broadinstitute.org/gatk/best-practices/). Then, we excluded somatic variants if they were found in: (1) a panel of normal built by the 113 matched normal samples; (2) the segmental duplication or simple Repeat regions marked by UCSC browser

(http://genome.ucsc.edu/); (3) the 1000 genomes project (the Phase III integrated variant set release, across 2,504 samples) with the same mutation direction.

Putative driver genes identification

We applied the MutSigCV algorithm (version 1.4)(Lawrence et al., 2013) and the IntOGen platform(Gonzalez-Perez et al., 2013) to identify significantly mutated genes (SMGs). MutSigCV was used to detect genes with higher mutation rates than the background after correcting for the nucleotide context, gene expression, replication time, silent mutations and mutations in surrounding regions. IntOGen included two algorithms (OncodriveCLUST(Tamborero et al., 2013) and OncodriveFM(Gonzalez-Perez and Lopez-Bigas, 2012)) to find genes with highly clustered mutations and non-random distribution of functional mutations, respectively. *P*-values were corrected from multiple testing using the false discovery rate (Benjamini-Hochberg FDR) and genes with a q value < 0.05 in any algorithm were reported in this study. Then, the putative driver genes were determined with one of the following conditions met: (1) the gene with a mutation rate > 5% in 113 OSCC or 89 AN-related OSCC patients; or (2) the gene was only mutated in 89 AN-related OSCC patients. We also downloaded somatic variants of 316 TCGA OSCC samples from the Firehose Broad GDAC (http://gdac.broadinstitute.org/, version 2016_01_28) and compared the mutational profiles between 89 AN-related patients from SCOCC and 340 AN-negative patients (316 samples form TCGA and 24 AN-negative samples from SCOCC). To further validate the identified AN-related genomic features, we retrieved the whole-exome sequencing data of 50 OSCC patients (43 AN-related and 7 AN-negative samples) from a Taiwan study(Chen et al., 2017) and performed the same analysis.

Mutational signature analysis

The Bioconductor package SomaticSignatures was used to perform mutational signature analysis(Gehring et al., 2015). The algorithm was developed based on the non-negative matrix factorization (NMF) method. Briefly, six possible substitutions (C>A, C>G, C>T, T>A, T>C and T>G) were considered in NMF, and SNVs together with their 5' and 3' neighboring bases could then be organized into the 96 trinucleotide contexts representing specific cancer-causing processes. Signatures identified in this study were compared to the COSMIC signatures

(http://cancer.sanger.ac.uk/cosmic/signatures) with cosine similarity distance. Then, we further estimated the contribution of 30 COSMIC mutational signatures to each sample's mutational spectrum with the R package deconstructSigs(Rosenthal et al., 2016), and mutational signatures attributed to the same cancer-causing process were combined. The signatures associated with mismatch repair deficiency (dMMR) include COSMIC Signature 6, Signature 15, Signature 20 and Signature 26. APOBEC related signatures include COSMIC signature 2 and Signature 13. The same analysis was applied to 316 TCGA and 50 Taiwan OSCC samples.

Copy number variations analysis

Somatic copy number variations (CNVs) were detected following the GATK best practice (https://software.broadinstitute.org/gatk/best-practices/) with default parameters. The copy number segments were analyzed with GISTIC2 to identify significantly amplified/deleted regions. FDR *q* values were computed for the aberrant regions. Peak regions with *q* values of <0.05 were defined as highly amplified/deleted regions.

RNA quantification and differential expression analysis

For the RNA-seq data of 46 tumor-normal pairs from SCOCC, the FastQC package (http://www.bioinformatics.babraham.ac.uk/projects/fastqc) was used to access the quality score distribution of the sequencing reads, and low-quality reads were removed prior to analysis. The remaining qualified reads were aligned to the GENCODE v19 genome assembly with STAR (version 2.4.1) and quantified with RSEM (version 1.2.25). Additionally, level 3 RNASeqV2 data of 320 OSCC samples in TCGA from the Firehose Broad GDAC (http://gdac.broadinstitute.org/, version 2016_01_28) was included. Differential expression analysis was performed with R package DESeq2(Love et al., 2014) in 40 AN-related patients from SCOCC and 326 AN-negative samples (320 samples from TCGA and 6 AN-negative samples from SCOCC). The batch effect was removed by sva package and raw read counts data was normalized by DESeq2. The magnitude (log2 transformed fold change) and significance (*P*-value) of differential expression between groups were calculated.

Statistical analysis

All statistical tests were performed using the Wilcoxon rank-sum test for continuous data, and the Spearman's rank correlation was used for the estimation of correlation. Student's *t* test was

used to analyze data from RT-qPCR, cell proliferation and migration. The Fisher's exact test was used to assess the difference in count data. The Kaplan-Meier method and the Log-rank test were used to estimate the difference in survival time between different subgroups categorized by the dMMR signature proportion. Multivariate Cox proportional hazard regression analysis was performed to estimate the HR and 95% CI, with the adjustment of age, smoking status, AN chewing status, clinical stage, recurrence and metastatic status. General statistical analysis was performed using R (R version 3.2.2). Two-sided *P* values less than 0.05 were considered statistically significant.

REFERENCES

Chen TW, Lee CC, Liu H, Wu CS, Pickering CR, Huang PJ *et al.* (2017). APOBEC3A is an oral cancer prognostic biomarker in Taiwanese carriers of an APOBEC deletion polymorphism. *Nature communications* 8(1)*:*465.

Gehring JS, Fischer B, Lawrence M, Huber W (2015). SomaticSignatures: inferring mutational signatures from single-nucleotide variants. *Bioinformatics* 31(22)*:*3673-3675.

Gonzalez-Perez A, Lopez-Bigas N (2012). Functional impact bias reveals cancer drivers. *Nucleic acids research* 40(21)*:*e169.

Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, Tamborero D, Schroeder MP, Jene-Sanz A *et al.* (2013). IntOGen-mutations identifies cancer drivers across tumor types. *Nature methods* 10(11)*:*1081-1082.

Humans IWGotEoCRt (2012). Personal habits and indoor combustions. Volume 100 E. A review of human carcinogens. *IARC monographs on the evaluation of carcinogenic risks to humans* 100(Pt E)*:*1-538.

Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A *et al.* (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499(7457)*:*214-218.

Li H, Durbin R (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14)*:*1754-1760.

Love MI, Huber W, Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15(12)*:*550.

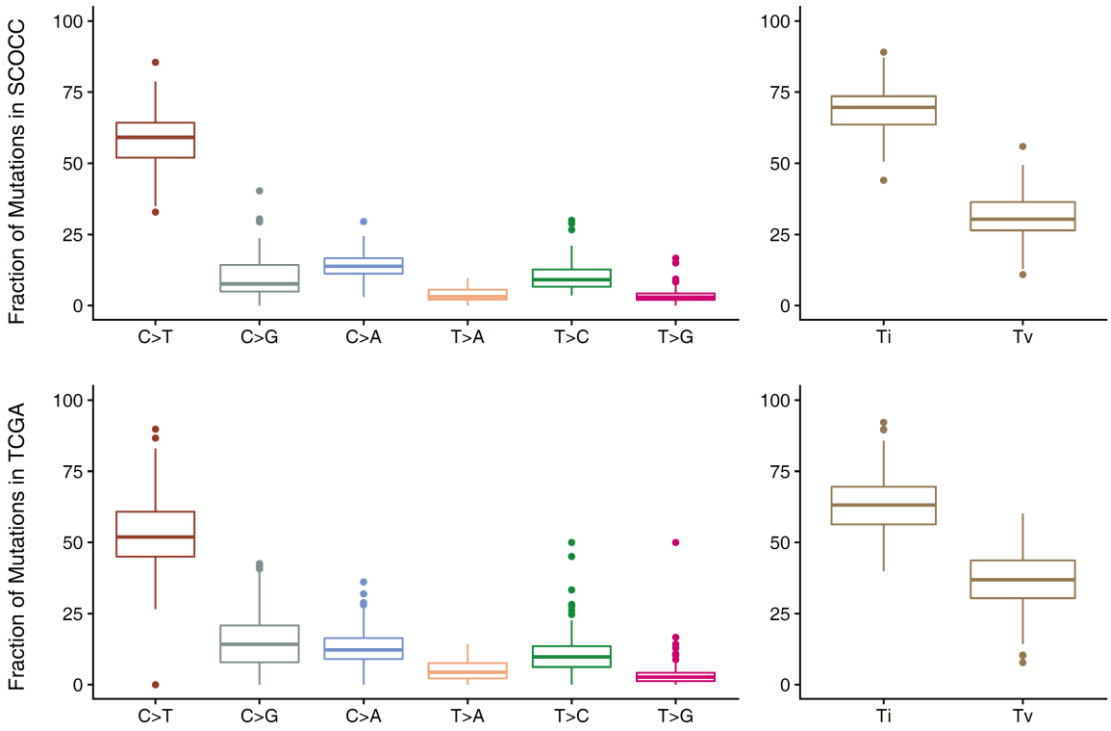
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A *et al.* (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* 20(9)*:*1297-1303.

Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C (2016). DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome biology* 17(31.

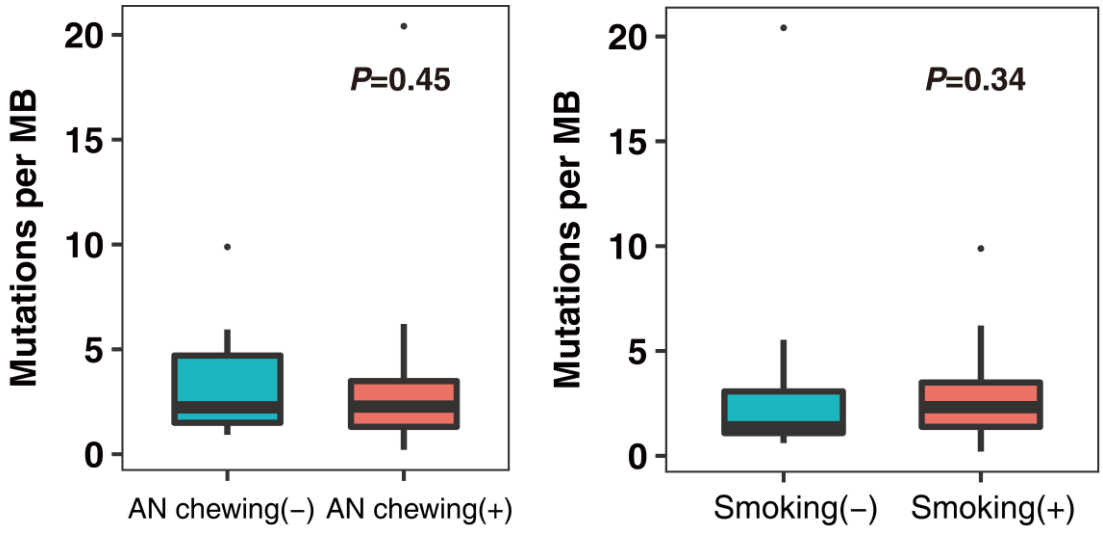
Tamborero D, Gonzalez-Perez A, Lopez-Bigas N (2013). OncodriveCLUST: exploiting the positional clustering of somatic mutations to identify cancer genes. *Bioinformatics* 29(18)*:*2238-2244.

Xu Z, Huang CM, Shao Z, Zhao XP, Wang M, Yan TL *et al.* (2017). Autophagy Induced by Areca Nut Extract Contributes to Decreasing Cisplatin Toxicity in Oral Squamous Cell Carcinoma Cells: Roles of Reactive Oxygen Species/AMPK Signaling. *International journal of molecular sciences* 18(3).

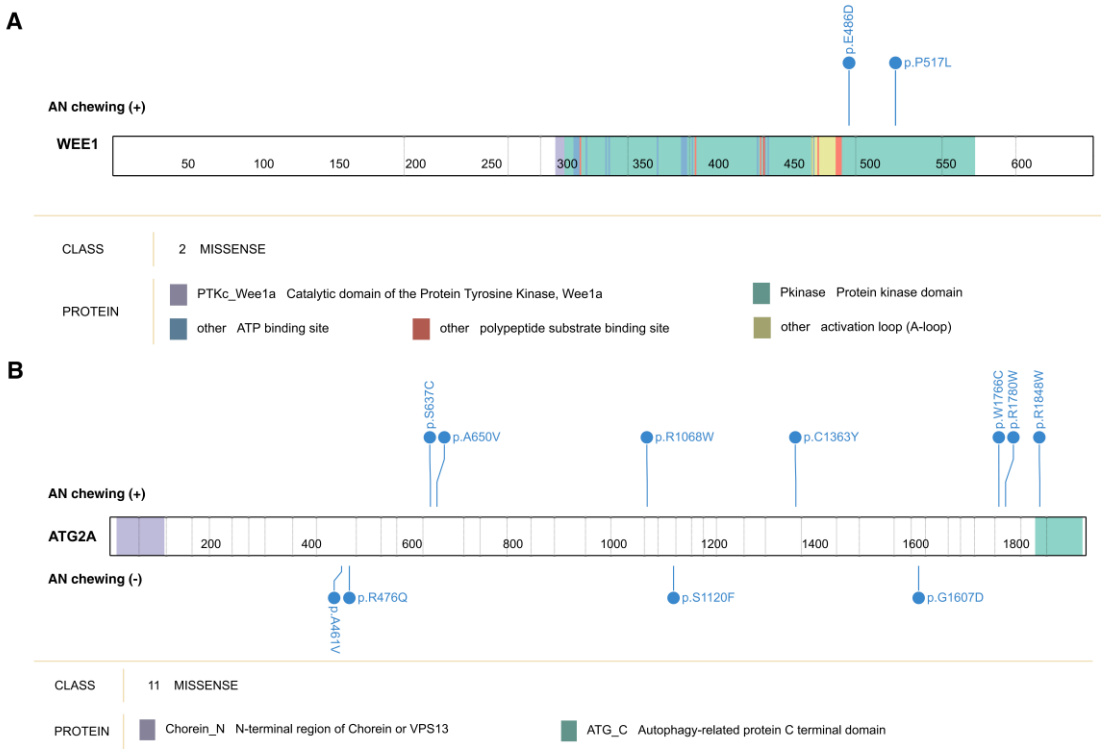
Appendix Figure 1. Mutational spectrum of 113 OSCC patients from our SCOCC study and 316 OSCC patients from the TCGA project.



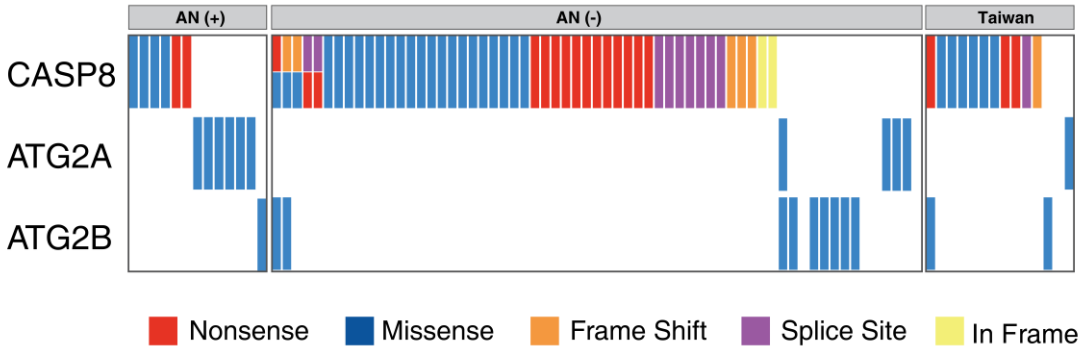
Appendix Figure 2. Mutation rates between patients with different exposure status.



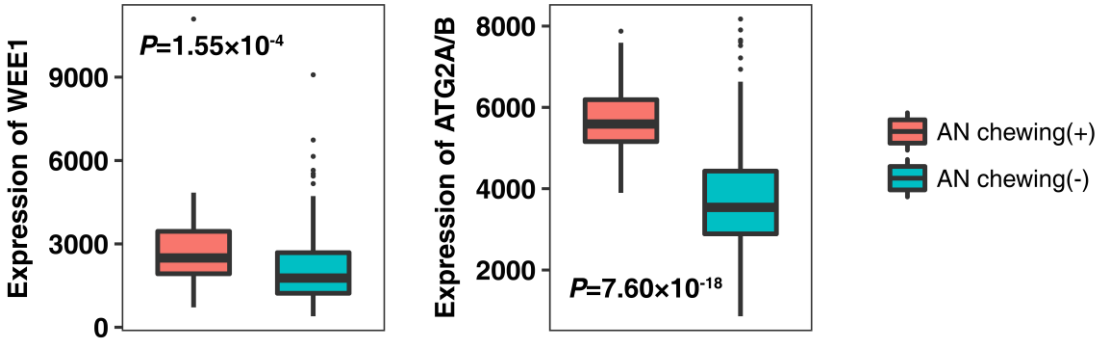
Appendix Figure 3. The mutational spectrum of *WEE1* and *ATG2A* in 89 AN-relative OSCC patients compared to 340 AN-negative OSCC patients.



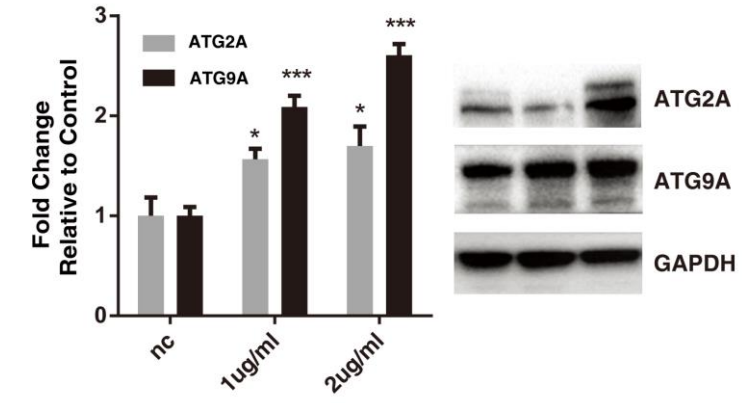
Appendix Figure 4. Mutual exclusivity pattern of mutations in *ATG2* homologues (*ATG2A* and *ATG2B*) and *CASP8* in SCOCC, TCGA and Taiwan studies.



Appendix Figure 5. Differential expression analysis of *WEE1* and ATG2 homologues (*ATG2A* and *ATG2B*) between 40 AN-related patients and 326 AN-negative patients.

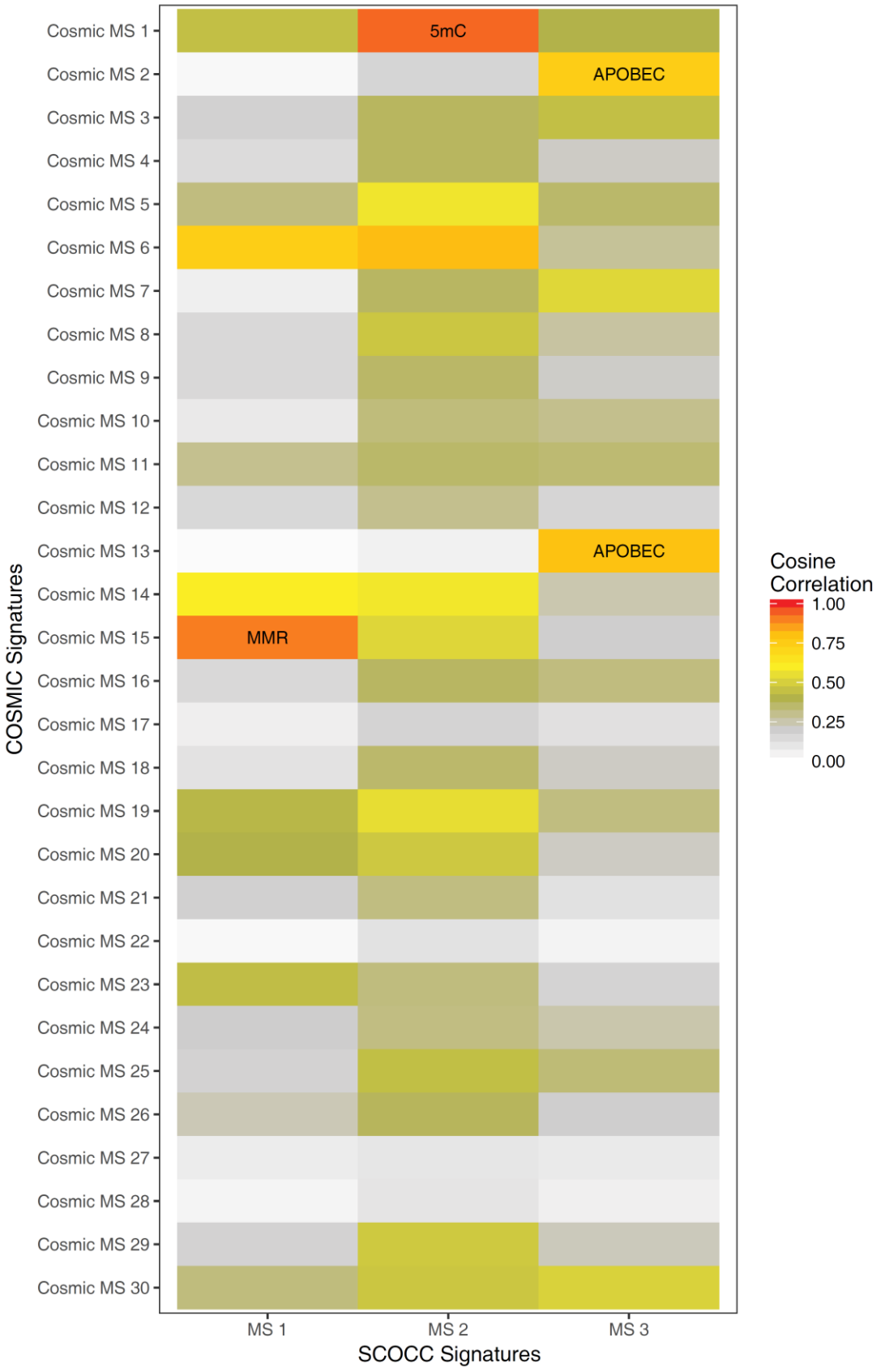


Appendix Figure 6. The expression levels of *ATG2A* and *ATG9A* in the CAL27 cell line treated with ANE (1 or 2 ug/mL) for 5 days.

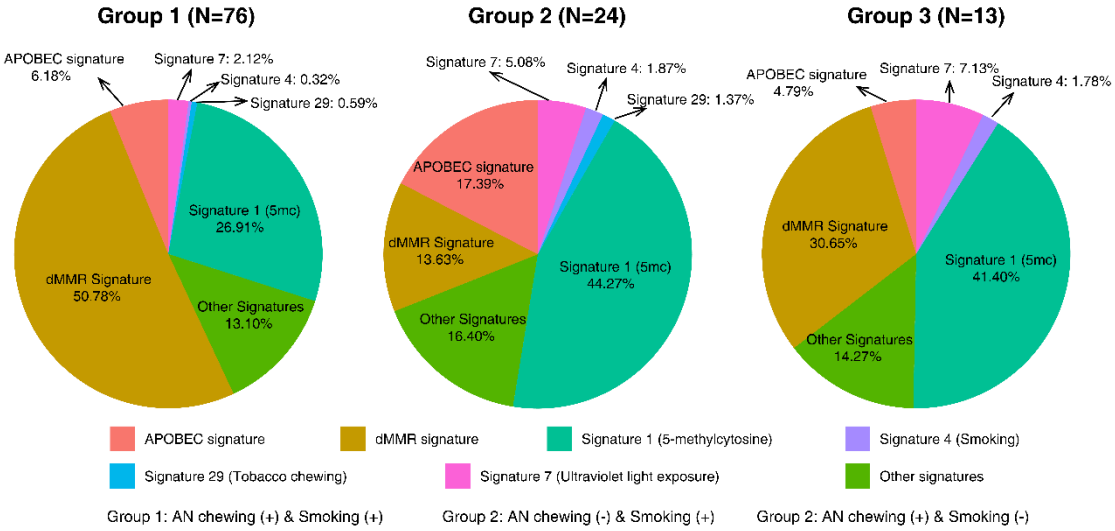


.

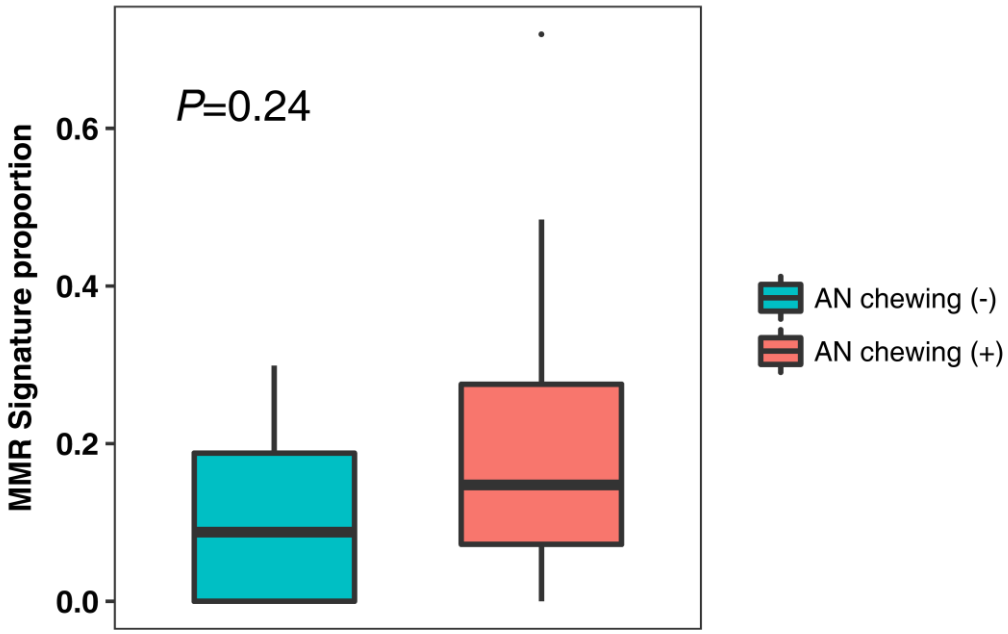
Appendix Figure 7. Correlation between mutational signatures derived in SCOCC study and previously defined signatures from COSMIC.



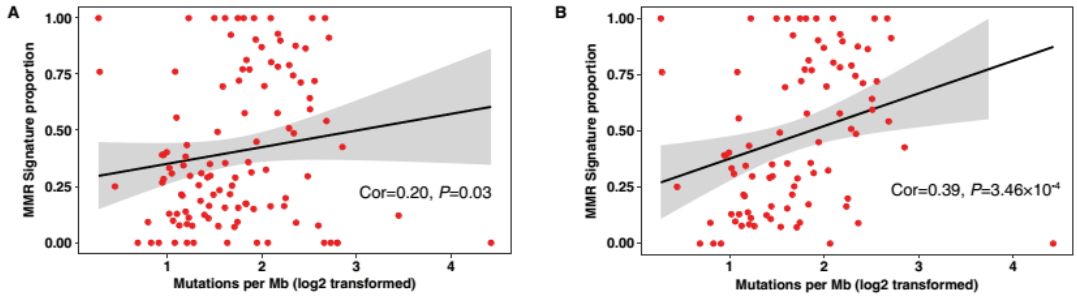
Appendix Figure 8. Proportions of different mutational signature types among three groups of OSCC patients from the SCOCC study.



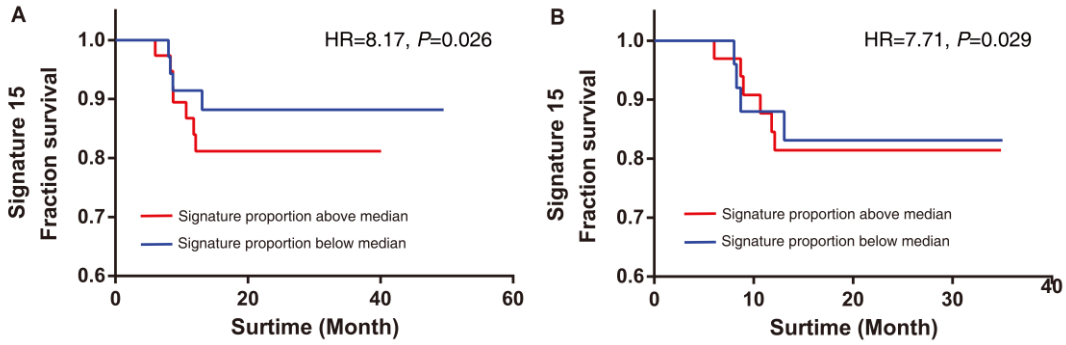
Appendix Figure 9. The fraction of MMR signature is higher, but not statistically significant, in AN-related patients than AN-negative patients in the Taiwanese study.



Appendix Figure 10. MMR signature is linearly associated with mutation burden of OSCC patients.



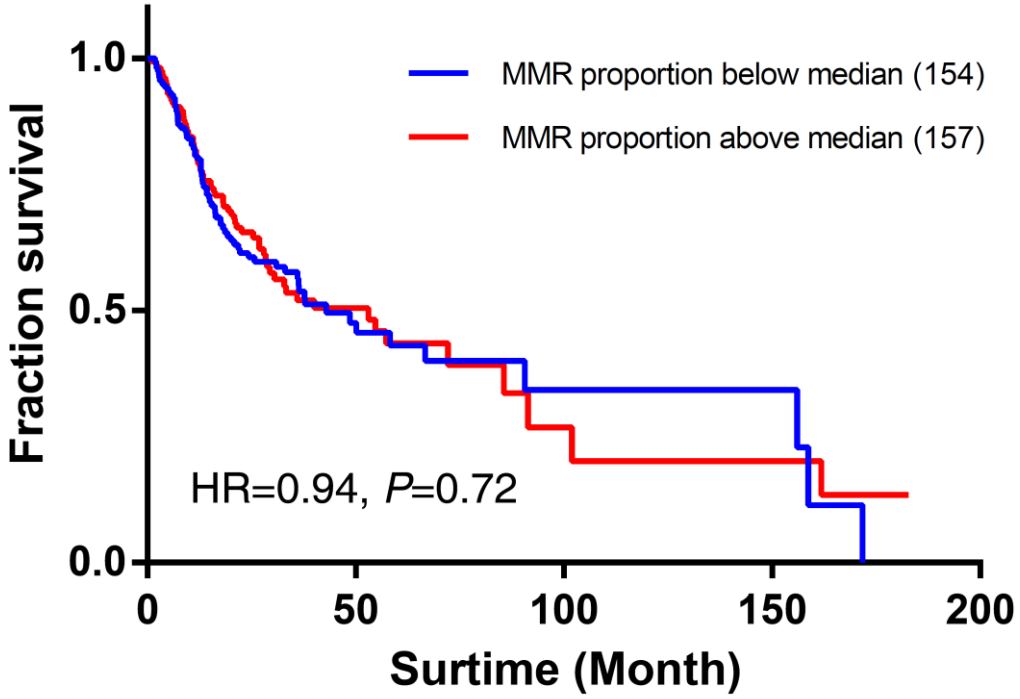
Appendix Figure 11. Kaplan-Meier plot for overall survival by the mutation signature for OSCC patients.



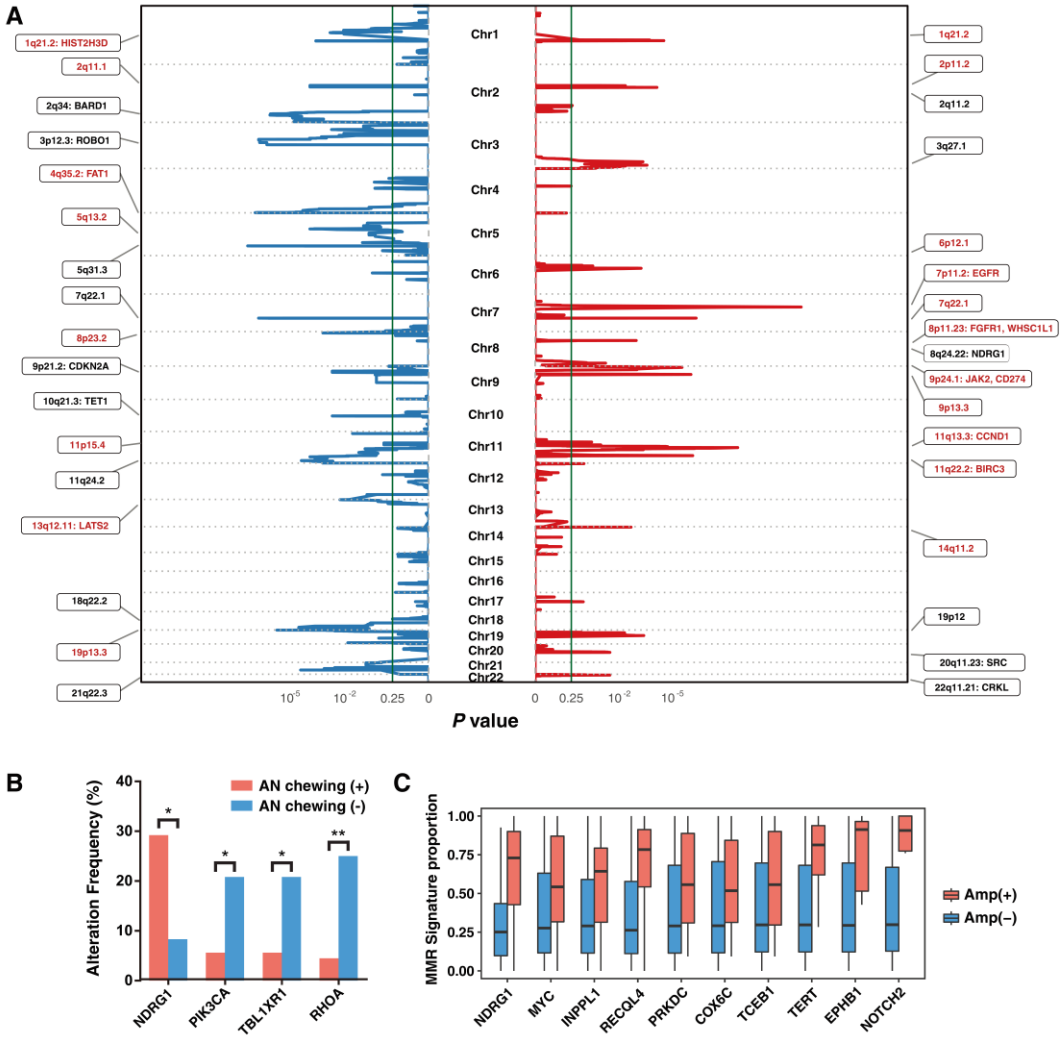
(A) Higher fraction of Signature 15 is significantly associated with worse overall survival in 75 OSCC patients.

(B) Higher fraction of Signature 15 is significantly associated with worse overall survival in 59 AN-related OSCC patients.

Appendix Figure 12. Kaplan-Meier plot for overall survival by the mutation signature for 311 OSCC patients from TCGA. No significant difference in overall survival was found for subgroups stratified according to the fraction of MMR signature.



Appendix Figure 13. Copy-number alterations in 113 OSCC patients from SCOCC study.

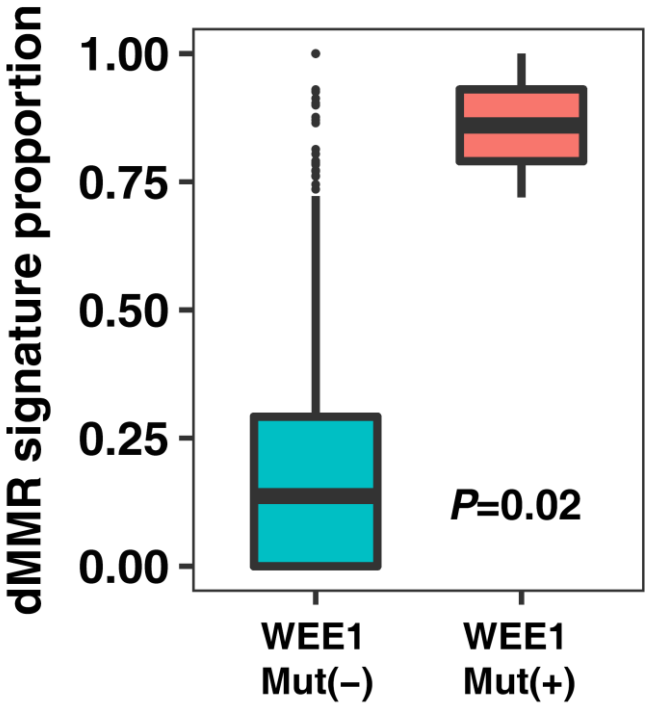


(A) GISTIC analysis of focal copy-number alterations.

(B) The copy number alteration status of cancer-related gene in 8q24.22 and 3q26.32 in AN-related and AN-negative OSCC patients.

(C) The association of cancer-related genes copy number alteration and the proportion of dMMR signature.

Appendix Figure 14. The proportion of dMMR signature was significantly higher in *WEE1* mutated samples.



Appendix Table 1. Demographic characteristics of 113 OSCC patients from SCOCC, 325 from the TCGA project, and 50 from the Taiwanese study.

	SCOCC (n=113)	TCGA (n=325)	Taiwan (n=50)
Age (median; range)	49·3 (27·8-78·9)	61·8 (20·0-90·1)	51·4 (31·1-79·7)
Gender (frequency; percentage)			
Male	111 (98·2)	221 (68·0)	44 (88·0)
Female	2 (1·8)	104 (32·0)	6 (12·0)
Smoking history (frequency; percentage)			
Ever smoker	100 (88·5)	232 (71·4)	41 (82·0)
Lifelong non-smoker	13 (11·5)	84 (25·8)	9 (18·0)
N/A	n/a	9 (2·8)	n/a
	Areca nut chewing history (frequency; percentage)		
Ever areca nut chewing	89 (78·8)	n/a	43 (86·0)
Lifelong non-areca nut chewing	24 (21·2)	n/a	7 (14·0)
N/A	n/a	325 (100·0)	n/a
Site (frequency; percentage)			
Tongue	57 (50·4)	148 (45·5)	19 (38·0)
Buccal mucosa	39 (34·5)	21 (6·5)	21 (42·0)
Floor of the mouth	5 (4·4)	59 (18·2)	1 (2·0)
Alveolar Ridge	11 (9·7)	17 (5·2)	7 (14·0)
Hard palate	1 (0·9)	7 (2·2)	2 (4·0)
Lip	n/a	2 (0·6)	n/a
Oral Cavity	n/a	71 (21·8)	n/a
Stage (frequency; percentage)			
I	15 (13·3)	19 (5·8)	1 (2·0)
II	44 (38·9)	50 (15·4)	12 (24·0)
III	30 (26·5)	59 (18·2)	1 (2·0)
IV	24 (21·2)	159 (48·9)	21 (42·0)
V	n/a	7 (2·2)	6 (12·0)
N/A	n/a	31 (9·5)	9 (18·0)

Appendix Table 2. Significantly mutated genes identified in 113 OSCC samples from SCOCC study.

							OncodriveCLUST						OncodriveFM					MutSigCV	
SYMBOL	GENE	SIGNALS	SIGNAL_COUNT	MUTS_CS_SAMPLES	KNOWN_DRIVER	CONNECTED_TO_DRIVER	MUTS_IN_CLUST	NUM_CLUSTERS	CLUST_COORDS	GENE_SCORE	ZSCORE	p-value	q-value	PPH2_SCORE_PVALUE	MA_SCORE_PVALUE	SIFT_SCORE_PVALUE	p-value	q-value	p-value	q-value
TP53	ENSG00000141510	CFR	3	86	TRUE	TRUE	64	10	[124,125]:4,[136,141]:3,[152,155]:3,[158,166]:7,[173,179]:13,[193,196]:6,[213,214]:4,[237,258]:16,[298,298]:2,[337,347]:6	0·62	2·63	4·22E-03	8·44E-03	1·00E-08	1·00E-08	1·00E-08	0·00E+00	0·00E+00	0·00E+00	0·00E+00
NOTCH1	ENSG00000148400	FR	2	23	TRUE	TRUE	2	1	[1418,1418]:2	0·13	-1·12	8·69E-01	8·69E-01	1·00E-08	1·00E-08	2·50E-03	2·22E-16	7·82E-14	4·48E-13	8·45E-10
CDKN2A	ENSG00000147889	CFR	3	20	TRUE	TRUE	7	1	[113,113]:7	1·00	5·55	1·46E-08	8·76E-08	1·00E-08	1·00E-08	1·00E-08	0·00E+00	0·00E+00	0·00E+00	0·00E+00
FAT1	ENSG00000083857	FR	2	20		TRUE								1·00E-08	1·00E-08	1·00E-04	0·00E+00	0v00E+00	5·55E-16	1·16E-12
DST	ENSG00000151914	F	1	10		TRUE								2·00E-03	1·69E-01	2·71E-02	7·33E-04	4·92E-02	1·20E-02	1·00E+00
HRAS	ENSG00000174775	CFR	3	8	TRUE	TRUE	7	2	[12,13]:5,[61,61]:2	0·84	4·30	8·43E-06	2·53E-05	1·08E-02	1·00E-04	1·60E-03	3·88E-07	7·82E-05	0·00E+00	0·00E+00
CASP8	ENSG00000064012	CR	2	7	TRUE	TRUE	4	2	[268,268]:2,[472,472]:2	0·57	2·25	1·22E-02	1·84E-02	5·83E-02	1·40E-02	1·07E-01	4·71E-03	1·19E-01	0·00E+00	0·00E+00
RASA1	ENSG00000145715	FR	2	6		TRUE								4·00E-04	9·00E-04	3·00E-03	2·54E-07	5·95E-05	8·64E-07	1·36E-03
TSC2	ENSG00000103197	F	1	6	TRUE	TRUE								1·40E-03	6·50E-02	2·40E-02	2·16E-04	2·03E-02	8·18E-02	1·00E+00

Appendix Table 3. Significantly mutated genes identified in 89 AN-related OSCC patients from SCOCC study.

SYMBOL	GENE	SIGNALS	SIGNAL_COUNT	MUTS_CS_SAMPLES	KNOWN_DRIVER	CONNECTED_TO_DRIVER	OncodriveCLUST							OncodriveFM					MutSigCV	
							MUTS_IN_CLUST	NUM_CLUSTERS	CLUST_COORDS	GENE_SCORE	ZSCORE	p-value	q-value	PPH2_SCORE_PVALUE	MA_SCORE_PVALUE	SIFT_SCORE_PVALUE	p-value	q-value	p-value	q-value
TP53	ENSG00000141510	CFR	3	66	TRUE	TRUE	48	9	[124,124]:2,[141,141]:2,[152,155]:3,[159,166]:6,[173,179]:10,[194,196]:5,[213,214]:3,[238,250]:13,[337,342]:4	0·64	2·81	2·45E-03	3·27E-03	1·00E-08	1·00E-08	1·00E-08	0·00E+00	0·00E+00	1·11E-16	2.·62E-13
CDKN2A	ENSG00000147889	CFR	3	14	TRUE	TRUE	5	1	[113,113]:5	1·00	5·55	1·46E-08	5·84E-08	1·00E-08	1·00E-08	1·00E-08	0·00E+00	0·00E+00	5·55E-12	1·16E-08
HRAS	ENSG00000174775	CFR	3	7	TRUE	TRUE	6	2	[12,13]:4,[61,61]:2	0·82	4·13	1·85E-05	3·70E-05	2·66E-02	3·00E-04	2·30E-03	3·26E-06	5·40E-04	6·81E-06	1·07E-02
NOTCH1	ENSG00000148400	FR	2	15	TRUE	TRUE								2·40E-03	1·00E-04	2·92E-02	1·37E-06	2·73E-04	6·91E-08	1·18E-04
FAT1	ENSG00000083857	FR	2	13		TRUE								1·00E-08	1·00E-08	1·00E-08	0·00E+00	0·00E+00	9·88E-11	1·86E-07
CASP8	ENSG00000064012	FR	2	6	TRUE	TRUE	2	1	[472,472]:2	0·33	0·42	3·38E-01	3·38E-01	1·29E-02	8·00E-03	7·02E-02	6·00E-04	4·27E-02	0·00E+00	0·00E+00
RASA1	ENSG00000145715	FR	2	5		TRUE								4·20E-03	3·50E-03	7·10E-03	1·53E-05	2·17E-03	1·48E-05	2·15E-02
ATG2A	ENSG00000110046	F	1	6										1·86E-02	5·02E-02	3·10E-03	2·75E-04	2·28E-02	1·57E-01	1·00E+00
WEE1	ENSG00000166483	F	1	2										1·68E-02	8·70E-03	7·99E-02	8·98E-04	3·89E-02	7·22E-03	1·00E+00

Appendix Table 4. Univariate analysis of overall survival (OS) in 75 OSCC patients from SCOCC study.

Characteristics	SCOCC (n=75) (frequency; percentage)	Hazards ratio	p-value
Age (median: 49·6)			0·34
< 49·6y	37 (49·3)	1·00 (Reference)	
≥ 49·6y	38 (50·7)	0·55	
Stage			0·35
I-II	42 (56·0)	1·00 (Reference)	
III-IV	33 (44·0)	1·76	
Smoking History			0·49
Lifelong non-smoker	9 (12·0)	1·00 (Reference)	
Ever smoker	66 (88·0)	0·58	
Areca nut chewing History			0·34
Lifelong non-areca nut chewing	16 (21·3)	1·00 (Reference)	
Ever areca nut chewing	59 (78·7)	2·75	
Recurrence			1·44E-04
No	51 (68·0)	1·00 (Reference)	
Yes	15 (20·0)	19·7999	
N/A	9 (12·0)		

Appendix Table 5. Differential expression analysis of MMR genes between AN-related and AN-negative OSCC patients.

Gene_ID	Gene_Name	baseMean	log2FoldChange	lfcSE	stat	pvalue	padj
ENSG00000062822	*POLD1*	1946·16	1·12	0·06	20·06	1·73E-89	3·14E-88
ENSG00000049541	*RFC2*	1666·38	1·24	0·06	19·99	6·78E-89	1·21E-87
ENSG00000119684	*MLH3*	691·47	-1·29	0·07	-18·05	7·25E-73	7·92E-72
ENSG00000132646	*PCNA*	6119·27	1·04	0·07	15·60	7·14E-55	4·81E-54
ENSG00000163918	*RFC4*	1230·68	0·75	0·06	11·74	8·15E-32	2·98E-31
ENSG00000064933	*PMS1*	483·10	-0·77	0·07	-10·90	1·09E-27	3·59E-27
ENSG00000113318	*MSH3*	784·03	-0·73	0·07	-10·68	1·32E-26	4·21E-26
ENSG00000204410	*MSH5*	272·05	-0·88	0·08	-10·43	1·78E-25	5·50E-25
ENSG00000057468	*MSH4*	2·32	-2·71	0·29	-9·37	7·57E-21	2·04E-20
ENSG00000174371	*EXO1*	607·62	0·50	0·06	7·72	1·13E-14	2·52E-14
ENSG00000111445	*RFC5*	924·45	0·40	0·05	7·35	2·01E-13	4·30E-13
ENSG00000122512	*PMS2*	1177·11	0·35	0·06	6·20	5·52E-10	1·05E-09
ENSG00000133119	*RFC3*	717·51	0·25	0·07	3·54	3·97E-04	5·71E-04
ENSG00000095002	*MSH2*	1420·91	0·15	0·05	2·77	5·69E-03	7·58E-03
ENSG00000035928	*RFC1*	2496·20	-0·17	0·06	-2·76	5·82E-03	7·75E-03
ENSG00000076242	*MLH1*	1077·14	0·13	0·05	2·71	6·78E-03	8·98E-03
ENSG00000116062	*MSH6*	3038·25	0·04	0·05	0·86	3·91E-01	4·38E-01

Appendix Table 6. Copy number alterations identified in 113 OSCC patients from SCOCC study.

Unique Name	cytoband[a]	wide peak boundaries	q value	residual q value	genes in wide peak
Amplification Peak 1	**7p11.2**	chr7:48560238-56119892	1·01E-31	1·01E-31	CCT6A,DDC,EGFR,GBAS,GRB10,HPVC1,PSPH,IKZF1,ZPBP,COBL,SEC61G,MRPS17,LANCL2,FIGNL1,VOPP1,ABCA13,CDC14C,VSTM2A,POM121L12,LOC285878,14-Sep,ZNF713,FKBP9L,VWC2, FLJ45974,LOC100129427,C7orf72,
Amplification Peak 2	**11q13.3**	chr11:68777784-71192730	1·75E-19	2·07E-13	hsa-mir-548k,hsa-mir-3164,CCND1,DHCR7,CTTN, FGF3,FGF4,PPFIA1,FADD,FGF19,SHANK2,MYEOV,ANO1,NADSYN1,MRGPRF,TPCN2,ORAOV1,FLJ42102,MIR548K,MIR3664,
Amplification Peak 3	**7q22.1**	chr7:102141437-102467866	7·14E-08	7·14E-08	RASA4,FAM185A,FBXL13,POLR2J2,SPDYE2,POLR2J3,UPK3BL,SPDYE2L,
Amplification Peak 4	**11q22.2**	chr11:100847633-103027792	1·62E-07	1·59E-03	BIRC2,BIRC3,MMP1,MMP3,MMP7,MMP8,MMP10,MMP12,MMP13,PGR,TRPC6,MMP20,YAP1,KIAA1377,MMP27,DYNC2H1,TMEM133,DCUN1D5,C11orf70,TMEM123,ARHGAP42,
Amplification Peak 5	**9p13.3**	chr9:35396680-35754480	2·14E-07	2·42E-07	ANGPTL5,LOC100288077,MIR3920, CA9,CD72,RMRP,TESK1,TLN1,TPM2,RGP1,RUSC2,CREB3,UNC13B,SIT1,GBA2,C9orf100,
Amplification Peak 6	**9p24.1**	chr9:4618369-7165643	1·67E-06	1·93E-06	ATP8B5P,CCDC107,MSMP,FAM166B,MIR4667, hsa-mir-101-2,MLANA,GLDC,INSL4,JAK2,RLN1, RLN2,RCL1,INSL6,KDM4C,RANBP6,CD274,AK3,C9orf68,CDC37L1,C9orf46,KIAA1432,ERMP1,PDCD1LG2,TPD52L3,IL33,UHRF2,KIAA2026,PPAPDC2,MIR101-2,MIR4665,
Amplification Peak 7	**1q21.2**	chr1:147126767-148589439	4·48E-05	4·48E-05	GJA5,GJA8,NBPF14,ACP6,GPR89B,PPIAL4A,NBPF11,NBPF15,FLJ39739,PPIAL4D,PPIAL4B,NBPF24,GPR89C,NBPF16,PDZK1P1,PPIAL4F,
Amplification Peak 8	**2q11.2**	chr2:97493697-98263772	1·29E-04	1·59E-04	COX5B,CNNM3,ANKRD39,FAM178B,SEMA4C,ANKRD36B,FAHD2B,ANKRD23,ANKRD36,LOC100506123

Amplification Peak 9	3q27.1	chr3:183775108-184289462	4·08E-04	4·08E-04	hsa-mir-1224,AP2M1,CLCN2,DVL3,EIF4G1, EPHB3,POLR2H,PSMD2,THPO,CHRD,EIF2B5,ECE2,ALG3,ABCF3,VWA5B2,CAMK2N2,FAM131A,HTR3C,HTR3E,SNORD66,MIR1224
Amplification Peak 10	19p12	chr19:23845269-30206014	7·60E-04	9·64E-04	UQCRFS1,ZNF254,POP4,PLEKHF1,C19orf12,LOC148145,LOC148189,ZNF681,ZNF675,LOC284395,VSTM2B,RPSAP58,ZNF726,LOC100101266,LOC100505835
Amplification Peak 11	**6p12.1**	chr6:53409662-55113889	1·05E-03	1·05E-03	GCLC,HCRTR2,TINAG,LRRC1,MLIP,FAM83B,KLHL31
Amplification Peak 12	**8p11.23**	chr8:38110875-38458507	1·85E-03	1·85E-03	FGFR1,DDHD2,WHSC1L1,PPAPDC1B,LETM2,RNF5P1,C8orf86
Amplification Peak 13	**14q11.2**	chr14:22539249-22991344	3·24E-03	3·24E-03	[DAD1]
Amplification Peak 14	2p11.2	chr2:89619431-95815146	4·68E-03	7·70E-03	MAL,MRPS5,ZNF514,LOC90499,TEKT4,ACTR3BP2,LOC442028,GGT8P,LOC654342,ANKRD20A8P
Amplification Peak 15	20q11.23	chr20:33370081-36488689	2·25E-02	2·25E-02	hsa-mir-1289-1,hsa-mir-499,EPB41L1,GGT7,GHRH, GSS,EIF6,NNAT,RBL1,RPN2,SPAG4,SRC,GDF5,CPNE1,NFS1,RBM39,RBM12,MYL9,PROCR,MMP24,BLCAP,CEP250,DLGAP4,NCOA6,SAMHD1,C20orf4,TRPC4AP,PHF20,SCAND1, ERGIC3,UQCC,EDEM2,ACSS2,C20orf24,CTNNBL1,NDRG3,MYH7B,TGIF2,MANBAL,DSN1,FER1L4,SLA2,HMGB3P1,FAM83C,C20orf132,KIAA0889,C20orf118,ROMO1,C20orf173,C20orf152,MIR499A,LOC647979,LOC100287792,TGIF2-C20ORF24,MIR499B
Amplification Peak 16	22q11.21	chr22:20891808-21965620	2·25E-02	2·25E-02	hsa-mir-649,CRKL,SERPIND1,PI4KA,SLC7A4, UBE2L3,LZTR1,P2RX6,SNAP29,HIC2,POM121L8P,MED15,THAP7,TMEM191A,MGC16703,AIFM3,RIMBP3C,POM121L4P,PI4KAP2,LOC400891,BCRP2,THAP7-AS1,P2RX6P,RIMBP3B,TMEM191C
Amplification Peak 17	8q24.22	chr8:133591056-134261297	2·62E-02	2·62E-02	SLA,TG,WISP1,NDRG1,LRRC6,PHF20L1,TMEM71
Deletion Peak 1	5q31.3	chr5:140056738-140167722	2·67E-10	8·29E-10	HARS2,VTRNA1-3,VTRNA1-2,VTRNA1-1,ZMAT2
Deletion Peak 2	**4q35.2**	chr4:186384468-191154276	2·54E-09	2·54E-09	F11,FAT1,FRG1,HSP90AA4P,KLKB1,MTNR1A,TLR3,SORBS2,DUX4,FAM149A,DUX2,PDLIM3,ZFP42,TRIML2,CYP4V2,LOC285441,LOC339975,TRIML1,LOC401164,DUX4L4,FRG2,FLJ38576,DUX4L6,DUX4L5,DUX4L3,DUX4L2,LOC100288255
Deletion Peak 3	7q22.1	chr7:102113787-102467866	5·02E-09	5·02E-09	RASA4,FAM185A,POLR2J2,SPDYE2,POLR2J3,UPK3BL,SPDYE2L

Deletion Peak 4	2q34	chr2:208576795-215812597	1·07E-07	1·07E-07	hsa-mir-548f-2,ACADL,BARD1,CPS1,CRYGA, CRYGB,CRYGC,CRYGD,ERBB4,IDH1,MAP2,MYL1,PTH2R,RPE,FZD5,LANCL1,IKZF2,CPS1-IT1,SPAG16,C2orf67,PIKFYVE,UNC80,PLEKHM3,C2orf80,VWC2L,LOC100130451,MIR548F2,LOC100507443,MIR4776-1,MIR4775,MIR4776-2
Deletion Peak 5	**19p13.3**	chr19:811020-871374	6·01E-07	6·01E-07	hsa-mir-3187,AZU1,CFD,ELANE,PRTN3,LPPR3, MIR3187
Deletion Peak 6	3p12.3	chr3:74536064-88175577	6·82E-09	3·63E-06	hsa-mir-4273,hsa-mir-1324,GBE1,HTR1F,POU1F1, ROBO1,ROBO2,CGGBP1,CHMP2B,CADM2,VGLL3,LOC401074,LOC440970,FLJ20518,FAM86DP,ZNF717,FRG2C,MIR1324,MIR4273,MIR4795,MIR4444-1
Deletion Peak 7	18q22.2	chr18:64172758-71740872	3·02E-05	3·02E-05	SOCS6,CD226,RTTN,TMX3,CCDC102B,NETO1,DSEL,CBLN2,DOK6,LOC400655,LOC643542,LOC100505776,LOC100505817
Deletion Peak 8	11q24.2	chr11:123448478-124493627	5·57E-05	1·01E-04	VWA5A,ZNF202,OR8G2,OR8B8,OR8G1,OR8B2,SCN3B,PANX3,OR8B12,OR8G5,OR10G8, OR10G9,OR10S1,OR6T1,OR4D5,OR8D1,OR8D2,OR8B4,TMEM225,OR8D4,OR6X1,OR6M1, OR10G4,OR10G7,OR8B3,OR8A1
Deletion Peak 9	**2q11.1**	chr2:89105239-95815146	2·01E-04	2·01E-04	MAL,MRPS5,LOC90499,TEKT4,ACTR3BP2,LOC442028,GGT8P,LOC654342,ANKRD20A8P,MIR4436A
Deletion Peak 10	21q22.3	chr21:47636729-47879407	4·41E-05	2·01E-04	PCNT,MCM3AP,C21orf58,YBEY,MCM3AP-AS1
Deletion Peak 11	**8p23.2**	chr8:2091648-6267164	3·46E-04	3·42E-04	CSMD1,LOC100287015
Deletion Peak 12	**1q21.2**	chr1:148591523-149804881	4·45E-04	6·24E-04	FCGR1A,HIST2H4A,FAM91A2,PPIAL4A,LOC388692,HIST2H2BF,HIST2H4B,PPIAL4D,LOC645166,PPIAL4B,PPIAL4C,HIST2H3D,LOC728855,NBPF16,PPIAL4F,PPIAL4E,FCGR1C
Deletion Peak 13	**5q13.2**	chr5:67591605-68651952	1·81E-04	2·23E-03	CCNB1,CDK7,SLC30A5,CENPH,MRPS36,CCDC125
Deletion Peak 14	10q21.3	chr10:70066913-70680063	2·84E-03	2·84E-03	hsa-mir-1254,DNA2,HNRNPH3,SLC25A16, RUFY2,CCAR1,TET1,STOX1,SNORD98
Deletion Peak 15	**13q12.11**	chr13:20660350-23903889	3·44E-03	3·44E-03	FGF9,GJA3,GJB2,SGCG,IFT88,SAP18,GJB6,LATS2,CRYL1,IL17D,XPO4,MRP63,N6AMT2, SKA3,EFHA1,ZDHHC20,BASP1P1,MIR4499

					hsa-mir-31,CDKN2A,CDKN2B,ELAVL2,
Deletion Peak 16	9p21.2	chr9:21305545-26917503	3·07E-03	4·54E-03	IFNA1,IFNA2,IFNA6,IFNA8,IFNA13,MTAP,C9orf53,KLHL9,DMRTA1,C9orf82,TUSC1,IFN
					E,MIR31,FLJ35282,MIR31HG,CDKN2B-AS1,LOC100506422
Deletion Peak 17	**11p15.4**	chr11:9445758-9595741	2·00E-02	2·00E-02	ZNF143,LOC644656,SNORA23
					hsa-mir-935,hsa-mir-373,NDUFA3,PRKCG,
Deletion Peak 18	19q13.42	chr19:54261807-54618960	1·35E-02	2·25E-02	CACNG8,CACNG7,CACNG6,NLRP12,MYADM,OSCAR,VSTM1,TARM1,MIR371A,MIR372
					,MIR373,MIR516A2,MIR519A2,MIR935,MIR371B

Appendix Table 7. Primers for quantitative real-time PCR (qPCR) assays and sequences of siRNAs.

Name	Primer sequence (5'-3')
ATG2A -qPCR_F	ACTCGCTGCTGAAGATGACC
ATG2A -qPCR_R	TCCGTGTACTCAGGCTCAGA
ATG9A -qPCR_F	GCTGTTCCTGAGGTGGTCAA
ATG9A -qPCR_R	GTGCAATACGGAAGGGCAGA
WEE1 -qPCR_F	CGATATTTCTCTGCGTGGGC
WEE1 -qPCR_R	CACATACCACTGTGAGGGCA
GAPDH -qPCR_F	AGCCACATCGCTCAGACAC
GAPDH -qPCR_R	GCCCAATACGACCAAATCC
MLH1 -qPCR_F	CTCTTCATCAACCATCGTCTGG
MLH1 -qPCR_R	GCAAATAGGCTGCATACACTGTT
MSH2 -qPCR_F	ATCCAGGCATGCTTGTGTTG
MSH2 -qPCR_R	CTTCACCTGATAAAGCATAG
WEE1 siRNA1	CCUCAGGACAGUGUCGUCGUAGAAA
WEE1 siRNA2	GGCAUGAAAUCAGACAGGGUAGAUU
ATG2A siRNA1	CCGAAGACCUGUGGCUGAUUGAGCA
ATG2A siRNA2	GAGCUGAAGCUAAAGCGGCUCUGUU

Appendix Table 8. Statistics of whole-exome sequencing for 113 tumor-blood paired OSCC patients from SCOCC study.

Sample ID	Sample Type	Mapped reads	Mapping Rate (%)	Reads in target rate (%)	Coverage (X)
Sample_N001	blood	73,617,465	99·87%	83·09%	120·144
Sample_N002	blood	67,795,991	99·65%	80·50%	105·237
Sample_N003	blood	66,765,050	99·68%	81·03%	104·548
Sample_N004	blood	73,865,458	99·92%	83·62%	121·591
Sample_N005	blood	77,001,535	99·57%	79·05%	116·460
Sample_N006	blood	73,851,151	99·89%	81·29%	116·707
Sample_N007	blood	74,194,394	99·90%	82·01%	118·272
Sample_N008	blood	73,845,454	99·62%	79·86%	113·225
Sample_N009	blood	77,981,160	99·91%	81·54%	124·101
Sample_N010	blood	75,323,087	99·88%	79·81%	115·233
Sample_N011	blood	74,587,459	99·90%	81·91%	118·843
Sample_N012	blood	74,208,412	99·68%	81·36%	117·388
Sample_N013	blood	74,711,346	99·87%	81·77%	119·136
Sample_N014	blood	74,965,301	99·90%	82·90%	122·708
Sample_N015	blood	78,328,582	99·86%	80·70%	121·713
Sample_N016	blood	76,137,524	99·69%	81·26%	120·018
Sample_N017	blood	69,141,017	99·69%	81·54%	109·542
Sample_N018	blood	75,077,599	99·86%	79·44%	114·761
Sample_N019	blood	77,995,718	99·92%	82·65%	126·615
Sample_N020	blood	77,986,213	99·88%	80·94%	121·799
Sample_N021	blood	75,552,763	99·90%	80·82%	118·186
Sample_N022	blood	74,065,319	99·88%	81·35%	117·026
Sample_N023	blood	72,117,579	99·74%	81·82%	115·108
Sample_N024	blood	70,207,834	99·72%	82·72%	113·773
Sample_N025	blood	75,731,715	99·87%	81·37%	119·475
Sample_N026	blood	76,109,407	99·90%	80·38%	118·471
Sample_N027	blood	76,139,605	99·87%	82·91%	123·886
Sample_N028	blood	67,668,610	99·70%	81·28%	106·953
Sample_N029	blood	75,350,151	99·67%	81·03%	118·274
Sample_N030	blood	76,677,968	99·87%	82·04%	122·679
Sample_N031	blood	76,942,014	99·87%	81·66%	122·482
Sample_N032	blood	76,727,178	99·85%	83·17%	125·449
Sample_N033	blood	72,779,610	99·91%	81·71%	116·086
Sample_N034	blood	66,863,523	99·66%	81·24%	105·188
Sample_N035	blood	77,246,687	99·87%	81·62%	122·313
Sample_N036	blood	79,043,428	99·86%	81·01%	123·890
Sample_N037	blood	77,460,696	99·76%	80·60%	120·561

Sample_N038	blood	72,168,086	99·71%	79·95%	111·088
Sample_N039	blood	77,505,703	99·91%	81·60%	123·036
Sample_N040	blood	101,330,380	99·87%	80·55%	156·492
Sample_N041	blood	96,532,469	99·87%	80·83%	150·231
Sample_N042	blood	106,197,504	99·87%	79·03%	162·975
Sample_N043	blood	95,814,598	99·76%	79·81%	144·852
Sample_N044	blood	99,681,300	99·77%	81·16%	153·524
Sample_N045	blood	73,300,742	99·71%	77·67%	108·160
Sample_N046	blood	91,551,299	99·75%	80·03%	138·226
Sample_N047	blood	101,578,083	99·74%	80·22%	153·682
Sample_N048	blood	107,028,563	99·77%	80·92%	164·504
Sample_N049	blood	107,095,543	99·76%	81·24%	165·597
Sample_N050	blood	92,111,506	99·74%	80·29%	139·102
Sample_N051	blood	115,272,590	99·75%	80·52%	175·075
Sample_N052	blood	86,117,546	99·71%	81·08%	131·581
Sample_N053	blood	97,998,792	99·75%	80·67%	149·674
Sample_N054	blood	99,656,094	99·77%	81·12%	153·420
Sample_N055	blood	102,561,303	99·76%	80·65%	156·992
Sample_N056	blood	94,951,885	99·77%	80·86%	146·139
Sample_N057	blood	102,964,309	99·77%	80·37%	156·721
Sample_N058	blood	97,236,476	99·77%	80·37%	148·145
Sample_N059	blood	114,087,072	99·81%	82·43%	181·931
Sample_N060	blood	105,901,364	99·77%	81·44%	165·159
Sample_N061	blood	110,815,591	99·79%	81·66%	173·572
Sample_N062	blood	110,642,167	99·77%	77·12%	166·554
Sample_N063	blood	114,246,077	99·78%	79·27%	177·062
Sample_N064	blood	97,249,091	99·77%	79·59%	149·087
Sample_N065	blood	84,213,416	99·77%	81·58%	132·686
Sample_N066	blood	94,849,701	99·76%	80·04%	145·348
Sample_N067	blood	92,888,127	99·88%	80·60%	142·518
Sample_N068	blood	91,495,511	99·88%	80·22%	138·750
Sample_N069	blood	95,864,708	99·86%	82·06%	149·939
Sample_N070	blood	91,367,181	99·75%	80·60%	141·699
Sample_N071	blood	101,440,204	99·74%	81·18%	159·123
Sample_N072	blood	114,780,380	99·80%	80·65%	176·385
Sample_N073	blood	104,260,861	99·80%	80·38%	158·774
Sample_N074	blood	100,811,800	99·78%	80·68%	154·507
Sample_N075	blood	105,752,261	99·82%	80·33%	161·540
Sample_N076	blood	84,366,375	99·76%	78·99%	124·586
Sample_N077	blood	90,465,266	99·82%	81·38%	140·680
Sample_N078	blood	113,038,880	99·83%	79·26%	170·682
Sample_N079	blood	120,436,683	99·82%	80·08%	182·918
Sample_N080	blood	102,785,101	99·82%	80·95%	158·564
Sample_N081	blood	110,128,573	99·80%	78·96%	163·230

Sample_N082	blood	98,118,390	99·75%	80·50%	148·619
Sample_N083	blood	82,353,016	99·74%	79·45%	121·912
Sample_N084	blood	111,699,689	99·78%	81·83%	173·956
Sample_N085	blood	98,344,459	99·74%	80·51%	149·650
Sample_N086	blood	109,684,249	99·77%	79·94%	165·277
Sample_N087	blood	99,271,355	99·77%	81·17%	153·133
Sample_N088	blood	97,212,737	99·79%	82·00%	152·468
Sample_N089	blood	99,791,982	99·78%	81·30%	154·302
Sample_N090	blood	104,819,256	99·76%	80·68%	160·142
Sample_N091	blood	99,840,053	99·86%	79·12%	153·702
Sample_N092	blood	97,923,740	99·87%	79·07%	150·344
Sample_N093	blood	99,799,420	99·86%	78·69%	151·300
Sample_N094	blood	103,161,277	99·79%	78·38%	154·963
Sample_N095	blood	91,284,792	99·81%	80·21%	141·474
Sample_N096	blood	105,122,562	99·80%	80·38%	163·303
Sample_N097	blood	104,707,663	99·80%	80·41%	162·188
Sample_N098	blood	105,182,866	99·78%	79·12%	158·381
Sample_N099	blood	144,785,433	99·82%	80·34%	223·837
Sample_N100	blood	131,993,559	99·73%	82·00%	208·629
Sample_N101	blood	104,480,008	99·80%	80·63%	162·092
Sample_N102	blood	113,480,077	99·80%	80·55%	175·847
Sample_N103	blood	85,374,410	99·79%	80·96%	133·223
Sample_N104	blood	103,199,243	99·79%	80·82%	160·559
Sample_N105	blood	81,165,044	99·78%	81·69%	127·866
Sample_N106	blood	85,671,076	99·78%	80·20%	131·573
Sample_N107	blood	115,137,233	99·79%	80·67%	178·942
Sample_N108	blood	109,741,093	99·74%	80·27%	168·694
Sample_N109	blood	102,826,071	99·76%	79·89%	156·806
Sample_N110	blood	101,823,365	99·77%	81·00%	155·617
Sample_N111	blood	97,072,253	99·82%	77·73%	141·443
Sample_N112	blood	104,491,378	99·75%	81·59%	163·341
Sample_N113	blood	92,133,005	99·73%	80·93%	141·189
Sample_T001	tumor	70,009,146	99·81%	81·61%	113·032
Sample_T002	tumor	64,543,384	99·77%	81·85%	103·756
Sample_T003	tumor	76,633,048	99·80%	80·53%	119·944
Sample_T004	tumor	66,149,072	99·74%	82·33%	107·015
Sample_T005	tumor	70,458,275	99·89%	81·35%	111·185
Sample_T006	tumor	72,387,599	99·72%	81·61%	115·420
Sample_T007	tumor	70,457,227	99·76%	82·55%	114·596
Sample_T008	tumor	71,053,168	99·90%	83·19%	116·938
Sample_T009	tumor	70,720,028	99·91%	83·50%	117·287
Sample_T010	tumor	73,489,334	99·81%	80·82%	115·102
Sample_T011	tumor	70,879,330	99·73%	81·20%	112·247
Sample_T012	tumor	70,679,940	99·74%	81·01%	110·898

Sample_T013	tumor	73,098,190	99·76%	80·84%	115·504
Sample_T014	tumor	68,044,762	99·78%	81·73%	109·192
Sample_T015	tumor	69,663,558	99·87%	82·61%	112·107
Sample_T016	tumor	70,832,493	99·76%	80·56%	112·081
Sample_T017	tumor	70,867,325	99·74%	81·65%	112·869
Sample_T018	tumor	71,637,312	99·90%	81·49%	114·579
Sample_T019	tumor	70,844,455	99·74%	82·73%	115·856
Sample_T020	tumor	72,083,178	99·90%	82·48%	118·434
Sample_T021	tumor	67,536,051	99·79%	80·85%	107·118
Sample_T022	tumor	70,391,238	99·81%	81·32%	112·106
Sample_T023	tumor	70,716,146	99·77%	81·63%	113·825
Sample_T024	tumor	72,282,785	99·87%	82·61%	116·223
Sample_T025	tumor	71,726,420	99·83%	81·96%	115·307
Sample_T026	tumor	71,340,144	99·78%	80·31%	110·911
Sample_T027	tumor	71,944,537	99·82%	81·94%	115·593
Sample_T028	tumor	73,096,105	99·91%	84·58%	124·050
Sample_T029	tumor	67,406,510	99·82%	81·73%	108·875
Sample_T030	tumor	72,842,400	99·91%	83·54%	119·691
Sample_T031	tumor	78,434,465	99·72%	81·30%	124·096
Sample_T032	tumor	75,604,621	99·81%	81·22%	120·306
Sample_T033	tumor	78,744,911	99·80%	82·17%	128·136
Sample_T034	tumor	74,682,622	99·89%	82·50%	120·551
Sample_T035	tumor	72,381,951	99·82%	81·25%	114·758
Sample_T036	tumor	75,293,660	99·67%	81·09%	118·376
Sample_T037	tumor	73,249,280	99·91%	83·07%	119·433
Sample_T038	tumor	73,325,480	99·90%	81·22%	115·763
Sample_T039	tumor	61,847,055	99·76%	81·02%	98·742
Sample_T040	tumor	90,286,688	99·85%	78·37%	132·348
Sample_T041	tumor	73,977,106	99·67%	73·42%	103·181
Sample_T042	tumor	107,835,222	99·87%	79·06%	160·286
Sample_T043	tumor	96,832,012	99·89%	78·15%	143·711
Sample_T044	tumor	106,139,404	99·84%	77·97%	155·774
Sample_T045	tumor	94,762,729	99·85%	78·30%	139·716
Sample_T046	tumor	112,125,617	99·86%	78·59%	172·141
Sample_T047	tumor	92,611,000	99·84%	77·90%	135·334
Sample_T048	tumor	104,951,750	99·83%	78·46%	156·368
Sample_T049	tumor	105,551,400	99·84%	78·51%	157·611
Sample_T050	tumor	99,718,159	99·86%	79·50%	149·414
Sample_T051	tumor	99,165,517	99·86%	78·01%	144·059
Sample_T052	tumor	105,183,652	99·86%	80·25%	161·869
Sample_T053	tumor	95,700,330	99·85%	78·55%	142·189
Sample_T054	tumor	96,497,431	99·85%	78·70%	142·110
Sample_T055	tumor	118,270,367	99·85%	78·60%	175·005
Sample_T056	tumor	99,773,399	99·85%	79·14%	149·729

Sample_T057	tumor	107,840,968	99·82%	78·85%	160·781
Sample_T058	tumor	97,395,037	99·86%	78·70%	143·069
Sample_T059	tumor	106,172,535	99·84%	79·59%	161·899
Sample_T060	tumor	96,783,893	99·86%	78·43%	144·178
Sample_T061	tumor	100,579,784	99·90%	77·22%	148·707
Sample_T062	tumor	93,676,035	99·81%	76·98%	136·278
Sample_T063	tumor	90,268,343	99·85%	77·32%	135·119
Sample_T064	tumor	95,439,787	99·82%	77·27%	139·029
Sample_T065	tumor	91,536,359	99·82%	77·43%	133·299
Sample_T066	tumor	103,945,396	99·83%	78·19%	153·327
Sample_T067	tumor	96,044,915	99·87%	79·76%	145·177
Sample_T068	tumor	88,318,791	99·87%	77·46%	129·010
Sample_T069	tumor	97,713,298	99·82%	77·94%	143·608
Sample_T070	tumor	86,821,769	99·86%	78·89%	128·080
Sample_T071	tumor	88,567,408	99·82%	79·61%	134·707
Sample_T072	tumor	90,621,515	99·87%	78·50%	134·791
Sample_T073	tumor	96,571,307	99·85%	79·41%	146·150
Sample_T074	tumor	108,922,968	99·84%	79·15%	164·365
Sample_T075	tumor	95,757,614	99·83%	77·69%	140·834
Sample_T076	tumor	92,350,943	99·85%	76·88%	137·537
Sample_T077	tumor	90,687,891	99·87%	78·45%	132·994
Sample_T078	tumor	105,352,836	99·88%	79·67%	159·837
Sample_T079	tumor	105,097,088	99·87%	77·07%	152·211
Sample_T080	tumor	94,571,341	99·82%	78·02%	138·636
Sample_T081	tumor	95,280,622	99·86%	77·09%	141·896
Sample_T082	tumor	92,447,438	99·84%	79·45%	139·359
Sample_T083	tumor	88,362,475	99·87%	79·17%	131·059
Sample_T084	tumor	73,002,478	99·87%	74·62%	103·995
Sample_T085	tumor	86,756,348	99·84%	77·73%	125·691
Sample_T086	tumor	101,903,738	99·85%	79·02%	153·519
Sample_T087	tumor	98,854,373	99·85%	76·95%	143·046
Sample_T088	tumor	87,125,257	99·88%	76·18%	123·161
Sample_T089	tumor	86,084,303	99·68%	77·02%	122·881
Sample_T090	tumor	96,274,201	99·82%	77·53%	140·610
Sample_T091	tumor	77,412,443	99·84%	79·19%	115·064
Sample_T092	tumor	94,956,157	99·71%	77·56%	136·974
Sample_T093	tumor	99,547,739	99·85%	79·05%	147·697
Sample_T094	tumor	89,500,765	99·84%	79·41%	133·263
Sample_T095	tumor	93,893,288	99·85%	79·79%	142·766
Sample_T096	tumor	115,731,444	99·85%	77·53%	168·585
Sample_T097	tumor	101,327,272	99·87%	78·84%	150·623
Sample_T098	tumor	95,847,507	99·86%	79·13%	141·981
Sample_T099	tumor	88,656,082	99·86%	77·59%	129·141
Sample_T100	tumor	105,221,953	99·86%	78·71%	156·813

Sample_T101	tumor	99,352,904	99·85%	79·09%	149·662
Sample_T102	tumor	89,272,305	99·85%	79·18%	131·786
Sample_T103	tumor	92,820,670	99·84%	79·51%	140·534
Sample_T104	tumor	81,854,228	99·85%	78·35%	119·220
Sample_T105	tumor	96,171,348	99·86%	77·04%	142·811
Sample_T106	tumor	107,870,541	99·84%	77·82%	157·789
Sample_T107	tumor	96,249,192	99·85%	77·10%	141·566
Sample_T108	tumor	97,118,677	99·88%	76·79%	140·690
Sample_T109	tumor	101,653,804	99·83%	77·68%	149·821
Sample_T110	tumor	98,552,591	99·85%	77·56%	144·957
Sample_T111	tumor	91,116,229	99·82%	78·05%	134·457
Sample_T112	tumor	104,518,348	99·84%	78·04%	151·473
Sample_T113	tumor	96,837,735	99·80%	71·14%	130·135