Genomic signature of mismatch repair deficiency in areca nut-related oral cancer

Wei-fa Yang[1#]; Na Qin[2,3#]; Xueyao Song[2,3#]; Canhua Jiang[4#]; Tingting Li[4]; Pei Ji[2,3]; Yuancheng Li[2,3]; Dongsheng Ding[2,3]; Cheng Wang[2,3,5]; Juncheng Dai[2,3]; Guangfu Jin[2,3]; Ting-Wen Chen[6]; Yu-Sun Chang[6, 7]; Dai-qiao Ouyang[8]; Gui-qing Liao[8]; Zhibin Hu[2,3]; Kai-Ping Chang[6, 7*]; Yu-xiong Su[1*]; Hongxia Ma[2,3*]

[1]Oral and Maxillofacial Surgery, Faculty of Dentistry, The University of Hong Kong, Hong Kong, China.

[2]Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing 211166, China.

[3]Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Medicine, Nanjing Medical University, Nanjing 211166, China.

[4]Department of Oral and Maxillofacial Surgery, Xiangya Hospital, Central South University, Changsha 410008, China.

[5]Department of Bioinformatics, School of Biomedical Engineering and Informatics, Nanjing Medical University, Nanjing 211116, China.

[6]Molecular Medicine Research Center, Chang Gung University, Taoyuan 33302, Taiwan.

[7]Department of Otolaryngology-Head and Neck Surgery, Chang Gung Memorial Hospital, Taoyuan 33305, Taiwan.

[8]Department of Oral and Maxillofacial Surgery, Guanghua School of Stomatology, Sun Yat-sen University, Guangzhou 510275, China.

[#]These authors contributed equally to this work.

Correspondence: Hongxia Ma, 818 Tianyuan East Road, Nanjing 211166, China, Tel: +86-25-8686-8440, E-mail: hongxiama@njmu.edu.cn; Yu-xiong Su, 2F, Prince Philip Dental Hospital, 34 Hospital Road, Hong Kong, Tel: +852-2859-0267, E-mail: richsu@hku.hk; and Kai-Ping Chang, 5 Fu-Hsing street, Kwei-Shan, Taoyuan 33305, Taiwan. Tel: +886-33281200-3967, E-mail: dr.kpchang@gmail.com.

Author Contribution

H.M., Y.S., and Z.H. designed the study; W.Y., C.J., T.L., P.J., Y.L., D.D., and X.S. assisted in, or provided support for, data collection and recruitment of study participants; G.J., Y.C., D.O., and G.L. assisted in, or supervised, sample collection and processing; P.J., Y.L., and X.S. supervised and/or performed DNA extraction, library preparation and sequencing; N.Q., C.W., J.D., and T.C. implemented sequencing analysis pipeline; N.Q., C.W., and J.D. developed the statistical analysis plan; N.Q. performed the statistical analyses; N.Q., H.M., and Y.S. wrote the first draft of the manuscript. All authors contributed to data interpretation and writing and approved the final manuscript for publication.

Acknowledgements

We thank all the subjects who took part in this study. The authors would like to thank all the research staffs and students who participated in this work. This work was supported by National Natural Science Foundation of China (81672678, 81473048, 81302361), Science Foundation for Distinguished Young Scholars of Jiangsu (BK20160046), Innovation Project for Graduate Student of Jiangsu Province (KYLX16_1121), Key Grant of Natural Science Foundation of Jiangsu Higher Education Institutions (16KJA330003), a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (Public Health and Preventive Medicine), Top-notch Academic Programs Project of Jiangsu Higher Education Institutions (PPZY2015A067), Hong Kong Research Grant Council General Research Fund (17120718), and HKU Seed Fund for Basic Research (201711159117).

Competing interests

The authors declare no potential conflicts of interest.

Running Head:

DNA Mismatch Repair Deficiency in Oral Cancer

Key words: oral cancer, areca nut chewing, whole-exome sequencing, mismatch repair deficiency, prognosis

Abstract word count: 230

Total word count: 3187

Total number of tables/figures: 5

Number of references: 40

Abstract

Areca nut (AN) chewing contributes to an increasing of oral squamous cell carcinoma (OSCC) cases in South and Southeast Asia; however, genomic events underlying the carcinogenesis process of AN-related OSCC remain unclear. Here, we comprehensively described the genomic and transcriptome alterations of 113 Chinese OSCC patients (89 AN-related and 24 AN-negative) by whole-exome sequencing and RNA sequencing, and compared the genomic differences between AN-related and AN-negative samples by integrating sequencing data of 325 OSCC patients from The Cancer Genome Atlas database and 50 from a published Taiwanese study. We identified eleven significantly mutated genes for OSCC, including four novel ones (*ATG2A*, *WEE1*, *DST*, and *TSC2*), of which *WEE1* and *ATG2A* mutated with significantly higher rates in AN-related samples (P=0.04 and P=0.003, respectively). Mutational signature analysis revealed that AN-related OSCC were specially characterized by the genomic signature of mismatch repair deficiency (dMMR), which could also predict the prognosis status of AN-related OSCC. In addition, an elevated *PD-L1* expression was also observed in both AN-related patients ($P=3.71\times10^{-11}$) and those with a high dMMR level ($P=1.99\times10^{-4}$). Further differential expression analysis and *in vitro* experiments confirmed the role of dMMR in the development of OSCC induced by AN exposure. Taken together, this study first revealed the molecular profiles and highlighted the role of dMMR in AN-related OSCC among Chinese population, and identified that AN-related OSCC may represent a potential cohort for effective anti-PD-1/L1 immunotherapy.

Introduction

Oral cancer is one of the most commonly diagnosed cancers and a major cause of cancer-related mortality worldwide. Squamous cell carcinoma (OSCC) is the main histological type of oral cancer, accounting for more than 90% of malignancies affecting the oral cavity. The prevalence of oral cancer is high globally, while Asia accounts for about 60% of oral cancer cases, especially in South and Southeast Asia (48.7%) (Shield et al. 2017). The varied incidences of oral cancer are affected by both genetic and exposure factors (Su et al. 2017). In addition to the main risk factors, including tobacco smoking, alcohol drinking, and human papillomavirus (HPV) infection, areca nut (AN) chewing is also considered an etiological factor of oral cancer, mostly prevalent in Southeast Asia and the Indian subcontinent (Shield et al. 2017). AN chewing has also been a popular dietary habit in some areas of China for many years, such as Hunan Province, Hainan Province and Taiwan, where AN-related oral cancer cases increase rapidly in recent years (Hu et al. 2017).

Areca nut is considered the most widely used addictive substance in the world after nicotine, ethanol, and caffeine. Studies have shown that areca nut extract (ANE) and its alkaloids can induce cytotoxicity and genotoxicity to provoke carcinogenic effects (Khan et al. 2015). Regular AN chewing causes chronic irritation and inflammation of oral epithelial cells and lead to oral submucous fibrosis , which is the premalignant condition specific to AN-related OSCC (Prabhu et al. 2014). OSCC patients with AN chewing habit often have a more metastatic phenotype (Mortazavi et al. 2014) and a higher mortality (Wu et al. 2015), suggesting a distinct molecular profile

of AN-related OSCC. Although the causal link between AN and OSCC has been well established in animal and epidemiological studies (Jeng et al. 2001), the underlying mechanisms and genomic alterations remain largely unknown.

In recent years, high-throughput technologies have accelerated the study of genomic alterations in oral carcinogenesis. Genes involved in the cell cycle (*TP53* and *CDKN2A*) (Pickering et al. 2013), Notch signaling (*NOTCH1*) (Agrawal et al. 2011), and apoptosis (*CASP8*) (Pickering et al. 2013) pathways, as well as APOBEC-associated mutational signature (Chen et al. 2017) have been identified to play key roles in the development of oral cancer. However, no study has ever gained insight into the genomic landscape of AN-related OSCC.

In this study, we performed a comprehensive study of the spectrum of genomic alterations in 113 (89 AN-related and 24 AN-negative) Chinese OSCC patients by whole-exome sequencing and RNA sequencing. We also accessed the difference of mutational profiles between AN-related and AN-negative patients by integrating data of 325 OSCC patients from The Cancer Genome Atlas (TCGA) and 50 OSCC patients from a published Taiwanese study (Chen et al. 2017).

Materials and Methods

Details of methods can be found in **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). Among all patients, 89 had a history of AN chewing, which was defined as the behavioral use of AN for at least three years. Sequencing data of 325 OSCC patients (AN-negative) from TCGA and 50 OSCC patients (43 AN-related) from a published Taiwan study (Chen et al. 2017) was integrated to compare the genomic and transcriptome features of patients with different AN exposure. Detailed information is provided in **Appendix Methods** and **Appendix Table 1**.

Whole-exome sequencing

Genomic DNA was extracted from 113 frozen oral tissues and paired blood samples. Exome capture was performed by Agilent SureSelect Human All Exon V6 followed by the sequencing on the Illumina HiSeq X Ten instrument. The sequencing reads mapping and following analyses were described in **Appendix Methods**.

RNA sequencing

Total RNA was extracted from 46 matched tumor-normal samples (40 AN-related and 6 AN-negative), and was subjected for paired-end sequencing. Reads alignment and following analyses were described in **Appendix Methods**.

Cell viability assessment in vitro

Cell viability and migration were measured with a Cell Counting Kit-8 (CCK8, Dojindo, Japan) and Costar Transwell plates (6.5mm diameter insert, 8.0 mm pore size,

polycarbonate membrane, Corning Sparks, MD), respectively. Details were provided in **Appendix Methods**

Results

Clinical characteristics and mutation burden of OSCC patients

Approximately 98.2% (111/113) of the patients were males, with a median age of 49.3 at diagnosis. Of all patients, 78.8% (89/113) reported a history of AN chewing, 88.5% (100/113) reported smoking, and 67.26% (76/113) reported both. Various anatomical sites were involved in this study, including tongue (50.4%), buccal mucosa (34.5%), floor of the mouth (4.4%), alveolar ridge (9.7%), and hard palate (0.9%) (**Appendix Table 1**).

For WES of 113 patients from SCOCC, the average sequence coverage was 132.92× for tumor samples and 142.54× for blood samples. A total of 19,613 somatic variants were detected, including 18,641 single-nucleotide variants and 972 small insertions and deletions. Of these variants, a median of 22 synonymous and 58 non-synonymous mutations were identified per tumor-blood pair (ranging from 16 to 1081). The average somatic mutation rate was 2.73 per megabase (MB), which was comparable to that in previous studies (Agrawal et al. 2011; McKay et al. 2017) (**Fig. 1A, Appendix Figure 1**). No significant difference in mutation rates was observed between patients with different AN chewing or smoking status (**Appendix Figure 2**).

Specific mutational spectrum of areca nut-related OSCC patients

To identify candidate driver genes for OSCC patients, we applied the MutSigCV and IntOGen frameworks to 113 OSCC patients and 89 AN-related OSCC patients, respectively. Overall, we identified eleven significantly mutated genes (SMGs), including seven genes (*TP53*, *NOTCH1*, *FAT1*, *HRAS*, *CASP8*, *CDKN2A*, and *RASA1*) reported in previous studies, two novel genes (*DST* and *TSC2*) with similar mutation frequency in AN-related and AN-negative patients, and two novel genes (*ATG2A* and *WEE1*) with significantly higher mutation frequencies in AN-related patients (*WEE1*: P=0.04; *ATG2A*: P=0.003) (**Fig. 1B, Appendix Tables 2-3**).

When comparing the mutational pattern of previously reported SMGs between 89 AN-related patients from SCOCC and 340 AN-negative patients (316 from TCGA and 24 from SCOCC), we observed distinct mutation patterns for three genes (*TP53*, *FAT1* and *CDKN2A*) (**Fig. 1B**). Although *TP53* bore comparable mutation rates in both two groups, mutations in AN-related patients were accumulated on the p53 binding domain (**Fig. 1C**). In addition, frequencies of two recurrent nonsense mutations of *TP53* were significantly higher among AN-related patients (p.C124*: P=0.04; p.R342*: P=0.03) (**Fig. 1C**), and p.C124* has never been reported in OSCC or head and neck squamous cell carcinoma (HNSCC). For *FAT1* and *CDKN2A*, although the overall mutation rates in AN-related patients were significantly lower (*FAT1*: P=0.01; *CDKN2A*: P=2.73×10^{-5}) (**Fig. 1B**), inactivating mutations were highly clustered. For example, all *CDKN2A* mutations in AN-related samples occurred at a single recurrent inactivating position p.R58*, while *CDKN2A* mutations in AN-negative samples were sporadic (**Fig. 1D**).

Of the two novel identified genes, two missense mutations (p.E486D and p.P517L

in the Pkinase domain) in *WEE1* (**Appendix Figure 3A**) and four mutations (p.R1068W, p.R1780W, p.C1363Y, and p.R1848W) in *ATG2A* (**Appendix Figure 3B**) were predicted to be deleterious by SIFT, Polyphen2, and MutationTaster. In addition, mutations in *ATG2A* (AN-related: P=0.002; AN-negative: P=3.42×10^{-6}) and another *ATG2* homologue *ATG2B* (AN-related: P=0.14; AN-negative: P=1.55×10^{-7}) occurred mutually exclusive with *CASP8* mutations in both AN-related and AN-negative patients (**Fig. 1B**, **Appendix Figure 4**). Further differential expression analysis revealed that the expression of *WEE1* and *ATG2* homologues were elevated in AN-related patients compared to AN-negative patients (*WEE1*: P=1.55×10^{-4}; *ATG2* homologues: P=7.60×10^{-18}) (**Appendix Figure 5**).

To validate above findings, we performed the same analysis in 50 Taiwanese OSCC samples (43 AN-related and 7 AN-negative). Because of the small sample size, we detected no *WEE1* mutation and only one deleterious variant in *ATG2A* (p.L1847Q) in AN-related patient; however, we also identified that mutations in *ATG2* homologues (*ATG2A* and *ATG2B*) occurred mutually exclusive with *CASP8* mutations (P=0.01) (**Appendix Figure 4**).

We then performed *in vitro* assays in the CAL27 cell line, and observed an attenuated proliferation and migration ability of cancer cells after knocking down the expression of *WEE1* (**Fig. 2A-C**) and *ATG2A* (**Fig. 2D-F**), suggesting a tumor-promoting role of these two genes in OSCC. Additionally, since Atg9 is an interacting factor for Atg2 in autophagosome completion (Gomez-Sanchez et al. 2018), we further accessed the effect of ANE on these two autophagy markers. The expression of *ATG2A*

and *ATG9A* was significantly elevated in the CAL27 cell line after a low-dose and non-toxic ANE (1 or 2 ug/mL) treatment for 5 days by RT-qPCR/ Western blot (**Appendix Figure 6**).

Mismatch repair deficiency (dMMR) mutational signature was predominant in areca nut-related OSCC patients and associated with the expression of *PD-L1*

We applied the non-negative matrix factorization (NMF) method on all mutations identified in 113 OSCC patients and detected three prominent mutational signatures: COSMIC Signature 1, which is caused by the spontaneous deamination of 5-methylcytosine; COSMIC Signature 2/13, an APOBEC-driven hypermutation phenotype; and COSMIC Signature 15, characterized by C>T at GpCpN trinucleotides and caused by dMMR (**Fig. 3A, Appendix Figure 7-8**). Given that different mutational signatures may be caused by the same etiological factor, we deconstructed all 30 COSMIC mutational signatures and combined four dMMR-related signatures (Signature 6, 15, 20 and 26) together. When compared to TCGA data, we found that the proportion of dMMR signature was significantly higher in 89 AN-related patients than that in 340 AN-negative patients (P=2.56×10^{-12}) (**Fig. 3B-D**). These findings were supported by data from 50 Taiwanese OSCC patients where the proportion of dMMR signature was also higher in AN-related samples (**Appendix Figure 9**), even though the significant level wasn't reached because of a small sample size.

Then, we evaluated the correlation between dMMR signature and the mutation burden. We observed a positive correlation between the proportion of dMMR signature

and mutation burden in 113 OSCC patients (Cor=0.20, *P*=0.03) (**Appendix Figure 10A**), and the correlation was stronger in 89 AN-related patients (Cor=0.39, $P=1.38\times10^{-4}$) (**Appendix Figure 10B**). As dMMR may lead to a microsatellite instability (MSI) phenotype, we further examined the MSI status in all 113 samples, and identified that the individual MSI status was significantly correlated with dMMR mutations (Cor=0.22, *P*=0.02).

Since recent studies have demonstrated the relationship between dMMR, expression of programmed death ligand 1 (PD-L1), and the response to programmed death 1 (PD-1) blockade (Lee and Le 2016), we then compared the expression level of *PD-L1* in patients with different AN exposure or dMMR levels. Interestingly, we found that *PD-L1* expression was significantly increased in patients with AN chewing habit ($P=3.71\times10^{-11}$, **Fig. 3E**) or with high dMMR levels (dMMR signature proportion≥50%) ($P=1.99\times10^{-4}$, **Fig. 3F**).

dMMR signature was a prognostic marker for OSCC

We performed survival analysis in 75 OSCC patients with available follow-up data from SCOCC study (**Appendix Table 4**). Eleven patients (14.67%) died from OSCC during the follow-up period and the median survival time (MST) was 22.87 months. Seventy-five patients were divided into either high- or low-dMMR group based on the median proportion of combined-dMMR signatures (Signature 6, 15, 20 and 26) or Signature 15, which was predominant in our samples. We identified that higher dMMR signature was associated with a worse overall survival (OS) (combined-dMMR

signatures: HR=8.90, *P*=0.038; Signature 15: HR=8.17, *P*=0.026) (**Fig. 4A, Appendix Figure 11A**). The results remained significant when only 59 AN-related patients were analyzed (combined-dMMR signatures: HR=8.33, *P*=0.042; Signature 15: HR=7.71, *P*=0.029) (**Fig. 4B, Appendix Figure 11B**). The same analysis was conducted in 311 TCGA OSCC patients with follow-up longer than one month. However, no correlation was observed because of the relatively low proportion of dMMR signatures (**Appendix Figure 12**).

DNA mismatch repair was repressed after arecoline treatment in human epithelial cell

To investigate whether the overrepresentation of dMMR mutational signature in AN-related OSCC samples was attributed to the alteration of DNA mismatch repair (MMR) genes, we performed differential expression analysis among 40 AN-related (SCOCC) and 326 AN-negative patients (320 form TCGA and 6 from SCOCC). The result revealed that ten MMR genes (*MLH1*, *MSH2*, *PMS2*, *POLD1*, *EXO1*, *PCNA*, and *RFC2-5*) were significantly down-regulated in AN-related samples (**Fig. 5A, Appendix Table 5**). Moreover, after 48 hours of treatment with arecoline (0.4 or 0.8 mM, Sigma), the expression level of MMR markers (MLH1 and MSH2) was significantly decreased in the CAL27 cell line (**Fig. 5B-C**).

Altered copy number of driver genes in areca nut-related OSCC patients

We conducted copy number variation analysis in 113 OSCC patients and detected 35 frequently altered regions (**Appendix Figure 13A, Appendix Table 6**), including 19

reported ones encompassing *EGFR* (7p11.2), *CCND1* and *FADD* (11q13.3), *FGFR1* (8p11.23), and *FAT1* (4q35.2), and six novel amplified peaks (3q27.1, 19p12, 2p11.2, 20q11.23, 22q11.21, and 8q24.22), and ten deleted regions (5q31.3, 7q22.1, 2q34, 3p12.1, 18q22.2, 11q24.2, 21q22.3, 10q21.3, 9p21.2, and 19q13.42). Further analysis suggested that the copy number alteration of cancer driver genes was associated with the AN chewing status. For example, *NDRG1* (8q24.22) was amplified at a significantly higher frequency in AN-related OSCC patients (P=0.03), while *PIK3CA* and *TBL1XR1* (3q26.32) were more frequently amplified in AN-negative patients (P=0.03) (**Appendix Figure 13B**). We also observed significant correlations between the amplification level of cancer-related genes and dMMR signature, including *NDRG1*, *MYC*, *NOTCH2* and *TERT* (**Appendix Figure 13C**).

Discussion

In this study, we comprehensively described the genomic and transcriptional profiles of OSCC, especially AN-related OSCC, among Chinese populations. We identified four novel driver genes for OSCC and provided evidence that dMMR mutational signature was predominant in AN-related OSCC, which could also serve as an independent prognostic factor for OSCC. We also unraveled different mutational spectrums of previously known driver genes (*TP53*, *FAT1*, *CASP8*, and *CDKN2A*) as well as novel ones (*ATG2A* and *WEE1*) in AN-related patients. Meanwhile, we found that *PD-L1* expression was significantly associated with AN exposure and dMMR

levels. Our findings not only enable a better understanding of the development of OSCC, but also provide clinical implications for the treatment of AN-related OSCC.

Previous studies identified that OSCC was dominated by sporadic mutations in tumor suppressor genes, and the most common ones were *TP53*, *NOTCH1*, *FAT1*, and *CDKN2A (Chen et al. 2017)*. In this study, we identified some new features for OSCC, such as the aggregated inactivating mutations of tumor suppressor genes within functional binding domains in AN-related patients, indicating a specific molecular process underlying AN exposure. One notable gene is *CDKN2A*, although only mutated in 7% of AN-related patients compared to 25% in AN-negative patients, the highly recurrent truncating mutation p.R58* confers the loss-of-function of CDKN2A/p16 (Parry and Peters 1996). CDKN2A/p16 inhibits checkpoint kinase inhibitors (CHKi) hypersensitivity by attenuating the catalytic activity of the cyclin-dependent kinase 4-6/cyclin D (Cyclin-CDK) complex and CDK2 (Gadhikar et al. 2018). Thus, the high mutation rate of p.R58* in AN-related samples provides clues for CHKi therapy for AN-related OSCC.

WEE1 kinase plays a crucial role in tumorigenesis (Music et al. 2016) by regulating the G2-M cell cycle checkpoint (Do et al. 2013). Consistent with previous findings that inactivation of WEE1 triggers DNA damage (Vriend et al. 2013), we also observed an elevated dMMR signature in *WEE1* mutation carriers (**Appendix Figure 14**), suggesting that *WEE1* alteration may promote the development of AN-related OSCC by hampering MMR function. In addition, it was recently reported that the combination of CHK1 and WEE1 inhibition leads to a synergistic antitumor effect in

tongue cancer and other HNSCC (Nojima et al. 2020; Sun et al. 2018; van Harten et al. 2019), which implies that *WEE1* may also play crucial roles in AN-negative OSCC.

As an important regulator of autophagosome completion (Velikkakath et al. 2012), ATG2A switches the transformation between autophagy and caspase-8 activation (Tang et al. 2017). Although caspase-8 (*CASP8*) was generally considered a driver gene for OSCC (Pickering et al. 2013), increasing evidence suggested that ANE-treated cells exhibit morphological changes of autophagy instead of apoptosis (Yen et al. 2014). Therefore, the mutually exclusive pattern between *CASP8* and *ATG* homologous mutations indicated that autophagy plays the major role in AN-related oral carcinogenesis. Additionally, previous studies identified that polymorphisms of autophagy genes were associated with the risk of OSCC (Fernandez-Mateos et al. 2017), which also proved the role of autophagy genes in oral carcinogenesis. Taken together, it is reasonable to propose that the up-regulated expression of autophagy genes exerted by AN exposure could switch cancer cells from caspase-8-dependent apoptosis to an autophagy addiction mode and thus contribute to the development of OSCC.

Areca nut and arecoline were generally considered cytotoxic and genotoxic to buccal epithelial cells by promoting oxidative DNA damage (Illeperuma et al. 2015) and inhibiting DNA repair (Tsai et al. 2008); however, the mechanisms remain largely unknown. In this study, we provided the first evidence that dMMR was prominent in AN-related OSCC, and could also predict the clinical outcomes, suggesting that the MMR pathway was involved in the occurrence and progression of AN-related OSCC. These findings are consistent with a recent study conducted in Taiwanese (Su et al.

2017). Recently, increasing evidence suggested that MMR is involved in the response to oxidative DNA damage accumulated during carcinogenesis (Bridge et al. 2014), particularly for DNA 8-oxouanine (8-oxoG) (Macpherson et al. 2005), which is the major product of oxidation damage and is specifically induced in MSH2- or MLH1-deficient cells (Colussi et al. 2002). Thus, the down-regulated MSH2 and MLH1 in arecoline-treated cell line suggested that AN exposure could lead to an aggregation of oxidative DNA damage (8-oxoG) by repressing the MMR pathway, and finally contribute to the development of OSCC. However, since previous studies identified that promoter hypermethylation could also alter the expression of *MLH1*, *MSH2*, and *PMS1* in OSCC (Czerninski et al. 2009; Shin et al. 2002; Wang et al. 2013), we could not exclude the possibility that the down-regulation of MMR genes in AN-related patients was caused by methylation alteration. Additionally, there may also exist other mechanisms underlying the development of AN-related OSCC, such as the cancer stemness conversion facilitated by chronic and long-term AN exposure, and the up-regulated stemness mediators exerted by chronic AN exposure (Li et al. 2016; Wang et al. 2016) (Chen et al. 2019; Yang et al. 2018). Thus, the exact molecular mechanism of AN during oral carcinogenesis remains to be fully studied.

Recently, accumulating evidence confirming the association between dMMR and responses to anti-PD-1/L1 therapy across several solid tumors (Calvete et al. 2017). Regarding OSCC, some clinical trials have demonstrated that although anti-PD-1/L1 immunotherapy showed overwhelming efficacy, the response rate was below 20% in recurrent or metastatic HNSCC (Ferris et al. 2016). Positive expression of PD-L1 was

the only viable marker for effective checkpoint immunotherapy in HNSCC. The up-regulated expression of *PD-L1* in AN-related patients and patients with high dMMR signature suggested that AN-related OSCC might represent a potential cohort for effective PD-1/L1 immunotherapy. Moreover, dMMR signature may also serve as a predictive biomarker for anti-PD-1/L1 treatment. These findings would expand the range of HNSCC patients who may benefit from PD-1 pathway blockade.

In conclusion, we identified specific genomic features of AN-related OSCC and emphasized the importance of dMMR in the oral carcinogenesis. We also proposed that dMMR signature might be used as a predictive marker for the anti-PD-1/L1 treatment of AN-related OSCC. These findings provide extensive evidence for understanding the molecular mechanisms underlying AN-induced OSCC and new clues to prevent and treat OSCC patients.

References

Agrawal N, Frederick MJ, Pickering CR, Bettegowda C, Chang K, Li RJ, Fakhry C, Xie TX, Zhang J, Wang J et al. 2011. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in notch1. Science. 333(6046):1154-1157.

Bridge G, Rashid S, Martin SA. 2014. DNA mismatch repair and oxidative DNA damage: Implications for cancer biology and treatment. Cancers. 6(3):1597-1614.

Calvete O, Garcia-Pavia P, Dominguez F, Bougeard G, Kunze K, Braeuninger A, Teule A, Lasa A, Ramon YCT, Llort G et al. 2017. The wide spectrum of pot1 gene variants correlates with multiple cancer types. European journal of human genetics : EJHG. 25(11):1278-1281.

Chen TW, Lee CC, Liu H, Wu CS, Pickering CR, Huang PJ, Wang J, Chang IY, Yeh YM, Chen CD et al. 2017. Apobec3a is an oral cancer prognostic biomarker in taiwanese carriers of an apobec deletion polymorphism. Nature communications. 8(1):465.

Chen YL, Liu KJ, Jang CW, Hsu CC, Yen YC, Liu YL, Chuang TH, Wang SH, Fu YK, Kuo CC et al. 2019. Erk activation modulates cancer stemness and motility of a novel mouse oral squamous cell carcinoma cell line. Cancers. 12(1).

Colussi C, Parlanti E, Degan P, Aquilina G, Barnes D, Macpherson P, Karran P, Crescenzi M, Dogliotti E, Bignami M. 2002. The mammalian mismatch repair pathway removes DNA 8-oxodgmp incorporated from the oxidized dntp pool.

Current biology : CB. 12(11):912-918.

Czerninski R, Krichevsky S, Ashhab Y, Gazit D, Patel V, Ben-Yehuda D. 2009. Promoter hypermethylation of mismatch repair genes, hmlh1 and hmsh2 in oral squamous cell carcinoma. Oral diseases. 15(3):206-213.

Do K, Doroshow JH, Kummar S. 2013. Wee1 kinase as a target for cancer therapy. Cell cycle. 12(19):3159-3164.

Fernandez-Mateos J, Seijas-Tamayo R, Klain JCA, Borgonon MP, Perez-Ruiz E, Mesia R, Del Barco E, Coloma CS, Dominguez AR, Daroqui JC et al. 2017. Analysis of autophagy gene polymorphisms in spanish patients with head and neck squamous cell carcinoma. Scientific reports. 7(1):6887.

Ferris RL, Blumenschein G, Jr., Fayette J, Guigay J, Colevas AD, Licitra L, Harrington K, Kasper S, Vokes EE, Even C et al. 2016. Nivolumab for recurrent squamous-cell carcinoma of the head and neck. The New England journal of medicine. 375(19):1856-1867.

Gadhikar MA, Zhang J, Shen L, Rao X, Wang J, Zhao M, Kalu NN, Johnson FM, Byers LA, Heymach J et al. 2018. Cdkn2a/p16 deletion in head and neck cancer cells is associated with cdk2 activation, replication stress, and vulnerability to chk1 inhibition. Cancer research. 78(3):781-797.

Gomez-Sanchez R, Rose J, Guimaraes R, Mari M, Papinski D, Rieter E, Geerts WJ, Hardenberg R, Kraft C, Ungermann C et al. 2018. Atg9 establishes atg2-dependent contact sites between the endoplasmic reticulum and phagophores. The Journal of cell biology. 217(8):2743-2763.

Hu YJ, Chen J, Zhong WS, Ling TY, Jian XC, Lu RH, Tang ZG, Tao L. 2017. Trend analysis of betel nut-associated oral cancer and health burden in china. The Chinese journal of dental research : the official journal of the Scientific Section of the Chinese Stomatological Association. 20(2):69-78.

Illeperuma RP, Kim DK, Park YJ, Son HK, Kim JY, Kim J, Lee DY, Kim KY, Jung DW, Tilakaratne WM et al. 2015. Areca nut exposure increases secretion of tumor-promoting cytokines in gingival fibroblasts that trigger DNA damage in oral keratinocytes. International journal of cancer. 137(11):2545-2557.

Jeng JH, Chang MC, Hahn LJ. 2001. Role of areca nut in betel quid-associated chemical carcinogenesis: Current awareness and future perspectives. Oral oncology. 37(6):477-492.

Khan I, Pant I, Narra S, Radhesh R, Ranganathan K, Rao SG, Kondaiah P. 2015. Epithelial atrophy in oral submucous fibrosis is mediated by copper (ii) and arecoline of areca nut. Journal of cellular and molecular medicine. 19(10):2397-2412.

Lee V, Le DT. 2016. Efficacy of pd-1 blockade in tumors with mmr deficiency. Immunotherapy. 8(1):1-3.

Li YC, Chang JT, Chiu C, Lu YC, Li YL, Chiang CH, You GR, Lee LY, Cheng AJ. 2016. Areca nut contributes to oral malignancy through facilitating the conversion of cancer stem cells. Molecular carcinogenesis. 55(5):1012-1023.

Macpherson P, Barone F, Maga G, Mazzei F, Karran P, Bignami M. 2005. 8-oxoguanine incorporation into DNA repeats in vitro and mismatch recognition

by mutsalpha. Nucleic acids research. 33(16):5094-5105.

McKay JD, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, Caporaso NE, Johansson M, Xiao X, Li Y et al. 2017. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. Nature genetics. 49(7):1126-1132.

Mortazavi H, Baharvand M, Mehdipour M. 2014. Oral potentially malignant disorders: An overview of more than 20 entities. Journal of dental research, dental clinics, dental prospects. 8(1):6-14.

Music D, Dahlrot RH, Hermansen SK, Hjelmborg J, de Stricker K, Hansen S, Kristensen BW. 2016. Expression and prognostic value of the wee1 kinase in gliomas. Journal of neuro-oncology. 127(2):381-389.

Nojima H, Homma H, Onozato Y, Kaida A, Harada H, Miura M. 2020. Differential properties of mitosis-associated events following chk1 and wee1 inhibitor treatments in human tongue carcinoma cells. Experimental cell research. 386(2):111720.

Parry D, Peters G. 1996. Temperature-sensitive mutants of p16cdkn2 associated with familial melanoma. Molecular and cellular biology. 16(7):3844-3852.

Pickering CR, Zhang J, Yoo SY, Bengtsson L, Moorthy S, Neskey DM, Zhao M, Ortega Alves MV, Chang K, Drummond J et al. 2013. Integrative genomic characterization of oral squamous cell carcinoma identifies frequent somatic drivers. Cancer discovery. 3(7):770-781.

Prabhu RV, Prabhu V, Chatra L, Shenai P, Suvarna N, Dandekeri S. 2014. Areca nut

and its role in oral submucous fibrosis. Journal of clinical and experimental dentistry. 6(5):e569-575.

Shield KD, Ferlay J, Jemal A, Sankaranarayanan R, Chaturvedi AK, Bray F, Soerjomataram I. 2017. The global incidence of lip, oral cavity, and pharyngeal cancers by subsite in 2012. CA: a cancer journal for clinicians. 67(1):51-64.

Shin KH, Park KH, Hong HJ, Kim JM, Oh JE, Choung PH, Min BM. 2002. Prevalence of microsatellite instability, inactivation of mismatch repair genes, p53 mutation, and human papillomavirus infection in korean oral cancer patients. International journal of oncology. 21(2):297-302.

Su SC, Lin CW, Liu YF, Fan WL, Chen MK, Yu CP, Yang WE, Su CW, Chuang CY, Li WH et al. 2017. Exome sequencing of oral squamous cell carcinoma reveals molecular subgroups and novel therapeutic opportunities. Theranostics. 7(5):1088-1099.

Sun L, Moore E, Berman R, Clavijo PE, Saleh A, Chen Z, Van Waes C, Davies J, Friedman J, Allen CT. 2018. Wee1 kinase inhibition reverses g2/m cell cycle checkpoint activation to sensitize cancer cells to immunotherapy. Oncoimmunology. 7(10):e1488359.

Tang Z, Takahashi Y, Chen C, Liu Y, He H, Tsotakos N, Serfass JM, Gebru MT, Chen H, Young MM et al. 2017. Atg2a/b deficiency switches cytoprotective autophagy to non-canonical caspase-8 activation and apoptosis. Cell death and differentiation. 24(12):2127-2138.

Tsai YS, Lee KW, Huang JL, Liu YS, Juo SH, Kuo WR, Chang JG, Lin CS, Jong YJ.

2008. Arecoline, a major alkaloid of areca nut, inhibits p53, represses DNA repair, and triggers DNA damage response in human epithelial cells. Toxicology. 249(2-3):230-237.

van Harten AM, Buijze M, van der Mast R, Rooimans MA, Martens-de Kemp SR, Bachas C, Brink A, Stigter-van Walsum M, Wolthuis RMF, Brakenhoff RH. 2019. Targeting the cell cycle in head and neck cancer by chk1 inhibition: A novel concept of bimodal cell death. Oncogenesis. 8(7):38.

Velikkakath AK, Nishimura T, Oita E, Ishihara N, Mizushima N. 2012. Mammalian atg2 proteins are essential for autophagosome formation and important for regulation of size and distribution of lipid droplets. Molecular biology of the cell. 23(5):896-909.

Vriend LE, De Witt Hamer PC, Van Noorden CJ, Wurdinger T. 2013. Wee1 inhibition and genomic instability in cancer. Biochimica et biophysica acta. 1836(2):227-235.

Wang TY, Peng CY, Lee SS, Chou MY, Yu CC, Chang YC. 2016. Acquisition cancer stemness, mesenchymal transdifferentiation, and chemoresistance properties by chronic exposure of oral epithelial cells to arecoline. Oncotarget. 7(51):84072-84081.

Wang Y, Zhou X, Song Y, Ji X, Zhang A, Zhang G, Gao Z. 2013. The mismatch repair gene hpms1 (human postmeiotic segregation1) is down regulated in oral squamous cell carcinoma. Gene. 524(1):28-34.

Wu F, Parvez F, Islam T, Ahmed A, Rakibuz-Zaman M, Hasan R, Argos M, Levy D,

Sarwar G, Ahsan H et al. 2015. Betel quid use and mortality in bangladesh: A cohort study. Bulletin of the World Health Organization. 93(10):684-692.

Yang SH, Lee TY, Ho CA, Yang CY, Huang WY, Lin YC, Nieh S, Lin YS, Chen SF, Lin FH. 2018. Exposure to nicotine-derived nitrosamine ketone and arecoline synergistically facilitates tumor aggressiveness via overexpression of epidermal growth factor receptor and its downstream signaling in head and neck squamous cell carcinoma. PloS one. 13(8):e0201267.

Yen CY, Chiang WF, Liu SY, Cheng PC, Lee SY, Hong WZ, Lin PY, Lin MH, Liu YC. 2014. Long-term stimulation of areca nut components results in increased chemoresistance through elevated autophagic activity. Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology. 43(2):91-96.

Figures Legends

Figure 1. Mutational spectrum of 429 OSCC patients from SCOCC and TCGA studies.

(A) Overall number of mutations (Synonymous, Non-synonymous and Indels) per megabase (MB) in 113 Chinese OSCC samples. Color coding indicates mutation types.

(B) Mutational spectrum of significantly mutated genes in 89 AN-related and 340 AN-negative OSCC samples.

Top: demographic characteristics and environmental exposures are shown.

Bottom: spectrum of mutations for each patient. Left, significantly mutated genes.

Right: mutation percentages in AN-related and AN-negative OSCC patients, respectively.

(C) The mutational spectrum of *CDKN2A* in 89 AN-related OSCC samples compared to 340 AN-negative OSCC samples.

Top: mutational spectrum of *CDKN2A* in 89 AN-related OSCC samples.

Bottom: mutational spectrum of *CDKN2A* in 340 AN-negative OSCC samples.

(D) The mutational spectrum of *TP53* in 89 AN-related OSCC samples compared to 340 AN-negative OSCC samples.

Top: mutational spectrum of *TP53* in 89 AN-related OSCC samples.

Bottom: mutational spectrum of *TP53* in 340 AN-negative OSCC samples.

Figure 2. The results of the RNAi functional assay of *WEE1*/*ATG2A* and low-dose ANE usage in the CAL27 cell line.

(A) Depleted *WEE1* levels after siRNA knockdown in the CAL27 cell line measured by RT-qPCR.

(B-C) Knockdown of *WEE1* attenuated the proliferation (growth curve and colony formation) and migration in the CAL27 cell line.

(D)Depleted *ATG2A* levels after siRNA knockdown in the CAL27 cell line measured by RT-qPCR.

(E-F) Knockdown of *ATG2A* attenuated the proliferation (growth curve and colony formation) and migration in the CAL27 cell line.

Error bars represent s.e.m, n=3. All experiments were performed at least three times and the statistical analyses were performed by two-sided t-test. **P*<0.05, ***P*<0.01, ****P*<0.001.

Figure 3. Mutational signature analysis.

(A) The pattern of mutational signatures in 113 OSCC samples from SCOCC study.

(B-C) The fraction of MMR signature is significantly higher in AN-related OSCC patients than AN-negative ones in 113 OSCC samples from SCOCC study.

(D) The fraction of MMR signature is significantly higher in 89 AN-related OSCC patients from SCOCC than 340 AN-negative ones (316 samples form TCGA and 24 AN-negative samples from SCOCC).

(E) The expression of *PD-L1* was significantly higher in 40 AN-related OSCC samples from SCOCC than 326 AN-negative ones (320 samples form TCGA and 6 AN-negative samples from SCOCC).

(F) The expression of PD-L1 was significantly higher in OSCC samples with a high dMMR signature proportion.

Figure 4. Kaplan-Meier plot for overall survival by dMMR signature for OSCC patients from SCOCC study.

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

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

Figure 5. Expression correlation of MMR genes with areca nut exposure.

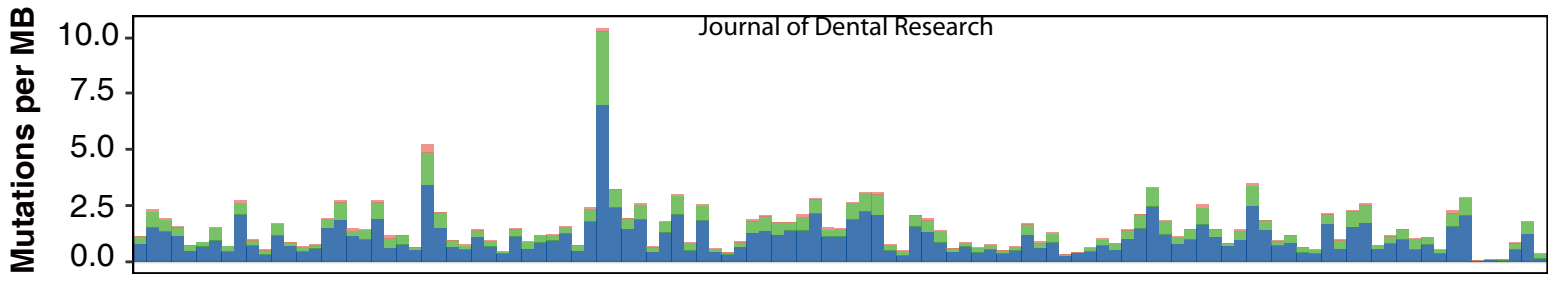
(A) Expression of MMR genes in 40 AN-related OSCC samples from SCOCC and 326 AN-negative ones (320 samples form TCGA and 6 AN-negative samples from SCOCC).

(B-C) The expression levels of MLH1 and MSH2 in the CAL27 cell line treated with arecoline (0.4 or 0.8 mM) for 48 h.

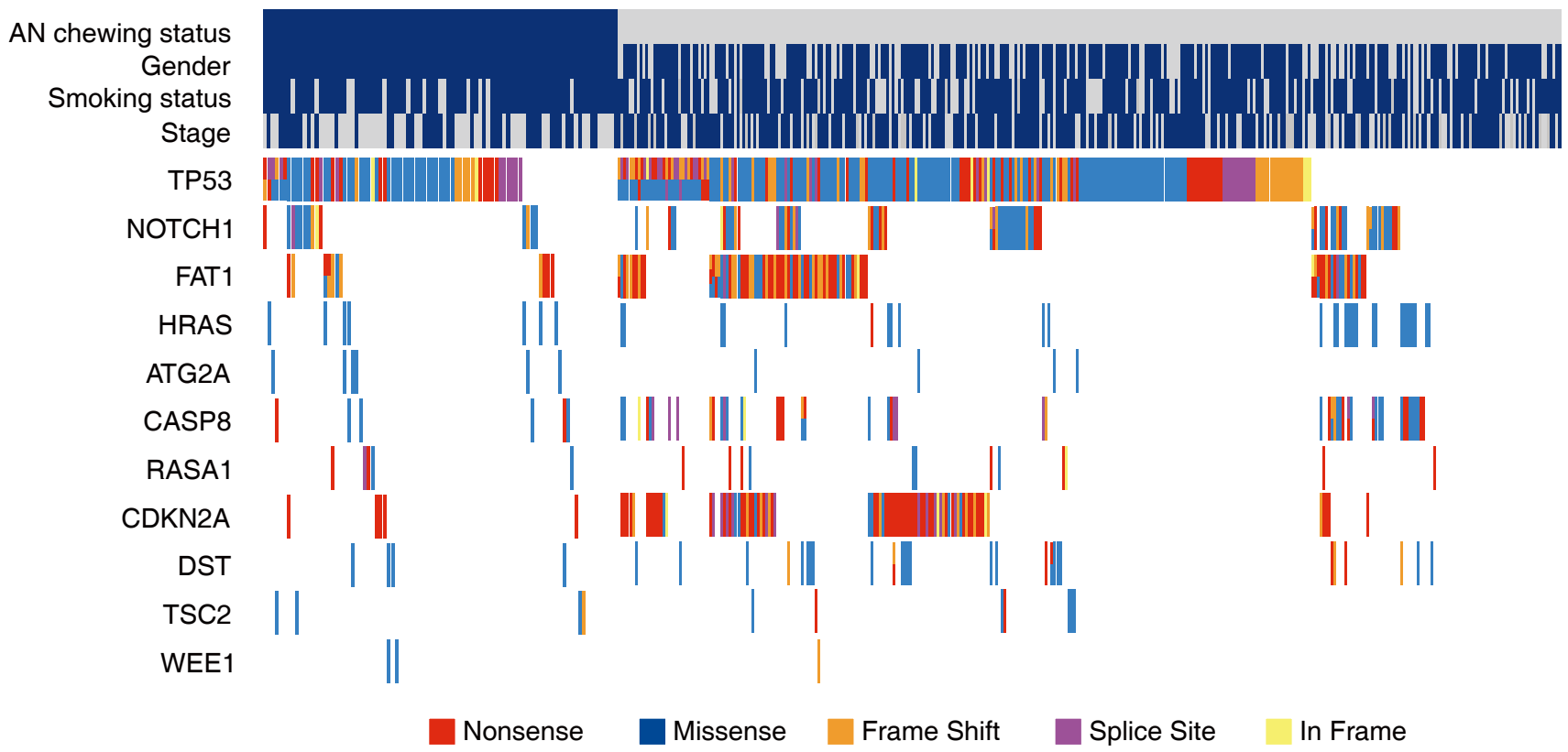
Error bars represent s.e.m, n=3. All experiments were performed at least three times and the statistical analyses were performed by two-sided t-test. **P*<0.05, ***P*<0.01, ****P*<0.001.

Journal of Dental Research

A

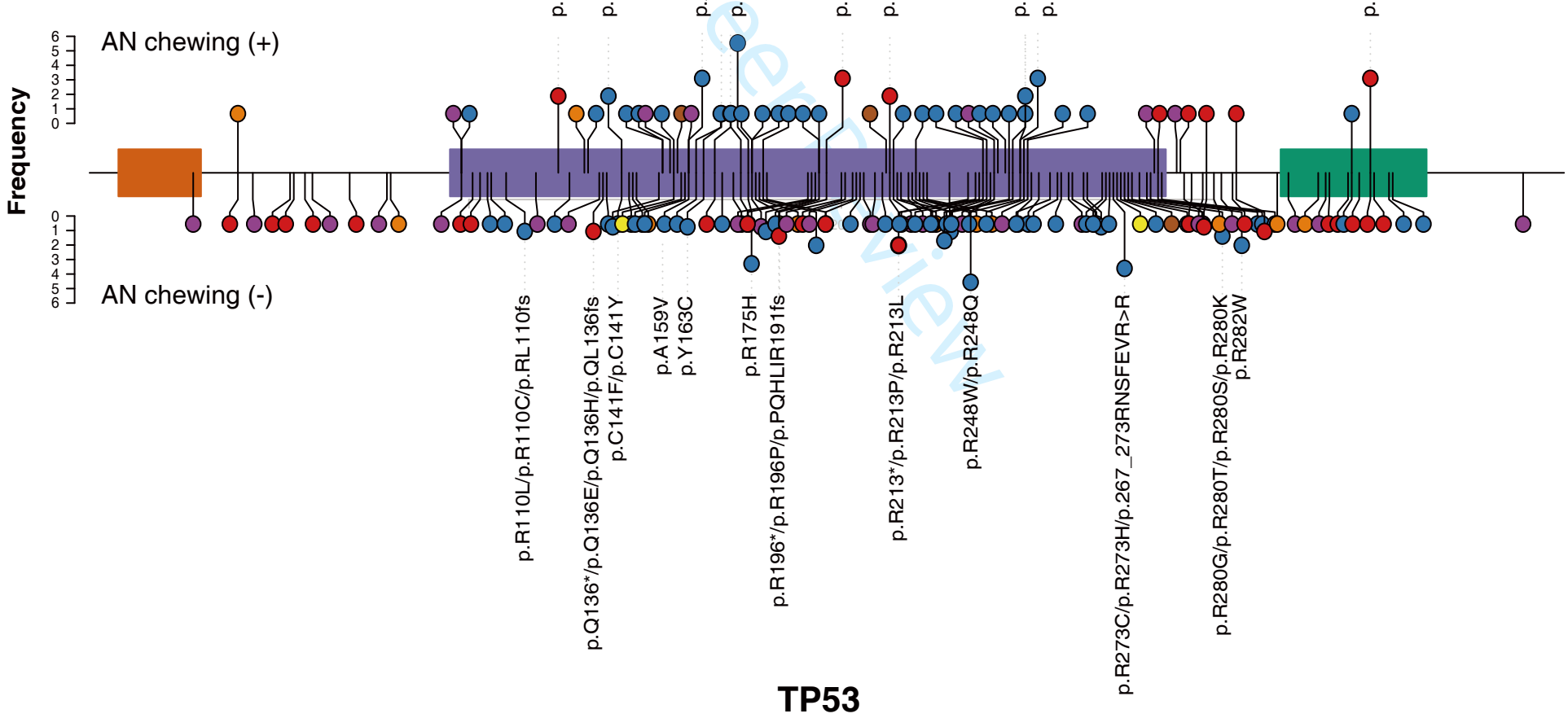


B

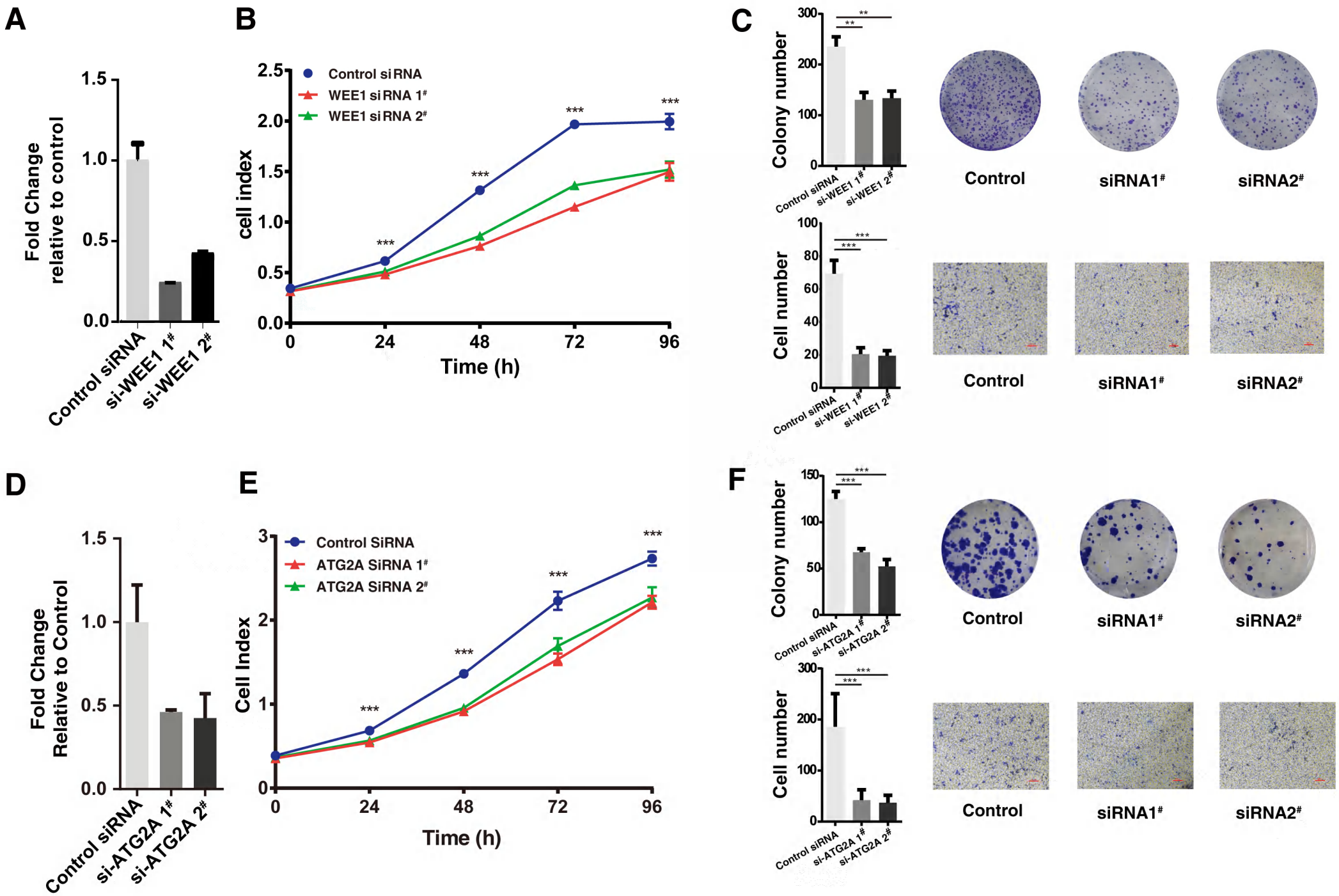


	AN chewing (+)	AN chewing (-)	*P*
TP53	73%	73%	1.00
NOTCH1	16%	21%	0.37
FAT1	12%	25%	0.01
HRAS	8%	9%	1.00
ATG2A	7%	1%	0.003
CASP8	7%	15%	0.07
RASA1	6%	4%	0.36
CDKN2A	6%	25%	<0.001
DST	4%	8%	0.36
TSC2	4%	2%	0.25
WEE1	2%	0%	0.04

C



D


A
Fold Change relative to control
1.5
1.0
0.5
0.0
Control siRNA
si-WEE1 1#
si-WEE1 2#
B
cell index
2.5
2.0
1.5
1.0
0.5
0.0
Control siRNA
WEE1 siRNA 1#
WEE1 siRNA 2#
***
0
24
48
72
96
Time (h)
C
Colony number
300
200
100
0
**
Control siRNA
si-WEE1 1#
si-WEE1 2#
Control
siRNA1#
siRNA2#
Cell number
100
80
60
40
20
0
***
Control siRNA
si-WEE1 1#
si-WEE1 2#
Control
siRNA1#
siRNA2#
D
Fold Change Relative to Control
1.5
1.0
0.5
0.0
Control siRNA
si-ATG2A 1#
si-ATG2A 2#
E
Cell Index
3
2
1
0
Control SiRNA
ATG2A SiRNA 1#
ATG2A SiRNA 2#
***
0
24
48
72
96
Time (h)
F
Colony number
150
100
50
0
***
Control siRNA
si-ATG2A 1#
si-ATG2A 2#
Control
siRNA1#
siRNA2#
Cell number
300
200
100
0
***
Control siRNA
si-ATG2A 1#
si-ATG2A 2#
Control
siRNA1#
siRNA2#