The immune microenvironment of HPV-negative oral squamous cell carcinoma from never-smokers and never-drinkers patients suggests higher clinical benefit of IDO1 and PD1/PD-L1 blockade

Jean-philippe Foy, Chloé Bertolus, Marie-Cécile Michallet, Sophie Deneuve, R. Incitti, Nathalie Bendriss-Vermare, Marie Alexandra Albaret, Sandra Ortiz-Cuaran, Emilie Thomas, A. Colombe, et al.

The immune microenvironment of HPV-negative oral squamous cell carcinoma from never-smokers and never-drinkers patients suggests higher clinical benefit of IDO1 and PD1/PD-L1 blockade

J-P. Foy[1,2,3], C. Bertolus[3], M-C. Michallet[1], S. Deneuve[4], R. Incitti[5], N. Bendriss-Vermare[1], M-A. Albaret[1,5], S. Ortiz-Cuaran[1,5], E. Thomas[5], A. Colombe[2], C. Py[6], N Gadot[2], J-P. Michot[6], J. Fayette[7], A. Viari[5], B. Van den Eynde[8], P. Goudot[3], M. Devouassoux-Shisheboran[9], A. Puisieux[1], C. Caux[1], P. Zrounba[4], S. Lantuejoul[2,6], P. Saintigny[1,2,7].

Authors affiliations: 1-Oncology, Univ Lyon, Université Claude Bernard Lyon 1, INSERM 1052, CNRS 5286, Centre Léon Bérard, Centre de recherche en cancérologie de Lyon, Lyon, 69008, France; 2-Department of Translational Research and Innovation, Centre Léon Bérard, Lyon, France; 3-Department of Oral and Maxillo-Facial Surgery, University of Paris 6, Pitié-Salpêtrière Hospital, Paris, France; 4-Department of Surgery, Centre Léon Bérard, Lyon, France; 5-Synergie Lyon Cancer-Platform of Bioinformatics-Gilles Thomas, Centre Léon Bérard, Lyon, France; 6-Department of Biopathology, Centre Léon Bérard, Lyon, France; 7-Department of Medicine, Centre Léon Bérard, Lyon, France; 8-Ludwig Institute for Cancer Research, Brussels Branch and de Duve Institute, Université catholique de Louvain, B-1200, Brussels, Belgium; 9-Department of Pathology, Croix-Rousse Hospital, Hospices Civils de Lyon, Claude Bernard University Lyon 1, Lyon, France.

Corresponding author: Dr Pierre Saintigny, Université Claude Bernard Lyon 1, INSERM 1052, CNRS 5286, Department of Medicine, Centre Léon Bérard, Cancer Research Center of Lyon; 28 rue Laennec, 69373 Lyon cedex 08, France; Tel : 33-(0)469856097; Fax : 33-(0)478782868; email: pierre.saintigny@lyon.unicancer.fr

Abstract (293 words)

Background: Never-smokers and never-drinkers patients (NSND) suffering from oral squamous cell carcinoma (OSCC) are epidemiologically different from smokers drinkers (SD). We therefore hypothesized that they harbored distinct targetable molecular alterations. **Patients and methods:** Data from The Cancer Genome Atlas (TCGA) (discovery set), Gene Expression Omnibus and Centre Léon Bérard (CLB) (3 validation sets) with available gene expression profiles of HPV-negative OSCC from NSND and SD were mined. Protein expression profiles and genomic alterations were also analyzed from TCGA, and a functional pathway enrichment analysis was performed. Formalin-fixed paraffin-embedded (FFPE) samples from 44 OSCC including 20 NSND and 24 SD treated at CLB were retrospectively collected to perform targeted-sequencing of 2,559 transcripts (HTG EdgeSeq system), and CD3, CD4, CD8, IDO1 and PD-L1 expression analyses by immunohistochemistry (IHC). Enrichment of a six-gene interferon-γ signature of clinical response to pembrozulimab (PD-1 inhibitor) was evaluated in each sample from all cohorts, using the single sample gene set enrichment analysis method. **Results**: A total of 854 genes and 29 proteins were found to be differentially expressed between NSND and SD in TCGA. Functional pathway analysis highlighted an overall enrichment for immune-related pathways in OSCC from NSND, especially involving T-cell activation. Interferon-γ response and PD1 signaling were strongly enriched in NSND. *IDO1* and *PD-L1* were overexpressed and the score of response to pembrolizumab was higher in NSND compared to SD, although the mutational load was lower in NSND. IHC analyses in the CLB cohort evidenced IDO1 and PD-L1 overexpression in tumor cells that was associated with a higher rate of tumor-infiltrating T-cells in NSND compared to SD. **Conclusion:** The main biological and actionable difference between OSCC from NSND and SD lies in the immune microenvironment, suggesting a higher clinical benefit of PD-L1 and IDO1 inhibition in OSCC from NSND.

Keywords: oral cancer; squamous cell carcinoma; never smokers never drinkers; immune microenvironment**,** programmed death-ligand 1, indoleamine 2,3-Dioxygenase 1

Key message

To our knowledge, this is the first report showing that the main biological and actionable difference between OSCC from NSND and SD lies in the immune microenvironment and suggesting a higher benefit of PD-1/PD-L1 and IDO1 inhibition in OSCC from NSND. Our report provides a strong rational to refine personalized strategies of immune-modulation based on smoking and alcohol habits.

Introduction

Head and Neck squamous cell carcinomas (HNSCC) are the seventh most common cancer worldwide and second to lung cancer as smoking-related cancer [1]. However, 10-15% HNSCC are diagnosed in never-smokers and never-drinkers patients (NSND). Three subgroups of NSND with HNSCC have been described: young to middle-aged men with oropharyngeal SCC, young women with oral tongue SCC (OTSCC) and elderly women with gingival SCC [2]. Although human papillomavirus (HPV) has been associated with the increasing incidence of oropharyngeal cancer in young men over the past decades [3], it is not involved in oral carcinogenesis [4] and therefore, does not account for the increasing incidence of OTSCC in young women [5]. Moreover, recent studies have not yet identified any other potentially oncogenic virus in oral cavity [6].

Despite the fact that OSCC arising in NSND and smokers drinkers (SD) are epidemiologically different, recent studies emphasized that they share most common genomic alterations, such as *TP53* and *FAT1* mutations, as well as similar miRNA expression profiles [6-8]. However, while a fraction of HNSCC (7%) may be attributable only to alcohol drinking [9], information in alcohol habits is missing in those studies. Moreover, some studies excluded elderly people [7], which represent another well-established subgroup of NSND with OSCC [2, 10].

Herein, after careful selection of patients based on smoking and alcohol habits independently of patients' age, we show in four independent cohorts that the main biological and actionable difference between OSCC from NSND and SD lies in the immune microenvironment. Notably, OSCC from NSND were characterized by the following: i-an

enrichment of interferon-γ and PD1 signaling pathways; ii-a higher intratumor CD8+ T-cell infiltrate; iii-an overexpression of IDO1 and PD-L1 by tumor cells and iv-higher score of response to pembrolizumab. These features suggest a higher clinical benefit of PD-1/PD-L1 and/or IDO1 inhibition in this subgroup of patients.

Material and Methods

Patients

Four independent cohorts including a total of 212 patients suffering from OSCC were established from a rigorous selection of patients based on alcohol and tobacco habits. The discovery cohort was identified from The Cancer Genome Atlas (TCGA). Three independent cohorts were used as validation cohorts: two cohorts, named GEO1 and GEO2, were established from GSE39366 and GSE65858 respectively, and downloaded from the "Gene Expression Omnibus" (GEO) repository; a third validation cohort included patients treated at Centre Léon Bérard cancer center (CLB, Lyon, France). Written informed consent was obtained from all patients from CLB and the study was approved by the institutional review board.

In each cohort, HPV-positive samples were excluded. The criteria used for selecting NSND and SD patients are detailed in Supplementary Table S1 and led to the inclusion of 117 (ratio NSND/SD: 57/60), 25 (9/16), 26 (7/19) and 45 (20/25) patients from the TCGA (discovery set), GEO1, GEO2 and CLB (validation sets) cohorts respectively.

Molecular profiling in oral samples

In the discovery cohort from TCGA, normalized gene-read counts generated from RNA-sequencing, somatic mutation and copy number alterations were downloaded using the TCGA2STAT R-package and cBioPortal, and were available in 114 (54 NSND; 60 SD), 47

(21 NSND; 26 SD) and 114 (56 NSND; 58 SD) patients respectively. Normalized expression data of 237 proteins and phosphoproteins generated by reverse-phase protein array (level 4) were retrieved from The Cancer Proteome Atlas (TCPA) for 82 patients (45 NSND and 37 SD).

In the three validation cohorts (GEO1, GEO2 and CLB), gene expression profiles were generated using microarray experiments (GEO1 and GEO2) as well as targeted-RNA sequencing in FFPE samples using the HTG EdgeSeq technology (CLB). Finally, array-based gene expression profiles of normal oral mucosa from 39 smokers and 40 never-smokers (GSE17913) and of 18 cell lines derived from human normal oral keratinocytes exposed or not to ethanol and nicotine (GSE57634).

Details on data processing are provided in the Supplementary Table S2.

Functional pathway analysis

Gene expression profiles of OSCC from TCGA were compared between NSND and SD using the EBSeq R-package. The resulting list of transcripts was used to perform a gene set enrichment analysis (GSEA). The single sample gene set enrichment analysis (ssGSEA) was used to compute individual enrichment scores. The search tool for the retrieval of interacting genes/proteins (STRING) was used to build a functional network using differentially expressed proteins between SD and NSND after downloading data from TCPA. Additional and detailed methodology is provided in the Supplementary Methods.

Immunohistochemistry (IHC)

Tissue sections (3µm thick) from 45 FFPE blocks of OSCC resected at CLB were used for IHC. IHC was performed using an automated immunostainer (Ventana Discovery XT, Roche, Meylan, France) with antibodies against PD-L1 (using clones SP142 and 28.8), CD3,

CD8, CD4, p16, and IDO1. Antibodies and procedure is detailed in Supplementary Methods. PD-L1 and IDO1 expression by tumor and immune cells were evaluated blindly to clinical information using 1, 5 and 10% thresholds, positive internal control being represented by T-cells and dendritic cells, respectively. The CD3, CD4 or CD8+ T-cell infiltrate was scored as absent or low (*i.e.* low) versus moderate to marked (*i.e.* high), and the presence of T-cells within the tumor lobules or located at the interface between lobules and stroma was recorded. An automated image analysis of IDO1, PDL1 (clone 28.8) and CD8 immunostaining was also performed independently (HistoWiz Inc., NY 11226) and was compared to the evaluation by the pathologist. Detection of HPV6 and HPV16 was done in p16-positive tumors, using DNA *in situ* hybridization (ISH) (Supplementary Methods).

Statistical analysis

Data statistical analysis was performed using GraphPad Prism version 6.00 (San Diego, SA). Unpaired Wilcoxon's test and Fisher's exact test were performed to compare continuous and categorical values respectively between SD and NSND. Disease-free survival (DFS) distributions were estimated using the Kaplan-Meier method and compared with the log-rank test between NSND and SD. DFS is defined for each cohort in Supplementary Methods. All statistical tests were two-sided, and *P*-values of 0.05 or less were considered to be statistically significant. For multiple testing, a false-discovery rate was computed in order to adjust P-value.

Results

OSCC in NSND are more common in elderly females and harbor less genomic alterations

Clinical and pathological characteristics of OSCC were compared between NSND and SD in four independent cohorts (TCGA, GEO1, GEO2 and CLB) (Supplementary Table S3). Markedly, female and elderly patients (>65years) were more common in NSND from TCGA (P<0.0001; P=0.0021), GEO1 (P=0.0003; P=0.0090), GEO2 (P=0.0138; P=0.0008) and CLB (P<0.0001; P=0.0006). No statistical difference in disease-free survival (DFS) was observed between NSND and SD (Supplementary Figure S1).

Overall, OSCC from NSND harbored significantly less mutations (P=0.0006; Figure 1A) and copy number alterations (P=0.0005; Figure 1B) compared to SD. To gain more insight into differences in genomic profiles between NSND and SD from TCGA, somatic mutation of 12,783 genes as well as copy number values of 22,618 genes were compared between NSND and SD. While the mutational profiles were not significantly different (Figure 1C), 46 genes presented different copy number values using a false discovery rate (FDR) < 0.01 (Wilcoxon test). Notably, amplification of 11q13, including CCND1, was less frequent in NSND compared to SD (P<0.0001) (Figures 1D-E).

The main biological difference between OSCC from NSND and SD lies in the immune microenvironment

In the TCGA dataset, we identified a set of 219 over-expressed and 635 under-expressed genes in NSND compared to SD (called from now on "NSND gene set") (FDR<0.05;$|\log_2 FC|$>1). To validate this NSND gene set, we computed its enrichment score in GEO1 (n=25) and GEO2 (n=26) cohorts, as well as in 79 oral normal mucosa and 18 cell lines derived from normal human oral keratinocytes. The score was significantly higher in OSCC from NSND compared to SD in both GEO1 (P=0.0053) and GEO2 (P=0.0089) (Figure 2A-B). Strikingly, no difference was observed in normal oral mucosa between smokers and non-smokers (P=0.3651) (Figure 2C), as well as in normal oral keratinocytes treated with

ethanol and/or nicotine versus untreated ones (Kruskall-Wallis; P=0.2667; Supplementary Figure 2). These results indicate that the difference we report in OSCC affecting NSND versus SD patients is not related to the effect of smoking and/or alcohol on normal mucosa or keratinocytes.

In order to gain more insight into the biological differences between NSND and SD, we performed a GSEA in the TCGA cohort. With a FDR<0.05, 28 canonical pathways as well as 49 biological processes from GO, were significantly enriched in NSND compared to SD. Many of these pathways/processes were immune-related, especially involving T-cell activation and differentiation. Markedly, OSCC from NSND were also strongly enriched for pathways related to interferon-γ and PD1-signaling (Figure 2D).

The expression of 237 proteins and phosphoproteins was compared between NSND and SD from TCPA, which resulted in the identification of 29 proteins differentially expressed (Wilcoxon, FDR<0.05). Using the STRING tool, 27 of them were interconnected in a functional network (Figure 2D) characterized by an enrichment for 519 biological processes from GO (FDR<0.05), especially for immune response-regulating processes such as "regulation of immune system process" (FDR<0.0001) and "regulation of lymphocyte activation" (FDR=0.0003).

Since our results pointed to the importance of the immune context in OSCC from NSND, we computed the enrichment score of 526 immune modules in the TCGA cohort. Each module has been previously described as a set of transcripts whose expression was most associated with a selected marker gene used to name the module [11]. The score of 150 modules was significantly different between NSND and SD (P<0.05) with 146/150 being higher in NSND. Using the previously published annotation for each module, the 25 most significant modules (Figure 2E) were related to interferon response and T-cells, such as the CD274 (*alias* PD-L1) module with a higher score in NSND (P=0.0048).

Overexpression of IDO1 and PD-L1, increased tumor-infiltrating CD8+ T-cells and higher score of response to pembrolizumab are features of OSCC from NSND

In the discovery cohort (TCGA), *IDO1* (FDR<0.0001) and *PD-L1* (FDR=0.0038) were overexpressed in NSND compared to SD. Of note, PD-L1 was also overexpressed in NSND at protein level (*P*=0.0152). Consistently, *IDO1* was significantly overexpressed in NSND compared to SD in GEO1 (*P*=0.0080) and GEO2 (*P*=0.0476) whereas *PD-L1* was statistically overexpressed in NSND in GEO1 (*P*=0.0116) but not in GEO2 (*P*=0.2086) (Figures 3A-B).

These immune-related differences between NSND and SD were further validated in the CLB cohort using targeted RNA-sequencing of 2,559 genes (HTG) as well as immunohistochemistry (IHC). Among the 5 tumors found to be p16 positive, one was HPV-positive by *in situ* hybridization and was therefore excluded from the analysis. Moreover, three SD samples did not pass HTG quality control and one NSND sample did not generate sufficient product to sequence, resulting in 44 and 40 samples for IHC and HTG data analysis respectively. *IDO1* and *PD-L1* were found to be overexpressed in NSND (*P*=0.0022 and *P*=0.0017 respectively) (Figures 3A-B). Expression of IDO1, PD-L1, CD3, CD4, and CD8 was then evaluated by IHC (Supplementary Table S4; Figure 4). Using a 1% expression threshold to define PD-L1-positive samples, we observed a significant overexpression of PD-L1 in tumor cells in NSND compared to SD (clone Sp142: *P*=0.0270; clone 28.8: *P*=0.0144). Of note, PD-L1-positive immune cells were mostly adjacent to tumor cells expressing PD-L1. Interestingly, two patterns of strong IDO1 staining (>10%) were observed in OSCC: some tumors had a strong and homogenous staining (NSND: 40% versus SD: 0%) whereas in other tumors, a strong staining was only found in delimited areas of the tumor (NSND: 60% versus SD: 17%) with a co-localization of IDO1-positive dendritic cells. Using these two patterns to

identify IDO1-positive samples, IDO1 was overexpressed in tumors from NSND compared to SD (P=0.0046). Scoring the T-cell infiltrate as low versus high, high scores were more common in tumors from NSND for CD3 (P=0.0112), CD4 (P=0.0154) and CD8 (P=0.0285), while no significant difference was found in peritumor areas. This evaluation was validated by an automated image analysis (Supplementary Figure S3-4).

The association of co-overexpression of *PD-L1*/*IDO1* with a high T-cell infiltrate suggested a higher clinical benefit of PD1/PD-L1 blockade in NSND [12, 13]. Thus, we computed the enrichment score of a six-gene response signature to pembrolizumab in HNSCC [14], in the four independent cohorts (Figure 3C). Interestingly, scores were higher in NSND compared to SD in TCGA (P=0.0012); GEO1 (P=0.0271); GEO2 (P=0.0181) and CLB cohorts (P=0.0002).

Discussion

Immunotherapy is providing unprecedented advances in management of HNSCC [12, 14, 15] but the challenge is now to understand in which clinical setting to use these agents in order to provide greater clinical benefit for patients [16]. Herein we report that HPV-negative OSCC in NSND is molecularly distinct from the disease affecting SD and is mainly characterized by its immune microenvironment. Notably, a significant enrichment for interferon-γ and PD1 pathways, an overexpression of IDO1 and PD-L1 as well as a higher CD8+ T-cell tumor infiltrate were observed in NSND compared to SD. Altogether, these features provide a strong rational for immunotherapy in NSND with OSCC, especially by targeting PD-L1 and IDO1.

Using a functional pathway analysis based on gene and protein expression to identify actionable pathways in NSND, we observed a significant enrichment in immune-related pathways in NSND compared to SD, especially involving T-cell signaling, which was

consistent with a higher rate of tumor infiltrating T-cells in NSND, including mostly T-cytotoxic CD8 cells. Immunologic changes induced by smoking are various and complex [17]. They may play a role during oral carcinogenesis through the inhibition of Langerhans cells maturation, by impairing T-cell recruitment and activation via the abrogation of their antigen presentation capability to CD4+ T cells [18], which is consistent with our result. In contrast, a higher rate of infiltrating CD8+ T-cells have been associated with smoking in other tumor settings, including non-small cell lung cancers [19]. These conflicting results across tumor types underlines that the effect of smoking on immunity response may depend on many variables including the tissue of origin [17]. Moreover, alcohol, a recognized risk factor for oral cancer but not for lung cancer [20], may also affect with the immune system and explain the differences between tumor types [21].

In line with the increased interferon-γ activation previously observed in HNSCC from NSND [22], we observed that OSCC from NSND were strongly enriched for IFN-γ response. In our study, we focused on *IDO1* and *PD-L1,* both induced by IFN-γ, because they are targetable and were consistently overexpressed in NSND in several cohorts, at the RNA and protein levels. IDO1 is an enzyme involved in tryptophan catabolism, resulting in an immunosuppressive local environment. Overexpression of *IDO1* is a mechanism of immune escape [23] and ID01 inhibition may exert antitumor effects via the activation of anticancer immunosurveillance [24]. Elevated expression of *IDO1* has been also observed in pre-treatment melanomas from responding patients to PD-L1 inhibition [13]. Together with our results, these observations suggest a rational for combination therapies in OSCC from NSND.

The higher rate of tumor infiltrating T-cells especially T-CD8, the overexpression of PD-L1 and IDO1 and the enrichment of IFN-γ transcriptomic signatures observed in OSCC from NSND, are features that have also been reported as predictors of response to PD1/PD-L1 inhibitors [12]. The lower mutational load in NSND versus SD is expected and linked to the

absence of exposure to carcinogenic effects of tobacco and alcohol. This might appear conflicting with the results of our study [12]. However, recent observations in melanomas indicate that a relatively low mutational load does not preclude a tumor response to PD1/PD-L1 inhibition [25]. In a recent study, non-synonymous mutational burden did not correlate with response to PD-1 blockade in patients with HNSCC treated by PD-1 blockade [26]. While the mutational load may contribute to the clinical response to immunotherapy through neo-antigen-specific T cell responses, this is not always the case [27]. HPV-positive HNSCC harbor fewer mutations [28] and present a higher response to anti-PD1 when compared to HPV-negative HNSCC [14, 15, 29]. In the case of viral-induced carcinogenesis, viral antigens likely represent dominant antigens that stimulate T-cells reactivity and contribute to the inflamed tumor phenotype and IFN-γ activation. These events are predictive of response to PD1/PDL1 blockade, without the need for mutation-derived neoantigens [30]. Thus, as observed in HPV-related diseases, we may suggest that the potent IFNγ mediated immune response in HPV-negative OSCC from NSND is linked to a dominant tumor antigen independent from the mutation load.

In addition to genomic-based biomarkers, immune signatures of response to PD1/PD-L1 inhibition have been recently proposed [25, 31]. In particular, a six IFN-γ-related genes signature of response to pembrozulimab has been established in metastatic HNSCC [14]. Using the HTG technology, that allows the generation of targeted gene expression profiling from a single FFPE section, we were able to show that this signature was significantly enriched in OSCC from NSND compared to SD, strengthening the potential benefit of PD1/PD-L1 inhibition in NSND.

In conclusion, the main biological and actionable difference between OSCC from NSND and SD lies in the immune microenvironment. Markedly, OSCC in NSND are characterized by an enrichment of interferon-γ and PD1 pathways, a higher intratumor T-cell

infiltrate, an overexpression of *IDO1* and *PD-L1,* and higher score of response signature to pembrolizumab. More studies are required to evaluate the impact of smoking and alcohol status on response to PD1/PD-L1 pathway inhibition.

Acknowledgments

We acknowledge Jean-Claude Gérard and Salim Aissaoui (HTG Molecular Diagnostics) for their assistance to perform RNA-targeted sequencing, as well as Elise Malandain (Department of Biopathology, Centre Léon Bérard, Lyon, France) for her technical assistance.

Funding

This work was supported by the *Cancéropôle Lyon Auvergne Rhône-Alpes* (CLARA) 2014-2016 Structured Program [Grant number CVPPRCAN000153 / International Head and Neck Prevention Act-IHNPACT to P.S.) and the LYric Grant INCa-DGOS-4664.

J.P.F. was supported by a fellowship grant from the *Fondation pour la Recherche Médicale* (FRM), the *Association Française pour le Développement de la Stomatologie* (AFDS) and from the *Agence Régionale de Santé* (*année-recherche 2015-2016*).

Disclosure

P.S. has received research funding from Roche; C.C. has received research funding from Roche and Bristol Myers Squibb (BMS); S.L. has served as a consultant/advisor for Bristol Myers Squibb (BMS), Merck Sharp and Dohme (MSD), Roche, AstraZeneca and Boerhinger-Ingelheim and has received research funding from Bristol Myers Squibb (BMS); J.F. has received honoraria from AstraZeneca; B.V.d.E. is co-founder and owns stock in iTeos

Therapeutics and has served as a consultant/advisor for Amgen (SAB member); R.I owns stock in OncoDiag. All remaining authors have declared no conflicts of interest.

References

1. Ferlay J, Soerjomataram I, Dikshit R et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015; 136: E359-386.
2. Dahlstrom KR, Little JA, Zafereo ME et al. Squamous cell carcinoma of the head and neck in never smoker-never drinkers: a descriptive epidemiologic study. Head Neck 2008; 30: 75-84.
3. Chaturvedi AK, Engels EA, Pfeiffer RM et al. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J Clin Oncol 2011; 29: 4294-4301.
4. Lingen MW, Xiao W, Schmitt A et al. Low etiologic fraction for high-risk human papillomavirus in oral cavity squamous cell carcinomas. Oral Oncol 2013; 49: 1-8.
5. Patel SC, Carpenter WR, Tyree S et al. Increasing incidence of oral tongue squamous cell carcinoma in young white women, age 18 to 44 years. J Clin Oncol 2011; 29: 1488-1494.
6. Li R, Faden DL, Fakhry C et al. Clinical, genomic, and metagenomic characterization of oral tongue squamous cell carcinoma in patients who do not smoke. Head Neck 2015; 37: 1642-1649.
7. Pickering CR, Zhang J, Neskey DM et al. Squamous cell carcinoma of the oral tongue in young non-smokers is genomically similar to tumors in older smokers. Clin Cancer Res 2014; 20: 3842-3848.
8. Kolokythas A, Zhou Y, Schwartz JL, Adami GR. Similar Squamous Cell Carcinoma Epithelium microRNA Expression in Never Smokers and Ever Smokers. PLoS One 2015; 10: e0141695.
9. Hashibe M, Brennan P, Benhamou S et al. Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. J Natl Cancer Inst 2007; 99: 777-789.
10. Koch WM, Lango M, Sewell D et al. Head and neck cancer in nonsmokers: a distinct clinical and molecular entity. Laryngoscope 1999; 109: 1544-1551.
11. Linsley PS, Chaussabel D, Speake C. The Relationship of Immune Cell Signatures to Patient Survival Varies within and between Tumor Types. PLoS One 2015; 10: e0138726.
12. Machiels JH, Coulie PG. The promise of immunostimulatory antibodies in head and neck cancer. Lancet Oncol 2016.
13. Herbst RS, Soria JC, Kowanetz M et al. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature 2014; 515: 563-567.
14. Seiwert TY, Burtness B, Mehra R et al. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. Lancet Oncol 2016.
15. Ferris RL, Blumenschein G, Jr., Fayette J et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med 2016.
16. Economopoulou P, Agelaki S, Perisanidis C et al. The promise of immunotherapy in head and neck squamous cell carcinoma. Ann Oncol 2016.
17. Lee J, Taneja V, Vassallo R. Cigarette smoking and inflammation: cellular and molecular mechanisms. J Dent Res 2012; 91: 142-149.
18. Schierl M, Patel D, Ding W et al. Tobacco smoke-induced immunologic changes may contribute to oral carcinogenesis. J Investig Med 2014; 62: 316-323.
19. Kinoshita T, Muramatsu R, Fujita T et al. Prognostic value of tumor-infiltrating lymphocytes differs depending on histological type and smoking habit in completely resected non-small-cell lung cancer. Ann Oncol 2016; 27: 2117-2123.
20. Boffetta P, Hashibe M. Alcohol and cancer. Lancet Oncol 2006; 7: 149-156.
21. Meadows GG, Zhang H. Effects of Alcohol on Tumor Growth, Metastasis, Immune Response, and Host Survival. Alcohol Res 2015; 37: 311-322.

22. Farshadpour F, Roepman P, Hordijk GJ et al. A gene expression profile for non-smoking and non-drinking patients with head and neck cancer. Oral Dis 2012; 18: 178-183.
23. Vigneron N, van Baren N, Van den Eynde BJ. Expression profile of the human IDO1 protein, a cancer drug target involved in tumoral immune resistance. Oncoimmunology 2015; 4: e1003012.
24. Vacchelli E, Aranda F, Eggermont A et al. Trial watch: IDO inhibitors in cancer therapy. Oncoimmunology 2014; 3: e957994.
25. Hugo W, Zaretsky JM, Sun L et al. Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. Cell 2016; 165: 35-44.
26. Hanna GJ, Liu H, Jones RE et al. Defining an inflamed tumor immunophenotype in recurrent, metastatic squamous cell carcinoma of the head and neck. Oral Oncol 2017; 67: 61-69.
27. Farkona S, Diamandis EP, Blasutig IM. Cancer immunotherapy: the beginning of the end of cancer? BMC Med 2016; 14: 73.
28. Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 2015; 517: 576-582.
29. Chow LQ, Haddad R, Gupta S et al. Antitumor Activity of Pembrolizumab in Biomarker-Unselected Patients With Recurrent and/or Metastatic Head and Neck Squamous Cell Carcinoma: Results From the Phase Ib KEYNOTE-012 Expansion Cohort. J Clin Oncol 2016.
30. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. Science 2015; 348: 69-74.
31. Chen PL, Roh W, Reuben A et al. Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. Cancer Discov 2016.

Figure legends

Figure 1: Genomic-profiles of OSCC in patients NSND versus SD. The mutation count (A) and the fraction of copy number alterations per sample (B) in OSCC from NSND and SD were compared in the TCGA cohort (Mann Whitney test). (C) The mutation status of genes commonly mutated is shown in each sample. Genes were selected based on an overall mutation frequency >10% in head and neck cancer from TCGA. (D) Copy number alterations established from GISTIC algorithm are shown for genes included in the 11q13 chromosomal region. (E) A Fisher-exact test was performed to compare the number of *CCND1* amplification between NSND and SD. Abbreviations: OSCC: oral squamous cell carcinomas; NSND: never-smokers never-drinkers; SD: smokers drinkers; TCGA: The Cancer Genome Atlas.

Figure 2: Functional pathway analysis in OSCC. (A-C) The enrichment score of the NSND gene set (defined in TCGA by 219 and 635 over- and under-expressed genes respectively in OSCC from NSND), was computed in two independent cohorts of OSCC [(A) GSE39366 and (B) GSE65858], as well as in a set of normal oral mucosa (GSE17913) (C); the score was compared in NSND versus SD using a Mann Whitney test. (D-E) In the TCGA cohort, NSND were characterized by a significant enrichment for the interferon-γ and PD1 signaling pathways using GSEA (D) as well as a functional network including 27 of 29 proteins differentially expressed between NSND and SD using STRING (E). (F) Functional immune landscape of OSCC in NSND and SD: the enrichment score for 526 different immune modules [11] was computed using ssGSEA and were compared between NSND and SD from

TCGA. The 25 most significant modules differentially enriched and their annotation (NK, T-cell, myeloid, interferon) are shown. Abbreviations: OSCC: oral squamous cell carcinomas; NSND: never-smokers never-drinkers; SD: smokers drinkers; S: smokers, NS: non-smokers; GSEA: gene set enrichment analysis; ssGSEA: single sample GSEA; STRING: search tool for the retrieval of interacting genes/proteins; TCGA: The Cancer Genome Atlas; NSND score: enrichment score of the NSND gene set.

Figure 3: OSCC from NSND are characterized by an overexpression of *IDO1*, *PD-L1* and higher score of response to pembrolizumab. *IDO1* (A) and *PD-L1* (B) expression are compared between NSND and SD patients in the TCGA discovery cohort and three validation series including GSE39366 and GSE65858 (array-based expression profiles) and the CLB cohort using the HTG EdgeSeq technology (Mann Whitney test). (C) In all four cohorts, we used the ssGSEA tool in order to compute an enrichment score of a six-gene response signature to pembrolizumab described in HNSCC [14] (score called Pembro S.). Tumors were ranked according to this score. *IDO1* and *PDL1* expression are shown in all cohorts. Additionally, the expression of PDL1 at protein level (as determined by RPPA), the mutation count (MUT. Count) and the percentage of copy number alteration (% CNA) are shown in TCGA as well as IHC data for PDL1, IDO1, CD3, CD4 and CD8, in CLB cohort. We used 1 and 10% thresholds to define PD-L1 and IDO1-positive samples respectively. The CD3, CD4 and CD8+ T cell infiltrate was scored as low (-) versus high (+). Abbreviations: OSCC: oral squamous cell carcinomas; NSND: never-smokers never-drinkers; SD: smokers drinkers; TCGA: The Cancer Genome Atlas; CLB: Centre Léon Bérard; IHC: immunohistochemistry; ssGSEA: single sample Gene Set Enrichment Analysis; RPPA: reverse phase protein array

Figure 4: IDO1, PD-L1 and CD3 expression in the CLB cohort by immunohistochemistry. (A-D) Well differentiated OSCC in a NSND patient; (A) Hematoxylin and Eosin stain (original magnification x 100); (B) PD-L1 expression by tumor cells and surrounding immune cells (SP142 clone; immunoperoxydase, x 100); (C) CD3 immunostaining of stromal T cells, at the interface between lobules and stroma and within the lobules (2GV6 clone; immunoperoxydase, x 100); (D) IDO1 expression by tumor cells and dendritic cells (4.16H1 clone; immunoperoxydase, x 100). (E-H) Well differentiated OSCC in a SD patient; (E) Hematoxylin and Eosin stain (original magnification x 100); (F) PD-L1 expression by less than 1% of tumor cells and immune cells (SP142 clone; immunoperoxydase, x 100); (G) CD3 expression by T cells, mainly located at the periphery of the lobules (2GV6 clone; immunoperoxydase, x 100); (H) IDO1 expression by very rare dendritic cells and no tumor cells (4.16H1 clone; immunoperoxydase, x 100). Abbreviations: OSCC: oral squamous cell carcinomas; NSND: never-smokers never-drinkers; SD: smokers drinkers

Supplementary Figure legends

Supplementary Figure S1: Disease-free survival (DFS) of NSND and SD patients treated for OSCC, in four independent cohorts including the TCGA (discovery cohort), GEO1, GEO2 and the CLB cohort (validation cohorts). In all datasets, disease-free survival (DFS) distributions were estimated using the Kaplan-Meier method and compared with the log-rank test between NSND and SD. Abbreviations: OSCC: oral squamous cell carcinomas; NSND: never-smokers never-drinkers; SD: smokers-drinkers; TCGA: The Cancer Genome Atlas; CLB: Centre Léon Bérard.

Supplementary Figure S2: Score of the NSND gene set in 18 cell lines derived from human normal oral keratinocytes. The enrichment score of the NSND gene set was computed in 18 cell lines derived from human normal oral keratinocytes treated with various doses of ethanol and/or nicotine. Cell lines were divided in four groups: “Eth- Nicot-”: controls not treated; “Eth- Nicot+”: not treated by ethanol but treated by various doses of nicotine; “Eth+ Nicot-”: treated by various doses of ethanol but not by nicotine; “Eth+ Nicot+”: treated by various doses of ethanol and nicotine. Scores were compared between the four groups, using a Kruskall-Wallis test.

Supplementary Figure S3: Automated image analysis and evaluation by the pathologist of IDO1, PDL1 (clone 28.8) and CD8 immunostaining in the CLB cohort. For each

antibody, the percentage of positive staining cells (automated image analysis) was compared between samples classified by the pathologist, using a Mann-Whitney test. In order to identify positive or negative samples, the thresholds used by the pathologist were 1% and 10% (diffuse or local) for PDL1 (28.8) and IDO1 respectively. For CD8, samples were classified as LOW (no or low infiltrate) or HIGH (moderate to marked infiltrate).
NEG: negative; POS: positive

Supplementary Figure S4: Expression of IDO1, PDL1 and CD8 at RNA (targeted RNASeq) and protein (immunohistochemistry) level in the CLB cohort. Using a Mann Whitney test, the expression of *PDL1*, *IDO1* and *CD8* (HTG Edge Seq technology) was compared between positive/high and negative/low samples for immunostaining of antibodies corresponding to these genes, as evaluated by the pathologist. The thresholds used by the pathologist were 1% and 10% (diffuse or local) for PDL1 (28.8) and IDO1 respectively. For CD8, samples were classified as LOW (no or low infiltrate) or HIGH (moderate to marked infiltrate).
NEG: negative; POS: positive

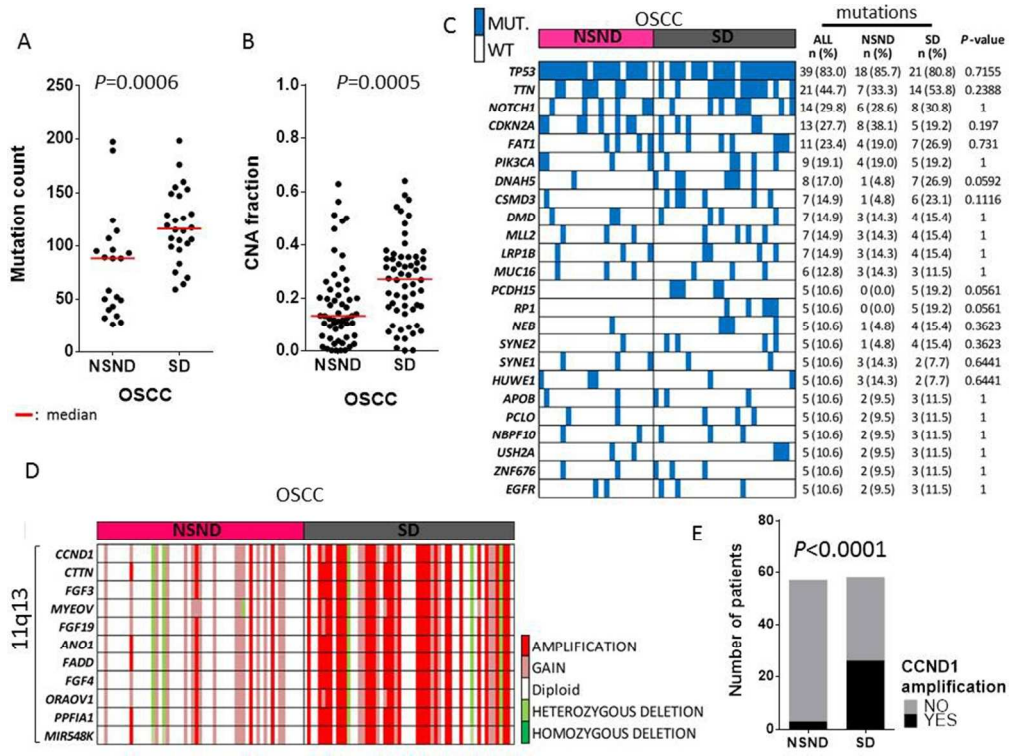

Figure 1

254x190mm (96 x 96 DPI)

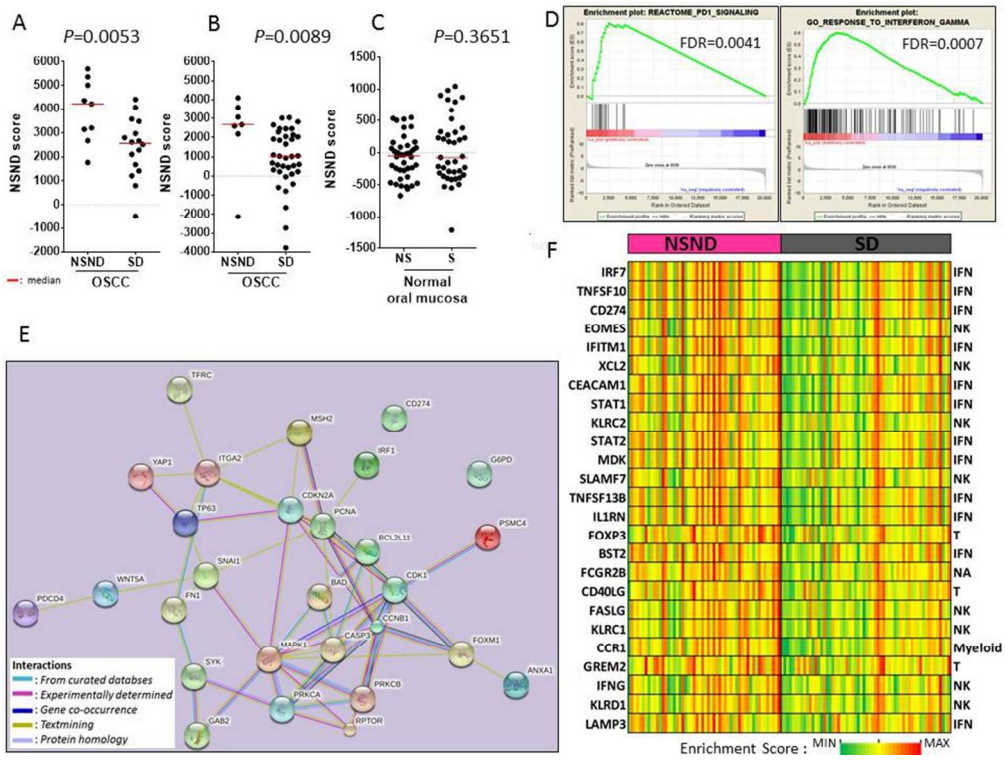

Figure 2

254x190mm (96 x 96 DPI)

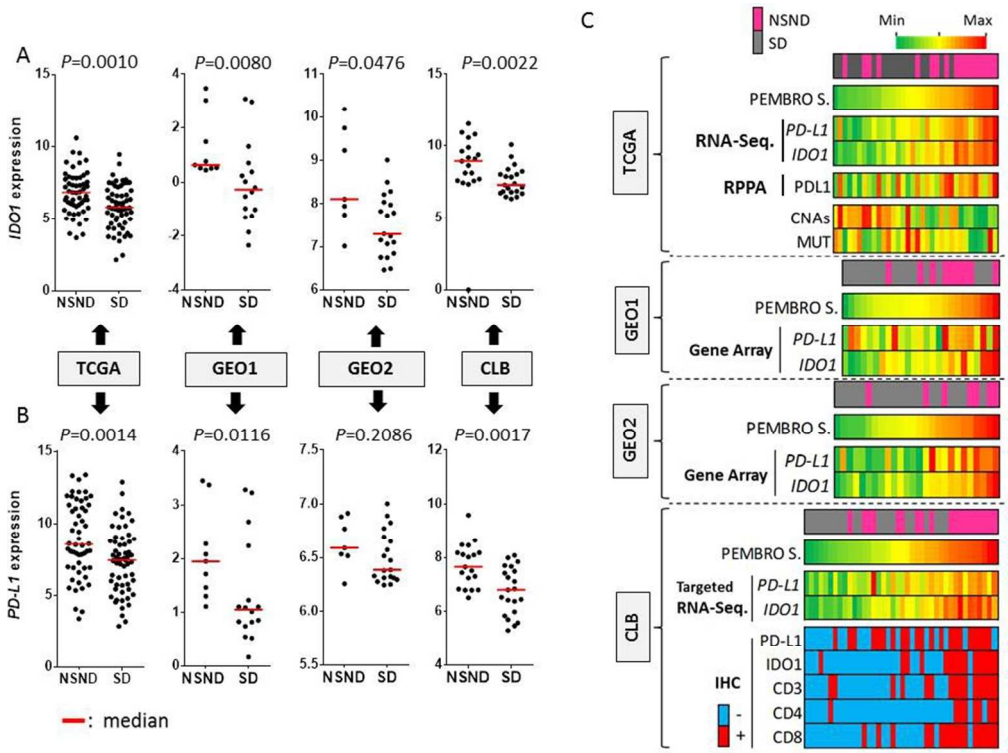

Figure 3

254x190mm (96 x 96 DPI)

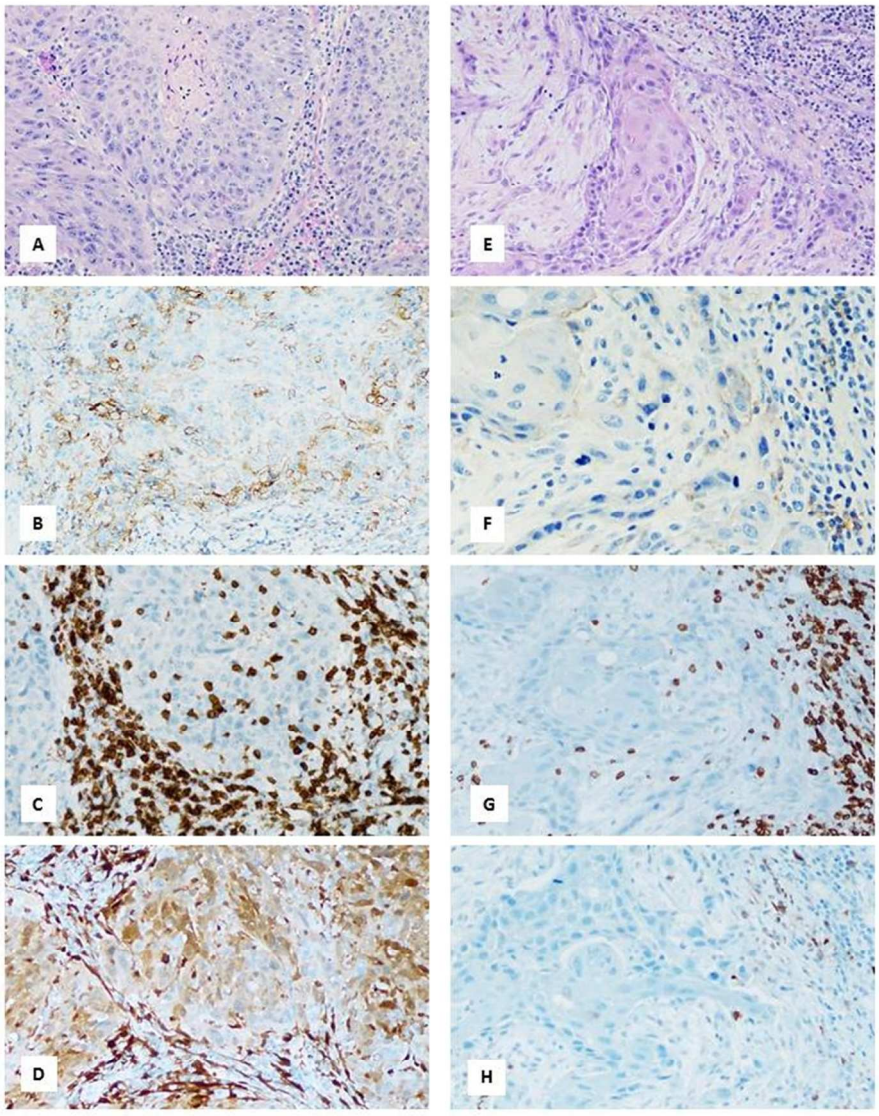

Figure 4

190x254mm (96 x 96 DPI)