

Submitted Manuscript: Confidential

Head and neck squamous cell carcinomas are characterized by a stable immune signature within the primary tumor over time

Authors:

Oliver Wood ^{†1}, James Clarke ^{†1}, Jeongmin Woo², Adal H Mirza^{1,4}, Christopher H Woelk ², Gareth J Thomas ¹, Pandurangan Vijayanand ^{2,3*}, Emma King ^{* 1,4}, Christian H Ottensmeier ^{1*}

Affiliations:

¹Cancer Sciences & NIHR and CRUK Experimental Cancer Sciences Unit, University of Southampton,

²Clinical & Experimental Sciences, University of Southampton

³La Jolla Institute for Allergy and Immunology, 9420 Athena Circle, La Jolla, CA 92037

⁴Department of Otolaryngology, Pool Hospital NHS Foundation Trust, Longfleet Road, Poole.

†Authors Contributed Equally to this work.

* Joint equal contribution

Keywords: RNA-Seq, HNSCC, Heterogeneity, Immunohistochemistry, Transcriptome analysis, Immune infiltrate, T-cell, Adaptive immune response, Immunology.

Financial Support: Funding was provided by Cancer Research UK grant number C491/A15951.

Correspondence should be addressed to:

Christian H Ottensmeier

University of Southampton, Faculty of Medicine

Cancer Sciences Unit

Somers Building MP824

Tremona Road

Southampton, UK SO16 6YD

Tel: 02381 205161

Email: C.H.Ottensmeier@soton.ac.uk

Conflict of interest: The authors declare no conflicts of interest.

Abstract word count: 247

Translational relevance word count: 145

Word count: 3167 (with changes)

Number of figures: 6

Number of references: 30

Supplementary material: 2 files (Supplementary Figures 1- 6; Supplementary Table 1)

Abstract

Purpose: Genetic and morphological heterogeneity is well-documented in solid cancers. Immune cells are also variably distributed within the tumor; this heterogeneity is difficult to assess in small biopsies, and may confound our understanding of the determinants of successful immunotherapy. We examined the transcriptomic variability of the immunological signature in head and neck squamous cell carcinoma (HNSCC) within individual tumors using transcriptomic and immunohistochemical assessments.

Experimental design: Forty-four tumor biopsies from 16 HNSCC patients, taken at diagnosis and later at resection, were analyzed using RNA-sequencing. Variance filtering was used to identify the top 4000 most variable genes. Principal component analysis, hierarchical clustering and correlation analysis were performed. Gene expression of *CD8A* was correlated to immunohistochemical analysis.

Results: Analysis of immunological gene expression was highly consistent in replicates from the same cancer. Across the cohort, samples from the same patient were most similar to each other, both spatially (at diagnosis) and notably, over time (diagnostic biopsy compared to resection); comparison of global gene expression by hierarchical clustering [$p < 0.0001$] and correlation analysis [median intrapatient $r = 0.82$; median interpatient $r = 0.63$]. *CD8A* gene transcript counts were highly correlated with CD8 T-cell counts by immunohistochemistry ($r = 0.82$).

Conclusion: Our data demonstrate that in HNSCC the global tumor and adaptive immune signatures are stable between discrete parts of the same tumor and also at different timepoints. This suggests that immunological heterogeneity may not be a key reason for failure of immunotherapy and underpins the use of transcriptomics for immunological evaluation of novel agents in HNSCC patients.

Translational relevance

Our aim was to evaluate and to quantify the immunological heterogeneity in head and neck squamous cell carcinoma (HNSCC). Understanding variability is important if immunological tumor assessment is used to inform treatment choice, to monitor response and to understand the variability of outcome after immunotherapy observed in clinical practice.

Multi-region tumor sampling at baseline, resection or both, were followed by RNA-Sequencing. We observed a high level of intertumoral heterogeneity but significant stability of intratumoral transcriptomes in individual patients; transcriptomes and immune cell signatures are stable in HNSCC across time and space in tumors that have not been exposed to treatment in the interim. The use of RNA-Seq accurately captures the immune landscape of an individual patient's tumor. Our data support that biopsies offer a representative insight into the immune characteristics of cancer tissue and provide a basis for interpreting changes that are observed following treatment.

INTRODUCTION

Morphological heterogeneity in cancer has long been appreciated by pathologists; more recent data show that tumor cell clones may undergo marked diversification over space and time as a result of mutational divergence on the one hand and selection by, for example, immunological attack on the other hand (1-7). It has further become evident that immune attack appears to vary according to its location in the cancer tissue: when quantifying the density of tumor-infiltrating immune cells by microscopy, the spatial distribution of immune cells with respect to tumor cells needs to be taken into account (8). Given that immunotherapeutics only benefit a minority of patients (9), heterogeneity of immune cell distribution might contribute to treatment failure; and in that case small biopsy samples might not accurately represent the tumor microenvironment and immune status of a patient's tumor. In clinical practice however, it is difficult to sample multiple tumor areas at any time other than at surgical resection.

Valuable prognostic information can be gained from simple enumeration of T-cell infiltrates in both human papillomavirus expressing (HPV(+)) and HPV independent (HPV(-)) head and neck squamous cell carcinoma (HNSCC) (10,11). Transcriptomic analysis of primary HNSCC using RNA-sequencing (RNA-Seq) allows further characterization of global molecular signatures (12) and detailed assessment of immune infiltrates (13). Such evaluation of gene expression provides unprecedented insight into biological processes that occur in tissue. Like morphological assessments, molecular data are shaped by tumor-driving mutations, tissue heterogeneity, disease progression or selective pressures from treatment (14-16). Technical issues may further affect interpretation: in breast cancer, significant differences were identified between diagnostic tissue core biopsies when compared to tumor excisions. Surgical ischemia was identified as a significant confounding factor in transcriptomic analysis (17).

Knowledge of variability is critical for the understanding and effective evaluation of immunotherapy interventions using checkpoint inhibitors (e.g. anti-CTLA4, anti-PD1 and anti-PDL-1 antibodies) (18,19). The mechanisms that underpin immunotherapeutic success as opposed to treatment failure remain poorly understood but must be linked to the particulars of the pre-existing immune infiltrate (5,20,21). To make sense of treatment failure therefore, and of any changes detected in longitudinal assessments, understanding of natural variability is critical,

as is the understanding of how results from different assays (e.g. histological and transcriptomic assessment) correlate with each other.

Using high resolution transcriptomic evaluation with RNA-Seq (22), we assessed variability in HNSCC within an individual cancer between multiple samples (replicates). Additionally, we wondered whether replicates at diagnosis would be similar or different to replicates from a resection specimen of the same cancer in one person, taken about a month later. We were particularly interested in immunological readouts and included both HPV(+) and HPV(-) HNSCC. Up to 6 tumor samples (replicates) per patient were collected from a total of 16 patients. Gene expression analysis using RNA-Seq revealed consistent hierarchical clustering and a high level of correlation between replicates in the same patient. Detailed analysis of immunological gene signatures showed clustering of replicates by patient. Overall our data show a stable transcriptional immunological signature between multiple samples from the same patient and between diagnosis and surgical resection. This suggests that if applied to treatment study, the approach has the potential to allow definition of treatment consequences. Our data further suggest that immunological heterogeneity within the primary tumor is not the cause for treatment failure in HNSCC.

MATERIALS AND METHODS

Sample acquisition & consent.

Following LREC approval and written informed consent, multi-region tumor samples (replicates) were collected under general anesthesia and were snap frozen immediately at diagnostic biopsy and surgical resection. Pre-resection replicates were taken before vascular ligation to minimize hypoxic time. Spatial heterogeneity was assessed by collection of multiple replicates at either diagnostic sampling or at resection. Samples were taken at least 10mm apart to maximize the chance of capturing spatial heterogeneity. The comparison between replicates collected at diagnostic sampling to those replicates from the same patient at resection were further used to assess change over time (temporal heterogeneity), in the recognition of the fact that this comparison may also be shaped by any spatial heterogeneity in any given cancer. Patient demographics, clinical details and number of replicates are shown in Supplementary Table 1. For RNA extraction, cryosections (10 μ m) were cut and RNA isolated with the RNeasy Mini Kit (Qiagen Ltd., Manchester, UK). From 16 patients, 44 samples were collected. Samples with transcript bias or RIN score <7 were removed (n=5); this led to the exclusion of two patients who as a result only had a single sample but no replicate. The full analysis included 14 patients and n=37 tumor replicates; from 6 patients 15 samples were collected at a single timepoint, 8 patients had replicates (n=22) from two timepoints, both diagnosis and resection, with an average of 26 days between them (Supplementary Table 1). Clinical information for an additional cohort of patients also assessed between diagnostic biopsy and surgical resection is shown in Supplementary table 1.

RNA Sequencing.

RNA quality was assessed using the Agilent 2100 Bioanalyzer generating an RNA integrity number (RIN) (Agilent Technologies UK Ltd., UK). Poly-A mRNA was purified from total RNA (100ng) using the Poly(A) Purist Mag Kit (Life Technologies Ltd., UK), according to the manufacturer's instructions. RNA was converted into a library for sequencing using the TruSeq™ stranded mRNA Sample Preparation Kit (Illumina Inc., USA) followed by hybridization to the flow cell for single-end (SE 35bp) sequencing on the HiSeq 2000 (Illumina Inc., USA). Low-quality SE read data from FASTQ files were trimmed or removed. High-quality

reads were mapped to the human genome (hg19) using TopHat (23) (version 2.0.9) and, following the removal of multi-mapping reads, converted to gene specific read counts for annotated genes using HTSeq-count (version 0.5.4) (24). A sample was considered positive for HPV if there were greater than 500 HPV assigned reads against the HPV reference genome (12) and confirmed with Illumina Pathseq (data not shown) (25). Data is available at: ArrayExpress accession E-MTAB-4546 (public upon manuscript publication).

Immunohistochemistry.

Tumor grade and differentiation were recorded from formalin fixed, paraffin embedded (FFPE) tissue. Immunohistochemical (IHC) staining for CD8a (anti-CD8a antibody (clone: C8/144B, DAKO) was performed on each frozen tumor tissue replicate for each case using a region immediately adjacent to that used for RNA-Seq analysis. An insufficient amount of material remained for cases #6, #11 and #13. An additional group of 9 patients were added in which FFPE from diagnostic material was compared to resection specimen again using anti-CD8a for IHC evaluation. Tumor-infiltrating lymphocytes (TILs) were quantified using an Olympus CKX41 with an average of 10 high-power (x400) fields; an average intratumoral score per high-power field (HPF) was calculated.

Statistical analysis.

Raw counts from RNA-Seq were processed in Bioconductor package DESeq, variance was estimated and size factor normalized. Correlations and statistics were performed in Corrplot 0.73, in an R statistical environment (3.3). Qlucore Omics software (3.2) was used to visualize the normalized RNA-Seq data using Principle component analysis and hierarchical clustering. These were performed on data using the setting: mean=0, variance=1 normalization in the Qlucore Omics software (3.2). Where the clustering is agglomerative, average linkage criterion was used to determine the distance between sets of samples and hierarchical clustering is based on Euclidean distances, where the Euclidean distance will be equal to $\sqrt{1-r^2}$, where r is the Pearson correlation coefficient. In order to identify the 4000 most variable genes, variance was set at 0.2879. Briefly genes with a standard deviation (SD) less than a specified variance cut

off from the maximal SD were removed using Qlucore Omics Explorer (3.2) as previously reported (22). Data were visualized in heat maps where each row represents normalized gene expression values for a given gene; each column represents the gene expression for a given tumor: red shading denotes greater gene expression; blue shading denotes lower gene expression. Hierarchical clustering (cluster method=average linkage; distance measure=Pearson correlation) of genes and tumors based on their expression profile is reflected in the dendrograms to the left and the top of the heatmap, respectively. A curated list of immune genes relating to immune cell markers, effector function and exhaustion/ regulatory function were derived from the gene ontology terms GO:0002250 adaptive immune response, GO:0002449 lymphocyte mediated immunity and GO:0002456 T cell mediated immunity. As a measure of similarity and dissimilarity for inpatient and outpatient samples, we quantified Euclidean distances for the 4000 most variable genes (mean=0, variance=1 normalized) in the R statistical environment (3.3) between all replicates from the same patient to all other replicates, where the Euclidean distance is equal to the $\sqrt{1-\text{pearson}}^2$. This measures distances between samples in the hierarchical tree, with smaller distances between replicates indicating a closer relationship. The median distances were calculated for all replicates to all other replicates (outpatient) and compared to the median distances from inpatient replicates, e.g. case #14 median distance for inpatient derived from n=6 and compared to median distance to all other cases n=31 (each replicate comparison is plotted). The Wilcoxon matched-pairs signed rank test was used to compare median distances between replicates from the same patient to those from different patients. Reads per kilobase per million mapped reads (RPKM) were used to display expression of genes between replicates in dot plot comparisons. Spearman correlation analysis between *CD8A* gene expression and CD8 T-Cell IHC count were performed; correlation analysis was also carried out on CD8 T-cell IHC count between diagnostic biopsy and surgical resection from FFPE samples in the second cohort.

Role of the funding source:

The funding for this work was provided by Cancer Research UK grant number C491/A15951.

RESULTS

Replicates from the same patient cluster hierarchically and are significantly related.

The top 4000 most variable genes across the cohort were displayed by variance filtering of the RNA-Seq data. These were hierarchically clustered and displayed in a heatmap (Figure 1A). The 37 replicates from 14 patients evaluated at either diagnosis or resection cluster precisely by patient in 9/14 cases. The same genes were visualized with principle component analysis (PCA) in Figure 1B showing grouping of replicates by patient, complementing the hierarchical clustering. In addition to this, the assessment of all genes (n=18979) and all cases was performed using Hierarchical clustering and PCA (Supplementary Figure 1A and B) again showing replicates grouping by case. In order to quantify the level of variation between replicates, the Euclidean distance was assessed for each replicate (Figure 2A). The graph displays each case and the median inpatient (circle) distance compared to the median interpatient distance (diamond). The length of the line connecting circles and diamonds is a measure of difference when replicates from one case are compared to all other samples. A long line indicates large differences (e.g. case #13), conversely a short line highlights cases where this difference is small. The minimum and maximum distance for each replicate is shown in Supplementary Figure 2. The individual Euclidean distances are presented as a scatter graph in Figure 2B demonstrating that tumor samples from the same patient (inpatient) are significantly closer to each other than to those from other patients (interpatient) ($p < 0.0001$).

Replicates from the same patient are highly correlated.

Correlation analysis of the top 4000 most variable genes was carried out across all samples (Figure 3). The r values in the correlation matrix range from $r=1$ (highly correlated, yellow) to $r=0.5$ (less correlated, blue with values <0.5 also shown as blue). It shows that inpatient tumor replicates are more correlated to each other than interpatient replicates. Case #14 shows a variable correlation between samples, but there is a higher correlation between its own replicates than to samples from other cases. The global median correlation coefficient for inpatient replicates was $r=0.82$ compared to an $r=0.63$ for interpatient replicates. When cases are displayed in a correlation matrix, cases #13 and #9 have similar gene expression; this was also visible in the heatmap (Figure 1A). Our data suggest that each patient's tumor has a distinct

transcriptomic landscape. Correlation analysis of all genes across all samples with the inclusion of the less variable genes increases the correlation between samples (r scale=0.9–1); again replicates from the same patient are most similar to each other (Supplementary Figure 3).

Immunological gene expression in tumor replicates from the same patient is consistent.

The prognostic importance of T-cell infiltration in HNSCC is well-documented (10-12). Gene expression for immune cell lineage markers, effector function, exhaustion and regulatory genes were therefore represented in a heatmap (Figure 4). Samples are ordered by case number and cases with replicates at a single timepoint were grouped, as were cases with replicates from two timepoints. We observe remarkable visual consistency in the immunological signature between replicates from the same patient including the markers of immune effector function (*IFNG*/*GZMA* (5)) and targets of cancer immunotherapy (*PDL1*/*CTLA4* (26)). In addition to this, the RPKM of *CD3E*, *GZMA*, *IFNG*, *CTLA4* and *PDL1* (CD274) are displayed as dot plots (Supplementary Figure 4A to E). Cases #1, #8, #5, #6 and #7 contain a consistent and low relative level of immune-gene expression, while cases #16, #9 and #13 have consistently high immune-gene expression. However, in some cases, replicates #2, #3, #6, #10 and one sample from case #14, the level of immune-gene expression is more variable, this is reflected in the hierarchical clustering of the tumor replicates for immune genes only (Supplementary Figure 5).

Correlation of CD8 by IHC and gene expression.

We evaluated how the gene expression for *CD8A* compared to the assessment of CD8 cell counts by standard IHC. Using tumor tissue sections that were immediately adjacent to those used for RNA-Seq, we were able to assess 11/14 cases. Gene expression for *CD8A* is presented as dot plots for cases grouped by whether we had replicates from a single timepoint or two timepoints (Figure 5A). CD8+ counts (average of 10 HPF) evaluated by IHC are shown in Figure 5B. Visually, the cases with high *CD8A* transcript count appear to be equally rich in CD8 T-cells. This can be quantified by assessing the Spearman correlation of the *CD8A* gene expression to CD8 T-cell IHC count with an $r=0.82$ (Supplementary Figure 6A). We asked if this degree of immunological homogeneity assessed by immunohistochemistry was unique to our cohort and

evaluated the immune cell counts in an additional tumor cohort of paired samples (Figure 6). The paired tumor replicates between diagnostic biopsy and surgical resection showed comparable levels of CD8+ T-cells with a Spearman correlation of $r=0.95$ (Supplementary Figure 6B).

DISCUSSION

In solid cancers, morphological heterogeneity is commonly observed and raises the concern of how representative of the whole cancer biopsy samples are. This is important clinically as biopsy samples are used to make treatment decisions and failure to appreciate heterogeneity could contribute to treatment failure. In addition, if tumor samples are used to gain detailed insights into molecular events resulting from therapeutic intervention(s), understanding transcriptomic heterogeneity becomes critical. Recent high-resolution studies have identified that beyond morphological features, genetic features can vary within the geography of an individual's cancer (4,6,7). Therefore, we wished to quantify tumor and immunological heterogeneity at high resolution using RNA sequencing in multiple tumor biopsy samples from individual patients. Our focus was to understand gene expression that reflects adaptive immune attack; as boosting this is the key for success of immunotherapy. We have previously shown that in HNSCC, TIL density is tightly linked to outcome, in both HPV(+) and HPV(-) cancers (11,13). Here we used RNA-Seq to quantitate the immune signatures and global tumor gene expression profiles from tumor biopsies, separated in space (more than one sample at the same time) and time (samples taken at diagnosis and compared to samples from resection). Molecular immune cell quantification was then correlated to data generated by the current 'gold standard', immunohistochemical assessment of HNSCC patients, which is known to be important for outlook (10,11,13,27).

The gene expression data generated using RNA-Seq for our patient cohort were assessed by PCA, hierarchical clustering and correlation analysis of the top 4000 most variable genes, as well as evaluation of specific immune-genes. Hierarchical clustering and Euclidean distance measures showed that multiple samples from the same patient were significantly more similar to each other than to those from other patients, this was also mirrored in the correlation analysis. Our data suggest a single sample is able to capture the key immune characteristics of the patient's primary HNSCC, irrespective of whether patient samples were taken at the same time or at different timepoints. Tumor biopsies clustered by patient and also displayed similar expression levels of the immune genes. In contrast, there is considerable variability between patients (11,13,28). We have previously identified that the global transcriptomic signature of T-cells is similar between HPV(+) and HPV(-) tumors, but that HPV(-) cancers have a lower overall density of immune

cell infiltrates (11,13). We find here that the similarity of tumor replicates is maintained in both HPV(+) and HPV(-) HNSCC, but recognize that our series is limited by its size.

Overall our data are suggestive of a consistent immune signature in an individual tumor across both location and time, at least within the window between diagnosis and resection; additionally, our data demonstrate that tumors can be grouped by the immune-gene expression profiles. The evaluation of gene expression of key markers for immune attack (*IFNG/ GZMA* (5)) and targets of cancer immunotherapy (*PDL1/ CTLA4* (26)) further confirm a stable immune signature in the individual patient.

Reassuringly, gene expression of *CD8A* correlated strongly ($r=0.82$) with the CD8 T-cell count assessed by manual counting of cells using IHC. Cell counts generated by IHC are also stable over time and space, this observation was confirmed in an independent cohort of paired samples from biopsy and resection specimens.

Nonetheless inpatient tumor replicates are not a perfect match with each other and in some cases (cases #2, #3, #10, #11, #14 and #15) the tumor samples did not cluster perfectly by patient: In case #10, gene expression differences were observed at the immune-gene level which were also reflected in the IHC of CD8. Conversely, case #2 displayed a stable CD8 and immune cell marker signature but variability in the expression of effector function genes. Cases #2, #6, #10 and #14 were laryngeal tumors, which have been highlighted as having a higher level of mutational heterogeneity compared to other sites in HNSCC (29). We are currently evaluating whether the degree of transcriptomic heterogeneity links to the genomic heterogeneity identified by other groups (1,30).

Our data support that in HNSCC the global tumor and adaptive immune signatures are stable across space and time between replicate samples from the same tumor. Our data have a number of implications. They suggest that immunological heterogeneity is not likely to be a key reason for immunotherapy failure for the primary cancer in HNSCC; paired analysis of primary and metastatic disease is necessary to evaluate this question further for patients with later stage disease. Our data also imply that it is important to examine in other tumors whether a consistent immunological ‘fingerprint’ is present even in cancers with known genetic heterogeneity. Finally, our data support that from a background of transcriptomic stability in untreated patients, RNA sequencing may be useful for the detection of change resulting from treatment effects.

REFERENCES

1. Alizadeh AA, Aranda V, Bardelli A, Blanpain C, Bock C, Borowski C, *et al.* Toward understanding and exploiting tumor heterogeneity. *Nature medicine* **2015**;21(8):846-53 doi 10.1038/nm.3915.
2. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoediting: from immunosurveillance to tumor escape. *Nature immunology* **2002**;3(11):991-8 doi 10.1038/ni1102-991.
3. Hensley CT, Faubert B, Yuan Q, Lev-Cohain N, Jin E, Kim J, *et al.* Metabolic Heterogeneity in Human Lung Tumors. *Cell* **2016**;164(4):681-94 doi 10.1016/j.cell.2015.12.034.
4. Jamal-Hanjani M, Quezada SA, Larkin J, Swanton C. Translational implications of tumor heterogeneity. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2015**;21(6):1258-66 doi 10.1158/1078-0432.CCR-14-1429.
5. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **2015**;160(1-2):48-61 doi 10.1016/j.cell.2014.12.033.
6. Dunne PD, McArt DG, Bradley CA, O'Reilly PG, Barrett HL, Cummins R, *et al.* Challenging the cancer molecular stratification dogma: Intratumoral heterogeneity undermines consensus molecular subtypes and potential diagnostic value in colorectal cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2016** doi 10.1158/1078-0432.CCR-16-0032.
7. Gyanchandani R, Lin Y, Lin HM, Cooper KL, Normolle DP, Brufsky AM, *et al.* Intra-tumor heterogeneity affects gene expression profile test prognostic risk stratification in early breast cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2016** doi 10.1158/1078-0432.CCR-15-2889.
8. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf AC, *et al.* Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* **2013**;39(4):782-95 doi 10.1016/j.immuni.2013.10.003.
9. Seiwert TY, Burtneß B, Mehra R, Weiss J, Berger R, Eder JP, *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. *The lancet oncology* **2016**;17(7):956-65 doi 10.1016/S1470-2045(16)30066-3.
10. Ward MJ, Mellows T, Harris S, Webb A, Patel NN, Cox HJ, *et al.* Staging and treatment of oropharyngeal cancer in the human papillomavirus era. *Head & neck* **2015**;37(7):1002-13 doi 10.1002/hed.23697.
 11. Ward MJ, Thirdborough SM, Mellows T, Riley C, Harris S, Suchak K, *et al.* Tumour-infiltrating lymphocytes predict for outcome in HPV-positive oropharyngeal cancer. *British journal of cancer* **2014**;110(2):489-500 doi 10.1038/bjc.2013.639.
 12. Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **2015**;517(7536):576-82 doi 10.1038/nature14129.
 13. Wood O, Woo J, Seumois G, Savelyeva N, McCann KJ, Singh D, *et al.* Gene expression analysis of TIL rich HPV-driven head and neck tumors reveals a distinct B-cell signature when compared to HPV independent tumors. *Oncotarget* **2016** doi 10.18632/oncotarget.10788.
 14. McBryan J, Fagan A, McCartan D, Bane FT, Vareslija D, Cocchiglia S, *et al.* Transcriptomic Profiling of Sequential Tumors from Breast Cancer Patients Provides a Global View of Metastatic Expression Changes Following Endocrine Therapy. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2015**;21(23):5371-9 doi 10.1158/1078-0432.CCR-14-2155.
 15. Wyatt AW, Mo F, Wang K, McConeghy B, Brahmabhatt S, Jong L, *et al.* Heterogeneity in the inter-tumor transcriptome of high risk prostate cancer. *Genome Biol* **2014**;15(8):426 doi 10.1186/s13059-014-0426-y.
 16. Zhang P, Mirani N, Baisre A, Fernandes H. Molecular heterogeneity of head and neck squamous cell carcinoma defined by next-generation sequencing. *The American journal of pathology* **2014**;184(5):1323-30 doi 10.1016/j.ajpath.2014.01.028.
 17. Pearce DA, Arthur LM, Turnbull AK, Renshaw L, Sabine VS, Thomas JS, *et al.* Tumour sampling method can significantly influence gene expression profiles derived from neoadjuvant window studies. *Scientific reports* **2016**;6:29434 doi 10.1038/srep29434.
 18. Garon EB, Rizvi NA, Hui R, Leigh N, Balmanoukian AS, Eder JP, *et al.* Pembrolizumab for the treatment of non-small-cell lung cancer. *The New England journal of medicine* **2015**;372(21):2018-28 doi 10.1056/NEJMoa1501824.

19. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, *et al.* Improved survival with ipilimumab in patients with metastatic melanoma. *The New England journal of medicine* **2010**;363(8):711-23 doi 10.1056/NEJMoa1003466.
20. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, *et al.* The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nature medicine* **2015**;21(8):938-45 doi 10.1038/nm.3909.
21. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJ, Robert L, *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **2014**;515(7528):568-71 doi 10.1038/nature13954.
22. Engel I, Seumois G, Chavez L, Samaniego-Castruita D, White B, Chawla A, *et al.* Innate-like functions of natural killer T cell subsets result from highly divergent gene programs. *Nature immunology* **2016**;17(6):728-39 doi 10.1038/ni.3437.
23. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **2009**;25(9):1105-11 doi 10.1093/bioinformatics/btp120.
24. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**;31(2):166-9 doi 10.1093/bioinformatics/btu638.
25. Kostic AD, Ojesina AI, Pedamallu CS, Jung J, Verhaak RG, Getz G, *et al.* PathSeq: software to identify or discover microbes by deep sequencing of human tissue. *Nature biotechnology* **2011**;29(5):393-6 doi 10.1038/nbt.1868.
26. Ahmad SS, Qian W, Ellis S, Mason E, Khattak MA, Gupta A, *et al.* Ipilimumab in the real world: the UK expanded access programme experience in previously treated advanced melanoma patients. *Melanoma research* **2015**;25(5):432-42 doi 10.1097/CMR.000000000000185.
27. King EV, Ottensmeier CH, Thomas GJ. The immune response in HPV+ oropharyngeal cancer. *Oncoimmunology* **2014**;3(1):e27254 doi 10.4161/onci.27254.
28. Ottensmeier CH, Perry KL, Harden EL, Stasakova J, Jenei V, Fleming J, *et al.* Upregulated Glucose Metabolism Correlates Inversely with CD8+ T-cell Infiltration and Survival in Squamous Cell Carcinoma. *Cancer research* **2016** doi 10.1158/0008-5472.CAN-15-3121.

29. Ledgerwood LG, Kumar D, Eterovic AK, Wick J, Chen K, Zhao H, *et al.* The degree of intratumor mutational heterogeneity varies by primary tumor sub-site. *Oncotarget* **2016**;7(19):27185-98 doi 10.18632/oncotarget.8448.
30. McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, *et al.* Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **2016**;351(6280):1463-9 doi 10.1126/science.aaf1490.

Acknowledgments:

The authors would like to thank the patients and clinical trials assistants for providing and handling of the samples, respectively. The authors acknowledge the use of the IRIDIS High Performance Computing Facility, and the research histology team with their associated support services at the University of Southampton, in the completion of this work. Finally, the authors would like to thank the La Jolla Bioinformatics and sequencing core for processing of the transcriptome libraries.

Authors' contributions:

CO, EK and VP designed, led and supervised the study. EK biopsied and surgically resected the samples analyzed in this study; OW undertook the sample processing. OW, JC , TW and JW undertook the bioinformatical analysis. VP, CO, OW, EK, JC generated the experimental and analysis design. GT and AM facilitated analysis of the IHC; counted by OW.

Figure legends:

Figure 1. Analysis of tumor replicates using hierarchical clustering and Principle Component Analysis (PCA). RNA-Seq analysis showing row wise z-scores of normalized read counts for the 4000 (variance filtered) most variable genes across the tumor replicates. Patient tumor replicates are color coded and displayed on the heatmap and principle component analysis (PCA) plot; HPV(+)=black and HPV(-)=beige; diagnostic biopsy (DB) and surgical resection (SR) replicates are annotated below. (A) Hierarchical clustering (distance measure=Pearson's correlation metric; clustering=average linkage method) of tumor replicates displayed as a heatmap shows close clustering of related samples. (B) PCA was used to visualize the sample to sample distances and also highlights the similarities between the tumor replicates.

Figure 2. Comparison of Euclidean distance from hierarchical clustering of tumor replicates. The distances between paired tumor replicates in the hierarchical tree were calculated and used to assess how closely related inpatient replicates are, compared to interpatient replicates. The median Euclidean distance ($\sqrt{1-\text{pearson}^2}$) for inpatient replicates (circles) was plotted, as was the median Euclidean distance for interpatient replicates (diamonds). (A) Displays the relationship between the intra and interpatient median Euclidean distances for each replicate in the cohort respectively. The length of the connecting line between circles and diamonds is a measure of the difference between an individual replicate and all samples from all other patients. (B) Represents the median inpatient distances compared to the median interpatient distances. There is a significant difference between inpatient and interpatient distances ($p < 0.0001$, Wilcoxon test).

Figure 3. Correlation analysis of tumor replicates. Single timepoint (sampling across space) and two timepoint (sampling across time between diagnosis and resection) tumor replicates were assessed using a correlation matrix of gene expression (Spearman correlation of top 4000 variance filtered genes); each tumor replicate's gene expression was correlated to itself and to each other sample. Inpatient tumor replicates were more correlated with a median correlation of $r=0.82$ compared to an interpatient median correlation $r=0.63$.

Figure 4. Gene expression of immune markers is consistent in tumor replicates from the same patient. RNA-Seq analysis showing row wise z-scores of normalized read counts for genes associated with immune lineage markers, cytotoxic function, exhaustion and regulatory function. Samples were grouped by patient and then by whether replicates had been collected at a single timepoint or at two timepoints. Expression profiles are consistent across tumor replicates in most cases.

Figure 5. Comparison of *CD8A* gene expression with CD8 immunohistochemistry (IHC) counts across time and space. (A) The gene expression of *CD8A* (RPKM – Reads per kilobase per million mapped reads) is shown for the cases grouped by single timepoint and two timepoint tumor replicates. (B) CD8 IHC counts are shown as the mean across 10 high power fields (HPF) on frozen tissue sections taken from the adjacent material used in the RNA-Seq analysis, tumor replicates are arranged as described above. The *CD8A* gene expression and CD8 IHC show a similar pattern of immune density and are consistent across time and space.

Figure 6. Assessment of CD8 immunohistochemistry (IHC) in tumor replicates between diagnostic biopsy and surgical resection in an additional sample cohort. The two timepoint tumor replicates were obtained from an independent patient cohort at diagnostic biopsy and then at surgical resection. Each data point represents CD8 counts shown as mean across 10 high power fields (HPF) from FFPE tissue blocks. In each case, the mean CD8 count in the biopsy and resection samples were closely mirrored.

Figure 1.

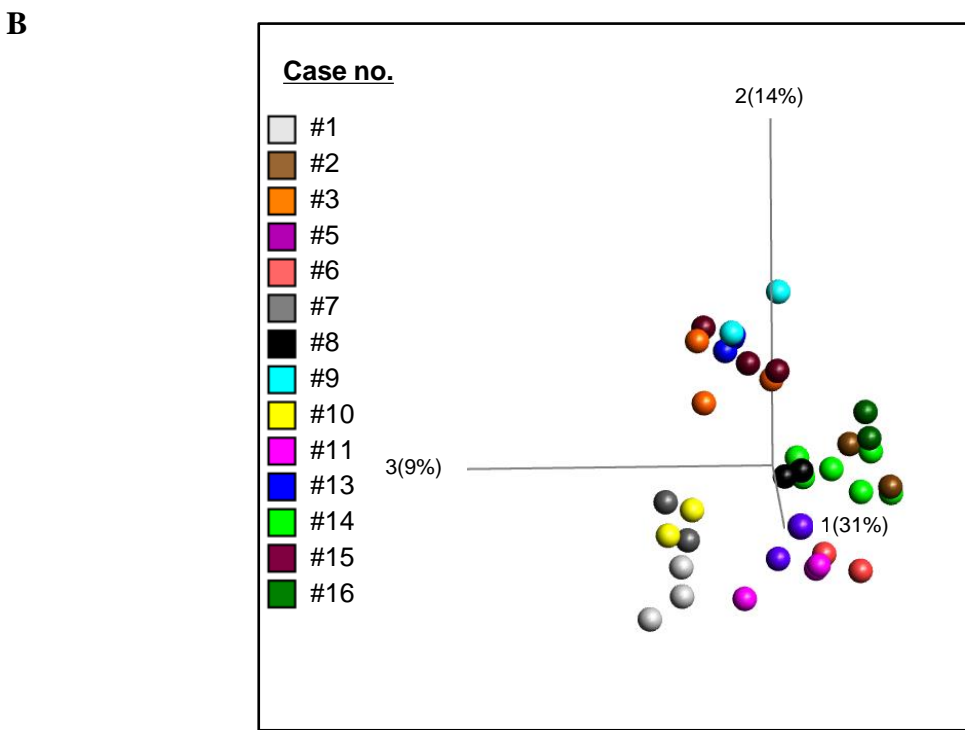
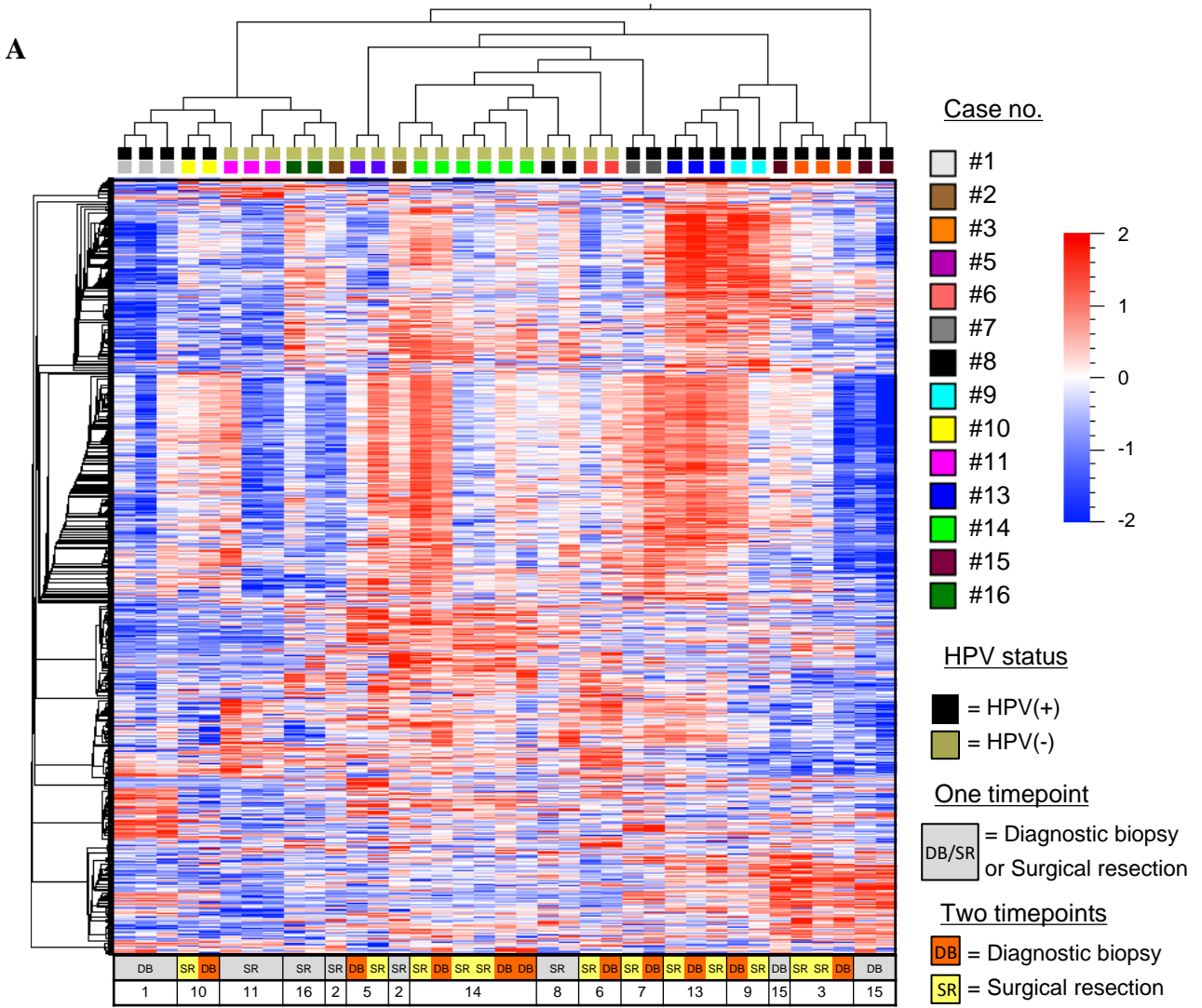


Figure 2.

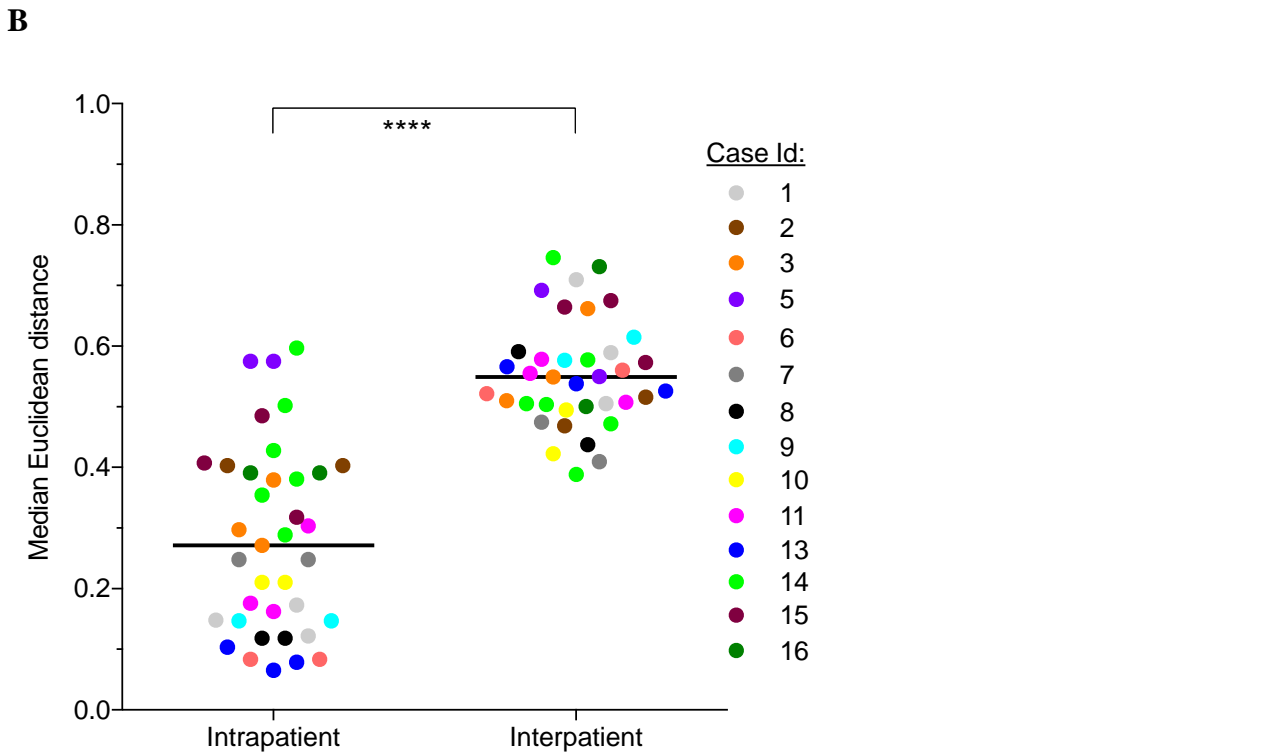
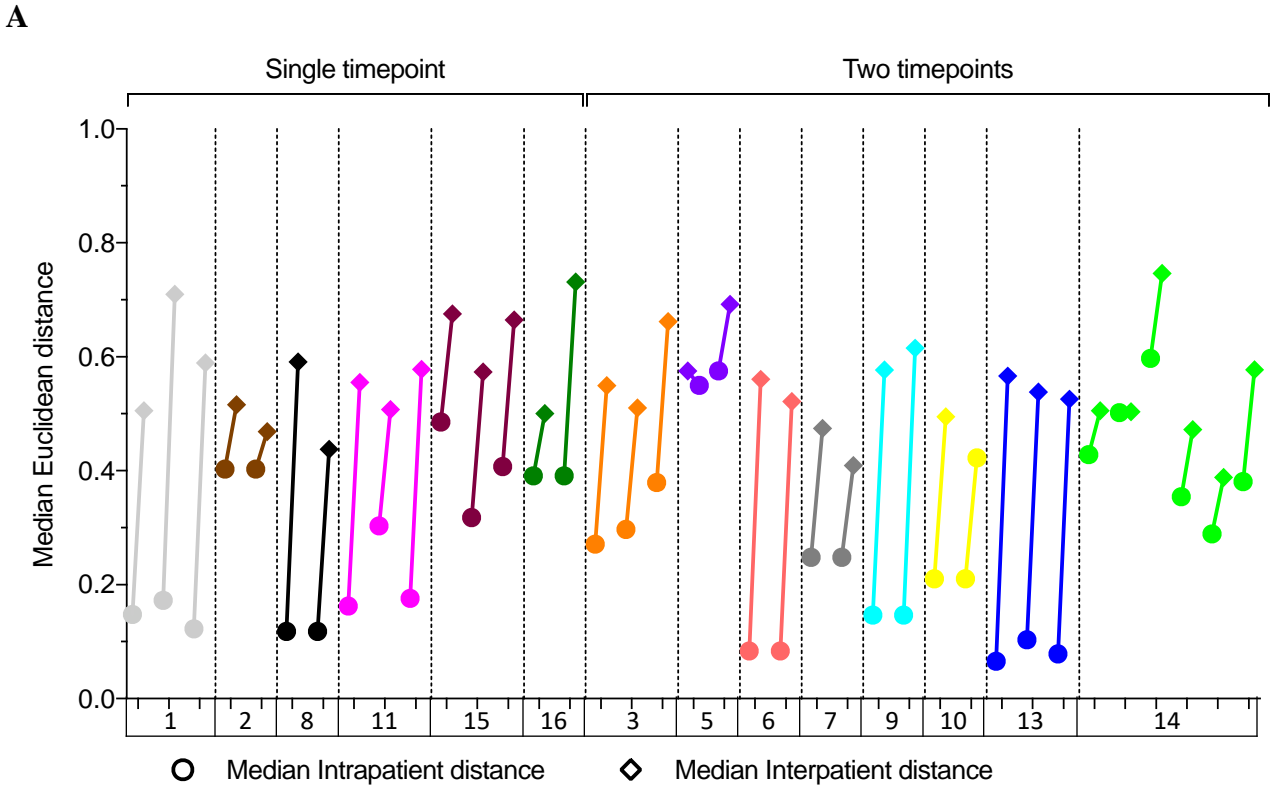
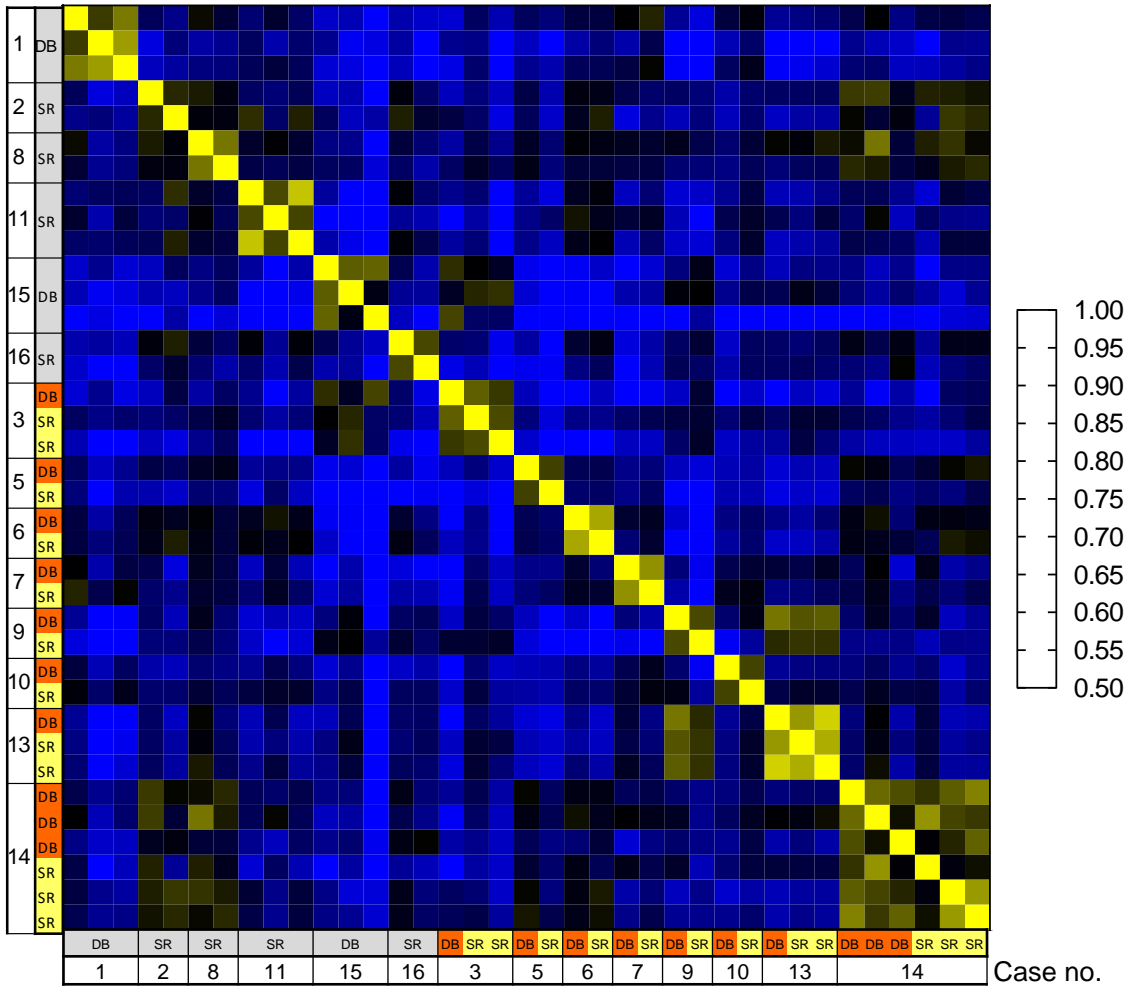


Figure 3.



Spearman correlation	r Inpatient	r Interpatient
Median	0.82	0.63
Min	0.76	0.26
Max	0.93	0.86

One timepoint

DB/SR = Diagnostic biopsy or Surgical resection

Two timepoints

DB = Diagnostic biopsy
SR = Surgical resection

Figure 4.

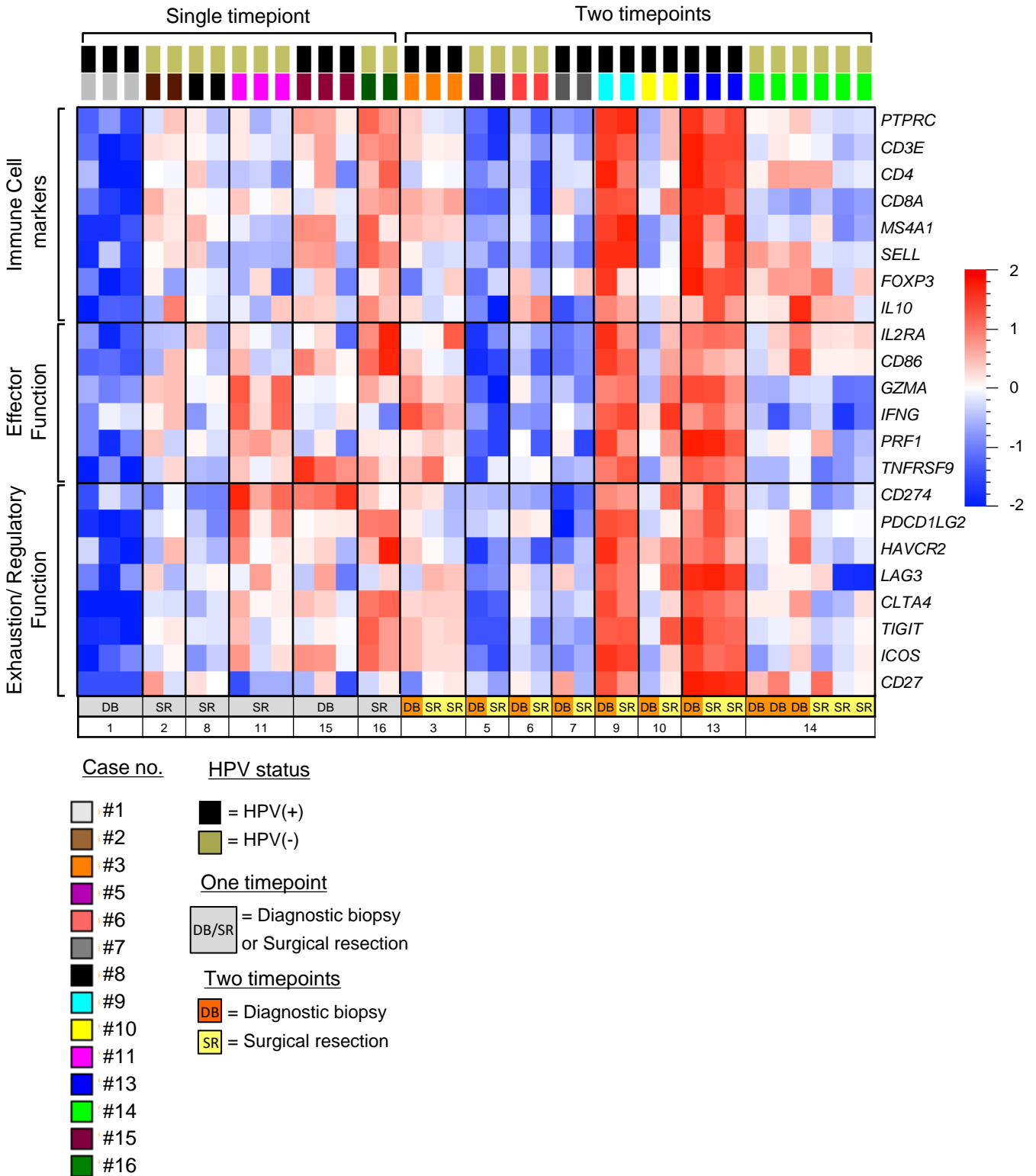
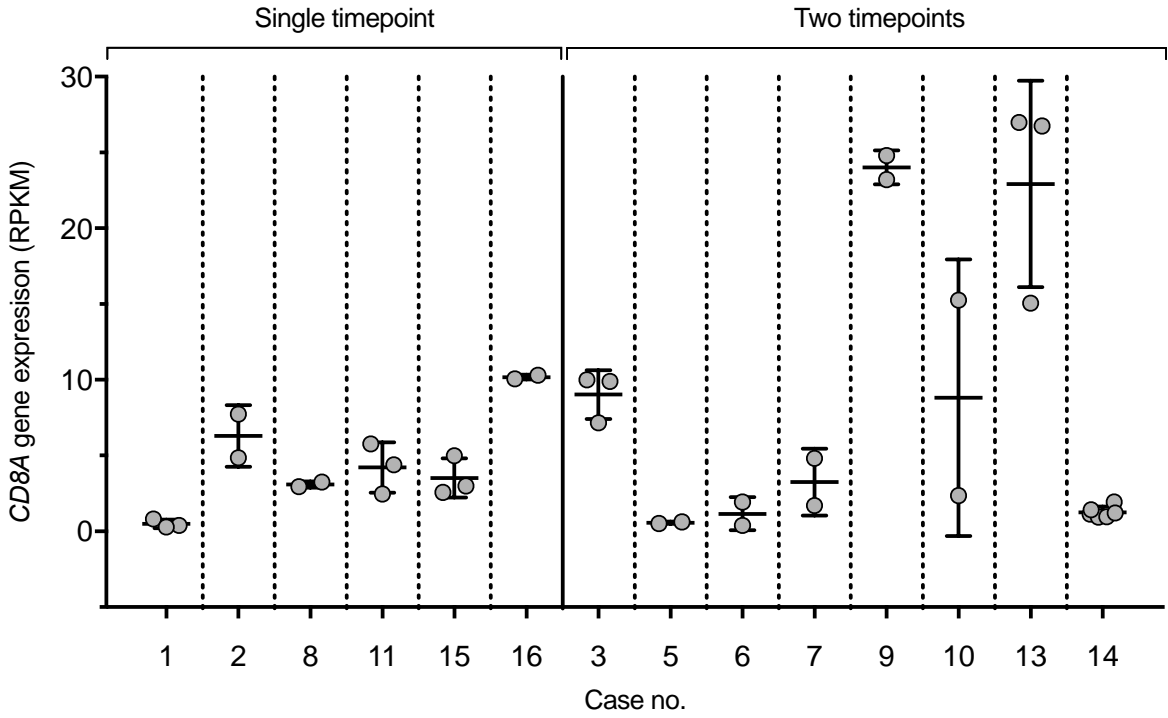


Figure 5.

A



B

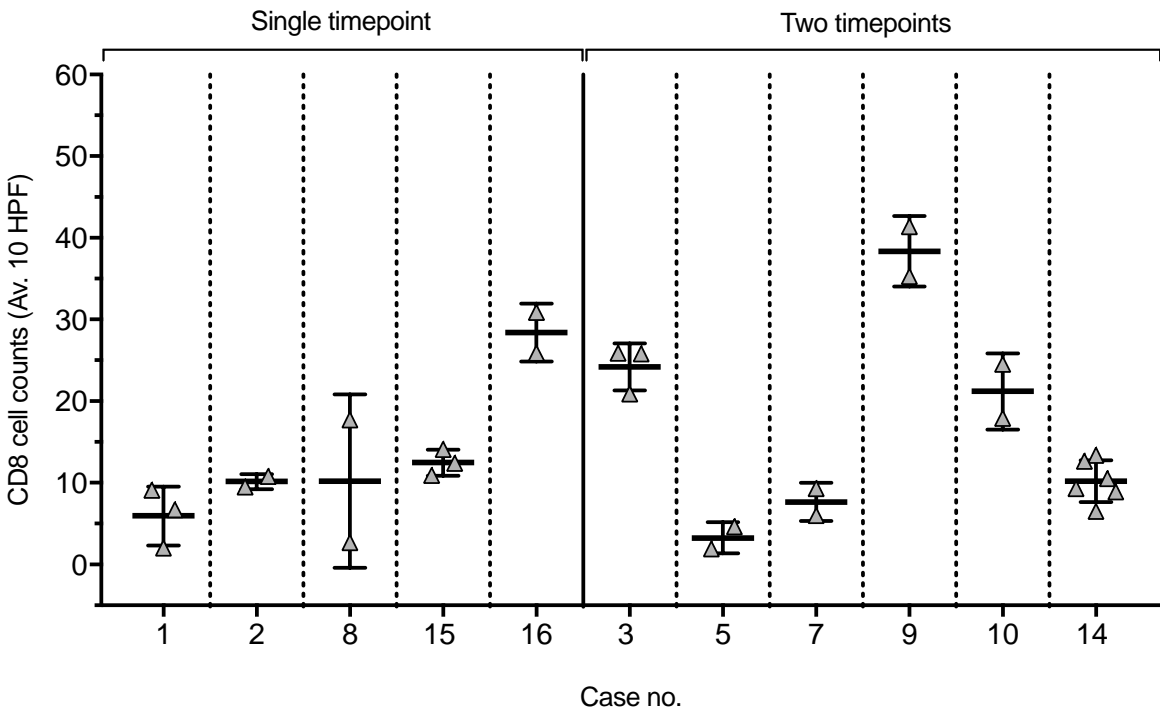
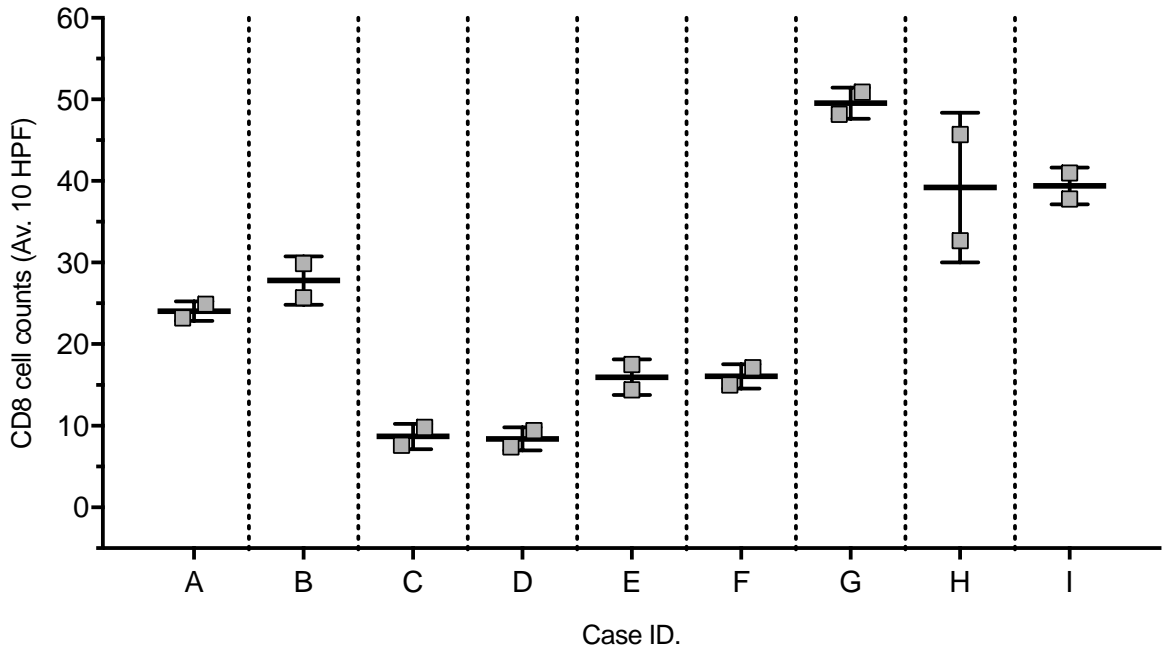


Figure 6.



Clinical Cancer Research

Head and neck squamous cell carcinomas are characterized by a stable immune signature within the primary tumor over time

Oliver Wood, James Clarke, Jeogmin Woo, et al.

Clin Cancer Res Published OnlineFirst September 26, 2017.

Updated version	Access the most recent version of this article at: doi: 10.1158/1078-0432.CCR-17-0373
Supplementary Material	Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2017/09/26/1078-0432.CCR-17-0373.DC1
Author Manuscript	Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.