

Persistent HPV16 E7 Expression Establishes a Preconditioned Epithelial State and Reveals Amplified Transcriptional Responses in Females During Oral Carcinogenesis

Kalani Hettiarachchi

The University of Queensland

Minwoo Kim

The University of Queensland

RLPR Liyanage

National Hospital Galle

Avi-Hai Hovav

Hebrew University <https://orcid.org/0000-0001-5118-571X>

Arutha Kulasinghe

The University of Queensland

Ian Frazer

The University of Queensland

Kelvin Tuong

The University of Queensland

Debottam Sinha

The University of Queensland

Janin Chandra

j.chandra@uq.edu.au

The University of Queensland <https://orcid.org/0000-0002-4456-9480>

Article

Keywords: HPV16 E7, Head and neck squamous cell carcinoma, Immune remodelling, 4-Nitroquinoline-1-oxide (4NQO)

Posted Date: October 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11147202/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: There is no conflict of interest

Abstract

Human papillomavirus (HPV)-associated and carcinogen-driven head and neck squamous cell carcinomas (HNSCC) arise through distinct biological routes, yet how each aetiology shapes the tissue state in which carcinogenesis begins remains unclear. We investigated how HPV16 E7 oncoprotein expression and biological sex influence the oral epithelial response before and during carcinogen exposure. Oral epithelium from male and female wild-type (WT) and HPV16 E7-expressing transgenic (K14E7) mice was analysed before and following 4-nitroquinoline-1-oxide (4NQO) induced oral carcinogenesis using flow cytometry, histopathology, immunohistochemistry, and RNA sequencing, with transcriptional programmes compared with human HPV⁺ and HPV⁻ HNSCC. Persistent HPV16 E7 expression established epithelial hyperplasia and a baseline transcriptional state enriched for cell-cycle, MYC, interferon response and inflammatory programmes, accompanied by selective remodelling of T cell and myeloid cells. HPV16 E7-driven transcriptional remodelling was more pronounced in females, although pathway enrichment was conserved between sexes. Following 4NQO exposure, WT epithelium underwent transcriptional remodelling that recapitulated programmes already active in untreated K14E7 epithelium, whereas K14E7 epithelium showed limited additional transcriptional change despite similar pathological endpoints. Cross-species analyses showed concordance between the murine WT and K14E7 4NQO model and human HPV⁻ and HPV⁺ HNSCC, with sex-associated pathways demonstrating significant, albeit more modest, cross-species concordance.

Together, these findings indicate that HPV16 E7 establishes a preconditioned proliferative and inflammatory epithelial state with early immune remodelling that alters responses to carcinogenic injury. HPV16 E7 may redirect rather than accelerate epithelial remodelling during carcinogenesis, with female transcriptional responsiveness to both oncogene expression and carcinogen exposure emerging as a feature of epithelial adaptation.

Introduction

Head and neck squamous cell carcinoma (HNSCC) represents a biologically heterogeneous malignancy arising from the stratified squamous epithelium of the upper aerodigestive tract, and accounts for approximately 890,000 new cases and more than 450,000 deaths annually ¹. Despite limited survival, HNSCC is increasingly recognised not as a single disease entity, but as a group of malignancies shaped by distinct carcinogenic drivers.

Two major aetiological pathways predominate: carcinogen-associated disease, principally linked to tobacco and alcohol exposure, and human papillomavirus (HPV)-associated disease, most commonly driven by the high-risk oncogenic genotype HPV16. These aetiologies differ substantially in tumour biology, immune organisation and clinical outcome, including differences in antigen-presenting cell composition, T cell states and immunoregulatory activity ². Persistent HPV infection and HPV-associated malignancy are associated with mechanisms that constrain effective T cell responses, including engagement of immune checkpoint pathways such as CTLA4 and PD1/PD-L1 ^{3,4}. Further, HPV⁺

HNSCC also exhibits distinct transcriptional programmes spanning proliferative, immune and tumour-associated processes^{5,6}. However, these aetiological influences are not mutually exclusive. Tobacco and alcohol exposure remain clinically relevant in HPV⁺ disease and are associated with poorer outcomes⁷, raising the possibility that carcinogen-induced epithelial injury interacts with persistent viral oncogene activity during tumour evolution. Most knowledge of these differences, however, derives from established tumours, while the tissue states that precede malignant transformation are considerably less well defined. Persistent oncogene expression may alter epithelial signalling and the local immune environment before histologically overt cancer develops, thereby changing the context in which subsequent carcinogenesis occurs.

Emerging evidence further suggests that biological sex influences tumour immune organisation. HNSCC occurs approximately three times more frequently in males than females, while associations between biological sex and clinical outcome vary across cohorts and appear to be influenced by HPV status^{8–11}. These epidemiological observations are accompanied by evidence that biological sex influences antiviral and antitumour immunity, with females exhibiting enhanced effector T cell function and reduced immunosuppressive activity¹². Nevertheless, whether biological sex modifies epithelial and immune responses to persistent HPV oncogene expression during early carcinogenesis remains unclear.

Defining these early events in human disease is challenging because premalignant evolution occurs over prolonged and largely subclinical periods. Preclinical models therefore provide an opportunity to resolve these processes experimentally. The K14E7 transgenic mouse expresses HPV16 E7 in stratified epithelium and develops epithelial hyperplasia and dysregulated differentiation, accompanied by local immune alterations^{13–17}. Complementarily, oral 4-nitroquinoline-1-oxide (4NQO) exposure produces DNA damage in epithelial cells similar to HNSCC¹⁸, leading to intraepithelial hyperplasia and progressive epithelial lesions, and culminating in squamous cell carcinoma within 3–6 months after exposure. 4NQO thereby models the naturally progressing, multistep carcinogen-driven oral carcinogenesis in humans^{18–20}. Combining these systems enables persistent viral oncogene expression to be examined both before and during carcinogen-driven transformation, providing a framework to distinguish pre-existing HPV16 E7-associated tissue states from responses induced by carcinogenic injury.

Here, we establish and characterise sex-stratified WT/4NQO and K14E7/4NQO models of oral carcinogenesis integrating histopathology, disease progression, high-dimensional immunophenotyping and transcriptomic profiling. We show that persistent HPV16 E7 expression establishes a proliferative and immune-remodelled oral epithelial state before carcinogen exposure, with sex-associated differences in local immune and transcriptional responses. Following 4NQO exposure, WT epithelium acquired many of the transcriptional programmes already present in K14E7 tissue, whereas K14E7 epithelium showed comparatively limited additional remodelling, and cross-species analyses demonstrated broad conservation of these programmes in human HPV⁺ HNSCC.

Results

HPV16 E7 promotes immune accumulation in the oral epithelium

We have previously demonstrated that the hyperproliferative cutaneous epidermis of HPV16 E7-expressing K14E7 mice contains a diverse immune infiltrate comprising lymphoid and myeloid populations and multiple immunosuppressive mechanisms^{13,21–24}. Whether persistent HPV16 E7 expression similarly alters immune organisation within the oral epithelium remained unclear. We therefore examined tongue epithelium from 12-14-week-old WT and K14E7 mice before carcinogen exposure. An initial cohort, comprising predominantly female mice, showed a significant increase in CD45⁺ immune-cell abundance in K14E7 compared with WT tongue epithelium (Fig. 1a). However, immune organisation within the oral mucosa, including potential sex-associated differences, has not been characterised in depth in either K14E7 or WT mice. We therefore performed spectral flow cytometric immunophenotyping of tongue epithelium from male and female WT and K14E7 mice.

In this sex-matched cohort, genotype-associated increases in CD45⁺ immune cell abundance showed a similar trend ($p = 0.0650$), with secondary comparisons indicating higher CD45⁺ frequency in female K14E7 mice than in WT females and male K14E7 mice (Fig. 1b). These findings suggest that persistent E7 expression increases immune occupancy of the oral epithelium, with the magnitude of this response varying according to biological sex.

We next asked whether this increased immune representation reflected broad redistribution of the epithelial immune compartment or selective changes in particular immune populations. High-dimensional spectral flow cytometry, visualised by t-SNE and supported by hierarchical manual gating, identified diverse lymphoid and myeloid populations across WT and K14E7 epithelium with broadly similar global immune architecture across genotypes and sexes (Fig. 1c, Suppl. Fig. S1). The lymphoid-to-myeloid ratio was similar between genotypes (Suppl. Fig. S2).

HPV16 E7 expression increases sex-associated CTLA4 expression across T cell compartments

TCR $\gamma\delta$ ⁺ T cells constituted a greater proportion of the CD45⁺ immune compartment in female than male WT oral epithelium (Fig. 1d). As these analyses quantify immune subsets as a proportion of CD45⁺ cells, they reflect differences in immune composition rather than absolute TCR $\gamma\delta$ ⁺ T cell abundance relative to epithelial cells (Data not shown). The proportion of TCR $\gamma\delta$ ⁺ T cells were reduced in K14E7 compared with WT epithelium, with the genotype-associated difference particularly evident in females. Phenotypically, the remaining TCR $\gamma\delta$ ⁺ population in K14E7 epithelium exhibited increased CTLA4 expression (Fig. 1e), particularly in males.

The overall abundance of TCR β ⁺ T cells were unchanged between genotypes. Separation of antigen-inexperienced and antigen-experienced TCR β ⁺ cells according to CD44 expression similarly revealed no differences in their overall abundance between WT and K14E7 epithelium (Suppl. Fig. S1 and S3). However, CTLA4⁺ cells were increased within the antigen-inexperienced CD44[–]TCR β ⁺ compartment of

K14E7 mice (Fig. 1d), with a more prominent frequency in female K14E7 mice compared to WT females and male K14E7 mice. Within this CTLA4⁺ population, both PD1 and CD25 expression were increased in K14E7 mice (Fig. 1f). CD25 expression additionally showed a male bias, with higher expression in males than females in both genotypes and a genotype-associated increase in males.

Within the antigen-experienced CD44⁺TCRβ⁺ compartment, NK1.1⁺NKT-like cells were increased in K14E7 epithelium (Fig. 1d). Within conventional NK1.1⁻CD44⁺TCRβ⁺ T cells, CTLA4⁺PD1⁻ cells constituted the largest checkpoint-expressing population. CTLA4⁺PD1⁻ cells were increased in K14E7 compared to WT epithelium (Fig. 1d), with the genotype-associated increase significant in males. Within this population, both CTLA4 and CD44 expression were increased in K14E7 compared to WT mice (Fig. 1g). Secondary comparisons indicated higher CTLA4 expression in female K14E7 mice compared to WT females and K14E7 males, whereas the genotype-associated increase in CD44 expression was significant in males.

Together, these findings reveal that persistent HPV16 E7 expression does not broadly expand the conventional T cell compartment, but instead selectively remodels T cell phenotype towards increased CTLA4 expression across multiple lymphocyte populations, with biological sex modifying the magnitude of selected response.

HPV16 E7 reshapes the epithelial myeloid landscape towards macrophage-like APC populations with increased CD86 and CD64 expression

Having identified selective T cell remodelling in K14E7 epithelium, we next asked whether persistent HPV16 E7 expression also altered local myeloid populations. Within the non-T/NK/neutrophil compartment, HPV16 E7 expression remodelled the myeloid landscape of the oral epithelium, characterised by contraction of EpCAM⁺ myeloid populations, particularly EpCAM⁺MHCII^{high} epithelial antigen-presenting cells (APCs), and expansion of EpCAM⁻CD64⁺MHCII⁺ macrophage-like APCs (Fig. 2a, b; Suppl. Fig. S1).

Among EpCAM⁺ myeloid cells, both MHCII^{high}XCR1^{low} and MHCII^{high}XCR1^{high} populations were reduced in K14E7 compared to WT epithelium (Fig. 2c). In contrast, the MHCII^{low/-}XCR1^{low} population showed a striking sex-associated distribution, being enriched in WT females but nearly undetectable in males of either genotype. These cells were CD24⁺CD11c⁺, supporting their myeloid identity and suggesting an immature or precursor epithelial APC phenotype. Secondary comparisons indicated that this population was markedly reduced in K14E7 females compared with WT females (Fig. 2c).

The proportion of EpCAM⁻CD64⁺MHCII⁺ macrophage-like APCs increased markedly in K14E7 epithelium (Fig. 2b). This expansion was driven predominantly by the CCR2^{low/-}XCR1^{low/-} subset, with a more modest increase in CCR2⁺XCR1^{high} macrophage-like APCs (Fig. 2c). Secondary comparisons also indicated sex-associated variation within these populations (Fig. 2c). Additionally, a CCR2⁺XCR1⁺ non-macrophage myeloid population was reduced in K14E7 epithelium, whereas overall neutrophil

abundance remained largely unchanged (Suppl. Fig. S4a-c). Within K14E7 mice, however, neutrophils were less abundant in females than males ($p = 0.0067$; Suppl. Fig. S4c).

Expression of the costimulatory molecule CD86 and the mononuclear phagocyte marker CD64 (FcγRI) was increased across multiple EpCAM⁺ myeloid and macrophage-like APC populations in K14E7 epithelium (Fig. 2d, e). CD86 was increased in the MHCII^{high}XCR1^{low}, MHCII^{high}XCR1^{high} and MHCII^{low/-}XCR1^{low} EpCAM⁺ populations, as well as the CCR2^{low/-}XCR1^{low} and CD64⁺MHCII⁺CCR2⁻XCR1⁻ macrophage-like populations (Fig. 2d). CD64 was similarly increased in several EpCAM⁺ and macrophage-like populations, including MHCII^{high}XCR1^{low}, MHCII^{high}XCR1^{high}, MHCII^{low/-}XCR1^{low}, CCR2^{low/-}XCR1^{low} and CD64⁺MHCII⁺CCR2⁻XCR1⁻ populations (Fig. 2e). PD-L1 expression remained largely unchanged (Suppl. Fig. S4d).

Together, these findings indicate that persistent HPV16 E7 expression reconfigures the oral myeloid compartment through contraction of EpCAM⁺ myeloid populations and expansion of macrophage-like APCs, accompanied by increased CD86 and CD64 expression.

HPV16 E7 establishes a conserved proliferative and inflammatory epithelial transcriptional programme with enhanced remodelling in females

To determine how HPV16 E7 expression remodels the oral epithelial transcriptome, RNA sequencing was performed on tongue epithelium from untreated WT and K14E7 mice aged approximately 12–14 weeks. Principal component analysis (PCA) demonstrated clear separation between K14E7 and WT samples, indicating that continuous HPV16 E7 expression establishes a distinct epithelial transcriptional state (Fig. 3a). Samples clustered primarily according to genotype, whereas male and female samples largely overlapped within each genotype, consistent with genotype being the dominant source of transcriptional variation.

Comparison of K14E7 and WT oral epithelium identified 1,066 differentially expressed genes (DEGs), comprising 842 upregulated and 224 downregulated genes in K14E7 epithelium (Fig. 3b). Hallmark Gene Set Enrichment Analysis (GSEA) demonstrated that HPV16 E7 expression established a transcriptional program dominated by proliferative and inflammatory signalling (Fig. 3c, d), including E2F targets, G2M checkpoint, MYC targets, interferon responses, inflammatory response, TNF signalling via NF-κB and IL6-JAK-STAT3 signalling. In contrast, WT epithelium showed relevant enrichment of EMT-associated and metabolic pathways. This EMT-associated signal was driven predominantly by extracellular matrix- and structural tissue-related genes that were comparatively reduced in the hyperproliferative K14E7 epithelium (Fig. 3d).

To determine whether the transcriptional response to HPV16 E7 differed between sexes, the effect of genotype was assessed separately within each sex by comparing the epithelial transcriptome of K14E7 females to WT females and that of K14E7 males to WT males (Fig. 3e, f). The K14E7-versus-WT comparison identified substantially more DEGs in females than in males, including 1,817 detected only in females and 147 detected only in males, while 485 DEGs were shared between the two comparisons.

Thus, HPV16 E7-associated transcriptional remodelling was substantially broader in females, although a shared core response was retained between sexes. Consistent with this, Hallmark GSEA showed highly concordant pathway enrichment between the female and male K14E7-versus-WT comparisons (Fig. 3f). Exploratory Hallmark over-representation analysis of the sex-partitioned DEG sets (Suppl. Fig. S5) showed that shared DEGs were enriched for core HPV16 E7-associated proliferative and interferon-response pathways, whereas DEGs detected only in females showed additional enrichment of inflammatory and tissue-remodelling pathways, including inflammatory response, KRAS activation, EMT and myogenesis. No Hallmark pathways were significantly enriched among DEGs detected only in males.

Together, our data establish that persistent HPV16 E7 expression profoundly alters the oral epithelial state, coupling cell-cycle activation with inflammatory and interferon-associated signalling. Biological sex modifies the breadth of this response, with females exhibiting more extensive transcriptional remodelling despite preservation of the core E7-associated programme across sexes.

WT and K14E7 mice reach similar pathological endpoints following 4NQO-induced carcinogenesis

To determine whether the preconditioned HPV16 E7-driven oral epithelial and immune remodelling influences the development of clinical pathology after carcinogen exposure, WT and K14E7 mice were treated with 4NQO in drinking water for 16 weeks followed by an 8-week observation period (Fig. 4a).

Despite their distinct epithelial phenotypes before 4NQO exposure, tumour-related disease events were observed in both WT and K14E7 mice following 4NQO treatment (Suppl. Table 1), with no significant differences in tumour-associated survival between genotypes or sexes (Fig. 4b). Body weight trajectories also remained comparable throughout the study, indicating similar tolerance of the treatment regimen (Suppl. Fig. S6).

Progressive lesions, ranging from focal epithelial thickening to papilloma-like exophytic growths, developed across all treated groups (Fig. 4c). Clinical disease severity did not differ significantly among genotype- and sex-stratified groups, with similar disease severity index (DSI) distributions across groups (Fig. 4e; Suppl. Table 2).

Histopathological analysis similarly demonstrated progression from epithelial dysplasia to invasive squamous cell carcinoma in both genotypes, with no genotype-associated difference in pathological grade distribution between WT and K14E7 mice (Fig. 4e, f). Although a single K14E7 female developed metastatic squamous cell carcinoma involving a regional lymph node, this represented an isolated event and was not observed elsewhere in the cohort (Suppl. Fig. S7).

Collectively, these findings demonstrate that, under the genotype-specific 4NQO exposure regimen used in this study, WT and K14E7 mice reached similar clinical and pathological disease endpoints. This

provided a common pathological context in which to examine how persistent HPV16 E7 expression alters the epithelial and immune programmes accompanying carcinogenesis.

HPV16 E7 establishes a hyperplastic, highly proliferative epithelium before carcinogen-induced remodelling

Prior to carcinogen exposure, K14E7 mice exhibited marked oral epithelial hyperplasia, with significantly greater epithelial thickness than WT controls (Fig. 5a, b). This architectural alteration was accompanied by increased epithelial proliferation, as evidenced by higher Ki-67 positivity in K14E7 controls (Fig. 5c, d).

At the experimental endpoint, 8 weeks after completion of the 16-week 4NQO exposure, epithelial thickness increased significantly in both genotypes, eliminating the pre-existing difference and resulting in comparable epithelial thickness between WT and K14E7 mice (Fig. 5a, b). Increased thickness in a subset of mice following 4NQO reflected papilloma-like exophytic lesions and was retained as a clinical feature of pathology.

Ki-67 positivity increased following 4NQO exposure in WT mice, whereas no significant additional increase was observed in K14E7 epithelium, resulting in comparable proliferative activity between genotypes following 4NQO exposure (Fig. 5c, d). No differences were observed between sexes.

Thus, carcinogenic challenge drove WT epithelium towards the proliferative phenotype already established by persistent E7 expression. No significant sex-associated differences in epithelial thickness or Ki-67 positivity were detected, indicating that the broader female transcriptional response observed earlier was not accompanied by a corresponding difference in these histological measures of epithelial remodelling.

HPV16 E7 preconditions the epithelial transcriptional response to carcinogen exposure

The global transcriptomic landscape differed primarily according to 4NQO treatment status, with genotype contributing a secondary source of variation (Fig. 6a). Consistent with this pattern, WT epithelium showed extensive transcriptional remodelling following 4NQO exposure, with 6,040 DEGs identified between 4NQO-treated and untreated control mice. In contrast, K14E7 epithelium showed substantially less additional transcriptional remodelling, with 757 DEGs identified between 4NQO-treated and control mice (Fig. 6b).

The transcriptional response of WT epithelium to 4NQO recapitulated several features of the HPV16 E7-associated epithelial programme present in untreated K14E7 mice. Representative proliferation and cell-cycle genes, including *Mki67*, *Top2a*, *Cdk1*, *Ccnb1*, *Birc5* and *Cdc20*, were strongly induced following 4NQO exposure (Fig. 6c). In contrast, control K14E7 epithelium already showed elevated expression of many of these genes relative to control WT epithelium, consistent with the hyperplastic proliferative state established by persistent HPV16 E7 expression. Interferon- and inflammatory-response genes were

likewise induced following 4NQO exposure in WT epithelium, whereas expression of these programmes changed less consistently in K14E7 mice after treatment (Fig. 6c).

This pattern was also evident at the pathway level. Hallmark pathway analysis demonstrated that the WT response to 4NQO largely recapitulated the epithelial transcriptional programme established by persistent HPV16 E7 expression alone, with positive enrichment of E2F targets, G2M checkpoint, MYC targets, interferon responses, TNF α signalling via NF- κ B, IL6-JAK-STAT3 signalling, glycolysis, hypoxia and EMT (Fig. 6d). These pathways were likewise positively enriched in control K14E7 epithelium relative to control WT, indicating that persistent HPV16 E7 establishes a primed epithelial transcriptional programme that is subsequently recapitulated during carcinogen-induced remodelling of the WT epithelium.

Consistent with the preconditioned transcriptional state established by persistent HPV16 E7 expression, K14E7 epithelium underwent comparatively little additional transcriptional remodelling following 4NQO exposure. Unlike WT tissue, which exhibited enrichment of proliferative and inflammatory programmes, no additional enrichment of these pathways was observed in K14E7 epithelium after 4NQO treatment. Instead, Hallmark analysis identified relative reduction of gene expression associated with interferon response, oxidative phosphorylation, PI3K-AKT-mTOR signalling, apoptosis and inflammatory response in 4NQO-treated compared with control K14E7 epithelium (Fig. 6d). Several pathways, including MYC targets V2, reactive oxygen species and KRAS signalling, showed positive enrichment trends following 4NQO exposure, although none remained significant after multiple-testing correction (Suppl. Fig. S8).

Despite converging on broadly comparable pathological disease severity, 4NQO-treated WT and K14E7 epithelium remained transcriptionally distinct, with 897 DEGs identified between genotypes (Fig. 6b). Gene Ontology analysis confirmed that these differences reflected distinct biological programmes. 4NQO-treated K14E7 epithelium preferentially enriched antiviral defence, interferon responses, innate immune signalling and DNA replication, whereas 4NQO-treated WT epithelium preferentially enriched epidermal development, keratinocyte differentiation, keratinisation, angiogenesis, vascular development and wound healing (Fig. 6e), consistent with an EMT-associated transcriptional programme identified by Hallmark GSEA (Fig. 6d).

Collectively, these findings indicate that persistent HPV16 E7 does not simply amplify the conventional epithelial transcriptional response to carcinogen exposure. Instead, it establishes an early reprogrammed proliferative and inflammatory transcriptional state, such that many of the programmes induced by 4NQO in WT epithelium are already active before exposure in K14E7 mice. Consequently, K14E7 epithelium undergoes comparatively limited additional transcriptional remodelling following carcinogen exposure. Nevertheless, antiviral, interferon-associated and innate immune programmes remained prominent in 4NQO-treated K14E7 epithelium, suggesting that immune-associated signalling remains a distinguishing feature of the HPV16 E7-expressing state, consistent with immune-related transcriptional features reported in human HPV-associated HNSCC⁶. Thus, although WT and K14E7

mice converge on similar pathological outcomes, the transcriptional programmes accompanying disease progression remain distinct.

Females exhibit broader transcriptional responses to carcinogen exposure

The broader E7-associated transcriptional response observed in untreated females prompted us to ask whether biological sex similarly influenced epithelial responses to carcinogenic challenge. Sex-stratified analyses revealed more extensive transcriptional remodelling following 4NQO exposure in females than males in both WT and K14E7 epithelium (Fig. 6f, g). Comparison of 4NQO-treated and control WT epithelium identified 3,840 DEGs in females compared with 1,440 in males, while the corresponding comparison in K14E7 epithelium identified 1,797 DEGs in females compared with 177 in males (Fig. 6f). In contrast, comparison of 4NQO-treated WT and K14E7 epithelium identified fewer genotype-associated DEGs in females than males (702 versus 1,459), indicating that greater female responsiveness was specific to the response to carcinogen exposure.

Despite these quantitative differences, the principal pathway-level responses to 4NQO were broadly conserved between sexes (Fig. 6g). Both female and male WT epithelium showed enrichment of proliferative, inflammatory and interferon response programmes following 4NQO exposure. K14E7 epithelium exhibited comparatively limited additional pathway remodelling in both sexes, characterised predominantly by a decrease of inflammatory, metabolic and stress-response programmes. Similarly, comparison of 4NQO-treated K14E7 and WT epithelium revealed broadly concordant pathway activity between females and males.

Collectively, these findings reveal a recurring pattern of greater transcriptional response in the female oral epithelium to 4NQO exposure in both WT and K14E7 mice, despite broadly shared pathway-level responses between sexes. Together with the stronger female HPV16 E7-associated transcriptional response observed at baseline, these findings identify enhanced female transcriptional responsiveness to both persistent viral oncogene expression and carcinogen exposure as a recurring feature of epithelial remodelling.

Murine HPV⁺ and HPV⁻ oral carcinogenesis recapitulates major transcriptional programmes of human HPV⁺ and HPV⁻ HNSCC

Genes differentially expressed between 4NQO-treated K14E7 and WT epithelium were mapped to their human orthologues and tested by pre-ranked GSEA against the TCGA HPV⁺ versus HPV⁻ HNSCC transcriptome. Genes upregulated in 4NQO-treated K14E7 epithelium were significantly enriched towards the HPV⁺ end of the human transcriptional ranking, whereas genes upregulated in 4NQO-treated WT epithelium showed reciprocal enrichment towards the HPV⁻ end (Fig. 7a). This bidirectional enrichment supports cross-species conservation of transcriptional differences associated with HPV16 E7 expression in 4NQO carcinogen-exposed murine epithelium and HPV status in human HNSCC.

To determine whether this gene-level correspondence extended to broader biological programmes, Hallmark pathway activity was compared between 4NQO-treated K14E7 and WT epithelium, and HPV⁺ and HPV⁻ human HNSCC. Pathway-level enrichment showed broad cross-species concordance, with the majority, 45 of 50 pathways (90%), of Hallmark pathways enriched in the same direction in the murine and human comparisons (Pearson $r = 0.67$, $p < 0.001$), (Fig. 7b; Supplementary Fig. S10). Proliferative programmes, including E2F targets and G2M checkpoint, were concordantly enriched in both murine K14E7 epithelium and human HPV⁺ HNSCC, whereas oxidative phosphorylation, xenobiotic metabolism, glycolysis and epithelial–mesenchymal transition were enriched in both murine WT epithelium and human HPV⁻ HNSCC. A small number of pathways showed opposing directions between species, including interferon-response signatures (Supplementary Data). Together, these findings demonstrate that the gene-level correspondence identified by direct enrichment analysis extends to broader HPV-associated transcriptional programmes.

To determine whether sex-associated transcriptional patterns observed in the murine model also extended to human HNSCC, Hallmark enrichment from female-versus-male comparisons was compared between 4NQO-treated murine epithelium and the overall human TCGA-HNSCC cohort. Sex-associated pathway activity showed more modest but significant cross-species concordance (Pearson $r = 0.32$, $p = 0.022$; Fig. 7c). Several inflammatory and immune-associated programmes were enriched towards females in both datasets. These findings suggest that aspects of sex-associated transcriptional variation observed during murine oral carcinogenesis are also evident in human HNSCC, although the overall correspondence was weaker than that associated with HPV status. An integrated analysis of HPV status and biological sex could not be performed as the HPV⁺ TCGA-HNSCC cohort was not statically powered to stratify by biological sex.

Collectively, these analyses demonstrate that the K14E7/4NQO and WT/4NQO models capture major transcriptional features associated with human HPV⁺ and HPV⁻ HNSCC at both gene and pathway levels. The reciprocal alignment of K14E7-associated gene signatures with human HPV status, together with broad conservation of pathway activity and modest conservation of sex-associated programmes in HNSCC, supports the translational relevance of these models.

Discussion

HPV-dependent and independent HNSCC arise through distinct aetiological pathways, yet how persistent viral oncogene expression shapes the epithelial and immune landscape before malignant transformation remains incompletely understood. Using the K14E7 model with and without 4NQO exposure, we show that HPV16 E7 established a distinct epithelial and immunological state before carcinogen challenge, characterised by epithelial hyperproliferation, broad transcriptional reprogramming and selective immune remodelling. Following 4NQO exposure, WT epithelium underwent extensive transcriptional change that substantially overlapped with programmes already active in K14E7 epithelium without 4NQO exposure, whereas K14E7 epithelium showed comparatively limited additional transcriptional change

after exposure. Across genotypes, females exhibited a broader transcriptional response to 4NQO exposure than males, extending the sex-associated differences already evident in the response to persistent E7 expression. These findings support a model in which persistent HPV16 E7 preconditions the epithelial and immune landscape and alters the subsequent response to carcinogenic injury.

HPV16 E7-mediated disruption of the retinoblastoma/E2F axis is a well-established driver of epithelial proliferation^{25,26}. Consistent with this mechanism, and previous K14E7 studies¹⁷, the K14E7 oral epithelium displays increased thickness and proliferation at baseline, together with enrichment of E2F targets, G2M checkpoint and MYC-associated programmes, paralleling molecular features of HPV-associated disease^{27,28}. Beyond cell-cycle deregulation, HPV16 E7 was associated with enrichment of inflammatory and immune-response programmes, including TNF/NF- κ B, IL6–JAK–STAT3 and interferon-response signatures. Thus, persistent HPV16 E7 expression alone establishes a proliferative and inflammatory epithelial state.

This epithelial state was accompanied by selective remodelling, characterised by enrichment of CTLA4-associated T cell phenotypes, depletion of MHCII^{high} EpCAM⁺ myeloid APCs and expansion of macrophage-like APCs. Altered Langerhans cell (LC) recruitment and function have been described in the cutaneous K14E7 epidermis and human cervical HPV-associated lesions^{15,29–31}, while HPV⁺ HNSCC shows reduced LC abundance and increased macrophage representation^{32,33}. Although these human observations derive from established premalignant high-grade lesions or SCC, their directional similarity suggests that HPV-associated myeloid remodelling may arise much earlier in disease evolution. Macrophage-like APCs at the intraepithelial hyperplastic stage showed increased CD64 and CD86 without increased PD-L1 expression (Suppl. Fig. S5), distinguishing them from an established immunosuppressive TAM phenotype. Similar mononuclear phagocyte remodelling occurs during early 4NQO carcinogenesis in WT mice³⁴, suggesting that HPV16 E7 may establish aspects of a myeloid response otherwise associated with an early epithelial hyperproliferation.

T cell remodelling was similarly selective. Overall TCR β ⁺ proportions and the distribution of antigen-inexperienced and antigen-experienced cells were largely preserved, while CTLA4-associated phenotypes were widely expanded in the K14E7 oral epithelium at baseline. Antigen-experienced T cells showed increased expression levels of CTLA4, without corresponding upregulation of PD1 or confidently detectable FoxP3. CTLA4 is strongly associated with regulatory T cell function and is among the inhibitory receptors expressed by exhausted T cells, CTLA4 is not specific to either state or can also be induced following conventional T cell activation^{35,36}. These findings therefore identify checkpoint-associated remodelling as an early feature of persistent HPV16 E7 expression. Reduced TCR $\gamma\delta$ ⁺ proportions and increased CTLA4 expression within TCR $\gamma\delta$ ⁺ cells extended this phenotype to an epithelial-surveillance pattern³⁷. Concurrent increases in APC-expressed CD86 and T cell-expressed CTLA4 suggest a potential interface between myeloid and lymphoid remodelling, although functional interaction cannot be inferred.

Biological sex was associated with variation in selected immune phenotypes, although these analyses were exploratory given the limited sex-stratified sample size. Female K14E7 mice showed more pronounced differences in selected EpCAM⁺ myeloid and conventional T cell populations, whereas increased CTLA4 expression within TCR γ δ ⁺ cells were more evident in males, indicating that sex-associated variation was diverse. This is consistent with broader evidence that sex influences antiviral and antitumour immunity^{38,39}, including experimental HNSCC models showing enhanced effector T cell activation and reduced immunosuppressive myeloid populations in females¹². These observations may be relevant to the marked male predominance of oral HPV infection and HPV-associated OPSCC^{40,41}. Nevertheless, the immune differences observed here did not translate into detectable differences in disease severity or survival.

Many proliferative and inflammatory programmes induced by 4NQO in WT tissue were already enriched in untreated K14E7 epithelium, indicating that persistent HPV16 E7 preconditions the transcriptional context in which carcinogenic injury occurs. The comparatively limited additional remodelling in K14E7 epithelium may therefore reflect prior establishment of these programmes rather than an attenuated response to 4NQO. This response was also modified by biological sex, with females showing more extensive transcriptional differences than males in both WT and K14E7 epithelium, alongside sex-associated differences in immune-related programmes. These findings suggest that the transcriptional context established by HPV16 E7 is further shaped by host sex. Importantly, 4NQO-treated WT and K14E7 epithelium remained transcriptionally distinct despite reaching similar pathological endpoints, demonstrating that similar disease phenotypes can arise through different molecular programmes. Previous studies have described progressive molecular and immune alterations during 4NQO carcinogenesis¹⁸ and enhanced carcinogenesis in the presence of HPV oncogenes⁴². Our findings extend these observations by showing that persistent HPV16 E7 alters the transcriptional context in which carcinogenic injury occurs, supporting a model in which HPV16 E7 redirects, rather than simply accelerates, epithelial remodelling during carcinogenesis.

Cross-species gene-level enrichment and broad pathway concordance with human HPV⁺ HNSCC support the translational relevance of the murine K14E7/4NQO and WT/4NQO models. This is consistent with previous genomic studies demonstrating extensive overlap in mutational signatures between 4NQO-induced murine OSCC and human disease, including recurrent alterations in key HNSCC-associated genes¹⁸. Together, these findings indicate that the 4NQO model recapitulates important molecular features of human HNSCC while enabling interrogation of epithelial responses during carcinogenesis.

Limitations include small sex-stratified subgroups, particularly the low number of HPV⁺ female TCGA cases (n = 5) that did not allow a comparison between sexes in human HPV⁺ HNSCC, and the inability of bulk RNA sequencing to resolve cell-specific transcriptional changes. Nevertheless, integration of epithelial, immune and transcriptional phenotypes supports HPV16 E7-associated preconditioning before overt malignancy.

Collectively, these findings show that persistent HPV16 E7 expression establishes a proliferative and immune-remodelled epithelial state before carcinogen exposure that alters the subsequent response to carcinogenic injury. This pre-existing state alters the subsequent response to carcinogenic injury, such that WT and K14E7 epithelium develop similar pathological endpoints through distinct transcriptional programmes. Biological sex further influences immune phenotypes in the normal oral epithelium, of the HPV16 E7-associated response, and demonstrated an overall increased transcriptional response to both HPV16 E7 and 4NQO in females. Conservation of major HPV16 E7-associated transcriptional programmes in human HPV⁺ HNSCC supports the translational relevance of the model. Together, our findings suggest that HPV16 E7 reshapes the tissue context in which carcinogenesis occurs, rather than simply accelerating a common pathway of disease progression.

Material and Methods

Animal studies

Male and female K14E7 transgenic mice, expressing HPV16 E7 under the keratin 14 promoter²⁰, and wild-type (WT; K14E7-negative) littermate controls were generated by breeding K14E7-positive male mice with C57BL/6 females (The Jackson Laboratory, strain #000664). Experiments were approved by the University of Queensland Animal Ethics Committee (Project No. 2023/AE000117), and mice were maintained under specific pathogen-free conditions at the Translational Research Institute Biological Research Facility (Brisbane, Australia).

4NQO-induced oral carcinogenesis

Oral carcinogenesis was induced using 4-nitroquinoline-1-oxide (4NQO; Sigma-Aldrich, N8141-5G) in drinking water. From 5 weeks of age, WT and K14E7 mice received 100µg/mL and 50µg/mL 4NQO, respectively, for 16 weeks, followed by untreated water for 8 weeks. Fresh 4NQO solution was prepared twice weekly and provided in light-protected bottles. The 4NQO concentration used in K14E7 mice was empirically optimised during model establishment. An initial concentration of 10µg/mL was selected based on previous studies demonstrating increased susceptibility of HPV16 oncogene-expressing mice to 4NQO-induced oral carcinogenesis at reduced carcinogen exposure⁴³. As this concentration resulted in limited lesion development in our K14E7 model, it was subsequently increased to 50µg/mL to achieve consistent disease progression.

Animal monitoring and survival analysis

Mice were monitored throughout the experiment according to predefined welfare criteria, and body weight was recorded. Survival time was defined from 4NQO initiation to euthanasia at predefined disease-related humane endpoints (a cumulative welfare score > 11 or a severe score in any individual welfare criterion). Animals euthanised for non-tumour-related welfare concerns were censored, including K14E7 mice euthanised due to strain-associated thymic hypertrophy⁴⁴.

Clinical assessment of tongue lesions

Following euthanasia and dissection at the experimental endpoint, tongues were photographed and macroscopic disease severity was assessed using a modified clinical scoring system adapted from Wagner et al. (2021) based on the number, size and morphology of visible tongue lesions (Table 1) ⁴⁵. Scores were grouped into Disease Severity Index (DSI) categories as described in Supplementary Table 2.

Table 1

Refined clinical scoring system for macroscopic tongue lesions Adapted from 45

Score	Surface Characteristic	Description
0	Normal surface	Smooth surface with no visible abnormalities; intact papillae and central groove.
1	Small papule	Small raised whitish papules (< 2mm) on the tongue surface.
2	Thick papule	Larger, more prominent papules with increased surface thickness (> 2mm).
3	Nodular	Multiple nodules or bumps; uneven surface architecture.
4	Single papilloma like exophytic lesion	A single raised, wart-like growth with a rough, exophytic surface (papilloma-like lesion) observed on the tongue.
5	Multiple papilloma like exophytic lesions	Multiple raised, wart-like exophytic growths (papilloma-like lesions).
6	Severe lesion with combined features	Combined features including papilloma-like growths, thickened, nodular surface or ulcerated/proliferative growths (> 2mm), indicative of advanced disease.

Tissue collection, histopathology and epithelial proliferation

Following clinical assessment, tongues were bisected longitudinally. One half was fixed overnight in 4% paraformaldehyde, paraffin-embedded and sectioned at 4µm for histological and immunofluorescence analyses. Histopathology was assessed independently by an experienced pathologist blinded to mouse identity and experimental group.

Epithelial thickness was quantified on haematoxylin and eosin (H&E) stained sections using QuPath (v0.5.1) at 500µm intervals along the dorsal epithelium, excluding the superficial keratin layer. Measurements from each animal were averaged for statistical analysis.

Cellular proliferation was assessed by Ki-67 immunofluorescence using the CONFIRM anti-Ki-67 (30 – 9; Roche Diagnostics), a FITC-conjugated anti-rabbit secondary antibody and DAPI counterstain on a Ventana Discovery Ultra platform. Sections were scanned using an Olympus VS200, and proliferation index was calculated as the Ki-67-positive nuclei as a percentage of DAPI-positive nuclei.

Single-cell preparation from murine tongue epithelium

Tongue epithelium was separated from underlying connective tissue following subepithelial injection of Dispace II (2.5 mg/mL; Roche Diagnostics) and incubation at 37°C for 30 min. Isolated epithelium was cut into small pieces dissociated in PBS containing Liberase TL (2.5 mg/mL; Roche Diagnostics) and DNase I (0.2 mg/mL; Thermo Fisher Scientific) at 37°C for 1 h with intermittent mixing. Cell suspensions were filtered through a 100µm strainer and used for flow cytometric analyses.

Flow cytometry

A 29-colour spectral flow cytometry panel was used to characterise lymphoid and myeloid populations (Suppl. Table 3). Single-cell suspensions were incubated with anti-mouse CD16/CD32 Fc receptor block (1:100; Cat. #553141, BD Biosciences) and Zombie NIR Fixable Viability Dye (1:1000; Cat. #423106, BioLegend) for 10 min on ice, followed by surface antibody staining for 20 min at 4°C. Cells were fixed and permeabilised using the FoxP3/Transcription Factor Staining Buffer Set (eBioscience) and stained intracellularly for FoxP3 and CTLA-4. Data were acquired on a Cytex Aurora spectral flow cytometer and analysed using FlowJo (v10.10).

Dimensionality reduction and population annotation

Live singlet CD45⁺ events were exported from each sample and uniformly down-sampled to 9,000 events per sample. Down-sampled datasets were concatenated and analysed using t-distributed stochastic neighbour embedding (t-SNE) for visualisation of immune-cell phenotypes. Immune populations were defined using a predefined manual hierarchical gating strategy and mapped onto the t-SNE embedding, with population identities supported by marker-expression patterns. For quantitative analysis, manually gated populations were back-gated to the corresponding individual samples within the down-sampled dataset. Population frequencies were calculated for each biological replicate relative to the total CD45⁺ compartment unless otherwise stated.

Immune cell annotations

Population definitions and representative gating strategies are shown in Supplementary Figs. S1 and S3. Lymphoid populations comprised TCRβ⁺ T cells, TCRγδ⁺ T cells, NK1.1⁺TCRβ⁻ natural killer (NK) cells, and NK1.1⁺TCRβ⁺ NKT-like cells. TCRβ⁺ cells were classified as antigen-inexperienced (CD5⁺CD44⁻) or antigen-experienced (CD5⁻CD44⁺). Antigen-inexperienced TCRβ⁺ cells were further resolved according to CTLA4 and PD1 expression into CTLA4⁺ (PD1[±]), CTLA4-PD1⁺ and CTLA4-PD1⁻ subsets, with CTLA4⁺ cells analysed as a single population irrespective of PD1 expression. Antigen-experienced TCRβ⁺ cells were subsequently separated according to NK1.1 expression into NK1.1⁻ conventional T cells and NK1.1⁺ NKT-like cells. Following exclusion of NKT-like cells, checkpoint phenotyping of the antigen-experienced NK1.1⁻ TCRβ⁺ compartment focused on the CTLA4⁺PD1⁻ subset. Other CTLA4/PD1-defined subsets within this compartment were not quantified separately because of low event numbers. FoxP3 expression was not detected with sufficient confidence; therefore, CTLA4⁺ T cells were not annotated as regulatory T cells.

Myeloid populations included Ly6G⁺Ly6C⁺ neutrophils and cells within the TCR⁻NK1.1⁻Ly6G⁻ compartment. EpCAM expression distinguished EpCAM⁺ and EpCAM⁻ myeloid populations. Within the EpCAM⁺ compartment, CD24⁺ cells were further resolved according to MHCII and XCR1 expression. MHCII^{high}XCR1^{low} and MHCII^{high}XCR1^{high} populations were conservatively designated as epithelial APCs, whereas the MHCII^{low/negative}XCR1^{low} population was designated an epithelial myeloid cell population. Although EpCAM and CD24 are commonly expressed by Langerhans cells, because CD207/Langerin was not included in the flow cytometry panel, EpCAM⁺CD24⁺ populations were conservatively designated as epithelial APCs or epithelial myeloid cells according to MHCII expression, rather than as definitive Langerhans cells. XCR1 expression within these populations is consistent with previous reports identifying XCR1-expressing cells within the oral mucosal Langerhans cell compartment⁴⁶.

Within the EpCAM⁻ myeloid compartment, CD64⁺MHCII⁺CD11c⁺ cells displaying F4/80 and SIRPα expression were conservatively annotated as macrophage-like APCs and subdivided into CCR2^{low/-}XCR1^{low/-} and CCR2⁺XCR1^{high} populations. A separate CD64⁻MHCII⁻ population contained a discrete CCR2⁺XCR1⁺ subset, which was annotated as monocytes, consistent with the established association of CCR2 with monocyte recruitment. No CD317⁺ pDCs or CD19⁺ B cells were identified.

RNA extraction

Total RNA was extracted from isolated tongue epithelium using the RNeasy Mini Kit (Qiagen) Epithelial tissue was lysed in RLT buffer supplemented with β-mercaptoethanol, homogenised using a Precellys® Evolution tissue homogeniser (Bertin Technologies) with 0.1 mm zirconia/silica beads (BioSpec Products), and purified using RNeasy spin columns. RNA concentration and integrity were assessed using NanoDrop spectrophotometer (Thermo Fisher Scientific), and Agilent TapeStation RNA ScreenTape assay (Agilent Technologies) respectively. Samples with an RNA integrity number (RIN) ≥ 8.5 were used for library preparation.

Whole-transcriptome RNA sequencing

Bulk RNA sequencing was performed on oral epithelial tissue from untreated and 4NQO-treated WT and K14E7 mice, including both sexes (n = 4 biological replicates per genotype, treatment, and sex). Whole-transcriptome libraries were prepared using the QIAseq UPXome RNA Library Kit HMR (Qiagen, Cat. No. 334705) and the QIAseq UX 12 Index Kit IL UDI (Qiagen, Cat. No. 331801) according to the manufacturer's protocol. Briefly, 100 ng of total RNA was fragmented and depleted of ribosomal RNA using Qiagen FastSelect chemistry, followed by cDNA synthesis using a combination of oligo(dT) and random primers to capture both polyadenylated and non-polyadenylated transcripts. Samples were individually barcoded, pooled into four libraries containing eight samples each, amplified by PCR (15 cycles), and purified. Library quality and concentration were assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA), and libraries were pooled at equimolar concentrations for sequencing.

Paired-end sequencing (2 × 151 bp) was performed on an Illumina NovaSeq X platform using a NovaSeq X Series 10B Reagent Kit (300 cycles; Illumina, San Diego, CA, USA). Library preparation was performed at the University of Queensland Sequencing Facility (Brisbane, Australia), and sequencing was carried out in collaboration with the Australian Genome Research Facility.

RNA-seq analysis

Raw sequencing reads were demultiplexed and the 16-bp sample barcode at the 5' end of Read 2 was removed using Cutadapt (v4.2), with Read 1 retained unchanged. Pre-processed paired-end reads were analysed using nf-core/rnaseq (v3.15.0), including rRNA removal with SortMeRNA and alignment and gene-level quantification against the *Mus musculus* GRCm39 reference genome using STAR/RSEM.

Differential expression was performed in R (v4.4.2) using DESeq2 (v1.46.0) with median-of-ratios normalisation. Principal component analysis (PCA) was performed on variance-stabilised counts. Differentially expressed genes (DEGs) were identified using Wald tests with Benjamini–Hochberg adjustment and defined as adjusted $P < 0.05$ and $|\log_2FC| \geq 1$. Analyses were performed using computational resources provided by The University of Queensland Research Computing Centre's Bunya supercomputer⁴⁷.

Gene Set Enrichment Analysis (GSEA)

Hallmark Gene Set Enrichment Analysis (GSEA) was performed using clusterProfiler (v4.14.4) with mouse Hallmark gene sets obtained from the Molecular Signatures Database (MSigDB) using msigdb (v7.5.1). Genes were ranked by the DESeq2 Wald statistic, and pathways with FDR < 0.05 were considered significantly enriched unless otherwise stated. NES direction was interpreted according to the order of groups in each comparison.

Gene Ontology (GO) enrichment analysis

GO Biological Process over-representation analysis was performed separately on DEGs with positive and negative $\log_2FC$ between 4NQO-treated K14E7 and WT epithelium using clusterProfiler (v4.14.4) enrichGO and org.Mm.eg.db. Ensembl identifiers were mapped to Entrez identifiers, and terms with adjusted $P < 0.05$ were considered significantly enriched. The 10 terms with the lowest adjusted P values in each direction were visualised.

Human TCGA-HNSC transcriptomic analysis

Raw gene-level RNA-SEQ counts from primary tumours in the Cancer Genome Atlas Head and Neck Squamous Cell Carcinoma (TCGA-HNSC) cohort were obtained from the National Cancer Institute Genomic Data Commons (GDC) using TCGAbiolinks (v2.34.0)⁴⁸. HPV subtype annotations were obtained from the TCGA-HNSC PanCancer Atlas dataset through cBioPortal and matched using TCGA patient identifiers. Tumours annotated as HNSC_HPV + or HNSC_HPV– were retained, yielding 72 HPV+

and 415 HPV⁺ tumours. These annotations reflect transcriptionally active HPV detected by viral RNA sequencing rather than p16 immunohistochemistry^{49,50}.

Raw count data were imported as a SummarizedExperiment object (v1.36.0), and the unstranded gene count assay was used for analysis. Genes were retained if they had at least 10 counts in at least 10 samples. Differential expression was assessed using DESeq2 Wald tests, with adjusted $P < 0.05$ and $|\log_2FC| > 1$ defining DEGs. Genes were annotated using org.Hs.eg.db for downstream comparative analyses.

Statistical analysis

Statistical analyses were performed using GraphPad Prism (v10.4.1) and R (v4.4.2). Data are presented as mean $\pm$ SD unless otherwise stated, with individual data points representing biologically independent animals. Exact sample sizes are provided in the corresponding figure legends. Pooled comparisons between genotypes were performed using Mann–Whitney U tests. Where genotype and biological sex were examined simultaneously, two-way ANOVA was performed with genotype and sex as factors, followed by Tukey's multiple-comparisons test. Kaplan–Meier survival curves were compared using the log-rank (Mantel–Cox) test. All statistical tests were two-sided where applicable. Exact P values are reported where appropriate, and $P < 0.05$ was considered statistically significant.

Data availability statement

The transcriptomics dataset generated in this study is available through UQ eSpace (<https://doi.org/10.48610/8e063d0>). TCGA-HNSC RNA-seq data were obtained from the Genomic Data Commons, with HPV subtype annotations obtained from the corresponding TCGA-HNSC PanCancer Atlas dataset available through cBioPortal (https://www.cbioportal.org/study/clinicalData?id=hnsk_tcga_pan_can_atlas_2018). No custom software or algorithms were developed for this study. Analyses were performed using publicly available R/Bioconductor packages as described in the methods.

Declarations

Conflict of interest

The authors declare no conflict of interest.

Supplementary information

Supplementary figures and table

Supplementary data

Funding statement

K.H. was supported by The University of Queensland Graduate School Scholarship. J.C. and A.H.H were supported by a Collaborative Project Grant by the Zelman Cowan Academic Initiative. J.C. was supported by a Mid-Career Fellowship from the Garnett Passe Rodney Williams Memorial Foundation. I.F. was supported by the National Health and Medical Research Council (NHMRC) Australia (APP1173927). D.S. is supported by the Garnett Passe Rodney Williams Memorial Foundation.

Acknowledgements

This research was carried out at the Translational Research Institute, Woolloongabba, QLD 4102, Australia. The Translational Research Institute is supported by a grant from the Australian Government. This research was undertaken with the assistance of resources and services from the University of Queensland Research Computing Centre (RCC), including the Bunya high-performance computing cluster: <https://rcc.uq.edu.au/about/acknowledging-rcc>

References

1. Barsouk, A., Aluru, J. S., Rawla, P., Saginala, K. & Barsouk, A. Epidemiology, Risk Factors, and Prevention of Head and Neck Squamous Cell Carcinoma. *Med Sci (Basel)* **11**, 42 (2023). <https://doi.org/10.3390/medsci11020042>
2. Canning, M. *et al.* Heterogeneity of the Head and Neck Squamous Cell Carcinoma Immune Landscape and Its Impact on Immunotherapy. *Frontiers in cell and developmental biology* **7**, 52 (2019). <https://doi.org/10.3389/fcell.2019.00052>
3. de Oliveira, M. C. *et al.* CTLA4 genetic variants influence immune regulation and susceptibility of HPV associated cervical cancer. *Discover. Oncology* **17**, 88 – 16 (2025). <https://doi.org/10.1007/s12672-025-04179-3>
4. Li, Y., Deng, J., Liu, Y. & Yu, S. HPV infection and the immune microenvironment in cervical cancer. *Frontiers in immunology* **16**, 1645019 (2025). <https://doi.org/10.3389/fimmu.2025.1645019>
5. Qin, T., Li, S., Henry, L. E., Liu, S. & Sartor, M. A. Molecular Tumor Subtypes of HPV-Positive Head and Neck Cancers: Biological Characteristics and Implications for Clinical Outcomes. *Cancers (Basel)* **13**, 2721 (2021). <https://doi.org/10.3390/cancers13112721>
6. Mehdi, A. M. *et al.* HPV status represents dominant trait driving delineation of survival-associated gene co-expression networks in head and neck cancer. *Cancer Gene Ther* **30**, 629–640 (2023). <https://doi.org/10.1038/s41417-022-00577-9>
7. Thakral, A. *et al.* Smoking and alcohol by HPV status in head and neck cancer: a Mendelian randomization study. *Nature communications* **15**, 7835–7839 (2024). <https://doi.org/10.1038/s41467-024-51679-x>

8. 8. Luo, Z., Huang, Y., Ye, R. & Yin, M. Global Burden and Gender Disparities in Head and Neck Cancers Among Adults Aged 40–64, 1990–2021: A Systematic Analysis From the Global Burden of Disease Study 2021. *Cancer reports* **8**, e70287-n/a (2025). <https://doi.org/10.1002/cnr2.70287>
9. 9. Du, M. et al. Clinicopathologic characteristics of HPV-associated head and neck squamous cell carcinoma in Southern China: long-term retrospective study of 400 cases. *Ther Adv Med Oncol* **16**, 17588359241242962–17588359241242962 (2024). <https://doi.org/10.1177/17588359241242962>
10. 10. Mundi, N. et al. Sex disparities in head & neck cancer driver genes: An analysis of the TCGA dataset. *Oral Oncol* **104**, 104614–104614 (2020). <https://doi.org/10.1016/j.oraloncology.2020.104614>
11. 11. Lortet-Tieulent, J. et al. Convergence of decreasing male and increasing female incidence rates in major tobacco-related cancers in Europe in 1988–2010. *European journal of cancer (1990)* **51**, 1144–1163 (2015). <https://doi.org/10.1016/j.ejca.2013.10.014>
12. 12. Jagadeesha, S. et al. Sex-based immune differences influence tumor progression in head and neck squamous cell carcinoma. *Oral oncology* **176**, 107929–107929 (2026). <https://doi.org/10.1016/j.oraloncology.2026.107929>
13. 13. Chandra, J., Miao, Y., Romoff, N., Frazer, I. H. & Akiyama, T. Epithelium expressing the E7 oncoprotein of HPV16 attracts immune-modulatory dendritic cells to the skin and suppresses their antigen-processing capacity. *PLoS One* **11**, e0152886-e0152886 (2016). <https://doi.org/10.1371/journal.pone.0152886>
14. 14. Bashaw, A. A. et al. HPV16 E7-Driven Epithelial Hyperplasia Promotes Impaired Antigen Presentation and Regulatory T-Cell Development. *Journal of investigative dermatology* **139**, 2467–2476.e2463 (2019). <https://doi.org/10.1016/j.jid.2019.03.1162>
15. 15. Tuong, Z. K. et al. A model of impaired Langerhans cell maturation associated with HPV induced epithelial hyperplasia. *iScience* **24**, 103326 (2021). <https://doi.org/10.1016/j.isci.2021.103326>
16. 16. Dang, T. V. T. et al. The Involvement of IL-34 in Inflammatory and Malignant Skin Diseases. *Journal of investigative dermatology* **145**, 2688–2700 (2025). <https://doi.org/10.1016/j.jid.2025.05.003>
17. 17. Tuong, Z. K. et al. Murine HPV16 E7-expressing transgenic skin effectively emulates the cellular and molecular features of human high-grade squamous intraepithelial lesions. *Papillomavirus Res* **5**, 6–20 (2018). <https://doi.org/10.1016/j.pvr.2017.10.001>
18. 18. Sequeira, I. et al. Genomic landscape and clonal architecture of mouse oral squamous cell carcinomas dictate tumour ecology. *Nat Commun* **11**, 5671–5671 (2020). <https://doi.org/10.1038/s41467-020-19401-9>
19. 19. Ge, S., Zhou, H., Zhou, Z., Liu, L. & Lou, J. Serum metabolite profiling of a 4-Nitroquinoline-1-oxide-induced experimental oral carcinogenesis model using gas chromatography-mass spectrometry. *PeerJ (San Francisco, CA)* **9**, e10619 (2021). <https://doi.org/10.7717/peerj.10619>
20. 20. Herber, R., Liem, A., Pitot, H. & Lambert, P. F. Squamous epithelial hyperplasia and carcinoma in mice transgenic for the human papillomavirus type 16 E7 oncogene. *J Virol* **70**, 1873–1881 (1996).

<https://doi.org/10.1128/jvi.70.3.1873-1881.1996>

21. 21. Mattarollo, S. R., Yong, M., Tan, L., Frazer, I. H. & Leggatt, G. R. Secretion of IFN- γ but Not IL-17 by CD1d-Restricted NKT Cells Enhances Rejection of Skin Grafts Expressing Epithelial Cell-Derived Antigen. *The Journal of immunology* (1950) **184**, 5663–5669 (2010).
<https://doi.org/10.4049/jimmunol.0903730>
22. 22. Mittal, D. et al. Indoleamine 2,3-Dioxygenase Activity Contributes to Local Immune Suppression in the Skin Expressing Human Papillomavirus Oncoprotein E7. *Journal of investigative dermatology* **133**, 2686–2694 (2013). <https://doi.org/10.1038/jid.2013.222>
23. 23. Bashaw, A. A. et al. Regulatory T Cells but Not IL-10 Impair Cell-Mediated Immunity in Human Papillomavirus E7 + Hyperplastic Epithelium. *Journal of investigative dermatology* **141**, 1264–1273.e1263 (2021). <https://doi.org/10.1016/j.jid.2020.10.011>
24. 24. Zhou, C., Tuong, Z. K., Lukowski, S. W., Chandra, J. & Frazer, I. H. Antigen Nonspecific Induction of Distinct Regulatory T Cell States in Oncogene-Driven Hyperproliferative Skin. *ImmunoHorizons* **5**, 102–116 (2021). <https://doi.org/10.4049/immunohorizons.2100006>
25. 25. McLaughlin-Drubin, M. E., Huh, K.-W. & Münger, K. Human Papillomavirus Type 16 E7 Oncoprotein Associates with E2F6. *Journal of Virology* **82**, 8695–8705 (2008).
<https://doi.org/10.1128/JVI.00579-08>
26. 26. McLaughlin-Drubin, M. E. & Munger, K. The human papillomavirus E7 oncoprotein. *Virology (New York, N.Y.)* **384**, 335–344 (2009). <https://doi.org/10.1016/j.virol.2008.10.006>
27. 27. Li, Y., Huang, Y., Li, B. & Yang, K. Roles of E2F family members in the diagnosis and prognosis of head and neck squamous cell carcinoma. *BMC medical genomics* **16**, 38 – 14 (2023).
<https://doi.org/10.1186/s12920-023-01470-6>
28. 28. Hazazi, A. et al. Signaling pathways in HPV-induced cervical cancer: Exploring the therapeutic promise of RNA modulation. *Pathology, research and practice* **263**, 155612 (2024).
<https://doi.org/10.1016/j.prp.2024.155612>
29. 29. Chandra, J., Miao, Y., Romoff, N. & Frazer, I. H. Epithelium Expressing the E7 Oncoprotein of HPV16 Attracts Immune-Modulatory Dendritic Cells to the Skin and Suppresses Their Antigen-Processing Capacity. *PLoS One* **11**, e0152886 (2016). <https://doi.org/10.1371/journal.pone.0152886>
30. 30. Dang, T. V. T. et al. Squamous cell carcinoma is associated with reduced IL34 expression, alterations in the Langerhans cell antigen-processing-presentation machinery and poor patient survival. *Clinical & translational immunology* **13**, e70018-n/a (2024).
<https://doi.org/10.1002/cti2.70018>
31. 31. Pavilion, G. et al. Spatial analysis of HPV-associated cervical intraepithelial neoplastic tissues demonstrate distinct immune signatures associated with cervical cancer progression. *The Journal of pathology* **268**, 200–214 (2026). <https://doi.org/10.1002/path.70002>
32. 32. Kindt, N. et al. Langerhans cell number is a strong and independent prognostic factor for head and neck squamous cell carcinomas. *Oral oncology* **62**, 1–10 (2016).
<https://doi.org/10.1016/j.oraloncology.2016.08.016>

33. 33. Seminerio, I. et al. High infiltration of CD68 + macrophages is associated with poor prognoses of head and neck squamous cell carcinoma patients and is influenced by human papillomavirus. *Oncotarget* **9**, 11046–11059 (2018). <https://doi.org/10.18632/oncotarget.24306>
34. 34. Saba, Y. et al. Early antitumor activity of oral Langerhans cells is compromised by a carcinogen. *Proceedings of the National Academy of Sciences - PNAS* **119**, 1–12 (2022). <https://doi.org/10.1073/pnas.2118424119>
35. 35. Trujillo-Ochoa, J. L., Kazemian, M. & Afzali, B. The role of transcription factors in shaping regulatory T cell identity. *Nature reviews. Immunology* **23**, 842–856 (2023). <https://doi.org/10.1038/s41577-023-00893-7>
36. 36. Negrini, S. et al. The CTLA-4 Pathway in Human Disease: Molecular Mechanisms and Therapeutic Targeting. *Genes* **17**, 574 (2026). <https://doi.org/10.3390/genes17050574>
37. 37. Dogan, S., Terzioglu, E. & Ucar, S. Innate immune response against HPV: Possible crosstalking with endocervical $\gamma\delta$ T cells. *Journal of reproductive immunology* **148**, 103435–103435 (2021). <https://doi.org/10.1016/j.jri.2021.103435>
38. 38. Windon, M. J. et al. Sex differences in HPV immunity among adults without cancer. *Human vaccines & immunotherapeutics* **15**, 1935–1941 (2019). <https://doi.org/10.1080/21645515.2019.1568157>
39. 39. Jacobsen, H. & Klein, S. L. Sex Differences in Immunity to Viral Infections. *Frontiers in immunology* **12**, 720952 (2021). <https://doi.org/10.3389/fimmu.2021.720952>
40. 40. Chaturvedi, A. K. et al. NHANES 2009–2012 Findings: Association of Sexual Behaviors with Higher Prevalence of Oral Oncogenic Human Papillomavirus Infections in U.S. Men. *Cancer research (Chicago, Ill.)* **75**, 2468–2477 (2015). <https://doi.org/10.1158/0008-5472.CAN-14-2843>
41. 41. D'Souza, G. et al. Sex Differences in Risk Factors and Natural History of Oral Human Papillomavirus Infection. *The Journal of infectious diseases* **213**, 1893–1896 (2016). <https://doi.org/10.1093/infdis/jiw063>
42. 42. Jabbar, S., Strati, K., Shin, M. K., Pitot, H. C. & Lambert, P. F. Human papillomavirus type 16 E6 and E7 oncoproteins act synergistically to cause head and neck cancer in mice. *Virology* **407**, 60–67 (2010). <https://doi.org/10.1016/j.virol.2010.08.003>
43. 43. Strati, K., Pitot, H. C. & Lambert, P. F. Identification of Biomarkers That Distinguish Human Papillomavirus (HPV)-Positive versus HPV-Negative Head and Neck Cancers in a Mouse Model. *Proceedings of the National Academy of Sciences - PNAS* **103**, 14152–14157 (2006). <https://doi.org/10.1073/pnas.0606698103>
44. 44. Malcolm, K. M. et al. Expression of the HPV16E7 Oncoprotein by Thymic Epithelium is Accompanied by Disrupted T Cell Maturation and a Failure of the Thymus to Involute with Age. *Clin Dev Immunol* **10**, 91–103 (2003). <https://doi.org/10.1080/10446670310001626562>
45. 45. Wagner, V. P. et al. Can propranolol act as a chemopreventive agent during oral carcinogenesis? An experimental animal study. *Eur J Cancer Prev* **30**, 315–321 (2021). <https://doi.org/10.1097/CEJ.0000000000000626>

46. 46. Kusumoto, Y. *et al.* Sublingual immune cell clusters and dendritic cell distribution in the oral cavity. *JCI insight* **9** (2024). <https://doi.org/10.1172/jci.insight.167373>
47. 47. Centre, T. U. o. Q. R. C. (2024).
48. 48. Colaprico, A. *et al.* TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic acids research* **44**, e71-e71 (2016). <https://doi.org/10.1093/nar/gkv1507>
49. 49. Shankar, K. & Walker, S. E. Analysis of divergent gene expression between HPV + and HPV- head and neck squamous cell carcinoma patients. *Infectious agents and cancer* **20**, 31–36 (2025). <https://doi.org/10.1186/s13027-025-00663-1>
50. 50. Ourailidis, I. *et al.* Multi-omics analysis to uncover the molecular basis of tumor budding in head and neck squamous cell carcinoma. *NPJ precision oncology* **9**, 73 – 18 (2025). <https://doi.org/10.1038/s41698-025-00856-2>

Figures

Figure 1

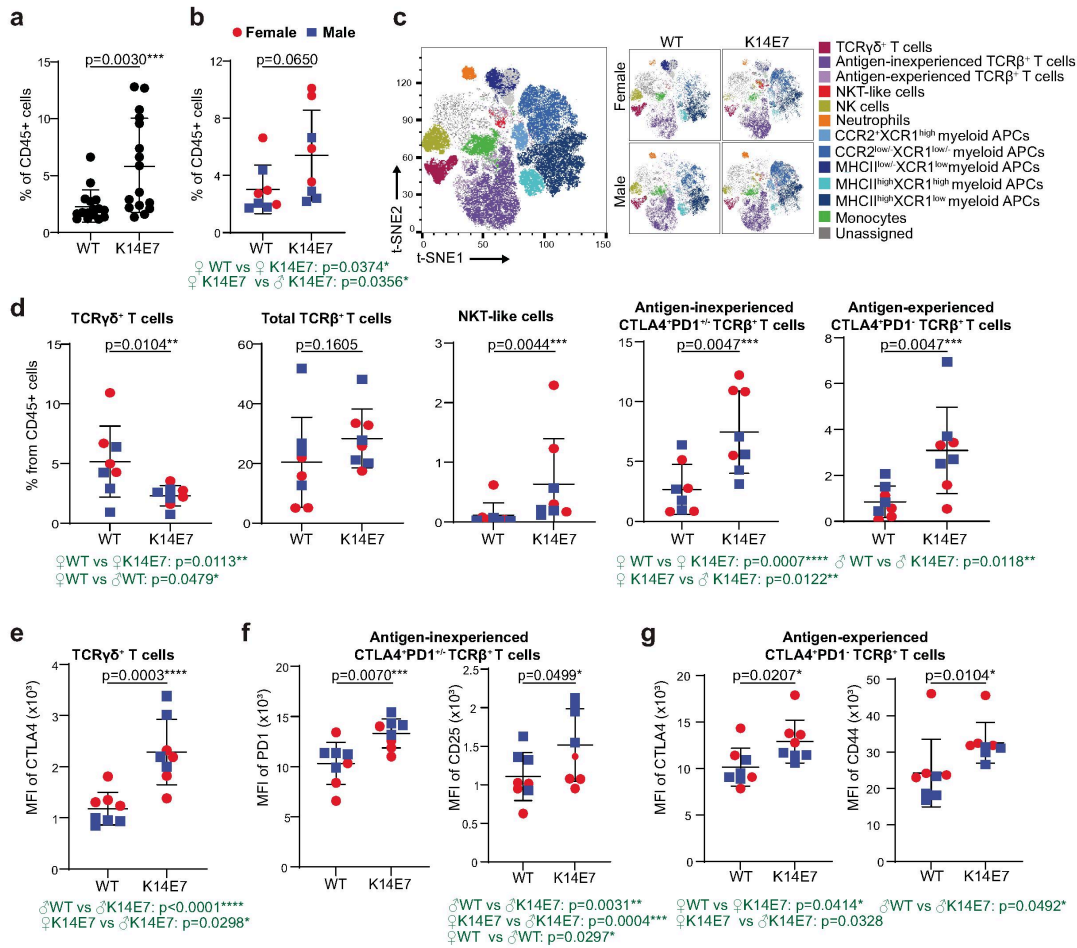


Figure 1

HPV16 E7 expression selectively remodels lymphoid populations and checkpoint-associated T cell phenotypes in the oral epithelium. (a) Frequency of CD45⁺ immune cells in WT and K14E7 tongue epithelium (n = 16 per genotype). (b) Frequency of CD45⁺ immune cells stratified by genotype and sex, with females shown as red circles and males as blue squares (n = 4 females and 4 males per genotype). (c) t-SNE visualisation of concatenated CD45⁺ immune cells from the sex-balanced cohort, showing the

major lymphoid and myeloid populations identified by high-dimensional flow cytometry, together with corresponding maps stratified by genotype and sex. **(d)** Frequencies of TCR $\gamma\delta^+$ cells, total TCR β^+ cells, NK1.1 $^+$ NKT-like cells, CTLA4 $^+$ PD1 $\pm$ antigen-inexperienced TCR β^+ cells, and CTLA4 $^+$ PD1 $^-$ antigen-experienced TCR β^+ cells. **(e)** Median fluorescence intensity (MFI) of CTLA4 within TCR $\gamma\delta^+$ cells. **(f)** MFI of PD1 and CD25 within CTLA4 $^+$ PD1 $\pm$ antigen-inexperienced TCR β^+ cells. **(g)** MFI of CTLA4 and CD44 within CTLA4 $^+$ PD1 $^-$ antigen-experienced TCR β^+ cells. Each symbol represents one biologically independent animal; bars show mean $\pm$ SD. Pooled genotype comparisons were performed using two-sided Mann-Whitney U tests. Sex- and genotype-specific effects were assessed by two-way ANOVA followed by Fisher's LSD tests for secondary pairwise comparisons.

Figure 2

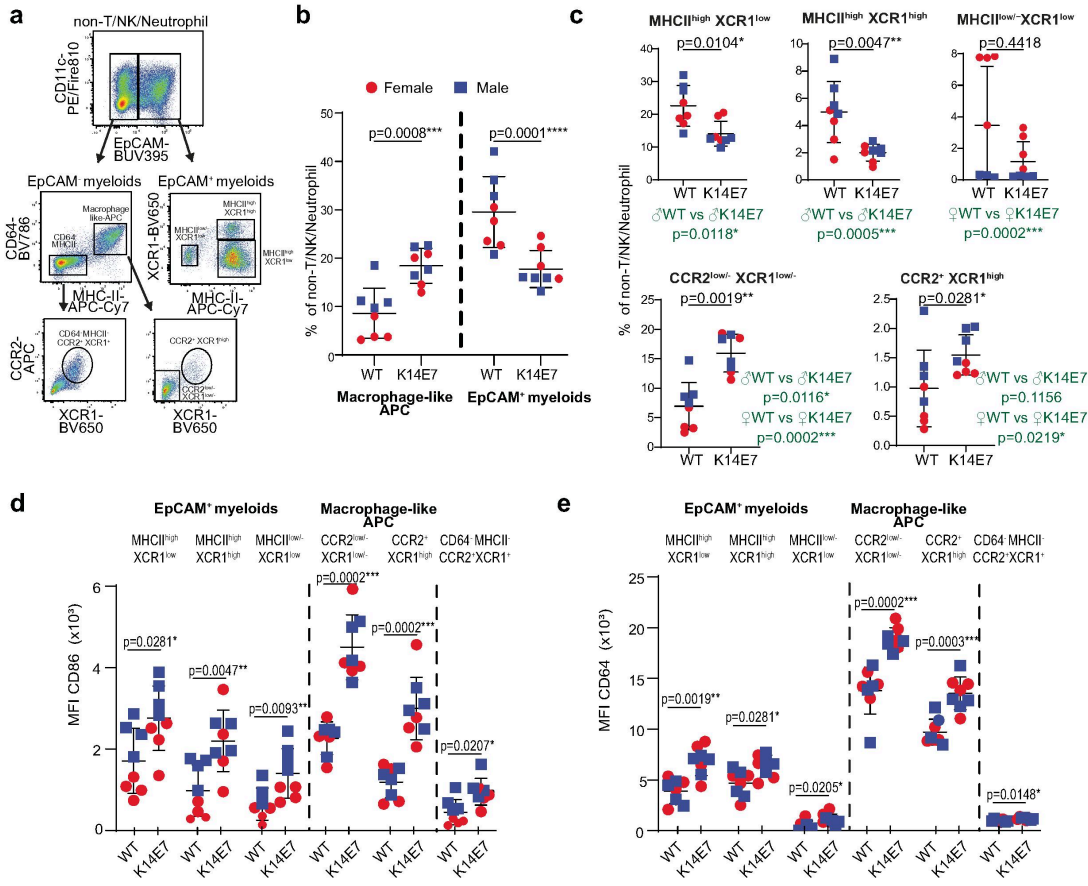


Figure 2

HPV16 E7 expression selectively remodels myeloid antigen-presenting cell compartments within tongue epithelium. (a) Representative gating strategy showing identification of major myeloid populations. EpCAM⁺ myeloid cells were resolved into MHCII^{high}XCR1^{low} and MHCII^{high}XCR1^{high} and MHCII^{low} XCR1^{low} populations. EpCAM⁺ myeloid cells were separated into CD64⁺MHCII⁺ macrophage-like APCs, which were further subdivided according to CCR2 and XCR1 expression. (b) Frequencies of

macrophage-like APCs and EpCAM⁺ myeloid cells expressed as a percentage of non-T/NK/neutrophil cells expressed as a percentage of non-T/NK/neutrophil cells. **(c)** Frequencies of MHCII^{high}XCR1^{low}, MHCII^{high}XCR1^{high}, and MHCII^{low/-}XCR1^{low} EpCAM⁺ myeloid populations, and CCR2^{low/-}XCR1^{low} and CCR2⁺XCR1^{high} macrophage-like APC populations. Secondary sex- and genotype-specific comparisons are indicated in green. **(d, e)** Median fluorescence intensity (MFI) of CD86 **(d)** and CD64 **(e)** across EpCAM⁺ myeloid and macrophage-like APC populations. Females are shown as red circles and males as blue squares. Each symbol represents one biologically independent animal; bars show mean $\pm$ SD ($n = 4$ females and 4 males per genotype). Pooled WT versus K14E7 comparisons were performed using two-sided Mann-Whitney U tests. Sex- and genotype-specific comparisons were analysed by two-way ANOVA followed by Fisher's LSD tests for secondary pairwise comparisons.

Figure 3

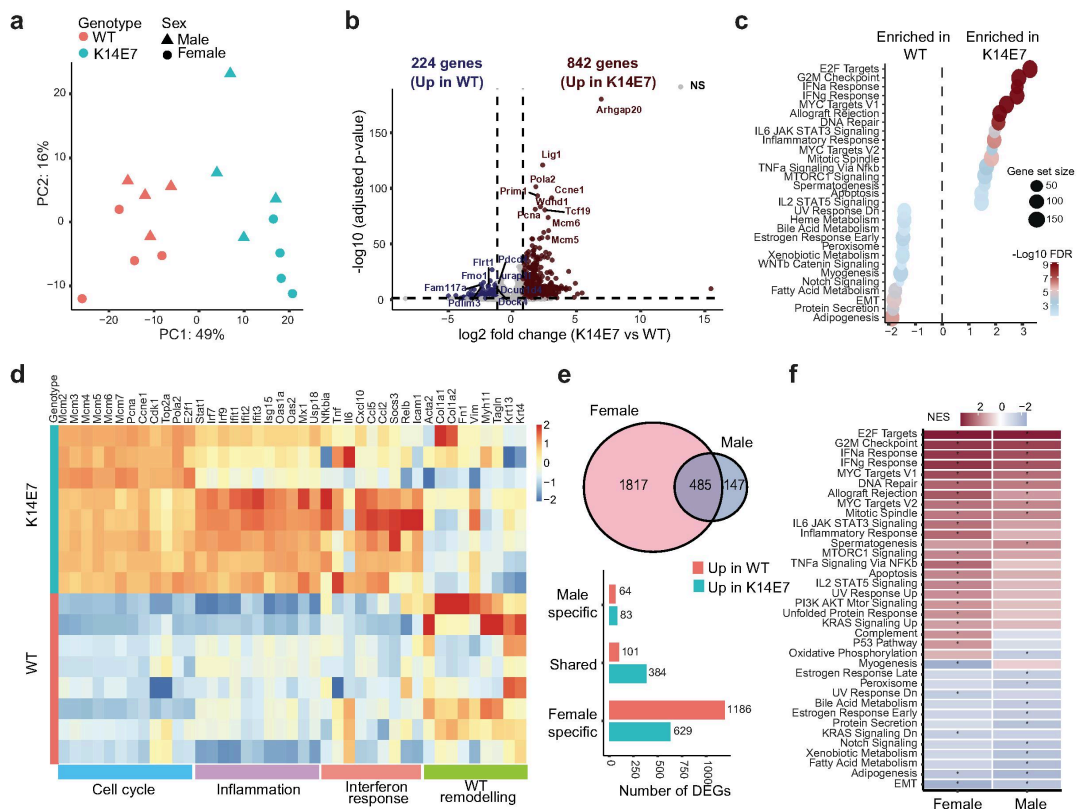


Figure 3

Transcriptional profiling of untreated WT and HPV16 E7-expressing oral epithelium. (a) Principal component analysis (PCA) of untreated WT and K14E7 tongue epithelium demonstrates separation according to genotype. Symbols indicate sex. (b) Volcano plot showing differentially expressed genes (DEGs) defined as Benjamini–Hochberg-adjusted $p < 0.05$ and $|\log_2 \text{fold change}| \geq 1$. Representative genes are labelled. (c) Hallmark Gene Set Enrichment Analysis (GSEA) comparing untreated K14E7 and

WT epithelium. Positive normalised enrichment scores (NES) indicate enrichment in K14E7, whereas negative NES indicate enrichment in WT. Dot size represents gene set size and colour denotes $-\log_{10}(\text{FDR})$. **(d)** Heatmap of representative genes contributing to selected biological programmes identified by GSEA. Genes are grouped according to functional categories including cell cycle, inflammation, interferon response and tissue remodelling. **(e)** Comparison of K14E7-versus-WT differential expression performed independently in female and male untreated epithelium. The upper panel shows the overlap between DEGs identified in the female and male comparisons. The lower panel shows the direction of regulation for DEGs detected only in females, shared between sexes, or detected only in males. DEGs were defined as Benjamini–Hochberg-adjusted $p < 0.05$ and $|\log_2 \text{fold change}| \geq 1$. **(f)** Hallmark GSEA of K14E7 versus WT epithelium performed independently in female and male mice. Tile colour represents the NES. Only Hallmark gene sets with $\text{FDR} < 0.01$ in at least one sex are shown; asterisks indicate $\text{FDR} < 0.01$ in the corresponding comparison. For RNA-seq analyses, $n = 4$ biologically independent animals per sex and genotype.

Figure 4

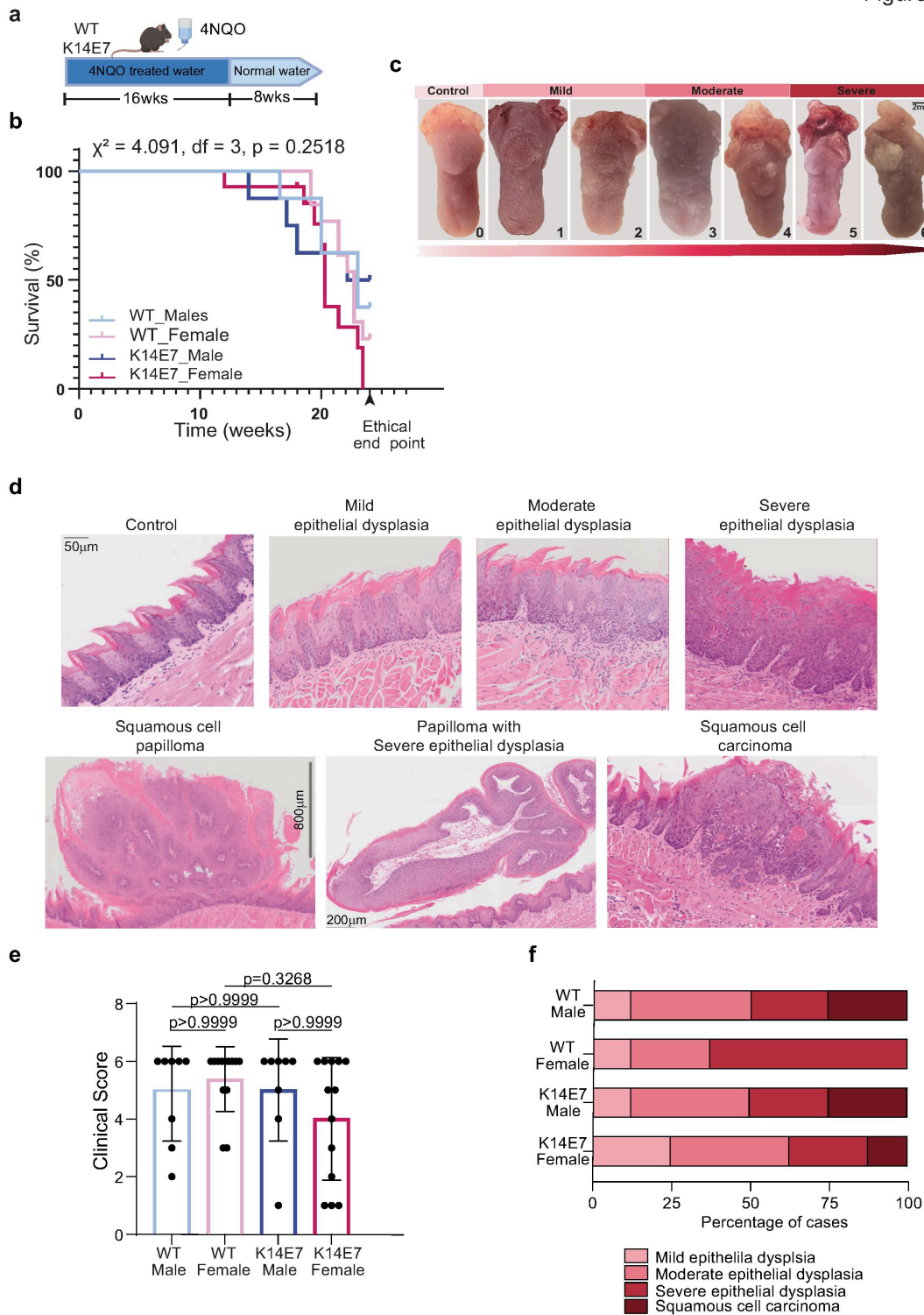


Figure 4

Disease progression following chronic 4NQO exposure in WT and K14E7 mice.

(a) Experimental design. WT and K14E7 mice received 4NQO in the drinking water for 16 weeks followed by an 8-week observation period on normal drinking water before tissue collection. (b) Kaplan–Meier survival analysis of WT and K14E7 mice stratified by sex throughout the experimental period. Survival distributions were compared using the log-rank (Mantel-Cox) test. (c) Representative macroscopic

images corresponding to clinical disease scores ranging from normal tongue (score 0) to severe lesions (score 6). **(d)** Representative haematoxylin and eosin (H&E)-stained tongue sections illustrating the histopathological categories assessed following 4NQO treatment: normal epithelium, mild epithelial dysplasia, moderate epithelial dysplasia, severe epithelial dysplasia, squamous cell papilloma, papilloma with severe epithelial dysplasia, and squamous cell carcinoma. Scale bars are as indicated. **(e)** Clinical disease scores at the experimental endpoint in WT and K14E7 mice stratified by sex. Group differences were assessed using the Kruskal-Wallis test followed by Dunn's multiple-comparisons test. Each symbol represents one biologically independent animal. **(f)** Distribution of histopathological diagnoses within each experimental group, expressed as the percentage of mice exhibiting mild epithelial dysplasia, moderate epithelial dysplasia, severe epithelial dysplasia, or squamous cell carcinoma.

Figure 5

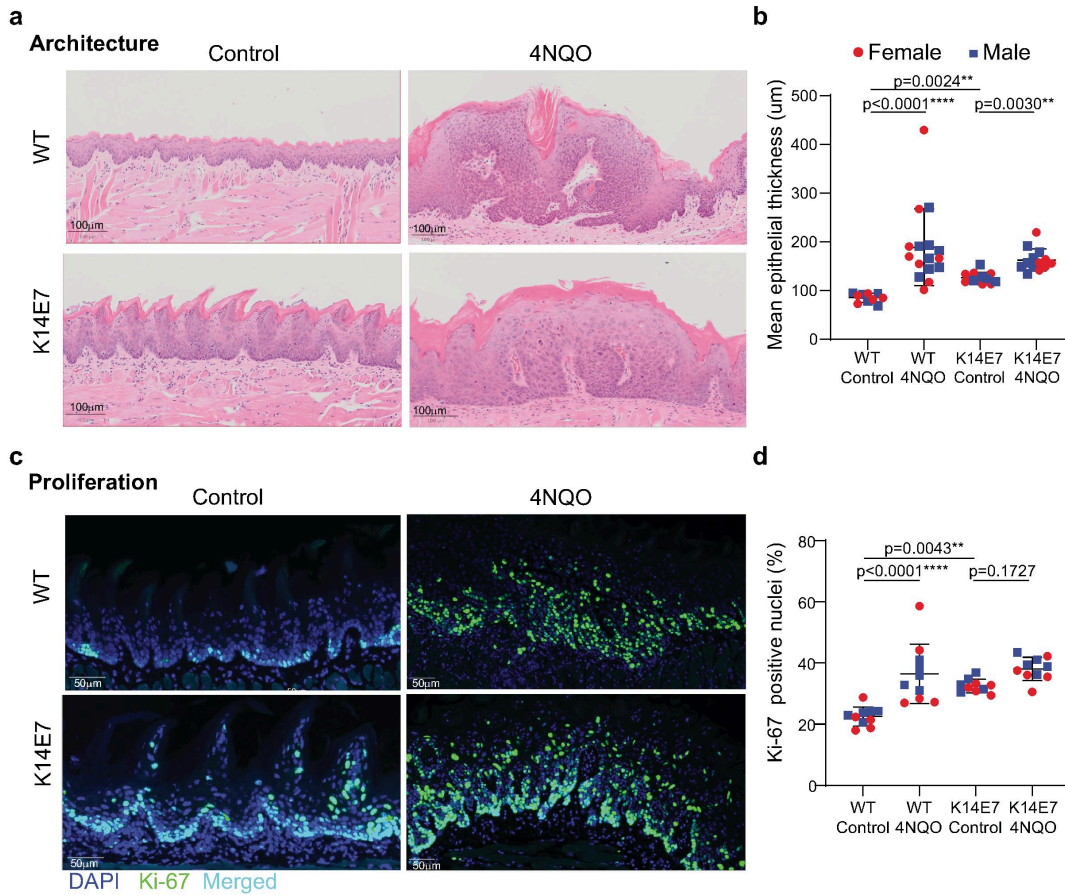


Figure 5

Epithelial thickness and proliferation in WT and K14E7 tongue epithelium before and following 4NQO exposure. (a) Representative haematoxylin and eosin (H&E)-stained tongue sections from control and following 4NQO exposed WT and K14E7 mice. Scale bars, 100-µm. (b) Quantification of mean epithelial thickness in untreated control and 4NQO-exposed WT and K14E7 mice. (c) Representative immunofluorescence images showing Ki-67 (green) and DAPI (blue) staining in untreated control and

4NQO-exposed WT and K14E7 tongue epithelium. Scale bars, 50-μm. **(d)** Quantification of Ki-67-positive nuclei as a percentage of total epithelial nuclei. Females are shown as red circles and males as blue squares. Each symbol represents one biologically independent animal; bars show mean ± SD. Pairwise comparisons were performed using Tukey's multiple-comparisons test.

Figure 6

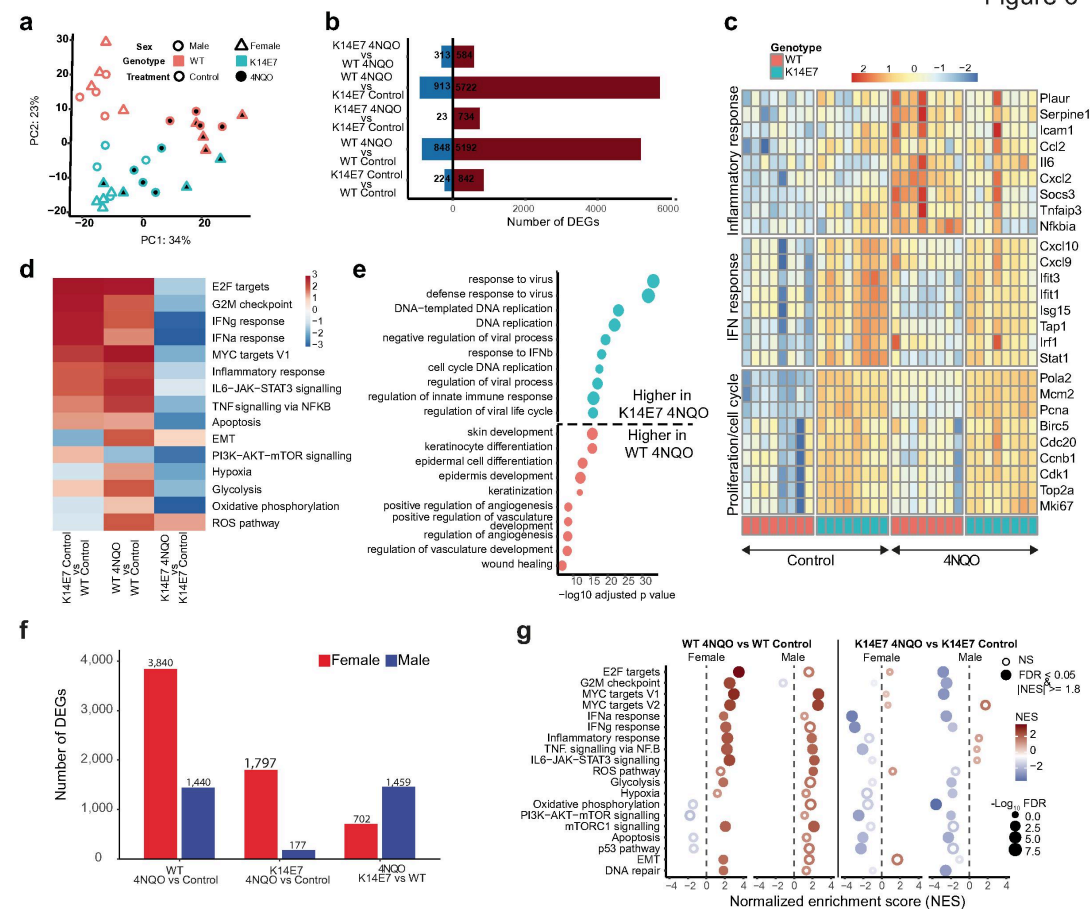


Figure 6

Transcriptional responses to 4NQO exposure in WT and HPV16 E7-expressing oral epithelium. (a)

Principal component analysis (PCA) of bulk RNA-sequencing data from WT and K14E7 tongue epithelium in untreated control mice and following 4NQO exposure ($n = 8$ per genotype/treatment group; 4 females and 4 males per group). Samples are coloured according to genotype and treatment, with symbols indicating sex. PC1 and PC2 explain 34% and 23% of the variance, respectively. **(b)** Numbers of differentially expressed genes (DEGs) across the indicated comparisons. DEGs were defined as Benjamini–Hochberg-adjusted $p < 0.05$ and $|\log_2 \text{fold change}| \geq 1$. Red and blue bars indicate genes expressed at higher levels in the first and second groups of each comparison, respectively. **(c)** Heatmap showing $\log_2$ -transformed normalised expression of representative genes associated with inflammatory response, interferon response and proliferation/cell cycle across experimental groups. Genes were selected as representative members of biological programmes identified by Hallmark GSEA and grouped by functional category. Colours indicate relative expression across samples, with warmer colours representing higher and cooler colours lower expression. **(d)** Hallmark Gene Set Enrichment Analysis (GSEA) across the indicated genotype and treatment comparisons. Tile colour represents the normalised enrichment score (NES), with positive and negative values indicating enrichment towards the first and second groups of each comparison, respectively. **(e)** Gene Ontology (GO) Biological Process enrichment analysis of genes differentially expressed between K14E7 and WT epithelium following 4NQO exposure, performed separately for genes expressed at higher levels in K14E7 and WT. Bubble size represents gene count and the x-axis indicates $-\log_{10}$ -adjusted p value. **(f)** UpSet plot showing overlap of DEGs across the indicated sex-stratified comparisons. DEGs were defined as Benjamini–Hochberg-adjusted $p < 0.05$ and $|\log_2 \text{fold change}| \geq 1$. Horizontal bars indicate the total number of DEGs in each comparison, vertical bars indicate intersection size, and connected dots identify the comparisons contributing to each intersection. **(g)** Hallmark GSEA performed separately in female and male mice for the indicated comparisons. Bubble position represents NES, bubble size represents $-\log_{10}(\text{FDR})$, filled circles indicate $\text{FDR} < 0.05$, and open circles indicate $\text{FDR} \geq 0.05$.

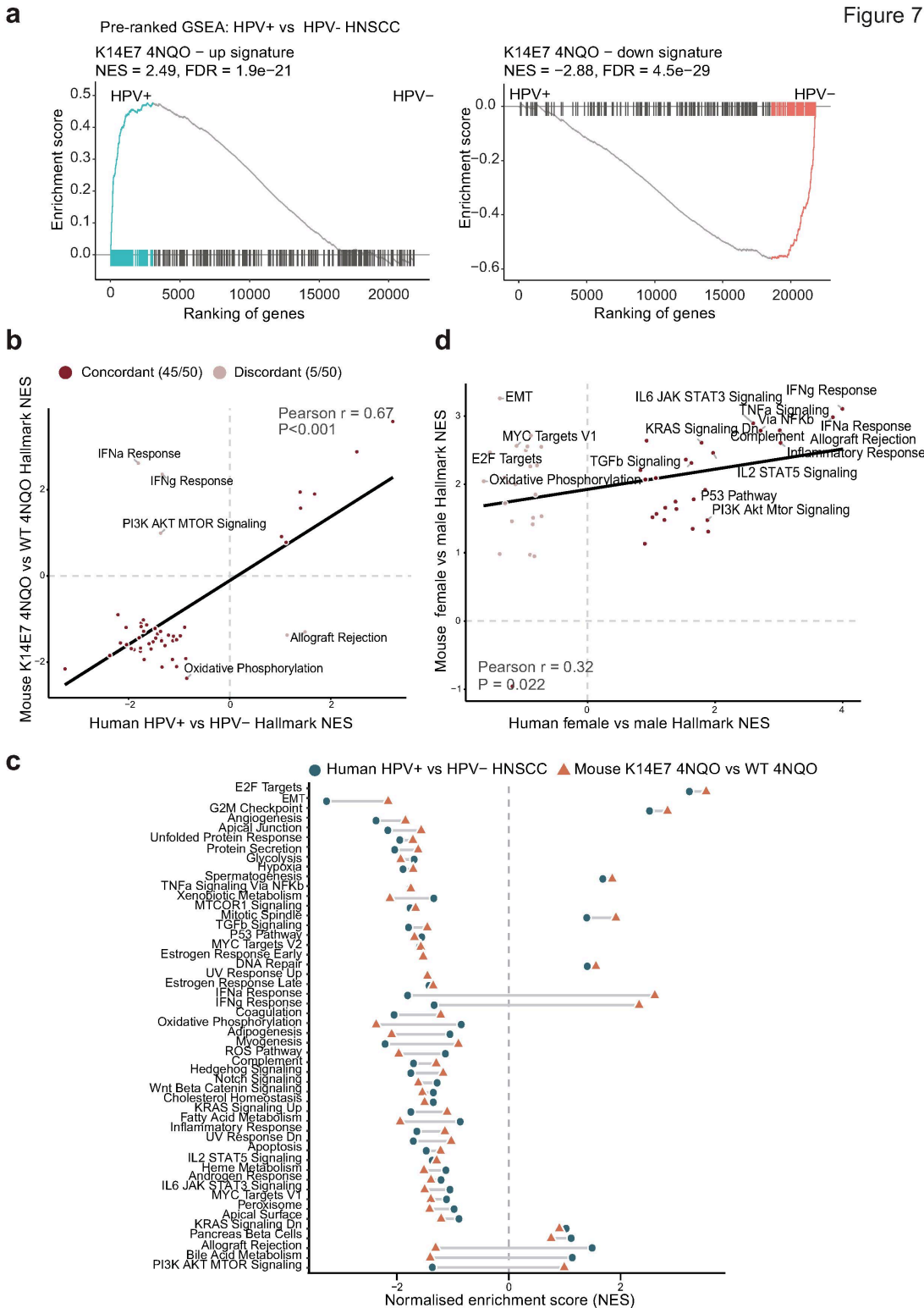


Figure 7

Cross-species conservation of transcriptional programmes between murine HPV16 E7-associated oral carcinogenesis and human HNSCC. (a) Pre-ranked gene set enrichment analysis (GSEA) testing transcriptional signatures derived from K14E7 versus WT epithelium following 4NQO exposure against the ranked human TCGA HPV⁺ versus HPV⁻ HNSCC transcriptome. Separate analyses were performed using genes expressed at higher and lower levels in K14E7 epithelium. Normalised enrichment scores

(NES) and false discovery rates (FDR) are shown. **(b)** Comparison of Hallmark pathway NES between K14E7 versus WT murine epithelium following 4NQO exposure and HPV-positive versus HPV-negative human HNSCC. Each point represents one of the 50 Hallmark gene sets and is classified according to concordant or discordant direction of enrichment between species. The solid line represents the linear regression fit and dashed lines indicate NES = 0. Pearson correlation coefficient and associated *p* value are shown **(c)** Comparison of Hallmark pathway NES between K14E7 versus WT murine epithelium following 4NQO exposure and HPV-positive versus HPV-negative human HNSCC. Circles represent human HPV⁺ versus HPV⁻ HNSCC and triangles represent 4NQO-treated K14E7 versus WT murine epithelium; connecting lines link the corresponding Hallmark gene sets across species. **(d)** Comparison of sex-associated Hallmark pathway NES between female versus male murine epithelium following 4NQO exposure and female versus male human HNSCC. Positive NES values indicate enrichment towards females and negative values enrichment towards males. Each point represents one of the 50 Hallmark gene sets. The solid line represents the linear regression fit and dashed lines indicate NES = 0. Pearson correlation coefficient and associated *p* value are shown.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Hettiarachchi et al Supplementary IJOS.docx](#)
- [Supplementary data.xlsx](#)