Supplementary

J.K. Lilja-Fischer et al: Characterization and radiosensitivity of HPV-related oropharyngeal squamous cell carcinoma patient-derived xenografts.

Supplementary methods

Histology and immunohistochemistry

Tumor differentiation of original patient samples as well as PDX tumors were classified according to WHO criteria. Immunohistochemistry for p16 was used as a surrogate marker of HPV, with strong nuclear and cytoplasmic staining in >70% of cells considered positive. Furthermore, immunohistochemistry for cytokeratin AE1/AE3 and CD45 was performed on PDX tumor samples. Lastly, examination for Epstein-Barr virus related proteins was performed using EBER in situ hybridization.

Quantitative RT-PCR

RNA was extracted from FFPE tumor tissue samples in an automated setup (Siemens Tissue Preparation System). Following cDNA generation, quantitative PCR was performed using TaqMan gene expression assays (Life Technologies). We examined expression of the 15 genes included in the hypoxia classifier described by Toustrup *et al*, and cases were classified as "less" or "more" hypoxic as described previously. Also, we investigated expression of 10 genes in the radiosensitivity index (RSI) described by Scott *et al.* by deriving the hypothetical genomic-adjusted radiation dose (GARD) using a standard of care radiation dose of 68 Gy in 34 fractions. Furthermore, we included a set of 6 genes previously described by Keck *et al* to distinguish an inflamed/mesenchymal (IM) or Classical subtype of HPV-positive tumors. Finally, we analyzed gene expression of *PD-1, PD-L1, TRIP-12, MET* and *SLC3A2*. A list of genes included in the assay is provided in **supplementary table 2**. Gene expression was quantified relative to household genes by the ΔCt-method as described by Toustrup *et al*. Gene expression in primary tumors and PDX tumors was compared with Pearson's correlation coefficient.

HPV classification

We defined tumors as HPV-positive if they were p16-positive by immunohistochemistry using a standardized classification, and positive for HPV *E6* or *E7* RNA or high-risk HPV DNA.
All tumors were examined for HPV-16 and HPV-18 *E6* and *E7* RNA, with PCR primers designed in-house. HPV-RNA negative, p16-positive samples were further examined for high-risk HPV DNA with the *INNO-LiPA® HPV Genotyping Extra II* assay (Fujirebio).

Cancer gene targeted sequencing

Targeted sequencing of 409 cancer-related genes included in the Ion AmpliSeq Comprehensive Cancer Panel (CCP) was performed using the Ion Torrent S5 platform running the Torrent Suite version 5.2.2 (Thermo Fisher Scientific). Briefly, formalin-fixed, paraffin embedded (FFPE) tumor samples were evaluated for tumor contents. After gross dissection, if necessary, DNA was isolated in an automated setup using a QIAsymphony SP and the QIAsymphony DSP DNA mini kit (ref. no: 937236). Library preparation (Ion AmpliSeq Library Kit 2.0, Thermo Fisher Scientific), amplification, and sequencing (using the Ion 540™ Kit-Chef and the Ion 540™ Chip Kit, Thermo Fisher Scientific) were performed according to the manufacturer's instructions.

Sequencing data analysis

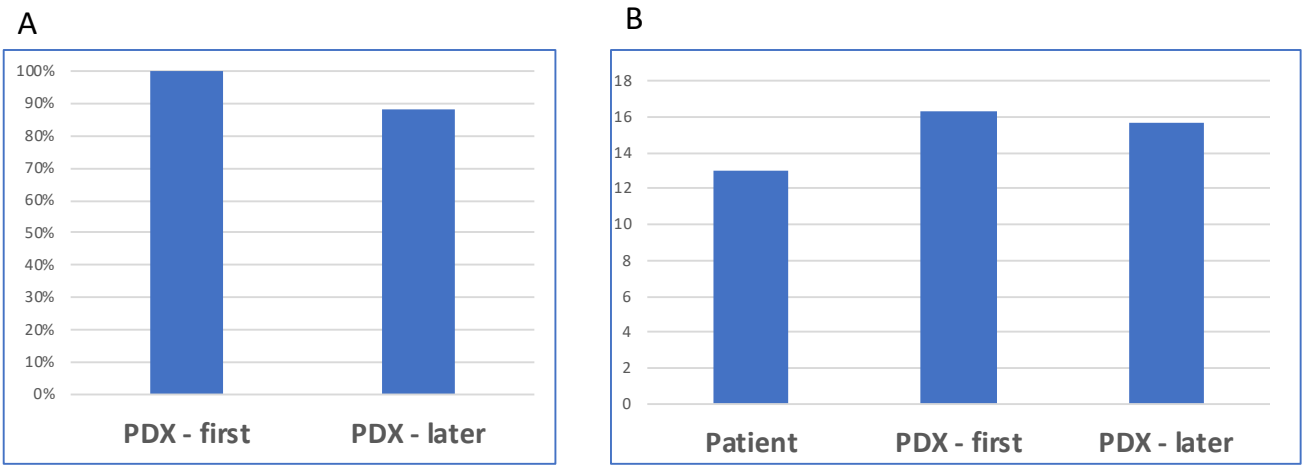
The Ion Reporter software 5.6 (Thermo Fisher Scientific) was used for variant calling using the Ampliseq CCP single sample workflow with additional filters. We excluded variants at known polymorphic sites, sites with homopolymeres of more than 4 bases, homozygous variants and variants <10 reads or <5% of total reads, or coverage less than 100x. Also, we excluded variants present in control samples or present in 1% or more in publicly available genomes (1000 genomes NHLBI ESP / ExAC). Further filtering was enabled to retain only non-synonymous variants with a probable somatic nature; and thus, variants described are presumed to be somatic mutations.

Radiosensitivity studies

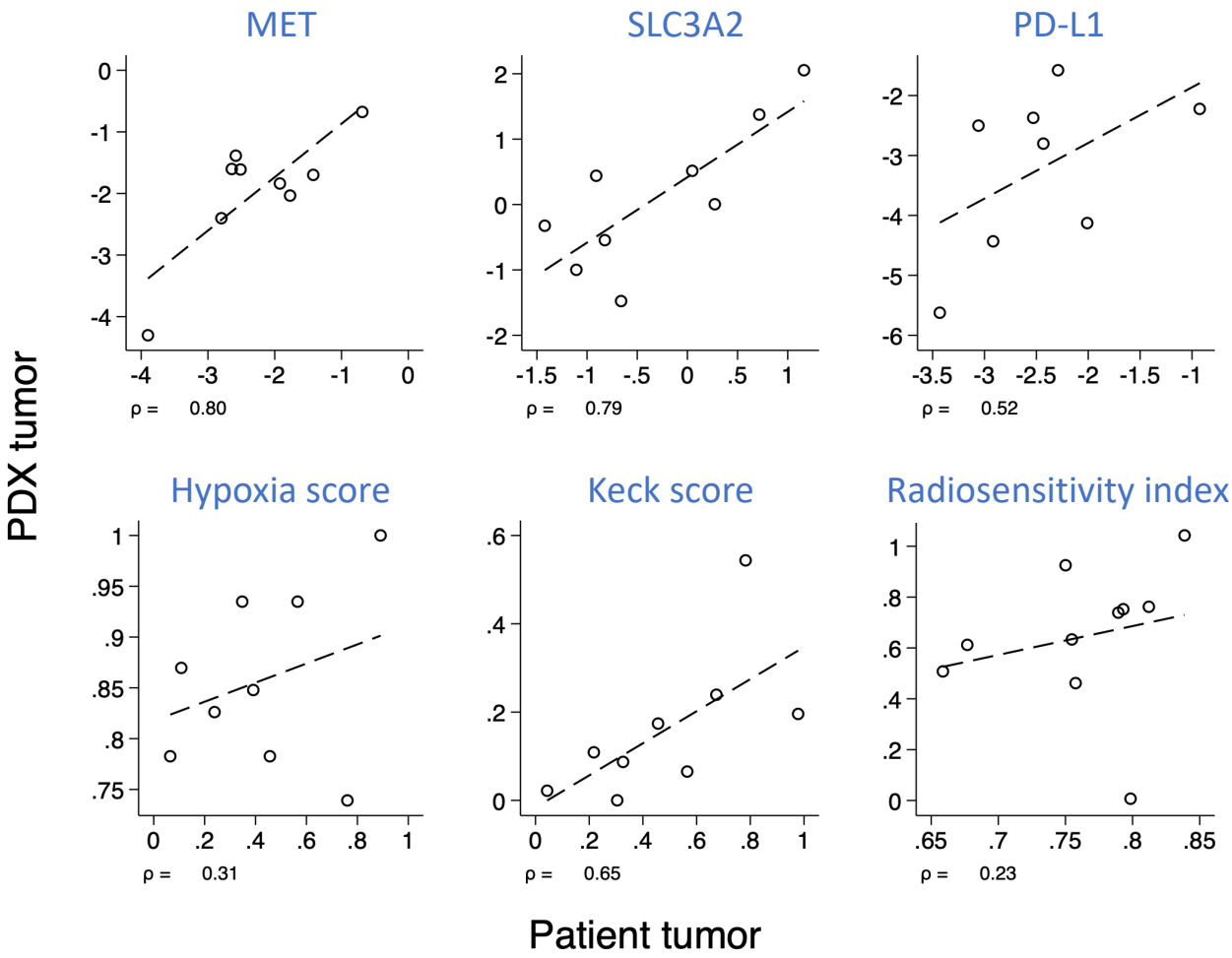
Tumor size was measured in three dimensions using calipers, and volume estimated as $A*B*C*\pi/6$. Tumor regrowth delay was estimated as median time to three times initial tumor volume (Tumor growth time; TGT). Specific growth delay (SGD) was estimated as $(TGT_{treated} - TGT_{controls})/ TGT_{controls}$, which allows comparison between different tumor models.

Supplementary figures

Supplementary figure 1. **A**. Proportion of cancer driver gene mutations from primary tumor samples, that were preserved in PDX samples. Driver genes were *TP53* (pathogenic), *PIK3CA* (hotspot), *KMT2D, PDE4DIP, SYNE1, KMT2C, EP300, NOTCH1, EGFR, CDKN2A,* and *NFE2L2*.
B. Mutation load expressed as number of variants present in the tumor sample. There was a mean of 13 variants in primary tumors versus approximately 16 variants per first and subsequent PDX tumor generation (p=0.11 and 0.13, paired t-test).

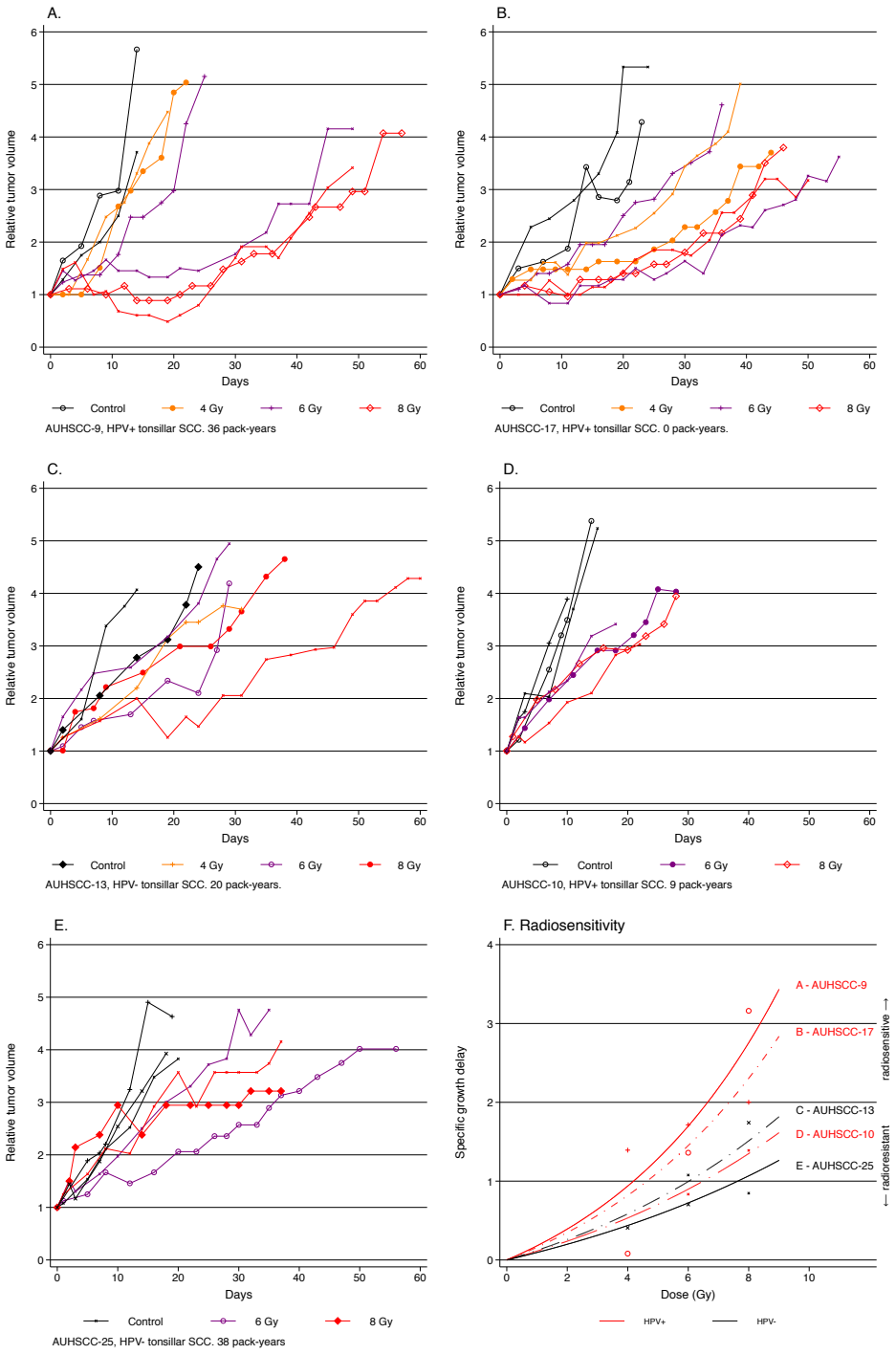


Supplementary figure 2. Expression of selected genes, quantified by qPCR. Top row shows expression of cancer stem cell markers *MET* and *SLC3A2*, as well as *PD-L1*. Bottom row shows scores derived from median expression of a number of genes. Hypoxia score is the ranked median expression of genes in the 15-gene hypoxia classifier described by Toustrup *et al*, and similarly the "Keck score" is the expression of 6 genes used to distinguish between the "Classical" and the "Inflamed/Mesenchymal" subtype of HPV-positive SCC described by Keck *et al*. The radiosensitivity index is similarly based on expression of 10 different genes described by Eschrich *et al*. ρ is Pearson's correlation coefficient.

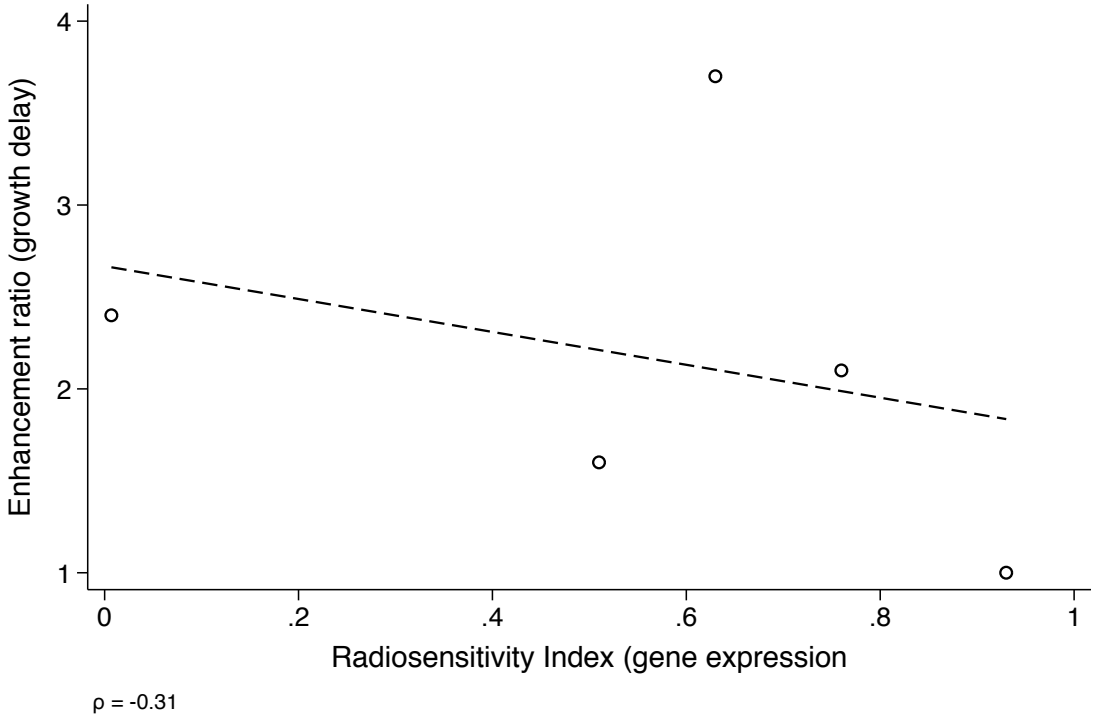


Supplementary figure 3. Normalized gene expression values, color indicates variation from minimum to maximum expression for each gene. Rows are genes, columns are tumor samples.

Supplementary figure 4. A-E: Tumor regrowth delay after irradiation in 5 individual PDX models derived from 5 patients. Each curve shows tumor volume, relative to volume at treatment start, which was approximately 200 mm^3. Colors indicate dose to the individual tumor. Time is days since treatment. **F:** Specific growth delay in the 5 different PDX models. Curves are fitted exponential functions. Data are also shown in **Figure 3**.



Supplementary figure 5. Correlation between *in vivo* radiosensitivity and Radiosensitivity Index derived from gene expression analyses.



Supplementary tables

Supplementary table 1. Characteristics of included patients, and patients with established PDX model.

	All patients (n=34)	PDX established (n=12)	*p*
Age, years			
Median (IQR)	64 (56-69)	62 (57-70)	0.9
Gender			0.7
Male	24 (71%)	8 (67%)	
Female	10 (29%)	4 (33%)	
Smoking			0.7
Non-smoker (<10 pack-years)	14 (41%)	4 (33%)	
Smoker (≥10 pack-years)	20 (59%)	8 (67%)	
Tumor site			0.04
Tonsil	21 (62%)	9 (75%)	
Base of tongue	8 (24%)	0 (0%)	
Oropharynx, other	4 (12%)	2 (17%)	
Hypopharynx	1 (3%)	1 (8%)	
T classification			0.5
T1	11 (32%)	3 (25%)	
T2	12 (35%)	3 (25%)	
T3	6 (18%)	3 (25%)	
T4	5 (15%)	3 (25%)	
N classification			0.04
N0	6 (18%)	0 (0%)	
N1	1 (3%)	1 (8%)	
N2a	3 (9%)	0 (0%)	
N2b	18 (53%)	7 (58%)	
N2c	6 (18%)	4 (33%)	
Stage			0.7
I	2 (6%)	0 (0%)	
II	2 (6%)	0 (0%)	
III	3 (9%)	1 (8%)	
IV	27 (79%)	11 (92%)	
Differentiation			0.02
High	3 (9%)	2 (17%)	
Moderate	16 (48%)	2 (17%)	
Poor	14 (41%)	8 (67%)	
p16 status			0.2
Positive	24 (71%)	7 (58%)	
Negative	10 (29%)	5 (42%)	
p16 and smoking combined			0.7
p16 negative	10 (29%)	5 (42%)	
p16 positive smoker	10 (29%)	3 (25%)	
p16 positive non-smoker	14 (41%)	4 (33%)	

IQR: Interquartile range. Staging is UICC 7th edition. p-values are from Fisher's exact test, except from age, which was tested with a rank-sum test.

Supplementary table 2. Genes included in qPCR gene expression assay.

15-gene hypoxia classifier (15 genes) (Toustrup *et al*, 2011)	*ADM, ALDOA, ANKRD37, BNIP3, BNIP3L, EGLN3, FAM162A, KCTD11, LOX, NDRG1, P4HA1, P4HA2, PDK1, PFKFB3, SLC2A1*
Gene-expression subtype (6 genes) (Keck *et al*, 2014)	*ICOS, CD8A, LAG3, HLA-DRA, VIM, MMP9*
Radiosensitivity Index (10 genes) (Eschrich *et al*, 2009)	*AR, JUN, STAT1, PRKCB, RELA, ABL1, SUMO1, PAK2, HDAC1, IRF1*
HPV (5 genes) (Marie Schou, unpublished)	*CDKN2A* (p16) HPV16: *E6, E7*. HPV18: *E6, E7*.
Other (5 genes)	*MET, SLC3A2, TRIP12, PDCD1* (PD-1), *CD274* (PD-L1)
Reference genes (3 genes) (Toustrup *et al*, 2011)	*ACTR3, NDFIP1, RPL37A*

41 genes, 3 reference genes.