Supplementary

Supplementary table 1. List of genes included in custom NGS panel (DAHANCA 19 cohort).

A. List of genes included in custom NGS panel (DAHANCA 19 cohort).

ATM, ATR, BCL6, CASP8, CCND1, CDKN2A, CDKN2B, CREBBP, CSMD3, CYLD, EGFR, EP300, FAT1, FBXW7, FGFR3, HLA-A, HLA-B, HRAS, KEAP1, KLLN, KMT2D, KRAS, MAPK1, MDM2, MET, NBN, NFE2L2, NOTCH1, NRAS, NSD1, PIK3CA, PTEN, RB1, SOX2, TBCD, TERC, TERT, TGFBR2, TP53, TRAF3, TSC1, ZNF750

B. List of genes included in RT-qPCR gene expression assays.

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

Supplementary table 2.

NGS variant analysis. Variants were filtered in the Ion Reporter software using this filter chain to exclude artefacts or normal variants.

- 10 <= Alternate Allele Count <= 100000
- 0 <= Homopolymer Length <= 4
- Variant Effect in unknown, missense, nonframeshiftInsertion, nonframeshiftDeletion, nonframeshiftBlockSubstitution, nonsense, stoploss, frameshiftInsertion, frameshiftDeletion, frameshiftBlockSubstitution
- 0.02 <= Allele Ratio <= 1.0
- Location in splicesite_5, splicesite_3, exonic
- UCSC Common SNPs = Not In
- Variant Type in SNV, INDEL, MNV, CNV, LONGDEL, FUSION, ASSAYS_5P_3P
- Filtered Coverage >= 100
- <= Minor Allele Frequency <= 0.03
- <= 5000Exomes Global MAF <= 0.01
- Variant Classification in Pathogenic, Likely Pathogenic, VUS, Uncertain Significance, Unknown, Suspected Deleterious, Deleterious

Supplementary table 3. Univariate analysis of prognostic factors in HPV+ OPSCC patients in the DAHANCA 19 cohort (n=128).

	Locoregional failure			**Overall survival**	
	n	Events	HR (95% CI)	Event	HR (95% CI)
All patients	128	17		37	
Age					
<60	76	10	Ref.	16	Ref.
≥60 years	52	7	0.95 (0.36 - 2.5)	21	1.2 (0.59 - 2.2)
Gender					
Male	104	16	3.7 (0.49 - 28)	34	1.7 (0.52 - 5.7)
Female	24	1	Ref.	3	Ref.
Performance status					
0	114	13	Ref.	30	Ref.
1	13	4	2.6 (0.84 - 8.0)	7	2.2 (0.95 - 5.2)
2	1	0	N/A	0	N/A
T					
T1	32	3	Ref.	7	Ref.
T2	68	8	1.2 (0.32 - 4.5)	19	1.7 (0.71 - 4.2)
T3	20	3	1.5 (0.31 - 7.7)	6	2.6 (0.84 - 8.4)
T4	8	3	4.2 (0.85 - 21)	5	5.6 (1.7 - 19)
N					
N0	5	0	N/A	4	Ref.
N1	27	6	Ref.	12	0.57 (0.18 - 1.8)
N2	90	8	0.38 (0.13 - 1.1)	18	0.38 (0.12 - 1.2)
N3	6	3	2.7 (0.65 - 11)	3	1.8 (0.37 - 8.8)
Tobacco status					
Never	32	2	Ref.	2	Ref.
Former	74	10	2.1 (0.47 - 9.8)	27	2.6 (0.59 - 12)
Current	22	5	3.6 (0.70 - 19)	8	2.9 (0.58 - 14)
Pack-years					
<10	50	7	Ref.	8	Ref.
≥10	78	10	0.84 (0.32 - 2.2)	29	1.1 (0.50 - 2.6)

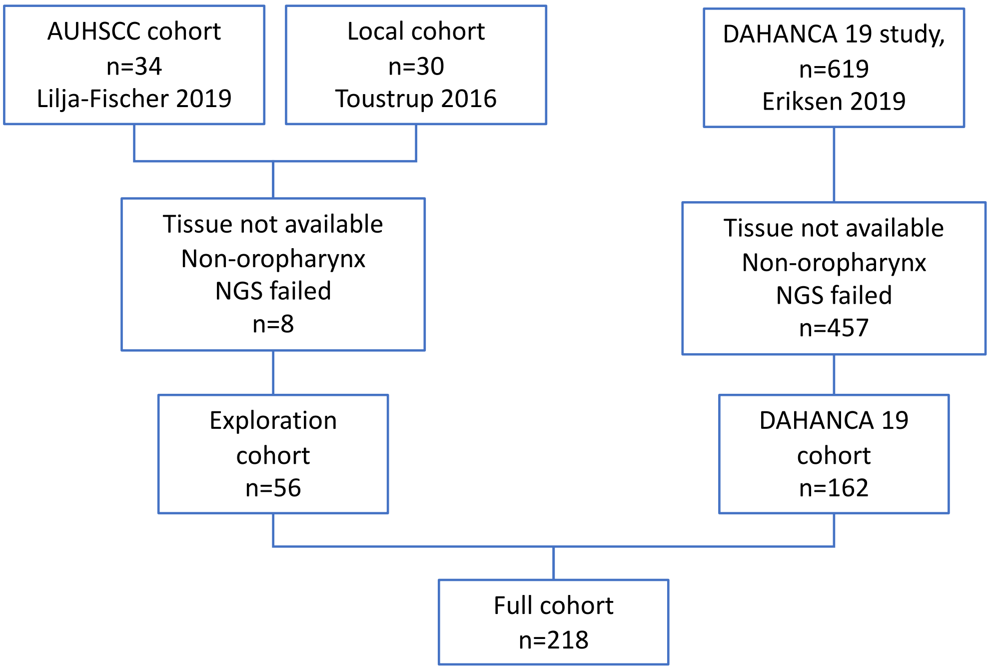
N/A: Not available.

Explained variation (R^2) for all variables combined in a multivariate model:
Locoregional failure: 0.33 (95% CI: 0.1 – 0.54).
Overall survival: 0.21 (95% CI: 0.04 – 0.41).

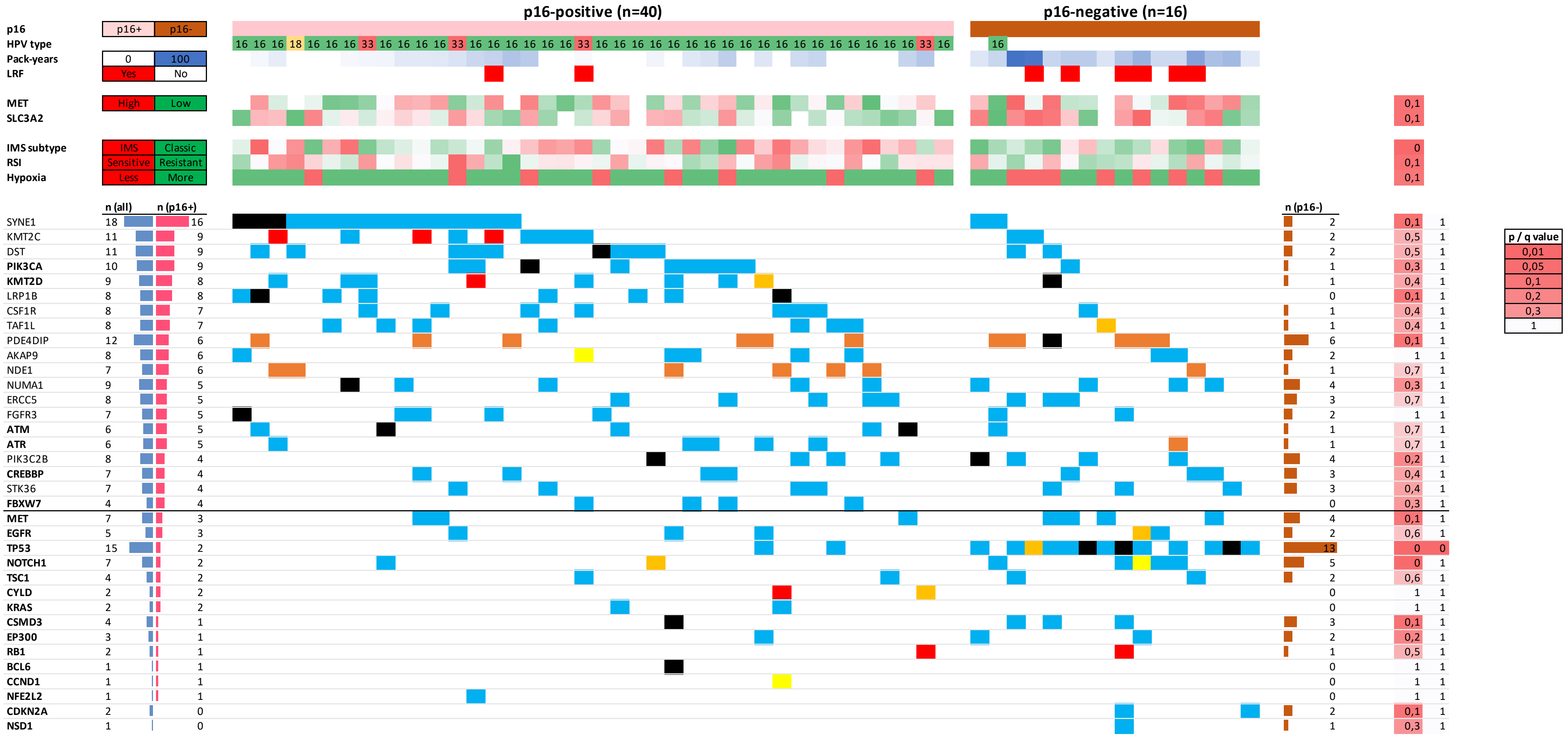
Supplementary figures.

Supplementary figure 1. Flowchart of included patients.

Cohorts are detailed in Lilja-Fischer (2019),[11] Toustrup (2016)[10] and Eriksen (2018).[12]

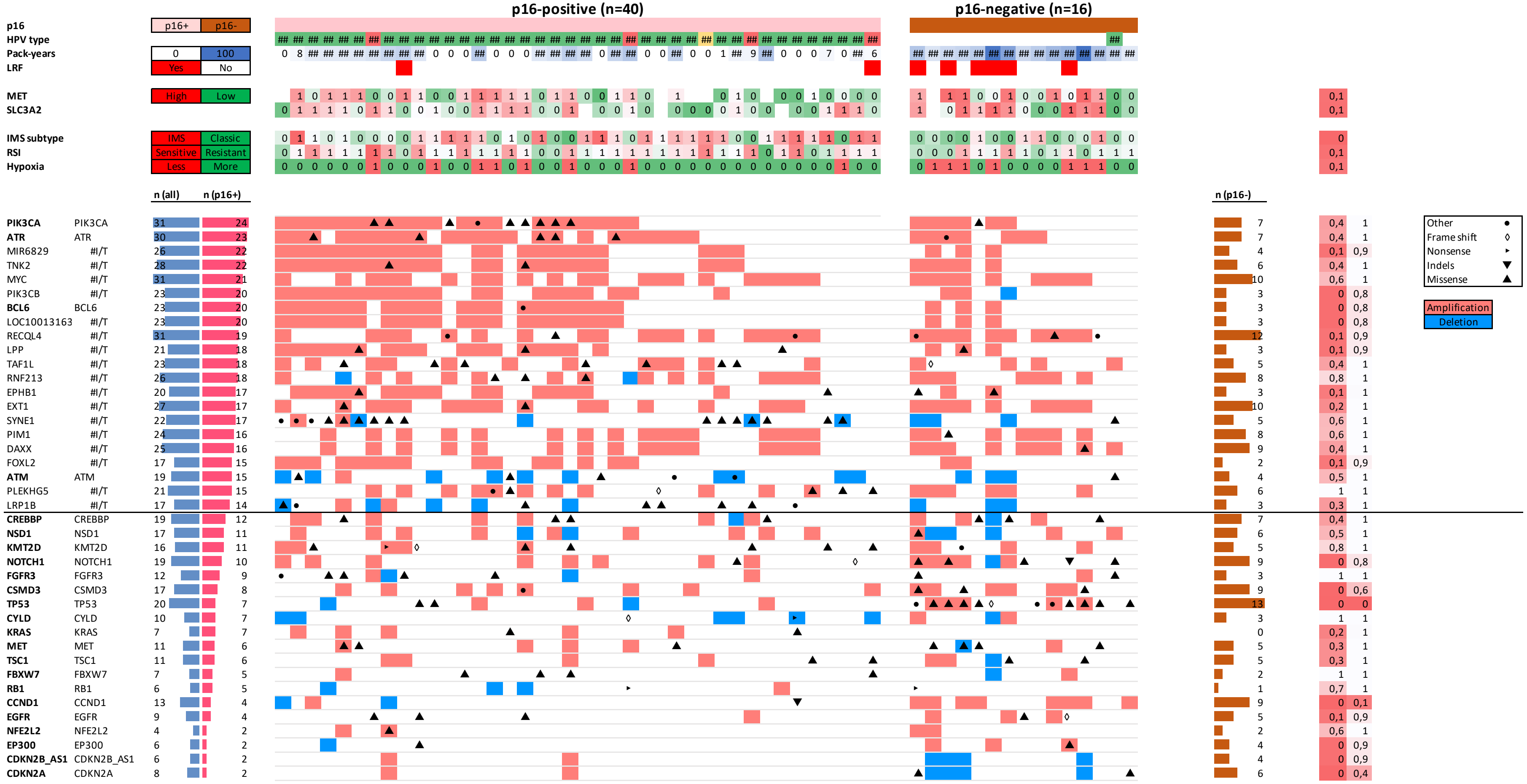


Supplementary figure 2. Top mutated genes (rows) from the CCP panel in 56 OPSCC tumors (columns). Top half shows tumors most frequently mutated, bottom half genes that were also selected for a custom gene panel. Genes are ordered by frequency among p16+ tumors.

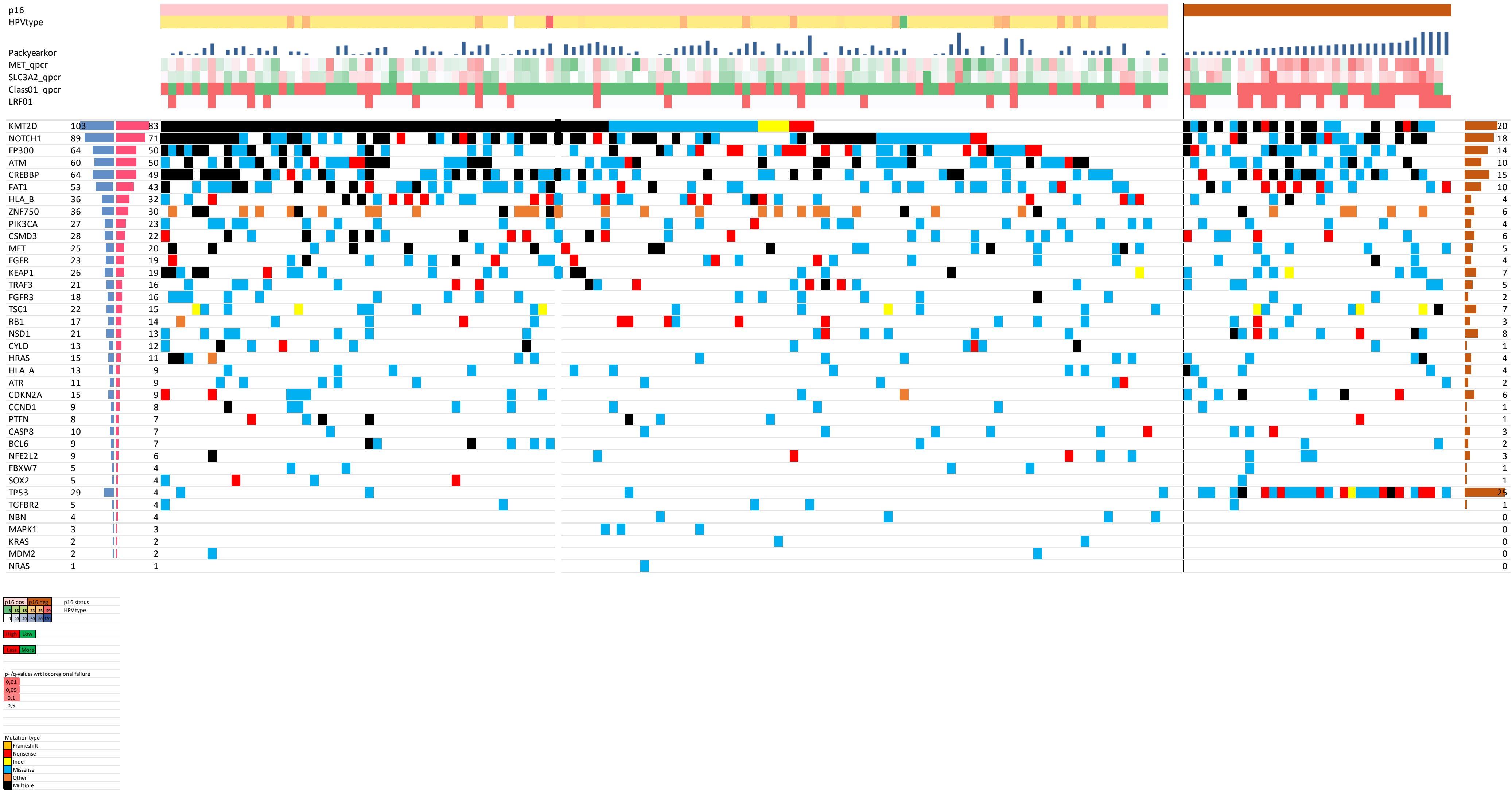


	n (all)	n (p16+)	n (p16-)		
SYNE1	18	16	2	0,1	1
KMT2C	11	9	2	0,5	1
DST	11	9	2	0,5	1
PIK3CA	10	9	1	0,3	1
KMT2D	9	8	1	0,4	1
LRP1B	8	8	0	0,1	1
CSF1R	8	7	1	0,4	1
TAF1L	8	7	1	0,4	1
PDE4DIP	12	6	6	0,1	1
AKAP9	8	6	2	1	1
NDE1	7	6	1	0,7	1
NUMA1	9	5	4	0,3	1
ERCC5	8	5	3	0,7	1
FGFR3	7	5	2	1	1
ATM	6	5	1	0,7	1
ATR	6	5	1	0,7	1
PIK3C2B	8	4	4	0,2	1
CREBBP	7	4	3	0,4	1
STK36	7	4	3	0,4	1
FBXW7	4	4	0	0,3	1
MET	7	3	4	0,1	1
EGFR	5	3	2	0,6	1
TP53	15	2	13	0	0
NOTCH1	7	2	5	0	1
TSC1	4	2	2	0,6	1
CYLD	2	2	0	1	1
KRAS	2	2	0	1	1
CSMD3	4	1	3	0,1	1
EP300	3	1	2	0,2	1
RB1	2	1	1	0,5	1
BCL6	1	1	0	1	1
CCND1	1	1	0	1	1
NFE2L2	1	1	0	1	1
CDKN2A	2	0	2	0,1	1
NSD1	1	0	1	0,3	1

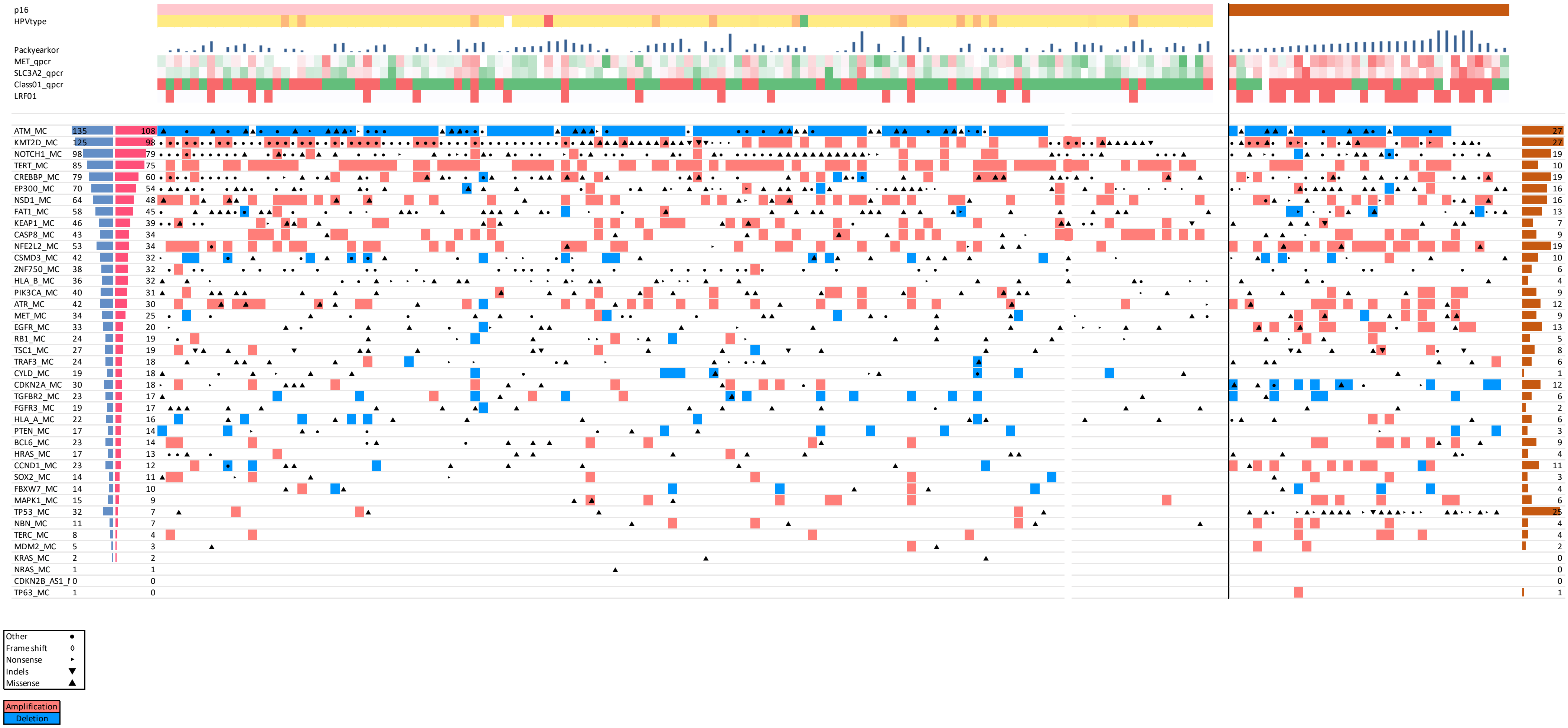
Supplementary figure 3. Top altered (mutation, amplification, or deletion) genes in rows from the CCP panel in 56 OPSCC tumors (columns). Top half shows genes most frequently altered, bottom half shows genes that were also selected for a custom gene panel. Genes are ordered by frequency among p16+ tumors.



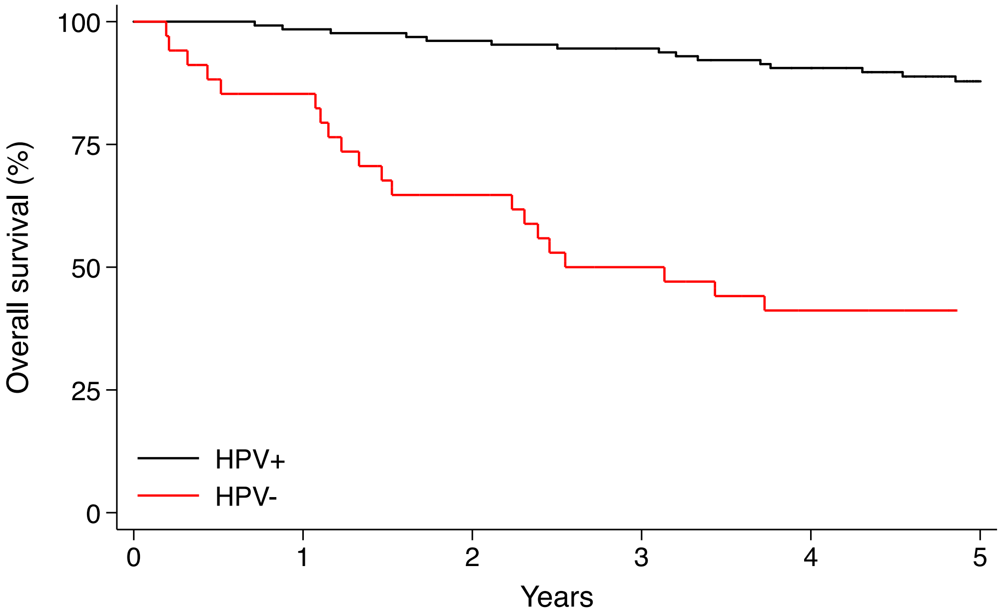
Supplementary figure 4. Top mutated genes in a custom gene panel in the DAHANCA 19 cohort.



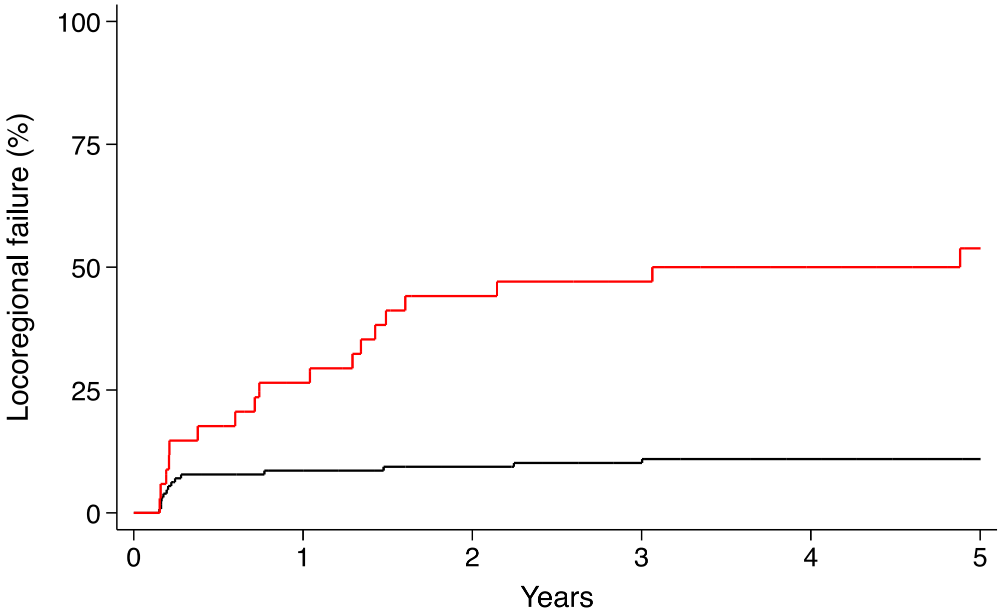
Supplementary Figure 5. Mutations and CNV in the DAHANCA 19 cohort (n=162) characterized with a custom NGS gene panel.



Supplementary figure 6. Overall survival and locoregional failure in the DAHANCA 19 cohort by HPV-status.

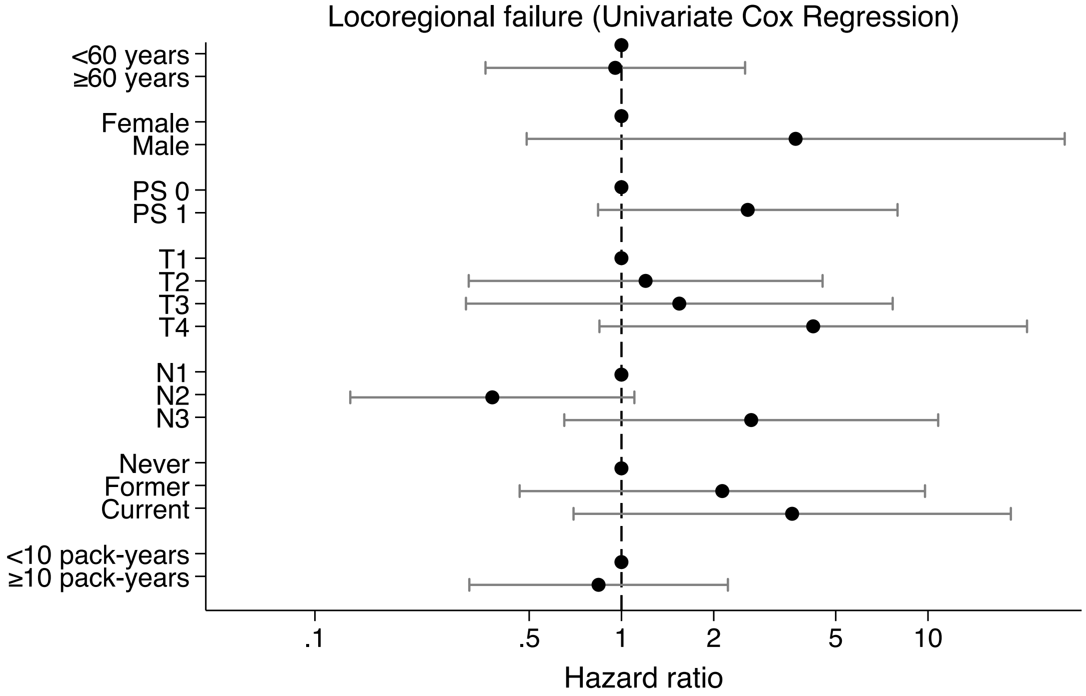


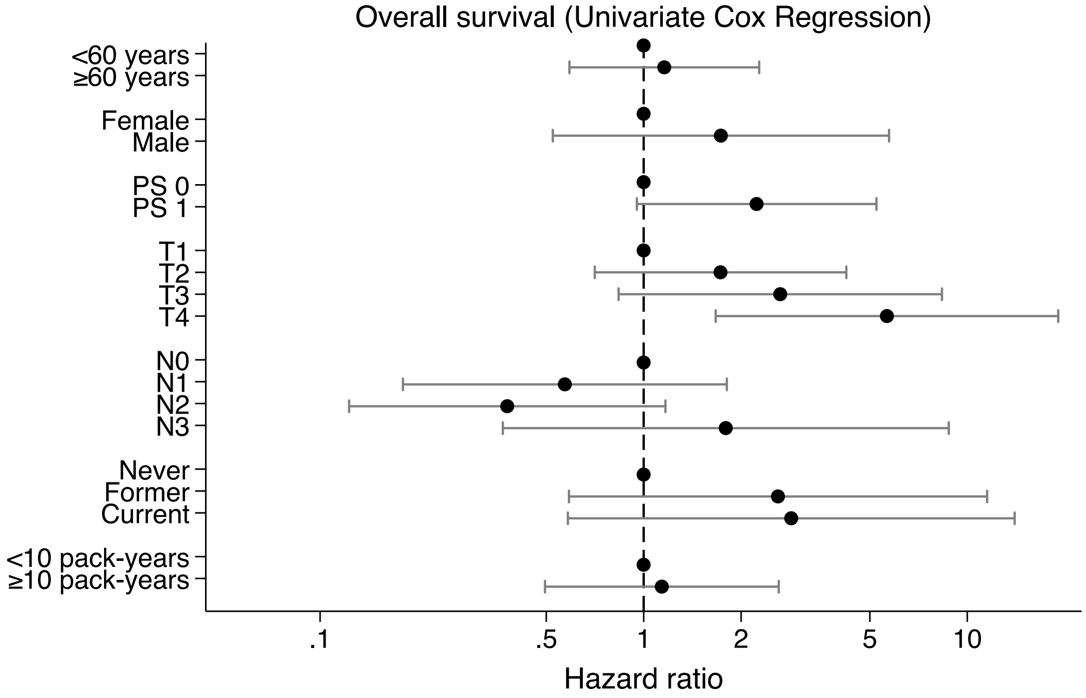
Number at risk						
HPV+	128	126	123	120	112	76
HPV-	34	29	22	17	14	12



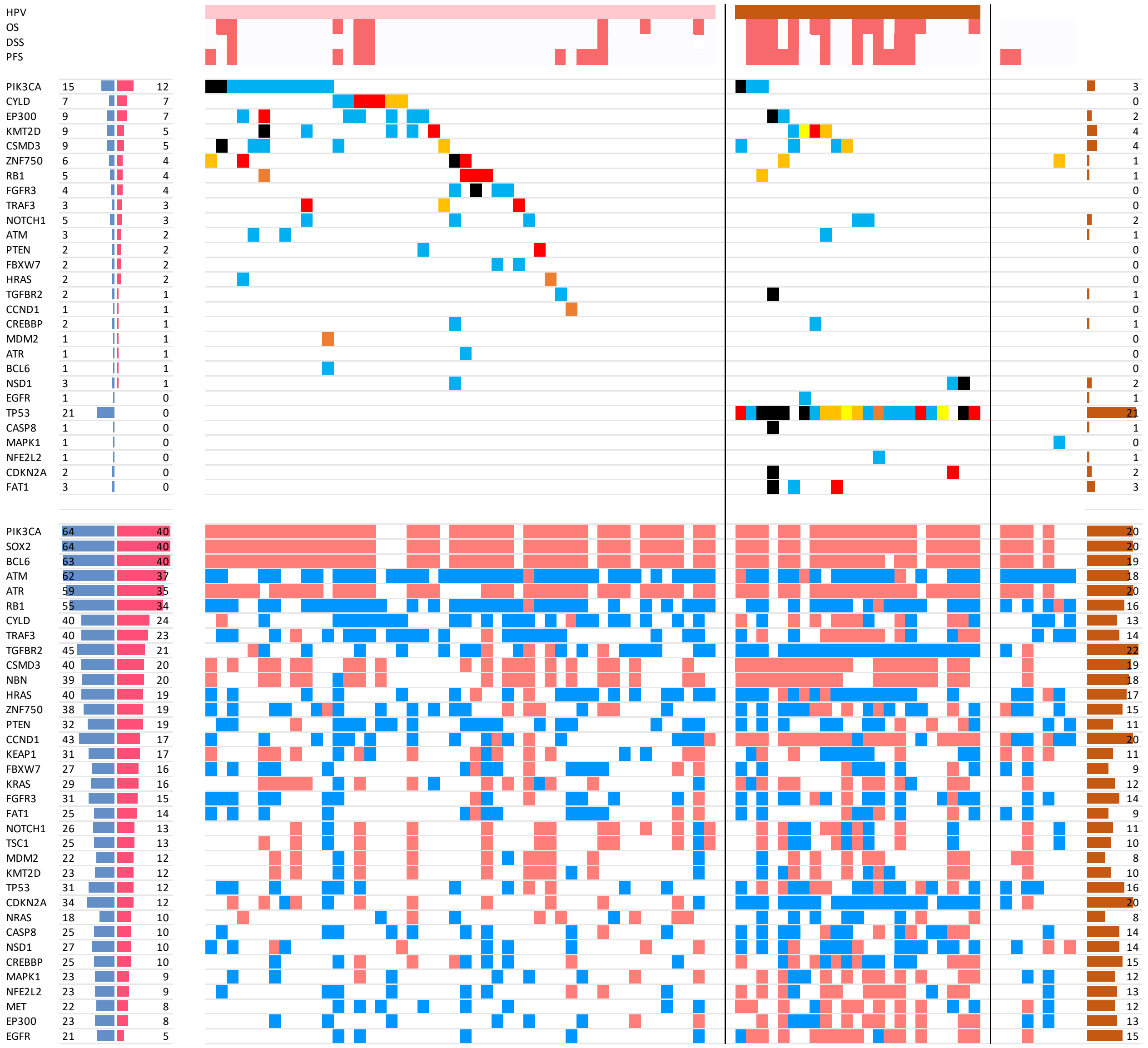
Number at risk						
HPV+	128	117	113	108	104	69
HPV-	34	23	16	15	13	9

Supplementary figure 7. Prognostic factors for locoregional failure and overall survival for HPV+ patients in the DAHANCA 19 cohort.

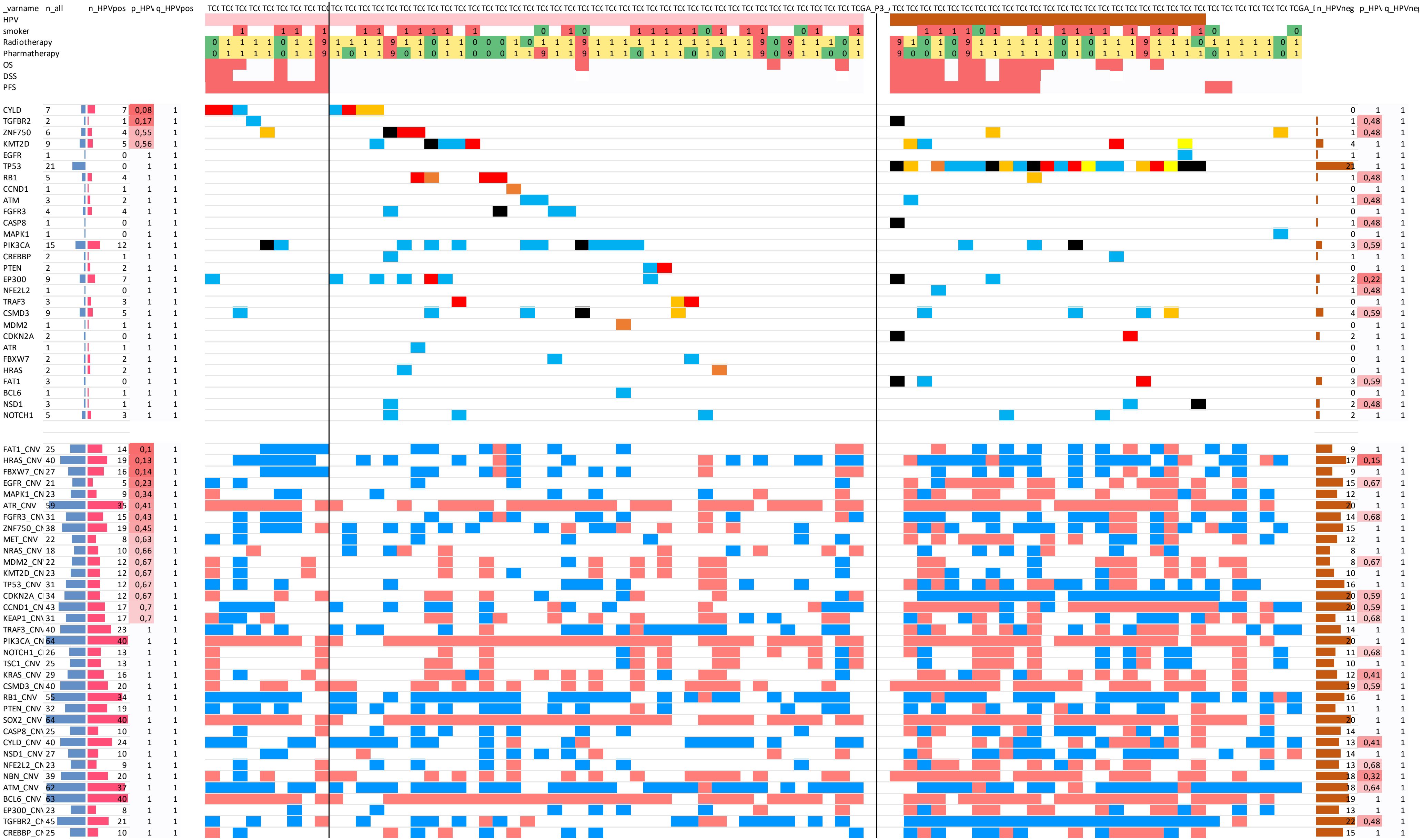




Supplementary figure 8. Distribution of mutations (top) and CNV (bottom) in HNSCC patients in the TCGA Pan Cancer Atlas with oropharynx tumor site, restricted to genes included in our custom gene panel.



Supplementary figure 9. Mutations (top) and CNV (bottom) of OPSCC in the TCGA Pan Cancer Atlas cohort. Cases are ordered by HPV-status and PFS-status. Genes are ordered by p-value for association with PFS within each category.



_varname	n_all	n_HPVpos	p_HPV	q_HPVpos	n_HPVneg	p_HPV	q_HPVne
HPV							
smoker							
Radiotherapy							
Pharmatherapy							
OS							
DSS							
PFS							
CYLD	7	7	0,08	1	0	1	1
TGFBR2	2	1	0,17	1	1	0,48	1
ZNF750	6	4	0,55	1	1	0,48	1
KMT2D	9	5	0,56	1	4	1	1
EGFR	1	0	1	1	1	1	1
TP53	21	0	1	1	21	1	1
RB1	5	4	1	1	1	0,48	1
CCND1	1	1	1	1	0	1	1
ATM	3	2	1	1	1	0,48	1
FGFR3	4	4	1	1	0	1	1
CASP8	1	0	1	1	1	0,48	1
MAPK1	1	0	1	1	0	1	1
PIK3CA	15	12	1	1	3	0,59	1
CREBBP	2	1	1	1	1	1	1
PTEN	2	2	1	1	0	1	1
EP300	9	7	1	1	2	0,22	1
NFE2L2	1	0	1	1	1	0,48	1
TRAF3	3	3	1	1	0	1	1
CSMD3	9	5	1	1	4	0,59	1
MDM2	1	1	1	1	0	1	1
CDKN2A	2	0	1	1	2	1	1
ATR	1	1	1	1	0	1	1
FBXW7	2	2	1	1	0	1	1
HRAS	2	2	1	1	0	1	1
FAT1	3	0	1	1	3	0,59	1
BCL6	1	1	1	1	0	1	1
NSD1	3	1	1	1	2	0,48	1
NOTCH1	5	3	1	1	2	1	1
FAT1_CNV	25	14	0,1	1	9	1	1
HRAS_CNV	40	19	0,13	1	17	0,15	1
FBXW7_CN	27	16	0,14	1	9	1	1
EGFR_CNV	21	5	0,23	1	15	0,67	1
MAPK1_CN	23	9	0,34	1	12	1	1
ATR_CNV	59	35	0,41	1	20	1	1
FGFR3_CN	31	15	0,43	1	14	0,68	1
ZNF750_CN	38	19	0,45	1	15	1	1
MET_CNV	22	8	0,63	1	12	1	1
NRAS_CNV	18	10	0,66	1	8	1	1
MDM2_CN	22	12	0,67	1	8	0,67	1
KMT2D_CN	23	12	0,67	1	10	1	1
TP53_CNV	31	12	0,67	1	16	1	1
CDKN2A_C	34	12	0,67	1	20	0,59	1
CCND1_CN	43	17	0,7	1	20	0,59	1
KEAP1_CN	31	17	0,7	1	11	0,68	1
TRAF3_CN	40	23	1	1	14	1	1
PIK3CA_CN	64	40	1	1	20	1	1
NOTCH1_C	26	13	1	1	11	0,68	1
TSC1_CNV	25	13	1	1	10	1	1
KRAS_CNV	29	16	1	1	12	0,41	1
CSMD3_CN	40	20	1	1	19	0,59	1
RB1_CNV	55	34	1	1	16	1	1
PTEN_CNV	32	19	1	1	11	1	1
SOX2_CNV	64	40	1	1	20	1	1
CASP8_CN	25	10	1	1	14	1	1
CYLD_CNV	40	24	1	1	13	0,41	1
NSD1_CNV	27	10	1	1	14	1	1
NFE2L2_CN	23	9	1	1	13	0,68	1
NBN_CNV	39	20	1	1	18	0,32	1
ATM_CNV	62	37	1	1	18	0,64	1
BCL6_CNV	63	40	1	1	19	1	1
EP300_CN	23	8	1	1	13	1	1
TGFBR2_CN	45	21	1	1	22	0,48	1
CREBBP_CN	25	10	1	1	15	1	1