Supplementary table 1

Sample name & characteristics	Sample type	RNA integrity number (RIN)	RNA concentration in ng/ul (qubit)	Average library size in bp (BioAnalyzer)	Library concentration in nM (qPCR-kappa)
P1	Lin-CD44-	6.7	23	323	78.29
(45F,	Lin-CD44+	7	19.1	317	103.68
P2	Lin-CD44-	7.6	19.2	338	121.46
(50F,	Lin-CD44+	9.5	36.1	364	209.65
P3	Lin-CD44-	8.6	46.6	346	115.4
(51M,	Lin-CD44+	9.5	48.4	358	256.26
P4	Lin-CD44-	5.6	13.7	349	277.35
(45F,	Lin-CD44+	6.6	14.6	348	429.08
P5	Lin-CD44-	5.8	12.85	354	147.16
(56M,	Lin-CD44+	6.9	47.4	360	488.40

Table S1. Sample characteristics- RNA-Seq libraries prepared from five OSCC-GB samples.
The age and sex (M-Male, F-Female) of the patients as well as the TNM staging information are in paranthese

T4aN2aM0)

T4aN1M0)

T4bN0M0)

T4aN2bMx)

T4aN0M0)

Supplementary table 2

Table S2. Lists of differently expressed genes.

Sample name	Down in CD44+Lin- (by log2 fold change<-2)	Up in CD44+Lin- (by log2 fold change>2)
P1	191	1423
P2	734	211
P3	169	800
P4	367	724
P5	1113	104

UP in at least 3 of 5 samples

Gene_Name	P1 foldchange	P2 foldchange	P3 foldchange	P4 foldchange	P5 foldchange
BTK	5.93101	2.02079	2.4142	4.71016	
COL4A4	4.92447	3.12431	2.06157		
MS4A2	9.43339	3.64762		7.64395	
KIT	8.30356		3.6753	4.13157	
CSF2RB	4.09471		3.35254	3.46342	
ITGA2B		3.33485	2.37145	5.57665	
HBB	3.88836	2.22549	10.4965		
FASLG	5.81255	2.18873		5.59586	
ADORA3	3.76016	2.38082		3.04142	
HPGD		3.78475	3.84217	8.06828	
NTRK1	6.10647		2.94736	2.47659	
C10orf128	4.85866		2.6285	3.35816	

ZMAT1	3.46723		2.10337	3.01621
SAMD3	6.55908	2.06289	2.09123	4.44847
LY9	3.9689	2.27045		5.51225
CD37	4.63436	2.12956		3.4545
SKAP1]	5.58524		3.02661	2.79474
ZAP70	5.15372		2.36215	3.82028
ZNF331	3.62127		2.64168	3.0161
ARHGAP9	3.68654		2.71413	2.73489
GPR18	4.07586	3.02772	2.06721	2.85554
TESPA1	5.57327	2.08519		4.48038
PLXNA4	3.03469	3.20177	2.74914	
KLRD1	5.74147	3.76102		5.94717
PLAC8	3.57554	4.021		2.98394
DOK3	2.35096		2.01028	2.40991
SNX20	4.22766		2.20753	2.68167
GATA2	8.00232		6.52888	4.91406
RASGRP4	5.54475		3.18994	4.3145
AFF2	4.94248		2.38511	2.99894
ITM2A	3.6855		3.83074	2.13637
PPP1R16B	4.96984	2.10688		3.82275
TYROBP	2.95074		2.03517	3.20299
IGLL5	4.46683		7.0031	4.27703
CD84	4.90397		3.52142	4.21804
CD22	6.78316		4.23335	3.95401
PTPN22	4.03002		2.23611	3.44852
FCRL5	3.66808	2.19209		3.13156
KLRC4-KLRK1	9.30271	2.14898		4.82455
ALOX5AP	2.86098		2.25083	3.23934
IKZF1	4.38345	2.08531		3.26501
CD48	4.51082	2.26526		2.81106
SAMSN1]	5.182		6.84149	5.31685
CD300A	3.85786		2.38722	2.7868
IKZF3	5.40308		3.29173	4.97827
CYSLTR1	4.83901	2.11526		2.35825
IL1RL1	6.93808	2.47182	3.72859	6.44154
PIK3CG	3.96688		3.09526	3.17571
PLCB2	3.49159		2.46672	2.28595
CD300LF	3.50947		2.74841	3.16376
ACKR2		2.07319	2.44297	4.28313
CABP4	3.28395	2.29648		2.10748
GNLY	8.61903	2.68992		4.96582
CD226	3.46722		2.31539	6.02516
CRTAM	8.44115	3.78569		6.99749
HDC	8.99133	3.53804		8.01355
ARHGEF6	4.68639		2.31545	2.09639
CTSW	7.87426	3.88681		5.97461
DUSP4	4.11507		2.07256	2.65273
CXCR5	2.04416	4.38195		3.39765

CD2	4.91994		2.4718	3.06124	
CD69	5.27996		5.05395	5.10147	
CD52	4.95037	2.26222	3.07622	3.84271	
CPA3	8.38364	5.55042	10.3184	8.50913	
CTSG	7.01336	3.48599		7.91099	
GZMK	4.47325	2.31248		4.77173	
IRF8	3.01328		3.55324	2.53595	
SRGN	3.25708		3.39409	3.58042	
RGS1	5.33496		2.98946	4.35285	
CCL22	4.19754	2.53271		3.08447	
TPSAB1	8.41823	4.97832		8.1063	
GPR65	4.68653		4.35884	2.56074	
FCER1G	3.08404		2.32361	2.83616	
FLT3	4.91391	2.27396		2.06248	
GZMB	5.64027	3.84187	2.22612	5.79394	
CYTIP	3.38031		3.67582	3.46719	
DOCK2	3.63755		2.01633	3.07997	
ATP2A3	5.50429		2.86509	2.2172	
NKG7	5.76996	2.19091		4.61052	
ST8SIA4	2.39974		2.65028	2.45307	
FGL2	2.10005		2.37594	2.38249	
PADI2	2.89562		2.99712	3.1264	
LAT2	4.01565		2.07791	2.99495	
HPGDS	7.7622	4.53425		5.27265	
RNF125	2.32053	2.03023		2.46592	
APOBR	5.12242		2.05545	2.6416	
SASH3	4.50308		2.20296	3.2197	
WDFY4	5.18491	2.97281		2.82592	
TPSB2	8.18641	4.2125		8.09134	
LRRC2	2.50042	2.31072		2.71482	
NUP210	4.93732	2.22413		2.44117	
MYO1G	5.2049		2.25865	2.83231	
FCRL3	4.62694	3.01869		3.3573	
TIGIT	2.75357		2.08147	3.5892	
MORN3	2.03836		2.38328	2.42086	
GPRIN3	2.91702		2.0909	2.75522	
NAPSB	4.03102	5.19871		3.09175	
FRG1JP	2.31708	2.22014		2.92959	
LINC00996		3.88375	2.78058	5.53155	
ANKRD31	2.93136			2.9514	2.25601
CHRDL2	2.14836	2.04413			3.21394
ERVW-1			2.03949	2.58243	2.54622

DOWN in at least 3 of 5 samples

Gene_Name	P1 foldchange	P2 foldchange	P3 foldchange	P4 foldchange	P5 foldchange

A2M	-3.45619	-2.93721	-4.85274
AGT	-2.10069	-4.30597	-4.46127
ARSE	-2.18717	-2.39183	-3.88417
LPL	-3.01453	-2.39739	-5.11802
VWF	-4.2806	-3.50636	-2.74997
AVPR1A	-3.53661	-2.47891	-6.34042
CACNA1D	-3.11382	-2.75968	-3.63737
GUCY1A2	-4.17073	-2.55297	-5.34881
MGP	-2.23089	-3.796	-5.08625
FAM13C	-2.60345	-2.07559	-4.00869
FMO3	-3.60614	-3.72076	-6.23483
PDE1A	-2.79205	-3.13866	-3.03515
VSTM4	-2.96242	-2.01145	-4.78468
TSPAN11	-3.14116	-2.10336	-4.96661
PRKG1	-4.04656	-2.30041	-3.3785
ADGRF5	-4.19167	-2.88409	-5.474
GGT5	-2.56382	-2.10144	-4.43667
ANGPT2	-4.39477	-2.50528	-4.54648
SPARCL1	-2.81571	-2.78732	-3.70744
SPON2	-3.10156	-2.09852	-2.69247
LRRC32	-2.67996	-2.0305	-3.81989
TMEM130	-3.31167	-2.03142	-2.15057
ACAN	-3.61271	-2.88887	-6.00646
NOX4	-3.72457	-2.1348	-5.21825
SLITRK2	-2.88143	-2.24014	-2.36901
C11orf96	-2.06965	-2.26761	-2.96645
ADAMTSL2	-3.98187	-2.63994	-2.81273
PLPPR4	-3.73073	-2.76734	-4.82844
ASPN	-2.44043	-3.14911	-4.79001
RGS5	-3.33076	-2.70194	-3.57137
ECM2	-2.27174	-2.11927	-3.40953
LAMP5	-4.09572	-2.21946	-5.2216
EGFLAM	-3.45145	-2.73236	-5.15263
F2RL2	-4.34593	-2.18119	-5.21063
WFDC1	-2.71384	-2.59278	-3.42227
FMO1	-2.12891	-2.00252	-5.46582
EBF1	-2.21707	-2.02879	-4.02712
OLFML2B	-2.89349	-2.06383	-4.49915
PCDH18	-3.41605	-2.11543	-2.12065
RASL12	-4.06232	-2.40649	-4.25645
SGIP1	-4.45486	-2.01137	-5.68974
F2R	-2.24307	-2.07343	-2.82342
ID4	-4.03092	-2.26382	-4.7293
COX7A1	-2.62941	-2.02614	-3.78235
CPE	-2.93193	-2.09413	-2.07435
FMOD	-2.70991	-2.35388	-4.9283
CCL21	-2.38005	-2.51157	-2.90178
TGFB3	-2.81294	-2.29993	-2.63909

CDH6		-2.48169		-3.1196	-3.69296
KCNJ8		-2.49372		-2.72608	-5.67801
EPHA3		-2.3101		-2.22	-3.39775
MEOX2		-2.40512		-2.18279	-3.76407
MMP16		-3.56711		-2.21236	-3.70089
TBX2		-3.56789		-2.07048	-3.43237
LAMC3		-4.39775		-2.44157	-3.74402
PPEF1		-6.40012		-2.99974	-4.39869
NID2		-2.84612		-2.014	-4.38934
TSPAN12		-2.53422		-3.34952	-4.74112
SHC2		-4.42042		-3.11466	-3.31942
NRXN2		-3.18537		-2.50656	-2.36579
GRID1		-3.00694		-2.25694	-2.69153
SUSD2		-2.52784		-2.28159	-4.12187
PLXDC1		-3.35399		-2.20119	-3.90111
ADGRL4		-2.76411		-2.25047	-5.17162
ESAM		-2.95774		-3.18986	-6.22874
LGI4		-2.4419		-3.0388	-3.09372
CNIH3		-4.21739		-2.87469	-5.10367
COL23A1		-4.70876		-2.31869	-2.38661
TMEM26		-3.42895		-2.12026	-3.25583
BMP8A		-3.4403		-2.09256	-3.041
ADAMTS9		-2.06841		-2.18234	-6.06493
TEX41		-2.54978		-2.09015	-2.13148
PRKG1-AS1		-3.992		-2.10713	-3.22823
LOC100507175		-3.20563		-2.86058	-2.58559
S100P	-2.57979	-3.37899	-2.02156		-3.65017
TEK	-2.49519	-2.04609			-2.60727
PLN	-2.34171	-2.70033			-2.66863
S100A7A	-2.15563	-3.66832			-4.0217
LRRC10B	-2.79143	-2.42811		-2.09311	
CDH5	-2.35047	-2.87489		-3.81996	
KRTDAP	-2.02012	-2.82503	-2.46825		
KLK13	-2.76901	-3.0716	-2.12152		
RNA 5.8S5	-2.03055	-3.4476	-2.28677		
DCLK3		-2.5926	-2.53546	-2.01617	
CD36		-2.54093	-2.48679		-2.28002

Supplementary table 3

Table S3. Total number of mapped reads against HPV16 and HPV18 virus database.

Patient	CD44-Lin-	CD44+Lin-
P1	0	7
P2	0	1
P3	5	7
P4	2	5
P5	2	5