

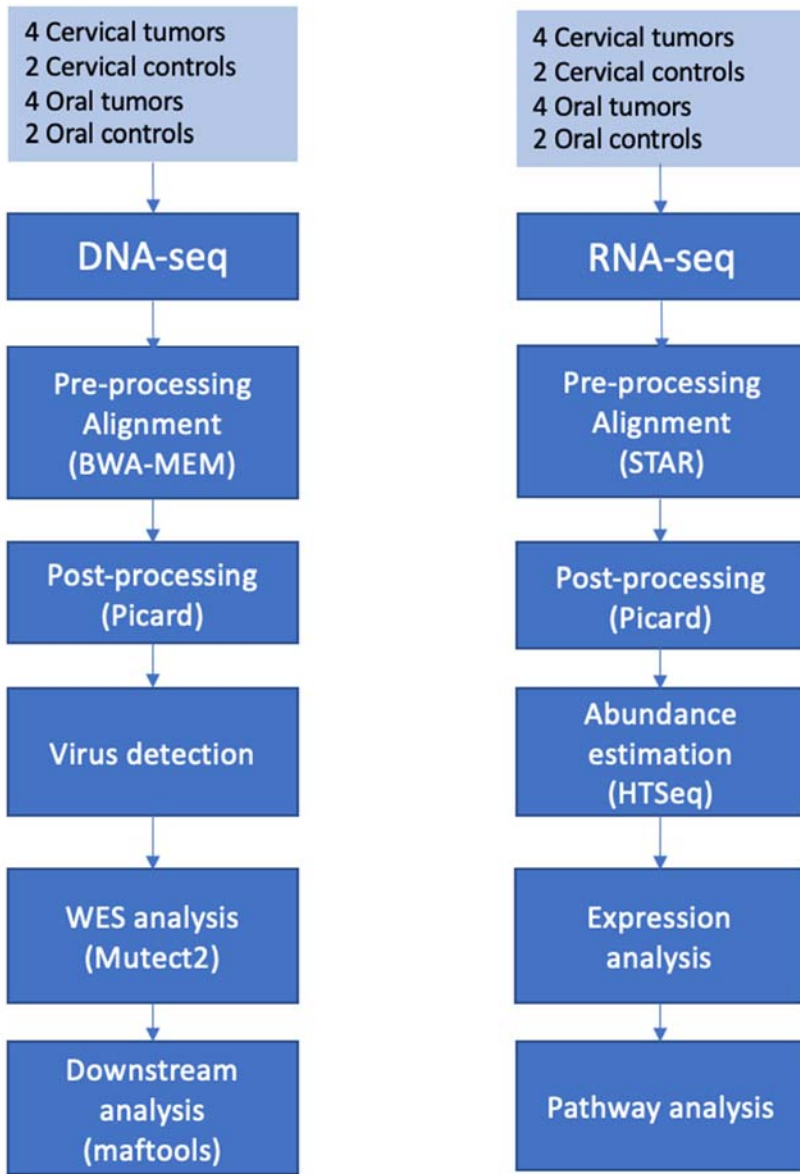
ID	CERVICAL			ORAL		
	summary	Mean	Median	Summary	Mean	Median
NCBI_BUILD	hg38	NA	NA	Hg38	NA	NA
CENTER	NA	NA	NA	NA	NA	NA
SAMPLES	4	NA	NA	4	NA	NA
NGENES	1617	NA	NA	1049	NA	NA
FRAME_SHIFT_DEL	118	29.5	24	27	6.75	4.5
FRAME_SHIFT_INS	5	1.25	1.5	4	1	1
IN_FRAME_DEL	22	5.5	5.5	3	0.75	0.5
MISSENSE_MUTATION	2702	675.5	704	1402	350.5	439
NONSENSE_MUTATION	185	46.25	47	103	25.75	24.5
NONSTOP_MUTATION	5	1.25	1.5	3	750.75	1
SPLICE_SITE	92	23	20	55	13.75	11.5
TOTAL	3129	782.25	815	1597	399.25	497

Supplementary Table 1. Somatic mutation types identified in CC and HNSCC tumor samples. Sum, mean and median somatic rates are shown in separate columns for both cervical and oral cancers.

METRICS	CERVICAL-CONTROL	CERVICAL-TUMOR	ORAL-CONTROL	ORAL-TUMOR
NUMBER OF INPUT READS 	6773956	3189636	6799184	3830650
Average input read length	100	125	104	114
UNIQUE READS:				
UNIQUELY MAPPED READS NUMBER	4911230	2862478	5907695	3007843
Uniquely mapped reads %	72.50%	89.74%	86.89%	78.52%
Average mapped length	95.52	121.67	101.53	111.43
Number of splices: Total	69751	18574	39496	29174
Number of splices: Annotated (sjdb)	15697	2955	558	6557
Number of splices: GT/AG	40899	10709	21532	16951
Number of splices: GC/AG	2940	675	1729	984
Number of splices: AT/AC	100	49	66	38
Number of splices: Non-canonical	25812	7141	16169	11201
Mismatch rate per base, %	1.57%	1.11%	1.26%	1.24%
Deletion rate per base	0.07%	0.06%	0.06%	0.06%
Deletion average length	2	1.77	1.79	1.82
Insertion rate per base	0.02%	0.02%	0.01%	0.01%
Insertion average length	2.07	1.91	1.86	1.94
MULTI-MAPPING READS:				
NUMBER OF READS MAPPED TO MULTIPLE LOCI 	688113	177029	406183	359142
% of reads mapped to multiple loci	10.16%	5.55%	5.97%	9.38%
Number of reads mapped to too many loci	52436	13128	57741	23077
% of reads mapped to too many loci	0.77%	0.41%	0.85%	0.60%
UNMAPPED READS:				
NUMBER OF READS UNMAPPED: TOO MANY MISMATCHES 	0	0	0	0
% of reads unmapped: too many mismatches	0.00%	0.00%	0.00%	0.00%
Number of reads unmapped: too short	1095410	128148	388378	423852
% of reads unmapped: too short	16.17%	4.02%	5.71%	11.06%
Number of reads unmapped: other	26767	8853	39187	16736
% of reads unmapped: other	0.40%	0.28%	0.58%	0.44%
CHIMERIC READS:				
NUMBER OF CHIMERIC READS 	0	0	0	0
% of chimeric reads	0.00%	0.00%	0.00%	0.00%

Supplementary Table 2. Mapping quality shown for the sequencing reads of the CC and HNSCC samples to hg38.

Fresh frozen tumors



Supplementary Figure 1. A schematic showing the work pipeline. The tumors and controls cohorts and next generation sequencing and analysis pipeline. DNA-seq analysis shown on the left and RNA-seq analysis is shown on the right.