

HPV integrations in cervical cancer

Table S1. Summary of 13 cancer samples analyzed in this study: HPV viral status and clinical information

| Sample name | HPV status | Sequencing depth | No. of validated integration sites on the basis of different reads | | | | | Pathology |
|-------------|----------------------------|------------------|--|-----|---|---|-----|--------------------|
| | | | Total | > 3 | 3 | 2 | 1 | |
| S1-T | HPV16, HPV82 | 60X | 46 | 2 | 1 | 3 | 40 | cervical carcinoma |
| S2-T | HPV16, HPV31, HPV82 | 60X | 37 | 5 | - | 2 | 30 | cervical carcinoma |
| S3-T | HPV16, HPV82 | 60X | 396 | 18 | 2 | 5 | 371 | cervical carcinoma |
| S4-T | HPV16, HPV31, HPV45, HPV82 | 60X | 85 | 1 | 1 | 6 | 77 | cervical carcinoma |
| S5-T | HPV82 | 60X | 28 | 2 | 1 | 1 | 24 | cervical carcinoma |
| S6-T | HPV16, HPV82 | 60X | 23 | - | 1 | 2 | 20 | cervical carcinoma |
| S7-T | HPV16, HPV82 | 60X | 32 | - | - | 1 | 31 | cervical carcinoma |
| S8-T | HPV16, HPV82 | 60X | 104 | 2 | - | - | 102 | cervical carcinoma |
| S9-T | HPV16, HPV82 | 60X | 26 | - | 1 | 4 | 21 | cervical carcinoma |
| S10-T | HPV82 | 60X | 15 | - | 1 | 1 | 13 | cervical carcinoma |
| S11-T | HPV82 | 60X | 21 | 1 | 1 | - | 19 | cervical carcinoma |
| S12-T | HPV58, HPV82 | 60X | 30 | 4 | - | 1 | 25 | cervical carcinoma |
| S13-T | HPV16, HPV82 | 60X | 27 | 3 | - | 3 | 21 | cervical carcinoma |

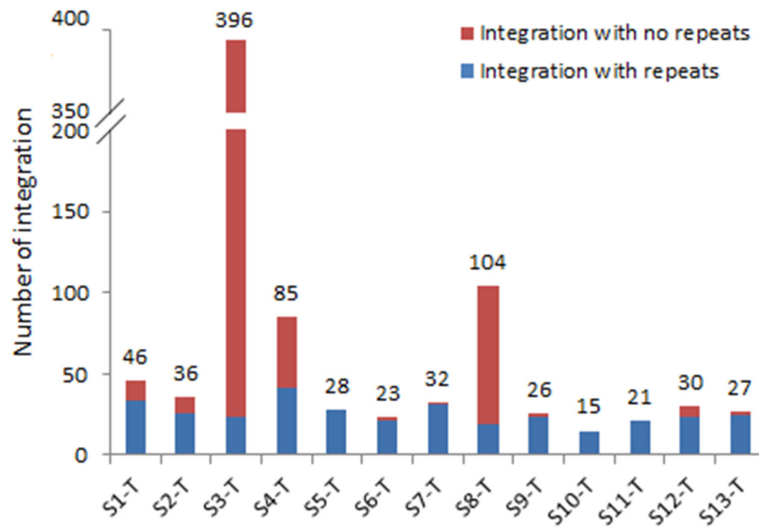


Figure S1. Distribution of integration in 13 samples.

HPV integrations in cervical cancer

Table S2. Production of whole-genome paired-end sequencing data

| Sample ID | Read length | Insert size | Reads PF*, M | Yield**, Gb | ≥ Q30 (%)** | # of HPV integrations |
|-----------|-------------|-------------|--------------|-------------|-------------|-----------------------|
| S1-T | 150 | 350 | 1820.68 | 273.1 | 91.7 | 46 |
| S2-T | 150 | 350 | 1626.75 | 244.01 | 91.7 | 37 |
| S3-T | 150 | 350 | 1821.71 | 273.26 | 92.38 | 396 |
| S4-T | 150 | 350 | 1695.22 | 254.28 | 92.27 | 85 |
| S5-T | 150 | 350 | 1538.42 | 230.76 | 95.53 | 28 |
| S6-T | 150 | 350 | 1580.1 | 237.01 | 95.32 | 23 |
| S7-T | 150 | 350 | 1590.99 | 238.65 | 95.42 | 32 |
| S8-T | 150 | 350 | 1593.53 | 239.03 | 95.34 | 103 |
| S9-T | 150 | 350 | 1589.14 | 238.37 | 95.02 | 26 |
| S10-T | 150 | 350 | 1571.88 | 235.78 | 94.93 | 15 |
| S11-T | 150 | 350 | 1354.19 | 203.13 | 95.37 | 21 |
| S12-T | 150 | 350 | 1543.42 | 231.51 | 94.94 | 30 |
| S13-T | 150 | 350 | 1743.54 | 261.53 | 94.94 | 27 |

1. Data separated by semicolon are corresponding to No. 1 end and No. 2 end of PE reads, respectively.

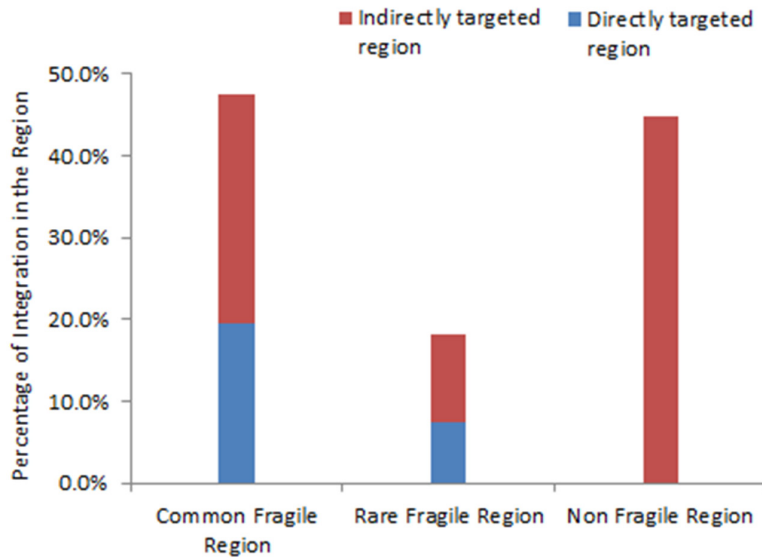


Figure S2. Distribution of integrations in human fragile regions.

HPV integrations in cervical cancer

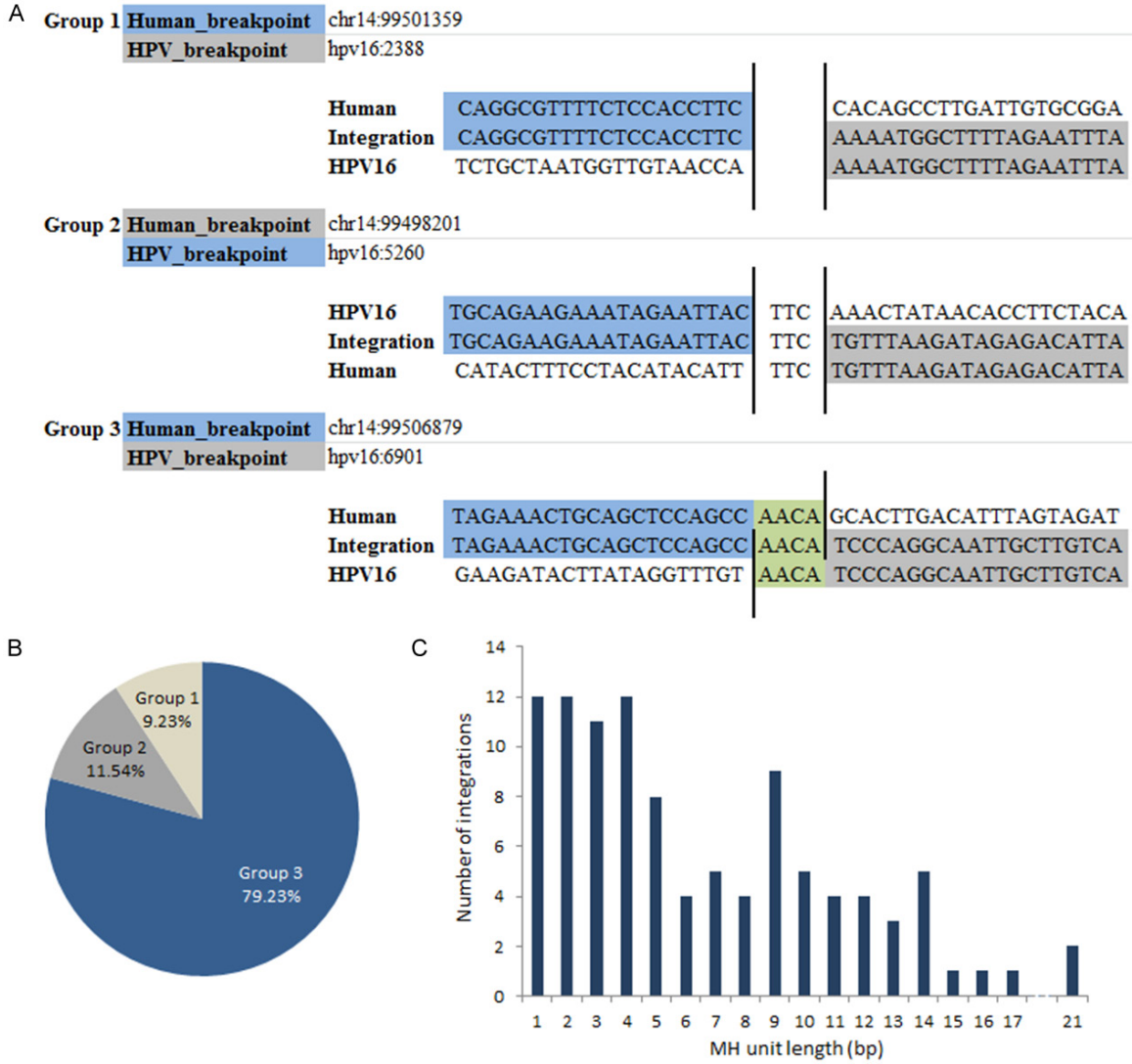


Figure S3. Sequence alignment around the integration site between the human genome and HPV genome.

HPV integrations in cervical cancer

Table S6. Viral breakpoints collection

| ORF | Breakpoint | Raw Count | Total Count | Band leghth | Density'-score |
|-----|------------|-----------|-------------|-------------|----------------|
| L1 | L1 | 96 | 109 | 296 | 29.1133 |
| | L2/L1 | 13 | | | |
| E1 | E1 | 132 | 151 | 476 | 25.0800 |
| | E1/E2 | 13 | | | |
| | E7/E1 | 6 | | | |
| E5 | E5 | 21 | 27 | 251 | 8.5045 |
| | E2/E5 | 5 | | | |
| | E5/L2 | 1 | | | |
| E4 | E4 | 29 | 29 | 287 | 7.9886 |
| E2 | E2 | 44 | 91 | 1097 | 6.5583 |
| | E1/E2 | 13 | | | |
| | E2/E5 | 5 | | | |
| | E4 | 29 | | | |
| LCR | LCR | 56 | 56 | 837 | 5.2896 |
| L2 | L2 | 83 | 97 | 1948 | 3.9368 |
| | L2/L1 | 13 | | | |
| | E5/L2 | 1 | | | |
| E7 | E7 | 14 | 25 | 1421 | 1.3909 |
| | E7/E1 | 6 | | | |
| | E6/E7 | 5 | | | |
| E6 | E6 | 18 | 23 | 1595 | 1.1401 |
| | E6/E7 | 5 | | | |

1. Raw count: N of intregation sites. 2. Density'-score: Total count/band leghth*7906/100.

HPV integrations in cervical cancer

Table S9. KEGG and GO pathways enriched by the ITGs

| Category | Term | Count | % | P Value | Genes | List Total | Pop Hits | Pop Total | Fold Enrichment | Bonferroni | Benjamini | FDR |
|------------------|--|-------|-------------|-------------|---|------------|----------|-----------|-----------------|------------|-----------|------------|
| GOTERM_MF_DIRECT | GO:0043565~sequence-specific DNA binding | 4 | 2.105263158 | 0.729338219 | ENSG00000057935, ENSG00000197587, ENSG00000166949, ENSG00000069667 | 119 | 519 | 16313 | 1.056524344 | 1 | 0.9999941 | 99.9999964 |
| GOTERM_BP_DIRECT | GO:0006955~immune response | 4 | 2.105263158 | 0.629988083 | ENSG00000166949, ENSG00000111536, ENSG00000111537, ENSG00000127318 | 130 | 420 | 16787 | 1.22981685 | 1 | 1 | 99.9999685 |
| GOTERM_MF_DIRECT | GO:0004674~protein serine/threonine kinase activity | 4 | 2.105263158 | 0.515670228 | ENSG00000010219, ENSG00000127334, ENSG00000079277, ENSG00000070759 | 119 | 377 | 16313 | 1.454472505 | 1 | 0.9999007 | 99.9926316 |
| GOTERM_MF_DIRECT | GO:0004674~protein serine/threonine kinase activity | 4 | 2.105263158 | 0.515670228 | ENSG00000010219, ENSG00000127334, ENSG00000079277, ENSG00000070759 | 119 | 377 | 16313 | 1.454472505 | 1 | 0.9999007 | 99.9926316 |
| GOTERM_BP_DIRECT | GO:0006468~protein phosphorylation | 5 | 2.631578947 | 0.467844366 | ENSG00000127334, ENSG00000106123, ENSG00000079277, ENSG00000070759, ENSG00000185974 | 130 | 457 | 16787 | 1.412809291 | 1 | 1 | 99.9925036 |
| GOTERM_MF_DIRECT | GO:0003700~transcription factor activity, sequence-specific DNA binding | 9 | 4.736842105 | 0.394827961 | ENSG00000057935, ENSG00000197587, ENSG00000166949, ENSG00000128604, ENSG00000069667, ENSG00000188786, ENSG00000066827, ENSG00000061337, ENSG00000030419 | 119 | 962 | 16313 | 1.282490959 | 1 | 0.9994185 | 99.8628847 |
| GOTERM_MF_DIRECT | GO:0005125~cytokine activity | 3 | 1.578947368 | 0.364482994 | ENSG00000111536, ENSG00000111537, ENSG00000127318 | 119 | 176 | 16313 | 2.336659664 | 1 | 0.9994816 | 99.7393981 |
| GOTERM_MF_DIRECT | GO:0000977~RNA polymerase II regulatory region sequence-specific DNA binding | 4 | 2.105263158 | 0.187376422 | ENSG00000197587, ENSG00000166949, ENSG00000069667, ENSG00000066827 | 119 | 206 | 16313 | 2.661825895 | 1 | 0.9978781 | 93.4344955 |
| KEGG_PATHWAY | hsa04512:ECM-receptor interaction | 3 | 1.578947368 | 0.143329244 | ENSG00000110799, ENSG00000134871, ENSG00000186340 | 54 | 87 | 6910 | 4.412515964 | 1 | 0.894787 | 83.5366551 |
| KEGG_PATHWAY | hsa04151:PI3K-Akt signaling pathway | 6 | 3.157894737 | 0.123304931 | ENSG00000080824, ENSG00000110799, ENSG00000134871, ENSG00000037280, ENSG00000186340, ENSG00000135930 | 54 | 345 | 6910 | 2.225442834 | 0.99999997 | 0.9147976 | 78.4456174 |

HPV integrations in cervical cancer

| | | | | | | | | | | | | |
|------------------|--|---|-------------|-------------|---|-----|-----|-------|-------------|------------|-----------|------------|
| GOTERM_MF_DIRECT | GO:0004672~protein kinase activity | 6 | 3.157894737 | 0.119180469 | ENSG00000010219, ENSG000000106123, ENSG000000183317, ENSG000000079277, ENSG000000070759, ENSG000000185974 | 119 | 359 | 16313 | 2.291098055 | 1 | 0.9920963 | 81.0930112 |
| KEGG_PATHWAY | hsa04060:Cytokine-cytokine receptor interaction | 5 | 2.631578947 | 0.09893686 | ENSG000000111536, ENSG000000111537, ENSG000000124334, ENSG000000037280, ENSG000000127318 | 54 | 230 | 6910 | 2.781803543 | 0.99999882 | 0.8971615 | 70.3254846 |
| GOTERM_CC_DIRECT | GO:0031012~extracellular matrix | 6 | 3.157894737 | 0.086093889 | ENSG000000080824, ENSG000000182492, ENSG000000110799, ENSG000000134871, ENSG000000197102, ENSG000000186340 | 143 | 300 | 18202 | 2.545734266 | 0.99999993 | 0.9629321 | 67.1090251 |
| KEGG_PATHWAY | hsa04510:Focal adhesion | 5 | 2.631578947 | 0.07241287 | ENSG000000110799, ENSG000000128591, ENSG000000134871, ENSG000000037280, ENSG000000186340 | 54 | 206 | 6910 | 3.105897159 | 0.9999471 | 0.9624604 | 58.3791417 |
| GOTERM_MF_DIRECT | GO:0004713~protein tyrosine kinase activity | 4 | 2.105263158 | 0.072032227 | ENSG000000080824, ENSG00000010219, ENSG000000127334, ENSG000000070759 | 119 | 133 | 16313 | 4.122828079 | 1 | 0.9987105 | 62.5149376 |
| GOTERM_MF_DIRECT | GO:0004712~protein serine/threonine/tyrosine kinase activity | 3 | 1.578947368 | 0.018545096 | ENSG00000010219, ENSG000000127334, ENSG000000070759 | 119 | 29 | 16313 | 14.18110693 | 0.99324876 | 0.9932488 | 21.7840035 |
| GOTERM_BP_DIRECT | GO:0018108~peptidyl-tyrosine phosphorylation | 6 | 3.157894737 | 0.006740315 | ENSG000000080824, ENSG00000010219, ENSG000000127334, ENSG000000037280, ENSG000000183317, ENSG000000070759 | 130 | 154 | 16787 | 5.031068931 | 0.9907218 | 0.9907218 | 9.68216959 |
