|  |
| --- |
|  |

**Table S3a. Sequences of site of viral integration with high score**

|  |  |  |
| --- | --- | --- |
| **Sample ID** | **Chromosomal loci** | **Sequence** |
| CT 700 | 1p | CTTTTCGGTTGTGCGTACAAGCACACACGTAGACTTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGATCCCCTATTGGACGAACCCTCCTTTACCCCTGACTGACTCACCAGTCCCTAATGCTGGTCAGATGTCCTGAACAAGGAAAGCCGTGATTTCTGCGACCAGAGAGAATCGGTCAGTCACGGTGTCCCTCAGCCAAAATGAACTCATCCTATTGTCTCCGGAGAGCTGGTAAATCTGCAGCTGAAAAACATACAGTTCTGTGTTTCAAAATGCAAAAACGCACATTTCCCTGACCTCGTGCTCCTGACATTCCCAACCAATGACTATAAATCTGCTCCTGGATTGTGAATTGAAATTAAACCAA |
| CT 702 | 1p | CTTCGGGTTGTGCGTACAAGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGATCCCCAGTTGGAGGAACCATCCTTTACAGCTGACTGACTCACCAGTCCCTAATGCTGGTCAGATGTCCTGAACAAGGAAAGCCGTGATTTCTGCGACCAGAGAGAATCGGTCAGTCACGGTGTCCCTCAGCCAAAATGAACTCAGCCTATTGTCTCCGGAGAGCTGGTAAATCGGCAGCTGAAGAAACATACAGTTCTGTGTTTCAAAATGCAAAAGACCACATTTCCCTGACCTCGGGTGCTCGCTGACACATTCCCAACCAATGACTATAAATCTGCTCCTGGATTGTGGTAAGTTGAGAAATTAAACACAAATATCCCATCTGATAAACAA |
| CT 709 | 1p | TKTTKSGWAAGCAMACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCMTCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTMCCAATGGGGAAGARGGTTCKGTTTGTAATGAYTGACTCACCAKTCCCTRMTGCTGGKCAAATGTCCTACACAAGGAAAGACGGACTTTATTCCTTATATAAAMTATGCRWCGKATTTCGTGTCCCTCARCCAAAATGAACTCWTCCTATTGTCTCCGGAGAATGGTAAATCGGYAGCTGAAKAAACWTACAKTTCTGTGTTTCCCCTGTAAAASAYGRCATTTCCCTGACCTCGGGTGCTCGCTGACGCATTCCCTYCCAATGAMTACCACTCTGCTCCTGGATTGTGGCCCSTTGASA |
| CT 712 | 1p | CTTCGGTTGTGCGTACAAGCACACAAACGGTATACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAG TCCCCCCAGCACCGAGTTTCCTCCAGCAGATTCGGGGCTCCCTCCTGCCAGCTTCTGCTCAGATCATGATTGGCGCCTATCCCACAGATGGTGATCAGAAGACCTATTGTGTGCCCGACACTGGATGAGGCATTTTATAGTCAGGGACAACATGACCGCTTGACTCGCCTAAAAAAGAAAAAAAAAAGTTCTAAGCCTAGGATCTGGGGCGTGAGGTGTGTGTTATGGGAGGAGGCTGGTAGGGGAGGTGGTTCCTTTATTATGTTATTTGGAGGAAAAAAGTCGGGGAATACTATATATTTTGCGTGAACCCCCCCTTTTAAATTGGGTGGGGCTCAAAAA |
| CT 809 | 1p | AKKSSGTWMAAGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTGTATTARCTGTCAAAAGCCACTGTGTCCTGAARAAAAGCAAAGACATCTGGACAAAARGMAAAGATTCCATAATATAAGGGGTCGGTGGACCGGTCGATGTATGTCTTGTTGCAGATCATCAAGAACACGTAGAGAAACCCAGCTGTAATCATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTTGCAACCAGAGACAACTGATCTCTACTGTTATGAGCAATTAAATGACAGCTCAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAAYCGGACAGAGCCCATTACAATCTCCAGCATTCCCCCTCAAACCTAAAAAAAAA |
| CT 866 | 1p | CCTTCGGTTGTGCGTACAAGCACACACGTAGACCTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTGTATTAACTGTCAAAAGCCACTGTGTCCTGAAGAAAAGCAAAGACATCTGGACAAAAAGCAAAGATTCCATAATATAAGGGGTCGGTGGACCGGTCGATGTATGTCTTGTTGCAGATCATCAAGAACACGTAGAGAAACCCAGCTGTAATCATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTTGCAACCAGAGACAACTGATCTCTACTGTTATGAGCAATTAAATGACAGCTCAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAATCGGACAGAGCCCATTACAATCTCCAGCATTCCCCCTCAAACCTAAAAAAAA |
| CT 1138 | 1p | CTTTCGGATTGTGCGTACAAGCACACACGTAGCCATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTGTATTAACTGTCAAAAGCCACTGTGTCCTGAAGAAAAGCAAAGACATCTGGACAAAAAGCAAAGATTCCATAATATAAGGGGTCGGTGGACCGGTCGATGTATGTCTTGTTGCAGATCATCAAGAACACGTAGAGAAACCCAGCTGTAATCATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTTGCAACCAGAGACAACTGATCTCTACTGTTATGAGCAATTAAATGACAGCTCAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAATCGGACAGAGCCCATTACAATCTCCAGCATTCCCCCTCAAACCTAAAAAAAAA |
| CT 864 | 1q | CTTCGTTGTGCGTACAAGCACACACGTATCCATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCGAGGACTAACATGGCTGATCCTGCAAGTGTATCAACTGTCAACCACCTGAGTGTTGTGATAATTCGCTTTACACCCTGGTGATATCAACAGCCCCCTAAGCTTTCGGGGCAACTGGTAGATGAAAATGTATGCCCCGATCTGCTCTCCTCCAACGTGCAAACCTACTCTAAATCCCAAATCACTACCCAGGAGGGTCCTTTCTATGACAATGGATTTGATTTAAAAAATGATTTTCTGGAGTTGCACCGGAAAACCTCACTAGCCCCACACGCCCACTTCCACAGCCCACAGGTGCAGGTTTACCTCTATACATACGAATATTTCATACATGCTTCTCCCCACTAACCTAGAAGCCCCAGCAGAAGTGAGGACCATGTTTCCTTGTGCATGCCACTGTAAACCCAGGGCCCAGCCCAGAATGAGGGCTCAACAGATGTTTAAGAAGGAGTGAAGATTTAGAGAAATCGAAAAAATTTTAGATCTTACAAGGCCCTTGCATCTAAACACTCCTCCAATTATGGCTAAAAAATGAGGAGCTATAGAAAACTGTCCTATTTAAATTTTGTTTCTAAAAACATTCATTTTGGCTTCTTTTAAAAAAGTTACTCAATTACTGAACAGATCACTGGGTGTGGCCTGTTGAACACAGTGCTGTAACTTCCTGCACAGAGACAAGTCTGAAAGTTTTGCTTTCCGCTGTGAAAGCATTGGTGACTGCTTGACAGACTCGCTTGAAGCAATAATGTGTTTTTGGTTAAAAAAAAAAAAACAGAGTCATCACAAAAAAAAAAAAAATTTTTTTTTACACAATTGGAGGCGAACACG |
| CT 755 | 3q | CTTCGGTTGTGCGTACAAGCACACACGTCGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGATGGAGTCTCCATGGTGCAATCATGGCTCACTGCAGCCTCAAACTCCTGGGCTCAAGCAATCCTCCCACCTCAGCCTCCCAGGTAGCTAGGATTATAGGCATGCATTTTTTATAGAGATTAAGTCTTGCAATGTTGCCCAGGCTGTTCTTAAACTCCTGGCCTCAAGCAATCCTCCTGCCCTGGCCTCTCAAAGCTCTGGGATTACAGGCATGATCCACTGCGCTCAGCCCTCAGCCTTATAAACCAGAATCCTCAAGTCAACCTTTCATTATTACCCCACCCCTATCCAATTCCCCAACCCCACAAGCCTAAGAATTGTGCCAGCTTGTGTTACAATTATAAGTGGAAAAAAATA |
| CT 785 | 3q | CTTCGTTGTGCGTACAAAGCAACACGTCCACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGATGGAGTCTCCATGGTGCACTCATGGCTCACTGCACCCTCAAACTCCTGGGCTCAAGCAATCCTCCCACCTCAGCCTCCCAGGTAGCTAGGATTATAGGCATGCATTTTTTATAGAGATTAAGTCTTGCAATGTTGCCCAGGCTGTTCTTAAACTCCTGGCCTCAAGCAATCCTCCTGCCCTGGCCTCTCAAAGCTCTGGGATTACAGGCATGATCCACTGCGCTCAGCCCTCAGCCTTATAAACCAAATCCTCAAGTCAACCTTTCATTATTACCCCACCCCTATCCAATTCCCCAACCCCACAAGCCTAAAATTGTGCCAGCTTGTGTTACAATTATAATGGAAAAAAAATAATATT |
| CT 723 | 3q | CGGKKKKKCGWMAASCACMCACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCAATTAAACTCCAAATGGTACTGCTGSMAGAASAACCCCTGRTGRCCCCRCTCTTCCTCCAAKACTCAATACATCCWCCKGAGAAGGASGMMCTGCTGCTCCTCTGCTGGCTCCTTTTTTTCAGGAAGAGGTGRCMAMRAATAATCRTCSWCCACCCCCCCMYARCAGTTRTTTTTTCTGCTCCTCAKGGGGGAAGAAATACWAGTGGTGCTTCTTAAAATTTTTTTATTARGAAAAAAAAGATAMAARGAGTACTTGTTGAKGMKTTTTTTTCTGTTAAAAAAAACCAACCAASAATTTCTTTTCTTAYAAAAAAAAAACTTGTTGGGSCGGCGGGSCAAGCTTTTATAATACRATAATGCCSCCCGTAAAAAATGGTCCCCCCAAWATGGCKACTCCCGTCTTCTTCTTGCCACGTGGGGTGACTGTCGTGGCGGCCACCTCAAAATAACACGTTGTSWTAACATCATGGCGACCAGTATTTTTTTCTTAATAAACAACTGAGGGAGG |
| CT 706 | 3q | TGGGKTKSSGWMAARGCACACACGTAGACATCCGTACGTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCAGCGACRAAGTATCCTCTGCTGAAATTACTAGGCAGCACTTGGCCAACCACTCCGCCGCGACCCATACCAAAGCCGTCGCCTTGGGCACCAAAGAAACACAGACGACTATCCAGCGACCAAGATCAGAGCCAGACACCGGAAACCCCTGCCACACCAACAAGTTGTTGCACAGAGACTCCGTGGACAGTGCTCCAATCCTCACTGCAGTTAACAGCTCACACAAAGGACGGAAGTGTGAAATGTGTCACTCACATAAAAGCCCCTGTTTAAAGAAAGGGKGGGACGGGGGAGTGCAGAGGAGTAACTATTTAGGGGTGATTTAAATATTTAGGTCTCCTTTTTTTCTCCTTCAAATTAATAMAAAGGCCTTCGKATATAGTTAARAGAKTTTGAAA |
| CT 892 | 3q | CTTCGGTGTGCGTACAGCACCACGTAGACATTCGTCCTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGATGAATAAAGCACGGAGAACCCGAGGGAGTGTGAAGATCACATCAGGTGTCTTTCGGTCAGTGGAGGAAACACCCGGAAGCAGTGAAACGCACTGGGCTTATTGCAACAACCCCAGCTACTGAAAATGCCACCCTCTCCACGACGAAGCACCGCCAGTCCCAAACTCCAATCAGCCGCTGGGCTGGGAGGCCGAAAAGGCAAGAAGACACAGAGCGGGGTTGAAATGAAGAAGGAACTCTATCTCAACGAGACTCATTACAGTATGCAAAACCTAAAAAACTGTACCCTGTGTTGTAATGAGAATATGACTGCATGGGAAATGTGGGTTACAACTCTTGACTTTACCTGCTTTTCACATATTTTCACATAGTGAGTAAAATCGCTGTGGGGAGGAAATGGGCAATGTGGAAATTTCTTCTTATTCCTCCCCTCCCAACTTTCACAACCCTCAGTGTCTATCTCCCCTCTTTGGTGTTCTCCCTCGGTCCGATGGGGTCTTGTGTCCCTGCGGTGCTCTCAAGATCGTTTATTTCGATGAAAAGACAAAAAAAAAAATCACTCTACATCCTACCTGTGAG |
| CT 1210 | 3q | CTTTCGGTTGTGCGTACAAGCACACACGTAGATCCTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTATCCATGCTTGTTAGCCAAGGGGAAGTACAATGGGGCAATCTCCGCTCACCACAACCTCCGCCTCCCAGGTTCAAGCGATTCCCCTGCCTCAGCCTCCCAACTAGTTAGTATTACGGGCATGCGCCACCACGCCTGGCTAATTTTGTATTTTCATTAGAGACGGGGTTTCTCTGTGCTGGTCAGGCTGGTCTCAAACTCCCGACCTCAGGTGATCTGCCTGCTTCAGCCTCCCAAATTGCTGGGATAACAGGTGTGAGGCACCGTGCCTGGCCCTAAATGGCTTTTCTACACGTCAGTTTCTTCAATACATTTGGATTACAAAATTAATGGAAATATTTCTCTTGACGCTTGAACAGGCCAAAAACCCCCCTTTTT |
| CT 711 | 3q | GTTATATAAAAACCAAATACCTTCGGTTGTGCGTAACAAGCCACCGTAGACATTCGTACTATATAGGAAGACCTGTTAATGGGTCCCACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGTCTGGTTATGAAGGTCGAGTCCCACTGAAGAGCGCTCGTCTGTTTTATGACATCCGCCAAAACGCTCAAAGGATTGTACAATCTTATTTTATGCTAAACTCAAATCAGTATTTTTCCTATACACACATGGTCTGCCAAACAACCCTGTCTGGTCACCAGGATATAATAAATGACCTCCTCCATCCCATCCATGCTGACAACTGTTTGTTGTATCCACACGCCAACAAATGCTATTAGGAGCCTCCTGCTTACTCCTTTCATAACTATATTGCTCTCCTATATATGAATGATGACTTTGAAGGAGGAGAATTCATATTCACAGAGATGGATGCTAAGACTGTGACTGCCTCTATAACATTGAACTGGACATGCTGGATCAATTTCTCATCTGGAGGAGAGAACCCTCATGGGGTGAAGGGAGTCACCAAAGGAAATTATGTCTGCTGTGGCTCTGTGGTTCACCTTGACCCCTTTATAAAAATTAAAAAA |
| CT 999 | 4q | CTTCGGTTGTGCGTACAAGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGCCAAGATGACAACGAGCCCAGCTGAAGCTGACATCCCAGCAAATTGCATGACAAATTGCAAAGACGACTAACCCACAACCTACTCTTCTGGAAAATACAATTTAAATAAAATAATTTTAAGTGAAAAAAAAAAAAAAACCCCCTGTCCCAAAAAAAAAAAAAATTTTTTTTCTA |
| CT 1114 | 4q | CTTCGGTTGTGCGTACAAGCACACACGTAGATCTTCGTACTTTGGAAGTTTTTTTTTTTTTTTTTATACAAAAGGATTTATTAAAAAACCAGTAAGACACTACTACATCATGACACTGTCACACTGGGCTTTTAACACAAGACTTGCTCTACAATACTGGGGGAAAGGGCATAAAACACAAATTGATTCTGAAGCATAGCAATTAAGAAATAAAACAATGAAAGCAAATTTCTTTTAATGAGAACTCAGAATTAAACTTCAGAGGGACCCAACGTCATACTTCCATTCAGGGACTTGATACAAAAAATTTAGTTTGAACTGCTATTAGCAGGTGGCAGGAGCCACCTTCAAATGAATCTTCAAATTGGAAAATACTGCTTCACCACCTGTTGGGGATAAGTTGCAAATGGAATAATTTAGTATGGTTTGTAGCTATTTTGATGACCACCTCGCCTGGATACCTTCCCATAACCAGTGCGGTACCACCCAATCGGATCCCGGGCCCGTCGACTGCAGAGGCCTGCATGCAAGCTTTCCCTATAGTGAGTCGTATTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAACCGTAAAAAGGCGC |
| CT 1122 | 6q | CCTTCGGGTTGTGCGTACAAGCACACACGTCGACATCCGTACAGTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGCTCAACTGGTGGAAACGGTATGCAAGCGGGGGCAAATCATGAGGAAAAAACTCTCTCACCATGCATAGAATCATCTTCACCCCATTCTCAAGCACGTGCATAACTAAATATGTACATGTGAGGCTACACTGATCCACCCCATGCACCAGTTACACCATCATCTTGATGCAATGCATCCCCCTGGACCGTGCCCTAATGGTTTCCCCTTCAATTCGCTACTTGAAAGGGCGGATTACCTGTAAGACTCCCACTACACCCATAGTATTCTTTAAGGTGATGCTAATTCTTTAAAATGTTTAGATACAAATGCATAAAGATTGTAAATTGTGTCCATCTTTGTCGTCCAATGGCATTGGACAGGACTTAATGTAAATATACTTTGCATTGTTACACTTCTATGGGCATAACATTTTTAGAAAACCAA |
| CT 846 | 6q | CTTCGTTGTGCGTACAAGCACACACGTAGCCATCCGTACGTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGAGAATGTAAGCTCCATAAGAGCAGGAACCTTAACAGCCTTGCTGACCAATGAATCTCCGGTACTTGAATCATCAGAATACCTGACACGCTGCAGGTGATTATGAAATACTTGTGAAATGAAGAAAGAATGCATCCACTTCCTACCAAGAACCAGGCTGAGATAACTGGTAAATACGTCAGTTCATACAAGATACAAATAATCTTTTAACTTCCTTCAGTGATCTAAACTCTGATCTTGATAAAACTATATTACTAGAAACTACATCCTTGGACCACCTTAAACAAGTGAATTGCATATCACTTTCACTGTGTTACATTATTCACTGTGTTACATCTTAATGTTATTGGAGTTCACAATTCATAAAAAAGGGGGG |
| CT 1117 | 6q | CTGAACGTACAAAGCACACCGTATCTCATTCGTACTTATAGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGATTGATACTGACACATAAAGAAGTTAACCTGATGCAGCATCTGAGTAAAACACAGTGTAATTATAGAACAGCAGACCCAACCATCAAATACATAATATGTTATAAACATAGTTTAATTCTCATCAACTAGAAAGTA |
| CT 893 | 6q | CTTCGGTTTGTAGCGGTACAAAGCACACACGTATCCATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGATTGATACTGACACATAAAGAAGTTAACCTGATGCAGCATCTGAGTAAAACACAGTGTAATTATAGAACAGCAGACCCAACCATCAAATACATAATATGTTATAAACATAGTTTAATTCTCATCAACTAAAAGTA |
| CT 1019 | 7q | GKKKKCSGWMAAAGCAMACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCMTCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTTTCCCACCTCCACTGCTTAAGGTTTAACTTAACAACTGACCTTCGACTGTAAAATGAAGCAAAGTATCTGCACTCAAGTAGTGAAAACCTAAGAAGTACTATTCTGCGCAACTTCCACAAGAGGGCAGTAGAACCCCTGAAGTAGTCCACAAAAGTAACAGACCTTGATTTTTCCTCTAATTTATTTTCATGCTAGCCTTCCTTAAAATCTTTTCCTAGAAAGCACTAACTAGCTGTGACCTGGCTAGGCCTCGTGGCTCATGCCTATCATCTCAGCACTTTGAGGGCCAAGGTGGGTGAATATCTTGAAGCCAGGAATTCAAGACCAGCCTGGACAAGATTGCAAGATCTTGTCTCTACAAAAAAATWAAATAAATTTTTTGGCTGGGAAAAAAAAAAAAAAAAAAAAAAAATTAAAATCCMCMKSWTCTCCCCTCCAAACTAAAAAAAAAAAAAATTAAKTAGATCATACAYAAAAAAATAAACAAAAATKASAAACTCTCACACTCGTGATGCGTGCGKWTTTACGTTGTTTGTCGWGAGTACTGAGCCCCG |
| CT 796 | 8p | CTTCGGTTGTGCGTACAGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGATGAGTAAGAAAGGTCACTTCCTCGGCAGACCTCAGGTATCATTGGGGACTCAAAAAACAC |
| CT 915 | 8p | TATTCTTTCAAAAAGATGGAAGACCTGTTATCCGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGATGAGTAAGAAAGGTCACTTCCTGGCAGACCTCAGGTATATTGGGGACATCAAGAAAAGAATAATTCATCAAAGTCTACAGATATTGCAGGCAAAGTTGAAACAAGCCAGACAAGTTCTCTTTGTATTCTTGAGAGACTTGGAAAGTCTAACCTGAGATTCCTTATTAAAAGTTCCAGCAAAACAGACTTA |
| CT 1160 | 8q | CTTCGTTGTGCGTACAAGCACACACGTAGCCATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGCTCAAGATGATTAAGATGCTGAAATGACGAAGAAGAGTCATGAACACTGTTCTTCTCCAAGTGACAGACCTTCAAAATTTGTGGGGTTTTTTTTTTCAATTTCCTTTTGTTTGACTATGCCCACAAAATCCCTCAGGGAAAGGTCAGCAAAGGTCCAAATGTCAATCAATCATGAACAAAGGGGTAAAGTGCCAAGGACCTGACCCGTCAAACAATGACTCAAGGGGACAGTTCTTTACAAAACAGCAGGACCTGCCCTTGAAAATTCCAAATACAAAAGGTTCTACAGGGAAAAAAAAAAAAAAAAAAAAAAGACCCCCCCCCCACAAAA |
| CT 1162 | 9q | CTTCGGGTTGTGCGTAAAAGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCAACCAAAGCATGGATAACCTAACGGTTTCCGGCATGAGCTGAACAAAGGAAACCGCCTACTGCCATACTGCTCTTCTAGGATTTTGCAAAGAAGAAAC |
| CT 714 | 9q | CTTCGGTTGTGCGACAAGAACCGTATACTCCGTACGTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTCCTGAAACACTATAACACTGATGACTGATTTGGTGACTATTCTATCCCTGGCGTCTTCACAATGTCTGAATTTAGTATGTACTAATATACATTTGATGAATAATCAAACGCAAACAACC |
| CT 753 | 9q | TGGTGCTACAAACACCTATAATCCGTACGTTGGAAGACCTGTTAATGGCCACCTAGGATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGAATCTTGCTCTGTCACCCAGGCTGGAGTTCTGAGAAGAATTGGCATTCAACCTCCCACCTATTGCACTGACCTCATTCATCACCCAAACCAGGACCCACATCCATCTGCATTCTCACTGCACAACCCGGCAATCCCTCTGCCGACACTATAATTCATTATCCCTACCACACCCCTCCAATCACGGTGATTGCCTGCATGAAGCCCGGCAATTCTGCAGGGAGGTGCAGGACTCACAAGCTCCAAATGAGGCTCAGCTATCACAGAATTCCATATCTGAGCTGGCTCTGGCTCTCAGTATGCTAAAAA |
| CT 821 | 10q | GAGGTCTGTTCCAAAAGGTAGCCGAGCGGGGAGGGCACGGGTTGTAACGGCTGGTTTTATGTACATGCTGGCAGCTCTAGATTATTAACTGTTGGTAATCCATATTTTAGGGTTCTGAGCTCTTTTGTTATTTTTATTATTTATCACTCTTTGGCTTGTCTTTTCACTCTTAATGGTATATATTGTACATCTTTTCTAGTTTGTTTTCAGTGAGAGGTTTCAGCACCATCTAGTCAACCACTGCCAGAGCAGAACTTGGTCTCCTATTTTAATGACTAAATTTTATTCCAGGCTGGGCATGGTGGCTCACGCCTGTAATCACAGCACTTTGGGAGGCTGAGGCAGGCAGATCACCTGAGGTCAGGAGTTCGAGACCAGCCTGACCACAATGGTGAAACCCCATTTCTACTTTTAAAAATACAAAATTTAGCCGGGCGTGATAGTGGGCTCCTGTGATCCCAGCTACTTAGGAGGCTGAGGCAGGAGAATCACTTGAACCGGGAAGCAAAAGATTGCAGTGAGCTTAGATTGCACCACTGCATTCCAGCCTGGGCAACAGAGGGAGACTTCATCTCAAAAAAAAAAAAAAA |
| CT 859 | 10q | GACGTTCATCTGAGTGTTTATTGGAAAAGCGTATCCCATTCGTACTTTGGAAGACCTGTTAATGGTCACACTAGGAATTGTGTGCCCCTTCTGTTCTCAGAAACCATATTCCCTTTGGAATGATCCTCCCCCCCACACCTAGTACTTTTCCCTCCCTTTTGTGAAGATCTTGACCAATGGCTAAGTGAAGATGACAATCATGTTGCACAATTCACTGTAAAGCTGGAAAGGGACAAACTGGTGTAATGATATGTGCATATTTATTACATCGGGGCAAATTTTTAAAGGCACAAGAGGCCCTAAATTTCTATGGGGAAGTAAGGACCAAAGACAAAAAGGTAAGTTATTTTTTGATGTTTTTCCTTTCCACTTCCTGGATCTGAAATTATTGAAAAA |
| CT 1169 | 11q | CTTCTCGAAGGAAAAAAAGACCTGTAATGGCCACCTAGAATTAAGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGATGAATCTGAGGACACAGATACTGATGGGGAAGAGGAGACATCACAACCCCCACCCCAGGCCAGCCACCCCTCTGCCCACTTTCAGAGCCCCCCGACACCCTTCCTGCCCTTCGCCTCTACTCTGCCTTTGCCCCCAGCGCCCCCGGGCCCCTCAGCACCTGATGAAGAGGACGAAGAAGATTACGACTCCTAGCGCCTTCTGCCCCCCAGACCATAGCCCCTTTTAGTTGGTTTTAGTTGCTCTGGGGGGAGGAGAGAAGGTAGAGCTGTTCTTAAATTTATTAAAAAAAAAAAAAAAAAAAATCGATGTCGACTC |
| CT 1183 | 11q | CTCTTCGGTTGTGCGTACAAGCACACACGTACGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGATGAATCTGAGGACACAGATACTGATGGGGAAGAGGAGACATCACAACCCCCACCCCAGGCCAGCCACCCCTCTGCCCACTTTCAGAGCCCCCCGACACCCTTCCTGCCCTTCGCCTCTACTCTGCCTTTGCCCCCAGCGCCCCCGGGCCCCTCAGCACCTGATGAAGAGGACGAAGAAGATTACGACTCCTAGCGCCTTCTGCCCCCCAGACCATAGCCCCTTTTAGTTGGTTTTAGTTGCTCTGGGGGGAGGAGAGAAGGTAGAGCTGTTCTTAAATTTATTAAAAAAAAAAAAAAAAAGGGATGGCGGCCCCGGCCAACAGATGTTCCCAGGGACCCAGCCTCAATTCACTGGGGGCGTTTTACACGGCGGGAAGGGGGAAAACTCGGAGGAATCCAACTTAGAGGCTATGGGAAAAA |
| CT 836 | 11q | CCTTCGGTTGTGCGTACAAGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCAGCAACGAAGTATCCTCTCCTGAAATTATTAGGCAGCACTTGGCCAACCACTCCGCCGCGACCCATACCAAAGCCGTCGCCTTGGGCACCGAAGAAACACAGACGACTATCCAGCGACCAAGATCAGAGCCAGACACCGGAAACCCCTGCCACACCACTAAGTTGTTGCACAGAGACTCAGTGGACAGTGCTCCAATCCTCACTGCATTTAACAGCTCACACAAAGGACGGATTAACTGTAATAGTAACACTACACCCATAGTACATTTAAAAGGTGATGCTAATACTTTAAAATGTTTAAATATAATTTAAAAAATTGTACATTGAAACCGCAGCCCACAGGCTGAAAGAGAAA |
| CT 871 | 11q | CTGTTTCTGAAACCCTGTCCTTTGTGTGTCCGTGGTGTGCATCCCAGCAGTAAGCAACAATGGCTGATCCAGAAGTGTTCATGAGCTTTGGAAAAGCTGCTAATAAGCCCTCAATCCGAGTTCGGGTCATTTCCACAAACTGTCGAGAAACAATAGCCTTTCCTGCTTTTGTGCAGACCGCTGCTGCCAACAGCACCTGCCAGGAACATATGCATAAGTACACTGGACGATGTTTTAAATCAAGTTATCTATTGACAAAAAAAAAAA |
| CT 896 | 13q | CTTCGGTTGTGCGTGCAAGCACACACGTATACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGACCCTTCTAGAGAACATTTGCCTTCGGCAACGGACTCTGAAGGTCACTGTCCTCTGGAGCTTAAGAACCAATATAATCATAGAATCCAGCCACTCAATAAATCCCAGCATCATACTCCAAAAGAGCTGCTGTTTCCAATTTCTTTATTCTTTGTAAAAAATATGGATAGGGATTTGTTTGATAACGAGCGCTGAAAAATTTTTTAAAAAAAAAAAAATTAAGAA |
| CT 976 | 13q | CTTCGTTGTGCGACAACCCCGTAACTTCGTACTTTGAAACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTTGCACAAAAGTTTATACCAAGTCTTCTCATTTAAAAGCTACCTGAGGACTACACTGGTA |
| CT 927 | 13q | CTTCGGTTGTGCGTACAAGCACACACGTAGATCTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTGCATTAGCTGAAAAAAGCCACTGTGTCCTGGGATGGTTTTAGACATCTGGACAAAAGGCAAAGATTCCATAATATAAGGGGTCGGTGGACCGGTCGATGTATGTCTTGTTGCAGATCATCAAGAACACGTATAGAAACCCAGCTGTAATCATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTTGCAACCAGAGACATCTGATCTCTACTGTTATGAGCAATTAAATGACAGCTCAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAACCGGACAGAGCCCATTACAATCTCCAGCATTCCCCCTCAAACCTAAAAAAAAAAAAAAAAAAAAAG |
| CT 1094 | 13q | CTTCGGTTGTGCGTGCAAGCACACACGTATACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGACCCTTCTAGAGAACATTTGCCTTCGGCAACGGACTCTGAAGGTCACTGTCCTCTGGAGCTTAAGAACCAATATAATCATAGAATCCAGCCACTCAATAAATCCCAGCATCATACTCCAAAAGAGCTGCTGTTTCCAATTTCTTTATTCTTTGTAAAAAATATGGATAGGGATTTGTTTGATAACGAGCGCTGAAAAATTTTTTAAAAAAAAAAAAATTAAGAA |
| CT 914 | 17q | GKGKGGGTMAGGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCMTCTGTTCTCAGAAACCATAA TCTACCATGGCTGATCCTGCAGCTGGAGTCTCACTGTCGCCCAGGCTGCAGTGAGTGAAGTGGTGCGATCTTGGCTCACTGCAACATCTGCCTCCCGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGACTACAGGCACACACCGCAACGCCCGGATGATTTTTTGTATTTTTAGTAGAGATAGAGTTTCACCGTGTTGCCCAGGCTGGTGTCAAACTCCTGAGCTTAGGCGATCCACCCGCCTCAGCCTCCCAAAGTGCTGGGATTACTGGCTTGAGCCGGTGCGCCAGGCCCAAATTTATTTTACTTTCAAAGCCACATGTATAAAACTTAGTTATATAGTTTTAACAGCTTTATGTAATAATTTTATTGAAAACATTATTCTTTTATTGTTCAAAATGTTCTATATGTTACGTGCAATTTTGCTTTCCTTGTTGCTCTGTAGTTTTAAAAGTTCTTTGTTTTTCTAAAAACAACTTCAGTATATTTTTTAATTATTTGTATAATTTGGCGGCTCTTATTGTGATTTTAGAATAATATTGTTTATGAACTAAATATATATATTCTCATTTGCCTTTAAAAAAAAAATAAATAAAKRRSWSYRYRCYATSAACATRS |
| CT 848 | 17q | CCTTCGGTTGTGCGTACAAGCACACACGTAGACCGTATTTCTCAATGTGAGAAGGACATGATATTTGGGAGGCCAGGGGTGGAATGATATGGTCTGGATGTGTGTCCCCTCCAAATCTCATGTTGAAATGTGATTCCCAGTGTTGGAGGTGGGGCTCAGTGGGAGGTAATGGATCATGGTGATGGGTGAGGAATTCATGCTTGAGTTAGTTCACGTGAGATCTGGCTGTTTAAAAGAGTCTAGGGCTAGGCATGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGCAGTTCACTTGAGCGCAGGAGTTTGAGACCAGCCTGGGCAACATGGCAAAACAAAAA |
| CT 1097 | 18q | CTTCGGTTGTGCGTACAAGCACACACGTAGACCTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGGGGAAGCTCATGCCACAGGACCTCTTCAAGAGCTGGAGACAAAAAGCCCCCTGTTCTGAGAATTGTTGAAAGGTTTCTTTCCCTTTCCCAGATTCTTTGCTCCTCGGTGCCCTGTCAAAACATCTGGCCCCTTTTGTCCTCAGCTTTCCCTGTGCTATTAAAGCCAGTCCTCATCACGAGCCACTTCCTGTGGACCTAAACATTTTTTTTATTATGCACCTTAACACTTCAAGATGATTAGAGCTGAAAGGGTAAGAACAAAAAATGAAGTTGCGCCTTTCTTTCACACTAATTTGATTC |
| CT 740 | 18q | CTGTTTCTAAACCCTGTCCTTTGTGTGTCCCTGGTGTGCATCCCAGCAGTAAGCAACAATGGCTGATCCAGAAGGTTAATGTTTGATCAAGATCTTCAGCTCGAAGGAGGAGAAGGAAGGGATCGGGTTTGCATGGATATGAGTGACCCACTTCACATGGAAA |
| CT 906 | 19p | CTTCGGTTGTGCGTACAAAGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCTACTCAGGAGTCTGAGGTGGGAGAATTGCTTGAACCCGGAAGGCAGAGACTGCTGTGAGCTGAGATTGTACCACTGTACTCTACCTGGGCCACAGACCAAGACTCTGACAAAAAAGAAAGAAAGAAAGAAAGTAAAGAATGAAAAAAAGAAAGAAGAGAA |
| CT 889 | 19q | CTTCGGTTGTGCGTACAAGCACACACGTATCACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTAGCTGAAGCAGAGAGAGAAGGCAGCATACATCAGCAATTTCTTCTCTGCACTTATAAGATCAAAGACTTTAAGACTTTCACTATTTCTTCTACCACTATCTACTACAAACTTCAAAGAGGAACCAGGAGTGTGAGAGGAGCATGAAAGTGGACCAGGAACGTGACCATTGAAACACTCCACCACAGGGAGGGGTTTCCGCCTCGAGATGACTGCCTCCAGGCCTAGATAGAATCCAGGCTTCCCAAGAATCCGGCGGAACAAATGTACCCCACTCCCCCAAGGAAC/ATGGAGACGTCCCTTT/ATGTGGTCTGCTAAGTAACGGTGCTGTCCCAAACACTAACGGTTAACCACTATGACCAAGGAGCCCTCCAAGCAGCTCCTTACTGTGAGAGTGTAACTAGAGGTCTCTCCTTGCTTCTAGGTGCCTACTATTCTGTCCACCACTGGTCCCCTTGCTAGCGACAACACGGTCTTAATTTCCCTTACAGGCTCATTATTACTAATTTAGCCGCCAACAACACACACACGAAACAGACGATTCCCCTCTAGAT |
| CT 839 | 20q | CTTTCGGGTTGCGCGTACAAGCACACACGTACAATTCGTAATATAGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCTGGCTGGATGTTGATCCTCAGGGAGACCTTGGAAGCTACCTGTTGCAGATGACTGAGCCCCCATCTGTCAGATTCCTGAGTGACTGACTGCATGAAGCAGAGCTCCTCCTCCTCATTGACCATCCTAACAATAAGAAATGTCTGCATTTACAAATGAAATGTTGTAATACTATGAAATTTTAAGAATTGAAATAAAAAATACAACTTATAAAAAAAAAAAAAAAAAAAAAAAAATCGATGTCCACTCCAGTCATCTTGCTGAAAAACTCCAGCCATCCGAAAGATCTGGCGGCCGCTCTCCCTATAGTGAGTCTTATTACGCCGGATGGATATGGTGTTCAGGCACAAGTGTTAAAGCAATTGATTTTATTCACTATGATGAAAAAAACAATGAATGGAGCCTGCTCCAAGTTAAAAATAGAAATAATACCGAAAACTCATCGAGTAGATTAATTAGAGATAATACAACAATAAAAGAATGGTTTATAACTTACTCACAGCGTGATGCTACTAGTTGGGACATTTTCCAGATGATGTATCATCTATTAATTTGAATGAAAAAGACTACCTGCTTTTGTTAAAATTATTTGACAAAAATAATATCATTCGGCTGCAGGGCGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATGATGGCTTCTAGACGTCAGGTGCTATTTTTCGGGGAATGTCGCAGCACCCCTATTTATTTTTTTTCTAATACATCAAGTATGTATCCGCTCATGAAACACTAGCTGATGATGCTCAATAGATGAAAAAGAAAAAGATGGATATCACTTCGGGTCCCTTCTTCTTTTTGCGATTGCCTCTGTCTGCCCACAAACGGAGAATAAGAGCTAACGTGGGCAAGGGTTACAAGTACCAGGAACTGGATCCAATCTTAAGAAACTTACGT  |
| CT 785 | 20q | CCTTCGGTTGTGCGTACAAGCACACACGTCGACATCCGTACGTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGTATTTCCTCCAAAGCCGATCACCTTGACGGACGAGCGGAAAGCCCACCGGCTGCAGCGCATGCTGGACATGAAGGTGAATCCTGTGCAGGGCCTGGCCTCCCGCTGGGACTATGAGAAGAAGCATTGGAACATGTGGTTTGCATCACCTTCTGACTCCCTGAGGCTCCGCCCTGGCTGGGAGCCTCTGGCGGC |
| CT 1170 | 20q | CGGTTGTGCGACAAGACCACGTACACATCCGTACGTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAAAACCATAATCTACCATGGCTGATCCTGCATATTTCCTCCAAACCGATCACCTTGACGGACACGGACCTGGCA |
| CT 918 | 20q | CCTTCGGTTGTGCGTACAAAGCACACACGTACTACATCCGTACGTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGTATTTCCTCCAAAGCCGATCACCTTGACGGACGAGCGGAAAGCCCAGCAGCTGCAGCGCATGCTGGACATGAAGGTGAATCCTGTGCAGGGCCTGGCCTCCCGCTGGGACTATGAGAAGAAGCAGTGGAACAATGACTTGCATCCCCACTGTCTCCCTGAGCTCCCCCTGGTGGACCTCTGGCCCTCCCCTCCCCTGCCCTAACCCCAAATCAA |

Key: HPV sequences are highlighted in blue.

Table S3b. Sequences of site of viral integration with low score

|  |  |
| --- | --- |
| **Sample ID** | **Sequence** |
| CT 1123 | CCTTTCGGTTGTGCGTACAAGCACACACGTACTACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGATCATCAAGAACACGTAGAGAAACCCAGCTGTAATCATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTTGCAACCAGAGACAACTGATCTCTACTGTTATGAGCAATTAAATGACAGCTCAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAATCGGACAGAGCCCATTACAATCTCCAGCATTCCCCCTCAAACCTAAAAAAAAAAAAAAAAAA |
| CT 1194 | CTTCGGGTTGTGCGTACAAGCACACACGATCAATTCGTACATAATGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGAGCAACAAAGTATCCTCTCCTGAAATTATTAGGCAGCACTTGGCCAACCACTCCGCCGCGACCCATACCAAAGCCGTCGCCTTGGGCACCGAAGAAACACAGACGACTATCCAGCGACCAAGATCAAAGCCGGACACCGGAAACCCCTGCCACACCACTAAGTTGTTGCCGGAGACTCGTGGACAATGCTCCAATCCTCACTGGATTTAACAGCTCCCGGAGGATTAAAAAAGGAGTAACCCCACACCCAAAAAAATT |
| CT 777 | TTTGCWKGGGAKSGAAGGSCCGGGGGCGRAAGYAACACGTARACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCMTCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGSAGCAWCRAAKKATCCTCTCCTGAAATTATTAGGCAGCACTTGGCCAACCACTCCGCCGCGACCCATACCAAAGCCGTCGCCTTGGGCACCGAAGAAACACAGACGACTATCCRGCGACCAAGATCARAGCCAGACACCGGAAACCCCTGCCACACCACTAAGTTGTTGCACAGAKACTCRGTGGACAGTGCTCCAATCCTCACTGCATTTAACAKCTCACACRAAGGACGGATTAACTGTAATAGTAACACTACACCCATAGTACATTTARRAGGTGATGCTAATACTTTAAAATGTTTAAGATATAGATTTAAAAAGCATTGTACATTGTATACTGCAGTGTCGTCTACATGGCATTGGACAGGACATAATGTAAAACATAAAAGTGCAATTGTTACACTTACATATGATAGTGAATGGCAACGTGACCAATTTTTGTCTCAAGTTAAAATACCAAAAACTAAAAAAAAAAAAAAAAAAAAMTAATGTCCTCCGGACCCCCAATYACCCTTGTCGRCCT |
| CT 733 | CTTTCAGTTGTGCGTACAAGCACACACGTACACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCAGCAACGAAGTATCCTCTCCTGAAATTATTAGGCAGCACTTGGCCAACCACTCCGCCGCGACCCATACCAAAGCCGTCGCCTTGGGCACCGAAGAAACACAGACGACTATCCAGCGACCAAGATCAGAGCCAGACACCGGAAACCCCTGCCACACCACTAAGTTGTGCACAAACTCGACAGTGCTCCTCCTCACTGATTTAACAGCTCACACAAAGGACGGATTAACGCTGTAATAGTATACACTACACCCCATGTACATTTAAAAGGACGATGCATAAACCTTAAAATTGCTTTAACCATAAAAATTTATAAAACACCATCAAATCTTCGAAATAACCTCGACAGTGTCCCTCCAAAAGTTTGCCTTCTAACTACAAACAAAACACATTAGACTCACACTCCATACATTAACATATTACCGATTCAAAATATAAAGAACAAATACATAAAAAAAATAATAGTAGATATGATCTAAATTTCTACC |
| CT 718 | CTTTCGGTTGTGCGTACAAGCACACACGTAGCCATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTGTATTAACTGTCAAAAGCCACTGTGTCCTGAAGAAAAGCAAAGACATCTGGACAAAAAGAAAAGATTCCATAATATAATGGGTCGGTGGACCGGTCAATGAATGTCTTGTGGCCGATCCCGGCAACACTGTTGGAACCCCAGGTGTAACCCTGCCTGGTGATACCCCTACATTCATTGCTGGTATGTTCAATTTGCCCCCCCAGACTACTGATCTCTACTGTTATGACCAATTACATGACAGCTCACTCTAGGAGGATGAAATAAATGGTCCAGCTGGACAAGCAGAATCGGACAGAGCCCATTACAATCTCCAGCATTCCCCCTCAAACCTACCACACCCCCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAAAGAACATGTTATTTATTCCTCCCGCAGACCCCCGAAAAACAACCATTTTAAACAACCCCACATAGAAAAGGAAAA |
| CT 707 | CTTCGGTTGTGCGTACAAGCACACACGTATACATCCGTACGTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCTATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCAGCGACGAAGTATCCTCTGCTGAAATTACTAGGCAGCACTTGGCCAACCACTCCGCCGCGACCCATACCAAAGCCGTCGCCTTGGGCACCAAAGAAACACAGACGACTATCCAGCGACCAAGATCAGAGCCAGACACCGGAAACCCCTGCCACACCAACAAGTTGTTGCACAGAGACTCCGTGGACAGTGCTCCAATCCTCACTGCAGTTAACAGCTCACACAAAGGACGGATTAACTGTAATAGTAACACTACACCCATAGTACATTTAAAAGGTGATGCTAATACTTTAAAATGTTTAAGATACAGATTTAAAAAGCATTGTAAATTGTATACAGCAGTGTCGTCCACATGGCATTGGACAGGACATAATGTAAAACATAAAAGTGCAATTGTTACACTTACATATGATAGTGAATGTCAACGTGAACAATTTTTGTCTCAAGTGAAAATACCAAAAACTAATACAGAAAAAAAAAATATATATTAGCTA |
| CT 819 | CCTTTCGGTTGTGCGTAAAAGCACACACGTATACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTGTATTAGCTGTCAAAAGCCACTGTGTCCTGAAAAAAAGCCAATACATCTGCTGGATTCACACCAATCCCCTGCCACACCTCCGGATGGATGGAACTATGTATGCCCGCCTGCACACCCGGTAATTTTTTTATAATCCCATATGAAACTGTGTGACCGATACCCGAGGGTGCATCTCCTGAACTCGAAACAAACACGATCTCCAAATGAAAGGCCCCAAACAACCA |
| CT 959 | GTKTKCSGTMAGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCAYCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGSAGCAACRAAKKATCCTCTCCTGAAATTATTARGCAGCACTTGGCCAACCACTCCGCCGCGACCCATACCAAAGCCGTCGCCTTGGGCACCGAAGAAACACAGACGACTATCCAGCGACCAAGATCARAGCCAGACACCGGAAACCCCTGCCACACCACTAAGTTGTTGCACAGAKACTCAGTGGACAGTGCTCCAATCCTCACTGCATTTAACAGCTCACACAAAGGACGGATTAACTGTAATAGTAACACTACACCCATAGTACATTTAAAAGGTGATGCTAATACTTTAAAATGTTTAAGATATAGATTTAAAAAGCRTTGTACATTGTATACTGCAGTGTCGTCTACATGGCATTGGACAGGACATAATGTAAAACATAAAAGTGCARTTGWGACACTTACATATGATAGTGAATGGCAACGTGACCAATTTTTGTCTCAAGTTAAAATACCAAAAACTAATAWAAAAAAAAAAAAACAAC |
| CT 912 | CTTCGTTGCGCGTACAAGCACACACGTCACATTCGTACTAATTGGAAGACCTGTTAATGCGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCACAACGAAGTATCCTCTCCTGAAATTATTAGCAGCACTTGGCCAACCACCCCGCCGCGACCCATACCAAAGCCGTCCCTTGGGCACCGAAAAACACAACACTATCCACGACCAAATACAACCCAACCCCTGACATATTGGAAACAAGAAG |
| CT 837 | CTTCGGTTGTGCGTACAAGCACACACGTAGTCCATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCAGCAACGAAGTATCCTCTCCTGAAATTATTAGGCAGCACTTGGCCAACCACTCCGCCGCGACCCATACCAAAGCCGTCGCCTTGGGCACCGAAGAAACACAGACGACTATCCAGCGACCAAGATCAGAGCCAGACACCGGAAACCCCTGCCACACCACTAAGTTGTTGCACAGAGACTCAGTGGACAGTGCTCCAATCCTCACTGCATTTAACAGCTCACACAAAGGACGGATTAACTGTAATAGTAACACTACACCCATAGTACATTTAAAAGGAGGTAGGAGGAGGTGAGAAAATTACTACTTGATGTCCTCCTCCTGCCGTATCACCGCTGGTGGTATATTCTCATCGGAGGGCGGCGACCATTGTGGACAGTGAATGAGGGGAATCTTCGATGCTTGATGCTTATTATGATCCGACGAGGAGTGCTGGATAATCTTCGAGGAGCCACCTCTTCTCAAAAAACCCCCCCCACCCTCCTTCCTTGGTGAAAAATTTGGGGAGGCCCCCCTGGGGGAGACAAGAGGGGGATGCTGCCCAAAAGGCGAAAAAAGGAAAAAAACTTCCTCCTTGTTTTTCTATCACCCACGTCCAAAAAATTAGCTCTGTCGTGTACGTGTGTGTACGTGTGCCACAACCCACTACACTATATAGAAAGAGATGAATGTACAATATACG |
| CT 744 | GTGGAAAGGCTTCGGTTGTGCGTAAGGCACACACGTAATCCTTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTACAAATGGGGAAGAGGGACGGGATGTAATGGATGGTTTTATGTGCAGGCTGTAATGAAAAAATTCTCCTGCCTCAATGGAGTATGGCTCCAGCCAACAAACAGATAGTTGGAAAA |
| CT 763 | CTTCGGTTGTGCGTACAAAGCAACACGTATAATTCGTGCTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTACCAATGGGGAAGAGGGTACGGGATGTAATGGATGGTTTTATGTAGAGGCTGTAGTGGAAAAAAAAAAAAAAAAATCAATGTCACTCGGGTCAAAAAATCGGAGAGCCCCTTAAAATCCCCCGCATTCCCCCCCCAACCTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA |
| CT 793 | CTTCGGTTGTGCGTACAAGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTACCAATGGGGAAGAGGGTACGGGATGTAATGGATGGTTTTATGTAGAGGCTGTAGTGGAAAAAAAAAAAAAAAAAACGGGGTCAATCCAGGCAATCTAAATGGCTTTCGAGGACACAAGACCCAAATCCGGGGTGTTTTTTTTACAACGCCCGAACGGGAAAAACTGGTGGTCCCCTTAATCCCCTGCAGCAAATCCCCCTTTATAATTGGGGGG |
| CT 1215 | CTTCGTTGTGCGTACAAGCACACACGTAGACATTCGTACTTTGGAAGACCTGTTATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTACCAATGGGGAAGAGGGTACGGGATGTAATGGATGGTTTTATGTAGAGGCTGTAGTGGAAAAAAAAAAAAAAAAATCGATGTCACCCCAGTCAAAAAAAATGCGTTTTCGAGGGGCCCAGATCAATTTTCTGGGCGGCGTTTTTTAACGGCGTGGCTGGGAAAAACCGGGTTTCCCCACTTTTCCCCTTGGGCCATCCCCCCTTAAAGTTGGAGGGGGCCCTGGGTTAAGGGCCGGGTCTGATTTGAAAGACC |
| CT 825 | CCTTCGGTTGTGCGTACAACGCACACACGTAGACATTTCCTACTTTGGAAGACCTGTTAATGGGCACACTAGAGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGCAGCAACAAAGTATCCTCTCCTGAAATTATTAGGCAGCACTTGGCCAACCACTCCGCCGCGACCCATACCAAAGCCGTCGCCTTGGGCACCGAAGAAACACAGACGACTATCCAGCGACCAAGATCAGAGCCAGACACCGGAAACCCCTGCCACACCACTAAGTTGTTGCACAGAGACTCAGTGGACAGTGCTCCAATCCTCACTGCATTTAACAGCTCACACAAAGGACGGATTAACTGTAATAGTAACACTACACCCATAGTACATTTAAAAGGTGATGCTAATACTTTAAAATGTTTAAATATAATTTAAAAAGCATTGTACATTGTATACTGCAGTGTCGCTACTGGATTGGACAGCGGATGTAAACATAAAAAGTGCAACTTGTCTACCCTTCATCATACGAGGTGGAACGCTCACTTTGTCCACAATTAACATCATACTAAAAAAAAA |
| CT 752 | CGTTTCTGAAACCCTGTCCTTTGTGTGTCCTCTGGTGTGCATCCCAGCAGTAAGCAACAATGGCTGATCCAGAAGGTACCTACGGGGACGGAACGGGTTGAAACGGCTGGTTTTATGTACAAGCTATTGTAGACAAAAAAACTGGAGATGTAATATCGGATGACGAGGAGGAAAATGCGGTTCGGGCTGGGTTGCTACACGACGGGAAGTTGCCACACACGGACATTTTGTGAGCAGGAAAATATAGAGGCTGACACGAATTGTTCCCTGCCAAGAGACCACAATGATGCACACTGTTGCATGTTTTAACACCAACCTTCCCCCACCCACAACACCCCCCCCCAAAAA |
| CT 1202 | CCTTCGGTTGTGCGTACAAAAGCACACACGTTCGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTGCATTAGCTGTCAAAAGCCCCTGTGTCCTGAAATGGTTTTATGCATAGGCTGAAATGGAAAAAAATCCCTGGTATAATGGGTCGGTGGACCGGACGATGTTTGTCTTGTTGCAGATCATCAAGAATTCGTACAGAAACCCAGCTGTAATCATGTGCAGATCCAACATAGCTGATTTGCCCATTCCCCCCCGACACACCGACCCACTGCAAACA |
| CT 716 | CGCTCGCGACTTACAAACAGATCCATAAACCCCGTCCTTCGGTTGTGCGTAACAAAGCACACACGCTAGTACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGAAATTATTCTCCTCCTGGTGGTGCTGGCAGAACAACCCCTGGTGGCCCCACTCTTCCTCCAATACTTATTTCATCCTCCTGAGGAGGACGCCCTGCTGCTCCCCCGCTGGCTGCTTTTTTTCAGGAAGTGGTGGCCAAAAATAATCATCCTCCACCCCCCCCCAACAGTTATGTTTGTTTCTTCTCTTGGGGGAAAAAATAAAAGTGGTGTTTATTAAAATTTTTTTTTGAAGAAAAAAAAGATAAAAAAAGGACTTGGTGATGATTTTTTTTTTGTTAAAAAAAACCCCCCAACAATTTTTTTTTTAACAAAAAAAAAACCTGAAGGGGGGGGGGGGGAAGAATTTATAATACAATGCTGCCCGCTATAAAAGATGCACCACACATGATGACTACTGCCGTCTTCTTGTCGCCGCGTGGGGCGAGAGTCGTGGCGACCACCACAAAATAATACGTGGTGATCACATCATGGTGATGAGTATTAGCAGCATATAGAGCAAAGGTGGGGAAATTCTTCACGATGTTCATCAGAAGTGAAGAAACATAAAAAACACTTCTTCCACCACCTCATATAACCACCACTTTTTCTGGCGCACACCGGGTTGCTCCTCGTCCTACTACCCCTTCTTCTGTCGGCTGTGCTCCGGAGCTAGCTGTCTTCTTCGCTGCCTCCTTGCTGACTAATAACTCCCCCTTCTAAAAAAAAAAAAAAAAAAATTTGTTCACCTAACCCCACACAACTCAGAAGACCCCTTTTTTCCTTTATTCCCCCCCCACTTTAAAAAAAAAAAAGAAAGGGGAAAC |
| CT 940 | CTTCGGGTTGTGCGTACAAGCACACACGTCGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCATCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGGCCAACGAGATATCCTCTCCTGAAATTATTAATGCAGCACTTGGCCAACCACCCCGCCGCGACCCATACCAAAGCCGTCGCCTTGGGCACCGAAGAAACACAGACGACTATCCAGCGACCAATATCAGAGCCAGACACCGGAAACCCCTGCCACACCACTAAGTTGTTGCACAGAGACTCAGTGGACAGTGCTCCAATCCTCACTGCATTTAACAGCTCACACAAAGGACGGATTAACTGTAATAGTAACACTACACCCATAGTACATTTAAAAGGTGATGCTAATACTTTAAAATGTTTAAGATATAAATTTAAAAACATTGTAATTGTATACTGCATTGTCGTCTACTGGATTGGACGAAATAATGTTCAACTCACTCTTCCCCCACCACAG |
| CT 922 | AGGGTKGKCGGWAAAGCAMACACGTAGACATTCGTACTTTGGAAGACCTGTTAATGGGCACACTAGGAATTGTGTGCCCCMTCTGTTCTCAGAAACCATAATCTACCATGGCTGATCCTGCAGGTGTATTAACTGTCAAAAGCCACTGTGTCCTGAAGAAAAGCAAAGACATCTGGACAAAAAGCAAAGATTCCATAATATAAGGGGTCGGTGGACCGGTCGATGTATGTCTTGTTGCAGATCATCAAGAACACGTAGAGAAACCCAGCTGTAATCATGCATGGAGATACACCTACATTGCATGAATATATGTTAGATTTGCAACCAGAGACAACTGATCTCTACTGTTATGAGCAATTACATGACAGCTCAGAGGAGGAGGATGAAATAGATGGTCCAGCTGGACAAGCAGAAYCGGACAGAGCCCATTACAATCTCCAGCATTCCCCCTCAAACCTAAAAAAAAAAAAAAAAAAARAARWYKAAWCYAGGSCAAAGG |

Key: HPV sequences are highlighted in blue.