**Supplemental Data**

We deep sequenced RT gene into 2 fragments: RT1 (413pb), RT2 (446pb) and gp120 gene (367pb). These genes were amplified using two rounds of PCR amplification.

Details of primers used for amplification are listed in table 1.

PCR1 Forward (for) and PCR1 Reverse (rev) for PCR round one, and PCR 2 for and PCR2 rev for PCR round 2 (nested PCR). To amplify RT1 and RT2 fragments, the following thermocycler parameters were used for PCR1: 50 °C for 30 minutes (mn), 94 °C for 7 mn, 94 °C for 10 seconds (s), 55 °C for 30s, 68 °C for 1 mn, 35 cycles of steps 3–5 and 68 °C for 7 mn. We used a touchdown PCR for round 2 with the following parameters: 98 °C for 1 mn, 3cycles (98°C for 10s; 66-64 °C for 30s; 72°C for 15s), 3cycles (98°C for 10s; 64-62 °C for 30s; 72°C for 15s), 3cycles (98°C for 10s; 62-60 °C for 30s; 72°C for 15s), 30 cycles (98°C for 10s; 60 °C for 30s; 72°C for 15s), and 72 °C for 7 mn.

To amplify ENV C2V3 region, the following thermocycler parameters were used for PCR1: 50 °C for 30 minutes (mn), 94 °C for 7 mn, 94 °C for 10 seconds (s), 53 °C for 30s,

68 °C for 1 mn, 35 cycles of steps 3–5 and 68 °C for 7 mn. Parameters used for PCR2 were:

98 °C for 1 mn, 98°C for 10s, 60 °C for 30s; 72°C for 15s, 40 cycles of steps 2-4 and 72 °C for 7 mn.

|  |  |
| --- | --- |
| **Primer Name** | **Primer Sequence (5’-3’ orientation)** |
| RT1-PCR1-for | TAGTCCTATTGARACTGTACCAGT |
| RT1- PCR1-rev | ATCCTACATACAARTCATCCATG |
| RT1- PCR2-for | ATGGCCATTGACAGAAGAAA |
| RT1- PCR2-rev | TGGAATATTGCTGGTGATCC |
| RT2- PCR1-for | GGGARGTYAATTAGGAATACC |
| RT2- PCR1-rev | AGTCTTTTGATGGGTCATAATA |
| RT2- PCR2-for | GATGTGGGkGATGCATATTT |
| RT2- PCR2-rev | CTGTATGTCATTGACAGTCCAG |
| V3- PCR1-for | CAG TAC AAT GTA CAC ATG G |
| V3- PCR1-rev | ATG GGA GGG GCA TAC ATT G |
| V3- PCR2-for | TTACAGTAGAAAAAT TCC CCT C |
| V3- PCR2-rev | AAT GGC AGTCTA GCAGAA G |

Table 1 : Primers used for Ultra Deep Sequencing

for=forward; rev=reverse