

## Additional materials

### Additional file 1: Table S1. Characteristics of study participants at time of pyrosequencing.

| PID <sup>a</sup> | Viral load          |                                  | Age/<br>length<br>of<br>infection<br>(years) | CD4<br>% <sup>d</sup> |
|------------------|---------------------|----------------------------------|--|-----------------------|
|                  | Plasma <sup>b</sup> | Cell-<br>associated <sup>c</sup> |  |                       |
| <b>S1</b>        | 4.0                 | 2.6                              | 4.5  | 24                    |
| <b>S2</b>        | 4.8                 | 4.5                              | 1.5  | 30                    |
| <b>S3</b>        | 4.9                 | 3.4                              | 4.8  | 22                    |
| <b>S4</b>        | 4.9                 | 3.3                              | 2.1  | 17                    |
| <b>S5</b>        | 5.2                 | 5.0                              | 4.2  | 22                    |
| <b>S6</b>        | 5.7                 | 5.5                              | 6.1  | 2                     |

<sup>a</sup> Patient identification. Subjects included in pyrosequencing study were infected perinatally by maternal transmission for more than one year, were naïve to combination antiretroviral therapy with viral RNA levels >4 log<sub>10</sub> copies per ml of plasma and genetic analysis of HIV-1 envelope by conventional clonal sequencing.

<sup>b</sup> Log<sub>10</sub> HIV-1 RNA copies per ml of plasma by measuring HIV-1 *gag* gene (COBAS® AMPLICOR HIV-1 MONITOR Test, v1.5, Roche, Pleasanton, CA, USA).

<sup>c</sup> Log<sub>10</sub> HIV-1 *gag* copies/10<sup>6</sup> CD4<sup>+</sup> T cells

<sup>d</sup> CD4% is used to normalize for the age-related lymphocytosis that occurs in pediatric subjects.