Supplementary data

Figure S1: Validation of methods by mixing BAC-derived HCMV strains Merlin and TB40E in various proportions to mimic a multiple strain infection. Sequencing libraries were prepared from four samples: Merlin and TB40E DNA mixed in 1:1, 1:10, 1:50 proportions, and unmixed Merlin. The latter was used as the reference for variant calling. A: Variants are plotted by frequency in the population and position on the HCMV genome. B: The frequency distribution of variants at each dilution illustrated in a histogram. Variants are binned at 1% intervals along the x-axis and the number of variants in each bin is shown on the y-axis. No variation was detected in the unmixed Merlin sample. The variant frequency was ~53%, ~14% and ~3% (median) for the 1:1, 1:10 and 1:50 mixtures, respectively.
Figure S2: Longitudinal HCMV sequence diversity in stem cell transplant recipient SCTR11 (D-/R+). The consensus sequence from day 88 was used as a reference for variant calling. A: Variants are plotted by position on the HCMV genome and frequency at each time point. B: The frequency distribution of variants at each time point is illustrated in a histogram. Variants are binned at 1% intervals along the x-axis and the number of variants in each bin are shown on the y-axis. C: HCMV DNA load (IU/ml) in peripheral blood is plotted against the time after transplantation (days) during the follow-up period. Assembled consensus sequences and identified variants from blood samples for which consensus sequences were assembled are indicated by blue arrows.
Figure S3: Longitudinal and compartmental HCMV sequence diversity in Child4. The consensus genome from day 92 was used as a reference for variant calling. A: Variants are plotted by position on the HCMV genome and frequency at each time point. B: The frequency distribution of variants at each time point is illustrated in a histogram. Variants are binned at 1% intervals along the x-axis and the number of variants in each bin are shown on the y-axis. Red arrows indicate resistance mutations detected at day 161. C: HCMV DNA load (IU/ml) in peripheral blood is plotted against the time after birth (days) during the follow-up period. Assembled consensus sequences and identified variants from blood and respiratory secretions for which consensus sequences were assembled are indicated by blue and purple, respectively.
Figure S4: Longitudinal and compartmental HCMV sequence diversity in SCTR9 (D+/R-). The consensus genome from day 83 was used as a reference for variant calling. A: Variants are plotted by position on the HCMV genome and frequency in each time point. B: The frequency distribution of variants at each time point is illustrated in a histogram. Variants are binned at 1% intervals along the x-axis and the number of variants in each bin are shown on the y-axis. C: HCMV DNA load (IU/ml) in peripheral blood is plotted against the time after transplantation (days) during the follow-up period. Assembled consensus sequences and identified variants from blood and intestinal biopsy samples for which consensus sequences were assembled are indicated by blue and brown arrows, respectively.