

```

Toledo 6 mcrreslrtpwlfwvllscprlleyssssfpfataDIAEKMWAENYETTSAPVPLVAEGEQVTIPCTVMTHSWPMVSIRARFCRSHDGSDELILDAVKG
W 1 .....a.....
U3 1 .....
3157 1 .....
AL 2 .....s.....
JP 2 .....
4119 1 .....
VR3216B 1 .....
Lisa 5 .....*.....
U4 1 ..h.....
6397 1 .....
711 1 .....f.....

```

```

Toledo HRLMNGLQYRLPYATWNFSQLHLGQIFSLTFNVSMDTAGMYECVLRNYSHGLIMQRFVILTQLETLSRPDEPCCTPALGRYSLGDQIWSPTPWRLRNHDC
W .....T.....
U3 .....
3157 .....
AL .....
JP .....
4119 .....T.....
VR3216B .....T.....
Lisa .....T.....
U4 .....T.....
6397 .....T.....
711 .....T.....

```

```

Toledo GTYRGFQRNYFYIGRADAEDCWKPACPDDEPDRCWTVIQRYRLPGDCYRSQPHPPKFLPVTAPPADIDTGMSPWATRGIAAFLGFWSIFTVCFLCYLCY
W .....
U3 .....A.....
3157 ..H.....
AL ..M.....
JP ..M.....
4119 ..M.....Y.....
VR3216B ..M.....
Lisa ..M.....
U4 ..M.....T.....
6397 ..M.....
711 ..M.....

```

```

Toledo LQCCGRWCPTPGRGRRGGEGYRRLPTYSYPGVRKMKR
W .....
U3 .....
3157 .....
AL .....
JP .....
4119 .....
VR3216B .....
Lisa .....K.....
U4 .....K.....
6397 .....H.....K.....
711 .....K.....

```

Supplementary Figure 1

Amino acid sequence alignment of UL141 from 23 HCMV strains showing a high degree of conservation. A representative of each sequence is shown, and the number of strains with that sequence is indicated. Dots denote residues identical to those in Toledo. The signal sequence is in lower case, the transmembrane domain is underlined thinly, and potential N-linked glycosylation sites are underlined thickly. The location of the frameshift mutation in Bar (related to Lisa, derived from strain TB40E) is marked by an asterisk.