|  |
| --- |
| **TABLE S1. HIV-1 specific CD8+ T cell responses in early infection: epitope specificity, MHC restriction, and frequency** |
| **PTID.** | **HLA Class I type** | **HIV protein** | **HXB2 sitea**  |  **Epitope Sequenceb** |  **HLA restrictionc** | **SFC/106 PBMC** |
| 11439 | A\*01:01, 29:02; B\*38:01, 44:03; C\*12:03, 16:01 | gp160 | 209 | 217 | SFEPIPIHY | A\*2902 | 1042 |
|  |  | gp160 | 104 | 112 | MHEDIISLW | B\*3801 | 3390 |
|  |  | gp160 | 376 | 384 | FNCGGEFFY | A29 | 1288 |
|  |  | gp160 | 620 | 628 | **NEIWDNMTW** | **B\*4403** | 3144 |
|  |  | Protease | 34 | 42 | EEMNLPGRW | B\*4403 | 1406 |
|  |  | Integrase | 10 | 19 | EEHEKYHSNW  | B44\*03 | 760 |
|  |  | Integrase | 227 | 235 | **YRDSRDPLW** | **B\*3801** | 1351 |
| 53617 | A\*02:01, 02:05; B\*27:05, 49:01; C\*02:02, 07:01 | Vpr | 31 | 39 | VRHFPRIWL | B27 | 408 |
|  |  | p17 | 11 | 19 | **GELDRWEKI** | **B\*4901** | 1033 |
|  |  | p17 | 15 | 23 | **RWEKIRLRP** | **A\*0205** | 1013 |
|  |  | p24 | 131 | 140 | KRWIILGLNK | B\*2705 | 1363 |
|  |  | gp160 | 777 | 785 | **IVTRIVELL** | **A\*0205** | 278 |
|  |  | gp160 | 786 | 794 | GRRGWEALK | B\*2705 | 103 |
|  |  | Protease | 76 | 84 | LVGPTPVNI | A\*0201 | 181 |
|  |  | Rnase | 60 | 68 | **QYALGIIQA** | **A\*0201** | 175 |
|  |  | Rnase | 77 | 85 | **LVSQIIEQL** | **A\*0205** | 1086 |
|  |  | Integrase | 186 | 194 | KRKGGIGGY | B\*2705 | 396 |
|  |  | Integrase | 165 | 173 | **VRDQAEHLK** | A\*0205 | 606 |
| 21746 | A\*01:01; 48:01; B\*40:01, 57:01; C\*03:04, 06:02 | p24 | 30 | 40 | KAFSPEVIPMF | B\*5701 | 412 |
|  |  | p24 | 15 | 23 | ISPRTLNAW | B\*5701 | 2355 |
|  |  | p24 | 108 | 117 | TSTLQEQIGW | B\*5701 | 1356 |
|  |  | p2p7p1p6 | 119 | 127 | **ELYPLASLR** |  | 278 |
|  |  | RT | 244 | 252 | IVLPEKDSW | B\*5701 | 1117 |
|  |  | Integrase | 123 | 132 | STTVKAACWW | B57 | 1138 |
| 71101 | A\*01:01; 03:01, B\*08:01, 35:03; C\*04:01, 07:01 | Nef | 68 | 76 | [FPVKPQVPL](http://www.hiv.lanl.gov/content/immunology/ctl_search?results=Search;id=54548) | B35 | 1874 |
|  |  | Nef | 90 | 97 | FLKEKGGL | B\*0801 | 255 |
|  |  | p24 | 195 | 202 | NPDCKTIL | B\*0801/B35 | 203 |
|  |  | p24 | 84 | 91 | HPVHAGPI | B35 | 128 |
|  |  | gp160 | 787 | 795 | RRGWEVLKY | A\*0101 | 2475 |
|  |  | gp160 | 848 | 856 | RQGLERALL | B8/B\*08 | 136 |
| 75688 | A\*02:01, 32:01; B\*08:01, 55:01; C\*03:03, 07:01 | p24 | 128 | 135 | EIYKRWII | B\*0801 | 538 |
|  |  | gp160 | 848 | 856 | RQGLERALL | B8 | 2918 |
|  |  | gp160 | 704 | 712 | **VINRVRQGY** |  | 693 |
| 41325 | A\*29:02, 32:01; B\*35G1, 44:03; C\*04G1, 16:01 | Nef | 128 | 137 | TPGPGVRYPL | B35 | 619 |
|  |  | gp160 | 209 | 217 | SFEPIPIHY | A\*2902 | 274 |
|  |  | gp160 | 704 | 712 | **VINRVRQGY** | B\*4403  | 77 |
| 94153 | A\*25:01, 68G1; B\*14:02, 44G1; C\*05:01, 08:02 | gp160 | 88 | 96 | NVTENFNMW | A\*2501 | 280 |
|  |  | gp160 | 584 | 592 | ERYLRDQQL | B\*14 | 396 |
|  |  | p24 | 13 | 23 | QAISPRTLNAW | A\*2501 | 220 |
|  |  | p24 | 71 | 80 | ETINEEAAEW | A\*2501 | 447 |
|  |  | p24 | 166 | 176 | DRFYKTLRAEQ | B\*1402 | 409 |
|  |  | Protease | 34 | 42 | EEMNLPGRW | B44 | 327 |
|  |  | Integrase | 10 | 19 | EEHEKYHSNW  | B\*4403  | 353 |
| 20786 | A\*02G1, 11:01; B\*27:05, 40:02; C\*02:02, 02:02 | p24 | 131 | 140 | KRWIILGLNK | B\*2705 | 1390 |
|  |  | p2p7p1p6 | 64 | 71 | TERQANFL | B\*4002 | 270 |
|  |  | Integrase | 186 | 194 | KRKGGIGGY | B\*2705 | 113 |
| 17543 | A\*01:01, 02:02; B\*07G1, 15:03; C\*02:10, 07:02 | Rev | 66 | 75 | RPAEPVPLQL | B7 | 798 |
|  |  | Nef | 183 | 191 | WRFDSRLAF | B\*1503 | 1656 |
|  |  | p17 | 77 | 85 | SLYNTVATL | A\*0202 | 443 |
|  |  | Integrase | 263 | 271 | RKAKIIRDY | B\*1503 | 1111 |
| 25122 | A\*01:01, 03G1; B\*07G1, 37:01; C\*06:02, 07:02 | Nef | 73 | 82 | QVPLRPMTYK | A\*0301  | 161 |
|  |  | Nef | 105 | 115 | [KRQDILDLWVY](http://www.hiv.lanl.gov/content/immunology/ctl_search?results=Search;id=56248) | Cw7  | 57 |
|  |  | p17 | 71 | 79 | [GTEELRSLY](http://www.hiv.lanl.gov/content/immunology/ctl_search?results=Search;id=53820) | A\*0101  | 150 |
|  |  | p17 | 87 | 95 | **CVHQRIEVK** | A\*0301 | 270 |
|  |  | gp160 | 32 | 40 | **EKLWVTVYY** | Cw\*0602,0702 | 72 |
|  |  | RT | 206 | 214 | **REHLLRWGF** | Cw\*0702 | 937 |
| 53653 | A\*24G1, 24G1; B\*35:02, 38:01; C\*04G1, 12:03 | Vif | 94 | 102 | **YSTQVDPDL** | B\*3801 | 725 |
|  |  | Nef | 134 | 143 | RYPLTFGWCF | A\*2402 | 895 |
|  |  | gp160 | 78 | 86 | DPNPQEVVL | B\*3502 | 460 |
|  |  | gp160 | 44 | 52 | **VWKDAETTL**  | B\*3801 | 458 |
|  |  | gp160 | 621 | 629 | **EIWDNMTWL** | B\*3801  | 148 |
| 51314 | A\*03G1, 11:01; B\*35:03, 53:01; C\*04G1, 04G1 | gp160 | 621 | 629 | **DIWDNMTWM** | B\*3503 | 515 |
|  |  | Nef | 73 | 82 | QVPLRPMTYK | A\*0301 | 285 |
|  |  | Nef | 84 | 92 | AVDLSHFLK | A\*0301/A\*1101 | 463 |
|  |  | Nef | 135 | 143 | YPLTFGWCF | B\*5301 | 169 |
|  |  | Nef | 175 | 183 | **DPEKEVLVW** | B\*5301 | 366 |
|  |  | p17 | 20 | 28 | RLRPGGKKK | A\*0301 | 487 |
|  |  | p17 | 83 | 91 | ATLYCVHQR  | A\*1101 | 1148 |
|  |  | p24 | 195 | 202 | NPDCKTIL | B\*3503 | 87 |
|  |  | p24 | 176 | 184 | QASQDVKNW | B\*5301 | 183 |
|  |  | RT | 269 | 277 | QIYAGIKVK | A\*03/A\*1101 | 243 |
|  |  | RT | 520 | 528 | QIIEQLIKK | A\*1101 | 88 |
| 44149 | A\*01:01, 24G1; B\*07G1, 08G1; C\*07G1, 07:02 | gp160 | 843 | 851 | IPRRIRQGL | B\*0702  | 717 |
|  |  | p17 | 93 | 101 | DIKDTKEAL | B\*0801 | 2316 |
|  |  | p24 | 195 | 202 | NPDCKTIL | B\*0801 | 83 |
|  |  | p24 | 128 | 135 | EIYKRWII | B\*0801 | 381 |
|  |  | Nef | 90 | 97 | FLKEKGGL | B\*0801 | 100 |
|  |  | Nef | 134 | 143 | RYPLTFGWCF | A\*2402  | 510 |
|  |  | RT | 156 | 164 | SPAIFQSSM | B7 | 917 |
|  |  | RT | 18 | 26 | GPKVKQWPL | B\*0801 | 97 |
| 10849 | A\*23:01, 68G1; B\*3543, 57:03; C\*01:02, 02:10 | Vif | 61 | 69 | **DARLVITTY** | B\*3501 | 277 |
|  |  | gp160 | 217 | 226 | YCAPAGFAIL | Cw\*0102 | 322 |
|  |  | p24 | 15 | 23 | ISPRTLNAW | B\*5701 | 644 |
|  |  | p24 | 145 | 153 | YSPVSILDI | Cw\*0102  | 252 |
|  |  | p24 | 30 | 40 | KAFSPEVIPMF | B\*5701 | 697 |
|  |  | p24 | 108 | 117 | TSTLQEQIAW | B\*5703 | 812 |
|  |  | Nef | 116 | 124 | HTQGYFPDW | B57 | 176 |
|  |  | Pol | 23 | 31 | **ANSPTSREL** |  | 286 |
|  |  | Integrase | 95 | 104 | QETAYFILKL | A\*6802 | 619 |
| 67200 | A\*01:01, 24G1; B\*08:01, 57:01; C\*06:02, 07G1 | gp160 | 805 | 814 | **QELKNSAVSL** | B\*0801 | 143 |
|  |  | p17 | 71 | 79 | GSEELRSLY | A\*0101/B\*57 | 983 |
|  |  | p17 | 87 | 95 | **CVHQRIEVK** | A\*02 | 633 |
|  |  | p24 | 128 | 136 | EIYKRWIIL | B\*0801 | 191 |
|  |  | p24 | 30 | 40 | KAFSPEVIPMF | B\*5701 | 716 |
|  |  | p24 | 108 | 117 | TSTLQEQIAW | B\*57 | 239 |
|  |  | p24 | 176 | 184 | QASQEVKNW | B\*5701 | 138 |
|  |  | Nef | 90 | 97 | FLKEKGGL | B\*0801 | 1198 |
|  |  | Nef | 116 | 124 | HTQGYFPDW | B\*5701 | 212 |
|  |  | RT | 127 | 135 | YTAFTIPSV | A2 | 1235 |
|  |  | RT | 244 | 252 | IVLPEKDSW | B\*5701 | 1306 |
|  |  | RT | 375 | 383 | ITTESIVIW | B\*5701 | 318 |
| 25327 | A\*23:01, 26:01; B\*38:01, 38:01; C\*04G1, 12:03 | Vif | 79 | 87 | [WHLGQGVSI](http://www.hiv.lanl.gov/content/immunology/ctl_search?results=Search;id=52563)  | B\*3801 | 2132 |
|  |  | gp160 | 104 | 112 | MHEDIISLW | B\*3801 | 1497 |
|  |  | gp160 | 53 | 61 | **FCASDAKSY** | **B\*3801** | 650 |
|  |  | Integrase | 11 | 19 | **EHEKYHNNW** | **B\*3801** | 973 |
|  |  | Integrase | 95 | 103 | **QETAYFILK** |  | 375 |
|  |  | Integrase | 227 | 235 | **YRDSRDPLW** | **B\*3801** | 1020 |
| 63794 | A\*02G1, 02G1; B\*07G1, 44G1; C\*07:02, 07:04 | Nef | 77 | 85 | RPMTYKAAV |  | 698 |
|  |  | Nef | 68 | 76 | [FPVKPQVPL](http://www.hiv.lanl.gov/content/immunology/ctl_search?results=Search;id=54548) | B7 | 1648 |
|  |  | p2p7p1p6 | 70 | 77 | FLGKIWPS | A\*0201 | 388 |
|  |  | p24 | 223 | 231 | GPSHKARVL | B\*0702 | 2035 |
|  |  | gp160 | 200 | 208 | **AITQACPKV** | **A\*02G1** | 1608 |
|  |  | gp160 | 298 | 307 | RPNNNTRKSI | B\*0702 | 303 |
|  |  | gp160 | 843 | 851 | IPRRIRQGL | B\*0702 | 3170 |
|  |  | gp160 | 700 | 708 | AVLSVVNRV | A2 | 125 |
|  |  | RT | 156 | 164 | SPAIFQSSM | B7 | 1310 |
| 79379 | A\*01:01, 01:01; B\*27:05, 52:01; C\*02:01, 12:02 | Tat | 39 | 47 | **ITKGLGISY** |  | 305 |
|  |  | Vpr | 31 | 39 | VRHFPRIWL | B27 | 73 |
|  |  | gp160 | 836 | 845 | **IGRAILHIPR** | B\*2705 | 638 |
|  |  | p17 | 19 | 27 | IRLRPGGKK | B\*2705 | 145 |
|  |  | p17 | 71 | 79 | [GTEELRSLY](http://www.hiv.lanl.gov/content/immunology/ctl_search?results=Search;id=53820) | A\*0101  | 1308 |
|  |  | p17 | 81 | 89 | **TVATLYCVH** |  | 345 |
|  |  | p24 | 131 | 140 | KRWIILGLNK | B\*2705 | 2265 |
|  |  | Integrase | 186 | 194 | KRKGGIGGY | B\*2705 | 1238 |
| 57604 | A\*02G1, 02G1; B\*13:02, 35G1; C\*04:19, 06:02 | Vif | 61 | 69 | **DARLVITTY** | B\*35 | 173 |
|  |  | Nef | 74 | 81 | VPLRPMTY | B\*3501 | 430 |
|  |  | Nef | 128 | 137 | TPGPGVRYPL | B35 | 1233 |
|  |  | Nef | 134 | 143 | RYPLTFGWCF | B\*35 | 80 |
|  |  | p17 | 77 | 85 | SLYNTVATL | A\*0201 | 683 |
|  |  | p24 | 122 | 130 | PPIPVGDIY | B\*3501 | 210 |
|  |  | p24 | 8 | 17 | **GQMVHQAISP** |  | 2090 |
| 98621 | A\*02G1, 02G1; B\*40G1, 44G1; C\*03:04, 05:01 | Vif | 102 | 111 | **LADQLIHLYY**  |  | 493 |
|  |  | Vpr | 12 | 20 | **REPYNEWTL** | B\*4002 | 731 |
|  |  | Nef | 92 | 100 | KEKGGLEGL | B\*4001 | 1290 |
|  |  | Nef | 83 | 91 | AAVDLSHFL | A\*02 | 100 |
|  |  | p17 | 119 | 128 | **AADTGNSSQV** |  | 1186 |
|  |  | p24 | 174 | 184 | AEQASQDVKNW | B\*4402  | 3836 |
|  |  | p24 | 118 | 126 | MTSNPPIPV | A\*0201 | 178 |
|  |  | p2p7p1p6 | 70 | 77 | FLGKIWPS | A\*201 | 561 |
|  |  | p6 | 33 | 41 | KELYPLASL | B\*40 | 298 |
|  |  | RT | 181 | 189 | YQYMDDLYV | A\*0201 | 518 |
|  |  | gp160 | 805 | 814 | QELKNSAVSL | B\*4001 | 1348 |
| 32645 | A\*01:01, 02G1; B\*08:01, 35:03; C\*04G1, 07G1 | Vif | 101 | 109 | GLADQLIHL | A\*0201 | 2047 |
|  |  | Vif | 149 | 157 | ALAALITPK | A2 | 929 |
|  |  | Nef | 90 | 97 | FLKEKGGL | B\*0801 | 3181 |
|  |  | p17 | 77 | 85 | SLYNTVATL | A\*0201 | 194 |
|  |  | p17 | 74 | 82 | ELRSLYNTV | B\*0801 | 626 |
|  |  | p24 | 128 | 135 | EIYKRWII | B\*0801 | 184 |
|  |  | p2 | 6 | 14 | **QVTNSATIM** | B\*3503 | 76 |
|  |  | RT | 107 | 115 | TVLDVGDAY | B\*35 | 1129 |
|  |  | Rnase | 21 | 29 | **RGRQKVVSL** | B\*0801 | 111 |
|  |  | gp160 | 78 | 86 | DPNPQEVVL | B\*35 | 2346 |
|  |  | gp160 | 606 | 614 | TAVPWNASW | B\*35 | 2300 |
|  |  | gp160 | 848 | 856 | **RQGFERALL** | A\*02 | 1423 |
|  |  | gp160 | 814 | 822 | **LLNTTAIVV**  | A\*0201 | 362 |
| 51729 | A\*24G1, 24G1; B\*35:05, 40G1; C\*03:04, 04G1 | Vif | 116 | 124 | **SESAIRNAI** | B\*4002 | 367 |
|  |  | Vpr | 12 | 20 | **REPYNEWTL** | B\*4002 | 187 |
|  |  | Nef | 92 | 100 | KEKGGLEGL | B\*4001 | 3405 |
|  |  | Nef | 134 | 143 | RYPLTFGWCF | A\*2402 | 904 |
|  |  | p17 | 28 | 36 | KYKLKHIVW | A\*2402 | 601 |
|  |  | p17 | 92 | 101 | IEIKDTKEAL | B\*4001 | 293 |
|  |  | p6 | 33 | 41 | KELYPLASL | B\*40 | 735 |
| 44091 | A\*02G1, 24G1; B\*44G1, 57:01; C\*05:01, 60:02 | Nef | 134 | 141 | RYPLTFGW | A\*2402 | 929 |
|  |  | Nef | 116 | 124 | HTQGYFPDW | B57 | 654 |
|  |  | p24 | 108 | 117 | TSTLQEQIGW | B\*5701 | 874 |
|  |  | RT | 244 | 252 | IVLPEKDSW | B\*5701 | 366 |
|  |  | Integrase | 114 | 123 | **HTDNGSNFTS** |  | 1146 |
|  |  | Integrase | 838 | 846 | STTVKAACW | B\*5701 | 274 |

**a**Epitope position (based on HXB2 amino acid sequence) in HIV-1 proteins

**b**Amino acid sequence of identified new T-cell epitopes, with newly defined HLA restriction shown in bold

**c**Predicted restricting HLA allele of the novel epitope, with the newly confirmed HLA restriction shown in bold

|  |
| --- |
| **TABLE S2. HIV-1 specific CD8+ T cell responses in early infection: comparison of CD8+ T cell epitope-conservation by different methods.** |
|  |  |  |  |  |  |  |  |  |  |
|  **Epitope Sequencea** | **HXB2 siteb**  | **HIV protein** | **bCSpc** | **mCSpd** | **bCSee** | **mCSef** | **bCSconseqg** | **mCSconseqh** |
| VINRVRQGY | 704 | 712 | Env | 0 | 0.042 | 0.342 | 0.437 | -0.463 | -0.368 |
| EIWDNMTWL | 621 | 629 | Env | 0 | 0.049 | 0.543 | 0.632 | 0.086 | 0.506 |
| VWKDAETTL  | 44 | 52 | Env | 0 | 0.077 | 0.149 | 0.432 | -0.465 | -0.443 |
| FCASDAKSY | 53 | 61 | Env | 0.005 | 0.086 | 0.096 | 0.156 | -0.503 | -0.617 |
| AITQACPKV | 200 | 208 | Env | 0.005 | 0.262 | 0.205 | 0.353 | -0.384 | -0.498 |
| IGRAILHIPR | 836 | 845 | Env | 0.015 | 0.023 | 0.648 | 0.729 | 0.022 | 0.308 |
| LLNTTAIVV  | 814 | 822 | Env | 0.015 | 0.061 | 0.572 | 0.726 | -0.135 | 0.559 |
| NEIWDNMTW | 620 | 628 | Env | 0.071 | 0.04 | 0.69 | 0.729 | 0.52 | 0.925 |
| DIWDNMTWM | 621 | 629 | Env | 0.086 | 0.09 | 0.543 | 0.632 | 0.086 | 0.506 |
| RQGFERALL | 848 | 856 | Env | 0.112 | 0.15 | 0.456 | 0.563 | -0.306 | 0.249 |
| EKLWVTVYY | 32 | 40 | Env | 0.218 | 0.073 | 0.313 | 0.386 | 0.014 | -0.339 |
| IVTRIVELL | 777 | 785 | Env | 0.335 | 0.101 | 0.575 | 0.792 | -0.213 | 0.206 |
| QELKNSAVSL | 805 | 814 | ENV | 0.563 | 0.157 | 0.452 | 0.644 | -0.381 | -0.18 |
| AVLSVVNRV | 700 | 708 | Env | 0.01 | 0.025 | 0.511 | 0.845 | -0.231 | -0.345 |
| RRGWEVLKY | 787 | 795 | Env | 0.147 | 0.049 | 0.479 | 0.984 | -0.062 | -0.17 |
| TAVPWNASW | 606 | 614 | Env | 0.289 | 0.075 | 0.375 | 0.458 | -0.242 | -0.245 |
| IPRRIRQGL | 843 | 851 | Env | 0.34 | 0.296 | 0.471 | 0.485 | -0.234 | 0.192 |
| ERYLRDQQL | 584 | 592 | Env | 0.34 | 0.176 | 0.222 | 0.351 | -0.407 | -0.343 |
| DPNPQEVVL | 78 | 86 | Env | 0.381 | 0.118 | 0.354 | 0.517 | -0.18 | -0.126 |
| GRRGWEALK | 786 | 794 | Env | 0.386 | 0.153 | 0.479 | 0.667 | -0.063 | 0.036 |
| RPNNNTRKSI | 298 | 307 | Env | 0.421 | 0.277 | 0.441 | 0.599 | -0.139 | -0.299 |
| RQGLERALL | 848 | 856 | Env | 0.533 | 0.33 | 0.456 | 0.563 | -0.306 | 0.249 |
| YCAPAGFAIL | 217 | 226 | Env | 0.599 | 0.464 | 0.182 | 0.196 | -0.369 | -0.589 |
| NVTENFNMW | 88 | 96 | Env | 0.721 | 0.66 | 0.166 | 0.18 | -0.424 | -0.589 |
| SFEPIPIHY | 209 | 217 | Env | 0.751 | 0.352 | 0.122 | 0.253 | -0.464 | -0.544 |
| MHEDIISLW | 104 | 112 | Env | 0.787 | 0.525 | 0.124 | 0.234 | -0.486 | -0.589 |
| FNCGGEFFY | 376 | 384 | Env | 0.802 | 0.512 | 0.147 | 0.226 | -0.441 | -0.595 |
| CVHQRIEVK | 87 | 95 | Gag | 0.152 | 0.094 | 0.359 | 0.674 | 0.688 | 0.908 |
| QVTNSATIM | 6 | 14 | Gag | 0.187 | 0.063 | 0.964 | 1.17 | 2.008 | 2.149 |
| AADTGNSSQV | 119 | 128 | Gag | 0.212 | 0.078 | 0.856 | 1.463 | 1.689 | 1.898 |
| ELYPLASLR | 119 | 127 | Gag | 0.247 | 0.079 | 0.514 | 0.702 | 0.868 | 1.161 |
| GELDRWEKI | 11 | 19 | Gag | 0.364 | 0.132 | 0.279 | 0.392 | 0.473 | 0.264 |
| RWEKIRLRP | 15 | 23 | Gag | 0.53 | 0.218 | 0.169 | 0.282 | 0.079 | 0.061 |
| GQMVHQAISP | 8 | 17 | Gag | 0.551 | 0.422 | 0.155 | 0.28 | -0.034 | 0.209 |
| TVATLYCVH | 81 | 89 | Gag | 0.551 | 0.604 | 0.184 | 0.253 | 0.1 | -0.012 |
| GTEELRSLY | 71 | 79 | Gag | 0.005 | 0.121 | 0.251 | 0.351 | 0.505 | 0.39 |
| DIKDTKEAL | 93 | 101 | Gag | 0.03 | 0.023 | 0.214 | 0.328 | 0.26 | 0.173 |
| TSTLQEQIAW | 108 | 117 | Gag | 0.066 | 0.244 | 0.154 | 0.253 | -0.118 | -0.016 |
| PPIPVGDIY | 122 | 130 | Gag | 0.086 | 0.241 | 0.109 | 0.187 | -0.323 | -0.222 |
| GPSHKARVL | 223 | 231 | Gag | 0.131 | 0.381 | 0.153 | 0.158 | -0.045 | -0.134 |
| MTSNPPIPV | 118 | 126 | Gag | 0.146 | 0.3 | 0.204 | 0.263 | -0.038 | -0.023 |
| IEIKDTKEAL | 92 | 101 | Gag | 0.167 | 0.103 | 0.201 | 0.32 | 0.189 | 0.128 |
| YSPVSILDI | 145 | 153 | Gag | 0.177 | 0.601 | 0.185 | 0.161 | -0.105 | -0.427 |
| ELRSLYNTV | 74 | 82 | Gag | 0.197 | 0.192 | 0.322 | 0.34 | 0.736 | 0.533 |
| KYKLKHIVW | 28 | 36 | Gag | 0.268 | 0.12 | 0.305 | 0.485 | 0.428 | 0.77 |
| AEQASQDVKNW | 174 | 184 | Gag | 0.278 | 0.162 | 0.164 | 0.23 | -0.08 | -0.193 |
| QASQDVKNW | 176 | 184 | Gag | 0.283 | 0.163 | 0.187 | 0.274 | 0.027 | -0.08 |
| GSEELRSLY | 71 | 79 | Gag | 0.288 | 0.124 | 0.251 | 0.351 | 0.505 | 0.39 |
| ATLYCVHQR  | 83 | 91 | Gag | 0.298 | 0.189 | 0.258 | 0.495 | 0.316 | 0.482 |
| SLYNTVATL | 77 | 85 | Gag | 0.318 | 0.326 | 0.286 | 0.287 | 0.57 | 0.25 |
| KELYPLASL | 33 | 41 | Gag | 0.455 | 0.174 | 0.462 | 0.748 | 0.588 | 1.158 |
| QASQEVKNW | 176 | 184 | Gag | 0.52 | 0.183 | 0.187 | 0.274 | 0.027 | -0.08 |
| QAISPRTLNAW | 13 | 23 | Gag | 0.54 | 0.425 | 0.15 | 0.208 | -0.03 | 0.056 |
| HPVHAGPI | 84 | 91 | Gag | 0.586 | 0.501 | 0.261 | 0.296 | 0.225 | 0.179 |
| TERQANFL | 64 | 71 | Gag | 0.601 | 0.792 | 0.25 | 0.193 | -0.251 | -0.44 |
| RLRPGGKKK | 20 | 28 | Gag | 0.646 | 0.446 | 0.16 | 0.253 | 0.013 | 0.039 |
| FLGKIWPS | 70 | 77 | Gag | 0.662 | 0.753 | 0.17 | 0.19 | -0.442 | -0.371 |
| ISPRTLNAW | 15 | 23 | Gag | 0.697 | 0.593 | 0.115 | 0.134 | -0.129 | -0.188 |
| TSTLQEQIGW | 108 | 117 | Gag | 0.727 | 0.397 | 0.154 | 0.253 | -0.118 | -0.016 |
| EIYKRWIIL | 128 | 136 | Gag | 0.773 | 0.528 | 0.137 | 0.165 | -0.142 | -0.261 |
| KRWIILGLNK | 131 | 140 | Gag | 0.823 | 0.866 | 0.108 | 0.092 | -0.29 | -0.487 |
| EIYKRWII | 128 | 135 | Gag | 0.864 | 0.559 | 0.099 | 0.155 | -0.345 | -0.33 |
| IRLRPGGKK | 19 | 27 | Gag | 0.869 | 0.822 | 0.063 | 0.102 | -0.37 | -0.348 |
| ETINEEAAEW | 71 | 80 | Gag | 0.869 | 0.372 | 0.065 | 0.137 | -0.413 | -0.45 |
| NPDCKTIL | 195 | 202 | Gag | 0.889 | 0.787 | 0.095 | 0.134 | -0.419 | -0.489 |
| DRFYKTLRAEQ | 166 | 176 | Gag | 0.899 | 0.396 | 0.069 | 0.143 | -0.51 | -0.515 |
| KAFSPEVIPMF | 30 | 40 | Gag | 0.97 | 0.84 | 0.027 | 0.096 | -0.536 | -0.514 |
| DPEKEVLVW | 175 | 183 | Nef | 0.181 | 0.036 | 0.347 | 0.659 | 0.176 | 0.648 |
| FPVKPQVPL | 68 | 76 | Nef | 0.115 | 0.086 | 0.126 | 0.143 | -0.39 | -0.507 |
| TPGPGVRYPL | 128 | 137 | Nef | 0.193 | 0.338 | 0.269 | 0.233 | 0.198 | -0.214 |
| AAVDLSHFL | 83 | 91 | Nef | 0.206 | 0.089 | 0.304 | 0.408 | 0.254 | 0.13 |
| RPMTYKAAV | 77 | 85 | Nef | 0.214 | 0.105 | 0.32 | 0.238 | 0.352 | -0.069 |
| WRFDSRLAF | 183 | 191 | Nef | 0.272 | 0.067 | 0.245 | 0.462 | -0.024 | -0.222 |
| AVDLSHFLK | 84 | 92 | Nef | 0.288 | 0.141 | 0.26 | 0.326 | -0.02 | -0.217 |
| KRQDILDLWVY | 105 | 115 | Nef | 0.383 | 0.224 | 0.207 | 0.227 | -0.158 | -0.325 |
| RYPLTFGWCF | 134 | 143 | Nef | 0.486 | 0.427 | 0.183 | 0.22 | -0.127 | -0.165 |
| YPLTFGWCF | 135 | 143 | Nef | 0.49 | 0.431 | 0.186 | 0.239 | -0.091 | -0.098 |
| RYPLTFGW | 134 | 141 | Nef | 0.547 | 0.582 | 0.201 | 0.187 | -0.041 | -0.254 |
| KEKGGLEGL | 92 | 100 | Nef | 0.679 | 0.497 | 0.159 | 0.167 | -0.324 | -0.513 |
| HTQGYFPDW | 116 | 124 | Nef | 0.741 | 0.359 | 0.125 | 0.185 | -0.388 | -0.239 |
| QVPLRPMTYK | 73 | 82 | Nef | 0.77 | 0.719 | 0.098 | 0.143 | -0.49 | -0.456 |
| FLKEKGGL | 90 | 97 | Nef | 0.778 | 0.908 | 0.09 | 0.077 | -0.491 | -0.715 |
| VPLRPMTY | 74 | 81 | Nef | 0.831 | 0.744 | 0.085 | 0.147 | -0.529 | -0.412 |
| REHLLRWGF | 206 | 214 | Pol | 0.021 | 0.135 | 0.226 | 0.392 | 0.659 | 0.773 |
| YTAFTIPSV | 127 | 135 | Pol | 0.113 | 0.122 | 0.093 | 0.148 | -0.047 | -0.124 |
| ANSPTSREL | 23 | 31 | Pol | 0.113 | 0.307 | 0.335 | 0.459 | 0.925 | 0.755 |
| EHEKYHNNW | 11 | 19 | Pol | 0.124 | 0.176 | 0.212 | 0.35 | 0.154 | 0.363 |
| RGRQKVVSL | 21 | 29 | Pol | 0.392 | 0.298 | 0.32 | 0.474 | 0.273 | 0.548 |
| QETAYFILK | 95 | 103 | Pol | 0.412 | 0.377 | 0.192 | 0.316 | 0.116 | 0.226 |
| LVSQIIEQL | 77 | 85 | Pol | 0.598 | 0.189 | 0.249 | 0.405 | 0.44 | 0.421 |
| HTDNGSNFTS | 114 | 123 | pol | 0.639 | 0.772 | 0.209 | 0.221 | -0.045 | -0.34 |
| YRDSRDPLW | 227 | 235 | Pol | 0.649 | 0.153 | 0.295 | 0.308 | 0.477 | -0.142 |
| QYALGIIQA | 60 | 68 | Pol | 0.887 | 0.898 | 0.114 | 0.154 | -0.325 | -0.442 |
| VRDQAEHLK | 165 | 173 | Pol | 0.887 | 0.62 | 0.158 | 0.269 | -0.376 | -0.149 |
| ITTESIVIW | 375 | 383 | Pol | 0.186 | 0.063 | 0.249 | 0.494 | 0.561 | 0.626 |
| QIYAGIKVK | 269 | 277 | Pol | 0.361 | 0.218 | 0.191 | 0.315 | 0.508 | 0.657 |
| EEMNLPGRW | 34 | 42 | Pol | 0.412 | 0.072 | 0.278 | 0.374 | 0.73 | 0.759 |
| QETAYFILKL | 95 | 104 | Pol | 0.412 | 0.377 | 0.18 | 0.298 | 0.054 | 0.149 |
| STTVKAACWW | 123 | 132 | Pol | 0.464 | 0.088 | 0.251 | 0.334 | 0.245 | 0.09 |
| STTVKAACW | 838 | 846 | Pol | 0.464 | 0.072 | 0.266 | 0.356 | 0.319 | 0.164 |
| EEHEKYHSNW  | 10 | 19 | Pol | 0.495 | 0.26 | 0.232 | 0.352 | 0.29 | 0.383 |
| IVLPEKDSW | 244 | 252 | Pol | 0.619 | 0.143 | 0.151 | 0.428 | 0.138 | 0.691 |
| SPAIFQSSM | 156 | 164 | Pol | 0.67 | 0.649 | 0.117 | 0.162 | -0.155 | -0.247 |
| RKAKIIRDY | 263 | 271 | Pol | 0.701 | 0.537 | 0.208 | 0.322 | 0.05 | 0.141 |
| QIIEQLIKK | 520 | 528 | Pol | 0.722 | 0.444 | 0.192 | 0.401 | 0.091 | 0.207 |
| LVGPTPVNI | 76 | 84 | Pol | 0.784 | 0.829 | 0.09 | 0.085 | -0.127 | -0.348 |
| KRKGGIGGY | 186 | 194 | Pol | 0.814 | 0.862 | 0.2 | 0.202 | -0.011 | -0.322 |
| YQYMDDLYV | 181 | 189 | Pol | 0.918 | 0.945 | 0.052 | 0.056 | -0.308 | -0.477 |
| TVLDVGDAY | 107 | 115 | Pol | 0.959 | 0.951 | 0.03 | 0.044 | -0.444 | -0.526 |
| GPKVKQWPL | 18 | 26 | Pol | 0.959 | 0.865 | 0.02 | 0.069 | -0.431 | -0.347 |
| RPAEPVPLQL | 66 | 75 | Rev | 0.251 | 0.157 | 0.394 | 0.436 | 0.242 | 0.211 |
| ITKGLGISY | 39 | 47 | Tat | 0.244 | 0.116 | 0.352 | 0.405 | 0.026 | -0.356 |
| DARLVITTY | 61 | 69 | Vif | 0.159 | 0.068 | 0.467 | 0.615 | 0.418 | 0.53 |
| YSTQVDPDL | 94 | 102 | Vif | 0.239 | 0.114 | 0.434 | 0.478 | 0.212 | -0.071 |
| SESAIRNAI | 116 | 124 | Vif | 0.294 | 0.116 | 0.463 | 0.5 | 0.199 | 0.019 |
| LADQLIHLYY  | 102 | 111 | Vif | 0.486 | 0.209 | 0.325 | 0.408 | -0.04 | -0.087 |
| ALAALITPK | 149 | 157 | Vif | 0.205 | 0.072 | 0.476 | 0.628 | 0.254 | 0.162 |
| GLADQLIHL | 101 | 109 | Vif | 0.486 | 0.059 | 0.292 | 0.362 | -0.155 | -0.224 |
| WHLGQGVSI  | 79 | 87 | Vif | 0.865 | 0.471 | 0.172 | 0.256 | -0.486 | -0.513 |
| REPYNEWTL | 12 | 20 | Vpr | 0.613 | 0.497 | 0.209 | 0.302 | 0.024 | 0.004 |
| VRHFPRIWL | 31 | 39 | Vpr | 0.333 | 0.121 | 0.289 | 0.277 | 0.148 | -0.005 |

a Amino acid sequence of identified new T-cell epitopes, with newly defined HLA restriction shown in bold

b Epitope position (based on HXB2 amino acid sequence) in HIV-1 proteins

c bCSp is defined as the frequency of the exact epitope matches in aligned clade-B sequences

d mCSp is defined as the frequency of the exact epitope matches in aligned group-M sequences

e bCSe is defined as the average entropy score of all the positions along the epitope based on aligned clade-B sequences.

f mCSe is defined as the average entropy score of all the positions along the epitope based on aligned group-M sequences

g bCSconseq is defined by averaging conseq scores computed from a phylogenetic tree built over a set of aligned group-B sequences.

h bCSconseq is defined by averaging conseq scores computed from a phylogenetic tree built over a set of aligned group-M sequences.