Preferential Targeting of Conserved Gag Regions after Vaccination with a Heterologous DNA prime - Modified Vaccinia Ankara (MVA) boost HIV-1 vaccine regimen

Asli Bauer[1,2], Lilli Podola[1,2], Philipp Mann[1,2], Marco Missanga[1], Antelmo Haule[1], Lwitiho Sudi[1], Charlotta Nilsson[5,6], Bahati Kaluwa[1†], Cornelia Lueer[1,2], Maria Mwakatima[1], Patricia J. Munseri[7], Leonard Maboko[1], Merlin L. Robb[8], Sodsai Tovanabutra[8], Gustavo Kijak[8], Mary Marovich[8], Sheena McCormack[9], Sarah Joseph[9], Eligius Lyamuya[4], Britta Wahren[11], Eric Sandström[10], Gunnel Biberfeld[5,11], Michael Hoelscher[2,3], Muhammad Bakari[4], Arne Kroidl[2,3] and Christof Geldmacher[2,3]

NIMR Mbeya Medical Research Center (MMRC), Mbeya, Tanzania[1]; Division of Infectious Diseases and Tropical Medicine, Medical Center of the University of Munich (LMU), Munich, Germany[2]; German Center for Infection Research (DZIF), partner site Munich, Munich, Germany[3]; Muhimbili University of Health and Allied Sciences (MUHAS), Dar es Salaam, Tanzania[4]; Public Health Agency of Sweden, Solna, Sweden[5]; Department of Laboratory Medicine, Karolinska Institutet, Huddinge, Sweden[6]; Department of Internal Medicine, Muhimbili University of Health and Allied Sciences, Dar es Salaam, Tanzania[7]; United States Military HIV Research Program (MHRP), Bethesda, USA[8]; MRC Clinical

trials Unit at UCL, London, United Kingdom[9]; Karolinska Institute, Stockholm, Sweden[10]; Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden[11]

Running title: Preferential targeting of conserved HIV Gag regions

†deceased

#Address correspondence to Christof Geldmacher, geldmacher@lrz.uni-muenchen.de

present address: Department for Infectious Diseases & Tropical Medicine, Medical Center of the University of Munich, LMU Leopoldstrasse 5, 80802 Munich, Germany

Tel: +40-89-2180 17614

Abstract: 250 words

Text: 3478 words

Abstract

Prime-boost vaccination strategies against HIV-1 often include multiple variants for a given immunogen for better coverage of the extensive viral diversity. To study the immunologic effects of this approach, we characterized breadth, phenotype, function and specificity of Gag-specific T cells induced by a DNA-prime Modified Vaccinia Ankara (MVA)-boost vaccination strategy, which uses mismatched Gag immunogens in the TamoVac 01 phase IIa trial. Healthy Tanzanian volunteers received three injections of the DNA-SMI vaccine encoding for a subtype B and AB-recombinant Gag_{p37} and two vaccinations with MVA-CMDR encoding subtype A Gag_{p55}. Gag-specific T-cell responses were studied in 42 vaccinees using fresh peripheral blood mononuclear cells. After the first MVA-CMDR boost, vaccine-induced IFN-γ^+ Gag-specific T cell responses were dominated by $CD4^+$ T cells (compared to $CD8^+$ T cells, $p<0.001$) that co-expressed IL-2 (66.4%) and/or TNFα (63.7%). A median of 3 antigenic regions were targeted with a higher median response magnitude to Gag_{p24} regions - more conserved between prime and boost - as compared to regions within Gag_{p15} (not primed) and Gag_{p17} (less conserved, both $p<0.0001$). Four regions within Gag_{p24} were each targeted by 45% to 74% of vaccinees upon restimulation with DNA-SMI-Gag matched peptides. The response rate to individual antigenic regions correlated with the sequence homology between the MVA and DNA Gag encoded immunogens

(p=0.04, r^2=0.47). In summary, after the first MVA-CMDR boost, the sequence-mismatched DNA-prime MVA-boost vaccine strategy induced a Gag-specific T cell response that was dominated by polyfunctional $CD4^+$ T cells and that targeted multiple antigenic regions within the conserved Gag_{p24} Protein.

Importance

Genetic diversity is a major challenge for the design of vaccines against variable viruses. While including multiple variants for a given immunogen in prime-boost vaccination strategies is one approach that aims to improve coverage for global virus variants, the immunologic consequences of this strategy have been poorly defined so far. It is unclear whether inclusion of multiple variants in prime-boost vaccination strategies improves recognition of variant viruses by T cells and by which mechanisms this would be achieved; either by improved cross-recogniton of multiple variants for a given antigenic region or rather through preferential targeting of antigenic regions more conserved between prime and boost. Engineering vaccines to induce adaptive immune responses that preferentially target conserved antigenic regions of viral vulnerability might facilitate better immune control after preventive and therapeutic vaccination for HIV and for other variable viruses.

86

87

Introduction

High antigenic variability of common viruses causing either chronic (e.g. Human Immunodeficiency Virus (HIV), Hepatitis C Virus (HCV)) or acute (e.g. Influenza Virus, Dengue Virus) disease complicates the design of efficacious vaccines. Vaccines against such variable viruses should optimally induce adaptive immune responses that target all variants of a vulnerable antigenic region to prevent infection or, if that cannot be achieved, at least facilitate immune control of viral replication to prevent disease progression. All virus proteomes, even the most variable ones, contain conserved regions, where functional constraints limit extensive sequence variability. Designing vaccines to focus immune recognition towards such conserved regions could be a viable strategy to improve vaccine efficacy against variable pathogens (1, 2).

HIV is a good example of a highly variable virus. It causes persistent infection and rapidly escapes the HIV envelope (Env)-specific antibody response. Extensive sequence variability and glycosylation of the Env protein complicate Env based vaccine design (3, 4). In contrast, sequence variability within the group-specific antigen (Gag) is more limited due to functional constraints, but still differs between the capsid protein p24, the matrix protein p17 and proteins of the p15-encoding region (nucleocapsid and virion assembly proteins). An extensive body of evidence supports the concept that vaccine-induction of Gag-

111 specific T-cell responses could be beneficial; Gag-specific T-cell
112 responses are linked to viral control during chronic HIV infection (5-10).
113 The breadth of Gag-recognition by $CD8^+$ T-cells is associated with
114 better viral control and slower disease progression (9, 10). Furthermore,
115 the protective mechanism mediated by HLA class I alleles has been
116 linked to $CD8^+$ T-cell recognition of defined Gag regions (10, 11).
117 Similarly, Gag-specific $CD4^+$ T-cell responses appear to contribute to
118 viral control, sharing similar features with the $CD8^+$ T-cell response (12-
119 14). In rhesus monkeys, vaccine induced Gag-specific T-cell responses
120 correlate with post-challenge immune control and prolonged survival
121 after SIV challenge (15). Together, these data support the concept that
122 induction of strong and broad Gag-specific T-cell responses targeting
123 common sequence variants could potentially improve HIV vaccine
124 efficacy against diverse HIV variants.

125 The multi-subtype TaMoVac DNA/MVA regimen used in the TaMoVac
126 01 phase IIa trial induced strong Gag-specific T-cell responses (16, 17).
127 This offered the opportunity to address the hypothesis that delivery of
128 non-identical, but related immunogens preferentially induces T cell
129 responses to antigenic regions more conserved between the
130 immunogens. We found that the sequence mismatched TaMoVac01
131 DNA-MVA vaccination induced broad recognition of conserved
132 antigenic regions within Gag_{p24}.

133

Results

Relative conservation of Gag p24 in a mixed subtype epidemic

In order to define the degree of conservation within Mbeyan HIV Gag sequences, we analyzed previously published sequences and determined the distribution of subtypes and recombinant forms (RF). Pure subtype C sequences were found most frequently (57.1%) followed by unique subtype A-containing Recombinant Forms (RFs) (20.9%), pure subtype A Gag sequences (19.8%) and CD RFs (2.2%, Figure 1). In order to estimate the variability at each Gag amino acid position, the Shannon entropy score, was then calculated (Fig. 1b). Gag p24 showed the highest conservation with 82.6% of aa positions with an entropy score below 0.5 compared to the less conserved p17 (55.7%) and the least conserved p15 (44.8%). Similarly, the median Shannon entropy score was lowest for p24 (0.14), followed by p17 (0.37) and p15 (0.49).

Preferential induction of T-cells targeting Gag_{p24} protein by sequence mismatched DNA/MVA vaccination

The DNA vaccine included two plasmids encoding for SMI-Gag_{p37} with an identical p17 subtype B sequence linked to subtype B or subtype A p24 sequences (Figure 2). The MVA-CMDR-Gag_{p55} boost encoded for a

155 subtype A Gag_{p55} and included the p15 region. Within the p17 region,
156 21.2% (28 of 132) of aa positions were mismatched between the DNA-
157 SMI prime and MVA-CMDR boost. The Gag_{p24} subtype A and subtype B
158 sequences included in the DNA-SMI prime differed from the MVA-
159 CMDR boost in 7.8% (18 mismatches in 231) and 11.7% (27 mismatches
160 in 231) of aa positions, respectively. The p15 region represented a 100%
161 mismatch to the DNA-SMI prime.

162

163 IFN-γ+ Gag-specific T cell responses were not detected during the pre-
164 vaccination visit upon MVA-CMDR-Gag_{p55} restimulation, but were
165 present in 15% of vaccinees upon DNA-SMI-Gag_{p37} restimulation (range
166 60-165 SFC/10^6 PBMC, data not shown). Gag-specific T cell numbers
167 already peaked after the first MVA-CMDR boost with a median of 228
168 SFC/10^6 PBMC and significantly lower numbers of Gag-specific T cells
169 after the second boost ($p<0.05$ for both peptide pools). Magnitude of T
170 cell responses against the control peptide pools CEF (median 83 and
171 90 SFC/10^6 PBMC, $p=0.58$) and CMVpp65 (median 1243 and 938
172 SFC/10^6 PBMC, $p=0.2$) were comparable for both visits.

173 Because vaccine-induced IFN-γ+ Gag-specific T cell responses peaked
174 after the first MVA-CMDR boost, we focus on this time point in the
175 subsequnt analyses. ICS data from 41 vaccine recipients were eligible
176 for further analyses. IFN-γ+ Gag-specific CD4 and $CD8^+$ T cell responses

were detected in 30 and 14 of 41 subjects, respectively after restimulation with the MVA-CMDR-Gag_{p55} peptide pool (Fisher`s Exact test; p=0.0008). Likewise, significantly higher frequencies of $CD4^+$ as compared to $CD8^+$ IFN-γ+ Gag-specific T cells were detected after re-stimulation with either the MVA-CMDR-Gag_{p55} (median; 0.04% versus 0.01%; p<0.0001, figure 3A) or the DNA-SMI-Gag_{p37} (median; 0.06% v 0.01% p=0.0002) peptide pool after the first MVA-CMDR boost. IFN-γ+ Gag-specific CD4 T cells frequently co-expressed IL-2 (66.4%, mean of both Gag peptide pools) or TNF-α (63.7%) with 49.9% co-expressing IL-2 and TNF-α and high concordance between the two tested Gag peptide pools (Figure 3B). The CCR5-ligand Mip-1β and degranulation marker CD107 were co-expressed by only 20% (CMDR-Gag_{p55}: 13%) and 11% (CMDR-Gag_{p55}: 7%) of IFN-γ+ Gag-specific $CD4^+$ T cells, respectively. By contrast high frequencies of IFN-γ+ $CD4^+$ T cells co-expressed Mip-1β (59%) and/or CD107 (53%) after re-stimulation with CMV_{pp65} peptides, including a significant proportion (16%) of polyfunctional cells expressing all assessed functional markers, which is concordant with previous reports (18). Almost 75% of IFN-γ+ Gag-specific CD8 T cells co-expressed Mip-1β and almost 50% were CD107+ indicating degranulation of cytotoxic granula.

Breadth and specificity of Gag-specific T cell recognition were determined in 42 vaccine recipients using the IFN-γ-ELISPOT on freshly

200 isolated PBMCs, after re-stimulation with linear consecutive peptide
201 pools matching for MVA-CMDR-Gag$_{p55}$ (n=9 pools) and DNA-SMI-
202 Gag$_{p37}$ (n=7 pools). Median breadth of Gag recognition was 3 peptide
203 pools (range 0 - 8, Fig. 3c). The majority of the 42 vaccine recipients
204 responded to at least one DNA-SMI-Gag$_{p37}$ (n=36, 85.7%) or MVA-
205 CMDR-Gag$_{p55}$ peptide pool (n=33, 78.6%). In order to compare
206 magnitude of T cell responses against the different Gag regions and
207 account for sequence lengths of p17, p24, and p15, we calculated
208 average SFC values as per 15_mer peptide for each Gag region.
209 Antigenic regions within p24 were recognized at a higher median
210 magnitude (14.55 SFC/10^6 PBMC/15mer_peptide) compared to those
211 in the more variable p17 and p15 regions (both 6.52 SFC/10^6
212 PBMC/15mer_peptide, $p<0.0001$, Fig. 4a). The magnitude of response
213 to p17 (median of 6.52 SFC/10^6 PBMC/15mer_peptide for MVA-CMDR
214 and 6.0 SFC/10^6 PBMC/15mer_peptide for DNA-SMI) was similar to the
215 response to p15 ($p=0.95$).

216 Ten of the 42 vaccine recipients (23.8%) mounted SMI Gag$_{p17}$ specific T-
217 cell responses and 36 volunteers (85.7%) responded to SMI-Gag$_{p24}$.
218 Eight vaccine recipients (19.0%) responded to CMDR-Gag$_{p17}$, 32
219 (76.2%) to CMDR-Gag$_{p24}$, and 8 (19.0%) to CMDR-Gag$_{p15}$. Recognition
220 of individual antigenic regions within Gag$_{p24}$ was comparable between
221 DNA and MVA encoded variants (Fig. 4b). Pools 5 (HxB aa 225-290)
222 and 6 (aa 275-347) were immunodominant with 57.1% (CMDR-Gag

peptides: 47.6%) and 73.8% (CMDR-Gag peptides: 57.1%) responders after restimulation with the SMI Gag peptides, respectively. Pool 5 was targeted with a median magnitude of 38.65 SFC/10^6 PBMCs/15mer (SMI Gag peptides) and 35.19 SFC/10^6 PBMCs/15mer (CMDR-Gag peptides), respectively. Pool 6 was targeted with a median magnitude 27.74 SFC/10^6 PBMCs/15mer (SMI-Gag peptides) and 29.79 SFC/10^6 PBMCs/15mer (CMDR-Gag peptides).

Next we mapped individual MVA-CMDR-Gag_{p55} peptide responses (Figure 4c) in 39 vaccine recipients using cryopreserved PBMC. Results from 3 vaccine recipients were excluded because of invalid ELISpot results. Positive peptide-specific ELISpot responses were detected in 23 of 39 (60%) vaccine recipients. The 23 responders almost exclusively recognized peptides located within the p24 region. Four individual peptide responses were detected in p17 and two within p15. When counting responses to consecutive peptides as a single antigenic region, these 23 responders recognized a mean of 4 antigenic regions - ranging from one (26% of subjects (6/23)) to 12 antigenic regions recognized (4.3% (1/23)). The peptide YVDRFYKTLRAEQAT (pool 6, HxB position aa 296-310) was the most frequently recognized with 39.13% (9 of 23) responders, followed by the peptide YKRWIILGLNKIVRMY (pool 5, aa 262-277, 30.43% responders). Four peptides within p24 were recognized with an equal frequency of 26.09%,; GATPQDLNMMLNIVGG (pool 4, aa 178-193), IAGTTSTLQEQIGWMT (pool 5, aa 236-251),

246 ILGLNKIVRMYSPVSI(9) (pool 5, aa 267-282) and WMTETLLVQNANPDCK
247 (pool 6, aa 316-331). As shown in Figure 4d, the six most frequently
248 recognized antigenic regions had only 0 or 1 aa difference between
249 the SMI-DNA-prime encoded Gag_{p37} sequences (DNA B or BA) and the
250 MVA-CMDR boost encoded Gag_{p55} sequences. These data indicate
251 that hotspots of the vaccine-induced Gag recognition by T cells were
252 exlusively located within p24, whereas recognition of antigenic regions
253 within p17 and p15 appeared attenuated.

254

255 We next compared the frequency of responders to a given antigenic
256 region with the region-specific amino acid mismatches between the
257 DNA- and the MVA-encoded Gag sequences using linear regression
258 analysis (Figure 5). Plotting the frequency of responders against the
259 sequence heterogeneity between the encoded aa sequence of the
260 DNA-SMI-Gag_{p37} prime and the MVA-CMDR-Gag_{p55} boost for each
261 linear peptide pool, we found a linear correlation between region-
262 specific aa mismatches and the frequency of responders to a given
263 antigenic region (r^2=0.69, p=0.04 (DNA-SMI-Gag_{p37}), Figure 5a) and
264 r^2=0.45, p=0.07 (MVA-CMDR-Gag_{p55}), suggesting that higher levels of
265 sequence conservation between DNA-SMI-Gag_{p37} prime and MVA-
266 CMDR-Gag_{p55} boost contribute to higher frequencies of recognition
267 within more conserved Gag regions.

268 The link between the level of sequence conservation within a given
269 antigenic region and its recognition could further be substantiated by
270 linear regression analysis. The number of aa mismatches between
271 priming DNA Gag_{p24} encoded subtype B sequence and boosting Gag
272 aa sequences (MVA) correlated with the frequency of responders
273 (Figure 5b, r^2=0.01 and p=0.0332) and the magnitude of response
274 (Figure 5c, p=0.04, r^2=0.04) to a given Gag_{p24} peptide. These data
275 support the concept that the sequence heterologous prime-boost
276 vaccination strategy applied during TaMoVac 01 contributed to
277 preferential recognition of the more conserved antigenic regions within
278 Gag.

279

280

281

Discussion

Genetic diversity poses a major challenge to the design of efficacious vaccines against HIV-1 and other variable viruses. Several vaccination strategies have been tested to address this extensive diversity of HIV-1 (2). The TaMoVac 01 study incorporated multiple variants of the same immunogen in the vaccine formulation (16, 17, 19). The immunologic consequences of such vaccination strategies incorporating multiple variants and in particular the impact on the pattern of antigen recognition - have only been poorly defined so far. Inclusion of three immunogen variants elicited strong Gag-specific T-cell responses in the TaMoVac 01 and HIVIS studies (16, 17) providing the opportunity to study parameters that potentially influence the immunodominance pattern of vaccine-induced Gag recognition by T cells. The analysis of 91 Gag sequences confirmed previous reports that the capsid antigen p24 shows the highest degree of conservation followed by the less conserved p17 and p15 (20).

In TaMoVac 01 vaccinees, antigenic regions within conserved Gag_{p24} regions were preferentially targeted after the first MVA-CMDR boost. A comparative analysis of HIV-1 rAd-5 T-cell based vaccine clinical trials Merck16, HVTN 054 and HVTN 502/Step, which included closely related Gag immunogen sequences and did not include DNA priming, found

304 considerable variation in the recognition of Gag regions (21). HVTN 054
305 showed an accumulation of hotspots within Gag_{p24}, whereas hotspots
306 of T cell recognition were more evenly distributed within Gag for
307 Merck16 and HVTN 502/Step. In comparison, the sequence
308 mismatched DNA/MVA TaMoVac vaccine "*focused*" T-cell responses
309 even more on highly conserved Gag_{p24} regions. The strength of
310 induced T cell responses against individual antigenic regions in Gag_{p24}
311 correlated with the degree of conservation between the SMI-DNA
312 encoded subtype B and MVA-CMDR encoded subtype A sequence
313 variants. While no such correlation was detected for the DNA encoded
314 subtype A variant, the six most frequently targeted peptides showed
315 only none or one aa mismatch when taking both DNA encoded
316 Gag_{p24} variants into account. Hence, immune recognition of peptides
317 with high degree of variability between MVA-CMDR and DNA-$Gag_{p37.1}$
318 (subtype A, Figure 2) were probably primed by the other DNA- Gag_{p37}
319 variant (subtype B).

320 What is the mechanism underlying this preferential targeting of more
321 conserved antigenic regions within Gag_{p24} in TaMoVac01 vaccinees?
322 One possibility is that induction of T-cell responses towards more
323 variable epitopes - more common in Gag_{p17} and Gag_{p15} - was
324 negatively affected by increasing numbers of mismatched amino acid
325 positions within a given epitope. Immunodominat Gag-specific $CD8^+$ T-
326 cell populations targeting the epitope variants TL9M7 (TPQDLNMML)

and TL9T7 (TPQDLNTML) during natural HIV subtype A and subtype C infections are completely different in their clonotypic composition(22). It is therefore plausible that certain T-cell clonotypes induced by one immunogen sequence cannot be boosted by certain other sequence variants or that boosting is suboptimal; however, if T-cell clonotypes partially overlap in their recognition of the two epitope variants, these cross-reactive clonotypes would be boosted stronger, resulting in improved variant cross-recognition and *"focusing" towards* such cross-reactive clonotypes. Our data show that recognition of epitopes with more than 2 mismatches is completely abrogated (Figure 5 C, D). A high number of epitope variant mismatches is therefore likely to abolish any efficient boosting. Of note, based on our analyses of 91 Mbeyan sequences, the tested CMDR-Gag sequence variants for the peptides P1 to P6 always closely matched the most frequent Mbeyan variants with ≤2 amino acid substitutions. In summary a variety of mechanisms probably contributed to preferential T cell recognition of more conserved Gag regions in the TaMoVac01 study using mismatched immunogen sequences.

All immunodominant regions recognized by the TaMoVac 01 vaccinees have also been identified in previous studies. P1 (YVDRFYKTLRAEQAT) is an immunodominant target for $CD4^+$ and $CD8^+$ T-cells during early and chronic infection (14, 23, 24) and was also a

hotspot in the HVTN502/Step and HVTN 054 trials (21). P2 (YKRWIILGLNKIVRMY) is frequently targeted by $CD4^+$ T cells during early and chronic HIV infection (14) and was a hotspot in the Step trial (21)(39). P3 (GATPQDLNMMLNIVGG) contains the highly immunodominant B42/B81 restricted TL9 epitope (22, 24-26). P4 to P6 also contained previously described epitopes (14, 23). Inherent immunogenicity of these peptide sequences P1-P6 could also contribute to their preferential recognition as these were also frequently recognized by T cells in natural infection or in other trials.

Most of individual peptide-specific T cell responses after the first MVA-CMDR boost were most likely mediated by $CD4^+$ compared to $CD8^+$ T cells, even though we were not able to phenotype individual vaccine-induced peptide-specific T-cell responses from cryopreserved PBMC. After the second MVA-CMDR boost, 73% (32 of 44) of vaccinees had higher frequencies of Gag-specific IFN-γ+ $CD8^+$ T-cell as compared to after the first MVA boost, suggesting that optimal induction of Gag-specific CD8 T cell responses probably *need* a second MVA boost.

One limitation of this study is that cryopreservation negatively affected our ability to detect vaccine-induced T cell responses. We did not observe this phenomenon during previous studies for natural HIV- or

MTB-specific T-cell responses(22, 27). However, this phenomenon is not unknown (28).

In conclusion, our results show that after one MVA-CMDR boost the sequence mismatched TaMoVac 01 DNA prime/MVA boost vaccine regimen induced Gag specific T-cell responses that were dominated by $CD4^+$ T cells co-expressing IL-2 and TNF-α and targeted multiple conserved antigenic regions within Gag_{p24}.

Methods

Study Design and samples

The TaMoVac 01 phase 2a trial was described in detail, previously (16). Briefly, this randomized controlled trial was performed at Muhimbili University of Health and Allied Science (MUHAS) and the National Institute for Medical Research -Mbeya Medical Research Center (NIMR-MMRC) in Mbeya, Tanzania with a total of 120 healthy, HIV-negative individuals, aged 18-40 years. Immunosuppressive medications were exclusion criteria. The TaMoVac 01 trial participants received two or five intradermal (ID) injections of DNA/Placebo at weeks 0, 4 and 12 and were boosted with two MVA-CMDR/Placebo intramuscular (IM) injections at weeks 30 and 46. The DNA/Placebo was administered intradermally in the skin over both deltoids by using a Biojector 2000® needleless device (Bioject Medical Technologies, Inc., Tualatin, OR, USA), the MVA injections were administered into the left deltoid muscle.

The DNA-SMI vaccine (Vecura, Huddinge, Stockholm Sweden) was composed of 7 plasmids encoding the HIV-1 genes Env (subtype A, B and C respectively), Rev clade B, Gag (subtype A and A/B) and reverse transcriptase subtype B (for details see (29)) and was administered at either 600µg or 1000µg total. The recombinant modified vaccinia virus Ankara (MVA) expressing HIV-1 gp150 clade E as well as gag and pol clade A (MVA-CMDR) was manufactured by

Walter Reed Army Institute of Research (WRAIR) (30) and administered at 10^8 pfu.

The study was reviewed and approved by the National Ethics Committee, the Institutional Review Board (IRB) at the NIMR-MMRC and the MUHAS IRB in compliance with national guidelines and institutional policies (Clinical Trials Registration: ATM2010050002122368), and informed consent was obtained in accordance with the Declaration of Helsinki.

In the present study, samples collected at two weeks after the first MVA-CMDR vaccination - when vaccine-induced T cell responses peaked - from 42 vaccine recipients at the National Institute for Medical Research – Mbeya Medical Research Center and who did not become HIV infected during the trial were used for analyses.

HIV genetic sequence analyses

HIV-1 Gag sequences included in the phylogenetic analyses were from 91 HIV-positive subjects from the Mbeya region and have the Genebank accession numbers FJ853501 to FJ85359 (22). The subtypes and recombinant forms were determined using the jpHMM-HIV tool (http://jphmm.gobics.de/submission_hiv).

423

424 **Peptide antigen and peptide pool design**

425 **DNA-SMI-Gag$_{p37}$ and MVA-CMDR-Gag$_{p55}$ peptide sets consisted of 15-**
426 **to 18-mer peptides overlapping by 11 amino acids had a purity of**
427 **>80% (JPT peptide technologies, Berlin). Individual peptide variants of**
428 **identical length for a given Gag region were used to allow for direct**
429 **comparison of T-cell responses targeting MVA or DNA-encoded**
430 **peptide variants and to prevent artifacts with linked to intrapeptide**
431 **epitope location (24, 31). Peptide maxipools including all peptides for**
432 **the DNA-SMI-Gag$_{p37}$ and MVA-CMDR-Gag$_{p55}$ immunogens were used**
433 **to study phenotypic and functional characteristics of vaccine-induced**
434 **Gag-specific T cell responses; Linear peptide pools were used to test T-**
435 **cell responses against distinct antigenic regions (Figure 2); nine pools**
436 **for MVA-CMDR-Gag$_{p55}$ and 7 pools for DNA-SMI-Gag$_{p37}$.**

437

438 **Intracellular cytokine staining (ICS)**

439 A 8-colour ICS assay was performed on fresh PBMC were stimulated in
440 the presence of Brefeldin A (5(g/ml Sigma-Aldrich) for 6 hours with
441 either MVA-CMDR-Gag$_{p55}$ or DNA-SMI-Gag$_{p37}$ maxipools
442 (1(g/ml/peptide), nothing (negative control) or the control antigens
443 (CMV$_{pp65}$ peptide pool (0.5(g/ml/peptide) and Staphylococcus

444 enterococcus Toxin B (1(g/ml, Sigma-Aldrich). Stimulated PBMCs were
445 then stained with α-CD3 APC-Cy7 (BD Biosciences Europe,
446 Erembodegem, Belgium), α-CD4 PerCp-Cy5.5 (ebioscience, San Diego,
447 CA), α-CD8 V500 (BD Bioscience), α–TNF-α Pe-Cy7 (BD), α-IFN-γ V450
448 (BD), α-IL-2 APC (BD), α-MIP-1β PE (BD), α-CD107 FITC (ebioscience).
449 Acquisition of samples was performed using a FACSCanto II flow
450 cytometer with acquisition-defined compensation using BD
451 CompBeads (BD). Flow Cytometry results were analysed using FlowJo
452 software, version 9.6.4 (Tree Star, Ashland, OR) and SPICE, version 5.35,
453 downloaded from http://exon.niaid.nih.gov (32). A minimum of 50.000
454 $CD3^+$ lymphocytes were required for a sample to be included in the
455 analysis. Background reactivity, defined by using unstimulated
456 negative controls, was subtracted for analyses of antigen-specific T cell
457 responses. IFN-γ+ T cell frequencies >0.025% were considered a positive
458 response.

459

460 **IFNγ ELISPOT assay**

461 The IFN-γ ELISPOT^{PLUS} kit (Mabtech, Nacka, Sweden) was used
462 according to the instructions of the manufacturer. Fresh Peripheral
463 Blood Mononuclear Cells (PBMC) were stimulated with linear peptide
464 pools matching DNA-SMI-Gag_{p37} and MVA-CMDR-Gag_{p55} (described
465 below, JPT, Berlin, Germany) (Figure 2). Frequencies of antigen-specific

466 spot-forming cells (SFC) were measured with an automated ELISPOT
467 reader (Immunospot, C.T.L., Bonn, Germany). Responses were
468 considered positive when the number of Spot Forming Cells (SFC) was
469 at least four times the medium control and >55 SFC/10^6 PBMCs (33).
470 Mapping of individual peptide responses was performed based on the
471 peptide pool matrix ELISPOT results and testing cryopreserved PBMC.
472 Responses of stimulated PBMCs with a SFC count higher than 2-fold the
473 unstimulated medium control and ≥50 SCF/10^6 PBMCs cells were
474 considered positive. Samples with a medium control of >60 SFC/PBMCs
475 were excluded from analyses.

476

477 **Statistical analyses**

478 Statistical analyses were performed using Prism version 6.0 (GraphPad,
479 Inc). Comparisons of two groups were performed with the Mann-
480 Whitney test. Magnitudes of CD4 and CD8 T cell responses were
481 performed with the Wilcoxon signed-rank test. Number of responders
482 with IFN-γ+ CD4 and CD8 T cell responses were compared using The
483 Fisher`s Exact test. The linear relationship between the number of
484 mismatched amino acid positions within a given antigenic region and
485 the corresponding frequency of recognition and response magnitude
486 was calculated using linear regression analyses. Tests used for statistical
487 analysis are mentioned in the figure legends.

Acknowledgements

We gratefully acknowledge funding from European and Developing Countries Clinical Trials Partnership (EDCTP) Grant CT.2006.33111.007 and by the German Ministry of Education and Research (BMBF, Grant ID: 01KA0802) and by the German Center for Infection Research (DZIF). This work was supported by the Medical Research Council (MRC_UU_12023/23)

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication

We declare no conflict of interests.

References

1. Hanke T. 2014. Conserved immunogens in prime-boost strategies for the next-generation HIV-1 vaccines. Expert Opin Biol Ther 14:601-16.
2. Barouch DH, Korber B. 2010. HIV-1 vaccine development after STEP. Annu Rev Med 61:153-67.
3. Wang W, Nie J, Prochnow C, Truong C, Jia Z, Wang S, Chen XS, Wang Y. 2013. A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. Retrovirology 10:14.
4. Raska M, Czernekova L, Moldoveanu Z, Zachova K, Elliott MC, Novak Z, Hall S, Hoelscher M, Maboko L, Brown R, Smith PD, Mestecky J, Novak J. 2014. Differential glycosylation of envelope gp120 is associated with differential recognition of HIV-1 by virus-specific antibodies and cell infection. AIDS Res Ther 11:23.
5. Goulder PJ, Walker BD. 2012. HIV and HLA class I: an evolving relationship. Immunity 37:426-40.
6. Rolland M, Heckerman D, Deng W, Rousseau CM, Coovadia H, Bishop K, Goulder PJ, Walker BD, Brander C, Mullins JI. 2008. Broad and Gag-biased HIV-1 epitope repertoires are associated with lower viral loads. PLoS One 3:e1424.
7. Borrow P, Lewicki H, Hahn BH, Shaw GM, Oldstone MB. 1994. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. J Virol 68:6103-10.
8. Chakraborty S, Rahman T, Chakravorty R. 2014. Characterization of the Protective HIV-1 CTL Epitopes and the Corresponding HLA Class I Alleles: A Step towards Designing CTL Based HIV-1 Vaccine. Adv Virol 2014:321974.
9. Kiepiela P, Ngumbela K, Thobakgale C, Ramduth D, Honeyborne I, Moodley E, Reddy S, de Pierres C, Mncube Z, Mkhwanazi N, Bishop K, van der Stok M, Nair K, Khan N, Crawford H, Payne R, Leslie A, Prado J, Prendergast A, Frater J, McCarthy N, Brander C, Learn GH, Nickle D, Rousseau C, Coovadia H, Mullins JI, Heckerman D, Walker BD, Goulder P. 2007. CD8+ T-cell responses to different HIV proteins have discordant associations with viral load. Nat Med 13:46-53.
10. Geldmacher C, Currier JR, Herrmann E, Haule A, Kuta E, McCutchan F, Njovu L, Geis S, Hoffmann O, Maboko L, Williamson C, Birx D, Meyerhans A, Cox J, Hoelscher M. 2007. CD8 T-cell recognition of multiple epitopes within specific Gag regions is associated with maintenance of a low steady-state viremia in human immunodeficiency virus type 1-seropositive patients. J Virol 81:2440-8.
11. Streeck H, Lichterfeld M, Alter G, Meier A, Teigen N, Yassine-Diab B, Sidhu HK, Little S, Kelleher A, Routy JP, Rosenberg ES, Sekaly RP, Walker BD, Altfeld M. 2007. Recognition of a defined region within p24 gag by CD8+ T cells during primary human immunodeficiency virus type 1 infection in individuals expressing protective HLA class I alleles. J Virol 81:7725-31.
12. Ranasinghe S, Cutler S, Davis I, Lu R, Soghoian DZ, Qi Y, Sidney J, Kranias G, Flanders MD, Lindqvist M, Kuhl B, Alter G, Deeks SG, Walker BD, Gao X, Sette A, Carrington M, Streeck H. 2013. Association of HLA-DRB1-restricted CD4(+) T cell responses with HIV immune control. Nat Med 19:930-3.
13. Ranasinghe S, Flanders M, Cutler S, Soghoian DZ, Ghebremichael M, Davis I, Lindqvist M, Pereyra F, Walker BD, Heckerman D, Streeck H. 2012. HIV-specific CD4 T cell responses to different viral proteins have discordant associations with viral load and clinical outcome. J Virol 86:277-83.
14. Schieffer M, Jessen HK, Oster AF, Pissani F, Soghoian DZ, Lu R, Jessen AB, Zedlack C, Schultz BT, Davis I, Ranasinghe S, Rosenberg ES, Alter G, Schumann RR, Streeck H. 2014.

Induction of Gag-specific CD4 T cell responses during acute HIV infection is associated with improved viral control. J Virol 88:7357-66.

15. Liu J, O'Brien KL, Lynch DM, Simmons NL, La Porte A, Riggs AM, Abbink P, Coffey RT, Grandpre LE, Seaman MS, Landucci G, Forthal DN, Montefiori DC, Carville A, Mansfield KG, Havenga MJ, Pau MG, Goudsmit J, Barouch DH. 2009. Immune control of an SIV challenge by a T-cell-based vaccine in rhesus monkeys. Nature 457:87-91.

16. Munseri PJ, Kroidl A, Nilsson C, Joachim A, Geldmacher C, Mann P, Moshiro C, Aboud S, Lyamuya E, Maboko L, Missanga M, Kaluwa B, Mfinanga S, Podola L, Bauer A, Godoy-Ramirez K, Marovich M, Moss B, Hoelscher M, Gotch F, Stohr W, Stout R, McCormack S, Wahren B, Mhalu F, Robb ML, Biberfeld G, Sandstrom E, Bakari M. 2015. Priming with a simplified intradermal HIV-1 DNA vaccine regimen followed by boosting with recombinant HIV-1 MVA vaccine is safe and immunogenic: a phase IIa randomized clinical trial. PLoS One 10:e0119629.

17. Sandstrom E, Nilsson C, Hejdeman B, Brave A, Bratt G, Robb M, Cox J, Vancott T, Marovich M, Stout R, Aboud S, Bakari M, Pallangyo K, Ljungberg K, Moss B, Earl P, Michael N, Birx D, Mhalu F, Wahren B, Biberfeld G, Team HIVIS. 2008. Broad immunogenicity of a multigene, multiclade HIV-1 DNA vaccine boosted with heterologous HIV-1 recombinant modified vaccinia virus Ankara. J Infect Dis 198:1482-90.

18. Casazza JP, Betts MR, Price DA, Precopio ML, Ruff LE, Brenchley JM, Hill BJ, Roederer M, Douek DC, Koup RA. 2006. Acquisition of direct antiviral effector functions by CMV-specific CD4+ T lymphocytes with cellular maturation. J Exp Med 203:2865-77.

19. Kibuuka H, Kimutai R, Maboko L, Sawe F, Schunk MS, Kroidl A, Shaffer D, Eller LA, Kibaya R, Eller MA, Schindler KB, Schuetz A, Millard M, Kroll J, Dally L, Hoelscher M, Bailer R, Cox JH, Marovich M, Birx DL, Graham BS, Michael NL, de Souza MS, Robb ML. 2010. A phase 1/2 study of a multiclade HIV-1 DNA plasmid prime and recombinant adenovirus serotype 5 boost vaccine in HIV-Uninfected East Africans (RV 172). J Infect Dis 201:600-7.

20. Li G, Verheyen J, Rhee SY, Voet A, Vandamme AM, Theys K. 2013. Functional conservation of HIV-1 Gag: implications for rational drug design. Retrovirology 10:126.

21. Hertz T, Ahmed H, Friedrich DP, Casimiro DR, Self SG, Corey L, McElrath MJ, Buchbinder S, Horton H, Frahm N, Robertson MN, Graham BS, Gilbert P. 2013. HIV-1 vaccine-induced T-cell responses cluster in epitope hotspots that differ from those induced in natural infection with HIV-1. PLoS Pathog 9:e1003404.

22. Geldmacher C, Metzler IS, Tovanabutra S, Asher TE, Gostick E, Ambrozak DR, Petrovas C, Schuetz A, Ngwenyama N, Kijak G, Maboko L, Hoelscher M, McCutchan F, Price DA, Douek DC, Koup RA. 2009. Minor viral and host genetic polymorphisms can dramatically impact the biologic outcome of an epitope-specific CD8 T-cell response. Blood 114:1553-62.

23. Kaufmann DE, Bailey PM, Sidney J, Wagner B, Norris PJ, Johnston MN, Cosimi LA, Addo MM, Lichterfeld M, Altfeld M, Frahm N, Brander C, Sette A, Walker BD, Rosenberg ES. 2004. Comprehensive analysis of human immunodeficiency virus type 1-specific CD4 responses reveals marked immunodominance of gag and nef and the presence of broadly recognized peptides. J Virol 78:4463-77.

24. Geldmacher C, Currier JR, Gerhardt M, Haule A, Maboko L, Birx D, Gray C, Meyerhans A, Cox J, Hoelscher M. 2007. In a mixed subtype epidemic, the HIV-1 Gag-specific T-cell response is biased towards the infecting subtype. AIDS 21:135-43.

25. Goulder PJ, Brander C, Annamalai K, Mngqundaniso N, Govender U, Tang Y, He S, Hartman KE, O'Callaghan CA, Ogg GS, Altfeld MA, Rosenberg ES, Cao H, Kalams SA, Hammond M, Bunce M, Pelton SI, Burchett SA, McIntosh K, Coovadia HM, Walker BD. 2000. Differential narrow focusing of immunodominant human immunodeficiency virus gag-specific cytotoxic T-lymphocyte responses in infected African and caucasoid adults and children. J Virol 74:5679-90.

26. Addo MM, Yu XG, Rathod A, Cohen D, Eldridge RL, Strick D, Johnston MN, Corcoran C, Wurcel AG, Fitzpatrick CA, Feeney ME, Rodriguez WR, Basgoz N, Draenert R, Stone DR, Brander C, Goulder PJ, Rosenberg ES, Altfeld M, Walker BD. 2003. Comprehensive epitope analysis of human immunodeficiency virus type 1 (HIV-1)-specific T-cell responses directed against the entire expressed HIV-1 genome demonstrate broadly directed responses, but no correlation to viral load. J Virol 77:2081-92.

27. Portevin D, Moukambi F, Clowes P, Bauer A, Chachage M, Ntinginya NE, Mfinanga E, Said K, Haraka F, Rachow A, Saathoff E, Mpina M, Jugheli L, Lwilla F, Marais BJ, Hoelscher M, Daubenberger C, Reither K, Geldmacher C. 2014. Assessment of the novel T-cell activation marker-tuberculosis assay for diagnosis of active tuberculosis in children: a prospective proof-of-concept study. Lancet Infect Dis 14:931-8.

28. Smith SG, Smits K, Joosten SA, van Meijgaarden KE, Satti I, Fletcher HA, Caccamo N, Dieli F, Mascart F, McShane H, Dockrell HM, Ottenhoff TH, Group TTBW. 2015. Intracellular Cytokine Staining and Flow Cytometry: Considerations for Application in Clinical Trials of Novel Tuberculosis Vaccines. PLoS One 10:e0138042.

29. Brave A, Ljungberg K, Boberg A, Rollman E, Engstrom G, Hinkula J, Wahren B. 2006. Reduced cellular immune responses following immunization with a multi-gene HIV-1 vaccine. Vaccine 24:4524-6.

30. Earl PL, Cotter C, Moss B, VanCott T, Currier J, Eller LA, McCutchan F, Birx DL, Michael NL, Marovich MA, Robb M, Cox JH. 2009. Design and evaluation of multi-gene, multi-clade HIV-1 MVA vaccines. Vaccine 27:5885-95.

31. Draenert R, Brander C, Yu XG, Altfeld M, Verrill CL, Feeney ME, Walker BD, Goulder PJ. 2004. Impact of intrapeptide epitope location on CD8 T cell recognition: implications for design of overlapping peptide panels. AIDS 18:871-6.

32. Roederer M, Nozzi JL, Nason MC. 2011. SPICE: Exploration and analysis of post-cytometric complex multivariate datasets. Cytometry Part A 79 A:167-174.

33. Nilsson C, Aboud S, Karlen K, Hejdeman B, Urassa W, Biberfeld G. 2008. Optimal blood mononuclear cell isolation procedures for gamma interferon enzyme-linked immunospot testing of healthy Swedish and Tanzanian subjects. Clin Vaccine Immunol 15:585-9.

632 **Figure legends**

633 **Figure 1. Diversity of HIV-1 Gag protein sequences originating from the**
634 **Mbeya Region.**

635 The distribution of subtypes and unique recombinant forms of Gag
636 polyprotein sequences from 91 HIV infected subjects from the Mbeya
637 Region is shown in the pie chart (A). A Shannon Entropy Plot generated
638 from these Gag sequences is shown in (B).

639

640 **Figure 2. Immunogen sequences included in the DNA-Gag prime and**
641 **Modified Vaccinia Ankara (MVA)-Gag bost and their coverage by**
642 **peptide pools**

643 The seven DNA peptide pools covered the Gag_{p37} region consisting of
644 p17 (blue) and p24 (grey), whereas the nine MVA-Gag peptide pools
645 covered the Gag_{p55} precursor protein including the p15 region (red) in
646 addition to p17 and p24. The p15 region was only covered by MVA-
647 Gag peptide pools 8 and 9.

648

649 **Figure 3. Phenotype, function and breadth of vaccine-induced Gag-**
650 **specific T cell responses**

651 **Representative Dot Plots for the analyses of Gag-specific CD4 and CD8**
652 **T cell functions are shown in (A). Frequency of IFN-γ+ $CD4^+$ and $CD8^+$ T**
653 **cell after stimulation of freshly isolated PBMC with whole MVA-CMDR-**
654 **Gag_{p55} (left panel) or DNA-SMI-Gag_{p37} peptide pools (right panel) in 41**
655 **vaccinees are shown in (B). Co-expression of additional functions**
656 **(TNFα, Mip-1β, IL2 and the degranulation marker CD107) is shown in (C)**
657 **for IFNγ+CD4 (left) and IFNγ+$CD8^+$ (right) T cells. The four colour coded**
658 **arcs indicate the proportion of cells co-expressing the four additional**
659 **functions. The colour-coded pies symbolize the 16 possible functional**
660 **combinations. Intracellular cytokine staining was performed using fresh**
661 **PBMC stimulated overnight with the indicated antigens as well as the**
662 **control antigens Staphylococcus enterotoxin B and CMV_{pp65}. The**
663 **number of different antigenic regions (linear peptide pools, x-axis)**
664 **recognized by TaMoVac vaccinees (n=42) is shown in (D) and was**
665 **determined using the IFN-γ ELISpot. The frequency of subjects with a**
666 **given Gag response breadth is indicated on the y-axis. Nine and 7**
667 **linear peptide pools matching MVA-CMDR-Gag_{p55} (gray bars) or DNA-**
668 **SMI-Gag_{p37} (black bars) subdivided Gag into distinct antigenic regions,**
669 **respectively. Statistical comparison in (B) was performed using**
670 **Wilcoxon-matched pairs-signed rank-test.**

671

672 **Figure 4. Recognition of Gag antigenic regions by vaccine induced T-**
673 **cells**

674 A comparison of the magnitude of Gag-specific T-cell responses (Y-
675 axis) targeting antigenic regions within p17, p24 and p15 normalized as
676 per 15mer peptide is shown in (**A**). The frequency of responders is
677 shown for different Gag regions (**B**). Vaccine-induced T-cell responses
678 were characterized using IFN-(ELISpot in 42 participants after
679 stimulation of fresh PBMCs with 9 and 7 peptide pools matching MVA-
680 CMDR-Gag_{p55} (grey bars) and DNA-$Gagp_{37}$ (black bars), respectively.
681 The frequency of responders for individual MVA-CMDR-Gag_{p55}
682 matched peptides is shown in (**C**) and is based on 23 subjects with at
683 least 1 detectable response against on individual peptide during fine
684 mapping using cryopreserved instead of fresh PBMC. The key data for
685 the most frequently recognized peptides are shown in (**D**). Cut off for a
686 positive response was 2-fold the unstimulated control. Corresponding
687 Gag regions p15, p17 and p24 are indicated in (**B**) and (**C**).

688

689 **Figure 5. Linear regression analyses between DNA-SMI-Gag_{p37} and**
690 **MVA-CMDR-Gag_{p55} sequence mismatches and induced T-cell**
691 **responses targeting specific antigenic regions.**

692 A linear regression analyses was performed to study the association of
693 immunogen-encoded amino acid mismatches and the respective T
694 cell response rate to the DNA-SMI-Gag_{p37} pools 1 to 6 (**A**). T cell
695 responses were detected using the IFN-γ ELISPOT and freshly isolated

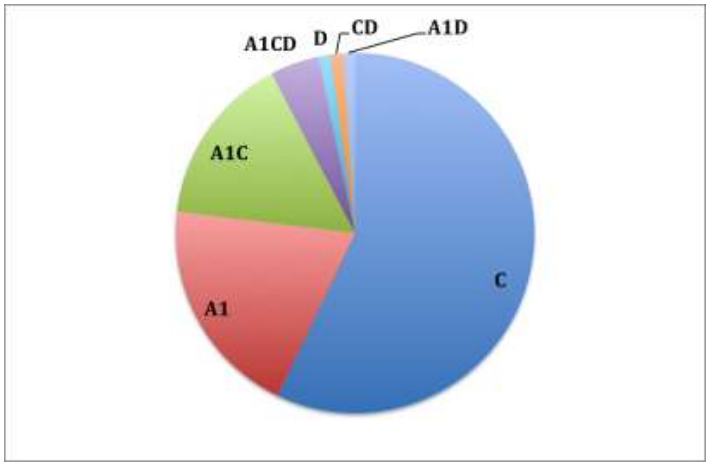
PBMC Pool 7 was excluded because it only contains 3 instead of 11 peptides. A comparison of mismatches between MVA-CMDR-Gag_{p55} to the DNA-SMI-Gag_{p37B} sequence and the respective T cell response rate (**B**) or the magnitude of responses (**C**) for single peptide-specific T cell responses as detected using the IFN-γ ELISPOT and cryopreserved PBMC.

703 **Figures**

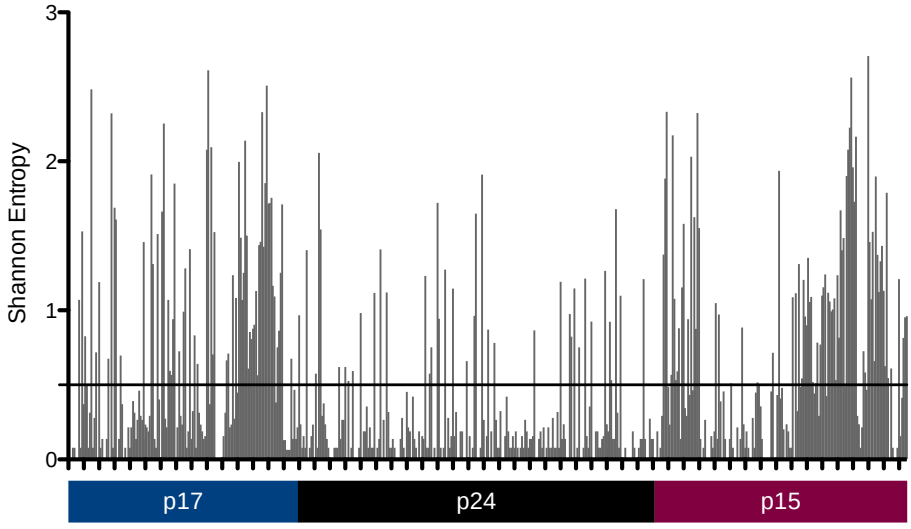
704 **Figure 1**

705

a



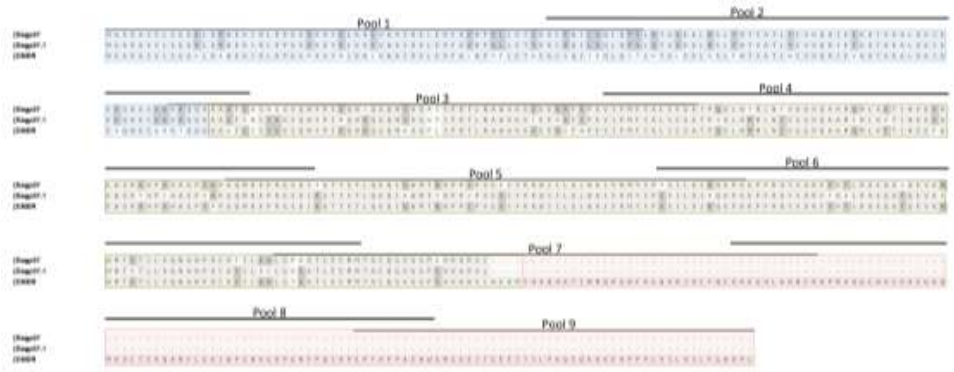
b



706

707

708 **Figure 2**

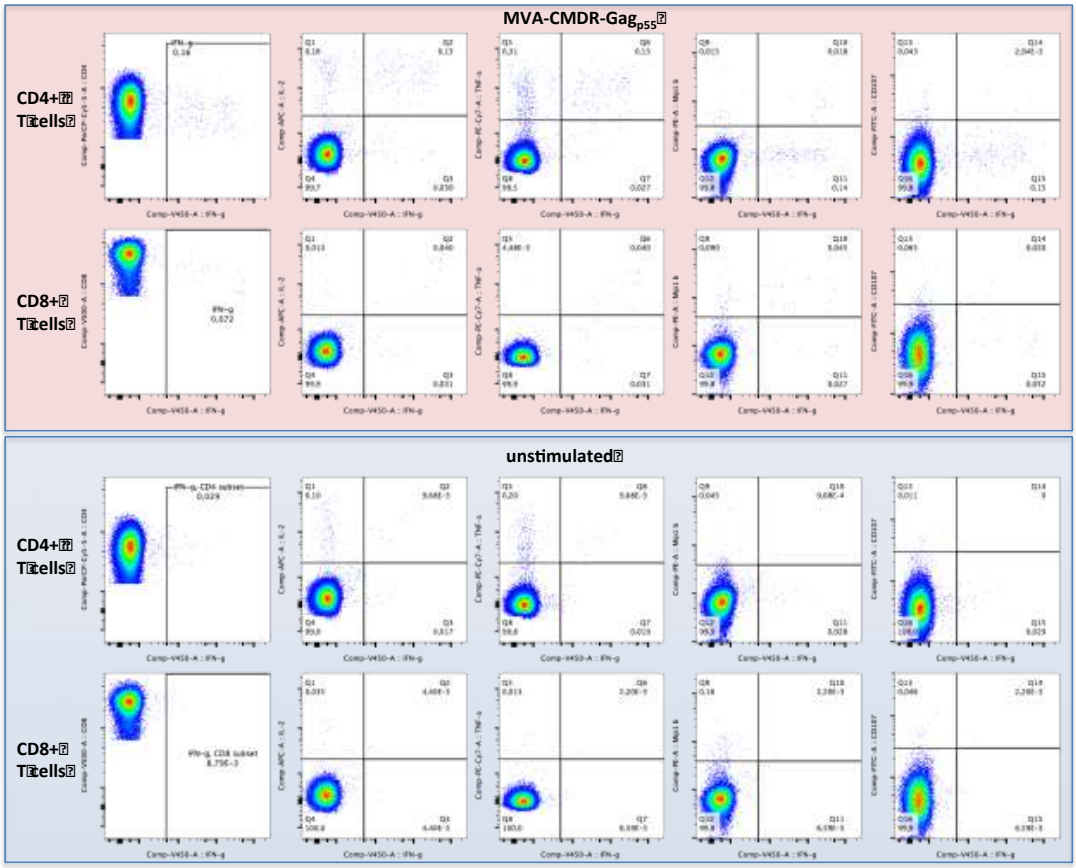


709

710

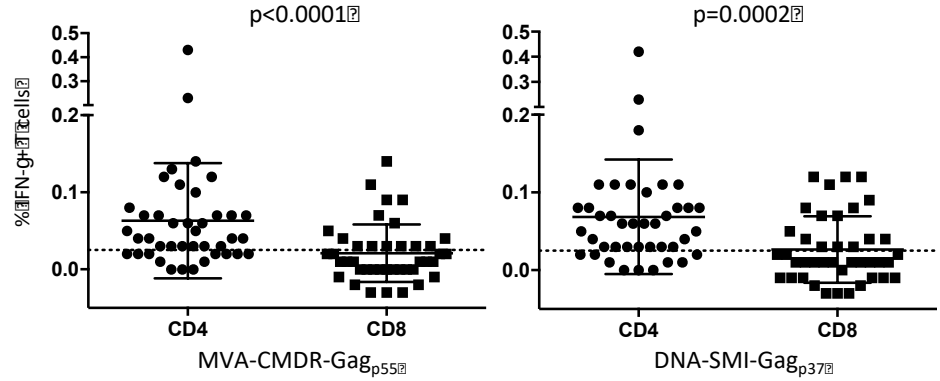
711

Figure 3

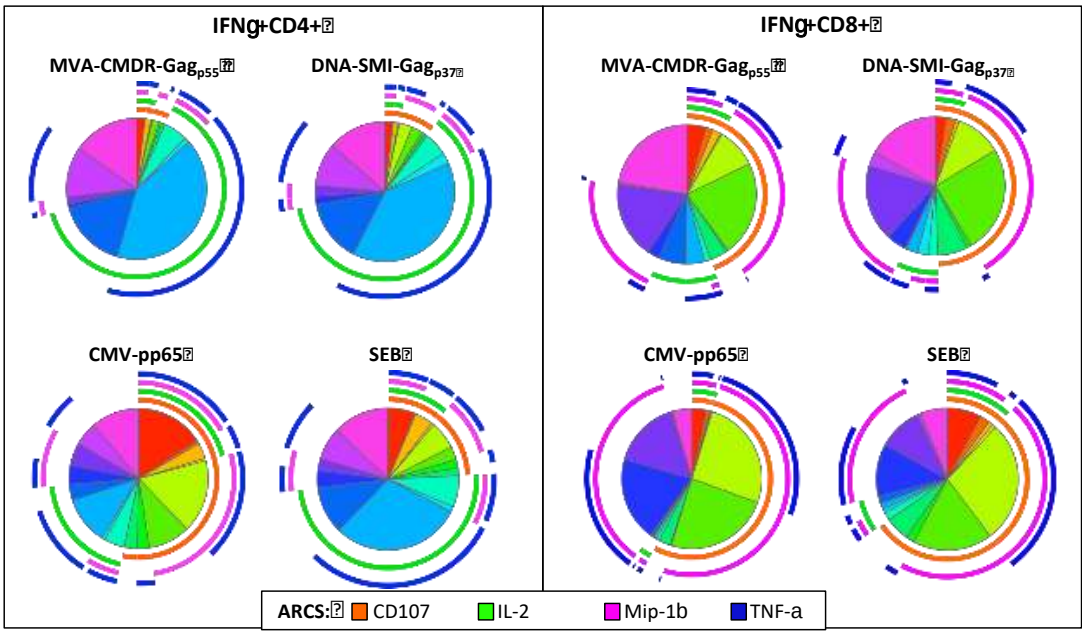


a

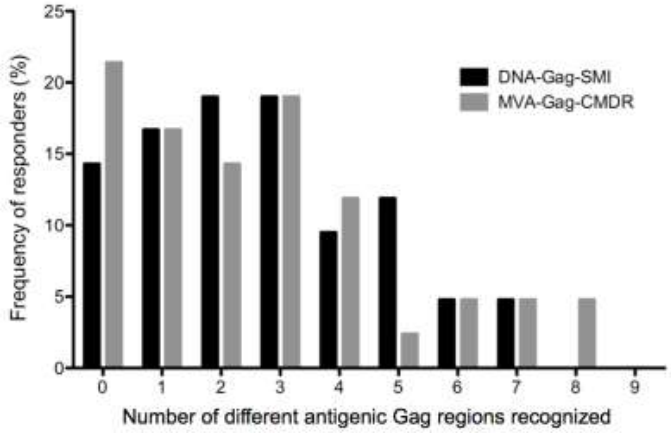
b



c



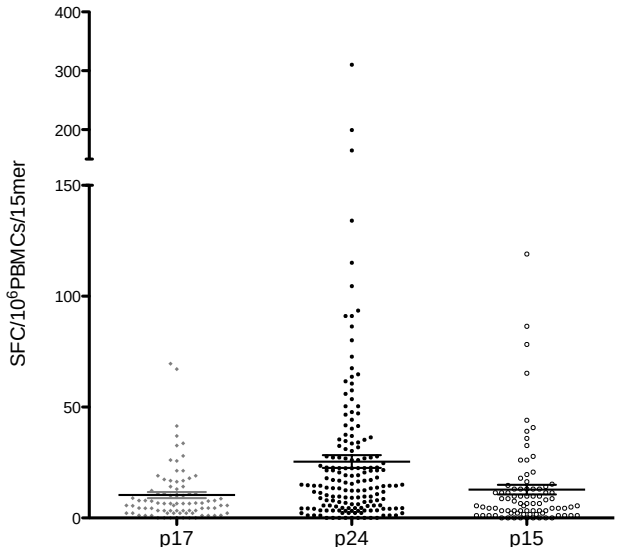
d



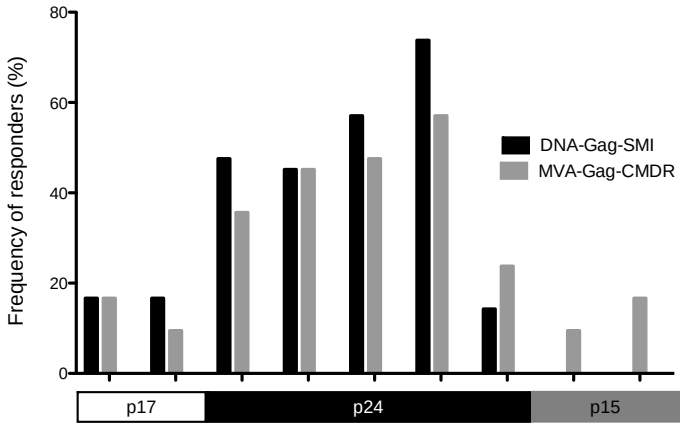
713

Figure 4

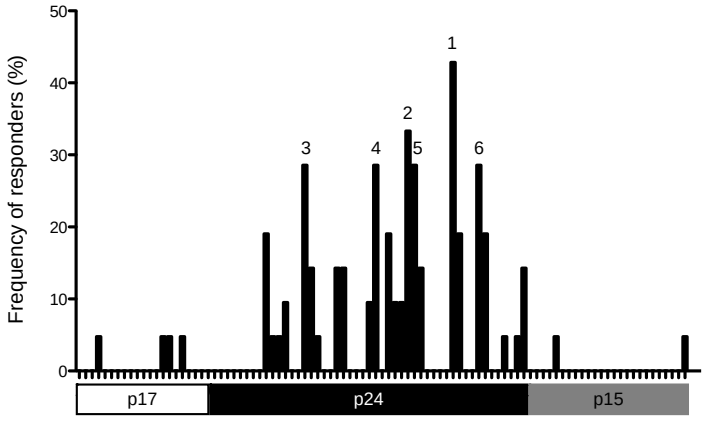
a



b



c



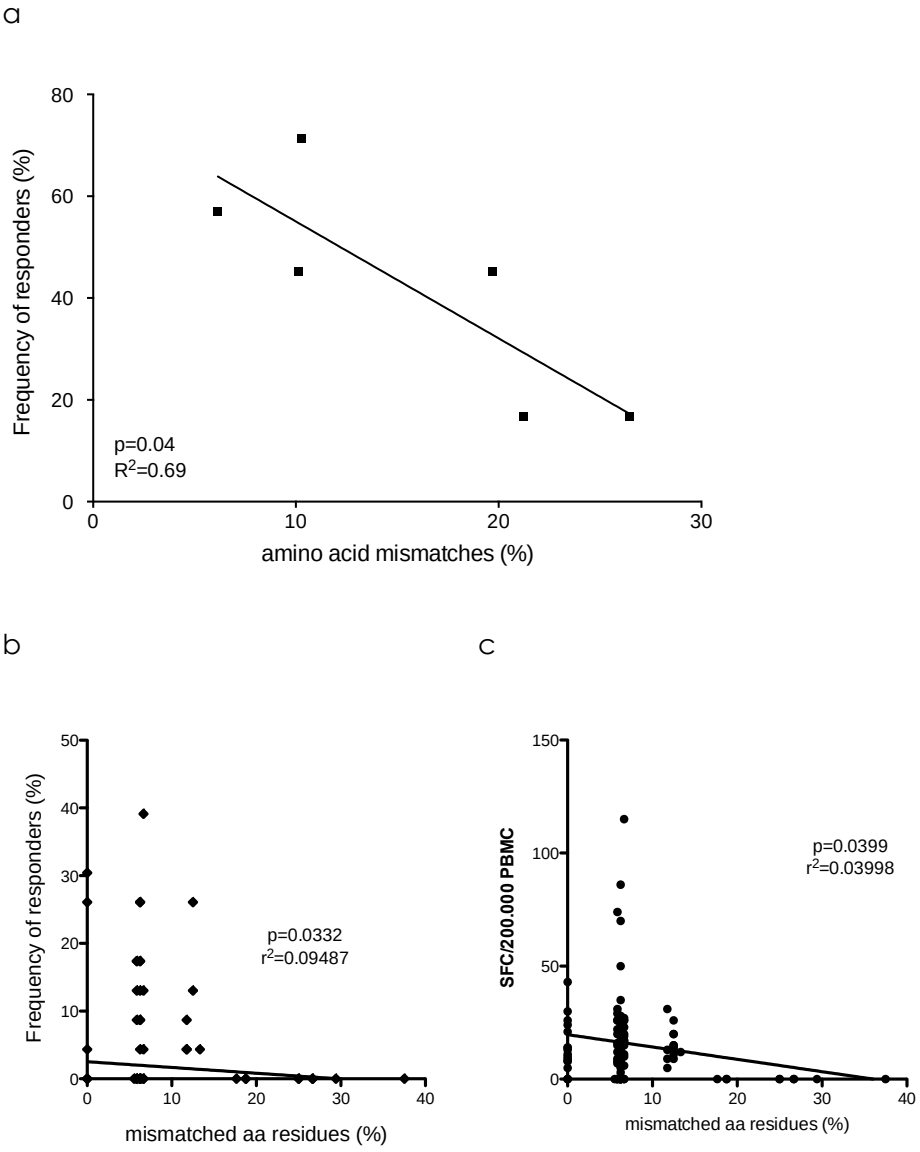
d

No.	Peptide sequence	HxB2 position	Frequency of recognition	aa mismatch MVA vs DNA B/DNA BA [%]
1	YVDRFYKTLRAEQAT	296-310	39.13%	6.67/13.33
2	YKRWIILGLNKIVRMY	262-277	30.43%	0/0
3	GATPQDLNMMLNIVGG	178-193	26.09%	12.5/6.25

4	IAGTTSTLQEQIGWMT	236-251	26.09%	6.25/6.25
5	ILGLNKIVRMYSPVSI	267-282	26.09%	6.25/0
6	WMTETLLVQNANPDCK	316-331	26.09%	0/6.25

715

716 **Figure 5**



717