Introduction
Recent findings of the RV144 vaccine trial imply that apart from vaccine regimen that do elicit neutralizing antibodies or cytotoxic T lymphocytes other mechanisms can confer protection against HIV-1 infection. As a consequence, attention has turned to activation of non-neutralizing antibodies and antibody-dependent cell-mediated cytotoxicity (ADCC) as possible mechanisms for vaccine-related protection. This also leads to an increased interest in defining new immunogenic regions presented by distinct HIV clades. Peptide microarrays are excellent tools for the identification of immunogenic protein regions and were also shown to enable monitoring of specific immune responses against HIV proteins (reported e.g. for the RV144 trial). The extreme variability of HIV protein sequences and the number of different clades make complex overlapping peptide libraries necessary to cover this diversity. The aim of this study was to design a peptide library presenting the immunodominant fractions of the HIV proteome while covering a high number of clades.

Input Sequences
- Based on alignment of HIV transcripts from the LANL database version 2009\(^1\).
- Full length of GAG (p17 and p24), TAT, ENV (gp120 and gp41), and NEF.
- Immunogenic regions of GAG p2p7p1p6, POL, VIF, and REV as published by LANL\(^2\).

Sequence Selection
- Consideration of most frequent clades A, B, C, D, G, CRF01_AE and CRF02_AG accounting for >90% of global prevalence\(^3\).
- From each of the frequent clades the best covering sequence was selected using MOSAIC\(^4\).
- A sequence cocktail containing these 7 best covering sequences was generated.
- 20 sequences with best coverage were identified without consideration of the clade and added to the sequence cocktail.
- Increase of coverage was calculated for each sequence and sequences gaining less than 0.75% in coverage were removed from the cocktail (Fig. 1).

Generation of MOSAIC Sequences
- MOSAIC sequences are pseudo-protein sequences assembled from parts of proteins of a given library to reach optimal coverage\(^4\).
- For each gene product 2 MOSAIC sequences were generated and added to the cocktail when the coverage gain was > 1%.

Final Library
- Unique overlapping peptides were generated.
- The final library consists of 5572 peptides with an average coverage of > 50%.

Application on Peptide Microarrays
- Peptide microarrays are excellent tools for monitoring humoral immune responses.
- Technology was successfully used in evaluation of patient samples in clinical trials\(^6\).
- Library (5572 peptides) fits onto one peptide microarray.

References
1. HIV Molecular Immunology: Maps of Ab Epitope Locations Plotted by Protein" in the version from March 5, 2010 from Theoretical Biology & Biophysics, Los Alamos National Laboratory was used (http://www.hiv.lanl.gov/content/immunology/maps/ab/ab.pdf).
2. "HIV Molecular Immunology: Maps of Ab Epitope Locations Plotted by Protein" in the version from March 5, 2010 from Theoretical Biology & Biophysics, Los Alamos National Laboratory was used (http://www.hiv.lanl.gov/content/immunology/maps/ab/ab.pdf).
3. "HIV Molecular Immunology: Maps of Ab Epitope Locations Plotted by Protein" in the version from March 5, 2010 from Theoretical Biology & Biophysics, Los Alamos National Laboratory was used (http://www.hiv.lanl.gov/content/immunology/maps/ab/ab.pdf).

Acknowledgements
We are indebted to Mike Schutkowski of Martin Luther University Halle/Saale and Holger Wenschuh, Petra Meyer, Johannes Zerweck and Nikolaus Pawlowski of JPT.