#1473

Discovery of serotype-specific epitope biomarkers for the diagnosis of Dengue virus infections

Renate Sekul[1], Felix F. Loeffler[2], Yuvaraj Mahendran[1], Thomas Jaenisch[3], Kirsten Heiss[1], Magelda Montoya Cruz[4], Josefina Coloma[4], Eva Harris[4] and Volker Stadler[1]

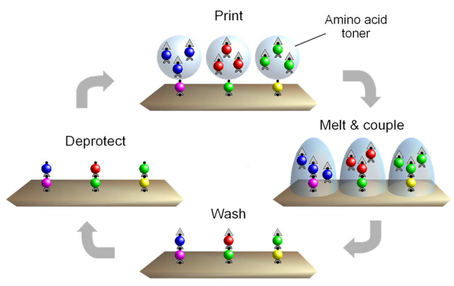
[1]PEPperPRINT GmbH, Heidelberg, Germany; [2]Max Planck Institute of Colloids and Interfaces, Potsdam, Germany; [3]Center for Infectious Diseases & Heidelberg Institute of Global Health, Germany; [4]University of California, Berkeley School of Public Health, Berkeley, USA

Abstract

Infections with dengue (DENV) circulating in tropical and subtropical regions of the world, have become a thread to public health. Severe disease is associated with infections with all DENV serotypes. However, secondary infections with a heterologous serotype are more correlated with fatal disease outcomes than primary infections. A precise and early diagnosis of the virus is hampered not only by similar clinical symptoms present in related diseases, but also by the serological cross-reactivity among DENV and other flaviviruses. Therefore, novel diagnostic tests are urgently needed to identify and discriminate co-circulating flaviviruses particularly in the acute phase, to predict severe disease outcomes and to distinguish primary and secondary infections.

Instead of using full length proteins, we intended to screen for epitope biomarkers for the differentiation of antibody responses against DENV peptides. We generated proteome-wide DENV peptide microarrays and screened them for IgG and IgM antibody responses with plasma samples of patients with acute and convalescent DENV infections of serotypes 1, 2 and 3. We observed heterogeneous antibody responses represented by distinctive and shared epitope recognition patterns and applied machine learning to predict discriminative peptides for acute and convalescent stage of infection, primary and secondary infection and serotypes, respectively. Here we present our data of machine learning for the classification of serotypes in the convalescent phase.

Technology



Benefits

- very low material consumption
- high spot density / high content (1,200 peptides/cm²)
- digital printing flexibility, fast production times
- high peptide quality with routine double couplings

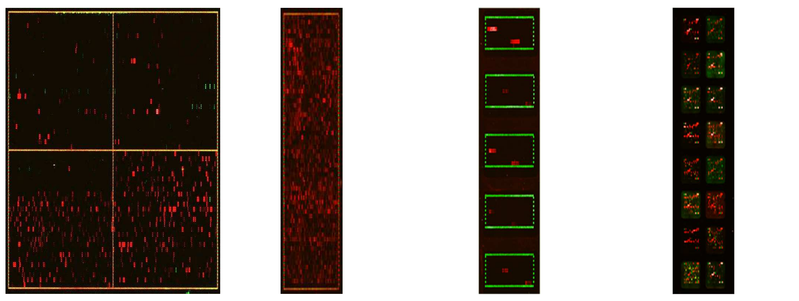
High-density PEPperCHIP® peptide microarrays are generated by digital laser printing on standard glass slides using a proprietary laser printer comprising 24 cartridges filled with individual amino acid toners. For array production, the amino acid toners are simultaneously printed with high precision on their respective positions on the glass slides.

Peptide synthesis is initiated by melting the toner particles at 90°C. Under these conditions, the amino acids are released and are available for coupling to the previous amino acid. The coupling cycle is completed by washing steps to remove excess building blocks and protecting groups. Finally, the array is ready for the next synthesis cycle with laser printing and coupling.

The benefits of this technique are an unique flexibility in terms of peptide content, a high spot density with up to 11,000 features per chip and low material consumption enabling the generation of customized peptide array at reasonable costs.

PEPperCHIP® Peptide Microarray Platform

Discovery Format	Standard Format	Mapping Format	Multiplexed Format

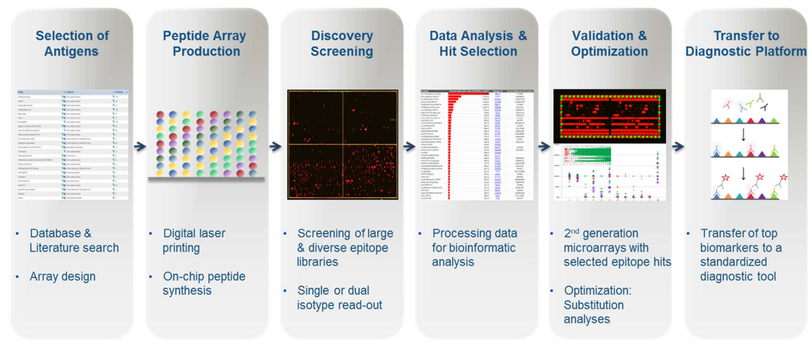


> 70,000 Peptides	~11,000 Peptides	e.g. 5 x 1,200 Peptides	16 x 200 Peptides

Antibody-peptide interactions are analyzed by immuno-type assays in a high-throughput fashion on peptide microarrays. Depending on the application, various microarray formats are available:

- **Discovery format**: > 70,000 individual peptides; suitable for screening of large, diverse epitope libraries covering up to 100 proteins; applied for biomarker and target binder discovery
- **Standard format**: ~ 11,000 individual peptides; routinely used for epitope mapping. Custom peptide microarrays or standard chips e.g. PEPperCHIP® Infectious Disease Epitope Microarray
- **Mapping format:** several identical array copies on a single chip; ideal for parallel screening of multiple samples; used for epitope mapping of single protein antigens, detailed epitope characterization or biomarker validation
- **Multiplexed format:** up to 16 array copies on a single chip; ideal for assay development or hit validation studies with sample cohorts

General Workflow Biomarker Discovery



Biomarker Discovery Approach:

- The process starts with the design of the antigen library. Antigen proteins are translated into sets of overlapping peptides and synthesized on linear or conformational discovery microarrays.
- The discovery screening is performed as immunoassay with single or dual isotype read-out.
- Most relevant biomarker candidates are selected for further validation and optimization on 2nd generation microarrays.
- Finally, the most comprehensive biomarkers are validated by standard diagnostic formats such as ELISA.

Distinct and shared antibody responses against DENV-derived epitopes

Primary Screening

Library Content

- 405 DENV 1-3 proteomes Nicaragua, 9 DENV 4 proteomes Columbia
- *In silico* analysis resulted in 5,522 individual peptides for screening
- DENV proteome microarray: **5,522 linear DENV peptides**

Study Outline

- 468 sera from Nicaragua including control samples
- **Patient stratification** by: DENV serotype (1-3), phase of infection (acute vs convalescent), immune response (primary vs secondary infection) and severity of disease
- Dual isotype screening: IgM and IgG

1,297 peptides identified

Secondary Screening

Library Content

- 1,297 hit peptides identified by the primary screening

Study Outline

- 457 sera from Nicaragua including control samples
- **Patient stratification** by: DENV serotype (1-3), phase of infection (acute vs convalescent), immune response (primary vs secondary infection) and severity of disease
- Dual isotype screening: IgM and IgG

350 validated hit peptides

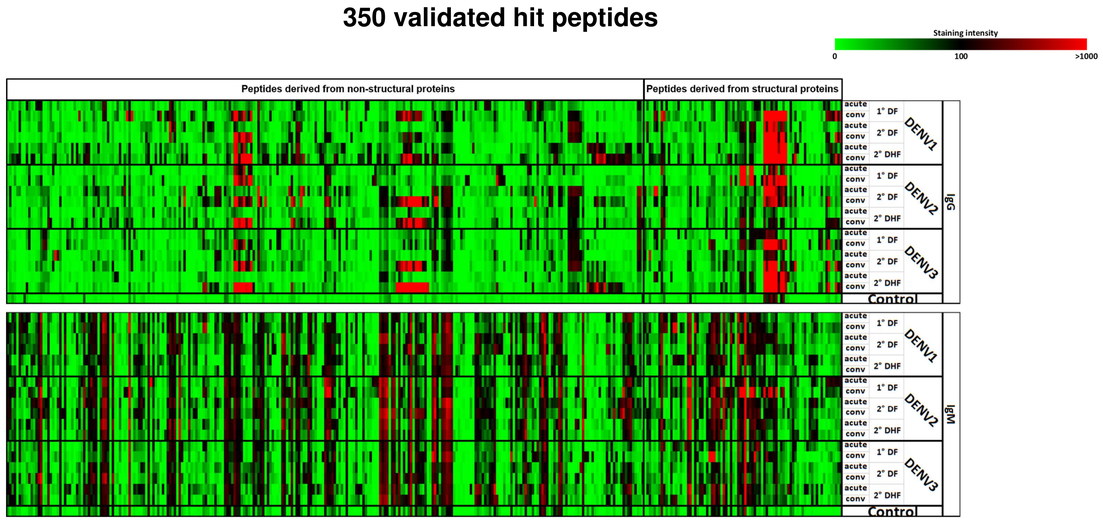


Figure 1. **IgG and IgM response against the selected 350 peptides**. The heat-map shows the mean fluorescence intensity for each individual peptide detected for the respective screening cohort. Peptides were selected based on the following criteria: (1) signal intensity above the baseline which is defined as average intensity of each peptide of the control group plus threefold SD (2) any patient group showing more than 25% of IgG values (for IgM 35%) above the baseline.

Applying machine learning to identify serotype-specific epitopes

Training

Outline Machine learning

- Rationale: identify classifiers of DENV serotypes
- **Training set**: building the model for antibody responses (IgG and IgM) against 350 validated hit peptides using **labeled** DENV samples from Nicaragua (convalescent phase)
- Applied models: Cross validation GLMNET –LASSO model

Test

Outline Machine learning

- Rationale: apply the best trained model from training data to predict the unlabeled serotype in test data
- **Test set 1**: antibody responses against 350 validated hit peptides using **unlabeled** DENV samples from Nicaragua
- **Test set 2**: antibody responses against 350 validated hit peptides using **unlabeled** DENV samples from travelers (returning from Asia)

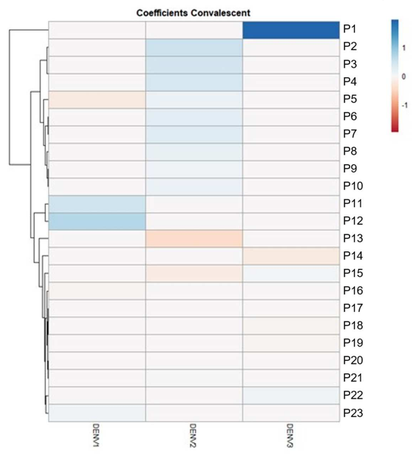


Figure 2: The fitted peptide coefficients are informative for discriminating between DENV1, DENV2 and DENV3 of the GLMNET model from training data.

Predicted	DENV1	DENV2	DENV3	Unclear
DENV1	17	17	17	0
DENV2	0	90	10	0
DENV3	0	33	67	0
Unclear	0	100	0	0

Figure 3a) Confusion matrix to evaluate the prediction accuracy in unlabeled Nicaragua convalescent samples. Overall, the model predicted 56% correctly. An accuracy of 90% for DENV2 and 67% for DENV3 was achieved.

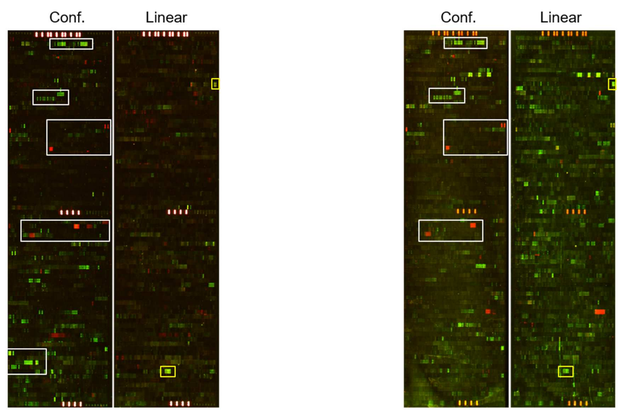
Prediction	DENV1	DENV2	DENV3
DENV1	0	0	0
DENV2	83	67	50
DENV3	17	33	50

Figure 3b) Confusion matrix to evaluate the prediction accuracy in the unlabeled travelers convalescent samples. Overall, the model predicted only 27% correctly. However, an accuracy of 67% for DENV2 and 50% for DENV3 was achieved.

DENV-specific antibody response against linear vs. conformational peptides highlighted new conformational epitopes

Library Content

- DENV proteome microarray with 5,522 linear or conformational DENV peptides



DENV Serum 473 (DENV1, Malaysia)

DENV Serum 481 (DENV2, Thailand)

Figure 4. **Screening of linear and conformational DENV proteome microarrays discovered novel conformational epitopes**.
Representative response pattern of DENV sera derived from travelers returning from Asia are shown. Microarrays were incubated with respective sera overnight at 4°C. Secondary detection was done using anti-human IgG DyLight680 and anti-human IgM DyLight800. Read-out was performed using a LI-COR Odyssey Imaging System. Polio peptide spots served as control.

Summary & Outlook

- The differentiation of DENV antibody responses on epitope instead of on protein level yielded a more comprehensive picture of DENV infections and peptide biomarkers.
- Our trained model can now be applied to predict for Dengue infection types and severity in additional test data sets.
- Combining high density peptide arrays and machine learning is a powerful tool to predict biomarkers for infectious diseases.

Acknowledgements

This work was supported by the U.S. Department of Homeland Security, Science & Technology Directorate (HSHQDC-15-C-B0010).

www.pepperprint.com

Please visit our booth # 518