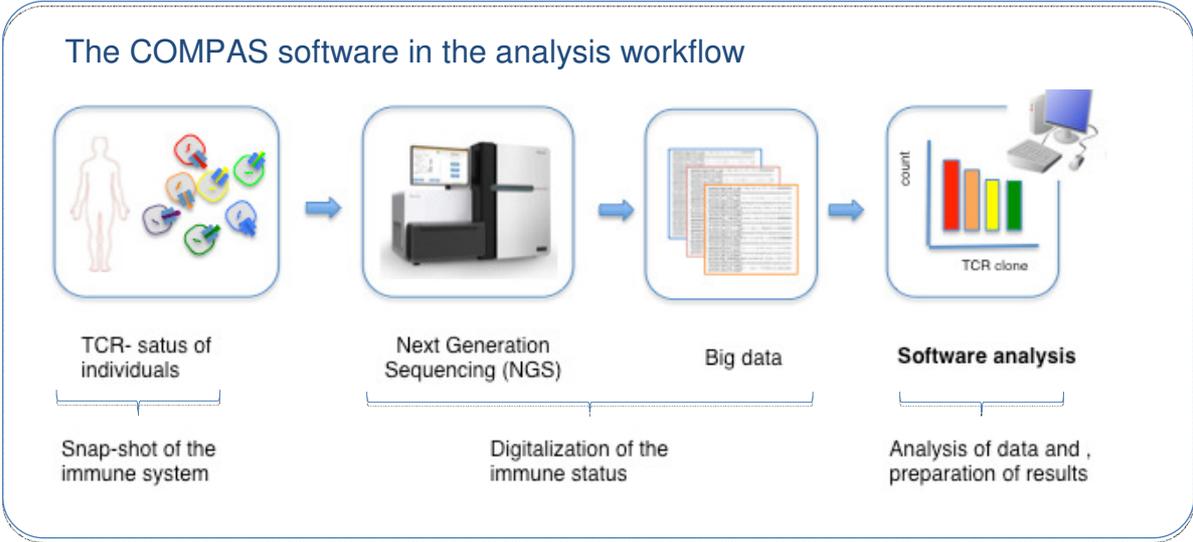apta IT

NGS Analysis of T-Cell Receptors
by COMPAS

The software tool COMPAS (COMmon PAtternS) harnesses next generation sequencing (NGS) data to monitor immune response on the level of T-cell receptors (TCRs).
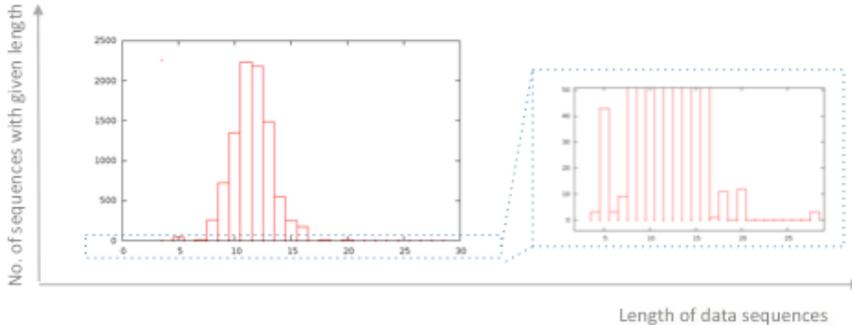
The COMPAS software in the analysis workflow

The amplicons from different time points after a clinical event (*i.e.* immunization) can be indexed and combined to enable economic NGS. COMPAS is subsequently used to sort data sets and to perform individual as well as comparative analysis of data sets.

Analysis pipeline
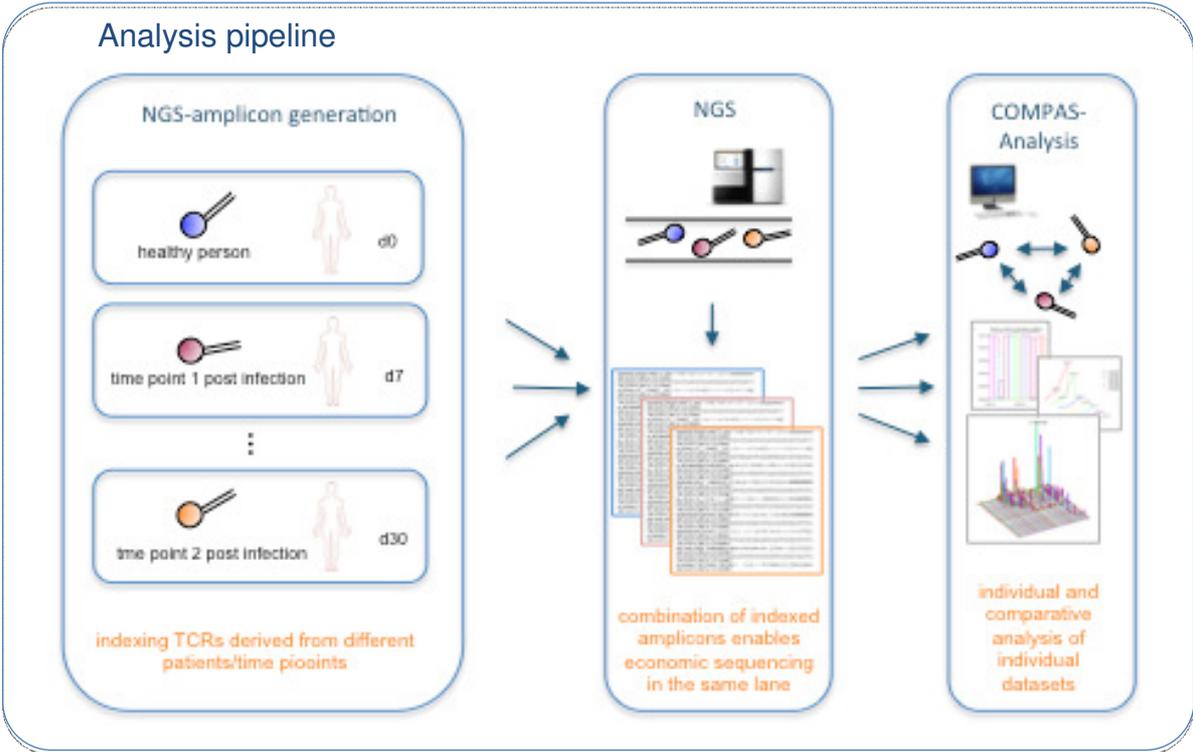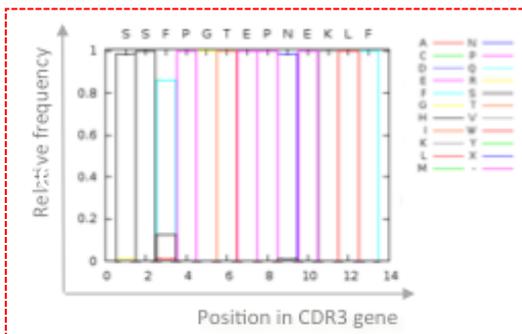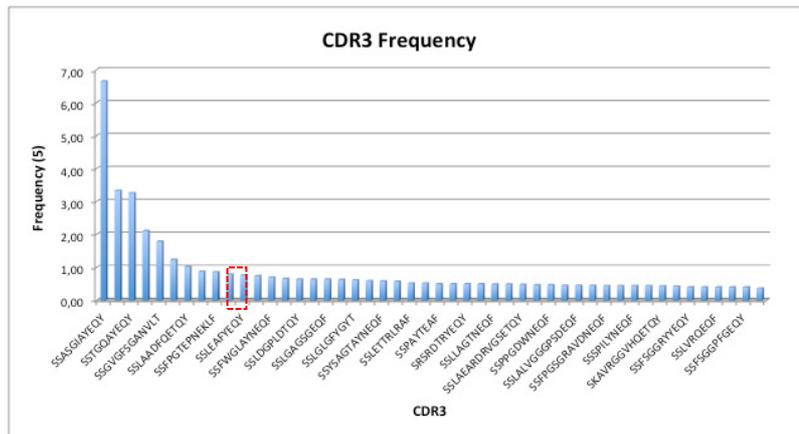
COMPAS - Results

Spectratyping

The distribution of lengths of the CDR3-genes can be analyzed at high resolution.
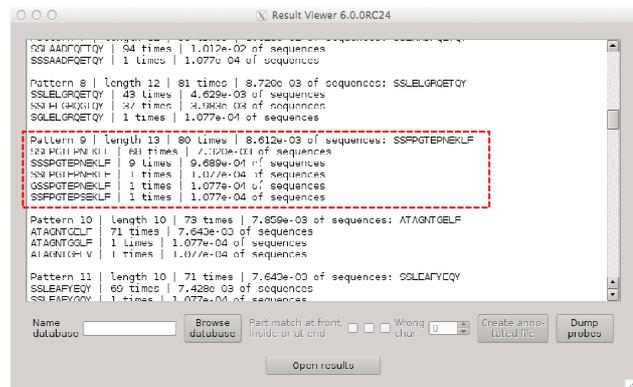
Frequency distribution of CDR3 genes

The frequency distribution of CDR3 is given in absolute and relative numbers.

Similar sequences can be clustered into CDR3 patterns.

A bar chart visualizes the distribution of aa over the position of the CDR3-gene (top).

CDR3-sequences for each CDR3-pattern are given including absolute and relative numbers (right).
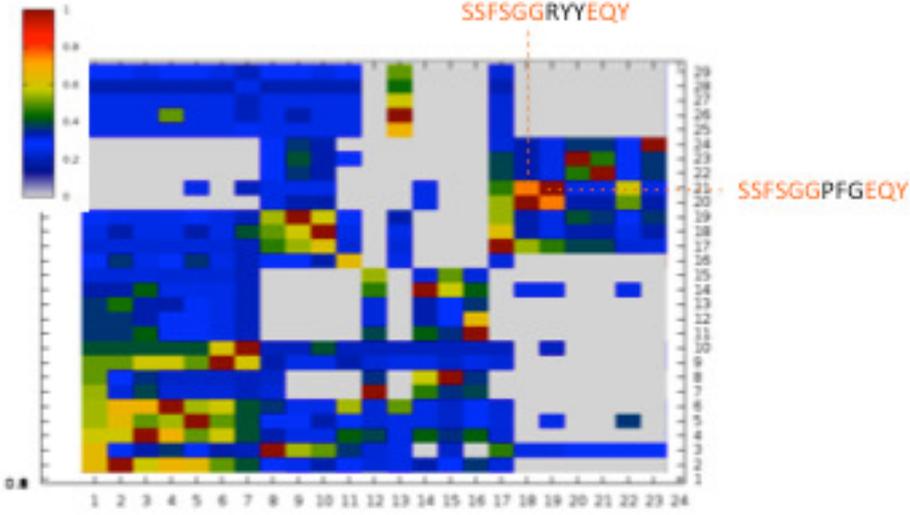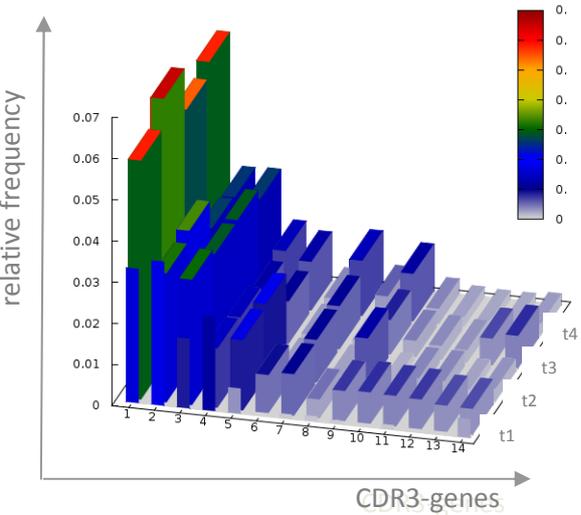
CDR3 relationships

The heat map indicates relationships of identified CDR3 genes. Orange fields represent CDR3 genes with 75% homology.

A pair of highly related CDR3 genes are given as an example on the right.

high similarity	
low similarity	

Tracing CDR3 genes over different time-points
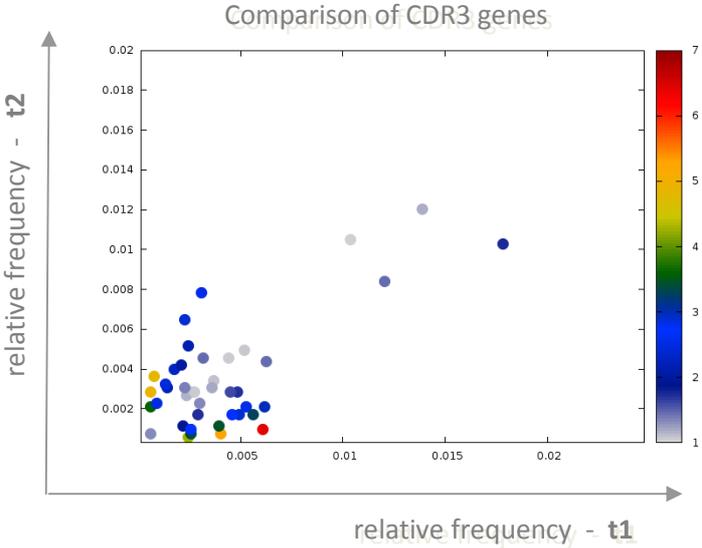

1: SSASGIAYEQY
2: STSLGSASTDTQY
3: SSTGQAYEQY
4: SSSHRVPNTEAF
5: SSGVGFSGANVLT
6: SSLVRQAYGYT
7: SSLAADFQETQY
8: SSLELGRQETQY
9: SSFPGTEPNEKLF
10: ATAGNTGELF
11: SSLEAFYEQY
12: SSLVGQGLHEQF
13: SSFWGLAYNEQF
14: SSLTANTEAF

The relative frequency of CDR3-sequences (figure on the left shows the 14 most frequent ones) are traced over different time points (t1–t4). This feature enables the identification of CDR3s as a response to a clinical event.

The development of CDR3 sequences can also be assessed by the criteria of enrichment factors. In the figure on the right, frequent CDR3s are compared between two time points (here t1 and t2). The position indicates the relative frequency. The color corresponds to the factor by which the respective CDR3-genes were enriched or decreased.

Comparison of CDR3 genes

high amplification factor	
low amplification factor	

V-gene and J-gene usage of CDR3 regions

COMAPS gives an overview of used V-genes as well as J-genes for identified CDR3-genes. The heat maps below summarize the used J- genes of the 14 most frequent CDR3 patterns.
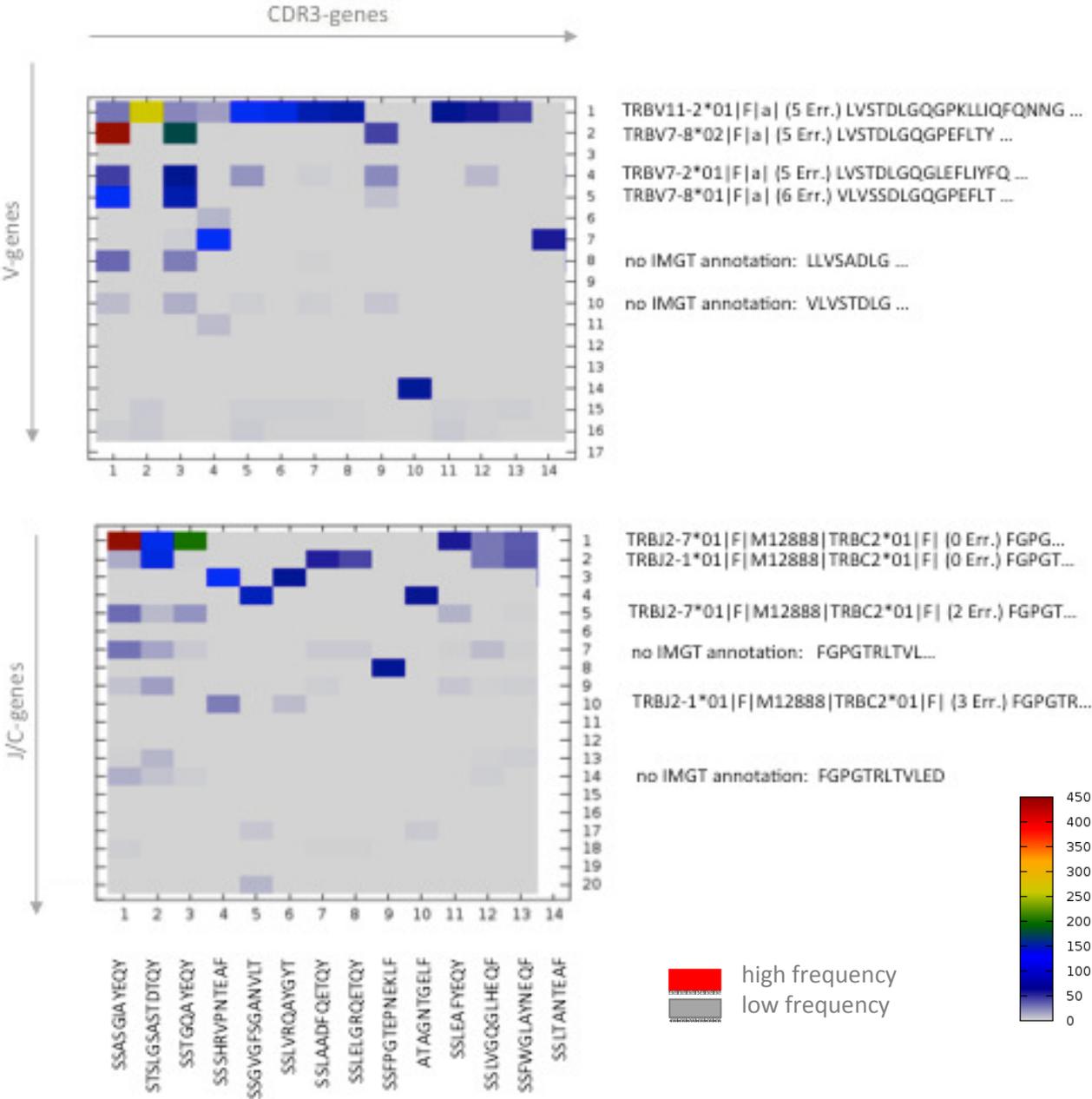
For the dominant CDR3 (1: SSASGIAYEQY) the most frequent V-gene is 2: TRBV7-8*02 and the most frequent J-gene is 1: TRBJ2-7*01. The Heat map makes clear that this CDR3-gene uses further V-and J-genes to some percentage. Annotation is performed according to IMGT-standard.