

# geno2pheno[coreceptor-hiv2]

## a new diagnostic tool for the genotypic determination of HIV-2 coreceptor usage

M. Döring<sup>1</sup>, P. Borrego<sup>2</sup>, J. Büch<sup>1</sup>, A. Martins<sup>2</sup>, G. Friedrich<sup>1</sup>, R. J. Camacho<sup>3</sup>, J. Eberle<sup>4</sup>, R. Kaiser<sup>5</sup>, T. Lengauer<sup>1</sup>, N. Taveira<sup>2,6</sup>, N. Pfeifer<sup>1</sup>

<sup>1</sup> Department for Computational Biology & Applied Algorithmics, Max Planck Institute for Informatics, Saarbrücken, Germany.

<sup>2</sup> Research Institute for Medicines (imed.ULisboa), Faculty of Pharmacy, University of Lisbon, Lisbon, Portugal.

<sup>3</sup> Department of Microbiology & Immunology, Rega Institute for Medical Research, KU Leuven, Belgium.

<sup>4</sup> Department of Virology, Max von Pettenkofer-Institut, Ludwig-Maximilians-University, Munich, Germany.

<sup>5</sup> Institute for Virology, University of Cologne, Cologne, Germany.

<sup>6</sup> Instituto Superior de Ciências da Saúde Egas Moniz (ISCSEM), Monte de Caparica, Portugal.

Contact: mdoering@mpi-inf.mpg.de

### Relevance of HIV-2 coreceptor usage

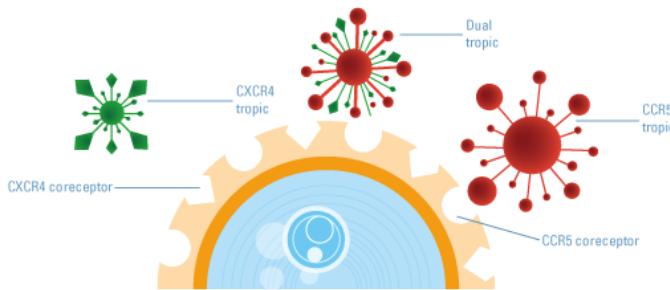


Figure 1: HIV coreceptors ([www.viivhcdxresource.com](http://www.viivhcdxresource.com))

- The selection of HIV-2 variants using the CXCR4 coreceptor (X4-capable) should be prevented because **X4-capable variants are harder to neutralize** than viruses using only CCR5 (R5)[1].
- Before prescribing CCR5-coreceptor antagonists to patients infected with HIV-2, **clinicians should rule out the existence of X4-capable variants**.
- Goal: differentiate R5 and X4-capable HIV-2 variants** based on the amino acid sequence of the V3 loop.

### Materials and methods

Support vector machines (SVMs) were trained on a data set of 73 R5 and 52 X4-capable samples to classify binary-encoded V3 amino acid sequences as either *R5* or *X4-capable*. Classifier performance was evaluated using 10-fold nested cross validation (CV). The predicted probabilities indicating whether a sequence originates from an X4-capable variant were transformed to false positive rates (FPRs).

We developed a visual representation of position-specific classifier weights to indicate amino acids associated with R5 and X4-capable variants (see Fig. 2). We evaluated established discriminatory sequence features from a rules-based approach by Visseaux et al. [2] and novel features detected by the SVM using Fisher's exact test with multiple testing correction (Benjamini and Hochberg).

### References

- [1] J. M. Marcelino et al. Resistance to antibody neutralization in HIV-2 infection occurs in late stage disease and is associated with X4 tropism. *AIDS*, 26(18):2275–2284, 2012.
- [2] B. Visseaux et al. Molecular determinants of HIV-2 R5-X4 tropism in the V3 loop: development of a new genotypic tool. *Journal of Infectious Diseases*, 205(1):111–120, 2012.

### Results

- A linear SVM (AUC=0.95) outperformed other models and was used in all subsequent analyses.
- For a set of 126 V3 sequences, the 10-fold nested CV sensitivity was 76.9% and the specificity was 97.3%.
- All samples from a set of nine, newly phenotyped V3 sequences were classified correctly by the SVM.
- We validated existing markers for X4-capability [2] and identified new, significant features ( $p \leq 0.05$ ): variants 27K, 15G, and 8S.

### Visualization of model weights

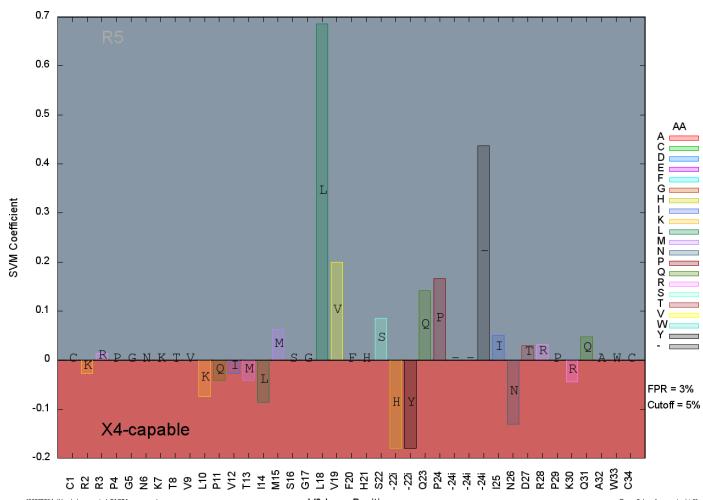


Figure 2: SVM weights for the V3 loop of a ROD10 isolate.

### Highlights of the tool

- Accuracy:** high sensitivity and specificity
- Interpretability:** visualization of sequence-specific weights and output of FPRs
- Availability:** an online web service is available at [coreceptor-hiv2.genopheno.org](http://coreceptor-hiv2.genopheno.org)
- Opportunities:** enables large-scale epidemiological studies on HIV-2 coreceptor usage