Heinrich Heine Universität Düsseldorf

Computational Pharmaceutical Chemistry Heinrich Heine University Düsseldorf

JÜLICH Forschungszentrum | JÜLICH SUPERCOMPUTING CENTRE

NIC (Project-ID: fic1)

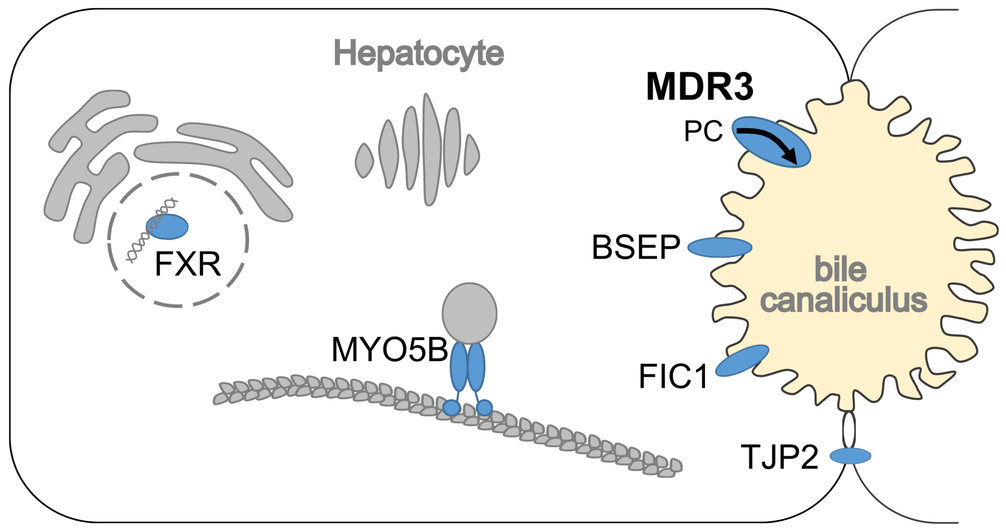
GEFÖRDERT VOM Bundesministerium für Bildung und Forschung

Vasor: accurately predicting single amino acid substitution impact within the MDR3 protein

Annika Behrendt[1], Pegah Golchin[2], Filip Koenig[1], Daniel Mulnaes[1], Amelie Stalke[3,4], Carola Dröge[5,6], Verena Keitel[5,6], Holger Gohlke[1,7]

1 Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf; 2 Department of Electrical Engineering and Information Technology, Technische Universität Darmstadt; 3 Department of Human Genetics, Hannover Medical School; 4 Department of Pediatric Gastroenterology and Hepatology, Division of Kidney, Liver and Metabolic Diseases, Hannover Medical School; 5 Department for Gastroenterology, Hepatology and Infectious Diseases, Medical Faculty, Otto von Guericke University Magdeburg; 6 Department for Gastroenterology, Hepatology and Infectious Diseases, University Hospital, Medical Faculty, Heinrich Heine University Düsseldorf; 7 John-von-Neumann-Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), Institute of Biological Information Processing (IBI-7: Structural Biochemistry), and Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich

Multidrug Resistance protein 3 (MDR3) is an essential hepatobiliary transport protein. MDR3 flops phosphatidylcholine (PC) from the inner to the outer membrane leaflet, where PC can get extracted into primary bile. Dysfunction of MDR3 thus can disturb the carefully regulated bile homeostasis. Phenotypically, this imbalance can result in transient manifestations such as intrahepatic cholestasis of pregnancy to severe forms of progressive familial intrahepatic cholestasis 3 (PFIC3), leading to liver failure in affected children. Genetically, the majority of disease-causing MDR3 variants are missense variations within the MDR3-coding *ABCB4* gene.
However, assessing the impact of variants identified in sequenced patients continues to be a challenge for clinicians and researchers.



Key protein players of bile homeostasis can be affected in PFIC diseases.

Aim: Predict effect of missense variants in MDR3 identified in patients into the categories benign and pathogenic

Our approach: Protein-specific prediction tool using XGBoost

Create MDR3-specific dataset

Literature (Manual filter) → Pathogenic variants

Literature (Manual filter) + gnomAD (VarSome) → Benign variants

Obtain features

General predictors:
- EVE
- PolyPhen-2
- I-Mutant2.0
- MUpro
- MAESTRO
- PON-P2
- EVmutation

Additional features:
- PTM site effect
- Secondary structure effect
- Residue solvent accessibility
- Half sphere exposure

18 features extracted

Machine Learning

Train model on selected features

↓

Evaluation on independent test set

↓

Comparison of performance to established predictors

Our dataset: 364 variants (279 benign, 85 pathogenic)

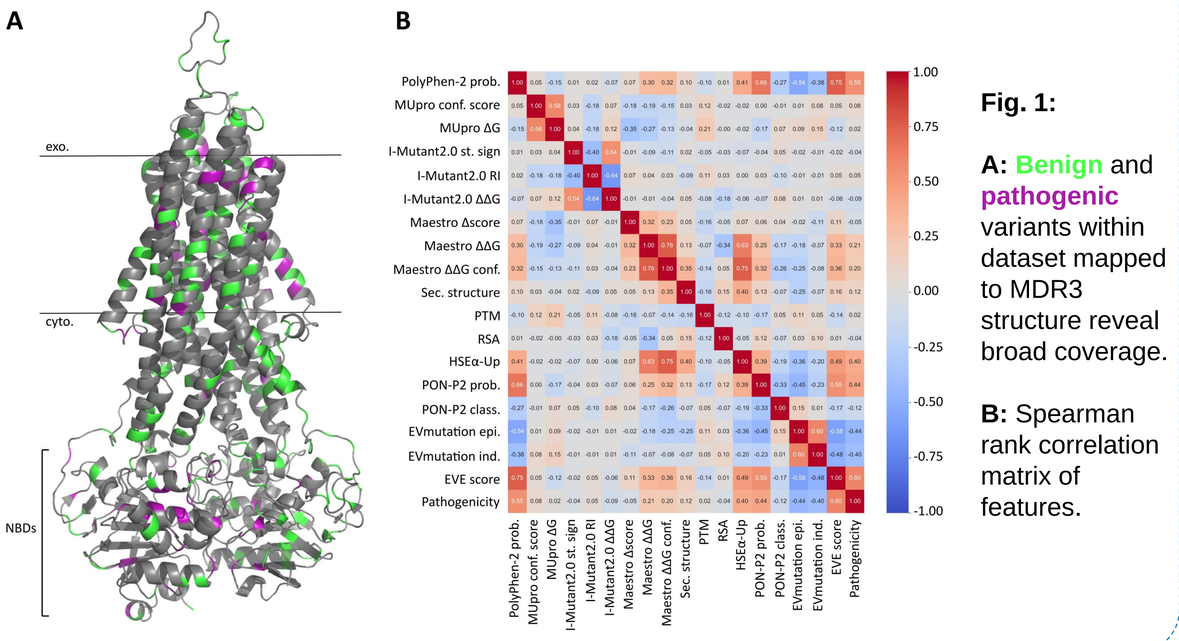


Fig. 1:

A: **Benign** and **pathogenic** variants within dataset mapped to MDR3 structure reveal broad coverage.

B: Spearman rank correlation matrix of features.

Variant assessor of MDR3 (Vasor) performance compared to established predictors

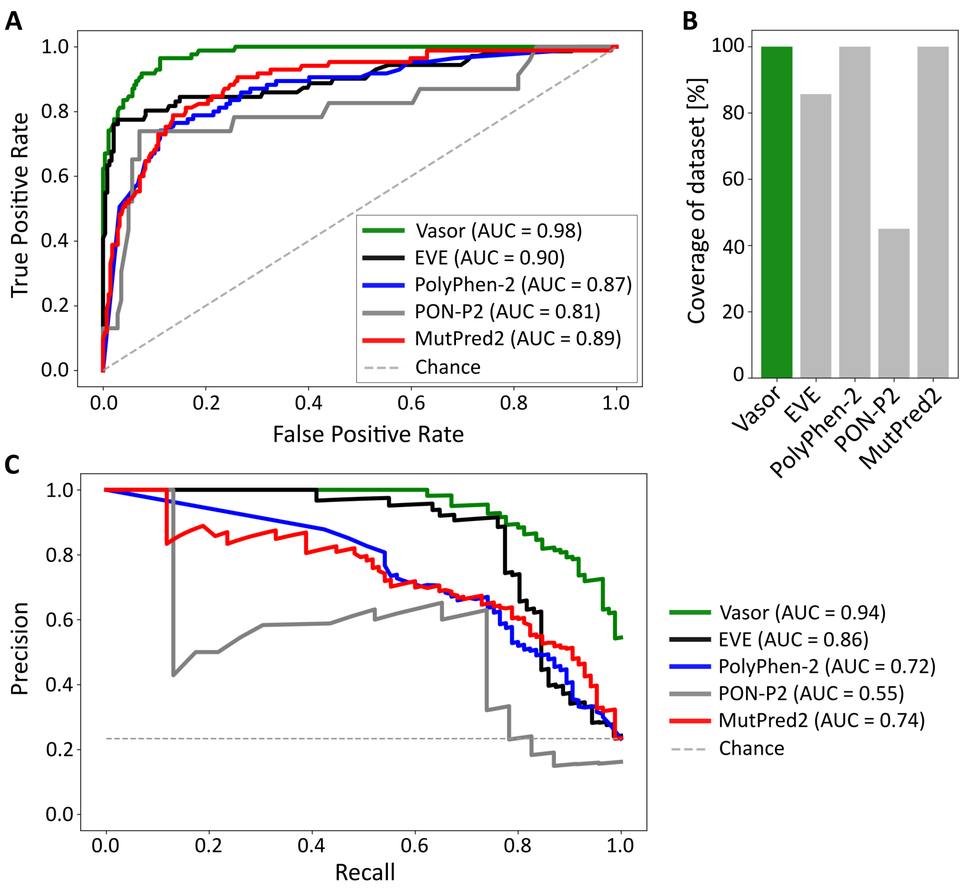


Fig. 2: A: ROC, **B:** coverage, and **C:** precision-recall curves of predictors on dataset.

Vasor's output: Probability of pathogenicity of variants

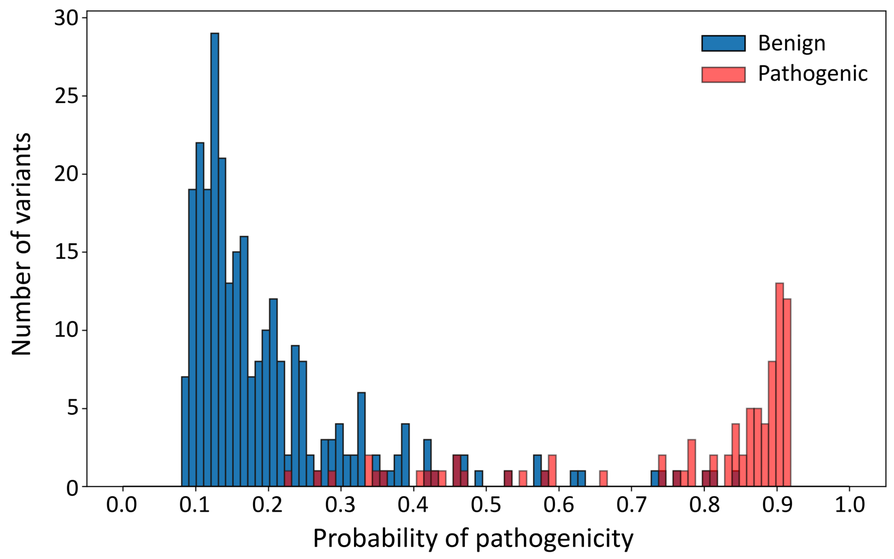


Fig. 3: Distribution of probability of pathogenicity output values of Vasor.

Vasor outperforms established predictors.

Vasor shows robust separation of probability of pathogenicity values for the classes.

Vasor allows calculation of every possible amino acid substitution.

Calculation of every possible substitution

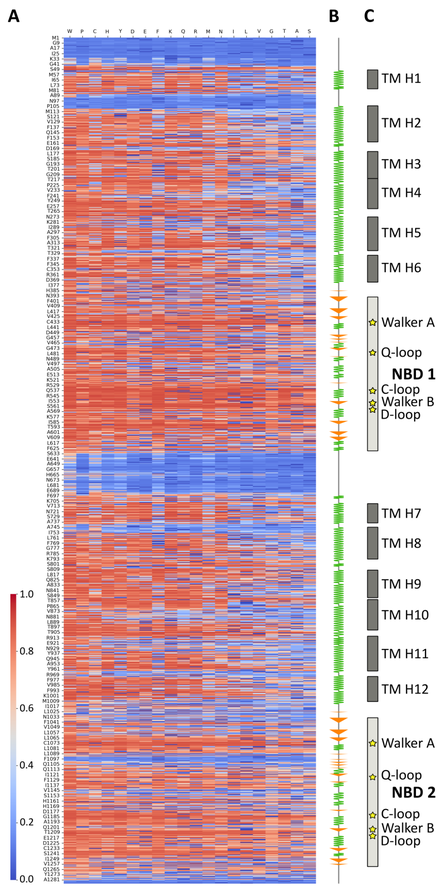
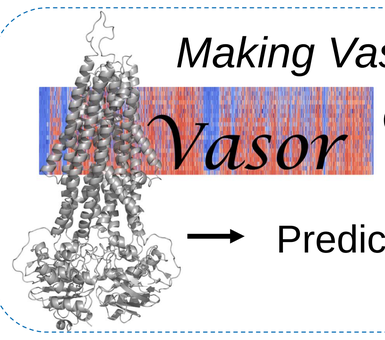


Fig. 4: A: Heatmap, **B:** sec. structure, **C:** motifs and domains of MDR3

Making Vasor easily accessible: Implementing a webserver



Clinicians and researchers can enter their variant of interest at https://cpclab.uni-duesseldorf.de/mdr3_predictor/

SCAN ME

→ Prediction output with probability of pathogenicity value and visualization of variant site within MDR3 protein structure

References & Funding

Behrendt A, Golchin P, König F, Mulnaes D, Stalke A, Dröge C, Keitel V, Gohlke H. Vasor: Accurate prediction of variant effects for amino acid substitutions in MDR3. *Hepatol. Commun.* **2022**, DOI: 10.1002/hep4.2088.

HiChol

This study was supported by the BMBF through HiChol (01GM1904A, 01GM1904B) and, in part, InCelluloProtStruct (031L0182).
We gratefully acknowledge the computing time granted by the John von Neumann Institute for Computing (NIC).

Annika.Behrendt@hhu.de | CPClab | https://cpclab.uni-duesseldorf.de/mdr3_predictor/

hhu.de