



XXX Symposium on Bioinformatics
and Computer-Aided Drug Discovery

Quantitative Prediction of Human Immunodeficiency Virus Drug Resistance

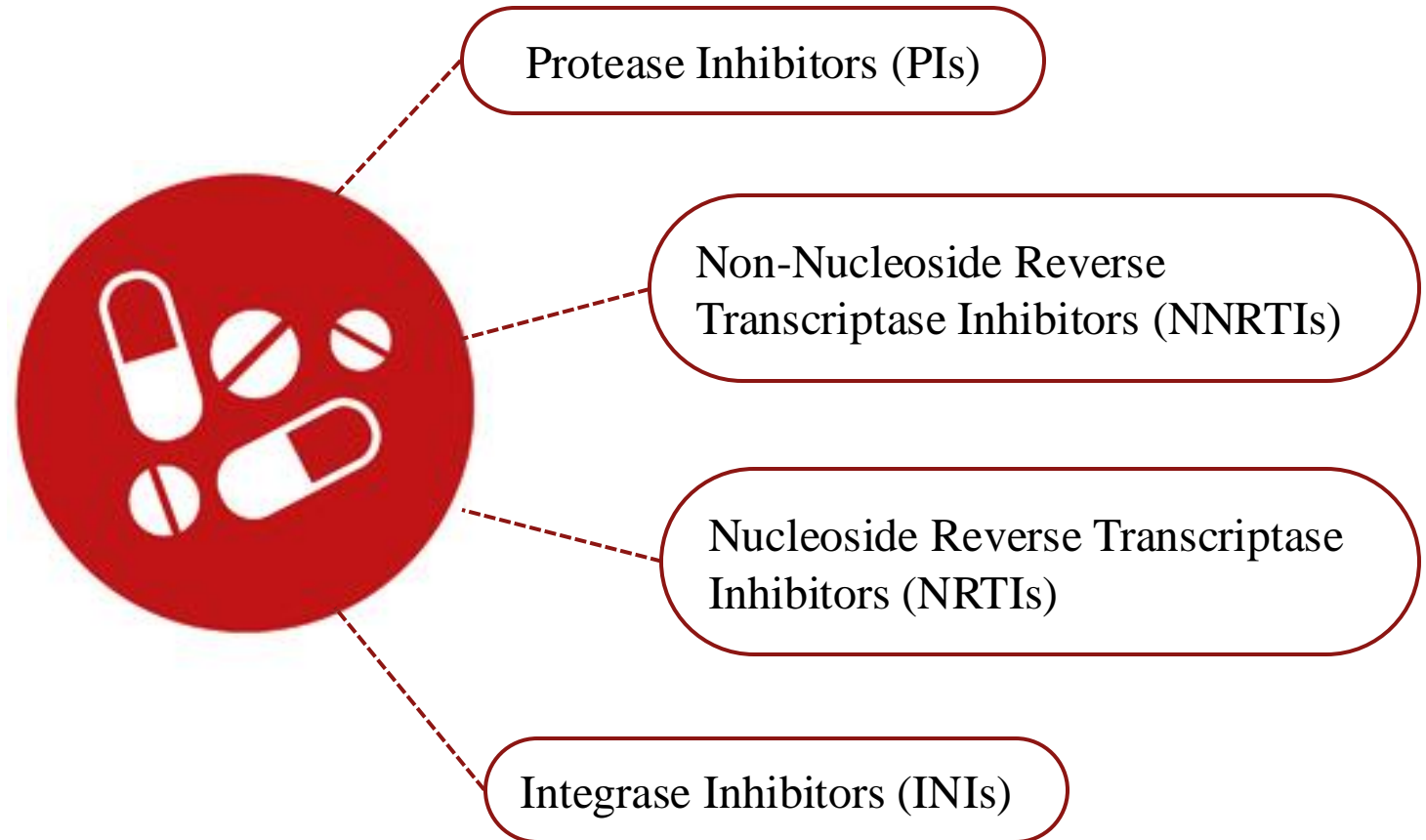
E. Stolbova^{1,2}, L. Stolbov¹, D. Filimonov¹, V. Poroikov¹, O. Tarasova¹

¹ *Institute of Biomedical Chemistry, Moscow, Russia;*

² *National Research University Higher School of Economics, Moscow, Russia*

High Active Antiretroviral therapy

- Some individuals experience poor therapeutic outcomes due to HIV drug resistance
- HIV resistance can lead to increased viral load, disease progression and higher rates of HIV transmission



Research objectives

In this study we aim to develop quantitative machine learning models that utilize genotype-phenotype data to predict HIV resistance to antiretroviral drugs:

- **NNRTIs** – efavirenz (EFV) and nevirapine (NVP)
- **NRTIs** – lamivudine (3TC), abacavir (ABC), zidovudine (AZT), stavudine (D4T), didanosine (DDI), tenofovir disoproxil fumarate (TDF)
- **INIs** - raltegravir (RAL) and elvitegravir (EVG)

Quantitative prediction of HIV drug resistance to **protease inhibitors**

Stolbova, E. A., Stolbov, L. A., Filimonov, D. A., Poroikov, V. V., & Tarasova, O. A. (2024). Quantitative Prediction of Human Immunodeficiency Virus Drug Resistance. Viruses, 16(7), 1132. <https://doi.org/10.3390/v16071132>

Materials & Methods



Genotype-phenotype data

HIV isolate sequence + Resistance (FR) to ART

Data filtering

- Missing values
- Sequences with inaccurately defined amino acids
- Duplicates

Seq PISPIETVPVKLKP...
A1 A3
A2 A4

peptide length = 5
peptide overlap = 2

Target variable	Log ₁₀ (FR)	Seq1	Seq2
Descriptors	A1	1	1
	A2	0	1
	A3	0	1
	A4	1	0
	⋮	⋮	⋮
	An	1	1

1 – presence of a specific peptide
0 – absence of specific peptide

Model evaluation
R² & RMSE

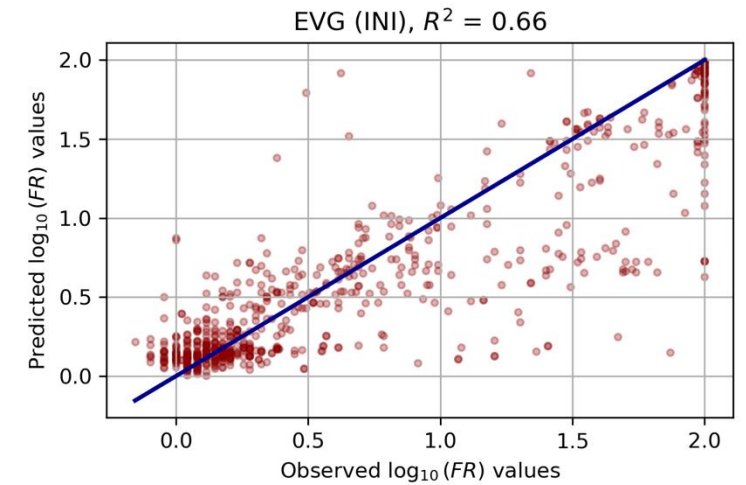
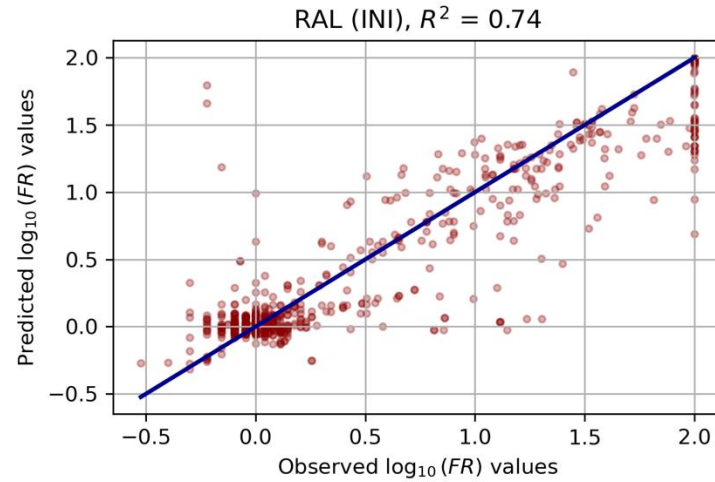
5-fold cross-validation

Random Forest Regression (RFR)

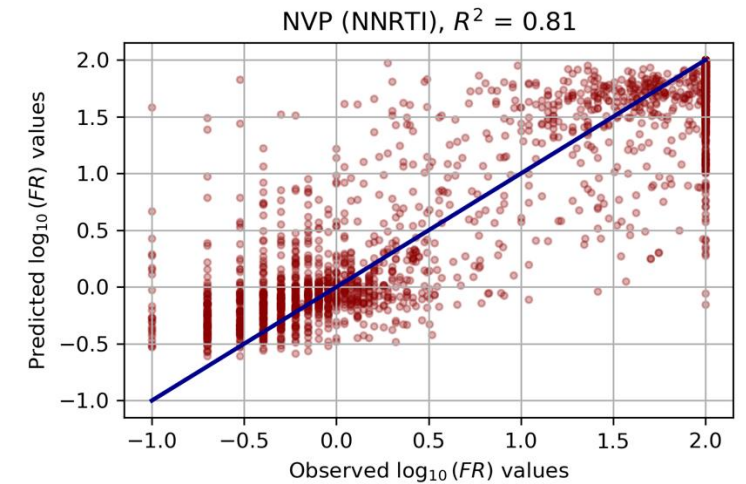
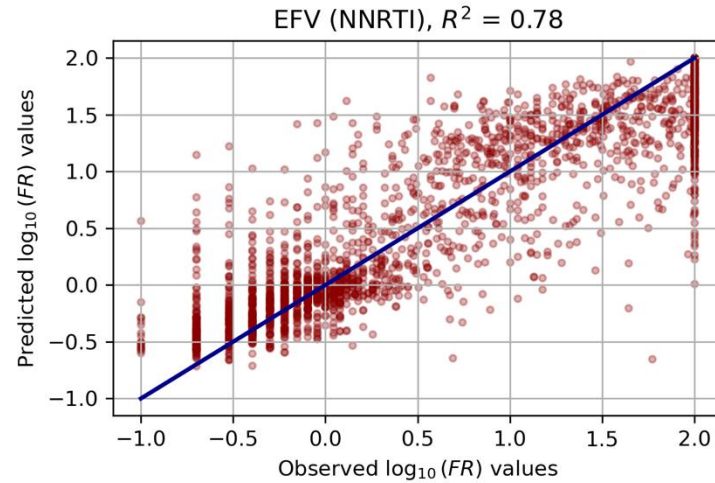
Support vector regression (SVR)

Results (INI & NNRTI)

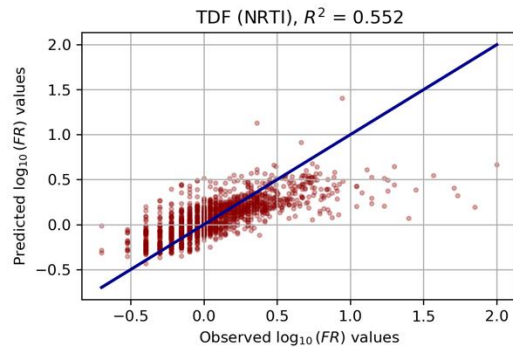
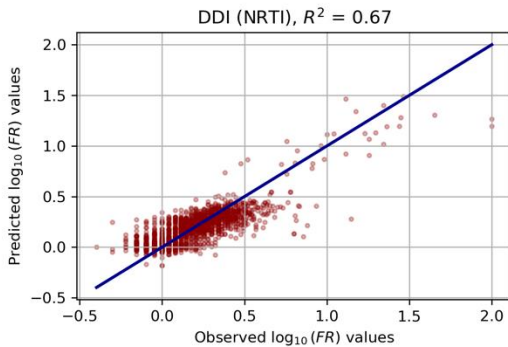
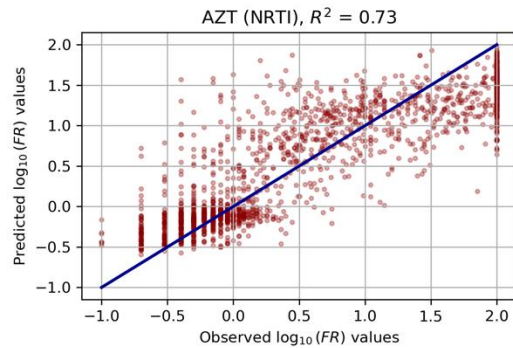
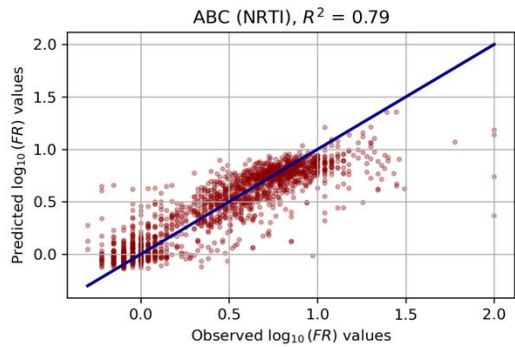
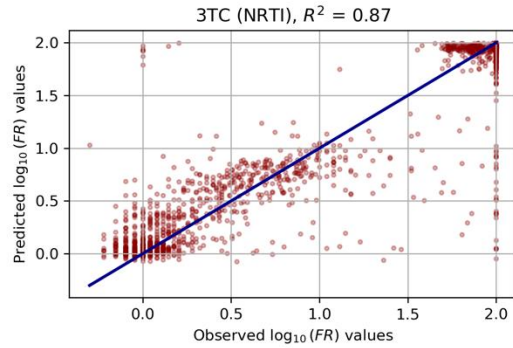
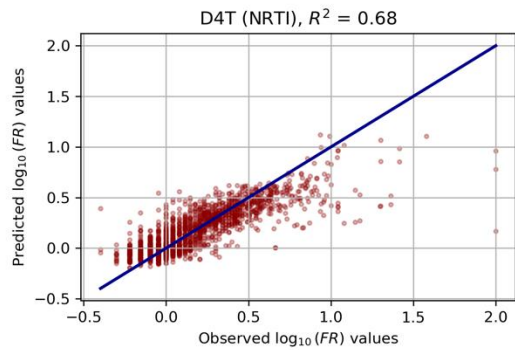
INI				
Drug	RFR		SVR	
	R ²	RMSE	R ²	RMSE
RAL	0,754	0,287	0,774	0,272
EVG	0,662	0,352	0,673	0,35



NNRTI				
Drug	RFR		SVR	
	R ²	RMSE	R ²	RMSE
EFV	0,785	0,43	0,793	0,42
NVP	0,805	0,452	0,776	0,489



Results (NRTI)



20 described resistance-associated mutation patterns*

NRTI				
Drug	RFR		SVR	
	R^2	RMSE	R^2	RMSE
3TC	0,871	0,311	0,83	0,359
ABC	0,79	0,186	0,792	0,186
AZT	0,732	0,449	0,763	0,424
D4T	0,684	0,169	0,689	0,17
DDI	0,674	0,133	0,636	0,140
TDF	0,552	0,212	0,595	0,199

12 described resistance-associated mutation patterns *

* Rhee, Soo-Yon et al. *Antimicrobial agents and chemotherapy* vol. 48,8 (2004): 3122-6. doi:10.1128/AAC.48.8.3122-3126.2004

Conclusions

We built quantitative models predicting HIV drug resistance to NNRTIs, NRTIs and INIs using RFR and SVR algorithms and genotype-phenotype data

These models may be useful for:

- analyzing HIV resistance to drugs used in HAART regimens
- predicting the efficacy of the specific drugs and drug combinations for a particular individual
- optimizing treatment strategies



This study is supported by the Program for Basic Research in the Russian Federation for a long-term period (2021–2030) (No. 124050800018-9)

**Thank you for
your attention**