Jesper Kjær
Copenhagen HIV Programme
University of Copenhagen,
Faculty of Health Sciences
Panum Institute/Building 21.1
Blegdamsvej 3B
DK-2200 CopenhagenN
+45 3545 5771
jkj@cphiv.dk

Are non-B subtypes less susceptible to antiretroviral drugs? – a bioinformatical approach to prediction of non-B subtype susceptibility.

J Kjær[1]; L Høj[1]; A Cozzi-Lepri[2]; Z Fox[1,2]; JD LundgrenD[1,3]

[1]Copenhagen HIV Programme (CHIP), Denmark; [2]Department of Infection & Population Health, Royal Free Campus, University College London, UK; and [3]Centre for Viral Diseases/KMA, Rigshospitalet, Copenhagen, Denmark

BACKGROUND

The first generation of antiretrovirals that have been developed and used to treat primarily subtype B infected patients are finally becoming available in African countries. The subtype B is however not the dominant subtype in Africa and question remains if these subtypes are less susceptible to these drugs. We will here try to answer this question by use of artificial neural networks (ANNs).

METHOD

We have previously showed that, based on the physiochemical properties of the amino acids, ANNs are able to extrapolate predictions to non-B subtypes with high accuracy (**see box below)** despite being trained only with subtype B sequence data and a matching IC50 fold change (FC).

The ANNs we use predict the log(IC 50 FC) susceptibility for protease or reserve transcriptase for a given drug based on the difference between physiochemical properties of the amino acids in HXB2 sequence (log(IC50 FC)=0) and any protease or reserve transcriptase presented to the ANN (**Figure 3**).

We applied the use of these ANNs to further predict the log(IC 50 FC) susceptibility for the most dominant subtypes in Africa: A, C, G and CRF_AG to the drugs: abacavir, amprenavir, atazanavir, didanosine, efavirenz, indinavir, lamivudine, lopinavir, nelfinavir, nevirapine, ritonavir, saquinavir, stavudine, tenofovir, zalcitabine, zidovudine.

We obtained reference sequences for these subtypes from the Los Alamos HIV sequence database. These subtypes where used for input to the ANNs to predict the log(IC50 FC) for each of the above drugs. We used an unpaired t-test to identify significant differences in mean of the log(IC50 FC) values between B-subtype and non-B subtypes.

RESULTS

From Los Alamos we extracted 6 sequences for subtype A, 3 sequences for subtype B and 4 sequences for each of the subtypes C, G and CRF_AG resulting in 96 log(IC50 FC) predictions for subtype A, 16 for subtype B and 64 for each of the subtypes C, G and CRF_AG.

The predicted mean IC50 FC across the 16 drugs for each subtype was:

Subtype	mean Log(IC50 FC)	mean IC50 FC
A:	0.16	1.18
B:	-0.13	0.90
C:	-0.30	0.74
G:	0.27	1.31
CRF_AG:	-0.14	0.87

Plots of the mean predictions for the individual drugs and for each subtype is shown in the **Figure 1**. Please note that these values have been transformed from the originally predicted log(IC50 FC) and each data point represents the mean per drug and subtype.

We found no significant difference in log(IC50 FC) for subtypes C or CRF_AG compared to the subtype B log(IC50 FC) mean. But there was a significant ($p<0.05$) difference between B and A ($p=0.03$) and B and G ($p=0.02$) predictions.

Detailed log(IC50 FC) predictions are listed in **Figure 2**. Drugs with a predicted IC50 FC › 2 fold are:

Subtype:	Drug (mean IC50 FC)
A:	nelfinavir (2.4) and zidovudine (2.7)
C:	nelfinavir (3.2)
G:	amprenavir (2.2), atazanavir (4.1), nelfinavir (4) and zidovudine (2.5)
CRF_AG:	atazanavir (2.5), nelfinavir (3.2) and zidovudine (2.8).

Excluding these drugs from the analysis resulted in only subtype C (mean IC50 FC=0.70) to be significantly different in susceptibility compared to subtype B ($p=0.04$).

LIMITATIONS

The extracted reference sequences from the Los Alamos data base constitutes a small and selected dataset. A few unusual polymorphisms in some of these sequences can easily have become a bias in the results presented here. Of note the low log(IC50 FC) values for subtype B which were expected closer to 0 for all drugs and for the subtypes C and CRF_AG hypersusceptibility to the NNRTIs are being predicted. But instead of using HXB2 (IC50 FC=1 for all drugs given the design of the ANNs) for comparison we decided to use the subtype B reference sequences from Los Alamos to allow for the same level of bias to the subtype B predictions as we would have to accept in regards to subtype A, C, G and CRF_AG. Only a larger dataset including larger subtype variation would allow to adjust for this and to investigate specific polymorphisms that cause significant change in IC50 FC values.

CONCLUSION

Predictions showed significant difference between the predicted mean IC50 FC values for the both A and G HIV-1 subtypes compared to IC50 FC for subtype B. This is in part driven by reduced susceptibility to amprenavir, atazanavir, nelfinavir and zidovudine and does also affect susceptibility for C and CRF_AG subtypes. To the majority of the drugs we did not find any reduced susceptibility to the drugs for the subtypes and C subtypes may even have an increased susceptibility to these drugs.

Kjær J, Høj L, Fox Z, Lundgren JD.: Prediction of phenotypic susceptibility to antiretroviral drugs using physiochemical properties of the primary enzymatic structure combined with artificial neural networks. HIV Medicin Vol. 9, issue 8, start page 642:

"RT sequence data was not available for non-B subtypes therefore we could not validate the use of ANNs for either NRTIs or NNRTIs. For PR we found 3 subtype C sequences for atazanavir; 15 subtype C and one subtype D sequences for amprenavir, indinavir, nelfinavir, ritonavir and saquinavir; and 16 subtype C and one subtype D sequence for lopinavir. This allowed for a total of 100 IC50 FC values to be predicted for primarily C subtypes by ANNs trained with unique subtype B sequences. The IC50 FC values were mainly in the susceptible range with a mean IC50 FC value of 1.25 (IQR: 0.70 – 5.84) and with 32% observed and 29% predicted above IC50 FC › x3 fold. The overall correlation coefficient between the observed and the predicted IC50 FC values was found to be 0.89 (0.77 for nelfinavir, 0.80 for lopinavir, 0.81 for amprenavir, 0.95 for indinavir and 0.96 for atazanavir, ritonavir and saquinavir, respectively). The overall PPV for all 100 predictions was 0.86. "

Figure 1

Plots of predicted mean log(IC50 FC) susceptibility for each subtype to the 16 drugs - see method section and Figure 2 for the list of drugs

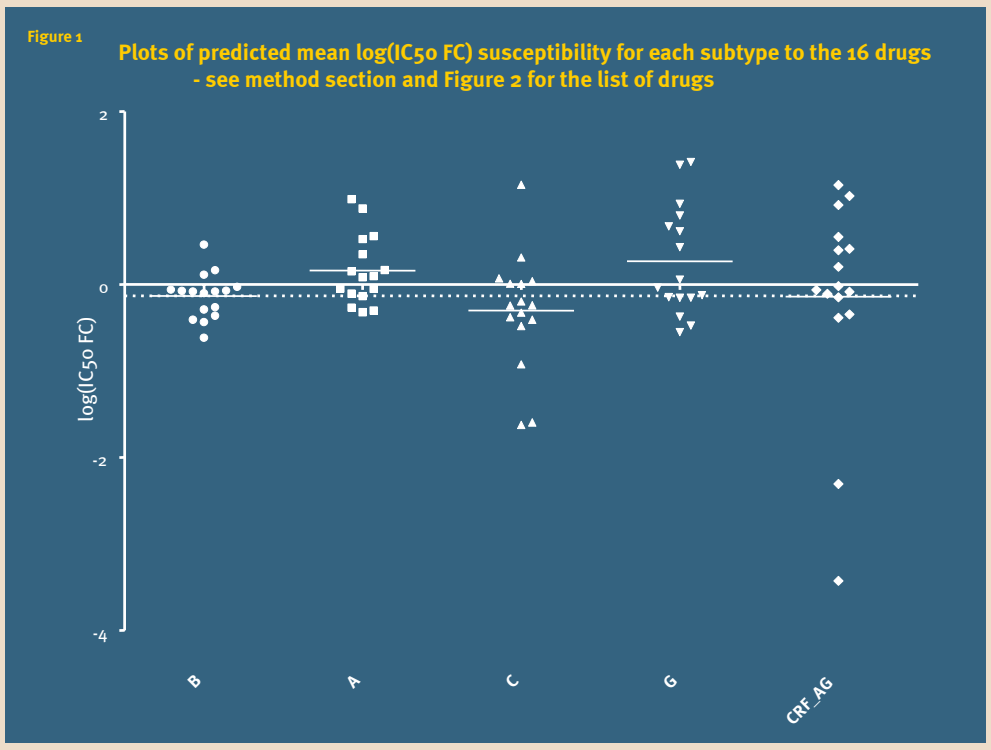


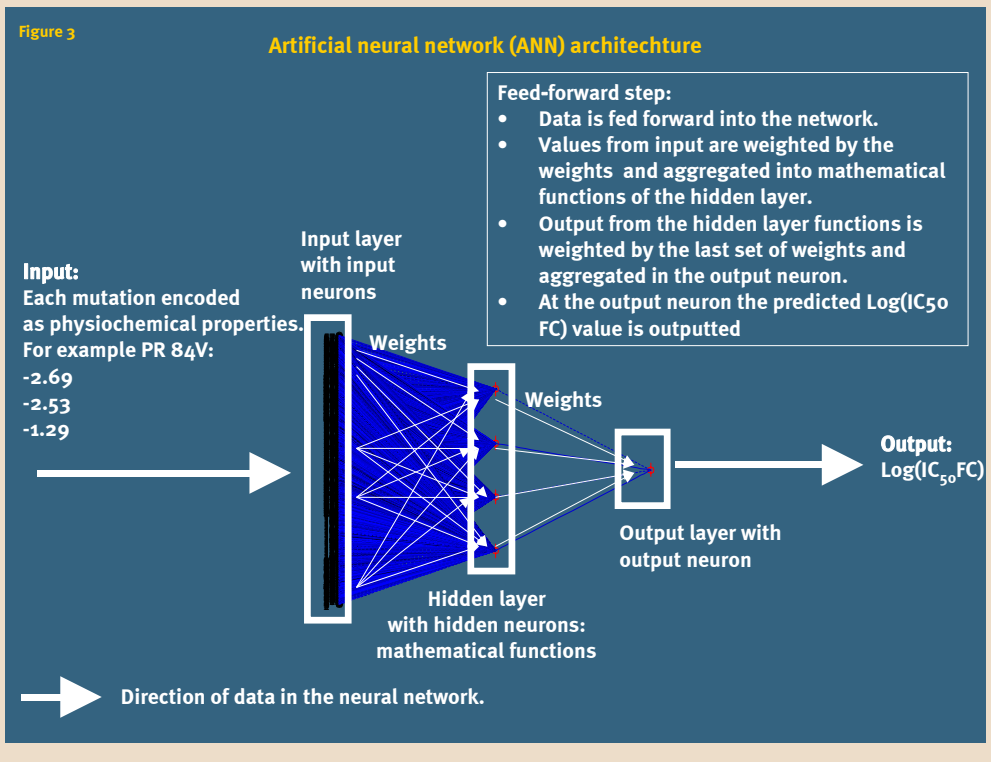
Figure 2

Predicted mean log(IC50 FC) susceptibility for each subtype and each of the 16 drugs - a value of zero equals IC50 FC = 1, anything above zero indicates reduced susceptibility

		Subtype				
		B	A	C	G	CRF_AG
NRTI	Abacavir	0.11	0.08	0.07	-0.04	0.20
	Didanosine	-0.06	-0.10	0.00	-0.15	-0.02
	Lamivudine	0.17	0.10	-0.20	-0.47	-0.38
	Stavudine	-0.08	-0.05	-0.33	-0.12	-0.11
	Tenofovir	-0.07	-0.30	-0.92	-0.37	-0.35
	Zalcitabine	-0.07	-0.05	0.01	-0.15	-0.08
	Ziduvudine	-0.03	0.99	0.04	0.93	1.02
NNRTI	Efavirenz	-0.10	-0.27	-1.60	0.43	-3.42
	Nevirapine	0.46	-0.32	-1.62	-0.55	-2.30
PI	Amprenavir	-0.41	0.17	-0.41	0.80	-0.06
	Atazanavir	-0.43	0.53	0.31	1.41	0.92
	Indinavir	-0.36	0.35	-0.24	-0.15	0.40
	Lopinavir	-0.26	-0.13	-0.38	0.62	0.55
	Nelfinavir	-0.08	0.88	1.15	1.39	1.15
	Ritonavir	-0.29	0.56	-0.24	0.67	0.41
	Saquinavir	-0.62	0.15	-0.48	0.06	-0.15

Figure 3

Artificial neural network (ANN) architechture



Download poster at: www.cphiv.dk

CHIP COPENHAGEN HIV PROGRAMME

FACULTY OF HEALTH SCIENCES UNIVERSITY OF COPENHAGEN

REGION H Rigshospitalet