NGS application for mutational stability monitoring of viral drug targets

Youri V. Kravatsky

Engelhardt Institute of Molecular Biology of Russian Academy of Sciences, Vavilov str., 32, Moscow, 119334, Russia, jiri@eimb.ru

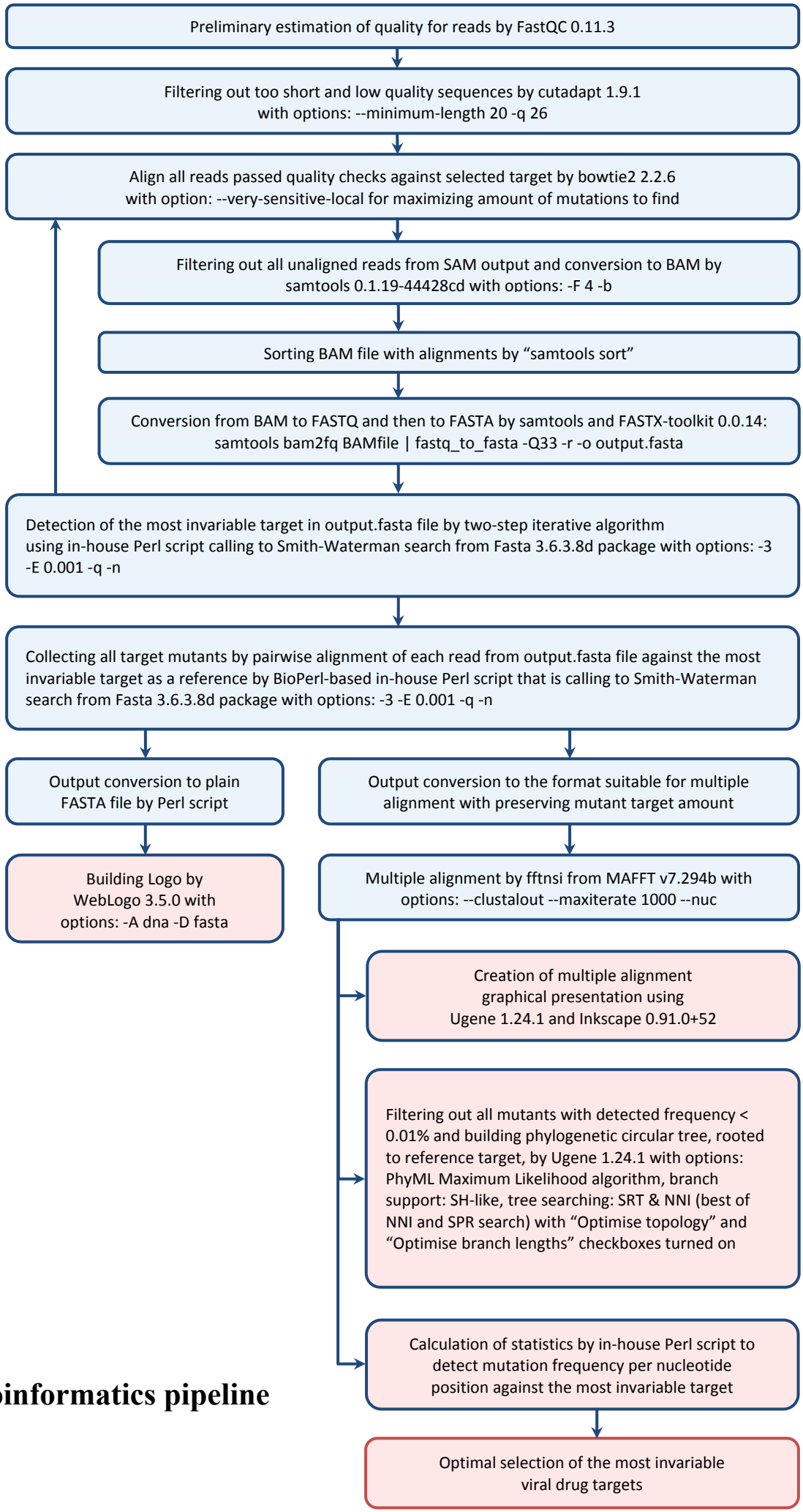
Vladimir R. Chechetkin

Engelhardt Institute of Molecular Biology of Russian Academy of Sciences, Vavilov str., 32, Moscow, 119334, Russia, vladimir_chechet@mail.ru

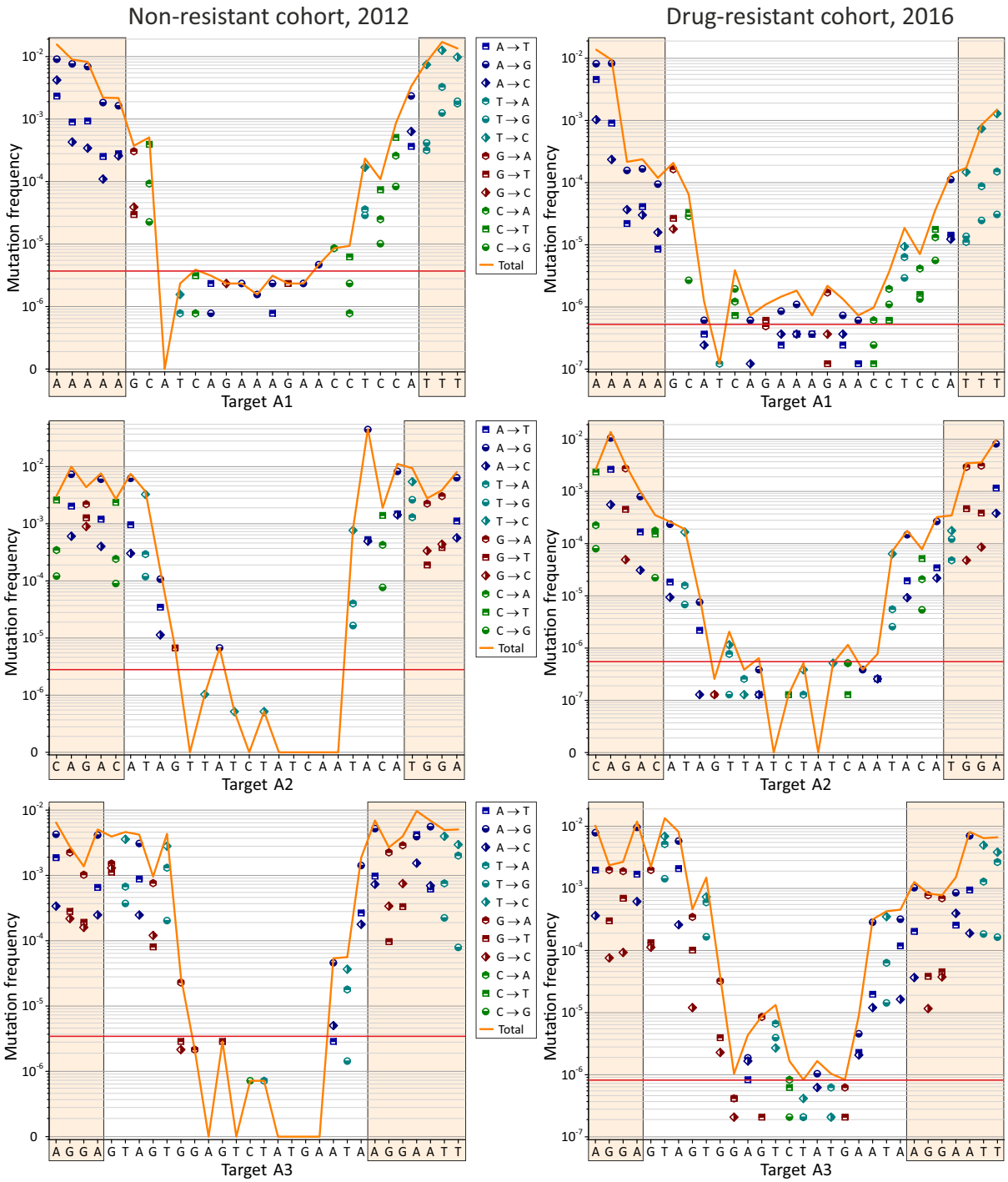
Nikolai A. Tchurikov

Engelhardt Institute of Molecular Biology of Russian Academy of Sciences, Vavilov str., 32, Moscow, 119334, Russia, tchurikov@eimb.ru

The efficient development and application of anti-viral drugs need the permanent monitoring of strict correspondence between a drug and a relevant highly variable viral DNA or RNA target. Such monitoring can be successfully performed by NGS (next generation sequencing) methods. NGS is able to provide both the assessment of general target invariability and the frequency of particular mutations in the different target sites. The computer processing of short reads corresponding to a drug targets obtained by ultra-deep sequencing remains a challenging task. Here we present the development and deployment of bioinformatics pipeline aimed to resolve this problem. The suggested bioinformatics pipeline combines the available programs and the original *ad hoc* scripts based on an original algorithm of the search for the most invariable targets in the ultra-deep sequencing data. We present also the statistical criteria for the threshold of reliable mutation detection and for the assessment of variations between counterpart data sets. The small variations in mutation frequencies detected by the NGS may have genetic significance and may be used for the early diagnostics of diseases. As an example, the bioinformatic pipeline is applied to the study of invariability of RNA interference (RNAi) targets in HIV-1 subtype A. We demonstrated that the Dicer substrates corresponding to the conserved HIV-1 targets efficiently attack the target RNA in a non-viral system using *luc* reporter genetic constructs in cultured human cells. The data are important for both the understanding of the patterns of HIV-1 mutability and properties of reverse transcriptase, and for the development of gene therapy approaches using RNAi for the treatment of HIV/AIDS.



Bioinformatics pipeline



Curves showing the frequencies of nucleotide substitutions along the 27–30-bp RNAi targets for both non-resistant and drug-resistant independent cohorts. The horizontal red lines correspond to the thresholds of reliable mutation detection. The frequencies were determined against the most invariable RNAi targets. Lower frequencies of mutations are observed in the middle part of the RNAi targets. Non-shaded regions of 19-bp length (core sequences) are proved experimentally to be effective RNAi targets [1,3].

Percentage of 19-bp target invariability for different targets and cohorts.

	Non-resistant cohort, 2012		Drug-resistant cohort, 2016	
Target	Aligned reads	Invariability, percent	Aligned reads	Invariability, percent
A1	1 291 242	99.17 ± 0.01	8 390 187	97.71 ± 0.01
A2	1 956 875	92.40 ± 0.02	7 816 744	99.58 ± 0.00
A3	1 418 446	96.84 ± 0.02	4 862 688	96.68 ± 0.01

We used two independent cohorts of patients from Russia. Non-resistant cohort includes five isolates taken in 2012 of HIV-1 subtype A from patients who were not receiving antiretroviral therapy. Drug-resistant cohort includes four isolates taken in 2016 from patients who received antiretroviral therapy for five years and still possessed high-titer viremia. Therefore, the cohort was considered as containing the drug-resistant HIV-1 strains. Such cohorts can mimic the variability of RNAi targets for multi-strain HIV-1 patients and/or the development of multiple strains over time for particular patient.

Bioinformatics pipeline diagram, links to all software used and source codes of all scripts are located at the website http://virmut.eimb.ru

This work was supported by a grant from the Russian Science Foundation (project no. 15-14-00005).

1. O.V.Kretova *et al.* (2017) Analysis of variability in HIV-1 subtype A strains in Russia suggests a combination of deep sequencing and multitarget RNA interference for silencing of the virus, *AIDS Res Hum Retroviruses,* **33(2):**194-201.

2. Y.V.Kravatsky *et al*. (2016) Mutation frequencies in HIV-1 subtype-A genome in regions containing efficient RNAi targets, *Mol Biol (MSK),* **50(3):**480-485.

3. N.A.Tchurikov *et al*. (2016) Conserved sequences in the current strains of HIV-1 subtype A in Russia are effectively targeted by artificial RNAi *in vitro*, *Gene,* **583(1):**78-83.