REVIEW

Of the Doctoral Dissertation of Gurevich Alexey Alexandrovich «Computational methods for analysis of error-prone metabologenomic data» submitted for defense of the degree of candidate of physico-mathematical sciences, specialization 03.01.09 – Mathematical biology, bioinformatics.

This Ph.D. thesis presents an ambitious computational approach to identifying peptide natural products from large mass spectra databases that have been collected by a consortium of metabolic physiologists. The thesis approaches two major challenges: genome assembly algorithmic improvements and discovery and mass spectra database screening for peptide natural products.

Issues are presented in four detailed chapters that consider:

1. Development and description of a bioinformatics tool kit, QUAST, that is designed to evaluate genome assemblies and assemblers by comparing de novo assemblies to a reference genome as a standard.
2. The extension and improvement of QUAST to MetaQuast specifically detecting metabolomics data development and discovery.
3. Application of the QUAST tools and derivative corrected assemblies for searching databases. The Dereplicator pipeline of mathematical and statistical evaluation is described and de demonstrated by searching the newly available Global Natural Products Social (GNPS) molecular networking infrastructure including a billion of mass spectra of natural products
4. The strength and weakness of the combined genome assembly improvement and GNPS screening for peptide natural product is discussed and evaluated and an in improved VARQuest algorithm for variable peptide natural product identification and characterization is proposed and described.

The project is ambitious and an important novel bioinformatics approach with multiple benefits for metabolite evaluation, function definition and even therapeutic potential. The thesis is well-organized, clearly presented, easy to read and provocative.

Most of the thesis chapters have been published in peer-reviewed reports cited as references 24, 25, 28 and 29. Overall the thesis is very strong, thorough, comprehensive and professionally presented.

вх 0912-270 от 04.12.18

Because the work has already gone through thorough copy-editing and peer review for the most part, I did not annotate typographical corrections. I have little substantive suggestions for improvement. I would suggest they include a summary suitable for a non-specialist so their work will be appreciated by a larger readership; this might be an expansion of the Conclusions in the front.

This is an excellent body of research wrk and I happily recommend the awarding of the Ph.D. degree to Alexey Alexandrovich Gurevich in light of this excellent and important presentation.

30.11.2018

Date, Signature

Stephen J. O'Brien, Ph.D.
Chief Scientific Officer
Theodosius Dobzhansky Center for Genome Bioinformatics
St. Petersburg State University
St. Petersburg, Russia

Подпись
ЗАВЕРЯЮ
30.11.2018