



Galaxy in teaching computational methods of genome analysis for master degree students in Medical Genetics Program at the Faculty of Medicine, Vilnius University



E.Pranckevičienė, L.Ambrozaitytė, I.Uktverytė, A.Utkus, V.Kučinskas,

Department of Human and Medical Genetics, Faculty of Medicine, Vilnius University

Background

Master program in Medical Genetics is offered by the Department of Human and Medical Genetics, Faculty of Medicine, Vilnius University. In this program a computational analysis of genomic data constitutes a considerable part of practical exercises. In “Biotechnology and fundamentals of bioinformatics analysis” course students are introduced to a computational pipeline of next generation sequencing (NGS) data analysis. For practicum in these courses we used a local Galaxy server running on Hardware-as-a-Service (2 Hexa core Intel Xeon CPU E5-2630L, 8 processing units, 8192 Mb RAM and 320 Gb disk space) and main Galaxy usegalaxy.org.

Practicum on computational analysis of NGS exome data

Main objectives of this practicum are (i) to learn steps of computational analysis of a sequenced exome; (ii) to understand the rationale of the analysis and (iii) to be able to interpret obtained results and formulate hypothesis for further analysis. Students used original publically available exome data (FASTQ, Illumina) [1]. They performed quality assessment and alignment of original data to a reference human genome on the local Galaxy server. Each student was assigned to analyze one chromosome. Using main Galaxy server [2] they on their own identified genomic variants specific to tumor exome and compared their findings with published results. Students annotated identified genomic variants and characterized the genes harboring mutations.

1. **Spector MS. et al.** Mast-cell leukemia exome sequencing reveals a mutation in the IgE mast-cell receptor β chain and KIT V654A. *Leukemia*. 2012 Jun; **26(6)**. PMID22173243.
2. **Blakenberg D. et al.** Galaxy: a web-based genome analysis tool for experimentalists. *Current Protocols in Molecular Biology*. 2010 Jan; Chapter 19:Unit 19.10.1-12.

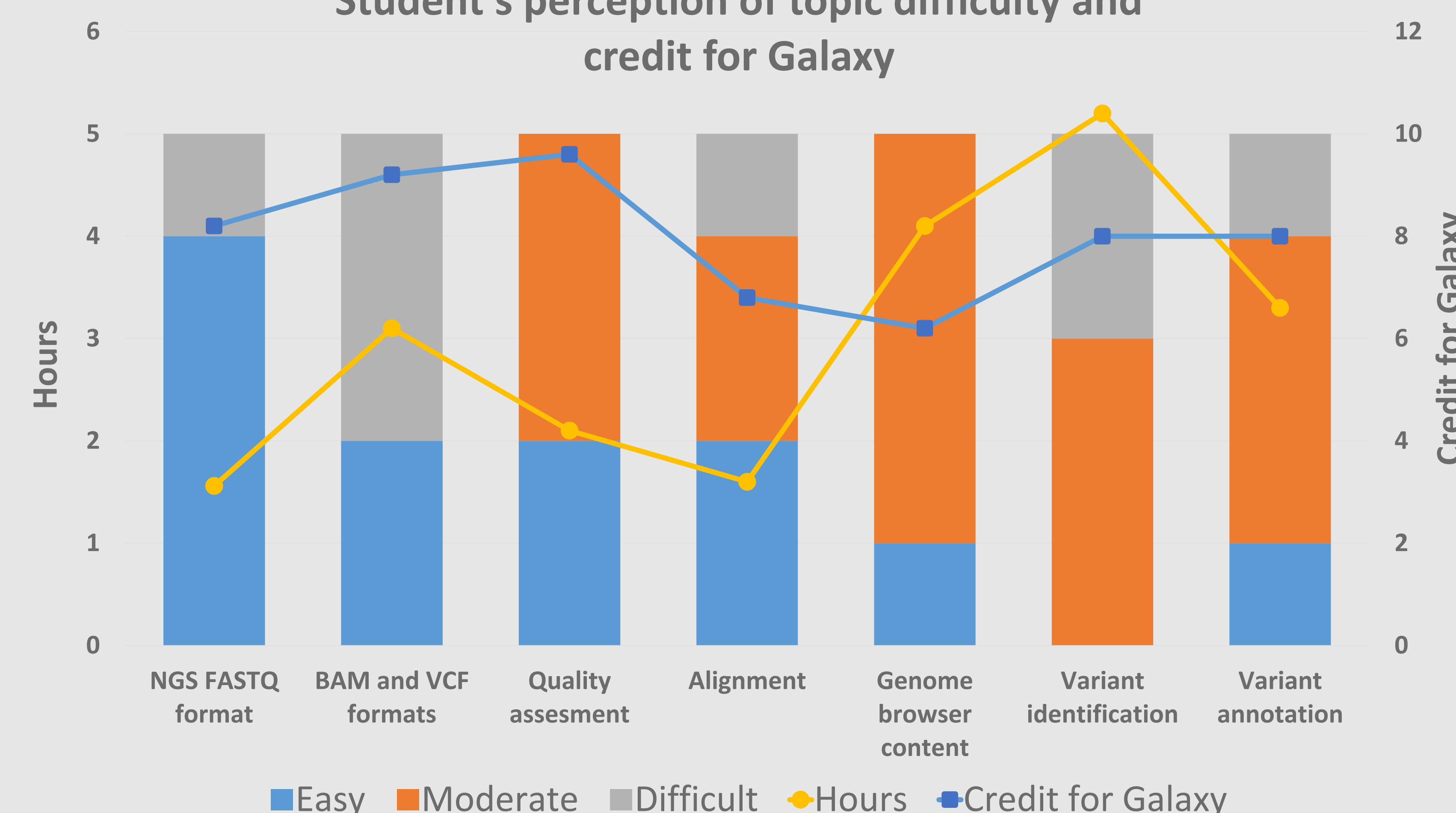
Concluding remarks

Using Galaxy in teaching and learning is a novel approach at the Department of Human and Medical Genetics. We aimed to understand a tendency towards how students perceive use of Galaxy: whether it helps them to better understand principles of computational analysis of exome data. It was noted that *introduction to Galaxy by a teacher is important* for successful completion of the practicum. Getting acquainted with Galaxy instruments students felt increasing interest in exome data analysis itself. In addition, a possibility to import the data from external sources into Galaxy environment boosted students curiosity in general. Current study has a very small scale. Nevertheless, we understood a level of complexity of computational tasks that can be taught and learned using Galaxy in 2.5 credit practicum. **Student’s view of Galaxy experience: absolutely new experience, very useful to know, will use in the future, understood analysis of exome data.**

Student’s self-identification

Future work	Wet lab , clinical research, data analysis, administration.
Education	Genetics , biochemistry, molecular biology.
Broader interest	Synthetic biology, stem cells, nutrition and health, public relations.
Prior knowledge	Reference genome, quality assessment, variant annotation.
Technical skills	Spreadsheets, basic statistics, R environment, no command line experience.

Student's perception of topic difficulty and credit for Galaxy



Lithuanian society of Human Genetics is greatly acknowledged for their support.

We thank L.Kasiliauskienė, K.Margevičiūtė, E.Gasperskaja, G.Olišauskaitė, J.Petraškaitė for their participation. J.Kasperīūnienė is acknowledged for her help with questionnaire.