

BITS :: Call for Abstracts 2019 - Oral communication

Type	Oral communication
Session	Bioinformatics challenge in precision medicine
Title	An overview of the clinical bioinformatics landscape in Italy
All Authors	Ceol A(1), Angeli D(2), Bucci G(3), Crisafulli G(4), De Summa S(5), Molineris I(4), Cereda M(6), Dugo M(7), Giacobbe L(1), Guida A(1), Fancello L(1), Frige' G(1), Pallocca M(8), Pozdnyakov I(1), Sardina DS(9), Scalera S(8), Zapparoli E(3), Mazzarella L(1), for the ACC genomics and bioinformatics working group.

Affiliation

- (1) IEO, European Institute of Oncology IRCCS, Milan, Italy
- (2) Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) IRCCS, Meldola, Italy
- (3) IRCCS Ospedale San Raffaele, Milan, Italy.
- (4) Candiolo Cancer Institute - IRCCS, Candiolo, Italy
- (5) Molecular Diagnostics and Pharmacogenetics Unit, IRCCS-Istituto Tumori "Giovanni Paolo II", Bari, Italy
- (6) Italian Institute for Genomic Medicine, Torino, Italy
- (7) Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy
- (8) SAFU Unit, IRCCS Regina Elena National Cancer Institute, Rome, Italy
- (9) AOUP "P. Giaccone", Cellular and Molecular Oncology Laboratory, Palermo, Italy

Motivation

Bioinformatics has become the backbone of health science, including proteomics, structural biology, imaging, genomics and epigenomics among others. Academic research facilitates the development and improvement of new, cutting-edge methods and algorithms, but research-oriented tools are often not the best choice in the clinical setting where reproducibility, automation, user-friendliness and computational parsimony are often more important than innovation. Clinical environments, however, differ greatly at both diagnostic and research levels. We here provide the results of a small pilot project to map the current landscape of Italian clinical bioinformatics in research-oriented clinical institutions, highlighting critical issues and identifying methods for improvement.

Methods

We surveyed research-oriented clinical centers affiliated with Alleanza Contro il Cancro [1], a network of 26 research-oriented clinical cancer centers, for multiple issues related with clinical bioinformatics including personnel, infrastructures, data management and the usage of Laboratory Information Management Systems (LIMS), pipelines and reproducibility, data privacy and security. We obtained data from 9 centers which were analyzed and discussed within the ACC genomics and bioinformatics working group (WG)

Results

Personnel: At least one (range 1-5) bioinformatician supports clinical diagnostics and research in each institute. Additional support is provided by researchers from within the institutes (for instance from research departments), or academy (e.g. local universities). The WG agreed that the growth of the clinical bioinformatics community can be limited by several factors, most notably i) the lack of an acknowledged working career, ii) the limited dedicated training, resulting in the difficulty of recruiting competitive bioinformaticians.

Infrastructures: Dedicated workstations and virtual machines are common. Most but not all institutes (4/9) have access to local high-performance computing (HPC) infrastructures. In 4 cases the computational weight is further supported by private cloud resources (like AWS, Google or Azure), or public resources such as GARR [2] and CINECA [3], outside of the resources allocated to the ACC WG, reserved to ACC projects. The WG highlighted that awareness of such publicly available HPC resources is still low. One of the reasons could be the lack of communication among researchers and their sys admins (Italian CEDs).

Data management: One institute has chosen a proprietary LIMS for managing the samples. 3 institutes have developed their own solution. In addition, the ACC workgroup as well has developed its own LIMS, whose usage recently started to be extended to single institutes. Finally, the WG reported a strong interest for sharing experience and solutions in order to leverage the time to be dedicated to software development and management.

Analytical pipelines: Standard pipelines (e.g. GATK-based [4]), and proprietary platforms (BaseSpace [5], Ion Reporter [6]) are preferred to novel methods and algorithms. The WG stressed reproducibility as a key

issue. Work within ACC relies on versioning system and continuous integration. The reproducibility is further improved by the usage of container-based technologies.

Privacy and data security: 3 institutes report increased awareness of sensible data protection after the advent of the GDPR [7] in 2018. Pseudonymization of samples, anonymization of meta data have been redefined, and cloud solutions have been often avoided or discontinued. The WG stressed that clear and standard directives on how to deal with patient data, defined at the national and/or European levels, are urgently needed.

Conclusions: This pilot project helped identify critical issues of the clinical bioinformatics landscape in Italy. This will serve as the basis for a more detailed and extensive survey that will help to standardize methodologies and facilitate data exchange across clinical institutes.

Info

[1] <https://www.alleanzacontroilcancro.it/>

[2] <https://www.garr.it>

[3] www.cineca.it

[4] The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA, 2010 GENOME RESEARCH 20:1297-303

[5] <https://basespace.illumina.com>

[6] <https://ionreporter.thermofisher.com>

[7] <https://eugdpr.org/>

Figure

-

Availability

-

Corresponding Author

Name, Surname Arnaud, Ceol

Email arnaud.ceol@ieo.it

Submitted on 29.04.2019

Società Italiana di Bioinformatica

C.F. / P.IVA 97319460586

E-mail bits@bioinformatics.it

Sede legale Viale G. Mazzini, 114/B - 00195 Roma

Website bioinformatics.it