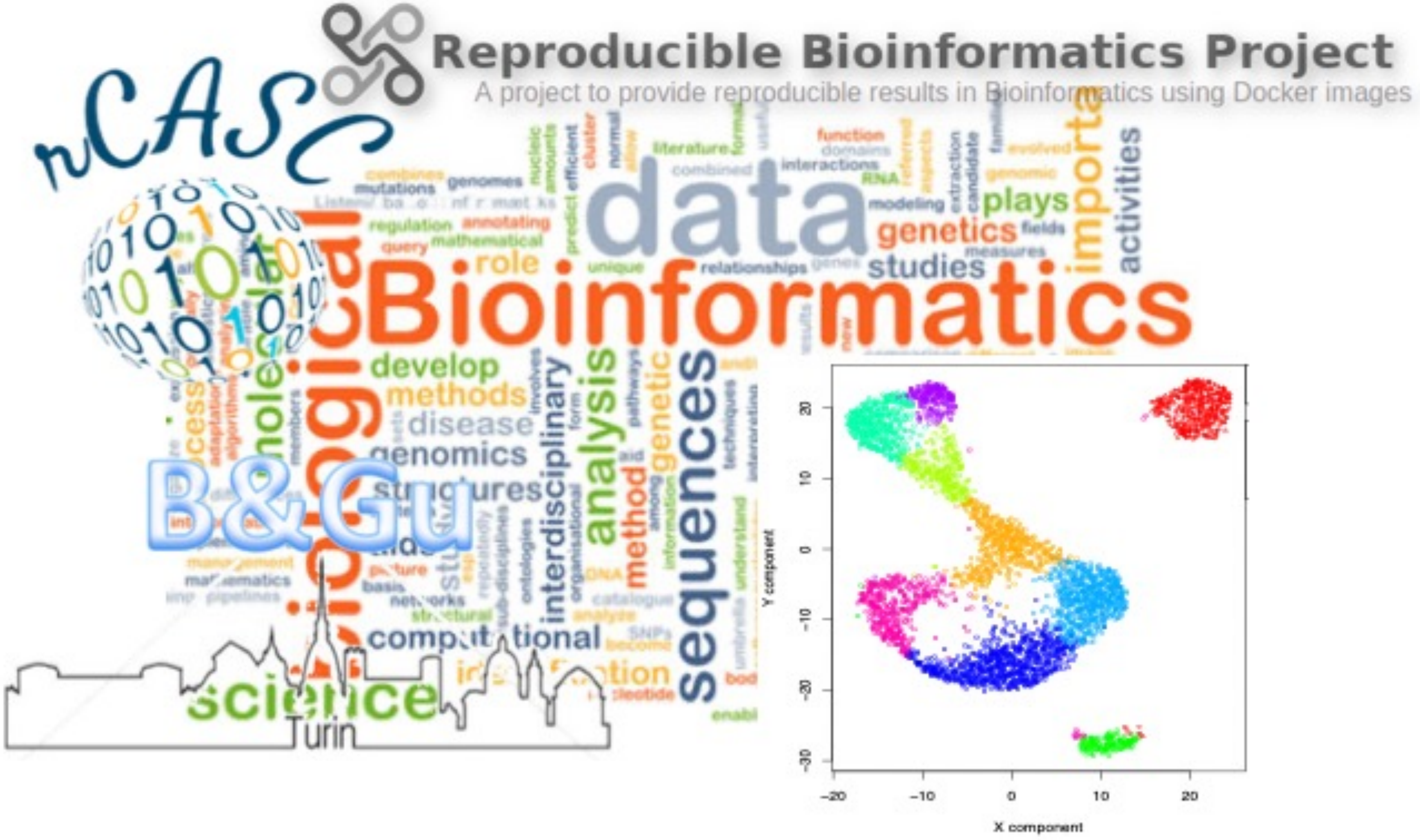
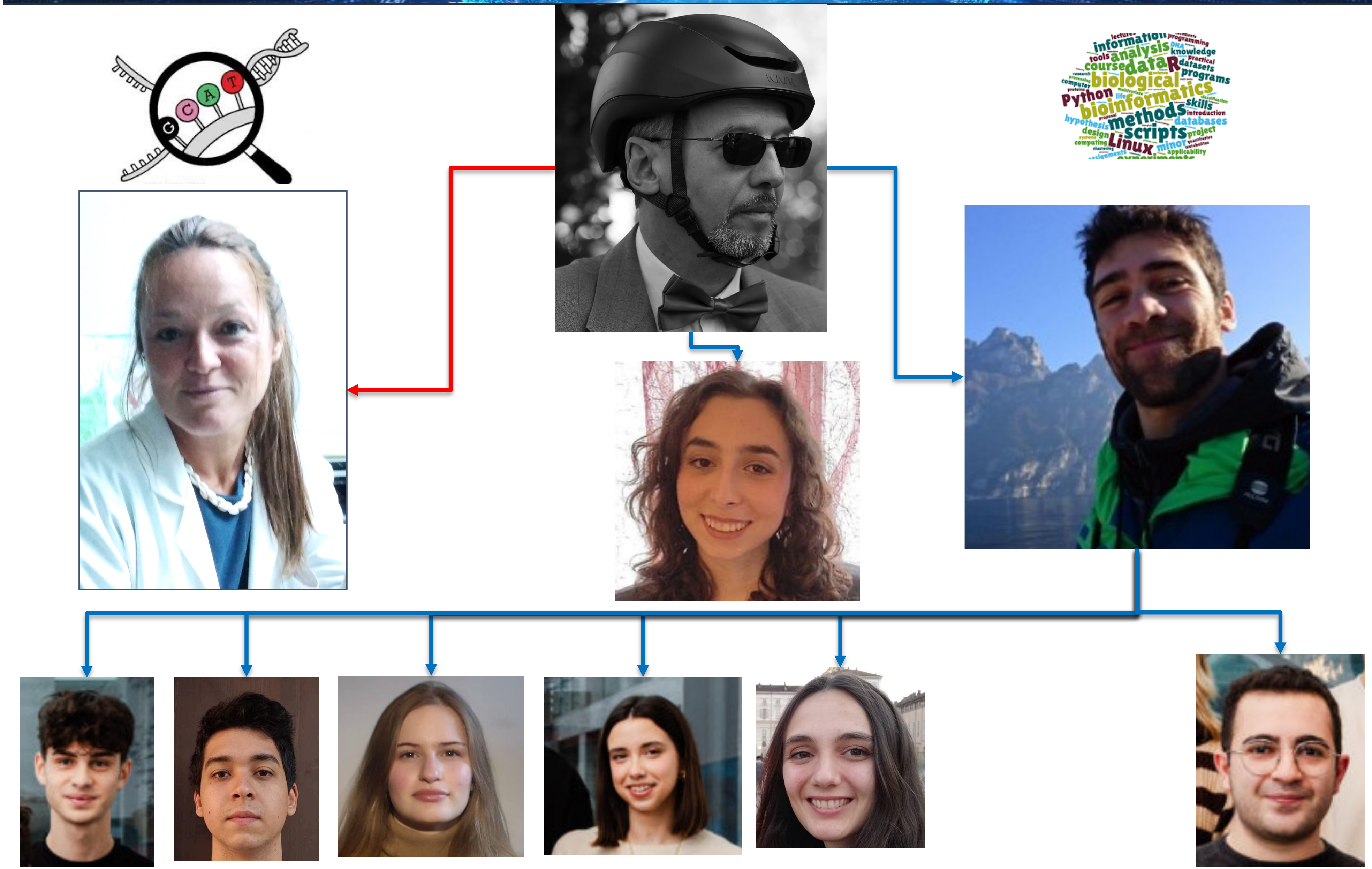
# Calogero’s Lab & BGcore



# BGcore

## Bioinformatics & Genomics core-lab

M B C



# Our activities

- Our research group is dedicated to pioneering advancements in data reproducibility and single-cell omics analysis.
- We're driving progress in bioinformatics and wet lab support for bulk and single-cell/spatial transcriptomics with a strong commitment to transparency, reliability, and innovation.

BGcore

CalogeroLab

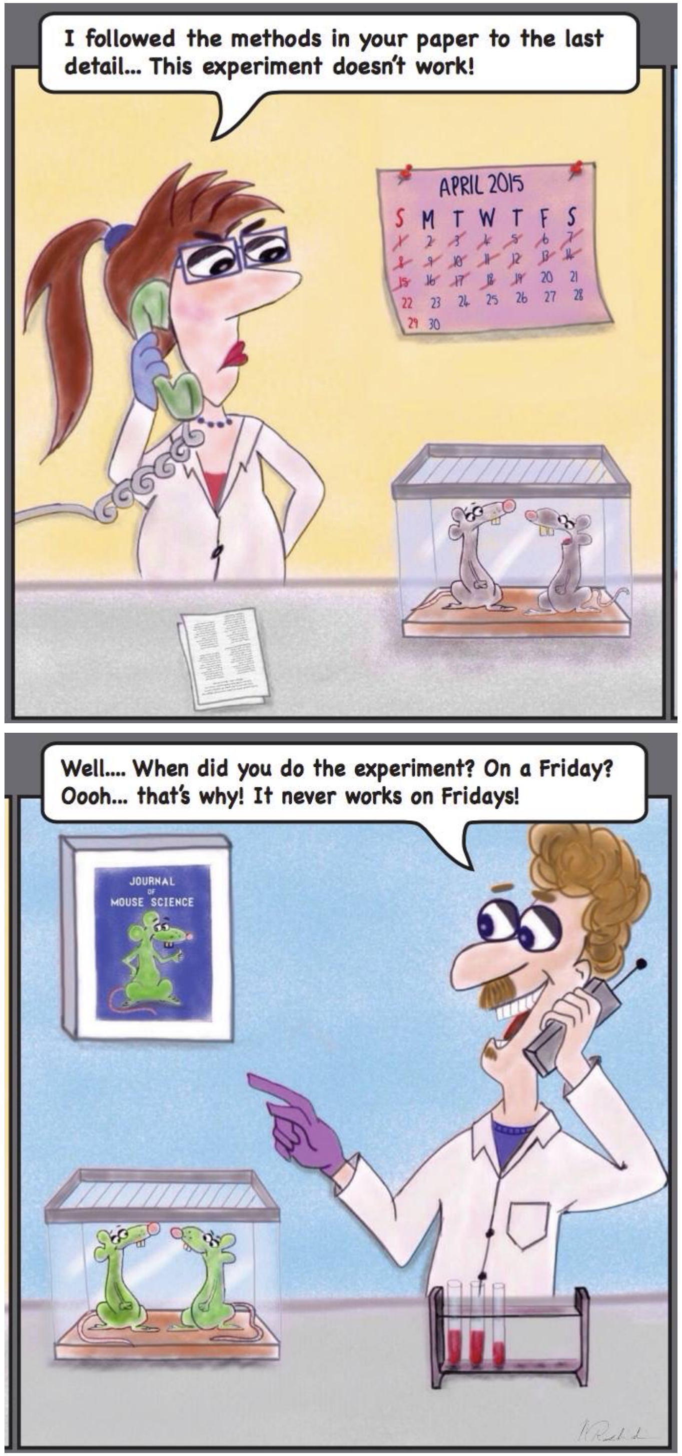
# Our activities

- **Reproducible Bioinformatics Project:**
  - We lead the charge in promoting reproducible research practices, ensuring that our methodologies are transparent and results are reliable, with the Reproducible Bioinformatic Project.
- **Docker-Based Tools:**
  - Our tools, including docker4seq for bulk RNAseq analysis and rCASC for single-cell RNAseq analysis, leverage Docker containers for efficient, scalable, and reproducible workflows.
- **Personalized Docker Containers:**
  - We specialize in building sophisticated Docker containers, exemplified by CREDOgui, facilitating seamless deployment and execution of bioinformatics pipelines.
- **Facility for Wet Lab and Bioinformatics Support:**
  - Through BGcore, we offer comprehensive support in experimental design, data generation, and analysis for bulk and single-cell/spatial transcriptomics.

# Top 5 Factors Affecting Reproducibility in Research

- Lack of Access to Raw Data and Methodologies
- Invalidated Biological Materials
- Lack of Knowledge to Analyze Data
- Incorrect Laboratory Practices
- Undervaluing Negative Results

https://www.enago.com/academy/top-5-factors-affecting-reproducibility-in-scientific-research/#5_Factors_Affecting_Reproducibility_in_Scientific_Research

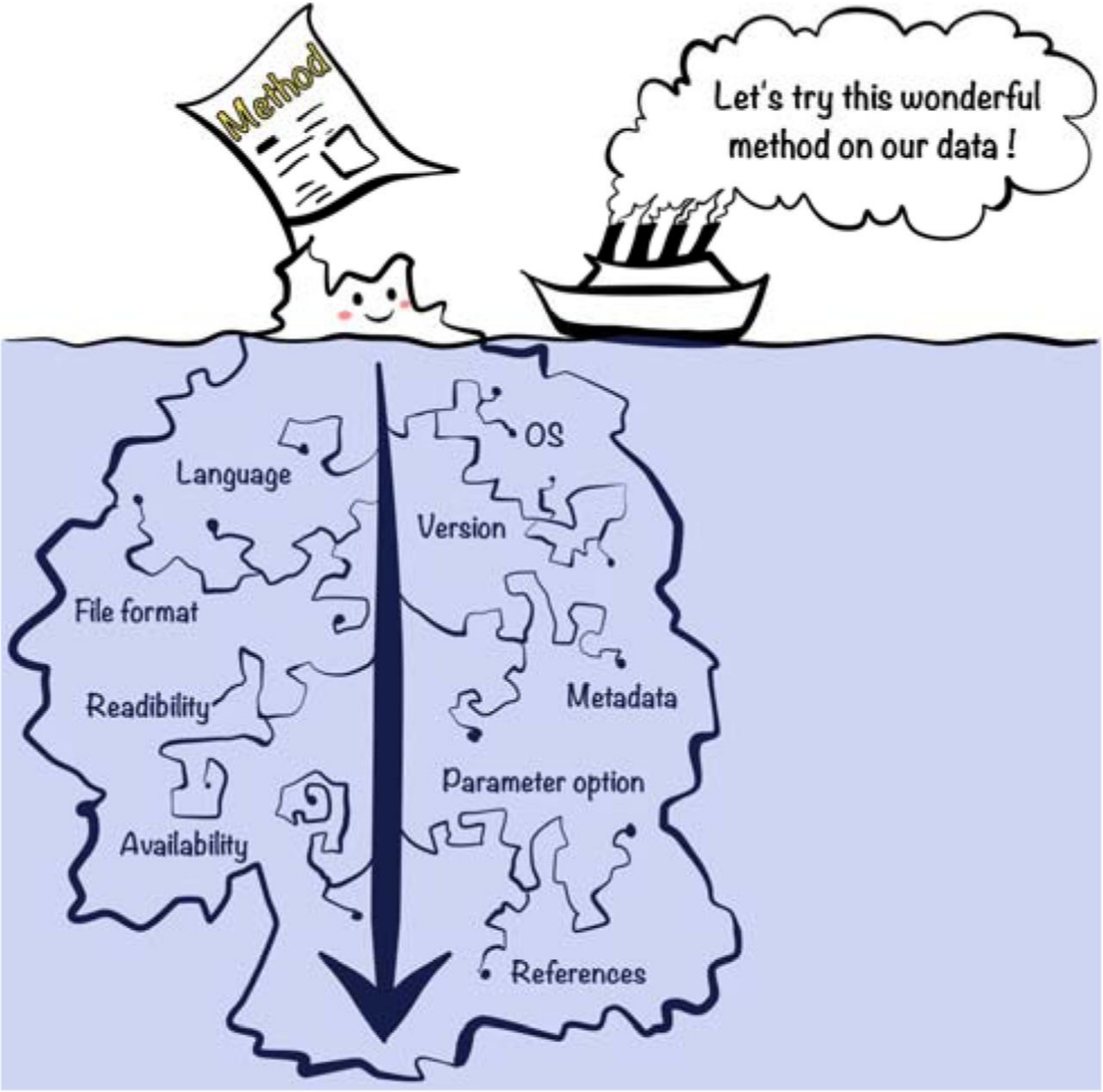


# Top 5 Factors Affecting Reproducibility in Research

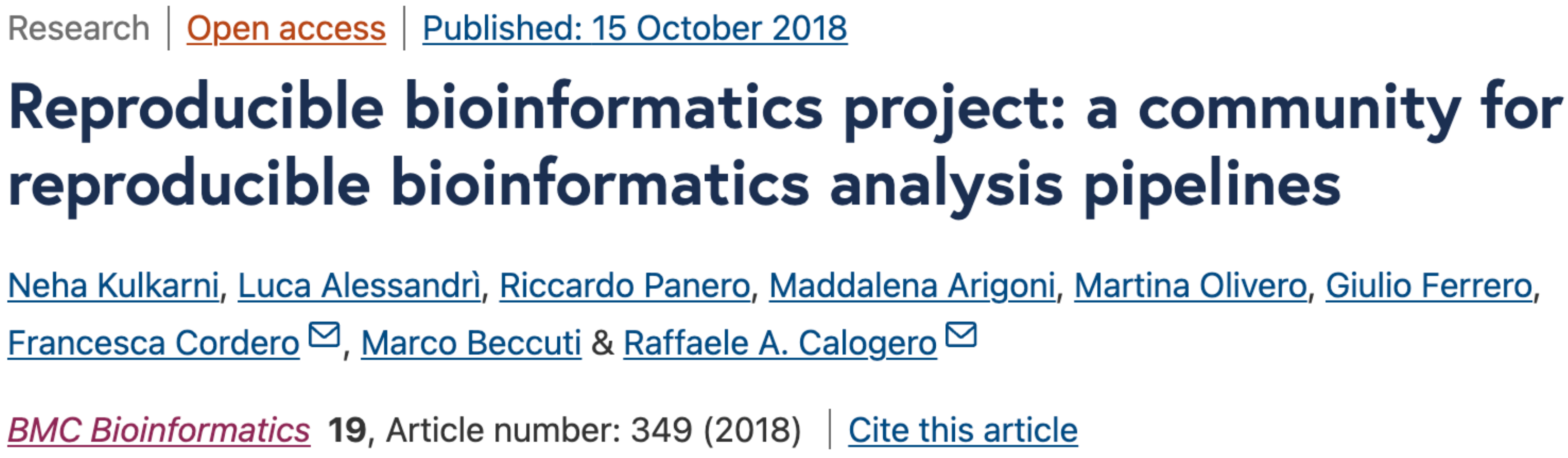
- **Lack of Access to Raw Data and Methodologies**
- Invalidated Biological Materials
- Lack of Knowledge to Analyze Data
- Incorrect Laboratory Practices
- Undervaluing Negative Results

https://www.enago.com/academy/top-5-factors-affecting-reproducibility-in-scientific-research/#5_Factors_Affecting_Reproducibility_in_Scientific_Research

# Reproducibility criticalities in Bioinformatics



# Reproducible Bioinformatics Project

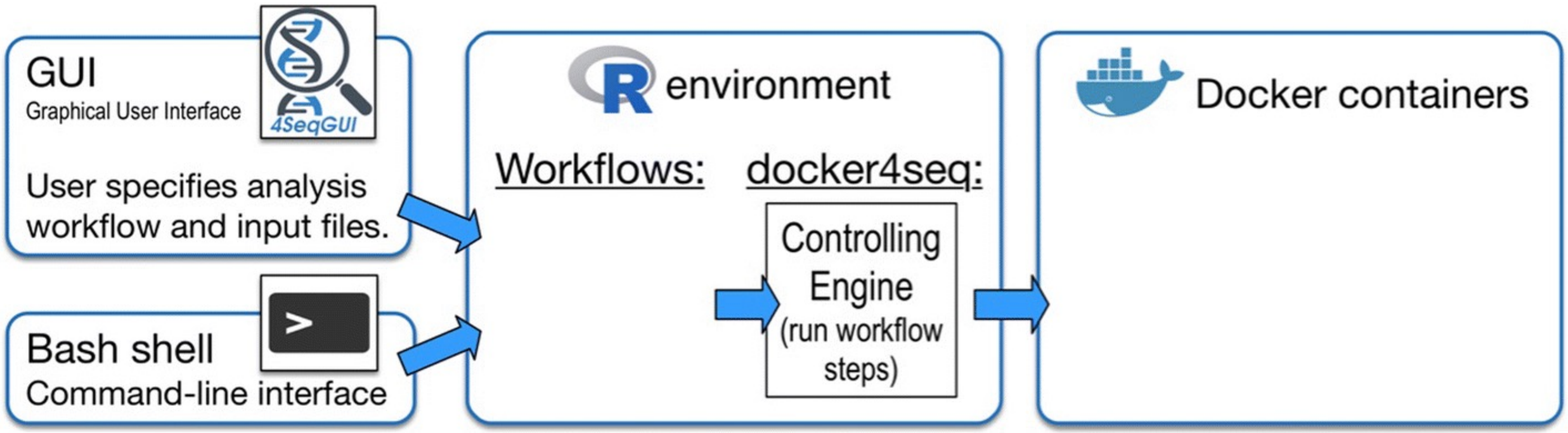


Research | Open access | Published: 15 October 2018

## Reproducible bioinformatics project: a community for reproducible bioinformatics analysis pipelines

Neha Kulkarni, Luca Alessandrì, Riccardo Panero, Maddalena Arigoni, Martina Olivero, Giulio Ferrero, Francesca Cordero ✉, Marco Beccuti & Raffaele A. Calogero ✉

*BMC Bioinformatics* **19**, Article number: 349 (2018) | Cite this article



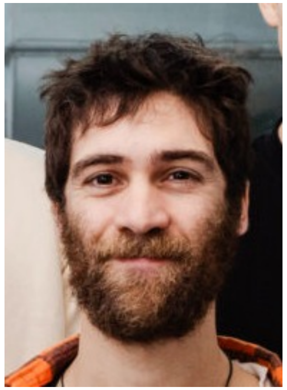
# Lack of Access to Raw Data and Methodologies

> Bioinformatics. 2018 Mar 1;34(5):871-872. doi: 10.1093/bioinformatics/btx674.

**SeqBox: RNAseq/ChIPseq reproducible analysis on a consumer game computer**

Marco Beccuti [1], Francesca Cordero [1], Maddalena Arigoni [2], Riccardo Panero [2], Elvio G Amparore [1], Susanna Donatelli [1], Raffaele A Calogero [2]

https://github.com/kendomaniac/docker4seq



L. Alessandri

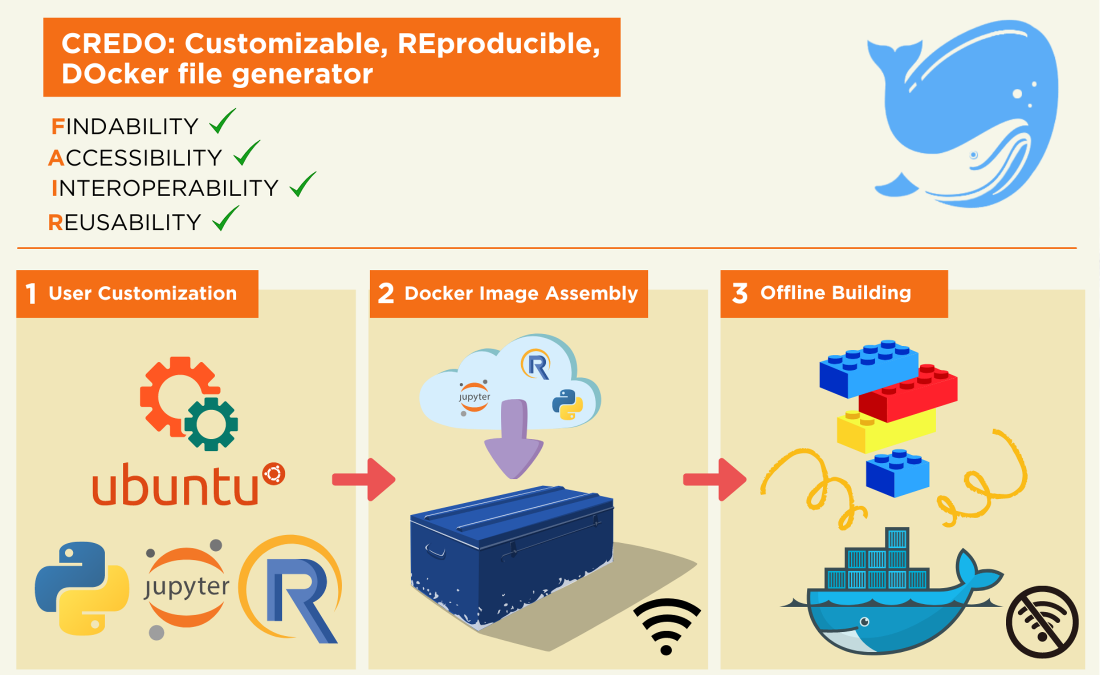
> Gigascience. 2019 Sep 1;8(9):giz105. doi: 10.1093/gigascience/giz105.

**rCASC: reproducible classification analysis of single-cell sequencing data**

Luca Alessandrì [1], Francesca Cordero [2], Marco Beccuti [2], Maddalena Arigoni [1], Martina Olivero [3], Greta Romano [2], Sergio Rabellino [2], Nicola Licheri [2], Gennaro De Libero [4], Luigia Pace [5], Raffaele A Calogero [1]

https://github.com/kendomaniac/rCASC

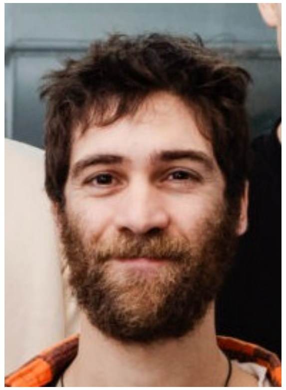
# Lack of Access to Raw Data and Methodologies



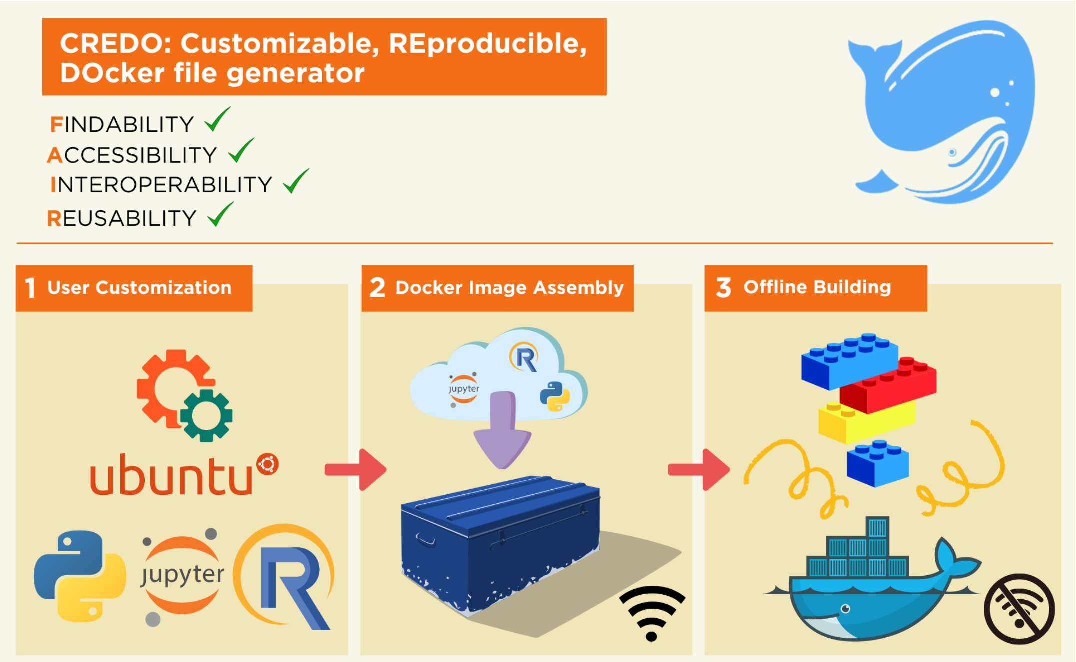
Alessandri et al. In Press BMC Bioinformatics

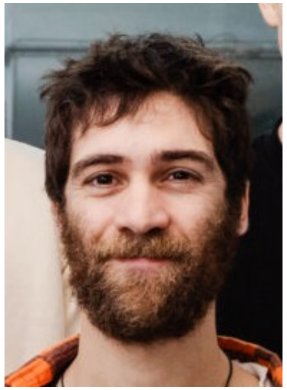
**A container is a computer science kit!**

≈



L. Alessandri

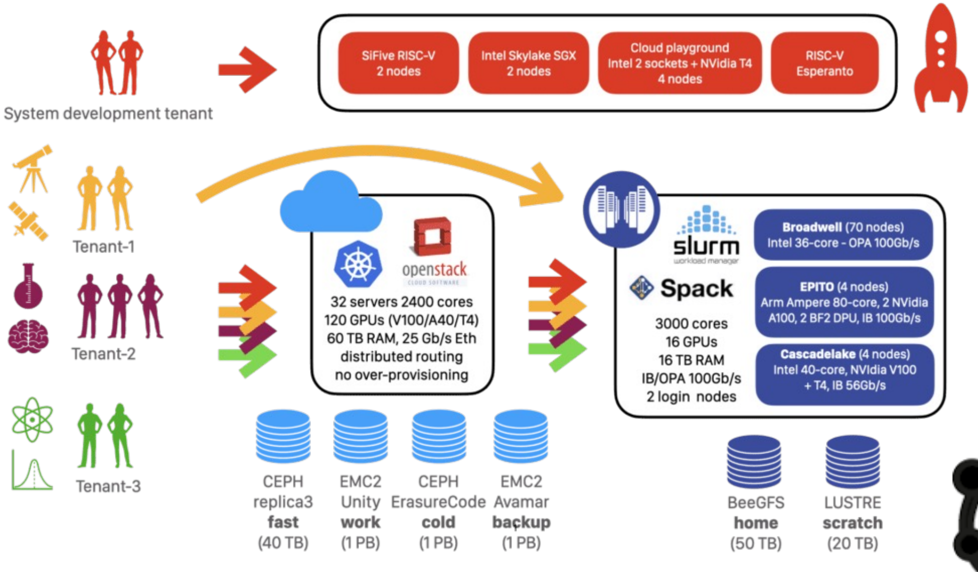




L. Alessandri

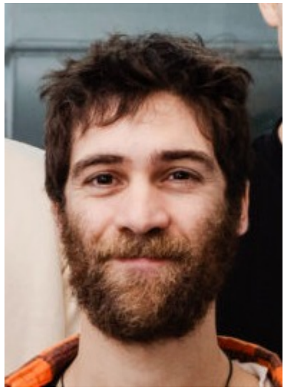
E. Martelli

Consortium GARR



HPC4AI

# Biologically driven dimensionality reduction

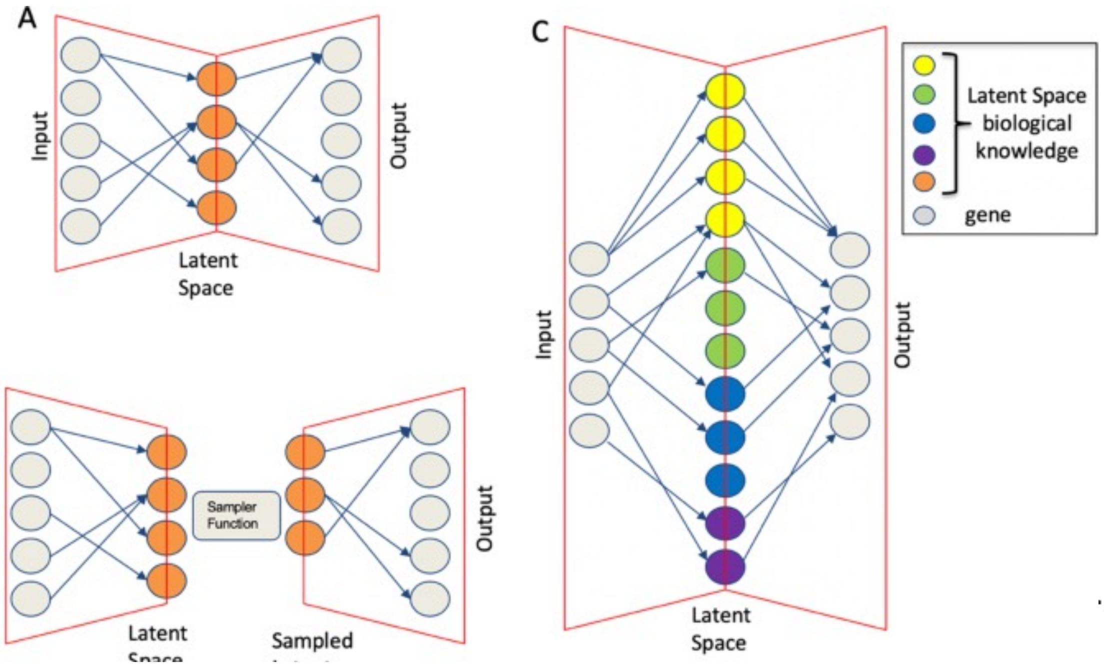


L. Alessandri

> NPJ Syst Biol Appl. 2021

**Sparsely-conn**
**RNAseq data n**

Luca Alessandri [1], Frances
Martina Olivero [3] [4], Maria

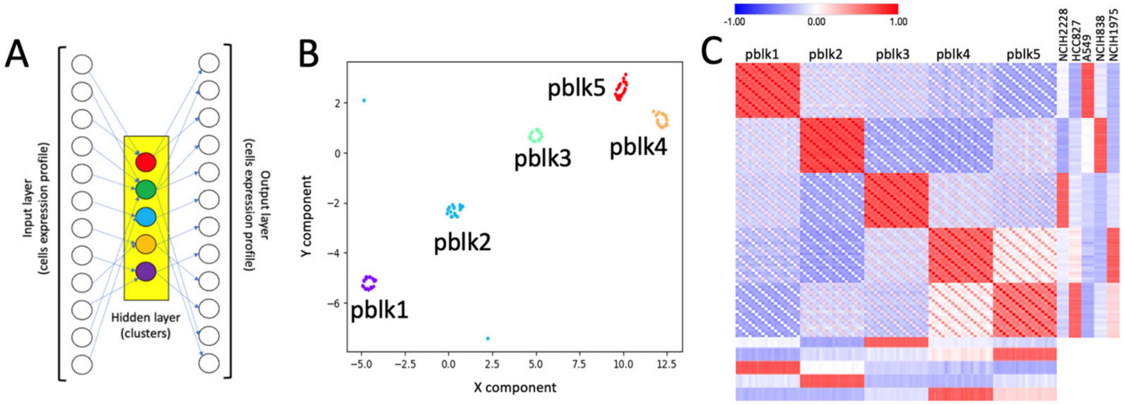


> Methods Mol B

**Function**
**Sparsel**

**g**

Luca Alessand



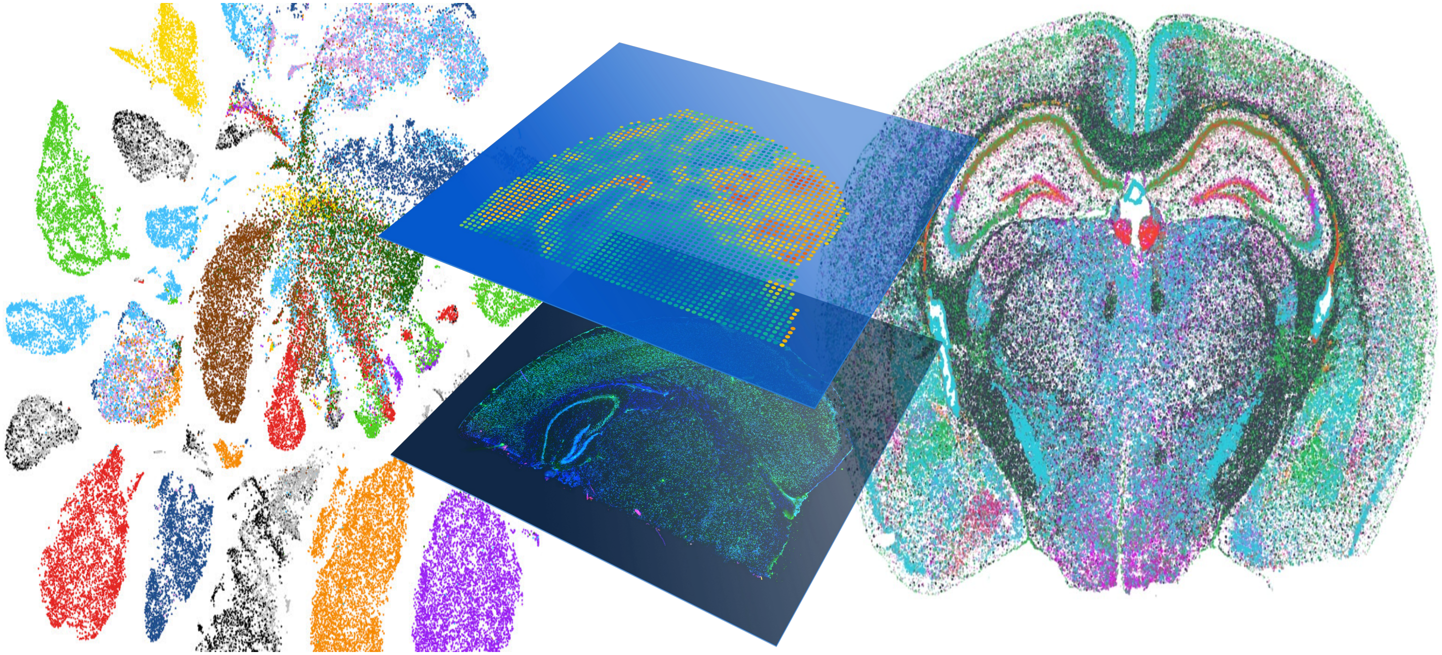
312755.

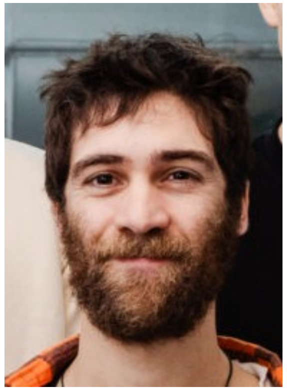
**A Multi-Purpose**

Luca Alessandri [1] [2], Maria Luisa Ratto [1], Sandro Gepiro Contaldo [2], Marco Beccuti [2], Francesca Cordero [2], Maddalena Arigoni [1], Raffaele A Calogero [1]

https://github.com/alessandriLuca/SC-ADMINE

# Spatial transcriptomics





L. Alessandri

> Gigascience. 2022 Aug 10:11:giac075. doi: 10.1093/gigascience/giac075.

**Stardust: improving spatial transcriptomics data analysis through space-aware modularity optimization-based clustering**

Simone Avesani [1], Eva Viesi [1], Luca Alessandrì [2], Giovanni Motterle [1], Vincenzo Bonnici [3], Marco Beccuti [4], Raffaele Calogero [2], Rosalba Giugno [1]

# Top 5 Factors Affecting Reproducibility in Research

- Lack of Access to Raw Data and Methodologies
- Invalidated Biological Materials
- Lack of Knowledge to Analyze Data
- **Incorrect Laboratory Practices**
- Undervaluing Negative Results

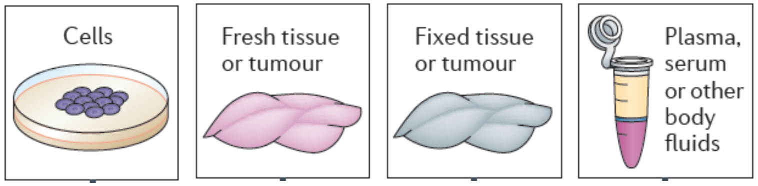
https://www.enago.com/academy/top-5-factors-affecting-reproducibility-in-scientific-research/#5_Factors_Affecting_Reproducibility_in_Scientific_Research

# BGcore
Bioinformatics & Genomics core-lab

M. Arigoni

# Services at BGcore

## I. Bulk RNA sequencing
## II. Small RNA sequencing



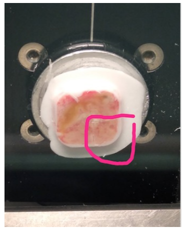
## III. Extracellular vesicles RNA and small RNA sequencing

HUMAN OR VETERINARY PATIENTS → EXOSOMES RELEASED INTO CIRCULATION → ISOLATION OF EXOSOMES RNA EXTRACTION → RNA SEQUENCING → DATA ANALYSIS → BIOMARKERS DISCOVERY →

I.DIAGNOSIS
II. RESPONSE TO TREATMENT
III. PROGNOSIS

## IV. Spatial Transcriptomics from FFPE and FF tissues

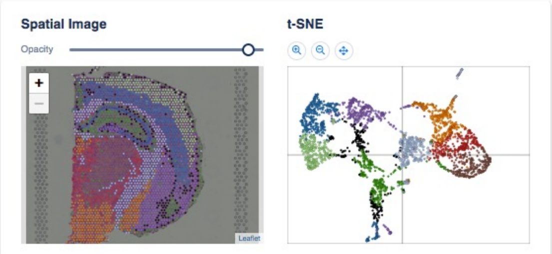
Spatial transcriptomics is a groundbreaking molecular profiling method that allows scientists to **measure all the gene activity** in a tissue sample and **map where the activity is occurring**.



IDENTIFICATION OF THE AREA OF INTEREST

SECTION PLACEMENT ON THE SLIDE

TRANSCRIPTOMIC PROFILING AND DATA ANALYSIS



GENE EXPRESSION PROFILING RETAINING SPATIAL CONTEXT INFORMATION

# Benchmark experiments

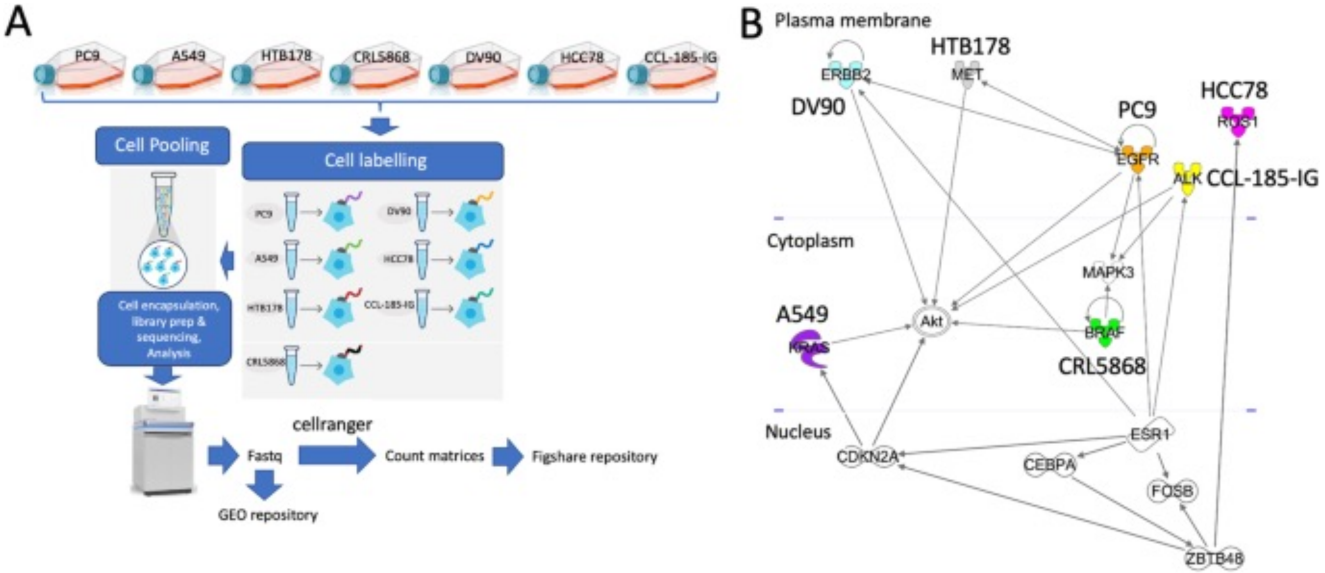
> Sci Data. 2024 Feb 2;11(1):159. doi: 10.1038/s41597-024-03002-y.

## A single cell RNAseq benchmark experiment embedding "controlled" cancer heterogeneity

Maddalena Arigoni # 1, Maria Luisa Ratto # 1, Federica Riccardo 1, Elisa Balmas 1, Lorenzo Calogero 2, Francesca Cordero 3, Marco Beccuti 3, Raffaele A Calogero 4, Luca Alessandri 1

M. Arigoni

ML. Ratto



- In preparation:
  - Multi-omics benchmark (scRNAseq, scATACseq)
  - Spatial transcriptomics benchmark (Visium 10XGenomics, Curio Bioscience)

# Top 5 Factors Affecting Reproducibility in Research

- Lack of Access to Raw Data and Methodologies
- Invalidated Biological Materials
- **Lack of Knowledge to Analyze Data**
- Incorrect Laboratory Practices
- Undervaluing Negative Results

https://www.enago.com/academy/top-5-factors-affecting-reproducibility-in-scientific-research/#5_Factors_Affecting_Reproducibility_in_Scientific_Research

# Hands-on data analysis courses for Life scientists

EMBL

DukeNUS
Medical School

FISV
FEDERAZIONE ITALIANA
SCIENZE DELLA VITA

Bits
Bioinformatics
ITalian
Society

IIB
INFRASTRUTTURA ITALIANA
DI BIOINFORMATICA

elixir
ITALY

Università di Torino
MBC
Molecular Biotechnology Center

# The side path to bioinformatics


A. Donofrio


S. Fasciolo

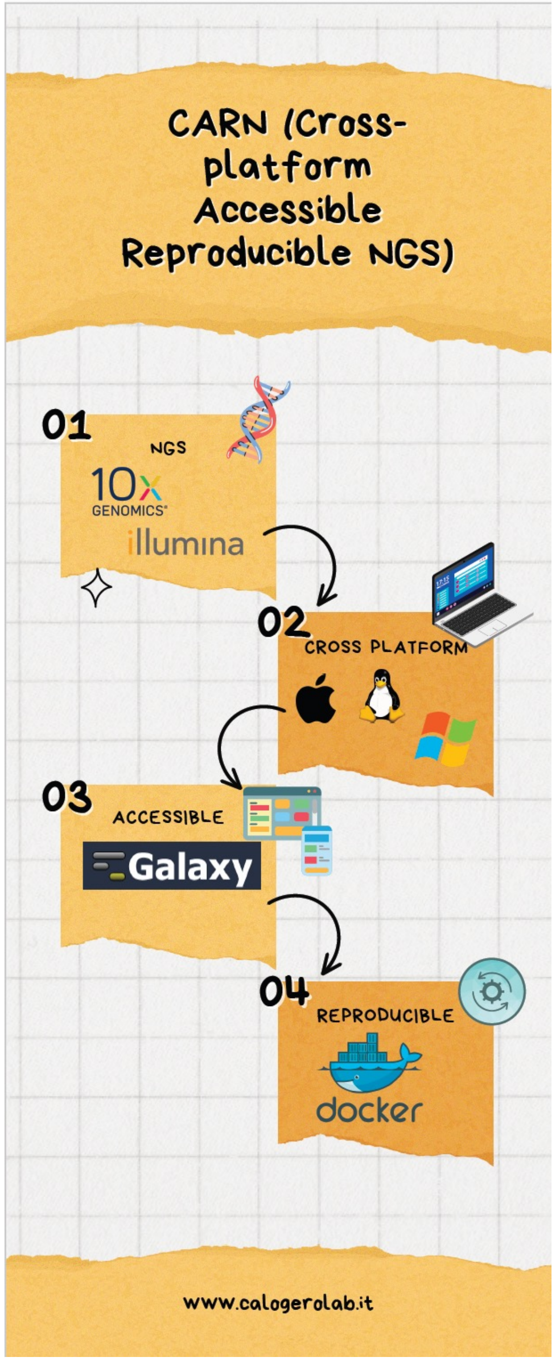

A. Lo Iacono


S. Bucatariu


J. Di Mauro


I. Castellano

- Becoming proficient in bioinformatics necessitates honing computing skills, a process that cannot be achieved within a short span of a few months.
- Mastery of scripting requires consistent practice over an extended period, typically spanning years.
- We provide ongoing support for students to develop the necessary expertise to pursue a career in bioinformatics starting from their first year in the Biotechnology course.


CARN (Cross-platform Accessible Reproducible NGS)

01 NGS

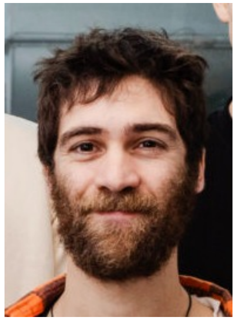
02 CROSS PLATFORM

03 ACCESSIBLE

04 REPRODUCIBLE

www.calogerolab.it

- Our goal is to make basic analyses accessible to any life scientist, thereby shortening analysis times.
- Life scientists will be able to perform initial analyses independently, freeing bioinformaticians to focus on more complex tasks.
- This approach not only democratizes omics analysis but also supports the principles of open science, ensuring that our software can be widely used, adapted, and built upon.
- Specifically, CARN provides a user-friendly interface via the Galaxy platform, CARN facilitates complex genomic analyses without necessitating command-line skills, integrating diverse NGS pipelines for a thorough analysis process from raw data to downstream results.
- Utilizing Docker containers, the project guarantees both functional and computational reproducibility, thus ensuring reliable outcomes.


L. Alessandri


A. Donofrio


S. Bucatariu

# What we offers for collaboration/service

- Experimental design support.
- Bulk RNAseq/miRNAseq library prep.
- Spatial transcriptomics (10Xgenomics/Curio Biosciences) from sections to data analysis.
- Bioinformatics support for data analysis (Bulk and single cell).

# On-going collaborations

Dipartimento di Biotecnologie Molecolari e Scienze per la Salute

2023 - 2027
DIPARTIMENTO DI ECCELLENZA
Ministero dell'Università e della Ricerca

UNIVERSITÀ DI TORINO

Università di Torino
MBC
Molecular Biotechnology Center

Molecular Biotechnology Center

Istituto Zooprofilattico Sperimentale
del Piemonte Liguria e Valle d'Aosta

Dipartimento di
**Medicina Veterinaria e Scienze Animali**

Dipartimento di
**Fisiopatologia Medico-Chirurgica e dei Trapianti**

Dipartimento di Eccellenza MUR
2023-2027

ISTITUTO DI CANDIOLO - IRCCS

HUMAN TECHNOPOLE

Computational biology
Iorio Group

IORIO LAB

Fondazione IRCCS Ca' Granda
Ospedale Maggiore Policlinico

Sistema Socio Sanitario
Regione Lombardia


Reproducible Bioinfor… Project
A project to provide reproducible results in Bio… Docker images
THE END
FINE
THE END
劇終
THE END
FINE
That's all Folks!
THE END
fine
The End
終り
THE END