Curriculum Vitae Prof. Raffaele A. Calogero

CONTACT INFORMATION

Name	Raffaele A. Calogero
Date and place of birth	March 3, 1960, Torino (Italy).
Department:	Dept of Molecular Biotechnology and Health Sciences, University of Torino
Address	Bioinformatics and Genomics Unit Molecular Biotechnology Center Via Nizza 52, Torino 10126
Tel	++39 0116706454
Fax	++39 0112366454
Email	raffaele.calogero@unito.it
www	www.reproducible-bioinformatics.org
Orcid	0000-0002-2848-628X
Scopus Author ID	7004035394
H-index scopus	43
Total citations (Scopus)	7069

EDUCATION

1984 Laurea in Biological Sciences summa cum laude from Naples University "Federico II" (Italy).

1984-1985 Post-degree training at Institute of Genetics, General and Molecular Biology (Naples University "Federico II").

1985-1988 Fellowship at Max Plank Institute fuer Molekulare Genetik, Berlin (Germany).

PROFESSIONAL EXPERIENCES

1989-1992 Researcher at SORIN Biomedica S.p.A. (I).

1992-1998 Associate Professor of Molecular Biology, at Naples University "Federico II".

1998-present Associate Professor of Molecular Biology at University of Torino.

CONSULTANCY ACTIVITIES

2004-2008 Consultant for the microarray program (AXXAM S.r.l, Milano, I)

2005-2009 Consultant for the microarray program (Nerviano Medical Sciences S.r.l, Nerviano, MI, I)

2010-2013 Consultant for Next generation sequencing application in QC (RBM SERONO, IVREA, I)

2016-2019: Consultant for Next generation sequencing application Galderma R&D, France.

2017 Consultant for Next generation sequencing application S. Raffaele Hospital (Milan,I).

OTHER ACTIVITIES

2002-2020 Part of the Faculty of the PhD Program in Complex Systems for Life Science, University of Torino.

2020-present Part of the Faculty of the PhD Program in Complex Systems for Quantitative Biomedicine, University of Torino.

2009-present Associated editor for BMC Bioinformatics.

2012-2019 Associated editor for PLOS ONE.

2017-present Cofounder of the Reproducible Bioinformatics Project (http://www.reproducible-bioinformatics.org/).

2017-present Member of the steering committee of the High-Performance Computing Center (c3s.unito.it), University of Torino.

2018-present Representative for University of Torino at the General Assembly of Elixir Italian Node.

2019-present Co-coordinator of Elixir IT single cell Omics community.

2021-present Member of the management committee of Elixir IT node.

2021-present Member of the scientific committee of CINI InfoLife lab.

2021-present President of the Italian Society of Bioinformatics (BITS)

INTERNATIONAL TRAINING COURSES

2002-2010 Introductory course on Microarray data analysis, supported by Affymetrix (University of Torino, Italy).

2013-2020 Introductory and Advance course on RNAseq data Analysis, sponsored by Illumina (University of

Torino, Italy).

2010-present Whole transcriptome data analysis course (EMBL – Heidelberg, Germany). Since 2016 the course doubled for high requests and it is run in spring and autumn.

2014 RNAseq data analysis course (Jagiellonian University Medical College, Krakow, Poland).

2015-present Analysis and Integration of Transcriptome and Proteome Data course (EMBL – Heidelberg, Germany)

2015-present Whole transcriptome data analysis course (DUKE-NUS – Singapore, Singapore)

2018 Whole transcriptome data analysis course (IIT – Genova, Italy)

2019-present Single-cell Omics EMBO course (EMBL – Heidelberg, Germany).

DIDATTICAL ACTIVITIES

1992-1998: Molecular Biology course for the degree in Biological Sciences at University of Naples "Federico II".

1996-2000: Molecular Biology laboratory course for the degree in Biological Sciences at University of Naples "Federico II".

1998-2000: Molecular Biology II course for the degree in Biotechnology at University of Naples "Federico II".

1998-present: Molecular Biology module within the integrated course of Biological, Molecular and Human Genetics bases for degree in Medicine and Surgery at University of Torino.

2003-2012: Bioinformatics course module B for the degree in Biomolecular Sciences at University of Torino.

2004-2005: Bioinformatics II course for the Master Program in Molecular Industrial Biotechnology at University of Verona.

2006-2009: Bioinformatics for the Master Program in Neurobiology and for the Master Program in Biomolecular Sciences at University of Torino.

2007-2012: Molecular Biology "special" course for the degree in Medicine and Surgery at University of Torino.

2016-2018: Image and Signal Analysis II for the Master Program in Molecular Biotechnology at University of Torino.

2016-2019: Molecular Biology module in the integrated course of Pharmacology and Molecular Biology for the Master Program in Medical Biotechnology at University of Torino.

2017-present: Data analysis for the Master Program in Molecular Biotechnology at University of Torino.

2018-present: Molecular Biology module for Master Program in Physics at University of Torino.

2019-present: Bioinformatics module in the integrated course of preclinical models and genomic methodologies for the Master Program in Medical Biotechnology at University of Torino.

2021-present: Informatics module in the integrated course of Physics and Informatics for the students of Biotechnology at University of Torino.

Master thesis 1995-present: 10 theses were supervised

PhD thesis 1995-present 9 PhD theses were supervised.

RESEARCH INTERESTS

My research experience focuses on the development, optimization of bioinformatics analysis workflows and on mining transcription-based experiments, mainly in the oncology framework. An important aspect of my past and present work is the constant interest on new technologies (e.g. today single-cell omics and spatial transcriptomics) and their integration in bioinformatics analysis workflows.

In 1992, I got the position as associate professor of Molecular Biology at University "Federico II" of Napoli. After few years, dedicated to finish the projects related to my previous position at Sorin Biomedica SPA (Saluggia, VC, Italy), I shifted my interest to the newborn technology of microarrays with a specific attention to data analysis.

In 1998, I moved to University of Torino, where my professional journey in Bioinformatics begun. I established the Genomics and Bioinformatics unit (B&Gu), which is an interdisciplinary group devoted to the study of multifactorial diseases by mean of high throughput technologies - i.e. microarray, Next Generation Sequencing – and bioinformatics.

From 1998 to 2007 the research activity of my group was focused on the identification of oncology biomarkers (Olivero et al. Nucl. Acid Res. 2003, Saviozzi et al. Methods Mol Med. 2004, Cavallo et al. BMC Bioinformatics 2005, Ceppi et al. Ann Oncol. 2006, Cavallo et al. Nat Rev Cancer. 2007) as part of the collaboration with the Immunology group headed by Prof. Forni first and subsequently by Prof. Cavallo.

From 2007 to 2010, my groups started to address the problem of characterizing the transcriptome of circulating extracellular vesicles, which represented the starting point of our interest in the use of circulating microvesicles as source of biomarkers (Bruno et al. J Am Soc Nephrol. 2009, Herrera et al. J Cell Mol Med. 2010). It is notable that, using gene expression microarrays, we were able to identify, for the first time, the presence of coding transcripts in circulating blood vesicles (Deregibus et al. Blood. 2007).

From 2010 to 2014, we started the development of bioinformatics tools devoted to the use of RNAseq to depict functional molecules and biomarkers from oncological specimens (Cordero et al. PLoS One. 2012, Cordero et al. Curr. Top Med. Chem. 2012, Carrara et al. Biomed Res. Int. 2013, Cordero et al. BMC Bioinformatics. 2013, Carrara et al. BMC

Bioinformatics. 2013). We also developed two protocols to generate total RNA data and miRNA data from circulating extracellular vesicles. These protocols were extensively used in the EU project NGS-PTL, in which we were involved as WP leader from 2013 to 2015.

From 2014 to 2020, we started a fruitful collaboration with Camargo' lab at Children Hospital (Harvard, Boston, USA). As part of this collaboration, we performed most of the bioinformatic analysis in various projects (Tremblay et al. Cancer Cell. 2014, Galli et al. Mol. Cell. 2015, Yuan et al. Nat. Commun. 2018, Maglic et al. EMBO J. 2018). In this period, we also started the development of tools for the analysis of single cell data (Alessandrì et al. GigaScience. 2019, Alessandrì et al. Methods Mol Biol. 2019, Ordoñez-Rueda et al. Cytometry. 2020). Exploiting these tools, we supported the Camargo lab in two high impact papers (Rodriguez-Fraticelli Nature. 2018, Christodoulou et al. Nature. 2020), in which single cell RNAseq data analysis was an important element. In this period, we also point our attention to the "reproducibility crisis" that was affecting various areas of biology. We founded the Reproducible Bioinformatics Project (reproducible-bioinformatics.org), which is a community of developers focusing on the production of reproducible bioinformatics workflows (Beccuti et al. Bioinformatics. 2018, Kulkarni et al. BMC Bioinformatics. 2018).

Since 2020, we continue working in the field of single cell Omics. Specifically, we are developing a new type of autoencoder, sparsely connected autoencoder, which is particularly suitable to extract biological information from single cell omics data (Alessandrì et al. NPJ Syst Biol Appl. 2021, Alessandrì et al. Methods Mol Biol. 2021, Alessandrì et al. Int J Mol Sci. 2021). Notably, this year, INTEL granted our project on the use of neuromorphic computing infrastructure to develop a more efficient implementation of the sparsely connected autoencoders. Furthermore, we extended our deep learning studies to the development of a tool for the detection of MET exon 14 skipping events in oncological samples (Nosi et al. Int J Mol Sci. 2021).

Since I started my research career as molecular biologist, I am convinced that genomics/transcriptomics data analysis, at least to some extent, should be placed back in the hands of wet biologists, thus it is more than twenty years that I am involved in training of PhD and Post-docs in the transcriptomics field, running courses at University of Naples first, then at University of Turin (Sponsored by Affimetrix first and later by lllumina), at EMBL-Heidelberg, DUKE-NUS in Singapore and Italian Institute of Technology in Genova. Thus, with my group, I have developed tools to simplify data analysis for microarray first (Sanges et al., 2007) and then for Next generation sequencing (Beccuti et al., 2017; Beccuti et al., 2018) and single cell RNAseq (Alessandrì et al. 2019).

Since the beginning of my carrier, I published 183 peer-reviewed papers. Since 2007, I deposited on GEO repository (www.ncbi.nlm.nih.gov/geo) a total of 927 samples as part of 84 different experiments involving the use of microarrays and RNAseq.

GRANTS

2022-2025	PNRR HPC Spoke 8, In Silico Medicine and Omics Data. Leader of the spoke for UNITO (594,000€)
2021	TCR data analysis of publicly available datasets. Research contract Matterhorn Biosciences (120,000€)
2012-2018	EPIGEN project, Bioinformatics work package. Research unit PI (200,000€).
2013-2015	Next Generation Sequencing platform for targeted Personalized Therapy of Leukemia (FP7-HEALTH-2012-INNOVATION-1). Responsible of work package (540,000 €).
2008-2010	Oncoantigens and microenvironment perturbation as a new antitumor vaccine strategy for inhibition of cancer (Italian Research program PRIN 2008). Research unit PI (64,000€).
2006-2009	Pipeline for Rapid Evaluation and Scoring of Targets in Oncology (Piedmont Region industrial research call CIPE 2006). Research unit PI (80,000€).
2006-2009	Innovative approaches to anticipate the diagnosis and to target therapies of colorectal cancer (AIRC Regione Piemonte 2006). Research unit PI (45,000€).
2006-2008	An integrated Bioinformatics and Genomics approach for oncoantigens identification (PRIN 2006). Research unit PI (65,000€).
2004-2006	Identification of new targets for the immunoprevention of cancer (Piedmont Region industrial research call CIPE 2004). Responsible of work package (80,000€).
2003-2005	Definition of gene expression profiles and prognostic predictors of response to endocrine therapy for breast cancer (Program for Health Science Italy 2003). Research unit PI (25,000€).
2002-2003	Isolation and functional characterization of p63 ligands involved in Hay-Wells Syndrome (AEC), (PRIN 2002) Research unit PI (35,000€).
2001-2002	Identification of p63 specific target genes by p63 isoforms-mediated transcriptional profiling (PRIN 2001). Research unit PI (52,000€).
2001-2003	Design and construction of expression vectors based on the translational reinitiation process (Progetto Finalizzato CNR 2001). Research unit PI (10,000€).
2001-2003	Improvement of the quality and quantity of production in horticulture and fruit: contribution of parthenocarpic fruit development and improvement of fertility, with genetic engineering techniques,

a variety cultivated tomato, eggplant, strawberry, raspberry, table grape, tangerine and lemon (FIRB 2001). Research unit PI (60,000€).

PATENT	TITLE	TYPE	YEAR	ROLE
TO2015000088.978	Molecole di acidi nucleici codificanti proteine chimeriche CSPG4 e relativi usi terapeutici	Italian/filed to Europe 2018	2015	coauthor
TO20100020	Costrutto di acido nucleico, vettore e vaccino a DNA includenti detto costrutto	Italian	2011	coauthor
TO20100028	Costrutto di acido nucleico, vettore e vaccino a DNA includenti detto costrutto	Italian	2011	coauthor
EP0571337 B1	Process to purify proteins from cell systems	European	1993	coauthor
EP0485347 A3	Recombinant hepatitis delta antigen, process for the purification and use thereof	European	1992	coauthor

HOBBIES

Kendo 5th Dan.
Kendo blogger (http://kendodream.blogspot.it/ ; http://www.youtube.com/user/kendomaniac)
Learning Japanese language: level beginner.

Prof. Raffaele A Calogero

Torino 09/02/2023