Luca Nanni, Ph.D.

Lausanne, 1003, Switzerland
Holder of Swiss work permit B

ORCID: 0000-0001-7021-2007
Website: lucananni93.github.io
Email: lucananni93@gmail.com
LinkedIn: linkedin.com/in/lucananni93
GitHub: github.com/lucananni93

About me

Experienced bioinformatician and data scientist with +7 years of expertise in designing, developing, and deploying robust algorithms and pipelines for analyzing next-generation sequencing data, clinical and biological, across multiple assays and technologies.

I have deep knowledge and expertise in cancer research, with a focus on understanding the molecular mechanisms of cancer development and progression, with a focus on genetic and epigenetic alterations. My work was published in high-impact journals and has contributed to advancing the field.

I have experience in industry-level genomic data analysis and interpretation, with a focus on translating complex biological questions into actionable insights. My work has involved close collaboration with cross-functional teams, including biologists, clinicians, and software engineers.

I have a proven track record in managing and analyzing complex multi-omics datasets (WGS, transcriptomics, epigenomics) at both bulk and single-cell levels, with a focus on oncology. I'm an expert in creating and implementing scalable, high-quality QC/QA protocols and transitioning novel analytical methods from R&D to commercial-grade, industry-level services.

I have a strong background in computer engineering, machine learning and statistics, with extensive experience in developing and applying advanced algorithms to solve complex biological problems.

I have experience in project and program management. I manage and lead cross-functional teams (+10 members) to deliver research results on time and within budget. I have experience in leading client-facing projects and ensuring high-quality deliverables.

Employment history

Program Manager, Genomics   Lausanne, Switzerland
New Biologix   11/2024–Current

- Promoted and extended my responsibilities to include program management, overseeing the development and execution of multiple genomics projects.

- Managed a portfolio of R&D and client projects, managing the entire project lifecycle from initiation to delivery, ensuring alignment with corporate goals and on-time, on-budget execution.

- Architect novel analytical methods for cell line and viral vector characterization using advanced technologies, including short- and long-read NGS and optical genome mapping.

- Managed the development of data pipelines and machine learning models to analyze multi-modal datasets (phenotype, genotype, transcriptomics, epigenomics), enhancing prediction of cell line performance.

R&D Director, Bioinformatics   Lausanne, Switzerland
New Biologix   11/2023–Current

- Developed novel sequencing and analytical protocols, transitioning them from R&D into both routine internal assays and commercial-grade services offered directly to clients.

- Engineered a novel viral genome quality control (QC) pipeline from scratch, integrating multi-modal data from long-read sequencing, mass photometry, and digital PCR (dPCR) to ensure vector integrity, purity, and safety.

- Developed deep learning models for annotation of genetic and epigenetic markers in engineered cell lines, enhancing the accuracy and scalability of cell characterization.

- Analyzed and interpreted large-scale genomics datasets (short- and long-read sequencing, optical mapping), correlating results with measured phenotypes
- Established rigorous standards for bioinformatics data reporting, quality control, and consistency to ensure industry-level data integrity.

Senior Bioinformatician
Lausanne, Switzerland
Alithea Genomics
09/2023–10/2023

- Design and optimization of transcriptomics data processing and analysis workflows
- Assisting in the design, benchmarking and testing of novel protocols
- Analysis, interpretation and reporting of research results and conversion into actionable insights
- Management of bioinformatics IT infrastructure, including cloud-based solutions

Postdoctoral Researcher in Computational Biology
Lausanne, Switzerland
Université de Lausanne, *Computational Systems Oncology Lab* of Prof. Ciriello
03/2021–08/2023

- Multi-omic data analysis of cancer genomes involving WGS, single and bulk RNA-seq, ChIP-seq, Cut&Tag, Hi-C and proteomics
- Design and implementation of novel bioinformatic pipelines for the analysis of Hi-C, ChIP-seq, Repli-seq and ATAC-seq data
- Identification of cancer genetic (bulk/single-cell RNA-seq) and epigenetic (ChIP-seq, Cut&Tag) drivers
- Chromatin conformation (bulk/single-cell Hi-C) of aneuploid cancer genomes
- Statistical modeling of DNA 3-D conformation from Hi-C data
- Variant calling, structural variants, and copy number analysis of cancer genomes from WGS data
- Unsupervised segmentation of Hi-C data based on epigenetic signals
- Management of students and young researchers, reporting to the head of the laboratory
- Grant writing, budgeting and submission

Data Scientist (part-time, remote)
Milan, Italy
Fondazione Human Technopole, *Health Data Science Center*
06/2023–08/2023

- Multi-omic approaches for the prediction of disease-genetic associations using machine learning

Research Assistant in Computational Biology
Lausanne, Switzerland
Université de Lausanne, *Computational Systems Oncology Lab* of Prof. Ciriello
11/2020–03/2021

- Design and implementation of novel bioinformatic pipelines for the analysis of Hi-C data
- Applying spatial models (HMMs, HMRFs) for the multi-omic analysis of Hi-C and ChIP-seq data

Intern in Computer Vision and AI
Bologna, Italy
Università di Bologna, *Computer Vision Lab* of Prof. Mattoccia
09/2014–07/2015

- Development of a deep learning and image analysis system for object recognition for the visually impaired

Education

Ph.D. in Computer Science and Engineering, *cum laude*
Milan, Italy
Politecnico di Milano, *Genomic Computing Lab* of Prof. Ceri
11/2017–11/2020

- Thesis: *Computational inference of DNA folding principles: from data management to machine learning*
- Modeling of chromatin conformation data (Hi□C) from epigenetic markers and transcription factors (ChIP□seq)
- Machine learning approaches for the analysis of cancer transcriptomic data (CNNs, Graph Neural Networks)
- Scalable software design and implementation for genomic data management and integration
- Genomics software deployment on cloud infrastructures

High Honors program   Milan/Turin, Italy
Alta Scuola Politecnica   10/2015–10/2017

- – Multidisciplinary school in Engineering, Architecture and Design
- – Project: *Synapse simulation and learning through a memristor network*

M.S. in Computer Science and Engineering, *cum laude*   Milan, Italy
Politecnico di Milano   10/2015–10/2017

- – Thesis: *A Python Data Analysis Library for Genomics and its Application to biology*

B.S. in Computer Engineering, *cum laude*   Bologna, Italy
University of Bologna   09/2012–07/2015

- – Thesis: *Convolutional neural networks for mobility aid for the visually impaired*

Languages

English	**French**	**Italian**
Fluent	Work-level	Native

Skills

- **Biology:** cancer genomics, genetics and epigenetics, chromatin conformation, single-cell genomics
- **Base calling and QC:** bcl2fastq, FastQC, MultiQC
- **Sequence alignment:** bwa, bowtie2, STAR, samtools
- **Genomic data analysis:** RNA-seq, Hi-C, scHi-C, WGS, scRNA-seq, ChIP-seq, Cut&Tag
- **Chromatin data analysis:** Juicer, Cooler, FitHic, ICE
- **Transcriptomic data analysis:** STAR, Salmon, Kallisto, DESEQ2, InferCNV, Cellranger
- **Bioinformatics workflows:** Nextflow, Snakemake
- **Variant and copy number calling:** mutect2, freebayes, ControlFREEC
- **Programming:** Python, R, C++, Scala, Java, C#
- **System administration:** Unix/Linux, Bash, Awk, shell scripting
- **Database management:** SQL
- **HPC:** Spark, PySpark, AWS, Google Cloud, Slurm, sbatch
- **DevOps:** conda, pip, Github, tox, testthat, Docker, Singularity
- **Data visualization:** matplotlib, seaborn, ggplot2, bokeh
- **Statistics and Machine Learning:** scikit-learn, pandas, numpy, scipy, Pytorch

Publications

[1] A. Iyer, M. Petrovic, D. Sesia, **L. Nanni**, M. Mina, and G. Ciriello, "Evolving patterns of co-mutations from tumor initiation to metastatic progression", *Nature Genetics*, Under revision.

[2] P. Li, A. Hatos, M. Petrovic, G. M. Franceschini, **L. Nanni**, D. Tavernari, and G. Ciriello, "Tadshop: Systematic benchmarking and identification of topologically associating domains.", *Nature Methods*, Under revision.

[3] R. A. Lambuta, **L. Nanni** (**co-first**), Y. Liu, J. Diaz-Miyar, A. Iyer, D. Tavernari, N. Katanayeva, G. Ciriello, and E. Oricchio, "Whole-genome doubling drives oncogenic loss of chromatin segregation", *Nature*, vol. 615, no. 7954, pp. 925–933, 2023.

[4] J. Schreiber *et al.*, "The encode imputation challenge: A critical assessment of methods for cross-cell type imputation of epigenomic profiles", *Genome Biology*, vol. 24, no. 1, pp. 1–22, 2023.

[5]   **L. Nanni**, "Computational inference of dna folding principles: From data management to machine learning", in *Special Topics in Information Technology*, L. Piroddi, Ed., Cham: Springer International Publishing, 2022, pp. 79–88, isbn: 978-3-030-85918-3. doi: 10.1007/978-3-030-85918-3_7.

[6]   A. Canakoglu, P. Pinoli, A. Gulino, **L. Nanni**, M. Masseroli, and S. Ceri, "Federated sharing and processing of genomic datasets for tertiary data analysis", *Briefings in Bioinformatics*, vol. 22, no. 3, bbaa091, 2021.

[7]   Y. Liu, **L. Nanni**, S. Sungalee, M. Zufferey, D. Tavernari, M. Mina, S. Ceri, E. Oricchio, and G. Ciriello, "Systematic inference and comparison of multi-scale chromatin sub-compartments connects spatial organization to cell phenotypes", *Nature communications*, vol. 12, no. 1, pp. 1–11, 2021.

[8]   A. Rouhi, **L. Nanni**, A. Canakoglu, P. Pinoli, and S. Ceri, "Ensemble feature selection for single cell chromatin conformation analysis", in *2021 13th International Conference on Bioinformatics and Biomedical Technology*, New York, NY, USA: Association for Computing Machinery, 2021, pp. 209–216, isbn: 9781450389655.

[9]   F. Cristovao, S. Cascianelli, A. Canakoglu, M. Carman, **L. Nanni**, P. Pinoli, and M. Masseroli, "Investigating deep learning based breast cancer subtyping using pan-cancer and multi-omic data", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2020.

[10] **L. Nanni**, S. Ceri, and C. Logie, "Spatial patterns of ctcf sites define the anatomy of tads and their boundaries", *Genome biology*, vol. 21, no. 1, pp. 1–25, 2020.

[11] M. Varrone, **L. Nanni** (**co-first**), G. Ciriello, and S. Ceri, "Exploring chromatin conformation and gene co-expression through graph embedding", *Bioinformatics*, vol. 36, no. Supplement_2, pp. i700–i708, 2020.

[12] S. Cascianelli, F. Cristovao, A. Canakoglu, M. Carman, **L. Nanni**, P. Pinoli, and M. Masseroli, "Evaluating deep semi-supervised learning for whole-transcriptome breast cancer subtyping", in *International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics*, Springer, 2019, pp. 232–244.

[13] M. Masseroli, A. Canakoglu, P. Pinoli, A. Kaitoua, A. Gulino, O. Horlova, **L. Nanni**, A. Bernasconi, S. Perna, E. Stamoulakatou, *et al.*, "Processing of big heterogeneous genomic datasets for tertiary analysis of next generation sequencing data", *Bioinformatics*, vol. 35, no. 5, pp. 729–736, 2019.

[14] **L. Nanni**, P. Pinoli, A. Canakoglu, and S. Ceri, "Pygmql: Scalable data extraction and analysis for heterogeneous genomic datasets", *BMC bioinformatics*, vol. 20, no. 1, pp. 1–11, 2019.

[15] P. Pinoli, S. Ceri, D. Martinenghi, and **L. Nanni**, "Metadata management for scientific databases", *Information Systems*, vol. 81, pp. 1–20, 2019.

[16] C. Wang, **L. Nanni**, B. Novakovic, W. Megchelenbrink, T. Kuznetsova, H. G. Stunnenberg, S. Ceri, and C. Logie, "Extensive epigenomic integration of the glucocorticoid response in primary human monocytes and in vitro derived macrophages", *Scientific reports*, vol. 9, no. 1, pp. 1–17, 2019.

[17] A. Canakoglu, **L. Nanni** (**co-first**), A. Sokolovsky, and S. Ceri, "Designing and evaluating deep learning models for cancer detection on gene expression data", in *International meeting on computational intelligence methods for bioinformatics and biostatistics*, Springer, 2018, pp. 249–261.

[18] S. Ceri, A. Canakoglu, A. Gulino, A. Kaitoua, M. Masseroli, **L. Nanni**, and P. Pinoli, "Demonstration of genometric query language", in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 2018, pp. 1887–1890.

[19] **L. Nanni**, P. Pinoli, A. Canakoglu, and S. Ceri, "Exploring genomic datasets: From batch to interactive and back", in *Proceedings of the 5th International Workshop on Exploratory Search in Databases and the Web*, 2018, pp. 1–6.

[20] S. Ceri, A. Bernasconi, A. Canakoglu, A. Gulino, A. Kaitoua, M. Masseroli, **L. Nanni**, and P. Pinoli, "Overview of geco: A project for exploring and integrating signals from the genome", in *International Conference on Data Analytics and Management in Data Intensive Domains*, Springer, 2017, pp. 46–57.

[21] M. Poggi, **L. Nanni**, and S. Mattoccia, "Crosswalk recognition through point-cloud processing and deep-learning suited to a wearable mobility aid for the visually impaired", in *International Conference on Image Analysis and Processing*, Springer, 2015, pp. 282–289.

Presentations

- Selected talk at the America Society of Gene and Cell Therapy (ASGCT) Annual Meeting   05/2025
- Talk at 18th Lausanne CompBio meeting   03/2025

- Selected talk at European Conference in Computational Biology (ECCB 2020)   09/2020
- Selected talk at SIGMOD PODS 2018   06/2018
- Invited talk at Broad Institute, Cambridge, Massachusetts (USA)   01/2018

Other scientific activities

- I was part of the organization committee of the **1st International Workshop on Web Applications for Life Sciences**, held on July 8th, 2022, in conjunction with the 22nd International Conference on Web Engineering (ICWE 2022)

- I am a **paper reviewer** for the following journals: *Bioinformatics*, *Cell Reports*, *Briefings in Bioinformatics*, *BMC Bioinformatics*, *PLOS Computational Biology*

Teaching

- **Teaching Assistant** at Université de Lausanne   A.Y. 2021, 2022
  Computational Thinking and Programming to Solve Biological Problems
- **Teaching Assistant** at Politecnico di Milano   A.Y. 2018
  Informatica B

Supervision

- Ph.D thesis of Karol Piera, *Deciphering the relationship between chromatin compartmentalization and accessibility*, Université de Lausanne   03/2023–Current
- Ph.D thesis of Pumin Li, *Benchmarking bioinformatics approaches for the study of topologically associating domains using Hi-C data*, Université de Lausanne   03/2023–Current
- M.S. thesis of Francesco Guzzo, *Imputation of biochemical activity associated with functional elements of the genome produced by epigenomic experiments*, Politecnico di Milano   04/2019–02/2020
- M.S. thesis of Amirreza Rouhi, *A machine learning-based method for feature selection in single-cell Hi-C contact maps*, Politecnico di Milano   03/2019–04/2020
- M.S. thesis of Marco Varrone, *Integration of biological knowledge in gene expression programs learning*, Politecnico di Milano   02/2019–02/2020
- M.S. thesis of Guido Borrelli, *Classification of genetic profiles using Transfer Learning*, Politecnico di Milano   03/2018–10/2018
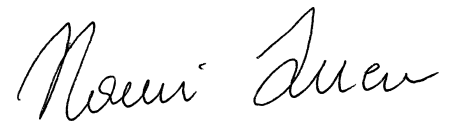
Awarded Grants

- **Amazon Machine Learning Research Award** (PI: Stefano Ceri)   2019-2021
  Data-Driven Machine and Deep Learning for Genomics

  I co-wrote a research grant to design and implement an integrated software environment for data-driven genomics research based on modern machine learning technologies. We were awarded 100.000 USD for cloud computing expenses and 50.000 USD for research expenses.

Scholarships and Awards

- Italian Government Scholarship for Ph.D. students   11/2017–01/2021
- High performing student award at University of Bologna   2014, 2015

Personal development

- Member of Yezers 2019–Current

  I work in the Education department of Yezers. Yezers is an italian startup dealing with important italian economic and social issues. Our objective is to bring novel and exciting proposals directly to the government desk. I am specifically involved in the design of novel educational strategies for the new generations to account for the new challenges our society is facing nowadays

- Music composition and execution 2008–Current

  I play the piano proficiently and I compose music. I studied at the Music Academy of Bologna for 3 years, specializing in Electronic Music Composition

- Fencing athlete 2000–2010

  I practiced fencing for almost 10 years, gaining the Italian champion title in 2005 in my age category

Luca Nanni