POLITECNICO DI MILANO

PhD in DATA ANALYTICS AND DECISION SCIENCES - 41st cycle

THEMATIC Research Field: DEVELOPMENT OF MACHINE LEARNING BASED METHODS FOR GENOTOXICITY DETECTION IN GENE THERAPY PATIENTS

Monthly net income of PhDscholarship (max 36 months)
1400.0
In case of a change of the welfare rates during the three-year period, the amount could be modified.

Context of the research activity	
Motivation and objectives of the research in this field	In hematopoetic stem cell gene therapy, integrating viral vectors insert therapeutic transgenes into semi-random genomic sites to correct genetic defects. While these vector have shown clinical potential, they also carry a risk of genotoxicity, through proto-oncogene activation or tumor suppressor gene inactivation. In fact, several clinical trials have reported adverse events such as leukemia caused by insertional mutagenesis. Traditional genotoxicity assessments rely on vector integration site (IS) analysis, focusing on the identification of common insertion sites and/or clonal dominance, as potential indicators of selective advantage. However these methods are often retrospective, lacking the sensitivity and predictive power. To enhance genotoxicity detection there is a growing need for innovative, data-driven approaches able to integrate multiple dimensions of insertional data into a more comprehensive framework. To address this need, we aim to develop and validate three novel readouts of genotoxicity: Abnormal Exon Targeting Frequency, Percentile-ranked integration Site (IS) Abundance, and Integration Orientation Bias. We will use these metrics to train *machine learning* models capable to classify each vector insertions as potentally genotoxic or benign, offering a more accurate and predictive tool for gene therapy safety.
Methods and techniques that will be	- Bioinformatics computational methods in genomics (for

developed and used to carry out the research	- Bioinformatics computational methods in genomics (for alignment, data filtering, data integration, etc.). - High performance computing (HPC) in distributed environments for large scale and big data computing. - Data collection and query in databases, integrating multiple public datasets. - Statistical comparisons of data (through generalized mixed model effects and common ANOVA tests). - Time series analysis.- Machine learning methods.
Educational objectives	We are aimed at developing specific competence, autonomy, research methodology, and skills in an interdisciplinary environment. Specific focus will be on data analysis with multidimensional data (genomics signatures and features) and machine learning modelling of genomics data.
Job opportunities	Academic positions in Bioinformatics, or tenured tracks in research centers. Alternatively, job placement in industries or private corporations in the field of Pharmaceutics, Genomics, and Health Data Science.
Composition of the research group	1 Full Professors 3 Associated Professors 1 Assistant Professors 8 PhD Students
Name of the research directors	Marco Masseroli; Eugenio Montini

Contacts
Marco Masseroli*, Associate ProfessorDEIB - Politecnico di Milano*   tel. +39 02 2399 3553   email marco.masseroli@polimi.it   https://www.deib.polimi.it/eng/people/details/266220   Eugenio Montini, *Group Leader at San Raffaele Telethon Institute for Gene Therapy, Milan, Italy*   tel. +39 02 2643 9301   email montini.eugenio@hsr.it   https://research.hsr.it/en/institutes/san-raffaele-telethon-institute-for-gene-therapy/safety-of-gene-therapy-and-insertional-mutagenesis.html

Additional support - Financial aid per PhD student per year (gross amount)	
Housing - Foreign Students	--
Housing - Out-of-town residents	--

Scholarship Increase for a period abroad	
Amount monthly	700.0 €
By number of months	6

Additional information: educational activity, teaching assistantship, computer availability, desk availability, any other information

Educational activities (purchase of study books and material, funding for participation in courses, summer schools, workshops and conferences):

financial aid per PhD student per year:

1st year: max 1.902,38 euro per student

2nd year: max 1.902,38 euro per student

3rd year: max 1.902,38 euro per student

Teaching and lab assistantship: there are various forms of financial aid for activities of support to the teaching practice.The PhD student is encouraged to take part in these activities, within the limits allowed by the regulations.

The PhD student is encouraged to take part in these activities, within the limits allowed by the regulations.

Computer availability:

1st year:individual use

2nd year: individual use

3rd year: individual use

Desk availability:

1st year:individual use

2nd year: individual use

3rd year: individual use