UNIVERSITÀ DEGLI STUDI DI MILANO

TO MAGNIFICO RETTORE OF UNIVERSITA' DEGLI STUDI DI MILANO   ID CODE __5704_____

I the undersigned asks to participate in the public selection, for qualifications and examinations, for the awarding of a type B fellowship at **Dipartimento di ___Scienze Agrarie e Ambientali - Produzione, Territorio, Agroenergia_____**

Scientist- in – charge: **_____Professor Laura Rossini_____**

[James Friel]

CURRICULUM VITAE

PERSONAL INFORMATION

Surname	Friel
Name	James

PRESENT OCCUPATION

Appointment	Structure
PhD student	Department of Bioscience, University of Milan

EDUCATION AND TRAINING

Degree	Course of studies	University	year of achievement of the degree
PhD	Molecular and Cellular Biology XXXV CYCLE	University of Milan	Expected to defend in May 2023
Master	Agricultural and Biological Sciences	National University of Ireland, Galway	2018
Degree	Higher Bachelor of Science (Finished with 1st class Honours)	National University of Ireland, Galway	2017

FOREIGN LANGUAGES

Languages	level of knowledge
English	Mother tongue
Italian	A1

![Università degli Studi di Milano logo] **UNIVERSITÀ DEGLI STUDI DI MILANO**

AWARDS, ACKNOWLEDGEMENTS, SCHOLARSHIPS

Year	Description of award
2019-2022	PhD research was fully funded by the project "Department of Excellence" awarded to the Department of Biosciences by the Italian Ministry of Research.
2022	Cornell University Triad Track 1: Foundational Research ($50,000 award) Exploring Diversity and Cold Adaptation in America's Forgotten Fruit, *Asimina triloba* (Common Pawpaw)

PROJECT ACTIVITY

Year	Project
2022	Exploring Diversity and Cold Adaptation in America's Forgotten Fruit, *Asimina triloba* (Common Pawpaw) Goals of the project: **Goal 1:** Develop and improve existing *A. triloba* genomic resources. **My role:** I was responsible for the genome assembly and annotation **Goal 2:** Sample genetic diversity of *A. triloba* populations from the extremes and midpoint of its native range to identify signatures of local adaptation. **My role:** I established a network of professional and mature botanists to collect samples from across the entire length and breadth of the species native range. I then prepared the GBS sequencing libraries and analyses the data.

TRAINING OR RESEARCH ACTIVITY

Oct 2019 – Current University of Milan, Milan, Italy (3 years)

Position: PhD Student

The goal of my PhD research project was to understand the factors influencing the evolutionary history, species distribution, and genetic variation in the species *Asimina triloba*.

The first steps in project involved the generation of a draft genome assembly using PacBio Sequel II long-reads and Illumina short-reads to identify genetic variants using genotype-by-sequencing (GBS) data. These variants were used to analyse the geographic features potentially affecting the genetic variation and geneflow in *A. triloba* populations of Virginia, USA.

This was followed by the generation of an improved reference genome and analysis of genetic diversity of the species across its entire native range to examine the factors influencing the species' migration into cooler climates. For the improved assembly we used the latest Hi-Fi sequencing technology along with Hi-C chromosome confirmation, and RNA-sequencing data to produce an annotated and chromosome-level genome assembly. The evolutionary history of the species was explored by calling variants on the improved genome using GBS data of samples from across 22 states in North America.

Overall, to study the genetic diversity of the species I needed to develop a combination of population genetics, deep biological knowledge of the species, statistical skills, and bioinformatic skills.

Skills acquired:

2

Università degli Studi di Milano – Direzione Trattamenti Economici e Lavoro Autonomo
Ufficio Contratti di formazione e Ricerca
Via Sant'Antonio 12 - 20122 Milano, Italia
assegni.ricerca@unimi.it DTELA_M_CVAssegniENG_rev. 00 del 02/09/2021

- **Genetic diversity and population structure analysis;** using both Whole Genome Sequencing and reduced representation methods (GBS)
- **Citizen Science and outreach**; Massive community recruitment to collect samples from across the US for use in my project
- **Time and resource management**

- **Public speaking**
- **Manuscript writing**

Bioinformatics skills acquired:

- **Data analysis;** using Linux, R, Python, Bash, and Perl.
- **Genotyping-by-sequencing;** experience in both the library preparation and data processing.
- **Genome assembly;** using both long and short read sequencing technologies, including the latest Hi-Fi sequencing technology from PacBio. Additional training in genome improvement using Hi-C chromosome contact map information to build pseudomolecules.
- **Genome annotation;** experience in training and use of multiple gene model predication tools
- **Gene expression analysis** using RNA-sequencing data.
- **RNA-seq data analysis** to generate transcripts and expressed sequence tags (EST) used in training gene prediction models.

Programming experience acquired in:

- **Linux**
- **Python**
- **Perl**
- **Bash**
- **R**

May 2022 – Jan 2023 IBMCP (UPV-CSIC), Valencia Spain (6 months)

Position: Visiting PhD student

Supervised by: Prof. Aureliano Bombarely

During my time in Valencia, I focused on genome assembly quality control and annotation of de novo assemblies. To ensure a high-quality genome assembly I utilized my expertise in Linux and Python scripting to evaluate various genome quality control metrics, including overall completeness, contamination, and sequence accuracy. By meticulously analysing these metrics, I was able to identify and address potential issues in genome assemblies and improve their overall quality. In addition to quality control, I also learned about gene prediction algorithms and how to utilise them to annotate assemblies using a combination of ab initio and evidence-based approaches. This involved developing an understanding of gene prediction models and proficiency in bioinformatic pipelines and tools such as MAKER and BRAKER to accurately annotation of key genomic features.

Bioinformatics skills acquired:

- **Genome quality controls** using assembly fragmentation assessment, k-mer distributions, gene space completeness, and repetitive element content.
- **Genome visualisation** with Juicebox pipeline

- **Genome annotation** using BRAKER and MAKER pipelines integrating multiple evidence sources (RNA transcripts, Expressed Sequences Tags, and homologous proteins)
- **Gene prediction** by training prediction algorithms
- **Bash scripting** Linux commands to improve and automate the demultiplexing and processing of raw GBS data.
- **Working with Hi-C data** to generate genetic contact map to assemble pseudomolecules and produce to highly contiguous chromosome-level genome assemblies.

Programming experience acquired in:

- **Linux**
- **Python**
- **Perl**
- **Bash**
- **R**

19th July 2022 – Manual genome curation – online

1-day workshop organised by the Vertebrate Genomes Project (VGP)

Bioinformatic skills acquired:

- **Manual genome curation** using genetic contact map information and the PretextView program.

Programming experience acquired in:

- **Linux**

Oct 2021 – Mar 2022 Cornell University, NY USA (6 months)

Position: Visiting PhD student

Supervised by: Dr Suzy Strickler

In the bioinformatics department of Cornell University, I received advanced training in de novo genome assembly using PacBio's High Fidelity (Hi-Fi) long-read sequencing technology. This involved mastering complex data analysis techniques, including genome assembly algorithms and quality control methods.

In addition, I was also responsible for conducting RNA extractions and transcriptomic analysis to generate evidence required to train gene prediction models for the annotation of a reference genome. This task required me to employ a range of analytical and technical skills, including proficiency in bioinformatics tools and programming languages such as Python and R. By contributing to the development of these gene prediction models, I was able to improve the accuracy of genomic annotations, paving the way for further research in the field.

Bioinformatics skills acquired:

- **Genome assembly** using the latest Hi-Fi long read sequencing technology from PacBio.
- **Gene expression analysis** using RNA-sequencing data.
- **RNA-seq data analysis** to generate transcripts and expressed sequence tags (EST) used in training gene prediction models.

Programming experience acquired in:

- **Linux**
- **Python**
- **Perl**
- **Bash**
- **R**

13ᵗʰ-16ᵗʰ Sept 2021 - K-mer workshop – online (1 ECT)

Organised by ForBio research school in biosystematics, University of Oslo, Norway

This workshop focused on the use of k-mers in genome sequencing projects in non-model species to better evaluate and improve genome sequencing projects

Bioinformatics skills acquired:

- Use of K-mers for:
    - **K-mer spectra analysis**
    - **Separation of sub-genomes**
    - **Analysing skimming data**

Programming experience acquired in:

- **Linux**
- **Bash**
- **R**

24ᵗʰ-26ᵗʰ Jun 2019 - GWAS workshop Virginia Tech, VA, USA (3 days)

Organised by Malachy Campbell and Gota Morota of Virginia Tech

This 3-day workshop provided me with the background, specifically in the use of Linear mixed models, to be able to successfully analysis GWAS data and associate genetic loci with qualitative traits even in long term field trials with multiple uncontrollable factors.

Bioinformatics skills acquired:

- **Single marker regression GWAS**
- **Bayesian whole-genome regression GWAS**
- **Genetic analysis of image-based high-throughput phenotyping data**

Programming experience acquired in:

- **Linux**
- **Bash**
- **R**

Feb 2018 – Jul 2019 Virginia Tech, VA, USA (5 months)

Position: Research Assistant

Supervised by: Prof. Song Li

In my previous role as a research assistant at VT, I conducted an analysis of Genome-Wide Association Study (GWAS) data from a 2-year Soybean field trial. My focus was in identifying Single Nucleotide Polymorphisms (SNPs) associated with the accumulation of protein and oil in seeds. To accomplish this, I learned about the management advanced data analysis techniques, including machine learning algorithms and statistical models, to identify relevant SNPs and develop a deeper understanding of their biological significance.

In addition, I also assisted in optimizing crop trial monitoring by utilizing LiDAR mounted on Unmanned Aerial Vehicles (UAVs). This approach allowed for precise measurements of plant height, canopy cover, and biomass accumulation, which provided valuable insights into crop performance and aided in the development of effective field management strategies.

Bioinformatics skills acquired:

- **GWAS data analysis** I learned to use the Genome Association and Prediction Integrated Tool (GAPIT), which was run in R, to analysis large SNP data sets, comparing multiple traits,

environments, and models to associate specific loci with phenotypic traits.

- **Large data management** using hapmap and PLINK formats, two file formats commonly used to storage and large amounts of genotypic data.

Programming experience acquired in:

- **Linux**
- **Python**
- **Perl**
- **R**
- **GAPIT (Genome Association and Prediction Integrated Tool)**

Other skills

- **UAV/Drone piloting**

Nov - Dec 2018 Virginia Tech, VA USA (2 months)

Position: Research Assistant

Supervised by: Aureliano Bombarely

Activity: As a research assistant, I successfully prepared numerous genotype-by-sequencing (GBS) libraries, preparing of over 700 samples for sequencing. My responsibilities included all steps from DNA extraction, through fragmentation, barcoding, and fragment size selection to quality controls.

skills acquired:

- **GBS library preparation**
- **DNA fragment size selection**
- **Library quality control** using both Bioanalyzer and Tapestation

---

Nov 2017 – May 2018 Virginia Tech, VA USA. (6 months)

Position: Visiting Masters student

Supervised by: Aureliano Bombarely

Training: During my master's program, I conducted research on Arabidopsis and focused on studying the triploid block. In this project, I was responsible for performing various laboratory tasks, including RNA extraction, preparing sequencing libraries, and conducting transcriptomic analysis. Additionally, I received specialized training in bioinformatics in Linux and R, which proved invaluable in analysing the vast amount of data generated by the transcriptomic analysis. This expertise enabled me to extract meaningful insights from complex data sets and accurately interpret the results of my experiments.

- **RNA-sequencing preparation**
- **Transcriptomic analysis**
- **Data visualisation**

Programming experience acquired in:

- **Linux**
- **Bash**
- **R**

UNIVERSITÀ DEGLI STUDI DI MILANO

CONGRESSES AND SEMINARS

Date	Title	Place
07/10/2022	**PhD school workshop: Oral Presentation** "Understanding the genetic diversity of potential new crop *Asimina triloba* (Pawpaw)"	University of Milan, Milan, IT
28/03/2022	**Invited Speaker: Oral Presentation** "Pawpaw, North America's tropical tasting mystery – A genetic diversity study of *Asimina triloba*"	Cornell university, Boyce Thompson Institute, NY, USA
26/10/2021	**Invited Speaker: Oral Presentation** "Comparative Analysis of Genotyping by Sequencing and Whole-Genome Sequencing Methods in Diversity Studies of *Olea europaea L.*"	Cornell university, BCBC department, NY, USA
09/10/2021	**PhD school workshop: Poster Presentation** "Comparative Analysis of Genotyping by Sequencing and Whole-Genome Sequencing Methods in Diversity Studies of *Olea europaea L.*"	University of Milan, Milan, IT

PUBLICATIONS

Articles in reviews
"Comparative Analysis of Genotyping by Sequencing and Whole-Genome Sequencing Methods in Diversity Studies of *Olea europaea L.*" **Friel, J.**, Bombarely, A., Fornell, C,D., Luque, F, and Fernández Ocaña, A,M., Plants, 2021, 10(11), 2514 **My Contribution:** Joint development of the methodology with other co-authors All the formal analysis of the WGS and GBS data in the study All preparation of original draft Joint reviewing and editing with co-authors
"Gene dosage compensation of rRNA transcript levels in *Arabidopsis thaliana* lines with reduced ribosomal gene copy number." Lopez, F,B., Fort, A., Tadini, L., Probst, A,V., McHale, M., **Friel, J.**, Ryder, P., Pontvianne, F., Pesaresi, P., Sulpice, R., McKeown, P., Brychkova, G., and Spillane, C. The Plant Cell, 2021, Volume 33, Issue 4, Pages 1135–1150 **My Contribution:** Performing CRISPR treatments and phenotyping of T3 and T4 generations Development of phenotyping methodologies
"Genetic insights into the modification of the pre-fertilization mechanisms during plant domestication." Manrique, S., **Friel, J.**, Gramazio, P., Hasing, T., Ezquer, I. and Bombarely, A. Journal of Experimental Botany, 2019., 70(11), pp.3007-3019. **My Contribution:** Joint preparation of original manuscript, reviewing and editing with co-authors
"Crop Height and Plot Estimation from Unmanned Aerial Vehicles using 3D LiDAR." Dhami, H., Yu, K., Xu,T., Zhu, Q., Dhakal, K., **Friel, J.**, Li, S., and Tokekar, P. Computing Research Repository (CoRR), 2019, Volume 1910.14031. **My Contribution:** Drone piloting

Manual data collection of crop height for comparison with LiDAR results

OTHER INFORMATION

Bioinformatic Skills:
De novo genome assembly
De novo genome annotation
RNA-seq data analysis
GBS data processing and analysis
WGS data processing and analysis
GWAS data analysis
Population genetics
Programming Languages:
Linux
R (Bioconductor)
Python
Perl
Society memberships:
Society for Molecular Biology and Evolution
Società Italiana di Biologia Evoluzionistica

Declarations given in the present curriculum must be considered released according to art. 46 and 47 of DPR n. 445/2000.

The present curriculum does not contain confidential and legal information according to art. 4, paragraph 1, points d) and e) of D.Lgs. 30.06.2003 n. 196.

Please note that CV WILL BE PUBLISHED on the University website and It is recommended that personal and sensitive data should not be included. This template is realized to satisfy the need of publication without personal and sensitive data.

Please DO NOT SIGN this form.

Place and date: ___Milano, Lombardia____, ____24/03/2023_____