

## OFERTA CONTRATO PREDOCTORAL:

PRE2023 - Desarrollo de Herramientas para Anotación de Genomas en Especies Cultivadas - PID2022-142508NB-I00

El contrato predoctoral estará asociado al proyecto titulado “Desarrollo de un Paquete de Herramientas de Análisis de *K*-mers para la Evaluación y Mejora de la Anotación de Genomas de Especies Cultivadas (KATULU)”. Este proyecto tiene como objetivo el desarrollo de herramientas bioinformáticas para mejorar la anotación de los genomas de las especies cultivadas usando metodologías de Aprendizaje Profundo y Teoría de la Información. La anotación de un genoma es el proceso por el cual se identifican las regiones de un genoma asociadas a propiedades como codificación de genes y elementos transponibles (anotación estructural) y se asocian funciones a las mismas (anotación funcional). Aunque existen varias herramientas para realizar anotaciones, casi no existen herramientas dirigidas a evaluar la calidad de estas. El proyecto se divide en dos partes. En la primera se estudiará las propiedades de *K*-mers en relación con distintos parámetros de calidad de las anotaciones estructurales de los genomas para desarrollar mejores herramientas de evaluación. La segunda parte está centrada en la integración de distintas capas de información a fin de mejorar la anotación funcional de los modelos génicos. El investigador predoctoral se formará en el uso de herramientas bioinformáticas de anotación de genomas, programación, inteligencia artificial y teoría de la información. También deberá desarrollar habilidades de interpretación crítica de los resultados y escritura de artículos científicos.

### Habilidades y Conocimientos Recomendados:

- Conocimientos básicos de biología, genética, y matemáticas.
- Conocimientos básicos de un lenguaje de programación (ej., Perl, Python).
- Conocimientos básicos de líneas de comandos en Linux.
- Conocimientos básicos de estadística.
- Buen nivel de inglés.
- Capacidad crítica.

### El Grupo de Bioinformática y Genómica Evolutiva:

El grupo de Bioinformática y Genómica Evolutiva es un equipo científico interdisciplinario localizado físicamente en el IBMCP (<https://ibmcp.upv.es/>) dirigido por Aureliano Bombarely (ver Anexo para más detalles). Aunque nuestros intereses se encuentran ampliamente distribuidos en los campos científicos de bioinformática, genómica, evolución, mejora y sistemática, estamos enfocados en dos actividades principales: 1- Desarrollo de herramientas bioinformáticas para procesar, administrar y analizar información genómica. Buenos ejemplos de estas herramientas son: Organelle\_PBA, una herramienta para ensamblado de genomas de cloroplastos y mitocondrias usando lecturas largas; o DeepTE, una herramienta para clasificar transposones usando Convolutional Neural Network (CNN); 2- Estudio de la evolución de los genomas durante los procesos de adaptación, domesticación y especiación. En este sentido tenemos varias líneas de investigación tales como la caracterización de cambios genómicos durante los procesos de diversificación en Solanáceas, elucidación de los mecanismos genéticos detrás de la producción de alelos nuevos durante la domesticación de la especie *Sinningia speciosa* y búsqueda de patrones de selección asociados a la adaptación a climas templados y fríos en poblaciones de *Asimina triloba*.

Contacto: Aureliano Bombarely, [abombarely@ibmcp.upv.es](mailto:abombarely@ibmcp.upv.es)

**CURRICULUM VITAE ABREVIADO (CVA)**

**IMPORTANT** – The Curriculum Vitae cannot exceed 4 pages. Instructions to fill this document are available in the website.

**Part A. PERSONAL INFORMATION**

First name	Aureliano		
Family name	Bombarely Gomez		
Gender (*)	Male	Birth date (dd/mm/yyyy)	28/12/1978
Social Security, Passport, ID number	75814059X		
e-mail	<a href="mailto:abombarely@ibmcp.upv.es">abombarely@ibmcp.upv.es</a>	URL Web	
Open Researcher and Contributor ID (ORCID) (*)	0000-0001-6257-8914		

(\*) Mandatory

**A.1. Current position**

Position	CSIC Professor		
Initial date	08/05/2021		
Institution	Instituto de Biología Molecular y Celular de Plantas (IBMCP)		
Department/Center	Biotecnología y Mejora Vegetal de Especies Cultivadas		
Country	Spain	Teleph. number	963877014
Key words	Genomics, Bioinformatics, Plant Evolution, Domestication		

**A.2. Previous positions (research activity interruptions, indicate total months)**

Period	Position/Institution/Country/Interruption cause
2018-2021	Associate Professor/University of Milan/Italy
2014-2018	Assistant Professor/Virginia Tech/USA
2012-2014	Research Associate/Cornell University/USA
2008-2013	Postdoctoral Associate/Boyce Thompson Institute/USA

**A.3. Education**

PhD, Licensed, Graduate	University/Country	Year
Ph.D. (Biochem. & Mol. Bio.)	University of Malaga/Spain	2007
B.Sc. (Chemistry)	University of Malaga/Spain	2001

(Include all the necessary rows)

**Part B. CV SUMMARY** (max. 5000 characters, including spaces)

I was an undergraduate student when the *A. thaliana* genome was published in 2000. At that time, I didn't know that I will dedicate my career to understand how the genetic information encoded in the genomes is linked to the plant evolution and domestication. Twenty-three years after, I feel confident that I have contributed to the plant genomic field in many ways.

I started participating in large sequencing consortiums (watermelon, tomato, 2012; mungbean, 2014) with small tasks: Verify assemblies, perform phylogenetic analysis, or annotate repeats. In 2012, I was responsible to produce a draft genome assembly for *N. benthamiana* proving that you don't need a big consortium to sequence a genome. It has been an unvaluable resource for the plant community. My work with this species drove to a collaboration with PMI to elucidate its evolutionary origin in 2017. I started to lead sequencing projects with our petunia work published in 2016. Since then, I led and collaborated in several international genome sequencing efforts like avocado (US, ZA), cherimoya (ES), pawpaw (ES, US), zucchini (ES), orchardgrass (CN) and mango (US, IL). The avocado sequencing is part of a \$4M USDA project in which I am a Co-PI.



In parallel, I have been working to make these genomes more accessible to the community through databases like the SGN. More recently, in collaboration with Noe Fernandez-Pozo, we developed a genomic database system called EasyGDB aiming to facilitate the generation of genomic portals. It is the software behind the Perseabase or OliveTreeDB. My group has developed other tools like Organelle\_PBA, to assemble organelle genomes from long reads, DeepTE, a tool to classify repetitive elements using Deep Learning and TEmarker, a tool to call variants associated to transposable elements.

Nevertheless, it is not the generation of these resources the main driver of my research, it is to contribute to the knowledge through the study of the evolution and domestication of plants. My involvement in the olive genome project has evolved to collaborations with researchers in Spain and Greece, with studies about the agronomical traits and the origin of an olive tree domestication. We proposed *Sinningia speciosa*, an ornamental plant, as a model for plant domestication thanks to its small genome size, its lower generation time and more importantly, because all the cultivars are descendent of a single wild population collected 200 years ago in Brazil. Our story published in 2019 got the attention of Prof. Friedman (Harvard University) who asked for material to use it for his botany class. The results of all these contributions are accessible through 76 peer-review publications (8 as corresponding) since 2008, with more than the half published as open access and 13 of them with more than 100 cites. They have been communicated in 40 conferences, and as international player in the science game, I have been invited as speaker 26 times in 12 different countries.

An unusual feature of my career has been the fast promotion through the different stages caused by the need of scientists with a computational/genomic profile. I was able to secure my current position as CSIC professor at the IBMCP seven years after I started as assistant professor at VT. Last year my department asked me to serve as department chair. Although, these events have been positive for my career, they have limited the funding that I have been able to secure and have impacted the number of postdoctoral researchers (2), PhD (3) and MS (4) students that I have mentored. Nevertheless, I think that they have been successful in their careers. My first postdoctoral researcher holds a scientist permanent position at CSIC and the second is leading a research department in a German company. My first PhD student has a permanent position in a US biotech company and my second has a successfully postdoctoral position at the UGA. 3 of the 4 MS students are pursuing an academic career. I have taught bioinformatics, genomics, and plant evolution for more than 10 years in 7 different countries, from workshops to regular courses at MS degrees.

My contributions to the community do not stop in the training of a next generation of scientists. I have served as reviewer for more than 100 articles being a usual reviewer for Nature Genetics. I am editor of BMC Genomics, BMC Plant Biology and Genes. I have served as evaluator of grant proposals for agencies like NSF, Human Frontiers and the *Agencia Española de Investigación*. I am ERGA Science Committee Representative for the Genome Annotation group since 2021. In recognition to my relevance in the Solanaceae research community, I was appointed Chair of the SOL Meetings in 2022. I received the VI International Olive Cultivation Research Award “Luis Vañó” for co-leading the sequencing of Picual olive tree genome in 2022. I am active in the social media, keeping a science focused Twitter account (@aubombarely) with more than 900 followers.

## Part C. RELEVANT MERITS (sorted by typology)

### C.1. Publications

*Note: Members of the Bombarely laboratory are in highlighted with **bold** font. Corresponding authors (CA) are marked with a \*.*

1. **Yan H**, Sun M, Zhang Z, ... **Bombarely A**, ... Huang L\* (19/26) (2023) Pangenomic analysis identifies structural variation associated with heat tolerance in pearl millet. Nature Genetics. Accepted on 19/01/2023. Available by request.
2. **Yan H**, Haak DC, Li S, Huang L, **Bombarely A\*** (2022) Transposable elements-based markers identify allelic variation among agronomic traits in rice. Plant Communications, 3(3):100207. doi:10.1016/j.xplc.2021.100270.



- Ziarsolo P, **Hasing T**, Hilario R, Garcia-Carpintero V, Blanca J, **Bombarely A\***, Cañizares J\* (2021) K-seq, an affordable, reliable, and open Klenow-NGS based genotyping technology. *Plant Methods* 17, 30. doi:10.1186/s13007-021-00733-6.
- Yan H**, **Bombarely A**, Li S\* (2020) DeepTE a computational method for de novo classification of transposons with convolutional neural network. *Bioinformatics*, btaa 519. doi:10.1093/bioinformatics/btaa519.
- Jimenez-Ruiz J, Ramirez-Tejero JA, Fernandez-Pozo N ... **Bombarely A\***, Luque F\* (14/15) (2020) The phenotypic diversity of the olive tree (*Olea europaea* L.) domestication is driven by transposon activity. *The Plant Genome* 13:e20010. doi:10.1002/tpg2.20010.
- Gramazio P, **Yan H**, **Hasing T**, Vilanova S, Prohens J, **Bombarely A\*** (2019) Whole-genome resequencing of seven eggplant (*Solanum melongena*) and one wild relative (*S. incanum*) accessions provides new insights and breeding tools for eggplant enhancement. *Frontiers in Plant Science* 10:1220. doi:10.3389/fpls.2019.01220.
- Hasing T**, **Rinaldi E**, Manrique S, Colombo L, Haak DC, Zaitlin D, **Bombarely A\*** (2019) Extensive phenotypic diversity in the cultivated Florist's Gloxinia, *Sinningia speciosa* (Lodd.) Hiern, is derived from the domestication of a single founder population. *Plants, People, Planet* 1: 363– 374. doi:10.1002/ppp3.10065.
- Soorni A**, Fatahi R, Haak D, Salami SA, **Bombarely A\***. (2017). Assessment of Genetic Diversity and Population Structure in Iranian Cannabis Germplasm by Genotyping-By-Sequencing Data. *Scientific Reports* 7:15668. doi:10.1038/s41598-017-15816-5.
- Soorni A**, Haak D, Zaitlin D, **Bombarely A\*** (2017) Organelle\_PBA, a pipeline for assembling chloroplast and mitochondrial genomes from PacBio DNA sequencing data. *BMC Genomics* 18 (1):49. doi:10.1186/s12864-016-3412-9.
- Bombarely A**, Moser M, Amrad A, ... Kuhlemeier C\* (1/56) (2016) Whole genome sequences of the wild parents of the garden petunia give insights into the evolution of Solanaceae genomes. *Nature Plants* 2: 16074. doi:10.1038/nplants.2016.74.

## C.2. Congress

*Note: Members of the Bombarely laboratory are in highlighted with bold font. Presenting authors (PA) are marked with a \*.*

- Oral presentation: Backer R, **Bombarely A\***, Chambers AH..., van den Berg N Impact of the genome assembly quality for allele mining in avocado. PAG30. 13/01/2023-19/01/2023. San Diego, California.
- Bombarely A\***, Di Marsico M, **Garcia-Carpintero V**... Matus JT. (2023) Grapevine takes it global: The GRAPEDIA Initiative. PAG30. 13/01/2023-19/01/2023. San Diego, California.
- Invited conference: **D'Andrea L**, Sierro N, Ouadi S, **Hasing T**, **Rinaldi E**, Ivanov NV, **Bombarely A\***. Insights of the complex origin of the Nicotiana genus, section Suaveolentes. XVII International Conference of the Plant Family Solanaceae 2022. 01/11/2022-05/11/2022 Thessaloniki, Greece.
- Invited conference: **Hasing T**, Zaitlin D, **Bombarely A\***. Linking Allele Variation to the Phenotypic Biodiversity in the Sinningia genus during Domestication. XXII International Congress of Genetics. 10/09/2018-14/10/2018. Foz do Iguazu, Brazil.
- Invited conference: **Bombarely A\***. Development of Community Driven Avocado Genomic Resources. Avocado Brainstorming 2018. 28/05/2018-01/06/2018. Tzaneen, South Africa.
- Oral presentation: **Bombarely A\***, Giudicelli GC, Hasing T, Petzold E, Freitas LB Petunia domestication through the analysis of the population structure and genetic



diversity. 16th World Petunia Days. 15/03/2018-18/03/2018. Amsterdam, The Netherlands.

7. Oral presentation: **Bombarely A\*** The *Nicotiana benthamiana* genome 2.0: from genes to pseudomolecules. 8th Solanaceae and 2nd Cucurbitaceae Genome Joint Conference. 03/09/2017-06/09/2017. Valencia, Spain.
8. Oral presentation: **Bombarely A\*** Insights of the genome architecture for the species *Begonia conchifolia*. International Botanical Conference 2017. 23/07/2017-29/07/2017. Shenzhen, China.
9. Oral presentation: **Bombarely A\*** Analysis of the Genomes Architectures for *Petunia hybrida* and its Wild Relatives. Plant & Animal Genome XXV. 14/01/2017-18/01/2017. San Diego, CA, USA.
10. Oral presentation: **Bombarely A\***, Zaitlin D (2017) Analysis of the *Sinningia speciosa* Diversity under Domestication. Plant & Animal Genome XXV. 14/01/2017-18/01/2017. San Diego, CA, USA.

### C.3. Research projects.

1. IG17111. *The Grapevine Genomics Encyclopedia: an innovative portal to integrate knowledge, resources and services for the grape scientific community and industry*. 01/11/2023-31/10/2024. €125,000. EU Cost Innovation Grant 2022. Dr. Tomas Matus, I2SysBio. Dr. Bombarely role: Team-member.
2. H2020-MSCA-RISE-2020-101007738. *EVOfruland - Evolution of genetic network required for fruit and fruit-like structures development of land plants*. 01/09/2021-31/08/2025. €662,400. H2020-MSCA-RISE-2020. Dr. Barbara Baldan, University of Padova. Dr. Bombarely role: Team-member.
3. CA-R-MPP-5217-CG. *Reducing avocado losses to major challenges by improving resistance selection and disease management using next generation technologies*. National Institute for Food and Agriculture (NIFA). Specialty Crop Research Initiative (SCRI) 2020. Dr. Patricia Manosalva, University of California, Riverside (CA, USA). 01/09/2020-31/08/2024. \$4,401,036. Dr. Bombarely role: Sub-PI funded with \$186.347.
4. None. *Sequencing the Asimina triloba (L.) Dun. (Annonaceae) genome*. Harvard Arboretum. Genomics Initiative and Sequencing Award 2017. Drs. Aureliano Bombarely and Jose I. Hormaza. 01/03/2017-30/06/2017. \$10,000. Dr. Bombarely role: PI.

### C.4. Contracts, technological or transfer merits

1. *Genome Evolution under Adaptation to Arid Environments in the Suaveolentes section (Nicotiana genus)*. Phillip Morris International (PMI). NA. Dr. Aureliano Bombarely, Virginia Tech (VA, USA). 01/10/2017-31/13/2017. \$66,000. Dr. Bombarely role: PI.