

MARCOS DÍAZ GAY

San Diego, CA 92122 (USA) • m.diaz.gay@gmail.com • +1 (858) 951-5811 • marcos-diazg.github.io

SUMMARY

I am a postdoctoral researcher in the group of Dr. Ludmil B. Alexandrov at the University of California San Diego. Originally trained in engineering, I have focused my career on computational biology and bioinformatics, with a strong interest in cancer genomic data analysis. My research focuses on the development of new computational methodologies that allow the discovery of the underlying mechanisms of cancer predisposition and progression.

EDUCATION

University of Barcelona
2019
Ph.D. | Medicine and Translational Research

University of Barcelona
2016
M.S. | Biomedicine

University of A Coruña
2015
B.S. & M.S. | Civil Engineering

TEACHING EXPERIENCE

University of California
San Diego
2023
Lecturer
Instructor of record for the Bioinformatics Laboratory (BIMM 143) undergraduate course offered by the School of Biological Sciences (Spring 2023 quarter – 60 hours).

Wellcome Connecting Science
2023 – 2022
Course Instructor
Cancer Genome Analysis Africa (40 hours – online) and Latin America and the Caribbean (60 hours – in person) global training courses. Led mutational signatures module, including lecture and practical hands-on session. Collaborated in other modules.

National Cancer Institute,
National Institutes of Health
2023
Course Instructor
Emerging Approaches for Tumor Analyses in Epidemiological Studies online course offered by the Division of Cancer Epidemiology and Genetics. Lectured on mutational signatures and assisted during practical session.

RESEARCH EXPERIENCE

University of California
San Diego
2020 – present
Postdoctoral Scholar
Advisor: Ludmil B. Alexandrov

- Analysis of the mutagenic forces driving lung cancer in never smokers (Sherlock project, in collaboration with the National Cancer Institute).
- Analysis of mutational signatures in colorectal cancer from 11 countries of varying incidence (Mutographs project, in collaboration with the Wellcome Sanger Institute and IARC – World Health Organization).
- Developed most advanced tools for extraction (published and cover of Cell Genomics) and assignment of mutational signatures (published in Bioinformatics).

August Pi i Sunyer Biomedical
Research Institute (IDIBAPS)
2016 – 2019
Ph.D. Student
Advisors: Sergi Castellví-Bel & Francesc Balaguer

- Identified novel genes involved in colorectal cancer predisposition.
- Developed web-based application for mutational signature analysis.

University of California
San Diego
2019

Visiting Graduate Student

Advisor: Ludmil B. Alexandrov

- Four months international research stay funded by a competitive grant.
- Developed somatic variant calling pipeline and performed mutational signature analysis.

August Pi i Sunyer Biomedical
Research Institute (IDIBAPS)
2015 – 2016

Master's Student

Advisor: Sergi Castellví-Bel

- Developed germline-somatic variant prioritization pipeline for novel predisposition genes involved in colorectal cancer.

University of A Coruña
2013 – 2015

Undergraduate Research Assistant

Advisor: Joaquín Suárez

- Collaborated in the Water and Environmental Engineering group

PUBLICATIONS

Total number of publications: 25 (2 submitted) | Citations: 520 ([Google scholar](#)) | h-index: 12 ([Google scholar](#))

Original articles as senior author

1. Assigning mutational signatures to individual samples and individual somatic mutations with SigProfilerAssignment
Díaz-Gay M.[^] *et al. **Bioinformatics.** 2023;39(12):btad756.* Citations: 4 (Google Scholar). Journal Impact Factor: 5.8 – Q1 (Clarivate Journal Citation Reports 2022).
[^] shared corresponding author

Original articles as first author

2. Uncovering novel mutational signatures by *de novo* extraction with SigProfilerExtractor
Islam S.M.A.*[,] **Díaz-Gay M.*** *et al. **Cell Genomics.** 2022;2(11):100179.* Cit: 121.
^{*} shared first authors
3. Using linkage studies combined with whole-exome sequencing to identify novel candidate genes for familial colorectal cancer
Toma C.*[,] **Díaz-Gay M.*** *et al. **International Journal of Cancer.** 2020;146(6):1568–1577.* Cit: 11. JIF: 7.396 – Q1 (Clarivate JCR 2020).
4. Identification of a novel candidate gene for serrated polyposis syndrome by performing linkage analysis combined with whole-exome sequencing
Toma C.*[,] **Díaz-Gay M.*** *et al. **Clinical and Translational Gastroenterology.** 2019;10(10):e00100.* Cit: 7. JIF: 3.968 – Q2.
5. Integrated analysis of germline and tumor DNA identifies new candidate genes involved in familial colorectal cancer
Díaz-Gay M. *et al. **Cancers.** 2019;11(3):362.* Cit: 19. JIF: 6.126 – Q1.
6. Mutational signatures in cancer (MuSiCa): a web application to implement mutational signatures analysis in cancer samples
Díaz-Gay M. *et al. **BMC Bioinformatics.** 2018;19(1):224.* Cit: 98. JIF: 2.511 – Q1.

Reviews and book chapters as first author

7. Unraveling the genomic landscape of colorectal cancer through mutational signatures [**book chapter**]
Díaz-Gay M. & Alexandrov L.B. ***Advances in Cancer Research: "Novel Approaches to Colorectal Cancer".** 2021;151:385-424.* Cit: 16. JIF: 5.767 – Q2.
8. Somatic mutational signatures in polyposis and colorectal cancer [**review**]
Grolleman J.*[,] **Díaz-Gay M.*** *et al. **Molecular Aspects of Medicine.** 2019;69:62–72.* Cit: 14. JIF: 9.577 – D1.

Original articles as second author

9. Geographic variation of mutagenic exposures in kidney cancer genomes
Senkin S.*, Moody S.*, **Díaz-Gay M. et al. *Nature*. 2023. In press (accepted in principle)**. Cit: 1. JIF: 64.8 – D1.
10. Topography of mutational signatures in human cancer
Otlu B., **Díaz-Gay M. et al. *Cell Reports*. 2023;42(8):112930**. Cit: 12. JIF: 8.8 – Q1.
11. Comprehensive genomic characterization of early-onset Lynch-like syndrome colorectal cancers
Golubicki M., **Díaz-Gay M. et al. *Cancers*. 2021;13(6):1259**. Cit: 5. JIF: 6.575 – Q1.

Original articles in collaboration

12. APOBEC shapes tumor evolution and age at onset of lung cancer in smokers
Zhang T. et al. ***Nature Communications*. 2023. Submitted (under review)**. JIF: 16.6 – D1.
13. Deep learning predicts HRD and platinum response from histology slides in breast and ovarian cancer
Bergstrom E. et al. ***Journal of Clinical Oncology*. 2023. Submitted (under review)**. JIF: 45.4 – D1.
14. Visualizing and exploring patterns of large mutational events with SigProfilerMatrixGenerator
Khandekar A. et al. ***BMC Genomics*. 2023;24(1):469**. Cit: 2. JIF: 4.4 – Q1.
15. Unraveling the impact of a germline heterozygous *POLD1* frameshift variant in serrated polyposis syndrome
Bonjoch L. et al. ***Frontiers in Molecular Biosciences*. 2023;10:1119900**. Cit: 4. JIF: 5.0 – Q2.
16. Germline mutations in *WNK2* could be associated with serrated polyposis syndrome
Soares de Lima Y. et al. ***Journal of Medical Genetics*. 2023;60(6):557-567**. Cit: 3. JIF: 4.0 – Q2.
17. Germline and somatic WES identifies new candidate genes involved in familial predisposition to serrated polyposis syndrome
Soares de Lima Y. et al. ***Cancers*. 2021;13(4):929**. Cit: 14. JIF: 6.575 – Q1.
18. Identification of new genes involved in germline predisposition to early-onset gastric cancer
Herrera-Pariente C. et al. ***International Journal of Molecular Sciences*. 2021;22(3):1310**. Cit: 8. JIF: 6.208–Q1
19. Germline biallelic mutations in *MCM8* are associated with early-onset lynch-like syndrome
Golubicki M. et al. ***JCI Insight*. 2020;5(18):e140698**. Cit: 21. JIF: 8.315 – Q1.
20. Germline mutations in *FAF1* are associated with hereditary colorectal cancer
Bonjoch L. et al. ***Gastroenterology*. 2020;159(1):227-240**. Cit: 19. JIF: 22.682 – D1.
21. Colorectal cancer genetic variants are also associated with serrated polyposis syndrome susceptibility
Arnau-Collell C. et al. ***Journal of Medical Genetics*. 2020;57(10):677-682**. Cit: 13. JIF: 6.318 – Q1.
22. CNApp, a tool for the quantification of copy number alterations and integrative analysis revealing clinical implications
Franch-Expósito S. et al. ***eLife*. 2020;9:e50267**. Cit: 51. JIF: 8.146 – D1.
23. Quantitative analysis of somatically-acquired and constitutive uniparental disomy in gastrointestinal cancers
Torabi K. et al. ***International Journal of Cancer*. 2019;144(3):513–524**. Cit: 8. JIF: 5.145 – Q1.
24. Rare germline copy number variants in colorectal cancer predisposition characterized by exome sequencing analysis
Franch-Expósito S. et al. ***Journal of Genetics and Genomics*. 2018;45(1):41–45**. Cit: 17. JIF: 4.65 – Q1.
25. *POLE* and *POLD1* screening in 155 patients with multiple polyps and early-onset colorectal cancer
Esteban-Jurado C. et al. ***Oncotarget*. 2017;8(16):26732–26743**. Cit: 49. JIF: 5.168 – Q1.

FUNDING

University of Barcelona
2018

Faculty of Medicine Short-Term International Research Stay Fellowship
Amount: 3,000€.

Regional Government of Catalunya 2017 – 2019	Agency for Management of University and Research Grants (AGAUR) Ph.D. Fellowship (FI-2017) Amount: 61,200€. 3 years fellowship covering salary and tuition.
Spanish Ministry of Education 2014	Department Collaboration Scholarship University of A Coruña. Department of Mathematical and Representation Methods. Amount: 2,000€.
University of A Coruña 2013	Lab Collaboration Scholarship School of Civil Engineering. Water and Environmental Engineering Group. Amount: 3,150€.

HONORS & AWARDS

Spanish Ministry of Universities 2022	National Faculty Habilitations for Assistant and Associate Professor (Profesor Ayudante Doctor / Profesor Contratado Doctor / Profesor de Universidad Privada) Spanish National Agency for Quality Assessment and Accreditation (ANECA) and Catalan University Quality Assurance Agency (AQU Catalunya).
University of Barcelona 2019	<i>Cum Laude</i> Honors & International Mention (Ph.D.) Faculty of Medicine.
University of Barcelona 2018	Extraordinary Master's Degree Award Faculty of Medicine.
University of A Coruña 2016	Extraordinary End-Of-Degree Award School of Civil Engineering.

MEMBERSHIP IN PROFESSIONAL ASSOCIATIONS

Professional Associations

European Association for Cancer Research (EACR) May 2021 – present
Spanish Association for Cancer Research (ASEICA) May 2021 – present
Spanish Scientists in the USA (ECUSA) Jan 2021 – present
American Association for Cancer Research (AACR) Nov 2019 – present

SERVICE

Peer-reviewed articles

BMC Bioinformatics (BMC) 1 article
Cancers (MDPI) 1 article
Cell Genomics (Cell Press) 1 article
Cell Reports Medicine (Cell Press) 1 article
Communications Biology (Springer Nature) 1 article (2 reviews)
Computational and Structural Biotechnology Journal (Elsevier) 1 article
Genes (MDPI) 2 articles (3 reviews)
Human Genomics (BMC) 1 article (3 reviews)

Heliyon (Elsevier) 1 article (2 reviews)

Journal of Personalized Medicine (MDPI) 1 article (2 reviews)

PLOS Computational Biology (Public Library of Science) 1 article (2 reviews)

RESEARCH PROJECTS

Collaborator

1. Mutographs

Cancer Grand Challenges (Cancer Research UK)

01/2017 – 12/2023

20,000,000£

PI: Mike Stratton (Leader work area 2: Ludmil B. Alexandrov)

2. Caracterización funcional de variantes en nuevos genes candidatos en la predisposición germinal al cáncer colorrectal y el síndrome de poliposis serrada (FUNCTION4GENE)

Instituto de Salud Carlos III

Program: Fondo de Investigación Sanitaria (FIS)

ID: PI17/00878

01/2018 – 06/2022

100,430€

PI: Sergi Castellví Bel

3. Identifying biomarkers through translational research for prevention and stratification of colorectal cancer (TRANSCOLONCAN)

European Union

Program: COST programme

ID: Action CA17118

11/2018 – 04/2022

234,580€

PI: Sergi Castellví Bel

4. Grupo de Investigación en Predisposición Genética al Cáncer Gastrointestinal

Agencia de Gestio d'Ajuts Universitaris i de Recerca - Generalitat De Catalunya

Programa: Grupo de Investigación Pre-consolidado Reconocido

ID: GRPRE 2017SGR21

01/2017 – 09/2021

PI: Sergi Castellví Bel

5. Prevenció del càncer colorrectal en la població de risc mitjà mitjançant biomarcadors genòmics i microbiòmics (CRIPREV)

Departament de Salut, Generalitat de Catalunya

Program: Pla estratègic de recerca i innovació en salut (PERIS)

ID: SLT002/16/00398

03/2017 – 12/2019

280,479€

PI: Antoni Castells Garangou

6. Identificación de nuevos biomarcadores para la prevención del cáncer colorrectal

Fundacion Científica Asociación Española Contra el Cáncer

ID: GCB13131592CAST

10/2013 – 09/2019

553,500€

PI: Antoni Castells Garangou

7. **Cooperation studies on inherited susceptibility to colorectal cancer**
 European Union
 Program: COST programme
 ID: CE_COST_2012 Action BM1206
 06/2013 – 05/2018
 244,350€
 PI: Sergi Castellví Bel
8. **Identificación de genes de predisposición germinal en un subgrupo de cáncer colorrectal: el síndrome de poliposis serrada en el punto de mira**
 Instituto de Salud Carlos III
 Program: Fondo de Investigación Sanitaria (FIS)
 ID: PI14/00173
 01/2015 – 12/2017
 94,380€
 PI: Sergi Castellví Bel
9. **Grupo de Investigación en Oncología Gastrointestinal y Pancreática**
 Agencia de Gestio d'Ajuts Universitaris i de Recerca - Generalitat De Catalunya
 ID: 2014SGR135
 01/2014 – 07/2017
 50,000€
 PI: Antoni Castells Garangou
10. **Grupo de Investigación en Predisposición Genética al Cáncer Colorrectal**
 Agencia de Gestio d'Ajuts Universitaris i de Recerca - Generalitat De Catalunya
 Programa: Grupo de Investigación Emergente Reconocido
 ID: GRE 2014SGR255
 01/2014 – 04/2017
 PI: Sergi Castellví Bel

CONFERENCE COMMUNICATIONS

Invited talks

1. Unravelling the genomic landscape of cancer through mutational signatures.
Curso de Entrenamiento Intensivo para el Manejo Interdisciplinario de los Tumores Digestivos, Intergupo Argentino para el Tratamiento de los Tumores Gastrointestinales (virtual), 09/2023.
2. Unravelling the genomic landscape of cancer through mutational signatures.
Future Leaders Virtual Seminar Series, Cancer Grand Challenges (CRUK & NCI) (virtual), 07/2023.
3. Uncovering novel mutational signatures of tobacco smoking with SigProfilerExtractor.
SoCal Genome Stability Symposium, UC Irvine, Irvine, CA, USA, 12/2022.
4. Unraveling the genomic landscape of cancer through mutational signature extraction and assignment.
Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain, 11/2022.
5. Unraveling the genomic landscape of cancer through *de novo* extraction of mutational signatures.
ISCB academy, International Society for Computational Biology (virtual), 06/2022.
6. Unraveling the genomic landscape of cancer through mutational signatures.
Universidade de Vigo (virtual), 02/2022.
7. A new user-friendly web application to implement mutational signatures analysis.
European Congress of Pathology, European Society of Pathology, Nice, France, 09/2019.

Oral communications

8. The mutagenic forces shaping the genomic landscape of lung cancer in never smokers.

AACR Annual Conference, American Association for Cancer Research, San Diego, CA, USA, 04/2024.

9. Genomic characterization of colorectal cancer across the world by de novo extraction of mutational signatures with SigProfilerExtractor.
European Association for Cancer Research Congress, Sevilla, Spain, 06/2022.
10. SigProfilerWeb: a user-friendly web application for mutational signatures refitting according to the new COSMIC framework.
European Hereditary Tumour Group Meeting, Barcelona, Spain, 10/2019.
11. Integrated analysis of germline and tumor DNA identifies new candidate genes involved in familial colorectal cancer.
Human Genome Meeting, Human Genome Organisation, Barcelona, Spain, 02/2017.
12. Integrated analysis of germline and tumor DNA for the identification of new genes involved in familial colorectal cancer.
Bioinformatics and Genomics Symposium, Societat Catalana de Biologia, Barcelona, Spain, 12/2016.
13. Highway runoff treatment evaluation using continuous turbidity measurements.
International Conference on Diffuse Pollution and Eutrophication, International Water Association, Berlin, Germany, 09/2015.

Posters

14. Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor.
Future Leaders Conference, Cancer Grand Challenges (CRUK & NCI), Barcelona, Spain, 11/2022.
15. Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor.
Moore's Cancer Center Delivering Discoveries Scientific Retreat, UC San Diego, La Jolla, CA, USA, 10/2022.
16. Enrichment of polymerase epsilon and delta exonuclease domain mutations in microsatellite unstable human cancers.
Structural and Functional Genomics Program Annual Retreat, UC San Diego, La Jolla, CA, USA, 03/2022.
17. Enrichment of polymerase epsilon and delta exonuclease domain mutations in microsatellite unstable human cancers.
Moore's Cancer Center Delivering Discoveries Scientific Retreat, UC San Diego (virtual), 11/2021.
18. Enrichment of polymerase epsilon and delta exonuclease domain mutations in microsatellite unstable human cancers.
SoCal Genome Stability Symposium, The Scripps Research Institute (virtual), 09/2021. (Best poster prize).
19. Integrated analysis of germline and tumor DNA identifies new candidate genes involved in familial colorectal cancer.
IDIBAPS PhD Day, Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain, 06/2019.
20. Integrated analysis of germline and tumor DNA for the identification of new candidate genes involved in familial colorectal cancer.
Bioinformatics and Genomics Symposium, Societat Catalana de Biologia, Barcelona, Spain, 12/2018.
21. Integrated analysis of germline and tumor DNA for the identification of new candidate genes involved in familial colorectal cancer.
Jornadas Científicas CIBERehd, Barcelona, Spain, 11/2018.
22. Mutational Signatures in Cancer (MuSiC): a web application to implement mutational signatures framework in cancer samples.
Bioinformatics and Genomics Symposium, Societat Catalana de Biologia, Barcelona, Spain, 12/2017.
23. Integrated analysis of germline and tumor DNA for the identification of new candidate genes involved in familial colorectal cancer.

European Human Genetics Conference, European Society of Human Genetics, Copenhagen, Denmark, 05/2017.

24. Integrated analysis of germline and tumor DNA identifies new candidate genes involved in familial colorectal cancer.

Translating Colorectal Cancer Research Workshop, COST EuCOLONGENE, Porto, Portugal, 02/2017.

25. Identification of new genes involved in familial colorectal cancer: integrated analysis of germline and tumor DNA.

Jornadas Científicas CIBERehd, Barcelona, Spain, 10/2016. (**Best poster prize**).

26. Constructed wetlands for the treatment of domestic sewage in rural areas of Galicia (Spain): case studies.

International Society for Environmental Biotechnology, Barcelona, Spain, 06/2016.

MENTORING EXPERIENCE

Zichen (Cardiff) Jiang
2021 – 2023

Undergraduate Student, Biology, UC San Diego

BISP 196 course (Senior Honors Thesis Program).

BENG 199 course (Independent Study for Undergraduates).

Development of an unmatched somatic variant calling pipeline.

Currently bioengineering master student at UC San Diego.

Xi (Sam) Wang
2021 – 2022

Undergraduate Student, Bioengineering, UC San Diego

BENG 199 course (Independent Study for Undergraduates).

Benchmarking of bioinformatics tools for mutational signature assignment.

Currently biomedical informatics master student at Harvard Medical School.

Yasmin Soares de Lima
2019

Ph.D. Student, Medicine and Translational Research, University of Barcelona

Identification of novel genes involved in predisposition to serrated polyposis syndrome by a germline-tumor integrated analysis.

First-author in publication 17. Collaborator in publications 3 and 4.

Roser Capó-García
2018

Master's Student, Biomedicine, University of Barcelona

Identification of novel predisposition genes in early-onset gastric cancer.

Collaborator in publication 18.

Mariano Golubicki
2017 – 2018

Ph.D. Student, Biochemistry, University of Buenos Aires (Argentina)

Genomic characterization of Lynch-like syndrome colorectal cancers.

First-author in publication 11. Collaborator in publication 18.

Paula Alejandra Sánchez-Rojas
2017

Master's Student, Biomedicine, University of Barcelona

Identification of novel predisposition genes in familial colorectal cancer.

Collaborator in publication 5.

TEACHING TRAINING

University of California
San Diego
2022

Introduction to College Teaching

Preparation to be an instructor of record by developing evidence-based effective teaching practices that support student learning (30 hours).

University of California
San Diego
2021

Pathways to Scientific Teaching

Life sciences postdoc training in scientific teaching. Development of learner-centered instructional materials and teaching strategies (18 hours).

COMPUTATIONAL SKILLS

Programming languages Proficient in R scripting programming (8 years of experience in NGS analysis, statistical analysis, data mining, and data visualization). Working knowledge of Python, Shiny web development environment, UNIX shell scripting, and HPC cluster environments. Basic knowledge of CSS, Fortran, HTML, JavaScript, MySQL, and Perl.

Software and websites Co-author and maintainer of the [SigProfilerAssignment](#) Shiny web application, [SigProfilerMatrixGenerator](#), [SigProfilerExtractor](#) and [SigProfilerAssignment](#) Python/R packages for mutational signature analysis, and [COSMIC Mutational Signatures](#) website.

Bioinformatic tools Bioconductor, Conda, IGV, ImageJ, PyPI, and GitHub user (github.com/marcos-diazg).

LANGUAGE SKILLS

Spanish Native speaker | **Galician** Native speaker | **Catalan** Limited working proficiency

English Full professional proficiency (C1 Cambridge CAE certificate) | **Portuguese** Limited working proficiency

French Limited working proficiency (B2 DELF certificate)

REFERENCES

Ludmil B. Alexandrov, Ph.D. (Postdoctoral Advisor)

Cellular and Molecular Medicine, and Bioengineering departments, Moores Cancer Center, University of California San Diego. 92093 La Jolla, CA (United States of America). (+1) 858 246 2747. l2alexandrov@ucsd.edu

Sergi Castellví-Bel, Ph.D. (Ph.D. Advisor)

Genetic Predisposition to Gastrointestinal Cancer group, August Pi i Sunyer Biomedical Research Institute (IDIBAPS), CIBERehd, Hospital Clínic. 08036 Barcelona (Spain). (+34) 93 227 54 00 ext.4183. sbel@clinic.cat

Francesc Balaguer, M.D., Ph.D. (Ph.D. Advisor)

Gastroenterology department, Hospital Clínic, CIBERehd, IDIBAPS. 08036 Barcelona (Spain). (+34) 93 227 54 00 ext.2802. fprunes@clinic.cat

Jordi Camps, Ph.D. (Collaborator)

Cellular Biology and Medical Genetics unit, Cellular Biology, Physiology and Immunology department, Universitat Autònoma de Barcelona. 08193 Bellaterra (Spain). (+34) 93 227 54 00 ext.2915. jordi.camps@uab.cat