Pablo S. C. Santos

Curriculum Vitae

Antonienstrasse 18
04229 - Leipzig, Germany
☎ *+49 176-28692001*
✉ *mail@psc-santos.com*
www.psc-santos.com
psc-santos
psc-santos

Last Updated: April 18, 2023

Computational Biologist with expertise in Molecular Pathology, Immunogenetics, Ecology, Data Science and Statistics. PhD in Human Genetics. Enthusiastic for Teaching and Open Data/Source/Science. Friends with the industry.

Education

2010	**Ph.D. in Human Genetics**, *Freie Universität Berlin and Charité Universitätsmedizin Berlin*, Germany Dissertation: Linkage Disequilibrium and Transmission Distortion Affecting Human Chromosome 6p (`https://refubium.fu-berlin.de/handle/fub188/9777`)
Honors:	Magna Cum Laude
2005	**Bachelor of Science in Genetics and Teacher's Certification**, *Universidade Federal do Paraná*, Curitiba, Brazil Thesis: Linkage Disequilibrium and MHC Haplotype Frequencies of 17,000 Stem Cell Donors from Curitiba/PR
Grade:	100/100

Experience

2022–Present	**Global Health Group, University of Halle-Wittenberg**, *Halle, Germany*, Senior Bioinformatician and Data Scientist Here I provide expertise in Bioinformatics, Data Science and Statistics by supervising MD and PhD projects in a team of medical doctors, epidemiologists, statisticians and social scientists. We aim at understanding cancer and improving preventive health care in the context of international colaboration.
2022–Present	**CQ Beratung+Bildung GmbH**, *Vocational School in Berlin, Germany*, Instructor Here I teach Bioinformatics online in a regular basis to a group of 14 to 16 stundents.
2020–2021	**PathoNext GmbH**, *Leipzig, Germany*, Senior Bioinformatician and Head of NGS Laboratory (Technischer Leiter) After working for 5 months as a Senior Bioinformatician, I assumed the full leadership of the NGS laboratory and the bioinformatics section of the company. I headed a highly motivated and international group of scientists and technical assistants providing sequencing services on the fields of Molecular Pathology and Oncology. My team and I established new products and improved compliance and quality management systems.
2020	**IAV GmbH**, *Gifhorn, Germany*, Data Scientist and Development Engineer After years in academia, this was my first leap into the private sector. Here I worked with an interdisciplinary team of data scientists developing products for the medical sector and applying data-driven methods into innovative industrial fields.

2014–2019 **Institute of Evolutionary Ecology and Conservation Genomics**, *University of Ulm, Germany*, Professor Assistant and Postdoc Scientist

Here I worked for 6 years as a computational and evolutionary biologist with focus on data analysis and bioinformatics. Apart from teaching (Bioinformatics, Genetics, Statistics, Ecology and Zoology), I have supervised and researched on projects tackling genomics, evolutionary ecology, disease transmission models and next-generation DNA sequencing. Data analysis and data-mining were among my primary methods, using tools such as Python (`pandas, scikit-learn, seaborn, requests, jupyter notebooks` and others), R (`ggplot2, lme4, glmer` and others) and Git/GitLab/GitHub.

2012–2013 **Immunogenetics and Histocompatibility Laboratory**, *UFPR University, Curitiba, Brazil*, Scientific Coordinator

Here I worked as a lecturer and scientific coordinator of a medium-sized workgroup focusing on immunogenetics, data mining, bioinformatics, organ transplant, miscarriage and cancer research. My workgroup genotyped voluntary stem cell donors for the National Bone Marrow Donor Registry.

2010–2012 **Leibniz Institute for Zoo and Wildlife Research**, *Berlin, Germany*, Postdoc Researcher

Here I was responsible for the network project *Olfactory choice of partners – immune system, smell receptors and their adaptive importance for mammal health.* Besides of networking with multiple collaborators, my tasks included experiment design, the interpretation of large data sets, statistics, data-mining and Illumina DNA sequencing

Awards/Fellowships

2007 – 2008 NaFöG Stipend from the Senate of Berlin, now called *Elsa-Neumann-Stipend*

2006 and 2009 Ernst-von-Leyden Stipend - Berliner Krebsgesellschaft Fellowship

Publications

Articles in Peer-Reviewed Journals

[1] Marcus Bauer, Martina Vetter, Kathrin Stückrath, Meron Yohannes, Zelalem Desalegn, Tewodros Yalew, Yonas Bekuretsion, Tariku W Kenea, Maureen Joffe amd Eunice J Van Den Berg, Julien Ilunga Nikulu, Kamaté Bakarou, Shyam S Manraj, Olufemi J Ogunbiyi, Ima-Obong Ekanem, Festus Igbinoba, Mohenou Diomande, Clement Adebamowo, Charles P Dzamalala, Angelica A Anele, Annelle Zietsman, Moses Galukande, Milena Foerster, Isabel Dos-Santos-Silva, Biying Liu, **Pablo S. C. Santos**, Ahmedin Jemal, Tamrat Abebe, Claudia Wickenhauser, Barbara Seliger, Valerie McCormack, and Eva J Kantelhardt. Regional variation in the tumor microenvironment, immune escape and prognostic factors in breast cancer in sub-saharan Africa. *Cancer Immunology Research*, Apr. 2023.
In Press.

[2] Zelalem Desalegn, Meron Yohannes, Martin Porsch, Kathrin Stückrath, Endale Anberber, **Pablo S. C. Santos**, Marcus Bauer, Adamu Addissie, Yonas Bekuretsion, Mathewos Assefa, Yasin Worku, Lesley Taylor, Tamrat Abebe, Eva Johanna Kantelhardt, and Martina Vetter. Intrinsic subtypes in Ethiopian breast cancer patient. *Breast Cancer Research and Treatment*, 196(3):495–504, Oct. 2022.

[3] Paulo Milet-Pinheiro, **Pablo S. C Santos**, Samuel Prieto-Benítez, Manfred Ayasse, and Stefan Dötterl. Differential evolutionary history in visual and olfactory floral cues of the bee-pollinated genus Campanula (Campanulaceae). *Plants*, 10(7), 2021.

[4] Mark A. F. Gillingham, B. Karina Montero, Kerstin Wihelm, Kara Grudzus, Simone Sommer, and **Pablo S. C. Santos**. A novel workflow to improve genotyping of multigene families in wildlife species: An experimental set-up with a known model system. *Molecular Ecology Resources*, 21(3):982–998, 2021.
Here, my colleagues and I developed an experimental setup to test and implement an analysis pipeline for NGS data on non-model organisms, witten by myself in Python: ACACIA (`https://gitlab.com/psc_santos/ACACIA`).

[5] Marianne Volleth, Stefan Müller, Simone Sommer, and **Pablo S. C. Santos**. Cytogenetic investigations in Emballonuroidea: Extensive chromosomal reorganization characterizes the karyotype of *Saccopteryx bilineata*. *Acta Chiropterologica*, 22(1):49–55, 2020.

[6] Juliana S. Schauren, Ana C. M. B. G. Torres, Rodrigo C. Almeida, **Pablo S. C. Santos**, Fabiane Mulinari-Brenner, Luiz H. Lima, Luiz A. Z. Filho, Naoye Shiokawa, Maria G. Bicalho, and Mario T. Sato. Novel *CDH3* variants in Brazilian families with hypotrichosis and juvenile macular dystrophy revealed by exome sequencing. *Clinical Genetics*, 97(3):529–531, Nov. 2019.

[7] **Pablo S. C. Santos**, Maja Mezger, Miriam Kolar, Frank-Uwe Michler, and Simone Sommer. The best smellers make the best choosers: mate choice is affected by female chemosensory receptor gene diversity in a mammal. *Proceedings of the Royal Society B: Biological Sciences*, 285(1893):20182426, Dec. 2018.

[8] **Pablo S. C. Santos**, Frank-Uwe Michler, and Simone Sommer. Can MHC-assortative partner choice promote offspring diversity? A new combination of MHC-dependent behaviours among sexes in a highly successful invasive mammal. *Molecular Ecology*, 26(8):2392–2404, Apr. 2017.
Here, my colleagues and I applied a statistical approach developed earlier by us (Santos et al., 2016) to disentangle ecological and evolutionary aspects of an invasive population of North American raccoons in Europe. This was a model to understand why some animal species turn invasive while others do not. The collaboration was carried out by myself and took place between our group and colleagues from the Technical University of Dresden.

[9] **Pablo S. C. Santos**, Alexandre Courtiol, Andrew J. Heidel, Oliver P. Höner, Ilja Heckmann, Martina Nagy, Frieder Mayer, Matthias Platzer, Christian C. Voigt, and Simone Sommer. MHC-dependent mate choice is linked to a trace-amine-associated receptor gene in a mammal. *Scientific Reports*, 6:38490, Dec. 2016.
Here, my colleagues and I used a 2-step statistical approach (Monte Carlo Randomizations and Generalized Linear Mixed Models) to detect and present evidence of an interaction between immune genes and olfactory receptor genes in a mammal (a Central American bat species). This was a large-scale collaboration project between our group and teams in Berlin, Jena and Nuremberg/Erlangen. My role was, apart of research, to coordinate the communication and collaboration among the partners. The statistical procedure used here has been well cited and proposes an improvement of the way researchers deal with the matter of mate-choice in ecology.

[10] José S. Silva, Pryscilla F. Wowk, Fabiana Poerner, **Pablo S. C. Santos**, and Maria G. Bicalho. Absence of strong linkage disequilibrium between odorant receptor alleles and the major histocompatibility complex. *Human Immunology*, 74(12):1619–1623, Dec. 2013.

[11] Eni P. Bompeixe, **Pablo S. C. Santos**, Rafael G. Vargas, Renate von Linsingen, Suelen. C. Zeck, Pryscilla F. Wowk, and Maria G. Bicalho. HLA class II polymorphisms and recurrent spontaneous abortion in a Southern Brazilian cohort. *International Journal of Immunogenetics*, 40(3):186–191, June 2013.

[12] **Pablo S. C. Santos**, Johannes Höhne, Fabiana Poerner, Maria G. Bicalho, Barbara Uchanska-Ziegler, and Andreas Ziegler. Does the new HapMap throw the baby out with the bath water? *European Journal of Human Genetics*, 19(7):733–734, July 2011.

[13] Andreas Ziegler, **Pablo S. C. Santos**, Thomas Kellermann, and Barbara Uchanska-Ziegler. Self/nonself perception, reproduction, and the extended MHC. *Self/Nonself*, 1(3):176–191, July 2010.

[14] **Pablo S. C. Santos**, Clineu J. S. Uehara, Andreas Ziegler, Barbara Uchanska-Ziegler, and Maria G. Bicalho. Variation and linkage disequilibrium within odorant receptor gene clusters linked to the human major histocompatibility complex. *Human Immunology*, 71(9):843–850, Sept. 2010.

[15] **Pablo S. C. Santos**, Thomas Kellermann, Barbara Uchanska-Ziegler, and Andreas Ziegler. Genomic architecture of MHC-linked odorant receptor gene repertoires among 16 vertebrate species. *Immunogenetics*, 62(9):569–584, Sept. 2010.

[16] Caroline. Sens-Abuázar, **Pablo S. C. Santos**, Maria G. Bicalho, M. L. Petzl-Erler, and V. Sperandio-Roxo. MHC microsatellites in a Southern Brazilian population. *International Journal of Immunogenetics*, 36(5):269–274, Oct. 2009.

[17] **Pablo S. C. Santos**, Johannes Höhne, Peter Schlattmann, Inke R. König, Andreas Ziegler, Barbara Uchanska-Ziegler, and Andreas Ziegler. Assessment of transmission distortion on chromosome 6p in healthy individuals using tagSNPs. *European Journal of Human Genetics*, 17(9):1182–1189, Sept. 2009.
In this article, my colleagues and I described an in silico approach to "scan" the human genome for areas that behave oddly concerning Mendelian inheriance. This made a methodological resource available for other researchers.

[18] **Pablo S. C. Santos**, George Füst, Zoltán Prohászka, Armin Volz, Roger Horton, Marcos Miretti, Chack-Yung Yu, Stephan Beck, Barbara Uchanska-Ziegler, and Andreas Ziegler. Association of smoking behavior with an odorant receptor allele telomeric to the human major histocompatibility complex. *Genetic Testing*, 12(4):481–486, Dec. 2008.

[19] **Pablo S. C. Santos**, Juliano A. Schinemann, Juarez Gabardo, and Maria G. Bicalho. New evidence that the MHC influences odor perception in humans: a study with 58 Southern Brazilian students. *Hormones and Behavior*, 47(4):384–388, Apr. 2005.

Teaching and Outreach

2023 **Principal Instructor**, *Statistics and Research Methods for Life Sciences ("R You With Me")*, University of Halle
This online course focuses on Statistics and Data Sciences. It is currently being developed and taught by me for a group of 20 life-science professionals in Germany, Ethiopia, Tanzania, South Africa and Nigeria.

2019 **Principal Instructor**, *Python for Biologists ("Python Time")*, University of Ulm
Developed and taught a new bioinformatics course focusing on Python for life-science researchers (using `Requests`, `Pandas`, `Seaborn`, `Matplotlib`, `Numpy` and others).

2019 **Principal Instructor**, *R for Ecologists ("The R Hour")*, University of Ulm
Developed and taught an introductory course to the R environment, taylored for researchers focusing in Ecology (using `ggplot2`, `effects`, `lme4` and others).

2019 **Instructor**, *Statistics for Biologists*, University of Ulm
Developed and taught the second half of an introductory course on statistical modeling (from ANOVAs to GLMMs) with R.

2014-2019 **Instructor and Developer**, *Master Module on Conservation Genomics*, University of Ulm
Development and teaching of an extensive masters course on Conservation Genomics methods such as NGS, Microbiome Analysis, MHC and Bioinformatic Pipelines.

2014-2019 **Teaching Assistant**, *Regular classes in Ornithology, Ecology and Population Genetics*, University of Ulm

2018 **Principal Instructor**, *Pre-Master's Course in Biology*, University of Ulm
Developed and taught a preparatory course for a heterogenic group of foreign students entering the Master's programme of the University of Ulm. The course had a strong focus on natural selection, genetics, statistics and bioinformatics, and was extremely well evaluated.

2012-2013 **Teaching Assistant**, *Regular classes in Immunogenetics, Genomics and Bioinformatics*, Postgraduate Program in Genetics of the Universidade Federal do Paraná, Curitiba, Brazil

2010-2012 **Teaching Assistant**, *Regular classes in Immunogenetics and Evolutionary Biology*, University of Potsdam

Software written by me

`ACACIA` Allele CAlling proCedure for Illumina Amplicon sequencing data: This Python-based workflow extracts allele information out of raw, paired-end Illumina sequencing data (or any groups of FASTQ files). It does so by first denoising, filtering and clustering sequences, and then by defining and calling alleles.
`https://gitlab.com/psc_santos/ACACIA`

Additional Relevant Skills and Interests

- Computer Languages/Methods/OS
  - `Linux`
  - `Bash`
  - `R`
  - `Python`
  - `git`
- Natural Languages (Full Proficiency)
  - German
  - English
  - Portuguese
  - Spanish
- Teaching
- Statistical Modeling
- Team Leadership
- Machine Learning
- Data Visualization
- Wikipedia Contributor
- Cycling, Tennis, Sewing and Art

Der Fachbereich Biologie, Chemie, Pharmazie der Freien Universität Berlin

verleiht dem Kandidaten der Naturwissenschaften

Herrn Lic. Pablo Sandro Carvalho Santos

geboren am 24.09.1978 in Brasília D.F., Brasilien

den akademischen Grad

eines Doktors der Naturwissenschaften (Dr. rer. nat.)

nachdem er in ordnungsgemäßem Promotionsverfahren
durch die Dissertation

**» Linkage Disequilibrium and Transmission Distortion
Affecting Human Chromosome 6p «**

Gutachter: Prof. Dr. A. Ziegler / Prof. Dr. H. Hofer

mit der Bewertung » magna cum laude «

und durch die Disputation am 09. 09.2010

mit der Bewertung » magna cum laude «

seine wissenschaftliche Befähigung erwiesen und als
Gesamturteil der Promotion

» magna cum laude «

erhalten hat.

Berlin, den 09. September 2010

Dekan
Univ.-Prof. Dr. Hartmut H. Hilger

104

NOTEN: summa cum laude (mit Auszeichnung) magna cum laude (sehr gut) cum laude (gut) rite (genügend)

Kommission zur Vergabe von Promotionsstipendien

Vergabekommission gemäß § 7 NaFöG

Die Kommission zur Vergabe von Promotionsstipendien spricht

Herrn
Pablo Sandro Carvalho Santos

ein Stipendium

nach dem Gesetz zur Förderung des wissenschaftlichen Nachwuchses

in der 43. Vergaberunde zu.

Berlin, am 03. November 2006

de Haan	M. Walter	A. Ziegler
Prof. Dr. Gerhard de Haan Freie Universität Berlin Vorsitzender	Prof. Dr. Monika Walter Technische Universität Berlin Stellv. Vorsitzende	Prof. Dr. Andreas Ziegler Humboldt-Universität zu Berlin Stellv. Vorsitzender

AKTION
GUTES
STUDIUM

Baden-Württemberg

Zertifikat
für
Hochschuldidaktik

Herrn

Dr. Pablo Santos

wird der Erwerb hochschuldidaktischer Kompetenzen
im Umfang von 200 Arbeitseinheiten bescheinigt.

Stuttgart, den 12. Oktober 2018

Theresia Bauer

Theresia Bauer MdL
Ministerin für Wissenschaft, Forschung
und Kunst Baden-Württemberg

Karin Amos

Professorin Dr. Karin Amos
Vorstandsvorsitzende des Hochschuldidaktikzentrums der Universitäten des
Landes Baden-Württemberg

UNIVERSITY OF
MICHIGAN

COURSE CERTIFICATE

08/25/2019

Pablo Sandro Carvalho Santos

has successfully completed

Applied Machine Learning in Python

an online non-credit course authorized by University of Michigan and offered through Coursera

Kevyn Collins-Thompson
Associate Professor
School of Information

Verify at coursera.org/verify/JG7TQX75VGP3

Coursera has confirmed the identity of this individual and their participation in the course.