The flux and flow of data: connecting large datasets with machine learning in a drug discovery envirionment

Bongers, B.J.

Part of this thesis

Bongers, B. J.; IJzerman, A. P.; Van Westen, G. J. P. Proteochemometrics – Recent Developments in Bioactivity and Selectivity Modeling. *Drug Discov Today Technol* **2019**, *32–33*, 89–98. https://doi.org/10.1016/J.DDTEC.2020.08.003.

Bongers, B. J. †; Sijben, H. J. †; Hartog, P. B. R.; Tarnovskiy, A.; IJzerman, A. P.; Heitman, L. H.; van Westen, G. J. P. Proteochemometric Modeling Identifies Chemically Diverse Norepinephrine Transporter Inhibitors. *J Chem Inf Model* **2023**, *63* (6), 1745–1755. https://doi.org/10.1021/acs.jcim.2c01645

Tuerkova, A. †; **Bongers, B. J.** †; Norinder, U.; Ungvári, O.; Székely, V.; Tarnovskiy, A.; Szakács, G.; Özvegy-Laczka, C.; Van Westen, G. J. P.; Zdrazil, B. Identifying Novel Inhibitors for Hepatic Organic Anion Transporting Polypeptides by Machine Learning-Based Virtual Screening. *J Chem Inf Model* **2022**, *62* (24), 6323–6335. https://doi.org/10.1021/acs.jcim.1c01460

Bongers, B. J. †; Gorostiola González, M. †; Wang, X.; van Vlijmen, H. W. T.; Jespers, W.; Gutiérrez-de-Terán, H.; Ye, K.; IJzerman, A. P.; Heitman, L. H.; van Westen, G. J. P. Pan-Cancer Functional Analysis of Somatic Mutations in G Protein-Coupled Receptors. Scientific Reports 2022 12:1 2022, 12 (1), 1–15. https://doi.org/10.1038/s41598-022-25323-x.

Béquignon, O. J. M. †; **Bongers, B. J**. †; Jespers, W.; IJzerman, A. P.; van der Water, B.; van Westen, G. J. P. Papyrus: A Large-Scale Curated Dataset Aimed at Bioactivity Predictions. *J Cheminform* **2023**, *15* (1), 1–11. https://doi.org/10.1186/s13321-022-00672-x

†These authors contributed equally

Other publications

Lenselink, E. B.; ten Dijke, N.; **Bongers, B.**; Papadatos, G.; van Vlijmen, H. W. T.; Kowalczyk, W.; IJzerman, A. P.; van Westen, G. J. P. Beyond the Hype: Deep Neural Networks Outperform Established Methods Using a ChEMBL Bioactivity Benchmark Set. *J Cheminform* **2017**, *9* (1), 45. https://doi.org/10.1186/s13321-017-0232-0.

Wang, X.; Jespers, W.; **Bongers, B. J.**; Habben Jansen, M. C. C.; Stangenberger, C. M.; Dilweg, M. A.; Gutiérrez-de-Terán, H.; IJzerman, A. P.; Heitman, L. H.; van Westen, G. J. P. Characterization of Cancer-Related Somatic Mutations in the Adenosine A2B Receptor. *Eur J Pharmacol* **2020**, *880*, 173126. https://doi.org/10.1016/J.EJPHAR.2020.173126.

Burggraaff, L.; Lenselink, E. B.; Jespers, W.; van Engelen, J.; **Bongers, B. J.**; Gorostiola González, M.; Liu, R.; Hoos, H. H.; van Vlijmen, H. W. T.; IJzerman, A. P.; van Westen, G. J. P. Successive Statistical and Structure-Based Modeling to Identify Chemically Novel Kinase Inhibitors. *J Chem Inf Model* **2020**, *60* (9), 4283–4295. https://doi.org/https://doi.org/10.1021/acs.jcim.9b01204.

Superti-Furga, G.; Lackner, D.; Wiedmer, T.; Ingles-Prieto, A.; Barbosa, B.; Girardi, E.; Goldmann, U.; Gürtl, B.; Klavins, K.; Klimek, C.; Lindinger, S.; Liñeiro-Retes, E.; Müller, A. C.; Onstein, S.; Redinger, G.;

Reil, D.; Sedlyarov, V.; Wolf, G.; Crawford, M.; Everley, R.; Hepworth, D.; Liu, S.; Noell, S.; Piotrowski, M.; Stanton, R.; Zhang, H.; Corallino, S.; Faedo, A.; Insidioso, M.; Maresca, G.; Redaelli, L.; Sassone, F.; Scarabottolo, L.; Stucchi, M.; Tarroni, P.; Tremolada, S.; Batoulis, H.; Becker, A.; Bender, E.; Chang, Y.-N.; Ehrmann, A.; Müller-Fahrnow, A.; Pütter, V.; Zindel, D.; Hamilton, B.; Lenter, M.; Santacruz, D.; Viollet, C.; Whitehurst, C.; Johnsson, K.; Leippe, P.; Baumgarten, B.; Chang, L.; Ibig, Y.; Pfeifer, M.; Reinhardt, J.; Schönbett, J.; Selzer, P.; Seuwen, K.; Bettembourg, C.; Biton, B.; Czech, J.; de Foucauld, H.; Didier, M.; Licher, T.; Mikol, V.; Pommereau, A.; Puech, F.; Yaligara, V.; Edwards, A.; **Bongers, B. J.**; Heitman, L. H.; IJzerman, A. P.; Sijben, H. J.; van Westen, G. J. P.; Grixti, J.; Kell, D. B.; Mughal, F.; Swainston, N.; Wright-Muelas, M.; Bohstedt, T.; Burgess-Brown, N.; Carpenter, L.; Dürr, K.; Hansen, J.; Scacioc, A.; Banci, G.; Colas, C.; Digles, D.; Ecker, G.; Füzi, B.; Gamsjäger, V.; Grandits, M.; Martini, R.; Troger, F.; Altermatt, P.; Doucerain, C.; Dürrenberger, F.; Manolova, V.; Steck, A.-L.; Sundström, H.; Wilhelm, M.; Steppan, C. M. The RESOLUTE Consortium: Unlocking SLC Transporters for Drug Discovery. *Nat. Rev. Drug Discov* **2020**, *19* (7), 429–430. https://doi.org/10.1038/d41573-020-00056-6.

Poster and oral communications

Event	Type
LACDR Spring Symposium (2018)	Poster
Awarded with the Poster Competition Prize	
FIGON Dutch Medicines Days (2018)	Poster
LACDR Spring Symposium (2019)	Poster
The Eigth Sheffield Chemoinformatics Conference (2019)	Poster
FIGON Dutch Medicines Days (2020)	Poster + Oral
Awarded with the Poster Competition Prize	
LACDR Spring Symposium (2021)	Poster
LACDR Spring Symposium (2022)	Poster
12th International Conference on Chemical Structures	Poster

Curriculum Vitae

Brandon Bongers was born on May 12th 1991 in Rotterdam, The Netherlands. He graduated from pre-university education at the Emmauscollege in Rotterdam in 2008, after which he started his degree in computer science at the TU Delft. After 1.5 years however, he swapped to the bachelor Bio-Pharmaceutical Sciences in February 2010, following a mixed-up schedule due to starting point. He completed his bachelor degree by finishing an internship under the supervision of dr. Bart Lenselink investigating homology models for the adenosine receptors. This was performed in the division of Medicinal Chemistry of the Leiden Academic Centre for Drug Research (LACDR). After a small gap, he started his master studies in Bio-Pharmaceutical Sciences, performing two internships at the LACDR and Leids Universitair Medisch Centrum (LUMC). The first internship was once more under the supervision of dr. Bart Lenselink, investigating machine learning benchmarks for bioactivity and selectivity modeling. The second internship was under the supervision of dr. Erik Schultes in the biosemantics group, performing a broad comparison and analysis on concept profiles that describe drug-disease relations, as well as making data findable, accessible, interoperable and reusable (FAIR).

After a successful completion of his master studies, Brandon started his PhD in 2017 at the division of Drug Discovery and Safety of the LACDR. Under the supervision of prof. dr. Gerard van Westen and prof. dr. Ad IJzerman, he sought to become an expert at statistical modeling in the area of medicinal chemistry and computational drug discovery. Knowledge about machine learning, and in particular proteochemometrics, accumulated in contributions to several projects. This included the investigation of mutations underneath a VENI project awarded, as well as contributing to the RESOLUTE consortium efforts to understand solute carriers. Brandon presented his work frequently at symposiums and conferences, including several appearances at the LACDR Spring Symposiums as well as the FIGON Dutch Medicines Days, awarded a poster prize twice in 2018 and 2020 respectively. He also presented his work at larger conferences, including the Eight Joint Sheffield Chemoinformatics Conference and the 12th International Conference on Chemical Structures (ICCS).

His interest for supervising master students during his PhD eventually transformed to a love for teaching and management, pursuing an extra year to fully focus on supervising multiple master students as well as taking courses to achieve a partial teaching license. With successfully navigating a difficult year during a global pandemic, he pursued steps to intensify the combination of research and teaching.

Brandon is currently a post-doctoral researcher at the LACDR, working with prof. dr. Gerard van Westen and dr. Willem Jespers to elevate the teaching capabilities of the computational division, as well as providing guidance to bachelor, master and occasionally PhD students. On top of that, he is looking to pursue certain research lines as well that are mainly focused on his strengths as a strong statistical modeler and jack-of-all-trades.

Acknowledgements

I began my PhD on the simple basis of 'why not?'. Being invited to come work at such a position still feels like an honour to this day and given the experiences, both past and present, I'd like to think I did the right thing. I am aware of my disposition to be on my own, a social battery that is strong (i like to think) but rapidly runs out of juice. Yet despite that, there are those around me that have accepted me for who I am, and supported me along the way. While this thesis is ultimately mine to present and defend, it is still a sum of many parts, and I would like to take this section to dedicate some words to anyone who has contributed to this PhD thesis in any way, whether to the scientific portion, or simply my mind state.

First, thank you Gerard and Ad for their guidance and support during this proces. Feedback was always readily provided, and in other times I could always lean in for some words of advice or to let some of my thoughts free. I am grateful for the positive environment that made me feel accepted, the patience you had with my sometimes inability to move forward, and the stimulation I needed for my ventures into research and teaching.

I would like to thank all collaborators that contributed to research in and outside of this thesis, in particular to Huub and Alzbeta for joining me in a vision of combining computational with experimental research. I am proud of the work that we put out, especially given the promising results.

To my myriad of students that I have supervised over during my PhD, thank you for your contributions. All the work that you have done have presented me with new insights, regardless if they ended up in written form in a journal. A special shout-out to my master students that I personally supervised: Marc, Peter, Meike, Kien, Yara and Tim. No project was the same, which made it that much more fun.

To all my colleagues, past or present, I would like to thank you for being a bunch of friendly and intelligent people. I cannot recall a single time where there was no laughter whenever we had our coffee breaks. Willem, to think that we both started our BOO at Bart, sitting right next to each other in the cramped office space, all the way to working closely together currently. To Bart and Lindsey, it was always good to hear your insights and I always looked up to both of you on how to do a PhD right. To my computational colleagues, seeing the group grow into the juggernaut that it has been now fills me with inspiration, to see so many bright minds working together, and it was especially nice to see former master students Linde and Remco working on their PhD now in the group. To Xuesong, Majlen, Jara, Rongfang, and likely many more, I enjoyed our talks during breaks or meetings, and felt there was no real barrier between us, whether it was about science or social.

However, I cannot understate the amount of support and love I shared with Marina and Olivier. Dubbing ourselves the Three Datateers, I felt like we are all a major part of this journey, and it has been my pleasure to work together with you on the scientific contributions found in this thesis and without you two, this thesis would not have been here, in any way you may interpret that. Marina, you've made me open up and pulled me into certain things that I would not have done solo. Olivier, you have been my office mate for the longest time, and I still remember us talking about research,

listening to music, discussing students, or panicking over certain abstract submissions.

Thank you to my friends and family who have supported me along the way. Special thanks to Alex, Michiel, Mark, Martin, Max, Nick, Sjoerd, Ferry, Juni, Nikki and Berry for hanging out with me, playing plenty of games over the years or supporting my growth over the years. Thank you Celeste for your support and in particular your help with the design of this thesis. Last but absolutely not least, thank you mom and dad for allowing me to find myself, multiple times even, and giving me all the unconditional love that I needed. You have always supported me whatever direction I went, and whether it was up or down. Thank you all.