Virginie RONDEAU
Director of Research INSERM, Ph.D.
INSERM CR1219 (Biostatistic team) - ISPED
146, Rue Léo Saignat
33076 Bordeaux Cedex, France

October 22nd, 2018

Dear Dr. Virginie RONDEAU,

I'm currently working at the CNRS UMR8199 - "Integrated Genomics and Metabolic Diseases Modelling" as the head of the biostatistic team.
I started working as a "Study Engineer" in September 2012. In October 2014, I was given the opportunity to start a Ph.D. thesis on joint models applied to the genetic field.
This was an opportunity to get back to the study of longitudinal data in which I am interested in since my internship in 2011 ("Development of a non-parametric (on time) clustering algorithm for longitudinal data").

During my six years at the CNRS UMR8199, I developed R packages [Yengo et al., 2016], web applications using the R package SHINY [Ndiaye et al., 2017, Verbanck et al., 2017] and wrote reproducible scripts from the data to the scientific article, using R and RMARKDOWN [Canouil et al., 2018].
Recently, I wrote a complete RMARKDOWN analysis plan for the European project "RHAPSODY WP3 - Pre-Diabetes progression". The analysis plan combined with the standard format (CDISC) used for the phenotypes was included in a DOCKER image. This allows to run the whole analysis on any cohorts' data within the same environment (Operating System, R version and R packages version).

Over the years, I improved my computing skills in different programming languages particularly R, with for example S4 object, RMARKDOWN and SHINY. This also includes code optimisation, version control using GIT and more recently code portability and environment control using Docker.

Working at INSERM CR1219 could be a great opportunity for me to continue to work on joint models. In particular for the high expertise showed by INSERM CR1219 in longitudinal and survival data analysis, with for instance, the R package LCMM and the work of Dr. Hélène JACMIN-GADDA et al. With the knowledge acquired during my Ph.D. on joint models and my computing skills, I believe joining your team could benefit both sides.
Thank you for considering my application. I would be happy to provide any additional information you might need.

References:

Ghislain ROCHELEAU Assistant Professor in Genetics and Genomic Sciences
Location Mount Sinai, New-York, United States
Mail ghislain.rocheleau@mssm.edu

Loïc YENGO-DEMBOU Senior Research Officer
Location Institute for Molecular Bioscience, Brisbane, Australia
Mail l.yengodimbou@imb.uq.edu.au
Phone +61 7 334 62095

Sincerely yours,

Mickaël CANOUIL

Attached: Curriculum Vitæ

Mickaël CANOUIL

Biostatistician, Ph.D.

The best thing about being a statistician is that you get to play in everyone's backyard
– John Tukey

Experience

Jan. 2016 – Nov. 2018 **Head of the Biostatistic Team (CNRS / Pasteur Institute Lille)**, *CNRS UMR8199 - Integrated Genomics and Metabolic Diseases Modeling*, Lille, Activities: Genome-wide association studies, experimental design, -omics data analysis, methodological developments, consortium work package lead analyst, team management.
Headed by Pr. P. Froguel.

Sept. 2012 – Dec. 2015 **Biostatistician (CNRS)**, *CNRS UMR8199 - Integrated Genomics and Metabolic Diseases Modeling*, Lille, Activities: Genome-wide association studies, experimental design, -omics data analysis, methodological developments.
Headed by Pr. P. Froguel.

Nov. 2011 – Dec. 2011 **Biostatistician**, *The French Institute of Science and Technology for Transport, Development and Networks (IFSTTAR) - UMRESTTE - UMR T 9405*, Bron, Activities: Analysis of a mobility and accident study in secondary-school pupils.
Supervised by Dr. M. Haddak.

Jan. 2011 – Jun. 2011 **Biostatistician (Internship)**, *Hospices Civiles de Lyon - Biostatistics Unit - CNRS UMR5558*, Lyon, Activities: Development of a non-parametric (on time) clustering algorithm for longitudinal data.
Supervised by Pr. R. Ecochard et Dr. C. Genolini.

Mar. 2010 – Jun. 2010 **Biostatistician (Internship)**, *Laboratory of Biometry and Evolutionary Biology (LBBE) - CNRS UMR5558*, Lyon, Activities: Mathematical modelling of nosocomial rotavirus infections in pediatric ward.
Supervised by Dr. C. Kribs-Zaleta.

138 Rue Pierre Mauroy – 59000 Lille, France
+33 (0)6 58 58 79 57 • mickael.canouil@cnrs.fr
http://mickael.canouil.fr • https://github.com/mcanouil

Main Activities In The Last Job Position

Data Analysis I've been analysing omics (SNP, CpG, expression, metabolites, etc.) data within numerous projects related to metabolic diseases, such as Type 2 Diabetes. These projects involve collaborations with national and international consortia like CKDgen, CHARGE, IMIDIA, DIRECT or more recently RHAPSODY. My contribution covers Genome-Wide Association Studies (GWAS), differential methylation/expression analyses, metabolomics analyses, rare variants analyses (clustering approach), disease progression modelling, genetic epidemiology and meta-analyses.

Team Management I've been managing a team of three junior statisticians. My role is to provide guidance regarding choices of statistical methodologies for analysing the data and code optimisation for implementing these methodologies in large scale omics data.

Research My main research interests are **mixed models**, more recently extended to joint models. **Joint model**, especially the **joint likelihood approach** implemented in the R package JM was studied in the context of Type 2 Diabetes incidence and fasting glucose progression (associated by diagnosis definition) using SNPs as biomarkers of interest [Canouil et al., 2018].

Education

Oct. 2014 – Sept. 2017 **Doctor of Philosophy (Ph.D.) in BioStatistics**, *University of Lille 2*, Lille, "Development and Application of Statistical Methods for Multi-Omics Studies in Type 2 Diabetes: Beyond the Genome-Wide Association Studies Era".
Supervised by Pr. P. Froguel and Dr. G. Rocheleau

Sept. 2009 – Jul. 2011 **Master's Degree in Biostatistics, Bioinformatics and Genomics**, *University Claude Bernard Lyon 1*, Lyon, Specialised in Biostatistics.

Sept. 2006 – Jul. 2009 **Bachelor's Degree in Biology**, *University Claude Bernard Lyon 1*, Lyon, Specialised in Mathematics and Informatics for Biology.

Computer Skills

Basic C/C++, SQL, NoSQL

Intermediate Julia, Lua, Perl, Python, SAS

Advanced R (Shiny, Rmarkdown, S4, etc.), HTML, CSS, LaTeX, markdown

Environment Docker, UNIX, Windows

Languages

French Native

English Fluent / Full Professional Proficiency

Spanish Elementary proficiency

Awards

2015 **Funding Allocation SFD-Lilly**, *French speaking Diabetes Society (SFD)*, Bordeaux, Detection of new genomic variants associated with fasting blood glucose and incidence of type 2 diabetes simultaneously.

138 Rue Pierre Mauroy – 59000 Lille, France
+33 (0)6 58 58 79 57 • *mickael.canouil@cnrs.fr*
http://mickael.canouil.fr • *https://github.com/mcanouil*

R Packages

snpEnrichment **R package implementing a method for calculating an enrichment statistic of a set of SNP within a GWA signal using minor allele frequency and linkage disequilibrium.**
Mickaël Canouil and Loïc Yengo (2013)
https://cran.r-project.org/package=snpEnrichment

clere **R package implementing the CLERE methodology.**
Loïc Yengo, Julien Jacques, Christophe Biernacki and Mickaël Canouil (2014)
https://cran.r-project.org/package=clere

Communications

Oral Presentations

- **Julia for Intensive Scientific Computing** (*half-day workshop available on GitHub, in French*)
  Mickaël Canouil
  "Journées nationales du DEVeloppement logiciel" - JDEV, Bordeaux, France (2015)
- **Longitudinal Genetic Modelling: Revisiting Associations of SNPs Associated with Blood Fasting Glucose in Normoglycemic Individuals**
  Mickaël Canouil, Ghislain Rocheleau, Loïc Yengo and Philippe Froguel
  Statistical Methods for Post Genomic Data - SMPGD, Lille, France (2016)
- **R and Databases** (*two-days training available on GitHub, in French*)
  Mickaël Canouil
  URFIST - University of Bordeaux, Bordeaux, France (2018)
- **Jointly Modelling SNPs with Survival & Longitudinal Trait?** (*available on GitHub, in French*)
  Mickaël Canouil
  Thematic Day "Statistic & Genomic" of the "Réseau Interdisciplinaire autour de la Statistique" - RIS, Paris, France (2018)

Poster Presentations

- **Application of Joint Models in Genetic Association Studies**
  Ghislain Rocheleau, Mickaël Canouil, Loïc Yengo and Philippe Froguel
  International Genetic Epidemiology Society - IGES, Baltimore, United-States (2015)
- **Single Nucleotide Polymorphisms Associated With Fasting Blood Glucose Trajectory And Type 2 Diabetes Incidence: A Joint Modelling Approach**
  Mickaël Canouil, Philippe Froguel and Ghislain Rocheleau
  International Genetic Epidemiology Society - IGES, Toronto, Canada (2016)
- **Single Nucleotide Polymorphisms Associated With Fasting Blood Glucose Trajectory And Type 2 Diabetes Incidence: A Joint Modelling Approach**
  Mickaël Canouil, Philippe Froguel and Ghislain Rocheleau
  4th Symposium European Genomic Institute for Diabetes (E.g.i.d), Lille, France (2016)
- **"Variants Génétiques Associés à la Trajectoire de la Glycémie à Jeun et à l'Incidence du Diabète de Type 2: Une Approche par Modèle Joint"** (CA-075)
  Mickaël Canouil, Philippe Froguel and Ghislain Rocheleau
  Annual Congress of "Société Francophone du Diabète" (SFD), Lille, France (2017)

Publications

[Abderrahmani et al., 2018] Abderrahmani, A., Yengo, L., Caiazzo, R., Canouil, M., Cauchi, S., Raverdy, V., Plaisance, V., Pawlowski, V., Lobbens, S., et al. (2018). Increased Hepatic PDGF-AA Signaling Mediates Liver Insulin Resistance in Obesity Associated Type 2 Diabetes.

[Baumeier et al., 2017] Baumeier, C., Saussenthaler, S., Kammel, A., Jähnert, M., Schlüter, L., Hesse, D., Canouil, M., Lobbens, S., Caiazzo, R., et al. (2017). Hepatic DPP4 DNA Methylation Associates With Fatty Liver. 66(1):25–35.

[Bonnefond et al., 2017] Bonnefond, A., Yengo, L., Dechaume, A., Canouil, M., Castelain, M., Roger, E., Allegaert, F., Caiazzo, R., Raverdy, V., et al. (2017). Relationship Between Salivary/Pancreatic Amylase and Body Mass Index: A Systems Biology Approach. 15(1):37.

[Canouil et al., 2018] Canouil, M., Balkau, B., Roussel, R., Froguel, P., and Rocheleau, G. (2018). Jointly Modelling Single Nucleotide Polymorphisms With Longitudinal and Time-to-Event Trait: An Application to Type 2 Diabetes and Fasting Plasma Glucose. 9.

[Carrat et al., 2017] Carrat, G. R., Hu, M., Nguyen-Tu, M.-S., Chabosseau, P., Gaulton, K. J., van de Bunt, M., Siddiq, A., Falchi, M., Thurner, M., et al. (2017). Decreased STARD10 Expression Is Associated with Defective Insulin Secretion in Humans and Mice. 100(2):238–256.

[Feitosa et al., 2018] Feitosa, M. F., Kraja, A. T., Chasman, D. I., Sung, Y. J., Winkler, T. W., Ntalla, I., Guo, X., Franceschini, N., Cheng, C.-Y., et al. (2018). Novel Genetic Associations for Blood Pressure Identified Via Gene-Alcohol Interaction In Up to 570K Individuals Across Multiple Ancestries. 13(6):e0198166.

[Karamitri et al., 2018] Karamitri, A., Plouffe, B., Bonnefond, A., Chen, M., Gallion, J., Guillaume, J.-L., Hegron, A., Boissel, M., Canouil, M., et al. (2018). Type 2 Diabetes–Associated Variants of The MT2 Melatonin Receptor Affect Distinct Modes of Signaling. 11(545):eaan6622.

[Mahajan et al., 2018a] Mahajan, A., Taliun, D., Thurner, M., Robertson, N. R., Torres, J. M., Rayner, N. W., Payne, A. J., Steinthorsdottir, V., Scott, R. A., et al. (2018a). Fine-Mapping Type 2 Diabetes Loci to Single-Variant Resolution Using High-Density Imputation and Islet-Specific Epigenome Maps. page 1.

[Mahajan et al., 2018b] Mahajan, A., Wessel, J., Willems, S. M., Zhao, W., Robertson, N. R., Chu, A. Y., Gan, W., Kitajima, H., Taliun, D., et al. (2018b). Refining the Accuracy of Validated Target Identification Through Coding Variant Fine-Mapping in Type 2 Diabetes. 50(4):559–571.

[Ndiaye et al., 2017] Ndiaye, F. K., Ortalli, A., Canouil, M., Huyvaert, M., Salazar-Cardozo, C., Lecoeur, C., Verbanck, M., Pawlowski, V., Boutry, R., et al. (2017). Expression and Functional Assessment of Candidate Type 2 Diabetes Susceptibility Genes Identify Four New Genes Contributing to Human Insulin Secretion. 6(6):459–470.

[Sung et al., 2018] Sung, Y. J., Winkler, T. W., de Las Fuentes, L., Bentley, A. R., Brown, M. R., Kraja, A. T., Schwander, K., Ntalla, I., Guo, X., et al. (2018). A Large-Scale Multi-Ancestry Genome-wide Study Accounting for Smoking Behavior Identifies Multiple Significant Loci for Blood Pressure. 102(3):375–400.

[Verbanck et al., 2017] Verbanck, M., Canouil, M., Leloire, A., Dhennin, V., Coumoul, X., Yengo, L., Froguel, P., and Poulain-Godefroy, O. (2017). Low-dose Exposure to Bisphenols A, F and S of Human Primary Adipocyte Impacts Coding and Non-coding RNA Profiles. 12(6):e0179583.

[Yengo et al., 2016] Yengo, L., Jacques, J., Biernacki, C., and Canouil, M. (2016). Variable Clustering in High-Dimensional Linear Regression: The R Package clere. 8(1):92–106. bibtex: RJ-2016-006.

As co-first author in the following articles:
- *[Ndiaye et al., 2017]*
- *[Verbanck et al., 2017]*
- *[Abderrahmani et al., 2018]*

Note: authors' lists have been shortened with "et al." and thus, might not contain my lastname.

Interests

- Movies / TV-Shows / Japanese Animation
- Data visualisation ("IMDbRating" on GitHub)
- Board games
- Hiking
- Archery
- A little bit of reading (fanstastic novels, e.g., Robin Hobb)
- Escape games