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 – Present
Head of the Biostatistic Team (Institut Pasteur de 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.

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.

R Packages

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

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.**
<u>Mickaël Canouil</u> and Loïc Yengo (2013)
https://cran.r-project.org/package=snpEnrichment

NACHO   **NACHO: NanoString Quality Control Dashboard.**
<u>Mickaël Canouil</u>, Gerard Bouland and Roderick Slieker (2019)
https://cran.r-project.org/package=NACHO

CARoT   **Centralised and Automated RepOrting Tools.**
<u>Mickaël Canouil</u>, Mathilde Boissel and Lijiao Ning (2019)
https://github.com/mcanouil/CARoT

Communications

Oral Presentations

○ **Julia for Intensive Scientific Computing** (*half-day workshop available on GitHub, in French*)
<u>Mickaël Canouil</u>
"Journées nationales du DEVeloppement logiciel" - JDEV, Bordeaux, France (2015)

○ **Longitudinal Genetic Modelling: Revisiting Associations of SNPs Associated with Blood Fasting Glucose in Normoglycemic Individuals**
<u>Mickaël Canouil</u>, 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*)
<u>Mickaël Canouil</u>
URFIST - University of Bordeaux, Bordeaux, France (2018)

○ **Jointly Modelling SNPs with Survival & Longitudinal Trait?** (*available on GitHub, in French*)
<u>Mickaël Canouil</u>
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, <u>Mickaël Canouil</u>, 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**
<u>Mickaël Canouil</u>, 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**
<u>Mickaël Canouil</u>, 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

[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. *The R Journal*, 8(1):92–106. bibtex: RJ-2016-006.

[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. *Diabetes*, 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. *BMC medicine*, 15(1):37.

[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. *American Journal of Human Genetics*, 100(2):238–256.

[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. *Molecular Metabolism*, 6(6):459–470.

[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. *PloS One*, 12(6):e0179583.

[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. *Diabetes*.

[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. *Frontiers in Genetics*, 9.

[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. *PLOS ONE*, 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. *Science 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. *Nature Genetics*, 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. *Nature Genetics*, 50(4):559–571.

[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. *American Journal of Human Genetics*, 102(3):375–400.

[Bentley et al., 2019] Bentley, A. R., Sung, Y. J., Brown, M. R., Winkler, T. W., Kraja, A. T., Ntalla, I., Schwander, K., Chasman, D. I., Lim, E., et al. (2019). Multi-ancestry genome-wide gene–smoking interaction study of 387,272 individuals identifies new loci associated with serum lipids. *Nature Genetics*, 51(4):636.

[de Vries et al., 2019] de Vries, P. S., Brown, M. R., Bentley, A. R., Sung, Y. J., Winkler, T. W., Ntalla, I., Schwander, K., Kraja, A. T., Guo, X., et al. (2019). Multi-Ancestry Genome-Wide Association Study of Lipid Levels Incorporating Gene-Alcohol Interactions. *American Journal of Epidemiology*.

[Khamis et al., 2019] Khamis, A., Canouil, M., Siddiq, A., Crouch, H., Falchi, M., Bulow, M. v., Ehehalt, F., Marselli, L., Distler, M., et al. (2019). Laser capture microdissection of human pancreatic islets reveals novel eQTLs associated with type 2 diabetes. *Molecular Metabolism*.

[Lifelines Cohort Study et al., 2019] Lifelines Cohort Study, Kilpeläinen, T. O., Bentley, A. R., Noordam, R., Sung, Y. J., Schwander, K., Winkler, T. W., Jakupović, H., Chasman, D. I., et al. (2019). Multi-ancestry study of blood lipid levels identifies four loci interacting with physical activity. *Nature Communications*, 10(1).

[Sung et al., 2019] Sung, Y. J., de las Fuentes, L., Winkler, T. W., Chasman, D. I., Bentley, A. R., Kraja, A. T., Ntalla, I., Warren, H. R., Guo, X., et al. (2019). A multi-ancestry genome-wide study incorporating gene-smoking interactions identifies multiple new loci for pulse pressure and mean arterial pressure. *Human Molecular Genetics*.

As co-first author in the following articles:
o *[Ndiaye et al., 2017]*
o *[Verbanck et al., 2017]*
o *[Abderrahmani et al., 2018]*
o *[Khamis et al., 2019]*

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

―――――― Interests

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

○ Escape games