Andrea **Rau**

Research Scientist • Chargée de Recherche

INRAE • French National Research Institute for Agriculture, Food and Environment

+33 1 34 65 22 82 | andrea.rau@inrae.fr | andrea-rau.com | 0000-0001-6469-488X | 9__bzfgAAAAJ | andreamrau | andreamrau

About me

My research is focused on developing and writing software for sound statistical methods for genomic and transcriptomic data analysis, including differential expression analyses, co-expression analyses, network inference, and integrative multi-omic analyses.

I currently belong to two INRAE labs:

- **Animal Genetics and Integrative Biology** (GABI) research unit (Jouy en Josas, France) in the Genomics, Biodiversity, Bioinformatics, Statistics (GiBBS) team
- Cross-border **BioEcoAgro** research unit (Estrées-Mons, France)

Keywords: Analysis of high-throughput sequencing data, mixture models, supervised classification methods, multi-omic integration, gene regulatory networks

Languages: English (maternal), French (fluent)

Education

HDR in Applied Mathematics — *2017*

Université d'Évry-Val-d'Essonne — *Évry, France*

- Title: "Statistical methods and software for the analysis of transcriptomic data"
  Note: An HDR is the French accreditation to supervise research and represents the highest French academic qualification level based on independent scholarship. It is reviewed by and defended before an academic committee.

PhD in Statistics — *2007-2010*

Purdue University — *West Lafayette, Indiana, USA*

- Title: Reverse engineering gene regulatory networks using genomic time-course data
  Advisors: Rebecca W. Doerge, Jean-Louis Foulley, and Florence Jaffrézic

MS in Applied Statistics — *2005-2007*

Purdue University — *West Lafayette, Indiana, USA*

- Internship: Time series modeling of advertising interventions on pharmacy sales (Walgreens; Deerfield, Illinois, USA)

BA in French and Mathematics (concentration in Statistics) — *2001-2005*

Saint Olaf College — *Northfield, Minnesota, USA*

- Internship: pharmacokinetic analysis of Phase I clinical trial data using a limited-sample model (Mayo Clinic; Rochester, Minnesota, USA)

Work experience

Research Scientist (Chargée de Recherche) — *2011-present*

INRAE — *Jouy-en-Josas, France*

Adjunct Assistant Professor — *2019*

Medical College of Wisconsin (4 months) — *Milwaukee, Wisconsin, USA*

AgreenSkills+ Visiting Scholar — *2017-2019*

Zilber School of Public Health, University of Wisconsin-Milwaukee (20 months) — *Milwaukee, Wisconsin, USA*

Visiting Scholar — *2016*

Zilber School of Public Health, University of Wisconsin-Milwaukee (6 weeks) — *Milwaukee, Wisconsin, USA*

Adjunct Assistant Professor — *2012-2017*

Ensai — *Rennes, France*

Post-doctoral researcher — *2010-2011*

Inria - Île-de-France — *Orsay, France*

Research assistant *2008-2010*

Department of Statistics, Purdue University (R. W. Doerge *West Lafayette, Indiana, USA*

Consultant in the Statistical Consulting Service *2007*

Department of Statistics, Purdue University *West Lafayette, Indiana, USA*

Awards

- → Graduate Women in Science Programs travel award *2010*
- → Student travel award, Conference on Applied Statistics in Agriculture at Kansas State University *2010*
- → Honorable mention, Gertrude M. Cox Scholarship *2009*
- → A.H. Ismail Interdisciplinary Program doctoral research travel award *2009*

Professional organizations

SFdS	Société Française de la Statistique	*2011-present*
ASA	American Statistical Association	*2005-present*

Dissertations, books & book chapters

1. Duranthon, V., Araújo, S., Palma, M., **Rau, A.**, Matzapetakis, M., and Almeida, A. (2021) Rabbit research in the post-genomic era: transcriptome, proteome, and metabolome analysis. *In: The Genetics and Genomics of the Rabbit.*, Ed. L. Fontanesi.
2. **Rau, A.** (2017) Statistical methods and software for the analysis of transcriptomic data. *HDR thesis*, Université d'Évry Val-d'Essonne.
3. Martin-Magniette, M.-L., Maugis-Rabusseau, C. and **Rau, A.** (2017) Clustering of co-expressed genes. *In: Model Choice and Model Aggregation*, Ed. F. Bertrand, J.-J. Droesbeke, G. Saporta, C. Thomas-Agnan.
4. Albert, I., Ancelet, S., David, O., Denis, J.-B., Makowski, D., Parent, É., **Rau, A.**, and Soubeyrand, S. (2015) Initiation à la statistique bayésienne : Bases théoriques et applications en alimentation, environnement, épidémiologie et génétique. *Éditions Ellipses*, collection références sciences.
5. **Rau, A.** (2010) Reverse engineering gene networks using genomic time-course data.. *PhD thesis*, Purdue University.

Peer-reviewed publications

1. **Rau, A.** (2021) Cooking up knowledge from big data using data science. *Frontiers in Young Minds*, 9:632923. `https://dx.doi.org/10.3389/frym.2021.632923`
2. Mollandin, F., **Rau, A.**, and Croiseau, P. (2021) An evaluation of the predictive performance and mapping power of the BayesR model for genomic prediction. *G3*, jkab225. `https://dx.doi.org/10.1093/g3journal/jkab225/6317672`
3. Sellem, E., Marthey, S., **Rau, A.**, Jouneau, L., Bonnet, A., Le Danvic, C., Kiefer, H., Jammes, H., and Schibler, L. (2021) Dynamics of cattle sperm sncRNAs during maturation, from testis to ejaculated sperm. *Epigenetics and Chromatin*, 14:24. `https://dx.doi.org/10.1186/s13072-021-00397-5`
4. Mach, N., Moroldo, M., **Rau, A.**, Lecardonnel, J., Le Moyec, L., Robert, C., and Barrey, E. (2021) Understanding the holobiont: crosstalk between gut microbiota and mitochondria during endurance. *Frontiers Molecular Biosciences*, 8:656204. `https://dx.doi.org/10.3389/fmolb.2021.656204`
5. Devogel, N., Auer, P. L., Manansala, R., **Rau, A.**, and Wang, T. (2020) A unified linear mixed model for familial relatedness and population structure in genetic association studies. *Genetic Epidemiology*, 45(3): 305-315. `https://dx.doi.org/10.1002/gepi.22371`
6. Cho, Y., **Rau, A.**, Reiner, A., Auer, P. L. (2020) Mendelian randomization analysis with survival outcomes. *Genetic Epidemiology*, 45(1): 16-23. `https://dx.doi.org/10.1002/gepi.22354`
7. Sellem, E., Marthey, S., **Rau, A.**, Jouneau, L., Bonnet, A., Perrier, J.-P., Fritz, S., Le Danvic, C. Boussaha, M., Kiefer, H., Jammes, H., Schiblier, L. (2020) A comprehensive overview of bull sperm-borne small non-coding RNAs and their diversity in six breeds. *Epigenetics and Chromatin*, 13:19. `https://dx.doi.org/10.1186/s13072-020-00340-0`
8. **Rau, A.**, Manansala, R., Flister, M. J., Rui, H., Jaffrézic, F., Laloë, D., and Auer, P. L. (2020) Individualized multi-omic pathway deviation scores using multiple factor analysis. *Biostatistics*, kxaa029. `https://dx.doi.org/10.1101/827022`

9. Godichon-Baggioni, A., Maugis-Rabusseau, C. and **Rau, A.** (2020) Multi-view cluster aggregation and splitting, with an application to multi-omic breast cancer data. *Annals of Applied Statistics*, 14:2, 752-767.
10. Jehl, F., Désert, C., Klopp, C., Brenet, M., **Rau, A.**, Leroux, S., Boutin, M., Muret, K., Blum, Y., Esquerré, D., Gourichon, D., Burlot, T., Collin, A., Pitel, F., Benani, A., Zerjal, T., Lagarrigue, S. (2019) Chicken adaptive response to low energy diet: main role of the hypothalamic lipid metabolism revealed by a phenotypic and multi-tissue transcriptomic approach. *BMC Genomics*, 20. https://dx.doi.org/10.1186/s12864-019-6384-8
11. Foissac, S., Djebali, S., Munyard, K., Villa-Vialaneix, N., **Rau, A.**, Muret, K., Esquerre, D., Zytnicki, M., Derrien, T., Bardou, P., Blanc, F., Cabau, C., Crisci, E., Dhorne-Pollet, S., Drouet, F., Gonzales, I., Goubil, A., Lacroix-Lamande, S., Laurent, F., Marthey, S., Marti-Marimon, M., Momal-Leisenring, R., Mompart, F., Quere, P., Robelin, D., San Cristobal, M., Tosser-Klopp, G., Vincent-Naulleau, S., Fabre, S., Pinard-Van der Laan, M.-H., Klopp, C., Tixier-Boichard, M., Acloque, H., Lagarrigue, S., Giuffra, E. (2019) Multi-species annotation of transcriptome and chromatin structure in domesticated animals. *BMC Biology*, 18:48.
12. Dhara, S., **Rau, A.**, Flister, M., Recka, N., Laiosa, M., Auer, P., and Udvadia, A. (2019) Cellular reprogramming for successful CNS axon regeneration is driven by a temporally changing cast of transcription factors. *Scientific Reports*, 9:14198. https://dx.doi.org/10.1038/s41598-019-50485-6
13. **Rau, A.**, Dhara, S., Udvadia, A., and Auer, P. (2019) Regeneration Rosetta: An interactive web application to explore regeneration-associated gene expression and chromatin accessibility. *G3: Genes|Genomes|Genetics*, 9(12): 3953-3959. https://dx.doi.org/10.1534/g3.119.400729
14. Plasterer, C., Tsaih, S.-W., Lemke, A., Schilling, R., Dwinell, M., **Rau, A.**, Auer, P., Rui, H., Flister, M.J. (2019) Identification of a rat mammary tumor risk locus that is syntenic with the commonly amplified 8q12.1 and 8q22.1 regions in human breast cancer patients. *G3: Genes|Genomes|Genetics*, 9(5): 1739-1743. https://dx.doi.org/10.1534/g3.118.200873
15. Ramayo-Caldas, Y., Zingaretti, L., Bernard, A., Estellé, J. Popova, M., Pons, N., Bellot, P., Mach, N., **Rau, A.**, Roume, H., Perez-Encisco, M., Faverdin, P., Edouard, N., Dusko, S., Morgavi, D.P. and Renand, G. (2019) Identification of rumen microbial biomarkers linked to methane emission in Holstein dairy cows. *Journal of Animal Breeding and Genetics*, 137:49-59. https://dx.doi.org/10.1111/jbg.12427
16. **Rau, A.**, Flister, M. J., Rui, H. and Livermore Auer, P. (2019) Exploring drivers of gene expression in The Cancer Genome Atlas. *Bioinformatics*, 35(1): 62-68. https://dx.doi.org/10.1093/bioinformatics/bty551
17. Godichon-Baggioni, A., Maugis-Rabusseau, C. and **Rau, A.** (2018) Clustering transformed compositional data using K-means, with applications in gene expression and bicycle sharing system data. *Journal of Applied Statistics*, 46(1):47-65.
18. **Rau, A.** and Maugis-Rabusseau, C. (2018) Transformation and model choice for RNA-seq co-expression analysis. *Briefings in Bioinformatics*, bbw128. https://dx.doi.org/10.1093/bib/bbw128
19. Verrier, E., Genet, C., Laloë, D., Jaffrézic, J., **Rau, A.**, Esquerre, D., Dechamp, N., Ciobataru, C., Hervet, C., Krieg, F., Quillet, E., Boudinot, P. (2018) Genetic and transcriptomic analyses provide new insights on the early antiviral response to VHSV in resistant and susceptible rainbow trout. *BMC Genomics*, 19:482.
20. Maroilley, T., Berri, M., Lemonnier, G., Esquerré, D., Chevaleyre, C., Mélo, S., Meurens, F., Coville, J.L., Leplat, J.J, **Rau, A.**, Bed'hom, B., Vincent-Naulleau, S., Mercat, M.J., Billon, Y., Lepage, P., Rogel-Gaillard, C., and Estellé, J. (2018) Immunome differences between porcine ileal and jejunal Peyer's patches revealed by global transcriptome sequencing of gut-associated lymphoid tissues. *Scientific Reports*, 8:9077.
21. Mondet, F., **Rau, A.**, Klopp, C., Rohmer, M. Severac, D., Le Conte, Y., and Alaux, C. (2018) Transcriptome profiling of the honeybee parasite Varroa destructor provides new biological insights into the mite adult life cycle. *BMC Genomics*, 19:328.
22. He, B., Tjhung, K., Bennett, N., Chou, Y., **Rau, A.**, Huang, J., and Derda, R. (2018) Compositional bias in naïve and chemically-modified phage-displayed libraries uncovered by paired-end deep sequencing. *Scientific Reports*, 8:1214.
23. Monneret, G., Jaffrézic, F., **Rau, A.**, Zerjal, T. and Nuel, G. (2017) Identification of marginal causal relationships in gene networks from observational and interventional expression data. *PLoS One*, 12(3): e0171142.
24. Sauvage, C., **Rau, A.**, Aichholz, C., Chadoeuf, J., Sarah, G., Ruiz, M., Santoni, S., Causse, M., David, J., Glémin, S. (2017) Domestication rewired gene expression and nucleotide diversity patterns in tomato. *The Plant Journal*, 91(4):631-645.
25. Rigaill, G., Balzergue, S., Brunaud, V., Blondet, E., **Rau, A.**, Rogier, O., Caius, J., Maugis-Rabusseau, C., Soubigou-Taconnat, L., Aubourg, S., Lurin, C., Martin-Magniette, M.-L., and Delannoy, E. (2016) Synthetic datasets for the identification of key ingredients for RNA-seq differential analysis. *Briefings in Bioinformatics*, 19(1):65-76. https://dx.doi.org/10.1093/bib/bbw092
26. Gallopin, M., Celeux, G., Jaffrézic, F., **Rau, A.** (2015) A model selection criterion for model-based clustering of

annotated gene expression data. *Statistical Applications in Genetics and Molecular Biology*, 14(5): 413-428.

27. Monneret, G., Jaffrézic, F., **Rau, A.**, Nuel, G. (2015) Estimation d'effets causaux dans les réseaux de régulation génique : vers la grande dimension. *Revue d'intelligence artificielle*, 29(2): 205-227.
28. **Rau, A.**, Maugis-Rabusseau, C., Martin-Magniette, M.-L., Celeux, G. (2015) Co-expression analysis of high-throughput transcriptome sequencing data with Poisson mixture models. *Bioinformatics*, 31(9): 1420-1427.
29. **Rau, A.**, Marot, G. and Jaffrézic, F. (2014) Differential meta-analysis of RNA-seq data from multiple studies. *BMC Bioinformatics*, 16:31.
30. Endale Ahanda, M.-L., Zerjal, T., Dhorne-Pollet, S., **Rau, A.**, Cooksey, A., and Giuffra, E. (2014) Impact of the genetic background on the composition of the chicken plasma miRNome in response to a stress. *PLoS One*, 9(12): e114598.
31. Nuel, G., **Rau, A.**, and Jaffrézic, F. (2013) Using pairwise ordering preferences to estimate causal effects in gene expression from a mixture of observational and intervention experiments.. *Quality Technology and Quantitative Management*, 11(1):23-37.
32. **Rau, A.**, Jaffrézic, F., and Nuel, G. (2013) Joint estimation of causal effects from observational and intervention gene expression data. *BMC Systems Biology*, 8:51.
33. Gallopin, M. **Rau, A.**, and Jaffrézic, F. (2013) A hierarchical Poisson log-normal model for network inference from RNA sequencing data. *PLoS One*, 8(10): e77503.
34. **Rau, A.**, Gallopin, M., Celeux, G., and Jaffrézic, F. (2013) Data-based filtering for replicated high-throughput transcriptome sequencing experiments. *Bioinformatics*, 29(17): 2146-2152.
35. Dillies, M.-A., **Rau, A.**, Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., Keime, C., Marot, G., Castel, D., Estelle, J., Guernec, G., Jagla, B., Jouneau, L., Laloë, D., Le Gall, C., Schaëffer, B., Charif, D., Le Crom, S., Guedj, M., and Jaffrézic, F. (2013) A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Briefings in Bioinformatics*, 14(6): 671-683. `https://dx.doi.org/10.1093/bib/bbs046`
36. Brenault, P., Lefevre, L. **Rau, A.**, Laloë, D., Pisoni, G., Moroni, P., Bevilacquia, C. and Martin, P. (2013) Contribution of mammary epithelial cells to the immune response during early stages of a bacterial infection to Staphylococcus aureus. *Veterinary Research*, 45:16.
37. **Rau, A.**, Jaffrézic, F., Foulley, J.-L., and Doerge, R. W. (2012) Reverse engineering gene regulatory networks using approximate Bayesian computation. *Statistics and Computing*, 22: 1257-1271.
38. **Rau, A.**, Jaffrézic, F., Foulley, J.-L., and Doerge, R. W. (2010) An empirical Bayesian method for estimating biological networks from temporal microarray data. *Statistical Applications in Genetics and Molecular Biology*, 9(1): 9.
39. Furth, A., Mandrekar, S., Tan, A. **Rau, A.**, Felten, S., Ames, M. Adjei, A. Erlichman, C. and Reid, J. (2008) A limited sample model to predict area under the drug concentration curve for 17-(allylamino)-17-demethoxygeldanamycin and its active metabolite 17-(amino)-17-demethoxygeldanomycin. *Cancer Chemotherapy Pharmacology*, 61(1): 39-45.

Pre-prints, technical reports, & other publications

1. **Rau, A.**, Passet, B., Castille, J., Asset, A., Lecardonnel, J., Moroldo, M., Jaffrézic, F., Laloë, D., Moazami-Goudarzi, K., and Vilotte, J.-L. (2021) Potential genetic robustness of Prnp and Sprn double knockout mouse embryos towards ShRNA-lentiviral inoculation… `https://www.biorxiv.org/content/10.1101/2021.10.22.465458v1`
2. Cazals, A., **Rau, A.**, Estellé, J., Bruneau, N., Coville, J.-L., Menanteau, P., Rossignol, M.-N., Jardet, D., Bevilacqua, C., Bed'Hom, B., Velge, P., and Calenge, F. (2021) Comparative analysis of the caecal tonsil transcriptome in two hen lines experimentally infected with Salmonella Enteritidis…
3. Mazo, G., Karlis, D., and **Rau, A.** (2021) A randomized pairwise likelihood method for complex statistical inferences. Submitted. `https://hal.archives-ouvertes.fr/hal-03126621`
4. Cazals, A., Estellé, J., Bruneau, N., Coville, J.-L., Menanteau, P., Rossignol, M.-N., Jardet, D., Bevilacqua, C., **Rau, A.**, Bed'Hom, B., Velge, P., and Calenge, F. (2020) Impact of host genetics on caecal microbiota composition and on Salmonella carriage in chicken. Submitted. `https://www.researchsquare.com/article/rs-76645/v1`
5. Bruford, M., Leroy, G., Orozco-terWengel, P., **Rau, A.**, and Simianer H. (2015) Section B: Molecular tools for exploring genetic diversity. *The Second Report on the State of the World's Animal Genetic Resources for Food and Agriculture* FAO Commission on Genetic Resources for Food and Agriculture.
6. Nuel, G., **Rau, A.**, and Jaffrézic, F. (2013) Joint likelihood calculation for intervention and observational data from a Gaussian Bayesian network. *arXiv* 1305.0709.
7. **Rau, A.**, Celeux, G., Martin-Magniette, M.-L., and Maugis-Rabusseau, C. (2011) Clustering high-throughput se-

quencing data with Poisson mixture models. *Inria Research Report* 7786.
8. **Rau, A.**, Jaffrézic, F., Foulley, J.-L., and Doerge, R. W. (2010) Approximate Bayesian approaches for reverse engineering biological networks. *Proceedings of the Kansas State University Conference on Applied Statistics in Agriculture* Manhattan, Kansas.
9. **Rau, A.** (2008) Success of Volunteer Statistical Consulting Service Leads to Expanded Network. *The Statistical Consultant* 25(1).
10. **Rau, A.** (2008) STATCOM Network Engages Growing Number of Student Volunteers. *Newsletter for the Section on Statistical Education* 13(1).
11. **Rau, A.** (2008) Success of Statistical Service Leads to Expanded Network. *Amstat News* April 2008.

Conference presentations

1. **Invited talk: Mixture models as a useful tool for identifying co-expressed genes from RNA-seq data**
MiMo Workshop on mixture models @ virtual (2021-04-08)
2. **Invited keynote: Integrative and interactive analyses of multi-omics data**
JOBIM 2020 @ virtual (2020-07-02)
3. **Invited talk: Individualized multi-omic pathway deviation scores using multiple factor analysis**
EuroBioc 2019 @ Brussels, Belgium (2019-12-09)
4. **Poster: Integrative methods for multi-omic data reveal multi-level gene and pathway regulation**
AgreenSkills+ annual meeting @ Brussels, Belgium (2019-04-12)
5. **coseq: An R/Bioconductor package for co-expression analyses of RNA-seq data**
Plant and Animal Genomes (PAG) XXVI @ San Diego, California, USA (2018-01-15)
6. **Invited talk: Model-based clustering to identify co-expressed genes from high-throughput sequencing data**
Working Group on Model-Based Clustering @ Perugia, Italy (2017-07-20)
7. **Clustering transformed compositional data using coseq**
useR!2017 @ Brussels, Belgium (2017-07-05)
8. **Invited talk (FAANG workshop): An update on the FAANG pilot project FR-AgENCODE**
Plant and Animal Genomes (PAG) XXVI @ San Diego, California, USA (2017-01-12)
9. **Invited talk: Statistical tools to identify and visualize co-expression clusers from RNA-seq data**
INRA RNA-seq day @ Avignon (2016-11-17)
10. **Identifying marginal causal relationships in gene networks from observational and interventional expression data**
Joint Statistical Meetings of the American Statistical Association @ Chicago (2016-07-31)
11. **Invited talk: Experimental design in 'omics studies**
2nd International Symposium on Microgenomics, Technical Workshop @ Jouy-en-Josas (2016-05-31)
12. **HTSCluster: a mixture-based approach for co-expression analyses of RNA-seq data**
15th Workshop: Statistical Methods for Post-Genomic Data @ Munich (2015-02-13)
13. **HTSDiff: More sensitive differential analysis of RNA-seq data**
Statistical analysis of RNA-seq data: Advances and challenges @ Paris (2013-11-26)
14. **HTSFilter: Data-based filtering for replicated high-throughput sequencing experiments**
Deuxièmes rencontres R @ Lyon (2013-06-28)
15. **Invited round table: Statistics applied to RNA-seq**
Journée de la transcriptome végétale de l'URGV-Genopole @ Evry (2013-05-16)
16. **Joint estimation of causal effects from observational and intervention gene expression data**
StatSeq meeting on genetical genomics @ Paris (2013-03-28)
17. **Invited talk: A comprehensive evaluation of normalization methods for high-throughput RNA sequencing data analysis**
Journée APLIBIO (Alliance des PLates-formes Île-de-France de BIOinformatique) @ Paris (2012-10-11)
18. **Clustering high-throughput sequencing data using Poisson mixture models**
Joint Statistical Meetings of the American Statistical Association @ San Diego, California (2012-07-31)
19. **Clustering high-throughput sequencing data using Poisson mixture models**
12th Workshop: Statistical Methods for Post-Genomic Data @ Lyon (2012-01-26)
20. **Reverse Engineering Gene Networks Using Approximate Bayesian Computation**
11th Workshop: Statistical Methods for Post-Genomic Data @ Paris (2011-01-27)
21. **Approximate Bayesian methods for reverse engineering biological networks**
Conference on Applied Statistics in Agriculture @ Manhattan, Kansas (2010-04-26)

22. **Reverse-Engineering Gene Networks from Microarray Data with Dynamic Bayesian Networks**
    GENESYS Satellite Meeting at the European Conference on Complex Systems @ Warwick, UK (2009-09-22)
23. **Using Dynamic Bayesian Networks with Hidden States to Infer Gene Regulatory Networks**
    Joint Statistical Meetings of the American Statistical Association @ Washington, DC (2009-08-05)
24. **Poster : Reverse-Engineering Genetic Regulatory Interactions from Transcriptomic Data using Dynamic Bayesian Networks**
    2nd Biennial Workshop on Statistical Bioinformatics and Stochastic Systems Biology @ Newcastle, UK (2009-05-18)
25. **Poster: An Empirical Bayes Approach to Inferring Genetic Regulatory Interactions with Dynamic Bayesian Networks**
    Conference on Applied Statistics in Agriculture @ Manhattan, Kansas (2009-04-19)
26. **Poster: An Empirical Bayes Approach to Inferring Genetic Regulatory Interactions with Dynamic Bayesian Networks**
    Gordon Conference on Quantitative Genetics and Genomics @ Galveston, Texas (2009-02-22)
27. **Poster: Seven Years of StatCom at Purdue: Managing a Growing Number of Student Volunteers**
    Joint Statistical Meetings of the American Statistical Association @ Denver, Colorado (2008-08-04)

Seminar & working group presentations

1. **Leveraging multi-omic data for integrative exploratory, predictive, and network analyses**
   NutriNeurO lab seminar @ virtual (2021-11-22)
2. **Multi-omic integration for enhanced interpretability in exploratory analyses**
   Grenoble Laboratoire Jean Kuntzmann seminar @ virtual (2021-04-29)
3. **Happy 20th Birthday, R!**
   INRAE GiBBS team meeting @ virtual (2020-05-18)
4. **Integrative methods for multi-omic data reveal multi-level gene regulation**
   AgroParisTech statistics seminar @ Paris, France (2020-01-20)
5. **Integrative multivariate methods for multi-omic data**
   Lundi de GABI seminar @ Jouy en Josas, France (2020-01-13)
6. **Integrative methods for multi-omic data reveal multi-level gene regulation**
   INRA MaIAGE research seminar @ Jouy en Josas, France (2019-11-18)
7. **Integrative methods for multi-omic data reveal multi-level gene regulation**
   Journée régionale Genotoul @ Toulouse, France (2019-10-04)
8. **Integrative methods for multi-omic data reveal multi-level gene regulation**
   EpiFun workshop @ Orléans, France (2019-09-17)
9. **Exploring drivers of gene expression in The Cancer Genome Atlas**
   Division of Biostatistics Seminar at MCW @ Milwaukee, Wisconsin (2018-12-04)
10. **Co-expression analyses of RNA-seq data in practice with the R/Bioconductor package coseq**
    MixStatSeq Workshop on Mixture Models: Theory and Application @ Paris (2018-06-22)
11. **Exploring drivers of gene expression in The Cancer Genome Atlas**
    Research seminar series, Joseph J. Zilber School of Public Health @ Milwaukee, WI (2018-04-09)
12. **Exploring drivers of gene expression in The Cancer Genome Atlas**
    Physiology Department Seminar at MCW @ Milwaukee, WI (2018-03-28)
13. **Easy interactivity in R with (gg)plotly and Shiny**
    INRA national bioinformatics workshop @ Dijon (2017-06-13)
14. **Challenges in data integration**
    SAPS doctoral school: Experimental animal biology and predictive modelisation @ Jouy en Josas, France (2017-03-17)
15. **Transformation, model choice, and visualization for RNA-seq co-expression**
    Seminar at the Human and Molecular Genetics Center, Milwaukee College of Medicine @ Milwaukee, WI (2016-09-10)
16. **Transformation, model choice, and visualization for RNA-seq co-expression**
    Seminar at the Zilber School of Public Health @ Milwaukee, WI (2016-09-09)
17. **Poisson mixtures with slope heuristics and visualization tools for RNA-seq co-expression**
    Groupe de travail de statistiques du LMRS @ Rouen (2016-05-12)
18. **From genotype to phenotype: what statistical methods to integrate heterogeeous data?**
    INRA national bioinformatics workshop @ Toulouse (2016-03-22)

19. **Integration of heterogeneous 'omics data**
    SAPS doctoral school: Experimental animal biology and predictive modelisation @ Jouy en Josas, France (2016-03-11)
20. **Poisson mixture models and visualization tools for RNA-seq co-expression**
    INRA NGS club @ Jouy en Josas, France (2016-03-08)
21. **Poisson mixtures with slope heuristics and visualization tools for RNA-seq co-expression**
    MAP5 seminar at Université Paris-Descartes @ Paris (2016-01-29)
22. **Model selection in mixture model based classification: Applications in biostatistics**
    4th Annual SFdS Young Statisticians and Probabilists Day @ Paris (2016-01-22)
23. **Statistical analysis of microarray and RNA-seq data**
    Seminar at Toulouse Mathematics Institute (IMT) @ Toulouse (2015-11-17)
24. **Integrative clustering and classification in multiple heterogeneous data**
    Statomique seminar @ Paris (2015-11-09)
25. **RNA-seq co-expression analysis using mixture models**
    NETBIO working group @ Paris (2015-09-29)
26. **HTSCluster: a mixture-based approach for co-expression analyses of RNA-seq data**
    Cirad seminar @ Montpellier (2015-09-25)
27. **Slope heuristics: the missing ingredient for identifying co-expressed genes from RNA-seq data**
    SELECT seminar @ Orsay (2014-10-16)
28. **HTSFilter: filtering replicated RNA-seq data using a data-driven approach**
    Statistics for Systems Biology (SSB) seminar @ Evry (2013-11-12)
29. **Reinforcing the biology-statistics feedback loop with tools for genomic data analysis**
    Seminar at INRA-GABI @ Jouy en Josas, France (2013-11-04)
30. **HTSAnalysis: a suite of R/Bioconductor packages for the analysis of RNA-seq data**
    Statistics for Integrative Biology (SIB) seminar @ Rennes (2013-10-29)
31. **Joint estimation of causal effects from observational and intervention gene expression data**
    NETBIO working group @ Paris (2013-09-20)
32. **Joint estimation of causal effects from observational and intervention gene expression data**
    Statistique et Santé working group @ Paris (2013-06-24)
33. **Joint estimation of causal effects from observational and intervention gene expression data**
    AppliBUGS Workshop @ Paris (2013-06-20)
34. **Joint estimation of causal effects from observational and intervention gene expression data**
    Statistics seminar @ Toulouse (2013-06-18)
35. **Joint estimation of causal effects from observational and intervention gene expression data**
    Statistics for Integrative Biology seminar @ Rennes (2013-02-22)
36. **Differential analysis of RNA-seq data by unsupervised classification**
    Assemblée générale PEPI IBIS @ Toulouse (2012-12-07)
37. **Independent data-based filtering for replicated high-throughput sequencing experiments**
    Statomique seminar @ Lyon (2012-11-27)
38. **Clustering high-throughput sequencing data using Poisson mixture models**
    LGC and SAGA seminar at INRA @ Toulouse (2012-06-25)
39. **Clustering high-throughput sequencing data using Poisson mixture models**
    SSB working group seminar @ Jouy en Josas, France (2012-06-19)
40. **Inferring gene regulatory networks with hidden variables using state space models**
    MIA Biological network inference methodological working group meeting @ Paris (2012-02-09)
41. **Exploring the identifiability of gene regulatory networks with approximate Bayesian computation**
    AppliBugs Workshop @ Paris (2011-12-09)
42. **Reverse Engineering Gene Networks Using Approximate Bayesian Computation (ABC)**
    Seminar at the Institut de Recherche Mathématique Avancée @ Strasbourg (2011-05-24)
43. **Reverse Engineering Gene Networks: A Statistician's Perspective**
    Seminar at the Unité de Recherche en Génomique Végétale @ Evry (2011-04-07)
44. **Reverse Engineering Gene Networks Using Approximate Bayesian Computation (ABC)**
    Seminar at the Institut de Mathématiques de Luminy @ Marseille (2011-04-04)
45. **Reverse Engineering Gene Networks Using Approximate Bayesian Computation (ABC)**
    Seminar at the Laboratoire Statistique et Génome @ Evry (2011-03-22)
46. **Reverse Engineering Gene Networks Using Approximate Bayesian Computation (ABC)**

Seminar at the équipe Génétique et Génomique Statistique @ Le Kremlin Bicêtre (2011-03-02)

47. **Reverse Engineering Gene Networks Using Approximate Bayesian Computation (ABC)**
    Rencontre de statistique autour des modèles hiérarchiques @ Strasbourg (2011-01-14)
48. **Reverse Engineering Gene Networks Using Approximate Bayesian Computation (ABC)**
    INA P-G, Paris Descartes, and SELECT working group @ Paris (2010-10-18)
49. **Approximate Bayesian methods for reverse engineering biological networks**
    Bioinformatics seminar at Purdue University @ West Lafayette, Indiana (2010-04-13)
50. **Inférence sur les réseaux génomiques par des modèles espace-état**
    Seminar at the AgroParisTech @ Paris (2009-06-22)
51. **Inférence sur les réseaux génomiques par des modèles espace-état**
    Seminar at the UMR GABI-INRA @ Jouy-en-Josas (2009-06-15)
52. **Reverse Engineering Gene Regulatory Networks**
    Ph.D. student seminar, INRA Département de Génétique Animale @ Jouy-en-Josas (2009-03-23)
53. **Poster: Inferring Gene Regulatory Network through Linear Feedback State Space Models**
    Ph.D. student seminar, INRA Département de Génétique Animale @ Toulouse (2008-03-20)

Participation in working groups

- Statomique (2009-present)
- Netbio (2014-present)

Software

1. **rpl**: Randomized pairwise likelihood method for complex statistical inferences, available at GitHub
2. **padma**: Pathway deviation scores using multiple factor analysis, available at BioC
3. **Invest Astuces**: An R/Shiny interactive web application for financial and real estate loan simulations, available as a Shiny web app
4. **Regeneration Rosetta**: An R/Shiny interactive web application to explore regeneration-associated gene expression and chromatin accessibility, available as a Shiny web app
5. **maskmeans**: Multi-view aggregation/splitting K-means clustering algorithm, available at GitHub
6. **Edge in TCGA**: An R/Shiny interactive web application for the exploration of drivers of gene expression in The Cancer Genome Atlas, available as a Shiny web app
7. **coseq**: Co-expression analysis of sequencing data, available at BioC
8. **ICAL**: Model selection for model based clustering of annotated data, available at GitHub
9. **metaRNASeq**: Meta-analysis of RNA-seq data, available at CRAN
10. **HTSDiff**: Differential analysis for RNA-seq data, available at R-Forge
11. **HTSFilter**: Filter for replicated high-throughput sequencing data, available at BioC
12. **HTSCluster**: Clustering high-throughput sequencing data with Poisson mixture models, available at CRAN
13. **ebdbNet**: Empirical Bayes estimation for dynamic Bayesian networks, available at CRAN

Administrative activities

Reviewer *2010-present*

Annals of Applied Statistics, Bioinformatics, BMC Bioinformatics, BMC Genomics, BMC Medical Genetics, Briefings in Bioinformatics, Computational and Structural Biotechnology Journal, F1000 Research, G3, Genetics Selection Evolution, Genome Biology, GigaScience, IEEE Proceedings, Journal of Computational Biology, Journal of Computational and Graphical Statistics, JRSS-C, Molecular Genetics and Genomics, Nature Computational Science, Nucleic Acids Research, Revue d'Intelligence Artificielle, RNA, Statistical Applications in Genetics and Molecular Biology, The Plant Journal

Elected member (substitute) *2021-2024*

INRAE Animal Genetics Department Conseil Scientifique

Appointed member *2021-2024*

INRAE Commission scientifique spécialisée (CSS) Mathématiques, Informatique, Scienes et Technologies du numérique, Intelligence artificielle et Robotique (MISTI

Reviewer *2021*

Saclay Plant Sciences Research Open Call

Member *2021*

INRAE GABI Scientific Communication and Mediation committee

Member *2021*

Evaluation committee for Assistant Professor position at Université Le Mans

Member *2021*

Evaluation committee for Research Scientist (chargé de recherche) position at INRAE

Associate Editor *2021*

BMC Genomics

Review Editor *2021*

Editorial Board of Statistical Genetics and Methodology, Frontiers in Genetics

Mentor *2021*

Université Paris-Saclay, Women and Science Mentoring Program

Reviewer *2020*

ANR MRSEI grant call

Scientific committee member *2019*

UseR!2019 international conference

Reviewer *2017*

University of Wisconsin-Milwaukee Research Growth Initiative grant call

Reviewer *2017*

INRA Selgen Metaprogramme grant call

Reviewer *2017*

Nantes Excellence Trajectory (NExT) Health and Engineering initiative "Internal interdisciplinary project" call

Member *2016*

Evaluation committee for Assistant Professor position at Université Rennes I, UMR Inserm IRSET 1085

Member *2016*

Evaluation committee for Assistant Professor position at Université Rennes I, IGDR (Institute of Genetics and Developmental biology of Rennes), CNRS UMR 6290

Scientific committee president *2016*

Rencontres R national conference

Member *2015-2019*

Conseil Scientifique des Utilisateurs (CSU) of the MIGALE Bioinformatics platform (INRA, Jouy en Josas)

Chair *2010*

Committee on Student Pro Bono Statistics of the ASA

Member *2009-2010*

Committee on Student Pro Bono Statistics of the American Statistical Association (ASA)

Organizer

2009

Invited round table (The Pros of Pro Bono Statistics) at the ASA Joint Statistical Meetings (Washington, DC, USA

Member

2006-2010

Statistics in the Community (StatCom) at Purdue University: StatCom is a volunteer organization of graduate students that provides free professional statistical consulting services to government and nonprofit groups

Funding

DINAMIC

2021-2023

Differential network analysis of mixed-type data with copulae, INRAE DIGIT-BIO metaprogram grant

- PI: Andrea Rau

GENE-SWitCH

2019-2023

The regulatory genome of swine and chicken: functional annotation during development, H2020 RIA grant

- PI: Elisabetta Giuffra and Herve Acloque; Role: Co-investigator and task leader

MiniSRegress

2019-2021

Characterisation of a minipig spontaneous regression model with no invalidating adverse effects, Inserm Plan cancer 2014-2019, New Experimental Models call

- PI: Giorgia Egidy-Maskos; Role: Co-investigator

LIPOMEC

2018-2022

Towards a better understanding of ruminant milk lipolysis through an integrative biology approach in milk and mammary epithelial cells, French National Research Agency (ANR) grant

- PI: Christelle Cebo; Role: Co-investigator

EpiFun

2018-2020

Systems biology for genomic selection, INRA Selgen metaprogram grant

- PI: Nathalie Vialaneix and Thomas Faraut; Role: Co-investigator

AgreenSkills+

2017-2019

Integrative analysis of multi-omics data for improved detection power of functional genetic variants, AgreenSkills+ mobility grant (University of Wisconsin-Milwaukee)

- PI: Andrea Rau

Microficient

2016-2019

Relationships between digestive microbiota and feed efficiency in cattle, AP-2016-007

- PI: Yuliaxis Ramayo and Gilles Renand; Role: Co-investigator

CARISTO-PF

2016-2019

Characterization and management of health and environmental risks linked to the development of ciguatera in phytobenthos in French Polynesia, French Polynesia Territory grant

- PI: Gregory Nuel; Role: Co-investigator

SalmoCar

2015-2017

Genetic and microbiotal control of Salmonella carriage in chicken and mice, Institut Carnot Pasteur Maladies Infectieuses (PMI) / Institut Carnot Santé Animale (ICSA) grant

- PI: Xavier Montagutelli; Role: Co-investigator

COSI-net

2015-2016

Using COmbinatorial gene Silencing & Inactivation to infer gene NETworks, INRA Animal Genetics Department internal grant

- PI: Andrea Rau

MixStatSeq

2014-2018

Mixture-based procedures for statistical analysis of RNA-seq data, French National Research Agency (ANR) grant (ANR-13-JS01-0001-01)

- PI: Cathy Maugis-Rabusseau; Role: Co-investigator

Causality

2014

Causal network inference, INRA Animal Genetics Department internal grant

- PI: Florence Jaffrézic; Role: Co-investigator

Advising

Maxime Guilleton *2021*

CDD (6 months)

- Bioinformatics research engineer, LIPOMEC grant (with Mylène Delosiere)

Smahane Chalabi *2021-2022*

Postdoc

- A diet x epigenetics study in pigs, GENE-SWitCH grant (with Elisabetta Giuffra and Sarah Djebali)

Alexandre Asset *2021*

L2 internship

- Inference of co-expression networks from intervention transcriptomic data (with Florence Jaffrézic and Denis Laloë)

Solène Pety *2021*

M1 internship

- Knowledge transfer using multivariate gene expression projections onto a large-scale reference database (with Catherine Giauffret)

Fanny Mollandin *2019-2022*

PhD

- Incorporating known functional annotations into Bayesian genomic prediction models (with Pascal Croiseau, co-funding from EU Horizon 2020 RIA grant GENE-SWitCH)

Raphaëlle Momal-Leisenring *2017*

M2 internship

- Integrative statistical analysis of multi-omics data

Frédéric Jehl *2017*

M2 internship

- Impact of heat stress on liver and blood transcriptomes of laying hens (with Tatiana Zerjal)

Dr. Manuel Revilla Sanchez *2016*

3-month PhD Erasmus+ Learning Mobility

- An integrative gene network analysis of the genetic determination of pig fatty acid composition (with Jordi Estelle and Yuliaxis Ramayo Caldas)

Babacar Ciss *2016*

M2 internship

- Constructing predictive models for ovine production data (with Eli Sellem, [Allice](http://www.allice.fr/))

Audrey Hulot *2015*

M1 internship

- Incorporating a priori biological knowledge into gene network inference from observational and intervention gene expression data (with Florence Jaffrézic)

Meriem Benabbas *2015*

M1 internship

- Identifying differentially expressed genes from RNA-seq data using mixture models

Gilles Monneret *2014-2018*

PhD

- Estimation of causal effects in gene networks from observational and intervention data (with Grégory Nuel and Florence Jaffrézic)

Marc Teissier and Chaoyu Dong *2014*

M1 STV/EM-ABG internship

- Power to detect significantly differential gene expression using RNA-seq data

Mélina Gallopin *2012-2015*

PhD

- Clustering and network inference for RNA-seq data (with Gilles Celeux and Florence Jaffrézic)

Rémi Bancal *2012*

M2 internship

- Gene network estimation by adaptive knockout experiments (with Grégory Nuel and Florence Jaffrézic)

Mélina Gallopin *2012*

M2 internship

- Gene network inference from RNA sequencing expression data (with Gilles Celeux and Florence Jaffrézic)

Ambre Giguelay *2021*

PhD evaluation committee member (rapportrice)

Leila Khajavi *2021*

PhD evaluation committee member (rapportrice)

Thibault Poinsignon *2021-2023*

PhD advisory committee member

Nicolas Jouvin *2021*

PhD evaluation committee member (rapportrice)

Lucile Broséus *2021*

PhD evaluation committee member (rapportrice)

Wilfried Heyse *2020-2022*

PhD advisory committee member

Antoine Leduc *2019-2022*

PhD advisory committee member

Alyssa Imbert *2018*

PhD evaluation committee member (rapportrice)

Frédéric Jehl *2017-2020*

PhD advisory committee member

Valentin Voillet *2016*

PhD evaluation committee member

Gabriel Guillocheau *2015-2018*

PhD advisory committee member

Teaching

Agrocampus Researcher School (23 June) *2021*

Instructor (with S. Lagarrigue and Y. Blum) @ Rennes

- Statistical analysis of RNA-seq data

Agrocampus Researcher School (6-7 February) *2020*

Instructor (with S. Lagarrigue and Y. Blum) @ Rennes

- Statistical analysis of RNA-seq data

Researcher training session: From gene expression to genomic networks (17-22 July) *2019*

Instructor (with M.-L. Martin-Magniette and E. Delannoy) @ INRAE, AgroImpact

- Differential analysis of RNA-seq data

Analysis of livestock metagenomics datasets (13-17 May) *2019*

Instructor (with J. Estellé @ INRA URZ, Guadeloupe

Physiological genomics (10 hours) *2019*

Instructor @ Medical College of Wisconsin

- R Bootcamp

Data management and visualization in R (3 course units) *2018*

Instructor @ University of Wisconsin-Milwaukee

PiGutNet Training School (3 hours) *2017*

Instructor @ INRA, Jouy en Josas

- Differential abundance analysis for microbial marker-gene surveys with metagenomeSeq

Bayesian statistics for genomics course (18 hours) *2017*

Instructor (coursework and labs) @ Université d'Evry Val d'Essonne

- Mathematics for the Life Sciences: Statistical Engineering and Genomics, M2

Agrocampus Researcher School (1-2 February) *2017*

Instructor (with S. Lagarrigue and Y. Blum) @ Rennes

- Statistical analysis of RNA-seq data

Genomics course (12 hours) *2017*

Instructor (coursework and labs) @ Ensai, Rennes

- Biostatistics M2

SPS Summer School: From gene expression to genomic networks (17-22 July) *2016*

Instructor (coursework and labs) @ Institute of Plant Sciences Paris-Saclay

- Co-expression analysis of RNA-seq data (3 hours

Bayesian statistics for genomics course (18 hours) *2016*

Instructor (coursework and labs) @ Université d'Evry Val d'Essonne

- Mathematics for the Life Sciences: Statistical Engineering and Genomics, M2

Agrocampus Researcher School (10-11 February) *2016*

Instructor (with S. Lagarrigue and Y. Blum) @ Rennes

- Statistical analysis of RNA-seq data

Genomics course (33 hours) *2016*

Instructor (coursework and labs) @ Ensai, Rennes

- Biostatistics M2

Mathematical Engineering for Life Sciences Master, M1: Case study (10 hours) *2015*

Instructor (coursework and labs) @ Université Paris Descartes

Genomics course (33 hours) *2015*

Instructor (coursework and labs) @ Ensai, Rennes

- Biostatistics M2. Note: The genomics course at Ensai was significantly expanded and re-developed in 2015 by myself and Mickaël Guedj.

BioBayes Researcher School (7-11 October) *2013*

Scientific committee member and instructor (coursework and labs) @ Cannes Mandelieu

- Bayesian statistical methods: Introduction to theory and applications in food, environment, epidemiology, and genetics

Genomics course (6 hours) *2013*

Instructor (coursework and labs) @ Ensai, Rennes

- Biostatistics M2

Training school on rabbit and pig genome analysis (6 hours) *2012*

Instructor @ COST action research school, Norwich, United Kingdom

Genomics course (6 hours) *2012*

Instructor (coursework and labs) @ Ensai, Rennes

- Biostatistics M2

Statistical Methods for Genome Enabled Prediction (2 hours) *2012*

Instructor @ European Graduate School in Animal Breeding and Genetics, Paris

- Approximate Bayesian methods: Application to gene regulatory networks

Next generation sequencing school for INRA researchers (2 hours) *2012*

Instructor @ INRA research school, Ecully

Computational biostatistics (6 hours) *2012*

Instructor (coursework and labs) @ UFR de Sciences, Université Paris-Sud 11

- Bioinformatics and Biostatistics / Mathematical engineering / Probability and Statistics M2

Statistical modeling (24 hours) *2011*

Lab assistant (R) @ UFR de Sciences, Université Paris-Sud 11

- Bioinformatics and Biostatistics Master, M1

Bioinformatics and Biostatistics / Mathematical engineering / Probability and Statistics M2 (12 hours) *2011*

Instructor (coursework and labs) @ UFR de Sciences, Université Paris-Sud 11

- Bioinformatics and Biostatistics / Mathematical engineering / Probability and Statistics M2

Elementary statistical methods (60 hours) *2006*

Lab assistant (SPSS) @ Department of Statistics, Purdue University (West Lafayette, Indiana, USA)

- 1st and 2nd year undergraduate