Pierre Neuvial

CNRS researcher at LaMME (CR1)
Primary affiliation: Mathematics and interactions (section 41)
Secondary affiliation: Modeling and analysis of biological systems (commission 51)

Laboratoire de Mathématiques et Modélisation d'Évry (LaMME), Équipe Statistique & Génome
Université d'Évry Val d'Essonne, UMR CNRS 8071 / ENSIIE / USC INRA
23, boulevard de France, 91000 EVRY, France
Tel: +33 1 64 85 35 44
pierre.neuvial@math.cnrs.fr
www.math-evry.cnrs.fr/members/pneuvial/

Born 1979
French citizenship

Research interests

- Multiple testing: theory and application to high-throughput genomic data
- Statistical methods for genomic data analyses, specifically DNA copy number analyses
- Applications to cancer research
- Reproducible research

Professional experience

since 2011	CNRS Researcher, LaMME, Évry (2011-2015: CR2; since fall 2015: CR1)
2010-2011	Post-doc, Statistique & Génome, UMR CNRS 8071, Évry
2008-2010	Post-doc, UC Berkeley, Department of Statistics (USA)
2003-2005	Research engineer in statistics and bioinformatics, Institut Curie, Biofinformatics
2001-2002	Intern statistician at Crédit Lyonnais, Groupe de Recherche Opérationnelle

Education

2004-2008	PhD in Applied Mathematics, Université Paris Diderot (Paris 7) Dissertation: "Contributions to the statistical analysis of DNA microarray data"
2002-2003	M. S. in Stochastic models, Université Paris Diderot (Paris 7)
1999-2003	ENSAE ParisTech

Teaching

For mathematicians/statisticians (M.Sc. level)

2015-2016	Statistical methods for genomic data analysis, 2x3h, Institut des Actuaires, Paris
2015-2016	Programming for genomics, 15h, Université d'Évry val d'Essonne
2011-2015	Statistical methods for molecular biology, 8h/y, ENSAE ParisTech
2011	Statistical methods for genomic data analysis, 10h, ENSAI, Rennes
2011-2012	Introduction to Statistical methods for molecular biology, 6h, École Centrale Paris
2005-2008	Statistical methods for molecular biology, 8h/year, ENSAE ParisTech
2004-2008	Mathematical Statistics, 20-32h/y, ENSAE ParisTech (practicals)

For biologists/bioinformaticians (M.Sc. level)

2013-2015	Statistical tests, 20h/year at Université d'Évry val d'Essonne
2011-2015	Statistical methods for genomic data analysis, 12-24h/y, Univ. Paris Sud / Orsay
2011-2012	Statistical tests, 16.5h/y, Université d'Évry val d'Essonne (practicals)
2004-2007	DNA copy number analysis, Université Paris Diderot / Paris 7

PhD students (with percentage of supervision)

- Guillermo Durand (2015-): "Multiple testing for structured biological data" (50%). With Etienne Roquain. Funded by École doctorale des Sciences Mathématiques de Paris Centre, University Paris 6.
- Benjamin Sadacca (2013-): "Tumoral microenvironment and treatment response in breast cancers" (50%). With Fabien Reyal. Funded by Institut Curie.
- Morgane Pierre-Jean (2013-): "Statistical methods for the analysis of structured genomic data" (100%). Applications in cancerology. With Catherine Matias. Funded by École doctorale GAO, University of Évry.
- Alia Dehman (2012-2015): "Structured sparse regression for Genome-Wide Association Studies" (50%). With Christophe Ambroise. Funded by École doctorale GAO, University of Évry. Alia defended her thesis in december 2015 and is now working for Altran (Toulouse).

Research funding (as PI or co-PI)

- 2015-2016: PEPS FaSciDo CNRS INSMI/INS2I (Fondements et Applications de la Science des Données): "Approches post-hoc pour les tests multiples à grande échelle" (12 K€ in 2015 + 5 K€ in 2016);
- 2012-2014: Cancéropôle Ile de France/Institut National du Cancer (INCa), "Data integration for cancer studies" (PI for UMR 8071; 10 K€)

Participation to other grants

2015-2018, INSERM/Cancer systems biology: "LIONS: Large-scale Integrative approach to unravel the complex relationships between differentiatiON and tumorigenesiS". PI: Mohamed Elati.

2013-2014, PEPS BMI, CNRS: "Comparaison de Réseaux de régulation par Enumération de Perturbations". PI: Etienne Birmelé;

2013, MIA network: INRA, Division of Applied Mathematics and Informatics, "Statistical methods for high-dimensional genomic data" (1.5 K€)

Fall 2012, Évry, FDIR: Université d'Évry Val d'Essonne, "Segmentation methods for DNA copy number analyses in cancer research" (travel grant, 2 K€)

Software development

Contribution to the development and technical support of R packages within open source and collaborative projects of statistical/bioinformatics tools, including:

- MANOR, DEGraph (Bioconductor)
- aroma.cn, aroma.cn.eval (the Aroma Project)
- jointseg (R-forge)
- tmle.npvi (CRAN)

Editorial boards

- Associate Editor of the International Journal of Biostatistics (since 2015)

Reviewer activities

- For journals in statistics/machine learning: American Statistician (2016), Annals of Statistics (2013, 2015), Annals of Applied Statistics (2009, 2011), Electronic Journal of Statistics (2012), Journal of Machine Learning Research (2008), Statistics (2013), Journal de la Société Française de Statistique (2014); Statistica Sinica (2015-2016), Scandinavian journal of Statistics (2015)
- For journals in applied statistics/bioinformatics: Advances in Bioinformatics (2012), Bioinformatics (2014 ×3), Biostatistics (2011), BMC Bioinformatics (2011, 2014×2), BMC Genomics (2012), EURASIP Journal on Bioinformatics and Systems Biology (2014), Genome Research (2009), PLoS One (2012), Statistics in Medicine (2009);
- For funding agencies: French National Research Agency (ANR), German-Israeli Fundation for Scientific Research and Development.

Organization of conferences and conference sessions

- Scientific commitee member of JOBIM 2016
- Scientific commitee member and local organizer for "Statistical analysis of Massive genomic data", Évry, Nov. 2015: `http://biomaths.genopole.fr`
- Invited session on "Multiple Testing", Journées MAS 2014

Hiring commitees

- 2016: member and president of a "Commission *ad hoc*" for a position of Temporary Lecturer (ATER) at Université d'Evry val d'Essonne.
- 2016: member of Institut Curie's IC3i PhD Program Review Board.
- 2016: member of a "Comite de sélection" for a position of Assistant Professor at Université de Bordeaux/Institut de santé publique, d'épidémiologie et de développement (ISPED).
- 2013: member of a "Comité de sélection" for a position of Assistant Professor at Université Paris-Sud: 26-67 MCF 2109.

Other responsibilities

- since 01/2014: "adjoint au directeur" (deputy director) of the Laboratoire de Mathématiques et Modèles Aléatoires d'Évry (LaMME)
- 2014-2016: member of the PhD thesis committee of 4 Phd students
- 2012: Organization of the "Fête de la science" for the Department of Mathematics of the University of Évry

Journal papers

[J1] A. Chambaz and P. Neuvial. "tmle.npvi: targeted, integrative search of associations between DNA copy number and gene expression, accounting for DNA methylation". *Bioinformatics* 31.18 (2015), pp. 3054–6.

[J2] A. Dehman, C. Ambroise, and P. Neuvial. "Performance of a Blockwise Approach in Variable Selection using Linkage Disequilibrium Information". *BMC Bioinformatics* (2015).

[J3] T. Picchetti et al. "A model for gene deregulation detection using expression data". *BMC Systems Biology* (Dec. 2015).

[J4] M. Pierre-Jean, G. J. Rigaill, and P. Neuvial. "Performance evaluation of DNA copy number segmentation methods". *Briefings in Bioinformatics* 4 (2015), pp. 600–615.

[J5] I. Brito et al. "Stability-based comparison of class discovery methods for array-CGH profiles". *PLoS One* 8.12 (Dec. 2013), e81458.

[J6] P. Neuvial. "Asymptotic Results on Adaptive False Discovery Rate Controlling Procedures Based on Kernel Estimators". *Journal of Machine Learning Research* 14 (2013), pp. 1423–1459.

[J7] A. Chambaz, P. Neuvial, and M. J. van der Laan. "Estimation of a Non-Parametric Variable Importance Measure of a Continuous Exposure". *Electron. J. Statist.* 6 (2012), pp. 1059–1099.

[J8] L. Heiser et al. "Subtype and pathway specific responses to anticancer compounds in breast cancer". *Proceedings of the National Academy of Sciences* 109.8 (Feb. 2012), pp. 2724–2729.

[J9] L. Jacob, P. Neuvial, and S. Dudoit. "More Power via Graph-Structured Tests for Differential Expression of Gene Networks". *Annals of Applied Statistics* 6.2 (2012), pp. 561–600.

[J10] P. Neuvial and E. Roquain. "On false discovery rate thresholding for classification under sparsity". *Annals of Statistics* 40.5 (2012), pp. 2572–2600.

[J11] M. Ortiz-Estevez et al. "CalMaTe: A Method and Software to Improve Allele-Specific Copy Number of SNP Arrays for Downstream Segmentation". *Bioinformatics* 28.13 (July 2012), pp. 1793–1794.

[J12] A. B. Olshen et al. "Parent-specific copy number in paired tumor-normal studies using circular binary segmentation". *Bioinformatics* 27.15 (Aug. 2011), pp. 2038–2046.

[J13] The Cancer Genome Atlas Research Network. "Integrated Genomic Analyses of Ovarian Carcinoma". *Nature* 474.7353 (June 2011), pp. 609–615.

[J14] H. Bengtsson, P. Neuvial, and T. P. Speed. "TumorBoost: Normalization of allele-specific tumor copy numbers from a single pair of tumor-normal genotyping microarrays". *BMC Bioinformatics* 11.1 (2010), p. 245.

[J15] H. Noushmehr et al. "Identification of a CpG Island Methylator Phenotype that Defines a Distinct Subgroup of Glioma". *Cancer Cell* 17.5 (Apr. 2010), pp. 510–522.

[J16] M. A. Bollet et al. "High-resolution mapping of DNA breakpoints to define true recurrences among ipsilateral breast cancers." *J Natl Cancer Inst* 100.1 (2008), pp. 48–58.

[J17] P. Neuvial. "Asymptotic properties of false discovery rate controlling procedures under independence". *Electron. J. Statist.* 2 (2008), pp. 1065–1110.

[J18] M. Elati et al. "LICORN: LearnIng COoperative Regulation Networks". *Bioinformatics* 23.18 (2007), pp. 2407–2414.

[J19] P. La Rosa et al. "VAMP: visualization and analysis of array-CGH, transcriptome and other molecular profiles." *Bioinformatics* 22.17 (Sept. 2006), pp. 2066–2073.

[J20] S. Liva et al. "CAPweb: a bioinformatics CGH array Analysis Platform." *Nucleic Acids Res* 34.Web Server issue (July 2006), pp. 477–481.

[J21] P. Neuvial et al. "Spatial normalization of array-CGH data." *BMC Bioinformatics* 7.1 (May 2006), p. 264.

Book chapters

[BC1] P. Neuvial, H. Bengtsson, and T. P. Speed. "Statistical analysis of Single Nucleotide Polymorphism microarrays in cancer studies". *Handbook of Statistical Bioinformatics*. Ed. by H. H.-S. Lu, B. Schölkopf, and H. Zhao. Springer Handbooks of Computational Statistics. Springer, 2011.

Technical reports

[TR1] P. Neuvial. "Contributions à l'analyse statistique des données de puces à ADN". PhD thesis. Institut Curie et Université Paris VII (France), 2008.

[TR2] E. Hauvuy, B. Lebrave, and P. Neuvial. "Analyse statistique du lien entre les plages homogènes de séquences d'ADN de différentes bactéries". MA thesis. ENSAE Paristech et Université Paris Diderot, 2003.

[TR3] R. Elie et al. *A Model of Prepayment for the French Residential Loan Market*. Tech. rep. Groupe de Recherche Opérationnelle, Crédit Lyonnais, France, 2002.

Popular science

[PS1] P. Neuvial. "Vers une médecine personnalisée grâce à la recherche en génomique". *Variances* 48 (Oct. 2013), pp. 31–33.

[PS2] P. Neuvial. "Tests multiples en génomique". *La gazette des mathématiciens* 130 (Oct. 2011), pp. 71–76.

[PS3] P. Neuvial and P.-Y. Bourguignon. "Problématiques statistiques à l'heure de la post-génomique". *Variances* 35 (Feb. 2009), pp. 56–60.

Invited talks

[IT1] *Performance evaluation of DNA copy number segmentation methods*. European Meeting of Statistians 2015, Amsterdam, the Netherlands. July 2015.

[IT2] *Performance of a blockwise approach to account for linkage disequilibrium in genome-wide association studies*. Colloque de la Société Française de Biologie Théorique, Poitiers, France. June 2015.

[IT3] *Post hoc inference for multiple testing*. Journées de Statistique de Rennes, 2015, Rennes, France. Oct. 2015.

[IT4] *Statistique en grande dimension; applications en génomique*. Nouveaux défis du big data, Evry Sciences et Innovation, France. Apr. 2015.

[IT5] *Evaluating DNA copy number analysis methods*. StatLearn 2014, Paris, France. Apr. 2014.

[IT6] *Evaluating DNA copy number segmentation methods*. Time Dynamic Change Point Models and its Applications, Göttingen. Oct. 2014.

[IT7] *Evaluating DNA copy number analysis methods*. Joint Statistical Meetings, Montréal, Canada. Aug. 2013.

[IT8] *Estimation of a Non-Parametric Variable Importance Measure of a Continuous Exposure*. Journées MAS 2012, session "Applications de la statistique semiparamétrique aux sciences du vivant", Clermont-Ferrand, France. Aug. 2012.

[IT9] *Estimation of a Non-Parametric Variable Importance Measure of a Continuous Exposure*. Journées de Statistiques 2012, Brussels, Belgium. May 2012.

[IT10] *More Power via Graph-Structured Tests for Differential Expression of Gene Networks*. Statistical Methods for Post-Genomic Data (SMPGD) workshop 2012, Lyon, France. Jan. 2012.

[IT11] *Statistical analysis of SNP array data in cancer studies*. École thématique "Analyse du génome tumoral", Maffliers, France. Mar. 2012.

Contributed talks

[CT1] *On false discovery rate thresholding for classification under sparsity.* Mathematical Statistics and Applications, Fréjus, France. Sept. 2011.

[CT2] *Greatly Improved Allele-specific Tumor Copy Numbers with DNA Microarrays when a Matched Normal is Available.* CSB 2010:International Conference on Computational Systems Bioinformatics (poster talk), Stanford, CA, USA. Aug. 2010.

[CT3] *Asymptotic properties of False Discovery Rate Controlling Procedures.* Mathematical Statistics and Applications, Fréjus, France. Sept. 2008.

[CT4] *Intrinsic bounds on the Benjamini-Hochberg multiple comparison procedure.* PASCAL workshop on Multiple Simultaneous Hypothesis Testing, Paris, France. May 2007.

Seminars

[Sem2] *Evaluating DNA copy number segmentation methods.* C3BI seminar, Institut Pasteur, Paris, France. Dec. 2015.

[Sem4] *Statistique en grande dimension; applications en génomique.* Nouveaux défis du big data, Evry Sciences et Innovation, France. Apr. 2015.

[Sem3] *Evaluating DNA copy number analysis methods.* Statistical Data Integration Challenges in Computational Biology: Regulatory Networks and Personalized Medicine, Banff International Research Station for Mathematical Innovation and Discovery, Canada. Aug. 2013.

[Sem4] *Evaluating DNA copy number analysis methods.* Séminaire de Statistique de Montpellier. June 2013.

[Sem5] *False Discovery Rate control for multiple testing.* Neyman Seminar, UC Berkeley, Department of Statistics. Sept. 2012.

[Sem6] *Joint segmentation of total copy numbers and allelic ratios in cancer samples using the group fused Lasso.* Séminaire Mathématiques-Évolution-Génome, Fédération de Recherche des Unités de Mathématiques de Marseille (FRUMAM). Nov. 2012.

[Sem7] *Joint segmentation of total copy numbers and allelic ratios in cancer samples using the group fused Lasso.* Speed/Dudoit/Purdom group meeting, UC Berkeley, Department of Statistics. Sept. 2012.

[Sem8] University of Postdam, Germany, Invited by Gilles Blanchard. Dec. 2011.

[Sem9] *Analyse statistique de données génomiques.* Séminaire professionnel à l'ENSAI, Rennes. Nov. 2011.

[Sem10] *On false discovery rate thresholding for classification under sparsity.* Séminaire parisien de statistique, Paris, France. May 2011.

[Sem11] *Tests multiples en génomique.* Inauguration de la Fédération de Mathématiques de l'Université d'Evry val d'Essonne, Evry. Nov. 2011.

[Sem1] *On detecting and calling DNA copy number alterations in cancer samples from genotyping microarrays.*

- Institut Curie, Paris. Invité par Emmanuel Barillot.
- Statistics for Systems Biology seminar, Paris. Invité par Sophie Schbath.
- AgroParisTech, Paris. Invité par Stéphane Robin.
- Laboratoire Statistique et Génome, Évry. Invité par Christophe Ambroise.

. 2010.

[Sem12] *On detecting and calling DNA copy number alterations in cancer samples from genotyping microarrays.* Stanford University, Zhang/Holmes group meeting. Invité par Nancy R. Zhang. Feb. 2010.

[Sem13] *Statistical Analysis of Single Nucleotide Polymorphism Microarrays in Cancer Studies.* Stanford Biostatistics Workshop, Invité par Chiara Sabatti. Sept. 2010.

[Sem14] *Targeted maximum likelihood estimation of the relationship between copy number and gene expression in cancer studies.* Statistical Genomics in Biomedical Research Workshop, Banff International Research Station for Mathematical Innovation and Discovery, Canada. July 2010.

[Sem15] *TumorBoost: Normalization of allele-specific tumor copy numbers in paired tumor/normal designs for genotyping microarrays.* University of California at Berkeley, Séminaire de statistique et génomique. Invité par Sandrine Dudoit. Oct. 2009.

[Sem16] *Asymptotic properties of False Discovery Rate Controlling Procedures.* Groupe de travail "Statistique des systèmes biologiques", Paris, France. Invité par Sophie Schbath. Feb. 2008.

[Sem17] *Defining true recurrences among ipsilateral breast tumor recurrences using DNA copy number data.* University of California at Berkeley, Séminaire de statistique et génomique. Invité par Sandrine Dudoit. Oct. 2008.

[Sem18] *Propriétés asymptotiques de procédures de contrôle du False Discovrey Rate.* Groupe de travail "Mathématiques Appliquées à Paris V", Paris, France. Invité par Avner Bar-Hen. Jan. 2008.

[Sem19] *Introduction aux applications de la statistique à la biologie et la médecine.* Séminaire de formation d'enseignants de mathématiques du secondaire (Université Paris Diderot). Invité par Jacqueline Mac Aleese. Mar. 2007.

[Sem20] *Statistical analysis of copy number and expression microarray data - Application to cancer study.* Netherlands Cancer Institute (Amsterdam, Pays-Bas). Invité par Lodewyk Wessels. Oct. 2007.

[Sem21] *Tools and methods for DNA copy number and expression microarray data analysis - Application to cancer study.* Free University Medical Center (Amsterdam, Pays-Bas). Invité par Mark Van de Wiel. Oct. 2007.

[Sem22] *Tools and methods for DNA copy number microarray data analysis.* Waterman Seminar, Leibniz Institute of Plant Genetics and Crop Plant Research and Bioinformatics Centre Gatersleben-Halle (Gatersleben, Allemagne). Invité par Ivo Grosse. Oct. 2007.

Software

[S1] A. Dehman and P. Neuvial. *BALD: a Block-wise Approach using Linkage Disequilibrium information.* R package, `http://www.math-evry.cnrs.fr/logiciels/bald`. 2015.

[S2] M. Pierre-Jean and P. Neuvial. *jointseg: joint segmentation of multivariate (copy number) signals.* R package, R-forge. 2013.

[S3] L. Jacob, P. Neuvial, and S. Dudoit. *DEGraph: Two-sample tests on a graph.* R package, Bioconductor. 2011.

[S4] A. Olshen et al. *PSCBS: Analysis of Parent-Specific DNA Copy Numbers.* R package, CRAN. 2011.

[S5] M. Ortiz-Estevez et al. *CalMaTe: A post-calibration process to improve allele-specific copy number estimates from SNP microarrays.* R package, CRAN. 2011.

[S6] H. Bengtsson and P. Neuvial. *aroma.cn: Analysis of copy-number estimates obtained from various platforms.* R package, aroma-project. 2010.

[S7] H. Bengtsson and P. Neuvial. *aroma.cn.eval: Evaluating copy-number estimates.* R package, aroma-project. 2010.

[S8] M. Elati and P. Neuvial. *LICORN: Learning Co-Operative Regulation Networks.* `CaML` program, `http://www.lri.fr/~elati/licorn.html`. 2008.

[S9] P. La Rosa et al. *VAMP: Visualisation and Analysis of Molecular Profiles.* `http://bioinfo.curie.fr/actudb`. 2006.

[S10] S. Liva et al. *CAPweb: Copy Number Microarray Analysis Platform.* `http://bioinfo.curie.fr/CAPweb`. 2006.

[S11] P. Neuvial and P. Hupé. *MANOR: Micro-Array data NORmalization.* R package, Bioconductor. 2006.