

# PIERRE NEUVIAL

Curriculum vitae

December 12, 2011

## CONTENTS

<b>Curriculum vitae</b>	<b>2</b>
Personal information . . . . .	2
Education . . . . .	2
Professional experience . . . . .	2
<b>Scientific activities</b>	<b>3</b>
Teaching . . . . .	3
Participation to research grants . . . . .	3
Software production . . . . .	3
Reviewer for journals . . . . .	4
<b>Scientific productions</b>	<b>5</b>
Papers . . . . .	5
Talks . . . . .	6
Software . . . . .	8

# CURRICULUM VITAE

## Personal information

---

Married, two children

Born: June 11, 1979

French

[pierre.neuvial@gmail.com](mailto:pierre.neuvial@gmail.com)

<http://neuvial.ensae.net>

Laboratoire Statistique & Génome

Université d'Evry Val d'Essonne

UMR CNRS 8071 — USC INRA

26, boulevard de France

91000 EVRY Cedex

## Education

---

<b>2004-2008</b>	<b>PhD in Applied Mathematics</b>
<i>Title</i>	Contributions to the statistical analysis of DNA microarray data
<i>Advisors</i>	Stéphane Boucheron, LPMA, Université Paris Diderot Emmanuel Barillot, Institut Curie/INSERM U900/Mines ParisTech
	Obtained with the highest honors
<b>2002-2003</b>	<b>M. S. in Stochastic models, Université Paris Diderot (Paris 7)</b>
<b>1999-2003</b>	<b>École Nationale de la Statistique et de l'Administration Économique (ENSAE)</b>

## Professional experience

---

<b>≥ 10/2011</b>	<b>Research scientist at CNRS</b>
<i>Lab</i>	Statistique & Génome (UMR CNRS 8071, Univ. Évry, INRA)
<b>2010-2011</b>	<b>Post-doc in statistics applied to genomics</b>
<i>Lab</i>	Statistique & Génome (UMR CNRS 8071, Univ. Évry, INRA)
<i>Funding</i>	CNRS
<b>10-12/2010</b>	<b>Scientific consultant in genomic data analysis</b>
<i>Client</i>	Johnson & Johnson (via <a href="#">OpenAnalytics</a> )
<i>Project</i>	Statistical analysis of SNP array data from liver and gastric cancers.
<b>2008-2010</b>	<b>Post-doc in statistics applied to genomics</b>
<i>Lab</i>	University of California at Berkeley, Department of Statistics (USA)
<i>Funding</i>	Project “The Cancer Genome Atlas” (TCGA) from NIH
<b>2003-2005</b>	<b>Research engineer in statistics and bioinformatics.</b>
<i>Lab</i>	Institut Curie, Bioinformatics
<b>2001-2002</b>	<b>Statistician (intern)</b>
<i>Company</i>	Crédit Lyonnais, Groupe de Recherche Opérationnelle
<b>2000</b>	<b>Statistician (summer intern)</b>
<i>Company</i>	Dendrite International (New Jersey, USA)

## SCIENTIFIC ACTIVITIES

### Teaching

---

<b>2011-2012</b>	<b>16.5h practicals at Université d'Évry val d'Essonne</b>
<i>Title</i>	Statistics: mostly on statistical tests (M. S. level in biology)
<b>2011-2012</b>	<b>6h course at École Centrale Paris</b>
<i>Title</i>	Statistical methods for molecular biology (M. S. level)
<b>2011-2012</b>	<b>20h course at ENSAE</b>
<i>Title</i>	Statistical methods for molecular biology (M. S. level in statistics) given with Catherine Matias
<b>Feb 2011</b>	<b>10h course at ENSAI, École Nationale de la Statistique et de l'Analyse de l'information</b>
<i>Title</i>	Statistical methods for genomic data analysis (M. S. level in biostatistics)
<b>2005-2008</b>	<b>16h course at ENSAE</b>
<i>Title</i>	Statistical methods for molecular biology (M. S. level in statistics) given with Pierre-Yves Bourguignon
<b>2004-2008</b>	<b>Practicals of mathematical statistics at ENSAE</b>
	M. S. level. 20 to 32 hours per year. Project supervision and grading.
<b>2004-2007</b>	<b>Practicals for M. S. in bioinformatics, Université Paris Diderot (Paris 7)</b>
<i>Subject</i>	DNA copy number analysis

### Participation to research grants

---

#### 2008-2010: The Cancer Genome Atlas (TCGA)

Genome Data Analysis Center (GDAC-A)

[http://cancergenome.nih.gov/wwd/program/research\\_network/gdac.asp](http://cancergenome.nih.gov/wwd/program/research_network/gdac.asp)

#### 2009-2010: "Stand-Up to Cancer" foundation

An Integrated Approach to Targeting Breast Cancer Molecular Subtypes and Their "Resistance" Phenotypes

[http://www.standup2cancer.org/su2c/about\\_us/scientific\\_dream\\_teams/breast](http://www.standup2cancer.org/su2c/about_us/scientific_dream_teams/breast)

#### 2006-2009: Agence Nationale pour la Recherche (ANR), programme blanc

TAMIS: adaptation, multiple testing, ranking and applications

### Software production

---

Contribution to the development and technical support of the two main open source and collaborative projects of bioinformatics tools in R: [Bioconductor](#) (packages MANOR et DEGraph) and [the Aroma Project](#) (packages aroma.cn et aroma.cn.eval).

## Reviewer for journals

---

- Annals of Applied Statistics (2009-2011)
- Journal of Machine Learning Research (2009)
- Statistics in Medicine (2009)
- Biostatistics (2011)
- BMC Bioinformatics (2011)

## Papers

---

### Submitted

- [P1] A. Chambaz, **P. Neuvial**, and M. J. van der Laan. Estimation of a non-parametric variable importance measure of a continuous exposure. Preprint at <http://hal.archives-ouvertes.fr/hal-00629899/en>, Oct. 2011.
- [P2] **P. Neuvial** and E. Roquain. On false discovery rate thresholding for classification under sparsity. Technical report, Preprint at <http://arXiv.org/abs/1106.6147>, June 2011.
- [P3] **P. Neuvial**. Intrinsic bounds and false discovery rate control in multiple testing problems. Preprint at <http://arxiv.org/abs/1003.0747>, Nov. 2009.

### Journal papers

- [P4] L. Jacob, **P. Neuvial**, and S. Dudoit. More power via graph-structured tests for differential expression of gene networks. *Annals of Applied Statistics*, to appear.
- [P5] L. Heiser, A. Sadanandam, W. Kuo, S. Benz, T. Goldstein, S. Ng, W. Gibb, N. Wang, S. Ziyad, F. Tong, N. Bayani, Z. Hu, J. I. Billig, A. Dueregger, S. Lewis, L. Jakkula, J. E. Korkola, S. Durinck, F. Pepin, Y. Guan, E. Purdom, **P. Neuvial**, H. Bengtsson, K. W. Wood, P. G. Smith, L. Vassilev, B. T. Hennessy, J. Greshock, K. E. Bachman, M. A. Hardwicke, J. W. Park, L. J. Marton, D. M. Wolf, E. A. Collisson, R. M. Neve, G. B. Mills, T. P. Speed, H. S. Feiler, R. F. Wooster, D. Haussler, J. M. Stuart, J. W. Gray, and P. T. Spellman. Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proceedings of the National Academy of Sciences*, 2011.
- [P6] The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature*, 474(7353):609–615, jun 2011.
- [P7] A. B. Olshen, H. Bengtsson, **P. Neuvial**, P. T. Spellman, R. A. Olshen, and V. E. Seshan. Parent-specific copy number in paired tumor-normal studies using circular binary segmentation. *Bioinformatics*, 27(15):2038–2046, Aug 2011.
- [P8] 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):245, 2010.
- [P9] H. Noushmehr, D. J. Weisenberger, K. Diefes, H. S. Phillips, K. Pujara, B. P. Berman, F. Pan, C. E. Pelloski, E. P. Sulman, K. P. Bhat, R. G. W. Verhaak, M. L. Meyerson, K. A. Hoadley, D. N. Hayes, C. M. Perou, H. K. Schmidt, L. Ding, R. K. Wilson, D. Van Den Berg, H. Shen, H. Bengtsson, **P. Neuvial**, L. M. Cope, J. Buckley, J. G. Herman, S. B. Baylin, P. W. Laird, and K. Aldape. Identification of a cpG island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell*, 17(5):510–522, April 2010.
- [P10] **P. Neuvial**. Asymptotic properties of false discovery rate controlling procedures under independence. *Electron. J. Statist.*, 2:1065–1110, 2008.
- [P11] M. A. Bollet, N. Servant, **P. Neuvial**, C. Decraene, I. Lebigot, J.-P. Meyniel, Y. De Rycke, A. Savignoni, G. Rigail, P. Hupé, A. Fourquet, B. Sigal-Zafrani, E. Barillot, and J.-P. Thiery. High-resolution mapping of DNA breakpoints to define true recurrences among ipsilateral breast cancers. *J Natl Cancer Inst*, 100(1):48–58, 2008.
- [P12] M. Elati, **P. Neuvial**, M. Bolotin-Fukuhara, E. Barillot, F. Radvanyi, and C. Rouveirol. LICORN: LearnIng COoperative Regulation Networks. *Bioinformatics*, 23(18):2407–2414, 2007.
- [P13] P. La Rosa, E. Viara, P. Hupé, G. Pierron, S. Liva, **P. Neuvial**, I. Brito, S. Lair, N. Servant, N. Robine, E. Manié, C. Brennetot, I. Janoueix-Lerosey, V. Raynal, N. Gruel, C. Rouveirol, N. Stransky, M.-H. Stern, O. Delattre, A. Aurias, F. Radvanyi, and E. Barillot. VAMP: visualization and analysis of array-CGH, transcriptome and other molecular profiles. *Bioinformatics*, 22(17):2066–2073, Sep 2006.

- [P14] S. Liva, P. Hupé, **P. Neuvial**, I. Brito, E. Viara, P. La Rosa, and E. Barillot. CAPweb: a bioinformatics CGH array Analysis Platform. *Nucleic Acids Res*, 34(Web Server issue):477–481, Jul 2006.
- [P15] **P. Neuvial**, P. Hupé, I. Brito, S. Liva, E. Manié, C. Brennetot, F. Radvanyi, A. Aurias, and E. Barillot. Spatial normalization of array-CGH data. *BMC Bioinformatics*, 7(1):264, May 2006.

## Book chapter

- [P16] **P. Neuvial**, H. Bengtsson, and T. P. Speed. Statistical analysis of single nucleotide polymorphism microarrays in cancer studies. In B. Z. H. Lu, Henry Horng-Shing; Schölkopf, editor, *Handbook of Statistical Bioinformatics*, Springer Handbooks of Computational Statistics. Springer, 1st edition, 2011.

## Technical reports

- [P17] I. Brito, P. Hupé, **P. Neuvial**, and E. Barillot. Stability-based comparison of class discovery methods for array-CGH profiles. Submitted to *PLoS One*, December 2010.
- [P18] **P. Neuvial**. *Contributions à l'analyse statistique des données de puces à ADN*. PhD thesis, Institut Curie et Université Paris VII (France), 2008.
- [P19] E. Hauvuy, B. Lebrave, and **P. Neuvial**. Analyse statistique du lien entre les plages homogènes de séquences d'ADN de différentes bactéries. Master's thesis, ENSAE Paristech et Université Paris Diderot, 2003.
- [P20] R. Elie, A. Frachot, P. Georges, and **P. Neuvial**. A model of prepayment for the French residential loan market. Technical report, Groupe de Recherche Opérationnelle, Crédit Lyonnais, France, 2002.

## Popular science

- [P21] **P. Neuvial**. Tests multiples en génomique. *La gazette des mathématiciens*, 130:71–76, Oct. 2011.
- [P22] **P. Neuvial** and P.-Y. Bourguignon. Problématiques statistiques à l'heure de la post-génomique. *Variations*, 35:56–60, February 2009.

## Talks

---

### Contributed talks

- [T23] On false discovery rate thresholding for classification under sparsity. Mathematical Statistics and Applications, Fréjus, France, Sept. 2011.
- [T24] 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.
- [T25] Asymptotic properties of false discovery rate controlling procedures. Mathematical Statistics and Applications, Fréjus, France, Sept. 2008.
- [T26] Intrinsic bounds on the Benjamini-Hochberg multiple comparison procedure. PASCAL workshop on Multiple Simultaneous Hypothesis Testing, Paris, France, May 2007.

### Seminars

- [T27] On false discovery rate thresholding for classification under sparsity. University of Postdam, Germany, Invited by Gilles Blanchard, Dec. 2011.
- [T28] Analyse statistique de données génomiques. Séminaire professionnel à l'ENSAI, Rennes, Nov. 2011.
- [T29] Tests multiples en génomique. Inauguration de la Fédération de Mathématiques de l'Université d'Evry val d'Essonne, Evry, Nov. 2011.
- [T30] On false discovery rate thresholding for classification under sparsity. Séminaire parisien de statistique, Paris, France, May 2011.

- [T31] Statistical analysis of single nucleotide polymorphism microarrays in cancer studies. Stanford Biostatistics Workshop, Invité par Chiara Sabatti, Sept. 2010.
- [T32] Targeted maximum likelihood estimation of the relationship between copy number and gene expression in cancer studies. Statistical Genomics in Biomedical Research Workshop, International, Banff International Research Station for Mathematical Innovation and Discovery, Canada, July 2010.
- [T33] 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.
- Spring 2010.
- [T34] 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.
- [T35] 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.
- [T36] 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.
- [T37] Asymptotic properties of false discovery rate controlling procedures. Groupe de travail "Statistique des systèmes biologiques", Paris, France. Invité par Sophie Schbath, Feb. 2008.
- [T38] Propriétés asymptotiques de procédures de contrôle du false discovery rate. Groupe de travail "Mathématiques Appliquées à Paris V", Paris, France. Invité par Avner Bar-Hen, Jan. 2008.
- [T39] 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.
- [T40] 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.
- [T41] Statistical analysis of copy number and expression microarray data - application to cancer study. Netherlands Cancer Institute (Amsterdam, Pays-Bas). Invité par Mark Van de Wiel, Oct. 2007.
- [T42] 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.

## Software

---

- [S43] L. Jacob, **P. Neuvial**, and S. Dudoit. DEGraph: Two-sample tests on a graph. R package, [Bioconductor](#).
- [S44] H. Bengtsson and **P. Neuvial**. aroma.cn: Analysis of copy-number estimates obtained from various platforms. R package, [aroma-project](#).
- [S45] H. Bengtsson and **P. Neuvial**. aroma.cn.eval: Evaluating copy-number estimates. R package, [aroma-project](#).
- [S46] **P. Neuvial** and P. Gestraud. GTCA: Genome-transcriptome correlation analysis.
- [S47] M. Elati and **P. Neuvial**. LICORN: Learning Co-Operative Regulation Networks. CaML program, <http://www.lri.fr/~elati/licorn.html>.
- [S48] P. La Rosa et al. VAMP: Visualisation and Analysis of Molecular Profiles. <http://bioinfo.curie.fr/actudb>.
- [S49] S. Liva et al. CAPweb: Copy Number Microarray Analysis Platform. <http://bioinfo.curie.fr/CAPweb>.
- [S50] **P. Neuvial** and P. Hupé. MANOR: Micro-Array data NORmalization. R package, [Bioconductor](#).