

## Jean-Philippe Vert

Professor, Department of Mathematics and Applications, Ecole normale supérieure, 45 rue d'Ulm, 75005 Paris  
Director, Centre for Computational Biology, Mines ParisTech, 60 bd Saint-Michel, 75006 Paris  
Team leader, Institut Curie, 26 rue d'Ulm, 75005 Paris

French nationality

Tel: +33 6 5087 6531

Born January 18th, 1973 in Paris (France)

E-mail: [Jean-Philippe.Vert@mines.org](mailto:Jean-Philippe.Vert@mines.org)

Married, three children

Web: <http://cbio.mines-paristech.fr/~jvert>

### EDUCATION

---

- 2004 Research habilitation (HDR), Paris 6 University, Paris, France.  
Dissertation “Kernel methods in computational biology”.
- 2001 PhD in Mathematics, Paris 6 University and Ecole normale supérieure de Paris, Paris, France.  
Dissertation: “Statistical methods for natural language modeling”.  
Advisor: Olivier Catoni, obtained with the highest honors.
- 1998 Master of Public Administration (M.P.A.), Corps des Mines, Paris, France.
- 1997 M.S. in Mathematics, Paris 6 University, Paris, France.
- 1995 B.S. in Mathematics, Ecole Polytechnique, Palaiseau, France.

### PROFESSIONAL EXPERIENCE

---

- 2016-present Professor, Department of Mathematics and Applications, Ecole Normale Supérieure, Paris, France
- 2014-present Senior researcher (*Directeur de recherche*, eq. research professor) and Director, Centre for Computational Biology, Mines ParisTech, Paris, France
- 2008-present Deputy director, “Cancer Computational Genomics, Bioinformatics, Biostatistics and Epidemiology” laboratory, Mines ParisTech / Institut Curie / INSERM U900, Paris, France
- 2006-2014 Senior researcher (*Maître de recherche*, eq. associate research professor) and director, Centre for Computational Biology, Mines ParisTech, Paris, France
- 2002 - 2005 Junior researcher (*Chargé de recherche*, eq. assistant research professor) and group leader, Ecole des Mines de Paris, Fontainebleau, France
- 2001 - 2002 Associate researcher (post-doc), Bioinformatics Center, Kyoto University, Kyoto, Japan
- 1999 - 2000 Scientific consultant (natural language processing and statistics), Sudimage S.A., Cachan, France
- 1996 - 1997 Research scientist (statistics), Elf Atochem North America, Philadelphia, USA
- 1995 - 1996 Consultant (strategy), Matra Automobile S.A., Romorantin, France
- 1994 Summer intern, Department of Mathematics, Kyoto University, Kyoto, Japan
- 1994 Summer intern, Hamamatsu Photonics, Hamamatsu, Japan
- 1992 - 1993 Military Service as officer and platoon leader, Monthlery, France

---

<sup>1</sup>Last update: Feb 2018

## HONORS AND AWARDS

---

- 2017 Pierre Faurre grand prize of the French Academy of Sciences
- 2017 F1000 Faculty member
- 2016 Fulbright Visiting Scholar, Department of Statistics, UC Berkeley
- 2015 Visiting Miller Research Professor, Department of Statistics, UC Berkeley
- 2013 Second best performer at the DREAM 8 NHIES-NCATS-UNC Toxicogenetics challenge (team leader)
- 2011 European Research Council (ERC) Starting Grant
- 2010 Honorable mention at the DREAM5 in silico network inference challenge (team leader)
- 2006 Bronze medal of the National Center for Scientific Research (CNRS)
- 2004 Simon Regnier prize of the Francophone Society for Classification

## PROFESSIONAL ACTIVITIES

---

### Editorial Boards:

- Journal of Machine Learning Research (since 2009)
- Machine Learning (Editorial Board 2010-2014, Action Editor since 2014)
- BMC Bioinformatics (since 2010, Section Editor since 2015)
- IEEE/ACM Transactions on Computational Biology and Bioinformatics (since 2011)
- PLoS Computational Biology (Guest Associate Editor, since 2014)
- De Gruyter series in mathematics and life sciences (Editor, since 2017)

### Area Chair and Senior Program Committee member:

- AISTATS 2012, 2014, 2015; ICANN 2011; ICML 2014-2015, 2018; IJCAI 2013; ISMB 2016; NIPS 2011-2012, 2014, 2016-2017; ECML-PKDD 2012, 2014.

### Program Committee member:

- AAAI 2010; ACML 2010; ACM SIGKDD 2013; AISTATS 2009,2013; APBC 2008-2011; BIBE 2009; BIOINFORMATICS 2010; BIRD 2007, 2008; CAP 2011-2013, 2018; CIBB 2007-2009, 2013; COLT 2003-2009; DREAM Idea Challenge 2017; ECCB 2005-2012; ECML/ PKDD 2006-2011; ESANN 2010-2011; GIW 2004-2009; HiTSeq 2015; ICML 2004-2013, 2017; ICLR 2018; IPG 2009-2010; ISBRA 2009-2010; ISMB 2004-2018; JOBIM 2009-2012; KDD 2011, 2013; KRBIO 2005; MLCB 2008-2010, 2015-2017; MLG 2007-2009; MLGG 2012; MLSB 2007-2012, 2014, 2017; NIPS 2003-2015; PMSB 2006; PRIB 2013; PSB 2015-2016; RECOMB 2014, 2018; RxDM 2009; SMPGD 2009.

### Reviewer for journals:

- Annals of Applied Statistics, Annals of Statistics, Applied and Computational Harmonic Analysis, Artificial Intelligence in Medicine, Bioinformatics, BMC Bioinformatics, BMC Genomics, BMC Systems Biology, Briefings in Bioinformatics, Combinatorial Chemistry and High Throughput Screening, Discrete Applied Mathematics, EURASIP Journal on Advances in Signal Processing, EURASIP Journal on Bioinformatics and Systems Biology, Genome Biology, IEEE Transactions on Information Theory, IEEE IEEE/ACM Transactions on Computational Biology and Bioinformatics, International Journal of Computer Vision, International Journal of Data Mining and Bioinformatics, International Journal of Knowledge Discovery in Bioinformatics, Journal of Applied Statistics, Journal of Bioinformatics and Computational Biology, Journal of Biomedical Informatics, Journal of Chemical Information and Modeling, Journal of Computational Biology, Journal of Computer Science and Technology, Journal of Machine Learning Research, Journal of the American Statistical Association, Journal of the Royal Statistical Society, Journal of Theoretical Biology, Machine Learning, Neural Networks, Neurocomputing, Neuroinformatics, Nucleic Acids Research, Pattern Recognition, PLoS Computational Biology, Proteins.

Reviewer for funding or evaluation agencies:

- Austrian Science Fund (FWF), Dutch National Science Foundation (NWO), European Innovative Medicines Initiative (IMI), European Research Council (ERC), French Agency for Evaluation of Research and Higher Education (AERES), French National Research Agency (ANR), Genome Canada, Israel Science Foundation (ISF), King Abdullah University of Science and Technology (KAUST), Luxembourg National Research Fund (FNR), Max Planck Society (MPS), Natural Sciences and Engineering Research Council of Canada (NSERC), Research Grant Council (RGC) of Hong Kong, Science Foundation Ireland (SFI), Swiss National Science Foundation (SNSF), Swiss Cancer League.

Workshop organization:

- RECOMB satellite workshop Kernel Methods in Computational Biology, Berlin, Germany (2003)
- NIPS workshop on Machine Learning in Computational Biology (MLCB), Whistler, Canada (2005, 2009, 2010), Granada, Spain (2011), Lake Tahoe, USA (2012, 2013)
- International workshop on Machine Learning in Systems Biology (MLSB), Berlin, Germany (2013)
- Cancer bioinformatics workshop, Cambridge, UK (2010).
- Paris Cancer Bioinformatics workshop, France (2012)
- Machine Learning for Personalized Medicine summer school, Paris, France (2014).

Other:

- Member of the European Network of Excellence PASCAL and PASCAL2 (2003-2013).
- Chairman of the Japan-France Frontier of Science (JFFoS) conference, Tokyo, Japan (2011).
- Member of the Scientific Committee, Paris Ile-de-France Canceropole (since 2010).
- Member of the Scientific Advisory Board, "Machine learning for big data" chair, Telecom ParisTech (since 2014).
- Member of the Scientific Advisory Board, FHU-OncoAge project (since 2016)
- Member of the Scientific Advisory Board, EPSRC-funded project "StatScale: Statistical Scalability for Streaming Data", Lancaster University, UK (since 2016).
- Scientific Advisor, SANCARE (since 2016)
- Scientific Advisor, OWKIN (since 2016)

- Member of the working group on Artificial Intelligence, French Ministry of Innovation and Digital Affairs (2017)
- Member of the scientific committee, Epidemium open science project for cancer epidemiology (since 2017)
- Member of the "Mathematics, Big data and Cancer" working group at the French Academy of Sciences and French Academy of Medicine (since 2017)
- Industry consulting: Alcon Labs, Genentech, Novartis (USA), ADCIS, WatchFrog (France).

## SELECTED RESEARCH FUNDING (AS PI OR CO-PI)

---

- 2015-2019 French National Research Agency (ANR) ANR-15-CE13-0012-03. CRESTNETMETABO: Novel Challenges in the Neural Crest Early Gene Regulatory Network. (own funding: 80 K€).
- 2012-2016 European Commission FP7-PEOPLE-2012-ITN. MLPM: Machine Learning for Personalised Medicine (own funding: 528 K€).
- 2012-2015 European Commission FP7-HEALTH-2012-Innovation-1. RADIANT: Rapid development and distribution of statistical tools for high-throughput sequencing (own funding: 315 K€).
- 2012-2016 French National Research Agency (ANR) "Investments for the future" program. ABS4NGS: Algorithms, Bioinformatics and Softwares for Next Generation Sequencing. (own funding: 243 K€).
- 2012-2014 Cancéropôle Ile-de-France program "Data integration for cancer research". (own funding: 50 K€).
- 2012-2014 French National Research Agency (ANR) ANR-11-BSV2-0013. CRESTNET: Building regulatory networks in neural crest induction: integrative approaches in vivo and in stem cells. (own funding: 91 K€).
- 2012-2013 French National Research Agency (ANR) Emergence 2011. AP'ONCALYPSE: Validation of an immune signature predicting a therapeutic response to anthracyclines in breast cancer. (own funding: 57 K€).
- 2011-2016 European Research Council (ERC) Starting Grant. SMAC: Statistical machine learning for complex biological data (own funding: 1,496 K€).
- 2011-2016 The French League Against Cancer. Integrated analysis of methylation profiles in breast cancer (own funding: 75 K€).
- 2010-2015 European Commission FP7-NMP-2009-LRAGE-3. NADINE: Nanosystems for early diagnosis of neurodegenerative diseases (own funding: 144 K€).
- 2009-2013 French National Research Agency ANR-09-BLAN-0051-04. CLARA: Clustering in high dimension, algorithms and applications (68 K€).
- 2008-2010 JSPS-INSERM Japan-France grant. Development of algorithms and databases in cancer informatics (40 K€).
- 2007-2011 French Ministry of Economy, Finance and Industry DGE-07-2-90-6473. RAMIS: High-resolution imaging for the screening of anti-cancer drugs (165 K€).
- 2007-2010 French National Research Agency ANR-07-BLAN-0311-03. MGA: Graphical models and applications (50 K€).
- 2007-2009 France-Berkeley fund. Inference and learning in dynamic graphical models, with applications in speech and bio-informatics (10 K\$).
- 2006-2009 French Ministry of Economy, Finance and Industry DGE-06-2-90-6056. BIOTYPE: Multidimensional molecular and cellular biotyping (123 K€).
- 2005-2008 European Commission FP6-2004-IST-NMP-2. INDIGO: Integrated highly sensitive fluorescence-based biosensor for diagnosis applications (64 K€).

- 2005-2007 European Commission LSH-2004-1.1.0-2. ESBIC-D: European Systems Biology Initiative for combating Complex Diseases (24 K€)
- 2004-2007 NIH R33HG003070-01: Detecting Relations Among Heterogeneous Datasets (215 K\$)

## SELECTED PUBLICATIONS IN JOURNALS AND PEER-REVIEWED CONFERENCES<sup>2</sup>

---

- [1] D. Marbach, J.C. Costello, R. Küffner, N.M. Vega, R.J. Prill, D.M. Camacho, K.R. Allison, The DREAM5 Consortium, M. Kellis, J.J. Collins and G. Stolovitzky. Wisdom of crowds for robust gene network inference. *Nature Methods*, 9:796-804, 2012. **772 citations**.
- [2] L. Jacob, G. Obozinski and J.-P. Vert. Group Lasso with overlap and graph Lasso. *Proceedings of the Twenty-Sixth International Conference on Machine Learning (ICML)*, 433–440. ACM Press, 2009. **630 citations**.
- [3] H. Saigo, J.-P. Vert, N. Ueda, and T. Akutsu. Protein homology detection using string alignment kernels. *Bioinformatics*, 20(11):1682–1689, 2004. **391 citations**.
- [4] L. Jacob, F. Bach and J.-P. Vert. Clustered multi-task learning: A convex formulation. *Adv. Neural Inform. Process. Syst. (NIPS)*, volume 20, 745–752. MIT Press, Cambridge, MA, 2008. **285 citations**.
- [5] F. Rapaport, A. Zynoviev, M. Dutreix, E. Barillot, and J.-P. Vert. Classification of microarray data using gene networks. *BMC Bioinformatics*, 8:35, 2007. **249 citations**.
- [6] M. Zaslavskiy, F. Bach and J.-P. Vert. A path following algorithm for the graph matching problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(12):2227-2242, 2009. **241 citations**.
- [7] Y. Yamanishi, J.-P. Vert, and M. Kanehisa. Protein network inference from multiple genomic data: a supervised approach. *Bioinformatics*, 20:i363–i370, 2004. **236 citations**.
- [8] L. Jacob and J.-P. Vert. Protein-ligand interaction prediction: an improved chemogenomics approach. *Bioinformatics*, 24(19):2149-2156, 2008. **223 citations**.
- [9] J. Abernethy, F. Bach, T. Evgeniou and J.-P. Vert. A New Approach to Collaborative Filtering: Operator Estimation with Spectral Regularization. *J. Mach. Learn. Res.*, 10:803–826, 2009. **220 citations**.
- [10] P. Mahé and J.-P. Vert. Graph kernels based on tree patterns for molecules. *Machine Learning*, 75(1):3-35, 2009. **206 citations**.
- [11] J.-P. Vert, N. Foveau, C. Lajaunie, and Y. Vandenbrouck. An accurate and interpretable model for siRNA efficacy prediction. *BMC Bioinformatics*, 7:520, 2006. **202 citations**.
- [12] J. C. Costello, L. M. Heiser, E. Georgii, M. Gönen, M. P Menden, N. J Wang, M. Bansal, M. Ammad-ud-din, P. Hintsanen, S. A Khan, J.-P. Mpindi, O. Kallioniemi, A. Honkela, T. Aittokallio, K. Wennerberg, NCI DREAM Community, J. J. Collins, D. Gallahan, D. Singer, J. Saez-Rodriguez, S. Kaski, J. W. Gray and G. Stolovitzky. A community effort to assess and improve drug sensitivity prediction algorithms *Nature Biotechnology*, 32:1202-1212, 2014. **195 citations**.
- [13] A.-C. Haury, P. Gestraud and J.-P. Vert. The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures. *PLoS One*, 6(12):e28210, 2011. **193 citations**.
- [14] P. Mahé, N. Ueda, T. Akutsu, J.-L. Perret, and J.-P. Vert. Graph kernels for molecular structure-activity relationship analysis with support vector machines. *J. Chem. Inf. Model.*, 45(4):939–51, 2005. **183 citations**.
- [15] P. Mahé, N. Ueda, T. Akutsu, J.-L. Perret, and J.-P. Vert. Extensions of marginalized graph kernels. *Proceedings of the Twenty-First International Conference on Machine Learning (ICML)*, 552–559. ACM Press, 2004. **178 citations**.
- [16] J.-P. Vert, K. Tsuda and B. Schölkopf. A primer on kernel methods. in B. Schölkopf, K. Tsuda and J.-P. Vert (Eds), *Kernel Methods in Computational Biology*, 35–70, 2003. **175 citations**.
- [17] M. Zaslavskiy, F. Bach and J.-P. Vert. Global alignment of protein-protein interaction networks by graph matching methods. *Bioinformatics*, 25(12):i259-1267, 2009. **174 citations**.
- [18] J.-P. Vert. A tree kernel to analyze phylogenetic profiles. *Bioinformatics*, 18:S276–S284, 2002. **171 citations**.
- [19] V. Boeva, A. Zinovyev, K. Bleakley, J.-P. Vert, I. Janoueix-Lerosey, O. Delattre and E. Barillot. Control-free calling of copy number alterations in deep-sequencing data using GC-content normalization. *Bioinformatics*,

<sup>2</sup>Out of **110 peer-reviewed publications**, with a total of **10,000+ citations (h-index=56)**. Complete list at <http://cbio.mines-paristech.fr/~jvert/publi>

27(2):267-268, 2011. **163 citations.**

[20] A.-C. Hauray, F. Mordelet, P. Vera-Licona and J.-P. Vert. TIGRESS: trustful inference of gene regulation using stability selection. *BMC Systems Biology*, 6:145, 2012. **147 citations.**

[21] M. Cuturi, J.-P. Vert, T. Birkenes and T. Matsui. A kernel for time series based on global alignment. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2:413–416, 2007. **144 citations.**

[22] R. Vert and J.-P. Vert. Consistency and convergence rates of one-class SVMs and related algorithms. *J. Mach. Learn. Res.*, 7:817–854, 2006. **144 citations.**

[23] A. Matsuda, J.-P. Vert, H. Saigo, N. Ueda, H. Toh, and T. Akutsu. A novel representation of protein sequences for prediction of subcellular location using support vector machines. *Protein Sci.*, 14(11):2804–2813, 2005. **142 citations.**

[24] F. Mordelet and J.-P. Vert. SIRENE: Supervised Inference of REgulatory NETworks. *Bioinformatics*, 24(16):i76–i82, 2008. **131 citations.**

[25] Y. Yamanishi, J.-P. Vert, A. Nakaya, and M. Kanehisa. Extraction of correlated gene clusters from multiple genomic data by generalized kernel canonical correlation analysis. *Bioinformatics*, 19:i323–i330, 2003. **123 citations.**

[26] J.-P. Vert and M. Kanehisa. Graph-driven features extraction from microarray data using diffusion kernels and kernel CCA. *Adv. Neural Inform. Process. Syst. (NIPS)*, 1449–1456. MIT Press, 2003. **111 citations.**

[27] C. Houdayer, ..., J.-P. Vert, ... and D. Stoppa-Lyonnet. Guidelines for splicing analysis in molecular diagnosis derived from a set of 327 combined in silico/in vitro studies on BRCA1 and BRCA2 variants. *Human mutation*, 33(8):1228–1238, 2012. **1004 citations.**

[28] K. Bleakley, G. Biau and J.-P. Vert. Supervised reconstruction of biological networks with local models. *Bioinformatics*, 23(13):i57–i65, 2007. **104 citations.**

[29] Y. Yamanishi, J.-P. Vert, and M. Kanehisa. Supervised enzyme network inference from the integration of genomic data and chemical information. *Bioinformatics*, 21:i468–i477, 2005. **101 citations.**

[30] J.-P. Vert and Y. Yamanishi. Supervised graph inference. *Adv. Neural Inform. Process. Syst. (NIPS)*, volume 17, 1433–1440. MIT Press, Cambridge, MA, 2005. **96 citations.**

[30] M. Cuturi, K. Fukumizu and J.-P. Vert. Semigroup Kernels on Measures. *J. Mach. Learn. Res.*, 6:1169–1198, 2005. **95 citations.**

## RESEARCH MONOGRAPHS

---

[31] B. Schölkopf, K. Tsuda, and J.-P. Vert. *Kernel Methods in Computational Biology*. MIT Press, Cambridge, Massachusetts, 2004. **770 citations.**

[32] E. Barillot, L. Calzone, P. Hupé, J.-P. Vert and A. Zinovyev. *Computational Systems Biology of Cancer*. CRC Press, 2012. **67 citations.**

## SELECTED INVITED PRESENTATIONS TO INTERNATIONAL WORKSHOPS, CONFERENCES, SUMMER SCHOOLS<sup>3</sup>

---

Statistics seminar, Imperial College, London, UK, 2017.

Emerging Topics in Biological Networks and Systems Biology symposium, Swedish Collegium for Advanced Study, Uppsala, Sweden, 2017.

1st France-Japan machine learning workshop, ENS Paris, France, 2017.

8th SFdS International Meeting on Statistical Methods in Biopharmacy, Paris, France, 2017.

”Mathematical methods of modern statistics” conference, CIRM, Luminy, France, 2017.

”Algorithmic challenges in genomics” reunion workshop, Simons Institute, Berkeley, USA, 2017.

---

<sup>3</sup>Complete list available at <http://cbio.ensmp.fr/~jvert/talks>

Bioinformatics and Systems Biology seminar, Bioinformatics Institute Ghent BIG N2N, Ghent, Belgium, 2017.  
 BIRS workshop on Statistical and Computational Challenges In Large-Scale Molecular Biology, Banff, Canada, 2017.  
 Distinguished speaker seminar, Department of Statistics, University of Oxford, Oxford, UK, 2017.  
 Distinguished speaker colloquium, Max Planck Institute for Informatics, Saarbrücken, Germany, 2017.  
 Krupp symposium From Machine Learning to Personalized Medicine, Max Planck Institute, Munich, Germany, 2016.  
 Data Science Colloquium, Ecole normale supérieure, Paris, France, 2016.  
 Machine learning for genome precision medicine workshop, KAIST, Daejeon, Korea, 2016.  
 Computational cancer biology workshop, Simons Institute, Berkeley, USA, 2016.  
 Nonparametric Methods for Large Scale Representation Learning NIPS workshop, Montreal, Canada, 2015.  
 Festival of Genomics, San Mateo, USA, 2015.  
 BIRS workshop on Statistical and Computational Challenges In Bridging Functional Genomics, Epigenomics, Molecular QTLs, and Disease Genetics, Banff, Canada, 2015.  
 Optimization and statistical learning workshop, Les Houches, France, 2015.  
 Sparse Representations, Numerical Linear Algebra, and Optimization workshop, Banff, Canada, 2014.  
 Genome Science: Biology, Technology and Bioinformatics workshop, Oxford, UK, 2014.  
 46th annual days of Statistics, Rennes, France, 2014.  
 Machine learning summer school (MLSS), Reykjavik, Iceland, 2014.  
 MascotNum workshop, ETH Zurich, Switzerland, 2014.  
 Kernel methods for big data workshop, Polytech'Lille, Lille, France, 2014.  
 Inference for Change-Point and Related Processes workshop, Cambridge, UK, 2014.  
 Machine Learning for Personalized Medicine summer school, Tuebingen, Germany, 2013.  
 International workshop on Machine Learning and Applications in Biology (MLAB), Sapporo, Japan, 2013.  
 13th Annual International workshop on Bioinformatics and Systems Biology (IBIS), Kyoto, Japan, 2013.  
 8th IAPR International Conference on Pattern Recognition in Bioinformatics (PRIB), Nice, France, 2013.  
 Congress of applied and industrial mathematics SMAI, Seignosse, France, 2013.  
 Optimization and statistical learning workshop, Les Houches, France, 2013.  
 15th Information-Based Inductive Science workshop (IBIS 2012), Tokyo, Japan, 2012 (keynote).  
 "From phenotypes to pathways" workshop, Cambridge, UK, 2012.  
 2012 Sapporo workshop on machine learning and applications to biology, Sapporo, Japan, 2012.  
 8th world congress in Probability and Statistics, Istanbul, Turkey, 2012.  
 IMA workshop on machine learning: Theory and computation, Minneapolis, USA, 2012.  
 Modeling and analysis in life science conference, Tokyo, Japan, 2011.  
 Machine Learning for Neuroimaging workshop, Marseille, France, 2011.  
 "From phenotypes to pathways" ESF exploratory workshop, Cambridge, UK, 2011.  
 Informatics and mathematics sciences: interactions with biomedical sciences, Paris, France, 2011.  
 StatLearn 11, workshop on Challenging Problems in Statistical Learning, Grenoble, France, 2011.  
 NIPS 2010 workshop on machine learning for computational chemistry, Whistler, Canada, 2010.  
 Malaysia Genome Institute, Kuala Lumpur, Malaysia, 2010.  
 50th workshop on optimization, machine learning and bioinformatics, Erice, Italy, 2010.  
 "From phenotypes to pathways" ESF exploratory workshop, Cambridge, UK, 2010.  
 10th annual International Workshop on Bioinformatics and Systems Biology (IBSB 2010), Kyoto, Japan, 2010.  
 Statistical Genomics in Biomedical Research workshop, Banff, Canada, 2010.  
 2nd Strasbourg Summer School on Chemoinformatics, Obernai, France, 2010.  
 1st spring school on machine learning (EPAT 2010), Cap Hornu, France, 2010.  
 NIPS 2009 Workshop: understanding multiple kernel learning methods, Whistler, Canada, 2009.  
 The third school on Analysis of Patterns, Cagliari, Italy, 2009.  
 27th European Meeting of Statisticians (EMS 2009), Toulouse, France, 2009.  
 "Advances in the Theory of Control, Signals and Systems" workshop, Lausanne, Switzerland, 2009.  
 6th International Workshop on Computational Systems Biology, (WCSB 2009), Aarhus, Denmark, 2009 (keynote).  
 "Statistical advances in Genome-scale Data Analysis" Workshop, Ascona, Switzerland, 2009.  
 5th Medchem Europe conference, Berlin, Germany, 2009.  
 3rd Japanese-French Frontiers of Science (JFFoS) Symposium, Tokyo, Japan, 2009.  
 2nd Canada-France congress of mathematics, Montreal, Canada, 2008.

10th European Symposium on Statistical Methods for the Food Industry, Louvain-la-Neuve, Belgium, 2008 (keynote).  
 The International Workshop on Data Mining and Statistical Science (DMSS 2007), Tokyo, Japan, 2007 (keynote).  
 VI Colloquium Chemiometricum Mediterraneum, Saint-Maximin, France, 2007.  
 French Conference on Bioinformatics (JOBIM 2007), Marseille, France, 2007 (keynote).  
 6th Workshop on Graph-based Representations in Pattern Recognition (Gbr 2007), Alicante, Spain, 2007 (keynote).  
 Machine Learning in Systems Biology Conference (MLSB 2007), Evry, France, 2007.  
 International conference on Embeddings of Graphs and Groups into Hilbert spaces, Lausanne, Switzerland, 2007.  
 Current Challenges in Kernel Methods workshop (CCKM'06), Brussel, Belgium, 2006.  
 International Conference on Grammatical Inference (ICGI 2006), Tokyo, Japan, 2006 (keynote).  
 Course at the Statistical Mathematics and Application Workshop, Luminy, France, 2006.  
 Course at the Machine Learning Summer School, Taipei, Taiwan, 2006.  
 Second Conference "Mathematical foundations of learning theory", Paris, France, 2006.  
 Workshop on Knowledge Discovery and Emergent Complexity in Bioinformatics, Ghent, Belgium, 2006.  
 "Kernel methods and structured domain" workshop, Whistler, Canada, 2005.  
 50th NIBB conference on "Structure and Dynamics of Complex Biological Networks, Okazaki, Japan, 2005.  
 Symposium on perspectives in computational and theoretical biology, Shanghai, P.R. China, 2004.  
 "Complex stochastic systems in biology and medicine" workshop, Munich, Germany, 2004.  
 "Advanced microarray data analysis" course, Center for Biological Sequence Analysis, Elsinore, Denmark, 2004.  
 The learning workshop, Snowbird, Utah, USA, 2004 (keynote).  
 Machine Learning in Bioinformatics conference, Brussels, Belgium, 2003.  
 AIM "Geometric models of biological phenomena" workshop, Palo Alto, CA, USA, 2003.  
 Workshop "Mathematical aspects of molecular biology : Towards new constructions", Nara, Japan, 2003.  
 Workshop "Statistical Learning in Classification and Model Selection", Eindhoven, The Netherlands, 2003.

## TEACHING

---

- 2017-2018: "Large-scale machine learning", 35h/year (organizer), MINES ParisTech, Paris, France.
- 2004-2015, 2017-2018: "Machine learning with kernels methods", 20h/year, ENS Cachan (MSc course), Cachan, France.
- 2012-2015, 2017: "Machine learning in computational biology", 16h/year, ENSAE (undergrad), Paris, France.
- 2010-2014: "Machine learning in computational biology", 4h/year, Mines ParisTech (undergrad), Paris, France.
- 2011-2014: "Kernel methods in drug design", 4h/year, Paris 7 University (MSc course), Paris, France.
- 2013: "Machine learning for personalized genomics", 3h, MLPM summer school (PhD course), Tübingen, Germany.
- 2011-2013 : "Machine learning in bioinformatics and drug design", 4h/year, ENS Paris (MSc course), Paris, France.
- 2011: "Statistical learning", 20h, Gaston Berger University (MSc course), Saint-Louis, Senegal.
- 2010: "Machine learning in computational systems biology", 30h, Malaysia Genome Institute, Malaysia.
- 2009: "Supervised classification for structured data", 3h, The analysis of pattern summer school, Cagliari, Italy.
- 2004-2007: "Kernel methods in computational biology", 12h/year, Paris 6 university (MSc course), Paris, France.
- 2006: "Kernel methods in computational biology", 8h, Statistical mathematics workshop, Luminy, France.
- 2006: "Kernel methods in computational biology", 4h, Machine Learning Summer School, Taipei, Taiwan.
- 2006: "Nonlinear programming", 20h, INSEAD (PhD course), Fontainebleau, France.
- 2004: "Support vector machines in computational genomics", 4h, Telecom Paris (undergrad), Paris, France.
- 2004: "Support vector machines in computational biology", 6h, Tokyo University, Tokyo, Japan.

## PHD STUDENTS (WITH PERCENTAGE OF SUPERVISION)

---

- Romain Ménégaux (2017-present): Large-scale machine learning for sequencing read classification (100%)
- Lotfi Slim (2016-present): Epistasis detection from GWAS studies (50%)
- Judith Abecassis (2016-present): Inference of intra-tumor heterogeneity in triple-negative breast cancer (50%)



- Beyrem Khalfaoui (2014-present): Machine learning for drug response prediction (100%)
- Marine Le Morvan (2014-now): Modeling of cancer somatic mutations (50%).
- Yunlong Jiao (2013-present): Robust feature selection for molecular prognosis in cancer (100%).
- Alice Schoenauer-Sebag (2012-2015): Toxicity detection from video-microscopy (10%).
- Elsa Bernard (2012-2016): Sparse machine learning for RNA-seq data analysis (100%).
- Nelle Varoquaux (2012-2015): Modelling DNA 3D structure (100%).
- Kevin Vervier (2011-2015): Structured output machine learning for metagenomics (100%).
- Matahi Moarii (2011-2015): Analysis of methylation in breast cancer (50%).
- Pierre Chiche (2010-2014): Methods to call genomic variations from next generation sequencing data (100%).
- Toby Dylan Hocking (2009-2012): Sparse structure methods for bioinformatics and computer vision (50%).
- Anne-Claire Haury (2009-2012): Feature selection from gene expression data (100%).
- Fantine Mordelet (2007-2010): Supervised inference of biological networks (100%).
- Mikhail Zaslavskiy (2006-2010): Graph matching for machine learning (50%).
- Laurent Jacob (2006-2009): Multitask learning in bio-, chemo- and immuno-informatics (100%).
- Martial Hue (2004-2011): Semi-supervised learning and classification of protein structures (100%).
- Franck Rapaport (2004-2008): Integration of gene networks and microarray data for cancer research (50%).
- Joannès Vermorel (2004-2006): Large-scale learning algorithms (100%).
- Pierre Mahé (2003-2006): Kernel methods in virtual screening (100%).
- Marco Cuturi (2002-2005): Learning from structured objects with semigroup kernels (100%).