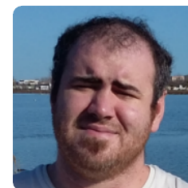


Bastien Cazaux

Curriculum Vitae

☎ (+33) 6 73 85 42 24
✉ bast.cazaux@gmail.com



Education

- 2020–2020 **Ingénieur biologiste en analyse de données du 25 mai 2020 au 9 août 2020**, CNRS, Montpellier, France.
- 2018–2019 **Post-doctorat en Informatique**, *University of Helsinki*, Helsinki, Finlande.
- 2017–2017 **Post-doctorat en Informatique**, *Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (LIRMM)*, Montpellier, France.
- 2014–2016 **Thèse en Bioinformatique**, *Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (LIRMM)*, Montpellier.
Intitulé : Approximation de superchaîne, indexation et assemblage de génome
Date : 7 Décembre 2016
Directeur de thèse : Eric Rivals
Rapporteurs :
— Irena Rusu
— Maxime Crochemore
Examineurs :
— Guillaume Blin
— Christophe Paul
- 2010–2013 **Master de Mathématiques option Informatique**, *Université de Montpellier 2*, Montpellier.
Intitulé : Reverse Engineering de l'arbre des suffixes
Directeur de stage : Eric Rivals
- 2008–2009 **Masters de Mathématiques et Probabilité**, *Université des sciences de Poitiers*, Poitiers.
- 2007–2008 **Licence de Mathématiques**, *Université des sciences de Poitiers*, Poitiers.
- 2006–2007 **Classe MP***, *Classe préparatoire Camille Guérin*, Poitiers.
- 2005–2006 **Classe MPSI**, *Classe préparatoire Camille Guérin*, Poitiers.
- 2005 **Baccalauréat section S option Biologie spécialité Mathématiques**, *Lycée Jean Moulin*, Thouars, Mention Assez Bien.

Visites de laboratoire

- Novembre 2019 **Travail avec Mathieu Raffinot et Raluca Uricaru**, *Université de Bordeaux*, Bordeaux, France.
- Mars 2019 **Travail avec Gonzalo Navarro et Travis Gagie**, *University of Chile*, Santiago, Chile.

Activités d'enseignement

Informatique	L1	22.5h de TD et 27 de TP (Français)
Algorithmes du texte	M1	8h de cours (Français)
Informatique	Master	36h de TD (Anglais)

Organisations

Organization	Seqbio 2013 Seqbio 2014 Workshop on Data Structures in Bioinformatics 2014
--------------	--

Co-supervision

Stage de Master	Katri Maljanen, 2019 University of Helsinki
Stage de M2	Samuel Juhel, 2017 Université de Montpellier
Stage de M2	Quentin Desmettre, 2015 ISIMA
Stage de L3	Jin Li, 2014, ENS Cachan

Logiciels

AQUA PONY	www.atgc-montpellier.fr/aquapony/
PONY-Tree	www.ponytree.org/

Liste des publications

Revue d'audience internationale avec comité de rédaction

- J1 **Bastien Cazaux** and Eric Rivals. Hierarchical Overlap Graph. *Information Processing Letters*, 105862, **2019**. DOI : [10.1016/j.ipl.2019.105862](https://doi.org/10.1016/j.ipl.2019.105862)
- J2 **[article joint au dossier]** Tuukka Norri, **Bastien Cazaux**, Dmitry Kosolobov, Veli Mäkinen. Linear time minimum segmentation enables scalable founder reconstruction. *Algorithms for Molecular Biology* 14, 12 :1–15, **2019**. DOI : [10.1186/s13015-019-0147-6](https://doi.org/10.1186/s13015-019-0147-6)
(version journal de la référence C5)
- J3 **Bastien Cazaux**, Guillaume Castel, Eric Rivals. AQUAPONY : visualization and interpretation of phylogeographic information on phylogenetic trees. *Bioinformatics* 35, 17 :3163–3165, **2019**. DOI : [10.1093/bioinformatics/btz011](https://doi.org/10.1093/bioinformatics/btz011)
- J4 **[article joint au dossier]** **Bastien Cazaux**, Thierry Lecroq, and Eric Rivals. Linking indexing data structures to de bruijn graphs : Construction and update. *Journal of Computer and System Sciences*, 104 :165–183, **2019**. DOI : [10.1016/j.jcss.2016.06.008](https://doi.org/10.1016/j.jcss.2016.06.008)
(version journal des références C11 et C9)

- J5 **Bastien Cazaux** and Eric Rivals. Relationship between superstring and compression measures : New insights on the greedy conjecture. *Discrete Applied Mathematics*, 245 :59–64, **2018**. DOI : [10.1016/j.dam.2017.04.017](https://doi.org/10.1016/j.dam.2017.04.017)
- J6 **Bastien Cazaux** and Eric Rivals. A linear time algorithm for Shortest Cyclic Cover of Strings. *Journal of Discrete Algorithms*, 37 :56–67, **2016**. DOI : [10.1016/j.jda.2016.05.001](https://doi.org/10.1016/j.jda.2016.05.001)
- J7 Antoine Limasset, **Bastien Cazaux**, Eric Rivals, and Pierre Peterlongo. Read mapping on de bruijn graphs. *BMC Bioinformatics*, 17 :237, **2016**. DOI : [10.1186/s12859-016-1103-9](https://doi.org/10.1186/s12859-016-1103-9)
- J8 Yvan Le Bras, Olivier Collin, Cyril Monjeaud, Vincent Lacroix, Éric Rivals, Claire Lemaitre, Vincent Miele, Gustavo Sacomoto, Camille Marchet, **Bastien Cazaux**, Amal Zine El Aabidine, Leena Salmela, Susete Alves-Carvalho, Alexan Andrieux, Raluca Uricaru and Pierre Peterlongo. Colib' read on galaxy : a tools suite dedicated to biological information extraction from raw NGS reads. *GigaScience*, 5 :1, **2016**. DOI : [10.1186/s13742-015-0105-2](https://doi.org/10.1186/s13742-015-0105-2)
- J9 **Bastien Cazaux** and Eric Rivals. The power of greedy algorithms for approximating Max-ATSP, Cyclic Cover, and superstrings. *Discrete Applied Mathematics*, 212 :48–60, **2016**. DOI : [10.1016/j.dam.2015.06.003](https://doi.org/10.1016/j.dam.2015.06.003)
(version journal de la référence **C10**)
- J10 **Bastien Cazaux** and Eric Rivals. Reverse engineering of compact suffix trees and links : A novel algorithm. *Journal of Discrete Algorithms*, 28 :9–22, **2014**. DOI : [10.1016/j.jda.2014.07.002](https://doi.org/10.1016/j.jda.2014.07.002)

Conférence d'audience internationale avec comité de sélection

- C1 Jarno Alanko, Hideo Bannai, **Bastien Cazaux**, Pierre Peterlongo, Jens Stoye. Finding All Maximal Perfect Haplotype Blocks in Linear Time. *19th International Workshop on Algorithms in Bioinformatics (WABI)*, 8 :1–8 :9, **2019**. DOI : [10.4230/LIPIcs.WABI.2019.8](https://doi.org/10.4230/LIPIcs.WABI.2019.8)
- C2 **Bastien Cazaux** and Eric Rivals. Linking BWT and XBW via Aho-Corasick Automaton : Applications to Run-Length Encoding. *30th Annual Symposium on Combinatorial Pattern Matching (CPM)*, 24 :1–24, **2019**. DOI : [10.4230/LIPIcs.CPM.2019.24](https://doi.org/10.4230/LIPIcs.CPM.2019.24)
- C3 **Bastien Cazaux**, Dmitry Kosolobov, Veli Mäkinen, Tuukka Norri. Linear Time Maximum Segmentation Problems in Column Stream Model. *String Processing and Information Retrieval - 26th International Symposium (SPIRE)*, 322–336, **2019**. DOI : [10.1007/978-3-030-32686-9_23](https://doi.org/10.1007/978-3-030-32686-9_23)
- C4 **Bastien Cazaux** and Eric Rivals. Superstrings with multiplicities. *Annual Symposium on Combinatorial Pattern Matching (CPM)*, 21 :1–21 :16, **2018**. DOI : [10.4230/LIPIcs.CPM.2018.21](https://doi.org/10.4230/LIPIcs.CPM.2018.21)
- C5 Tuukka Norri, **Bastien Cazaux**, Dmitry Kosolobov, Veli Mäkinen. Minimum Segmentation for Pan-genomic Founder Reconstruction in Linear Time. *18th International Workshop on Algorithms in Bioinformatics (WABI)*, 15 :1–15 :15, **2018**. DOI : [10.4230/LIPIcs.WABI.2018.15](https://doi.org/10.4230/LIPIcs.WABI.2018.15)

- C6 **Bastien Cazaux**, Samuel Juhel, Eric Rivals. Practical lower and upper bounds for the Shortest Linear Superstring. *17th International Symposium on Experimental Algorithms (SEA)*, 18 :1–18 :14, **2018**. DOI : [10.4230/LIPIcs.SEA.2018.18](https://doi.org/10.4230/LIPIcs.SEA.2018.18)
- C7 **Bastien Cazaux**, Gustavo Sacomoto, and Eric Rivals. Superstring Graph : a new approach for genome assembly. In *Algorithmic Aspects in Information and Management - 11th International Conference (AAIM)*, 39–52, **2016**. DOI : [10.1007/978-3-319-41168-2_4](https://doi.org/10.1007/978-3-319-41168-2_4)
- C8 **[article joint au dossier] Bastien Cazaux**, Rodrigo Cánovas, and Eric Rivals. Shortest DNA cyclic cover in compressed space. In *Data Compression Conference (DCC)*, 536–545, **2016**. DOI : [10.1109/DCC.2016.79](https://doi.org/10.1109/DCC.2016.79)
- C9 **Bastien Cazaux**, Thierry Lecroq, and Eric Rivals. Construction of a de Bruijn Graph for Assembly from a Truncated Suffix Tree. *Language and Automata Theory and Applications - 9th International Conference (LATA)*, 109–120, **2015**. DOI : [10.1007/978-3-319-15579-1_8](https://doi.org/10.1007/978-3-319-15579-1_8)
- C10 **Bastien Cazaux** and Eric Rivals. Approximation of greedy algorithms for Max-ATSP, Maximal Compression, Maximal Cycle Cover, and Shortest Cyclic Cover of Strings. In *Proc. of Prague Stringology Conference (PSC)*, 148–161, **2014**. DOI : [978-80-01-05547-2](https://doi.org/10.978-80-01-05547-2)
- C11 **Bastien Cazaux**, Thierry Lecroq, and Eric Rivals. From Indexing Data Structures to de Bruijn Graphs. *Proc. of the 25th Annual Symposium on Combinatorial Pattern Matching (CPM)*, 89–99, **2014**. DOI : [10.1007/978-3-319-07566-2_10](https://doi.org/10.1007/978-3-319-07566-2_10)

Visibilité

Communications invitées

- 2014 Arbres des suffixes et applications, Séminaire des doctorants, **Montpellier**
- 2015 Shortest superstring problems and graphs, Séminaire d'équipe ALGCo, **Montpellier**
The Superstring Graph, Séminaire d'équipe MAB, **Montpellier**
- 2016 Superstring Problems, Séminaire d'équipe MAB, **Montpellier**
The Superstring Graph, Séminaire d'équipe MEB, **Marseille**
Superstring Graph, Séminaire d'équipe ALGCo, **Montpellier**
- 2018 Superstring Graph, Séminaire d'équipe GSA, **Helsinki**
- 2019 Applications of the pBWT Pan-genomic Founder Reconstruction and Finding all maximal perfect haplotype blocks, Séminaire d'équipe MAB, **Montpellier**
Applications of the pBWT Pan-genomic Founder Reconstruction and Finding all maximal perfect haplotype blocks, Séminaire d'équipe MABioVis, **Bordeaux**

Communications dans des conférences et workshops nationaux

- 2016 Le graphe glouton pour les problèmes de superchaîne, Journées Graphes et Algorithmes, **Paris**
Superstring Graph : a new approach for genome assembly, SeqBio, **Nantes**

- 2017 Superstring Graph : a new approach for genome assembly, journées IBC, **Montpellier**
- 2018 Minimum Segmentation for Pan-genomic Founder Reconstruction in Linear Time, SeqBio, **Rouen**
- 2019 Minimum Segmentation for Pan-genomic Founder Reconstruction in Linear Time, Bioinformatics Day, **Helsinki (Finlande)**
- Communications dans des conférences et workshops internationaux
- 2014 Approximation of greedy algorithms for Max- ATSP, Maximal Compression, Maximal Cycle Cover, and Shortest Cyclic Cover of Strings, PSC, **Prague (République tchèque)**
- 2015 Construction of a de Bruijn Graph for Assembly from a Truncated Suffix Tree, LATA, **Nice**
- 2016 The Superstring Graph, DSB, **Bielefeld (Allemagne)**
Superstring Graph : a new approach for genome assembly, AAIM, **Bergamo (Italie)**
Shortest DNA cyclic cover in compressed space, DCC, **Snowbird (USA)**
- 2018 Superstrings with multiplicities, CPM, **Qingdao (Chine)**
- 2019 Range Minimum Deque for Linear Time Maximum Segmentation in Column Stream Model, DSB, **Dortmund (Allemagne)**
Linking BWT and XBW via Aho-Corasick Automaton : Applications to Run-Length Encoding, CPM, **Pisa (Italie)**