# Yuri Pirola / Curriculum Vitae

yuri.pirola@unimib.it

https://algolab.eu/pirola

#### **Research Interests**

Algorithms and data structures in Bioinformatics / Computational and parameterized complexity / Genomic and transcriptomic sequence analysis and assembly / Haplotype inference and reconstruction / Evolutionary metaheuristics and genetic programming

#### **Positions**

- > Assistant Professor / Oct. 2019–current
  - Dept. of Informatics, Systems and Communications, Univ. degli Studi di Milano-Bicocca
  - Member of the Bioinformatics and Experimental Algorithmic lab and of the Interdisciplinary Research Centre "Bicocca Bioinformatics Biostatistics and Bioimaging centre" (B4) at Univ. degli Studi di Milano-Bicocca
- > Post-Doctoral Fellow / May 2012–Dec. 2015 / "Algorithmic methods for Next-Generation Sequencing data analysis"
  - Funding: Univ. degli Studi di Milano-Bicocca
- > Post-Doctoral Fellow / Jan. 2011–Apr. 2012 / "Efficient haplotype inference in livestock from high-density SNP chips"
  - Funding: Univ. degli Studi di Milano-Bicocca / Lombardy Region / Parco Tecnologico Padano (Lodi)
- Post-Doctoral Researcher / Mar. 2010–Dec. 2010 / "PROZOO Project"
   Funding: Parco Tecnologico Padano (Lodi)

#### **Education**

- > Ph.D. in Computer Science / Univ. degli Studi di Milano-Bicocca / February 3, 2010 "Combinatorial Problems in Studies of Genetic Variations: Haplotyping and Transcript Analysis" Supervisor: Prof. Paola Bonizzoni
- > M.Sc. in Computer Science, magna cum laude / Univ. degli Studi di Milano-Bicocca / February 23, 2006 "Analisi della Neutralità degli Spazi di Ricerca Booleani in Programmazione Genetica" Supervisors: Dr. Leonardo Vanneschi & Prof. Giancarlo Mauri

## **Research projects**

- > "PANGAIA: Pan-genome Graph Algorithms and Data Integration"
  Funding: EU Horizon 2020, RISE Marie Skłodowska-Curie Action / Role: participant
- > "Modulation of anti-cancer immune response by regulatory non-coding RNAs" Funding: Cariplo Foundation / Role: participant
- "Automata and formal languages: mathematical and applicative aspects"
   Funding: MIUR (Italian Ministry of Education), PRIN 2010/11 Action / Role: participant
- > "Next Generation methods to preserve farm animal biodiversity by optimizing present and future breeding options" Funding: European Commission, 7th Framework Programme / Role: participant
- > "PROZOO: Applications of genomics to fertility, disease resistance, and product quality assurance in cattle and pigs" Funding: Cariplo Foundation & Lombardy Region / Role: participant

**Fellowships** 

- > Ph.D. fellowship / MIUR (Italian Ministry of Education) / Nov. 2006–Oct. 2009.
- > Scholarship for graduate students / Univ. degli Studi di Milano / May-Oct. 2006

**Publications** 

- > 27 articles on international peer-reviewed scientific journals (indexed on Scopus and/or WoS)
- > 15 papers on proceedings of international peer-reviewed conferences (indexed on Scopus and/or WoS)
- > 2 invited book chapters (indexed on Scopus and/or WoS)
- > Several posters presented at national and international conferences

## International peer-reviewed journal articles

- [J27] Bonizzoni, P., Della Vedova, G., **Pirola, Y.**, Previtali, M., and Rizzi, R. "Computing the multi-string BWT and LCP array in external memory". *Theor. Comput. Sci.* 862 (2021), 42–58. DOI: 10.1016/j.tcs.2020.11.041.
- [J26] Denti, L., **Pirola, Y.**, Previtali, M., Ceccato, T., Della Vedova, G., Rizzi, R., and Bonizzoni, P. "Shark: fishing relevant reads in an RNA-Seq sample". *Bioinformatics* 37.4 (2021), 464–472.

  DOI: 10.1093/bioinformatics/btaa779.
- [J25] Rizzi, R., Beretta, S., Patterson, M., **Pirola, Y.**, Previtali, M., Della Vedova, G., and Bonizzoni, P. "Overlap graphs and de Bruijn graphs: data structures for de novo genome assembly in the big data era". *Quant. Biol.* 7.4 (2019), 278–292. DOI: 10.1007/s40484-019-0181-x.
- [J24] Calabria, A., Beretta, S., Merelli, I., Spinozzi, G., Brasca, S., **Pirola, Y.**, Benedicenti, F., Tenderini, E., Bonizzoni, P., Milanesi, L., and Montini, E. "γ-TRIS: a graph-algorithm for comprehensive identification of vector genomic insertion sites". *Bioinformatics* 36.5 (2020), 1622–1624. DOI: 10.1093/bioinformatics/btz747.
- [J23] Bonizzoni, P., Della Vedova, G., Pirola, Y., Previtali, M., and Rizzi, R. "Multithread Multistring Burrows–Wheeler Transform and Longest Common Prefix Array". J. Comput. Biol. 26.9 (2019), 948–961.
  DOI: 10.1089/cmb.2018.0230.
- [J22] Grüning, B., Dale, R., Sjödin, A., Chapman, B. A., Rowe, J., Tomkins-Tinch, C. H., Valieris, R., Köster, J., and The Bioconda Team (including **Pirola, Y.**) "Bioconda: sustainable and comprehensive software distribution for the life sciences". *Nature Methods* 15.7 (2018), 475–476. DOI: 10.1038/s41592-018-0046-7.
- [J21] Bonizzoni, P., Della Vedova, G., **Pirola, Y.**, Previtali, M., and Rizzi, R. "FSG: Fast String Graph Construction for De Novo Assembly". *J. Comput. Biol.* 24.10 (2017), 953–968. DOI: 10.1089/cmb.2017.0089.
- [J20] Bonizzoni, P., Della Vedova, G., **Pirola, Y.**, Previtali, M., and Rizzi, R. "An External-Memory Algorithm for String Graph Construction". *Algorithmica* 78.2 (2017), 394–424. DOI: 10.1007/s00453-016-0165-4.
- [J19] Biscarini, F., Schwarzenbacher, H., Pausch, H., Nicolazzi, E. L., **Pirola, Y.**, and Biffani, S. "Use of SNP genotypes to identify carriers of harmful recessive mutations in cattle populations". *BMC Genomics* 17 (2016), 857. DOI: 10.1186/s12864-016-3218-9.
- [J18] Chiaradonna, F., **Pirola, Y.**, Ricciardiello, F., and Palorini, R. "Transcriptional profiling of immortalized and K-ras-transformed mouse fibroblasts upon PKA stimulation by forskolin in low glucose availability". *Genomics Data* 9 (2016), 100–104. DOI: 10.1016/j.gdata.2016.07.004.
- [J17] Bonizzoni, P., Dondi, R., Klau, G. W., **Pirola, Y.**, Pisanti, N., and Zaccaria, S. "On the Minimum Error Correction Problem for Haplotype Assembly in Diploid and Polyploid Genomes". *J. Comput. Biol.* 23.9 (2016), 718–736. DOI: 10.1089/cmb.2015.0220.
- [J16] Palorini, R., Votta, G., **Pirola, Y.**, De Vitto, H., De Palma, S., Airoldi, C., Vasso, M., Ricciardiello, F., Lombardi, P. P., Cirulli, C., Rizzi, R., Nicotra, F., Hiller, K., Gelfi, C., Alberghina, L., and Chiaradonna, F. "Protein Kinase A Activation Promotes Cancer Cell Resistance to Glucose Starvation and *Anoikis*". *PLoS Genet.* 12.3 (2016), 1–41. DOI: 10.1371/journal.pgen.1005931.

- [J15] Bonizzoni, P., Della Vedova, G., Pirola, Y., Previtali, M., and Rizzi, R. "LSG: An External-Memory Tool to Compute String Graphs for NGS Data Assembly". J. Comput. Biol. 23.3 (2016), 137–149.
  DOI: 10.1089/cmb.2015.0172.
- [J14] **Pirola, Y.**, Zaccaria, S., Dondi, R., Klau, G. W., Pisanti, N., and Bonizzoni, P. "HapCol: Accurate and Memory-Efficient Haplotype Assembly from Long Reads". *Bioinformatics* 32.11 (2016), 1610–1617. DOI: 10.1093/bioinformatics/btv495.
- [J13] Beerenwinkel, N., Beretta, S., Bonizzoni, P., Dondi, R., and **Pirola, Y.** "Covering Pairs in Directed Acyclic Graphs". *Comput. J.* 58.7 (2015), 1673–1686. DOI: 10.1093/comjnl/bxu116.
- [J12] Batini, C., Bonizzoni, P., Comerio, M., Dondi, R., Pirola, Y., and Salandra, F. "A Clustering Algorithm for Planning the Integration Process of a Large Number of Conceptual Schemas". J. Comput. Sci. Technol. 30.1 (2015), 214–224. DOI: 10.1007/s11390-015-1514-5.
- [J11] Beretta, S., Bonizzoni, P., Della Vedova, G., **Pirola, Y.**, and Rizzi, R. "Modeling Alternative Splicing Variants from RNA-Seq Data with Isoform Graphs". *J. Comput. Biol.* 21.1 (2014), 16–40.

  DOI: 10.1089/cmb.2013.0112.
- [J10] Bonizzoni, P., Della Vedova, G., Dondi, R., and **Pirola, Y.** "Parameterized Complexity of *k*-Anonymity: Hardness and Tractability". *J. Comb. Optim.* 26.1 (2013), 19–43. DOI: 10.1007/s10878-011-9428-9.
- [J9] Bonizzoni, P., Dondi, R., and **Pirola, Y.** "Maximum Disjoint Paths on Edge-Colored Graphs: Approximability and Tractability". *Algorithms* 6.1 (2013), 1–11. DOI: 10.3390/a6010001.
- [J8] **Pirola, Y.**, Della Vedova, G., Biffani, S., Stella, A., and Bonizzoni, P. "A Fast and Practical Approach to Genotype Phasing and Imputation on a Pedigree with Erroneous and Incomplete Information". *IEEE/ACM Trans. Comput. Biol. Bioinform.* 9.6 (2012), 1582–1594. DOI: 10.1109/TCBB.2012.100.
- [J7] **Pirola, Y.**, Rizzi, R., Picardi, E., Pesole, G., Della Vedova, G., and Bonizzoni, P. "PIntron: A Fast Method for Detecting the Gene Structure Due to Alternative Splicing Via Maximal Pairings of a Pattern and a Text". *BMC Bioinformatics* 13.S5 (2012), S2. DOI: 10.1186/1471-2105-13-S5-S2.
- [J6] Vanneschi, L., **Pirola, Y.**, Mauri, G., Tomassini, M., Collard, P., and Verel, S. "A Study of Neutrality of Boolean Function Landscapes in Genetic Programming". *Theor. Comput. Sci.* 425 (2012), 34–57. DOI: 10.1016/j.tcs.2011.03.011.
- [J5] **Pirola, Y.**, Bonizzoni, P., and Jiang, T. "An Efficient Algorithm for Haplotype Inference on Pedigrees with Recombinations and Mutations". *IEEE/ACM Trans. Comput. Biol. Bioinform.* 9.1 (2012), 12–25. DOI: 10.1109/TCBB.2011.51.
- [J4] Bonizzoni, P., Della Vedova, G., Dondi, R., and **Pirola, Y.** "Variants of Constrained Longest Common Subsequence". *Inf. Process. Lett.* 110.20 (2010), 877–881. DOI: 10.1016/j.ipl.2010.07.015.
- [J3] Bonizzoni, P., Della Vedova, G., Dondi, R., **Pirola, Y.**, and Rizzi, R. "Pure Parsimony Xor Haplotyping". *IEEE/ACM Trans. Comput. Biol. Bioinform.* 7.4 (2010), 598–610. DOI: 10.1109/TCBB.2010.52.
- [J2] Della Vedova, G., Dondi, R., Jiang, T., Pavesi, G., **Pirola, Y.**, and Wang, L. "Beyond Evolutionary Trees". *Nat. Comput.* 9.2 (2010), 421–435. DOI: 10.1007/s11047-009-9156-6.
- [J1] Bonizzoni, P., Mauri, G., Pesole, G., Picardi, E., **Pirola, Y.**, and Rizzi, R. "Detecting Alternative Gene Structures from Spliced ESTs: A Computational Approach". *J. Comput. Biol.* 16.1 (2009), 43–66. DOI: 10.1089/cmb.2008.0028.

## International peer-reviewed conference papers

- [C15] Bonizzoni, P., De Felice, C., Petescia, A., **Pirola, Y.**, Rizzi, R., Stoye, J., Zaccagnino, R., and Zizza, R. "Can We Replace Reads by Numeric Signatures? Lyndon Fingerprints as Representations of Sequencing Reads for Machine Learning". In: *Algorithms for Computational Biology (AlCoB)*. Vol. 12715. LNCS. Springer, 2021, 16–28. DOI: 10.1007/978-3-030-74432-8\_2.
- [C14] Bonizzoni, P., Della Vedova, G., Nicosia, S., **Pirola, Y.**, Previtali, M., and Rizzi, R. "Divide and Conquer Computation of the Multi-string BWT and LCP Array". In: *Computability in Europe (CiE)*. Vol. 10936. LNCS. Springer, 2018, 107–117. DOI: 10.1007/978-3-319-94418-0\_11.

- [C13] Bonizzoni, P., Della Vedova, G., **Pirola, Y.**, Previtali, M., and Rizzi, R. "FSG: Fast String Graph Construction for De Novo Assembly of reads data". In: *Bioinformatics Research and Applications (ISBRA)*. Vol. 9683. LNCS. Springer, 2016, 27–39. DOI: 10.1007/978-3-319-38782-6\_3.
- [C12] Bonizzoni, P., Dondi, R., Klau, G. W., **Pirola, Y.**, Pisanti, N., and Zaccaria, S. "On the Fixed Parameter Tractability and Approximability of the Minimum Error Correction problem". In: *Combinatorial Pattern Matching (CPM)*. Vol. 9133. LNCS. Springer, 2015, 100–113. DOI: 10.1007/978-3-319-19929-0\_9.
- [C11] Bonizzoni, P., Della Vedova, G., **Pirola, Y.**, Previtali, M., and Rizzi, R. "Constructing String Graphs in External Memory". In: *Algorithms in Bioinformatics (WABI)*. Vol. 8701. LNCS. Springer, 2014, 311–325. DOI: 10.1007/978-3-662-44753-6\_23.
- [C10] Beerenwinkel, N., Beretta, S., Bonizzoni, P., Dondi, R., and **Pirola, Y.** "Covering Pairs in Directed Acyclic Graphs". In: *Language and Automata Theory and Applications (LATA)*. Vol. 8370. LNCS. Springer, 2014, 126–137. DOI: 10.1007/978-3-319-04921-2\_10.
- [C9] **Pirola, Y.**, Della Vedova, G., Bonizzoni, P., Stella, A., and Biscarini, F. "Haplotype-based prediction of gene alleles using pedigrees and SNP genotypes". In: *Bioinformatics, Computational Biology, and Biomedical Informatics (ACM BCB)*. ACM, 2013, 33–41. DOI: 10.1145/2506583.2506592.
- [C8] **Pirola, Y.**, Della Vedova, G., Biffani, S., Stella, A., and Bonizzoni, P. "A fast and practical approach to genotype phasing and imputation on a pedigree with erroneous and incomplete information". In: *Computational Advances in Bio and medical Sciences (ICCABS)*. IEEE, 2012. DOI: 10.1109/ICCABS.2012.6182643.
- [C7] Bonizzoni, P., Della Vedova, G., **Pirola, Y.**, and Rizzi, R. "PIntron: a fast method for gene structure prediction via maximal pairings of a pattern and a text". In: *Computational Advances in Bio and medical Sciences (ICCABS)*. IEEE, 2011, 33–39. DOI: 10.1109/ICCABS.2011.5729935.
- [C6] **Pirola, Y.**, Bonizzoni, P., and Jiang, T. "Haplotype Inference on Pedigrees with Recombinations and Mutations". In: *Algorithms in Bioinformatics (WABI)*. Vol. 6293. LNCS. Springer, 2010, 148–161. DOI: 10.1007/978-3-642-15294-8\_13.
- [C5] Bonizzoni, P., Della Vedova, G., Dondi, R., and **Pirola, Y.** "Parameterized Complexity of k-Anonymity: Hardness and Tractability". In: *Combinatorial Algorithms (IWOCA)*. Vol. 6460. LNCS. Springer, 2011, 242–255. DOI: 10.1007/978-3-642-19222-7\_25.
- [C4] Bonizzoni, P., Della Vedova, G., Dondi, R., **Pirola, Y.**, and Rizzi, R. "Minimum Factorization Agreement of Spliced ESTs". In: *Algorithms in Bioinformatics (WABI)*. Vol. 5724. LNCS. Springer, 2009, 1–12. DOI: 10.1007/978-3-642-04241-6\_1.
- [C3] Bonizzoni, P., Della Vedova, G., Dondi, R., **Pirola, Y.**, and Rizzi, R. "Pure Parsimony Xor Haplotyping". In: *Bioinformatics Research and Applications (ISBRA)*. Vol. 5542. LNCS. Springer, 2009, 186–197. DOI: 10.1007/978-3-642-01551-9\_19.
- [C2] Vanneschi, L., Tomassini, M., Collard, P., Verel, S., **Pirola, Y.**, and Mauri, G. "A Comprehensive View of Fitness Landscapes with Neutrality and Fitness Clouds". In: *Genetic Programming (EuroGP)*. Vol. 4445. LNCS. Springer, 2007, 241–250. DOI: 10.1007/978-3-540-71605-1\_22.
- [C1] Vanneschi, L., **Pirola, Y.**, and Collard, P. "A quantitative study of neutrality in GP boolean landscapes". In: *Genetic and Evolutionary Computation (GECCO)*. ACM, 2006, 895–902. DOI: 10.1145/1143997.1144152.

# **Book chapters**

- [B2] Dondi, R. and **Pirola, Y.** "Beyond Evolutionary Trees". In: *Encyclopedia of Algorithms*. Ed. by M.-Y. Kao. Springer, 2016, 183–189. ISBN: 978-3-642-27848-8. DOI: 10.1007/978-3-642-27848-8\_599-1.
- [B1] Bonizzoni, P., Della Vedova, G., Pesole, G., Picardi, E., **Pirola, Y.**, and Rizzi, R. "Transcriptome Assembly and Alternative Splicing Analysis". In: *RNA Bioinformatics*. Ed. by E. Picardi. Vol. 1269. Methods in Molecular Biology. Springer, 2015, 173–188. ISBN: 978-1-4939-2290-1. DOI: 10.1007/978-1-4939-2291-8\_11.

  In: ISI WoS, Scopus.

## Talks, schools, and visits

## **Invited talks**

- > Università degli Studi di Milano / invited by Prof. Giovanni Righini / Feb. 8, 2010
- > Parco Tecnologico Padano (Lodi) / invited by Dr. Alessandra Stella / Mar. 25, 2010

#### **Talks**

- > Bioinformatics and Computational Biology Conference (BBCC) / Naples, Italy / Nov. 2020
- > Bioinformatics Open Source Conference (BOSC) / Toronto, Canada / Jul. 2020
- > Int. Workshop on Data Structures in Bioinformatics (DSB) / Rennes, France / Feb. 2020
- > Int. Conf. on Language and Automata Theory and Applications (LATA) / Madrid, Spain / Mar. 2014
- > Workshop "Combinatorial structures for sequence analysis in bioinformatics" / Milan, Italy / Nov. 2013
- > ACM Int. Conf. on Bioinf., Computational Biol., and Biomedical Inform. (ACM BCB) / Washington DC, USA / Sep. 2013
- > Italian Conf. on Theoretical Computer Science (ICTCS) / Palermo, Italy / Sep. 2013
- > IEEE Int. Conf. on Computational Advances in Bio and medical Sciences (ICCABS) / Las Vegas NV, USA / Feb. 2012.
- > IEEE Int. Conf. on Computational Advances in Bio and medical Sciences (ICCABS) / Orlando FL, USA / Feb. 2011
- > Int. Workshop on Algorithms in Bioinformatics (WABI) / Liverpool, UK / Sep. 2010
- > Int. Workshop on Algorithms in Bioinformatics (WABI) / Philadelphia PA, USA / Sep. 2009
- > Int. Symp. on Bioinformatics Research and Applications (ISBRA) / Ft. Lauderdale FL, USA / May 2009

## Schools

- > Lipari Int. Summer School on Bioinformatics and Computational Biology / Univ. degli Studi di Catania / 2008
- > Summer school on Parallel and Scientific Computing / CINECA / 2006

## **Research visits**

- > Centrum Wiskunde & Informatica / Amsterdam, The Netherlands / Dr. Gunnar W. Klau / Jul. 14–17, 2015
- > Université Paris-Est Marne-la-Vallée / Paris, France / Dr. Gregory Kucherov / Jan. 12-18, 2014
- > University of California, Riverside / Riverside CA, USA / Prof. Tao Jiang / Feb.-Jun. 2009

## **Professional activities**

- > Co-organizer (with S. Pissis, CWI) of a satellite workshop of ECCB 2020
- > Program committee member of: BICOB 2020 / HPC4COVID-19 2020 (part of IEEE BIBM) / IWBBIO 2020 / BBC 2015–2021 (part of ICCS) / PDP 2015–2017
- > Organizing committee member of: DSB 2021 / CiE 2013 / "Unconventional Models of Computation" 2009 / ASWorkshop 2008
- > Topic editor in the *editorial board* of Algorithms
- > Review editor in the *editorial board* of Frontiers in Bioinformatics

- > Reviewer for the following journals: Discrete Applied Mathematics / Bioinformatics / Briefings in Bioinformatics / IEEE-ACM Trans. Computational Biology and Bioinformatics / IEEE Access / J. Computational Science / J. Biomedical Informatics / Computers & Operations Research / Mathematics / Genes / Processes / Systems / Technologies / Concurrency and Computation: Practice and Experience / BMC Genomics / Int. J. Bioinformatics Research and Applications / Int. J. Molecular Sciences / BioMed Research Int.
- > Reviewer for the following international conferences: WABI 2021 / APBC 2020 / SODA 2018 / AAIM 2016 / WALCOM 2016 / IWOCA 2016, 2015 / ISBRA 2015, 2014 / WABI 2014, 2012 / CIBB 2019, 2014 / IEEE BIBM 2020, 2013, 2012, 2009 / BSB 2013 / CiE 2013 / CATS 2013 / ECCB 2012 / ICCABS 2011

### **Supervision activities**

- > Supervisor for 1 Ph.D. thesis in CS at Univ. degli Studi di Milano-Bicocca
- > Supervisor for 4 M.Sc. theses in CS at Univ. degli Studi di Milano-Bicocca
- > Supervisor for 8 B.Sc. theses in CS at Univ. degli Studi di Milano-Bicocca

### **Teaching**

- > Member of the board of the Ph.D. program in Computer Science at Univ. degli Studi di Milano-Bicocca
- > Teacher/Teaching assistant, at Univ. degli Studi di Milano-Bicocca, of:

Algorithms and DS	20h, class	B.Sc.	School of Computer Science	2019/20-20/21
Algorithms and DS	20h, lab.	B.Sc.	School of Computer Science	2019/20-20/21
Software design	12h, class	M.Sc.	School of Computer Science	2019/20-20/21
Computer science	16h, class	MD	School of Medicine	2018/19-19/20
Elements of Bioinformatics	16h, lab.	B.Sc.	School of Computer Science	2013/14
Algorithms and DS 2	12h, class	B.Sc.	School of Computer Science	2010/11
Bioinformatics	4h, lectures	M.Sc.	School of Computer Science	2009/10
Bioinformatics	8h, class	M.Sc.	School of Computer Science	2008/09
Analysis of algorithms	12h, class	M.Sc.	School of Computer Science	2007/08
Bioinformatics	4h, lectures	B.Sc.	School of Computer Science	2007/08
Programming languages	24h, lab.	B.Sc.	School of Computer Science	2006/07
Algorithms and DS	12h, lab.	B.Sc.	School of Computer Science	2005/06
Operating systems	24h, lab.	B.Sc.	School of Computer Science	2005/06
Operating systems	48h, lab.	B.Sc.	School of Computer Science	2004/05

- > *Teaching assistant*, at Univ. degli Studi di Bergamo, of: Introduction to Informatics 30h, class B.A.
- School of Foreign Languages 2007/08–08/09
  - m in Computer Science
- > *Lecturer* for the course "Do we need data structures?" (4h, Oct. 2020) of the Ph.D. program in Computer Science at Univ. degli Studi di Milano-Bicocca
- > Lecturer on the topic "Introduction to programming in C" (16h) for post-graduate LS researchers at PTP, Lodi