

Can Alkan

CONTACT INFORMATION

Department of Computer Engineering
Bilkent University
Engineering Building, EA 509
Bilkent, Ankara 06800
Turkey

Voice: (+90) 312-290-2912
Fax: (+90) 312-266-4047
E-mail: calkan@cs.bilkent.edu.tr
ORCID: 0000-0002-5443-0706
ResearcherID: D-2982-2009
Web: <http://www.alkanlab.org>



RESEARCH INTERESTS

Combinatorial algorithms for the analysis of next-generation sequencing data, genomic structural variation, human and primate segmental duplications.

EXPERIENCE

Department of Computer Engineering, Bilkent University, Bilkent, Ankara, Turkey

Assistant Professor

January 2012 - present

Department of Genome Sciences, University of Washington, Seattle, Washington, USA

Acting Assistant Professor

June 2011 - December 2011

Senior Fellow

October 2005 - May 2011

School of Computing Science, Simon Fraser University, Burnaby, British Columbia, Canada

Visiting Researcher

January, 2004 - September 2005

EDUCATION

University of Washington, Seattle, Washington, USA

Department of Genome Sciences and Howard Hughes Medical Institute

Postdoctoral Fellow, Genomics, October 2005 - December 2011

- Worked on computational methods to reconstruct the evolutionary history of alpha-satellite DNA, *de novo* alphoid sequence prediction from whole-genome shotgun sequence data, and detection of human structural variation and segmental duplications using next-generation sequencing technologies.
- Advisor: Evan E. Eichler

Case Western Reserve University, Cleveland, Ohio, USA

Department of EECS

Ph.D., Computer Science, August 2005

- Dissertation Topic: "Computational Studies on Evolution and Functionality of Genomic Repeats"
- Advisor: S. Cenk Şahinalp

Bilkent University, Ankara, Turkey

Department of Computer Engineering

B.Sc., Computer Science, May, 2000

HONORS AND AWARDS

- TÜBİTAK Incentive Award, 2015.
- Thomson Reuters Highly Cited Researcher, 2014.
- Young Investigator Award (BAGEP), Science Academy of Turkey, 2013.
- AAAS Newcomb Cleveland Prize for the Most Outstanding Paper Published in Science, American Association for the Advancement of Science, USA, 2010.
- Case Prime Fellowship, Case Western Reserve University, Cleveland, Ohio, USA, 2000.
- Full Scholarship Awarded by Bilkent University, 1996-2000.

PROFESSIONAL MEMBERSHIPS

- Member of the International Society for Computational Biology since 2002.

SUPERVISED
STUDENTS AND
INTERNS

CURRENT GRADUATE STUDENTS (PH.D.)

- **Pinar Kavak:** Ph.D. student, Computer Engineering, Fall 2010 - present. Boğaziçi University, İstanbul, Turkey. Co-supervised with Tunga Güngör.
- **Shatlyk Asyralyyev:** Ph.D. student, Computer Engineering, Fall 2013 - present. Bilkent University, Ankara, Turkey. Co-supervised with Cevdet Aykanat.
- **Volkan Yazar:** Ph.D. student, Molecular Biology and Genetics, Fall 2013 - present. Bilkent University, Ankara, Turkey. Co-supervised with İhsan Gürsel.
- **Fatma Kahveci:** Ph.D. student, Computer Engineering, Fall 2014 - present. Bilkent University, Ankara, Turkey.
- **Mohammed Alser:** Ph.D. student, Computer Engineering, Fall 2014 - present. Bilkent University, Ankara, Turkey. Co-supervised with Özcan Öztürk.
- **Handan Kulan:** Ph.D. student, Computer Engineering, Fall 2014 - present. Bilkent University, Ankara, Turkey.
- **Arda Söylev:** Ph.D. student, Computer Engineering, Fall 2014 - present. Bilkent University, Ankara, Turkey.
- **Marzieh Eslami Rasekh:** Ph.D. student, Computer Engineering, Fall 2015 - present. Bilkent University, Ankara, Turkey.

CURRENT GRADUATE STUDENTS (M.Sc.)

- **Azita Nouri:** M.Sc. student, Computer Engineering, Spring 2014 - present. Bilkent University, Ankara, Turkey. Co-supervised with Özcan Öztürk.
- **Gülfem Demir:** M.Sc. student, Computer Engineering, Fall 2014 - present. Bilkent University, Ankara, Turkey.
- **Can Fırtına:** M.Sc. student, Computer Engineering, Fall 2015 - present. Bilkent University, Ankara, Turkey.
- **F. Tuğba Doğan:** M.Sc. student, Computer Engineering, Fall 2015 - present. Bilkent University, Ankara, Turkey.

GRADUATE STUDENTS (ALUMNI)

- **Mustafa Korkmaz:** M.Sc., Computer Engineering, graduated in Fall 2012. Bilkent University, Ankara, Turkey. Co-supervised with Cevdet Aykanat.
Thesis Title: “Massively parallel mapping of next generation sequence reads using GPU”
now: Ph.D. student at University of Waterloo.
- **Fatma Kahveci (née Balcı):** M.Sc., Computer Engineering, graduated in Spring 2014. Bilkent University, Ankara, Turkey.
Thesis Title: “Bias correction in finding copy number variation using read depth based methods in exome sequencing data”
now: Ph.D. student at Bilkent University
- **Elif Dal:** M.Sc., Computer Engineering, graduated in Fall 2014. Bilkent University, Ankara, Turkey.
Thesis Title: “Genome scaffolding using pooled clone sequencing”
now: software engineer at TÜBİTAK Software Technologies Research Institute
- **Marzieh Eslami Rasekh:** M.Sc., Computer Engineering, graduated in Spring 2015. Bilkent University, Ankara, Turkey.
Thesis Title: “Algorithms for the discovery of large genomic inversions using pooled clone sequencing”
now: Ph.D. student at Bilkent University

THESIS COMMITTEES (CURRENT)

- **Arif Yılmaz:** Ph.D. candidate, Bioinformatics, since December 2012. Advisor: Yeşim Aydın Son. Middle East Technical University, Ankara, Turkey.
- **Evrin Tekeli:** Ph.D. candidate, Anthropology, since Fall 2013. Advisor: Timur Gültekin. Ankara University, Ankara, Turkey.
- **Fahrettin Can Koyuncu:** Ph.D. candidate, Computer Engineering, since Spring 2014. Advisor: Çiğdem Gündüz Demir. Bilkent University, Ankara, Turkey.

- **Abdullah Atmaca:** Ph.D. candidate, Computer Engineering, since Spring 2015. Advisors: Yavuz Oruç and Cevdet Aykanat. Bilkent University, Ankara, Turkey.

THESIS COMMITTEES (ALUMNI)

- **Fahrettin Can Koyuncu:** M.Sc. Computer Engineering, completed in Fall 2012. Advisor: Çiğdem Gündüz Demir. Bilkent University, Ankara, Turkey. *now: Ph.D. student at Bilkent University, Ankara, Turkey.*
- **Merve Çakır:** M.Sc. Computer Engineering, completed in Spring 2013. Advisor: Uğur Doğrusöz. Bilkent University, Ankara, Turkey. *now: Ph.D. student at Duke University, Durham, NC, USA.*
- **Mehmet Başaran:** M.Sc. Computer Engineering, completed in Fall 2014. Advisor: Cevdet Aykanat. Bilkent University, Ankara, Turkey.
- **Gülşah Merve Dal:** Ph.D. candidate, Molecular Biology and Genetics, completed in Fall 2014. Advisor: Tayfun Özçelik. Bilkent University, Ankara, Turkey. *now: Postdoctoral fellow at National Nanotechnology Research Center, Ankara, Turkey.*
- **Yogesh Paudel:** Ph.D. candidate, Biostatistics and Informatics, completed in January 2015. Advisor: Martien A.M. Groenen. Institut für Biostatistik und Informatik in Medizin und Alternsforschung, Wageningen University, Wageningen, The Netherlands.
- **Özlem Tufanlı:** Ph.D. candidate, Molecular Biology and Genetics, since Fall 2013. Left the committee in Spring 2015. Advisor: Ebru Erbay. Bilkent University, Ankara, Turkey.

UNDERGRADUATE VOLUNTEERS (CURRENT)

- **Abdullah Alperen:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. Since Spring 2015.
- **Halil İbrahim Özercan:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. Since Fall 2015.
- **Mert İnan:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. Since Fall 2015.
- **Ege Berkay Gülcan:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. Since Fall 2015.
- **Ertuğrul Akay:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. Since Fall 2015.
- **Burcu Çanakçı:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. Since Fall 2015.

UNDERGRADUATE VOLUNTEERS (ALUMNI)

- **Serhat Kıyak:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. 2013-2014. *now: Ph.D. student at Carnegie Mellon University.*
- **Yiğit Küçük:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. 2012-2014. *now: Ph.D. student at Case Western Reserve University.*
- **Can Koçkan:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. 2012-2014. *now: M.Sc. student at Simon Fraser University.*
- **Gülfem Demir:** B.Sc. student in Computer Engineering, Middle East Technical University, Ankara, Turkey. 2013-2014. *now: M.Sc. student at Bilkent University.*
- **Atalay Mert İleri:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. 2014. *now: Ph.D. student at MIT.*
- **Can Fırtına:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. Spring 2015. *now: M.Sc. student at Bilkent University.*
- **Damla Şenol:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. Spring 2015. *now: Ph.D. student at Carnegie Mellon University.*
- **Ayhun Tekat:** B.Sc. student in Computer Engineering, Bilkent University, Ankara, Turkey. Summer 2015. *now: M.Sc. student at Technische Universität München.*

OTHERS

- **Kerry Hall:** Intern at UW in 2006.
- **Fereydoun Hormozdiari:** Intern between February to July 2008 at UW. *now: Postdoctoral Fellow at University of Washington.*

- **Iman Hajirasouliha:** Intern UW between February to July 2009 at UW. *now: Postdoctoral Fellow at Brown University.*
- **Jacob O. Kitzman:** Rotation student at the Eichler Lab between January to April 2009. *now: Assistant Professor at University of Michigan.*
- **Peter H. Sudmant:** Ph.D. student at the Eichler Lab between May 2009 to December 2011. *now: Postdoctoral Fellow at University of Washington.*
- **Pietro D’Addabbo:** Visiting scientist at the Eichler Lab between August to November 2009. *now: Staff scientist at University of Bari.*
- **Michael Duyzend:** Rotation student at the Eichler Lab between June to August 2010. *now: Ph.D. student at University of Washington.*
- **Farhad Hormozdiari:** Research Scientist at the Eichler Lab between February to September 2010. *now: Ph.D. student at University of California, Los Angeles.*
- **Sònia Casillas:** Visiting scientist at the Eichler Lab between February to July 2011.
- **Vineet Bhakhar:** Intern between May to July 2014 at Bilkent.
- **Farnush Farhadi:** Intern between July to September 2014 at Bilkent.
- **Dilek Gengç:** M.Sc. student, Computer Engineering, Fall 2015. Bilkent University, Ankara, Turkey.

SERVICE ACTIVITIES

UNIVERSITY DUTIES

- Academic Coordinator for Erasmus+ and International Exchange Programs, Department of Computer Engineering, Bilkent University, since Summer 2014.

EDITORIAL

- Associate Editor, BMC Bioinformatics, 2010-*present*
- PC and OC Member, 7th Conference on High Throughput Sequencing Analysis and Algorithms (HiTSeq 2016), July 6-7, 2016, Orlando, FL, USA.
- PC Member, the 3rd International Conference on Algorithms for Computational Biology (AlCoB 2016), June 21-23, 2016, Trujillo, Spain.
- PC Member, the 20th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2016), April 17-21, 2015, Santa Monica, CA, USA.
- PC Member, 6th Conference on High Throughput Sequencing Analysis and Algorithms (HiTSeq 2015), July 10-11, 2015, Dublin, Ireland.
- Co-Chair, Bertinoro Computational Biology Meeting, June 14-17, 2015, Bertinoro, Italy.
- PC Member, 23rd Annual International Conference on Intelligent Systems for Molecular Biology and 14th European Conference on Computational Biology (ISMB/ECCB 2013), July 10-14, 2015, Dublin, Ireland.
- PC Member, 22nd Annual International Conference on Intelligent Systems for Molecular Biology (ISMB 2014), July 13-15, 2014, Boston, MA, USA.
- PC and OC Member, 8th International Symposium on Health Informatics and Bioinformatics, (HIBIT), September 25-27, 2013, Ankara, Turkey.
- PC Member, 21st Annual International Conference on Intelligent Systems for Molecular Biology and 12th European Conference on Computational Biology (ISMB/ECCB 2013), July 19-23, 2013, Berlin, Germany.
- PC Member, Third Annual RECOMB Workshop on Massively Parallel Sequencing (RECOMB-seq), April 11-12, 2013, Beijing, China.
- PC Member, 20th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB 2012), July 15-17, 2012, Long Beach, CA, USA.
- PC Member, Second Annual RECOMB Workshop on Massively Parallel Sequencing (RECOMB-seq), April 19-20, 2012, Barcelona, Spain.
- PC Member, 19th Annual International Conference on Intelligent Systems for Molecular Biology and 10th European Conference on Computational Biology (ISMB/ECCB 2011), July 15-19, 2011, Vienna, Austria.
- Chair, First Annual RECOMB Workshop on Massively Parallel Sequencing (RECOMB-seq), March 26-27, 2011, Vancouver, BC, Canada.
- PC Member, the 15th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2011), March 28-31, 2011, Vancouver, BC, Canada.

- Session Co-Chair, Session on Personal Genomics at the Pacific Symposium for Biocomputing (PSB 2011), January 3-7, 2011, The Big Island of Hawaii, USA.
- PC Member, 18th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB 2010), July 9-14, 2010, Boston, MA, USA.
- Session Co-Chair, Session on Personal Genomics at the Pacific Symposium for Biocomputing (PSB 2010), January 4-8, 2010, The Big Island of Hawaii, USA.
- PC Member, The IEEE 22nd International Conference on Advanced Information Networking and Applications 2008 (AINA-08).
- PC Member, The IEEE 21st International Conference on Advanced Information Networking and Applications 2007 (AINA-07).

CONSORTIUM MEMBERSHIPS

- 1000 Genomes Project Structural Variation Analysis Group, 2008-2015.
- Bonobo Genome Consortium, 2008-2012.
- Orangutan Genome Consortium, 2008-2011.
- Neandertal Genome Consortium, 2009-2012.
- Gorilla Genome Consortium, 2009-2013.
- Denisovan Genome Consortium, 2010-2012.
- Great Ape Diversity Project, 2011-2014.
- Genome in a Bottle Consortium, 2015-*present*

AD-HOC REVIEWER

- Nature Genetics, Nucleic Acids Research, Nature Reviews Genetics, Genome Research, PLoS Genetics, PLoS Computational Biology, Nature Biotechnology, BMC Genomics, Genome Biology, Methods, Transactions of Computational Biology and Bioinformatics, Bioinformatics, Briefings in Bioinformatics, RECOMB, RECOMB-seq, SODA, FOCS, ICDE, CPM, AINA, PSB, ISMB, HiTSeq.

TEACHING

- CS 476 Automata Theory and Formal Languages, Spring 2014, Fall 2014, Spring 2015, Fall 2015, Spring 2016.
- CS 481 Bioinformatics Algorithms, Fall 2012, Fall 2013.
- CS 319 Object Oriented Software Engineering, Summer 2012, Spring 2013, Fall 2013, Fall 2014, Fall 2015.
- CS 681 Advanced Topics in Computational Biology, Spring 2012, Spring 2013.
- CS 202 Fundamental Structures of Computer Science II, Spring 2013.
- CS 590/690 Research Topics (coordinator), Spring 2012, Spring 2014.

SENIOR PROJECT SUPERVISION

- BioPeer: Secure research data sharing platform, 2014-2015.
- Coinami: Cryptocurrency mining system for scientific computation, 2014-2015.
- Anotice: Anonymized social network platform, 2014-2015.
- Touravel: Travel logging application, 2014-2015.
- SUN (Social University Network): Study group game application, 2014-2015.
- SmartVote: Electronic Voting System for Mobile Devices, 2013-2014.
Recipient of the Best Senior Project Award in 2014.
- Daycept: Concept-based idea sharing platform, 2013-2014.
- Betcha: Social platform that enables betting among friends, 2013-2014.
Recipient of the Usability Award in 2014.
- SocioDiary: Digital diary system, 2013-2014.
- CMPS101: Activity planning for university students, 2013-2014.
- chkDNA: Genomic variant analysis tool, 2012-2013.

OTHER

- Member of the Scientific Advisory Board for the Center for Advanced Genomics and Bioinformatics (İGBAM) in TÜBİTAK Marmara Research Center, Gebze, Turkey. October 2012 - *present*.
- Observer to ELIXIR for Turkey. April 2013 - *present*.
- Individual Member, Global Alliance for Genomic Health. November 2014 - *present*.

RESEARCH GRANTS

ACTIVE

- National Institutes of Health (R01 HG006004), 2011-2015
Title: Novel algorithms and hardware designs for ultra-fast next-gen sequence analysis
Goal: Developing specialized hardware architectures to accelerate mapping reads generated with the high throughput sequencing platforms.
PI: Onur Mutlu
Co-PI: Can Alkan
- European Union Marie Curie Actions Career Integration Grant (PCIG10-GA-2011-303772), 2012-2016
Title: Integrated approaches for genomic variation discovery using high throughput sequencing
Goal: Developing novel combinatorial algorithms to comprehensively and quickly discover genomic variation.
PI: Can Alkan
- Scientific and Technical Research Council of Turkey (TÜBİTAK-1001-112E135), 2012-2016
Title: Development and application of novel genome assembly algorithms that use multiple data sources
Goal: Developing assembly algorithms to more reliably construct de novo genome assemblies using data from multiple sources.
PI: Can Alkan
- European Molecular Biology Organization Installation Grant (IG-2521), 2013-2016
Title: Development and application of computational methods to analyze next generation sequence data to characterize both normal and disease causing variation, and build de novo genome assemblies
Goal: Analysis of biological sequences generated with the next-gen sequencing platforms.
PI: Can Alkan

CONFERENCE AND INVITED TALKS

- Invited Speaker, **A cryptocurrency with DNA Sequence Alignment as Proof-of-work**. Middle East Technical University, Ankara, Turkey, October 21, 2015.
- Invited Speaker, **A cryptocurrency with DNA Sequence Alignment as Proof-of-work**. 9th International Symposium on Health Informatics and Bioinformatics (HIBIT 2015), Muğla, Turkey, October 16, 2015.
- Workshop Talk, **Can you really anonymize the donors of genomic data in today's digital world?** 10th International Workshop on Data Privacy Management (DPM 2015), September 21, 2015, Vienna, Austria.
- Workshop Short Talk, **Characterization of structural variation and segmental duplications for the GiaB Project**. Genome in a Bottle Workshop, National Institute of Standards and Technology, Gaithersburg, MD, USA, August 27, 2015.
- Invited Speaker, **Discovery of large genomic inversions using pooled clone sequencing**. Georgia Institute of Technology, Atlanta, GA, USA, August 26, 2015.
- Conference Talk, **Coinami: A cryptocurrency with DNA Sequence Alignment as Proof-of-work**. High Throughput Sequencing Algorithms and Applications (HiTSeq 2015), Dublin, Ireland, July 10, 2015.
- Invited Speaker, **Discovery of large genomic inversions using pooled clone sequencing**. İzmir Biomedicine and Genome Center, Dokuz Eylül University, İzmir, Turkey, April 20, 2015.
- Invited Speaker, **High Throughput Sequencing and Bioinformatics for characterization of diseases of genomic origin**. 24th Conference of the Asian Pacific Association for the Study of the Liver (APASL 2015), İstanbul, Turkey, March 11, 2015.
- Invited Speaker, **Genetik Hastalıkların Karakterizasyonunda Yüksek Ölçekli Dizileme ve Biyoenformatik**. (*in Turkish*), İzmir Biomedicine and Genome Center, Dokuz Eylül University, İzmir, Turkey, October 8, 2014

- Invited Speaker, **Characterization of genome structural variation and large inversions using NGS**. IU Bioinformatics Clinic, July 17, 2014.
- Conference Talk, **Characterization of large inversions using pooled clone sequencing**. European Molecular Biology Organization, Young Investigator Meeting, Heidelberg, Germany, May 14-16, 2014.
- Invited Lecturer, **Characterization of mobile element insertions using NGS**. ALLBIO Course on Next Generation Sequencing (NGS) methods for identification of mutations and large structural variants, Lausanne, Switzerland, March 11-12, 2014.
- Invited Speaker, **Next-generation sequence characterization of complex genome structural variation**. İzmir Biomedicine and Genome Center, Dokuz Eylül University, İzmir, Turkey, October 8, 2014.
- Instructor, **EMBO Practical Course: Computational biology: From genomes to cells and systems**. Cappadocia, Nevşehir, Turkey. September 29 - October 4, 2013.
- Invited Speaker, **Next-generation sequence characterization of complex genome structural variation**. Sabancı University, Tuzla, İstanbul, Turkey. April 17, 2013.
- Invited Speaker, **Yeni nesil dizileme ile karmaşık genom yapısal farklılıkların karakterizasyonu**. (*in Turkish*), MBG Weekend Seminars VIII, Boğaziçi University, March 23, 2013.
- Invited Speaker, **Genetik Hastalıkların Keşfinde Yüksek Ölçekli Dizileme ve Biyoinformatik**. (*in Turkish*), 10th National Medical Genetics Congress, Bursa, Turkey, December 20, 2012.
- Invited Speaker, **Next-generation sequence characterization of complex genome structural variation**. Swiss Institute of Allergy and Asthma Research (SIAF), Davos, Switzerland, September 3, 2012.
- Invited Speaker, **Characterization of mobile element insertions using high throughput sequencing**. 63rd Fujihara Seminar: A new horizon of retroposon research, Kyoto, Japan, August 2, 2012.
- Invited Speaker, **Next-generation sequence characterization of complex genome structural variation**. School of Medicine, Koç University, İstanbul, Turkey, May 21, 2012.
- Invited Speaker, **Next-generation sequence characterization of complex genome structural variation**. Department of Molecular Biology and Genetics, Bilkent University, Ankara, Turkey, March 14, 2012.
- Invited Speaker, **Initial analysis results of the Turkish Genome Project**. Boğaziçi University, İstanbul, Turkey, January 20-21, 2012.
- Invited Speaker, **Alu repeat discovery and characterization within human genomes**. FASEB Meeting on Mobile DNA in Mammalian Genomes, Snowmass Village, CO, USA, August 12, 2011.
- Invited Speaker, **Next-generation sequence characterization of complex genome structural variation**. Department of Genetics, Yale University, New Haven, CT, USA, May 11, 2011.
- Invited Speaker, **Dark side of genomes: what is missing in current sequence assemblies?** Genome 10K Project Genome Assembly Workshop, Santa Cruz, CA, USA, March 16, 2011.
- Invited Speaker, **Next-generation sequence characterization of complex genome structural variation**. Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA, February 22, 2011.
- Invited Speaker, **Next-generation sequence characterization of complex genome structural variation**. Keystone Symposium on Functional Consequences of Structural Variation in the Genome, Steamboat Springs, CO, USA, January 10, 2011.
- Tutorial Presenter, **Session tutorial: Personal Genomics**. Pacific Symposium for Biocomputing (PSB 2011), The Big Island of Hawaii, USA, January 3, 2011.
- Invited Speaker, **Discovery and characterization of copy-number variants with next-gen sequencing technologies**. TÜBİTAK Marmara Research Center, Gebze, Turkey, August 6, 2010.
- Invited Speaker, **Characterization of novel sequence insertions through both clone-based and clone-free sequencing**. Illumina Sequencing Expert Panel Meeting, Toronto, ON, Canada, July 22, 2010.
- Tutorial Presenter, **Session tutorial: Personal Genomics**. Pacific Symposium for Biocomputing (PSB 2010), January 4, 2010, The Big Island of Hawaii, USA.

- Invited Speaker, **Structural variation discovery and characterization of segmental duplications with next-gen sequencing technologies**. Simon Fraser University, Vancouver, BC, Canada, November 19, 2009.
- Invited Speaker, **Structural variation discovery and characterization of segmental duplications with next-gen sequencing technologies**. Universitat Pompeu Fabra, Barcelona, Spain, November 16, 2009.
- Invited Speaker, **Structural variation discovery and characterization of segmental duplications with next-gen sequencing technologies**. Università di Bologna, Bologna, Italy, November 11, 2009.
- Invited Speaker, **Structural variation discovery and characterization of segmental duplications with next-gen sequencing technologies**. Università degli Studi di Bari, Bari, Italy, November 9, 2009.
- Keynote Speaker, **Structural variation discovery and characterization of segmental Duplications with next-gen sequencing technologies**. Spanish and Portuguese Bioinformatics Conference (Jornadas de Bioinformática), Lisbon, Portugal, November 6, 2009.
- Invited Speaker, **Characterization of structural variation and segmental duplications using next-gen sequencing technology**. The Mediterranean Medical Genetics Meeting 2009, Ankara, Turkey, June 30, 2009.
- Conference Talk, **Organization and evolution of primate centromeric DNA from whole genome shotgun sequence data**. 16th Annual International Conference Intelligent Systems for Molecular Biology (ISMB 2008), *Highlights Track*, Toronto, ON, Canada, July 22, 2008.
- Workshop Talk, **Personalized copy-number and segmental duplication maps using next-gen sequencing technology**. Special Interest Group on Algorithms for Short Read Assembly, Alignment & Variation Analysis at the 16th Annual International Conference Intelligent Systems for Molecular Biology (ISMB 2008), Toronto, ON, Canada, July 19, 2008.
- Invited Speaker, **Organization and evolution of primate centromeric DNA from whole genome shotgun sequence data**. Bilkent University Center for Bioinformatics, Ankara, Turkey, July 20, 2007.
- Invited Speaker, **Towards a model for the evolution of alpha-satellite DNA**. BC Genome Sciences Centre, Vancouver, BC, Canada April 12, 2006.
- Conference Talk, **RNA-RNA interaction prediction and antisense RNA target search**. Ninth Annual Conference on Research in Computational Molecular Biology (RECOMB 2005), Cambridge, MA, USA, May 14, 2005.
- Workshop Talk, **RNA-RNA interactions and miRNA target prediction**. MITACS-PIMS Pacific Northwest Bioinformatics Day, Simon Fraser University, Burnaby, BC, Canada, May 20, 2004.

PUBLICATIONS

JOURNAL PUBLICATIONS

Joint first authors are marked with * in case of equal contribution. Joint last authors are marked with ‡.

Determining the origin of synchronous multifocal bladder cancer by exome sequencing. Ö. Acar*, E. Özkurt*, G. Demir, H. Saraç, **C. Alkan**, T. Esen, M. Somel‡, Nathan A Lack‡. *BMC Genomics*, to appear.

A global reference for human genetic variation. The 1000 Genomes Project Consortium. *Nature*, Oct 1; 526 (7571):98-74, 2015.

An integrated map of structural variation in 2,504 human genomes. P.H. Sudmant, T. Rausch, E.J. Gardner, R.E. Handsaker, A. Abyzov, J. Huddleston, Y. Zhang, K. Ye, G. Jun, M.H.-Y. Fritz, M.K. Konkel, A. Malhotra, A.M. Stütz, X. Shi, F.P. Casale, J. Chen, F. Hormozdiari, G. Dayama, K. Chen, M. Malig, M.J.P. Chaisson, K. Walter, S. Meiers, S. Kashin, E. Garrison, A. Auton, H.Y.K. Lam, X. J. Mu, **C. Alkan**, D. Antaki, T. Bae, E. Cerveira, P. Chines, Z. Chong, L. Clarke, E. Dal, L. Ding, S. Emery, X. Fan, M. Gujral, F. Kahveci, J.M. Kidd, Y. Kong, E-W. Lameijer, S. McCarthy, P. Flicek, R.A. Gibbs, G. Marth, C.E. Mason, A. Menelaou, D.M. Muzny, B.J. Nelson, A. Noor, N.F. Parrish, M. Pendleton, A. Quitadamo, B. Raeder, E.E. Schadt, M. Romanovitch, A. Schlattl, R. Sebra, A.A. Shabalín, A. Untergasser, J.A. Walker, M. Wang, F. Yu, C. Zhang, J. Zhang, X. Zheng-Bradley, W. Zhou, T. Zichner, J. Sebat, M.A. Batzer, S.A. McCarroll, The 1000 Genomes Project Consortium, R.E. Mills, M.B. Gerstein, A. Bashir, O. Stegle, S.E. Devine, C. Lee, E.E. Eichler, J.O. Korbel. *Nature*, Oct 1; 526 (7571):75–81, 2015.

Robustness of massively parallel sequencing platforms. P. Kavak, B. Yüksel, S. Aksu, M.O. Kulekci, T. Güngör, F. Hach, S.C. Sahinalp, Turkish Human Genome Project, C. Alkan[†], M.S. Sağiroğlu[‡]. *PLoS ONE*, Sep 18;10(9):e0138259, 2015.

Fast and accurate mapping of Complete Genomics reads. D. Lee, F. Hormozdiari, H. Xin, F. Hach, O. Mutlu[†], C. Alkan[†]. *Methods*, Jun;79-80:3-10, 2015.

Shifted Hamming Distance: a fast and accurate SIMD-friendly filter to accelerate alignment verification in read mapping. H. Xin, J. Greth, J. Emmons, G. Pekhimenko, C. Kingsford, C. Alkan[†], O. Mutlu[†]. *Bioinformatics*, May 15;31(10):1553-60, 2015.

Activating mutations of STAT5B and STAT3 in lymphomas derived from $\gamma\delta$ -T or NK cells. C. Kucuk, B. Jiang, X. Hu, W. Zhang, J. Chan, W. Xiao, N. Lack, C. Alkan, J. Williams, K. Avery, P. Kavak, A. Scuto, E. Sen, P. Gaulard, L. Staudt, J. Iqbal, W. Zhang, A. Cornish, Q. Gong, Q. Yang, H. Sun, F. d'Amore, S. Leppä, W. Liu, K. Fu, L. de Leval, T. McKeithan. *Nature Communications*, Jan 14;6:6025, 2015.

Whole genome sequencing of Turkish genomes reveals functional private alleles and impact of genetic interactions with Europe, Asia and Africa. C. Alkan, P. Kavak, M. Somel, O. Gokcumen, S. Ugurlu, E. Dal, K. Buğra-Bilge, T. Güngör, S.C. Sahinalp, N. Özören, C. Bekpen. *BMC Genomics*, 15 (1):963, 2014.

Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. M.J. Montague, G. Li, B. Gandolfi, R. Khan, B.L. Aken, S.M.J. Searle, P. Minx, L. Hillier, D.C. Koblodt, B.W. Davis, C.A. Driscoll, C.S. Barr, K. Blackistone, J. Quilez, B. Lorente-Galdos, T. Marques-Bonet, C. Alkan, G.W.C. Thomas, M. W. Hahn, M. Menotti-Raymond, S.J. O'Brien, R. Wilson, L.A. Lyons, W.J. Murphy, W.C. Warren. *Proc Natl Acad Sci*, Dec 2;111(48):17230-17235, 2014.

Annotated features of domestic cat – *Felis catus* genome. G. Tamazian, S. Simonov, P. Dobrynin, A. Makunin, A. Logachev, A. Komissarov, A. Shevchenko, V. Brukhin, N. Cherkasov, A. Svitin, K-P. Koepfli, J. Pontius, C. A Driscoll, K. Blackistone, C. Barr, D. Goldman, A. Antunes, J. Quilez, B. Lorente-Galdos, C. Alkan, T. Marques-Bonet, M. Menotti-Raymond, V.A. David, K. Narfström, S.J. O'Brien. *GigaScience*, Aug 5; 3:(13), 2014.

mrsFAST-Ultra: a compact, SNP-aware mapper for high performance sequencing applications. F. Hach*, I. Sarrafi*, F. Hormozdiari, C. Alkan, E.E. Eichler, S.C. Sahinalp. *Nucl Acids Research*, Jul;42(Web Server issue):W494-500, 2014.

Early postzygotic mutations contribute to de novo variation in a healthy monozygotic twin pair. G.M. Dal, B. Ergüner, M.S. Sağiroğlu, B. Yüksel, O.E. Onat, C. Alkan, T. Özçelik. *J Med Genet*, 51(7):455-459, 2014.

Genome Sequencing Highlights the Dynamic Early History of Dogs. A.H. Freedman, I. Gronau, R.M. Schweizer, D. Ortega-Del Vecchyo, E. Han, P.M. Silva, M. Galaverni, Z. Fan, P. Marx, B. Lorente-Galdos, H. Beale, O. Ramirez, F. Hormozdiari, C. Alkan, C. Vilà, K. Squire, E. Geffen, J. Kusak, A.R. Boyko, H.G. Parker, C. Lee, V. Tadigotla, A. Siepel, C.D. Bustamante, T.T. Harkins, S.F. Nelson, E.A. Ostrander, T. Marques-Bonet, R.K. Wayne, J. Novembre. *PLoS Genetics*, 10(1): e1004016, 2014.

Reconstructing complex regions of genomes using long-read sequencing technology. J. Huddleston, S. Ranade, M. Malig, F. Antonacci, M. Chaisson, L. Hon, P.H. Sudmant, T.A. Graves, C. Alkan, M.Y. Dennis, R.K. Wilson, S.W. Turner, J. Korlach, E.E. Eichler. *Genome Research*, 24(4):688-96, 2014.

Genome Sequencing Highlights the Dynamic Early History of Dogs. A.H. Freedman, I. Gronau, R.M. Schweizer, D. Ortega-Del Vecchyo, E. Han, P.M. Silva, M. Galaverni, Z. Fan, P. Marx, B. Lorente-Galdos, H. Beale, O. Ramirez, F. Hormozdiari, C. Alkan, C. Vilà, K. Squire, E. Geffen, J. Kusak, A.R. Boyko, H.G. Parker, C. Lee, V. Tadigotla, A. Siepel, C.D. Bustamante, T.T. Harkins, S.F. Nelson, E.A. Ostrander, T. Marques-Bonet, R.K. Wayne, J. Novembre. *PLoS Genetics*, 10(1): e1004016, 2014.

Rates and patterns of great ape retrotransposition. F. Hormozdiari, M.K. Konkel, J. Prado-Martinez, G. Chiatante, I. Hernando-Herraez, J.A. Walker, B. Nelson, C. Alkan, P.H. Sudmant, J. Huddleston, C.R. Catacchio, A. Ko, M. Malig, C. Baker, T. Marques-Bonet, M. Ventura, M.A. Batzer, and E.E. Eichler. *Proc Natl Acad Sci*, Aug 13;110(33):13457-62, 2013.

Great ape genetic diversity and population history. J. Prado-Martinez, P.H. Sudmant, J.M. Kidd, H. Li, J.L. Kelley, B. Lorente-Galdos, K.R. Veeramah, A.E. Woerner, T.D. O'Connor, G. Santpere, A. Cagan, C. Theunert, F. Casals, H. Laayouni, K. Munch, A. Hobolth, A.E. Halager, M. Malig, J. Hernandez-Rodriguez, I. Hernando-Herraez, K. Prüfer, M. Pybus, L. Johnstone, M. Lachmann, **C. Alkan**, D. Twigg, N. Petit, C. Baker, F. Hormozdiari, M. Fernandez-Callejo, M. Dabad, M.L. Wilson, L. Stevison, C. Camprubí, T. Carvalho, A. Ruiz-Herrera, L. Vives, M. Mele, T. Abello, I. Kondova, R.E. Bontrop, A. Pusey, F. Lankester, J.A. Kiyang, R.A. Bergl, E. Lonsdorf, S. Myers, M. Ventura, P. Gagneux, D. Comas, H. Siegmund, J. Blanc, L. Agueda-Calpena, M. Gut, L. Fulton, S.A. Tishkoff, J.C. Mullikin, R.K. Wilson, I.G. Gut, M. Katherine Gonder, O.A. Ryder, B.H. Hahn, A. Navarro, J.M. Akey, J. Bertranpetit, D. Reich, T. Mailund, M.H. Schierup, C. Hvilsom, A.M. Andrés, J.D. Wall, C.D. Bustamante, M.F. Hammer, E.E. Eichler, T. Marques-Bonet. *Nature*, 499(7459):471-475, 2013.

The genome sequencing of an albino Western lowland gorilla reveals inbreeding in the wild. J. Prado-Martinez, I. Hernando-Herraez, B. Lorente-Galdos, M. Dabad, O. Ramirez, C. Baeza-Delgado, C. Morcillo-Suarez, **C. Alkan**, F. Hormozdiari, E. Raineri, J. Estellé, M. Fernandez-Callejo, M. Valles, L. Ritscher, T. Schöneberg, E. de la Calle-Mustienes, S. Casillas, R. Rubio-Acero, M. Melé, J. Engelken, M. Caceres, J.L. Gomez-Skarmeta, M. Gut, J. Bertranpetit, I.G. Gut, T. Abello, E.E. Eichler, I. Mingarro, C. Lalueza-Fox, A. Navarro, T. Marques-Bonet. *BMC Genomics*, May 31;14(1):363, 2013.

Refinement and Discovery of New Hotspots of Copy-Number Variation Associated with Autism Spectrum Disorder. S. Girirajan*, M.Y. Dennis*, C. Baker, M. Malig, B.P. Coe, C.D. Campbell, K. Mark, T.H. Vu, **C. Alkan**, Z. Cheng, L.G. Biesecker, R. Bernier, E.E. Eichler. *Am J Hum Genet.*, Feb 7;92(2):221-37, 2013.

Accelerating read mapping with FastHASH. H. Xin, D. Lee, F. Hormozdiari, S. Yedkar, O. Mutlu[‡], **C. Alkan**[‡]. *BMC Genomics*, 14(Suppl 1):S13, 2013.

Special issue for the 11th Asia Pacific Bioinformatics Conference, Jan. 21-23, 2013, Vancouver, BC, Canada

An integrated map of genetic variation from 1,092 human genomes. The 1000 Genomes Project Consortium. *Nature*, Nov 1;491(7422):56-65, 2012.

SCALCE: boosting Sequence Compression Algorithms using Locally Consistent Encoding. F. Hach, I. Numanagic, **C. Alkan**, S. Cenik Sahinalp. *Bioinformatics*, Dec 1;28(23):3051-57, 2012.

A high-coverage genome sequence from an archaic Denisovan individual. M. Meyer, M. Kircher, M.-T. Gansauge, H. Li, F. Racimo, S. Mallick, J.G. Schraiber, F. Jay, K. Prüfer, C. de Filippo, P.H. Sudmant, **C. Alkan**, Q. Fu, R. Do, N. Rohland, A. Tandon, M. Siebauer, R.E. Green, K. Bryc, A.W. Briggs, U. Stenzel, J. Dabney, J. Shendure, J. Kitzman, M.F. Hammer, M.V. Shunkov, A.P. Derevianko, N. Patterson, A.M. Andrés, E.E. Eichler, M. Slatkin, D. Reich, J. Kelso, S. Pääbo. *Science*, 338(6014):222-226, 2012.

The bonobo genome compared with the chimpanzee and human genomes. The International Gorilla Genome Sequencing and Analysis Consortium. *Nature*, 486(7404):527-531, 2012.

Insights into hominid evolution from the gorilla genome sequence. The International Gorilla Genome Sequencing and Analysis Consortium. *Nature*, 483(7388): 169-175, 2012.

Copy number variation of individual cattle genomes using next-generation sequencing. D.M. Bickhart, Y. Hou, S.G. Schroeder, **C. Alkan**, M.F. Cardone, L.K. Matukumalli, J. Song, R.D. Schnabel, M. Ventura, J.F. Taylor, J.F. Garcia, C.P. Van Tassell, T.S. Sonstegard, E.E. Eichler, G.E. Liu. *Genome Research*, Apr;22(4):778-90, 2012.

Detection of structural variants and indels within exome data. E. Karakoc, **C. Alkan**, B.J. O'Roak, M.Y. Dennis, L. Vives, K. Mark, M.J. Rieder, D.A. Nickerson, E.E. Eichler. *Nature Methods*, 9(2): 176-178, 2012.

Identification and validation of a novel mature microRNA encoded by the Merkel cell polyomavirus in human Merkel cell carcinomas. S. Lee, K.G. Paulson, E.P. Murchison, O.K. Afanasiev, **C. Alkan**, J.H. Leonard, D.R. Byrd, G.J. Hannon, P. Nghiem. *J Clin Virol*. Nov; 52(3):272-275, 2011.

A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. A.E. Renton, E. Majounie, A. Waite, J. Simón-Sánchez, S. Rollinson, J.R. Gibbs, J.C. Schymick, H. Laaksovirta, J.C. van Swieten, L. Myllykangas, H. Kalimo, A. Paetau, Y. Abramzon, A.M. Remes, A. Kaganovich, S.W. Scholz, J. Duckworth, J. Ding, D.W. Harmer, D.G. Hernandez, J.O. Johnson, K. Mok, M. Rytten, D. Trabzuni, R.J. Guerreiro, R.W. Orrell, J. Neal, A. Murray, J. Pearson, I.E. Jansen, D. Sondervan, H. Seelaar, D. Blake, K. Young, N. Halliwell, J.B. Callister, G. Toulson, A. Richardson, A. Gerhard, J. Snowden, D. Mann, D. Neary, M.A. Nalls, T. Peuralinna, L. Jansson, V.M. Isoviita, A.L. Kaivorinne, M. Hölttä-Vuori, E. Ikonen, R. Sulkava, M. Benatar, J. Wu, A. Chiò, G. Restagno, G. Borghero, M. Sabatelli, The ITALSGEN Consortium, D. Heckerman, E. Rogaeva, L. Zinman, J.D. Rothstein, M. Sendtner, C. Drepper, E.E. Eichler, **C. Alkan**, Z. Abdullaev, S.D. Pack, A. Dutra, E. Pak, J. Hardy, A. Singleton, N.M. Williams, P. Heutink, S. Pickering-Brown, H.R. Morris, P.J. Tienari, B.J. Traynor. *Neuron*, Oct 20; 72(2):257-268, 2011.

Gorilla genome structural variation reveals evolutionary parallelisms with chimpanzee. M. Ventura, C.R. Catacchio, **C. Alkan**, T. Marques-Bonet, S. Sajjadian, T.A. Graves, F. Hormozdiari, A. Navarro, M. Malig, C. Baker, C. Lee, E.H. Turner, L. Chen, J.M. Kidd, N. Archidiacono, J. Shendure, R.K. Wilson, E.E. Eichler. *Genome Research*, Oct;21(10):1640-9, 2011.

Sensitive and fast mapping of di-base encoded reads. F. Hormozdiari*, F. Hach*, S.C. Sahinalp, E.E. Eichler, **C. Alkan**. *Bioinformatics*, Jul 15;27(14):1915-21, 2011.

Alu repeat discovery and characterization within human genomes. F. Hormozdiari*, **C. Alkan***, M. Ventura*, I. Hajirasouliha, M. Malig, F. Hach, D. Yorukoglu, P. Dao, M. Bakshi, S.C. Sahinalp, E.E. Eichler. *Genome Research*, Jun;21(6):840-9, 2011.

Genome structural variation discovery and genotyping. **C. Alkan**, B.P. Coe, E.E. Eichler. *Nature Reviews Genetics*, 12:363-376, 2011.

Cln4-2 genomic structure differs between the X locus in Mus spretus and the autosomal locus in Mus musculus: AT motif enrichment on the X. D.K. Nguyen, F. Yang, R. Kaul, **C. Alkan**, A. Antonellis, K.F. Friery, B. Zhu, P.J. de Jong, C.M. Disteche. *Genome Research*, 21(3):402-409, 2011.

Mapping copy number variation at fine scale by population scale genome sequencing. .E. Mills*, K. Walter*, C. Stewart*, R.E. Handsaker*, K. Chen*, **C. Alkan***, A. Abyzov*, S.C. Yoon*, K. Ye*, R.K. Cheetham, A. Chinwalla, D.F. Conrad, Y. Fu, F. Grubert, I. Hajirasouliha, F. Hormozdiari, L.M. Iakouchava, Z. Iqbal, S. Kang, J.M. Kidd, M.K. Konkel, J. Korn, E. Khurana, D. Kural, H.Y.K. Lam, J. Leng, R. Li, Y. Li, C-Y. Lin, R. Luo, X.J. Mu, J. Nemesh, H.E. Peckham, T. Rausch, A. Scally, X. Shi, M.P. Stromberg, A.M. Stütz, A. E. Urban, J.A. Walker, J. Wu, Y. Zhang, Z.D. Zhang, M.A. Batzer, L. Ding, G.T. Marth, G. McVean, J. Sebat, M. Snyder, J. Wang, K. Ye, E.E. Eichler, M.B. Gerstein, M.E. Hurles, C. Lee, S.A. McCarroll, J.O. Korb. *Nature*, 470(7332):56-65, 2011.

Comparative and demographic analysis of orangutan genomes. International Orangutan Genome Sequencing and Analysis Consortium. *Nature*, 469(7331):529-533, 2011.

Haplotype resolved genome sequencing of a Gujarati Indian individual. J.O. Kitzman, A.P. MacKenzie, A. Adey, J.B. Hiatt, R.P. Patwardhan, P.H. Sudmant, S.B. Ng, **C. Alkan**, R. Qiu, E.E. Eichler, J. Shendure. *Nature Biotechnology*, 29(1):59-63, 2011.

Limitations of next-generation genome assembly. **C. Alkan**, S. Sajjadian, E.E. Eichler. *Nature Methods*, 8(1):61-65, 2011.

Highlighted in "Assemblies: the good, the bad, the ugly", E. Birney, Nature Methods, 8(1):59-60, 2011.

Genome-wide characterization of centromeric satellites from multiple mammalian genomes. **C. Alkan***, M.F. Cardone*, C.R. Catacchio, F. Antonacci, S.J. O'Brien, O.A. Ryder, S. Purgato, M. Zoli, G. Della Valle, E.E. Eichler, M. Ventura. *Genome Research*, 21(1): 137-145, 2011.

Genetic history of an archaic hominin group from Denisova Cave in Siberia. D. Reich, R.E. Green, M. Kircher, J. Krause, N. Patterson, E.Y. Durand, B. Viola, A.W. Briggs, U. Stenzel, P.L.F. Johnson, T. Maricic, J.M. Good, T. Marques-Bonet, **C. Alkan**, Q. Fu, S. Mallick, H. Li, M. Meyer, E.E. Eichler, M. Stoneking, M. Richards, S. Talamo, M.V. Shunkov, A.P. Derevianko, J.-J. Hublin, J. Kelso, M. Slatkin, S. Pääbo. *Nature*, Dec; 468(7327):1053-1060, 2010.

Diversity of human copy number variation and multicopy genes. P.H. Sudmant, J.O. Kitzman, F. Antonacci, **C. Alkan**, M. Malig, A. Tsalenko, N. Sampas, L. Bruhn, J. Shendure, The 1000 Genomes Project Consortium, E.E. Eichler. *Science*, Oct; 330(6004):641-646, 2010.

A map of human genome variation from population scale sequencing. The 1000 Genomes Project Consortium. *Nature*, Oct; 467(7319):1061-1073, 2010.

A large and complex structural polymorphism at 16p12.1 underlies microdeletion disease risk. F. Antonacci, J.M. Kidd, T. Marques-Bonet, B. Teague, M. Ventura, S. Girirajan, **C. Alkan**, C.D. Campbell, L. Vives, M. Malig, J.A. Rosenfeld, B.C. Ballif, L.G. Shaffer, T.A. Graves, R.K. Wilson, D.C. Schwartz, E.E. Eichler. *Nature Genetics*, Sep; 42(9):745-750, 2010.

mrsFAST: a cache-oblivious algorithm for short-read mapping. F. Hach, F. Hormozdiari, **C. Alkan**, F. Hormozdiari, I. Birol, E.E. Eichler, S.C. Sahinalp. *Nature Methods*, Aug;7(8):576-7, 2010.

A draft sequence of the Neandertal genome. R.E. Green, J. Krause, A.W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M.H.-Y. Fritz, N.F. Hansen, E.Y. Durand, A.-S. Malaspina, J.D. Jensen, T. Marques-Bonet, **C. Alkan**, K. Prüfer, M. Meyer, H.A. Burbano, J.M. Good, R. Schultz, A. Aximu-Petri, A. Butthof, B. Höber, B. Höffner, M. Siegemund, A. Weihmann, C. Nusbaum, E.S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C. Verna, P. Rudan, D. Brajkovic, Z. Kucan, I. Gusic, V.B. Doronichev, L.V. Golovanova, C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R.W. Schmitz, P.L.F. Johnson, E.E. Eichler, D. Falush, E. Birney, J.C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M. Lachmann, D. Reich, S. Pääbo. *Science*, 7 May, 328 (5979):710-722, 2010.

Recipient of the 2010 AAAS Newcomb Cleveland Prize.

Characterization of missing human genome sequences and copy-number polymorphic insertions. J.M. Kidd, N. Sampas, F. Antonacci, T. Graves, R. Fulton, H.S. Hayden, **C. Alkan**, M. Malig, M. Ventura, G. Giannuzzi, J. Kalicki, P. Anderson, A. Tsalenko, N.A. Yamada, P. Tsang, R. Kaul, R.K. Wilson, L. Bruhn, E.E. Eichler. *Nature Methods*, May, 7 (5):365-371, 2010.

Next-generation VariationHunter: combinatorial algorithms for transposon insertion discovery. F. Hormozdiari, I. Hajirasouliha, P. Dao, F. Hach, D. Yorukoglu, **C. Alkan**, E.E. Eichler, S.C. Sahinalp. *Bioinformatics*, Jun 15; 26(12):i350-i357, 2010.

Detection and characterization of novel sequence insertions using paired-end next-generation sequencing. I. Hajirasouliha*, F. Hormozdiari*, **C. Alkan***, J.M. Kidd, I. Birol, E.E. Eichler, S.C. Sahinalp. *Bioinformatics*, May 15;26(10):1277-83, 2010.

Complete Khoisan and Bantu genomes from southern Africa. S.C. Schuster, W. Miller, A. Ratan, L.P. Tomsho, B. Giardine, L.R. Kason, R.S. Harris, D.C. Petersen, F. Zhao, J. Qi, **C. Alkan**, J.M. Kidd, Y. Sun, D.I. Drautz, P. Bouffard, D.M. Muzny, J.G. Reid, L.V. Nazareth, Q. Wang, R. Burhans, C. Riemer, N.E. Wittekindt, P. Moorjani, E.A. Tindall, C.G. Danko, W.S. Teo, A.M. Buboltz, Z. Zhang, Q. Ma, A. Oosthuisen, A.W. Steenkamp, H. Oostuisen, P. Venter, J. Gajewski, Y. Zhang, B.F. Pugh, K.D. Makova, A. Nekrutenko, E.R. Mardis, N. Patterson, T.H. Pringle, F. Chiaromonte, J.C. Mullikin, E.E. Eichler, R.C. Hardison, R.A. Gibbs, T.T. Harkins, V.M. Hayes. *Nature*, Feb, 463(7283):943-947, 2010.

Personalized copy-number and segmental duplication maps using next-generation sequencing. **C. Alkan**, J.M. Kidd, T. Marques-Bonet, G. Aksay, F. Antonacci, F. Hormozdiari, J. O. Kitzman, C. Baker, M. Malig, O. Mutlu, S.C. Sahinalp, R. A. Gibbs, E.E. Eichler. *Nature Genetics*, Oct, 41(10):1061-1067, 2009.

Highlighted in "Mapping duplicated sequences", DY Chiang and SA McCarroll, Nature Biotechnology, Nov; 27(11):1001-2, 2009.

Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two base encoding. K.J. McKernan, H.E. Peckham, G.L. Costa, S.F. McLaughlin, Y. Fu, E.F. Tsung, C.R. Clouser, C. Duncan, J.K. Ichikawa, C.C. Lee, Z. Zhang, S.S. Ranade, E.T. Dimalanta, F.C. Hyland, T.D. Sokolsky, L. Zhang, J.A. Sheridan, H. Fu, C.L. Hendrickson, B. Li, L. Kotler, J.R. Stuart, J.A. Malek, J.M. Manning, A.A. Antipova, D.S. Perez, M.P. Moore, K.C. Hayashibara, M.R. Lyons, R.E. Beaudoin, B.E. Coleman, M.W. Laptewicz, A.E. Sannicandro, M.D. Rhodes, R.K. Gottimukkala, S. Yang, V. Bafna, A. Bashir, A. MacBride, **C. Alkan**, J.M. Kidd, E.E. Eichler, M.G. Reese, F.M. De La Vega, A.P. Blanchard. *Genome Research*, Sep, 19(9): 1527-1541, 2009.

New insights into centromere organization and evolution from the white-cheeked gibbon and marmoset. A. Cellamare, C.R. Catacchio, **C. Alkan**, G. Giannuzzi, F. Antonacci, M.F. Cardone, G. Della Valle, M. Malig, M. Rocchi, E.E. Eichler, M. Ventura. *Mol Biol Evol*, Aug; 26(8):1889-1900, 2009.

MoDIL: detecting small indels from clone-end sequencing with mixtures of distributions. S. Lee, F. Hormozdiari, **C. Alkan**, M. Brudno. *Nature Methods*, Jul, 6(7):473-4, 2009.

Combinatorial algorithms for structural variation detection in high throughput sequenced genomes. F. Hormozdiari*, **C. Alkan***, E.E. Eichler, S.C. Sahinalp. *Genome Research*, Jul, 19(7):1270-8, 2009.

Comparative analysis of Alu repeats in primate genomes. G.E. Liu, **C. Alkan**, L. Jiang, S. Zhao, E.E. Eichler. *Genome Research*, 19(5):876-885, 2009.

Death and resurrection of the human *IRGM* gene. C. Bekpen, T. Marques-Bonet, **C. Alkan**, F. Antonacci, M. B. Leogrande, M. Ventura, J.M. Kidd, P. Siswara, J. C. Howard, E.E. Eichler. *PLoS Genetics*, March 2009; 5(3):e1000403, 2009.

A burst of segmental duplications in the genome of the African great ape ancestor. T. Marques-Bonet, J.M. Kidd, M. Ventura, T.A. Graves, Z. Cheng, L.W. Hillier, Z. Jiang, C. Baker, R. Malfavon-Borja, L.A. Fulton, **C. Alkan**, G. Aksay, S. Girirajan, P. Siswara, L. Chen, M.F. Cardone, A. Navarro, E.R. Mardis, R.K. Wilson, E.E. Eichler. *Nature*, Feb 12 2009; 457(7231):877-881, 2009.

Mapping and sequencing of structural variation from eight human genomes. J.M. Kidd, G.M. Cooper, W.F. Donahue, H.S. Hayden, N. Sampas, T. Graves, N. Hansen, B. Teague, **C. Alkan**, F. Antonacci, E. Haugen, T. Zerr, N.A. Yamada, P. Tsang, T.L. Newman, E. Tüzün, Z. Cheng, H.M. Ebling, N. Tusneem, R. David, W. Gillett, K.A. Phelps, D. Saranga, A. Brand, W. Tao, E. Gustafson, K. McKernan, L. Chen, M. Malig, J.D. Smith, S.A. McCarroll, D.A. Altshuler, D.A. Peiffer, M. Dorschner, J. Stamatoyannopoulos, D. Schwartz, D.A. Nickerson, J.C. Mullikin, R.K. Wilson, L. Bruhn, M.V. Olson, R. Kaul, D.R. Smith, E.E. Eichler. *Nature*, May 1, 2008; 453(7191):56-64, 2008.

Optimal design of oligonucleotide microarrays for measurement of DNA copy number. A.J. Sharp, A. Itsara, Z. Cheng, **C. Alkan**, S. Schwartz, E.E. Eichler. *Human Molecular Genetics*, Nov 15; 16(22):2770-2779, 2007.

Organization and evolution of primate centromeric DNA from whole genome shotgun sequence data. **C. Alkan**, M. Ventura, N. Archidiacono, M. Rocchi, S.C. Şahinalp, E.E. Eichler. *PLoS Computational Biology*, 3(9): e181, September 2007.

taveRNA: a web suite for RNA algorithms and applications. C. Aksay, R. Salari, E. Karakoç, **C. Alkan**, S.C. Şahinalp. *Nucleic Acids Research*, July 1, 2007; Vol. 35 (Web Server Issue): W325-329, 2007.

RNA-RNA interaction prediction and antisense RNA target search. **C. Alkan**, E. Karakoç, J. Nadeau, S.C. Şahinalp, K. Zhang. *Journal of Computational Biology*, March 2006; Vol 13, No 2: 267-282, 2006.

Manipulating multiple sequence alignments via MaM and WebMaM. **C. Alkan**, E. Tüzün, J. Buard, F. Lethiec, E.E. Eichler, J.A. Bailey, S.C. Şahinalp. *Nucleic Acids Research*, July 1, 2005; Vol. 33 (Web Server issue): W295-W298, 2005.

The role of unequal crossover in alpha-satellite DNA evolution: a computational analysis. **C. Alkan**, E.E. Eichler, J.A. Bailey, S.C. Şahinalp, E. Tüzün. *Journal of Computational Biology*, Vol. 11, No.5: 933-944, 2004.

The structure and evolution of centromeric transition regions within the human genome. X. She, J.E. Horvath, Z. Jiang, G. Liu, T. S. Furey, L. Christ, R. Clark, T. Graves, C.L. Gulden, **C. Alkan**, J.A. Bailey, S.C. Şahinalp, M. Rocchi, D. Haussler, R.K. Wilson, W. Miller, S. Schwartz, E.E. Eichler. *Nature*, Aug 19, 2004; 430(7002):857-64, 2004.

Analysis of primate genomic variation reveals a repeat-driven expansion of the human genome. G. Liu, NISC Comparative Sequencing Program, S. Zhao, J.A. Bailey, S.C. Şahinalp, **C. Alkan**, E. Tüzün, E.D. Green, E.E. Eichler. *Genome Research*, 13(3):358-368, 2003.

Divergent origins and concerted expansion of two segmental duplications on chromosome 16. E.E. Eichler, M.E. Johnson, **C. Alkan**, E. Tüzün, C. Şahinalp, D. Misceo, N. Archidiacono, and M. Rocchi. *Journal of Heredity*, 92: 462-468, 2001.

PEER-REVIEWED CONFERENCE PUBLICATIONS

Note: In CS disciplines papers in highly ranked conferences are fully refereed and are regarded as highly as the top journal papers. In computational molecular biology RECOMB is generally accepted as the top conference, with 10-15% acceptance rate. Joint first authors are marked with * in case of equal contribution. Joint last authors are marked with ‡.

Can you really anonymize the donors of genomic data in today's digital world? M. Alser, N. Almadhoun, A. Nouri, **C. Alkan**[‡], and E. Ayday[†]. *10th International Workshop on Data Privacy Management (DPM 2015)*, September 21-22, 2015, Vienna, Austria.

Improving genome assemblies using multi-platform sequence data. P. Kavak, B. Ergüner, D. Üstek, B. Yüksel, M.Ş. Sağıroğlu, T. Güngör and **C. Alkan**. *12th Computational Intelligence methods for Bioinformatics and Biostatistics (CIBB 2015)*, September 10-12, 2015, Naples, Italy.

Accelerating read mapping with FastHASH. H. Xin, D. Lee, F. Hormozdiari, S. Yedkar, O. Mutlu[‡], C. Alkan[‡]. *Proc. of the 11th Asia Pacific Bioinformatics Conference, Jan. 21-23, 2013, Vancouver, BC, Canada*

Detection of characterization of novel sequence insertions using paired-end next-generation sequencing, I. Hajirasouliha*, F. Hormozdiari*, C. Alkan*, J.M. Kidd, I. Birol, E.E. Eichler, S.C. Sahinalp. *Proc. of the 11th Conference on High Throughput Sequencing Analysis and Algorithms, Special Interest Group of ISMB 2010 (HiTSeq 2010): July 9-10 2010, Boston, MA, USA*

Next-generation VariationHunter: combinatorial algorithms for transposon insertion discovery. F. Hormozdiari, I. Hajirasouliha, P. Dao, F. Hach, D. Yorukoglu, C. Alkan, E.E. Eichler, S.C. Sahinalp. *Proc. of the 18th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB 2010), July 11-13, 2010, Boston, MA, USA*

Combinatorial algorithms for structural variation detection in high throughput sequenced genomes. F. Hormozdiari*, C. Alkan*, E.E. Eichler, S.C. Sahinalp. *Proc. of the Thirteenth Annual International Conference on Research in Computational Molecular Biology (RECOMB 2009), May 18-21, 2009 Tucson, AZ, USA*

RNA secondary structure prediction via energy density minimization. C. Alkan*, E. Karakoç*, S.C. Sahinalp, P. Unrau, H. A. Ebhardt, K. Zhang, J. Buhler. *Proc. of the Tenth Annual International Conference on Research in Computational Molecular Biology (RECOMB 2006 - LNBI 3909), pp. 130-142, April 2-5, 2006 Venice, Italy, 2006.*

RNA-RNA interaction prediction and antisense RNA target search. C. Alkan, E. Karakoç, J. Nadeau, S.C. Sahinalp, K. Zhang. *Proc. of the Ninth Annual International Conference on Research in Computational Molecular Biology (RECOMB 2005 - LNBI 3500), pp 152-171, May 14-18, 2005 Cambridge, MA, USA, 2005.*

An algorithmic analysis of the role of unequal crossover in alpha-satellite DNA evolution. C. Alkan, J.A. Bailey, E.E. Eichler, S.C. Sahinalp, E. Tüzün. *The 13th International Conference on Genome Informatics, December 16-18, 2002, Tokyo, Japan, vol.18 pp 93-102, 2002.*

UNPUBLISHED CONFERENCE TALKS AND POSTERS

Massively Parallel Mapping of Next Generation Sequence Reads Using GPUs.. A. Nouri, R.O. Selvitopi, O. Ozturk, O. Mutlu[‡], C. Alkan[‡]. *The 20th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2015), March 14-18, 2015, İstanbul, Turkey. Short talk and poster presented by Azita Nouri.*

BioPeer: A Fast and Secure Peer-to-Peer Data Sharing Tool.. C. Oge, F.T. Dogan, G. Goktepe, F. Koc, C. Sevim, C. Alkan. *The 6th Conference on High Throughput Sequencing Analysis and Algorithms (HiTSeq 2015), July 10-11, 2015, Dublin, Ireland. Poster presented by Cihad Öge.*

Coinami: A Cryptocurrency with DNA Sequence Alignment as Proof-of-work., A.M. Ileri, H.I. Ozercan, A. Gundogdu, A.K. Senol, M.Y. Ozkaya, C. Alkan. *The 6th Conference on High Throughput Sequencing Analysis and Algorithms (HiTSeq 2015), July 10-11, 2015, Dublin, Ireland. Oral and poster presentation by H. İbrahim Özercan.*

Discovery of large genomic inversions using pooled clone sequencing. M.E. Rasekh, G. Chiatante, M. Miroballo, J. Tang, M. Ventura, C.T. Amemiya, E.E. Eichler, F. Antonacci[‡], C. Alkan[‡]. *The 6th Conference on High Throughput Sequencing Analysis and Algorithms (HiTSeq 2015), July 10-11, 2015, Dublin, Ireland. Poster presentation by Marzieh Eslami Rasekh.*

A new inference attack against kin genomic privacy. F. Balci, H. Kulan, C. Alkan[‡], and E. Ayday[‡]. *Privacy-aware computational genomics (PRIVAGEN 2015), September 8, 2015, Tokyo, Japan. Oral and poster presentation by Fatma Balci.*

Identifying anonymous donors of genetic information. M. Alser, N. Almadhoun, A. Nouri, C. Alkan[‡], and E. Ayday[‡]. *Privacy-aware computational genomics (PRIVAGEN 2015), September 8, 2015, Tokyo, Japan. Poster presentation by Mohammed Alser.*

SOFTWARE

- Multiple alignment Manipulator (MaM).
<http://mam-bio.sourceforge.net>
- taveRNA: RNA suite for RNA folding, RNA-RNA interaction prediction and search.
<http://compbio.cs.sfu.ca/taverna>
- micro-read Fast Alignment Search Tools (mrFAST, mrsFAST, drFAST, and sirFAST).
<http://mrfast.sourceforge.net>, <http://mrsfast.sourceforge.net>, <http://drfast.sourceforge.net>,
<http://sirfast.sourceforge.net>
- mrCaNaVaR: characterization of segmental duplications and absolute copy number prediction from read depth.
<http://mrcanavar.sourceforge.net>
- VariationHunter: Structural variation calling algorithm for paired-end, next generation sequencing data.
<http://variationhunter.sourceforge.net>
- NovelSeq: computational pipeline fo detect novel sequence insertions using second generation sequencing.
<http://novelseq.sourceforge.net>
- RepeatNet: algorithm to characterize centromeric satellite sequences from paired end sequence data.
<http://eichlerlab.gs.washington.edu/software/repeatnet>
- SCALCE: Tool to compress FASTQ files.
<http://scalce.sourceforge.net>