## Brad Chapman


| | | |
|---|---|---|
| Contact Information | 35 Partridge Avenue<br>Somerville, MA 02145, USA | 1-617-447-8586<br>chapmanb@50mail.com |


**Education**

Ph.D., University of Georgia, Plant Biology, August 2004
B.S., Michigan State University, Botany and Plant Pathology, May 1999

**Community and code**

- Open source community involvement:

  Blog: Code sharing, scientific discussion and documentation
  Main organizer for community collaboration working groups (Codefest) since 2010
  Bioinformatics open source conference organizer since 2011

- Develop bcbio, providing validated, scalable, community developed variant calling and RNA-seq analysis
- Daily programming in Python, Clojure, R and Javascript. Familiarity with Perl, Java and C++. (code repositories: GitHub, Bitbucket).
- Comfortable analyzing large datasets in local clusters and on containerized cloud environments

**Experience**

**Harvard T.H. Chan School of Public Health** Research Scientist. Mar 2011-Present

- Established external collaborations with multiple partners in both academia and industry to enable critical infrastructure development
- Embeded bcbio platform within multiple open source and commercial analysis platforms using community workflow standards
- Provided scaling and analysis for large scale sequencing projects
- Building reference standards and validation to ensure accuracy and reproducibility of analyses
- Custom research support in collaboration with researchers, including variant calling, RNA-seq, detection of low frequency HIV populations, shRNA, and transposon insertion analysis. (code for recent projects).

**Massachusetts General Hospital** Bioinformatics Specialist II. Sep 2008-Mar 2011

- Custom next-generation sequencing analysis in collaboration with researchers in the hospital; RNA-seq, ChIP-seq, short RNA, and SNP analyses for a wide variety of organisms including Human, Mouse, Arabidopsis, C elegans, and Drosophila. (code for recent projects).
- Developed LIMS integrating our Illumina sequencing core with the Galaxy platform (code). Included an automated pipeline to process sequencing data, allowing our group to focus on custom analyses while providing researchers with rapid access to their raw sequences and alignments (code).

**Gen9** Bioinformatics consultant. Nov 2009-Oct 2010

Part time consultant at synthetic biology startup. Created a server automating a novel synthetic construct design approach. Provided documentation, training and support for transferring system maintenance to internal programming team.

**Codon Devices** Research Scientist, Bioinformatics. 2005-2008

- Designed and implemented automated oligo-based gene synthesis strategies that drove all revenue generating business. Combined existing academic approaches with internal development efforts using continuous feedback from production pipeline.
- Provided informatics platform for unique synthesis products, including large variant libraries and long construct assemblies.
- Directed remote programming team in Bangladesh to provide an e-commerce website.

**Plate Genome Mapping Laboratory, University of Georgia** PhD student. 1999-2004

- Developed phylogentic algorithms for dating duplication events in plant species using whole genome comparative data.
- Evaluated duplicate gene evolution through analysis of SNP accumulation, generating a new hypothesis explaining the prevalence of large scale duplications in plants.

**University of Georgia** Teaching assistant. 2002-2003

Teaching Assistant – Introduction to Gene Technology; Bioinformatics Applications.

**USDA Beltsville Agricultural Research Center, Autar Mattoo** Summer research scientist. 1999

RNA isolation and detection of transgenic tomatoes by RT-PCR and northern hybridization.

**Cornell University, Susan McCouch** Summer research scientist. 1998

Sub-cloning and sequencing of rice bacterial artificial chromosomes.

**University of Minnesota, Neil Olszewski** Summer research scientist. 1997

Cloning and protein expression in *Escherichia coli*.

**Michigan State University, Mariam Sticklen** Research scientist. 1996-1999

Development of transgenic maize by particle bombardment and detection via PCR and Southern hybridization.

## Publications


Li A, Hooli B, Mullin K, Tate RE, Bubnys A, Kirchner R, Chapman B, Hofmann O, Hide W, Tanzi RE. Silencing of the Drosophila ortholog of SOX5 leads to abnormal neuronal development and behavioral impairment. 2017 *Human Molecular Genetics*

Ahdesmaki M, Chapman BA, Cingolani P, Hofmann O, Sidoruk A, Lai Z, Kazkarov G, Rodichenko, Alperovich M, Jenkins D, Carr TH, Stetson D, Doughery B, Barrett JC, Johnson JH. Prioritisation of structural variant calls in cancer genomes. 2017. *PeerJ*

Filippova A, Chapman B, Geiger RS, Herbsleb JD, Kalyanasundaram A, Trainer E, Moser A, Stolzfus A. Hacking and Making at Time-Bounded Events: Current Trends and Next Steps in Research and Event Design. 2017. *Computer Supported Cooperative Work and Social Computing*

Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, McEwen R, Johnson J, Dougherty B, Barrett JC, Dry JR. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. 2016. *Nucleic Acids Research*

Lindstrom S, Ablorh A, Chapman B, Gusev A, Chen G, Tuman C, Eliassen H, Price AL, Henderson BE, Le Marchand L, Hofmann O, Haiman CA, Kraft P. Deep targeted sequencing of 12 breast cancer susceptibility regions in 4611 women across four different ethnicities. 2016. *Breast Cancer Research*

Harris NL, Cock PJA, Chapman B, Fields CJ, Hokamp K, Lapp H, Munoz-Torres M, Wiencko H. The 2016 Bioinformatics Open Source Conference (BOSC). 2016. *F1000 Research*

Harris NL, Cock PJ, Chapman BA, Goecks J, Hotz HR, Lapp H. The Bioinformatics Open Source Conference (BOSC). 2015. *Bioinformatics*

Zook JM, Chapman B, Wang J, Mittelman D, Hofmann O, Hide W, Salit M. Integrating human sequence data sets provides a resource of benchmark SNP and indel genotype calls. 2014. *Nature Biotechnology*

Sun J, Ramos A, Chapman B, Johnnidis JB, Le L, Ho YJ, Klein A, Hofmann O, Camargo FD. Clonal dynamics of native haematopoiesis. 2014. *Nature*

Li JZ, Chapman B, Charlebois P, Hofmann O, Weiner B, Porter AJ, Samuel R, Vardhanabhuti S, Zheng L, Eron J, Taiwo B, Zody MC, Henn MR, Kuritzkes DR, Hide W; ACTG A5262 Study Team, Wilson CC, Berzins BI, Acosta EP, Bastow B, Kim PS, Read SW, Janik J, Meres DS, Lederman MM, Mong-Kryspin L, Shaw KE, Zimmerman LG, Leavitt R, De La Rosa G, Jennings A. omparison of illumina and 454 deep sequencing in participants failing raltegravir-based antiretroviral therapy. 2014. *PLoS One*

Mller S, Afgan E, Banck M, Bonnal RJ, Booth T, Chilton J, Cock PJ, Gumbel M, Harris N, Holland R, Kala M, Kajn L, Kibukawa E, Powel DR, Prins P, Quinn J, Sallou O, Strozzi F, Seemann T, Sloggett C, Soiland-Reyes S, Spooner W, Steinbiss S, Tille A, Travis AJ, Guimera R, Katayama T, Chapman BA. Community-driven development for computational biology at Sprints, Hackathons and Codefests. 2014. *BMC Bioinformatics*

Paila U, Chapman BA, Kirchner R, Quinlan AR. GEMINI: integrative exploration of genetic variation and genome annotations. 2013. *PLoS Comput Biol*

Tabach Y, Billi AC, Hayes GD, Newman MA, Zuk O, Gabel H, Kamath R, Yacoby K, Chapman B, Garcia SM, Borowsky M, Kim JK, Ruvkun G. Identification of small RNA pathway genes using patterns of phylogenetic conservation and divergence. 2013. *Nature*

Lieber DS, Calvo SE, Shanahan K, Slate NG, Liu S, Hershman SG, Gold NB, Chapman BA, Thorburn DR, Berry GT, Schmahmann JD, Borowsky ML, Mueller DM, Sims KB, Mootha VK. Targeted exome sequencing of suspected mitochondrial disorders. 2013. *Neurology*

Katayama T et al. The 3rd DBCLS BioHackathon: improving life science data integration with Semantic Web technologies. 2013. *J Biomed Semantics*

Talevich E, Invergo BM, Cock PJ, Chapman BA. Bio.Phylo: A unified toolkit for processing, analyzing and visualizing phylogenetic trees in Biopython. 2012. *BMC Bioinformatics.*

Afgan E, Chapman B, Taylor J. CloudMan as a platform for tool, data, and analysis distribution. 2012. *BMC Bioinformatics*

Afgan E, Chapman B, Jadan M, Franke V, Taylor J. Using cloud computing infrastructure with CloudBioLinux, CloudMan, and Galaxy. 2012. *Curr Protoc Bioinformatics.*

Krampis K, Booth T, Chapman B, Tiwari B, Bicak M, Field D, and Nelson K. Cloud BioLinux: pre-configured and on-demand bioinformatics computing for the genomics community. 2012. *BMC Bioinformatics.*

Sansone SA et al. Toward interoperable bioscience data. 2012. *Nature Genetics.*

Ho Sui SJ, Begley K, Reilly D, Chapman B, McGovern R, Rocca-Sera P, Maguire E, Altschuler GM, Hansen TA, Sompallae R, Krivtsov A, Shivdasani RA, Armstrong SA, Culhane AC, Correll M, Sansone SA, Hofmann O, and Hide W. The Stem Cell Discovery Engine: an integrated repository and analysis system for cancer stem cell comparisons. 2012. *Nucleic Acids Res.*

Yuan CC, Matthews AGW, Jin Y, Chen CF, Chapman BA, Ohsumi TK, Glass KC, Kutateladze TG, Borowsky ML, Struhl K and, Oettinger MA. Histone H3R2 Symmetric Dimethylation and Histone H3K4 Trimethylation Are Tightly Correlated in Eukaryotic Genomes. 2012. *Cell Reports.*

Simon MD, Wang CI, Kharchenko PV, West JA, Chapman BA, Alekseyenko AA, Borowsky ML, Kuroda MI, and Kingston RE. The genomic binding sites of a noncoding RNA. 2011. *Proc Natl Acad Sci U S A.*.

Grau DJ, Chapman BA, Garlick JD, Borowsky M, Francis NJ, and Kingston RE. Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. 2011. *Genes Dev.*

Afgan E, Baker D, Coraor N, Chapman B, Nekrutenko A, and J Taylor. Galaxy CloudMan: delivering cloud compute clusters. 2010. *BMC Bioinformatics.*

Lippow SM, Moon TS, Basu S, Yoon S, Li X, Chapman BA, Robison K, Lipovek D and KLJ Prather. Engineering enzyme specificity using computational design of a defined-sequence library. 2010. *Chem Biol*

Blake WJ, Chapman BA, Zindal A, Lee ME, Lippow SM and BM Baynes. Pairwise selection assembly for sequence-independent construction of long-length DNA. 2010. *Nucleic Acids Res*

Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B and MJL de Hoon. Biopython: freely available Python tools for computational molecular biology and bioinformatics. 2009. *Bioinformatics*

Afeyan N, Church G, Jacobson J, Baynes BM, Nesmith KG, Chapman BA and B Strack-Louge. Methods for Assembly of High Fidelity Synthetic Polynucleotides. 2006. WO/2006/044956

Chapman BA, Bowers JE, Feltus FA and AH Paterson. Buffering of crucial functions by paleologous duplicated genes may contribute cyclicality to angiosperm genome duplication. 2006. *Proc Natl Acad Sci USA*

Paterson AH, Chapman BA, Kissinger JC, Bower JE, Feltus FA and JC Estill. Many gene and domain families have convergent fates following independent whole-genome duplication events in *Arabidopsis*, *Oryza*, *Saccharomyces* and *Tetraodon.* 2006. *Trends Genet*

Chapman BA, Bowers JE, Schulze SR and AH Paterson. A comparative phylogenetic approach for dating whole genome duplication events. 2004. *Bioinformatics*

Paterson AH, Bowers JE and BA Chapman. Ancient polyplodization predating divergence of the cereals, and its consequences for comparative genomics. 2004. *Proc Natl Acad Sci USA*

Paterson AH, Bowers JE, Chapman BA, Peterson DG, Rong J and TM Wicker. Comparative genome analysis of monocots and dicots, towards characterization of angiosperm diversity. 2004. *Curr Opin Biotech*

Bowers JE, Chapman BA, Rong J and AH Paterson. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. 2003. *Nature*

Paterson AH, Bowers JE, Peterson DG, Estill JC and BA Chapman. Structure and evolution of cereal genomes. 2003. *Curr Opin Genet Devel*

Zhong H, Teymouri F, Chapman B, Maqbool S, Sabzikar R, El-Maghraby Y, Dale B and MB Sticklen. The dicot pea (*Possum sativum* L.) rbcS transit peptide directs the *Alcaligenes eutrophus* polyhydroxybutyrate enzymes into the monocot maize (*Zea mays* L.) chloroplasts. 2003. *Plant Sci*

Blair MW, Garris AJ, Iyer AS, Chapman BA, Kresovich S and SR McCouch. High resolution genetic mapping and candidate gene identification at the *xa5* locus for bacterial blight resistance in rice (*Oryza sativa* L.). 2003. *Theor Appl Genet*

Chapman BA, Jacobo CM, Bunch RA and AH Paterson. Cactus breeding and biotechnology. 2002. in *Cacti: Biology and Uses*