

VIVEK KRISHNAKUMARhello@vivekkrish.com | [Google Scholar](https://scholar.google.com/citations?user=VIVEK) | vivekkrish.com**EDUCATION**

School of Informatics, Indiana University, Bloomington

(Aug '07 - Aug '09)

Masters of Science (M.S), Bioinformatics

School of Biotechnology, SRM University (Tamil Nadu, India)

(Jul '03 – May '07)

Bachelor of Technology (B.Tech), Biotechnology**SKILLS**

Prog Languages	C#, Python, Perl, PHP, Java
Bioinformatics	Annotation MAKER, Trinity RNA-Seq, PASA, Cufflinks, TopHat, Salmon, BUSCO
	Assembly Opgen MapSolver, ALLPATHS-LG, Velvet, ABySS, SOAPdenovo
Tools	Alignment GMAP, NCBI-BLAST+, BLAT, MUMmer, LAST, Bowtie, BWA, FastQC
	Data/Viz InterMine, BioMart, JBrowse, GBrowse, IGV
	Toolsets SAMTools, bedtools, GenomeTools, UCSC Kent Tools
Operating Systems	Linux/UNIX, Mac OS X, Windows
Database Systems	PostgreSQL, MySQL, Sybase, SQLite
Web Development	CGI/Perl/Python, PHP, HTML5, CSS3, JavaScript, jQuery

WORK EXPERIENCE

Sr. Staff Bioinformatics Engineer

(Jan '21 - Present)

Staff Bioinformatics Engineer

(April '20 - Dec '20)

Sr. Bioinformatics Engineer

(Jun '19 – Mar '20)

Bioinformatics Engineer II

(May '18 – May '19)

Responsibilities:

- Lead the testing effort for data analysis software.
- Partner with software, bioinformatics developers and domain experts in designing, performing, and improving verification tests
- Design and Implement automated software verification framework and scripts
- Independently develop bioinformatics utilities in order to facilitate software testing and automation
- Improve and maintain in-house bioinformatics libraries, test automation frameworks and tools
- Analyze test results to ensure existing functionality and recommend corrective action

J. Craig Venter Institute (JCVI)

Lead Bioinformatics Engineer

(September '17 - Apr '18)

Sr. Bioinformatics Engineer

(Aug '12 – Aug '17)

Bioinformatics Engineer

(Jun '09 – Jul '12)

Responsibilities:

- Contribute to analysis infrastructure by developing implementations of new/existing algorithms and analysis tools for scientific discovery
- Provide research support to projects by contributing to data analysis, documenting results and providing informatics expertise

- Prepare, present and/or provide input as necessary for progress reports (either internal institute presentations or external conferences/meetings)
- Senior member of the Plant Genomics group, leading efforts related to model plant species such as *Arabidopsis thaliana*, *Medicago truncatula*, etc.

PROJECTS (JCVI)

Legume Federation (LegFed)

(Apr '15 – Present)

NSF funded collaborative project involving Iowa State University (ISU), National Center for Genome Resources (NCGR), CyVerse @ University of Arizona, and JCVI, to help foster increased use of data standards, distributed development, and comparative analysis across the legume family

Website: <https://www.legumefederation.org>

- Primary stakeholder & developer representing the *Medicago truncatula* genome database
- Member of working group tasked with establishing standards for data, metadata & web services
- Developed & contributed custom ETL components for the open-source InterMine Java/PostgreSQL data warehouse, and extensions to enable Synteny-based automated interlinking of compatible resources

Arabidopsis Information Portal (Araport)

(Sep '13 – Present)

NSF funded and BBSRC co-funded collaborative project involving JCVI, Texas Advanced Computing Center (TACC), and Cambridge University, UK, responsible for developing a flexible and community-extensible portal to catalyze the next generation of Arabidopsis research

Website: <https://www.araport.org>

- Involved in the planning, design, implementation and downstream data analysis efforts relating to the *Arabidopsis thaliana* col-0 genome annotation update pipeline
- Involved in key aspects of the Araport data warehouse implementation: data integration into InterMine and JBrowse, new feature development, setup and administration of tools
- Involved in the planning and implementation of genomic data migration from TAIR to Araport into the CHADO open-source biological database schema
- Presented work done by the Araport team and promoted the portal to the research community via invited conference talks and posters presentations
- Involved in the Araport user and developer community outreach efforts
- Led the manuscript writing efforts culminating in a publication describing the Araport infrastructure, portal, data warehouse and supporting tools, reported in *Nucleic Acids Research* (2015)

Reference Genome Curation of the model legume, *Medicago truncatula*

(Jun '09 – Aug '14)

NSF funded project geared towards completing the *Medicago truncatula* genome and its annotation

Website: <http://www.medicagogenome.org>

- Managed the genome annotation and data distribution for the *Medicago* genome. Collaborated with leading community scientists in data analysis and development of landmark publication in *Nature* (2011).
- Involved in periodic genome sequencing and assembly update efforts aimed toward improving the genome, leveraging key technologies like Optical maps and Genotyping-by-sequencing based genetic maps to help guide the genome assembly
- Managed the planning, design, implementation and downstream data analysis efforts relating to periodic structural and functional annotation updates

- Co-led the manuscript writing efforts culminating in a publication describing genome assembly and annotation improvements reported in *BMC Genomics* (2014)
- Managed in-house data storage infrastructure (databases) catering to the needs of the project
- Involved in the setting up and maintenance of a data distribution platform, *Medicago truncatula* Genome Database (MTGD) consisting of the project website, data mining and visualization interfaces, etc., accessible by collaborating groups and the research community
- Led the manuscript writing efforts culminating in a publication describing MTGD, reported in *Plant & Cell Physiology* (2015)
- Developed and maintained software & resources for community annotation initiatives
- Managed submission of assembly and annotation data to public data repositories (NCBI)

***Medicago truncatula* Bioinformatics Outreach**

(Jun '10 – Aug '14)

- Participated in planning and coordination of yearly workshop aimed to educate the research community on various aspects of bioinformatic data analysis
- Created didactic material and made presentations for several workshop components like genome assembly, annotation techniques, data management, unix command-line usage, etc.
- Setup and coordinated IT infrastructure requirements for local/remote attendee and presenter participation

Maize Cell Genomics

(Jan '10 – Present)

NSF-funded collaborative project between CSHL, UWYO and JCVI to generate reporter lines that express proteins tagged with fluorescent markers, using Confocal Laser Scanning Microscopy (CLSM) to visualize how the tagged proteins may be interacting and functioning.

Website: <http://maize.jcvi.org/cellgenomics>

- Provided informatics support for the collaborating scientists on the project
- Maintained and improved the public Maize Cell Genomics (MCG) project website and web-based resources catering to the Maize research community, exposing gene construct and CLSM imaging data
- Led the manuscript writing efforts culminating in a publication describing MCG, reported in *Plant & Cell Physiology* (2015)

Center for Genomics and Bioinformatics (CGB), IU Bloomington

Graduate Research Assistant

(Apr '08 – May '09)

Responsibilities:

- Involved in planning, design and development of web-based genomic data visualization tools
- Setup and maintained databases for the tools and software packages developed at the CGB
- Performed computational analysis & interpretation of sequence data produced by in-house labs
- Worked on Masters thesis titled “Developing a Computational Method to help solve the Sequencing Gap Closure problem”, aimed towards helping finish genomes with no closely sequenced relatives

PROJECTS (CGB)

Comparative Gene Cluster Viewer (CGCV)

(Oct '08 – Feb '09)

Website: <http://cas-bioinfo.cas.unt.edu/cgcv/>

- Involved in the development and testing of a tool for gene cluster visualization against multiple reference genomes (both prokaryotic and eukaryotic)

- Developed the front-end web interface using CGI/Perl and backend MySQL database that synchronized on a nightly basis with NCBI GenBank

Male Urethral Microbiome (MUM) Project

(Aug '08 – Sep '08)

Website: <http://www.microbiota.org/cgi-bin/mum/index.cgi>

- Analyzed data derived from 16s rRNA phylogeny, qPCR and metagenomic sequencing
- Performed sequence similarity and community diversity studies, utilizing tools like BLAST, PHYLIP, MEGA and Distance-based OTU and Richness Estimation (DOTUR)
- Setup and maintained project website to share data with community

School of Informatics and Computing, IU Bloomington

Associate Instructor (Teaching Assistant)

(Sep '07 – May '09)

Responsibilities:

- Instructed sophomore level course titled “Information Infrastructure: Programming in C”
- In charge of teaching the laboratory component of the course
- Graded weekly homework assignments, tests and exams
- Assisted students with coursework during scheduled office hours
- Worked in a team of 4 instructors, handling around 200 students

PUBLICATIONS

Peer Reviewed Manuscripts

ORCID: <https://orcid.org/0000-0002-5227-0200>

1. Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., ... **Krishnakumar, V.**, ... & Provart, N. J. (2017). ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *The Plant Cell*, 29(7), tpc.00073.2017
2. **Krishnakumar, V.**, Contrino, S., Cheng, C. Y., Belyaeva, I., Ferlanti, E. S., Miller, J. R., ... & Chan, A. P. (2016). ThaleMine: A Warehouse for Arabidopsis Data Integration and Discovery. *Plant and Cell Physiology*, pcw200.
3. Hamilton, E. P., Kapusta, A., Huvos, P. E., Bidwell, S. L., Zafar, N., Tang, H., Hadjithomas, M., **Krishnakumar, V.**, ... & Russ, C. (2016). Structure of the germline genome of Tetrahymena thermophila and relationship to the massively rearranged somatic genome. *eLife*, 5, e19090.
4. Cheng, C. Y., **Krishnakumar, V.**, Chan, A., Thibaud-Nissen, F., Schobel, S., & Town, C. D. (2016). Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. *The Plant Journal*.
5. Hanlon, M. R., Mock, S., Dooley, R., Moreira, W., Stubbs, J., Pence, E., Ferlanti, E., **Krishnakumar, V.**, Miller, J., Town, C., Vaughn, M. (2015) Araport: an Application Platform for Data Discovery. *Concurrency Computat.: Pract. Exper.*, 27(16): 4412-4422
6. Chamala, S., Garcia, N., Godden, G., **Krishnakumar, V.**, ..., Barbazuk, W. B., Soltis, D. E., Soltis, P. S. (2015) MarkerMiner 1.0: A New Bioinformatic Workflow and Application for Phylogenetic Marker Development Using Angiosperm Transcriptomes, *Applications in Plant Sciences*, 3(4):1400115. 2015
7. Zhang, X., Rosen, B. D., Tang, H., **Krishnakumar, V.**, Town, C. D. (2015) Polyribosomal RNA-seq reveals the decreased complexity and diversity of the Arabidopsis translatome, *PLoS One*, 10(2):e0117699
8. **Krishnakumar, V.**, Choi, Y., Beck, E., Wu, Q., Luo, A., Sylvester, A., Jackson, D., Chan, A. P. (2015) A Maize Database Resource that Captures Tissue-Specific and Subcellular-Localized Gene

Expression, via Fluorescent Tags and Confocal Imaging (Maize Cell Genomics Database), *Plant & Cell Physiology*, 56(1): e12

9. **Krishnakumar, V.**, Kim, M., Rosen, B. D., Karamycheva, S., Bidwell, S. L., Tang, H., Town, C. D. (2015) MTGD: the Medicago truncatula Genome Database, *Plant & Cell Physiology*, 56(1): e(1)
10. **Krishnakumar, V.**, Hanlon, M. R., Contrino, S., ..., Micklem, G., Vaughn, M. W., Town, C. D. (2015) Araport: the Arabidopsis Information Portal, *Nucleic Acids Res*, 43(D1), D1003-D1009
11. Tang, H., **Krishnakumar, V.**, Bidwell, S. L., ..., Schwartz, D. C., Town, C. D. (2014) An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*, *BMC Genomics*, 15(1), 312
12. Parkin, I. A. P., Koh, C., Tang, H., Robinson, S. J., Kagale, S., Clarke, W. E., Town, C. D., Chalhoub, B., Nixon, J., **Krishnakumar, V.**, ..., Paterson, A. H., King, G., Bancroft, I., Sharpe, A. G. (2014) Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*, *Genome Biology*, 15(6):R77
13. Young, N. D., Debellé, F., Oldroyd, G. E. D., Guerts, R., Cannon, S. B., Udvardi, M. K., Benedito, V. A., Mayer, K. F. X., Gouzy, J., Schoof, H., de Peer, Y. V., Proost, S., Cook, D. R., Meyers, B. C., Spannagl, M., Cheung, F., de Mita, S., **Krishnakumar, V.**, ..., Town, C. D., Roe, B. A. (2011) The *Medicago* genome provides insight into the evolution of rhizobial symbioses, *Nature*, 480, 520–524
14. Revanna, K. V., **Krishnakumar, V.**, Dong, Q. (2009) A web-based software system for dynamic gene cluster comparison across multiple genomes, *Bioinformatics*, 25 (7), 956-957

Conference Papers

1. Hanlon, M. R., Vaughn, M., Mock, S., Dooley, R., Moreira, W., Stubbs, J., Town, C., Miller, J., **Krishnakumar, V.**, Ferlanti, E. and Pence, E. (2014) The Arabidopsis Information Portal: an application platform for data discovery. In *Proceedings of the 9th Gateway Computing Environments Workshop* (GCE '14). IEEE Press, 38-41

PRESENTATIONS

Conference talks

SpeakerDeck: <https://speakerdeck.com/vivekkrish>

1. Presenter in the *New and Updated Bioinformatics Datasets, Tools and Resources* workshop at the 28th International Conference on Arabidopsis Research (2017), St. Louis, MO.
Talk title: “What’s new at Araport”
2. Presenter in the 2017 InterMine Developers Workshop and Hackathon at the Joint Genome Institute (JGI), Walnut Creek, CA. **Talk title:** “JBrowse and Inter-’Mine’ Communication”
3. Invited presenter in the *Teaching Genetics, Genomics, Biotechnology, and Bioinformatics* workshop at the 2016 International Plant and Animal Genome XXIII Conference, San Diego, CA.
Talk title: “Teaching Bioinformatics Data Analysis Using Cloud Computing Technology”
4. Invited presenter in the *Tripal Database Network and Initiatives* workshop at the 2015 International Plant and Animal Genome XXIII Conference, San Diego, CA.
Talk title: “Tripal within the Arabidopsis Information Portal”
5. Presenter in the *GMOD Workshop* at the 2015 International Plant and Animal Genome XXIII Conference, San Diego, CA. **Talk title:** “JBrowse within the Arabidopsis Information Portal”

Training workshops

1. Invited instructor in the 2015 *Summer Workshop in Bioinformatics* held at Oklahoma State University, Stillwater, OK
2. Instructor in the 2014 *Araport Developer Workshop* held at the Texas Advanced Computing Center, University of Texas, Austin, TX

3. Invited instructor in the 2014 *Bioinformatics of Entangled Genomes Workshop* held at Oklahoma State University, Stillwater, OK

POSTERS

1. **Krishnakumar, V.**, Chan, A., ... Town, C.D. (2017) Araport: A Community Platform for Data Sharing, Discovery and Integration. *International Conference on Arabidopsis Research*, #310
2. Cheng, C.Y., **Krishnakumar, V.**, .. Town, C.D. (2015) A refined annotation release (Araport11) for *Arabidopsis thaliana*. *ASPB Mid-Atlantic Section*.
3. **Krishnakumar, V.**, Cheng, C.Y., ... Town, C.D. (2015) Araport: Data Integration for the Arabidopsis Research Community. *16th Annual Plant Biology Minisymposium*.
4. Carlson, C., Tang, H., **Krishnakumar, V.**, ... Smart, L. (2015) Assembly and Functional Annotation of the *Salix purpurea* L. Chloroplast Genome. *Plant and Animal Genome XXIII Conference*. P0003
5. **Krishnakumar, V.**, Tang, H., Bidwell, S. L., Rosen, B. D., Zhou, S., Schwartz, D., Chan, A. & Town, C. D. (2014) An improved genome release (Mt4.0) for the model legume *Medicago truncatula*. *Plant and Animal Genome XXII Conference*, P347
6. Wu, Q., Luo, A., Zadrozny, T., **Krishnakumar, V.**, Chan, A. P, Sylvester, A. W., Jackson, D. P. (2013) Maize Cell Genomics: Developing a two component transactivation system, *Maize Genetics Conference Abstracts*, 55, P200
7. **Krishnakumar, V.**, Tang, H., Bidwell, S. L., Rosen, B. D., Zhou, S., Schwartz, D., Chan, A. & Town, C. D. (2013) *Medicago truncatula* Genome Release v4. *Plant and Animal Genome XXI Conference*, P314
8. Sylvester, A. W., Luo, A., DeBlasio, S., Zadrozny, T., Yang, F., Goldschmidt, A., Rasmussen, C. G., **Krishnakumar, V.**, Chan, A. P., Jackson, D. P. (2012) Maize Cell Genomics: Developing a two component transactivation system, *Maize Genetics Conference Abstracts*, 54, P128
9. Tang, H., **Krishnakumar, V.**, Bidwell, S., Town, C. D. (2012) Curating the *Medicago truncatula* Genome. *Plant and Animal Genome XX Conference*, P0410
10. Revanna, K. V., **Krishnakumar, V.**, and Dong, Q. (2009) Comparative Gene Cluster Viewer. *RECOMB Conference*.

PROFESSIONAL ACTIVITIES

1. Invited Peer Reviewer for scientific journals (2014 - Present, [Publons Reviewer Record](#)): BMC Genomics, Plant Methods, Nature Horticulture Research, Oxford Journals (Plant and Cell Physiology, Nucleic Acids Research), PLoS ONE, The Scientific World Journal, Functional and Integrative Genomics, Physiological Genomics.
2. Participant in the 2017 InterMine Developers Workshop and Hackathon conducted at the Joint Genome Institute (JGI), Walnut Creek, CA
3. Participant in the 2014 WebApollo Developers Hackathon conducted at Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA