

## Saket Choudhary

---

CONTACT INFORMATION	1076W 30 <sup>th</sup> Street Apartment #13 Los Angeles, California 90007	<i>Email:</i> skchoudh@usc.edu <i>Homepage:</i> <a href="http://saket-choudhary.me">http://saket-choudhary.me</a> Google Scholar                      Github
EDUCATION	<b>University Of Southern California[USC]</b> , Los Angeles, USA <i>PhD Student</i> , Computational Biology and Bioinformatics [2014–] Advisor: Prof. Anton Valouev Project: Tools for Motif conservation Analysis <a href="http://moca.usc.edu">http://moca.usc.edu</a> Poster            Abstract  <b>Indian Institute of Technology Bombay[IITB]</b> , Mumbai India <i>Bachelor of Technology, Master of Technology</i> , Chemical Engineering [2009 – 2014] Dissertation Topic: Pattern Recognition in Clinical Data <b>GPA:</b> 8.47/10	
HONORS AND AWARDS	<ul style="list-style-type: none"><li>■ Provost Fellowship[USC] awarded to outstanding incoming PhD Students [2014]</li><li>■ Gandhian Young Technological Innovation Award for designing a low cost spectrophotometer for testing water impurity by Indian Institute of Management Ahmadabad [2013]</li><li>■ Institute Organizational Color[IITB] for excellence in Organizational activities as a Web Manager for the UG Academic Council [2013]</li><li>■ Institute Technical Special Mention[IITB] for contributing actively to technical activities [2012]</li><li>■ Undergraduate Research Award[IITB] for developing 'Scilab on Cloud' [2012]</li><li>■ Institute Technical Special Mention[IITB] for improving student participation in technical activities, while working as a core team member of the Student Technical Body [2011]</li><li>■ Institute Technical Special Mention[IITB] for contributing to Institute's Technical activities as a freshman [2010]</li><li>■ Kishor Vagnyanik Protsahan Yojana(KVPY) Fellowship by Indian Institute of Science, awarded to top 125 students in India to pursue Science [2007]</li><li>■ Homi Bhabha Young Scientists' Gold Medal by Bombay's Science Teacher Association [2005]</li></ul>	
OLYMPIADS	<ul style="list-style-type: none"><li>■ Among <b>Top 6</b> and <b>Top 30</b> to be selected for <b>Indian National Mathematics Olympiad(INMO)</b>, selection level exam for International Mathematical Olympiad(IMO) [2008,09]</li><li>■ <b>Top 250</b> in <b>Indian National Physics Olympiad (INPhO)</b> [2009]</li><li>■ <b>Top 300</b> in <b>Indian National Astronomy Olympiad (INAO)</b> [2009]</li></ul>	

## PUBLICATIONS

Syed, Parvez, Shabarni Gupta, **Saket Choudhary**, Narendra Goud Pandala, Apurva Atak, Annie Richharia, Heng Zhu et al. *Autoantibody Profiling of Glioma Serum Samples to Identify Biomarkers Using Human Proteome Arrays* Scientific reports 5 (2015). [Link](#)

Yachdav, Guy, Tatyana Goldberg, Sebastian Wilzbach, David Dao, Iris Shih, **Saket Choudhary**, Steve Crouch et al. *Anatomy of BioJS, an open source community for the life sciences*. eLife 4 (2015): e07009. [Link](#)

**Choudhary, Saket**, and Santosh B. Noronha. *GalDrive: Pipeline for comparative identification of driver mutations using the Galaxy framework*. bioRxiv (2014): 010538. [Link](#)

**Choudhary, Saket**, Vishnu Raj, K. Sanmugasundaram, Gyan Singh Patel, and Kannan Moudgalya. *Scilab on Cloud and Textbook Companion Project: A Web 2.0 Service for Open Source Education*. In 2013 International Conference on Cloud Computing and Big Data. [Link](#)

Gatkine, Pradip, Swati Gatkine, Sushanth Poojary, **Saket Choudhary**, and Santosh Noronha. *Development of piezo-electric sensor based noninvasive low cost Arterial Pulse Analyzer*. In Biomedical Engineering International Conference (BMEiCON), 2013 6th, pp. 1-4. IEEE, 2013. [Link](#)

Dilip Save, Yogesh, R. Rakhi, N. D. Shambhulingayya, Amit Srivastava, Manas Ranjan Das, **Saket Choudhary**, and Kannan M. Moudgalya. *Oscad: An open source EDA tool for circuit design, simulation, analysis and PCB design*. In Electronics, Circuits, and Systems (ICECS), 2013 IEEE 20th International Conference on, pp. 851-854. IEEE, 2013. [Link](#)

## RESEARCH EXPERIENCE

### **Tools for Motif Conservation Analysis, PhD Project**

May, 2015 - Ongoing

Guide: Prof. Anton Valouev

Dept. of Preventive Medicine, Keck School of Medicine, USC

Motif analysis of ChIP-Seq datasets often reports multiple motifs. However, determining the quality of a reported motif is hard. Motifs predicted by motif discovery tools can often not be the 'true motifs' and can have significant p-value(or E-values) for even 'false motifs'. We hypothesized that, a 'true motif' should exhibit high evolutionary conservation scores. MoCA makes use of the PhyloP and Gerp scores to assess the conservation profile of the motif bases and compares it with flanking bases and by searching for motifs in random genomic regions. We performed analysis on various ENCODE Chip-Seq datasets and found that the 'true motifs', validated experimentally do exhibit high conservation scores.

MoCA is available as a web service at <http://moca.usc.edu>

### **Patter Recognition in Clinical Data, Masters Thesis**

April, 2013 - July, 2013

Guide: Prof. Santosh Noronha

Dept. of Chemical Engineering, IIT Bombay

### ***Awarded Outstanding Thesis Award***

Cancer is a disease known to be affected by mutations. These mutations however may not all be significant. Distinguishing driver mutations from passengers is a non-trivial problem. We are trying to explore integrative approaches in significant mutation discovery by integrating analysis at multiple levels. The current methods are either focused on functional pathway analysis or machine learning. We developed a Galaxy based toolbox to run multiple such prediction tools at once, thus removing the need to convert data formats, in a reproducible manner. The end results were displayed as a heatmap, to give an insight into those mutations which are predicted to be drivers, by all the tools, thus possibly reducing the set of mutations to the 'real' drivers.

We use Correspondence Analysis and recursive feature elimination to predict a smaller set of marker genes that can be used to differentiate marker genes of Glioblastoma.

The Galaxy tools are hosted on Testtoolshed [here](#)

[Publication] Autoantibody Profiling of Glioma Serum Samples to Identify Biomarkers Using Human Proteome Arrays

[Thesis] Pattern Recognition in Clinical Data

[Defense Presentation]

[Preprint] GalDrive: Pipeline for comparative identification of driver mutations using the Galaxy framework

### **Automated Mining of Reaction Patterns**

May 2012 - Jul 2012

*Guide: Dr. Syed Asad Rahman*

*Janet Thornton Lab, EMBL-EBI, Cambridge(UK)*

EC-BLAST is a novel tool to compare enzymes and map reactions. I used Machine Learning (clustering) based approaches to point out misclassified enzymes in the established classification system. I also developed a RESTful web-service to allow automated job submissions that would allow the users to keep track of all submitted jobs and retrieve results on demand.

### **Next Generation Sequencing, Supervised Learning Project**

Jul,2012-Dec,2012

*Guide: Prof. Santosh Noronha*

*Dept. of Chemical Engineering, IIT Bombay*

This project was in collaboration with Advanced Centre for Treatment, Research and Education in Cancer (ACTREC). Surveyed literature on Next Generation Sequencing techniques and developed automated pipelines using Python to analyze whole genome data of cancer tumors. As part of the project I contributed open source modules for BWA and samtools for Biopython, a Python based open source library for bioinformatics.

### **Scilab On Cloud**

May 2012-Jul 2012

*Guide: Prof. Kannan Moudagalya*

*Dept. of Chemical Engineering, IIT Bombay*

Scilab is an open source software for numerical computation and is primarily command line/GUI based. We developed a back end that allowed running Scilab through browser much like the modern day IPython notebooks.

Presented at IEEE Conference Cloud Computing and Big Data (CloudCom-Asia), 2013

## **PROFESSIONAL EXPERIENCE**

**Google Summer of Code 2015 | Mixed Effect Models for *statsmodels*** May 2015-Jul 2015  
Student Contract Developer

- Implemented new IPython based notebooks using Mixed Effects Models
- Implemented Likelihood ratio tests, performance improvements

**SlideShare**  
Software Engineering Intern

May 2011-Jul 2011  
New Delhi, India

- SlideShare <http://www.slideshare.net> is an online platform to view and upload presentations
- Deployed a Ruby on Rails module to allow administrators to delete/suspend defunct users and slide-shows using filters
- Tool is being used currently at SlideShare internally for management of users and slide-shows

**Google Summer of Code | BioJavascript**  
Student Contract Developer

Jul, 2014-Sep, 2014

- Worked with BioJavascript, an open source library of javascript components to represent biological data
- Developed Human Genetic Variation Viewer, a d3.js based component to visualize genetic variations in humans
- Demo: <http://saketkc.github.io/biojs>
- Screencast: [http://youtu.be/jd6S\\_xnCGwU](http://youtu.be/jd6S_xnCGwU)

**Google Summer of Code-Penn State University | Galaxy Project**  
Student Contract Developer

July, 2013 - Sep, 2013

- Worked with Galaxy Project <http://galaxyproject.org/>, an open source web-based platform for data intensive biomedical research
- Implemented 'nested workflows' that allows users to run a workflow inside a workflow, obviating the need of replicating steps, <https://bitbucket.org/galaxy/galaxy-central/pull-request/229/nested-workflows>
- Added 'edit on the go' functionality to edit default parameters before runtime <https://bitbucket.org/galaxy/galaxy-central/pull-request/232/editable-workflows-gsoc2013>
- Proof Of Concept: <http://galaxy-gsoc2013.blogspot.com/2013/09/and-it-comes-to-end.html>

**Google Summer of Code-Rice University | Connexions Project**  
Student Contract Developer

July, 2012 - Sep, 2012

- Developed a Python module to aid conversion of slide-shows to online published notebooks using SlideShare API
- Implemented the functionality to add user defined quiz as an additional achievement
- Demo: [http://youtu.be/jfU0uj\\_ipY8](http://youtu.be/jfU0uj_ipY8)

#### OTHER PROJECTS

**Image Analysis of Tuberculosis samples**

Jan, 2013 - Apr, 2013

*Supervised Learning Project, Collaborator: Hinduja Hospital, Mumbai*

- Implemented **image processing** algorithms for to automatically identify true positives in TB sputum images
- Developed a wxPython based **GUI** and OpenCV based **image processor**

	<b>Pratham, Student Satellite Programme</b> May, 2010 - Oct, 2010 <i>Part of First Student's Satellite Team IIT Bombay</i> <ul style="list-style-type: none"> <li>■ Pratham is an interdisciplinary program aiming to build a cubesat, to measure the Total Electron Count in the atmosphere</li> <li>■ Implemented ADC sub-module on ATMEGA32, performed <b>hardware testing</b></li> </ul>
	<b>IIT Bombay Grading System on SMS</b> Oct, 2011-Dec, 2011 <i>Course Project</i> <ul style="list-style-type: none"> <li>■ Developed a Python/Flask based app to automatically <b>scrape</b> the IITB web pages</li> <li>■ Developed an SMS based API to receive queries and send back the fetched grades from the IITB servers, received <b>500+</b> SMS requests</li> <li>■ Deployed a Gmail based chat bot to perform similar operations, received more than <b>1000</b> chat requests</li> </ul>
	<b>Digital Audio Player</b> IIT Bombay <i>Electronics Club Summer Project, IIT Bombay May, 2010 - June, 2010</i> <ul style="list-style-type: none"> <li>■ Implemented the <i>FAT32</i> file system reading code for reading the .wav files stored on a SD card</li> <li>■ <b>ATMEGA16</b> was used to play pre-recorded songs on SD card through <i>Pulse Width Modulation</i></li> </ul>
Teaching Experience	<b>Teaching Assistant, Computer Programming and Utilization</b> Autumn 2011 <ul style="list-style-type: none"> <li>■ Selected as an Undergraduate Teaching Assistant on the basis of past performance</li> <li>■ Mentored students for Lab Sessions and Projects assist in conduct of examination and evaluation</li> </ul>
	<b>Teaching Assistant, Artificial Intelligence in Process Engineering</b> Autumn 2011 <ul style="list-style-type: none"> <li>■ Involved in contributing to forum posts, evaluating answer scripts</li> </ul>
POSITIONS OF RESPONSIBILITY	<b>Web Manager, UG Academic Council</b> July, 2012 - April, 2013 <ul style="list-style-type: none"> <li>■ Initiated a number of web portals for improving the accessibility of academic resources to the students</li> <li>■ Implemented Online Notice Board System via Institute's Gymkhana website</li> <li>■ Awarded <b>Institute Organizational Color</b></li> </ul>
	<b>Institute Internship Coordinator, Practical Training Cell</b> July, 2011 - April, 2012 <ul style="list-style-type: none"> <li>■ Initiated Industry &amp; University Interaction for internships</li> <li>■ Number of internships registered a 40 percent growth</li> </ul>
	<b>TechniC, Core Group Member</b> July, 2010 - April, 2011 <ul style="list-style-type: none"> <li>■ Member of IIT Bombay's Technical events organizing body</li> <li>■ Organized technical events inside the institute along with nine other members</li> <li>■ Mentored students for various technical competitions</li> </ul>
STANDARDIZED TEST SCORES	<ul style="list-style-type: none"> <li>• GRE: Quantitative: 170/170 Verbal: 153/170 Analytical Writing: 3.5/6</li> <li>• TOEFL: Reading: 29/30 Listening: 28/30 Speaking: 24/30 Writing: 28/30 Total: 109/120</li> </ul>