

Arunkumar Srinivasan

arunkumar.sriniv@gmail.com • 0176 621 88 098
Koernerstrasse 11 • 50823 Cologne • GERMANY

Application for Scientist / Project Manager Bioinformatics.

Education

- | | |
|------------------|---|
| Feb'11 - present | <i>PhD student, Bioinformatics, MPI For Plant Breeding Research, MIPZ (Cologne, Germany). I will be defending this year.</i>
Thesis: “Development and application of a computational pipeline to detect differential splicing events from RNA-seq data.” |
| Oct'07 - Dec'09 | <i>M.Sc, Bioinformatics, Zentrum fuer Bioinformatik, University of Tuebingen (Germany)</i>
Thesis: “Kernel methods for Genome-Wide Association Studies (GWAS).” |
| 2001 - 2005 | <i>B. Tech, Electronics Engineering, Anna University (Tamilnadu, India)</i> |

Experience

Bioinformatics

Main PhD project:

1. Developed a computational pipeline to **detect and statistically validate** differential splicing events between two or more RNA-seq datasets.

Collaborative projects:

2. Analysis of **large RNA-seq data** and **statistical testing** of the effect of RPS4 mediated immune response on alternative splicing patterns in *Arabidopsis thaliana*.
3. Transcriptional profiling of *smg 7-1 pad 4-1* double mutant and *pad 4-1* single mutant in *Arabidopsis thaliana* to identify primary targets of NMD that are essential for immune response.
4. Differences in the distribution of alternative splicing events and potential impact of these splicing differences between dry and developing seeds in *Arabidopsis thaliana*.

Outcome:

- Experience in managing and efficiently handling analysis of very large data sets.
 - Develop user-friendly tools for reproducible analysis through code reuse.
 - Designing appropriate statistical tests to test hypotheses, carefully.
 - Communicating clearly, appreciating better ideas, and working in harmony.

I'm fortunate to have worked on these projects, because all of them together cover not *just* Bioinformatics but also Computational biology, Computer science and (Bio)statistics.

Other efforts

data.table is an R package designed to handle analysis of large (ex: NGS) data very efficiently, for faster development. I've been involved in the development since mid 2013.

Outcome:

- Experience in the process of software release and maintenance.
 - Importance of unit testing towards software development productivity.
 - Communicating with other developers over long distances clearly.
 - Exhaustive documentation, helping users with issues, prioritise bug fixes.
 - Constant learning and improving, being aware and staying ahead of competition.

Technical

R, C/C++, Perl	Co-developer and contributor of R's data.table package. Authored various <i>in-house</i> R-packages such as readCounts, gffutils, splicer to facilitate analysis of Next Generation Sequencing (NGS) data through reusable code . Developed various scripts and programs for automating pipelines for data preprocessing, filtering, analysis and visualisation using R and Perl . Worked on various part-time jobs during M.Sc coding in C and C++ .
Project Management	Worked on three different collaborative Bioinformatics projects during my PhD, amongst various <i>other smaller side projects</i> . <div>Defined objectives for projects; discussion, organisation and completion of tasks; communication of results to collaborators (mostly Geneticists / Plant biologists).</div>
Reproducibility	Git, Markdown, RMarkdown, Knitr, Pandoc, Bootstrap, Jekyll (workflow for reproducible analysis).
Other	Bioconductor, HTML, CSS, Javascript, d3.js, Bash, Makefile, Matlab, Latex, Adobe Illustrator, Microsoft Office, iWorks, Openoffice.
OS	My daily operating system is Mac OS X. But I'm equally adept with Linux and Windows.

Languages

Proficient	Tamil, English
Other	German - Cleared TestDaF qualification exam in October 2010 with level 3

Activities and Interests

Hobbies	Listening to Music, Reading, Board games.
Sports	Cricket, Badminton, Chess and occasionally running.
Activities	Social coding, Contributing on Stackoverflow.

References

Available upon request.