Jean Qin
Bioinformatics Scientist - Computational Life Science, Bayer Crop Science

Cary, NC - Email me on Indeed: indeed.com/r/Jean-Qin/185b8e97aee7e11c

WORK EXPERIENCE

Bioinformatics Scientist

Computational Life Science, Bayer Crop Science   -   Research Triangle Park, NC - January 2015 to Present

Responsibilities: provide bioinformatics support for trait research at Bayer Crop Science including to understand project requirements, provide scientific guideline, develop data analysis plan, perform large scale data analysis for target discovery, evaluate and develop methods and pipelines to handle large scale of data analysis in HPC environment, lead pest genomics and research projects, support target discovery and MoA studies.

Experience and accomplishment: since joined Bayer in 2015, I've worked closely with project stakeholders on a number of NGS projects for target discovery for every step of projects from understanding & collecting stakeholders' project requirement, develop analysis plan to perform data analysis and result presentation. I successfully led and completed several important projects and made significant contribution to target discovery for Trait Research at Bayer. I successfully led to completion of several RNA-Seq projects of pests and bacteria species and identified candidate targets and pathways for MoA studies as well as biomarkers for toxin surrogate assay. As the technical lead for the metagenomics initiative, I developed novel quantitative methods and pipelines that have been successfully used for our target gene discovery as well as gene and species diversity profiling for comparative metagenomics studies using public available metagenomics data, which led to discover the first active target for next phase studies. The patent and publication of the work is in progress. As the lead of pest genome initiatives, I provide the evaluation and recommendation of sequencing technologies and analysis strategies for whole genome DNA-Seq and whole transcriptome RNA-Seq, developed analysis pipelines and perform genome and transcriptome analysis including genome/ transcriptome assembly, gene prediction, functional annotation and validation of several key pest genomes which have been successfully used for receptor discovery and MoA studies among others. The publication of the work is in progress. I successfully developed methods and pipelines for genomics analysis, RNA-Seq expression analysis, transcriptome analysis, metagenomics as well as curation and integration of public and internal databases using a combination of unix shell scripts, perl and R packages. I document my work and share with others when necessary. I provided project updates, analysis results regularly to stakeholders as well as significant discovery to site-wide scientific meetings. I also provided consultant & training for lab scientists on bioinformatics courses including Sequencing history & RNA-Seq. Received a number of employee awards and trait research award nomination in 2016

Senior Bioinformatician

Duke Center for Genomic and Computational Biology, Duke University - January 2011 to December 2014

Responsibilities: provide bioinformatics support for the Duke Center for Genomic and Computational Biology, with focus on NGS including the project requirement & consultant, scientific guidance, data analysis plan & data analysis for research projects (both Duke and none Duke), developing methods and pipelines for large dataset processing & analysis in HPC environment, technical evaluation and implementation of NGS databases and software, providing consultation, tutorial, customer training and workshop for research communities.

Experience and accomplishment: Since joined Duke in Jan 2011, I've made significant contribution to establish new bioinformatics services to support data analysis for our research community and data management for Duke Genomics center. I worked closely with our scientists and project stakeholders to understand/collect

project requirements, develop analysis plan and execute large scale data analysis in HPC/AWS environment. I've successfully completed the 50+ projects including RNA-Seq expression analysis, ChIP-Seq protein/DNA interaction, gene regulation, motif discovery, methylation, pathway analysis, exome-seq variant analysis and whole genome association studies, transcriptome and genome sequence analysis, metagenomics analysis, protein and gene sequence and function analysis in a broad medical research area including cancer research, other diseases and drug discovery with several peer viewed papers. I developed novel methods for custom projects where there's no existing tool available, for instance, qualitative analysis of allelic specific expression & annotation, aptamer target selection analysis, metagenomics analysis, genotyping & genome wide association analysis, DNA recombination and gene conversion. I provided project presentation for each project and helped stakeholders to understand/interpret results. I also provide tutorial and training (e.g., one on one and step by step tutorial for many of researchers and colleague to design experiment & analysis plan, explore data and results using NGS tools) and workshop to NGS research communities including duke faculties, postdoc and students as well as none duke customers for their research projects (e.g., 3 day NGS analysis workshop for research communities, invited presentation of my exome-seq capture and analysis pipeline and results at Agilent sponsored seminar). I've been actively participated in collaborative projects in a broad research field from project planning to end publication, for instance, the ChIP-Seq for gene regulation and pathway analysis and RNA-Seq work for gene expression analysis in cancer research (PNAS, 2012) & (Oncotarget, 2013), exome-seq capture and variant analysis of genome editing for novel drug discovery in DMD (Molecular Therapy, 2013), ChIP-Seq, RNA-Seq & pathway analysis in arthritis (J Invest Dermatol 2015), aptmer analysis for drug discovery (Nature Chemical Biology 2016). I also helped to setup computational infrastructure, evaluate, test, deploy and develop bioinformatics tools and databases for NGS data processing and analysis in HPC environment. I developed a number of pipelines in HPC environments for most of the routine NGS analysis work with combination of open sources and in house developed tools. I also successfully developed a number of workflow and automation pipelines for large scale data parallel processing for all NGS platforms (Hi-Seq, MiSeq, 454, Ion Torrent, PacBio) which dramatically saved hands-on time and reduced potential manual mistakes for our core.

Investigator & Data Management Analyst

GlaxoSmithKline - Research Triangle Park, NC - July 2002 to December 2010

Responsibilities: support omics project stakeholders for their project planning & data analysis; develop and manage applications and pipelines for large-scale data processing & analysis; develop, support and manage internal and public domain biological databases for bioinformatics applications and research projects; working with IT and R&D groups to manage projects and resources and provide recommendation/requirements to design, plan and implement bioinformatics software and databases for R&D scientists; provide support, training, consultation and recommendation for R&D projects and scientists on bioinformatics and computational software

Experience and accomplishment: I worked closely with our scientists and project stakeholders to collect business requirement, develop analysis strategy and plan and execute data analysis for a number of projects for disease & drug target discovery including large scale sequence analysis, expression analysis, proteomics, variant analysis & genotyping, network & pathway, diseases, literature search & text mining; support complex data processing, curation, management and integration for internal and public domain biological databases for various diseases; supported tools and pipelines for large scale data handling & analysis on HPC environment; designed, implemented and supported relational databases in oracle and mysql in the area of sequence, gene, literature & network pathway; worked closely with IT and science groups, managed projects and resources (external and internal developers and scientists) to collect business requirement, develop, customize and support various omics LIMS systems for sequencing, qPCR, proteomics, microarray, genotyping and sample management; deployed, supported and managed the bioinformatics software for R&D scientists and bioinformatics projects; provided support, scientific and bioinformatics consultation & training courses for R&D

projects and scientists on data analysis and data mining using bioinformatics tools; received a number of impact awards and high annual performance rating.

Scientist

GlaxoSmithKline   -   Durham, NC - July 2001 to July 2002

Responsibilities: provide bioinformatics support for science projects; develop, deploy and manage biological data, databases and bioinformatics applications; evaluate, deploy, customize and support commercial available bioinformatics software; provide training, consultation and recommendation for bioinformatics tools

Experience and accomplishment: actively participated in scientific projects and worked closely with science groups and team members to design plans, develop methods and solutions to address their bioinformatics needs for target discovery such as high throughput large scale genomic and gene sequence data analysis, gene annotation, disease gene mining and target prioritization for neurodegenerative diseases; developed, deployed and support bioinformatics and computation tools and public domain biological databases for sequence data analysis, genome mapping, result query and visualization; for example, developed methods and automation tools for public domain biological database deployment and management, large scale sequencing data processing and analysis for high throughput biology projects and helped scientists with their sequence analysis and gene annotation effort and provided their cheery-picking ability for candidate genes selection and received high praise from their program head; developed novel genomic sequence mapping methods to help gene regulation group to prioritize their screen targets with BAC clones; provided training, consultation, recommendation, deployment & support for scientists on internal, open sources and commercial bioinformatics/computational/molecular biology software; designed, implemented and maintained relational oracle database for central data repository and information search; received an annual high performance rating.

Scientist

GlaxoSmithKline   -   Research Triangle Park, NC - January 2001 to June 2001

Responsibilities: provide support for large scale functional genomic data analysis on phytopathogenic fungal species for the mechanism of pathogen and host relationship and infection disease target discovery

Experience and accomplishment: developed tools and pipeline incorporated with existing tools for de novo fungal EST (transcriptome) and genomics sequence assembly, annotation & functional analysis; completed an uni-gene set for lab microarray chip design; supported high throughput sequence analysis effort for gene annotation and structural & function discovery; developed novel algorithm and methods for expression analysis of high throughput EST and genomic data at different conditions (infected vs normal and different developmental and environment conditions); discovered the candidate genes expressed at infectious stage; finished a draft internal report for the fungal expression profile and the candidate infectious targets; provided consultation for expression data analysis; participated in developing automation pipeline for data process, data analysis and data mining

North Carolina State University - January 2000 to June 2001

Courses for certificate in Computer Science, North Carolina State University, courses include Java programming, Perl, relational database design and management courses in bioinformatics and computer science

Research Fellow

Dept of Cell Biology, Duke Univ. Medical Center - October 1998 to December 1999

Responsibility: structure and function of genes & proteins involved in cell receptor adhesion for wound-healing

Accomplishment: finished a draft paper identifying the domains of proteins and genes that involve in cell fate during would-healing stage using molecular biology/gene expression tools, bioinformatics software and protein techniques

Research Fellow
Chem. & Biochem., University of Delaware - June 1995 to October 1998

Responsibility: structural & functional analysis on a family of genes & proteins involved in adhesion in aqueous environment using molecular & protein techniques and bioinformatics tools.

Accomplishment: published several peer reviewed papers (Science, JBC, PNAS, Biochem)

EDUCATION

certificate
NC State University Statistical Summer Institute
2005

Certificate
NC State University
1999 to 2001

Ph.D. in Biochemistry
University of Delaware
1995

MS
Nanjing Agricultural University
1986

SKILLS

Bioinformatics (10+ years), data analysis (10+ years), databases (10+ years), MARKETING ANALYSIS (10+ years), training (10+ years)

ADDITIONAL INFORMATION

Technical skills/experience
Computer Skills
• Languages: perl, R, unix shell scripting, CGI, Java, JavaScript, JSP, HTML, XML, python
• Databases: Oracle, MySQL; fluent in SQL, PL/SQL; experience in design and implement relational databases in Oracle and MySQL; Certificate in Oracle 8i admin; experience in developing automation tools for database schema creation, large batch data loading, updates and validation
• OS: Unix/Linux/Window/Mac: training in Solaris 8 (Unix) Admin. Fluent with unix/linux environment and scripting, HPC, AWS
• Data, databases and data analysis: expertise in handling large biological data and data integration including both experimental data and public domain databases: data QC, processing, curation, analysis, annotation and downstream functional mining in the area of high throughput classic and next generation sequencing analysis, microarray, proteomics, genotyping, network pathway & literature.

• Tool development: expertise in developing novel methods and automation pipelines for handling large data set including data QC, processing, troubleshooting, analysis & annotation as well as developing novel methods, scripting and tools when necessary. Experience with HPC environment.

• Bioinformatics and statistics software: Broad knowledge and experience in bioinformatics software and algorithms including both classic and NGS analysis for RNA-Seq transcriptome expression analysis, ChIP-Seq, gene regulation, motif discovery, methylation, pathway analysis, smRNA analysis, whole transcriptome/ genome sequencing assembly, annotation and functional analysis, whole genome and exome variant detection, genome wide association studies, comparative genomics, metagenomics, protein & gene sequence and function analysis, microarray, proteomics and literature mining; familiar with common public databases and tools, R packages and BioPerl for large data set processing & analysis

• Application/user support: experience in customer support environment and work with IT and science group for software lifecycle support with documentation in electronic note book. Provide scientific, technical support, consultation, recommendation and training on bioinformatics tools for scientists.

• Management: experience in management of a variety of science projects and resources with external and internal developers and scientists from requirement collection, software development to application support and life cycle management

• Platforms: NGS--- illumina, iontorrent/ion proton, PacBio, Omics---microarray, proteomics, genotyping, sequencing, Sanger

Molecular biology: Solid molecular biology background & skills with peer-viewed publications (3 yrs postdoc) for routine molecular cloning, construct manipulation, DNA sequencing, DNA sequence assembly and analysis, protein expression in a variety of cell types (bacteria, baculovirus, mammalian), DNA structural & functional characterization and analysis

Protein techniques: Solid protein biochemistry background & skills for protein structural & functional characterization and analysis

Award

Impact awards (2009) for significant contribution to a data analysis project

Impact Bronze award (2008) for excellence in support for data analysis projects

Exceptional High performance awarded for 2007 for developing novel methods and automated data analysis pipeline, designing, validating, implementing relational database and delivering candidate targets for high throughput sequence analysis project, proactively working with multiple groups to provide support to science projects/customer groups

Impact Awards (2003) for excellence in support & training for bioinformatics tools

Publication

Kahsai et al., (2016) Conformationally selective RNA aptamers allosterically modulate the β2-adrenoceptor, Nature Chemical Biology […]

Zhang et al., (2015) RNA-Seq and ChIP-Seq reveals SQSTM1/p62 as a key mediator of JunB suppression of NF-κB-dependent inflammation, J Invest Dermatol, […]

Guo et al., (2013) KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation, Oncotarget, Oncotarget […]

Ousterout et al., (2013) Reading Frame Correction by Targeted Genome Editing Restores Dystrophin Expression in Cells From Duchenne Muscular Dystrophy Patients, Molecular Therapy, […]

Guo et al., (2012) Global identification of MLL2-targeted loci reveals MLL2's role in diverse signaling pathways, Proc. Natl. Acad. Sci. doi: […]

Waite, J.H. & Qin, X. (2001) Polyphosphoprotein from the Adhesive Pads of Mytilus edulis, Biochemistry, […]

Qin, XX, and Waite, J.H. (1998) A potential mediator of collagenous block copolymer gradients in mussel byssal threads, Proc. Natl. Acad. Sci. 95 (18), […]

Qin, XX, Coyne, K. J. and Waite, J. H. (1997) Tough tendons: Mussel byssus has collagen with silk-like domains, J. Biol. Chem. […]

Coyne, K. J., Qin, XX and Waite, J. H. (1997) Extensible collagen in mussel byssus: A natural block copolymer, Science, […]

Waite, J. H., Qin, XX and Coyne, K. J. (1998) The peculiar collagens of mussel byssus, Invitation review for Matrix Biology, […]

Qin, XX and Waite, J. H. (1998) A novel natural copolymer, Presented at Society for Integrative and Comparative Biology (SICB) Meeting, Boston

Qin, X. & Waite, J.H. (1995) Exotic collagen gradients in the byssus of the mussel Mytilus edulis. The Journal of Experimental Biology, […]