

Dr. Alan E. Yocca



January 23, 2024

To Dr. Rebecca Lowdon:

I am writing to express my interest in the position of Genomics Data Scientist at Bayer Crop Science. Having a PhD in Plant Biology, I am searching for a position in which I can contribute to agricultural advancements through genetic and genomic means. This job and company align with my goals.

I want to highlight my diverse experience with machine learning and large-scale genomics. I am the lead author on a manuscript in which I deployed multiple machine learning classification algorithms to classify genes as core or variable (dispensable) in pangenomes given only a reference genome. I am also involved in optimizing machine learning models to identify gene expression biomarkers to predict post harvest outcomes for different apple cultivars. This will allow growers to optimize harvest date as well as indicate which batches of fruit will last in storage and which batches must be immediately shipped, boosting farmer profits and reducing food waste.

I've been involved in the assembly or annotation of dozens of plant genomes including Strawberry, Blueberry, Teff, Apple, and Pear. My first dissertation chapter leveraged publicly available NGS data to assemble dozens of diverse *Arabidopsis thaliana* genomes to investigate patterns of nonprotein coding evolution. I have extensive experience with pangenomics (two first-author publications) which routinely involve integration and synthesis of diverse data types.

During my experience as a Postdoctoral Researcher I've gained experience with several large-scale bioinformatics projects. I am leveraging whole-genome alignments of long-read genomes across Rosaceae including seven cultivated species to identify conserved noncoding sequences. These sequences likely contain *cis*-regulatory elements critical for core and species-specific functions. My other main project is developing methods to study allele-specific expression in fully-phased genome assemblies with pangenome graphs.

Additionally I am proficient in the programming languages of Python, Perl, and R and have experience with high performance computing environments with SLURM as well as LSF workload managers and the workflow manager Nextflow. I am eager to further discuss my fit for this position.

Thank you for your consideration,

Alan E. Yocca