

December 1, 2017

BioMed Central  
Floor 6, 236 Gray's Inn Road  
London  
WC1X 8HB  
United Kingdom

Dear Editor,

Please consider our submission entitled “***Integrating Co-Expression Networks with GWAS Detects Genes Driving Elemental Accumulation in Maize Seeds***” for consideration as a Research article to *Genome Biology*. This work is not under consideration elsewhere, but has been posted to bioRxiv (DOI:10.1101/221655). All authors have approved this submission and do not have any competing interests with its content.

Co-expression networks are a powerful tool for inferring the biological function of a gene when little is known about it. With the advent of next-generation sequencing technologies, the construction and utilization of co-expression networks is now possible in many agricultural species. In this manuscript, we used co-expression network analysis to tackle the important problem of identifying causal genes from loci that have been associated with traits. Powerful genetic mapping populations in crops have been highly successful in that they have found hundreds of links between genetic markers and agronomic traits of interest. However, the causal gene or allele responsible for these associations often remains unknown due to many genes being in linkage disequilibrium (LD) with the genetic marker. To address this problem, we developed a computational approach that integrates results from genome wide association studies (GWAS) with co-expression networks in order to prioritize the loci associated with agro-economically important traits. The key motivation for our approach is that the set of causal genes related to a trait will exhibit strong co-expression with each other while the inconsequential genes in LD with the marker will exhibit random patterns of co-expression.

To test our approach, we built, characterized and compared three maize (*Zea mays*) co-expression networks. Two of these networks were built from general purpose, publically available datasets, and the third was built specifically for our phenotype of interest (and also made publically available). We performed a rigorous GWAS simulation under different conditions and show that our approach is robust to noise that is introduced through LD. Our simulations also allowed us to identify important aspects of experimental design and parameter choice that are critical for this approach to be successful. We then use our approach to annotate the output of a large scale GWAS in maize characterizing the accumulation of 17 non-organic, elements and ions in the seed (the ionome) reducing the number of likely candidate genes by several orders of magnitude. Furthermore, we performed a validation experiment using mutant lines for a gene predicted by our approach to have an ionic phenotype, which was not previously associated with elemental accumulation.

In addition to demonstrating the usefulness of this approach in the maize grain ionome, our work establishes a paradigm for using co-expression networks for other important traits. The high quality, crop co-expression networks and tools utilized here will be applicable to both existing and future maize GWAS studies. Furthermore, this software is free and open source allowing this approach to be

extended to any species where GWAS and gene expression studies are being performed. We provide the necessary tools to build, analyze and integrate co-expression networks but also establish clear limitations and discussion on scenarios where we expect this approach to excel. We believe that our approach will be of broad interest to the readers of *Genome Biology*. Any scientists who use GWAS approaches to study the genetic basis of traits will be interesting in reading about our method.

Please feel free to contact me should you have any questions about our manuscript.

Sincerely,

A handwritten signature in black ink, appearing to read 'Chad Myers', with a long horizontal flourish extending to the right.

Chad Myers  
Associate Professor  
Department of Computer Science and Engineering  
Director of Graduate Studies, Program in Bioinformatics and Computational Biology  
University of Minnesota  
cmyers@cs.umn.edu  
(612) 624-8306