

Simplifying SNP discovery in the cotton genome

April 1 2015

The term "single-nucleotide polymorphism" (SNP) refers to a single base change in DNA sequence between two individuals. SNPs are the most common type of genetic variation in plant and animal genomes and are, thus, an important resource to biologists. The ubiquity of these markers and the fact that these polymorphisms show variation at such a fine scale (i.e., at the individual level) makes them ideal markers for many applications, such as population-level genetic diversity studies and genetic mapping in plants.

The growing popularity of next-generation sequencing has made SNPs a pervasive genetic marker in many areas of plant biology. The everincreasing throughput of sequencing platforms has resulted in the ability to easily identify and genotype thousands of SNPs across numerous individuals to uncover genetic variation among and within populations. This technique, however, becomes quite challenging when the species of interest has undergone whole genome duplication events (i.e., polyploidy), as is common in many plant lineages.

Researchers at Texas A&M and the Southern Plains Agricultural Research Center have developed a strategy that simplifies the discovery of useful SNPs within the complex genome of cotton. The protocol is freely available in a recent issue of <u>Applications in Plant Sciences</u>.

"Cotton presents a challenge for SNP marker discovery due to the polyploid origin of the two most widely grown species," says Dr. Alan Pepper, an author of the study. "All plants have duplicated sequences,



whether due to whole genome duplication, duplication of segments of chromosomes, duplication by retroviruses, or duplication by unequal crossing over. When you are looking for potential SNPs, particularly without a reference genome, you run the risk of identifying sequence differences between duplicated sequences rather than differences between individuals. This problem is particularly acute in recent allopolyploids."

Allopolyploid species are the product of hybridization between two divergent taxa. The genomes of these plants, therefore, contain two very similar copies of their genes—one from each parent.

According to Pepper, "A problem arises when our computational methods accidentally align DNA regions that are duplicated within the genomes of the plants being studied, rather than mapping the orthologous regions between the plants."

Enter the strategy presented by Pepper and colleagues.

Using the Illumina <u>next-generation sequencing</u> platform, over 50 million DNA reads were collected from restriction enzyme-digested DNA from four *Gossypium* species. The team then filtered these reads to enrich for orthologous DNA fragments.

Pepper explains, "One of the exciting things about this approach is that it employs a widely used, well-supported, off-the-shelf bioinformatics software known as Stacks (written by Julian Catchen at the University of Oregon) as a "filter" to enrich for pairs of fragments that are likely to be alleles of a single, orthologous region, rather than paralogs or homeologs."

The new method allows for the detection of polymorphisms between <u>individuals</u>, which will be useful for downstream applications such as



marker-assisted selection, linkage and QTL mapping, and <u>genetic</u> <u>diversity</u> studies.

Pepper concludes, "The overall strategy for genotyping-by-sequencing, marker discovery, and annotation that we have provided in this study will be useful for researchers working with the many economically important allotetraploid species (such as the crop brassicas), but can be extended to any <u>species</u>, including those that do not currently have a reference genome."

More information: Carlo Jo Logan-Young, John Z. Yu, Surender K. Verma, Richard G. Percy, and Alan E. Pepper. 2015. SNP discovery in complex allotetraploid genomes (Gossypium spp., Malvaceae) using genotyping by sequencing. *Applications in Plant Sciences* 3(3): 1400077. DOI: 10.3732/apps.1400077

Provided by Botanical Society of America

Citation: Simplifying SNP discovery in the cotton genome (2015, April 1) retrieved 3 May 2024 from <u>https://phys.org/news/2015-04-snp-discovery-cotton-genome.html</u>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.