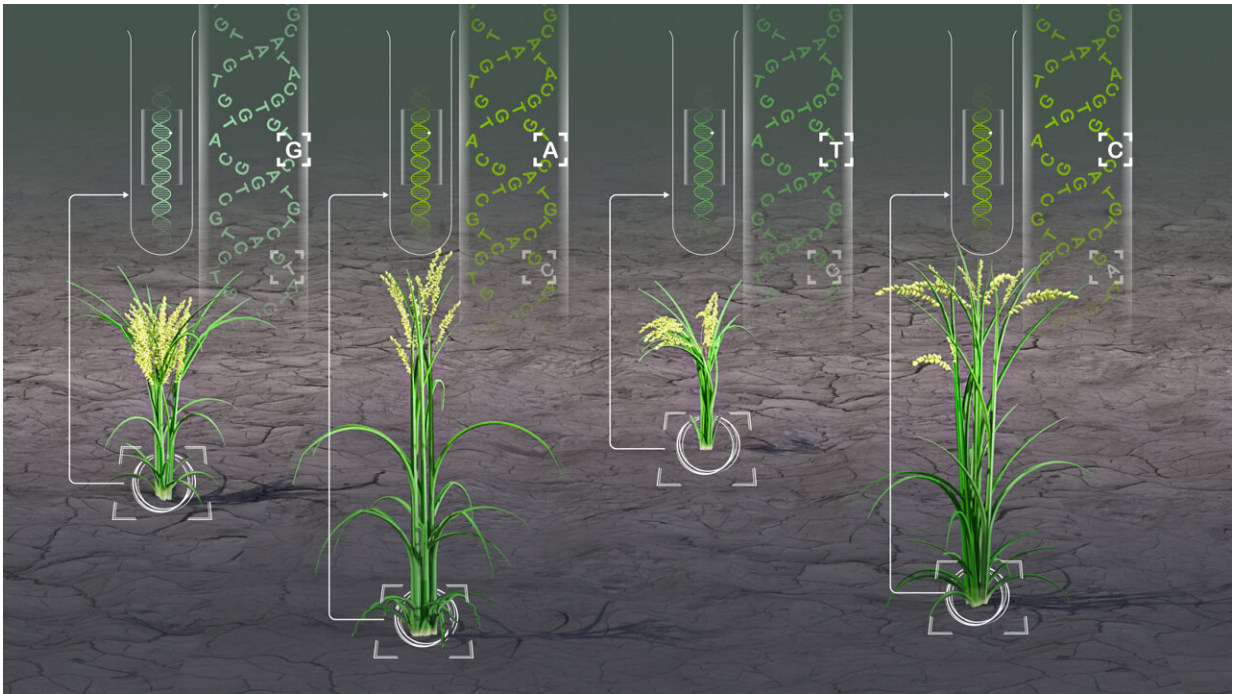


Analytic tool rapidly reveals genetic diversity for next-generation crop breeding

March 20 2024



HPC-GVCW is a computational tool that quickly identifies genetic variations in several plant species. This helps to improve crops like rice, maize, soybean and sorghum. Credit: 2024 KAUST; Heno Hwang.

In a major advance for agricultural science, researchers have developed a new computational tool designed to swiftly and efficiently expose genetic diversity within DNA databases of various plant species.

The open-source platform is poised to accelerate the discovery of genetic variations that are key to developing crops with improved resilience, yield, and nutritional value.

Harnessing advanced algorithms and the capabilities of high-performance computing (HPC), the KAUST team, led by plant genomicist Rod Wing, demonstrated the tool's ability to detect small DNA differences—so-called single nucleotide variants (SNPs)—across various strains of [rice](#), maize, soybean, and sorghum.

In the case of the rice investigation, for instance, the team employed the tool on a complex genetic dataset of DNA sequences from thousands of distinct accessions—a comprehensive "pan-genome" that the researchers had previously helped to assemble for Asian rice (*Oryza sativa*).

Using this dataset along with the group's novel analytical method, the KAUST researchers uncovered more than 2 million genetic variants previously overlooked by conventional interrogations of a single reference rice genome.

This marks an initial step towards unlocking new avenues in crop enhancement and [sustainable agriculture](#), notes plant geneticist and study co-author Yong Zhou. "These hidden SNPs could now be utilized for breeding programs immediately and also to identify novel functional genes for agricultural traits," he says.

The discovery of SNPs in this manner can also help to reveal genetic and evolutionary connections among different rice lineages. Recently, Wing and Zhou spearheaded the creation of a high-quality reference genome for Hassawi red rice, a crop indigenous to Saudi Arabia known for its resilience to local drought and high-salinity conditions.

Using the tool, the researchers were able to establish a genetic link

between Hassawi rice and a subgroup of rice that includes varieties originating from Australia, India, and parts of Southeast Asia.

Key to the performance of the tool—named the high-performance computing genome variant calling workflow, or HPC-GVCW—is the ability to divide large chunks of the genome into discrete bits and then rely on parallel processing technologies to solve complex computing problems on large-scale multidimensional genomics data.

"This reduces the execution time massively," says study co-author Nagarajan Kathiresan, a computational scientist, "making it able to process 3,000 genomes within 24 hours."

With more genomes now getting sequenced than ever before, Zhou adds, the new tool should prove invaluable for streamlining their analysis to empower next-generation crop breeding.

The work is [published](#) in the journal *BMC Biology*.

More information: Yong Zhou et al, A high-performance computational workflow to accelerate GATK SNP detection across a 25-genome dataset, *BMC Biology* (2024). [DOI: 10.1186/s12915-024-01820-5](#)

Provided by King Abdullah University of Science and Technology

Citation: Analytic tool rapidly reveals genetic diversity for next-generation crop breeding (2024, March 20) retrieved 27 April 2024 from <https://phys.org/news/2024-03-analytic-tool-rapidly-reveals-genetic.html>

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.