Leibniz Institute of Plant Genetics and Crop Plant Research (IPK)
Foundation under Public Law

AI & big data: IPK research team improves predictions for ‘tailor-made’ wheat

Press Release

Scientific Contact
Prof. Dr. Jochen Reif
Phone: +49 39482 5840
reif@ipk-gatersleben.de

Media Contact
Christian Schafmeister
Phone: +49 39482 5461
schafmeister@ipk-gatersleben.de

Gatersleben, 09.02.2026 **Climate change and evolving growing conditions present new challenges for breeding. It is important to take local environmental conditions into account. An international team led by the IPK Leibniz Institute of Plant Genetics and Crop Plant Research has used AI and big data to develop a method of determining which winter wheat varieties are best suited to specific locations. The study’s results have been published in the journal Genome Biology.**

The interaction between genotype and environmental conditions is crucial for a plant’s performance and yield. For instance, a wheat variety may produce a high yield in one location but perform poorly in another with distinct environmental conditions. Therefore, the environment affects the performance of the genotype. Given the increasing diversification of cultivation environments, it is crucial, in the context of climate change, to provide varieties tailored to specific local conditions. The research team, therefore, focused on modelling the interactions between genotype and environment as precisely as possible. This is essential for accurately predicting yields in specific locations.

First, the scientists analysed large amounts of data on winter wheat. Grain yield data from over 13,200 genotypes (lines and hybrids) grown and tested at 31 locations in Central Europe between 2010 and 2022 were collected for this purpose. This phenotypic data was then combined with genomic data (approximately 10,000 genetic markers) and environmental information, such as daily temperature and precipitation. The researchers built and compared different prediction models, including statistical and deep learning approaches. They used the best model to forecast wheat line performance across 117 environments and to identify varieties suited to specific conditions.

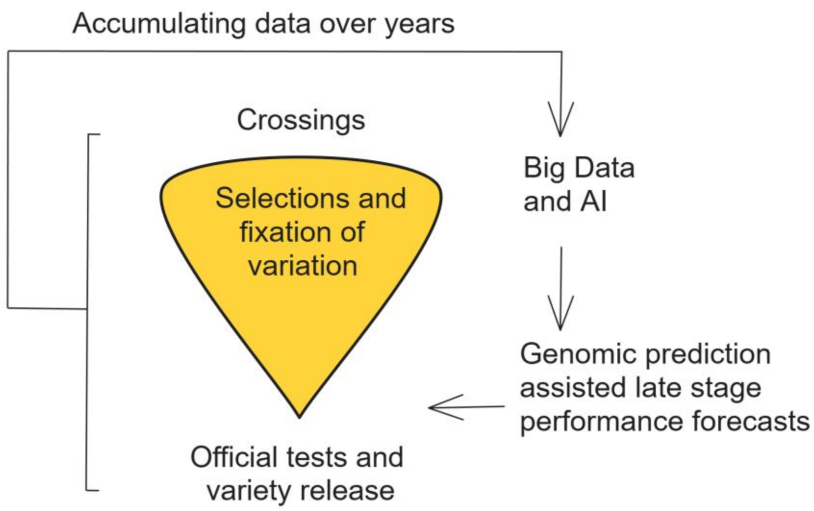
“Our study shows that interactions between genes and environmental conditions are key to significantly improving yield forecasts,” explains Abhishek Gogna, the study’s first author. By considering how genotype and environment interact, the research team was able to predict how new hybrids would perform in specific environments with greater accuracy, achieving improvements as high as 23 percent. This can be compared to buying a new suit: instead of a standard size that fits most people on average (traditional prediction), you get a suit that is tailored exactly to your body shape (environmentally adapted prediction).

Furthermore, by selecting the top ten per cent of environmentally adapted genotypes for each specific environment, an additional yield increase of almost four quintals per hectare was achieved compared to winter wheat varieties selected solely for average performance. “This additional yield is equivalent to the success of up to twelve years of conventional breeding progress in Germany,” says Prof. Dr. Jochen Reif, head of the ‘Breeding Research’ department at the IPK. “This demonstrates the enormous, previously hidden yield potential of breeding programmes.” The study’s practical relevance is also emphasised by KWS SAAT SE & Co. KGaA’s participation.

Original publication:

Gogna *et al.* (2026): Predicting enviromically adapted varieties for refining candidate selection in advanced breeding stages. Genome Biology. DOI: 10.1186/s13059-025-03914-x

Graphic:



Breeding programmes begin with crosses between superior parents, followed by several generations of selection to fix favourable genetic variation and produce selection candidates. Only the most promising candidates reach the late breeding stages, where they are tested across multiple environments and evaluated in official variety trials before release.

The current study pools data accumulated over many years from these late stages into large datasets ("big data") and applies AI-based genomic prediction models to forecast late-stage performance.