

Big data in plant breeding: IPK researchers double accuracy in predicting wheat yields

Gatersleben, 11.06.2021 Continuing to ensure stable yields of cereals such as wheat in the face of rapidly changing environmental conditions and a growing world population is one of the key challenges for science and breeding. In future, it will be increasingly important to be able to predict the yields of individual varieties as accurately as possible in a given environment. An international research team led by the IPK Leibniz Institute has compiled, processed and analysed extensive data sets for this purpose. Ultimately, Big Data was able to double the prediction accuracy for yield. The results, which could herald a new era for plant breeding, have now been published in the magazine "Science Advances".

The enormous potential of Big Data has already been demonstrated in areas such as financial services and telecommunications. An international team of researchers led by the IPK Leibniz Institute has now tapped the potential of big data for the first time on a large scale for plant research. To this end, data from three projects were used to increase the predictive accuracy for yield in hybrid varieties of wheat. "We were able to draw on the largest dataset published to date, which contains information from almost a decade of wheat research and development," says Prof. Dr. Jochen Reif, Head of the Breeding Research Department at IPK.

Finally, data on more than 13,000 genotypes tested in 125,000 yield plots were analysed. For comparison: In a breeding programme, plants are tested in 20,000 yield plots every year. "It was clear to us that we would have to increase the population sizes in order to ultimately develop robust predictive models for yield," says Prof. Dr. Jochen Reif, "so in this case it was really once: 'a lot goes a long way'". The effort was worth it, he said. "We were able to double the predictive accuracy for yield in our study."

The research team used data from the two previous projects HYWHEAT (funded by the Federal Ministry of Research and Education) and Zuchtwert (funded by the Federal Ministry of Food and Agriculture) as well as from a programme of the seed producer KWS. Basically, the challenge in such studies is to prepare the information to a uniform quality level and thus enable a common analysis. "Since we were responsible for the designs of the experiments from the start, we were able to plan them in such a way that a small proportion of the same genotypes were always tested across the projects, thus enabling an integrated analysis in the first place," says Prof. Dr. Jochen Reif.

The scientist is firmly convinced that it pays off to use Big Data for plant breeding and research. "We have ultimately worked on the future of all of us", says the IPK scientist. "We have succeeded in showing the potential of Big Data for breeding yield-stable varieties in times of climate change."

According to Prof. Dr. Jochen Reif, the current model study has a significance that goes far beyond one crop type and hopefully heralds a cultural change in breeding. "We were able to show the great benefits of Big Data for plant breeding. However, the possibilities for this



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

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

are only possible through a trusting cooperation of all stakeholders to share data and master the challenges of the future together." Ultimately, this is also the entry point for the use of artificial intelligence (AI). "The successful use of AI also stands and falls in plant breeding and research with curated and comprehensive data. Our current study is an important door opener for this path."

Original publication: Zhao *et al.* (2021), Unlocking big data doubled the accuracy in predicting the grain yield in hybrid wheat. *Science Advances*.

DOI: 10.1126/sciadv.abf9106



By increasing population sizes, an international team of scientists led by IPK Leibniz-Institute was able to double the prediction accuracy for wheat yield. Photo: IPK/ Christoph Martin