

Advances in crop phenotyping and multi-environment trials

Zhe LIU, Fan ZHANG, Qin MA, Dong AN, Lin LI, Xiaodong ZHANG, Dehai ZHU, Shaoming LI (✉)

College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

Abstract Efficient evaluation of crop phenotypes is a prerequisite for breeding, cultivar adoption, genomics and phenomics study. Plant genotyping is developing rapidly through the use of high-throughput sequencing techniques, while plant phenotyping has lagged far behind and it has become the rate-limiting factor in genetics, large-scale breeding and development of new cultivars. In this paper, we consider crop phenotyping technology under three categories. The first is high-throughput phenotyping techniques in controlled environments such as greenhouses or specifically designed platforms. The second is a phenotypic strengthening test in semi-controlled environments, especially for traits that are difficult to be tested in multi-environment trials (MET), such as lodging, drought and disease resistance. The third is MET in uncontrolled environments, in which crop plants are managed according to farmer's cultural practices. Research and application of these phenotyping techniques are reviewed and methods for MET improvement proposed.

Keywords crop breeding, genotyping, phenotyping, genotype-environment interaction, cultivar regional test

1 Phenotyping is an important means of gaining insight into crop cultivars

In this paper, “crop cultivar” refers to hybrid and inbred lines used in agricultural production, as well as elite and improved lines at different stages of breeding. Genotypic characterization or genotyping refers to the process of obtaining information on molecular marker polymorphism of cultivars. Phenotypic characterization or phenotyping refers to the evaluation of agronomic traits such as yield, grain weight per year, growth habit, plant height, morphology, as well as lodging, disease, insect, and drought tolerance. Modern phenotyping also includes

phenotypic characterization of cells, tissues and organs, i.e., phenome, as well as transcriptome, proteome, metabolome, etc.^[1]. The scientific significance of crop characterization techniques lies in accurately and rapidly acquiring phenotypic and genotypic data for the discovery of intrinsic connection between various characteristics, whereas the practical significance is to develop elite cultivars for production^[2–8].

In recent years, various cognitive techniques for crop phenotyping and especially genotyping have been developed^[9,10]. With the development of high-throughput sequencing techniques and instruments^[11,12], costs of genotyping have been sharply reduced and its efficiency greatly improved^[13,14]. For example, in 2001, the human genome sequencing project was completed at the cost of 437 million \$. While the same task could be done at the cost of merely 1000 \$, within a single day^[15] with the sequencing instrument Ion Proton™ from Life Technologies. In terms of plant genome sequencing, more than 10 crops including corn and rice have had their whole genome sequencing completed. Starting from 2010, Beijing Genomics Institute and other agencies have jointly engaged in more than 1000 key projects sequencing important plant and animal species, with 104 species completed^[16]. Monsanto and DuPont Pioneer have also created the high-throughput automatic corn seed chipper system, which realized the automatic processing of genotyping from sample preparation to molecular detection, without affecting seed viability. This technology has enabled the two transcontinental seed companies to expand the scale of their breeding activities by more than five times^[17,18].

Compared to genotyping, phenotyping is closer to breeding and production practices. However, due to the influence of environmental factors and genotype-environment interaction, phenotyping is more complicated and difficult to be precisely evaluated. Additionally, phenotyping techniques have developed slowly and have been the rate-limiting factor in genetics, large-scale breeding and development of new cultivars.

Common phenotyping techniques include: high-

Received February 15, 2015; accepted March 24, 2015

Correspondence: lshaoming@sina.com

throughput phenotyping techniques in controlled environments; field experiments in semi-controlled environments; and multi-environment trials (MET) in uncontrolled environments. These three techniques will be discussed later in this paper, followed by more detailed consideration of the application of MET in the context of China.

2 Advances in crop phenotyping techniques

2.1 High-throughput phenotyping techniques in controlled environments

High-throughput phenotyping in controlled environments has developed rapidly, and in recent years, has become a focus for research. It uses machine vision technology to closely observe individual plants and to analyze growth information and phenotypic parameters. It takes mass observations for analysis with the aid of an automated assembly line and intelligent image processing system^[19–22]. Representative examples are the plant phenotyping platforms developed by CropDesign and LemnaTec, called TraitMill and Scanalyzer, respectively^[23,24]. The former was purchased by BASF in 2006 and used by the core team in charge of high-throughput phenotyping by BASF and Monsanto. The latter has developed into the main technical support to create an international and national platform of phenomics research, such as the greenhouse automatic observation system of the Australian Centre for Plant Functional Genomics, Institute National de la Recherche Agronomique (INRA) and KeyGene^[25].

The complete system includes a conveyor belt, imaging system, darkroom, transporters, watering, weighing devices and control system (Fig. 1). The core component, imaging system, includes visible light, near infrared, fluorescence and other imaging systems. Visible light imaging is used to measure plant structure, width, density and symmetry, as well as leaf length, width, area, angle, color, scab and other parameters. Near infrared imaging is used to analyze moisture distribution in plant roots and earth pillar, and to study plant transpiration and drought

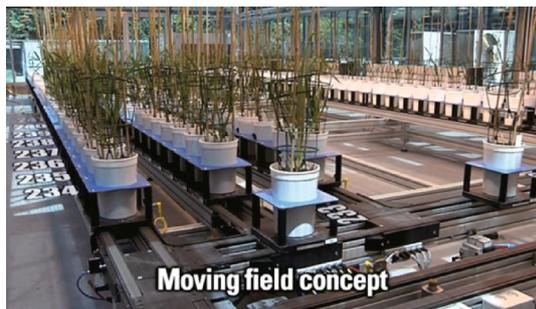


Fig. 1 LemnaTec platform of automated greenhouse plant phenomics^[24]

stress. Fluorescence imaging is used to analyze physiological status of plants. All plants are marked with barcode or radio frequency identification, their dynamic distribution on the conveyor belt is controlled by the software, and their phenotypic data regularly measured in the growth stage. The translocation by conveyor belt, transporter and plants in the greenhouse avoids the impact caused by uneven distribution of light, temperature and moisture. Imaging module, darkroom, watering and weighing devices are installed in an independent air conditioning room, and are connected to a greenhouse through via the conveyor belt. The system is a combination of greenhouse automation, high-throughput imaging technology, robot technology, image analysis and large-scale computing capacity, and is able to carry out full-automatic and high-throughput 3D imaging from seedlings to mature plants. Special software is used to analyze the imaging results, and to carry out high-throughput screening of plants. The system is applicable to research on plant functional genomics and phenomics, and is also a powerful tool for genetic breeding, mutant strain screening and phenotypic screening^[24,25]. The main problem of the system is the large upfront investment, and domestic seed companies cannot afford high construction costs afforded by several large multinational seed companies. In December 2014, the Biotechnology Research Institute, Chinese Academy of Agricultural Sciences (CAAS) established the first research platform of full-automatic high-throughput 3D imaging plant phenomics in China^[26].

Besides automatic plant phenotypic detection, with the constant breakthroughs in various imaging techniques, the phenotypic detecting technique for grains and ears has also been extensively developed. For example, Dang et al. adopted the atomic force micro-imaging technique for micro-imaging of grains, with resolution ratio of 3–7 μm ^[27], and Ogawa et al. adopted the 3D visualized technique for chipper scanning layer by layer, and also used a 3D reconstructing algorithm for reconstruction and analysis of three-dimensional structure of grains^[28]. Jayas's team from the University of Manitoba adopted a simple means of machine vision for gray processing and gray level classification of wheat sample pictures taken through the black-white area-array camera, and has developed prototype equipment and wheat was ultimately classified according to the statistical results. This team also used an X-ray imaging technique for 3D reconstruction of pea pore structure to study the micro-morphological structure of grain^[29], and to detect the internal quality of grain to study diseases and insect pests of wheat seeds^[30]. Most of these phenotyping techniques are in the laboratory stage, and still cannot be used in routine breeding.

In China, there are also some teams conducting automated, high-throughput testing phenotyping of plants, grains and ears. For example, the platform of automatic measurement and analysis of rice phenotypic parameters during growth (jointly developed by the rice phenomic

research team from the National Key Laboratory of Crop Genetic Improvement of Huazhong Agricultural University and the Center for BioMedical Photonics of Huazhong University of Science and Technology) can automatically extract 15 trait parameters, including height, leaf area, tiller number, biomass, and yield, at a rate of 1920 pots per day^[31–33].

Based on the requirements for information technology and engineering equipment for crop cultivar selection, testing, production, processing and other linked processes, the Key Laboratory for Agricultural Information Acquisition Technology of China Agricultural University has carried out technical research and software and hardware development. This team has proposed a series of phenotypic parameter acquisition technologies using visible light, infrared, X-ray and laser radar image of grains, ears and plant populations. A rapid grain detecting technique and equipment developed by this team are able to determine a seed's identity, vigor and other information within 5 min, with a degree of accuracy of above 80%, and without affecting seed viability. DuPont Pioneer, China Agricultural University and Beijing Academy of Agriculture and Forestry Sciences have respectively developed the corn grain test system (Fig. 2), which is capable of helping breeders to analyze geometrical parameters in a breeding plot such as individual plant's ear length and barren tip, as well as yield traits such as grain number and number of sterile grain, to sharply increase the efficiency of corn test, data accuracy and scale of breeding plots. Aimed at practical problems in crop breeding in China, these phenotyping techniques have been developed, in line with the actual demands.



Fig. 2 Corn ear test system developed by China Agricultural University

2.2 Phenotyping techniques in semi-controlled environments

Identification of disease resistance, and phenotypic tests for lodging, drought and poor soil resistance in partly-

controlled environments still require some subjective assessment, and are an important supplement to detect phenotypic traits, which is difficult to obtain in MET. Resistance identification is commonly carried out in the national crop cultivar regional trials. The identified traits prescribed in the technical specification for rice regional trials focus on rice blast and bacterial leaf blight, and are increased or decreased according to need in different rice regions. Resistance identification in corn production areas includes plant diseases and insect pests such as corn southern leaf blight, gray speck disease, curvularia leaf spot, *Sporisorium reilianum*, sheath blight, stem rot and *Ostrinia nubilalis*. In addition to identification of resistance to diseases such as stripe rust, wheat powdery mildew, brown leaf rust, head scab, sheath blight, it also attaches great importance to the identification of cold and drought resistance^[34–36]. For most crops, only 1–2 testing sites of resistance identification are established in a trial region. At these sites, physiological races applied by artificial inoculation, induce plant lesions, and then a quantitative or qualitative evaluation is conducted for resistance of each cultivar. The main problem arising from these tests for resistance is that physiological races are often taken from local diseased plants from the previous crop or even previous crops, and are likely to be different from disease races in the current year. In addition, due to management differences, the results can be different, and even show great variation, from those of field tests.

To improve tests for drought, desert and lodging resistance in partly-controlled environments, DuPont Pioneer has established a network of managed stress environments^[37], for example, the drought-resistant corn breeding centers established in Woodland, California, Viluco and Chile. In these centers, there is little rainfall in the corn growing season, so irrigation is used. Thus, during the growth period, researchers can accurately take control of irrigation water, which is conducive to accurate and rapid discovery of drought resistant genes and breeding drought-resistant cultivars. The Spectrum Seed Company (a company specializing in non-GM corn in Indiana, USA) found that the southern shore of Puerto Rico Island is also an ideal location for detection of stress resistance^[38]. This island is hot and dry all the year round, with the average annual minimum temperature of 21.1°C, so corn needs complete field drought management measures. Therefore, the Company has established the drip irrigation system (Fig. 3), taking precise control of water, nutrients, pesticides and sterilizing agent respectively.

In addition to simulating and controlling the change of water and nutrients to detect drought, desert, disease and insect resistance in new lines, lodging resistance is another trait needing close attention during the breeding process. Losses from corn lodging caused by wind in North America are more than USD 1 billion a year. However, lodging resistance of new cultivars cannot be effectively



Fig. 3 Puerto Rico drip irrigation and breeding system of Spectrum Seed^[19]

detected by relying only on lodging and stem thrust tests in the natural environment. DuPont Pioneer, therefore, developed a large mobile artificial wind machine (Boreas Mobile Wind Machine^[22], Fig. 4). Through simulating high-speed turbulence resulting in corn lodging, lodging resistance of plants during growth is manually tested in several common breeding sites, so that lodging resistance of cultivars has been improved significantly. This equipment is an important part of the Accelerated Yield Technology™ System^[22] of DuPont Pioneer.



Fig. 4 DuPont Pioneer's large mobile artificial wind tunnel^[22]

2.3 Phenotyping techniques in uncontrolled environments

Phenotyping in uncontrolled environments uses MET. In China, general MET is a cultivar regional test of national and provincial organizations, seed companies (hereinafter referred to as regional test). Large-scale trial and demons to MET. It is characterized by having no auxiliary measures for testing cultivars except normal field management. Ultimately, the superior cultivar is selected and the inferior ones eliminated according to phenotypic traits of each cultivar such as yield and resistance. MET is an important part of the modern commercial breeding system, and domestic and foreign scholars have carried out extensive research on the problems of MET design and management, data processing, analysis and decision-making.

2.3.1 MET design methods

Selecting testing sites is important for evaluating test sufficiency and reliability, and is also the key to trial design. As Troyer^[39] pointed out, 200 replicate tests can accurately predict the performance of maize cultivar, and the greater the number of sites, years and regions assessed, the more accurate the predictions. Starting from variation estimation of sample variance, Piepho and McCulloch^[40,41] proposed a formula for the number of tests required to give reliable results and when the coefficient of variance variation is 0.1, the number of sites is about 200. The main problem with this method is that it gives the same number of tests for all crops, trial regions and phenotypic traits, which is not conducive to optimizing the configuration of test resources.

Kempton et al.^[42] conducted extensive research on cultivars, year, location, repetition, random error and interaction effect. Troyer^[39] pointed out that in the MET in the USA, factors affecting corn cultivar were year > testing site > density > seeding time > block in the testing site > plot repetition. Zhang et al.^[43] proposed that nonlinear programming could be used to solve the number of optimal years, the number of testing sites and the number of repetitions for optimization of regional test plans, so as to achieve the optimum testing accuracy using limited funds. The cultivar comparison formula of the regional test proposed by Kong et al.^[44] could be used for calculating the required testing year, location and the number of repetitions. The location and year obtained through these two methods are quite different from the 200 tests proposed by Troyer, which may be related to the fact that plot repetition in the methods is considered as the basic unit rather than as a testing site.

Zhang et al.^[45] proposed a regional trial evaluation system covering testing evaluation and cultivar evaluation. He compared the concrete evaluation method and indicators, and pointed out unresolved difficulties, for example, a lack of an exact analytical method for layout optimization and effectiveness of testing sites as well as other indicators in the system, but both of them had a great impact on the reliability of regional tests. Liu et al.^[46,47] studied the quantitative analysis of environmental stress including corn lodging based on wind probability and southern leaf blight based on cumulative temperature and humidity according to the environmental impact factors, which can give rise to corn lodging and southern leaf blight, so as to provide methods and data support for selection of test environment and layout research.

2.3.2 MET analytical method

For a long time, researchers around the world have introduced and proposed many methods, and the national and provincial regional tests use common research methods, such as analysis of variance (ANOVA), stability

analysis^[48–50], parameter estimation of non-equilibrium data^[51,52], rank analysis^[53–58], and Additive Main Effects and Multiplicative Interaction (AMMI) models^[59,60] for the yield trait; as well as fuzzy comprehensive evaluation^[61–66], similarity and difference analytical method^[67,68], and Technique for Order Preference by Similarity to Ideal Solution (DTOPSIS) method^[69] for multi-traits analysis.

The cultivar evaluation method has developed comprehensively from inter varietal difference comparison based on ANOVA to an in-depth analysis of gene-environment interaction (GEI) as well as understanding environmental adaptability of cultivars^[70]. At present, the AMMI-type Mixed Model and BiPLOT analytical techniques have become the advanced methods for cultivar evaluation^[71]. At the same time, some scholars have studied the precision of various evaluation analytical models. For example, Zhang et al. compared prediction accuracy of various models through cross validation of data, and the results showed that precision rank of various models was Linear Regression (LR) - Principal Components Analysis (PCA) composite model > AMMI model > PCA model > mean value disposing model > regression model > ANOVA additive main effect model^[60].

Most of the above experimental and analytical methods and models are only aimed at yield per unit. And, to get the ideal analytical results, there is a high demand for mathematical statistic and analytical tools, but most cultivar testing workers have difficulty in mastering those tools. In addition, the complex analytical method is difficult to adapt to rapid, simple and easy analysis of mass test data, so that test data mining is insufficient. At present, there is no analytical method available to effectively reduce the GEI of cultivar evaluation, which is easy to understand and apply to a wide range of traits.

2.3.3 Development history of MET in the United States

The modern seed industry mostly developed in the USA, and various international seed companies there have conducted systematic research on MET technique, which contributes insight to the evolution and problems of the techniques and provides reference for MET improvement of China. An analysis of the literature^[72–79], summarized here, identifies the main characteristics of MET in the USA at each stage of its development (Table 1).

In general, its development history can be divided into three stages, namely: before 1980, when attention was paid to accurate estimation of cultivar yield, from 1980 to 2000, when attention was paid to accurate testing of cultivar stress resistance, and after 2000, when attention was paid to performance prediction of cultivars in target promotional environments (TPE) with the innovation of biological technology. With the changes in breeding targets at the different stages, breeding strategies are

constantly adjusted, and various innovative technologies have been put forward and introduced, such as visualized analysis of test results, special testing analysis software and large databases. Driven by these innovative techniques, breeding improvements maintain momentum and steady growth in the USA.

3 Status of multi-environment trials in China and suggestions for improvement

3.1 Status of multi-environment trials of crop cultivars in China

In China, before approval and widespread promotion, new cultivars must pass seed companies MET, and provincial or national regional trials. After each round of testing, yield potential and adaptability of tested cultivars is evaluated, and based on these results, superior cultivars are selected and inferior ones eliminated. New superior cultivars selected through various tests, however, are always average in performance when planted in the field, and those widely planted are usually less than 15% of the total^[80]. The performance of new cultivars in the promotion stage is inconsistent with their yielding capability and adaptability in the testing stage, showing that the reliability of testing results is on the low side. The reasons for poor testing quality are insufficient understanding of the importance of MET and inadequate investment^[81]. Kong et al. made an accurate analysis of multiple years of regional testing in China, and found that in more than 50% of the 331 tests studied there, it cannot identify yield differences of less than 10% between the cultivar tested and the check cultivar^[44,82]. In addition, with rapid development of the seed market, the value of approved new cultivars has greatly increased, and many non-technical factors have also seriously interfered with the objectivity of provincial or national MET.

At present, MET cannot meet the demands of modern breeding, as highlighted by the four considerations that follow.

3.1.1 Test design

There is a serious shortage of testing sites, and there is no quantitative standard for selection and arrangement of testing sites, leading to a high level of randomness. Like cultivation experiments, MET uses strict plot techniques (including randomized block, plot repetition and control plots) and the evaluation results are reliable in each testing site^[83,84]. However, MET uses only limited samples from the vast number of regional target promotional environments (TPE). If the number of MET samples is insufficient (i.e., inadequate testing), the general characteristics estimated on the basis of sample characteristics will be

Table 1 Development history of multi-environment trials in the United States

Stages	Before 1980	1980–2000	After 2000
Main characteristics	Accurate yield estimation	Accurate stress-resistance evaluation	Field presentation prediction
Test purposes	Select high-yield cultivars	Select high-yield and widely adaptive cultivars	Predict the performance of cultivar in the target promotional environment (TPE)
Test means	Field test, manual harvesting	Field test, little natural stress environment, manually and by machinery	Field test, natural and artificial stress environment, by machinery, molecular detection
Field testing key points	Accuracy of repetitions in each test station	Accuracy of repetitions and precision of multi environments are equally important	Precision of multi environments
Field testing systems	Primary test, several rounds of advanced tests, cultivar promotion test	Primary test, advanced test, strip test	Advanced test, strip test
Selection of testing sites	A dozen testing sites, high-yield environment	Dozens of testing sites represent the main environmental types, including low yield and stress environment	Hundreds of testing sites represent all the environmental types
Plot designs	Latin square, multiple repetition, low density, large area	Randomized block, 2–3 repetitions, medium density, medium area	Interval arrangement, no repetition, high density, small area
Error controls	Costly, abandonment of abnormal values	Medium cost, and summary of revised abnormal values	Costs less, direct summary of abnormal values
Data analysis	Variance analysis, significance test, no across-year comparison, great importance attached to the analysis in each test station	Pair comparison, <i>t</i> test and stability analysis, more attention paid to across-year, multi-environment, across-system data integration analysis, as well as to gene-environment interactions (GEI)	Comparison of multiple references, performance prediction based on across-year, multi-environment and across-system data, utilize GEI
Data management tools	Calculator	Mainframe computers, databases, statistical software	Micro-computers, databases, special software
Data presentations	Professional statistical reports	Structured reports, using special font and symbol, statistical diagram	Visualized analysis software, brief reports
Decision-making groups	Breeders	Breeders, senior management	Breeders, senior management, farmers
Presentation of cultivars	Narrow TPEs and long economic life	Medium TPEs, stress resistance, medium economic life	Wide TPEs, stress resistance, suitability for users, short economic life

unreliable, and the test design will lack the analytical methods to calculate the sample size and test sufficiency. In addition, in the cultivar promotion stage, abnormalities, such as significant crop failure, is often caused by one fatal flaw in stress resistance. Due to historical reasons, in China, most environmental conditions at the testing sites are moderate, and the probability of suffering from severe stress is low^[46], so it is difficult to obtain a reliable evaluation of cultivar stress-resistance. Meanwhile, because of the lack of methods and tools, there is no systematic research on quantitative selection and spatial arrangement of testing environment.

3.1.2 Test analysis

First, there is a lack of analytical methods which are simple and easy to use, and accurately and reliably reflect test results and cultivar genetic characteristics. Secondly, there is a lack of visualized diagram analytical technique appropriate for MET data, and most scholars still focus

on complex statistical methods and models, so that the mining, transmission and utilization of test data are relatively limited. In addition, another reason for inaccurate evaluation is treating the check cultivar as the evaluation and reference system^[85].

3.1.3 Phenotypic acquisition

Whether in field plant phenotypic measurement or later-stage indoor ear test, data acquisition for most observed phenotypic traits relies on manual measurement, which gives rise to problems such as phenotypic observation, needing considerable time and effort, and a high degree of artificial data deviation and poor traceability. With the rapid increase of labor costs and expansion of modern breeding test scale, those conflicts and pressures are increasingly problematic, so there is an urgent need to develop various high-throughput, automatic field phenotyping technique and equipment, replacing phenotypic measurement based on manual observation.

3.1.4 Data management

Software tools are used to manage and analyze MET data. Test analysis software, either developed in China or imported, has its problems. The first is the difficulty in conducting a comprehensive analysis of across-year and across-trial-region test data. The second is the large number of data quality problems in test data; furthermore, the existing software has barely any ability to validate problem data, so a lot of manual error-checking work is required. The third is that results from analyses are not user-friendly, namely, the results of a certain statistical methods are often output directly, hard to read, greatly restricting the number of researchers able to interpret these results. Lastly, the software only manages text data, so it lacks effective management and analysis of test photos, even though visualized information has become important for decision-making.

3.2 Technical essentials for optimization of MET

MET in uncontrolled environments is mainly aimed at screening new cultivars with high yield and broad adaptability, and to predict future performance. Based on this objective, the system design and data analysis include two levels. The first is testing site and the second the trial region.

When designing a field experiment in a testing site, some experimental factors usually need to be considered including fertilization, irrigation, density and sowing time. Each factor is divided into different levels of quantity value, and each factor and its corresponding level constitute the experimental treatment number. When designing a field experiment, according to actual needs, costs and other considerations, experimental treatments are completely or incompletely implemented, and the final field arrangement needs to consider the application of plot technique such as randomized block, plot repetition and control plots.

In general, the experimental design technique at the level of testing sites is mature, but the test design technique at the level of trial region lacks support by relevant technical methods and tools. For some trial regions, the basic problems related to the test design are rarely studied. For example, how many testing sites are required for the MET system, in order to guarantee accurate evaluation of tested cultivars, and how to select and arrange each testing site, in order to obtain high test efficiency? Although most experimental factors at the level of testing sites are controllable, those at the level of trial region (such as yield, lodging, drought, plant diseases and insect pests, etc.) are uncontrollable. Although the occurrence of each of these experimental factors has a certain probability range, they are intrinsically uncertain. In addition, unlike the previous designs for sampling, cultivar performance in MET is affected by environmental effects and GEI, so it is

difficult to accurately estimate cultivar genetic characteristics. As a result, special consideration should be given to confirmation number of tests in the MET design and the method of selection and arrangement of testing sites^[46,47,86].

3.3 Analysis of test results

The statistical analysis techniques for cultivar testing has evolved from earlier statistical methods for cultivation testing^[55] and requires balance of experimental data. However, the MET is aimed at screening of new cultivars, and in each round of tests, the superior cultivar or combination is selected but the inferior ones are eliminated, so that across-year or annual test results are not balanced. Therefore, the MET data analytical methods needs to be adaptable to unbalanced data. In addition, the analysis of test results also needs an effective method to separate environmental effect and GEI, and its visualized analysis should take into consideration the multi-trait and large size of MET data^[87–89].

4 Conclusions

In recent years, high-throughput phenotyping in controllable and half-controlled environments has developed rapidly, and has enriched and improved the system of crop breeding phenotyping techniques. The MET is still of great significance as an important link to check the actual field performance and market prospects of new cultivars. Breeding institutions and seed companies around the world have attached great importance to MET, for example in Monsanto, the number of testing plot increased by 1.8 times from 2003 to 2007^[17]. China has large farmland areas, rich in ecological types and abundant crop cultivars, thus the MET is especially important. In general, the MET system used in China is only comparable with level operating in the 1980s in the USA, focusing on cultivar yield and precision testing of stress resistance. In the next five to ten years, it is anticipated that different levels of government and seed enterprises will set up and improve MET system significantly.

Therefore, research on methods and tools for test design and analysis, phenotypic acquisition and management, are urgently needed so as to provide support for the establishment of a reliable crop cultivar MET system, improvement of testing efficiency and reliability of the selected cultivars and reduction of risk in selection and introduction of cultivars.

Acknowledgements This work was supported by the National Natural Science Foundation (Spatial Distribution of Multi-environment Trial Stations for Maize Cultivar, 41301075), and the National Science-technology Support Plan Projects (Research and Demonstration of North China Corn Commercialized Breeding Technique, 2014BAD01B01) and Key Laboratory of Agricultural Information Acquisition Technology, Ministry of Agriculture.

We appreciate language assistance and suggestions from Hongshuo Wang at the Ohio State University.

Compliance with ethics guidelines Zhe Liu, Fan Zhang, Qin Ma, Dong An, Lin Li, Xiaodong Zhang, Dehai Zhu and Shaoming Li declare that they have no conflict of interest or financial conflicts to disclose.

This article is a review and does not contain any studies with human or animal subjects performed by any of the authors.

References

1. Yu G H, Fang X J. Concept of phenomics and its development in plant science. *Molecular Plant Breeding*, 2009, **7**(4): 639–645
2. Thakur A K, Rath S, Patil D U, Kumar A. Effects on rice plant morphology and physiology of water and associated management practices of the system of rice intensification and their implications for crop performance. *Paddy and Water Environment*, 2011, **9**(1): 13–24
3. Pask A, Joshi A K, Manès Y, Sharma I, Chatrath R, Singh G P, Sohu V S, Mavi G S, Sakuru V S P, Kalappanavar I K, Mishra V K, Arun B, Mujahid M Y, Hussain M, Gautam N R, Barma N C D, Hakim A, Hoppitt W, Trethowan R, Reynolds M P. A wheat phenotyping network to incorporate physiological traits for climate change in South Asia. *Field Crops Research*, 2014, **168**: 156–167
4. Klassen S P, Villa J, Adamchuk V, Serraj R. Soil mapping for improved phenotyping of drought resistance in lowland rice fields. *Field Crops Research*, 2014, **167**: 112–118
5. Jha U C, Bohra A, Singh N P. Heat stress in crop plants: its nature, impacts and integrated breeding strategies to improve heat tolerance. *Plant Breeding*, 2014, **133**(6): 679–701
6. Nogales A, Muñoz-Sanhueza L, Hansen L D, Arnoldt-Schmitt B. Phenotyping carrot (*Daucus carota* L.) for yield-determining temperature response by calorimetry. *Planta*, 2015, **241**(2): 525–538
7. Junker A, Muraya M M, Weigelt-Fischer K, Arana-Ceballos F, Klukas C, Melchinger A E, Meyer RC, Riewe D, Altmann T. Optimizing experimental procedures for quantitative evaluation of crop plant performance in high throughput phenotyping systems. *Frontiers in Plant Science*, 2015, **5**: 770
8. Mutka A M, Bart R S. Image-based phenotyping of plant disease symptoms. *Frontiers in Plant Science*, 2015, **5**: 734
9. Cockram J, Jones H, Norris C, O'Sullivan D M. Evaluation of diagnostic molecular markers for DUS phenotypic assessment in the cereal crop, barley (*Hordeum vulgare* ssp. *vulgare* L.). *Theoretical and Applied Genetics*, 2012, **125**(8): 1735–1749
10. Bucksch A, Burrige J, York L M, Das A, Nord E, Weitz J S, Lynch J P. Image-based high-throughput field phenotyping of crop roots. *Plant Physiology*, 2014: 1–49
11. Seiffert U, Bollenbeck F, Mock H-P, Matros A. Clustering of crop phenotypes by means of hyperspectral signatures using artificial neural networks. *Hyperspectral Image and Signal Processing: Evolution in Remote sensing*, 2010: 1–4
12. Harbinson J, Prinzenberg A E, Kruijjer W, Aarts M G. High throughput screening with chlorophyll fluorescence imaging and its use in crop improvement. *Current Opinion in Biotechnology*, 2012, **23**(2): 221–226
13. Barabaschi D, Guerra D, Lacrima K, Laino P, Michelotti V, Urso S, Valè G, Cattivelli L. Emerging knowledge from genome sequencing of crop species. *Molecular Biotechnology*, 2012, **50**(3): 250–266
14. Cobb J N, Declerck G, Greenberg A, Clark R, McCouch S. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics*, 2013, **126**(4): 867–887
15. Life Technologies. Life Technologies introduces the Benchtop Ion Proton™ Sequencer. Available at Life Technologies website on January 10, 2012
16. BGI. 1000 plant and animal reference genomes project. Available at LDL Genomics website on April 5, 2012
17. Monsanto. Research and development pipeline. Available at Monsanto website on January 5, 2012
18. Pioneer. Technologies brochures & specification sheets. Available at Pioneer website on March 1, 2012
19. Nanni L, Lumini A. A reliable method for cell phenotype image classification. *Artificial Intelligence in Medicine*, 2008, **43**(2): 87–97
20. Zhang L, Grift T E. A LIDAR-based crop height measurement system for *Miscanthus giganteus*. *Computers and Electronics in Agriculture*, 2012, **85**: 70–76
21. Busemeyer L, Ruckelshausen A, Möller K, Melchinger A E, Alheit K V, Maurer H P, Hahn V, Weissmann E A, Reif J C, Würschum T. Precision phenotyping of biomass accumulation in triticale reveals temporal genetic patterns of regulation. *Scientific Reports*, 2013, **3**: 2442
22. Jay S, Rabatel G, Hadoux X, Moura D, Gorretta N. In-field crop row phenotyping from 3D modeling performed using Structure from Motion. *Computers and Electronics in Agriculture*, 2015, **110**: 70–77
23. CropDesign. Traitmill- platform and process. Available at CropDesign website on April 16, 2015
24. LemnaTec. Available at LemnaTec website on April 16, 2015
25. ShangHai ZeQuan Technology Limited. Fully automated high-throughput plant 3D imaging system. Available at Zealquest website on December 8, 2009
26. Biotechnology Research Institute of Chinese Academy of Agricultural Sciences (CAAS). CAAS completed China's first fully automated high-throughput 3D imaging phenomics research platform. Available at CAAS website on April 16, 2015
27. Dang J M C, Copeland L. Imaging rice grains using atomic force microscopy. *Journal of Cereal Science*, 2003, **37**(2): 165–170
28. Ogawa Y, Kuensting H, Sugiyama J, Ohtani T, Liu X Q, Kokubo M, Kudoh K, Higuchi T. Structure of a rice grain represented by a new three-dimensional visualisation technique. *Journal of Cereal Science*, 2002, **36**(1): 1–7
29. Neethirajan S, Jayas D S, White N D G, Zhang H. Investigation of 3 D geometry of bulk wheat and pea pores using X-ray computed tomography images. *Computers and Electronics in Agriculture*, 2008, **63**(2): 104–111
30. Narvankar D S, Singh C B, Jayas D S, White N D G. Assessment of soft X-ray imaging for detection of fungal infection in wheat. *Biosystems Engineering*, 2009, **103**(1): 49–56
31. Yang W N. Development of automatic and digital extraction

- instrument for rice yield-related traits. Dissertation for the Doctoral Degree. Wuhan: *Huazhong University of Science and Technology*, 2011 (in Chinese)
32. Bi K. Study on some parameter measurement system for rice phenotype. Dissertation for the Doctoral Degree. Wuhan: *Huazhong University of Science and Technology*, 2009 (in Chinese)
 33. Yang W, Guo Z, Huang C, Duan L, Chen G, Jiang N, Fang W, Feng H, Xie W, Lian X, Wang G, Luo Q, Zhang Q, Liu Q, Xiong L. Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nature Communications*, 2014, **5**: 5087
 34. Technical Regulations for Crop cultivar Regional Testing-Maize. China's Ministry of Agriculture, NY/T 1209–2006
 35. Technical Regulations for Crop cultivar Regional Testing-Rice. China's Ministry of Agriculture, NY/T 1300–2007
 36. Technical Regulations for Crop cultivar Regional Testing-Wheat. China's Ministry of Agriculture, NY/T 1301–2007
 37. Pioneer. Pioneer research to develop drought-tolerant corn hybrids. Available at Pioneer website on April 16, 2015
 38. Spectrum Seed Company (SSC). Spectrum managed stress environments. Available at SSC website on April 16, 2015
 39. Troyer A F. Phenotypic selection and evaluation of maize inbreds for adaptedness. *Plant Breeding Reviews*, 2007, **29**(28): 101–123
 40. Piepho H P. Methods for comparing the yield stability of cropping system—a review. *Journal Agronomy & Crop Science*, 1998, **180**(4): 193–213
 41. Piepho H P, McCulloch C E. Transformations in mixed models: application to risk analysis for a multi-environment trial. *Journal of Agricultural Biological & Environmental Statistics*, 2004, **9**(2): 123–137
 42. Kempton R A, Fox P N. Statistical methods for plant cultivar evaluation. London: *Chapman and Hall*, 1997
 43. Zhang Q Y, Kong F L, Yang F X. Optimizing the experimental design of regional crop trial by nonlinear program. *Scientia Agricultura Sinica*, 2000, **33**(4): 25–30 (in Chinese)
 44. Kong F L, Zhang Q Y, Ge Z N, Yang F X, Zhang D G, Liu W X. Experiment precision and cultivar comparison precision in regional crop trials. *Scientia Agricultura Sinica*, 2001, **34**(3): 266–271 (in Chinese)
 45. Zhang Q Y, Kong F L, Liao Q, Yang F X, Wang L. Evaluation system and methods for regional crop trials. *System Sciences and Comprehensive Studies in Agriculture*, 2000, **16**(2): 81–86 (in Chinese)
 46. Liu Z, Li S M, Yang J Y, Yang Y, Mi C Q, Wang H, Zhang X D, Zhu D H. Method of test environments selection for corn lodging resistance. *Transactions of the CSAE*, 2010, **26**(10): 167–171 (in Chinese)
 47. Liu Z, Guo J, Li S M, Yang Y, Zhang X D, Chen B, Zhao Z, Yang J Y, Zhu D H. Verification and quantitative expression on southern corn leaf blight. *Transactions of the CSAE*, 2011, **27**(11): 160–163 (in Chinese)
 48. Mo H D. Agricultural experiment statistics. Shanghai: *Shanghai Science and Technology Publishing House*, 1992
 49. Wang F T, Guo R L. Analysis of agricultural experiment design and statistics. Beijing: *The Rural Books Publishing House*, 1993
 50. Fasoulas A C. Rating cultivars and trials in applied plant breeding. *Euphytica*, 1983, **32**(3): 939–943
 51. Zhu J, Xu Y H, Lai M G. Analysis methods for unbalanced data from regional trial of crop cultivar: analysis for single trait. *Journal of Zhejiang Agricultural University*, 1993, **19**(1): 7–13 (in Chinese)
 52. Zhao F C, Duan D F, Tan H P, Wang G Y. Analysis of yield character and stability for unbalanced data in sweet corn regional trial. *Journal of Maize Sciences*, 2007, **15**(Z1): 142–145 (in Chinese)
 53. Jin W L, Bai Q Y. Validity of appreciation of crop yield capability with the rank analysis method by cultivar regional trials. *Chinese Journal of Oil Crop Sciences*, 2004, **26**(3): 18–23 (in Chinese)
 54. Jin W L, Bai Q Y. The analysis based on ranks of crop cultivars in regional trials. *Acta Agronomica Sinica*, 1999, **25**(5): 632–638 (in Chinese)
 55. Jin W L. The rank analysis model of evaluating crop cultivars yield stability in regional trials. *Acta Agronomica Sinica*, 2000, **26**(6): 925–930 (in Chinese)
 56. Jin W L, Bai Q Y. The rank analysis method of non-balance data in mid-long run cultivar comparative trials of rolling way. *Acta Agronomica Sinica*, 2001, **27**(6): 946–952 (in Chinese)
 57. Wang Y J. The evaluation method of new corn cultivar by using rank analysis. *Inner Mongolia Agricultural Science and Technology*, 2008, (1): 31–33 (in Chinese)
 58. Liu H Y, Zhao Y Z. Evaluation of sesame yield potential in the national trial by using rank analysis. *Chinese Journal of Oil Crop Sciences*, 2008, **30**(1): 51–55 (in Chinese)
 59. An Y W, Gao X N, Ge W D. Application of AMMI method in the analysis of spring wheat regional trial date. *Journal of Anhui Agricultural Sciences*, 2006, **34**(17): 4199–4200, 4203 (in Chinese)
 60. Zhang Q Y, Kong F L. Comparison of statistical models for regional crop trial analysis. *Scientia Agricultura Sinica*, 2002, **35**(4): 365–371 (in Chinese)
 61. Guo R L. Fuzzy sets in agriculture. Zhengzhou: *He'nan Science and Technology Press*, 1991
 62. Wang R Z, Zhao J. Application of fuzzy assessment method in the evaluation of rice cultivars. *Modern Agriculture Science and Technology*, 2007, (10): 112 (in Chinese)
 63. Guo R L. Crop grey breeding science. Beijing: *China Agriculture Science and Technology Press*, 1995
 64. Zhou C, Liu X L, Shen J H, Zhang X W. The Application of computer in the association analysis in grey system theory. *Seed*, 2000, (4): 42–44
 65. Guo R L. The test of connection trend for the similarity-difference analysis and its application in regional test of wheat cultivar. *Journal of Triticeae Crops*, 2004, **24**(1): 63–65
 66. Guo R L, Zhang J Z, Zhang A Q. Multidimensional matter element analysis method of crop cultivar. *Mathematics in Practice and Theory*, 2006, **36**(1): 115–121
 67. Zhao K Q. Set pair analysis and its preliminary application. Hangzhou: *Zhejiang Science and Technology Press*, 2000
 68. Guo R L, Yang C. L., Guan L, Hou J H, Wang K, Song Z J, He M L, Jia H Q. Study on identical and different analysis method of wheat cultivar regional test. *Journal of Triticeae Crops*, 2001, **21**(3): 60–63
 69. Zhou X R, Kong X L. The comprehensive evaluation method of

- maize regional test cultivars with the DTOPSIS. *Journal of Maize Sciences*, 2005, (Z1): 32–33,36 (in Chinese)
70. Gauch H G Jr. Statistical analysis of yield trials by AMMI and GGE. *Crop Science*, 2006, **46**(4): 1488–1500
71. Yan W K, Kang M S, Ma B L, Woods S L, Cornelius P L. GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Science*, 2007, **47**(2): 643–653
72. Bradley J P, Knittle K H, Troyer A F. Statistical methods in seed corn product selection. *Journal of Production Agriculture*, 1988, **1** (1): 34–38
73. Eberhart S A, Russell W A. Stability parameters for comparing cultivars. *Crop Science*, 1966, **6**(1): 36–40
74. Eskridge K M, Johnson B E. Expected utility maximization and selection of stable plant cultivars. *Theoretical and Applied Genetics*, 1991, **81**(6): 825–832
75. Eskridge K M, Mumm R F. Choosing plant cultivars based on the probability of outperforming a check. *Theoretical and Applied Genetics*, 1992, **84**(3–4): 494–500
76. Eskridge K M, Smith O S, Byrne P F. Comparing test cultivars using reliability functions of test-check differences from on-farm trials. *Theoretical and Applied Genetics*, 1993, **87**(1–2): 60–64
77. Gauch H G Jr, Zobel R W. Accuracy and selection success in yield trial analyses. *Theoretical and Applied Genetics*, 1989, **77**(4): 473–481
78. Gauch H G Jr. Full and reduced models for yield trials. *Theoretical and Applied Genetics*, 1990, **80**(2): 153–160
79. China Corn Blog. Available at Chinamaize Blog on April 16, 2015
80. Liu S. The debate about commercial breeding. Available at Sina Blog on February 12, 2012
81. Sun S X. Discussion on the national crop regional test problem. *Journal of Crop*, 1996, (5): 14–15 (in Chinese)
82. Zhang Q Y, Kong F L, Yang F X. Evaluation of the precision of regional crop trials in China. *Journal of China Agricultural University*, 2001, **6**(1): 43–50 (in Chinese)
83. Ming D X. Analysis of field test and statistical, 2nd ed. Beijing: Science Press, 2008
84. Nanjing Agricultural University. Analysis of field test and statistical. 2nd ed. Beijing: Agricultural Press, 1988
85. Yang G H, Tang S W, Ji G D, Sun S X. Change and development of check in maize cultivars plot trial at state-level. *Journal of Marine Science*, 2010, **18**(4): 49–51
86. Liu Z, Li S M, Zhang X D. Environmental sample size estimation based on cultivar means estimation and means comparison for multi-environment trial. *2nd International Conference on Agro-Geoinformatics*, 2013: 460–465
87. Liu Z, Yang J Y, Li S M, Wang H, Li L, Zhang X D, Zhu D H. Optimal method of transforming observables into relative values for multi-environment trials in maize. *Transactions of the CSAE*, 2011, **27**(7): 205–209 (in Chinese)
88. Liu Z, Wang H, Yang J Y, Li S M, Ma Q, Li L, Zhang X D, Zhu D H. Analytical graphics for multi-environment trials of breed selection. *Transactions of the CSAE*, 2011, **27**(10): 142–147 (in Chinese)
89. Liu Z, Li S M, Mi C Q, Wang H, Ma Q, Li L, Zhang X D, Zhu D H. A spatial-temporal diffusion study of maize cultivar in china using a new two-dimensional representation. *Journal of Food Agriculture and Environment*, 2012, **10**(1): 331–337