OPINION PAPER

Towards recommendations for metadata and data handling in plant phenotyping

Paweł Krajewski[1],* Dijun Chen[2], Hanna Ćwiek[1], Aalt D.J. van Dijk[3], Fabio Fiorani[4], Paul Kersey[5], Christian Klukas[2], Matthias Lange[2], Augustyn Markiewicz[6], Jan Peter Nap[3], Jan van Oeveren[7], Cyril Pommier[8], Uwe Scholz[2], Marco van Schriek[7], Björn Usadel[4,9] and Stephan Weise[2]

[1] Institute of Plant Genetics, Polish Academy of Sciences, ul. Strzeszyńska 34, Poznań, Poland

[2] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), OT Gatersleben, Corrensstrasse 3, D-06466 Stadt Seeland, Germany

[3] Applied Bioinformatics, Bioscience, Plant Sciences Group, Wageningen University and Research Centre (WUR), Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[4] Forschungszentrum Jülich, IBG-2 Plant Sciences, Jülich, Germany

[5] The European Molecular Biology Laboratory–The European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire CB10 1SD, UK

[6] Poznań University of Life Sciences, ul. Wojska Polskiego 28, Poznań, Poland

[7] Keygene N.V., Agro Business Park 90, 6708 PW Wageningen, The Netherlands

[8] INRA-URGI, Route de Saint Cyr, Versaille, France

[9] RWTH Aachen, Worringer Weg 3, Institute of Biology I, Aachen, Germany

* To whom correspondence should be addressed. E-mail: pkra@igr.poznan.pl
All authors contributed equally to this work.

Abstract

Recent methodological developments in plant phenotyping, as well as the growing importance of its applications in plant science and breeding, are resulting in a fast accumulation of multidimensional data. There is great potential for expediting both discovery and application if these data are made publicly available for analysis. However, collection and storage of phenotypic observations is not yet sufficiently governed by standards that would ensure interoperability among data providers and precisely link specific phenotypes and associated genomic sequence information. This lack of standards is mainly a result of a large variability of phenotyping protocols, the multitude of phenotypic traits that are measured, and the dependence of these traits on the environment. This paper discusses the current situation of standardization in the area of phenomics, points out the problems and shortages, and presents the areas that would benefit from improvement in this field. In addition, the foundations of the work that could revise the situation are proposed, and practical solutions developed by the authors are introduced.

Key words: Data formatting, data interoperability, metadata content, minimum information recommendations, phenotyping, standardization.

Introduction

Plant phenotyping, a procedure leading to understanding of structural and functional plant traits and the relationships between them, has a long history. Mostly based on visual observations and scoring systems or on simple instrumental

inspection of organisms and their parts, it was practised first by farmers and then by specialized breeders to select plants better suited to their environments, leading to increased productivity and net economic benefit. It helped to build the foundations of plant taxonomy, namely of description, identification, and classification of species by major differences in appearance, structure, or behaviour. As phenotyping has developed into a part of plant science, the need for increased throughput and precision has arisen to bridge the gap between genomics and phenomics, as well as for application in agronomy, ecology, and digital agriculture (Fiorani and Schurr, 2013). Protocols used initially for the selection of genotypes with favourable traits have now been incorporated into research programmes, and also implemented in routine official assessments such as variety registration trials. Currently, minute differences between plants belonging to different populations, characterized by different genotypes or subjected to different treatments, can also be of interest, requiring the development of increasingly precise and repeatable protocols and measurement methods.

Quantitative phenotypic traits emerge from complex interactions between heredity (genome and epigenome) and various environmental factors. This calls for observations of plants in multiple environmental contexts, if broadly applicable conclusions are to be drawn, and for application of multifactorial experimental designs in both field and greenhouse trials. Moreover, novel non-invasive methods, such as high-throughput imaging, emphasizing the dimension 'time', offer unprecedented possibilities to model dynamic processes, functional mechanisms, and growth in association with ecophysiological studies (Lynch *et al.*, 1997; Nagel *et al.*, 2012; Honsdorf *et al.*, 2014; Yang *et al.*, 2014). To disentangle interactions of genotype and environment and to interpret the growth models, a detailed characterization of the environment in which the experiments are conducted is necessary (Tardieu and Tuberosa, 2010; Poorter *et al.*, 2012). Such characterizations must be expressed in a standardized way to be of value for the community.

Recently, the term 'molecular phenotype' (Ménard *et al.*, 2013) has been coined for observations of chromatin structure features, transcripts, proteins, or metabolites. These can be seen either as interesting descriptors in their own right or as markers of associated physiological traits. Thus, the term 'phenotyping' effectively now covers any procedure of measuring plant characteristics that can be expressed quantitatively or qualitatively, at the level ranging from single cells, through whole plants (Dhondt *et al.*, 2013), to field plots and ecosystems, and with a consequently broad range of experimental designs. The protocols for different measurements performed even in a single project may require different sampling schemes and device-specific pre-processing algorithms. However, the data on all traits always have to be aggregated at a clearly defined level (of 'experimental units' or 'samples') to allow their integration with existing knowledge about the system under study. Moreover, the growing number of experiments investigating similar plant systems on a large scale calls for the development and deployment of technology that supports integration of data coming from different sources. Such integrations are feasible only if standards concerning definition and recording of the phenotypes are agreed upon and widely used.

There have been previous projects aiming at recommendations for standards for phenotypic observations. Phenotypic data models for plant research were proposed and profitably used by, for example, the DROPS (http://dropsproject.eu) and PODD (http://plantphenomics.org.au/projects/podd; Li *et al.*, 2010) projects. It is a common practice that each existing repository of phenotypic data defines the set of metadata and data format in the instructions for data submission; this was done, for example, at MaizeGDB, Triticeae Toolbox, Phenopsis DB, and Ephesis databases. Notable results concerning standards have been achieved in various '-omics' studies such as genomics (and variation), transcriptomics, proteomics, and metabolomics. However, these advances do not cover all relevant traits related to plant morphology, yield, quality, and stress resistance, nor do they describe the particular assays that generate each relevant data type. Therefore, we—two European infrastructural projects, transPLANT (Trans-national Infrastructure for Plant Genomic Science, http://transplantdb.eu) and EPPN (European Plant Phenotyping Network, http://plant-phenotyping-network.eu)—are developing new recommendations and seeking how to combine them with existing standards to address the content, meaning (semantics), and format of the old and novel phenotypic data types, and to support the integrative analysis of multiple types of phenotypic data. We aim to promote these recommendations through implementations in publicly available databases, web services, and data analysis tools. We draw on the experience of other standardization initiatives for life sciences research to obtain a good balance between pragmatism and formal correctness. Current results of our work are published at a dedicated website http://cropnet.pl/phenotypes, and have been registered at the Biosharing platform (http://biosharing.org) as bsg-000543 entry.

The present paper describes our approach and challenges that have to be addressed to obtain a broad agreement on detailed recommendations concerning phenotypic data handling. We summarize the main goals of standardization and discuss how to reach these in the context of the new data types associated with plant phenotyping. We outline the more important applications of phenotyping and discuss their specificity and its impact on standardization. We think that these sections broaden the treatment of the matter in comparison with other review papers, in particular with that of Cobb *et al.* (2013). Other issues, such as the influence of phenotyping methodologies, the need for good ontologies to achieve semantic interoperability of data commons, and the clear demand for standard(s)-aware tools and services, are outlined at the end.

Thus, this paper presents our point of view on standardized capture and description of phenotypic information. It results from numerous discussions with stakeholders both within and outside of our projects. We hope that our effort will stimulate debate leading to a broad acceptance of certain fundamental principles, and thus allow development of the practical implementations that are urgently needed.

Approach

By a 'plant phenotyping experiment' (PPE), we understand a set of experimental units (fields, plots, pots, boxes) with assigned levels of factors (classifications) that differentiate and identify their role in the investigation. Thus, the units can be assigned to different biological objects (germplasm types, genotypes), and to different parts of the experiment (blocks, sub-blocks, rows, columns), and can be treated by agents that modify the environment or that are thought to disturb or enhance the plants' responses. In the general theory of experimental design, according to terminology originating from Fisher (1947) and Cochran and Cox (1957), the assignment of experimental units to parts of the design is called 'block structure', and we use this meaning without a change. However, in the same theory, assignment of the units to the levels of all experimental factors is called 'treatment structure', with all factors being treated equally, including the biological ones. We recommend distinguishing among the factors those related to the studied 'biosources' and those related to the real 'treatments', so that our use of the general term 'treatment structure' is not fully compatible with the earlier theory. Thus, in the description of a phenotyping experiment, we recognize the descriptors of the experimental design (layout) and two types of factors: biological source and treatments. The treatments can be not only agronomic practices but also environmental challenges, such as pathogen introduction or rootstock interaction. In addition to experimental design, we introduce the term 'plant phenotyping observational study' (PPOS), by which we understand an experiment that has all the features of a PPE, except that the assignment of experimental units to treatments is beyond the control of the investigator. In such a study, the samples representing biosources are collected and their classification with respect to treatments (usually environmental variables) is recognized and taken into account in the data analysis.

By 'observation' we understand any association of a variable (trait, property, feature) with an observed value at a defined time. By 'phenotypic trait' we understand any quantitative or qualitative characteristic (external, internal, molecular—e.g. transcript level, protein level, metabolite level, morphological—obtained by imaging or measurement, yield component, resistance score, etc.) that can be observed on a biosource or a sample extracted from it. In addition to phenotypic traits, we define 'environmental variables' to characterize the environment in which the phenotypic traits are recorded. We include here environmental measurements such as temperature and soil properties (e.g. water content, composition), observed at different dates on experimental units (Hannemann *et al.*, 2009; Poorter *et al.*, 2012). By 'metadata' we understand all information that describes the observations (NISO Press, 2004). Some metadata are constant across units; they must include environmental descriptors concerning the complete experiment, in addition to the aforementioned environmental observation related to particular units.

Our aim is to develop recommendations that pertain to the content and the format of the phenotypic data. The content consists of: metadata describing the study and environment; description of biosources; description of treatments and of their levels (variants); description of the experimental design; description of sampling methods and protocols, namely collection, processing, management, and storage of samples; and observations: (i) phenotypic (trait names and types, measurement protocols, processing protocols, notation scales with the associated decision rules, units, and values) and (ii) environmental (variable names and types, measurement protocols, processing protocols, units, and values).

To provide recommendations, we follow the successful approach of similar projects in other experimental life sciences, and re-use existing standards as much as possible, with special preference given to standards that have already been successfully implemented. In our opinion, this is possible for description of biosources and treatments, less for description of sampling schemes, and only to some extent for experimental designs. The main challenge that remains is standardization of definitions and descriptions of the observations—phenotypic traits and environmental variables.

Hereby we propose recommendations to address three different aspects of standardization.

(i) To define which attributes make the content of the data set, we follow the 'minimum information' (MI) approach. It defines a 'checklist' of attributes that may be necessary to describe each experimental unit (Fig. 1), appropriate for deposition and storage in an archive or publication. Not all attributes from the checklist have to be recorded; however, the list should be used by any person (or the system) depositing or publishing data to minimize the
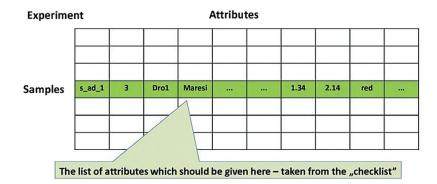

Fig. 1. Attributes that describe each sample or experimental unit.

likelihood of missing any important type of information. Numerous MI initiatives are registered at the Biosharing web portal (http://biosharing.org). The MI initiatives relevant to our work are listed in Table 1. Our checklist registered at Biosharing is called Minimum Information about Plant Phenotyping Experiment (MIAPPE); it can be found at http://cropnet.pl/phenotypes.

(ii) To ensure proper understanding of the content, we rely on its annotation with respect to publicly available ontologies and vocabularies. Numerous initiatives provide annotation tools for biologists; plant biology is a specific target of the work of Plant Ontology (http://plantontology.org) and Crop Ontology (http://cropontology.org) consortia. We intend to collaborate with the projects that develop ontologies which can be used for phenotypic experiments (see later).

(iii) To ensure proper use and interoperability of data sets prepared according to our recommendations, we work on choosing the proper format. The formats that are used in similar applications and could be applied here include CSV, XML (used, for example, by Hannemann *et al.*, 2009), RDF (http://w3.org), MAGE-TAB (Rayner *et al.*, 2006), or ISA-TAB (Sansone *et al.*, 2012). Our implementation of ISA-TAB format for phenotypic data can be found at http://cropnet.pl/phenotypes; it comprises the configuration file allowing to format data using the ISAcreator software (http://isa-tools.org/software-suite) and examples of formatted data.

The documents referred to above, published at our website, present generic solutions that can be applied, in our opinion, in a broad range of plant phenotyping experiments. They should be considered as starting points for developing more specific recommendations and formats for particular use cases; some outlook towards this is contained in the rest of this paper.

The goals of standardization

To improve biological interpretation of experimental results

A prerequisite for the improved biological interpretation of phenotypic data is to quantify the genotype–environment (GE) interactions (Annicchiarico, 2002; Tardieu and Tuberosa, 2010). Such interactions can be observed even in experiments run at different locations in—theoretically homogeneous—growth chambers due to the incomplete adoption of complex protocols across different laboratories and to difficulties in the management of environmental factors and their description (Massonet *et al.*, 2010). These issues are more acute in experiments conducted in greenhouses, characterized by an overall lower level of climate control, and obviously even more in field trials. Therefore, it is crucial to define a common set of environmental variables and describe experimental protocols to enable interpretation of plant responses to the environment (Tardieu and Tuberosa, 2010; Poorter *et al.*, 2012). We propose that a minimum set of environmental and

experiment management information constructed around the factors identified by both Hanneman *et al.* (2009) and Poorter *et al.* (2012) should be widely adopted by the plant biology community to enable interpretation and modelling of plant phenotypic responses. Following these approaches, the EPPN consortium is conducting a multilocation reference experiment measuring growth and biomass of a set of reference genotypes to demonstrate the applicability and added value of these environmental variables. These considerations can also be extended to field trials, where each trial location should also be defined in terms of climate, edaphic parameters, and management practice.

The quantification of GE interactions is also crucial for the genotype–phenotype association studies such as linkage mapping, association mapping, and breeding value estimation. Although in many studies the environment has been found to influence the location and effects of quantitative trait loci (QTLs), comparisons of QTL effects over locations is usually done without proper recognition of the importance and involvement of metadata that describe environmental conditions. Methods appropriate for studying the GE interactions (see, for example, Malosetti *et al.*, 2013) are less or not applicable if important environmental explanatory variables are not recorded. In general, the current level of standardization of QTL reports in the scientific literature, also in terms of their annotation with respect to biosources, plant organs, measurement methods, and scales, is disappointingly low. The situation improves after curation (e.g. Gramene QTL Database; Ni *et al.*, 2009). To our knowledge, the proposed standard MIQAS (Minimum Information for QTLs and Association Studies, http://miqas.sourceforge.net) has not been implemented. In our opinion, the recommendations for standardized descriptions of genotype–phenotype associations should be compatible with recommendations for genomic sequences and their variability (Yilmaz *et al.*, 2011); they should also allow for integration of data provided by breeders, and, vice versa, for transfer of the results to breeding practice.

Proper biological interpretation often requires appropriate integration of different types of phenotypic data. Currently there are very few research projects that aim at measuring traditionally observable phenotypes without also taking additional measurements made possible using new technologies. Also in applied areas, such as plant breeding, attempts are made to gather as much molecular information about genotypes as possible. The recommendations concerning various '-omics' data based on the MI approach that currently exist could be helpful for data integration, but their implementation is not fully compatible at the data level. For example, the formats of processed quantitative data aggregated at the level of assay or sample in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) and in the Metabolights database (http://www.ebi.ac.uk/metabolights) are different, and not standardized or documented in a way allowing for universal conversion. We think that this could be changed if the general rules represented in MIAPPE and in our ISA-TAB implementation are used.

Table 1. *Minimum information standards relevant to plant phenotyping*

Minimum information document (initiative)	Microarrays, MIAME (MGED)	Metabolomics, CIMR (MSI)	Sequence MIxS (GSC)	Proteomics, MIAPE (PSI)
Full name (initiative), publication, document	Minimum information about a microarray experiment (Microarray Gene Expression Database Group) Brazma *et al.* (2001) hhttp://mibbi.sourceforge.net/ projects/MIAME.shtml	Core information for metabolomics reporting (Metabolomics Standards Initiative) Fiehn *et al.* (2007a) http://msi-workgroups. sourceforge.net/	Minimum information about any sequence specifications (Genomic Standards Initiative) Yilmaz *et al.* (2011) http://gensc.org/gc_wiki/index. php/MIxS	Minimum information about a proteomics experiment (Proteomics Standards Initiative) Taylor *et al.* (2007) http://www.psidev.info/groups/ miape
Relevant extensions, publication, checklists	MIAME/Plant Zimmerman *et al.* (2006) http://mibbi.sourceforge.net/ projects/MIAME-Plant.shtml	CIMR: Plant Biology Context, Fiehn *et al.* (2007b) http://msi-workgroups. sourceforge.net/bio-metadata/ reporting/pbc/doc.rtf CIMR: EnvironmentalAnalysisContext Morrison *et al.* (2007) http://msi-workgroups. sourceforge.net/bio-metadata/ reporting/env/reporting-requirements/ECWSG_ reporting_requirements_v1.rtf	MIxSPlant-associated environmental package Yilmaz *et al.* (2011) http://wiki.gensc.org/index. php?title=MIMARKS	None (only assay-specific documents)
MIBBI project http://mibbi.sourceforge.net/ projects	Checklists stored	Checklists stored	MIGS, MIMIS checklists partially stored, link to full MIMARKS list at GSC site	No own checklists, checklists of MIBBI used
Compliant data exchange format	MAGE-TAB Rayner *et al.* (2006)	ISA-TAB Rocca-Serra *et al.* (2010)	Several, depending on the database	PRIDE XML
Databases able to store compliant information	ArrayExpress (EBI)—MAGE-TAB, spreadsheet submission, online tool submission Gene Expression Omnibus (NCBI)—various formats, and tools	MetaboLights (EBI)—ISA-TAB	Genebank (NCBI)—various formats and tools Sequence Read Archive (NCBI)—SRA XML, various tools	PRIDE (EBI), PRIDE XML, PRIDE Converter 2 (no quantitative data)

To facilitate queries in heterogeneous data

There is a growing demand for finding, extracting, merging, and synthesizing information from multiple, disparate sources. Developments in biology, computer science, and information technology will accelerate progress in this area. The rapid access to and retrieval of the most relevant biological data, and the ability to query across heterogeneous database systems and/ or data is essential (Lacroix and Critchlow, 2003). Database integration plays an important role in this area. Two classical approaches to database integration are the construction of centralized data warehouses, and federated systems. However, linking and discovery of data could also be achieved by exploring the relationships between existing distributed databases (Stevens *et al.*, 2001; Stein, 2002). In the case of phenotyping experiments, the prerequisite for such exploration is a structured set of semantic metadata comprising all experimental units, factors, biological objects, and treatments in an unambiguous manner. In particular, the use of persistent identifiers and vocabularies as well as their homogeneous assignment to a defined set of metadata elements and attributes is essential to ensure that relevant linkages are discoverable.

Improvements in querying heterogeneous data sets would create the possibility of comparing experiments with respect to their goals and results. Such comparisons are now not possible. They would help to evaluate new data sets in terms of novelty and complementarity to already existing data, which in turn is important for the efficiency of phenotypic studies and the sustainability of data storage. They would require a combination of the comparison of textual metadata and numerical data, using appropriate similarity measures [based on a semantic distance for metadata (e.g. Pesquita *et al.*, 2009; Oellrich *et al.*, 2015) and on Euclidean distance for quantitative traits]. It is clear that both metadata and data must be semantically described and standardized to allow matching and computations. In our opinion, the hierarchical organization of metadata (possible, for example, in the ISA-TAB format) would help distinguish different metadata types and levels, and ensure that redundant, meaningless or incorrect similarity measures are not taken into account; for example, those resulting from the traits reported in different scales or measured on different organs.

To facilitate statistical analysis and meta-analysis

Each phenotyping experiment has its own peculiarities that have to be taken care of in data analysis if this stage is to be

successful in terms of knowledge extraction. Fortunately, there are also some general common features that may be used to formulate a general model. For example, the multienvironment series of experiments run by breeders for selection of breeding lines suitable for further screening are performed in standard designs, usually in (incomplete) blocks, and a linear mixed model is mostly used to analyse data (Caliński *et al.*, 2005; Smith *et al.*, 2005). We think that the use of a standard structure for the experiment description and for the annotation of experimental factors would allow for automatic identification of factors with fixed effects and factors with random effects, and the functions of these effects (parameters), which are to be estimated, tested for significance, and interpreted. For automation of the analysis, it is also important that the annotation and formatting systems recommended for phenotypic data be able to describe specific features such as representation of sources of biological and random error variation, and missing data.

A particular form of statistical analysis, called meta-analysis, often follows the data querying and extraction. The value of meta-analysis is under debate. Some authors state that this method may imply ignoring important differences between studies, leading to erroneous conclusions (Borenstein *et al.*, 2009; Stegenga, 2011). Others point out that meta-analysis, if used with care, can broaden the inference even if the data sets differ with respect to metadata and data (see, for example, the Cochrane Collaboration project at http://cochrane.org). Combining results of several phenotyping experiments performed in different environments can be done, for example, using 'additive main effects multiplicative interaction' (AMMI) or mixed-model approaches (Malosetti *et al.*, 2013). However, in our opinion, the meta-analysis models should not be constructed based on the metadata itself, but rather on their annotation to standard vocabularies or ontologies. The same applies to observations, standardized (normalized) with respect to the measurement protocols and scales via ontologies such as Trait Ontology or Crop Ontology. To our knowledge, algorithms of statistical analysis working in this way are not available at the moment, but we think that they will need to be constructed in the future.

To facilitate data publication

Some publishers allow for inclusion of supplementary information, including data tables, in the articles, or require submission of data (e.g. DNA sequences) to public repositories. However, we claim that currently the scope of experimental data that the authors are encouraged to submit and that are accepted in a standardized way, even by leading publishers (see, for example, http://nature.com/nature/authors/submissions/final/suppinfo.html), is inadequate in comparison with the variety of data being obtained in plant experiments. The situation may be improved by launching of the new Nature Publishing Group's journal *Scientific Data*, in which metadata submission is not restricted to any special area of science and is based on clearly defined specifications. Currently, however, the list of data repositories recommended by *Scientific Data* does not include any that addresses phenotyping. This implies that there is a need to create data commons allowing researchers to publish and share phenotypic data. Moreover, the policy of publication of the supplementary phenotypic metadata and data should be more broadly accepted by scientific publishers. Journal-specific instructions for authors should clearly define the minimum reporting requirements (as, for example, *Metabolomics* referring to the Metabolomics Standards Initiative, MSI), and even go further to provide a link to a compliant data repository. As many papers combine information obtained by different '-omics' protocols, the requirements should be as universal as, for example, those presented in our implementations.

The data life cycle, from experiments to scientific publications, generally follows the schema described by Arend *et al.* (2014; Fig. 1), starting with the 'data in drawers and on disks' and ending with 'data in research papers'. Several billion dollars have been invested at the bottom level, in the collection of phenotype data. Condensed and enriched with metadata, phenotype data published on their own are, in our opinion, not less valuable than selected data in paper supplements. This is the basis for efforts such as the Open Archival Information System (OAIS), which aims to preserve primary data and provide associated information to relevant communities. Such comprehensive models are, however, expensive and sometimes too demanding for short-term and mid-term research projects. Therefore, (free) data-sharing platforms and cloud storage (e.g. Dropbox and Google Drive) are becoming very popular as economic alternatives to custom project-level data infrastructure. These platforms tend to have limited security management and do not support proper biological metadata management. More comprehensive data-sharing and publication services such as e!DAL (http://edal.ipk-gatersleben.de) and figshare.com provide professional support not only for file handling, but also for consistent data identification (e.g. through DOI and URN). We hope that our work will draw the attention of the developers of these services to the problems of phenotypic data.

One should note here that data publishing and their long-term preservation is supported by the Open Science or Open Data policy promoted by scientific funding agencies and organizations in different countries, such as the BBSRC (UK), INRA (France), or NWO (The Netherlands), and by governments (Code de la recherche: Loi ESR Article L112-1). Open data are also strongly recommended at the European level (http://europa.eu/rapid/press-release_IP-13-1257_en.htm) and the US level (https://congress.gov/bill/112th-congress/house-bill/3699). As community awareness of these needs has grown, new resources do facilitate open data sharing, for example repositories such as Dryad (http://datadryad.org) or data journals such as Giga Science (http://gigasciencejournal.com). International initiatives such as ELIXIR (http://elixir-europe.org) with FAIR data stewardship and DataCite (http://datacite.org) are exploring sustainable research networks and infrastructures. All these initiatives should be and to a great extent are built on standards; in the case of phenotypic data, our approach could be taken.

Applications of phenotyping

For modelling in systems biology

In the modern approach of systems biology, experimentation and computational analysis are intertwined in cycles of integration of new experimental data based on hypotheses generated from computations (Kitano, 2002). In the context of the impact of standardization, a distinction between bottom-up and top-down systems biology is relevant. Bottom-up systems biology examines the mechanisms through which functional properties arise in the interactions of known components, whereas top-down systems biology identifies interaction networks on the basis of patterns observed in genome-wide data sets. In the bottom-up approach, integration could typically involve models, where various submodels describing aspects of a systems are combined; in particular, molecular, cellular, or tissue levels can be combined with whole-plant (crop) models, or classical crop models with systems biology models (Chew et al., 2014). In the top-down approach, integration mainly involves different molecular and phenotypic data types. It seems that the recommendations that we consider herein are more relevant to the top-down approach. However, the bottom-up approach also needs standards for the description of the predicted functions and, if applicable, the phenotypic effects.

Assessment of existing infrastructures and initiatives related to systems biology indicates that standardization is indeed a bottleneck. For example, the BioModels database (http://www.ebi.ac.uk/biomodels-main), a repository of systems biology models, contains in total >1000 curated models, including several dozens for plants. Model components in this database are linked to appropriate ontologies (e.g. Gene Ontology). The models are described in SBML, the current standard for model description (Hucka et al., 2003). However, no links are provided for the associated data and phenotyping experiments on which those models are based, except to the scientific paper describing the model. Although a standard is available describing how to report on associations between systems biology models and data (Dada, 2010), as far as we know this standard is not widely implemented. The ESFRI initiative Infrastructure for Systems Biology Europe (ISBE) aims at developing a distributed infrastructure for the integration and synthesis of systems biology, including the development of standards. More focused on the development of standards for computational models is the COMBINE project (http://co.mbine.org). We think that the recommendations that we work on will also be valuable for those initiatives.

Another example of the potential benefits of better standardization for systems biology involves the proper description of mutant phenotypes. Knockout or overexpression of genes is used for assessing the predictive performance of systems biology models by generating experimental data that can be compared with model predictions. There are good examples of using the qualitative controlled terminology or ontology-based annotation that highlight the utility of curated and standardized data sets for the modelling of a set of phenotypic traits and the underlying gene networks (e.g. Lloyd and Meike, 2012; Szakonyi et al., 2015). However, for many experiments described in the literature, especially for the model plant *Arabidopsis thaliana*, such interpretations are difficult because both a proper description of the associated experiment and a standardized description of the resulting phenotype are lacking.

For the evaluation of gene bank accessions

Many applications of plant products (e.g. Metzger and Bornscheuer, 2006; Tilman et al., 2006) require the preservation of biological diversity; that is, collection, maintenance, and characterization of crop plants and their wild relatives. Gene banks play an important role in the long-term conservation of plant genetic resources for food and agriculture (PGRFA). Their focus is not only on pure conservation, but also to provide new impulses to plant breeding, for example through the addition of new alleles to existing breeding programmes (Hoisington et al., 1999). An indispensable part of the activities of gene banks is the phenotypic characterization of accessions. Phenotypic data about gene bank material are generated during each regeneration cycle and would ideally be preserved and published. Seed regeneration is driven by different factors, such as the availability of material or germination capacity. As a result, the phenotypic data available are highly incomplete (non-orthogonal). The analyses of such data allow, for example, the identification of promising new alleles (Keilwagen et al., 2014). Around the world, there are ~1800 collections that conserve plant genetic resources for food and agriculture. About 625 collections comprising >2 million accessions are maintained in Europe (Engels and Maggioni, 2012).

The European Cooperative Programme for Plant Genetic Resources (ECPGR) has an ongoing initiative aiming at providing information about European plant collections via a central entry point—EURISCO (the European Search Catalogue for Plant Genetic Resources, http://eurisco.ipk-gatersleben.de). EURISCO is an information system presently comprising data on ~1.1 million accessions (>5600 genera) maintained *ex situ* in European collections; it is based on a network of 43 National Inventories (NIs) co-ordinated by so-called National Focal Points (NFPs). Although currently there are only passport data available, an extension of the system for the management of characterization and evaluation information will be in place soon. The large numbers of accessions indicate the challenges EURISCO is facing regarding the integration of phenotypic data of plant collections. In addition to non-orthogonality, phenotypic data of PGRFA collections are often characterized by many different formats, different trait names or synonyms, different scales, and different amounts and quality of meta-information. For meaningful exploitation of these data, proper standardization beyond current standards is essential; we hope that the proposed solutions will prove useful also in this area.

For variety testing and registration

Plant variety registration is a system developed to organize the protection of plant breeders' intellectual property

rights. This system is highly regulated by both international and national laws and is based on standards developed by specialized bodies. It is carried out by national registration systems (centres, breeders organizations) according to the recommendations produced by the International Union for the Protection of New Varieties of Plants (UPOV). Variety registration is generally based on the assessment of distinctness, uniformity, and stability (DUS). For some species, the assessment of the value for cultivation and use (VCU) is also obligatory. An additional element of the variety testing system is phenotyping for variety regionalization.

The EU variety registration system is centralized in the sense that varieties tested and approved at national level also enter the 'Common catalogue of varieties … that can be marketed in the EU'. The DUS testing system has an extensive set of standardization documents. There are standards concerning the design of the trials, a representative selection of plant material, a set of characteristics to be observed, and the observational methods or scales, both in general terms (see, for example, UPOV DUS recommendations TG/1/3, http://upov.int) and for a particular species (see, for example, UPOV DUS recommendations TG/19/10 for barley). VCU and regionalization tests are usually based on documents and schemes developed and approved at a national level. So, in general, the variety registration and testing area is relatively well covered by standards for single phenotyping experiments. It is fairly clear how the experiments must be organized, which characteristics should be measured, and which observational protocol should be used to provide unbiased measurements. However, much less attention is paid to the metadata with respect to the general conditions of the trials. The lack of proper annotation is partly due to the fact that the series of experiments are organized by national agencies at the set of locations that have been used for a long time, and it is thought that their local conditions are known. Although soil, climate, and weather observations are collected, they are not always included in data sets, which makes immediate comparison of different experiments difficult. This may not be required for DUS trials performed in one or two similar locations with the requirement that the environmental influence on the behaviour of the plants is minimized. However, in VCU and regional assessments, joint analyses of data coming from several trials are crucial. The problem is alleviated when the VCU data are to be combined with other types of phenotyping or genotyping records.

Another issue in variety registration is a lack of a common data format used by national agencies. This is largely due to historical reasons: the agencies use different processing methods originating from local data processing or statistics specialists. We think that the EU variety registration system could benefit considerably from better standardization at the level of experimental data exchange, possibly using our solutions. In fact, this may be necessary to achieve one of the goals described in the 'Strategic Plan 2010–2015' published by Community Plant Variety Office, of 'harmonization of the processing … of candidate varieties on a world-wide scale' (http://cpvo.europa.eu/documents/News/CPVO_Strategic_Plan_2010.pdf).

Other aspects of standardization

Phenotyping methodologies

As we already stated, the recommendations and guidelines for management of plant experiments and protocols presented by Poorter *et al.* (2012) and Hannemann *et al.* (2009), and the experimental variables that are currently investigated in the EU consortia DROPS and EPPN, form—in our opinion—the basis to capture the relevant environmental characteristics at the required spatial and temporal resolution. This knowledge can be applied to different experimental scenarios, ranging from controlled (growth chambers) to semi-controlled (greenhouse) growth conditions, and should be further adapted to field experiments. The location of the experiment is therefore part of the minimum set of information that should be captured along with weather data (Poorter *et al.*, 2012). These data are particularly important for observational and destructive phenotyping. However, they can also take into account the requirements of phenotyping platforms that focus on non-invasive or minimally invasive phenotyping protocols (for a description of the various data- and analysis-related issues of such platforms, see Granier *et al.*, 2005; Arvidsson *et al.*, 2011; Zhang *et al.*, 2012; Lièvre *et al.*, 2013; Tisné *et al.*, 2013; Chen *et al.*, 2014). When applying these protocols, it is important to capture information about the imaging technique used, sensors and calibration, the specific pre-processing, and the feature extraction algorithm. It is also important to document a number of other features, for example whether the plant is moving to an imaging station, or the imaging device is moving towards the plant (Fiorani and Schurr, 2013). A similar approach should be used in molecular laboratory-based phenotyping of samples using low- or high-throughput protocols. We think that the solutions that we develop are capable of dealing with this challenge.

Ontologies

To integrate data properly, we need to identify the objects we are working with and the observations made on those objects. In this way, we define interoperability pivots which allow the comparison of trials from different data sets or the construction of integrative data sets for meta-analysis or genetic analysis. The interoperability of the observations (phenotypic and environmental variables) can be achieved thanks to many ontologies, some of which are linked, and which are under active development (see, for example, the Phenotype RCN project at http://phenotypercn.org or the Planteome project at http://planteome.org). The pivot object (i.e. the observation variable) consists of a trait, a method, and a scale. The phenotypic trait or environment variable can itself be decomposed into an entity and a quality (EQ model; Mungall *et al.*, 2010; Arnaud *et al.*, 2012; Deans *et al.*, 2015). For instance, the trait 'leaf area' is related to the entity 'leaf' and the quality 'area'. Therefore, we need strong references for those entity and quality, and they are provided by the Plant Ontology and the PATO ontology (http://bioportal.bioontology.org/ontologies/PATO). We can also share trait dictionaries/lists,

which is a role fulfilled by the Crop Ontology (http://cropontology.org) that offers a good repository of species-specific trait ontologies, including some measurement protocols and notation scales. Such ontologies assist multiple communities, as a simple trait, such as 'spike density', can have different measurement methods and units in different contexts and/or projects. Furthermore, it could be classified as a morphological trait in one project and as a grain quality trait in another. We think that while there are ontologies that can be recommended for immediate use for phenotypic metadata, there are areas where existing ontologies do not fit phenotyping needs. This is especially obvious for environmental descriptions, for which further development is urgently needed and strongly recommended. For example, the physical and the chemical properties of the environment should be describable, with units, definitions of terms, and comprehensive and easily repeatable measurement protocols.

Tools and services

The growing volume of data and the growing number of data types create many challenges for repositories of biological data and for scientists filling these repositories. On both sides, considerably more time is needed to prepare data in the proper format for submission. Moreover, the resources available per submission to curate for quality and interoperability are declining, whereas such resources become increasingly important as more complex multifactorial experiments are performed. The range of data models needed to represent the data increases. This is requiring more expertise from downstream users and continuous attention from those maintaining these. The increased data volumes are therefore an important factor in ongoing discussions about the need, desirability, methodology, sustainability, and scale of material to be archived.

We claim that standards have an important role to play in addressing all these issues. A widely adopted standard can be implemented by the manufacturer of an experimental device, which can directly produce data in a format accepted by the laboratory information management system (Arend *et al.*, 2014); while databases can implement tools that facilitate direct submissions. Formal definitions of standards will promote the development of programs for automatic verification of data submissions, reducing the need for human involvement, and ensuring that all accepted submissions are compliant with the standards that are considered necessary. Moreover, community-agreed standards are vital to enable the distribution of efforts across the entire scientific community. A good standard means that one implementation—of a submission tool, a file parser, a user interface, etc.—is useful for many implementations. If multiple databases are established in a single area, the standard serves as a data exchange format, allowing different resources to present their collective data to users as if they were a single entity. A sufficiently well-defined standard allows submission of richer data that can be made available and verified automatically through standards-aware interfaces. Likewise, maintenance of data models can be distributed among members of the community working in

a given area, and the standards may evolve in line with considerations of what data are sufficiently worthwhile to store.

In the short term, the major advantages of the development of standards for data types and databases are the simpler, more reliable, cheaper integration of common data types into well-established platforms; for example, many genome browsers (e.g. the Ensembl genome browser) can dynamically integrate data supplied in common file formats such as GFF (General Feature Format), SAM/BAM (Sequence Alignment Map/Binary Alignment Map), or VCF (Variant Call Format), or made available through standard client–server protocols (e.g. the Distributed Annotation Server protocols). Submission tools and data for persistent archives have also accommodated standards such as for microarray data or MIxS for sequence data (Table 1), or GO for gene annotation. In our opinion, the adoption of standards for phenotypic data and of controlled vocabularies for trait descriptions should have similar effects. It is likely that much phenotypic data will be held locally, by the researchers and institutions generating them. Without standardized representation, these data will be of limited value to the wider community.

Conclusions

The main aim of standardization is to improve data utilization for better insight into biological phenomena. The formal and technical aspects of standards play a supporting role. However, decisions and agreements with respect to practical implementations seem to be more difficult to accomplish than consensus on general importance and principles. A path towards better standardization practices is currently taken by various projects that are ongoing, have been submitted, or are in the making. Technical papers based on the recommendations published by us at http://cropnet.pl/phenotypes are in preparation and will demonstrate the feasibility of our approach with exemplary use cases and phenotypic data sets subjected to standardization. Discussion on technical details of the recommendations that we develop, necessary for clear presentation of such examples, is beyond the scope of this paper. We invite all parties interested in the subject to visit our webpage and to comment on our solutions.

When considerable progress in standardization is made, some further applications can be discussed. For example, the compliance with the official standards and rules for data integration and preservation should be required as a necessary condition for acceptance of applications for new biological projects financed from public funds. Our experience says that the solutions currently proposed by the bioinformatics and statistics teams involved in such projects are not always optimal in the sense of interoperability and sustainability.

Some of us are responsible for data integration in large projects in which a systems biology approach has been taken and data are collected in different types of assays (DROPS, http://dropsproject.eu; POLAPGEN, http://polapgen.pl; GnpIS, PMID: 23959375; Breedwheat, http://breedwheat.fr; Amaizing, http://amaizing.fr; Aker, http://www.aker-better-ave.fr/en; Rapsodyn, http://rapsodyn.fr/en; PeaMust, http://

peamust-project.fr/peamust_eng, others). For many molecular data types, we are able to assist the experimenters in using proper recommendations and existing public databases for storing their raw and processed data together with metadata. This is not yet the case for phenotypic data. Well-developed phenotypic database systems such as Ephesis (http://urgi.versailles.inra.fr/ephesis) or the Triticeae Toolbox (http://triticeaetoolbox.org) are projects that go in a proper direction and serve some communities well. Initiatives such as WheatIS (http://wheatis.org) go further and aim to build internationally accepted solutions. However, in most projects, the phenotypic information is still stored locally in the structures that are not compatible with molecular-level databases and with structures used at other locations. The probability that such non-standardized data will never be used and vanish without any contribution to the better understanding of plant systems is very large; this probability can be decreased by acceptance of opinions presented herein and by increased efforts to build and implement good recommendations.

Availability of good ontologies is a necessary condition for interoperability and semantic links. An important effort in building the Trait and Environment Ontologies has already been made. This work must go on. There are still fields which are not covered sufficiently by ontologies, domains which must be reinforced, and there will always be new traits to be added. We recommend a bottom-up approach by gathering all useful experimental parameters used by researchers, agronomists, and breeders. A full capture of these will improve the content and use of ontologies.

Standardization of phenotypes is timely. The new non-destructive protocols, now attracting a lot of attention, can—through the implementation of standards and the release of standardized data—play a major role in the promotion of standardization. If standardization fails, poorly annotated and formatted data may generate a tidal wave of noise and confusion, and obscure the inherent information present in these data and others. Wide publication of rich data in standardized formats—whether in organized databases, or in publishers' or institutional repositories—is a necessary step towards the establishment of a truly semantic web. This will be a network in which data can be indexed and searched using generic frameworks and in which developers will be able to produce genuinely flexible tools.

Acknowledgements

References

Annicchiarico P. 2002. *Genotype environment interaction: challenges and opportunities for plant breeding and cultivar recommendations* . FAO plant production and protection paper **174**.

Arend D, Lange M, Chen J, Colmsee C, Flemming S, Hecht D, Scholz U. 2014. e!DAL—a framework to store, share and publish research data. *BMC Bioinformatics* **15**, 214.

Arnaud E, Cooper L, Shrestha R, *et al.* 2012. Towards a reference plant trait ontology for modelling knowledge of plant traits and phenotypes. *Proceedings of the International Conference on Knowledge Engineering and Ontology Development*, 220–225. 4–7 October 2012, Barcelona, Spain.

Arvidsson S, Pérez-Rodríguez P, Mueller-Roeber B. 2011. A growth phenotyping pipeline for *Arabidopsis thaliana* integrating image analysis and rosette area modeling for robust quantification of genotype effects. *New Phytologist* **191**, 895–907.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. 2009. *Introduction to meta-analysis*. Chichester: Wiley.

Brazma A, Hingamp P, Quackenbush J, *et al*. 2001. Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. *Nature Genetics* **29**, 365–371.

Caliński T, Czajka S, Kaczmarek Z, Krajewski P, Pilarczyk W. 2005. Analyzing multi-environment variety trials using randomization-derived mixed models. *Biometrics* **61**, 448–455.

Chen D, Neumann K, Friedel S, Kilian B, Chen M, Altmann T, Klukas C. 2014. Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *The Plant Cell* **26**, 4636–4655.

Chew YH, Smith RW, Jones HJ, Seaton DD, Grima R, Halliday KJ. 2014. Mathematical models light up plant signaling. *The Plant Cell* **26**, 5–20.

Cobb JN, DeClerck G, Greenberg A, Clark R, McCouch S. 2013. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. *Theoretical and Applied Genetics* **126**, 867–887.

Cochran WG, Cox GM. 1957. *Experimental designs*. New York: John Wiley & Sons.

Dada JO, Spasić I, Paton NW, Mendes P. 2010. SBRML: a markup language for associating systems biology data with models. *Bioinformatics* **26**, 932–938.

Deans AR, Lewis SE, Huala E, *et al*. 2015. Finding our way through phenotypes. *PLoS Biology* **13**, e1002033.

Dhondt S, Wuyts N, Inzé D. 2013. Cell to whole-plant phenotyping: the best is yet to come. *Trends in Plant Science* **18**, 8.

Engels JMM, Maggioni L. 2012. AEGIS: a regionally based approach to PGR conservation. In: Maxted N, Dulloo ME, Ford-Lloyd BV, Frese L, Iriondo JM, Pinheiro de Carvalho MAA, eds. *Agrobiodiversity conservation: securing the diversity of crop wild relatives and landraces*. Wallingford, UK: CABI Publishers, 321–326.

Fiehn O, Robertson D, Griffin J, *et al*. 2007*a*. The Metabolomics Standards Initiative (MSI). *Metabolomics* **3**, 175–178.

Fiehn O, Sumner LW, Rhee SY, *et al*. 2007*b*. Minimum reporting standards for plant biology context information in metabolomic studies. *Metabolomics* **3**, 195–201.

Fisher RA. 1947. *The design of experiment*. Edinburgh: Oliver and Boyd.

Fiorani F, Schurr U. 2013. Future scenarios for plant phenotyping. *Annual Review of Plant Biology* **64**, 267–291.

Granier C, Hamard P, Muller B, *et al*. 2005. PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in *Arabidopsis thaliana* permitted the identification of an accession with low sensitivity to soil water deficit. *New Phytologist* **169**, 623–635.

Hannemann J, Poorter H, Usadel B, Bläsing OE, Finck A, Tardieu F, Atkin O, Pons T, Stitt M, Gibon Y. 2009. Xeml Lab: a software suite for a standardised description of the growth environment of plants. *Plant, Cell and Environment* **32**, 1185–1200.

Hoisington D, Khairallah M, Reeves T, Ribaut J-M, Skovmand B, Taba S, Warburton M. 1999. Plant genetic resources: what can they contribute toward increased crop productivity? *Proceedings of the National Academy of Sciences, USA* **96**, 5937–5943.

Honsdorf N, March TJ, Berger B, Tester M, Pillen K. 2014. High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. *PLoS One* **9**, e97047.

Hucka M, Finney A, Sauro HM, et al. 2003. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19,** 524–531.

Keilwagen J, Kilian B, Özkan H, et al. 2014. Separating the wheat from the chaff—a strategy to utilize plant genetic resources from ex situ genebanks. *Scientific Reports* **4,** 5231.

Kitano H. 2002. Systems biology: a brief overview. *Science* **295,** 1662–1664.

Lacroix Z, Critchlow T. 2003. *Bioinformatics: managing scientific data*. Amsterdam: Elsevier.

Li YF, Kennedy G, Davies F, Hunter J. 2010. *PODD: an ontology-driven data repository for collaborative phenomics research. The role of digital libraries in a time of global change*. Berlin: Springer, 179–188.

Lièvre M, Wuyts N, Cookson SJ, et al. 2013. Phenotyping the kinematics of leaf development in flowering plants: recommendations and pitfalls. *Wiley Interdisciplinary Reviews: Developmental Biology* **2,** 809–821.

Lloyd J, Meinke D. 2012. A comprehensive dataset of genes with a loss-of-function mutant phenotype in Arabidopsis. *Plant Physiology* **158,** 1115–1129.

Lynch JP, Nielsen KL, Davis RD, Jablokow AG. 1997. SimRoot: modelling and visualization of root systems. *Plant and Soil* **188,** 139–151.

Malosetti M, Ribaut J-M, van Eeuwijk FA. 2013. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Frontiers in Physiology* **4,** 44.

Massonnet C, Vile D, Fabre J, et al. 2010. Probing the reproducibility of leaf growth and molecular phenotypes: a comparison of three Arabidopsis accessions cultivated in ten laboratories. *Plant Physiology* **152,** 2142–2157.

Ménard G, Biais B, Prodhomme D, Ballias P, Petit J, Just D, Rothan C, Rolin D, Gibon Y. 2013. High throughput biochemical phenotyping for plants. In: Rolin D, Gadal P, Jacquot JP, eds. *Metabolomics coming of age with its technological diversity. Advances in Botanical Research*. Amsterdam: Elsevier, 407–439.

Metzger JO, Bornscheuer U. 2006. Lipids as renewable resources: current state of chemical and biotechnological conversion and diversification. *Applied Microbiology and Biotechnology* **71,** 13–22.

Morrison N, Bearden D, Bundy J, et al. 2007. Standard reporting requirements for biological samples in metabolomics experiments: environmental context. *Metabolomics* **3,** 203–210

Mungall CJ, Gkoutos GV, Smith CL, Haendel MA, Lewis SE, Ashburner M. 2010. Integrating phenotype ontologies across multiple species. *Genome Biology* **11,** R2.

Nagel KA, Putz A, Gilmer F, et al. 2012. GROWSCREEN-Rhizo is a novel phenotyping robot enabling simultaneous measurements of root and shoot growth for plants grown in soil-filled rhizotrons. *Functional Plant Biology* **39,** 891–904.

Ni J, Pujar A, Youens-Clark K, et al. 2009. *Gramene QTL database: development, content and applications* . Database **2009,** bap005.

NISO Press. 2004. *Understanding metadata*, ISBN: 978-1-880124-62-8, available online at http://www.niso.org/publications/press.

Oellrich A, Walls RL, Cannon EKS, et al. 2015. An ontology approach to comparative phenomics in plants. *Plant Methods* **11,** 10.

Pesquita C, Faria D, Falcão AO, Lord P, Couto FM. 2009. Semantic similarity in biomedical ontologies. *PLoS Computational Biology* **5,** e1000443.

Poorter H, Fiorani F, Stitt M, et al. 2012. The art of growing plants for experimental purposes: a practical guide for the plant biologist. *Functional Plant Biology* **39,** 821–838.

Rayner TF, Rocca-Serra P, Spellman PT, et al. 2006. A simple spreadsheet-based, MIAME-supportive format for microarray data: MAGE-TAB. *BMC Bioinformatics* **7,** 489.

Rocca-Serra P, Brandizi M, Maguire E, et al. 2010. ISA software suite: supporting standards-compliant experimental annotation and enabling curation at the community level. *Bioinformatics* **26,** 2354–2356.

Sansone S-A, Rocca-Serra P, Field D, et al. 2012. Toward interoperable bioscience data. *Nature Genetics* **44,** 121–126.

Smith AB, Cullis BR, Thompson R. 2005. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *Journal of Agricultural Science* **143,** 449–462.

Stegenga J. 2011. Is meta-analysis the platinum standard of evidence? *Studies in History and Philosophy of Biological and Biomedical Sciences* **42,** 497–507.

Stein L. 2002. Creating a bioinformatics nation. *Nature* **417,** 119–120.

Stevens R, Goble C, Baker P, Brass A. 2001. A classification of tasks in bioinformatics. *Bioinformatics* **17,** 180–188.

Szakonyi D, Van Landeghem S, Baerenfaller K, et al. 2015. The KnownLeaf literature curation system captures knowledge about Arabidopsis leaf growth and development and facilitates integrated data mining. *Current Plant Biology* **2,** 1–11.

Tardieu F, Tubeorsa R. 2010. Dissection and modelling of abiotic stress tolerance in plants. *Current Opinion in Plant Biology* **13,** 206–212.

Taylor CF, Paton NW, Lilley KS, et al. 2007. The minimum information about a proteomics experiment (MIAPE). *Nature Biotechnology* **25,** 887–893.

Tilman D, Hill J, Lehman C. 2006. Carbon-negative biofuels from low-input high-diversity grassland biomass. *Science* **314,** 1598–1600.

Tisné S, Serrand Y, Bach L, et al. 2013. Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. *The Plant Journal* **74,** 534–544.

Yang W, Guo Z, Huang C, et al. 2014. Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nature Communications* **5,** 5087.

Yilmaz P, Kottmann R, Field D, et al. 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications. *Nature Biotechnology* **29,** 415–420.

Zimmermann P, Schildknecht B, Craigon D. 2006. MIAME/Plant—adding value to plant microarrray experiments. *Plant Methods* **2,** 1.

Zhang X, Hause RJ Jr, Borevitz JO. 2012. Natural genetic variation for growth and development revealed by high-throughput phenotyping in *Arabidopsis thaliana*. *G3* **2,** 29–34.