Science and Public Policy, 2025, **52**, 883–888
DOI: https://doi.org/10.1093/scipol/scaf039
Advance Access Publication Date: 25 August 2025
Perspective

The need for robust, FAIR phenomic databases supporting agricultural efficiency and resiliency

Jodi Callwood[1], Burcu Celebioglu[2], Nicholas Gladman[3,4], Jinha Jung[5], Jennifer Lachowiec[6], Elsa H. Quezada Rodriguez[7,8], John P. McNamara[9], Jennifer Clarke[10,*]

[1]PhD Program in Bioinformatics and Computational Biology, 2014 Molecular Biology Building, Iowa State University, Ames, IA 50011, United States
[2]Department of Plant Breeding and Genetics, 430 Strand Agriculture Hall, Oregon State University, Corvallis, OR 97331, United States
[3]Agricultural Research Service, US Department of Agriculture, 538 Tower Road, Ithaca, NY 14853, United States
[4]Cold Spring Harbor Laboratory, 1 Bungtown Rd, Cold Spring Harbor, NY 11724, United States
[5]School of Civil and Construction Engineering, 610 Purdue Mall, Purdue University, West Lafayette, IN 47907, United States
[6]Department of Plant Sciences and Plant Pathology, Plant Bioscience Building, Montana State University, Bozeman, MT 59715, United States
[7]Departamento de Producción Agrícola y Animal, Universidad Autónoma Metropolitana-Xochimilco, Calz. del Hueso 1100, Coapa, Villa Quietud, Coyoacán, 04960 Ciudad de México, México
[8]Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, Circuito, Mario de La Cueva 20, Insurgentes Cuicuilco, Coyoacán Ciudad de México, México
[9]Department of Animal Sciences (Emeritus), Washington State University, PO Box 646310, Pullman, WA 99164, United States
[10]Department of Statistics, University of Nebraska–Lincoln, 340 Hardin Hall North, Lincoln, NE 68583, United States

*Corresponding author. Department of Statistics, University of Nebraska–Lincoln, 2200 Vine Street Suite 230, Lincoln, NE, United States.
E-mail: jclarke3@unl.edu

The US agriculture and food systems research and education system remains the envy of the world, and the US Department of Agriculture and the Land-Grant University system lead the public and private partnerships that have improved agricultural productivity and human health phenomenally for over 160 years. The continuation of these improvements relies on equitable access to trustworthy data—particularly in genetics and phenomics—and the ability to leverage such data to address future scientific challenges. In this article, we discuss the growing need in agriculture for phenomic databases that follow findable, accessible, interoperable, and reproducible data (FAIR) guidelines, as well as the need for public policy supporting a sustainable funding model for these databases.

Keywords: agriculture; phenomics; FAIR; databases; sustainability.

1. Needs for phenomic databases in agriculture

The US Department of Agriculture (USDA) and the Land-Grant University system lead the public and private partnerships that have improved agricultural productivity and human health. The increases in crop yields, animal production, safety, healthfulness, and food security for people in the USA and around the world are astounding. Our system was built on cooperation and coordination between the public sector and private enterprise. The US agriculture and food systems research and education system remains the envy of the world and attracts the best from around the globe. We can produce food in sufficient amounts for global needs, yet there are gains to be made in productive efficiency. The system is immensely complicated, but a foundational requirement is gathering, validating, curating, storing, and using data. An underlying knowledge base of genetics (genome) and phenotypes (phenome) is critical. Therefore, this article focuses on one key part of that complex system, the phenomics of agricultural plants and animals. Phenomics refers to the systematic study of the traits of organisms. It is just one part of the need for findable, accessible, interoperable, and reproducible data (FAIR) (Wilkinson et al. 2016) in support of agriculture and food systems.

It is paramount to understand that not only scientists benefit from data sharing. We are serving people, all people. The growing requests from farmers and farm groups, businesses, and consumers for access to publicly supported research and information, and the recent Congress and Administration policy and funding support for such access, demonstrate broad potential benefit. This article will focus on the work on phenomics, but the needs and approach reach across all aspects of food and agriculture.

1.1 Addressing challenges in food production

Improved breeding methods have been phenomenally successful (Schultz et al. 2020; Kaniuka 2023); however, these approaches are slow, resource intensive, and limited in scope. It remains a challenge to predict how a specific crop or breed will perform under different environmental conditions (Rahman et al. 2015). The cost of including multiple genotypes and environments has, in many cases, resulted in studies with a decrease in the number of samples and an increase in the data collection per sample. This has reduced the power of comparison and the applicability of the results. Research to solve today's and tomorrow's problems requires serious investment in larger scale studies for usable knowledge.

In addition to phenotypic data acquisition and correlation of phenotypes with genetic data, more complex analyses of both genomic and phenotypic data will require additional resources and skills than currently available (Houle et al. 2010). The integration of phenotype data with genomic information enables more precise selection of desirable traits and easier reuse and validation of previous data, accelerating the development of high-performing crop varieties and livestock breeds. Phenomic databases enable the systematic collection and use of data on crop and livestock traits such as plant growth patterns, root structure, drought tolerance, post-harvest quality, milk yield, and meat quality, with metadata meeting current scientific standards. These traits have significant economic value; breeding to optimize these traits can contribute significantly to food and feed security.

1.2 Understanding genotype–environment interactions

The interaction between an organism's genetics and its environment (G × E) is fundamental to agriculture and our understanding of ecological and environmental adaptability. There is often not a direct link between a variant in a genome and an observable phenotypic trait; most key traits are polygenic, resulting in several phenotypic variations. The use of phenomics databases, integrated with genetic/genomic databases, is imperative to determine the key G × E relationships between soil quality, temperature, light, nutrition, humidity, and disease pressures that influence plant and animal traits. Using high-throughput tools such as drones, sensors, and artificial intelligence/machine learning (AI/ML) for large-scale, automated trait collection will be required (Kopeć, 2024). Sample processing remains slow in relation to the speed of application and improvement (Houle et al. 2010). Making phenotyping data FAIR is the best way to maximize the scientific and social return on the investment in phenotyping research, teaching, and application.

1.3 Enhancing agricultural sustainability and resiliency

Phenomic databases can significantly increase agricultural sustainability, maintaining productive and resilient systems regardless of changes in production or climate patterns or economic vagaries. By analysing how plants and animals respond to input levels in different environments, varieties that need less water or nutrients, or are more resistant to stress or disease, can be identified. This contributes to agricultural sustainability by reducing the use of natural or chemical inputs (Kant et al. 2011; Seleiman et al. 2021). One impressive example of the application of phenotypic knowledge is a camera-guided, laser-equipped weeding machine (Carbon Robotics 2025). The machine stores data on the weed phenotype and then moves down the field, using the camera and the database to direct a laser beam to the weed with the precision of the actual size of the seed. This represents how direct on-farm application of FAIR phenomic data is here, now.

Using high-throughput data from satellites and sensors, phenomic-based technologies can facilitate the management of livestock and crops specific to a geographic area or environment and increase productivity by reducing the amount of input. Phenomic tools increase water use efficiency in plants and reduce methane emission levels in animals (Islam et al. 2020). By monitoring plant and animal responses in real time, farmers can make data-driven decisions on irrigation, fertilization, feed, and pest control, reducing environmental impact while maintaining productivity (Ghamkhar et al. 2025). AI models contribute to biodiversity conservation and sustainability by analysing historical phenotypic data and environmental conditions to predict various parameters such as seed viability rate, disease resistance, reproductive efficiency, and adaptability of indigenous breeds to climate stressors (Tuia et al. 2022).

2. Development of phenomic databases

In this section, we present summary details on the technical breadth and expertise and methods needed for such a successful system. Developing robust phenomic databases necessitates multidisciplinary coordination that integrates cutting-edge technologies such as data curation, geospatial data science, computational infrastructure, and agricultural science.

2.1 Data acquisition technologies

Modern phenomic databases rely on diverse technologies to capture phenotypic traits at scale. Technologies for generating high-throughput phenotyping data in recent years include field-based platforms such as Uncrewed Aerial Vehicles (Chang et al. 2017) and autonomous rovers (Madec et al. 2017), controlled-environment systems (Jiang and Yang 2022), and lab-based sensors (Bogati et al. 2023). These systems generate multimodal data including RGB (red, green, blue), multispectral/hyperspectral, light detection and ranging point clouds, and X-ray computed tomography scan data. Airborne and spaceborne remote sensing further enable macro-scale phenotyping measurements. The Internet of Things devices facilitate monitoring of field conditions such as temperature and soil moisture; these data are often boosted by edge computing to enable real-time field management in conjunction with robotics to automate data collection for consistency.

2.2 Database processing, architecture, and storage solutions

Seamless data use requires thorough processing to ensure compatibility across the system, including sensor calibration and noise reduction, geometric and radiometric corrections, and normalization across devices or environments. It is essential to have well-documented processing steps included in the metadata to maintain contextual relevance and prevent potential confusion in the use of phenotypic data.

Scalable storage architectures are essential to handle large-volume datasets. Early adopters of these technologies stored the collected phenomic data in local storage solutions, but this option is quickly becoming obsolete. Recently, hybrid solutions that combine cloud-based systems (e.g. AWS, Google Cloud) for flexibility with on-premises servers for sensitive data are gaining popularity. Various database architectures are utilized to store big phenomic data. For example, relational databases (e.g. PostgreSQL) manage structured metadata, whereas NoSQL systems (e.g. MongoDB) store heterogeneous data types, such as images and time series. Alternatively, data lakes are often used to aggregate raw and processed datasets, with hierarchical storage tiers optimizing cost and access speed.

2.3 Data standardization

Interoperability demands adherence to standards. The 'phenotype community' is still in the early stages of establishing such standards. The Minimum Information About Plant Phenotyping Experiments (MIAPPE) initiative (Papoutsoglou et al. 2020) is the first of its kind, being an open and community-driven data standard aimed at harmonizing data from plant phenotyping experiments. The MIAPPE provides a specification that includes a checklist and a data model of the metadata required to adequately describe plant phenotyping experiments. FAIRsharing (Sansone et al. 2019) is a platform for up-to-date information about relevant standards (like MIAPPE), databases, and policies that promote the discoverability and reuse of research data (Goble et al. 2020).

Many factors can limit standards adoption, including insufficient communication with researchers, poor language definitions, resistance to change, and non-enforcement by journals and scientific societies. Information availability and education around best data practices remain a significant obstacle to data publication and public access. Creating a helpful, dependable data service that community members desire will also improve FAIR goals and standards adoption (Wilkinson et al. 2016). A centralized platform built on quality data standards encourages the community to establish and adopt robust biocuration practices. Research conducted through accessing these databases will enable researchers to add data to this platform, thus increasing the quality of available data to the community (e.g. Brown et al. 2021; Woodhouse et al. 2021; Gladman et al. 2022; Peng et al. 2023; Hislop et al. 2024). Educating the community about standard practices and citation when accessing public data increases participation from data generators (Shanahan et al. 2021) and stresses the importance and consistency of FAIR standards in research, workforce development and application.

2.4 Collaborative framework and open-source platforms

Open-source ecosystems accelerate responsible and relevant database development. Projects like CyVerse (Swetnam et al. 2024), PHENET (Wibberg 2024), and Data to Science (Jung et al. 2024) provide shared infrastructure, while platforms such as Breeding API (BrAPI) (Selby et al. 2019) enable interoperable data exchange. The Agricultural Genome to Phenome Initiative (AG2PI) (Tuggle et al. 2022) is a major research effort funded by the USDA National Institute of Food and Agriculture (NIFA) aimed at developing a framework for integrating genomic and phenomic data across diverse agricultural species. The initiative fosters collaboration between plant and animal scientists, ensuring that data collection methodologies and analytical tools are standardized. Furthermore, the International Plant Phenotyping Network and its regional partners are working together to increase the visibility and impact of plant phenotyping while enabling cooperation by fostering communication between stakeholders in academia and industry.

2.5 Genotyping platforms

Decreases in cost have increased integration of genotyping data, and combining genotyping and phenotyping has emerged as a popular multi-omics approach to studying genotyping. The integration of genotype and phenotype is a core requirement of solving complex problems in food production. The creation of mid- and high-density genotyping panels across various agricultural commodities permits rapid identification and selection of accessions, strains, cultivars, and stocks for breeding programmes and basic functional research. Single nucleotide polymorphism (SNP) or microsatellite genotyping arrays have been generated for blueberries (Clare et al. 2024), alfalfa (Medina et al. 2025), potato (Ellis et al. 2018), and cattle (Reverter et al. 2020). Development of these tools required collecting decentralized information from multiple sources, including classification of sample phenotypes using potentially non-standardized ontologies. Centralizing phenotypic data allows data standardization, enabling relationships to existing genotypic database samples, removing complications associated with data processing, and streamlining the process.

3. Benefits of improved phenotyping and trait discovery

Database interoperability and reproducibility rely on standardized ontologies, sample collection and processing, and metadata documentation. Creating phenomic databases that embody these features enables trait discovery programmes to use the data for AI/ML approaches. Genetic markers were identified through machine learning approaches using standardized morphological data from genome-wide association studies and a SNP genotyping panel (Medina et al. 2024). Open data access stimulates innovation in phenotyping, machine learning, data mining, and network analysis by reducing the cost of research and leveraging previous investments for novel discoveries. Bioinformatic and computer science tools rely on publicly available data. Benchmark datasets provide essential performance metrics and openly available datasets provide experimentally confirmed datasets that otherwise would be lacking in many research areas (Friedrich and Friede 2023). Tools for exploring genomics data platforms provide utility and ease of targeting specific information whose development would not be possible without existing data repositories (Shiell et al. 2025).

Integration of curated phenotyping data advances and strengthens many areas of science. With pangenomics, having quantitative and reproducible phenotypic data allows researchers to properly subset populations for trait discovery and breeding programmes. In crop trait discovery, numerous diversity and breeding populations are developed to optimize traits such as cold tolerance, disease resistance, biomass, flowering time, and/or grain/fruit traits (Visakh et al. 2024). Understanding the metadata of the original experiments and identifying these traits provide essential information for correctly identifying and studying phenotypic traits. Providing and maintaining a centralized phenotyping platform offers benefits to the broader scientific community to better serve the wider global community.

3.1 Existing phenotypic databases' impact on the scientific community

Phenotypic databases provide an invaluable resource for the scientific community, which supports the application in the field. The cellular phenomics databases supported by European Molecular Biology Laboratory (EMBL) provide a centralized database for *Homo sapiens* and *Drosophila melanogaster* with open access and indexable data (Kirsanova et al. 2015). The existence of this phenomics database enables shared phenotypic results in the research community, provides a maintained platform for authors to submit their data, provides credits to authors who submit their data, and encourages sharing between researchers, which overall improves the data quality within the scientific community (Burel et al. 2015). Generating phenotypic databases provides a resource for cataloguing known and unknown phenotypes, spurring innovation (Brookes and Robinson 2015). Access to supported and maintained phenotypic databases relevant to agriculture, with standardized formats, regular updates, and user support is required for the innovation we need.

Reliable and consistent data collection, trait descriptors, standardization, and metadata improve the agricultural framework for stakeholders. Germplasm repositories, such as the USDA National Plant Germplasm System and the USDA Animal Germplasm Program, ingest standardized ontologies and phenotypic data for various germplasm (Blackburn et al. 2019; Volk et al. 2021). These types of features improve prebreeding and gene discovery through genomics studies. Dense morphological data benefits collection management, providing curators with additional methods to check processes and ensure accurate germplasm validation and classification during storage and propagation. Genotyping panels improve repository collections by ensuring proper sampling of genetic variation (Park et al. 2021; Berry and Spangler 2023).

3.2 Integration with other databases

The problems of the agricultural and food systems are global, and thus the research and education needs are complex. We already coordinate several databases in agricultural, business, and medical contexts, so there are examples to learn from and build on. Areas such as food yield and production, food distribution and waste, and consumer preferences and spending are all examples of connected systems. FAIR access to these connected databases will be a requirement. The authors assume as given that the present uses of AI will develop into further and more rapid expansion of uses of AI across the board: AI/ML will be essential to fast and economic use of integrated databases (NASEM 2024a,b; National Academy of Medicine, 2025).

Phenomic databases translate complex genetic information into observable traits and are a hub for the integration of environmental metrics that affect target traits such as soil properties, weather metrics, and greenhouse gas concentrations. Integrating these datasets helps identify gene-trait associations and develop predictive models for breeding and trait selection. By linking phenotypic data with production databases, researchers and industry professionals can refine best management practices. Phenomic databases can contribute to post-harvest management by improving quality assessment techniques. Machine learning models can predict fruit ripeness, optimize storage conditions, and reduce food spoilage throughout the supply chain.

Innovation in phenotypic data repositories would improve the overall computational landscape by enabling genotypic by phenotypic analysis. Improvements to phenotypic data collection such as imaging, segmentation, and data processing have increased access to collecting quality phenotypic data (Onnela 2021). The increase in phenotypic data collection has highlighted the lack of appropriate database structures such as those that exist for other data types such as genomics, metabolomics, and transcriptomics. Integrating phenotypic data with microbiome, transcriptomics, and genomics data provides a more comprehensive picture of the organism's expression (Yang et al. 2021; Zhang et al. 2025). Database integration has long been an issue due to differences in nomenclature (Feng et al. 2022). The Universal Protein Resource (UniProt) (Bairoch et al. 2005; Rigden and Fernández 2022) provides an example of the data alignment between nomenclature done well. Nuances in database integration require meticulous curation by dedicated personnel; done correctly, this would be a massive benefit to our key beneficiaries: people who eat.

4. Policy to support and sustain agricultural phenomic databases

To successfully integrate databases to support agriculture, we have identified several obstacles and limitations including a lack of data, standardization, and quality; infrastructure limitations for accessibility or sharing; and inadequate policies, funding and management. A continued and expanded public and private partnership, combining government policy that supports research, acquisition, curation, storage and free use of data with the expertise of application and commercialization, has been the cornerstone of agricultural and food systems improvement. No single part of society, nor single country, did it alone, nor can they in the future. Plant and animal breeding organizations and commercial businesses, American agriculture businesses, veterinary medical associations and pharmaceutical businesses, chemicals and machinery improvements, and global cooperatives have all been the result of publicly funded research resulting in free use data, which can then be leveraged as appropriate to bring a product or service into use. Federally funded agricultural research and education have stimulated many jobs and economic activity, and supported towns, schools, counties and families. One successful example of development of critical research infrastructure in areas such as phenomics is the Horizon Europe strategy for research and technology infrastructures (European Commission: Directorate-General for Research and Innovation 2024). This strategy and its associated funding opportunities promote and support research data and services for national competitiveness.

The National Agricultural Statistics Service gathers information on a broad scale and is in high demand across the agricultural research, education, application and management system. The NIH NCBI is the prime example of a knowledge database (Pubmed) and specific research databases (Genbank) which are funded by taxpayer dollars at a very low cost relative to return. The success of crop and animal improvement is all based on a combination of forward-thinking government policies and funding; national and international commodity associations and economic development organizations; user fees (checkoffs); and commercial development and sales. Policy that can adequately support the rate of knowledge acquisition and application for the food systems of the future must be supported by cogent laws, regulations, tax structures, research spending and education of the workforce. This is consistent with strengthening the Land Grant University System, streamlining regulations, encouraging and supporting coordination and collaboration among university scientists, private industry, breed associations, and non-governmental organizations as was called for in the National Academies Board on Agriculture and Natural Resources Consensus Study (NASEM 2022).

Another key step forward was the 2022 Office of Science and Technology Policy (OSTP) 'Nelson' memorandum which directed that federally funded research data should be made openly accessible. This policy fosters transparency, encourages collaboration, and ensures that phenomic data can be leveraged by researchers, industry professionals, and policymakers to drive innovation. It is paramount that we, as scientists and teachers and consumers, continue to encourage Congress and the Administration to support these plans to their completion and implementation.

5. Conclusions and recommendations

Phenomic databases are transforming agriculture by providing a comprehensive understanding of how genetic and environmental factors shape plant and animal traits. Their integration with genomics, precision agriculture, and food distribution systems offers numerous benefits, from improving breeding programmes to reducing food waste. The development of phenomic databases is not without significant hurdles. Key challenges include data heterogeneity, computational costs, and open data accessibility. Recent advancements in computational frameworks and interdisciplinary collaboration are paving the way for scalable, equitable solutions. High-performance computing and open access phenomic databases may provide breakthroughs in processing speed and enable advanced AI/ML model development. These techniques, processes, and systems are not 'extras that would be nice to have' but absolute requirements to study, predict, prevent, and respond to our changing agriculture and food systems.

As our data technology evolves, the potential of phenomic databases will only grow. Public policies, such as the OSTP memorandum, play a crucial role in ensuring that phenomic data remains accessible, fostering collaboration across scientific disciplines and supporting sustainable agriculture. The expansion and refinement of phenomic databases will be essential for addressing global food security challenges and advancing a more resilient and efficient food production system.

Acknowledgements

The authors would like to acknowledge the helpful comments from all the members of the AgBioData Working Group on Phenomics Data.

Author contributions

Jodi Callwood (Methodology, Writing—original draft, Writing—review & editing), Burcu Celebioglu (Conceptualization, Writing—original draft, Writing—review & editing), Nicholas Gladman (Conceptualization, Writing—original draft, Writing—review & editing), Jinha Jung (Writing—original draft, Writing—review & editing), Jennifer Lachowiec (Conceptualization, Writing—original draft, Writing—review & editing), Elsa H. Quezada Rodriguez (Methodology, Writing—original draft, Writing—review & editing), and John McNamara (Conceptualization, Methodology, Project administration, Supervision, Writing—original draft, Writing—review & editing)

Conflict of interest: None declared.

Funding

This work was supported by US Department of Agriculture (USDA)-National Institute of Food and Agriculture (NIFA) [2024-77039-43724 to J.C.], US National Science Foundation (NSF)-Division of Integrative Organismal Systems (IOS) [2126334 to J.M.], Department of Energy (DOE)-Biological and Environmental Research (BER) [DE-SC0021369 to J.L.], and USDA-Agricultural Research Service (ARS) [8062-21000-051-000D to N.G.].

Data availability

There are no new data associated with this article.

References

Agricultural Genome to Phenome Initiative (AG2PI) https://www.nifa.usda.gov/grants/programs/agricultural-genome-phenome-initiative-ag2pi, accessed 8 June 2025. USDA NIFA.

Bairoch, A. et al. (2005) 'The Universal Protein Resource (UniProt)', *Nucleic Acids Research*, **33**: D154–9.

Berry, D., and Spangler, M. (2023) 'Animal Board Invited Review: Practical Applications of Genomic Information in Livestock', *Animal*, **11**: 100996.

Blackburn, H., Wilson, C., and Krehbiel, B. (2019) 'Conservation and Utilization of Livestock Genetic Diversity in the United States of America through Gene Banking', *Diversity*, **11**: 244.

Bogati, S. et al. (2023) *Examining Root System Variability in Soybean Using Imaging Technologies Authorea*. November 14, 2023. https://doi.org/10.22541/essoar.170000395.52860395/v1.

Brookes, A., and Robinson, P. (2015) 'Human Genotype–Phenotype Databases: Aims, Challenges and Opportunities', *Nature Reviews Genetics*, **16**: 702–15.

Brown, A. et al. (2021) 'A New Decade and New Data at SoyBase, the USDA-ARS Soybean Genetics and Genomics Database', *Nucleic Acids Research*, **49**: D1496–501.

Burel, J.-M. et al. (2015) 'Publishing and Sharing Multi-dimensional Image Data with OMERO', *Mammalian Genome*, **26**: 441–7.

CARBON Robotics (2025), https://carbonrobotics.com/, accessed 18 June 2025.

Chang, A. et al. (2017) 'Crop Height Monitoring with Digital Imagery from Unmanned Aerial System (UAS)', *Computers and Electronics in Agriculture*, **141**: 232–7.

Clare, S. et al. (2024) 'Development of a Targeted Genotyping Platform for Reproducible Results within Tetraploid and Hexaploid Blueberry', *Frontiers in Horticulture*, **2**: 1339310.

Ellis, D. et al. (2018) 'Genetic Identity in Genebanks: Application of the SolCAP 12K SNP Array in Fingerprinting and Diversity Analysis in the Global in Trust Potato Collection', *Genome*, **61**: 523–37.

europa.eu, 'Horizon Europe', https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe/strategic-plan_en, accessed 9 June 2025.

European Commission: Directorate-General for Research and Innovation (2024) *Horizon Europe Strategic Plan 2025-2027*. Publications Office of the European Union, 0, rue de Reims, L-2417 Luxembourg, Luxembourg.

Feng, X. et al. (2022) 'A Review of the Heterogeneous Landscape of Biodiversity Databases: Opportunities and Challenges for a Synthesized Biodiversity Knowledge Base', *Global Ecology and Biogeography*, **31**: 1242–60.

Friedrich, S., and Friede, T. (2023) 'On the Role of Benchmarking Data Sets and Simulations in Method Comparison Studies', *Biometrical Journal*, **66**: 2200212.

Ghamkhar, K. et al. (2025) 'Realizing the Potential of Plant Genetic Resources: The Use of Phenomics for Genebanks', *Plants, People, Planet*, **7**: 23–32.

Gladman, N. et al. (2022) 'SorghumBase: A Web-Based Portal for Sorghum Genetic Information and Community Advancement', *Planta*, **255**: 35.

Goble, C. et al. (2020) 'FAIR Computational Workflows', *Data Intelligence*, **2**: 108–21.

Hislop, L. et al. (2024) 'A Blueberry (Vaccinium L.) Crop Ontology to Enable Standardized Phenotyping for Blueberry Breeding and Research', *HortScience*, **59**: 1433–42.

Houle, D., Govindaraju, D., and Omholt, S. (2010) 'Phenomics: The Next Challenge', *Nature Reviews Genetics*, **11**: 855–66.

Islam, S. et al. (2020) 'Mitigation of Greenhouse Gas Emissions and Reduced Irrigation Water Use in Rice Production Through Water-Saving Irrigation Scheduling, Reduced Tillage and Fertiliser Application Strategies', *Science of the Total Environment*, **739**: 140215.

Jiang, Y., and Yang, Y. (2022) 'High-Throughput Phenotyping for Plant Growth and Biomass Yield of Switchgrass Under a Controlled Environment', *Grass Research*, **2**: 1–7.

Jung, J. et al. (2024) 'Data to Science: An Open-Source Online Platform for Managing, Visualizing, and Publishing UAS Data', in Thomasson, J. A., and Bauer, C., (eds.) *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping IX*, **13053**. pp. 12–5. Bellingham, WA, USA: Proceedings of SPIE.

Kaniuka, R. P. (2023) 'Fueling the Green Revolution', https://www.ars.usda.gov/oc/timeline/green/, accessed 9 June 2025.

Kant, S., Bi, Y., and Rothstein, S. J. (2011) 'Understanding Plant Response to Nitrogen Limitation for the Improvement of Crop Nitrogen Use Efficiency', *Journal of Experimental Botany*, **62**: 1499–509.

Kirsanova, C. et al. (2015) 'Cellular Phenotype Database: A Repository for Systems Microscopy Data', *Bioinformatics*, **31**: 2736–40.

Kopeć, P. (2024) 'Climate Change—The Rise of Climate-Resilient Crops', *Plants*, **13**: 490.

Madec, S. et al. (2017) 'High-Throughput Phenotyping of Plant Height: Comparing Unmanned Aerial Vehicles and Ground LiDAR Estimates', *Frontiers in Plant Science*, **8**: 2002.

Medina, C. et al. (2024) 'Multi-trait Modeling and Machine Learning Discover New Markers Associated with Stem Traits in Alfalfa', *Frontiers in Plant Science*, **15**: 1429976.

Medina, C. et al. (2025) 'Pre-breeding in Alfalfa Germplasm Develops Highly Differentiated Populations, as Revealed by Genome-Wide Microhaplotype Markers', *Scientific Reports*, **15**: 1253.

National Academies of Sciences, Engineering, and Medicine (2022) *Enhancing Coordination and Collaboration Across the Land-Grant System*, Washington, DC: The National Academies Press.

National Academies of Sciences, Engineering, and Medicine (2024a) *Artificial Intelligence and Automated Laboratories for Biotechnology: Leveraging Opportunities and Mitigating Risks: Proceedings of a Workshop—in Brief*, Washington, DC: The National Academies Press.

National Academies of Sciences, Engineering, and Medicine (2024b) *Artificial Intelligence at the Nexus of Collaboration, Competition, and Change: Proceedings of a Workshop—in Brief*, Washington, DC: The National Academies Press.

National Academy of Medicine (2025) *Generative Artificial Intelligence in Health and Medicine: Opportunities and Responsibilities for Transformative Innovation*, Washington, DC: The National Academies Press.

Onnela, J.-P. (2021) 'Opportunities and Challenges in the Collection and Analysis of Digital Phenotyping Data', *Neuropsychopharmacology*, **46**: 45–54.

Papoutsoglou, E. A. et al. (2020) 'Enabling Reusability of Plant Phenomic Datasets with MIAPPE 1.1', *New Phytologist*, **227**: 260–73.

Park, S. et al. (2021) 'Population Genetics and Genome-Wide Association Studies Provide Insights into the Influence of Selective Breeding on Genetic Variation in Lettuce', *The Plant Genome*, **14**: e20086.

Peng, S. et al. (2023) 'Functional Annotation of the Animal Genomes: An Integrated Annotation Resource for the Horse', *PLOS Genetics*, **19**: e1010468.

Rahman, H. et al. (2015) 'Phenomics: Technologies and Applications in Plant and Agriculture', in Barh, D., Khan, M., and Davies, E., (eds.) *PlantOmics: The Omics of Plant Science*, New Delhi: Springer.

Reverter, A. et al. (2020) 'A Low-Density SNP Genotyping Panel for the Accurate Prediction of Cattle Breeds', *Journal of Animal Science*, **98**: skaa337.

Rigden, D. J., and Fernández, X. M. (2022) 'The 2022 Nucleic Acids Research Database Issue and the Online Molecular Biology Database Collection', *Nucleic Acids Research*, **50**: D1–10.

Sansona, S.-A. et al. (2019) 'FAIRsharing as a Community Approach to Standards, Repositories and Policies', *Nature Biotechnology*, **37**: 358–67.

Schultz, B., Serão, N., and Ross, J. W. (2020) 'Chapter 23—Genetic Improvement of Livestock, from Conventional Breeding to Biotechnological Approaches', in Bazer, F. W., Lamb, G. C., and Wu, G., (eds.) *Animal Agriculture*. pp. 393–405. Cambridge, MA: Academic Press.

Selby, P. et al. (2019) 'BrAPI—An Application Programming Interface for Plant Breeding Applications', *Bioinformatics*, **35**: 4147–55.

Seleiman, M. et al. (2021) 'Drought Stress Impacts on Plants and Different Approaches to Alleviate Its Adverse Effects', *Plants*, **10**: 259.

Shanahan, H., Hoebelheinrich, N., and Whyte, A. (2021) 'Progress Toward a Comprehensive Teaching Approach to the FAIR Data Principles', *Patterns*, **2**: 10.

Shiell, M. et al. (2025) 'Overture: An Open-Source Genomics Data Platform', *GigaScience*, **14**: giaf038.

Swetnam, T. L. et al. (2024) 'CyVerse: Cyberinfrastructure for Open Science', *PLoS Computational Biology*, **20**: e1011270.

Tuggle, C. et al. (2022) 'The Agricultural Genome to Phenome Initiative (AG2PI): Creating a Shared Vision Across Crop and Livestock Research Communities', *Genome Biology*, **23**: 3.

Tuia, D. et al. (2022) 'Perspectives in Machine Learning for Wildlife Conservation', *Nature Communications*, **13**: 792.

Visakh, R. et al. (2024) 'Precision Phenotyping in Crop Science: From Plant Traits to Gene Discovery for Climate-Smart Agriculture', *Plant Breeding*, **0**: 1–29.

Volk, G. et al. (2021) 'Integrating Genomic and Phenomic Approaches to Support Plant Genetic Resources Conservation and Use', *Plants*, **10**: 2260.

Wibberg, D. (2024) 'The PHENET EU Project—Advancing Training, Fostering Innovation and Streamlining Data Management for Plant Phenotyping and Envirotyping [Version 1]', *F1000Research*, **13**(ELIXIR): 1047.

Wilkinson, M. D. et al. (2016) 'The Fair Guiding Principles for Scientific Data Management and Stewardship', *Scientific Data*, **3**: 160018.

Woodhouse, M. et al. (2021) 'A Pan-Genomic Approach to Genome Databases Using Maize as a Model System', *BMC Plant Biology*, **21**: 385.

Yang, Y. et al. (2021) 'Applications of Multi-omics Technologies for Crop Improvement', *Frontiers in Plant Science*, **12**: 563953.

Zhang, N. et al. (2025) 'Integration of Multi-omics Data Accelerates Molecular Analysis of Common Wheat Traits', *Nature Communications*, **16**: 2200.

Science and Public Policy, 2025, 52, 883–888
https://doi.org/10.1093/scipol/scaf039
Perspective