eLife

FEATURE ARTICLE

SCIENCE FORUM

Vision, challenges and opportunities for a Plant Cell Atlas

Abstract With growing populations and pressing environmental problems, future economies will be increasingly plant-based. Now is the time to reimagine plant science as a critical component of fundamental science, agriculture, environmental stewardship, energy, technology and healthcare. This effort requires a conceptual and technological framework to identify and map all cell types, and to comprehensively annotate the localization and organization of molecules at cellular and tissue levels. This framework, called the Plant Cell Atlas (PCA), will be critical for understanding and engineering plant development, physiology and environmental responses. A workshop was convened to discuss the purpose and utility of such an initiative, resulting in a roadmap that acknowledges the current knowledge gaps and technical challenges, and underscores how the PCA initiative can help to overcome them.

PLANT CELL ATLAS CONSORTIUM, SURYATAPA GHOSH JHA†, ALEXANDER T BOROWSKY†, BENJAMIN J COLE†, NOAH FAHLGREN†, ANDREW FARMER†, SHAO-SHAN CAROL HUANG†, PURVA KARIA†, MARC LIBAULT†, NICHOLAS J PROVART†, SELENA L RICE†, MAITE SAURA-SANCHEZ†, PINKY AGARWAL, AMIR H AHKAMI, CHRISTOPHER R ANDERTON, STEVEN P BRIGGS, JENNIFER AN BROPHY, PETER DENOLF, LUIGI F DI COSTANZO, MOISES EXPOSITO-ALONSO, STEFANIA GIACOMELLO, FABIO GOMEZ-CANO, KERSTIN KAUFMANN, DAE KWAN KO, SAGAR KUMAR, ANDREY V MALKOVSKIY, NAOMI NAKAYAMA, TOSHIHIRO OBATA, MARISA S OTEGUI, GERGO PALFALVI, ELSA H QUEZADA-RODRÍGUEZ, RAJVEER SINGH, R GLEN UHRIG, JAMIE WAESE, KLAAS VAN WIJK, R CLAY WRIGHT, DAVID W EHRHARDT*, KENNETH D BIRNBAUM* AND SEUNG Y RHEE*

***For correspondence:** dehrhardt@carnegiescience.edu (DWE); kdb5@nyu.edu (KDB); srhee@carnegiescience.edu (SYR)

†These authors contributed equally to this work

Competing interests: The authors declare that no competing interests exist.

Funding: See page 20

Reviewing editor: Elsa Loissel, eLife, United Kingdom

© Copyright Jha et al. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Location-to-function knowledge can unlock new discoveries in plant science

The relationship between the sequence, structure and function of a protein has been explored from many angles, from biophysics to drug discovery (***Brausemann et al., 2017***; ***Rajendran et al., 2018***; ***Shang et al., 2020***). This long-standing paradigm has driven the development of innovative technologies in imaging, mass spectrometry, genomics and bioinformatics (***Camp et al., 2019***; ***Mallis et al., 2020***; ***Nakane et al., 2020***). Yet an evolving paradigm stating that the location of molecules is essential for their function at all scales, from molecular complexes to cells, organs, organisms and whole communities, also holds true (***Figure 1***).

At the molecular level, compelling examples of the location-to-function paradigm include the multi-enzyme complexes that work as molecular machines (so-called 'metabolons') to channel intermediate metabolites and produce specific molecular products in a highly controlled manner (***Obata, 2019***).

At the cellular level, the partitioning of proteins and lipids into organelles – and even

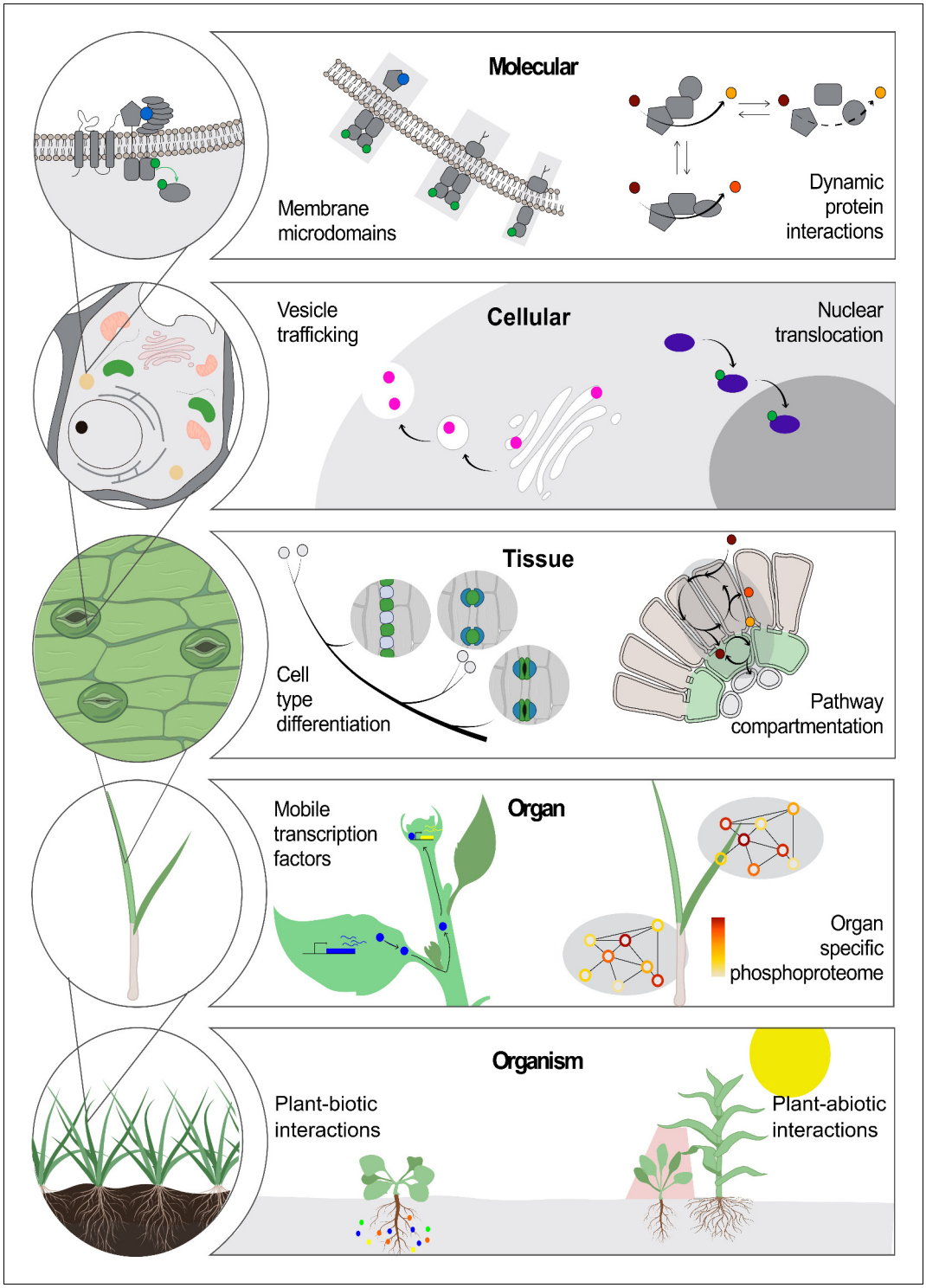


Figure 1. Location-to-function paradigm. The illustration shows the levels of organization of a plant (left) and a few examples highlight how location determines function at each level (right). At the molecular level, functions of protein complexes can be determined by their localization in membrane microdomains (***Jarsch et al., 2014***) or by dynamic protein interactions (***Obata, 2019***). At the cellular level, a protein can be located differentially through transport mechanisms, including vesicle trafficking (***Goring and Di Sansebastiano, 2018***) and nuclear translocation (***Marchive et al., 2013***), regulating its function. At the tissue level, cell position can drive its fate into a specialized cell type (***Shao and Dong, 2016***). Metabolic pathways can operate specifically in specialized cell types across tissues (***Marchive et al., 2013***; ***Schlüter and Weber, 2020***). At the next level, the existence of non-

Figure 1 continued on next page

Figure 1 continued

cell-autonomous transcription factors can transverse intercellular scales across plant organs (**Han et al., 2014**). Also, an organ-dependent post-translational proteome has been described as a mechanism of protein function regulation (**Uhrig et al., 2019**). At the organism level, plant interaction with biotic (**Harrison et al., 2002**) and abiotic factors (**Michaud et al., 2017**) can occur through a localized cue perception.

membrane microdomains within organelles – plays crucial roles in molecular trafficking and signal transduction, which have great impacts in many biological processes such as plant development and defence responses (**Jarsch et al., 2014**; **Xing et al., 2019**; **Yu et al., 2020**; **Heinze et al., 2020**; **Shimizu et al., 2021**). The differential location of a single molecule can determine one of several alternative functions (see, for example, **Smirnoff and Arnaud, 2019**). Protein function can also be dynamically regulated by a change in location, such as the activation of transcription factors by translocation into the nucleus (**Jiang et al., 2020**; **Marchive et al., 2013**; **Ramon et al., 2019**; **Pastor-Cantizano et al., 2020**). Furthermore, protein translocation to different cellular domains results in different functional outcomes (see for instance, **Kim et al., 2020**).

At the tissue level, C_4 photosynthesis is a classic example of how differential molecular localization between cell types enhances plant performance: differential expression of enzymes in bundle sheath and mesophyll cells in C_4 plants can concentrate CO_2 in close proximity to Rubisco, which is not possible in C_3 plants where the enzymes are present, but not differentially expressed between the cells (**Schlüter and Weber, 2020**).

Considering the opportunities in science and technology today, we contend that this 'location-to-function' paradigm will likely usher in conceptual leaps and transformative technologies, as was the case for the structure-to-function paradigm.

One particular area of potential success is agriculture, as the success of molecular approaches for crop improvement depends on a detailed understanding of the spatio-temporal regulation of genes and proteins and the production of metabolites (see **Wang et al., 2008** for an example of how controlling the location of gene expression impacts rice grain size and yield). The timing of expression is also a key determinant in imparting functional specification and influencing traits. For instance, engineered early expression of *WRINKLED1* (*WRI1*, an AP2 class transcription factor) during maturation increased *Arabidopsis* seed size and oil content (**Kanai et al., 2016**). More recently, an ambitious goal to globally sequester extra CO_2 into roots was established, which predicates on understanding what controls the development of a specific tissue, periderm, and the cell-type-specific expression of suberin, a lipophilic polymer found in plant cell walls. These examples illustrate how a thorough understanding of the cell- or tissue-specific expression of plant genes may

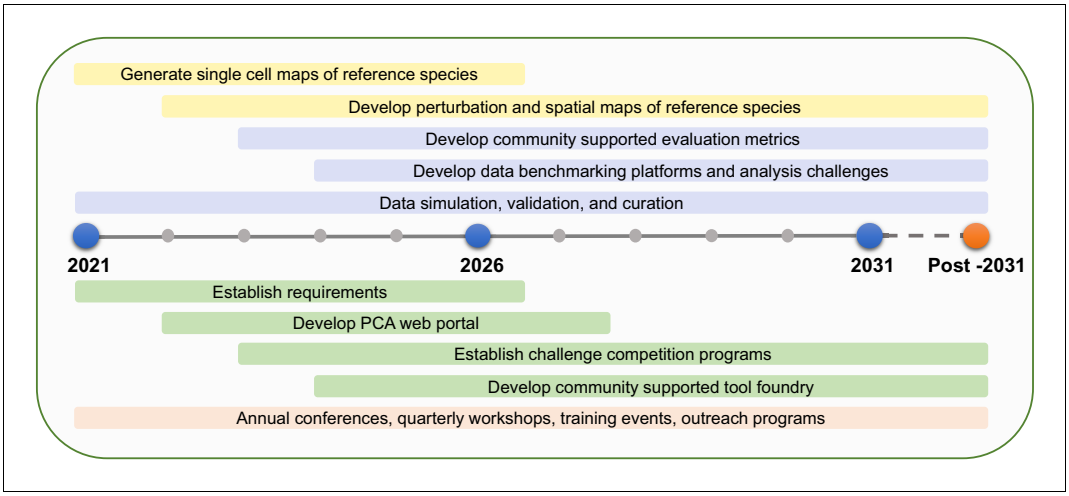


Figure 2. PCA milestones. PCA milestones for the next 10 years and beyond in data generation (yellow bars), data analysis (purple bars), software development (green bars), and building PCA community (orange bar).

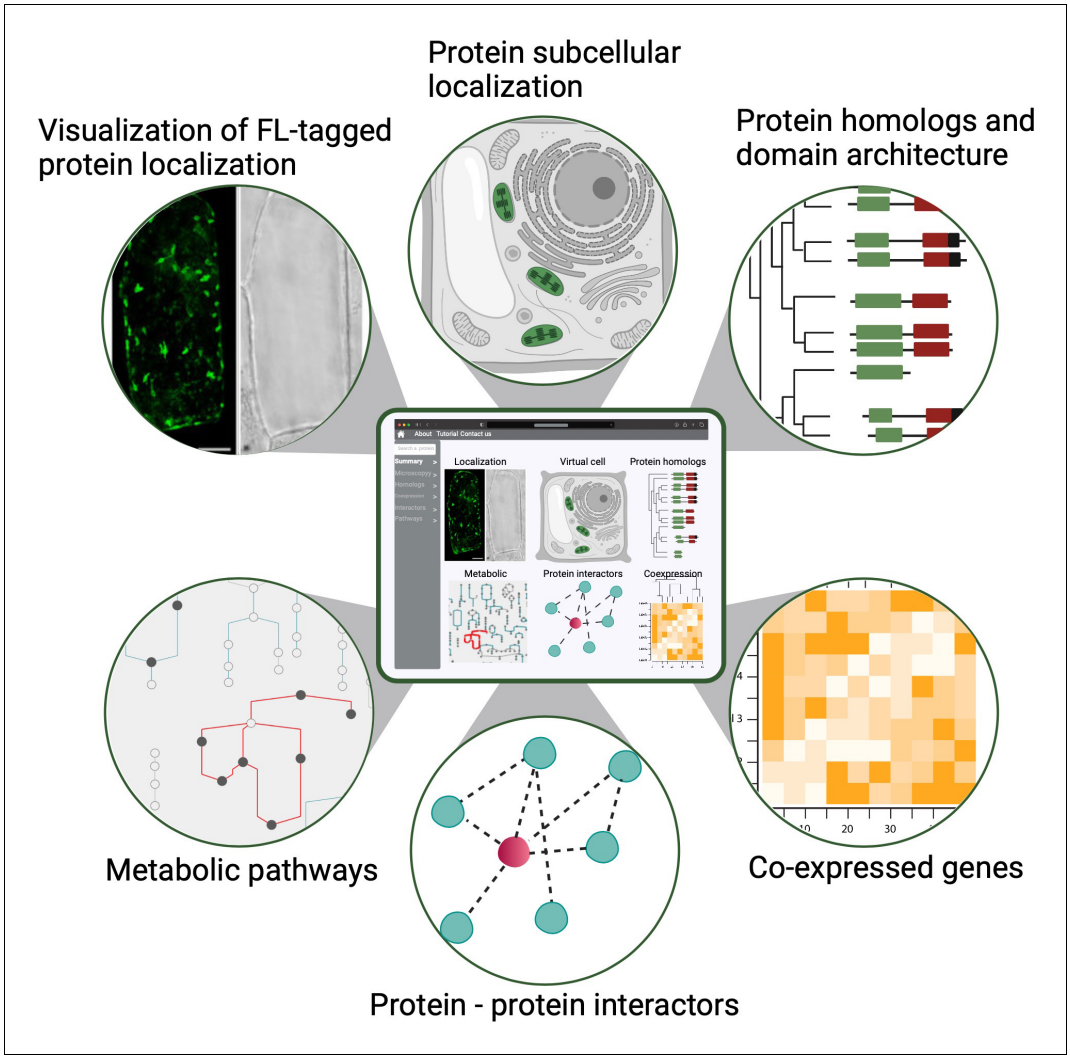


Figure 3. A conceptual diagram of a PCA user interface. An example of a PCA user interface that integrates various data types from molecular, biochemical, cellular and evolutionary contexts to connect location to function. Shown are example data types that would be seamlessly connected to enable easy navigation and discovery (FL: fluorescence).

enable the development of more effective biotechnological tools to facilitate crop improvement, plant resilience and climate change mitigation. Currently, these spatially- or cell-type-resolved expression data are missing or incomplete for most crop plants and there is no centralized repository for their integration. These shortcomings limit the rapid development of improved crops based on location-to-function data.

Unlocking location-to-function: The first Plant Cell Atlas (PCA)

The PCA workshop

Establishing a comprehensive picture of where molecules are located and how their locations inform function in higher plants will require concerted efforts from diverse groups of scientists. To catalyze such a community, a series of workshops were organized by Carnegie Institution for Science, New York University and Stanford University scientists around the topic of a cell atlas for plants, a framework to generate and synthesize molecular and cellular data on the development, dynamic functions and specialization of plant cells. This effort would also strive to

determine the state of the field, identify necessary components and people and discuss a path forward. Three workshops were held virtually during May and June of 2020. Over 400 scientists participated, of which 70% were early-career researchers, and 33% non-US based (***Rice et al., 2020***). Attendees of this first series of workshops became the basis of the PCA community. Following the success of the inaugural meeting, two technology-oriented workshops on single-cell sequencing and spatial proteomics and a career planning workshop were held virtually in 2021, respectively attracting 273, 253 and 42 participants. Additional workshops and symposia will be announced on the PCA website.

The PCA goals

The goal of the PCA is to bridge gaps in knowledge, providing critical localization, dynamics and interaction information at the cellular and subcellular levels. The participants in the PCA Research Coordination Network envision a set of concrete milestones over the next decade to reach that vision, based on community goals set during these meetings in 2020 and 2021 (***Figure 2***).

1. Identify and quantify relative abundances of all definable cell types in every organ of several strategically selected reference species (e.g. a single-celled organism, a non-vascular plant, a eudicot, a monocot).
2. Profile the developmental trajectory of every specialized cell type in each organ in the reference species through single-cell sequencing.
3. Develop a 'reference' clearinghouse where new experiments based on single-cell profiling can be mapped onto comprehensive atlases representing the diversity of known transcriptional states among cell types.
4. Profile the responses of major organ systems to a common set of environmental perturbations at a single-cell resolution.
5. Determine the intraspecific variation of developmental trajectories and environmental responses at a single-cell resolution.
6. Establish data integration platforms and community-supported data standards to better enable data integration.
7. Develop a community-supported data benchmarking platform that will accomplish simulation and validation of data and evaluation metrics.
8. Determine protein subcellular localization, interactions and dynamics in the major cell types of leaves, roots and flowers for every gene in reference species.
9. Curate empirical datasets and develop data analysis and visualization tools through community-driven challenges (e.g., the DREAM single-cell transcriptomics challenge; ***Alonso et al., 2019***; ***Pham et al., 2020***).
10. Establish mechanisms for fostering new collaborations and funding for data generation, data integration as well as new technology and method development or adoption for plants.

Some of these goals can be accomplished quickly, as already existing technologies such as single-cell and single-nucleus RNA-seq make them realistic. Others require the development of technologies to collect information at the single-cell level. As technology and discovery reshape the vision of the PCA, new aspirations should emerge from the community. Nonetheless, these ambitious milestones represent key gaps in knowledge that will accelerate plant research (if addressed) and are feasible to achieve given the opportunities presented by recent technological advances. The PCA community can enable this vision in several ways.

The PCA vision

First, the participants in the PCA Research Coordination Network envision a one-stop, user-friendly website with illustrated and graphical details about cellular organization across different developmental stages. This information, modeled after findable, accessible, interoperable and reusable (FAIR) data principles, could be widely accessed and used to advance research and develop educational tools (***Figure 3***; ***Wilkinson et al., 2016***). Although single-cell genomics data (mostly transcriptomic) have been generated and analyzed in individual studies for several years, how well the broader community can use these datasets in comparative analyses has been limited. This is largely due to the difficulty in integrating and comparing data generated by different methods, growth conditions and collection methods. The PCA will be critical in harnessing the power of single-cell data by establishing and promoting community-supported best practices in sample preparation, data generation, analysis and integration – including data benchmarking platforms, evaluation metrics and reference cell maps for major

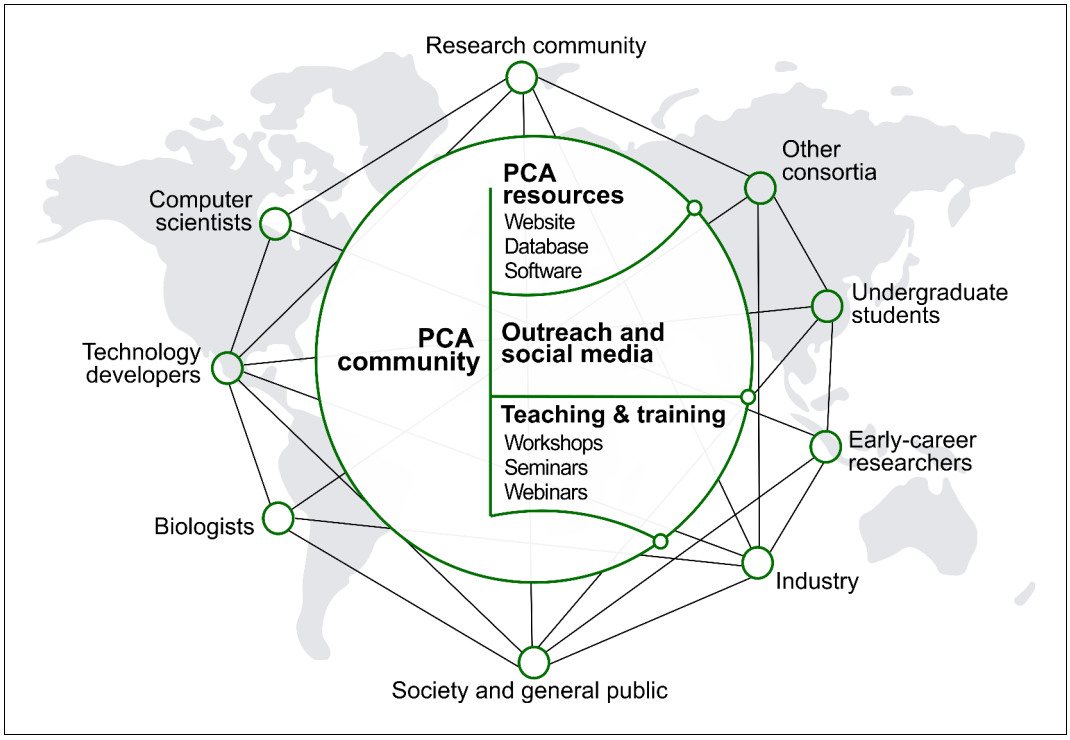


Figure 4. People and culture. Major stakeholders of the PCA are described. The goal is to establish a broad network of collaboration between developers and users through the creation of an accessible platform, educational tools and outreach activities. The PCA community should strive to be inclusive, diverse and transparent.

organ/cell types in model species. Important considerations for such infrastructure include how data will be gathered, curated and stored in order to integrate resources and contributions from different laboratories. It is necessary to develop and incentivize standards for both data collection (i.e. experimental protocols) and curation to allow integration of disparate datasets into a unified resource such as the PCA. This should include application programming interfaces (APIs) for programmatic access to ensure that PCA is accessible to all in a FAIR format. Similar standards have been produced for instance for plant phenotypic data (***Papoutsoglou et al., 2020***).

Second, we challenge and support the community to create a 4D representation of a developing plant from root to fruit (***Henkhaus et al., 2020***) with data collected from single-cell omics platforms (genomics, epigenomics, transcriptomics, proteomics and metabolomics) for each cell or cell type that are mapped onto tissue atlases of organ development. The data collected from the single-cell platforms form the basis of these atlases and serve as a reference for mapping further experimental data (e.g. from perturbation experiments). The PCA efforts include integration of not only single-cell omics data, but also spatiotemporal dynamics of single cells and their connection to cell types and developmental trajectories. It therefore facilitates a comprehensive understanding of how different complex tissues and organs are formed during plant growth and development.

Third, we wish to generate excitement and awareness about the potential for plant science research to address society's most pressing challenges in agriculture, food security, bioenergy, resource management, as well as ecosystem stewardship and robustness in response to climate change. The PCA community needs to overcome the growing gap between ongoing research and societal perception with public communication and education. The PCA efforts should include outreach, training and education of the general public to understand the motivations, successes and limitations of this research. The initiative will help to build public awareness of how translational research that arises from model organisms can positively impact society and help feed the world. It will also demonstrate how understanding plants as foundational

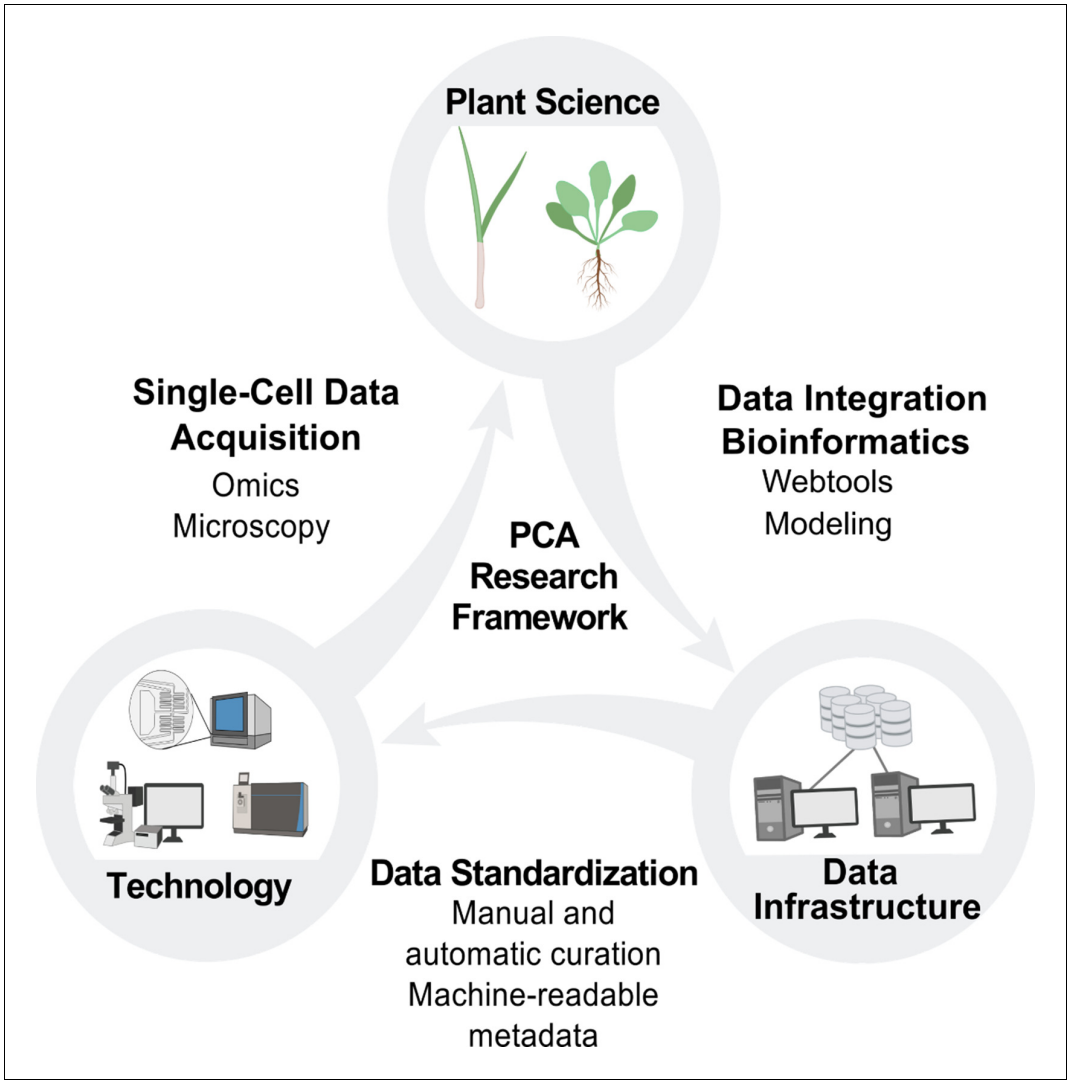


Figure 5. PCA research framework. Building the PCA will require cooperation and coordination between plant science, technology and data infrastructure. Plant scientists will use new technologies to generate data at the single-cell level, which will be integrated by data scientists into the PCA infrastructure through manual and automated curation. This data integration will enable the development of new technologies and drive new hypotheses and investigations for plant scientists, who will then contribute additional data to the PCA resource.

components of the ecosystem helps to predict and mitigate the consequences of climate change, and to be better stewards of the planet. Without these efforts, we fear that growing skepticism about science will impair the ability to aid society with the power of plant research, as apparent in the non-evidence-based regulation of gene editing in the European Union (***Jorasch, 2020***; ***Ruffell, 2018***; ***Zimny et al., 2019***).

Fourth, we see the PCA embracing new technologies and fostering multidisciplinary collaborations. The PCA will aim to build a community by introducing research tools to new users, and by facilitating interactions among scientists, engineers and breeders worldwide. These goals can be met by organizing in-person events and online workshops with presentations from domain experts, followed by breakout rooms for targeted discussions and networking. The multidisciplinary nature of the PCA has enormous potential for accelerating the fundamental understanding of plant cells and how they connect to the whole organism, and for translating fundamental discoveries to agriculturally important crops across the globe. The dream of the PCA initiative in this regard is to actively engage applied plant scientists and breeders by facilitating the formation of joint national and international funding schemes across scientific disciplines. This will allow diverse communities

to benefit from each other's expertise and for ideas and technologies to be transferred from lab to land. Enhancement in research translatability will address globally relevant societal questions that are aligned with the UN Sustainable Development Goals (**Biermann et al., 2017**; **Costanza et al., 2016**; **Griggs et al., 2013**), including food sustainability (Goals # 1–3), resources to tackle climate change (Goal #13) and the overall survival of the biosphere (Goals # 14–15). We envision that starting with a comprehensive cell atlas of model organisms will help to better understand non-model species and eventually expand to an ecosystem-scale analysis of interactions between organisms.

Inspiration from similar projects

What the participants in the PCA Research Coordination Network envisage is ambitious and not yet available for any system. However, examples of components of our vision will serve as guides. Among the animal whole organism atlases (**Lähnemann et al., 2020**), the mouse brain and adult mouse cell atlases are models for comprehensive coverage and community accessibility (**Rosenberg et al., 2018**; **Saunders et al., 2018**; **Zeisel et al., 2018**; **Tabula Muris Consortium, 2018**; **Han et al., 2018**). The Human Cell Atlas' data portal already includes data from 55 organs and 12 million cells, and it is becoming a model for making community-generated single-cell data available at a single portal, along with APIs and data analysis pipelines made easily accessible. In addition, the OpenWorm project provides an excellent example for modeling, visualizing and simulating various aspects of the biology and behavior of an organism (**Szigeti et al., 2014**). For plants, response to the environment would serve as an analogous level to behavior. EMBL-EBI's Single Cell Expression Atlas recently started hosting plant single-cell transcriptome data, which is an exciting development (**Papatheodorou et al., 2020**). Finally, the Global Natural Product Social Molecular Networking initiative provides a model for a one-stop-shop web infrastructure to enable data generation, integration, analysis and publication from mass spectrometry (**Wang et al., 2016**).

The envisioned PCA community

The PCA will be developed, maintained and used by a global community with diverse scientific, technical, creative and educational backgrounds (**Figure 4**). Helped by rapidly evolving online tools that increase inclusivity and transparency, the initiative should seek ways to facilitate the engagement, communication and collective decision-making among all groups involved. This work will aim to be grounded in active listening, psychological safety, teamwork, mutual respect and integrity. An inclusive and bottom-up ethos will underpin the development of global biological databases built with scientific rigor and transparency.

The participants in the PCA Research Coordination Network anticipate that the success of the PCA will hinge on an open and frequent dialogue between developers and users, with facilitation and guidance from a steering group. Recruiting scientists from diverse disciplines as well as from different career stages will be crucial in the successful development of the PCA. In particular, the initiative should continue to provide an opportunity for early-career researchers and foster the development of future leaders in plant biology.

Developers

The mission of the PCA community will be achieved by groups who will contribute to the development and maintenance of the database, as well as those who generate or use the data and provide inputs for improvement. Developing the PCA database will be done with the expertise of scientists, software engineers, modelers, scientific illustrators and animators coming together to fine-tune the display of cellular organization and collecting information. Data and computational scientists will have to be involved in the curation and annotation of data. Engagement will extend to consortia pursuing similar activities in different biological systems (e.g., Human Cell Atlas). As the PCA will heavily depend on data visualization, user interface and experience design will also benefit from advisors in art and design fields. Additionally, collaboration with cartographers will enhance the understanding of spatial mapping for developing the PCA database. These influences from outside the field of plant research will complement and bolster the expertise of plant scientists in the PCA community.

Users

The participants in the PCA Research Coordination Network expect several types of users. One of the major user groups will be the data generators, which will involve plant scientists from many disciplines, including single-cell profiling, protein-protein interactions, cell and

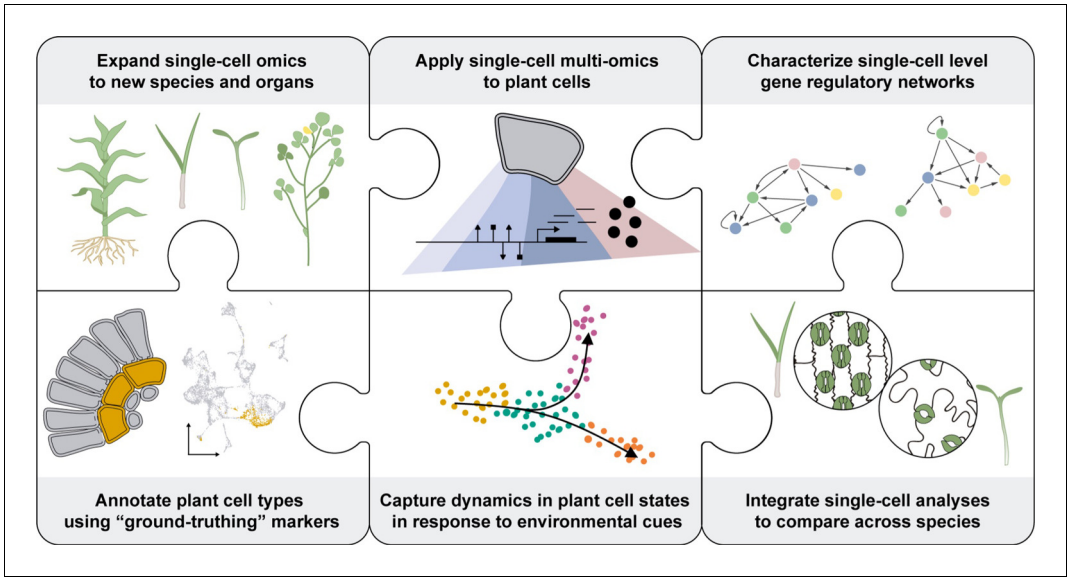


Figure 6. Knowledge gaps to fill for the PCA. Several limitations inherent to plants (shown as pieces of a jigsaw puzzle) must be overcome to enhance the understanding of plant cells as biological systems. To characterize the unique molecular attributes of each cell type composing a plant, plant biologists will need to develop reliable methods to broadly access the different cell types across various plant species. Also, single-cell multi-omics technologies will be necessary to get a deeper understanding of plant cell processes across all layers of molecular regulation. This data must be gained in the context of functional annotation of the cells. This would require the identification of reliable marker genes for each type of plant cell-. Ultimately, the spatiotemporal distribution of the molecular attributes of each cell will help to understand their dynamic regulation during plant cell development and in response to environmental stresses. Integrating the information collected using bioinformatics tools will enable the characterization of regulatory networks at a single-cell resolution and comparative analyses across plant species at the cellular level. This data will fuel the establishment of new synthetic biology strategies to enhance plant biology.

developmental biology, imaging or synthetic biology. To engage scientists from different career stages and areas of plant science, it will be important to develop mechanisms attributing and giving credit to data contributors. Another type of users will be researchers who wish to explore the PCA data to make discoveries. Engaging these users through a bottom-up process driven by their needs will be required for a vital and thriving PCA. Finally, the third group of users, both within and outside of the research community, are those who will access the PCA for educational purposes. Simple and intuitive interfaces combined with an attractive display of information and compelling visualizations will attract a large number and range of educators seeking information for teaching.

Integrating the PCA in the wider bioinformatics community

Robust computational infrastructure development will be essential to support the research, community-building and outreach activities of the PCA (***Figure 5***). We envision the PCA infrastructure (that is, the technologies, standards and computational platforms used by developers to build a framework to house and make available the PCA-relevant data) to have a seamless integration of multi-omics data with localization data at multiple scales of the plant. A governance structure, contribution guide and code of conduct for participants in infrastructure development should be established to define and support community values. These values include open-source development, FAIR principles and outcomes-based planning and implementation. Stakeholders should be included in the decision-making process. To implement these values, existing efforts with similar values and goals should be leveraged. For example, the DREAM challenges group could be leveraged for developing data analysis, visualization and access tools (***Pham et al., 2020***).

While new tools and resources will be needed to build a PCA community platform, existing resources should be strong guiding principles when developing the PCA infrastructure. The PCA should make efforts to liaison with data

repositories, knowledge bases and data analysis hubs such as ATTED, eFP Browser, ePlant, Ensembl Plants, Gramene, MaizeGDB, TAIR, The 1001 Arabidopsis Genomes phenotype database, NCBI's GEO, EMBL-EBI's Expression Atlas, AraPheno and others (**Obayashi et al., 2007**; **Sullivan et al., 2019**; **Howe et al., 2020**; **Tello-Ruiz et al., 2021**; **Portwood, 2018**; **Berardini et al., 2015**; **Seren et al., 2016**; **Barrett et al., 2005**; **Papatheodorou et al., 2020**; **Waese et al., 2017**). For example, OMERO is a global consortium effort aiming to standardize biological imaging data, which includes data depository as well as image processing software to translate image files into a universal format. Cross-species information may be mapped to the latest phylogenetic tree updated by the Tree of Life Web Project and Bioschemas. Engagement should also extend to consortia that are undertaking related work in other organisms, such as the Human Cell Atlas. To create a collaborative research environment, projects such as Synapse, Cyverse and KBase could be evaluated for strategic partnership. To develop a framework to represent the PCA at multiple scales, it will be important to implement conceptual frameworks such as the common coordinate framework (**Rood et al., 2019**), developmental trajectories ('pseudo-time') and spatial gradients ('pseudo-space'; **Stewart et al., 2019**).

Driving the community forward

The PCA will receive diverse sets of data from different types of techniques and experiments, many of which are constantly pushing the boundaries of the state of the art. As such, the PCA needs to position itself and its community at the frontier of plant biology, looking out for emerging findings and technologies. This could be achieved through online workshops that enable pioneering researchers and developers to introduce emerging technologies to the community. Other training activities might include hands-on workshops on single-cell, spatially resolved experimental techniques, data annotation and analysis tutorials, as well as hackathons and virtual office hours to help novice code developers, data scientists and other researchers.

Gaps in knowledge: How the PCA could address unresolved questions in plant science

Many unresolved questions in plant biology require new technologies and approaches that the PCA could drive forward, as well as benefit from. Examples of such major knowledge gaps are described here (see also **Figure 6**).

Plant development

The PCA could unlock answers to questions that require understanding the molecular changes in cells during development. The potential to characterize embryonic development comprehensively at single-cell resolution and in the multicellular context of complex tissues could help decipher the 'rules' underlying the earliest stages of plant tissue formation (**ten Hove et al., 2015**). In post-embryonic development, plant meristems are remarkable multicellular machines, continually producing new cells that must be specified to fate and then guided through phases of maturation (**Esau, 1953**; **Sinnott, 1960**). A major challenge is to understand the mechanisms that coordinate these cell maturation and fate specification events within meristems. Furthermore, plants present a fascinating contrast to mammalian systems with respect to the plasticity they maintain (**Borges, 2008**). For example, perennial plants often show meristematic plasticity, where only a subset of meristems per season are activated, and with the meristematic activity frequently modulated in response to external and internal cues (**Tylewicz et al., 2018**). The power of single-cell omics to capture the dynamic molecular coordination of tissue maturation at fine scales promises to help advance these classic questions in plant development (**McFaline-Figueroa et al., 2020**).

Response to the environment

As for problems in development, how plants respond to environmental stimuli requires understanding the molecular changes taking place in individual cells and how these changes are communicated cell to cell, across tissues and among organisms. An example can be found in how the threat of invasion of foreign biota into cellular and intercellular spaces often induces localized hypersensitive responses of programmed cell death (**Jones and Dangl, 2006**; **Ngou et al., 2021**). The question remains as to how the signals induced at the attack sites are communicated to neighboring cells and propagated to

Box 1. Major challenges in building a PCA.

1. Isolating plant cells and subcellular compartments. Efficient strategies for dissociating plant cells away from their structural context are needed to understand biological processes at the single-cell level. The fidelity and utility of methods requires extensive validation for collecting single-cell information across tissues and species.
2. Lack of well-characterized, cell-type-specific marker genes in non-model species. The functional characterization of plant cell types based on their transcriptomes suffers from the small pool of plant marker genes. The major limitation of the most widely applicable technology (Visium, 10x Genomics) is its relatively coarse resolution that would obscure the boundaries between several cell layers in plant tissues.
3. Incomplete annotation of plant genomes. Approaches that account for gene duplication, gene loss, and the potential for functional swapping among paralogs are required in plants. Comparison of individual transcript isoforms is difficult with current technologies. Efforts are needed to develop long-read technology to single-cell applications, or other strategies that can provide sequencing coverage to whole transcripts.
4. Need to integrate multiple components of the cell, such as nucleic acids (DNAs and RNAs), proteins, lipids, and metabolites. Addressing this challenge requires the development of multi-omics capabilities at cellular resolution, coupled with computational approaches that can integrate the data (***Libault et al., 2017***).
5. Plasticity of plant phenotypes. Plant phenotypes can be highly plastic according to changes in the environment. Efforts to establish reference atlases should ensure that standards in metadata description for growth conditions are established and adhered to.
6. Need to develop innovative strategies to functionally characterize plant genes at the single-cell level. Single-cell genetic technologies should be adapted to plant cells for high-throughput characterization of gene function in specific cells and cell types. For example, combining CRISPR-Cas9 technology with cell-type-specific inducible systems enabled the creation of genetic mosaics (***Decaestecker et al., 2019***; ***Wang et al., 2020***). In addition, studies of proteins are limited to the whole plant-based tissue analyses. Approaches are needed for investigating single-cell or cell-type-specific proteomics in a high-throughput manner (***Balasubramanian et al., 2021***).
7. Limited resources. The plant science community is small compared to its animal or human counterparts and having access to more resources is a long-term challenge. Despite this, this community is relatively open and collaborative, making efforts to standardize aspects of single-cell experimental and analytical strategies potentially an easier barrier to surmount.

distal regions. Do hidden pre-patterned states among seemingly homogenous cells help the plant anticipate attacks? Or is there heterogeneity in cellular responses that could reveal a more complex response to infection? How do these processes differ between susceptible and resistant plant varieties? Heterogeneous responses to infection around invasion sites, as well as differences in molecular signals that correlate with cellular immunity, could provide new insights into how successful defense responses are coordinated.

A parallel set of questions emerge from biotic interactions involved in symbioses, such as the ancient association of land plants with mycorrhizal fungi, or more recent symbioses like those formed by members of the nitrogen-fixing Fabids clade (***Wang and Qiu, 2006***). In the latter case, infection by the symbiont through the root hairs takes place asynchronously, so bulk sampling often leads to profiles that are a heterogeneous mixture of cell states (***Libault et al., 2010***). Can new technologies help to gain a discrete view of how individual cells respond to the symbiont over time? Can fresh insights be gained into plant-symbiont interactions by studying root cells and specialized symbiont structures (i.e., arbuscules, nodules) simultaneously? In addition, can the chain of cellular communication in the plant throughout the stages of microbial colonization be better understood? Lastly, by comparing these processes across plants with differential symbiotic specificities, is it possible to start understanding the hallmarks of a robust symbiotic relationship?

Cellular evolution

From rhizoids to pollen, plants have evolved many new specialized cell types, allowing adaptation to new ecological niches. In addition, specific cell types in the root harbor molecular mechanisms to perceive and respond to environmental stimuli, suggesting they may act as 'gatekeepers' for these responses (**Henderson and Gilliham, 2015**). What new perspectives on cellular evolution could we gain by producing comparable single-cell omics data sets on strategically selected species within the plant kingdom? Previous attempts to use expression patterns of homologous genes in distantly related species to assess functional orthology have provided insights (**Adamski et al., 2020**). However, bulk tissues collected in different species may have substantially different composition of cell types (**Huang and Schiefelbein, 2015**). Furthermore, the presence of novel cell types in some species may be obscured in whole organ samples (**Kajala et al., 2021**). By focusing on the fundamental unit of organs – cell types and cell states – comparative studies of species representing key plant lineages could identify minimal molecular toolkits driving conserved plant cellular functions. In addition, these single-cell approaches could highlight the differences in developmental pathways that lead to phenotypic diversity and adaptation to a myriad of environmental conditions.

Technical challenges faced in the development of the PCA

Critical challenges must be overcome to achieve the data generation, analysis and software development milestones of the PCA (**Figure 2**). Many mirror those faced by the larger single-cell initiatives that include animal models (**Lähnemann et al., 2020**), but we focus here on those that specifically affect the plant research community (summarized in **Box 1**).

Isolating individual plant cells

A major challenge is dissociating plant cells: due to their rigid cell walls, some tissues and many species are recalcitrant to digestion (**Himmel et al., 2007**; **Yoo et al., 2007**). This feature limits the ability to sample mature tissues that often mediate environmental stresses. In addition, it can be difficult to isolate even young, meristematic cells in some key species that present cellular innovations. One emerging approach that has the potential to address this issue in plants is single-nucleus profiling, which avoids the need for cell wall digestion (**Giacomello, 2021**). Such nuclear isolation methods have already shown promise (**Farmer et al., 2021**; **Marand et al., 2021**). Recently developed proximity labeling approaches bypass the need for isolating cells or organelles by tagging proteins and RNAs near a target protein in living cells and organisms. These methods can profile components of protein complexes, dynamic protein-protein interactions, specific cell types and subcellular locations of proteins or RNAs (**Zhang et al., 2019a**; **Mair et al., 2019**; **Huang et al., 2020**; **Kim et al., 2019**; **Wang et al., 2019**).

Tracking cell provenance

Another challenge is mapping cells (or nuclei) back to their tissue location in the whole plant. Methods that can rapidly and conveniently generate 'ground truth' markers of cell types in situ will be especially important as more tissues and species with few existing markers are profiled. Adapting so-called spatial transcriptomic approaches to plants will be a critical tool. In particular, untargeted methods (that is, methods based on unbiased sequencing rather than hybridization of a limited set of predefined probes) have the potential for the resolution and throughput needed (**Giacomello et al., 2017**). In addition, nomenclature systems will be necessary to link the reference maps from different species that have varying organ anatomy and specialized cell types. The hierarchical approach proposed by the common coordinate framework for humans can serve as a starting point (**Rood et al., 2019**). However, unlike the coordinate systems used in animal organ atlases which focus on accounting for differences between individuals, plant coordinate systems should instead be geared toward discovering differences among representative species and populations. Any such system should be able to accommodate differences in cell type composition (**Huang and Schiefelbein, 2015**) and the development of novel cell types (**Kajala et al., 2021**).

Gene homology

Another set of issues surrounds gene homology in comparisons across populations and species. Plants have a well-known capacity for duplication at many scales, ranging from single gene to whole genome. An informative molecular profiling of cells will require high-quality genome annotations that consider the gene family

expansion, gene loss and complex orthology relationships which frequently occur when extensive gene duplication follows speciation. Subfunctionalization and neofunctionalization can complicate assumptions about functional orthologs. On the other hand, difficulty identifying distant homologs that share low sequence similarity could obscure functional conservation. Advancement in sequence alignment methods, such as incorporating structural information, could help to discover these distant homologs (**Morton et al., 2020**). This will probably be an iterative process, where increased resolution of gene expression at the highly resolved cellular or subcellular level could allow inferences about functional orthology across species in large gene families.

Building the database and infrastructure of the PCA

Collection, standardization, curation, integration and visualization of data are key elements of a PCA platform, and the core PCA infrastructure should function as a data resource onto which services and tools are built.

Experimental and analytical standards

Establishing a vetted and community-approved set of experimental and analytical standards will be important. Such an approach was key to the success of the Encyclopedia of DNA Elements project, which ensured that data from multiple international laboratories were easily compared and integrated (**ENCODE Project Consortium, 2011**). Likewise, the 1001 Arabidopsis Genomes website provided a user-friendly interface to download all sequencing data as phases were completed – allowing the community to begin analyses and discoveries even before publication – and was complemented with an interactive tool to explore phenotypes (**Seren et al., 2016**) and their genomic associations (**Togninalli et al., 2018**). The standards were made easily available and updated over time as technologies advanced.

These platforms serve as data infrastructure blueprints as well as knowledge banks and inspiration. This would allow the PCA to be extended from an atlas of a single reference species and clone, to a set of atlases capturing natural variation (within and between species) in cell types and functions. For example, the interactive PCA platform could hyperlink the gene found in each tissue or cell type to the web portal of that gene in natural variation panels of the species of interest (such as the AraGWAS Catalog, **Togninalli et al., 2020**). Data structure scaffolds in the PCA could also be created so that the template can be replicated across genotypes of a species (e.g. different ecotypes) or different species, and allow users to select their desired organism or compare between the atlases.

Data visualization and exploration tools

In addition to data analysis and interpretation, the PCA infrastructure should include tools that provide effective ways to visualize real or virtual plant cells, and the dynamic processes that form and define them. This could enable hypothesis-generating in silico experiments that predict how plant cell types would respond to stimuli, or the properties of cells with a set of desired features. Creating these digital cells, tissues and other structures would require developing modeling, simulation, artificial intelligence, as well as data visualization and exploration tools similar to those being set up for other projects such as the OpenWorm (**Sarma et al., 2018**). Computational modeling and simulation-based predictions could then inform experimental work, the results of which could feed back into the PCA to improve models and predictions (**Radivojević et al., 2020**; **Zhang et al., 2020**).

Curating the PCA databases

To achieve the PCA's vision of building a 4D representation of a developing plant, a wealth of existing data, resources and tools will need to be leveraged. For example, it is now routine to deposit next-generation sequencing, processed expression data, proteomics, metabolomics and protein structures in numerous repositories worldwide (**Deutsch et al., 2017**; **Kaminuma et al., 2010**; **Leinonen et al., 2011**; **NCBI Resource Coordinators, 2018**; **wwPDB consortium, 2019**). A combination of automated searches and manual curation of datasets from existing data repositories can be used to build the PCA data infrastructure. Reference genome databases such as TAIR, Phytozome, Gramene, EnsemblPlants and others that provide gene and pathway annotation resources (e.g. Gene Ontology) could enable an initial assessment of gene structure, location and function (**Berardini et al., 2015**; **Goodstein et al., 2012**; **Howe et al., 2020**; **Tello-Ruiz et al., 2021**; **The Gene Ontology Consortium, 2019**).

Building a platform that brings these data together with data exploration and visualization tools that allow a user to explore at multiple scales, from molecular to cellular to organismal, is a challenge. To this end, single-cell datasets should be made available in both raw and processed data formats and linked from the public-facing PCA website. This resource could additionally be populated using automated literature search tools, followed by manual curation by the PCA community. A selection of high-quality 'reference' single-cell datasets directly hosted by the PCA may also prove to be useful for benchmarking purposes. For all datasets, curation standards should be imposed to include minimal metadata falling into several categories: (1) experimental metadata should include a project name, details regarding how plants were grown, treated and sampled, as well as quality control metrics and definitions used to evaluate sample integrity; (2) cell metadata should include a cell identifier (where applicable), cell type annotations and other cell-level data that have come from standard or specific analysis of the dataset; and (3) gene metadata that include structural and functional annotations should be required.

In addition to providing links to data repositories, the PCA portal should include web-based tools to generate, access and query relevant data and metadata using a standardized ontology for metadata, similar to the Single Cell Expression Atlas (***Papatheodorou et al., 2020***) and the Human Cell Atlas implementation of metadata types. This will facilitate interoperability of datasets. Depending on the data type, scripts that implement standardized pipelines for single-cell data analysis could also be made available that reflect the current best practices, including those that implement methods for cross-platform and multimodal data integration as well as label transfer (e.g., ***Lotfollahi et al., 2020***; ***Hie et al., 2019***; ***Korsunsky et al., 2019***; ***Kang et al., 2021***; ***Stuart et al., 2019***). Version-controlled source code that demonstrates analyses using the information in the PCA as a starting point should also be made available. These resources should promote greater use of the rapidly growing number of single-cell datasets.

Multiscale networks can provide an intuitive method to discretize and integrate different data types (***Duran-Nebreda and Bassel, 2017***). Other informatics frontiers that will be useful to explore include multi-dimensional image visualization and analysis tools such as Napari and Squidpy (***Sofroniew et al., 2021***; ***Palla et al., 2021***), graph-based knowledgebases for integrating datasets (e.g., Reactome; ***Fabregat et al., 2018***), natural language processing for automating curation (***Braun and Lawrence-Dill, 2020***), as well as machine learning and artificial intelligence methods for integrating heterogeneous data (***Ma et al., 2020***).

Challenges to the PCA infrastructure

Despite the widespread availability of data types described above, additional data that need to be integrated to achieve the PCA vision (metabolomics, imaging and phenotypic observations) are currently available in a variety of heterogeneous sources such as data sharing platforms (e.g. EMBL-EBI, Dryad, Figshare, and Zenodo), institutional data repositories, researcher websites or embedded within scientific publication records. A lack of centrality in data deposition makes these data less accessible.

In addition, methods for data standardization and curation will be critical tools to mitigate challenges with existing data, create best practices for collecting new data and ensure consistency to allow for data exchange and interoperability (***Sansone et al., 2012***). Experiments, datasets and other resources should be described in sufficient detail to be easily accessible and reproducible. Validating data with resource description frameworks is a means to ensure data compliance against the conceptual model it follows, data consistency and completeness. Data records and metadata should be machine-readable by using standardized vocabularies (i.e. cell type ontologies; ***Bard et al., 2005***) to enable interoperability. Existing solutions and concepts of such ontologies are broadly applied in vertebrates and are available for plants, though they need to be widely adopted by the PCA community (***Avraham et al., 2008***; ***Bard et al., 2005***; ***Diehl et al., 2016***; ***Ilic et al., 2007***; ***Xia and Yanai, 2019***). Mapping ontologies to the 4D representation of the PCA will facilitate linking and visualizing the growing knowledge of plant cells.

Originality of the PCA infrastructure

The long-term vision for the PCA infrastructure is similar in scope to other moonshot ideas such as the Human Cell Atlas (***Regev et al., 2017***), the Transparent Plant (***Henkhaus et al., 2020***) and the Virtual Plant (***Kost, 2001***). However, the PCA vision is different in several important ways. The Human Cell Atlas is for a single organism

Box 2. The chloroplast structure, function and dynamics illustrate PCA needs.

As highlighted here, technical solutions are needed to overcome key questions and challenges in chloroplast biology. The chloroplast is an important organelle in the plant kingdom, which is responsible for harvesting and converting light into chemical energy and sequestering carbon dioxide. Whereas much is known about chloroplast form and function, many aspects of chloroplast biology remain to be elucidated, especially in the context of the 'location-to-function' paradigm and proteostasis (***Figure 7***). The chloroplast is organized into several compartments (***Solymosi et al., 2018***; ***Figure 7A***). This suborganellar compartmentalization allows for highly efficient and specialized metabolism, optimized light-driven electron transport, redox reactions and more (***Rolland et al., 2018***). The function and organization of chloroplasts greatly varies across the plant lineage and cell types (***Pinard and Mizrachi, 2018***; ***Figure 7B***). For example, microalgae typically have a single chloroplast with a carbon-concentrating organelle called the pyrenoid (***Barrett et al., 2021***). In plants that can perform C_4 photosynthesis, the chloroplasts in the bundle sheath and mesophyll cells have different compositions and functions (***Edwards et al., 2004***). A few plants even perform C_4 photosynthesis within individual chlorenchyma cells where C_4 function is obtained by spatial separation of function between dimorphic chloroplasts located in central and peripheral cytoplasmic compartments within the same cells (***Erlinghaeuser et al., 2016***; ***Edwards et al., 2004***). In most land plants, the epidermal cells of leaves (except in guard cells) are devoid of chloroplasts but instead contain non-photosynthetic plastids, whereas aquatic plants are enriched in chloroplasts in their epidermal cell layer (***Larkum et al., 2017***; ***Han et al., 2020***; ***Figure 7C***). Recently *Arabidopsis* epidermal plastids have been shown to act in defense equipped with plant immune components (***Irieda and Takano, 2021***). Furthermore, different environmental conditions change the chloroplast proteome as well as relative positions of chloroplasts and their parts at various scales (***Kong and Wada, 2014***; ***Figure 7D***). Finally, the chloroplast life cycle includes various transformations in form and function from birth to differentiation to senescence, and how these transformations occur is an understudied frontier (***Llorente et al., 2020***; ***Solymosi et al., 2018***; ***Figure 7E***). It is still not fully understood how chloroplast and other types of plastids differentiate, adapt, function and integrate their activities within specific cell types. Single-cell technologies bring promise for solving key challenges in cell-type and species-specific differentiation of plastids. Examples of key technologies and how they could contribute to a range of PCA-related question are listed below:

- Affinity tagging of plastids using cell-type-specific promotors driving the expression of tagged proteins that associate to the plastid surface. Recently this approach has been developed for *Arabidopsis thaliana* mitochondria (***Niehaus et al., 2020***; ***Kuhnert et al., 2020***). Successful cell-type-specific plastid affinity purification would resolve molecular features, properties and associated functions of plastid with high precision. This could include determination of plastid RNA, protein and metabolite populations through high sensitivity, state of the art RNA-seq profiling and mass spectrometry techniques.
- Time-resolved plastid-targeted proximity labeling techniques to understand intra-plastid protein organization and dynamics. Recently, proximity labeling approaches have been successfully implemented in plants. They enable discovery of proteins that interact with specific regions of a bait protein, as well as identify localized protein interactions at high spatial and time-resolved resolutions (***Huang et al., 2020***; ***Mair et al., 2019***; ***Zhang et al., 2019b***). This novel approach would enable resolving intra-plastid protein organization and dynamics for different cell types and photosynthetic and non-photosynthetic plastids.
- Technology to detect and track retrograde signaling pathways from plastids to nuclei in response to, for example, singlet O_2, protein folding stress, redox and metabolites. These retrograde pathways involve a range of small molecules and likely also protein signaling cascades (***Jiang and Dehesh, 2021***; ***Muñoz and Munné-Bosch, 2020***; ***Pesaresi and Kim, 2019***). Non-invasive, high-resolution and high-sensitivity imaging technologies should be developed or adapted to address this challenge.
- Plastid-localized reporter systems (biosensors) to monitor pH, calcium, ATP and redox potential. There are now many such reporters (***Isoda et al., 2021***; ***Okumoto and Versaw, 2017***) and a fluorescent ATP sensor protein that allows the measure of ATP concentrations in the plastid and cytosol (***Voon et al., 2018***). Functional relationships between biophysical and biochemical parameters could be probed through the systematic application of reporter systems in plastids.
- Plastid protein lifetime reporter technologies to understand cell-type-specific proteostasis. A novel protein lifetime reporter system has only been applied to address N-degron pathways in the cytosol (***Zhang et al., 2019b***). Adaption to plastids would resolve the complexities of plastid proteome remodeling in response to development and (a)biotic changes (***Bouchnak and van Wijk, 2019***).
- Synthetic biology to engineer novel plastid metabolic pathways. Plastids are excellent targets for metabolic engineering through synthetic biology: they have their own genome and generally tolerate high levels of 'foreign' protein without significantly impacting function or proteostasis (***Boehm and Bock, 2019***; ***Jensen and Scharff, 2019***).

with over 200 cell types, whereas the PCA community embraces cell atlas data from a diverse array of plants. A typical plant has about 100 cell types (https://www.ebi.ac.uk/ols/ontologies/po; ***Ilic et al., 2007***). Understanding the diversity and origin of cell types across the plant lineage will be a central part of the initiative, as illustrated in ***Box 2*** with the chloroplast as an example. The PCA is an essential step towards realizing the vision of a Transparent Plant, a simulated environment where virtual plants can be built from parts and behaviors (***Henkhaus et al., 2020***). In order to materialize such a simulation environment, where the parts are and how they work together must be understood. Finally, we see the Virtual Plant database (***Katari et al., 2010***) as an essential component of the PCA vision and infrastructure.

Potential sources of funding for the PCA

The PCA will require adequate funding to achieve its full potential – including translational impact for agriculture (***Bartuska, 2017***; ***Bol et al., 2018***; ***Gök et al., 2016***; ***Hu, 2020***; ***Jaffe et al., 2015***; ***Rosenbloom et al., 2015***). As the PCA has an international reach, programs that facilitate collaboration across borders will be valuable in advancing the goals of the initiative, such as partnering programs by the Biotechnology and Biological Sciences Research Council (BBSRC) in the United Kingdom, European Commission's Horizon Europe and the United States' Fulbright scholars program. Similar programs exist that foster alliances between scientists from the US with those in Japan, Israel and Germany; and between India and other countries. As big data, nanotechnology, computation and artificial intelligence have now become an integral part of biological research, funds to support the interdisciplinary vision of PCA will be important to foster integration of these disciplines.

While governments are the major sources of funding for basic research, their level of funding has been steadily decreasing (***Khan et al., 2020***). Foundations are the major source of private basic research funding and are less constrained by time, politics and topic. They have a high potential for funding initiatives in emerging areas such as the PCA. However, it is challenging for foundations and individuals to find worthy initiatives. To fill this gap, the Science Philanthropy Alliance was established. One of their success stories is the Chan Zuckerberg Initiative providing substantial funds to develop technologies for the Human Cell Atlas initiative. Such a partnership for the PCA would take its vision to new heights.

Industry is another potential source of funding. Over 70% of applied research and development in the US is funded by industry (***Khan et al., 2020***). A thorough understanding of plant systems is essential for designing effective agro-biotech solutions leading to new crop varieties, and to innovative crop protection products that can leverage sustainable food production. A partnership with the PCA initiative would accelerate development of new solutions for customers of the agbiotech industry. Thus, funding and collaborations, ranging over multiple disciplines and countries, and multiway interactions amongst academia, industry and philanthropy will be essential to realize the vision of the PCA.

Conclusion and outlook

Future economies will be increasingly plant-based. Biomass is projected to be a major source of primary energy by 2050 (***Reid et al., 2020***), plant-based meat and dairy products are already transforming the food industry (***Shepon et al., 2018***), and plant-based vaccines against several diseases are increasingly being researched (reviewed by ***Shahid and Daniell, 2016***). With the rapidly changing climate, economy and values in our society, now is the time to reimagine plant science as a critical component of the future not only for agriculture, but also the environment, energy, health care, manufacturing and technology.

Today, plant science constitutes a minority of life sciences in terms of funding and workforce. To meet the demands of the future economy, we need to strengthen plant science, attract talent from other fields and train the next generation of plant scientists (***Henkhaus et al., 2020***). In 2004, the evolutionary biologist John Avise wrote that ''despite nearly three decades of experience with recombinant DNA techniques, the ultimate contribution to the broader human enterprise remains uncertain'' (***Avise, 2004***). Two decades after this statement, in 2020, recombinant DNA technology ushered in the beginning of an end to a global pandemic that cost millions of lives, through a record-breaking runtime of development and deployment of mRNA and DNA-based viral vaccines. While still at its infancy, we hope that the PCA initiative has the potential to serve as a nucleator,

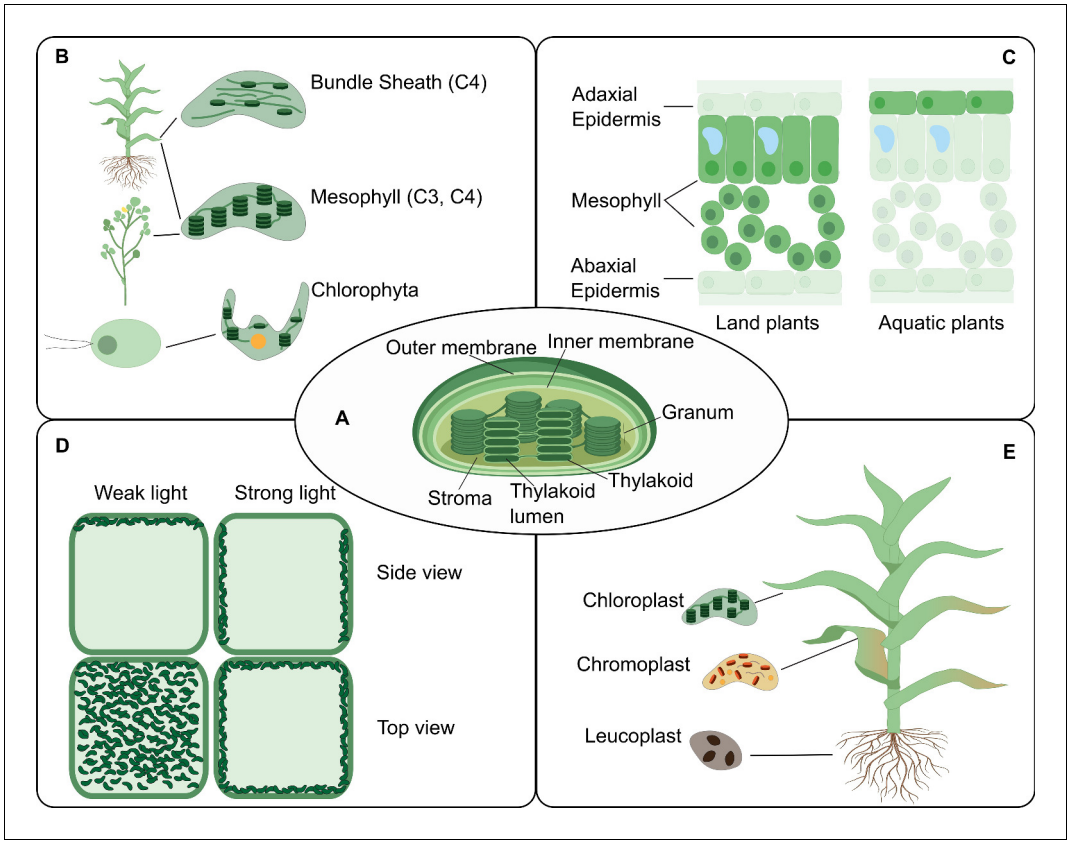


Figure 7. The chloroplast structure and function in connection to location. The PCA could facilitate an increased understanding of the biology of the chloroplast by providing accessible, integrative and spatially-resolved data for transcripts, proteins and metabolites across cell types (***Box 2***). (**A**) A representation of a typical chloroplast structure and its compartments. (**B**) The chloroplast structural differences in some plant lineage. (**C**) A leaf cross-section showing different chloroplast location at the tissue level in land and aquatic plants. (**D**) The chloroplast location can change within the cell upon environmental cues such as light intensity. (**E**) Various forms of chloroplasts in different organs.

bringing together scientists and engineers from a wide range of fields to solve fundamental problems in plant biology with innovative solutions from emerging technologies.

Acknowledgements

The PCA community-building activities are funded in part by the National Science Foundation grant numbers MCB-1916797 and MCB-2052590, Carnegie Institution for Science, and BASF. We thank Emily Fryer, Nick Melosh, Heather Meyer, Jason Thomas, Terri Tippets, Renate Weizbauer, Zhiyong Wang and Kangmei Zhao for helping with organizing the first workshop. We thank Emily Fryer and Julie Gosse at the Science Editors Network for developing the PCA website. We are grateful to the first PCA workshop steering committee members Jim Haseloff, David Jackson, Edward Marcotte, John Marioni, Marisa Otegui, Alberto Salleo, Waltraud Schulze, Edgar Spalding, Michael Sussman, Marja Timmermans and HS Philip Wong for their guidance and support. We thank Rachel Shahan, Shouling Xu, Kevin Cox and Erin Zess for their input in developing the manuscript. Some images in the figures were created with BioRender.com.

Plant Cell Atlas Consortium

Jahed Ahmed: Louvain Institute of Biomolecular Science and Technology, UCLouvain, Louvain La Neuve, Belgium; **Oluwafemi Alaba**: School of Biology and Ecology, University of Maine, Orono, United States; **Gazala Ameen**: Washington State University, Pullman, United States; **Vaishali Arora**: University of Delhi, Delhi, India; **Mario A Arteaga-Vazquez**: Universidad Veracruzana, Xalapa, Mexico; **Alok Arun**: Inter American University of Puerto Rico, Barranquitas, United States; **Julia Bailey-Serres**: Center for Plant Cell Biology and Department of Botany and Plant Sciences, University of California, Riverside, Riverside, United

States; **Laura E Bartley**: Washington State University, Pullman, United States; **George W Bassel**: University of Warwick, Coventry, United Kingdom; **Dominique C Bergmann**: Stanford University and Howard Hughes Medical Institute, Stanford, United States; **Edoardo Bertolini**: Donald Danforth Plant Science Center, St. Louis, United States; **Kaushal Kumar Bhati**: Louvain Institute of Biomolecular Sciences, Catholic University of Louvain, Louvain La Neuve, Belgium; **Noel Blanco-Touriñán**: Plant Vascular Development Group, ETH Zürich, Zurich, Switzerland; **Steven P Briggs**: University of California San Diego, La Jolla, United States; **Javier Brumos**: Instituto de Biologia Molecular y Celular de Plantas (IBMCP) Consejo Superior de Investigaciones Científicas (CSIC), Valencia, Spain; **Benjamin Buer**: Bayer AG, Monheim am Rhein, Germany; **Adrien Burlaocot**: HHMI, Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, United States; **Sergio Alan Cervantes-Pérez**: LANGEBIO-CINVESTAV, Irapuato, Mexico; **Sixue Chen**: Department of Biology, University of Florida, Gainesville, United States; **Bruno Contreras-Moreira**: European Bioinformatics Institute - EMBL, Hinxton, United Kingdom; **Francisco J CORPAS**: Estación Experimental del Zaidin (CSIC), Granada, Spain; **Alfredo Cruz-Ramirez**: National Laboratory of Genomics for Biodiversity CINVESTAV, Irapuato, Mexico; **Cesar L Cuevas-Velazquez**: Departamento de Bioquímica, Facultad de Química, Universidad Nacional Autónoma de México, Cd., Irapuato, Mexico; **Josh T Cuperus**: University of Washington, Seattle, United States; **Lisa I David**: North Carolina State University, Raleigh, United States; **Stefan de Folter**: Unidad de Genómica Avanzada (UGA-LANGEBIO), Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN), Irapuato, Mexico; **Peter H Denolf**: BASF Innovation Center Gent, Zwijnaarde, Belgium; **Pingtao Ding**: Institute of Biology Leiden, Leiden University, Leiden, Netherlands; **William P Dwyer**: Department of Plant Biology, Carnegie Institution for Science, Stanford, United States; **Matthew MS Evans**: Department of Plant Biology, Carnegie Institution for Science, Stanford, United States; **Nancy George**: European Bioinformatics Institute, Cambridge, United Kingdom; **Pubudu P Handakumbura**: Environmental Molecular Sciences Division, Pacific Northwest National Laboratory, Richland, United States; **Maria J Harrison**: Boyce Thompson Institute, Ithaca, United States; **Elizabeth S Haswell**: Washington University in St. Louis, St. Louis, United States; **Venura Herath**: University of Peradeniya, Peradeniya, Sri Lanka; **Yuling Jiao**: Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China; **Robert E Jinkerson**: University of California, Riverside, Riverside, United States; **Uwe John**: Alfrd-Wegener-Institute Helmholtz centre for Polar and Marine Research, Bremerhaven, Germany; **Sanjay Joshi**: University of Kentucky, Lexington, United States; **Abhishek Joshi**: Mohanlal Sukhadia University, Udaipur, India; **Lydia-Marie Joubert**: Stanford University, Stanford, United States; **Ramesh Katam**: Florida A&M University, Tallahassee, United States; **Harmanpreet Kaur**: Punjab Agricultural University, Ludhiana, India; **Yana Kazachkova**: Weizmann Institute of Science, Rehovot, Israel; **Sunil K Kenchanmane Raju**: Michigan State University, East Lansing, United States; **Mather A Khan**: Heinrich Heine University, Dusseldorf, Germany; **Rajdeep Khangura**: Purdue University, West Lafayette, United States; **Ajay Kumar**: ICAR-NIPB, New Delhi, India; **Arun Kumar**: Biotechnology Division, CSIR-Institute of Himalayan Bioresource Technology, Palampur, India; **Pankaj Kumar**: Dr Yashwant Singh Parmar University of Horticulture and Forestry Nauni Solan Himachal Pradesh, Solan, India; **Pradeep Kumar**: Centre for Cellular & Molecular Biology, Hyderabad, India; **Dhruv Lavania**: University of Alberta, Edmonton, Canada; **Tedrick Thomas Salim Lew**: Institute of Materials Research and Engineering, Singapore, Singapore; **Mathew G Lewsey**: La Trobe University, Bundoora, Australia; **Chien-Yuan Lin**: Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Emeryville, United States; **Dianyi Liu**: Donald Danforth Plant Science Center, Saint Louis, United States; **Le Liu**: Department of Biology, University of Massachusetts Amherst, Amherst, United States; **Tie Liu**: University of Florida, Gainesville, United States; **Ansul Lokdarshi**: University of Tennessee, Knoxville, United States; **Ai My Luong**: Louvain Institute of Biomolecular Science and Technology, Louvain-la-Neuve, Belgium; **Iain C Macaulay**: Earlham Institute, Norwich, United Kingdom; **Sakil Mahmud**: Plant Cell Biology, IZMB, University of Bonn, Bonn, Germany; **Ari Pekka Mähönen**: Institute of Biotechnology, HiLIFE, University of Helsinki, Helsinki, Finland; **Kamal Kumar Malukani**: Molecular Biology, Hyderabad, India; **Alexandre P Marand**: University of Georgia, Athens, United States; **Carly A Martin**: MIT, Cambridge, United States; **Claire D McWhite**: Princeton University, Princeton, United States; **Devang Mehta**: University of Alberta, Edmonton, Canada; **Miguel Miñambres Martín**: Heinrich Heine University, Düsseldorf, Germany; **Jenny C Mortimer**: University of Adelaide, Adelaide, Australia; **Lachezar A Nikolov**: University of California, Los Angeles, Los, Angeles, United States; **Tatsuya Nobori**: Salk Institute for Biological Studies, San, Diego, United States; **Trevor M Nolan**: Duke University, Durham, United States; **Aaron J Ogden**: Pacific Northwest National Laboratory, Richland, United States; **Marisa S Otegui**: University of Wisconsin-Madison, Madison, United States; **Mark-Christoph Ott**: Bayer AG, Monheim am Rhein, Germany; **José M Palma**: Estación Experimental del Zaidín, CSIC, Granada, Spain; **Puneet Paul**: University of Nebraska-Lincoln, Lincoln, United States; **Atique U Rehman**: Department of Agronomy, Bahauddin Zakariya University Multan, Multan, Pakistan; **Maida Romera-Branchat**: Institute of Plant Biology and Biotechnology, University of Münster, Muenster (Germany), Germany; **Luis C Romero**: Instituto de Bioquimica Vegetal y Fotosintesis, CSIC-US, Seville, Spain; **Ronelle Roth**: Department of Plant Sciences, University of Oxford, Oxford, United Kingdom; **Saroj**

K Sah: The Pennsylvania State University, University Park, United States; **Rachel Shahan**: Duke University, Durham, United States; **Shyam Solanki**: Washington State University, Pullman, United States; **Bao-Hua Song**: University of North Carolina at Charlotte, Charlotte, United States; **Rosangela Sozzani**: North Carolina State University, Raleigh, United States; **Gary Stacey**: University of Missouri, Columbia, United States; **Anna N Stepanova**: North Carolina State University, Raleigh, United States; **Nicolas L Taylor**: The University of Western Australia, Perth, Australia; **Marcela K Tello-Ruiz**: Cold Spring Harbor Laboratory, Cold Spring Harbor, United States; **Tuan M Tran**: Nanyang Technological University, Singapore, Singapore; **Rajiv Kumar Tripathi**: University of Manitoba, Winnipeg, Canada; **Batthula Vijaya Lakshmi Vadde Vadde**: Cornell University, Ithaca, United States; **Tamas Varga**: Environmental Molecular Sciences Division, Pacific Northwest National Laboratory, (PNNL), Richland, United States; **Marija Vidovic**: Institute of Molecular Genetics and Genetic Engineering, University of Belgrade, Belgrade, Serbia; **Justin W Walley**: Iowa State University, Ames, United States; **Zhiyong Wang**: Carnegie Institution for Science, Dept. of Plant Biology, Stanford, United States; **Renate A Weizbauer**: Carnegie Institution for Science, Dept. of Plant Biology, Stanford, United States; **James Whelan**: La Trobe University, Melbourne, Australia; **Asela J Wijeratne**: Arkansas State University, Jonesboro, United States; **Tingting Xiang**: University of North Carolina at Charlotte, Charlotte, United States; **Shouling Xu**: Carnegie Institution for Science, Stanford, United States; **Ramin Yadegari**: University of Arizona, Tucson, United States; **Houlin Yu**: Department of Biochemistry and Molecular Biology, University of Massachusetts Amherst, Amherst, United States; **Hai Ying Yuan**: University of Saskatchewan, Saskatoon, Canada; **Fabio Zanini**: UNSW, Sydney, Australia; **Feng Zhao**: Laboratoire Reproduction et Développement des Plantes, Université de Lyon, Lyon, France; **Jie Zhu**: University of California, Davis, Davis, United States; **Xiaohong Zhuang**: Centre for Cell and Developmental Biology, State Key Laboratory of Agrobiotechnology, School of Life Sciences, The Chinese University of Hong Kong, Hong Kong SAR, Shatin, Hong Kong;

Suryatapa Ghosh Jha is in the Department of Plant Biology, Carnegie Institution for Science, Stanford, United States
https://orcid.org/0000-0001-7976-1272

Alexander T Borowsky is in the Department of Botany and Plant Sciences, University of California, Riverside, Riverside, United States
https://orcid.org/0000-0002-2510-8136

Benjamin J Cole is in the Joint Genome Institute, Lawrence Berkeley National Laboratory, Walnut Creek, United States
https://orcid.org/0000-0001-9652-624X

Noah Fahlgren is in the Donald Danforth Plant Science Center, St. Louis, United States
https://orcid.org/0000-0002-5597-4537

Andrew Farmer is in the National Center for Genome Resources, Santa Fe, United States
https://orcid.org/0000-0002-4224-2433

Shao-shan Carol Huang is in the Center for Genomics and Systems Biology, New York University, New York, United States

Purva Karia is in the Department of Plant Biology, Carnegie Institution for Science, Stanford, United States and the Department of Cell and Systems Biology, University of Toronto, Toronto, Canada
https://orcid.org/0000-0002-4263-4867

Marc Libault is in the Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, United States
https://orcid.org/0000-0001-7419-9129

Nicholas J Provart is in the Department of Cell and Systems Biology and the Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Canada
https://orcid.org/0000-0001-5551-7232

Selena L Rice is in the Department of Plant Biology, Carnegie Institution for Science, Stanford, United States
https://orcid.org/0000-0001-8403-5785

Maite Saura-Sanchez is in the Consejo Nacional de Investigaciones Científicas y Técnicas, Instituto de Investigaciones Fisiológicas y Ecológicas Vinculadas a la Agricultura, Facultad de Agronomía, Universidad de Buenos Aires, Buenos Aires, Argentina
https://orcid.org/0000-0001-8115-3540

Pinky Agarwal is in the National Institute of Plant Genome Research, New Delhi, India

Amir H Ahkami is in the Environmental Molecular Sciences Division, Pacific Northwest National Laboratory, Richland, United States
https://orcid.org/0000-0002-0545-1236

Christopher R Anderton is in the Environmental Molecular Sciences Division, Pacific Northwest National Laboratory, Richland, United States

Steven P Briggs is in the Department of Biological Sciences, University of California, San Diego, United States
https://orcid.org/0000-0002-7226-8618

Jennifer AN Brophy is in the Department of Biology, Stanford University, Stanford, United States
https://orcid.org/0000-0001-7808-4281

Peter Denolf is in the BASF Seeds & Traits, Ghent, Belgium
https://orcid.org/0000-0003-4567-2700

Luigi F Di Costanzo is in the Department of Agricultural Sciences, University of Naples Federico II, Napoli, Italy
https://orcid.org/0000-0002-4795-2573

Moises Exposito-Alonso is in the Department of Plant Biology, Carnegie Institution for Science, Stanford, United States
https://orcid.org/0000-0001-5711-0700

Stefania Giacomello is in the SciLifeLab, KTH Royal Institute of Technology, Solna, Sweden
https://orcid.org/0000-0003-0738-1574

Fabio Gomez-Cano is in the Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, United States
https://orcid.org/0000-0002-2624-0112

Kerstin Kaufmann is in the Department for Plant Cell and Molecular Biology, Institute for Biology, Humboldt-Universitaet zu Berlin, Berlin, Germany
https://orcid.org/0000-0001-7960-6256

Dae Kwan Ko is in the Great Lakes Bioenergy Research Center, Michigan State University, East Lansing, United States
https://orcid.org/0000-0002-9720-5138

Sagar Kumar is in the Department of Plant Breeding & Genetics, Mata Gujri College, Fatehgarh Sahib, Punjabi University, Patiala, India
https://orcid.org/0000-0002-0725-0059

Andrey V Malkovskiy is in the Department of Plant Biology, Carnegie Institution for Science, Stanford, United States
https://orcid.org/0000-0002-5648-8602

Naomi Nakayama is in the Department of Bioengineering, Imperial College London, London, United Kingdom

Toshihiro Obata is in the Department of Biochemistry, University of Nebraska-Lincoln, Lincoln, United States
https://orcid.org/0000-0001-8931-7722

Marisa S Otegui is in the Department of Botany, University of Wisconsin-Madison, Madison, United States
https://orcid.org/0000-0003-4699-6950

Gergo Palfalvi is in the Division of Evolutionary Biology, National Institute for Basic Biology, Okazaki, Japan

Elsa H Quezada-Rodríguez is in the Ciencias Agrogenómicas, Escuela Nacional de Estudios Superiores Unidad León, Universidad Nacional Autónoma de México, León, Mexico
https://orcid.org/0000-0001-7789-4987

Rajveer Singh is in the School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, India
https://orcid.org/0000-0002-2922-1371

R Glen Uhrig is in the Department of Science, University of Alberta, Edmonton, Canada
https://orcid.org/0000-0003-2773-4381

Jamie Waese is in the Department of Cell and Systems Biology/Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Canada
https://orcid.org/0000-0001-8783-9914

Klaas Van Wijk is in the School of Integrated Plant Science, Plant Biology Section, Cornell University, Ithaca, United States
https://orcid.org/0000-0001-9536-0487

R Clay Wright is in the Department of Biological Systems Engineering, Virginia Tech, Blacksburg, United States

David W Ehrhardt is in the Department of Plant Biology, Carnegie Institution for Science, Stanford, United States
dehrhardt@carnegiescience.edu

Kenneth D Birnbaum is in the Center for Genomics and Systems Biology, New York University, New York, United States
kdb5@nyu.edu

Seung Y Rhee is in the Department of Plant Biology, Carnegie Institution for Science, Stanford, United States
srhee@carnegiescience.edu
https://orcid.org/0000-0002-7572-4762

Author contributions: Plant Cell Atlas Consortium, Moises Exposito-Alonso, Klaas Van Wijk, Writing - review and editing; Suryatapa Ghosh Jha, Noah Fahlgren, Andrew Farmer, Shao-shan Carol Huang, Nicholas J Provart, Toshihiro Obata, Supervision, Writing - original draft, Writing - review and editing; Alexander T Borowsky, Supervision, Visualization, Writing - original draft, Writing - review and editing; Benjamin J Cole, Pinky Agarwal, Amir H Ahkami, Christopher R Anderton, Steven P Briggs, Jennifer AN Brophy, Peter Denolf, Luigi F Di Costanzo, Stefania Giacomello, Fabio Gomez-Cano, Kerstin Kaufmann, Dae Kwan Ko, Sagar Kumar, Naomi Nakayama, Marisa S Otegui, Gergo Palfalvi, Elsa H Quezada-Rodríguez, Rajveer Singh, R Glen Uhrig, R Clay Wright, Writing - original draft, Writing - review and editing; Purva Karia, Marc Libault, Maite Saura-Sanchez, Andrey V Malkovskiy, Visualization, Writing - original draft, Writing - review and editing; Selena L Rice, Supervision, Writing - original draft, Project administration, Writing - review and editing; Jamie Waese, Visualization, Writing - review and editing; David W Ehrhardt, Kenneth D Birnbaum, Conceptualization, Funding acquisition, Writing - original draft, Writing - review and editing; Seung Y Rhee, Conceptualization, Supervision, Funding acquisition, Writing - original draft, Project administration, Writing - review and editing

Competing interests: The authors declare that no competing interests exist.

Received 25 January 2021
Accepted 26 August 2021
Published 07 September 2021

Funding

Funder	Grant reference number	Author
National Science Foundation	1916797	David W Ehrhardt Kenneth D Birnbaum Seung Yon Rhee
National Science Foundation	2052590	Seung Yon Rhee

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.66877.sa1
Author response https://doi.org/10.7554/eLife.66877.sa2

Additional files

Data availability

No associated datasets.

References

Adamski NM, Borrill P, Brinton J, Harrington SA, Marchal C, Bentley AR, Bovill WD, Cattivelli L, Cockram J, Contreras-Moreira B, Ford B, Ghosh S, Harwood W, Hassani-Pak K, Hayta S, Hickey LT, Kanyuka K, King J, Maccaferri M, Naamati G, et al. 2020. A roadmap for gene functional characterisation in crops with large genomes: lessons from polyploid wheat. *eLife* **9**:e55646. DOI: https://doi.org/10.7554/eLife.55646, PMID: 32208137

Alonso AM, Carrea A, Diambra L. 2019. Prediction of cell position using single-cell transcriptomic data: an iterative procedure. *F1000Research* **8**:1775. DOI: https://doi.org/10.12688/f1000research.20715.1, PMID: 32399185

Avise J. 2004. The hope, hype, and reality of genetic engineering: remarkable stories from agriculture, industry, medicine, and the environment. *Choice* **1**: 0283. DOI: https://doi.org/10.5860/choice.42-0283

Avraham S, Tung C-W, Ilic K, Jaiswal P, Kellogg EA, McCouch S, Pujar A, Reiser L, Rhee SY, Sachs MM, Schaeffer M, Stein L, Stevens P, Vincent L, Zapata F, Ware D. 2008. The plant ontology database: a community resource for plant structure and developmental stages controlled vocabulary and annotations. *Nucleic Acids Research* **36**:D449–D454. DOI: https://doi.org/10.1093/nar/gkm908

Balasubramanian VK, Purvine SO, Liang Y, Kelly RT, Pasa-Tolic L, Chrisler WB, Blumwald E, Stewart CN, Zhu Y, Ahkami AH. 2021. Cell-type-specific proteomics analysis of a small number of plant cells by integrating laser capture microdissection with a nanodroplet sample processing platform. *Current Protocols* **1**:e153. DOI: https://doi.org/10.1002/cpz1.153, PMID: 34043287

Bard J, Rhee SY, Ashburner M. 2005. An ontology for cell types. *Genome Biology* **6**:R21. DOI: https://doi.org/10.1186/gb-2005-6-2-r21, PMID: 15693950

Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau WC, Ledoux P, Rudnev D, Lash AE, Fujibuchi W, Edgar R. 2005. NCBI GEO: mining millions of expression profiles-database and tools. *Nucleic Acids Research* **33**:D562–D566. DOI: https://doi.org/10.1093/nar/gki022, PMID: 15608262

Barrett J, Girr P, Mackinder LCM. 2021. Pyrenoids: CO2-fixing phase separated liquid organelles. *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research* **1868**:118949. DOI: https://doi.org/10.1016/j.bbamcr.2021.118949

Bartuska A. 2017. *USDA Testimony*: U.S. Department of Agriculture. https://www.usda.gov/media/testimony.

Berardini TZ, Reiser L, Li D, Mezheritsky Y, Muller R, Strait E, Huala E. 2015. The Arabidopsis information resource: making and mining the "gold standard" annotated reference plant genome. *Genesis* **53**:474–485. DOI: https://doi.org/10.1002/dvg.22877, PMID: 26201819

Biermann F, Kanie N, Kim RE. 2017. Global governance by goal-setting: the novel approach of the UN sustainable development goals. *Current Opinion in Environmental Sustainability* **26-27**:26–31. DOI: https://doi.org/10.1016/j.cosust.2017.01.010

Boehm CR, Bock R. 2019. Recent advances and current challenges in synthetic biology of the plastid genetic system and metabolism. *Plant Physiology* **179**: 794–802. DOI: https://doi.org/10.1104/pp.18.00767, PMID: 30181342

Bol T, de Vaan M, van de Rijt A. 2018. The Matthew effect in science funding. *PNAS* **115**:4887–4890. DOI: https://doi.org/10.1073/pnas.1719557115, PMID: 29686094

Borges RM. 2008. Plasticity comparisons between plants and animals: concepts and mechanisms. *Plant Signaling & Behavior* **3**:367–375. DOI: https://doi.org/10.4161/psb.3.6.5823, PMID: 19513224

Bouchnak I, van Wijk KJ. 2019. N-Degron pathways in plastids. *Trends in Plant Science* **24**:917–926. DOI: https://doi.org/10.1016/j.tplants.2019.06.013, PMID: 31300194

Braun IR, Lawrence-Dill CJ. 2020. Automated methods enable direct computation on phenotypic descriptions for novel candidate gene prediction. *Frontiers in Plant Science* **10**:1629. DOI: https://doi.org/10.3389/fpls.2019.01629, PMID: 31998331

Brausemann A, Gemmecker S, Koschmieder J, Ghisla S, Beyer P, Einsle O. 2017. Structure of phytoene desaturase provides insights into herbicide binding and reaction mechanisms involved in carotene desaturation. *Structure* **25**:1222–1232. DOI: https://doi.org/10.1016/j.str.2017.06.002, PMID: 28669634

Camp JG, Platt R, Treutlein B. 2019. Mapping human cell phenotypes to genotypes with single-cell genomics. *Science* **365**:1401–1405. DOI: https://doi.org/10.1126/science.aax6648, PMID: 31604266

Costanza R, Fioramonti L, Kubiszewski I. 2016. The UN sustainable development goals and the dynamics of well-being. *Frontiers in Ecology and the Environment* **14**:59. DOI: https://doi.org/10.1002/fee.1231

Decaestecker W, Buono RA, Pfeiffer ML, Vangheluwe N, Jourquin J, Karimi M, Van Isterdael G, Beeckman T, Nowack MK, Jacobs TB. 2019. CRISPR-TSKO: a technique for efficient mutagenesis in specific cell types, tissues, or organs in *Arabidopsis*. *The Plant Cell* **31**:2868–2887. DOI: https://doi.org/10.1105/tpc.19.00454, PMID: 31562216

Deutsch EW, Csordas A, Sun Z, Jarnuczak A, Perez-Riverol Y, Ternent T, Campbell DS, Bernal-Llinares M, Okuda S, Kawano S, Moritz RL, Carver JJ, Wang M, Ishihama Y, Bandeira N, Hermjakob H, Vizcaíno JA. 2017. The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. *Nucleic Acids Research* **45**:D1100–D1106. DOI: https://doi.org/10.1093/nar/gkw936, PMID: 27924013

Diehl AD, Meehan TF, Bradford YM, Brush MH, Dahdul WM, Dougall DS, He Y, Osumi-Sutherland D, Ruttenberg A, Sarntivijai S, Van Slyke CE, Vasilevsky NA, Haendel MA, Blake JA, Mungall CJ. 2016. The cell ontology 2016: enhanced content, modularization, and ontology interoperability. *Journal of Biomedical Semantics* **7**:44. DOI: https://doi.org/10.1186/s13326-016-0088-7, PMID: 27377652

Duran-Nebreda S, Bassel GW. 2017. Bridging scales in plant biology using network science. *Trends in Plant Science* **22**:1001–1003. DOI: https://doi.org/10.1016/j.tplants.2017.09.017, PMID: 29029828

Edwards GE, Franceschi VR, Voznesenskaya EV. 2004. Single-cell C(4) photosynthesis versus the dual-cell (Kranz) paradigm. *Annual Review of Plant Biology* **55**: 173–196. DOI: https://doi.org/10.1146/annurev.arplant.55.031903.141725, PMID: 15377218

ENCODE Project Consortium. 2011. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLOS Biology* **9**:e1001046. DOI: https://doi.org/10.1371/journal.pbio.1001046, PMID: 21526222

Erlinghaeuser M, Hagenau L, Wimmer D, Offermann S. 2016. Development, subcellular positioning and selective protein accumulation in the dimorphic chloroplasts of single-cell C4 species. *Current Opinion in Plant Biology* **31**:76–82. DOI: https://doi.org/10.1016/j.pbi.2016.03.017, PMID: 27061048

Esau K. 1953. Plant anatomy. *Soil Science* **75**:407. DOI: https://doi.org/10.1097/00010694-195305000-00014

Fabregat A, Korninger F, Viteri G, Sidiropoulos K, Marin-Garcia P, Ping P, Wu G, Stein L, D'Eustachio P, Hermjakob H. 2018. Reactome graph database: efficient access to complex pathway data. *PLOS Computational Biology* **14**:e1005968. DOI: https://doi.org/10.1371/journal.pcbi.1005968, PMID: 29377902

Farmer A, Thibivilliers S, Ryu KH, Schiefelbein J, Libault M. 2021. Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in *Arabidopsis* roots at the single-cell level. *Molecular Plant* **14**:372–383. DOI: https://doi.org/10.1016/j.molp.2021.01.001, PMID: 33422696

Giacomello S, Salmén F, Terebieniec BK, Vickovic S, Navarro JF, Alexeyenko A, Reimegård J, McKee LS, Mannapperuma C, Bulone V, Ståhl PL, Sundström JF, Street NR, Lundeberg J. 2017. Spatially resolved transcriptome profiling in model plant species. *Nature Plants* **3**:17061. DOI: https://doi.org/10.1038/nplants.2017.61, PMID: 28481330

Giacomello S. 2021. A new era for plant science: spatial single-cell transcriptomics. *Current Opinion in Plant Biology* **60**:102041. DOI: https://doi.org/10.1016/j.pbi.2021.102041, PMID: 33915520

Gök A, Rigby J, Shapira P. 2016. The impact of research funding on scientific outputs: evidence from six smaller european countries. *Journal of the Association for Information Science and Technology* **67**:715–730. DOI: https://doi.org/10.1002/asi.23406

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS. 2012. Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* **40**: D1178–D1186. DOI: https://doi.org/10.1093/nar/gkr944

Goring DR, Di Sansebastiano GP. 2018. Protein and membrane trafficking routes in plants: conventional or unconventional? *Journal of Experimental Botany* **69**:1–5. DOI: https://doi.org/10.1093/jxb/erx435

Griggs D, Stafford-Smith M, Gaffney O, Rockström J, Ohman MC, Shyamsundar P, Steffen W, Glaser G, Kanie N, Noble I. 2013. Policy: sustainable development goals for people and planet. *Nature* **495**: 305–307. DOI: https://doi.org/10.1038/495305a, PMID: 23518546

Han X, Kumar D, Chen H, Wu S, Kim JY. 2014. Transcription factor-mediated cell-to-cell signalling in plants. *Journal of Experimental Botany* **65**:1737–1749. DOI: https://doi.org/10.1093/jxb/ert422, PMID: 24347464

Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, Saadatpour A, Zhou Z, Chen H, Ye F, Huang D, Xu Y, Huang W, Jiang M, Jiang X, Mao J, Chen Y, Lu C, Xie J, Fang Q, et al. 2018. Mapping the mouse cell atlas by Microwell-Seq. *Cell* **172**:1091–1107. DOI: https://doi.org/10.1016/j.cell.2018.02.001

Han S, Maberly SC, Gontero B, Xing Z, Li W, Jiang H, Huang W. 2020. Structural basis for C4 photosynthesis without Kranz anatomy in leaves of the submerged freshwater plant *Ottelia alismoides*. *Annals of Botany* **125**:869–879. DOI: https://doi.org/10.1093/aob/mcaa005, PMID: 31942934

Harrison MJ, Dewbre GR, Liu J. 2002. A phosphate transporter from *Medicago truncatula* involved in the acquisition of phosphate released by arbuscular mycorrhizal fungi. *The Plant Cell* **14**:2413–2429. DOI: https://doi.org/10.1105/tpc.004861, PMID: 12368495

Heinze L, Freimuth N, Rößling AK, Hahnke R, Riebschläger S, Fröhlich A, Sampathkumar A, McFarlane HE, Sauer M. 2020. EPSIN1 and MTV1 define functionally overlapping but molecularly distinct trans-Golgi network subdomains in *Arabidopsis*. *PNAS* **117**:25880–25889. DOI: https://doi.org/10.1073/pnas.2004822117, PMID: 32989160

Henderson SW, Gilliham M. 2015. The "gatekeeper" concept: cell-type specific molecular mechanisms of plant adaptation to abiotic stress. In: In: Laitinen R. A. E (Ed). *Molecular Mechanisms in Plant Adaptation*. John Wiley & Sons, Inc. p. 83–115. DOI: https://doi.org/10.1002/9781118860526.ch4

Henkhaus N, Bartlett M, Gang D, Grumet R, Jordon-Thaden I, Lorence A, Lyons E, Miller S, Murray S, Nelson A, Specht C, Tyler B, Wentworth T, Ackerly D, Baltensperger D, Benfey P, Birchler J, Chellamma S, Crowder R, Donoghue M, et al. 2020. Plant science decadal vision 2020-2030: reimagining the potential of plants for a healthy and sustainable future. *Plant Direct* **4**:e00252. DOI: https://doi.org/10.1002/pld3.252, PMID: 32904806

Hie B, Bryson B, Berger B. 2019. Efficient integration of heterogeneous single-cell transcriptomes using scanorama. *Nature Biotechnology* **37**:685–691. DOI: https://doi.org/10.1038/s41587-019-0113-3, PMID: 31061482

Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD. 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**:804–807. DOI: https://doi.org/10.1126/science.1137016, PMID: 17289988

Howe KL, Contreras-Moreira B, De Silva N, Maslen G, Akanni W, Allen J, Alvarez-Jarreta J, Barba M, Bolser DM, Cambell L, Carbajo M, Chakiachvili M, Christensen M, Cummins C, Cuzick A, Davis P, Fexova

S, Gall A, George N, Gil L, et al. 2020. Ensembl genomes 2020-enabling non-vertebrate genomic research. *Nucleic Acids Research* **48**:D689–D695. DOI: https://doi.org/10.1093/nar/gkz890, PMID: 31598706

Hu AGZ. 2020. Public funding and the ascent of Chinese science: evidence from the national natural science foundation of China. *Research Policy* **49**: 103983. DOI: https://doi.org/10.1016/j.respol.2020.103983

Huang A, Tang Y, Shi X, Jia M, Zhu J, Yan X, Chen H, Gu Y. 2020. Proximity labeling proteomics reveals critical regulators for inner nuclear membrane protein degradation in plants. *Nature Communications* **11**: 3284. DOI: https://doi.org/10.1038/s41467-020-16744-1, PMID: 32601292

Huang L, Schiefelbein J. 2015. Conserved gene expression programs in developing roots from diverse plants. *The Plant Cell* **27**:2119–2132. DOI: https://doi.org/10.1105/tpc.15.00328, PMID: 26265761

Ilic K, Kellogg EA, Jaiswal P, Zapata F, Stevens PF, Vincent LP, Avraham S, Reiser L, Pujar A, Sachs MM, Whitman NT, McCouch SR, Schaeffer ML, Ware DH, Stein LD, Rhee SY. 2007. The plant structure ontology, a unified vocabulary of anatomy and morphology of a flowering plant. *Plant Physiology* **143**:587–599. DOI: https://doi.org/10.1104/pp.106.092825, PMID: 17142475

Irieda H, Takano Y. 2021. Epidermal chloroplasts are defense-related motile organelles equipped with plant immune components. *Nature Communications* **12**: 2739. DOI: https://doi.org/10.1038/s41467-021-22977-5, PMID: 34016974

Isoda R, Yoshinari A, Ishikawa Y, Sadoine M, Simon R, Frommer WB, Nakamura M. 2021. Sensors for the quantification, localization and analysis of the dynamics of plant hormones. *The Plant Journal* **105**:542–557. DOI: https://doi.org/10.1111/tpj.15096, PMID: 33231903

Jaffe AB, Gush J, Larsen V, Laws A. 2015. The effect of public funding on research output. *The New Zealand Marsden Fund* **52**:1–22. DOI: https://doi.org/10.1080/00779954.2017.1325921

Jarsch IK, Konrad SS, Stratil TF, Urbanus SL, Szymanski W, Braun P, Braun KH, Ott T. 2014. Plasma membranes are subcompartmentalized into a plethora of coexisting and diverse microdomains in *Arabidopsis* and *Nicotiana benthamiana*. *The Plant Cell* **26**:1698–1711. DOI: https://doi.org/10.1105/tpc.114.124446, PMID: 24714763

Jensen PE, Scharff LB. 2019. Engineering of plastids to optimize the production of high-value metabolites and proteins. *Current Opinion in Biotechnology* **59**:8–15. DOI: https://doi.org/10.1016/j.copbio.2019.01.009, PMID: 30798145

Jiang X, Hoehenwarter W, Scheel D, Lee J. 2020. Phosphorylation of the CAMTA3 transcription factor triggers its destabilization and nuclear export. *Plant Physiology* **184**:1056–1071. DOI: https://doi.org/10.1104/pp.20.00795, PMID: 32769161

Jiang J, Dehesh K. 2021. Plastidial retrograde modulation of light and hormonal signaling: an odyssey. *New Phytologist* **230**:931–937. DOI: https://doi.org/10.1111/nph.17192, PMID: 33452833

Jones JD, Dangl JL. 2006. The plant immune system. *Nature* **444**:323–329. DOI: https://doi.org/10.1038/nature05286, PMID: 17108957

Jorasch P. 2020. Will the EU stay out of step with science and the rest of the world on plant breeding innovation? *Plant Cell Reports* **39**:163–167. DOI: https://doi.org/10.1007/s00299-019-02482-2, PMID: 31754780

Kajala K, Gouran M, Shaar-Moshe L, Mason GA, Rodriguez-Medina J, Kawa D, Pauluzzi G, Reynoso M, Canto-Pastor A, Manzano C, Lau V, Artur MAS, West DA, Gray SB, Borowsky AT, Moore BP, Yao AI, Morimoto KW, Bajic M, Formentin E, et al. 2021. Innovation, conservation, and repurposing of gene function in root cell type development. *Cell* **184**:3333–3348. DOI: https://doi.org/10.1016/j.cell.2021.04.024, PMID: 34010619

Kaminuma E, Mashima J, Kodama Y, Gojobori T, Ogasawara O, Okubo K, Takagi T, Nakamura Y. 2010. DDBJ launches a new archive database with analytical tools for next-generation sequence data. *Nucleic Acids Research* **38**:D33–D38. DOI: https://doi.org/10.1093/nar/gkp847, PMID: 19850725

Kanai M, Mano S, Kondo M, Hayashi M, Nishimura M. 2016. Extension of oil biosynthesis during the mid-phase of seed development enhances oil content in *Arabidopsis* seeds . *Plant Biotechnology Journal* **14**: 1241–1250. DOI: https://doi.org/10.1111/pbi.12489

Kang JB, Nathan A, Zhang F, Millard N, Rumker L, Branch Moody D, Korsunsky I, Raychaudhuri S. 2021. Efficient and precise single-cell reference atlas mapping with symphony. *bioRxiv*. DOI: https://doi.org/10.1101/2020.11.18.389189

Katari MS, Nowicki SD, Aceituno FF, Nero D, Kelfer J, Thompson LP, Cabello JM, Davidson RS, Goldberg AP, Shasha DE, Coruzzi GM, Gutiérrez RA. 2010. VirtualPlant: a software platform to support systems biology research. *Plant Physiology* **152**:500–515. DOI: https://doi.org/10.1104/pp.109.147025, PMID: 20007449

Khan B, Robbins C, Okrent A. 2020. The state of US. science and engineering. https://ncses.nsf.gov/pubs/nsb20201/u-s-r-d-performance-and-funding [Accessed September 18, 2020].

Kim T-W, Park CH, Hsu C-C, Zhu J-Y, Hsiao Y, Branon T, Xu S-L, Ting AY, Wang Z-Y. 2019. Application of TurboID-mediated proximity labeling for mapping a GSK3 kinase signaling network in *Arabidopsis*. *bioRxiv*. DOI: https://doi.org/10.1101/636324

Kim SC, Guo L, Wang X. 2020. Nuclear moonlighting of cytosolic glyceraldehyde-3-phosphate dehydrogenase regulates *Arabidopsis* response to heat stress. *Nature Communications* **11**:3439. DOI: https://doi.org/10.1038/s41467-020-17311-4, PMID: 32651385

Kong SG, Wada M. 2014. Recent advances in understanding the molecular mechanism of chloroplast photorelocation movement. *Biochimica Et Biophysica Acta (BBA) - Bioenergetics* **1837**:522–530. DOI: https://doi.org/10.1016/j.bbabio.2013.12.004, PMID: 24333784

Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh PR, Raychaudhuri S. 2019. Fast, sensitive and accurate integration of single-cell data with harmony. *Nature Methods* **16**: 1289–1296. DOI: https://doi.org/10.1038/s41592-019-0619-0, PMID: 31740819

Kost B. 2001. Towards a virtual *Arabidopsis* plant. *Genome Biology* **2**:REPORTS4019. DOI: https://doi.

org/10.1186/gb-2001-2-8-reports4019, PMID: 11532210
Kuhnert F, Stefanski A, Overbeck N, Drews L, Reichert AS, Stühler K, Weber APM. 2020. Rapid single-step affinity purification of HA-tagged plant mitochondria. *Plant Physiology* **182**:692–706. DOI: https://doi.org/10.1104/pp.19.00732, PMID: 31818904
Lähnemann D, Köster J, Szczurek E, McCarthy DJ, Hicks SC, Robinson MD, Vallejos CA, Campbell KR, Beerenwinkel N, Mahfouz A, Pinello L, Skums P, Stamatakis A, Attolini CS, Aparicio S, Baaijens J, Balvert M, Barbanson B, Cappuccio A, Corleone G, et al. 2020. Eleven grand challenges in single-cell data science. *Genome Biology* **21**:31. DOI: https://doi.org/10.1186/s13059-020-1926-6, PMID: 32033589
Larkum AWD, Davey PA, Kuo J, Ralph PJ, Raven JA. 2017. Carbon-concentrating mechanisms in seagrasses. *Journal of Experimental Botany* **68**:3773–3784. DOI: https://doi.org/10.1093/jxb/erx206, PMID: 28911056
Leinonen R, Akhtar R, Birney E, Bower L, Cerdeno-Tárraga A, Cheng Y, Cleland I, Faruque N, Goodgame N, Gibson R, Hoad G, Jang M, Pakseresht N, Plaister S, Radhakrishnan R, Reddy K, Sobhany S, Ten Hoopen P, Vaughan R, Zalunin V, et al. 2011. The European Nucleotide Archive. *Nucleic Acids Research* **39**:D28–D31. DOI: https://doi.org/10.1093/nar/gkq967, PMID: 20972220
Libault M, Brechenmacher L, Cheng J, Xu D, Stacey G. 2010. Root hair systems biology. *Trends in Plant Science* **15**:641–650. DOI: https://doi.org/10.1016/j.tplants.2010.08.010, PMID: 20851035
Libault M, Pingault L, Zogli P, Schiefelbein J. 2017. Plant systems biology at the single-cell level. *Trends in Plant Science* **22**:949–960. DOI: https://doi.org/10.1016/j.tplants.2017.08.006, PMID: 28970001
Llorente B, Torres-Montilla S, Morelli L, Florez-Sarasa I, Matus JT, Ezquerro M, D'Andrea L, Houhou F, Majer E, Picó B, Cebolla J, Troncoso A, Fernie AR, Daròs JA, Rodriguez-Concepcion M. 2020. Synthetic conversion of leaf chloroplasts into carotenoid-rich plastids reveals mechanistic basis of natural chromoplast development. *PNAS* **117**:21796–21803. DOI: https://doi.org/10.1073/pnas.2004405117, PMID: 32817419
Lotfollahi M, Naghipourfar M, Luecken MD, Khajavi M, Büttner M, Avsec Z, Misharin AV, Theis FJ. 2020. Query to reference single-cell integration with transfer learning. *bioRxiv*. DOI: https://doi.org/10.1101/2020.07.16.205997
Ma A, McDermaid A, Xu J, Chang Y, Ma Q. 2020. Integrative methods and practical challenges for single-cell multi-omics. *Trends in Biotechnology* **38**:1007–1022. DOI: https://doi.org/10.1016/j.tibtech.2020.02.013, PMID: 32818441
Mair A, Xu SL, Branon TC, Ting AY, Bergmann DC. 2019. Proximity labeling of protein complexes and cell-type-specific organellar proteomes in *Arabidopsis* enabled by TurboID. *eLife* **8**:e47864. DOI: https://doi.org/10.7554/eLife.47864, PMID: 31535972
Mallis CS, Zheng X, Qiu X, McCabe JW, Shirzadeh M, Lyu J, Laganowsky A, Russell DH. 2020. Development of native MS capabilities on an extended mass range Q-TOF MS. *International Journal of Mass Spectrometry* **458**:116451. DOI: https://doi.org/10.1016/j.ijms.2020.116451, PMID: 33162786
Marand AP, Chen Z, Gallavotti A, Schmitz RJ. 2021. A cis-regulatory atlas in maize at single-cell resolution. *Cell* **184**:3041–3055. DOI: https://doi.org/10.1016/j.cell.2021.04.014, PMID: 33964211
Marchive C, Roudier F, Castaings L, Bréhaut V, Blondet E, Colot V, Meyer C, Krapp A. 2013. Nuclear retention of the transcription factor NLP7 orchestrates the early response to nitrate in plants. *Nature Communications* **4**:1713. DOI: https://doi.org/10.1038/ncomms2650, PMID: 23591880
McFaline-Figueroa JL, Trapnell C, Cuperus JT. 2020. The promise of single-cell genomics in plants. *Current Opinion in Plant Biology* **54**:114–121. DOI: https://doi.org/10.1016/j.pbi.2020.04.002, PMID: 32388018
Michaud O, Fiorucci AS, Xenarios I, Fankhauser C. 2017. Local auxin production underlies a spatially restricted neighbor-detection response in *Arabidopsis*. *PNAS* **114**:7444–7449. DOI: https://doi.org/10.1073/pnas.1702276114, PMID: 28652343
Morton JT, Strauss CEM, Blackwell R, Berenberg D, Gligorijevic V, Bonneau R. 2020. Protein structural alignments from sequence. *bioRxiv*. DOI: https://doi.org/10.1101/2020.11.03.365932
Muñoz P, Munné-Bosch S. 2020. Oxylipins in plastidial retrograde signaling. *Redox Biology* **37**:101717. DOI: https://doi.org/10.1016/j.redox.2020.101717, PMID: 32979794
Nakane T, Kotecha A, Sente A, McMullan G, Masiulis S, Brown P, Grigoras IT, Malinauskaite L, Malinauskas T, Miehling J, Uchański T, Yu L, Karia D, Pechnikova EV, de Jong E, Keizer J, Bischoff M, McCormack J, Tiemeijer P, Hardwick SW, et al. 2020. Single-particle cryo-EM at atomic resolution. *Nature* **587**:152–156. DOI: https://doi.org/10.1038/s41586-020-2829-0, PMID: 33087931
NCBI Resource Coordinators. 2018. Database resources of the national center for biotechnology information. *Nucleic Acids Research* **46**:D8–D13. DOI: https://doi.org/10.1093/nar/gkx1095, PMID: 29140470
Ngou BPM, Ahn HK, Ding P, Jones JDG. 2021. Mutual potentiation of plant immunity by cell-surface and intracellular receptors. *Nature* **592**:110–115. DOI: https://doi.org/10.1038/s41586-021-03315-7, PMID: 33692545
Niehaus M, Straube H, Künzler P, Rugen N, Hegermann J, Giavalisco P, Eubel H, Witte CP, Herde M. 2020. Rapid affinity purification of tagged plant mitochondria (Mito-AP) for metabolome and proteome analyses. *Plant Physiology* **182**:1194–1210. DOI: https://doi.org/10.1104/pp.19.00736, PMID: 31911558
Obata T. 2019. Metabolons in plant primary and secondary metabolism. *Phytochemistry Reviews* **18**:1483–1507. DOI: https://doi.org/10.1007/s11101-019-09619-x
Obayashi T, Kinoshita K, Nakai K, Shibaoka M, Hayashi S, Saeki M, Shibata D, Saito K, Ohta H. 2007. ATTED-II: a database of co-expressed genes and cis elements for identifying co-regulated gene groups in *Arabidopsis*. *Nucleic Acids Research* **35**:D863–D869. DOI: https://doi.org/10.1093/nar/gkl783, PMID: 17130150
Okumoto S, Versaw W. 2017. Genetically encoded sensors for monitoring the transport and concentration of nitrogen-containing and phosphorus-containing molecules in plants. *Current Opinion in Plant Biology* **39**:129–135. DOI: https://doi.org/10.1016/j.pbi.2017.07.004, PMID: 28750256

Palla G, Spitzer H, Klein M, Fischer D, Schaar AC, Kuemmerle LB, Rybakov S, Ibarra IL, Holmberg O, Virshup I, Lotfollahi M, Richter S, Theis FJ. 2021. Squidpy: a scalable framework for spatial single cell analysis. *bioRxiv*. DOI: https://doi.org/10.1101/2021.02.19.431994

Papatheodorou I, Moreno P, Manning J, Fuentes AM, George N, Fexova S, Fonseca NA, Füllgrabe A, Green M, Huang N, Huerta L, Iqbal H, Jianu M, Mohammed S, Zhao L, Jarnuczak AF, Jupp S, Marioni J, Meyer K, Petryszak R, et al. 2020. Expression atlas update: from tissues to single cells. *Nucleic Acids Research* **48**:D77–D83. DOI: https://doi.org/10.1093/nar/gkz947, PMID: 31665515

Papoutsoglou EA, Faria D, Arend D, Arnaud E, Athanasiadis IN, Chaves I, Coppens F, Cornut G, Costa BV, Ćwiek-Kupczyńska H, Droesbeke B, Finkers R, Gruden K, Junker A, King GJ, Krajewski P, Lange M, Laporte MA, Michotey C, Oppermann M, et al. 2020. Enabling reusability of plant phenomic datasets with MIAPPE 1.1. *New Phytologist* **227**:260–273. DOI: https://doi.org/10.1111/nph.16544, PMID: 32171029

Pastor-Cantizano N, Ko DK, Angelos E, Pu Y, Brandizzi F. 2020. Functional diversification of ER stress responses in *Arabidopsis*. *Trends in Biochemical Sciences* **45**:123–136. DOI: https://doi.org/10.1016/j.tibs.2019.10.008, PMID: 31753702

Pesaresi P, Kim C. 2019. Current understanding of GUN1: a key mediator involved in biogenic retrograde signaling. *Plant Cell Reports* **38**:819–823. DOI: https://doi.org/10.1007/s00299-019-02383-4, PMID: 30671650

Pham VVH, Li X, Truong B, Nguyen T, Liu L, Li J, Le TD. 2020. The winning methods for predicting cellular position in the DREAM single cell transcriptomics challenge. *bioRxiv*. DOI: https://doi.org/10.1101/2020.05.09.086397

Pinard D, Mizrachi E. 2018. Unsung and understudied: plastids involved in secondary growth. *Current Opinion in Plant Biology* **42**:30–36. DOI: https://doi.org/10.1016/j.pbi.2018.01.011, PMID: 29459221

Portwood JL. 2018. The maize multi-genome genetics and genomics database. *Nucleic Acids Research* **1**:D1146–D1154. DOI: https://doi.org/10.1093/nar/gky1046

Radivojević T, Costello Z, Workman K, Garcia Martin H. 2020. A machine learning automated recommendation tool for synthetic biology. *Nature Communications* **11**:4879. DOI: https://doi.org/10.1038/s41467-020-18008-4, PMID: 32978379

Rajendran V, Kalita P, Shukla H, Kumar A, Tripathi T. 2018. Aminoacyl-tRNA synthetases: structure, function, and drug discovery. *International Journal of Biological Macromolecules* **111**:400–414. DOI: https://doi.org/10.1016/j.ijbiomac.2017.12.157, PMID: 29305884

Ramon M, Dang TVT, Broeckx T, Hulsmans S, Crepin N, Sheen J, Rolland F. 2019. Default activation and nuclear translocation of the plant cellular energy sensor SnRK1 regulate metabolic stress responses and development. *The Plant Cell* **31**:1614–1632. DOI: https://doi.org/10.1105/tpc.18.00500, PMID: 31123051

Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, Bodenmiller B, Campbell P, Carninci P, Clatworthy M, Clevers H, Deplancke B, Dunham I, Eberwine J, Eils R, Enard W, Farmer A, Fugger L, Göttgens B, Hacohen N, et al. 2017. The human cell atlas. *eLife* **6**:e27041. DOI: https://doi.org/10.7554/eLife.27041, PMID: 29206104

Reid WV, Ali MK, Field CB. 2020. The future of bioenergy. *Global Change Biology* **26**:274–286. DOI: https://doi.org/10.1111/gcb.14883, PMID: 31642554

Rice S, Fryer E, Ghosh Jha S, Malkovskiy A, Meyer H, Thomas J, Weizbauer R, Zhao K, Birnbaum K, Ehrhardt D, Wang Z, Rhee SY, Plant Cell Atlas Consortium. 2020. First plant cell atlas workshop report. *Plant Direct* **4**:e00271. DOI: https://doi.org/10.1002/pld3.271, PMID: 33083684

Rolland N, Bouchnak I, Moyet L, Salvi D, Kuntz M. 2018. The main functions of plastids. *Methods in Molecular Biology* **1829**:73–85. DOI: https://doi.org/10.1007/978-1-4939-8654-5_5, PMID: 29987715

Rood JE, Stuart T, Ghazanfar S, Biancalani T, Fisher E, Butler A, Hupalowska A, Gaffney L, Mauck W, Eraslan G, Marioni JC, Regev A, Satija R. 2019. Toward a common coordinate framework for the human body. *Cell* **179**:1455–1467. DOI: https://doi.org/10.1016/j.cell.2019.11.019, PMID: 31835027

Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, Graybuck LT, Peeler DJ, Mukherjee S, Chen W, Pun SH, Sellers DL, Tasic B, Seelig G. 2018. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**:176–182. DOI: https://doi.org/10.1126/science.aam8999

Rosenbloom JL, Ginther DK, Juhl T, Heppert JA. 2015. The effects of research and development funding on scientific productivity: academic chemistry, 1990-2009. *PLOS ONE* **10**:e0138176. DOI: https://doi.org/10.1371/journal.pone.0138176

Ruffell D. 2018. The EU court of justice extends the GMO directive to gene-edited organisms. *FEBS Letters* **592**:3653–3657. DOI: https://doi.org/10.1002/1873-3468.13293, PMID: 30426476

Sansone S-A, Rocca-Serra P, Field D, Maguire E, Taylor C, Hofmann O, Fang H, Neumann S, Tong W, Amaral-Zettler L, Begley K, Booth T, Bougueleret L, Burns G, Chapman B, Clark T, Coleman L-A, Copeland J, Das S, de Daruvar A, et al. 2012. Toward interoperable bioscience data. *Nature Genetics* **44**:121–126. DOI: https://doi.org/10.1038/ng.1054

Sarma GP, Lee CW, Portegys T, Ghayoomie V, Jacobs T, Alicea B, Cantarelli M, Currie M, Gerkin RC, Gingell S, Gleeson P, Gordon R, Hasani RM, Idili G, Khayrulin S, Lung D, Palyanov A, Watts M, Larson SD. 2018. OpenWorm: overview and recent advances in integrative biological simulation of *Caenorhabditis elegans*. *Philosophical Transactions of the Royal Society B: Biological Sciences* **373**:20170382. DOI: https://doi.org/10.1098/rstb.2017.0382

Saunders A, Macosko EZ, Wysoker A, Goldman M, Krienen FM, de Rivera H, Bien E, Baum M, Bortolin L, Wang S, Goeva A, Nemesh J, Kamitaki N, Brumbaugh S, Kulp D, McCarroll SA. 2018. Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**:1015–1030. DOI: https://doi.org/10.1016/j.cell.2018.07.028, PMID: 30096299

Schlüter U, Weber APM. 2020. Regulation and evolution of C_4 photosynthesis. *Annual Review of Plant Biology* **71**:183–215. DOI: https://doi.org/10.1146/annurev-arplant-042916-040915, PMID: 32131603

Seren Ü, Grimm D, Fitz J, Weigel D, Nordborg M, Borgwardt K, Korte A. 2016. AraPheno: a public database for *Arabidopsis thaliana* phenotypes. *Nucleic Acids Research* **1**:D1054–D1059. DOI: https://doi.org/10.1093/nar/gkw986
Shahid N, Daniell H. 2016. Plant-based oral vaccines against zoonotic and non-zoonotic diseases. *Plant Biotechnology Journal* **14**:2079–2099. DOI: https://doi.org/10.1111/pbi.12604, PMID: 27442628
Shang J, Ye G, Shi K, Wan Y, Luo C, Aihara H, Geng Q, Auerbach A, Li F. 2020. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **581**:221–224. DOI: https://doi.org/10.1038/s41586-020-2179-y, PMID: 32225175
Shao W, Dong J. 2016. Polarity in plant asymmetric cell division: division orientation and cell fate differentiation. *Developmental Biology* **419**:121–131. DOI: https://doi.org/10.1016/j.ydbio.2016.07.020, PMID: 27475487
Shepon A, Eshel G, Noor E, Milo R. 2018. The opportunity cost of animal based diets exceeds all food losses. *PNAS* **115**:3804–3809. DOI: https://doi.org/10.1073/pnas.1713820115, PMID: 29581251
Shimizu Y, Takagi J, Ito E, Ito Y, Ebine K, Komatsu Y, Goto Y, Sato M, Toyooka K, Ueda T, Kurokawa K, Uemura T, Nakano A. 2021. Cargo sorting zones in the trans-Golgi network visualized by super-resolution confocal live imaging microscopy in plants. *Nature Communications* **12**:1901. DOI: https://doi.org/10.1038/s41467-021-22267-0, PMID: 33772008
Sinnott EW. 1960. Plant morphogenesis. *Nature* **189**:697–699. DOI: https://doi.org/10.1038/189697a0
Smirnoff N, Arnaud D. 2019. Hydrogen peroxide metabolism and functions in plants. *New Phytologist* **221**:1197–1214. DOI: https://doi.org/10.1111/nph.15488
Sofroniew N, Lambert T, Evans K, Nunez-Iglesias J, Winston P, Bokota G, Yamauchi K, Solak AC, Peña-Castellanos G, ziyangczi BM, Buckley G, Pop DD. 2021. Napari. *Zenodo*. 0.4.8rc6. https://doi.org/10.5281/zenodo.3555620
Solymosi K, Lethin J, Aronsson H. 2018. Diversity and plasticity of plastids in land plants. *Methods in Molecular Biology* **1829**:55–72. DOI: https://doi.org/10.1007/978-1-4939-8654-5_4, PMID: 29987714
Stewart BJ, Ferdinand JR, Young MD, Mitchell TJ, Loudon KW, Riding AM, Richoz N, Frazer GL, Staniforth JUL, Vieira Braga FA, Botting RA, Popescu DM, Vento-Tormo R, Stephenson E, Cagan A, Farndon SJ, Polanski K, Efremova M, Green K, Del Castillo Velasco-Herrera M, et al. 2019. Spatiotemporal immune zonation of the human kidney. *Science* **365**:1461–1466. DOI: https://doi.org/10.1126/science.aat5031, PMID: 31604275
Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R. 2019. Comprehensive integration of single-cell data. *Cell* **177**:1888–1902. DOI: https://doi.org/10.1016/j.cell.2019.05.031, PMID: 31178118
Sullivan A, Purohit PK, Freese NH, Pasha A, Esteban E, Waese J, Wu A, Chen M, Chin CY, Song R, Watharkar SR, Chan AP, Krishnakumar V, Vaughn MW, Town C, Loraine AE, Provart NJ. 2019. An 'eFP-Seq Browser' for visualizing and exploring RNA sequencing data. *The Plant Journal : For Cell and Molecular Biology* **100**:641–654. DOI: https://doi.org/10.1111/tpj.14468, PMID: 31350781
Szigeti B, Gleeson P, Vella M, Khayrulin S, Palyanov A, Hokanson J, Currie M, Cantarelli M, Idili G, Larson S. 2014. OpenWorm: an open-science approach to modeling *Caenorhabditis elegans*. *Frontiers in Computational Neuroscience* **8**:137. DOI: https://doi.org/10.3389/fncom.2014.00137, PMID: 25404913
Tabula Muris Consortium. 2018. Single-cell transcriptomics of 20 mouse organs creates a *Tabula muris*. *Nature* **562**:367–372. DOI: https://doi.org/10.1038/s41586-018-0590-4, PMID: 30283141
Tello-Ruiz MK, Naithani S, Gupta P, Olson A, Wei S, Preece J, Jiao Y, Wang B, Chougule K, Garg P, Elser J, Kumari S, Kumar V, Contreras-Moreira B, Naamati G, George N, Cook J, Bolser D, D'Eustachio P, Stein LD, et al. 2021. Gramene 2021: harnessing the power of comparative genomics and pathways for plant research. *Nucleic Acids Research* **49**:D1452–D1463. DOI: https://doi.org/10.1093/nar/gkaa979, PMID: 33170273
ten Hove CA, Lu KJ, Weijers D. 2015. Building a plant: cell fate specification in the early *Arabidopsis* embryo. *Development* **142**:420–430. DOI: https://doi.org/10.1242/dev.111500, PMID: 25605778
The Gene Ontology Consortium. 2019. The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Research* **47**:D330–D338. DOI: https://doi.org/10.1093/nar/gky1055, PMID: 30395331
Togninalli M, Seren Ü, Meng D, Fitz J, Nordborg M, Weigel D, Borgwardt K, Korte A, Grimm DG. 2018. The AraGWAS catalog: a curated and standardized *Arabidopsis thaliana* GWAS catalog. *Nucleic Acids Research* **46**:D1150–D1156. DOI: https://doi.org/10.1093/nar/gkx954, PMID: 29059333
Togninalli M, Seren Ü, Freudenthal JA, Monroe JG, Meng D, Nordborg M, Weigel D, Borgwardt K, Korte A, Grimm DG. 2020. AraPheno and the AraGWAS catalog 2020: a major database update including RNA-Seq and knockout mutation data for *Arabidopsis thaliana*. *Nucleic Acids Research* **48**:D1063–D1068. DOI: https://doi.org/10.1093/nar/gkz925, PMID: 31642487
Tylewicz S, Petterle A, Marttila S, Miskolczi P, Azeez A, Singh RK, Immanen J, Mähler N, Hvidsten TR, Eklund DM, Bowman JL, Helariutta Y, Bhalerao RP. 2018. Photoperiodic control of seasonal growth is mediated by ABA acting on cell-cell communication. *Science* **360**:212–215. DOI: https://doi.org/10.1126/science.aan8576, PMID: 29519919
Uhrig RG, Schläpfer P, Roschitzki B, Hirsch-Hoffmann M, Gruissem W. 2019. Diurnal changes in concerted plant protein phosphorylation and acetylation in *Arabidopsis* organs and seedlings. *The Plant Journal* **99**:176–194. DOI: https://doi.org/10.1111/tpj.14315, PMID: 30920011
Voon CP, Guan X, Sun Y, Sahu A, Chan MN, Gardeström P, Wagner S, Fuchs P, Nietzel T, Versaw WK, Schwarzländer M, Lim BL. 2018. ATP compartmentation in plastids and cytosol of *Arabidopsis thaliana* revealed by fluorescent protein sensing. *PNAS* **115**:E10778–E10787. DOI: https://doi.org/10.1073/pnas.1711497115, PMID: 30352850
Waese J, Fan J, Pasha A, Yu H, Fucile G, Shi R, Cumming M, Kelley LA, Sternberg MJ, Krishnakumar V, Ferlanti E, Miller J, Town C, Stuerzlinger W, Provart NJ. 2017. ePlant: visualizing and exploring multiple levels of data for hypothesis generation in plant

biology. *The Plant Cell* **29**:1806–1821. DOI: https://doi.org/10.1105/tpc.17.00073, PMID: 28808136

Wang E, Wang J, Zhu X, Hao W, Wang L, Li Q, Zhang L, He W, Lu B, Lin H, Ma H, Zhang G, He Z. 2008. Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nature Genetics* **40**:1370–1374. DOI: https://doi.org/10.1038/ng.220, PMID: 18820698

Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J, Kapono CA, Luzzatto-Knaan T, Porto C, Bouslimani A, Melnik AV, Meehan MJ, Liu WT, Crüsemann M, Boudreau PD, Esquenazi E, Sandoval-Calderón M, Kersten RD, et al. 2016. Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nature Biotechnology* **34**:828–837. DOI: https://doi.org/10.1038/nbt.3597, PMID: 27504778

Wang P, Tang W, Li Z, Zou Z, Zhou Y, Li R, Xiong T, Wang J, Zou P. 2019. Mapping spatial transcriptome with light-activated proximity-dependent RNA labeling. *Nature Chemical Biology* **15**:1110–1119. DOI: https://doi.org/10.1038/s41589-019-0368-5, PMID: 31591565

Wang X, Ye L, Lyu M, Ursache R, Löytynoja A, Mähönen AP. 2020. An inducible genome editing system for plants. *Nature Plants* **6**:766–772. DOI: https://doi.org/10.1038/s41477-020-0695-2, PMID: 32601420

Wang B, Qiu YL. 2006. Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* **16**: 299–363. DOI: https://doi.org/10.1007/s00572-005-0033-6, PMID: 16845554

Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, et al. 2016. The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* **3**: 160018. DOI: https://doi.org/10.1038/sdata.2016.18, PMID: 26978244

wwPDB consortium. 2019. Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Research* **47**:D520–D528. DOI: https://doi.org/10.1093/nar/gky949, PMID: 30357364

Xia B, Yanai I. 2019. A periodic table of cell types. *Development* **146**:dev.169854. DOI: https://doi.org/10.1242/dev.169854

Xing J, Li X, Wang X, Lv X, Wang L, Zhang L, Zhu Y, Shen Q, Baluška F, Šamaj J, Lin J. 2019. Secretion of phospholipase Dδ functions as a regulatory mechanism in plant innate immunity. *The Plant Cell* **31**:3015–3032. DOI: https://doi.org/10.1105/tpc.19.00534, PMID: 31597687

Yoo SD, Cho YH, Sheen J. 2007. Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nature Protocols* **2**: 1565–1572. DOI: https://doi.org/10.1038/nprot.2007.199, PMID: 17585298

Yu M, Cui Y, Zhang X, Li R, Lin J. 2020. Organization and dynamics of functional plant membrane microdomains. *Cellular and Molecular Life Sciences* **77**: 275–287. DOI: https://doi.org/10.1007/s00018-019-03270-7, PMID: 31422442

Zeisel A, Hochgerner H, Lönnerberg P, Johnsson A, Memic F, van der Zwan J, Häring M, Braun E, Borm LE, La Manno G, Codeluppi S, Furlan A, Lee K, Skene N, Harris KD, Hjerling-Leffler J, Arenas E, Ernfors P, Marklund U, Linnarsson S. 2018. Molecular architecture of the mouse nervous system. *Cell* **174**:999–1014. DOI: https://doi.org/10.1016/j.cell.2018.06.021, PMID: 30096314

Zhang Y, Song G, Lal NK, Nagalakshmi U, Li Y, Zheng W, Huang PJ, Branon TC, Ting AY, Walley JW, Dinesh-Kumar SP. 2019a. TurboID-based proximity labeling reveals that UBR7 is a regulator of N NLR immune receptor-mediated immunity. *Nature Communications* **10**:3252. DOI: https://doi.org/10.1038/s41467-019-11202-z, PMID: 31324801

Zhang TQ, Xu ZG, Shang GD, Wang JW. 2019b. A single-cell RNA sequencing profiles the developmental landscape of *Arabidopsis* root. *Molecular Plant* **12**: 648–660. DOI: https://doi.org/10.1016/j.molp.2019.04.004, PMID: 31004836

Zhang J, Petersen SD, Radivojevic T, Ramirez A, Pérez-Manríquez A, Abeliuk E, Sánchez BJ, Costello Z, Chen Y, Fero MJ, Martin HG, Nielsen J, Keasling JD, Jensen MK. 2020. Combining mechanistic and machine learning models for predictive engineering and optimization of tryptophan metabolism. *Nature Communications* **11**:4880. DOI: https://doi.org/10.1038/s41467-020-17910-1, PMID: 32978375

Zimny T, Sowa S, Tyczewska A, Twardowski T. 2019. Certain new plant breeding techniques and their marketability in the context of EU GMO legislation - recent developments. *New Biotechnology* **51**:49–56. DOI: https://doi.org/10.1016/j.nbt.2019.02.003, PMID: 30779963