Volume 1, 2025, Pages 23-33 *Journal of Smart and Sustainable Farming*
| Regular Article | Open Access

Advancements and Emerging Trends in Crop Phenomics

Ketema Niguse Baye
Department of Plant Sciences, Mattu University, Metu, Ethiopia.
ketema.niguse@mau.edu.et

Article Info
Journal of Smart and Sustainable Farming
https://www.ansispublications.com/journals/jssf/jssf.html

Received 20 December 2024
Revised from 26 January 2025
Accepted 02 February 2025
Available online 26 February 2025
Published by Ansis Publications

https://doi.org/10.64026/JSSF/2025003

Corresponding author(s):
Ketema Niguse Baye, Department of Plant Sciences, Mattu University, Metu, Ethiopia.
Email: ketema.niguse@mau.edu.et

Abstract – In 1911, Wilhelm Johannsen provided a definition for the term "phenotype," stating that it encompasses the many observable characteristics of organisms, which may be discerned by direct visual examination or more precise techniques including measurement or description. The term "phenotype" is derived from the Greek words "phainein" and "typos," which translate to "show" and "type" respectively. The area of research is now through a transformative phase known as 'phenomics,' which may be attributed to the rapid development in high-throughput technology of phenotyping. The agricultural phenotyping community is required to advance the field of bioinformatics in order to effectively extract valuable information from the extensive omics data. Additionally, it is essential to engage in research pertaining to technological systems that can accurately detect and characterize phenotypic features. This paper provides an outline of the topic of research including the gathering of phenotypic information using various sensors and the subsequent analysis of phenomics. In conclusion, we conducted an analysis of the challenges and possibilities associated with agricultural phenomics. Our objective was to provide suggestions for enhancing gene mining methodologies pertaining to essential agronomic traits, as well as implementing innovative intelligent approaches for precision breeding.

Keywords – Crop Phenomics, Model-Assisted Phenotyping, Crop Phenotypic Data Collection, Precision Feeding, Image-Based Phenotyping.

I.INTRODUCTION

The term "genome" was introduced in 1949 to refer to the genetic material underlying the genotype. Subsequently, the term "phenome" emerged to denote the ensemble of phenotypes arising from the cell's extragenic, non-auto reproductive constituents. Phenomics emerged as a nascent discipline throughout the 1990s, aiming to establish associations between genetic variations and clinical manifestations in the investigation of intricate human ailments. According to the study conducted by Rahaman, Chen, Gillani, Klukas, and Chen [1], phenomics is characterized as the methodical collection of multiset data pertaining to the phenotype of an organism. Zhang, Kong, Wu, Guan, Ding, and Chen [2] define "plant phenotyping" as the compilation of methodologies and protocols used to quantitatively assess several attributes of plants, including dimensions, morphology, and chemical composition. Furthermore, we choose to define crop phenomics as an interdisciplinary field that focuses on the efficient acquisition and analysis of multidimensional phenotypic data at a large scale, including the whole development cycle of crops.

Over the past decade, the field of plant phenomics has experienced significant growth, transitioning from prevailing domain to a booming research area. Plant phenomics involves the comprehensive collection of multi-dimensional phenotypic dataset at different dimensions, including the population, plant, organs, and cell levels. The intricate characteristics of crop phenotypes may be attributed to their emergence from the interactions between the genotype and a diverse array of environmental factors. The interaction between many factors influences the development and growth of crops, which may be seen via the structural characteristics at several levels of organization, including plants, organs, tissue and cellular levels. This interaction also affects the operation of the plant, as shown by its physiological qualities. The determination of external phenotypes, such as biomass, performance, and shape of a crop, is reliant on internal phenotypes.

The primary objective of crop phenomics research is to develop a proficient technical system capable of phenotyping crops in a manner that is multi-dimensional, high-throughput, big-data-driven, automated and intelligent. This research field integrates various disciplines such as life sciences, agronomy, mathematics, engineering sciences, and information science, alongside artificial intelligence technology and high-performance computing. Its aim is to investigate the diverse phenotypic data about the growth of a crop within intricate environmental context.

This study focuses on two primary components of agricultural phenomics research: the acquisition of phenotypic data using sensors and the subsequent analysis of such data. An extensive examination of crop phenotyping techniques was conducted, including several scales ranging from the molecular to the ecological, as well as spanning laboratory-based and field-based approaches. The investigation next delved into the processes of phenotyping extraction, phenotype information processing, and knowledge storage, with a particular focus on the substantial quantities of phenotypic data that are already accessible. It is emphasized that the big-data era, which phenomics is entering, is characterized by the presence of multi-level, multi-scale, and multi-domain attributes. The importance of advancing bioinformatics technologies that incorporate genomic and phenotypic data is underscored. This includes investigating E-feature depth evaluation frameworks and complex infrastructures to achieve precise detection of the crops phenotype. Additionally, the pursuit of functional-structure plant modeling according to phenomics is emphasized. Lastly, the establishment of a comprehensive, extensive, and cross-regional database of crop phenotypes is addressed.

Here is how this article has been organized: Section II discusses crop phenotypic data collection. In Section III a discussion of crop phenomics is provided in details. Section IV reviews model-assisted phenotyping. Lastly, Section V draws final remarks to the article as well as directions for future research.

II.CROP PHENOTYPIC DATA COLLECTION

The process of choosing unique genotypes in different treatments and conditions, as seen in **Fig. 1**, enables a control phenotyping station to generate experimental samples with great efficiency. In the course of the experiment, it is possible to precisely define and modify the initiation and intensity of certain conditions. Managing and combining experimental materials in a dynamic process poses difficulties, since the analysis of data gathering must include the influence of microclimate and ambient conditions. The gathering of phenotypic data, which involves observing and measuring observable traits of plants, may now be undertaken in an automated manner at regular intervals during the plant's life cycle. This has been made feasible due to recent advancements in the automation of plant phenotypic monitoring, which includes the use of robotics and sensors.

While it is possible to acquire high-throughput phenotyping facilities from commercial sources, a number of academic institutes are opting to develop their own facilities. Several more research organizations are now using these resources, such as IPK Gatersleben in Germany, Crop Design in Belgium, The Plant Accelerator in Australia, and PhenoArch in Montpellier, France. The phenotyping platform developed by LemnaTec is highly advanced and well recognized in the industry. It encompasses a range of applications that enable the analysis of plant pictures and the identification of various phenotypes.

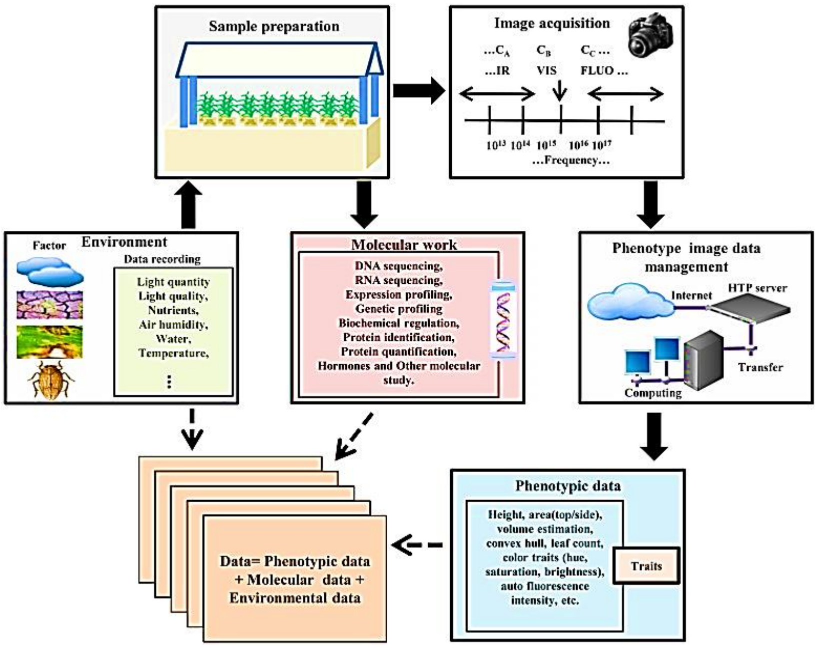

Fig 1. The Process of High-Throughput Phenotyping and Data Collecting from Plants

Plants are cultivated inside a controlled environment using a robotic control house system within phenotyping systems designed for high-throughput phenotypic imaging. Every individual plant is subjected to a distinct treatment, such as stress or mutant therapy, and is contained in a separate vessel with a properly regulated nutrient supply. Furthermore, each plant is separately transported down the conveyor belt. The technique autonomously captures a multitude of aerial and lateral images of germplasm resources and populations. After the acquisition of photos, it is necessary to transmit them to a data management system, wherein supplementary data such as environmental factors and genetic composition may be documented. Phenotypic features and qualities are then derived via the use of image processing methods. Data mining methods are used to get the values of extracted characteristics or statistically model and simulate phenotypic data with the purpose of developing phenotype-genotype models in different environmental contexts.

According to Xue, Elnashef, Jin, and Filin [3], there is a division within society of plant phenotyping, with a number of members advocating for phenotyping of a higher throughput that offers poorer resolution, while others prefer complex phenotyping that provides low throughput but high resolution. Automatic phenotyping models in controlled environments and high-throughput techniques in the field prioritize the aspect of higher throughput. However, phenotyping at cellular, tissue, and organ levels places emphasis on complex phenotyping and achieving better resolutions. This paper provides a complete survey of the methodologies used for the characterization of crops across several scales ranging from the molecular level to the field level. We emphasize the advantages and drawbacks associated with different technological approaches utilized in this context.

Micro-Phenotyping Bottleneck Alleviation

The attainment of higher temporal and spatial resolutions in tissues and cellular size phenotyping is a greater challenge when compared to the technologies used for whole-plant phenotyping. The efficiency of high resolution samples imaging is hindered by the micron-level output, while the organ or plant samples pre-treatment is often detrimental when conducted using a laborious and multi-step process. There is an urgent need for more sophisticated and automated image analysis tools to reliably evaluate the cell and tissue characteristics of crops inside a larger, higher-quality image. Over the past few decades, there has been a significant development in the establishment of various algorithms and approaches to effectively handle microscopic pictures obtained by different methods like as hand-slicing, micro-CT, fluorescence, laser, and paraffin section imaging.

The capacity of a plant to extract water and nutrients from the soil and transport them to the aboveground structures is contingent upon the anatomical characteristics of its roots. In recent years, there has been a notable advancement in micro-image capture technology and computer vision, which have significantly enhanced our knowledge of functional and anatomical aspects of roots. The Agricultural Science Lab at the Penn State College [4] used a computer-assisted technique to retrieve phenotypic data from sequence pictures of paraffin sections of wheat roots. Subsequently, the researchers have devised a phenotyping platform that exhibits a high resolution and high throughput. The platform integrates serial imaging, and laser optics alongside 3D imaging and reconstruction (RootSlice), including techniques to comprehensively comprehend root composition.

The semi-automated RootScan method, as proposed by Burton, Williams, Lynch, and Brown [5], is used in this context. To enhance the effectiveness of image segmentation and ensure a high level of accuracy, Chopin, Laga, Huang, Heuer, and Miklavcic [6] developed a fully automated software program known as RootAnalyzer. This tool was specifically designed for extracting microscopic phenotypic information related to roots. Fidan [7] developed an image processing system using X-ray micro-CT inside the Beijing Key Laboratory of Digital Plant. This system was designed specifically for the purpose of three-dimensional segmentation of metaxylem vessels in maize roots. The recommended methodologies greatly enhanced the ability to quantitatively assess phenotypic characteristics at the micron scale, as opposed to the traditional approach of physically measuring vascular bundles in maize roots.

The complexity of the microstructure of crop stalks surpasses that of the root system. The diversity and complexity of microscopic image information provide issues in the identification and detection of stalk tissue microscopic phenotypes. Ilakiya, Parameswari, Davamani, Swetha, and Prakash [8] conducted a semi-automated analysis of blemished stalk sections microscopic images, whereas Maddalena et al. [9] performed an automated analysis of the same images. The ability to assess anatomical features in a high-throughput manner was facilitated by the use of an image analysis tool that was created by Ning [10]. The application demonstrated the capability to process images of manually dissected stem cross-sections. Although significant advancements have been achieved in the assessment of the vascular bundle, challenges persist in accurately characterizing the anatomical properties of the rind and assuring precise detection.

Zhang, Ma, Pan, Wang, Guo, and Du [11], from the Digital Plant at the Beijing Key Laboratory, based on micro-CT (micro-computer tomography) technological advancement to examine the phenotypic characteristics of vascular bundles in whole cross-sections of maize stalks. Additionally, they devised the VesselParser 1.0 program for the purpose of analyzing the obtained images. A novel approach has been devised by experts from Beijing's Key Laboratory of Digital Plant to enhance the absorption difference of X-ray of maize tissues for conventional micro-CT scanning. Shi, Wang, Zhou, and Ren [12] devised a collection of image processing techniques specifically tailored for CT images of maize stalks, leaves, roots. These techniques efficiently obtained microscopic phenotypic vascular bundles characteristics.

The field of tissue and cellular-scale phenotyping still requires sophisticated methodologies. In order to enhance the efficiency of microscopic phenotyping research, it is necessary to optimize the preparation protocol sample and discover sophisticated imaging approaches. The processing of pictures is a notable challenge in micro-phenotyping investigations. The development of micro-image analysis approaches has mostly focused on specific biological testing, owing to the unique phenotypic characteristics of crop organs and cells across diverse variations.

3D Phenotyping at the Organ Level

In the realm of plant phenotyping, the majority of systems prioritize high-throughput capabilities for individual plants. Hence, the level of phenotypic accuracy shown by plant organs has consistently fallen short of perfection. Phenotyping systems have facilitated the acquisition of widely used phenotyping indices, including as leaf area, fruit volume, and leaf length. The mobile application platform was specifically developed to facilitate field phenotyping via the use of still photos and in vivo image processing techniques. When seen from a two-dimensional viewpoint, the application facilitates the process of measuring leaf angles and leaf length. The plants may be ethically and non-destructively sampled inside a controlled setting. Many phenotyping platforms already use inexpensive, two-dimensional camera systems. These camera systems have shown to be very effective in capturing phenotypic data for plant species that possess intricate branch structures. They are particularly useful for monitoring the dynamic growth and development of plant organs.

However, it is necessary to calibrate some morphological traits that have been anticipated, as well as acknowledge that 2D photographs lack information in the third dimension. The multi-view stereo (MVS) approach is an additional cost-effective alternative that is widely used for organ level phenotyping. The phenotypic characteristics were obtained by dividing the organs of individual plants into segments, followed by the reconstruction of three-dimensional point clouds utilizing structure from motion methodologies. The presented 3D reconstruction approach, which is cost-effective, demonstrates a viable alternative to using an expensive laser scanner in situations where automated processes are applicable. In a comprehensive omics analysis, it is possible that minor mistakes might be disregarded because to the significant inter-cultivar variation seen in the organ characteristics.

However, acquiring morphological data and statistically characterizing distinctions among plant organs that extend beyond measurements of length, area, and volume, such as blade folds, blade profiles, leaf hues, and vein curves, presents a formidable task. In order to facilitate comprehensive investigations into phenotypic traits via the use of precise organ-level phenotyping data, it is essential to develop a range of mathematical techniques that can effectively quantify and describe these variations. High-resolution 3D scanners are used by scientists for the aim of examining plant anatomy. Although high-resolution 3D scanners may be rather expensive, they provide a higher level of precision in capturing morphological data compared to organs reconstructed using multi-view stereo (MVS) techniques, especially for plants with non-planar surfaces. Phenotypic characteristics were assessed by using a mix of 2D LiDAR scanners, depth sensors, turn tables, and translation devices to collect and reconstruct point clouds in three dimensions.

In general, the use of 2D or 3D photography may assist in addressing the challenges associated with data collection in the context of 3D organ scale phenotyping, except in the case of tall plants. The use of computer graphics methods, including as skeleton extraction, surface reconstruction, and feature-preserving remeshing, is often applied in the processing of 3D high-throughput plant organs point clouds for the purpose of extracting 3D phenotypic traits. This is due to the intricate nature of the processing required in this task. In order to address the difficulties associated with the gathering and analysis of comprehensive organ data, Kurbatova et al. [13] put forward a standardized approach for data collection. Additionally, they developed a database that contains plant organ resources derived from in-situ morphological measurements. The main purpose of this initiative is to augment the integration and dissemination of organ data of superior quality.

Automated Phenotyping Platforms in Controllable Environment

A diverse array of methods and equipment has been developed to assess the phenotypic response of plants to particular environmental conditions. This is important given the considerable levels of environmental variability seen in field settings. A wide range of facility designs are available that provide high-quality and comprehensive resolution in gathering phenotyping data, while also allowing for flexible levels of environmental control. The range of facilities available for experimental control may vary, ranging from those that provide manipulation of a single component, such as rainout shelters, to those that allow manipulation of many factors, such as greenhouses, to those that provide complete regulation and can replicate specific environmental conditions seen in the field.

Mobile carrier platforms often use cameras and sensors for phenotyping in dynamic field settings. However, these platforms are susceptible to the effect of environmental conditions, such as sunshine intensity and wind speed, which may affect the accuracy of measurements. The efficacy of field phenotyping, when compared to indoor phenotyping conducted in controlled environments, may be compromised in accurately discerning the influence of genotypic variation on a captured phenotype image. This limitation arises from unfavorable environmental factors, such as the presence of strong wind causing leaf and whole plant canopy movement, as well as the obstruction of direct sunlight by clouds, leading to color distortions during the imaging process. Despite this susceptibility, field phenotyping methods provide valuable data and significantly contribute to our understanding of the interaction between a crop and its surrounding environment.

The utilization of camera systems installed on mobile tractors or unmanned aerial vehicles facilitates the non-intrusive and efficient assessment of field stands, allowing for the swift and precise measurement of plant height and the extraction of color-related characteristics from crop canopies. Satellite imagery, with its ability to capture an entire field in a single image, holds promise for achieving optimal efficiency in measuring features such as the normalized difference vegetation index. Research conducted by Chan, Gomez, Kothikar, and Baiz [14] has demonstrated that satellite imagery exhibits the highest correlation with biomass and yield when compared to alternative methodologies. The use of stationary equipment strategically positioned around a field stand enables the attainment of higher-resolution phenotyping in field-based research. The researchers used a stationary crane equipped with cameras to conduct high-throughput phenotyping of a field stand.

This method proved effective in accurately evaluating the wheat heading and flowering stages at the Rothamsted Research facility located in the United Kingdom.

The development of automated and robotic systems, advancements in imaging, and sensors techniques, including both software and hardware, has enabled the introduction of HTPP (high-throughput plant phenotyping platforms). Over the last decade, there have been notable developments in the research and development of HTPP. HTPP utilized in growth chambers or greenhouses can be categorized into two main types: plant-to-sensor and sensor-to-plant. The sensor-to-plant models integrate plants that remain stationary while an imaging setup is moved to capture data from each individual plant. On the other hand, the plant-to-sensor models involve transporting the plants to an imaging station for data collection. The use of various approaches in HTPPs inside growth chambers or greenhouses is seen collectively, as shown in **Table 1**.

Table 1. Approaches in HTPPs within growth chambers or greenhouses

RGB imaging	The phenotypes of plant color, texture, and form are acquired.
Hyperspectral imaging	Determines characteristics such pigmentation, biochemical make-up, nitrogen concentration, and water content.
Chlorophyll fluorescence imaging	Acquires phenotypic traits related to photosynthesis
Thermal imaging	Identifies a plant's stomatal conductivity, transpiration rate, and surface temperature distribution
Lidar	Determines a plant's phenotype based on its three-dimensional structure

Furthermore, advanced imaging techniques often used in the field of medicine, such as Positron Emission Tomography (PET), Computed Tomography (CT), and Magnetic Resonance Imaging (MRI), have been incorporated into High-Throughput Phenotyping Platforms (HTPPs) inside growth chambers or greenhouses. The following table provides a comprehensive overview of the prevalent imaging techniques used in high-throughput plant phenotyping (HTPP) inside greenhouse or growth chamber settings. Sun et al. [15] presents a comprehensive compilation of High-Throughput Phenotyping Platforms (HTPPs) that have shown excellent phenotyping outcomes inside controlled environments such as growth chambers or greenhouses, facilitated by the use of digital technology, platform equipment, and information technology. Automated, high-throughput, and precise phenotyping methods have been shown to enhance the productivity of crop breeding by enabling more accurate and efficient data collection from plants.

High-Throughput Approaches for Crop Phenotyping Within the Field Environments

Field-based phenotyping (FBP) plays a crucial role in the assessment of environmental factors, genetic factors, and their interplay on significant production characteristics such as yield potentials, and tolerance to biotic/abiotic stresses. Consequently, FBP holds immense significance in the realm of crop enhancement via genetic means. The FBP technique is becoming identified as the major strategy able to attain the needed throughput and provided precise depiction of trait manifestation in practical agricultural contexts. Measuring the efficacy of breeding programs on agricultural production and productivity within their respective natural habitats is of utmost importance. The majority of current field-based phenotyping platforms (FBPPs) rely on aerial vehicles or ground wheeled vehicles, which are integrated with diverse sensors to efficiently measure plant characteristics within a short time frame per plot. The process of gathering data from a substantial quantity of plots may be arduous for ground-based Forest Biodiversity Plot Programs (FBPPs).

As an example, the process of measuring attributes on each individual row required around 40 hours while using a solitary vehicle traveling at a speed of 2 kilometers per hour through a total of 20,000 plots. Collecting data simultaneously in all plots using several vehicles and sets of sensors would result in higher costs. Furthermore, the restricted mobility of terrestrial Fixed Base Power Platforms (FBPPs) renders them unsuitable for undertaking projects that include many areas. In recent years, the development of the cable-suspended field phenotyping platform has emerged as a valuable tool for efficiently and non-invasively evaluating agricultural qualities. The cable-suspended field-based phenotyping system has many advantages, including little physical contact with plants, freedom from soil-related constraints, and high precision. Cable-suspended field phenotyping devices provide the capacity for extensive phenotyping on a broad scale; nevertheless, their use is limited due to their restricted placement options. The use of satellites or aircrafts for the purpose of remote sensing has the potential to mitigate certain limitations.

The increasing need for precise temporal, spectral, and spatial resolutions in crop phenotyping conducted in field settings has resulted in the emergence of UAV-RSPs (unmanned aerial vehicle remote sensing platforms) [16] and manned aircraft as an efficient and rapid tool. For example, the temperature of canopies was examined by the use of thermal sensors that were installed on manned aircraft. The sensors commonly carried by UAV-RSPs have various applications, such as modeling the surface of canopies and estimating crop biomass using visible imaging, monitoring the physiological status of crops through high-resolution hyperspectral imaging, and visible-near-infrared spectroscopy, and identifying plant water status using infrared thermography and hyperspectral imaging, among others. UAV-RSPs provide many advantages, including portability, high monitoring efficiency, cost-effectiveness, and the flexibility to operate in various field situations.

Nevertheless, UAV-RSPs are vulnerable to factors such as fluctuating meteorological conditions, rigorous limitations on airspace, and the lack of expedient and automated instruments for data processing and modeling. In recent times, there has been a notable advancement in the enhancement of phenotyping capabilities via the integration of aerial and ground-oriented platforms. To illustrate, the integration of a tractor-oriented proximate crop-sensing model with an unmanned aerial vehicle-based model was utilized to address intricate factors such as growth of sorghum.

III.CROP PHENOMICS

The presence of several sensor models, including factors such as morphology, structure, color, and physiological data, together with multiple phenotypic data scales ranging from population to cellular dimension, and the inherent unpredictability in the environment provide challenges in the direct replication of phenomic research. Phenomic investigations need a comprehensive approach beyond the mere acquisition of phenotypic data pertaining to agricultural crops. The identification of phenotypic traits within raw data. What strategies may be used to attain centralized storage of phenotypic data? This study explores the methodologies used in doing meta-analyses that include several scales and dimensions. In conclusion, this inquiry explores the potential implementation of model-assisted phenotyping and phenotypic-genomic selection methodologies using extensive datasets pertaining to phenotypic traits. Thus, this part focuses on the previously stated contemporary developments in agricultural phenomics research. It is posited that the current stage of development in this particular area of research is seeing a notable expansion due to advancements in artificial intelligence methodologies, particularly in the realm of deep learning. These advancements have the potential to aid researchers in effectively transforming vast quantities of omic-data into valuable insights and understanding.

Phenotype Extraction

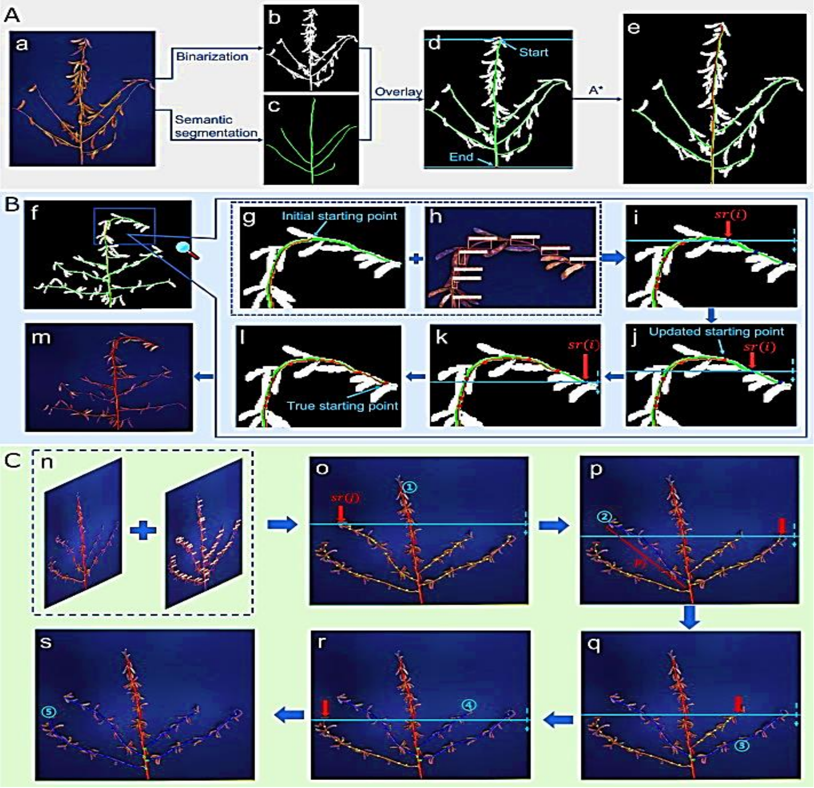

Fig 2. Procedures for Extracting Phenotypes Linked with Stems

Image-based phenotyping has been used in laboratory, greenhouse, and field settings to assess and measure visual and model-based characteristics of plants. Currently, a plethora of image evaluation systems and methods have been designed for the purpose of picture-based plant phenotyping. From a perspective of image analysis, the phenotypic characteristics of plants may be dissected into linear and non-linear components that are linked to the pixel representation. These components can further be categorized into four distinct groups: quantity, geometry, color, and texture. The use of image attributes may also serve as a means to deduce significant agronomic and physiological qualities.

Typically, imaging methodologies are customized to suit the specific crops and phenotypic characteristics of interest, necessitating prior knowledge of the subjects under investigation, and including a certain level of human involvement. Phenotypic data pertaining to biomass, normalized difference vegetation index (NDVI), chlorophyll responses, compactness, and other relevant parameters might be deemed appropriate within the context of using the conventional image processing methodology inside a laboratory setting. The processing pipeline for basic image analysis still faces challenges when it comes to non-geometric and non-linear phenotyping tasks. The development of an intelligent and fully automated image processing system will need a significant amount of time due to the vast array of species and habitats involved.

The determination of plant height involves the measurement of the vertical distance between the highest point of the plant and the cotyledon node located on the primary stem. Similarly, branch length is measured by measuring the distance from the uppermost point of a branch to the point at which it intersects with the primary stem. The problem at hand might be conceptualized as the task of determining the most efficient route between two designated locations on the pixel-based representation of an image. The generation of the pixel map may be accomplished by two distinct methods: binarization, as previously mentioned, or semantic segmentation. **Fig. 2** A illustrates the process of creating a composite map using the SPP-extractor, which involves the fusion of binarized and semantically annotated maps: (A) Using a map overlay to identify the main branches. For (a), the primary photo. The resulting binary image (b). That which has been semantically annotated is (c). Path finding composite map (d). (e) The picture with the main stem removed and annotated. Option (B): The curved top. (f) The original, unfinished course of the main stem. (g) Establishing a base line. Obtaining useful stem parts is step (h). Expansion of the route, items (i) through (l). The ultimate, unbroken route of the main stem. Identifying the length of a branch and measuring it (C). The merged map used to pinpoint individual branches. (o-s) Take a measurement of all four limbs after they've been cut off.

Phenotype Data Standardization and Storage

Deriving combinable phenotypic variables using algorithmic techniques is a straightforward process, achieved by constructing common weight categories based on continuous data. However, the issue of effectively standardizing tests of intelligence and other complex qualities has been given comparatively less attention. In order to get a deeper comprehension of the present implementation of harmonization in the realm of cognition, we conducted a comprehensive examination of meta-analyses. Our analysis revealed that a significant proportion of aggregate data meta-analyses utilised the technique of standardization to integrate cognitive measurements across several studies. The prevalence of standardization approaches may be linked to their little reliance on complex modeling methods, such as latent variable analysis.

All the aforementioned phenotypic technologies possess the capacity to generate substantial quantities of intricate data, including the amalgamation of diverse picture, spectral, and environmental data, sometimes reaching the scale of petabytes or gigabytes, instituting unstructured "Big Data." Consequently, the management, storage and access of phenotypic data have arisen as significant challenges. Currently, consensus exists about three fundamental components of information standardization. It is advisable to utilize the "minimum information" (MI) method in order to establish the dataset content. Additionally, the application of ontology terms should be employed for the purpose of annotating data in a manner that is both unique and repeatable. Furthermore, it is recommended to select appropriate formats of data, such as MAGE-TAB, XML, RDF, and CSV, etc., for datasets construction. Numerous phenotyping resources, including both single-species and multi-data-type approaches, have been established so far. The PHENOPSIS DB serves as an information system specifically designed for the acquisition of phenotypic data related to Arabidopsis thaliana. This data is obtained through the utilization of the PHENOPSIS phenotyping platform. Bisque, on the other hand, represents a pioneering web-based, cross-platform tool that has been developed to function as a repository. Its primary purpose is to analyze, organize, visualize, and store images within a cloud-based environment. Lastly, PODD was established in 2010 with the aim of facilitating the capture, management, annotation, and distribution of data. Its intended beneficiaries include both the Australian and international biological research communities.

PhenoFront refers to a dataset that was made accessible to the LemnaTec Phenotyper platform, containing information on above-ground plant tissue. AraPheno, on the other hand, served as the initial all-encompassing repository for phenotypic data on a population scale. It successfully integrated over 250 openly accessible phenotypes from 6 distinct researches, focusing on inbred lines of A. thaliana. Lastly, the plant genomics and phenomics database (PGPDB) was established in 2016 with the purpose of effectively storing, managing, and retrieving leaf images and associated phenotypic data. The use of cloud technology as a storage scheme is increasingly gaining popularity in the field of plant phenotypic data storage due to its ability to facilitated efficient storage and standardization of database data. The use of cloud storage systems may enhance the plant phenotype platform by leveraging their capacity to improve system infrastructure, high-speed cache, file structure, and other relevant components. Currently, there exists a dearth of centralized infrastructure for the collection of phenotypic data at the regional, national, or continental scale. The development of an agricultural phenotype database using AI technology and including multi-layer phenotypic data, similar to the Human Genome Database of GDB, would attract significant attention from many stakeholders.

IV.MODEL-ASSISTED PHENOTYPING

The morphological attributes and physiological responses of plants exhibit significant plasticity in relation to genetic factors, environmental conditions, and cultivation practices. The phenotypic development of plants is influenced by intricate interactions including genetics, environment, and management practices at different scales. However, the precise significance of each element in determining the phenotype remains little understood. The examination of abiotic and biotic factors that affect plant phenotypes may be effectively analyzed via the use of dynamic models. Functional-structural plant (FSP) models are used to quantitatively analyze the complex relationship between a plant's anatomical characteristics and its physiological functions. These models simulate the development and growth of a plant through time and in a three-dimensional context. The use of phenotyping in combination with FSP modeling has been employed due to its simplicity as a strategy for addressing research inquiries.

Initially, FSP models serve as a valuable tool for understanding the underlying mechanisms associated with a certain characteristic. Blanc, Barbillon, Fournier, Lecarpentier, Pradal, and Enjalbert [17] used Functional-Structural Plant (FSP) model to simulate the capture of light by the whole vegetation, with the aim of distinguishing the impact caused by interspecific trait variations from the impact caused by phenotypic plasticity. This was achieved by comparing cases with and without phenotypic plasticity, utilizing experimental trait data of plants. The assessment of the impact of individual architectural characteristics on light interception and dry mass production was done based on dynamic FSP framework for tomato to simulate canopy growth. The FSP framework was utilized to evaluate light interception for individual maize plants. This allowed us to separate the factors of radiation interference and radiations consumption efficiency into distinct components, namely a genetic term and environmental term. This analysis was conducted using simulated multi-genotype canopies. The use of FSP modeling has shown significant utility in assessing the relative impact of separate processes on the overall system.

Furthermore, the use of phenotyping methodologies with high-throughput enables the automatic and accurate FSP model calibration. The use of daily picture analysis was employed in the reconstruction of maize plant architecture in order to determine fresh plant weight and leaf area, as well as to measure light interception using an FSP model. The intricate structure of the tree canopy was rebuilt using the use of terrestrial LiDAR scanning technology. This enabled the estimation of many microclimate-related parameters such as evapotranspiration, surface temperature, and net radiation in three-dimensional space. The OPENSIMROOT project [18] integrated three-dimensional phenotyping methodologies, such as CT and MRI, with a root model capable of simulating the growth and development of a root system. Phenotyping methodologies are often used by researchers to evaluate the model's ability to accurately replicate plant architecture and geometry, as well as to provide valuable insights about functional responses via the analysis of images.

Functional-structural plant models (FSPMs), sometimes referred to as virtual plant models, include mathematical representations of plants that elucidate the temporal dynamics of their three-dimensional architecture or structure. These models capture the intricate interplay between physiological processes and environmental variables, since the former influences the latter and vice versa. According to this perspective, plants are shown as a complex arrangement of basic units referred to as phytomers see **Fig. 3**, which are alternatively known as metamers. Based on the FSPM paradigm, plants have the ability to modify the environmental conditions, including light, in which their physiological processes occur. This is achieved through adjustments in both their functional activities, such as photosynthesis, transpiration, and nitrogen allocation, as well as their structural attributes, such as bud development and dormancy, and the morphology and orientation of their organs. This is precisely why Functional-Structural Plant Modeling (FSPM) is so advantageous, since it enables the direct capturing of feedback mechanisms between plant structure and function. Furthermore, the interplay of activities taking place inside an individual organ, referred to as the "local level," and the overall functioning of plants or plants stands as a whole, known as the "global level," may also be examined.

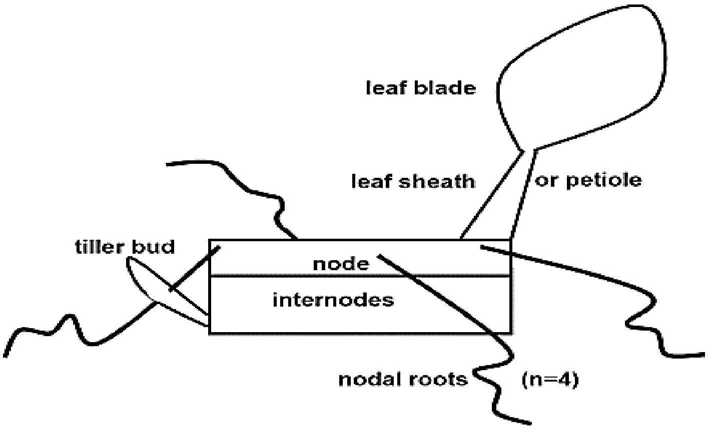

Fig 3. The Phytomer, Which Serves as the Rudimentary Unit for Modeling Plant Configuration

The term "leaf sheath" is often used in reference to cereals and grasses, whereas the term "petiole" is typically used to describe broad-leaf plants. Similar to wheat, these pictures illustrate the process of implanting four nodal roots. The establishment of an FSPM raises not just questions about computational issues and application, but also presents chances to expand the repertoire of tools accessible to plant experts.

The focus of the simulation in the conventional Plant-Based Model (PBM) pertains to the yield per unit area, as shown by Giampieri et al. [19]. Although the primary emphasis of Functional-Structural Plant Modeling (FSPM) is in examining the internal mechanisms of an individual plant, it is feasible to construct a model including a community of plants and their interrelationships. The structural composition of each model inherently varies based on the specific objectives of the investigation. Dynamic frameworks utilize the available condition of the system and change rate visualized during the time phase t to calculate the anticipated system state at consequent time phase, t+t. Additionally, there are "static structural models" that depict a three-dimensional configuration at a certain point in silico, but lack methods for determining the structure's temporal progression. This research use dynamic modeling as the primary assumption, while also using static structural models in various scenarios.

Phenotype–Genotype Association Analysis

The first examination of a connection acts as a significant tool for discovery. It seems improbable that a single study would establish a conclusive genotype-phenotype association without the need for further replication in the foreseeable future. When determining the need of doing follow-up replication studies for a result derived from a preliminary analysis of the whole genome or specific candidate genes, it is crucial to thoroughly evaluate several aspects pertaining to the design and reporting of the research. The absence of comprehensive technical details in the first report or the lack of methodological rigor in the original study may sometimes provide challenges when attempting to replicate a claimed association.

In the absence of stringent statistical criteria, the prevalence of false positives will significantly surpass that of true positives in genome-wide research, given the extensive examination of genotype-phenotype connections. P-values that approach 0 indicate a heightened degree of correlation between the two variables. However, it is important to note that there is no universally defined threshold for determining statistical significance. Nevertheless, it is important to exercise caution in interpreting high P-values until replication studies have been conducted. This is because such values might potentially be influenced by technical artifacts or genetic mistakes that may be unevenly distributed across the cases and controls. When seeking for markers of high significance, it is important to closely observe the cluster plots. In order to establish the accuracy of the reported genotypes, it might be advantageous to include corroborating data obtained via a secondary genotyping approach. The analysis of confounding population stratification necessitates the careful selection of cases and controls from cohorts that exhibit comparable genetic backgrounds and environmental exposures.

Diabetes and age-related macular degeneration are two well-established phenotypes that have been widely used in research investigating genotype-phenotype interactions, which have been replicated several times. The use of dependable criteria is expected to result in a decrease in misclassification rates. Endophenotypes, also known as intermediate phenotypes, have been documented in specific association studies. However, the available research has not provided sufficient evidence to ascertain their validity or reliability. In the absence of established and universally accepted standards, it is crucial to provide sufficient information to clearly identify the phenotypes under investigation and assess their validity and comparability across different research studies.

In order to achieve a successful replication of a genotype-phenotype relationship, it is essential to adhere to similar parameters as those seen in genome-wide association or candidate-gene research. Replication studies are of utmost importance in ensuring the validity and reliability of scientific research. In order to do this, it is essential that these studies focus on phenotypes that align with those investigated in the original study. In order to assure the reliability of findings, replication studies, similar to their original counterparts, must use statistical methodologies that effectively address population stratification. Numerous studies have been undertaken and subsequently published pertaining to individuals of European descent, hence posing challenges in extrapolating the results to include other populations. The TCF7L2 gene variant among populations of West Africa and East Asia. Conducting further investigations in some instances might provide challenges owing to the atypical characteristics of the study sample or a dearth of appropriate replication subjects. In the event that replication is unattainable, laboratory-derived biological insights may be used to bolster the interpretation of association findings.

In the era of omics, establishing a strong connection between genomic discoveries and plant phenomics is essential, notwithstanding the significant contribution of genomic data to advancements in crop genetics and breeding initiatives. In recent times, there has been a surge in phenomic research across several species. This research involves the integration of genomic data with quantitative variations in phenotypes. By doing so, it enables the rapid decoding of the function of numerous unidentified genes and enhances our understanding of the genotype-phenotype map.

Multiple genes with moderate effects regulate a variety of agronomically significant traits. Genome-wide association studies (GWAS) and Quantitative trait locus (QTL) mapping have been extensively used in the field of agricultural plants to investigate the genetic basis of biomass accumulation. These approaches include the correlation of phenotypic features related to small grain cereals with genomic data, enabling the dissection of the underlying genetic architecture. In their study, Young and Scott [20] used genome-wide association studies (GWAS) to detect 141 associated loci in rice, using a set of 13 conventional agronomic parameters and two recently created traits. In the year 2015, a group of researchers using a combination of Genome-Wide Association Studies (GWAS) and high-throughput leaf scoring (HLS) techniques to discover a total of 73 previously unidentified genetic regions linked to leaf size, 123 genetic regions linked to leaf color, and 177 genetic regions linked to leaf shape. These associations were seen across 29 distinct leaf characteristics at three different stages of development.

In 2017, a comprehensive analysis was conducted on the inbred line of maize, by Shahzad, Gul, Ahsan, Wang, and Fahad [21], covering the seedling phase to the tasseling phase. This analysis included the mapping of a total of 106 agronomic traits, resulting in the discovery of 988 quantitative trait loci (QTLs). A total of twelve marker-trait associations pertaining to major effects were developed by measuring the growth of plants over eleven developmental stages in a population of 252 different maize inbred lines.

V.CONCLUSION AND FUTURE PROSPECTS

Given the current era of "Big Data" in phenomics, it is crucial for the crop science society to utilize the technology of artificial intelligence and foster national and international collaboration. This will facilitate the development of a novel framework for studying crop phenotypic data, establishing an efficient technical model capable of multi-dimensional, high-throughput big-data, automated and intelligent crop phenotyping, and designing a comprehensive tool that seamlessly integrates big data. Overcoming various challenges will be essential for the field of agricultural phenomics in the next five to 10 years.

The use of multi-dimensional, multi-scale, and high-throughput phenotyping techniques is propelling the study of phenomics into the era of big data. In this study, we present an extensive array of phenotyping techniques utilized for the evaluation of crop structure, physiological data, and morphology. These methods exhibit 3 key attributes: multi-domain capabilities encompassing phenomics, genomics, and other domains; multi-level capacities ranging from traditional medium and small scale evaluations to large-scaled omics evaluations; and multi-scale capability enabling the examination of cellular to whole-plant morphology, structure, and physiological data. The use of data derived from systematic and comprehensive phenomics research will play a crucial role in future investigations. This is due to the fact that relying just on single and individual phenotypic information will no longer enough for association analysis in the emerging era of "-omics."

In order to address emerging challenges, the area of image-based phenotyping will be advanced via the use of novel AI-based methodologies and methods. The automated phenotyping method and platform generate many digital characteristics, which need validation via statistical analysis with large samples and association analysis with traditional agronomic variables in order to determine their value. One of the primary obstacles in the advancement of plant phenotyping is in the determination of effective methods for the precise and expeditious evaluation, comprehension, and interpretation of digital image-derived attributes, as well as the extraction of pertinent quantitative traits for functional genomics. There is an urgent requirement to utilize the latest advancements in artificial intelligence (AI) encompassing data fusion, deep learning, swarm intelligence, and hybrid intelligence. The purpose is to advance big-data management systems that can effectively facilitate the integration, interoperability, ontologies, shareability, and globalization of data. This should be based on the incorporation of multi-level, multi-scale, and multi-domain phenotypic data.

CRediT Author Statement

The author reviewed the results and approved the final version of the manuscript.

Data Availability

The datasets generated during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interests

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Funding

No funding was received for conducting this research.

Competing Interests

The authors declare no competing interests.

References

[1]. M. M. Rahaman, D. Chen, Z. Gillani, C. Klukas, and M. Chen, “Advanced phenotyping and phenotype data analysis for the study of plant growth and development,” Front. Plant Sci., vol. 6, p. 619, 2015.

[2]. C. Zhang, J. Kong, D. Wu, Z. Guan, B. Ding, and F. Chen, “Wearable sensor: An emerging data collection tool for plant phenotyping,” Plant Phenomics, vol. 5, p. 0051, 2023.

[3]. F. Xue, B. Elnashef, W. Jin, and S. Filin, “Polarimetric monocular leaf normal estimation model for plant phenotyping,” ISPRS J. Photogramm. Remote Sens., vol. 202, pp. 142–157, 2023.

[4]. “Agricultural analytical services lab (Penn State College of agricultural sciences),” Agricultural Analytical Services Lab (Penn State College of Agricultural Sciences). [Online]. Available: https://agsci.psu.edu/aasl. [Accessed: 30-Jul-2023].

[5]. A. L. Burton, M. Williams, J. P. Lynch, and K. M. Brown, “RootScan: Software for high-throughput analysis of root anatomical traits,” Plant Soil, vol. 357, no. 1–2, pp. 189–203, 2012.

[6]. J. Chopin, H. Laga, C. Y. Huang, S. Heuer, and S. J. Miklavcic, “RootAnalyzer: A cross-section image analysis tool for automated characterization of root cells and tissues,” PLoS One, vol. 10, no. 9, p. e0137655, 2015.

[7]. S. Fidan, “The use of micro-CT in materials science and aerospace engineering,” in Micro-computed Tomography (micro-CT) in Medicine and Engineering, Cham: Springer International Publishing, 2020, pp. 267–276.

[8]. T. Ilakiya, E. Parameswari, V. Davamani, D. Swetha, and E. Prakash, “High-throughput crop phenotyping in vegetable crops,” Pharma Innov., vol. 9, no. 8, pp. 184–191, 2020.

[9]. R. Maddalena et al., “Nanoparticle characterization by automated acquisition and analysis of images and EDS data in the TEM,” Microsc. Microanal., vol. 27, no. S1, pp. 1810–1810, 2021.

[10]. C. Ning, “RETRACTED ARTICLE: Design and research of motion video image analysis system in sports training,” Multimed. Tools Appl., vol. 82, no. 18, pp. 28673–28673, 2023.

[11]. Y. Zhang, L. Ma, X. Pan, J. Wang, X. Guo, and J. Du, “Micron-scale phenotyping techniques of maize vascular bundles based on X-ray microcomputed tomography,” J. Vis. Exp., no. 140, 2018.
[12]. Z. Shi, Y. Wang, Z. Zhou, and W. Ren, “Integrating deep learning and traditional image enhancement techniques for underwater image enhancement,” IET Image Process., vol. 16, no. 13, pp. 3471–3484, 2022.
[13]. E. V. Kurbatova et al., “A standardized approach for collection of objective data to support outcome determination for late-phase tuberculosis clinical trials,” Am. J. Respir. Crit. Care Med., vol. 207, no. 10, pp. 1376–1382, 2023.
[14]. T. Chan, C. A. Gomez, A. Kothikar, and P. Baiz, “Joint study of Above Ground Biomass and Soil Organic Carbon for total carbon estimation using satellite imagery in Scotland,” arXiv [stat.AP], 2022.
[15]. J. Sun et al., “High-throughput phenotyping platforms enhance genomic selection for wheat grain yield across populations and cycles in early stage,” Züchter Genet. Breed. Res., vol. 132, no. 6, pp. 1705–1720, 2019.
[16]. J. Pfeiffer, T. Zieher, M. Rutzinger, M. Bremer, and V. Wichmann, “Comparison and time series analysis of landslide displacement mapped by airborne, terrestrial and Unmanned aerial vehicle based platforms,” ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci., vol. IV-2/W5, pp. 421–428, 2019.
[17]. E. Blanc, P. Barbillon, C. Fournier, C. Lecarpentier, C. Pradal, and J. Enjalbert, “Functional-structural plant modeling highlights how diversity in leaf dimensions and tillering capability could promote the efficiency of wheat cultivar mixtures,” Front. Plant Sci., vol. 12, p. 734056, 2021.
[18]. J. A. Postma et al., “O pen S im R oot : widening the scope and application of root architectural models,” New Phytol., vol. 215, no. 3, pp. 1274–1286, 2017.
[19]. F. Giampieri et al., “Organic vs conventional plant-based foods: A review,” Food Chem., vol. 383, no. 132352, p. 132352, 2022.
[20]. R. P. Young and R. J. Scott, “Gene-gene interaction in ever smokers with lung cancer: Is there confounding by chronic obstructive pulmonary disease in genome-wide association studies?,” J. Thorac. Oncol., vol. 18, no. 3, pp. e23–e24, 2023.
[21]. A. Shahzad, H. Gul, M. Ahsan, D. Wang, and S. Fahad, “Comparative genetic evaluation of maize inbred lines at seedling and maturity stages under drought stress,” J. Plant Growth Regul., vol. 42, no. 2, pp. 989–1005, 2023.