ACTA BIOLÓGICA COLOMBIANA
http://www.revistas.unal.edu.co/index.php/actabiol

Facultad de Ciencias
Departamento de Biología
Sede Bogotá

UNIVERSIDAD NACIONAL DE COLOMBIA

ARTÍCULO DE REVISIÓN / REVIEW ARTICLE

UNRAVELING THE GENETIC ARCHITECTURE OF COMPLEX TRAITS IN PLANTS

Descifrando la arquitectura genética de rasgos complejos en plantas

Laura Vanessa CHIVATÁ-PEÑA[1], Laura M. PERILLA-HENAO[1], Johana Carolina SOTO SEDANO[2*]

[1] Department of Plant Pathology, University of California, One Shields Avenue, Davis, U.S.A.
[2] Manihot Biotec Laboratory, Biology department, Universidad Nacional de Colombia, Carrera 30 n°. 45-03, Bogotá D.C., Colombia
** For correspondence: jcsotos@unal.edu.co*

Received: 10th June 2021. **Revised**: 20th February 2023. **Aceptado**: 17th April 2022
Editor asociado: Hernán Mauricio Romero

Citation/ citar este artículo como: Chivatá-Peña, L.V., Perilla-Henao, L.M, Soto Sedano, and J.C. (2023). Unraveling the Genetic Architecture of Complex Traits in Plants. *Acta Biol Colomb*, *28*(3), 350-367. https://doi.org/10.15446/abc.v28n3.98891

ABSTRACT

Complex traits are those whose inheritance does not follow simple and predictable patterns. They are not governed by a single locus, instead, they are determined by several loci and are influenced by the environment. Most of the traits with agronomic interest and economic importance such as resistance to biotic and abiotic stress, and yield, among others, are quantitative traits and their study is based on dissecting the underlying genetic architecture, the number of loci responsible for the variance of a quantitative trait, the relevant contribution made by each locus and their interaction with the environment. This review provides the most relevant conceptual bases for the study of the genetic architecture of complex quantitative traits in plants. The methodologies that allow identifying the loci and candidate genes that govern this type of traits are described, such as QTL mapping by linkage and association mapping. In addition, the incorporation of these loci in phenotype prediction strategies such as marker-assisted selection and genomic selection, exhibits the benefits and limitations of these approaches. Finally, the challenges and perspectives facing the study of the genetic architecture of complex traits in plants are discussed.

Keywords: Genomic selection, linkage disequilibrium, polygenic trait, quantitative trait *loci*.

RESUMEN

Los rasgos complejos son aquellos cuya herencia no sigue patrones simples y predecibles. No están gobernados por un solo locus, sino que están determinados por varios loci y, además, están influenciados por el entorno. La mayoría de los rasgos de interés agronómico como la resistencia al estrés biótico y abiótico, el rendimiento, entre otros, son rasgos complejos, gobernados por múltiples genes a lo largo del genoma. El estudio de la arquitectura genética de rasgos complejos se basa en la identificación del número de loci asociados a un rasgo, la contribución individual de cada loci al rasgo, la heredabilidad y el grado de influencia que del ambiente en el fenotipo. Esta revisión proporciona los conceptos más relevantes para el estudio de la arquitectura genética de rasgos complejos en plantas. Se describen las metodologías que permiten identificar los loci y genes candidatos, que gobiernan este tipo de rasgos como el mapeo QTL por ligamiento y el mapeo por asociación. Además, la incorporación de estos loci en estrategias de predicción del fenotipo como la selección asistida por marcadores y la selección genómica, presentando los beneficios y limitaciones de estos enfoques. Finalmente, se presentan los desafíos y perspectivas que enfrenta el estudio de la arquitectura genética de rasgos complejos en plantas.

Palabras clave: Desequilibrio de ligamiento, *loci* de rasgos cuantitativos, rasgo poligénico, selección genómica.

INTRODUCTION

Since the classic Mendel´s work in 1856, the research of "simple" inheritance traits, or monogenic, led the research in plant genetics in the XX century. The characterization of several monogenic traits has been achieved, identifying their governing locus, in model and non-model plant species. The study of monogenetic traits is relatively simple, phenotyping is performed in discrete qualitative categories and rarely presents ambiguities, which are a reflection of the underlying genotype (St.clair, 2010). On the other hand, the inheritance of quantitative traits is polygenic and complex in nature, as they are controlled by many genes and their interaction with each other and with the environment. The research on quantitative traits has bloomed in recent decades, in part because of the economic importance of these traits in agriculture, which become the interest and the focus of many plant breeding programs around the world. Several studies in model species and of agronomic interest have described phenotypes with continuous ranges, variable expressivity, and in most cases highly dependent on the environment, such as yield (Cai et al., 2016), resistance to biotics (Thoen et al., 2017), and abiotics factors (Pan et al., 2015) they owe their dynamics to the orchestrated action of multiple loci.

A challenge for contemporary genetics is the understanding of how allelic variation causes phenotypic variation in complex traits. As well as describing the number of quantitative trait loci or QTL, that underlie their genetics, the possible interactions between these loci and their individual and synergistic effects. At the end, to describe its genetic architecture, including all those genetic contributions to the trait (Goddard et al., 2016). The challenge of dissecting complex traits does not stop at simply understanding and describing loci and candidate genes that govern traits, on the contrary, it goes further. Being able to use the molecular markers associated with QTLs in order to select individuals with greater effects, and thus, from the genotype, to predict the phenotype. Approaches such as marker-assisted selection (MAS) (Francia et al., 2005) and more recently genomic selection (GS) (Crossa et al., 2017) have made progress in this regard for plant species of economic interest (Foolad and Panthee, 2012; Zhao et al., 2014)

Quantitative genetics is the branch of genetics that studies the factors that govern complex traits (Hill, 2010). This is supported by the development of different strategies that use models based on parametric and non-parametric statistics, as well as methodologies for genotyping and obtaining polymorphisms in the study of populations, and phenotyping of the trait, to identify the regions of the genome that are associated with the phenotype of interest (Hill, 2010). The most popular strategy is genetic mapping by QTL linkage and more recently association mapping or AM (Breseghello and Sorrells, 2006).

A decade ago, the greatest challenge we had in QTL identification studies in both, plants, and animals, was to increase the number of polymorphisms obtained by genotyping in the populations, in order to capture all the allelic variants of the loci that govern the complex traits. However, thanks to the advances in next-generation sequencing technologies, this limitation has been overcome (Elshire et al., 2011). However, advances in genotyping and phenotyping techniques have not been synchronous. This has generated an urgent need to achieve robust, precise phenotypes and, if possible, obtained under different environmental conditions, since the complex traits are highly dependent on the environment.

This review presents the concepts and theorical principles of the quantitative genetics of complex traits in plants. Also, the current methodologies to decipher the genetic architecture of this type of traits are described, exposing their benefits and limitations. Finally, we discuss the challenges and the opportunities ahead in the field of studying complex traits in plants.

HISTORIC BACKGROUND

From a historical perspective, studying the genetics of heritable traits has two distinct paths according to the number of loci controlling the character, monogenic or polygenic. A division that reflects two distinct schools of geneticists since the early 1900's (Robinson, 1996). The term for simple inheritance pattern -*monogenic*- derives from classical Mendel's work with *qualitative* plant characters whose variation can be rated as present or absent. Given that each gene could represent two states, dominant or recessive, when an individual that has the same alleles, both dominant or recessive is considered homozygous, while an individual with one dominant allele and one recessive allele is heterozygous. The identification of monogenic traits in discrete categories, represents the underlying genotype and facilitates their study (St.clair, 2010). Traits such the monogenic potato resistance to some pathotypes of *Synchytrium endobioticum*, conferred by the *Sen3* locus (Bartkiewicz et al., 2018) and the rice red seed color (Waghmode et al., 2017), are just two typical examples of agronomical monogenic traits and they are also referred commonly as mendelian traits. A distinct approach was considered when many genes contribute to one phenotypic trait - *polygenic*-, the characters that are *quantitatively* variable, those traits are not clearly separated in discrete classes. Geneticists interested in quantitative traits such as Charles Darwin, Thomas Huxley, and Francis Galton used continuous scales of measurements that when plotted, produced a bell-shaped curve, or normal distribution. Pearson later developed regression and correlation techniques for analysis of quantitative traits (Lynch and Walsh, 1998). In the early 1900's the Mendelian school was working on single-gene characters, while the school of biometricians was dealing with quantitative traits (Lynch

and Walsh, 1998; Robinson, 1996). Among those early works in quantitative traits, the study of wheat seed coat color inheritance patterns by Nilsson-Ehle in 1908 (Nilsson-Ehle, 1908), revealed a complex or 'non-mendelian' pattern, representing a trait value that varies continuously with contributions by multiple loci (Meunier, 2016).

The conceptual framework of quantitative genetics was developed after Fisher in 1919, who deduced that multiple loci contribute to one trait, producing continuously varying quantitative phenotypes when measured in populations (Boyle et al., 2017). The 'infinitesimal model' developed by Fisher states that when a large number of loci are associated to a trait, each loci contributes individually in an additive manner (Boyle et al., 2017; Fisher, 1919). At the time, was not possible to know how many genes were responsible for a specific trait value. A contrasting model was proposed by Wright (1931), whose premise indicates that complex traits are the outcome of multiple loci interacting, and the effect varies importantly according to the genetic background. Both models are considered nowadays relevant in understanding the genetic architecture of complex traits.

During the second half of the XX century, the understanding of DNA as genetic unit, and the advances in molecular biology techniques allowed progress in the field of quantitative genetics. The last few decades have provided new technologies and methodologies, such that nowadays it is understood that complex genetic traits, and their corresponding phenotypes- are those governed by multiple *loci*, their interaction, the environment, and the genetic context (Goddard et al., 2016).

GENETIC ARCHITECTURE: STUDY OF THE COMPOSITION AND DYNAMICS OF COMPLEX TRAITS

The modern field of quantitative genetics uses a collection of approaches aiming to understand the nature of genetic variation underlying quantitative traits (Bazakos et al., 2017). The objective of studying a trait genetic architecture largely depend on whether it is a monogenic or polygenic trait. This, due to the variability in the way genes contribute to the phenotype, whether it is a single gene or multiple genes (Hansen, 2006). In the case of monogenic traits, the genetic architecture seeks to identify the functions of individual genes and the pathway that controls phenotypic variation (Hansen, 2006). In contrast, in a polygenic trait, the study of its genetic architecture seeks to identify the number of genes involved, the contribution of the different alleles of these genes to the trait, and the genetic interactions between them (Sella and Barton, 2019).

The studies of the genetic architecture of complex traits increased in the 1970s, with the rise of molecular biology tools and the development of genotyping techniques based on the detection of DNA polymorphisms. These made it possible, through molecular markers, to develop linkage maps and to identify the QTLs of interest (Edwards et al., 1992; Remington, 2015). The first studies that were carried out on QTL identified chromosomal regions with great effects on agronomic traits such as grain yield in corn (Edwards et al., 1987), insect resistance mechanisms in tomato (Nienhuis et al., 1987) and the content of soluble solids in tomato fruits (Osborn et al., 1987), among others.

Currently, the study of genetic architecture continues to be oriented towards the identification of QTLs associated with the phenotypic variability of the traits of interest. It integrates approaches such as linkage QTL mapping and association mapping. Successful cases of QTL mapping for yield have been identified for species such as rice (Zaw et al., 2019), soybean (Diers et al., 2018), and wheat (Li et al., 2019) among others. QTL mapping identified tolerance to abiotic stress in *Arabidopsis thaliana* (Thoen et al., 2017), cold tolerance in sorghum (Marla et al., 2019), and iron deficiency in soybean (Assefa et al., 2020). Also, resistance to biotic stress, such as disease resistance in *A. thaliana* (Rajarammohan et al., 2017), cassava- *Manihot esculenta* (Soto et al., 2017) and corn-*Zea mays* (Ju et al., 2017), among many others.

The increasing interest of understanding quantitative traits in multiple fields of biology such as evolutionary biology, plant physiology, crop breeding, among others implies that the extent of detail obtained for a quantitative trait can vary depending on the purpose and field of study. Within the context of evolutionary biology, genetic architecture is a tool to understand the genetic basis of adaptive traits, the phylogenetic constrains of polymorphic traits, and the genetic variation that arises from evolutionary forces such as selection and random drift (Morris et al., 2019). It also reveals the mechanisms for speciation as consequences of inbreeding and outcrossing, and develops predictive models for evolutionary change (Lynch and Walsh, 1998). In the field of plant physiology, genetic architecture has made possible to describe the genomic regions involved in the regulation of physiological processes such as photosynthesis and respiration (de Oliveira Silva et al., 2018).

This review considers four practical aspects that build up the conceptual framework of genetic architecture with focus on crops and plant breeding (Fig. 1). The first aspect is identifying the number and distribution of loci underlying the QTLs associated with the trait of interest, the gene content or copy number is useful for determining a genotypic value. The second aspect are the allelic effects, given by the additive genetic variance and the dominance genetic variance for each locus. Both genotypic values and variance partitioning are computed using statistical models accounting for all the loci and the interactions involved in the trait. The third consideration are the gene interactions, and the dynamics of the identified loci/genes. In some cases, two traits governed by different QTLs, have a genetic correlation due to linkage or pleiotropy. When multiple loci are linked, they are in physical proximity in the same chromosome,

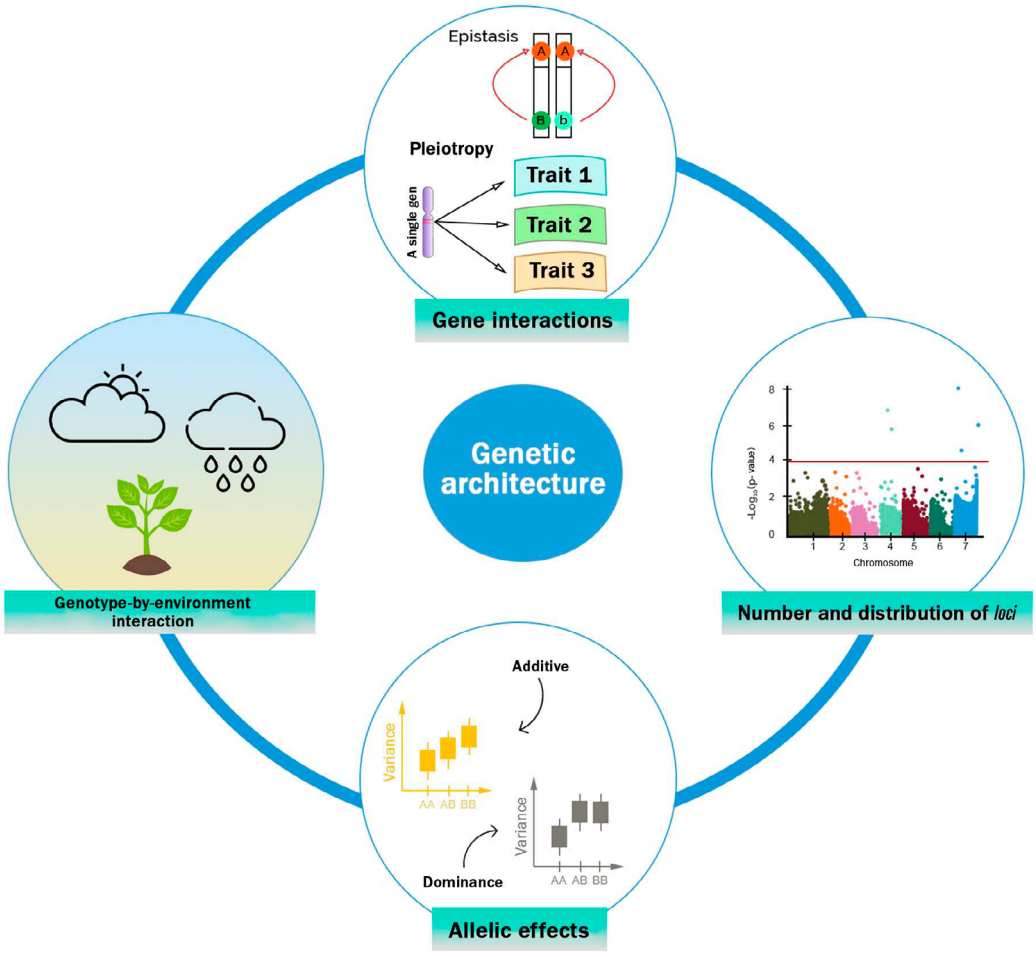

Figure 1. Composition of the genetic architecture of a complex trait. Components of the genetic architecture of complex traits in plants explain the complexity of genetic architecture in plants and how multiple factors interact to influence trait expression These components include the influence of environmental factors and phenotypic plasticity, gene interactions such as epistasis and pleiotropy, allelic effects related to dominance and additivity, and the number and distribution of *loci* contributing to the trait.

by using repeated cycles of meiosis (selfing) those linked traits can be separated. On the other hand, pleiotropy of two traits controlled by the same loci cannot be changed by repeated cycles of meiosis and often has a physiological basis (Bernardo, 2010). A different type of gene interactions is epistasis, that occur between genes located at different loci (Yadav and Sinha, 2018). Unraveling the effects of epistatic and pleiotropic genes involved in genetic architecture goes beyond QTL linkage mapping and association mapping approaches. These techniques hardly detect interaction between genes, since they do not map a specific gene, but rather a region in which many genes are located (Flint and Mott, 2001), thereby representing a challenge within the techniques presented here.

The fourth consideration is the interaction of genotype and environment (G x E), since the effect of the environment on complex traits has been one of the key points in the dissection of the genetic architecture of complex traits in plants (Diouf et al., 2020). The ability of the genotype to give variability to phenotypes in response to the environment under which it is influenced, which is known as genotype-by-environment interaction (G × E) (Zakir, 2018).

QTL IDENTIFICATION: CORNERSTONES OF THE GENETIC ARCHITECTURE

QTLs are genomic regions (loci) that correlate with the variation of a quantitative or polygenic trait in a population of individuals of a species. The identification of QTL has been relevant for the scientific community because it provides: 1) a way to genetically dissect the quantitative variation that governs the traits of interest, 2) are informative about the selection of traits and parents in breeding programs, and 3) it allows to propose hypotheses about candidate genes that control a trait of interest (Hill, 2010). The approaches for QTL identification are based on associating the quantified

phenotype (phenotyping) in a specific population with genetic markers/variation (genotyping) present within the individuals in the study population(s). The most relevant considerations of each of these processes are presented below.

Genotyping

After the explosion in the knowledge of both the physical and chemical properties of the DNA molecule between the 50's and 70's, and consequently the recognition of the existence of differences in DNA sequences between individuals of a population (polymorphisms), a new era was born for the study of genetic architecture. Molecular markers are defined as a particular segment of DNA that differs between individuals at the genome level, with a known location on a chromosome, and that can be associated with a particular gene or trait (Hayward et al., 2015).

Molecular markers can be in linkage disequilibrium (LD) or segregating non-independently with the loci that influence the trait of interest. It is expected that the greater the linkage between them, the more frequently they segregate with the values of the phenotype of interest. While the unlinked markers will not show a significant association with the phenotype (Gupta et al., 2005).

The first studies on complex traits in plants were carried out by Sax (1923) in beans, whose available genetic markers at that time where morphological, histological, and biochemical, and were obtained with laborious process. Sax methodologies were the basis for establishing experimental QTL mapping (Sax, 1923).

Later on, Restriction Fragment Length Polymorphism (RFLP) were the first type of molecular markers distributed along the genome with sufficient density, that permitted the beginning of identification and characterization of QTL's (Lander & Botstein, 1989). Besides RFLPs, Random Amplified Polymorphic DNA (RAPD) and simple sequence repeats (SSR's) were the methods available and used during three decades (70's to 90's) for developing genetic maps and QTL mapping (Williams et al., 1990). These types of molecular markers had several limitations, they were obtained in low quantities, the distribution in the genome was limited and not of highly density, also the position in the genome was unknown (or anonymous).

The development in the early 2000's of Next Generation Sequencing (NGS), reduced dramatically the cost and time of sequencing, and equally important allowed the massive and simultaneous detection of thousands of single nucleotide polymorphisms (SNP's). SNP's markers are widely distributed along the genome, and can be deployed in *n* numbers of individuals representing populations (Elshire et al., 2011). When a reference genome is available for the species of interest, the SNPs are mapped to an exact location in the genome, with non-anonymous markers.

Several genotyping methods have been developed and their use would vary according to the stage in the study of the crop/species, and available funding. Typically, *de novo* SNP identification pipelines use NGS raw read data, quality control, filtering and trimming, mapping or alignment to a reference genome, SNP calling, filtering and validation which requires several steps of bioinformatic processing (Pavan et al., 2020). Currently the most used methods are reduced representation sequencing technologies including Genotyping by Sequencing (GBS) (Elshire et al., 2011), Restriction Site Associated DNA Sequencing (RAD-seq), Specific Locus Amplified Fragment (SLAF), T*ype IIB Endonucleases Restriction-site Associated DNA* (2b-RAD) (Davey et al., 2011). These methods use restrictions enzymes that fragment the genome and allow adapter ligation for sequencing, they are popular because have reduced cost, and can be customized with different combination of enzymes (Pavan et al., 2020). Some limitations from NGS technologies are associated with labor-intensive library preparation, genotyping errors in low coverage runs, which can be improved using deeper sequencing or reducing multiplexing. NGS data analysis is complex because of missing data and sequencing errors, and imputation models could introduce bias, also heterozygote miscalling produce incorrect genotype-phenotype associations (Pavan et al., 2020; Rasheed et al., 2017).

In about 25 crops with extensive genomic resources, genotyping SNP arrays have been developed, using a chip manufactured from companies such as Affymetrix and Illumina, that contains probes to detect from 3K to 5 million SNPs in one sample (Pavan et al., 2020; Rasheed et al., 2017). The use of SNPs arrays is cost-effective in large breeding programs, produce accurate genotype call, with reduced missing data, and is computationally faster. They have been mainly used for genotyping polyploid species with better accuracy than NGS (Rasheed et al., 2017).

Phenotyping

The concept of *phenotype* refers to the actual measurable trait value for an individual, it represents both the genetic and environmental effects. The phenotypic variation in organisms arises at multiple biological levels, structural or functional, and is dynamic along individual developmental time, representing a challenge for a single definition of phenotype.

In plants, there are different methods and approaches for phenotyping depending on the spatial scale and the species such as crop phenotyping in the field, medium or large-scale greenhouse experimental units, and smaller scale with artificial media, substrates, or special designs to facilitate data acquisition. In each approach and scale, multiple considerations and limitations (technical or financial) can constrain data acquisition and experimental design. In all cases, the objective is to understand the plant- environment

interactions and quantify the trait of interest (Pieruschka and Schurr, 2019).

The concept of functional units for a trait- "phene"- refer to a set of parameters that provide a function, it has been particularly useful to select root system architecture *phenes* for adaptation to low fertility soils (Lynch, 2011; York, 2019). Phenotyping is the measure of multiple characteristics in the organism, it can be performed at multiple scales: field level, individual whole plants, at organ level or at subcellular level. It can include the biochemical profile, metabolites, volatile compounds, RNA profile, or protein profile (Yang et al., 2020). Field phenotyping is the basis of plant genetics and crop improvement, is the scale that provide more realistic information regarding crop performance, testing plant genetic material in soils and multiple geographic locations captures the genotype- environment variability to the external conditions, biotic and abiotic stresses that are dynamic in space and time (Pieruschka and Schurr, 2019). Traditionally, field phenotyping relies mainly on the visual observations, breeder criteria, and human experience for a few targeted traits for selection and directed crossing. Grain yield is the main traditional quantitative value for field crops, it encompasses all the developmental, physiologic, and metabolic aspects of a genotype (Bucksch et al., 2014; Kirchgessner et al., 2016). Other measurements at a single timepoints require destructive or invasive methods that are labor intensive (Biomass dry weight, Digging roots). Non-destructive measurements for physiological canopy traits

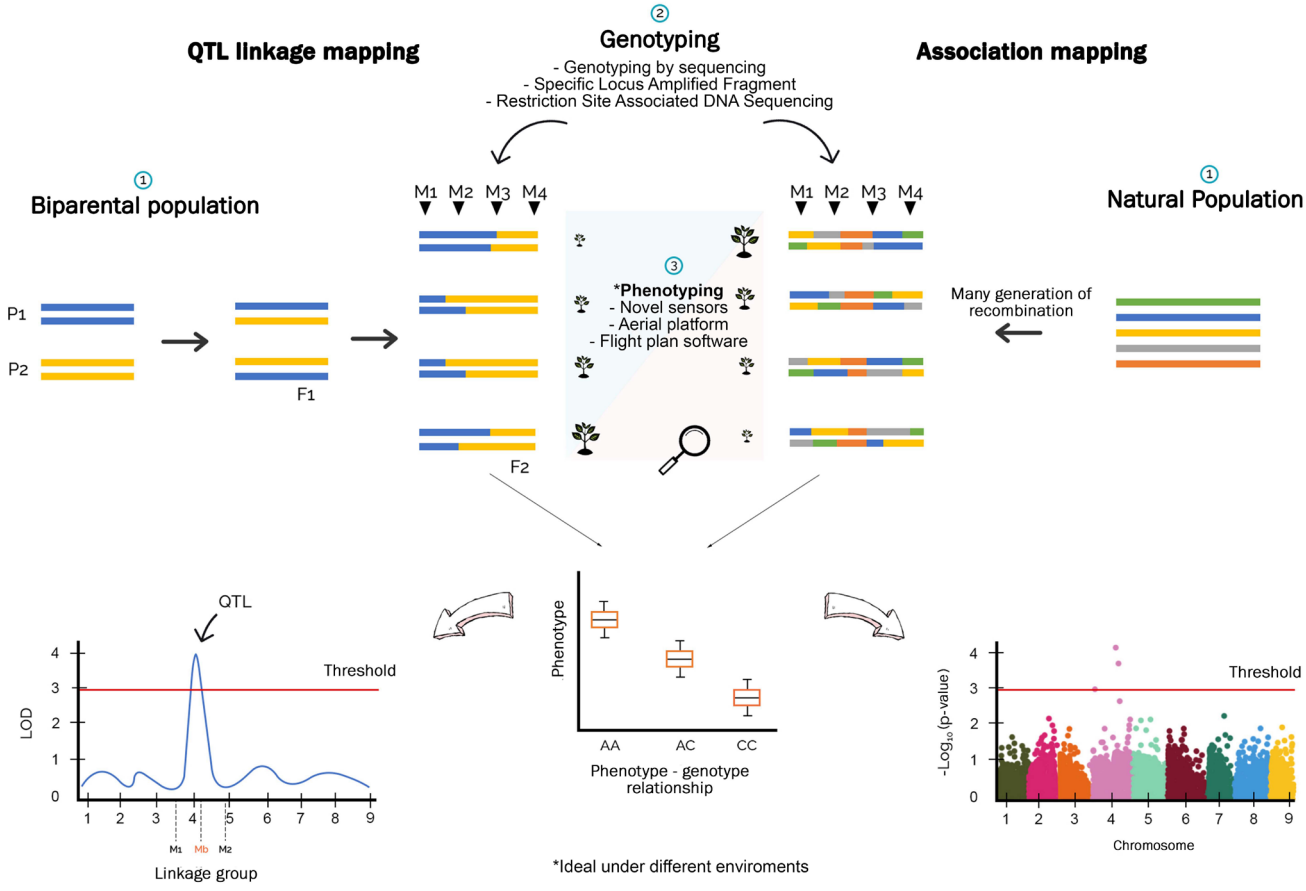

Figure 2. Graphic summary of the approaches used to identify QTLs. Linkage mapping approach shown on the left and the association mapping (AM) approach shown on the right. It depicts the essential steps involved in identifying QTLs, including the type of population used for linkage mapping and AM (1), genotyping using molecular markers (M1, M2, M3, M4) (2), and phenotyping (3). In both approaches, the relationship between phenotypic and genotypic data is crucial and is depicted using a box plot. The linkage mapping identifies QTLs in an interval of flanking markers (M1 and M2) and the peak marker (Mb) within a linkage group on a genetic map. In contrast, QTLs in AM are located at a specific position in the genome. The figure also emphasizes that a significance threshold must be set (represented by a red line) for both approaches.

such as transpiration and photosynthesis are measures that occur one plant at the time, requiring multiple sets of equipment and operators.

Greenhouse or growth chamber experiments have a comparative advantage by reducing the dimension of external variation that occur in the field, and controlling environmental conditions, permits to study aspects of the phenotype (Metabolites, RNAseq), as well as increasing growing cycles off-season (Rapid breeding). Specialized growth systems such as rhizotrons allow to acquire digital imaging without perturbing the root system for phenotypic quantification (Bontpart et al., 2020; Bucksch et al., 2014). Other cases that could not be managed in field situations include controlled drought experiments requiring multiple measures of soil water content along the plant cycle (Khan et al., 2020).

While powerful as experimental setting for phenotyping, the genotypes selected based on artificial growth systems could perform different under field conditions. For example, a comparison of *A. thaliana* grown in controlled conditions vs. field conditions, showed a difference of leaf size and chlorophyll content among different conditions (Fiorani and Schurr, 2013). In fact, a characteristic of quantitative traits is the phenotypic plasticity and G X E interactions. Therefore, performing field trials is considered more relevant and critical for quantitative traits that are more influenced by the environment.

Independent of the scale of the experimental system, there is a growing need for optimizing data acquisition that reflects the nature of the quantitative trait variation and that is useful for breeding programs. One source of genetic diversity for crop improvement are gene bank collections, and combining high throughput genotyping with selective phenotyping with a subset of 290/962 sorghum accessions were phenotyped in quantitative traits such as yield, plant height and more (Yu et al., 2016).

The limitations and bottlenecks of field phenotyping and other settings are being addressed by the emerging discipline of *phenomics*. The goal of phenomics is to characterize accurately a full set of phenotypes of an individual, in a fast way (high throughput), achieving large scale and automation to reduce manual labor. There is a bloom of phenotyping platforms developed using tools such as digital imaging, spectroscopy, remote sensing (Yang et al., 2020), digital sensors, robotics and high throughput data centers (Pieruschka and Schurr, 2019). Some platforms are commercial and the software of analysis proprietary, requiring licenses, while other initiatives for open source hardware and open source software that can be customized (Yang et al., 2020). The outcomes of phenomic experiments can be stored and shared in open source platforms, with good practices of data management (Wilkinson et al., 2016) including standardized information storage and interoperable data usage (Krajewski et al., 2015).

However, the development of phenomic platforms is limited, since the technical expertise to install, maintain and operate is sophisticated and can result expensive for individual research groups. Despite the challenge, a goal for breeding programs in developing countries is the implementation of accessible, precise, fast, and cost-permissive phenotyping.

METHODOLOGICAL APPROACHES FOR QTL IDENTIFICATION

QTL linkage mapping

QTL linkage mapping is a statistical method that compares phenotypic data in quantitative measures, with genotypic data (molecular markers), considering the recombination frequencies in a genetic map (linkage map). This method allows positioning the phenotypic variation in a specific region within a chromosome (Giri et al., 2018). The phenotypic mean values of the quantitative traits will be different with different genotypes at the marker locus (Mackay, 2001). Linkage mapping is based on the premise that the markers associated with genetic loci responsible for quantitative variation are segregating during meiotic chromosomal recombination (Boopathi, 2020). During recombination the genes and markers that are tightly linked, or in linkage disequilibrium, will be passed to the progeny with more frequency than those genes and markers that are further apart (Boopathi, 2020).

Populations used for QTL linkage mapping.

Linkage mapping identification is performed in structured populations, using contrasting parental lines for the trait in directed crosses (Rajpal et al., 2016). A linkage mapping population size range from 50 to 250 individuals (Collard et al., 2005), or even more (Li et al., 2010). There are different ways to build up a linkage mapping population, and the resolution of the QTL interval is contingent to the recombination events within the population (Mackay, 2001). Among the most used populations are the F1 segregants of full siblings, F2 populations derived from F1 hybrids, and recurrent backcrossing (BC) from the hybrid F1 with one of the parental lines, in order to increase the effect of the QTL (Collard et al., 2005; Mackay, 2001).

Other type of populations widely used that require successive self-pollination are Recombinant Inbred Lines (RIL), where a F2 population undergo successive and repetitive cycles of self-pollination to produce endogamic lines (Ordas et al., 2010). By using RIL advanced generations, there will be more recombination events and can increase the precision mapping (Mackay, 2001).

Near-Isogenic Lines (NIL) populations are obtained by crossing a line with the trait of interest -or molecular marker- and continuous backcrossing with a donor line. NIL populations have identical genetic background with the exception of the small interval of introgression representing

few loci (Szalma et al., 2007). One of the advantages of NIL populations, is the possibility of multiple measurements of the trait in different environments or locations, using the same genetic background.

QTL identification by linkage methods

The methods for the detection of QTL by linkage are based on detecting significant correlations between the variation of the phenotype, given by continuous values, and the genotype (polymorphism or molecular markers) in a segregating and contrasting population for the trait of interest. The three most popular models for these analyzes are single marker analysis, Simple Interval Mapping (SIM), and Composite Interval Mapping (CIM) (Liu, 1998; Tanksley, 1993).

Single marker analysis is a simple method that allows identifying QTLs associated with a single marker, does not require the generation of a linkage map and can be performed using basic statistical analyzes such as t-test, analysis of variance (ANOVA) and lineal regression. This is most often used because the marker's coefficient of determination (R^2) explains the percentage of phenotypic variation that arises from the QTL. Since this analysis does not use a genetic map, it does not provide information on the chromosomal position of the identified QTL (Collard et al., 2005).

On the other hand, the interval mapping approach uses genetic maps as a positional reference of the polymorphisms for the detection of QTLs. In addition, the result is presented with a polymorphism or peak marker (QTL), which will present the highest level of significance, and two flanking markers, one to the right and one to the left of the peak marker, generating an interval where it is expected that the candidate loci / genes responsible for the phenotypic variation of the trait under study will be found (Collard et al., 2005) .

The SIM, tests for the presence of significant QTLs at many positions between two mapped markers. Therefore, the most probable position of a QTL and the size of its effects are estimated with greater precision than in single point analysis. The presence of a significant QTL is calculated using the Logarithmic of Odds Score. This statistical estimate indicates the probability that a significant QTL exists or not at that position. The scores of each marker are plotted along the genetic map, those that exceed the threshold of significance, example LOD 3 (1000 times more probable) then suggest the presence of a QTL in that position. This approach has some limitations, among them that it can be difficult to separate the effects of nearby QTLs, as well as that the position of one QTL can be influenced by other QTLs (Rajpal et al., 2016).

An evolution of SIM analysis, which is more robust, is the CIM method. The basis of this method is to isolate individual QTL effects by combining interval mapping with multiple regression. CIM controls genetic variation in other regions of the genome, thus reducing background "noise" that can affect QTL identification (Jansen, 1993). For this, cofactors are selected, these are representative markers of the identified QTL. Cofactors make it possible to eliminate variations in other QTLs located in other parts of the genome, thereby reducing the effect of those other QTLs (Jansen, 1993). The number of markers that are selected as cofactors is essential because it will affect the detection of QTL. If few cofactors are selected, the variation of the other QTLs will not decrease, but, if too many cofactors are selected, the detection power of the QTL will decrease (Silva et al., 2012).

Regardless of whether the model used for the detection of QTL is SIM or CIM, a typical result is a graph in which the molecular markers are positioned by linkage group on the "x" axis, while the "y" axis represents the values of the LOD test statistic, and their threshold are presented. The point(s) where the molecular markers reach the highest LOD values are identified as the most likely positions for the presence of a QTL. The point at which the marker reaches a high LOD value is called the "peak marker", when it exceeds the specified level of significance, it indicates that the QTL is statistically significant (Collard et al., 2005). The determination of the thresholds of significance is carried out a priori by the investigator, normally using a LOD of 3, or by means of permutation tests, generally between 500 and 1000 (Collard et al., 2005). These tests evaluate the phenotypic values given for the individuals of the population with respect to all the markers, in order to determine robust levels of significance of the marker-phenotype associations.

The QTL linkage mapping has been a widely used tool and has allowed the identification of QTL for important and diverse traits of interest in plant species and despite its utility, for linkage QTL mapping some disadvantages are recognized. The first is the population design and development. The need to obtain a genetic map and the limited number of recombination present in the mapping population, which in turn will result in a low QTL detection resolution. Finally, that the result is expressed in an interval measured in cM, where depending on the genomic region they can physically reside in the number of genes. This last limitation has been overcome in recent years, with the possibility of obtaining highly dense genetic maps with genotyping from NGS, which has finally led to the reduction of the length of the QTL interval to a few cM, increasing the resolution in QTL detection (Jaganathan et al., 2020; Soto et al., 2015). As an alternative to linkage QTL mapping, the AM approach was generated (Risch and Merikangas, 1996). A more robust method than linkage mapping and which has made it possible to overcome the limitations of the latter (Zhu et al., 2008).

Association mapping

Association mapping is a robust method that leverage on the historical recombination events in natural populations

(Mackay, 2001). It can include a wide range of genetic variation from diverse collection of families, inbred, races and multiple sources of germplasm. It offers three advantages over linkage mapping: better mapping resolution, more allele numbers, and the option to study simultaneously several genomic regions without using genetic maps or spending time in directed crosses for mapping populations (Zhu et al., 2008). However, the mapping resolution power is influenced by linkage disequilibrium at the relevant genomic region, the species, and the type and density of molecular markers (Risch and Merikangas, 1996; Zhu et al., 2008).

It was first introduced for finding the genetic basis for human diseases (Lohmueller et al., 2003), and later expanded successfully to other animal species and crop plants (Ersoz et al., 2007). The approach of *Genome Wide Association Studies* (GWAS) in some cases produce better results in plants than in human quantitative traits (Anderson et al., 2019). Generally the studies performed in humans only explain a small fraction of the phenotypic variability, while in plants it has explained a large part of the phenotypic variability of the traits, possibly explained by a higher hereditability in plant traits (Korte and Farlow, 2013).

In association mapping it is critical to evaluate the population structure, and the degree of genetic differentiation among groups, to avoid false marker-trait associations. It is also critical to test the level of relatedness among pairs of individuals in a sample (Abdurakhmonov and Abdukarimov, 2008). This can be achieve calculating the similarities between genotypes or individuals, using the whole genome wide SNPs, in a matrix called genomic relationship matrix (GRM), build from a pairwise kinship analysis among individuals and its genomic inbreeding coefficients (Villanueva et al., 2021). Then, with the information obtained from phenotypic and genotypic data, it is proceeded to quantify the LD and population structure and perform the statistical analysis for associating the markers with the phenotypic value.

Populations used for association mapping

Contrary to the populations used in QTL linkage mapping, AM requires populations with little or no population structure (Pritchard et al., 2000). One type of population used in AM are *unstructured populations*, where the kinship relationship is unknown, as occurs in natural populations or diverse panels. In plants, they are groups of lines of diverse origin that represent the variability of the species under study with historical recombination events, resulting in greater allelic richness (Collard et al., 2005). Compared to populations derived from biparental crosses, natural unstructured populations such as elite cultivars, local varieties, landraces, wild relatives and exotic accessions have greater power to identify QTL (Collard et al., 2005). This approach avoids the need to build mapping populations for each trait of interest and instead uses recombination events that have occurred throughout the evolutionary history of the species. It has been applied in corn, finding candidate genes for starch content regulation (Liu et al., 2016), in *A. thaliana* for traits such as flowering time (Zan and Carlborg, 2018). In rice, to identify candidate genes associated with resistance to drought (Guo et al., 2018), and to improved yield (Yano et al., 2019). In potato, genes related to resistance to *Phytophthora infestans* (Juyo Rojas et al., 2019), among many others.

In contrast to natural populations *multi-parent populations* such as *Nested Association Mapping* or NAM (Gage et al., 2020), and the *Multiparent Advanced Generation Intercross* or MAGIC (Cavanagh et al., 2008), have been designed in order to take advantage of the benefits of both linkage mapping and association mapping. By having multiple parents, the genetic diversity of these leads to a great phenotypic diversity, which increases the allelic diversity and the QTL detection power (Flint- Garcia et al., 2005).

Analytic considerations in Genome wide Association Studies.

The first step for association mapping analysis is a selection of a diverse association panel, it can be a natural population, a germplasm collection with broad genetic diversity or a multiparent population (Ibrahim et al., 2020). The second step is phenotyping the population for the trait of interest under different environments and genotyping to obtain molecular markers, perform the kinship analysis through the GRM and analyzed an eventual population structure. The association among phenotypic variance and each allelic variant in the genome in a diverse panel of individuals is preformed using linear models. Next, the interpretation of the association can lead to identify a narrow genetic region or ideally candidate genes responsible of the phenotypic variation (Alqudah et al., 2020). The linkage disequilibrium (LD) degree is quantified based on the genotypic data (Ersoz et al., 2007) and is specific within the studied populations. It is expected that in an AM approach, blocks in the genome occur with less distance in the genome (Davey et al., 2011).

The degree of resolution and accuracy of GWAS depend on the LD decay in the studied population. The LD decay can be plotted with the "y" axis representing the R^2 values and the "x" axis representing the genetic distance or physical distance of the SNP's along the genome (Alqudah et al., 2020). When the LD decay occurs over a short distance it is expected a high resolution- the capacity of associating a SNP within a causal locus-, in that case it is required to have a high density of SNPs. On the contrary, if the LD decay occurs over a long distance, the mapping resolution is reduced even in the case of having low density of molecular markers (Ibrahim et al., 2020). For example, in Barley when the LD decay is around 5 kbp (Kilobase pair), it is required around ~1million SNPs to cover the genome 5.100 Mb (Mega bases). While when the LD decay occurs at 100

kbp, the analysis can use only 57.000 SNPs (Semagn et al., 2010). The rate of LD decay allow to identify the number of markers required for the study, the value is obtained in a species dividing the genome size by the distance of LD decay (Alqudah et al., 2020). Hence, the number of informative molecular markers depend both on genome size and rate of LD decay.

For AM, the population structure is an important parameter when selecting the work material, since its presence can result in false positives -false trait-markers associations-, especially if the population structure induces LD (large blocks at a great distance, throughout the genome) (Platt et al., 2010). False positives occur when highly significant associations are present between a marker and a phenotype, even though that marker is not physically linked to the causal allele responsible for the phenotypic variation of the trait of interest (Zhu et al., 2008).

There are different approaches that allow determining and correcting the population structure, thereby ensuring that the identified associations are real. Approaches include Principal Component Analysis (PCA), Mixed Model Approach, Structured Association (SA), and Genomic Control (Price et al., 2006), among others.

Currently there are different software that compute GWAS, such as *Trait Analysis by Association, Evolution, and Linkage* (TASSEL) (Bradbury et al., 2007), as well as R package such as *Genome Association and Prediction Integrated Tool (*GAPIT) (Lipka et al., 2012) and the Whole genome association analysis toolset Plink (Purcell et al., 2007), among others. The statistical test that are used to identify the association among phenotype and genotype are based in *p-value*. The *p-value* is plotted in logarithmic scale -$\log_{10}$ e in the "*y*" axes, and the genomic position of the molecular markers in the "*x*" axes, the produced graph is a Manhattan plot (Fig. 2). The genomic regions with a peak above a threshold level are the genomic positions with the strongest signals that contribute to the phenotypic variation (Burghardt et al., 2020). With the purpose of using a p-value that reduces the risk of false associations with multiple testing, statistical test such as *False Discovery Rate* (FDR) (Kaler and Purcell, 2019) are used. However, there are several multiple testing corrections applied to linear mixed models, such as Bonferroni correction among others (Joo et al., 2016). Routinely, SNPs with p-values <0.05 and FDR <0.1 are considered significantly associated, and if they reside in a coding gene, those can be candidate genes responsible for a percentage of phenotypic variability for the trait.

APPLICATIONS IN AGRICULTURE

Several applications in agriculture are promising once a QTL is identified, delimited by molecular markers with a significant association with a phenotype. The routine detection of those markers is known as directed genotyping, which is useful in marker-assisted selection, background selection, marker-assisted backcrossing, and gene tagging. Methodologies for detection of specific alleles of individuals in a population are based on quantitative PCR (qPCR) such as Kompetitive Allele Specific PCR (KASP), a cost-effective method, other PCR-based methods are also available for single marker genotyping (Rasheed et al., 2017).

Marker-Assisted Selection (MAS), aims at detecting markers associated with a QTL to keep the trait of interest in new populations or individuals in early developmental stages. The commercial breeding programs for horticultural crop routinely use MAS in the seedling stages, before planting the material in the field that will be available for breeders for selection. In that way the new germplasm maintains traits that are desirable. A potential limitation is that multiple loci of small effect could not be easily maintained in a population. Although the MAS strategy has been successful in some cases, for traits governed by multiple genes of small effect, this strategy has not been entirely effective.

A different and more accurate approach is genomic selection *Genomic Selection* (GS), a strategy for selection of individuals based on predicting the phenotype using markers that represent the whole genome (Jannink et al., 2010). The rationale for GS, also known as genome wide selection, is to omit significance test (such as the ones used for QTL mapping or association mapping) and use a large set of random markers along the genome in marker-based selection. This type of selection maintains multiple QTL loci with small contributions, rather than focusing on one single QTL with a large effect (Bernardo, 2010). It requires using cheap genotyping methods that can be followed in multiple growth cycles. Calculating the kinship and GRM then effects of markers in genome wide selection can be performed with BLUP, a simplification calculation when the variance and random effects are known. Training populations are used for modeling the genotype with the phenotypes measured, to produce a predictive value for the phenotype in new individuals that are only genotyped. The model for genomic selection can be more robust using prior information, and incorporating it with Bayesian methods (Bernardo, 2010; Jannink et al., 2010).

Both MAS and GS can be subsequent and complementary approaches to QTL linkage and GWAS analysis, insofar as the SNPs/loci associated with the trait of interest can strengthen the prediction by being incorporated into the phenotype prediction equation.

Candidate Gene Approach

Candidate gene (CG) approach assumes that a gene with specific function is contributing to the phenotype of a trait. The SNPs polymorphisms within the candidate gene can occur in the coding part of the gene or in the regulatory elements, for plant geneticist CG are genes with polymorphisms linked to a QTL or statistically associated with phenotypic trait variation (Pflieger et al., 2001).

A candidate gene can be identified based on literature background, the outcomes of QTL mapping, fine-mapping, or GWAS in a different species where the candidate gene is proposed. The hypothesis of a gene associated with a trait can arise from prior information for gene structure, known expression patterns, gene product, and regulatory elements that can be found in databases such as Entrez Gene and Ensembl (Patnala et al., 2013). An additional source of information is provided with tools such as RNAseq, that provide expression profile to specific experimental or environmental conditions (Giri and Mohapatra, 2017). Another tool for selecting CG is comparative map analysis, by using positional information and comparative mapping from another related species. Several grasses have highly conserved gene order, thus for species such as rice, millet, sugarcane, sorghum, maize, and oat, proposing candidate genes based on syntenic regions and QTLs, is a way to economize resources and optimize time for analysis (Bernardo, 2010; Pflieger et al., 2001). The main advantage of CG approach is that the analysis requires less time and investment, given that they test previously reported genes that are potentially valuable. However, it is also a biased approach towards the hypothetical association of a known gene, omitting genes or genetic regions that can also contribute to the phenotype.

Once the CG are selected, the SNPs in the genes need to be screened and re-analyzed with independent association experiments. fine- mapping experiments, or correlations among the genetic polymorphisms and trait values in a different organisms (Patnala et al., 2013; Pflieger et al., 2001), or differential gene expression analysis.

CHOOSING THE BEST ONES: PROMISING CANDIDATE GENES

Typically, QTL mapping and AM outcomes correspond to an interval in the genome that contains several genes, but not all the genes in the region are the responsible for the association of phenotype-genotype. Hence, a further selection is necessary to narrow-down the most promissory ones.

The alternatives for selection of candidate genes that directly co-localize with the QTL, as well as the co-segregation of markers in the candidate loci, should fulfill at least one of the following criteria: 1) Those who explain a high percentage of phenotypic variance, indicating a major role in the genetic control of the trait. 2) Genes localized in stable QTL, that is that the gene has a strong effect with low environmental effect. 3) Genes that code for proteins with predicted functions or predicted domains relevant to the trait of interest. 4) Genes with mRNA profile that is differential in contrasting phenotypes for the trait.

After a curated selection of promising candidate genes, the next step is the confirmation of the effect in the trait of interest. The polymorphism can be found in a panel of accessions, different to the population where the QTL was initially found and perform a new round of phenotyping and directed genotyping. In all cases, will be required an experimental validation of the candidate gene for ultimate confirmation.

VALIDATION OF CANDIDATE GENES

When a gene is a candidate, via QTL mapping, association mapping, or comparative CG selection, it is relevant to validate their function, since that is the gold standard for a trait and the phenotype are controlled by that gene product.

One approach is by physiological analysis associated with the gene-product of the candidate gene. An important parameter that can be quantified is the mRNA and the dynamic changes in abundance, this was traditionally done with Northern Blot and currently with quantitative RT-PCR or RNAseq analysis. In addition, measures of protein level or enzyme activity can be contrasting in lines with or without the polymorphism in the candidate gene (Pflieger et al., 2001).

A different approach is the genetic transformation with gene gain or loss of function, which are the gold standard for attributing the function to a gene. For gene silencing, it can be used *Interfering RNA (RNAi)* which uses an RNA template, the *Virus-Induced Gene Silencing (*VIGS) which is useful in solanaceous species (Unver and Budak, 2009) and less used in other taxonomic groups (Robertson, 2004), or with artificial microRNA. Another technique uses the genome editing platform CRISPR / Cas9 (Ran et al., 2013). All these genetic transformation methods require an established protocol of successful transformation, in *vitro* tissue culture and *in vitro* regeneration, all of which represent an extended time, as well as high cost and limitation in the technology access.

CHALLENGES

The study of genetic architecture provides the principles for understanding the genetic basis of complex traits. Current genotyping technologies are highly accurate, thus allowing the identification of QTLs associated with the trait of interest in crops. Fine mapping, validation, and confirmation of candidate genes within QTLs as causal genes responsible for the variation of the trait of interest is a subsequent challenge.

Although the methodologies presented here allow the identification of QTL for complex traits, there are still methodological challenges for QTL linkage mapping and association mapping. One is the difficulty in studying the interactions that takes place in the genetic architecture, such as complex gene by gene or gene by environment interactions, which makes determining all the variables involved a very difficult and sometimes impossible task to accomplish (Tam et al., 2019). With GWAS CG controlling several traits of interest have been identified. However, a gap remains of the biological functions of these loci. It is not known exactly which variant of the identified loci give rise to

the association or if these loci alter the genetic regulation of other genes (Gallagher and Chen-Plotkin, 2018). Epigenetic and expression studies could provide the necessary information to know in a functional way the effect of the loci identified in the trait of interest (Albert and Kruglyak, 2015).

Finally, phenotyping also represents a challenge, to ensure accuracy and precision of the phenotypic data obtained. Accuracy ensures data reliability (Albert and Kruglyak, 2015), and is fidelity between the phenotypic values taken and the true value of the phenotype. While precision refers to the variability in the measurements between the different repetitions. If the phenotypic data are not reliable, the genotype-phenotype relationship can become imprecise. Besides technical challenges, there are logistical and financial constrains to access phenomic platforms. The obstacles to access high-throughput screening methods, precision phenotyping, and phenomic platforms is more unequal in middle and low-income countries, which are the locations where their local crops and breeding programs will benefit the most.

FUTURE PERSPECTIVES

In recent years there has been a high increase in the identification of QTL involved in diverse traits of agronomic importance, in a wide variety of crops. It is relevant that the QTLs found in previous studies, can be validated, and advanced along the new advanced lines. With the creation of public databases, it is possible to find genetic and genomic information, among which the QTLs identified in different studies are included. Among the available databases we can mention *Cassava Genome Hub*, *Legume Information System*, *Maize Genetics and Genomics Database*, *Solanaceae Genomics Network*, among other databases. Such networks of multiple research centers help to establish collaborations in infrastructure or resources and build up long-term germplasm and genomics exchange platforms, for agriculture and crop improvement.

As mentioned, complex plant traits have agronomic and economic importance. Therefore, they are traits that are widely studied in genetic breeding programs. In a world with a constant population growth and climate change, efforts are required to improve crop production to provide resistant and resilient crops with higher yields. Increasing productivity in a context of limited arable land, changes in weather patterns, low soil fertility, and climatic changes in biotic interactions. The tools and examples of quantitative genetics discussed here can be extended to more local crops to provide elite genetic materials to small-holder farmers, increasing the efforts of developing cost-effective systems for phenotyping in developing countries (Bontpart et al., 2020). Allowing technology transfer, with open-source hardware devices, and open-source software for digital imaging and analysis, democratizing phenomic platforms to improve speed, precision and reliability in data collection and harnessing collaborations to make phenomics a reality for advancing crops resilient and in response to future challenges in agriculture.

AUTHOR'S PARTICIPATION

In this review article, the authors' contributions were complementary and collaborative. Soto-Johana played a pivotal role in the conception and design of the study, outlining the review's scope and objectives. Chivatá-Laura conducted an extensive literature search, systematically collected relevant articles, and managed the reference database. Additionally, Perilla-Laura took the lead in drafting the manuscript and contributed significantly to the synthesis and interpretation of the findings. Both Soto-Johana and Perrilla-Laura engaged in critical discussions and revisions, enhancing the clarity and coherence of the manuscript. Ultimately, all authors provided their final approval for submission.

ACKNOWLEDGMENTS

The authors express acknowledgment to the "Fundación para la promoción de la investigación y la tecnología, del Banco de la República", cooperation agreement 2019 of 2020, to the "Dirección de Investigaciones Sede Bogotá de la Universidad Nacional de Colombia (Hermes (44921)", and to the Fulbright-COLCIENCIAS (Departamento Administrativo de Ciencia, Tecnología e Innovación) doctoral fellowship to LP-H.

CONFLICT OF INTEREST

The authors of this review declare that they have no conflict of interest that could influence the objectivity or integrity of the content presented in this review article. There is no financial or personal relationship that could give rise to a conflict of interest with respect to the topics, authors, sources of information or conclusions discussed in this work. This review has been conducted in an unbiased manner and for the sole purpose of providing an objective synthesis of the existing scientific literature.

REFERENCES

Abdurakhmonov, I. Y., and Abdukarimov, A. (2008). Application of association mapping to understanding the genetic diversity of plant germplasm resources. *International Journal of Plant Genomics*, *2008*. https://doi.org/10.1155/2008/574927

Albert, F. W., and Kruglyak, L. (2015). The role of regulatory variation in complex traits and disease. *Nature Reviews Genetics*, *16*(4), 197–212. https://doi.org/10.1038/nrg3891

Alqudah, A. M., Sallam, A., Baenziger, P. S., and Börner, A. (2020). GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: lessons from Barley – A review. *Journal of Advanced Research*, *22*, 119–135. https://doi.org/10.1016/j.jare.2019.10.013

Anderson, R., Edwards, D., Batley, J., and Bayer, P. E. (2019). Genome-Wide Association Studies in Plants. *ELS*, 1–7. https://doi.org/10.1002/9780470015902.a0027950

Assefa, T., Zhang, J., Chowda-Reddy, R. V., Moran Lauter, A. N., Singh, A., O'Rourke, J. A., Graham, M. A., and Singh, A. K. (2020). Deconstructing the genetic architecture of iron deficiency chlorosis in soybean using genome-wide approaches. *BMC Plant Biology*, *20*(42), 1–13. https://doi.org/10.1186/s12870-020-2237-5

Bartkiewicz, A., Chilla, F., Terefe-Ayana, D., Lübeck, J., Strahwald, J., Tacke, E., Hofferbert, H. R., Flath, K., Linde, M., and Debener, T. (2018). Improved genetic resolution for linkage mapping of resistance to potato wart in monoparental dihaploids with potential diagnostic value in tetraploid potato varieties. *Theoretical and Applied Genetics*, *131*(12), 2555–2566. https://doi.org/10.1007/s00122-018-3172-9

Bazakos, C., Hanemian, M., Trontin, C., Jimenéz-Gómez, J. M., and Loudet, O. (2017). New Strategies and Tools in Quantitative Genetics : How to Go from the Phenotype to the Genotype. *Annual Review of Plant Biology*, *68*, 1–21. https://doi.org/10.1146/annurev-arplant-042916-040820

Bernardo, R. N. (2010). *Breeding for Quantitative Traits in Plants* (Stemma Press (ed.); Second Edi).

Cavanagh, C., Morell, M., Mackay, I., and Powell, W. (2008). From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Current Opinion in Plant Biology*, *11*(2), 215–221. https://doi.org/10.1016/j.pbi.2008.01.002

Bontpart, T., Concha, C., Giuffrida, M. V., Robertson, I., Admkie, K., Degefu, T., Girma, N., Tesfaye, K., Haileselassie, T., Fikre, A., Fetene, M., Tsaftaris, S. A., and Doerner, P. (2020). Affordable and robust phenotyping framework to analyse root system architecture of soil-grown plants. *Plant Journal*, *103*(6), 2330–2343. https://doi.org/10.1111/tpj.14877

Boopathi, N. M. (2020). Genetic Mapping and Marker Assisted Selection. In *Springer*. https://doi.org/10.1007/978-981-15-2949-8

Boyle, E. A., Li, Y. I., and Pritchard, J. K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell*, *169*(7), 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19). https://doi.org/10.1093/bioinformatics/btm308

Breseghello, F., and Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) Cultivars. *Genetics*, *172*(2), 1165–1177. https://doi.org/10.1534/genetics.105.044586

Bucksch, A., Burridge, J., York, L. M., Das, A., Nord, E., Weitz, J. S., and Lynch, J. P. (2014). Image-based high-throughput field phenotyping of crop roots. *Plant Physiology*, *166*(2), 470–486. https://doi.org/10.1104/pp.114.243519

Burghardt, L. T., Young, N. D., and Tiffin, P. (2020). A Guide to Genome-Wide Association Mapping in Plants. *Current Protocols in Plant Biology*, *2*(1), 22–38. https://doi.org/10.1002/cppb.20041

Cai, G., Yang, Q., Chen, H., Yang, Q., Zhang, C., Fan, C., and Zhou, Y. (2016). Genetic dissection of plant architecture and yield-related traits in *Brassica napus*. *Scientific Reports*, *6*(21625), 1–16. https://doi.org/10.1038/srep21625

Collard, B. C. Y., Jahufer, M. Z. Z., Brouwer, J. B., and Pang, E. C. K. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, *142*(1–2), 169–196. https://doi.org/10.1007/s10681-005-1681-5

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., Burgueño, J., González-Camacho, J. M., Pérez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., and Varshney, R. K. (2017). Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends in Plant Science*, *22*(11), 961–975. https://doi.org/10.1016/j.tplants.2017.08.011

Silva, L., Wang, S., and Zeng, Z. B. (2012). Composite interval mapping and multiple interval mapping: Procedures and guidelines for using windows QTL cartographer. In *Methods in Molecular Biology* (Vol. 871). https://doi.org/10.1007/978-1-61779-785-9_6

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., and Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*(7), 499–510. https://doi.org/10.1038/nrg3012

de Oliveira Silva, F. M., Lichtenstein, G., Alseekh, S., Rosado-Souza, L., Conte, M., Fuentes V., Silvestre B., Fanourakis, D., Usadel, B., Lopes L., DaMatta, F. M., Sulpice, R., Araújo, W. L., Rossi, M., de Setta, N., Fernie, A. R., Carrari, F., and Nunes-Nesi, A. (2018). The genetic architecture of photosynthesis and plant growth-related traits in tomato. *Plant Cell and Environment*, *41*(2), 327–341. https://doi.org/10.1111/pce.13084

Diers, B. W., Specht, J., Rainey, K. M., Cregan, P., Song, Q., Ramasubramanian, V., Graef, G., Nelson, R., Schapaugh, W., Wang, D., Shannon, G., Mchale, L., Kantartzi, S. K., Xavier, A., Mian, R., Stupar, R. M., Michno, J. M., An, Y. Q. C., Goettel, W., Ward, R., Fox, C., Lipka, A. E., Hyten, D., Cary, T., and Beavis, W. D. (2018). Genetic architecture of soybean yield and agronomic traits. *G3:*

Genes, Genomes, Genetics, *8*(10), 3367–3375. https://doi.org/10.1534/g3.118.200332

Diouf, I., Derivot, L., Koussevitzky, S., Carretero, Y., Bitton, F., Moreau, L., and Causse, M. (2020). Genetic basis of phenotypic plasticity and genotype × environment interactions in a multi-parental tomato population. *Journal of Experimental Botany*, *71*, 5365–5376. https://doi.org/10.1101/2020.02.07.938456

Edwards, M. D., Helentjaris, T., Wright, S., and Stuber, C. W. (1992). Molecular-marker-facilitated investigations of quantitative trait loci in maize. *Theoretical and Applied Genetics*, *83* (6–7), 765–774. https://doi.org/10.1007/bf00226696

Edwards, M. D., Stuber, C. W., and Wendel, J. F. (1987). Molecular-marker-facilitated investigations of quantitative-trait loci in maize. I. Numbers, genomic distribution and types of gene action. *Genetics*, *116*(1), 113–125.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., and Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, *6*(5), 1–10. https://doi.org/10.1371/journal.pone.0019379

Ersoz, E. S., Yu, J., and Buckler, E. S. (2007). Applications of linkage disequilibrium and association mapping in crop plants. In *Genomics-Assisted Crop Improvement* (Vol. 1). https://doi.org/10.1007/978-1-4020-6295-7_5

Fiorani, F., and Schurr, U. (2013). Future scenarios for plant phenotyping. *Annual Review of Plant Biology*, *64*, 267–291. https://doi.org/10.1146/annurev-arplant-050312-120137

Fisher, R. A. (1919). XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh*, *52*(2), 399–433. https://doi.org/10.1017/S0080456800012163

Flint-Garcia, S. A., Thuillet, A. C., Yu, J., Pressoir, G., Romero, S. M., Mitchell, S. E., Doebley, J., Kresovich, S., Goodman, M. M., and Buckler, E. S. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *The Plant Journal*, *44*(6), 1054–1064. https://doi.org/10.1111/j.1365-313X.2005.02591.x

Flint, J., and Mott, R. (2001). Finding the molecular basis of quantitative traits: Successes and pitfalls. *Nature Reviews Genetics*, *2*(6), 437–445. https://doi.org/10.1038/35076585

Foolad, M. R., and Panthee, D. R. (2012). Marker-Assisted Selection in Tomato Breeding. *Critical Reviews in Plant Sciences*, *31*(2), 93–123. https://doi.org/10.1080/07352689.2011.616057

Francia, E., Tacconi, G., Crosatti, C., Barabaschi, D., Bulgarelli, D., Dall'Aglio, E., and Valè, G. (2005). Marker assisted selection in crop plants. *Plant Cell, Tissue and Organ Culture*, *82*(3), 317–342. https://doi.org/10.1007/s11240-005-2387-z

Gage, J. L., Monier, B., Giri, A., and Buckler, E. S. (2020). Ten years of the maize nested association mapping population: Impact, limitations, and future directions. *Plant Cell*, *32*(7), 2083–2093. https://doi.org/10.1105/tpc.19.00951

Gallagher, M. D., and Chen-Plotkin, A. S. (2018). The Post-GWAS Era: From Association to Function. *American Journal of Human Genetics*, *102*(5), 717–730. https://doi.org/10.1016/j.ajhg.2018.04.002

Giri, P., Yadav, M. L., and Mohapatra, B. (2018). *QTL Linkage Analysis*. Springer International Publishing. https://doi.org/10.1007/978-3-319-47829-6_161-1

Giri, P., and Mohapatra, B. (2017). Candidate Gene. *Encyclopedia of Animal Cognition and Behavior*, 1–4. https://doi.org/10.1007/978-3-319-47829-6_1-1

Goddard, M. E., Kemper, K. E., MacLeod, I. M., Chamberlain, A. J., and Hayes, B. J. (2016). Genetics of complex traits: Prediction of phenotype, identification of causal polymorphisms and genetic architecture. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1835). https://doi.org/10.1098/rspb.2016.0569

Guo, Z., Yang, W., Chang, Y., Ma, X., Tu, H., Xiong, F., Jiang, N., Feng, H., Huang, C., Yang, P., Zhao, H., Chen, G., Liu, H., Luo, L., Hu, H., Liu, Q., and Xiong, L.. (2018). Genome-wide association studies of image traits reveal genetic architecture of drought resistance in rice. *Molecular Plant*, *11*(6), 789–805. https://doi.org/10.1016/j.molp.2018.03.018

Gupta, P. K., Rustgi, S., and Kulwal, P. L. (2005). Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Molecular Biology*, *57*(4), 461–485. https://doi.org/10.1007/s11103-005-0257-z

Hansen, T. E. (2006). The evolution of genetic architecture. *Annual Review of Ecology, Evolution, and Systematics*, *37*(8), 123–157. https://doi.org/10.1146/annurev.ecolsys.37.091305.110224

Hayward, A. C., Tollenaere, R., Dalton-Morgan, J., and Batley, J. (2015). Molecular marker applications in plants. *Methods in Molecular Biology*, (pp. 13–27). https://doi.org/10.1007/978-1-4939-1966-6_2

Hill, W. G. (2010). Understanding and using quantitative genetic variation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1537), 73–85. https://doi.org/10.1098/rstb.2009.0203

Ibrahim, A. K., Zhang, L., Niyitanga, S., Afzal, M. Z., Xu, Y., Zhang, L., Zhang, L., and Qi, J. (2020). Principles and approaches of association mapping in plant breeding. *Tropical Plant Biology*, *13*(3), 212–224. https://doi.org/10.1007/s12042-020-09261-4

Jaganathan, D., Bohra, A., Thudi, M., and Varshney, R. K. (2020). Fine mapping and gene cloning in the post-NGS era: advances and prospects. *Theoretical and Applied Genetics*, *133*(5), 1791–1810. https://doi.org/10.1007/s00122-020-03560-w

Jannink, J. L., Lorenz, A. J., and Iwata, H. (2010). Genomic selection in plant breeding: From theory to practice. *Briefings in Functional Genomics and Proteomics*, *9*(2), 166–177. https://doi.org/10.1093/bfgp/elq001

Jansen, R. C. (1993). Interval mapping of multiple quantitative trait loci. *Genetics*, *135*(1), 205–211. https://doi.org/10.1093/genetics/135.1.205

Joo, J. W. J., Hormozdiari, F., Han, B., and Eskin, E. (2016). Multiple testing correction in linear mixed models. *Genome Biology*, *17*, 62. https://doi.org/10.1186/s13059-016-0903-6

Ju, M., Zhou, Z., Mu, C., Zhang, X., Gao, J., Liang, Y., Chen, J., Wu, Y., Li, X., Wang, S., Wen, J., Yang, L., and Wu, J. (2017). Dissecting the genetic architecture of *Fusarium verticillioides seed* rot resistance in maize by combining QTL mapping and genome-wide association analysis. *Scientific Reports*, *7*, 1–11. https://doi.org/10.1038/srep46446

Juyo Rojas, D. K., Soto Sedano, J. C., Ballvora, A., Léon, J., and Mosquera Vásquez, T. (2019). Novel Organ-Specific Genetic Factors for Quantitative Resistance to Late Blight in Potato. *PLoS ONE*, *14*, 1–15. https://doi.org/10.1371/journal.pone.0213818

Kaler, A. S., and Purcell, L. C. (2019). Estimation of a significance threshold for genome-wide association studies. *BMC Genomics*, *20*(1), 1–8. https://doi.org/10.1186/s12864-019-5992-7

Khan, R., Ma, X., Shah, S., Wu, X., Shaheen, A., Xiao, L., Wu, Y., and Wang, S. (2020). Drought-hardening improves drought tolerance in *Nicotiana tabacum* at physiological, biochemical, and molecular levels. *BMC Plant Biology*, *20*(1), 1–19. https://doi.org/10.1186/s12870-020-02688-7

Kirchgessner, N., Liebisch, F., Yu, K., Pfeifer, J., Friedli, M., Hund, A., and Walter, A. (2016). The ETH field phenotyping platform FIP: A cable-suspended multi-sensor system. *Functional Plant Biology*, *44*(1), 154–168. https://doi.org/10.1071/FP16165

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods*, *9*(1), 1. https://doi.org/10.1186/1746-4811-9-29

Krajewski, P., Chen, D., - wiek, H., Van Dijk, A. D. J., Fiorani, F., Kersey, P., Klukas, C., Lange, M., Markiewicz, A., Nap, J. P., Van Oeveren, J., Pommier, C., Scholz, U., Van Schriek, M., Usadel, B., and Weise, S. (2015). Towards recommendations for metadata and data handling in plant phenotyping. *Journal of Experimental Botany*, *66*(18), 5417–5427. https://doi.org/10.1093/jxb/erv271

Lander, E. S., and Botstein, D. (1989). Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps. *Genetics*, *121*(1), 185–199. https://doi.org/10.1093/genetics/121.1.185

Li, F., Wen, W., Liu, J., Zhang, Y., Cao, S., He, Z., Rasheed, A., Jin, H., Zhang, C., Yan, J., Zhang, P., Wan, Y., and Xia, X. (2019). Genetic architecture of grain yield in bread wheat based on genome-wide association studies. *BMC Plant Biology*, *19*(168), 1–19. https://doi.org/10.1186/s12870-019-1781-3

Li, H., Hearne, S., Bänziger, M., Li, Z., and Wang, J. (2010). Statistical properties of QTL linkage mapping in biparental genetic populations. *Heredity*, *105*(3), 257–267. https://doi.org/10.1038/hdy.2010.56

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., and Zhang, Z. (2012). GAPIT: Genome association and prediction integrated tool. *Bioinformatics*, *28*(18), 2397–2399. https://doi.org/10.1093/bioinformatics/bts444

Liu, B. H. (1998). *Statistical Genomics: Linkage, Mapping and QTL Analysis*. CRC Press.

Liu, N., Xue, Y., Guo, Z., Li, W., and Tang, J. (2016). Genome-wide association study identifies candidate genes for starch content regulation in maize kernels. *Frontiers in Plant Science*, *7*, 1046. https://doi.org/10.3389/fpls.2016.01046

Lohmueller, K. E., Pearce, C. L., Pike, M., Lander, E. S., and Hirschhorn, J. N. (2003). Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nature Genetics*, *33*, 177–182. https://doi.org/10.1038 / ng1071

Lynch, J. P. (2011). Root phenes for enhanced soil exploration and phosphorus acquisition: Tools for future crops. *Plant Physiology*, *156*(3), 1041–1049. https://doi.org/10.1104/pp.111.175414

Lynch, M., and Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Inc.

Mackay, T. F. C. (2001). The genetic architecture of quantitative traits. *Annual Review of Genetics*, *35*(1), 303–339. https://doi.org/10.1146/annurev.genet.35.102401.090633

Marla, S. R., Burow, G., Chopra, R., Hayes, C., Olatoye, M. O., Felderhoff, T., Hu, Z., Raymundo, R., Perumal, R., and Morris, G. P. (2019). Genetic Architecture of Chilling Tolerance in Sorghum Dissected with a Nested Association Mapping Population. *Genetics*, *9*(12), 4045–4057. https://doi.org/https://doi.org/10.1534/g3.119.400353

Meunier, R. (2016). The many lives of experiments: Wilhelm Johannsen, selection, hybridization, and the complex relations of genes and characters. *History and Philosophy of the Life Sciences*, *38*(1), 42–64. https://doi.org/10.1007/s40656-015-0093-7

Morris, J., Navarro, N., Rastas, P., Rawlins, L. D., Sammy, J., Mallet, J., and Dasmahapatra, K. K. (2019). The genetic architecture of adaptation: convergence and pleiotropy in Heliconius wing pattern evolution. *Heredity*, *123*(2), 138–152. https://doi.org/10.1038/s41437-018-0180-0

Nienhuis, J., Helentjaris, T., Slocum, M., Ruggero, B., and Schaefer, A. (1987). Restriction Fragment Length Polymorphism Analysis of Loci Associated with Insect Resistance in Tomato[1] . *Crop Science*, *27*(4), 797–803. https://doi.org/10.2135/cropsci1987.0011183x002700040039x

Nilsson-Ehle, H. (1908). Einige Ergebnisse von Kreuzungen bei Hafer und Weizen. *Botaniska Notiser*, 301–350.

Ordas, B., Malvar, R. A., Santiago, R., and Butron, A. (2010). QTL mapping for Mediterranean corn borer resistance in European flint germplasm using recombinant inbred lines. *BMC Genomics*, *11*(174). https://doi.org/10.1186/1471-2164-11-174

Osborn, T. C., Alexander, D. C., and Fobes, J. F. (1987). Identification of restriction fragment length polymorphisms linked to genes controlling soluble solids content in tomato fruit. *Theoretical and Applied Genetics*, *73*(3), 350–356. https://doi.org/10.1007/BF00262500

Pan, Y., Zhang, H., Zhang, D., Li, J., Xiong, H., Yu, J., Li, J., Rashid, M. A. R., Li, G., Ma, X., Cao, G., Han, L., and Li, Z. (2015). Genetic analysis of cold tolerance at the germination and booting stages in rice by association mapping. *PLoS ONE*, *10*(3). https://doi.org/10.1371/journal.pone.0120590

Patnala, R., Clements, J., and Batra, J. (2013). Candidate gene association studies: a comprehensive guide to useful in *silicotools*. *Molecular Breeding*, *14*(39). https://doi.org/10.1186/1471-2156-14-39

Pavan, S., Delvento, C., Ricciardi, L., Lotti, C., Ciani, E., and D'Agostino, N. (2020). Recommendations for Choosing the Genotyping Method and Best Practices for Quality Control in Crop Genome-Wide Association Studies. *Frontiers in Genetics*, *11*(). https://doi.org/10.3389/fgene.2020.00447

Pflieger, S., Lefebvre, V., and Causse, M. (2001). The candidate gene approach in plant genetics: a review. *Molecular Breeding*, *7*(4), 275–291. https://doi.org/10.1023/A:1011605013259

Pieruschka, R., and Schurr, U. (2019). Plant Phenotyping: Past, Present, and Future. *Plant Phenomics*, *2019*, 1–6. https://doi.org/10.34133/2019/7507131

Platt, A., Vilhjálmsson, B. J., and Nordborg, M. (2010). Conditions under which genome-wide association studies will be positively misleading. *Genetics*, *186*(3), 1045–1052. https://doi.org/10.1534/genetics.110.121665

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. https://doi.org/10.1038/ng1847

Pritchard, J. K., Stephens, M., Rosenberg, N. A., and Donnelly, P. (2000). Association mapping in structured populations. *The American Journal of Human Genetics*, *67*(1), 170–181. https://doi.org/10.1086 / 302959

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., and Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Rajarammohan, S., Kumar, A., Gupta, V., Pental, D., Pradhan, A. K., and Kaur, J. (2017). Genetic architecture of resistance to alternaria brassicae in Arabidopsis thaliana: QTL mapping reveals two major resistance-conferring loci. *Frontiers in Plant Science*, *8*, 1–9. https://doi.org/10.3389/fpls.2017.00260

Rajpal, V. R., Dreisigacke, S., Sukumaran, S., Guzmán, C., He, X., Lan, C., Bonnett, D., and Crossa, J. (2016). Molecular Breeding for Sustainable Crop Improvement. In *Springer* (Vol. 11, Issue ii). https://doi.org/10.1007/978-3-319-27090-6

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, *8*(11), 2281–2308. https://doi.org/10.1038/nprot.2013.143

Rasheed, A., Hao, Y., Xia, X., Khan, A., Xu, Y., Varshney, R. K., and He, Z. (2017). Crop Breeding Chips and Genotyping Platforms: Progress, Challenges, and Perspectives. *Molecular Plant*, *10*(8), 1047–1064. https://doi.org/10.1016/j.molp.2017.06.008

Remington, D. L. (2015). Alleles versus mutations: Understanding the evolution of genetic architecture requires a molecular perspective on allelic origins. *Evolution*, *69*(12), 3025–3038. https://doi.org/10.1111/evo.12775

Risch, N., and Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science*, *273*(5281), 1516–1517. https://doi.org/10.1126/science.273.5281.1516

Robertson, D. (2004). VIGS Vectors for Gene Silencing: Many targets, many tools. *Annual Review of Plant Biology*, *55*, 495–519. https://doi.org/10.1146/annurev.arplant.55.031903.141803

Robinson, R. A. (1996). *Return to resistance. Breeding crops to Reduce Pesticide Dependence*. *Australasian Plant Pathology*, 216-217. https://doi.org/10.1007/BF03213684

Sax, K. (1923). The Association of Size Differences with Seed-Coat Pattern and Pigmentation in *Phaseolus Vulgaris*. *Genetics*, *8*(6), 552–560. https://doi.org/10.1093/genetics/8.6.552

Sella, G., and Barton, N. H. (2019). Thinking about the Evolution of Complex Traits in the Era of Genome-Wide Association Studies. *Annual Review of Genomics and Human Genetics*, *20*, 461–493. https://doi.org/10.1146/annurev-genom-083115-022316

Semagn, K., Bjørnstad, Å., and Xu, Y. (2010). The genetic dissection of quantitative traits in crops. *Electronic Journal of Biotechnology*, *13*(5). https://doi.org/10.2225/vol13-issue5-fulltext-14

Soto, J. C., Ortiz, J. F., Perlaza-Jiménez, L., Vásquez, A. X., Lopez-Lavalle, L. A. B., Mathew, B., Léon, J., Bernal, A. J., Ballvora, A., and López, C. E. (2015). A genetic map of cassava (*Manihot esculenta* Crantz) with integrated physical mapping of immunity-related genes. *BMC Genomics*, *16*(1), 1–16. https://doi.org/10.1186/s12864-015-1397-4

Soto , J. C., Mora , R. E., Mathew, B., Léon, J., Gómez , F. A., Ballvora, A., and López , C. E. (2017). Major novel QTL for resistance to Cassava bacterial blight identified through a multi-environmental analysis. *Frontiers in Plant Science*, *8*(), 1–13. https://doi.org/10.3389/fpls.2017.01169

St.clair, D. A. (2010). Quantitative disease resistance and quantitative resistance loci in breeding. *Annual Review of Phytopathology*, *48*, 247–268. https://doi.org/10.1146/annurev-phyto-080508-081904

Szalma, S. J., Hostert, B. M., LeDeaux, J. R., Stuber, C. W., and Holland, J. B. (2007). QTL mapping with near-isogenic lines in maize. *Theoretical and Applied Genetics*, *114*(7), 1211–1228. https://doi.org/10.1007/s00122-007-0512-6

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., and Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*, *20*(8), 467–484. https://doi.org/10.1038/s41576-019-0127-1

Tanksley, S. D. (1993). Mapping polygenes. *Annual Review of Genetics*, *27*(1), 205–233. https://doi.org/10.1146/annurev.ge.27.120193.001225

Thoen, M. P. M., Davila, N. H., Kloth, K. J., Coolen, S., Huang, P. P., Aarts, M. G. M., Bac-Molenaar, J. A., Bakker, J., Bouwmeester, H. J., Broekgaarden, C., Bucher, J., Busscher-Lange, J., Cheng, X., Fradin, E. F., Jongsma, M. A., Julkowska, M. M., Keurentjes, J. J. B., Ligterink, W., Pieterse, C. M. J., Ruyter-Spira, C., Smant, G., Testerink, C., Usadel, B., van Loon, J. J. A., van Pelt, J. A., van Schaik, C. C., van Wees, S. C. M., Visser, R. G. F., Voorrips, R., Vosman, B., Vreugdenhil, D., Warmerdam, S., Wiegers, G. J., van Heerwaarden, J., Kruijer, W., van Eeuwijk, F. A., andDicke, M. (2017). Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping. *New Phytologist*, *213*(3), 1346–1362. https://doi.org/10.1111/nph.14220

Unver, T., and Budak, H. (2009). Virus-induced gene silencing, a post transcriptional gene silencing method. *International Journal of Plant Genomics*, *2009*. https://doi.org/10.1155/2009/198680

Villanueva, B., Fernández, A., Saura, M., Caballero, A., Fernández, J., Morales-González, E., Toro, M. A., and Pong-Wong, R. (2021). The value of genomic relationship matrices to estimate levels of inbreeding. *Genetics Selection Evolution*, *53*(42), 1–17. https://doi.org/10.1186/s12711-021-00635-0

Waghmode, B. D., Sabnis, G. R., Navhale, V. C., and Thaware, B. L. (2017). Inheritance studies in red kernel rice (*Oryza sativa* L.). *Electronic Journal of Plant Breeding*, *8*(2), 521–527. https://doi.org/10.5958/0975-928X.2017.00078.3

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., and Mons, B. (2016). Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*, 1–9. https://doi.org/10.1038/sdata.2016.18

Williams, J. G. K., Kubelik, A. R., Livak, K. J., and Rafalski, J. A., T. S. (1990). DNA polymorphisms amplified methods to plant systematic and evolutionary biology. In *Molecular systematic and evolutionary biology* (pp. 43–86).

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, *16*(3). https://doi.org/10.1093/genetics/16.3.290

Yadav, A., and Sinha, H. (2018). Gene–gene and gene–environment interactions in complex traits in yeast. *Yeast*, *35*(6), 403–416. https://doi.org/10.1002/yea.3304

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., Xiong, L., and Yan, J. (2020). Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives. *Molecular Plant*, *13*(2), 187–214. https://doi.org/10.1016/j.molp.2020.01.008

Yano, K., Morinaka, Y., Wang, F., Huang, P., Takehara, S., Hirai, T., Ito, A., Koketsu, E., Kawamura, M., Kotake, K., Yoshida, S., Endo, M., Tamiya, G., Kitano, H., Ueguchi-Tanaka, M., Hirano, K., and Matsuoka, M. (2019). GWAS with principal component analysis identifies a gene comprehensively controlling rice architecture. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(42), 2162–21267. https://doi.org/10.1073/pnas.1904964116

York, L. M. (2019). Functional phenomics: An emerging field integrating high-throughput phenotyping, physiology, and bioinformatics. *Journal of Experimental Botany*, *70*(2), 379–386. https://doi.org/10.1093/jxb/ery379

Yu, X., Li, X., Guo, T., Zhu, C., Wu, Y., Mitchell, S. E., Roozeboom, K. L., Wang, D., Wang, M. L., Pederson, G. A., Tesso, T. T., Schnable, P. S., Bernardo, R., and Yu, J. (2016). Genomic prediction contributing to a promising global strategy to turbocharge gene banks. *Nature Plants*, *2*(10), 1–7. https://doi.org/10.1038/nplants.2016.150

Zakir, M. (2018). Review on Genotype X Environment Interaction in Plant Breeding and Agronomic Stability of Crops. *Journal of Biology, Agriculture and Healthcare*, *8*(12), 14–21. https://www.iiste.org/Journals/index.php/JBAH/article/view/43065

Zan, Y., and Carlborg, Ö. (2018). A multilocus association analysis method integrating phenotype and expression data reveals multiple novel associations to flowering time variation in wild- collected Arabidopsis thaliana. *Molecular Ecology Resources*, *18*(4), 798–808. https://doi.org/10.1111/1755-0998.12757

Zaw, H., Raghavan, C., Pocsedio, A., Swamy, B. P. M., Jubay, M. L., Singh, R. K., Bonifacio, J., Mauleon, R., Hernandez, J. E., Mendioro, M. S., Gregorio, G. B., and Hei, L. (2019). Exploring genetic architecture of grain yield and quality traits in a 16- way indica by *japonica* rice MAGIC global population. *Scientific Reports*, *9*, 1–11. https://doi.org/10.1038/s41598-019-55357-7

Zhao, Y., Mette, M. F., Gowda, M., Longin, C. F. H., and Reif, J. C. (2014). Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. *Heredity*, *112*(6), 638–645. https://doi.org/10.1038/hdy.2014.1

Zhu, C., Gore, M., Buckler, E. S., and Yu, J. (2008). Status and Prospects of Association Mapping in Plants. *The Plant Genome*, *1*(1), 5–20. https://doi.org/10.3835/plantgenome2008.02.0089