Taylor & Francis
Taylor & Francis Group

Critical Reviews in Plant Sciences

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/bpts20

Nested Association Mapping (NAM) Populations: Present Status and Future Prospects in the Genomics Era

Channappa Gireesh, Raman. M. Sundaram, Siddaiah M. Anantha, Manish K. Pandey, Maganti S. Madhav, Santosha Rathod, Kondajji R. Yathish, Ponnuvel Senguttuvel, Barbadikar M. Kalyani, Ellur Ranjith, Lella Venkata Subbarao, Tapan Kumar Mondal, Mallikarjuna Swamy & Sujay Rakshit

To cite this article: Channappa Gireesh, Raman. M. Sundaram, Siddaiah M. Anantha, Manish K. Pandey, Maganti S. Madhav, Santosha Rathod, Kondajji R. Yathish, Ponnuvel Senguttuvel, Barbadikar M. Kalyani, Ellur Ranjith, Lella Venkata Subbarao, Tapan Kumar Mondal, Mallikarjuna Swamy & Sujay Rakshit (2021): Nested Association Mapping (NAM) Populations: Present Status and Future Prospects in the Genomics Era, Critical Reviews in Plant Sciences, DOI: 10.1080/07352689.2021.1880019

To link to this article: https://doi.org/10.1080/07352689.2021.1880019

Published online: 06 Mar 2021.

CRITICAL REVIEWS IN PLANT SCIENCES
https://doi.org/10.1080/07352689.2021.1880019

Taylor & Francis
Taylor & Francis Group

Nested Association Mapping (NAM) Populations: Present Status and Future Prospects in the Genomics Era

Channappa Gireesh[a], Raman. M. Sundaram[a], Siddaiah M. Anantha[a], Manish K. Pandey[b], Maganti S. Madhav[a], Santosha Rathod[a], Kondajji R. Yathish[c], Ponnuvel Senguttuvel[a], Barbadikar M. Kalyani[a], Ellur Ranjith[d], Lella Venkata Subbarao[a], Tapan Kumar Mondal[e], Mallikarjuna Swamy[f], and Sujay Rakshit[c]

[a]ICAR-Indian Institute of Rice Research, Hyderabad, India; [b]International Crops Research Institute for the Semi-Arid Tropics, Hyderabad, India; [c]ICAR-Indian Institute of Maize Research, Ludhiana, India; [d]ICAR-Indian Agricultural Research Institute, New Delhi, India; [e]ICAR-National Institute for Plant Biotechnology, New Delhi, India; [f]International Rice Research Institute, Manila, Philippines

ABSTRACT

Development of mapping population is a prerequisite for genetic dissection of genomic regions underlying complex traits. Nested Association Mapping (NAM) is an integrated multi-parent population approach that combines the advantages of linkage mapping and association mapping for high resolution and high-power mapping of complex traits. The NAM population is constituent of independent RIL populations derived from crossing several diverse donor parents with a common founder parent. The first NAM population was developed in maize and later on in several crops like barley, sorghum, wheat, rice, soybean, etc. This review provides an overview of NAM population development, its features, advantages over the other mapping populations, availability of high density genotyping platforms, key considerations for their development, applications and future prospects. We propose that the recent high-throughput analytical tools including high-end genotyping will accelerate utilization of NAM population for prediction of genomic estimated breeding value and genomic assisted selection in crop improvement program.

KEYWORDS

Combined linkage analysis and linkage disequilibrium; genomic prediction; GWAS; quantitative traits; multi-parent mapping population

I. Introduction

Most of the economically important traits in crop plants are quantitative in nature. These traits are influenced by several genes, environmental factors, epistatic interactions of loci, and thus exhibit continuous variations. With the recent advances in DNA marker technology, it is now feasible to identify the genomic regions controlling the quantitative traits. The term "Quantitative Trait Loci" (QTLs), coined by Gelderman (1975), refers to the physical location of causative factors in the genome, associated with the expression of quantitative traits. The concept of QTL mapping depends on finding an association between DNA sequence variation (genetic marker) and measurable phenotypic variations of a complex trait. The prerequisite for identification of precise QTLs associated with any quantitative traits is mapping population along with their phenotypic variations and molecular marker data. Understanding the genetic architecture of quantitative traits is further strengthened and accelerated by rapid advances in sequencing and genotyping technologies, coupled with improvement in phenotyping techniques (Choudhary *et al.*, 2019).

II. Evolution of mapping population – gradual progress in precision and resolution of mapping

Mapping population is a group of individuals suitable for linkage mapping of genetic markers using principles of Mendelian inheritance. Over the years with a better understanding of the breeding objectives, several types of mapping populations have been developed. These populations have been categorized into first generation, second generation, and third generation mapping populations which are described in Table 1. Comparative features of different mapping populations employed in crop plants are provided in Table 2.

CONTACT Sujay Rakshit s.rakshit@icar.gov.in ICAR-Indian Institute of Maize Research, Ludhiana, India.

Table 1. Evolution of genetic populations for molecular mapping, gene discovery and marker development for target traits in crop plants.

Generation of populations	Development of populations	Examples of mapping populations	Features
First generation mapping population	Population derived by crossing two parents (Biparental mapping population) with contrasting features for target traits	F_2, $F_{2:3}$, backcross population (BC_1, BC_2), Advanced backcross population, NILs, RILs, DH population and CSSLs	These populations were used for mapping single to few target traits based on the phenotype diversity available in parental genotypes used for making cross.
Second generation mapping population	Population is comprised set of genetically variable lines containing phenotypic variability for target traits. Such populations were more useful for plant species where making crosses and developing biparental population was difficult or not possible.	Association panel and Mutant population	It is based on the genetic variability present among the natural populations that are fixed through historical recombination (Association panel) and the artificial inducted genetic variations through mutations (Mutant population).
Third generation mapping population	Population is derived by crossing multiple parents in different mating designs to facilitate mapping for multiple target traits in single effort. Besides trait mapping, these populations also leads to development of highly diverse breeding lines for further selection and evaluation toward varietal development.	MAGIC and NAM population, Four way hybrid population, diallele crossed population	These populations involve utilization of multiple parents to create highly diverse genetic populations for performing high resolution trait mapping and gene discovery.

NILs: Near Isogenic Lines; RILs: Recombinant Inbred Lines; DH: Double Haploids; CSSLs: Chromosome Segment Substitution Lines; MAGIC: Multiparent Advanced Generation Inter Cross; NAM: Nested Association Mapping.

The first-generation mapping population refers to biparental populations developed by crossing two contrasting parents. It involves F_2, $F_{2:3}$, backcross populations (BC_1, BC_2), Advanced backcross populations, NILs (Near Isogenic Lines), RILs (Recombinant Inbred Lines), DH population (Double Haploids) and CSSLs (Chromosome Segment Substitution Lines). Using linkage analysis, these populations have been extensively utilized for identification and mapping of qualitative traits like disease resistance (controlled by single or few genes) and also quantitative traits like salinity tolerance, drought tolerance, etc. Although linkage mapping is a powerful mapping strategy for the detection of QTLs associated with complex traits, its mapping resolution is poor due to the limited amount of recombinations (Xu *et al.*, 2017) and possible loss of recombinants during population advancement. Increasing the density of markers (beyond one marker per 15 cM) may not be fruitful without increasing its population size to improve the resolution of QTL mapping in biparental populations (Darvasi and Soller, 1995; Kearsey and Farquhar, 1998). Besides, the biparental population will have a low level of allelic variations (as there are two possible alleles for each locus) and such populations show segregation for only a few traits, which are contrasting between the two parents (Rakshit, *et al.*, 2012).

Thus, to overcome the shortfalls of first-generation mapping populations, researchers utilized natural populations and mutant populations as second generation mapping populations for mapping complex traits. In the recent past, association mapping is gaining much more prominence as it promises to overcome the limitations of linkage mapping. In the case of association mapping, panels containing a diverse set of genotypes with phenotypic diversity for target traits are used for establishing marker-trait association using high-density genotyping and multi-season/multi-environment phenotyping data. It takes advantage of historical recombination that occurred during evolution, enabling higher precision and resolution for mapping complex traits (Mackay and Powell, 2007). Association mapping has low power than linkage mapping for the detection of QTLs. The power of QTL detection in association mapping depends on population size, population structure, phenotyping of population, genome-wide markers, and suitable statistical methods employed (Xu *et al.*, 2017). Further, the accuracy of detection of QTLs with the minor effect is low in the association mapping approach as compared to linkage mapping (Cockram and Mackay, 2018). Besides, rare alleles are detected in much lesser frequency (maf

Table 2. Comparative features of different mapping populations in plants.

S. No.	Characters	NAM	MAGIC	Association mapping	Biparental mapping
1.	Parents involved	>2	8	>100	2
2.	Power of mapping	High	Moderate	Low	High
3.	Resolution of mapping	High	High	High	Low
4.	Historical LD	Yes	Yes	Yes	No
5.	Recombination derived LD	Yes	Yes	No	Yes
6.	Detection of multiple alleles	Yes	Yes	Yes	2 Alleles
7.	Power of detection of rare QTLs	High	Moderate	Less	High
8.	Population structure	Low	No	Yes	No
9.	Genetic base and diversity of parents	Broad	Moderate	Broad	Narrow
10.	Hybridization for development of population	Required	Required	No	Required

QTL: quantitative trait locus; LD: linkage disequilibrium.

threshold level set as to 0.01, 0.02, or 0.5 to reduce false positives) during association mapping, even if their effects are large (Cockram and Mackay, 2018). However, while analyzing such alleles are filtered out in order to reduce the chances of false positives (Cockram and Mackay, 2018). Population structure in the association mapping panel is another important consideration, which can lead to the spurious marker-trait association (Knowler *et al.*, 1988). Filtering of rare variations and the presence of inherent population structure in association panels tend to reduce the power of QTL detection. It is pertinent to note that rare variants could be of economic interest for breeders and if population structure is associated with such rare variants, it is hardly be detected in GWAS (Genome Wide Association Studies). In this scenario, multiparent populations are suitable to reduce such spurious association and to increase the power of rare allele detection.

The third-generation mapping populations are developed to combine the principles of linkage mapping and association mapping for QTL detection by developing family-based populations using several diverse parents by breeding designs. Rebai and Goffinet (Rebai and Goffinet, 1993) proposed the concept of utilizing multi-parents for molecular mapping of QTLs in six F_2 populations derived from diallel crosses using high-density RFLP. Later, Jannink and Jansen (2000) employed double haploids populations derived from three parents in diallel crosses for mapping higher-order epistasis (involves multiple QTLs) and genetic background interaction in one-dimensional genome search wherein QTL allelic values are either nested within or fixed over populations. The four kinds of multiparent mapping populations developed in animals and plants are four way crosses, diallel cross population, multiparent advanced generation inter-cross (MAGIC) population and Nested Association Mapping population (Mackay and Powell, 2007; Cavanagh *et al.*, 2008). Four-way crosses populations are derived from a cross between two F_1 hybrids [F_1 (A1/A2)/F_1 (A3/A4] and developed in maize (Anderson *et al.*, 2018). Several multiparent intercross mapping populations have been developed in mice (Churchil et al., 2004), Drosophila (MacDonald and Long, 2007), Arabidopsis (Kover *et al.*, 2009), maize (Anderson *et al.*, 2018).

In mice, MAGIC population was reported for development of the first multiparental inter-mated population to constitute collaborative cross, a community resource for the genetic analysis of complex traits (Churchil *et al.*, 2004). The multiparent mapping populations have been utilized for mapping candidate genes for serum cholesterol and coat color traits (Cavanagh *et al.*, 2008; Svenson *et al.*, 2012). Development of the MAGIC population involves crossing eight diverse founder parents in a designed manner and further intermating of siblings (2-way crosses, 4-way crosses) to develop eight independent RIL populations through SSD (Single Seed Descent) method which represent mosaic genome of eight founder lines. MAGIC populations have been developed in *Arabidopsis*, tomato, barley, maize, sorghum, wheat, and rice (Ladejobi *et al.*, 2016).

Although linkage analysis offers high power of QTL detection, its mapping resolution is low, whereas, the association mapping gives a higher resolution of QTL mapping with lower power of QTL detection. Consequently, an integrative mapping strategy combining the principles of both linkage analysis and association mapping for high resolution and high power mapping is indeed imperative to unearth the genetic architecture of complex traits. The development of such an integrative mapping strategy is known as Nested Association Mapping (NAM). The deployment of multiple parents in the development of NAM population augment accurate trait mapping as well as gainful utilization of allelic diversity. It enhances the resolution power of mapping complex traits and broadens the genetic base of the breeding population by shuffling the diverse genomic regions controlling quantitative traits.

Table 3. Details of NAM populations developed in crops.

Crop	Common founder parent	Number of donor founder parents	Population size	Traits mapped/studied	References
Maize	B73	25 inbred lines	5000	Developed US-NAM population in maize and used for studying genetic design and statistical power of population.	Yu *et al.*, 2008
			5000	Genetic properties of population-recombination events, segregation distortion and residual heterozygosity	McMullen *et al.*, 2009
			5000	Flowering time	Buckler *et al.*, 2009
			5000	Southern leaf blight	Kump *et al.*, 2011
			4630	Northern leaf blight	Poland *et al.*, 2011
			5000	Leaf architecture	Tian *et al.*, 2011
			4699	Kernel composition (kernel starch, protein, and oil)	Cook *et al.*, 2012
			4692	Genetic architecture of maize stalk strength	Peiffer *et al.*, 2013
			5000	Genetic architecture of maize height	Peiffer *et al.*, 2014
			5000	Carbon and nitrogen metobolism	Zhang *et al.*, 2015
			5000	Drought tolerance	Li, Sun, *et al.*, 2016
			5000	Tocochromanols (Tocopherols, tocotrienols and plastochromanols)	Diepenbrock *et al.*, 2017
	Huangzaosi	11 diverse lines	2000	Developed Chinese NAM population (CN-NAM) in maize.	Li *et al.*, 2015
	Both US-NAM and CN-NAM		7000	Drought tolerance	Li, Sun, *et al.*, 2016
	W22	5 teosinte inbreds	1257	Domestication and agronomic trait analysis	Chen *et al.*, 2019
Barley	Barke	HEB-5:Five accessions of wild barely *(Hordeum vulgare* ssp. *Spontaneum)*	295	Leaf rust seedling resistance (*Puccinia hordei*)	Schnaithmann *et al.*, 2014
	Barke	HEB-25: A 25 accessions of wild barley (24 wild barley accessions of *H. Vulgare* ssp. *spontaneum* (Hsp), the progenitor of domesticated barley, and one Tibetian *H. Vulgare* ssp. *agriocrithon* (Hag) accession)	1420	Flowering time	Maurer *et al.*, 2015
			1336	Salinity tolerance and flowering time, yield and the harvest index under salinity stress condition.	Saade *et al.*, 2016
			1403	Net blotch (*Pyrenophora teres f.* teres)	Vatter *et al.*, 2017
			1420	Leaf sheath hairiness	Saade *et al.*, 2017
			1420	Yield related traits (Thousand grain weight, grain area, grain length, grain width, grain roundness, grains yield)	Sharma *et al.*, 2018
	Rasmusson	25 Diverse wild barley lines [*Hordeum vulgare* L. subsp. *Spontaneu*]	796	Heading date, plant height, test height, productive tillers per plant and yield	Nice *et al.*, 2017
Rice	IR64	10 diverse tropical japonica lines (Including Azucena)	1879	Days to heading, recombination events, segregation distortion,	Fragoso *et al.*, 2017
	IR64	21 Accessions of *O. glaberrima*	2400	Abiotic stress	Gireesh (persl. Comm.)
Sorghum	RTx430	10 Diverse global lines	2214	Adaptive traits (flowering time and plant height) and recombination events	Bouchet *et al.*, 2017
Soybean	IA3023	40 Diverse soybean genotypes	5600	Genetic properties, segregation distortion, recombination events and residual heterozygosity	Song *et al.*, 2017
				Yield and agronomic traits	Diers *et al.*, 2018
	JS335	20 Diverse soybean	900	Yield attributing traits	Shivakumar *et al.*, 2019
	M8206	Four (Linhe, Zhengyang Tongshan WSB)	623	Flowering date	Li *et al.*, 2017
Wheat	LMPG 6	10 stem rust resistance spring wheat cultivars	852	Stem rust resistance	Bajgain *et al.*, 2016
	Berkut	29 diverse (four cultivars and 25 landraces)	2100	Genetic properties, recombination QTL, cross over frequency and distribution	Jordan *et al.*, 2018

(*Continued*)

Table 3. Continued.

Crop	Common founder parent	Number of donor founder parents	Population size	Traits mapped/studied	References
	Avocet-YrA	2 Lines (Francolin and Quaiu)	391	powdery meldew resistance	Ren *et al.*, 2017
Durum wheat	Asassa	50 Ethiopian farmers varieties	6280	Phenology traits, plant height, flowering time	Kidane *et al.*, 2019
Common bean	Merlo	Three accessions	160	Agronomic traits (days to flowering, days to maturity, lodging resistance, canopy height, seed size and yield)	Hoyos-Villegas *et al.*, 2016
Rapeseed (*Brassica napus*)	PBY061	51 accessions of world-wide diverse *Brassica napus* germplasm	2900	Developed *Brassica* NAM population. Genetic properties and power of QTL detection	Li, Anja, *et al.*, 2016
	Zhongshuang 11 (ZS11)	15	2425	Developed *Brassica* NAM population. Constructed high-resolution linkage maps and genetic properties of the NAM population	Hu *et al.*, 2018
Groundnut (*Arachis hypogaea*)	Tifrunner and Florida-07	N08082oIJCT, C76-16, NC 3033 and GP-NC WS16 (SPT 06-06)	581 (NAM-Tifrunner) and 496 (NAM-Florida-07)	Genetic analysis and QTL discovery for seed traits in groundnut	Gangurde *et al.*, 2020
	ICGV 91114 (Spanish) and ICGS 76 (Virginia)	22 diverse testers crossed with ICGV 91114 and 21 diverse testers with ICGS 76	>2000 NAM lines for Spanish-NAM and Virginia-NAM populations	Two NAM populations each belonging to Spanish and Virginia types available for extensive phenotyping and trait mapping.	Pandey *et al.*, 2016
Pigeonpea (*Cajanus cajan*)	Asha (ICPL 87119)	10 different elite lines.	>2500 lines	Available for mapping for Fusarium wilt, sterility mosaic disease, yield and yield related traits and seed protein content.	Pandey *et al.*, 2016

NAM population was first developed in maize by crossing 25 diverse founder inbred parents with B73, the common founder parent, and 5000 RILs were developed to constitute maize NAM population (Yu *et al.*, 2008). Subsequently, NAM populations were developed in barley, wheat, rice, sorghum, rapeseed, soybean, groundnut etc. (Table 3). Once populations are developed, many studies can be undertaken using a single NAM population since diverse parents with divergent traits are involved in the process. For example, the NAM population in maize by Yu *et al.* (2008) or in barley by Maurer *et al.* (2015) have lead to 12 and 6 studies independent from the base populations, respectively. Since, more number of crosses is to be made for construction of NAM population; it is much easier to develop them in cross pollinated crops like maize as against highly self-pollinated crops like chickpea.

Considering the importance of NAM and its utility, various aspects of NAM population development, their application in crop improvement, current status, and prospects of NAM strategies in conjunction with the genomics are reviewed and discussed in present review.

III. Strategy for developing Nested Association Mapping population

The NAM population is a multiparent mapping strategy wherein multiple diverse donor parents (named founder donor parents) are crossed with a single parent (named common founder parent). From each cross of founder donor parent with common founder parent is subjected to create independent RIL populations. Together these individual RIL populations constitute the NAM population. Various steps in the development of NAM population and chromosomal inheritance are depicted in Figures 1 and 2. Each line of the NAM population represents a mosaic segment of genomes from respective donor founder parents and a common founder parent (Yu *et al.*, 2008; McMullen *et al.*, 2009; Bazakos *et al.*, 2017). The NAM design nests the historical Linkage Disequilibirum (LD) within the new recombinations and exploits both historical LD (due to the use of a large number of diverse founder parents) and recombination derived LD (during the development of RILs) (Yu *et al.*, 2008; Guo *et al.*, 2013; Nice *et al.*, 2016). During the development of NAM populations, the genome of founder parents gets reshuffled causing segregation of loci in the populations (Buckler *et al.*, 2009). Owing to the use of diverse founder parents

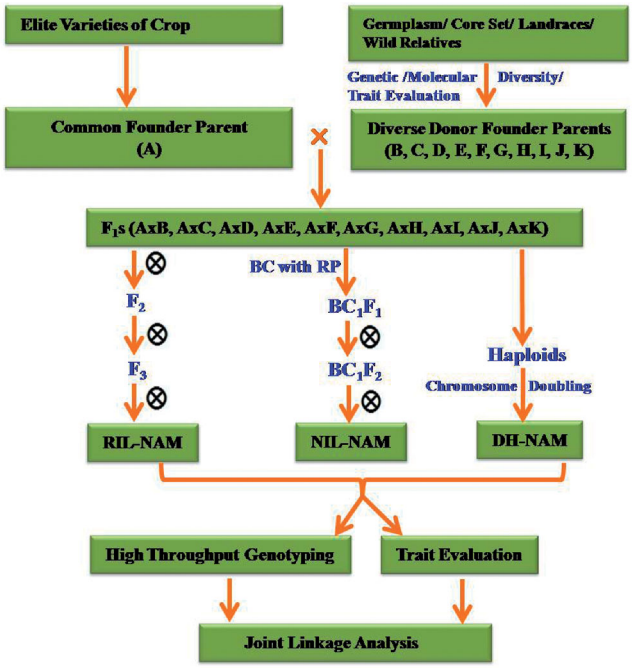

Figure 1. Flow diagram depicting development of different stages of activities involved in nested association mapping. Figure 1 depicts selection of common founder parent (for example, genotype A) from crop cultivar to cross with donor diverse founder parents (for example, I0 donor parents, namely, B,C,D,E,F,G,H,I,J,K) drawn from germplasm/landraces/wild relatives. It further explains development of NAM population through, SSD, backcross and DH technique to develop RIL-NAM, NIL-NAM, and DH NAM, respectively. It also shows the utilizing NAM population high throughput genotyping and phenotyping and QTL analysis.

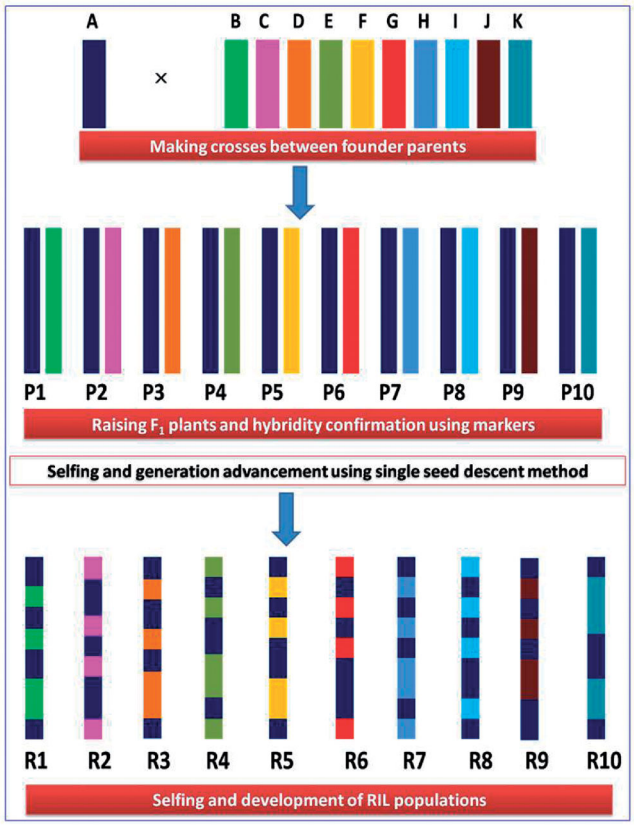

Figure 2. Chromosomal segments of NAM population that looks in each generation: it depicts genome composition of common founder parents and donor founder parents in different generation. P1, P2, P3, P4, P5, P6, P7, P8, P9, and PIO are 10 different F_1 populations derived from 10 different crosses A × B, A × C, A × D, A × E, A × F, A × G, A × H, A × I, A × J, and A × K, respectively. R1, R2, R3 … R10 are the 10 different RIL populations of NAM population derived from 10 different crosses as mentioned above. These 10 RIL populations show mosaic segment of donor founder parents in the background of common founder parents.

for the development of NAM population, multiple alleles and rare alleles are enriched in the population (Yu *et al.*, 2008).

IV. Major considerations in the NAM development approach

Several factors need to be considered in NAM such as a selection of founder parents, the methodology followed in population development (crossing and selfing), mating design, size of the population, the heritability of traits, number of QTLs governing the trait, genetic components of traits (additive effects and epistatic interactions), method of phenotyping, genotyping and statistical tools. These are briefly described below.

A. Selection of common founder parent

The selection of a common founder parent is governed by its morpho-physiological attributes and also target traits under consideration. An elite cultivar of the crop is generally selected as the common founder parent for developing the NAM population. This not only facilitates the mapping of novel QTLs but also leads to the genetic improvement of the elite cultivar. Further, a widely adopted cultivar of a crop as a founder parent makes it possible to evaluate the derived NAM population across multiple locations in different ecologies. In a wide variety of crops, the elite cultivars have been deployed as common founder parents. For example, in maize, *B73* was used as a common founder parent, as it is a reference cultivar used for maize genome sequence and also widely deployed in breeding, genetic and genomic studies (Yu *et al.*, 2008). In barely, *Rasmusson,* a high-yielding and six-rowed spring malting cultivar was used as a common founder parent (Nice *et al.*, 2016). Similarly, IR64 in rice (Fragoso *et al.*, 2017), RTx430 in sorghum (Bouchet *et al.*, 2017), IA3023 (Song *et al.*, 2017) and JS335 in soybean (Shivakumar *et al.*, 2019), LMPG 6 (Bajgain *et al.*, 2016), Berkut (Jordan *et al.*, 2018), Avocet-YrA (Ren *et al.*, 2017) in wheat, Merlo in common bean (Hoyos-Villegas *et al.*, 2016), PBY061 and Zhongshuang 11 (ZS11) in rapeseed (Li, Anja, *et al.*, 2016) were used as common founder parents for the development of NAM population. In addition to mapping of complex traits, the use of popular and well-adapted varieties as common parent opens up the possibility for developing highly diverse breeding material which is a great resource for selecting better performing lines than the common parent for potential replacement in farmer's field.

B. Selection of donor founder parents

The selection of donor founder parents is a critical factor in the development of the NAM population. Donor founder parents should be as diverse as possible and should represent the maximum genetic diversity for target traits in the crop. Further, the donor parents can be selected from adopted cultivars, germplasm, landraces and wild relatives of crops. The donor founder lines can be selected based on genetic diversity, pedigree information, morphological data, agronomic data, adaptability, geographical and ecological data.

In rice, IR64 as common founder parent was taken to represent *indica*, whereas 10 diverse donor lines (including Azucena) were drawn from tropical japonica based on their resistance/tolerance against biotic/abiotic stresses, physiological traits, and genetic diversity (Fragoso *et al.*, 2017). In the case of rapeseed, Hu *et al.* (2018) collected 307 inbred lines representing adaptations to major rapeseed producing countries and evaluated the growth adaptability to the local climate in Wuhan, China. Based on findings, 192 lines were selected for genotyping using 451 SSR markers. Finally, a total of 15 inbred lines representing genetic diversity were chosen as diverse donor parents for the development of the rapeseed NAM population. Similarly, in soybean, 120 lines were selected from the United States, China, Korea, Japan, and other countries based on their high yield, diverse ancestry, drought tolerance features, and were genotyped with Illumina GoldenGate assay containing 1536 SNPs. Based on cluster analysis, 40 lines were selected as donor founder parents for the development of the soybean NAM population (Song *et al.,* 2017). The mini core developed from the core collection represents the maximum genetic diversity of crop plants and paves the way for the selection of diverse lines as donor founder parents for the development of NAM populations (Gireesh *et al.*, 2015; Rakshit and Swapna, 2015). For instance, in the barely NAM population, 25 diverse founder parents were selected from 318 accessions of germplasm collection using SNP markers (Nice *et al.*, 2016). In another study, 10 diverse lines were selected as donor founder parents from the sorghum association panel representing global sorghum diversity for the development of NAM population (Bouchet *et al.*, 2017). Alternatively, donor founder lines can also be chosen based on traits of interest, for instance, stem rust resistant lines were used as donor founder parents for the NAM population in wheat (Bajgain *et al.*, 2016). Depending on the objectives of the geneticist/breeder for developing the NAM

population for multiple agronomic traits or few traits containing the maximum allelic trait diversity, the founder parent is chosen. The NAM populations can also be developed for different subgroups/subspecies separately as was the case in groundnut where separate NAM populations were developed for two market types namely Spanish and Virginia (Pandey *et al.*, 2016).

The number of donor lines to be selected for the development of NAM populations depends on the purpose of the development of NAM. It has been reported to vary from four donor lines in soybean (Li *et al.*, 2017) to 51 donor lines in rapeseed (Li, Anja, *et al.*, 2016). A study of Stich (2009) in maize and *Arabidopsis* has revealed that with a higher number of founder parents used for the development of the NAM population, the power of QTL detection (1-β^*) increases. However, what should be the optimum number of donor lines to be used for the development of NAM population is still a point of further deliberation and the decision lies with the geneticist/breeder considering the target traits and available resources.

C. Population development

Breeding design, size of individual RIL population, and total NAM populations are important criteria to be considered. By increasing the population size, the number of recombination increases, which in turn improves the power of QTL detection (1-β^*) and also ensures precise estimation of allelic effect (Schon *et al.*, 2004; Stich, 2009; Cockram and Mackay, 2018). Nevertheless, the use of an unequal size of sub-populations (individual RIL population) will undermine the power of QTL detection in the NAM approach, because it leads to unbalanced allelic frequencies across the sub-populations. This will in turn reduce the power of QTL detection and therefore, it is important to develop proper and optimum populations/sub-populations for efficient use of the NAM approach (Li, Anja, *et al.*, 2016). Three kinds of NAM populations can be developed based on the breeding design employed (Li, Anja, *et al.*, 2016) which are discussed below.

1. Recombinant inbred Line-NAM (RIL-NAM) populations

The founder and donor parents, selected from germplasm are crossed with common founder parent. Resultant F_1s and their subsequent segregation populations are advanced through the single seed descent (SSD) method until attaining homozygosity to constitute individual RIL populations. This approach was initially employed in maize, wherein a common founder parent B73 was crossed with 25 donor founder parents and developed 25 RIL population to constitute maize NAM population, having 200 RILs from each cross (Yu *et al.*, 2008; McMullen *et al.*, 2009). This approach is suitable for NAM population wherein the donor parents are selected from a crossable set of germplasm and cultivars. This is the most common and popular approach for NAM population development in crops.

2. Backcross-NAM/advanced Backcross-NAM (BC-NAM/AB-NAM) populations

In this approach, donor founder parents selected from unadapted germplasm (wild species or landraces) possessing some special characters (for e.g., resistance/tolerance to stresses) are used as male parents to cross with the elite cultivar (common founder parent) which serve as the female parent. The resultants F_1s are backcrossed with the recurrent parent (common founder parent) to generate backcross progenies. In a subsequent generation, the backcross progenies are advanced through SSD till the population reaches homozygosity (for e.g., BC_1F_6). Such populations are known as BC-NAM or AB-NAM populations. This approach was employed in barley, wherein, 25 wild barley accessions were crossed to elite cultivar Rasmusson to develop 796 $BC_2F_{4:6}$ lines (Nice *et al.*, 2016).

3. Development of Double Haploid-NAM (DH-NAM) populations

The NAM populations can also be developed using the double haploidy approach. The common founder parent is crossed with donor founder parents and resultant F_1s are utilized for the development of DH populations in each cross separately. This approach helps in the development of homozygous lines in a single generation. DH-NAM approach reduces the breeding duration enormously while at the same time, significantly increasing the genetic gain. However, this approach is suitable only for those crops where the well-established protocols for DH development are available.

D. Experimental designs and phenotyping of the NAM population

Appropriate experimental designs should be employed for accurate and efficient estimation of genotypic effects in the mapping population. The most

commonly used experimental design in a field experiment is Randomized Complete Block Design (RCBD), which is useful for a small number of genotypes (<25) (Gomez and Gomez, 1984). However, a large number of genotypes in RCBD will lead to more heterogeneity within the block. Un-replicated designs with systematically spaced checks are the most prominent experimental designs for phenotyping of a large number of genotypes (Federer and Crossa, 2012). Augmented designs are the class of incomplete block designs, useful for controlling variability and estimation of experimental errors. The NAM populations will have a large number of genotypes, hence augmented designs should be employed for phenotyping. Different types of augmented experimental designs were discussed by Federer in series of papers; Augmented RCBD (Federer, 1961), Row-Column Augmented Design (Federer and Raghavarao, 1975), and Augmented split-plot (Fédérer and Arguillas, 2006). Augmented RCBD is useful when we have a large number of genotypes with one-way elimination of heterogeneity. When the field layout is in row-column or in a square shape, row-column augmented design can be used to eliminate the two-way heterogeneity. To compare the two different factors with varying importance such as genotypes and nutrient dosages, genotypes and insecticides, genotypes and tillage practices etc., an augmented split-plot design is most appropriate. The efficiency of augmented design lies in the selection of several checks, the general recommendation is to use between 10 and 15% of the treatments as checks (Burgueño *et al.* 2018). Yates (1936) proposed that number of checks in an augmented design should be the order of the square root of the genotypes under testing. Therefore, based on the size of the NAM population and the environmental condition of the experimental field, a suitable design with the appropriate number of checks should be employed.

The quality of the phenotyping data generated from mapping populations in target environments will influence the output of mapping studies. Therefore, the deployment of robust and precise phenotyping techniques is one of the most important considerations. It is a well-known fact that the expression of quantitative traits is highly influenced by environmental factors, which makes it necessary for generating phenotypic data from across years (target seasons) and locations (hot spot locations in target geographies) to understand the complexity of trait. Despite advances in agricultural technologies, precise estimation of quantitative traits remains a difficult task and this considerably affects the accurate estimation of the effects of individual QTLs. High throughput phenotyping of mapping populations has become an imperative in the era of high-throughput genomics to maximize gains from vast genomic data generated. High-throughput phenotyping can be undertaken under controlled conditions like glasshouse and growth chambers. Field-based high throughput phenotyping is also an important tool that can be employed for phenotyping a large number of lines if proper controls are included in the phenotyping exercise. In general, the size of the NAM population is higher than the usual biparental populations and association panels, therefore, applications of high throughput phenotypic techniques are indeed essential to realize its full potential. Employing high-throughput phenotypic techniques in NAM for trait characterization saves time, labor, and inputs required (Bouchet *et al.*, 2017; Barbadikar *et al.*, 2019).

E. High throughput genotyping and transcriptomics analysis of NAM populations

High throughput genotyping (HTG) using next-generation sequencing (NGS) technology *viz.* whole genome sequencing, Genotyping-by-sequencing (GBS) whole-genome resequencing, RNA-sequencing or transcriptome sequencing, epigenetic sequencing have been deployed for cost-effectively generating large genomic resources for dissecting complex trait-phenotype relationship (https://nam-genomes.org/). With the advent of time, the NGS tools and algorithms have been improved for genotyping large populations like NAM. The HTG make it possible to genotype a large number of lines with a higher density of markers and increased mapping resolution. GBS has been employed for genotyping several NAM populations of maize, wheat, chickpea resulting in millions of SNPs and increases the mapping resolution (Varshney *et al.*, 2014). The SNP-based arrays can be well utilized for genotyping the NAM populations and the significant SNP-trait associations physically linked to the QTLs can be identified.

Ploidy, large genome size, heterozygosity, repetitive sequences and cost of genotyping are some of the important determining factors for applications of HTG in NAM (Ray and Satya, 2014). Common founder and donor founder lines have been sequenced using NGS platforms for maize, wheat, soyabean, barley, oilseed rape, chickpea (https://www.maizegdb.org/NAM_project, https://www.soybase.org/SoyNAM/, http://hickeylab.com/wheat-nam-population/).

The RNA-seq or transcriptome sequencing captures the spatial and temporal expression of genes of tissues in a particular developmental stage, condition or stress. This enables a better understanding of the genes underlying the QTLs/candidate genes identified through GWAS. The genomic data combined with the transcriptome data is a powerful and efficient approach for the identification of causal candidate genes with the phenotypic trait. The SNPs identified from founder lines subjected to different conditions can be effectively used to genotype the NAM population. Moreover, such data can be combined with genetic data for understanding the basis of a stress response. Lin *et al.* (2017) sequenced the transcriptomes of maize tissues (seedling shoot apex, immature unpollinated ears, immature tassels, seedling shoots, and roots) from 27 diverse NAM founder lines to identify the gene expression patterns and variations. This study reported development of a complementary method to SNP-based GWAS called eRD-GWAS, a Bayesian-based method for GWAS. The gene expression of transcription factors was compared with other genes within and among the NAM founder parents and significant association was noticed with phenotypic traits retrieved from a panel of 369 maize diverse inbreds. Li, Sun, *et al.*, 2016 used a combined approach for dissecting the genomic regions associated with drought tolerance traits in maize NAM populations using GWAS, joint QTL mapping, GBS, and transcriptome sequencing. The mapping resolution was increased by constructing high-density recombination maps using GBS data. The GBS generated SNPs underlying the identified genomic regions were associated with drought tolerance traits and certain genes harboring SNPs were highly upregulated in the inbred line B73 under drought stress. The combined strategy of high throughput genotyping, phenotyping, and transcriptome sequencing can detect the causal genes as well as decipher the regulatory variations operating for a specific trait of interest. It is suggested that a combination of two or more technologies like genotyping by sequencing and re-sequencing can detect the causal region with higher resolution and accuracy.

F. Statistical tools for analyzing NAM populations

The QTL analysis for understanding the genetic architecture of complex traits is generally carried out by linkage analysis in bi-parental mapping populations and GWAS in the natural population (Li, Sun, *et al.*, 2016). Statistical approaches *viz.*, Single Marker Analysis (SMA), Interval Mapping (IM), Composite Interval Mapping (CIM), and Inclusive Composite Interval Mapping (ICIM) have been employed for QTL analysis (Soller *et al.*, 1976; Lander and Botstein, 1989; Haley and Knott, 1992; Zeng, 1994; Li *et al.*, 2007), but these methods are mainly designed for individual bi-parental mapping populations. Linkage analysis in bi-parental populations will only detect two alleles per locus and has poor mapping resolution (Li, Sun, *et al.*, 2016). Whereas, population structure and genetic relatedness will often lead to spurious associations in GWAS. The NAM is primarily designed to address some of the inherent limitations of linkage analysis and GWAS.

The nested association mapping requires specially designed statistical tools for achieving the high power and high-resolution mapping of complex traits. While the NAM population is derived from multiple parents, all the families need to be taken together for analysis to realize the advantages of the breeding design. Three approaches currently being employed in the NAM design are joint linkage analysis and NAM-GWAS. These are briefly described below.

1. Joint linkage analysis

Joint linkage analysis refers to QTL analysis of complex traits across all families of the NAM population. Two methods viz., Joint Composite Interval Mapping (JCIM) and Joint Inclusive Composite Interval Mapping (JICIM) are being employed for joint linkage analysis of complex traits in various NAM studies (Table 4).

The joint QTL analysis has a high power of QTL detection (Buckler *et al.*, 2009; Li *et al.*, 2011) and also identifies effects of more than two alleles per locus as compared to individual family analysis (Sneller *et al.*, 2009; Ogut *et al.*, 2015). For instance, joint stepwise regression and JICIM were employed for mapping flowering time in the maize NAM population. The joint QTL analysis has identified nearly twice as many QTL effects as identified by individual family-wise (Buckler *et al.*, 2009). In the case of *Arabidopsis*, JICIM has identified nine QTLs contributing 83% to the phenotypic variance of flowering time wherein, LOD scores were much higher than the LOD score of these nine QTL identified from individual families (Li *et al.*, 2011). Further, among the nine QTLs identified in *Arabidopsis* for flowering time, four QTLs were significant only in one family indicating the higher power of detecting such rare QTL using joint linkage QTL analysis (Li *et al.*, 2011). Li, Anja, *et al.* (2016) developed a JCIM algorithm with a combination of LASSO (Least absolute shrinkage and selection operator) regression model for the selection

Table 4. QTL analysis methods in NAM population.

S No.	Mapping design/Method	Cofactor selection method	Crop	Traits	References
1.	Joint composite interval mapping (JCIM)	LASSO regression model	Rapeseed (*Brassica napus* L.)		Li, Anja, *et al.*, 2016
2.	Joint inclusive composite interval mapping (JICIM)	Regression model	Common bean	Agronomic traits	Hoyos-Villegas *et al.*, 2016.
		Joint stepwise regression model	Maize	Flowering time	Buckler *et al.*, 2009
			Groundnut	Seed weight	Gangurde *et al.*, 2020
		Forward stepwise regression/ principle component regression	Maize	Carbon and nitrogen metabolism in maize	Zhang *et al.*, 2015

of co-factors (nontargeted QTLs), while carrying out QTL mapping in rapeseed. The JCIM model showed higher power of QTL detection than the existing JICIM model due to the higher efficiency of the cofactor selection method.

2. NAM-GWAS

The GWAS is a popular technique employed in natural populations for high-resolution mapping of complex traits (Rafalski, 2002). However, the genetic relatedness and population structure exist in the natural population will lead to the spurious marker-trait association (Yu and Buckler, 2006). Besides, rare alleles and small effect loci are difficult to detect in GWAS (Lu *et al.*, 2010). These limitations of GWAS in a natural population can be addressed by employing GWAS in the NAM population (known as NAM-GWAS). Population structure is reduced in the NAM population due to recombination during the development of populations. A study in sorghum showed that the presence of a low level of population structure in the NAM population as compared to the association panel (Bouchet *et al.*, 2017). While in Brassica, the population stratification/clustering was significantly low in the NAM population while it was completely absent in the MAGIC population (Hu *et al.*, 2018). Although, recombination events during the development of NAM population breakdown traits correlations due to LD, the NAM population still retains low population structure (Bouchet *et al.*, 2017).

The frequency of rare alleles derived from one or few diverse donor parents will be enriched in the individual population to a detectable level. Consequently, such rare alleles could be detectable by joint QTL analysis and NAM-GWAS. The NAM has a higher power of QTL detection than GWAS under high and low heritability, small population size, low effects of QTLs. However, GWAS has a high power of QTL detection than NAM under large sample size and high heritability which could be due to loss of allelic diversity in the NAM population as compared to association panel (Bouchet *et al.*, 2017). By increasing the common founder parents, we can increase the power of QTL detection and reduce the FDR (false discovery rate) in NAM (Bouchet *et al.*, 2017). High-resolution mapping can be achieved due to preserving historical LD within the novel recombinations. The LD could be decayed during the development of the population due to recombination between loci and QTLs. Small LD blocks offer a better resolution for identifying causal candidate genes, while the contrary is the case for large LD blocks. In linkage analysis, due to limited recombination during development of the population, the closely spaced QTLs or alleles will lead to a fused QTL signal, while NAM-GWAS will separate into different components due to historical recombinants present in the founder parents (Tian *et al.*, 2011). Several studies have employed NAM-GWAS for mapping complex traits in NAM populations *viz.*, maize (Kump *et al.*, 2011; Poland *et al.*, 2011; Tian *et al.*, 2011; Peiffer *et al.*, 2014; Zhang *et al.*, 2015; Lin *et al.*, 2017); barely (Schnaithmann *et al.*, 2014; Maurer *et al.*, 2015, Nice *et al.*, 2016; Saade *et al.*, 2017; Vatter *et al.*, 2017) Sorghum (Bouchet, *et al.*, 2017) and soybean (Li *et al.*, 2017; Diers *et al.*, 2018).

Several studies have employed both joint linkage analysis and NAM-GWAS independently for mapping complex traits using NAM population. However, the findings of joint linkage analysis and NAM-GWAS in some cases are not similar (Tian *et al.*, 2011). For instance, Tian *et al.*, 2011 employed both joint linkage analysis and NAM-GWAS for mapping leaf architecture traits in maize. The joint linkage analysis with 1106 markers identified 30, 34, and 36 QTLs for upper leaf angle, leaf width, and leaf length explaining 74, 80, and 77% phenotypic variations respectively. While NAM-GWAS studies using 1.6 million tested SNPs have detected 203, 287, and 295 significant SNPs for upper leaf angle, leaf length and leaf width, respectively. But studies in the maize NAM population (Kump *et al.*, 2011, Olukolu *et al.*, 2014) reveal that GWAS-associated SNPs mostly localize to the joint linkage QTL. The QTLs have been shown to contain multiple causal genes in these NAM studies and other biparental studies. The joint linkage analysis has been shown to narrow QTL interval as compared to single linkage analysis.

3. Combined linkage analysis and linkage disequilibrium mapping approach

To further improve the statistical power of QTL analysis, several studies employed combined linkage analysis and linkage disequilibrium mapping strategy to overcome the limitations of using linkage analysis and LD mapping independently. The combined linkage analysis and linkage disequilibrium mapping approach is a very powerful statistical tool for mapping complex traits wherein both linkage and LD between markers and QTLs will be estimated. It captures information about the linkage of markers based on recombination frequency and LD created historically. Consequently, it greatly enhances the mapping resolution and power of QTL detection (Wu and Zeng, 2001; Wu *et al.*, 2002; Bardol *et al.*, 2013). The integrated approaches have been employed in humans and animals for mapping complex traits (Allison 1997; Rabinowitz 1997; Meuwissen and Goddard, 2001; Wu and Zeng, 2001; Meuwissen *et al.*, 2002).

In case of plants, Lu *et al.*, 2010 employed a combined linkage and LD mapping strategy in three independent RIL populations and 305 inbred lines for mapping drought tolerance in maize. The integrated mapping identified 18 additional QTLs that were not detected either by linkage analysis or LD mapping indicating the higher mapping efficiency of the integrated mapping strategy. Further, the study also showed that the effect of QTL detected in integrated mapping is larger as compared to linkage mapping and LD mapping. For instance, linkage mapping and haplotype-based LD mapping could able to detect QTL with PVE up to 22.7% while integrated mapping strategy identified QTL with PVE (phenotypic variations expalined) up to 34.7%. While Bardol *et al.*, 2013 compared different mapping strategies and reported the higher efficiency of integrated mapping strategy in multi-parent mapping population for unraveling the complex traits (Lu *et al.*, 2010). It is a pertinent note that combined linkage analysis and linkage disequilibrium mapping strategy should be an appropriate statistical strategy to be employed in NAM population for realizing higher resolution and higher power of QTL detection.

V. Applications of NAM populations in crop improvement

A. Understanding the genetic architecture of complex traits

Quantitative traits are controlled by several genomic regions with small effects. By employing the principles of linkage analysis and GWAS, the NAM enables the identification of genomic regions associated with complex traits at higher resolution with a higher power of QTLs detection. The use of several diverse founder parents enables the detection of multiple alleles and rare alleles. The large-sized NAM population provides high-quality phenotypic data that helps in understanding genetic variations underlying complex traits. For example in maize, flowering time is a complex trait that controls the adaptability of maize cultivars to different environments. The NAM in maize (Buckler *et al.*, 2009) revealed that unlike other crops flowering time in maize is controlled by numerous small additive QTLs with genetic and environmental interactions. In another study, a total of 208 SNPs associated with northern leaf blight resistance in maize NAM population was identified, of which, five SNPs were located in or adjacent to LRR receptor-like-kinase genes (Poland *et al.*, 2011). In groundnut, using two NAM populations, 11 and 8 major effect QTLs were detected for pod weight (PW) and seed weight (SW), respectively in NAM_Tifrunner, while 13 and 11 major effect QTLs for PW and SW, respectively, in another NAM population (called NAM_Florida-07) (Gangurde *et al.*, 2020). Most of the QTLs associated with PW and SW were co-localized suggesting these traits regulated by a common set of candidate genes.

B. Discovery of novel alleles/QTLs/genes from unadapted germplasm, landraces and wild species of crops and broadening the genetic base of crop plants

Unadapted germplasm, landraces, and wild species of crops are known to possess novel genes/QTLs for several important traits and valuable resource for broadening the genetic base of crop plants. By using principles of the AB-QTL approach and NAM, the AB-NAM enable the simultaneous introgression and detection of genomic regions associated with complex traits at high resolution and high power mapping. The AB-NAM population is a robust resource for exploring unadapted germplasm for crop improvement. The barley AB-NAM population developed from 25 accessions of wild species were utilized for mapping of the glossy spike, glossy sheath, black hull color, and grain protein content (Nice *et al.*, 2016). Further, a major QTL allele on 2H chromosome derived from wild barley was identified which increases 37% yield under saline stress conditions. The study also identified SNP (BOPA2_12_30822) in the alpha-glucosidase gene (AK375658) associated with the major QTL (Saade *et al.*, 2016). The TeoNAM (Teosinte NAM)

population in maize developed from the wild progenitor of maize, Teosinte (Chen *et al.*, 2019). Genetic studies in TeoNAM revealed 255 QTLs for 22 domestication and agronomic traits, many of these either located at earlier reported regions or novel candidate genes. The AB-NAM helps in bringing diverse novel genomic regions and traits in the background of an elite cultivar (i.e., common founder parent) and enhances the genetic base of the crop plants.

C. Identification of candidate genes

Applications of omics in NAM has tremendous ability to unearth the candidates genes underlying complex traits. Using the NAM, four candidate genes on chromosome 3, *viz.*, CCCH-type zinc finger protein, MIKC-type MADS-box protein (*OsMADS50*), DNA-binding with one finger 12 gene (OsDof12) and Rice phytochrome B (*OsPhyB*) responsible for days to heading in rice were identified (Fragoso *et al.*, 2017). Similarly, in groundnut using two NAM populations and SNP-based array, candidate genes having SNP-trait associations for 100-pod and 100-seed weight were identified (Gangurde *et al.,* 2020). With the availability of low-cost and affordable sequencing, such results are expected to increase in coming years and expected to accelerate gene discovery and marker development.

D. Delineation of traits for pleiotropy vis-à-vis linkage

Co-inheritance of QTLs/traits is caused by either due to linkage (physically linked) or pleiotropy (multiple effects of single gene). The occurrence of recombination events in large-sized NAM population will help in delineating linkage from pleiotropy. The negative correlation between flowering time and northern leaf blight resistance was detected among maize founder parents. This negative correlation was broken in the NAM population due to genome reshuffling. This showed that the strong phenotypic association between flowering time and northern leaf blight resistance among founder parents is due to confounding population structure coupled with weak linkage within families rather than pleiotropy (Poland *et al.*, 2011).

E. Mapping of recombination events and segregation distortions

The SNP data of founder parents and NAM population can be utilized for estimation of segregation distortion and recombination events. As NAM involves diverse parents and wild species, segregation distortion will be evident in the population. The common segregation distortion loci across the population and the loci specific to a particular population can be estimated in NAM population. In bi-parental mapping population, if loci have the same allele in both the parents, it is not possible to estimate recombination events, while in the case of the NAM population, due to the use of diverse founder parents, if such loci have allelic variations, the recombination events can be estimated through joint linkage analysis. Allelic frequency in each family can be estimated which is expected to segregate in a 1:1 ratio. Chi-square test (χ^2) can be carried out to determine whether observed allelic frequency segregates as per expected frequency. Through a joint analysis of the NAM population, the segregation distortion can be dissected and characterized.

A rice NAM population developed from IR64 and tropical japonica (10 accessions) identified segregation distortion on chromosomes 3, 6, 7 and 9. A detailed study of the regions on Chromosome 3 associated with segregation distortions showed that three genes namely ras-related protein (regulates pollen tube growth in *Arabidopsis*), actin-depolymerizing factor (essential for pollen tube growth elongation) and another gene region homologous to maize pollen gene aberrant pollen 1 contribute for segregation distortions (Fragoso *et al.*, 2017). Further, it was utilized for the estimation of recombination and an average number of recombination was found to be 18.9 with a standard deviation of 10.9 (Fragoso *et al.*, 2017).

F. Detection of rare alleles and minor QTLs

The NAM approach maximizes the detection of rare alleles and minor QTLs due to larger population sizes derived from diverse parents. It increases the frequency of rare alleles to a detectable level as compared to the natural population. However, this depends on the number of founder parents possessing the rare allele and their frequencies in the corresponding NAM populations (Nice *et al.*, 2016). A simulation study by Guo *et al.,* 2010 showed that the NAM approach has adequate power to precisely identify the functional markers associated with minor QTLs contributing at least 5% of the phenotypic variation and segregating in at least five RIL families of the NAM population derived from 28 crosses. In the maize NAM population, 32 QTLs for southern leaf blight resistance were identified with small effects and the absolute values of allelic effects ranged from 0.09 to

0.39 on the nine-point scale (Kump *et al.*, 2011). In another study, Sharma *et al.* (2018) identified 96 QTLs for yield-related traits under varying nitrogen levels in barley NAM population.

VI. NAM populations for genomic selection

Genomic selection predicts the phenotype of selected individuals (in nonphenotyped breeding population) using GEBV (genomic estimated breeding value) derived from genome-wide marker loci and phenotypic data of training population (Meuwissen *et al.*, 2001). It is gaining importance for genetic improvement of quantitative traits in plant species owing to its advantage over conventional breeding and marker-assisted selection approaches (Nakaya and Isobe, 2012). The effect of an individual marker is estimated in the training population using appropriate prediction models. The training population is a set of individuals that is phenotyped for quantitative traits and genotyped using genome-wide markers. Size and genetic relatedness between training population and breeding population, the population structure of training population are important factors that influence the accuracy of genomic prediction (Habier *et al.*, 2007; Nakaya and Isobe, 2012). Investigations have been undertaken in the NAM population of soybean and maize to study the impact of training population size, marker density, and prediction models on the accuracy of genomic prediction (Xavier *et al.*, 2016; Bian and Holland, 2017; Rincent *et al.*, 2017). NAM population is a structured population that can be employed for prediction across families. Thus, the NAM population can be considered an important genetic resource for genomic selection.

VII. Constraints in deploying NAM approach

The NAM offers great prospects to unravel the genetic architecture of complex traits as has been demonstrated well in maize, barley, sorghum, rice, groundnut, and in few other crops as described in earlier sections. However, several constraints need to be addressed for better utilization of the approach.

The selection of donor founder parents, that harbors functionally distinct alleles as compared to the common founder parent is pivotal for the development of NAM families. It is necessary to select subsets of founder lines with the maximum diversity based on previous simulation algorithms and genealogy of core sets. Hence, a thorough understanding of core set, whole germplasm based on large morphological and molecular data is required for deciding the number of donor founder parents and also for a selection of appropriate diverse founder parents. The cross-compatibility of diverse founder lines with common founder parents is a prerequisite for the development of NAM populations. The species originating in different ecologies have different adaptation mechanisms and heterogeneity, which can affect the cross-compatibility. The utilization of wild relatives as donor founder parents leads to segregation distortion and requires more time for population stabilization. It also involves carrying forward large segregating populations of individual crosses and multi-location evaluation, requiring a large experimental field area and resource inputs. It is crucial to standardize the optimum number of diverse founder parents, size of the NAM population depending on crop species, ploidy level and the number of SNPs required for high throughput genotyping.

Haplotype diversity in the NAM population depends on the number of diverse parents used and is higher than the biparental population. However, haplotype diversity of the NAM population is lower as compared to the association panel. This could be addressed by cautious selection of diverse parents and also by increasing the number of parents (Guo *et al.*, 2010; Cockram and Mackay, 2018). For enhancing the haplotype diversity, it has been suggested to increase the number of common founder (Guo *et al.*, 2010), Multi-location evaluation of the NAM populations in replications enables higher power of QTL detection (Stich *et al.,* 2010). Based on the availability of resources, one has to decide whether to use RIL-NAM, AB-NAM or DH-NAM. We suggest the speed breeding technique for accelerating breeding cycles during the development of NAM populations.

VIII. Prospects of NAM populations in crop plants

Recent advances witnessed in genomic technologies are improving our understanding of genetic architecture of complex traits and they are increasingly being utilized for genetic improvement of various crops. Next-generation sequencing tools can be better employed in NAM populations to identify the diverse QTLs/genes and their favorable alleles and further use them for functional characterization, gene identification and finally for trait improvement through breeding. Further, NAM population has the potential to be utilized for genomic selection for rapid genetic improvement of quantitative traits. The breeding process of development of NAM populations generates a

large number of breeding resources that would be a valuable resource for plant breeders and geneticists. Due to the unparallel importance of NAM approaches in molecular breeding and genomics assisted breeding, higher emphasis should be given for the development of NAM populations in many other crops and intensify such efforts in all the major crops. However, in this regard, the ease in making more number of crosses, population stabilization times etc. are important considerations and may be a limiting factor in many crops. In crops like maize where DH technology is well established, developing countries also need to invest in this to generate NAM populations, which will serve not only as an important resource for gene/QTL identification but also as a valuable resource for breeding and genetic studies. However, in the process of the development of NAM resources in hybrid based crops, care must be taken so that the heterotic groups do not get mixed up while selecting diverse parents. Creating separate NAM populations for contrasting heterotic groups is indeed a good strategy. Global and national platforms should accelerate efforts in the development of NAM populations so that the NAM approach can be better exploited for genetic improvement of various major food and commercial crops. Finally, we would also like to reiterate that NAM populations are community resources that may be freely shared with researchers across the globe for their gainful utilization.

Acknowledgments

The authors express sincere gratitude and immensely thank the reviewers of the journal for critical suggestions for improving the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Author contributions

CG, SMA, PS, ER, RS: Conceptualization of manuscript and drafted the manuscript; CG, SMA, PS, BMK, RS: Section-wise drafted the manuscript; MS, LVS, RMS, MSM, MS, KRY, BMK, TKM, SR: Critically edited the manuscript. All authors have read the manuscript and critically evaluated.

Funding

Authors are grateful to the Indian Council of Agricultural Research (ICAR) for the support received for the preparation of the manuscript.

References

Allison, D. B. 1997. Transmission-disequilibrium tests for quantitative traits. *Am. J. Hum. Genet.* **60**: 676–690.

Anderson, S. L., Mahan, A. L., Murray, S. C., and Klein, P. E. 2018. Four parent maize (FPM) population: effects of mating designs on linkage disequilibrium and mapping quantitative traits. *Plant Genome.* **11**: 170102.

Bajgain, P., Rouse, M. N., Tsilo, T. J., Macharia, G. K., Bhavani, S., Jin, Y., and Anderson, J. A. 2016. Nested association mapping of stem rust resistance in wheat using genotyping by sequencing. *PLOS One.* **11**: e0155760.

Barbadikar, K. M., Balakrishnan, D., Gireesh, C., Kardile, H., Bosamia, T., and Mishra, A. 2019. High-throughput phenotyping: potential tool for genomics. In *OMICS Based Approaches in Plant Biotechnology*. Srivener Publishing: Beverly, Massachusetts, pp 303–321.

Bardol, N., Ventelon, M., Mangin, B., Jasson, S., Loywick, V., Couton, F., Derue, C., Blanchard, P., Charcosset, A., and Moreau, L. 2013. Combined linkage and linkage disequilibrium QTL mapping in multiple families of maize (*Zea mays* L.) line crosses highlights complementarities between models based on parental haplotype and single locus polymorphism. *Theor. Appl. Genet.* **126**: 2717–2736.

Bazakos, C., Hanemian, M., Trontin, C., Jiménez-Gómez, J. M., and Loudet, O. 2017. New strategies and tools in quantitative genetics: how to go from the phenotype to the genotype. *Annu. Rev. Plant Biol.* **68**: 435–455.

Bian, Y., and Holland, J. B. 2017. Enhancing genomic prediction with genome-wide association studies in multi-parental maize populations. *Heredity.* **118**: 585–593.

Bouchet, S., Olatoye, M. O., Marla, S. R., Perumal, R., Tesso, T., Yu, J., Tuinstra, M., and Morris, G. P. 2017. Increased power to dissect adaptive traits in global sorghum diversity using a nested association mapping population. *Genetics.* **206**: 573–585.

Buckler, E. S., Holland, J. B., Bradbury, P. J., Acharya, C. B., Brown, P. J., Browne, C., Ersoz, E., Garcia, F. S., Garcia, A., Glaubitz, J. C., Goodman, M. M., Harjes, C., Guill, K., Kroon, D. E., Larsson, S., Lepak, N. K., Li, H., Mitchell, S. E., Pressoir, G., Peiffer, J. A., Rosas, M. O., Rocheford, T. R., Romay, M. C., Romero, S., Salvo, S., Villeda, S. H., da Silva, H. S., Sun, Q., Tian, F., Upadyayula, N., Ware, D., Yates, H., Yu, J., Zhang, Z., Kresovich, S., and McMullen, M. D. 2009. The genetic architecture of maize flowering time. *Science.* **325**: 714–718.

Burgueño, J., Crossa, J., Rodríguez, F., Yeater, K. M., 2018. Chapter 13: Augmented designs–experimental designs in which all treatments are not replicated. In *Applied Statistics in Agricultural, Biological, and Environmental Sciences*, Glaz, B. and Yeater, K. M. Eds. John Wiley & Sons.

Cavanagh, C., Morell, M., Mackay, I., and Powell, W. 2008. From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Curr. Opin. Plant Biol.* **11**: 215–221.

Chen, Q., Yang, C. J., York, A. M., Xue, W., Daskalska, L. L., DeValk, C. A., Krueger, K. W., Lawton, S. B., Spiegelberg, B. G., Schnell, J. M., Neumeyer, M. A., Perry, J. S., Peterson, A. C., Kim, B., Bergstrom, L., Yang,

L., Barber, I. C., Tian, F., and Doebley, J. F. 2019. TeoNAM: a nested association mapping population for domestication and agronomic trait analysis in maize. *Genetics.* **213**: 1065–1078.

Choudhary, M., Wani, S. H., Kumar, P., Bagaria, P. K., Rakshit, S., Roorkiwal, M., and Varshney, R. K. 2019. QTLian breeding for climate resilience in cereals: progress and prospects. *Funct. Integr. Genomics.* **19**: 685–701.

Churchill, G. A., Airey, D. C., Allayee, H., Angel, J. M., Attie, A. D., Beatty, J., Beavis, W. D., Belknap, J. K., Bennett, B., Berrettini, W., Bleich, A., Bogue, M., Broman, K. W., Buck, K. J., Buckler, E., Burmeister, M., Chesler, E. J., Cheverud, J. M., Clapcote, S., Cook, M. N., Cox, R. D., Crabbe, J. C., Crusio, W. E., Darvasi, A., Deschepper, C. F., Doerge, R. W., Farber, C. R., Forejt, J., Gaile, D., Garlow, S. J., Geiger, H., Gershenfeld, H., Gordon, T., Gu, J., Gu, W., de Haan, G., Hayes, N. L., Heller, C., Himmelbauer, H., Hitzemann, R., Hunter, K., Hsu, H.-C., Iraqi, F. A., Ivandic, B., Jacob, H. J., Jansen, R. C., Jepsen, K. J., Johnson, D. K., Johnson, T. E., Kempermann, G., Kendziorski, C., Kotb, M., Kooy, R. F., Llamas, B., Lammert, F., Lassalle, J.-M., Lowenstein, P. R., Lu, L., Lusis, A., Manly, K. F., Marcucio, R., Matthews, D., Medrano, J. F., Miller, D. R., Mittleman, G., Mock, B. A., Mogil, J. S., Montagutelli, X., Morahan, G., Morris, D. G., Mott, R., Nadeau, J. H., Nagase, H., Nowakowski, R. S., O'Hara, B. F., Osadchuk, A. V., Page, G. P., Paigen, B., Paigen, K., Palmer, A. A., Pan, H.-J., Peltonen-Palotie, L., Peirce, J., Pomp, D., Pravenec, M., Prows, D. R., Qi, Z., Reeves, R. H., Roder, J., Rosen, G. D., Schadt, E. E., Schalkwyk, L. C., Seltzer, Z., Shimomura, K., Shou, S., Sillanpää, M. J., Siracusa, L. D., Snoeck, H.-W., Spearow, J. L., Svenson, K., Tarantino, L. M., Threadgill, D., Toth, L. A., Valdar, W., de Villena, F. P.-M., Warden, C., Whatley, S., Williams, R. W., Wiltshire, T., Yi, N., Zhang, D., Zhang, M., and Zou, F., Complex Trait Consortium 2004. The collaborative cross, a community resource for the genetic analysis of complex traits. *Nat. Genet.* **36**: 1133–1137.

Cockram, J., and Mackay, I. 2018. Genetic mapping populations for conducting high-resolution trait mapping in plants. *Adv Biochem. Eng. Biotechnol.* **164**: 109–138.

Cook, J. P., McMullen, M. D., Holland, J. B., Tian, F., Bradbury, P., Ross-Ibarra, J., Buckler, E. S., Flint-Garcia, S. A. 2012. Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiol.* **158**: 824–834.

Darvasi, A., and Soller, M. 1995. Advanced intercross lines, an experimental population for fine genetic mapping. *Genetics.* **141**: 1199–1207.

Diepenbrock, C. H., Kandianis, C. B., Lipka, A. E., Lundback, M. M., Vaillancourt, B., Castillo, E. G., Wallace, J. G., Cepela, J., Mesberg, A., Bradbury, P. J., Ilut, D. C., Hernandez, M. M., Hamilton, J., Owens, B. F.,Tiede, T., Buckler, E. S., Rocheford, T., Buell,C. R., Gore, M. A., Penna, D. 2017. Novel Loci Underlie Natural Variation in Vitamin E Levels in Maize Grain. *Plant Cell.* **29**: 2374–2392.

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.* **8**: 3367–3375.

Federer, W. T. 1961. Augmented designs with one-way elimination of heterogeneity. *Biometrics.* **17**: 447–473.

Fédérer, W. T., Jr., and Arguillas, F. O. 2006. Augmented split plot experiment design. *J. Crop Improv.* **15**:81–96.

Federer, W. T., and Crossa, J. 2012. I. 4 Screening experimental designs for quantitative trait loci, association mapping, genotype-by environment interaction, and other investigations. *Front. Physiol.* **3**: 1–8.

Federer, W. T., and Raghavarao, D. 1975. On augmented designs. *Biometrics.* **31**:29–35.

Fragoso, C. A., Maria, M., Zuoheng, W., Christopher, H., Lady, J. A., John, A. A., Natalia, F., Luz, E. R., Karine, L., Zhao, H., Stephen, L. D., and Mathias, L. 2017. Genetic architecture of a rice nested association mapping population. *G3.* **7**:1913–1926.

Gangurde, S. S., Wang, H., Yaduru, S., Pandey, M. K., Fountain, J. C., Chu, Y., Isleib, T., Holbrook, C. C., Xavier, A., Culbreath, A. K., Ozias-Akins, P., Varshney, R. K., and Guo, B. 2020. Nested-association mapping (NAM)-based genetic dissection uncovers candidate genes for seed and pod weights in peanut (*Arachis hypogaea*). *Plant Biotechnol. J.* **18**: 1457–1471.

Gelderman, H. 1975. Investigations on inheritance of quantitative characters in animals by gene markers. *I. Methods. Theor. Appl. Genet.* **46**: 319–330.

Gireesh, C., Husain, S. M., Shivakumar, M., Satpute, G., Giriraj, K., Arya, M., Agarwal, D. K., and Bhatia, V. S. 2015. Integrating principal component score strategy with power core method for development of core collection in Indian soybean germplasm. *P. Gen. Res.* **15**: 1–9.

Gomez, K. A., and Gomez, A. A. 1984. *Statistical Procedures for Agricultural Research*, 2nd ed. NewYork: John Wiley and Sons, pp 680.

Guo, B. H., Sleper, D. A., and Beavis, W. D. 2010. Nested association mapping for identification of functional markers. *Genetics.* **186**: 373–383.

Guo, Z., Tucker, D. M., Wang, D., Basten, C. J., Ersoz, E., Briggs, W. H., Lu, J., Li, M., and Gay, G. 2013. Accuracy of across-environment genome-wide prediction in maize nested association mapping populations. *Genes Genom. Genet.* **3**: 263–272.

Habier, D., Fernando, R. L., and Dekkers, J. C. 2007. The impact of genetic relationship information on genome-assisted breeding values. *Genetics.* **177**: 2389–2397.

Haley, C. S., and Knott, S.A. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity.* **69**: 315–324.

Hoyos-Villegas, V., Song, Q., Wright, E. M., Beebe, S. E., and Kelly, J. D. 2016. Joint linkage QTL mapping for yield and agronomic traits in a composite map of three common bean RIL populations. *Crop Sci.* **56**: 2546–2563.

Hu, J., Guo, C., Wang, B., Ye, J., Liu, M., Wu, Z., Xiao, Y., Zhang, Q., Li, H., King, G. J., and Liu, K. 2018. Genetic properties of a nested association mapping population constructed with semi-winter and spring oilseed rapes. *Front Plant Sci.* **9**: 1740.

Jannink, J. L., and Jansen, R. C. 2000. The diallel mating design for mapping interacting QTLs. In *Quantitative Genetics and Breeding Methods: The Way Ahead*; INRA, ed. Paris: Plant Research International, Centre for Biometry, pp 81–88.

Jordan, K. W., Wang, S., He, F., Chao, S., Lun, Y., Paux, E., Sourdille, P., Sherman, J., Akhunova, A., Blake, N. K., Pumphrey, M. O., Glover, K., Dubcovsky, J., Talbert, L., and Akhunov, E. D. 2018. The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping. *Plant J.* **95**: 1039–1054.

Kearsey, M. J., and Farquhar, A. G. 1998. QTL analysis in plants; where are we now? *Heredity.* **80**: 137–142.

Kidane, Y. G., Gesesse, C. A., Hailemariam, B. N., Desta, E. A., Mengistu, D. K., Fadda, C., Pè, M. E., and Dell'Acqua, M. 2019. A large nested association mapping population for breeding and quantitative trait locus mapping in Ethiopian durum wheat. *Plant Biotechnol. J.* **17**: 1380–1393.

Knowler, W. C., Williams, R. C., Pettitt, D. J., and Steinberg, A. G. 1988. Gm3;5,13,14 and type 2 diabetes mellitus: an association in American Indians with genetic admixture. *Am. J. Hum. Genet.* **43**: 520–526.

Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., Durrant, C., and Mott, R. 2009. A multiparent advanced generation intercross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* **5**: e1000551.

Kump, K. L., Bradbury, P. J., Wisser, R. J., Buckler, E. S., Belcher, A. R., Oropeza-Rosas, M. A., Zwonitzer, J. C., Kresovich, S., McMullen, M. D., Ware, D., Balint-Kurti, P. J., and Holland, J. B. 2011. Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat. Genet.* **43**: 163–168.

Ladejobi, O., Elderfield, J., Gardner, K. A., Gaynor, R. C., Hickey, J., Hibberd, J. M., Mackay, I. J., and Bentley, A. R. 2016. Maximizing the potential of multi-parental crop populations. *Appl. Transl. Genom.* **11**: 9–17.

Lander, E.S., and Botstein, D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics.* **121**: 185–199.

Li, C. H., Li, Y. X., Bradbury, P. J., Wu, X., Shi, Y. S., Song, Y. C. 2015. Construction of high-quality recombination maps with low-coverage genomic sequencing for joint linkage analysis in maize. *BMC Biol.* **13**: 78–90.

Li, C., Sun, B., Li, Y., Liu, C., Wu, X., Zhang, D., Shi, Y., Song, Y., Buckler, E. S., Zhang, Z., Wang, T., and Li, Y. 2016. Numerous genetic loci identified for drought tolerance in the maize nested association mapping populations. *BMC Genomics.* **17**: 894.

Li, H., Bradbury, P., Ersoz, E., Buckler, E.S., and Wang, J. 2011. Joint QTL Linkage Mapping for Multiple-Cross Mating Design Sharing One Common Parent. *PLOS One.* **6**: e17573.

Li, H., Ye, G., and Wang, J. 2007. A modified algorithm for the improvement of composite interval mapping. *Genetics.* **175**: 361–374.

Li, J., Anja, B., Viola, S., and Benjamin, S. 2016. Comparison of statistical models for nested association mapping in rapeseed (*Brassica napus* L.) through computer simulations. *BMC Plant Biol.* **16**: 26.

Li, S., Yongce, C., He, J., Zhao, T., and Gai, J. 2017. Detecting the QTL-allele system conferring flowering date in a nested association mapping population of soybean using a novel procedure. *Theor. Appl. Genet.* **130**: 2297–2314.

Lin, H. Y., Liu, Q., Li, X., Yang, J., Liu, S., Huang, Y., Scanlon, M. J., Nettleton, D., and Schnable, P. S. 2017. Substantial contribution of genetic variation in the expression of transcription factors to phenotypic variation revealed by eRD-GWAS. *Genome Biol.* **18**: 14.

Lu, Y., Zhang, S., Shah, T., Xie, C., Hao, Z., Li, X., Farkhari, M., Ribaut, J.-M., Cao, M., Rong, T., and Xu, Y. 2010. Joint linkage–linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. *Proc. Natl. Acad. Sci.* **107**: 19585–19590.

Mackay, I., and Powell, W. 2007. Methods for linkage disequilibrium mapping in crops. *Trends Plant Sci.* **12**: 57–63.

Maurer, A., Draba, V., Jiang, Y., Schnaithmann, F., Sharma, R., Schumann, E., Kilian, B., Reif, J. C., and Pillen, K. 2015. Modelling the genetic architecture of flowering time control in barley through nested association mapping. *BMC Genomics.* **16**: 290.

Macdonald, S. J., and Long, A. D. 2007. Joint estimates of quantitative trait locus effect and frequency using synthetic recombinant populations of Drosophila *melanogaster*. *Genetics.* **176**: 1261–1281.

McMullen, M. D., Kresovich, S., Villeda, H. S., Bradbury, P., Li, H., Sun, Q., Flint-Garcia, S., Thornsberry, J., Acharya, C., Bottoms, C., Brown, P., Browne, C., Eller, M., Guill, K., Harjes, C., Kroon, D., Lepak, N., Mitchell, S. E., Peterson, B., Pressoir, G., Romero, S., Rosas, O. M., Salvo, S., Yates, H., Hanson, M., Jones, E., Smith, S., Glaubitz, J. C., Goodman, M., Ware, D., Holland, J. B., and Buckler, E. S. 2009. Genetic properties of the maize nested association mapping population. *Science.* **325**: 737–740.

Meuwissen, T.H., and Goddard, M.E. 2001. Prediction of identity by descent probabilities from marker-haplotypes. *Genet. Sel. Evol.* **33**: 605–634.

Meuwissen, T.H., Hayes, B.J., and Goddard, M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics.* **157**: 1819–1829.

Meuwissen, T.H., Karlsen, A., Lien, S., Olsaker, I., and Goddard, M.E. 2002. Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. *Genetics.* **161**: 373–379.

Nakaya, A., and Isobe, S. N. 2012. Will genomic selection be a practical method for plant breeding? *Ann. Bot.* **110**: 1303–1316.

Nice, L. M., Brian, J. S., Gina, L. B., Eduard, D. A., Liu, C., Thomas, J. Y. K., Peter, L. M., Thomas, K. B., Richard, D. H., Kevin, P. S., and Gary, J. M. 2016. Development and genetic characterization of an advanced backcross-nested association mapping (AB-NAM) population of wild × cultivated barley. *Genetics.* **203**: 1453–1467.

Nice, L. M., Steffenson, B. J., Blake, T. K., Horsley, R. D., Smith, K. P., Muehlbauer, G. R. 2017. Mapping

agronomic traits in a wild barley advanced backcross--nested association mapping population. *Crop Sci.* **57**: 1199–1210.

Ogut, F., Bian, Y., Bradbury, P.J., and Holland, J.B. 2015. Joint-multiple family linkage analysis predicts within-family variation better than single-family analysis of the maize nested association mapping population. *Heredity.* **114**: 552–563.

Olukolu, B. A., Wang, G., Vijay ,V, Bala, P. V., Sandeep, M., Jiabing, J., Emma, G., Kevin, C., Adisu, N., Jacqueline, B., Rebecca, N., Peter, B., Dahlia, N., James, B. H., Peter, J. B., Gurmukh, J. A. 2014. Genome-wide association study of the maize hypersensitive defense response identifies genes that cluster in related pathways. *Plos Genetics* **10**: 1–16.

Pandey, M. K., Roorkiwal, M., Singh, V. K., Ramalingam, A., Kudapa, H., Thudi, M., Chitikineni, A., Rathore, A., and Varshney, R. K. 2016. Emerging genomic tools for legume breeding: current status and future prospects. *Front Plant Sci.* **7**: 455.

Peiffer, J. A., Flint-Garcia, S. A., De-Leon, N., McMullen, M. D., Kaeppler, S. M., Buckler, E. S. 2013. The genetic architecture of maize stalk strength. *PLoS ONE.* **8**: e67066.

Peiffer, J. A., Romay, M. C., Gore, M. A., Garcia, F. S. A., Zhang, Z., Millard, M. J., Gardner, C. A., McMullen, M. D., Holland, J. B., Bradbury, P. J., and Buckler, E. S. 2014. The genetic architecture of maize height. *Genetics.* **196**: 1337–1356.

Poland, J. A., Bradbury, P. J., Buckler, E. S., and Nelson, R. J. 2011. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc. Natl. Acad. Sci. USA.* **108**: 6893–6898.

Rabinowitz, D. 1997. A transmission disequilibrium test for quantitative trait loci. *Hum. Hered.* **47**: 342–350.

Rafalski, A. 2002. Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* **5**: 94–100.

Rakshit, S., Rakshit, A., and Patil, J. 2012. Multiparent intercross populations in analysis of quantitative traits. *J. Genet.* **91**: 111–117.

Rakshit, S., and Swapna, M. 2015. DNA markers in diversity analysis. In *Sorghum Molecular Breeding.* New Delhi: Springer, pp 23–46.

Ray, S., and Satya, P. 2014. Next generation sequencing technologies for next generation plant breeding. *Front. Plant Sci.* **5**: 1–4.

Rebai, A., and Goffinet, B. 1993. Power of tests for QTL detection using replicated progenies derived from a diallel cross. *Theor. Appl. Genet.* **86**: 1014–1022.

Ren, Y., Hou, W., Lan, C., Basnet, B. R., Singh, R. P., Zhu, W., Cheng, X., Cui, D., and Chen, F. 2017. QTL analysis and nested association mapping for adult plant resistance to powdery mildew in two bread wheat populations. *Front. Plant Sci.* **8**: 1212.

Rincent, R., Charcosset, A., and Moreau, L. 2017. Predicting genomic selection efficiency to optimize calibration set and to assess prediction accuracy in highly structured populations. *Theor. Appl. Genet.* **130**: 2231–2247.

Saade, S., Kutlu, B., Draba, V., Förster, K., Schumann, E., Tester, M., Pillen, K., and Maurer, A. 2017. A donor-specific QTL, exhibiting allelic variation for leaf sheath hairiness in a nested association mapping population, is located on barley chromosome 4H. *PLOS One.* **12**: e0189446.

Saade, S., Maurer, A., Shahid, M., Oakey, H., Schmöckel, S. M., Negrão, S., Pillen, K., and Tester, M. 2016. Yield-related salinity tolerance traits identified in a nested association mapping (NAM) population of wild barley. *Sci. Rep.* **6**: 32586.

Schnaithmann, F., Doris, K., and Klaus, P. 2014. A first step toward the development of a barley NAM population and its utilization to detect QTLs conferring leaf rust seedling resistance. *Theor. Appl. Genet.* **127**: 1513–1525.

Schon, C., Utz, H.F., Groh, S., Truberg, B., Openshaw, S., and Melchinger, A. E. 2004. Quantitative trait locus mapping based on resampling in a vast maize testcross experiment and its relevance to quantitative genetics for complex traits. *Genetics.* **167**: 485–498.

Sharma, R., Draicchio, F., Bull, H., Herzig, P., Maurer, A., Pillen, K. W., Thomas, T. B., and Andrew, J. F. 2018. Genome-wide association of yield traits in a nested association mapping population of barley reveals new gene diversity for future breeding. *J. Exp. Bot.* **69**: 3811–3822.

Shivakumar, M., Kumawat, G., Nataraj, V., Gireesh, C., Gupta, S., Satpute, G. K., Ratnaparkhe, M. B., and Yadav, D. P. 2019. NAM population–a novel genetic resource for soybean improvement: development and characterization for yield and attributing traits. *Plant Genet. Resour.* **17**: 545–553.

Sneller, C. H., Mather, D. E., and Crepieux, S. 2009. Analytical approaches and population types for finding and utilizing QTL in complex plant populations. *Crop Sci.* **49**: 363–380.

Soller, M., Brody, T., and Genizi, A. 1976. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theor. Appl. Genet.* **47**: 35–39.

Song, Q., Yan, L., Quigley, C., Jordan, B. D., Fickus, E., Schroeder, S., Song, B. H., Charles An, Y. Q., Hyten, D., Nelson, R., Rainey, K., Beavis, W. D., Specht, J., Diers, B., and Cregan, P. 2017. Genetic characterization of the soybean nested association mapping population. *Plant Genome* **10**: 1–14.

Stich, B. 2009. Comparison of mating designs for establishing nested association mapping populations in maize and Arabidopsis thaliana. *Genetics.* **183**: 1525–1534.

Stich, B., Utz, H. F., Piepho, H. P., Maurer, H. P., and Melchinger, A. E. 2010. Optimum allocation of resources for QTL detection using a nested association mapping strategy in maize. *Theor. Appl. Genet.* **120**: 553–561.

Svenson, K. L., Gatti, D. M., Valdar, W., Welsh, C. E., Cheng, R., Chesler, E. J., Palmer, A. A., McMillan, L., and Churchill, G. A. 2012. High-resolution genetic mapping using the mouse diversity outbred population. *Genetics.* **190**: 437–447.

Tian, F., Bradbury, P. J., Brown, P. J., Hung, H., Sun, Q., Flint-Garcia, S., Rocheford, T. R., McMullen, M. D., Holland, J. B., and Buckler, E. S. 2011. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat. Genet.* **43**: 159–162.

Varshney, R. K., Terauchi, R., and McCouch, S. R. 2014. Harvesting the promising fruits of genomics: applying

genome sequencing technologies to crop breeding. *PLOS Biol.* **12**: e1001883.

Vatter, T., Maurer, A., Kopahnke, D., Perovic, D., Ordon, F., and Pillen, K. 2017. A nested association mapping population identifies multiple small effect QTL conferring resistance against net blotch (*Pyrenophora teres f. teres) in wild barley*. *PLOS One*. **12**: e0186803.

Wu, R., Ma, C. X., and Casella, G. 2002. Joint linkage and linkage disequilibrium mapping of quantitative trait loci in natural populations. *Genetics.* **160**: 779–792.

Wu, R., and Zeng, Z. B. 2001. Joint linkage and linkage disequilibrium mapping in natural populations. *Genetics.* **157**: 899–909.

Xavier, A., Muir, W. M., and Rainey, K. M. 2016. Assessing predictive properties of genome-wide selection in soybeans. *G3.* **6**: 2611–2616.

Xu, Y., Li, P., Yang, Z., and Xu, C. 2017. Genetic mapping of quantitative trait loci in crops. *Crop J.* **5**: 175–184.

Yates, F. 1936. A new method of arranging variety trials involving a large number of varieties. *J. Agric. Sci.* **26**: 424–455.

Yu, J., and Buckler, S. E. 2006. Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* **17**: 155–160.

Yu, J., Holland, J. B., McMullen, M. D., and Buckler, E. S. 2008. Genetic design and statistical power of nested association mapping in maize. *Genetics.* **178**: 539–551.

Zeng, Z. B. 1994. Precision mapping of quantitative trait loci. *Genetics.* **136**: 1457–1468.

Zhang, N., Gibon, Y., Wallace, J. G., Lepak, N., Li, P., Dedow, L., Chen, C., So, Y.-S., Kremling, K., Bradbury, P. J., Brutnell, T., Stitt, M., and Buckler, E. S. 2015. Genome-wide association of carbon and nitrogen metabolism in the maize nested association mapping population. *Plant Physiol.* **168**: 575–583.