Nucleotide Polymorphism in *Zea*: Patterns and Influences on Crop Traits

Lan Zhou [1], Yan Bao [1] ✉, Jiaen Wang [2], Shuling Wang [1], Weixin Zhong [1], Xiangrui Sun [1]

1 College of Agriculture, Jilin Agricultural Science and Technology University, Jilin, 132101, Jilin, China

2 Meihekou City Farmers' Science and Technology Education Center, Meihekou City, 135000, Jilin, China

✉ Corresponding email: baoyan0302@126.com

Abstract Nucleotide polymorphisms, particularly single nucleotide polymorphisms (SNPs), play a crucial role in the genetic diversity and trait development of maize (*Zea mays*). This study synthesizes current research on the patterns of nucleotide polymorphism in maize and its wild relatives, and their influence on crop traits. Studies have shown that maize exhibits a high level of nucleotide diversity, with significant variations between landraces and inbred lines. Population structure and genetic subdivision significantly affect the observed patterns of nucleotide polymorphism, which is crucial for understanding the genetic basis of important agronomic traits. Advances in sequencing technology have facilitated the discovery and application of SNPs in crop genetics, enabling more precise association analyses and marker-assisted breeding. Specific polymorphisms have been linked to key traits such as root development and adaptation to environmental conditions, highlighting their potential for crop improvement. This study underscores the importance of integrating genetic diversity studies with modern genomic tools to enhance our understanding of trait development and improve maize breeding strategies.

Keywords Nucleotide polymorphism; *Zea mays*; Genetic diversity; SNPs; Crop traits

1 Introduction

Maize (*Zea mays* L.), commonly known as corn, is one of the world's most significant crops, serving as a staple food, animal feed, and industrial raw material. Originating in Central America, maize was introduced to Europe and other continents following Columbus's voyages at the end of the 15th century (Revilla et al., 2022). Its adaptability has led to a wide variety of genetic resources, making it a crucial crop in diverse agricultural systems. Over the centuries, maize has become integral to traditional food specialties in various regions, particularly in Europe, where it is predominantly used for animal feed but also for numerous traditional food products (Revilla et al., 2022).

Research on nucleotide polymorphism in plant genetics has advanced significantly, particularly with the advent of genome-wide association studies (GWAS) and quantitative trait locus (QTL) mapping. These techniques have enabled the identification of genetic loci associated with important traits in crops like maize. For instance, studies have refined numerous QTL and quantitative trait nucleotide (QTN) clusters related to ear traits in maize, which are critical for yield improvement (Dong et al., 2023). Additionally, the construction of detailed physical and genetic maps of the maize genome has provided insights into its complex evolutionary history and structural variations (Wei et al., 2007).

Exploring nucleotide polymorphism in maize involves understanding the genetic diversity and evolutionary history of this crop. Studies have shown significant phenotypic and genetic differences between ancient and modern maize, suggesting human selection and the introduction of new varieties over time. Furthermore, the introgression of genes from wild relatives, such as teosinte, into cultivated maize has been documented, highlighting the role of hybridization in maize evolution. These findings underscore the importance of nucleotide polymorphism in shaping the genetic makeup and adaptability of maize.

This study analyzes the impact of nucleotide polymorphism on crop traits involves examining how genetic variations influence phenotypic characteristics and overall crop performance. Specific genetic loci have been linked to ear traits that determine maize yield, with some regions showing pleiotropic effects on multiple traits. Additionally, the introgression of adaptive alleles from wild relatives has been shown to enhance the adaptability of maize to different environmental conditions. Understanding these genetic influences is crucial for developing breeding strategies aimed at improving crop traits and ensuring food security. The study of nucleotide polymorphism in maize provides valuable insights into the genetic factors that influence crop traits, explores the historical and current genetic diversity of maize, researchers can develop strategies to enhance its yield, adaptability, and overall performance, thereby supporting global food production and sustainability.

2 Definition and Overview of Nucleotide Polymorphism

2.1 Basic concept of nucleotide polymorphism (SNPs)

Nucleotide polymorphisms, particularly single nucleotide polymorphisms (SNPs), are the most common form of genetic variation within a species. SNPs represent a single base pair change in the DNA sequence and can occur throughout the genome. These variations can be found in both coding and non-coding regions of the genome and may or may not affect the function of genes. SNPs are highly abundant and evenly distributed across the genomes of many organisms, making them ideal markers for genetic studies (Rafalski, 2002; Duran et al., 2020; Morgil et al., 2020). The discovery and application of SNPs have been significantly advanced by next-generation sequencing technologies, which allow for the rapid and cost-effective identification of these polymorphisms (Ganal et al., 2009; Morgil et al., 2020). SNPs can be used in various applications, including genetic mapping, population structure analysis, and marker-assisted breeding, due to their high resolution and low mutation rate (Ganal et al., 2009; Duran et al., 2020).

2.2 Distribution and characteristics of nucleotide polymorphisms in plants

In plants, SNPs are distributed throughout the genome and are particularly useful for studying genetic diversity and crop improvement. The distribution of SNPs can vary between species and even within different populations of the same species. For example, in maize, SNPs are abundant and evenly distributed, making them valuable for genetic research and breeding programs (Yan et al., 2010). The identification of SNPs in plants has been facilitated by the availability of extensive genomic resources, such as expressed sequence tags (ESTs) and whole-genome sequences (Rafalski, 2002; Duran et al., 2020). These resources allow for the direct readout of SNP haplotypes, which can provide more informative analyses than individual SNPs alone. Additionally, SNPs can be used to identify regions of the genome associated with specific phenotypes, aiding in the selection of desirable traits in breeding programs (Figure 1) (Rafalski, 2002; Morgil et al., 2020).

2.3 Characteristics of nucleotide polymorphism in the maize genome

The maize genome is characterized by a high level of nucleotide polymorphism, which has been extensively studied using SNP markers. Maize exhibits a high degree of intraspecific nucleotide diversity, making it an excellent model for genetic studies (Rafalski, 2002; Yan et al., 2010). SNPs in maize are not only abundant but also evenly distributed across the genome, which facilitates their use in high-throughput genotyping platforms such as the GoldenGate assay. This assay allows for the simultaneous genotyping of thousands of SNP markers, providing a comprehensive view of genetic variation in maize populations (Yan et al., 2010).

In addition to SNPs, other types of nucleotide polymorphisms, such as insertions and deletions (indels), are also prevalent in the maize genome. These indels can serve as highly informative genetic markers and can be used for genetic mapping and diagnostics (Bhattramakki et al., 2002). The presence of large-scale structural variations, such as chromosomal inversions, further adds to the complexity of the maize genome. For instance, a megabase-scale inversion polymorphism has been identified on chromosome 1 of maize, which captures over 700 genes and shows evidence of adaptive evolution. This inversion is present in the wild ancestors of maize but is absent in domesticated varieties, highlighting the impact of structural variations on the genetic makeup of maize (Fang et al., 2012).
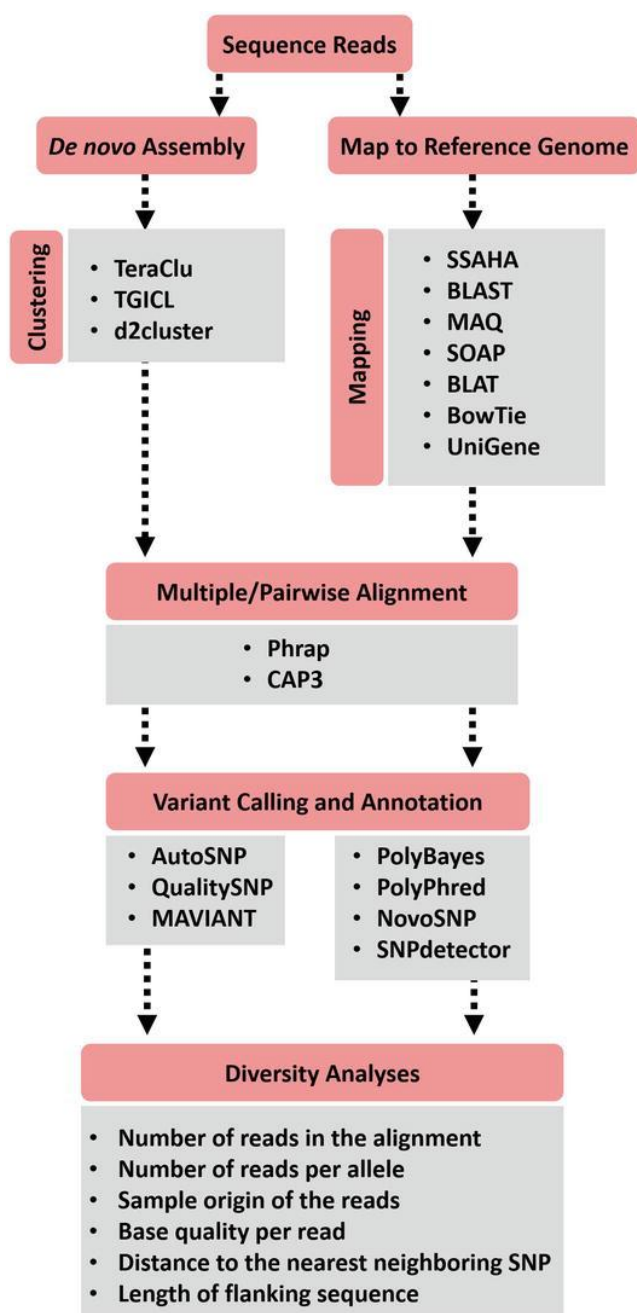
Figure 1 A common workflow schema for in silico SNP mining (Adopted from Morgil et al., 2020)

Image caption: According to the input data type, the steps and the algorithms change. If the data set is de novo assembly output, clustering step is needed. If the data set is based on a reference mapped output, mapping step is required. In order to mine possible SNPs, after clustering and mapping steps, alignment, variant calling, annotation, and diversity analysis should be applied, respectively (Adopted from Morgil et al., 2020)

The study of nucleotide polymorphisms in maize provides valuable insights into the genetic diversity and evolutionary history of this important crop. The identification and characterization of these polymorphisms are crucial for understanding the genetic basis of phenotypic traits and for the development of improved maize varieties through marker-assisted breeding (Rafalski, 2002; Yan et al., 2010; Morgil et al., 2020).

3 Mechanisms of Nucleotide Polymorphism Formation in Maize
3.1 Gene mutation and recombination
Gene mutation and recombination are fundamental mechanisms driving nucleotide polymorphism in maize. Mutations introduce new genetic variations by altering the DNA sequence, while recombination shuffles existing

genetic material to create new allele combinations. In maize, the rate of single nucleotide substitutions has been estimated using data from the domestication locus *teosinte branched1* (*tb1*). This locus, a major target of human selection during maize domestication, shows that polymorphism in the region is consistent with new mutations following fixation for a small number of haplotypes during domestication (Clark et al., 2005). The estimated nucleotide substitution rates for the *tb1* intergenic region are approximately 2.9×10^{-8} and 3.3×10^{-8} substitutions per site per year, highlighting the role of mutation in generating genetic diversity.

Recombination also plays a crucial role in shaping genetic diversity in maize. Genome-wide studies have mapped recombination events across the maize genome, revealing substantial variation in recombination frequency and distribution among different populations. These studies identified 143 recombination hot regions and showed that intragenic recombination events are associated with variation in gene expression and agronomic traits, suggesting that recombination contributes significantly to phenotypic diversity (Pan et al., 2016). Additionally, the correlation between sequence diversity and recombination rates indicates that recombination can break down linkage disequilibrium among single nucleotide polymorphisms (SNPs), further promoting genetic variation (Tenaillon et al., 2001; Tenaillon et al., 2002).

3.2 Gene flow and natural selection

Gene flow and natural selection are critical processes influencing nucleotide polymorphism in maize. Gene flow, the transfer of genetic material between populations, introduces new alleles and increases genetic diversity. In maize, gene flow from wild relatives such as teosinte has contributed to the genetic makeup of domesticated maize. This exchange of genetic material has facilitated the adaptation of maize to diverse environments and agricultural practices.

Natural selection acts on genetic variation, favoring alleles that confer a selective advantage. In maize, patterns of polymorphism along chromosome 1 show that natural selection shapes genetic diversity. For instance, regions with high gene density and frequent crossing-over exhibit lower levels of nucleotide variability, suggesting that selection on linked variation reduces genetic diversity in these areas (Flowers et al., 2012). Additionally, chromosomal inversions, which suppress recombination, have been identified in wild maize ancestors. These inversions capture locally adapted alleles and show evidence of adaptive evolution, including associations with environmental variables and phenotypic traits (Fang et al., 2012).

3.3 Influence of artificial selection on polymorphism

Artificial selection has profoundly influenced nucleotide polymorphism in maize. During domestication, humans selected for desirable traits, leading to rapid phenotypic evolution. Analysis of single-nucleotide polymorphisms in maize genes indicates that 2 to 4% of these genes experienced artificial selection, affecting approximately 1 200 genes throughout the maize genome (Wright et al., 2005). These selected genes are often clustered near quantitative trait loci that contribute to phenotypic differences between maize and its wild ancestor, teosinte.

Recent studies have also identified genomic regions that underwent positive artificial selection during maize improvement. For example, a gene-oriented haplotype comparison revealed over 1 100 genomic regions selected during the improvement of temperate and tropical maize germplasm. These regions include regulatory genes and key genes with visible mutant phenotypes, highlighting the role of artificial selection in shaping the maize genome (He et al., 2017). Furthermore, selection for specific traits, such as starch metabolism, has led to low genetic diversity in critical genes involved in these pathways, suggesting that artificial selection can significantly reduce genetic variation in targeted loci (Whitt et al., 2002).

The interplay between gene mutation, recombination, gene flow, natural selection, and artificial selection has created a complex landscape of nucleotide polymorphism in maize. Understanding these mechanisms provides valuable insights into the genetic basis of crop traits and informs strategies for future maize breeding and improvement.

4 Distribution Patterns of Nucleotide Polymorphism in Maize

4.1 Genome-wide distribution of SNPs

Single nucleotide polymorphisms (SNPs) are the most abundant form of genetic variation in the maize genome, and their distribution is widespread across the genome. Studies have shown that maize exhibits a high level of nucleotide diversity, with an average of one SNP every 104 base pairs (bp) between two randomly sampled sequences (Tenaillon et al., 2001). This level of diversity is higher than that observed in humans or Drosophila melanogaster. The distribution of SNPs is not uniform, with some regions exhibiting higher diversity than others. For instance, a study using RNA-sequencing (RNA-seq) identified 351 710 polymorphic loci distributed throughout the maize genome, covering 22 830 annotated genes (Hansey et al., 2012). This extensive distribution of SNPs across the genome facilitates their use in genome-wide association studies (GWAS) and marker-assisted selection in breeding programs.

The development of high-throughput genotyping platforms, such as the maize 55K SNP array, has further enhanced our ability to map SNPs across the genome. This array includes 55 229 SNPs evenly distributed across the genome, with a significant portion located in exonic and intronic regions (Figure 2) (Xu et al., 2017). Such comprehensive coverage ensures that SNP markers can be effectively used for various genetic analyses, including the identification of genomic regions associated with important agronomic traits.
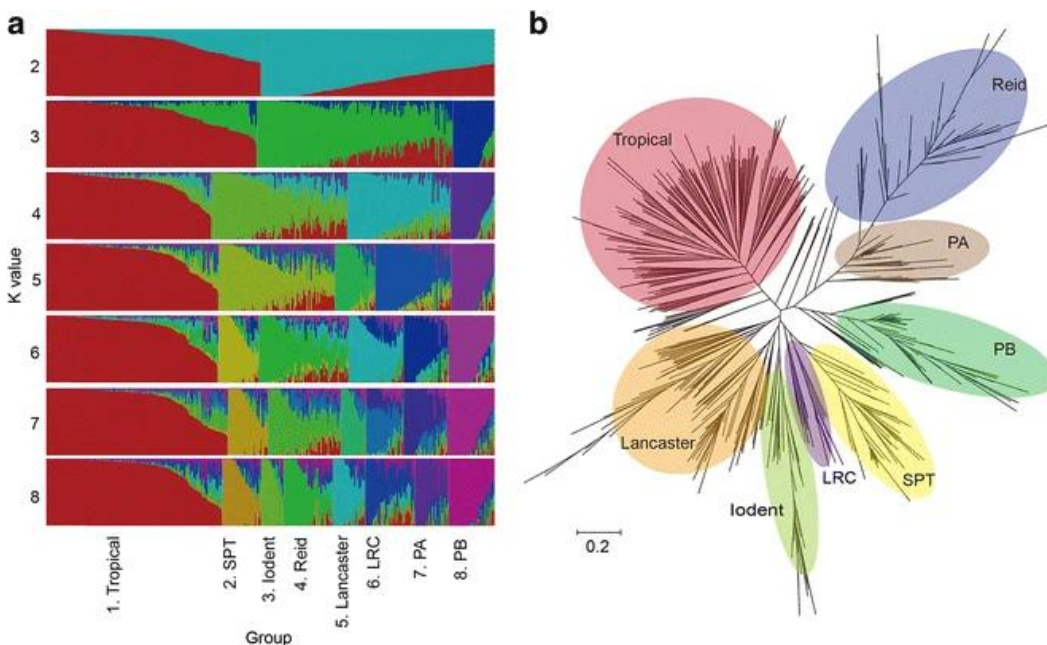

Figure 2 a: Inferred population structure for 593 diverse *maize inbred lines* with different total number of populations (*k* = 2 to 8); b: Maximum likelihood phylogenetic tree for 593 *maize inbred lines* (Adopted from Xu et al., 2017)

Image caption: a: Each *vertical line* represents one sample. Membership coefficients (Q values) were summarized from 20 replications performed by fastSTRUCTURE. b: Major population groups were determined by population structure (Adopted from Xu et al., 2017)

4.2 Differences in polymorphism between and within genes

The level of nucleotide polymorphism in maize varies significantly between and within genes. In a study examining 21 loci along chromosome 1, it was found that sequence diversity was heterogeneous among loci, with no clear pattern of diversity along the genetic map (Tenaillon et al., 2001). This heterogeneity is influenced by several factors, including recombination rates and selection pressures. For example, recombination rates were found to be correlated with sequence diversity (r=0.65), suggesting that regions with higher recombination rates tend to have higher levels of polymorphism (Tenaillon et al., 2001).

RNA-seq studies have also revealed differences in polymorphism within genes. For instance, transcript abundance analysis across 21 maize inbred lines showed that while the total number of genes expressed was relatively

consistent, the specific set of transcribed genes varied significantly among the lines (Hansey et al., 2012). This variation in gene expression, coupled with the presence of novel transcripts in some lines but not others, indicates that polymorphism within genes can contribute to phenotypic diversity and adaptation.

4.3 Diversity in maize germplasm resources

Maize germplasm resources exhibit a remarkable level of genetic diversity, which is crucial for crop improvement and adaptation to different environments. A global germplasm collection, including 527 inbred lines with tropical, subtropical, and temperate backgrounds, revealed broad phenotypic diversity and complex genetic relatedness among the lines. This diversity is reflected in the presence of numerous SNPs, with 926 SNPs having minor allele frequencies of ≥0.1 used to estimate genetic diversity and relatedness (Yang et al., 2011).

The maize pan-genome, which includes genes present in every individual (core) and genes in a subset of individuals (dispensable), further highlights the extent of genetic diversity in maize. Transcriptome sequencing of 503 maize inbred lines identified 8 681 representative transcript assemblies (RTAs), with 16.4% expressed in all lines and 82.7% expressed in subsets of the lines (Hirsch et al., 2014). This dynamic nature of the maize genome suggests that a substantial portion of genetic variation lies outside the single reference genome, emphasizing the importance of exploring diverse germplasm resources for crop improvement.

High-throughput SNP genotyping platforms, such as the GoldenGate assay, have been instrumental in characterizing genetic diversity in maize. A custom GoldenGate assay containing 1 536 SNPs was used to genotype a panel of 154 diverse inbred lines, revealing a high level of polymorphism and providing valuable insights into the genetic structure of maize populations (Yan et al., 2010). Such tools are essential for conducting association mapping studies and understanding the genetic basis of complex traits in maize.

The distribution patterns of nucleotide polymorphism in maize are influenced by various factors, including recombination rates, selection pressures, and the dynamic nature of the maize genome. The extensive genetic diversity present in maize germplasm resources provides a rich foundation for crop improvement and adaptation to diverse environments.

5 Influence of Nucleotide Polymorphism on Maize Crop Traits

5.1 Impact on stress resistance traits

Nucleotide polymorphisms play a significant role in enhancing maize's resistance to various environmental stresses. For instance, a genome-wide association study (GWAS) identified several single-nucleotide polymorphisms (SNPs) associated with alkaline stress tolerance in maize seedlings. This study revealed nine SNPs and their associated candidate genes that significantly contribute to alkaline tolerance, which is crucial for improving maize growth in salt-alkalized soils (Li et al., 2022). Additionally, another study focused on drought and aflatoxin resistance in maize hybrids in sub-tropical regions. This research identified ten quantitative trait variants for grain yield, plant height, and other agronomic traits under both irrigated and non-irrigated conditions, demonstrating the potential of genetic diversity to improve stress resistance in maize breeding programs (Farfan et al., 2015).

Moreover, a study on the genetic architecture of stalk lodging resistance-related traits in maize identified 423 significant quantitative trait nucleotides (QTNs) associated with stem diameter, stalk bending strength, and rind penetrometer resistance. These traits are crucial for improving lodging resistance, which directly impacts maize's ability to withstand environmental stresses such as strong winds and heavy rains (Zhang et al., 2018). Collectively, these findings underscore the importance of nucleotide polymorphisms in enhancing maize's resilience to various environmental stresses, thereby contributing to more stable and sustainable crop production.

5.2 Impact on yield-related traits

Nucleotide polymorphisms also significantly influence yield-related traits in maize. A study on the genetic analysis of tropical quality protein maize (QPM) germplasm revealed that both additive and non-additive genetic effects are important for the inheritance of grain yield and other agronomic traits under both stress and non-stress

conditions. This research highlighted the potential of specific QPM lines with desirable general combining ability (GCA) effects for grain yield, which can be utilized in breeding programs to enhance maize yield (Njeri et al., 2017).

Furthermore, a fine mapping study of quantitative trait loci (QTL) for plant and ear height in a maize nested-association mapping population identified 105 SNPs and 22 QTL significantly associated with these traits. The study found that variations in the promoter region of a candidate gene, *Zm00001d031938*, could decrease plant and ear height, which are closely related to lodging resistance and planting density, ultimately affecting yield (Yin et al., 2022). Another GWAS on tropical maize germplasm under terminal drought and combined heat and drought conditions identified several SNPs associated with grain yield and related traits. This study provided valuable insights into the genetic architecture of yield-related traits under stress conditions, which is essential for breeding climate-smart maize varieties (Osuman et al., 2022).

5.3 Impact on disease resistance traits

Nucleotide polymorphisms are crucial for improving disease resistance in maize. A meta-analysis and co-expression analysis identified stable QTL and candidate genes conferring resistance to Fusarium and Gibberella ear rots, two devastating diseases that reduce maize yield and grain quality. The study identified 40 meta-QTL, with 29 associated with multiple disease-related traits, and 59 candidate genes responsive to these diseases. These findings provide a foundation for genomics-assisted breeding strategies to enhance disease resistance in maize (Table 1) (Akohoue et al., 2022).

Table 1 Selected meta-QTL (MQTL) and corresponding candidate genes (CG) (Adopted from Akohoue et al., 2022)

MQTLa	Number of QTL	Disease and trait		Number of populations	PVE (%)	CI 95% (cM)	Physical distance (Mbp)	Number of CG
		FER	GER					
ZmMQTL1.2	5	KR	KR, SR	4	10.60	4.72	3.04	10
ZmMQTL1.4	5	KR	HC, KR, SR	5	14.00	5.85	7.00	146
ZmMQTL1.5	2	KR	-	2	11.50	14.80	15.55	331
ZmMQTL1.7	2	KR	DON	2	11.00	8.00	7.28	226
ZmMQTL2.1	4	SR	DON, SR	3	11.75	3.02	0.63	30
ZmMQTL2.2	2	-	KR, SR	2	13.00	9.74	7.28	201
ZmMQTL2.3	7	KR	KR, SR	5	10.00	2.65	6.18	68
ZmMQTL3.3	3	KR,SR	SR	2	10.00	3.75	3.98	77
ZmMQTL4.3	2	KR	SR	2	17.00	11.51	14.50	342
ZmMQTL4.4	5	KR	KR	2	13.40	8.89	6.75	155
ZmMQTL7.1	5	KR	SR	2	15.20	7.75	9.85	143
ZmMQTL7.3	3	FUM	SR	2	29.67	3.89	0.75	37
ZmMQTL9.2	5	KR	SR	3	10.40	8.00	15.08	304
ZmMQTL9.4	2	FUM	DON	2	13.50	11.71	5.94	202

Note: CL, confdeanceinteral; FER, Fusarium earot; GER, Giberel earot SR, silkresistance; KR: kerne resistance; DON: deoynivalenol acumulation; FUM: fumonisinacunulation; KD: kernel dry-down rate; HC, husk coverage; PVE, phenotypic variance explained; Meta-QTL name refered to *Zea mays* abbreviatedas Zm, followed by MQTL, the corresponding chromosome, and identfication number on the chromosome (Adopted from Akohoue et al., 2022)

Additionally, a study on the role of a dual-subcellular localized β-glucosidase gene, *ZmBGLU17*, demonstrated its importance in conferring resistance to both pathogens and insect pests without a yield penalty. The research showed that structural variations and a SNP in the *ZmBGLU17* gene affect its expression and function, leading to enhanced resistance against the oomycete pathogen *Pythium aphanidermatum* and the Asian corn borer. This gene's overexpression lines exhibited normal growth and yield, highlighting the potential of nucleotide polymorphisms in improving disease resistance without compromising yield (Liu et al., 2023).

Nucleotide polymorphisms significantly influence maize crop traits, including stress resistance, yield, and disease resistance. These genetic variations provide valuable resources for breeding programs aimed at developing more resilient and high-yielding maize varieties. By leveraging these polymorphisms, researchers and breeders can enhance maize's adaptability to various environmental conditions and improve overall crop performance.

6 Application of Nucleotide Polymorphism in Maize Breeding

6.1 Marker-assisted selection

Marker-assisted selection (MAS) leverages molecular markers to enhance the efficiency and accuracy of selecting desirable traits in crop breeding. In maize, single nucleotide polymorphisms (SNPs) have become a pivotal tool in MAS due to their abundance and stability. The development of SNP markers has significantly advanced the ability to identify and select for traits such as yield, disease resistance, and stress tolerance. For instance, the use of SNP and SilicoDArT markers has been shown to predict heterosis effects for yield traits in maize, thereby aiding in the selection of superior parental lines for hybrid production (Tomkowiak et al., 2019). Additionally, the application of next-generation sequencing (NGS) technologies, such as genotyping-by-sequencing (GBS), has revolutionized plant genotyping and breeding by providing high-throughput sequences that facilitate the discovery and genotyping of SNPs across large crop genomes like maize (He et al., 2014). This approach has been successfully implemented in genome-wide association studies (GWAS), genetic linkage analysis, and molecular marker discovery, making it an ultimate MAS tool (He et al., 2014).

6.2 Genomic selection techniques

Genomic selection (GS) techniques involve the use of genome-wide markers to predict the genetic value of breeding candidates. This method has gained traction in maize breeding due to its ability to enhance the accuracy of selection and reduce the breeding cycle time. Recent studies have demonstrated the efficacy of deep learning models in improving the prediction accuracy of flowering-related traits in maize. These models outperformed traditional Bayesian models, showing a 14.4% increase in prediction accuracy when employing multi-trait models compared to single-trait approaches (Mora-Poblete et al., 2023). The integration of SNP markers in GS has also been facilitated by the development of high-quality SNP markers that meet stringent criteria for polymorphism and genetic stability (Mammadov et al., 2010). These markers are essential for the accurate prediction of quantitative traits and the identification of key genomic regions associated with agronomic traits (Mora-Poblete et al., 2023). Moreover, the use of SNP arrays in GS has proven to be a robust and cost-effective tool for generating high-throughput genotype data, which is crucial for the selection of superior genotypes in large breeding populations (Weber et al., 2023).

6.3 Association analysis of SNPs with breeding targets

Association analysis, particularly genome-wide association studies (GWAS), has been instrumental in identifying SNPs linked to important breeding targets in maize. This method involves scanning the genome to find associations between SNPs and phenotypic traits, thereby uncovering the genetic basis of complex traits. For example, a study on maize kernel size traits identified 21 SNPs significantly associated with kernel length, width, and thickness, providing insights into the genetic architecture of these yield-related traits (Liu et al., 2019). Similarly, association analyses of SNPs in candidate genes have revealed polymorphisms linked to root traits in maize seedlings, which are crucial for improving nutrient uptake and plant growth under nitrogen-deficient conditions (Kumar et al., 2014). These findings highlight the potential of SNP-based association analysis in enhancing the selection of maize lines with desirable traits.

The application of nucleotide polymorphism in maize breeding through MAS, GS, and association analysis has significantly advanced the field. The development and utilization of SNP markers have enabled more precise and efficient selection of breeding targets, ultimately contributing to the improvement of maize varieties with enhanced yield, stress tolerance, and other agronomic traits (Mammadov et al., 2010; He et al., 2014; Kumar et al., 2014; Liu et al., 2019; Tomkowiak et al., 2019; Mora-Poblete et al., 2023; Weber et al., 2023).

7 Future Research Directions and Challenges

7.1 Complex relationships between nucleotide polymorphism and phenotype

Understanding the complex relationships between nucleotide polymorphisms and phenotypic traits in *Zea mays* remains a significant challenge. The genetic architecture of complex traits often involves thousands of polymorphisms, each contributing a small effect to the overall phenotype (Goddard et al., 2016). This complexity is further compounded by the presence of linkage disequilibrium (LD) and the intricate interplay between different genetic loci. For instance, the identification of specific subpopulations within a maize panel has shown that population structure and familial kinship can significantly influence marker-trait associations, necessitating sophisticated statistical models to reduce false positives (Yang et al., 2011). Moreover, chromosomal inversions, such as the 50-Mb inversion on chromosome 1 in wild maize ancestors, add another layer of complexity by maintaining locally adapted alleles and influencing phenotypic traits (Fang et al., 2012).

To unravel these complex relationships, future research should focus on integrating multi-omics data, including genomics, transcriptomics, and metabolomics, to provide a holistic view of how nucleotide polymorphisms influence phenotypes. Additionally, advanced statistical methods that can simultaneously fit all SNP effects, as advocated in genome-wide association studies (GWAS), should be further developed and refined (Goddard et al., 2016). These approaches will help in predicting future phenotypes and identifying causal mutations, thereby enhancing our understanding of the genetic architecture of complex traits.

7.2 Application of emerging technologies in polymorphism research

The advent of high-throughput sequencing technologies and advanced computational methods has revolutionized polymorphism research in crops like maize. Techniques such as next-generation sequencing (NGS) have accelerated the discovery of single nucleotide polymorphisms (SNPs) and other genetic variations, enabling more detailed genetic analyses (Morgil et al., 2020). Moreover, the use of deep learning models in genomic prediction has shown significant improvements in the accuracy of predicting complex traits, such as flowering time in maize (Mora-Poblete et al., 2023). These models outperform traditional Bayesian approaches, highlighting the potential of machine learning in enhancing genomic selection.

Emerging technologies like SNP arrays have also been instrumental in genotyping polyploid crops, despite the inherent challenges posed by their complex genomes (You et al., 2018). These arrays offer a high-throughput, cost-efficient, and automated approach to genotyping, which is crucial for large-scale genetic studies and molecular breeding. Additionally, the integration of phenomics-high-throughput phenotyping technologies-into crop research is paving the way for more precise and rapid genetic gain in breeding programs (Zhao et al., 2019). By building comprehensive phenotypic databases and developing bioinformatics tools for data analysis, researchers can better understand the genotype-phenotype relationship and identify key genetic loci associated with important agronomic traits.

7.3 Practical challenges in breeding applications

Despite the advancements in nucleotide polymorphism research, several practical challenges remain in applying these findings to breeding programs. One major challenge is the accurate prediction of phenotypes based on genotypic data, especially for complex traits influenced by multiple genetic and environmental factors (Goddard et al., 2016). The high level of genetic diversity in crops like maize further complicates this task, as it requires extensive genotyping and phenotyping efforts to capture the full spectrum of genetic variation (Rafalski, 2002; Huang and Hong, 2024).

Another challenge is the development and validation of SNP markers for use in marker-assisted selection (MAS). While SNP arrays and other genotyping tools have made significant strides, the identification of reliable markers that can be consistently associated with desired traits across different environments and genetic backgrounds remains difficult (You et al., 2018). This is particularly true for polyploid crops, where the complexity of the genome can hinder SNP discovery and validation (Clevenger et al., 2015).

Moreover, the practical implementation of genomic selection in breeding programs requires substantial investment in infrastructure, including high-throughput sequencing and phenotyping facilities, as well as computational resources for data analysis (Zhao et al., 2019). Training breeders and researchers in the use of these advanced technologies and statistical methods is also essential to fully realize their potential in crop improvement.

While significant progress has been made in understanding nucleotide polymorphisms and their influence on crop traits, future research must address the complex relationships between genotype and phenotype, leverage emerging technologies for more accurate genetic analyses, and overcome practical challenges in breeding applications. By doing so, we can enhance the efficiency and effectiveness of breeding programs, ultimately leading to the development of superior crop varieties.

8 Concluding Remarks

The study of nucleotide polymorphism in *Zea mays* has revealed significant insights into the genetic diversity and evolutionary mechanisms that shape this important crop. Chromosomal inversions, such as the ~50-Mb region on chromosome 1 identified in wild *Zea mays* subspecies, play a crucial role in local adaptation by suppressing recombination and maintaining locally adapted alleles. Association mapping in a diverse maize panel has demonstrated the complex genetic architecture underlying phenotypic traits, highlighting the importance of accounting for population structure to reduce false positives in marker-trait associations. Additionally, population genetic studies in teosinte, the wild ancestor of maize, have shown that population structure significantly influences patterns of nucleotide polymorphism, which is essential for understanding the genetic basis of important traits.

Forward genetics approaches integrating genome-wide association studies (GWAS) and expression quantitative trait locus (eQTL) mapping have identified candidate genes involved in leaf development, emphasizing the role of specific functional categories such as vacuolar proton pumps and cell wall effectors in trait determination. The study of crop-wild introgression has revealed widespread gene flow between maize and its wild relative, teosinte, with evidence of adaptive introgression contributing to maize's adaptation to highland environments. Furthermore, research on DNA methylation variation in maize populations has shown that both genetic and epigenetic factors influence methylation patterns, which can affect gene expression and contribute to phenotypic diversity.

Nucleotide polymorphism is a cornerstone of genetic diversity, which is vital for the genetic improvement of maize. The identification of chromosomal inversions and their role in local adaptation provides valuable information for breeding programs aimed at developing maize varieties suited to specific environmental conditions. The use of association mapping panels that capture the global diversity of maize allows for the identification of genetic variants associated with complex traits, facilitating the development of improved maize varieties with desirable agronomic characteristics. Understanding the population structure and its impact on nucleotide polymorphism is crucial for accurately identifying loci under selection and for the effective use of genetic resources in breeding.

The integration of GWAS and eQTL mapping in forward genetics approaches enables the identification of key genes and regulatory networks involved in important traits such as leaf development, which can be targeted in precision breeding efforts to enhance maize productivity. The study of introgression between maize and its wild relatives provides insights into the incorporation of adaptive alleles, which can be harnessed to improve stress tolerance and adaptability in maize. Additionally, the exploration of DNA methylation variation and its genetic and epigenetic determinants offers new avenues for manipulating gene expression and trait development, further expanding the toolkit available for maize genetic improvement.

The comprehensive study of nucleotide polymorphism in *Zea mays* not only enhances our understanding of the genetic and evolutionary processes underlying maize diversity but also provides practical insights and tools for the genetic improvement of this vital crop. By leveraging the genetic and epigenetic variation present in maize and its wild relatives, breeders can develop more resilient and productive maize varieties to meet the growing demands of global agriculture.

Acknowledgments

Funding

Conflict of Interest Disclosure

The authors affirm that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Akohoue F., and Miedaner T., 2022, Meta-analysis and co-expression analysis revealed stable QTL and candidate genes conferring resistances to Fusarium and Gibberella ear rots while reducing mycotoxin contamination in maize, Frontiers in Plant Science, 13: 1050891.
https://doi.org/10.3389/fpls.2022.1050891

Bhattramakki D., Dolan M., Hanafey M., Wineland R., Vaske D., Register J., Tingey S., and Rafalski A., 2002, Insertion-deletion polymorphisms in 3' regions of maize genes occur frequently and can be used as highly informative genetic markers, Plant Molecular Biology, 48: 539-547.
https://doi.org/10.1023/A:1014841612043

Clark R., Tavaré S., and Doebley J., 2005, Estimating a nucleotide substitution rate for maize from polymorphism at a major domestication locus, Molecular Biology and Evolution, 22(11): 2304-2312.
https://doi.org/10.1093/molbev/msi228

Clevenger J., Chavarro C., Pearl S., Ozias-Akins P., and Jackson S., 2015, Single nucleotide polymorphism identification in polyploids: a review, example, and recommendations, Molecular Plant, 8(6): 831-846.
https://doi.org/10.1016/j.molp.2015.02.002

Dong Z., Wang Y., Bao J., Li Y., Yin Z., Long Y., and Wan X., 2023, The genetic structures and molecular mechanisms underlying ear traits in maize (*Zea mays* L.), Cells, 12(14): 1900.
https://doi.org/10.3390/cells12141900

Duran C., Appleby N., Clark T., Wood D., Imelfort M., Batley J., and Edwards D., 2008, AutoSNPdb: an annotated single nucleotide polymorphism database for crop plants, Nucleic Acids Research, 37: D951-D953.
https://doi.org/10.1093/nar/gkn650

Fang Z., Pyhäjärvi T., Weber A., Weber A., Dawe R., Glaubitz J., González J., Ross-Ibarra C., Doebley J., Morrell P., and Ross-Ibarra J., 2012, Megabase-scale inversion polymorphism in the wild ancestor of maize, Genetics, 191: 883-894.
https://doi.org/10.1534/genetics.112.138578

Farfan I., Fuente G., Murray S., Isakeit T., Huang P., Warburton M., Williams P., Windham G., and Kolomiets M., 2015, Genome wide association study for drought, aflatoxin resistance, and important agronomic traits of maize hybrids in the sub-tropics, PLoS One, 10(2): e0117737.
https://doi.org/10.1371/journal.pone.0117737

Flowers J., Molina J., Rubinstein S., Huang P., Schaal B., and Purugganan M., 2012, Natural selection in gene-dense regions shapes the genomic pattern of polymorphism in wild and domesticated rice, Molecular Biology and Evolution, 29(2): 675-687.
https://doi.org/10.1093/molbev/msr225

Ganal M., Altmann T., and Röder M., 2009, SNP identification in crop plants, Current Opinion in Plant Biology, 12(2): 211-217.
https://doi.org/10.1016/j.pbi.2008.12.009

Goddard M., Kemper K., Macleod I., Chamberlain A., and Hayes B., 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): 20160569.
https://doi.org/10.1098/rspb.2016.0569

Hansey C., Vaillancourt B., Sekhon R., León N., Kaeppler S., Buell C., and Moustafa A., 2012, Maize (*Zea mays* L.) genome diversity as revealed by RNA-sequencing, PLoS One, 7(3): e33071.
https://doi.org/10.1371/journal.pone.0033071

He C., Fu J., Zhang J., Li Y., Zheng J., Zhang H., Yang X., Wang J., and Wang G., 2017, A gene-oriented haplotype comparison reveals recently selected genomic regions in temperate and tropical maize germplasm, PLoS One, 12(1): e0169806.
https://doi.org/10.1371/journal.pone.0169806

He J., Zhao X., Laroche A., Lu Z., Liu H., and Li Z., 2014, Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding, Frontiers in Plant Science, 5: 484.
https://doi.org/10.3389/fpls.2014.00484

Hirsch C., Foerster J., Johnson J., Sekhon R., Muttoni G., Vaillancourt B., Peñagaricano F., Lindquist E., Pedraza M., Barry K., León N., Kaeppler S., and Buell C., 2014, Insights into the maize pan-genome and pan-transcriptome, Plant Cell, 26: 121-135.
https://doi.org/10.1105/tpc.113.119982

Huang W.Z., and Hong Z.M., 2024, Marker-assisted selection in cassava: from theory to practice, Plant Gene and Trait, 15(1): 33-43.
https://doi.org/10.5376/pgt.2024.15.0005

Jiang C., 2024, Genetic mechanisms of crop disease resistance: new advances in GWAS, Plant Gene and Trait, 15(1): 15-22.
https://doi.org/10.5376/pgt.2024.15.0003

Kumar B., Abdel-Ghani A., Pace J., Reyes-Matamoros J., Hochholdinger F., and Lübberstedt T., 2014, Association analysis of single nucleotide polymorphisms in candidate genes with root traits in maize (*Zea mays* L.) seedlings, Plant Science, 224: 9-19.
https://doi.org/10.1016/j.plantsci.2014.03.019

Li C., Jia Y., Zhou R., Liu L., Cao M., Zhou Y., Wang Z., and Di H., 2022, GWAS and RNA-seq analysis uncover candidate genes associated with alkaline stress tolerance in maize (*Zea mays* L.) seedlings, Frontiers in Plant Science, 13: 963874.
https://doi.org/10.3389/fpls.2022.963874

Liu C., He S., Chen J., Wang M., Li Z., Wei L., Chen Y., Du M., Liu D., Li C., An C., Bhadauria V., Lai J., and Zhu W., 2023, A dual-subcellular localized β-glucosidase confers pathogen and insect resistance without a yield penalty in maize, Plant Biotechnology Journal, 22(4): 1017-1032.
https://doi.org/10.1111/pbi.14242

Liu M., Tan X., Yang Y., Liu P., Zhang X., Zhang Y., Wang L., Hu Y., Ma L., Li Z., Zhang Y., Zou C., Lin H., Gao S., Lee M., Lübberstedt T., Pan G., and Shen Y., 2019, Analysis of the genetic architecture of maize kernel size traits by combined linkage and association mapping, Plant Biotechnology Journal, 18: 207-221.
https://doi.org/10.1111/pbi.13188

Mammadov J., Chen W., Ren R., Pai R., Marchione W., Yalçin F., Witsenboer H., Greene T., Thompson S., and Kumpatla S., 2010, Development of highly polymorphic SNP markers from the complexity reduced portion of maize [*Zea mays* L.] genome for use in marker-assisted breeding, Theoretical and Applied Genetics, 121: 577-588.
https://doi.org/10.1007/s00122-010-1331-8

Mora-Poblete F., Maldonado C., Henrique L., Uhdre R., Scapim C., and Mangolim C., 2023, Multi-trait and multi-environment genomic prediction for flowering traits in maize: a deep learning approach, Frontiers in Plant Science, 14: 1153040.
https://doi.org/10.3389/fpls.2023.1153040

Morgil H., Gerçek Y., and Tulum I., 2020, Single nucleotide polymorphisms (SNPs) in plant genetics and breeding, In: Çalışkan M., Erol O., and Öz G.C. (eds.), The recent topics in genetic polymorphisms, IntechOpen, London, UK, pp.148.
https://doi.org/10.5772/intechopen.91886

Njeri S., Makumbi D., Warburton M., Diallo A., Jumbo M., and Chemining'wa G., 2017, Genetic analysis of tropical quality protein maize (*Zea mays* L.) germplasm, Euphytica, 213: 261.
https://doi.org/10.1007/s10681-017-2048-4

Osuman A., Badu-Apraku B., Karikari B., Ifie B., Tongoona P., and Danquah E., 2022, Genome-wide association study reveals genetic architecture and candidate genes for yield and related traits under terminal drought, combined heat and drought in tropical maize germplasm, Genes, 13(2): 349.
https://doi.org/10.3390/genes13020349

Pan Q., Li L., Yang X., Tong H., Xu S., Li Z., Li W., Muehlbauer G., Li J., and Yan J., 2016, Genome-wide recombination dynamics are associated with phenotypic variation in maize, The New Phytologist, 210(3): 1083-1094.
https://doi.org/10.1111/nph.13810

Rafalski A., 2002, Applications of single nucleotide polymorphisms in crop genetics, Current Opinion in Plant Biology, 5(2): 94-100.
https://doi.org/10.1016/S1369-5266(02)00240-6

Revilla P., Alves M., Andelković V., Balconi C., Dinis I., Mendes-Moreira P., Redaelli R., Galarreta J., Patto M., Žilić S., and Malvar R., 2022, Traditional foods from maize (*Zea mays* L.) in Europe, Frontiers in Nutrition, 8: 683399.
https://doi.org/10.3389/fnut.2021.683399

Tenaillon M., Sawkins M., Long A., Gaut R., Doebley J., and Gaut B., 2001, Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.), Proceedings of the National Academy of Sciences of the United States of America, 98: 9161-9166.
https://doi.org/10.1073/pnas.151244298

Tomkowiak A., Bocianowski J., Radzikowska D., and Kowalczewski P., 2019, Selection of parental material to maximize heterosis using SNP and SilicoDarT markers in maize, Plants, 8(9): 349.

https://doi.org/10.3390/plants8090349

Weber S., Chawla H., Ehrig L., Hickey L., Frisch M., and Snowdon R., 2023, Accurate prediction of quantitative traits with failed SNP calls in canola and maize, Frontiers in Plant Science, 14: 1221750.

https://doi.org/10.3389/fpls.2023.1221750

Wei F., Coe E., Nelson W., Bharti A., Engler F., Butler E., Kim H., Goicoechea J., Chen M., Lee S., Fuks G., Sanchez-Villeda H., Schroeder S., Fang Z., McMullen M., Davis G., Bowers J., Paterson A., Schaeffer M., Gardiner J., Cone K., Messing J., Soderlund C., and Wing R., 2007, Physical and genetic structure of the maize genome reflects its complex evolutionary history, PLoS Genetics, 3(7): e123.

https://doi.org/10.1371/journal.pgen.0030123

Whitt S., Wilson L., Tenaillon M., Gaut B., and Buckler E., 2002, Genetic diversity and selection in the maize starch pathway, Proceedings of the National Academy of Sciences of the United States of America, 99: 12959-12962.

https://doi.org/10.1073/pnas.202476999

Wright S., Bi I., Schroeder S., Yamasaki M., Doebley J., McMullen M., and Gaut B., 2005, The effects of artificial selection on the maize genome, Science, 308: 1310-1314.

https://doi.org/10.1126/science.1107891

Xu C., Ren Y., Jian Y., Guo Z., Zhang Y., Xie C., Fu J., Wang H., Wang G., Xu Y., Li P., and Zou C., 2017, Development of a maize 55 K SNP array with improved genome coverage for molecular breeding, Molecular Breeding, 37: 20.

https://doi.org/10.1007/s11032-017-0622-z

Yan J., Yang X., Shah T., Sanchez-Villeda H., Li J., Warburton M., Zhou Y., Crouch J., and Xu Y., 2010, High-throughput SNP genotyping with the GoldenGate assay in maize, Molecular Breeding, 25: 441-451.

https://doi.org/10.1007/s11032-009-9343-2

Yang X., Gao S., Xu S., Zhang Z., Prasanna B., Li L., Li J., and Yan J., 2011, Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize, Molecular Breeding, 28: 511-526.

https://doi.org/10.1007/s11032-010-9500-7

Yin X., Bi Y., Jiang F., Guo R., Zhang Y., Fan J., Kang M., and Fan X., 2022, Fine mapping of candidate quantitative trait loci for plant and ear height in a maize nested-association mapping population, Frontiers in Plant Science, 13: 963985.

https://doi.org/10.3389/fpls.2022.963985

You Q., Yang X., Peng Z., Xu L., and Wang J., 2018, Development and applications of a high throughput genotyping tool for polyploid crops: single nucleotide polymorphism (SNP) array, Frontiers in Plant Science, 9: 104.

https://doi.org/10.3389/fpls.2018.00104

Zhang Y., Liu P., Zhang X., Zheng Q., Chen M., Ge F., Li Z., Sun W., Guan Z., Liang T., Zheng Y., Tan X., Zou C., Peng H., Pan G., and Shen Y., 2018, Multi-locus genome-wide association study reveals the genetic architecture of stalk lodging resistance-related traits in maize, Frontiers in Plant Science, 9: 611.

https://doi.org/10.3389/fpls.2018.00611

Zhao C., Zhang Y., Du J., Guo X., Wen W., Gu S., Wang J., and Fan J., 2019, Crop phenomics: current status and perspectives, Frontiers in Plant Science, 10: 714.

https://doi.org/10.3389/fpls.2019.00714