UNIVERSIDADE ESTADUAL PAULISTA - UNESP

CÂMPUS DE JABOTICABAL

GENOMIC PREDICTION IN SWEET CORN (*Zea mays saccharata*)

Guilherme Repeza Marquez

MSc. Agriculture Engineer

2023

SÃO PAULO STATE UNIVERSITY - UNESP

JABOTICABAL CAMPUS

GENOMIC PREDICTION IN SWEET CORN (*Zea mays saccharata*)

Student: Guilherme Repeza Marquez

Advisor: Prof. PhD Gustavo Vitti Moro

Thesis presented to the São Paulo State University – UNESP, Jaboticabal campus, as part of the requirements for obtaining the title of Doctor of Philosophy in Agronomy (Genetics and Plant Breeding).

2023

M357g

Marquez, Guilherme Repeza
Genomic prediction in sweet corn (Zea mays saccharata) / Guilherme Repeza Marquez. -- Jaboticabal, 2023
57 p. : tabs., fotos

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal
Orientador: Gustavo Vitti Moro

1. Sweet corn. 2. Genomics. 3. Genetic Markers. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: GENOMIC PREDICTION IN SWEET CORN (*Zea mays saccharata*)

AUTOR: GUILHERME REPEZA MARQUEZ
ORIENTADOR: GUSTAVO VITTI MÔRO

Aprovado como parte das exigências para obtenção do Título de Doutor em Agronomia (Genética e Melhoramento de Plantas), pela Comissão Examinadora:

Prof. Dr. GUSTAVO VITTI MÔRO (Participaçao Virtual)
Departamento de Ciencias da Producao Agricola / FCAV UNESP Jaboticabal

gov.br Documento assinado digitalmente
GUSTAVO VITTI MORO
Data: 27/10/2023 10:06:18-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. GUSTAVO HUGO FERREIRA DE OLIVEIRA (Participaçao Virtual)
Universidade Federal de Sergipe (UFS) / Nossa Senhora da Glória/SE

gov.br Documento assinado digitalmente
GUSTAVO HUGO FERREIRA DE OLIVEIRA
Data: 25/10/2023 07:34:42-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. RAFAEL RESENDE FINZI (Participaçao Virtual)
Departamento de Ciências Agrárias / Universidade Federal de Uberlândia (UFU) - U

gov.br Documento assinado digitalmente
RAFAEL RESENDE FINZI
Data: 25/10/2023 17:53:58-0300
Verifique em https://validar.iti.gov.br

Dr. RAMON MACEDO RANGEL (Participaçao Virtual)
Departamento de Pesquisa e Desenvolvimento / Syngenta Seeds LLC. - Uberlândia

gov.br Documento assinado digitalmente
RAMON MACEDO RANGEL
Data: 26/10/2023 00:59:13-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. RINALDO CESAR DE PAULA (Participaçao Virtual)
Departamento de Ciencias da Producao Agricola / FCAV UNESP Jaboticabal

gov.br Documento assinado digitalmente
RINALDO CESAR DE PAULA
Data: 26/10/2023 14:55:08-0300
Verifique em https://validar.iti.gov.br

Jaboticabal, 24 de outubro de 2023

CURRICULUM DATA OF THE AUTHOR

GUILHERME REPEZA MARQUEZ – Born in Uberlândia, Minas Gerais state, on September 3rd of 1992. Joined the Agronomy course at the Federal University of Uberlândia - UFU in 2010. Worked as an intern for four years at the Department of Olericulture at UFU, assisting in laboratory activities and field research. During this period, a scientific initiation project, with potato crop, focusing on fertilization and nutritional efficiency, was performed and also a collaboration in the class named Fertilization and Mineral Nutrition of Plants as a monitor for new students. During graduation, performed internships at Bayer Vegetable Seeds (Nunhems) in 2013 (Watermelon Breeding Program) and in 2015 (Tomato Breeding Program) and was a visiting student at Texas Wesleyan University (Texas, US, 2014), where he graduated with excellence in English. In 2015, graduated and got the title of Agricultural Engineer and in 2016 started a master's degree in plant production, focusing on genetics and vegetables breeding at UFU, developing tomato genotypes tolerant to the spider mite and the whitefly throughout the introgression of wild alleles. In January 2017, joined Syngenta Seeds LLC. as a Product Development Technical Assistant and became a Vegetable Researcher in May 2017, when he also obtained his master's degree at UFU. In 2019 joined the Sweet Corn Breeding team, where played many roles as Assistant Breeder and Jr. Breeder. In 2020, joined the doctoral course in Agronomy (Genetics and Plant Breeding), at the São Paulo State University - UNESP - Jaboticabal campus as a Doctor of Philosophy candidate.

“It is not worth living a soft and easy life if it damages the fiber of the brain, heart and muscles. We must dare to be great, and we must realize that greatness is the fruit of work, sacrifice and great courage...for us it is a life of action, of arduous fulfillment of duty. Let us live fighting and choosing the risk of wearing ourselves out before we get rusty..."

Theodore Roosevelt

“I know perfectly how precious time is. Enjoy the moment”.

Stephen Hawking

To my parents, Rosana and Robson, to
my second mother, Geni, to my
grandfather Roberto and to my fiancé,
Gabriela

I DEDICATE

ACKNOWLEDGEMENT

I first thank God for the gift of life and for the blessings destined for me, as a life full of happiness, health and spectacular people around. I am so blessed.

I would like to express my gratitude to my parents for their affection, care and commitment to always providing me with the best tools to face the challenges that life would present. They are responsible for this achievement and never measured efforts to guarantee a good education, health and happiness. It is with great affection and pride that I thank you.

To my dear family, I thank you for the care, guidance and affection that you have always shown me, both in difficult times and in successes. I also thank you for always showing me the best path to follow and for the unconditional support in the decisions I made.

To my fiancée, I thank you for the companionship, care and love you always showed, which were the foundation of our daily lives. I also thank you for your patience and understanding at times when greater concentration, work and studies were needed.

To my friends for their affection, support and care. I want to extend this gratitude to everyone who fits in the definition of a friend and who, without a doubt, were essential for this achievement.

To Syngenta LLC. and my managers, Ramon Rangel, Andrea Cardinal, Ryan Walker, Shichen Zhang-Biehn, Bryant Long and all colleagues that made this dream possible. Meetings, evaluations, daily activities were changed, modified and even missed so I could attend the classes, work on my PhD activities and develop myself. My eternal gratitude to all of them.

I would like to express my gratitude to Professor Dr. Gustavo Moro for his exceptional guidance, patience and for allowing me to fulfill one of my dreams, even in the face of my inability to be present or contribute with other colleagues on a daily basis. I will be eternally grateful to the professor and to UNESP for this achievement.

To the colleagues of NEGEMM, who were always willing to help and support me in my studies, especially during a pandemic situation where we were not physically able to discuss and interact but always found a way to help and support each other.

Thank you very much!

SUMMARY

Page

GENOMIC PREDICTION IN SWEET CORN (*Zea mays saccharata*)

ABSTRACT – Corn is a key crop all over the world and has been exploited by man for different purposes, including human consumption, where cultivars with distinct composition of the endosperm have been gained prominence. Among them, the sweet specie, *Zea mays sacharatta*, whose singularity is caused by the inhibition of the total conversion of glucose into starch, during the grain maturation process, stands out for having a better palatability. Due to these particularities, sweet corn consumption has been increasing over the years and, in 2023, its seeds market was valued in US$ 820.3 million. Even though highly consumed worldwide, sweet corn consumption is still recent in Brazil and predominates in the form of canned grains, being produced by few growers and with little information available about its market, forms of production and benefits. In order to meet the agro-industrial needs, sweet corn varieties must account for many traits, as yield, industrial behavior, tolerance to the main pathogens and organoleptic characteristics, which make its phenotyping a costly practice and requires complementary strategies to be feasible. In this sense, the use of predictive models, based on the molecular information, presents as an excellent alternative, as they combined genotypic and phenotypic data to build a statistical model to estimate the effects of all markers and consequently the genetic value of individuals, shortening the breeding cycle and phenotyping processes. However, the accuracy of the prediction is highly influenced by characteristics such as data quality and marker density, which impact on project costs. On this way, several studies regarding the pre-selection of markers have been carried out to verify minimum densities without the significant loss in prediction accuracies. They can be done based on blocks of haplotypes or based on marker's selections evenly spaced across the genome and were already implemented on other crops and animal breeding, but never in sweet corn genotypes. In this study, the objectives were to perform a simple literature review about sweet corn and its benefits, as well as assess and validate different strategies to pre-select markers for genomic selection and to find the minimum density in a prediction of yield traits in sweet corn. Initially, the prediction was performed with a high-density chip and then, using a pre-selection strategy of clustering markers into haplotype blocks. Furthermore, a third strategy was tested, where markers were selected evenly across the genome. In general, all traits showed a significant reduction in accuracy as the number of markers decreased. However, the relationship between marker's increment and accuracy did not remain constant and reached a plateau after a certain point. Applying marker pre-selection can be a good option for a cost-efficient implementation of genomic selection in sweet corn for yield traits, as they can be predicted with a significant accuracy using a panel of ~8k quality markers that are evenly across the genome. Furthermore, using one marker per haplotype block appears to be a better cost-effective strategy for carrying out genomic selection in sweet corn, for yield traits.

Key-words: Sweet Corn; genomic selection; marker density; accuracy.

PREDIÇÃO GENÔMICA EM MILHO DOCE (Zea mays saccharata)

RESUMO – O milho é uma cultura-chave globalmente e tem sido explorado para diversos fins, inclusive para o consumo humano, onde cultivares com composição distinta do endosperma tem se destacado. Dentre elas, a espécie doce, *Zea mays sacharatta*, cuja singularidade é causada pela inibição da conversão total da glicose em amido, durante a maturação do grão, destaca-se por possuir melhor palatabilidade. Devido a essas particularidades, o consumo de milho doce vem aumentando ao longo dos anos e, em 2023, seu mercado de sementes foi avaliado em US$ 820,3 milhões. Apesar de muito consumido, o milho doce ainda é recente no Brasil e predomina na forma enlatada, sendo produzido por poucos agricultores e com poucas informações sobre seu mercado, produção e benefícios. Para atender às necessidades agroindustriais, genótipos de milho doce devem possuir produtividade e comportamento industrial competitivos, tolerância aos principais patógenos e boa palatabilidade, tornando sua fenotipagem onerosa e requerendo estratégias complementares. Nesse sentido, a utilização de modelos preditivos, apresenta-se como uma boa alternativa ao combinar dados genotípicos e fenotípicos para construir um modelo estatístico que estima os efeitos dos marcadores e, consequentemente, o valor genético dos indivíduos, encurtando o ciclo reprodutivo e os processos de fenotipagem. A precisão do modelo é altamente influenciada por fatores como qualidade dos dados e densidade de marcadores, que impactam nos custos do projeto. Dessa forma, vários estudos de pré-seleção de marcadores têm sido realizados para verificar densidades mínimas sem perdas significativas nas acurácias de predição. Eles podem ser realizados com base em blocos de haplótipos ou com base em seleções de marcadores espaçados uniformemente ao longo do genoma e, embora tenham sido implementados em outras culturas, nunca se testou em genótipos de milho doce. Neste estudo, objetivou-se realizar uma revisão de literatura do milho doce no Brasil, bem como avaliar e validar diferentes estratégias de pré-seleção de marcadores e encontrar a densidade mínima para se predizer caracteres de produtividade em milho doce. Inicialmente, a predição foi realizada com um chip de alta densidade e, em seguida, utilizando-se uma estratégia de agrupamento de marcadores em blocos de haplótipos. Além disso, uma terceira estratégia foi testada, onde os marcadores foram selecionados uniformemente em todo o genoma. Em geral, a acurácia de todas as características reduziu-se à medida que o número de marcadores diminuiu. No entanto, essa relação não permaneceu constante e atingiu um platô após um certo ponto. A aplicação da pré-seleção de marcadores pode ser uma boa opção para uma implementação econômica da seleção genômica em milho doce para caracteres de produção, por apresentarem uma acuracidade significativa usando-se um painel de ~ 8k marcadores. Além disso, usar um marcador por bloco de haplótipo parece ser uma estratégia de melhor custo-benefício para realizar a seleção genômica em milho doce, para características de rendimento.

Palavras-chave: Milho doce; seleção genômica; densidade de marcadores; acurácia.

CHAPTER 1 – OVERALL CONSIDERATIONS

1.1 INTRODUCTION

Corn is a key crop all over the world, due to its social, economic and environmental aspect. Having a broad genetic base, it has been exploited by man for different purposes and, in this sense, the cultivation of special corns, characterized by a distinct composition of the endosperm, has been gained prominence for being destinated for human consumption (Tracy 2001).

Among special corns, the sweet specie *Zea mays sacharatta's* distinction is caused by the action of one or more genes in recessive homozygosis (generally, sugary-su1 and/or shrunken-sh2), inhibiting the total conversion of glucose into starch during the grain maturation process and resulting in better palatability (Dodson-Swenson and Tracy, 2015). Due to these particularities, sweet corn consumption has been increasing, and in 2023 the sweet corn seeds market (global) was valued in US$ 820.3 million, with a perspective of achieve US$ 1.19 billion of market value until the end of the year (Global Market Insights, 2023).

Regarding Brazilian production, about 40,000 hectares were destined, in 2016, for sweet corn cultivation, most of which allocated for industrial processing (EMBRAPA, 2016) and under controlled irrigation conditions (pivots), emphasizing the growing of the market and the importance Brazilian farmers and customers are seeing on this crop. Furthermore, more than 135 genotypes, between inbreeds and hybrids, can be found registered on the official national portal (MAPA, 2023), by public and private sectors, also showing a greater interest of private and public companies in the sector.

In order to meet the agro-industrial needs, sweet corn varieties must present acceptable organoleptic characteristics as good flavor, pericarp tenderness and a bright color in addition to a high productivity, a package of disease resistance and an acceptable recovery rate (proportion of the green yield that will fill the cans) (Barbieri, 2010). By accounting for several traits, under different genetic effects, sweet corn phenotyping is a costly practice for breeding, which make complementary strategies necessary. In this sense, the use of predictive models, based on the molecular

information of the genotypes, presents itself as an excellent alternative for breeding programs.

In genomic prediction, genotypic and phenotypic data are combined to build a model to estimate marker's effects and, also, the genomic estimate of genetic value (GEGV) of individuals, based on molecular data, which shortens the breeding cycle. The prediction accuracy of the model is highly influenced by the proximity between training and validation sets and their respective characteristics (Habier et al., 2007; Solberg et al., 2008; Hayes et al., 2009; Zhong et al., 2009). It is also impacted by the size of the training set, quantity and quality of phenotypic data and marker density (Hayes et al., 2009), which are directly related to project costs.

Marker density plays a role in genomic selection (GS), affecting project costs and the prediction accuracy, which by the way, plateaus when the marker density reaches a certain threshold (Heffner et al., 2011; de Los Campos et al., 2013). Several studies have been carried out to find minimum densities of markers that would, in parallel, bring significant accuracies and cost saving (Habier et al., 2009; Wellmann et al., 2013; Ma et al., 2016; Werner et al., 2018; Sousa et al., 2019; Lee et al., 2021), but none in the sweet corn crop.

The previous choice of markers to be used in a GS can be done based on blocks of haplotypes (Cuyabano et al., 2015; Ma et al., 2016), which are specific combinations of alleles observed in a population (Gabriel et al., 2002). As they are clusters of alleles, haplotypes are expected to be in greater linkage disequilibrium with the quantitative trait loci (QTL) of interest compared to individual single nucleotide polymorphisms (SNPs) and, consequently, to be better predictors in GS (Zondervan and Cardon, 2004). Besides that, another strategy of pre-selection is the choice of markers evenly spaced across the genome, where the density to be used and the consequent expected coverage along the genome are predetermined (Ma et al., 2016). Although simple, these methodologies have never been used in the sweet corn crop.

Based on the positive results found in animal breeding, in other crops and the costly phenotyping that sweet corn presents, the pre-selection of markers and the creation of low-density panels for carrying out GS presents as an excellent alternative to reduce computational time and costs in sweet corn breeding programs.

1.2 LITERATURE REVIEW

1.2.1 History and Origen of Sweet Corn (*Zea mays saccharata*)

Sweet corn (*Zea mays saccharata*) is a naturally variety of maize that is also known as sugar corn and has been reported in history since pre-Colombian period, especially cultivars with endosperm *su1* (Hendry, 1930). Its center of origin is Latin America, but the relationship between the first cultivars and modern sweet corn is unclear (Wellhausen et., 1952). For some researchers, current genotypes are descendant of the cultivar Maiz Dulce (Manglesdorf, 1974) and for others, they have independent origin (Erwin, 1934). Nevertheless, there is an agreement on the needing not only for the mutation of the *su1/sh2* genes, but also for the accumulation of modifying genes that improve the viability of the seeds and, consequently, contributed to the development of modern cultivars.

Despite its occurrence throughout the American territory, sweet corn was not widely cultivated, initially, and/or consumed in its fresh form due to its “gummy” consistency but used in the preparation of other traditional American dishes, such as *pinole* and *kanocha*, and in the manufacture of alcoholic beverages (Carter, 1948). The scenario was only changed thanks to the loss of a gene responsible for the initial gummy consistency (Galinat, 1971) and the crossing with adapted starchy varieties, which contributed to increase the acceptance of the product.

Modern genotypes differ from common maize (*Zea mays*) due to its lower starch concentrations (<25-30%) and higher sugar contents (>11-20%), which affect the organoleptic properties and interfere in its grain texture and flavor (Tracy, 2001). This peculiarity is due to the action of one or more genes in recessive homozygosity, that act in the conversion of glucose into starch during the grain maturation stage, inhibiting the full process and culminating in an endosperm with differentiated composition (Dodson-Swenson & Tracy, 2015). This distinction made sweet corn to present a better palatability and, consequently, acceptance for human consumption, which made it to be included in the class of vegetables (Tracy, 2001).

Being destinated for human consumption, sweet corn genotypes must present, in addition to good palatability, high productivity, resistance package and/or tolerance

to the main phytopathogenic organisms that affect the crop, high industrial recovery rate and good looking ears, with a length and diameter of ~20cm and ~5cm, respectively, adequate husk length to protect the ear and other important traits that make their phenotyping a costly practice for breeding programs.

1.2.2 Socioeconomic importance of Sweet Corn (*Zea mays saccharata*)

Sweet corn is consumed at the milk stage of the endosperm development, in industrialized and fresh forms, especially in the Americas and Pacific Asia regions. It is often used in traditional dishes worldwide, as fresh corn, special soups, corn creams, salads and even barbecues. Sweet corn kernels taste sweet, especially, at 20 (earlier varieties) to 50 days after pollination and the cultivars super sweets or extra sweets, that have a shrunken (*sh*) gene, present their original levels of sugar increased and their conversion process (sugar into starch) delayed even more (Haynes et al., 2018). However, it is important to emphasize that the quality of fresh sweet corn decreases rapidly if its storage is inadequately, mainly at high temperatures, facing an increment in respiration rates and losses of sugar (Xie et al., 2017).

Aside from a high sugar concentration during its commercial form, sweet corn provides important phytochemicals such as tryptophan (Revilla et al., 2021), carotenoids (Song et al., 2016a; 2016b), anthocyanin (Hong et al., 2020), flavonoids (Zhang et al., 2018) and phenolic compounds (Zhang et al., 2017; Das & Singh, 2016). Furthermore, it is rich in vitamins A, B, C and contains calcium, phosphorus, iron, manganese and potassium (Khan et al., 2018). The nutritional value of sweet corn kernels relates to its high-water content and to the total amount of solid parts, that includes hydrocarbons (81%), having starch as the main one; proteins (13%); lipids (3.5%) and others (2.5%) (Szymanek, 2012).

With a market value estimated in $1.4 billion dollars, in 2015, where 26% referred to the processed consumption (USDA, 2015), sweet corn highlights as one of the main important vegetables worldwide, being an important source of income and employment for millions of workers. The United States of America is the highest sweet corn producer, having reached 2237 million pounds in 2018, with California state highlighting in the fresh market and Minnesota state in the frozen/processed one

(Becerra- Sanchez & Taylor, 2021). Furthermore, only the seeds market (global) of sweet corn was estimated in US$ 820.3 million in 2023, with a perspective of achieve US$ 1.19 billion until the end of the year (Global Market Insights, 2023) and emphasizing the economic importance of this special crop.

Regarding Brazilian market, 40,000 hectares were used for growing sweet corn in 2016, most of which were assigned for industrial processing, in the form of canned grains (EMBRAPA, 2016). Even though processing is responsible for almost all production in Brazil, fresh consumption has been gaining supporters in the national territory, thanks to the advent of tropicalized genotypes with better organoleptic characteristics and the constant search, on the part of consumers, for novelties in the food sector. Furthermore, more than 135 genotypes, between inbreeds and hybrids, can be found registered on the in the Brazilian Department of Agriculture (MAPA, 2023), also showing an increment on the interest of private and public companies in the sector. Although it has increased, information about sweet corn cultivation and particularities are still not widespread in Brazilian territory.

1.2.3 Production of Sweet Corn (*Zea mays saccharata*)

1.2.3.1 Starting in sweet corn field.

Sweet corn differs from field corn in the importance attributed to characteristics related to both husked and unhusked ears, in addition to those affecting palatability. Their importance varies according to the destination of the vegetable and impact on the necessary cares during sweet corn's cultivation. For fresh market, in addition to yield traits, ears must present a good shape, adequate husking coverage, straightness of rows, brightness of husk leaves and kernel colors and good palatability to stimulate the consumption. On the other hand, for process market, brightness is not so important as recovery rate or husking ability, which will greatly impact on the operation and industrial results. A commercial variety in Brazil, targeting process market, must produce at least 20 t. ha^{-1} of green yield, be tolerant to the main phytopathogens that affect the crop, especially against corn stunt and ear rot, and guarantee more than 40% of recovery rate, in addition to good husking ability, long kernel's depth and eating

quality. Unlike field corn varieties, which are harvested when the kernels are dry and fully mature, sweet corn is harvested when immature (milk stage), around 75-110 days after planting, when presenting 74%-79% of moisture and is consumed as a vegetable (Schultheis, 1994).

Being controlled by recessive genes (*su1/sh2),* sweet corn is highly affected by xenia, which is the immediate effect of foreign pollen in the endosperm of some angiosperms (Waller, 1917), resulting in cultivation impacts and requiring isolation between sweet and common maize, in order to avoid drastic changes on its organoleptic characteristics.

Even though the crop might have different destinations, to start a sweet corn field, aiming seed production or final market, the first thing to attempt is high quality seed with good emergence and getting that is harder than for other types of corn. The mutants change the starch pathway and, consequently, the endosperm composition, culminating in nearly all cases reduced starch levels. Kernels are more angular and present a thinner pericarp, which may split or blister away from the endosperm (Tracy, 2001). Furthermore, having a lower concentration of starch, a major composer of the endosperm and the kernel, sweet corn grains tend to present depressions caused by their non-filling, as a result of the metabolic process that occurs. Thus, it is normal to find germination and seedling vigor lower than standard field corn, even more if different mutants are combined (Tracy, 2001). Nonetheless, breeding and improvements in seed production and treatment have greatly reduced this issue (Tracy et al., 2020a; Tracy et al., 2020b) and now commercial sweet corn genotypes present germination and vigor as high as standard field corn.

1.2.3.2 Plant Nutrition

Having acquired a seed with good quality, next step is to attempt to soil nutrition as sweet corn performs best in fertile, well-drained soils, with a pH of 6.0-6.5, receiving at least six hours of direct sun daily (Haynes et al., 2018). It is also important to emphasize that sweet corn genotypes present a decrease in emergence and seedling vigor on cold and wet soils (Revilla et al., 2021), having a higher impact when soil temperatures are below to 13 ˚C (Haynes et al., 2018).

In order to achieve the genetic potential, sweet corn genotypes require adequate mineral nutrition, especially nitrogen (N), which is the most important element for the crop, impacting on productivity and response to environment effects (Revilla et al., 2021). Although it has a huge importance for sweet corn, N is highly and easily leached and its deficiency may culminate in slightly decreases in phosphorus (P) content of kernels (Zucareli et al., 2018), in addition to the reduction of yield. Furthermore, nitrogen impacts directly on plant height, ear length, ear diameter, crude protein ratio (Revilla et al., 2021) and in the eating quality (Mohammed et al., 2017). To reduce these impacts, several authors studied the mineral nutrition and fertilization in sweet corn. According to Zucareli et al. (2018), N must be applied at a rate of 120 kg. ha^{-1} at V6 stage to guarantee an adequate yield and protein maintenance in the kernels. Liu et al. (2019), studying the effects of urea on fresh ear yield and nitrogen efficiency in sweet corn, found a linear relationship between urea utilization and higher root growth, better leaf physiological functions and an increment in soil nitrogen. Pangaribuan et al. (2018) recommended the integrated use of organic fertilizer and urea in sweet corn, aiming a better postharvest quality and soil health.

Sweet corn is a crop that requires a moderate amount of phosphorus and potassium (K), as the first is responsible for energy production, energy transportation and is an important structural component of DNA, RNA and cell membranes and the second highlights for play essential roles in processes like stomatal opening, water supply during cell elongation, protein synthesis, organization of enzymes and photosynthesis (Tripathi et al., 2014). Additionally, calcium (Ca), another important macronutrient plays significant roles during the formation of cell membrane and cell wall elongation (Kadir et al., 2004). It also potentially enhances the tolerance against several viral and bacterial diseases (Fouda, 2017) and support canopy growth along with dynamic root and leaf expansion (Amor & Marcelis, 2003). Finally, magnesium (Mg) and sulfur (S) are also important for corn and sweet corn development, for being key players of the photosynthesis's process, activating RNA polymerase (Mg) and participating on the absorption of nitrogen (S).

Similarly, essential micronutrients such as molybdenum (Mo), boron (B), chlorine (Cl), copper (Cu), manganese (Mn), zinc (Zn) and iron (Fe) are known to have several important structural and functional roles in plant development. Molybdenum

helps in effective utilization of soil nitrogen, as it is an important element for the nitrate reductase and, consequently, in the nitrogen metabolism (Kaiser et al., 2005). Boron is an important player for reproductive development and consequently for pollination and fecundation. It also acts on the production and manufacturing of carbohydrates and sugars that are essential for the development of fruits and seeds (Hänsch & Mendel, 2009). Cl is found to be plentiful in the soil and is essential for the development of flora. Being highly found, it is simply taken by plants in large quantities and deficiencies are rarely found. So, in order to fulfill the physiological requirements of plants, this element is only required in a very low concentration of few ppm (Garner et al., 1930). Cu is an essential element for reproductive growth, root metabolism and utilization of essential proteins (Karakurt & Aslantas, 2010). Mn is required for important plant processes, such as activation of antioxidative enzymes, respiration, photosynthesis (Santandrea et al., 2000) and is also known for stimulating highly valuable enzymes, predominately those that retard the manufacturing of nucleotides and fatty acids (Nadeem et al., 2018). Zn is a vital plant nutrient that helps in several cellular processes, regulates the consumption of sugars, being an indispensable part of the enzyme system (Nadeem et al., 2018) and a ribossomal component. Last but not least, Fe is an essential element that helps in formation of chlorophyll (Kim & Guerinot, 2007), impacting directly in the metabolism and plant's life.

Having on mind the importance to provide all the essential nutrients for an adequate development, it is essential to carry out a soil analysis, before sowing, to verify the present levels of each mineral and to correct the deficiencies and excesses before moving forward with a sweet corn production.

1.2.3.3 Cultural management

Even though it is considered a vegetable and has key differences when compared to field corn, the cultivation of the sweet specie is quite similar to the field corn. However, knowing that sweet corn has a final product with higher aggregate value and that requires a more elaborate strategy to achieve its final consumer, good practices and differentiated managements need to be implemented (Pereira Filho & da Costa, 2021).

In Brazil, sweet corn is being produced in the spacings of 0.5-0.7m between rows and ~0.25-0.33m between plants, in order to achieve populations of ~60k plants, variating according to the technology available. This target population is quite similar to what American growers are implementing (Williams, 2012; Dhaliwal & Williams, 2019). With the breeding advances, growers are not double sowing as in the past, due to the improvements achieved on the germination of new varieties and are only correcting the percentage of lost (vigor/emergence). Another important aspect of sweet corn cultivation is the period for weed control, since the competition affects the development of the crop and directly impact on plant height, leaf area and green yield of sweet corn genotypes. According to Simarmata et al. (2018), the critical period for weed control is from 3 to 53 days after sowing, when the crop is still developing, emitting a root system and the canopy is not yet closed, culminating in an ample competition. As hand weeding has high labor costs and sweet corn presents to be more sensible to some herbicides, as Nicossulfuron (Choe & Williams, 2020), it is highly important to previously reduce the population of weeds. Furthermore, new varieties are being developed using transgenes and traits as GA21, which confers resistance against glyphosate, helping growers to overcome this obstacle. In countries where transgenic food is not accepted, breeding will also have to look for genotypes with tolerance to the main herbicides available in the market, as tembotrione.

1.2.3.4 Water management and its importance for sweet corn

Almost all sweet corn cultivation, in Brazil and in several countries, is done under constant irrigation, either via drips, sprinklers and mainly via pivots, due to the high-water dependence that the crop has and knowing that its cultivation under water stress results in a lower biomass accumulation (Stone et al., 2001), that allied with a lower absorption of nutrients, will impact the physiology and metabolism of sweet corn plants, as well as the final production (Traore et al., 2000; Payero et al., 2009; Singh et al., 2019; Karvar et al., 2023). The effect of a water deficit may be aggravated if the crop is in the vegetative stage or during the period of exposure of the style-stigmata (female flowering), resulting in early dehydration and a consequent failure in fecundation, culminating in a lower commercial production (Cordner, 1942; McGillivray, 1949).

Water deficit is the most important abiotic stress for sweet corn and can impact directly not only in final yield, but also on the nutritional quality of kernels, since sweet corn genotypes tend to accumulate more protein and sugars under such conditions (Ertek & Kara, 2013), which changes their organoleptic properties and, consequently, the consumer acceptance. Moreover, the lack of water can impact, significantly, on industrial logistics, disturbing the entire production chain and vegetable consumption, by affecting the flowering period, vegetative growth, stomatal and osmotic activity (Tambussi et al., 2007; Richards, 2014) and affecting seed moisture contents, seed germination, vigor index, seedling growth and, consequently, impacting on the processors planning and strategy (Li et al., 2017).

Genetic variability for drought tolerance is limited in sweet corn and the use of agronomic strategies to optimize irrigation is essential for a sustainable agriculture. To overcome this issue, different techniques have been implemented, as evapotranspiration maps, that assess water stress and apparent soil electrical conductivity (Nocco et al., 2019). They allow a precise recommendation, even more when regionalized due to a better understanding of climate variability and has already been implemented in US (Nocco et al., 2019). Moteva et al. (2016) established irrigation paraments for sweet corn cultivation using drips and sprinklers, aiming to optimize yield components and found that drips provided better conditions for green biomass accumulation, due to a more homogenous water supply and sprinklers improved productivity, by improving the microclimate in the field (lower temperatures and higher relative moisture) and, consequently, supporting the fecundation of plants.

In any of the methods used, attention must be paid to the needs of the crop and its less rusticity, thus, many farmers start their cultivation with a heavy water slide, followed by supplementation twice a day (or daily), until the pre-flowering stage, where it is slightly decreased (can be done daily). Attention is increased during full pollination, since the lack of water might result in a lower fecundation rate (Cordner, 1942; MacGillivray, 1949). Subsequently, irrigation is reduced on alternate days and as the need for supplementation is verified. According to Haynes et al. (2018), a plant of sweet corn requires approximately 25.4mm of water (either from rainfall or irrigation) per week for a normal growth and attention has to be increased at flowering and also during ear development. Last but not least, it is important to emphasize that almost all sweet corn

production in Brazil, for processing, is done under central pivot, which ends up favoring its cultivation throughout the year and the consequent supply of the consumer market, but also makes the production system more expensive.

1.2.3.5 Phytosanitary management

Sweet corn is a crop that suffers with phytopathogenic organisms from early stage until close to harvest. They can be bacteria, fungi, insects and even all of them. The initial pathogens are considered the most important due to their ability to kill the plant, reducing the final stand and, consequently, impacting on production. To overcome them, several techniques are implemented, as seed treatment, biological and chemical spraying, breeding and trait introgression.

Among the insects that attack the crop, highlights in Brazil corn earworm (*Helicoverpa zea* Boddie), fall armyworm (*Spodoptera frugiperda* J. E. Smith) and corn-silk fly (*Euxesta* sp.), causing damages that include the consumption of style-stigmata, reducing pollination; absence of grains at the tip of the ear; destruction of developing grains; increment of ear vulnerability to pathogens; quality reduction and even plant death (Pereira Filho & da Costa, 2021). These insects can be naturally controlled using parasitoids as *Trichogramma* spp., fungi as *Metarhizium* sp., bacteria as *Bacillus thuringiensis* and even virus as *Baculovirus* sp.; biologically by introducing genes that act on plant resistance to insects, such as Mir162 and Bt11; and chemically, using different molecules as clorantraniliprole, lambda-cyhalothrin, indoxacarb, emamectin benzoate and others.

Regarding the diseases that affect sweet corn crop in Brazil, corn stunt highlights as the most dangerous and important one. It results in stunted plants that produce small ears with a lower kernel content. Corn stunt is caused by *Spiroplasma kunkelii*, a bacterial pathogen that is transmitted singly or in combination with *Maize bushy stunt phytoplasma* (MBSP) and/or *Maize rayado fino virus* (MRFV) to healthy sweet corn plants by the corn leafhopper, *Dalbulus maidis*. During the 2017-2018 period, Brazilian corn production was severe damaged by corn stunt and efforts to control the leafhopper led to an increment of 85% in pesticide usage compared to the previous seasons, resulting in economic constraints to local growers (Gottems, 2018).

There is no current strategy established for managing corn stunt disease directly (Jones & Medina, 2020) and the control methods are relying on suppress or completely eradicate the vector (*D. maidis*), resulting in indirect pathogen control (Pérez-López et al., 2018), mainly with insecticides as bifenthrin, zeta-cypermethrin, carbosulfan, plinazolin and others. Furthermore, the development of varieties with acceptable tolerance against corn stunt is also a key strategy to overcome this challenging scenario, making this trait one of the most important in a breeding program for maize and sweet corn.

Bacteria as *Pseudomonas* sp. and *Erwinia* sp. might cause stalk root in sweet corn genotypes and can led to a seriously reduction on yield when present. They usually start on the lower part of the plants and rapidly reach the top. High temperatures and relative moisture favor their appearance and progression. So, to avoid or reduce their impacts, it is important to seek for an adequate water management and improvements on soil drainage, aligned with chemical spraying using products rich in casugamicine.

Sweet corn is committed by a series of fungi disease, highlighting northern corn leaf blight (*Exserohilum turcicum*), southern corn leaf blight (*Bipolaris maydis*), *Fusarium* sp. and *Phaeosphaeria maydis* leaf spot, for their ability to cause damage and the complete destruction or failure to reach the final product at all stages of the plant. They can appear at early phases and in the rotting of the ears (*Fusarium, Bipolaris*), which particularly makes the ears unfeasible for human consumption due to the production of fumonisins (Blacutt et al., 2018) and other mycotoxins. Leaf diseases are being well controlled by Brazilian farmers using chemicals as strobilurin, triazoles and carboxamides. Nonetheless, controlling ear rot in sweet corn has been a huge challenge for growers and processors, due to the constant presence of silk flies and ear worm, that increase the vulnerability to the fungi and allied with a constant irrigation, create the perfect macroclimate for ear rotting (Pereira Filho & da Costa, 2021). To overcome these issue, companies and public sectors are improving their cultivars, especially their husking coverage and tolerance to the main ear pathogens.

1.2.3.6 Harvest and post-harvest of sweet corn

Sweet corn is harvested when immature (milk stage), around 75-110 days after planting, or 15-40 days after flowering and consumed as a vegetable (Schultheis, 1994). At this time, the silks are brown and dry at the ear tip, moisture is around 70-79% and all the plant parts are showing a high concentration of water, being a potential source of subproducts that can improve the agricultural waste management and profitability (Revilla et al., 2021), when used as forage (Chaudhary et al., 2016; Habibpor et al., 2016; Lauriault et al., 2018).

Sweet corn matures faster in hot weather and slower when climate conditions are cool, remaining in the milk stage for a short time and facing a rapidly decreasing in its quality following the peak (Szymanek, 2009; Haynes et al. 2018). An accurate determination of its maturity, for harvesting, can lead to a higher crop yield and better quality (Ruan et al., 1999), as the kernels are plump, sweet, milky, tender, and nearly of maximum sizes. However, it is important to emphasize that when sweet corn is harvested before full maturation, it shows a small diameter, a poor cob fill and kernels that are watery with a lack of sweetness (Tracy, 2001; Szymanek, 2009). Furthermore, overmature sweet corn is rather starchy than sweet, with kernels often dented (Motes et al., 2007) and also needs to be avoided due to a faster decrement on eating quality (Tracy, 2001).

Harvest can be done manually or mechanically, depending on the final destination of the product, cultivar and specially the environmental, economic and cultural conditions. For fresh market, a homogeneous maturity is not so important as for process market, where cultivars must be adapted to husking, cutting and machine harvest, that are improved when we have plants with similar moisture content (Tracy, 2001).

Due to the high perishability of sweet corn ears and kernels, the correct strategies of harvest, transportation, processing and storage are essential to reduce losses and ensure a long and high-quality shelf life (Lum et al., 2016; Becerra-Sanchez & Taylor, 2021). When transportation is done inadequately and in the hottest periods of the day, it can lead to a significantly reduction in moisture and water content, due to a high transpiration rate (Boyette et al., 1990) and, consequently, to a decrement in eating quality. Cooling, handling, packaging and shipment are some of the main action

points, as transportation, when working with fresh crops, aiming an optimization of the supply chain (Watkins, 2016).

For fresh market and targeting for high quality, sweet corn should be immediately cooled down, after harvest, reducing the environment interaction and the heat produced by cob respiration (Becerra-Sanchez & Taylor, 2021). After that, ears can be stored at 0 °C (Shao & Li, 2011; Liplap et al., 2013; Xie et al., 2017). According to Xie et al. (2017), the optimum retailing storage condition was under 4 °C. Besides that, the authors found a significantly variation in organoleptic characteristics, weight, soluble sugar content, vitamin C content and soluble protein, under different retail conditions. Standard sweet corn cultivars may lose 50 percent of their sugar within 12 hours of harvest if not immediately refrigerated (Haynes et al., 2018) and, knowing that a loss of 2% in moisture can lead to significant kernel denting, which compromises the freshness appearance (Ashrae, 2010), their refrigeration needs special attention. Haynes et al. (2018) also emphasize that the ears can be stored in the refrigerator at 0 °C, for 4-8 days, without significantly losses, if they are kept unhusked. However, freezing injury is also possible if they are kept for a long time under refrigeration and/or in a very cold temperature (Becerra-Sanchez & Taylor, 2021). So, timing is crucial to guarantee a high-quality product, even though new cultivars are being developed to slow a little bit the conversion of glucose into starch, seeking for a higher shelf-life.

Targeting process market, sweet corn ears must be processed as rapidly as possible after harvest, avoiding a moisture decrement and all processes described above caused by water loosing. Furthermore, a correct time for processing culminates in a reduction of mechanical damage that negatively impact on the taste and eating quality of the finished product (Papusha et al., 2020). Moreover, a longer time between harvesting and processing can result in obstacles in the husking process, affecting not only the quality of the ear/kernels, but also increasing the industry's rate of return and, consequently, making the production more expensive. As described above, a genotype must be able to undergo the processing steps, showing a good husk coverage, but an easy husking ability, culminating in a lower rate of return; furthermore, it has to present a good maturation homogeneity, to guarantee a better functionality and cost benefit (Tracy, 2001).

Sweet corn is a popular food in the US, Asia and Europe and its popularity is spreading to countries where consumers are not familiar with it, as Brazil. By accounting for several traits, as good flavor and pericarp tenderness, bright color, high productivity, a package of disease resistance and acceptable recovery rate, sweet corn phenotyping is a costly practice, making complementary strategies necessary. In this sense, the use of predictive models, based on the molecular information of the genotypes, presents itself as an excellent alternative in sweet corn breeding programs.

1.2.4 Genomic Selection

Thomas Roderick, in 1986, attributed the term genomics to the study of the structure, function and evolution of genomes, which consist of an individual's genetic information and determines all heritable traits. However, researchers were already fascinated with this new science in 1977, when Frederick Sanger, Allan Maxam and Walter Gilbert developed a technique that allows the reading of nucleotides in small fragments of DNA. Although it was a milestone in scientific history, this sequencing technique required a lot of resources and time, to generate an average of 200 sequenced bases per working day, which restricted its use only to genes and proteins of great interest. After 1986, with the advancement of technology and the automation of this technique, sequencing complete genomes, such as the human genome itself, became possible. Genomics has developed rapidly over the years and in 2015 the second phase of its revolution began, also called next generation sequencing (NGS), which relies on the use of sophisticated equipment that can determine the sequence of 1 billion bases in 1 hour. Compared to the description initially made by Sanger, these new technologies represent a significant increase in the capacity to generate genetic information (Vitorello et al., 2017).

Nowadays, it is possible to sequence entire organisms in a few hours and with costs much lower than the initial ones, allowing scientific programs to use this technology to base and guide their choices of methods, drugs, genitors (animals and plants) etc. The first completely sequenced plant was *Arabidopsis thaliana* (The Arabidopsis Genome Initiative, 2000), using the “old methodology” but today crops such as potato, rice, soy, sorghum, apple and maize are among those whose genomes

have already been sequenced and with the advance of technology, hundreds and thousands of progenies are being sequenced all over the world for breeding purpose.

Initially, genomic selection (GS) for hybrid prediction in plants was based on the genetic distance of the parents and on the existing linear relationship between that and the heterosis of the progenies (Falconer & Mackay, 1996). However, the strategy was not disseminated in the scientific community due to its low methodological efficiency, but it remained an excellent approach to maximize the use of available germplasm. With the development of high-density genotyping technologies, such as single nucleotide polymorphism (SNP) panels and genotyping sequencing (GBS), the feasibility and efficiency using molecular and GS data has increased, contributing to overcome existing biases (Melchinger et al., 1998). According to Meuwissen et al. (2001), the objective of genomic selection is to explore the linkage disequilibrium (LD) existing between certain quantitative trait loci (QTLs) and molecular markers throughout the entire genome of interest.

In genomic selection models, genotypic and phenotypic data from a given population are combined to build a statistical model, where the effects of all markers are estimated and, a posteriori, used to determine the genomic genetic value (GEBVs) of individuals, making feasible their selections based only on molecular data and the consequent shortening of generations. Model validation is usually done by dividing the sample in two, one population for training and other for validation, where the first one is used to estimate marker effects (Goddard & Hayes, 2007).

Genomic Selection can be performed using different methodologies, linear (RR-BLUP and GBLUP) and non-linear (Bayesian), being the ideal procedure the one that best accommodates the character structure and the number of markers (Resende et al., 2012). Among them, the genomic model of best linear unbiased prediction (GBLUP) is the most applied for using genetic marker's information to associate individuals based on their Genomic Relationship Matrix (GRM), resulting in satisfactory data (Habier et al., 2007). This and other methodologies were developed to overcome statistical and computational limitations such as model dimensionality, collinearity between marks and the complexity of quantitative traits, as well as to overcome the limitation existing regarding the number of degrees of freedom available to adjust all marker effects in high-density panels (Lande & Thompson, 1990).

The accuracy of the genomic prediction model is highly influenced by the proximity between training and validation populations and their respective characteristics, such as heritability, extent of linkage disequilibrium (LD) and genetic distance (Habier et al., 2007; Solberg et al., 2008; Hayes et al., 2009; Zhong et al., 2009). It is also impacted by the effective sample size, quantity and quality of phenotypic data, and marker density (Hayes et al., 2009), which are directly related to project costs.

1.2.5 Markers Density

Being directly responsible for the accuracy and cost of the process, marker density plays a central role in genomic selection. High-density panels have a greater predisposition to obtain significant accuracies, due to a greater LD between the QTL of interest and the markers (Heffner et al., 2009; Desta & Ortiz 2014). However, the use of robust panels presents, in addition to the economic obstacle, the non-guarantee of high accuracy, as well as the high probability of reaching a *plateau* after a certain increase in the number of marks (Heffner et al., 2011; de Los Campos et al., 2013). This moment of stability is determined by the genetic diversity of the population under study and by the similarity between the training and validation sets, where the closer they are, the smaller the need for use robust panels (Hickey et al., 2014).

Several studies have been carried out to find minimum densities of markers that would, in parallel, bring significant accuracies and cost saving (Habier et al., 2009; Wellmann et al., 2013; Ma et al., 2016; Werner et al. al., 2018; Sousa et al., 2019; Lee et al., 2021) using different methodologies, but none in the sweet corn crop.

Even though several studies have found appropriate results to corroborate the reduction of markers, many authors suggested that low alleles frequency markers with potential larger effects may not work for genomic prediction due to their low precision estimating effects in a finite size training population (Lettre, 2011; Park et al., 2011).

The previous choice of markers to be used in GS can be done based on blocks of haplotypes (Ma et al., 2016), which are specific combinations of alleles observed in a population (Gabriel et al., 2002). As these are clusters of alleles, haplotypes are expected to be in greater linkage disequilibrium with the QTLs of interest compared to

individual SNPs and, consequently, to be better predictors in genomic selection (Zondervan & Cardon, 2004). The construction of haplotype blocks based on LD, using reliable thresholds, is also an excellent operational and economical alternative, as it considerably reduces computational time (Cuyabano et al., 2015).

Another selection strategy is the choice of markers evenly spaced along the genome, where the density to be used and the consequent expected coverage/distribution along the genome are predetermined (Ma et al., 2016). Although simple, this methodology has never been used in the sweet corn crop.

Although some authors do not agree with the strategy of reducing markers and using anchors, it is clear the need to seek alternatives to reduce computational time and costs. Furthermore, based on the numerous positive results found in animal breeding and in other cultures, the previous marker's selection and the creation of low-density panels for carrying out the genomic selection in the sweet corn crop, is presented as an excellent alternative.

1.3 REFERENCES

Amor FD, Marcelis L (2003) Regulation of nutrient uptake, water uptake and growth under calcium starvation and recovery. **The Journal of Horticultural Science and Biotechnology**, v.78, n.3, p.343-349. Accessed on August 17th, 2023.

Ashrae (2010) ASHRAE Handbook: Refrigeration (SI Edition). **American Society of Heating Refrigerating and Air-Conditioning Engineers**. Accessed on August 23th, 2023.

Barbieri VHB (2010) **Mapeamento de QTL em testecrosses de milho doce com diferentes testadores e ambientes**. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, University of São Paulo, Piracicaba. Accessed on July 2nd, 2020.

Becerra- Sanchez F, Taylor G (2021) Reducing post- harvest losses and improving quality in sweet corn (Zea mays L.): challenges and solutions for less food waste and improved food security. **Food and Energy Security**, v.10, n.3. Accessed on August 17th, 2023.

Bernardo R (1995) Genetic models for predicting maize performance in unbalanced yield trial data. **Crop Science**, Madison, v.35, n.1, p.141-147. Accessed on July 2nd, 2020.

Blacutt AA, Gold SE, Voss KA, Gao M, Glenn AE (2018) *Fusarium verticillioides*: Advancements in Understanding the Toxicity, Virulence, and Niche Adaptations of a Model Mycotoxigenic Pathogen of Maize. **Phytopathology**, v.108, n.3, p.312-326. Accessed on August 22th, 2023.

Boyette MD, Wilson LG, Estes EA (1990) Postharvest cooling and handling of sweet corn in North Carolina, AG-413-4. N.C. **Agricultural Extension Service**. Accessed on August 22th, 2023.

Carter GF (1948) Sweet corn among the Indians. **Geographical Review**, v.38, p.206-22. Accessed on April 26th, 2022.

Chaudhary DP, Kumar A, Kumar R, Singode A, Mukri G, Sah RP, Tiwana US, Kumar B (2016) Evaluation of normal and specialty corn for fodder yield and quality traits. **Range Management and Agroforestry**, v.37, p.79–83. Accessed on August 20th, 2023.

Choe E, Williams MM (2020) Expression and comparison of sweet corn CYP81A9s in relation to nicosulfuron sensitivity. **Pest Management Science**, v.76, p.3012–3019. Accessed on August 22th, 2023.

Cordner HB (1942) The influence of irrigation water on the yield and quality of sweet corn and tomatoes with special reference to the time and number of applications. **Proceedings of the American Society for Horticultural Science**, v.40, p.475-481. Acessed on March 27th, 2023.

Cuyabano BC, Su G, Lund MS (2015) Selection of haplotype variables from a high-density marker map for genomic prediction. **Genetics Selection Evolution**, v.47, p.61.

Accessed on October 23th, 2022.

Dhaliwal DS, Williams MMII (2019) Optimum plant density for crowding stress tolerant processing sweet corn. **PLoS One**, v.14, n.9. Accessed on August 2nd, 2023.

Das AK, Singh V (2016) Antioxidative free and bound phenolic constituents in botanical fractions of Indian specialty maize (Zea maysL.) genotypes. **Food Chemistry**, v.201, p.298-306. Accessed on September 28th, 2021.

De Los Campos G, Hickey JM, Pong-Wong R, Daetwyler HD, Calus M MPL (2013) Whole-genome regression and prediction methods applied to plant and animal breeding. **Genetics**, v.193, p.327–345. Accessed on August 15th, 2022.

Desta ZA, Ortiz R (2014) Genomic selection: genome-wide prediction in plant improvement. **Trends in Plant Science**, v.19, p.592–601. Accessed on August 15th, 2022.

Dodson-Swenson HG, Tracy WF (2015) Endosperm carbohydrate composition and kernel characteristics of shrunken2- intermediate (sh2-i/sh2-i Su1/Su1) and shrunken2-intermediate-sugary1-reference (sh2-i/sh2-i su1-r/su1-r) in sweet corn. **Crop Science**, Madison, v.55, p.2647-2656. Accessed on July 3rd, 2020.

Empresa Brasileira de Pesquisa e Agropecuária **EMBRAPA** (2016) O cultivo do Milho Doce. EMBRAPA Milho e Sorgo, Brasília, DF. p.289. Accessed on July 3rd, 2020.

Endelman JB (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. **The Plant Genome**, v.4, p.250-255. Accessed on July 2nd, 2020.

Ertek A, Kara B (2013) Yield and quality of sweet corn under deficit irrigation. **Agricultural Water Management**, v.129, p.138-144. Acessed on April 8th, 2023.

Erwin AT (1934) Sweet corn — its origin and importance as an Indian food plant in the United States, **Iowa State College Journal of Science**, v.8, p.385. Accessed on January 20th, 2022.

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. **Longman Scientific and Technical**, 4th ed., 464p. Accessed on July 10th, 2022.

Fouda K (2017). Effect of Interaction among N Forms and Calcium Sources on Quality and Chemical Composition of Tomato (Lycopersicon). **Egyptian Journal of Soil Science**, v.57, n.2, p.61 -71. Accessed on August 21th, 2023.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structureof haplotype blocks in the human genome. **Science**, v.296, n.5576, p.2225–2229. Accessed on October 24th, 2022.

Galinat WC (1971) The evolution of sweet corn. **Bulletin-Massachusetts Agricultural Experiment Station**, v.591. Accessed on April 26th, 2022.

Ganal MW, Durstewitz G, Polley A, Berard A, Buckler ES, Charcosset A, Clark JD, Graner EM, Hansen M, Joets J, Le Paslier MC, McMullen MD, Montalent P, Rose M, Schön CC, Sun Q, Walter H, Martin OC, Falque M (2011) A large maize (Zea mays

L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. **PloS one**, v.6, n.12. Accessed on August 14th, 2022.

Garner W, McMurtrey Jr J, Bowling J, Moss E (1930) Role of chlorine in nutrition and growth of the tobacco plant and its effect on the quality of the cured leaf. **Journal of Agriculture Research**, v.40, p.627-648. Accessed on August 21th, 2023.

Global Market Insights (2023) Sweet corn seeds maerket by seed type (Hybrid certified seeds, farm saved seeds, open pollinated certified seeds), by seed category (yellow, white, bicolor), by farming type (organic, conventional), by sales channel & by region. **Fact.MR, Food and Beverage**, p.170.

Goddard ME, Hayes BJ (2007) Genomic Selection. **Journal of Animal Breeding and Genetics**, Oxford, v.124, p.323-330. Accessed on July 25th, 2021.

Gottems L (2018) Sales of Insecticides for Corn Leafhopper Control Jump in Brazil. **Agro News**, v.3. Accessed on July 30th, 2023.

Habibpor SS, Naderi A, Lak S, Faraji H, Mojaddam M (2016) Effects of salicylic acid on morphological and physiological characteristics of sweet corn hybrids under water stress conditions. **Journal of Fundamental and Applied Sciences**, v.8, p.522–543. Accessed on August 20th, 2023.

Habier D, Fernando R, Dekkers J (2007) Genomic Selection Using Low-Density Marker Panels. **Genetics**, v.177, n.4, p.2389–97. Accessed on August 14th, 2022.

Habier D, Fernando R, Dekkers J (2009) The impact of genetic relationship information on genome-assisted breeding values. **Genetics**, v.182, n.1, p.343–353. Accessed on September 17th, 2022.

Hayes B, Bowman P, Chamberlain A, Goddard M (2009) Invited review: Genomic selection in dairy cattle: Progress and challenges. **Journal of dairy science**, v.92, n.2, p.433–443. Accessed on August 14th, 2022.

Hänsch R, Mendel RR (2009) Physiological functions of mineral micronutrients (Cu, Zn, Mn, Fe, Ni, Mo, B, Cl). **Current opinion in plant biology**, v.12, n.3, p.259-266. Accessed on August 21th, 2023.

Haynes C, Everhart E, Jauron R, Nair A (2018) Sweet Corn. **Iowa State University: Extension and Outreach**. HORT 3089. Accessed on August 14th, 2022.

Heffner EL, Sorrells ME, Jannink JC (2009) Genomic selection for crop improvement. **Crop Science**, v.49, p.1–12. Accessed on August 15th, 2022.

Heffner EL, Jannink JC , Sorrells ME (2011) Genomic selection accuracy using multifamily prediction models in a wheat breeding program. **Plant Genome**, v.4, p.65–75. Accessed on August 15th, 2022.

Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. **Biometrics**, Arlington, v.31, p.423-447. Accessed on July 4th, 2020.

Hendry GW (1930) Archaeological evidence concerning the origin of sweet maize. **Journal of the American Society of Agronomy**, v.22, p.508. Accessed on April 26th,

2022.

Hickey JM, Dreisigacker S, Crossa J, Hearne S, Babu R, Prasanna BM, Grondona M, Zambelli A, Windhausen VS, Mathews K, Gorjanc G (2014) Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. **Crop Science**, v.54, p.1476–1488. Accessed on August 15th, 2022.

Hong HT, Netzel ME, O'Hare TJ (2020) Optimization of extraction procedure and development of LC-DAD-MS methodology for anthocyanin analysis in anthocyanin-pigmented corn kernels. **Food Chemistry**, v.319. Accessed on August 19th, 2021.

Jacobson A, Lian L, Zhong S, Bernardo R (2014) General combining ability model for genomewide selection in a biparental cross. **Crop Science**, v.54, p.895-905. Accessed on July 7th, 2020.

Jones T-kL, Medina RF (2020) Corn Stunt Disease: An Ideal Insect–Microbial–Plant Pathosystem for Comprehensive Studies of Vector-Borne Plant Diseases of Corn. **Plants**, v.9, n.6, p.747. Accessed on April 25th, 2022.

Kadir T, Zisserman A, Brady M (2004). An Affine Invariant Salient Region Detector. In: Pajdla, T., Matas, J. (eds) European Conference on Computer Vision – ECCV. **Lecture Notes in Computer Science**, v.3021. Accessed on August 8th, 2023.

Kaiser BN, Gridley KL, Ngaire Brady J, Phillips T, Tyerman SD (2005) The role of molybdenum in agricultural plant production. **Annals of Botany**, v.96, n.5, p.745-754. Accessed on August 21th, 2023.

Karakurt H, Aslantas R (2010) Effects of some plant growth promoting rhizobacteria (PGPR) strains on plant growth and leaf nutrient content of apple. **Journal of Fruit and Ornamental Plant Research**, v.18, n.1, p.101-110. Accessed on August 21th, 2023.

Karvar M, Azaro A, Rahimi A, Maddah-Hosseini S, Ahmadi-Lahijani MJ (2023) Potassium silicate reduces water consumption, improves drought tolerance, and enhances the productivity of sweet corn (Zea mays) under deficit irrigation. **Acta Physiologiae Plantarum**, v.45, n.38. Acessed on April 10th, 2023.

Khan AA, Hussain A, Ganai MA, Sofi NR, Hussain ST (2018) Yield, nutrient uptake and quality of sweet corn as influenced by transplanting dates and nitrogen levels. **Journal of Pharmacognosy and Phytochemistry**, v. 7, n.2, p.3567-3571. Accessed on July 10th, 2023.

Kim SA, Guerinot ML (2007) Mining iron: iron uptake and transport in plants. **FEBS letters**, v.581, n.12, p.2273-2280. Accessed on August 21th, 2023.

Lande R, Thompson R (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. **Genetics**, v.124, n.3, p.743–56. Accessed on August 14th, 2022.

Lauriault LM, Guldan SJ, Popiel-Powers FG, Steiner RL, Martin CA, Heyduck RF, Falk CL, Petersen MK, May T (2018) Relay Intercropping with Cover Crops Improved

Autumn Forage Potential of Sweet Maize Stover. **Agriculture**, v.8, p.103. Accessed on August 20th, 2023.

Lee D, Kim Y, Chung Y, Lee D, Seo D, Choi TJ, Lim D, Yoon D, Lee SH (2021) Accuracy of genotype imputation based on reference population size and marker density in Hanwoo cattle. **Journal of Animal Science and Technology**, v.63, n.6, p.1232–1246. Accessed on September 18th, 2022.

Lettre G (2011). Recent progress in the study of the genetics of height. **Human Genetics**, v.129, p.465–472. Accessed on July 8th, 2020.

Li W, Zhang XL, Ashraf U, Mo ZW, Suo HC, Li GK (2017) Dynamics of seed germination, seedling growth and physiological responses of sweet corn under peg-induced water stress. **Pakistan Journal of Botany**, v.49, p.639–646. Accessed on July 18th, 2023.

Lian L, Jacobson A, Zhong S, Bernardo R (2014) Genomewide prediction accuracy within 969 maize biparental populations. **Crop Science**, v.54, p.1514-1522. Accessed on July 8th, 2020.

Liplap P, Vigneault C, Toivonen P, Boutin J, Raghavan G S V (2013) Effect of hyperbaric treatment on respiration rates and quality attributes of sweet corn. **International Journal of Postharvest Technology and Innovation**, v.3, n.3, p.257–271. Accessed on August 21th, 2023.

Liu W, Xiong YS, Xu XY, Xu FS (2019) Hussain, S.; Xiong, H.F.; Yuan, J.F. Deep placement of controlled-release urea effectively enhanced nitrogen use efficiency and fresh ear yield of sweet corn in fluvo-aquic soil. **Scientific Reports**, v.9. Accessed on August 21th, 2023.

Lum GB, Shelp BJ, DeEll, J R, Bozzo GG (2016) Oxidative metabolism is associated with physiological disorders in fruits stored under multiple environmental stresses. **Plant Science**, v.245, p.143–152. Accessed on August 23th, 2023.

Ma Y, Reif JC, Jiang Y, Wen Z, Wang D, Liu Z, Guo Y, Wei S, Wang S, Yang C, Wang H, Yang C, Lu W, Xu R, Zhou R, Wang R, Sun Z, Chen H, Zhang W, Wu J, Hu G, Liu C, Luan X, Fu Y, Guo T, Han T, Zhang M, Sun B, Zhang L, Chen W, Wu C, Sun S, Yuan B, Zhou X, Han D, Yan H, Li W, Qiu L (2016) Potential of marker selection to increase prediction accuracy of genomic selection in soybean (Glycine max L.). **Molecular Breeding**, v.36, n.113. Accessed on September 24th, 2022.

MacGillivray JH (1949) Effect of irrigation on the growth and yield of sweet corn. **Proceedings of the American Society for Horticultural Science**, v.54, p.330-338. Acessed on April 10th, 2023.

Manglesdorf PC (1974) Corn, Its Origin, Evolution, and Improvement. **Belknap/Harvard University Press**. Accessed on September 3rd, 2021.

MAPA - MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO (2023) **Registro Nacional de Cultivares**. Accessed on May 2nd, 2023.

Melchinger AE, Utz HF, Schön CC (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of

QTL detection and large bias in estimates of QTL effects. **Genetics**, v.149, n.1. Accessed on July 8th, 2020.

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v.157, p.1819–1829. Accessed on July 8th, 2020.

Mendes MP (2015) **Predição genômica de híbridos de milho**. Tese (Doutorado em Genética e Melhoramento de Plantas) – Escola Superior de Agricultura Luiz de Queiroz, University of São Paulo, Piracicaba. Accessed on July 8th, 2020.

Mohammed AA, Majid ZM, Kasnazany SAS, Salih SJ, Mustafa SB, Salih OA (2017) Growth and yield quality of sweet corn, as influenced by nitrogen fertilization levels in Sulaimani region. **Iraqi Journal of Agriculture Sciences**, v.48, p.1582–1589. Accessed on August 8th, 2023.

Motes JE, Roberts W, Cartwright B (2007) HLA6021-Sweet corn production. **Oklahoma Cooperative Extension Service**, p.1-4. Accessed on August 23th, 2023.

Moteva M, Gadjalska N, Kancheva V, Tashev T, Georgieva V, Koleva N, Mortev I, Petrova-Brahicheva V (2016) Irrigation scheduling and the impact of irrigation on the yield and yield components of sweet corn. **Scientific Papers. Series A. Agronomy**, v. 59, p.332–339. Accessed on August 22th, 2023.

Nadeem F, Hanif MA, Mushtaq Z (2018) Role of Macronutrients and Micronutrients in the Growth and Development of Plants and Prevention of Deleterious Plant Diseases – A Comprehensive Review. **International Journal of Chemical and Biochemical Sciences**, v.13, p.31-52. Accessed on August 21th, 2023.

Nocco MA, Zipper SC, Booth EG, Cummings CR, Loheide SP, Kucharik CJ (2019) Combining Evapotranspiration and Soil Apparent Electrical Conductivity Mapping to Identify Potential Precision Irrigation Benefits. **Remote Sensing**, v.11, p.2460. Accessed on August 21th, 2023.

Pangaribuan DH, Hendarto K, Elzhivago SR, Yulistiani A (2018) The effect of organic fertilizer and urea fertilizer on growth, yield and quality of sweet corn and soil health. **Asian Journal of Agriculture and Biology**, v.6, p.335–344. Accessed on August 8th, 2023.

Papusha SK, Nikitenko NA, Zhadko VV (2020) Parameters of the knife-drum type device for separating sweet corn cobs. **E3S Web of Conferences**, v.193, p.3. Accessed on August 23th, 2023.

Park JH, Gail MH, Weinberg CR, Carroll RJ, Chung CC, Wang Z, Chanock SJ, Fraumeni JF, Chatterjee N (2011) Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants. **Proceedings of the National Academy of Sciences of the United States of America**, v.108, n.44, p.18026–18031. Accessed on July 8th, 2020.

Payero JO, Tarkalson DD, Irmak S, Davison D, Petersen JL (2009) Effect of timing of a deficit-irrigation allocation on corn evapotranspiration, yield, water use efficiency, and dry mass. **Agricultural Water Management**, v.96, n.9, p.1387–1397. Acessed on April 10th, 2023.

Pereira Filho IA, da Costa RV (2021) **Milho Doce**. In: Milho. EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária. Accessed on August 28th, 2023.

Pérez-López E, Wist T, Rodríguez Y, Luna-Rodríguez M, Olivier CY (2018) Maize bushy stunt in native corn: Implications for Mexican “subsistence farmers”. **Environment, Development and Sustainability**, v.20, p.1797–1805. Accessed on April 25th, 2022.

Powell JE, Visscher PM, Goddard ME (2010) Reconciling the analysis of IBS in complex trait studies. **Nature Review Genetics**, v.11, p.800-805. Accessed on July 8th, 2020.

R Development core team (2012) **A language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing. 3604p. Accessed on July 3rd, 2020.

Reif JC, Melchinger AE, Frisch M (2005) Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. **Crop Science**, Madison, v.45, n.1, p.1-7. Accessed on May 5th, 2022.

Resende MDV, Silva FF, Lopes PS, Azevedo CF (2012) **Seleção Genômica Ampla (GWS) via Modelos Mistos (REML/BLUP), Inferência Bayesiana (MCMC), Regressão Aleatória Multivariada (RRM) e Estatística Espacial**. Viçosa: Universidade Federal de Viçosa. 291p. Accessed on July 3rd, 2020.

Revilla P, Anibas CM, Tracy WF (2021) Sweet corn research around the world 2015-2020. **Agronomy**, v.11, n.534. Accessed on August 5th, 2022.

Richards JHL (2014) Pre-dawn stomatal opening does not substantially enhance early-morning photosynthesis in Helianthus annuus. **Plant, Cell & Environment**, v.37, p.1364–1370. Acessed on March 13th, 2023.

Ruan RR, Chen PL, Almaer S (1999) Nondestructive analysis of sweet corn maturity using NMR. **HortScience**, v.34, p.319–321. Acessed on August 23th, 2023.

Santandrea G, Pandolfini T, Bennici A (2000) Aphysiological characterization of Mn-tolerant tobacco plants selected by in vitro culture. **Plant Science**, v.150, n.2, p.163-170. Accessed on August 21th, 2023.

SAS Institute (2008) **Statistical analysis software**: SAS 9.2. Accessed on July 10th, 2022.

Schultheis RJ (1994) **Sweet corn production**. North Carolina Cooperative Extension Service, North Carolina State university. Accessed on August 15th, 2022.

Shao X, Li Y (2011) Quality control of fresh sweet corn in controlled freezing-point storage. **African Journal of Biotechnology**, v.10, n.65, p.14534–14542. Accessed on August 23th, 2023.

Simarmata M, Nurjanah U, Setyowati N (2018) Determination of the critical period for weed control of sweet corn under tropical organic farming system. **Asian Journal of Agriculture and Biology**, v.6, p.447–454. Accessed on August 15th, 2023.

Singh M, Saini RK, Sing S, Sharma SP (2019) Potential of integrating biochar and

deficit irrigation strategies for sustaining vegetable production in water-limited regions: a review. **HortScience horts**, v.54, n.11, p.1872-1878. Acessed on July 5th, 2023.

Solberg T, Sonesson A, Woolliams J (2008) Genomic selection using different marker types and densities. **Journal of animal science**, v.86, n.10, p.2447–2454. Accessed on August 14th, 2022.

Song JF, Li DJ, He MJ, Chen JQ, Liu CQ (2016a) Comparison of carotenoid composition in immature and mature grains of corn (Zea maysL.) varieties. **International Journal of Food Properties**, v.19, p.351-358. Accessed on August 14th, 2022.

Song JF, Li DJ, Liu NY, Liu CQ, He MJ, Zhang Y (2016b) Carotenoid composition and changes in sweet and field corn (Zea mays) during kernel development. **Cereal Chemistry**, v.93, p.409-413. Accessed on August 14th, 2022.

Sousa MB, Galli G, Lyra DH, Granato ISC, Matias FI, Alves FC, Fritsche-Neto R (2019) Increasing accuracy and reducing costs of genomic prediction by marker selection. **Euphytica**, v.215, n.18. Accessed on September 24th, 2022.

Stone PJ, Wilson DR, Reid JB, Gillespie RN (2001) Water deficit effects on sweet corn. I. Water use, radiation use efficiency, growth, and yield. **Australian Journal of Agricultural Research,** v.52, p.103-113. Accessed on March26th, 2023.

Szymanek M (2009) Influence of sweet corn harvest date on kernels quality. **Research in Agricultural Engineering**, v.55, n.1, p.10–17. Accessed on August 23th, 2023.

Szymanek M (2012) Processing of Sweet Corn: Trends in Vital Food and Control Engineering. **InTech**. Accessed on August 17th, 2023.

Tambussi EA, Bort J, Araus JL (2007) Water use efficiency in C3 cereals under mediterranean conditions: a review of physiological aspects. **Annals of Applied Biology**, v.150, p.307–321. Acessed on April 9th, 2023.

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. **Nature**, v.408, p.796–815. Accessed on July 10th, 2022.

Tracy WF (2001) Sweet Corn. In: **Specialty corn. 2nd edition**. Boca Raton, CRC Press. p.155-198. Accessed on March 15th, 2020.

Tracy WF, Shuler SL, Dodson-Swenson H (2020a) The Use of Endosperm Genes for Sweet Corn Improvement: A review of developments in endosperm genes in sweet corn since the seminal. Plant Breeding Reviews, Volume 1, by Charles Boyer and Jack Shannon (1984). **Plant Breeding Reviews**, v.43, p.215–241. Accessed on August, 20th, 2023.

Tracy WF, Shuler SL, Dodson-Swenson H (2020b) Sweet Corn. In: **The Physiology of Vegetable Crops**, p.320–335. Accessed on August, 20th, 2023.

Traore SB, Carlson RE, Pilcher CD, Rice ME (2000) Bt and non-Bt maize growth and development as affected by temperature and drought stress. **Agronomy Journal**, v.92, n.5, p.1027–1035. Acessed on March 26th, 2023.

Tripathi DK, Singh VP, Chauhan DK, Prasad SM, Dubey N (2014) Role of Macronutrients in Plant Growth and Acclimation: Recent Advances and Future Prospective. In: **Improvement of Crops in the Era of Climatic Changes**, p.197-216. Accessed on September 15th, 2021.

USDA-NASS (2019) **United States Department of Agriculture–National Agricultural Statistics Service QuickStats**. Accessed on September 15th, 2021.

Vitorello CBM, Jorge EC, Camargo LEA (2017) Genômica: a revolução na análise dos genes. **Genética molecular básica: dos genes aos genomas**. Tradução Rio de Janeiro: Guanabara Koogan. Accessed on January 08th, 2023.

Waller AE (1917) Xenia and other influences following fertilization. **Ohio Journal of Science**, v. 17, n.8. Accessed on October 30th, 2023.

Watkins CB (2016) Postharvest physiological disorders of fresh crops. **Encyclopedia of applied plant sciences** (2nd ed.). Elsevier. Accessed on August 23th, 2023.

Wellhausen EJ, Roberts LM, Hernandez XE (1952) **Races of Maize in Mexico**. Bussey Institution, Harvard University, Cambridge, MA. Accessed on April 26th, 2022.

Wellmann R, Preuß S, Tholen E, Heinkel J, Wimmers K, Bennewitz J (2013) Genomic selection using low density marker panels with application to a sire line in pigs. **Genetics Selection Evolution**, v.45, n.28. Accessed on September 17th, 2022.

Werner CR, Voss-Fels KP, Miller CN, Qian W, Hua W, Guan CY, Snowdon RJ, Qian L (2018) Effective Genomic Selection in a Narrow-Genepool Crop with Low-Density Markers: Asian Rapeseed as an Example. **The Plant Genome**, v.11, n.2. Accessed on September 18th, 2022.

White TL, Hodge GR (1989) **Predicting breeding values with applications in forest tree improvement**. Dordrecht: Kluwer Academic, 376p. Accessed on July 2nd, 2020.

Whright S (1922) Coefficient of inbreeding and relationship. **The American Naturalist**, Chicago, v.56, n.645, p.330-338. Accessed on July 2nd, 2020.

Williams MM II (2012) Agronomics and economics of plant population density on processing sweet corn. **Field Crops Research**, v.128, p.55–61. Accessed on August 2nd, 2023.

Xie Y, Liu S, Jia L, Gao E, Song H (2017) Effect of different storage temperatures on respiration and marketable quality of sweet corn. **Advanced Engineering Technology**, v.3, p.219–224. Accessed on August 23rd, 2023.

Zhang RF, Huang L, Deng YY, Chi JW, Zhang Y, Wei ZC, Zhang MW (2017) Phenolic content and antioxidant activity of eight representative sweet corn varieties grown in South China. **International Journal of Food Properties**, v.20, p.3043-3055. Accessed on August 23rd, 2023.

Zhang ZB, Liang ZB, Yin LF, Li QX, Wu ZY (2018) Distribution of four bioactive flavonoids in maize tissues of five varieties and correlation with expression of the biosynthetic genes. **Journal of Agricultural and Food Chemistry**, v.66, p.10431-10437. Accessed on May 1st, 2021.

Zondervan KT, Cardon LR (2004) The comples interplay among factors that influence allelic association. **Nature Reviews Genetics**, v. 5, p.89–101. Accessed on October 24th, 2022.

Zhong S, Dekkers JC, Fernando RL, Jannink J- L (2009) Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a barley case study. **Genetics**, v.182, n.1, p.355–64. Accessed on August 14th, 2022.

Zucareli C, Bazzo JHB, Silva JB, Costa DS, Fonseca ICB (2018) Nitrogen rates and side-dressing timing on sweet corn seed production and physiological potential. **Revista Caatinga**, v.31, p.344–351. Accessed on August 21th, 2023.

CHAPTER 2 – EFFECTS OF MARKER DENSITY ON GENOMIC PREDICTION FOR YIELD TRAITS IN SWEET CORN

ABSTRACT – By accounting for many traits, phenotyping sweet corn is a costly practice, making complementary strategies necessary. Thus, predictive methods present as an excellent alternative for the prediction and selection of the traits. The accuracy of the prediction is highly influenced by characteristics such as phenotypic data quality and marker density, which impact on project costs. Several studies have been carried out to verify minimum densities without the significant loss in prediction accuracies, but none with sweet corn. In this study, the objectives were to test, assess and validate different strategies to pre-select markers for genomic selection and to find the minimum density in a prediction of yield traits in sweet corn. Initially, the prediction was performed with a high-density chip and then, using a pre-selection strategy of clustering markers into haplotype blocks. Furthermore, a third strategy was tested, where markers were selected evenly across the genome. In general, all traits showed a significant reduction in accuracy as the number of markers decreased. However, the relationship between marker's increment and accuracy did not remain constant and reached a plateau after a certain point. Applying marker pre-selection can be a good option for a cost-efficient implementation of genomic selection in sweet corn for yield traits, as they can be predicted with a significant accuracy using a panel of ~8k quality markers that are evenly across the genome. Furthermore, using one marker per haplotype block appears to be a better cost-effective strategy for carrying out genomic selection in sweet corn, for yield traits.

Keywords: Sweet Corn; genomic selection; marker density; accuracy.
Submitted in Euphytica international journal.

2.1 INTRODUCTION

Corn is a key crop all over the world, due to its social, economic and environmental aspect. Having a broad genetic base, it has been exploited by man for different purposes and, in this sense, the cultivation of special corns, characterized by a distinct composition of the endosperm, has been gained prominence for being destinated for human consumption (Tracy 2001).

Among special corns, the sweet specie *Zea mays sacharatta's* distinction is caused by the action of one or more genes in recessive homozygosis (generally, sugary-su1 and/or shrunken-sh2), inhibiting the total conversion of glucose into starch during the grain maturation process and resulting in better palatability (Dodson-Swenson and Tracy 2015). Due to these particularities, sweet corn consumption has been increasing over the years and, in 2023 the sweet corn seeds market (global) was valued in US$ 820.3 million, with a perspective of achieve US$ 1.19 billion of market value until the end of the 2023 year (Global Market Insights 2023).

Regarding Brazilian production, about 40,000 hectares were destined, in 2016, for sweet corn cultivation, most of which allocated for industrial processing (EMBRAPA 2016) and under controlled irrigation conditions (pivots), emphasizing the growing of the market and the importance Brazilian farmers and customers are seeing on this crop. Furthermore, more than 135 genotypes, between inbreeds and hybrids, can be found registered on the official national portal (MAPA 2023), by public and private sectors, also showing a greater interest of private and public companies in the sector.

In order to meet the agro-industrial needs, sweet corn varieties must present acceptable organoleptic characteristics as good flavor, pericarp tenderness and a bright color in addition to a high productivity, a package of disease resistance and an acceptable recovery rate (proportion of the green yield that will fill the cans) (Barbieri 2010). By accounting for several traits, under different genetic effects, sweet corn phenotyping is a costly practice for breeding programs, which make complementary strategies necessary. In this sense, the use of predictive models, based on the molecular information of the genotypes, presents itself as an excellent alternative for the costly phenotyping in sweet corn breeding programs.

In genomic prediction, genotypic and phenotypic data are combined to build a statistical model to estimate the effects of all markers and consequently the genomic estimate of genetic value (GEGV) of individuals, based on molecular data, which shortens the breeding cycle. The prediction accuracy of the model is highly influenced by the proximity between training and validation sets and their respective characteristics (Habier et al. 2007; Solberg et al. 2008; Hayes et al. 2009; Zhong et al. 2009). It is also impacted by the size of the training set, quantity and quality of phenotypic data and marker density (Hayes et al. 2009), which are directly related to project costs. Marker density plays a role in genomic selection (GS) as it affects project costs and the prediction accuracy, which by the way, plateaus when the marker density reaches a certain threshold (Heffner et al. 2011; de Los Campos et al. 2013). Several studies have been carried out to find minimum densities of markers that would, in parallel, bring significant accuracies and cost saving (Habier et al. 2009; Wellmann et al. 2013; Ma et al. 2016; Werner et al. 2018; Sousa et al. 2019; Lee et al. 2021), but none in the sweet corn crop.

The previous choice of markers to be used in a GS can be done based on blocks of haplotypes (Meuwissen et al. 2014; Cuyabano et al. 2015; Ma et al. 2016; Mayer et al. 2020; Vojgani et al. 2023), which are specific combinations of alleles observed in a population (Gabriel et al. 2002). As they are clusters of alleles, haplotypes are expected to be in greater linkage disequilibrium with the quantitative trait loci (QTL) of interest compared to individual single nucleotide polymorphisms (SNPs), making them more informative and, consequently, better predictors in GS (Zhang et al. 2002; Zondervan and Cardon 2004). Recent studies have confirmed the power and benefits of using haplotypes in different analyzes (Jiang et al. 2018; Haberer et al. 2019; Pook et al 2019; Vojgani et al 2023) and works such as those by Vojgani et al (2023) have been modifying and increasing the art of the field by confirming that the use of haplotypes can significantly reduce computational costs in genomic predictions and, consequently, contribute to a greater dissemination of the technology. Even though it is an emergent and revolutionary approach, clustering into haplotypes to improve reliability and/or reduce project costs is not fully widespread in vegetable's research, although it has already presented great results in some crops as tomato (Yamamoto et al 2017), cabbage (Liu et al. 2021), rapeseed (Hu et al. 2022) and even in sweet

corn (Baseggio et al 2020). Besides that, another selection strategy is the choice of markers evenly spaced across the genome, where the density to be used and the consequent expected coverage/distribution along the genome are predetermined (Ma et al. 2016). Even though these are simple strategies, they have never been used in the sweet corn crop aiming to find a minimum density of markers with a good prediction ability.

Based on the numerous positive results found in animal breeding, in other crops and the costly phenotyping that sweet corn presents, the pre-selection on markers and the creation of low-density panels for carrying out genomic selection presents as an excellent alternative to reduce computational time and costs in sweet corn breeding programs. In this study, the objectives were to test and compare different strategies of marker's pre-selection and to find a minimum density to predict yield traits in sweet corn.

2.2 MATERIAL AND METHODS

2.2.1 Germplasm

To perform the present study, 770 lines from the sweet corn breeding program of Syngenta Seeds LLC were used. These lines were developed from different populations and were separated into female and male heterotic groups to create 1244 single hybrids (F1) that also contributed with phenotypic data to perform the experiment.

2.2.2 Field Experiments & Phenotyping

The 1244 Hybrids (F1) were phenotyped in 2018, 2019, 2020 and 2021 in a total of 113 trials, mainly in Minnesota and Wisconsin states (U.S.), using augmented design. It is important to emphasize that not all hybrids were evaluated in every year, but common checks were used across years and locations. All treatments were automatically randomized, grown in a single-row plot (5.5 m, 0.53 m alleys and 0.76 m row width) and double seeded in order to achieve an appropriate population of 28

plants/plot (density of 66k plants/hectare). Due to a good germination, 20 days after sowing (DAS), exceeded plants were thinned and final stand count registered for later weight conversion.

Plots were harvested when achieved R3-R4 stages (milky), demanded by the processing market for consumption (Figure 1). When harvesting, ear diameter (ED), ear length (EL), husk length (HL), cut weight (CW) and green weight (GW) were collected in all hybrids. ED and EL were measured, in centimeter (cm), as an average of the diameter and length, respectively, of 5 ears, collected in the middle of the plot, with an accuracy of 0.1 cm, where the same ears were used to collect both traits. Furthermore, HL, that is the mensuration of the husk leaf extending above the tip of the ear (distance between ear and husk tips), was also measured (cm) as an average of 5 ears (Figure 2). In order to estimate CW and GW, plots were harvested, weighted (kg), processed (kernels separated from husk leaves and cobs) and weighted again (cut weight). In possession of the weights and final stand count, CW and GW were estimated in hectares.

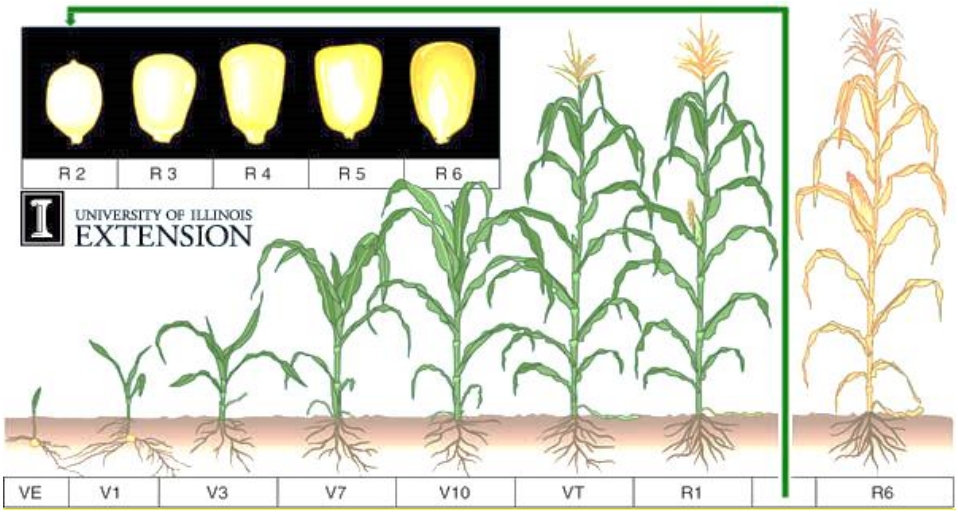


Figure 1. Sweet corn growth stages. University of Illinois extension, 2022.

Figure 2. Husk length measurement, Holambra, São Paulo, Brazil, 2020.

Phenotyped data were analyzed using SAS software (SAS Institute 2008) and best linear unbiased predictors (BLUPs) generated as:

$$y=Xb+Z_u+e\ (1)$$

where y is the vector of trial observations; X is the design matrix for fixed effects, such as years and treatments; b is the vector of fixed parameters, such as intercept and treatment effects; Z is the design matrix for random effects, such as genetic and environmental effects; *u* is the vector of random effects, which follows a multivariate normal distribution with mean zero and covariance G; e is the vector of random errors, which follows a normal distribution with mean zero and covariance R. The BLUPs obtained were used as response variables in the genomic selection model.

2.2.3 Genotyping and genomic selection

The inbreds were genotyped, in the United States, using a high-density Affymetrix Axiom® ~23K chip from 2018 to 2021. Since the quality of genotypic data is critical in genomic selection accuracy, a quality control (QC) was performed, where sites with more than 0.15 of missing data or sites with minor allele frequency (MAF) less than 0.05 or with a heterozygosity higher than 0.08 were recorded (during GS) and removed, culminating in a final chip with ~14K SNPs that was used in the genomic prediction studies.

The genomic best linear unbiased prediction (GBLUP) was used for training as below, according to the additive and dominance model (Zhao et al., 2021):

$$y = 1\mu + Z_1\mathrm{A} + Z_2\mathrm{D} + \mathrm{e}\ (1)$$

where y is an n × 1 vector of phenotypic data with n hybrids; 1 is an n × 1 vector of ones; μ is the overall mean; Z_1 is an n × n incidence matrix with additive genetic values; A are the additive genetic values following a normal distribution N(0, $GA_{(train\text{-}train)}\sigma A_2$) with 0 an n × 1 vector of zeros, $GA_{(train\text{-}train)}$ the additive GRM between any two hybrids in the training data set based on the Astle & Balding method (Astle and Balding 2009), and σA_2 the additive genetic variance; Z_2 is an n × n incidence matrix with dominance genetic values; D are the dominance genetic values following a normal distribution N(0, $GD_{(train\text{-}train)}\sigma D_2$) with 0 an n × 1 vector of zeros, $GD_{(train\text{-}train)}$ the

dominance GRM between any two hybrids in the training data set, also based on Astle balding method and σD_2 the dominance genetic variance and “e” are the residuals of the model. The Astle & Balding method was used in this study due to a higher reliability of the predictions when compared to other techniques, as identity-by-state (IBS) (Xia et al., 2015).

2.2.3.1 Prediction Process

In possession of the GBLUPs, training and prediction were done in a single step. Given the population mean and variance components estimated in model (1), each hybrid was predicted, based on their genomic estimated breeding value (GEBV) and GRM as:

$$\hat{y} = G_{A(tested\text{-}trained)} \times G_{A(trained\text{-}trained)}^{-1} \times GEBV_{A(trained)} + G_{D(tested\text{-}trained)} \times G_{D(trained\text{-}trained)}^{-1} \times GEBV_{D(trained)}$$

where $G_{A(tested\text{-}trained)}$ is the additive GRM between tested individuals and individuals in the training population; $G_{A(trained\text{-}trained)}^{-1}$ is the inverse of the additive GRM between the individuals in the training population; $GEBV_{A(trained)}$ is the additive GEBV of the training individuals; $G_{D(tested\text{-}trained)}$ is the dominance GRM between tested individuals and individuals in the training population; $G_{D(trained\text{-}trained)}^{-1}$ is the inverse of the dominance GRM between the individuals in the training population and $GEBV_{D(trained)}$ is the dominance GEBV of the training individuals.

2.2.3.2 Cross-Validation and prediction accuracy calculation

The cross-validation method used in our experiment was an x-fold design that randomly assigns individuals to folds. Each of the x-folds was used in turn as a validation sample and the remaining folds were combined to build the training set. In order to perform our study, a 5-fold method was implemented, where the dataset was divided into 5 equal parts. The model was trained on 4 of the folds and tested on the remaining one. This process was repeated 5 times, with each fold serving as the test set once. The results were then averaged to obtain an estimate of the model's performance.

2.2.4 Marker selection

Genomic selection was done, initially, with the full marker set following the strategy described above. Additionally, to reasonably reduce the density of SNPs and test the new prediction abilities, two methodologies were used, where the first one was based on linkage disequilibrium (LD) and haplotypes block estimation and the second one was based on even distribution across the genome.

The first methodology used to select markers and reduce the initial set was selecting markers evenly across the genome, where we choose evenly spaced SNPs for subsets of 0.05 Megabase (Mb), 0.09Mb, 0.1Mb, 0.2Mb, 0.3Mb, 0.4Mb, 0.5Mb, 1.0Mb, 1.5Mb, 2.0Mb, 2.5Mb, 5.0Mb, 5.5Mb, 7.0Mb, 7.5Mb and 10.0Mb and ran a genomic selection for each group (Figure 3) A Megabase refers to a unit of genetic information that contains approximately one million DNA base pairs. Furthermore, it is important to emphasize that the extent of LD was estimated for all pairs of SNPs on the same chromosome.

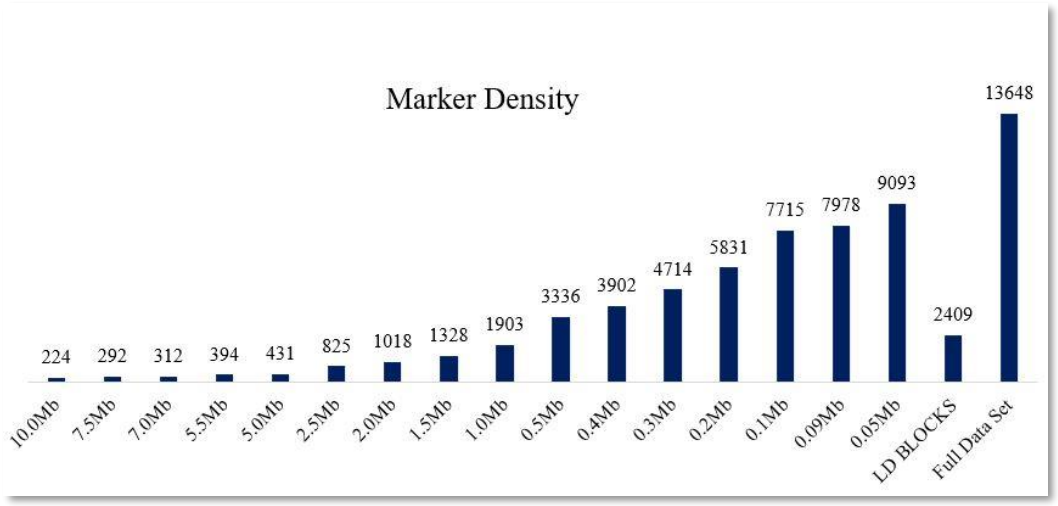


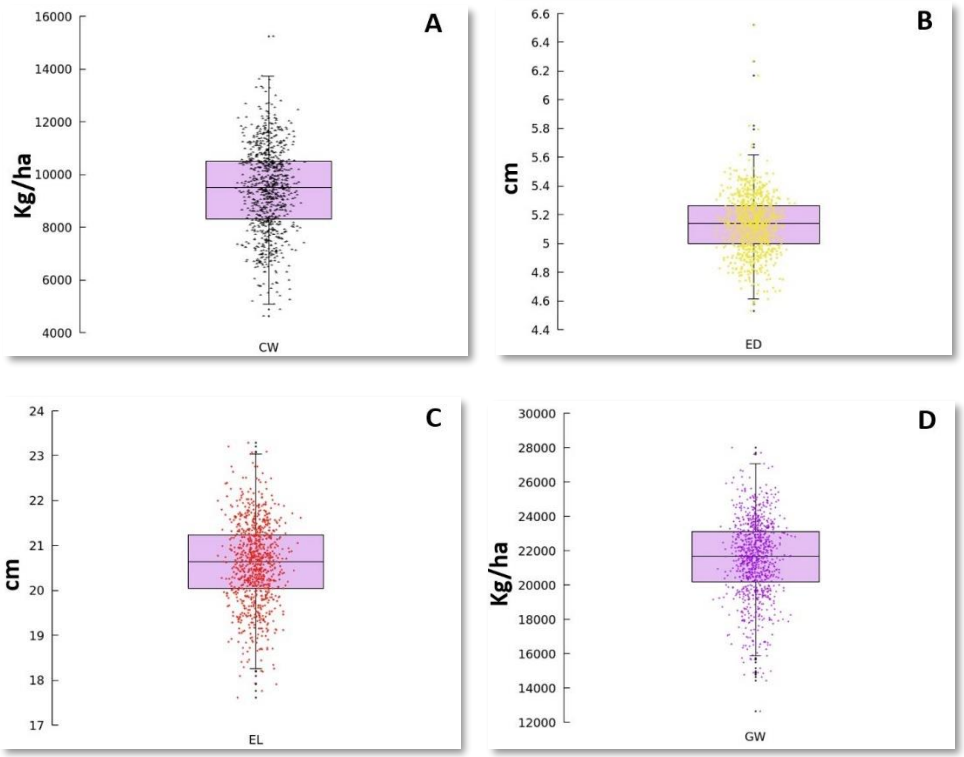
Figure 3. Different marker density strategies and their respective dimensions to predict production traits in sweet corn genotypes, based on additive plus dominance effects.

Having the coordinates of each marker from the full set and aiming to create haplotype blocks, the program PLINK v. 1.9 (Chang et al. 2015) was chosen to reduce the initial marker set, considering a LD window of 200kb. All markers with minor allele frequency (MAF) less than 0.05 were ignored during the process. This new version of

PLINK improved and made faster robust jobs as: identify-by-state matrix computation, distance-based clustering, LD-base pruning, haplotype block identification. This process resulted in a data panel containing 7440 SNPs that were grouped into 2409 haplotype blocks using the software Galaxy (Basten 2022). With the blocks formed and having the MAF and heterozygosity for all markers, an anchor was selected for each block in order to be, posteriorly, used in a new genomic selection (Figure 3).

2.3 RESULTS

After the field phenotypic evaluations, focusing on yield traits, the sweet corn genotypes presented, for ED, BLUPs ranging from 4.5 to 6.5 cm, with an average of 5.13cm; for ear length (EL), values ranged from 17.6 to 23.3 cm, with an average of 20.59 cm; for husk length (HL), BLUPs ranged from -1.6 to 6.2 cm, with an average of 1.83; regarding cut weight (CW), values ranged from 4600 to 15200 kg/ha, with an average of 9440 kg/ha and for green weight (GW), BLUPs ranged from 12700 to 28000 kg/ha, with an average of 21480 kg/ha. Results and can be seen in figure 4.



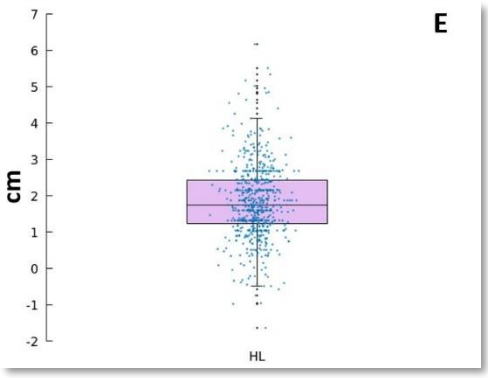


Figure 4. Prediction mean (BLUPs) for yield traits, as cut weight (CW) (A), ear diameter (ED) (B), ear length (EL) (C), green weight (GW) (D) and husk length (HL) (E), in sweet corn genotypes, phenotyped in 113 trials, between 2018 and 2021.

Having a marker set with high quality and phenotypic BLUPs, a genomic selection (GS) was performed considering, initially, a full density marker set (~14k SNPs). Furthermore, 100 replications were done for cross-validation and accuracy estimation.

During the process, GBLUPs were generated and used to obtain the heritability (h^2) of the traits (narrow), that can be seen in Table 1. Ear diameter and ear length showed the highest heritabilities, which were 0.9 and 0.73, respectively. On the other hand, green and cut weight presented the lowest heritabilities, being 0.4 and 0.56, respectively. Furthermore, on the same table, it is possible to verify the variance components (additive and dominance).

Table 1. Genomic selection outputs: variance components, heritability in a narrow sense and phenotypic adjusted means for cut weight (CW), ear diameter (ED), ear length (EL), green weight (GW) and husk length (HL).

GS Output resources	CW (V)	ED(VI)	EL (V)	GW (V)	HL (V)
Additive	748172.90	0.07	0.62	903282.75	0.68
Dominance	542791.07	0.01	0.14	896125.8	0.21
Residual	1029755.72	0.01	0.28	2736923.15	0.39
h^2	0.56	0.90	0.73	0.40	0.69
Phenotypic adjusted mean	9300 kg/ha	5.1 cm	20.54 cm	21300 kg/ha	1.86 cm

Moreover, the prediction ability (mean) of the yield traits, that can be seen on Figure 5, showed a quite similarity between them and a higher accuracy to predict CW (0.65), EL (0.66) and HL (0.63). Even though ED and GW were predicted with a lower precision than the others, having an accuracy of 0.59 and 0.57, respectively, their ability was quite high.

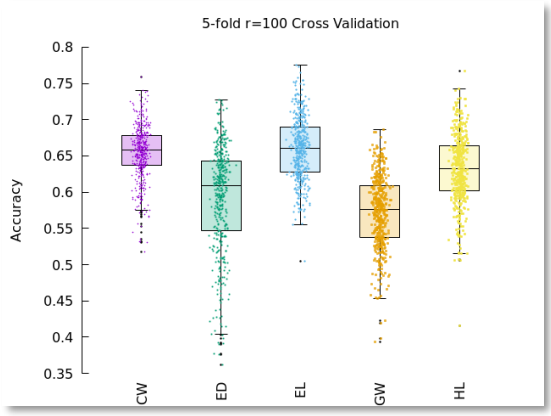


Figure 5. Comparison of prediction accuracies (mean), based on additive plus dominance effects, in sweet corn genotypes, using a high-density Axiom® ~14K SNP panel, for yield traits as ear diameter (ED), husk length (HL), ear length (EL), cut weight (CW) and green weigh (GW).

In order to reduce the initial marker set, PLINK (Chang et al. 2015) was used to generate 2409 haplotype blocks, considering the linkage distance between SNPs, across the 10 sweet corn chromosomes. The number of SNPs composing haplotype blocks ranged from 2 to 17 and their distribution across all chromosomes was quite homogeneous, culminating in haplotypes well distributed too. The chromosome with lowest coverage (10) presented 156 haplotypes and the one with highest coverage (1) had 349 haplotypes, with an average of 249 blocks per chromosome. Once the haplotypes were assembled, a new genomic selection was carried out, using them instead of the initial set and the result obtained was quite similar to that found initially (Figure 6). For all traits, the prediction accuracy (mean) was very similar to that found using the initial marker set (high density). CW and EL were predicted with an accuracy of 0.65; for ED it was 0.58; GW and HL were predicted with an accuracy of 0.57 and 0.63, respectively.

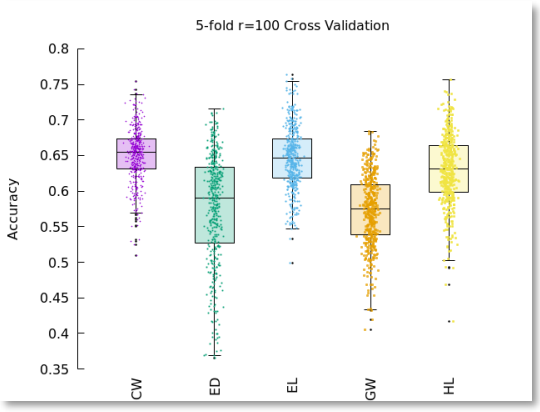


Figure 6. Comparison of prediction accuracies (mean), based on additive plus dominance effects, in sweet corn genotypes, using pre-selected SNPs, based on haplotype clustering, for yield traits as cut weight (CW), ear diameter (ED), ear length (EL), green weight (GW) and husk length (HL).

The second approach to select SNPs for genomic selection was evenly across the genome, where we predeterminate marker densities of 0.05Mb, 0.09Mb, 0.1Mb, 0.2Mb, 0.3Mb, 0.4Mb, 0.5Mb, 1.0Mb, 1.5Mb, 2.0Mb, 2.5Mb, 5.0Mb, 5.5Mb, 7.0Mb, 7.5Mb, 10.0Mb and used the selected markers (Figure 3) to run new predictions.

The results (means) and comparisons can be seen in Figure 7. Regarding ED and EL, the prediction ability achieved with the initial marker set (high density) was very similar to what was found using 0.1Mb, 0.09Mb and 0.05Mb, having a slightly decrease when using the haplotypes and 0.2Mb for prediction. For HL, as mentioned above, the accuracy using blocks of haplotypes was quite similar to that found using the complete panel (0.63), which, by the way, were also similar to the accuracies found using 0.1Mb, 0.09Mb and 0.05Mb. The same concept prevailed when comparing the accuracies of CW and GW traits, finding 0.65 and 0.57, respectively.

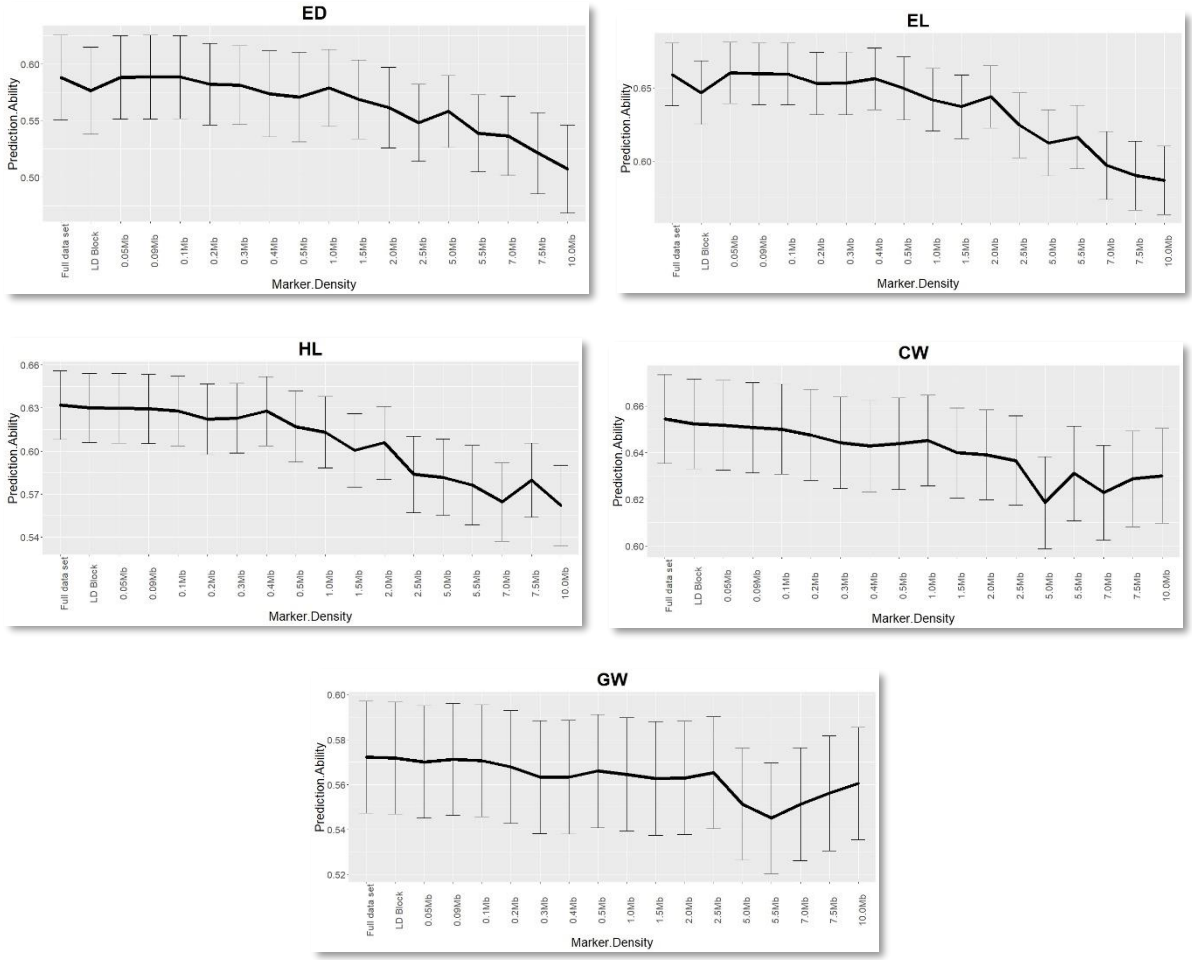


Figure 7. Different marker density strategies and their respective accuracies (mean) to predict yield traits, as ear diameter (ED), husk length (HL), ear length (EL), cut weight (CW) and green weight (GW), in sweet corn genotypes, based on additive plus dominance effects.

2.4 DISCUSSION

A QC was performed on the initial Axiom® chip to improve the study reliability and to reduce the risk of bias caused by poor quality genotyping data, as subtle differences in DNA quality among individuals can pollute true hits considerably (Weale 2010). This strategy has already been successfully implemented by other authors (Zhang et al., 2015; Crossa et al. 2017; Zhang et al., 2019), aiming to remove low-quality samples, to check for population stratification and to access the accuracy of imputation.

Heritability is the proportion of the total phenotypic variance that is attributable to the effect of genes, that in turn determines the degree of resemblance between relatives (Falconer and Mackay, 1996). Furthermore, it investigates the relationship between phenotypic and genotypic values (Holland et al., 2002) or the percentage of the total variation that is due to genetic factors and, thus, indicates the proportion of the variance that will be transmitted to the following progenies. On the present study, ear diameter (ED) and ear length (EL) highlighted, showing the highest heritabilities (0.9 and 0.73, respectively) that were also superior to what was found by Abe and Adelegan (2019) (0.67 and 0.66, respectively), when evaluating the heritability in shrunken-2 sweet corn populations. This difference can be explained by the use of different methods, since in the present study GBLUPs, that can increase the accuracy of predicting the genetic value of an individual, particularly for traits that are difficult or expensive to measure directly (Goddard et al., 2010), were applied to estimate the heritabilities. Abe and Adelegan (2019) also evaluated husk length (HL), achieving a heritability of 0.78, higher than the current study, but being explained by the methodology applied to collect this trait as the authors did not measure the husk leaves, but graded them according to their scale. In general, heritabilities were quite acceptable, rating from 0.4 to 0.9, being lower for traits controlled by genes with minor effects, as cut and green weight (CW, GW).

For all traits, the prediction abilities achieved were higher than 0.53, indicating a greater annual genetic gain compared to pedigree selection (PS) in maize (Heffner et

al. 2010). For ED, the accuracy achieved in the present study was similar to what was found by Endelman (2011), using different methods, as gaussian (0.53), exponential (0.54), ridge regression (0.52) and Bayesian (0.53) models. Regarding HL, Cui et al. (2020), when working to understand the potential of GS to improve husk traits in maize subpopulations, found an accuracy close to 0.5 in tropical-subtropical genotypes and in mixed populations, which was lower than the result found in the present study. Similarly, EL was predicted, in the present study, with a higher accuracy when compared to the experiment carried out by Liu et al. (2019), who evaluated the effect of trait-relevant markers (TRMs) in maize genotypes (~0.45). On the other hand, Zystro et al. (2021), achieved accuracies ranging from 0.38 to 0.9 when predicting ear length, in sweet corn genotypes, basing on additive and additive plus dominance effects, respectively, indicating that the difference can be explained by the structure of population, by the type of endosperm, by the predicted effects and also due to the quality of markers.

Cut weight (CW) prediction ability was 0.65 and even though there are not enough studies regarding the prediction ability of this trait in sweet corn, similar approaches were used in maize. Zhao et al. (2012), in a genomic selection study for European maize elite breeding populations, achieved an accuracy of 0.58 for grain yield and concluded that this average of accuracy corresponds to the precision of phenotyping without replications in 3-4 locations, which culminates in a higher selection gain per unit time and cost saving. Furthermore, Beyene et al. (2015), working with sub-Saharan Africa genotypes, demonstrated that the overall gain in average grain yield, using GS, was twice to four times higher than the reported gain, under drought stress, using conventional phenotypic selection.

Grouping markers into haplotypes, based on LD, is not an unknown strategy and has already been implemented by other authors in different crops. This approach was utilized by Ma et al. (2016), when evaluating the potential of a pre-marker selection to increase prediction accuracy in soybean. On that occasion, the authors were able to distribute their haplotypes across all soybean chromosomes, with a variation of 1.3% to 67.3%, and reached the conclusion that clustering SNPs is a cost-efficient implementation for genomic selection. In the present study and using an anchor for each of the 2409 haplotypes, a new GS was performed and for all traits there was no

significant difference between initial and selection sets in accuracy, corroborating with Ma et al. (2016) and reinforcing the idea that one marker per LD block would be sufficient for GS, especially in a crop like sweetcorn that has a narrow genetic basis. Therefore, the cost of genotyping will be reduced significantly without compromising the prediction ability. According to the estimation of the cost of HD genotyping ($15) and the cost of LD genotyping ($12) provided by Kriaridou et al. (2023), we can obtain a reduction of 20% in genotyping costs when working with 700 individuals, demonstrating the financial importance that marker's pre-selection has in a genomic selection project.

The second approach to previously select SNPs for genomic selection was evenly across the genome, where we predeterminate marker densities and used the selected markers to run new predictions. Regarding the length (EL) and diameter (ED) of the ear, important for the consumer market and consequently for breeding programs, results indicated that they can be predicted using only ~8k quality markers. For HL, as mentioned above, the accuracy using blocks of haplotypes was quite similar to that found using the complete panel and also similar to that found using 0.1Mb, 0.09Mb and 0.05Mb, indicating that this trait can be predicted with a smaller number of markers and with different pre-selection strategies. The same concept prevailed when comparing the accuracies of CW and GW.

In general, all traits showed a significant reduction, in prediction accuracy, as the number of SNPs decreased (Figure 8), demonstrating the impact that marker density has on prediction accuracy (Hayes et al. 2009; Heffner et al. 2009; Desta and Ortiz 2014). On the other hand, the relationship between marker's increment and accuracy did not remain linear and reached a plateau after a certain point, for all traits, corroborating with Heffner et al. (2011) and de Los Campos et al. (2013).

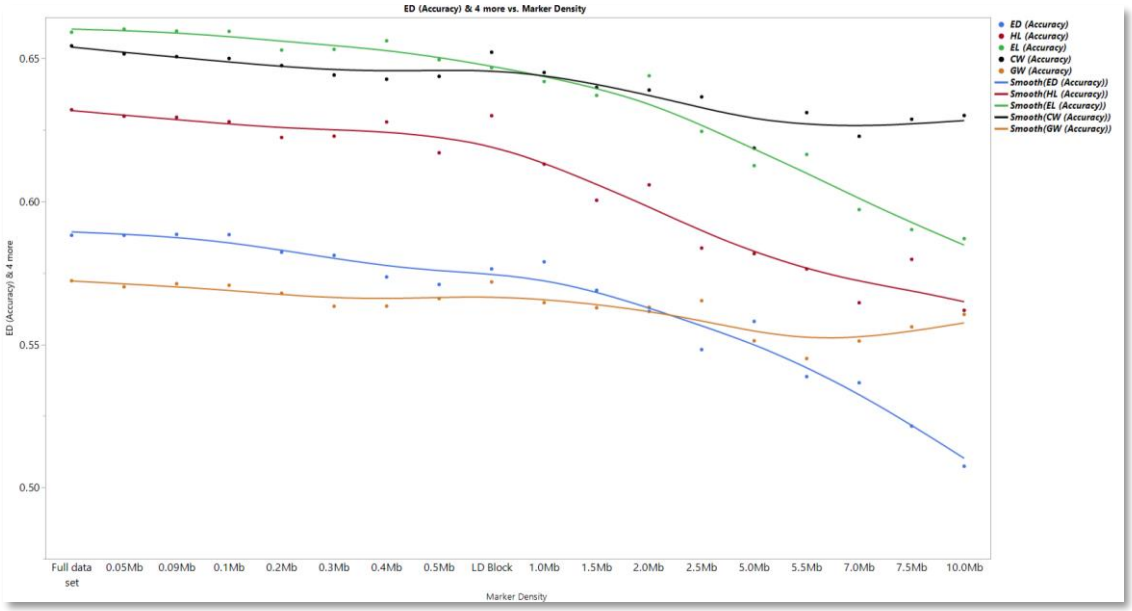


Figure 8. Different marker density strategies and their respective accuracies, grouped, to predict yield traits, as ear diameter (ED), husk length (HL), ear length (EL), cut weight (CW) and green weight (GW), in sweet corn genotypes, based on additive plus dominance effects.

2.5 CONCLUSIONS

Applying marker pre-selection based on haplotype blocks and evenly across the genome are good options for a cost-efficient implementation of genomic selection in sweet corn genotypes for yield traits, with desired prediction abilities.

Yield traits, in sweet corn genotypes, can be predicted using a panel of ~8K quality markers, achieving the same prediction ability as using the full marker set (~14K).

In general and considering the number of final markers used in the predictions, clustering into haplotypes appears to be a better strategy for the pre-selection of markers before carrying out a genomic selection in sweet corn, for yield traits, as similar prediction abilities were achieved as using the full marker set and cost is significantly reduced.

2.6 REFERENCES

Abe A, Adelegan C A (2019) Genetic variability, heritability and genetic advance in shrunken-2 super-sweet corn (Zea mays L. saccharata) populations. **Journal of Plant Breeding and Crop Science**, v.11,n.4, p.100-105. Accessed on August 23rd, 2023.

Astle W, Balding DJ (2009) Population Structure and Cryptic Relatedness in Genetic Association Studies. **Statistical Science**, v.24, n.4, p.451-471. Accessed on October 23rd, 2022.

Barbieri VHB (2010) **Mapeamento de QTL em testecrosses de milho doce com diferentes testadores e ambientes**. Tese (Doutorado em Genética e Melhoramento de Plantas) - Escola Superior de Agricultura Luiz de Queiroz, University of São Paulo, Piracicaba. Accessed on July 2nd, 2020.

Beyene Y, Semagn K, Mugo S, Tarekegne A, Babu R, Meisel B, Sehabiague P, Makumbi D, Magorokosho C, Oikeh S, Gakunga J, Vargas M, Olsen M, Prasanna BM, Banziger M, Crossa J (2015) Genetic gains in grain yield through genomic selection in eight bi-parental maize populations under drought stress. **Crop Science**, v.55, p.154-163. Accessed on March 4th, 2023.

Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. **GigaScience**, v.4. Accessed on September 22th, 2022.

Crossa J, Pérez-Rodríguez P, Cuevas J, Montesinos-López O, Jarquín D, de Los Campos G, Burgueño J (2017) Genomic selection in plant breeding: methods, models, and perspectives. **Trends in Plant Science**, v.22, n.11. Accessed on May 09th, 2023.

Cui Z, Dong H, Zhang A, Ruan Y, He Y, Zhang Z (2020) Assessment of the Potential for Genomic Selection To Improve Husk Traits in Maize. **G3 Genes|Genomes|Genetics**, V.10, p.3741–3749. Accessed on March 04th, 2023.

Cuyabano BC, Su G, Lund MS (2015) Selection of haplotype variables from a high-density marker map for genomic prediction. **Genetics Selection Evolution**, v.47, p.61. Accessed on October 23th, 2022.

de Los Campos G, Hickey JM, Pong-Wong R, Daetwyler HD, Calus M MPL (2013) Whole-genome regression and prediction methods applied to plant and animal breeding. **Genetics**, v.193, p.327–345. Accessed on August 15th, 2022.

Desta ZA, Ortiz R (2014) Genomic selection: genome-wide prediction in plant improvement. **Trends in Plant Science**, v.19, p.592–601. Accessed on August 15th, 2022.

Dodson-Swenson HG, Tracy WF (2015) Endosperm carbohydrate composition and kernel characteristics of shrunken2- intermediate (sh2-i/sh2-i Su1/Su1) and shrunken2-intermediate-sugary1-reference (sh2-i/sh2-i su1-r/su1-r) in sweet corn. **Crop Science**, Madison, v.55, p.2647-2656. Accessed on July 3rd, 2020.

Empresa Brasileira de Pesquisa e Agropecuária **EMBRAPA** (2016) O cultivo do Milho Doce. EMBRAPA Milho e Sorgo, Brasília, DF. p.289. Accessed on July 3rd, 2020.

Endelman JB (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. **The Plant Genome**, v.4, p.250-255. Accessed on July 2nd, 2020.

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. **Longman Scientific and Technical**, 4th ed., 464p. Accessed on July 10th, 2022.

Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D (2002) The structureof haplotype blocks in the human genome. **Science**, v.296, n.5576, p.2225–2229. Accessed on October 24th, 2022.

Global Market Insights (2023) Sweet corn seeds maerket by seed type (Hybrid certified seeds, farm saved seeds, open pollinated certified seeds), by seed category (yellow, white, bicolor), by farming type (organic, conventional), by sales channel & by region. **Fact.MR, Food and Beverage**, p.170. Accessed on July 25th, 2023.

Goddard ME, Hayes BJ, Meuwissen THE (2010) Genomic selection in livestock breeding. **Genetics Research**, v.92, n.5-6, p.413-421. Accessed on July 25th, 2023.

Habier D, Fernando R, Dekkers J (2007) Genomic Selection Using Low-Density Marker Panels. **Genetics**, v.177, n.4, p.2389–97. Accessed on August 14th, 2022.

Habier D, Fernando R, Dekkers J (2009) The impact of genetic relationship information on genome-assisted breeding values. **Genetics**, v.182, n.1, p.343–353. Accessed on September 17th, 2022.

Hayes B, Bowman P, Chamberlain A, Goddard M (2009) Invited review: Genomic selection in dairy cattle: Progress and challenges. **Journal of dairy science**, v.92, n.2, p.433–443. Accessed on August 14th, 2022.

Heffner EL, Sorrells ME, Jannink JC (2009) Genomic selection for cropimprovement. **Crop Science**, v.49, p.1–12. Accessed on August 15th, 2022.

Heffner EL, Lorenz AJ, Jannink JC, Sorrels ME (2010) Plant Breeding with genomic selection: gain per unit time and cost. **Crop Science**, v.50, p.1681-1690. Accessed on February 28th, 2023.

Heffner EL, Jannink JC , Sorrells ME (2011) Genomic selection accuracy using multifamily prediction models in a wheat breeding program. **Plant Genome**, v.4, p.65–75. Accessed on August 15th, 2022.

Holland JB, Nyquist WE, Cervantes-Martínez CT (2002) Estimating and Interpreting Heritability for Plant Breeding: An Update. In: **Plant Breeding Reviews**, p.9-112. Accessed on August 24th, 2021.

Kriaridou C, Tsairidou S, Fraslin C, Gorjanc G, Looseley ME, Johnston IA, Houston RD, Robledo D (2023) Evaluation of low-density SNP panels and imputation for cost-effective genomic selection in four aquaculture species. **Frontiers in Genetics**, v.14. Accessed on July 23th, 2023.

Lee D, Kim Y, Chung Y, Lee D, Seo D, Choi TJ, Lim D, Yoon D, Lee SH (2021) Accuracy of genotype imputation based on reference population size and marker density in Hanwoo cattle. **Journal of Animal Science and Technology**, v.63, n.6,

p.1232–1246. Accessed on September 18th, 2022.

Liu X, Wang H, Hu X, Li K, Liu Z, Wu Y, Huang C (2019) Improving Genomic Selection With Quantitative Trait Loci and Nonadditive Effects Revealed by Empirical Evidence in Maize. **Frontiers in Plant Science**, v.10, p.1129. Accessed on February 28th, 2023.

Ma Y, Reif JC, Jiang Y, Wen Z, Wang D, Liu Z, Guo Y, Wei S, Wang S, Yang C, Wang H, Yang C, Lu W, Xu R, Zhou R, Wang R, Sun Z, Chen H, Zhang W, Wu J, Hu G, Liu C, Luan X, Fu Y, Guo T, Han T, Zhang M, Sun B, Zhang L, Chen W, Wu C, Sun S, Yuan B, Zhou X, Han D, Yan H, Li W, Qiu L (2016) Potential of marker selection to increase prediction accuracy of genomic selection in soybean (Glycine max L.). **Molecular Breeding**, v.36, n.113. Accessed on September 24th, 2022.

MAPA - MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO (2023) Registro Nacional de Cultivares. Accessed on May 2nd, 2023.

SAS Institute (2008) Statistical analysis software: **SAS 9.2**. Accessed on July 10th, 2022.

Solberg T, Sonesson A, Woolliams J (2008) Genomic selection using different marker types and densities. **Journal of animal science**, v.86, n.10, p.2447–2454. Accessed on August 14th, 2022.

Sousa MB, Galli G, Lyra DH, Granato ISC, Matias FI, Alves FC, Fritsche-Neto R (2019) Increasing accuracy and reducing costs of genomic prediction by marker selection. **Euphytica**, v.215, n.18. Accessed on September 24th, 2022.

Tracy WF (2001) Sweet Corn. In: Hallauer AR (Ed.) **Specialty corn. 2nd edition**. Boca Raton, CRC Press. p.155-198. Accessed on March 15th, 2020.

Weale ME (2010) Quality Control for Genome-Wide Association Studies. In: Barnes, M., Breen, G. (eds) Genetic Variation. **Methods in Molecular Biology**, v.628. Humana Press, Totowa, NJ. Accessed on February 18th, 2023.

Wellmann R, Preuß S, Tholen E, Heinkel J, Wimmers K, Bennewitz J (2013) Genomic selection using low density marker panels with application to a sire line in pigs. **Genetics Selection Evolution**, v.45, n.28. Accessed on September 17th, 2022.

Werner CR, Voss-Fels KP, Miller CN, Qian W, Hua W, Guan CY, Snowdon RJ, Qian L (2018) Effective Genomic Selection in a Narrow-Genepool Crop with Low-Density Markers: Asian Rapeseed as an Example. **The Plant Genome**, v.11, n.2. Accessed on September 18th, 2022.

Zhang Z, Erbe M, He J, Ober U, Gao N, Zhang H, Simianer H, Li J (2015) Accuracy of whole-genome prediction using a genetic architecture-enhanced variance-covariance matrix. **G3 Genes, Genomes, Genetics**, v.5, p.615–627. Accessed on December 4th, 2022.

Zhang H, Yin L, Wang M, Yuan X, Liu X (2019) Factors Affecting the Accuracy of Genomic Selection for Agricultural Economic Traits in Maize, Cattle, and Pig Populations. **Frontiers in Genetics**, v.10, p.189. Accessed on December 4th, 2022.

Zhao Y, Gowda M, Liu W, Würschum T, Maurer HP, Longin FH, Ranc N, Reif JC (2012) Accuracy of genomic selection in European maize elite breeding populations.

Theoretical and Applied Genetics, v.124, p.769-776. Accessed in March 4th, 2023.

Zhao Y, Thorwarth P, Jiang Y, Philipp N, Schulthess AW, Gils M, Boeven PHG, Longin CFH, Schacht J, Ebmeyer E, Korzun V, Mirdita V, Dörnte J, Avenhaus U, Horbach R, Cöster H, Holzapfel J, Ramgraber L, Künle S, Varenne P, Starke A, Schürmann F, Beier S, Scholz U, Liu F, Schmidt RH, Reif JC (2021) Unlocking big data doubled the accuracy in predicting the grain yield in hybrid wheat. **Sciences Advances**, v.7, n.24. Accessed on December 4th, 2022.

Zhong S, Dekkers JC, Fernando RL, Jannink J- L (2009) Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a barley case study. **Genetics**, v.182, n.1, p.355–64. Accessed on August 14th, 2022.

Zondervan KT, Cardon LR (2004) The comples interplay among factors that influence allelic association. **Nature Reviews Genetics**, v. 5, p.89–101. Accessed on October 24th, 2022.

Zystro J, Peters T, Miller K, Tracy WF (2021) Classical and genomic prediction of hybrid sweet corn performance in organic environments. **Crop Science**, v.61, p.1698-1708. Accessed on March 4th, 2023.