Doubled haploid breeding methods in maize and soybean

by

Jarrod Friederich

A creative component submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Agronomy

Program of Study Committee:
Dr. Lübberstedt, Major Professor
Dr. Lenssen
Dr. Knapp

Iowa State University

Ames, Iowa

2020

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Lübberstedt, and my committee members, Dr. Lenssen, and Dr. Knapp, for their guidance and support throughout the course of this research.

I would also like to thank the staff and faculty at Iowa State University for making my experience in this program one of the most rewarding. Additionally, I would also like to thank the distance agronomy program administrative staff for their support and guidance in accessing and navigating course materials without their dedication and support distance graduate programs such as this would not be possible.

LIST OF ACRONYMS

APM: Amiprophos methyl

CMS**:** Cytoplasmic male sterility

CRISPR: (clustered regularly interspaced short palindromic repeats)

DH: Doubled haploid

DHL: DH line

MATL: Matrilineal

QTL: quantitative trait loci

LIST OF FIGURES AND TABLES

ABSTRACT

The use of haploids in breeding programs has many benefits and has been a useful tool in plant breeding programs for decades. Since the initial discovery of haploids in the 1920's, haploid technology has allowed for an increased rate of genetic gain in crops such as maize by allowing efficient integration of desirable agronomic, phenotypic, and genomic traits into elite germplasm quickly and efficiently. The objectives of this creative component were to review the use of doubled haploid (DH) breeding technology in established plant breeding programs such as maize (*Zea mays* L.) and the genes currently known to be involved in DH induction. The second objective was to discuss the requirements necessary to establish a viable DH breeding program in soybean [*Glycine max* (L.) Merr].

Due to the highly successful use of doubled haploid (DH) technology in maize DHs have been researched and developed for use in several crops with varying rates of success. Wheat has been among the more successful recent adoptions of commercial use of DHs. Rice (*Oryza sativa* L.) and soybean have been met with limited results. However, there has been promising research conducted recently that may help make this technology more commercially viable for these crops.

Introduction

The use of haploids in breeding programs has many benefits and has been a useful tool in plant breeding programs for decades (Liu et al. 2016). Doubled haploid technology has allowed for an increased rate of genetic gain in crops such as maize (*Zea mays* L.) by allowing efficient integration of desirable agronomic traits into elite germplasm. Haploids in maize were first discovered in the 1920's (Randolph and Fischer 1939). Haploid kernels occur naturally at very low rates, approximately 1 in 100,000, in native populations (Chase 1949). Early research was conducted on using haploids for breeding purposes (Chase 1949), but due to low naturally occurring haploid induction rates, unreliable kernel identification methods, and inefficient genome doubling capabilities, DH technology was not widely used in maize breeding until the 1990s. A major breakthrough was the discovery of the Stock 6 haploid inducer (Coe 1959).

The Stock 6 inducer line reported in 1959 (Coe 1959) had a haploid induction rate of 3%, which was significantly higher than naturally occurring induction rates (Liu et al. 2016). This higher induction rate along with more efficient methods of genome doubling enabled DH breeding methods to be used more extensively, and initiated development of more efficient inducer lines (Liu et al. 2016). The increase in efficiency allowed DH breeding to be used reliably on a commercial scale (Liu et al. 2016).

The use of DH breeding systems allows breeders to produce homozygous inbred lines in two generations rather than six to eight generations (Liu et al. 2016). This method is very popular in current inbred line development programs due to the increase in efficiency in producing elite inbred lines that are homozygous for a desired trait (Liu et al. 2016). Inbred lines are developed and commercialized to be used as varieties (in case of line breeding such as for barley (*Hordeum vulgare* L.) and wheat (*Triticum aestivum* L.), or as parents of hybrid varieties (such as in maize).

The first widely planted commercial hybrid developed from a DH program was Dekalb 640 developed in the 1950's from a double cross of three DH and one conventional inbred parent lines (Liu et al. 2016). Dekalb 640 was very popular in the eastern U.S. and in Europe. Its tremendous success helped pave the way for widespread use of DHs for inbred line development. The main reason for using haploid technology is to reduce the time it takes to produce inbred lines, saving costs associated with plot numbers, number of generations grown, pollination numbers, laboratory analyses, and time needed for registration by allowing the regulatory process to be started much sooner (Liu et al. 2016). The faster regulatory approval process allows a new variety to be approved for market use faster, thus generating a revenue stream in a much shorter time frame.

The objectives of this creative component were to review literature on doubled haploid seed production in maize and other important agronomic crops, including a survey of genes currently known to be involved in doubled haploid seed production. The 2nd main objective was to evaluate the feasibility of developing a doubled haploid breeding system for soybeans [*Glycine max* (L.) Merr.].

Chapter 1

In vivo haploid induction

In vivo haploid induction uses interspecific or intraspecific haploid inducers. Interspecific induction is when a crop species such as wheat is pollinated with a different species such as maize to induce haploids (Wu et al. 2017). In this case, wheat plants are pollinated with maize pollen (Wu et al. 2017). Upon successful pollination the maize genome is then eliminated through a process called centromere mediated chromosome elimination (Wu et al. 2017). The

removal of the chromosomes allows for successful development of haploid wheat plants. Doubled haploid wheat lines produced in this way will only contain the wheat genome (Wu et al. 2017). Interspecific induction with maize is also used in rice, oats (*Avena sativa* L.) and rye (*Secale cereale* L.) (Wu et al. 2017).

Table 1.1 Interspecific Induction and Inducers.

Crop species	Induction type	Crop used as Inducer
Wheat	Interspecific	Maize
Rice	Interspecific	Maize
Rye	Interspecific	Maize
Oats	Interspecific	Maize

Table 1.2 Intraspecific induction and Paternal vs. Maternal genes.

Crop	Induction type	Inducer method	Genes Involved
Maize	Intraspecific	Paternal	Ig1, CENH3
Maize	Intraspecific	Maternal	MATL, ZmDMP, MTL genes
Rice	Intraspecific	Maternal	OsMATL, MATL

Intraspecies induction is used in maize breeding programs due to its successful track record and relatively high percentage of successfully induced haploid kernels. This method involves using an inducer line from the same crop species to produce haploid seed (Liu et al. 2016). The inducer line can be either a male line (maternal haploid induction) or a female line (paternal haploid induction) that have a genetic trait that allows it to induce the development of haploid kernels in the donor line at a much higher rate than would occur spontaneously. This method has also been used in rice and attempted soybeans.

Haploids are produced in maize mainly by in vivo haploid induction. In vivo induction occurs either maternally or paternally (Liu et al. 2016). The parent plant with desired agronomic traits is the donor (Liu et al. 2016). The inducer line is a specific line either male or female that

promotes a high rate of haploid induction. Most modern inducer lines can be traced back to Stock 6 (Liu et al. 2016). Stock 6 is an inducer line developed in the 1950's and has helped make modern DH programs in maize possible (Liu et al. 2016). It has a haploid induction rate of 3% in maternal lines (Zhong et al. 2019). In paternal haploid induction the donor is the male and the inducer line is the female. The resulting DHs inherit the cytoplasm of the inducer line and the chromosomes are from the male donor (Liu et al. 2016).

Maternal Induction

In maternal haploid induction the donor is the female and the inducer is the male (Liu et al. 2016). The resulting DH lines inherit both the cytoplasm and chromosomes from the female donor (Liu et al. 2016). Maternal induction is the most common form of haploid induction used in maize breeding programs (Liu et al. 2016). Maternal haploid induction is the preferred method in DH breeding programs today due to the higher haploid induction frequency (10%) of maternally derived DH lines.

The seed resulting from haploid induction crosses must be screened for haploids, as the majority of the seed will be regular diploid seed and only a fraction of the seed is haploid. Screening can be accomplished by several methods such as using color marker genes, phenotypic sorting, analyzing oil content, or seed weight. The most commonly used method for identifying haploid seeds is the R1-nj gene (Liu et al. 2016). R1-nj gene is a dominant color marker gene that can be incorporated into an inducer line. If the resulting kernel has been successfully induced and is haploid, the kernel crown will appear red/purple, while the embryo will be uncolored (Liu et al. 2016). Diploid kernels will have a purple kernel crown and embryo coloration, because they received the genome from the inducer carrying the dominant R1-nj

gene, in contrast to haploid kernels. Kernels that are either outcrosses or contaminated with pollen not originating from the inducer will not be colored (Geiger et al. 2009).

There are several genes of interest that have been identified in different species as being important to in vivo induction systems. In maize the R1-nj gene is used as a color marker to assist with identifying haploid kernels more efficiently and accurately (Liu et al. 2016). The haploid kernels will have a purple hue to the embryo and the non-haploid embryos will not. The RI-nj color marker makes it possible to utilize automated color sorting for selection of haploid kernels. Automated color sorting allows large amounts of kernels to be sorted by machine. The machine is calibrated to sort through large amounts of kernels using infrared light. The infrared light detects all kernels with a specific color and will discard all other kernels. This technology allows for larger population sizes to be sorted more efficiently and accurately than by traditional visual and manual selection.

The selected haploid seed then undergoes a process to induce doubling of the genome. Since haploid plants are normally sterile, genome doubling is required to produce a plant that is capable of being self-pollinated to produce viable offspring (Liu et al. 2016), producing a DH plant with two complete sets of chromosomes. Haploid genome doubling in maize programs is mostly done by treating the seed with colchicine. Colchicine is a chemical that is commonly used to double the genome in an embryo or seedlings. Due to its high toxicity to humans and other animals, and mutagenic properties several alternative methods have been studied such as N_2O applications, herbicide applications, and APM (amiprophos methyl) (Liu et al. 2016). Currently colchicine applications are the most inexpensive and successful methods used to double haploid genomes in maize programs.

Colchicine-treated seedlings (D0 plants) are then planted in the field or greenhouse. Plants can be tissue sampled and tested using marker technology to ensure that the desired traits are present and the genotype is homozygous as expected. The plant is then self-pollinated and the resulting seed is genetically homozygous and homogeneous (D1 generation; DH line). After genomic testing confirms the zygosity of the trait in the seed the planted kernels are grown in plots and either self-pollinated or for hybrid performance testing (Liu et al. 2016). The specific process and use of DH lines will vary depending on the needs of the specific breeding program.

Maternal haploid induction genes

MATL. Matrilineal (MATL) is an important gene in maize as it codes for a pollen specific phospholipase and influences the number of successful haploid inductions (Yao et al. 2018). MATL is expressed in the cytoplasm of pollen cells (Wu et al. 2017). The MATL gene is believed to affect the haploid induction rate of maternal inducer lines (Wu et al. 2017). The discovery of this gene has helped to increase the efficiency of modern high efficiency inducer lines and can potentially be of use in other crop species such as rice to increase haploid induction frequencies (Wu et al. 2017). MATL is also known by several different names describing the same gene, such as MTL, ZmPLA1, and NLD and are the consequence of three different studies published in the same year (Gilles et al. 2017; Kellihner et al. 2017; Liu et al. 2017).

The three studies used very similar methods to identify the MATL gene. The studies first identified QTL in inducer lines derived from Stock6 such as RWK, CAU5, CAUHOI by mapping studies. Fine mapping was used to ultimately identify a 4 bp insertion in a region on chromosome 1, shown to affect haploid induction rates. CRISPR-Cas9 was used to confirm the function of MATL by inserting the gene into a non-inducer inbred line [a 4 bp insertion leads to a frameshift mutation. Is MATL an active gene or is it the result of a knock-out mutation (loss of

function) This is important, when trying to use MATL in other species]. Increased haploid induction rates were observed in these lines proving the function of the gene in haploid induction (Gilles et al. 2017; Kellihner et al. 2017; Liu et al. 2017).

Two major quantitative trait loci (QTL) have been discovered to affect maternal haploid induction rates in maize. These QTL were named qhir1 and qhir8 (Giles et al. 2017). MATL is located in the qhir1 region on chromosome 1 and increases the haploid induction rate many times compared to rates in wild type plants (Zhong et al. 2019). The MATL gene has the largest effect on the haploid induction rate of a genome (Giles et al. 2017). Other genes affecting haploid induction rates (HIRs) on chromosome 9 have been observed to further increase the haploid induction rates when present in addition to MATL (Giles et al. 2017; Zhong et al. 2019). One such gene is ZmDMP. ZmDMP has a lesser effect on haploid induction rates compared to MATL and increases the haploid induction minimally over wildtype when present in addition to the MATL gene (Giles, et al. 2017; Zhong et al. 2019). CRISPR (clustered regularly interspaced short palindromic repeats) are used to determine the effect of the MATL gene by knocking the gene out of the inducer line and inserting the gene into a non-inducer line to observe the effect on the haploid induction rate (Giles et al. 2017; Kelliner et al. 2017; Liu et al. 2017). CRISPR may also be used to increase the haploid induction rate efficiency of modern induction lines (Zhong et al. 2019) by allowing a more precise way to alter the expression of specific genes that are known to influence haploid induction rates in inducer lines.

The MATL gene has also been identified and used successfully for haploid induction in rice using a rice orthologue of the maize MATL gene to induce an increase in haploid seed production of 2-6% haploid kernels in rice (Yao et al. 2018). The orthologous gene found in rice is known as OsMATL (Yao et al. 2018). CRISPR technology was used in an experiment

conducted to create the mutations necessary to induce haploid seed set (Yao et al. 2018). In the experiment two constructs 23843 and 23845 were transformed into cultivar IR58025B, to conduct CRISPR Cas9 gene editing (Yao et al. 2018). Each construct had a specific region in the genome that it was designed to target.

Construct 23843 targeted the amino terminal region in exon 1 (Yao et al. 2018). 23845 targeted exon 4 with the restriction site 4 bp away from the native maize inducer allele frameshift site (Yao et al. 2018). Haploid induction was observed with a rate of 6% on average, which is below the industry standard (Yao et al. 2018). As haploid induction rates may be female germplasm dependent, further experiments on different donors are conducted to ascertain which genes are involved (Yao et al. 2018).

Paternal haploid induction

In paternal haploid induction the donor is used as male and the inducer is the female line (Liu et al. 2016). The resulting DH lines inherit the cytoplasm from the inducer line and the chromosomes from the donor (Liu et al. 2016). Paternal induction is used in maize breeding programs, though not as frequently as maternal haploid induction due to lower haploid induction rates. CMS (cytoplasmic male sterility) is used in breeding programs to increase the efficiency in producing hybrid maize seed (Havey 1982). CMS in maize is cytoplasmatically inherited. A paternal inducer carrying the CMS cytoplasm results in plants sterile tassels.

Converting a male fertile line into a CMS line is accomplished by using a paternal inducer line with CMS. CMS plants can be easily pollinated by a male fertile line (Havey 1982), without need to remove tassels. CMS technology allows hybrid seed to be efficiently produced in

a production field by using a male sterile line as the female line and a male fertile line as the pollinator (Havey 1982).

The increased use of CMS technology in commercial seed production has proven extremely valuable in modern commercial breeding programs by reducing the time and resources involved in detasseling and scouting female rows for shedding pollen. It also reduces the likelihood of a female plant self-pollinating, which could potentially result in the loss of the entire field if genetic purity is compromised.

As useful as CMS technology has been for the maize seed production industry the technology has not been without challenges. CMS technology has been used on a large scale in the past. In the 1950s the Texas cytoplasm, T-cytoplasm was discovered (Levings 1993). The T-cytoplasm carried the cytoplasmic male sterile (CMS) trait. This germplasm became so widely used that it comprised over 85 percent of the hybrid corn grown in the United States by 1970 (Levings 1993). Around this time in 1969 an outbreak of Southern corn leaf blight caused by *Bipolaris maydis*, t-race was beginning (Levings 1993).

By 1970 the Southern corn leaf blight pathogen had infected a large portion of the hybrid corn in the United States. It was determined that the T-cytoplasm responsible for the CMS gene was also conferring susceptibility to the Southern corn leaf blight pathogen compared to native Maize with no T-cytoplasm (Levings 1993). This susceptibility caused the T-cytoplasm to be discontinued and detasseling to be used more widely in the industry (Levings 1993).

Table 1.3: Induction types with donor and inducer

Induction type	Donor	Inducer
Maternal	Female	Male
Paternal	Male	Female

Paternal haploid induction genes

Ig1 (Indeterminate gametophyte) is an important gene as it influences paternal induction by increasing the number of haploids occurring after pollination (Kermicle 1969). The Ig1 gene was isolated and identified by sequencing a mutant Ig1 gene (Ig1-O) (Evans 2007). The Ig1-O mutant has a phenotype with higher amounts of aborted kernels compared to a wild type line indicating a sterility gene. The Ig1-O mutant was then backcrossed into several inbred lines. The inbred lines showed varying degrees of paternal haploid induction based on the genotype of the inbred line. A second mutation was discovered: Ig1-mum. Fine mapping was used to sequence the gene and identify the location (Evans 2007).

Centromere mediated chromosome elimination is believed to be influenced by the CenH3 gene (Wu et al. 2017). This gene is believed to affect haploid induction rates and is a paternal gene. CenH3 was first identified and isolated in Arabidopsis and barley by Ravi and Chan (2010). The key finding of this study was identification of a GFP-tail swap mutant. The mutant was crossed with a wildtype parent. The offspring were sequenced and found to be haploid only containing the wildtype genome (Ravi and Chan, 2010). The haploid plants were also sterile and less vigorous than the diploid plants. When self-pollinated the GFP-tail swap mutant did not produce haploids (Ravi and Chan, 2010).

The CenH3 gene was also isolated in barley in 2011 by (Sanei et al. 2011). Some key findings of this study were haploids that were created by crossing barley with *H. bulbosum*. Sequencing was conducted on the resulting offspring and it was found that only the barley

genome remained and the *H. bulbosum* genome had been eliminated (Sanei et al. 2011). A low temperature of 18 C promoted chromosome elimination.

A test was conducted with antibodies to determine the centromere activity of the cells undergoing chromosome elimination (Sanei et al. 2011). It was found that the CenH3 protein activity was reduced in the eliminated chromosomes compared to the barley chromosomes (Sanei et al. 2011). The study concluded that the loss of CenH3 activity plays a role in chromosome elimination and subsequent elimination of the *H. bulbosum* genome from the barley haploid (Sanei et al. 2011).

A CenH3 study in maize was conducted by Kelliher et al. (2016). In this study, several CenH3 RNAi lines, tail swap or CenH3 transgenes were used. These resulting crosses were then compared to testcrosses with wildtype plants (Kelliher et al. 2016). The RNAi lines were found to not produce haploid plants. Some other CenH3 transgene lines that were used did produce haploids, but at low rates (Kelliher et al. 2016). The hemizygous CenH3-tail swap lines showed a significant increase in the haploid induction rates when backcrossed to wildtype plants as males. CenH3-tail swap was shown to have an influence on the haploid induction rates in maternal haploid induction (Kelliher et al. 2016).

In vitro haploid production

Haploid plants can be obtained in vitro when immature pollen (androgenesis) or ovules (oviculture) from a donor plant are used to develop haploid plants by tissue culture methods (Liu et al. 2016). Oviculture is when immature ovules are selected from a donor plant and placed on a media to form plant calli (Atanassov et al. 1995), but it is not nearly as widely used as androgenesis due to the inconsistency of reliably producing haploid callus tissue.

Androgenesis is most commonly used for heterozygous donor plants producing a segregating population. The best performing lines from the segregating population may then be selected for advancement in the breeding program. Usually immature anthers are selected from a donor plant with desirable agronomic traits. Immature pollen within those immature anthers represent gametophytic tissue, which is not fully developed and thus capable of differentiating into callus tissue (Liu et al. 2016). The anthers are then grown on media to produce plant calli resulting in haploid plants (Atanasov et al. 1995).

Haploid plant calli are exposed to a doubling agent such as Colchicine to double their genome creating a “doubled haploid” diploid plant (Liu et al. 2016). The DH plant is then transplanted and self-pollinated. The genome in the haploid plant must be doubled to create a fertile diploid plant, which is capable of being successfully self-pollinated to produce viable seed. Haploid plants are generally sterile, and not capable of producing offspring. This method has been used very successfully in wheat breeding programs. In vitro haploid induction has also been used in maize breeding programs, but not as successfully due to difficulty in finding a method that overcomes genotype specificity for response to anther-culture (Liu et al. 2016).

In vitro haploid production can be useful for crops such as soybean that may not have inducer lines available. However, this method tends to be very tedious, time consuming, and can be costly compared to in vivo induction. In vivo haploid induction is useful because it can be used as a way to induce haploid development in the field by the use of an inducer line. This method eliminates the need to culture callus tissue in the laboratory. However, in vivo haploid induction does require the use of reliable inducer lines, which may be difficult to obtain for some crops.

Current DH maize breeding programs have become increasingly efficient in recent years due to the development of highly efficient inducer lines. High efficiency inducer lines were developed by Iowa State University to overcome the shortcomings of earlier high frequency inducer lines that were well adapted to Europe, but not well adapted to the climate in the Midwest (Liu et al. 2016). Lines such as BHI 201 were developed (Liu et al. 2016). BHI 201 is a B73 based inducer line developed by Iowa State University in 2016 (Liu et al. 2016). It has an induction rate of 12-14% (Liu et al. 2016) and was developed from a genetic background that is well suited to the climate in the Midwest.

BHI 301 was also developed by Iowa State University in 2016 and has a similar induction rate to BHI 201 (Liu et al. 2016). These and other high efficiency induction lines have haploid induction rates of at least 10% (Liu et al. 2016), depending on environmental conditions and the inbred line being induced. Higher haploid induction rates increase efficiency by increasing the percentage of haploid kernels produced by each plant. The higher number of haploids produced allows a larger percentage of doubling to occur which then decreases the population size required and allows for more successful and efficient selection for inbred line development.

Chapter 2

In vivo haploid induction in soybeans

To date there are no efficient methods to produce DH soybeans on a commercial scale. Development of an efficient DH breeding program in soybeans would prove to be extremely valuable to modern agriculture. DH systems have been used in maize breeding programs for decades and have greatly increased the genetic gain of the crop. DH breeding programs have also

allowed new agronomic and phenotypic traits to be integrated into the maize genome much faster than traditional methods.

Haploid induction is not commonly used in soybean breeding programs as of recently, due to the inefficiency and low success rate of creating successful haploid plants. There have been few successful in vitro anther culture experiments conducted in soybean that have led to haploid induction (Lulsdorf et al. 2011). However, these anther-culture experiments have yielded mixed results and the haploid induction rates were quite low compared to commonly used commercial methods of haploid induction in other species. Formation of callus tissue was not consistent and would not be practical on a commercial scale. There has been some research done on in vivo haploid induction in the Fabaceae family of plants and other legume species such as lentils, chickpea, pigeon pea, and common bean utilizing interspecific crossing and centromere-mediated chromosome reduction (Wu et al. 2017). But this research is in its infancy and has not been reliably adapted for use in soybeans.

Table 2.1: Success factors for in vivo DH production

Crop	Flower type	Number of kernels/pollination	Availability of CMS	Success rate of hand pollinations	Seed per plant	High Frequency Inducers
Maize	Allogamous	300-500	Yes	90% +	300-500	Yes
Soybean	Autogamous	2-3	Yes	50%	50-100	No

Studies were conducted by Lewis et al. (1996), Ortiz-Perez et al. (2006), and Havey (1982) comparing different methods used to create hybrid soybean varieties on a commercial scale. Soybeans are autogamous and thus self-pollinating species (Lewis et al. 1996). The flowers are usually self-pollinated before they fully open by the anthers contained within the floral structure (Lewis et al. 1996). Natural soybean hybridization does occasionally occur, but it usually involves an insect vector or a neighboring plant that is in very close proximity to the cross-pollinated plant (Lewis et al. 1996). The rate is very low and would require an infeasibly large population size to produce a large enough quantity of seed for commercial production.

Commercial hybridization of soybeans is not very widely used by industry as it is very inefficient. Due to the autogamous nature of soybeans, hybridization requires cross-pollination by hand, which is very tedious. Hand pollination requires highly trained staff, who will open the flower of the female plant with a pair of forceps and extract anthers before they begin to shed pollen. The anthers from the male plant will then be extracted and deposited on the pistil of the female plant. If all goes well, hybrid seed will be formed.

This method is very time consuming, requires extensive training, and has a low success rate averaging approximately 50% seed set (Ortiz- Perez et al. 2006). Successful soybean hand pollinations are influenced by many factors such as temperature, humidity, available water to the plant, sunlight, and mechanical damage done by the researcher to the flowering structure. An average soybean pod also has a relatively low average seed number of two seeds per pod compared to other crops such as maize that are commonly hand pollinated (Ortiz- Perez et al. 2006).

Another method that has been explored is the use of insect vectors to cross pollinate soybeans (Lewis et al. 1996). This method requires plots to be planted in close proximity to each

other which allows insect vectors such as bees to easily fly between rows of plants (Lewis et al. 1996). In order for an insect vector mediated cross pollination to occur successfully, male sterility genes must be used (Ortiz-Perez et al. 2006). Male sterility genes create flowers that do not have fertile pollen and are thus incapable of self-pollinating. This allows the pistil of the flower to be fully exposed to be pollinated from a flower from the desired plot without contamination from its own anthers (Ortiz-Perez et al. 2006).

Male sterility genes are often combined with phenotypic marker traits that allow one to distinguish the truly male sterile plants from non-male sterile plants (Lewis et al. 1996; Ortiz-Perez et al. 2006). These phenotypic traits can be color markers such as purple hypocotyl and purple flowers and are linked to genic ms genes (Lewis et al. 1996; Ortiz-Perez et al. 2006). Genic male sterile plants differ from true CMS male sterile plants as restorer lines are not required to produce fertile offspring (Jin, 1999). The genic male sterile plants will also have a segregation rate of 50% male sterile and 50% fertile plants. The male sterile plants in the row will be kept and other non-male sterile plants in the row will be rogued to reduce the amount of contamination and self-pollination within the row (Lewis et al. 1996; Ortiz- Perez et al. 2006).

There are several genic male sterility genes (ms) that have been used to create successful cross pollinations (Ortiz-Perez et al. 2006) such as ms6. Ms6 is a male sterility gene that has been used in hybridization experiments (Ortiz-Perez et al. 2006). Insects are used to carry pollen from the male row and deposit it onto the male sterile female row creating a hybrid seed (Ortiz-Perez et al. 2006).

If a DH soybean breeding program were to be initiated, it can be developed using a combination of methods described previously. Based on the information presented, a paternal haploid system would most likely be the best choice for a soybean DH program. Paternal haploid

induction systems would be preferable in soybeans because self-pollination can be greatly reduced, if not eliminated entirely. This would save significant time and resources (Table 2). In comparison to maize, castration of flowers would be a major cost factor for seed production. As male sterility is available, a system combining ms and paternal haploid induction appears to be most economically feasible.

Table 2.2 Maternal vs. Paternal Induction System in Soybean

System	Maternal	Paternal
Induction Frequency	10%	10%
Doubling Frequency	10%	10%
Number of Crosses	10	10
Number of DH seed/cross	100	100
Inducer	Maternal Inducer	Male Sterile
Number of castrations/cross	10,000	0
Number of successful pollinations needed	5,000	5,000
Total number of castrations	100,000	0
Total number of successful pollinations	50,000	50,000

The haploid embryo would inherit the cytoplasm from the inducer and chromosomes of the donor (Liu et al. 2016). The non-male sterile line would be used as a donor line and would be grown in rows on opposite sides of the female inducer with a buffer of two rows of male sterile control materials on either side and between different plots. The buffer would prevent contamination from pollen of a different background or trait. Insects such as honeybees would be brought into the field during pollination to increase the efficiency of the pollinations.

The resulting seed can be tested for successful induction by utilizing a color marker similar to R-nj in maize or by seed chipping to determine, if the seed is haploid or diploid. After screening and genome doubling, the haploid plants can be grown in a greenhouse or growth

chamber and allowed to self-pollinate. Greenhouses or growth chambers would be preferable because the DH plants may be too fragile for field conditions. Greenhouses or growth chambers would also make it easier to track individual plants in pots and rogue any segregating material before flowering to prevent contamination from occurring. The resulting seeds may then be tested to ensure homozygosity for the given trait. The F2 plants can also be tissue sampled at an early growth stage in the greenhouse to ensure the plant is indeed homozygous for the desired trait and is not segregating.

For a soybean DH breeding program a CMS sterility system would be preferable over a genic sterility system because genic systems can only be successfully maintained through backcrossing. The system would also be very labor intensive due to the large amount of rouging that would be required. The CMS system can be successfully used if a female line with a fertility restorer gene is used. The resulting DH off-spring would then be fertile (Jin, 1999).

Once a stable line is obtained the seed can then go to field testing in short rows to test for agronomic, and phenotypic traits. The desired plants can then be selected for seed increase. The seed increases can occur in a normal field layout for soybeans. Since the desired traits will be homozygous in the DH line there is no need for cross pollination. The stable line can now eventually be commercialized as a new hybrid line.

Some key genes that have been shown to affect paternal haploid induction in other crops such as Ig1 and CenH3 may be beneficial for soy as well. Mutations for these genes may be discovered by the use of sequencing and backcrossing mutant lines to wildtype soy lines to confirm the mutation and the effect on paternal induction rates. Such as described by Evans (2007) in discovering Ig1 mutations in Maize. It may also be possible to introduce a CenH3- tail

swap gene using CRISPR or a transformation method such as described by Kelliher et al. (2016) to increase paternal haploid induction frequencies.

Summary and Conclusion

DH programs have been extremely beneficial to modern agriculture by allowing a homozygous inbred line to be developed much faster and more efficiently than traditional breeding methods. DH technology saves time and field resources by allowing for smaller population sizes to produce a homozygous (fixed) trait. This increase in efficiency in inbred line production has led to a substantial increase in genetic gain for maize over the past several decades. In vitro induction has been successfully used to produce haploid plants in several crop species such as rice, and soybeans. This system utilizes processes like androgenesis to produce plant calli from gametophytic tissue.

In vivo haploid induction is used to produce haploid plants in many commercial maize and wheat breeding programs. This method involves either interspecies or intraspecies induction. Interspecies induction methods are used in wheat DH programs and involve using a similar crop such as maize as a pollen donor and wheat as donor. The genome from the inducer is removed by centromere mediated elimination resulting in only the wheat genome remaining. Intraspecies haploid induction is used in crops such as maize on a commercial scale. This process involves either maternal or paternal induction.

Maternal induction is the most commonly used and preferred method in modern maize breeding programs due to its greater efficiency and more reliable source of stable inducer lines than paternal induction. High frequency inducer lines are also predominately maternal inducers. Genes controlling haploid induction have been identified in maize and other species, such as

MATL, ig1, CENH3, and ZmDMP. These genes are either affecting maternal (MATL, ZmDMP) or paternal haploid induction (ig1, CENH3).

Research has been conducted in soybean to develop efficient haploid induction programs. There have been promising results, but no breakthroughs to date. If a program would be developed towards development of soybean DH lines using in vivo haploid induction, paternal induction would be the preferred method due to the possibility use of male sterile lines. Use of a color marker would be required to discern which plants in the row are indeed male sterile and can be used as inducer plants. Pollination would be insect mediated by the use of honeybees to pollinate inducer plants with donor pollen.

Development of an efficient method to produce soybean DHs on a commercial scale would be extremely beneficial to modern agriculture as soybeans are a major agronomic crop worldwide. Increasing yield potential and allowing improved genetic traits to be quickly integrated into elite germplasm will help improve the yield potential of modern soybeans and will also allow new agronomic and output traits such as disease resistance, high oil and high protein lines to be developed and deployed more efficiently.

References

Atanassov, A. Zagorska, N. Boyadjiev, P. and Djilianov, D. 1995. In vitro production of haploid plants. World Journal of Microbiology and Biotechnology 11:400-405.

Coe, E.H. Jr. 1959. A line of maize with high haploid frequency. Naturalist 93: 873: 381– 382.

Chase, S.S. 1949. Monoploid frequencies in a commercial double cross hybrid maize, and in its component single cross hybrids and inbred lines. Genetics magazine 34. 3. 328–332. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1209449/ Accessed, March 2020.

Evans, M.S.M. 2007. The indeterminate gametophyte1 Gene of Maize Encodes a LOB Domain Protein Required for Embryo Sac and Leaf Development. Plant Cell (2007). 1; 46-62. doi: 10.1105/tpc.106.047506 Accessed, May 2020.

Geiger, H.H., and Gordillo, G.A. 2009. Doubled haploids in hybrid maize breeding. Maydica 54:485-499.

Geiger, H.H. 2009. Doubled Haploids. Maize Handbook. Volume 2. (2009) 641-657. Springer Science + Business Media LLC, Stuttart, Germany Accessed, March 2020.

Giles, L.M., Khaled, A., Laffaire, J-B., Chaignon, S., Gendrot, G., Laplaige, J., Bergès, H., Beydon, G., Bayle, V., Barret, P., Comadran, J., Martinant, J., Rogowsky, P.M., and Wildietz T. 2017. Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. The EMBO Journal (2017). 36; 6; 707-717. doi: 10.15252/embj.201796603 Accessed, April 2020.

Havey, M. 1982. The use of cytoplasmic male sterility for hybrid seed production. P.h.D dissertation. USDA -ARS, Department of Horticulture, University of Wisconsin, 1575 Linden Drive, Madison, WI USA.

Jin, W., Horner, H.T., and Palmer, R.G. 1997. Genetics and cytology of a new genic male-sterile soybean [Glycine max (L.) Merr.]. Sexual Plant Reproduction (1997);.10, 13-21 Accessed, April 2020.

Kelliher, T., et al. 2017. MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. Nature (2017). 000; 1-18. doi: 10.1038/nature20827 Accessed, May 2020.

Kelliher, T., Starr, D., Wang, W., McCuiston, J., Zhong, H., Nuccio, M.J., and Martin, Barry. 2016. Maternal Haploids Are Preferentially Induced by CENH3-tailswap Transgenic Complementation in Maize. Frontiers in Plant Science. (2016). 7; 114; 1-11. doi: 10.3389/fpls.2016.00414 Accessed, May 2020.

Kermicle, J.L. 1969. Androgenesis conditioned by a mutation in maize. Science magazine (1969) 166: 1422– 1424. doi: 10.1126/science.166.3911.1422 Accessed, February 2020

Levings, C.S. III. 1993. Thoughts on Cytoplasmic Male Sterility in cms-T Maize. The Plant Cell (1993) 5; 1285-1290. doi: 10.1105/tpc.5.10.1285 Accessed, May 2020.

Lewis, K.S. St. Martin, S.K. Widrlechner, M.P. Palmer, R.G. Hedges, B.R. 1996. Hybrid Soybean Seed Production: Comparison of Three Methods. Crop Science 36:1560–1567.

Liu, Z. 2017. A 4-bp Insertion at ZmPLA1 Encoding a Putative Phospholipase A Generates Haploid Induction in Maize. Molecular Plant (2017). 10; 520-522. https://www.cell.com/molecular-plant/pdf/S1674-2052(17)30012-6.pdf Accessed, May 2020.

Liu, Zhixin. and Wang, Yanbo, Ren, Jiaojiao, Mei, Mei, Frei, Ursula K., Trampe, Benjamin, Lübberstedt, Thomas. 2016. Maize DHs. Plant Breeding Reviews. Volume 40. (First Edition):123-166.

Lulsdorf, M.M. Croser, J.S and Ochatt, S. 2011. Androgenesis and Doubled-Haploid Production in Food Legumes. Cab International, Wallingford, UK (2011). Biology and Breeding of Food Legumes

159-177 Accessed, Nov 2019.

Ortiz-Perez, E. Horner, H.T. Hanlin, S.J. Palmer, R.G. 2006. Evaluation of insect-mediated seed set among soybean lines segregating for male sterility at the ms6 locus. Field Crops Research 97:353–362.

Randolph, L.F., and Fisher, H.E. 1939. The occurrence of parthenogenetic diploids in tetraploid maize.PNAS. (1939) 25. 4. 161–164.

Doi: https://dx.doi.org/10.1073%2Fpnas.25.4.161 Accessed, March 2020.

Ravi, M., and S.W.L. Chan. 2010. Haploid plants produced by centromere-mediated genome elimination. Nature 464:615–618.

Ren, J. Wu, P. Trampe, B. Tian, X. Lübberstedt, T. Chen, S. 2017. Novel technologies in DH line development. Plant Biotechnology Journal 15:1361–1370.

Sanei, M., R. Pickering, K. Kumke, S. Nasuda, and A. Houben. 2011. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. PNAS. 108:E498–E505. doi: 10.1073/pnas.1103190108 Accessed, May 2020.

Vanous, A.E. 2011. Optimization of DH production in maize (*Zea mays* L.). M.S. diss. Iowa State University. Ames, Iowa.

Yao, Li. 2018. OsMATL mutation induces haploid seed formation in indica rice. Nature Plants, 4: 530-533https://doi.org/10.1038/s41477-018-0193-y Date Accessed: May 2019.

Zhong, Yu, Liu, Chenchu, Qi, Xiaolong, Jiao, Yanyan, Wang, Dong, Wang, Yuwen, Liu, Zongkai, Chen, Chen, Chen, Baojian, Tian, Xiaolong, Li, Jinlong, Chen, Ming, Dong, Xin, Xu, Xiaowei, Li, Liang, Li, Wei, Liu, Wenxin, Jin, Weiwei, Lai, Jinsheng, Chen, Shaojiang. 2019. Mutation of ZmDMP enhances haploid induction in maize. Nature Plants. 5:. JUNE 2019: 575–580.