Title

The Genetics of Potato Haploid Induction, and The Potential Evolutionary Significance of Haploid Induction in Natural Plant Populations

Author

Nganga, Mwaura Livingstone

Publication Date

2022

Peer reviewed|Thesis/dissertation

The Genetics of Potato Haploid Induction, and The Potential Evolutionary Significance of Haploid Induction in Natural Plant Populations

By

MWAURA LIVINGSTONE wa NGANGA na WANJIKU

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PLANT BIOLOGY

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Luca Comai, Chair

John Harada

Julin N. Maloof

Committee in Charge

2022

Abstract

Haploid induction (HI) has been a cornerstone of potato and maize breeding for more than half a century. Haploids, organisms with gametic number of chromosomes, are obtained in these crops by crossing cultivars to special lines called haploid inducers. Among the outcomes of these crosses are seeds that lack a paternal genome. These result in embryos that, different from regular crosses, only inherit half the maternal chromosomes and none from the paternal parent. Fascinating as this phenomenon is, it remains relatively understudied, particularly in potato. The aim of my studies was to review the literature on maize and potato haploid induction mechanisms, formulate hypotheses that explain the trait, and test some of these hypotheses.

Beyond breeding interests, haploid induction is rarely studied. Haploids in plants occur in nature across dozens of seed plants that have been surveyed. However, haploids in nature are rare occurrences, perhaps due to their low fitness. In potato, many diploid accessions induce haploids when crossed to tetraploid cultivars. This common occurrence of haploids in potato crosses suggests that this is not the result of an evolutionary fluke that has fixed *deleterious* haploid induction alleles in wild potato populations. Here, I put forth a hypothesis that explains the presence of these haploid induction alleles in populations, and their evolutionary significance.

To further our understanding of potato haploid induction, I developed a diploid potato population segregating for the haploid inducer trait. This population was used to map a genetic locus that is associated with haploid induction in potato. Haploid induction was also found to be positively associated with seed death.

Finally, I observed that the population developed above was segregating for traits other than haploid induction. These included color, seed size, inheritance of the embryo spot, and the growth habit. I attempted to map two of these traits - seed size and the presence of the embryo

spot. A region on chromosome ten that has previously been shown to harbor numerous anthocyanin biosynthesis genes was found to be associated with the seed embryo (anthocyanin) spot trait. This confirmed that my method and population is robust enough to pick up QTLs. As for the seed size, no association was found between the trait and a genomic locus. However, this was not surprising given the experimental setup.

Overall, I was able to map for traits that have huge implications for diploid potato biology, breeding, and that might have evolutionary biology implications.

Matemo

Ũthuthuria ũyũ wanjĩrĩrie maambara-inĩ (*laboratory*) ya mbũrobeca Luca Comai cukuru-inĩ mũnene wa Njimbo ya Karibonia itũũra-inĩ rĩa Davis (*University of California - Davis*). Matanya makwa maarĩ gũcaria njiini (*gene* kana *genes*) cia kuuna merĩ (*reduce by half*) namba ya kurũmũcũmu mĩmera-inĩ. Bata wa kũmenya njiini ici nĩ atĩ tene ũrĩũka nĩcirĩteithia kũhũthia gũciarithania kwa mĩmera. Gwa kahinda gaka, gũciarithania mĩmera kũbataraga imera nyingĩ (kuuma imera mũgwanja nginya kenda) kamwĩira mũmera gathirange kana karĩng'a ũcĩarwo. Mũmera karĩng'a (karĩng'a; *i.e pure breed* kana *inbred*) ũrutaga maciaro mahanaine o-kĩmera. Maciaro ma karĩng'a mahananaga nĩ gũkorwo atĩ icunjĩ cierĩ cia kurũmũcũmu (*chromosome pair*) cia karĩng'a itirĩ ũtiganu kaimana. Kwa ũgwo, karĩng'a yeciarithia (*plant selfing*), mbegũ ciagwo ikoragwo na kurũmũcũmu itarĩ ũtĩganu, na citarie ta cia mũciari mbeũ icio. Mũhianĩre na mĩtugo ya kĩmera gia karĩng'a ĩkoragwo ĩiganaine na ĩtarĩ ũtiganu. Arĩmi oothe mabataraga mĩmera ĩhuanaine nĩgwo ũramati wa mĩgũnda ũhũthe - arĩmĩ matienda, kwa mũhano, mĩmera ĩmwe ĩkĩruta mĩthuka rĩrĩa ĩngĩ nĩrĩkĩtie maciaro, kana mbeũ imwe igĩthundũka ciumia thutha wa iria ingĩ. Karĩng'a nĩcio ihũthagĩrwo nĩgwo wĩĩgananu wa kĩmera woneke mĩgũnda-inĩ.

Hihi notũhote gũthondeka karĩng'a hatarĩ gwĩciarithia kwa mĩmera imera nyingĩ? No tũhote na njĩra ĩno ngũtarĩria. Tũkwanjia na mũmera mũtukanu (*heterozygous*) - gũtukana, haha, nĩ kuuga atĩ ti karĩng'a. Tũgũcoka tuune merĩ namba ya kurũmũcũmu mbeũ-inĩ iria cigũciarwo nĩ mũmera ũcio. Mĩmera ĩrĩa ĩgĩthundũka kuuma mbeũ-inĩ icio ĩgũkorwo na gĩcunjĩ kĩmwe, handũ ha igĩrĩ, gĩa o-kurũmũcũmu (*i.e we obtain haploid plants*). Thutha ũcio, no tũrutithie o-gĩcunjĩ gĩa kurũmũcũmu nĩgwo gĩcunjĩ kiume kĩngĩ thiinie wa cero (*plant cell*) cia mũmera ũcio (*chromosome doubling*), nĩgwo icunjĩ cia kurũmũcũmu cihinge igĩrĩ rĩngĩ. Tondũ twarutithia gĩcunjĩ kuuma kĩngĩ (na njĩra ĩngĩ nĩtatwahũra maita na igĩri o-gĩcunjĩ), tũgũkorwo twathondeka karĩng'a (*doubled-haploid, which is homozygous just like an inbred*). Icunjĩ cierĩ cia kurũmũcũmu igũkorwo cihuanaine kaimana. Njĩra ĩno ya gũthondeka karĩng'a no ĩhoteke na

kĩmera kĩmwe handũ ha imera nyingĩ iria ibataragwo kũhũthĩra njĩra ya gwĩciarithia kwa mũmera. Thĩĩna ũrĩho nĩ atĩ gũtirĩ njĩra nyingĩ cia kuuna merĩ namba ya kurũmũcũmu (*haploid induction*). Njĩra iria cirĩ ho citingĩhũthĩka na mĩmera miingĩ. Kuuna merĩ namba ya kurũmũcũmu ti ũhũthũ.

Nĩkũrĩ waru mũthemba ũmwe, ĩrĩa irũa (*pollen*) rĩayo rĩngĩhũthĩrwo gũciarithia waru ingĩ, rĩtũmaga waru icio ciciare mbeũ *haploid* (nyune). Mbeũ ici citigayaga (*inherit*) kurũmũcũmu cia irũa (*pollen*). Kwa ũgwo, mbeũ icio ikoragwo na kurũmũcũmu cia ũka (*female* kana *maternal chromosomes*) tu. Kurũmũcũmu cia nyune (mũmera wĩna nuthu ya namba ya kawainda ya kurũmũcũmu, *i.e haploid*) ciahũrithanio maita (*chromosome doubling*), mũmera ũcio muune ũtuĩkaga karĩng'a (*double haploid*). Mworoto wakwa nĩkũmenya njiini iria itũmaga waru ũyu muuni (*haploid inducer*) uune (*haploidize*) waru iria ingĩ. Nĩndahotire kũmenya atĩ nĩ kurũmũcũmu ya gatatũ ĩna njiini nyuni kurũmũcũmu, ona gũkorwo ndĩahotire gũtigithũkania nĩ njiini ĩrĩkũ yuunaga kurũmũcũmu.

Wĩra ũyũ nĩ ũkũhotithia njiini ĩyo yoneke, na thutha mũnene, gũtũmĩra njiini ĩyo kuuna kurũmũcũmu mĩmera-inĩ ĩrĩa ĩngĩ na kũhũthia ũciarithanio wa mĩmera. Ningĩ kũmenya njiini ĩyo, na umenyo wa ũrĩa ĩrutaga wĩra cero-inĩ, nĩigũtũteithia kũmenya ũria mũtũũrĩre wa mĩmera ũcenjagia na mahinda (*evolution of plants*) na ũrĩa macĩgĩrĩra mangĩcũngĩrĩra gũcenjia gũkũ na njĩra ya kuuna kana kũhũra maita namba ya kurũmũcũmu. Maya mothe nĩ matemanie na magabacĩrĩrwo riboti-inĩ ĩno ya ndicatĩconi (*dissertation*).

Ngaatho

Wa mbere nĩ gũcokeria awa na ayu - Ng'ang'a (wa Mbũgua na Wanjirũ) na Wanjikũ (wa Mwaũra na Wamaitha) - ngaatho nyingĩ nĩkũndera hatarĩ kũndoreria njĩra o-yoothe hoka nĩ maũndũ ma gĩthomo. Ũrerĩ ũyũ nĩwaheire mweke na wĩyathi wa gũtũmĩra meciria makwa gwĩtemera njĩra o-yothe ya ũgo ingĩenda kuuma ndĩo mũnyinyi. Na o-ho, njokerie gũũka, mũtiga irĩ na ĩrĩri, Mbũgua (wa Ng'ang'a 'Munge' na Ngonyo), na cũcũ, Wanjirũ (wa Wambu Kĩnũthia na Mũkuhĩ) - tondũ wa gũkorwo marĩ areri akwa a kerĩ. Nĩmandonyire ũhoro wa ũrĩmi wa mĩmera na mahiũ, na ningĩ makĩnduta bata wa kwĩrutanĩria mũtũrĩre-inĩ na gĩthomo-inĩ. Ũrutani wao nĩgwo ũhotethĩtie gũkinya nginya handũ ha kũheo thũmbĩ ya *PhD.*

Ningĩ nĩngwenda gũcokeria ngaatho nyingĩ mbarĩ yoothe ya Ng'ang'a (Kamithi) na ya Mwaũra Ndiho (Ndegwa) - arĩa me mwoyo na arĩa mathiĩte ũkiriinĩ. Makĩria ngwenda gũcokia ngaatho cia mwanya kũrĩ acio - aingĩ a o ngomi na ngoma - marũ-ĩire wĩyathi na kĩhoto kĩa mũndũ mwĩirũ hĩndĩ ya wathi wa nyakerũ bũrũriinĩ wa Gĩkũyũ, na Kenya yoothe. Nĩmaimirwo na makĩgirio gĩthomo gĩa kĩĩmerera na gĩa ũthũngũ, magĩtunywo indo na mĩgũnda yao, makĩherithio icagi-inĩ (*concentration camps*) na aingĩ makĩonja meciria na mwiri kana magĩkua. Wĩrutĩri, ũmĩrĩria na mworoto wao ndwathererire na rũĩ, kwona atĩ nĩ niĩ ũyũ mwana wao – Mwaũra wa Ng'ang'a na Wanjikũ - waheo thũmbĩ ya *PhD.* Ngaatho ciakwa ingĩcokia tu na njĩra ĩmwe - gũtigĩrĩra atĩ mĩoroto na iroto cia njamba icio nĩciakinyanĩra na itinahuka. Makĩria nĩ ngwĩheana gũtwarithia na mbere wĩra ũcio maambirie gũtigĩrĩra atĩ ũkombo wa meciria na wa mĩĩrĩ ya ndũrĩrĩ ciothe na andũ othe - kana nĩ, kana ti, aka na arũme - nĩwathira.

Ningĩ nĩngũcokeria ngaatho mwarĩ na aariũ a awa na ayu arĩo Mbũgua, Wambu, Ndegwa, na Wanjirũ - na tũihwa twakwa, Ng'ang'a na Wanjikũ, nĩũndũ wa ũrata wao na kũũhe ngoro manahe.

Ningĩ nĩngwenda gũcokia ngaatho cia mwanya kũrĩ ndũũrĩrĩ cia Patwin, eene mĩgũnda na macĩgĩrĩra maria itũũra rĩa Davis na cukuru mũnene wa Njimbo ya California itũũra-inĩ rĩa Davis wakĩtwo. Ndũrĩrĩ cia Patwin nĩ ciagereire maũndũ maritũ moko-inĩ ma wathi wa nyakerũ o-ta ũrĩa andũ a ndĩra ciitũ magereire. Matigari ma Patwin, o-haria mũrĩ, nĩndamũtambũra nĩũndũ wa ũcamba na ũmĩrĩru wanyu, na hamwe nĩtũgũthie na mbere kwĩruta ũkombo-inĩ.

Mũthia, njokerie ngaatho ataari (*mentors*) akwa arĩa manahe mweke wa kũruta wĩra mambara-inĩ ciao, na arĩa mananyitĩrĩra na kũhe ngoro rũgendo-inĩ rũrũ. Aya nĩ Bethany Zolman wa cukuru mũnene wa Njimbo ya Micauri itũũra-inĩ rĩa Saint Louis (*University of Missouri - St. Louis*), Todd Mockler wa *Donald Danforth Plant Science Center* (itũũra rĩa *Saint Louis, Missouri*), na Luca Comai wa cukuru mũnene wa Njimbo ya California itũũra-inĩ rĩa Davis. Ũthuthuria wakwa ndekĩire mambara-inĩ ya Luca Comai. Ũrutani wake nĩũnginyĩthĩtie makinya manene. Na o-ho njokerie arĩa twanarutithania wĩra nao mambara-inĩ ya Luca Comai. Na ndĩkie na gũcokeria ngaatho kamĩtĩ ya ndicatĩconi yakwa (Luca Comai, Julin Maloof na John Harada) nĩ kũnũnga mananũnga.

Ũtarĩria wa mwandĩki

Mwaũra Ndiho “Kahiga Livingstone” Ng’ang’a aciarĩirwo itũũra-inĩ rĩa Cura, mwena wa Kabete ya Kĩanda (*Lower Kabete*) hakuhĩ na taũni ya Wangige mwaka wa 1990, mweri wa kanana, mũthĩ wa 31. Aciari aake nĩ Ng’ang’a wa Mbũgua na Wanjirũ, na Wanjikũ wa Mwaũra na Wamaitha. Nĩ wa mũhĩrĩga wa Ethaga (Ambura kana Akĩũru), na mbarĩ ya Ng’ang’a kuuma Kabete. Athomeire cukuru-inĩ wa Ndũrarua Primary School, Lower Kabete. Gĩthomo gĩake gĩa cekondarĩ athomeire cukuru wa Mũrũria Secondary School, Gatũndũ (mwaka wa mbere) na Kĩambu High School, Kĩambu. Thutha ũcio nĩathomeire ndigirii cia Mbayo-Kemicitĩrĩ na Mbayo-Tekinoronjĩ (*Bachelor of Science in Biochemistry, and Bachelor of Science in Biotechnology*) cukuru-inĩ mũnene wa Njimbo ya Micauri, itũũra-inĩ ria Saint Louis (*University of Missouri - St. Louis*), na akĩrĩkĩrĩria na githomo gĩa *PhD* ya Mbayoronjĩ ya Mimera (*Plant Biology PhD*) cukuru-inĩ mũnene wa Njimbo ya Karibonia itũũra-inĩ rĩa Davis (*University of California - Davis*).

Mworoto wake nĩ gũtũmĩra ũũgĩ wa Cayanci na Tekinoronjĩ kũhũrana na ũkĩara na mĩrimũ, na gũtigĩrĩra atĩ mũndũ o-wothe nĩagĩa na mweke wa gũtũmĩra taranda na iheo ciake kwa gĩkĩro kĩa igũrũ o ta ũrĩa kũngĩhoteka.

Ũrongoreria

Ndicatĩconi (*dissertation*) ĩno ĩrongoreirio:

- Ndĩra ciitũ - arĩa makirĩte, arĩa marĩho na arĩa magooka. Mũtikanage, na ngoma cianyu o-haria cirĩ nĩ-ikene, cĩĩganĩre na itikanaage mahinda mothe. Na mũtũũre mwendaine, na o mboco yagwa, mũkenyũrana. Nagwo mũrĩnga wanyu ndũkanathire, mbeũ cianyu na mahiũ manyu matikanahuke, na mwonage kĩmera.
- Arĩa oothe marũagĩra ĩhoto cia o-mũndũ, na makĩria cia andũ arĩa mahinyĩrĩirio nĩ mambemberũ (akoroni) na matigari ma mbemberũ.

Acknowledgments

I would like to acknowledge and thank everyone who has contributed to the success of this work. In particular, I thank Professor Luca Comai and Dr. Isabelle Henry for their mentorship, supervision, and guidance during my PhD research. I thank Ms. Benny Ordoñez and Dr. Kirk Richard Amundson, with whom we spent countless hours discussing potato biology. I also appreciate the many late night hours Dr. Mohan Anand Prem Marimuthu, Dr. Peter Gregory Lynagh, and I spent time in the lab discussing various ideas. Dr. Marimuthu was very helpful in helping me understand the basics of haploid induction. In addition to these, I thank my former undergraduate assistants who helped me conduct research. They are Ms. Melodie Najarro, Ms. Julie Ngov, Ms. Rutu Vyas and Ms. Victoria Stewart. Without their help in the lab and in the greenhouse, this work would not have been accomplished. The greenhouse staff - Messrs Ronald Lane, William R. Werner, Jianguo Chen, Davith Hin and Tony Lopez - contributed a great deal to success by making sure that my greenhouse plants received the best of care. The United States Department of Agriculture (USDA) National Genetic Resources Program (NGRP) provided me with the plant material that was instrumental in my studies.

I am grateful that the National Science Foundation funded (NSF grant #1650042) a three-year predoctoral fellowship (The NSF Graduate Research Fellowship) that supported my research from 2016-2019. In addition, I was in part supported by NSF grant PGRP IOS-2055260 "*Mechanisms of haploid induction in potato*". Other entities that supported me include UC Davis Mike and Renee Child Institute for Innovation and Entrepreneurship through the Business Development Fellowship (2020 - 2021). The UC Davis Innovation Institute for Food and Health through the IIFH Innovator Fellowship (2021 - 2022).

I would also like to thank my past mentors who gave me the confidence, and the opportunity, to pursue research. These are Dr. Bethany Zoolman (University of Missouri - St. Louis), and Dr. Todd Mockler (Donald Danforth Plant Science Center).

Outside the academy, I would like to thank my parents, grandparents, siblings (Mbũgua, Wambu, Ndegwa and Wanjirũ), and my nephew and niece (Ng'ang'a and Wanjikũ) for their steadfast support. Their love and support have been the wind under my wings. Lastly, I thank all my friends and the 2014 Plant Biology graduate class. They have been there for me. They helped me party away all the mental pressure of completing a PhD. This cannot be overstated.

And finally, I acknowledge the true owners and stewards of the land where all my work was done. The Patwin Nation dutifully managed the lands currently occupied by The University of California - Davis, and The City of Davis for centuries. The Patwin Nation, like my people, suffered a great deal of loss and oppression under the yoke of colonialism and imperialism. I admire the resilience, the courage, and the grace with which they have borne their tribulations. Together, we will continue rebuilding our societies.

Dedication

This work is dedicated to my family, my ancestors and their descendants, as well as all that spent and continue to spend their time and resources to fight colonialism, imperialism, and their legacies.

I would like to especially single out, among my lineage, my grandmother and grandfather - Wanjiru wa Wambu Kĩnũthia na Mũkũhĩ and the late Mbũgua wa Ng'ang'a na Ngonyo - who were stripped of most their culture, identity and religion, and were barred from getting both the indigenous and western education by racist British colonial government laws. Nonetheless, they persevered and fought back the injustices, and taught me much indigenous science and knowledge. Their strong work ethic and dedication to truth and right laid for me a strong foundation. I would also like to thank my mother and father - Wanjikũ wa Mwaũra na Wamaitha, and Ng'ang'a wa Mbũgua na Wanjirũ - who were denied their chance to pursue studies in physics and chemistry, respectively, by imperialistic policies and the corruption of a repressive regime supported by global powers during the cold-war. Despite not achieving their dreams, they turned to small-holder farming, worked hard and did all they could to ensure that I and my siblings went further than they did. They denied themself of life's pleasures so that my siblings and I receive a good education. Their trust in my ability to make informed career choices as far back as primary school allowed me freely choose a career in science. Their sacrifices have paid off.

I also dedicate this PhD to my descendants and community at large. May you take a cue from my parents and grandparents. *Kagũrũ karĩ ime gakĩrĩte karĩ mũhu*.

And lastly, to everyone in the struggle to rid the world of all forms of oppression, particularly the overt and covert colonial and imperial domination of indigenous people around the world. May we all decolonize our minds, and continue the works of Frederick Douglass,

Mekatilili wa Menza, Marcus Garvey, Mũthoni Nyanjirũ, The Mau Mau and Kĩmathi Waciũri, Malcolm X, Rev. Martin Luther King, Wangarĩ Maathai, my grandmother Wanjirũ wa Wambu na Mũkuhĩ, and many other sung and unsung heroes.

Table of Contents

Chapter 2: A Review of Potato Haploid Induction (HI) & the Evolutionary Significance of HI

Chapter 3: Mapping Potato Haploid Induction gene(s)

Chapter 4 - Seed Traits

Chapter 5: Dissertation Conclusion and Summary

Chapter 1: Review of Maize Haploid Induction

1. Importance of haploid induction (HI)

The first report of a haploid plant is usually credited to Blakeslee [1]. Since then, haploid plants have been discovered and produced in hundreds of species [2]. Artificially, haploids are obtained through anther culture, ovule culture, interspecific or inter-genera crosses, or through special genetic crosses to strains known as 'haploid inducers' [2]. The number of chromosomes in the haploid is doubled in breeding programs to produce what is known as a *doubled haploid* (DH). DHs are quick alternatives to inbred lines that take many generations to develop. It is possible to obtain an DH "inbred plant" in a single generation compared to many generations using inbreeding [3].

For those reasons, haploids have become central in the breeding of some crops, particularly potato and maize. Maize and potato haploid induction are widely adopted in breeding programs due to the ease of use of these haploid induction systems. An interest in genetic haploid induction surged in 2010 when a method using manipulation of centromere proteins was used to obtain haploids in *Arabidopsis thaliana* [3–5]. In the following review, I will look first at haploid induction in maize because of its similarity to haploid induction in potato. My aim is to discuss possible haploid induction mechanisms that could be involved in both of these systems. In Chapter 2, I will discuss potato haploid induction, and in general, the evolutionary significance of haploid induction in angiosperms beyond the often-cited function in establishment of species barriers.

2. Haploid induction in maize

2.1. Definition of important terms

- HI - Haploid induction
- HI Cross - a genetic cross suitable for obtaining haploids in the resulting progeny.
- HIR - haploid induction rate. In the case of maize or Arabidopsis, it is the percentage of live seeds that develop into haploids. In the case of potato, it is the number of haploids obtained divided by the number of fruits that were analyzed for the presence of haploids.
- Inducer - a haploid inducer line
- Polyembryony (also known as twin embryos or twinning) - when a seed contains two or more embryos. This occurs sometimes in the maize HI cross and numerous other HI systems [6–10]
- Twinning-rate - the percentage of seeds that have more than one embryo
- (Proper) Uniparental genome elimination - specific loss of one parental chromosome set following successful fertilization. Typically, these chromosomes are lost during early embryonic cell divisions. E.g, chromosomes that have defective/incompatible centromeres such as those contributed by a CENH3-based haploid inducer [4,5]
- Parthenogenesis - Development of an unfertilized egg cell into a zygote without any input from a male gamete. E.g during ectopic expression of the *BabyBoom* gene in the egg cells of rice [11]. It is a mechanism of apomixis in pennisetum [12].
- Heterofertilization - independent fertilization of egg and central cell in the same ovule by sperms that are derived from different pollen tubes [13,14].
- Double fertilization - the dominant mode of sexual reproduction in angiosperms where an egg and central cell are fertilized by a sperm each.
- Single fertilization - In plants, the abnormal fertilization of only the egg or central cell instead of both as happens during double-fertilization.

2.2. Brief introduction to maize HI

In maize, haploids occur spontaneously at a rate of about 0.01-0.1% [15–18]. In 1959, E.H Coe [19] identified Stock 6, a maize line with relatively efficient (~3% HIR) haploid induction, either when selfing or when outcrossing [19,20]. The authors noted that the line might have had a higher than normal degree of heterofertilization [21]. It was suggested that "haploids result primarily from failure of fertilization caused by an abnormal condition, either inherent or induced, in a male or female gamete, and subsequent development of the reduced egg into the embryo" [16]. The rate of haploid induction either spontaneously or by use of haploid inducer line is dependent on both the paternal and maternal genetic background [8,17,19,22].

All modern maize haploid inducers are derived from the Stock 6 line, and have been bred for high HIR, which can be as high as 15% [23,24]. This involves the combination of at least two QTL, *qhir1* (from the stock 6 line) and *qhir8* (of unknown origin). The suggestions made by Ed Coe [19] about the characteristics and mechanism of haploid induction were prophetic.
Maize haploid induction is associated with the following events and traits pertaining to pollen development and function:

- twin-embryony or twinning [6,7]
- single fertilization events and heterofertilization in ovules [14,21,25–27]
- embryo and endosperm abortion [28]
- Segregation distortion of the *qhir1* HI QTL (*mtl* gene) - the major HI locus [28–31] .

Notwithstanding the above pollen abnormalities observed in Stock 6, pollen fertility appears normal [6,32,33], although the viability of inducer pollen might be slightly lower than the wild-type pollen [32]. The inducer pollen germinates normally, but it grows slower and is out-competed by non-inducer pollen [28,30]. Each of the listed factors is discussed below.

3. Nature of haploid induction

3.1. Twin Embryony (Polyembryony/Twinning)

Twin embryony during maize haploid induction was first reported by Sarkar & Coe [16], but no definitive relationship between HIR and twin embryony was drawn [16]. Interestingly, X-ray and UV radiation of pollen have been reported to cause HI in maize, and to increase twin-embryony rate [17,34]. Elevated twin embryo-rate was observed in maize HI crosses by Qiu *et al* 6, which was reportedly a confirmation of Li *et al*'s [35] results. It has further been shown that HIR is positively correlated with twin-embryony 7. These twin embryos have been found in various combinations: haploid-haploid, haploid-diploid, and diploid-diploid twins that can also include aneuploids [6,7]. This is not surprising as twin embryony is associated with haploids in more than 24 other species 8.

3.2. Maize Haploid Inducer Pollen abnormalities

The commercially viable HI trait, which is dependent on the joint action of two QTLs, *qhir1* and *qhir8* loci (next section), is paternally determined [19]. Single fertilization events and abnormal pollen morphology (unequal sperm size, small sperm, or one big sperm) have been observed in haploid inducers [36]. Haploid inducer pollen is associated with sperm cells of altered morphology, which was hypothesized to be the cause of HI and which may explain the observation of single fertilization events [36]. Haploid inducer pollen can also display aneuploidy as evidenced by cytology and by single cell sequencing experiments [6,32,37]. This aneuploidy starts as early as microsporogenesis, where more than or less than 10 bivalents are observed in haploid inducers but not in normal inbreds [38]. Qiu et al. 6 goes as far as suggesting that the pollen mother cells are aneuploid themselves. Li et al [32], by single nuclei sequencing, observed the formation of fragmented chromosomes in the sperms of the haploid inducer and suggested that this genomic instability may lead to the HI trait. In conclusion, the pollen karyotype may be defective, but its characterization is incomplete.

Recent studies identified a burst in ROS (reactive oxygen species) after the generative cell division (PM II) in the haploid inducer pollen [39,40]. A regulated ROS burst is a classical signature of programmed apoptotic cell death or cellular response to biotic conditions (e.g viral or bacterial attacks) [41–43]. Additionally, harsh environments and chemicals such as UV light and ozone that damage DNA cause a ROS burst in response to the DNA damage [41]. On the other hand, the unexpected accumulation of ROS such as by photosynthesis during high light intensity also induces cellular and DNA damage [41,44].

It is possible the ROS burst seen in haploid inducer pollen is in response to DNA damage and genome instability in the mutants. The damaged chromosomes are incapable of participating in cell division. However, is this DNA damage a response or cause of the haploid induction process? That is, is HI caused by damaged DNA or are "unincorporated chromosomes" being destroyed as *foreign* DNA? Other ROS studies in different systems have similar confounding observations where ". . . one cannot rule out that some of the DNA damage observed in plants after exposure to stress is a consequence of the onset of PCD rather than actual stress-induced DNA damage" [45]. It is likely that ROS is the cause of HI given that knocking-out of a pollen specific peroxidase or treatment of pollen pre-pollination with chemicals (Methimazole) that induce ROS production leads to HI [40]. Peroxidases are involved in both production and in response to excessive ROS production [46].

3.3. Partial transmission of paternal DNA and cellular content during HI

In multiple haploid induction systems where haploidization results from genome elimination, it is common to find chromosomes or chromosomal fragments from the haploid inducer genome persisting in the progeny [47–49]. This has been interpreted as occasional incomplete genome elimination, and has been observed in oat haploids from oat x maize crosses, and barley haploids from inter-specific barley crosses [47,50–52]. Independent studies that I was not able to access, but cited by others, claimed evidence of genomic instability, during

maize HI, which includes laggards, micronuclei, chromosome bridges, aneuploidy and mixoploidy, and rearranged chromosomes such a ring-chromosomes [6,20,51,53–55]. Paternal DNA transmission into an otherwise haploid zygote can in theory take place via addition of independent paternal chromosomes (chromosome addition), or introgression (recombination of paternal DNA with the maternal genome). Introgression has not been clearly demonstrated, while multiple studies are consistent with addition.

Often, the maize haploid inducer addition chromosomes appeared unstable resulting in mosaic (mixoploid) endosperm or radicles [6,20,51]. Endosperm aneuploidy increased gradually from 5 DAP (day after pollination) to 13 DAP [51]. Mixoploidy can be ascribed to different endosperm sectors experiencing different rates of paternal chromosome elimination. Most cells of HI cross progeny are either haploid (2n=10) or diploid (2n=20), with aneuploids at low frequencies [20]. Curiously, I have been unable to find published evidence on the persistence of addition chromosomes in adult maize plants. The maize HI system might be efficient at generating euploid plants (either haploid or diploid). It was proposed that HI takes place very early in embryogenesis in the form of paternal genome elimination. During early embryo development, the aneuploid cells are outcompeted and their proportion decreases although aneuploidy and mixoploidy might persist into the adult progeny [6,20,51].

The use of CRISPR has shown that haploid induction in *Stock6* derived-lines involves, to some extent, the fusion of the egg and the sperm. By crossing wild-type lines to a haploid inducer that expresses CRISPR-CAS9 and gRNA, CRISPR-CAS9-edited haploid plants can be obtained [56]. This could imply transient expression of the haploid inducer genome, or transcription of haploid inducer pollen transcripts in the zygote or early embryo, or the transfer of Cas9 nucleoprotein via the cytoplasm. The gene-edited haploids obtained from these HI crosses had no paternal DNA contamination [56]. Together, the occasional presence of residual paternal chromosomes and the transfer of paternally-expressed CRISPR reagents in the haploids suggest some form of karyogamy.

3.4. Heterofertilization

Heterofertilization as a phenomenon in maize was first reported by Sprague in 1929 and described as the process that results in a genetically different embryo-endosperm pair caused by delivery of genetically different sperm cells to the ovule [13]. Heterofertilization is widespread in maize and occurs naturally at different rates, dependent on genetic background [13,14].

The association of heterofertilization with maize HI was first suggested by Coe [21] when studying Stock 6, the original haploid inducer line [21]. Haploid inducer pollen exhibits dramatically increased heterofertilization rates compared to non-inducer pollen [25,27,57]. In one elegant experiment, one maternal parent and two paternal parents (an inducer and a non-inducer) were used. Pollen from both paternal parents was applied to the same silks but with a 4 hour delay between the two pollinations. When inducer pollen was first applied, the heterofertilization rate was much higher than when non-inducer pollen was applied first [25]. More direct evidence that HIR is positively correlated with heterofertilization was obtained by comparing the heterofertilization rates between related haploid inducers with different HIR [58].

The correlation between heterofertilization and haploid induction is probably related to single-fertilization events (see next section). In Arabidopsis, single fertilization events cause heterofertilization where a second pollen tube is attracted into an ovule as a fertilization recovery mechanism [59–62]. Interestingly, heterofertilization and haploid induction can also be triggered by application of mitotic inhibitors on maize pollen, suggesting that these chemicals affect pathways that are defective in maize haploid inducers [26,27,58,63].

3.5. Single fertilization & role of fusogen proteins

Improper egg-sperm cell fusion has been suggested as the cause of maize haploid induction. Ovules and seeds lacking either embryo or endosperm are observed following a HI cross [6,25]. This echoes a 1990 study that reportedly showed single fertilization events and delayed development of either egg or central cell following pollination with HI pollen [36,64]. This prompted the hypothesis that haploids are formed when a single sperm fertilizes the central cell

and the egg cell is stimulated to develop unfertilized [26]. Unfused sperm cells have been observed in the ovules with developing endosperms in the haploid induction crosses [33]. Similar heterofertilization events are observed in various Arabidopsis egg-sperm fusion mutants (some that are haploid inducers), in in *Solanum nigrum* x *Solanum luteum* haploid induction cross, and during apogamy in *Zephyranthes texana* (now known as *Habranthus tubispathus* or Atamosco Lily)[65,66] [61,65–67]. In all of these cases, the pollen tube delivers the sperm pair to the ovule, after which the sperm attaches to the egg without proper fusion/karyogamy.

3.6. Segregation distortion, and embryo & endosperm abortion

Seed abortion (no kernel development), embryo abortion (normal endosperm without an embryo), and endosperm abortion (normal embryo without an endosperm) are all observed in the maize HI crosses [28,30,31,68,69]. The direct link between HIR and embryo/endosperm abortion is also observed when either of the HI causal genes (*Matrilineal, MTL*; and *DUF679 domain membrane protein, DMP*. See below) is knocked out in non-inducer lines [30,69–71]. Genotyping endosperm tissue of defective kernels and haploid kernels from a heterozygous haploid inducer showed that these kernels always result from pollen grains that inherited the haploid induction allele [28,31].

Haploidy and kernel abortion reduce genetic transmission of the haploid inducer allele. In a heterozygous background, where the wild-type and haploid inducer allele competition can be measured, there is strong segregation distortion in seedlings against the haploid inducer allele. This segregation distortion is present only when the haploid inducer is used as male and absent when used as female [28,30,31,68]. HI rate is variable depending on the female line used and the amount of segregation distortion observed is correlated with observed HIR [31]. Therefore, segregation distortion is dependent on the efficiency of haploid formation. This distortion is also observed when the HI causal gene (*MTL*) in the major QTL is artificially knocked out in non-inducer background [30]. However, segregation distortion is not observed at the *qhir8* QTL that contains the *DMP* gene [68]. Why *dmp* segregation is not distorted requires further investigation.

However, I hypothesize that *dmp* segregation distortion occurs but remains undetected because the *dmp* effect (0.1-0.3% HIR) is much lower compared to the *mtl* effect (3% HIR) [70,71]. The distortion would be observed in a study with larger sample size.

4. Mapping maize HI genes

4.1. *qhir1*, the major HI QTL

Deimling et al. [72] identified two maize HI QTL, the major one being on chromosome 1 [31]. QTL analysis of a different mapping population identified the same region on chromosome 1 (the *gynogenesis inducer 1* (*ggi1*) QTL; renamed *qhir1* in subsequent studies) as responsible for HI. A different study used new populations derived from inducer x inducer, and non-inducer x inducer crosses to identify eight QTLs with the two most important QTL, *qhir1* on chromosome 1 and *qhir8* on chromosome 9, explaining 66% and 20% of HI variation respectively [68]. *qhir1* was then fine-mapped to a 243 kb region harboring the HI locus [29]. Using a novel technique that was named “Conditional Haplotype Extension”, the *qhir1* region was partitioned into two regions - *qhir11 and qhir12* [24]. This was done by looking in the *qhir1* region for a haplotype shared by 53 inducers derived from 29 breeding programs. In all cases, they traced some ancestry to the original Stock 6 inducer [24]. Whereas this study correctly partitioned the *qhir1*, their conclusion that subregion *qhir12* is responsible for HI was erroneous [73]. The *qhir11* subregion of *qhir1* (with 13 genes & pseudogenes) was ultimately identified as responsible for HI [30,69,73,74]. Three independent genetic complementation studies identified a 4bp insertion in one of these genes, encoding a presumed phospholipase, as the cause of HI in maize. The gene was named *NOT-LIKE-DAD (NLD)*, *Phospholipase A 1* (*PLA1*), and *Matrilineal* (*MTL*) [30,69,74]. The name *MTL* is used here.

4.2. *qhir8*

The second QTL, *qhir8*, could explain 12-34% of genetic variance in populations derived from crosses between two haploid inducers that are fixed for *qhir1* [68]. The *qhir8* QTL was fine-mapped to a 789kb region that increased HIR and embryo abortion rate, but unlike *qhir1*, it did not cause segregation distortion [70]. HIR and embryo abortion were inextricably linked. The *qhir8* was further fine-mapped to a 138kb region and, finally, narrowed down to a 318bp region harboring a non-synonymous amino acid change in the *ZmDMP* (DUF679 domain membrane protein) gene. This lesion was shown to be responsible for HI, and verified using CRISPR/Cas9 knock-out of *ZmDMP,* and complementation [71]. On its own, *dmp* knockout induced haploids in the B73 inbred genetic background at a HIR of 0.1-0.3%. On the other hand, *mtl* HIR is 2-3% but it increases to ~10% when combined with *dmp* [71]. The synergistic increase in HIR caused by the *dmp-mtl* double mutations suggest that these mutations affect different pathways converging on (or affecting) the same HI trait.

5. Maize HI Genes

5.1. Phospholipase A1 (Matrilineal or NOT-LIKE-DAD)

The MTL (a.k.a NLD or PLA1) gene in *qhir1* HI QTL is expressed in pollen [30,69,74,75]. Introgressing *mtl* from Stock 6 background into the non-inducer background of the A188 inbred line through backcrosses and selfing gives a paltry, but significant induction rate of ~0.81% in BC2S1 plants [69]. This was attributed to the change in genetic background, i.e. in the loss during backcrossing of other favorable Stock6 HI QTLs. Thus, on its own, *mtl* causes low rates of haploid induction in maize.

The effect of genetic background was also demonstrated when transfering HIR to other species. Knocking out the *MTL* ortholog in rice and wheat causes seed abortion and haploid induction HIR as high as 6 -19% [37,76,77]. As in maize, rice pollen fertility and germination were

not dramatically altered [6,28,32,76]. *mtl* knock-out in wheat caused 19% HIR, reduced seed set, and occurrence of seeds lacking the embryo or both the embryo and endosperm [77].

Preliminary studies had suggested that *MTL* localized to the sperm cell membrane and is expressed after pollen mitosis II (PM II) and during pollen tube growth [69,74,75]. However, a more thorough subsequent study has shown that it is expressed in the vegetative cell, and that the protein is localized on the membrane (endo-plasma membrane) that surrounds the sperm cells [75]. This was shown using fluorescently-labeled *MTL*, and by separating sperm cells from the endomembrane. The tight association between sperm cell-membrane and endo-plasma membrane might explain why early studies erroneously concluded that *MTL* was localized to the sperm cell membrane [75].

MTL localizes to a membrane but does not have a membrane targeting signal. Truncated *mtl* protein fails to localize to the endo-plasma membrane and remains cytosolic [69,75]. It is thought that positively charged protein domains interact with the negatively charged endo-plasma membrane to anchor the protein in or on the membrane [75]. Further, it was suggested that chemical, physical or genetic disruption of the endo-plasma membrane might be a good target for development of HI in other plants [75]. It is possible that *MTL* is involved in the process of gamete interaction, and *MTL* localization on the endo-plasma membrane surface is essential for completion of signaling, adhesion or fusion of egg and sperm cells during fertilization, or for separation of sperms from the male germ unit [78–80]. Recently, a reverse genetic study showed that mutating a *MTL*-like pollen gene, *ZmPLD3* (Phospholipase D3) results in HI [81].

5.2. DMP

DMP proteins (***D****UF679 domain* ***m****embrane* ***p****rotein*) are conserved, plant-specific proteins first discovered in Arabidopsis [82]. *DMPs* have four transmembrane domains and localize in membranes with both the carboxy and amino termini on the cytosolic side [62,82–84]. *Chlamydomonas reinhardtii* and *Physcomitrella patens* genomes have one DMP gene while

there are 5-13 *DMP* genes in dicots, and 11-16 in monocots. Arabidopsis has ten *DMP* genes - *DMP1* to *DMP10* - with tissue specific expression patterns [82].

While Arabidopsis *DMP9* pollen-specific expression had been shown by Kasaras and Kunze [82], its sperm-cell specificity and role in double-fertilization was first identified in a comparative proteomic analysis study of *Lilium longiflorum* mature pollen and generative cell microsomal fractions [62]. The *DMP9* transcripts accumulate after pollen mitosis I in the generative cell and then in the sperms, with the protein likely localizing on the sperm-surface rather than on the endo-membrane [62]. This study showed that RNAi knockdown of *DMP9* causes single fertilization events. A completely fertilized central cell was observed accompanied by an unfertilized egg cell with the compact sperm cell(s) adhering on its surface in ovules [62]. This caused seed abortion. Sometimes, more than two sperm cells are attached to an egg, suggesting polytubey [62]. *dmp9* single-fertilization events phenocopy other single-fertilization Arabidopsis mutants such as *DTA, CDKA1, FBL17, CAF1,* etc [85,86].

The above observations in Arabidopsis stimulated the hypothesis that *DMP9* acts after sperm attachment to the egg cell, and that it might be involved in the egg's recognition of sperm attachment [62]. *DMP9* is co-expressed with the sperm-specific sperm-egg adhesion protein, *GEX2 (Gamete Expressed 2*) [83]. *DMP8* and *DMP9* localize in the sperm cell membrane [83]. However, it might be that the protein is in the endo-plasma membrane as fluorescently tagged *DMP* identified "signals at the sperm plasma membrane, including the membranous extension associated with the vegetative nucleus" [83]. This uncertainty echoes the previous misidentification of *MTL* localization on the plasma membrane instead of the endo-plasma membrane, which was due to poor resolution [75]. The above studies with *dmp9* mutants did not look for haploids in their progeny.

The role of DMP in haploid induction was first studied in maize following the mapping of HI to a QTL that harbored DMP [71]. Knocking out DMP in non-inducer inbred maize resulted in low but significant HIR (0.1-0.3%), compared to 0% for the respective wild-type control. *dmp*

knockouts, which cause haploid induction in both maize and Arabidopsis, and accompanying complementation experiments confirmed the role of *dmp* in HI [71,84]. The Arabidopsis *dmp9* mutant induced haploids at a HIR of 0.4% while *dmp8* induced haploids at a HIR 0.03%. Combining these two mutations resulted in a HIR of 2% in *dmp8,dmp9* double mutant [84]. Arabidopsis *dmp8,dmp9* double mutants display normal pollen development and pollen tube growth, and reduced seed set, but only when used as male [83].

In maize HI, *dmp* utility results from its synergistic interaction with *mtl* to increase HIR [71]. Similar to maize haploid inducers, frequent single fertilization events were observed in Arabidopsis *dmp8-dmp9*: 24.5% of ovules had a developing endosperm but lacked an embryo [83]. This was due to a dramatic reduction of sperm-egg fusion, as only 43% of ovules were fertilized 42 hours after pollination (compared to 98% in controls). A large proportion of the ovules (57%) contained unfused sperms, sometimes with 2-4 sperms in the ovules with the central-cell almost always fertilized [83]. Normally, the double fertilization events happen within minutes of each other in Arabidopsis. On the other hand, Arabidopsis *dmp8-dmp9* sperm cells linger in the ovule long after (>1.5 hours) the central cell has been fertilized and the endosperm starts developing [62,83]. Curiously, these observations are consistent with those reported for *Solanum nigrum* x *Solanum luteum* HI nearly a century ago [65]. All these are signs of impaired sperm fusion that are reminiscent of Enaleeva's [64] observations of single fertilization events in maize [25,33,58,64].

In Arabidopsis *dmp8-dmp9,* the majority (59%) of unfused sperms that attached to the eggs were in pairs, leading to the hypothesis that there might be a problem with the two sperm cells separating after adhering to the egg [83]. Might it be that the endo-plasma membrane does not re-organize to release sperm cells for fusion? Might the fact that the sperm-pair does not separate explain the observation of mixoploidy and aneuploidy in endosperms in maize HI crosses?

The striking similarities between maize and Arabidopsis *DMP*-based HI argue that the cause of maize HI is aberrant double-fertilization, which also results in single-fertilization, polytubey and heterofertilization. The disruption of double fertilization is through impairment of egg-sperm fusion. This impairment of egg-sperm fusion might be physical (the membranes are unable to fuse), communicative (proper signals between egg and sperm are not relayed or detected), or both. It is unclear if *DMP*-based HI causes inducer chromosome addition. All the maize and Arabidopsis studies that have sequenced presumed haploid plants have not observed haploid inducer chromosome addition or introgression [56,84]. However, we can confidently conclude that both *DMP* and *MTL* are localized in sperm-membranes, likely the endo-plasma membrane, which helps explain why combining *qhir1 (mtl)* and *qhir8 (dmp)* increases maize HIR synergistically. Different pathways converge on the membrane.

5.3. Phospholipase D3 (PLD3)

To test the role of cell lipid membranes in the process of haploid induction, Li *et al* [81] performed a reverse genetics study. They identified a pollen specific phospholipase, Phospholipase D3 (PLD3), expressed highly in mature pollen grains. Knocking out the PLD3 resulted in self haploid induction in a non-inducing genetic background at a HIR of 0.9%. This rate is similar to that of the *mtl* mutant (HIR=1.19%). A double knockout of *mtl* and *pld3* resulted in synergistic haploid induction rate from around 1% for the single mutants to 4.13% for the double mutant. The authors could not obtain a triple mutant of *mtl-pld3-dmp*. However, the double mutant *mtl-dmp* in a heterozygous *pld3* (+/-) background resulted in a HIR over 7%. The PLD3 gene is not within a region known to have a HI QTL, but this may be explained by lack of variation at this locus [81]. This is the first reverse-genetics study to identify a novel haploid induction gene. The fact that a triple mutant could not be obtained for the three genes points to their essentiality in double fertilization.

6. Possible mechanisms for maize HI

As suggested here and in the literature, maize HI is likely the result of defective interaction between the sperm and egg membranes [30,69,74,87]. To understand how this might result in haploid induction, we need to first consider the mechanism of double fertilization and fertilization recovery using Arabidopsis as a model [59]. The ovule contains two synergid cells whose signals attract the pollen tube to the ovule. The pollen tube enters the ovule through the micropyle and bursts to release the two sperm cells leading to subsequent degeneration of one of the synergid cells. This synergid is called the receptive synergid while the other intact synergid is called the persistent synergid.

Within 4-10 minutes of each other, one sperm cell fuses with the egg cell and the other one fuses with the central cell [75,88] . The binary fusion events, specifically egg-sperm fusion, prevent attraction of additional pollen tubes to the ovule [67]. The surviving synergid persists for some time, but is incapable of pollen tube attraction. This suggests that successful egg-sperm fusion or events following fusion emit signals that block the persistent synergid's ability to attract pollen tubes. Either production of the pollen tube attractants by synergids is blocked or the signals are neutralized. The fertilized egg, now a zygote, gains competency for division and growth into an embryo. Expression of male specific or imprinted genes such as BABYBOOM (BBM) is known to activate the egg to develop into a growth-competent zygote [11].

However, lack of gametic fusion (as in Arabidopsis *duo1, dmp8-dmp9, hap2* (*gcs1*), *gex2*, or *duo3* mutants), triggers a process called fertilization recovery. After a long delay, the persistent synergid attracts a second pollen tube [59,60,67,89]. This second pollen tube delivers two sperm cells to the ovule, one of which will fuse with the egg. The result is a heterofertilized ovule. Following successful egg-sperm fusion in the second attempt, the persistent synergid degenerates and there is no further attraction of pollen tubes. However, if fertilization by the second pollen tube fails, the persistent synergid will continue attracting more pollen tubes [61,67].

Fertilization recovery is thought to be an evolutionary mechanism that increases the chances of successful formation of a seed.

For successful egg-sperm fusion, there must be communication between the egg and the sperm. The egg secretes small-peptides, called *Egg Cell 1* (*EC1*), that are recognized by the sperm. The sperm responds by expressing the surface fusogen protein *HAP2* (a.k.a GCS1; Hapless 2 or Generative Cell Specific 1) [88,90], which is required for the sperm's fusion with the egg [59,91]. Normally, the sperm cell pair delivered by the pollen tube is enclosed in an endo-plasma membrane and is deposited in a region between the egg and the central cell [92]. Reception of *EC1* by the sperm cells results in sperm membrane reorganization. As part of this reorganization, the sperm *HAP2* protein localizes from the endomembrane to the sperm cell membrane, and activates sperm for fusion [88,93,94]. The five *EC1* genes in Arabidopsis are functionally redundant. Knocking out any one of them does not result in a change in phenotype. However, simultaneously knocking down these five EC1 genes results in deposition of multiple sperm cells that do not fuse with the egg in the ovule [8893]. This outcome phenocopies *hap2, duo1*, *duo3, dta, cdka1, fbl17, caf1* or *dmp8-dmp9* mutants whose sperms are incapable of fusing with the egg [59,83,91,95,96].

In summary, the above observations strongly suggest that disruption of communication between the egg and the sperm can block gamete fusion. In turn, failure of gamete fusion stimulates heterofertilization. As already discussed, maize HI is caused by mutations in an egg-sperm fusogen (*DMP*) and in a sperm endo-membrane-localized protein (*MTL*). Knocking-out DMP in Arabidopsis results, as in maize haploid inducers, in protracted persistence of unfused sperms in the ovules [33,62,71,83,84]. We can also reconcile HI and HI-associated traits that include single fertilization events, polytubey, heterofertilization, seed death, and segregation distortion [21,25–27,64]. Polyembryony is more difficult to explain. Perhaps, failure of egg-cell fusion can activate remedial pathways that recruit somatic or egg sac cells into embryogeny. More

importantly, how does HI happen given the knowledge of events leading to HI? Specifically, is maize HI due to genome elimination, parthenogenesis, or neither?

I hypothesize the following. First, *dmp* and *mtl* knockouts contribute in different ways to the function and performance of the membrane in gamete fusion, and are thus synergistic. This is similar to the Arabidopsis single *dmp8* and *dmp9* mutations that result in low HIRs of 0.03% and 0.4%, respectively. The HIR elevates synergistically to 2.11% in the *dmp8-dmp9* double mutant due to the redundancy of the two DMP genes [84]. The maize *dmp* and *mtl* interaction is not due to redundancy, but to the likely convergence of two different pathways on the same trait. A single mutant is less effective at HI due to compensatory mechanisms. The interaction between the two mutations (*mtl* and *dmp*) might occur as follows. If *mtl* triggered ROS damage the genome and prevent its transmission, such damage may be time-dependent. The *dmp* mutations could slow down fertilization allowing excessive damage and increasing the rate of chromosome transmission failure. Other compensatory mechanisms may function through the multi-layered regulation in gamete signaling, gamete recognition, gamete adhesion, and gamete fusion.

Second, defective *dmp* or *mtl* proteins lead to lack of proper communication or physical interaction between the egg and sperm, which results in delayed or stalled gamete activation, fertilization, or zygotic activation. Third, events following these defects lead to haploid induction according to one of the models discussed below. These models have some overlap. A good HI model must account for and explain the following:

1. Abnormal fertilization (single fertilization and heterofertilization)
2. Occasional chromosome transfer
3. Transmission of CRISPR factors (either as RNAs or nucleoprotein) to the egg during HI

The following models are proposed or considered.

1. Abnormal egg fertilization

    1.1. Zygotic induction

1.1.1. Delay-triggered (time-dependent egg proliferation)

1.1.2. Pseudo fertilization (sperm leakage of activating factors)

1.1.3. Polytubey and synergid degeneration (synergid egg-block release)

2. True fertilization with subsequent genome elimination

3. Semigamy

As explained below, I favor the egg zygotic induction model through pseudofertilization (Figure 1 and Figure 2)

6.1. Abnormal egg fertilization

Abnormal egg fertilization with or without post-zygotic induction accidental transfer of cytoplasm and chromosomes could occur in one of the following ways.

6.1.1. Model 1: Delay-triggered (Time-dependent Egg proliferation)

In certain cases, even with polytubey mediated by fertilization recovery through the persistent synergid, the egg-sperm fusion does not take place, presumably due to the sperms being defective in their fusion or signaling proteins. Lack of proper fusion delays the acquisition of zygotic fate by the egg. As a final attempt to form a seed, the egg acquires zygotic fate dependent on the extended time. Thus the egg acquires zygotic identity independent of sperm contribution and undergo mitosis to form a parthenogenetic haploid embryo. At the same time, fertilization-dependent endosperm development proceeds while the egg-sperm fusion is stalled.

Once zygotic division starts, if a sperm is adhering to the egg, the parthenogenetic zygote may capture sperm components via membrane fusion during cell division-based membrane reorganization. It may be even possible that male chromosomes remain in the zygotic cytoplasm and form micronuclei. These extranuclear paternal chromosomes would be incompetent for cell division (are improperly condensed, have inactive centromeres, or are

unstable in the cytoplasmic environment). Within the micronuclei, they are susceptible to attack by nucleases and either undergo elimination or occasionally are rescued into the nucleus proper. Occasionally, if paternal chromosomes are rescued, aneuploids or mixoploids would result. In this model, sperm attachment to the egg is not a prerequisite for the zygotic activation. Rather, the central feature to this model is the protracted time the egg is unfertilized, and the autonomous acquisition of zygotic fate Figure 1. This model explains single-fertilization events (and heterofertilization), transfer of pollen-derived CRISPR components, and occasional addition chromosomes in otherwise haploid backgrounds.

6.1.2. Model 2: Pseudo-fertilization (sperm leakage of activating factors)

In the pseudo-fertilization model, the sperm attaches to the egg cell and provides a zygotic activation factor. The sperm-membrane and egg membrane interact sub-optimally, mimicking *proper* plasmogamy or karyogamy. The transfer of sperm- signals (e.g mRNA, small RNAs, proteins, peptides, or small molecules) resembles that taking place during successful gamete fusion. Thus, the egg acquires zygotic fate. Stalled, improper, or incomplete egg-sperm fusion results in partial transfer of sperm cytoplasm or sperm membrane components to the egg during zygotic cell division (Figure 1 and Figure 2). These events and outcomes resemble those in Model 1 and 2. This model explains abnormal fertilization events, transfer of CRISPR components and occasional inheritance of haploid inducer chromosomes.

The central feature of this model is the incomplete plasmogamy and possibly karyogamy triggering signals for zygote activation.

6.1.3. Model 3: Polytubey and synergid degeneration (synergid egg-block release)

In a third model, in addition to attracting pollen tubes to the ovule, synergids may release factors that block the egg from acquiring zygotic fate: I call these “synergid egg-block”. This

block is only overcome when the egg is fertilized. Successful fertilization blocks the attraction of a second tube to the ovule and causes the degeneration of the synergid [61,67,97].

With a persistent synergid, the egg cell does not develop unless it is fertilized. Both synergids have been shown to degenerate even without egg fertilization [67]. After attraction of multiple pollen tubes, without subsequent egg fertilization, both synergids may degenerate and the synergid egg-block signal dissipates. The egg could then acquire zygotic fate and develop into a parthenogenetic haploid embryo (Figure 1 and Figure 2). Testing this model might help determine if this mechanism exists and functions in ensuring formation of a seed (after unsuccessful fertilization) as opposed to having no seed at all. Even though haploid plants have lower fitness, they are known to produce viable seeds through either somatic tissue doubling or all univalents segregating to the same pole in meiosis I [4,98–100]. Dihaploids from tetraploid plants like potato produce viable seeds.

The central feature to this model is programmed synergid death due to attraction of multiple pollen tubes with concurrent loss of synergid egg-block. The synergid egg-block release model is not mutually exclusive from the time-dependent fertilization recovery.

6.2. Model 4: Uniparental chromosome elimination (Genome elimination)

This is the simplest explanation for maize haploid induction. Defective *MTL* or *DMP* proteins lead to unstable membranes, which can lead to genome instability before or after gamete fusion via i) direct DNA damage or ii) malfunction of sperm proteins required for genome stability, such as kinetochores, or DNA replication factors, or nuclear membrane function. Unstable membranes can expose DNA to the DNA-toxic cytoplasmic and extracellular environments (including in the ovules where the mutant sperm linger). Damaged chromosomes would fail replication and/or missegregate during zygotic and embryonic cell division.

Additionally, if the timing of zygotic activation is off with respect to karyogamy, late incorporation into the egg nucleus would cause the paternal chromosomes to be *late* for zygotic divisions, leading to instability of these chromosomes and their eventual elimination. This would be similar to haploid induction in Arabidopsis through *CENH3* manipulation (Figure 1) [4,101,102].

The central feature of this model is genome instability as a result of membrane abnormalities or delayed karyogamy. This model does not account for the single fertilization and heterofertilization events sufficiently. It could be argued that sperms that are undergoing DNA-damage and membrane instability have compromised cellular functions including transcription, translation, and signaling, are incompeptent of carrying out their biological functions including fertilization.

6.3. Model 5: Semigamy

Semigamy is "the phenomenon of independent division of both the sperm nucleus and the egg nucleus in the egg cytoplasm" that has been reported in several genera [103,104]. Semigamy is a mechanism of HI that does not seem applicable to maize. HI due to semigamy was considered a possible mechanism by Xu *et al*. [28] but the authors concluded that chromosome elimination was the most plausible mechanism. Semigamy gives rise to diploid, haploid, or sector diploid-haploid plants in cotton. The haploid from a single cross can be either paternal or maternal, or maternal-paternal haploid chimera [104,105]. Similar phenomenon has been observed in cacao [106]. Based on the extensive data discussed above, semigamy does not seem like a plausible mechanism for maize HI. However, it is a good explanation for polyembryony.

7. Conclusion and Future directions

The set of events between parental meiosis and filial embryo are concatenated processes along a continuum that involves precise signaling and the interaction of matched

components. Any deviation from the normal trajectory and disruption of part of the process could result in improper chromosome inheritance, such as during HI. In this chapter, I have reviewed 60 years of literature on maize HI and created models that attempt to explain maize Stock6-based haploid induction. Models 1-4 have considerable overlaps and fulfill the criteria set above for a good model explaining haploid induction in maize. These models are applicable to both potato and maize haploid induction systems. Maize HI (discussed here) and potato HI (next section) have striking similarities; specifically, a fertilized endosperm being accompanied by an unfertilized egg or haploid embryo early in seed development, abnormal pollen morphologies that do not affect pollen germination, and very low rates of paternal DNA contamination [49,107–109].

The pseudo-fertilization model is the most attractive. It is bolstered by the observation of sperms attached to the egg in both the maize haploid inducers and *dmp8-dmp9* Arabidopsis mutants [64,83,110]. The time-dependent egg proliferation and synergid egg-block release models need to be explored further. These two models predict that abolishment of egg fertilization (with or without sperm-egg adhesion) would lead to haploid induction. This has not been observed yet. Studying egg-fertilization mutants, female or male mediated, would help explore the models [88,92,93,111].

All the models discussed above can guide future experiments to study gamete dynamics, double fertilization, and haploid induction. The knowledge gained therein can be used to make more efficient haploid inducers in maize and other plant systems. One low-hanging fruit is to combine known single-fertilization mutants such as *cdka1, gex2, hap2, et cetera* with *mtl* and *dmp* both in monocots and dicots in attempts to increase haploid induction.

The knowledge of how these systems work could be expanded by transcriptomics, proteomics, and protein-complex analysis in pollen generative nuclei and sperm, and ovules obtained within the normal fertilization windows (temporal and developmental) following pollinations using inducer and wild-type pollen. *MTL*, *DMP*, or *PLD3*-based antibodies or

protein-tags would allow us to isolate interaction partners and complexes that are involved in the process of double fertilization. Transcriptomics work can be used to create co-expression networks that can help predict *DMP* and *MTL* protein partners, as well as identify differentially regulated genes between inducers and non-inducers.

Another simple, but very effective method is a suppressor screen experiment. Stock6-based haploid inducers have high rate of embryo and kernel abortion. After EMS mutagenesis on these haploid inducers, one could search in the M1 and M2 populations for mutants that restore kernel and embryo development . In maize, full cob analysis would be effective. These relatively simple and cost-effective experiments might allow identification of both male and female factors that influence haploid induction.

Chapter 2: A Review of Potato Haploid Induction (HI) & the Evolutionary Significance of HI

1. Brief introduction

Potato is an important food crop. Wild potato was consumed in South America as early as 12,500 years ago and was domesticated from wild tuber-bearing Solanum species about 7,000 years ago [112,113]. The domestication took place in the Andes mountain highlands in the region between Peru and Bolivia, with a possible second domestication in Chile[112,113]. Potato has more than 200 wild species, the most of any crop plant [112]. The wild relatives occupy multiple highly diverse environments of South America and are usually found in mixed-ploidy populations that range from diploid to hexaploid species [112,113,114 a,115 b].

Modern commercial potato varieties are autotetraploids derived from a few plants that were imported from South America to Europe, and then introduced around the world [116,117]. Commercial potato varieties, therefore, have a very narrow genetic base. Additionally, some of the most important varieties such as Russet Burbank, the world's foremost french fries potato variety, have not been improved since they were first released more than a century ago [118]. Potato varieties are heterozygous, with a high genetic load, and are propagated vegetatively through tubers to maintain this heterozygosity and agronomically important trait combinations [119,120]. Most of the wild potato accessions and species are diploid, and contain a wealth of genetic resources that can most readily be accessed through introgression at the 2X level [121,122]. Potato breeders have come up with ingenious methods of breeding potato that overcome ploidy barriers and avoid formation of triploids that are expected from tetraploid x diploid potato crosses. This involves use of unreduced gametes in diploid backgrounds to make tetraploid hybrids. However, the most practical method is the use of potato strain strains called haploid inducers. Haploid inducers are used to obtain dihaploids, which are haploids derived from

tetraploid parents. Dihaploids are diploid, and are used as genetic bridges between the cultivated and diploid wild accessions. The dihaploids are critical because they also allow breeding at the diploid level, which helps avoid the tetrasomic inheritance connected to tetraploid breeding. Haploids from polyploids are called polyhaploids. Specifically, haploids from tetraploids and hexaploids are called dihaploids and trihaploids, respectively, while those from diploids or dihaploids are called monoploids (monohaploids) [123–125].

In comparison to the well-studied maize haploid induction system, very little is known about potato haploid induction despite the two systems being discovered around the same time in the 1950s, and their universal adoption in breeding programs. One of the earliest reports of a potato haploid was by Ivanovskaja [126]. This haploid was obtained by pollinating a tetraploid potato variety with a diploid male [127]. Additional reports of haploid plants derived from genetic crosses involving hexaploid or tetraploid potato species appeared in the 1950s [127–130]. Following t[125]hese developments, particularly with one tetraploid variety (Katahdin) with 10% haploids among its progeny, it was suggested that breeding potato at the diploid level might be advantageous and should be prioritized [130,131]. The resulting efforts led to identification of selected diploid clones derived from a single *S. tuberosum* group Phureja accession (P.I. 225682) that were efficient at haploid induction [132].

Most of the literature prior to 1970 that reported potato haploids used various diploid potato accessions or species as haploid inducers. In particular, Phureja accessions have been cited extensively as capable of inducing haploids. Due to taxonomic classification challenges, different Phureja accessions that were used to produce haploids bore, prior to 1970, different names such as *S. rybinii*, *S. ascasabii, S. stenotomum*, *S. tuberosum* Group Stenotomum, or *S. tuberosum* Group Phureja [*128,129,133–136*]. Other names used to refer to Phureja in the past include *S. boyacense, S. canariense, S. cardenasii, S. kesselbrenneri* and *S. phureja* [137]. According to the US Department of Agriculture taxonomists additional names include *S. apurimaceme*, *S .goniocalyx*, *S. herrerae*, *S. hygrothermicum*, *S .mammilliferum*, *S. chaucha,* and *S. chiloense* [138],.

Recently, the entire Phureja group has been collapsed into *Solanum tuberosum* supsp. Andigenum according to taxonomists at the United States Department of Agriculture Potato Germplasm Repository [138]. All these accessions will be referred to as Phureja in this publication (despite previous or current classification). Additionally, other distinct, diploid non-Phureja wild potato species such as *S. torlapanum* (*S. toralapanum*), *S. microdontum* (formerly *S. simplicifolium*), and *S. chomatophilum* are also known to induce haploids [128,129,139–141].

Haploid induction in potato is not confined to polyploid x diploid crosses. For example, a haploid plant has been derived from tetraploid *Solanum chaucha* (now classified as *Solanum tuberosum* supsp. Andigenum) by crossing it to a tetraploid *Solanum tuberosum* [142]. However, this is not a typical haploid induction cross, and the haploid obtained was a polyembryonic twin of a tetraploid. That a haploid was found in a polyembryonic seed is not surprising. Polyembryony is associated with HI in maize, and *Gossypium* inter-specific haploid induction, and in many other species during genetic, chemical or physical haploid induction 8.

Haploid induction in potato, just like maize haploid induction, is dependent on both the female and male genotype [132,140,143,144]. To fully utilize haploid breeding, a rapid screening procedure for potential haploids using dominant genetic markers was suggested by Peloquin and Hougas [145] and implemented by Hermsen and Verdenius [146]. Some of the resulting haploid inducers, IvP-35, 48 and 101, have since then been used in potato breeding programs. They are diploid Phureja lines homozygous for genes that deposit anthocyanin on internodes. This trait appears on the embryo as a spot that is visible on the seed thus providing a simple, cheap, and quick screening tool for putative haploids (Figure 3) [144,146] .

Tetraploid varieties that are haploidized are usually *S. tuberosum* group Tuberosum or *S. tuberosum* group Andigena. These tetraploids are recessive for the seed embryo spot trait. Following a tetraploid x haploid inducer cross, all spotted seeds are presumed hybrid, and are therefore discarded. The non-spotted seeds are grown out and tested for haploidy to distinguish haploids from accidental selfs or hybrid escapes.

Even though haploid inducers are typically used with the tetraploid varieties, haploid induction is not restricted to tetraploid x (diploid) haploid inducers. The haploid inducers can also induce haploids on diploid (including dihaploid) potatoes and on some wild potato species [147–151]. Similarly, HI can occur in hexaploids such as *S. demissum,* and octoploids (chromosome doubled tetraploids) *S. stoloniferum* and *S. fendleri* [128,129,152]. In this section, the term 'HI cross' will refer to the 4X x 2X haploid induction cross. In all these potato crosses, haploid induction rate (HIR) is measured as the number of haploids per berry (h/b).

The mechanism of haploid induction in potatoes is poorly understood. Two competing hypotheses have been put forth to explain haploid induction in potatoes: parthenogenesis [107,109], and genome elimination [153–155]. We will discuss these, as well as propose new alternative hypotheses.

2. The Genetics of potato Haploid Induction

The haploid induction trait is genetically determined and can therefore be inherited and bred into various backgrounds [132,140,146]. Gabert [132], while studying the ancestor of modern haploid inducers, concluded based on limited data that haploid induction is a recessive trait controlled by a single gene. He studied seventy seedlings of heirloom variety, PI 225682, collected in a local market in Nariño, Colombia in 1948 [156]. He identified three plants named Phureja (Phu) 1.1, 1.13, and 1.22 (or plants 225682.1, 225682.13 and 225682.22 respectively) with HIR that were 10-63 fold higher than in the rest of the material. He called these "superior" haploid inducers. Phu 1.22 has been clonally maintained to this day and is deposited at the USDA Germplasm Collection. It is the ancestor to all modern haploid inducers [144,146]. Based on the fact that only three of the seventy plants analyzed (less than 5%) were superior haploid inducers, I conclude that both low and very high rates of haploid induction segregate in this accession [132]. This trait (high rates of HI) is heritable and therefore genetically controlled.

Van Breukelen [151] studied a population bred by Hermsen [136] for the haploid-screening embryo spot trait. He concluded that haploid induction is controlled by multiple genes, and that the genes were semi-dominant. This population was derived from a cross between two accessions - the embryo spot parent (accession PI 225702, low HIR = 0.08 h/b) and the HI parent (Phu 1.22, relatively high HIR = 0.4 h/b) [146]. PI 225702 was collected in 1948 at a market in Valle del Cauca, Colombia [157]. HIR within the F2 population varied widely. Some of the F2 individuals in the population exhibited higher HI rates than either of the parents.

In yet a different experiment, Hermsen [136] crossed a moderate haploid inducer (HIR = 0.125 h/b) to a wild type (HIR = 0.0 h/b) and obtained both superior (HIR= 0.90 and 0.462 h/b) and inferior (HIR = 0.024 and 0.055 h/b) haploid inducers in the F1 [136]. Similarly, Van Breukelen [151] crossed plants with different rates of HI and obtained analogous results. I, too, observed similar outcomes in the population I characterized (see chapter 3). These observations suggest that more than one gene in Phureja causes haploid induction. A weaker HI locus (or loci) of PI 225702, potentially heterozygous, caused the low haploid induction [132,136,146,151]. It is possible that the very high rates of HI in some of the IvP lines involve synergism with this second, weaker HI gene. Table 1 lists these historical lines.

Besides PI 225702 and PI 225682, there are several other accessions that have been reported to induce haploids at low rates including Phureja accessions PI 225672, PI 195198, PI 225683, PI 273453, PI 225698, PI 225675, and several other poorly defined accessions; and finally other close wild potato species such as *Solanum simplicifolium* (now reclassified as *S. microdontum*; PI 218222), *S. vernei, S. megistacrolobum, S. macmillanii (S. millanii?), S. canasense and S. tarijense* [128–130,133,139,140,158–163]. It seems that the genes favoring haploid induction exist in various apparently unrelated Phureja accessions. Whether these genes and alleles are the same is unknown. It is also not clear whether these Phureja HI genes and alleles are related to the ones that cause HI in the wild potato species.

Phureja accessions are maintained by sib intercrossing. Phureja accession PI 225682, the source of modern high HIR haploid inducers, had 4.2% of individuals displaying very high rates of HI unseen in any other accessions. The majority of the individuals displayed low HIR comparable to HI in other accessions. I hypothesize that accession PI 225682 is polymorphic for the widespread HI gene(s) that cause low rates of haploid induction. The HI alleles at this locus are ancestral and common in Phureja. I call this HI locus Basal (B). The superior haploid inducers, on the other hand, are caused by an uncommon allele at a different gene. I call this high HIR gene H. In summary, I hypothesize that there are two major-effect QTLs, and that one of the genes (H) is semi-dominant with larger effect (hh=0, Hh=some HI, HH= highest HI). The B locus action is unclear, but could be recessive.

3. Possible Mechanisms for potato haploid induction

Considering the available information, the nature of haploid induction in potato remains an open question whose answer has multiple implications. Is haploid induction caused by a rare genetic mutation, a low-frequency allele that exists in Phureja populations, a common allele in Phureja populations, or a combination of these? Is there an evolutionary explanation for the presence of these haploid inducing alleles in populations? Is the occurrence of haploid inducing alleles in wild populations common in other genera and families? Here, I discuss two hypotheses that explain the presence of HI alleles in many potato accessions, and the nature of HI in general: The fertilization barrier hypothesis and the Ploidy-Flux hypothesis.

3.1. Evolution of D-M (Dobzhansky-Muller) incompatibilities

The fertilization barrier hypothesis is based on the Dobzhansky-Muller incompatibilities (DMI) model [164]. The fusion of an egg and sperm, which is essential for sexual reproduction, is precisely co-ordinated. The DMI hypothesis postulates that dsyfunctional interactions between diverged factors (from individuals with different evolutionary histories) involved in sexual

reproduction compromise plasmogamy, karyogamy or other sexual development programs. While compromised, the interactions could still succeed in activating the zygotic phase of the egg and in production of haploids. The interlocus contest evolution (ICE) theory hypothesizes that male and female factors evolve rapidly within a population [165]. To maintain species barriers, co-evolution of reproductive traits within any species is expected to be rapid compared to non-reproductive traits [166]. The male and female signals and proteins that are involved in gamete interaction co-evolve to result in an optimal gamete fusion process within a species, but may fail when interspecific gametes interact due to divergence and mismatch of cognate interacting factors [166]. As divergence increases, failing interactions between gametic factors may be sufficient to trigger egg development through transmission of egg-activating factors or by ligands binding receptors that activate the egg. This would result in pseudo-fertilization (Figure 1 and Figure 2). Another example of the outcomes of suboptimal interaction of cognate factors from different populations is the control of seed (endosperm) size of hybrids [167].

It is well-established that interspecific crosses can result in haploid induction. This has been observed in the grasses including wheat (*Triticum*) species, barley (*Hordeum*) species, wheat x maize, triticale x maize, rye x maize, oats x maize, sorghum (*Sorghum*) species, *Aegilops longissima x Triticum durum*, wheat x pearl millet, wheat x Sorghum, wheat x *Aegilops cylindrica,* and wheat x *Imperata cylindrica*; and in eudicots including in *Datura* species, *Oenothera* species, *Solanum nigrum* x *Solanum Luteum*, *Nicotiana* species, *Crepis* species, and Cotton (*Gossypium*) species; and close potato relatives *S. polytrichon* x *S. stoloniferum*, *S. chacoense* x Phureja, and *S. acaule* x Phureja, *S. verrucosum* x Phureja, *S. etuberosum* x Phureja, *S. bulbocastanum* x Phureja [136,139,168–177]. The fact that this is such a common occurrence in different inter-specific crosses, including many between Solanaceae species, suggests that the mechanism underlying HI in different potato crosses might be similar. In this scenario, reproductive factors between different species or populations have diverged. In potato, the diploid Phureja lineage and the tetraploid populations may be speciating but have

not evolved enough to have robust species barriers to prevent hybridization. Potato species and accessions coexist in a series of ploidies including diploids, tetraploids, and hexaploids with extensive gene flow that blurs taxonomic ("*accession*") boundaries [137,178–180].

In the above scenarios, emerging interspecific reproductive barriers result in high seed lethality and haploids from selected crosses. A candidate for this type of haploid induction is observed in wide-crosses in *Solanum nigrum* x *Solanum luteum .* Phenotypically, it is similar to apogamy in *Zephyranthes texana* (now known as *Habranthus tubispathus* or Atamosco Lily) [65,66]. Some features of haploid induction in wide crosses may be similar to haploid induction in maize and potato [33,36,109]. Old reports suggested that the pollen tube enters the ovule, delivers sperm that attaches to the egg, but either never fuses with the egg or with the nucleus [65,168]. This could be explained if the species differ in ligand-receptor specificities responsible for proper communication and fusion between egg and sperm. In the case of Atamosco Lily apogamy and of several other species, the sperm delivers its nucleus in the egg but karyogamy does not take place [66]. The connected ligand-receptor interaction may be such that even without karyogamy, there is sufficient egg-sperm interaction to activate the egg's zygotic fate during Atamosco Lily's apogamy. Similar occurrences could theoretically take place in haploid inducing crosses as a result of faulty gamete interactions that can sometimes result in haploids, hybrids and undeveloped seed.

Consistent with progressive divergence, wide-crosses that result in haploid induction seem to be more efficient between closely related species (usually in the same genus) such as among *Gossypium* (cotton), *Hordeum, Datura*, *Nicotiana* species, *Crepis*, and *Oenothera* species [8,65,136,139,168,171,172,174,181,182].

The higher the parental divergence, the less effective haploid induction or hybridization should be due to weakening of interactions. Indeed, haploid induction in all reported inter-genera crosses requires embryo rescue, usually due to endosperm failure 2. Rescue is not required for some of the intra-genera interspecific crosses. The co-evolution of mating factors

would also be expected in the central and sperm cells interaction, and hence affect endosperm formation. This is not a new idea. According to Sophia Satina, “It must be confessed that the more distinct two taxonomic groups are one from the other, the more difficult it generally is to secure hybrids between them” [183].

Haploid induction, in these cases, is not caused by a random mutation that is either neutral or deleterious. Instead, it is caused by divergence of coevolved factors that cannot combine between populations or that interact suboptimally.

The above considerations explain the occurrence of HI in distant matings. But, why would haploid induction persist in mixed-ploidy and sympatric potato populations that regularly have inter-ploidy and inter-*accession* matings? Why are haploid induction alleles not purged by negative selection? I hypothesize that weak, suboptimal egg-sperm interactions are under negative frequency-dependent selection in some species with mixed-ploidies such as potato. During negative frequency-dependent selection, "relatively rare variants have a selective advantage over more common variants and thus tend to increase in frequency and avoid local extinction” [184]. Note that the increase in the frequency of variants under negative frequency-dependent selection is inversely correlated with fitness. Alleles that compromise fertilization will be purged by purifying selection in uniform ploidy species or populations such as the diploid tomato. On the other hand, the HI alleles allow gene flow during the inter-ploidy intraspecific and inter-specific potato matings by ensuring formation of fertile euploid (diploid or tetraploid) progeny (Figure 4, and Figure 5). This property will be discussed below in my “Ploidy Flux hypothesis”.

3.2. Parthenogenesis and Pseudo-fertilization Hypotheses

The parthenogenesis hypothesis postulates that an unfertilized egg can initiate zygotic growth in the presence of a functional endosperm [109]. Wangenheim et al [109] was the first to suggest that the potato haploids are formed parthenogenetically and after a restitution 2n sperm

fuses with the central cell in the 4X x 2X HI cross . Nuclei restitution (restituted cell) is the formation of nuclei (cell) with double the normal chromosome number due to incomplete cell division [185–187]. Of about a hundred ovules analyzed following a potato HI cross, only one had a pentaploid endosperm and it was accompanied by a triploid embryo. 99% (96 of 97 ovules) of well-developed embryos had hexaploid endosperm (4X central cell + 2X sperm) showing that 2X=2n sperms from the diploid were responsible for most of the progeny [109]. Five ovules contained haploid embryos accompanied by hexaploid endosperm. This is consistent with central cell fertilization by a 2n sperm. Further, more than thirty ovules contained no developing embryo, but displayed a developing hexaploid endosperm [109]. This is reminiscent of maize haploid induction and Arabidopsis single-fertilization mutants where the endosperm initiates development while the egg remains unfertilized [33,58,62,64,83].

The above observations spurred the hypothesis that HI is caused by a mutation that results in restitution sperm formation during the second pollen mitosis (PM II) of a 1n pollen. The 2n sperm disrupts double fertilization, thereby inducing haploids [108,109]. Subsequent studies of pollen development in the haploid inducers offered further insight. Bender [141] observed restitution nuclei and chromosome bridges between the sperm nuclei after generative nuclei mitosis in the haploid inducer pollen [107], an observation confirmed by Montezuma de Carvalho [108].

Further support for the hypothesis that restitution sperm underlie HI comes from Montelongo-Escobedo and Rowe [107] , who studied the second pollen mitosis (PM II) in the Phureja haploid inducer. Normal pollen development results in formation of two sperm cells at the end of PM II. In contrast, up to 38% of the germinated haploid inducer pollen had a single PM II restitution nucleus compared to 2.8% for the non-inducer pollen. Interestingly, when the pollen from both the inducer and non-inducer was treated with colchicine and used for pollinations, the haploid inducer HIR increased more than 2X from 0.14 h/b to 0.33 h/b while that of the non-inducer increased from 0 h/b to 0.10 [107]. A restitution sperm was formed in 100%

of the treated pollen, associating it with haploid induction. This also provides a good explanation for the observation of hexaploid endosperms (maternal 4x central cell+ paternal 2x sperm) with dihaploid embryos in the 4X x 2X HI cross [107–109].

Pointing to a potentially conserved developmental program, blocking the second pollen mitosis or interfering with DNA integrity using chemicals or genetic mutations causes haploid induction whose disruption. Additional examples of haploid induction using chemicals exist. Toluidine blue, a dye that strongly binds DNA and RNA, some proteins, glycans and microtubules, and disrupts membranes induces haploids in poplar species, frogs and mice [188–201]. Treating pollen with Trifluralin pre-pollination prevents pollen mitosis II and leads to bicellular pollen that induce haploids in maize [202]. This pollen is similar to the one observed in potato haploid inducers, or potato wild-type pollen that induces haploids following colchicine or N_20 (nitrous oxide) treatment [107,108]. Thus, genetic or pharmacological effects might induce haploids through alterations of the sperm cellular structures (such as membranes) or the sperm genome. The parthenogenesis hypothesis is based on similar genetically-derived defects.

In summary, the parthenogenesis hypothesis states that normal (1n) haploid inducer pollen forms a single 2n restitution sperm, the pollen tube delivers a single sperm-like cell to the ovule, the restitution sperm fuses with the central cell, and the egg develops unfertilized.

I propose a substantial modification of this hypothesis. I call this alternative the **pseudo-fertilization** model. Instead of a single restitution sperm fertilizing the central cell and concomitantly inducing egg zygotic-induction and parthenogenesis, the delivery of a single restitution sperm induces polytubey as part of fertilization recovery [59–62]. A faulty interaction of the restitution sperm with the egg prevents proper fusion but could be sufficient to trigger zygotic fate in the egg cell. In the same ovule, the central cell is fertilized by a normal sperm from the second pollen tube. For tetraploid female potato, fertilization of the central cell by 2n (x=24) sperm is *ideal* as it restores the 3:1 maternal:paternal endosperm genome ratio.

Formation of the restitution sperm may not be causal for HI *per se.* Instead, a HI allele that affects PM II, also affects chromosome or membrane integrity, or other critical components involved in proper egg-sperm fusion. The pseudo-fertilization model is thus similar to the mechanism I proposed for maize *MTL* and *DMP*-based HI (Figure 1 and Figure 2). It assumes that HI is a result of random, deleterious mutations that affect critical cellular processes during sexual reproduction, and that have not been eliminated yet by negative selection.

3.3. Uniparental Genome Elimination Hypothesis

The uniparental genome elimination postulates that potato fertilization takes place but the haploid inducer chromosomes are destroyed or eliminated within the zygote or embryo nuclei. This is a relatively recent hypothesis. Until 1991, it was assumed that dihaploid or near-dihaploid progeny were devoid of haploid inducer chromosomes. By analyzing root tips of assumed haploid progeny from a potato HI cross, it was found that some cells were aneuploid with more than the 24 chromosomes expected in a dihaploid. AFLP analysis of such plants showed that upto 8.6% of thirty-five dihaploid or near-dihaploid plants had haploid inducer DNA [153]. This provided the first molecular evidence of residual haploid inducer DNA following haploid induction. Further research showed that some of these near-dihaploids had variable number of chromosomes within the same plant body, and expressed some of the haploid inducer isozymes, and had haploid inducer DNA as detected using PCR or sequencing [153–155,203–208].

It is difficult to evaluate the reports on presence of haploid inducer DNA in dihaploids because of their often conflicting nature and small sample size (8 to 35 dihaploids were analyzed). More recently, whole genome sequencing of HI cross progeny, including 923 dihaploid/near dihaploid individuals, produced at the International Potato Center and in the Comai lab have affirmed the observation of addition (introgression is less likely, but theoretically possible) of haploid inducer DNA in otherwise dihaploid plants [49]. Out of the 923 dihaploid/near dihaploid individuals, only 8 had haploid inducer chromosomes - two individuals had additional chromosome segments while the other six had additional whole chromosomes. This represents

a low rate (0.87%) of haploid inducer chromosome addition to otherwise dihaploid individuals [49]. This is 10-fold lower than the (possibly erroneous) 8.6% when a small sample (n=35) of dihaploids was studied using AFLP [205]. Observation of haploid inducer chromosomes in otherwise haploid individuals is usually attributed to incomplete genome elimination such as is observed in *CenH3*-based or interspecific crosses.

The most interesting aspect of the above WGS dihaploid study, however, is the genotype of the hybrid individuals[49]. Both triploid and tetraploid hybrids can be obtained in 4X x 2X potato crosses involving Phureja diploids [209–211]. In the analysis of triploid and tetraploid hybrids derived from the HI cross, it was found that tetraploid, but not triploid progeny displayed high levels of haploid inducer chromosome genome instability [49]. The observation of genome instability in progeny from unreduced pollen (tetraploids) but not from reduced pollen (i.e triploids) suggests involvement of 2n sperms (either 1n pollen's PM II restitution sperm or 2n pollen's 2n sperm). However, Amundson *et al* found no evidence for the involvement of the 2n restitution sperm nuclei postulated by Wangeheim et al [109] as the cause of HI [49]. Instead, analysis of centromere homozygosity in tetraploids' haploid inducer chromosome set was consistent with first meiotic division restitution[49]. Wangeheim ***et al***'s [109] PM II restitution sperm (or second meiotic division restitution sperm) would result in homozygosity of the haploid inducer chromosome set in tetraploid hybrids. This was not the case in the Amundson *et al* analysis [49]. Whether the paternal chromosome instability observed in both near-dihaploid plants and tetraploid aneuploids are related remains unclear as no controls (tetraploids from non-haploid inducers) were included in the analysis.

4. Evolutionary Significance of Haploid Induction

Cultivated and wild potato populations display three common traits affecting sexual reproduction - self-incompatibility, unreduced gametes, and haploid induction. Here, I

hypothesize that these three traits act cooperatively to increase genetic diversity and promote sexual reproduction in potato. I call this phenomenon “ploidy-flux".

4.1. Self-incompatibility

Solanum species display high diversity at the self-incompatibility (SI) locus. This locus is a haplotype composed of a maternally-expressed S-RNAse gene and multiple pollen-expressed SLF (S-Locus F-Box) genes. There are 20 or more alleles of the S-Loci [212–214]. The SI system works at the diploid level with each SLF protein (S-allele) recognizing non-self S-RNAse (i.e S1 recognizes S2 S-RNAse but not S1 S-RNAse). Recognized S-RNAses are degraded via the ubiquitin-proteasome pathway, allowing pollen tube growth. Unrecognized S-RNAses inhibit pollen tube growth through degradation of pollen RNA [212–214]. Consider an S1S2 pistil: while S1 pollen recognizes S2 S-RNAses and results in its degradation, it will be blocked by S1 S-RNAse. In the same way, S2 pollen will be blocked by S2 S-RNAse. Both S2 and S1 pollen will be inhibited. However, if S3 pollen lands on this pistil, it will recognize both S1 and S2 RNases and degrade them. As a result, the S3 pollen will grow and succeed in fertilization.

Thus, a given S-allele will prevent a plant from self-fertilizing and from fertilizing plants of similar genotype. Promotion of outcrossing by this system in Solanaceae (tomato, potato, petunia, etc) is thought to increase genetic diversity and hence reproductive success. The function of the S-alleles is conducive to evolution of additional S-alleles, which promote outcrossing and ensure wide matings. A similar incompatibility system exists in nearly half of all flowering plant species, and is thought to have evolved before eudicots [215,216]. There are many other self-incompatibility mechanisms (e.g the Brassicaceae sporophytic SI system) not dependent on the classical Solanaceae S-Locus [216].

While the Solanaceae SI system can be advantageous, founder effects, population bottlenecks, or a genetic sweep linked to a particular S-allele can have disastrous effects if they reduce S-locus genetic diversity to a couple of alleles. This would reduce reproductive success as most of the individuals will not produce seed. Interestingly, SI partially breaks down after the

genome doubling of self-incompatible plants, such as one with S1S2 genotype, due to formation of 2x pollen [136]. Three gamete-types are formed by the corresponding tetraploid: S1S1, S2S2 and S1S2 pollen. The S1S2 pollen is able to self-fertilize their duplex parent (S1S1S2S2) since both S1 and S2 S-RNAses can be recognized and degraded. S1S1 and S2S2 pollen can only pollinate the quadruplex genotypes (S2S2S2S2 and S1S1S1S1, respectively), but not the heterozygotes [216]. In conclusion, polyploidization can overcome SI-dependent infertility.

4.2. Unreduced gametes

In many potato wild and cultivated populations, there exist many loci that result in formation of unreduced gametes - both 2n eggs and 2n pollen [217–219]. While these loci are widespread and heavily used in potato breeding programs, their evolutionary function is not entirely understood. I hypothesize that these exist to promote reproductive success by doubling the genome of diploids through fertilization of a 2n egg by a 2n sperm. Genome doubling overcomes the lethal effects of a population with a fixed S genotype. However, it seems that the partial breakdown upon polyploidization is prevalent in species, such as Solanum, that have gametophytic SI systems, and less so in those with sporophytic or other SI systems [220]. In addition to fertilization between unreduced gametes, the doubling can occur through pollination of the diploid plants' unreduced egg with pollen from an existing tetraploid population, or the diploid-plants' unreduced pollen pollinating the existing tetraploid population. This promotes reproductive success in the following ways:

1. Doubling the genome of a diploid potato partially disarms the gametophytic self-incompatibility mechanism due to presence of heterozygous pollen thereby facilitating a switch from an outcrossing mode to selfing mode.
2. Switching from diploid to tetraploid state allows a population gene flow from other tetraploid populations. This introduces genetic variation, and critically, allows for inflow of other S-haplotypes that may be fixed or segregating in these established tetraploid populations.

3. Evolution of new alleles and genes is faster after whole genome duplication (WGD). Tetraploids that arose from a diploid population could evolve new alleles and genes, which can later be shuttled into the diploid population through HI.

Unreduced gametes in plants can form through multiple mechanisms, which include first division restitution (FDR), second division restitution (SDR), pre-meiotic doubling, post-meiotic doubling, etc. Unlike the other mechanisms, FDR maintains most of the original parent heterozygosity. For this reason, FDR has been proposed to be the most advantageous mechanism of polyploidization, and making a highly heterozygous and heterotic tetra-hybrid tetraploid (ABCD) from diploid (AB) hybrid x diploid (CD) hybrid [114 a,115 b,221,222]. Both FDR and SDR are well-documented in potato [114 a,115 b,218,223–229].

Polyploidization can be primary or secondary depending on its origin. Primary polyploidization occurs when an individual within a population autopolyploidizes while secondary polyploidization occurs when allopatric diploids meet and form an allopolyploid during hybridization [230,231].

4.3. Haploid Induction

Tetraploid x diploid crosses can result in diploid (dihaploid) seeds at very low frequencies (0.01 to 0.1 haploids/berry). Such rates are higher than it appears given that, due to high seed lethality, only five or fewer seeds per berry are formed from potato interploid crosses. Thus, haploids can easily constitute 5-20% of progeny. As previously mentioned, many accessions of diploid potato induce haploids when crossed to tetraploid females. While this has not been observed in the wild, it is well documented in crosses using tetraploid (or hexaploid) female x diploid males collected from the wild. Haploid induction occurs at very low rates in many other species, but the evolutionary significance of this has not been explored. I hypothesize that haploid induction exists as a complementary trait to unreduced gamete formation in a phenomenon I call ploidy flux.

4.4. Triploid Block

Cultivated tetraploid cultivars can produce triploids when crossed to diploid individuals [209]. However, triploids rarely occur in wild potato populations but can comprise a significant proportion of indigenous cultivated populations [209,232]. It is not surprising that cultivated populations have a high proportion of triploid. Since flowering and seed set is antagonistic to tuber formation [233,234], triploid tubers might be selected by farmers due to reduced triploid fertility, which promotes tuberization.

Unsurprisingly, natural or hand-pollinated tetraploid x diploid crosses are not as productive. A diploid potato ovary has about 600 ovules [235]. Upon normal pollination and fertilization, a diploid potato plant produces an average of 300 seeds per berry. On the other hand, the 4X x 2X interploidy crosses will have an average of about 1-10 seeds per berry while 2X x 4X crosses generally fail or produce very few seeds [135,136,209,236–241]. Interploidy cross progeny can be triploid or tetraploid. Surprisingly, tetraploid seeds tend to dominate the output of the 4X x 2X inter-ploidy crosses [135,235–237]. The triploid state is deleterious to seed development: for example, it leads to endosperm gene dosage imbalance. Endosperm gene dosage imbalance tolerance, however, varies with genotype and triploid seed can be common in some crosses. Interploidy crosses have an extremely high failure rate as only few of the potential ~300 seeds complete development [235]. While the role of the triploid block as a hybridization barrier is well recognized in the literature, its potential function in population biology and mating-strategy optimization has not been considered [242].

While the triploid block causes widespread lethality in interploid crosses, it facilitates production of both dihaploid (diploid) and tetraploid progeny by reducing competition from the predominant triploids on two fronts. First, reduction of competition for resources within the 600-ovule fruit, and later after seed dispersal and germination. With less competition and more resources, dihaploid and tetraploid seed are more likely to succeed in the fruit and can

contribute to the ploidy flux. This allows an isolated diploid population to auto-tetraploidize through unreduced gametes, and to later dihaploidize tetraploid individuals that might exist in the population without having to go through the triploid intermediate.

The evolutionary importance of triploids and pentaploids as genetic bridges (including in potato) between polyploid and diploid populations through the rare chance formation of 1X and 2X gametes is acknowledged, but will not be discussed here [113,114 a,115 b,221,243,244]. Going through an intermediate (triploid or pentaploids) that compete with the even chromosome-number euploids (diploids and tetraploids) for scarce resources may be less efficient by reducing the entire population's sexual-reproductive success.

5. Synthesis: Ploidy Flux.

Self incompatibility, unreduced gametes, triploid block, and haploid induction as means of increasing genetic diversity and reproductive success.

The Ploidy Flux hypothesis postulates that mixed-ploidy populations or populations with mixed ploidy potential can rapidly adapt and achieve the best evolutionary mating strategy by ploidy transition. Ploidy Flux optimizes fitness by switching between a primarily outcrossing mode for diploids and selfing modes for tetraploids in species with gametophytic self-incompatibility systems such as Solanaceae species. When genetic diversity is low, the population favors polyploids plants that have partially broken-down gametophytic self-incompatibility to allow the sexual reproduction. High genetic diversity favors diploids that are self-incompatible and that primarily outcross. Haploid induction, triploids, the SI-system and unreduced gametes are critical components of Ploidy Flux. I hypothesize that newly established populations would be primarily expected to be polyploid (due to low genetic diversity) while long-established populations with high genetic diversity would be expected to be mostly diploid.

Ploidy flux could work as follows: in a diploid population with two S alleles, formation of a tetraploid avoids the sterility-dependent extinction. This resulting tetraploid population can

persist and may eventually gain S-alleles by mutation or by gene flow. A tetraploid population with high diversity of the S-allele could return to diploidy through dihaploid formation bringing sufficient S-allele diversity to form a viable, outcrossing diploid population. In summary, the two traits, 2n gametes and haploid induction, while individually deleterious, can act in concert to provide flexible and reversible lifestyles that would increase the overall adaptive potential of the species.

For illustration, reproductive success of an apparently doomed single self-incompatible diploid S_1S_2 (AB for simplicity) individual that is beyond the population range of its parent population can occur as follows. Unreduced pollen and egg from this individual can result in a self-fertile tetraploid capable of colonizing new areas. If individuals in this AABB (i.e $S_1S_1S_2S_2$) duplex population mate with CCDD tetraploids during the new population range expansion or as a result of gene flow, the resulting hybrid progeny (ABCD) can undergo dihaploidization to give as many as seven diploid types, of which four (AC, AD, BC, and BD) can mate with any of the original diploid AB (S_1S_2) plants. Thus, in two generations, tetraploidization followed by di-haploidization can increase genetic diversity and restore outcrossing in the diploid population (Figure 4 and Figure 5).

In an extreme scenario, homozygous S-haplotype individuals can also arise as follows. A single AB diploid plant can form three types of unreduced eggs - AB, AA, and BB. If unreduced pollen has the potential to induce haploids, an unreduced AB pollen grain could pollinate its parent plant and induce dihaploids on unreduced ovules leading to AA, BB or AB seed. Two dihaploid individuals from this plant, an AA and a BB individual, each with hundreds of flowers, would be cross fertile and could generate thousands of progeny. While this scenario appears improbable, considering the high reproductive potential of potato (each potato plant makes hundreds of flowers, each flower has ~600 ovules of which about 200-300 develop into seeds in a regular diploid x diploid cross) makes the scenario more plausible. Additionally, haploid induction and unreduced gametes are relatively common in wild diploid potato

populations. They are formed in some potato dihaploid at very high frequencies of upto 45% [219,245,246]. As a result, the HI system combined with unreduced gamete production could allow frequent rescue of homozygous S-haplotype diploid individuals that would otherwise be at a reproductive dead end.

Interestingly, the above system could work in the absence of polyploids. Low-rates of unreduced gametes and spontaneous haploid induction are known to exist in many species. The above scenario may have occurred when Abdalla & Hermsen [172] pollinated a diploid (2n=2x=24) *S. verrucosum* plant heterozygous for *Phytophthora infestans* resistance with *S. etuberosum* (2n=2x=24, a wild species different from *S. tuberosum*). One of 5 seeds germinated to form a diploid (2n=2x=24) plant displaying inbreeding depression. This plant was phenotypically like *S. verrucosum* and upon selfing formed depressed uniform plants that were 100% resistant to *Phytophthora infestans*. Likely, this wide cross resulted in zygotic induction of an unreduced egg. Another possibility is that a true haploid was induced and that it spontaneously doubled during somatic growth. Similar results were obtained by crossing *S. etuberosum x* Phureja and *S. bulbocastanum x* Phureja [172]. Analogous results, albeit with the spontaneous doubled haploid coming from the male, were obtained in other crosses [172]. Spontaneous haploid doubling has been observed in Arabidopsis, potato, tomato, and maize [4,98–100,172].

Strategies to overcome self-incompatibility should be critical in harsh environments, such as extreme temperatures. These environments may foster speciation and may also facilitate tetraploidization by inducing unreduced gametes , and twin embryos (polyembryony), which are connected to haploid induction [247]. As stated by Webber [10]: “polyembryony is not abnormal. . . . in many genera polyembryony plays an important role both in the origin and the perpetuation of new forms”. Haploid-induction through wide-crosses may be part of this system. While the triploid block is ancient, this barrier could theoretically be strengthened by ploidy flux in mixed-ploidy populations. The presence of haploid inducing alleles would favor formation of euploid

seeds and decrease triploid seed in interploidy matings. Further, in mixed-ploidy populations without haploid-inducing alleles, the triploid block might be more relaxed to allow for the formation of triploids that are important gene-flow bridges between diploids and tetraploids [242,244]. Asexual reproduction might participate in this mechanism by increasing the probability of ploidy flux. Ploidy flux could be favored by clonal habits such as production of tubers or stolons. Tuberization, which is antagonistic to flowering, could ensure clonal survival in a population that cannot make seeds due to fixed self-incompatibility. The availability of genetic tools should make these (or my) assumptions testable and help address the role of the ploidy flux in mixed ploidy populations.

As a note on terminology, Mendiburu and Peloquin [221] have extensively discussed terminology and mechanisms applied to polyploidization and de-polyploidization. Briefly, potato polyploidization and de-polyploidization (haploidization) can occur unilaterally or bilaterally. Dihaploid production in potato is a form of unilateral sexual de-polyploidization. Unilateral sexual de-polyploidization is the production of progeny with reduced ploidy due to inheritance of less than the expected chromosome set from one of the parents such as during potato or maize haploid induction. The analogous term when both parents contribute fewer than expected chromosome sets is called bilateral sexual depolyploidization. Eg. Formation of a diploid individual from a triploid x triploid cross [221].

On the other hand, unilateral polyploidization occurs when, for example, a diploid plant's unreduced gamete fuses with a reduced one to produce progeny with higher ploidy than either parent; a triploid in this example. Bilateral polyploidization occurs, for example, if two unreduced gametes from a diploid were to form a tetraploid [114,221,243]. These terms are applied without regard to parent ploidy, or reproduction mode (selfing or outcrossing) [221]. Specific terms are applied to such polyploid progeny depending on the unreduced gamete parent, and the aforementioned's ploidy. For example, tetraploids from 4X x 2X potato crosses are diplandorous

tetraploids while those from a 2X x 2X cross are diplandrogynous tetraploids. See Mendiburu and Peloquin [221] for a fuller and extensive discussion on these mechanisms.

5.1. Implications for mixed-ploidy populations

Mixed ploidy populations are known to exist in many species. These are composed of different cytotypes (ploidies) which can include diploids, triploids, tetraploids, pentaploids, hexaploids or other polyploids. For unknown reasons, one cytotype usually dominates a population [243]. Dominant cytotypes were defined as "having an overall frequency exceeding 20% and sometimes observed in single-cytotype populations" [243]. The cytotypes in sexual species tend to have even chromosome set numbers while asexual species have additional cytotypes with odd-numbered chromosome sets [243]. The lack of cytotypes with odd-numbered chromosome sets (*'odd-ploids'*) in sexual species is most likely due to their being sexually selected against. In asexual species, these can be maintained through tubers, stolons, or other methods. The *'odd-ploids'* are, however, disadvantaged in the long-term survival due to sexual selection, but can transiently be powerful agents of polyploid evolution.

Kolar [243] asks "How stable are mixed-ploidy populations? Do they arise and disappear transiently or are they stable over long periods of time?" Ploidy flux can help answer questions, particularly in species with gametophytic SI mechanisms. I hypothesize that the dominant cytotype is determined by the amount of genetic variation present in the population. When genetic diversity is high, the diploid cytotype will dominate because outcrossing increases fitness. On the other hand, self-fertilization in tetraploids will lead to inbreeding depression, particularly because half of the tetraploid pollen could be rejected because of self-incompatible alleles. Meiosis in tetraploids, especially recent autotetraploids, is usually imperfect due to multivalent formation at prophase, which through missegregation produces defective gametes (and progeny). As a result, aneuploidy is frequent in tetraploids [248]. In potato, tetraploids

transmit extra or fewer chromosomes than the gametic number [49]. When genetic variation is low, the diploids will have fewer mates and risk extinction due to SI. In these populations, the tetraploids will dominate due to their sexual advantage through selfing, which, however, can eventually lead to inbreeding depression.

Ploidy flux, therefore, acts as a balance between selfing and outcrossing to promote sexual reproduction and to increase genetic diversity. Indeed, it has been observed that extreme temperatures, likely to be found outside a species range, lead to production of unreduced gametes and formation of tetraploids [222,230,249,250]. Polyploids are important for colonization of new, disturbed or marginal environments because of, among other things, self-fertility [251–253]. They have been shown to prosper in such new environments better than their diploid analogs [253]. Polyploids are also likely to be found in diploid population contact zones [230,253]. In addition, formation of new polyploids has been associated with (geological) mass extinctions events affecting vegetation, animals, and insects (including pollinators), such as in the Cretaceous–Paleogene (K-Pg) boundary [254]. Polyploidization might have enabled colonization, overcoming extinction of pollinators and climate perturbation due to the impact of the Chicxulub asteroid that wiped out dinosaurs [253,255,256]. Per Levin and Soltis (2018), “the demise of diploids afforded opportunities for polyploid establishment and expansion into novel habitats. Large populations of polyploids ostensibly were more likely achieved during the post K–Pg period than during most other times, because the extinction and depletion of a significant fraction of plant species liberated habitats and resources which could be exploited by the survivors. Under these circumstances, newly emergent species that were self-fertile would have a reproductive advantage” [253]. The role of polyploidization in plant survival is further suggested by analysis of other mass extinction events, such as the Cretaceous-Tertiary (K-T) mass extinction [256].

Mechanisms and processes of polyploidization and de-polyploidization, including through the triploid bridge, have been discussed extensively by Stebbins [257], De Wet [258], and

Husband *et al* [243,244,259]. Haploid induction is a more direct mechanism of de-polyploidization that has received little consideration, potentially because it is poorly understood . For example, HI genes and mechanisms have only been identified recently, compared to the long standing knowledge of unreduced gametes and interploid mating. Additionally, without a good understanding of the genomic structure, it is difficult to distinguish dihaploids (and trihaploids) from true diploids (and triploids) by cytology and flow cytometric analysis. The Ploidy Flux hypothesis predicts the existence of haploid inducing alleles in mixed-ploidy populations. The identification of such alleles could be co-opted for improving or developing haploid induction. The hypothesis also predicts the existence of polyploids in small or new populations, consistent with the literature [250].

In summary, ploidy flux describes a novel mechanism that could have a significant role during polyploidization and de-polyploidization mediating the interaction of populations of different ploidy and the evolution of, mixed-ploidy populations. Ploidy flux predicts that dominance of autopolyploids in mixed-ploidy populations might be a signal of reduced genetic diversity in that population. Consideration of ploidy flux may help advance our understanding of the ecology and evolution of plants.

6. Discussion and future directions

I have discussed the history of potato haploid induction, and highlighted some of the abnormalities that are observed in potato pollen and embryo development. HI in potato appears similar to maize haploid induction, which in turn is very similar to apogamy in species like Atamosco Lily. In this species, pollen delivers sperm cells to the ovule. One sperm fuses with the central cell to form the endosperm, while the other sperm seems to adhere to the egg but does not fuse with it. The egg develops to form an embryo that is haploid, as it may be the case in maize and in potato.

In the case of potato haploid induction, more work is required to further our understanding. For example, the process of double fertilization and embryo development following the HI cross have not been well characterized as in maize and Arabidopsis. Microscopic, proteomic, and biochemical analysis of the events leading up to and immediately following double fertilization in potato will be essential to test hypotheses on the mode of HI and to identify candidate factors controlling double fertilization and haploid induction.

The Ploidy Flux model provides a novel explanation for the existence of haploid induction in nature. In this model, haploid induction, unreduced gametes, and self-incompatibility play evolutionary roles by tailoring the mating strategy of a species to fit varying genetic and ecological scenarios. When there is a dearth of mates, these phenomena can lead to a switch from diploidy to tetraploidy. This allows for selfing in a single flower to produce future mates. However, when there is an abundance of mates, the ploidy can be reduced from tetraploid to diploid, promoting outcrossing and increasing genetic diversity. This model can be tested in other species that like potato, have ploidy series in sympatry and can interbreed such as *Solanum nigrum, Pilosella echioides, Senecio carniolicus, Solidago altissima,* and *Senecio carniolicus* [222,230,244,259–261].

For breeding applications of HI, new HI loci should be identified by studying wild hybridizing mixed-ploidy populations, especially ones with low-incidences of triploids. More loci that behave like the ones identified in potato and maize might be found and co-opted into breeding programs.

Chapter 3: Mapping Potato Haploid Induction gene(s)

1. Abstract

Production of potato haploids by use of genetic haploid inducers is an important method in potato breeding programs. The haploid inducers are diploid lines that, when crossed to commercial tetraploid lines, generate dihaploids (diploid haploids of the tetraploid variety). The dihaploids are important bridges for the introgression of various traits from diploid germplasm into tetraploids cultivars. However, neither the mechanism nor the genes involved in potato haploid induction (HI) have been elucidated. This chapter presents my effort to map the locus or loci associated with haploid induction in potato. I used two parents, a non-haploid inducer diploid Bolivian Phureja accession (GND; PI 258855), and the haploid inducer IvP 35, to generate a diploid mapping population. Haploid induction was tested by crossing F1 and F2 individuals from this population to a tetraploid accession (PI 310467) related to tetraploid cultivar Desiree. Amongst the F1 and F2 progeny of GND and IvP 35, I observed a wide range of haploid induction rates. Specifically, 41% of the progeny were non-haploid inducers and 9% were excellent haploid inducers. HI rate was positively correlated with seed abortion. I also identified a 6.3 Mb region on chromosome 3 that is highly associated with HI rate and potentially contains the major HI locus. A list of candidate HI genes are presented.

2. Introduction

The formation of haploid (monoploid, dihaploid or trihaploid) plants in potato crosses has been known for close to a century [126]. Haploid induction (HI) is usually observed when tetraploid or hexaploid plants are used as female and crossed to diploid plants, particularly Phureja strains [128–130,133,139,140,158,159]. These crosses result in variable percentage of dihaploid (a haploid of a tetraploid) or trihaploid (a haploid of hexaploid) progeny. This phenomenon has been co-opted by potato breeders for the production of dihaploid plants. Dihaploid plants are critical in potato breeding for two main reasons. First, they allow easy introgression of traits of interest from wild accessions and species into commercially cultivated varieties that are almost always tetraploid. Second, dihaploids allow breeding at the diploid level which involves disomic instead of tetrasomic inheritance [130,131].

Efficient dihaploid breeding in potato became possible with the identification and subsequent breeding of diploid potato lines that exhibited very high rates of haploid induction when crossed to tetraploid cultivars [132]. These HI strains showed much higher HI rates (HIR) than typical in the tetraploid x diploid inter-ploidy crosses [144,151]. Virtually all potato breeding programs employed haploid inducers after 1973 when the high HIR trait was combined with a haploid screening trait (by use of an embryo spot) in one genetic background by Hermsen and Verdenius [262]. These haploid inducers are called IvP lines and the most widely used lines are two full-sibling plants, IvP 35 and IvP 48, and a derivative of initial IvP lines called IvP 101 [144]. These lines are maintained asexually and can be obtained from potato germplasm banks including at the USDA (US Department of Agriculture) and in Peru at CIP (Centro Internacional de la Papa i.e The International Potato Center).

The universal adoption of IvP haploid inducers is due to their ease of use. A normal haploid induction cross is made using IvP as male. The resulting progeny will include hybrid and dihaploid seed whose germination does not require any special technique or treatment such as embryo rescue or hormone applications. IvP plants are homozygous dominant for the

production and deposition of anthocyanin between cotyledons. The resulting anthocyanin embryo spot is visible to the naked eye through the seed coat [262]. Hybrid seed displays the embryo spot while the dihaploid seed or accidental selfs do not. This provides a simple and convenient method for screening hybrids without the need of seed germination.

Despite the wide adoption of IvP haploid inducers in potato breeding programs, there are very few studies on the mechanism and genetics of potato haploid induction. Three studies from the 1960s concluded that haploid induction is a result of parthenogenesis [107–109]. These studies suggested that the haploid inducer pollen produces defective sperm that is unable to fertilize the egg. However, the defective sperm results in conversion of the unfertilized egg into a zygote. After the 1990s, a few studies reported detection of haploid inducer DNA in near-dihaploid plants using PCR, cytology, Southern blot analysis, isozyme expression, and whole genome sequencing [153–155,203–206,208]. These researchers suggested that haploid induction happens through genome elimination instead, although at least one other study reported that no haploid inducer DNA was detected in a substantial number of dihaploids [263].

Because the frequency of haploid inducer DNA addition reported by these studies varied greatly, our laboratory sequenced over 900 dihaploids and found that a haploid inducer chromosome or chromosome fragment was incorporated in less than 1% of flow cytometry-identified dihaploids. The remaining 99.3% of these were entirely free of haploid inducer chromosomes and therefore were true dihaploids [49]. This observation suggests that haploid induction takes place very early during embryo development, and that it is very efficient at eliminating inducer chromosomes. However, the mechanism remains to be elucidated. To address this lacuna, I undertook the genetic mapping of the HI trait.

Low rates of haploid induction are observed when many Phureja diploid accessions are crossed to tetraploids [128–130,133,139,140,158,159,262]. This suggests the existence of haploid induction allele(s) in various Phureja accessions. The commercial haploid inducers are derived from one of these accessions - PI 225682. In a 1963 study of this accession, three out seventy plants

studied exhibited unusually high rates of HI - up to 63-fold higher than the regular background rates [132]. Improvements in HI rate were attained by further breeding [144,151,262]. The observation of widespread, low-efficiency HI in many Phureja accessions, and the gains obtained by breeding into several backgrounds, suggest that HI in potato is a multi-genic trait. The inheritance pattern of HI also suggests that HI can be mapped. To produce a mapping population, a non-inducing plant (referred to as GND here) from a Bolivian accession (GND s.n; PI 258855) was used as the wild-type parent and IvP 35 (PI 584995) was used as the HI parent.

I hypothesized, based on the previous literature and my observations [132,144,151,262], that there are at least two HI genes that govern haploid induction in potato. The major HI gene (*H*, for haploid induction) is expected to be either semi-dominant or gametophytic. Data reported in various literature weakly suggests that the second gene (*e*, for enhancer) is recessive [132,136,146,151]. Van Breukelen [151] observed that HI rate was continuous and inferred that it could be broken down to seven HI rate levels, and that these levels could be explained by five to seven loci. Dominance and recessivity of the HI loci was ruled out, and it was assumed that all HI loci have “intermediate inheritance” (semi-dominant or gametophytic). However, this cannot explain observations where the progeny of two low HIR parents have very high HI [136]. To me, this suggests recessivity or epistasis. Their reasoning does not seem to take into account how the combination of various genotypes at a few loci can result in a wide range of phenotypic observations. Instead, I hypothesize that two genes proposed here are sufficient to explain seven levels of HIR based on allelic combination with the HHee having the highest HIR and *hhEE* being wild-type. Locus *H* has a larger effect than locus *E*. There are nine allelic combinations - ordered here based on expected HIR levels - *HHee*, *HHEe*, *HHEE*, *Hhee*, *HhEe*, *HhEE*, *hhee*, *hhEe*, *hhEE*, two or three of which might have HIRs (e.g *HhEE* = *hhee* or *Hhee* = *HHEE*) that are similar enough that it is impossible categorize them as distinct. However, it is possible that there are a few additional loci that might explain minor variations, as is the case in maize where HIR is dependent on the genetic background [68,71]. However, the effects of these

minor-effect loci are likely not to be detected given environmental variance and small population sizes in both this and other published studies.

3. Methods

3.1. Plant material

The haploid induction tester plant PI 310467, a close relative of Desiree (hereafter called Red Polenta), the haploid inducer mapping parent (IVP 35; PI 584995), and wild-type mapping parent (GND s.n., PI 258855) were obtained from the USDA Plant Germplasm center as *in vitro* plantlets for the first two, and as seed for GND. Seeds were germinated and propagated aseptically in half-strength MS media. *In vitro* plantlets were transplanted in the greenhouse and GND s.n plant # 4, with purple stems and underleafs, was chosen to be used as a wild-type parent in the mapping population. GND s.n plant # 4 (GND SN4 a.k.a GND 4) will be referred to as GND from here onwards. It will be deposited in the USDA Plant Germplasm repository.
Red Polenta has green stems and leaves. IvP 35 is homozygous for the embryo spot marker, a trait well suited for screening hybrids and resulting from anthocyanin deposition between the cotyledons that is visible through the seed coat [262]. Even though IvP 101's HI rate is three-fold that of IvP 35, it was not used as the HI parent because it has some Tuberosum (non-Phureja) genetic background [144]. IvP 35 and GND have red stems and leaves. F1 and F2 seeds were obtained as described below.

To obtain the mapping population, GND was pollinated with IvP 35 . Fifty-two of the resulting F1 plants (all diploid) were grown in the greenhouse and intercrossed to make an F2 population. Pollen pooled from these 52 F1 plants was used to pollinate emasculated flowers of the same 52 plants. This was done to reduce chances of biased marker inheritance caused by the presence of less than four self-incompatibility (SI) alleles. At least one berry from the intercrossed F1 flowers was collected from each of the 52 plants and its seeds extracted to obtain the F2 lines.

Additional haploid inducer lines were obtained from the USDA Germplasm collection. These include Phu 1.22, IvP 35, IvP 48, IvP 101, PL4, and Herm 18. All these lines share ancestry (Figure 6). IvP 35 and IvP 48 are siblings. Phu 1.22 is the grandparent to IvP 35 and IvP 48. IvP 101 and PL4 are derived from IvP 35, IvP 48, or their siblings that were produced by Hermsen & Verdinius [144,262,264]. Herm 18 ancestry is not very clear but was deposited at the USDA potato germplasm bank by Shally Jansky. Jansky inherited it from Robert Hanneman Jr, who likely inherited it from his co-researcher, John Hermsen, who developed the IvP populations [262,265].

3.2. Seed germination and potato line propagation

It is recommended that seeds are obtained from ripe berries and are aged six months or more after seed extraction to break dormancy [266–269] . However, I was able to germinate seeds that were at least three months old without any problem.

To increase germination rate, seeds were soaked in 1500ppm GA3 (gibberellic acid) for 24-48 hours at room temperature to break any residual dormancy [270,271]. The seeds were then rinsed in water, and washed in soapy bleach solution (50% bleach with either 0.5% Tween-20 or Triton X-100) for up to 10 minutes. The bleach was washed off and the seeds plated on half-strength MS media (0.5x MS with vitamins, 0.5% sucrose & 0.7% phytoagar) in magenta boxes. Germination was done under cool and dark conditions where the highest germination rates are obtained [270,271]. F1, F2, IVP 35, GND and Red Polenta seedlings or *in vitro* plantlets were maintained *in vitro* by propagating the material using cuttings.

3.3. Crosses

Pollen was extracted by harvesting anthers just before or right after light browning at the tip. Anthers were placed on a filter paper and left to dry for 24-48 hrs. Pollen was then extracted using a vibrating rod (VegieBee™) on the folded filter paper. The pollen was then collected in 1.5 ml tubes and stored at 4° C or used immediately. Only pollen less than two weeks old was used for pollination.

For pollination, female flowers were emasculated by removing the anthers at the start of yellowing. Pollen was then placed on the stigmas of these flowers either immediately or up to 3 days after emasculation, depending on flower maturity. Fruits were then harvested at around 30 days, and further ripened on the lab bench for up to 30 more days. Seeds were extracted from ripe fruits, rinsed with tap water, washed in 25% bleach for 10 minutes, rinsed, and placed on a filter paper to dry. The bleach wash bleached the seed coat to increase visibility of the seed embryo spot. Seeds were stored for at least three months on the lab bench before sowing to partially break dormancy.

3.4. Phenotyping

13 F1 seeds and 100 F2 seeds were germinated and the plants phenotyped for HI ability. Red Polenta was used as the tester plant. It has completely green stems and leaves, and exhibits no anthocyanin production on the body. The Phureja clones, on the other hand, have red stems, leaves and an embryo spot. Phenotyping was done by pollinating emasculated Red Polenta (female) flowers using pollen from a HI mapping (HIM) plant, GND or IVP35. This was called a Haploid Induction test cross. HI test-cross seeds were harvested in batches of 5 fruits or less.

Only unspotted seeds from each test cross were germinated. Purple seedlings and spotted seeds were assumed hybrids and therefore were not evaluated further. The ploidy of green seedlings was determined using a flow cytometer. An F1 or F2 individual was classified as a haploid inducer if any of the seedlings produced from the test cross was dihaploid. HIR (Haploid induction rate) was calculated as dihaploids per berry (h/b) in line with all prior literature. I aimed to have at least 5 berries per plant analyzed in line with literature [151]. However, a few of the plants fell below that threshold, but were still informative if they induced haploids. A plant that produces at least one haploid from any one berry is categorized as a haploid inducer (observation of haploids is mutually exclusive with being wild-type). There were no replicates and hence standard errors could not be calculated, and environmental effects

could not be determined. However, it is known that haploid induction is robust across environments [132,144,151]. Three of the 23 wild-type plants for which three or fewer berries were analyzed could be problematic, but were sequenced nonetheless and used.

Ploidy was determined using a modified protocol based on the method of Rayburn *et al* [272] and Bino *et al* [273]. Young leaves of *in vitro* plants were chopped in chopping buffer to release nuclei and the tissue debris filtered into a test tube by passing the chopped leaf suspension liquid through the Cell Strainer (35µm mesh size) Snap Cap of a Falcon™ Round-Bottom Polystyrene Test Tubes (catalog number 352235). A BD Biosciences *FACScan* flow cytometer or Beckman Coulter *Cytoflex* flow cytometer were used for ploidy analysis. Red Polenta was used as the 4X control while GND and IvP 35 were used as the 2X control. At least 5000-10000 events were used to call the ploidy.

3.5. Sequence analysis and QTL mapping

GND and IVP 35 were previously sequenced to around 30X coverage [49]. The genomic DNA extraction was done as previously described by Ghislain *et al* [274] and library prep was done using 300bp sheared input DNA with KAPA Hyper Prep kit (cat. No KK8504). Sequencing was performed at the University of California, Davis DNA Technologies Core on Illumina Novaseq™ 6000 lane using paired-end dual index 150bp reads. A full description of the DNA extraction, library prep, and sequencing, and demultiplexing is described in Amundson *et al* [49].

Leaf tissue from eight F1 lines, seventy-three F2 lines, Phu 1.22, and HERM 18 were shipped in dry ice to Novogene Corporation in Sacramento, CA for DNA extraction, library prep, and sequencing using Novogene Corporation's internal protocols. Paired-ended 150 bp reads were obtained on Illumina NovaSeq 6000 platform for an average coverage of 10X read coverage per sample. Custom Linux scripts were used to demultiplex the reads and remove the adapter sequences. These scripts can be downloaded at: http://comailab.genomecenter.ucdavis.edu/index.php/Barcoded_data_preparation_tools.

Custom Linux and R scripts were used for downstream analyses. The reads were mapped to the potato reference genome assembly version DM v6.1 [275] using the BWA software [276](BWA mem algorithm) and variants called using BCFtools version 1.10.2 [277]. Only reads with mapping quality greater than 30 were used and all duplicate variants removed (bcftools norm -d all). GND and IvP 35 aligned reads were compared to find genomic variants between the two parents. Sites with genotype inference score qualities less than 100, or read depth higher than 70 (2.3x genomic average), or less than 25 (0.83x genomic average) were excluded. Three sets of variants were identified: SNPs that are homozygous in GND but heterozygous in IvP 35 (900,866 positions), SNPS that were heterozygous in GND but homozygous in IvP 35 (1,935,345), and finally positions that were homozygous in both parents but different between GND and IVP 35 (618,599 positions). Mapping analysis was attempted using these three SNP-types. The genotypes of F1 and F2 samples at the chosen sites were then determined as below. To determine the marker genotype in F1 and F2 lines, the genome was first partitioned into 100kb non-overlapping bins.

For each individual, genotype was based on information from all SNPs located with each bin. The minimum total number of informative reads covering SNPs in a bin was set at 50. For each individual, the genotype of bins for which fewer than 50 informative reads were available were left unassigned (NA). For bins with sufficient informative reads, a locus was called homozygous when more than 95% of reads covered a single allele and heterozygous when more than 5% but less than 95% covered both alleles. These bins were used as markers for the bulked segregant analysis (BSA) or *r/qtl* analysis [278–283].

For the BSA, 60 F2 plants were sequenced and divided into three bulks (sub-populations) based on HIR for preliminary analysis. These were divided into the wild-type bulk (HIR = 0), the intermediate bulk (HIR ranging 0.2 and 0.4 h/b) and high HIR bulk (HIR => 0.6). Plants with intermediate HIR were removed from this analysis. Based on encouraging results, I

sequenced an additional twenty-one F1 and F2 plants, and performed both BSA analysis and QTL mapping on the full population.

3.6. Phasing IvP35 chromosomes

To conduct the r/qtl analysis, I phased the IvP35 haplotypes using heterozygous SNPs, using both the F1 and F2 individuals. Sites where 3 or more of the 81 individuals had a high-quality read coverage of less than seven (average read depth coverage across the whole genome was ~9) were not used for phasing.

All F1s carry one IvP 35 haplotype while the F2s could have zero, one, or two IvP35 haplotypes. Phasing a particular genomic region was done as explained here and in Figure 7. Only F2 individuals that carried a single IvP haplotype in that region either as heterozygotes (one GND chromosome and one IvP 35 allele), or as a homozygote (two copies of the same IvP35 allele) were used. Individuals with both IvP35 haplotypes in that region, i.e heterozygous (haplotype1 / haplotype2), or only GND haplotype, i.e homozygous for GND chromosome, could not be used for phasing. The genotype information obtained as described as in Figure 7 was used to determine which regions of which F2 individuals to retain.

To determine if two markers were in-phase, I counted the number of times two loci were in phase or in repulsion i.e for adjacent loci that are AB/AB, the number of BA/BA, AB/AB, or BB/BB individuals in F1 and F2 determine the phase. At a particular locus, an individual could have GND chromosomes only and therefore could not be used for haplotyping. To ensure proper haplotyping, a minimum of 30 individuals in each locus were used for haplotyping. The markers are in phase if more than 85% of individuals are BB, and in-repulsion if more than 85% of individuals are AB or BA. The haplotype is built progressively by doing the same for the 2nd loci and the third loci, then the 3rd loci and the next one, and so on. There should be no recombination between any two adjacent markers for the overwhelming majority of markers and individuals, and only a maximum of one is expected if it occurs.

464,574 (51.6%) of the 900,866 IvP35 heterozygous positions were phased into haplotypes. F1 and F2 genotypes were assigned for each of these positions by comparison to haplotype. For regions homozygous for one IvP35 haplotype, these sites should always be homozygous. The same is true for regions heterozygous for IvP35 and GND chromosomes. Regions homozygous for IvP, but with both IvP35 chromosomes will be heterozygous for all sites. Thus we can assign genotype based on which haplotype a region has. Those homozygous for GND chromosomes (i.e do not have either IvP haplotype) were made equivalent to either of the IvP35 haplotype and analysis done both ways i.e both haplotype 1 and GND chromosomes are assigned genotype A for mapping purposes while haplotype 2 is assigned genotype B (data set 1). The haplotype assignment can then be switched so that haplotype 2 = GND = A while haplotype 1 = B (data set 2).

These SNPs were put in 3276 100kb bins across the whole genome. A bin was used as a marker if it had at least five of these SNP positions. If 80% (⅘) or more of the SNPs corresponded to the same genotype, that genotype was assigned to the bin (marker). If not, the bin was assigned a missing genotype. These data were used for rQTL analysis.

The GND heterozygous SNPs (GND is AB and IvP35 is AA) were also phased analogously. 1,196,238 (61%) of 1,935,345 GND heterozygous SNPs were phased and assigned to 6511 bins (markers).

3.7. QTL mapping

QTL mapping was performed in R statistical software version 3.2.3 [284] using the *qtl* (*r/qtl*) package version 1.41-6 [280]. Standard data cleanup outlined in the *r/qtl* manual [278]. Removal of duplicated markers and markers genotyped in less than 60 individuals reduced my marker number from 3276 to 781.

r/qtl uses interval mapping to identify QTLs. Three mapping algorithms implemented in *r/qtl* were used: the EM algorithm, Haley-Knott regression, and multiple imputation [278,285–288].

SnpEff version 4.3 [289] was used to predict the effects of genic SNPs on protein function. This analysis using all IvP35 heterozygous SNPS within the QTL interval.

4. Results

4.1. A general description of the mapping population

The F1 and F2 plants were segregating for many traits as would be expected of progeny from highly heterozygous outcrossing populations. These traits include pollen yield, anther shape, pistil size, style length, organ colors, growth habit, stolon formation, tuber yield, tuber morphology (color, shape, and size), stem shape and thickness, flowering ability, self incompatibility, unreduced pollen (as assessed by seed yield in 4X x 2X crosses), berry size, seed yield, and of course, haploid induction.

IvP35 has few, relatively thicker stems that senesce at maturity. This necessitates continual planting of IvP35 (and other IvP or IvP-like plants, Herm 18, and Phu 1.22). GND on the other hand, has a bushy growth habit resulting from many underground runners that produce many relatively thinner stems. It grows *indefinitely* and flowers very profusely. I have been able to maintain a GND plant (and some of the F1 and F2 plants) in a single pot for more than three years. GND produces very small tubers with white skin and flesh. It also appears very resistant to spider-mites, a green-house potato pest that decimates the IvP plants.

The GND x IVP 35 F1 progeny had unexpected phenotypes. First, not every F1 plant was spotted, despite the expectation given that IvP35 is homozygous for the seed spot [146,151]. Additionally, the seeds were a binary mixture of very small seeds and very large seeds with the large seeds having more than twice the volume of the small ones (Figure 8). These unusual traits prompted me to test seedlings from these seeds for haploidy (monoploidy). All were diploid.

The seed size and embryo spot phenotypes segregated in the F2. Overall, the plants from small non-spotted (SNS) seeds were morphologically similar to GND whereas those from

small spotted (SSP) seed and large non-spotted (LNS) seed were intermediate. Plants from large spotted (LSP) seed resembled IVP 35.

F2 plants with high haploid induction rates (HIR equal to or better than IvP 101 HIR, the best IvP HI plant) that inherited the GND morphology could be useful haploid inducers because, compared to the IvP, they require lower-maintenance during greenhouse growth, might be resistant to common pests, and flower abundantly and persistently.

4.2. The Haploid Induction Phenotype

All haploid induction phenotyping was done by crossing the test subjects as male to Red Polenta. The Red Polenta tetraploid tester-line is a relative of tetraploid Desiree cultivar. Lines that yield at least one dihaploid following the crossing to the tetraploid tester line are classified as haploid inducers. Haploid induction rate (HIR) is measured as haploids per berry (h/b). Wild-type lines do not induce any haploids (i.e are non-inducers)

The HIR of various IvP inducers was measured. The HIRs of these haploid inducers (IvPs) were consistent with published literature [132,136,146,151]. IvP35 and IvP48 siblings induced haploids at a HIR of 0.32 h/b and 0.27 h/b, respectively. HERM 18 (it is unclear if this is IvP18, a sibling to IvP35 and IvP48) induced haploids at 0.17 h/b. IvP 101, a far superior haploid inducer derived from IvP 48 and its siblings, induced haploids at 1.0 h/b. Phu 1.22 (GS 29) induced haploids at a rate of 0.5 h/b. Phu 1.22 is the ancestor to all modern potato haploid inducers. It is one of three *'superior'* (i.e individuals with unusually high HIR) haploid inducers identified by Gabert [132] by phenotyping for HIR in seventy plants grown from accession PI 225682 seed [132,144,151,262]. The HIR of GND, the wild-type parent used in creating our mapping population is 0.0 h/b.

The haploid induction rate of eight F1 plants varied. Five of these plants were haploid inducers and three were wild-type (i.e non-inducers) (Table 2). One of the F1 haploid inducers was more than three-fold more efficient (1.0 h/b) than IvP 35 (0.3 h/b), its inducer parent. The

observation that an F1's HIR is three-fold higher that of its HI parent seems surprising at first. However, this is not unusual. Similar results were observed in other wild-type x haploid inducer F1 populations [136,151]. I speculate that this is linked to the presence of more than one HI gene in Phureja accessions in general, some of which might be epistatic or synergistic. As a result, a haploid inducer's genetic background plays a big role in determining HIR in potato, just as in maize.

The HI trait segregated in the phenotyped F2 plants (n=75) with both wild-type and haploid-inducing plants present (Table 3 and Figure 9). 41% of the tested F2 plants were wild-type, and 9% were excellent haploid inducers with HIRs comparable to those of best haploid inducers (IvP101 and PL4) used in breeding programs. The rest were intermediate between the two extreme HI phenotypes.

4.3. HI rate is correlated with seed death

The tetraploid x haploid inducer crosses in potato result in a wide variety of offspring: tetraploid, triploid, dihaploid, aneu-dihaploid, aneuploid, and aborted or dead seed [49]. The incidence rate of these progeny types vary depending on the female line and the haploid inducer [144]. My study material offered a chance to study correlations between any of these traits. Among these is the correlation between HI rate and seed abortion, and between HI rate and seed set. Seed set in 4X x 2X crosses in potato can be a proxy for formation of 2n pollen [290]. No correlation between seed set and HI rate was observed in my material. This is in line with previous reports [144].

I then looked at the seed death. HI rate is known to have a strongly positive-correlation with seed abortion in *CenH3*-based, *DMP*-based, and *Phospholipase A1* (*Not-like-Dad* or *Matrilineal*)-based haploid induction, as previously observed in Arabidopsis and maize [28,30,57,69–71,291]. Therefore, I tested whether the same is true for potato HI.

There was a strong positive Pearson's correlation (0.61; p-value = 6.6e-10) between HI rate and seed abortion (Table 4 and Figure 10) in 86 diploid lines (both haploid inducers and

wild-type) that include F1 and F2 individuals; IvPs 18 (HERM 18), 35, 48 and 101; PL4 and GND. The correlation is 0.62 (p-value = 1.527e-10) when HI rate is normalized by taking its square root. When the correlation test is performed for the haploid inducers only (n=54), the correlation is 0.47 (p-value = 3.04e-04) or 0.48 (p-value = 2.67e-04) when HI rate is normalized (Figure 11). Potato haploid induction, therefore, is similar to both maize HI and CenH3 with respect to seed death [30,291,292].

4.4. Sequencing & mapping

Single nucleotide polymorphism (SNPs) between the mapping population parental lines were identified. Two sets of SNPs were used for mapping. SNPs in the first set were homozygous, that is, fixed (true breeding) in both GND and IvP35 but different between the two parents. For simplicity, I represent the SNP information with letters. I used CC and DD as the genotype notation for one locus with homozygous SNP. The second set consists of heterozygous SNPs - SNPs that are fixed in GND, but heterozygous in IvP35, with one of the two IvP35 alleles being identical to the GND allele. This is represented below as AA and AB, respectively. Both heterozygous and homozygous SNPs sets were used for mapping but independent of each other. The unphased heterozygous IvP35 genotypes will be represented as (AB) to denote the two alleles at a single locus. Phased SNPs will use the conventional representation A/B or $\frac{A}{B}$.

IvP35 had 900,866 IvP35 heterozygous SNPs (i.e GND is AA and IvP35 is AB) of which 464,574 (51.6%) were phased. GND on the other hand had 1,935,345 GND heterozygous SNPs. This is twice as many as IvP35 heterozygous SNPs. This is not surprising given that the IvP lines were obtained by several rounds of selfing during the breeding of the HI lines. GND, on the other hand, was obtained from a seed collected from a wild outcrossing population. 1,196,238 (61%) of 1,935,345 1,194,291 GND heterozygous SNPs were phased.

I first attempted mapping the HI QTL by using the homozygous SNPs and the 73 F2 plants, which were considered a true F2 generation. However, I did not find any significant QTL

loci associated with HIR (data not shown). This was not surprising given the small population size and my hypothesis that IvP35 is heterozygous for the HI gene.

With the hypothesis that the haploid induction locus H is heterozygous in IvP35, I expected the mapping experiment to have more power when performed under the assumption that the mapping population resulted from intercrossing BC1 (backcross 1) individuals. In this instance, I used heterozygous SNPs for mapping. GND, considered as the "*recurrent parent*", is homozygous (AA)(AA)(AA)(AA) or $\frac{AAAA}{AAAA}$at four linked loci while IvP 35, considered as the *"F1 parent*", is unphased and heterozygous (AB)(AB)(AB)(AB) (Figure 7). Therefore, the initial GND x IvP35 cross is considered the "*BC1* cross". The inter-crossing of fifty-two GND x IvP 35 F1 individuals produced F2 that are equivalent to a BC1-intercross 1 (BC1-I1). Unlike a true BC1 or BC1-I1 population, three haplotypes (instead of two) are inherited in my mapping population: one from GND (AAAA/AAAA), and two from IvP 35 (e.g BAAB/ABBA; or any other such combination). Summarizing, my mapping population is considered an inter-cross for the purpose of mapping using the *r/qtl* package in the R Statistical Software. Even though IvP35 is the "*F1 parent*", it will not be referred to as such. I retain the terms F1 and F2 to refer to material derived from the GND x IvP 35 cross in the first and second generation, respectively.

My hypothesis is that the H gene is heterozygous in IvP35. Only one of the two IvP35 haplotypes carries the haploid inducing allele if my assumption that the HI locus (H) is heterozygous in IvP 35 is correct. This can be represented as BAABDDDD**H** and ABBADDDD**h,** with H being the dominant (or gametophytic) haploid inducing allele. The GND haplotype is AAAACCCC**h.** All the chromosomes were haplotyped and the haplotyped markers used for mapping. Below is the *r/qtl*'s summarization of the data:

```
F2 intercross
No. individuals:     81
No. phenotypes:      1
Percent phenotyped:  100
No. chromosomes:     12
Autosomes:           1 2 3 4 5 6 7 8 9 10 11 12
```

Total markers:	781
No. markers:	39 83 53 38 37 65 54 40 178 42 26 126
Percent genotyped:	89.9
Genotypes (%):	AA:64.0 AB:30.9 BB:5.1 not BB:0.0 not AA:0.0

I was able to detect a QTL after running the *r/qtl* analysis. The QTL could still be detected, albeit with reduced power, if the GND homozygous regions were equated to missing genotype. However, no QTL was detected if the GND haplotype (AAAACCCC**h)** was equated to haploid inducing haplotype H1 (BAABDDDD**H)**. I also attempted to identify any secondary or interacting QTL by running the *scantwo* function in *r/qtl,* but none was detected, possibly because our population is too small to identify any interacting loci (results not presented).

QTL mapping was conducted using three of the interval mapping algorithms (EM algorithm, Haley-Knott regression, and multiple imputation) implemented in the *r/qtl* package [278,285–288]. All three algorithms detected a QTL with the peak on chr 03 at 37.9 Mb with LOD values above 8 (Figure 12), and an LOD threshold around 6. The LOD threshold was set by running permutation tests as recommended in the *r/qtl* manual. The predicted LOD interval is 6.3 Mb between 37.3 Mb to 43.6 Mb. The presence of a QTL on chromosome 3 is in agreement with my preliminary bulked segregant analysis of sixty F2 individuals using homozygous parental SNPs (Figure 13 and Figure 14) that showed that HI QTL likely resided on the second half chromosome 03. The recombination breaks for each of the F2 is shown in Figure 15 and Figure 16. A boxplot of a representative genotype at the QTL versus HI rate is shown in Figure 17 and Figure 18. As shown in Figure 19, the F1 haploid inducers have a different haplotype compared to the F1 wild-type (non-inducers) across the QTL interval.

To check whether GND contributed a HI QTL to the population, mapping was performed as before except for the use of GND heterozygous SNPs (GND is AB and IvP35 is AA). Mapping for HI using these markers did not yield any QTL (data not shown). There were 3656 markers after *r/qtl* data cleanup down from the initial 6511. Below is the data summary.

No. individuals: 81
No. phenotypes: 1
Percent phenotyped: 100
No. chromosomes: 12
Autosomes: 1 2 3 4 5 6 7 8 9 10 11 12
Total markers: 3656
No. markers: 332 245 350 342 302 284 335 336 313 340 236 241
Percent genotyped: 100
Genotypes (%): AA:0.1 AB:77.0 BB:22.8 not BB:0.0 not AA:0.0

4.5. Narrowing down candidates

There are 476 genes within the LOD interval. Between chr03:37.2 - 43.7 Mb, there are 7319 SNPs which are homozygous in GND and heterozygous in IvP 35. Of these, I was interested in the SNPs that were within genes and therefore more likely to cause a change in protein function. This, of course, risks missing the candidate gene if the causative variant is within a regulatory region such as the promoter.

I narrowed down the list using the following assumptions: First, the HI gene encodes a conserved factor present at a minimum within Solanum species, and present in all potato species. Second, the *H* allele acts gametophytically (and hence seems dominant), most likely because of a loss of function mutation. This gametophytic expression appears additive. Third, the gene is actively transcribed in mature pollen, perhaps during or right after male gametogenesis. Fourth, most potato accessions have the functionally wild-type allele *h*. Fifth, most inducers including the IvP and Phu1.22 inducers should have the H allele. The genotype of H allele should predict the HIR. For example, IvP101, an excellent inducer, should be HH, IvP48 and HERM 18 (IvP 18) that are intermediate inducers should be Hh, and Phureja 1.22 - the ancestor of all the inducers - should be either Hh or HH. The reality, however, might be more complex than this given that there seems to be additional HI loci.

To apply these criteria as filters, I used a reference set of potato variants in 71 diploid or tetraploid potato accessions, varieties, and wild species (Panel-K71) Table 5. None of the individuals in this set has been reported to be a haploid inducer. In addition, the set features

only two Phureja accessions (of unknown HI potential). I compared all genic SNPs (regardless of genotype in GND and IvP 35) within the LOD interval using the SnpEff program to predict the effect of these SNPs on protein function. Any gene displaying a severe mutation (truncation, stop or start codon loss, splice site acceptor or donor mutation, or frameshift) in more than 10 potato accessions in the variation panel was considered a non-essential gene and disqualified. This narrowed the candidate list from 476 genes to 153 genes that I loosely called "*conserved*" genes.

4.6. SNP effect analysis

I then carried out SNP effect (using the SnpEff program) analysis for my mapping population using all the SNPs (genic or not) that are homozygous in GND but heterozygous in IvP35 that were in the QTL interval. There were 7319 positions that had such variants. The majority (5472) of all mutations are not predicted to affect the gene body. 1847 were predicted to have effect on protein function, coding sequence, expression or non-coding sequence (intron and UTRs). 310 mutations were predicted to affect CDS and alter protein structure or function with 292 having moderate effect (missense) and 18 having severe effect (stop gain, stop loss, etc) (Table 6). These 1847 SNPs fall in 97 genes of which 29 are in the conserved list as identified from the K-71 panel, i.e these genes are knocked out (not necessarily by related mutations) in fewer than ten lines in the 71-accession panel. These are my candidate genes (Table 7). One of these genes, Soltu.DM.03G016230, has five unique missense mutations that are not found in any other potato line in the K71 panel, and are present in the original Phu1.22 haploid inducer and its descendant extant-IvP haploid inducers.

5. Other result from a preliminary analysis

I had also considered one other gene (Soltu.DM.03G015780) that did not make the high priority list, but had been the top candidate before all the F2 lines had been sequenced. Soltu.DM.03G015780, a DUFF1442 gene that our preliminary studies had identified as a possible candidate, sits at the QTL peak and has a protein truncating mutation in thirteen of the

seventy-one accessions in the K71-panel. It has five close homologs in Arabidopsis, two of which (At5g62280 and At3g60780) had interesting properties. At3g60780 is a sperm-specific gene that is highly enriched in sperm-cells. At5g62280 expression changes 12-fold when Arabidopsis siblings are treated with MMS (Methyl Methanesulfonate), a mutagen. Arabidopsis mutants of these genes had, therefore, been selected to test their ability to induce haploids when selfed. Some of the seedlings, particularly for At5g62280, showed altered phenotypes similar to those observed in chromosome missegregation mutants *sds* (*solo dancers*), such as having three-cotyledons (instead of two), serrated leaves, stunted-growth, among others [293] . While some of these phenotypes are associated with chromosome missegregation, further analysis on these two genes were not done due to lack of haploids in the 300 plants analyzed for haploidy by search of haploid phenotypes (small infertile plants with small organs).

5.1. Genotypes of related haploid inducers

The pedigree of some of the extant haploid inducers is shown on Figure 6 [132,144,262,265]. Phu 1.22 is the original 'superior' haploid inducer identified by Gabert [132]. It has a moderate HIR compared to the two modern haploid inducers (IvP101 and PL4) with the highest HIR [144,264,294]. It is the grandparent to IvP 35 and IvP 48 - two widely used inducers that have moderate HIR. IvP101 and PL4 HIR is about two to three-fold that of IvP35, IvP48, and Phu 1.22 [262]. All these haploid inducers would be expected to share alleles with IvP35 based on their phenotype and genetic history. At the minimum, heterozygosity is expected. Homozygosity of the IvP35 allele (as opposed to the GND allele) would be expected for the high HIR lines given my hypothesis that HI is caused by a dominant or gametophytic factor. Consistent with the hypothesis, all these inducers have the IvP 35-specific Soltu.DM.03G016230 alleles. IvP 35, IvP 48, and Phu 1.22 are heterozygous. IvP 101 and PL4 are homozygous. However, Herm 18, a haploid inducer whose origin is not clear but is likely a sibling to IvP 35, has moderate rates of HI but homozygous for IvP SNP [265]. This plant makes a lot of 2n pollen evidenced by observation of numerous seeds per berry produced in the HI cross.

6. Conclusion

This is the first attempt (to my knowledge) to map the haploid induction genes in potato. The fact that I was able to identify a QTL for the haploid induction trait will add impetus to the study of haploid induction in potato and other crops. As a first attempt, this study was limited but has provided resources to develop a larger and more comprehensive study of potato haploid induction genetics. Particularly, it will not only allow us to identify the major haploid induction gene but potentially the basal haploid induction gene by fixing the major haploid induction locus in subsequent mapping material.

My studies identified the major QTL controlling haploid induction. Even though I attempted to reduce my candidate HI gene set using various criteria, I realize that I could have missed the HI gene because of the assumptions used. For example, it could be that HI is caused by changes in the promoter sequence of the gene instead of a direct change in protein structure. Such changes could not be determined based on the assumptions that I used for my analysis and thus my method could have missed the most important clues. However, I believe this to be a reasonable start, particularly given the large QTL interval with hundreds of genes.

The HI phenotype corresponded with the genotype in other individuals such as IvP haploid inducers and Phu 1.22 that are not in my mapping population family. The only exception was Herm 18 whose HIR is low despite being homozygous for the HI markers. I cannot speculate whether this line's production of numerous 2n pollen or interaction with other loci accounts for the reduction in HI rate of this inducer despite the HH genotype.

The study of haploid induction will not only help create strategies for inducing haploid induction in other crops but will also advance our understanding of chromosome inheritance, double fertilization, and zygotic and early embryonic development. For example, how does an embryo develop without the required (assumed) paternal-imprinted genes? If an embryo can develop without imprinted paternal genes, does that suggest that the imprinted genes are

actually transcribed in the sperm such that a fusion with the egg of a sperm shell (no chromosomes are deposited in the egg) is sufficient to trigger development as these transcripts get translated? What are the various communication signals between the sperm and egg that mark egg-sperm adhesion, plasmogamy, complete karyogamy, and that trigger embryonic fate? Understanding how the haploid embryos are formed in potato might help solve or provide tools for studying these and other biological processes.

Chapter 4 - Seed Traits

1. Abstract

Seeds can either be planted, or consumed as human food or animal feed . For these reasons, seed traits are important and play a central role in agriculture and food production. Physical, physiological, and molecular traits such as color, nutritional value, dormancy, size, shape, and volume are selected by breeders for different purposes. For example, size and shape are selected to allow automation of farm or processing machinery. Color can be a phenotype relevant to consumers, or a genetic marker. Additionally, multiple seed traits are of interest to scientists who conduct basic scientific research. Seeds are central to the evolution and success of spermatophyta (seed plants) due to the fact they house and nurture the next generation of a species. There are many genes, expressed biparentally or biased, that control seed development. One interesting outcome of plant seed evolution is sibling rivalry, and maternal-paternal (parent-offspring) conflict during seed development. The study of these phenomena increases our understanding of the process of evolution at the molecular level.

In a population derived from two diploid potato accessions, I observed dramatic variation in seed size in both the F1 and F2 generations. This change in seed size is reminiscent of the outcome of parental conflict. Additionally, I observed suppression of dominant anthocyanin genes that are usually expressed in the embryo. This trait is used as a selection tool during potato breeding. Potato color is also an important consumer and agronomic trait. I attempted to map the genes responsible for these two traits due to their evolutionary and practical importance.

2. Introduction

This chapter is an extension of the analysis of the population described in the third chapter. This was a corollary objective following the unexpected observations in the F1

population developed for mapping haploid induction. As a result, the experimental setup is not optimal and thus any results obtained here will require additional work. However, we can still obtain some information that might inform future studies.

2.1. Color genes and the seed embryo-spot

An unexpected observation was the absence of the embryo spot in some of the diploid F1 seeds from a cross of GND x IvP 35 (male). IvP 35 is homozygous for embryo-spot genes [262]. The embryo spot results from deposition of anthocyanins between the cotyledons. These pigments can be blue, red or purple depending on the allelic combination of several loci [295]. Color in potatoes is not a simple trait. It is determined by several genes that produce various color phenotypes [295,296]. In addition to color, there are genes that determine the patterning of the color genes expression. For example, it can be restricted to tuber eyes, tuber flesh, tuber skin, petals, floral abscission zone, embryo SAM (shoot apical meristem) area (i.e the seed embryo spot), nodes, stems, and abaxial or adaxial leaf surfaces [295–297].

The embryo spot is determined by two dominant color genes, B and P or R. P and R result in purple and red colored anthocyanins, respectively, throughout the plant body [295,298]. P is epistatic to R. B on the other hand results in the deposition of anthocyanins in the floral abscission zone, embryo spot, tuber eye spots and stem nodes [295].

The color traits discussed above are important not only from a genetic or evolutionary point of view, but also for commercial and human-health purposes. Tuber flesh or skin color is one of the major determinants of whether various potato varieties are processed for the french-fry market, for chipping (potato chips/crisps), retail store sale (i.e fresh market), specialty markets, or for other uses [299–302]. Thus, the inheritance of color from wild potato accession or species - these tend to be heavily colored - can create a major bottleneck during breeding. Some of the important potato agronomic traits, such as the aforementioned, are introgressed into the cultivated varieties from the wild accessions, landrace varieties, or wild species. A potato plant that is resistant to all potato diseases, has ideal size tubers, and is resistant to

major pests and harsh climatic conditions, can be rejected by growers and processing companies if its tuber flesh is red or light purple instead of the preferred white color. This, of course, is a major challenge for plant breeders.

Additionally, potato color might be important for human health. Some studies have indicated that pigmented potato that have a high anthocyanin content improve health by increasing the amount of dietary antioxidants [300]. Therefore, understanding the inheritance of color in potato might have huge implications in the future. The presence of a gene that blocked expression of color in the F1 seeds of GND (female) x IvP35 provided an opportunity to map the color loci.

2.2. Seed size

I also observed strong variation in seed size in the F1 and F2 seed. This was a binary trait with the seeds being big or small as determined by visual examination. Seed size is another important agronomic trait. Seeds are the source of nutritional calories or industrial raw material for some seed crops. Additionally, farm and industrial equipment used for seed planting, harvesting, and processing are customized to defined seed sizes. Seeds smaller than the standard size might be overplanted, lost, or damage equipment during harvesting and processing.

Biologically, seeds were one of the most important evolutionary adaptations that led to the success of seed plants. The seed is a complex and unique adaptive organ made of tissues with one of three unique genetic makeup that include the seed coat (maternal genetic makeup), the embryo (1:1 maternal:paternal DNA) and the endosperm (2:1 maternal:paternal DNA). All these tissues are interdependent, and have complex communication and signaling among them, and with the endosperm providing nutrition to the embryo [303,304].

One of the most fascinating evolutionary outcomes of propagation of species through seeds is the well documented paternal-maternal conflict where the male and female parents "*attempt*" to control seed size to increase their competitive edge [167,305]. The paternal-maternal

conflict hypothesis postulates that there are male genes that have evolved the function of increasing the size of the endosperm to attract more nutrients for their offspring. This leads to large endosperms. However, this is detrimental to the success of the half-sibs from the female and different males as bigger seeds deplete resources that would otherwise be used to develop these half-sibs from a different male. Therefore, plant have evolved maternal parent genes to counteract the male genes that increase seed size, at least within a population [167,304–306]. This advantages the female by ensuring that she produces more offspring each with equal chance of success as all seeds she produces have her genetic makeup even though they might be half-sibs.

Differences in seed size tend to be observed when individuals from different natural populations cross [305,307]. This is hypothesized to be due to the evolutionary arms-race between paternal and maternal control of seed size. Within a population, the maternal and paternal seed size control genes have reached equilibrium. However, upon outcrossing to a different population, the paternal alleles “escape” the repression by the female genes, and thus result in larger seeds. However, it is generally recognized and shown that maternal parents influence seed size more than paternal parents [165,303,306]. The size of the seed can accurately be predicted in some cases [304]. Breeders can take advantage of this phenomenon to control seed size to suit agriculture by selecting alleles that give uniform seed size.

3. Methods

The population described in chapter 3 was used in this analysis. It was obtained by crossing GND (female) to IvP 35. This population as well as the analytical methods used were described in the previous chapter. The seed phenotypes (size, spotting) were determined before the seeds were sowed. The individuals used in this study thus came from one of four possible seed categories: large spotted (LSP), large non-spotted (LNS), small spotted (SSP) and small

non-spotted (SNS). I did not phenotype the progeny of the individuals that came from these seeds.

4. Results

4.1. Phenotypes - seed spot and seed size

IvP35 was bred by Hermsen & Verdenius [146] to be homozygous for the embryo spot for the purpose of identifying haploids following haploid induction crosses. IvP35 and other IvP lines were tested and confirmed to be homozygous for embryo spot genes (BBPP) by crosses to a tester line that is recessive for both color genes (bbpp) [262]. It also has deep purple flowers, flower/fruit abscission zone, tuber flesh and skin, internodes, and stem. GND on the other hand has light purple flowers, white tuber flesh, white tuber skin, purple-red abaxial leaf surfaces, and lightly colored green stems.

I expected 100% of the F1 seeds from GND x IvP35 to be spotted. However, some of the seeds in the F1 were non-spotted. These non-spotted seeds produced diploid seedlings (confirmed using flow cytometry) as did the spotted. Thus the lack of spots was not due to the formation of haploids. Most likely, this is a result of blocking of the embryo-spot character by a gene in GND. A count of 1505 F1 seeds showed that only 57.3% of them were spotted.

Additionally, the F1 and F2 plants produced from this cross exhibited diverse flower colors that included purple, deep purple, light purple, and sometimes purple-tinted white, white, or white with speckles. Variations in the rest of the plant body were also observed including in the tuber flesh and skin color, the stem color, and leaf color. I hypothesize that GND is heterozygous for a negative regulator of one or more of the embryo-spot genes. This gene also affects colors in other organs [295,308]. From my casual observations, and in line with literature, SSP and LSP (spotted seeds) plants had deeper colors compared to the LNS and SNS plants.

A similar phenomenon was observed for seed size with 38.2% of the F2 seeds being small. All four phenotypic categories i.e large spotted, small spotted, large non-spotted, and

small non-spotted appear in both F1 and F2. The size of the seeds that gave rise to the parental lines cannot be determined. Of the 1128 F2 seeds counted, 38.4% were spotted and 31% were small. Seed size was binary with seeds either being large or small. By visual observation, 318 F2 seeds were separated into large and small seeds. To quantify and statistically confirm the size difference, the length (longest axis) each seed was measured under a Zeiss dissecting stereoscope and images processed using the Zeiss Zen Imaging software.

The mean seed size for large seeds was 30.31 mm and that of small seed size was 25.8mm (Table 8 and Figure 20). These seed size differences are statistically significant (p-value of pairwise t-test of seed size is less than 2e-16 (Table 9 and Table 10). There is no difference in seed size between spotted and non-spotted seeds. Similar observations (Table 11) were obtained in the F1 generation (n=177).

The first thought upon observation of small seeds from crossing GND (a non-haploid inducing line) to the IvP35 (a haploid inducer) was that they were a result of formation of haploid embryos. However, flow cytometric analysis determined both the large and small seeds produced diploid seedlings. With haploid induction ruled out, the alternate hypothesis I could formulate to explain the difference in seed size was maternal-paternal conflict.

4.2. Sequencing and Haplotyping

The GND allele frequency (at positions fixed for parental SNPs i.e GND is AA and IvP35 is aa) was plotted for the spotted and non-spotted bulks (Figure 21). A region associated with the seed spot suppression trait will have an increased (>0.5) GND allele frequency in the non-spotted bulk and a corresponding decrease in the spotted bulk. A region on chromosome 10 had the highest allele frequency difference (Figure 22).

I proceeded to attempt mapping using these SNPs in the *r/qtl*. However, as with prior experience, this did not yield any significant results. A much bigger population would be needed to detect any signal. My hypothesis is that GND is heterozygous for the allele that regulates

anthocyanin biosynthesis. It is therefore better to map for these genes using SNPs that are heterozygous in GND (Aa) and homozygous in IvP 35 (AA). The Ivp35 SNP has to be one of the two alleles in GND. As such, there will be three haplotypes segregating in the F2 - two from GND and one from IvP35. The mapping strategy was the same as the one presented in chapter 3.

There were 1,935,345 GND heterozygous positions. Of those, I was able to phased 1,196,238 (61%) of SNPs for mapping. To genotype the F1 and F2, the genome was partitioned into 6,511 100 kb non-overlapping bins and the genotype in each bin determined for each sample as previously discussed in Chapter 3 of this thesis. These bins were used as markers for mapping. Mapping was performed using three of the algorithms implemented in the r*/qtl* package using default settings. These algorithms are the EM, HK and multiple imputation algorithms [278,285–288]. Since the two seed traits are binary, mapping was also done using a logistic regression model by specifying the model (model="binary") in the call to the *scanone* mapping function of the *r/qtl* package.

4.3. Mapping results for the seed spot

The HK algorithm detected a QTL on the distal end of chromosome 10 (p-value = 0.025) and a minor one that is not robust on chromosome 8 (p=0.047). The QTL interval is chr10:48500000-56600000 with the peak at chr10:56200000. On the other hand, the less powerful EM algorithm did not detect any QTLs while the multiple imputation algorithm results identified many peaks, which is too noisy for any meaningful interpretation. The plot of the LOD scores from these three algorithms are in Figure 23. Similar results were obtained when the mapping was done using a logistic regression where one QTL was obtained on chromosome 10 with a peak in the bin 60500000. The interval is between chr10:55400000-60800000 (p-val = 0.004). The distal arm of chromosome 10 is known to harbor several anthocyanin biosynthesis genes and transcription factors [298]. This result is encouraging. With an expanded population, the mapping of the seed spot trait is likely to give clearer results.

Attempts to map using the homozygous SNPs (GND AA and IvP aa) or by IvP35 heterozygous SNPs (GND AA and IvP35 Aa) did not identify any QTLs (results not shown).

4.4. Mapping seed size results

Mapping seed size genes was attempted in similar methods as seed spot was mapped above. Preliminary analysis was first done by plotting the GND allele frequency for the large seed and small seed bulks. There were minor differences in allele frequency as would be expected of any two populations. However, these were much smaller than the typical allele frequency differences (0.3 to 0.7) observed in BSA analysis (Figure 24) [281,283,309]. QTL mapping using the three SNP-types used for seed spot mapping above and in chapter 3 did not identify any QTL as was expected from the preliminary analysis (Figure 25). A deliberate and proper experimental design using individuals from reciprocal crosses is needed to map for seed size genes and to identify QTLs that are likely to be involved in maternal-paternal conflict.

5. Discussion

This study, even though not comprehensive, was able to identify a region on the distal end of chromosome 10 that potentially harbors a gene that controls deposition of color on the embryo SAM region, and most likely other organs that are analogous to the embryo spot such as the stem nodes and flower abscission zone [295]. This region has already been shown in several mapping studies to have at least three and up to nine anthocyanin biosynthesis genes including a R2R3-MYB transcription factor and anthocyanin biochemical pathway genes [298,299,301,308,310,311]. These loci were described in the early 1900s and the identity of some of the genes underlying these loci has only been revealed recently [297,301,311,312].

Variation in color is usually controlled by gene expression regulation rather than by the absence or presence of the basic anthocyanin biochemical pathway genes [298,299,301,308,310,311]. The gene responsible for the lack of seed embryo-spot in my material most likely acts in trans to regulate the expression of anthocyanin biosynthesis in these organs. The plants are not

colorless but rather have varying expressions of color in different organs. Alternatively, the gene could be epistatic to the biosynthesis genes either converting the precursor to the colored anthocyanins to other biochemical products or altering the anthocyanins themselves to a different non-colored product. Either way, identifying these genes is critical.

The attempt to map the seed-size genes was not successful, most likely because of the sub-optimal design. Crosses between GND x IvP35 were not reciprocal. The F2 was obtained by pooling pollen from 50 plants and using it to pollinate the same plants. Thus, there is no way to determine which F1 produced which F2. When GND (male) was crossed to a different accession, DM 1-3 (female), all the seeds were large. Seed size within natural populations is usually uniform. However, crossing individuals from two or more populations can lead to differences in seed size most likely due to relaxation of maternal control over paternal genes [304,305,307].

The experimental setup above is not ideal for testing whether the small seed size is a male or female-determined trait. Diallel crosses between individuals in this population would be ideal for this type of study [305]. Nonetheless, I attempted mapping the seed size gene using the population.

I cannot confidently say that seed size differences were due to cross-direction and therefore maternal-paternal conflict. However, even if we were sure that the seed size differences were due to parental conflict, the power to detect QTLs controlling seed size would be low in my setup as these are only expressed from either male or female. Thus the effect of maternally-expressed seed size genes that are transmitted through the male are not apparent; these lead to phenotypic differences only when expressed from the female parent. The paternally-expressed seed size genes act analogously in the opposite direction.

6. Conclusion

The biggest change in commercial potato genetic studies, breeding, and agronomy has been proposed by several leading potato geneticists and breeders [313,314]. The proposal calls for the development of diploid inbred potato varieties that can be used to make hybrids that can be propagated through true potato seed (TPS). The proposal has spawned several startup companies such as Aardevo B.V (USA and Germany) and Solynta B.V (Netherlands) with tens of millions of dollars in funding, and has shifted some focus of major breeding companies and universities towards study and implementation of true-hybrid diploid potato farming and production away from the traditional non-sexual, tuber-propagated tetraploid varieties. This proposal is very ambitious and is almost equivalent to domestication of a new crop. Among things that will have to be considered, is grower and consumer tuber-color preference. Most of the genetic variation that is set to be acquired from the hundreds of diploid wild and cultivated accessions will have genetic drag associated with anthocyanin biosynthesis genes. Understanding and selecting for genes such as the one responsible for turning off the anthocyanin production in the embryo spot will be of paramount importance [295,300,308]. Indeed, the diploid hybrid potato proposal has called for a fixation of major color genes in the various inbreds that will be used for making hybrids [313].

Seed size might also be an important factor. In potato, limited studies suggest that seed size in potato, and other crops, affects the seedling sturdiness which might have important implications for establishment of planted seeds in a field [315]. Seed size has also been shown to affect the tuber yield, with larger seeds having higher yields [315,316]. In other species, larger seeds have been shown to have better emergence from great soil depth and to better tolerate stressful conditions than small seeds [304]. Thus it is important to consider seed size during making of diploid commercial potato varieties not only for design of mechanical seed handling but also for potential impact on tuber yield.

My study, using a small population with non-optimal design, has been able to identify a potential region responsible for pigmentation on chromosome 10. My results were not very conclusive. However, that the region I identified has also been identified in various other studies using different accessions shows that my method of study is robust, save experimental design. A proper and larger study will unequivocally identify the responsible region. The seed size gene might be identified using the material developed in this study.

Chapter 5: Dissertation Conclusion and Summary

The major aim of my study was to further the understanding of potato haploid induction genetics, particularly by mapping the haploid induction gene. Efficient genetic haploid induction systems are implemented in the breeding of two crops only - maize and potato. Even though the maize haploid induction genes have been identified, attempts to utilize them in eudicots have not been successful. The identification of potato HI gene(s) might bridge this gap.

To my knowledge, this is the first study that has established a locus associated with haploid induction in potato. The locus is on the distal end of chromosome three. Identifying the responsible mutation might help transform the HI in other crops such as tomato, soybeans, and fruit trees that currently employ the laborious, and often recalcitrant, haploid production by use of microspore culture.

This study has established a large F2 population that is segregating for HI. The data and insights generated here lay a firm foundation for fine-mapping and eventual identification of the haploid induction gene. Identifying the HI gene is not only important for haploid induction, but also for study of pollen biology. The hallmark of sexual reproduction is the congregation of male and female chromosomes in the egg nucleus. An understanding of how the deviation of this process can occur will be of interest to geneticists, molecular, and evolutionary biologists.

A review of potato literature shows that HI is very common in potato interploidy crosses. I generated the ploidy flux hypothesis to explain why HI is common in these crosses. This hypothesis is testable in other species that, like potato, have mixed ploidy populations. If future studies support this hypothesis, then it is possible that new genes that might be important for haploid induction and other applications such as artificial apomixis or apogamy might be identified [317].

Finally, this study looked at two seed traits - the seed embryo spot and seed size - that were also segregating in the population generated above. I identified a region on chromosome 10 that is derived from GND that suppresses color expression in the embryo shoot apical meristem. A more robust experiment is needed to confirm this. This region has already shown to harbor several other anthocyanin biosynthetic and transcription factors genes.

On the other hand, I did not identify any region associated with seed size variation most likely because experimental setup was subpar. Seed size is usually cross direction dependent. My sequenced individuals did not account for this.

Funding

This work was supported by the United States of America National Science Foundation (NSF) Plant Genome Integrative Organismal Systems Grant 1444612 (*Rapid and Targeted Introgression of Traits via Genome Elimination*) to Luca Comai and Isabelle Marie Henry, the National Science Foundation Graduate Student Fellowship grant 1650042 to me, the University of California Davis Innovation Institute for Food and Health Fellowship (2021) to me, and by the University of California Davis Flow Cytometry Shared Resource Laboratory with funding from the NCI P30 CA093373 (Cancer Center), and S10 OD018223 (Astrios Cell Sorter), and S10 RR 026825 (Fortessa Cytometer) grants, with technical assistance from Ms. Bridget McLaughlin and Mr. Jonathan Van Dyke".

Tables

USDA PI # (accession #)	Name	Note	Source or Collection point	Reference
225682	-	Original accession collection segregating for high HIR	Market in Nariño, Colombia	Hermsen & Verdinius, 1973
GS 29 (Genetic stock 29)	Phu 1.22 (or 225682.22)	Ancestor of IvP haploid inducers (IvP18, 35, 48, 101), and PL4. Identified as a progeny of 225682	PI 225682	Gabert 1963 Hermsen & Verdinius 1973 Hutten et al 1994
PI 225702	-	Accession with embryo spot, and source of IvP embryo spot. Induces haploids at low rates	Valle del Cauca, Colombia	Hermsen & Verdinius 1973

Table 1: Information on the historical lines that were bred for high HIR.

F1 Line	sequencing ID	Berries	Number of haploids	Haploids as percentage of healthy seed (n healthy seed)	Haploids as % of healthy and shriveled seeds (n all seed)	Haploids per berry
SSP 12	ZR330856	9	9	18% (49)	7% (132)	1.00
LSP 04	ZR330849	6	4	19% (21)	6% (62)	0.67
LNS 15	ZR330844	6	3	6% (54)	3% (110)	0.50
LNS 09	NA*	12	5	(NA)	6% (77)	0.42
SNS 01	ZR330827	11	2	5% (43)	1% (145)	0.18
SNS 08	ZR330826	6	1	8% (13)	1% (67)	0.17
SNS 03	ZR33082	6	0	0% (9)	0%(29)	0.00
SSP 01	ZR330814	5	0	0% (78)	0% (102)	0.00
SSP 11	ZR33085	5	0	0% (13)	0% (19)	0.00

Table 2. Haploid induction of F1 plants.

*Individual not sequenced but data is reported to show the variation in F1 HI rates.

F2 Line	sequencing ID	Berries	Number of haploids	Haploids as percentage of healthy seed (n healthy seed)	Haploids as % of healthy and shriveled seeds (n all seed)	HI Rate per berry	HI Category
SSP 345	ZR330860	6	11	31% (35)	13% (86)	1.83	Great
LSP 371	ZR330854	8	10	31.3 % (32)	10% (103)	1.3	Great
LNS 327	ZR330858	5	6	20% (30)	9% (70)	1.20	Great
SNS 517	ZR330859	5	6	38% (16)	7% (87)	1.20	Great
SSP 343	ZR330857	6	7	35% (20)	8% (93)	1.17	Great
LNS 394	ZR330853	6	6	38% (16)	6% (95)	1.00	Great
SSP 530	ZR330855	5	5	6.3% (80)	4% (131)	1.00	Great
LSP 541	ZR330877	6	5	10% (48)	7% (76)	0.83	Very Good
SSP 382	ZR330851	6	5	13% (39)	5% (99)	0.83	Very Good
SSP 344	ZR330852	5	4	11% (36)	6% (67)	0.80	Very Good
SNS 350	ZR330850	4	3	37.5% (8)	11% (27)	0.75	Very Good
SSP 339	ZR330848	6	4	67% (6)	9% (46)	0.67	Very Good
LNS 328	ZR330847	11	7	22% (32)	9% (76)	0.64	Good
SSP 532	ZR330846	5	3	6% (38)	3% (98)	0.60	Good
SNS 353	ZR330843	6	3	33% (9)	9% (35)	0.50	Good
SNS 547	ZR330871	4	2	11% (8)	6% (33)	0.50	Good
SNS 551	ZR330845	4	2	25% (8)	5% (41)	0.50	Good
SSP 340	ZR330873	4	2	11% (19)	2% (85)	0.50	Good
LNS 329	ZR330842	5	2	13% (15)	4% (47)	0.40	Good
LSP 503	ZR330841	5	2	15% (13)	7% (30)	0.40	Good
LSP 565	ZR330840	5	2	3% (59)	2% (81)	0.40	Good
SNS 384	ZR330869	5	2	1% (224)	1% (>224)	0.40	Good
SSP 504	ZR330875	5	2	11% (19)	3% (67)	0.40	Good
SSP 507	ZR330837	9	3	12.5% (24)	4.8% (63)	0.33	OK
LNS 543	ZR330838	6	2	7% (27)	4% (21)	0.33	OK
LSP 538	ZR330839	6	2	1% (162)	1% (227)	0.33	OK
SSP 381	ZR330836	7	2	11% (18)	5% (41)	0.29	OK
LNS 393	ZR330861	4	1	7% (14)	4% (24)	0.25	OK
LSP 361	ZR330864	4	1	7% (14)	2% (52)	0.25	OK
LSP 563	ZR330866	4	1	33% (3)	6% (17)	0.25	OK
SNS 388	ZR330834	4	1	17% (6)	4% (28)	0.25	OK
SSP 379	ZR330835	4	1	17% (6)	5% (18)	0.25	OK
LNS 390	ZR330828	5	1	5% (21)	2% (62)	0.20	OK
LNS 511	ZR330832	5	1	11% (9)	2% (46)	0.20	OK
LSP 525	ZR330829	5	1	5% (19)	2% (63)	0.20	OK
SNS 521	ZR330830	5	1	25% (4)	9% (11)	0.20	OK
SSP 506	ZR330831	5	1	1% (78)	1% (19)	0.20	OK
SSP 531	ZR330833	5	1	5% (20)	2% (51)	0.20	OK
LNS 330	ZR330824	6	1	3% (29)	2% (48)	0.17	OK
LNS 333	ZR330822	6	1	17% (6)	2% (42)	0.17	OK
LNS 334	ZR330876	6	1	5% (19)	2% (65)	0.17	OK
LSP 537	ZR330825	6	1	8% (12)	3% (31)	0.17	OK
SNS 351	ZR330823	6	1	10% (10)	2% (41)	0.17	OK
LSP 354	ZR330821	9	1	4% (24)	3% (37)	0.11	OK
LNS 311	ZR330816	5	0	0% (5)	0% (34)	0.00	WT
LNS 332	ZR33081	5	0	0% (12)	0% (27)	0.00	WT
LNS 391	ZR330820	5	0	0% (10)	0% (21)	0.00	WT
LNS 508	ZR33086	6	0	0% (18)	0% (28)	0.00	WT
LNS 509	**ZR330862**	4	0	0% (16)	0% (20)	0.00	WT
LSP 304	ZR33087	6	0	0% (10)	0% (25)	0.00	WT
LSP 360	**ZR330863**	5	0	0% (TMTC)	0% (TMTC)	0.00	WT

LSP 365	**ZR33088**	6	0	0% (6)	0% (34)	0.00	WT
LSP 366	**ZR33089**	5	0	0% (10)	0% (17)	0.00	WT
LSP 373	**ZR330878**	6	0	0% (22)	0% (47)	0.00	WT
LSP 374	ZR330810	7	0	0% (33)	0% (40)	0.00	WT
LSP 522	**ZR330879**	5	0	0% (13)	0% (31)	0.00	WT
LSP 526	ZR330818	5	0	0% (6)	0% (11)	0.00	WT
LSP 536	**ZR330865**	6	0	0% (10)	0% (23)	0.00	WT
LSP 540	ZR330811	6	0	0% (52)	0% (80)	0.00	WT
SNS 322	**ZR330867**	5	0	0% (22)	0% (30)	0.00	WT
SNS 352	ZR330812	6	0	0% (10)	0% (30)	0.00	WT
SNS 357	ZR330868	3	0	0% (419)	0% (437)	0.00	WT
SNS 358	ZR330817	7	0	0% (18)	0% (40)	0.00	WT
SNS 515	**ZR330880**	5	0	0% (61)	0% (83)	0.00	WT
SNS 518	**ZR330881**	1	0	0% (TMTC)	0% (TMTC)	0.00	WT
SNS 520	**ZR330870**	5	0	0% (22)	0% (40)	0.00	WT
SNS 546	ZR330819	5	0	0% (4)	0% (25)	0.00	WT
SNS 549	ZR330813	5	0	0% (16)	0% (33)	0.00	WT
SSP 338	**ZR330872**	4	0	0% (11)	0% (57)	0.00	WT
SSP 341	ZR330815	8	0	0% (13)	0% (31)	0.00	WT
SSP 380	ZR33083	5	0	0% (11)	0% (15)	0.00	WT
SSP 501	**ZR330874**	1	0	0% (27)	0% (29)	0.00	WT
SSP 502	ZR33084	6	0	0% (19)	0% (28)	0.00	WT

Table 3. Haploid induction rates of the sequenced F2 plants.

TMTC is too-many-to-count. Bold *sequencing IDs are* not the actual sequencing IDs but codes during analysis to simplify parsing of my data.

Line	Berries	Dead seed (n)	HI rate	Seed death rate Dead seed/berry
GND	4	13	0	3.25
GS 29 (Phu 1.22)	6	21	0.5	3.50
IVP 101	5	52	1	10.40
IVP 35	7	47	0.1	6.71
IVP 48	4	18	0.3	4.50
LNS 15	6	56	0.5	9.33
LNS 311	5	34	0	6.80
LNS 327	5	40	1.2	8.00
LNS 328	11	44	0.6	4.00
LNS 329	5	32	0.4	6.40
LNS 330	6	19	0.2	3.17
LNS 332	5	15	0	3.00
LNS 333	6	36	0.2	6.00
LNS 334	6	46	0.2	7.67
LNS 344	5	28	0	5.60
LNS 390	5	41	0.2	8.20
LNS 393	4	10	0.3	2.50
LNS 394	6	79	1	13.17
LNS 395	4	32	0.5	8.00
LNS 508	6	10	0	1.67
LNS 509	4	4	0	1.00
LNS 511	5	37	0.2	7.40
LNS 543	6	21	0.3	3.50
LSP 04	6	41	0.7	6.83
LSP 304	6	15	0	2.50
LSP 354	9	13	0.1	1.44
LSP 361	4	38	0.3	9.50
LSP 365	6	28	0	4.67
LSP 366	5	7	0	1.40

LSP 371	8	81	1.3	10.13
LSP 373	6	25	0	4.17
LSP 374	7	7	0	1.00
LSP 377	5	10	0.2	2.00
LSP 503	5	17	0.4	3.40
LSP 506	5	4	0	0.80
LSP 522	5	18	0	3.60
LSP 525	5	44	0.2	8.80
LSP 526	5	5	0	1.00
LSP 536	6	13	0	2.17
LSP 537	6	19	0.2	3.17
LSP 538	6	65	0.3	10.83
LSP 540	6	28	0	4.67
LSP 541	6	28	0.8	4.67
LSP 563	4	14	0.3	3.50
LSP 565	5	22	0.4	4.40
SNS 01	11	102	0.2	9.27
SNS 03	6	20	0	3.33
SNS 08	6	67	0.2	11.17
SNS 322	5	8	0	1.60
SNS 350	4	19	0.8	4.75
SNS 351	6	31	0.2	5.17
SNS 352	12	20	0	1.67
SNS 353	6	26	0.5	4.33
SNS 357	3	18	0	6.00
SNS 358	7	22	0	3.14
SNS 388	4	22	0.25	5.50
SNS 515	5	22	0	4.40
SNS 517	5	71	1.2	14.20
SNS 520	5	18	0	3.60
SNS 521	5	7	0.2	1.40

SNS 546	5	21	0	4.20
SNS 547	4	15	0.5	3.75
SNS 549	5	17	0	3.40
SNS 551	4	33	0.5	8.25
SSP 01	5	24	0	4.80
SSP 11	5	6	0	1.20
SSP 12	9	83	1	9.22
SSP 338	4	46	0	11.50
SSP 339	6	40	0.7	6.67
SSP 340	4	66	0.5	16.50
SSP 341	8	18	0	2.25
SSP 343	6	73	1.2	12.17
SSP 344	5	31	0.8	6.20
SSP 345	6	51	1.8	8.50
SSP 379	4	15	0.25	3.75
SSP 380	5	4	0	0.80
SSP 381	7	23	0.3	3.29
SSP 382	6	60	0.8	10.00
SSP 501	1	2	0	2.00
SSP 502	6	9	0	1.50
SSP 504	5	48	0.4	9.60
SSP 506	5	19	0.2	3.80
SSP 507	9	39	0.3	4.33
SSP 530	5	51	1	10.20
SSP 531	5	31	0.2	6.20
SSP 532	5	50	0.6	10.00

Table 4. HIR and seed death rate for F1 and F2 lines, GND, IvPs and Phu 1.22.

HIR (haploid induction rate) and seed death are measured as the number of haploids per berry and the number of dead seeds per berry, respectively. Seeds from each harvest were counted for presence of spotted, non-spotted, and dead seeds.

1	Atlantic
2	Burbank
3	DM 1-3
4	Dakota Diamond
5	Early Rose
6	Garnet Chili
7	IPK Desiree
8	Irish Cobbler
9	Kalkaska
10	Katahdin
11	Kennebec
12	LOP868
13	Missaukee
14	Mountain Rose
15	Norland
16	PI 195191 Phureja
17	PI 195204 Stenotomum
18	PI 210044 Solanum candolleanum
19	PI 214421 Andigena
20	PI 225710 Phureja
21	PI 230512 Stenotomum
22	PI 234011 Stenotomum
23	PI 234013 Stenotomum
24	PI 243469 Phureja
25	PI 245847
26	PI 245935 Andigena
27	PI 245940 Andigena
28	PI 258874 Andigena
29	PI 258885
30	PI 265863 Solanum candolleanum
31	PI 265872 Solanum medians

32	PI 275139 Solanum chacoense
33	PI 275216 Solanum ehrenbergii
34	PI 275260 Solanum verrucosum
35	PI 292110 Stenotomum
36	PI 296126 Solanum raphanifolium
37	PI 320355 Phureja
38	PI 365339 Solanum chomatophilum
39	PI 365344 Stenotomum
40	PI 365345 Andigena
41	PI 458324 Solanum infundibuliforme
42	PI 458355 Solanum microdontum
43	PI 458365 Solanum berthaultii
44	PI 458368 Solanum okadae
45	PI 458424 Solanum jamesii
46	PI 472978 Solanum spegazzinii
47	PI 473065 Solanum gourlayi
48	PI 473305 Solanum vernei
49	PI 473385 Solanum sparsipilum
50	PI 498112 Solanum brevicaule
51	PI 498359 Solanum kurtzianum
52	PI 545964 Solanum boliviense
53	PI 545987 Solanum leptophyes
54	PI 546000 Solanum megistacrolobum
55	PI 546023 Andigena
56	PI 558050 Solanum commersonii
57	PI 558142 Andigena
58	PI 558288 Solanum etuberosum
59	PI 607886 Andigena
60	Premier Russet
61	Purple Majesty
62	Rio Grande Russet

63	Russet Burbank
64	Russet Norkotah
65	Shepody
66	Sierra Gold
67	Snowden
68	Spunta
69	Superior
70	USDA Desiree
71	Yukon Gold

Table 5 - Potato varieties, wild accessions and wild species in the K71-variant panel.

Mutation type	Predicted effect	Number of mutations
5 prime UTR variant	Modifier	27
3 prime UTR variant	Modifier	95
Intron variant	Modifier	1176
Splice region variant	Low	2
Splice region variant & synonymous variant	Low	4
5 prime UTR premature start codon gain variant	Low	7
Splice region variant & intron variant	Low	18
Synonymous variant	Low	208
Missense variant & splice region variant	Moderate	7
Missense variant	Moderate	285
Splice donor variant & intron variant	High	1
Start lost	High	1
Stop gained & splice region variant	High	1
Stop lost & splice region variant	High	1
Splice acceptor variant & intron variant	High	3
Stop gained	High	11
Total		**1847**

Table 6. The effect of IvP35 heterozygous SNPs in the QTL on protein function.

Only 1847 of the 7319 SNPs within the QTL region were predicted to have an effect on protein.

Gene	Annotation
Soltu.DM.03G014820	RNI-like superfamily protein
Soltu.DM.03G014870	NAD(P)-linked oxidoreductase superfamily protein
Soltu.DM.03G014910	P-loop containing nucleoside triphosphate hydrolases superfamily protein
Soltu.DM.03G014920	Tonoplast intrinsic protein 5;1
Soltu.DM.03G015010	Tetratricopeptide repeat (TPR)-like superfamily protein
Soltu.DM.03G015280	Glycosyl hydrolases family 32 protein
Soltu.DM.03G015300	SKU5 similar
Soltu.DM.03G015350	Plant invertase/pectin methylesterase inhibitor superfamily
Soltu.DM.03G015370	Plant invertase/pectin methylesterase inhibitor superfamily
Soltu.DM.03G015440	Plant invertase/pectin methylesterase inhibitor superfamily protein
Soltu.DM.03G015560	Plant invertase/pectin methylesterase inhibitor superfamily protein
Soltu.DM.03G015610	UDP-glucosyl transferase 85A2
Soltu.DM.03G016060	AP2/B3-like transcriptional factor family protein
Soltu.DM.03G016210	DEAD box RNA helicase (PRH75)
Soltu.DM.03G016220	Peptidase family M48 family protein
Soltu.DM.03G016230	Tetratricopeptide repeat (TPR)-like superfamily protein
Soltu.DM.03G016270	Subtilase 1.3
Soltu.DM.03G017120	Leucine-rich repeat transmembrane protein kinase
Soltu.DM.03G017150	Predicted transmembrane protein 161AB domain containing protein
Soltu.DM.03G017530	Remorin family protein
Soltu.DM.03G017570	CAP160 protein
Soltu.DM.03G017600	Arabidopsis thaliana protein of unknown function (DUF821)
Soltu.DM.03G018690	Kunitz family trypsin and protease inhibitor protein
Soltu.DM.03G018830	Hypothetical protein
Soltu.DM.03G018860	Kinase interacting (KIP1-like) family protein
Soltu.DM.03G018880	Xyloglucan endotransglucosylase/hydrolase
Soltu.DM.03G018990	Hypothetical protein
Soltu.DM.03G019010	Conserved hypothetical protein
Soltu.DM.03G019030	Homeodomain-like superfamily protein

Table 7. Candidate gene list.

These are “conserved” genes within the LOD interval that have a missense or knock-out allele in IvP35.

Mean seed size (mm)	LNS	LSP	SNS	SSP
B2	30.79395	30.971	26.22964	25.5168
B3	29.1328	29.62103	25.96011	25.69775
B4	30.53041	30.1758	25.46889	26.3745
B5	29.6576	-	-	25.889
Mean total (mm)	30.10659	30.18125	25.74019	25.88346
Total seeds	100	85	73	61
mean standard error	0.227252	0.2708613	0.2774411	0.2157598

Table 8. F2 Seed size measurements.

Several batches (B1-B5) of F2 seeds were scooped with a spatula from the total of F2 seed stock. Seed from each of these batches was then separated into large non-spotted (LNS), small non-spotted, large spotted (LSP) and small non-spotted (SSP) categories. Seed size was considered a binary trait (large or small). The size of each individual seed was then measured to quantify the magnitude of those differences. The data from B1 was omitted. The measurements seem mixed up (i.e SSP measurements were larger than LSP) and so the data set was thrown out (n=60). In total, 318 seeds were analyzed. The values for batch 5 LSP and SNS were not measured. B5 LSP and SNS were not measured.

	B2_LNS	B2_LSP	B2_SNS	B2_SSP	B3_SNS	B3_LSP	B3_SSP	B3_LNS	B4_SNS	B4_LNS	B4_SSP	B4_LSP	B5_LNS
B2_LSP	0.83473	-	-	-	-	-	-	-	-	-	-	-	-
B2_SNS	1.60E-06	3.20E-06	-	-	-	-	-	-	-	-	-	-	-
B2_SSP	2.90E-10	1.40E-08	0.46482	-	-	-	-	-	-	-	-	-	-
B3_SNS	4.10E-12	1.20E-08	0.55929	0.80442	-	-	-	-	-	-	-	-	-
B3_LSP	0.07495	0.10309	0.0001	8.80E-08	1.40E-08	-	-	-	-	-	-	-	-
B3_SSP	2.40E-10	3.00E-08	0.76126	0.55929	0.70286	1.10E-07	-	-	-	-	-	-	-
B3_LNS	0.01283	0.02559	0.00078	2.40E-06	1.10E-06	0.5565	5.40E-06	-	-	-	-	-	-
B4_SNS	1.40E-08	8.40E-08	0.46827	0.95316	0.78828	1.60E-06	0.56914	1.70E-05	-	-	-	-	-
B4_LNS	0.72229	0.62798	5.90E-06	1.40E-08	2.80E-09	0.21189	1.70E-08	0.05149	1.10E-07	-	-	-	-
B4_SSP	1.90E-06	3.80E-06	0.88358	0.34545	0.44971	0.00012	0.62518	0.00096	0.36952	7.60E-06	-	-	-
B4_LSP	0.3966	0.38024	1.40E-05	1.40E-08	1.10E-09	0.46827	1.40E-08	0.14286	1.90E-07	0.63281	1.70E-05	-	-
B5_LNS	0.10781	0.12751	0.00013	4.00E-07	1.60E-07	0.95316	9.70E-07	0.54939	3.10E-06	0.26514	0.00017	0.53554	-
B5_SSP	2.90E-10	2.90E-08	0.69413	0.61207	0.7686	8.80E-08	0.90593	3.60E-06	0.61982	2.40E-08	0.55929	1.40E-08	8.20E-07

Table 9. P-values pairwise comparison for the F2 seed size categories.

The green highlight shows the expected trend (i.e large seeds are not significantly different from large seeds, spotted or not; ditto small seeds). This was true for all comparisons except for Batch 3 LNS seeds when compared to Batch 2 LNS and LSP. Even then, note that B3 LNS mean seed size of 29.13280 mm is (slightly lower than that of B2 LSP and B2 LNS), whereas the mean seed size for all large seeds is 30.1 mm and mean small seed size is 25.8 (Table 8).

F2 seed (Mean seed size)	LNS (30.10 mm)	LSP (30.18 mm)	SNS (25.74 mm)
LSP (30.18 mm)	0.83	-	-
SNS (25.74 mm)	<2e-16	<2e-16	-
SSP (25.88 mm)	<2e-16	<2e-16	0.82

Table 10. Pairwise t-test comparison of F2 seed categories.

P-values for comparison for all the F2 seed categories (n=318) in the bulked batches (B2-B5, Table 8)

F1 seed category (mean seed size)	F1_SSP (21.83 mm)	F1_SNS (23.47 mm)	F1_LSP (29.64 mm)
F1_SNS (23.47 mm)	0.057	-	-
F1_LSP (29.64 mm)	3.50E-11	2.00E-16	-
F1_LNS (29.04 mm)	1.90E-10	2.00E-16	0.071

Table 11. P-value for pairwise t-tests for the F1 seed categories (n=177).

Mean seed size values are shown in brackets. Levene's test for homogeneity of variance is 0.3049 and the p-value for an ANOVA test is p-value < 2.2e-16.

Figures

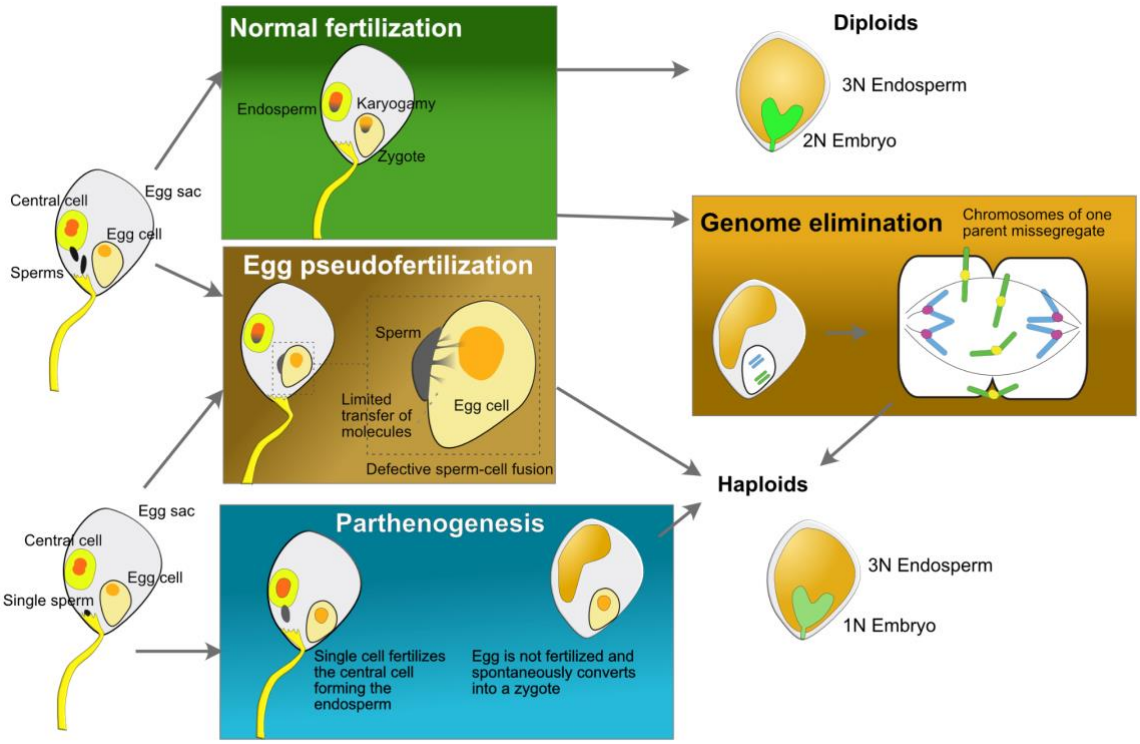

Figure 1. Models for haploid induction in maize.

Parthenogenesis, genome elimination, pseudo-fertilization models for haploid induction are represented for a conceptual diploid. For simplicity, polytubey is not represented here and these processes are explained in the context of a regular diploid x diploid cross. A well-documented example of genome elimination is *CenH3*-based haploid induction and wheat interspecific haploid inducing crosses [4,47,101,102]. Figure created by Professor Luca Comai and myself.

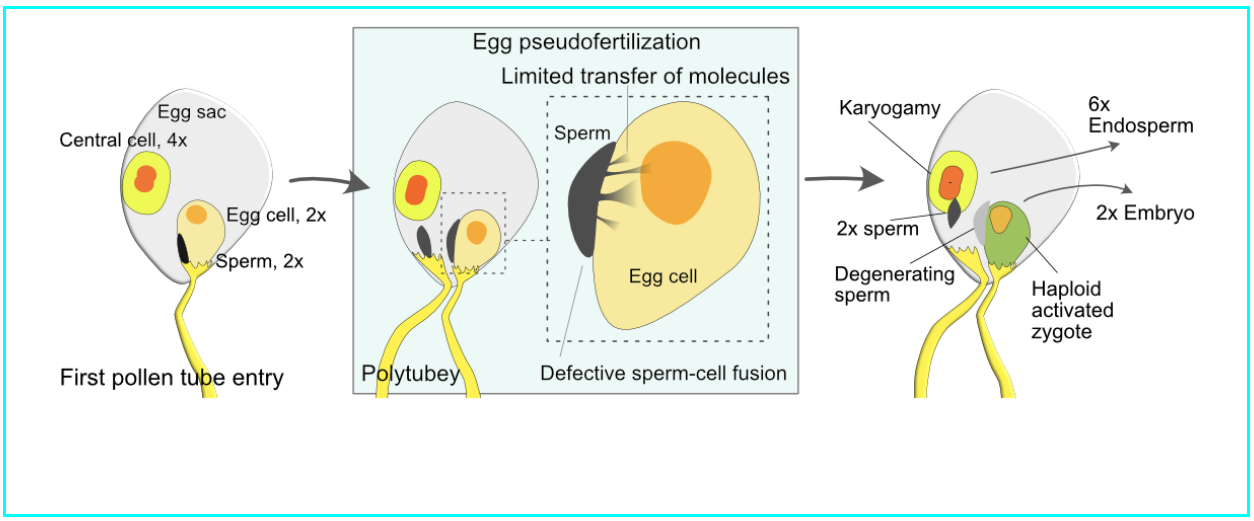

Figure 2. Potato pseudo-fertilization model following a 4X x 2X haploid induction cross.

In the pseudo-fertilization, the first pollen tube sperm does not fuse to the egg resulting in attraction of more pollen tubes. Sperms that attach to the egg can transfer limited material (egg-sperm cell recognition signals, RNA, and other cytoplasmic molecules. These activate zygotic transition in the unfertilized egg. An unreduced pollen sperm fertilizes the central cell to form the 6X endosperm. It is conceivable that a reduced pollen sperm can fertilize the central cell but a pentaploid endosperm is not well tolerated to triploid block. Figure created by Professor Luca Comai and myself.

Figure 3. Spotted, non-spotted, and dead seeds seeds from a haploid induction cross.

The seeds were obtained from eight berries resulting from a cross of cv. Red Polenta (Desiree) to the IvP35 haploid inducer. The spotted seeds (the group on the lower left corner) have the haploid inducer genome. The non-spotted seeds (the group on the upper right corner) could be either accidental selfs (rare for an experienced student) or haploids. The dead seeds (the group on the right) are flat and without embryo and endosperm.

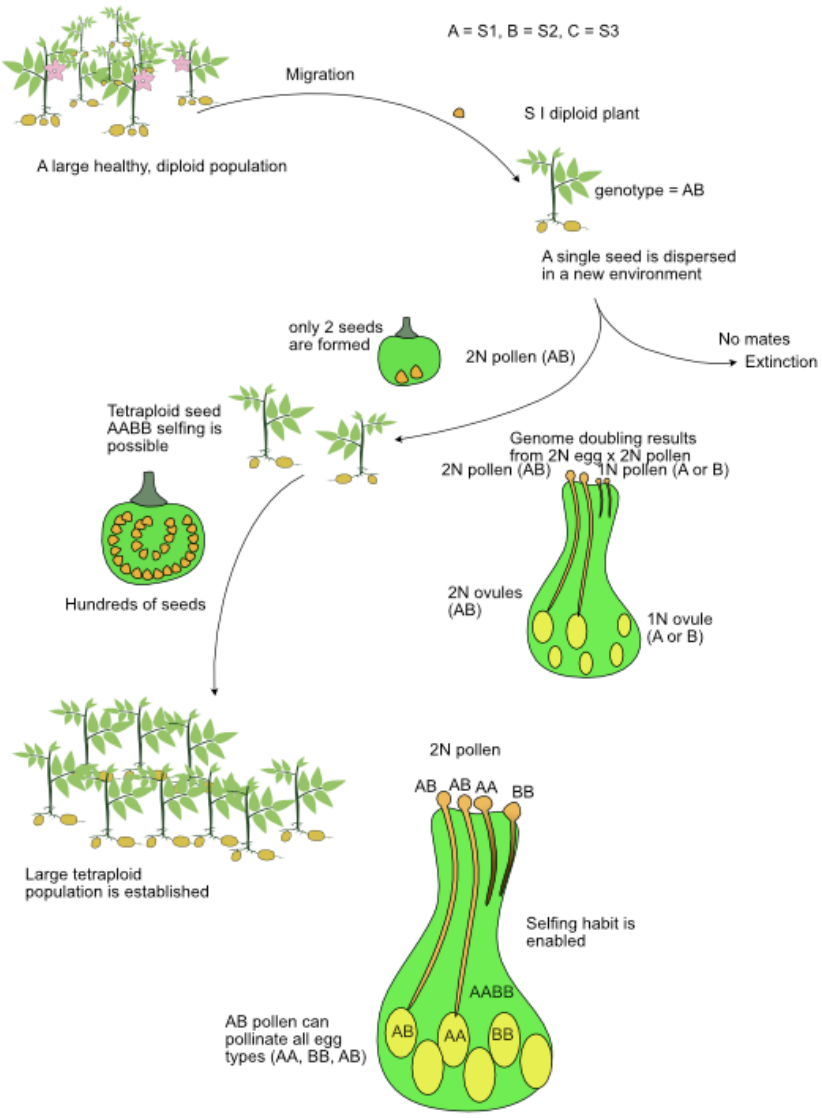

Figure 4a. Tetraploidization of a diploid population.

I designed Figures 4a and 4b but they were drawn by Professor Luca Comai.

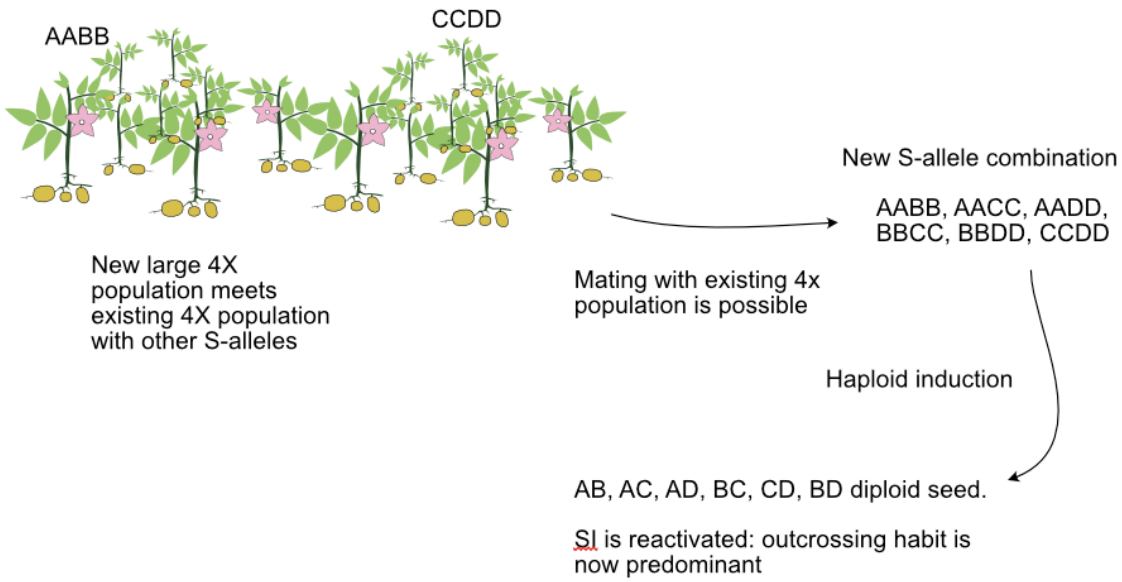

Figure 4b. Tetraploidization of a diploid population.

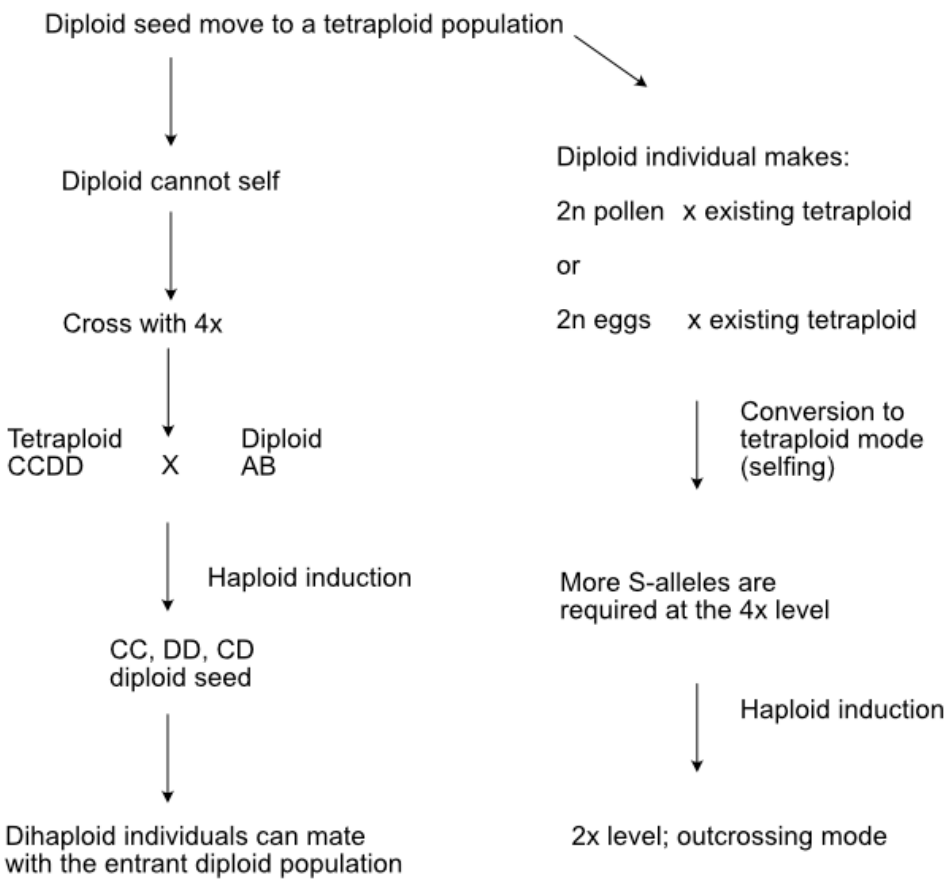

Figure 5: Mating mechanism depending on the population history.

The number S-alleles are a proxy for genetic diversity. In an abundance of S-alleles, diploid level outcrossing mode is favorable as it reduces inbreeding depression. When S-allele diversity is low, selfing and clonal propagation are favored mode of reproduction as these prevent extinction.

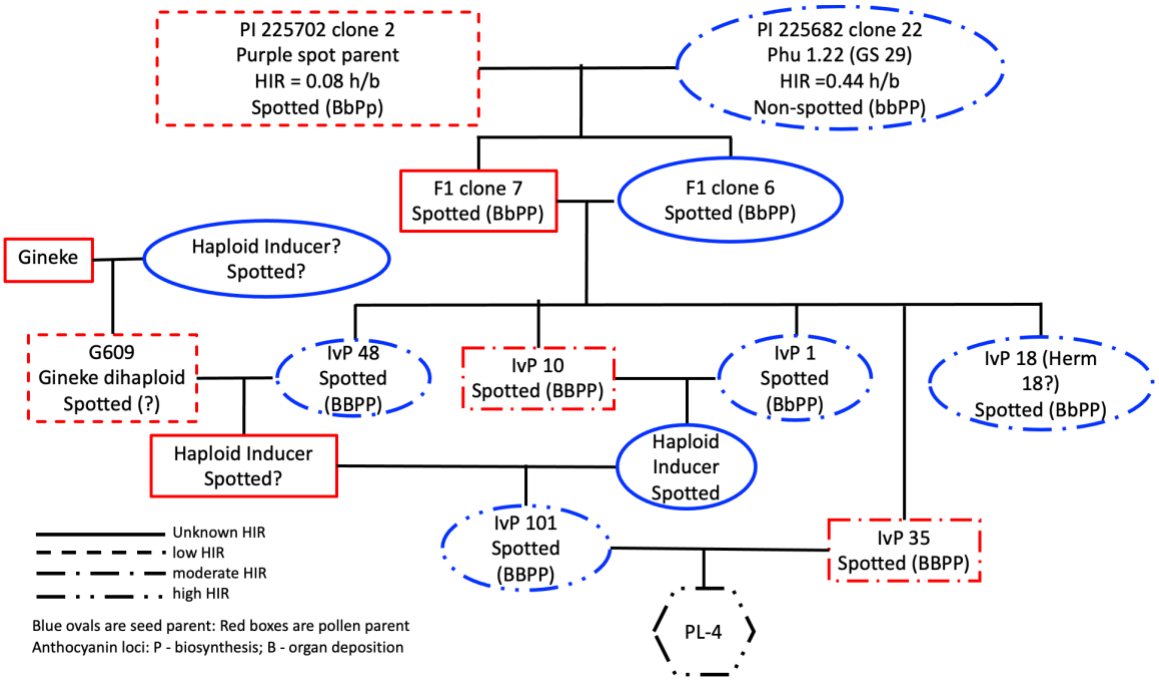

Figure 6. A pedigree of modern haploid inducers.

The B and P loci in this case are anthocyanin biosynthesis genes that give the embryo spot. Homozygous BBPP were bred into the modern widely used inducers to identify as many hybrid seeds as possible. Hybrid seeds will be spotted. The HI rate classification is arbitrary and is by comparing the rate of a line to the rest in the population from which the line was obtained. Blue ovals represent the seed parent while the red boxes are the pollen parent.

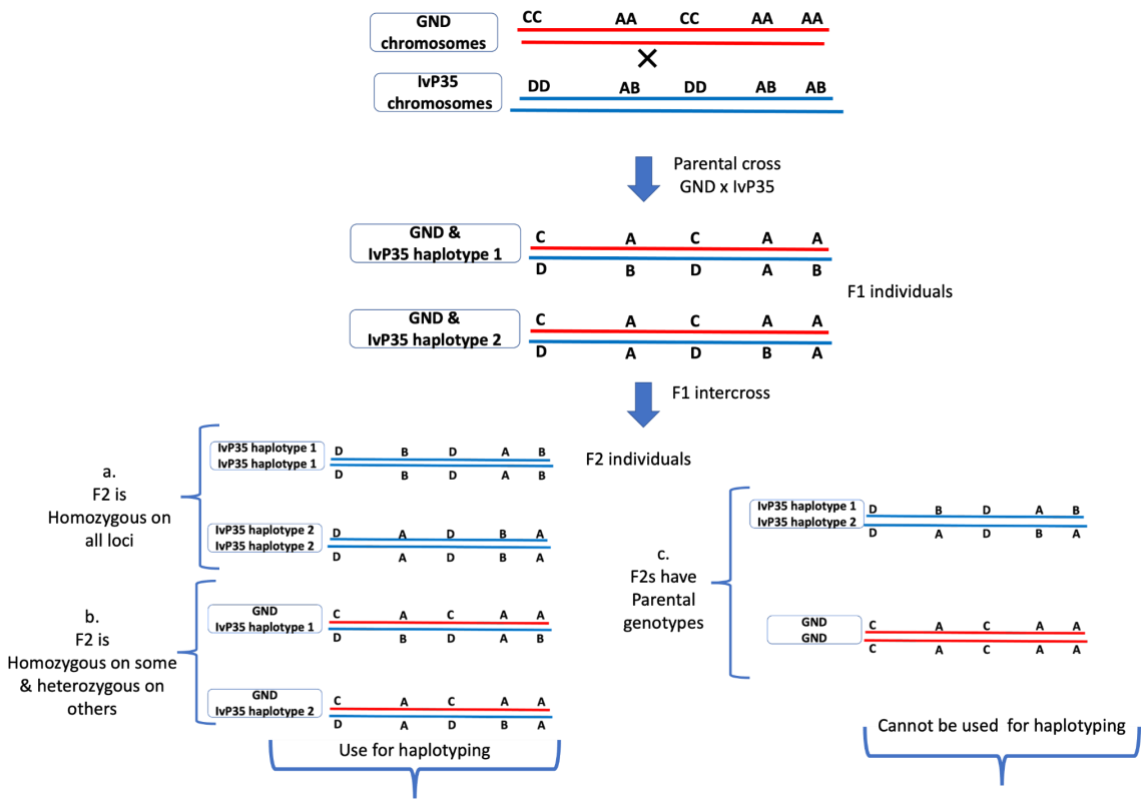

Figure 7. Haplotyping strategy.

To determine if a region in F2 has two chromosomes from GND or IvP35, the IvP35-GND homozygous SNPs (CC and DD respectively) are used.
(a.) A region that is DD and has one IvP35 haplotype will be homozygous at all IvP35-GND heterozygous sites.
(b.) Much like an F1, a region that is CD has one chromosome from GND and the other one from IvP35. It can be used for haplotyping. In this case, some of the GND-IvP35 heterozygous sites will be homozygous (AA) and the others will be heterozygous (AB).
(c.) A region that is DD and has both IvP35 haplotypes or one that is CC (has GND chromosomes only) reconstitute the parental genotypes. These cannot be used for haplotyping.

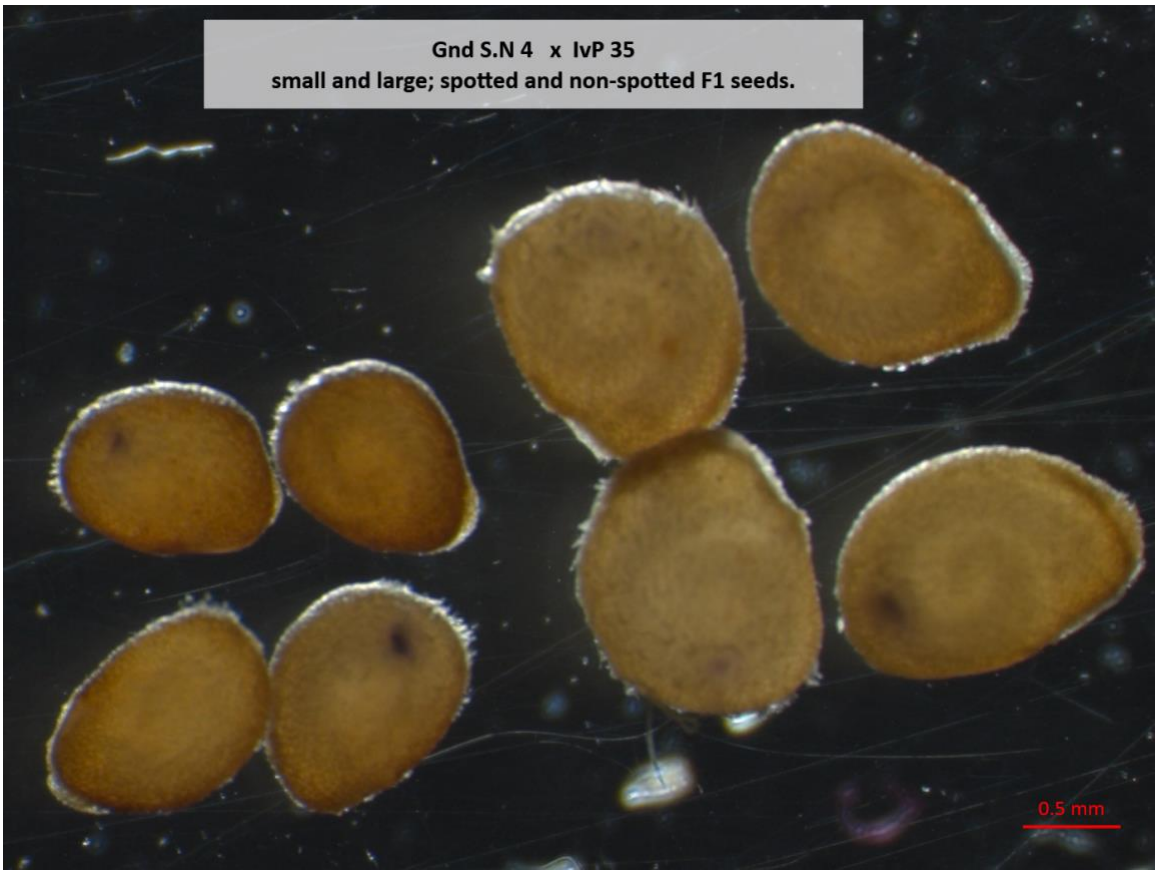

Figure 8. Seed dimorphism - The four types of F1 and F2 seed.

Small spotted (SSP), small non-spotted (SNS), large non-spotted (LNS), and large spotted (LSP).

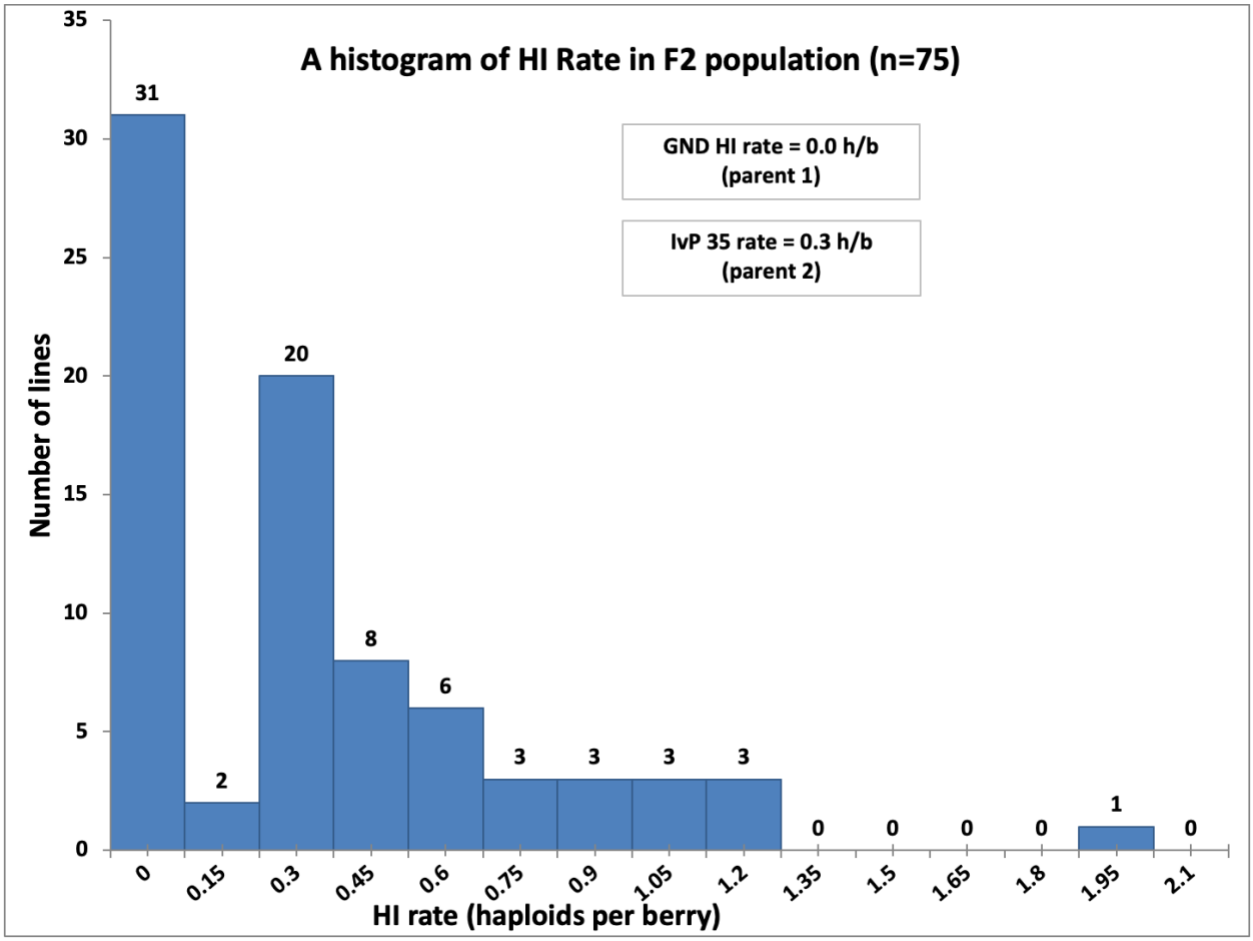

Figure 9. Distribution of HIR in the F2 generation (n=75)

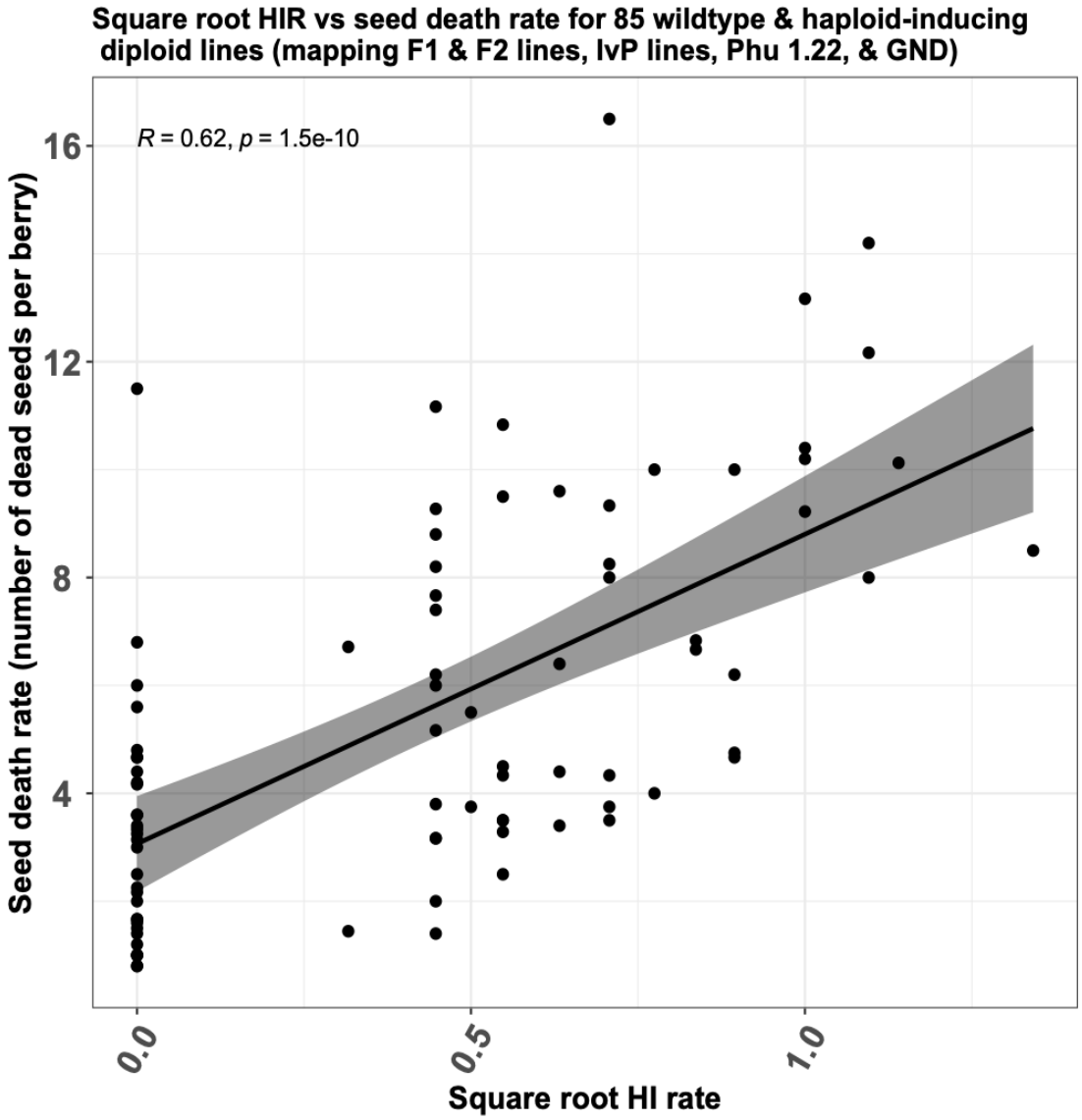

Figure 10. Relationship between HIR and seed death rate.

The population consists of both haploid inducing and wild-type (non-inducing) diploid Phureja lines that include my F1 and F2 mapping population, Phu1.22, IvP lines (IvP18/Herm 18, IvP35, IvP48, and IvP101), and GND. Population-wide HIR (haploid induction rate) is normalized by taking the square root of the HIR. Seed death rate is the number of dead seeds per berry. I am curious if the wild-types with high seed death rate are mischaracterized haploid inducers.

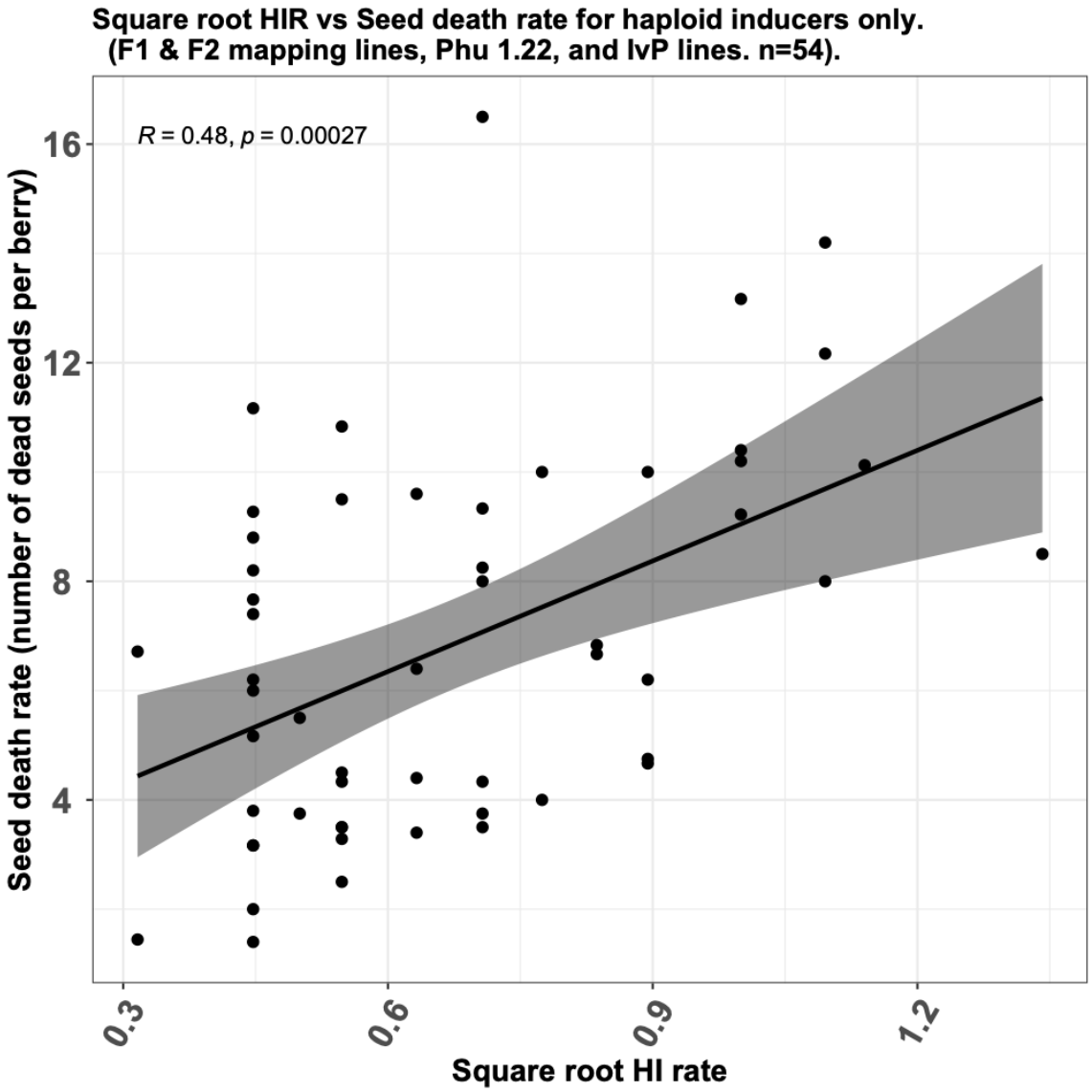

Figure 11. Relationship between (normalized) HIR and seed death rate.

This is for haploid inducing lines only, which subset of individuals analyzed in the previous figure. HIR (haploid induction rate) is in haploids per berry (h/b) while seed death rate is measured as dead seeds per berry.

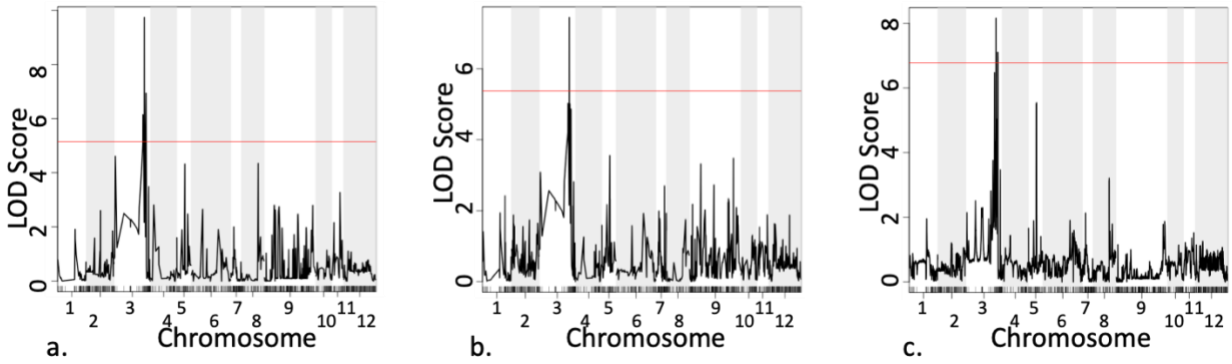

Figure 12. The results of the QTL analysis across the whole genome.

The analysis was performed using heterozygous SNP markers. A QTL was identified on the same region on chromosome 3 using the three algorithms (EM, HK and multiple imputation, respectively) implemented in the r/qtl program. The haploid induction H allele was hypothesized to be on this chromosome using BSA analysis.
No interacting or additive loci were detected (not shown). The peak on chromosome 9 is likely a misplaced or a problematic marker. The inheritance pattern in this region (Figure 13), and a single marker peak are not consistent with a QTL.

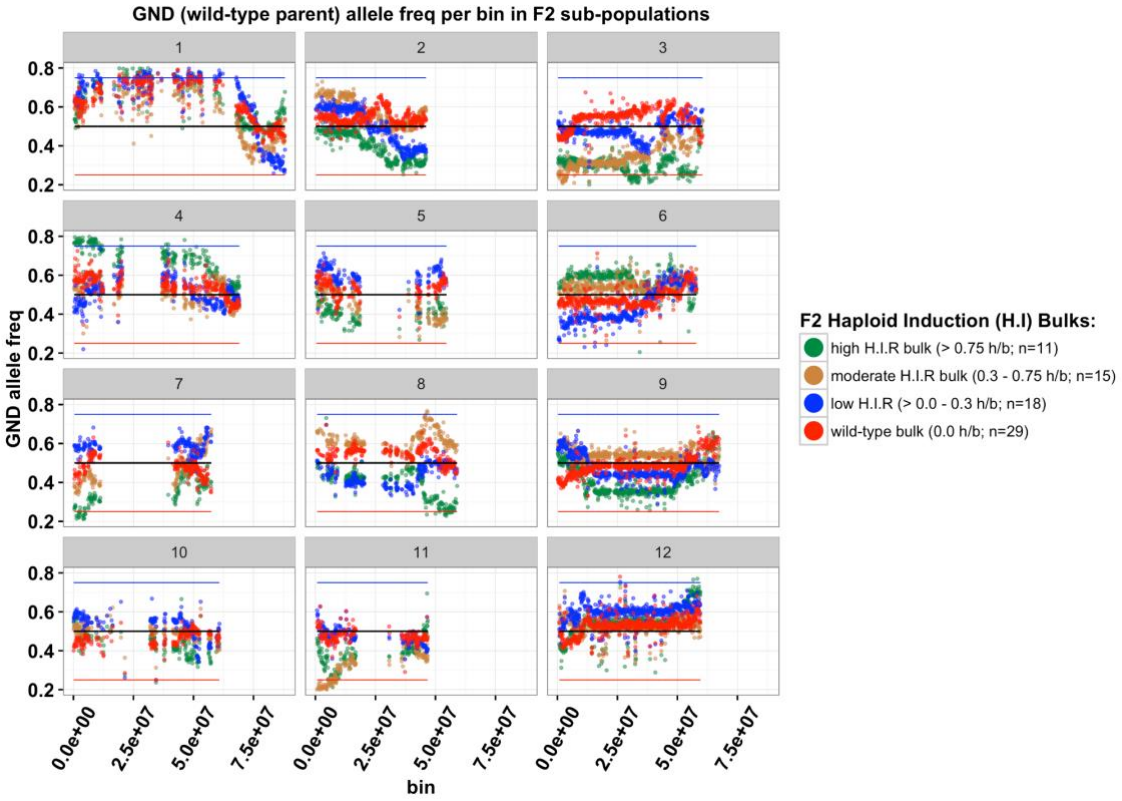

Figure 13. BSA analysis of the 73 F2 plants.

Homozygous SNPs (GND AA and IVP aa) were used. The expected inheritance pattern of GND alleles is as follows: GND allele frequency is high above 0.5 in the zero HIR (WT bulk), just below or around 0.5 in the low HIR bulk, lower in the moderate HI bulk, and very low tending towards zero in the high HIR bulk. This pattern is observed in chromosome 3 only. The scatter plot pattern from the top should be Red, Blue, Brown, and Green for a region associated with HI. The blue, black and red lines mark a frequency of 0.75, 0.5, and 0.25, respectively.

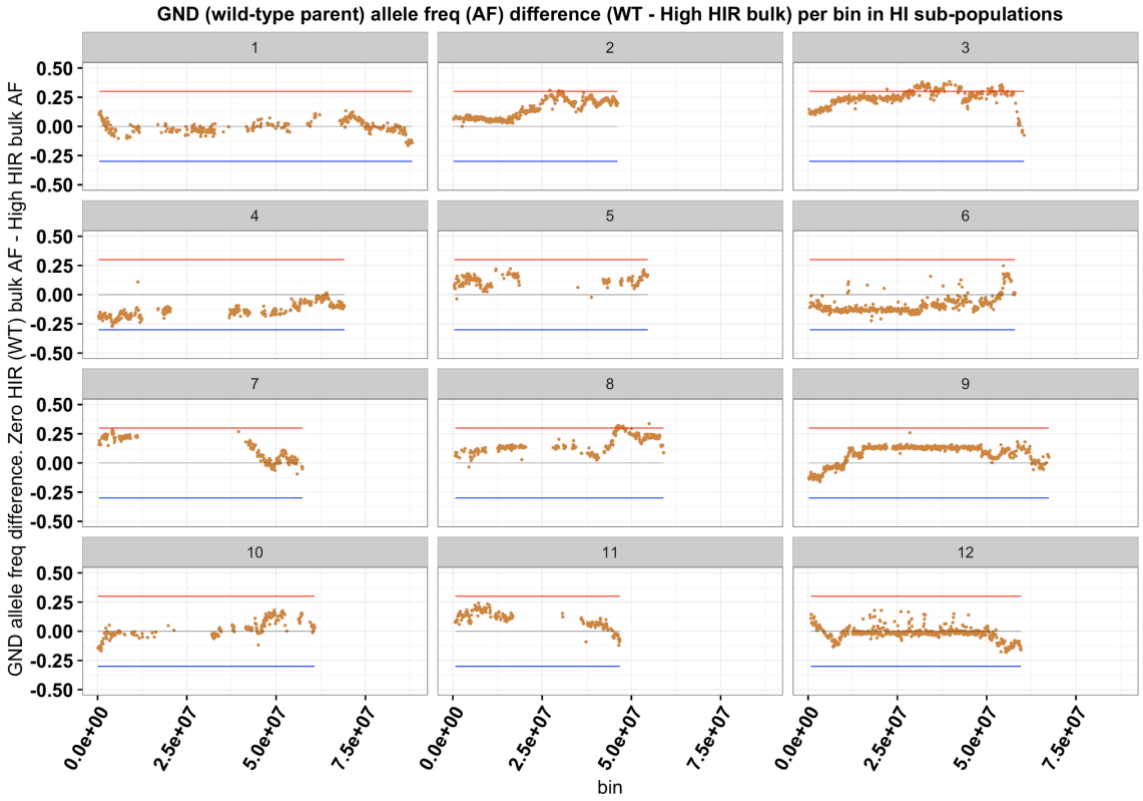

Figure 14. Allele frequency difference between the wild-type bulk and the high HIR bulk.

The highest difference is on chromosome 3. The allele frequency difference between regions on chromosome 2 and on chromosome 8 are almost as large as that on chromosome 3, the pattern of this allele frequency in the different bulks (figure 13) is not as hypothesized i.e wild-type bulk has highest GND allele frequency while high HIR bulk has the lowest frequency.

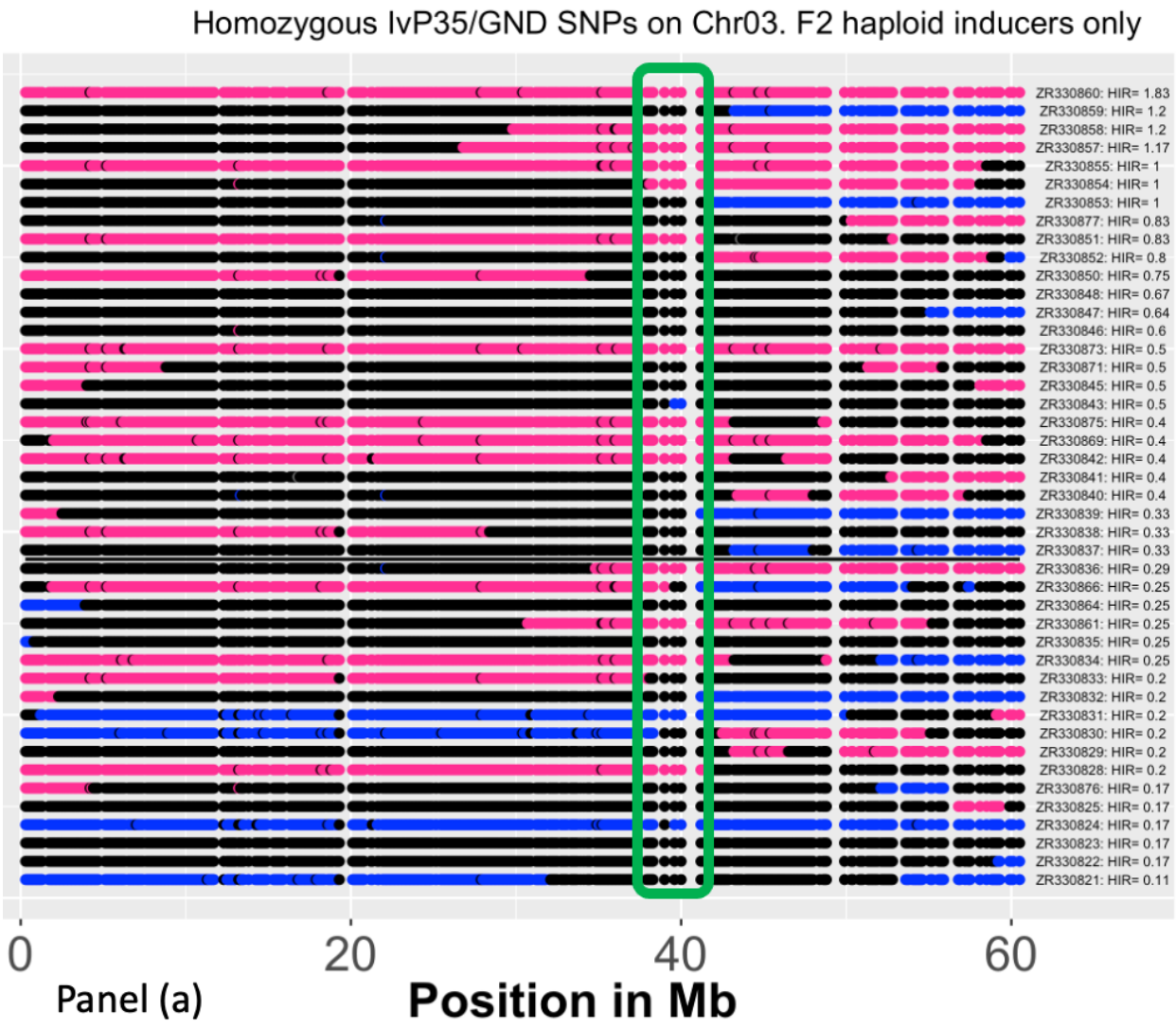

Figure 15. Recombination breaks across chromosome 3 for each F2 haploid inducer.

Homozygous SNPs were used to generate the figure. The blue dots represent GND homozygous regions (AA), red ones represent IvP35 homozygous (BB), while the black ones represent heterozygous (AB) regions. The specific IvP35 haplotype is not specified. Panel (a.) shows the F2 haploid inducers arranged by increasing HI rate. The black line between individuals 18 and 19 (HIR=0.29h/b and HIR=0.33 h/b) shows the IvP35 HI rate (0.3 h/b). Panel (.b) shows the wild-type (non-haploid inducer) individuals.

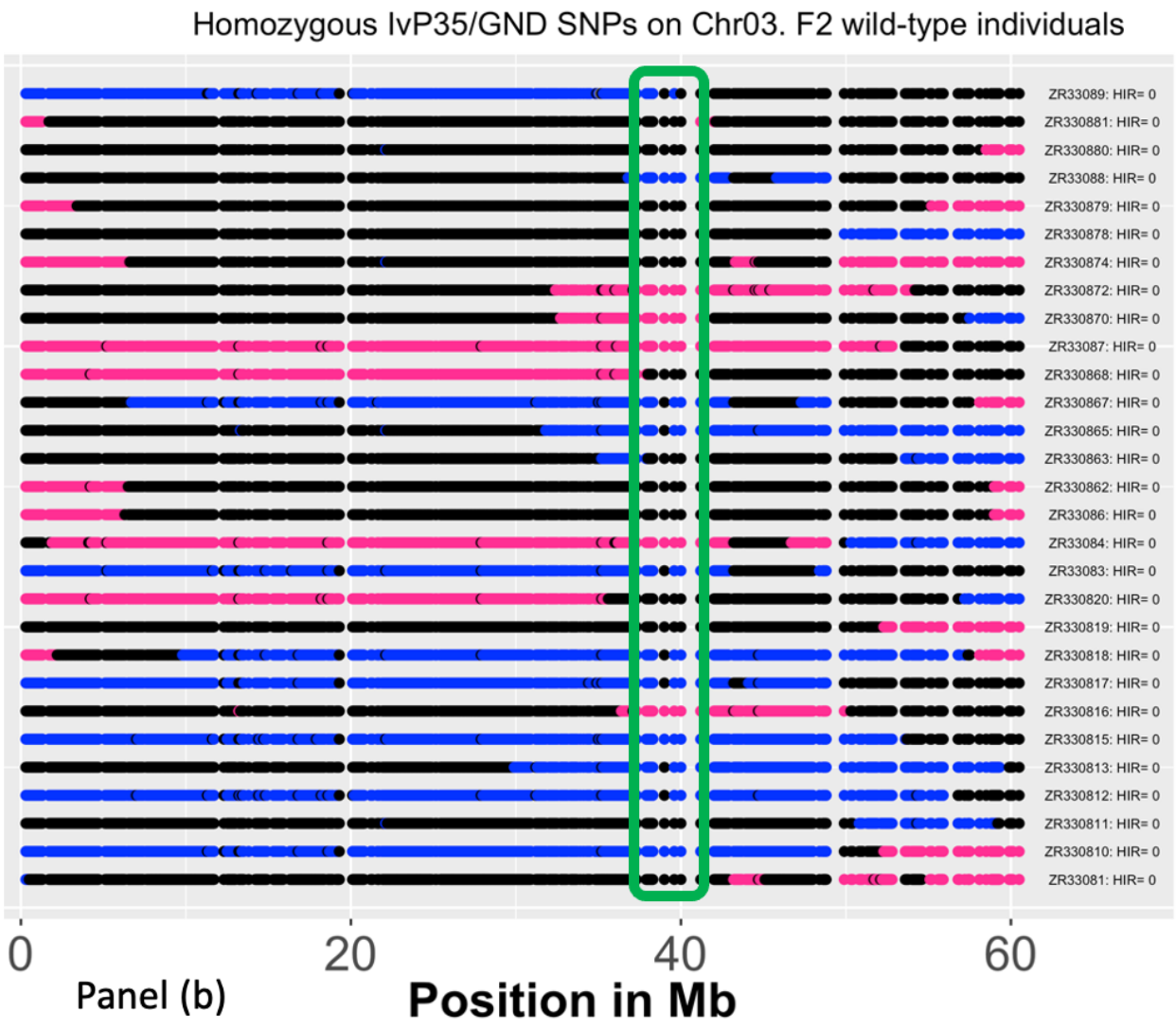

Figure 16. Chromosome 3 recombination breaks for wild-type (non-inducing) F2 plants.

The figure was generated as explained in text and in figure 15.

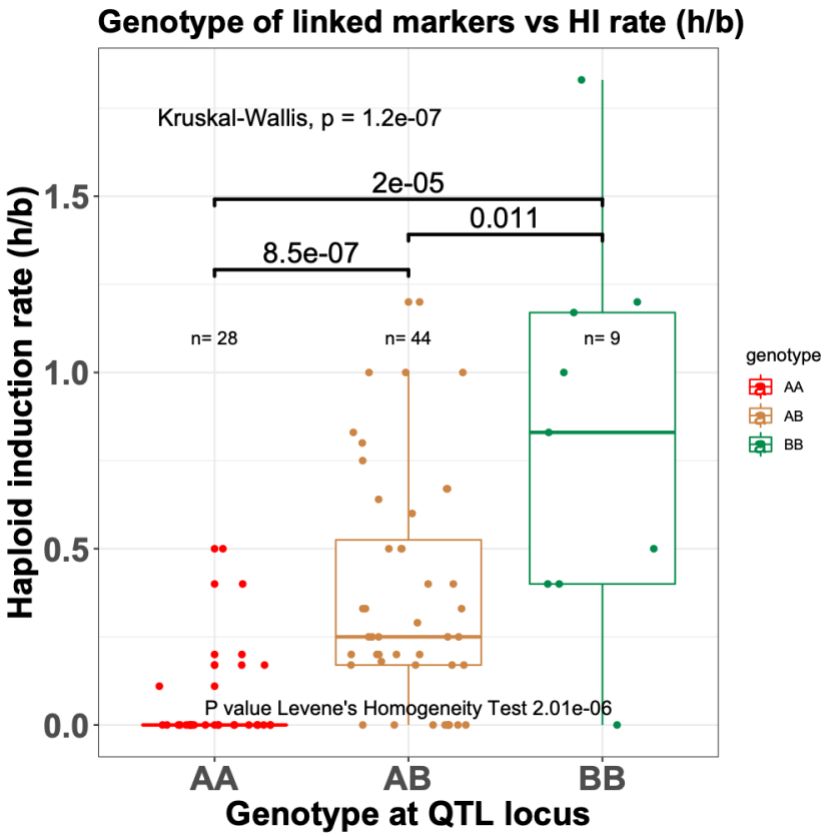

Figure 17. Relationship between HIR (h/b) and genotypes at the H locus.

The variances between the two groups are unequal (p-value Levene's test < 0.05). The non-parametric Kruskal-Wallis test for analysis of variance was used to detect differences between the three genotypes and the non-parametric Wilcox test used for group comparisons. The p-values for the group comparisons are shown on the plot

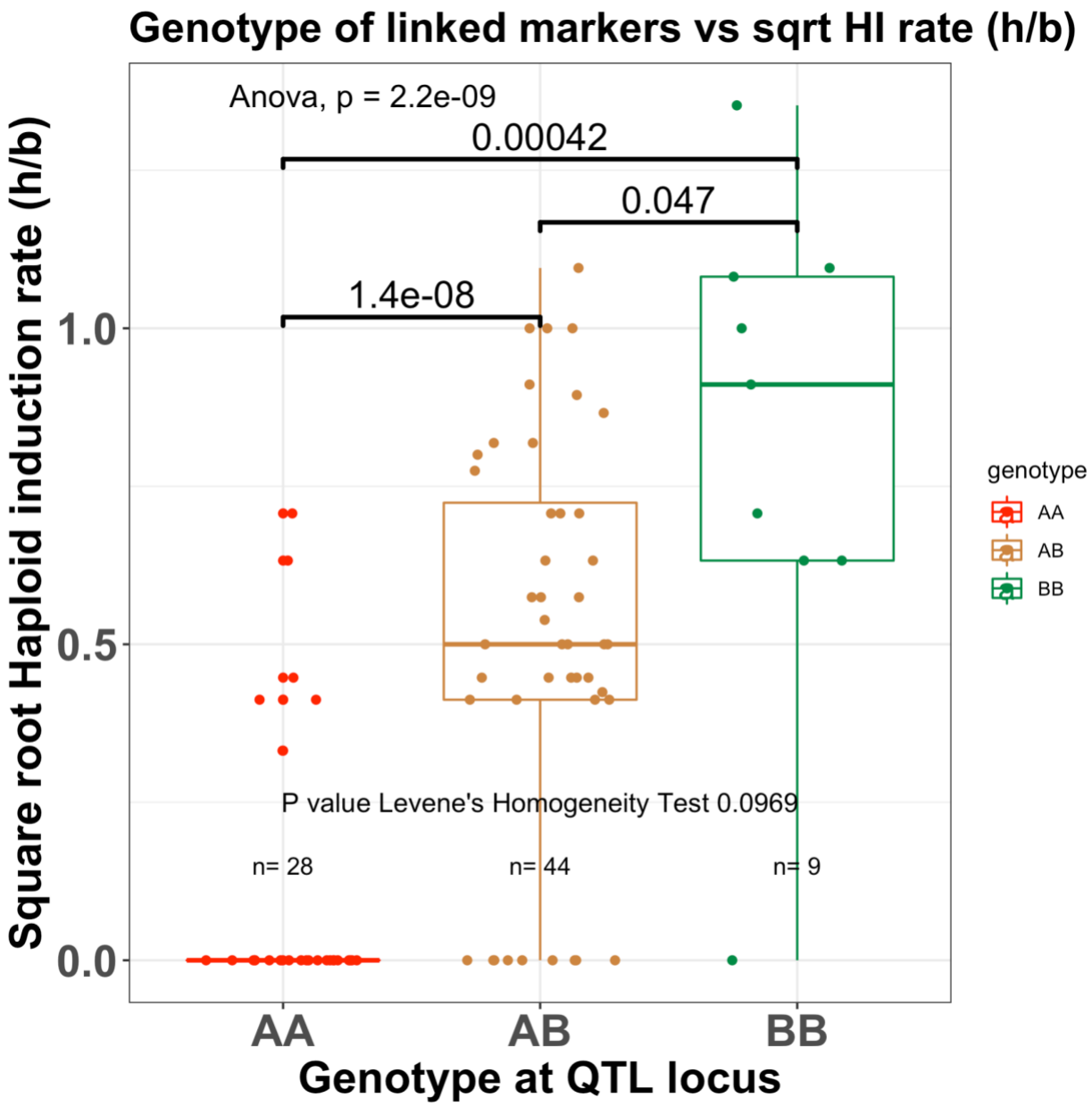

Figure 18. Relationship between normalized HIR (h/b) and genotypes at the H locus.

HI rate data set was normalized by taking the square root. The variance is not significant after the transformation. An ANOVA test was performed and t-tests used for group comparisons. The p-values are shown as above.

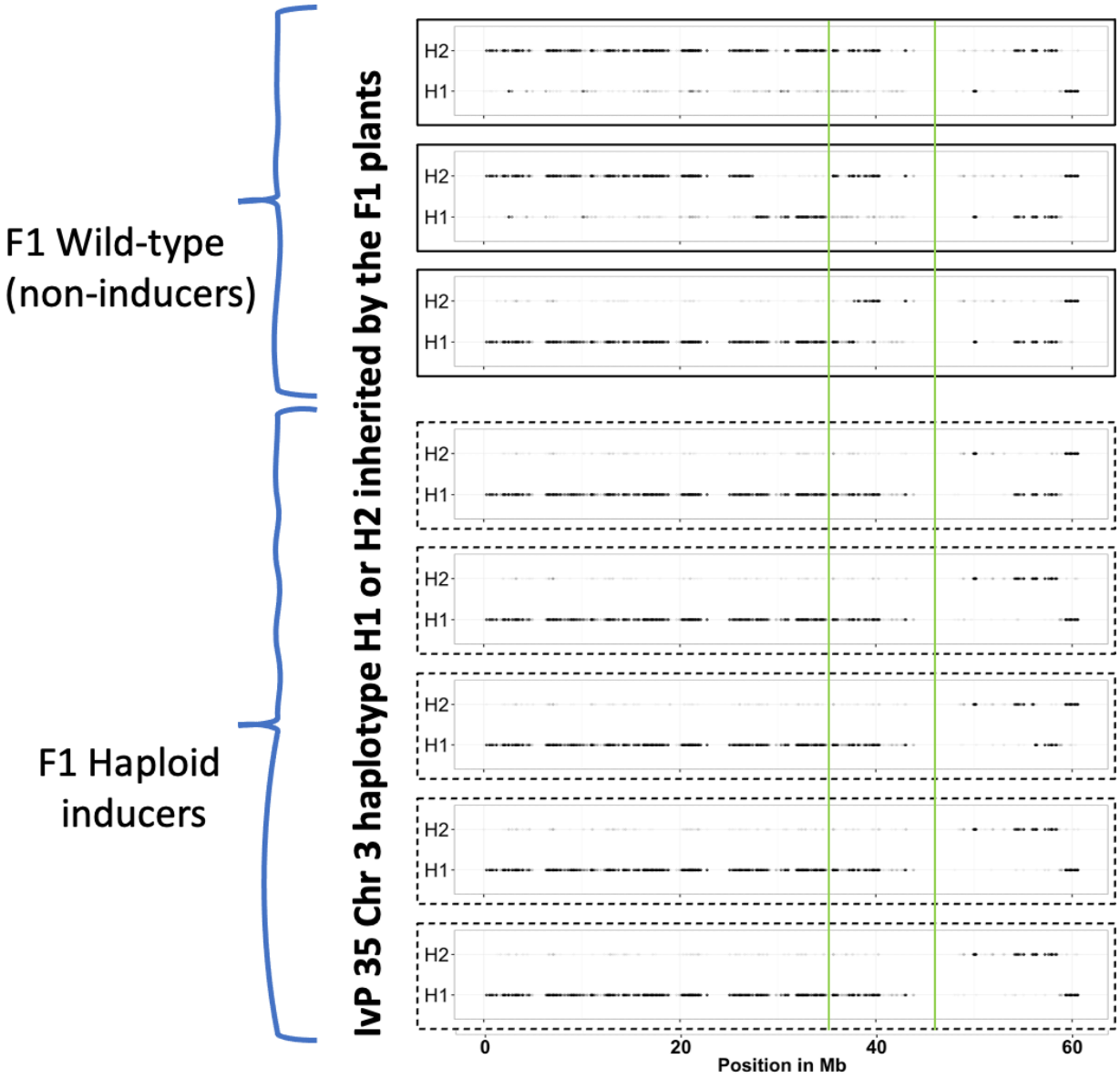

Figure 19. Chromosome 3 IvP35 haplotypes 1 and 2 (H1 and H2) in the F1 lines.

The haplotype that was inherited by each sequenced F1 plant is shown. Note that the three wild-type plants have inherited the H2 haplotype from 37.3 Mb to 43.6 Mb while all inducers inherited the H1 haplotype.

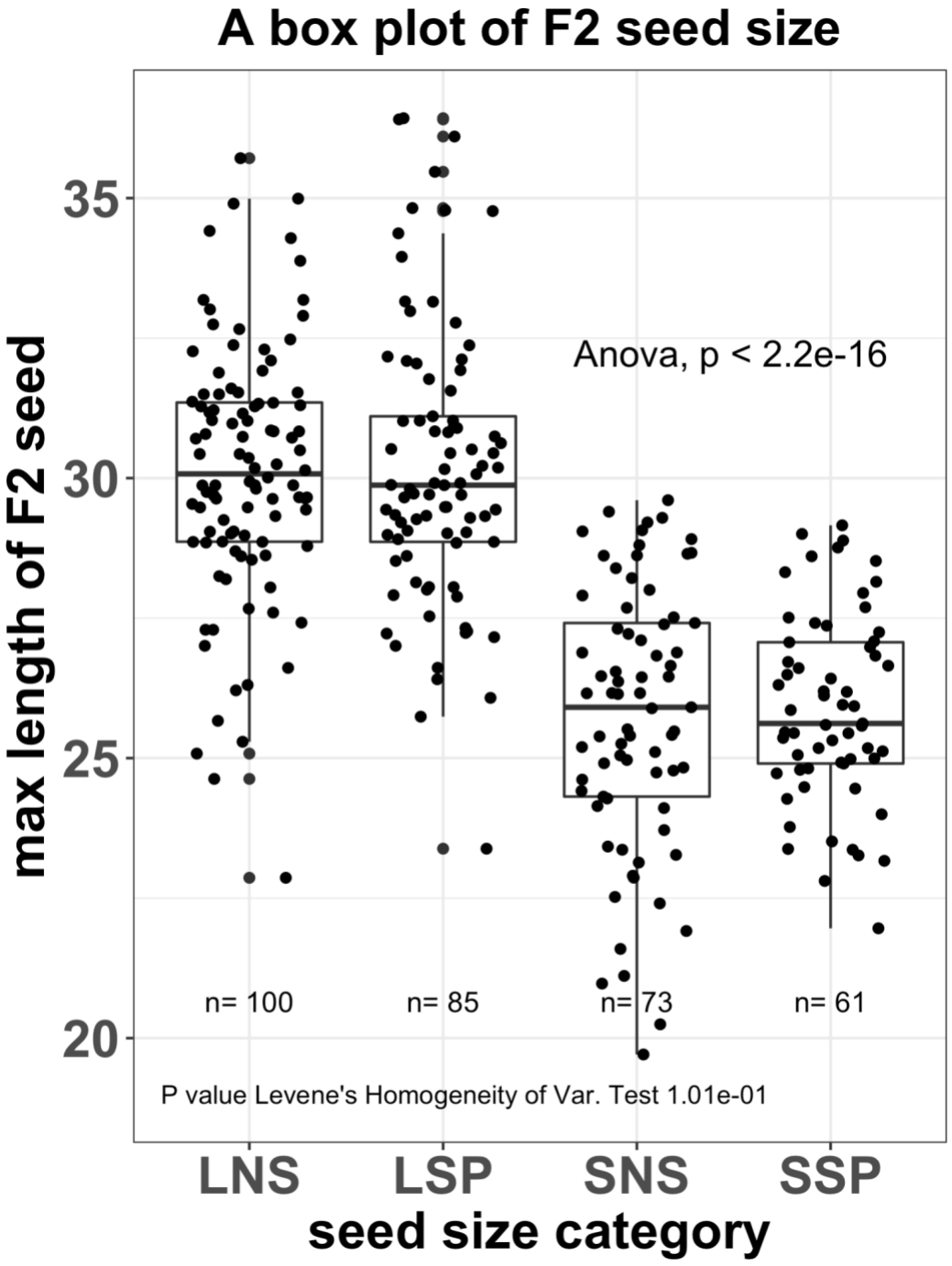

Figure 20. A box plot of F2 seed size.

Seeds were separated into large non-spotted seeds (LNS), large spotted (LSP), small non-spotted (SNS) and small spotted (SSP) seeds by visual examination. The longest axis of the seed was then measured.

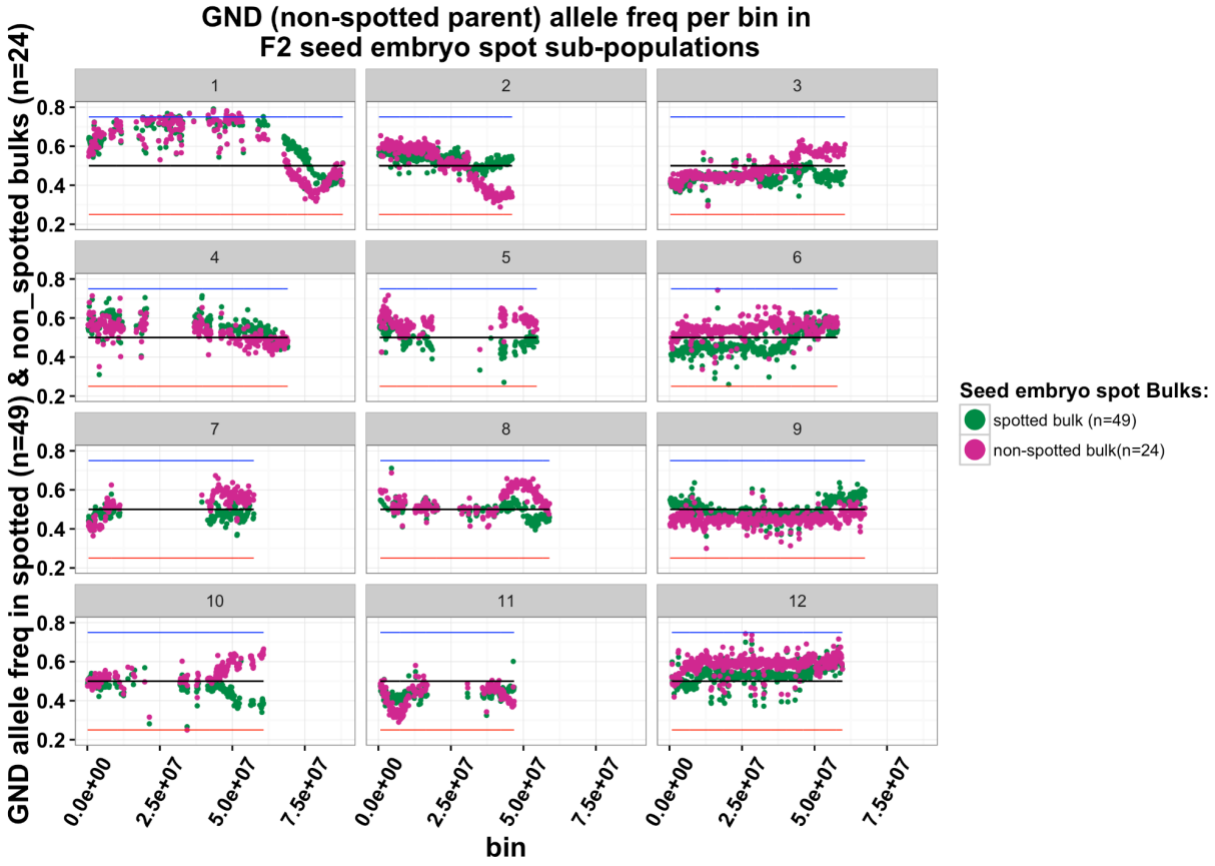

Figure 21: Allele frequencies in spotted and non-spotted seed bulks.

The graph shows the GND allele frequency across the genome. Each dot represents a 100kb bin. The bin genotypes were called by counting the total number of reads covering the parents. A bin was assigned homozygous if >95% of reads came from one parent. The minimum number of informative reads in the bin was 100 reads. Only SNPs that were homozygous in both parents were used. The number of F2 plants in the non-spotted bulk is 24 while the spotted bulk has 49 individuals.

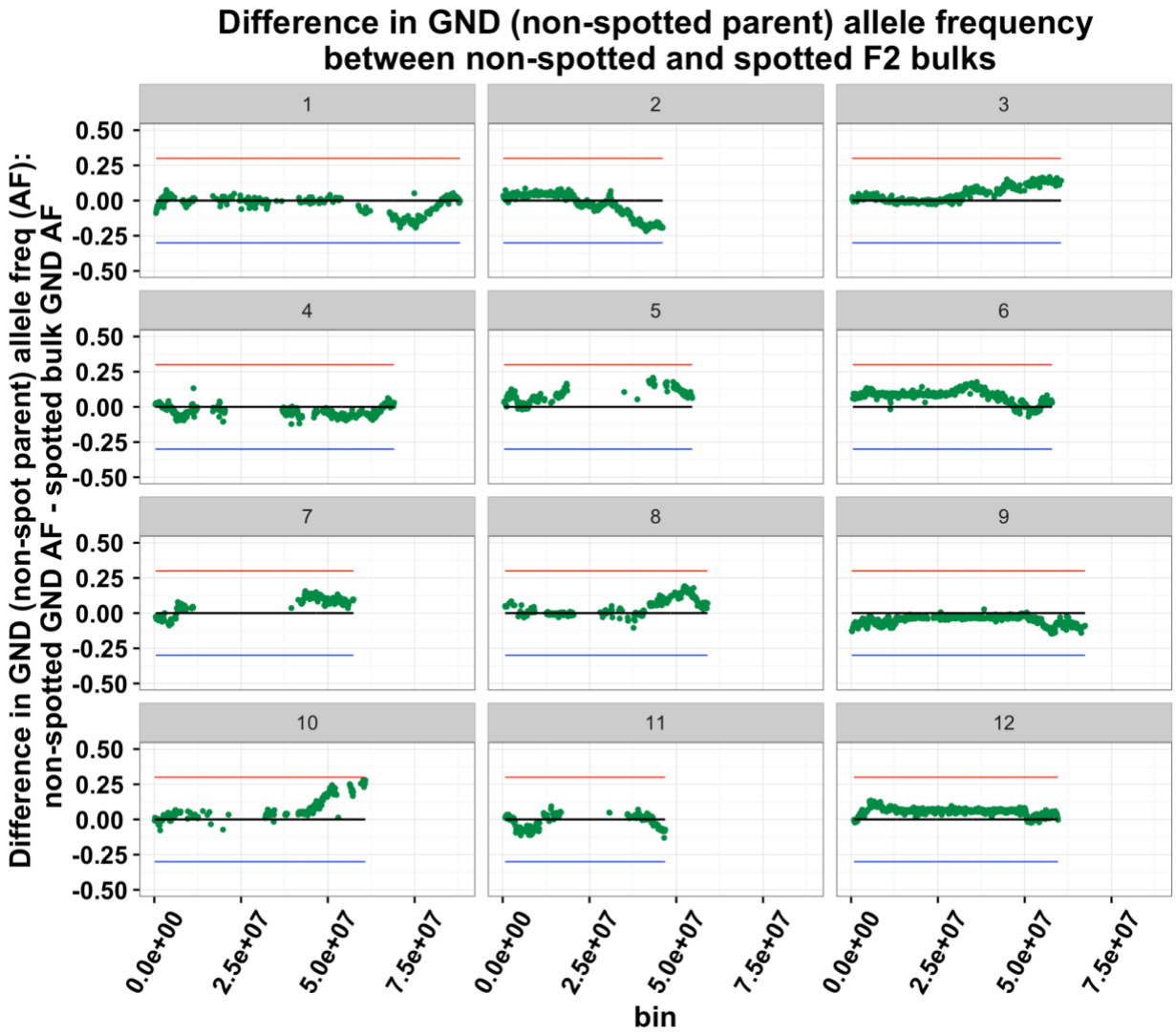

Figure 22. GND allele frequency difference between the two F2 sub-populations.

The GND allele frequency in the spotted bulk minus the GND allele frequency in the non-spotted bulk. The black line marks points at which allele frequency difference between the two bulks is zero. The blue and the red lines are points when allele frequency difference is 0.3 in either diretion. The highest allele frequency difference is on chromosome 10 where the maximum allele frequency difference is 0.3. The allele frequency of the non-spotted parent (GND) is elevated in the non-spotted bulk and reduced in the spotted bulk.

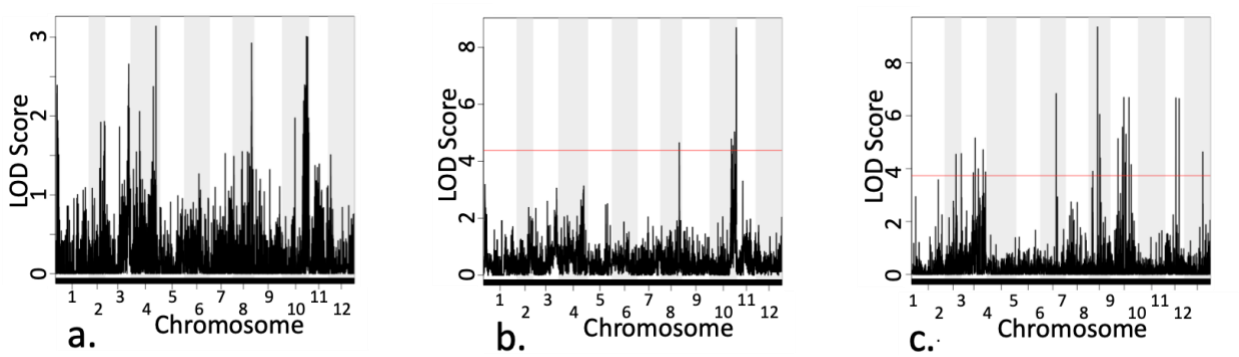

Figure 23: Detection of loci associated with seed spotting.

The three plots show the results of QTL mapping using the r/qtl package in R software using three algorithms i.e (a) standard interval mapping using the E.M algorithm, (b) Haley-Knot algorithm and (c) multiple imputation method. SNPs that were heterozygous in GND but homozygous in IvP 35 were used to construct the bins that were used as markers. There were 3646 markers.

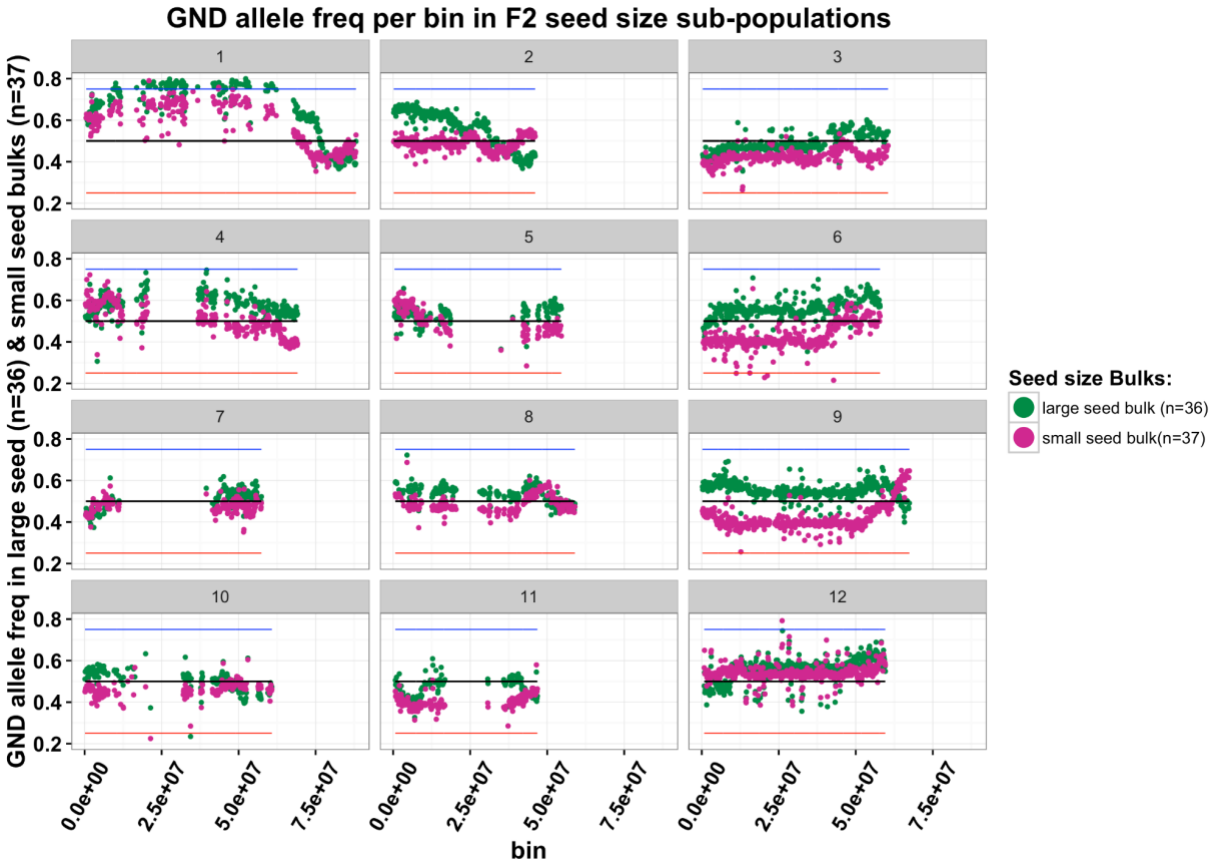

Figure 24. GND allele frequency in the F2 seed size subpopulations.

Each individual in the population is derived from a small seed (n=37) or from a large seed (n=36). It was not determined how the trait would segregate from these individuals. The blue, black and red lines mark a frequency of 0.75, 0.5, and 0.25, respectively.

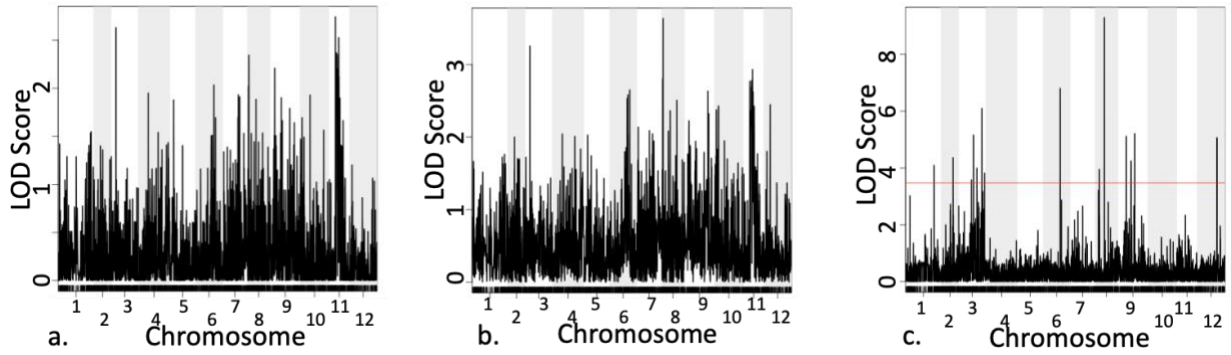

Figure 25. QTL mapping for seed size, using the r/qtl package in R software.

Three algorithms implemented in the package were used. i.e (a) standard interval mapping using the E.M algorithm, (b) Haley-Knot algorithm and (c) multiple imputation method. SNPs that GND heterozygous but homozygous in IvP 35 were used to construct the bins that were used as markers. There were 3646 markers.

References

1. Blakeslee, A. F., Belling, J., Farnham, M. E. & Bergner, A. D. A Haploid Mutant in the Jimson Weed, ``Datura Stramonium''. *Science* **55**, 646–647 (1922).

2. Forster, B. P., Heberle-Bors, E., Kasha, K. J. & Touraev, A. The resurgence of haploids in higher plants. *Trends Plant Sci.* **12**, 368–375 (2007).

3. Seymour, D. K. *et al.* Rapid creation of Arabidopsis doubled haploid lines for quantitative trait locus mapping. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 4227–4232 (2012).

4. Ravi, M. & Chan, S. W. L. Haploid plants produced by centromere-mediated genome elimination. *Nature* **464**, 615–618 (2010).

5. Ravi, M. *et al.* A haploid genetics toolbox for Arabidopsis thaliana. *Nat. Commun.* 5, 5334 (2014).

6. Qiu, F. *et al.* Morphological, cellular and molecular evidences of chromosome random elimination in vivo upon haploid induction in maize. *Current Plant Biology* 1, 83–90 (2014).

7. Liu, L. *et al.* In vivo haploid induction leads to increased frequency of twin-embryo and abnormal fertilization in maize. *BMC Plant Biol.* **18**, 313 (2018).

8. Magoon, M. L. & Khanna, K. R. Haploids. *Caryologia* **16**, 191–235 (1963).

9. Lakshmanan, K. K. & Ambegaokar, K. B. Polyembryony. *Embryology of Angiosperms* 445–474 (1984) doi:10.1007/978-3-642-69302-1_9.

10. Webber, J. M. Polyembryony. *Bot. Rev.* 6, 575–598 (1940).

11. Khanday, I., Skinner, D., Yang, B., Mercier, R. & Sundaresan, V. A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. *Nature* **565**, 91–95 (2019).

12. Ozias-Akins, P., Akiyama, Y. & Hanna, W. W. Molecular characterization of the genomic region linked with apomixis in Pennisetum/Cenchrus. *Funct. Integr. Genomics* 3, 94–104 (2003).

13. Sprague, G. F. The Nature and Extent of Hetero-Fertilization in Maize. *Genetics* **17**, 358–368 (1932).

14. Gao, S. *et al.* Revisiting the hetero-fertilization phenomenon in maize. *PLoS One* 6, e16101 (2011).

15. Randolph, L. F. Note on haploid frequencies. *Maize Genet Coop Newsl* **14**, 23–24 (1940).

16. Sarkar, K. R. & Coe, E. H. A genetic analysis of the origin of maternal haploids in maize. *Genetics* **54**, 453–464 (1966).

17. Chase, S. S. Monoploids and monoploid-derivatives of maize (Zea mays L.). *Bot. Rev.* **35**, 117–168 (1969).

18. Lashermes, P. & Beckert, M. Genetic control of maternal haploidy in maize (Zea mays L.) and selection of haploid inducing lines. *Theor. Appl. Genet.* **76**, 405–410 (1988).

19. Coe, E. H. A Line of Maize with High Haploid Frequency. *Am. Nat.* **93**, 381–382 (1959).

20. Zhang, Z. *et al.* Chromosome elimination and in vivo haploid production induced by Stock 6-derived inducer line in maize (Zea mays L.). *Plant Cell Rep.* **27**, 1851–1860 (2008).

21. Coe, E. H. High-haploid line. in *Maize Genetics Cooperation Newsletter (MNL)* 138–139 (1957).

22. Chase, S. S. & Others. Monoploids in maize. *Heterosis* 389–399 (1952).

23. Rotarenco, V., Dicu, G. & Fuia, S. New inducers of maternal haploids in maize. *Maize Genet. Coop. News Lett.* (2010).

24. Hu, H. *et al.* The Genetic Basis of Haploid Induction in Maize Identified with a Novel Genome-Wide Association Method. *Genetics* **202**, 1267–1276 (2016).

25. Tian, X. *et al.* Hetero-fertilization together with failed egg-sperm cell fusion supports single fertilization involved in in vivo haploid induction in maize. *J. Exp. Bot.* **69**, 4689–4701 (2018).

26. Chang, M. T. Preferential fertilization induced from Stock 6. *Maize Genet Coop Newslett* **66**, 99–100 (1992).

27. Rotarenco, V. & Eder, J. Possible effects of heterofertilization on the induction of maternal haploids in maize. *Maize Genetics Cooperation Newsletter (MNL)* **77**, 30 (2003).

28. Xu, X. *et al.* Gametophytic and zygotic selection leads to segregation distortion through in vivo induction of a maternal haploid in maize. *J. Exp. Bot.* **64**, 1083–1096 (2013).

29. Dong, X. *et al.* Fine mapping of qhir1 influencing in vivo haploid induction in maize. *Theor. Appl. Genet.* **126**, 1713–1720 (2013).

30. Kelliher, T. *et al.* MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. *Nature* **542**, 105–109 (2017).

31. Barret, P., Brinkmann, M. & Beckert, M. A major locus expressed in the male gametophyte with incomplete penetrance is responsible for in situ gynogenesis in maize. *Theor. Appl. Genet.* **117**, 581–594 (2008).

32. Li, X. *et al.* Single nucleus sequencing reveals spermatid chromosome fragmentation as a possible cause of maize haploid induction. *Nat. Commun.* 8, 991 (2017).

33. Swapna, M. & Sarkar, K. R. Anomalous fertilization in haploidy inducer lines in maize (Zea mays L). *Maydica* **56**, (2012).

34. Morgan, D. T., Jr & Rappleye, R. D. Polyembryony in maize and lily; following X-irradiation of the pollen. *J. Hered.* **42**, 90–93 (1951).

35. Li, L., Dong, X., Xu, X. W., Liu, L. W. & Liu, C. X. Observation of twin seedling in maternal haploid induction in maize. *Journal of China* (2012).

36. Bylich, V. G. & Chalyk, S. T. Existence of pollen grains with a pair of morphologically different sperm nuclei as a possible cause of the haploid-inducing capacity in ZMS line. *Maize Genetics Cooperation Newsletter* **70**, 33–33 (1996).

37. Sun, G. *et al.* Matrilineal empowers wheat pollen with haploid induction potency via triggering post-mitosis ROS activity. *New Phytol.* (2022) doi:10.1111/nph.17963.

38. Chalyk, S., Baumann, A., Daniel, G. & Eder, J. Aneuploidy as a possible cause of haploid-

induction in maize. *Maize Genetics Cooperation Newsletter* **77**, 29–29 (2003).

39. Sun, G., Geng, S., Zhang, H., Jia, M. & Wang, Z. Matrilineal empowers wheat pollen with haploid induction potency via triggering post-mitosis ROS activity. *New* (2022).

40. Jiang, C. *et al.* A reactive oxygen species burst causes haploid induction in maize. *Molecular Plant* (2022) doi:10.1016/j.molp.2022.04.001.

41. Roldán-Arjona, T. & Ariza, R. R. Repair and tolerance of oxidative DNA damage in plants. *Mutat. Res.* **681**, 169–179 (2009).

42. Srinivas, U. S., Tan, B. W. Q., Vellayappan, B. A. & Jeyasekharan, A. D. ROS and the DNA damage response in cancer. *Redox Biol* **25**, 101084 (2019).

43. Rowe, L. A., Degtyareva, N. & Doetsch, P. W. DNA damage-induced reactive oxygen species (ROS) stress response in Saccharomyces cerevisiae. *Free Radic. Biol. Med.* **45**, 1167–1177 (2008).

44. Choudhary, A., Kumar, A. & Kaur, N. ROS and oxidative burst: Roots in plant development. *Plant Divers* **42**, 33–43 (2020).

45. Nisa, M.-U., Huang, Y., Benhamed, M. & Raynaud, C. The Plant DNA Damage Response: Signaling Pathways Leading to Growth Inhibition and Putative Role in Response to Stress Conditions. *Front. Plant Sci.* **10**, 653 (2019).

46. Mittler, R. ROS Are Good. *Trends Plant Sci.* **22**, 11–19 (2017).

47. Sanei, M., Pickering, R., Kumke, K., Nasuda, S. & Houben, A. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. *Proc. Natl. Acad. Sci. U. S. A.* **108**, E498–505 (2011).

48. Comai, L. & Tan, E. H. Haploid Induction and Genome Instability. *Trends Genet.* **35**, 791–803 (2019).

49. Amundson, K. R. *et al.* Rare instances of haploid inducer DNA in potato dihaploids and ploidy-dependent genome instability. *Plant Cell* (2021) doi:10.1093/plcell/koab100.

50. Li, L., Xu, X., Jin, W. & Chen, S. Morphological and molecular evidences for DNA introgression in haploid induction via a high oil inducer CAUHOI in maize. *Planta* **230**, 367–376 (2009).

51. Zhao, X., Xu, X., Xie, H., Chen, S. & Jin, W. Fertilization and uniparental chromosome elimination during crosses with maize haploid inducers. *Plant Physiol.* **163**, 721–731 (2013).

52. Riera-Lizarazu, O., Rines, H. W. & Phillips, R. L. Cytological and molecular characterization of oat x maize partial hybrids. *Theor. Appl. Genet.* **93**, 123–135 (1996).

53. Wedzony, M., Röber, F. K. & Geiger, H. H. Chromosome elimination observed in selfed progenies of maize inducer line RWS. in *XVIIth International Congress on Sex Plant Reports Maria Curie--Sklodowska University Press, Lublin* vol. 173 (2002).

54. Röber, F. K., Gordillo, G. A. & Geiger, H. H. In vivo haploid induction in maize. Performance of new inducers and significance of doubled haploid lines in hybrid breeding [Zea mays L.]. *Maydica* (2005).

55. Fischer, E. *Molekulargenetische Untersuchungen zum Vorkommen paternaler DNA-Übertragung bei der in-vivo-Haploideninduktion bei Mais (Zea mays L.)*. (Grauer, 2004).

56. Kelliher, T. *et al.* One-step genome editing of elite crop germplasm during haploid induction. *Nat. Biotechnol.* **37**, 287–292 (2019).

57. Rotarenco, V. A., Dicu, G., Sarmaniuc, M. & Others. Induction of maternal haploids in maize. *Maize Genetics Cooperation Newsletter* **83**, 21 (2009).

58. Liu, C. *et al.* New insight into the mechanism of heterofertilization during maize haploid induction. *Euphytica* **213**, 174 (2017).

59. Kasahara, R. D. *et al.* Fertilization recovery after defective sperm cell release in Arabidopsis. *Curr. Biol.* **22**, 1084–1089 (2012).

60. Kasahara, R. D., Maruyama, D. & Higashiyama, T. Fertilization recovery system is dependent on the number of pollen grains for efficient reproduction in plants. *Plant Signal. Behav.* 8, e23690 (2013).

61. Higashiyama, T. & Takeuchi, H. The mechanism and key molecules involved in pollen tube guidance. *Annu. Rev. Plant Biol.* **66**, 393–413 (2015).

62. Takahashi, T. *et al.* The male gamete membrane protein DMP9/DAU2 is required for double fertilization in flowering plants. *Development* **145**, (2018).

63. Kato, A. Induced single fertilization in maize. *Sex. Plant Reprod.* **10**, 96–100 (1997).

64. Enaleeva, N. K., Tyrnov, V. S., Selivanova, L. P. & Zavalishina, A. N. Single fertilization and the problem of haploidy induction in maize. in *Doklady Biological Sciences* vol. 353 225 (New York: Consultants Bureau, c1965-1992., 1997).

65. Jorgensen, G. A. The experimental formation of heteroploid plants in the genussolanum. *J. Genet.* **19**, 133 (1928).

66. Pace, L. Apogamy in Atamosco. *Bot. Gaz.* **56**, 376–394 (1913).

67. Beale, K. M., Leydon, A. R. & Johnson, M. A. Gamete fusion is required to block multiple pollen tubes from entering an Arabidopsis ovule. *Curr. Biol.* **22**, 1090–1094 (2012).

68. Prigge, V. *et al.* New insights into the genetics of in vivo induction of maternal haploids, the backbone of doubled haploid technology in maize. *Genetics* **190**, 781–793 (2012).

69. Gilles, L. M. *et al.* Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. *EMBO J.* **36**, 707–717 (2017).

70. Liu, C. *et al.* Fine mapping of qhir8 affecting in vivo haploid induction in maize. *Theor. Appl. Genet.* **128**, 2507–2515 (2015).

71. Zhong, Y. *et al.* Mutation of ZmDMP enhances haploid induction in maize. *Nat Plants* 5, 575–580 (2019).

72. Deimling, S., Röber, F. K. & Geiger, H. H. Methodology and genetics of in vivo haploid induction in maize. *Vortr. Pflanzenzüchtg* **38**, 203–224 (1997).

73. Nair, S. K. *et al.* Dissection of a major QTL qhir1 conferring maternal haploid induction ability in maize. *Theor. Appl. Genet.* **130**, 1113–1122 (2017).

74. Liu, C. *et al.* A 4-bp Insertion at ZmPLA1 Encoding a Putative Phospholipase A Generates

Haploid Induction in Maize. *Mol. Plant* **10**, 520–522 (2017).

75. Gilles, L. M. *et al.* Lipid anchoring and electrostatic interactions target the phospholipase NOT-LIKE-DAD to pollen endo-plasma membrane. *Cold Spring Harbor Laboratory* 2020.10.05.326157 (2020) doi:10.1101/2020.10.05.326157.

76. Yao, L. *et al.* OsMATL mutation induces haploid seed formation in indica rice. *Nat Plants* 4, 530–533 (2018).

77. Liu, H. *et al.* Efficient induction of haploid plants in wheat by editing of TaMTL using an optimized Agrobacterium-mediated CRISPR system. *J. Exp. Bot.* **71**, 1337–1349 (2020).

78. Hafidh, S., Fíla, J. & Honys, D. Male gametophyte development and function in angiosperms: a general concept. *Plant Reprod.* **29**, 31–51 (2016).

79. Lalanne, E. & Twell, D. Genetic control of male germ unit organization in Arabidopsis. *Plant Physiol.* **129**, 865–875 (2002).

80. Taimur, N. A study of molecular and genetic mechanisms mediating the formation of twin sperm cells in Arabidopsis thaliana. (2016).

81. Li, Y. *et al.* Loss-of-function alleles of ZmPLD3 cause haploid induction in maize. *Nat Plants* 7, 1579–1588 (2021).

82. Kasaras, A. & Kunze, R. Expression, localisation and phylogeny of a novel family of plant-specific membrane proteins. *Plant Biol.* **12 Suppl 1**, 140–152 (2010).

83. Cyprys, P., Lindemeier, M. & Sprunck, S. Gamete fusion is facilitated by two sperm cell-expressed DUF679 membrane proteins. *Nat Plants* 5, 253–257 (2019).

84. Zhong, Y. *et al.* A DMP-triggered in vivo maternal haploid induction system in the dicotyledonous Arabidopsis. *Nat Plants* 6, 466–472 (2020).

85. Berger, F. & Twell, D. Germline specification and function in plants. *Annu. Rev. Plant Biol.* **62**, 461–484 (2011).

86. Frank, A. C. & Johnson, M. A. Expressing the diphtheria toxin A subunit from the

HAP2(GCS1) promoter blocks sperm maturation and produces single sperm-like cells capable of fertilization. *Plant Physiol.* **151**, 1390–1400 (2009).

87. Chumakov, M. I. Matroclinic Haploidy and Gamete Interaction in Maize. *Russ. J. Genet.* **54**, 1137–1141 (2018).

88. Sprunck, S. *et al.* Egg cell-secreted EC1 triggers sperm cell activation during double fertilization. *Science* **338**, 1093–1097 (2012).

89. Mori, T., Igawa, T., Tamiya, G., Miyagishima, S.-Y. & Berger, F. Gamete attachment requires GEX2 for successful fertilization in Arabidopsis. *Curr. Biol.* **24**, 170–175 (2014).

90. Dresselhaus, T. & Sprunck, S. Plant fertilization: maximizing reproductive success. *Current biology: CB* vol. 22 R487–9 (2012).

91. von Besser, K., Frank, A. C., Johnson, M. A. & Preuss, D. Arabidopsis HAP2 (GCS1) is a sperm-specific gene required for pollen tube guidance and fertilization. *Development* **133**, 4761–4769 (2006).

92. Hafidh, S. & Honys, D. Reproduction Multitasking: The Male Gametophyte. *Annu. Rev. Plant Biol.* **72**, 581–614 (2021).

93. Englhart, M. *Direct Gamete Interactions and EGG CELL1-mediated Sperm Cell Activation in Arabidopsis Thaliana*. (Universitätsbibliothek Regensburg, 2018).

94. Huang, J., Ju, Y., Wang, X., Zhang, Q. & Sodmergen. A one-step rectification of sperm cell targeting ensures the success of double fertilization. *J. Integr. Plant Biol.* **57**, 496–503 (2015).

95. Gusti, A. *et al.* The Arabidopsis thaliana F-box protein FBL17 is essential for progression through the second mitosis during pollen development. *PLoS One* 4, e4780 (2009).

96. Borg, M. *et al.* The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in Arabidopsis. *Plant Cell* **23**, 534–549 (2011).

97. Higashiyama, T. *et al.* Pollen tube attraction by the synergid cell. *Science* **293**, 1480–1483

(2001).

98. Boerman, N. A., Frei, U. K. & Lübberstedt, T. Impact of Spontaneous Haploid Genome Doubling in Maize Breeding. *Plants* vol. 9 369 (2020).

99. Chaikam, V., Gowda, M., Nair, S. K., Melchinger, A. E. & Boddupalli, P. M. Genome-wide association study to identify genomic regions influencing spontaneous fertility in maize haploids. *Euphytica* **215**, 138 (2019).

100. Zhong, Y. *et al.* A genotype independent DMP-HI system in dicot crops. *bioRxiv* 2021.06.21.449224 (2021) doi:10.1101/2021.06.21.449224.

101. Wang, N., Gent, J. I. & Dawe, R. K. Haploid induction by a maize cenh3 null mutant. *Sci Adv* 7, (2021).

102. Marimuthu, M. P. A. *et al.* Biased removal and loading of centromeric histone H3 during reproduction underlies uniparental genome elimination. *bioRxiv* 2021.02.24.432754 (2021) doi:10.1101/2021.02.24.432754.

103. Battaglia, E. Unusual Cytological Features in the Apomictic Rudbeckia Sullivantii Boynton et Beadle. *Caryologia* 8, 1–32 (1955).

104. Chaudhari, H. K. Use of Semigamy in the Production of Cotton Haploids. *Bull. Torrey Bot. Club* **105**, 98–103 (1978).

105. Turcotte, E. L. & Feaster, C. V. Semigamy in Pima Cotton. *J. Hered.* **58**, 55–57 (1967).

106. Lanaud, C. Origin of haploids and semigamy in Theobroma cacao L. *Euphytica* **38**, 221–228 (1988).

107. Montelongo-Escobedo, H. & Rowe, P. R. Haploid induction in potato: cytological basis for the pollinator effect. *Euphytica* **18**, 116–123 (1969).

108. Montezuma-de-Carvalho, J. The effect of N2O on pollen tube mitosis in styles and its potential significance for inducing haploidy in potato. *Euphytica* **16**, 190–198 (1967).

109. Wangenheim, K.-H. v., Peloquin, S. J. & Hougas, R. W. Embryological investigations on the

formation of haploids in the potato (Solanum tuberosum). *Zeitschrift für Vererbungslehre* **91**, 391–399 (1960).

110. Eder, J. & Chalyk, S. In vivo haploid induction in maize. *Theor. Appl. Genet.* **104**, 703–708 (2002).

111. Sprunck, S., Hackenberg, T., Englhart, M. & Vogler, F. Same same but different: sperm-activating EC1 and ECA1 gametogenesis-related family proteins. *Biochem. Soc. Trans.* **42**, 401–407 (2014).

112. Huamán, Z., Golmirzaie, A. & Amoros, W. The potato. *Biodiversity in trust: conservation and use of plant genetic resources in CGIAR Centres* 21–28 (1997).

113. Bradshaw, J. E. & Ramsay, G. Potato Origin and Production. *Advances in Potato Chemistry and Technology* 1–26 (2009) doi:10.1016/b978-0-12-374349-7.00001-5.

114. Den Nijs, T. P. M. & Peloquin, S. J. Polyploid evolution via 2n gametes. *Am. Potato J.* **54**, 377–386 (1977).

115. den Nijs, T. P. M. & Peloquin, S. J. 2n gametes in potato species and their function in sexual polyploidization. *Euphytica* **26**, 585–600 (1977).

116. Plaisted, R. L. Advances and limitations in the utilization of Neotuberosum in potato breeding. *The production of new potato varieties: technological advances* 186–196 doi:10.1017/cbo9780511895463.041.

117. Glendinning, D. R. Neo-Tuberosum: new potato breeding material. 1. The origin, composition, and development of the Tuberosum and Neo-Tuberosum gene pools. *Potato Research* vol. 18 256–261 (1975).

118. Bethke, P. C. *et al.* History and origin of russet Burbank (netted gem) a sport of Burbank. *Am. J. Potato Res.* **91**, 594–609 (2014).

119. Ames, M. & Spooner, D. M. DNA from herbarium specimens settles a controversy about origins of the European potato. *Am. J. Bot.* **95**, 252–257 (2008).

120. Spooner, D. M., Ghislain, M., Simon, R., Jansky, S. H. & Gavrilenko, T. Systematics, Diversity, Genetics, and Evolution of Wild and Cultivated Potatoes. *Bot. Rev.* **80**, 283–383 (2014).

121. Hanneman, R. E. The potato germplasm resource. *Am. Potato J.* **66**, 655–667 (1989).

122. Ovchinnikova, A. *et al.* Taxonomy of cultivated potatoes (Solanum section Petota: Solanaceae). *Bot. J. Linn. Soc.* **165**, 107–155 (2011).

123. Murovec, J. & Bohanec, B. Haploids and Doubled Haploids in Plant Breeding. *Plant Breeding* (2012) doi:10.5772/29982.

124. Sanetomo, R. & Hosaka, K. Re-evaluation of Monohaploid Solanum verrucosum and S. bulbocastanum (2n = x = 12) and Dihaploid S. stoloniferum and S. acaule (2n = 2x = 24), All Derived from Anther Culture. *American Journal of Potato Research* vol. 98 333–343 (2021).

125. Bothmer, R. von, von Bothmer, R. & Subrahmanyam, N. C. Assessment of chromosome associations in haploids and their parental accessions in *Hordeum*. *Genome* vol. 30 204–210 (1988).

126. Ivanovskaja, E. V. A haploid plant of Solatium tuberosum LCR. *Doklady Acad. Sei. USSR* **24**, 517–520 (1939).

127. Hougas, R. W. & Peloquin, S. J. A haploid plant of the potato variety katahdin. *Nature* **180**, 1209–1210 (1957).

128. Bains, G. S. & Howard, H. W. Haploid plants of Solanum demissum. *Nature* **166**, 795 (1950).

129. Dodds, K. S. Polyhaploids of Solanum demissum. *Nature* **166**, 795 (1950).

130. Hougas, R. W., Peloquin, S. J. & Ross, R. W. HAPLOIDS OF THE COMMON POTATO. *J. Hered.* **49**, 103–106 (1958).

131. Hougas, R. W., Peloquin, S. J. & Ross, R. W. The potential of haploid potatoes for research and breeding. *Am. J. Potato Res.* **35**, 442 (1958).

132. Gabert, A. C. *Factors influencing the frequency of haploids in the common potato (Solanum tuberosum L.)*. (University of Wisconsin--Madison, 1963).

133. Kawakami, K. & Mastubayashi, M. Studies on the Haploid Plants of Solanum tuberosum : I. Morphological Characterisitics of the Polyhaploid Plants. *Japanese Journal of Breeding* vol. 10 10–18 (1960).

134. Matsubayashi, M. Studies on the Haploid Plants of Solanum tuberosum: II. Meiotic chromosome pairing in the polyhaploid plants. *Japanese Journal of Breeding* **10**, 195–202 (1960).

135. Hanneman, R. E. & Peloquin, S. J. Ploidy levels of progeny from diploid-tetraploid crosses in the potato. *Am. Potato J.* **45**, 255–261 (1968).

136. Hermsen, J. G. T. Induction of haploids and aneuhaploids in colchicine-induced tetraploid Solanum chacoense Bitt. *Euphytica* **18**, 183–189 (1969).

137. Huamán, Z. & Ross, R. W. Updated listing of potato species names, abbreviations and taxonomic status. *American Potato Journal* vol. 62 629–641 (1985).

138. USDA United States Department of Agriculture. Solanum tuberosum L. subsp. andigenum (Juz. & Bukasov) Hawkes GRIN-Global. https://web.archive.org/web/20210405112021/https://npgsweb.ars-grin.gov/gringlobal/taxon/taxonomydetail?id=101304.

139. Howard, H. W. & Swaminathan, M. S. The cytology of haploid plants ofSolanum demissum. *Genetica* **26**, 381–391 (1953).

140. Hougas, R. W., Peloquin, S. J. & Gabert, A. C. Effect of seed-parent and pollinator on frequency of haploids in Solanum tuberosum 1. *Crop Sci.* 4, 593–595 (1964).

141. Bender, K. The production and origin of dihaploid plants in Solanum tuberosum. *Zeitschrift fuer Pflanzenzuechtung (West Germany)* **: 50**, (1963).

142. Lamm, R. Notes on a haploid potato hybrid. *Hereditas* **24**, 391–396 (2010).

143. Caligari, P. D. S., Powell, W., Liddell, K., Maine, M. J. D. E. & Swan, G. E. L. Methods and strategies for detecting Solanum tuberosum dihaploids in interspecific crosses with S. phureja. *Ann. Appl. Biol.* **112**, 323–328 (1988).

144. Hutten, R. C. B., E J M, Huigen, D. J., Th. Hermsen, J. G. & Jacobsen, E. Analysis of dihaploid induction and production ability and seed parent x pollinator interaction in potato. *Euphytica* vol. 72 61–64 (1994).

145. Peloquin, S. J. & Hougas, R. W. Decapitation and genetic markers as related to haploidy inSolanum tuberosum. *European Potato Journal* vol. 2 176–183 (1959).

146. Hermsen, J. G. T. & Verdenius, J. Selection from Solanum tuberosum group phureja of genotypes combining high-frequency haploid induction with homozygosity for embryo-spot. *Euphytica* **22**, 244–259 (1973).

147. van Breukelen, E., Ramanna, M. S. & Hermsen, J. G. T. Parthenogenetic monohaploids (2n= x= 12) from Solanum tuberosum L. and S. verrucosum Schlechtd. and the production of homozygous potato diploids. *Euphytica* (1977).

148. Uijtewaal, B. A., Huigen, D. J. & Th. Hermsen, J. G. Production of potato monohapioids (2n = x = 12) through prickle pollination. *Theor. Appl. Genet.* **73**, 751–758 (1987).

149. Lough, R. C., Varrieur, J. M. & Veilleux, R. E. Selection inherent in monoploid derivation mechanisms for potato. *Züchter Genet. Breed. Res.* **103**, 178–184 (2001).

150. Van Breukelen, E. W. M., Ramanna, M. S. & Hermsen, J. G. T. Monohaploids (n=x=12) from autotetraploid Solanum tuberosum (2n=4x=48) through two successive cycles of female parthenogenesis. *Euphytica* (1975).

151. Van Breukelen, E. W. M. Pseudogamic production of dihaploids and monoploids in Solanum tuberosum and some related species. (1981).

152. Singsit, C. & Hanneman, R. E. Haploids of tetraploid (2n=4x=48) Mexican potato species—Their extraction, cytology and crossability. *Am. Potato J.* **64**, 469–482 (1987).

153. Clulow, S. A. *et al.* Cytological and molecular observations on Solanum phureja-induced

dihaploid potatoes. *Theor. Appl. Genet.* **82**, 545–551 (1991).

154. Clulow, S. A., Wilkinson, M. J. & Burch, L. R. Solanum phureja genes are expressed in the leaves and tubers of aneusomatic potato dihaploids. *Euphytica* **69**, 1–6 (1993).

155. Clulow, S. A. & Rousselle-Bourgeois, F. Widespread introgression of Solatium phureja DNA in potato (S. tuberosum) dihaploids. *Plant Breed.* **116**, 347–351 (1997).

156. PI 225682 GRIN-Global. https://web.archive.org/web/20210607182642/https://npgsweb.ars-grin.gov/gringlobal/accessiondetail?id=1181411.

157. PI 225702 GRIN-Global. https://web.archive.org/web/20210607185000/https://npgsweb.ars-grin.gov/gringlobal/accessiondetail?id=1181427.

158. Cooper, D. C. & Rieman, G. H. Diploid plants in a seedling population of the cultivated potato. *Am. Potato J.* **35**, 642–642 (1958).

159. Peloquin, S. J., Gabert, A. C. & Ortiz, R. Nature of ` Pollinator ' Effect in Potato (Solanum tuberosum L.) Haploid Production. *Ann. Bot.* **77**, 539–542 (1996).

160. Wasilewicz-Flis, I. & Jakuczun, H. Induction and application of dihaploids of potato (Solanum tuberosum L.). *Plant Breeding and Seed Science* **52**, 47–64 (2005).

161. Carroll, C. P. & Low, R. J. Aspects of male fertility in group Tuberosum dihaploids. *Potato Res.* **19**, 109–121 (1976).

162. Carroll, C. P. & Low, R. J. Flowering behaviour and seed fertility in dihaploidSolanum tuberosum. *Potato Res.* **18**, 416–427 (1975).

163. Buketova, M. I. Choice of pollinators for producing dihaploids. *Kartofel'i Ovoshchi* (1970).

164. Moyle, L. C. & Nakazato, T. Hybrid incompatibility 'snowballs' between Solanum species. *Science* **329**, 1521–1523 (2010).

165. Rice, W. R. & Holland, B. The enemies within: intergenomic conflict, interlocus contest

evolution (ICE), and the intraspecific Red Queen. *Behav. Ecol. Sociobiol.* **41**, 1–10 (1997).

166. Clark, N. L., Aagaard, J. E. & Swanson, W. J. Evolution of reproductive proteins from animals and plants. *Reproduction* **131**, 11–22 (2006).

167. Willi, Y. The battle of the sexes over seed size: support for both kinship genomic imprinting and interlocus contest evolution. *Am. Nat.* **181**, 787–798 (2013).

168. Ivanov, M. A. Experimental production of haploids in nicotiana rustical L. *Genetica* **20**, 295–397 (1938).

169. Lindstrom, E. W. A HAPLOID MUTANT IN THE TOMATO*. *Journal of Heredity* vol. 20 23–30 (1929).

170. Krishnaswamy, N. Cytological studies in a haploid plant of Triticum vulgare. *Hereditas* **25**, 77–86 (1939).

171. Marks, G. E. A Polyhaploid Plant of Solanum polytrichon Rydb. *Nature* **175**, 469–469 (1955).

172. Abdalla, M. M. F. & Th. Hermsen, J. G. Diploid parthenogenesis and androgenesis in diploid Solanum. *Euphytica* vol. 21 426–431 (1972).

173. Laurie, D. A. & Bennett, M. D. Wheat × maize hybridization. *Canadian Journal of Genetics and Cytology* vol. 28 313–316 (1986).

174. Camadro, E. L., Masuelli, R. W. & Cortés, M. C. Haploids of the wild tetraploid potato Solanum acaule ssp. acaule: generation, meiotic behavior, and electrophoretic pattern for the aspartate aminotransferase system. *Genome* vol. 35 431–435 (1992).

175. Inagaki, M. & Mujeeb-Kazi, A. Comparison of Polyhaploid Production Frequencies in Crosses of Hexaploid Wheat with Maize, Pearl Millet and Sorghuml. *Japanese Journal of Breeding* vol. 45 157–161 (1995).

176. Chaudhary, H. K., Sethi, G. S., Singh, S., Pratap, A. & Sharma, S. Efficient haploid induction in wheat by using pollen of Imperata cylindrica. *Plant Breed.* **124**, 96–98 (2005).

177. Ren, J. *et al.* Novel technologies in doubled haploid line development. *Plant Biotechnology*

Journal vol. 15 1361–1370 (2017).

178. Camadro, E. L., Erazzú, L. E., Maune, J. F. & Bedogni, M. C. A genetic approach to the species problem in wild potato. *Plant Biol.* **14**, 543–554 (2012).

179. Huaman, Z. [Systematic botany, identification, distribution and evaluation of cultivated potato [Solanum tuberosum]]. *Documento de Capacitacion (CIP)* (1984).

180. Huamán, Z. & Spooner, D. M. Reclassification of landrace populations of cultivated potatoes (Solanum sect. Petota). *Am. J. Bot.* **89**, 947–965 (2002).

181. Cersosimo, A. & Sorrentino, C. Haploidy in Nicotiana spp. in *In Vitro Haploid Production in Higher Plants: Volume 5 — Oil, Ornamental and Miscellaneous Plants* (eds. Jain, S. M., Sopory, S. K. & Veilleux, R. E.) 135–163 (Springer Netherlands, 1997).

182. Burk, L. G., Gerstel, D. U. & Wernsman, E. A. Maternal Haploids of Nicotiana tabacum L. from Seed. *Science* **206**, 585 (1979).

183. Satina, S., Rappaport, J. & Blakeslee, A. F. Ovular Tumors Connected with Incompatible Crosses in Datura. *Am. J. Bot.* **37**, 576–586 (1950).

184. Brisson, D. Negative Frequency-Dependent Selection Is Frequently Confounding. *Front Ecol Evol* 6, (2018).

185. Barber, H. N. The suppression of meiosis and the origin of diplochromosomes. *Proceedings of the Royal Society of London. Series B - Biological Sciences* **128**, 170–185 (1940).

186. Gastony, G. J. & Windham, M. D. Species Concepts in Pteridophytes: The Treatment and Definition of Agamosporous Species. *American Fern Journal* vol. 79 65 (1989).

187. Shimamura, T. Cytological Studies of Polyploidy Induced by Colchicine. *CYTOLOGIA* vol. 9 486–494 (1939).

188. Briggs, R. An analysis of the inactivation of the frog sperm nucleus by toluidine blue. *J. Gen. Physiol.* **35**, 761–780 (1952).

189. Chimoyo, H. M. The effect of ultraviolet irradiation, toluidine blue, and environment on maternal haploid frequencies from the cross between Nicotiana tabacum and N. africana.

(Clemson University, 1988).

190. Edwards, R. G. Colchicine-induced heteroploidy in early mouse embryos. *Nature* **174**, 276–277 (1954).

191. Edwards, R. G. The experimental induction of pseudogamy in early mouse embryos. *Experientia* **10**, 499–500 (1954).

192. Edwards, R. G. The experimental induction of gynogenesis in the mouse. III. Treatment of sperm with trypaflavine, toluidine blue, or nitrogen mustard. *Proc. R. Soc. Lond. B Biol. Sci.* **149**, 117–129 (1958).

193. Illies, Z. M. Induction of haploid parthenogenesis in aspen by post-pollination treatment with Toluidine-blue. *Silvae Genet.* **23**, 221–226 (1974).

194. NsokA, M. & KoNDo, Y. Production of gynogenetic haploids by treatment of spermatozoa with toluidine blue in Rana japonica. *Sci. Rep. Lab. Amphibian Biol. Hiroshima Univ.* 3, 385–398 (1978).

195. Stockert, J. C., Juarranz, A., Villanueva, A. & Cañete, M. Photodynamic damage to HeLa cell microtubules induced by thiazine dyes. *Cancer Chemother. Pharmacol.* **39**, 167–169 (1996).

196. Bacellar, I. O. L. *et al.* Membrane damage efficiency of phenothiazinium photosensitizers. *Photochem. Photobiol.* **90**, 801–813 (2014).

197. Paardekooper, M. *et al.* Photodynamic treatment of yeast cells with the dye toluidine blue: all-or-none loss of plasma membrane barrier properties. *Biochim. Biophys. Acta* **1108**, 86–90 (1992).

198. Vázquez-Ortega, F. *et al.* Photodynamic treatment induced membrane cell damage in Acanthamoeba castellanii Neff. *Dyes and Pigments* vol. 180 108481 (2020).

199. Paardekooper, M., De Bruijne, A. W., Van Steveninck, J. & Van den Broek, P. J. Inhibition of transport systems in yeast by photodynamic treatment with toluidine blue. *Biochim. Biophys. Acta* **1151**, 143–148 (1993).

200. Tuite, E. M. & Kelly, J. M. New trends in photobiology: Photochemical interactions of methylene blue and analogues with DNA and other biological substrates. *J. Photochem. Photobiol. B* **21**, 103–124 (1993).

201. Sridharan, G. & Shankar, A. A. Toluidine blue: A review of its chemistry and clinical utility. *J. Oral Maxillofac. Pathol.* **16**, 251–255 (2012).

202. Kato, A. Induction of bicellular pollen and dihaploidization of tetraploid maize. *MAIZE GENETICS COOPERATION NEWSLETTER* **70**, 25–25 (1996).

203. Waugh, R., Baird, E. & Powell, W. The use of RAPD markers for the detection of gene introgression in potato. *Plant Cell Rep.* **11**, 466–469 (1992).

204. Wilkinson, M. J. *et al.* Evidence for somatic translocation during potato dihaploid induction. *Heredity* **74 (Pt 2)**, 146–151 (1995).

205. Straadt, I. K. & Rasmussen, O. S. AFLP analysis of Solanum phureja DNA introgressed into potato dihaploids. *Plant Breed.* **122**, 352–356 (2003).

206. Ercolano, M. R. *et al.* Assessment of genetic variability of haploids extracted from tetraploid (2n = 4x = 48) Solanum tuberosum. *Genome* **47**, 633–638 (2004).

207. Bartkiewicz, A. M. *et al.* Maximization of Markers Linked in Coupling for Tetraploid Potatoes via Monoparental Haploids. *Front. Plant Sci.* 9, 620 (2018).

208. Pham, G. M. *et al.* Genome-wide inference of somatic translocation events during potato dihaploid production. *Plant Genome* **12**, 180079 (2019).

209. Marks, G. E. The enigma of triploid potatoes. *Euphytica* **15**, 285–290 (1966).

210. Van Suchtelen, N. Triploids of the common potato,Solanum tuberosum L. *Potato Research* vol. 19 377–380 (1976).

211. Graebner, R. C., Chen, H., Contreras, R. N., Haynes, K. G. & Sathuvalli, V. Identification of the High Frequency of Triploid Potato Resulting from Tetraploid × Diploid Crosses. *HortScience* vol. 54 1159–1163 (2019).

212. Franklin-Tong, V. E. *Self-Incompatibility in Flowering Plants: Evolution, Diversity, and*

Mechanisms. (Springer Science & Business Media, 2008).

213. Igic, B., Smith, W. A., Robertson, K. A., Schaal, B. A. & Kohn, J. R. Studies of self-incompatibility in wild tomatoes: I. S-allele diversity in Solanum chilense (Dun.) Reiche [corrected] (Solanaceae). *Heredity* **99**, 553–561 (2007).

214. Dzidzienyo, D. K., Bryan, G. J., Wilde, G. & Robbins, T. P. Allelic diversity of S-RNase alleles in diploid potato species. *Theor. Appl. Genet.* **129**, 1985–2001 (2016).

215. Newbigin, E., Anderson, M. A. & Clarke, A. E. Gametophytic Self-Incompatibility Systems. *The Plant Cell* vol. 5 1315 (1993).

216. Fujii, S., Kubo, K.-I. & Takayama, S. Non-self- and self-recognition models in plant self-incompatibility. *Nat Plants* 2, 16130 (2016).

217. Puul, S. 2n GAMETES IN DIPLOID SOLANUM: FREQUENCY AND TYPES OF SPINDLE ABNORMALITIES1. *Can. J. Genet. Cytol.* **24**, 301–314 (1982).

218. Ramanna, M. S. A re-examination of the mechanisms of 2n gamete formation in potato and its implications for breeding. *Euphytica* **28**, 537–561 (1979).

219. Carputo, D., Barone, A. & Frusciante, L. 2n gametes in the potato: essential ingredients for breeding and germplasm transfer. *Theor. Appl. Genet.* **101**, 805–813 (2000).

220. Mable, B. K. Polyploidy and self-compatibility: is there an association? *New Phytol.* **162**, 803–811 (2004).

221. Mendiburu, A. O. & Peloquin, S. J. Sexual polyploidization and depolyploidization: some terminology and definitions. *Theoretical and Applied Genetics* vol. 48 137–143 (1976).

222. Bretagnolle, F. & Thompson, J. D. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. *New Phytol.* **129**, 1–22 (1995).

223. Iwanaga, M. & Peloquin, S. J. Origin and evolution of cultivated tetraploid potatoes via 2n gametes. *Theor. Appl. Genet.* **61**, 161–169 (1982).

224. Quinn, A. A., Mok, D. W. S. & Peloquin, S. J. Distribution and significance of diplandroids

among the diploid solanums. *Am. Potato J.* **51**, 16–21 (1974).

225. Watanabe, K. & Peloquin, S. J. Cytological basis of 2n pollen formation in a wide range of 2x, 4x, and 6x taxa from tuber-bearing Solanum species. *Genome* **36**, 8–13 (1993).

226. Stelly, D. M. & Peloquin, S. J. FORMATION OF 2N MEGAGAMETOPHYTES IN DIPLOID TUBER-BEARING SOLANUMS. *American Journal of Botany* vol. 73 1351–1363 (1986).

227. Camadro, E. L. & Peloquin, S. J. The occurrence and frequency of 2n pollen in three diploid solanums from Northwest Argentina. *Theoretical and Applied Genetics* vols 56-56 11–15 (1980).

228. Jongedijk, E. Desynapsis and FDR 2N-megaspore formation in diploid potato : potentials and limitations for breeding and for the induction of diplosporic apomixis. (Jongedijk, 1991).

229. Werner, J. E. & Peloquin, S. J. Occurrence and mechanisms of 2n egg formation in 2x potato. *Genome* vol. 34 975–982 (1991).

230. Petit, C., Bretagnolle, F. & Felber, F. Evolutionary consequences of diploid–polyploid hybrid zones in wild species. *Trends Ecol. Evol.* **14**, 306–311 (1999).

231. Hanzl, M., Kolář, F., Nováková, D. & Suda, J. Nonadaptive processes governing early stages of polyploid evolution: Insights from a primary contact zone of relict serpentine Knautia arvensis (Caprifoliaceae). *Am. J. Bot.* **101**, 935–945 (2014).

232. Correll, D. S. & Others. The potato and its wild relatives. *The potato and its wild relatives.* (1962).

233. Navarro, C. *et al.* Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. *Nature* **478**, 119–122 (2011).

234. Abelenda, J. A., Navarro, C. & Prat, S. Flowering and tuberization: a tale of two nightshades. *Trends Plant Sci.* **19**, 115–122 (2014).

235. Dnyansagar, V. R. & Cooper, D. C. Development of the Seed of Solanum phureja. *Am. J. Bot.* **47**, 176–186 (1960).

236. Hanneman, R. E. & Peloquin, S. J. Croosability of 24-chromosome potato hybrids with 48-

chromosome cultivars. *European Potato Journal* **10**, 62–73 (1967).

237. Johnston, S. A. & Hanneman, R. E., Jr. The genetics of triploid formation and its relationship to endosperm balance number in potato. *Genome* **38**, 60–67 (1995).

238. Masuelli, R. W. & Camadro, E. L. Crossability relationships among wild potato species with different ploidies and Endosperm Balance Numbers (EBN). *Euphytica* **94**, 227–235 (1997).

239. Jackson, M. T., Rowe, P. R. & Hawkes, J. G. Crossability relationships of andean potato varieties of three ploidy levels. *Euphytica* **27**, 541–551 (1978).

240. Tarn, T. R. & Hawkes, J. G. Cytogenetic studies and the occurrence of triploidy in the wild potato species Solanum commersonii Dun. *Euphytica* **35**, 293–302 (1986).

241. Graebner, R. C., Chen, H., Contreras, R. N., Haynes, K. G. & Sathuvalli, V. Identification of the High Frequency of Triploid Potato Resulting from Tetraploid× Diploid Crosses. *HortScience* **54**, 1159–1163 (2019).

242. Köhler, C., Mittelsten Scheid, O. & Erilova, A. The impact of the triploid block on the origin and evolution of polyploid plants. *Trends Genet.* **26**, 142–148 (2010).

243. Kolář, F., Čertner, M., Suda, J., Schönswetter, P. & Husband, B. C. Mixed-Ploidy Species: Progress and Opportunities in Polyploid Research. *Trends Plant Sci.* **22**, 1041–1055 (2017).

244. Husband, B. C. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. *Biol. J. Linn. Soc. Lond.* **82**, 537–546 (2004).

245. Alfano, F., Cammareri, M., Errico, A., Frusciante, L. & Conicella, C. 2n gametes inSolanum tuberosum dihaploids. *Am. J. Potato Res.* **76**, 281–285 (1999).

246. Conicella, C., Genualdo, G., Errico, A., Frusciante, L. & Monti, L. M. Meiotic restitution mechanisms and 2n pollen formation in a Solanum tuberosum dihaploid and in dihaploid x wild species hybrids*. *Plant Breeding* vol. 115 157–161 (1996).

247. Sax, K. THE EXPERIMENTAL PRODUCTION OF POLYPLOIDY. *J. Arnold Arbor.* **17**, 153–159 (1936).

248. Comai, L. The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* 6, 836–846 (2005).

249. McHale, N. A. Environmental induction of high frequency 2n pollen formation in diploid Solanum. *Can. J. Genet. Cytol.* **25**, 609–615 (1983).

250. Fawcett, J. A. & Van de Peer, Y. Angiosperm polyploids and their road to evolutionary success. *Trends Evol. Biol.* 2, e3–e3 (2010).

251. Castro, M. O. Evolutionary ecology of polyploids: understanding species coexistence at the contact zones. (00500::Universidade de Coimbra, 2019).

252. Buggs, R. J. A. & Pannell, J. R. Ecological differentiation and diploid superiority across a moving ploidy contact zone. *Evolution* **61**, 125–140 (2007).

253. Levin, D. A. & Soltis, D. E. Factors promoting polyploid persistence and diversification and limiting diploid speciation during the K–Pg interlude. *Current Opinion in Plant Biology* vol. 42 1–7 (2018).

254. Soltis, P. S., Marchant, D. B., Van de Peer, Y. & Soltis, D. E. Polyploidy and genome evolution in plants. *Curr. Opin. Genet. Dev.* **35**, 119–125 (2015).

255. Vajda, V. & Bercovici, A. The global vegetation pattern across the Cretaceous–Paleogene mass extinction interval: A template for other extinction events. *Glob. Planet. Change* **122**, 29–49 (2014).

256. Soltis, D. E. & Gordon Burleigh, J. Surviving the K-T mass extinction: New perspectives of polyploidization in angiosperms. *Proceedings of the National Academy of Sciences* vol. 106 5455–5456 (2009).

257. Stebbins, G. L. Apomixis in the angiosperms. *Bot. Rev.* 7, 507–542 (1941).

258. de Wet, J. M. J. REVERSIBLE TETRAPLOIDY AS AN EVOLUTIONARY MECHANISM. *Evolution* **25**, 545–548 (1971).

259. Husband, B. C., Baldwin, S. J. & Suda, J. The Incidence of Polyploidy in Natural Plant

Populations: Major Patterns and Evolutionary Processes. in *Plant Genome Diversity Volume 2: Physical Structure, Behaviour and Evolution of Plant Genomes* (eds. Greilhuber, J., Dolezel, J. & Wendel, J. F.) 255–276 (Springer Vienna, 2013).

260. Tandon, S. L. & Rao, G. R. Cytogenetical Investigations in Relation to the Mechanism of Evolution in Hexaploid Solanum nigrum L. *Nature* **201**, 1348–1349 (1964).

261. Melo, C. A. F., Martins, M. I. G., Oliveira, M. B. M., Benko-Iseppon, A. M. & Carvalho, R. Karyotype analysis for diploid and polyploid species of the Solanum L. *Osterr. bot. Z.* **293**, 227–235 (2011).

262. Hermsen, J. G. T. & Verdenius, J. Selection from Solanum tuberosum group phureja of genotypes combining high-frequency haploid induction with homozygosity for embryo-spot. *Euphytica* **22**, 244–259 (1973).

263. Samitsu, Y. & Hosaka, K. Molecular marker analysis of 24- and 25-chromosome plants obtained from Solanum tuberosum L. subsp. andigena (2n = 4x = 48) pollinated with a Solanum phureja haploid inducer. *Genome* **45**, 577–583 (2002).

264. Ordoñez, B. *et al.* PL-4 (CIP596131.4): an Improved Potato Haploid Inducer. *Am. J. Potato Res.* **98**, 255–262 (2021).

265. Jansky, S. (2021).

266. Pallais, N. Osmotic priming of true potato seed: Effects of seed age. *Potato Res.* **32**, 235–244 (1989).

267. D'Antonio, V. L. & McHale, N. A. Effect of storage temperature and extraction methods on dormancy and germination of true potato seed. *Am. Potato J.* **65**, 573–581 (1988).

268. Simmonds, N. W. Experiments on the germination of potato seeds. I. *European Potato Journal* vol. 6 45–60 (1963).

269. Simmonds, N. W. Experiments on the germination of potato seeds. II. *Europ. Potato J.* 6, 69–76 (1963).

270. Lam, S.-L. & Erickson, H. T. Interaction of light and gibberellin on potato seed germination.

Am. Potato J. **43**, 443–449 (1966).

271. Lam, S.-L. Interaction of Temperature and Gibberellin on Potato Seed Germination. *Am. J. Bot.* **55**, 193–198 (1968).

272. Rayburn, A. L., Auger, J. A., Benzinger, E. A. & Hepburn, A. G. Detection of Intraspecific DNA Content Variation in Zea mays L. by Flow Cytometry. *J. Exp. Bot.* **40**, 1179–1183 (1989).

273. Bino, R. J., Lanteri, S., Verhoeven, H. A. & Kraak, H. L. Flow Cytometric Determination of Nuclear Replication Stage in Seed Tissues. *Ann. Bot.* **72**, 181–187 (1993).

274. Ghislain, M., Zhang, D. P. ;. & Herrera, M. R. *Molecular Biology Laboratory Protocols Plant Genotyping: Training Manual*. (International Potato Center, 1999).

275. Pham, G. M. *et al.* Construction of a chromosome-scale long-read reference genome assembly for potato. *Gigascience* 9, (2020).

276. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bio.GN]* (2013).

277. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, (2021).

278. Broman, K. W. & Sen, S. *A Guide to QTL Mapping with R/qtl*. (Springer, New York, NY, 2009).

279. Arends, D., Prins, P., Jansen, R. C. & Broman, K. W. R/qtl: high-throughput multiple QTL mapping. *Bioinformatics* **26**, 2990–2992 (2010).

280. Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889–890 (2003).

281. Huo, H. *et al.* Rapid identification of lettuce seed germination mutants by bulked segregant analysis and whole genome sequencing. *Plant J.* **88**, 345–360 (2016).

282. Michelmore, R. W., Paran, I. & Kesseli, R. V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. U. S. A.*

88, 9828–9832 (1991).

283. Abe, A. *et al.* Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat. Biotechnol.* **30**, 174–178 (2012).

284. Team, R. C. & Others. R: A language and environment for statistical computing. (2013).

285. Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum Likelihood from Incomplete Data Via the*EM*Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* vol. 39 1–22 (1977).

286. Lander, E. S. & Botstein, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185–199 (1989).

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

288. Chen, Z. The Full EM Algorithm for the MLEs of QTL Effects and Positions and Their Estimated Variances in Multiple-Interval Mapping. *Biometrics* vol. 61 474–480 (2005).

289. Ruden, D. *et al.* Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. *Front. Genet.* 3, 35 (2012).

290. Stelly, D. M. & Peloquin, S. J. Formation of 2n megagametophytes in diploid tuber-bearing Solanums. *Am. J. Bot.* (1986).

291. Kuppu, S. *et al.* A variety of changes, including CRISPR/Cas9-mediated deletions, in CENH3 lead to haploid induction on outcrossing. *Plant Biotechnology Journal* vol. 18 2068–2080 (2020).

292. Kuppu, S. *et al.* Point Mutations in Centromeric Histone Induce Post-zygotic Incompatibility and Uniparental Inheritance. *PLoS Genet.* **11**, e1005494 (2015).

293. Azumi, Y. *et al.* Homolog interaction during meiotic prophase I in Arabidopsis requires the SOLO DANCERS gene encoding a novel cyclin-like protein. *EMBO J.* **21**, 3081–3095 (2002).

294. Amundson, K. R. *et al.* Rare instances of haploid inducer DNA in potato dihaploids and ploidy-dependent genome instability. *Plant Cell* (2021) doi:10.1093/plcell/koab100.

295. Dodds, K. S. & Long, D. H. The inheritance of colour in diploid potatoes. *J. Genet.* **53**, 136–149 (1955).

296. Jong, H. D. Inheritance of anthocyanin pigmentation in the cultivated potato: A critical review. *Am. Potato J.* **68**, 585–593 (1991).

297. Salaman, R. N. The inheritance of colour and other characters in the potato. *J. Genet.* (1910).

298. Endelman, J. B. & Jansky, S. H. Genetic mapping with an inbred line-derived F2 population in potato. *Theor. Appl. Genet.* **129**, 935–943 (2016).

299. Jung, C. S. *Identities of the P and D loci in potato (Solanum tuberosum L.)*. (Cornell University, Jan, 2005).

300. Zhang, Y., Jung, C. S. & De Jong, W. S. Genetic analysis of pigmented tuber flesh in potato. *Theor. Appl. Genet.* **119**, 143–150 (2009).

301. Jung, C. S. *et al.* The potato developer (D) locus encodes an R2R3 MYB transcription factor that regulates expression of multiple anthocyanin structural genes in tuber skin. *Theor. Appl. Genet.* **120**, 45–57 (2009).

302. Ortiz, R. & M Golmirzaie, A. Genetic parameters for agronomic characteristics. I. Early and intermediate breeding populations of true potato seed. *Hereditas* **139**, 212–216 (2003).

303. Baskin, J. M. & Baskin, C. C. How much influence does the paternal parent have on seed germination? *Seed Sci. Res.* **29**, 1–11 (2019).

304. De Jong, T. J., Van Dijk, H. & Klinkhamer, P. G. L. Hamilton's rule, imprinting and parent-offspring conflict over seed mass in partially selfing plants. *J. Evol. Biol.* **18**, 676–682 (2005).

305. de Jong, T. J. & Scott, R. J. Parental conflict does not necessarily lead to the evolution of imprinting. *Trends Plant Sci.* **12**, 439–443 (2007).

306. de Jong, T. J., Hermans, C. M. & van der Veen-van Wijk, K. C. A. M. Paternal effects on seed mass in Arabidopsis thaliana. *Plant Biol.* **13 Suppl 1**, 71–77 (2011).

307. Raunsgard, A. *et al.* Intersexual conflict over seed size is stronger in more outcrossed populations of a mixed-mating plant. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 11561–11566 (2018).

308. van Eck, H. J., Jacobs, J. M. E., van den Berg, P. M. M. M., Stiekema, W. J. & Jacobsen, E. The inheritance of anthocyanin pigmentation in potato (Solanum tuberosum L.) and mapping of tuber skin colour loci using RFLPs. *Heredity* **73**, 410–421 (1994).

309. Hazen, S. P. *et al.* Rapid array mapping of circadian clock and developmental mutations in Arabidopsis. *Plant Physiol.* **138**, 990–997 (2005).

310. Tengkun, N., Dongdong, W., Xiaohui, M., Yue, C. & Qin, C. Analysis of Key Genes Involved in Potato Anthocyanin Biosynthesis Based on Genomics and Transcriptomics Data. *Frontiers in Plant Science* vol. 10 (2019).

311. Parra-Galindo, M. A., Soto-Sedano, J. C., Mosquera-Vásquez, T. & Roda, F. Pathway-based analysis of anthocyanin diversity in diploid potato. *PLoS One* **16**, e0250861 (2021).

312. Jung, C. S. *et al.* The potato P locus codes for flavonoid 3',5'-hydroxylase. *Züchter Genet. Breed. Res.* **110**, 269–275 (2005).

313. Jansky, S. H. *et al.* Reinventing potato as a diploid inbred line-based crop. *Crop Sci.* **56**, 1412–1422 (2016).

314. Lindhout, P. *et al.* Towards F1 Hybrid Seed Potato Breeding. *Potato Res.* **54**, 301–312 (2011).

315. Simmonds, N. W. Seed size in the cultivated potatoes. *Euphytica* **14**, 143–152 (1965).

316. Arndt, G. C. & Peloquin, S. J. The identification and evaluation of hybrid plants among open pollinated true seed families. *Am. Potato J.* **67**, 393–404 (1990).

317. d'Erfurth, I. *et al.* Turning meiosis into mitosis. *PLoS Biol.* 7, e1000124 (2009).

1. Blakeslee, A. F., Belling, J., Farnham, M. E. & Bergner, A. D. A Haploid Mutant in the Jimson Weed, ``Datura Stramonium''. Science 55, 646–647 (1922).
2. Forster, B. P., Heberle-Bors, E., Kasha, K. J. & Touraev, A. The resurgence of haploids in higher plants. Trends Plant Sci. 12, 368–375 (2007).
3. Seymour, D. K. et al. Rapid creation of Arabidopsis doubled haploid lines for quantitative trait locus mapping. Proc. Natl. Acad. Sci. U. S. A. 109, 4227–4232 (2012).
4. Ravi, M. & Chan, S. W. L. Haploid plants produced by centromere-mediated genome elimination. Nature 464, 615–618 (2010).
5. Ravi, M. et al. A haploid genetics toolbox for Arabidopsis thaliana. Nat. Commun. 5, 5334 (2014).
6. Qiu, F. et al. Morphological, cellular and molecular evidences of chromosome random elimination in vivo upon haploid induction in maize. Current Plant Biology 1, 83–90 (2014).
7. Liu, L. et al. In vivo haploid induction leads to increased frequency of twin-embryo and abnormal fertilization in maize. BMC Plant Biol. 18, 313 (2018).
8. Magoon, M. L. & Khanna, K. R. Haploids. Caryologia 16, 191–235 (1963).
9. Lakshmanan, K. K. & Ambegaokar, K. B. Polyembryony. Embryology of Angiosperms 445–474 (1984) doi:10.1007/978-3-642-69302-1_9.
10. Webber, J. M. Polyembryony. Bot. Rev. 6, 575–598 (1940).
11. Khanday, I., Skinner, D., Yang, B., Mercier, R. & Sundaresan, V. A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. Nature 565, 91–95 (2019).
12. Ozias-Akins, P., Akiyama, Y. & Hanna, W. W. Molecular characterization of the genomic region linked with apomixis in Pennisetum/Cenchrus. Funct. Integr. Genomics 3, 94–104 (2003).
13. Sprague, G. F. The Nature and Extent of Hetero-Fertilization in Maize. Genetics 17, 358–368 (1932).

14. Gao, S. et al. Revisiting the hetero-fertilization phenomenon in maize. PLoS One 6, e16101 (2011).
15. Randolph, L. F. Note on haploid frequencies. Maize Genet Coop Newsl 14, 23–24 (1940).
16. Sarkar, K. R. & Coe, E. H. A genetic analysis of the origin of maternal haploids in maize. Genetics 54, 453–464 (1966).
17. Chase, S. S. Monoploids and monoploid-derivatives of maize (Zea mays L.). Bot. Rev. 35, 117–168 (1969).
18. Lashermes, P. & Beckert, M. Genetic control of maternal haploidy in maize (Zea mays L.) and selection of haploid inducing lines. Theor. Appl. Genet. 76, 405–410 (1988).
19. Coe, E. H. A Line of Maize with High Haploid Frequency. Am. Nat. 93, 381–382 (1959).
20. Zhang, Z. et al. Chromosome elimination and in vivo haploid production induced by Stock 6-derived inducer line in maize (Zea mays L.). Plant Cell Rep. 27, 1851–1860 (2008).
21. Coe, E. H. High-haploid line. in Maize Genetics Cooperation Newsletter (MNL) 138–139 (1957).
22. Chase, S. S. & Others. Monoploids in maize. Heterosis 389–399 (1952).
23. Rotarenco, V., Dicu, G. & Fuia, S. New inducers of maternal haploids in maize. Maize Genet. Coop. News Lett. (2010).
24. Hu, H. et al. The Genetic Basis of Haploid Induction in Maize Identified with a Novel Genome-Wide Association Method. Genetics 202, 1267–1276 (2016).
25. Tian, X. et al. Hetero-fertilization together with failed egg-sperm cell fusion supports single fertilization involved in in vivo haploid induction in maize. J. Exp. Bot. 69, 4689–4701 (2018).
26. Chang, M. T. Preferential fertilization induced from Stock 6. Maize Genet Coop Newslett 66, 99–100 (1992).
27. Rotarenco, V. & Eder, J. Possible effects of heterofertilization on the induction of maternal haploids in maize. Maize Genetics Cooperation Newsletter (MNL) 77, 30 (2003).

28. Xu, X. et al. Gametophytic and zygotic selection leads to segregation distortion through in vivo induction of a maternal haploid in maize. J. Exp. Bot. 64, 1083–1096 (2013).
29. Dong, X. et al. Fine mapping of qhir1 influencing in vivo haploid induction in maize. Theor. Appl. Genet. 126, 1713–1720 (2013).
30. Kelliher, T. et al. MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. Nature 542, 105–109 (2017).
31. Barret, P., Brinkmann, M. & Beckert, M. A major locus expressed in the male gametophyte with incomplete penetrance is responsible for in situ gynogenesis in maize. Theor. Appl. Genet. 117, 581–594 (2008).
32. Li, X. et al. Single nucleus sequencing reveals spermatid chromosome fragmentation as a possible cause of maize haploid induction. Nat. Commun. 8, 991 (2017).
33. Swapna, M. & Sarkar, K. R. Anomalous fertilization in haploidy inducer lines in maize (Zea mays L). Maydica 56, (2012).
34. Morgan, D. T., Jr & Rappleye, R. D. Polyembryony in maize and lily; following X-irradiation of the pollen. J. Hered. 42, 90–93 (1951).
35. Li, L., Dong, X., Xu, X. W., Liu, L. W. & Liu, C. X. Observation of twin seedling in maternal haploid induction in maize. Journal of China (2012).
36. Bylich, V. G. & Chalyk, S. T. Existence of pollen grains with a pair of morphologically different sperm nuclei as a possible cause of the haploid-inducing capacity in ZMS line. Maize Genetics Cooperation Newsletter 70, 33–33 (1996).
37. Sun, G. et al. Matrilineal empowers wheat pollen with haploid induction potency via triggering post-mitosis ROS activity. New Phytol. (2022) doi:10.1111/nph.17963.
38. Chalyk, S., Baumann, A., Daniel, G. & Eder, J. Aneuploidy as a possible cause of haploid-induction in maize. Maize Genetics Cooperation Newsletter 77, 29–29 (2003).
39. Sun, G., Geng, S., Zhang, H., Jia, M. & Wang, Z. Matrilineal empowers wheat pollen with haploid induction potency via triggering post-mitosis ROS activity. New (2022).

40. Jiang, C. et al. A reactive oxygen species burst causes haploid induction in maize. Molecular Plant (2022) doi:10.1016/j.molp.2022.04.001.
41. Roldán-Arjona, T. & Ariza, R. R. Repair and tolerance of oxidative DNA damage in plants. Mutat. Res. 681, 169–179 (2009).
42. Srinivas, U. S., Tan, B. W. Q., Vellayappan, B. A. & Jeyasekharan, A. D. ROS and the DNA damage response in cancer. Redox Biol 25, 101084 (2019).
43. Rowe, L. A., Degtyareva, N. & Doetsch, P. W. DNA damage-induced reactive oxygen species (ROS) stress response in Saccharomyces cerevisiae. Free Radic. Biol. Med. 45, 1167–1177 (2008).
44. Choudhary, A., Kumar, A. & Kaur, N. ROS and oxidative burst: Roots in plant development. Plant Divers 42, 33–43 (2020).
45. Nisa, M.-U., Huang, Y., Benhamed, M. & Raynaud, C. The Plant DNA Damage Response: Signaling Pathways Leading to Growth Inhibition and Putative Role in Response to Stress Conditions. Front. Plant Sci. 10, 653 (2019).
46. Sanei, M., Pickering, R., Kumke, K., Nasuda, S. & Houben, A. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc. Natl. Acad. Sci. U. S. A. 108, E498–505 (2011).
47. Comai, L. & Tan, E. H. Haploid Induction and Genome Instability. Trends Genet. 35, 791–803 (2019).
48. Amundson, K. R. et al. Rare instances of haploid inducer DNA in potato dihaploids and ploidy-dependent genome instability. Plant Cell (2021) doi:10.1093/plcell/koab100.
49. Li, L., Xu, X., Jin, W. & Chen, S. Morphological and molecular evidences for DNA introgression in haploid induction via a high oil inducer CAUHOI in maize. Planta 230, 367–376 (2009).
50. Zhao, X., Xu, X., Xie, H., Chen, S. & Jin, W. Fertilization and uniparental chromosome elimination during crosses with maize haploid inducers. Plant Physiol. 163, 721–731

(2013).

51. Riera-Lizarazu, O., Rines, H. W. & Phillips, R. L. Cytological and molecular characterization of oat x maize partial hybrids. Theor. Appl. Genet. 93, 123–135 (1996).

52. Wedzony, M., Röber, F. K. & Geiger, H. H. Chromosome elimination observed in selfed progenies of maize inducer line RWS. in XVIIth International Congress on Sex Plant Reports Maria Curie--Sklodowska University Press, Lublin vol. 173 (2002).

53. Röber, F. K., Gordillo, G. A. & Geiger, H. H. In vivo haploid induction in maize. Performance of new inducers and significance of doubled haploid lines in hybrid breeding [Zea mays L.]. Maydica (2005).

54. Fischer, E. Molekulargenetische Untersuchungen zum Vorkommen paternaler DNA-Übertragung bei der in-vivo-Haploideninduktion bei Mais (Zea mays L.). (Grauer, 2004).

55. Kelliher, T. et al. One-step genome editing of elite crop germplasm during haploid induction. Nat. Biotechnol. 37, 287–292 (2019).

56. Rotarenco, V. A., Dicu, G., Sarmaniuc, M. & Others. Induction of maternal haploids in maize. Maize Genetics Cooperation Newsletter 83, 21 (2009).

57. Liu, C. et al. New insight into the mechanism of heterofertilization during maize haploid induction. Euphytica 213, 174 (2017).

58. Kasahara, R. D. et al. Fertilization recovery after defective sperm cell release in Arabidopsis. Curr. Biol. 22, 1084–1089 (2012).

59. Kasahara, R. D., Maruyama, D. & Higashiyama, T. Fertilization recovery system is dependent on the number of pollen grains for efficient reproduction in plants. Plant Signal. Behav. 8, e23690 (2013).

60. Higashiyama, T. & Takeuchi, H. The mechanism and key molecules involved in pollen tube guidance. Annu. Rev. Plant Biol. 66, 393–413 (2015).

61. Takahashi, T. et al. The male gamete membrane protein DMP9/DAU2 is required for double fertilization in flowering plants. Development 145, (2018).

62. Kato, A. Induced single fertilization in maize. Sex. Plant Reprod. 10, 96–100 (1997).

63. Enaleeva, N. K., Tyrnov, V. S., Selivanova, L. P. & Zavalishina, A. N. Single fertilization and the problem of haploidy induction in maize. in Doklady Biological Sciences vol. 353 225 (New York: Consultants Bureau, c1965-1992., 1997).

64. Jorgensen, G. A. The experimental formation of heteroploid plants in the genussolanum. J. Genet. 19, 133 (1928).

65. Pace, L. Apogamy in Atamosco. Bot. Gaz. 56, 376–394 (1913).

66. Beale, K. M., Leydon, A. R. & Johnson, M. A. Gamete fusion is required to block multiple pollen tubes from entering an Arabidopsis ovule. Curr. Biol. 22, 1090–1094 (2012).

67. Prigge, V. et al. New insights into the genetics of in vivo induction of maternal haploids, the backbone of doubled haploid technology in maize. Genetics 190, 781–793 (2012).

68. Gilles, L. M. et al. Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. EMBO J. 36, 707–717 (2017).

69. Liu, C. et al. Fine mapping of qhir8 affecting in vivo haploid induction in maize. Theor. Appl. Genet. 128, 2507–2515 (2015).

70. Zhong, Y. et al. Mutation of ZmDMP enhances haploid induction in maize. Nat Plants 5, 575–580 (2019).

71. Deimling, S., Röber, F. K. & Geiger, H. H. Methodology and genetics of in vivo haploid induction in maize. Vortr. Pflanzenzüchtg 38, 203–224 (1997).

72. Nair, S. K. et al. Dissection of a major QTL qhir1 conferring maternal haploid induction ability in maize. Theor. Appl. Genet. 130, 1113–1122 (2017).

73. Liu, C. et al. A 4-bp Insertion at ZmPLA1 Encoding a Putative Phospholipase A Generates Haploid Induction in Maize. Mol. Plant 10, 520–522 (2017).

74. Gilles, L. M. et al. Lipid anchoring and electrostatic interactions target the phospholipase NOT-LIKE-DAD to pollen endo-plasma membrane. Cold Spring Harbor Laboratory 2020.10.05.326157 (2020) doi:10.1101/2020.10.05.326157.

75. Yao, L. et al. OsMATL mutation induces haploid seed formation in indica rice. Nat Plants 4, 530–533 (2018).
76. Liu, H. et al. Efficient induction of haploid plants in wheat by editing of TaMTL using an optimized Agrobacterium-mediated CRISPR system. J. Exp. Bot. 71, 1337–1349 (2020).
77. Hafidh, S., Fíla, J. & Honys, D. Male gametophyte development and function in angiosperms: a general concept. Plant Reprod. 29, 31–51 (2016).
78. Lalanne, E. & Twell, D. Genetic control of male germ unit organization in Arabidopsis. Plant Physiol. 129, 865–875 (2002).
79. Taimur, N. A study of molecular and genetic mechanisms mediating the formation of twin sperm cells in Arabidopsis thaliana. (2016).
80. Li, Y. et al. Loss-of-function alleles of ZmPLD3 cause haploid induction in maize. Nat Plants 7, 1579–1588 (2021).
81. Kasaras, A. & Kunze, R. Expression, localisation and phylogeny of a novel family of plant-specific membrane proteins. Plant Biol. 12 Suppl 1, 140–152 (2010).
82. Cyprys, P., Lindemeier, M. & Sprunck, S. Gamete fusion is facilitated by two sperm cell-expressed DUF679 membrane proteins. Nat Plants 5, 253–257 (2019).
83. Zhong, Y. et al. A DMP-triggered in vivo maternal haploid induction system in the dicotyledonous Arabidopsis. Nat Plants 6, 466–472 (2020).
84. Berger, F. & Twell, D. Germline specification and function in plants. Annu. Rev. Plant Biol. 62, 461–484 (2011).
85. Frank, A. C. & Johnson, M. A. Expressing the diphtheria toxin A subunit from the HAP2(GCS1) promoter blocks sperm maturation and produces single sperm-like cells capable of fertilization. Plant Physiol. 151, 1390–1400 (2009).
86. Chumakov, M. I. Matroclinic Haploidy and Gamete Interaction in Maize. Russ. J. Genet. 54, 1137–1141 (2018).
87. Sprunck, S. et al. Egg cell-secreted EC1 triggers sperm cell activation during double

fertilization. Science 338, 1093–1097 (2012).

88. Mori, T., Igawa, T., Tamiya, G., Miyagishima, S.-Y. & Berger, F. Gamete attachment requires GEX2 for successful fertilization in Arabidopsis. Curr. Biol. 24, 170–175 (2014).

89. Dresselhaus, T. & Sprunck, S. Plant fertilization: maximizing reproductive success. Current biology: CB vol. 22 R487–9 (2012).

90. von Besser, K., Frank, A. C., Johnson, M. A. & Preuss, D. Arabidopsis HAP2 (GCS1) is a sperm-specific gene required for pollen tube guidance and fertilization. Development 133, 4761–4769 (2006).

91. Hafidh, S. & Honys, D. Reproduction Multitasking: The Male Gametophyte. Annu. Rev. Plant Biol. 72, 581–614 (2021).

92. Englhart, M. Direct Gamete Interactions and EGG CELL1-mediated Sperm Cell Activation in Arabidopsis Thaliana. (Universitätsbibliothek Regensburg, 2018).

93. Huang, J., Ju, Y., Wang, X., Zhang, Q. & Sodmergen. A one-step rectification of sperm cell targeting ensures the success of double fertilization. J. Integr. Plant Biol. 57, 496–503 (2015).

94. Gusti, A. et al. The Arabidopsis thaliana F-box protein FBL17 is essential for progression through the second mitosis during pollen development. PLoS One 4, e4780 (2009).

95. Borg, M. et al. The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in Arabidopsis. Plant Cell 23, 534–549 (2011).

96. Higashiyama, T. et al. Pollen tube attraction by the synergid cell. Science 293, 1480–1483 (2001).

97. Boerman, N. A., Frei, U. K. & Lübberstedt, T. Impact of Spontaneous Haploid Genome Doubling in Maize Breeding. Plants vol. 9 369 (2020).

98. Chaikam, V., Gowda, M., Nair, S. K., Melchinger, A. E. & Boddupalli, P. M. Genome-wide association study to identify genomic regions influencing spontaneous fertility in maize

haploids. Euphytica 215, 138 (2019).

99. Zhong, Y. et al. A genotype independent DMP-HI system in dicot crops. bioRxiv 2021.06.21.449224 (2021) doi:10.1101/2021.06.21.449224.

100. Wang, N., Gent, J. I. & Dawe, R. K. Haploid induction by a maize cenh3 null mutant. Sci Adv 7, (2021).

101. Marimuthu, M. P. A. et al. Biased removal and loading of centromeric histone H3 during reproduction underlies uniparental genome elimination. bioRxiv 2021.02.24.432754 (2021) doi:10.1101/2021.02.24.432754.

102. Battaglia, E. Unusual Cytological Features in the Apomictic Rudbeckia Sullivantii Boynton et Beadle. Caryologia 8, 1–32 (1955).

103. Chaudhari, H. K. Use of Semigamy in the Production of Cotton Haploids. Bull. Torrey Bot. Club 105, 98–103 (1978).

104. Turcotte, E. L. & Feaster, C. V. Semigamy in Pima Cotton. J. Hered. 58, 55–57 (1967).

105. Lanaud, C. Origin of haploids and semigamy in Theobroma cacao L. Euphytica 38, 221–228 (1988).

106. Montelongo-Escobedo, H. & Rowe, P. R. Haploid induction in potato: cytological basis for the pollinator effect. Euphytica 18, 116–123 (1969).

107. Montezuma-de-Carvalho, J. The effect of N2O on pollen tube mitosis in styles and its potential significance for inducing haploidy in potato. Euphytica 16, 190–198 (1967).

108. Wangenheim, K.-H. v., Peloquin, S. J. & Hougas, R. W. Embryological investigations on the formation of haploids in the potato (Solanum tuberosum). Zeitschrift für Vererbungslehre 91, 391–399 (1960).

109. Eder, J. & Chalyk, S. In vivo haploid induction in maize. Theor. Appl. Genet. 104, 703–708 (2002).

110. Sprunck, S., Hackenberg, T., Englhart, M. & Vogler, F. Same same but different: sperm-activating EC1 and ECA1 gametogenesis-related family proteins. Biochem. Soc. Trans. 42,

401–407 (2014).

111. Huamán, Z., Golmirzaie, A. & Amoros, W. The potato. Biodiversity in trust: conservation and use of plant genetic resources in CGIAR Centres 21–28 (1997).

112. Bradshaw, J. E. & Ramsay, G. Potato Origin and Production. Advances in Potato Chemistry and Technology 1–26 (2009) doi:10.1016/b978-0-12-374349-7.00001-5.

113. Den Nijs, T. P. M. & Peloquin, S. J. Polyploid evolution via 2n gametes. Am. Potato J. 54, 377–386 (1977).

114. den Nijs, T. P. M. & Peloquin, S. J. 2n gametes in potato species and their function in sexual polyploidization. Euphytica 26, 585–600 (1977).

115. Plaisted, R. L. Advances and limitations in the utilization of Neotuberosum in potato breeding. The production of new potato varieties: technological advances 186–196 doi:10.1017/cbo9780511895463.041.

116. Glendinning, D. R. Neo-Tuberosum: new potato breeding material. 1. The origin, composition, and development of the Tuberosum and Neo-Tuberosum gene pools. Potato Research vol. 18 256–261 (1975).

117. Bethke, P. C. et al. History and origin of russet Burbank (netted gem) a sport of Burbank. Am. J. Potato Res. 91, 594–609 (2014).

118. Ames, M. & Spooner, D. M. DNA from herbarium specimens settles a controversy about origins of the European potato. Am. J. Bot. 95, 252–257 (2008).

119. Spooner, D. M., Ghislain, M., Simon, R., Jansky, S. H. & Gavrilenko, T. Systematics, Diversity, Genetics, and Evolution of Wild and Cultivated Potatoes. Bot. Rev. 80, 283–383 (2014).

120. Hanneman, R. E. The potato germplasm resource. Am. Potato J. 66, 655–667 (1989).

121. Ovchinnikova, A. et al. Taxonomy of cultivated potatoes (Solanum section Petota: Solanaceae). Bot. J. Linn. Soc. 165, 107–155 (2011).

122. Murovec, J. & Bohanec, B. Haploids and Doubled Haploids in Plant Breeding. Plant

Breeding (2012) doi:10.5772/29982.

123. Sanetomo, R. & Hosaka, K. Re-evaluation of Monohaploid Solanum verrucosum and S. bulbocastanum (2n = x = 12) and Dihaploid S. stoloniferum and S. acaule (2n = 2x = 24), All Derived from Anther Culture. American Journal of Potato Research vol. 98 333–343 (2021).

124. Ivanovskaja, E. V. A haploid plant of Solatium tuberosum LCR. Doklady Acad. Sei. USSR 24, 517–520 (1939).

125. Hougas, R. W. & Peloquin, S. J. A haploid plant of the potato variety katahdin. Nature 180, 1209–1210 (1957).

126. Bains, G. S. & Howard, H. W. Haploid plants of Solanum demissum. Nature 166, 795 (1950).

127. Dodds, K. S. Polyhaploids of Solanum demissum. Nature 166, 795 (1950).

128. Hougas, R. W., Peloquin, S. J. & Ross, R. W. HAPLOIDS OF THE COMMON POTATO. J. Hered. 49, 103–106 (1958).

129. Hougas, R. W., Peloquin, S. J. & Ross, R. W. The potential of haploid potatoes for research and breeding. Am. J. Potato Res. 35, 442 (1958).

130. Gabert, A. C. Factors influencing the frequency of haploids in the common potato (Solanum tuberosum L.). (University of Wisconsin--Madison, 1963).

131. Kawakami, K. & Mastubayashi, M. Studies on the Haploid Plants of Solanum tuberosum : I. Morphological Characterisitics of the Polyhaploid Plants. Japanese Journal of Breeding vol. 10 10–18 (1960).

132. Matsubayashi, M. Studies on the Haploid Plants of Solanum tuberosum: II. Meiotic chromosome pairing in the polyhaploid plants. Japanese Journal of Breeding 10, 195–202 (1960).

133. Hanneman, R. E. & Peloquin, S. J. Ploidy levels of progeny from diploid-tetraploid crosses in the potato. Am. Potato J. 45, 255–261 (1968).

134. Hermsen, J. G. T. Induction of haploids and aneuhaploids in colchicine-induced tetraploid Solanum chacoense Bitt. Euphytica 18, 183–189 (1969).
135. Huamán, Z. & Ross, R. W. Updated listing of potato species names, abbreviations and taxonomic status. American Potato Journal vol. 62 629–641 (1985).
136. USDA United States Department of Agriculture. Solanum tuberosum L. subsp. andigenum (Juz. & Bukasov) Hawkes GRIN-Global. https://web.archive.org/web/20210405112021/https://npgsweb.ars-grin.gov/gringlobal/taxon/taxonomydetail?id=101304.
137. Howard, H. W. & Swaminathan, M. S. The cytology of haploid plants ofSolanum demissum. Genetica 26, 381–391 (1953).
138. Hougas, R. W., Peloquin, S. J. & Gabert, A. C. Effect of seed-parent and pollinator on frequency of haploids in Solanum tuberosum 1. Crop Sci. 4, 593–595 (1964).
139. Bender, K. The production and origin of dihaploid plants in Solanum tuberosum. Zeitschrift fuer Pflanzenzuechtung (West Germany) : 50, (1963).
140. Lamm, R. Notes on a haploid potato hybrid. Hereditas 24, 391–396 (2010).
141. Caligari, P. D. S., Powell, W., Liddell, K., Maine, M. J. D. E. & Swan, G. E. L. Methods and strategies for detecting Solanum tuberosum dihaploids in interspecific crosses with S. phureja. Ann. Appl. Biol. 112, 323–328 (1988).
142. Hutten, R. C. B., E J M, Huigen, D. J., Th. Hermsen, J. G. & Jacobsen, E. Analysis of dihaploid induction and production ability and seed parent x pollinator interaction in potato. Euphytica vol. 72 61–64 (1994).
143. Peloquin, S. J. & Hougas, R. W. Decapitation and genetic markers as related to haploidy inSolanum tuberosum. European Potato Journal vol. 2 176–183 (1959).
144. Hermsen, J. G. T. & Verdenius, J. Selection from Solanum tuberosum group phureja of genotypes combining high-frequency haploid induction with homozygosity for embryo-spot. Euphytica 22, 244–259 (1973).

145. van Breukelen, E., Ramanna, M. S. & Hermsen, J. G. T. Parthenogenetic monohaploids (2n= x= 12) from Solanum tuberosum L. and S. verrucosum Schlechtd. and the production of homozygous potato diploids. Euphytica (1977).

146. Uijtewaal, B. A., Huigen, D. J. & Th. Hermsen, J. G. Production of potato monohapioids (2n = x = 12) through prickle pollination. Theor. Appl. Genet. 73, 751–758 (1987).

147. Lough, R. C., Varrieur, J. M. & Veilleux, R. E. Selection inherent in monoploid derivation mechanisms for potato. Züchter Genet. Breed. Res. 103, 178–184 (2001).

148. Van Breukelen, E. W. M., Ramanna, M. S. & Hermsen, J. G. T. Monohaploids (n=x=12) from autotetraploid Solanum tuberosum (2n=4x=48) through two successive cycles of female parthenogenesis. Euphytica (1975).

149. Van Breukelen, E. W. M. Pseudogamic production of dihaploids and monoploids in Solanum tuberosum and some related species. (1981).

150. Singsit, C. & Hanneman, R. E. Haploids of tetraploid (2n=4x=48) Mexican potato species—Their extraction, cytology and crossability. Am. Potato J. 64, 469–482 (1987).

151. Clulow, S. A. et al. Cytological and molecular observations on Solanum phureja-induced dihaploid potatoes. Theor. Appl. Genet. 82, 545–551 (1991).

152. Clulow, S. A., Wilkinson, M. J. & Burch, L. R. Solanum phureja genes are expressed in the leaves and tubers of aneusomatic potato dihaploids. Euphytica 69, 1–6 (1993).

153. Clulow, S. A. & Rousselle-Bourgeois, F. Widespread introgression of Solatium phureja DNA in potato (S. tuberosum) dihaploids. Plant Breed. 116, 347–351 (1997).

154. PI 225682 GRIN-Global. https://web.archive.org/web/20210607182642/https://npgsweb.ars-grin.gov/gringlobal/accessiondetail?id=1181411.

155. PI 225702 GRIN-Global. https://web.archive.org/web/20210607185000/https://npgsweb.ars-grin.gov/gringlobal/accessiondetail?id=1181427.

156. Cooper, D. C. & Rieman, G. H. Diploid plants in a seedling population of the cultivated potato. Am. Potato J. 35, 642–642 (1958).

157. Peloquin, S. J., Gabert, A. C. & Ortiz, R. Nature of ` Pollinator ' Effect in Potato (Solanum tuberosum L.) Haploid Production. Ann. Bot. 77, 539–542 (1996).

158. Wasilewicz-Flis, I. & Jakuczun, H. Induction and application of dihaploids of potato (Solanum tuberosum L.). Plant Breeding and Seed Science 52, 47–64 (2005).

159. Carroll, C. P. & Low, R. J. Aspects of male fertility in group Tuberosum dihaploids. Potato Res. 19, 109–121 (1976).

160. Carroll, C. P. & Low, R. J. Flowering behaviour and seed fertility in dihaploidSolanum tuberosum. Potato Res. 18, 416–427 (1975).

161. Buketova, M. I. Choice of pollinators for producing dihaploids. Kartofel'i Ovoshchi (1970).

162. Moyle, L. C. & Nakazato, T. Hybrid incompatibility 'snowballs' between Solanum species. Science 329, 1521–1523 (2010).

163. Rice, W. R. & Holland, B. The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. Behav. Ecol. Sociobiol. 41, 1–10 (1997).

164. Clark, N. L., Aagaard, J. E. & Swanson, W. J. Evolution of reproductive proteins from animals and plants. Reproduction 131, 11–22 (2006).

165. Willi, Y. The battle of the sexes over seed size: support for both kinship genomic imprinting and interlocus contest evolution. Am. Nat. 181, 787–798 (2013).

166. Ivanov, M. A. Experimental production of haploids in nicotiana rustical L. Genetica 20, 295–397 (1938).

167. Lindstrom, E. W. A HAPLOID MUTANT IN THE TOMATO*. Journal of Heredity vol. 20 23–30 (1929).

168. Krishnaswamy, N. Cytological studies in a haploid plant of Triticum vulgare. Hereditas 25, 77–86 (1939).

169. Marks, G. E. A Polyhaploid Plant of Solanum polytrichon Rydb. Nature 175, 469–469

(1955).

170. Abdalla, M. M. F. & Th. Hermsen, J. G. Diploid parthenogenesis and androgenesis in diploid Solanum. Euphytica vol. 21 426–431 (1972).

171. Laurie, D. A. & Bennett, M. D. Wheat × maize hybridization. Canadian Journal of Genetics and Cytology vol. 28 313–316 (1986).

172. Camadro, E. L., Masuelli, R. W. & Cortés, M. C. Haploids of the wild tetraploid potato Solanum acaule ssp. acaule: generation, meiotic behavior, and electrophoretic pattern for the aspartate aminotransferase system. Genome vol. 35 431–435 (1992).

173. Inagaki, M. & Mujeeb-Kazi, A. Comparison of Polyhaploid Production Frequencies in Crosses of Hexaploid Wheat with Maize, Pearl Millet and Sorghuml. Japanese Journal of Breeding vol. 45 157–161 (1995).

174. Chaudhary, H. K., Sethi, G. S., Singh, S., Pratap, A. & Sharma, S. Efficient haploid induction in wheat by using pollen of Imperata cylindrica. Plant Breed. 124, 96–98 (2005).

175. Ren, J. et al. Novel technologies in doubled haploid line development. Plant Biotechnology Journal vol. 15 1361–1370 (2017).

176. Camadro, E. L., Erazzú, L. E., Maune, J. F. & Bedogni, M. C. A genetic approach to the species problem in wild potato. Plant Biol. 14, 543–554 (2012).

177. Huaman, Z. [Systematic botany, identification, distribution and evaluation of cultivated potato [Solanum tuberosum]]. Documento de Capacitacion (CIP) (1984).

178. Huamán, Z. & Spooner, D. M. Reclassification of landrace populations of cultivated potatoes (Solanum sect. Petota). Am. J. Bot. 89, 947–965 (2002).

179. Cersosimo, A. & Sorrentino, C. Haploidy in Nicotiana spp. in In Vitro Haploid Production in Higher Plants: Volume 5 — Oil, Ornamental and Miscellaneous Plants (eds. Jain, S. M., Sopory, S. K. & Veilleux, R. E.) 135–163 (Springer Netherlands, 1997).

180. Burk, L. G., Gerstel, D. U. & Wernsman, E. A. Maternal Haploids of Nicotiana tabacum L. from Seed. Science 206, 585 (1979).

181. Satina, S., Rappaport, J. & Blakeslee, A. F. Ovular Tumors Connected with Incompatible Crosses in Datura. Am. J. Bot. 37, 576–586 (1950).

182. Brisson, D. Negative Frequency-Dependent Selection Is Frequently Confounding. Front Ecol Evol 6, (2018).

183. Barber, H. N. The suppression of meiosis and the origin of diplochromosomes. Proceedings of the Royal Society of London. Series B - Biological Sciences 128, 170–185 (1940).

184. Gastony, G. J. & Windham, M. D. Species Concepts in Pteridophytes: The Treatment and Definition of Agamosporous Species. American Fern Journal vol. 79 65 (1989).

185. Shimamura, T. Cytological Studies of Polyploidy Induced by Colchicine. CYTOLOGIA vol. 9 486–494 (1939).

186. Briggs, R. An analysis of the inactivation of the frog sperm nucleus by toluidine blue. J. Gen. Physiol. 35, 761–780 (1952).

187. Chimoyo, H. M. The effect of ultraviolet irradiation, toluidine blue, and environment on maternal haploid frequencies from the cross between Nicotiana tabacum and N. africana. (Clemson University, 1988).

188. Edwards, R. G. Colchicine-induced heteroploidy in early mouse embryos. Nature 174, 276–277 (1954).

189. Edwards, R. G. The experimental induction of pseudogamy in early mouse embryos. Experientia 10, 499–500 (1954).

190. Edwards, R. G. The experimental induction of gynogenesis in the mouse. III. Treatment of sperm with trypaflavine, toluidine blue, or nitrogen mustard. Proc. R. Soc. Lond. B Biol. Sci. 149, 117–129 (1958).

191. Illies, Z. M. Induction of haploid parthenogenesis in aspen by post-pollination treatment with Toluidine-blue. Silvae Genet. 23, 221–226 (1974).

192. NsokA, M. & KoNDo, Y. Production of gynogenetic haploids by treatment of spermatozoa with toluidine blue in Rana japonica. Sci. Rep. Lab. Amphibian Biol. Hiroshima Univ. 3,

385–398 (1978).

193. Stockert, J. C., Juarranz, A., Villanueva, A. & Cañete, M. Photodynamic damage to HeLa cell microtubules induced by thiazine dyes. Cancer Chemother. Pharmacol. 39, 167–169 (1996).

194. Bacellar, I. O. L. et al. Membrane damage efficiency of phenothiazinium photosensitizers. Photochem. Photobiol. 90, 801–813 (2014).

195. Paardekooper, M. et al. Photodynamic treatment of yeast cells with the dye toluidine blue: all-or-none loss of plasma membrane barrier properties. Biochim. Biophys. Acta 1108, 86–90 (1992).

196. Vázquez-Ortega, F. et al. Photodynamic treatment induced membrane cell damage in Acanthamoeba castellanii Neff. Dyes and Pigments vol. 180 108481 (2020).

197. Paardekooper, M., De Bruijne, A. W., Van Steveninck, J. & Van den Broek, P. J. Inhibition of transport systems in yeast by photodynamic treatment with toluidine blue. Biochim. Biophys. Acta 1151, 143–148 (1993).

198. Tuite, E. M. & Kelly, J. M. New trends in photobiology: Photochemical interactions of methylene blue and analogues with DNA and other biological substrates. J. Photochem. Photobiol. B 21, 103–124 (1993).

199. Sridharan, G. & Shankar, A. A. Toluidine blue: A review of its chemistry and clinical utility. J. Oral Maxillofac. Pathol. 16, 251–255 (2012).

200. Kato, A. Induction of bicellular pollen and dihaploidization of tetraploid maize. MAIZE GENETICS COOPERATION NEWSLETTER 70, 25–25 (1996).

201. Waugh, R., Baird, E. & Powell, W. The use of RAPD markers for the detection of gene introgression in potato. Plant Cell Rep. 11, 466–469 (1992).

202. Wilkinson, M. J. et al. Evidence for somatic translocation during potato dihaploid induction. Heredity 74 (Pt 2), 146–151 (1995).

203. Straadt, I. K. & Rasmussen, O. S. AFLP analysis of Solanum phureja DNA introgressed

into potato dihaploids. Plant Breed. 122, 352–356 (2003).

204. Ercolano, M. R. et al. Assessment of genetic variability of haploids extracted from tetraploid (2n = 4x = 48) Solanum tuberosum. Genome 47, 633–638 (2004).

205. Bartkiewicz, A. M. et al. Maximization of Markers Linked in Coupling for Tetraploid Potatoes via Monoparental Haploids. Front. Plant Sci. 9, 620 (2018).

206. Pham, G. M. et al. Genome-wide inference of somatic translocation events during potato dihaploid production. Plant Genome 12, 180079 (2019).

207. Marks, G. E. The enigma of triploid potatoes. Euphytica 15, 285–290 (1966).

208. Van Suchtelen, N. Triploids of the common potato,Solanum tuberosum L. Potato Research vol. 19 377–380 (1976).

209. Graebner, R. C., Chen, H., Contreras, R. N., Haynes, K. G. & Sathuvalli, V. Identification of the High Frequency of Triploid Potato Resulting from Tetraploid × Diploid Crosses. HortScience vol. 54 1159–1163 (2019).

210. Franklin-Tong, V. E. Self-Incompatibility in Flowering Plants: Evolution, Diversity, and Mechanisms. (Springer Science & Business Media, 2008).

211. Igic, B., Smith, W. A., Robertson, K. A., Schaal, B. A. & Kohn, J. R. Studies of self-incompatibility in wild tomatoes: I. S-allele diversity in Solanum chilense (Dun.) Reiche [corrected] (Solanaceae). Heredity 99, 553–561 (2007).

212. Dzidzienyo, D. K., Bryan, G. J., Wilde, G. & Robbins, T. P. Allelic diversity of S-RNase alleles in diploid potato species. Theor. Appl. Genet. 129, 1985–2001 (2016).

213. Newbigin, E., Anderson, M. A. & Clarke, A. E. Gametophytic Self-Incompatibility Systems. The Plant Cell vol. 5 1315 (1993).

214. Fujii, S., Kubo, K.-I. & Takayama, S. Non-self- and self-recognition models in plant self-incompatibility. Nat Plants 2, 16130 (2016).

215. Puul, S. 2n GAMETES IN DIPLOID SOLANUM: FREQUENCY AND TYPES OF SPINDLE ABNORMALITIES1. Can. J. Genet. Cytol. 24, 301–314 (1982).

216. Ramanna, M. S. A re-examination of the mechanisms of 2n gamete formation in potato and its implications for breeding. Euphytica 28, 537–561 (1979).

217. Carputo, D., Barone, A. & Frusciante, L. 2n gametes in the potato: essential ingredients for breeding and germplasm transfer. Theor. Appl. Genet. 101, 805–813 (2000).

218. Mable, B. K. Polyploidy and self-compatibility: is there an association? New Phytol. 162, 803–811 (2004).

219. Mendiburu, A. O. & Peloquin, S. J. Sexual polyploidization and depolyploidization: some terminology and definitions. Theoretical and Applied Genetics vol. 48 137–143 (1976).

220. Bretagnolle, F. & Thompson, J. D. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. New Phytol. 129, 1–22 (1995).

221. Iwanaga, M. & Peloquin, S. J. Origin and evolution of cultivated tetraploid potatoes via 2n gametes. Theor. Appl. Genet. 61, 161–169 (1982).

222. Quinn, A. A., Mok, D. W. S. & Peloquin, S. J. Distribution and significance of diplandroids among the diploid solanums. Am. Potato J. 51, 16–21 (1974).

223. Watanabe, K. & Peloquin, S. J. Cytological basis of 2n pollen formation in a wide range of 2x, 4x, and 6x taxa from tuber-bearing Solanum species. Genome 36, 8–13 (1993).

224. Stelly, D. M. & Peloquin, S. J. FORMATION OF 2N MEGAGAMETOPHYTES IN DIPLOID TUBER-BEARING SOLANUMS. American Journal of Botany vol. 73 1351–1363 (1986).

225. Camadro, E. L. & Peloquin, S. J. The occurrence and frequency of 2n pollen in three diploid solanums from Northwest Argentina. Theoretical and Applied Genetics vols 56-56 11–15 (1980).

226. Jongedijk, E. Desynapsis and FDR 2N-megaspore formation in diploid potato : potentials and limitations for breeding and for the induction of diplosporic apomixis. (Jongedijk, 1991).

227. Werner, J. E. & Peloquin, S. J. Occurrence and mechanisms of 2n egg formation in 2x potato. Genome vol. 34 975–982 (1991).

228. Correll, D. S. & Others. The potato and its wild relatives. The potato and its wild relatives. (1962).

229. Navarro, C. et al. Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. Nature 478, 119–122 (2011).

230. Abelenda, J. A., Navarro, C. & Prat, S. Flowering and tuberization: a tale of two nightshades. Trends Plant Sci. 19, 115–122 (2014).

231. Dnyansagar, V. R. & Cooper, D. C. Development of the Seed of Solanum phureja. Am. J. Bot. 47, 176–186 (1960).

232. Hanneman, R. E. & Peloquin, S. J. Croosability of 24-chromosome potato hybrids with 48-chromosome cultivars. European Potato Journal 10, 62–73 (1967).

233. Johnston, S. A. & Hanneman, R. E., Jr. The genetics of triploid formation and its relationship to endosperm balance number in potato. Genome 38, 60–67 (1995).

234. Masuelli, R. W. & Camadro, E. L. Crossability relationships among wild potato species with different ploidies and Endosperm Balance Numbers (EBN). Euphytica 94, 227–235 (1997).

235. Jackson, M. T., Rowe, P. R. & Hawkes, J. G. Crossability relationships of andean potato varieties of three ploidy levels. Euphytica 27, 541–551 (1978).

236. Tarn, T. R. & Hawkes, J. G. Cytogenetic studies and the occurrence of triploidy in the wild potato species Solanum commersonii Dun. Euphytica 35, 293–302 (1986).

237. Graebner, R. C., Chen, H., Contreras, R. N., Haynes, K. G. & Sathuvalli, V. Identification of the High Frequency of Triploid Potato Resulting from Tetraploid× Diploid Crosses. HortScience 54, 1159–1163 (2019).

238. Köhler, C., Mittelsten Scheid, O. & Erilova, A. The impact of the triploid block on the origin and evolution of polyploid plants. Trends Genet. 26, 142–148 (2010).

239. Kolář, F., Čertner, M., Suda, J., Schönswetter, P. & Husband, B. C. Mixed-Ploidy Species: Progress and Opportunities in Polyploid Research. Trends Plant Sci. 22, 1041–1055 (2017).

240. Husband, B. C. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. Biol. J. Linn. Soc. Lond. 82, 537–546 (2004).

241. Alfano, F., Cammareri, M., Errico, A., Frusciante, L. & Conicella, C. 2n gametes inSolanum tuberosum dihaploids. Am. J. Potato Res. 76, 281–285 (1999).

242. Conicella, C., Genualdo, G., Errico, A., Frusciante, L. & Monti, L. M. Meiotic restitution mechanisms and 2n pollen formation in a Solanum tuberosum dihaploid and in dihaploid x wild species hybrids*. Plant Breeding vol. 115 157–161 (1996).

243. Sax, K. THE EXPERIMENTAL PRODUCTION OF POLYPLOIDY. J. Arnold Arbor. 17, 153–159 (1936).

244. Comai, L. The advantages and disadvantages of being polyploid. Nat. Rev. Genet. 6, 836–846 (2005).

245. Petit, C., Bretagnolle, F. & Felber, F. Evolutionary consequences of diploid–polyploid hybrid zones in wild species. Trends Ecol. Evol. 14, 306–311 (1999).

246. McHale, N. A. Environmental induction of high frequency 2n pollen formation in diploid Solanum. Can. J. Genet. Cytol. 25, 609–615 (1983).

247. Castro, M. O. Evolutionary ecology of polyploids: understanding species coexistence at the contact zones. (00500::Universidade de Coimbra, 2019).

248. Buggs, R. J. A. & Pannell, J. R. Ecological differentiation and diploid superiority across a moving ploidy contact zone. Evolution 61, 125–140 (2007).

249. Soltis, P. S., Marchant, D. B., Van de Peer, Y. & Soltis, D. E. Polyploidy and genome evolution in plants. Curr. Opin. Genet. Dev. 35, 119–125 (2015).

250. Levin, D. A. & Soltis, D. E. Factors promoting polyploid persistence and diversification and limiting diploid speciation during the K–Pg interlude. Current Opinion in Plant Biology vol. 42 1–7 (2018).

251. Stebbins, G. L. Apomixis in the angiosperms. Bot. Rev. 7, 507–542 (1941).

252. de Wet, J. M. J. REVERSIBLE TETRAPLOIDY AS AN EVOLUTIONARY MECHANISM.

Evolution 25, 545–548 (1971).

253. Husband, B. C., Baldwin, S. J. & Suda, J. The Incidence of Polyploidy in Natural Plant Populations: Major Patterns and Evolutionary Processes. in Plant Genome Diversity Volume 2: Physical Structure, Behaviour and Evolution of Plant Genomes (eds. Greilhuber, J., Dolezel, J. & Wendel, J. F.) 255–276 (Springer Vienna, 2013).

254. Hanzl, M., Kolář, F., Nováková, D. & Suda, J. Nonadaptive processes governing early stages of polyploid evolution: Insights from a primary contact zone of relict serpentine Knautia arvensis (Caprifoliaceae). Am. J. Bot. 101, 935–945 (2014).

255. Tandon, S. L. & Rao, G. R. Cytogenetical Investigations in Relation to the Mechanism of Evolution in Hexaploid Solanum nigrum L. Nature 201, 1348–1349 (1964).

256. Melo, C. A. F., Martins, M. I. G., Oliveira, M. B. M., Benko-Iseppon, A. M. & Carvalho, R. Karyotype analysis for diploid and polyploid species of the Solanum L. Osterr. bot. Z. 293, 227–235 (2011).

257. Hermsen, J. G. T. & Verdenius, J. Selection from Solanum tuberosum group phureja of genotypes combining high-frequency haploid induction with homozygosity for embryo-spot. Euphytica 22, 244–259 (1973).

258. Samitsu, Y. & Hosaka, K. Molecular marker analysis of 24- and 25-chromosome plants obtained from Solanum tuberosum L. subsp. andigena (2n = 4x = 48) pollinated with a Solanum phureja haploid inducer. Genome 45, 577–583 (2002).

259. Ordoñez, B. et al. PL-4 (CIP596131.4): an Improved Potato Haploid Inducer. Am. J. Potato Res. 98, 255–262 (2021).

260. Jansky, S. (2021).

261. Pallais, N. Osmotic priming of true potato seed: Effects of seed age. Potato Res. 32, 235–244 (1989).

262. D'Antonio, V. L. & McHale, N. A. Effect of storage temperature and extraction methods on dormancy and germination of true potato seed. Am. Potato J. 65, 573–581 (1988).

263. Simmonds, N. W. Experiments on the germination of potato seeds. I. European Potato Journal vol. 6 45–60 (1963).

264. Simmonds, N. W. Experiments on the germination of potato seeds. II. Europ. Potato J. 6, 69–76 (1963).

265. Lam, S.-L. & Erickson, H. T. Interaction of light and gibberellin on potato seed germination. Am. Potato J. 43, 443–449 (1966).

266. Lam, S.-L. Interaction of Temperature and Gibberellin on Potato Seed Germination. Am. J. Bot. 55, 193–198 (1968).

267. Rayburn, A. L., Auger, J. A., Benzinger, E. A. & Hepburn, A. G. Detection of Intraspecific DNA Content Variation in Zea mays L. by Flow Cytometry. J. Exp. Bot. 40, 1179–1183 (1989).

268. Bino, R. J., Lanteri, S., Verhoeven, H. A. & Kraak, H. L. Flow Cytometric Determination of Nuclear Replication Stage in Seed Tissues. Ann. Bot. 72, 181–187 (1993).

269. Ghislain, M., Zhang, D. P. ;. & Herrera, M. R. Molecular Biology Laboratory Protocols Plant Genotyping: Training Manual. (International Potato Center, 1999).

270. Pham, G. M. et al. Construction of a chromosome-scale long-read reference genome assembly for potato. Gigascience 9, (2020).

271. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv [q-bio.GN] (2013).

272. Danecek, P. et al. Twelve years of SAMtools and BCFtools. Gigascience 10, (2021).

273. Broman, K. W. & Sen, S. A Guide to QTL Mapping with R/qtl. (Springer, New York, NY, 2009).

274. Arends, D., Prins, P., Jansen, R. C. & Broman, K. W. R/qtl: high-throughput multiple QTL mapping. Bioinformatics 26, 2990–2992 (2010).

275. Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19, 889–890 (2003).

276. Huo, H. et al. Rapid identification of lettuce seed germination mutants by bulked segregant analysis and whole genome sequencing. Plant J. 88, 345–360 (2016).
277. Michelmore, R. W., Paran, I. & Kesseli, R. V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. U. S. A. 88, 9828–9832 (1991).
278. Abe, A. et al. Genome sequencing reveals agronomically important loci in rice using MutMap. Nat. Biotechnol. 30, 174–178 (2012).
279. Team, R. C. & Others. R: A language and environment for statistical computing. (2013).
280. Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum Likelihood from Incomplete Data Via theEMAlgorithm. Journal of the Royal Statistical Society: Series B (Methodological) vol. 39 1–22 (1977).
281. Lander, E. S. & Botstein, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121, 185–199 (1989).
282. Haley, C. S. & Knott, S. A. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity vol. 69 315–324 (1992).
283. Chen, Z. The Full EM Algorithm for the MLEs of QTL Effects and Positions and Their Estimated Variances in Multiple-Interval Mapping. Biometrics vol. 61 474–480 (2005).
284. Ruden, D. et al. Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front. Genet. 3, 35 (2012).
285. Stelly, D. M. & Peloquin, S. J. Formation of 2n megagametophytes in diploid tuber-bearing Solanums. Am. J. Bot. (1986).
286. Kuppu, S. et al. A variety of changes, including CRISPR/Cas9-mediated deletions, in CENH3 lead to haploid induction on outcrossing. Plant Biotechnology Journal vol. 18 2068–2080 (2020).
287. Kuppu, S. et al. Point Mutations in Centromeric Histone Induce Post-zygotic Incompatibility

and Uniparental Inheritance. PLoS Genet. 11, e1005494 (2015).

288. Azumi, Y. et al. Homolog interaction during meiotic prophase I in Arabidopsis requires the SOLO DANCERS gene encoding a novel cyclin-like protein. EMBO J. 21, 3081–3095 (2002).

289. Amundson, K. R. et al. Rare instances of haploid inducer DNA in potato dihaploids and ploidy-dependent genome instability. Plant Cell (2021) doi:10.1093/plcell/koab100.

290. Dodds, K. S. & Long, D. H. The inheritance of colour in diploid potatoes. J. Genet. 53, 136–149 (1955).

291. Jong, H. D. Inheritance of anthocyanin pigmentation in the cultivated potato: A critical review. Am. Potato J. 68, 585–593 (1991).

292. Salaman, R. N. The inheritance of colour and other characters in the potato. J. Genet. (1910).

293. Endelman, J. B. & Jansky, S. H. Genetic mapping with an inbred line-derived F2 population in potato. Theor. Appl. Genet. 129, 935–943 (2016).

294. Jung, C. S. Identities of the P and D loci in potato (Solanum tuberosum L.). (Cornell University, Jan, 2005).

295. Zhang, Y., Jung, C. S. & De Jong, W. S. Genetic analysis of pigmented tuber flesh in potato. Theor. Appl. Genet. 119, 143–150 (2009).

296. Jung, C. S. et al. The potato developer (D) locus encodes an R2R3 MYB transcription factor that regulates expression of multiple anthocyanin structural genes in tuber skin. Theor. Appl. Genet. 120, 45–57 (2009).

297. Ortiz, R. & M Golmirzaie, A. Genetic parameters for agronomic characteristics. I. Early and intermediate breeding populations of true potato seed. Hereditas 139, 212–216 (2003).

298. Baskin, J. M. & Baskin, C. C. How much influence does the paternal parent have on seed germination? Seed Sci. Res. 29, 1–11 (2019).

299. De Jong, T. J., Van Dijk, H. & Klinkhamer, P. G. L. Hamilton's rule, imprinting and parent-

offspring conflict over seed mass in partially selfing plants. J. Evol. Biol. 18, 676–682 (2005).

300. de Jong, T. J. & Scott, R. J. Parental conflict does not necessarily lead to the evolution of imprinting. Trends Plant Sci. 12, 439–443 (2007).

301. de Jong, T. J., Hermans, C. M. & van der Veen-van Wijk, K. C. A. M. Paternal effects on seed mass in Arabidopsis thaliana. Plant Biol. 13 Suppl 1, 71–77 (2011).

302. Raunsgard, A. et al. Intersexual conflict over seed size is stronger in more outcrossed populations of a mixed-mating plant. Proc. Natl. Acad. Sci. U. S. A. 115, 11561–11566 (2018).

303. van Eck, H. J., Jacobs, J. M. E., van den Berg, P. M. M. M., Stiekema, W. J. & Jacobsen, E. The inheritance of anthocyanin pigmentation in potato (Solanum tuberosum L.) and mapping of tuber skin colour loci using RFLPs. Heredity 73, 410–421 (1994).

304. Hazen, S. P. et al. Rapid array mapping of circadian clock and developmental mutations in Arabidopsis. Plant Physiol. 138, 990–997 (2005).

305. Tengkun, N., Dongdong, W., Xiaohui, M., Yue, C. & Qin, C. Analysis of Key Genes Involved in Potato Anthocyanin Biosynthesis Based on Genomics and Transcriptomics Data. Frontiers in Plant Science vol. 10 (2019).

306. Parra-Galindo, M. A., Soto-Sedano, J. C., Mosquera-Vásquez, T. & Roda, F. Pathway-based analysis of anthocyanin diversity in diploid potato. PLoS One 16, e0250861 (2021).

307. Jung, C. S. et al. The potato P locus codes for flavonoid 3',5'-hydroxylase. Züchter Genet. Breed. Res. 110, 269–275 (2005).

308. Jansky, S. H. et al. Reinventing potato as a diploid inbred line-based crop. Crop Sci. 56, 1412–1422 (2016).

309. Lindhout, P. et al. Towards F1 Hybrid Seed Potato Breeding. Potato Res. 54, 301–312 (2011).

310. Simmonds, N. W. Seed size in the cultivated potatoes. Euphytica 14, 143–152 (1965).

311. Arndt, G. C. & Peloquin, S. J. The identification and evaluation of hybrid plants among open pollinated true seed families. Am. Potato J. 67, 393–404 (1990).

312. d'Erfurth, I. et al. Turning meiosis into mitosis. PLoS Biol. 7, e1000124 (2009).

1. Blakeslee, A. F., Belling, J., Farnham, M. E. & Bergner, A. D. A Haploid Mutant in the Jimson Weed, ``Datura Stramonium''. Science 55, 646–647 (1922).
2. Forster, B. P., Heberle-Bors, E., Kasha, K. J. & Touraev, A. The resurgence of haploids in higher plants. Trends Plant Sci. 12, 368–375 (2007).
3. Seymour, D. K. et al. Rapid creation of Arabidopsis doubled haploid lines for quantitative trait locus mapping. Proc. Natl. Acad. Sci. U. S. A. 109, 4227–4232 (2012).
4. Ravi, M. & Chan, S. W. L. Haploid plants produced by centromere-mediated genome elimination. Nature 464, 615–618 (2010).
5. Ravi, M. et al. A haploid genetics toolbox for Arabidopsis thaliana. Nat. Commun. 5, 5334 (2014).
6. Qiu, F. et al. Morphological, cellular and molecular evidences of chromosome random elimination in vivo upon haploid induction in maize. Current Plant Biology 1, 83–90 (2014).
7. Liu, L. et al. In vivo haploid induction leads to increased frequency of twin-embryo and abnormal fertilization in maize. BMC Plant Biol. 18, 313 (2018).
8. Magoon, M. L. & Khanna, K. R. Haploids. Caryologia 16, 191–235 (1963).
9. Lakshmanan, K. K. & Ambegaokar, K. B. Polyembryony. Embryology of Angiosperms 445–474 (1984) doi:10.1007/978-3-642-69302-1_9.
10. Webber, J. M. Polyembryony. Bot. Rev. 6, 575–598 (1940).
11. Khanday, I., Skinner, D., Yang, B., Mercier, R. & Sundaresan, V. A male-expressed rice embryogenic trigger redirected for asexual propagation through seeds. Nature 565, 91–95 (2019).
12. Ozias-Akins, P., Akiyama, Y. & Hanna, W. W. Molecular characterization of the genomic

region linked with apomixis in Pennisetum/Cenchrus. Funct. Integr. Genomics 3, 94–104 (2003).

13. Sprague, G. F. The Nature and Extent of Hetero-Fertilization in Maize. Genetics 17, 358–368 (1932).
14. Gao, S. et al. Revisiting the hetero-fertilization phenomenon in maize. PLoS One 6, e16101 (2011).
15. Randolph, L. F. Note on haploid frequencies. Maize Genet Coop Newsl 14, 23–24 (1940).
16. Sarkar, K. R. & Coe, E. H. A genetic analysis of the origin of maternal haploids in maize. Genetics 54, 453–464 (1966).
17. Chase, S. S. Monoploids and monoploid-derivatives of maize (Zea mays L.). Bot. Rev. 35, 117–168 (1969).
18. Lashermes, P. & Beckert, M. Genetic control of maternal haploidy in maize (Zea mays L.) and selection of haploid inducing lines. Theor. Appl. Genet. 76, 405–410 (1988).
19. Coe, E. H. A Line of Maize with High Haploid Frequency. Am. Nat. 93, 381–382 (1959).
20. Zhang, Z. et al. Chromosome elimination and in vivo haploid production induced by Stock 6-derived inducer line in maize (Zea mays L.). Plant Cell Rep. 27, 1851–1860 (2008).
21. Coe, E. H. High-haploid line. in Maize Genetics Cooperation Newsletter (MNL) 138–139 (1957).
22. Chase, S. S. & Others. Monoploids in maize. Heterosis 389–399 (1952).
23. Rotarenco, V., Dicu, G. & Fuia, S. New inducers of maternal haploids in maize. Maize Genet. Coop. News Lett. (2010).
24. Hu, H. et al. The Genetic Basis of Haploid Induction in Maize Identified with a Novel Genome-Wide Association Method. Genetics 202, 1267–1276 (2016).
25. Tian, X. et al. Hetero-fertilization together with failed egg-sperm cell fusion supports single fertilization involved in in vivo haploid induction in maize. J. Exp. Bot. 69, 4689–4701 (2018).

26. Chang, M. T. Preferential fertilization induced from Stock 6. Maize Genet Coop Newslett 66, 99–100 (1992).
27. Rotarenco, V. & Eder, J. Possible effects of heterofertilization on the induction of maternal haploids in maize. Maize Genetics Cooperation Newsletter (MNL) 77, 30 (2003).
28. Xu, X. et al. Gametophytic and zygotic selection leads to segregation distortion through in vivo induction of a maternal haploid in maize. J. Exp. Bot. 64, 1083–1096 (2013).
29. Dong, X. et al. Fine mapping of qhir1 influencing in vivo haploid induction in maize. Theor. Appl. Genet. 126, 1713–1720 (2013).
30. Kelliher, T. et al. MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. Nature 542, 105–109 (2017).
31. Barret, P., Brinkmann, M. & Beckert, M. A major locus expressed in the male gametophyte with incomplete penetrance is responsible for in situ gynogenesis in maize. Theor. Appl. Genet. 117, 581–594 (2008).
32. Li, X. et al. Single nucleus sequencing reveals spermatid chromosome fragmentation as a possible cause of maize haploid induction. Nat. Commun. 8, 991 (2017).
33. Swapna, M. & Sarkar, K. R. Anomalous fertilization in haploidy inducer lines in maize (Zea mays L). Maydica 56, (2012).
34. Morgan, D. T., Jr & Rappleye, R. D. Polyembryony in maize and lily; following X-irradiation of the pollen. J. Hered. 42, 90–93 (1951).
35. Li, L., Dong, X., Xu, X. W., Liu, L. W. & Liu, C. X. Observation of twin seedling in maternal haploid induction in maize. Journal of China (2012).
36. Bylich, V. G. & Chalyk, S. T. Existence of pollen grains with a pair of morphologically different sperm nuclei as a possible cause of the haploid-inducing capacity in ZMS line. Maize Genetics Cooperation Newsletter 70, 33–33 (1996).
37. Sun, G. et al. Matrilineal empowers wheat pollen with haploid induction potency via triggering post-mitosis ROS activity. New Phytol. (2022) doi:10.1111/nph.17963.

38. Chalyk, S., Baumann, A., Daniel, G. & Eder, J. Aneuploidy as a possible cause of haploid-induction in maize. Maize Genetics Cooperation Newsletter 77, 29–29 (2003).
39. Sun, G., Geng, S., Zhang, H., Jia, M. & Wang, Z. Matrilineal empowers wheat pollen with haploid induction potency via triggering post-mitosis ROS activity. New (2022).
40. Jiang, C. et al. A reactive oxygen species burst causes haploid induction in maize. Molecular Plant (2022) doi:10.1016/j.molp.2022.04.001.
41. Roldán-Arjona, T. & Ariza, R. R. Repair and tolerance of oxidative DNA damage in plants. Mutat. Res. 681, 169–179 (2009).
42. Srinivas, U. S., Tan, B. W. Q., Vellayappan, B. A. & Jeyasekharan, A. D. ROS and the DNA damage response in cancer. Redox Biol 25, 101084 (2019).
43. Rowe, L. A., Degtyareva, N. & Doetsch, P. W. DNA damage-induced reactive oxygen species (ROS) stress response in Saccharomyces cerevisiae. Free Radic. Biol. Med. 45, 1167–1177 (2008).
44. Choudhary, A., Kumar, A. & Kaur, N. ROS and oxidative burst: Roots in plant development. Plant Divers 42, 33–43 (2020).
45. Nisa, M.-U., Huang, Y., Benhamed, M. & Raynaud, C. The Plant DNA Damage Response: Signaling Pathways Leading to Growth Inhibition and Putative Role in Response to Stress Conditions. Front. Plant Sci. 10, 653 (2019).
46. Sanei, M., Pickering, R., Kumke, K., Nasuda, S. & Houben, A. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc. Natl. Acad. Sci. U. S. A. 108, E498–505 (2011).
47. Comai, L. & Tan, E. H. Haploid Induction and Genome Instability. Trends Genet. 35, 791–803 (2019).
48. Amundson, K. R. et al. Rare instances of haploid inducer DNA in potato dihaploids and ploidy-dependent genome instability. Plant Cell (2021) doi:10.1093/plcell/koab100.
49. Li, L., Xu, X., Jin, W. & Chen, S. Morphological and molecular evidences for DNA

introgression in haploid induction via a high oil inducer CAUHOI in maize. Planta 230, 367–376 (2009).

50. Zhao, X., Xu, X., Xie, H., Chen, S. & Jin, W. Fertilization and uniparental chromosome elimination during crosses with maize haploid inducers. Plant Physiol. 163, 721–731 (2013).

51. Riera-Lizarazu, O., Rines, H. W. & Phillips, R. L. Cytological and molecular characterization of oat x maize partial hybrids. Theor. Appl. Genet. 93, 123–135 (1996).

52. Wedzony, M., Röber, F. K. & Geiger, H. H. Chromosome elimination observed in selfed progenies of maize inducer line RWS. in XVIIth International Congress on Sex Plant Reports Maria Curie--Sklodowska University Press, Lublin vol. 173 (2002).

53. Röber, F. K., Gordillo, G. A. & Geiger, H. H. In vivo haploid induction in maize. Performance of new inducers and significance of doubled haploid lines in hybrid breeding [Zea mays L.]. Maydica (2005).

54. Fischer, E. Molekulargenetische Untersuchungen zum Vorkommen paternaler DNA-Übertragung bei der in-vivo-Haploideninduktion bei Mais (Zea mays L.). (Grauer, 2004).

55. Kelliher, T. et al. One-step genome editing of elite crop germplasm during haploid induction. Nat. Biotechnol. 37, 287–292 (2019).

56. Rotarenco, V. A., Dicu, G., Sarmaniuc, M. & Others. Induction of maternal haploids in maize. Maize Genetics Cooperation Newsletter 83, 21 (2009).

57. Liu, C. et al. New insight into the mechanism of heterofertilization during maize haploid induction. Euphytica 213, 174 (2017).

58. Kasahara, R. D. et al. Fertilization recovery after defective sperm cell release in Arabidopsis. Curr. Biol. 22, 1084–1089 (2012).

59. Kasahara, R. D., Maruyama, D. & Higashiyama, T. Fertilization recovery system is dependent on the number of pollen grains for efficient reproduction in plants. Plant Signal. Behav. 8, e23690 (2013).

60. Higashiyama, T. & Takeuchi, H. The mechanism and key molecules involved in pollen tube guidance. Annu. Rev. Plant Biol. 66, 393–413 (2015).
61. Takahashi, T. et al. The male gamete membrane protein DMP9/DAU2 is required for double fertilization in flowering plants. Development 145, (2018).
62. Kato, A. Induced single fertilization in maize. Sex. Plant Reprod. 10, 96–100 (1997).
63. Enaleeva, N. K., Tyrnov, V. S., Selivanova, L. P. & Zavalishina, A. N. Single fertilization and the problem of haploidy induction in maize. in Doklady Biological Sciences vol. 353 225 (New York: Consultants Bureau, c1965-1992., 1997).
64. Jorgensen, G. A. The experimental formation of heteroploid plants in the genussolanum. J. Genet. 19, 133 (1928).
65. Pace, L. Apogamy in Atamosco. Bot. Gaz. 56, 376–394 (1913).
66. Beale, K. M., Leydon, A. R. & Johnson, M. A. Gamete fusion is required to block multiple pollen tubes from entering an Arabidopsis ovule. Curr. Biol. 22, 1090–1094 (2012).
67. Prigge, V. et al. New insights into the genetics of in vivo induction of maternal haploids, the backbone of doubled haploid technology in maize. Genetics 190, 781–793 (2012).
68. Gilles, L. M. et al. Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. EMBO J. 36, 707–717 (2017).
69. Liu, C. et al. Fine mapping of qhir8 affecting in vivo haploid induction in maize. Theor. Appl. Genet. 128, 2507–2515 (2015).
70. Zhong, Y. et al. Mutation of ZmDMP enhances haploid induction in maize. Nat Plants 5, 575–580 (2019).
71. Deimling, S., Röber, F. K. & Geiger, H. H. Methodology and genetics of in vivo haploid induction in maize. Vortr. Pflanzenzüchtg 38, 203–224 (1997).
72. Nair, S. K. et al. Dissection of a major QTL qhir1 conferring maternal haploid induction ability in maize. Theor. Appl. Genet. 130, 1113–1122 (2017).
73. Liu, C. et al. A 4-bp Insertion at ZmPLA1 Encoding a Putative Phospholipase A Generates

Haploid Induction in Maize. Mol. Plant 10, 520–522 (2017).

74. Gilles, L. M. et al. Lipid anchoring and electrostatic interactions target the phospholipase NOT-LIKE-DAD to pollen endo-plasma membrane. Cold Spring Harbor Laboratory 2020.10.05.326157 (2020) doi:10.1101/2020.10.05.326157.

75. Yao, L. et al. OsMATL mutation induces haploid seed formation in indica rice. Nat Plants 4, 530–533 (2018).

76. Liu, H. et al. Efficient induction of haploid plants in wheat by editing of TaMTL using an optimized Agrobacterium-mediated CRISPR system. J. Exp. Bot. 71, 1337–1349 (2020).

77. Hafidh, S., Fíla, J. & Honys, D. Male gametophyte development and function in angiosperms: a general concept. Plant Reprod. 29, 31–51 (2016).

78. Lalanne, E. & Twell, D. Genetic control of male germ unit organization in Arabidopsis. Plant Physiol. 129, 865–875 (2002).

79. Taimur, N. A study of molecular and genetic mechanisms mediating the formation of twin sperm cells in Arabidopsis thaliana. (2016).

80. Li, Y. et al. Loss-of-function alleles of ZmPLD3 cause haploid induction in maize. Nat Plants 7, 1579–1588 (2021).

81. Kasaras, A. & Kunze, R. Expression, localisation and phylogeny of a novel family of plant-specific membrane proteins. Plant Biol. 12 Suppl 1, 140–152 (2010).

82. Cyprys, P., Lindemeier, M. & Sprunck, S. Gamete fusion is facilitated by two sperm cell-expressed DUF679 membrane proteins. Nat Plants 5, 253–257 (2019).

83. Zhong, Y. et al. A DMP-triggered in vivo maternal haploid induction system in the dicotyledonous Arabidopsis. Nat Plants 6, 466–472 (2020).

84. Berger, F. & Twell, D. Germline specification and function in plants. Annu. Rev. Plant Biol. 62, 461–484 (2011).

85. Frank, A. C. & Johnson, M. A. Expressing the diphtheria toxin A subunit from the HAP2(GCS1) promoter blocks sperm maturation and produces single sperm-like cells

capable of fertilization. Plant Physiol. 151, 1390–1400 (2009).

86. Chumakov, M. I. Matroclinic Haploidy and Gamete Interaction in Maize. Russ. J. Genet. 54, 1137–1141 (2018).

87. Sprunck, S. et al. Egg cell-secreted EC1 triggers sperm cell activation during double fertilization. Science 338, 1093–1097 (2012).

88. Mori, T., Igawa, T., Tamiya, G., Miyagishima, S.-Y. & Berger, F. Gamete attachment requires GEX2 for successful fertilization in Arabidopsis. Curr. Biol. 24, 170–175 (2014).

89. Dresselhaus, T. & Sprunck, S. Plant fertilization: maximizing reproductive success. Current biology: CB vol. 22 R487–9 (2012).

90. von Besser, K., Frank, A. C., Johnson, M. A. & Preuss, D. Arabidopsis HAP2 (GCS1) is a sperm-specific gene required for pollen tube guidance and fertilization. Development 133, 4761–4769 (2006).

91. Hafidh, S. & Honys, D. Reproduction Multitasking: The Male Gametophyte. Annu. Rev. Plant Biol. 72, 581–614 (2021).

92. Englhart, M. Direct Gamete Interactions and EGG CELL1-mediated Sperm Cell Activation in Arabidopsis Thaliana. (Universitätsbibliothek Regensburg, 2018).

93. Huang, J., Ju, Y., Wang, X., Zhang, Q. & Sodmergen. A one-step rectification of sperm cell targeting ensures the success of double fertilization. J. Integr. Plant Biol. 57, 496–503 (2015).

94. Gusti, A. et al. The Arabidopsis thaliana F-box protein FBL17 is essential for progression through the second mitosis during pollen development. PLoS One 4, e4780 (2009).

95. Borg, M. et al. The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in Arabidopsis. Plant Cell 23, 534–549 (2011).

96. Higashiyama, T. et al. Pollen tube attraction by the synergid cell. Science 293, 1480–1483 (2001).

97. Boerman, N. A., Frei, U. K. & Lübberstedt, T. Impact of Spontaneous Haploid Genome Doubling in Maize Breeding. Plants vol. 9 369 (2020).
98. Chaikam, V., Gowda, M., Nair, S. K., Melchinger, A. E. & Boddupalli, P. M. Genome-wide association study to identify genomic regions influencing spontaneous fertility in maize haploids. Euphytica 215, 138 (2019).
99. Zhong, Y. et al. A genotype independent DMP-HI system in dicot crops. bioRxiv 2021.06.21.449224 (2021) doi:10.1101/2021.06.21.449224.
100. Wang, N., Gent, J. I. & Dawe, R. K. Haploid induction by a maize cenh3 null mutant. Sci Adv 7, (2021).
101. Marimuthu, M. P. A. et al. Biased removal and loading of centromeric histone H3 during reproduction underlies uniparental genome elimination. bioRxiv 2021.02.24.432754 (2021) doi:10.1101/2021.02.24.432754.
102. Battaglia, E. Unusual Cytological Features in the Apomictic Rudbeckia Sullivantii Boynton et Beadle. Caryologia 8, 1–32 (1955).
103. Chaudhari, H. K. Use of Semigamy in the Production of Cotton Haploids. Bull. Torrey Bot. Club 105, 98–103 (1978).
104. Turcotte, E. L. & Feaster, C. V. Semigamy in Pima Cotton. J. Hered. 58, 55–57 (1967).
105. Lanaud, C. Origin of haploids and semigamy in Theobroma cacao L. Euphytica 38, 221–228 (1988).
106. Montelongo-Escobedo, H. & Rowe, P. R. Haploid induction in potato: cytological basis for the pollinator effect. Euphytica 18, 116–123 (1969).
107. Montezuma-de-Carvalho, J. The effect of N2O on pollen tube mitosis in styles and its potential significance for inducing haploidy in potato. Euphytica 16, 190–198 (1967).
108. Wangenheim, K.-H. v., Peloquin, S. J. & Hougas, R. W. Embryological investigations on the formation of haploids in the potato (Solanum tuberosum). Zeitschrift für Vererbungslehre 91, 391–399 (1960).

109. Eder, J. & Chalyk, S. In vivo haploid induction in maize. Theor. Appl. Genet. 104, 703–708 (2002).

110. Sprunck, S., Hackenberg, T., Englhart, M. & Vogler, F. Same same but different: sperm-activating EC1 and ECA1 gametogenesis-related family proteins. Biochem. Soc. Trans. 42, 401–407 (2014).

111. Huamán, Z., Golmirzaie, A. & Amoros, W. The potato. Biodiversity in trust: conservation and use of plant genetic resources in CGIAR Centres 21–28 (1997).

112. Bradshaw, J. E. & Ramsay, G. Potato Origin and Production. Advances in Potato Chemistry and Technology 1–26 (2009) doi:10.1016/b978-0-12-374349-7.00001-5.

113. Den Nijs, T. P. M. & Peloquin, S. J. Polyploid evolution via 2n gametes. Am. Potato J. 54, 377–386 (1977).

114. den Nijs, T. P. M. & Peloquin, S. J. 2n gametes in potato species and their function in sexual polyploidization. Euphytica 26, 585–600 (1977).

115. Plaisted, R. L. Advances and limitations in the utilization of Neotuberosum in potato breeding. The production of new potato varieties: technological advances 186–196 doi:10.1017/cbo9780511895463.041.

116. Glendinning, D. R. Neo-Tuberosum: new potato breeding material. 1. The origin, composition, and development of the Tuberosum and Neo-Tuberosum gene pools. Potato Research vol. 18 256–261 (1975).

117. Bethke, P. C. et al. History and origin of russet Burbank (netted gem) a sport of Burbank. Am. J. Potato Res. 91, 594–609 (2014).

118. Ames, M. & Spooner, D. M. DNA from herbarium specimens settles a controversy about origins of the European potato. Am. J. Bot. 95, 252–257 (2008).

119. Spooner, D. M., Ghislain, M., Simon, R., Jansky, S. H. & Gavrilenko, T. Systematics, Diversity, Genetics, and Evolution of Wild and Cultivated Potatoes. Bot. Rev. 80, 283–383 (2014).

120. Hanneman, R. E. The potato germplasm resource. Am. Potato J. 66, 655–667 (1989).

121. Ovchinnikova, A. et al. Taxonomy of cultivated potatoes (Solanum section Petota: Solanaceae). Bot. J. Linn. Soc. 165, 107–155 (2011).

122. Murovec, J. & Bohanec, B. Haploids and Doubled Haploids in Plant Breeding. Plant Breeding (2012) doi:10.5772/29982.

123. Sanetomo, R. & Hosaka, K. Re-evaluation of Monohaploid Solanum verrucosum and S. bulbocastanum (2n = x = 12) and Dihaploid S. stoloniferum and S. acaule (2n = 2x = 24), All Derived from Anther Culture. American Journal of Potato Research vol. 98 333–343 (2021).

124. Ivanovskaja, E. V. A haploid plant of Solatium tuberosum LCR. Doklady Acad. Sei. USSR 24, 517–520 (1939).

125. Hougas, R. W. & Peloquin, S. J. A haploid plant of the potato variety katahdin. Nature 180, 1209–1210 (1957).

126. Bains, G. S. & Howard, H. W. Haploid plants of Solanum demissum. Nature 166, 795 (1950).

127. Dodds, K. S. Polyhaploids of Solanum demissum. Nature 166, 795 (1950).

128. Hougas, R. W., Peloquin, S. J. & Ross, R. W. HAPLOIDS OF THE COMMON POTATO. J. Hered. 49, 103–106 (1958).

129. Hougas, R. W., Peloquin, S. J. & Ross, R. W. The potential of haploid potatoes for research and breeding. Am. J. Potato Res. 35, 442 (1958).

130. Gabert, A. C. Factors influencing the frequency of haploids in the common potato (Solanum tuberosum L.). (University of Wisconsin--Madison, 1963).

131. Kawakami, K. & Mastubayashi, M. Studies on the Haploid Plants of Solanum tuberosum : I. Morphological Characterisitics of the Polyhaploid Plants. Japanese Journal of Breeding vol. 10 10–18 (1960).

132. Matsubayashi, M. Studies on the Haploid Plants of Solanum tuberosum: II. Meiotic

chromosome pairing in the polyhaploid plants. Japanese Journal of Breeding 10, 195–202 (1960).

133. Hanneman, R. E. & Peloquin, S. J. Ploidy levels of progeny from diploid-tetraploid crosses in the potato. Am. Potato J. 45, 255–261 (1968).

134. Hermsen, J. G. T. Induction of haploids and aneuhaploids in colchicine-induced tetraploid Solanum chacoense Bitt. Euphytica 18, 183–189 (1969).

135. Huamán, Z. & Ross, R. W. Updated listing of potato species names, abbreviations and taxonomic status. American Potato Journal vol. 62 629–641 (1985).

136. USDA United States Department of Agriculture. Solanum tuberosum L. subsp. andigenum (Juz. & Bukasov) Hawkes GRIN-Global. https://web.archive.org/web/20210405112021/https://npgsweb.ars-grin.gov/gringlobal/taxon/taxonomydetail?id=101304.

137. Howard, H. W. & Swaminathan, M. S. The cytology of haploid plants ofSolanum demissum. Genetica 26, 381–391 (1953).

138. Hougas, R. W., Peloquin, S. J. & Gabert, A. C. Effect of seed-parent and pollinator on frequency of haploids in Solanum tuberosum 1. Crop Sci. 4, 593–595 (1964).

139. Bender, K. The production and origin of dihaploid plants in Solanum tuberosum. Zeitschrift fuer Pflanzenzuechtung (West Germany) : 50, (1963).

140. Lamm, R. Notes on a haploid potato hybrid. Hereditas 24, 391–396 (2010).

141. Caligari, P. D. S., Powell, W., Liddell, K., Maine, M. J. D. E. & Swan, G. E. L. Methods and strategies for detecting Solanum tuberosum dihaploids in interspecific crosses with S. phureja. Ann. Appl. Biol. 112, 323–328 (1988).

142. Hutten, R. C. B., E J M, Huigen, D. J., Th. Hermsen, J. G. & Jacobsen, E. Analysis of dihaploid induction and production ability and seed parent x pollinator interaction in potato. Euphytica vol. 72 61–64 (1994).

143. Peloquin, S. J. & Hougas, R. W. Decapitation and genetic markers as related to haploidy

inSolanum tuberosum. European Potato Journal vol. 2 176–183 (1959).

144. Hermsen, J. G. T. & Verdenius, J. Selection from Solanum tuberosum group phureja of genotypes combining high-frequency haploid induction with homozygosity for embryo-spot. Euphytica 22, 244–259 (1973).

145. van Breukelen, E., Ramanna, M. S. & Hermsen, J. G. T. Parthenogenetic monohaploids (2n= x= 12) from Solanum tuberosum L. and S. verrucosum Schlechtd. and the production of homozygous potato diploids. Euphytica (1977).

146. Uijtewaal, B. A., Huigen, D. J. & Th. Hermsen, J. G. Production of potato monohapioids (2n = x = 12) through prickle pollination. Theor. Appl. Genet. 73, 751–758 (1987).

147. Lough, R. C., Varrieur, J. M. & Veilleux, R. E. Selection inherent in monoploid derivation mechanisms for potato. Züchter Genet. Breed. Res. 103, 178–184 (2001).

148. Van Breukelen, E. W. M., Ramanna, M. S. & Hermsen, J. G. T. Monohaploids (n=x=12) from autotetraploid Solanum tuberosum (2n=4x=48) through two successive cycles of female parthenogenesis. Euphytica (1975).

149. Van Breukelen, E. W. M. Pseudogamic production of dihaploids and monoploids in Solanum tuberosum and some related species. (1981).

150. Singsit, C. & Hanneman, R. E. Haploids of tetraploid (2n=4x=48) Mexican potato species—Their extraction, cytology and crossability. Am. Potato J. 64, 469–482 (1987).

151. Clulow, S. A. et al. Cytological and molecular observations on Solanum phureja-induced dihaploid potatoes. Theor. Appl. Genet. 82, 545–551 (1991).

152. Clulow, S. A., Wilkinson, M. J. & Burch, L. R. Solanum phureja genes are expressed in the leaves and tubers of aneusomatic potato dihaploids. Euphytica 69, 1–6 (1993).

153. Clulow, S. A. & Rousselle-Bourgeois, F. Widespread introgression of Solatium phureja DNA in potato (S. tuberosum) dihaploids. Plant Breed. 116, 347–351 (1997).

154. PI 225682 GRIN-Global. https://web.archive.org/web/20210607182642/https://npgsweb.ars-

grin.gov/gringlobal/accessiondetail?id=1181411.

155. PI 225702 GRIN-Global. https://web.archive.org/web/20210607185000/https://npgsweb.ars-grin.gov/gringlobal/accessiondetail?id=1181427.

156. Cooper, D. C. & Rieman, G. H. Diploid plants in a seedling population of the cultivated potato. Am. Potato J. 35, 642–642 (1958).

157. Peloquin, S. J., Gabert, A. C. & Ortiz, R. Nature of ` Pollinator ' Effect in Potato (Solanum tuberosum L.) Haploid Production. Ann. Bot. 77, 539–542 (1996).

158. Wasilewicz-Flis, I. & Jakuczun, H. Induction and application of dihaploids of potato (Solanum tuberosum L.). Plant Breeding and Seed Science 52, 47–64 (2005).

159. Carroll, C. P. & Low, R. J. Aspects of male fertility in group Tuberosum dihaploids. Potato Res. 19, 109–121 (1976).

160. Carroll, C. P. & Low, R. J. Flowering behaviour and seed fertility in dihaploidSolanum tuberosum. Potato Res. 18, 416–427 (1975).

161. Buketova, M. I. Choice of pollinators for producing dihaploids. Kartofel'i Ovoshchi (1970).

162. Moyle, L. C. & Nakazato, T. Hybrid incompatibility 'snowballs' between Solanum species. Science 329, 1521–1523 (2010).

163. Rice, W. R. & Holland, B. The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. Behav. Ecol. Sociobiol. 41, 1–10 (1997).

164. Clark, N. L., Aagaard, J. E. & Swanson, W. J. Evolution of reproductive proteins from animals and plants. Reproduction 131, 11–22 (2006).

165. Willi, Y. The battle of the sexes over seed size: support for both kinship genomic imprinting and interlocus contest evolution. Am. Nat. 181, 787–798 (2013).

166. Ivanov, M. A. Experimental production of haploids in nicotiana rustical L. Genetica 20, 295–397 (1938).

167. Lindstrom, E. W. A HAPLOID MUTANT IN THE TOMATO*. Journal of Heredity vol. 20 23–

30 (1929).

168. Krishnaswamy, N. Cytological studies in a haploid plant of Triticum vulgare. Hereditas 25, 77–86 (1939).

169. Marks, G. E. A Polyhaploid Plant of Solanum polytrichon Rydb. Nature 175, 469–469 (1955).

170. Abdalla, M. M. F. & Th. Hermsen, J. G. Diploid parthenogenesis and androgenesis in diploid Solanum. Euphytica vol. 21 426–431 (1972).

171. Laurie, D. A. & Bennett, M. D. Wheat × maize hybridization. Canadian Journal of Genetics and Cytology vol. 28 313–316 (1986).

172. Camadro, E. L., Masuelli, R. W. & Cortés, M. C. Haploids of the wild tetraploid potato Solanum acaule ssp. acaule: generation, meiotic behavior, and electrophoretic pattern for the aspartate aminotransferase system. Genome vol. 35 431–435 (1992).

173. Inagaki, M. & Mujeeb-Kazi, A. Comparison of Polyhaploid Production Frequencies in Crosses of Hexaploid Wheat with Maize, Pearl Millet and Sorghuml. Japanese Journal of Breeding vol. 45 157–161 (1995).

174. Chaudhary, H. K., Sethi, G. S., Singh, S., Pratap, A. & Sharma, S. Efficient haploid induction in wheat by using pollen of Imperata cylindrica. Plant Breed. 124, 96–98 (2005).

175. Ren, J. et al. Novel technologies in doubled haploid line development. Plant Biotechnology Journal vol. 15 1361–1370 (2017).

176. Camadro, E. L., Erazzú, L. E., Maune, J. F. & Bedogni, M. C. A genetic approach to the species problem in wild potato. Plant Biol. 14, 543–554 (2012).

177. Huaman, Z. [Systematic botany, identification, distribution and evaluation of cultivated potato [Solanum tuberosum]]. Documento de Capacitacion (CIP) (1984).

178. Huamán, Z. & Spooner, D. M. Reclassification of landrace populations of cultivated potatoes (Solanum sect. Petota). Am. J. Bot. 89, 947–965 (2002).

179. Cersosimo, A. & Sorrentino, C. Haploidy in Nicotiana spp. in In Vitro Haploid Production in

Higher Plants: Volume 5 — Oil, Ornamental and Miscellaneous Plants (eds. Jain, S. M., Sopory, S. K. & Veilleux, R. E.) 135–163 (Springer Netherlands, 1997).

180. Burk, L. G., Gerstel, D. U. & Wernsman, E. A. Maternal Haploids of Nicotiana tabacum L. from Seed. Science 206, 585 (1979).

181. Satina, S., Rappaport, J. & Blakeslee, A. F. Ovular Tumors Connected with Incompatible Crosses in Datura. Am. J. Bot. 37, 576–586 (1950).

182. Brisson, D. Negative Frequency-Dependent Selection Is Frequently Confounding. Front Ecol Evol 6, (2018).

183. Barber, H. N. The suppression of meiosis and the origin of diplochromosomes. Proceedings of the Royal Society of London. Series B - Biological Sciences 128, 170–185 (1940).

184. Gastony, G. J. & Windham, M. D. Species Concepts in Pteridophytes: The Treatment and Definition of Agamosporous Species. American Fern Journal vol. 79 65 (1989).

185. Shimamura, T. Cytological Studies of Polyploidy Induced by Colchicine. CYTOLOGIA vol. 9 486–494 (1939).

186. Briggs, R. An analysis of the inactivation of the frog sperm nucleus by toluidine blue. J. Gen. Physiol. 35, 761–780 (1952).

187. Chimoyo, H. M. The effect of ultraviolet irradiation, toluidine blue, and environment on maternal haploid frequencies from the cross between Nicotiana tabacum and N. africana. (Clemson University, 1988).

188. Edwards, R. G. Colchicine-induced heteroploidy in early mouse embryos. Nature 174, 276–277 (1954).

189. Edwards, R. G. The experimental induction of pseudogamy in early mouse embryos. Experientia 10, 499–500 (1954).

190. Edwards, R. G. The experimental induction of gynogenesis in the mouse. III. Treatment of sperm with trypaflavine, toluidine blue, or nitrogen mustard. Proc. R. Soc. Lond. B Biol. Sci. 149, 117–129 (1958).

191. Illies, Z. M. Induction of haploid parthenogenesis in aspen by post-pollination treatment with Toluidine-blue. Silvae Genet. 23, 221–226 (1974).

192. NsokA, M. & KoNDo, Y. Production of gynogenetic haploids by treatment of spermatozoa with toluidine blue in Rana japonica. Sci. Rep. Lab. Amphibian Biol. Hiroshima Univ. 3, 385–398 (1978).

193. Stockert, J. C., Juarranz, A., Villanueva, A. & Cañete, M. Photodynamic damage to HeLa cell microtubules induced by thiazine dyes. Cancer Chemother. Pharmacol. 39, 167–169 (1996).

194. Bacellar, I. O. L. et al. Membrane damage efficiency of phenothiazinium photosensitizers. Photochem. Photobiol. 90, 801–813 (2014).

195. Paardekooper, M. et al. Photodynamic treatment of yeast cells with the dye toluidine blue: all-or-none loss of plasma membrane barrier properties. Biochim. Biophys. Acta 1108, 86–90 (1992).

196. Vázquez-Ortega, F. et al. Photodynamic treatment induced membrane cell damage in Acanthamoeba castellanii Neff. Dyes and Pigments vol. 180 108481 (2020).

197. Paardekooper, M., De Bruijne, A. W., Van Steveninck, J. & Van den Broek, P. J. Inhibition of transport systems in yeast by photodynamic treatment with toluidine blue. Biochim. Biophys. Acta 1151, 143–148 (1993).

198. Tuite, E. M. & Kelly, J. M. New trends in photobiology: Photochemical interactions of methylene blue and analogues with DNA and other biological substrates. J. Photochem. Photobiol. B 21, 103–124 (1993).

199. Sridharan, G. & Shankar, A. A. Toluidine blue: A review of its chemistry and clinical utility. J. Oral Maxillofac. Pathol. 16, 251–255 (2012).

200. Kato, A. Induction of bicellular pollen and dihaploidization of tetraploid maize. MAIZE GENETICS COOPERATION NEWSLETTER 70, 25–25 (1996).

201. Waugh, R., Baird, E. & Powell, W. The use of RAPD markers for the detection of gene

introgression in potato. Plant Cell Rep. 11, 466–469 (1992).

202. Wilkinson, M. J. et al. Evidence for somatic translocation during potato dihaploid induction. Heredity 74 (Pt 2), 146–151 (1995).

203. Straadt, I. K. & Rasmussen, O. S. AFLP analysis of Solanum phureja DNA introgressed into potato dihaploids. Plant Breed. 122, 352–356 (2003).

204. Ercolano, M. R. et al. Assessment of genetic variability of haploids extracted from tetraploid (2n = 4x = 48) Solanum tuberosum. Genome 47, 633–638 (2004).

205. Bartkiewicz, A. M. et al. Maximization of Markers Linked in Coupling for Tetraploid Potatoes via Monoparental Haploids. Front. Plant Sci. 9, 620 (2018).

206. Pham, G. M. et al. Genome-wide inference of somatic translocation events during potato dihaploid production. Plant Genome 12, 180079 (2019).

207. Marks, G. E. The enigma of triploid potatoes. Euphytica 15, 285–290 (1966).

208. Van Suchtelen, N. Triploids of the common potato,Solanum tuberosum L. Potato Research vol. 19 377–380 (1976).

209. Graebner, R. C., Chen, H., Contreras, R. N., Haynes, K. G. & Sathuvalli, V. Identification of the High Frequency of Triploid Potato Resulting from Tetraploid × Diploid Crosses. HortScience vol. 54 1159–1163 (2019).

210. Franklin-Tong, V. E. Self-Incompatibility in Flowering Plants: Evolution, Diversity, and Mechanisms. (Springer Science & Business Media, 2008).

211. Igic, B., Smith, W. A., Robertson, K. A., Schaal, B. A. & Kohn, J. R. Studies of self-incompatibility in wild tomatoes: I. S-allele diversity in Solanum chilense (Dun.) Reiche [corrected] (Solanaceae). Heredity 99, 553–561 (2007).

212. Dzidzienyo, D. K., Bryan, G. J., Wilde, G. & Robbins, T. P. Allelic diversity of S-RNase alleles in diploid potato species. Theor. Appl. Genet. 129, 1985–2001 (2016).

213. Newbigin, E., Anderson, M. A. & Clarke, A. E. Gametophytic Self-Incompatibility Systems. The Plant Cell vol. 5 1315 (1993).

214. Fujii, S., Kubo, K.-I. & Takayama, S. Non-self- and self-recognition models in plant self-incompatibility. Nat Plants 2, 16130 (2016).

215. Puul, S. 2n GAMETES IN DIPLOID SOLANUM: FREQUENCY AND TYPES OF SPINDLE ABNORMALITIES1. Can. J. Genet. Cytol. 24, 301–314 (1982).

216. Ramanna, M. S. A re-examination of the mechanisms of 2n gamete formation in potato and its implications for breeding. Euphytica 28, 537–561 (1979).

217. Carputo, D., Barone, A. & Frusciante, L. 2n gametes in the potato: essential ingredients for breeding and germplasm transfer. Theor. Appl. Genet. 101, 805–813 (2000).

218. Mable, B. K. Polyploidy and self-compatibility: is there an association? New Phytol. 162, 803–811 (2004).

219. Mendiburu, A. O. & Peloquin, S. J. Sexual polyploidization and depolyploidization: some terminology and definitions. Theoretical and Applied Genetics vol. 48 137–143 (1976).

220. Bretagnolle, F. & Thompson, J. D. Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants. New Phytol. 129, 1–22 (1995).

221. Iwanaga, M. & Peloquin, S. J. Origin and evolution of cultivated tetraploid potatoes via 2n gametes. Theor. Appl. Genet. 61, 161–169 (1982).

222. Quinn, A. A., Mok, D. W. S. & Peloquin, S. J. Distribution and significance of diplandroids among the diploid solanums. Am. Potato J. 51, 16–21 (1974).

223. Watanabe, K. & Peloquin, S. J. Cytological basis of 2n pollen formation in a wide range of 2x, 4x, and 6x taxa from tuber-bearing Solanum species. Genome 36, 8–13 (1993).

224. Stelly, D. M. & Peloquin, S. J. FORMATION OF 2N MEGAGAMETOPHYTES IN DIPLOID TUBER-BEARING SOLANUMS. American Journal of Botany vol. 73 1351–1363 (1986).

225. Camadro, E. L. & Peloquin, S. J. The occurrence and frequency of 2n pollen in three diploid solanums from Northwest Argentina. Theoretical and Applied Genetics vols 56-56 11–15 (1980).

226. Jongedijk, E. Desynapsis and FDR 2N-megaspore formation in diploid potato : potentials and limitations for breeding and for the induction of diplosporic apomixis. (Jongedijk, 1991).
227. Werner, J. E. & Peloquin, S. J. Occurrence and mechanisms of 2n egg formation in 2x potato. Genome vol. 34 975–982 (1991).
228. Correll, D. S. & Others. The potato and its wild relatives. The potato and its wild relatives. (1962).
229. Navarro, C. et al. Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. Nature 478, 119–122 (2011).
230. Abelenda, J. A., Navarro, C. & Prat, S. Flowering and tuberization: a tale of two nightshades. Trends Plant Sci. 19, 115–122 (2014).
231. Dnyansagar, V. R. & Cooper, D. C. Development of the Seed of Solanum phureja. Am. J. Bot. 47, 176–186 (1960).
232. Hanneman, R. E. & Peloquin, S. J. Croosability of 24-chromosome potato hybrids with 48-chromosome cultivars. European Potato Journal 10, 62–73 (1967).
233. Johnston, S. A. & Hanneman, R. E., Jr. The genetics of triploid formation and its relationship to endosperm balance number in potato. Genome 38, 60–67 (1995).
234. Masuelli, R. W. & Camadro, E. L. Crossability relationships among wild potato species with different ploidies and Endosperm Balance Numbers (EBN). Euphytica 94, 227–235 (1997).
235. Jackson, M. T., Rowe, P. R. & Hawkes, J. G. Crossability relationships of andean potato varieties of three ploidy levels. Euphytica 27, 541–551 (1978).
236. Tarn, T. R. & Hawkes, J. G. Cytogenetic studies and the occurrence of triploidy in the wild potato species Solanum commersonii Dun. Euphytica 35, 293–302 (1986).
237. Graebner, R. C., Chen, H., Contreras, R. N., Haynes, K. G. & Sathuvalli, V. Identification of the High Frequency of Triploid Potato Resulting from Tetraploid× Diploid Crosses. HortScience 54, 1159–1163 (2019).
238. Köhler, C., Mittelsten Scheid, O. & Erilova, A. The impact of the triploid block on the origin

and evolution of polyploid plants. Trends Genet. 26, 142–148 (2010).

239. Kolář, F., Čertner, M., Suda, J., Schönswetter, P. & Husband, B. C. Mixed-Ploidy Species: Progress and Opportunities in Polyploid Research. Trends Plant Sci. 22, 1041–1055 (2017).

240. Husband, B. C. The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations. Biol. J. Linn. Soc. Lond. 82, 537–546 (2004).

241. Alfano, F., Cammareri, M., Errico, A., Frusciante, L. & Conicella, C. 2n gametes inSolanum tuberosum dihaploids. Am. J. Potato Res. 76, 281–285 (1999).

242. Conicella, C., Genualdo, G., Errico, A., Frusciante, L. & Monti, L. M. Meiotic restitution mechanisms and 2n pollen formation in a Solanum tuberosum dihaploid and in dihaploid x wild species hybrids*. Plant Breeding vol. 115 157–161 (1996).

243. Sax, K. THE EXPERIMENTAL PRODUCTION OF POLYPLOIDY. J. Arnold Arbor. 17, 153–159 (1936).

244. Comai, L. The advantages and disadvantages of being polyploid. Nat. Rev. Genet. 6, 836–846 (2005).

245. Petit, C., Bretagnolle, F. & Felber, F. Evolutionary consequences of diploid–polyploid hybrid zones in wild species. Trends Ecol. Evol. 14, 306–311 (1999).

246. McHale, N. A. Environmental induction of high frequency 2n pollen formation in diploid Solanum. Can. J. Genet. Cytol. 25, 609–615 (1983).

247. Castro, M. O. Evolutionary ecology of polyploids: understanding species coexistence at the contact zones. (00500::Universidade de Coimbra, 2019).

248. Buggs, R. J. A. & Pannell, J. R. Ecological differentiation and diploid superiority across a moving ploidy contact zone. Evolution 61, 125–140 (2007).

249. Soltis, P. S., Marchant, D. B., Van de Peer, Y. & Soltis, D. E. Polyploidy and genome evolution in plants. Curr. Opin. Genet. Dev. 35, 119–125 (2015).

250. Levin, D. A. & Soltis, D. E. Factors promoting polyploid persistence and diversification and

limiting diploid speciation during the K–Pg interlude. Current Opinion in Plant Biology vol. 42 1–7 (2018).

251. Stebbins, G. L. Apomixis in the angiosperms. Bot. Rev. 7, 507–542 (1941).

252. de Wet, J. M. J. REVERSIBLE TETRAPLOIDY AS AN EVOLUTIONARY MECHANISM. Evolution 25, 545–548 (1971).

253. Husband, B. C., Baldwin, S. J. & Suda, J. The Incidence of Polyploidy in Natural Plant Populations: Major Patterns and Evolutionary Processes. in Plant Genome Diversity Volume 2: Physical Structure, Behaviour and Evolution of Plant Genomes (eds. Greilhuber, J., Dolezel, J. & Wendel, J. F.) 255–276 (Springer Vienna, 2013).

254. Hanzl, M., Kolář, F., Nováková, D. & Suda, J. Nonadaptive processes governing early stages of polyploid evolution: Insights from a primary contact zone of relict serpentine Knautia arvensis (Caprifoliaceae). Am. J. Bot. 101, 935–945 (2014).

255. Tandon, S. L. & Rao, G. R. Cytogenetical Investigations in Relation to the Mechanism of Evolution in Hexaploid Solanum nigrum L. Nature 201, 1348–1349 (1964).

256. Melo, C. A. F., Martins, M. I. G., Oliveira, M. B. M., Benko-Iseppon, A. M. & Carvalho, R. Karyotype analysis for diploid and polyploid species of the Solanum L. Osterr. bot. Z. 293, 227–235 (2011).

257. Hermsen, J. G. T. & Verdenius, J. Selection from Solanum tuberosum group phureja of genotypes combining high-frequency haploid induction with homozygosity for embryo-spot. Euphytica 22, 244–259 (1973).

258. Samitsu, Y. & Hosaka, K. Molecular marker analysis of 24- and 25-chromosome plants obtained from Solanum tuberosum L. subsp. andigena (2n = 4x = 48) pollinated with a Solanum phureja haploid inducer. Genome 45, 577–583 (2002).

259. Ordoñez, B. et al. PL-4 (CIP596131.4): an Improved Potato Haploid Inducer. Am. J. Potato Res. 98, 255–262 (2021).

260. Jansky, S. (2021).

261. Pallais, N. Osmotic priming of true potato seed: Effects of seed age. Potato Res. 32, 235–244 (1989).

262. D'Antonio, V. L. & McHale, N. A. Effect of storage temperature and extraction methods on dormancy and germination of true potato seed. Am. Potato J. 65, 573–581 (1988).

263. Simmonds, N. W. Experiments on the germination of potato seeds. I. European Potato Journal vol. 6 45–60 (1963).

264. Simmonds, N. W. Experiments on the germination of potato seeds. II. Europ. Potato J. 6, 69–76 (1963).

265. Lam, S.-L. & Erickson, H. T. Interaction of light and gibberellin on potato seed germination. Am. Potato J. 43, 443–449 (1966).

266. Lam, S.-L. Interaction of Temperature and Gibberellin on Potato Seed Germination. Am. J. Bot. 55, 193–198 (1968).

267. Rayburn, A. L., Auger, J. A., Benzinger, E. A. & Hepburn, A. G. Detection of Intraspecific DNA Content Variation in Zea mays L. by Flow Cytometry. J. Exp. Bot. 40, 1179–1183 (1989).

268. Bino, R. J., Lanteri, S., Verhoeven, H. A. & Kraak, H. L. Flow Cytometric Determination of Nuclear Replication Stage in Seed Tissues. Ann. Bot. 72, 181–187 (1993).

269. Ghislain, M., Zhang, D. P. ;. & Herrera, M. R. Molecular Biology Laboratory Protocols Plant Genotyping: Training Manual. (International Potato Center, 1999).

270. Pham, G. M. et al. Construction of a chromosome-scale long-read reference genome assembly for potato. Gigascience 9, (2020).

271. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv [q-bio.GN] (2013).

272. Danecek, P. et al. Twelve years of SAMtools and BCFtools. Gigascience 10, (2021).

273. Broman, K. W. & Sen, S. A Guide to QTL Mapping with R/qtl. (Springer, New York, NY, 2009).

274. Arends, D., Prins, P., Jansen, R. C. & Broman, K. W. R/qtl: high-throughput multiple QTL mapping. Bioinformatics 26, 2990–2992 (2010).

275. Broman, K. W., Wu, H., Sen, S. & Churchill, G. A. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19, 889–890 (2003).

276. Huo, H. et al. Rapid identification of lettuce seed germination mutants by bulked segregant analysis and whole genome sequencing. Plant J. 88, 345–360 (2016).

277. Michelmore, R. W., Paran, I. & Kesseli, R. V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. U. S. A. 88, 9828–9832 (1991).

278. Abe, A. et al. Genome sequencing reveals agronomically important loci in rice using MutMap. Nat. Biotechnol. 30, 174–178 (2012).

279. Team, R. C. & Others. R: A language and environment for statistical computing. (2013).

280. Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum Likelihood from Incomplete Data Via theEMAlgorithm. Journal of the Royal Statistical Society: Series B (Methodological) vol. 39 1–22 (1977).

281. Lander, E. S. & Botstein, D. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121, 185–199 (1989).

282. Haley, C. S. & Knott, S. A. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity vol. 69 315–324 (1992).

283. Chen, Z. The Full EM Algorithm for the MLEs of QTL Effects and Positions and Their Estimated Variances in Multiple-Interval Mapping. Biometrics vol. 61 474–480 (2005).

284. Ruden, D. et al. Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front. Genet. 3, 35 (2012).

285. Stelly, D. M. & Peloquin, S. J. Formation of 2n megagametophytes in diploid tuber-bearing Solanums. Am. J. Bot. (1986).

286. Kuppu, S. et al. A variety of changes, including CRISPR/Cas9-mediated deletions, in CENH3 lead to haploid induction on outcrossing. Plant Biotechnology Journal vol. 18 2068–2080 (2020).

287. Kuppu, S. et al. Point Mutations in Centromeric Histone Induce Post-zygotic Incompatibility and Uniparental Inheritance. PLoS Genet. 11, e1005494 (2015).

288. Azumi, Y. et al. Homolog interaction during meiotic prophase I in Arabidopsis requires the SOLO DANCERS gene encoding a novel cyclin-like protein. EMBO J. 21, 3081–3095 (2002).

289. Amundson, K. R. et al. Rare instances of haploid inducer DNA in potato dihaploids and ploidy-dependent genome instability. Plant Cell (2021) doi:10.1093/plcell/koab100.

290. Dodds, K. S. & Long, D. H. The inheritance of colour in diploid potatoes. J. Genet. 53, 136–149 (1955).

291. Jong, H. D. Inheritance of anthocyanin pigmentation in the cultivated potato: A critical review. Am. Potato J. 68, 585–593 (1991).

292. Salaman, R. N. The inheritance of colour and other characters in the potato. J. Genet. (1910).

293. Endelman, J. B. & Jansky, S. H. Genetic mapping with an inbred line-derived F2 population in potato. Theor. Appl. Genet. 129, 935–943 (2016).

294. Jung, C. S. Identities of the P and D loci in potato (Solanum tuberosum L.). (Cornell University, Jan, 2005).

295. Zhang, Y., Jung, C. S. & De Jong, W. S. Genetic analysis of pigmented tuber flesh in potato. Theor. Appl. Genet. 119, 143–150 (2009).

296. Jung, C. S. et al. The potato developer (D) locus encodes an R2R3 MYB transcription factor that regulates expression of multiple anthocyanin structural genes in tuber skin. Theor. Appl. Genet. 120, 45–57 (2009).

297. Ortiz, R. & M Golmirzaie, A. Genetic parameters for agronomic characteristics. I. Early and

intermediate breeding populations of true potato seed. Hereditas 139, 212–216 (2003).

298. Baskin, J. M. & Baskin, C. C. How much influence does the paternal parent have on seed germination? Seed Sci. Res. 29, 1–11 (2019).

299. De Jong, T. J., Van Dijk, H. & Klinkhamer, P. G. L. Hamilton's rule, imprinting and parent-offspring conflict over seed mass in partially selfing plants. J. Evol. Biol. 18, 676–682 (2005).

300. de Jong, T. J. & Scott, R. J. Parental conflict does not necessarily lead to the evolution of imprinting. Trends Plant Sci. 12, 439–443 (2007).

301. de Jong, T. J., Hermans, C. M. & van der Veen-van Wijk, K. C. A. M. Paternal effects on seed mass in Arabidopsis thaliana. Plant Biol. 13 Suppl 1, 71–77 (2011).

302. Raunsgard, A. et al. Intersexual conflict over seed size is stronger in more outcrossed populations of a mixed-mating plant. Proc. Natl. Acad. Sci. U. S. A. 115, 11561–11566 (2018).

303. van Eck, H. J., Jacobs, J. M. E., van den Berg, P. M. M. M., Stiekema, W. J. & Jacobsen, E. The inheritance of anthocyanin pigmentation in potato (Solanum tuberosum L.) and mapping of tuber skin colour loci using RFLPs. Heredity 73, 410–421 (1994).

304. Hazen, S. P. et al. Rapid array mapping of circadian clock and developmental mutations in Arabidopsis. Plant Physiol. 138, 990–997 (2005).

305. Tengkun, N., Dongdong, W., Xiaohui, M., Yue, C. & Qin, C. Analysis of Key Genes Involved in Potato Anthocyanin Biosynthesis Based on Genomics and Transcriptomics Data. Frontiers in Plant Science vol. 10 (2019).

306. Parra-Galindo, M. A., Soto-Sedano, J. C., Mosquera-Vásquez, T. & Roda, F. Pathway-based analysis of anthocyanin diversity in diploid potato. PLoS One 16, e0250861 (2021).

307. Jung, C. S. et al. The potato P locus codes for flavonoid 3',5'-hydroxylase. Züchter Genet. Breed. Res. 110, 269–275 (2005).

308. Jansky, S. H. et al. Reinventing potato as a diploid inbred line-based crop. Crop Sci. 56,

1412–1422 (2016).

309. Lindhout, P. et al. Towards F1 Hybrid Seed Potato Breeding. Potato Res. 54, 301–312 (2011).

310. Simmonds, N. W. Seed size in the cultivated potatoes. Euphytica 14, 143–152 (1965).

311. Arndt, G. C. & Peloquin, S. J. The identification and evaluation of hybrid plants among open pollinated true seed families. Am. Potato J. 67, 393–404 (1990).

312. d'Erfurth, I. et al. Turning meiosis into mitosis. PLoS Biol. 7, e1000124 (2009).