2017

Misregulation of Genomic Imprinting Drives Abnormal Seed Development in Hybrid Monkeyflowers (Mimulus)

Taliesin Kinser
College of William and Mary, tjkinser@email.wm.edu

Misregulation of Genomic Imprinting Drives Abnormal Seed Development in Hybrid Monkeyflowers (*Mimulus*)

Taliesin Joaquin Kinser

Dallas, Texas

Bachelor of Arts, Austin College, 2014

A Thesis presented to the Graduate Faculty of The College of William & Mary in Candidacy for the Degree of
Master of Science

Department of Biology

College of William & Mary
August 2017

APPROVAL PAGE

This Thesis is submitted in partial fulfillment of
the requirements for the degree of

Master of Science

Taliesin Joaquin Kinser

Approved by the Committee, May, 2017

Committee Chair
Assistant Professor Joshua Puzey, Biology
College of William & Mary

Associate Professor Diane Shakes, Biology
College of William & Mary

Assistant Professor Harmony Dalgleish, Biology
College of William & Mary

ABSTRACT

Genomic imprinting is the preferential expression of one allele over the other. It is an epigenetic phenomenon that occurs in the placentas of mammals and the endosperm of angiosperms. Endosperm, like placentas, is a nutrient rich tissue that supports the growing embryo within the seed. All grains are predominantly composed of this tissue. It is the product of a second fertilization event, resulting in both maternal and paternal alleles. Some alleles are regulated differentially, resulting in imprinted genes. There are both paternally expressed imprinted genes (PEGs) and maternally expressed imprinted genes (MEGs) in the endosperm. In general PEGs tend to have functions that induce the proliferation of endosperm (and the placenta in mammals) and MEGs tend to regulate or limit proliferation. There are many theories on the evolution of imprinting and parent-specific functions in such diverged taxa. Interploidy hybridization systems are often used to study these parent-specific effects. Such systems occur when a diploid is crossed with a polyploid, typically a tetraploid. By switching the parentage, parent specific genome dosage can be altered: if the tetraploid is the mother, then the offspring, or endosperm, has maternal genomic excess, and if the tetraploid is the father, then the endosperm has paternal excess. Maternal excess is typically characterized by endosperm underproliferation and paternal excess is characterized by endosperm overproliferation, as predicted by MEG and PEG functions. While significant progress has been made in genomic imprinting, there is still much unknown. For example, in plants, maternal excess is predicted to be more stable for evolutionary and functional reasons, yet there are many cases where the opposite occurs. By using a system in *Mimulus* that is both interploidy and interspecies (*M. guttatus* is diploid and *M. luteus* is tetraploid) and where paternal excess is favored in offspring viability, we aim to uncover further clues behind the mechanisms and evolutionary drivers of genomic imprinting. Here we show that the paternal excess hybrid suffers from endosperm underproliferation, opposite of what is predicted, and the maternal excess hybrid suffers from complete endosperm and embryo failure. We show that smaller endosperm results in failed or delayed germination. Furthermore, using genomic techniques, we show that *M. luteus* is genomically dominant in the hybrids regardless of crossing direction, likely interfering with imprinting patterns. We identify new PEGs involved in cellular proliferation. We show an overall paternal bias in *M. luteus*, which is unexpected and uncommon – potentially suggesting other adaptive drivers in imprinting. We suggest that abnormalities in the hybrids may be due to this genomic dominance and potentially other genetic and developmental differences between the two species that interferes with MEG and PEG roles.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This writer wishes to express his appreciation to Professor Joshua Puzey, under whose guidance this investigation was conducted, for his patience, guidance and criticism throughout the investigation. The author is also indebted to Professors Diane Shakes and Harmony Dalgleish for their careful reading and criticism of the manuscript.

Furthermore, the author would like to express his gratitude to Sarah Meier, Colleen Flynn, Nora Flynn, and Scott Teresi for their technical support in the laboratory; to Abigail Reft for her training and support for the histology of this project; to Renee Peace and Charnel Thornton for their official work; to Professor Helen Murphy for her sharing of materials; and to Ron Smith for his significant contribution to the genomics of this project.

This M.S. is dedicated to Dr. Lina Abi Mosleh, who gave me a chance when no one else would, who saw in me what others did not, who patiently guided me beyond what I could envision for myself, and who never once lost faith, confidence in and support for me.

LIST OF TABLES

LIST OF FIGURES

Misregulation of genomic imprinting drives abnormal seed development in hybrid monkeyflowers (*Mimulus*)

Taliesin J. Kinser

Advisor: Joshua Puzey

May 2017

1 Introduction

Plant productivity provides the base of terrestrial ecosystems, introducing chemical energy through photosynthesis to all organisms within these systems [Ågren and Andersson, 2011]. Angiosperms (flowering plants) are the most specious of green plants [Crepet and Niklas, 2009], following their widespread adaptive radiation during the Cretaceous [Berendse and Scheffer, 2009], and thus have become critical drivers in terrestrial ecosystems. Furthermore, almost all agricultural and horticultural crops are angiosperms [Forbes and Watson, 1992]. While there are many reasons and theories attributed to the overwhelming dominance of angiosperms [Crepet and Niklas, 2009], the evolution of the angiosperm seed has provided clear advantages. These include release from the maternal plant at a relatively advanced stage in development with food stored from the parent and the ability to retain dormancy to germinate when conditions are favorable [Harper et al., 1970].

The angiosperm seed can be divided into three genetically distinct tissues. The seed coat is diploid maternal tissue that fully encompasses the egg. The diploid embryo is the product of fertilization between a sperm cell delivered by the pollen grain and the egg within the seed coat. Finally, the triploid endosperm is the product of a second fertilization event between an additional sperm cell and two polar nuclei within the seed coat that are genetically identical to the egg. Nutrients from the maternal plant are allocated into the endosperm, which is the nutritive tissue that supports the development of the embryo. The development of the endosperm proceeds the majority of embryo development and is of vital importance to the viability of a seed [Kozlowski, 1972]. Gymnosperm seeds also contain a nutritive tissue, but instead of sexual endosperm, it is the maternal megagame-

tophyte [Baroux et al., 2002]. The sexual endosperm is an angiosperm-specific tissue, a major distinguisher from other seed plants [Haig and Westoby, 1989], and may provide certain advantages over the female gametophyte. These advantages may include control of resource allocation from both maternal and paternal genomes and a measure against wasting resources by allocating them only to fertilized seeds [Baroux et al., 2002]. An analogous organ has been compared in therian mammals: the placenta [Harper et al., 1970]. While angiosperms possess an organ named the placenta, which directs sustenance through vascular bundles from the ovary to the ovule [Endress, 2011], the endosperm shares some striking similarities with the mammalian placenta. Both shuttle nutrients from the maternal progenitor to the growing embryo proper, both have vital hormone signaling and influence on embryo size and development, both have paternal alleles, and both exhibit genomic imprinting [Garnica and Chan, 1996, Monk, 2015, Gehring and Satyaki, 2017]

Genomic imprinting is an epigenetic phenomenon where alleles are regulated differentially depending on their parent of origin. One manifestation of this regulation is parent-specific patterns of gene expression. RNA and protein production from an imprinted gene is contributed predominately by **one of** its alleles, the maternal or the paternal allele [Kinoshita, 2007]. Expression of the given gene can either be differential between the two alleles (one allele is expressed more than the other) or fully monoallelic (expression is only from one allele) [Gehring and Satyaki, 2017, Monk, 2015]. Genomic imprinting is perhaps most easily understood through studies on its prevention of parthenogensis in mammals [McGrath and Solter, 1984]. In the mouse, parthenogenesis of the oocyte is prevented due to repressive methylation in critical development genes. In a bi-maternal mouse oocyte, development will eventually fail. However, if the unmethylated paternal allele from a sperm cell is present, or if one of the bi-maternal alleles is experimentally demethylated, then

full development can occur [Kawahara et al., 2007]. While parthenogenesis is certainly affected by genomic imprinting, at least in mammals, it is not often considered an evolutionary driver of imprinting. There are many theories behind the evolution of genomic imprinting in angiosperms and placental mammals, but the most common one, particularly in angiosperms, is the kin conflict model [Spencer and Clark, 2014].

Under the kin conflict (or parental-conflict) theory, conflict arises between the mother and offspring, and between siblings. Selection favors mothers to distribute resources among their offspring to maximize the number of viable offspring [Haig, 2013]. Offspring, however, are less related to their siblings, particularly in a multiple mating system, and therefore are selected to allocate more resources than their siblings. Therefore, it is predicted that the paternal allele should pull for resources into the fertilized seed and the maternal allele should regulate this [Haig and Westoby, 1989]. The peculiar two maternal to one paternal genome in the endosperm has been argued to have evolved in response to this kin conflict and genomic imprinting [Stewart-Cox et al., 2004], and that it may allow for greater maternal control of resource distribution among seeds [Westoby and Rice, 1982]. Departures from this 2m:1p parental genome dosage results in abnormal endosperm development likely due to misregulation of genomic imprinting [Haig and Westoby, 1991].

A common method of studying parent-of-origin effects in the endosperm is by altering the genome dosage of a specific parent through interploidy hybridization. Two parents of different ploidy levels (often a diploid and a tetraploid) are crossed reciprocally so that offspring either have a greater number of maternal alleles or a greater number of paternal alleles in the endosperm. Such crosses are performed either by naturally occurring diploid and tetraploid variants of a given species [Scott et al., 1998] or with experimentally induced polyploids [Li and Dickinson, 2009]. By crossing a diploid maternal progenitor with

a tetraploid paternal progenitor, the genome dosage in the endosperm becomes 2m:2p, a scenario coined as paternal excess. The opposite cross provides a 4m:1p ratio and is considered maternal excess [Haig and Westoby, 1991]. Genomic imprinting is characterized by specific maternal excess and paternal excess traits [Scott et al., 1998].

By altering parent specific genome dosage, certain maternal and paternal roles in endosperm development have been observed. Increasing the dosage of maternal alleles results in earlier endosperm differentiation and decreased proliferation producing less tissue, and a greater paternal dosage results in later differentiation and increased proliferation, thus increasing the seed size [Scott et al., 1998]. In maize such misregulation in imprinting is highly fatal to the misregulated seed [Pennington et al., 2008]. A paternally expressed imprinted gene has been found in *Arabidopsis* that reduces seed fatality of paternal excess crosses when deleted [Kradolfer et al., 2013]. Similar phenotypes have been observed in the placenta when parental genome dosage is altered. Cases of abnormal fertilization that result in complete (0m:2p) or incomplete (1m:2p) hydatiform moles (resulting respectively from fertilization of an egg that lost DNA and from an egg fertilized by two sperm cells) and in digynic triploids (unreduced egg cell is fertilized - 2m:1p) reveal similar abnormal growth patterns. A lack of fetal development and excessive trophoblast (part of embryo where placenta is derived from) growth is characterized by the hydatiform moles, and an abnormally small placenta forms from digynic triploids. Such morphologies suggest a placental growth promoting role of paternal alleles and a growth inhibiting role of maternal alleles [Hoffner and Surti, 2012].

Further evidence for a division of roles between maternally and paternally expressed alleles is found in functional genetic studies of imprinted genes. While the function or relevance of many maternally expressed genes (MEGs) and paternally expressed genes

(PEGs) is unknown or inconclusive, a number of genes have revealed specific roles. In plants many MEGs and PEGs have been associated with epigenetic regulators. Some PEGs have been associated with inducing repressive regulation of genes and some MEGs maintain imprinted expression during endosperm development along with other epigenetic regulation [Gehring and Satyaki, 2017]. Certain PEGs are responsible for auxin production and thus growth of the endosperm. Others have also been associated with regulating cytokinesis [Wang and Köhler, 2017]. The same epigenetic regulatory MEGs that maintain imprinted expression thus prevent endosperm proliferation [Gehring and Satyaki, 2017]. In mammals the MEG, Phlda2, regulates glycogen storage and placental growth. Decreased dosage results in placental overgrowth [Frank et al., 2002], and increased dosage reduces glycogen storage and size of the junctional zone (where maternal and embryonic portion of the placenta meet) [Tunster et al., 2010]. Overall, MEGs are predicted to reduce or regulate endosperm proliferation, and PEGs are predicted to promote it [Gehring and Satyaki, 2017].

There are multiple mechanisms that produce the imprinted status of MEGs and PEGs. These mechanisms can be both pre-zygotic and post-zygotic. Protein coding genes, non-coding RNA, and transposable elements (TEs) can all be imprinted [Gehring and Satyaki, 2017]. The precursor to the endosperm within the female megagametophyte, the central cell, undergoes cytosine demethylation by the plant DNA glycosylase, DME (and other similar enzymes or DME orthologs) [Wang and Köhler, 2017]. This results in large regions of maternal euchromatin in the endosperm. The sperm cell also has low levels of CHH methylation and a large number of 21-22 nucleotide siRNAs, which may invoke RNA-directed DNA methylation (RdDM) [Wang and Köhler, 2017]. Many imprinted genes are associated with differentially methylated regions (DMRs). The hypomethylation of the ma-

ternal allele in the endosperm can result in MEGs. Some MEGs are the product of RdDM on or nearby the paternal allele. In contrast, many PEGs are methylated on the paternal allele, thus protecting them from repressive H3K27me3 histone modification that will silence the hypomethylated maternal allele (intriguingly the molecular machinery that produces this histone modification includes two MEGs) [Gehring and Satyaki, 2017]. PEGs in some species are also associated with hypermethylation of the maternal allele [Klosinska et al., 2016]. The large number of 21-22 nucleotide siRNAs in the sperm cell may repress TEs postfertilization, though their exact role is still largely unknown [Martinez and Köhler, 2017]. Imprinted genes in the mammalian placenta are also largely driven by DMRs and repressive histone modifications (which can be induced my non-coding RNA) [Monk, 2015]. Imprinted regions and the mechanisms that imprint them can vary quite drastically between species in plants and mammals [Gehring and Satyaki, 2017, Monk, 2015]. A note of interest is that, at least in plants, PEGs tend to be more conserved among species, and MEGs tend to contain more endosperm-specific expressed genes [Klosinska et al., 2016, Waters et al., 2013]. Perhaps this demonstrates a general role of cellular growth and proliferation from PEGs and a regulatory role of MEGs depending on species specific epigenetic mechanisms in the endosperm.

Interspecies hybridization also results in endosperm abnormalities often associated with misregulation of genomic imprinting. These abnormalities may be caused by rapidly derived differences in epigenetic regulation and imprinting. Hybridization between recently diverged outcrossing and self-fertilizing species produce similar phenotypes as observed in interploidy crosses. Hybrids with the outcrosser as the maternal progenitor suffer from under-proliferated endosperm, while those with the outcrosser as the paternal progenitor suffer from a failure of cytokinesis or over-proliferated endosperm, presumably because

outcrossers are more susceptible to parent-offspring conflict and have stronger imprinting mechanisms [Rebernig et al., 2015, Josefsson et al., 2006]. In other cases, systematic shifts in gene expression or incompatibilities of imprinted loci appear to enforce hybrid seed failure, suggesting species specific divergence of imprinting and regulatory mechanisms [Florez-Rueda et al., 2016, Wolff et al., 2015, Garner et al., 2016]

The kin conflict theory is supported by failures in crosses between species with different mating strategies (outcrossing and self-fertilizing) that exhibit interploidy endosperm abnormalities. Certain imprinted genes also support this theory. p57*kip2* is an MEG and a negative regulator of cell proliferation in the placenta. Lack of expression in this MEG can induce preeclampsia-like symptoms in the mother, a pregnancy complication in which blood pressure is elevated and organ damage occurs [Kanayama et al., 2002]. Additionally, some placenta specific paternally derived miRNAs are associated with the expansion of the embryonic portion of the placenta into the maternal portion, and increased circulation of them in the maternal blood system has been linked to preeclampsia [Xie et al., 2014, Hromadnikova et al., 2013]. However other scenarios such as the seemingly random epigenetic incompatibilities between species or paternally induced imprinting on its own allele [Monk, 2015] may fit more closely to other theories on the evolution of imprinting such as adaptive advantage of a sex-specific allele or co-adaptation of the offspring to the mother to match her physiology [Patten et al., 2014].

Clearly there is a great deal unknown about genomic imprinting and are likely many mechanisms and avenues for it to evolve. Areas of further research include understanding functional roles and significance of imprinted genes and the selection pressures, if any, driving them, the convergence and conservation of imprinted loci among species and mechanisms differentiating them, the role of genome dosage in traits of the develop-

ing endosperm, and distinguishing pre-zygotic and post-zygotic mechanisms of imprinting, among others [Gehring and Satyaki, 2017, Wang and Köhler, 2017].

To approach some of these questions, we use a recent, natural interploidy hybridization system between two diverged species in *Mimulus* that exhibits asymmetry in hybrid seed viability depending on crossing direction [Vallejo-Marin et al., 2016]. *M. guttatus* is diploid and is widespread throughout the North American west coast [Oneal et al., 2014]. *M. luteus* is a tetraploid native to Andean Chile [Medel et al., 2003]. Each species was separately introduced to the United Kingdom as ornamentals and subsequently naturalized in the region. Hybrids were first recognized as early as 1872. While sterile, the triploid hybrid, *M. robertsii*, is highly clonal and widespread across the UK [Vallejo-Marin and Lye, 2013]. It was recently determined that hybridization between the two species has occurred many times and predominantly in the direction of *M. guttatus* as the maternal progenitor [Vallejo-Marin et al., 2016]. A fertile and naturally reproducing hexaploid, *M. perigrinus*, derived by multiple occasions of whole genome duplication in the triploid hybrid has also recently been described [Vallejo-Marin, 2012]. It was originally postulated that asymmetry of seed viability in interploidy crosses favors a maternal progenitor of higher ploidy [Ramsey and Schemske, 1998], though a more recent survey of the literature finds comparable frequencies of paternal progenitors with higher ploidy, as seen in this system [Vallejo-Marin et al., 2016].

Here, we investigate the morphological and genetic drivers behind this observed asymmetry as well as genomic imprinting patterns within *M. luteus* - the first imprinting study on a *Mimulus* species. We use electron microscopy, histology, and image analysis software to elucidate the development of the parent species and reciprocal hybrid crosses. Through germination experiments, we link developmental abnormalities to seed viability.

Using RNA-seq we investigate transcriptome and potential genomic imprinting patterns and incompatibilities within the reciprocal hybrid crosses, as well as identify patterns in *M. luteus*. We identify endosperm abnormalities that differ from other interploidy and interspecific hybridization studies and unusual patterns of imprinting. To the best of our knowledge, this is the first study on genomic imprinting in a hybrid system that is both interploidy and interspecific.

2 Results

2.1 Seed morphology and germination rate by cross direction

In order to test for abnormalities in seed development of the reciprocal hybrids, four crosses were set up using *M. luteus* and *M. guttatus* plants. Seeds were produced for each of the parents and for the two interploid hybrid crosses. Crosses are always denoted as seed parent x pollen donor. The two reciprocal hybrid crosses are *M. guttatus* x *M. luteus* (denoted as 2n x 4n) and *M. luteus* x *M. guttatus* (denoted as 4n x 2n). Seeds were characterized through area and aspect ratio of mature seeds, SEM imaging of mature seeds, histology of developing seeds, and quantitative measurements of histology.

2.1.1 2n x 2n

M. guttatus seeds are relatively large, round, and plump (Fig. 3). They have a mean area (Fig. 1) of 0.146 +/- 0.003 (n =314), and are the roundest of the four crosses (Fig. 2) with a mean aspect ratio of 1.589 +/- 0.026 (1 is a perfect circle). While some seeds are smaller, oblong, and more shriveled, the majority are smooth and spherical. Additionally, they have a high probability of germination. Of the seeds we planted on filter paper, 98.7% germinated. We considered germination as the full penetration of the radical through the

seed coat. We classify normal, parent seeds with this round and plump morphology.

2.1.2 4n x 4n

Mature seeds of *M. luteus* have a very similar morphology to those of *M. guttatus* (Fig. 3). They are indistinguishable in size (0.143 +/- 0.005; Fig. 1), though are not quite as round (1.701 +/- 0.034; n = 187; Fig. 2). Additionally, they have a somewhat lower germination rate (86.1%) and take longer to germinate, however, the lower germination rate may be due to insufficient time in the experiment. Again, we consider these parent seeds to also have the round and plump morphology.

2.1.3 2n x 4n

The viable hybrid seeds with *M. guttatus* as the maternal progenitor are smaller, less round, and more shriveled. They are smaller than each of the parent seeds with an area of 0.123 +/- 0.004 (Fig. 1) and are the least round of all the crosses (Fig. 2) with an aspect ratio of 1.859 +/- 0.029 (n = 253). Morphologically, they appear oblong, ridged, constricted, and contorted (Fig. 3). While still fairly successful, these seeds have a lower germination rate than the parents (70%). We classify these seeds with a shrivelled morphology.

2.1.4 4n x 2n

The inviable hybrid seeds with *M. luteus* as the maternal progenitor are the smallest of all the crosses (Fig. 1). They have a mean area of 0.108 +/- 0.003. While rounder than the 2n x 4n crosses and sharing a similar aspect ratio to *M. luteus* (1.674 +/- 0.026; Fig. 2), their qualitative morphology reveals more insight into the character of these seeds. On closer examination, the seeds appear thin and cup shaped, with the vast majority of seed tissue belonging to the seed coat (Fig. 3). There is little to no endosperm or embryo tissue or even

open space beneath the seed coat. We found no seeds that germinated (n = 316), though other studies have seen a small percentage of germinates [Vallejo-Marin et al., 2016]. We classify these seeds with a flattened morphology.

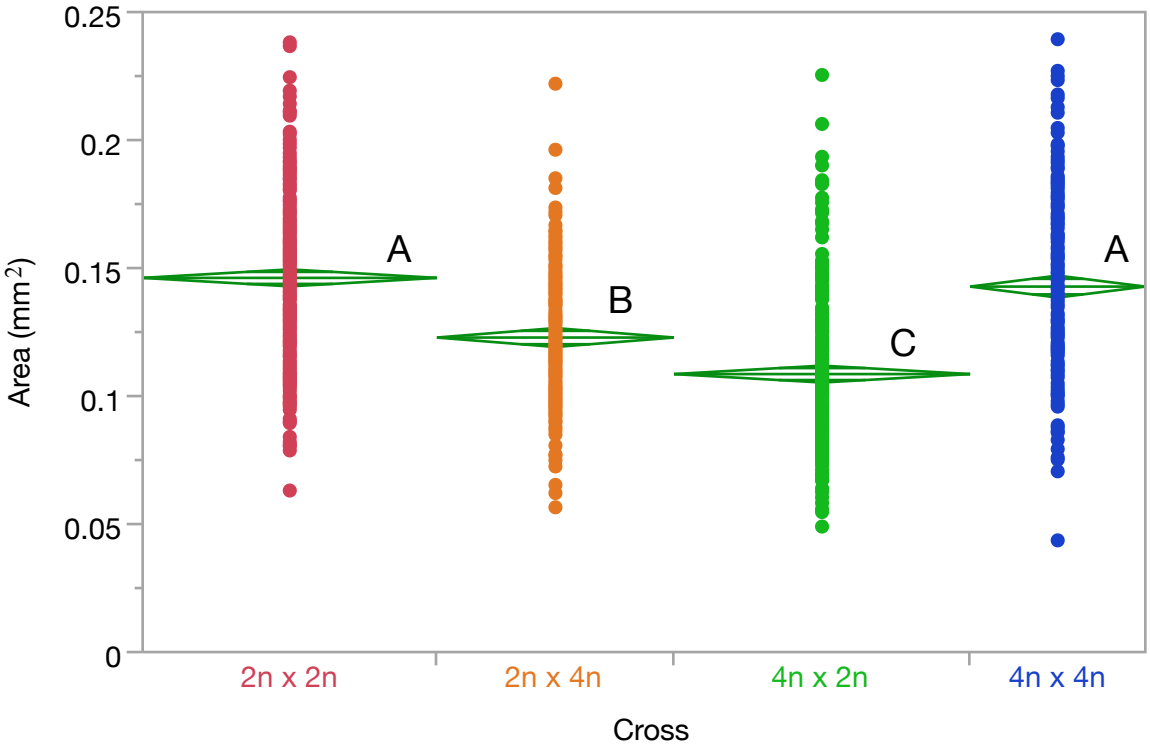

Figure 1: **Area of seeds**. Area of seeds in each cross were measured. Letters represent Tukey-Kramer results following an ANOVA. Groups with different letters are statistically different in area.

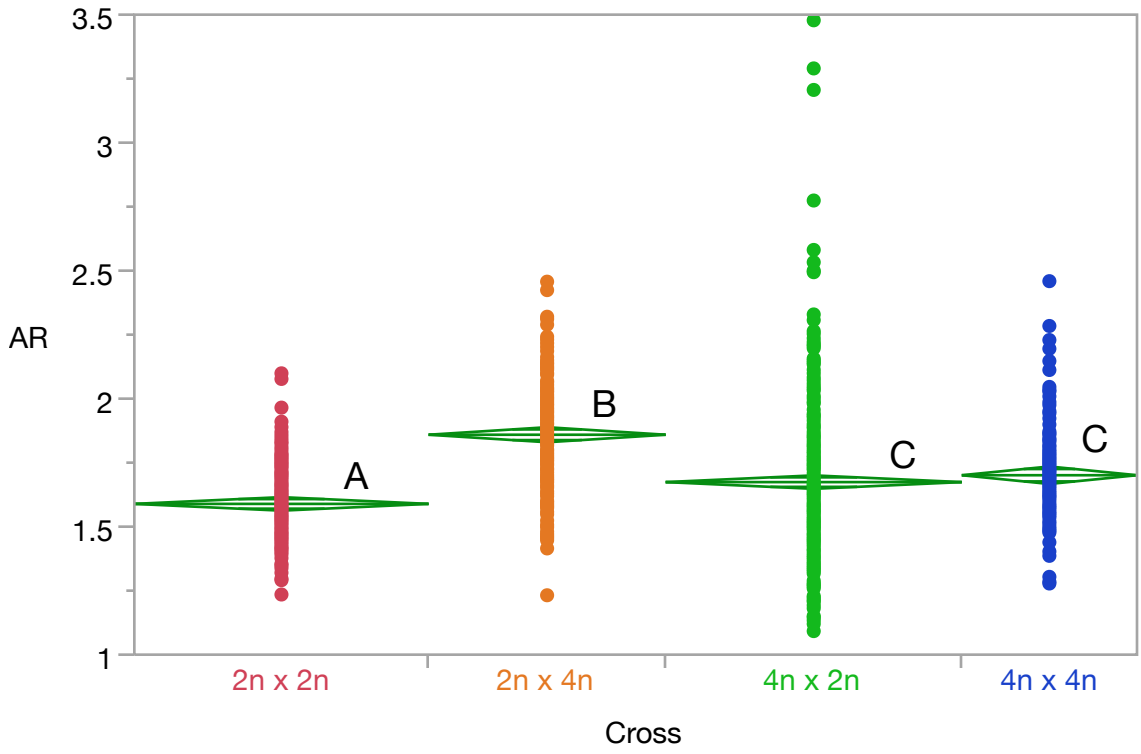

Figure 2: **Aspect ratio of seeds**. Aspect ratio of seeds in each cross were measured. 1 is a perfect circle. Letters represent Tukey-Kramer results following an ANOVA. Groups with different letters are statistically different in aspect ratio.

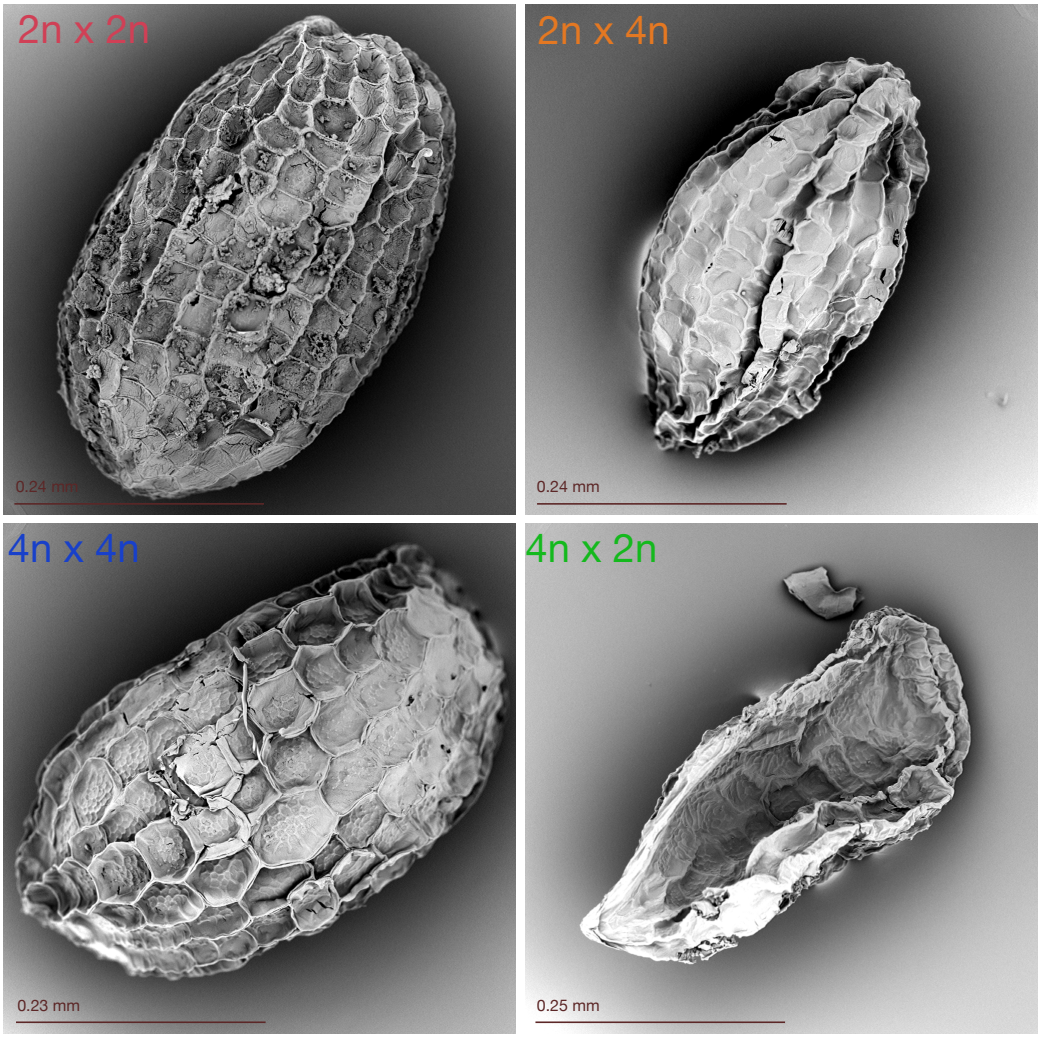

Figure 3: **SEM images of seeds**. Mature seeds were collected from each of the crosses between 15 and 18 DAP (when fruit dehisces). Images were taken using a scanning electron microscope. Seeds of the parents, *M. guttatus* (2n x 2n) and *M. luteus* (4n x 4n), have a round and plump morphology. The viable hybrid seeds (2n x 4n) are smaller with a shriveled seed coat, and the inviable hybrid seeds (4n x 2n) are small and thin, forming a bowl shape.

2.2 Endosperm and embryo development are abnormal in reciprocal hybrids

After observing the external morphologies of the seeds, we aimed to identify their internal morphologies to understand their underlying abnormalities. Using seed histology, we produced images throughout the development of the parent and reciprocal hybrid crosses, which are provided in Fig. 4. Images are color coded by tissue type and qualitative descriptions are provided below. Quantitative measurements were made from histological data, and four ANOVAs with post-hoc Tukey-Kramer tests were performed: embryo area among the four crosses at 8 DAP, embryo area at 11 DAP, endosperm area at 8 DAP, and endosperm area at 11 DAP (Fig. 5). A table of p-values for individual two sample t-tests between the embryo or endosperm area of each cross at 8 DAP and its area at 11 DAP are provided in Table 1 to demonstrate growth or lack of growth of the tissue types over time.

2.3 Description of seed histology

Refer to Fig 4: At 3 DAP, the central cell has gone through some devisions in all four crosses. Embryos are not visible in these images. By 8 DAP, the embryos of *M. guttatus* and 2n x 4n have entered the globular stage. The endosperm is well proliferated in *M. guttatus*, but 2n x 4n is beginning to reveal signs of endosperm underproliferation (it potentially is at 3 DAP as well). Embryos of *M. luteus* and 4n x 2n have not entered the globular stage, and, while *M. luteus* has less proliferated endosperm than *M. guttatus*, 4n x 2n is showing endosperm deficiencies as well. At 8 DAP, *M. guttatus* and 2n x 4n embryos are well into the heart stage and entering the torpedo stage. Endosperm underproliferation is still clear in 2n x 4n as compared to *M. guttatus*. *M. luteus* has well proliferated endosperm and the embryo is in the globular stage. 4n x 2n is small, does not appear to have entered

the globular stage, and has abnormal, less cellularized and compact endosperm. Finally, by 11 DAP, the embryos of *M. guttatus*, 2n x 4n, and *M. luteus* are all in the torpedo stage. The whole seed, endosperm, and embryo are clearly smaller in 2n x 4n than *M. guttatus*. Endosperm and embryo tissues also appear to be more crowded and compacted. *M. luteus* still seems to be slower developing than *M. guttatus*, though this is not verified quantitatively. Endosperm can be seen entering the cotyledons in this image. 4n x 2n has negligible endosperm development, and the embryo does not appear any further developed than it was at 8 DAP. The majority of the seed is composed of maternal seed coat tissue. It should be noted that only images with visible embryos for 4n x 2n are shown, and the phenotype of abnormal and negligible endosperm is even more dramatic in many other images. Histology results of *M. guttatus* seed development closely correspond to confocal images and analyses produced by [Oneal et al., 2016].

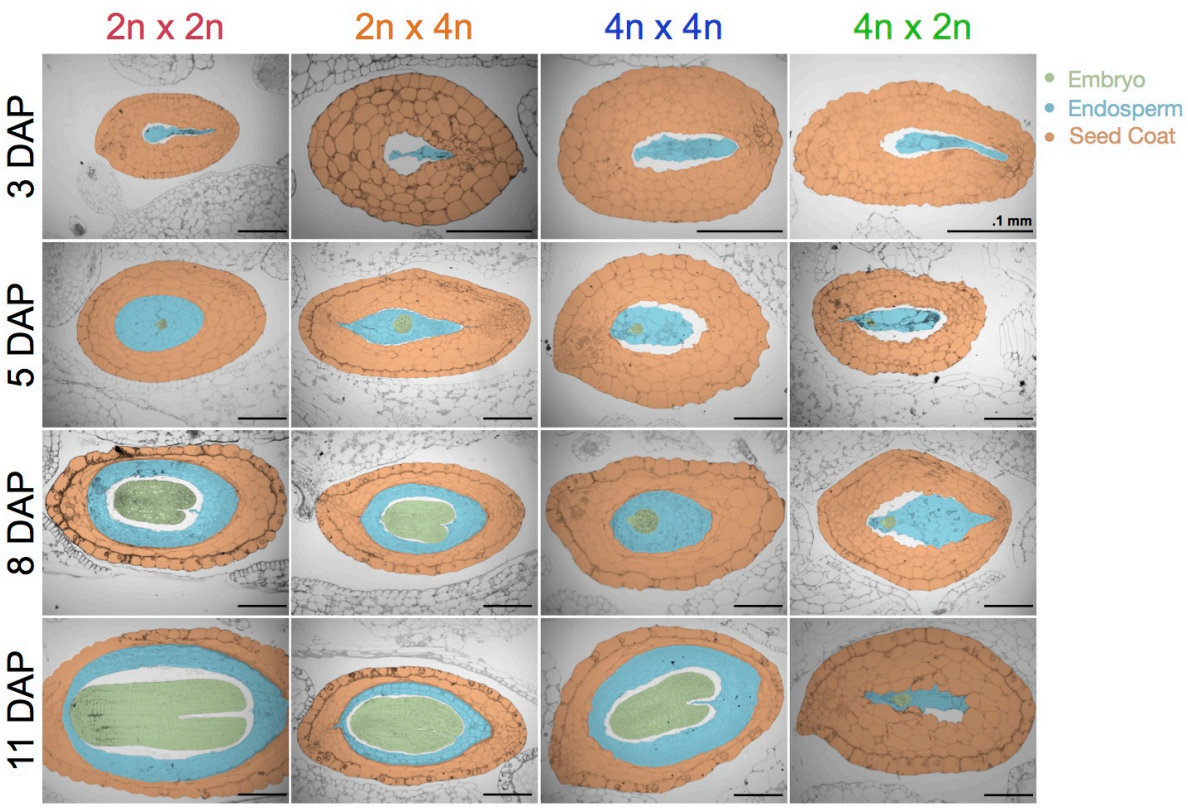

Figure 4: **Development of parent and reciprocal hybrid seeds**. Sections were made of each cross through a developmental progression of 3, 5, 8, and 11 DAP. Crosses (*M. guttatus* - 2n x 2n; 2n x 4n; *M. luteus* - 4n x 4n; 4n x 2n) are displayed in columns and the days of development are displayed in rows.

2.3.1 Endosperm failure and embryo arrest in inviable maternal excess (4n x 2n)

The embryo of the 4n x 2n hybrid is arrested by 8 DAP, showing no further growth at 11 DAP (Table 1). It is similar in growth and size to its maternal progenitor, *M. luteus*, at 8 DAP, and is smaller than *M. guttatus* and the 2n x 4n hybrid (Fig. 5a). By 11 DAP, the embryos of all the other crosses have continued to grow and are significantly larger than the 4n x 2n embryo (Fig. 5b). Its endosperm is similar in size to 2n x 4n and *M. luteus* at 8 DAP but is significantly smaller than *M. guttatus* (Fig. 5c). No further growth is detectable

by 11 DAP (Table 1; Fig. 5d). Qualitatively, the 4n x 2n endosperm exhibits even more abnormalities than its failed growth. Endosperm tissue is either uncellularized or in large, non-condensed cells (Fig. 4). Issues with both endosperm proliferation and cellularization may have occured, thus resulting in aborted embryos. This data alone, however, does not rule out issues with the embryo as well.

2.3.2 Endosperm under-proliferation and immature embryo in viable paternal excess (2n x 4n)

In the 2n x 4n hybrid, the embryo follows the developmental timing of its maternal progenitor, *M. guttatus* (Fig. 5a), but growth is slowed by 11 DAP (it is still significantly larger than it was at 8 DAP, Table 1). The *M. luteus* embryo appears to develop at a slower rate than *M. gutatus* (Fig. 5a) but catches up to 2n x 4n at 11 DAP. 2n x 4n's embryo is still significantly larger than that of the 4n x 2n embryo at this point (Fig. 5b). While its endosperm is similar in size to the other crosses at 8 DAP (Fig. 5c), it does not continue to expand (Table 1) and is as small as the 4n x 2n endosperm at 11 DAP (Fig. 5d). It likely is still growing, but at a rate slower or similar to that of which it is absorbed by the cotyledons. The slow growth of the endosperm may drive the stunted growth of the embryo seen at 11 DAP (Fig. 5b). At the torpedo stage, the embryo is no longer attached to the suspensor and is now fully dependent on the endosperm for further growth. The embryo may not receive enough nutrients from the endosperm to grow to the size of the *M. guttatus* embryo. Furthermore, endosperm cells appear to be more densely packed and compartmentalized than those of the parent's endosperm at 11 DAP (Fig. 4). The decreased rate of growth and compact cells match the phenotype of endosperm under-proliferation typically described in the maternal excess direction [Scott et al., 1998]. Endosperm under-proliferation may result

in the decreased germination rates observed in the 2n x 4n cross (described above).

2.3.3 Endosperm and embryo morphologies explain the three classes of seed morphology

The three classes of external morphologies categorized among the parental and hybrid crosses (Fig. 3) match the internal characters we have described (Fig. 4). *M. guttatus* and *M. luteus* have round and plump seeds with large and full embryo and endosperm. The viable hybrid, 2n x 4n, has shriveled seeds with constricted, condensed endosperm and subsequent stunted embryonic growth. The inviable hybrid, 4n x 2n, has flat seeds with very little (if any) abnormal endosperm and aborted pre-globular embryos. From these observations, we can link characteristics observed based off of external morphology to the phenotype of the endosperm and embryo.

Cross	Embryo	Endosperm
2n x 2n	0.0016	ns
2n x 4n	0.0231	ns
4n x 4n	0.0013	0.034
4n x 2n	ns	ns

Table 1: **T-tests for embryo and endosperm growth**. P-values are provided from t-tests comparing embryo and endosperm size at 8 DAP to 11 DAP for each cross. A significant p-value represents a significant shift in size from 8 to 11 DAP for the given tissue. Alpha is set at 0.05.

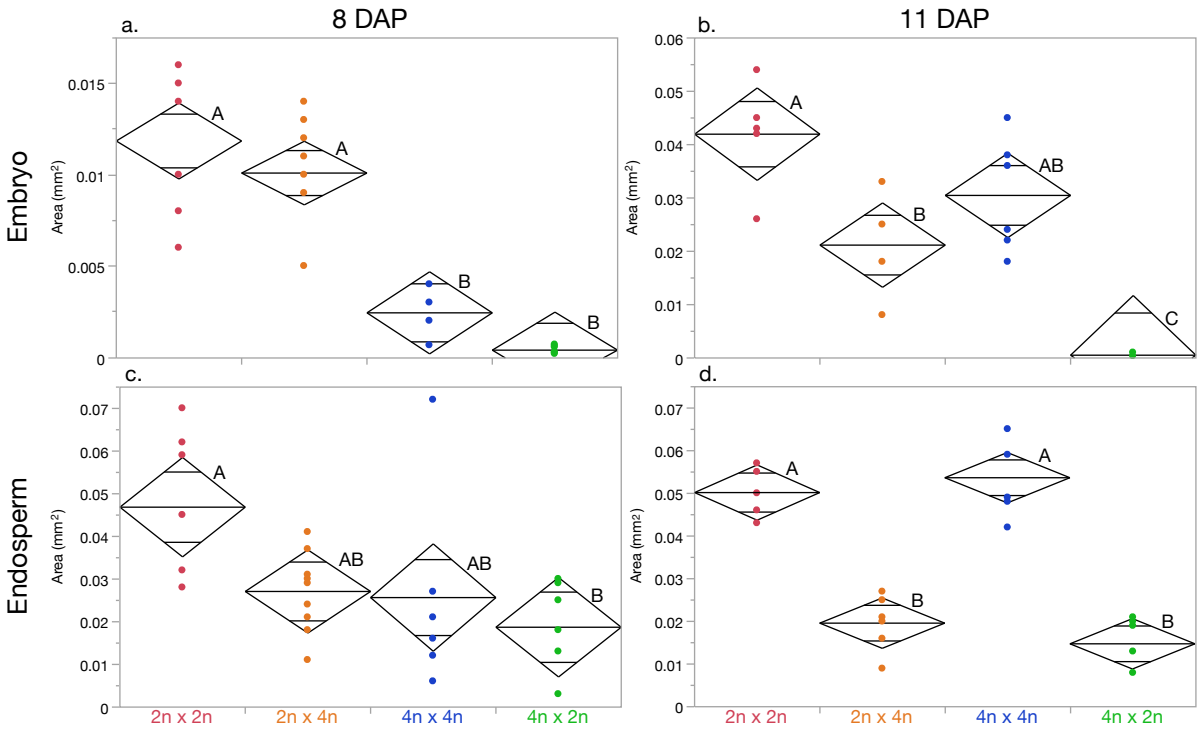

Figure 5: **Quantitative patterns of seed development**. The area of embryos (a and b) and endosperm (c and d) were measured at 8 DAP and 11 DAP. ANOVAs were performed for each of the plots. Letters associated with each group represent subsequent Tukey-Kramer tests: shared letters display no significant difference in area, while different letters display a difference with alpha of 0.05. (a) At 8 DAP, embryo size of *M. guttatus* (2n x 2n) is greater than that of *M. luteus* (4n x 4n). The reciprocal hybrid embryos are the same size as their maternal progenitors. (b) By 11 DAP embryos from all crosses except 4n x 2n have grown. However, 2n x 4n has grown to a lesser extent than its maternal progenitor, *M. guttatus*. *M. luteus* embryo size is catching up with *M. guttatus*. (c) At 8 DAP the endosperm size of *M. guttatus* is significantly larger than that of 4n x 2n. *M. luteus* and 2n x 4n endosperm do not differ significantly from any cross. (d) At 11 DAP the endosperm area of neither of the reciprocal hybrid crosses has increased, while the endosperm areas of *M. guttatus* and *M. luteus* have and are larger than the hybrids.

2.4 Abnormal endosperm and embryo growth results in delayed or failed seed germination

2.4.1 Endosperm area is highly correlated with whole seed area

Since we cannot directly observe seed tissue phenotypes and measure germination simultaneously, we measured correlations between area of the seed tissues and the whole seed, which is measurable in a germination experiment. Among the four crosses, endosperm size is highly correlated with the area of the entire seed with an R^2 of 0.69 (Fig. 6b). The embryo has a lower correlation with an R^2 of 0.40 (Fig. 6a). Considering individual crosses, endosperm shows high correlation with seed size of each cross except for the 2n x 4n hybrid, which is unsurprising since we found no increase in endosperm area from 8 to 11 DAP (Table 1). The embryo shows no correlation with whole seed area of any individual cross except for *M. guttatus*, with which it is highly correlated. This may be due to the quick growth of *M. guttatus* as compared to any of the other crosses and since the cotyledons absorb the endosperm in mature eudicot seeds. The seed coat has fairly high correlation with seed area (R^2 = 0.48), but individual crosses do not have significant p-values in their regressions. In general, the endosperm appears to be the most reliable predictor of seed size. Since the size of seed tissues determine the size of the seed, we can predict that trends in germination observed by seed size are driven by the phenotypes we observed within the seeds, particularly that of the endosperm.

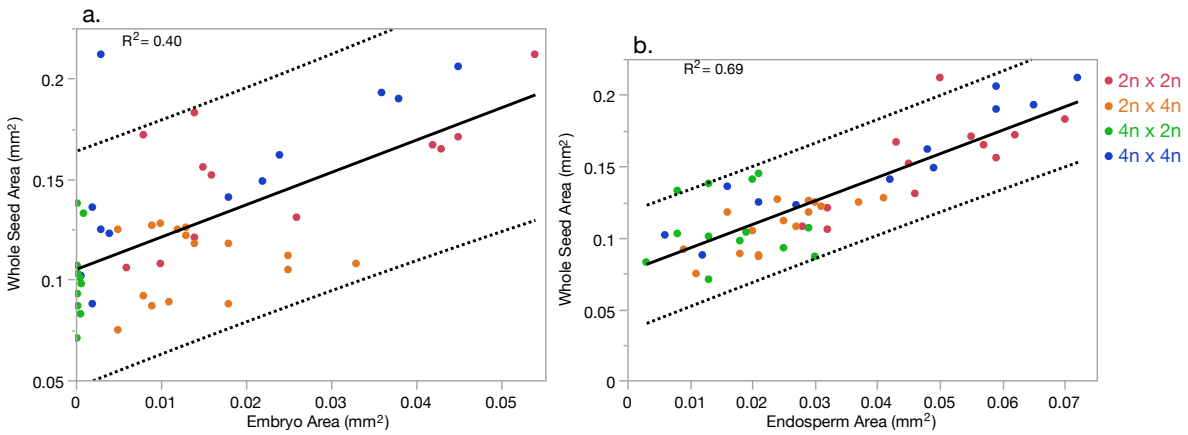

Figure 6: **Correlations between seed tissues and the whole seed**. In order to compare endosperm size to germination success, linear regressions were performed between seed tissues and the whole seed. Data points represent individual replicates color coded by the four crosses as seen in the legend. (a) Embryo size is fairly correlated with seed size (R^2 = 0.40). (b) Endosperm size is more tightly correlated with seed size (R^2 = 0.69).

2.4.2 Germinated seeds are larger

The area of seeds that germinated is significantly greater than that of seeds that failed to germinate (P <0.0001, Fig. 7). Mean area of germinated seeds was 0.14 +- 0.002 (n = 648), and mean area of ungerminated seeds was 0.12 +- 0.005 (n = 106). We did not include 4n x 2n seeds for fear of skewing the data since none of them germinated, and they are the smallest group, though the difference is still highly significant when included. Comparing individual crosses, only *M. luteus* shows no significant difference in size between germinated and ungerminated seeds. Furthermore, ungerminated *M. guttatus* seeds, ungerminated 2n x 4n seeds, and 4n x 2n seeds are all similar in size, while germinated seeds have differences among the crosses (except *M. luteus*). Given correlations within internal phenotypes, these comparisons suggest that ungerminated seeds, including 4n x 2n, suffer

form underdeveloped endosperm and embryos (although, for unknown reasons we do not see the same trends with *M. luteus*).

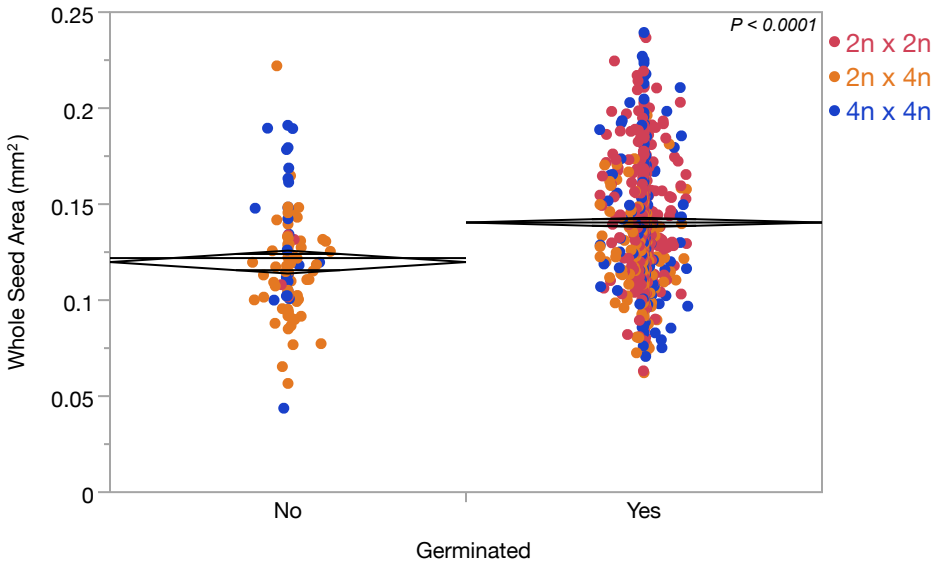

Figure 7: **Germinated seeds are larger than ungerminated seeds**. Area of seeds that failed to germinate is compared to successfully germinated seeds among the crosses. *M. guttatus* (2n x 2n) is depicted in red, 2n x 4n in orange, and *M. luteus* (4n x 4n) in blue. 4n x 2n seeds are not included because there were no germinates, and they are the smallest of the four crosses, so we did not want to bias the comparison. Germinating seeds are significantly larger than inviable seeds (P <0.0001).

2.4.3 Early germinating seeds are larger

Of seeds that did germinate, the area of seeds that germinated at 2 DAP is greater than those that germinated at subsequent dates (Fig. 8). Mean area of seeds germinating at 2 DAP was 0.148 +- 0.004 (n = 231), at 3 DAP was 0.138 +- 0.003 (n = 324), at 4 DAP was 0.130 +- .009 (n = 48), and at 5 DAP was 0.126 +- 0.011 (n = 29). However, given that *M. guttatus* seeds germinate more quickly than the other crosses [Vallejo-Marin et al., 2016]

and are larger than 2n x 4n, this comparison may be misleading. ANOVAs of individual crosses reveal differences among the germination dates for *M. guttatus* and 2n x 4n, but subsequent Tukey-Kramer pairwise tests are insignificant. Regressions reveal negative correlations between seed size and germination date, but R^2 values are very low. Seed, and thus endosperm, area likely influences the timing of germination - delaying germination of smaller seeds - but perhaps only weakly. Overall, these results suggest that the patterns of endosperm and embryonic growth observed in the cross-sections determine the outcome of hybrid seed viability.

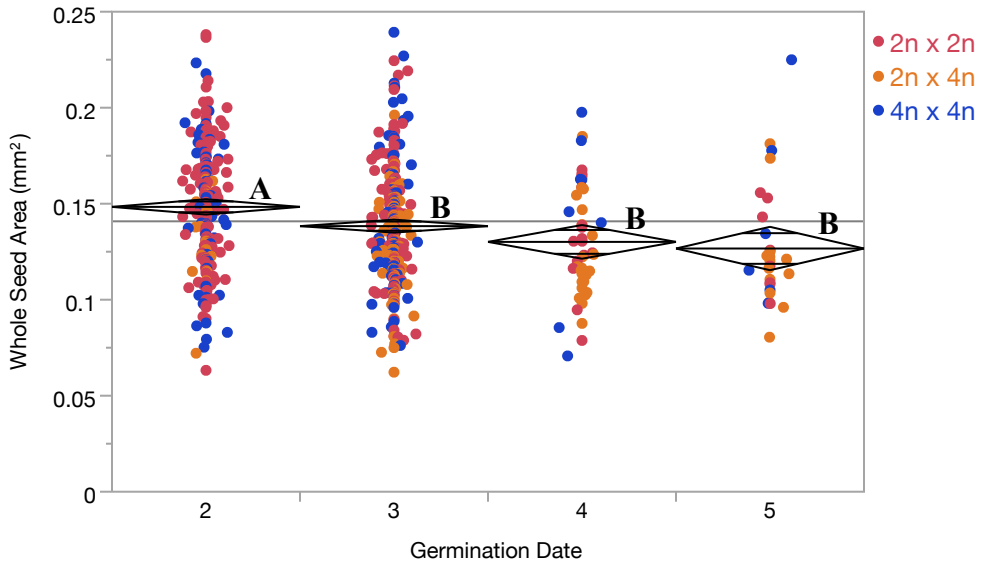

Figure 8: **Earliest germinating seeds are largest**. Of seeds that did germinate, those that germinated earlier are somewhat larger. ANOVA groups represent seeds that germinated at 2, 3, 4, and 5 days after the start of the experiment. *M. guttatus* (2n x 2n) is depicted in red, 2n x 4n in orange, and *M. luteus* (4n x 4n) in blue. 4n x 2n is not included because they had no germinates. Letters correspond to Tukey-Kramer tests. Shared letters indicate that no significance in differences were found.

2.5 Genomic imprinting and genomic expression dominance reveal unexpected patterns in *Mimulus* seeds

In order to investigate patterns of genomic imprinting in seed tissues, seeds were produced for the two reciprocal hybrid crosses and for *M. luteus*. Seeds were collected from 10-11 DAP, and endosperm and embryos were dissected and pooled separately. RNA was extracted and libraries were constructed and sequenced. Allele specific gene expression from endosperm for the two reciprocal crosses and for endosperm and embryo from *M. luteus* are reported. Following polyploidy, duplicated genes may evolve independently from one another resulting in separate ”sub”genomes [Cronn et al., 1999]. The tetraploid, *M. luteus* is composed of two independently segregating subgenomes. Two sets of chromosomes belong to subgenome A, and the other two sets belong to subgenome B [Edgar et al., 2016]. We refer to corresponding genes between the two subgenomes as homeologs. To test for subgenome specific patterns of imprinting in the hybrid crosses, subgenome specific gene expression was measured as well.

2.5.1 Bioinformatic method

Using two distinct inbred lines of *M. luteus* (M11 [Stanton et al., 2017] and C5 [Vallejo-Marin et al., 2016]) allows the identification of line specific SNPs for use in allele specific expression quantification. High-throughput RNA-seq was used to quantify allele specific expression. Following sequencing, reads were mapped to a ’hybrid’ reference genome that contained SNP corrected alleles. Thus, high-throughput sequencing reads could be mapped to specific alleles. We used the number of allele specific mappings as a proxy for the expression of that allele. These expression values were normalized by the length of the allele and total number of reads mapped for that given library.

After allele specific expression was determined, the next goal was to determine the expression ratio of each allele pair. Using methods previously established in the Puzey lab [Smith et al., 2017], we calculated the expression ratio of two alleles as $log2(allele1/allele2) = AEB$. AEB is the allele expression bias. Using a LRT we tested whether the alleles were expressed at unequal levels.

Next, using an extended LRT we asked whether the expression ratio (AEB) changed depending on crossing direction. This new term, AEBS (allele expression ratio shift) is quantified as $AEBS = log(1t/2t) - log(1s/2s)$, where t and s represent different conditions.

2.5.2 *M. luteus* endosperm is biased towards excessive paternal expression

A highly inbred line of (*M. luteus*), M11, was pollinated by a genetically distinguishable line, C5 (M11 x C5). AEB was measured in the endosperm from gene pairs with at least one read from each line. There were 2112 testable gene pairs in this cross of which 98 showed significant bias in either direction given the AEB likelihood ratio test (Fig. 9). Bias in the maternal direction is negative, and bias in the paternal direction is positive. After controlling for differences in genome dosage between maternal and paternal genes in the endosperm, the overall mean AEB was 0.50, indicating a general bias towards expression of paternal alleles. Given this data, it appears that either maternal alleles are generally down-regulated or paternal alleles are generally up-regulated in the endosperm of *M. luteus*. This may suggest genomic imprinting with overall paternal expression bias.

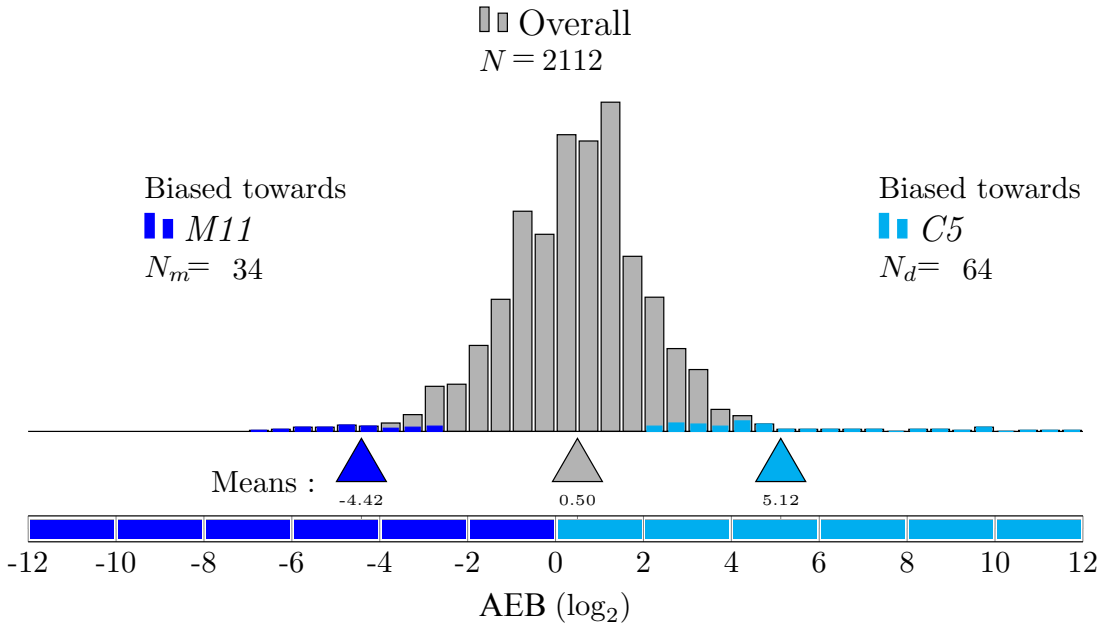

Figure 9: **Allele expression is biased towards paternal genome, C5, in the endosperm of M11 x C5**. Allele expression bias (AEB) is compared between two genetically distinguishable lines of *M. luteus*, M11 and C5, in the endosperm at 10 DAP. Counts represent testable gene pairs between M11 and C5 that had at least 1 read each. Differences in genome dosage are standardized. An AEB of 0 represents an equal number of reads from both parents. A negative AEB represents a bias in read number towards the maternal progenitor (M11), and a positive AEB is towards the paternal progenitor (C5). Gene pairs that passed the LRT test and had significant biases are highlighted (means shown by triangles). The overall mean AEB is 0.50.

2.5.3 *M. luteus* embryo is slightly biased towards excessive maternal expression

AEB was measured in the same manner from embryos of the same cross (M11 x C5) as described above. There were a total of 3690 testable gene pairs and 185 with significant bias in either direction (Fig. 10) The overall mean AEB was -0.37 suggesting a general

bias towards maternal expression in the *M. luteus* embryo. Again, this data suggests that either maternal alleles are up-regulated or paternal alleles are down-regulated, potentially revealing genomic imprinting in the embryo.

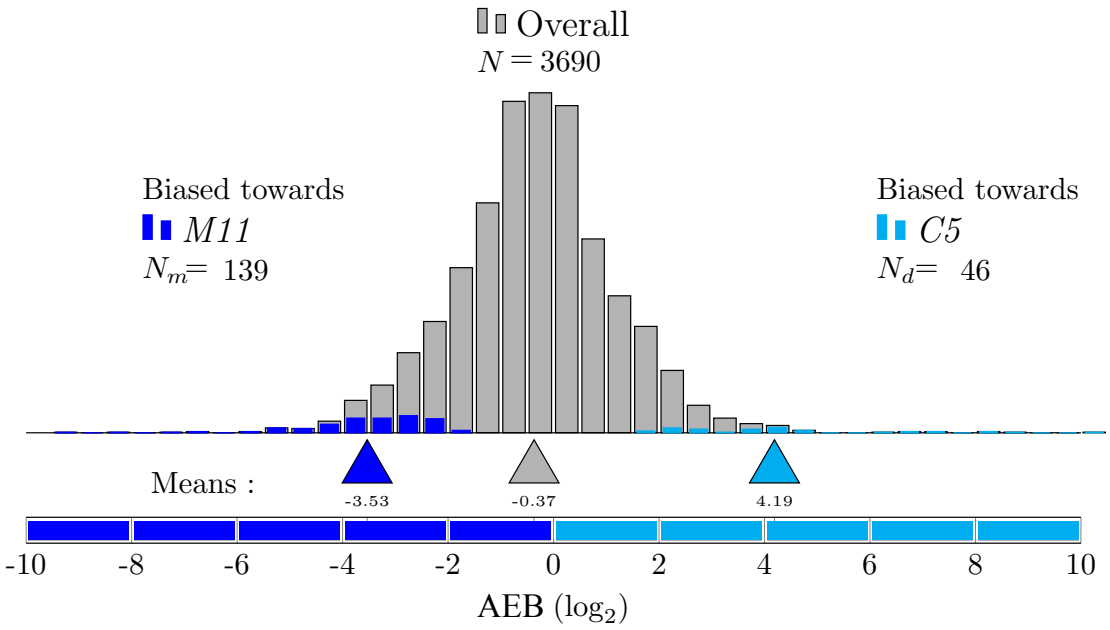

Figure 10: **Allele expression is biased towards maternal genome, M11, in the embryo of M11 x C5**. Allele expression bias (AEB) is compared between two genetically distinguishable lines of *M. luteus*, M11 and C5, in the embryo at 10 DAP. Counts represent testable gene pairs between M11 and C5 that had at least 1 read each. Differences in genome dosage are standardized. An AEB of 0 represents an equal number of reads from both parents. A negative AEB represents a bias in read number towards the maternal progenitor (M11), and a positive AEB is towards the paternal progenitor (C5). Gene pairs that passed the LRT test and had significant biases are highlighted (means shown by triangles). The overall mean AEB is -0.37.

2.5.4 2n x 4n shows excessive bias towards *M. luteus*, paternal tetraploid

A third generation inbred line of CG (*M. guttatus*) was crossed with M11 to produce the viable hybrid seeds (2n x 4n). In the endosperm, there were 5850 testable pairs between the homeologs of CG and M11 (Fig. 11). 3743 showed significant bias towards either species. The overall mean AEB was 1.06, showing bias towards the paternal *M. luteus*.

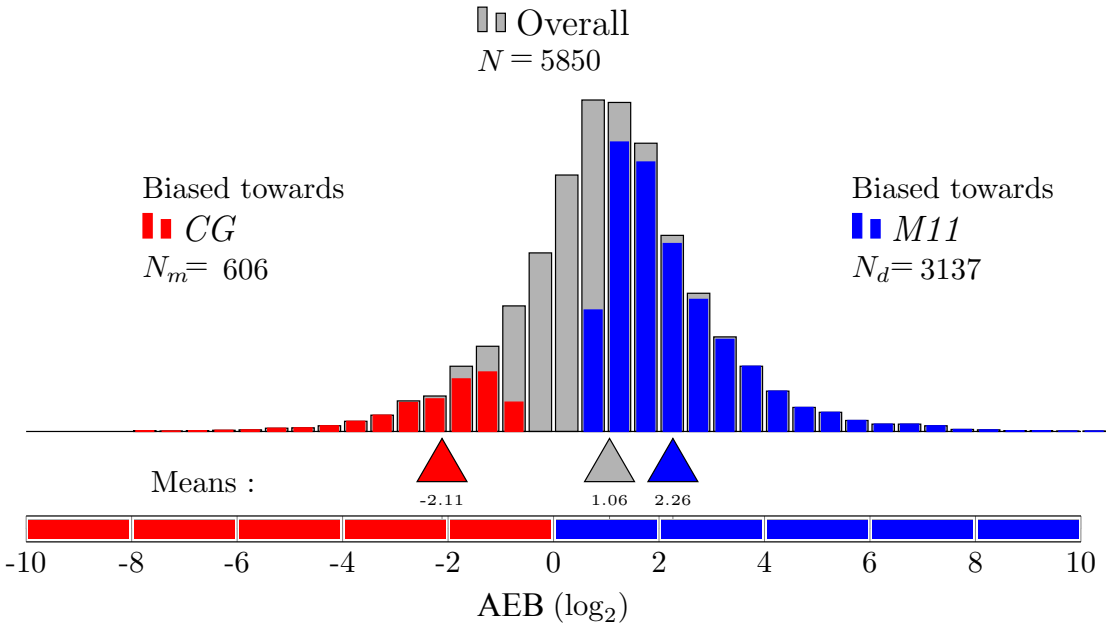

Figure 11: **Allele expression is biased towards paternal genome, M11, in the endosperm of CG x M11**. Allele expression bias (AEB) is compared between CG (*M. gutatus*) and M11 (*M. luteus*) in the endosperm at 11 DAP. Counts represent testable gene pairs between CG and M11 that had at least 1 read each. Differences in genome dosage are standardized. An AEB of 0 represents an equal number of reads from both parents. A negative AEB represents a bias in read number towards the maternal progenitor (CG), and a positive AEB is towards the paternal progenitor (M11). Gene pairs that passed the LRT test and had significant biases are highlighted. The overall mean AEB is 1.06.

2.5.5 4n x 2n shows excessive bias towards *M. luteus*, maternal tetraploid

The same lines as above were crossed reciprocally to produce the inviable hybrid seeds (4n x 2n). In the endosperm, there were 5767 testable pairs between the homeologs of CG and M11 (Fig. 12). 443 showed significant bias towards either species. The overall mean AEB was -0.97, showing bias towards the maternal *M. luteus*.

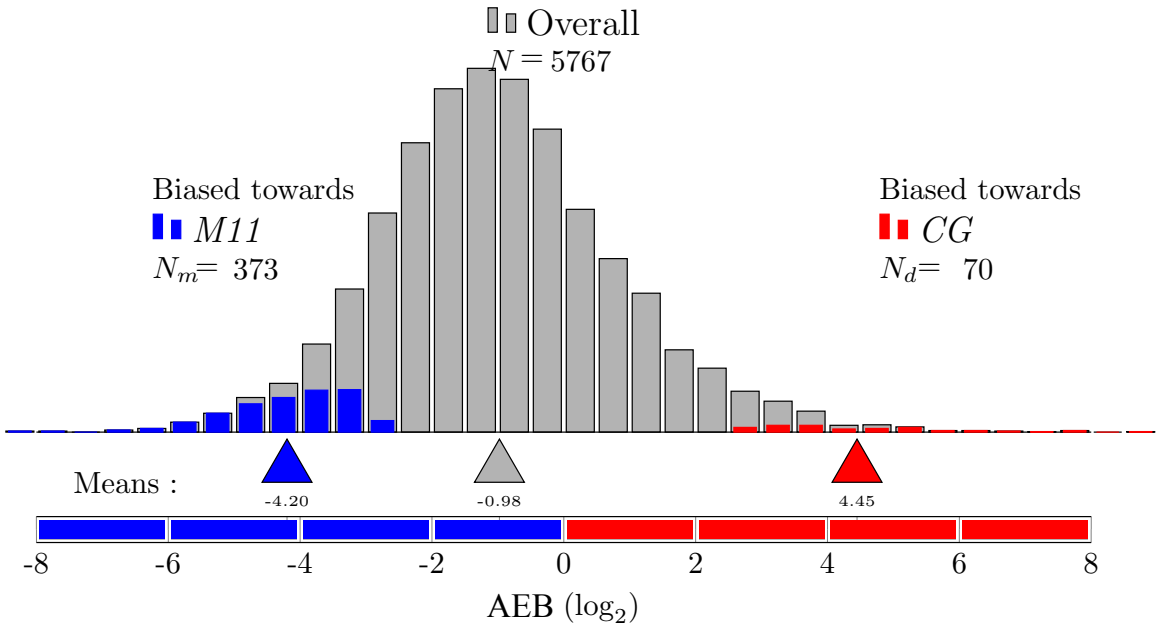

Figure 12: **Allele expression is biased towards maternal genome, M11, in the endosperm of M11 x CG**. Allele expression bias (AEB) is compared between M11 (*M. luteus*) and CG (*M. gutatus*) in the endosperm at 11 DAP. Counts represent testable gene pairs between M11 and CG that had at least 1 read each. Differences in genome dosage are standardized. An AEB of 0 represents an equal number of reads from both parents. A negative AEB represents a bias in read number towards the maternal progenitor (M11), and a positive AEB is towards the paternal progenitor (CG). Gene pairs that passed the LRT test and had significant biases are highlighted. The overall mean AEB is -0.97.

2.5.6 Overall expression is similar between each subgenome in the hybrids

In the viable hybrid (2n x 4n), when comparing CG to each M11 subgenome separately, the number of testable pairs with subgenome A was 358, with 251 showing significant AEB, and an overall mean of 1.28 (Fig. 13A). The number of testable pairs with subgenome B was 340, with 226 showing significant AEB, and an overall mean of 1.22 (Fig. 13B). There is strong bias towards *M. luteus*, the paternal progenitor, and the dominance appears to be driven equally by both subgenomes. Similarly in the inviable hybrid (4n x 2n), when comparing CG to each M11 subgenome separately, the number of testable pairs with subgenome A was 353, with 36 showing significant AEB, and an overall mean of -1.28 (Fig. 14A). The number of testable pairs with subgenome B was 328, with 27 showing significant AEB, and an overall mean of -1.38 (Fig. 14B). Again, after controlling for differences in genome dosage, there is strong bias towards *M. luteus*, which in this case, is the maternal progenitor. As in 2n x 4n endosperm, expression dominance appears to be driven equally by both subgenomes. Due to the almost equal bias towards *M. luteus* in either crossing direction, *M. luteus* may be more transcriptionally active in hybrid endosperm, exhibiting genomic expression dominance from each subgenome.

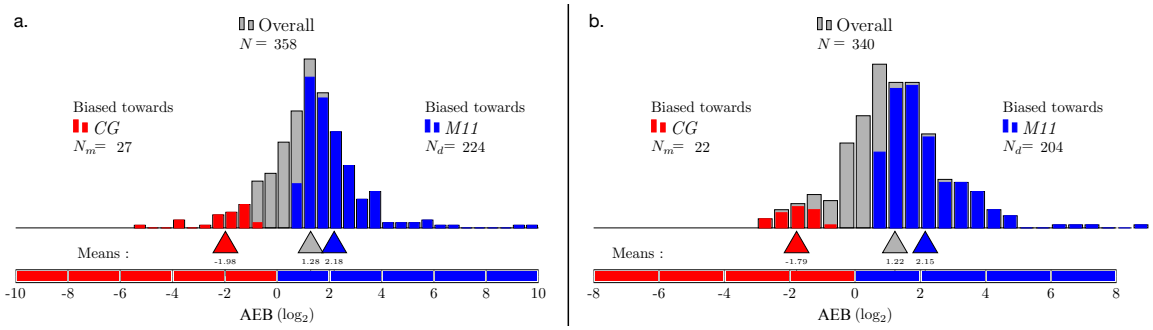

Figure 13: **Allele expression is equally biased towards each paternal subgenome of M11 in the endosperm of CG x M11**. Allele expression bias (AEB) is compared between CG (*M. gutatus*) and M11 (*M. luteus*) in the endosperm at 11 DAP. Counts represent testable gene pairs between CG and M11 that had at least 1 read each. Differences in genome dosage are standardized. An AEB of 0 represents an equal number of reads from both parents. A negative AEB represents a bias in read number towards the maternal progenitor (CG), and a positive AEB is towards the paternal progenitor (M11). Gene pairs that passed the LRT test and had significant biases are highlighted (means shown by triangles). Counts are separated by comparisons between CG and corresponding homeologs in the M11 subgenome A (A), and between CG and corresponding homeologs in the M11 subgenome B (B). The overall mean AEB for (A) is 1.28 and for (B) is 1.22. There is no apparent difference in the amount of bias towards either of the subgenomes.

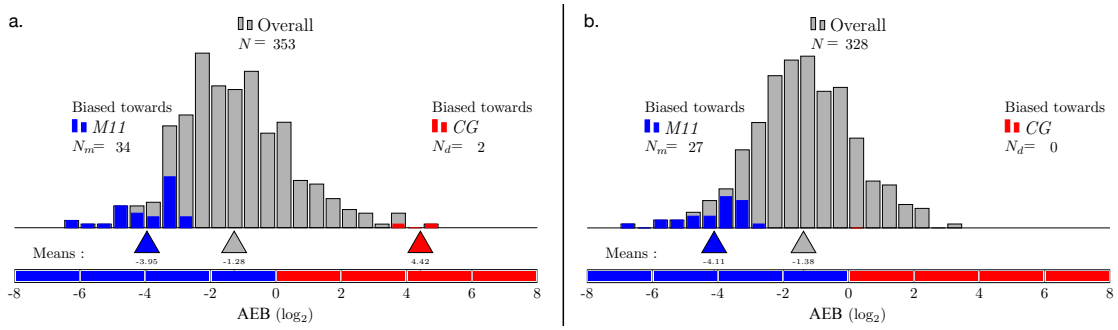

Figure 14: **Allele expression is equally biased towards each maternal subgenome of M11 in the endosperm of M11 x CG**. Allele expression bias (AEB) is compared between M11 (*M. luteus*) and CG (*M. gutatus*) in the endosperm at 11 DAP. Counts represent testable gene pairs between M11 and CG that had at least 1 read each. Differences in genome dosage are standardized. An AEB of 0 represents an equal number of reads from both parents. A negative AEB represents a bias in read number towards the maternal progenitor (M11), and a positive AEB is towards the paternal progenitor (CG). Gene pairs that passed the LRT test and had significant biases are highlighted (means shown by triangles). Counts are separated by comparisons between CG and corresponding homeologs in the M11 subgenome A (A), and between CG and corresponding homeologs in the M11 subgenome B (B). The overall mean AEB for (A) is -1.28 and for (B) is -1.38. There is no apparent difference in the amount of bias towards either of the subgenomes.

2.5.7 Gene-specific regulation differs between subgenomes in the hybrids

While overall bias is similar between each subgenome in both of the hybrids, bias may differ between subgenomes for individual genes. In order to test for this, we plotted AEB values for all testable gene pairs between CG and Subgenome B against the AEB of all matching gene pairs between CG and Subgenome A in the viable hybrid (2n x 4n) (Fig. 15). We

would expect a very high relationship between homeologs for AEB if they were regulated in the same way. However we find a fairly low R^2 of 0.37, suggesting dissimilarities of AEB values between homeologs on the subgenomes. This suggests that gene regulation or imprinting can be specific to or differ between subgenomes for a given gene in *M. luteus* and the hybrid.

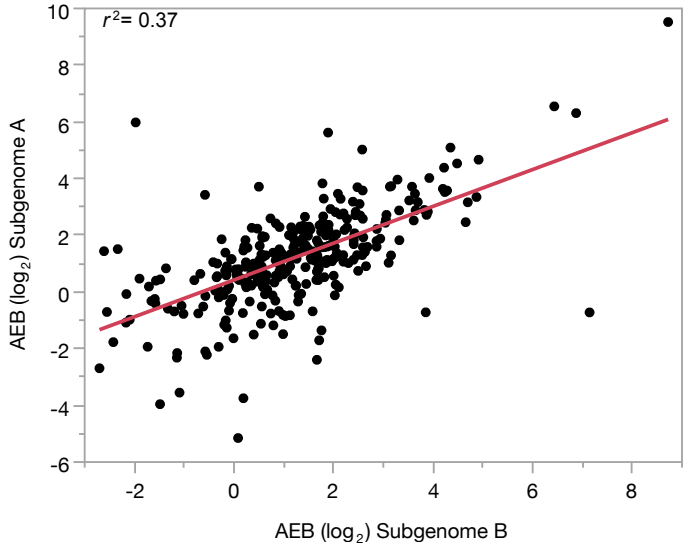

Figure 15: **Relationship between AEB in Subgenome A and Subgenome B**. AEB values of Subgenome B were compared to corresponding homeologs in Subgenome A in the viable hybrid (2n x 4n). A regression was performed with an R^2 value of 0.37.

2.5.8 Genomic dominance by *M. luteus* in the hybrids' endosperm

In order to test for crossing effect and identify true imprinted genes, we tested for shifts in AEB of gene pairs between the reciprocal hybrid crosses, 2n x 4n and 4n x 2n. This shift in AEB between two different conditions (in this case crossing direction) is referred to as AEBS (described above). We found that, regardless of the crossing direction, there is a shift in distribution towards expression by *M. luteus*, suggesting genomic dominance in the

hybrids by *M. luteus*. Furthermore, we identify 20 genes that always have significant biases towards the mother, which are true MEGs, and 28 genes that always have significant biases towards the father, which are true PEGs (Fig. 16).

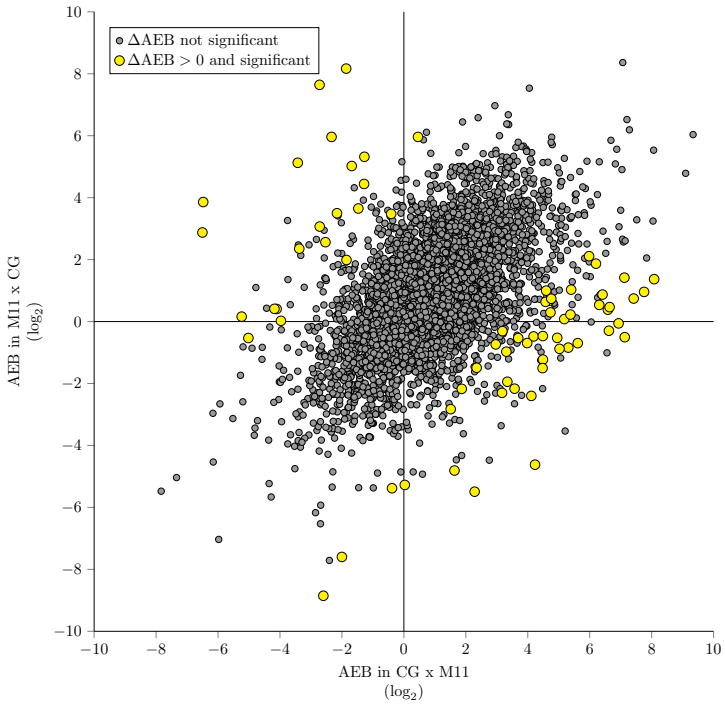

Figure 16: **AEBS for reciprocal hybrid crosses**. The shift in AEB between CG x M11 (the 2n x 4n hybrid) and M11 x CG (the 4n x 2n hybrid) is plotted. AEB values for gene pairs of CG x M11 are given by the x-axis, and AEB values for the same gene pairs of M11 x CG are given by the y-axis. Gene pairs that had a significant shift in AEB (AEBS) are colored in yellow. Gene pairs that fall within the top left quadrant are always biased towards the mother, and those in the bottom right are always biased towards the father. Significant values within these quadrants are true MEGs and true PEGs respectively. Gene pairs in the top right quadrant are always biased towards *M. luteus*, and those in the bottom left are always biased towards *M. guttatus*. Significant values here represent an even greater bias towards *M. luteus* or *M. guttatus* in one cross over the other.

2.5.9 Function of MEGs and PEGs in *Mimulus*

We selected a few random MEGs and PEGs to BLAST for gene identity and function. MEGs are involved in protein modification, aging, and immune response. PEGs are involved in processes of cell growth and division, which matches the endosperm proliferation of PEGs as well as the cellular type of endosperm growth seen in *Mimulus* and other Lamiales [Geeta, 2003] (Table 2).

Gene	**Allelic Expression**	**Function**
Hpgt1	MEG	protein modification
Aalp	MEG	aging, immune response
Atrbl1	PEG	proteolysis
Casp5	PEG	cell wall modification and junction assembly
Atgh 9b13	PEG	carbohydrate metabolism, cell wall organization, cellulolysis
Bglu16	PEG	carbohydrate metabolism, hormone response

Table 2: **MEGs and PEGs**. A selection of genes identified from AEBS were BLASTed to assess gene functions found in MEGs and PEGs.

3 Discussion

The use of interploidy hybrid systems has long been associated with naturally altering parental genome dosage to study parent of origin effects in developing seeds. While such effects could be driven by other mechanisms, such as cytonuclear interactions, imprinting has been a key target of focus [Gehring and Satyaki, 2017]. Such studies have revealed a trend of endosperm underproliferation from maternal excess crosses and endosperm overproliferation from paternal excess crosses in Arabadopsis [Scott et al., 1998]

and maize [Pennington et al., 2008]. These alterations in genome balance also affect the differentiation of cell types in the complex grain endosperm of maize and affect the timing of cellularization. Maternal excess expedites cellularization while paternal excess delays or even prevents it. In species that undergo nuclear, or syncytial, endosperm development, where multiple rounds of nuclear proliferation proceed cytokinesis [Olsen, 2004], the process and timing of endosperm cellularization is critical in seed viability [Hehenberger et al., 2012]. Some genes involved in endosperm cellularization have been identified as targets of imprinting within the endosperm [Gehring and Satyaki, 2017]. While not all angiosperm species undergo this method of endosperm development, it appears to have had strong selective advantages in angiosperms, evolving independently in many clades, with basal angiosperm and eudicot species exhibiting the ancestral cellular proliferation trait (normal cellular division in the endosperm) [Geeta, 2003]. Endosperm proliferation and cellularization have been the predominant parent of origin phenotypes described in interploidy hybrid systems, as well as interspecies hybrid systems [Rebernig et al., 2015] in plants.

The interploidy hybrid system we have studied in *Mimulus* reveals striking differences from those observed in maize and Arabidopsis [Pennington et al., 2008, Scott et al., 1998]. The paternal excess hybrid, 2n x 4n (*M. guttatus* x *M. luteus*) produces a phenotype opposite to what is expected. Paternal excess is predicted to induce increased proliferation of endosperm cells (Lamiales, the order to which *Mimulus* belongs, undergoes cellular endosperm development [Geeta, 2003]). However, this cross suffers from underproliferation of the endosperm, which likely reduces growth of the dependent embryo as well. Furthermore, the maternal excess hybrid, 4n x 2n, which is predicted to have underproliferated endosperm and reduced embryonic growth, suffers from near complete failure of both the endosperm and the embryo. Such unexpected phenotypes as compared to other taxa may

be due to differences in misregulation or due to differences in imprinting patterns of the species.

There are varying patterns of overall parental bias in different species where imprinting has been studied. Early imprinting studies in maize revealed an overall bias of MEGs over PEGs, though more recent studies reveal a more even division [Waters et al., 2011] or even a greater proportion of PEGs [Waters et al., 2013]. The latter study did however show that a higher proportion of MEGs were endosperm specific and that imprinting for many maize genes may be of limited functional consequence. Imprinting studies in another monocot model, rice, revealed overall more MEGs than PEGs [Luo et al., 2011] as well as in the monocot, sorghum [Zhang et al., 2015]. Genomic imprinting has been studied in three eudicot taxa: Arabidopsis (and closely related outcrossing species), *Solanum* (tomatoes), and *Ricinus* (castor bean). A strong bias towards MEGs had previously been described in Arabidopsis [Gehring et al., 2011], though a recent study on maternal seed coat contamination in Arabidopsis embryo and endosperm transcriptome data suggested that, after accounting for this contamination, no overall proportional bias towards MEGs or PEGs could be detected [Schon and Nodine, 2017]. A large proportion of imprinted genes in *Solanum peruvianum* are MEGs, while *S. chilense*, where far fewer imprinted genes were found, does not share this pattern [Florez-Rueda et al., 2016]. Maternal bias is very strong in the castor bean endosperm [Xu et al., 2014]. Broad imprinting patterns do not appear to be widespread or quite generalizable across different taxa, though a preference for some extent of maternal bias seems common.

M. luteus has an unusual pattern of imprinting. While distinguishable genetic differences between the two lines of *M. luteus* were limited, there were still many testable gene pairs between the two alleles with a number exhibiting significant bias in either direction. *M.*

luteus appears to demonstrate an overall paternal bias within the endosperm, though the reciprocal cross will still need to be analyzed to verify this. Compared to other taxa studied, this pattern is unexpected. In some cases, this bias is quite strong, with one putative PEG showing an almost 3000 fold bias over its homeolog. Furthermore, *M. luteus* has an overall maternal bias in the embryo with several maternally and paternally biased genes, some with very strong bias. While we cannot yet verify true genomic imprinting, if so, this will be the first documentation of significant imprinting within the embryo of any angiosperm. It is quite interesting that the endosperm and embryo of *M. luteus* exhibit opposite biases.

Patterns of imprinting can differ between even recently diverged species. *A. lyrata*, an outcrossing relative to Arabidopsis, shares 50% of its PEGs and 35% MEGs with Arabidopsis. In addition, PEGs possessed greater CHG methylation on the maternal allele than Arabidopsis, resulting in even more biased expression [Klosinska et al., 2016]. According to the conflict hypothesis, it is unsurprising that *A. lyrata* has greater regulatory patterns and increased bias in its PEGs since Arabidopsis is predominately selfing, and *A. lyrata* is an outcrossing species, which may, furthermore, allow it to better respond to selection pressures [Lafon-Placette et al., 2017]. Such differences in mating strategies and responsiveness to selection pressures could explain these rather striking differences in imprinting patterns between related species. *M. luteus* is also a predominantly outcrossing species with highly conspicuous and variable flowers, perhaps explaining the observed bias towards its paternal allele in the endosperm [Medel et al., 2003].

In addition to mating strategies, studied species also have varying strategies in seed and endosperm development. Monocot seeds are albuminous, or endosperm persistent. The endosperm is retained in mature seeds and during seed dormancy, and the embryo metabolizes the large starch reserves during seed germination. Arabidopsis and *Mimu-*

lus, like most eudicots, are exalbuminous seeds, where the majority of the endosperm is absorbed into the cotyledons at seed maturity and stored there during dormancy until germination [Yan et al., 2014]. Patterns of imprinting may differ based off such strategies for nutrient storage. *A. lyrata* and *M. luteus* both display some sort of bias towards PEGs or paternal expression and have potentially fewer total imprinted genes than other species. There may be relaxed conflict or drive for imprinting when the endosperm is so transient. Alternatively, or in addition to this, quick proliferation of endosperm to quickly be absorbed into the cotyledons may select for stronger paternal influence. The albuminous monocots, rice and sorghum, have many maternally biased loci [Luo et al., 2011, Zhang et al., 2015]. Maize has many MEGs and PEGs, though it also has a very complicated endosperm and is highly sensitive to changes in genome dosage [Leblanc et al., 2002]. Intriguingly, the eudicot, castor bean, has highly persistent endosperm unlike Arabidopsis and *Mimulus*. At seed maturity, the cotyledons are thin, and the majority of energy storage is within the endosperm [Greenwood et al., 2005]. In contrast to Arabidopsis and *Mimulus*, the castor bean endosperm is highly maternally biased. Many of these MEGs are involved in metabolic processes and may have strong influence over endosperm development [Xu et al., 2014]. Since seeds are so large, take a long time to develop, and are dominated by endosperm, maternal control may be favored, which follows a couple of imprinting theories on maternal-embryo coadaptation or coordination [Spencer and Clark, 2014]. *Solanum*, is also a eudicot with a persistent endosperm, though not quite as dominant as in castor. At seed maturity, the endosperm is hard and thick and is broken down upon germination [Yan et al., 2014]. The species studied ranged from no maternal bias (though total imprinted genes were low here) to fairly large bias [Florez-Rueda et al., 2016]. To investigate the role of endosperm persistence on genomic imprinting further, more species will need to be investigated includ-

ing other albuminous eudicots, such as coffee where 99% of the mature seed is composed of endosperm [Joët et al., 2009], exalbuminous eudicots, such as garden peas were the entirety of endosperm is absorbed into the cotyledons at maturity [Yan et al., 2014], more albuminous monocots, and exalbuminous monocots, such as *Vallisneria*, where starch is stored in the one cotyledon and the large hypocotyl [Witmer, 1937]. Perhaps a further line of interest and potentially evidence, albuminous seeds with an underdeveloped embryo upon seed maturity and dormancy is likely the ancestral state in angiosperms [Forbis et al., 2002], when imprinting was likely first evolving in the endosperm. Genomic imprinting studies on albuminous basal angiosperms may provide insight into early forces of imprinting, and comparison to differing derived seed development systems may help elucidate other drivers behind variable imprinting patterns. Overall, mating strategies, differences in seed development, and a variety of other drivers likely influence roles and the effects of the pollen parent and the seed parent and thus, subsequent patterns of imprinting.

In our *Mimulus* system, we have observed unusual patterns of parent-specific expression as well as unusual mechanisms of abnormal seed development based off of differences in genome dosage. Clear patterns of imprinting appear to be over-shadowed by *M. luteus* expression in the reciprocal hybrid crosses, perhaps describing a lack of inter-genomic post-zygotic regulation due to genetic divergence, or an imbalance of transcription due to perturbation of or differences in regulatory machinery (similar to patterns observed in *Solanum* hybrids [Florez-Rueda et al., 2016]). Genomic dominance from *M. luteus* is unsurprising since it has been observed in vegetative tissue from the hybrid [Edgar et al., 2016]. From PEGs we identified involved in cell growth and replication, observations we have made on very minimal endosperm development in the flat 4n x 2n seeds, and the strong bias towards *M. luteus*, the maternal progenitor, in this cross, we conclude that the

inviable hybrid (*M. luteus* x *M. guttatus*) suffers from insufficient expression or a misregulation of these PEGs necessary for endosperm proliferation.

The narrative of the viable hybrid, 2n x 4n (*M. guttatus* x *M. luteus*) is less evident. While sufficient endosperm is often produced, the seeds appear to suffer from underproliferation, a phenotype typically observed in maternal excess crosses [Scott et al., 1998]. Furthermore, *M. luteus*, again in this cross, is dominating in genomic expression. Since it is the paternal progenitor, a maternal excess phenotype is surprising. The uniqueness and utility of this system is that it not only represents an interploidy hybrid system, but also a fairly diverged interspecies hybrid system. *M. guttatus* and *M. luteus* may represent a point of divergence where epigenetic regulatory mechanisms are not well recognized or as functional between the two species' genomes, preventing the appropriate proliferation PEGs from *M. luteus* from activating to the necessary extent. Alternatively, these proteins or regulatory mechanisms may not function as well within the *M. guttatus* nucleus/cell. Imprinted genes and imprinting mechanisms can differ quite substantially and rapidly between different, even closely related, species in plants and mammals [Gehring and Satyaki, 2017, Kinoshita, 2007, Monk, 2015]. Furthermore, we still don't know what role the differences in imprinting patterns for individual genes between subgenomes plays on the phenotype of these hybrids.

Endosperm failure has been associated with other developmental differences between species as well. There is a strong hybridization barrier between two closely related *Mimulus* species, *M. guttatus* and *M. nudatus*, an annual species within the *M. guttatus* species complex that occurs within close proximity to *M. guttatus* [Oneal et al., 2014]. Hybrid seeds suffer from the same shriveled and flat morphologies as seen in our study [Oneal et al., 2016]. Furthermore, like *M. guttatus* and *M. luteus*, *M. guttatus* and *M. nudatus* differ in

development patterns and timing of endosperm and embryo development, and reciprocal hybrid seed development appears to follow the developmental timing of its maternal progenitor, at least in earlier stages. One important difference between this study and ours is that, while there is some asymmetry between reciprocal hybrid crosses, it is not as pronounced as the asymmetry in seed morphology and viability as in *M. guttatus* and *M. luteus* hybrids. Additionally, overall germination success of hybrids is very low and does not exhibit the frequent hybridization between *M. guttatus* and *M. luteus* in the wild and widespread *M. robertsii* [Vallejo-Marin et al., 2016]. Oneal et al. predict that early arrested seed development in *M. guttatus* and *M. nudatus* hybrids results in the flat seed morphology and the delayed development results in the shriveled morphology and that this is due to differences in development patterns. We cannot rule out the possibility that differences in developmental timing between *M. guttatus* and *M. luteus* is, at least in part, driving the morphologies and asymmetry seen in the hybrids. The 4n x 2n hybrid seems to have clear deficiencies in dosage expression, but it could very well be that the 2n x 4n hybrid suffers from a delayed contribution of proliferating PEGs since, as we have shown, *M. luteus* has slower developmental timing than *M. guttatus*.

4 Future Directions

Further work will be added to supplement and expand the findings from this project. As reciprocal crosses were used in the hybrid to perform AEBS and verify MEGs and PEGs, a reciprocal cross for M11 x C5 (C5 x M11) will be sequenced to also perform this analysis in *M. luteus*. Furthermore, reciprocal crosses of genetically distinguishable lines of *M. guttatus* will be sequenced for the same analysis. By identifying MEGs and PEGs in *M. guttatus* and *M. luteus*, we can understand how well conserved imprinted genes are between the two

species and which genes maintain their imprinted status and function in the hybrid. Using GO Categories, we can further identify functions of MEGs, PEGs, subgenome specific imprinted genes in *M. luteus*, and imprinted genes that are and are not maintained in the hybrids. This information will be useful for understanding the role of imprinted genes in these *Mimulus* species, how imprinted genes are partitioned between subgenomes, and their implications on the developmental patterns observed in the hybrids. To understand the role of genetic divergence on genomic recognition and imprinting in the developing hybrids, data on genetic relatedness between the subgenomes of *M. luteus*, and between *M. luteus* and *M. guttatus* will be used. Finally, Methyl-Seq will be performed to identify mechanisms of imprinting and to compare overall methylation patterns in the different crosses. These patterns will be used to better understand how parent genomes of these two species interact within the developing hybrids.

5 Methods

5.1 Crossing design

Four different crosses were set up for all morphology experiments. *M. luteus* individuals, line C5, and *M. guttatus* individuals, line CG, were used [Vallejo-Marin et al., 2016]. To produce *M. luteus* seeds, C5 plants were outcrossed with each other, and CG seeds were outcrossed with each other to produce *M. guttatus* seeds. Viable hybrid seeds, 2n x 4n, were produced by fertilizing CG flowers with C5 pollen, and inviable hybrid seeds, 4n x 2n, were produced by the opposite pollination.

5.2 Seed morphology

5.2.1 External Morphology

Seeds from each of the four crosses were collected at 15-16 days after pollination (DAP). Four ovaries with 100-200 seeds each were collected for each of the four crosses. Images were taken of all seeds from each cross under a dissecting microscope. Images were then run through a program in ImageJ that produced different measurements from the outline of the seeds. Average area and aspect ratio was measured from all seeds pooled for each cross. ANOVAs were performed using JMP® among the four crosses to measure differences in area and aspect ratio.

5.2.2 SEM images

Ten seeds from each of the four crosses were fixed in 2% glutaraldehyde with Phosphate Buffered Saline (PBS) for two hours. After fixation, seeds were washed with PBS and put through a dehydration series at concentrations of 50%, 70%, 85%, 95%, and 100% ethanol:PBS for one hour at each step. Seeds were left in 100% ethanol overnight. One hour before critical point drying, 100% ethanol was renewed. After critical point drying with a Samdri®-PVT-3B Critical Point Drier (Tousimis® Research Corporation), seeds were sputter-coated with gold-paladium using a Hummer® sputtering system from Anatech Ltd. and mounted on pin stubs for Scanning Electron Microscopy (SEM). All images were acquired using Phenom Tabletop Scanning Electron Microscope.

5.2.3 Seed histology

Three replicates of the above crosses were collected for each of the following DAP: 3, 5, 8, and 11 DAP. Ovaries were immediately placed into 4% paraformaldehyde and vacuum

infiltrated for 15-20 minutes and kept at 4°C for 48 hours. They were then washed with PBS for 30 minutes three times. They were left in fresh PBS at 4°C overnight. Ovaries were then dehydrated using the following dehydration series: 10% ethanol for 15 minutes, 30% for 15 minutes, 50% for 15 minutes, 70% for 15 minutes, 95% for 20 minutes, and two times of 100% for 30 minutes, all on a shaker. Ovaries were next infiltrated through the following infiltration series: 100% propylene oxide three times for 20 minutes each, 2:1 propylene oxide to Spurr's resin mixture for one hour, and 1:1 propylene oxide to Spurr's for one hour, all on a shaker. They were embedded into full resin, changed after two hours, and left overnight in fresh resin at room temperature on a shaker. The next day, resin was changed twice every two hours. Ovaries were then placed into individual molds and left in an oven at 60°C for 48 hours. Using an ultramicrotome, 0.5 micrometer sections were produced from the resin molds. Sections were adhered to slides in water at 50°C on a heat block. Slides were next stained with Azure II for 5 minutes at 50°C on a heat block. Coverslips were mounted onto slides using acrylamide. Images were next taken on a compound microscope using SPOT™.

5.2.4 Quantitative internal morphology

Seed histology images from three replicates of each of the four crosses at 8 and 11 DAP were analyzed in ImageJ. Endosperm, embryo, seed coat, and the whole seed were traced and measured for area. ANOVAs were performed comparing the four crosses at 8 DAP and 11 DAP using JMP®. T-tests were performed to demonstrate endosperm and embryo growth from 8 to 11 DAP.

5.3 Seed germination

In order to link the observed seed morphologies to germination rates, a germination experiment was performed in which morphology was measured. Seeds from each cross were pollinated at the same time and collected at 15 - 17 DAP (when mature and dihescing). Seeds were soaked in 3% calcium hypochlorite for 10 minutes on a shaker. They were then rinsed in PBS three times for 5 minutes on a shaker. Seeds were placed onto 60 mm petri dishes with gridded filter paper on top of additional filter paper to retain moisture. The petri dishes and filter paper had been sterilized with UV for 30 minutes first. Images of seeds on the plates were taken, and seed germination was tracked specifically for each seed based off its position on the grid. 5 plates were used per cross with 16-32 seeds per plate. Seed images were then analyzed in ImageJ. Area of seeds were then matched to the germination of each given seed. Statistical comparisons of area and germination rate, and area and germination date were performed in JMP®.

5.4 RNA-Seq

5.4.1 Plant material

Two genetically distinguishable lines of *M. luteus* were crossed to identify parental genomes in the progeny. The highly inbred M11 [Stanton et al., 2017] was pollinated by a different line of *M. luteus*, C5 [Vallejo-Marin et al., 2016]. Seeds were collected at 10 DAP. To reproduce the hybrids, M11 was crossed with a third generation inbred line of *M. guttatus*, CG [Vallejo-Marin et al., 2016]. Seeds of each reciprocal hybrid (CG x M11 and M11 x CG) were collected at 11 DAP.

5.4.2 Endosperm isolation and RNA extraction

In order to separate endosperm tissue from the rest of the seeds, a modified protocol was used from [Gehring et al., 2009]. After collection, ovaries were stored in RNAlater® (Ambion™) at 4°C. Endosperm was dissected from seeds in a 0.3 M sorbitol/5mM MES (2-[N-morpholino] ethanesulfonic acid) solution [Gehring et al., 2009] on a glass slide under a dissecting microscope using 30-gauge hypodermic needles. Endosperm from 20-30 seeds were pooled per replicate (an ovary from one pollination event) and rinsed with the same solution 5-10 times. Pooled endosperm was then placed into the Lysis solution of RNAqueous®-Micro Total RNA Isolation Kit (Ambion™) and disrupted with a pestle (rotating 50-60 times). The remainder of the RNA extraction protocol from the kit was performed.

5.4.3 Library preparation and sequencing

RNA was converted into cDNA and libraries were constructed using KAPA Stranded mRNA-Seq Kit. During library construction, sequence specific Illumina TruSeq® adapters were added to distinguish each library. Using an Agilent 2100 Bioanalyzer, average fragment lengths were determined to be between 250 and 300 bp. Libraries were then pooled and sequenced by the Duke Center for Genomic and Computational Biology on an Illumina® HiSeq 2500 instrument.

5.4.4 Genome assembly and AEB/AEBS

Parental genomes were SNP corrected by first mapping paired end WGS data to previously assembled reference genomes [Edgar et al., 2016, Hellsten et al., 2013] using bowtie2 in –very-sensitive-local mode [Langmead and Salzberg, 2012]. Picardtools were used to prepare the files for the GATK toolbox by invoking FixMateInformation to verify and fix mate

information for paired reads, MarkDuplicates to remove PCR duplicates, and AddOrReplaceReadGroups to add read groups. The GATK UnifiedGenotyper was used to create VCF files with the following options:

java -Xmx5g -jar GenomeAnalysisTK.jar -R genome -T UnifiedGenotyper -rf MaxInsertSize –maxInsertSize 10000 -rf DuplicateRead -rf BadMate -rf BadCigar –minbasequalityscore 25 -rf MappingQuality –minmappingqualityscore 25 -ploidy 2 –genotypelikelihoodsmodel BOTH –outputmode EMITALLSITES –maxalternatealleles 2 str -o out

and the GATK FastaAlternateReferenceMaker (http://www.broadinstitute.org/gatk/) was used to generate the new SNP corrected fasta files.

For each cross, the SNP corrected coding regions of the parental genomes were combined into a single fasta file, and the fastq reads were aligned to these references using bowtie2 in –very-fast-local mode. Counts were generated using HTSeq-count with the default options.

The statistical tests for AEB [Smith et al., 2017] used a 5% significance level, and the Benajmini-Hochberg correction [Benjamini and Yekutieli, 2001] was made to correct for multiple testing error.

6 Bibliography

References

[Ågren and Andersson, 2011] Ågren, G. I. and Andersson, F. O. (2011). *Terrestrial ecosystem ecology: principles and applications*. Cambridge University Press.

[Baroux et al., 2002] Baroux, C., Spillane, C., and Grossniklaus, U. (2002). Evolutionary origins of the endosperm in flowering plants. *Genome Biology*, 3(9):reviews1026–1.

[Benjamini and Yekutieli, 2001] Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of statistics*, pages 1165–1188.

[Berendse and Scheffer, 2009] Berendse, F. and Scheffer, M. (2009). The angiosperm radiation revisited, an ecological explanation for darwin's 'abominable mystery'. *Ecology Letters*, 12(9):865–872.

[Crepet and Niklas, 2009] Crepet, W. L. and Niklas, K. J. (2009). Darwin's second "abominable mystery": Why are there so many angiosperm species? *American Journal of Botany*, 96(1):366–381.

[Cronn et al., 1999] Cronn, R. C., Small, R. L., and Wendel, J. F. (1999). Duplicated genes evolve independently after polyploid formation in cotton. *Proceedings of the National Academy of Sciences*, 96(25):14406–14411.

[Edgar et al., 2016] Edgar, P. P., Smith, R. D., McKain, M. R., Cooley, A. M., Vallejo-Marin, M., Yuan, Y., Bewick, A. J., Ji, L., Platts, A. E., Bowman, M. J., Childs, K. L., Schmitz, R. J., Smith, G. D., Pires, J. C., and Puzey, J. R. (2016). Subgenome dominance in

an interspecific hybrid, synthetic allopolyploid, and a 140 year old naturally established neo-allopolyploid monkeyflower. *bioRxiv*.

[Endress, 2011] Endress, P. K. (2011). Angiosperm ovules: diversity, development, evolution. *Annals of Botany*, 107(9):1465–1489.

[Florez-Rueda et al., 2016] Florez-Rueda, A. M., Paris, M., Schmidt, A., Widmer, A., Grossniklaus, U., and Städler, T. (2016). Genomic imprinting in the endosperm is systematically perturbed in abortive hybrid tomato seeds. *Molecular Biology and Evolution*, 33(11):2935.

[Forbes and Watson, 1992] Forbes, J. C. and Watson, D. (1992). *Plants in agriculture*. Cambridge University Press.

[Forbis et al., 2002] Forbis, T. A., Floyd, S. K., de Queiroz, A., and Baum, D. (2002). The evolution of embryo size in angiosperms and other seed plants: implications for the evolution of seed dormancy. *Evolution*, 56(11):2112–2125.

[Frank et al., 2002] Frank, D., Fortino, W., Clark, L., Musalo, R., Wang, W., Saxena, A., Li, C.-M., Reik, W., Ludwig, T., and Tycko, B. (2002). Placental overgrowth in mice lacking the imprinted gene ipl. *Proceedings of the National Academy of Sciences*, 99(11):7490–7495.

[Garner et al., 2016] Garner, A. G., Kenney, A. M., Fishman, L., and Sweigart, A. L. (2016). Genetic loci with parent-of-origin effects cause hybrid seed lethality in crosses between mimulus species. *New Phytologist*.

[Garnica and Chan, 1996] Garnica, A. D. and Chan, W.-Y. (1996). The role of the placenta in fetal nutrition and growth. *Journal of the American College of Nutrition*, 15(3):206–222.

[Geeta, 2003] Geeta, R. (2003). The origin and maintenance of nuclear endosperms: viewing development through a phylogenetic lens. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1510):29–35.

[Gehring et al., 2009] Gehring, M., Bubb, K. L., and Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science*, 324(5933):1447–1451.

[Gehring et al., 2011] Gehring, M., Missirian, V., and Henikoff, S. (2011). Genomic analysis of parent-of-origin allelic expression in arabidopsis thaliana seeds. *PLoS One*, 6(8):e23687.

[Gehring and Satyaki, 2017] Gehring, M. and Satyaki, P. (2017). Endosperm and imprinting, inextricably linked. *Plant Physiology*, 173(1):143–154.

[Greenwood et al., 2005] Greenwood, J. S., Helm, M., and Gietl, C. (2005). Ricinosomes and endosperm transfer cell structure in programmed cell death of the nucellus during ricinus seed development. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6):2238–2243.

[Haig, 2013] Haig, D. (2013). Kin conflict in seed development: an interdependent but fractious collective. *Annual Review of Cell and Developmental Biology*, 29:189–211.

[Haig and Westoby, 1989] Haig, D. and Westoby, M. (1989). Parent-specific gene expression and the triploid endosperm. *The American Naturalist*, 134(1):147–155.

[Haig and Westoby, 1991] Haig, D. and Westoby, M. (1991). Genomic imprinting in endosperm: its effect on seed development in crosses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis.

Philosophical Transactions of the Royal Society of London B: Biological Sciences, 333(1266):1–13.

[Harper et al., 1970] Harper, J. L., Lovell, P., and Moore, K. (1970). The shapes and sizes of seeds. *Annual review of ecology and systematics*, 1(1):327–356.

[Hehenberger et al., 2012] Hehenberger, E., Kradolfer, D., and Köhler, C. (2012). Endosperm cellularization defines an important developmental transition for embryo development. *Development*, 139(11):2031–2039.

[Hellsten et al., 2013] Hellsten, U., Wright, K. M., Jenkins, J., Shu, S., Yuan, Y., Wessler, S. R., Schmutz, J., Willis, J. H., and Rokhsar, D. S. (2013). Fine-scale variation in meiotic recombination in mimulus inferred from population shotgun sequencing. *Proceedings of the National Academy of Sciences*, 110(48):19478–19482.

[Hoffner and Surti, 2012] Hoffner, L. and Surti, U. (2012). The genetics of gestational trophoblastic disease: a rare complication of pregnancy. *Cancer genetics*, 205(3):63–77.

[Hromadnikova et al., 2013] Hromadnikova, I., Kotlabova, K., Ondrackova, M., Kestlerova, A., Novotna, V., Hympanova, L., Doucha, J., and Krofta, L. (2013). Circulating c19mc microrna in preeclampsia, gestational hypertension, and fetal growth restriction. *Mediators of inflammation*, 2013.

[Joët et al., 2009] Joët, T., Laffargue, A., Salmona, J., Doulbeau, S., Descroix, F., Bertrand, B., De Kochko, A., and Dussert, S. (2009). Metabolic pathways in tropical dicotyledonous albuminous seeds: Coffea arabica as a case study. *New Phytologist*, 182(1):146–162.

[Josefsson et al., 2006] Josefsson, C., Dilkes, B., and Comai, L. (2006). Parent-dependent loss of gene silencing during interspecies hybridization. *Current Biology*, 16(13):1322–1328.

[Kanayama et al., 2002] Kanayama, N., Takahashi, K., Matsuura, T., Sugimura, M., Kobayashi, T., Moniwa, N., Tomita, M., and Nakayama, K. (2002). Deficiency in p57kip2 expression induces preeclampsia-like symptoms in mice. *Molecular human reproduction*, 8(12):1129–1135.

[Kawahara et al., 2007] Kawahara, M., Wu, Q., Takahashi, N., Morita, S., Yamada, K., Ito, M., Ferguson-Smith, A. C., and Kono, T. (2007). High-frequency generation of viable mice from engineered bi-maternal embryos. *Nature biotechnology*, 25(9):1045–1050.

[Kinoshita, 2007] Kinoshita, T. (2007). Reproductive barrier and genomic imprinting in the endosperm of flowering plants. *Genes & genetic systems*, 82(3):177–186.

[Klosinska et al., 2016] Klosinska, M., Picard, C. L., and Gehring, M. (2016). Conserved imprinting associated with unique epigenetic signatures in the arabidopsis genus. *Nature Plants*, 2:16145.

[Kozlowski, 1972] Kozlowski, T. (1972). *Seed Biology Vol. 1*. New York: Academic Press.

[Kradolfer et al., 2013] Kradolfer, D., Wolff, P., Jiang, H., Siretskiy, A., and Köhler, C. (2013). An imprinted gene underlies postzygotic reproductive isolation in arabidopsis thaliana. *Developmental cell*, 26(5):525–535.

[Lafon-Placette et al., 2017] Lafon-Placette, C., Johannessen, I. M., Hornslien, K. S., Ali, M. F., Bjerkan, K. N., Bramsiepe, J., Glöckle, B. M., Rebernig, C. A., Brysting, A. K., Grini, P. E., et al. (2017). Endosperm-based hybridization barriers explain the pattern of gene

flow between arabidopsis lyrata and arabidopsis arenosa in central europe. *Proceedings of the National Academy of Sciences*, page 201615123.

[Langmead and Salzberg, 2012] Langmead, B. and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nature methods*, 9(4):357–359.

[Leblanc et al., 2002] Leblanc, O., Pointe, C., and Hernandez, M. (2002). Cell cycle progression during endosperm development in zea mays depends on parental dosage effects. *The Plant Journal*, 32(6):1057–1066.

[Li and Dickinson, 2009] Li, N. and Dickinson, H. G. (2009). Balance between maternal and paternal alleles sets the timing of resource accumulation in the maize endosperm. *Proceedings of the Royal Society of London B: Biological Sciences*, page rspb20091209.

[Luo et al., 2011] Luo, M., Taylor, J. M., Spriggs, A., Zhang, H., Wu, X., Russell, S., Singh, M., and Koltunow, A. (2011). A genome-wide survey of imprinted genes in rice seeds reveals imprinting primarily occurs in the endosperm. *PLoS Genet*, 7(6):e1002125.

[Martinez and Köhler, 2017] Martinez, G. and Köhler, C. (2017). Role of small rnas in epigenetic reprogramming during plant sexual reproduction. *Current Opinion in Plant Biology*, 36:22–28.

[McGrath and Solter, 1984] McGrath, J. and Solter, D. (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell*, 37(1):179–183.

[Medel et al., 2003] Medel, R., Botto-Mahan, C., and Kalin-Arroyo, M. (2003). Pollinator-mediated selection on the nectar guide phenotype in the andean monkey flower, mimulus luteus. *Ecology*, 84(7):1721–1732.

[Monk, 2015] Monk, D. (2015). Genomic imprinting in the human placenta. *American journal of obstetrics and gynecology*, 213(4):S152–S162.

[Olsen, 2004] Olsen, O.-A. (2004). Nuclear endosperm development in cereals and arabidopsis thaliana. *The Plant Cell*, 16(suppl 1):S214–S227.

[Oneal et al., 2014] Oneal, E., Lowry, D. B., Wright, K. M., Zhu, Z., and Willis, J. H. (2014). Divergent population structure and climate associations of a chromosomal inversion polymorphism across the mimulus guttatus species complex. *Molecular ecology*, 23(11):2844–2860.

[Oneal et al., 2016] Oneal, E., Willis, J. H., and Franks, R. G. (2016). Disruption of endosperm development is a major cause of hybrid seed inviability between mimulus guttatus and mimulus nudatus. *New Phytologist*.

[Patten et al., 2014] Patten, M., Ross, L., Curley, J., Queller, D. C., Bonduriansky, R., and Wolf, J. (2014). The evolution of genomic imprinting: theories, predictions and empirical tests. *Heredity*, 113(2):119–128.

[Pennington et al., 2008] Pennington, P. D., Costa, L. M., Gutierrez-Marcos, J. F., Greenland, A. J., and Dickinson, H. G. (2008). When genomes collide: aberrant seed development following maize interploidy crosses. *Annals of botany*, 101(6):833–843.

[Ramsey and Schemske, 1998] Ramsey, J. and Schemske, D. W. (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, 29(1):467–501.

[Rebernig et al., 2015] Rebernig, C. A., Lafon-Placette, C., Hatorangan, M. R., Slotte, T., and Köhler, C. (2015). Non-reciprocal interspecies hybridization barriers in the capsella genus are established in the endosperm. *PLoS Genet*, 11(6):e1005295.

[Schon and Nodine, 2017] Schon, M. A. and Nodine, M. (2017). Widespread contamination of arabidopsis embryo and endosperm transcriptome datasets. *The Plant Cell*, pages tpc–00845.

[Scott et al., 1998] Scott, R. J., Spielman, M., Bailey, J., and Dickinson, H. G. (1998). Parent-of-origin effects on seed development in arabidopsis thaliana. *Development*, 125(17):3329–3341.

[Smith et al., 2017] Smith, R. D., Kinser, T. J., Smith, G. D., and Puzey, J. R. (2017). A likelihood ratio test for changes in homeolog expression bias. *bioRxiv*, page 119438.

[Spencer and Clark, 2014] Spencer, H. and Clark, A. (2014). Non-conflict theories for the evolution of genomic imprinting. *Heredity*, 113(2):112–118.

[Stanton et al., 2017] Stanton, K. A., Edger, P. P., Puzey, J. R., Kinser, T., Cheng, P., Vernon, D. M., Forsthoefel, N. R., and Cooley, A. M. (2017). A whole-transcriptome approach to evaluating reference genes for quantitative gene expression studies: A case study in mimulus. *G3: Genes, Genomes, Genetics*, 7(4):1085–1095.

[Stewart-Cox et al., 2004] Stewart-Cox, J., Britton, N. F., and Mogie, M. (2004). Endosperm triploidy has a selective advantage during ongoing parental conflict by imprinting. *PROCEEDINGS-ROYAL SOCIETY OF LONDON B*, 271:1737–1744.

[Tunster et al., 2010] Tunster, S. J., Tycko, B., and John, R. M. (2010). The imprinted phlda2 gene regulates extraembryonic energy stores. *Molecular and cellular biology*, 30(1):295–306.

[Vallejo-Marin, 2012] Vallejo-Marin, M. (2012). Mimulus peregrinus (phrymaceae): A new british allopolyploid species. *PhytoKeys*, 14:1.

[Vallejo-Marin et al., 2016] Vallejo-Marin, M., Cooley, A. M., Lee, M. Y., Folmer, M., McKain, M. R., and Puzey, J. R. (2016). Strongly asymmetric hybridization barriers shape the origin of a new polyploid species and its hybrid ancestor. *American journal of botany*, 103(7):1272–1288.

[Vallejo-Marin and Lye, 2013] Vallejo-Marin, M. and Lye, G. C. (2013). Hybridisation and genetic diversity in introduced mimulus (phrymaceae). *Heredity*, 110(2):111–122.

[Wang and Köhler, 2017] Wang, G. and Köhler, C. (2017). Epigenetic processes in flowering plant reproduction. *Journal of Experimental Botany*, page erw486.

[Waters et al., 2013] Waters, A. J., Bilinski, P., Eichten, S. R., Vaughn, M. W., Ross-Ibarra, J., Gehring, M., and Springer, N. M. (2013). Comprehensive analysis of imprinted genes in maize reveals allelic variation for imprinting and limited conservation with other species. *Proceedings of the National Academy of Sciences*, 110(48):19639–19644.

[Waters et al., 2011] Waters, A. J., Makarevitch, I., Eichten, S. R., Swanson-Wagner, R. A., Yeh, C.-T., Xu, W., Schnable, P. S., Vaughn, M. W., Gehring, M., and Springer, N. M. (2011). Parent-of-origin effects on gene expression and dna methylation in the maize endosperm. *The Plant Cell*, 23(12):4221–4233.

[Westoby and Rice, 1982] Westoby, M. and Rice, B. (1982). Evolution of the seed plants and inclusive fitness of plant tissues. *Evolution*, pages 713–724.

[Witmer, 1937] Witmer, S. W. (1937). Morphology and cytology of vallisneria spiralis l. *American Midland Naturalist*, pages 309–333.

[Wolff et al., 2015] Wolff, P., Jiang, H., Wang, G., Santos-González, J., and Köhler, C. (2015). Paternally expressed imprinted genes establish postzygotic hybridization barriers in arabidopsis thaliana. *Elife*, 4:e10074.

[Xie et al., 2014] Xie, L., Mouillet, J.-F., Chu, T., Parks, W. T., Sadovsky, E., Knöfler, M., and Sadovsky, Y. (2014). C19mc micrornas regulate the migration of human trophoblasts. *Endocrinology*, 155(12):4975–4985.

[Xu et al., 2014] Xu, W., Dai, M., Li, F., and Liu, A. (2014). Genomic imprinting, methylation and parent-of-origin effects in reciprocal hybrid endosperm of castor bean. *Nucleic Acids Research*, 42(11):6987.

[Yan et al., 2014] Yan, D., Duermeyer, L., Leoveanu, C., and Nambara, E. (2014). The functions of the endosperm during seed germination. *Plant and Cell Physiology*, page pcu089.

[Zhang et al., 2015] Zhang, M., Li, N., He, W., Zhang, H., Yang, W., and Liu, B. (2015). Genome-wide screen of genes imprinted in sorghum endosperm and the roles of allelic differential cytosine methylation. *The Plant Journal*.