Ollscoil na Gaillimhe
University of Galway

Provided by the author(s) and University of Galway in accordance with publisher policies. Please cite the published version when available.

Title	Genome dosage, genetic hybridity and crop improvement
Author(s)	Hallahan, Brendan
Publication Date	2020-03-02
Publisher	NUI Galway
Item record	http://hdl.handle.net/10379/15833

Downloaded 2024-04-17T20:14:01Z

Some rights reserved. For more information, please see the item record link above.

Genome Dosage, Genetic Hybridity and Crop Improvement

Volume 1 of 1

Brendan Hallahan, BSc., MSc.

A thesis submitted to National University of Ireland Galway for the degree of

Doctor of Philosophy

Plant and AgriBioscience Research Centre (PABC), Ryan Institute,

College of Science, School of Natural Sciences

Supervisor: Prof Charles Spillane

Head of School: Prof Ciaran Morrison

Acknowledgements

I wish to thank the Irish Research Council for their financial support through a Government of Ireland Postgraduate Scholarship (Grant No GOIPG/2014/1021).

I wish to thank Prof Charles Spillane for giving me the opportunity to carry out research in the Genetics and Biotechnology Laboratory of the Plant and AgriBioscience Research Centre (PABC). Your support and advice were invaluable over the past 4 years. Likewise, I want to extend sincere thanks to my Graduate Research Committee who offered excellent feedback on my research and progress over the course of my PhD.

Thank you to all lab mates past and present. Your friendship and assistance made all this possible and it was a pleasure working with you every day.

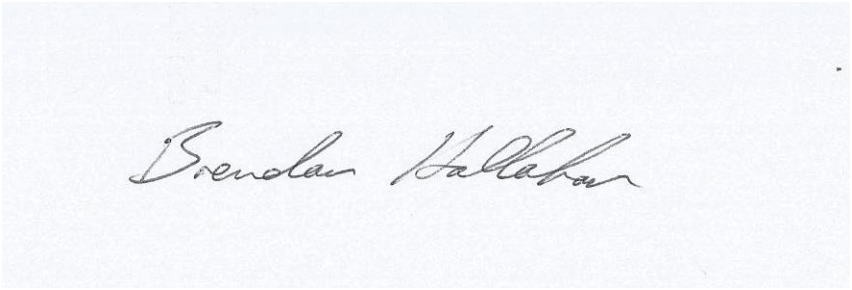
Finally, a huge thank you to my parents, Carol and Tom, my brothers, Eoin and Cian, and my extended family. You are always available for work- and life-advice, loving praise or fair criticism, and instilled in me a sense of purpose throughout my time in Galway. Thank you, I will never forget it.

Table of Contents

Author's declaration

I declare this thesis has not been submitted as part of another degree, either at NUI Galway or another University, and that all data generated along with analyses undertaken are my own unless explicitly stated otherwise.

Signature:



Brendan Hallahan

List of publications generated, and oral presentations, workshops, and conferences attended

Publications Generated

Brendan F. Hallahan, Eva Fernandez-Tendero, Antoine Fort, Peter Ryder, Gilles Dupouy, Marc Deletre, Edna Curley, Galina Brychkova, Britta Schulz and Charles Spillane (2018) Hybridity has a greater effect than paternal genome dosage on heterosis in sugar beet (*Beta vulgaris*). *BMC Plant Biology* **18**(1):120. doi: 10.1186/s12870-018-1338-x.

Marijke Hummel, **Brendan F. Hallahan**, Galina Brychkova, Julian Ramirez-Villegas, Veronica Guwela, Bartholomew Chataika, Edna Curley, Liam Morrison, Elise Talsma, Steve Beebe, Andy Jarvis, Rowland Chirwa, and Charles Spillane (2018) Decline in nutritional quality of common bean under climate change induced drought stress in Africa. *Nature Scientific Reports* **8**:16187. doi: 10.1038/s41598-018-33952-4.

Oral Presentations

Scientific Conferences

"Investigating the relative contributions of hybridity and genome dosage to heterosis in sugar beet (*Beta vulgaris*)". Irish Plant Scientists Association Meeting (IPSAM), April 2016, Trinity College Dublin, Ireland.

Non-scientific conferences

"Environmental Clichés" NUI Galway Climate Congress 2015, March 2015, NUI Galway, Ireland

"Climate change and Ireland – is anyone listening?" NUI Galway Climate Congress 2016, April 2016, NUI Galway, Ireland

"Divestment – what's the point?" NUI Galway Climate Congress 2017, March 2017, NUI Galway, Ireland

"How useful is the common bean crop in fighting hidden hunger in sub-Saharan Africa?" Pint of Science Festival, May 2018, Campbell's Tavern, Cloughanover, Headford, Co. Galway, Ireland

Workshops attended

European Plant Phenotyping Network (EPPN) Plant Phenotyping Workshop. March 2015, Institute of Biological Environmental and Rural Sciences (IBERS), Aberystwyth University, Wales

Conferences attended

International Association for Plant Biotechnology (IAPB) Congress. August 2018, Convention Centre Dublin, Ireland

Summary

Plant breeding is focussed on introducing novel traits in plants, as well as enhancing or repressing traits already present. A noble objective in agriculture is to produce more food per unit area with overall fewer resources. Plant breeders have a central role to play in achieving this objective. My PhD research investigates the genetic and physiological consequences of two different plant breeding phenomena: *heterosis* and *polyploidy*.

Heterosis occurs when different parents of the same species are interbred and the offspring display enhanced characteristics, such as higher yield. *Polyploidy* is when an organism possesses more than two sets of chromosomes. Humans, for example, are *diploid* organisms with just two sets of chromosomes (we inherit one set from our mother and one from our father), but *polyploidy* is common in both wild and cultivated plants.

In **chapter one** of this thesis I discuss the most relevant and most recent literature. Next, I present five research chapters composed of varied research methodologies, from classical botany to the latest gene sequencing technologies. In **chapter two** I investigate if the *heterosis* phenomenon in sugar beet can be improved by increasing the number of paternally inherited chromosome sets. In **chapter three** I utilise *tetraploid* plants to investigate for the presence of *progressive heterosis* in *Arabidopsis thaliana*. In **chapter four** I investigate if a novel method of whole-genome elimination in plant breeding is influenced by nuclear and cytoplasm effects. In **chapter five** I investigate whether the induction of *polyploidy*, including whether a chromosome set is inherited maternally or paternally, can enhance salt stress tolerance in *Arabidopsis thaliana*. In **chapter six** I model the impact of climate change on land suitability of related lower and higher ploidy grasses.

Chapter 1. Introduction to polyploidy in plants

> "Polyploidy is not only one of the best known of evolutionary processes; in addition, it is the most rapid method known of producing radically different but nevertheless vigorous and well-adapted genotypes. For these reasons, it is the large-scale evolutionary process most suitable for development as a technique in plant breeding."
>
> - G. Ledyard Stebbins, Jr., 1950. *Variation and Evolution in Plants*, p.369

Polyploidy is a state of possessing more than two sets of chromosomes within the nucleus. The number of chromosome sets can be determined directly by viewing chromosomes via chromosome spreading/karyotyping (Fukui et al. 1998; Deng et al. 2003) or indirectly via flow cytometry (Doležel et al. 2007). The polyploidy phenomenon is considerably more widespread among plants than animals (Vandel 1938; Otto and Whitton 2000; Davoli and Lange 2011). Polyploidy can arise from abnormalities in either meiosis, mitosis, or fertilisation. However, such abnormalities (1) are a major evolutionary force in plants, and (2) can be utilised by plant breeders to improve crops for human consumption. What follows is a description of how polyploids are formed, the consequences for growth in a newly formed polyploid (a 'neopolyploid'), and the application of polyploidy for crop improvement, including as an aid to genetic hybridisation.

Polyploid formation in plants

A diploid plant, denoted 2x (where 'x' indicates the ploidy level in the nucleus of a standard somatic cell), can become a polyploid plant through two chief mechanisms: sexual reproduction from unreduced gametes (2*n* instead of simply *n*, where '*n*' indicates the number of chromosomes in the gametes) (Kreiner et al. 2017) or somatic chromosome doubling. A third mechanism, polyspermy, when an egg is fertilised by more than one sperm cell, is another (albeit rarer) possibility (Grossniklaus 2017; Nakel et al. 2017). See *Box 1* for an explanation of the different mechanisms of polyploid formation in plants. Regardless of origin, this duplication of an entire genome is given the moniker a 'whole-genome duplication event', or, WGD event for short. All plant genomes investigated in detail with genome sequencing show evidence of recurrent WGD (Jiao et al. 2011; Li et al. 2015; Van de Peer et al. 2017), the key evidence being the high proportion of duplicated genes as a proportion of total genes within the genome and tendency to increased chromosome numbers in some lineages (Haldane 1933; Otto and Whitton 2000; Blanc and Wolfe 2004). Polyploid plants may give the appearance of

being evolutionary 'dead ends' as the majority of polyploid lineages appear to go extinct (Arrigo and Barker 2012), but (1) polyploid plants may have a short-term survival rate and hybridise with diploid progenitors, potentially passing on novel traits into the population, and (2) newly formed polyploids may possess traits that allow expansion and establishments into new ecological niches.

Before further discussion it is necessary to define the two classes of polyploid organisms – autopolyploids and allopolyploids. The former are polyploids with identical genomes, whereas the latter are polyploids with different genomes (hence, allopolyploids are also hybrids by definition).

Box 1. Mechanisms of polyploid formation in plants

Unreduced gametes: Cell division through meiosis is necessary for sexual reproduction. Through DNA division and replication, a sex cell is produced with half the chromosome number of the parent cell and, in a heterozygous plant, will contain new genetic variability. Defects in the meiotic process can lead to unviable meiocyte cells and thus sterile gametes. However, some meiotic abnormalities in plants have been documented which produce viable, unreduced gametes. For example, the *dyad* meiotic mutant in *Arabidopsis thaliana* fails to correctly separate sister chromatids in meiosis I during female gamete formation, resulting in a small number of viable, unreduced egg cells (2*n*) which can be crossed with pollen (*n*) to produce triploid offspring (Mercier et al. 2001; Ravi et al. 2008). Another meiotic mutant example in *A. thaliana,* but for male gamete formation, is the *Atps1* mutant. The physical separation of chromosomes in meiosis II is organised by spindles and the final product should be four haploid (*n*) cells (microspores) joined together as a tetrad. Mutations at the *parallel spindles* locus results in abnormal orientation of spindles in meiosis II and the result is a mixture of viable *n* and 2*n* pollen which can produce diploid and triploid plants (d'Erfurth et al. 2008). Finally, observations of unreduced gamete formation have been documented in newly formed interspecific F1 hybrids (Ramsey and Schemske 1998) and as a response to environmental stimuli such as cold temperatures (Belling 1925) and viral infections (Kostoff 1933). Thus, genetic hybridity, genetic mutations and plant responses to natural environmental factors can influence unreduced gamete formation.

Somatic chromosome doubling: Endopolyploidy – DNA replication inside a nucleus without cytokinesis – occurs at different frequencies in different tissues, suggesting larger cell size is

advantageous for certain functions (Van de Peer et al. 2017). Endopolyploidy in non-meristematic tissue could be a source of new polyploid plants, e.g. diploid broad bean (*Vicia faba*) stem tissue has been shown to contain extensive endopolyploidy (Coleman 1950) and induction of tumours on the stems encourages high-polyploid cell growth (Therman 1956). New growth from such wounds or tumours could be polyploid shoots, hence an otherwise diploid plant may support a tetraploid organ (Ramsey and Schemske 1998). The first product of fertilisation – the zygote – may also experience spontaneous genome doubling. This has been induced artificially in maize (Randolph 1932) and various wheat species (*Triticum*) and rye (*Secale cereale*) (Dorsey 1936) by exposing pollinated female organs to high temperatures (38-48 °C) for a period of 30 minutes to 1 hour. The unicellular, diploid, rapidly dividing zygote is affected by the heat stress and abnormalities in mitosis can produce polyploid cells. The contribution of this heat effect to the evolutionary formation of polyploids remains uncertain because such a short, sharp temperature increase in natural settings is unlikely, except for perhaps a nearby fire.

Polyspermy: In animals, fertilisation of an egg by multiple sperm (polyspermy) is thought to be lethal. Evolutionary selective pressures to ensure fertilisation through monospermy means a 'block' around the egg cell wall instantly emerges upon fertilisation with a sperm cell (Wong and Wessel 2005). In plants, identifying a similar mechanism has been difficult to determine *in vivo*, yet polyspermy events were deemed unlikely. However, two recent publications demonstrate that polyspermy events are possible in maize (Grossniklaus 2017) and *A. thaliana* (Nakel et al. 2017).

10 years ago, research documented that the central cell of *A. thaliana* can fuse with more than one sperm (Scott et al. 2008). This was achieved by pollinating a wild type female with pollen from *tetraspore* mutant. This *tes* mutant can deliver pollen tubes with multiple sperm, instead of the traditional two. Multiple fertilisation of the egg cell was not found. Polyspermy of the egg cell has now been documented to occur in *A. thaliana* at a rate of 0.012% (Nakel et al. 2017). The authors succeeded by pollinating a wild type female with two pollen donors that each harbour a different T-DNA insertion, only tri-parental (triploid) progeny that inherit both constructs are resistant to BASTA herbicide. Pollen delivery from just one of the two male parents produces seedlings susceptible to the herbicide.

In maize, it is known that more than one pollen tube can occasionally reach an ovule, leading to hetero-fertilisation (Sprague 1929) where the egg cell and the central cell may fuse with different sperm from different pollen tubes. Grossniklaus' experiment used the

phenomenon of hetero-fertilisation plus knowledge of the strict 2:1 maternal:paternal genome ratio in the endosperm to examine the likelihood of polyspermy in maize. A tetraploid maize line (W23), which does not produce anthocyanin (*r*), was used as a female parent with a pollen mixture from two diploid male parents *(R-stippled* and *R-navajo*) which possess distinct endosperm pigment markers. Single fertilisation of the central cell would produce a 4:1 maternal:paternal genome ratio in the endosperm, and hence would abort. Double fertilisation of the central cell from both parents, however, would produce 4:2 maternal:paternal genome ratio and produce a viable kernel with both pigment markers visible. Of the 23 kernels which had both pigment markers present, a single tri-parent (tetraploid) progeny was discovered. These two recent publications suggest that polyspermy events giving rise to tri-parental offspring may contribute to plant polyploidy evolution.

The consequences for gene expression, growth and fitness when a plant undergoes polyploidisation

Increasing the genome content will lead to an immediate increase in cell volume. This volume increase is believed to be a DNA-content-mediated change rather than any regulatory pathway response (Gregory 2001; Storchova and Pellman 2004). Polyploids reliably show an increase in cell size (particularly stomatal cell size), flower size, seed size and a delay in flowering time across taxa (Kondorosi et al. 2000; Ramsey and Schemske 2002; Soltis et al. 2004). Increased cell volume could change the shape of the cell and affect intra-cellular interactions, for example between chromatin and the nuclear envelope. Immediately upon a WGD event, a plant may suffer from a fitness perspective due to instability in mitosis and meiosis, and changes in gene expression (this assumes that these processes are optimised or near-optimised at the diploid level).

Gene expression and growth

Assuming expression patterns of diploid parent(s) have been optimised through selection, new gene expression patterns will be disadvantageous. Increasing the number of chromosome sets will equally increase all gene copies, however a corresponding, linear increase in gene expression is not guaranteed. For example, using a ploidy series of closely related (but not genetically identical) maize plants, Guo and colleagues demonstrated that leaf gene expression (for 18 genes) at the haploid, diploid, triploid and tetraploid level were approximately equal (Guo et al. 1996). In cabbage (*Brassica oleracea*) fully homozygous haploid, corresponding

diploid and autotetraploid genotypes display similar gene expression in photosynthetic tissues (Albertin et al. 2005). Likewise, in potato (*Solanum tuberosum*) O37 series of haploid, diploid and tetraploid plants, leaf and root tip gene expression was markedly similar across ploidy levels (Stupar et al. 2007). In contrast, for one of two *A. thaliana* genotypes, the induction of autopolyploidy caused differential expression of several hundred genes (Yu et al. 2010). These findings suggest that upon WGD, when alleles are derived from a single species, some gene expression is under regulatory control that does not correspond to simple increases in copy number and may depend on the genotype of the diploid progenitor. In newly formed allotetraploids, gene expression changes have been shown to be more prominent. In *Arabidopsis* for example, gene expression changes occur immediately and stochastically following allopolyploidisation and some homoeologous gene silencing is only displayed in later generations. Reactivation of a few silenced genes was achieved by manipulation of DNA methylation, suggesting that epigenetic remodelling of homoeologous genes in allopolyploids may play an important evolutionary role (Wang et al. 2004). Further research using *Arabidopsis* as a model allopolyploid revealed significant instability among certain transposable elements in allotetraploids (Madlung et al. 2005). Such 'jumping genes' are considered to be an important mechanism in plant genome evolution (Bennetzen 2005). For example, methylation changes to genome regions flanking transposable elements have been implicated in *Spartina* response to allopolyploidisation, also (Christian et al. 2009).

Duplicated gene copies are ready to be acted upon by natural selection. The most likely outcome of duplicate genes appears to be gene loss (Maere et al. 2005), however there is also potential for neofunctionalization or subfunctionalization. That is, a gene copy may take on a new function (neofunctionalization), or, retain only part of its original function (subfunctionalization) in the polyploid plant (Comai 2005). See *Box 2* for an example of research (Adams et al. 2003) showing evidence of immediate subfunctionalization upon WGD. Of note, finding evidence for a "new" gene function in neofunctionalization is difficult. When one gene of a gene pair is said to have evolved a "new" function, it may simply be a re-expression of an ancestral function (Cheng et al. 2018). There appears to be bias in what duplicated genes are retained and what are lost: proteins involved in transcription are more likely to be retained in duplicate whereas genes encoding enzymes are often lost (Seoighe and Gehring 2004). In some cases, there can be whole-genome expression bias following allopolyploidisation. In *Arabidopsis* allopolyploids there is genome-wide expression dominance of *A. arenosa* over *A. thaliana* (Wang et al. 2006). Similar genome-wide transcriptome analysis in allohexaploid (AABBDD) wheat (*T. aestivum*) does not show strong

evidence for genome-wide expression dominance; instead it reveals that a small percentage of genes show non-additive expression (Chagué et al. 2010), which is interesting considering that domesticated wheat is said to phenotypically resemble its A-genome progenitor einkorn wheat (*T. urartu*). Thus, retention of parental genes, additive and non-additive expression of these genes, and favouring one genome over the other in allopolyploid plants appears to be parental genome-specific.

Assuming there is no fatal intergenomic incompatibility, a newly formed allopolyploid plant will establish and will have to compete with its parent(s). An allopolyploid plant may display heterosis (a.k.a. hybrid vigour). This is the phenomenon of an F1 hybrid displaying enhanced characteristics over one or both parents. An improvement in size, growth rate or biotic/abiotic stress tolerance could allow a new allopolyploid to establish in its current ecological niche or extend to a new one. For example, the vigorous allopolyploid grass hybrid *Spartina townsendii* has a wider range than its progenitors (Baumel et al. 2001). Furthermore, starting with the assumption that diploidy is the ancestral state of plants, Pandit and colleagues found that invasive plant species are disproportionately polyploid while endangered plant species are more likely to be diploid. Thus, increasing the ploidy level of a plant increases its invasiveness potential (Pandit et al. 2011). The genotypic, and potentially phenotypic, diversity given by WGD may not be useful in a plant's current ecological niche – and thus the evolutionary success of a neopolyploid plant may depend upon the availability of a new ecological niche (Van de Peer et al. 2017). The argument that environments which experience frequent disturbances may favour the establishment of polyploids was first articulated over 40 years ago (Stebbins 1971). There is some evidence to support this. For example, autotetraploid cytotypes of the Mediterranean shrub *Santolina pectinata* occupy more ecologically disturbed areas than the diploid cytotype, which tends to occupy a wider range (Rivero-Guerra 2008). The mirror image of this argument is that long-term environmental stability selects against polyploid establishment – this has been demonstrated at the southern tip of Africa, which experiences a very stable environment all the while maintaining high species richness through primarily diploid lineages and a very low level of polyploidy (Oberlander et al. 2016).

Plant Reproduction

The fitness of an organism is dependent upon its reproductive success (Michod 2000). A plant's transition from diploidy to polyploidy presents challenges for meiosis and reproduction (and hence fitness). Thus a neopolyploid could face reproductive isolation from diploids (Ramsey and Schemske 2002). For example, consider segregation of chromosomes during meiosis. As

diploids contain two homologous chromosomes, the meiotic association of chromosomes during meiosis will produce a bivalent structure (the nomenclature maintains the Latin prefix *bi* even though four sister chromatids are present). In autopolyploid meiosis, however, the meiotic association may lead to more than two chromosomes (multivalents) or a single chromosome (univalent), both of which are considered meiotic irregularities which could lead to non-functional gametes. Similarly, in triploids trivalents may be produced (Jackson 1976). Random segregation of multiple chromosomes from unbalanced univalents, trivalents, and multivalents will give rise to an aneuploid gamete. The viability of the resulting zygote is species-specific, for example *A. thaliana* triploids can establish and are somewhat fertile (Steinitz-Sears 1963; Henry et al. 2005; Duszynska et al. 2013; Fort et al. 2016), giving credence to the 'triploid bridge' hypothesis that triploids may act as intermediaries between the diploid and tetraploid state (Ramsey and Schemske 1998; Köhler et al. 2010). In *Primula*, however, diploid plants crossed with tetraploid plants fail to produce seed (Woodell and Valentine 1961), and likewise in *Poinsettia* most crosses between diploid and tetraploid plants are unsuccessful (Milbocker and Sink 1969). Some polyploids may overcome the sexual hybridisation barrier by asexual reproduction, i.e. through vegetative reproduction or apomictic reproduction (embryo formation from maternal tissue without fertilisation) (Richards 2003; Bicknell and Koltunow 2004). The invasiveness of certain polyploid plants has been attributed to such asexual means of reproduction (Chapman et al. 2000; Elvira Hörandl 2008; Bailey et al. 2009). Curiously, the problem of aneuploid gametes in polyploids may be overcome through time - some autotetraploid *A. thaliana* genotypes begin to show diploid-like meiosis after several generations, i.e. bivalent formation of chromosomes. This is a potential mechanism for 'diploidisation' in the evolution of plants – a polyploid returning to the diploid level (Santos et al. 2003).

In autopolyploids, chromosomes are just as likely to pair and recombine with any of the available homologs, i.e. alleles at a certain locus on the homologous chromosomes should segregate at random. This contrasts with allopolyploids, which display distinct inheritance patterns during meiosis: chromosomes tend to pair and crossover with the homolog(s) from the same parental species (or the mostly closely related homolog), i.e. it is non-random (Ramsey and Schemske 2002; Stift et al. 2008; Lloyd and Bomblies 2016). Thus, an allotetraploid plant will largely display 'disomic' inheritance (independent segregation of sub-genomes), whereas an autotetraploid plant will largely display 'tetrasomic' (or 'multisomic') inheritance (equal frequencies of all potential allelic combinations are possible). A genetic mechanism for homologous chromosome pairing has been identified in allohexaploid wheat (Sears 1976;

Greer et al. 2012). Regarding filial frequencies, heterozygous autotetraploids will maintain higher heterozygosity levels than their diploid progenitors due to tetrasomic inheritance (Moody et al. 1993). The flip side to this fact is that attaining genetic purity in an autotetraploid line is more difficult than in a diploid line due to the presence of four different alleles at each locus. This has consequences for plant evolution as well breeding programs (see the Punnett Squares in *Box 3*).

WGD may change a plant's reproduction preference from out-crossing to self-pollination. For a number of genera it has been shown that WGD can lead to a transition from self-incompatibility to self-compatibility suggesting that a newly formed tetraploid in a diploid population may not need to mate with its diploid ancestor to survive (avoiding a potentially unstable triploid embryo) (Miller and Venable 2000). A more detailed analysis of this phenomenon revealed that allopolyploid flowering plants do indeed incline toward selfing, but autopolyploid flowering plants display higher outcrossing rates (Husband et al. 2008). Insect-pollinated plants may benefit from polyploidisation if floral morphology changes to encourage more generalist pollinators, as suggested in *Nicotiana* allopolyploids (McCarthy et al. 2016), thus contributing to plant speciation.

Box 2. Example of polyploidy leading to immediate, broad changes in gene expression, including an example of subfunctionalization

Cotton (*Gossypium*) is an allotetraploid plant (AADD), which originated from the hybridisation of two diploid ancestors *G. herbaceum* (AA) and *G. raimondii* (DD) (Wendel and Cronn 2003). Pairs of genes in the same species that originated by speciation and brought together through allopolyploidisation are called homoeologous genes (Glover et al. 2016). Adams and colleagues were the first to study the expression of homoeologous gene pairs in different plant organs between natural polyploid cotton and a newly formed ('synthetic') polyploid genotype. First, as discovered in *Arabidopsis* (Wang et al. 2004) and wheat (Kashkush et al. 2002), Adams *et al.* found some genes duplicated through polyploidy were silenced or experienced differential expression. This was identified in ovules of varying ages. Of note, there was no preferential expression of genes derived from either the A-genome or D-genome in allotetraploid ovules. Next, the authors studied homoeologous gene pair expression in various organs (root, stem, leaf, petal, etc.) of both natural and synthetic allotetraploids. Interestingly, they uncovered evidence of reciprocal silencing of some genes, i.e. there is silencing of one member of a duplicated gene pair in some organs while the other gene copy is silenced in other organs. For example, in both the natural and synthetic allotetraploid genotypes, *alcohol*

dehydrogenase A (*adhA*) expression in the carpel was almost exclusively derived from the A-genome homoelog whereas *adhA* expression in the stamens was derived from the D-genome homoelog. This finding suggests that (1) the genomic response to WGD can differ between tissues, (2) rapid subfunctionalization is a potential consequence of allopolyploidisation, and (3) this subfunctionalization can persist through the generations.

Box 3. Disomic and tetrasomic inheritance in diploid and polyploid plants

Assuming the breeding objective in a certain crop is to maximise heterozygosity, the utilisation of tetrasomic inheritance can be beneficial. For example, selfing a heterozygous diploid plant, with genotype (AB), will give rise to progeny in the ratio of 1(AA):2(AB):1(BB) assuming full disomic inheritance:

♀ \ ♂	**A**	**B**
A	AA	AB
B	AB	BB

However, the same heterozygous genotype at the autotetraploid level (AABB), assuming full tetrasomic inheritance, will produce a large increase in the number of heterozygous progeny of various genotypes, specifically 1(AAAA):34 heterozygotes:1(BBBB):

♀ \ ♂	**AA**	**AB**	**AB**	**AB**	**AB**	**BB**
AA	AAAA	AAAB	AAAB	AAAB	AAAB	AABB
AB	ABAA	ABAB	ABAB	ABAB	ABAB	ABBB
AB	ABAA	ABAB	ABAB	ABAB	ABAB	ABBB
AB	ABAA	ABAB	ABAB	ABAB	ABAB	ABBB
AB	ABAA	ABAB	ABAB	ABAB	ABAB	ABBB
BB	BBAA	BBAB	BBAB	BBAB	BBAB	BBBB

Assuming the breeding objective in a certain crop is to remove a deleterious allele at a certain locus (i.e. achieve complete dominance) tetrasomic inheritance is now a disadvantage. For example, let us assume dominant allele *A* shows incomplete dominance over recessive allele *a*.

Upon selfing, segregation at a heterozygous gene locus *Aa* at the diploid level will segregate in the ratio of 3:1 dominant:recessive. ¼ of the progeny will show complete dominance:

♀ \ ♂	**A**	**a**
A	AA	Aa
a	aA	aa

However, the same heterozygous gene locus at the autotetraploid level *AAaa* would produce far more undesirable genotypes harbouring the recessive allele, making selection more difficult. $^{1}/_{36}$ of the progeny will show complete dominance:

♀ \ ♂	**AA**	**Aa**	**Aa**	**Aa**	**Aa**	**aa**
AA	AAAA	AAAa	AAAa	AAAa	AAAa	AAaa
Aa	AaAA	AaAa	AaAa	AaAa	AaAa	Aaaa
Aa	AaAA	AaAa	AaAa	AaAa	AaAa	Aaaa
Aa	AaAA	AaAa	AaAa	AaAa	AaAa	Aaaa
Aa	AaAA	AaAa	AaAa	AaAa	AaAa	Aaaa
aa	aaAA	aaAa	aaAa	aaAa	aaAa	aaaa

Thus, polysomic inheritance has important consequences for genetic variance and inbreeding depression.

Contribution of Polyploidy to Hybrid Crop Improvement

Increasing the genome dosage by WGD and increasing the genetic variance by hybridisation are two mechanisms which contribute to plant speciation in the natural world. These mechanisms can act individually (autopolyploid speciation, homoploid speciation) or together (allopolyploid speciation) (Grant 1981; Soltis et al. 1995; Rieseberg 1997; Ferguson and Sang 2001; Maki and Murata 2001; Christian et al. 2003; Rieseberg et al. 2003; Christian et al. 2004; Mallet 2007; Frajman et al. 2009; Crosby et al. 2014). Plant breeders can use these same mechanisms – manipulating the genome dosage and/or genetic hybridity by intraspecific hybridisation – to improve desirable crop traits. The application of polyploids to plant breeding

has, from a practical perspective, produced some of our most valuable crops (Table 1), and from a biological perspective, contributed to our understanding of heterosis.

Table 1. Some commonly cultivated polyploid crops.

Ploidy	Common name	Scientific name	Reference
3x	Banana	*Musa*	(Christelová et al. 2017)
	Poplar	*Populus*	(Chen et al. 2004; Zhang et al. 2004)
	Tahiti Lime	*Citrus latifolia*	(Rouiss et al. 2018)
	Turfgrass	*Cynodon*	(Hanna and Anderson 2008)
	Watermelon	*Citrullus*	(Wehner 2008)
	Onion	*Allium* × *cornutum*	(Fredotović et al. 2014)
4x	Alfalfa	*Medicago sativa*	(Brouwer and Osborn 1999)
	Arabica coffee	*Coffea arabica*	(Pearl et al. 2004)
	Brown mustard Ethiopian mustard Oilseed rape	*Brassica juncea* *Brassica carinata* *Brassica napus*	(Nagaharu 1935; Gupta et al. 2015; Li et al. 2017)
	Cherry	*Prunus cerasus*	(Hauck et al. 2002)
	Durum wheat	*turgidum* subsp. *durum*	(Trebbi et al. 2011)
	Emmer wheat	*turgidum* subsp. *diccocon*	(Feldman and Kislev 2007)
	Elephant grass	*Pennisetum purpureum*	(Techio et al. 2006)
	Extra-long-staple cotton Upland cotton	*Gossypium barbadense* *Gossypium hirsutum*	(Lacape et al. 2003)
	Peanut	*Arachis hypogaea*	(Burow et al. 2001)
	Potato	*Solanum tuberosum*	(Jansky 2009)
	Roselle	*Hibiscus sabdariffa*	(Wilson and Menzel 1964; Menzel and Wilson 1966)
	Tobacco	*Nicotiana tobaccum*	(Wu et al. 2010)
6x	Bread wheat	*Triticum aestivum* subp. *aestivum*	(M. et al. 2011)
	Jerusalem artichoke	*Helianthus tuberosus*	(Sujatha and Prabakaran 2006)
	Oat	*Avena sativa*	(Loskutov and Rines 2011)
	Plum	*Prunus domestica*	(Decroocq et al. 2004)
	Triticale	× *Triticosecale*	(Ma et al. 2004)
8x	Strawberry	*Fragaria*	(Liston et al. 2014)
	Sugarcane	*Saccharum officinarum*	(Grivet and Arruda 2002)

Crop improvement through heterosis

An F1 hybrid offspring plant is a genetic composite of the maternal cytoplasmic genome and the biparental nuclear genomes inherited from each parent (Hedtke and Hillis 2011). Heterosis has occurred, as described previously, when the F1 hybrid displays enhanced characteristics over one or both parents. The practical definition of heterosis (i.e. from a commercial viewpoint) is the superior performance of an F1 hybrid over the best parent (Acquaah 2009). The application of heterosis for the improvement of specific traits in crops is highlighted by the widespread development of F1 hybrid varieties and their widespread adoption by farmers (Schnable and Springer 2013). While biologists recognise the importance and ubiquity of heterosis to plants, the precise molecular mechanisms behind it are not fully understood. What follows is a description of the prevailing theories that have been experimentally demonstrated to contribute to the genetic basis of heterosis.

Role of Dominance in Heterosis

The dominance theory of heterosis is perhaps best understood as the opposite of inbreeding depression. It posits that hybridising genetically distinct genomes will lead to heterosis. Assuming multiple recessive, slightly deleterious, alleles are present in different parents these will be complemented by a dominant, superior copy from the other parent in the F1 hybrid. Hence, the deleterious effect of these recessive alleles will be masked.

Role of Over-dominance in Heterosis

It was recognised in the early literature that the dominance theory does not account for all heterosis effects observed (East 1936). The over-dominance theory proposes that heterozygosity at a specific locus or loci will lead to a heterotic effect, whereas either homozygous condition will not. In tomato (*Solanum lycopersicum*), considerable yield heterosis has been demonstrated when there is heterozygosity for loss-of-function alleles of *SINGLE FLOWER TRUSS* (*SFT*). Plants homozygous for the *sft* mutation do not display heterosis (Krieger et al. 2010). Linked genes can make determining dominant from over-dominant mechanisms difficult. For example, in the early 1990s over-dominant effects at several quantitative trait loci (QTLs) were attributed to maize (*Zea mays*) heterosis (Stuber et al. 1992). These may in fact be dominant effects. The QTL with the largest effect identified by Stuber and colleagues was later found to be an effect called “pseudo-overdominance”: the QTL was further dissected into (at least) two distinct genes that were in a state of repulsion phase

linkage. These linked loci both contribute to heterosis via dominant mechanisms (Graham et al. 1997; Schnable and Springer 2013).

Role of Epistasis in Heterosis

Further nuance is required when considering the epistasis theory of heterosis. This theory is defined as novel interactions between alleles at two or more different loci leading to a heterotic effect. For example, an over-dominant acting locus may interact with a dominant-acting locus elsewhere in the genome of an F1 hybrid, thus contributing to heterosis (Springer and Stupar 2007). Epistasis mechanisms have been proposed to account for the majority of the heterotic effects in certain crosses of wild tomato *Lycopersicon hirsutum* through the use of near-isogenic lines (Monforte and Tanksley 2000) and in rice through chromosome segment substitution lines (Yu et al. 2005).

Role of Epigenetics in Heterosis

While genetically disparate individuals may induce heterosis through genetic mechanisms, it is highly likely that increased genetic differences are also associated with increased epigenetic differences (Schmitz et al. 2013). In this context, epigenetic effects can be defined as heritable changes in gene activity that are not related to underlying changes in DNA sequences (McKeown and Spillane 2014). To determine if parental epigenetic effects alone can cause heterosis in plants, it is necessary to construct a population of inbred lines that vary only for a segregating epigenetic marker. Such epigenetic recombinant inbred lines (epiRILs) in *A. thaliana* have been created (Johannes et al. 2009; Reinders et al. 2009). These can be crossed to the relevant wild-type (WT) accession to separate parental genetic effects from parental epigenetic effects on heterosis. Recently, epiRIL *X* WT crosses have demonstrated that maternal transmission of a hypomethylated genome is sufficient to trigger rosette size heterosis (Dapp et al. 2015). Attempts to create epiRILs in agronomically important crops like maize (Li et al. 2014) and rice (*Oryza sativa*) (Hu et al. 2014; Yamauchi et al. 2014) have been met with obstacles such as strong lethality following alterations to the DNA methylation network.

The genome dosage component of polyploid heterosis

Increasing the number of chromosome sets in the nucleus, including whether a genome is inherited maternally or paternally, has provided some novel insights into the mechanisms of heterosis and further tested the strict definitions of dominance, over-dominance, and epistasis. For example, there is the phenomenon known as ‘progressive heterosis’. It has been

demonstrated that increasing the heterozygosity through a double cross (beyond a single cross) substantially increases the heterosis effect in tetraploids while having a minimal effect on diploids. In stylised form, with hypothetical genotypes A, B, C, and D, this can be summarised as follows:

Cross	2x	Heterosis	4x	Heterosis
Single	AA *X* BB = AB	Yes	AAAA *X* BBBB = AABB	Yes
	CC *X* DD = CD	Yes	CCCC *X* DDDD = CCDD	Yes
Double	AB *X* CD = A/B/C/D	No	AABB *X* CCDD = A/B/C/D	Yes

Progressive heterosis effects for yield were observed in autotetraploids of potato (Mok and Peloquin 1975), alfalfa (Groose et al. 1989; Bingham et al. 1994), and maize (Sockness and Dudley 1989) as well as limited evidence for the same phenomenon in rice (Tu et al. 2007; Wu et al. 2013). The progressive heterosis effects in potato have been attributed to epistasis mechanisms due to greater allelic diversity in tetraploids than in diploids (Jansky 2009), and this gives further credence to the notion that the simple dominance theory is inadequate to explain observed heterosis effects. The barrier to widespread adoption of double cross tetraploid hybrids is species-specific. While potato and alfalfa are commercially grown at the tetraploid level (Table 1), maize is commercially grown at the diploid level. Tetraploid maize has lower fertility compared with diploids (Birchler 2014) and they frequently exhibit 'double reduction' in meiosis, where sister chromatids are not correctly separated during the first meiotic division and the resulting gamete contains both sister chromatids (Bingham et al. 1968). Thus, the progeny from a double cross tetraploid maize likely possess different genotypes.

Beyond enhancing the heterosis effect by doubling the ploidy level, both genome dosage and parent-of-origin effects on heterosis have been observed at the triploid level. In *A. thaliana*, for example, the separation of genome dosage *versus* genetic hybridity effects on heterosis have been determined. F1 triploid hybrids with two paternal genomes exhibit seed size and leaf area heterosis greater than that observed in F1 diploid hybrids. However, if an F1 triploid hybrid inherits two maternal genomes, the heterosis effect is absent or indeed reversed ('negative heterosis') (Miller et al. 2012; Fort et al. 2016). Thus, genome dosage effects on heterosis in *A. thaliana* triploids are parent-of-origin specific. In maize, paternal genome dosage effects on F1 triploid heterosis have been shown to be genotype-dependent (Yao et al.

2013). In this instance the differential heterosis observed between equivalent F1 triploid hybrids reveals, once again, the limitations of the dominance theory of heterosis.

Conclusions and Objectives

In this opening chapter I have outlined how polyploidy in plants can arise and summarised the consequences for plant gene expression, growth and reproduction. There is an abundance of literature to read, as it is now well-established (through detailed analysis of plant genomes, transcriptomes, and physiology) that WGD events are central actors in the story of plant evolution. Furthermore, I have summarised what contribution polyploidy can make to crop improvement. However, the published literature regarding the genome dosage component of heterosis is scattered and many knowledge gaps remain (see next paragraph). I agree with the 2014 review paper that asks those interested in crop improvement, particularly through the application of heterosis, not to ignore polyploids (Washburn and Birchler 2014). The novel heterotic phenotypes observed above the diploid level will contribute to solving the heterotic puzzle, and ultimately the noble goal of increasing the amount of food that can be produced from arable land while using resources more efficiently.

The knowledge gaps I have addressed in my research are as follows:

In chapter two I investigate if paternal genome dosage, genetic hybridity, or multi-locus nuclear heterozygosity is the main contributor to polyploid heterosis in the root crop sugar beet. I believe the novel findings of this research can inform sugar beet breeding programmes.

In chapter three I investigate if the phenomenon known as ‘progressive heterosis’ is present in the model organism *A. thaliana*. I believe the novel findings of this research can help answer fundamental questions of what is driving the heterosis effect.

In chapter four I investigate uniparental genome elimination in *A. thaliana*. I believe the novel findings of this research have potential to explain why breeding for haploid plants via centromere-mediated chromosome elimination may succeed or fail.

In chapter five I investigate if salt stress tolerance across several *A. thaliana* accessions is influenced by genome dosage or parent-of-origin effects. I believe the novel findings of this research can stand out among the published works searching for correlations between abiotic stress tolerance and genome dosage.

In chapter six I investigate if the ploidy level of wild and domesticated grass species is a reliable indicator of tolerance to temperature and precipitation fluctuation from climate change in the

year 2050. I believe the novel findings of this research raise important questions regarding polyploid establishment and agriculture's ability to meet future demands for cereal products.

To resolve these knowledge gaps, I undertook a range of research approaches including crop modelling, classical botany, crop field trials, and molecular genetics. I intentionally pursued both fundamental research and practical research for two reasons. One, the research tools and collaborators at my disposal in the Spillane Lab of the PABC, Ryan Institute, represented an excellent opportunity for me to "up-skill" and generate quality data from multiple, complementary sources. Two, in my short research career to date I have been inspired by Norman Borlaug's challenge of how research should move beyond a single scientific discipline if it is to have a positive impact on crop production (Borlaug 1986, 1996). The emergence of new technologies in plant science, and new disciplines and sub-disciplines dedicated to these technologies, are a boon to plant breeders (Morrell et al. 2012), yet how will these advancements fare as they leave the laboratory? For example, translating the interdisciplinary knowledge from a model species (e.g. *Arabidopsis thaliana*) to an agronomic crop, and then trying to make new crop varieties or new farming practices relevant to real-world scenarios, is a challenge (Wollenweber et al. 2005; Leonelli 2013). I believe there is room for plant scientists who can think broadly across disciplines and beyond my PhD I hope to be part of this discussion.

References

Acquaah, G., 2009 *Principles of plant genetics and breeding*: John Wiley & Sons.

Adams, K.L., R. Cronn, R. Percifield, and J.F. Wendel, 2003 Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. *Proceedings of the National Academy of Sciences* 100 (8):4649-4654.

Albertin, W., P. Brabant, O. Catrice, F. Eber, E. Jenczewski *et al.*, 2005 Autopolyploidy in cabbage (Brassica oleracea L.) does not alter significantly the proteomes of green tissues. *Proteomics* 5 (8):2131-2139.

Arrigo, N., and M.S. Barker, 2012 Rarely successful polyploids and their legacy in plant genomes. *Current Opinion in Plant Biology* 15 (2):140-146.

Bailey, J.P., K. Bímová, and B. Mandák, 2009 Asexual spread versus sexual reproduction and evolution in Japanese Knotweed s.l. sets the stage for the "Battle of the Clones". *Biological Invasions* 11 (5):1189-1203.

Baumel, A., M.L. Ainouche, and J.E. Levasseur, 2001 Molecular investigations in populations of Spartina anglica C.E. Hubbard (Poaceae) invading coastal Brittany (France). *Molecular Ecology* 10 (7):1689-1701.

Belling, J., 1925 The origin of chromosomal mutations inUvularia. *Journal of Genetics* 15 (3):245-266.

Bennetzen, J.L., 2005 Transposable elements, gene creation and genome rearrangement in flowering plants. *Curr Opin Genet Dev* 15 (6):621-627.

Bicknell, R.A., and A.M. Koltunow, 2004 Understanding Apomixis: Recent Advances and Remaining Conundrums. *The Plant Cell* 16 (suppl 1):S228.

Bingham, E.T., C.R. Burnham, and C.E. Gates, 1968 Double and single backcross linkage estimates in autotetraploid maize. *Genetics* 59 (3):399.

Bingham, E.T., R.W. Groose, D.R. Woodfield, and K.K. Kidwell, 1994 Complementary Gene Interactions in Alfalfa are Greater in Autotetraploids than Diploids. *Crop science* 34 (4):823-829.

Birchler, J.A., 2014 Interploidy hybridization barrier of endosperm as a dosage interaction. *Frontiers in Plant Science* 5 (281).

Blanc, G., and K.H. Wolfe, 2004 Functional Divergence of Duplicated Genes Formed by Polyploidy during Arabidopsis Evolution. *The Plant Cell* 16 (7):1679-1691.

Borlaug, N.E., 1986 Accelerating agricultural research and production in the third world: A scientist's viewpoint. *Agriculture and Human Values* 3 (3):5-14.

Borlaug, N.E., 1996 Biotechnology: Scientific panacea or research bandwagon.

Brouwer, D.J., and T.C. Osborn, 1999 A molecular marker linkage map of tetraploid alfalfa (Medicago sativa L.). *Theoretical and Applied Genetics* 99 (7):1194-1200.

Burow, M.D., C.E. Simpson, J.L. Starr, and A.H. Paterson, 2001 Transmission Genetics of Chromatin From a Synthetic Amphidiploid to Cultivated Peanut (Arachis hypogaea L.): Broadening the Gene Pool of a Monophyletic Polyploid Species. *Genetics* 159 (2):823-837.

Chagué, V., J. Just, I. Mestiri, S. Balzergue, A.M. Tanguy *et al.*, 2010 Genome-wide gene expression changes in genetically stable synthetic and natural wheat allohexaploids. *New Phytologist* 187 (4):1181-1194.

Chapman, H.M., D. Parh, and N. Oraguzie, 2000 Genetic structure and colonizing success of a clonal, weedy species, Pilosella officinarum (Asteraceae). *Heredity* 84:401.

Chen, C., L. Qi, S. Zhang, S. Han, X. Li *et al.*, 2004 The karyptype analysis of triploid poplar. *Wuhan botanical research* 22 (6):565-567.

Cheng, F., J. Wu, X. Cai, J. Liang, M. Freeling *et al.*, 2018 Gene retention, fractionation and subgenome differences in polyploid plants. *Nature plants* 4 (5):258-268.

Christelová, P., E. De Langhe, E. Hřibová, J. Čížková, J. Sardos *et al.*, 2017 Molecular and cytological characterization of the global Musa germplasm collection provides insights into the treasure of banana diversity. *Biodiversity and Conservation* 26 (4):801-824.

Christian, L., W.M. E., R. Olivier, and R.L. H., 2003 THE ORIGIN OF ECOLOGICAL DIVERGENCE IN HELIANTHUS PARADOXUS (ASTERACEAE): SELECTION ON TRANSGRESSIVE CHARACTERS IN A NOVEL HYBRID HABITAT. *Evolution* 57 (9):1989-2000.

Christian, L., L. Zhao, and R.L. H., 2004 Candidate gene polymorphisms associated with salt tolerance in wild sunflower hybrids: implications for the origin of Helianthus paradoxus, a diploid hybrid species. *New Phytologist* 161 (1):225-233.

Christian, P., S. Armel, Z. Tatiana, T. Maud, G. Marie-Angèle *et al.*, 2009 Rapid structural and epigenetic reorganization near transposable elements in hybrid and allopolyploid genomes in Spartina. *New Phytologist* 184 (4):1003-1015.

Coleman, L.C., 1950 Nuclear conditions in normal stem tissue of Vicia faba. *Canadian Journal of Research* 28 (3):382-391.

Comai, L., 2005 The advantages and disadvantages of being polyploid. *Nature reviews. Genetics* 6 (11):836.

Crosby, K., T.O. Stokes, and R.G. Latta, 2014 Evolving California genotypes of Avena barbata are derived from multiple introductions but still maintain substantial population structure. *PeerJ* 2:e633.

d'Erfurth, I., S. Jolivet, N. Froger, O. Catrice, M. Novatchkova *et al.*, 2008 Mutations in AtPS1 (Arabidopsis thaliana Parallel Spindle 1) Lead to the Production of Diploid Pollen Grains. *PLOS Genetics* 4 (11):e1000274.

Dapp, M., J. Reinders, A. Bediee, C. Balsera, E. Bucher *et al.*, 2015 Heterosis and inbreeding depression of epigenetic Arabidopsis hybrids. *Nature plants* 1 (7):15092.

Davoli, T., and T.d. Lange, 2011 The Causes and Consequences of Polyploidy in Normal Development and Cancer. *Annual Review of Cell and Developmental Biology* 27 (1):585-610.

Decroocq, V., L.S. Hagen, M.G. Favé, J.P. Eyquard, and A. Pierronnet, 2004 Microsatellite markers in the hexaploid Prunus domestica species and parentage lineage of three European plum cultivars using nuclear and chloroplast simple-sequence repeats. *Molecular breeding* 13 (2):135-142.

Deng, W., S.W. Tsao, J.N. Lucas, C.S. Leung, and A.L. Cheung, 2003 A new method for improving metaphase chromosome spreading. *Cytometry A* 51 (1):46-51.

Doležel, J., J. Greilhuber, and J. Suda, 2007 Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols* 2:2233.

Dorsey, E., 1936 INDUCED POLYPLOIDY IN WHEAT AND RYE: Chromosome Doubling in Triticum, Secale and Triticum-Secale Hybrids Produced by Temperature Changes. *Journal of Heredity* 27 (4):155-160.

Duszynska, D., P.C. McKeown, T.E. Juenger, A. Pietraszewska-Bogiel, D. Geelen *et al.*, 2013 Gamete fertility and ovule number variation in selfed reciprocal F1 hybrid triploid plants are heritable and display epigenetic parent-of-origin effects. *New Phytologist* 198 (1):71-81.

East, E.M., 1936 Heterosis. *Genetics* 21 (4):375.

Elvira Hörandl, 2008 Evolutionary Implications of Self-Compatibility and Reproductive Fitness in the Apomictic Ranunculus auricomus Polyploid Complex (Ranunculaceae). *International Journal of Plant Sciences* 169 (9):1219-1228.

Feldman, M., and M.E. Kislev, 2007 Domestication of emmer wheat and evolution of free-threshing tetraploid wheat. *Israel Journal of Plant Sciences* 55 (3-4):207-221.

Ferguson, D., and T. Sang, 2001 Speciation through homoploid hybridization between allotetraploids in peonies (Paeonia). *Proceedings of the National Academy of Sciences* 98 (7):3915-3919.

Fort, A., P. Ryder, P.C. McKeown, C. Wijnen, M.G. Aarts *et al.*, 2016 Disaggregating polyploidy, parental genome dosage and hybridity contributions to heterosis in Arabidopsis thaliana. *New Phytologist* 209 (2):590-599.

Frajman, B., F. Eggens, and B. Oxelman, 2009 Hybrid Origins and Homoploid Reticulate Evolution within Heliosperma (Sileneae, Caryophyllaceae)—A Multigene Phylogenetic Approach with Relative Dating. *Systematic Biology* 58 (3):328-345.

Fredotović, Ž., I. Šamanić, H. Weiss-Schneeweiss, J. Kamenjarin, T.-S. Jang *et al.*, 2014 Triparental origin of triploid onion, Allium × cornutum (Clementi ex Visiani, 1842), as evidenced by molecular, phylogenetic and cytogenetic analyses. *BMC plant biology* 14 (1):24.

Fukui, K., S. Nakayama, N. Ohmido, H. Yoshiaki, and M. Yamabe, 1998 Quantitative karyotyping of three diploid Brassica species by imaging methods and localization of 45s rDNA loci on the identified chromosomes. *Theoretical and Applied Genetics* 96 (3):325-330.

Glover, N.M., H. Redestig, and C. Dessimoz, 2016 Homoeologs: What Are They and How Do We Infer Them? *Trends in Plant Science* 21 (7):609-621.

Graham, G.I., D.W. Wolff, and C.W. Stuber, 1997 Characterization of a Yield Quantitative Trait Locus on Chromosome Five of Maize by Fine Mapping. *Crop Science* 37 (5):1601-1610.

Grant, V., 1981 Plant speciation. Columbia University Press.

Greer, E., A.C. Martín, A. Pendle, I. Colas, A.M.E. Jones *et al.*, 2012 The Ph1 Locus Suppresses Cdk2-Type Activity during Premeiosis and Meiosis in Wheat. *The Plant Cell* 24 (1):152-162.

Gregory, T.R., 2001 Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biological Reviews* 76 (1):65-101.

Grivet, L., and P. Arruda, 2002 Sugarcane genomics: depicting the complex genome of an important tropical crop. *Current Opinion in Plant Biology* 5 (2):122-127.

Groose, R.W., L.E. Talbert, W.P. Kojis, and E.T. Bingham, 1989 Progressive heterosis in autotetraploid alfalfa: studies using two types of inbreds. *Crop science* 29 (5):1173-1177.

Grossniklaus, U., 2017 Polyspermy produces tri-parental seeds in maize. *Current Biology* 27 (24):R1300-R1302.

Guo, M., D. Davis, and J.A. Birchler, 1996 Dosage Effects on Gene Expression in a Maize Ploidy Series. *Genetics* 142 (4):1349-1355.

Gupta, M., S. Gupta, H. Kumar, N. Kumar, and S.S. Banga, 2015 Population structure and breeding value of a new type of Brassica juncea created by combining A and B genomes from related allotetraploids. *Theoretical and Applied Genetics* 128 (2):221-234.

Haldane, J.B.S., 1933 The part played by recurrent mutation in evolution. *The American Naturalist* 67 (708):5-19.

Hanna, W.W., and W.F. Anderson, 2008 Development and impact of vegetative propagation in forage and turf bermudagrasses. *Agronomy journal* 100 (Supplement_3):S-103.

Hauck, N.R., H. Yamane, R. Tao, and A.F. Iezzoni, 2002 Self-compatibility and incompatibility in tetraploid sour cherry (Prunus cerasus L.). *Sexual Plant Reproduction* 15 (1):39-46.

Hedtke, S.M., and D.M. Hillis, 2011 The Potential Role of Androgenesis in Cytoplasmic–Nuclear Phylogenetic Discordance. *Systematic Biology* 60 (1):87-96.

Henry, I.M., B.P. Dilkes, K. Young, B. Watson, H. Wu *et al.*, 2005 Aneuploidy and genetic variation in the Arabidopsis thaliana triploid response. *Genetics* 170 (4):1979-1988.

Hu, L., N. Li, C. Xu, S. Zhong, X. Lin *et al.*, 2014 Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. *Proceedings of the National Academy of Sciences*.

Husband, B.C., B. Ozimec, S.L. Martin, and L. Pollock, 2008 Mating consequences of polyploid evolution in flowering plants: current trends and insights from synthetic polyploids. *International Journal of Plant Sciences* 169 (1):195-206.

Jackson, R.C., 1976 Evolution and Systematic Significance of Polyploidy. *Annual Review of Ecology and Systematics* 7:209-234.

Jansky, S., 2009 Breeding, genetics, and cultivar development, pp. 27-62 in *Advances in potato chemistry and technology*. Elsevier.

Jiao, Y., N.J. Wickett, S. Ayyampalayam, A.S. Chanderbali, L. Landherr *et al.*, 2011 Ancestral polyploidy in seed plants and angiosperms. *Nature* 473:97.

Johannes, F., E. Porcher, F.K. Teixeira, V. Saliba-Colombani, M. Simon *et al.*, 2009 Assessing the impact of transgenerational epigenetic variation on complex traits. *PLOS Genetics* 5 (6):e1000530.

Kashkush, K., M. Feldman, and A.A. Levy, 2002 Gene Loss, Silencing and Activation in a Newly Synthesized Wheat Allotetraploid. *Genetics* 160 (4):1651-1659.

Köhler, C., O. Mittelsten Scheid, and A. Erilova, 2010 The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics* 26 (3):142-148.

Kondorosi, E., F. Roudier, and E. Gendreau, 2000 Plant cell-size control: growing by ploidy? *Current Opinion in Plant Biology* 3 (6):488-492.

Kostoff, D., 1933 A contribution to the sterility and irregularities in the meiotic processes caused by virus diseases. *Genetica* 15 (1-2):103-114.

Kreiner, J.M., P. Kron, and B.C. Husband, 2017 Frequency and maintenance of unreduced gametes in natural plant populations: associations with reproductive mode, life history and genome size. *New Phytologist* 214 (2):879-889.

Krieger, U., Z.B. Lippman, and D. Zamir, 2010 The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nature Genetics* 42:459.

Lacape, J.M., T.B. Nguyen, S. Thibivilliers, B. Bojinov, B. Courtois *et al.*, 2003 A combined RFLP SSR AFLP map of tetraploid cotton based on a Gossypium hirsutum× Gossypium barbadense backcross population. *Genome* 46 (4):612-626.

Leonelli, S., 2013 Integrating data to acquire new knowledge: Three modes of integration in plant science. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 44 (4, Part A):503-514.

Li, P., S. Zhang, F. Li, S. Zhang, H. Zhang *et al.*, 2017 A Phylogenetic Analysis of Chloroplast Genomes Elucidates the Relationships of the Six Economically Important Brassica Species Comprising the Triangle of U. *Frontiers in Plant Science* 8 (111).

Li, Q., S.R. Eichten, P.J. Hermanson, V.M. Zaunbrecher, J. Song *et al.*, 2014 Genetic Perturbation of the Maize Methylome. *The Plant Cell Online*.

Li, Z., A.E. Baniaga, E.B. Sessa, M. Scascitelli, S.W. Graham *et al.*, 2015 Early genome duplications in conifers and other seed plants. *Science Advances* 1 (10).

Liston, A., R. Cronn, and T.L. Ashman, 2014 Fragaria: A genus with deep historical roots and ripe for evolutionary and ecological insights. *American Journal of Botany* 101 (10):1686-1699.

Lloyd, A., and K. Bomblies, 2016 Meiosis in autopolyploid and allopolyploid Arabidopsis. *Current Opinion in Plant Biology* 30:116-122.

Loskutov, I.G., and H.W. Rines, 2011 Avena, pp. 109-183 in *Wild Crop Relatives: Genomic and Breeding Resources*. Springer.

M., A.A., B.G. L.A., B.S. T., C.J. A., G. Rhian *et al.*, 2011 Transcript-specific, single-nucleotide polymorphism discovery and linkage analysis in hexaploid bread wheat (Triticum aestivum L.). *Plant Biotechnology Journal* 9 (9):1086-1099.

Ma, X.-F., P. Fang, and J.P. Gustafson, 2004 Polyploidization-induced genome variation in triticale. *Genome* 47 (5):839-848.

Madlung, A., A.P. Tyagi, B. Watson, H. Jiang, T. Kagochi *et al.*, 2005 Genomic changes in synthetic Arabidopsis polyploids. *The Plant Journal* 41 (2):221-230.

Maere, S., S. De Bodt, J. Raes, T. Casneuf, M. Van Montagu *et al.*, 2005 Modeling gene and genome duplications in eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America* 102 (15):5454-5459.

Maki, M., and J. Murata, 2001 Allozyme analysis of the hybrid origin of Arisaema ehimense (Araceae). *Heredity* 86:87.

Mallet, J., 2007 Hybrid speciation. *Nature* 446 (7133):279.

McCarthy, E.W., M.W. Chase, S. Knapp, A. Litt, A.R. Leitch *et al.*, 2016 Transgressive phenotypes and generalist pollination in the floral evolution of Nicotiana polyploids. *Nat Plants* 2:16119.

McKeown, P.C., and C. Spillane, 2014 Landscaping Plant Epigenetics, pp. 1-24 in *Plant Epigenetics and Epigenomics: Methods and Protocols*, edited by C. Spillane and P.C. McKeown. Humana Press, Totowa, NJ.

Menzel, M.Y., and F.D. Wilson, 1966 Hybrids and genome relations of Hibiscus sabdariffa, H. meeusei, H. radiatus and H. acetosella. *American Journal of Botany* 53 (3):270-275.

Mercier, R., D. Vezon, E. Bullier, J.C. Motamayor, A. Sellier *et al.*, 2001 SWITCH1 (SWI1): a novel protein required for the establishment of sister chromatid cohesion and for bivalent formation at meiosis. *Genes & Development* 15 (14):1859-1871.

Michod, R.E., 2000 *Darwinian dynamics: evolutionary transitions in fitness and individuality*: Princeton University Press.

Milbocker, D., and K. Sink, 1969 Embryology of diploid× diploid and diploid× tetraploid crosses in poinsettia. *Canadian Journal of Genetics and Cytology* 11 (3):598-601.

Miller, J.S., and D.L. Venable, 2000 Polyploidy and the Evolution of Gender Dimorphism in Plants. *Science* 289 (5488):2335-2338.

Miller, M., C. Zhang, and Z.J. Chen, 2012 Ploidy and hybridity effects on growth vigor and gene expression in Arabidopsis thaliana hybrids and their parents. *G3: Genes, Genomes, Genetics* 2 (4):505-513.

Mok, D.W., and S.J. Peloquin, 1975 Breeding value of 2n pollen (diplandroids) in tetraploid x diploid crosses in potatoes. *Theor Appl Genet* 46 (6):307-314.

Monforte, A.J., and S.D. Tanksley, 2000 Fine mapping of a quantitative trait locus (QTL) from Lycopersicon hirsutum chromosome 1 affecting fruit characteristics and agronomic traits: breaking linkage among QTLs affecting different traits and dissection of heterosis for yield. *Theoretical and Applied Genetics* 100 (3):471-479.

Moody, M.E., L.D. Mueller, and D.E. Soltis, 1993 Genetic variation and random drift in autotetraploid populations. *Genetics* 134 (2):649-657.

Morrell, P.L., E.S. Buckler, and J. Ross-Ibarra, 2012 Crop genomics: advances and applications. *Nature Reviews Genetics* 13 (2):85-96.

Nagaharu, U., 1935 Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. *Jpn J Bot* 7 (7):389-452.

Nakel, T., D.G. Tekleyohans, Y. Mao, G. Fuchert, D. Vo *et al.*, 2017 Triparental plants provide direct evidence for polyspermy induced polyploidy. *Nature Communications* 8 (1):1033.

Oberlander, K.C., L.L. Dreyer, P. Goldblatt, J. Suda, and H.P. Linder, 2016 Species-rich and polyploid-poor: Insights into the evolutionary role of whole-genome duplication from the Cape flora biodiversity hotspot. *American Journal of Botany* 103 (7):1336-1347.

Otto, S.P., and J. Whitton, 2000 Polyploid Incidence and Evolution. *Annual Review of Genetics* 34 (1):401-437.

Pandit, M.K., M.J.O. Pocock, and W.E. Kunin, 2011 Ploidy influences rarity and invasiveness in plants. *Journal of Ecology* 99 (5):1108-1115.

Pearl, H.M., C. Nagai, P.H. Moore, D.L. Steiger, R.V. Osgood *et al.*, 2004 Construction of a genetic map for arabica coffee. *Theoretical and Applied Genetics* 108 (5):829-835.

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

Ramsey, J., and D.W. Schemske, 2002 Neopolyploidy in Flowering Plants. *Annual Review of Ecology and Systematics* 33 (1):589-639.

Randolph, L.F., 1932 Some effects of high temperature on polyploidy and other variations in maize. *Proceedings of the National Academy of Sciences* 18 (3):222-229.

Ravi, M., M.P.A. Marimuthu, and I. Siddiqi, 2008 Gamete formation without meiosis in Arabidopsis. *Nature* 451:1121.

Reinders, J., B.B.H. Wulff, M. Mirouze, A. Marí-Ordóñez, M. Dapp *et al.*, 2009 Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. *Genes & Development* 23 (8):939-950.

Richards, A.J., 2003 Apomixis in flowering plants: an overview. *Philosophical Transactions of the Royal Society B: Biological Sciences* 358 (1434):1085-1093.

Rieseberg, L.H., 1997 Hybrid origins of plant species. *Annual Review of Ecology and Systematics* 28 (1):359-389.

Rieseberg, L.H., O. Raymond, D.M. Rosenthal, Z. Lai, K. Livingstone *et al.*, 2003 Major Ecological Transitions in Wild Sunflowers Facilitated by Hybridization. *Science* 301 (5637):1211-1216.

Rivero-Guerra, A.O., 2008 Cytogenetics, geographical distribution, and pollen fertility of diploid and tetraploid cytotypes of Santolina pectinata Lag. (Asteraceae: Anthemideae). *Botanical Journal of the Linnean Society* 156 (4):657-667.

Rouiss, H., F. Bakry, Y. Froelicher, L. Navarro, P. Aleza *et al.*, 2018 Origin of C. latifolia and C. aurantiifolia triploid limes: the preferential disomic inheritance of doubled-diploid 'Mexican' lime is consistent with an interploid hybridization hypothesis. *Annals of Botany* 121 (3):571-585.

Santos, J.L., D. Alfaro, E. Sanchez-Moran, S.J. Armstrong, F.C.H. Franklin *et al.*, 2003 Partial Diploidization of Meiosis in Autotetraploid Arabidopsis thaliana. *Genetics* 165 (3):1533-1540.

Schmitz, R.J., M.D. Schultz, M.A. Urich, J.R. Nery, M. Pelizzola *et al.*, 2013 Patterns of population epigenomic diversity. *Nature* 495 (7440):193.

Schnable, P.S., and N.M. Springer, 2013 Progress toward understanding heterosis in crop plants. *Annual review of plant biology* 64:71-88.

Scott, R.J., S.J. Armstrong, J. Doughty, and M. Spielman, 2008 Double fertilization in Arabidopsis thaliana involves a polyspermy block on the egg but not the central cell. *Mol Plant* 1 (4):611-619.

Sears, E.R., 1976 Genetic Control of Chromosome Pairing in Wheat. *Annual Review of Genetics* 10 (1):31-51.

Seoighe, C., and C. Gehring, 2004 Genome duplication led to highly selective expansion of the Arabidopsis thaliana proteome. *Trends in Genetics* 20 (10):461-464.

Sockness, B.A., and J.W. Dudley, 1989 Performance of Single and Double Cross Autotetraploid Maize Hybrids with Different Levels of Inbreeding. *Crop science* 29 (4):875-879.

Soltis, D.E., P.S. Soltis, and J.A. Tate, 2004 Advances in the study of polyploidy since Plant speciation. *New Phytologist* 161 (1):173-191.

Soltis, P., G. Plunkett, S. Novak, and D. Soltis, 1995 GENETIC VARIATION IN TRAGOPOGON SPECIES: ADDITIONAL ORIGINS OF THE ALLOTETRAPLOIDS T. MIRUS AND T. MISCELLUS (COMPOSITAE). *American Journal of Botany* 82 (10):1329-1341.

Sprague, G.F., 1929 Hetero-fertilization in maize. *Science* 69 (1794):526-527.

Springer, N.M., and R.M. Stupar, 2007 Allelic variation and heterosis in maize: how do two halves make more than a whole? *Genome research* 17 (3):000-000.

Stebbins, G.L., 1971 Chromosomal evolution in higher plants. *Chromosomal evolution in higher plants.*

Steinitz-Sears, L.M., 1963 Chromosome studies in Arabidopsis thaliana. *Genetics* 48 (4):483.

Stift, M., C. Berenos, P. Kuperus, and P.H. van Tienderen, 2008 Segregation Models for Disomic, Tetrasomic and Intermediate Inheritance in Tetraploids: A General Procedure Applied to Rorippa (Yellow Cress) Microsatellite Data. *Genetics* 179 (4):2113-2123.

Storchova, Z., and D. Pellman, 2004 From polyploidy to aneuploidy, genome instability and cancer. *Nature Reviews Molecular Cell Biology* 5:45.

Stuber, C.W., S.E. Lincoln, D.W. Wolff, T. Helentjaris, and E.S. Lander, 1992 Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics* 132 (3):823-839.

Stupar, R.M., P.B. Bhaskar, B.S. Yandell, W.A. Rensink, A.L. Hart *et al.*, 2007 Phenotypic and Transcriptomic Changes Associated With Potato Autopolyploidization. *Genetics* 176 (4):2055-2067.

Sujatha, M., and A.J. Prabakaran, 2006 Ploidy manipulation and introgression of resistance to Alternaria helianthi from wild hexaploid Helianthus species to cultivated sunflower (H. annuus L.) aided by anther culture. *Euphytica* 152 (2):201-215.

Techio, V.H., L.C. Davide, and A.V. Pereira, 2006 Meiosis in elephant grass (Pennisetum purpureum), pearl millet (Pennisetum glaucum)(Poaceae, Poales) and their interspecific hybrids. *Genetics and Molecular Biology* 29 (2):353-362.

Therman, E., 1956 Dedifferentiation and Differentiation of Cells in Crown Gall of Vicia Faba. *Caryologia* 8 (3):325-348.

Trebbi, D., M. Maccaferri, P. de Heer, A. Sørensen, S. Giuliani *et al.*, 2011 High-throughput SNP discovery and genotyping in durum wheat (Triticum durum Desf.). *Theoretical and Applied Genetics* 123 (4):555-569.

Tu, S., L. Luan, Y. Liu, W. Long, F. Kong *et al.*, 2007 Production and Heterosis Analysis of Rice Autotetraploid Hybrids. *Crop science* 47 (6):2356-2363.

Van de Peer, Y., E. Mizrachi, and K. Marchal, 2017 The evolutionary significance of polyploidy. *Nature Reviews Genetics* 18 (7):411.

Vandel, A., 1938 Chromosome Number, Polyploidy, and Sex in the Animal Kingdom*. *Proceedings of the Zoological Society of London* A107 (4):519-541.

Wang, J., L. Tian, H.-S. Lee, N.E. Wei, H. Jiang *et al.*, 2006 Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. *Genetics* 172 (1):507-517.

Wang, J., L. Tian, A. Madlung, H.-S. Lee, M. Chen *et al.*, 2004 Stochastic and Epigenetic Changes of Gene Expression in Arabidopsis Polyploids. *Genetics* 167 (4):1961-1973.

Washburn, J.D., and J.A. Birchler, 2014 Polyploids as a "model system" for the study of heterosis. *Plant Reprod* 27 (1):1-5.

Wehner, T.C., 2008 Watermelon, pp. 381-418 in *Vegetables I*. Springer.

Wendel, J.F., and R.C. Cronn, 2003 Polyploidy and the evolutionary history of cotton. *Advances in agronomy* 78:139.

Wilson, F.D., and M.Y. Menzel, 1964 Kenaf (Hibiscus cannabinus), roselle (Hibiscus sabdariffa). *Economic Botany* 18 (1):80-91.

Wollenweber, B., J.R. Porter, and T. Lübberstedt, 2005 Need for multidisciplinary research towards a second green revolution. *Current Opinion in Plant Biology* 8 (3):337-341.

Wong, J.L., and G.M. Wessel, 2005 Defending the Zygote: Search for the Ancestral Animal Block to Polyspermy, pp. 1-151 in *Current Topics in Developmental Biology*. Academic Press.

Woodell, S.R.J., and D.H. Valentine, 1961 STUDIES IN BRITISH PRIMULAS: IX. SEED INCOMPATIBILITY IN DIPLOID-AUTOTETRAPLOID CROSSES. *New Phytologist* 60 (3):282-294.

Wu, F., N.T. Eannetta, Y. Xu, J. Plieske, M. Ganal *et al.*, 2010 COSII genetic maps of two diploid Nicotiana species provide a detailed picture of synteny with tomato and insights into chromosome evolution in tetraploid N. tabacum. *Theoretical and Applied Genetics* 120 (4):809-827.

Wu, J.-W., C.-Y. Hu, M.Q. Shahid, H.-B. Guo, Y.-X. Zeng *et al.*, 2013 Analysis on genetic diversification and heterosis in autotetraploid rice. *SpringerPlus* 2 (1):439.

Yamauchi, T., Y. Johzuka-Hisatomi, R. Terada, I. Nakamura, and S. Iida, 2014 The MET1b gene encoding a maintenance DNA methyltransferase is indispensable for normal development in rice. *Plant Molecular Biology* 85 (3):219-232.

Yao, H., A. Dogra Gray, D.L. Auger, and J.A. Birchler, 2013 Genomic dosage effects on heterosis in triploid maize. *Proceedings of the National Academy of Sciences* 110 (7):2665-2669.

Yu, C., J. Wan, H. Zhai, C. Wang, L. Jiang *et al.*, 2005 Study on heterosis of inter-subspecies between indica and japonica rice (Oryza sativa L.) using chromosome segment substitution lines. *Chinese Science Bulletin* 50 (2):131-136.

Yu, Z., G. Haberer, M. Matthes, T. Rattei, K.F.X. Mayer *et al.*, 2010 Impact of natural genetic variation on the transcriptome of autotetraploid Arabidopsis thaliana. *Proceedings of the National Academy of Sciences* 107 (41):17809-17814.

Zhang, S., L. Qi, C. Chen, X. Li, W. Song *et al.*, 2004 A report of triploid Populus of the section Aigeiros. *Silvae Genetica* 53 (1-6):69-75.

Chapter 2. Hybridity has a greater effect than paternal genome dosage on heterosis in sugar beet (*Beta vulgaris*)

Note: this chapter was published in *BMC Plant Biology* (https://doi.org/10.1186/s12870-018-1338-x) with the following author contributions – Charles Spillane conceived and designed the study with inputs from Brendan Hallahan, Edna Curley, Galina Brychkova and Britta Schulz. Britta Schulz provided the sugar beet material. Brendan Hallahan and Eva Fernandez managed the field trials with input from Antoine Fort, Peter Ryder, Gilles Dupouy, Marc Deletre and Edna Curley. Brendan Hallahan and Eva Fernandez analysed seed and fruit traits, field trial data, and conducted the statistical analysis. Charles Spillane and Brendan Hallahan wrote the manuscript. All authors read and approved the final manuscript.

Introduction

Heterosis can be described as an increase in size or other desirable characteristics (e.g. grain yield) in the F1 offspring beyond that observed in the parental lines (Schnable and Springer 2013). The application of heterosis for the improvement of specific traits in crops is highlighted by the widespread development of F1 hybrid varieties and their widespread adoption by farmers (Schnable and Springer 2013). Most crop plants are polyploids, either autopolyploids or allopolyploids (Comai 2005). Polyploidy (i.e. genome dosage changes) can also be harnessed for crop improvement, as some polyploids can display fitness advantages over progenitor or parental lines. During crop domestication, and subsequent artificial selection, polyploids may have been selected for due to desirable traits such as gigantism (Paterson and Wendel 2015). Genome dosage increases in newly formed polyploids can elicit novel phenotypes, while genetic redundancy within polyploid genomes can allow for neo- or subfunctionalization of gene functions (Roulin et al. 2013; Tan et al. 2016).

Heterosis effects can arise as a result of both gene and genome dosage effects (Chen 2010; Jiang et al. 2013; Ryder et al. 2014; Washburn and Birchler 2014), particularly in allopolyploid plants (which are also by definition hybrids) (Wang et al. 2006; Chen 2010). The disentanglement of parental genome dosage *versus* hybridity contributions to heterosis requires the use of genome dosage series lines (genotypes) which are ideally genetically identical F1 hybrids, yet contain additional chromosome sets from either one or the other parent (Duszynska et al. 2013a; Donoghue et al. 2014; Fort et al. 2016; Fort et al. 2017). Such experiments have been done in *Arabidopsis thaliana* (Miller et al. 2012; Donoghue et al. 2013; Duszynska et al. 2013b; Fort et al. 2016) and in maize (*Zea mays*) (Guo et al. 1996; Yao et al. 2013) revealing significant parental genome dosage effects on heterosis.

Parental lines which elicit significant heterosis effects are identified empirically through crossing experiments to identify parental germplasm pools that display good combining ability to generate heterosis effects in F1 hybrids (Schnable and Springer 2013). While it could be considered that the genetic distance between parental lines could be used as basis to select parental lines to generate superior heterozygous F1 hybrids, there is rather limited evidence to support this approach. Early investigations into parental genetic distance and F1 hybrid performance in maize found a general correlation between genetic distance and heterosis up to a certain threshold. By grouping parental germplasm based on regional adaptation, F1 heterosis increased with increased parental divergence within a range: parents from different parts of the USA when crossed together produced F1 hybrids with considerable

heterosis, but when parents from the USA were crossed with Mexican varieties, which represents a wider cross, there was less heterosis (Moll et al. 1965). More recent experiments utilizing molecular markers have found no correlation between parental genetic distance and heterosis in maize (Benchimol et al. 2000; Frisch et al. 2010; Reif et al. 2010). Other crops where this has been investigated (e.g., bread wheat, rice, pepper, oilseed rape) and in models such as *Arabidopsis thaliana*, have also shown limited or no evidence supporting this approach to selecting parental lines for triggering heterosis effects in F1 hybrids (Ali et al. 1995; Martin et al. 1995; Zhang et al. 1995; Diers et al. 1996; Riaz et al. 2001; Geleta et al. 2004; Meyer et al. 2004; Stokes et al. 2007).

Sugar beet is a crop which is amenable to heterosis comparisons between diploid and triploid genotypes, including whether heterosis effects can be augmented by changing paternal genome dosage or multi-locus heterozygosity in the F1 hybrid. Indeed, commercial sugar beet breeding first began exploiting genome dosage effects in the 1930s (Draycott 2008), where early triploid varieties displayed many favorable agronomic traits, including high yield (Peto and Boyes 1940). Following the advent of cytoplasmic male sterile lines and the introduction of the monogerm seed character (i.e. fruits which produce a single seed) (Savitsky 1950), more efficient and reliable F1 hybrid sugar beet production became possible. In North America, sugar beet breeding has largely focused on diploid F1 hybrid varieties whereas in Europe triploid F1 hybrids are historically more popular. However, in recent years European sugar beet breeding programs have increasingly moved toward diploid F1 hybrid breeding (Draycott 2008; Biancardi et al. 2010). Sugar beet provides a useful model for investigating the contributions of paternal genome dosage *versus* hybridity to heterosis in a commercial crop.

In this study, I consider three mechanisms that could potentially contribute to sugar beet heterosis, namely ploidy (genome dosage), hybridity and increasing multi-locus nuclear heterozygosity. While clearly interlinked, hybridity and heterozygosity are not synonymous. An F1 hybrid offspring plant is a genetic composite of the maternal cytoplasmic genome and the biparental nuclear genomes inherited from each parents (Hedtke and Hillis 2011). Heterozygosity is a measure of nuclear genetic diversity, which can be determined by the extent of multi-locus single nucleotide polymorphisms (SNPs) across the nuclear genome (Van Dyke 2008). Sugar beet lines of different ploidy levels were crossed to generate F1 hybrids allowing investigation of: (1) the effect of paternal genome dosage increase on F1 heterosis, and; (2) the effect of homozygous versus heterozygous tetraploid male parents on F1 hybrid triploid heterosis. My findings indicate that (1) agronomically important traits such as total yield, root yield and sugar yield are more influenced by hybridity than paternal genome dosage increases,

and (2) F1 triploid hybrids with greater levels of multi-locus nuclear heterozygosity do not display improvements for total yield, root yield, and sugar yield. I consider that my findings have relevance to the design of future hybrid breeding programs for sugar beet improvement.

Materials and Methods

Sugar Beet Lines and Crossing Design

Sugar beet lines/genotypes obtained from KWS SAAT consisted of: (a) isogenic and hybrid diploid parental lines; (b) isogenic and hybrid tetraploid parental lines; and (c) F1 progeny (diploid and triploid) offspring generated from the parental lines (**Table *2*** and **Table *3***). Double haploids were generated by ovule isolation from an F1 seed, followed by colchicine treatment. Homozygous tetraploids were generated when the double haploids underwent spontaneous doubling during colchicine treatment. The heterozygous tetraploids were selected by KWS SAAT for good performance over many years within the KWS SAAT sugar beet breeding program, where these tetraploids ($C^1C^2C^3C^4$ and $D^1D^2D^3D^4$) harbor high levels of heterozygosity at each locus across the nuclear genome, hereafter referred to as (CCCC) and (DDDD). All three diploid tester lines (Tester 1, 2, and 3) of sugar beet that were used as female parents were sourced from the monogerm seed parent pool, i.e. two monogerm lines were crossed and the F1 hybrid was back-crossed several times with a cytoplasmic male sterile (CMS) line. The F1 hybrids were created using a female tester and pollinator line in the same plot at KWS SAAT. The F1 hybrid seed was harvested from the male sterile female tester lines.

Sugar Beet Germination and Seed Viability Test

Sugar beet fruits can contain one seed (monogerm) or more than one seed (multigerm). Sugar beet fruits (containing botanical seed) were sown and seed germination was investigated in accordance with International Rules for Seed Testing (ISTA, 2010). Fifty randomly selected fruits of each line were placed is 50ml distilled water and covered with foil to block any light. These were left for 2 hours, rinsed, and placed in Grade 3236 pleated cellulose filtered paper (GE Healthcare, Fairfield, CT, USA) inside a 100mm x 100mm x 20mm petri dish (Sarstedt AG & Co, Nümbrecht, Germany). The paper was cut to size and 6.4ml of distilled water was added. The petri dishes were sealed with parafilm and placed in a growth chamber (Snijders Scientific, Tilburg, Netherlands) at the beginning of the dark cycle. The growth conditions were 8 hours day/16 hours night @30°C/20°C respectively, in accordance with ISTA guidelines. Germination was visually scored on day 4 and 14 post sowing date. A seed was categorized as ‘germinated’ when the radicle had emerged from the operculum of the fruit. Ungerminated seeds were examined under a Leica MZ microscope (Leica, Wetzlar, Germany) and were categorized as alive if they were plump, while seeds that were wrinkled and poorly formed

were classified as dead (Supplementary Results, Figure S1). Fruits with only dead seeds were excluded from final calculations. The seed germination count of each line was calculated as follows:

$$\frac{No.of\ fruits\ containing\ at\ least\ one\ germinated\ seed}{No.of\ fruits\ containing\ only\ ungerminated, healthy\ seeds}$$

The results are of the seed germination analysis are provided in Supplementary Results (Figures S2 and S3).

Sugar Beet Seed and Fruit Analysis

For analysis of the seed and empty fruits of sugar beet, twenty-five randomly selected fruits of each line were placed in 50ml distilled water for 48 hours and covered with foil to block any light. The operculum and seed were removed from the fruit and weighed individually on a weighing scale (Mettler Toledo, Switzerland). Seed size data generated is for alive seeds. The entire experiment was replicated 4 times giving a total of 100 seed measurements for each line. For the analysis of the seed cross-sections, thirty randomly selected fruits of each line were placed in 50ml distilled water and covered with foil. After 36 hours, the operculum and the seed testa (seed coat) were removed with a razor blade to make the embryo and perisperm visible. The cross-section was imaged under a Leica MZ dissecting microscope (Leica, Wetzlar, Germany), and the embryo and perisperm size (i.e. area) were determined using IMAGEJ (US ImageJ, Bethesda, MD, USA) software. See **Figure *1*** for a representative example of a cross section of a sugar beet fruit containing a seed.

Ploidy Analysis of Sugar Beet Seedlings

Seedlings of each sugar beet line were grown in individual pots of soil (5:1:1 mixture of peat soil:perlite:vermiculite) and placed in a growth chamber (16 hours day/8 hours night @21°C/18°C). A destructive harvest of first true leaves was performed, where 400µl of nuclei extraction buffer (Sysmex, Kobe, Japan) was added to the leaf material which was chopped with a razor blade. The chopped mixture was left for 5 minutes. The mixture was strained through a 30µm CellTrics® filter (Sysmex, Kobe, Japan) into a 3.5ml Röhren tube. 40µl of 1% v/v polyvinylpyrrolidone (PVP) was added and left for 5 minutes. Finally, 1ml of UV-stain was added to the tube. The sample was analyzed on a Partec Ploidy Analyzer (Sysmex Kobe, Japan). The ploidy level of each line was confirmed (Supplementary Results).

DNA Extraction

Seedlings of each sugar beet line were grown in individual pots as described previously. A destructive harvest of first true leaves was performed. Genomic DNA (gDNA) was extracted from three biological replicates of each line with a NucleoSpin® Plant II kit (Düren, Germany), according to the manufacturer's instructions with one revision: incubation time with RNase A was increased to 30 minutes. The quality and quantity of gDNA was evaluated via agarose gel electrophoresis, NanoDrop® 2000 spectrophotometer (Thermo Fisher Scientific, MA, USA), and Qubit® 2.0 fluorometer (Thermo Fisher Scientific, MA, USA). Sample DNA, with OD_{260}/OD_{280} ratio of 1.8 to 2.0 and total amount of more than 1.5µg, was used for library construction.

Library Preparation, Illumina Sequencing and SNP Detection for measuring nuclear multi-locus heterozygosity

Circa 0.3~0.6µg of gDNA was digested with restriction enzymes and the resulting digested fragments were ligated to two barcoded adapters, the universal adapter (5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGA TCT 3') and the indexed adapter (5' GATCGGAAGAGCACACGTCTGAACTCCAGTCAC-NNNNNN-ATCTCGTATGCCGTCTTCTGCTTG 3') which when annealed generated compatible sticky ends corresponding to the restriction digestion enzyme. Following PCR amplification, all samples were pooled and size-selected for the required fragments to complete the library construction. Qubit® 2.0 fluorometer was used to determine the DNA concentration of the prepared libraries. After dilution to 1 ng DNA/µl, an Agilent® 2100 bioanalyzer (Agilent Technologies, CA, USA) was used to determine the insert size. Finally, quantitative real-time PCR (qPCR) was performed to determine the effective concentration of each library. If the library with appropriate insert size had an effective concentration of > 2nM, the constructed libraries were deemed sufficient quality and used for Illumina® high-throughput sequencing.

The DNA libraries were pooled according to their effective concentration as well as the expected data production. Paired-end sequencing was performed on an Illumina®HiSeq platform (Illumina, CA, USA) at Novogene Technology Co., Ltd. (Beijing, China).

Raw sequencing reads were processed with CASAVA software (version 1.8) and sequencing data was assessed for quality distribution, sequencing errors and adapter contamination. In total 29.068Gbit of raw data were sequenced, with 29.065Gbit clean data generated after quality control. The clean sequencing data was aligned with the reference

sequence using Burrows-Wheeler Aligner (BWA) software (Li and Durbin 2009) and the mapping rate and coverage was assessed. Sorting of the BAM files was performed with SAMtools software (Li et al. 2009). For the detection and filtration of SNPs and InDels SAMtools software was used, also (Li et al. 2009). Annotation of detected SNPs was performed with ANNOVAR software (Wang et al. 2010). The genome-wide multi-locus heterozygosity rate was calculated by the ratio of heterozygous SNPs to the total number of genome bases. A detailed description of the genotyping-by-sequencing methods are provided in Supplementary Materials and Methods.

Sugar Beet Field Trial

A field trial of the sugar beet lines was conducted in Cobh, Cork, Ireland (51.851°N 8.2967°W) as a complete randomized block design with 4 replicates. Fruits (seeds) were sown in the first week of April 2015 and harvested in the last week of October 2015. The mean air temperature over the season was 13°C and mean precipitation was 68.4mm (Cobh Weather station 51°51'.18N 008°17'.40W). Trial plots consisted of 20 rows, each 9 meters in length with an inter-row spacing of 61 cm. Each sugar beet line (i.e. genotype) was sowed once per plot. The position of each line was randomized in each plot with a commercial sugar beet variety (cv. Rosalinda from KWS SAAT) occupying the five remaining rows. 300 fruits per row were sown manually and thinned out (after homogenous emergence four weeks post sowing date) to a population density of 50 plants per row, thus leaving approximately 18 cm of space between plants.

Sugar Beet Harvest Procedure

The harvesting of the sugar beet was performed manually. All weight measurements were recorded with a ‘Defender 3000’ weighing scale (Ohaus Corporation, NJ, USA). The total yield (kg/ha) of each line was calculated, plants were left to dry overnight in the field, and thereafter 25 plants from each line were randomly selected and weighed individually. The crown and leaves were both removed by cutting below the lowest leaf scar, and both above- and below-ground parts were weighed separately. Root length was recorded from the top of the root to the root tip and circumference was recorded at the widest part of the root. Twenty-five roots from each line were stored in polypropylene bags and transported under controlled conditions @ 2-7°C to KWS SAAT SE, Germany for compositional analysis.

Sugar Beet Chemical Compositional Analysis

Raw sugar beet roots were stored according to best practice (Kenter et al. 2006) and analyzed following internationally agreed protocols (De Whalley 2013). Chemical composition was determined with an automatic beet laboratory system (Venema, Groningen, Netherlands). Briefly, roots were washed, weighed and processed into brei. Subsequently, brei was stored at -26°C until analysis. The brei sample was clarified with 0.3% w/v aluminium sulfate solution. Sucrose was measured polarimetrically and Sodium and Potassium by flame photometry. α-amino Nitrogen was analyzed by the fluorometric OPA-method.

Formulae and Statistical Analysis

Following chemical analysis, the "new Braunschweig formula" was applied to calculate the loss of sugar to impurities, the resulting sugar content and subsequent sugar yield of each line (Buchholz et al. 1995):

$$Standard\ Molasses\ Loss\ (SML) = 0.12(Na + K) + 0.24(\alpha - N) + 0.48$$

$$Corrected\ Sugar\ Content\ (CSC) = SC - SML - SFL$$

$$Corrected\ Sugar\ yield = beet\ yield\ \times CSC$$

Where Na + K is the sum of Sodium and Potassium in mmol/100g of beet, α-N is α-amino Nitrogen in mmol/100g of beet, SC is sugar content of beet and SFL is Standard Factory Loss of 0.6.

The mid-parent value for all traits was calculated as follows:

$$\frac{Mean\ of\ Parent\ 1}{(Mean\ of\ Parent\ 1\ and\ Parent\ 2)}$$

All data points for Parent 1 were normalized around this mid-parent value. A one-tailed independent samples *t*-test was used to determine whether F1 hybrid means were significantly higher or lower than the best-parent or mid-parent mean. A two-way analysis of variance (ANOVA) with a post-hoc Tukey's HSD test was used to determine the influence of ploidy level and hybridity across different traits in female and male parent lines (Genotypes (EE), (AA), (BB), (AAAA), (BBBB)) and F1 hybrids (Genotypes (EA), (EB), (EAA), (EBB)). A one-way ANOVA with a post-hoc Tukey's HSD test was used to determine whether the F1 3x hybrids (Genotypes (EAA), (EBB), (FCC), (GDD)) differed for important agronomic traits and for nuclear multi-locus heterozygosity levels.

Table 2. Sugar beet germplasm.

	Genotype	Hybridity status	Experiment ID
2x	Double Haploid (AA)	Homozygous	2x DH (AA)
	Double Haploid (BB)	Homozygous	2x DH (BB)
	CMS line 'Tester 1' (EE)	Heterozygous	2x hybrid ('Tester 1') (EE)
	CMS line 'Tester 2' (FF)	Heterozygous	2x hybrid ('Tester 2') (FF)
	CMS line 'Tester 3' (GG)	Heterozygous	2x hybrid ('Tester 2') (GG)
4x	Tetraploid (AAAA)	Homozygous	4x (AAAA)
	Tetraploid (BBBB)	Homozygous	4x (BBBB)
	Tetraploid (CCCC)	Heterozygous	4x hybrid (CCCC)
	Tetraploid (DDDD)	Heterozygous	4x hybrid (DDDD)

CMS = Cytoplasmic Male Sterile; DH = Double Haploid

Table 3. Sugar beet parental lines used to generate F1 diploid and F1 triploid hybrid offspring.

♀ \ ♂	2x DH (AA)	4x (AAAA)	2x DH (BB)	4x (BBBB)	4x hybrid (CCCC)	4x hybrid (DDDD)
2x hybrid ('Tester 1') (EE)	F1 2x hybrid (EA)	F1 3x hybrid (EAA)	F1 2x hybrid (EB)	F1 3x hybrid (EBB)		
2x hybrid ('Tester 2') (FF)					F1 3x hybrid (FCC)	
2x hybrid ('Tester 3') (GG)						F1 3x hybrid (GDD)

DH = Double Haploid. Different genotypes specified in parentheses. A = A genotype; B = B genotype; C = C genotype; D = D genotype; E = genotype E; F = genotype F; G = genotype G.

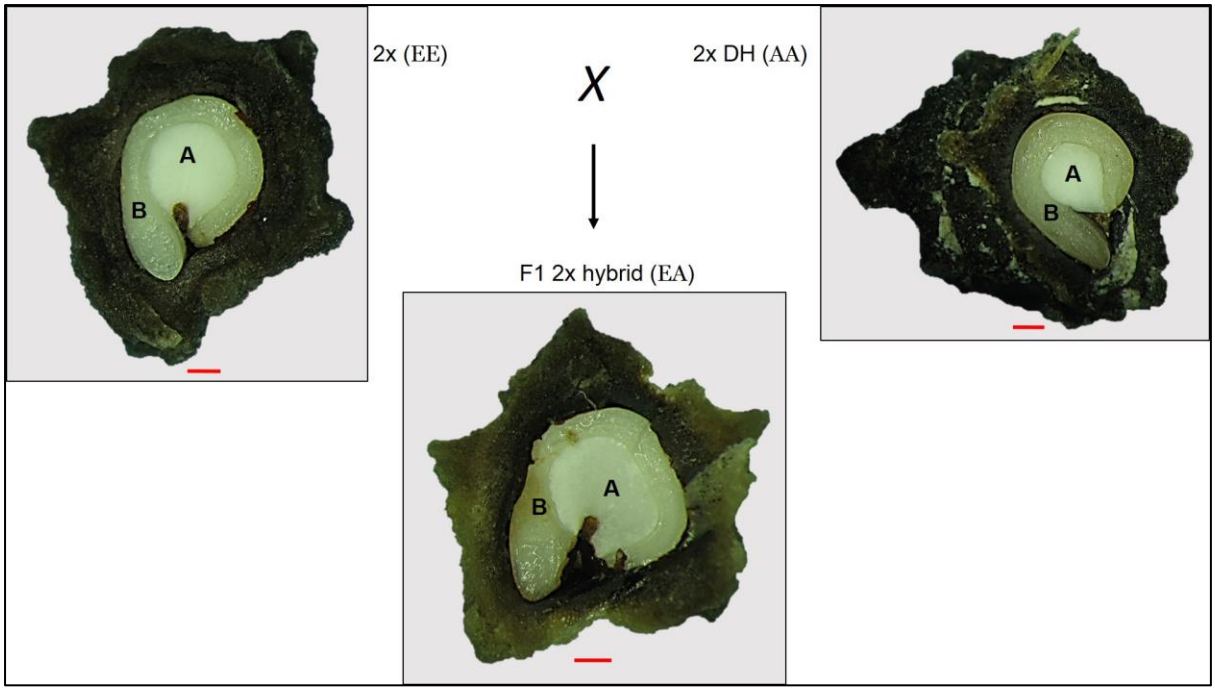


Figure 1. Cross-section of sugar beet fruits revealing seed inside. Representative F1 2x hybrid (EA) and its parent lines. Seed tissue is labelled A perisperm, and, B embryo. Red line is scale of 0.5mm.

Results

F1 diploid hybrids of sugar beet exhibit positive heterosis effects on seed traits

To determine the extent of heterosis effects on sugar beet seed traits at the diploid level, the viability and size of sugar beet seeds of the parental and F1 generations were analyzed (**Table *2*** and **Table *3***). For this analysis sugar beet seeds were removed from their fruits (**Figure *1***). The F1 2x hybrids (EA) and (EB) display heterosis effects with respect to seed viability ($P \leq 0.05$) and seed size ($P \leq 0.05$) (**Figure *2***A&B). To determine the extent of heterosis effects at the diploid level on tissue characteristics of the F1 hybrid seeds, seed cross-sections were investigated. This revealed that the F1 2x hybrid (EA) and (EB) seeds display heterosis ($P \leq 0.05$) for embryo size (**Figure *2***C). While the F1 2x hybrid (EA) seed display heterosis ($P \leq 0.05$) for perisperm size, there was no significant heterosis effect in perisperm size observed for the F1 2x hybrid (EB) ($P > 0.05$) (**Figure *2***D).

F1 diploid hybrids of sugar beet exhibit positive heterosis for root morphology, yield and sugar yield

To investigate possible heterosis effects on agronomic traits at the diploid level, a field trial of parental and F1 hybrid lines was conducted. At 121 T/ha, the female parent 2x (EE) has the highest yield of the parent lines. Both the F1 2x hybrid (EA) and F1 2x hybrid (EB) display heterosis ($P \leq 0.05$) for total yield (**Figure *3***A). However, both F1 2x hybrids (EA) and (EB) have a harvest plant density which is not significantly different ($P > 0.05$) from their mid-parent values. The F1 2x hybrid (EA) displays heterosis ($P \leq 0.05$) for above-ground biomass (i.e. the fresh weight of the leaves plus the root crown), whereas the F1 2x hybrid (EB) has an above-ground biomass not significantly different ($P = 0.15$) from the mid-parent value (Table S2). At 71 T/ha, the female parent 2x (EE) has the highest root yield of the parent lines. Both the F1 2x hybrid (EA) and F1 2x hybrid (EB) display heterosis for root yield ($P \leq 0.05$) (**Figure *3***B). Both of the F1 2x hybrids display heterosis ($P \leq 0.05$) for root circumference. For root length, only the F1 2x hybrid (EB) shows heterosis ($P \leq 0.05$) while the F1 2x hybrid (EA) does not have a significantly longer tap root ($P = 0.10$) than the mid-parent value (Table S2). To determine heterosis effects on root quality traits, the harvested sugar beets were analyzed for chemical composition. The corrected sugar content for all lines ranges between 13 and 14%; the F1 hybrids do not have significantly different sugar content than the mid-parent values ($P > 0.05$) (**Figure *3***C). At 10T/ha, the female parent 2x (EE) has the highest corrected sugar yield

of the parent lines. Both the F1 2x hybrid (EA) and F1 2x hybrid (EB) display heterosis for corrected sugar yield (**Figure *3***D).

Paternally-inherited genome dosage increase does not enhance heterosis effects in F1 triploid hybrids relative to F1 diploids hybrids

To determine whether a paternally-inherited genome dosage increase influences heterosis in F1 hybrids, F1 triploids were generated. The sterile mother, genotype (EE), was crossed with a tetraploid pollen parent (AAAA) and (BBBB) to generate F1 triploid offspring. This differed from use of a diploid pollen parent, (AA) and (BB), as used for the F1 diploid hybrids. These F1 triploids, (EAA) and (EBB), are genetically identical (at the DNA sequence level) to the F1 diploids, (EA) and (EB), apart from the triploids having an extra paternally-inherited chromosome set (**Table *3***).

F1 triploid hybrids of sugar beet exhibit positive heterosis effects on seed traits

Similar to their equivalent F1 2x hybrids, the F1 3x hybrid (EAA) and (EBB) display heterosis for seed viability ($P \leq 0.05$) and seed size ($P \leq 0.05$). The homozygous tetraploid male parents, 4x (AAAA) and (BBBB), have the lowest percentage of alive seeds among the parent lines (23% and 15%, respectively) (**Figure *2***A & B). With an embryo size area of 4mm^2 and 4.7mm^2 respectively, the homozygous tetraploid male parents 4x (AAAA) and (BBBB) are the best performing parents for F1 embryo size. Both the F1 3x hybrid (EAA) and F1 3x hybrid (EBB) seeds display heterosis ($P \leq 0.05$) for embryo size (**Figure *2***C). The F1 3x hybrid (EAA) exhibits heterosis ($P \leq 0.05$) for perisperm size, but there is no significant difference between F1 3x hybrid (EBB) and its parent lines for perisperm size ($P > 0.05$). The heterosis effect on perisperm size recorded in F1 triploids mimics that seen in their equivalent F1 diploids (**Figure *2***D).

F1 triploid hybrids of sugar beet exhibit positive heterosis for root morphology, yield and sugar yield

Both F1 3x hybrid (EAA) and F1 3x hybrid (EBB) display heterosis for total yield and root yield ($P \leq 0.05$) (**Figure *3***A & B). Both of the F1 3x hybrids display heterosis ($P \leq 0.05$) for harvest plant density, likely due to the low field emergence of parents 4x (AAAA) and (BBBB). This contrasts with their equivalent F1 2x hybrids which do not show a greater plant density than their parents. Both F1 3x hybrid (EAA) and F1 3x hybrid (EBB) display heterosis ($P \leq$

0.05) for above-ground biomass. Like their equivalent F1 2x hybrids, both F1 3x hybrids display heterosis ($P \leq 0.05$) for root circumference. For root length, both F1 3x hybrid (EAA) and F1 3x hybrid (EBB) display heterosis ($P \leq 0.05$) (Table S2). As also seen for both F1 2x hybrids, both F1 3x hybrids do not have significantly different sugar content than their mid-parent values ($P > 0.05$) (**Figure *3***C). Both F1 3x hybrids display heterosis ($P \leq 0.05$) for corrected sugar yield (**Figure *3***D).

F1 hybrids exhibit heterosis for important agronomic traits regardless of ploidy level

Comparisons of the F1 2x and 3x hybrids reveal that there are some differences in the levels of heterosis for certain important agronomic traits. For example, the F1 2x hybrids (EA) and (EB) have close to 100% seed viability whereas the F1 3x hybrid (EAA) and 3x hybrid (EBB) have circa 75% seed viability (**Figure *2***A). Also, both F1 2x hybrid (EA) and (EB) show best parent heterosis for root yield and sugar yield while both F1 3x hybrid (EAA) and 3x hybrid (EBB) show only mid-parent heterosis for these traits (**Figure *3***B & D).

Using the different levels of heterosis recorded across both F1 2x and 3x hybrids, a two-way analysis of variance (ANOVA) was performed to determine the relative influence of ploidy level and hybridity on important agronomic traits. Both the F1 2x hybrids (EA) and (EB) and both F1 3x hybrids (EAA) and (EBB) and their parent lines (EE), (AA), (BB), (AAAA) and (BBBB) were grouped according to ploidy level (diploid, triploid, tetraploid; factor 1) or hybridity (isogenic, hybrid; factor 2). A Tukey's HSD test detected any statistically significant difference between ploidy levels. Both ploidy level and hybridity significantly affect all important agronomic traits examined ($P \leq 0.05$) (**Table *4***). A Tukey's HSD test reveals ploidy level has a significant effect on seed viability between diploid and triploid lines ($P \leq 0.05$), whereas for all other traits examined – total yield, root yield, corrected sugar content, corrected sugar yield – there is no significant effect from ploidy level on diploid and triploid lines ($P > 0.05$) (**Table *5***). My data indicates that F1 hybrids perform better than their parents, regardless of their ploidy (genome dosage) level.

Homozygous and heterozygous tetraploid male parents produce F1 triploids with different nuclear multi-locus heterozygosity levels which exhibit largely equivalent heterosis

To investigate whether increased nuclear multi-locus heterozygosity in F1 hybrids in a polyploid system could affect heterosis in sugar beet, I compared the performance of F1 triploid

hybrids generated from test crosses using homozygous *versus* highly heterozygous male parents. To determine the effect of homozygous *versus* heterozygous tetraploid male parent on F1 triploid heterosis, a set of F1 triploid hybrids were analysed: F1 3x hybrid (EAA) and (EBB) share the same female parent, 2x (EE), and have homozygous tetraploid male parents derived from doubled haploidy and chromosome doubling, 4x (AAAA) and (BBBB), while F1 3x hybrid (FCC) and (GDD) have different female parents, 2x (FF) and (GG), and highly heterozygous tetraploid male parents, 4x (CCCC) and (DDDD) (**Table *2*** and **Table *3***). The F1 3x hybrid (FCC) and (GDD) have a higher extent of multi-locus nuclear heterozygosity than the F1 3x hybrid (EAA) and (EBB), as confirmed through genotyping-by-sequencing (**Figure *4***).

F1 triploid hybrids with heterozygous tetraploid male parents exhibit both positive and negative heterosis effects on seed traits

My results indicate that 100% of the highly heterozygous F1 3x hybrid (FCC) and (GDD) fruits are monogerm and these fruits contain 100% viable seeds, indicating heterosis for seed viability ($P \leq 0.05$) (**Figure *5***A, Table S3). The F1 3x hybrid (FCC) displays positive heterosis ($P \leq 0.05$) for seed size, while the F1 3x hybrid (GDD) displays negative heterosis ($P \leq 0.05$) for seed size (**Figure *5***B). The F1 3x hybrid (FCC) displays heterosis ($P \leq 0.05$) for embryo size. In contrast, the F1 3x hybrid (GDD) displays negative heterosis ($P \leq 0.05$) for embryo size (**Figure *5***C). Similarly, the F1 3x hybrid (FCC) displays heterosis ($P \leq 0.05$) for perisperm size but F1 3x hybrid (GDD) has a perisperm size not significantly different ($P = 0.10$) from the mid-parent value (**Figure *5***D).

F1 triploid hybrids with heterozygous tetraploid male parents do not exhibit a uniform heterotic response in relation to both root morphology, yield and sugar yield

The F1 3x hybrid (FCC) has a total yield and root yield not significantly different ($P = 0.37$, $P = 0.31$) than the mid-parent value, whereas F1 3x hybrid (GDD) displays heterosis for total yield and root yield ($P \leq 0.05$) (**Figure *6***A & B). The heterozygous male parents, 4x (CCCC) and (DDDD), both have a harvest plant density of 44.75 (Table S4), the highest recorded for any male parent in this experiment. The corresponding F1 3x hybrids have a similar harvest plant density to their mid-parent values ($P > 0.05$). For above-ground biomass, both the F1 3x hybrid (FCC) and F1 3x hybrid (GDD) do not significantly differ from their mid-parent values ($P > 0.05$). Both the F1 3x hybrid (FCC) and F1 3x hybrid (GDD) show no significant

difference ($P > 0.05$) in root length or root circumference from their mid-parent values (Table S4). The corrected sugar content for all F1 3x hybrids does not significantly differ from their mid-parent values ($P > 0.05$) (**Figure *6***C). The F1 3x hybrid (FCC) does not have a significantly different corrected sugar yield than the mid-parent value ($P = 0.31$). In contrast, the F1 3x hybrid (GDD) displays heterosis for corrected sugar yield ($P \leq 0.05$) (**Figure *6***D).

Triploid F1 hybrids do not differ for several important agronomic traits

To determine if there is a difference between triploid F1 hybrids with homozygous and heterozygous tetraploid male parents for important agronomic traits, a one-way ANOVA was performed to compare all four F1 triploid hybrids. A post-hoc Tukey's HSD test revealed means which are significantly different from one another. Seed viability is significantly different between groups ($P \leq 0.05$). The F1 3x hybrid (FCC) and F1 3x hybrid (GDD) have greater seed viability than F1 3x hybrid (EAA) and F1 3x hybrid (EBB). However, total yield, root yield, corrected sugar content, and corrected sugar yield are not significantly different ($P > 0.05$) (**Table *6*** and **Table *7***).

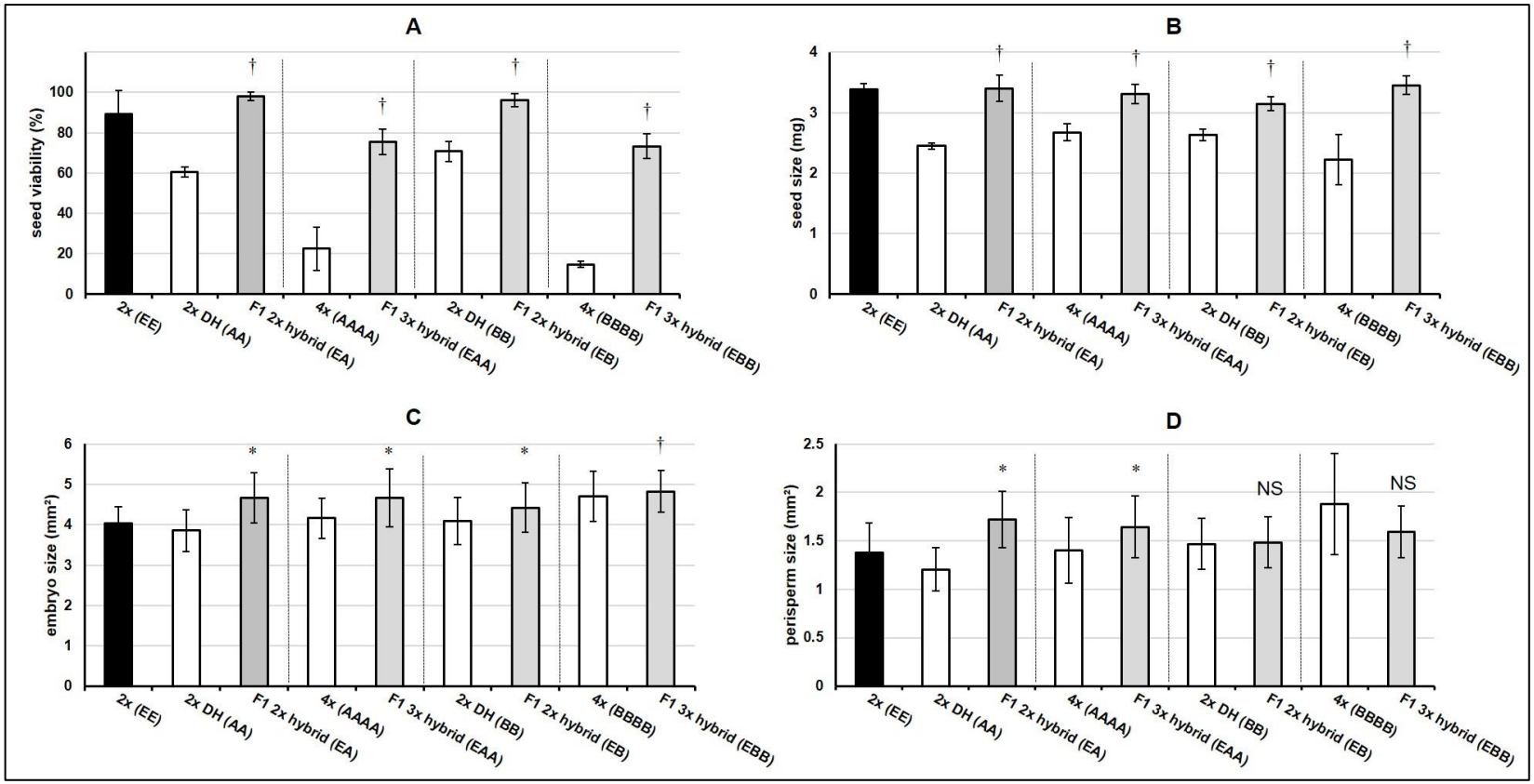


Figure 2. Sugar beet seed characteristics of diploid and triploid F1 hybrids and their parent lines. Data are mean of four replicates (± SD). A Seed viability, B seed size, C embryo size, D perisperm size. Different genotypes specified in parentheses. DH = double haploid. * Best parent heterosis (P ≤ 0.05), † Mid-parent heterosis (P ≤ 0.05), NS Not significantly different (P > 0.05).

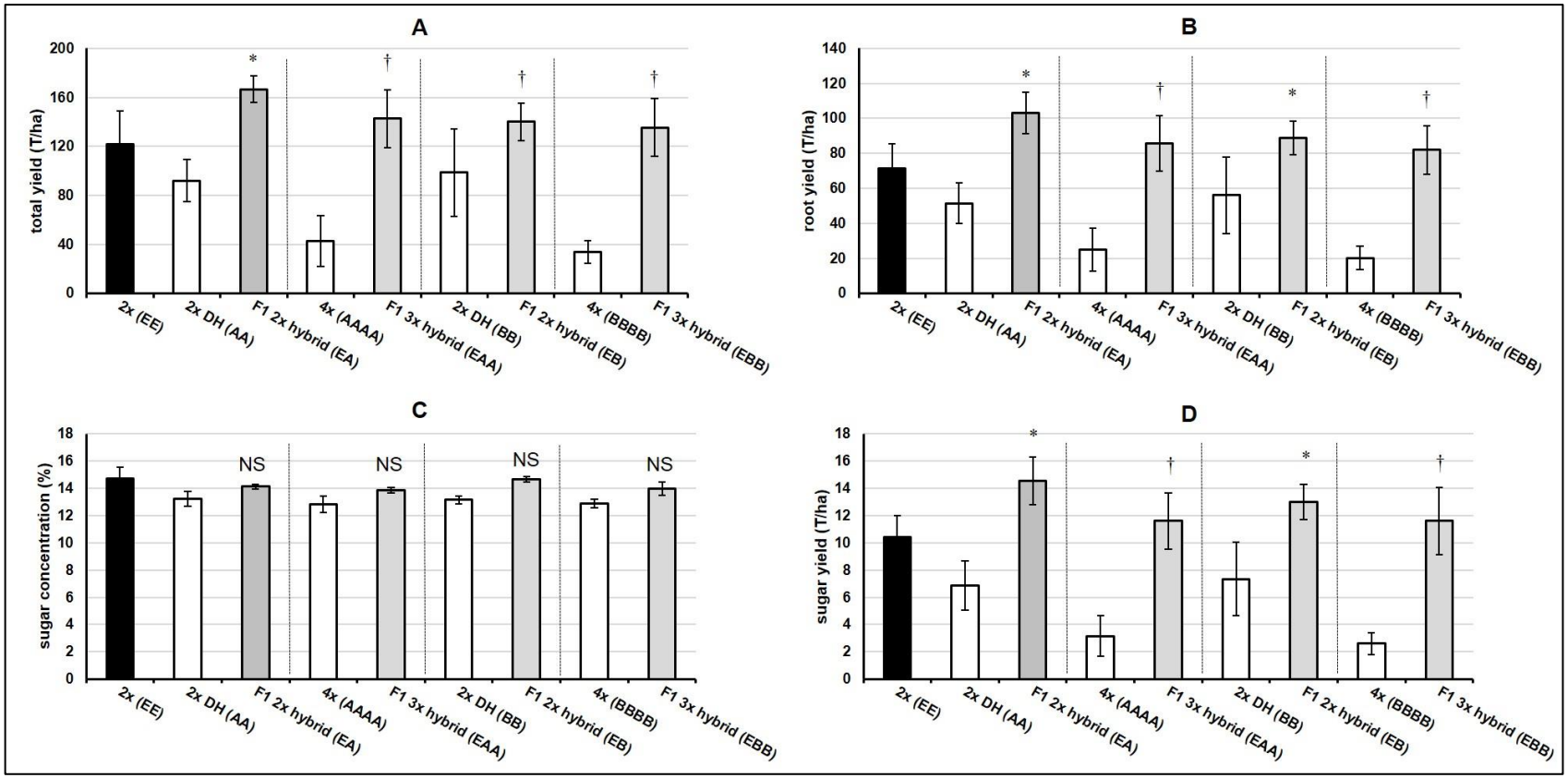


Figure 3. Agronomic and root quality traits of diploid and triploid F1 hybrids of sugar beet and their parental lines. Data are mean of four replicates (± SD). A Total yield, B root yield, C corrected sugar content, D corrected sugar yield. Different genotypes specified in parentheses. DH = double haploid. * Best parent heterosis ($P \leq 0.05$), † Mid-parent heterosis ($P \leq 0.05$), NS Not significantly different ($P > 0.05$).

Table 4. Two-way ANOVA results displaying significant effects of ploidy level and hybridity on important agronomic traits. Both F1 2x hybrids (EA) and (EB) and F1 3x hybrids (EAA) and (EBB) and their parent lines (EE), (AA), (BB), (AAAA) and (BBBB) were grouped according to ploidy level (diploid, triploid, tetraploid; factor 1) or hybridity (isogenic, hybrid; factor 2).

Source	Trait	Type III Sum of Squares	df	Mean Square	F	Sig.
Ploidy	Total yield	11600.86	2	5800.42	12.99	< 0.001
	Root yield	3418.44	2	1709.22	8.97	0.001
	Sugar content	2.00	2	1.00	3.80	0.033
	Sugar yield	93.92	2	46.96	9.98	< 0.001
	Seed viability	10781.19	2	5390.59	104.14	< 0.001
Hybridity	Total yield	12437.99	1	12437.99	27.85	< 0.001
	Root yield	6293.95	1	6293.95	33.03	< 0.001
	Sugar content	7.98	1	7.98	30.27	< 0.001
	Sugar yield	207.82	1	207.82	44.16	< 0.001
	Seed viability	3980.95	1	3980.95	76.91	< 0.001

Table 5. Tukey HSD results displaying significant differences between ploidy levels for important agronomic traits. Both F1 2x hybrids (EA) and (EB) and F1 3x hybrids (EAA) and (EBB) and their parent lines (EE), (AA), (BB), (AAAA) and (BBBB) were analyzed according to ploidy level (diploid, triploid, tetraploid).

Trait	(I) Ploidy	(J) Ploidy	Mean Difference (I-J)	Std. Error	Sig.
Total yield	Diploid	Triploid	-16.58	8.84	0.162
		Tetraploid	84.24	8.84	< 0.001
	Triploid	Tetraploid	100.81	10.57	< 0.001
Root yield	Diploid	Triploid	-10.70	5.77	0.169
		Tetraploid	50.67	5.77	< 0.001
	Triploid	Tetraploid	61.36	6.90	< 0.001
Sugar content	Diploid	Triploid	0.04	0.22	0.986
		Tetraploid	1.14	0.22	< 0.001
	Triploid	Tetraploid	1.11	0.26	< 0.001
Sugar yield	Diploid	Triploid	-1.59	0.91	0.202
		Tetraploid	8.66	0.91	< 0.001
	Triploid	Tetraploid	10.25	1.09	< 0.001
Seed viability	Diploid	Triploid	8.56	3.01	0.021
		Tetraploid	64.31	3.01	< 0.001
	Triploid	Tetraploid	55.75	3.60	< 0.001

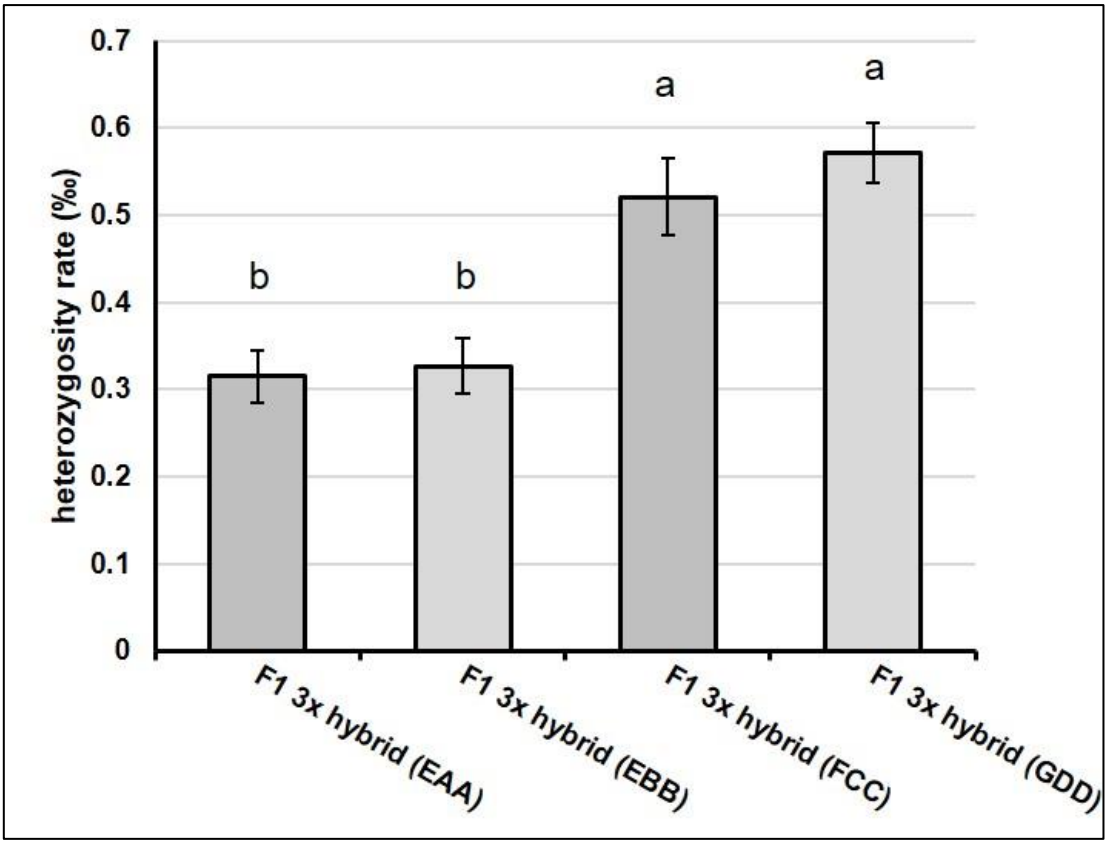


Figure 4. Mean genome-wide heterozygosity rate of F1 3x hybrids of sugar beet generated in this experiment. Heterozygosity rate is calculated by the ratio of heterozygous SNPs to the total number of genome bases. Data are mean of three replicates (± SD). F1 3x hybrids with heterozygous male parents, F1 3x hybrid (FCC) and (GDD), are significantly more heterozygous than F1 3x hybrids with homozygous male parents, F1 3x hybrid (EAA) and (EBB). Statistical differences were determined with a one-way ANOVA and Tukey's HSD test. Means assigned different letters are statistically different ($P < 0.05$).

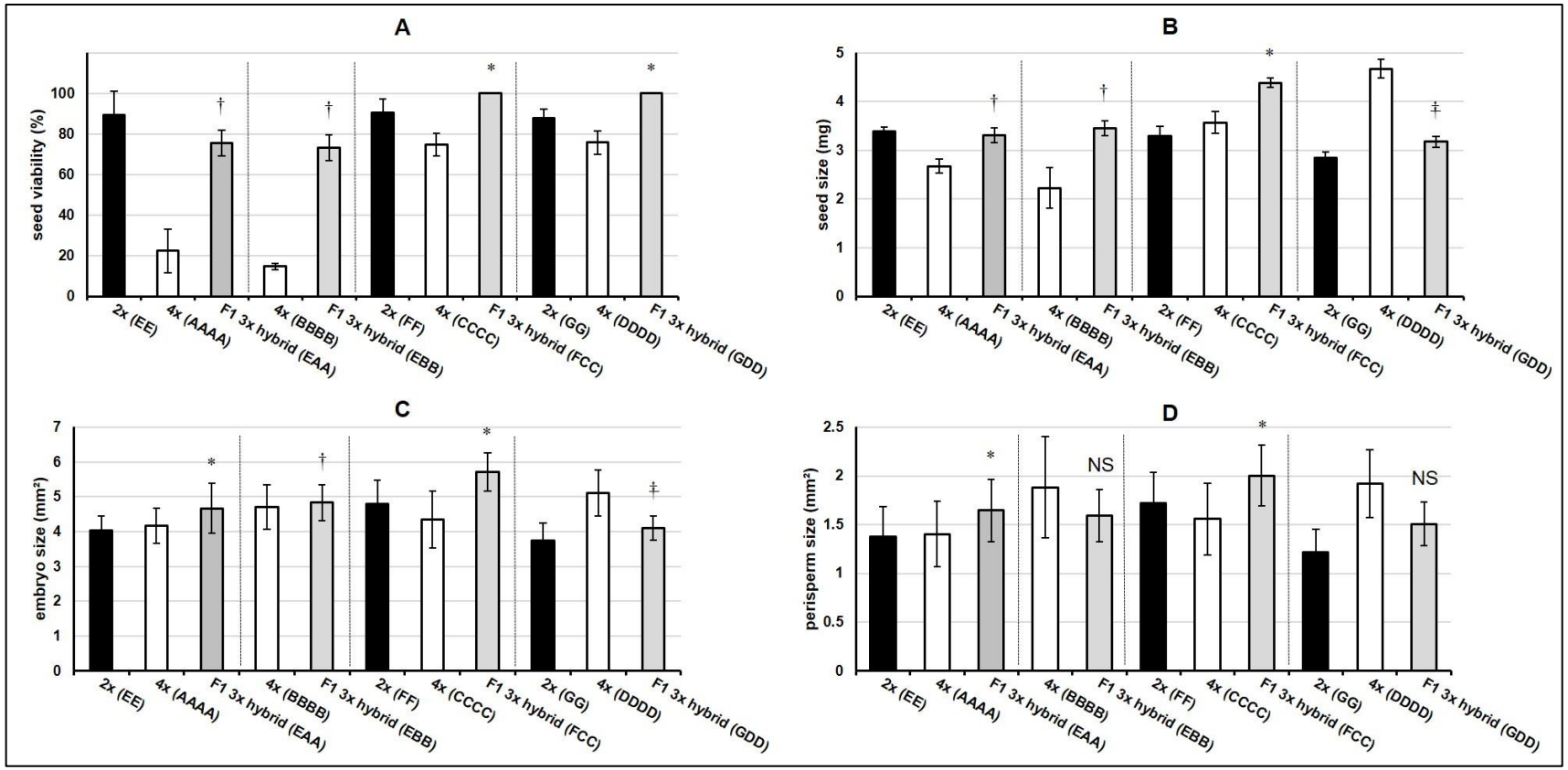


Figure 5. Sugar beer fruit and seed characteristics of FI triploid hybrids and their parent lines. Data are mean of four replicates (± SD). **A Seed viability, B seed size, C embryo size, D perisperm size. Different genotypes specified in parentheses.** * Best parent heterosis ($P \leq 0.05$), † Mid-parent heterosis ($P \leq 0.05$), NS Not significantly different ($P > 0.05$), ǂ Below mid-parent value ($P \leq 0.05$).

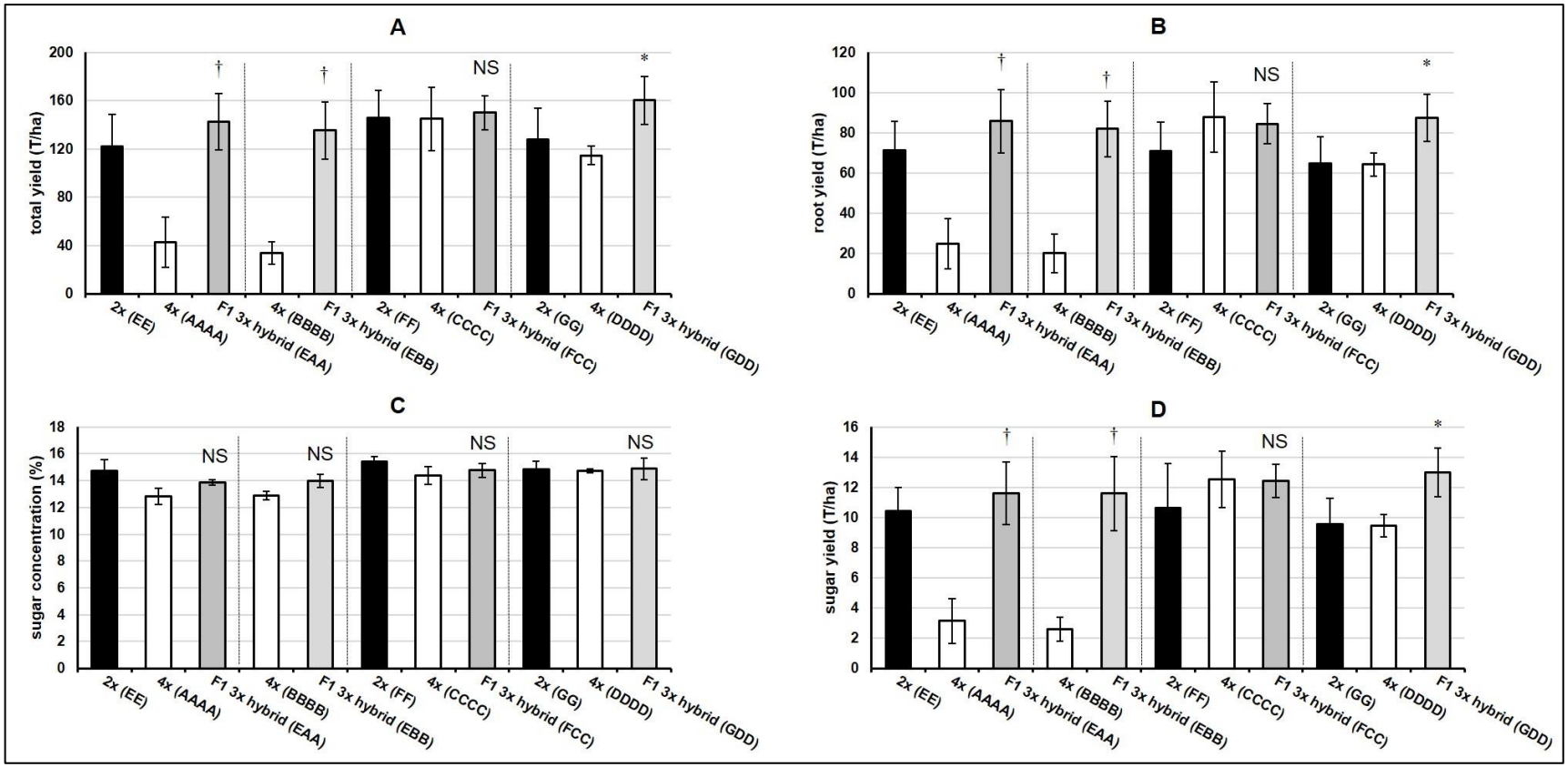


Figure 6. Agronomic and root quality traits of F1 triploid hybrids of sugar beet and their parental lines. Data are mean of four replicates (± SD). **A Total yield, B root yield, C corrected sugar content, D corrected sugar yield**. Different genotypes specified in parentheses. * Best parent heterosis ($P \leq 0.05$), † Mid-parent heterosis ($P \leq 0.05$), NS Not significantly different ($P > 0.05$).

Table 6. One-way ANOVA results displaying significant differences between F1 3x hybrids (EAA), (EBB), (FCC), and (GDD) for important agronomic traits.

Trait	Source of variation	Sum of Squares	df	Mean Square	F	Sig.
Total yield	Between Groups	1358.47	3	452.82	1.06	0.401
	Within Groups	5111.07	12	425.92		
	Total	6469.54	15			
Root yield	Between Groups	64.29	3	21.423	0.13	0.942
	Within Groups	2019.70	12	168.31		
	Total	2083.99	15			
Sugar content	Between Groups	3.04	3	1.01	2.69	0.093
	Within Groups	4.51	12	0.38		
	Total	7.56	15			
Sugar yield	Between Groups	5.26	3	1.75	0.43	0.734
	Within Groups	48.81	12	4.07		
	Total	54.07	15			
Seed viability	Between Groups	2638.25	3	879.412	44.37	< 0.001
	Within Groups	237.83	12	19.82		
	Total	2876.08	15			

Table 7. Tukey HSD result displaying significant differences between F1 3x hybrids (EAA), (EBB), (FCC), and (GDD) for seed viability.

Trait	(I) Genotype	(J) Genotype	Mean Difference (I-J)	Std. Error	Sig.
Seed viability	F1 3x hybrid (EAA)	F1 3x hybrid (EBB)	2.28	3.15	0.886
		F1 3x hybrid (FCC)	-24.49	3.15	< 0.001
		F1 3x hybrid (GDD)	-24.49	3.15	< 0.001
	F1 3x hybrid (EBB)	F1 3x hybrid (FCC)	-26.77	3.15	< 0.001
		F1 3x hybrid (GDD)	-26.77	3.15	< 0.001
	F1 3x hybrid (FCC)	F1 3x hybrid (GDD)	0.00	3.15	1.000

Different genotypes specified in parentheses.

Discussion

Increases in paternal genome dosage in F1 hybrids of sugar beet does not significantly enhance heterosis effects

In this study diploid and triploid F1 hybrids of sugar beet were generated which had either one paternally inherited nuclear genome set (i.e. EA, EB) or two paternally inherited nuclear genome sets (i.e. EAA, EBB) (**Table *3***). When the male sterile diploid line (Genotype EE) is crossed with pollen from a diploid homozygous doubled haploid (DH, Genotype AA), significant heterosis effects are seen in the diploid F1 hybrid (EA): mid-parent heterosis is regularly seen and in some instances best-parent heterosis (**Figure *2*** and **Figure *3***). However, when a genetically identical homozygous tetraploid pollen donor is used in the same cross (i.e. AAAA) there is no additional effect on heterosis in the triploid F1 hybrids (EAA) generated (**Figure *2***, **Figure *3*** and **Table *4***). Similarly, when a different diploid homozygous doubled haploid (DH, Genotype BB) is used, there is again no differential heterosis effect observed due to paternal genome dosage increase in the F1 hybrid triploid (EBB) relative to the F1 hybrid diploid (EB).

My results indicate a greater impact of hybridity over paternal genome dosage on heterosis effects for important agronomic traits in sugar beet (**Table *4*** and **Table *5***). Indeed, a paternal genome dosage increase effect is not statistically significant for total yield, root yield, and sugar yield in the pooled comparison of diploid and triploid lines of genotypes (A) and (B), suggesting that the difference between them is mainly due to the effects of hybridity. The only exception is seed viability, where the effect of paternal genome dosage increase is significant between diploids and triploids ($P \leq 0.05$) (**Table *5***): both F1 2x hybrids (EA) and (EB) have ~23% greater seed viability than their equivalent F1 3x hybrids (EAA) and (EBB), indicating a negative consequence of paternal genome dosage increase in these F1 hybrids for this trait (**Figure *2***).

Investigations into the effects of genome addition in sugar beet hybrids were conducted in the early literature (Hecker et al. 1970). In such studies, eight diploid inbred lines were converted to tetraploids, and crosses were performed to generate diploid and reciprocal triploid F1 hybrids. It was concluded that diploid and triploid F1 hybrids were largely equivalent with regards to root yield and sucrose content. However, the authors indicated that their study had experimental design issues such as low hybridization success, the absence of CMS lines, and poor field germination. Using the same inbred diploid and tetraploid lines, it was later reported that the resulting diploid and reciprocal triploid hybrids were “low-producing” and “not high

yielding", i.e. there was no heterosis (Smith et al. 1979). Another study, consisting of 120 hybrids, reported that F1 triploid hybrids averaged a 9% higher root yield than F1 diploid hybrids (Lasa et al. 1989). However, the lines used in this study were genetically diverse: the fifteen male parents originated from different European countries. In such studies, it was not possible to disaggregate genome dosage effects from hybridity effects on heterosis in sugar beet. In contrast, using my crossing design it is possible to investigate paternal genome dosage *versus* hybridity contributions to heterosis effects, through the use of the same CMS line, equivalent homozygous double haploid (AA, BB) male parents and homozygous tetraploid (AAAA, BBBB) male parents, reliable hybridization, and field emergence of F1 hybrids and parental lines. My results indicate that increasing paternal genome dosage does not enhance heterosis in sugar beet F1 hybrids. Indeed, I demonstrate that for some genotypes there is a negative heterosis effect on seed viability in sugar beet F1 hybrids.

Parental genome dosage effects on heterosis have previously been investigated in maize (Yao et al. 2013). Using the maize inbred lines B73 and Mo17, diploid F1 hybrids were generated, while a trifluralin procedure was used to generate triploid F1 hybrids with two paternally derived genome sets. The diploid *versus* triploid F1 hybrids differed only in relation to paternal genome dosage: reciprocal F1 diploid hybrids inherited one genome from each parent, while the F1 triploid hybrids inherited either one maternal genome from B73 and two paternal genomes from Mo17, or, one maternal genome from Mo17 and two paternal genomes from B73. Analogous to my results in sugar beet, the F1 diploid and F1 triploid maize hybrids (with two paternally derived genome sets) exhibit largely equivalent heterosis over their parents, e.g. the F1 diploid and F1 triploid hybrids displayed equivalent mid-parent heterosis for plant height, leaf length and no. of tassel branches (Yao et al. 2013). The maize F1 triploid hybrids showed a different heterotic response for a number of agronomic traits relative to their corresponding F1 diploid hybrids, e.g. F1 triploid hybrids had reduced ear length and took a longer period to reach anther emergence (Yao et al. 2013). Unlike my study however, the authors were able to perform reciprocal crosses at the diploid level, thus allowing them to differentiate parent-of-origin effects from genome dosage effects. One of the F1 triploid hybrids exhibited higher levels of heterosis than the other F1 triploid hybrid, while the reciprocal F1 diploid hybrids displayed equal levels of heterosis (Yao et al. 2013). The authors concluded that the difference between F1 triploid hybrids is not due to a parent-of-origin effect but rather due to genome dosage effects that can depend on the genotype in question (Yao et al. 2013).

Increasing the paternal genome dosage in *A. thaliana* has revealed a paternal genome dosage effect on heterosis for seed size and leaf area (Miller et al. 2012; Fort et al. 2016). Unlike maize, this is not a genotype-dependent genome dosage effect, as the paternal genome dosage effects on heterosis have been demonstrated across multiple genotypes, where F1 triploid hybrids with two paternal genomes have larger seeds and leaf area than F1 diploid hybrids (Miller et al. 2012; Fort et al. 2016). In my experiments with sugar beet, a genotype-dependent paternal genome dosage effect on F1 hybrids was not observed.

Increasing heterozygosity in F1 triploid hybrids does not enhance heterosis effects

Heterozygosity is one measure of genetic variation. At a given locus in a plant genome, heterozygosity refers to the presence of different alleles for the same gene (e.g. if diploid parents are homozygous for different pairs of the same gene, say A_1/A_1 and A_2/A_2, their offspring will inherit allele A_1 and allele A_2 and display heterozygosity at this locus, A_1/A_2). Thus, the level of heterozygosity at all nuclear gene loci cumulatively determines the extent of multi-locus heterozygosity for any genotype. Genotyping-by-sequencing (GBS) analysis can be used to identify single nucleotide polymorphisms (SNPs) across the nuclear genome, and has been used to calculate rates of multi-locus heterozygosity of genotypes in cotton (Islam et al. 2015), soybean (*Glycine max*) (Sonah et al. 2013), *Miscanthus sinensis* (Ma et al. 2012), maize (Azmach et al. 2013), and yams (*Dioscorea*) (Girma et al. 2014).

The F1 3x hybrids generated in this study have different tetraploid male parents. The homozygous 4x (AAAA) and 4x (BBBB) genotypes have been generated by spontaneous chromosome doubling during double-haploid production and each have four identical sets of chromosomes that are homozygous. In contrast, the 4x (CCCC) and 4x (DDDD) genotypes are highly heterozygous lines (i.e. $C^1C^2C^3C^4$ and $D^1D^2D^3D^4$) used in sugar beet breeding. The resulting F1 3x hybrids can therefore paternally inherit either two homozygous genomes (to generate F1 3x hybrids (EAA) and (EBB)) or two heterozygous genomes (to generate F1 3x hybrids (FCC) and (GDD)). As a result, the heterozygosity rate (i.e. the ratio of heterozygous SNPs to the total number of genome bases) is higher for the F1 3x hybrids (FCC) and (GDD), when compared to the F1 3x hybrids (EAA) and (EBB) (**Figure *4***).

The one-way ANOVA and post-hoc test results (**Table *6*** and **Table *7***) reveal that the F1 3x hybrids are similar for the majority of traits measured in the field. The only difference observed between them is for seed viability, where the fruits of F1 3x hybrid (FCC) and (GDD) show improved seed viability (**Table *7***). Due to the unrelated parents of the F1 3x hybrids used

in this experiment it is difficult to determine whether increased nuclear heterozygosity or genotype is responsible for this effect, although it is noteworthy that the fruits of F1 3x hybrids (FCC) and (GDD) are both 100% monogerm which could be a contributing factor (Table S3).

My findings indicate that there may be no positive relationship between levels of nuclear multi-locus heterozygosity and heterosis effects for traits such as total yield and root yield in sugar beet. Some researchers have found that higher levels of multi-locus heterozygosity in oilseed rape (*Brassica napus*) leads to greater heterosis i.e. the more genetically divergent the parent lines, the greater the expression of heterosis in intraspecific crosses (Ali et al. 1995; Riaz et al. 2001). However, in another study in oilseed rape no relationship between genetic distance and heterosis was found (Diers et al. 1996). Likewise, there is conflicting evidence in maize, where early studies suggested heterosis increases with increasing parental genetic distance up to a certain threshold (Moll et al. 1965), but recent investigations have not found a correlation between extent of heterosis effects and parental genetic distance in maize (Benchimol et al. 2000; Frisch et al. 2010; Reif et al. 2010). A significant body of research has concluded that genetic distance does not correlate well with heterosis for intraspecific crosses, as seen in bread wheat (*Triticum aestivum)* (Martin et al. 1995), rice (*Oryza sativa*) (Zhang et al. 1995), and pepper (*Capsicum annum*) (Geleta et al. 2004). My findings in sugar beet (albeit based on a limited number of parental genotypes) suggest that there is no axiomatic relationship between parental genetic distance (and corresponding multi-locus heterozygosity levels) and heterosis effects for the agronomic traits analyzed.

Genome dosage effects on sugar beet seed biology

The seed biology of sugar beet differs from the cereal crop maize and the model organism *A. thaliana* (**Table *8***). In sugar beet seed development, the maternal nucellus is not fully digested during maturation and leads to the perisperm tissue which accumulates starch reserves (Hermann et al. 2007). The perisperm starch reserve plays a nutritive role in relation to seed germination and early (approx. first 7 days) growth (Lawrence et al. 1990; Elamrani et al. 1992; Catusse et al. 2008). While sugar beet seeds arising from crosses between diploid parents contain triploid endosperm tissue (which is a fertilization product), the endosperm tissue is not as extensive as the diploid maternal perisperm tissue (Hermann et al. 2007). In maize and *A. thaliana* the main nutritive source for the seed embryo is the endosperm tissue, which is persistent in maize and transient in *A. thaliana* (Baroux et al. 2002; Costa et al. 2004). Paternal

genome dosage increases can only directly impact on the fertilization products (i.e. the embryo and endosperm) and not on the maternal perisperm (**Table *8***). If the maternal perisperm is the main nutritive tissue supporting embryo development in sugar beet this could explain why paternal genome dosage heterosis effects are not evident in sugar beet. In maize and *A. thaliana* seed, in contrast, the main nutritive tissue (the endosperm) genome dosage can be increased from 2m:1p in F1 diploids to 2m:2p in F1 triploids, with greater potential for genome dosage effects on the main nutritive tissue supporting embryo growth and development.

Hybridity can trigger increased root size without reductions in sugar concentrations in sugar beet

Heterosis breeding in sugar beet is challenged by the inverse relationship between root size and sugar content. Many studies have shown that varieties with a high root yield have low sucrose concentrations (Bergen 1967; Oldemeyer 1975; Carter 1987). An ideal sugar beet variety is one that combines high root yield with high sucrose concentration (plus other characteristics such as low levels of impurities and strong abiotic and biotic stress tolerance). My study demonstrates that there is no heterotic effect on sugar concentration in the sugar beet F1 hybrids (**Figure *3***C and **Figure *6***C). The moderate negative relationship (Pearson's correlation coefficient, r = -0.47) between root size and sugar concentration in a pooled analysis of all F1 hybrids suggests that root size can be increased while sugar concentration remains stable. The most extreme example of this is F1 2x hybrid (EA) which has a root yield of ~32 T/ha greater than that of the best parent, while the sucrose concentration has not changed (**Figure *3***C). The identification of the molecular mechanisms which can allow increased root size without reductions in sugar concentrations could provide new avenues for sustainable intensification by increasing sugar yield per unit area.

Table 8. Parental genome dosage of seed tissue in F1 diploid and F1 triploid hybrids generated in this experiment.

Tissue	♀ ♂ **2x *X* 2x**	♀ ♂ **2x *X* 4x**
Embryo	1m:1p	1m:2p
Endosperm	2m:1p	2m:2p
Perisperm	2m:0p	2m:0p

m = maternal genome contribution; p = paternal genome contribution.

Conclusions

My results demonstrate that increasing the paternal genome dosage in F1 triploid hybrids of sugar beet does not enhance heterosis effects beyond what can be achieved in F1 diploid hybrids. Furthermore, increasing the extent of paternally inherited nuclear multi-locus heterozygosity in F1 triploid hybrids also does not enhance the heterosis effect, suggesting there is no axiomatic relationship between heterozygosity levels and heterosis in sugar beet F1 triploids. My results suggest that heterosis gains for important agronomic traits (e.g. root yield, sugar yield) in sugar beet can largely be achieved at the diploid breeding level.

References

Ali, M., L.O. Copeland, S.G. Elias, and J.D. Kelly, 1995 Relationship between genetic distance and heterosis for yield and morphological traits in winter canola (Brassica napus L.). *Theoretical and Applied Genetics* 91 (1):118-121.

Azmach, G., M. Gedil, A. Menkir, and C. Spillane, 2013 Marker-trait association analysis of functional gene markers for provitamin A levels across diverse tropical yellow maize inbred lines. *BMC plant biology* 13 (1):227.

Baroux, C., C. Spillane, and U. Grossniklaus, 2002 Evolutionary origins of the endosperm in flowering plants. *Genome Biol* 3 (9):reviews1026.

Benchimol, L.L., C.L. de Souza, A.A.F. Garcia, P.M.S. Kono, C.A. Mangolin *et al.*, 2000 Genetic diversity in tropical maize inbred lines: heterotic group assignment and hybrid performance determined by RFLP markers. *Plant Breeding* 119 (6):491-496.

Bergen, P., 1967 Seasonal patterns of sucrose accumulation and weight increase in sugar beets. *J Am Soc Sugar Beet Technol* 14:538-545.

Biancardi, E., J.M. McGrath, L.W. Panella, R.T. Lewellen, and P. Stevanato, 2010 Sugar beet, pp. 173-219 in *Root and tuber crops*. Springer.

Buchholz, K., B. Märländer, H. Puke, H. Glattkowski, and K. Thielecke, 1995 Neubewertung des technischen Wertes von Zuckerrueben. *Zuckerindustrie*.

Carter, J., 1987 Sucrose production as affected by root yield and sucrose concentration of sugarbeets. *Journal of the American Society of Sugar Beet Technologists* 24 (1):14-31.

Catusse, J., J.M. Strub, C. Job, A. Van Dorsselaer, and D. Job, 2008 Proteome-wide characterization of sugarbeet seed vigor and its tissue specific expression. *Proc Natl Acad Sci U S A* 105 (29):10262-10267.

Chen, Z.J., 2010 Molecular mechanisms of polyploidy and hybrid vigor. *Trends in Plant Science* 15 (2):57-71.

Comai, L., 2005 The advantages and disadvantages of being polyploid. *Nature reviews. Genetics* 6 (11):836.

Costa, L.M., J.F. Gutièrrez-Marcos, and H.G. Dickinson, 2004 More than a yolk: the short life and complex times of the plant endosperm. *Trends in Plant Science* 9 (10):507-514.

De Whalley, H.C.S., 2013 *ICUMSA methods of sugar analysis: official and tentative methods recommended by the International Commission for Uniform Methods of Sugar Analysis (ICUMSA)*: Elsevier.

Diers, B., P. McVetty, and T. Osborn, 1996 Relationship between heterosis and genetic distance based on restriction fragment length polymorphism markers in oilseed rape (Brassica napus L.). *Crop Science* 36 (1):79-83.

Donoghue, M.T., A. Fort, R. Clifton, X. Zhang, P.C. McKeown *et al.*, 2013 CmCGG methylation-independent parent-of-origin effects on genome-wide transcript levels in isogenic reciprocal F1 triploid plants. *DNA research* 21 (2):141-151.

Donoghue, M.T., A. Fort, R. Clifton, X. Zhang, P.C. McKeown *et al.*, 2014 C(m)CGG methylation-independent parent-of-origin effects on genome-wide transcript levels in isogenic reciprocal F1 triploid plants. *DNA Res* 21 (2):141-151.

Draycott, A.P., 2008 *Sugar beet*: John Wiley & Sons.

Duszynska, D., P.C. McKeown, T.E. Juenger, A. Pietraszewska-Bogiel, D. Geelen *et al.*, 2013a Gamete fertility and ovule number variation in selfed reciprocal F1 hybrid triploid plants are heritable and display epigenetic parent-of-origin effects. *New Phytol* 198 (1):71-81.

Duszynska, D., P.C. McKeown, T.E. Juenger, A. Pietraszewska-Bogiel, D. Geelen *et al.*, 2013b Gamete fertility and ovule number variation in selfed reciprocal F1 hybrid triploid

plants are heritable and display epigenetic parent-of-origin effects. *New Phytologist* 198 (1):71-81.

Elamrani, A., P. Raymond, and P. Saglio, 1992 Nature and utilization of seed reserves during germination and heterotrophic growth of young sugar beet seedlings. *Seed Science Research* 2 (1):1-8.

Fort, A., P. Ryder, P.C. McKeown, C. Wijnen, M.G. Aarts *et al.*, 2016 Disaggregating polyploidy, parental genome dosage and hybridity contributions to heterosis in Arabidopsis thaliana. *New Phytologist* 209 (2):590-599.

Fort, A., R. Tuteja, M. Braud, P.C. McKeown, and C. Spillane, 2017 Parental-genome dosage effects on the transcriptome of F1 hybrid triploid embryos of Arabidopsis thaliana. *Plant J* 92 (6):1044-1058.

Frisch, M., A. Thiemann, J. Fu, T.A. Schrag, S. Scholten *et al.*, 2010 Transcriptome-based distance measures for grouping of germplasm and prediction of hybrid performance in maize. *Theoretical and Applied Genetics* 120 (2):441-450.

Geleta, L., M. Labuschagne, and C. Viljoen, 2004 Relationship between heterosis and genetic distance based on morphological traits and AFLP markers in pepper. *Plant Breeding* 123 (5):467-473.

Girma, G., K.E. Hyma, R. Asiedu, S.E. Mitchell, M. Gedil *et al.*, 2014 Next-generation sequencing based genotyping, cytometry and phenotyping for understanding diversity and evolution of guinea yams. *Theoretical and Applied Genetics* 127 (8):1783-1794.

Guo, M., D. Davis, and J.A. Birchler, 1996 Dosage Effects on Gene Expression in a Maize Ploidy Series. *Genetics* 142 (4):1349-1355.

Hecker, R., R. Stafford, R. Helmerick, and G. Maag, 1970 Comparison of the same sugarbeet F1 hybrids as diploids, triploids and tetraploids. *Journal of the American Society of Sugar Beet Technologists* 16 (2):106-116.

Hedtke, S.M., and D.M. Hillis, 2011 The Potential Role of Androgenesis in Cytoplasmic–Nuclear Phylogenetic Discordance. *Systematic Biology* 60 (1):87-96.

Hermann, K., J. Meinhard, P. Dobrev, A. Linkies, B. Pesek *et al.*, 2007 1-Aminocyclopropane-1-carboxylic acid and abscisic acid during the germination of sugar beet (Beta vulgaris L.): a comparative study of fruits and seeds. *Journal of Experimental Botany* 58 (11):3047-3060.

Islam, M.S., G.N. Thyssen, J.N. Jenkins, and D.D. Fang, 2015 Detection, validation, and application of genotyping-by-sequencing based single nucleotide polymorphisms in upland cotton. *The Plant Genome* 8 (1).

Jiang, K., K.L. Liberatore, S.J. Park, J.P. Alvarez, and Z.B. Lippman, 2013 Tomato yield heterosis is triggered by a dosage sensitivity of the florigen pathway that fine-tunes shoot architecture. *PLoS Genet* 9 (12):e1004043.

Kenter, C., C. Hoffmann, and B. Märländer, 2006 Sugarbeet as raw material–Advanced storage management to gain good processing quality/Optimierung der Rohstoffqualität von Zuckerrüben durch verbessertes Lagerungsmanagement. *Zuckerindustrie* 131:706-720.

Lasa, J., I. Romagosa, R. Hecker, and J. Sanz, 1989 Combining ability in diploid and triploid sugarbeet hybrids from diverse parents. *Journal of sugar beet research* 26:10-18.

Lawrence, D.M., P. Halmer, and D.J. Bowles, 1990 Mobilisation of storage reserves during germination and early seedling growth of sugar beet. *Physiologia Plantarum* 78 (3):421-429.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25 (14):1754-1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* 25 (16):2078-2079.

Ma, X.-F., E. Jensen, N. Alexandrov, M. Troukhan, L. Zhang *et al.*, 2012 High resolution genetic mapping by genome sequencing reveals genome duplication and tetraploid genetic structure of the diploid Miscanthus sinensis. *PloS one* 7 (3):e33821.

Martin, J., L. Talbert, S. Lanning, and N. Blake, 1995 Hybrid performance in wheat as related to parental diversity. *Crop Science* 35 (1):104-108.

Meyer, R.C., O. Törjék, M. Becher, and T. Altmann, 2004 Heterosis of Biomass Production in Arabidopsis. Establishment during Early Development. *Plant Physiology* 134 (4):1813-1823.

Miller, M., C. Zhang, and Z.J. Chen, 2012 Ploidy and hybridity effects on growth vigor and gene expression in Arabidopsis thaliana hybrids and their parents. *G3: Genes, Genomes, Genetics* 2 (4):505-513.

Moll, R., J. Lonnquist, J.V. Fortuno, and E. Johnson, 1965 The relationship of heterosis and genetic divergence in maize. *Genetics* 52 (1):139.

Oldemeyer, R., 1975 Introgressive hybridization as a breeding method in Beta vulgaris. *J Am Soc Sugar Beet Technol* 18:269-273.

Paterson, A.H., and J.F. Wendel, 2015 Unraveling the fabric of polyploidy. *Nature biotechnology* 33 (5):491-493.

Peto, F., and J. Boyes, 1940 Comparison of diploid and triploid sugar beets. *Canadian Journal of Research* 18 (7):273-282.

Reif, J.C., S. Fischer, T.A. Schrag, K.R. Lamkey, D. Klein *et al.*, 2010 Broadening the genetic base of European maize heterotic pools with US Cornbelt germplasm using field and molecular marker data. *Theoretical and Applied Genetics* 120 (2):301-310.

Riaz, A., G. Li, Z. Quresh, M.S. Swati, and C.F. Quiros, 2001 Genetic diversity of oilseed Brassica napus inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. *Plant Breeding* 120 (5):411-415.

Roulin, A., P.L. Auer, M. Libault, J. Schlueter, A. Farmer *et al.*, 2013 The fate of duplicated genes in a polyploid plant genome. *Plant J* 73 (1):143-153.

Ryder, P., P.C. McKeown, A. Fort, and C. Spillane, 2014 Epigenetics and heterosis in crop plants, pp. 13-31 in *Epigenetics in Plants of Agronomic Importance: Fundamentals and Applications*. Springer.

Savitsky, V., 1950 Monogerm sugar beets in the United States. *Proc. Amer. Soc. Sugar Beet Tech*:156-159.

Schnable, P.S., and N.M. Springer, 2013 Progress toward understanding heterosis in crop plants. *Annual review of plant biology* 64:71-88.

Smith, G., R. Hecker, and S. Martin, 1979 Effects of ploidy level on the components of sucrose yield and quality in sugarbeet. *Crop Science* 19 (3):319-323.

Sonah, H., M. Bastien, E. Iquira, A. Tardivel, G. Légaré *et al.*, 2013 An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PloS one* 8 (1):e54603.

Stokes, D., C. Morgan, C. O'Neill, and I. Bancroft, 2007 Evaluating the utility of Arabidopsis thaliana as a model for understanding heterosis in hybrid crops. *Euphytica* 156 (1-2):157-171.

Tan, C., Q. Pan, C. Cui, Y. Xiang, X. Ge *et al.*, 2016 Genome-Wide Gene/Genome Dosage Imbalance Regulates Gene Expressions in Synthetic Brassica napus and Derivatives (AC, AAC, CCA, CCAA). *Front Plant Sci* 7:1432.

Van Dyke, F., 2008 *Conservation biology: foundations, concepts, applications*: Springer Science & Business Media.

Wang, J., L. Tian, H.-S. Lee, N.E. Wei, H. Jiang *et al.*, 2006 Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. *Genetics* 172 (1):507-517.

Wang, K., M. Li, and H. Hakonarson, 2010 ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research* 38 (16):e164-e164.

Washburn, J.D., and J.A. Birchler, 2014 Polyploids as a "model system" for the study of heterosis. *Plant Reprod* 27 (1):1-5.

Yao, H., A. Dogra Gray, D.L. Auger, and J.A. Birchler, 2013 Genomic dosage effects on heterosis in triploid maize. *Proceedings of the National Academy of Sciences* 110 (7):2665-2669.

Zhang, Q., Y. Gao, M.S. Maroof, S. Yang, and J. Li, 1995 Molecular divergence and hybrid performance in rice. *Molecular breeding* 1 (2):133-142.

Chapter 3. Evidence for progressive heterosis in *Arabidopsis thaliana*

Note: crosses were performed by Dr Marcus McHale and Ms. Rachel Soden

Introduction

Heterosis can be described as the phenomenon of an increase in quantitative traits (e.g. plant size, fertility, yield) in a first-generation offspring plant (filial 1; F1) beyond that observed in parental lines (East 1936; Shull 1948; Hayes et al. 1955). Commercial F1 hybrid seed is of enormous importance to farmers worldwide. F1 hybrid varieties possessing agronomic traits of interest have contributed between 10 – 50% of the yield increases witnessed since the mid-20th century for crops such as maize (*Zea mays*), sorghum (*Sorghum bicolor*), sugar beet (*Beta vulgaris*), upland cotton (*Gossypium hirsutum*) soybean (*Glycine max*) and bread wheat (*Triticum aestivum*) among others (Miller and Kebede 1984; Russell 1991; Brancourt-Hulmel et al. 2003; Duvick 2005; Egli 2008; Campbell et al. 2012; Loel et al. 2014). However, current and predicted shifts in atmospheric carbon dioxide, temperature and precipitation, combined with recent 'yield plateaus' in some of the our most valuable crops, have highlighted the need for continued improvement of F1 hybrid breeding programmes (Fischer 1996; Lobell and Field 2007). Furthermore, a 'business as usual' approach to crop production which includes land expansion and a heavy reliance on artificial fertilisers and pesticides is undesirable due to the associated environmental degradation, and regardless such an approach will likely fail to meet future rising food demands (Tester and Langridge 2010). To further exploit the heterosis phenomenon in hybrid breeding will require investigations into the different modes under which it can be enhanced or ablated.

Integrating polyploidy with the heterosis phenomenon has been largely ignored in the heterosis literature (Washburn and Birchler 2014). Polyploid plants are those that contain more than two chromosome sets in their nucleus. There is evidence of a genome dosage component to heterosis, i.e. the heterotic response at the diploid level is not necessarily a reliable predictor of the heterotic response at higher ploidy levels. For example, consider the F1 heterosis exhibited in comparable F1 diploid and triploid hybrids of both *Arabidopsis thaliana* and maize. F1 triploid hybrids of *A. thaliana* exhibit greater heterosis for both seed size and leaf area than F1 diploid hybrids providing the extra genome is inherited paternally (Miller et al. 2012; Fort et al. 2016). In maize, similar paternal genome dosage effects on F1 triploid hybrids have been demonstrated, although unlike *A. thaliana* appear to be genotype-dependent (Yao et al. 2013).

Furthermore, there is the phenomenon known as 'progressive heterosis'. Utilising autopolyploid plants, that is, polyploid plants with identical genomes, it has been demonstrated that increasing the heterozygosity through a double-cross substantially increases the heterosis

effect in tetraploids while having a minimal effect on diploids. A double-cross tetraploid hybrid plant can potentially harbour four different alleles at a locus, whereas a single cross tetraploid hybrid can harbour a maximum of two (**Figure 7**). This suggests that maximising multi-locus nuclear heterozygosity (e.g. tetra-allelic locus *abcd*) maximises the heterotic response, and heterosis declines thereafter (e.g. *abcd* > *abcc* > *aabb* > *aaaa*) (Fehr 1991; Riddle and Birchler 2008). Progressive heterosis effects for yield were observed in autotetraploids of potato (*Solanum tuberosum*), alfalfa (*Medicago sativa*), and maize (Mok and Peloquin 1975; Groose et al. 1989b; Sockness and Dudley 1989; Bingham et al. 1994; Riddle and Birchler 2008) as well as limited evidence for the same phenomenon in rice (*Oryza sativa*) (Tu et al. 2007; Wu et al. 2013).

The use of *A. thaliana*, maize, rice and tomato (*Solanum lycopersicum*) in heterosis research has substantially contributed to our understanding of the potential genetic and epigenetic mechanisms driving heterosis (Ryder et al. 2014). However, a full account of the precise mechanisms explaining all current heterosis observations is absent (Birchler et al. 2003; Birchler 2016). Furthermore, most discussions on the mechanism of heterosis only consider diploid plants, thus failing to account for the novel heterotic responses at the polyploid level (Washburn and Birchler 2014). In this experiment, we investigate autotetraploid *A. thaliana* for evidence of progressive heterosis, which up until now has never been addressed in the literature. It is demonstrated that approximately 40% of three-way cross F1 hybrids show progressive heterosis for leaf area. In addition, it is further demonstrated that no double-cross F1 hybrid (0%) shows progressive heterosis for leaf area. We conclude that increasing the multi-locus heterozygosity beyond the homozygous state can increase the observed heterosis, but only up to a threshold. These findings have implications for our understanding of the effects of nuclear heterozygosity and the design of F1 hybrid breeding programs, particularly polyploid breeding programs.

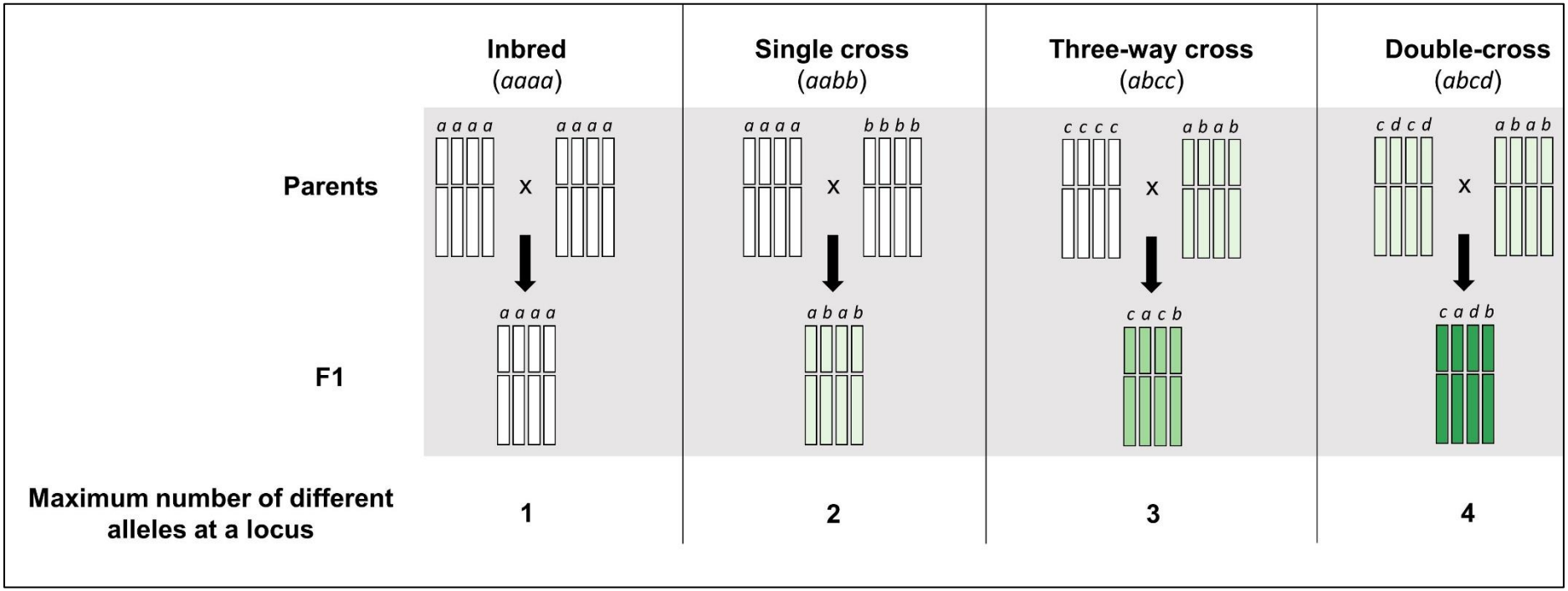


Figure 7. Crossing design to test for progressive heterosis in tetraploid plants. An autotetraploid contains four identical chromosome sets within its nucleus. Let us propose a hypothetical gene at a certain locus has four different alleles, denoted a, b, c, and d. Inbred (100% homozygous) lines will possess a mono-allelic locus for this gene, i.e. four identical copies of one allele (*aaaa* and *bbbb*). Following cross-fertilisation, the single cross F1 hybrid will possess a di-allelic locus for this gene, i.e. two copies of two different alleles (*aabb*). This single cross F1 hybrid can be crossed to a different inbred line (*cccc*) and the resulting three-way cross F1 hybrid will possess a tri-allelic locus for this gene, i.e. one copy of two different alleles and two identical copies of one allele (*abcc*). Alternatively, the single cross F1 hybrid can be crossed to a different single cross F1 hybrid (*ccdd*) and the resulting double-cross F1 hybrid will possess a tetra-allelic locus, i.e. one copy of four different alleles (*abcd*). The increasing intensity of the colour green reflects the increasing levels of heterozygosity.

Materials and Methods

Plant material

Four *Arabidopsis thaliana* accessions at the tetraploid (4x) level were the kind gift of Luca Comai (UC Davis, CA, USA). These were C24, Colombia-0 (Col), Landsberg *erecta*-0 (L*er*), and Zurich-0 (Zu). Accessions were colchicine-doubled diploid (2x) plants which have been maintained for several generations in our laboratory (Donoghue et al. 2014; Fort et al. 2016; Fort et al. 2017; Ryder et al. 2017).

Crossing design

For all crosses, four female plants were used. Five flowers on side branches were emasculated with tweezers under a dissection microscope. Two days later siliques were manually pollinated. Plants were grown in a growth chamber with 16 hours light at 22°C and 8 hours dark at 20°C. Seed was harvested three weeks following pollination.

Tetraploid plants were crossed to obtain inbred plants, single cross F1 hybrids, three-way cross F1 hybrids and double-cross F1 hybrids as follows. To obtain inbred plants, all inbred accessions were self-pollinated, e.g. C24 X C24, Col X Col, etc. To obtain single cross F1 hybrids, all inbred accessions were reciprocally crossed with one another, e.g. C24 X Col, Col X C24, etc. To obtain three-way cross F1 hybrids, eight different single cross F1 hybrids were crossed as a male parent to an unrelated inbred accession, e.g. L*er* X (C24 X Col), Zu X (C24 X Col), etc. To obtain double-cross F1 hybrids, different single cross F1 hybrids were crossed with an unrelated single cross F1 hybrid, e.g. (C24 X Col) X (L*er* X Zu), (C24 X Col) X (Zu X L*er*), etc (**Table *9***).

Phenotyping platform for high-throughput leaf area measurement

Seeds were stratified for two days in tap water at 4°C in the dark. 7 X 7 X 6.5cm pots (Modiform, Leusden, Netherlands) in were prepared with a soil mixture (5:1:1 peat:perlite:vermiculite with a final thin layer of peat soil on top). Four seeds of the same genotype were sown in a pot, evenly spaced apart. Pots were placed in growth chambers (Snijders Labs, Tilburg, Netherlands) with 16 hours light at 21°C and 8 hours dark at 18°C. Humidity was maintained with a single sheet of cling film for four days to facilitate germination. The position of genotypes within growth chambers was randomised. The experiment was conducted over two runs. The first run contained inbred lines and single cross

F1 hybrids. These plants were subsequently used for further crosses. The second run contained inbred lines, single cross F1 hybrids, three-way cross F1 hybrids, and double-cross F1 hybrids. The mean leaf area 10 days after sowing of inbred lines and single cross F1 hybrids was not significantly different between runs ($P = 0.66$; independent sample t-test; data not shown) and so it was determined that the data from both runs could be combined for analysis.

Four cameras (Autofocus USB camera with Ov5640 sensor, Ailipu Technology Co., Ltd, China) were positioned at the top of each growth chamber to capture images of plant growth at noon each day. Image correction and segmentation were performed with Plant Computer Vision software (Fahlgren et al. 2015).

Table 9. List of tetraploid *A. thaliana* crosses performed in this study.

Inbred	Single cross F1 hybrid	Three-way cross F1 hybrid	Double-cross F1 hybrid
C24 X C24	C24 X Col	C24 X (Col X Zu)	(C24 X Col) X (L*er* X Zu)
Col X Col	Col X C24	C24 X (Zu X Col)	(C24 X Col) X (Zu X L*er*)
L*er* X L*er*	C24 X L*er*	L*er* X (Col X Zu)	(Col X C24) X (L*er* X Zu)
Zu X Zu	L*er* X C24	L*er* X (Zu X Col)	(Col X C24) X (Zu X L*er*)
	C24 X Zu	L*er* X (C24 X Col)	(C24 X L*er*) X (Col X Zu)
	Zu X C24	L*er* X (Col X C24)	(C24 X L*er*) X (Zu X Col)
	Col X L*er*	Zu X (C24 X Col)	(L*er* X C24) X (Col X Zu)
	L*er* X Col	Zu X (Col X C24)	(L*er* X C24) X (Zu X Col)
	Col X Zu		(Col X L*er*) X (Zu X C24)
	Zu X Col		(Col X L*er*) X (C24 X Zu)
	L*er* X Zu		(L*er* X Col) X (Zu X C24)
	Zu X L*er*		(L*er* X Col) X (C24 X Zu)
			(Col X Zu) X (C24 X L*er*)
			(Col X Zu) X (L*er* X C24)
			(Zu X Col) X (C24 X L*er*)
			(Zu X Col) X (L*er* X C24)
			(C24 X Zu) X (Col X L*er*)
			(Zu X C24) X (Col X L*er*)
			(Zu X L*er*) X (C24 X Col)
			(L*er* X Zu) X (Col X C24)

Results

Single cross F1 hybrids exhibit heterosis for leaf area in ten out of twelve different combinations

The leaf area of twelve single cross F1 hybrids were examined for heterosis ten days after sowing. In six combinations heterosis over the best parent is observed: C24 X Col, Col X C24, C24 X L*er*, Zu X C24, Zu X Col, and Col X L*er*. In four combinations heterosis over the worst parent is observed: C24 X Zu, Col X Zu, L*er* X Col, and L*er* X Zu. One single cross F1 hybrid exhibits no heterosis: L*er* X C24. One single cross F1 hybrid exhibits below best-parent heterosis (i.e. negative heterosis): Zu X L*er* (**Figure *8***).

Three-way cross F1 hybrids exhibit progressive heterosis for leaf area in four out of eleven different combinations

The leaf area of eleven three-way cross F1 hybrids were examined for progressive heterosis ten days after sowing. In four combinations progressive heterosis is observed: L*er* X (Col X C24), Zu X (Col X C24), Zu X (Col X C24), and C24 X (Ler X Col) exhibit heterosis over the best parent (the single cross F1 hybrid). In four combinations there is no heterosis: L*er* X (Col X Zu), C24 X (Zu X Col), Zu X (Ler X Col), and Zu X (Col X Ler). In three combinations there is negative heterosis: L*er* X (C24 X Col), L*er* X (Zu X Col), and Zu X (C24 X Col) (**Figure *9***).

Double-cross F1 hybrids do not exhibit progressive heterosis for leaf area in any twenty different combinations

The leaf area of twenty double-cross F1 hybrids were examined for progressive heterosis ten days after sowing. No combinations exhibited progressive heterosis. In four combinations heterosis over the worst parent is observed: (C24 X Col) X (L*er* X Zu), (Col X Zu) X (L*er* X C24), (Zu X L*er*) X (C24 X Col) and (L*er* X Zu) X (Col X C24). In six combinations there is negative heterosis: (C24 X Col) X (Zu X L*er*), (Col X C24) X (Zu X L*er*), (Col X C24) X (Zu X L*er*), (C24 X L*er*) X (Col X Zu), (C24 X L*er*) X (Zu X Col), and (L*er* X Col) X (Zu X C24). In the remaining ten combinations there is no heterosis (**Figure *10***).

Greater levels of multi-locus nuclear heterozygosity are associated with increased leaf area up to a threshold

Varying levels of heterosis are observed across tetraploid F1 hybrids possessing different levels of multi-locus nuclear heterozygosity, with certain three-way cross F1 hybrids exhibiting evidence for progressive heterosis (**Figure *8***, **Figure *9***, and **Figure *10***). According to the progressive heterosis literature, maximising allelic diversity in tetraploids will maximise the observed heterosis. An analysis was performed to determine the increase in leaf area for each stepwise progression in heterozygosity relative to the homozygous state. The leaf area of each F1 hybrid as a proportion of its inbred lines was determined and F1 hybrids were grouped according to their locus heterozygosity for analysis (di-allelic, tri-allelic or tetra-allelic). The mean leaf area of tetraploid plants possessing a maximum of two different alleles at all loci is 1.35 times larger than homozygous tetraploid plants. The mean leaf area of tetraploid plants possessing a maximum of three different alleles at all loci is 1.5 times larger than homozygous tetraploid plants. The mean leaf area of tetraploid plants possessing a maximum of four different alleles at all loci is 1.3 times larger homozygous tetraploid plants. Thus, leaf area increases with increasing levels of multi-locus nuclear heterozygosity up to the tri-allelic level, whereupon the effect on leaf area increase levels off (*aaaa* < *aabb* < *abcc* > *abcd*) (**Figure *11***).

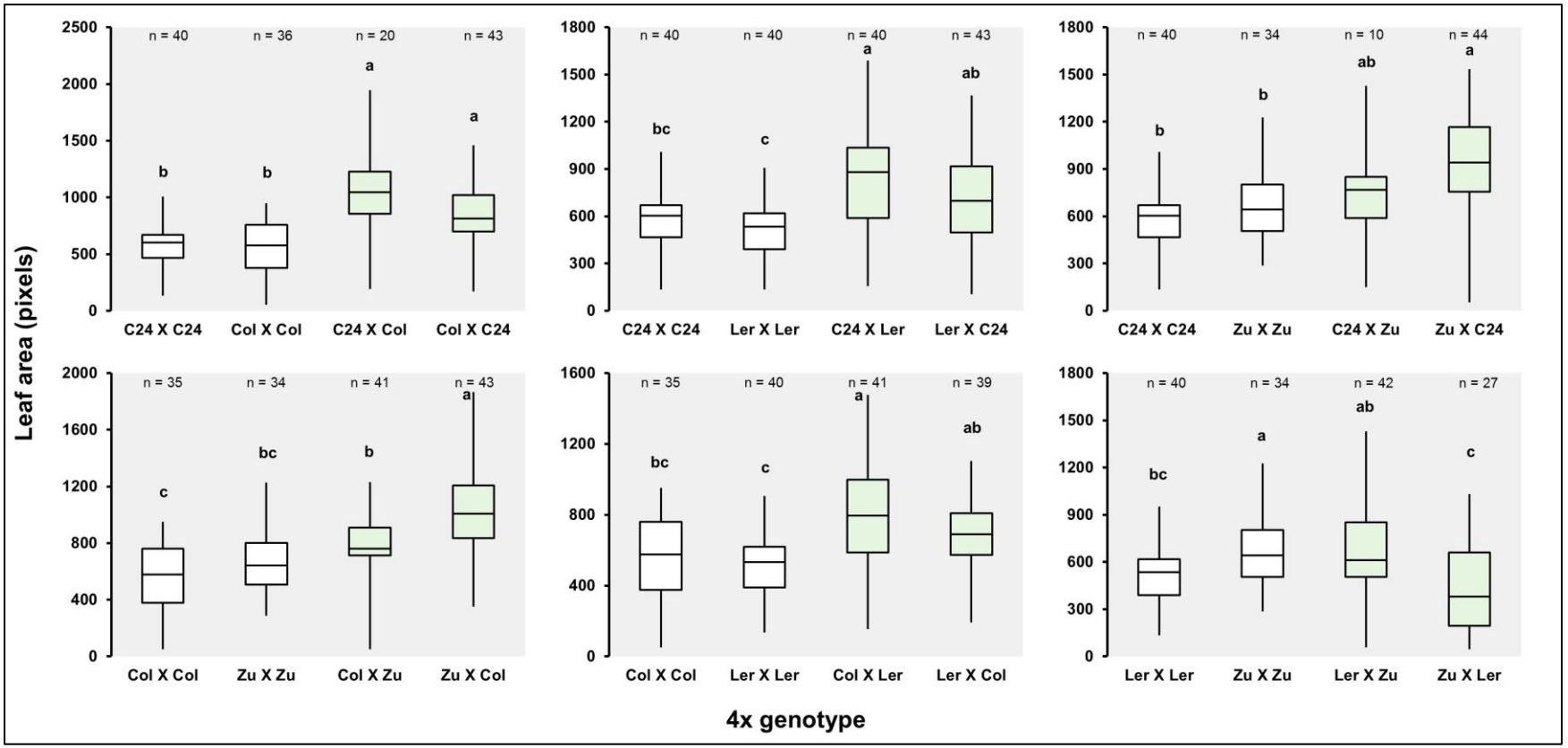


Figure 8. Box and whisker plot showing presence of leaf area heterosis in tetraploid single cross F1 hybrids. Leaf area was determined 10 days after sowing for parents (inbred lines) and F1 offspring (single cross F1 hybrid). A one-way ANOVA was performed to determine any statistically significant differences among genotypes. Those assigned the same letter are not significantly different from one another ($P > 0.05$) according to Tukey's HSD test. Sample size (n) indicated for each genotype.

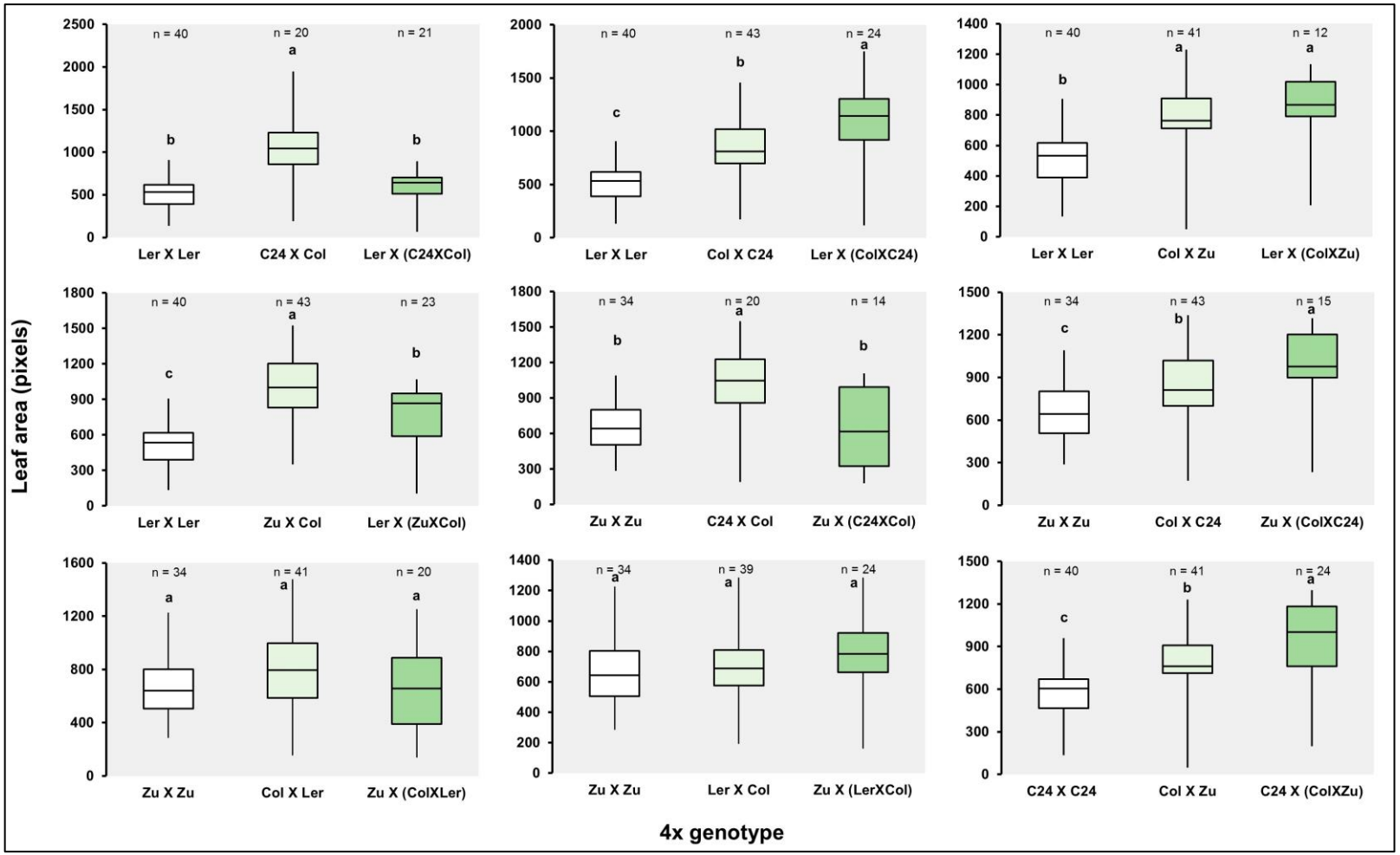


Figure 9. Box and whisker plot showing presence of leaf area progressive heterosis in tetraploid three-way cross F1 hybrids. Leaf area was determined 10 days after sowing for parents (inbred and single cross F1 hybrid) and F1 offspring (three-way cross F1 hybrid). A one-way ANOVA was performed to determine any statistically significant differences among genotypes. Those assigned the same letter are not significantly different from one another ($P > 0.05$) according to Tukey's HSD test. Sample size (n) indicated for each genotype.

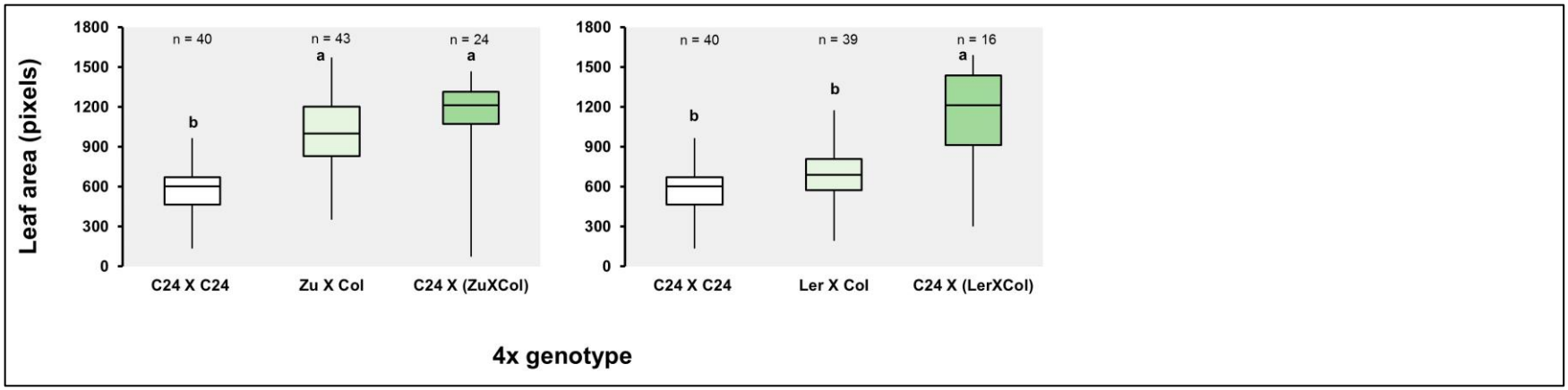


Figure 9 (continued).

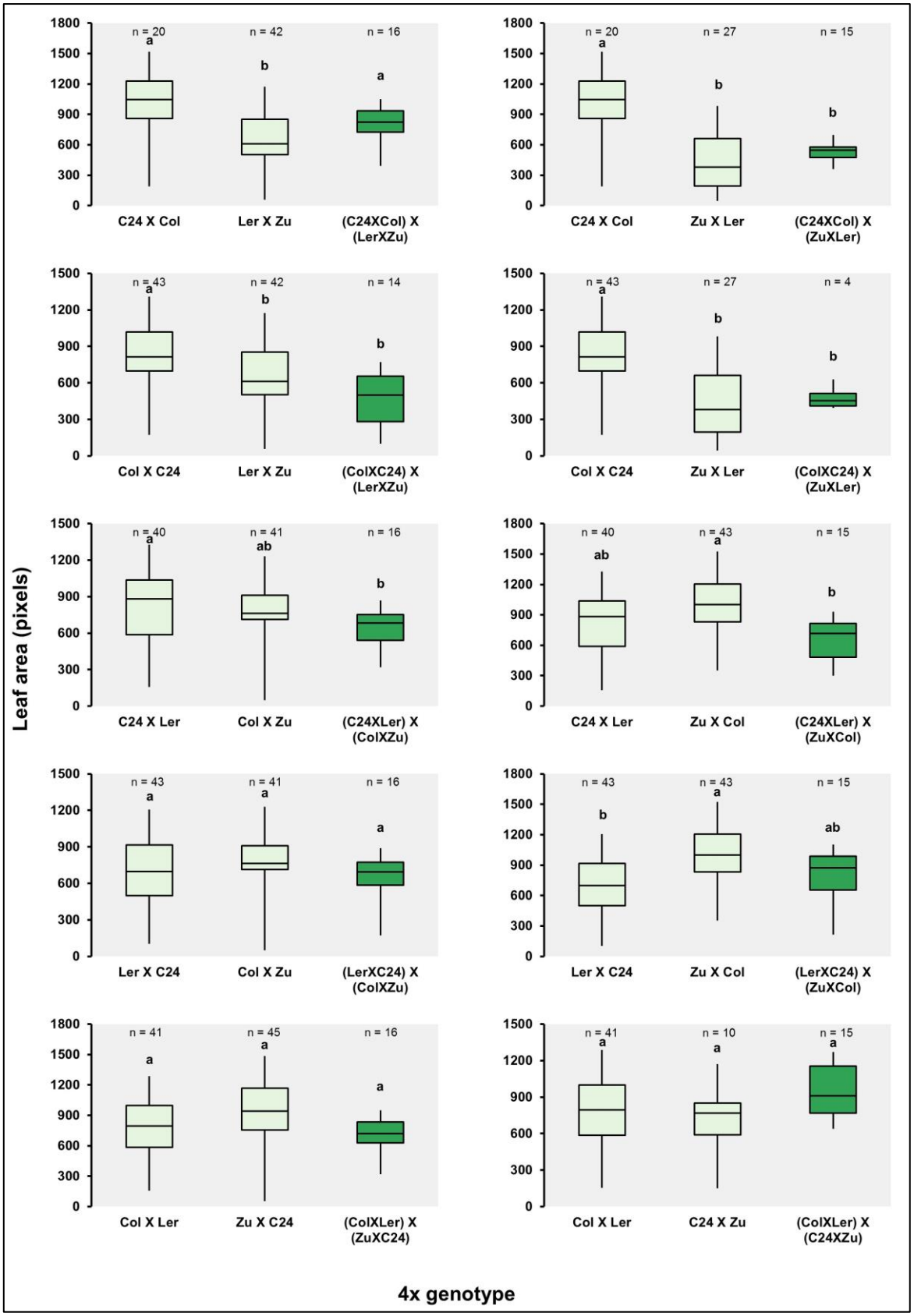


Figure 10. Box and whisker plot showing absence of leaf area progressive heterosis in tetraploid double-cross F1 hybrids. Leaf area was determined 10 days after sowing for parents (single cross F1 hybrids) and F1 offspring (double-cross F1 hybrid). A one-way

ANOVA was performed to determine any statistically significant differences among genotypes. Those assigned the same letter are not significantly different from one another ($P > 0.05$) according to Tukey's HSD test. Sample size (n) indicated for each genotype.

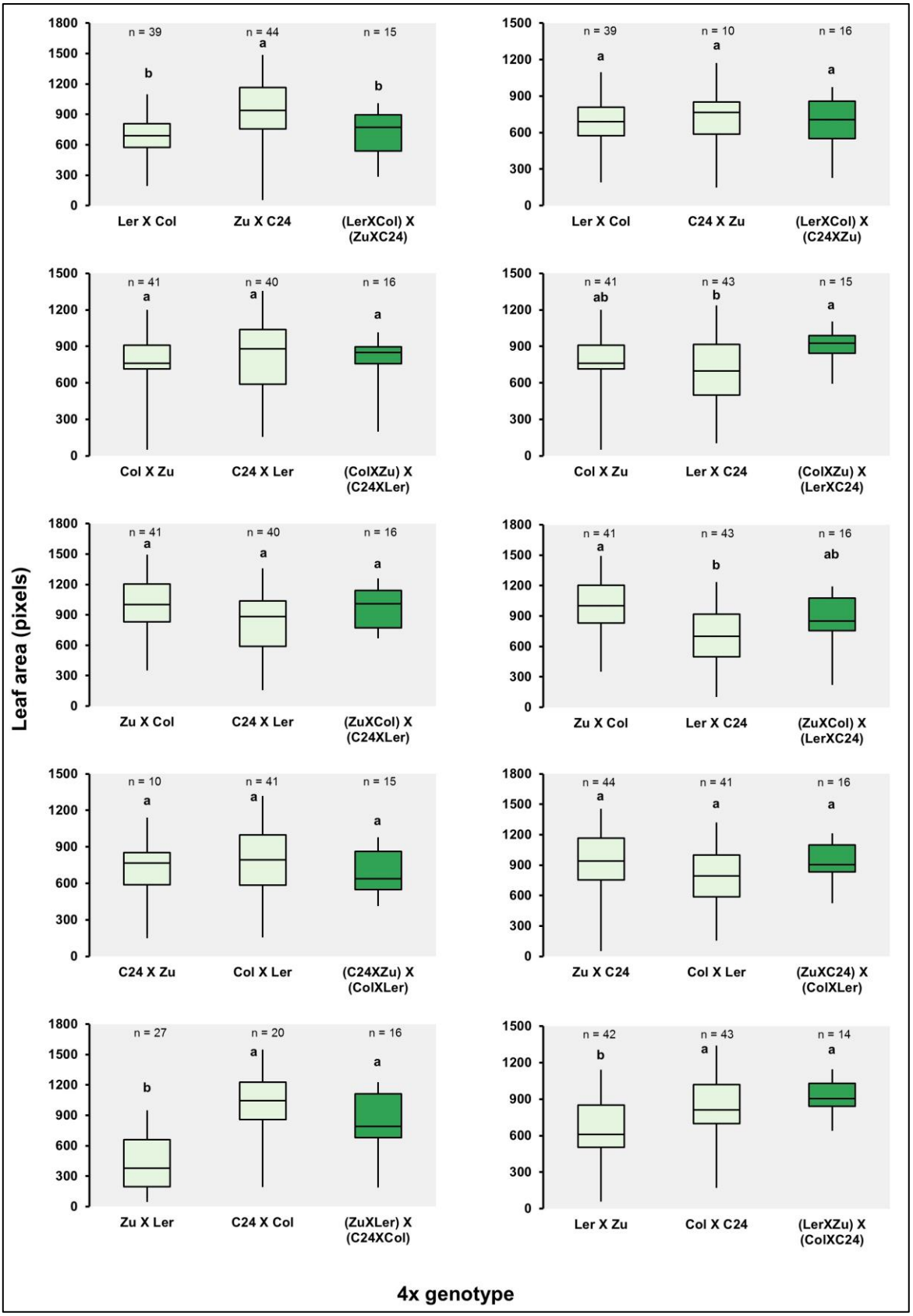


Figure 10 (continued).

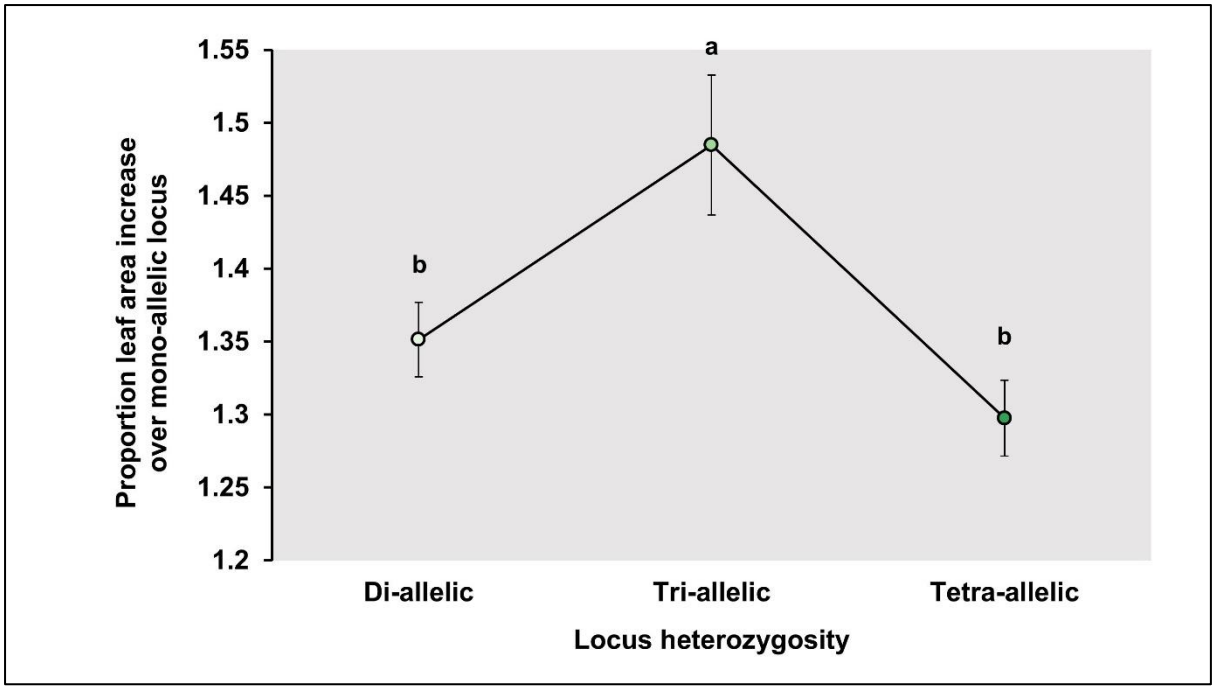


Figure 11. Line chart showing tetraploid leaf area increase over inbred lines with increasing levels of multi-locus nuclear heterozygosity. Leaf area of single cross F1 hybrids was calculated as a proportion of two corresponding inbred lines. Leaf area of three-way cross F1 hybrids was calculated as a proportion of three corresponding inbred lines. Leaf area of double-cross F1 hybrids was calculated as a proportion of all four corresponding inbred lines. F1 hybrids were grouped according to the maximum number of different alleles at a locus and mean values are displayed (± standard error). A one-way ANOVA was performed to determine any statistically significant differences among F1 hybrids. Those assigned the same letter are not significantly different from one another ($P > 0.05$) according to Tukey's HSD test.

Discussion

Combining diverse genomes in hybrid plants to increase heterosis

Double-cross autotetraploid F1 hybrids of potato, alfalfa and maize have exhibited heterosis over single cross F1 hybrids (Mok and Peloquin 1975; Groose et al. 1989a; Sockness and Dudley 1989; Bingham et al. 1994; Riddle and Birchler 2008). Such heterosis gains are believed to be not possible at the diploid level because the maximum number of alleles at any one locus is two, whereas a tetraploid can harbour up to four different alleles at the same locus. This suggests that greater allelic diversity can be captured at the polyploid level in comparison to the diploid level and this leads to a greater number of intra-allelic interactions. Indeed, the success of polyploid establishment and their contribution to plant speciation has been attributed to the potential for novel gene expression (as well as to their relatively large size and changes in reproductive behaviour) (Stebbins 1971; Ramsey and Schemske 1998; Van de Peer et al. 2017).

The findings from potato, alfalfa and maize suggest that polyploids can harness increasing levels of multi-locus nuclear heterozygosity but diploids cannot. If so, then by crossing genetically diverse parental lines it should be possible to enhance heterosis in polyploid plants in a single cross. Furthermore, crossing genetically diverse parental lines at the diploid level should have a minimal effect on heterosis. Interestingly, there is some evidence for this proposition in the literature. Heterosis in diploid maize, *A. thaliana*, rice, and pepper (*Capsicum annum*) is unrelated to the genetic distance between parents (Zhang et al. 1995; Benchimol et al. 2000; Geleta et al. 2004; Meyer et al. 2004; Stokes et al. 2007; Frisch et al. 2010; Reif et al. 2010). Heterosis in autopolyploid plants, however, is somewhat correlated with the genetic distance between parents. Using controlled crosses, it has been demonstrated that there is a small but significant association in autotetraploid rice between the parental genetic distance and the expression of heterosis (Tu et al. 2007; Wu et al. 2013). In contrast, significantly increasing multi-locus nuclear heterozygosity in autotriploid sugar beet (*Beta vulgaris*) does not increase the expression of heterosis (Chapter 2 of this thesis). The results of our experiment in autotetraploid *A. thaliana* demonstrate that heterosis effects can be enhanced with increasing levels of multi-locus nuclear heterozygosity, although a threshold is reached where no additional enhancement is evident. If a greater number of intra-allelic interactions are responsible for the progressive heterosis effect (Fehr 1991; Riddle and Birchler 2008), then the maximum beneficial intra-allelic interactions in *A. thaliana* are reached at the

tri-allelic level (e.g. *abcc*) and any additional heterozygosity (e.g. *abcd*) does not enhance heterosis (**Figure *11***).

The most extreme example of increasing heterozygosity in an F1 hybrid is seen in allopolyploid plants. Allopolyploids are polyploid plants with chromosome sets from different species, and thus represent the merger of genetically distinct genomes. Increasing the genetic distance among parents and the expression of heterosis has been shown to correlate in allopolyploid *Triticale* and oilseed rape (*Brassica napus*) (Ali et al. 1995; Riaz et al. 2001; Góral et al. 2005). Other studies in allopolyploid oilseed rape and bread wheat, however, demonstrate that the heterosis effect is not enhanced upon drastically increasing heterozygosity levels (Martin et al. 1995; Diers et al. 1996). Intriguingly, novel allopolyploid F1 hybrids which exhibit considerable heterosis have been created between *A. thaliana* and *A. arenosa*, between radish (*Raphanus sativus*) and cabbage (*Brassica oleracea*), as well as anecdotal evidence in tobacco (*Nicotiana*) (Pack 1927; East 1936; Ni et al. 2008)

Interpreting allopolyploid heterosis is less straightforward than autopolyploid heterosis, though, particularly among novel allopolyploid F1 hybrids. Allopolyploid plants are the hybrid of distinct species genomes which represent a far 'wider' cross than anything at that could be achieved at the autopolyploid level. Allopolyploid F1 hybrids exhibit a range of large-scale genome changes, such as stochastic gene expression, considerable epigenetic remodelling of homoeologous genes, instability of transposable elements, as well as genome-wide expression dominance of one parent over another (Wang et al. 2004; Madlung et al. 2005; Wang et al. 2006). Nonetheless, it is noteworthy that if no fatal intergenomic incompatibility exists, allopolyploids are capable of harnessing diverse genomes which can lead to heterosis.

Leaf area and the mechanisms of heterosis

Leaf growth in *Arabidopsis thaliana* is a sequential sequence of cell division and cell expansion. New leaves from the leaf primordium are initiated through rapid cell division and, as it grows, leaf development switches from a process of cell division to one of cell expansion (Donnelly et al. 1999; Kazama et al. 2010; Andriankaja et al. 2012; Claeys et al. 2012). Several independent molecular pathways responsible for final leaf size have been identified, with the key genetic regulators appearing to be those involved in cell division – the early stages of cell division are critical in determining final leaf size (Anastasiou et al. 2007; Gonzalez et al. 2010; Andriankaja et al. 2012). Presumably, a larger leaf will facilitate more light interception and hence provide the opportunity for increased photosynthesis. The carbon produced from

photosynthesis is partitioned and used for tissue growth as well as to maintain respiration (O'Leary 1988; Reich et al. 1998; Hanson et al. 2016). Thus, a noble goal in F1 hybrid breeding programs would be to discover what is driving the heterotic effect for larger leaves.

There is heterosis for leaf area in tetraploid single cross F1 hybrids of *A. thaliana* (**Figure *8***). In addition, there is further heterosis in certain combinations of three-way cross F1 hybrids, so called 'progressive heterosis' (**Figure *9***). The majority of the leaf area of ten-day old *A. thaliana* plants can be attributed to cell division rather than cell expansion (Pyke et al. 1991) and so the heterotic effect in operation is likely influencing cell division. Any proposed mechanisms of heterosis that seek to explain increased cell division of single cross F1 hybrids should also go a considerable way to explaining the increase in cell division of three-way cross F1 hybrids.

The proposed, nonexclusive genetic hypotheses for heterosis are dominance, over-dominance and epistasis. Two hypotheses (dominance, over-dominance) are concerned with interactions between alleles at the same locus whereas one of the hypotheses (epistasis) is concerned with interactions between alleles among two or more different loci. In addition to these proposed genetic mechanisms, epigenetic effects such as hypomethylation can contribute to heterosis (Dapp et al. 2015). Most studies of the heterosis phenomenon in plants address these various hypotheses, and what unites them is that the mechanisms driving heterosis can be species, accession, and trait specific. For example, the latest genomic sequencing investigations in large-scale hybrid rice and hybrid bread wheat populations have revealed that positive, dominance effects can explain the majority of yield heterosis in rice (although the candidate genes in question appear to be context dependent), but positive, dominance effects on yield heterosis in bread wheat are less influential than epistasis effects (Huang et al. 2016; Jiang et al. 2017). This highlights the elusive nature of heterosis and that old definitions can be a constraint to data interpretation. Quantitative traits (e.g. yield, leaf architecture) are affected by multiple genes making the identification of a single genomic region that can trigger heterosis a rare occurrence (Reymond et al. 2003; El-Lithy et al. 2004; Buckler et al. 2009; Mackay et al. 2009; McMullen et al. 2009; Birchler et al. 2010; Krieger et al. 2010).

Conclusions

The goal of this research project was to identify progressive heterosis in the model organism *A. thaliana*. The ubiquitous use of *A. thaliana* in the heterosis literature has aided considerably to our understanding of the mechanisms behind various heterosis effects (Meyer et al. 2004;

Stokes et al. 2007; Shen et al. 2012; Wang et al. 2015; van Hulten et al. 2018), however these experiments have all been performed at the diploid level. This research chapter is the first to report the presence of 'progressive heterosis' in *A. thaliana*. Some three-way cross combinations can further enhance the heterosis effect on leaf area. However, increasing heterozygosity in any double-cross combination consistently fails to enhance the heterosis effect. The positive finding in relation to three-way cross combinations has direct relevance to polyploid hybrid breeding programs, such as tetraploid oilseed rape (*Brassica napus*) and triploid sugar beet, which regularly use single-cross F1 hybrids in further crosses before reaching the desired commercial product (Möhring et al. 1999; Rahman 2005; Draycott 2008; Biancardi et al. 2010; Rahman 2013). The negative finding in relation to double-cross combinations has direct relevance to investigations of genetic incompatibility and outbreeding depression. Outbreeding depression, where upon certain allele combinations that positively contribute to plant fitness in one genetic background have a neutral or negative effect upon hybridisation, has been attributed to negative epistatic incompatibilities and is an important factor in plant speciation (Lynch 1991; Turelli and Orr 2000; Edmands 2002; Bomblies et al. 2007; Oakley et al. 2015).

Future directions

A manuscript has been written for submission to *Genetics*. The categorisation of progressive heterosis effects in *A. thaliana* has, so far, gone unreported in the literature and I believe the novel findings presented here in relation to multi-locus nuclear heterozygosity are appropriate for a high-quality genetics journal that regularly publishes on the topic of heterosis.

References

Ali, M., L.O. Copeland, S.G. Elias, and J.D. Kelly, 1995 Relationship between genetic distance and heterosis for yield and morphological traits in winter canola (Brassica napus L.). *Theoretical and Applied Genetics* 91 (1):118-121.

Anastasiou, E., S. Kenz, M. Gerstung, D. MacLean, J. Timmer *et al.*, 2007 Control of Plant Organ Size by KLUH/CYP78A5-Dependent Intercellular Signaling. *Developmental Cell* 13 (6):843-856.

Andriankaja, M., S. Dhondt, S. De Bodt, H. Vanhaeren, F. Coppens *et al.*, 2012 Exit from Proliferation during Leaf Development in Arabidopsis thaliana: A Not-So-Gradual Process. *Developmental Cell* 22 (1):64-78.

Benchimol, L.L., C.L. de Souza, A.A.F. Garcia, P.M.S. Kono, C.A. Mangolin *et al.*, 2000 Genetic diversity in tropical maize inbred lines: heterotic group assignment and hybrid performance determined by RFLP markers. *Plant Breeding* 119 (6):491-496.

Biancardi, E., J.M. McGrath, L.W. Panella, R.T. Lewellen, and P. Stevanato, 2010 Sugar beet, pp. 173-219 in *Root and tuber crops*. Springer.

Bingham, E.T., R.W. Groose, D.R. Woodfield, and K.K. Kidwell, 1994 Complementary Gene Interactions in Alfalfa are Greater in Autotetraploids than Diploids. *Crop Science* 34 (4):823-829.

Birchler, J.A., 2016 Hybrid vigour characterized. *Nature* 537:620.

Birchler, J.A., D.L. Auger, and N.C. Riddle, 2003 In Search of the Molecular Basis of Heterosis. *The Plant Cell* 15 (10):2236-2239.

Birchler, J.A., H. Yao, S. Chudalayandi, D. Vaiman, and R.A. Veitia, 2010 Heterosis. *The Plant Cell* 22 (7):2105-2112.

Bomblies, K., J. Lempe, P. Epple, N. Warthmann, C. Lanz *et al.*, 2007 Autoimmune Response as a Mechanism for a Dobzhansky-Muller-Type Incompatibility Syndrome in Plants. *PLOS Biology* 5 (9):e236.

Brancourt-Hulmel, M., G. Doussinault, C. Lecomte, P. Berard, B. Le Buanec *et al.*, 2003 Genetic improvement of agronomic traits of winter wheat cultivars released in France from 1946 to 1992. *Crop Science* 43 (1):37-45.

Buckler, E.S., J.B. Holland, P.J. Bradbury, C.B. Acharya, P.J. Brown *et al.*, 2009 The Genetic Architecture of Maize Flowering Time. *Science* 325 (5941):714-718.

Campbell, B., P. Chee, E. Lubbers, D. Bowman, W. Meredith *et al.*, 2012 Dissecting genotype× environment interactions and trait correlations present in the Pee Dee cotton germplasm collection following seventy years of plant breeding. *Crop Science* 52 (2):690-699.

Claeys, H., A. Skirycz, K. Maleux, and D. Inzé, 2012 DELLA Signaling Mediates Stress-Induced Cell Differentiation in Arabidopsis Leaves through Modulation of Anaphase-Promoting Complex/Cyclosome Activity. *Plant Physiology* 159 (2):739-747.

Dapp, M., J. Reinders, A. Bediee, C. Balsera, E. Bucher *et al.*, 2015 Heterosis and inbreeding depression of epigenetic Arabidopsis hybrids. *Nature plants* 1 (7):15092.

Diers, B., P. McVetty, and T. Osborn, 1996 Relationship between heterosis and genetic distance based on restriction fragment length polymorphism markers in oilseed rape (Brassica napus L.). *Crop Science* 36 (1):79-83.

Donnelly, P.M., D. Bonetta, H. Tsukaya, R.E. Dengler, and N.G. Dengler, 1999 Cell Cycling and Cell Enlargement in Developing Leaves of Arabidopsis. *Developmental Biology* 215 (2):407-419.

Donoghue, M.T.A., A. Fort, R. Clifton, X. Zhang, P.C. McKeown *et al.*, 2014 CmCGG Methylation-Independent Parent-of-Origin Effects on Genome-Wide Transcript Levels in Isogenic Reciprocal F1 Triploid Plants. *DNA research* 21 (2):141-151.

Draycott, A.P., 2008 *Sugar beet*: John Wiley & Sons.

Duvick, D.N., 2005 The Contribution of Breeding to Yield Advances in maize (Zea mays L.), pp. 83-145 in *Advances in agronomy*. Academic Press.

East, E.M., 1936 Heterosis. *Genetics* 21 (4):375.

Edmands, S., 2002 Does parental divergence predict reproductive compatibility? *Trends in Ecology & Evolution* 17 (11):520-527.

Egli, D.B., 2008 Comparison of Corn and Soybean Yields in the United States: Historical Trends and Future Prospects. *Agronomy journal* 100 (Supplement_3):S-79-S-88.

El-Lithy, M.E., E.J.M. Clerkx, G.J. Ruys, M. Koornneef, and D. Vreugdenhil, 2004 Quantitative Trait Locus Analysis of Growth-Related Traits in a New Arabidopsis Recombinant Inbred Population. *Plant Physiology* 135 (1):444-458.

Fahlgren, N., M. Feldman, Malia A. Gehan, Melinda S. Wilson, C. Shyu *et al.*, 2015 A Versatile Phenotyping System and Analytics Platform Reveals Diverse Temporal Responses to Water Availability in Setaria. *Mol Plant* 8 (10):1520-1535.

Fehr, W., 1991 *Principles of cultivar development: theory and technique*. IA, USA: Macmillian Publishing Company.

Fischer, R., 1996 Wheat Physiology at CIMMYT and Rainsing the Yield Plateau, pp. 191-202 in *Increasing Yield Potential in Wheat: Breaking the Barriers : Proceedings of a Workshop Held in Ciudad Obregón, Sonora, Mexico*, edited by M.P. Reynolds, S. Rajaram and A. McNab. CIMMYT.

Fort, A., P. Ryder, P.C. McKeown, C. Wijnen, M.G. Aarts *et al.*, 2016 Disaggregating polyploidy, parental genome dosage and hybridity contributions to heterosis in Arabidopsis thaliana. *New Phytologist* 209 (2):590-599.

Fort, A., R. Tuteja, M. Braud, P.C. McKeown, and C. Spillane, 2017 Parental-genome dosage effects on the transcriptome of F1 hybrid triploid embryos of Arabidopsis thaliana. *The Plant Journal* 92 (6):1044-1058.

Frisch, M., A. Thiemann, J. Fu, T.A. Schrag, S. Scholten *et al.*, 2010 Transcriptome-based distance measures for grouping of germplasm and prediction of hybrid performance in maize. *Theoretical and Applied Genetics* 120 (2):441-450.

Geleta, L., M. Labuschagne, and C. Viljoen, 2004 Relationship between heterosis and genetic distance based on morphological traits and AFLP markers in pepper. *Plant Breeding* 123 (5):467-473.

Gonzalez, N., S. De Bodt, R. Sulpice, Y. Jikumaru, E. Chae *et al.*, 2010 Increased Leaf Size: Different Means to an End. *Plant Physiology* 153 (3):1261.

Góral, H., M. Tyrka, and L. Spiss, 2005 Assessing genetic variation to predict the breeding value of winter triticale cultivars and lines. *Journal of Applied Genetics* 46 (2):125-131.

Groose, R., L. Talbert, W. Kojis, and E. Bingham, 1989a Progressive heterosis in autotetraploid alfalfa: studies using two types of inbreds. *Crop Science* 29 (5):1173-1177.

Groose, R.W., L.E. Talbert, W.P. Kojis, and E.T. Bingham, 1989b Progressive heterosis in autotetraploid alfalfa: studies using two types of inbreds. *Crop Science* 29 (5):1173-1177.

Hanson, D.T., S.S. Stutz, and J.S. Boyer, 2016 Why small fluxes matter: the case and approaches for improving measurements of photosynthesis and (photo)respiration. *Journal of Experimental Botany* 67 (10):3027-3039.

Hayes, H.K., R.F. Immer, and D.C. Smith, 1955 *Methods of plant breeding*. CA, USA: McGraw-Hill.

Huang, X., S. Yang, J. Gong, Q. Zhao, Q. Feng *et al.*, 2016 Genomic architecture of heterosis for yield traits in rice. *Nature* 537:629.

Jiang, Y., R.H. Schmidt, Y. Zhao, and J.C. Reif, 2017 A quantitative genetic framework highlights the role of epistatic effects for grain-yield heterosis in bread wheat. *Nature Genetics* 49:1741.

Kazama, T., Y. Ichihashi, S. Murata, and H. Tsukaya, 2010 The Mechanism of Cell Cycle Arrest Front Progression Explained by a KLUH/CYP78A5-dependent Mobile Growth Factor in Developing Leaves of Arabidopsis thaliana. *Plant and Cell Physiology* 51 (6):1046-1054.

Krieger, U., Z.B. Lippman, and D. Zamir, 2010 The flowering gene SINGLE FLOWER TRUSS drives heterosis for yield in tomato. *Nature Genetics* 42:459.

Lobell, D., and C. Field, 2007 Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental Research Letters* 2 (1):014002.

Loel, J., C. Kenter, B. Märländer, and C.M. Hoffmann, 2014 Assessment of breeding progress in sugar beet by testing old and new varieties under greenhouse and field conditions. *European Journal of Agronomy* 52, Part B:146-156.

Lynch, M., 1991 THE GENETIC INTERPRETATION OF INBREEDING DEPRESSION AND OUTBREEDING DEPRESSION. *Evolution* 45 (3):622-629.

Mackay, T.F.C., E.A. Stone, and J.F. Ayroles, 2009 The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics* 10:565.

Madlung, A., A.P. Tyagi, B. Watson, H. Jiang, T. Kagochi *et al.*, 2005 Genomic changes in synthetic Arabidopsis polyploids. *The Plant Journal* 41 (2):221-230.

Martin, J., L. Talbert, S. Lanning, and N. Blake, 1995 Hybrid performance in wheat as related to parental diversity. *Crop Science* 35 (1):104-108.

McMullen, M.D., S. Kresovich, H.S. Villeda, P. Bradbury, H. Li *et al.*, 2009 Genetic Properties of the Maize Nested Association Mapping Population. *Science* 325 (5941):737-740.

Meyer, R.C., O. Törjék, M. Becher, and T. Altmann, 2004 Heterosis of Biomass Production in Arabidopsis. Establishment during Early Development. *Plant Physiology* 134 (4):1813-1823.

Miller, F.R., and Y. Kebede, 1984 Genetic Contributions to Yield Gains in Sorghum, 1950 to 19801, pp. 1-14 in *Genetic Contributions to Yield Gains of Five Major Crop Plants*, edited by W.R. Fehr. Crop Science Society of America and American Society of Agronomy, Madison, WI.

Miller, M., C. Zhang, and Z.J. Chen, 2012 Ploidy and hybridity effects on growth vigor and gene expression in Arabidopsis thaliana hybrids and their parents. *G3: Genes, Genomes, Genetics* 2 (4):505-513.

Möhring, S., E. Esch, and G. Wricke, 1999 Breeding hybrid varieties in winter rapeseed using recessive self-incompatibility, pp. 228-229 in *Proceedings of the 10th International Rapeseed Congress, Canberra, Australia*.

Mok, D.W., and S.J. Peloquin, 1975 Breeding value of 2n pollen (diplandroids) in tetraploid x diploid crosses in potatoes. *Theor Appl Genet* 46 (6):307-314.

Ni, Z., E.-D. Kim, M. Ha, E. Lackey, J. Liu *et al.*, 2008 Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature* 457:327.

O'Leary, M.H., 1988 Carbon Isotopes in Photosynthesis. *BioScience* 38 (5):328-336.

Oakley, C.G., J. Ågren, and D.W. Schemske, 2015 Heterosis and outbreeding depression in crosses between natural populations of Arabidopsis thaliana. *Heredity* 115:73.

Pack, D.A., 1927 Ring density of sugar beets as a character for selection. *American Journal of Botany*:238-245.

Pyke, K.A., J.L. Marrison, and A.M. Leech, 1991 Temporal and Spatial Development of the Cells of the Expanding First Leaf of Arabidopsis thaliana (L.) Heynh. *Journal of Experimental Botany* 42 (11):1407-1416.

Rahman, H., 2013 *Review: Breeding spring canola (Brassica napus L.) by the use of exotic germplasm*: SPIE.

Rahman, M.H., 2005 Resynthesis of Brassica napus L. for self-incompatibility: self-incompatibility reaction, inheritance and breeding potential. *Plant Breeding* 124 (1):13-19.

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

Reich, P.B., M.B. Walters, M.G. Tjoelker, D. Vanderklein, and C. Buschena, 1998 Photosynthesis and respiration rates depend on leaf and root morphology and nitrogen concentration in nine boreal tree species differing in relative growth rate. *Functional Ecology* 12 (3):395-405.

Reif, J.C., S. Fischer, T.A. Schrag, K.R. Lamkey, D. Klein *et al.*, 2010 Broadening the genetic base of European maize heterotic pools with US Cornbelt germplasm using field and molecular marker data. *Theoretical and Applied Genetics* 120 (2):301-310.

Reymond, M., B. Muller, A. Leonardi, A. Charcosset, and F. Tardieu, 2003 Combining Quantitative Trait Loci Analysis and an Ecophysiological Model to Analyze the Genetic Variability of the Responses of Maize Leaf Growth to Temperature and Water Deficit. *Plant Physiology* 131 (2):664-675.

Riaz, A., G. Li, Z. Quresh, M.S. Swati, and C.F. Quiros, 2001 Genetic diversity of oilseed Brassica napus inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. *Plant Breeding* 120 (5):411-415.

Riddle, N.C., and J.A. Birchler, 2008 Comparative analysis of inbred and hybrid maize at the diploid and tetraploid levels. *Theoretical and Applied Genetics* 116 (4):563-576.

Russell, W., 1991 Genetic improvement of maize yields, pp. 245-298 in *Advances in agronomy*, edited by D. Sparks. Academic Press, CA, USA.

Ryder, P., M. McHale, A. Fort, and C. Spillane, 2017 Generation of stable nulliplex autopolyploid lines of Arabidopsis thaliana using CRISPR/Cas9 genome editing. *Plant Cell Reports* 36 (6):1005-1008.

Ryder, P., P.C. McKeown, A. Fort, and C. Spillane, 2014 Epigenetics and heterosis in crop plants, pp. 13-31 in *Epigenetics in Plants of Agronomic Importance: Fundamentals and Applications*. Springer.

Shen, H., H. He, J. Li, W. Chen, X. Wang *et al.*, 2012 Genome-wide analysis of DNA methylation and gene expression changes in two Arabidopsis ecotypes and their reciprocal hybrids. *The Plant Cell* 24 (3):875-892.

Shull, G.H., 1948 WHAT IS "HETEROSIS"? *Genetics* 33 (5):439-446.

Sockness, B.A., and J.W. Dudley, 1989 Performance of Single and Double Cross Autotetraploid Maize Hybrids with Different Levels of Inbreeding. *Crop Science* 29 (4):875-879.

Stebbins, G.L., 1971 Chromosomal evolution in higher plants. *Chromosomal evolution in higher plants.*

Stokes, D., C. Morgan, C. O'Neill, and I. Bancroft, 2007 Evaluating the utility of Arabidopsis thaliana as a model for understanding heterosis in hybrid crops. *Euphytica* 156 (1-2):157-171.

Tester, M., and P. Langridge, 2010 Breeding Technologies to Increase Crop Production in a Changing World. *Science* 327 (5967):818-822.

Tu, S., L. Luan, Y. Liu, W. Long, F. Kong *et al.*, 2007 Production and Heterosis Analysis of Rice Autotetraploid Hybrids. *Crop Science* 47 (6):2356-2363.

Turelli, M., and H.A. Orr, 2000 Dominance, Epistasis and the Genetics of Postzygotic Isolation. *Genetics* 154 (4):1663-1679.

Van de Peer, Y., E. Mizrachi, and K. Marchal, 2017 The evolutionary significance of polyploidy. *Nature Reviews Genetics* 18 (7):411.

van Hulten, M.H.A., M.-J. Paulo, W. Kruijer, H. Blankestijn-De Vries, B. Kemperman *et al.*, 2018 Assessment of heterosis in two Arabidopsis thaliana common-reference mapping populations. *PloS one* 13 (10):e0205564.

Wang, J., L. Tian, H.-S. Lee, N.E. Wei, H. Jiang *et al.*, 2006 Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. *Genetics* 172 (1):507-517.

Wang, J., L. Tian, A. Madlung, H.-S. Lee, M. Chen *et al.*, 2004 Stochastic and Epigenetic Changes of Gene Expression in Arabidopsis Polyploids. *Genetics* 167 (4):1961-1973.

Wang, L., I.K. Greaves, M. Groszmann, L.M. Wu, E.S. Dennis *et al.*, 2015 Hybrid mimics and hybrid vigor in Arabidopsis. *Proceedings of the National Academy of Sciences* 112 (35):E4959-E4967.

Washburn, J.D., and J.A. Birchler, 2014 Polyploids as a "model system" for the study of heterosis. *Plant Reprod* 27 (1):1-5.

Wu, J.-W., C.-Y. Hu, M.Q. Shahid, H.-B. Guo, Y.-X. Zeng *et al.*, 2013 Analysis on genetic diversification and heterosis in autotetraploid rice. *SpringerPlus* 2 (1):439.

Yao, H., A. Dogra Gray, D.L. Auger, and J.A. Birchler, 2013 Genomic dosage effects on heterosis in triploid maize. *Proceedings of the National Academy of Sciences* 110 (7):2665-2669.

Zhang, Q., Y. Gao, M.S. Maroof, S. Yang, and J. Li, 1995 Molecular divergence and hybrid performance in rice. *Molecular breeding* 1 (2):133-142.

Chapter 4. Investigating hybrid nucleus and cytoplasm effects on haploid induction through centromere-mediated chromosome elimination in *A. thaliana*

Note: cross to generate "HI Zu/Col^{Zu}" was performed by Dr Antoine Fort

Introduction

Haploid plants are those that contain a gametic chromosome number. Haploid plants with a single copy of their chromosome – denoted 1x – are sterile due to the absence of chromosome pairing during meiosis. However, these plants are useful to plant breeders when the chromosome number can be doubled to create a 'double haploid' (DH), consequently restoring plant fertility. DHs are homozygous at all loci and thus can function as (1) a rapid method for production of homozygous lines instead of undertaking several generations of self-pollination, and, (2) a way to develop inbred lines where it would otherwise not be possible, such as in self-incompatible species, dioecious species or species that are vulnerable to severe inbreeding depression upon self-pollination (Murovec and Bohanec 2012).

Genome elimination for haploid induction in crops has been achieved *in vivo* through wide species crosses, i.e. interspecies crosses (Ishii et al. 2016). A landmark paper in 2010 described a new haploid induction approach through genetic engineering of the centromeric region in *Arabidopsis thaliana* (Ravi and Chan 2010). The centromere is a chromosomal position that contains tandem repeats of short DNA sequences which serve as a binding site for kinetochore proteins (Henikoff et al. 2001). Microtubules bind to the kinetochore complex during both mitosis and meiosis (**Figure *12***). The establishment of this kinetochore complex is dependent on the centromeric histone H3 variant 'CENH3'. Ravi and Chan (2010) found that a *cenh3*-/- null mutant complemented with an altered version of *CENH3* led to partial or complete uniparental genome elimination (aneuploid and haploid offspring) when crossed with a WT plant. A green fluorescent protein fused to the N-terminal tail domain of a H3 variant, replacing the N-terminal tail of CENH3, improved the frequency of uniparental genome elimination. It was noted that haploid induction efficiency was higher when the mutant was used as a female parent than as a male parent.

Creating haploids *in vivo* by manipulating the centromeric histone CENH3 in *Arabidopsis thaliana* has great relevance for haploid induction efforts in commercial crops because the histone-fold domain (HFD) of CENH3 is highly conserved among plant species, while the N-terminal tail region is hypervariable (Malik and Henikoff 2003; Ravi et al. 2010; Kuppu et al. 2015; Britt and Kuppu 2016). It has been demonstrated that a single point mutation in the HFD of CENH3 is sufficient to impair centromere function (in this case, CENH3 assembly) in *A. thaliana*, barley (*Hordeum vulgare*) and sugar beet, and the *A. thaliana* mutant in this instance (*Atcenh3* L130F-1) can induce haploids upon crossing with WT pollen (Karimi-Ashtiyani et al. 2015). In a similar but independent experiment it was reported that a single

point mutation in the HFD, in four out of six instances, was sufficient to generate haploid inducers in *A. thaliana* (Kuppu et al. 2015). Furthermore, the N-terminal tail of CENH3 in *A. thaliana* can be replaced with the N-terminal tail from *Lepidium oleraceum* causing no obvious impacts on self-fertilisation but triggers uniparental genome elimination upon cross-fertilisation with WT pollen (Maheshwari et al. 2015). Beyond the model organism Arabidopsis, the first example of manipulating the centromeric region for chromosome elimination in a crop species was shown in maize (Kelliher et al. 2016). An RNA-interference (RNAi) strategy and gene knockout strategy were tested to impair native *CENH3*. Like the pioneering work in Arabidopsis (Ravi and Chan 2010), the authors constructed transgenes with an altered N-terminal tail of CENH3 plus the inclusion of a green fluorescent protein. Both hemizygous and homozygous *AcGREEN-tailswap-CENH3* mutants were the best-performing haploid inducers, albeit at very low frequencies (from 1 – 4%).

The precise mechanisms behind centromere-mediated chromosome elimination are not fully understood. Valuable clues are available from the classic haploid induction method in barley, where interspecific crosses between cultivated barley (*H. vulgare*) and a wild-relative (*H. bulbosum*) frequently leads to elimination of *H. bulbosum* chromosomes (Kasha and Kao 1970). In this instance zygotic mitosis can be observed in viable and in unviable *H. vulgare X H. bulbosum* crosses. This has revealed that *H. bulbosum CENH3* genes are not silenced following fertilisation, but rather are present in the embryo and are even transcribed. However, during embryogenesis *H. bulbosum* centromeres show delayed loading of CENH3, and in a genome elimination event this leads to a failure to capture microtubules (which is necessary for mitosis) causing missegregation. Hence, the *H. bulbosum* genome is lost/degraded (Sanei et al. 2011). Why the lagging *H. bulbosum* chromosomes cannot load CENH3 is unclear.

An innovative use for the *GFP-tailswap cenh3* mutant in *A. thaliana* is the ability to create plants with the nucleus from a single parent and with the cytoplasm from a different parent. For maximum efficiency this first requires transferring the *GFP-tailswap cenh3* mutation from its current genotype (Columbia-0, thereafter Col) to a new genotype with a new cytoplasm. The genetically engineered CENH3 is now present in a hybrid nucleus and a completely novel cytoplasm. This offers a framework to study centromere-mediated chromosome elimination in a new context to determine any cytonuclear interactions. In this experiment, the *GFP-tailswap cenh3* mutation was transferred to a new genotype (Zurich-0, thereafter Zu) and crosses with WT were performed with the expectation that this would lead to uniparental genome elimination. The objective was to determine if centromere-mediated

chromosome elimination in *Arabidopsis thaliana* is influenced by the presence of a hybrid nucleus and a different cytoplasm.

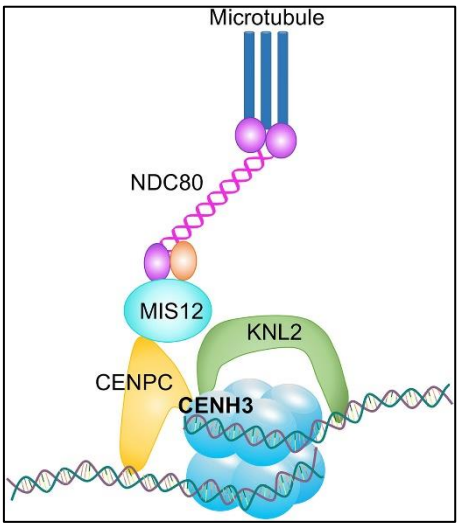


Figure 12. The kinetochore complex in plants. At plant centromeres the centromeric specific histone H3 variant 'CENH3' assembles several proteins to form the kinetochore complex. The kinetochore captures microtubules in both mitosis and meiosis. Taken from Wang & Dawe (2018).

Materials & Methods

Mutant lines

The haploid inducer (HI) mutant, named *GFP-tailswap* elsewhere (Ravi and Chan 2010; Ravi et al. 2014) but HI Col^{Col} for this experiment, is a *cenh3-1* null mutant complemented with a green-fluorescent protein fused to the tail of a H3 variant, inserted as a transgene replacing the N-terminal tail of CENH3. Plants heterozygous for the *cenh3-1* mutation were the kind gift of Luca Comai (UC Davis, CA, USA).

High-throughput DNA extraction method

A fast, high-throughput method of plant DNA extraction was performed for genotyping. This encompasses a homemade extraction buffer of salt water supplemented with sodium dodecyl sulphate (SDS) and handling samples in a 96-well plate normally used for PCR. For detailed description including the primers used on extracted DNA see Supplementary Materials and Methods.

Generating haploid inducer lines with new cytoplasm

The HI mutant is maintained as heterozygous for the *cenh3-1* mutation. This produces ample pollen and can be used as a male parent to move the *GFP-tailswap cenh3* mutation to another genotype, as described previously (Ravi et al. 2014). The original HI line is in a Col-0 background. 10 seeds of heterozygous HI Col^{Col} and 10 seeds of WT Zu were stratified for 3 days in the dark at 4°C. These were sown to 7 X 7 X 6.5cm pots (Modiform, Leusden, Netherlands) in a soil mix consisting of peat, perlite, and vermiculite in a ratio of 5:1:1. Plants were grown in a plant growth chamber (Snijders Scientific, Tilburg, Netherlands) with 16/8hr day/night at 21/18°C. HI plants were genotyped for the transgene and mutation (Ravi and Chan 2010). WT Zu inflorescences were manually emasculated with a Dumont no. 5 tweezer (Electron Microscopy Sciences, PA, USA) and were pollinated with the heterozygous HI Col^{Col}. The F1 seed from this cross ("heterozygous HI F1 Zu/Col^{Zu}") was harvested 3 weeks later and grown under the same conditions. All plants from this cross are heterozygous for the *GFP-tailswap* transgene and were genotyped to find plants heterozygous for the *cenh3-1* mutation. These plants were allowed to self and resulting seed was genotyped to find plants homozygous for the *GFP-tailswap* transgene and heterozygous for the *cenh3-1* mutation ("heterozygous HI Zu/Col^{Zu}"). Thus, a HI line in a new cytoplasm is available and can be maintained in a heterozygous state until use as a homozygote ("HI Zu/Col^{Zu}") (**Figure *13***). The new HI line was maintained as a heterozygote for three generations before selecting for the homozygous condition.

Generating haploids

100 seeds each of the heterozygous HI Col^{Col} and HI Zu/Col^{Zu} plants were stratified as before and sown four seeds to a pot. 25 seeds of WT Col and WT Zu were stratified, also, and sown one seed per pot. Plants were grown in the same growth chamber as before. When plants reached growth stage 1.10 (Boyes et al. 2001) the true HI plants (homozygous for *cenh3-1* mutation) became evident due to their curly leaf rosettes (Ravi and Chan 2010). Plants were genotyped to confirm presence of transgene and mutation and heterozygous HI plants were manually removed from the soil and thrown away.

At the on-set of flowering, the absence of self-fertilisation in the HI lines is evident (**Figure *14***). Three HI lines were selected and three siliques on a lateral branch were pollinated with either WT Col pollen or WT Zu pollen. The F1 seed from this cross was harvested 3 weeks later and stored for one week in a seed storage chamber (Weiss Technik, Schunk Group, Heuchelheim, Germany) at 12°C. All seed was sterilised with 70% methanol, then a seed sterilization solution consisting of 5% sodium hypochlorite solution (NaClO) with 0.01%

Triton X-100, followed by 5 times wash with sterile, distilled water. Seeds were stratified as before and sown to sterile media *in vitro*, as recommended (Ravi and Bondada 2016), in a plant growth chamber (CLF Plant Climatics, Emersacker, Germany) with 16/8hr day/night at 22°C/18°C. The media contained ½ strength MS media, with 0.5% sucrose, 0.8% agar and pH adjusted to 5.8 with KOH dropwise. Seed shape was determined visually, with WT seed as a reference. Plump seeds were scored as normal-shaped and dark coloured, wrinkled seed scored as misshapen. It has been noted that F1 seeds from HI *X* WT crosses may take up to two weeks to germinate, and these late-germinating seeds are often haploids (Ravi and Bondada 2016). After 10 days, the first germination count was recorded and plants at growth stage 1.02 were removed from agar plates inside the laminar hood and transferred to soil. The plate was resealed and returned to the growth chamber. After 1 additional week, the plates were examined again and no further germination was seen. Pots with seedlings were enclosed with a clear plastic cover to maintain humidity for 5 days. When plants had at least 5 rosette leaves (growth stage 1.05) they were genotyped for the presence of the HI transgene and WT histone tail. Samples showing amplification of just the WT band were scored as '1x', amplification of just the transgene as 'Selfed', and amplification of both as '2x or aneuploid'.

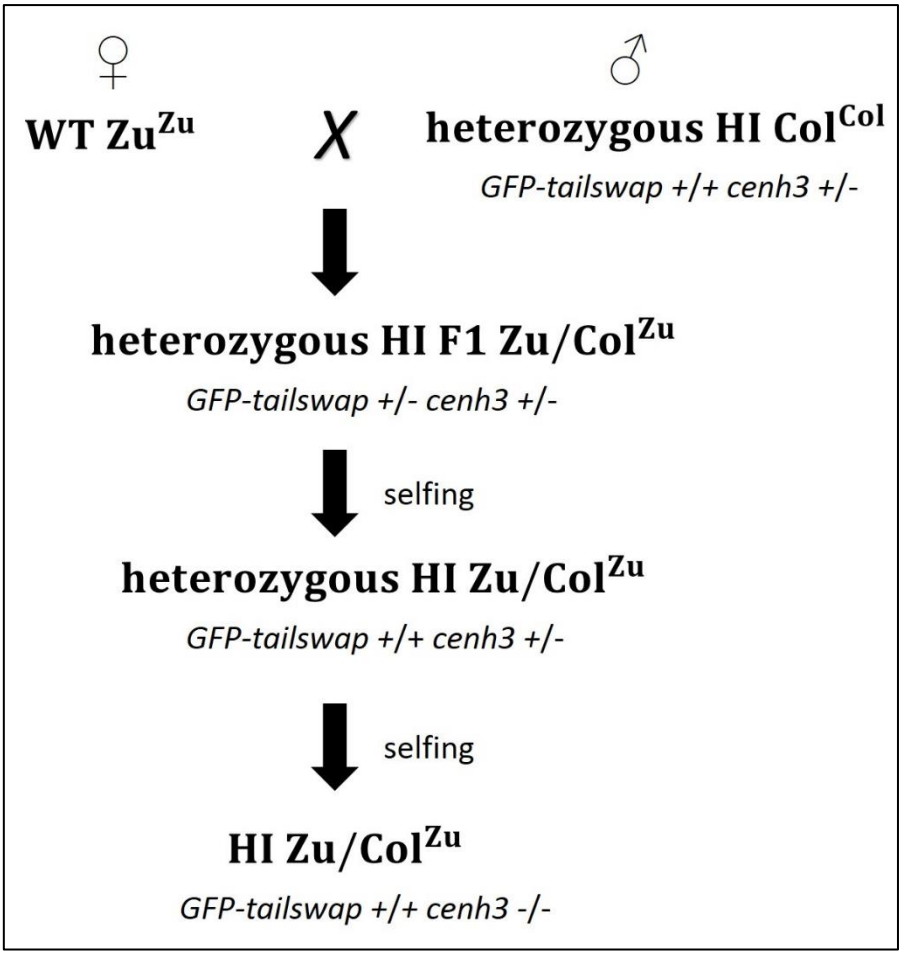


Figure 13. Generating HI Zu/Col^{Zu}. A new HI line with a hybrid nucleus and a different cytoplasm is created by crossing the heterozygous HI Col^{Col} as a male parent and selecting correctly through the generations. Superscript text indicates the cytoplasm.

Figure 14. WT plants and equivalent HI mutant plants. HI Col^{Col} and HI Zu/Col^{Zu} have curly leaf rosettes and exhibit reduced self-fertilisation compared with WT plants as identified through unpollinated pistils.

Results

HI Zu/Col^{Zu} displays high male sterility

A standout phenotype in the original HI line is the differential effect on male and female meiosis. When left to grow, HI Col^{Col} exhibits very low self-fertilisation. This is not attributed to female gametogenesis, but rather caused by almost complete male sterility (Ravi and Chan 2010; Ravi et al. 2011; Ravi et al. 2014). Likewise, when left to grow the new HI Zu/Col^{Zu} also exhibits very low self-fertilisation (**Figure *14***). Upon out-crossing, siliques contain viable seed (**Table 10**). Thus, the modified centromere in the new context of a hybrid nucleus and Zu cytoplasm is also associated with low self-fertilisation due to high male sterility.

F1 seeds from HI *X* WT crosses are variable in shape with low germination

Following crosses with WT Col and Zu pollen, both HI lines yield approx. 20-30 seeds per silique. F1 seeds from HI *X* WT crosses are variable in shape, with normal shaped (plump) and misshapen (wrinkled and dark) seeds both present (Figure 15;Table 10). Regardless of shape, all F1 seeds display low germination. HI Col^{Col} *X* WT crosses germinate at ≤ 30% while HI Zu/Col^{Zu} *X* WT crosses germinate at ≤ 55%. This left a final sample size of 44 F1 plants from crosses to HI Col^{Col} and 99 F1 plants from crosses to HI Zu/Col^{Zu} (**Table *10***).

Haploid plants can be identified by a single-step genotyping method

There are different genotyping options to determine if an F1 plant is the product of a genome elimination event (**Figure *16***). Plants can be genotyped for the presence of the WT histone tail and the modified transgene histone tail – F1 plants expressing only the WT histone tail from the male parent are the product of maternal genome elimination event during zygotic mitosis (**Figure *16***, sample #100). F1 plants expressing both are the product of fertilisation (**Figure *16***, sample #116). Also, plants can be genotyped for the presence of a gene (*AT4G01310*) encoding a ribosomal LP5 family protein. This gene is present in both Col-0 and Zu genomes and gives a slightly larger product from Zu. Overnight digestion with SacI will cut the Col-0 sequence while a single nucleotide polymorphism (SNP) at this locus in the Zu sequence will not cut. F1 plants expressing the gene with a digested fragment are the product of a fertilisation event (**Figure *16***, sample #116). However, the restriction digestion genotyping method is only suitable for HI Col^{Col} *X* WT Zu crosses. Therefore, amplification of the WT histone tail and the transgene *tailswap* histone tail is a faster and more reliable method of genotyping.

Maternal genome elimination with isogenic nucleus and Col-0 cytoplasm is higher than with hybrid nucleus and Zu cytoplasm

All F1 plants were genotyped for the presence of both the WT histone tail and the transgene *tailswap* histone tail (Supplementary Figure 1). The frequency of genome elimination differs between the HI lines, independent of male parent in the cross. Both HI Col^{Col} *X* WT Col and HI Col^{Col} *X* WT Zu crosses preferentially produce paternal haploid plants, on average 68%. In contrast, both HI Zu/Col^{Zu} *X* WT Col and HI Zu/Col^{Zu} *X* WT Zu preferentially produce hybrid plants, on average 80% with only 11% paternal haploid plants (**Figure *17***). The haploid induction by HI Col^{Col} is significantly greater than HI Zu/Col^{Zu} ($P < 0.05$ according to Tukey's test).

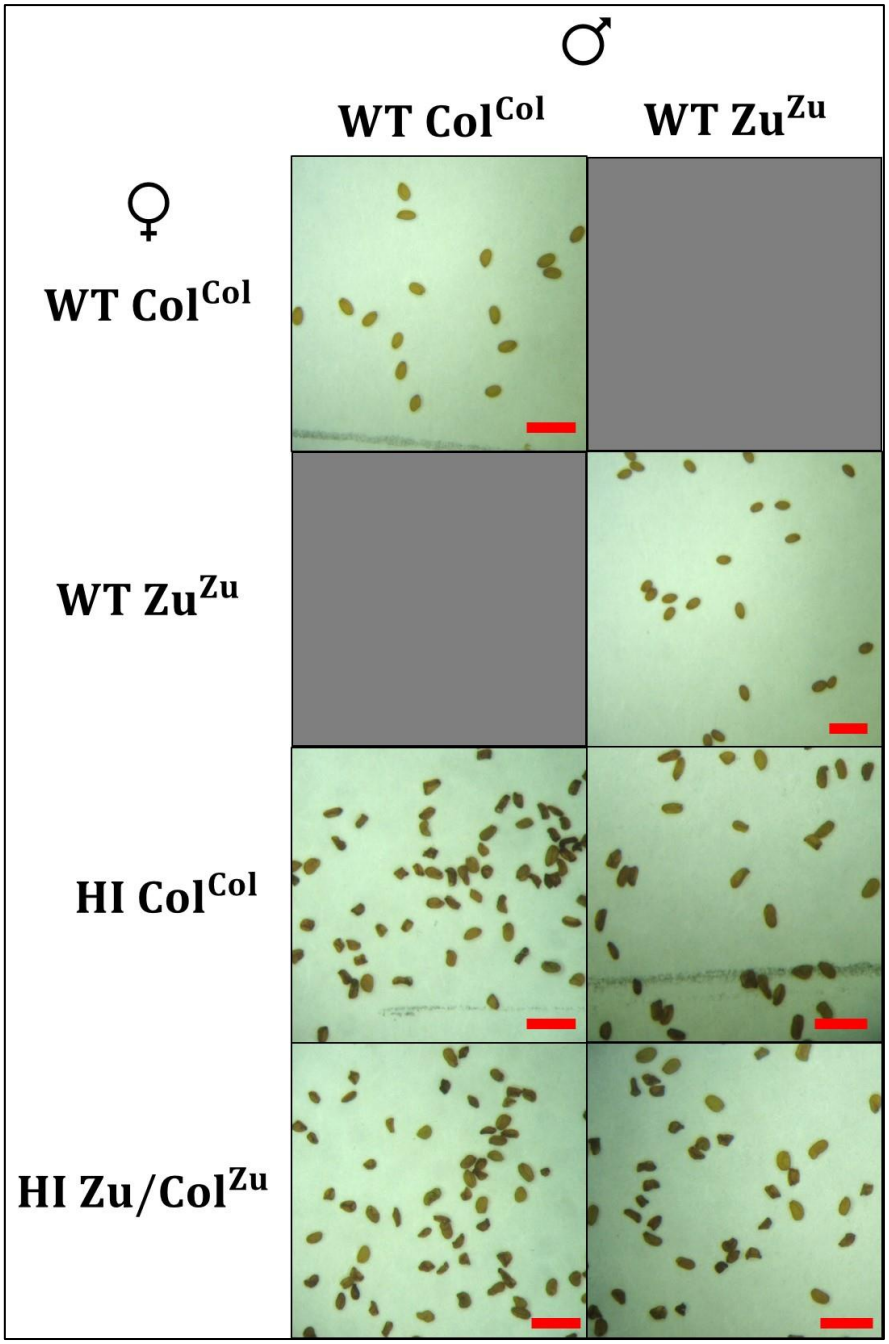


Figure 15. F1 seeds from HI X WT crosses are variable in shape. Both HI lines produce a mixture of normal shaped (plump) and misshapen (wrinkled) seeds following pollination with WT. Isogenic, control crosses produce only normal shaped seed. Red line = 1mm.

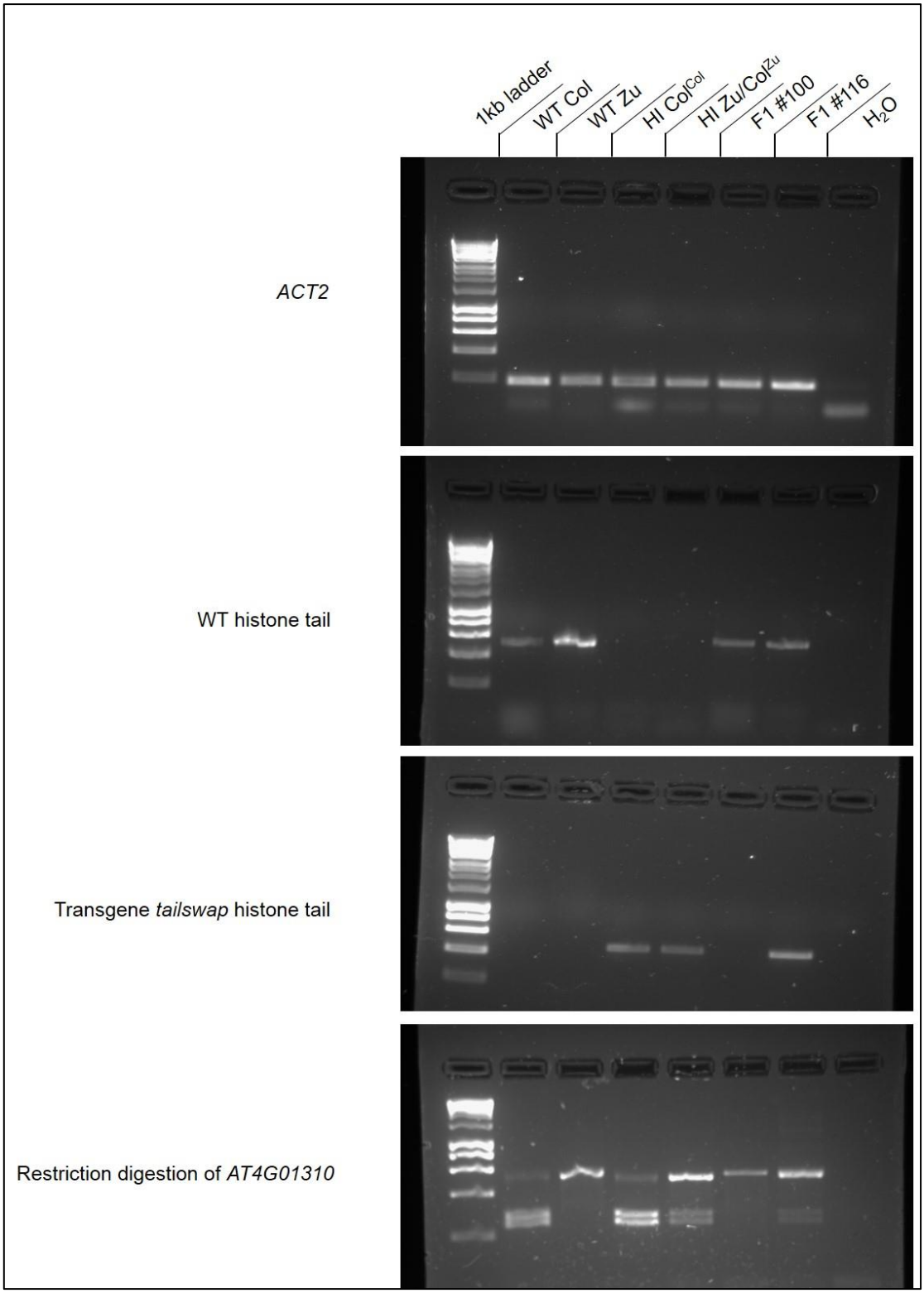


Figure 16. Identification of maternal genome elimination in F1 plants through genotyping. Amplification of *ACT2* was performed as a PCR control. F1 plants #100 and #116 are the product of a HI Col^{Col} X WT Zu cross. Samples expressing only the WT histone tail and no cut of AT4G01310 product (#100) possess only the Zu nuclear genome and thus identified as haploids. Samples expressing both WT histone tail and the transgene tailswap histone tail as well as a cut fragment of AT4G01310 product (#116) possess both the Col and Zu genome, and thus identified as diploid or anueploid.

Table 10. F1 seed shape, seed germination and plant survival following transfer to soil of HI X WT crosses. Three biological replicates of HI mutants were used as female parents in each cross, indicated by subscript numbers.

Cross	Normal-shaped seeds	Germinated		Misshapen seeds	Germinated		Viable seedlings	
♀ ♂	#	#	%	#	#	%	#	%
WT Col *X* WT Col	165	162	98.18	0	n/a	n/a	10	100*
WT Zu *X* WT Zu	174	172	98.85	0	n/a	n/a	10	100*
HI Col^{Col} *X* WT Col $_1$	30	4	13.33	40	3	7.5	7	100
$_2$	23	4	17.39	30	3	10	7	100
$_3$	26	6	23.08	28	1	3.57	5	71.43
HI Col^{Col} *X* WT Zu $_1$	64	14	21.88	8	1	12.5	5	33.33
$_2$	32	12	37.5	26	1	3.85	9	69.23
$_3$	33	10	30.3	19	3	15.79	11	84.62
HI Zu/Col^{Zu} *X* WT Zu $_1$	36	20	55.56	14	3	21.43	23	100
$_2$	42	22	52.38	14	1	7.14	22	95.65
$_3$	43	24	55.81	28	3	10.71	25	92.59
HI Zu/Col^{Zu} *X* WT Col $_1$	33	13	39.39	22	2	9.09	7	46.67
$_2$	59	14	23.73	26	1	3.85	6	40
$_3$	51	12	23.53	36	4	11.11	16	100

* Of the several hundred WT seedlings, only 10 were transferred to soil.

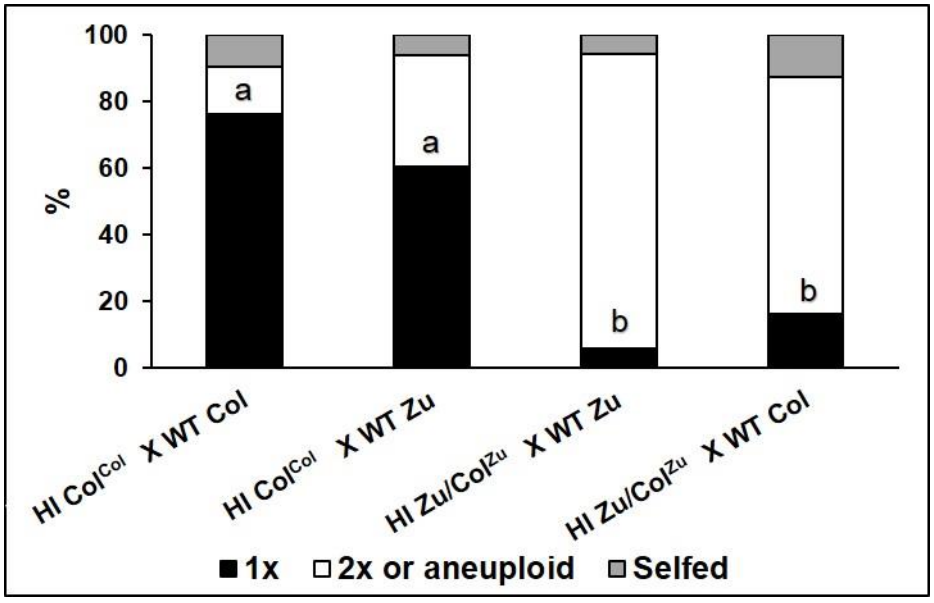


Figure 17. Maternal genome elimination is higher with an isogenic nucleus and Col-0 cytoplasm than with a hybrid nucleus and Zu cytoplasm. HI Col^{Col} produces significantly more paternal haploids than HI Zu/Col^{Zu}, as determined with one-way ANOVA and Tukey's test. Means assigned different letters are significantly different ($P < 0.05$).

Discussion

Variation in the frequency of haploid induction has been observed following switches to the N-terminal tail (Maheshwari et al. 2015) and following single point mutations in the HFD of CENH3 in *A. thaliana* (Karimi-Ashtiyani et al. 2015; Kuppu et al. 2015). However, my result is the first to demonstrate differences in the frequency of centromere-mediated chromosome elimination between identical, modified centromeric histone H3 variants. The *GFP-tailswap cenh3* are present in different maternal environments: one is in a Col-0 nucleus with a Col-0 cytoplasm while the other is in a recombined Zu/Col-0 nucleus with a Zu cytoplasm.

Large frequency of hybrid genomes from HI Zu/Col^{Zu} *X* WT crosses suggests that zygotic mitosis has been partially rescued

Male gametogenesis is considerably impaired in the haploid inducer mutants. Both HI Col^{Col} and HI Zu/Col^{Zu} exhibit reduced self-fertilisation in comparison to WT (**Figure *14***) but produce seed upon out-crossing (Table 10). Low pollen production in HI Col^{Col} has been attributed to incorrect loading of CENH3 to the centromere in metaphase I. This causes a depletion of the essential kinetochore protein MIS12 (Figure 12) (Ravi et al. 2011). Through visually assessing the extent of male sterility (i.e. without immunostaining) (**Figure *14***) it appears HI Zu/Col^{Zu} also suffers from depletion of CENH3 and faulty kinetochore assembly

during male meiosis. Intriguingly, zygotic mitosis in the HI mutants differs. The haploid inducer genome from HI Col^{Col} is more frequently eliminated than HI Zu/Col^{Zu} upon pollination with WT (Figure 17). Differences in the frequency of uniparental genome elimination could be explained by numerous factors.

Kinetochore assembly-dependent factors: It has been proposed that the modified *GFP-tailswap cenh3* centromere malfunctions when it must "compete" with WT centromere for certain factors necessary for kinetochore assembly, and thus the *GFP-tailswap cenh3* centromere missegregates in a genome elimination event during zygotic mitosis (Ravi and Chan 2010; Dwivedi et al. 2015). Ravi et al. (2011) propose that CENH3 loading in meiosis and mitosis are regulated differently. HI Zu/Col^{Zu} may be more efficient at CENH3 loading, and hence kinetochore assembly, in zygotic mitosis than HI Col^{Col}. This enables HI Zu/Col^{Zu} centromeres to compete to a reasonable extent with WT centromeres during zygotic mitosis, as evident by the high percentage of hybrid genomes in the F1 generation (**Figure *17***).

How can centromeres lose the race to assemble a kinetochore? Recently, a simple centromere size model has been proposed to account for differences in CENH3 loading (Wang and Dawe 2018). The authors cite 24 interspecies genome elimination examples which they believe can be attributed to centromere size dimorphism: the species with the smaller centromere (using chromosome size as a proxy) loses out. This suggests that HI mutants support a smaller-than-usual centromere with a lower loading capacity; and that larger centromeres can support more CENH3. This may explain why HI mutants do not induce haploids upon self-fertilisation. If it is true that HI Zu/Col^{Zu} can load more CENH3 than HI Col^{Col} onto the centromere, this may simply be a function of centromere size.

The centromere size model is compatible with the intriguing observation that centromeric DNA sequences are highly variable and appear to be rapidly evolving both within and between species – suggesting centromere stability is more influenced by epigenetic factors rather than a strict DNA sequence (Musacchio and Desai 2017). Also, centromere position and neocentromeres (new centromeres forming in a completely novel genome region) can be triggered by genetic hybridity (Topp et al. 2009; Birchler et al. 2011) which could further influence centromere stability and hence kinetochore assembly.

Kinetochore assembly-independent factors: It has been demonstrated *in vitro* that the bipolar spindle, which is composed of the microtubules and associated proteins, can assemble and orientate around DNA completely independent of the kinetochore complex (Heald et al. 1996).

Microtubules are moved and organised by a number of motor proteins (Bannigan et al. 2007; Yamada and Goshima 2017). Thus, microtubule movement and bipolar spindle orientation required for mitosis may be stronger in one context (e.g. HI Zu/Col^{Zu}) than in another (e.g. HI Col^{Col}) and capable of compensating for the weak kinetochores in question.

Genome elimination *in vivo* potentially unrelated to centromeres

As well as genetic engineering of *CENH3*, there have been other genes and QTL implicated in uniparental genome elimination *in vivo*. In maize breeding, the haploid inducer is traditionally the male parent, and a knockout mutation in a pollen-specific gene (*MATRILINEAL*) on chromosome 1 (Gilles et al. 2017; Kelliher et al. 2017; Liu et al. 2017) and a QTL on chromosome 9 (Liu et al. 2015) have been linked to paternal genome elimination. Knocking out an ortholog of *MATRILINEAL* in rice (*Oryza sativa*) *OsMATL* has recently been demonstrated to lead to haploid induction (Yao et al. 2018). The precise molecular mechanisms causing uniparental genome elimination in this instance are not resolved, as maternal haploid induction in maize through a knocking out *MATRILINEAL* is associated with lower pollen viability (Li et al. 2017) but this is not true for rice (Yao et al. 2018).

Conclusion

The widespread use of efficient haploid inducers in commercial breeding can facilitate the fixation of desirable traits in DH lines as well contribute to genome sequencing efforts in highly heterozygous crops (Murovec and Bohanec 2012). A promising development is the potential for centromere-mediated chromosome elimination to be adopted in a range of crop species (Henikoff et al. 2001; Ravi et al. 2010; Kuppu et al. 2015; Britt and Kuppu 2016; Kelliher et al. 2016). However, while first identified in the 1970s in barley (Kasha and Kao 1970) and redefined 8 years ago in mutant *Arabidopsis thaliana* (Ravi and Chan 2010), the precise mechanisms behind uniparental genome elimination remain to be fully explained. To maximise the use of haploid inducers in commercial breeding it will be necessary to better understand why uniparental genome elimination succeeds or fails in different instances. I have demonstrated that uniparental genome elimination caused by identical, modified centromeric histone H3 variants is affected by the maternal environment in which it resides. Whether the maternal influence in this instance is nuclear, cytoplasmic, or a combination of both remains to be fully elucidated (Sloan et al. 2018).

Future Directions

Generating additional haploid inducer lines with different Col-0 and Zu contributions to disaggregate cytonuclear influences

To determine a mechanism for differential uniparental genome elimination frequencies between HI Col^{Col} and HI Zu/Col^{Zu}, it is necessary to create additional HI lines. These can disaggregate whether the effect on haploid induction is related to a Zu/Col-0 hybrid nucleus or related to a Zu cytoplasm. Unfortunately, an identical HI line in the accession Zurich is not available ("HI Zu^{Zu}"). However, it is possible through hybridisation to create two additional HI lines, one maintaining the Col-0 cytoplasm with a hybrid nucleus ("HI Col/Zu^{Col}") and one maintaining the Zu/Col-0 hybrid nucleus with a Col-0 cytoplasm ("HI Zu/Col^{Col}") (**Figure *18***).

The new HI Col/Zu^{Col} is equivalent to HI Col^{Col}, the only difference being a recombined nucleus. If the frequency of haploid induction changes it can be attributed to the hybrid nucleus effect. The new HI Zu/Col^{Col} is equivalent to HI Zu/Col^{Zu}, the only difference being the hybrid nucleus genome proportion has slightly changed (from 50:50 to approx. 33:66 ratio) and the cytoplasm has been swapped. If the frequency of haploid induction changes it can be attributed to the cytoplasm effect.

The generation of these additional HI lines is currently underway. Following crosses to WT and analysing the results it will be necessary to differentiate between normal hybridisation (diploid offspring) and abnormal hybridisation (aneuploid offspring) in the genetic hybrids via flow cytometry. Finally, to investigate a potential mechanism at play, it is possible to validate the simple centromere size model (Wang and Dawe 2018) by chromatin immunoprecipitation with sequencing (ChIP-seq) of all HI and WT lines used in the crosses. To investigate mechanisms independent of kinetochore assembly will require microtubule movement and motor protein expression analysis.

Using different haploid inducer lines to determine genotype specificity

While generating new HI lines with altered Col-0 and Zu contributions can disaggregate cytonuclear influences on uniparental genome elimination, this may only be relevant in a single genotype context. To determine a nuclear or cytoplasm genotype specificity it will be necessary to assess the haploid induction when the modified centromeric histone H3 variant is present in other Arabidopsis accessions. The SpillaneLab also possess HI lines with a hybrid nucleus from C24 with a C24 cytoplasm (HI $C24/Col^{C24}$) and a hybrid nucleus from L*er* with a L*er*

cytoplasm (HI L*er*/Col^{L*er*}). By crossing these different HI lines with WT plants, it will be possible to determine if there is a general hybrid nucleus effect on haploid induction efficiency. Of note, there is a recent preprint on bioRxiv which investigates the interaction between the nuclear genome and cytoplasm across seven *A. thaliana* accessions (Flood et al. 2018). The authors generated different HI lines, although the accession Zurich was not one of them and there is no detail regarding which genotypes are more likely to be part of a successful genome elimination event.

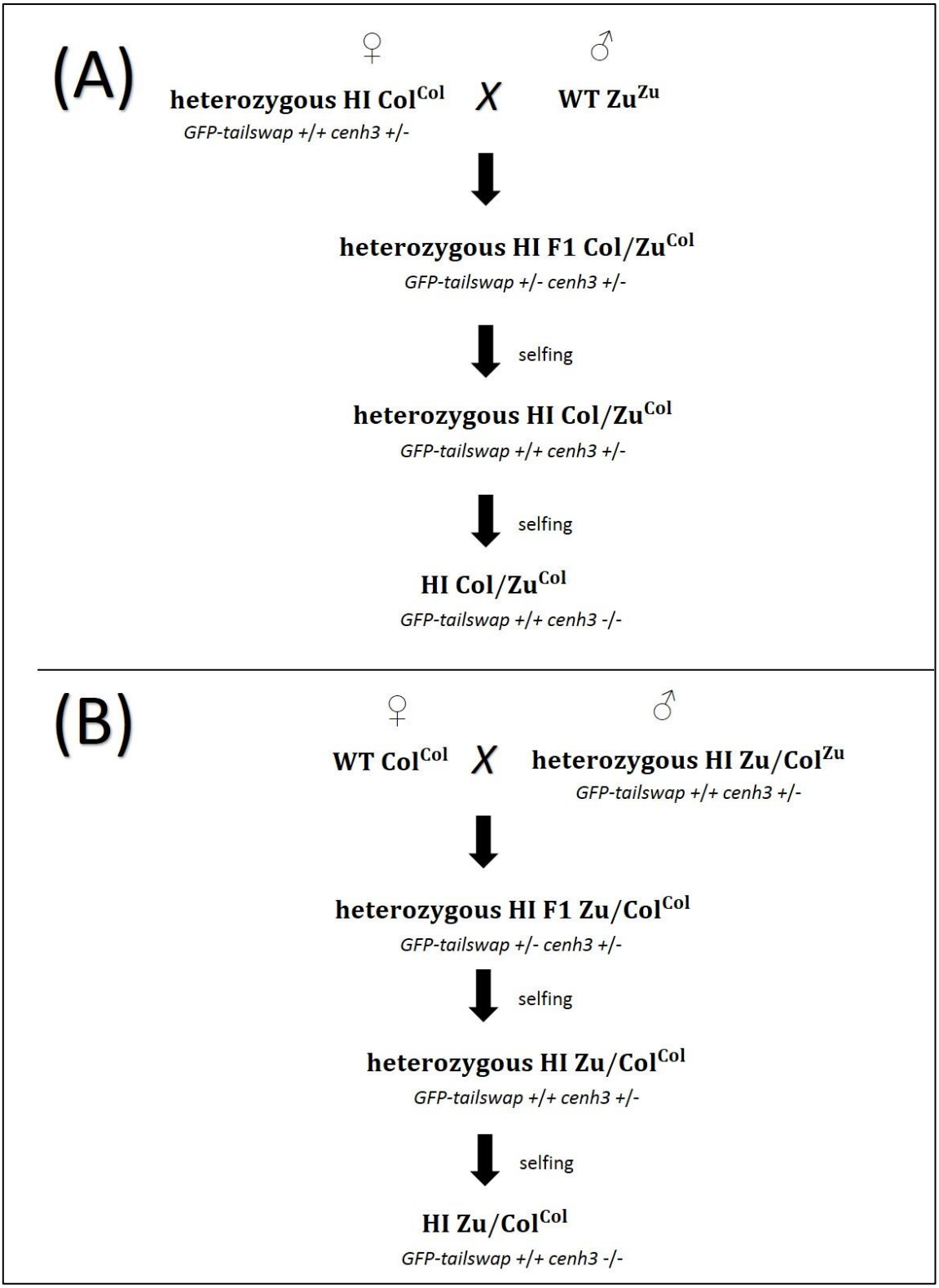


Figure 18. Generating HI Col/Zu^{Zu} and HI Zu/Col^{Col}. (**A**) A new HI line based on HI Col^{Col} with a different nucleus is created by crossing WT Zu as a male parent to the heterozygous HI Col^{Col} as a female parent and selecting correctly through the generations. (**B**) A new HI line based on HI Zu/Col^{Col} with a different cytoplasm is created by crossing the heterozygous HI Zu/Col^{Zu} as a male parent to WT Col as a female parent and selecting correctly through the generations.

References

Bannigan, A., W.-R. Scheible, W. Lukowitz, C. Fagerstrom, P. Wadsworth *et al.*, 2007 A conserved role for kinesin-5 in plant mitosis. *Journal of Cell Science* 120 (16):2819-2827.

Birchler, J.A., Z. Gao, A. Sharma, G.G. Presting, and F. Han, 2011 Epigenetic aspects of centromere function in plants. *Current Opinion in Plant Biology* 14 (2):217-222.

Boyes, D.C., A.M. Zayed, R. Ascenzi, A.J. McCaskill, N.E. Hoffman *et al.*, 2001 Growth Stage–Based Phenotypic Analysis of Arabidopsis: A Model for High Throughput Functional Genomics in Plants. *The Plant Cell* 13 (7):1499-1510.

Britt, A.B., and S. Kuppu, 2016 Cenh3: An Emerging Player in Haploid Induction Technology. *Frontiers in Plant Science* 7 (357).

Dwivedi, S.L., A.B. Britt, L. Tripathi, S. Sharma, H.D. Upadhyaya *et al.*, 2015 Haploids: Constraints and opportunities in plant breeding. *Biotechnology Advances* 33 (6, Part 1):812-829.

Flood, P.J., T.P.J.M. Theeuwen, K. Schneeberger, P. Keizer, W. Kruijer *et al.*, 2018 Reciprocal cybrids reveal how organellar genomes affect plant phenotypes. *bioRxiv*:477687.

Gilles, L.M., A. Khaled, J.B. Laffaire, S. Chaignon, G. Gendrot *et al.*, 2017 Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. *The EMBO Journal*.

Heald, R., R. Tournebize, T. Blank, R. Sandaltzopoulos, P. Becker *et al.*, 1996 Self-organization of microtubules into bipolar spindles around artificial chromosomes in Xenopus egg extracts. *Nature* 382:420.

Henikoff, S., K. Ahmad, and H.S. Malik, 2001 The Centromere Paradox: Stable Inheritance with Rapidly Evolving DNA. *Science* 293 (5532):1098-1102.

Ishii, T., R. Karimi-Ashtiyani, and A. Houben, 2016 Haploidization via Chromosome Elimination: Means and Mechanisms. *Annu Rev Plant Biol* 67:421-438.

Karimi-Ashtiyani, R., T. Ishii, M. Niessen, N. Stein, S. Heckmann *et al.*, 2015 Point mutation impairs centromeric CENH3 loading and induces haploid plants. *Proceedings of the National Academy of Sciences* 112 (36):11211-11216.

Kasha, K.J., and K.N. Kao, 1970 High Frequency Haploid Production in Barley (Hordeum vulgare L.). *Nature* 225:874.

Kelliher, T., D. Starr, L. Richbourg, S. Chintamanani, B. Delzer *et al.*, 2017 MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. *Nature* 542:105.

Kelliher, T., D. Starr, W. Wang, J. McCuiston, H. Zhong *et al.*, 2016 Maternal Haploids Are Preferentially Induced by CENH3-tailswap Transgenic Complementation in Maize. *Frontiers in Plant Science* 7 (414).

Kuppu, S., E.H. Tan, H. Nguyen, A. Rodgers, L. Comai *et al.*, 2015 Point Mutations in Centromeric Histone Induce Post-zygotic Incompatibility and Uniparental Inheritance. *PLOS Genetics* 11 (9):e1005494.

Li, X., D. Meng, S. Chen, H. Luo, Q. Zhang *et al.*, 2017 Single nucleus sequencing reveals spermatid chromosome fragmentation as a possible cause of maize haploid induction. *Nature Communications* 8 (1):991.

Liu, C., W. Li, Y. Zhong, X. Dong, H. Hu *et al.*, 2015 Fine mapping of qhir8 affecting in vivo haploid induction in maize. *Theoretical and Applied Genetics* 128 (12):2507-2515.

Liu, C., X. Li, D. Meng, Y. Zhong, C. Chen *et al.*, 2017 A 4-bp Insertion at ZmPLA1 Encoding a Putative Phospholipase A Generates Haploid Induction in Maize. *Mol Plant* 10 (3):520-522.

Maheshwari, S., E.H. Tan, A. West, F.C.H. Franklin, L. Comai *et al.*, 2015 Naturally Occurring Differences in CENH3 Affect Chromosome Segregation in Zygotic Mitosis of Hybrids. *PLOS Genetics* 11 (1):e1004970.

Malik, H.S., and S. Henikoff, 2003 Phylogenomics of the nucleosome. *Nature Structural Biology* 10:882.

Murovec, J., and B. Bohanec, 2012 Haploids and Doubled Haploids in Plant Breeding, pp. 87-106 in *Plant Breeding*, edited by I. Abdurakhmonov. INTECH Open Access Publisher.

Musacchio, A., and A. Desai, 2017 A Molecular View of Kinetochore Assembly and Function. *Biology* 6 (1):5.

Ravi, M., and R. Bondada, 2016 Genome Elimination by Tailswap CenH3: In Vivo Haploid Production in Arabidopsis thaliana, pp. 77-99 in *Chromosome and Genomic Engineering in Plants: Methods and Protocols*, edited by M. Murata. Springer New York, New York, NY.

Ravi, M., and S.W. Chan, 2010 Haploid plants produced by centromere-mediated genome elimination. *Nature* 464 (7288):615-618.

Ravi, M., P.N. Kwong, R.M.G. Menorca, J.T. Valencia, J.S. Ramahi *et al.*, 2010 The rapidly evolving centromere-specific histone has stringent functional requirements in Arabidopsis thaliana. *Genetics* 186 (2):461-471.

Ravi, M., M.P.A. Marimuthu, E.H. Tan, S. Maheshwari, I.M. Henry *et al.*, 2014 A haploid genetics toolbox for Arabidopsis thaliana. *Nature Communications* 5.

Ravi, M., F. Shibata, J.S. Ramahi, K. Nagaki, C. Chen *et al.*, 2011 Meiosis-Specific Loading of the Centromere-Specific Histone CENH3 in Arabidopsis thaliana. *PLOS Genetics* 7 (6):e1002121.

Sanei, M., R. Pickering, K. Kumke, S. Nasuda, and A. Houben, 2011 Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. *Proceedings of the National Academy of Sciences* 108 (33):E498-E505.

Sloan, D.B., J.M. Warren, A.M. Williams, Z. Wu, S.E. Abdel-Ghany *et al.*, 2018 Cytonuclear integration and co-evolution. *Nature Reviews Genetics* 19 (10):635-648.

Topp, C.N., R.J. Okagaki, J.R. Melo, R.G. Kynast, R.L. Phillips *et al.*, 2009 Identification of a Maize Neocentromere in an Oat-Maize Addition Line. *Cytogenetic and Genome Research* 124 (3-4):228-238.

Wang, N., and R.K. Dawe, 2018 Centromere Size and Its Relationship to Haploid Formation in Plants. *Mol Plant* 11 (3):398-406.

Yamada, M., and G. Goshima, 2017 Mitotic Spindle Assembly in Land Plants: Molecules and Mechanisms. *Biology* 6 (1):6.

Yao, L., Y. Zhang, C. Liu, Y. Liu, Y. Wang *et al.*, 2018 OsMATL mutation induces haploid seed formation in indica rice. *Nature plants* 4 (8):530-533.

Chapter 5. Salt stress tolerance in *Arabidopsis thaliana* is not influenced by genome dosage or parent-of-origin effects *in vitro*

Introduction

Whole genome duplication events are a major mechanism of plant evolution and speciation (Blanc and Wolfe 2004; Jiao et al. 2011; Li et al. 2015). Genome dosage increases in newly formed polyploids can give rise to new phenotypes, while genetic redundancy within polyploid genomes can allow duplicated genes to take on a new function (neofunctionalization) or retain different components of an original function (subfunctionalization) (Blanc and Wolfe 2004; Comai 2005; Jiao et al. 2011; Roulin et al. 2013). The survival and reproduction of polyploid plants, as compared to their diploid predecessor, is proposed to be linked with altered gene expression and physiology.

Gene and genome dosage effects on plant growth have been reported in allopolyploids (polyploids with divergent genomes) (Wang et al. 2006; Chen 2010), but these plants are genetic hybrids and so determining the contribution of genome dosage *versus* genetic hybridity on plant growth is difficult. In contrast, by using autopolyploids (polyploids with genomes of the same type) it is possible to create an isogenic ploidy series which differ only in the number of chromosome copies in the nucleus. Autopolyploid research in *Arabidopsis thaliana* (Miller et al. 2012; Duszynska et al. 2013; Donoghue et al. 2014; Fort et al. 2016) and maize (*Zea mays*) (Guo et al. 1996; Yao et al. 2013) has shown both genome dosage and parent-of-origin effects (whether a chromosome is maternally or paternally inherited) on plant growth and development. The induction of autopolyploidy in commercial crops for improvement in yield and quality is seen in potato (*S. tuberosum*) (Jansky 2009), sugarcane (*Saccharum officinarum*) (Ming et al. 2001), perennial ryegrass (*Lolium perenne*) (Wilkins and Humphreys 2003), blueberry (*Vaccinium corymbosum*) (McCallum et al. 2016), and alfalfa (*Medicago sativa*) (Brouwer and Osborn 1999) to name just five. As well as increasing yield and quality, improving the salt stress tolerance of crops is an important objective in plant breeding (Pitman and Läuchli 2002). However, research investigating genome dosage effects on abiotic stress tolerance is not as widespread in the literature, and research investigating parent-of-origin effects is absent.

Research into the abiotic stress response of autopolyploid plants and their natural ploidy-level equivalents have revealed positive, negative and neutral genome dosage effects. This holds for both annual and perennial plants. In diploid turnip (*Brassica rapa*) the relatively salt stress-tolerant cultivar ‘Aijiaohuang’ exhibits further salt stress tolerance at the tetraploid level (Meng et al. 2011). In *A. thaliana* it has been reported, for a small number of accessions, that plants are more salt tolerant at the tetraploid level than at the diploid level: as measured by

days-to-death, seed yield, and levels of anthocyanin (Chao et al. 2013; Pozo and Ramirez-Parra 2014). In Chrysanthemum (*Chrysanthemum indicum*), for a single genotype, it was demonstrated that cold, salt and drought tolerance were improved upon induction of tetraploidy, but heat tolerance was greater at the diploid level (Liu et al. 2011). Using field transplant experiments of wild yarrow (*Achillea borealis*), hexaploid plants are more likely to survive sand dune environments than tetraploid plants, although population effects are also significant (Ramsey 2011). The enhanced drought tolerance of wild willowherb (*Chamerion angustifolium*) at the tetraploid level over the diploid level has been attributed to evolution through natural selection, rather than an intrinsic response to WGD (Maherali et al. 2009). Taken together, data from diverse plant species of different growing habits suggest that genome dosage generally has a positive impact on abiotic stress response.

In this experiment I utilized NaCl as an ionic stress. I define this stress as a "distress", i.e. a condition leading to a major imbalance between the environment and physiology (Jansen 2017). I did not allow the stress to lead to plant death. The aim of this experiment was to investigate if there was (1) a genome dosage effect, and (2) a parent-of-origin effect on salt stress tolerance across a range of *A. thaliana* accessions. I define "salt stress tolerance" as the plant biomass produced in saline conditions *versus* non-saline conditions (Munns 2010).

To effectively answer my research questions, the first step was to identify the "optimal" concentration at which salt stress can be applied to all ten accessions at their natural ploidy level. A "low" salt concentration will fail to induce stress, while a "high" salt concentration may induce leaf bleaching and cease growth completely. Following experiments with NaCl concentration gradients *in vitro*, I proceeded to investigate the salt stress tolerance of all ten accessions in an isogenic ploidy series (i.e. at the diploid, tetraploid and reciprocal triploid level). While there were some noteworthy accession-specific results, I conclude that, in general, salt stress tolerance in *A. thaliana* is not influenced by genome dosage and parent-of-origin effects.

Materials and Methods

Plant material and crossing design

Ten *A. thaliana* accessions were kindly provided both at the diploid and tetraploid level. C24, Col-0, L*er*-0, Zu-0 were the kind gift of Luca Comai (UC Davis, CA, USA) and Bur-0, Cvi, Sorbo, T910, TAL07, Wilna were the kind gift of Ortrun Mittelsten Scheid (GmbH, Vienna, Austria). These accessions have their ecological origin assigned to eight different countries

(**Figure *19***). Plants were grown in 7 X 7 X 6.5cm pots (Modiform, Leusden, Netherlands) in soil (5:1:1 mix of peat:vermiculite:perlite). Growth room (Cambridge HOK, East Yorkshire, UK) conditions were 16/8hr light/dark @22/20°C. Plants were maintained for at least six generations before crossing. Reciprocal triploids were created by manually emasculating flowers on lateral stems with a Dumont no. 5 tweezers (Electron Microscopy Sciences, PA, USA) and reciprocally crossing the diploid and tetraploid lines in both directions: 2x ♀ X 4x ♂ crosses produced paternal-excess triploids – labelled 3x(P), while 4x ♀ X 2x ♂ crosses produced maternal-excess triploids – labelled 3x(M). A single maternal and single paternal plant were used for all crosses, except for Col-0 and Zu-0, where three maternal plants were used to create 3x(P) seed due to the strong triploid block. Self-pollinated diploid and tetraploid siliques were harvested from the maternal parent on the same lateral stem used for generating triploids.

Growth media

All chemicals were obtained from Sigma Aldrich, Ireland. Growth media was prepared by adding ½ strength Murashige and Skoog Basal Medium and sucrose at 0.5% to distilled water. The required amount of NaCl was added. The solution was brought to pH 5.7 using 1M KOH dropwise. Lastly, agar at 0.8% was added before sterilization.

Plant stress experiments

Seeds were first surface sterilized with 70% methanol, then a seed sterilization solution consisting of 5% sodium hypochlorite solution (NaClO) with 0.01% Triton X-100, followed by 5 times wash with sterile, distilled water. Seeds were stratified for 3 days in the dark at 4°C. Seeds were sown to fresh, stress-free media in 100 X 100 X 20mm petri-dishes (Sarstedt, Nümbrecht, Germany) sealed with Micropore™ tape (3M, MN, USA). Plates were horizontally positioned in a growth chamber (CLF Plant Climatics, Emersacker, Germany) with 16/8hr light/dark @22/20°C. After 2 days plates were positioned vertically. 9 day-old plants (growth stage 1.02 (Boyes et al. 2001)) were transferred to fresh media with or without NaCl. Diploid and tetraploid plants were grown together in the same vertical plates to minimize any plate effect, as were reciprocal triploid plants. Stress plates contained max. 5 biological replicates per ploidy level while stress-free plates contained max. 3 biological replicates. 7 days later plants were destructively harvested for fresh weight measurements. Plants were dried on a paper towel and weighed on a NewClassic MF weighing scale (METTLER TOLEDO,

Greifensee, Switzerland) to the ten-thousandth decimal value. The percentage plant biomass produced in saline conditions *versus* non-saline conditions was calculated. Data was inspected with a Shapiro-Wilk test to determine if it was normally distributed and statistical differences were determined with a two-way independent samples t-test.

Flow cytometry

Plants were grown in a growth room as before. All chemicals were obtained from Sysmex, Kobe, Japan. Approx. 3cm^2 of (new) leaf material from 1-month old plants was removed and chopped with a razor blade in the presence of 400µl nuclei extraction buffer. After 5 minutes the mixture was strained into a 3.5ml Röhren tube through a 30µm CellTrics® filter. 1ml of UV-stain was added before the sample was analyzed on a Partec Ploidy Analyzer (Sysmex, Kobe, Japan). Ploidy levels of diploid, tetraploid and reciprocal triploid plants were confirmed (Supplementary Results 1).

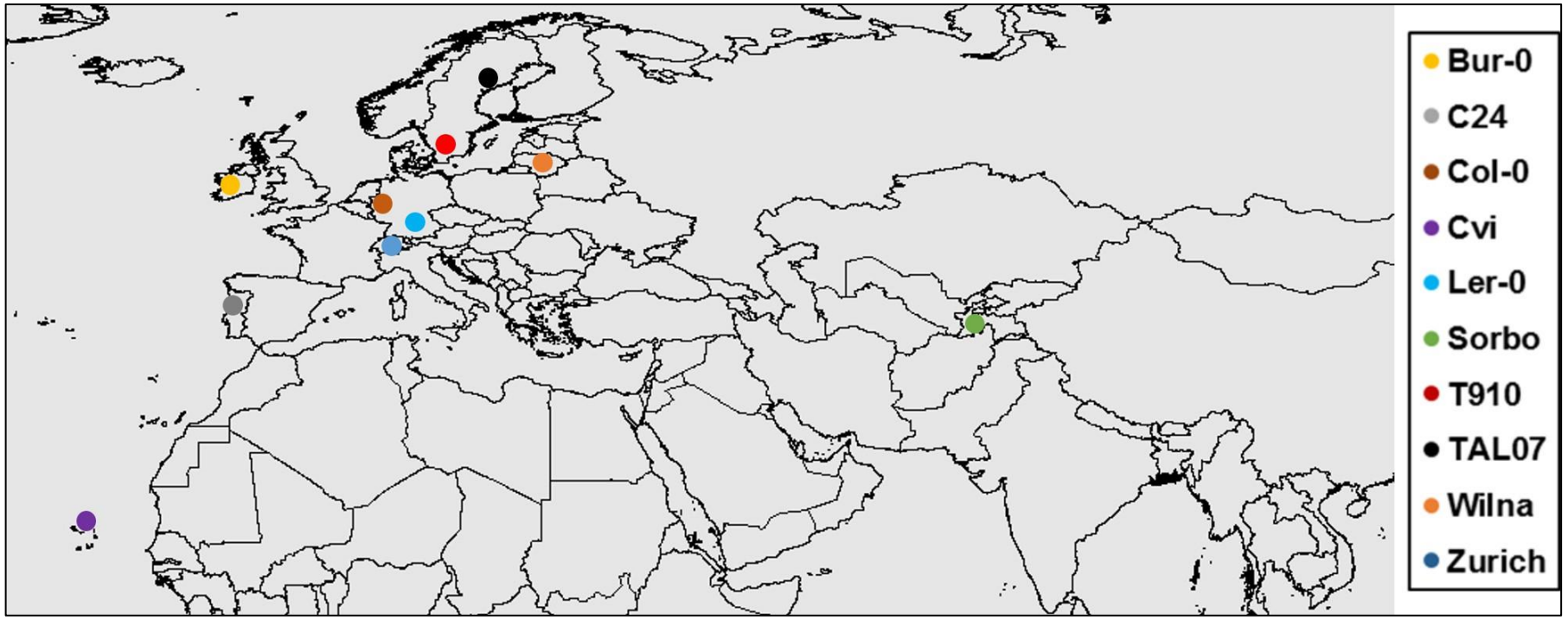


Figure 19. Ecological origin of *A. thaliana* accessions used in this experiment. The country of origin of each accession is as follows: Bur-0, Ireland; C24, Portugal; Col-0 and L*er*-0, Germany; Cvi, Cape Verde Islands; Sorbo, Tajikistan, T910 and TAL07, Sweden; Wilna, Lithuania; Zurich, Switzerland.

Results

Concentration of 125mM NaCl is optimal to induce stress in most *A. thaliana* accessions *in vitro*

All ten accessions, at their natural ploidy level (diploid), were grown for 9-days in stress-free media before transfer to different media supplemented with a range of NaCl concentrations. The objective was to determine the highest salt stress concentration that did not stop growth nor induce bleaching. Response to low NaCl concentrations (50mM) varies between accessions. In Col-0, for example, it stimulates growth whereas in TAL07 it has a considerable inhibitory effect (**Figure *20***). The highest of NaCl concentration (150mM) leads to bleaching in all accessions and cessation of growth in most accessions. For most accessions it was determined that 125mM NaCl is optimal as it inhibits growth significantly (but not completely) and widespread bleaching is absent. Accession Cvi was identified as particularly salt-sensitive, as reported by others (Borsani et al. 2001), and thus I determined should be tested at max. 100mM NaCl.

Genome dosage effect on salt stress tolerance in two accessions

Both in stress and stress-free media tetraploid plants are larger than diploid plants, as measured for both above- and below-ground biomass (**Figure *21*** A&C). For two accessions there is a difference in salt stress tolerance between diploid and tetraploid plants. In the accession Bur-0, diploid plants accumulate significantly ($P < 0.05$) more above-ground biomass than tetraploid plants. In the accession Cvi, tetraploid plants accumulate significantly ($P < 0.05$) more below-ground biomass than diploid plants (Supplementary Results 2). For all other accessions, diploid and tetraploid plants accumulate similar above- and below-ground biomass under salt stress. On average diploid and tetraploid plants accumulate equal biomass under salt stress: above-ground biomass is reduced by ~50% and below ground biomass is reduced by ~70% (**Figure *21*** D&E).

Parent-of-origin effect on salt stress tolerance in four accessions

Both in stress and stress-free media paternal-excess triploid plants are larger than maternal-excess triploid plants, as measured for both above- and below-ground biomass (**Figure *21*** E&G). For four accessions there is a difference in salt stress response between reciprocal triploid plants. In the accession Wilna maternal-excess triploid plants accumulate significantly

($P < 0.05$) more above- and below-ground biomass than paternal-excess triploid plants. In the accession L*er*-0 maternal-excess triploid plants accumulate significantly ($P < 0.05$) more above-ground biomass than paternal-excess triploid plants. In the accessions C24 and Col-0, paternal-excess triploid plants accumulate significantly ($P < 0.05$) more below-ground biomass than maternal-excess triploid plants (Supplementary Results 2). For all other accessions, reciprocal triploid plants accumulate similar above- and below-ground biomass under salt stress. On average reciprocal triploid plants accumulate equal biomass under salt stress: above-ground biomass is reduced by ~50% and below ground biomass is reduced by ~75% (**Figure *21*** F&H).

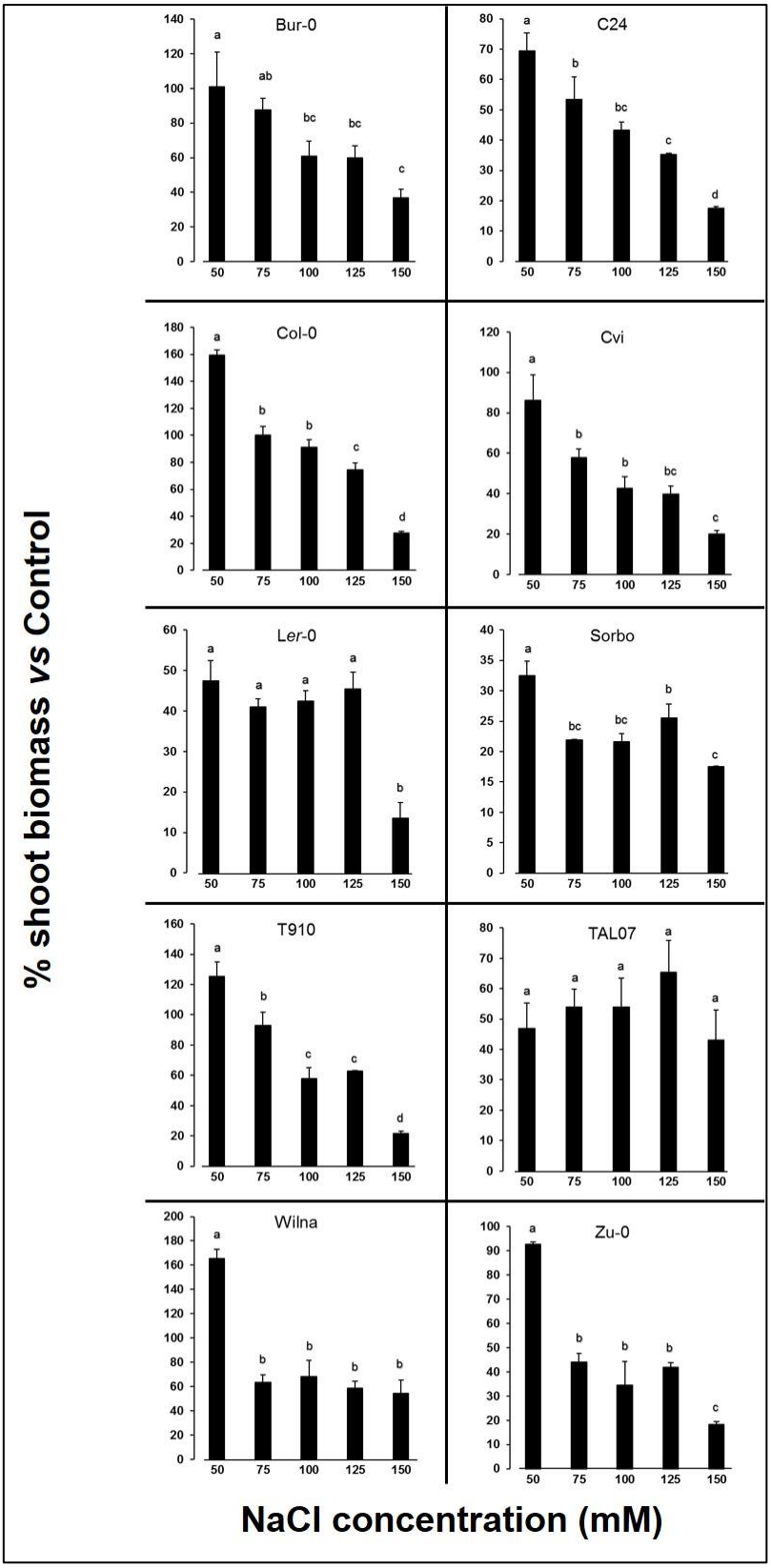


Figure 20. Above-ground biomass accumulation of ten *A. thaliana* accessions under a range of salt concentrations. Plants were germinated in stress-free artificial growth media and after 9 days transferred to fresh media with or without NaCl. Plants were destructively harvested after 7 days. Each accession was analyzed with a one-way ANOVA and means assigned different letters are statistically different ($P < 0.05$) according to Tukey's HSD test.

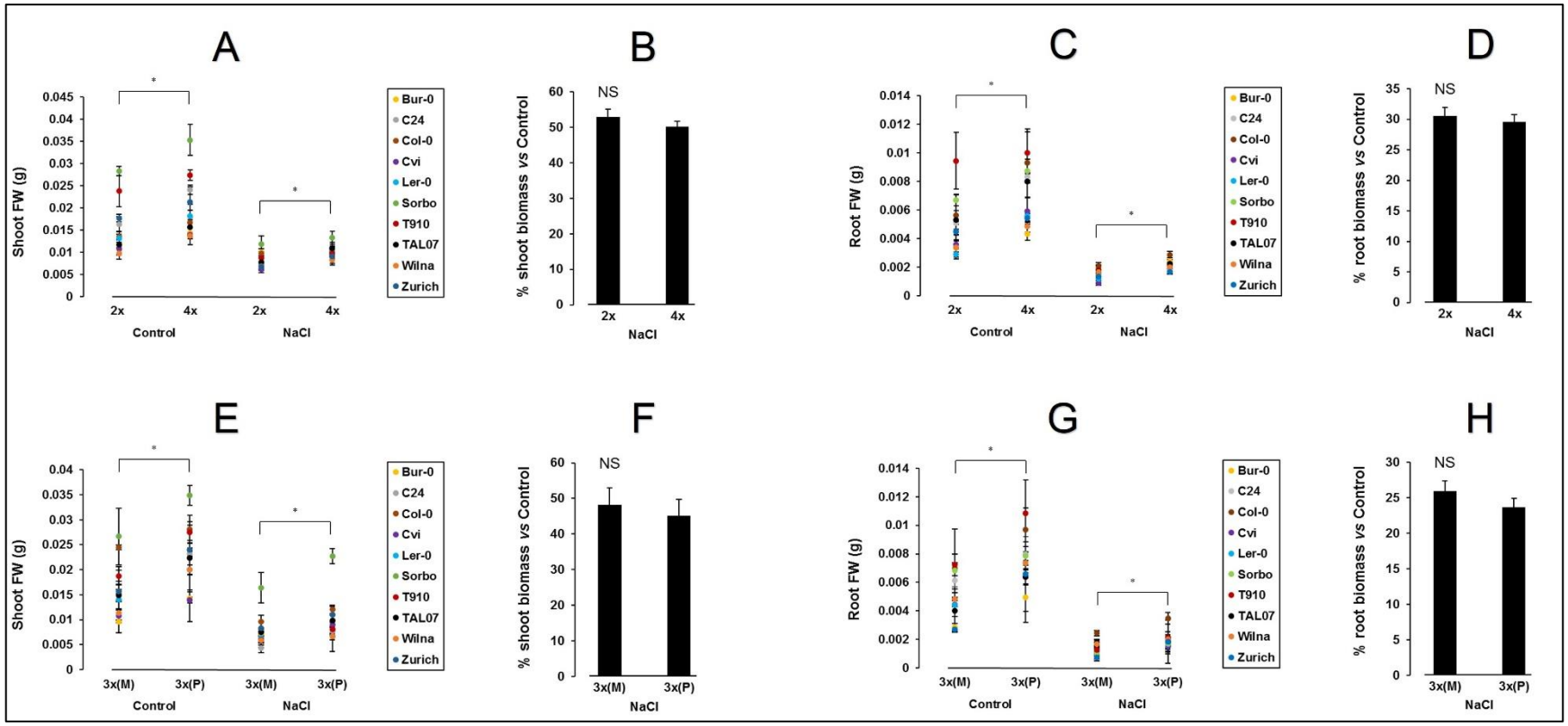


Figure 21. There is no genome dosage or parent-of-origin effect on salt stress tolerance. A Tetraploid plants accumulate more above-ground biomass than diploid plants. **B** Diploid and tetraploid plants accumulate equal above-ground biomass under saline conditions. **C** Tetraploid plants accumulate more below-ground biomass than diploid plants. **D** Diploid and tetraploid plants accumulate equal below-ground biomass under saline conditions. **E** Paternal-excess triploid plants accumulate more above-ground biomass than maternal-excess triploid plants. **F** Maternal- and paternal-excess triploid plants accumulate equal above-ground biomass under saline conditions. **G** Paternal-excess triploid plants accumulate more below-ground biomass than maternal-excess triploid plants. **H** Maternal- and paternal-excess triploid plants accumulate equal below-ground biomass under saline conditions. Plants were germinated in stress-free artificial growth media and after 9 days transferred to fresh media with or without NaCl. Plants were destructively harvested after 7 days. * statistically different ($P < 0.05$); NS not statistically different ($P > 0.05$) according to a two-tailed independent samples t-test.

Discussion

Salt stress in plants

Saline conditions impose physiological constraints on plants through, firstly, oxidative stress, and secondly, ionic stress (Shabala 2017). Salt accumulating in the substrate surrounding the roots inhibits the capacity of roots to take up water (an osmotic effect). Lower water availability at this stage will lead to reduced plant growth, as has been shown for *A. thaliana* (Luo et al. 2017), rice (Yeo et al. 1991), maize (Frensch and Hsiao 1994; Rodriguez et al. 1997), and barley (Munns et al. 2000). Plant roots use the ionic composition of the substrate in which they are grown for turgor recovery, and thus can soon begin to take in water after the initial osmotic stress (Shabala and Lew 2002). However, this water contains very high Na^+ and Cl^- concentrations leading to the second major physiological constraint on growth: ionic stress. Na^+ and K^+ ions possess similar physio-chemical properties: Na^+ can compete with K^+ for important binding sites within the cell, impairing enzyme activity (Shabala and Lew 2002; Flowers et al. 2010) . In addition, Na^+ and Cl^- ions can accumulate in the cell wall causing cell dehydration (Munns and Passioura 1984). It appears the inability to compartmentalize harmful ions inside the cell (e.g. in the vacuole) inhibits regular cell function and leads to cell death from either toxicity or dehydration (Munns 2002). This internal injury inhibits new leaf growth, reduces overall plant photosynthesis and thus reduces supply of carbohydrates to new cells (Shabala 2017).

Plants have evolved different ionic stress tolerance mechanisms. These mechanisms vary across species and can depend on local environmental conditions as well as the length of salinity exposure (Munns and Tester 2008). These are (1) the ability of roots to recognize Na^+ ions and exclude them from accumulating within the plant (Matsushita and Matoh 1991; Alberico and Cramer 1993; Fortmeir and Schubert 1995; Byrt et al. 2007), and (2) tissue tolerance of Na^+ and Cl^- ions through compartmentalization (Apse et al. 1999; Mühling and Läuchli 2002; Flowers and Colmer 2008; Flowers et al. 2010).

Arabidopsis thaliana accessions display a large variation in tolerance to low concentrations of NaCl, but not high NaCl concentrations

The accessions used in this experiment originate from across the northern hemisphere (**Figure *19***). While *A. thaliana* accessions show only small differences in nucleotide sequences they can display large genetic variation for phenotypic characteristics, e.g. flowering time

(Kowalski et al. 1994). At low concentrations of NaCl there is notable variation between accessions in this experiment. Four accessions (Bur-0, Col-0, T910, Wilna) display normal or above-normal above-ground biomass accumulation at 50mM NaCl, while others (L*er*-0, Sorbo, TAL07) display ~50% reduction in above-ground biomass accumulation at this concentration (**Figure *20***).

Previous work has shown large differences between *A. thaliana* accessions for NaCl tolerance, as measured using days-to-death (Katori et al. 2010) and leaf rosette area coupled with electrolyte leakage (Julkowska et al. 2016). My experiment did not find a large variation in NaCl tolerance between accessions at high NaCl concentrations (**Figure *20***), even though there is some overlap with the accessions used in this study and in the previously cited literature. It is noteworthy that these experiments used much higher concentrations of NaCl than my study, both in soil and in artificial growth media *in vitro* experiments, and plants at both younger and older growth stages.

Increasing the genome dosage, or whether a genome is paternally or maternally inherited, does not influence salt stress tolerance

As reported previously (Fort et al. 2016), young *A. thaliana* tetraploid plants accumulate more above-ground biomass than their diploid equivalents, indicating a dosage effect. Likewise, young paternal-excess triploid plants also accumulate more above-ground biomass than their diploid equivalent, as well as maternal-excess triploid plants. The same dosage effect and parent-of-origin effect on plant growth is shown in this experiment, including new, equivalent data for below-ground biomass (**Figure *21*** A, C, E, &G). Despite some accession-specific results, on average there is no difference in salt stress tolerance between diploid and tetraploid plants (**Figure *21*** B&D). Likewise, on average there is no difference in salt stress tolerance between paternal- and maternal-excess triploid plants (**Figure *21*** F&H). This indicates that there is no genome dosage, nor parent-of-origin, effect on salt stress tolerance under these conditions.

Assessing plant growth under saline conditions

The absence of a genome dosage effect in my experiment conflicts with the literature. Chao et al. (2013) used three *A. thaliana* accessions and reported that tetraploid plants lived ~5 days longer than diploid plants when irrigated continuously with salt water unto death, while tetraploid plants also had a higher seed set when irrigated continuously with a lower

concentration of salt water until senescence. Higher salt tolerance in tetraploid plants was attributed to K^+ ion accumulation, and Na^+ exclusion, in the leaves. Del Pozo and Ramirez-Parra (2014) used two *A. thaliana* accessions and reported that tetraploid plants recovered to a greater degree than diploid plants after a period of salt-water irrigation followed by normal irrigation. In addition, tetraploid plants showed lower levels of anthocyanin (a stress-induced pigment) than diploid plants under salt water irrigation. Higher salt tolerance in tetraploid plants was attributed to enhanced abscisic acid signalling. The inconsistencies between my results and the published literature can be explained by differences in experimental design.

The term "stress" is often used arbitrarily in plant physiology literature and is rarely defined. Depending on the severity of the "stressor", whether it be light, heat, physical disturbance, etc., it can act as an essential environmental signal for the plant (a "eustress") or can lead to destabilization of cellular functions, metabolism, and plant physiology (a "distress") (Jansen 2017). I designed this experiment to specifically induce a distress caused by salt. The choice of measurements for assessing tolerance to this distress is theoretically endless. Salt stress experiments generally measure, at the physiological level, transpiration, biomass, yield, or survival; at the biochemical level, hormones, reactive oxygen species, or enzymatic activity; and at the genetic level, gene expression. Of all these, the most common measurement for assessing salt tolerance is the amount of biomass added (Munns 2010).

I chose to measure the amount of fresh biomass added in an easy-to-use and replicable manner on 16-day old plants: artificial growth media *in vitro*. The advantages of working *in vitro* is that seeds and young plants can be exposed to an accurate, uniform and reproducible chemical concentration of macronutrients, micronutrients, as well as any other dissolved chemical in the growth media; making them useful in fundamental research to investigate plant growth and identify genetic loci in response to growth hormones (Staswick et al. 1992) or abiotic stresses such as salt (Burssens et al. 2000; Shi et al. 2000; Kim et al. 2008; Zolla et al. 2009). The biomass data I generated *in vitro* is not easily related to plant survival data, seed yield data, nor biochemical data generated *in vivo* (Chao et al. 2013; Del Pozo and Ramirez-Parra 2014), but crucially root growth in my experiment is reduced to a greater extent than shoot growth under stress (**Figure *21*** B, D, F &H) – consistent with the ionic effect, rather than the osmotic effect, of the salt (Verslues et al. 2006).

Conclusions

The effects of soil salinization are greatest in arid and semi-arid agricultural settings. Response to soil salinization can be physical (e.g. drainage management) or biological (breeding crops for enhanced salt tolerance) (Acosta-Motos et al. 2017). The induction of polyploidy is one method of crop improvement, and the literature suggests abiotic stress tolerance may correlate with genome dosage. Using the model plant species *Arabidopsis thaliana*, I have shown that salt stress tolerance – as measured using the fundamental unit of biomass added – is not influenced by doubling the genome dosage, nor by inheriting an additional genome maternally or paternally.

Future Directions

This research chapter has been written up as a paper for submission to *BMC Research Notes*. I feel this work is suitable for this journal because it publishes null results from scientifically valid experiments.

References

Acosta-Motos, J., M. Ortuño, A. Bernal-Vicente, P. Diaz-Vivancos, M. Sanchez-Blanco *et al.*, 2017 Plant Responses to Salt Stress: Adaptive Mechanisms. *Agronomy* 7 (1):18.

Alberico, G.J., and G.R. Cramer, 1993 Is the salt tolerance of maize related to sodium exclusion? I. Preliminary screening of seven cultivars. *Journal of Plant Nutrition* 16 (11):2289-2303.

Apse, M.P., Gilad S. Aharon, Wayne A. Snedden, and E. Blumwald, 1999 Salt tolerance conferred by overexpression of a vacuolar Na^+/H^+ antiport in Arabidopsis. *Science* 285:1256-1258.

Blanc, G., and K.H. Wolfe, 2004 Widespread Paleopolyploidy in Model Plant Species Inferred from Age Distributions of Duplicate Genes. *The Plant Cell* 16 (7):1667-1678.

Borsani, O., V. Valpuesta, and M.A. Botella, 2001 Evidence for a role of salicylic acid in the oxidative damage generated by NaCl and osmotic stress in Arabidopsis seedlings. *Plant Physiology* 126 (3):1024-1030.

Boyes, D.C., A.M. Zayed, R. Ascenzi, A.J. McCaskill, N.E. Hoffman *et al.*, 2001 Growth Stage–Based Phenotypic Analysis of Arabidopsis. *A Model for High Throughput Functional Genomics in Plants* 13 (7):1499-1510.

Brouwer, D.J., and T.C. Osborn, 1999 A molecular marker linkage map of tetraploid alfalfa (Medicago sativa L.). *Theoretical and Applied Genetics* 99 (7):1194-1200.

Burssens, S., K. Himanen, B. van de Cotte, T. Beeckman, M. Van Montagu *et al.*, 2000 Expression of cell cycle regulatory genes and morphological alterations in response to salt stress in Arabidopsis thaliana. *Planta* 211 (5):632-640.

Byrt, C.S., J.D. Platten, W. Spielmeyer, R.A. James, E.S. Lagudah *et al.*, 2007 HKT1;5-Like Cation Transporters Linked to Na(+) Exclusion Loci in Wheat, Nax2 and Kna1. *Plant Physiology* 143 (4):1918-1928.

Chao, D.Y., B. Dilkes, H. Luo, A. Douglas, E. Yakubova *et al.*, 2013 Polyploids exhibit higher potassium uptake and salinity tolerance in Arabidopsis. *Science* 341 (6146):658-659.

Chen, Z.J., 2010 Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci* 15 (2):57-71.

Comai, L., 2005 The advantages and disadvantages of being polyploid. *Nature reviews genetics* 6 (11):836-846.

Del Pozo, J.C., and E. Ramirez-Parra, 2014 Deciphering the molecular bases for drought tolerance in Arabidopsis autotetraploids. *Plant, Cell & Environment* 37 (12):2722-2737.

Donoghue, M.T., A. Fort, R. Clifton, X. Zhang, P.C. McKeown *et al.*, 2014 C(m)CGG methylation-independent parent-of-origin effects on genome-wide transcript levels in isogenic reciprocal F1 triploid plants. *DNA Res* 21 (2):141-151.

Duszynska, D., P.C. McKeown, T.E. Juenger, A. Pietraszewska-Bogiel, D. Geelen *et al.*, 2013 Gamete fertility and ovule number variation in selfed reciprocal F1 hybrid triploid plants are heritable and display epigenetic parent-of-origin effects. *New Phytol* 198 (1):71-81.

Flowers, T.J., and T.D. Colmer, 2008 Salinity tolerance in halophytes. *New Phytologist* 179 (4):945-963.

Flowers, T.J., H.K. Galal, and L. Bromham, 2010 Evolution of halophytes: multiple origins of salt tolerance in land plants. *Functional Plant Biology* 37 (7):604-612.

Fort, A., P. Ryder, P.C. McKeown, C. Wijnen, M.G. Aarts *et al.*, 2016 Disaggregating polyploidy, parental genome dosage and hybridity contributions to heterosis in Arabidopsis thaliana. *New Phytol* 209 (2):590-599.

Fortmeir, R., and S. Schubert, 1995 Salt tolerance of maize (Zea mays L.): the role of sodium exclusion. *Plant, Cell & Environment* 18 (9):1041-1047.

Frensch, J., and T.C. Hsiao, 1994 Transient Responses of Cell Turgor and Growth of Maize Roots as Affected by Changes in Water Potential. *Plant Physiology* 104 (1):247-254.

Guo, M., D. Davis, and J.A. Birchler, 1996 Dosage effects on gene expression in a maize ploidy series. *Genetics* 142 (4):1349-1355.

Jansen, M.A.K., Potters, G., 2017 Stress: The Way of Life, pp. ix - xiv in *Plant Stress Physiology, 2nd Edition*, edited by S. Shabala. CABI, Croydon, UK.

Jansky, S., 2009 Breeding, genetics, and cultivar development, pp. 27-62 in *Advances in potato chemistry and technology*. Elsevier.

Jiao, Y., N.J. Wickett, S. Ayyampalayam, A.S. Chanderbali, L. Landherr *et al.*, 2011 Ancestral polyploidy in seed plants and angiosperms. *Nature* 473 (7345):97-100.

Julkowska, M.M., K. Klei, L. Fokkens, M.A. Haring, M.E. Schranz *et al.*, 2016 Natural variation in rosette size under salt stress conditions corresponds to developmental differences between Arabidopsis accessions and allelic variation in the LRR-KISS gene. *J Exp Bot* 67 (8):2127-2138.

Katori, T., A. Ikeda, S. Iuchi, M. Kobayashi, K. Shinozaki *et al.*, 2010 Dissecting the genetic control of natural variation in salt tolerance of Arabidopsis thaliana accessions. *Journal of Experimental Botany* 61 (4):1125-1138.

Kim, J.S., K.A. Kim, T.R. Oh, C.M. Park, and H. Kang, 2008 Functional Characterization of DEAD-Box RNA Helicases in Arabidopsis thaliana under Abiotic Stress Conditions. *Plant and Cell Physiology* 49 (10):1563-1571.

Kowalski, S.P., T.-H. Lan, K.A. Feldmann, and A.H. Paterson, 1994 QTL mapping of naturally-occurring variation in flowering time of Arabidopsis thaliana. *Molecular and General Genetics MGG* 245 (5):548-555.

Li, Z., A.E. Baniaga, E.B. Sessa, M. Scascitelli, S.W. Graham *et al.*, 2015 Early genome duplications in conifers and other seed plants. *Science Advances* 1 (10):e1501084.

Liu, S., S. Chen, Y. Chen, Z. Guan, D. Yin *et al.*, 2011 In vitro induced tetraploid of Dendranthema nankingense (Nakai) Tzvel. shows an improved level of abiotic stress tolerance. *Scientia Horticulturae* 127 (3):411-419.

Luo, L., P. Zhang, R. Zhu, J. Fu, J. Su *et al.*, 2017 Autophagy Is Rapidly Induced by Salt Stress and Is Required for Salt Tolerance in Arabidopsis. *Frontiers in plant science* 8:1459.

Maherali, H., A.E. Walden, and B.C. Husband, 2009 Genome duplication and the evolution of physiological responses to water stress. *New Phytologist* 184 (3):721-731.

Matsushita, N., and T. Matoh, 1991 Characterization of Na+ exclusion mechanisms of salt-tolerant reed plants in comparison with salt-sensitive rice plants. *Physiologia Plantarum* 83 (1):170-176.

McCallum, S., J. Graham, L. Jorgensen, L.J. Rowland, N.V. Bassil *et al.*, 2016 Construction of a SNP and SSR linkage map in autotetraploid blueberry using genotyping by sequencing. *Molecular breeding* 36 (4):41.

Meng, H.-b., S.-s. Jiang, S.-j. Hua, X.-y. Lin, Y.-l. Li *et al.*, 2011 Comparison Between a Tetraploid Turnip and Its Diploid Progenitor (Brassica rapa L.): The Adaptation to Salinity Stress. *Agricultural Sciences in China* 10 (3):363-375.

Miller, M., C. Zhang, and Z.J. Chen, 2012 Ploidy and Hybridity Effects on Growth Vigor and Gene Expression in Arabidopsis thaliana Hybrids and Their Parents. *G3 (Bethesda)* 2 (4):505-513.

Ming, R., S.C. Liu, P.H. Moore, J.E. Irvine, and A.H. Paterson, 2001 QTL analysis in a complex autopolyploid: genetic control of sugar content in sugarcane. *Genome research* 11 (12):2075-2084.

Mühling, K.H., and A. Läuchli, 2002 Effect of salt stress on growth and cation compartmentation in leaves of two plant species differing in salt tolerance. *Journal of Plant Physiology* 159 (2):137-146.

Munns, R., 2002 Comparative physiology of salt and water stress. *Plant, Cell & Environment* 25 (2):239-250.

Munns, R., 2010 Approaches to identifying genes for salinity tolerance and the importance of timescale, pp. 25-38 in *Plant stress tolerance*, edited by R. Sunkar. Springer.

Munns, R., and J. Passioura, 1984 Effect of Prolonged Exposure to Nacl on the Osmotic Pressure of Leaf Xylem Sap From Intact, Transpiring Barley Plants. *Functional Plant Biology* 11 (6):497-507.

Munns, R., J.B. Passioura, J. Guo, O. Chazen, and G.R. Cramer, 2000 Water relations and leaf expansion: importance of time scale. *Journal of Experimental Botany* 51 (350):1495-1504.

Munns, R., and M. Tester, 2008 Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.* 59:651-681.

Pitman, M.G., and A. Läuchli, 2002 Global Impact of Salinity and Agricultural Ecosystems, pp. 3-20 in *Salinity: Environment - Plants - Molecules*, edited by A. Läuchli and U. Lüttge. Springer Netherlands, Dordrecht.

Pozo, J.C., and E. Ramirez-Parra, 2014 Deciphering the molecular bases for drought tolerance in Arabidopsis autotetraploids. *Plant, Cell & Environment* 37 (12):2722-2737.

Ramsey, J., 2011 Polyploidy and ecological adaptation in wild yarrow. *Proceedings of the National Academy of Sciences* 108 (17):7096-7101.

Rodriguez, H.G., J. Roberts, W.R. Jordan, and M.C. Drew, 1997 Growth, Water Relations, and Accumulation of Organic and Inorganic Solutes in Roots of Maize Seedlings during Salt Stress. *Plant Physiology* 113 (3):881-893.

Roulin, A., P.L. Auer, M. Libault, J. Schlueter, A. Farmer *et al.*, 2013 The fate of duplicated genes in a polyploid plant genome. *Plant J* 73 (1):143-153.

Shabala, S., Munns, R., 2017 Salinity Stress: Physiological Constraints and Adaptive Mechanisms, pp. 24 - 63 in *Plant Stress Physiology, 2nd Edition*, edited by S. Shabala. CABI.

Shabala, S.N., and R.R. Lew, 2002 Turgor Regulation in Osmotically Stressed Arabidopsis Epidermal Root Cells. Direct Support for the Role of Inorganic Ion Uptake as Revealed by Concurrent Flux and Cell Turgor Measurements. *Plant Physiology* 129 (1):290-299.

Shi, H., M. Ishitani, C. Kim, and J.-K. Zhu, 2000 The Arabidopsis thaliana salt tolerance gene SOS1 encodes a putative Na^+/H^+ antiporter. *Proceedings of the National Academy of Sciences* 97 (12):6896-6901.

Staswick, P.E., W. Su, and S.H. Howell, 1992 Methyl jasmonate inhibition of root growth and induction of a leaf protein are decreased in an Arabidopsis thaliana mutant. *Proceedings of the National Academy of Sciences* 89 (15):6837-6840.

Verslues, P.E., M. Agarwal, S. Katiyar-Agarwal, J. Zhu, and J.-K. Zhu, 2006 Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. *The Plant Journal* 45 (4):523-539.

Wang, J., L. Tian, H.S. Lee, N.E. Wei, H. Jiang *et al.*, 2006 Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. *Genetics* 172 (1):507-517.

Wilkins, P.W., and M.O. Humphreys, 2003 Progress in breeding perennial forage grasses for temperate agriculture. *The Journal of Agricultural Science* 140 (2):129-150.

Yao, H., A. Dogra Gray, D.L. Auger, and J.A. Birchler, 2013 Genomic dosage effects on heterosis in triploid maize. *Proc Natl Acad Sci U S A* 110 (7):2665-2669.

Yeo, A.R., λ.S. Lee, P. Izard, P.J. Boursier, and T.J. Flowers, 1991 Short- and Long-Term Effects of Salinity on Leaf Growth in Rice (Oryza sativa L.). *Journal of Experimental Botany* 42 (7):881-889.

Zolla, G., Y.M. Heimer, and S. Barak, 2009 Mild salinity stimulates a stress-induced morphogenic response in Arabidopsis thaliana roots. *Journal of Experimental Botany* 61 (1):211-224.

Chapter 6. Using the EcoCrop niche model to forecast the impacts of climate change on diploid, autopolyploid, and allopolyploid wheat (*Triticum*), wheatgrass (*Agropyron*) and wildrye (*Elymus*)

Introduction

Global food demand and climate change

A key issue facing crop production in the 21st century is the effect of anthropogenic climate change on yield and quality. As well as an increase in atmospheric carbon dioxide (CO_2), climate change will affect the intensity and frequency of temperature and precipitation fluctuations (Alexander et al. 2006; Liu et al. 2009; Trenberth 2011; Chou et al. 2012). These changes in climatic conditions can influence crop growth directly (e.g. below optimal precipitation during the growing season leading to drought) or indirectly (e.g. temperature changes allowing for the proliferation of crop pests leading to biotic stress).

Food demand is expected to increase in both developed and developing countries between now and the year 2050. Most of this demand is for cereals and animal-based products, and the increase in demand is predicted to be notably larger in the developing world (Edgerton 2009; Bodirsky et al. 2015). The driving factors behind this demand are population increases, urbanisation, rising per capita income, changes in food prices, globalisation of the food industry, and changes in dietary preferences (Popkin 1993; Drewnowski and Popkin 1997; Smil 2000; Popkin 2017). To meet this demand, agricultural production could increase by land intensification (e.g. increasing inputs of fertiliser, shorter fallow periods) or land expansion (converting uncultivated land to agricultural land). An important tool to increase crop production will be the development of 'improved' crops varieties, i.e. varieties that show better yield or quality under climate change stresses, thus alleviating the need for land intensification or expansion. The predicted impact of climate change on some of our most valuable crops has highlighted the urgent need for faster breeding programs and the introduction of new germplasm (Reynolds et al. 1999; McCouch 2004; Luck et al. 2011; Tilman et al. 2011; Pingali 2012; Ray et al. 2013).

Niche shift of polyploid plants

Studies analysing the ranges of lower and higher ploidy plants and the emergence of whole genome duplication in certain lineages do so *ex post facto* through bioinformatic analysis. These analyses regularly identify possible evidence of whole genome duplication events that correlate with major climatic shifts in a certain geological epoch (Soltis et al. 2008; Kagale et al. 2014; Macqueen and Johnston 2014; Vanneste et al. 2014; Tank et al. 2015). There is some evidence to support the hypothesis that higher ploidy plants can tolerate fluctuations in climatic conditions better than lower ploidy cytotypes, leading to the establishment of polyploid

populations within or adjacent to diploid populations (Stebbins 1971; Pandit et al. 2011). However, the impracticality of performing controlled experiments over a meaningful timeframe (evolutionary scale) or distribution (biogeographic scale) has been recognised and any climate-based niche shifts between diploid-polyploid pairs attributed to genomic properties could be attributed to stochastic establishment (Pandit et al. 2011; Glennon et al. 2014; Visger et al. 2016). However, there are examples of introduced polyploids becoming invasive through phenotypic plasticity that is absent at the diploid level, e.g. spotted knapweed (*Centaurea stoebe*) in North America (Hahn et al. 2012).

To investigate the hypothesis that polyploid plant populations can establish and prosper under changing climatic conditions to a greater extent than diploid plant populations on more biologically relevant temporal and geographical scales, computer modelling software has been employed. These investigations have characterised ecological niches of largely equivalent diploid and polyploid wild plants in Northern Europe (Theodoridis et al. 2013), North America (Blaine Marchant et al. 2016; Visger et al. 2016), and a combination of both regions (Glennon et al. 2014). The conclusions of these investigations are not alike: in some cases widespread niche overlap has been found between diploid and polyploid plant populations suggesting "microclimate" effects are minimal, but occasionally polyploids occupy separate or wider ecological niches. When a neopolyploid plant establishes in an ecological niche a number of pre- and postzygotic barriers can isolate it from its diploid progenitor (Levin 1975; Husband and Sabara 2004). The current and predicted shifting weather patterns attributed to climate change presents an opportunity for novel research to further investigate the hypothesis that, for a given flora, polyploidy can confer an enhanced ability to tolerate, adapt and proliferate within a current ecological niche or perhaps enable occupation of a new niche.

Objectives

In this PhD research chapter, I investigate if the impacts of climate change on wild and domesticated grass species will be affected to different extents based on ploidy level. It was hypothesised that (1) lower ploidy plants within a species will be at a greater risk to the impacts of climate change than higher ploidy plants, and (2) domesticated species will be at a greater risk to the impacts of climate change than wild species. To do this I use computer modelling software EcoCrop. Such 'crop modelling' has previously been used to predict crop yields and growing area suitability due to climate change effects (Rötter et al. 2013; Challinor et al. 2015; Fraga et al. 2016; Newbery et al. 2016; Hummel et al. 2018).

Many important cultivated crops are present at different ploidy levels (see Introduction to this Thesis), including wheat (*Triticum*). Wheat was domesticated from wild grass species approximately 10,000 years ago and has become the most widely cultivated cereal crop in agriculture through intense commercial selection for improvement of various characteristics (Calderini et al. 1995; Bushuk 1997; Shiferaw et al. 2013). Cultivated wheat is present at the allotetraploid and allohexaploid level (**Table *11***) The most commonly cultivated species are the allotetraploid durum wheat (*T. turgidum* subsp. *durum*) for pasta and the allohexaploid bread wheat (*T. aestivum* subsp. *aestivum*). To a lesser extent the allotetraploid emmer wheat (*T. turgidum* subsp. *diccocon*) and allohexaploid spelt wheat (*T. aestivum* subsp. *spelta*) are also commercially bred and cultivated (Campbell 1997; Zaharieva et al. 2010).

Raising the "yield plateau" (Fischer 1996) and meeting future demand for wheat is limited by the narrow genetic base among cultivated wheat from a loss of variation during domestication and selection (Dubcovsky and Dvorak 2007; van de Wouw et al. 2009; Hall and Richards 2013; Shiferaw et al. 2013). Thus, there is a need to explore wild relatives of wheat as a source of novel alleles. Two such wild relatives are wheatgrass (*Agropyron*) and wildrye (*Elymus*).

Wheatgrass is a Eurasian perennial grass which has become naturalised in Canada, the United States and Mexico (Westover 1932; Potter 1948; Lorenz 1986). Some wheatgrass species are used as a forage crop or as a soil stabiliser in arid and semi-arid regions, but commercial improvement is minimal with publicly-funded breeding efforts contributing to registered varieties (Lorenz 1986; Casler 2001; Cox et al. 2002). Wheatgrass can act as weed in many agricultural settings and in the wild can out-compete native plants, reducing biodiversity (D'Antonio and Vitousek 1992; Sutter and Brigham 1998; Heidinga and Wilson 2002; Vaness and Wilson 2007). Different *Agropyron* species are present at diploid, autopolyploid and allopolyploid levels (**Table *11***) and as members of the Triticeae tribe are capable of sexual hybridisation with wheat (Sharma and Gill 1983; Dewey 1984; Friebe et al. 1992; Wu et al. 2006; Li et al. 2008; Wang and Jensen 2009). All wheatgrass species are based on the P-genome (Dewey 1964, 1974, 1975, 1976, 1983; Hsiao et al. 1989; Wang et al. 1994) with some authors suggesting previously identified wheatgrass assigned to the E-genome be moved from *Agropyron* to *Thinopyrum* (Wang 2011). As both *Agropyron* and *Thinopyrum* have a similar growth habit and are capable of hybridisation with wheat they will be considered together in this research chapter.

Wildrye is a perennial grass which is prevalent throughout the Northern hemisphere. It is considered a native species in United States and Canada with its use as a forage crop in pre-

20th Century United States documented (Lesperance et al. 1978). Presently it is often used as a soil stabiliser and occasionally as a fodder crop but has been largely replaced in commercial settings as fodder by other plants (Erickson et al. 2004). Different *Elymus* species are present at the allotetraploid and allohexaploid level (**Table *11***) and many species have been sexually hybridised with wheat *in vivo* or *in vitro* via embryo cultures (Smith 1942; Yen and Liu 1987; Lu and von Bothmer 1991). All wildrye species are based on the combination of the St-genome, believed to originate from the wild grass *Pseudogeneria*, and the H-genome from *Hordeum* (Jiang et al. 1993; Wang et al. 1994; Mason-Gamer 2001; Wang 2011).

I investigate the hypothesis that the response of equivalent lower and higher ploidy grasses to predicted fluctuations in temperature and precipitation under climate change will be different. The choice of plants in this instance allows investigation of ploidy effects in both wild and domesticated grass species. I demonstrate that the growing area of all wheat species will be considerably reduced in the year 2050 under RCP 6.0, independent of ploidy level. For wild grass species, there will be a mixture of niche expansion and contraction in the year 2050 under RCP 6.0. Ploidy level is positively associated with future suitable growing areas for wheatgrass, but negatively associated with future suitable growing areas for wildrye. The potential implications for wheat breeding are discussed.

Table 11. Wheat (*Triticum*), wheatgrass (*Agropyron*), and wildrye (*Elymus*) species analysed in this study, their associated common name, ploidy level and designated genome.

Genus	Species	Common name	Ploidy	Genome*
Triticum	*turgidum* subsp. *durum*	Durum wheat	4x	AABB
	turgidum subsp. *diccocon*	Emmer wheat	4x	AABB
	aestivum subsp. *aestivum*	Bread wheat	6x	AABBDD
	aestivum subsp. *spelta*	Spelt wheat	6x	AABBDD
Agropyron	*cristatum*	Crested wheatgrass	2x	PP
	spicatum	Bluebunch wheatgrass	2x	PP
	tauri	-	2x	PP
	mongolicum	Mongolian wheatgrass	2x	PP
	elongatum	Tall wheatgrass	2x	EE
	desertorum	Standard crested wheatgrass	4x	PPPP
	fragile	Siberian wheatgrass	4x	PPPP
	dasystachyum	Thickspike wheatgrass	4x	PPHH
	trachycaulum	Slender wheatgrass	4x	PPHH
	riparium	Streambank wheatgrass	4x	P_1P_1XX
	repens	Couch grass	6x	$P_1P_1P_2P_2XX$
	intermedium	Intermediate wheatgrass	6x	$E_1E_1E_2E_2XX$
	trichophorum	-	6x	$E_1E_1E_2E_2XX$
	smithii	Western wheatgrass	8x	PPHHEEXX
Elymus	*glaucus*	Blue wildrye	4x	StStHH
	canadensis	Canada wildrye	4x	StStHH
	dahuricus	-	6x	StStHHYY

*Genome designation follows Dewey (1964, 1974, 1975, 1976, 1983, 1984); Hsiao et al. (1989); Jiang et al. (1993); Wang et al. (1994); Mason-Gamer (2001); Dubcovsky and Dvorak (2007); Wang (2011) where 'X' indicates presently undetermined.

Materials and Methods

The EcoCrop model and Geographic Information System to determine growing area suitability

To determine growing area suitability at a meaningful scale, global crop production and trading statistics from the years 1994 - 2016 of wheat from the Food and Agricultural Organisation (FAO) of the United Nations was cross-referenced online (FAO 2018), as well as a description of agricultural practices from the International Maize and Wheat Improvement Center (CIMMYT) (Shiferaw et al. 2013). This revealed that (1) the single land mass comprising of Canada, United States, and Mexico is the second-largest wheat producer and the largest wheat exporter in the world, and (2) this is achieved primarily through rainfed agriculture. It was decided that growing area suitability of wheat and associated wild wheatgrass and wildrye should be determined for Canada, United States and Mexico.

The Geographic Information System (GIS) software 'DIVA-GIS' version 7.5 was downloaded from http://www.diva-gis.org/ and used to map spatial data. A detailed description of how the EcoCrop model works has been described elsewhere (Hijmans et al. 2001; Ramirez-Villegas et al. 2013). In brief, using the FAO-EcoCrop database of environmental requirements for different plant species (FAO 2007), a suitability index for a given plant over a geographic area is determined by temperature and precipitation ranges. The temperature parameters used are: absolute temperature that will kill the plant, minimum temperature at which the plant can grow, optimum temperature at which the plant can grow, maximum temperature at which the plant will grow, and maximum temperature at which the plant ceases growth. The precipitation ranges used are: minimum precipitation during the growing season, optimal minimum precipitation during the growing season, optimal maximum precipitation during the growing season, and maximum precipitation during the growing season. Historical and projected climate data for a region is available at http://www.worldclim.org/version1 (Hijmans et al. 2005). Thus, for a specific plant, the climatic suitability of a region is determined from the interaction of temperature and precipitation. For my analysis, current suitability simulations are based on the historical period 1960-1990 (baseline) and future suitability simulations are based on the year 2050 under Representative Concentration Pathway (RCP) 6.0. Simulations are carried out by EcoCrop using 19 Global Circulation Models (GCMs) with appropriate downscaling and bias-correction included (Ramirez-Villegas and Jarvis 2010). A plant is assumed to be not viable when suitability index falls below 43% (Rippke et al. 2016; Hummel et al. 2018).

Statistical Analysis

Climate data and growing area suitability data under baseline (i.e. current) and future scenarios were mapped to a defined area. A suitability index for a plant species in a cell is calculated by

EcoCrop. Next, all grid cell values containing climate data and growing area suitability data were extracted for each simulation (current and future) as a text file (.txt) and organised in Microsoft Excel (Microsoft, WA, United States) (**Figure 22**). To investigate changes in growing areas, the change in suitability (future-current) was calculated for currently 'suitable' growing areas (cells where growing area suitability ≥ 43%) for all wheat, wheatgrass and wildrye species respectively. To investigate the potential to exploit new land, the change in suitability (future-current) was calculated for all currently 'unsuitable' growing areas (cells where growing area suitability < 43%) for all wheat, wheatgrass and wildrye species respectively. A one-way ANOVA followed by Tukey's test was employed to determine any statistically significant differences between species for (1) average change in current suitable growing areas, and (2) average change in current unsuitable growing areas. The term "niche contraction" was defined as total growing area suitability decreasing under climate change, while "niche expansion" was defined as total growing area suitability increasing under climate change. To investigate an association between ploidy level and predicted future 'suitable' growing areas (cells where growing area suitability ≥ 43%) for all wheat, wheatgrass and wildrye species respectively, a Pearson's correlation analysis was employed. Statistical analysis was performed on computer software SPSS v. 24 (IBM, NY, USA).

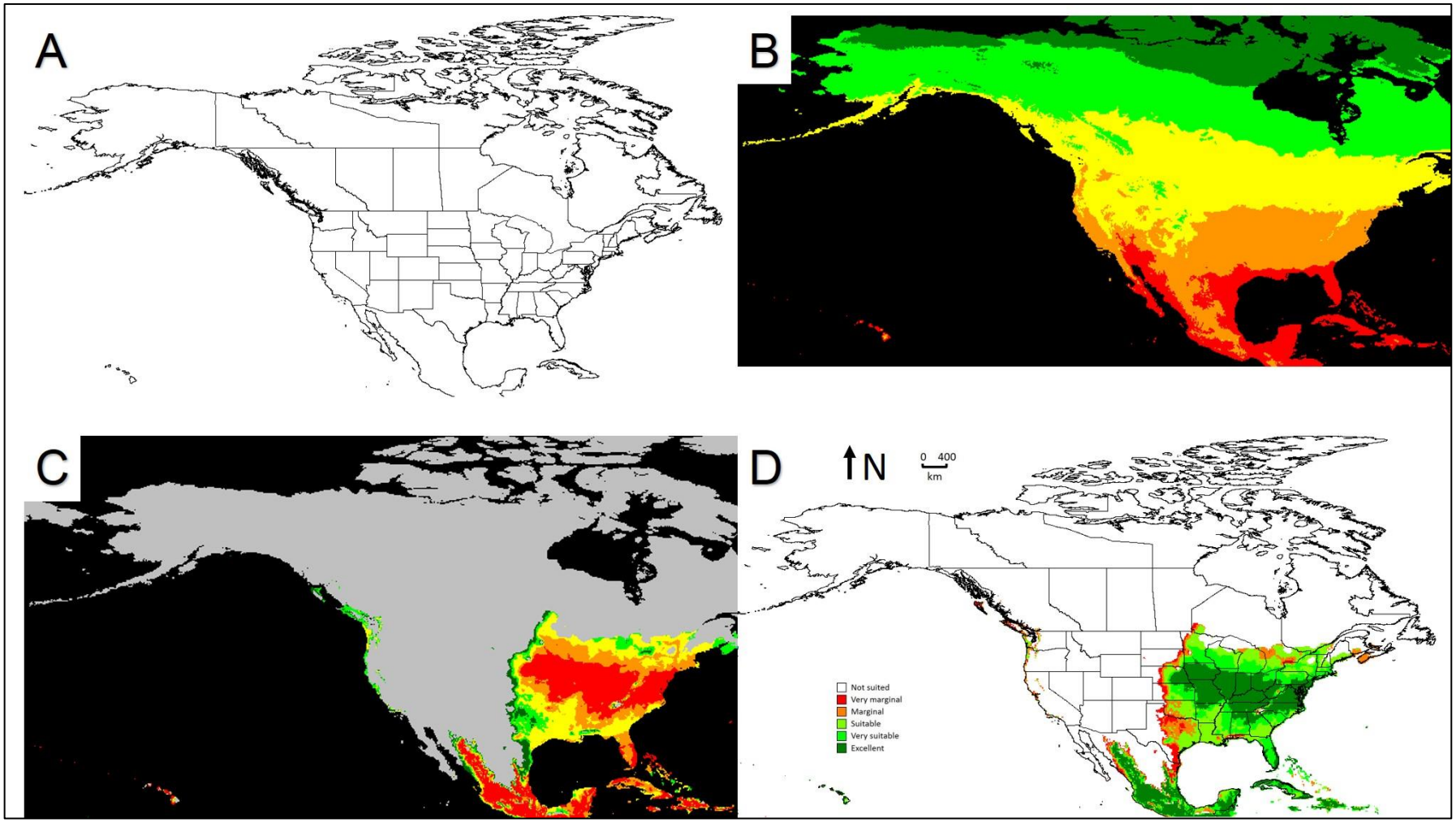


Figure 22. Steps taken for modelling and extracting data from EcoCrop model. A Global boundary data defined for Canada, United States and Mexico in DIVA-GIS. **B** Minimum, maximum and average temperature and precipitation data at each grid cell is defined for baseline (i.e. current) and future years. **C** For a given plant, suitability of land area is calculated for baseline and future years independently. **D** End result is map with growing area suitability calculated for current or future simulation, and categories defined from 'Not suited' to 'Excellent' at each grid cell.

Results

Canada, United States and Mexico will experience an increase in average temperature and an increase in precipitation fluctuations

For each grid cell the minimum, maximum and average temperature and precipitation was determined for current (baseline) and future (2050, RCP 6.0) years. The predicted climate data for the year 2050 in these regions under RCP 6.0 as determined by the EcoCrop model is an increase in average temperature by 2.33 °C and an increase in average precipitation by 42.48 mm accompanied by increased precipitation fluctuations (**Figure 23**).

All wheat species will undergo a niche contraction under RCP 6.0

Large areas across the United States and Mexico presently include regions from 'Excellent' suitability to 'Marginal' suitability for wheat. Only south-eastern regions of Canada presently include regions suitable for wheat cultivation (Supplementary Results). Under future climate change scenario RCP 6.0, growing areas currently designated as suitable for all wheat species will experience an average decrease. The largest decrease will be 16% for both the allotetraploid durum wheat (*T. turgidum* subsp. *diccocon*) and the allohexaploid spelt wheat (*T. aestivum* subsp. *spelta*) (**Figure *24***, closed bars). Growing areas currently designated as unsuitable for all wheat species will experience an average increase. The largest increase will be for the allotetraploid durum wheat (*T. turgidum* subsp. *diccocon*) at 6% and smallest for the allohexaploid bread wheat (*T. aestivum* subsp. *aestivum*) (**Figure *24***, open bars). In total, all wheat species are predicted to undergo a niche contraction (**Table *12***).

Five wheatgrass species possessing the P-genome will undergo a niche contraction under RCP 6.0, while five wheatgrass species will undergo a niche expansion and one species will not change

Most wheatgrass species possessing the P-genome can currently be grown across large areas of Canada, United States and Mexico. *A. mongolicum* is particularly suited to all countries, while in contrast *A. tauri* is only suitable to central Mexico, south, south-east and the west coast of the United States (Supplementary Results). Under future climate change scenario RCP 6.0, growing areas currently designated as suitable for many wheatgrass species possessing the P-genome will experience an average decrease. The largest decrease will be for the autodiploid *A. tauri* at 17%. The smallest decrease will be for the allotetraploid *A. riparium* at 3%. Growing areas currently designated as suitable for five wheatgrass species possessing the P-genome will

experience an average increase. The largest increase will be for the allohexaploid *A. repens* at 4%, followed by a 3% increase each for the allooctaploid *A. smithii* and the allotetraploid *A. dasystachyum*, and the smallest increase will be for the autodiploid *A. mongolicum* at 0.5% (**Figure *25***, closed bars). Growing areas currently designated as unsuitable for all wheatgrass species possessing the P-genome will experience an average increase. The largest increase will be for the autodiploid *A. mongolicum* at 6% and the smallest for the autodiploid *A. tauri* at 2% (**Figure *25***, open bars). In total, five wheatgrass species possessing the P-genome are predicted to undergo a niche contraction, while five species will undergo a niche expansion and one species will not change (**Table *13***).

Two wheatgrass species possessing the E-genome will undergo a niche contraction under RCP 6.0, while two wheatgrass species will undergo a niche expansion

Most wheatgrass species possessing the E-genome can currently be grown across large areas of United States and Mexico. *A. elongatum, A. intermedium* and *A. smithii* can currently be grown across large areas of United States and Mexico while *A. smithii* can additionally be grown across large areas of Canada. *A. trichophorum* is only suitable to central Mexico, south, south-east and the west coast of United States (Supplementary Results). Under future climate change scenario RCP 6.0, growing areas currently designated as suitable for three wheatgrass species possessing the E-genome will decrease. The largest decrease will be for the allohexaploid *A. trichophorum* at 16%, followed by the autodiploid *A. elongatum* at 12%, and allohexaploid *A. intermedium* at 1.5%. Growing areas currently designated as suitable for one wheatgrass species possessing the E-genome will increase. The allooctaploid *A. smithii* suitability will increase by 3% (**Figure *26***, closed bars). Growing areas currently designated as unsuitable for all wheatgrass species possessing the E-genome will experience an average increase. The largest increase will be for the allohexaploid *A. intermedium* at 5% and the smallest increase will be for the allohexaploid *A. trichophorum* at 2% (**Figure *26***, open bars). In total, two wheatgrass species possessing the E-genome are predicted to undergo a niche contraction, while two will undergo a niche expansion (**Table *14***).

Two wildrye species will undergo a niche contraction under RCP 6.0, while one wildrye species will undergo a niche expansion

Wildrye species can currently be grown across different areas of Canada, United States and Mexico. *E. canadensis* can currently be grown across large areas of east Canada, United States and Mexico. *E. glaucus* can currently be grown across east Canada, parts of south United States

and the west coast. *E. dahuricus* can currently be grown across large areas of United States and Mexico (Supplementary Results). Under future climate change scenario RCP 6.0, growing areas currently designated as suitable for two wildrye species will experience an average decrease. The largest decrease will be for the allotetraploid *E. glaucus* at 12%, followed by the allohexaploid *E. dahuricus* at 9%. Growing areas currently designated as suitable for one wildrye species will experience an average increase. The allotetraploid *E. canadensis* suitability will increase by 1.5% (**Figure *27***, closed bars). Growing areas currently designated as unsuitable for all wildrye species will experience an average increase. The largest increase will be for the allotetraploid *E. canadensis* at 3% and the smallest increase for *E. glaucus* at 1% (**Figure *27***, open bars). In total, two wildrye species are predicted to undergo a niche contraction, while one species will undergo an expansion (**Table *15***).

The correlation between ploidy level and future suitable growing area is species-specific

There is potential for all grass species in this study to exploit new growing areas by the year 2050 under RCP 6.0 (**Figure *24***, Figure 25, Figure 26, and Figure 27, open bars). Additionally, some species will experience an increase in average suitability on currently suitable growing areas, such as the autodiploid *A. mongolicum* and the allotetraploid *A. dasystachyum* (**Figure *25***, closed bars). The niche change is a relative statistic; for a given species it is the aggregate of change in currently suitable and currently unsuitable growing areas (Table ***12***, Table ***13***, Table ***14***, and Table ***15***). This does not determine whether the entire future suitable growing areas of Canada, United States and Mexico are likely to be associated with lower or higher ploidy grass species, so a correlation analysis was performed to determine if all future suitable growing areas (cells where growing area suitability ≥ 43%) are positively or negatively associated with ploidy level.

For wheatgrass (*Agropyron*), there is a positive association between temperature (r = 0.2; $P < 0.001$), precipitation (r = 0.1; $P < 0.001$) and ploidy level (r = 0.02; $P < 0.001$) with future suitable growing area. For wildrye (*Elymus*), there is a positive association between temperature (r = 0.2; $P < 0.001$) and precipitation (r = 0.2; $P < 0.001$) with future suitable growing area, but ploidy level is negatively associated (r = -0.04; $P < 0.001$) with future suitable growing area. For wheat (*Triticum*), there is a positive association between temperature (r = 0.2; $P < 0.001$) and precipitation (r = 0.3; $P < 0.001$) with future suitable growing area, but ploidy level is negatively associated (r = -0.1; $P < 0.001$) with future suitable growing area (Table ***16***). Therefore, the correlations between future climatic conditions and future suitable growing areas are consistent across both wild and domesticated grass species, whereas the

correlations between ploidy level and future suitable growing areas are species-specific. Higher ploidy wheatgrass (*Agropyron*) and lower ploidy wildrye (*Elymus*) and wheat (*Triticum*) are associated with future suitable growing areas across Canada, United States, and Mexico.

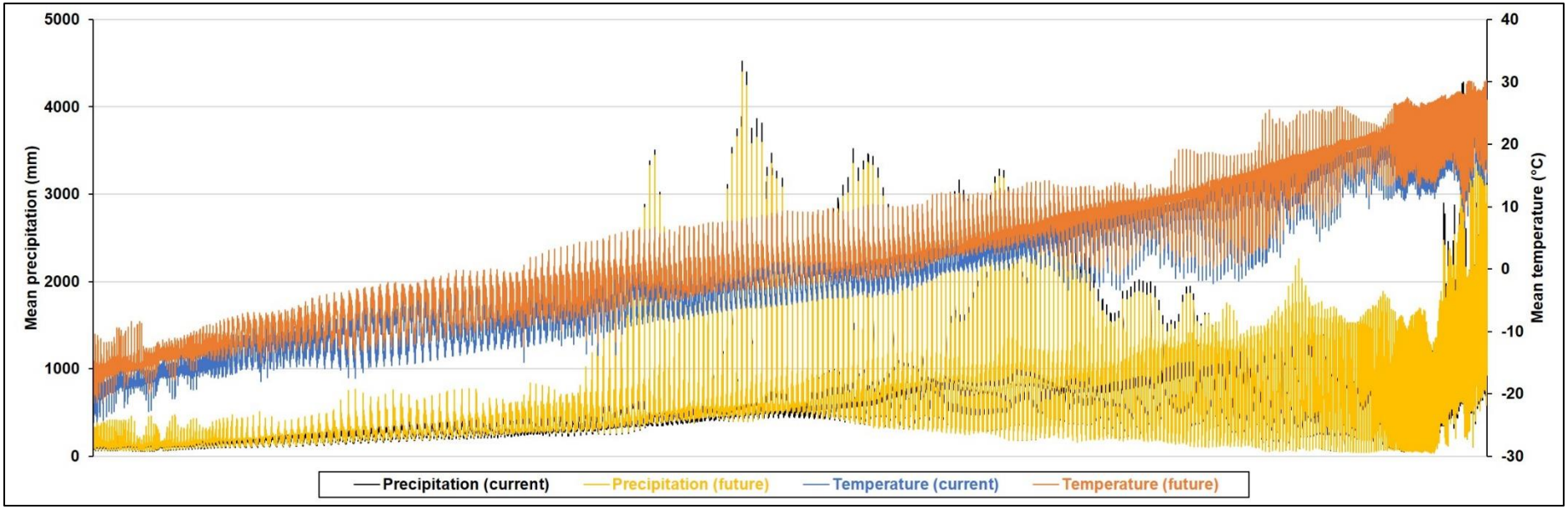


Figure 23. Extracted climate data from EcoCrop model for Canada, United States, and Mexico reveals higher mean temperatures and greater precipitation fluctuations in the year 2050 under RCP6.0. For each grid cell, starting with Canada and moving south, current mean temperature (blue), future mean temperature (orange), current mean precipitation (black), and future mean precipitation (yellow) are displayed.

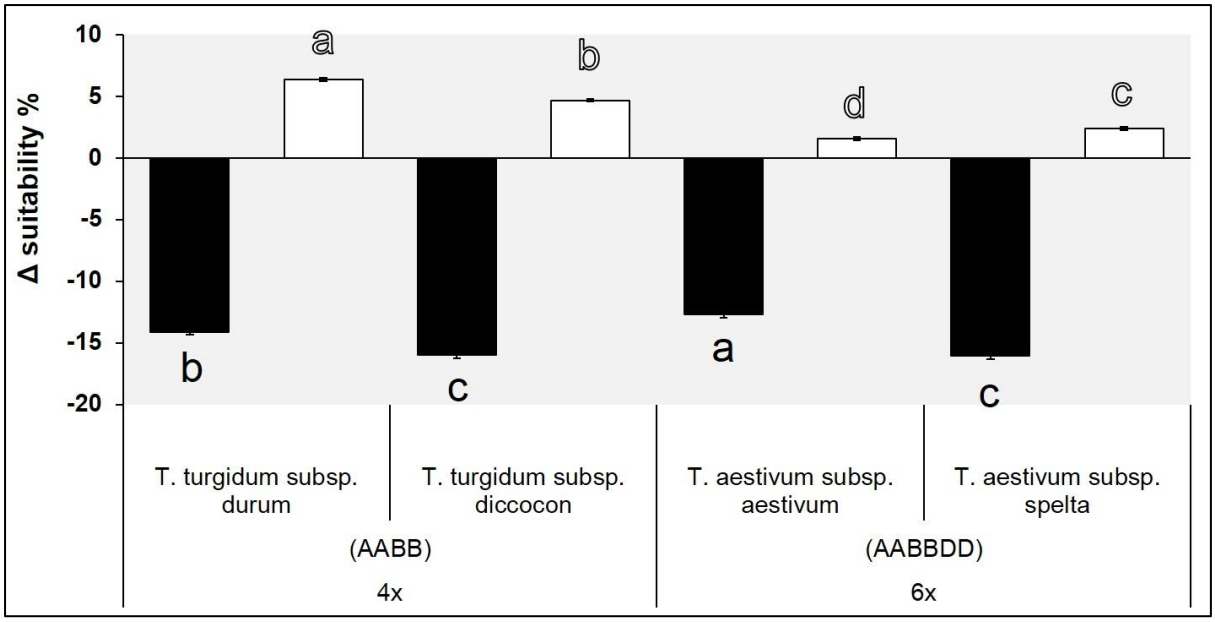


Figure 24. Mean change in current suitable growing areas and current unsuitable growing areas of allotetraploid and allohexaploid wheat (*Triticum*) species under RCP 6.0. The difference between current suitable growing areas and the same areas in the year 2050 under RCP 6.0 was calculated (closed bars, ± SE). The difference between current unsuitable growing areas and the same areas in the year 2050 under RCP 6.0 was calculated (open bars, ± SE). One-way ANOVA calculated for change in current suitable growing areas and current unsuitable growing areas independently. Means assigned different letters are significantly different ($P < 0.05$) according to Tukey's test.

Table 12. Summary of niche change of allotetraploid and allohexaploid wheat (*Triticum*) species under RCP 6.0.

Species	Ploidy	Genome	Total Δ suitability	Niche change
turgidum subsp. *durum*	4x	AABB	-8%	Contraction
turgidum subsp. *diccocon*			-11%	Contraction
aestivum subsp. *aestivum*	6x	AABBDD	-11%	Contraction
aestivum subsp. *spelta*			-14%	Contraction

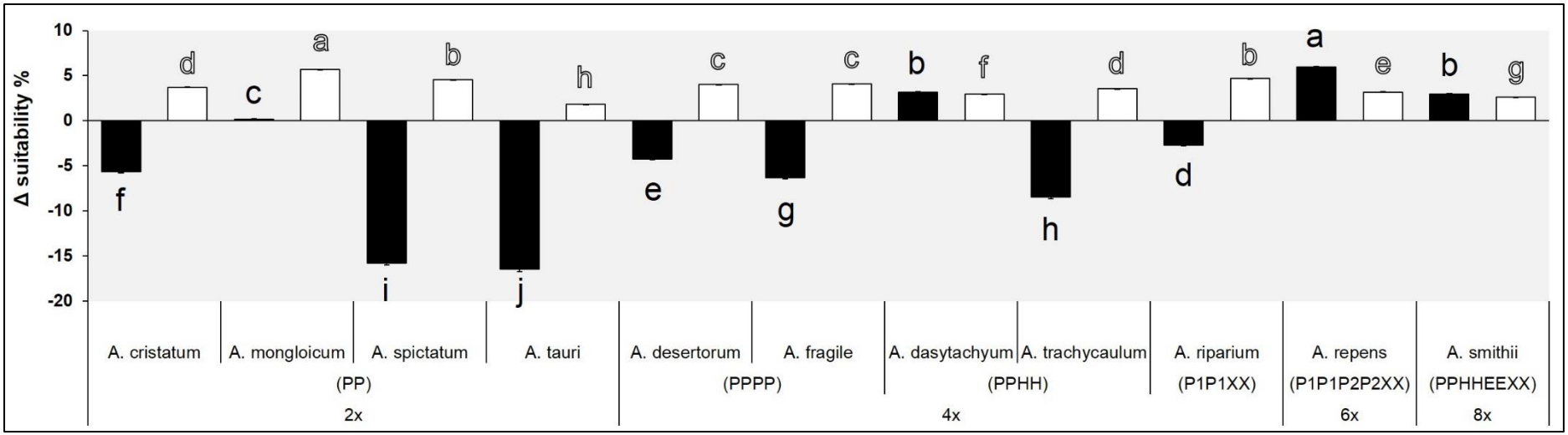


Figure 25. Mean change in current suitable growing areas and current unsuitable growing areas of autodiploid, autotetraploid, allotetraploid, allohexaploid, and allooctaploid wheatgrass (*Agropyron*) species possessing the P-genome under RCP 6.0. The difference between current suitable growing areas and the same areas in the year 2050 under RCP 6.0 was calculated (closed bars, ± SE). The difference between current unsuitable growing areas and the same areas in the year 2050 under RCP 6.0 was calculated (open bars, ± SE). One-way ANOVA calculated for change in current suitable growing areas and current unsuitable growing areas independently. Means assigned different letters are significantly different ($P < 0.05$) according to Tukey's test.

Table 13. Summary of niche change of autodiploid, autotetraploid, allotetraploid, allohexaploid, and allooctaploid wheatgrass (*Agropyron*) species possessing the P-genome under RCP 6.0.

Species	Ploidy	Genome	Total Δ suitability	Niche change
cristatum	2x	PP	-2%	Contraction
spicatum			-10%	Contraction
tauri			-15%	Contraction
mongolicum			+7%	Expansion
desertorum	4x	PPPP	0%	No change
fragile			-2%	Contraction
dasystachyum		PPHH	+6%	Expansion
trachycaulum			-5%	Contraction
riparium		P_1P_1XX	+2%	Expansion
repens	6x	$P_1P_1P_2P_2XX$	+7%	Expansion
smithii	8x	PPHHEEXX	+6%	Expansion

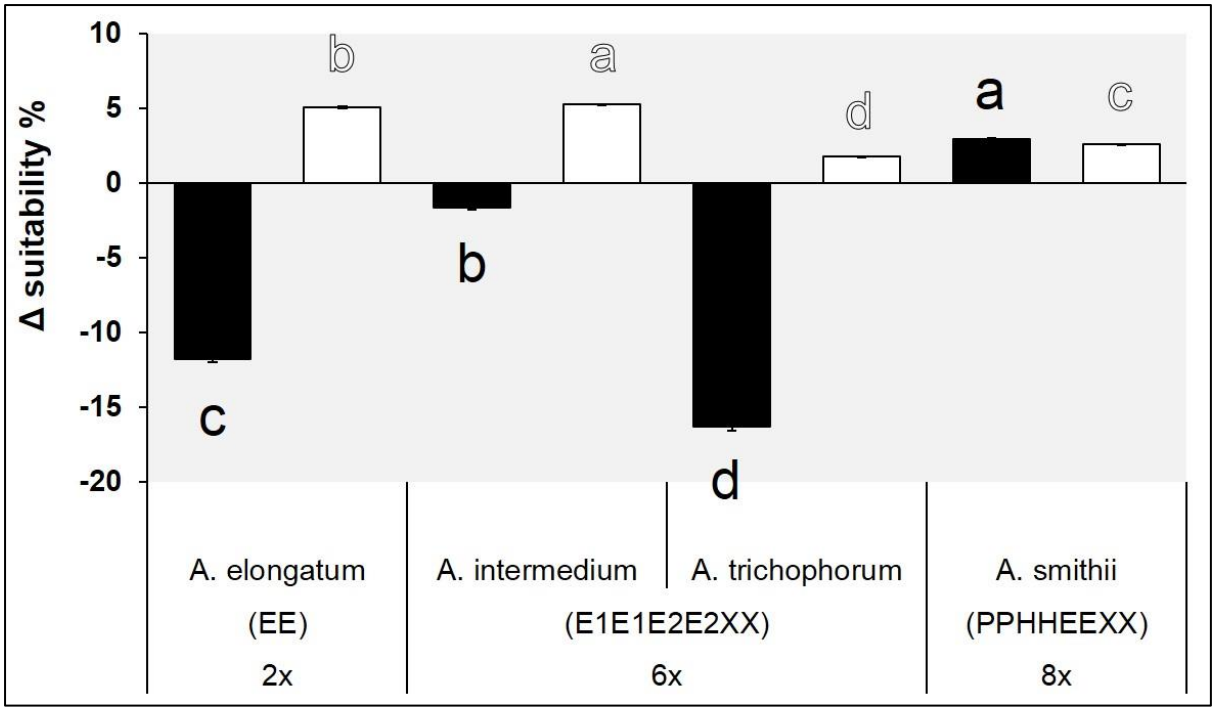


Figure 26. Mean change in current suitable growing areas and current unsuitable growing areas of autodiploid, allohexaploid and allooctaploid wheatgrass (*Agropyron*) species possessing the E-genome under RCP 6.0. The difference between current suitable growing areas and the same areas in the year 2050 under RCP 6.0 was calculated (closed bars, ± SE). The difference between current unsuitable growing areas and the same areas in the year 2050 under RCP 6.0 was calculated (open bars, ± SE). One-way ANOVA calculated for change in current suitable growing areas and current unsuitable growing areas independently. Means assigned different letters are significantly different ($P < 0.05$) according to Tukey's test.

Table 14. Summary of niche change of autodiploid, allohexaploid and allooctaploid wheatgrass (*Agropyron*) species possessing the E-genome under RCP 6.0.

Species	Ploidy	Genome	Total Δ suitability	Niche change
elongatum	2x	EE	-7%	Contraction
intermedium	6x	$E^1E^1E^2E^2XX$	+3%	Expansion
trichophorum			-14%	Contraction
smithii	8x	PPHHEEXX	+6%	Expansion

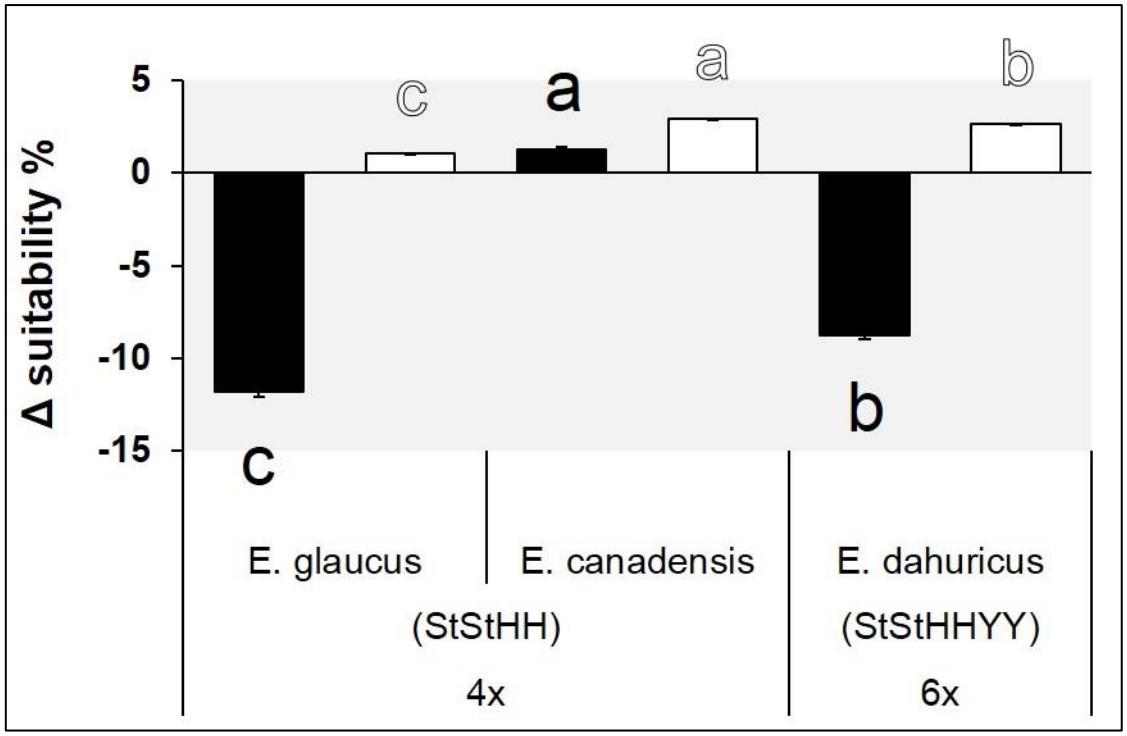


Figure 27. Mean change in current suitable growing areas and current unsuitable growing areas for allotetraploid and allohexaploid wildrye (*Elymus*) species under RCP 6.0. The difference between current suitable growing areas and the same areas in the year 2050 under RCP 6.0 was calculated (closed bars, ± SE). The difference between current unsuitable growing areas and the same areas in the year 2050 under RCP 6.0 was calculated (open bars, ± SE). One-way ANOVA calculated for change in current suitable growing areas and current unsuitable growing areas independently. Means assigned different letters are significantly different ($P < 0.05$) according to Tukey's test.

Table 15. Summary of niche change of allotetraploid and allohexaploid wildrye (*Elymus*) species under RCP 6.0.

Species	Ploidy	Genome	Total Δ suitability	Niche change
glaucus	4x	StStHH	-11%	Contraction
canadensis			+4.5%	Expansion
dahuricus	6x	StStHHYY	-6%	Contraction

Table 16. Association between future suitable growing areas, future mean temperature and precipitation, and ploidy level for wild and cultivated grass species.

		Future suitable growing areas		
		Wheatgrass	**Wildrye**	**Wheat**
Temperature	Pearson Correlation	0.236	0.208	0.199
	Significance	$P < 0.001$	$P < 0.001$	$P < 0.001$
	N	435813	55233	47344
Precipitation	Pearson Correlation	0.111	0.210	0.030
	Significance	$P < 0.001$	$P < 0.001$	$P < 0.001$
	N	435813	55233	47344
Ploidy	Pearson Correlation	0.022	-0.044	-0.143
	Significance	$P < 0.001$	$P < 0.001$	$P < 0.001$
	N	435813	55233	47344

Discussion

The capacity of grass genera to exploit future temperature and precipitation fluctuations is not always associated with higher levels of ploidy and hybridity

Polyploid plants are abundant in nature and this partially accounts for their high representation among domesticated crops, as well as plant breeders harnessing the "gigas" effect and the higher multi-locus heterozygosity possible with polyploid plants (Sattler et al. 2016). Indeed, many of our most important crop plants are (evolutionarily recent) allopolyploids, and those that are diploid have undergone at least one round of polyploidisation in their lineage (Renny-Byfield and Wendel 2014). The success of polyploid plants in nature has been attributed to their relatively large size and the potential for novel gene expression. In addition, field sampling exercises and computer modelling investigations have further suggested that polyploid plant populations are more likely than diploid plant populations to establish in disturbed or extreme climatic niches, but determining the relative effects of stochastic establishment *versus* niche divergence from evolutionary pressures is difficult (Soltis et al. 1995; Ramsey 2011; Blaine Marchant et al. 2016; Visger et al. 2016).

My results indicate that there will be differences in growing area suitability among related species in response to climate change (**Figure *24***, **Figure *25***, **Figure *26***, and **Figure *27***). These related grass species can presently be grown across the same geographic area (Supplementary Results). Some species have the potential to expand their current niche, such as the autodiploid *A. mongolicum* and the allohexaploid *A. repens (***Table *13***), while other species are at risk of niche contraction, such as the autodiploid *A. tauri* or allotetraploid *E. glaucus* (**Table *13*** and **Table *15***). This is an agreement with other research that finds widespread niche overlap between lower and higher ploidy plants (Glennon et al. 2014; Blaine Marchant et al. 2016).

The link between ploidy level and all future suitable growing areas across Canada, United States, and Mexico is inconsistent across genera (**Table *16***). Future growing area suitability in wild *Elymus* is negatively correlated with ploidy level. This suggests that the allohexaploid *E. dahuricus* could be out-competed by the allotetraploid *E. canadensis*. Other investigations using georeferencing data have reported similar results, such as examples where higher ploidy primrose (*Primula* sect. *Aleurita*) in Europe occupies a narrower geographic space than its lower ploidy relative, which is considerably more abundant (Theodoridis et al. 2013). A related, interesting finding comes from the Mediterranean shrub *Santolina pectinata*. Autotetraploid cytotypes are notably less abundant than diploid cytotypes, yet autotetraploids

tend to occupy more ecologically disturbed areas (Rivero-Guerra 2008). In contrast, other investigations have found that higher ploidy ferns (*Asplenium*), *Spartina* grass, and birch trees (*Betula*) have colonised a considerably larger range than their lower ploidy equivalents – although such colonisation success has been at least partially attributed to differences in mating habit between cytotypes rather than solely due to an increase in phenotypic plasticity (Vogel et al. 1999; Ainouche et al. 2004; Zohren et al. 2016). Future growing area suitability in wild *Agropyron* is positively correlated with ploidy level. This suggests that while autodiploid *A. mongolicum* can potentially expand its current niche (**Table *13***), it will be in competition with a greater number of higher ploidy *Agropyron* species assuming they compete for the same ecological niche. As mentioned, *Agropyron* is a recognised invasive plant across the geographic area of this study, and Pandit et al. (2011) conclude that polyploidy is positively correlated with invasiveness potential. The direct causes behind these correlations cannot be fully described by this analysis, as the question remains, does polyploidy beget abundance or does abundance beget polyploidy (Kunin 1997)? The influence of the climate in mixed-ploidy plant species is perhaps best investigated at "hybrid zones" or "contact zones" (Petit et al. 1999; Mráz et al. 2012; Kolář et al. 2017; Čertner et al. 2018; Gomes et al. 2018), which have highlighted species- and context-dependent instances where diploid and polyploid plants respond differently to microclimates, including examples of intermediate cytotypes. Considering the present and future impacts of climate change to natural ecosystems, these microclimate effects could amplify niche shifts among mixed-ploidy plant species at contact zones.

An ideal crop variety will maintain average suitability under climate change without land use changes

The regions chosen for this analysis in the global west are the leading exporter of wheat for the world food system. The predicted climate for the year 2050 in these regions (under RCP 6.0) is an increase in average temperature and an increase in precipitation fluctuations (**Figure 23**). Previous research investigating wheat responses to heat stress have observed reduced fertility (Saini et al. 1983, 1984) and genotypes moving more rapidly through growth stages (Reynolds et al. 1994; Mohammed et al. 2014), all of which contribute to lower yield. The future growing area suitability for wheat, as predicted by the EcoCrop model, is for a general decrease (**Table *12***), suggesting that expansion into all available land cannot maintain current average suitability levels.

The EcoCrop simulations calculate growing area suitability on all land. However, for calculations of growing area suitability to have practical relevance it is necessary to determine where land use changes are feasible (Lambin et al. 2013; Johnson et al. 2014). For example, the suitable growing area for wheat in the United States will experience a shift either directly northward (emmer, spelt) or north-eastward (durum, bread) (Supplementary Results). To maintain current acreage of wheat may not be possible due to land use conflicts (e.g. forestry, other crops, protected land, domestic or commercial use). To minimise these conflicts, therefore, an ideal crop variety is one that can maintain suitability on current land. No *Triticum* species meets this standard. Indeed, higher levels of ploidy and hybridity in *Triticum* are negatively associated with future growing area suitability (**Table *12***). Therefore, research into climate change adaptation measures for wheat should be an immediate priority.

Plant abiotic stresses linked to anthropogenic climate change are predicted to be concurrent, i.e. changes in temperature and precipitation are likely to occur simultaneously and these changes could have a synergistic or antagonist relationship (Mittler 2006; Mickelbart et al. 2015). Therefore, it will be necessary to undertake climate change adaptations that are context-specific. The feasibility of adaptation must consider location, crop species in question, and available resources. Climate change adaptations for crop production can be biological (e.g. breeding for improved varieties) or physical (e.g. changing fertiliser or irrigation practices).

Breeding for improved wheat varieties within the *Triticum* genus is limited by the narrow genetic basis among domesticated wheat and the fact that the number of genotypes available within a breeding programme is finite and is itself a product of all the selection that has come before (Able et al. 2007). Yet, breeding for improved wheat varieties through wide crosses with other members of the Triticeae tribe is possible (Smith 1942; Sharma and Gill 1983; Dewey 1984; Yen and Liu 1987; Lu and von Bothmer 1991; Friebe et al. 1992; Sharma 1995; Wu et al. 2006; Li et al. 2008; Wang and Jensen 2009). Considerable success has been achieved with one of the wild ancestors of wheat, goatgrass (*Aegilops*), where many abiotic and biotic stress resistance genes have been successfully transferred to wheat (Schneider et al. 2008). My crop modelling analysis suggests that the wild grass species *A. mongolicum*, *A. dasystachyum*, *A. repens*, *A. smithii*, and *E. canadensis* are worthy candidates for further investigation as these species show niche expansion under climate change. Validation under experimental conditions (in growth chambers or field analogue sites) could further identify novel phenotypes and their genetic control. In crop wild relatives, however, strong abiotic stress tolerance often manifests as a trade-off with other agricultural traits of interest such as shattering resistance or yield (Des Marais and Juenger 2010; Koehler et al. 2012), meaning

targeted selection of molecular markers and gene transfer with the latest breeding technologies will likely be necessary to minimise the length of breeding programs (Zamir 2001; Tester and Langridge 2010; Puchta 2017).

Limitations of the EcoCrop model

Investigating niche shifts of a species is primarily carried out over a large distributional range or over a long time (Pearman et al. 2008). To anticipate niche shifts at a meaningful scale for any location requires that the species' complete current geographic range is included as well as any potential new ranges that could be reached via reasonable assumptions of dispersal capacity (Barve et al. 2011). Both requirements are met in the current investigation. However, while growing area suitability calculations of the EcoCrop model are accurate for the combined, direct effect of temperature and precipitation, the model cannot process elevated atmospheric CO_2 levels, increased frequency of extreme climatic events, and changes in crop pests and diseases, all of which are predicted to occur by the year 2050 under RCP 6.0 (IPCC 2014).

Furthermore, the climatic suitability parameters used by EcoCrop to calculate growing area suitability may need to be manually adjusted in some instances. For example, while the southeast of Canada is correctly identified as a suitable growing region for wheat by EcoCrop, the south of the country is currently designated as "unsuitable" (Supplementary Results). This is in stark contrast to common knowledge that the south of Canada is a vibrant wheat growing region. Cross-referencing agricultural economic output data with the EcoCrop growing area suitability calculations will improve the accuracy of the model.

Conclusion

For the regions and species under investigation, it is shown that (1) lower ploidy plants within a species are not always at a greater risk to the impacts of climate change simply due to their low number of chromosome sets, and, (2) some wild grass species are at an equivalent risk to the impacts of climate change to that of domesticated grass species. Furthermore, I demonstrate that ploidy level is positively associated with future suitable growing areas for wild wheatgrass, but negatively associated with future suitable growing areas for wildrye and domesticated wheat. Research has shown that variation in seasonal temperature and precipitation have negatively impacted cereal yields since the 1980s, and this negative effect is projected to continue depending on adaptation measures (Lobell and Field 2007). Thus, breeding for improved varieties that can withstand these temperature and precipitation changes is a necessity

if we are to meet current and future food demands. The EcoCrop model is an excellent starting point to identify regions and plant species that could form part of potential climate change adaptation measures.

Future Directions

This research chapter sought to test the hypothesis that the impacts of climate change on wild and domesticated grass species will be affected to different extents based on ploidy level. To further test the hypothesis and prepare this research chapter as a novel publication requires the inclusion of additional plant species. The EcoCrop database has 2,568 plant species which could be further interrogated to identify additional lower and higher ploidy pairs. Unfortunately, there is no categorisation via ploidy level so whole genera will have to be searched for suitable plant species. In addition, while the EcoCrop model can only predict the future climate change for a given region in the year 2050 under RCP 6.0, specific climatic changes could be identified and categorised for analysis (e.g. regions facing rising temperature and falling precipitation, rising temperature and rising precipitation, or rising temperature and uniform precipitation).

References

Able, J.A., P. Langridge, and A.S. Milligan, 2007 Capturing diversity in the cereals: many options but little promiscuity. *Trends in Plant Science* 12 (2):71-79.

Ainouche, M.L., A. Baumel, A. Salmon, and G. Yannic, 2004 Hybridization, polyploidy and speciation in Spartina (Poaceae). *New Phytologist* 161 (1):165-172.

Alexander, L.V., X. Zhang, T.C. Peterson, J. Caesar, B. Gleason *et al.*, 2006 Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres* 111 (D5).

Barve, N., V. Barve, A. Jiménez-Valverde, A. Lira-Noriega, S.P. Maher *et al.*, 2011 The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling* 222 (11):1810-1819.

Blaine Marchant, D., D.E. Soltis, and P.S. Soltis, 2016 Patterns of abiotic niche shifts in allopolyploids relative to their progenitors. *New Phytologist* 212 (3):708-718.

Bodirsky, B.L., S. Rolinski, A. Biewald, I. Weindl, A. Popp *et al.*, 2015 Global Food Demand Scenarios for the 21st Century. *PloS one* 10 (11):e0139201.

Bushuk, W., 1997 Wheat breeding for end-product use, pp. 203-211 in *Wheat: Prospects for Global Improvement: Proceedings of the 5th International Wheat Conference, 10–14 June, 1996, Ankara, Turkey*, edited by H.J. Braun, F. Altay, W.E. Kronstad, S.P.S. Beniwal and A. McNab. Springer Netherlands, Dordrecht.

Calderini, D.F., M.F. Dreccer, and G.A. Slafer, 1995 Genetic improvement in wheat yield and associated traits. A re-examination of previous results and the latest trends. *Plant Breeding* 114 (2):108-112.

Campbell, K.G., 1997 Spelt: agronomy, genetics, and breeding in *Plant Breeding Reviews*, edited by J. Janick. John Wiley & Sons, New York, NY.

Casler, M.D., 2001 Breeding forage crops for increased nutritional value, pp. 51-107 in *Advances in agronomy*. Academic Press.

Čertner, M., R. Sudová, M. Weiser, J. Suda, and F. Kolář, 2018 Ploidy-altered phenotype interacts with local environment and may enhance polyploid establishment in Knautia serpentinicola (Caprifoliaceae). *New Phytologist* 0 (0).

Challinor, A.J., B. Parkes, and J. Ramirez-Villegas, 2015 Crop yield response to climate change varies with cropping intensity. *Global Change Biology* 21 (4):1679-1688.

Chou, C., C.-A. Chen, P.-H. Tan, and K.T. Chen, 2012 Mechanisms for Global Warming Impacts on Precipitation Frequency and Intensity. *Journal of Climate* 25 (9):3291-3306.

Cox, T.S., M. Bender, C. Picone, D.L.V. Tassel, J.B. Holland *et al.*, 2002 Breeding Perennial Grain Crops. *Critical Reviews in Plant Sciences* 21 (2):59-91.

D'Antonio, C.M., and P.M. Vitousek, 1992 Biological Invasions by Exotic Grasses, the Grass/Fire Cycle, and Global Change. *Annual Review of Ecology and Systematics* 23:63-87.

Des Marais, D.L., and T.E. Juenger, 2010 Pleiotropy, plasticity, and the evolution of plant abiotic stress tolerance. *Annals of the New York Academy of Sciences* 1206 (1):56-79.

Dewey, D.R., 1964 GENOME ANALYSIS OF AGROPYRON REPENS × AGROPYRON CRISTATUM SYNTHETIC HYBRIDS. *American Journal of Botany* 51 (10):1062-1068.

Dewey, D.R., 1974 Cytogenetics of Elymus sibiricus and Its Hybrids with Agropyron tauri, Elymus canadensis, and Agropyron caninum. *Botanical Gazette* 135 (1):80-87.

Dewey, D.R., 1975 THE ORIGIN OF AGROPYRON SMITHII. *American Journal of Botany* 62 (5):524-530.

Dewey, D.R., 1976 Derivation of a New Forage Grass from Agropyron repens ‘ Agropyron spicatum Hybrids1. *Crop Science* 16 (2):175-180.

Dewey, D.R., 1983 Historical and Current Taxonomic Perspectives of Agropyron, Elymus, and Related Genera1. *Crop Science* 23 (4):637-642.

Dewey, D.R., 1984 The Genomic System of Classification as a Guide to Intergeneric Hybridization with the Perennial Triticeae, pp. 209-279 in *Gene Manipulation in Plant Improvement: 16th Stadler Genetics Symposium*, edited by J.P. Gustafson. Springer US, Boston, MA.

Drewnowski, A., and B.M. Popkin, 1997 The nutrition transition: new trends in the global diet. *Nutrition Reviews* 55 (2):31-43.

Dubcovsky, J., and J. Dvorak, 2007 Genome Plasticity a Key Factor in the Success of Polyploid Wheat Under Domestication. *Science* 316 (5833):1862-1866.

Edgerton, M.D., 2009 Increasing crop productivity to meet global needs for feed, food, and fuel. *Plant Physiology* 149 (1):7-13.

Erickson, V.J., N.L. Mandel, and F.C. Sorensen, 2004 Landscape patterns of phenotypic variation and population structuring in a selfing grass, Elymus glaucus (blue wildrye). *Canadian Journal of Botany* 82 (12):1776-1789.

FAO, 2007 EcoCrop.

FAO, 2018 FAOSTAT.

Fischer, R., 1996 Wheat Physiology at CIMMYT and Rainsing the Yield Plateau, pp. 191-202 in *Increasing Yield Potential in Wheat: Breaking the Barriers : Proceedings of a Workshop Held in Ciudad Obregón, Sonora, Mexico*, edited by M.P. Reynolds, S. Rajaram and A. McNab. CIMMYT.

Fraga, H., I. García de Cortázar Atauri, A.C. Malheiro, and J.A. Santos, 2016 Modelling climate change impacts on viticultural yield, phenology and stress conditions in Europe. *Global Change Biology* 22 (11):3774-3788.

Friebe, B., Y. Mukai, B.S. Gill, and Y. Cauderon, 1992 C-banding and in-situ hybridization analyses of Agropyron intermedium, a partial wheat x Ag. intermedium amphiploid, and six derived chromosome addition lines. *Theoretical and Applied Genetics* 84 (7):899-905.

Glennon, K.L., M.E. Ritchie, and K.A. Segraves, 2014 Evidence for shared broad-scale climatic niches of diploid and polyploid plants. *Ecology Letters* 17 (5):574-582.

Gomes, S.S.L., J.D. Vidal, C.S. Neves, C. Zorzatto, T.V.S. Campacci *et al.*, 2018 Genome size and climate segregation suggest distinct colonization histories of an orchid species from Neotropical high-elevation rocky complexes. *Biological Journal of the Linnean Society* 124 (3):456-465.

Hahn, M.A., M. van Kleunen, and H. Müller-Schärer, 2012 Increased Phenotypic Plasticity to Climate May Have Boosted the Invasion Success of Polyploid Centaurea stoebe. *PloS one* 7 (11):e50284.

Hall, A.J., and R.A. Richards, 2013 Prognosis for genetic improvement of yield potential and water-limited yield of major grain crops. *Field crops research* 143:18-33.

Heidinga, L., and S.D. Wilson, 2002 The impact of an invading alien grass (Agropyron cristatum) on species turnover in native prairie. *Diversity and Distributions* 8 (5):249-258.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis, 2005 Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25 (15):1965-1978.

Hijmans, R.J., L. Guarino, M. Cruz, and E. Rojas, 2001 Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter*:15-19.

Hsiao, C., K.H. Asay, and D.R. Dewey, 1989 Cytogenetic analysis of interspecific hybrids and amphiploids between two diploid crested wheatgrasses, Agropyron mongolicum and A. cristatum. *Genome* 32 (6):1079-1084.

Hummel, M., B.F. Hallahan, G. Brychkova, J. Ramirez-Villegas, V. Guwela *et al.*, 2018 Reduction in nutritional quality and growing area suitability of common bean under climate change induced drought stress in Africa. *Scientific Reports* 8 (1):16187.

Husband, B.C., and H.A. Sabara, 2004 Reproductive isolation between autotetraploids and their diploid progenitors in fireweed, Chamerion angustifolium (Onagraceae). *New Phytologist* 161 (3):703-713.

IPCC, 2014 *Climate Change 2014 – Impacts, Adaptation and Vulnerability: Global and Sectoral Aspects*: Cambridge University Press.

Jiang, J., B. Friebe, and B.S. Gill, 1993 Recent advances in alien gene transfer in wheat. *Euphytica* 73 (3):199-212.

Johnson, J.A., C.F. Runge, B. Senauer, J. Foley, and S. Polasky, 2014 Global agriculture and carbon trade-offs. *Proceedings of the National Academy of Sciences* 111 (34):12342-12347.

Kagale, S., S.J. Robinson, J. Nixon, R. Xiao, T. Huebert *et al.*, 2014 Polyploid Evolution of the Brassicaceae during the Cenozoic Era. *The Plant Cell* 26 (7):2777-2791.

Koehler, K., A. Center, and J. Cavender-Bares, 2012 Evidence for a freezing tolerance–growth rate trade-off in the live oaks (Quercus series Virentes) across the tropical–temperate divide. *New Phytologist* 193 (3):730-744.

Kolář, F., M. Čertner, J. Suda, P. Schönswetter, and B.C. Husband, 2017 Mixed-Ploidy Species: Progress and Opportunities in Polyploid Research. *Trends in Plant Science* 22 (12):1041-1055.

Kunin, W.E., 1997 Introduction: on the causes and consequences of rare—common differences, pp. 3-11 in *The Biology of Rarity: Causes and consequences of rare—common differences*, edited by W.E. Kunin and K.J. Gaston. Springer Netherlands, Dordrecht.

Lambin, E.F., H.K. Gibbs, L. Ferreira, R. Grau, P. Mayaux *et al.*, 2013 Estimating the world's potentially available cropland using a bottom-up approach. *Global Environmental Change* 23 (5):892-901.

Lesperance, A.L., J.A. Young, R.E. Eckert, and R.A. Evans, 1978 Great Basin Wildrye. *Rangeman's Journal* 5 (4):125-127.

Levin, D.A., 1975 Minority Cytotype Exclusion in Local Plant Populations. *Taxon* 24 (1):35-43.

Li, Z., B. Li, and Y. Tong, 2008 The contribution of distant hybridization with decaploid Agropyron elongatum to wheat improvement in China. *Journal of Genetics and Genomics* 35 (8):451-456.

Liu, S.C., C. Fu, C.-J. Shiu, J.-P. Chen, and F. Wu, 2009 Temperature dependence of global precipitation extremes. *Geophysical Research Letters* 36 (17).

Lobell, D., and C. Field, 2007 Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental Research Letters* 2 (1):014002.

Lorenz, R.J., 1986 Introduction and early use of crested wheatgrass in the Northern Great Plains.

Lu, B.-R., and R. von Bothmer, 1991 Production and cytogenetic analysis of the intergeneric hybrids between nine Elymus species and common wheat (Triticum aestivum L.). *Euphytica* 58 (1):81-95.

Luck, J., M. Spackman, A. Freeman, P. Tre¸bicki, W. Griffiths *et al.*, 2011 Climate change and diseases of food crops. *Plant Pathology* 60 (1):113-121.

Macqueen, D.J., and I.A. Johnston, 2014 A well-constrained estimate for the timing of the salmonid whole genome duplication reveals major decoupling from species diversification. *Proceedings of the Royal Society B: Biological Sciences* 281 (1778).

Mason-Gamer, R.J., 2001 Origin of North American Elymus (Poaceae: Triticeae) Allotetraploids Based on Granule-Bound Starch Synthase Gene Sequences. *Systematic Botany* 26 (4):757-768.

McCouch, S., 2004 Diversifying Selection in Plant Breeding. *PLOS Biology* 2 (10):e347.

Mickelbart, M.V., P.M. Hasegawa, and J. Bailey-Serres, 2015 Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. *Nature Reviews Genetics* 16:237.

Mittler, R., 2006 Abiotic stress, the field environment and stress combination. *Trends in Plant Science* 11 (1):15-19.

Mohammed, Y.S.A., I.S.A. Tahir, N.M. Kamal, A.E. Eltayeb, A.M. Ali *et al.*, 2014 Impact of wheat-Leymus racemosus added chromosomes on wheat adaptation and tolerance to heat stress. *Breeding Science* 63 (5):450-460.

Mráz, P., S. Španiel, A. Keller, G. Bowmann, A. Farkas *et al.*, 2012 Anthropogenic disturbance as a driver of microspatial and microhabitat segregation of cytotypes of Centaurea stoebe and cytotype interactions in secondary contact zones. *Annals of Botany* 110 (3):615-627.

Newbery, F., A. Qi, and B.D.L. Fitt, 2016 Modelling impacts of climate change on arable crop diseases: progress, challenges and applications. *Current Opinion in Plant Biology* 32:101-109.

Pandit, M.K., M.J.O. Pocock, and W.E. Kunin, 2011 Ploidy influences rarity and invasiveness in plants. *Journal of Ecology* 99 (5):1108-1115.

Pearman, P.B., A. Guisan, O. Broennimann, and C.F. Randin, 2008 Niche dynamics in space and time. *Trends in Ecology and Evolution* 23 (3):149-158.

Petit, C., F. Bretagnolle, and F. Felber, 1999 Evolutionary consequences of diploid–polyploid hybrid zones in wild species. *Trends in Ecology & Evolution* 14 (8):306-311.

Pingali, P.L., 2012 Green Revolution: Impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences* 109 (31):12302-12308.

Popkin, B.M., 1993 Nutritional patterns and transitions. *Population and development review*:138-157.

Popkin, B.M., 2017 Relationship between shifts in food system dynamics and acceleration of the global nutrition transition. *Nutrition Reviews* 75 (2):73-82.

Potter, L.D., 1948 *Comparative Seedling Characteristics of Standard Crested Wheatgrass (Agropyron Cristatum (L.) Gaertn.) and Russian Wild-rye (Elymus Junceus Fisch.)*: University of Minnesota.

Puchta, H., 2017 Applying CRISPR/Cas for genome engineering in plants: the best is yet to come. *Current Opinion in Plant Biology* 36:1-8.

Ramirez-Villegas, J., and A. Jarvis, 2010 Downscaling global circulation model outputs: the delta method decision and policy analysis working paper no. 1. Centro Internacional de Agricultura Tropical (CIAT).

Ramirez-Villegas, J., A. Jarvis, and P. Läderach, 2013 Empirical approaches for assessing impacts of climate change on agriculture: The EcoCrop model and a case study with grain sorghum. *Agricultural and Forest Meteorology* 170:67-78.

Ramsey, J., 2011 Polyploidy and ecological adaptation in wild yarrow. *Proceedings of the National Academy of Sciences* 108 (17):7096-7101.

Ray, D.K., N.D. Mueller, P.C. West, and J.A. Foley, 2013 Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PloS one* 8 (6):e66428.

Renny-Byfield, S., and J.F. Wendel, 2014 Doubling down on genomes: Polyploidy and crop plants. *American Journal of Botany* 101 (10):1711-1725.

Reynolds, M., M. Balota, M. Delgado, I. Amani, and R. Fischer, 1994 Physiological and Morphological Traits Associated With Spring Wheat Yield Under Hot, Irrigated Conditions. *Functional Plant Biology* 21 (6):717-730.

Reynolds, M.P., S. Rajaram, and K.D. Sayre, 1999 Physiological and Genetic Changes of Irrigated Wheat in the Post–Green Revolution Period and Approaches for Meeting Projected Global Demand. *Crop Science* 39 (6):1611-1621.

Rippke, U., J. Ramirez-Villegas, A. Jarvis, S.J. Vermeulen, L. Parker *et al.*, 2016 Timescales of transformational climate change adaptation in sub-Saharan African agriculture. *Nature Climate Change* 6:605.

Rivero-Guerra, A.O., 2008 Cytogenetics, geographical distribution, and pollen fertility of diploid and tetraploid cytotypes of Santolina pectinata Lag. (Asteraceae: Anthemideae). *Botanical Journal of the Linnean Society* 156 (4):657-667.

Rötter, R.P., J. Höhn, M. Trnka, S. Fronzek, T.R. Carter *et al.*, 2013 Modelling shifts in agroclimate and crop cultivar response under climate change. *Ecology and Evolution* 3 (12):4197-4214.

Saini, H., M. Sedgley, and D. Aspinall, 1983 Effect of Heat Stress During Floral Development on Pollen Tube Growth and Ovary Anatomy in Wheat (<I>Triticum aestivum</I> L.). *Functional Plant Biology* 10 (2):137-144.

Saini, H., M. Sedgley, and D. Aspinall, 1984 Development Anatomy in Wheat of Male Sterility Induced by Heat Stress, Water Deficit or Abscisic Acid. *Functional Plant Biology* 11 (4):243-253.

Sattler, M.C., C.R. Carvalho, and W.R. Clarindo, 2016 The polyploidy and its key role in plant breeding. *Planta* 243 (2):281-296.

Schneider, A., I. Molnár, and M. Molnár-Láng, 2008 Utilisation of Aegilops (goatgrass) species to widen the genetic diversity of cultivated wheat. *Euphytica* 163 (1):1-19.

Sharma, H.C., 1995 How wide can a wide cross be? *Euphytica* 82 (1):43-64.

Sharma, H.C., and B.S. Gill, 1983 Current status of wide hybridization in wheat. *Euphytica* 32 (1):17-31.

Shiferaw, B., M. Smale, H.-J. Braun, E. Duveiller, M. Reynolds *et al.*, 2013 Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Security* 5 (3):291-317.

Smil, V., 2000 Feeding the world: A challenge for the 21st century. MIT Press, Cambridge, MA.

Smith, D.C., 1942 Intergeneric hybridization of cereals and other grasses. *J. Agric. Res* 64:33-47.

Soltis, D.E., C.D. Bell, S. Kim, and P.S. Soltis, 2008 Origin and Early Evolution of Angiosperms. *Annals of the New York Academy of Sciences* 1133 (1):3-25.

Soltis, P., G. Plunkett, S. Novak, and D. Soltis, 1995 GENETIC VARIATION IN TRAGOPOGON SPECIES: ADDITIONAL ORIGINS OF THE ALLOTETRAPLOIDS T. MIRUS AND T. MISCELLUS (COMPOSITAE). *American Journal of Botany* 82 (10):1329-1341.

Stebbins, G.L., 1971 Chromosomal evolution in higher plants. *Chromosomal evolution in higher plants.*

Sutter, G.C., and R.M. Brigham, 1998 Avifaunal and habitat changes resulting from conversion of native prairie to crested wheat grass: patterns at songbird community and species levels. *Canadian Journal of Zoology* 76 (5):869-875.

Tank, D.C., J.M. Eastman, M.W. Pennell, P.S. Soltis, D.E. Soltis *et al.*, 2015 Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. *New Phytologist* 207 (2):454-467.

Tester, M., and P. Langridge, 2010 Breeding Technologies to Increase Crop Production in a Changing World. *Science* 327 (5967):818-822.

Theodoridis, S., C. Randin, O. Broennimann, T. Patsiou, and E. Conti, 2013 Divergent and narrower climatic niches characterize polyploid species of European primroses in Primula sect. Aleuritia. *Journal of Biogeography* 40 (7):1278-1289.

Tilman, D., C. Balzer, J. Hill, and B.L. Befort, 2011 Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences* 108 (50):20260-20264.

Trenberth, K.E., 2011 Changes in precipitation with climate change. *Climate Research* 47 (1-2):123-138.

van de Wouw, M., C. Kik, T. van Hintum, R. van Treuren, and B. Visser, 2009 Genetic erosion in crops: concept, research results and challenges. *Plant Genetic Resources* 8 (1):1-15.

Vaness, B.M., and S.D. Wilson, 2007 Impact and management of crested wheatgrass (Agropyron cristatum) in the northern Great Plains. *Canadian Journal of Plant Science* 87 (5):1023-1028.

Vanneste, K., S. Maere, and Y. Van de Peer, 2014 Tangled up in two: a burst of genome duplications at the end of the Cretaceous and the consequences for plant evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369 (1648).

Visger, C.J., C.C. Germain-Aubrey, M. Patel, E.B. Sessa, P.S. Soltis *et al.*, 2016 Niche divergence between diploid and autotetraploid Tolmiea. *American Journal of Botany* 103 (8):1396-1406.

Vogel, J.C., F.J. Rumsey, J.J. Scheller, J.A. Barrett, and M. Gibby, 1999 Where are the glacial refugia in Europe? Evidence from pteridophytes. *Biological Journal of the Linnean Society* 66 (1):23-37.

Wang, R., R. Von Bothmer, J. Dvorak, G. Fedak, I. Linde-Laursen *et al.*, 1994 Genome symbols in the Triticeae (Poaceae).

Wang, R.C., 2011 Agropyron and Psathyrostachys, pp. 77-108 in *Wild Crop Relatives: Genomic and Breeding Resources: Cereals*, edited by C. Kole. Springer Berlin Heidelberg.

Wang, R.C., and K.B. Jensen, 2009 Wheatgrasses and Wildrye Grasses, pp. 41-80 in *Genetic Resources, Chromosome Engineering, and Crop Improvement: Forage Crops*, edited by R.J. Singh. CRC Press.

Westover, H.L., 1932 *Crested wheatgrass as compared with bromegrass, slender wheatgrass, and other hay and pasture crops for the northern Great Plains*: U.S. Dept. of Agriculture.

Wu, J., X. Yang, H. Wang, H. Li, L. Li *et al.*, 2006 The introgression of chromosome 6P specifying for increased numbers of florets and kernels from Agropyron cristatum into wheat. *Theoretical and Applied Genetics* 114 (1):13-20.

Yen, Y., and D. Liu, 1987 Production, morphology, and cytogenetics of intergeneric hybrids of Elymus L. species with Triticum aestivum L. and their backcross derivatives. *Genome* 29 (5):689-694.

Zaharieva, M., N.G. Ayana, A.A. Hakimi, S.C. Misra, and P. Monneveux, 2010 Cultivated emmer wheat (Triticum dicoccon Schrank), an old crop with promising future: a review. *Genetic Resources and Crop Evolution* 57 (6):937-962.

Zamir, D., 2001 Improving plant breeding with exotic genetic libraries. *Nature Reviews Genetics* 2:983.

Zohren, J., N. Wang, I. Kardailsky, J.S. Borrell, A. Joecker *et al.*, 2016 Unidirectional diploid–tetraploid introgression among British birch trees with shifting ranges shown by restriction site-associated markers. *Molecular Ecology* 25 (11):2413-2426.

Overall Conclusions

This thesis contains five individual pieces of research into genome dosage and genetic hybridity effects on plant growth. While they exist as standalone chapters, they are undoubtably linked. In this final section I will explain their relevance to one another and end with some concluding messages.

To begin, I introduced the agriculturally important root crop sugar beet. Commercial breeding companies are working with sugar beet to improve its use as a source of food, feed, fuel, and novel nutraceuticals. The crop is grown at different ploidy levels and F1 hybrid varieties are commonly available on national registered variety lists. This makes it a valuable resource to test hypotheses relating to the interaction between polyploidy and heterosis. My research built upon the well-established paternal genome dosage effect on heterosis in the model organism *Arabidopsis thaliana* and the maize crop (*Zea mays*). I discovered the absence of any paternal genome dosage effects on heterosis in sugar beet. This holds true even when the extra paternal genome is highly heterozygous. This suggests that, firstly, most heterotic gains in sugar beet can likely be achieved at the diploid level and, secondly, increasing multi-locus nuclear heterozygosity in a polyploid context is not correlated with increasing heterotic gains for agronomically important traits. This point regarding heterozygosity leads us to my next research chapter.

In order to maximise heterozygosity in an F1 hybrid, ideally one would use a polyploid plant. For example, let us assume a hypothetical gene at a certain locus has four alleles; *a*, *b*, *c*, and *d*. A diploid plant will, at most, possess a di-allelic locus for this gene (*ab*, *ac*, *ad*, *bc*, *bd* or *cd*). However, a tetraploid plant can in theory harbour all four alleles at once in a tetra-allelic locus (*abcd*). The phenomenon of 'progressive heterosis' suggests that heterosis is maximised in polyploid plants because they can achieve high levels of multi-locus nuclear heterozygosity. The model organism *Arabidopsis thaliana* is an ideal candidate to investigate progressive heterosis due to the wealth of heterosis knowledge about the plant. The number of tetraploid *A. thaliana* crosses performed in my second research chapter is considerably larger in scope than seen in my opening research chapter in sugar beet. Using a stepwise progression, it is demonstrated that increasing multi-locus heterozygosity in *A. thaliana* can enhance leaf area heterosis, but a threshold is reached where heterosis gains effectively stop. These novel findings should be a springboard for further research into what is driving the heterosis effect for leaf growth and can have direct relevance to polyploid hybrid breeding programs. Likewise, my next research chapter also has important links with commercial plant breeding.

In chapter four I focus on haploid plant breeding. Haploid plants are those which contain a gametic chromosome number. They are commonly generated to have a single copy of their chromosome, i.e. a '1x' plant derived from a '2x' plant. While sterile, 1x plants can be incredibly useful to breeders upon doubling their chromosome number because 'double haploids' regain fertility and are homozygous at all loci. Achieving 100% homozygosity (and thus fixing traits of interest) in just two steps is a huge improvement over several rounds of inbreeding and in awkward scenarios involving self-incompatible or dioecious species. The ability to generate haploid plants, however, can be cumbersome and unreliable when plant breeders are depending upon whole-plant regeneration from gametes *in vitro*, or, a series of ever-widening inter-species crosses. Genetic engineering of the centromeric region in *A. thaliana* and crop species represents a new avenue for haploid plant breeding. I investigated different genetic contexts in which centromere-mediated chromosome elimination in *A. thaliana* succeeds or fails. My key finding is that uniparental genome elimination caused by a genetically engineered centromere is affected by the maternal environment in which it resides. Further research is needed to disaggregate the cytonuclear influences at play, and my hope is that this can benefit future haploid breeding efforts in commercial crops. Another important plant breeding goal is salt tolerance, which is addressed in my next research chapter.

In chapter five I investigate two potential plant breeding mechanisms that could influence salt stress tolerance in *A. thaliana*: one is doubling the genome dosage (from a 2x plant to a 4x plant) and the other is whether an extra genome is inherited maternally or paternally (a maternal-excess 3x plant or a paternal-excess 3x plant). Both genome dosage and parent-of-origin effects on plant growth in non-stress conditions are well established in *A. thaliana.* Furthermore, there is some evidence to suggest that tetraploid plants display greater salt stress tolerance than equivalent diploid plants. There is an absence of literature on parent-of-origin effects on salt stress tolerance. Thus, this chapter neatly fits in with the classic literature that seeks to investigate the immediate consequences of polyploidisation. Using 10 genotypes, I demonstrated that there is generally an absence of genome dosage and parent-of-origin effects on salt stress tolerance in *A. thaliana*. This is relevant to arid and semi-arid agricultural settings which are seeking to grow crops on soil that is at risk from severe salinization. Abiotic stresses on crop growth are further addressed in my final research chapter.

In chapter six, I took several concepts regarding polyploid plants together, and investigated the future of cereal growth. First, I raise an important food security topic – the current and future demand for cereal products, particularly wheat (*Triticum*). Second, I consider the ubiquity of wild diploid and polyploid grass species to ask (a) Will lower ploidy plants

within a species be at a greater risk to climate change than higher ploidy plants? (b) Will domesticated wheat be at a greater risk to climate change than related, wild wheatgrass (*Agropyron*) and wildrye (*Elymus*)?

I began this thesis with a quote from renowned botanist George Ledyard Stebbins, who wrote forcefully in the mid-twentieth century on the influence of polyploidy in the evolution of plants. Indeed, the potential for harnessing polyploidy as a crop improvement tool is now firmly established, considering the multitude of polyploid crops which form part of our daily diets, e.g. wheat, oats, potatoes, coffee, peanuts (Table 1). The research undertaken during my thesis was a deliberate effort to weave the topic of polyploidy together with another crop improvement tool, heterosis. I considered the impacts of genome dosage alone on plant growth (chapters 4, 5 and 6), as well as the impacts of merging divergent genomes in a diploid and polyploid context (chapters 2, 3 and 6). While I am taken by G. Ledyard Stebbins' optimism that auto- and allopolyploid plant breeding represents an opportunity to produce new economically valuable crops, I conclude from my own research that more genomes is not necessarily always better. Modern plant breeding companies can produce equally heterotic genotypes at the diploid level (chapter 2). Maximising the allelic diversity in autotetraploid plants does not maximise the heterotic effect (chapter 3). Shifting the genome dosage of young plants has no effect on salt stress tolerance (chapter 5). Yet, these are general conclusions and there are interesting exceptions therein. My wish is that the un-published research chapters in this thesis can each form part of a future, peer-reviewed publication that clarifies the role of genome dosage and genetic hybridity in crop improvement.

Supplementary Information

Appendix 1

Chapter 2. Supplementary Materials and Methods.

Detailed methods of genotyping-by-sequencing analysis. The steps taken for library preparation, high-throughput DNA sequencing, bioinformatic analysis of sequencing data, and software used are explained herein.

Library Preparation and Sequencing

Throughout the whole process of sequencing from the DNA sample to the final data, each step, including sample test, library preparation, and sequencing procedures, influences the quality of data production, while the data quality further impacts on the analysis results directly. To guarantee the accuracy and reliability of the sequencing data, we utilized stringent quality control (QC) procedures. The workflow was as follows: 1. DNA quantification and qualification, 2. Library construction, 3. Library quality control, High-throughput DNA sequencing, 5. Quality control (QC).

1. DNA quantification and qualification

The three major QC methods for DNA sample qualification were as follows: 1. Agarose gel electrophoresis analysis for DNA purity and integrity, 2. NanoDrop® 2000 spectrophotometer measurement for DNA purity by assessing the OD_{260}/OD_{280} ratio, 3. Qubit® 2.0 fluorometer quantitation for accurate measurement of DNA concentration. Sample DNA, with OD_{260}/OD_{280} ratio of 1.8 to 2.0 and total amount of more than 1.5 μg, was used for library construction.

2. Library Construction

The genomic DNA of each sample was digested with restriction enzymes based on the *in silico* evaluation results, and the obtained fragments were ligated with two barcoded adapters with a compatible sticky end corresponding to the primary restriction digestion enzyme and the Illumina P5 or P7 universal sequence. Following several rounds PCR amplification, all the samples were pooled and size-selected for the required fragments to complete the library construction. The experimental procedures are as follows: 1. *In/Ex silico* digestion evaluation: The genome assembly was subjected to *in silico* digestion analysis to aid the optimization of enzyme sets and fragment size, ensure even genome coverage, and identification of repeated regions. The combination of experimental digestion assay with the computational approach

ensures high reliable and reproducible data production. 2. Restriction enzyme digestion: 0.3~0.6 μg genomic DNA was digested completely with the optimized restriction enzyme set, in order to obtain a suitable marker density. 3. Ligating P5 and P7 adapter: each end of digested fragment was respectively ligated with P5 and P7 barcoded-adapter (with complementarily sticky ends to the digested DNA). 4. PCR enrichment and fragment selection: tags containing both P5 and P7 adapters were amplified through PCR. After DNA fragments of different samples were pooled, the desired fragments of DNA were recovered from gel electrophoresis. 5. High-throughput DNA sequencing: After cluster preparation, high-throughput DNA sequencing was performed on Illumina HiSeq platform

The experimental procedures of DNA library preparation are shown in Figure 1.

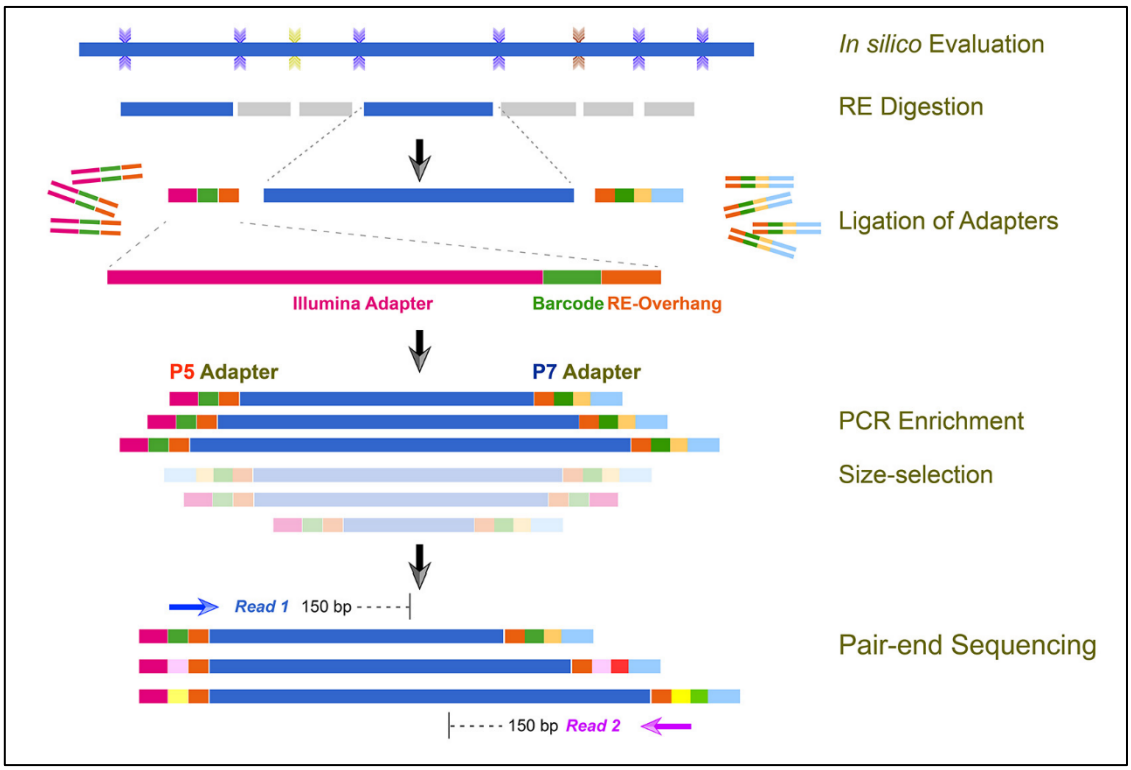


Figure 1. Experimental procedures of library preparation.

3. Library Quality Control

To check the prepared DNA libraries, Qubit® 2.0 fluorometer was used to determine the concentration of the library. After dilution to 1 ng/ul, the Agilent® 2100 bioanalyzer was used to assess the insert size. Finally, quantitative real-time PCR (qPCR) was performed to detect the effective concentration of each library. If the library with appropriate insert size has an effective concentration of more than 2 nM, the constructed libraries are of sufficient quality and ready for Illumina® high-throughput sequencing.

4. High-throughput DNA Sequencing

The DNA libraries were pooled according to their effective concentration as well as the expected data production. Pair-end sequencing was performed on Illumina®HiSeq platform, with the read length of 144 bp at each end.

Bioinformatics Analysis

The bioinformatic analysis procedures are as follows: 1. Creation of FASTQ files 2. Quality control of raw sequencing data for clean data filtration, 3. Mapping clean reads to reference genome, 4. SNP and InDel detection and annotation according to the reference genome mapping results.

1. Creation of FASTQ files

The original data acquired by Illumina®HiSeq platform in image files are firstly converted into sequence data by base calling with the CASAVA software (version 1.8). The sequences and corresponding sequencing quality information are stored in a FASTQ file.

Every read in FASTQ format is stored in four lines as follows:

```
@K00124:82:H2MH5BBXX:1:1101:31389:1158                                  2:N:0:0
TAGCCACATAGAAACCAACAGCCATATAACTGGTAGCTTTAAGCGGCTCACCTTTAGCATCAACAGGCCACAACCAACC
AGAACGTGAAAAAGCGTCCTGCGTGTAGCGAACTGCGATGGGCATACAGATCGGAAGAGCGTCGTGTAGGG
+
AAFFFKKKKKKKFKKKFFKKAAFKKKKKFKKKKFKKA,FKKKKKKKKKAKKFKKKKKKKAKKKKKKFFKKKKF<FFKK
KKKKKKKKKKKFKKFKKF7FFFFFKFKKKFKKKKKKKKF<FFKKKKFKKKKKFKFKFKKFK<<F,A7,AFK
```

Line 1 begins with an '@' character and is followed by Illumina sequence identifiers, and an optional description (such as a FASTA title line). Line 2 is the sequence of a sequencing read. Line 3 begins with a '+' character and is optionally followed by Illumina sequence identifier and description. Line 4 encodes the quality values for the sequence in Line 2, and must contain the same number of characters as the bases in the sequence. The per base sequencing quality score could be calculated by the ASCII value of each character in Line 4 minus a constant 33.

EAS139	Unique instrument name
136	Run ID
FC706VJ	Flowcell ID
2	Flowcell lane
2104	Tile number within the flowcell lane
15343	'x'-coordinate of the cluster within the tile

197393	'y'-coordinate of the cluster within the tile
1	Member of a pair, 1 or 2 (paired-end or mate-pair reads only)
Y	Y if the read fails filter (read is bad), N otherwise
18	0 when none of the control bits are on, otherwise it is an even number
ATCACG	Index sequence

The raw sequence FASTQ files are available on the NCBI Sequence Read Archive, data set SRP125454.

2. Quality Control of Sequencing Data

2.1 Sequencing Quality Distribution

If the sequencing error rate is represented by e, and Illumina HiSeq^{TM} 2500 sequencing quality by Q_{phred}, the quality score of a base (Phred score) is calculated by the following equation: $Q_{phred} = -10\log_{10}(e)$. The correspondence relationship between Illunima sequencing quality and Phred score in base calling by CASAVA is listed as follows:

Phred score	**Error Rate**	**Correct Rate**	**Q-score**
10	1/10	90%	Q10
20	1/100	99%	Q20
30	1/1000	99.9%	Q30
40	1/10000	99.99%	Q40

For next-generation sequencing (NGS), the sequencing platform, chemical reactants, and sample quality can influence sequencing quality and base error rate. Sequencing quality distribution is examined over the full length of all sequences to detect any sites (base positions) with an unusually low sequencing quality and where incorrect bases may be incorporated at abnormally high levels. For detailed sequencing quality distribution, please refer https://figshare.com/collections/Hybridity_has_a_greater_effect_than_paternal_genome_dosage_on_heterosis_in_Sugar_Beet/4024687.

2.2 Distribution of Sequencing Errors

Sequencing error rate is related to the base quality of the obtained sequence. The sequencing platform, chemical reactants, and sample quality can all influence sequencing error rate and herein the base quality. For NGS with sequencing-by-synthesis strategy, sequencing error rate distribution shows two common features: 1. Error rate increases with extending of the

sequencing reads due to the consumption of chemical reagents, damage of the DNA template by laser irradiation, and possible accumulation of errors during the sequencing cycles. All the Illumina high-throughput sequencing platforms have this feature. 2. The sequencing error rate is higher for the first several bases than at other positions, which is likely the result of reading errors during the first few cycles after calibration of the optical instruments.

Sequencing error rate distribution is examined over the full length of all sequences, to detect any sites (base positions) with an unusually high error rate, where incorrect bases may be incorporated at abnormally high levels. For detailed sequencing error distribution, please refer to https://figshare.com/collections/Hybridity_has_a_greater_effect_than_paternal_genome_dosage_on_heterosis_in_Sugar_Beet/4024687.

2.3 Sequencing Data Filtration

Raw data obtained from sequencing contains adapter contamination and low-quality reads. These sequencing artefacts may increase the complexity of downstream analyses, and therefore, we utilize quality control steps to remove them. Consequently, all the downstream analyses are based on the clean reads. The quality control steps are as follows: 1. Discard the paired reads when either read contains adapter contamination. 2. Discard the paired reads when uncertain nucleotides (N) constitute more than 10 percent of either read. 3. Discard the paired reads when low quality nucleotides (base quality less than 5, $Q \leq 5$) constitute more than 50 percent of either read. For detailed sequencing read classification, please refer to https://figshare.com/collections/Hybridity_has_a_greater_effect_than_paternal_genome_dosage_on_heterosis_in_Sugar_Beet/4024687.

2.4 Statistics of Sequencing Data

Statistics of sequencing data are listed in https://figshare.com/collections/Hybridity_has_a_greater_effect_than_paternal_genome_dosage_on_heterosis_in_Sugar_Beet/4024687.

2.5 Sequencing Evaluation Summary

In total 29.068Gbit raw data were sequenced from this run, with 29.065Gbit clean data generated after filtering low-quality data. The raw data production for each sample ranged from 519.515 M to 784.908 M, indicating a sufficient amount of data production. As the Q20 and Q30 reached 94.02% and 85.94%, respectively, the sequencing quality meets the proper

analysis requirements. The GC content of 35.92% to 38.22% is also in the normal distribution range, fulfilling the quality standard.

In conclusion, the library construction and sequencing procedures are successful and highly reliable.

3. Mapping Statistics

The sequencing data was aligned with the reference sequence through BWA (Li and Durbin 2009) software (parameters: mem -t 4 -k 32 -M), and the mapping rate and coverage was counted according to the alignment results (see Table 4.3). The BAM files were handled by SAMtools (Li et al. 2009).

3.1 Statistics of Reference Genome

Reference genome is available at: http://bvseq.molgen.mpg.de/Genome/Download/RefBeet-1.2/RefBeet-1.2.fna.gz

Statistics of the reference genome are listed in https://figshare.com/collections/Hybridity_has_a_greater_effect_than_paternal_genome_dosage_on_heterosis_in_Sugar_Beet/4024687.

3.2 Mapping Statistics with Reference Genome and Tag Summary

The mapping rates of samples reflect the similarity between each sample and the reference genome. The depth and coverage are indicators of the evenness and homology with the reference genome. With GBS, tag-related statistics are also calculated.

Statistics of mapping rate, depth and coverage, as well as tag-related statistics are listed in https://figshare.com/collections/Hybridity_has_a_greater_effect_than_paternal_genome_dosage_on_heterosis_in_Sugar_Beet/4024687.

3.3 Mapping Summary

For the current 566,571,340 bp reference genome, the mapping rate of each sample ranges from 97.59% to 98.31%. The average depth on the reference genome (without Ns) is in 5.19X to 9.7X range, while the more than 1X coverage exceeds 12.17%. This result is in the qualified normal range and may serve in the subsequent variation detection and related analyses.

3.4 Enzymatic Digestion Summary

Among pairs of clean reads, those containing the exact conserved sequence of the first restriction enzyme at the beginning ends of both Read1 and Read2 are considered as successfully enzyme-catched reads, while those containing no recognition sequence of both the primary and additional restriction enzyme(s) are considered as completely cut reads. In this experiment, the ratio of enzyme-catched reads is among 97.2% to 99.3%, while the enzyme-digestion ratio ranges from 80.6% to 93.7%. All enzymatic digestion statistics are listed in https://figshare.com/collections/Hybridity_has_a_greater_effect_than_paternal_genome_dosage_on_heterosis_in_Sugar_Beet/4024687.

4. SNP Detection & Annotation

Single nucleotide polymorphism (SNP) refers to a variation in a single nucleotide which may occur at some specific position in the genome, including transition and transversion of a single nucleotide. We detected the individual SNP variations using SAMtools (Li et al. 2009) with the following parameter: 'mpileup -m 2 -F 0.002 -d 1000'.

To reduce the error rate in SNP detection, we filtered the results with the criterion as follows: 1. The number of support reads for each SNP should be more than 4. 2. The mapping quality (MQ) of each SNP should be higher than 20.

4.1 Statistics of SNP Detection & Annotation

ANNOVAR (Wang et al. 2010) is a widely used software in variation annotation with multiple capabilities, including gene-based annotation, region-based annotation, filter-based annotation as well as other functionalities. We used ANNOVAR to perform annotation of detected SNPs. The UCSC known genes (https://genome.ucsc.edu/) were used for gene and region annotations.

Statistics of SNP detection and annotation are listed in https://figshare.com/collections/Hybridity_has_a_greater_effect_than_paternal_genome_dosage_on_heterosis_in_Sugar_Beet/4024687.

4.2 SNP Quality Distribution

To assess the credibility of detected SNPs, we checked the distribution of support reads number, SNP quality, as well as the distance between adjacent SNPs.

4.3 SNP Mutation Frequency

Taking the T:A>C:G mutations as an example, this category includes mutations from T to C and A to G. When T>C mutation appears on either of the double-strand, the A>G mutation will

be found in the same position of the other chain. Therefore, the T>C and A>G mutations are classified into one category. Accordingly, the whole-genome SNP mutations could be classified into six categories. The frequency of each type is shown in Figure 2.

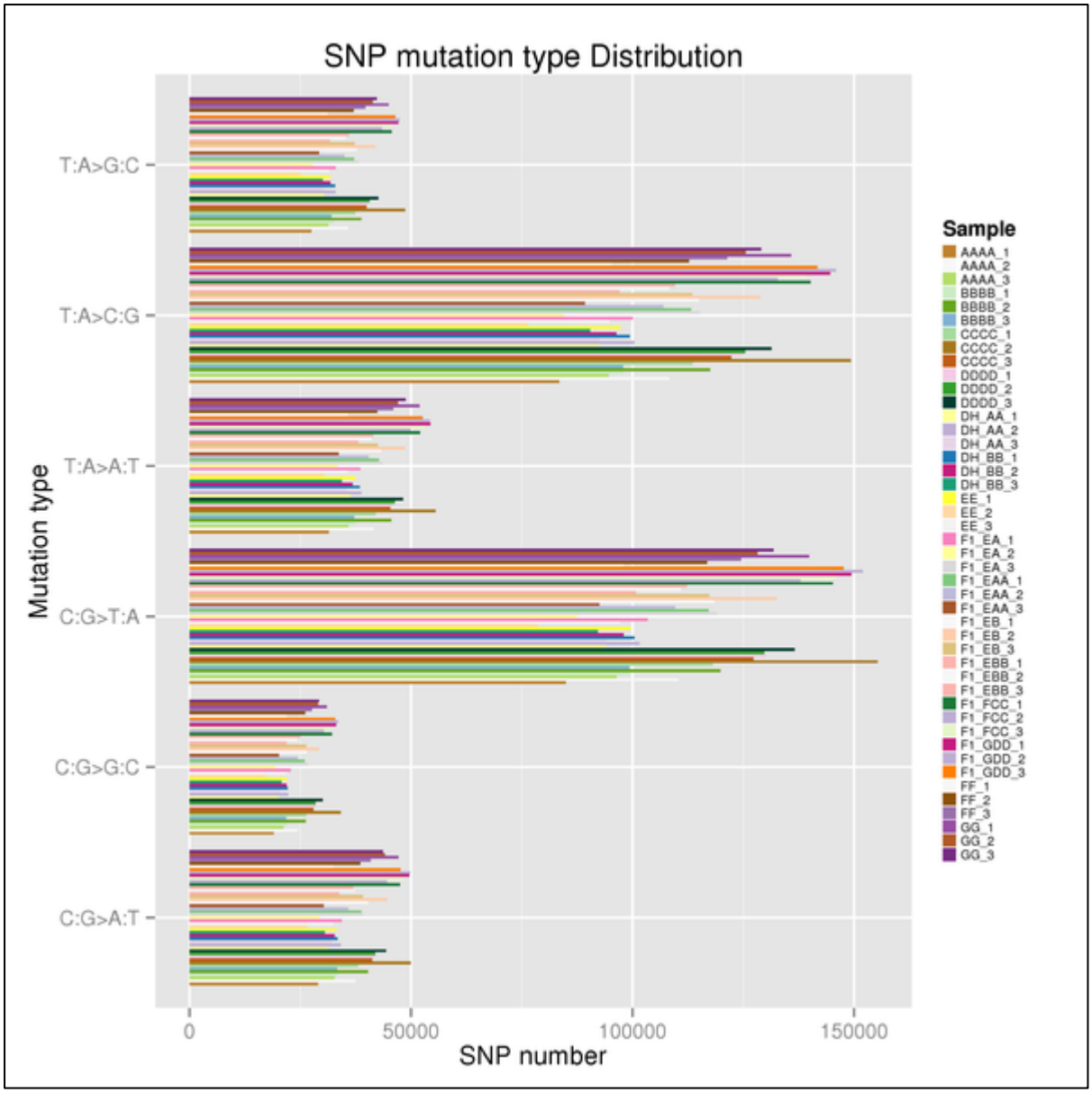


Figure 2. Frequency of SNP mutations.

The x-axis represents the number of the SNPs, and y-axis indicates the mutation types.

References

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25 (14):1754-1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan *et al.*, 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* 25 (16):2078-2079.

Wang, K., M. Li, and H. Hakonarson, 2010 ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research* 38 (16):e164-e164.

List of Software

Analysis	**Software**	**Usage**	**Version**
Mapping	BWA	Mapping clean reads to the reference genome and generation of bam result files.	0.7.8-r455
	SAMtools	Sorting the bam files.	0.1.19-44428cd
SNP/InDel Detection	SAMtools	Detection and filtration of SNPs and InDels.	0.1.19-44428cd
Variation Annotation	ANNOVAR	Annotation of the detected variations.	2013Aug23

Chapter 2. Supplementary Results.

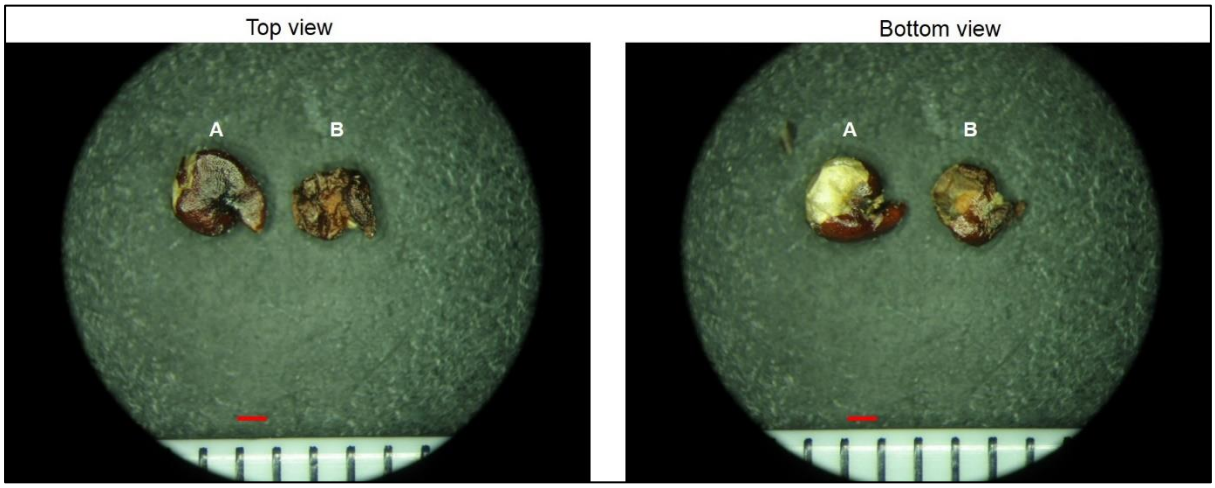
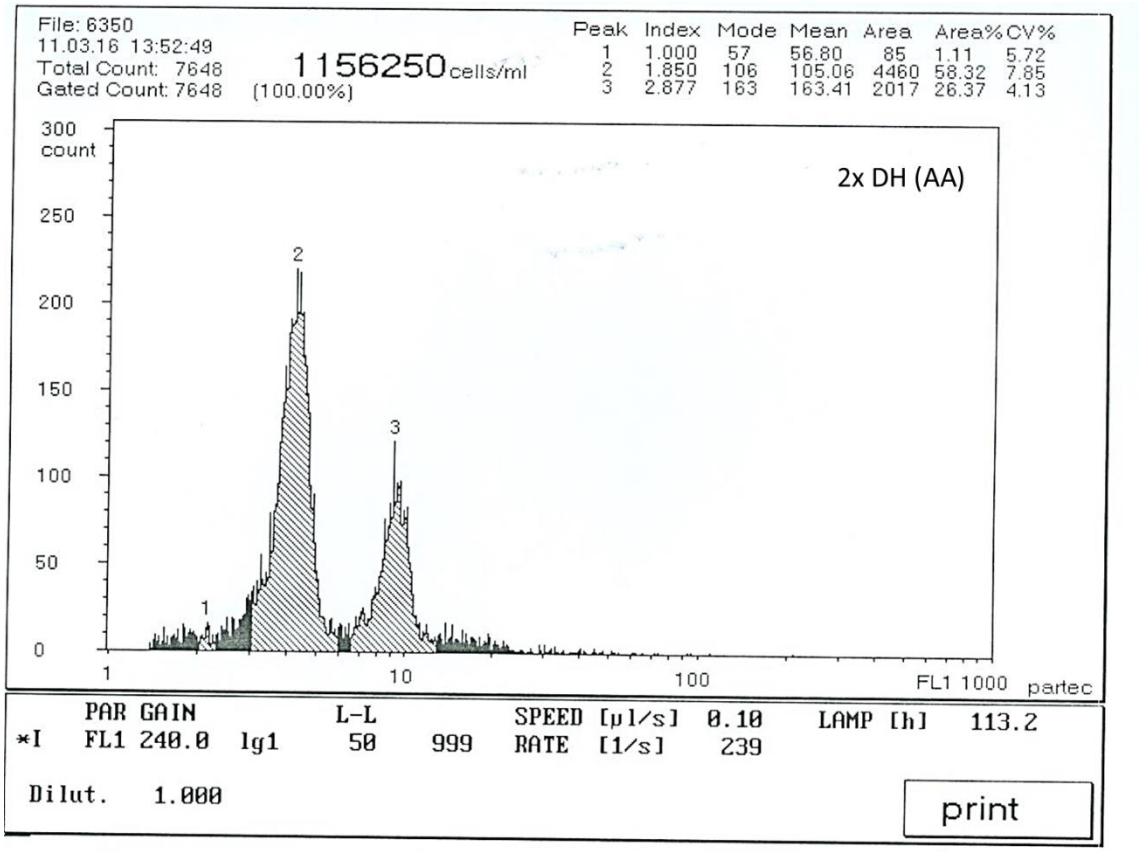
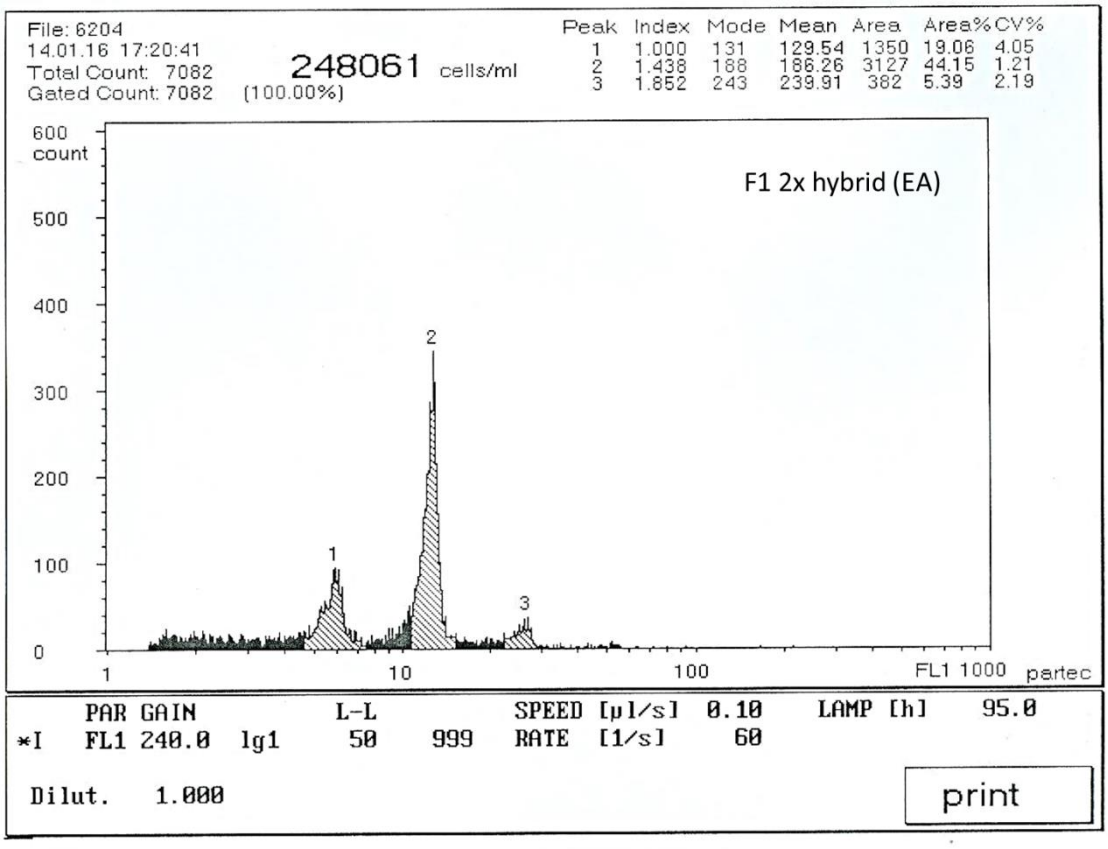
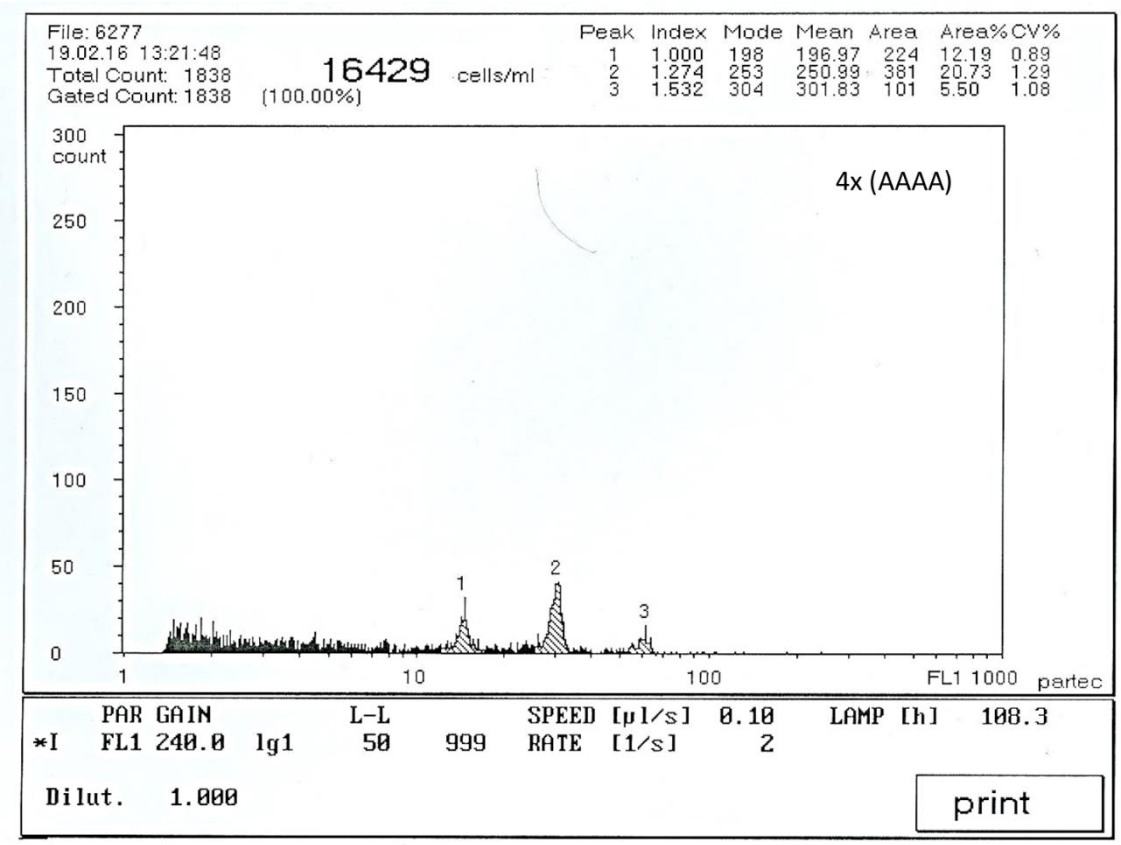


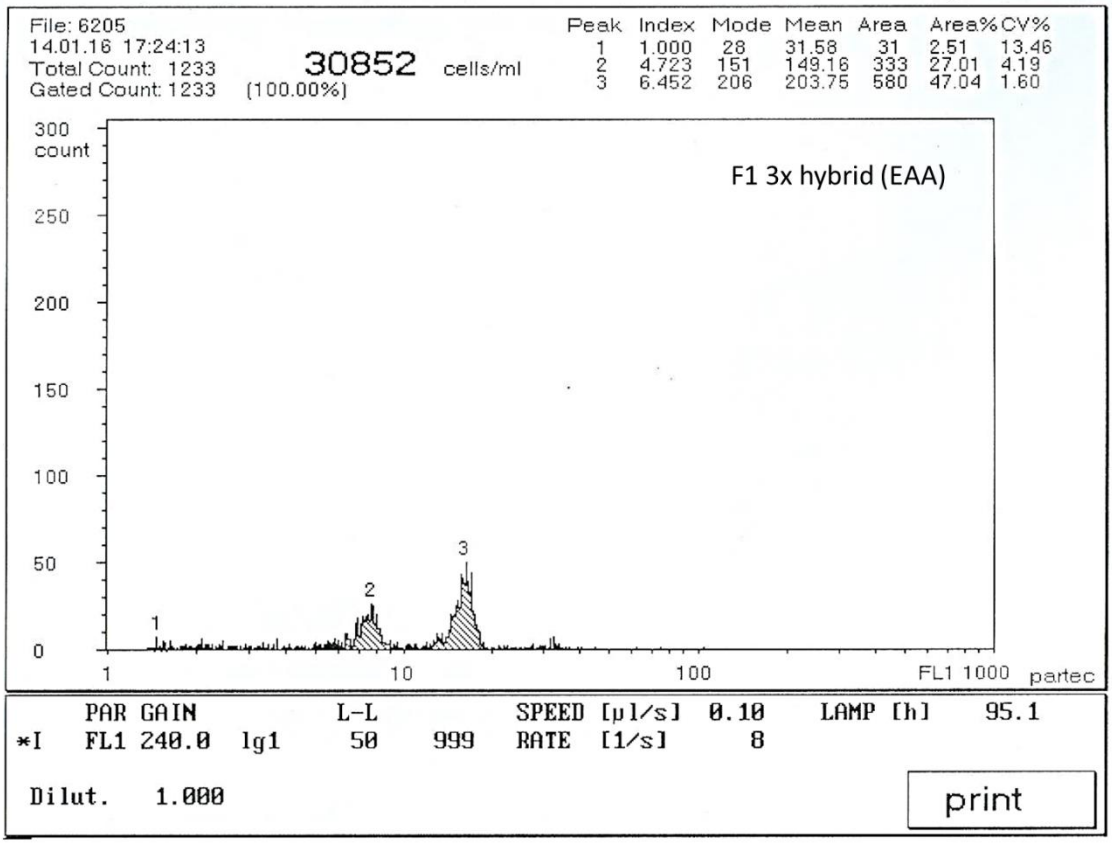
Figure S1. Top and bottom view of representative alive and dead sugar been seeds.

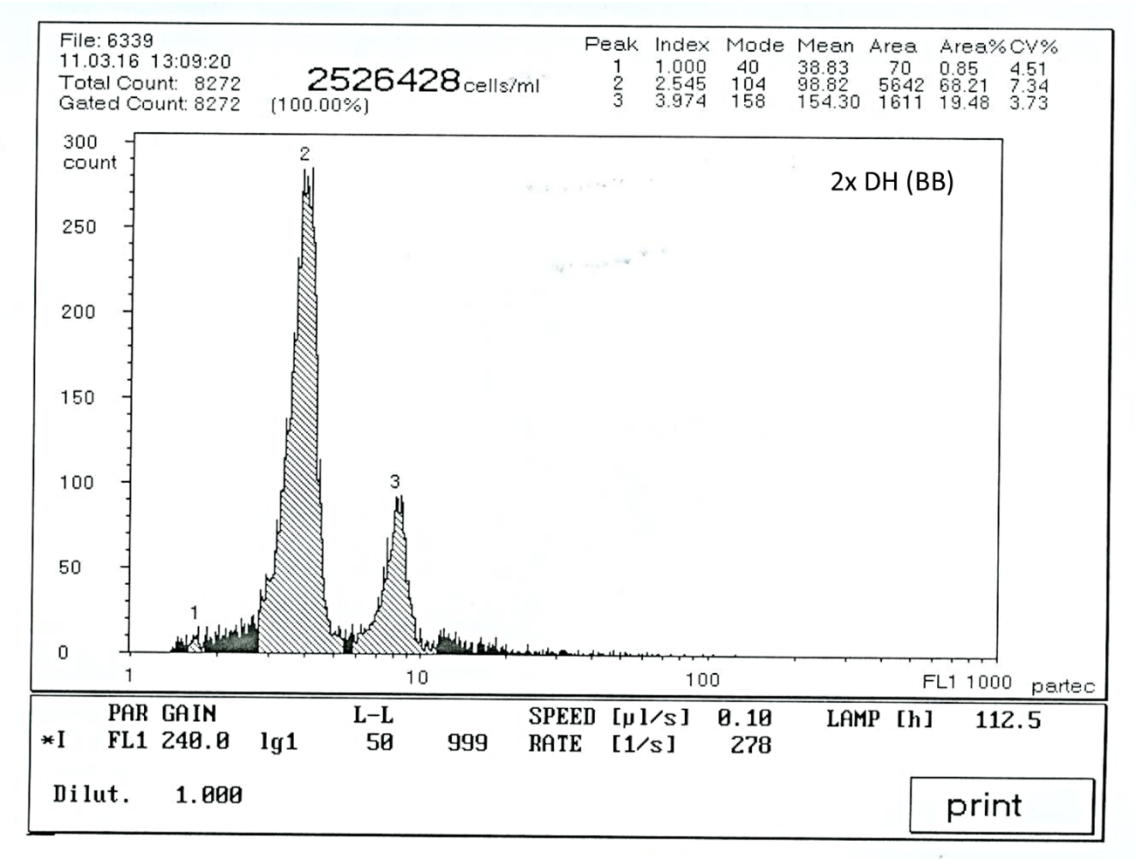
Scanned images of flow cytometry data. Analysis of nuclei from leaves confirms ploidy level of each line.

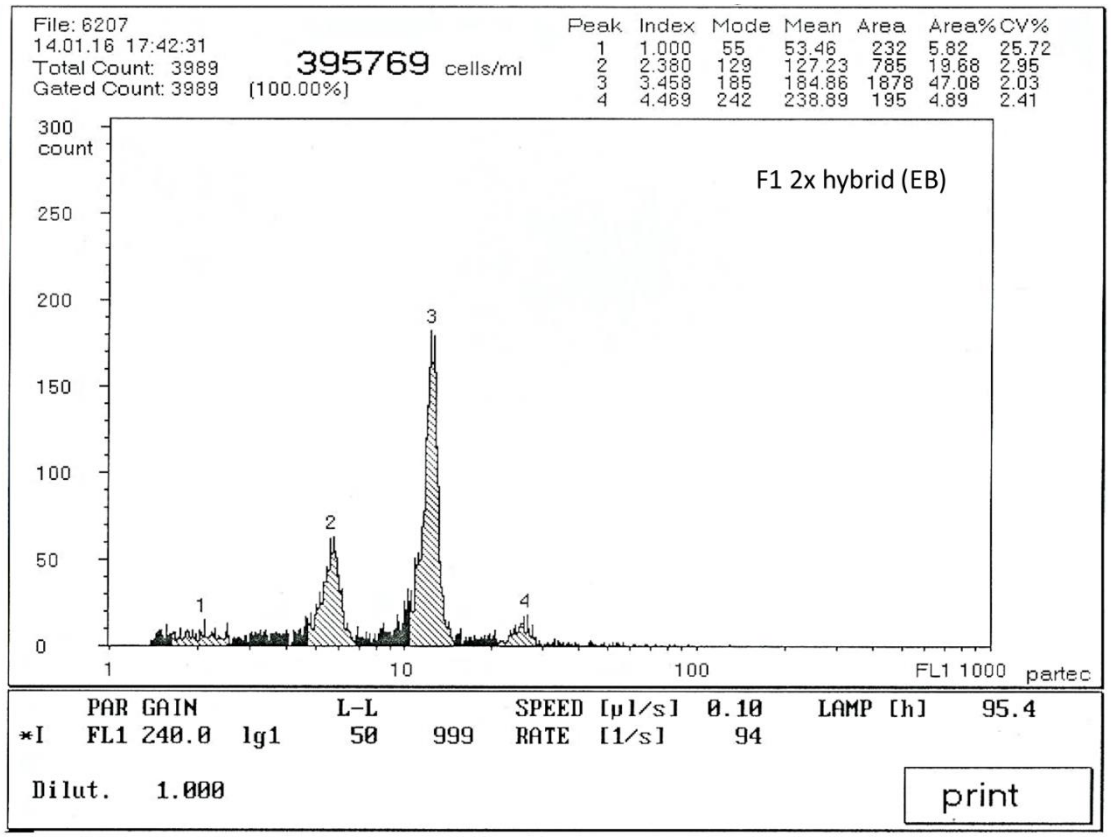

File: 6350
11.03.16 13:52:49
Total Count: 7648
Gated Count: 7648
1156250 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 57 56.80 85 1.11 5.72
2 1.850 106 105.06 4460 58.32 7.85
3 2.877 163 163.41 2017 26.37 4.13
300 count
250
200
150
100
50
0
2x DH (AA)
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 113.2
*I FL1 240.0 lg1 50 999 RATE [1/s] 239
Dilut. 1.000
print

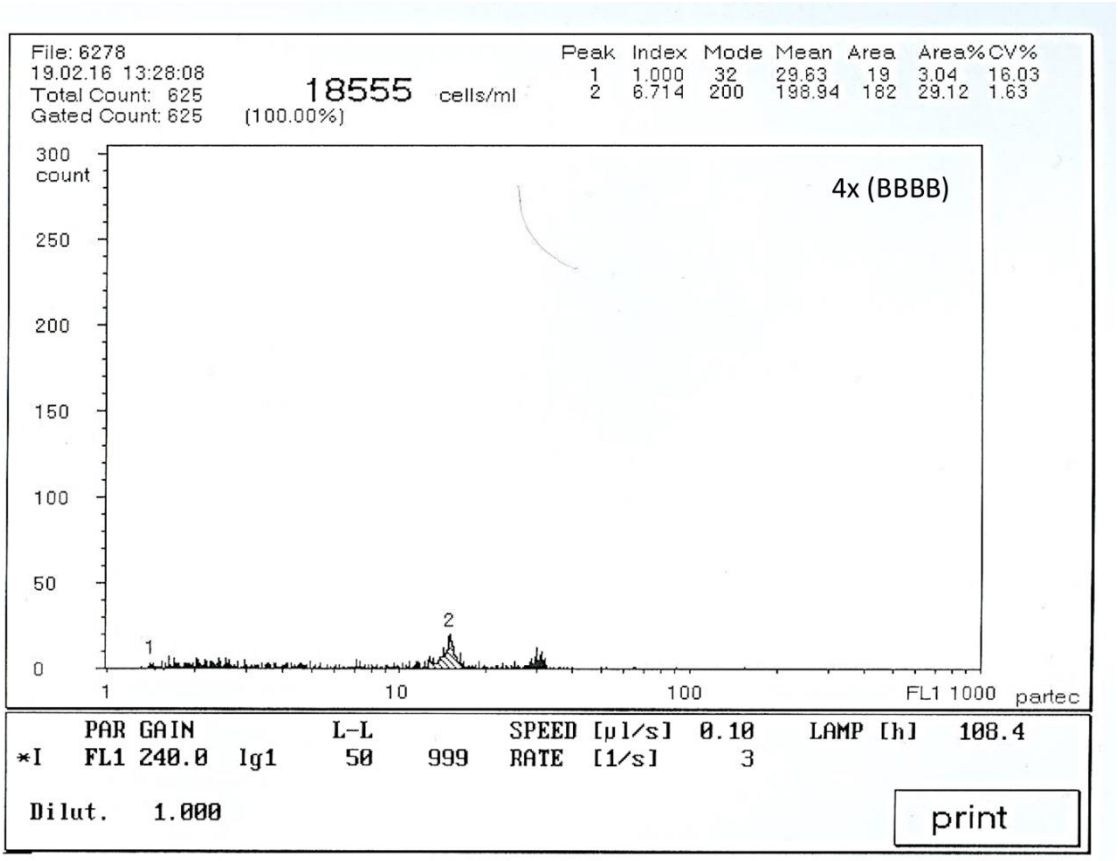

File: 6204
14.01.16 17:20:41
Total Count: 7082
Gated Count: 7082
248061 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 131 129.54 1350 19.06 4.05
2 1.438 188 186.26 3127 44.15 1.21
3 1.852 243 239.91 382 5.39 2.19
600 count
500
400
300
200
100
0
F1 2x hybrid (EA)
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 95.0
*I FL1 240.0 lg1 50 999 RATE [1/s] 60
Dilut. 1.000
print

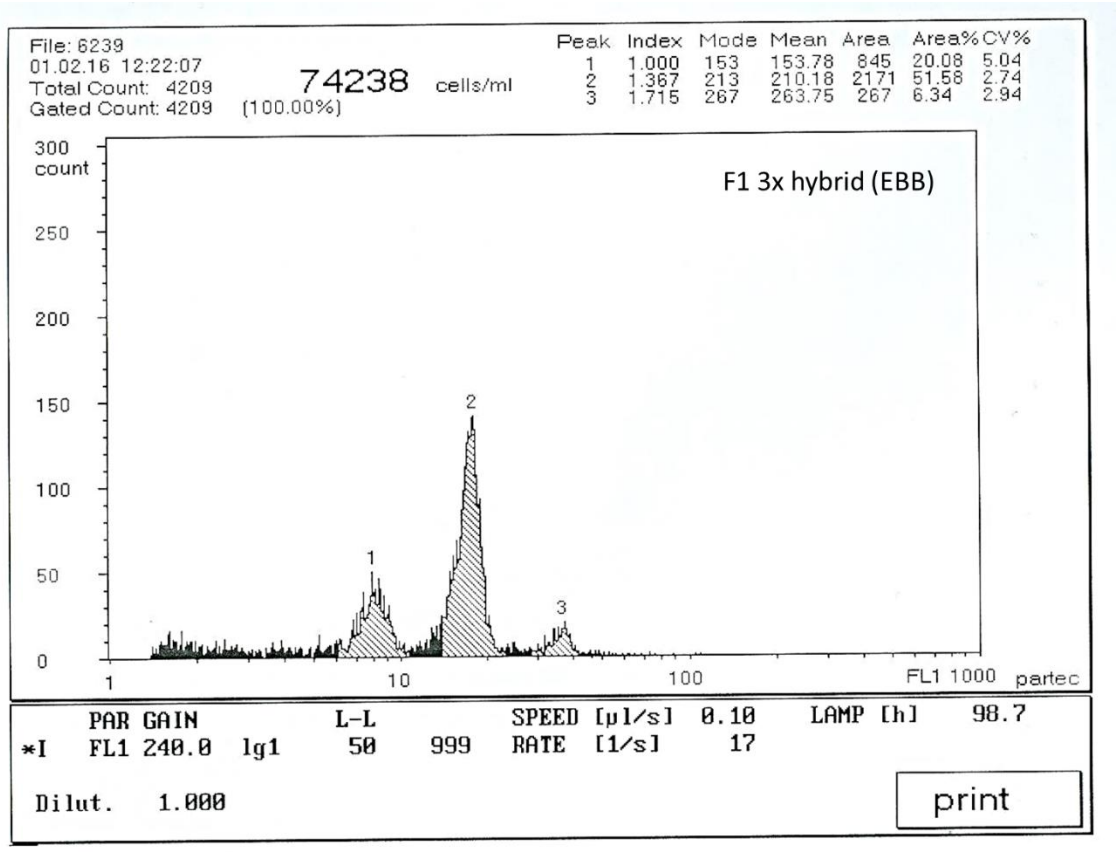

File: 6277
19.02.16 13:21:48
Total Count: 1838
Gated Count: 1838
16429 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 198 196.97 224 12.19 0.89
2 1.274 253 250.99 381 20.73 1.29
3 1.532 304 301.83 101 5.50 1.08
300 count
250
200
150
100
50
0
4x (AAAA)
1
2
3
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 108.3
*I FL1 240.0 lg1 50 999 RATE [1/s] 2
Dilut. 1.000
print

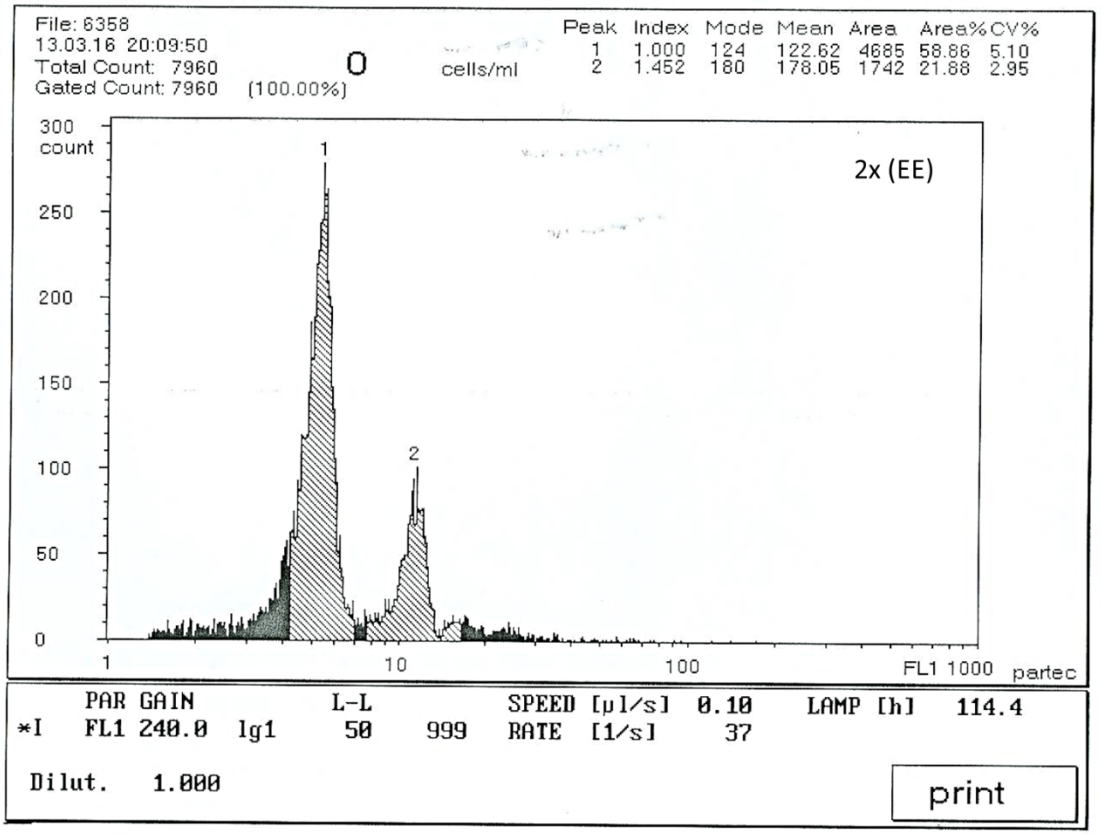

File: 6205
14.01.16 17:24:13
Total Count: 1233
Gated Count: 1233
30852 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 28 31.58 31 2.51 13.46
2 4.723 151 149.16 333 27.01 4.19
3 6.452 206 203.75 580 47.04 1.60
300 count
250
200
150
100
50
0
F1 3x hybrid (EAA)
1
2
3
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 95.1
*I FL1 240.0 lg1 50 999 RATE [1/s] 8
Dilut. 1.000
print

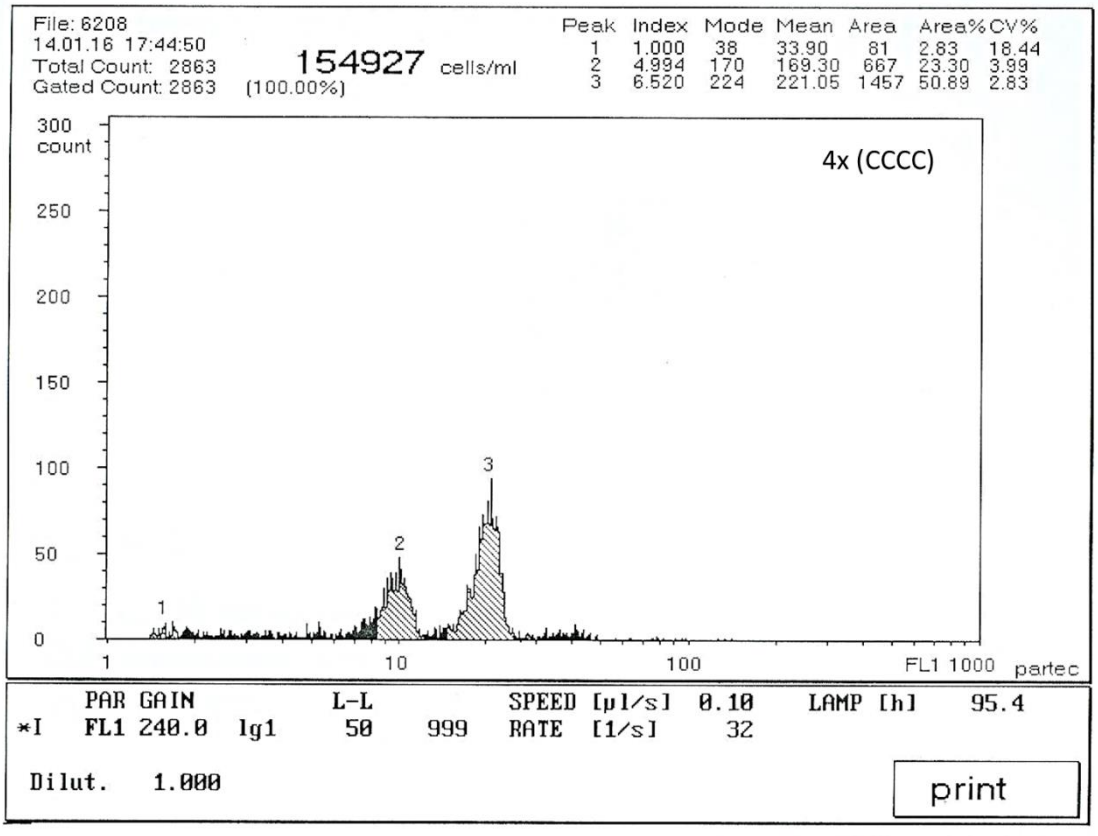

File: 6339
11.03.16 13:09:20
Total Count: 8272
Gated Count: 8272
2526428 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 40 38.83 70 0.85 4.51
2 2.545 104 98.82 5642 68.21 7.34
3 3.974 158 154.30 1611 19.48 3.73
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
2x DH (BB)
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 112.5
*I FL1 240.0 lg1 50 999 RATE [1/s] 278
Dilut. 1.000
print

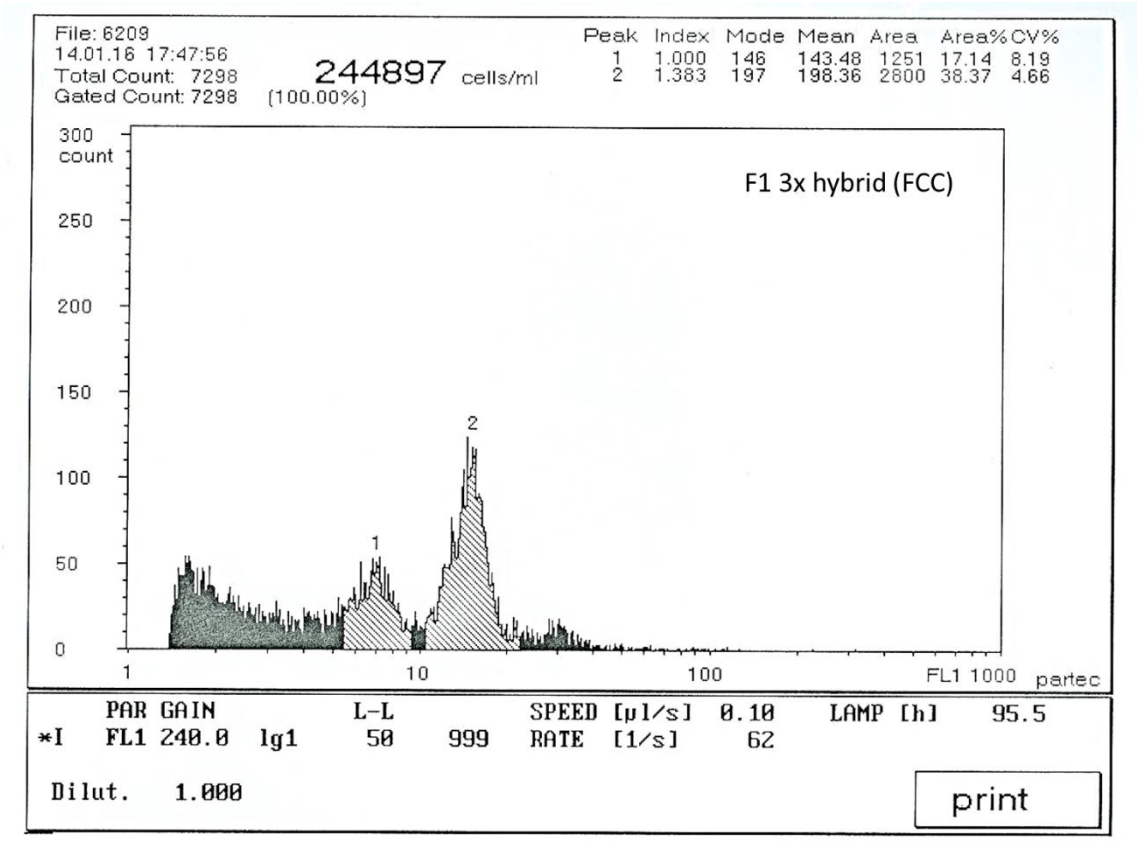

File: 6207
14.01.16 17:42:31
Total Count: 3989
Gated Count: 3989
395769 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 55 53.46 232 5.82 25.72
2 2.380 129 127.23 785 19.68 2.95
3 3.458 185 184.86 1878 47.08 2.03
4 4.469 242 238.89 195 4.89 2.41
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
F1 2x hybrid (EB)
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 95.4
*I FL1 240.0 lg1 50 999 RATE [1/s] 94
Dilut. 1.000
print

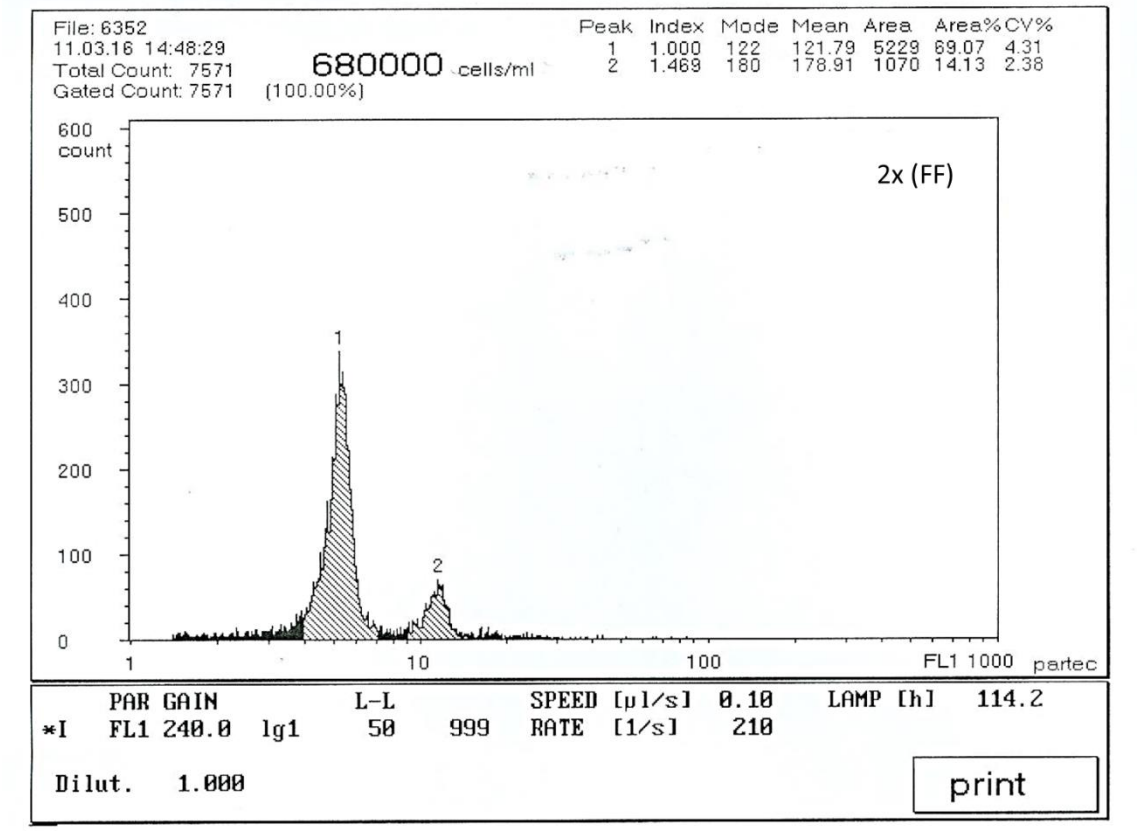

File: 6278
19.02.16 13:28:08
Total Count: 625
Gated Count: 625
18555 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 32 29.63 19 3.04 16.03
2 6.714 200 198.94 182 29.12 1.63
300 count
250
200
150
100
50
0
1
2
4x (BBBB)
1 10 100 FL1 1000 partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 108.4
*I FL1 240.0 lg1 50 999 RATE [1/s] 3
Dilut. 1.000
print

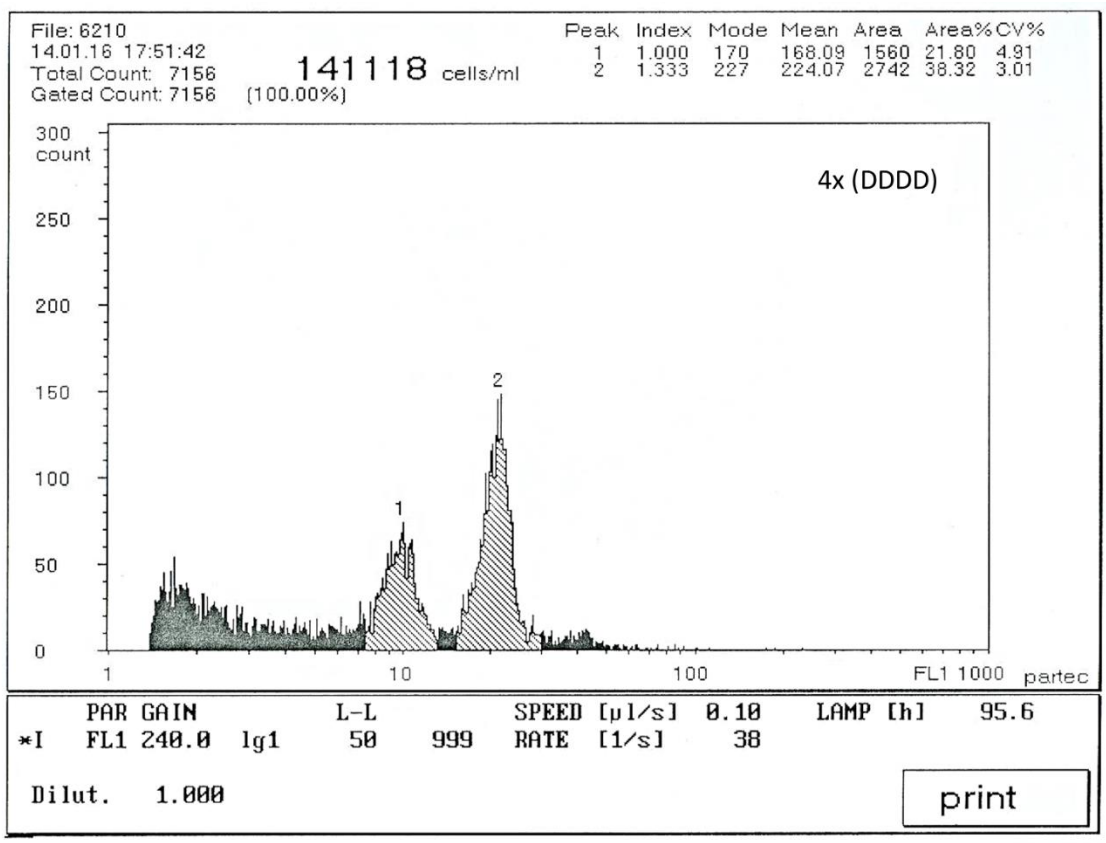

File: 6239
01.02.16 12:22:07
Total Count: 4209
Gated Count: 4209
74238 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 153 153.78 845 20.08 5.04
2 1.367 213 210.18 2171 51.58 2.74
3 1.715 267 263.75 267 6.34 2.94
300 count
250
200
150
100
50
0
1
2
3
F1 3x hybrid (EBB)
1 10 100 FL1 1000 partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 98.7
*I FL1 240.0 lg1 50 999 RATE [1/s] 17
Dilut. 1.000
print

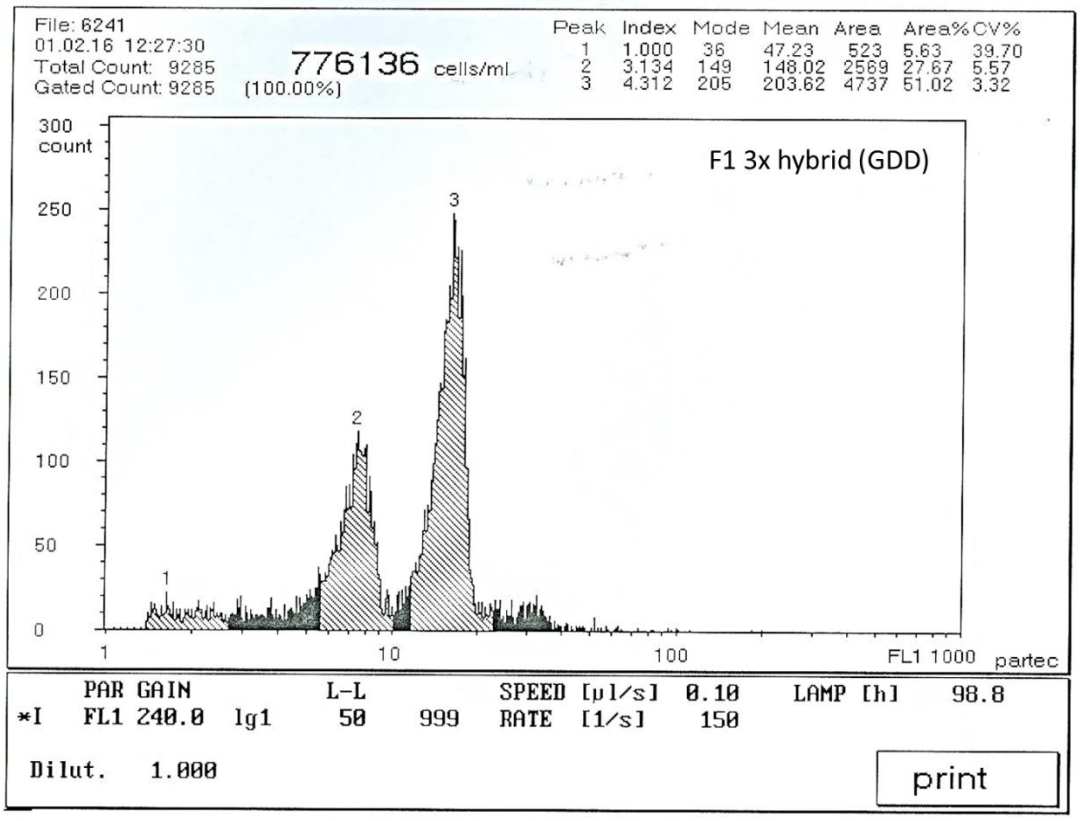

File: 6358
13.03.16 20:09:50
Total Count: 7960
Gated Count: 7960
0
cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 124 122.62 4685 58.86 5.10
2 1.452 180 178.05 1742 21.88 2.95
300
count
250
200
150
100
50
0
1
2
2x (EE)
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 114.4
*I FL1 240.0 lg1 50 999 RATE [1/s] 37
Dilut. 1.000
print

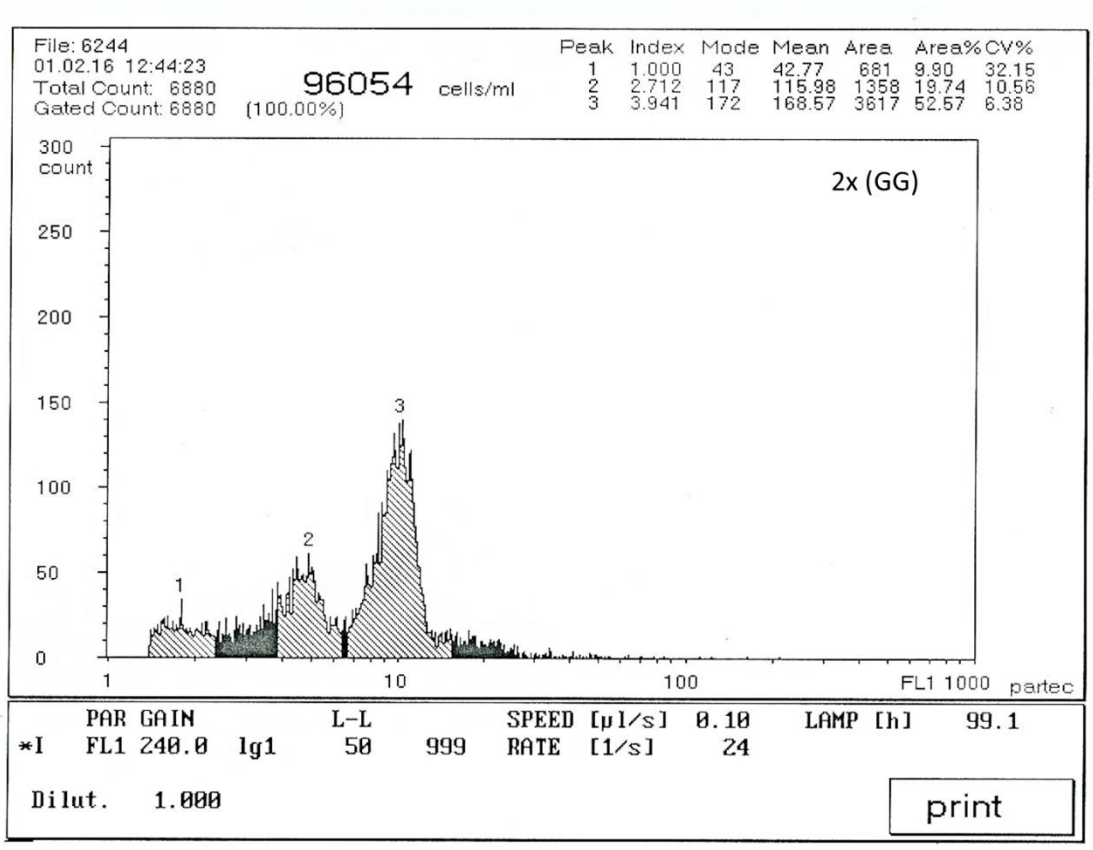

File: 6208
14.01.16 17:44:50
Total Count: 2863
Gated Count: 2863
154927
cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 38 33.90 81 2.83 18.44
2 4.994 170 169.30 667 23.30 3.99
3 6.520 224 221.05 1457 50.89 2.83
300
count
250
200
150
100
50
0
1
2
3
4x (CCCC)
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 95.4
*I FL1 240.0 lg1 50 999 RATE [1/s] 32
Dilut. 1.000
print


File: 6209
14.01.16 17:47:56
Total Count: 7298
Gated Count: 7298
244897 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 146 143.48 1251 17.14 8.19
2 1.383 197 198.36 2800 38.37 4.66
300 count
250
200
150
100
50
0
1
2
F1 3x hybrid (FCC)
1 10 100 FL1 1000 partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 95.5
*I FL1 240.0 lg1 50 999 RATE [1/s] 62
Dilut. 1.000
print


File: 6352
11.03.16 14:48:29
Total Count: 7571
Gated Count: 7571
680000 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 122 121.79 5229 69.07 4.31
2 1.469 180 178.91 1070 14.13 2.38
600 count
500
400
300
200
100
0
1
2
2x (FF)
1 10 100 FL1 1000 partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 114.2
*I FL1 240.0 lg1 50 999 RATE [1/s] 210
Dilut. 1.000
print


File: 6210
14.01.16 17:51:42
Total Count: 7156
Gated Count: 7156
141118 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 170 168.09 1560 21.80 4.91
2 1.333 227 224.07 2742 38.32 3.01
300 count
250
200
150
100
50
0
1
2
4x (DDDD)
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 95.6
*I FL1 240.0 lg1 50 999 RATE [1/s] 38
Dilut. 1.000
print


File: 6241
01.02.16 12:27:30
Total Count: 9285
Gated Count: 9285
776136 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 36 47.23 523 5.63 39.70
2 3.134 149 148.02 2569 27.67 5.57
3 4.312 205 203.62 4737 51.02 3.32
300 count
250
200
150
100
50
0
1
2
3
F1 3x hybrid (GDD)
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 0.10 LAMP [h] 98.8
*I FL1 240.0 lg1 50 999 RATE [1/s] 150
Dilut. 1.000
print


File: 6244
01.02.16 12:44:23
Total Count: 6880
Gated Count: 6880
96054 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 43 42.77 681 9.90 32.15
2 2.712 117 115.98 1358 19.74 10.56
3 3.941 172 168.57 3617 52.57 6.38
300
count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
2x (GG)
PAR GAIN L-L SPEED [pl/s] 0.10 LAMP [h] 99.1
*I FL1 240.0 lg1 50 999 RATE [1/s] 24
Dilut. 1.000
print

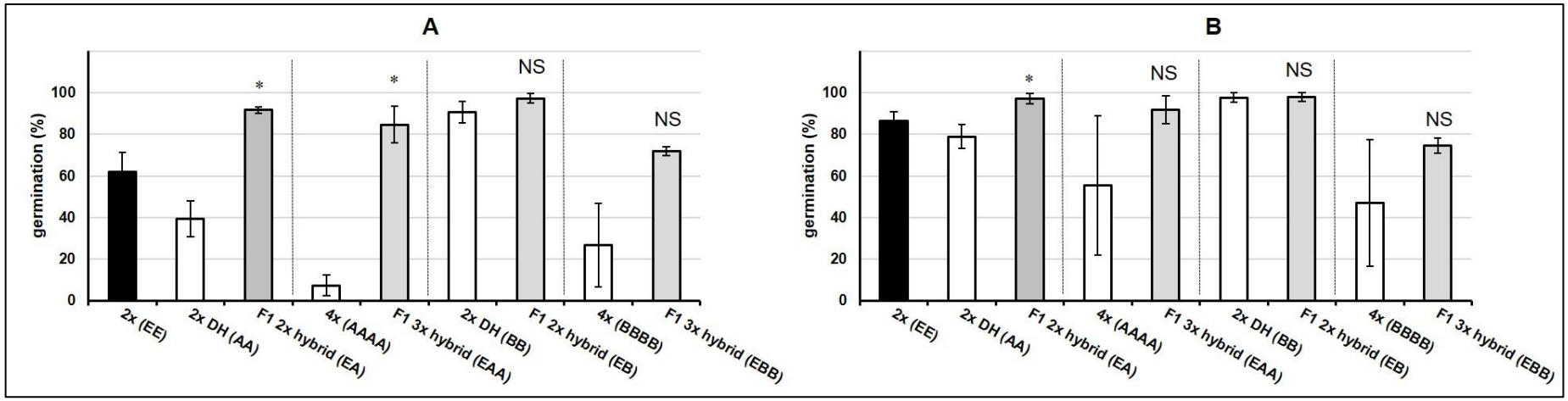


Figure S2. Germination percentage of F1 diploid and triploid hybrids and their parent lines. Data are mean of three replicates (± SD). **A** Day 4, **B** day 14. Different genotypes specified in parentheses. DH = double haploid. * Best parent heterosis ($P \leq 0.05$), NS Not significantly different ($P > 0.05$).

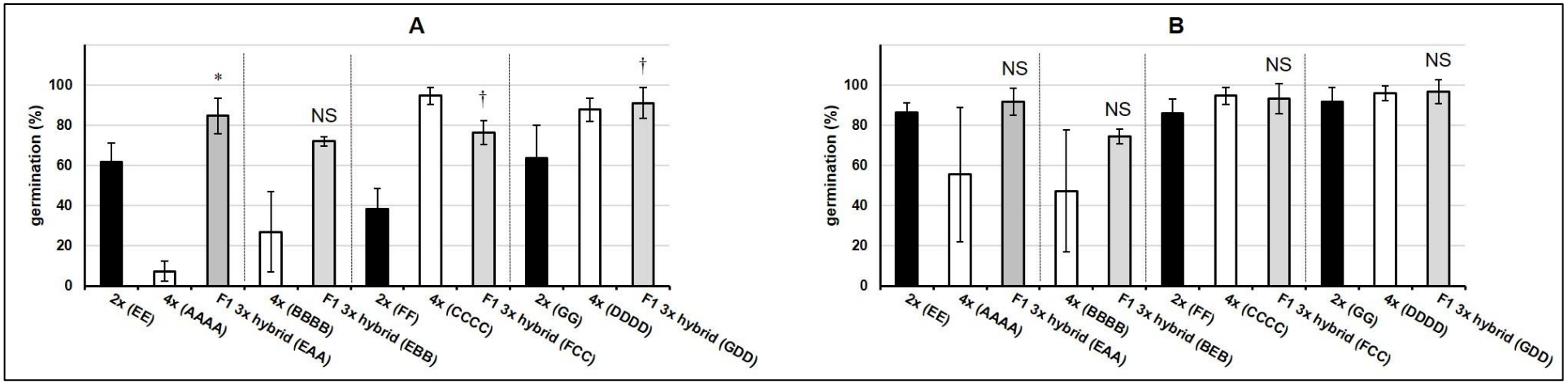


Figure S3. Germination percentage of F1 triploid hybrids and their parent lines. Data are mean of three replicates (± SD). **A** Day 4, **B** day 14. Different genotypes specified in parentheses. * Best parent heterosis ($P \leq 0.05$), † Mid-parent heterosis ($P \leq 0.05$), NS Not significantly different ($P > 0.05$).

Table S1. Fruit and seed characteristics of F1 diploid and triploid hybrids and parent lines. Data are mean of four replicates (± SE).

Trait	♀ 2x (EE)	♂ 2x DH (AA)	F1 2x hybrid (EA)	♂ 4x (AAAA)	F1 3x hybrid (EAA)	♂ 2x DH (BB)	F1 2x hybrid (EB)	♂ 4x (BBBB)	F1 3x hybrid (EBB)
Monogermity (%)	89.00 (± 3.00)	2.00 (± <0.01)	99.04 (± 0.96)	7.92 (± 5.44)	50.89 (± 26.06)	0.00 (± 0.00)	99.00 (± 1.00)	0.00 (± 0.00)	77.00 (± 21.69)
Fruit weight (mg)	7.45 (± 0.23)	13.12 (± 0.59)	7.16 (± 0.27)	22.14 (± 0.90)	8.50 (± 0.29)	9.17 (± 0.35)	6.59 (± 0.24)	12.87 (± 0.58)	9.96 (± 0.32)
Seed viability (%)	89.26 (± 5.91)	60.56 (± 1.10)	98.08 (± 0.96) †	22.49 (± 5.56)	75.51 (± 4.44) †	70.73 (± 2.45)	96.00 (± 1.92) †	14.75 (± 1.97)	73.20 (± 3.78) †
Seed weight (mg)	3.39 (± 0.08)	2.45 (± 0.05)	3.40 (± 0.22) †	2.68 (± 0.14)	3.31 (±0.16) †	2.63 (± 0.09)	3.15 (± 0.012) †	2.23 (± 0.42)	3.45 (± 0.16) †
Seed:Fruit weight ratio	0.47 (± 0.02)	0.23 (± 0.02)	0.50 (± 0.04) *	0.09 (± 0.01)	0.39 (± 0.10) ^{NS}	0.33 (± 0.02)	0.49 (± 0.05) †	0.18 (± 0.04)	0.36 (± 0.02) ǂ
Embryo size (mm^2)	4.04 (± 0.73)	3.86 (± 0.94)	4.67 (± 1.15) *	4.17 (± 0.91)	4.67 (± 1.31) *	4.10 (± 1.06)	4.42 (± 1.13) *	4.71 (± 1.15)	4.83 (0.95) †
Perisperm size (mm^2)	1.38 (± 0.57)	1.20 (± 0.41)	1.72 (± 0.53) *	1.40 (± 0.61)	1.64 (± 0.58) *	1.47 (± 0.48)	1.48 (± 0.49) ^{NS}	1.88 (± 0.95)	1.59 (± 0.49) ^{NS}

DH = double haploid. Different genotypes specified in parentheses. * Best parent heterosis ($P \leq 0.05$), † Mid-parent heterosis ($P \leq 0.05$), ^{NS} Not significantly different ($P > 0.05$), ǂ Below mid-parent value ($P \leq 0.05$)

Table S2. Agronomic and root quality traits of F1 diploid and triploid hybrids and their parent lines. Data are mean of four replicates (± SE).

Agronomic Traits	♀ 2x (EE)	♂ 2x DH (AA)	F1 2x hybrid (EA)	♂ 4x (AAAA)	F1 3x hybrid (EAA)	♂ 2x DH (BB)	F1 2x hybrid (EB)	♂ 4x (BBBB)	F1 3x hybrid (EBB)
Harvest plant density [1]	46.5 (± 1.71)	39.75 (± 2.78)	46.5 (± 2.56) NS	19.5 (± 3.20)	43.75 (± 1.09) †	43.75 (± 1.12)	46.5 (± 1.50) NS	18.5 (± 4.72)	44.5 (± 2.06) †
Total yield (T/ha)	121.79 (± 13.46)	91.96 (± 8.68)	166.64 (± 5.42) *	42.84 (± 10.39)	142.68 (± 11.77) †	98.75 (± 17.86)	140.16 (± 7.70) †	33.68 (± 4.74)	135.47 (± 11.83) †
Above-ground biomass (g) [2]	515.2 (± 63.79)	576.04 (± 34.20)	718.71 (± 62.34) *	463.63 (± 58.42)	684.6 (± 7.79) *	464.08 (± 58.15)	575.48 (± 47.20) NS	352.13 (± 28.90)	608.8 (± 67.97) †
Root length (cm)	17.25 (± 0.58)	18.23 (± 1.06)	18.73 (± 0.31) NS	15.85 (± 0.95)	18.88 (± 0.60) *	15.74 (± 0.41)	18.80 (± 0.80) †	14.73 (± 0.42)	17.91 (± 0.61) †
Root circumference (cm)	32.04 (± 1.36)	31.42 (± 0.78)	37.54 (± 0.65) *	29.98 (± 1.48)	37.32 (± 0.65) *	32.31 (± 1.97)	36.72 (± 0.73) *	29.52 (± 1.75)	35.73 (± 1.24) *
Root yield (T/ha)	71.44 (± 7.13)	51.60 (± 5.84)	103.06 (± 5.84) *	24.87 (± 6.19)	85.81 (± 7.92) †	56.06 (± 10.92)	88.63 (± 4.79) *	20.24 (± 3.26)	82.03 (± 6.90) †
Root Quality Traits									
Corrected sugar content (%)	14.74 (± 0.40)	13.23 (± 0.27)	14.13 (± 0.08) NS	12.84 (± 0.30)	13.87 (± 0.09) NS	13.16 (± 0.14)	14.68 (± 0.10) NS	12.87 (± 0.16)	13.98 (± 0.24) NS
Standard molasses loss (%)	1.85 (± 0.07)	1.99 (± 0.07)	1.80 (± 0.03) ≛	2.19 (± 0.07)	1.86 (± 0.02) †	1.85 (± 0.07)	1.71 (± 0.03) NS	1.93 (± 0.10)	1.81 (± 0.06) NS
Corrected sugar yield (T/ha)	10.44 (± 0.78)	6.86 (± 0.91)	14.57 (± 0.88) *	3.14 (± 0.74)	11.89 (± 1.04) †	7.34 (± 1.36)	13.00 (± 0.65) *	2.59 (± 0.40)	11.60 (± 1.23) †

DH = double haploid. Different genotypes specified in parentheses. * Best parent heterosis ($P \leq 0.05$), † Mid-parent heterosis ($P \leq 0.05$), NS Not significantly different ($P > 0.05$), ≛ Below best parent value ($P \leq 0.05$), ǂ Below mid-parent value ($P \leq 0.05$)

Table S3. Fruit and seed characteristics of FI triploid hybrids and parent lines. Data are mean of four replicates (± SE).

Trait	♀ 2x (EE)	♂ 4x (AAAA)	F1 3x hybrid (EAA)	♂ 4x (BBBB)	F1 3x hybrid (EBB)	♀ 2x (FF)	♂ 4x (CCCC)	F1 3x hybrid (FCC)	♀ 2x (GG)	♂ 4x (DDDD)	F1 3x hybrid (GDD)
Monogermity (%)	89.00 (± 3.00)	7.92 (± 5.44)	50.89 (± 26.06)	0.00 (± 0.00)	77.00 (± 21.69)	95.08 (± 0.92)	0.00 (± 0.00)	100 (± 0.00)	71.00 (± 11.71)	3.92 (± 1.57)	100 (± 0.00)
Fruit weight (mg)	7.45 (± 0.23)	22.14 (± 0.90)	8.50 (± 0.29)	12.87 (± 0.58)	9.96 (± 0.32)	17.90 (± 0.70)	21.14 (± 0.66)	10.08 (± 0.21)	9.78 (± 0.47)	22.45 (± 0.85)	7.81 (± 0.11)
Seed viability (%)	89.26 (± 5.91)	22.49 (± 5.56)	75.51 (± 4.44) †	14.75 (± 1.97)	73.20 (± 3.78) †	90.46 (± 2.41)	74.84 (± 2.91)	100 (± 0.00) *	87.98 (± 1.49)	75.75 (± 5.64)	100 (± 0.00) *
Seed weight (mg)	3.39 (± 0.08)	2.68 (± 0.14)	3.31 (± 0.16) †	2.23 (± 0.42)	3.45 (± 0.16) †	3.29 (± 0.20)	3.57 (± 0.23)	4.39 (± 0.09) *	2.86 (± 0.12)	4.67 (± 0.20)	3.17 (± 0.11) ǂ
Seed:Fruit weight ratio	0.47 (± 0.02)	0.09 (± 0.01)	0.39 (± 0.10) ^{NS}	0.18 (± 0.04)	0.36 (± 0.02) ǂ	0.15 (± 0.01)	0.22 (± 0.01)	0.44 (± <0.01) *	0.35 (± 0.01)	0.23 (± 0.04)	0.52 (± <0.01) *
Embryo size (mm^2)	4.04 (± 0.73)	4.17 (± 0.91)	4.67 (± 1.31) *	4.71 (± 1.15)	4.83 (± 0.95) †	4.80 (± 1.23)	4.35 (± 1.49)	5.71 (± 1.00) *	3.75 (± 0.88)	5.12 (± 1.20)	4.10 (± 0.64) ǂ
Perisperm size (mm^2)	1.38 (± 0.57)	1.40 (± 0.61)	1.64 (± 0.58) *	1.88 (± 0.95)	1.59 (± 0.49) ^{NS}	1.72 (± 0.57)	1.56 (± 0.67)	2.00 (± 0.57) *	1.22 (± 0.44)	1.92 (± 0.63)	1.51 (± 0.41) ^{NS}

Different genotypes specified in parentheses. * Best parent heterosis ($P \leq 0.05$), † Mid-parent heterosis ($P \leq 0.05$), ^{NS} Not significantly different ($P > 0.05$), ǂ Below mid-parent value ($P \leq 0.05$)

Table S4. Agronomic and root quality traits of F1 triploid hybrids and their parent lines. Data are mean of four replicates (± SE).

Agronomic Traits	♀ 2x (EE)	♂ 4x (AAAA)	F1 3x hybrid (EAA)	♂ 4x (BBBB)	F1 3x hybrid (EBB)	♀ 2x (FF)	♂ 4x (CCCC)	F1 3x hybrid (FCC)	♀ 2x (GG)	♂ 4x (DDDD)	F1 3x hybrid (GDD)
Harvest plant density [1]	46.5 (± 1.71)	19.5 (± 3.20)	43.75 (± 1.09) †	18.5 (± 4.72)	44.5 (± 2.06) †	41.75 (± 2.50)	44.75 (± 1.49)	42.5 (± 1.44) NS	43.25 (± 1.60)	44.75 (± 0.25)	44.5 (± 1.19) NS
Total yield (T/ha)	121.79 (± 13.46)	42.84 (± 10.39)	142.68 (± 11.77) †	33.68 (± 4.74)	135.47 (± 11.83) †	145.56 (± 11.32)	144.98 (± 13.08)	150.07 (± 7.02) NS	128.10 (± 12.72)	114.55 (± 3.87)	160.37 (± 9.91) *
Above-ground biomass (g) [2]	515.2 (± 63.79)	463.63 (± 58.42)	684.6 (± 7.79) *	352.13 (± 28.90)	608.8 (± 67.97) †	935.6 (± 18.22)	629.33 (± 52.42)	779.4 (± 47.07) ±	795.4 (± 108.98)	532.83 (± 19.65)	764.48 (± 68.07) NS
Root length (cm)	17.25 (± 0.58)	15.85 (± 0.95)	18.88 (± 0.60) *	14.73 (± 0.42)	17.91 (± 0.61) †	17.97 (± 0.19)	19.34 (± 0.16)	18.15 (± 0.56) ±	16.67 (± 0.35)	16.51 (± 0.57)	17.64 (± 0.52) NS
Root circumference (cm)	32.04 (± 1.36)	29.98 (± 1.48)	37.32 (± 0.65) *	29.52 (± 1.75)	35.73 (± 1.24) *	34.33 (± 0.96)	33.71 (± 1.05)	34.94 (± 0.91) NS	33.35 (± 2.11)	29.29 (± 0.53)	34.54 (± 1.05) NS
Root yield (T/ha)	71.44 (± 7.13)	24.87 (± 6.19)	85.81 (± 7.92) †	20.24 (± 3.26)	82.03 (± 6.90) †	70.87 (± 7.18)	87.92 (± 8.72)	84.46 (± 4.97) NS	64.90 (± 6.54)	64.31 (± 2.82)	87.51 (± 5.78) *
Root Quality Traits											
Corrected sugar content (%)	14.74 (± 0.40)	12.84 (± 0.30)	13.87 (± 0.09) NS	12.87 (± 0.16)	13.98 (± 0.24) NS	15.40 (± 0.19)	14.36 (± 0.33)	14.77 (± 0.26) NS	14.85 (± 0.31)	14.71 (± 0.07)	14.87 (± 0.39) NS
Standard molasses loss (%)	1.85 (± 0.07)	2.19 (± 0.07)	1.86 (± 0.02) †	1.93 (± 0.10)	1.81 (± 0.06) NS	1.73 (± 0.05)	1.90 (± 0.09)	1.77 (± 0.03) NS	1.84 (± 0.06)	1.68 (± 0.05)	1.69 (± 0.06) NS
Corrected sugar yield (T/ha)	10.44 (± 0.78)	3.14 (± 0.74)	11.89 (± 1.04) †	2.59 (± 0.40)	11.60 (± 1.23) †	10.62 (± 1.47)	12.53 (± 0.92)	12.44 (± 0.55) NS	9.57 (± 0.85)	9.46 (± 0.40)	12.99 (± 0.81) *

Different genotypes specified in parentheses. * Best parent heterosis ($P \leq 0.05$), † Mid-parent heterosis ($P \leq 0.05$), NS Not significantly different ($P > 0.05$), ± Below best parent value ($P \leq 0.05$), ǂ Below mid-parent value ($P \leq 0.05$)

Appendix 2

Chapter 4. Supplementary Materials and Methods.

High-throughput DNA extraction method

A ball-bearing (Steel Shot Company Ltd., Hull, UK) was placed in each well of a 96-well plate (Thermo Scientific, MA, USA). 200µl of extraction buffer (500mM NaCl + 1% SDS) was added to the wells with a multi-channel pipette. A piece of leaf tissue (~3mm^2) was removed from each sample and placed in a well. The plate was capped and put in a QIAGEN shaker (Hilden, Germany) at 30hz for 1 minute. The plate was turned upside down and returned to the shaker for an additional 1 minute. The plate was centrifuged for 15 minutes at 4000rpm (model 5810 R, Eppendorf, Hamburg, Germany). Next, 100µl of supernatant was transferred to a new plate with the multi-channel pipette and 100µl of ice-cold isopropanol (Merck, NJ, USA) was added. The new plate was spun as before, and the supernatant was gently removed by briefly turning upside down. 100µl of 70% ethanol was added. After 1 minute, the supernatant is gently removed as before. The plate is spun upside down on a piece of tissue in the centrifuge for 20 seconds at 300rpm. The plate is left to air dry for approx. 5mins or until all the ethanol has visibly evaporated. The DNA is resuspended in 40µl of distilled water.

Genotyping of samples via PCR

The *tailswap* transgene, WT histone tail, *AT4G01310,* and *cenh3-1* mutation can be amplified under the same PCR conditions. Amplification for a single reaction was as follows:

Reagent	Volume in µl	Volume in µl	Volume in µl	Volume in µl
2X BIOLINE My Taq Red Mix	5.0	5.0	5.0	5.0
GFP-tailswap F	0.1			
GFP-tailswap R	0.1	0.1		
GFP-tailswap T		0.1		
cenh3-1 F			0.1	
cenh3-1 R			0.1	
SL10 F				0.1
SL10 R				0.1
H_2O	2.8	2.8	2.8	2.8
DNA sample	2.0	2.0	2.0	2.0

PCR was performed in a 10µl reaction with a heated lid at 105°C.

Step 1 (once): 94°C for 1 minute.

Step 2 (repeated 35 times): 94°C for 30 seconds, 60°C for 30 seconds, 72°C for 30seconds.

Step 3 (once): 72°C for 5 minutes, end.

Restriction digestion of the cenh3-1 amplification product is necessary to identify plants homozygous for the mutant allele (true HI plants). Restriction digestion of the *AT4G01310* product is necessary to identify plants carrying the Col or Zu nuclear genome. Digestion for a single product is as follows:

Reagent	**Volume in µl**	**Volume in µl**
CutSmart Buffer	0.5	0.5
XbaI enzyme	0.1	
SacI enzyme		0.1
H_2O	4.4	4.4
PCR product	5.0	5.0

Samples are incubated @ 37°C overnight in a thermocycler with a heated lid.

List of primers

Name	**Sequence (5' → 3')**	**Purpose**
GFP-tailswap F	CTGAAGCTGAACCTTCGTCTCG	WT histone tail amplified by F + R primers, transgene amplified by T + R primers, from Ravi et al. (2014)
GFP-tailswap R	CACATACTCGCTACTGGTCAGAGAATC	
GFP-tailswap T	AATCCAGATCCCCCGAATTA	

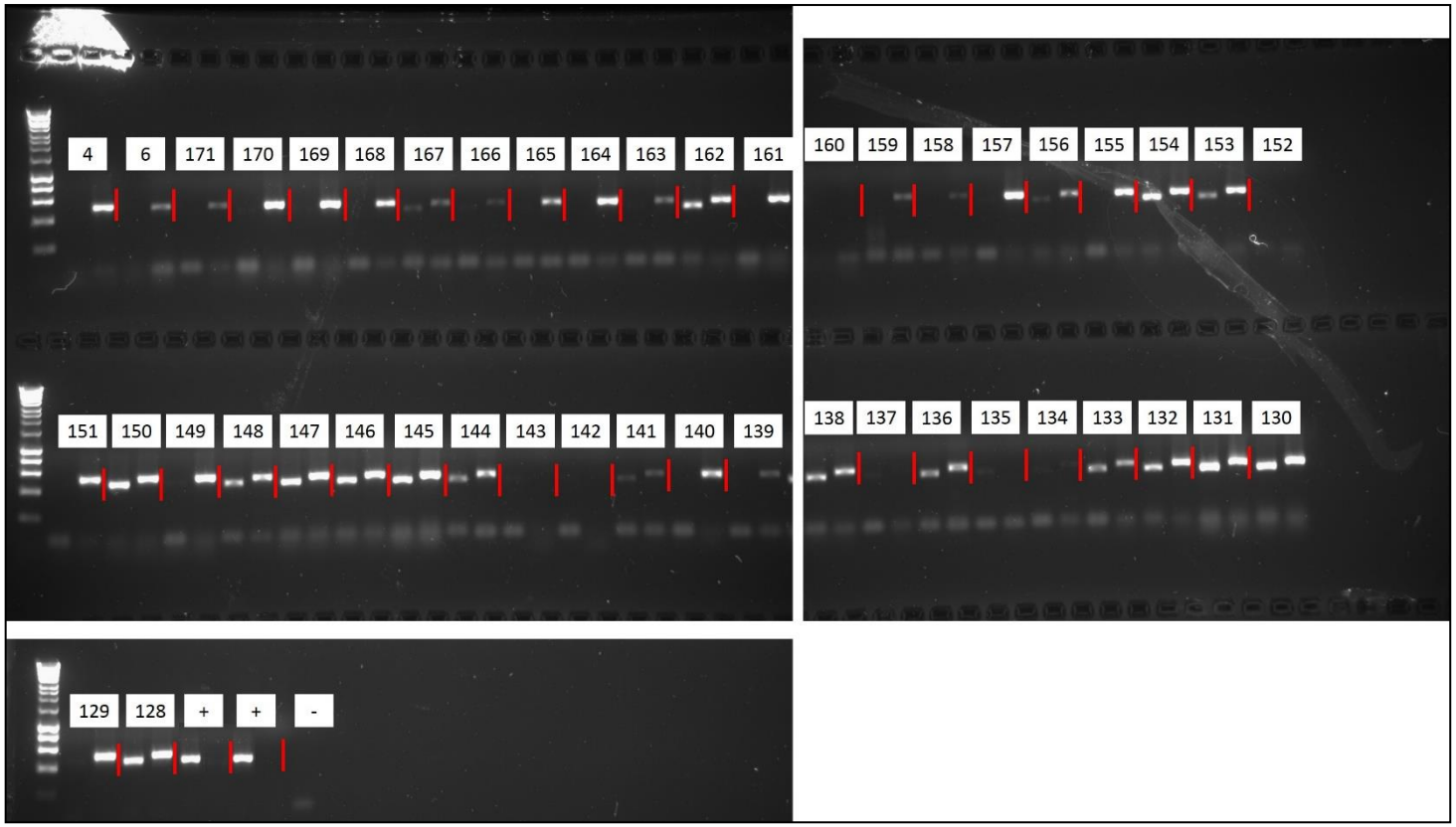
cenh3-1 F cenh3-1 R	GGTGCGATTTCTCCAGCAGTAAAAATC CTGAGAAGATGAAGCACCGGCGATAT	Amplification of region spanning the *cenh3-1* point mutation, from Ravi et al. (2014). Product is cleaved with XbaI.
SL10 F SL10 R	ATCCTCTCGAGGTAAGCGGT TCTTCTCAGGTCGGTGTGGA	Amplification of *AT4G01310*. Product is cleaved with SacI.

References

Ravi, M., M.P.A. Marimuthu, E.H. Tan, S. Maheshwari, I.M. Henry *et al.*, 2014 A haploid genetics toolbox for Arabidopsis thaliana. *Nature Communications* 5.

Chapter 4. Supplementary Results.


4 6 21 22 24 25 26 30 31 33 34 35 38 39
40 41 42 43 48 49 50 65 66 67
73 75 76 77 78 79 80 81 82 83 86 87 88 89
90 93 94 95 97 106 107 108 117 119
120 121 122 123 124 125 126 127 + + -
1 5 11 12 13 14 16 17 18 19 20 25 27 28
29 32 36 37 44 45 46 52 53 54
55 56 57 58 59 60 61 62 63 64 68 69 70 71
72 74 84 85 91 92 98 99 100 101
102 104 105 109 110 112 113 114 115 116 + + -



Supplementary Figure 1. Gel electrophoresis images of genotyped samples following HI *X* WT crosses. Extracted DNA from each sample was genotyped for the presence of the HI transgene and the WT histone tail. Sample numbers are indicated in text box. The first well of each sample is the PCR product for amplification of the HI transgene, the second well is the PCR product for amplification of the WT histone tail. The positive controls for the WT histone tail are sample numbers 1, 4, 5, and 6. The positive controls for the transgene are the HI lines, indicated by '+' sign. The negative control is water in replace of DNA in the PCR, indicated by '-'sign.

Supplementary Table 5. Record of genotyped samples following HI *X* WT crosses.

Sample no.	Plant ID	Amplified band present?	
		***tailswap* transgene**	**WT tail**
1	WT Col	×	✓
2	WT Col		
3	WT Col		
4	WT Col	×	✓
5	WT Zu	×	✓
6	WT Zu	×	✓
7	WT Zu		
8	WT Zu		
9	WT Zu		
10			
11	HI Zu/Col^{Zu} *X* WT Col (3)	×	✓
12	HI Zu/Col^{Zu} *X* WT Col (3)	✓	×
13	HI Zu/Col^{Zu} *X* WT Col (3)	×	✓
14	HI Zu/Col^{Zu} *X* WT Col (3)		
15			
16	HI Zu/Col^{Zu} *X* WT Col (1)	✓	×
17	HI Zu/Col^{Zu} *X* WT Col (2)	✓	✓
18	HI Zu/Col^{Zu} *X* WT Col (2)	✓	✓
19	HI Zu/Col^{Zu} *X* WT Col (2)	✓	×
20	HI Zu/Col^{Zu} *X* WT Col (2)	×	✓
21	HI Col^{Col} *X* WT Col (3)	×	✓
22	HI Col^{Col} *X* WT Col (3)	×	✓
23	HI Col^{Col} *X* WT Col (3)	×	✓
24	HI Col^{Col} *X* WT Col (3)	×	✓
25	HI Col^{Col} *X* WT Col (3)	×	✓
26	HI Col^{Col} *X* WT Col (2)	×	✓
27	HI Col^{Col} *X* WT Col (2)	✓	×
28	HI Col^{Col} *X* WT Col (2)	✓	✓
29	HI Col^{Col} *X* WT Col (1)	×	✓
30	HI Col^{Col} *X* WT Col (2)	✓	×
31	HI Col^{Col} *X* WT Col (2)	×	✓
32	HI Col^{Col} *X* WT Col (1)	×	✓

33	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
34	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
35	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
36	HI Zu/Col^{Zu} *X* WT Zu (1)	×	✓
37	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
38	HI Zu/Col^{Zu} *X* WT Zu (1)		
39	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
40	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
41	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
42	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
43	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
44	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
45	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
46	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
47			
48	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
49	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
50	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
51			
52	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
53	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
54	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
55	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
56	HI Zu/Col^{Zu} *X* WT Zu (2)	×	✓
57	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
58	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
59	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
60	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
61	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
62	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
63	HI Zu/Col^{Zu} *X* WT Zu (2)		
64	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
65	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
66	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓

67	HI Zu/Col^{Zu} *X* WT Zu (2)	×	✓
68	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	×
69	HI Zu/Col^{Zu} *X* WT Zu (2)		
70	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
71	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
72	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
73	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
74	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
75	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
76	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	×
77	HI Zu/Col^{Zu} *X* WT Zu (3)	×	✓
78	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
79	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
80	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
81	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
82	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	×
83	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
84	HI Zu/Col^{Zu} *X* WT Zu (3)		
85	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
86	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
87	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
88	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
89	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
90	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	×
91	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
92	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
93	HI Col^{Col} *X* WT Zu (1)	×	✓
94	HI Col^{Col} *X* WT Zu (1)	×	✓
95	HI Col^{Col} *X* WT Zu (1)	×	✓
96			
97	HI Col^{Col} *X* WT Zu (1)	✓	✓
98	HI Col^{Col} *X* WT Zu (3)	×	✓
99	HI Col^{Col} *X* WT Zu (3)	✓	×
100	HI Col^{Col} *X* WT Zu (3)	✓	✓

101	HI Col^{Col} *X* WT Zu (3)		
102	HI Col^{Col} *X* WT Zu (3)	×	✓
103			
104	HI Col^{Col} *X* WT Zu (2)	×	✓
105	HI Col^{Col} *X* WT Zu (2)		
106	HI Col^{Col} *X* WT Zu (2)	✓	✓
107	HI Col^{Col} *X* WT Zu (2)	✓	✓
108	HI Col^{Col} *X* WT Zu (2)	×	✓
109	HI Col^{Col} *X* WT Zu (2)	✓	✓
110	HI Col^{Col} *X* WT Zu (2)	✓	✓
111			
112	HI Col^{Col} *X* WT Zu (3)		
113	HI Col^{Col} *X* WT Zu (3)	✓	×
114	HI Col^{Col} *X* WT Zu (3)	×	✓
115	HI Col^{Col} *X* WT Zu (3)	×	✓
116	HI Col^{Col} *X* WT Zu (3)	✓	✓
117	HI Col^{Col} *X* WT Zu (3)	✓	✓
118			
119	HI Col^{Col} *X* WT Zu (3)	✓	✓
120	HI Col^{Col} *X* WT Zu (3)	×	✓
121	HI Zu/Col^{Zu} *X* WT Col (3)	✓	✓
122	HI Zu/Col^{Zu} *X* WT Col (3)	✓	✓
123	HI Zu/Col^{Zu} *X* WT Col (3)	×	✓
124	HI Zu/Col^{Zu} *X* WT Col (3)	×	✓
125	HI Zu/Col^{Zu} *X* WT Col (3)	×	✓
126	HI Zu/Col^{Zu} *X* WT Col (1)	✓	✓
127	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
128	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
129	HI Col^{Col} *X* WT Zu (2)	×	✓
130	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
131	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
132	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
133	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
134			

135			
136	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
137			
138	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
139	HI Col^{Col} *X* WT Zu (2)	×	✓
140	HI Col^{Col} *X* WT Col (1)	×	✓
141	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
142			
143			
144	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
145	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
146	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
147	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
148	HI Zu/Col^{Zu} *X* WT Zu (1)	✓	✓
149	HI Col^{Col} *X* WT Col (1)	×	✓
150	HI Zu/Col^{Zu} *X* WT Zu (2)	✓	✓
151	HI Col^{Col} *X* WT Col (1)	×	✓
152			
153	HI Zu/Col^{Zu} *X* WT Zu (3)	✓	✓
154	HI Col^{Col} *X* WT Col (1)	✓	✓
155	HI Col^{Col} *X* WT Zu (1)	×	✓
156	HI Col^{Col} *X* WT Col (1)	✓	✓
157	HI Col^{Col} *X* WT Zu (2)	×	✓
158	HI Col^{Col} *X* WT Col (2)	×	✓
159	HI Col^{Col} *X* WT Col (2)	×	✓

Supplementary Table 6. Summary of number of F1 genotypes following HI *X* WT crosses. Three biological replicates of HI mutants were used as female parents in each cross, indicated by subscript number.

Cross ♀ ♂	1x	2x or aneuploid	Selfed
HI Col^{Col} *X* WT Col $_1$	5	2	0
$_2$	4	1	2
$_3$	5	0	0
HI Col^{Col} *X* WT Zu $_1$	4	1	0
$_2$	5	4	0
$_3$	5	4	2
HI Zu/Col^{Zu} *X* WT Zu $_1$	1	22	0
$_2$	2	19	1
$_3$	1	21	3
HI Zu/Col^{Zu} *X* WT Col $_1$	0	6	1
$_2$	1	4	1
$_3$	5	10	1

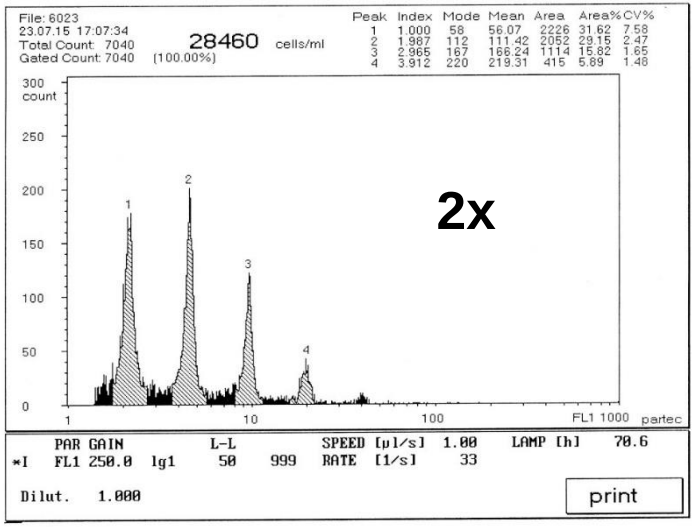
Appendix 3

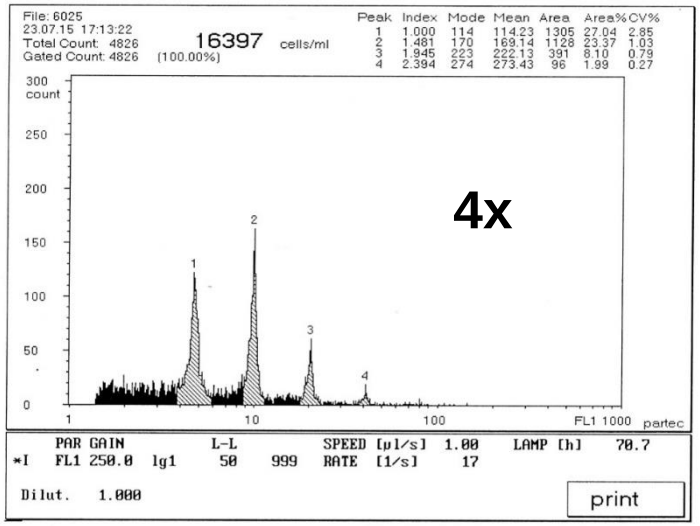
Chapter 5. Supplementary Results.

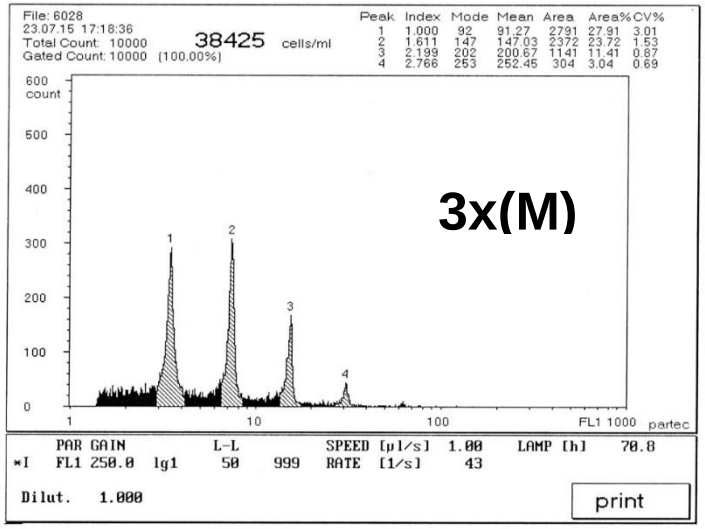
Scanned images of flow cytometry data.

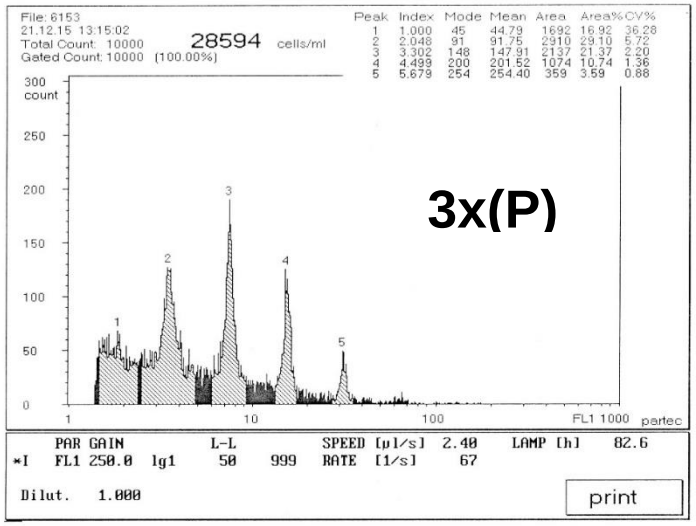
Ploidy level of diploid, tetraploid and reciprocal triploid plants of each accession confirmed through flow cytometry.

Bur-0

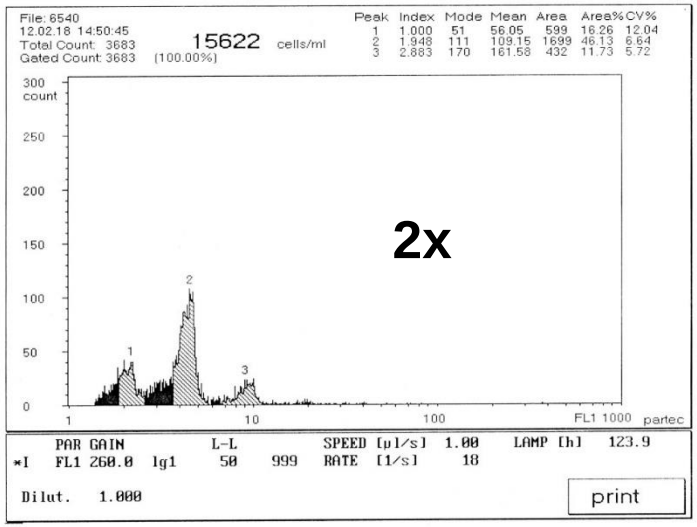

File: 6023
23.07.15 17:07:34
Total Count: 7040
Gated Count: 7040
28460 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 58 56.07 2226 31.62 7.58
2 1.987 112 111.42 2052 29.15 2.47
3 2.965 167 166.24 1114 15.82 1.65
4 3.912 220 219.31 415 5.89 1.48
2x
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 70.6
*I FL1 250.0 lg1 50 999 RATE [1/s] 33
Dilut. 1.000
print

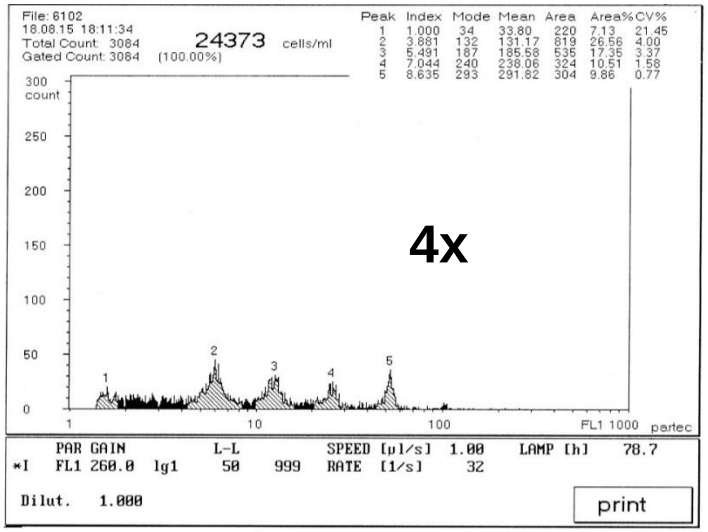

File: 6025
23.07.15 17:13:22
Total Count: 4826
Gated Count: 4826
16397 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 114 114.23 1305 27.04 2.85
2 1.481 170 169.14 1128 23.37 1.03
3 1.945 223 222.13 391 8.10 0.79
4 2.394 274 273.43 96 1.99 0.27
4x
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 70.7
*I FL1 250.0 lg1 50 999 RATE [1/s] 17
Dilut. 1.000
print

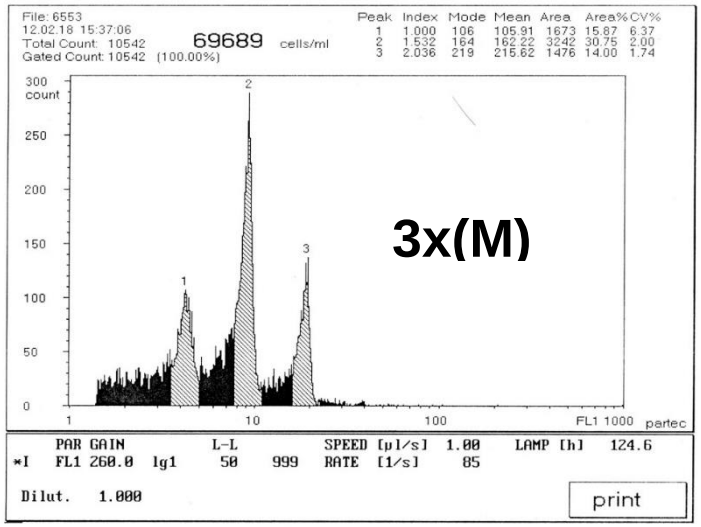

File: 6028
23.07.15 17:18:36
Total Count: 10000
Gated Count: 10000
38425 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 92 91.27 2791 27.91 3.01
2 1.611 147 147.03 2372 23.72 1.53
3 2.199 202 200.67 1141 11.41 0.87
4 2.766 253 252.45 304 3.04 0.69
3x(M)
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 70.8
*I FL1 250.0 lg1 50 999 RATE [1/s] 43
Dilut. 1.000
print

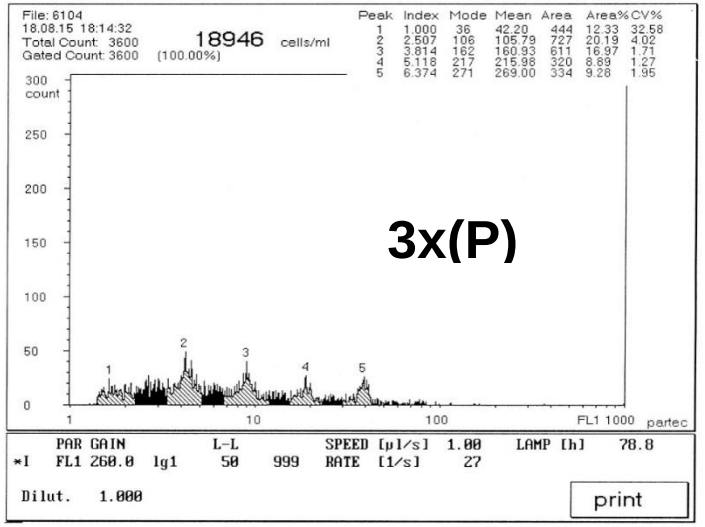

File: 6153
21.12.15 13:15:02
Total Count: 10000
Gated Count: 10000
28594 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 45 44.79 1692 16.92 36.28
2 2.048 91 91.75 2910 29.10 5.72
3 3.302 148 147.91 2137 21.37 2.20
4 4.499 200 201.52 1074 10.74 1.36
5 5.679 254 254.40 359 3.59 0.88
3x(P)
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 2.40 LAMP [h] 82.6
*I FL1 250.0 lg1 50 999 RATE [1/s] 67
Dilut. 1.000
print

C24

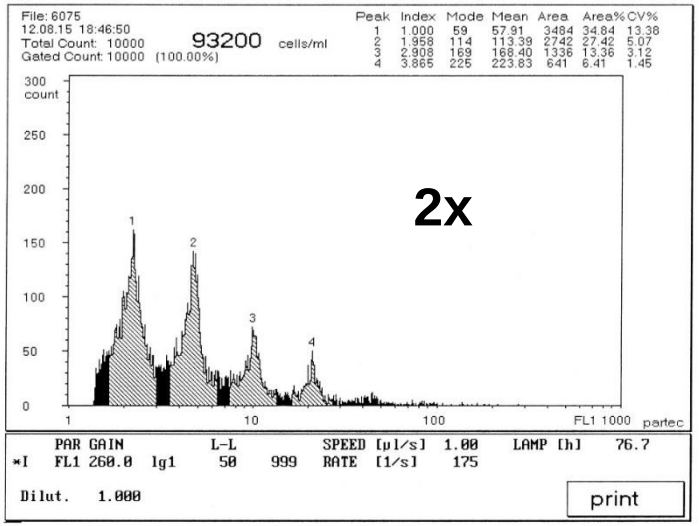

File: 6540
12.02.18 14:50:45
Total Count: 3683
Gated Count: 3683
15622 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 51 56.05 599 16.26 12.04
2 1.948 111 109.15 1699 46.13 6.64
3 2.883 170 161.58 432 11.73 5.72
2x
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 123.9
*I FL1 260.0 lg1 50 999 RATE [1/s] 18
Dilut. 1.000
print

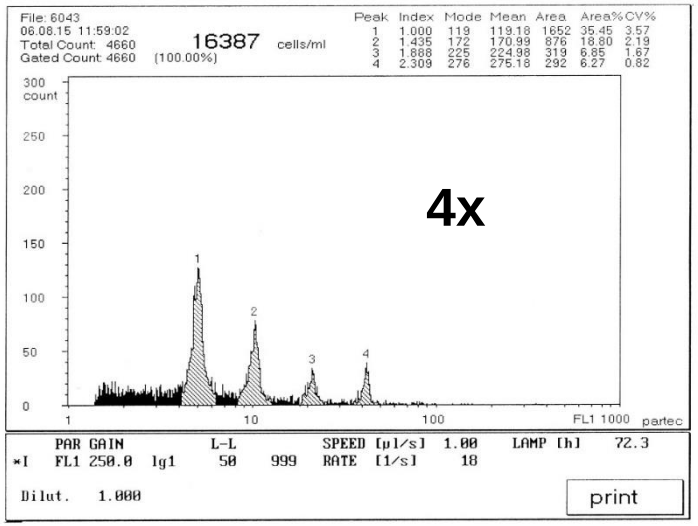

File: 6102
18.08.15 18:11:34
Total Count: 3084
Gated Count: 3084
24373 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 34 33.80 220 7.13 21.45
2 3.881 132 131.17 819 26.56 4.00
3 5.491 187 185.58 535 17.35 3.37
4 7.044 240 238.06 324 10.51 1.58
5 8.635 293 291.82 304 9.86 0.77
4x
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 78.7
*I FL1 260.0 lg1 50 999 RATE [1/s] 32
Dilut. 1.000
print

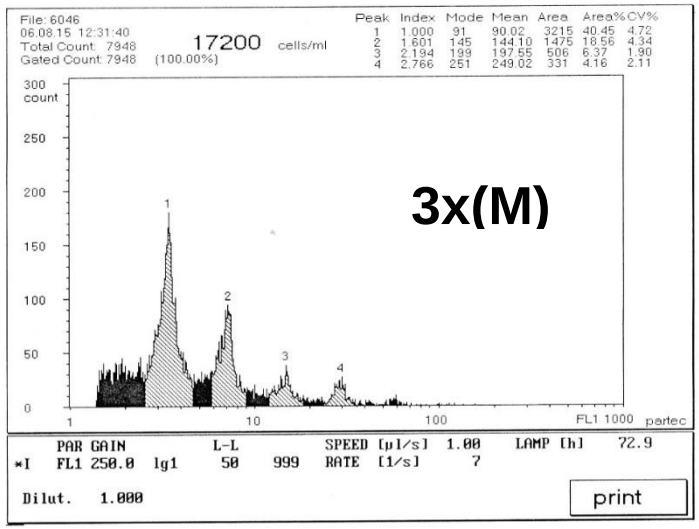

File: 6553
12.02.18 15:37:06
Total Count: 10542
Gated Count: 10542
69689 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 106 105.91 1673 15.87 6.37
2 1.532 164 162.22 3242 30.75 2.00
3 2.036 219 215.62 1476 14.00 1.74
3x(M)
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 124.6
*I FL1 260.0 lg1 50 999 RATE [1/s] 85
Dilut. 1.000
print

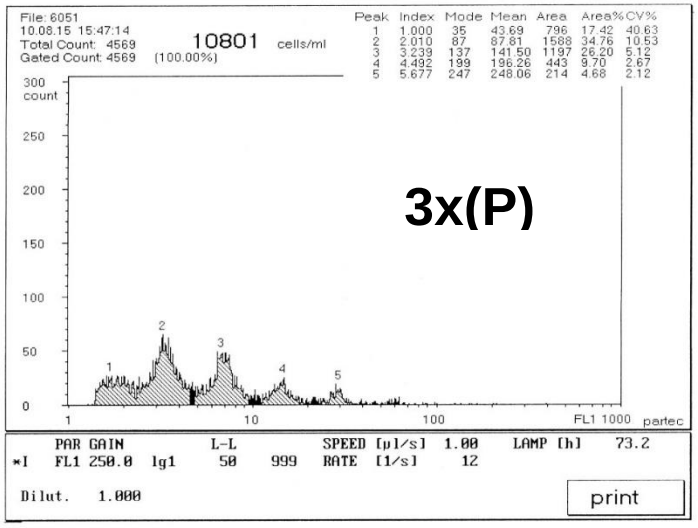

File: 6104
18.08.15 18:14:32
Total Count: 3600
Gated Count: 3600
18946 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 36 42.20 444 12.33 32.58
2 2.507 106 105.79 727 20.19 4.02
3 3.814 162 160.93 611 16.97 1.71
4 5.118 217 215.98 320 8.89 1.27
5 6.374 271 269.00 334 9.28 1.95
3x(P)
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 78.8
*I FL1 260.0 lg1 50 999 RATE [1/s] 27
Dilut. 1.000
print

Col-0

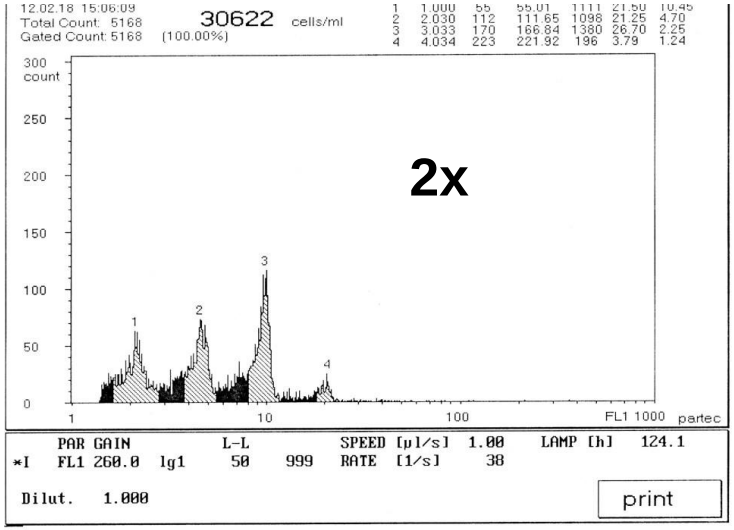

File: 6075
12.08.15 18:46:50
Total Count: 10000
Gated Count: 10000
93200 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 59 57.91 3484 34.84 13.38
2 1.958 114 113.39 2742 27.42 5.07
3 2.908 169 168.40 1336 13.36 3.12
4 3.865 225 223.83 641 6.41 1.45
2x
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 76.7
*I FL1 260.0 lg1 50 999 RATE [1/s] 175
Dilut. 1.000
print

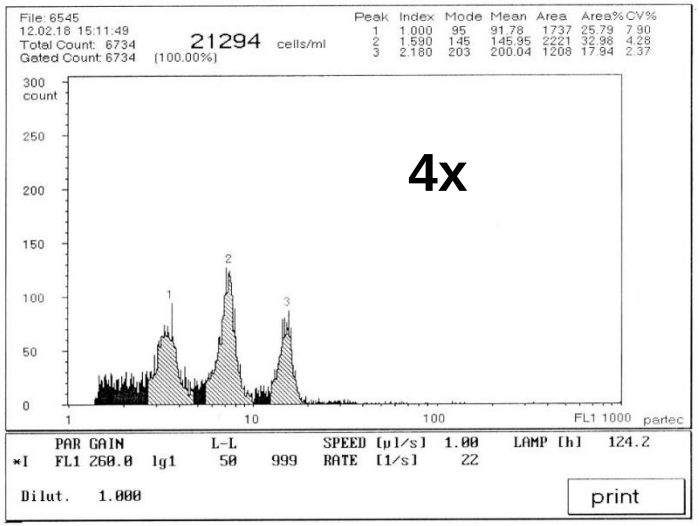

File: 6043
06.08.15 11:59:02
Total Count: 4660
Gated Count: 4660
16387 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 119 119.18 1652 35.45 3.57
2 1.435 172 170.99 876 18.80 2.19
3 1.888 225 224.98 319 6.85 1.67
4 2.309 276 275.18 292 6.27 0.82
4x
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 72.3
*I FL1 250.0 lg1 50 999 RATE [1/s] 18
Dilut. 1.000
print

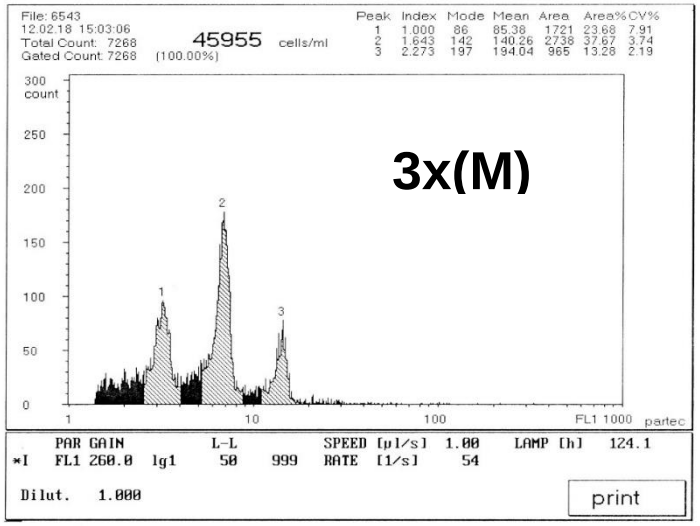

File: 6046
06.08.15 12:31:40
Total Count: 7948
Gated Count: 7948
17200 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 91 90.02 3215 40.45 4.72
2 1.601 145 144.10 1475 18.56 4.34
3 2.194 199 197.55 506 6.37 1.90
4 2.766 251 249.02 331 4.16 2.11
3x(M)
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 72.9
*I FL1 250.0 lg1 50 999 RATE [1/s] 7
Dilut. 1.000
print

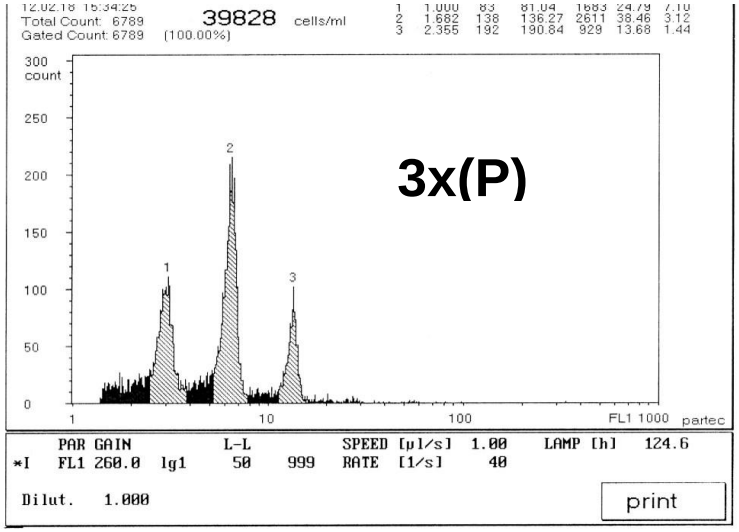

File: 6051
10.08.15 15:47:14
Total Count: 4569
Gated Count: 4569
10801 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 35 43.69 796 17.42 40.63
2 2.010 87 87.81 1588 34.76 10.53
3 3.239 137 141.50 1197 26.20 5.12
4 4.492 199 196.26 443 9.70 2.67
5 5.677 247 248.06 214 4.68 2.12
3x(P)
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 73.2
*I FL1 250.0 lg1 50 999 RATE [1/s] 12
Dilut. 1.000
print

Cvi

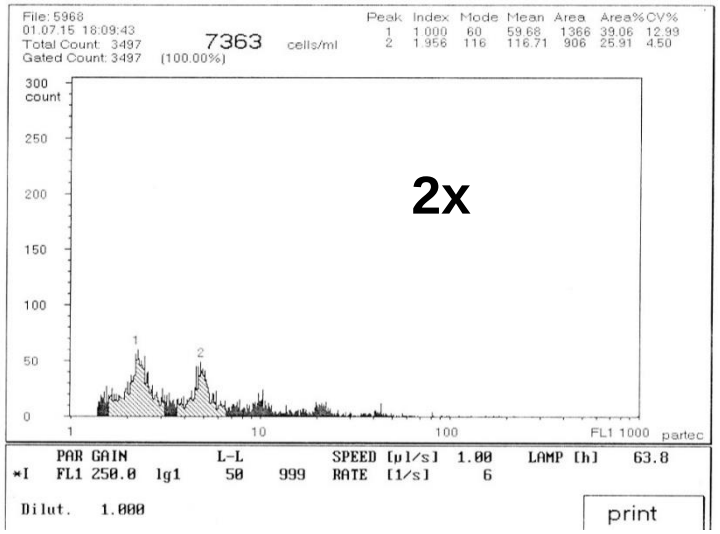

12.02.18 15:06:09
Total Count: 5168
Gated Count: 5168
30622 cells/ml
(100.00%)
1 1.000 55 55.01 1111 21.50 10.45
2 2.030 112 111.65 1098 21.25 4.70
3 3.033 170 166.84 1380 26.70 2.25
4 4.034 223 221.92 196 3.79 1.24
2x
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [pl/s] 1.00 LAMP [h] 124.1
*I FL1 260.0 lg1 50 999 RATE [1/s] 38
Dilut. 1.000
print

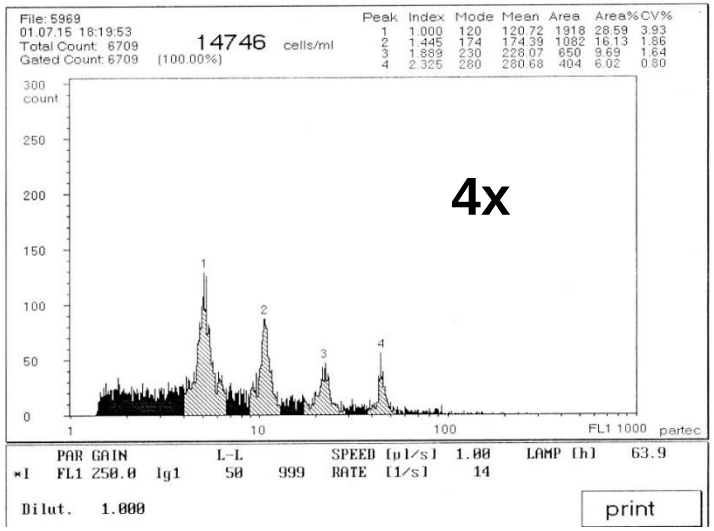

File: 6545
12.02.18 15:11:49
Total Count: 6734
Gated Count: 6734
21294 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 95 91.78 1737 25.79 7.90
2 1.590 145 145.95 2221 32.98 4.28
3 2.180 203 200.04 1208 17.94 2.37
4x
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [pl/s] 1.00 LAMP [h] 124.2
*I FL1 260.0 lg1 50 999 RATE [1/s] 22
Dilut. 1.000
print

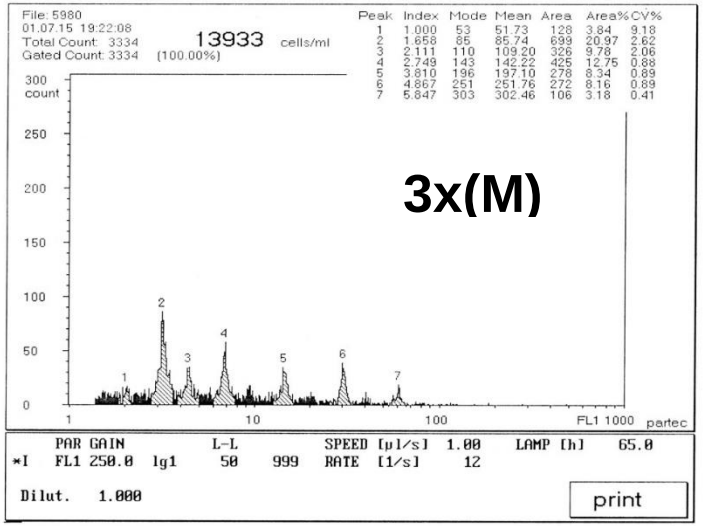

File: 6543
12.02.18 15:03:06
Total Count: 7268
Gated Count: 7268
45955 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 86 85.38 1721 23.68 7.91
2 1.643 142 140.26 2738 37.67 3.74
3 2.273 197 194.04 965 13.28 2.19
3x(M)
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [pl/s] 1.00 LAMP [h] 124.1
*I FL1 260.0 lg1 50 999 RATE [1/s] 54
Dilut. 1.000
print

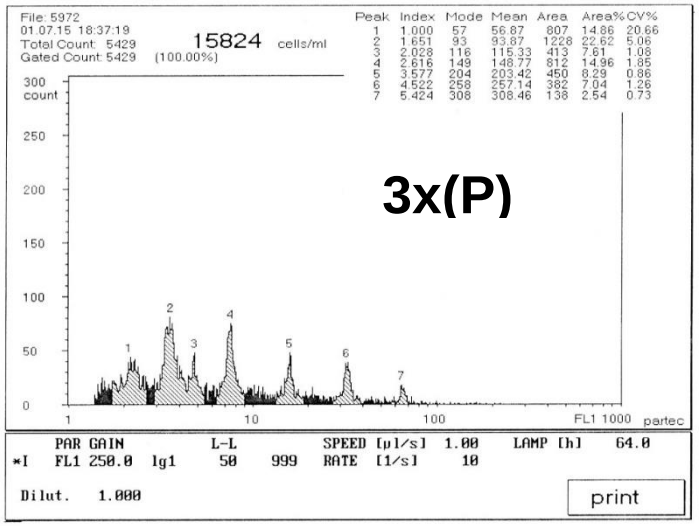

Total Count: 6789
Gated Count: 6789
39828 cells/ml
(100.00%)
1 1.000 83 81.04 1883 24.79 7.10
2 1.682 138 136.27 2611 38.46 3.12
3 2.355 192 190.84 929 13.68 1.44
3x(P)
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [pl/s] 1.00 LAMP [h] 124.6
*I FL1 260.0 lg1 50 999 RATE [1/s] 40
Dilut. 1.000
print

L*er*-0

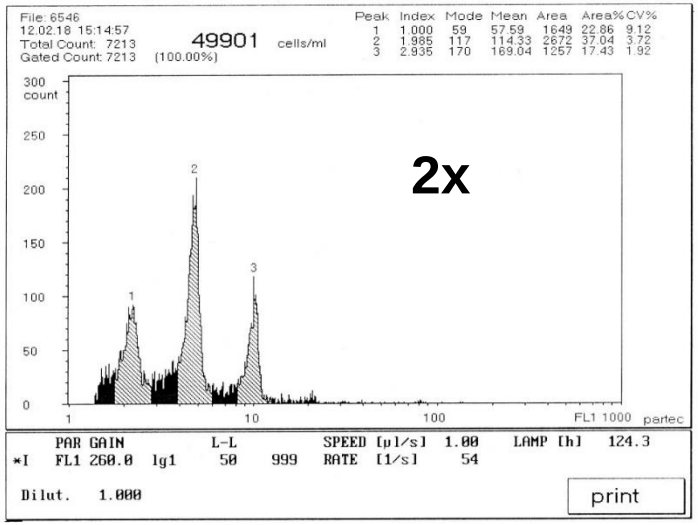

File: 5968
01.07.15 18:09:43
Total Count: 3497
Gated Count: 3497
7363 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 60 59.68 1366 39.06 12.99
2 1.956 116 116.71 906 25.91 4.50
2x
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 63.8
*I FL1 250.0 lg1 50 999 RATE [1/s] 6
Dilut. 1.000
print

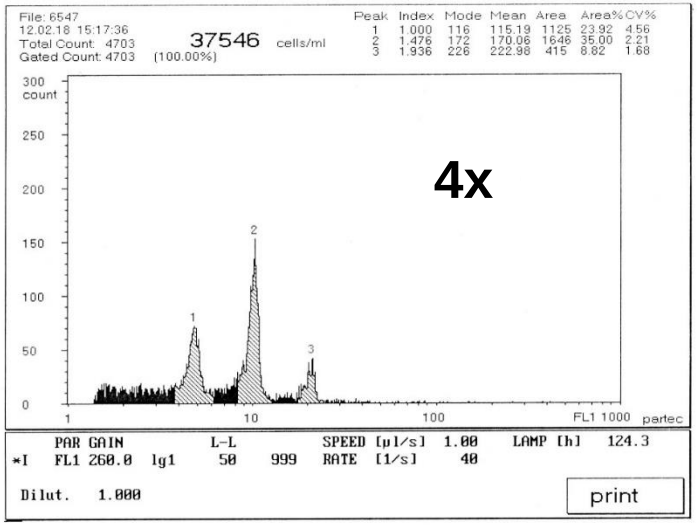

File: 5969
01.07.15 18:19:53
Total Count: 6709
Gated Count: 6709
14746 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 120 120.72 1918 28.59 3.93
2 1.445 174 174.39 1082 16.13 1.86
3 1.889 230 228.07 650 9.69 1.64
4 2.325 280 280.68 404 6.02 0.80
4x
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 63.9
*I FL1 250.0 lg1 50 999 RATE [1/s] 14
Dilut. 1.000
print

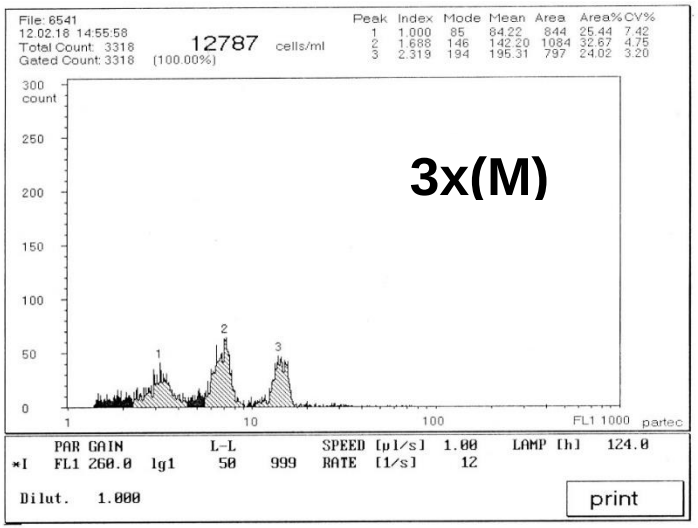

File: 5980
01.07.15 19:22:08
Total Count: 3334
Gated Count: 3334
13933 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 53 51.73 128 3.84 9.18
2 1.658 85 85.74 699 20.97 2.62
3 2.111 110 109.20 326 9.78 2.06
4 2.749 143 142.22 425 12.75 0.88
5 3.810 196 197.10 278 8.34 0.89
6 4.867 251 251.76 272 8.16 0.89
7 5.847 303 302.46 106 3.18 0.41
3x(M)
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 65.0
*I FL1 250.0 lg1 50 999 RATE [1/s] 12
Dilut. 1.000
print

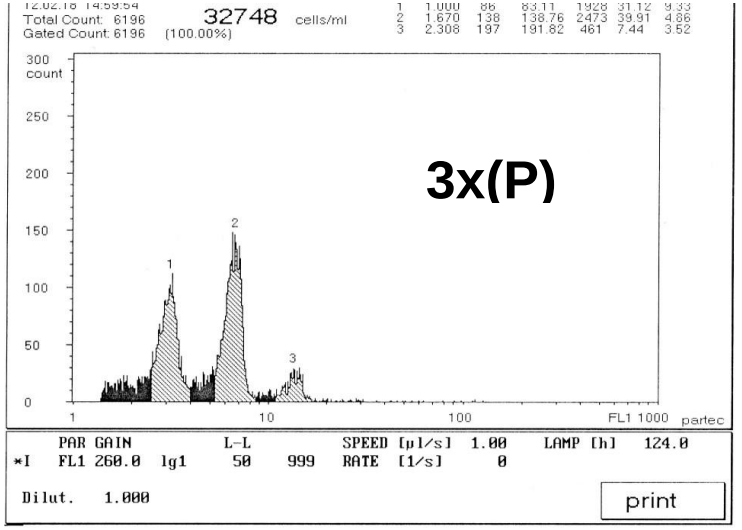

File: 5972
01.07.15 18:37:19
Total Count: 5429
Gated Count: 5429
15824 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 57 56.87 807 14.86 20.66
2 1.651 93 93.87 1228 22.62 5.06
3 2.028 116 115.33 413 7.61 1.08
4 2.616 149 148.77 812 14.96 1.85
5 3.577 204 203.42 450 8.29 0.86
6 4.522 258 257.14 382 7.04 1.26
7 5.424 308 308.46 138 2.54 0.73
3x(P)
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 64.0
*I FL1 250.0 lg1 50 999 RATE [1/s] 10
Dilut. 1.000
print

Sorbo

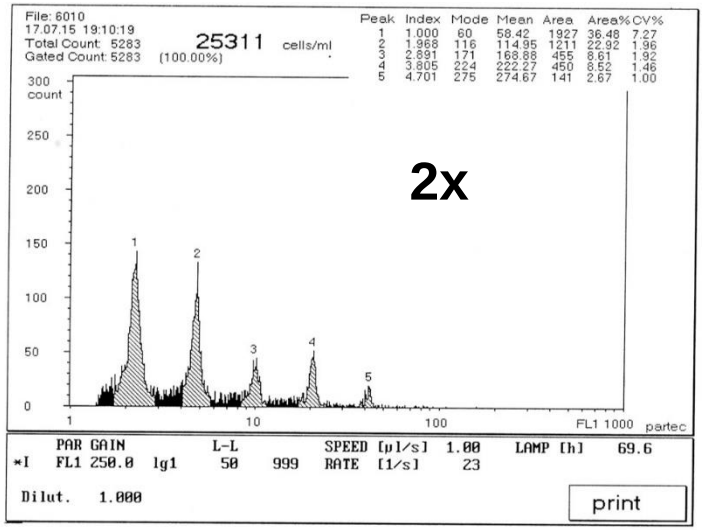

File: 6546
12.02.18 15:14:57
Total Count: 7213
Gated Count: 7213
49901 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 59 57.59 1649 22.86 9.12
2 1.985 117 114.33 2672 37.04 3.72
3 2.935 170 169.04 1257 17.43 1.92
2x
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 124.3
*I FL1 260.0 lg1 50 999 RATE [1/s] 54
Dilut. 1.000
print

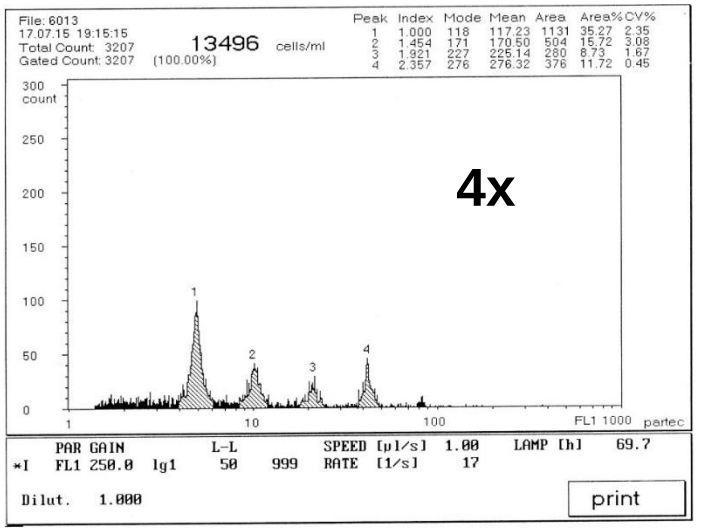

File: 6547
12.02.18 15:17:36
Total Count: 4703
Gated Count: 4703
37546 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 116 115.19 1125 23.92 4.56
2 1.476 172 170.06 1646 35.00 2.21
3 1.936 226 222.98 415 8.82 1.68
4x
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 124.3
*I FL1 260.0 lg1 50 999 RATE [1/s] 40
Dilut. 1.000
print

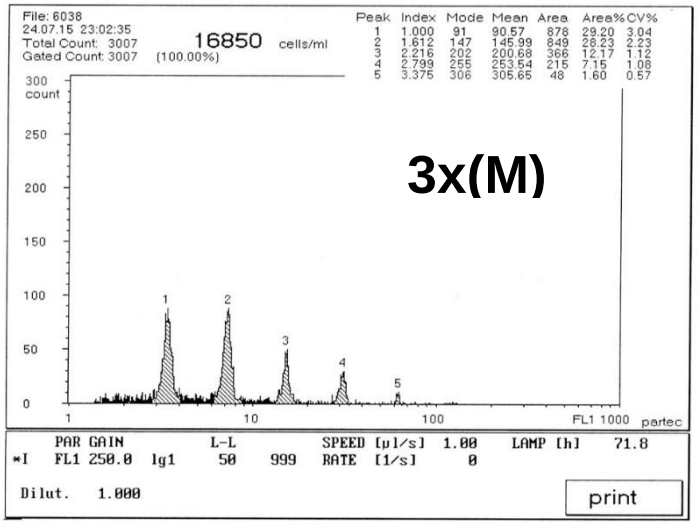

File: 6541
12.02.18 14:55:58
Total Count: 3318
Gated Count: 3318
12787 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 85 84.22 844 25.44 7.42
2 1.688 146 142.20 1084 32.67 4.75
3 2.319 194 195.31 797 24.02 3.20
3x(M)
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 124.0
*I FL1 260.0 lg1 50 999 RATE [1/s] 12
Dilut. 1.000
print

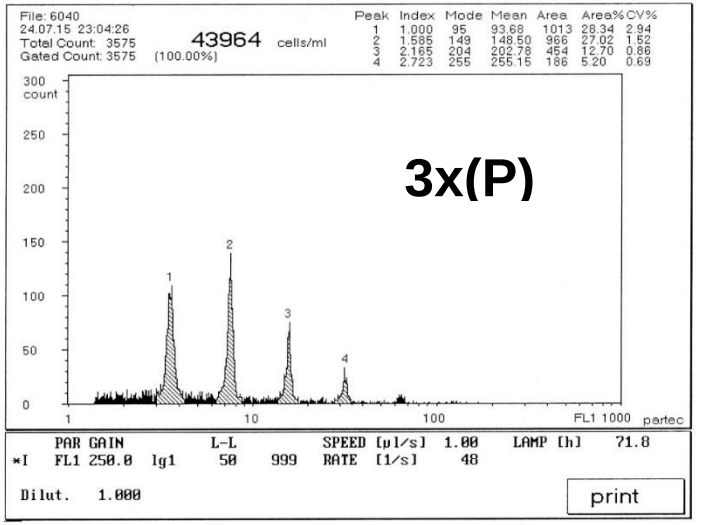

Total Count: 6196
Gated Count: 6196
32748 cells/ml
(100.00%)
1 1.000 86 83.11 1928 31.12 9.33
2 1.670 138 138.76 2473 39.91 4.86
3 2.308 197 191.82 461 7.44 3.52
3x(P)
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 124.0
*I FL1 260.0 lg1 50 999 RATE [1/s] 0
Dilut. 1.000
print

T910

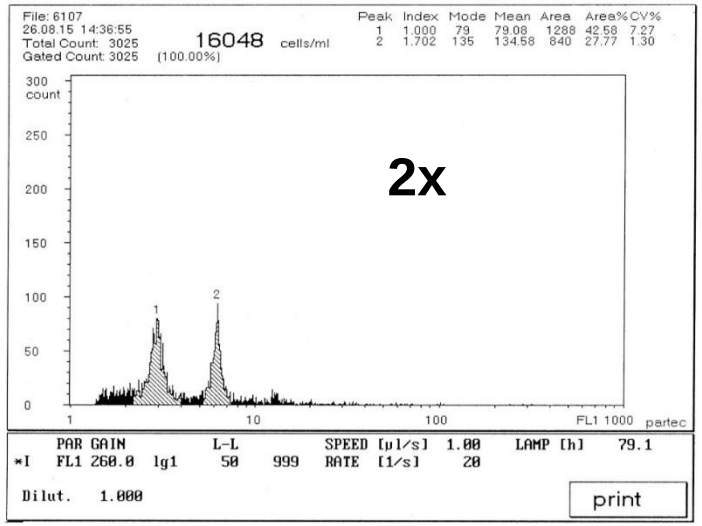

File: 6010
17.07.15 19:10:19
Total Count: 5283
Gated Count: 5283
25311 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 60 58.42 1927 36.48 7.27
2 1.968 116 114.95 1211 22.92 1.96
3 2.891 171 168.88 455 8.61 1.92
4 3.805 224 222.27 450 8.52 1.46
5 4.701 275 274.67 141 2.67 1.00
2x
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 69.6
*I FL1 250.0 lg1 50 999 RATE [1/s] 23
Dilut. 1.000
print

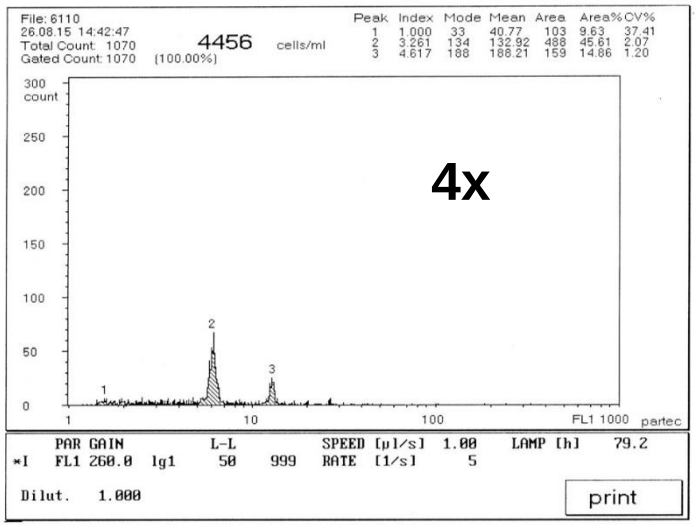

File: 6013
17.07.15 19:15:15
Total Count: 3207
Gated Count: 3207
13496 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 118 117.23 1131 35.27 2.35
2 1.454 171 170.50 504 15.72 3.08
3 1.921 227 225.14 280 8.73 1.67
4 2.357 276 276.32 376 11.72 0.45
4x
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 69.7
*I FL1 250.0 lg1 50 999 RATE [1/s] 17
Dilut. 1.000
print

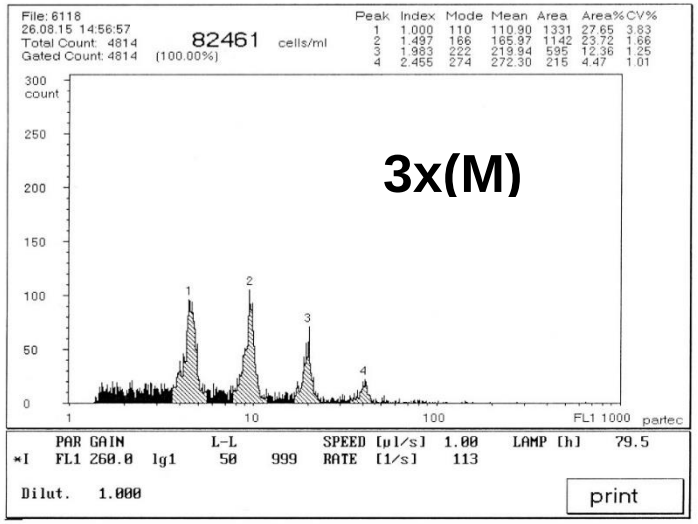

File: 6038
24.07.15 23:02:35
Total Count: 3007
Gated Count: 3007
16850 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 91 90.57 878 29.20 3.04
2 1.612 147 145.99 849 28.23 2.23
3 2.216 202 200.68 366 12.17 1.12
4 2.799 255 253.54 215 7.15 1.08
5 3.375 306 305.65 48 1.60 0.57
3x(M)
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 71.8
*I FL1 250.0 lg1 50 999 RATE [1/s] 0
Dilut. 1.000
print

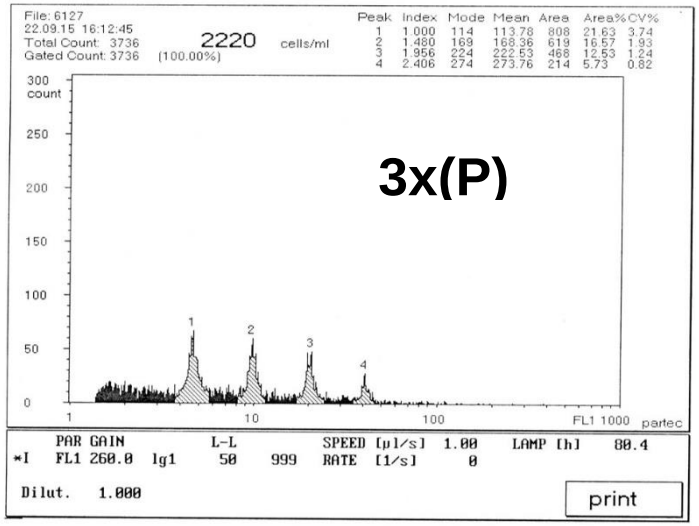

File: 6040
24.07.15 23:04:26
Total Count: 3575
Gated Count: 3575
43964 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 95 93.68 1013 28.34 2.94
2 1.585 149 148.50 966 27.02 1.52
3 2.165 204 202.78 454 12.70 0.86
4 2.723 255 255.15 186 5.20 0.69
3x(P)
300 count
250
200
150
100
50
0
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 71.8
*I FL1 250.0 lg1 50 999 RATE [1/s] 48
Dilut. 1.000
print

TAL07

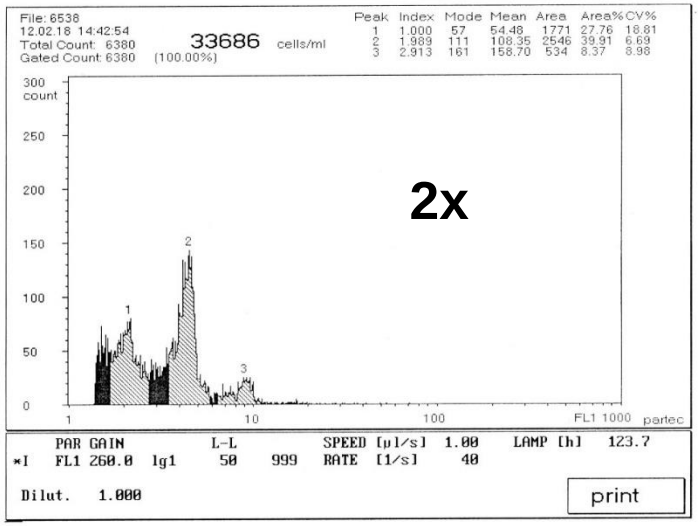

File: 6107
26.08.15 14:36:55
Total Count: 3025
Gated Count: 3025
16048 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 79 79.08 1288 42.58 7.27
2 1.702 135 134.58 840 27.77 1.30
2x
300 count
250
200
150
100
50
0
1
2
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 79.1
*I FL1 260.0 lg1 50 999 RATE [1/s] 20
Dilut. 1.000
print

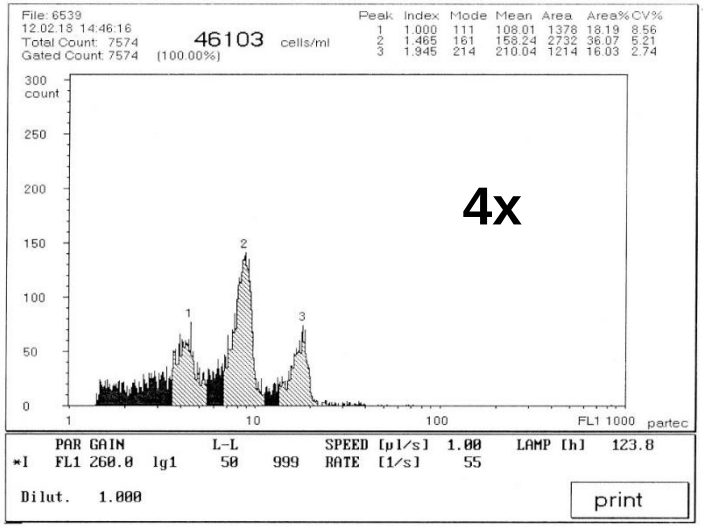

File: 6110
26.08.15 14:42:47
Total Count: 1070
Gated Count: 1070
4456 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 33 40.77 103 9.63 37.41
2 3.261 134 132.92 488 45.81 2.07
3 4.617 188 188.21 159 14.86 1.20
4x
300 count
250
200
150
100
50
0
1
2
3
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 79.2
*I FL1 260.0 lg1 50 999 RATE [1/s] 5
Dilut. 1.000
print

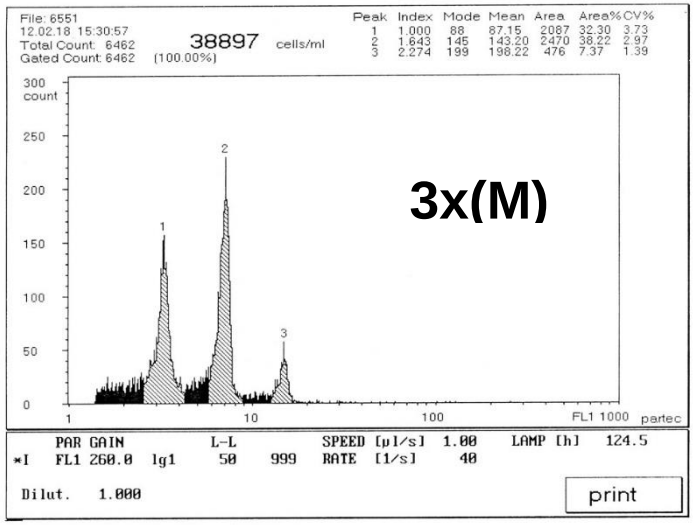

File: 6118
26.08.15 14:56:57
Total Count: 4814
Gated Count: 4814
82461 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 110 110.90 1331 27.65 3.83
2 1.497 166 165.97 1142 23.72 1.66
3 1.983 222 219.94 595 12.36 1.25
4 2.455 274 272.30 215 4.47 1.01
3x(M)
300 count
250
200
150
100
50
0
1
2
3
4
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 79.5
*I FL1 260.0 lg1 50 999 RATE [1/s] 113
Dilut. 1.000
print

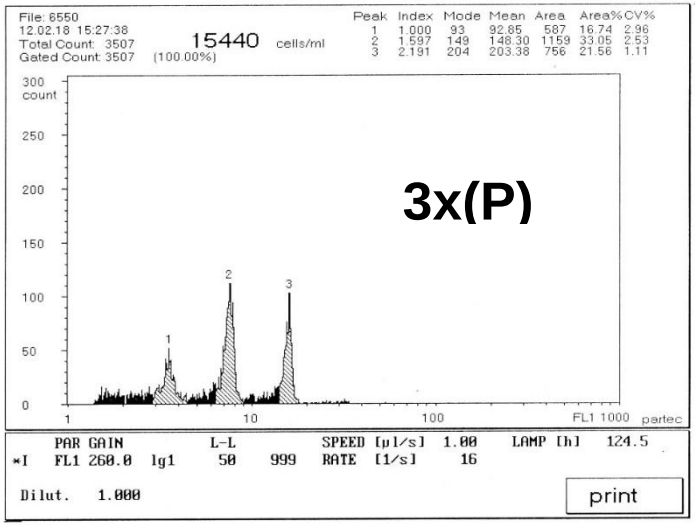

File: 6127
22.09.15 16:12:45
Total Count: 3736
Gated Count: 3736
2220 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 114 113.78 808 21.63 3.74
2 1.480 169 168.36 619 16.57 1.93
3 1.956 224 222.53 468 12.53 1.24
4 2.406 274 273.76 214 5.73 0.82
3x(P)
300 count
250
200
150
100
50
0
1
2
3
4
1
10
100
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 80.4
*I FL1 260.0 lg1 50 999 RATE [1/s] 0
Dilut. 1.000
print

Wilna

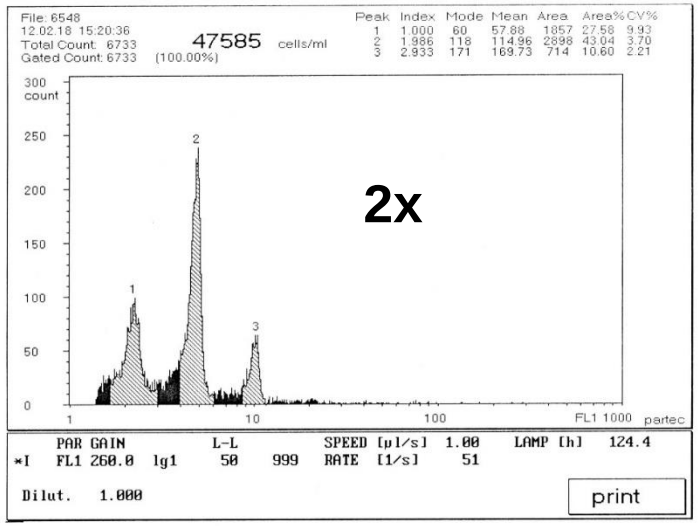

File: 6538
12.02.18 14:42:54
Total Count: 6380
Gated Count: 6380
33686 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 57 54.48 1771 27.76 18.81
2 1.989 111 108.35 2546 39.91 6.69
3 2.913 161 158.70 534 8.37 8.98
2x
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 123.7
*I FL1 260.0 lg1 50 999 RATE [1/s] 40
Dilut. 1.000
print

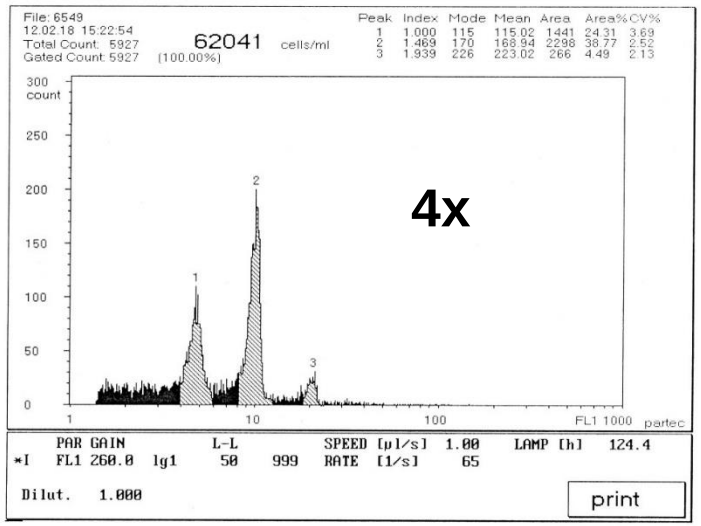

File: 6539
12.02.18 14:46:16
Total Count: 7574
Gated Count: 7574
46103 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 111 108.01 1378 18.19 8.56
2 1.465 161 158.24 2732 36.07 5.21
3 1.945 214 210.04 1214 16.03 2.74
4x
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 123.8
*I FL1 260.0 lg1 50 999 RATE [1/s] 55
Dilut. 1.000
print

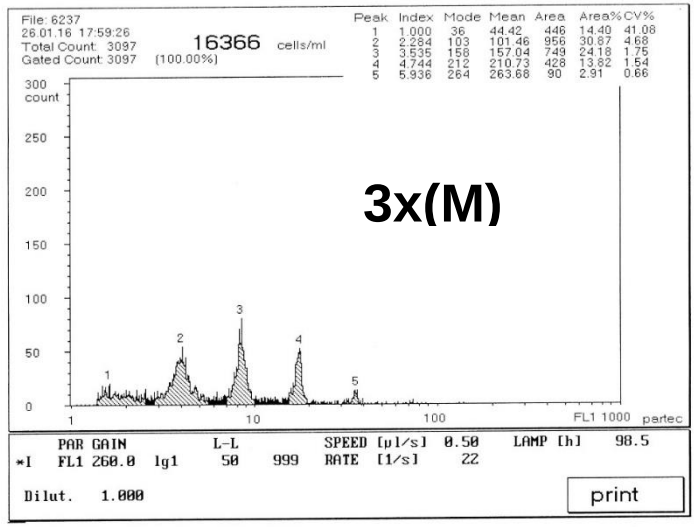

File: 6551
12.02.18 15:30:57
Total Count: 6462
Gated Count: 6462
38897 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 88 87.15 2087 32.30 3.73
2 1.643 145 143.20 2470 38.22 2.97
3 2.274 199 198.22 476 7.37 1.39
3x(M)
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 124.5
*I FL1 260.0 lg1 50 999 RATE [1/s] 40
Dilut. 1.000
print

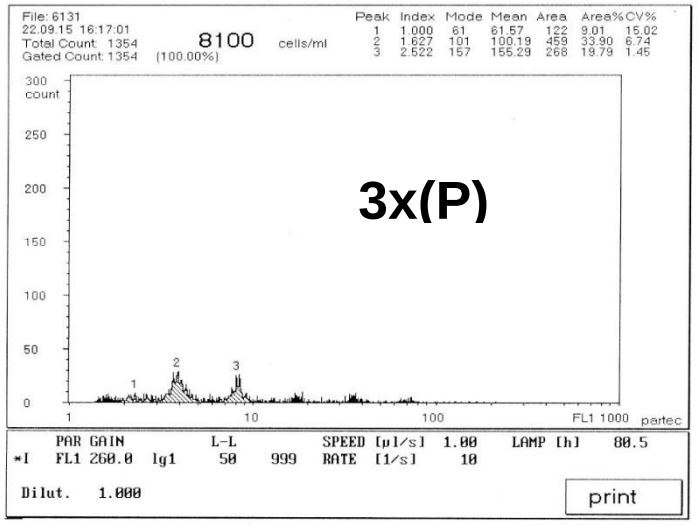

File: 6550
12.02.18 15:27:38
Total Count: 3507
Gated Count: 3507
15440 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 93 92.85 587 16.74 2.96
2 1.597 149 148.30 1159 33.05 2.53
3 2.191 204 203.38 756 21.56 1.11
3x(P)
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 124.5
*I FL1 260.0 lg1 50 999 RATE [1/s] 16
Dilut. 1.000
print

Zurich


File: 6548
12.02.18 15:20:36
Total Count: 6733
Gated Count: 6733
47585 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 60 57.88 1857 27.58 9.93
2 1.986 118 114.96 2898 43.04 3.70
3 2.933 171 169.73 714 10.60 2.21
2x
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 124.4
*I FL1 260.0 lg1 50 999 RATE [1/s] 51
Dilut. 1.000
print


File: 6549
12.02.18 15:22:54
Total Count: 5927
Gated Count: 5927
62041 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 115 115.02 1441 24.31 3.69
2 1.469 170 168.94 2298 38.77 2.52
3 1.939 226 223.02 266 4.49 2.13
4x
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 124.4
*I FL1 260.0 lg1 50 999 RATE [1/s] 65
Dilut. 1.000
print


File: 6237
26.01.16 17:59:26
Total Count: 3097
Gated Count: 3097
16366 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 36 44.42 446 14.40 41.08
2 2.284 103 101.46 956 30.87 4.68
3 3.535 158 157.04 749 24.18 1.75
4 4.744 212 210.73 428 13.82 1.54
5 5.936 264 263.68 90 2.91 0.66
3x(M)
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 0.50 LAMP [h] 98.5
*I FL1 260.0 lg1 50 999 RATE [1/s] 22
Dilut. 1.000
print

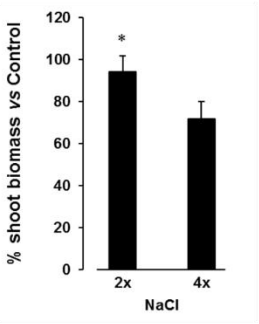

File: 6131
22.09.15 16:17:01
Total Count: 1354
Gated Count: 1354
8100 cells/ml
(100.00%)
Peak Index Mode Mean Area Area% CV%
1 1.000 61 61.57 122 9.01 15.02
2 1.627 101 100.19 459 33.90 6.74
3 2.522 157 155.29 268 19.79 1.45
3x(P)
count
FL1 1000
partec
PAR GAIN L-L SPEED [µl/s] 1.00 LAMP [h] 80.5
*I FL1 260.0 lg1 50 999 RATE [1/s] 10
Dilut. 1.000
print

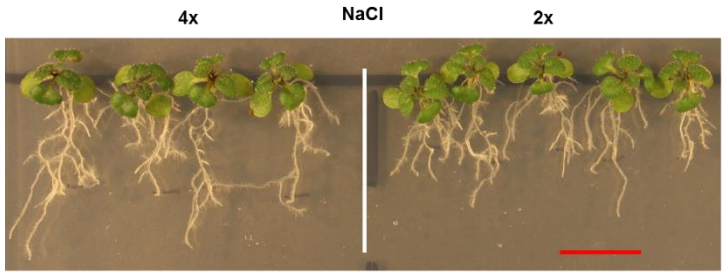
The genome dosage and parent-of-origin effect on salt stress tolerance of each accession.

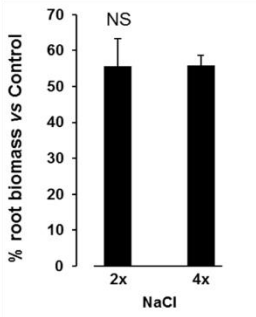
Salt stress tolerance of diploid, tetraploid and reciprocal triploids plants of each accession, including representative pictures. * statistically different ($P < 0.05$); NS not statistically different ($P > 0.05$) according to a two-tailed independent samples t-test.

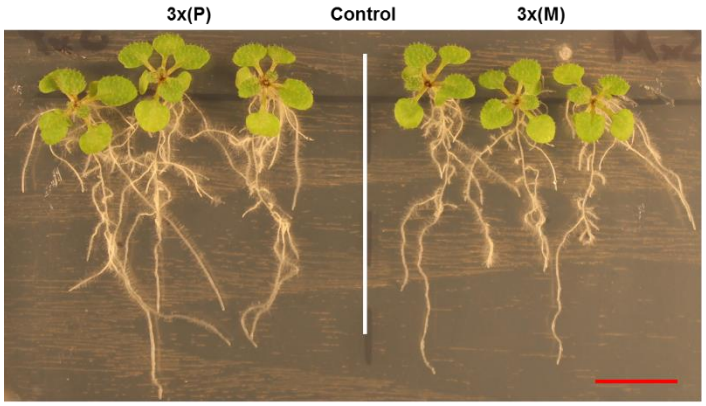
Bur-0

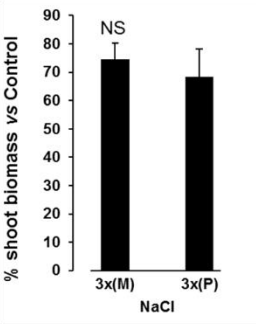

Control
4x
2x

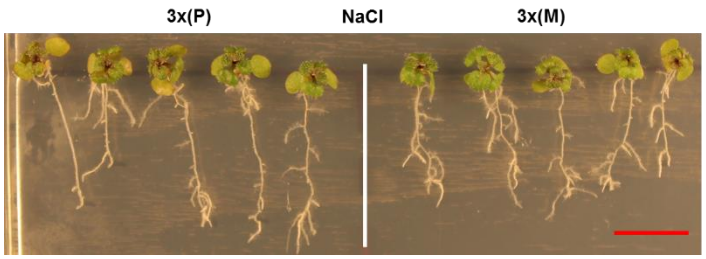

% shoot biomass vs Control
120
100
80
60
40
20
0
*
2x
4x
NaCl

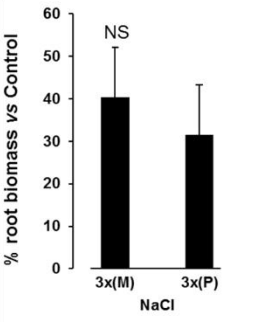

4x
NaCl
2x


% root biomass vs Control
70
60
50
40
30
20
10
0
NS
2x
4x
NaCl

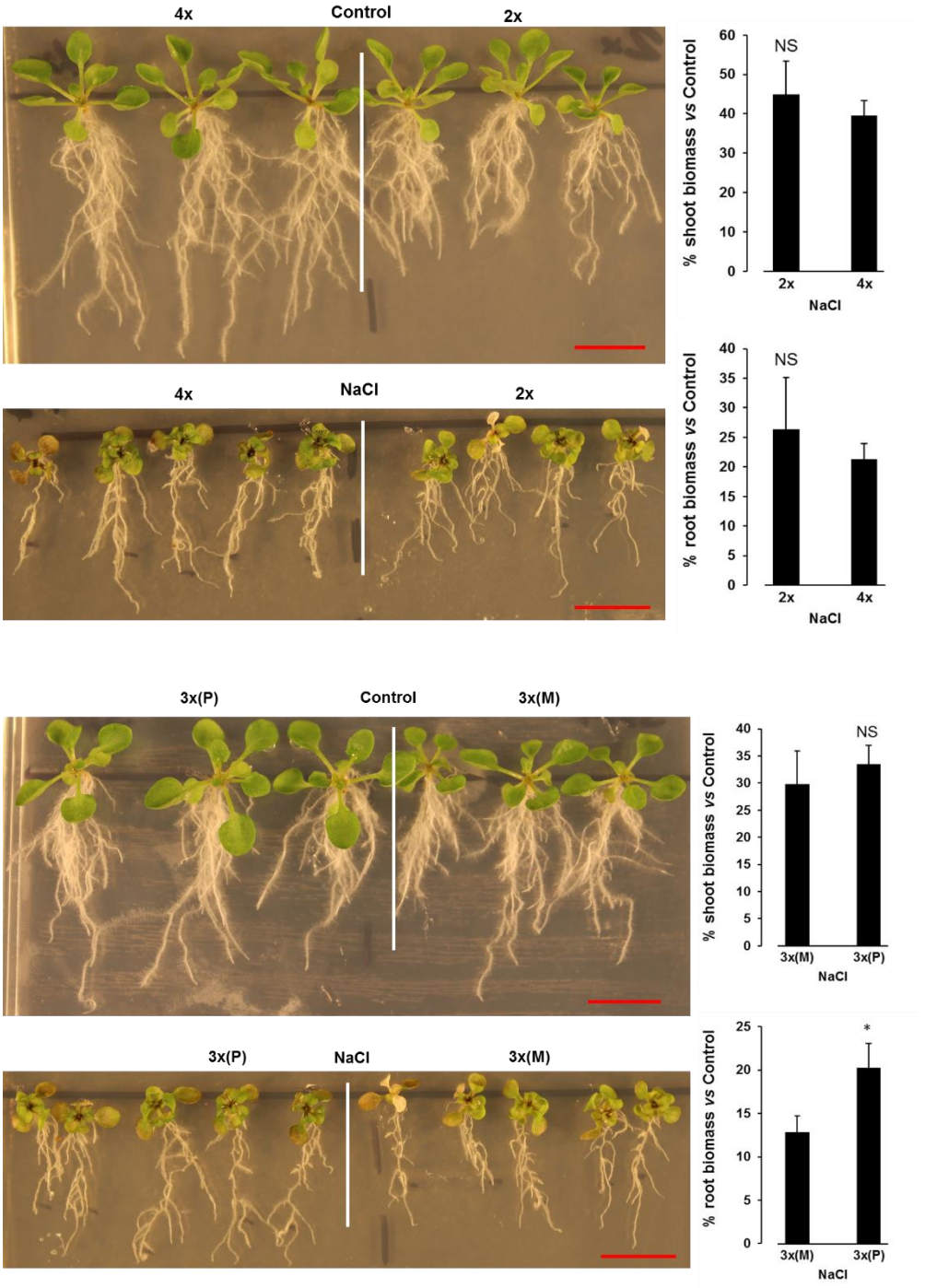

3x(P)
Control
3x(M)


% shoot biomass vs Control
90
80
70
60
50
40
30
20
10
0
NS
3x(M)
3x(P)
NaCl

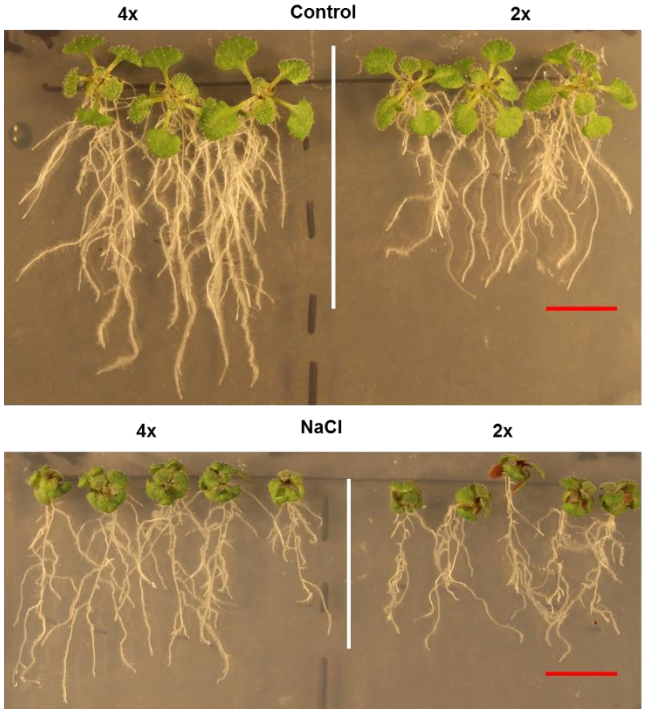

3x(P)
NaCl
3x(M)

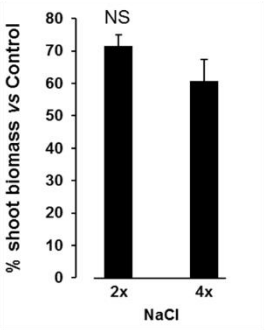

% root biomass vs Control
60
50
40
30
20
10
0
NS
3x(M)
3x(P)
NaCl

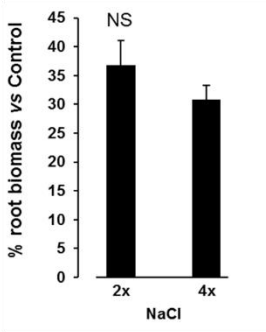
C24

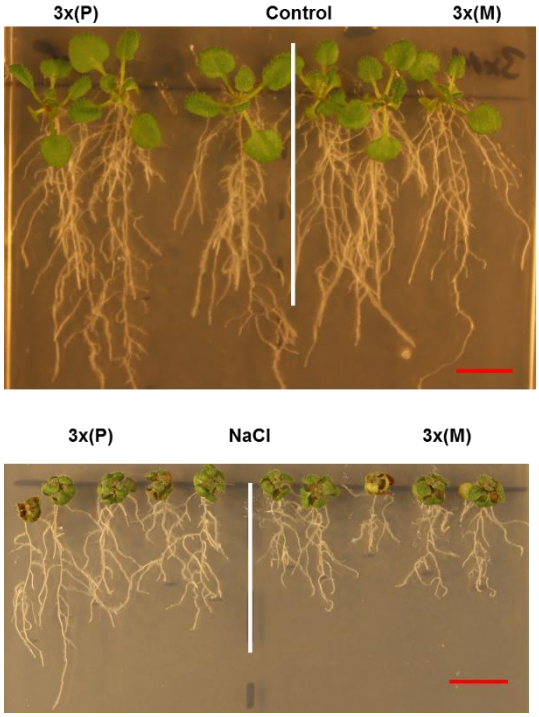

4x
Control
2x
% shoot biomass vs Control
NS
2x
4x
NaCl
4x
NaCl
2x
% root biomass vs Control
NS
2x
4x
NaCl
3x(P)
Control
3x(M)
% shoot biomass vs Control
NS
3x(M)
3x(P)
NaCl
3x(P)
NaCl
3x(M)
% root biomass vs Control
*
3x(M)
3x(P)
NaCl

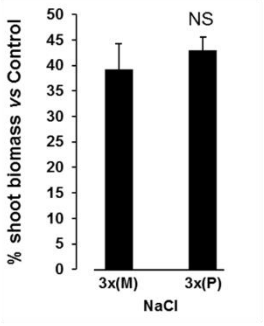
Col-0

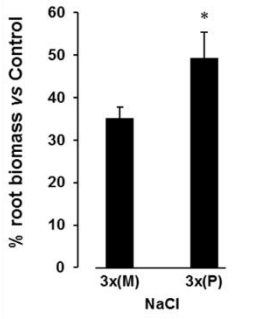

4x
Control
2x
4x
NaCl
2x


% shoot biomass vs Control
NS
80
70
60
50
40
30
20
10
0
2x
4x
NaCl

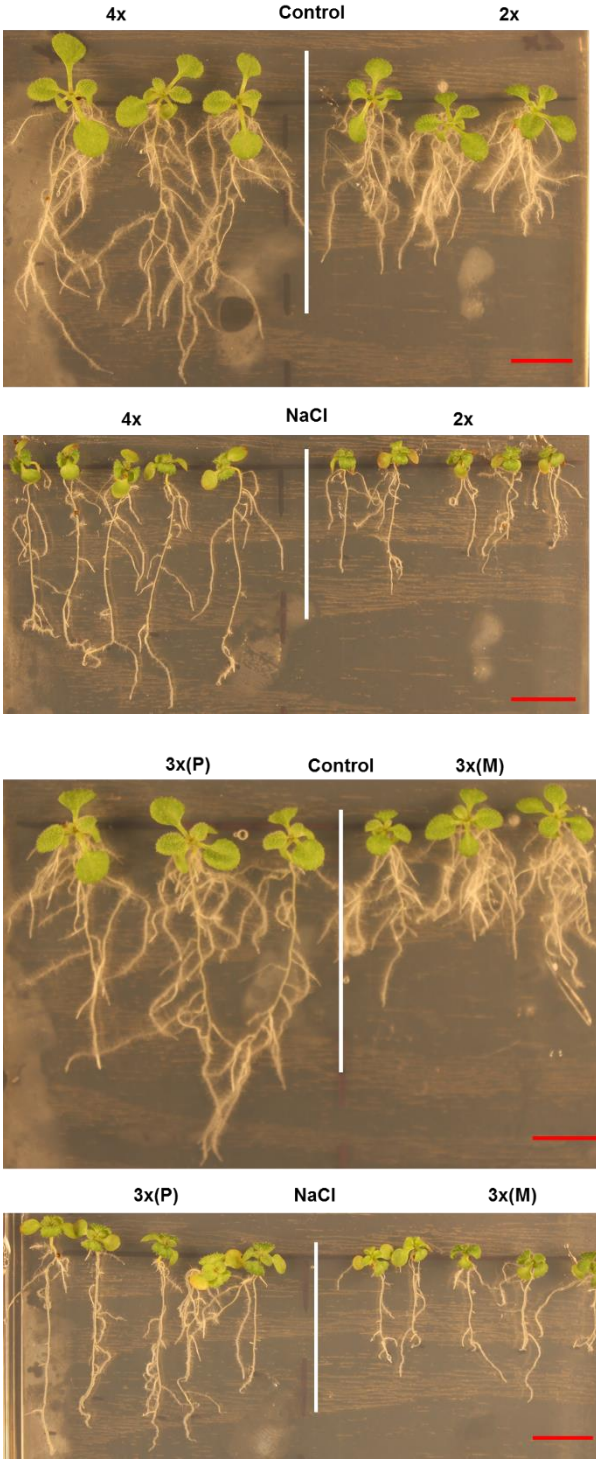

% root biomass vs Control
NS
45
40
35
30
25
20
15
10
5
0
2x
4x
NaCl

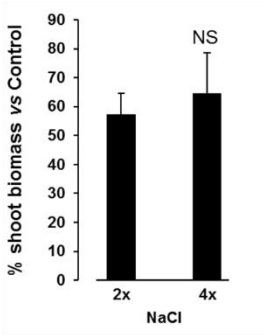

3x(P)
Control
3x(M)
3x(P)
NaCl
3x(M)

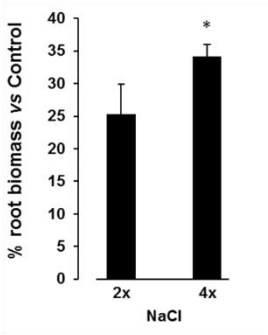

% shoot biomass vs Control
NS
50
45
40
35
30
25
20
15
10
5
0
3x(M)
3x(P)
NaCl

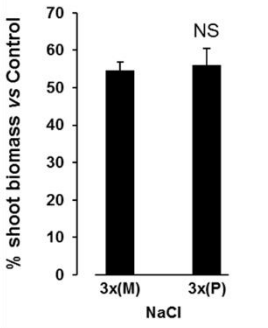

% root biomass vs Control
*
60
50
40
30
20
10
0
3x(M)
3x(P)
NaCl

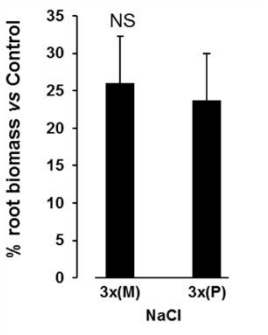
Cvi


4x
Control
2x
4x
NaCl
2x
3x(P)
Control
3x(M)
3x(P)
NaCl
3x(M)

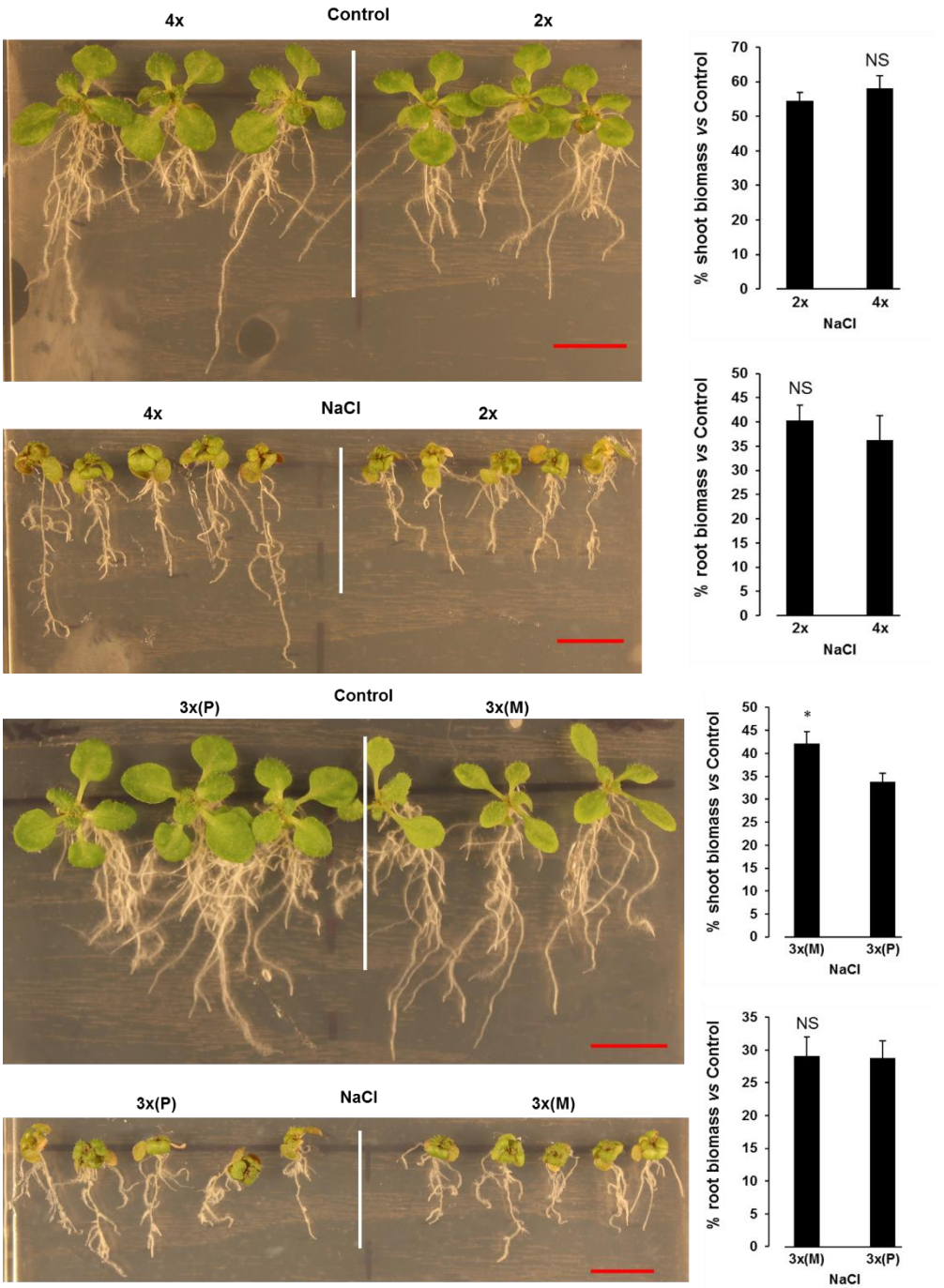

% shoot biomass vs Control
NS
2x
4x
NaCl


% root biomass vs Control
*
2x
4x
NaCl

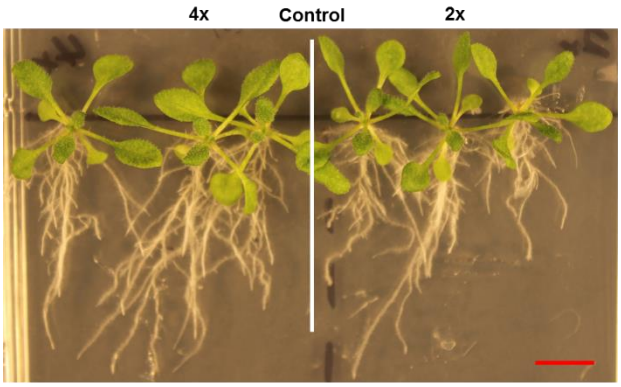

% shoot biomass vs Control
NS
3x(M)
3x(P)
NaCl

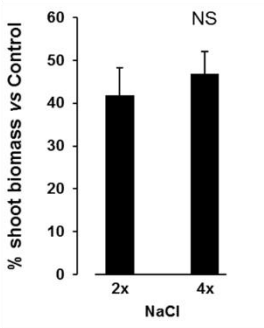

% root biomass vs Control
NS
3x(M)
3x(P)
NaCl

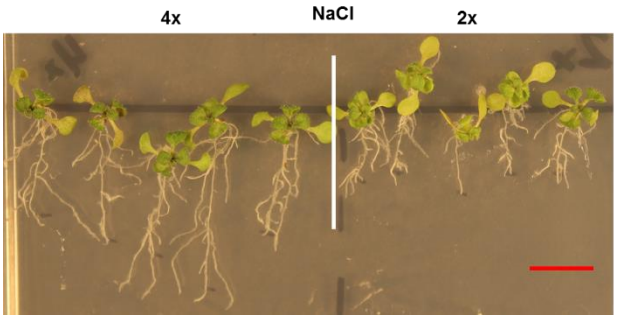
L*er*-0

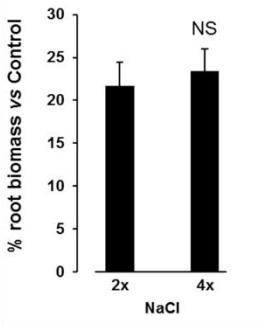

4x
Control
2x
% shoot biomass vs Control
NS
2x
4x
NaCl
4x
NaCl
2x
% root biomass vs Control
NS
2x
4x
NaCl
3x(P)
Control
3x(M)
% shoot biomass vs Control
*
3x(M)
3x(P)
NaCl
3x(P)
NaCl
3x(M)
% root biomass vs Control
NS
3x(M)
3x(P)
NaCl

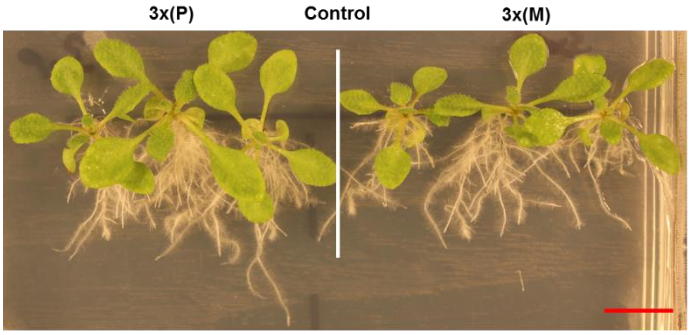
Sorbo

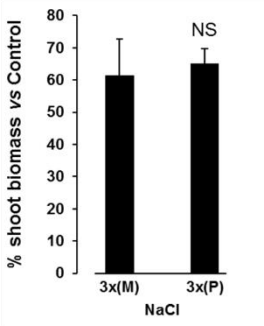

4x
Control
2x

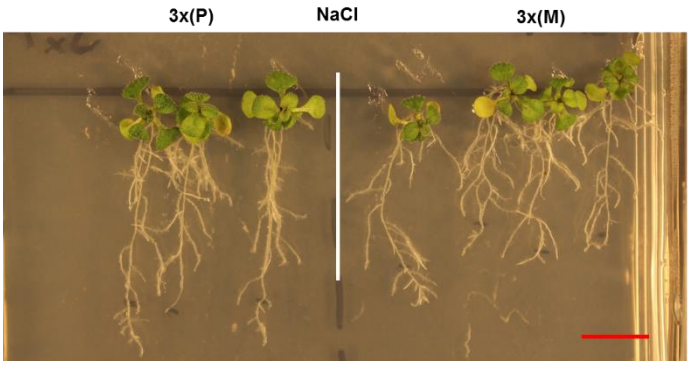

% shoot biomass vs Control
NS
60
50
40
30
20
10
0
2x
4x
NaCl

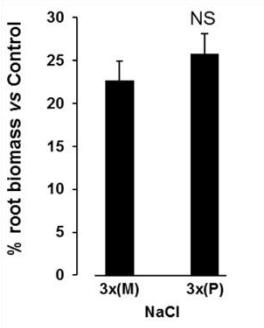

4x
NaCl
2x


% root biomass vs Control
NS
30
25
20
15
10
5
0
2x
4x
NaCl


3x(P)
Control
3x(M)

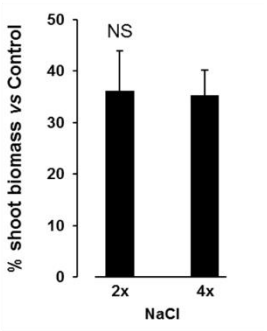

% shoot biomass vs Control
NS
80
70
60
50
40
30
20
10
0
3x(M)
3x(P)
NaCl

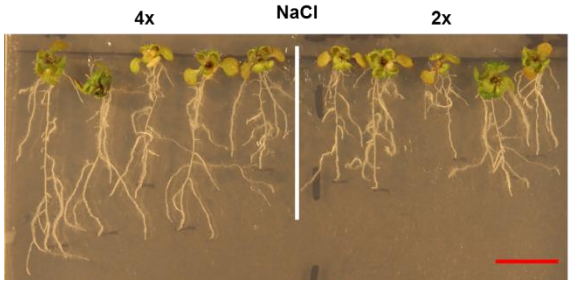

3x(P)
NaCl
3x(M)

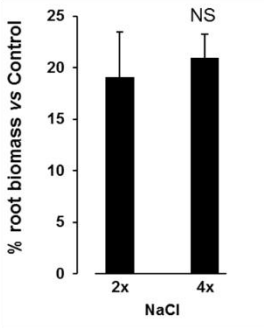

% root biomass vs Control
NS
30
25
20
15
10
5
0
3x(M)
3x(P)
NaCl

T910

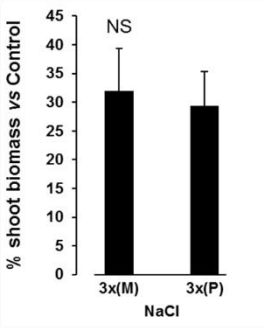

4x
Control
2x

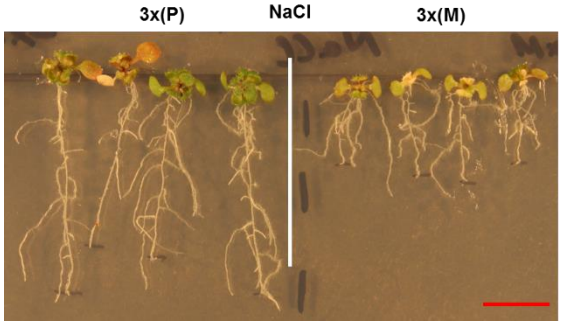

% shoot biomass vs Control
50
40
30
20
10
0
NS
2x
4x
NaCl

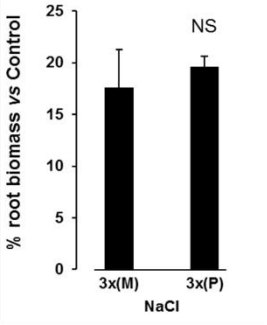

4x
NaCl
2x


% root biomass vs Control
25
20
15
10
5
0
NS
2x
4x
NaCl

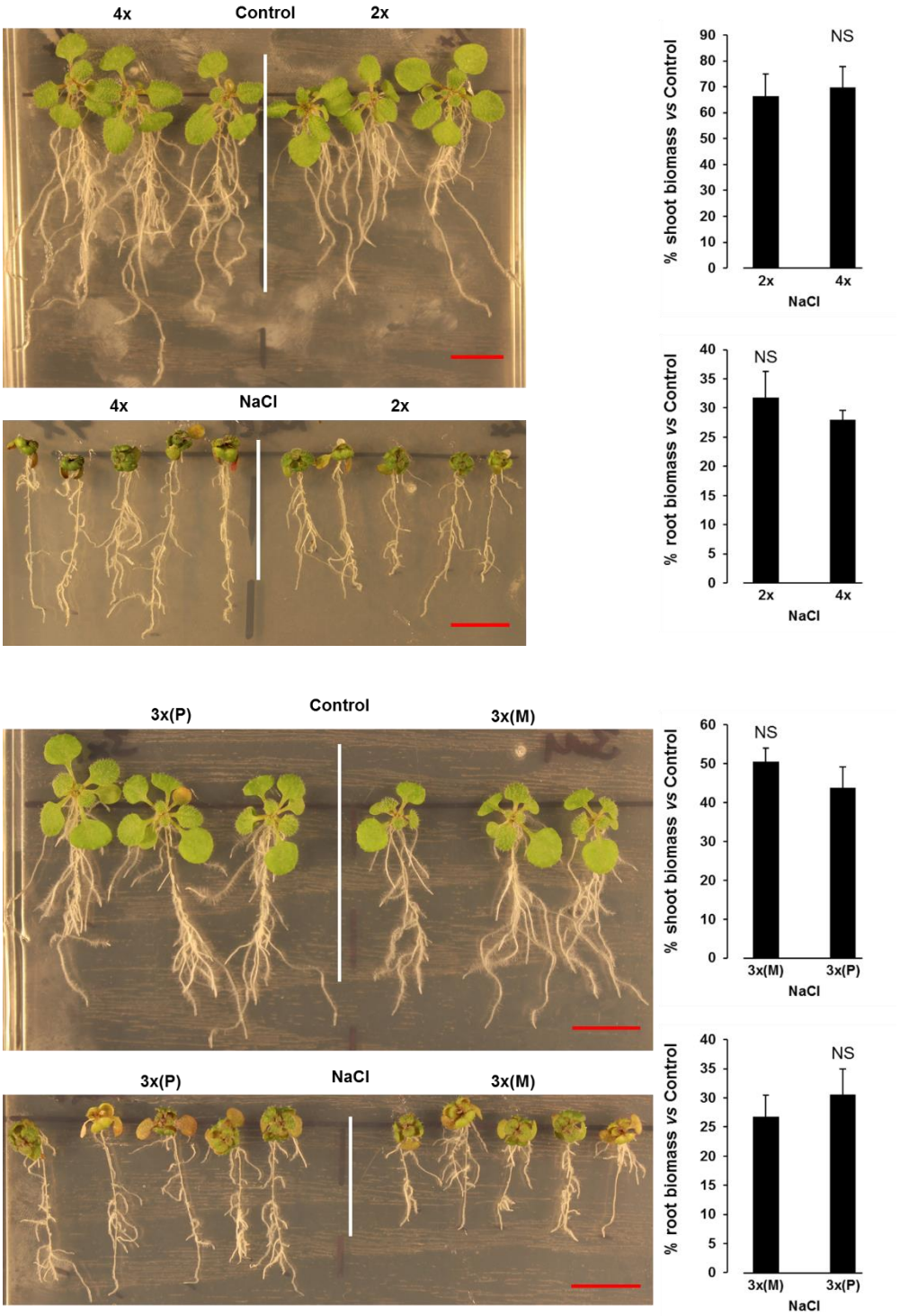

3x(P)
Control
3x(M)


% shoot biomass vs Control
45
40
35
30
25
20
15
10
5
0
NS
3x(M)
3x(P)
NaCl

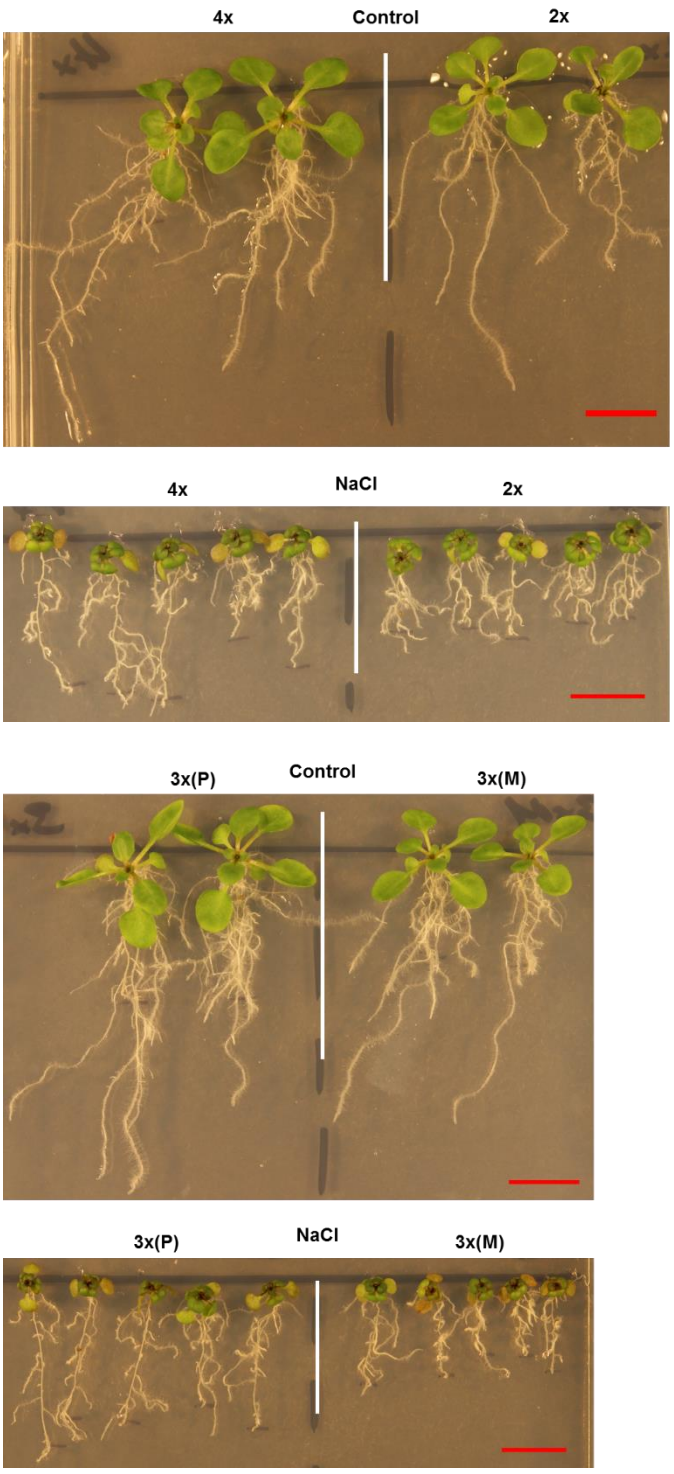

3x(P)
NaCl
3x(M)

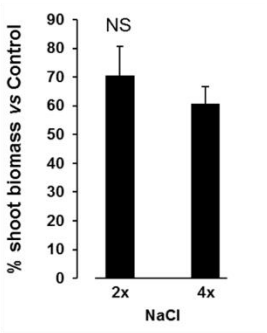

% root biomass vs Control
25
20
15
10
5
0
NS
3x(M)
3x(P)
NaCl

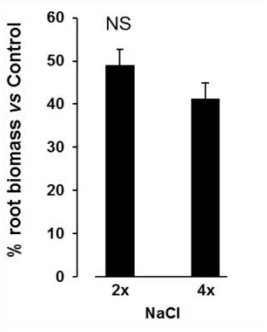
TAL07

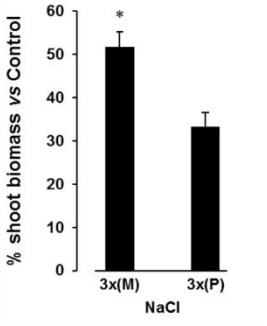

4x
Control
2x
4x
NaCl
2x
3x(P)
Control
3x(M)
3x(P)
NaCl
3x(M)
% shoot biomass vs Control
% root biomass vs Control
NS
2x
4x
NaCl
3x(M)
3x(P)

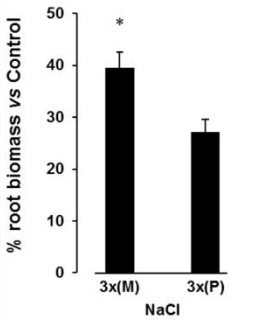
Wilna


4x
Control
2x
4x
NaCl
2x
3x(P)
Control
3x(M)
3x(P)
NaCl
3x(M)

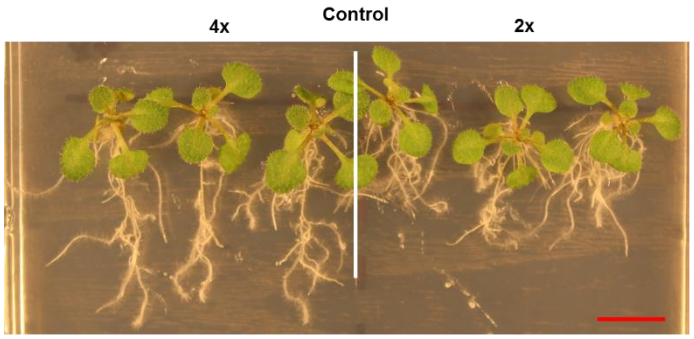

% shoot biomass vs Control
NS
2x
4x
NaCl

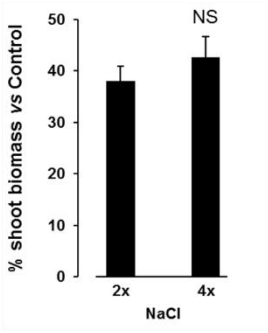

% root biomass vs Control
NS
2x
4x
NaCl

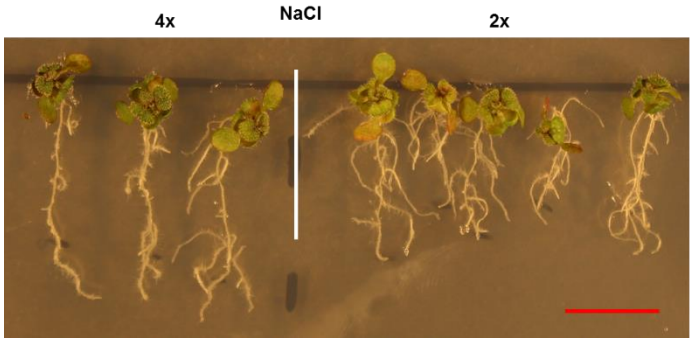

% shoot biomass vs Control
*
3x(M)
3x(P)
NaCl

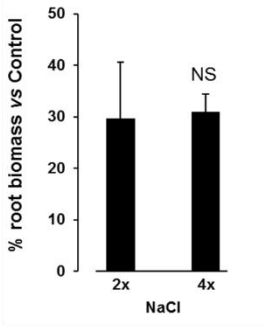

% root biomass vs Control
*
3x(M)
3x(P)
NaCl

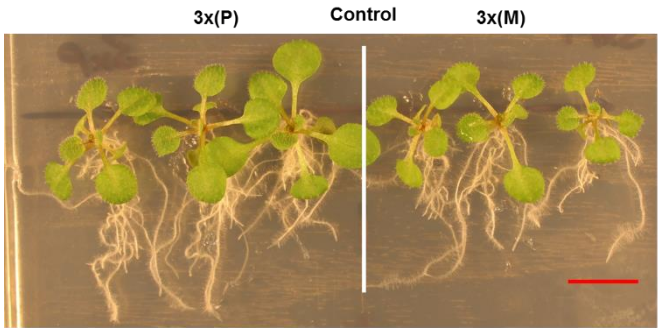
Zurich

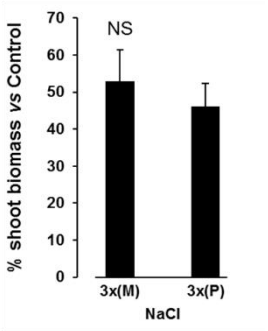

4x
Control
2x

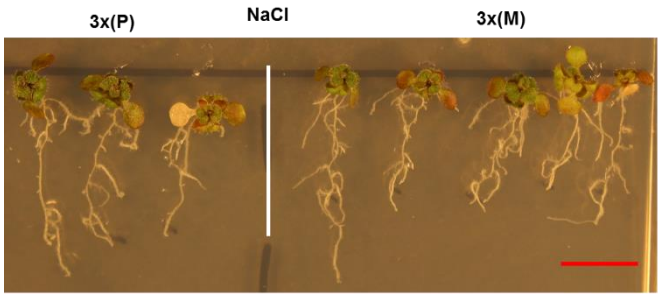

% shoot biomass vs Control
NS
50
40
30
20
10
0
2x
4x
NaCl

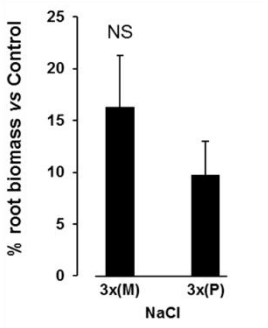

4x
NaCl
2x


% root biomass vs Control
NS
50
40
30
20
10
0
2x
4x
NaCl


3x(P)
Control
3x(M)


% shoot biomass vs Control
NS
70
60
50
40
30
20
10
0
3x(M)
3x(P)
NaCl

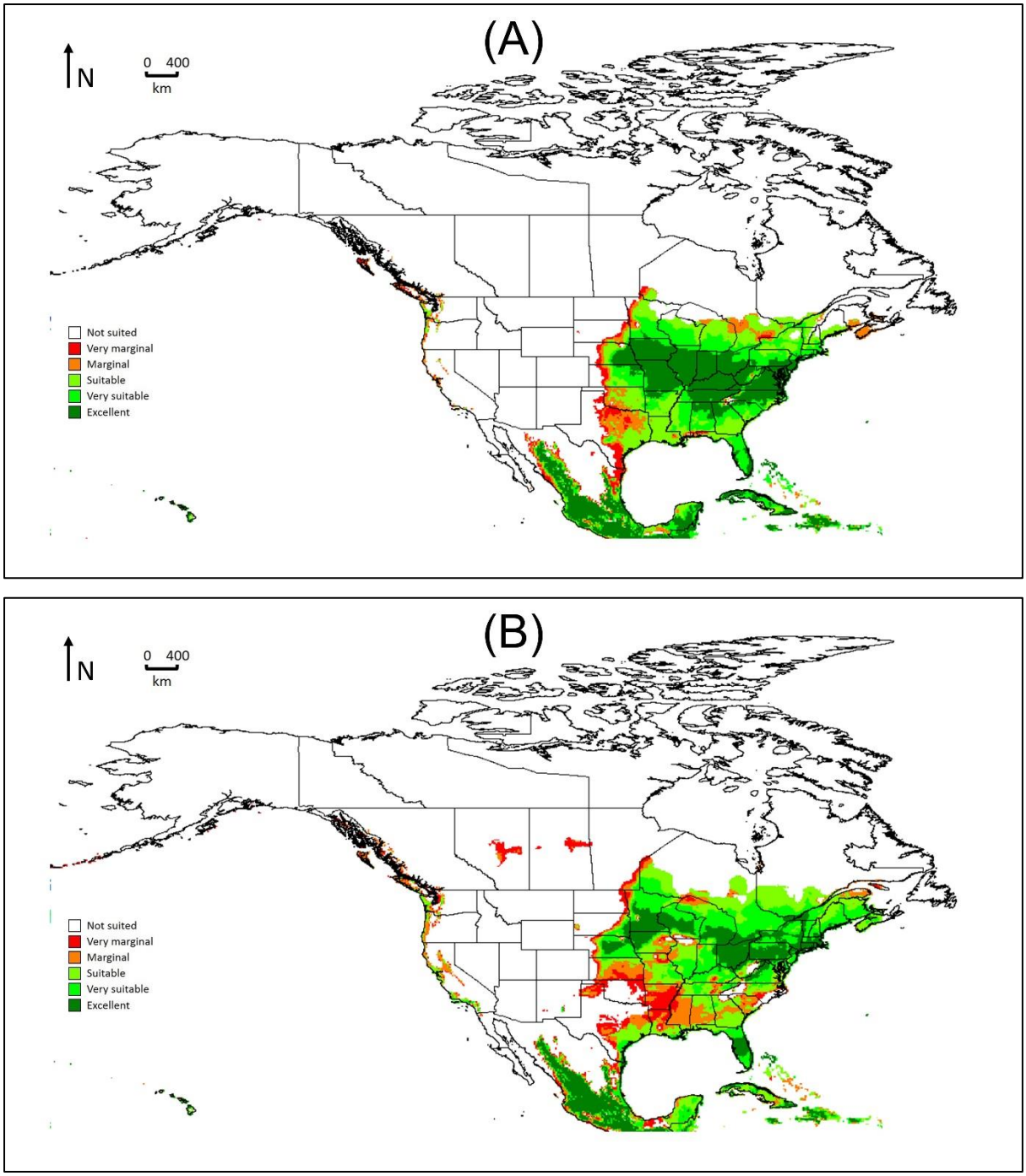

3x(P)
NaCl
3x(M)


% root biomass vs Control
NS
25
20
15
10
5
0
3x(M)
3x(P)
NaCl

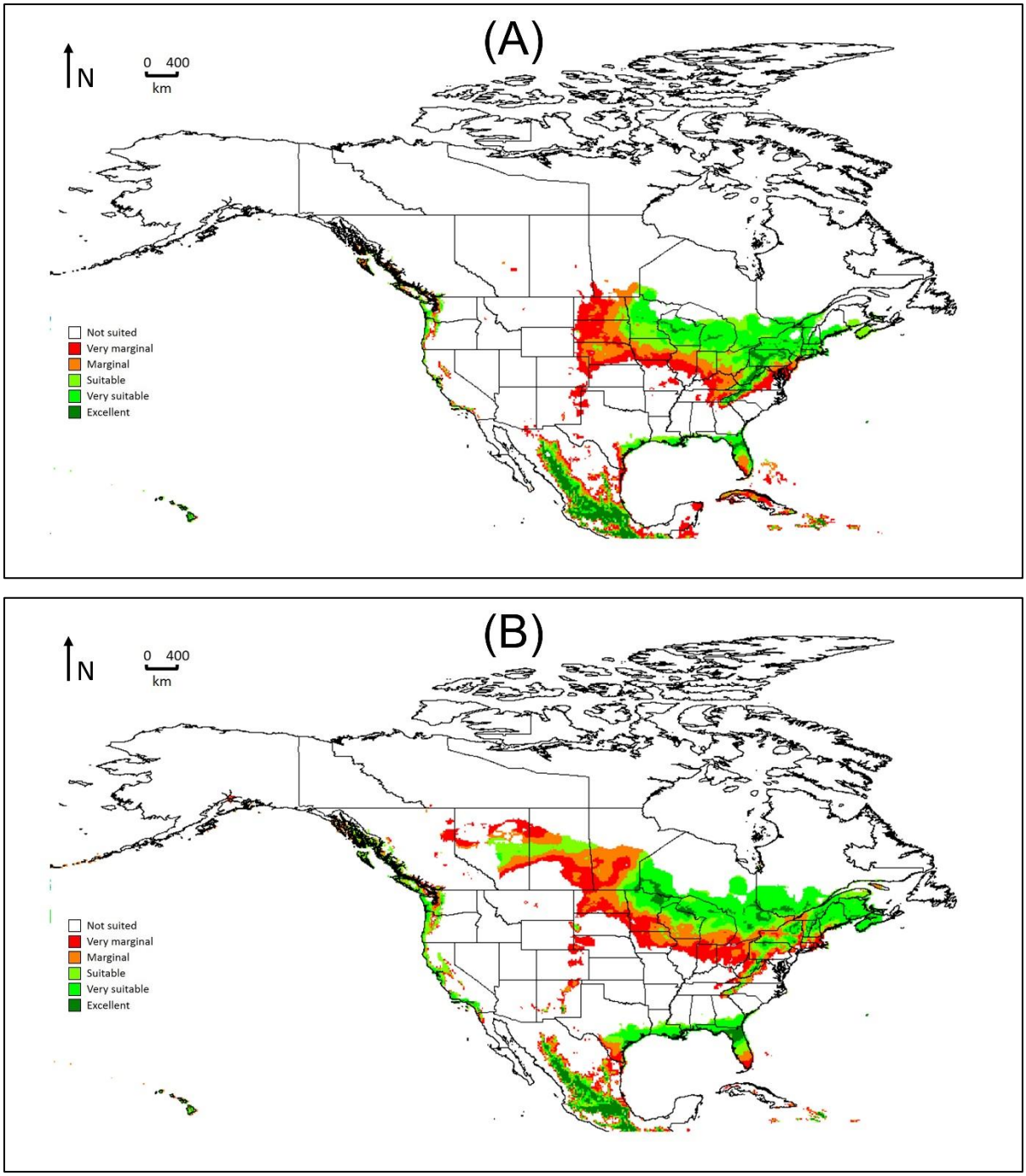
Appendix 4

Chapter 6. Supplementary Results.

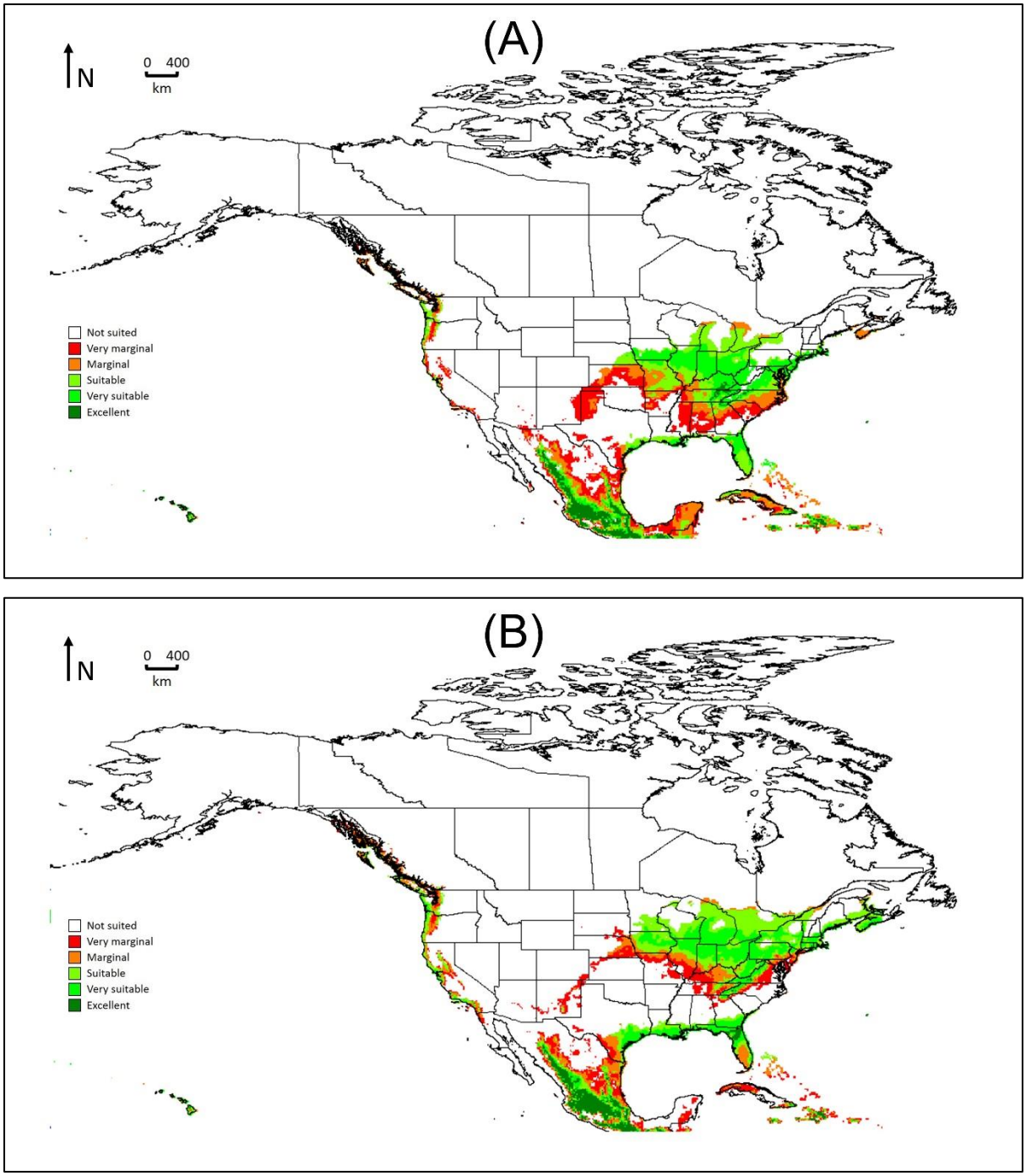
Figures of EcoCrop maps generated to analyse baseline (1960-1990) and future (2050, RCP 6.0) growing area suitability.



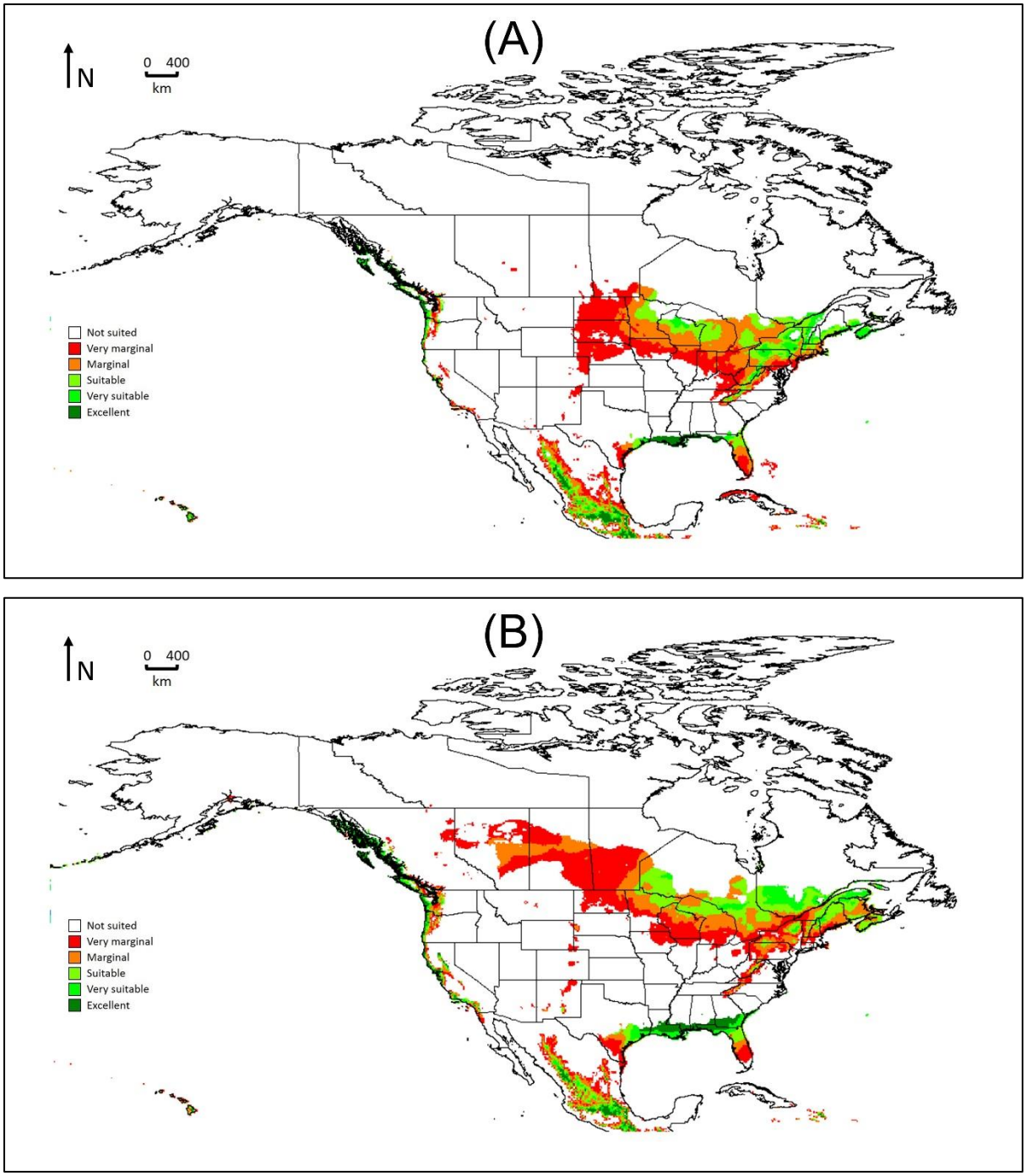
Supplementary Figure 1. Durum wheat (*Triticum turgidum* subsp. *durum*) suitable growing areas for baseline (A) and future (B) scenarios.



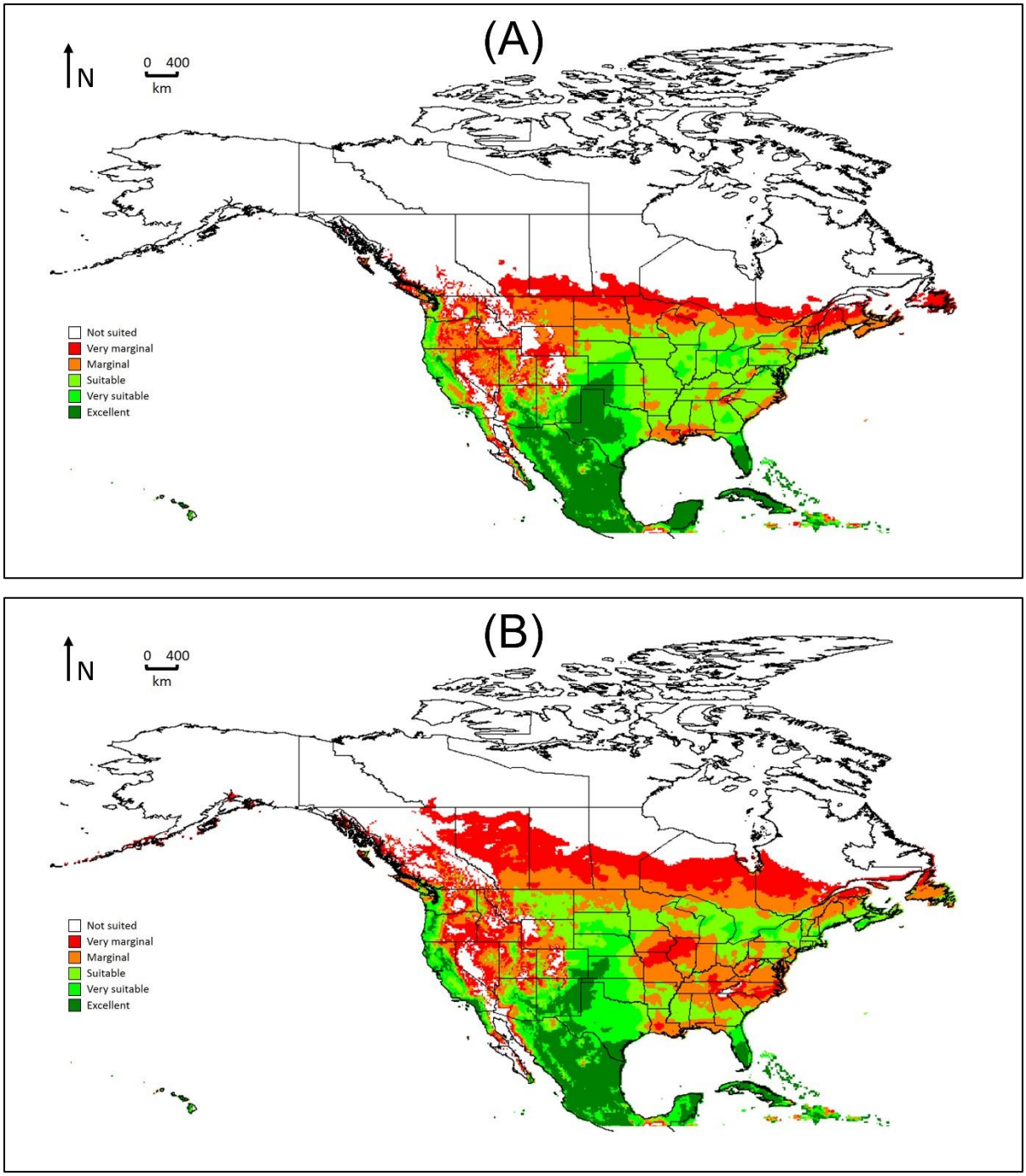
Supplementary Figure 2. Emmer wheat (*Triticum turgidum* subsp. *diccocon*) suitable growing areas for baseline (A) and future (B) scenarios.



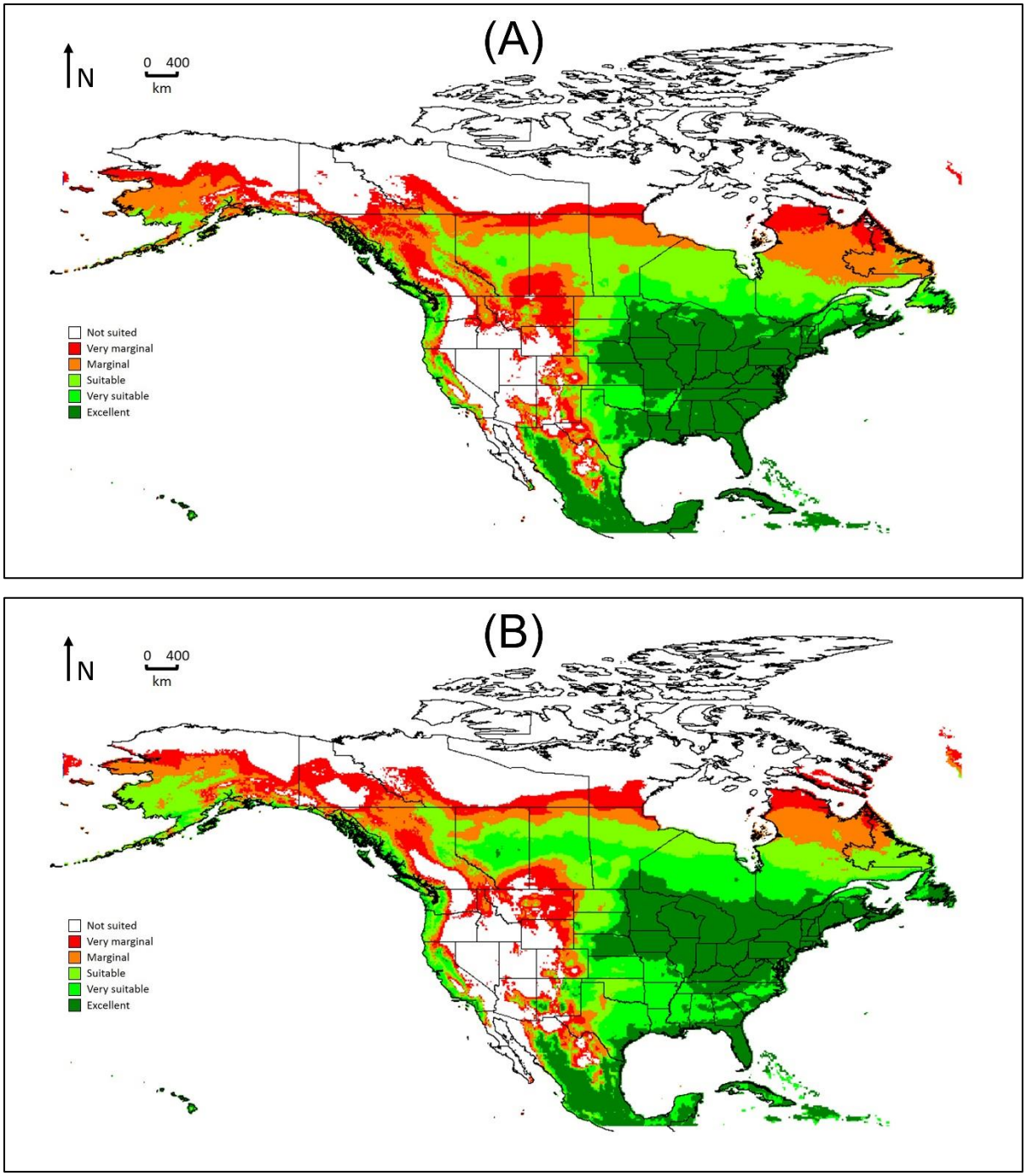
Supplementary Figure 3. Bread wheat (*Triticum aestivum* subsp. *aestivum*) suitable growing areas for baseline (A) and future (B) scenarios.



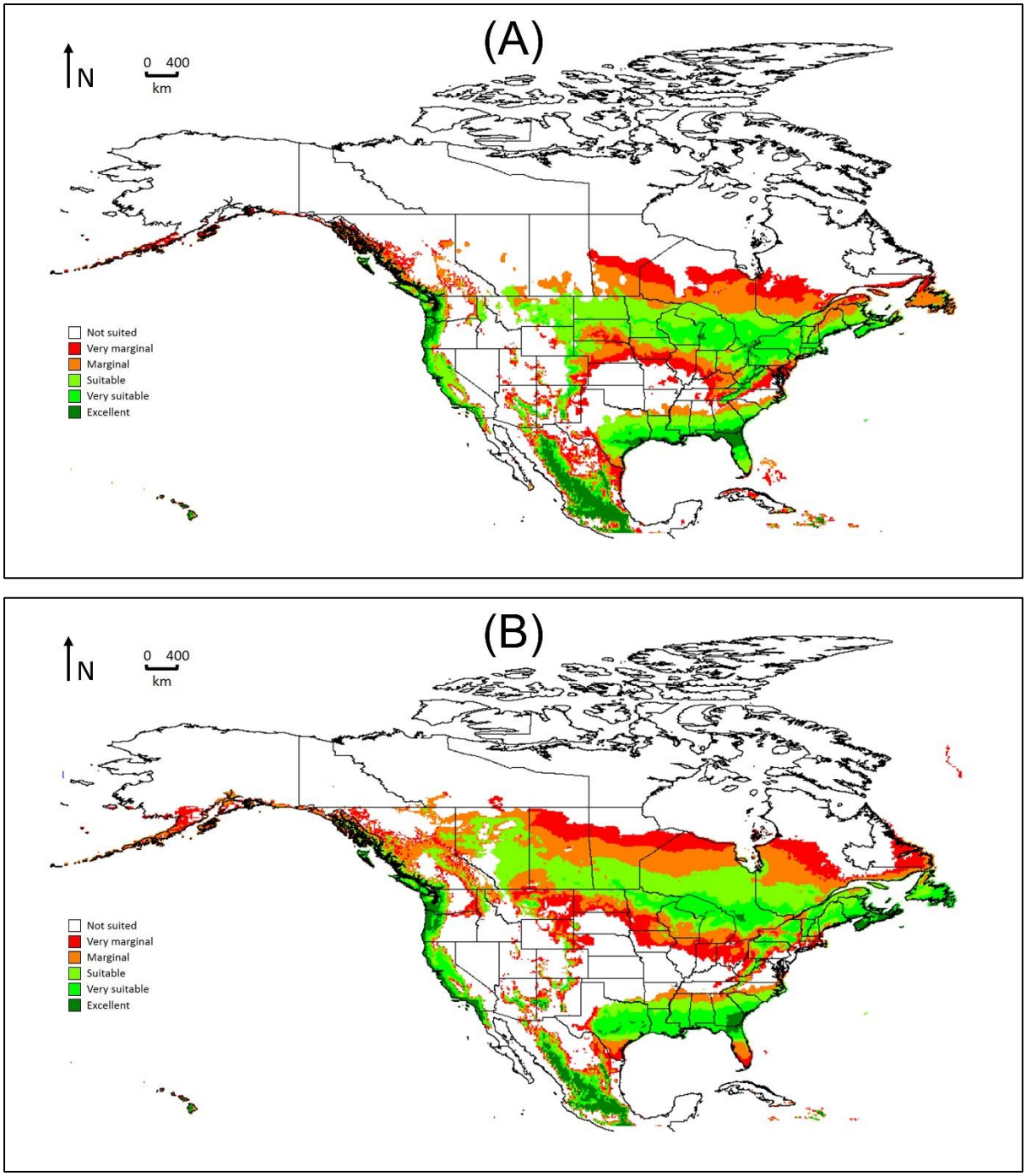
Supplementary Figure 4. Spelt wheat (*Triticum aestivum* subsp. *spelta*) suitable growing areas for baseline (A) and future (B) scenarios.



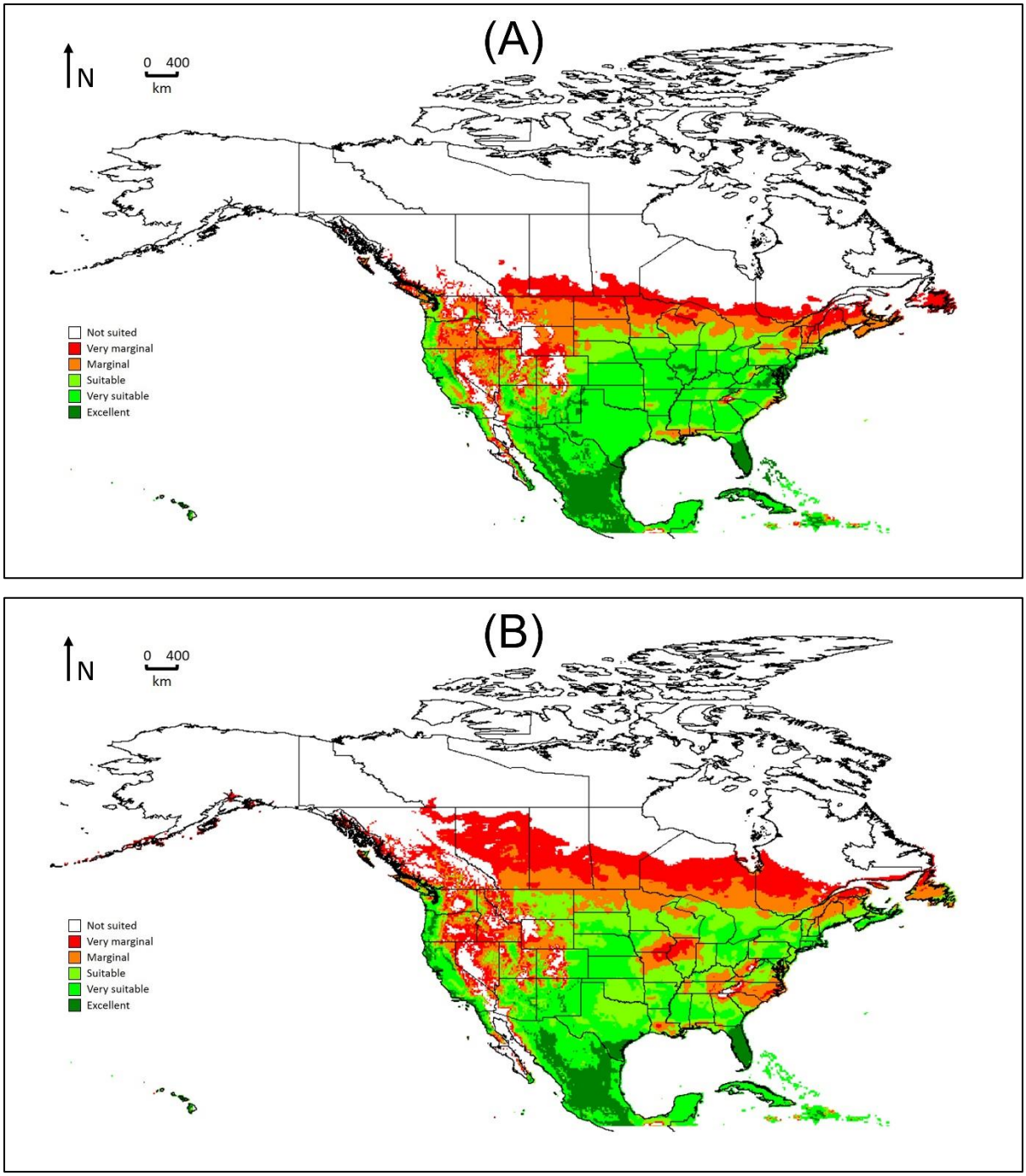
Supplementary Figure 5. Crested wheatgrass (*Agropyron cristatum*) suitable growing areas for baseline (A) and future (B) scenarios.



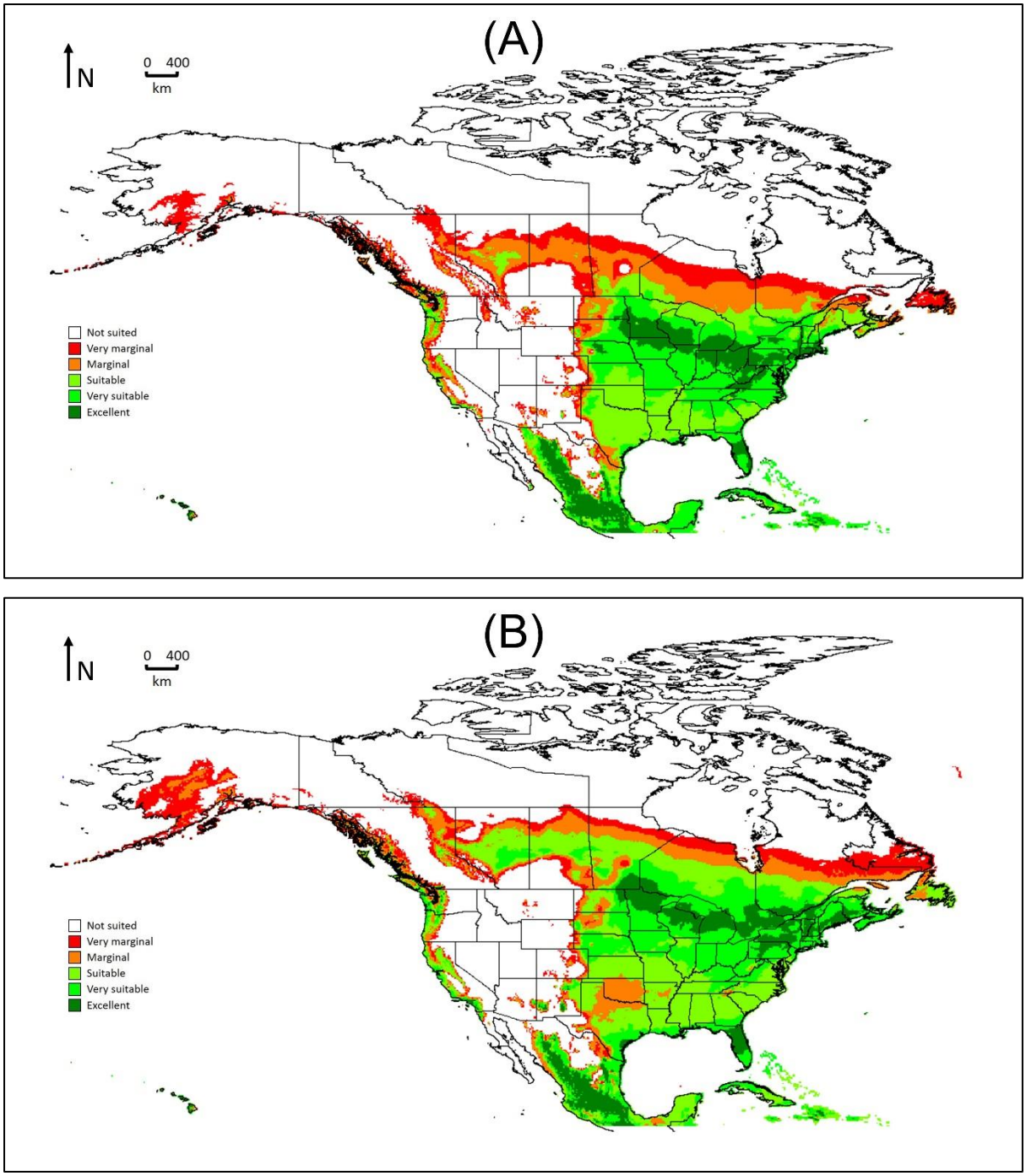
Supplementary Figure 6. Standard wheatgrass (*Agropyron desertorum*) suitable growing areas for baseline (A) and future (B) scenarios.



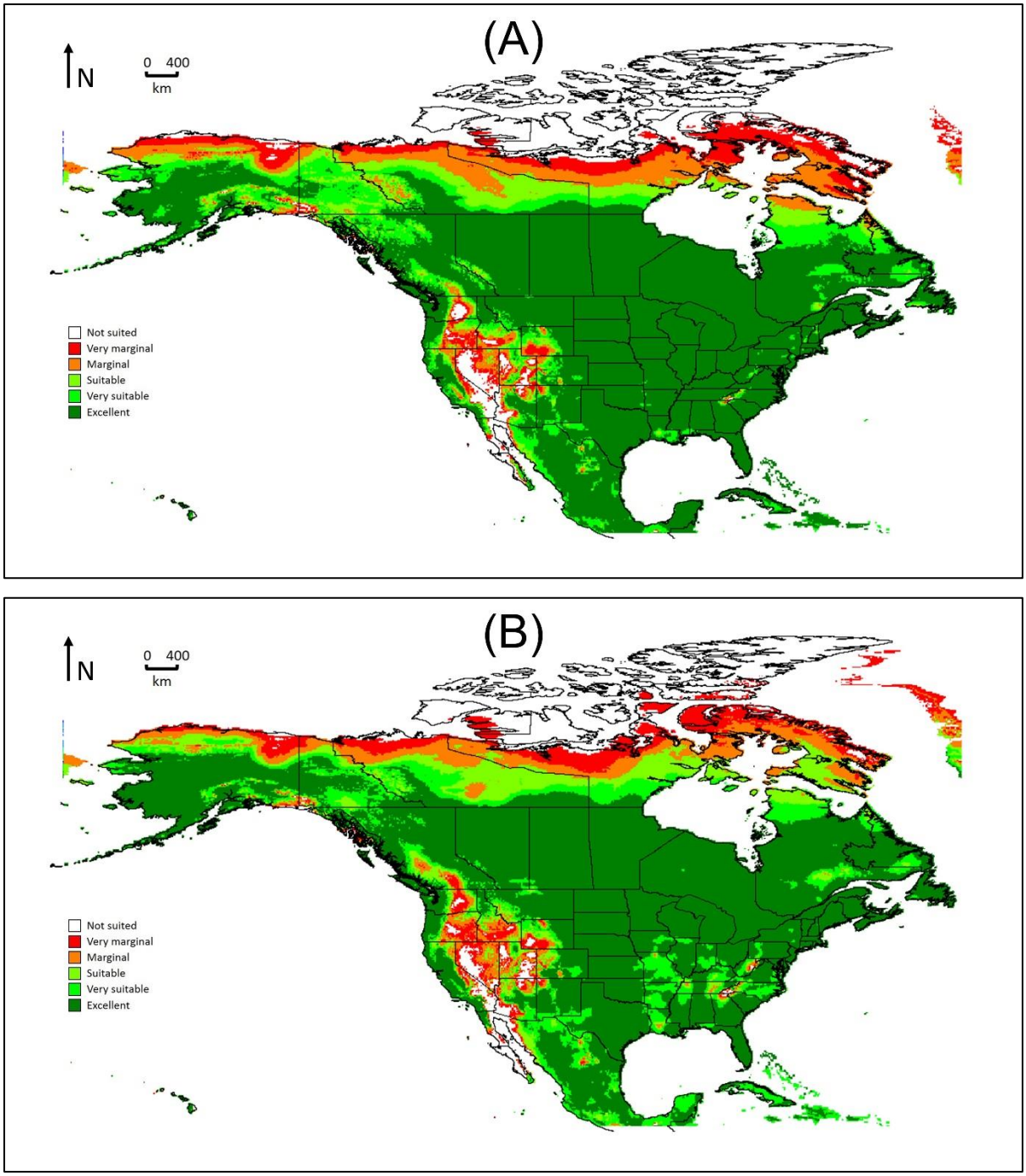
Supplementary Figure 7. Tall wheatgrass (*Agropyron elongatum*) suitable growing areas for baseline (A) and future (B) scenarios.



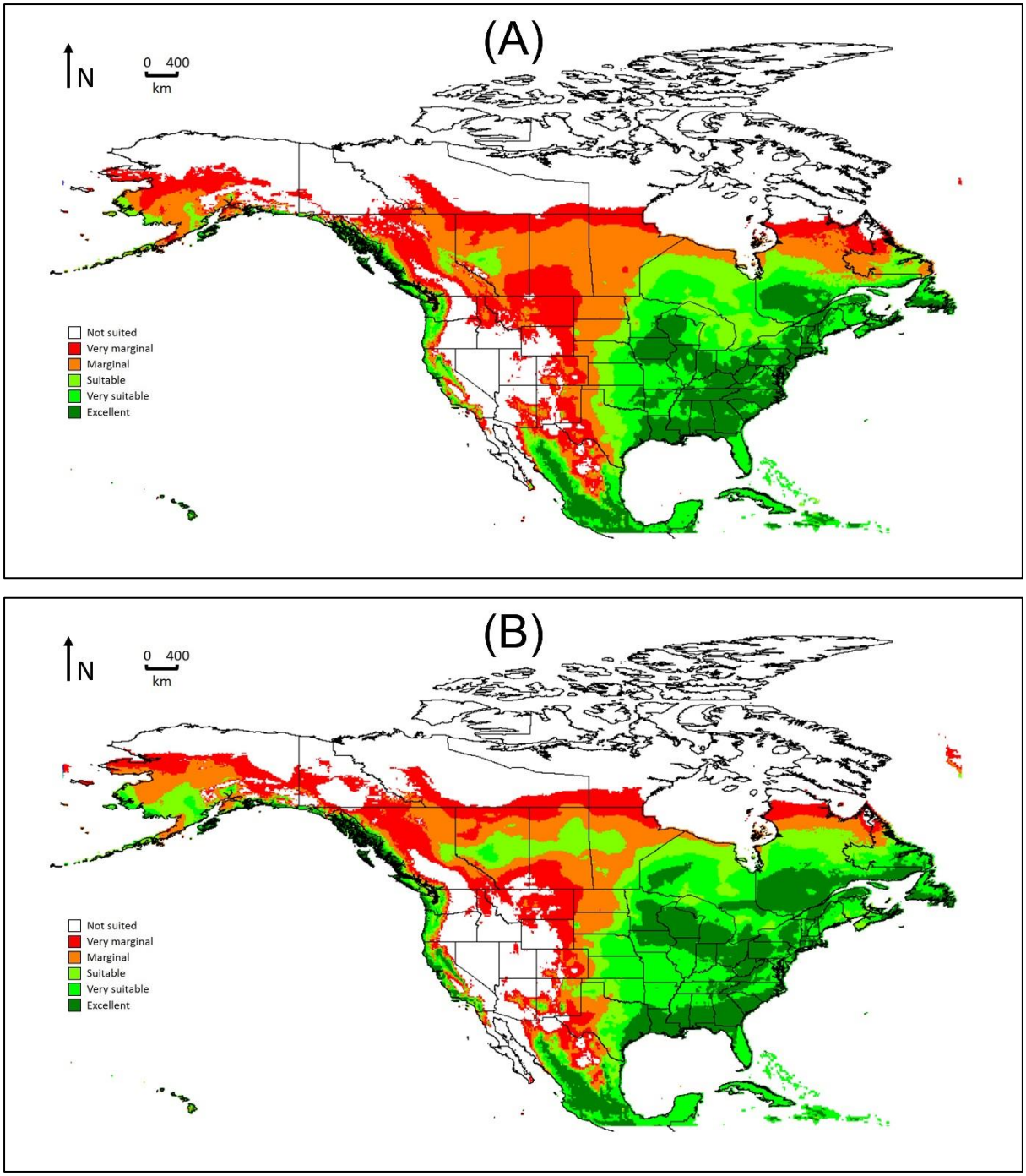
Supplementary Figure 8. Siberian wheatgrass (*Agropyron fragile*) suitable growing areas for baseline (A) and future (B) scenarios.



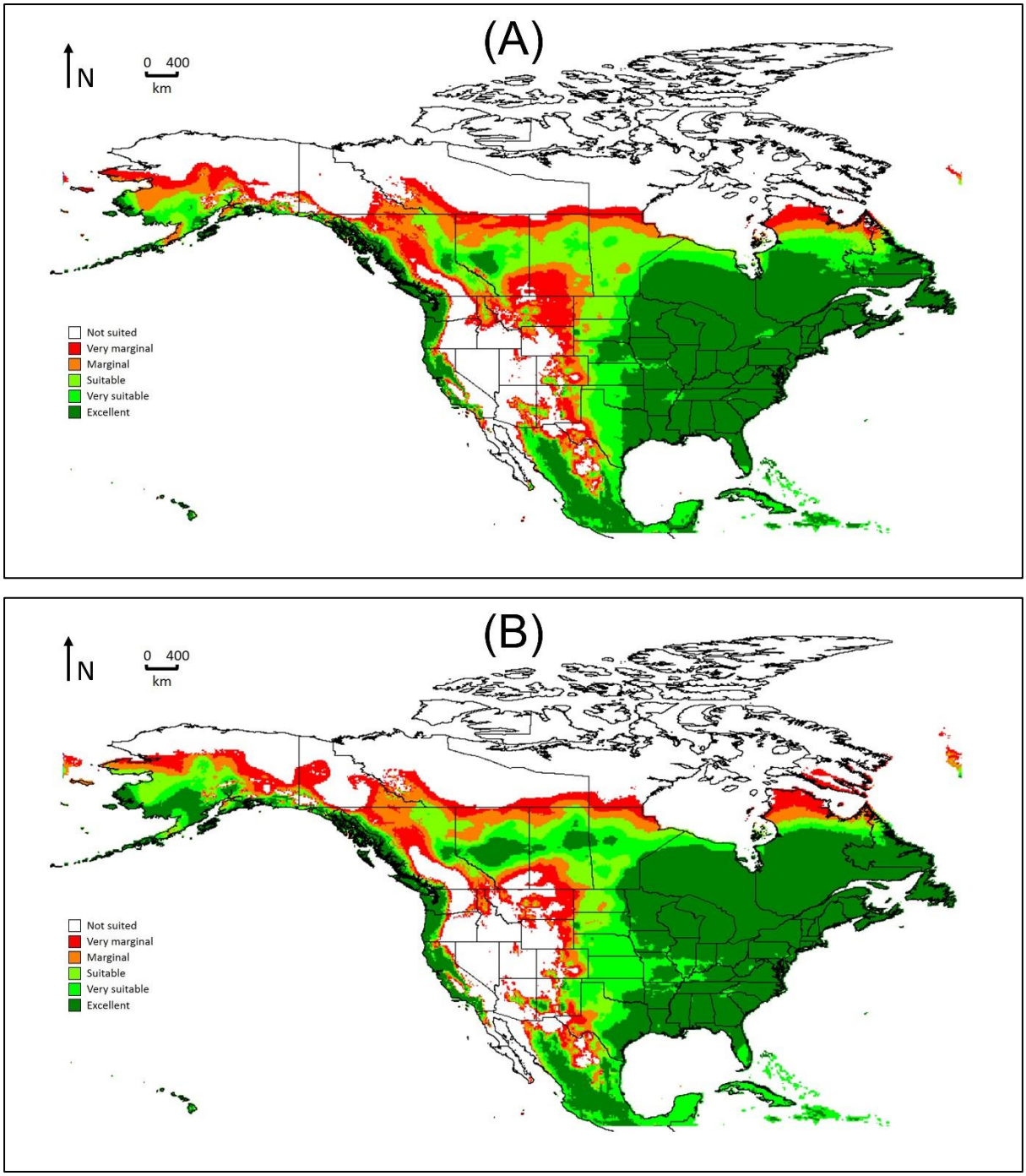
Supplementary Figure 9. Intermediate wheatgrass (*Agropyron intermedium*) suitable growing areas for baseline (A) and future (B) scenarios.



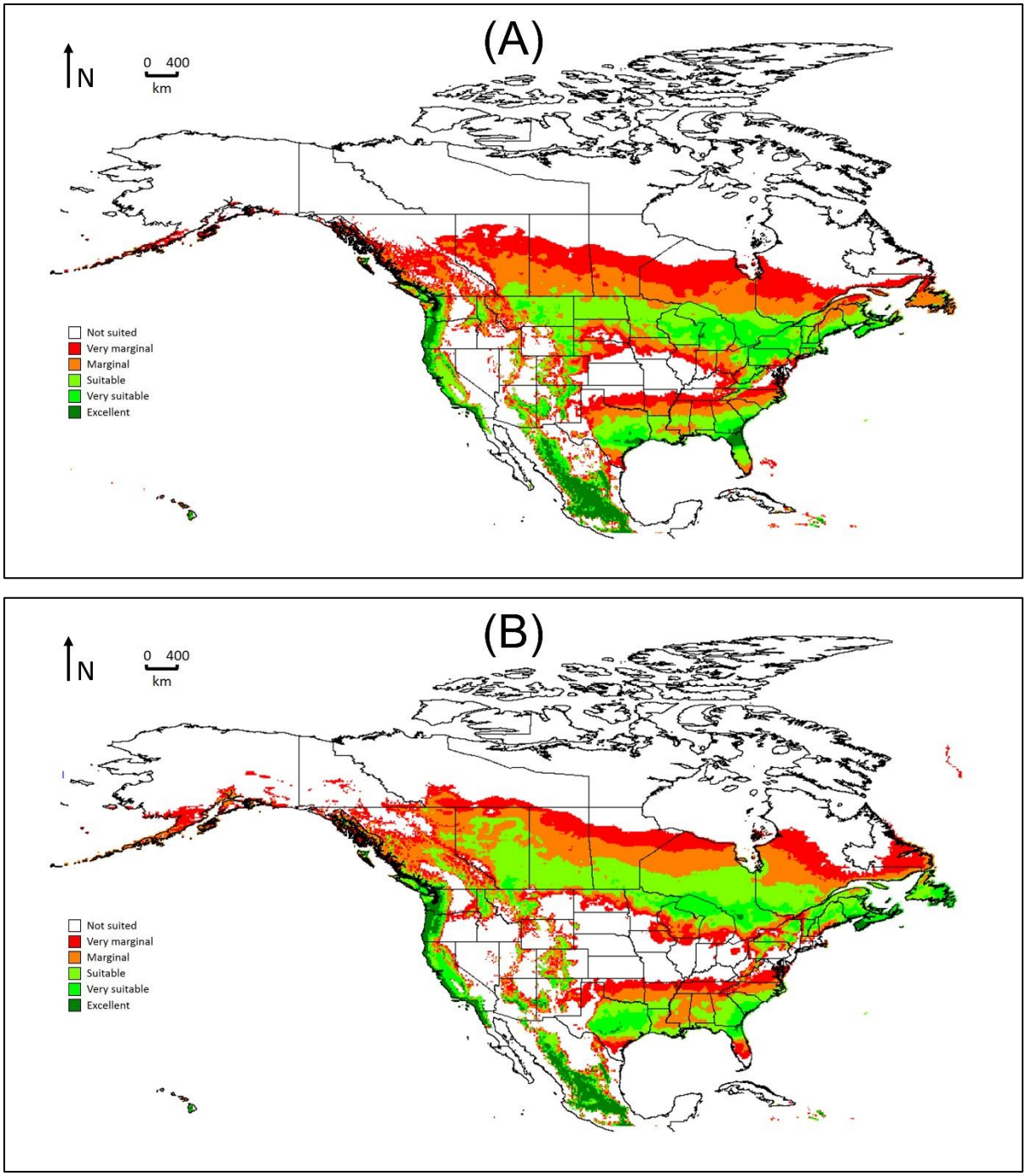
Supplementary Figure 10. Mongolian wheatgrass (*Agropyron mongolicum*) suitable growing areas for baseline (A) and future (B) scenarios.



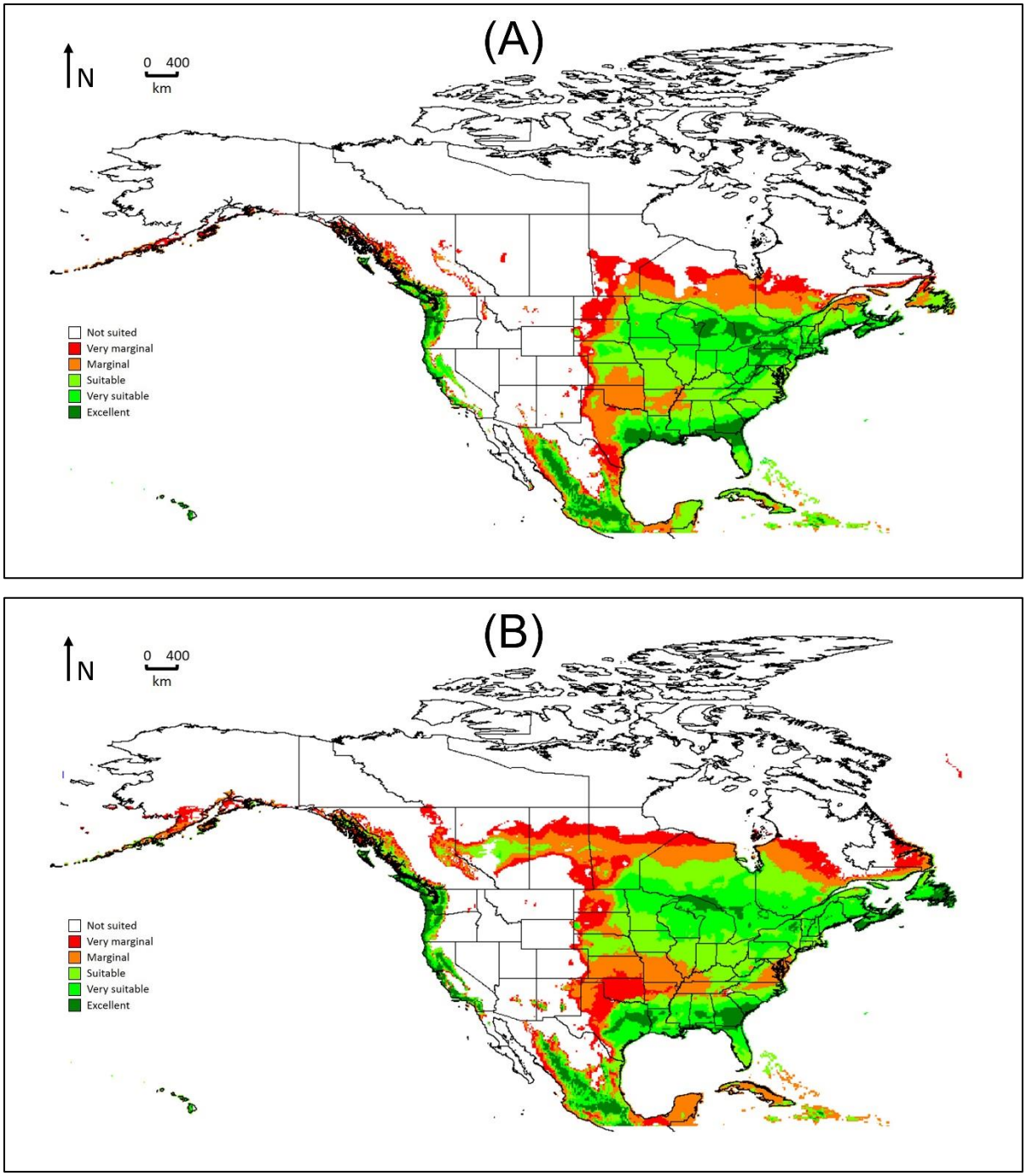
Supplementary Figure 11. Couch grass (*Agropyron repens*) suitable growing areas for baseline (A) and future (B) scenarios.



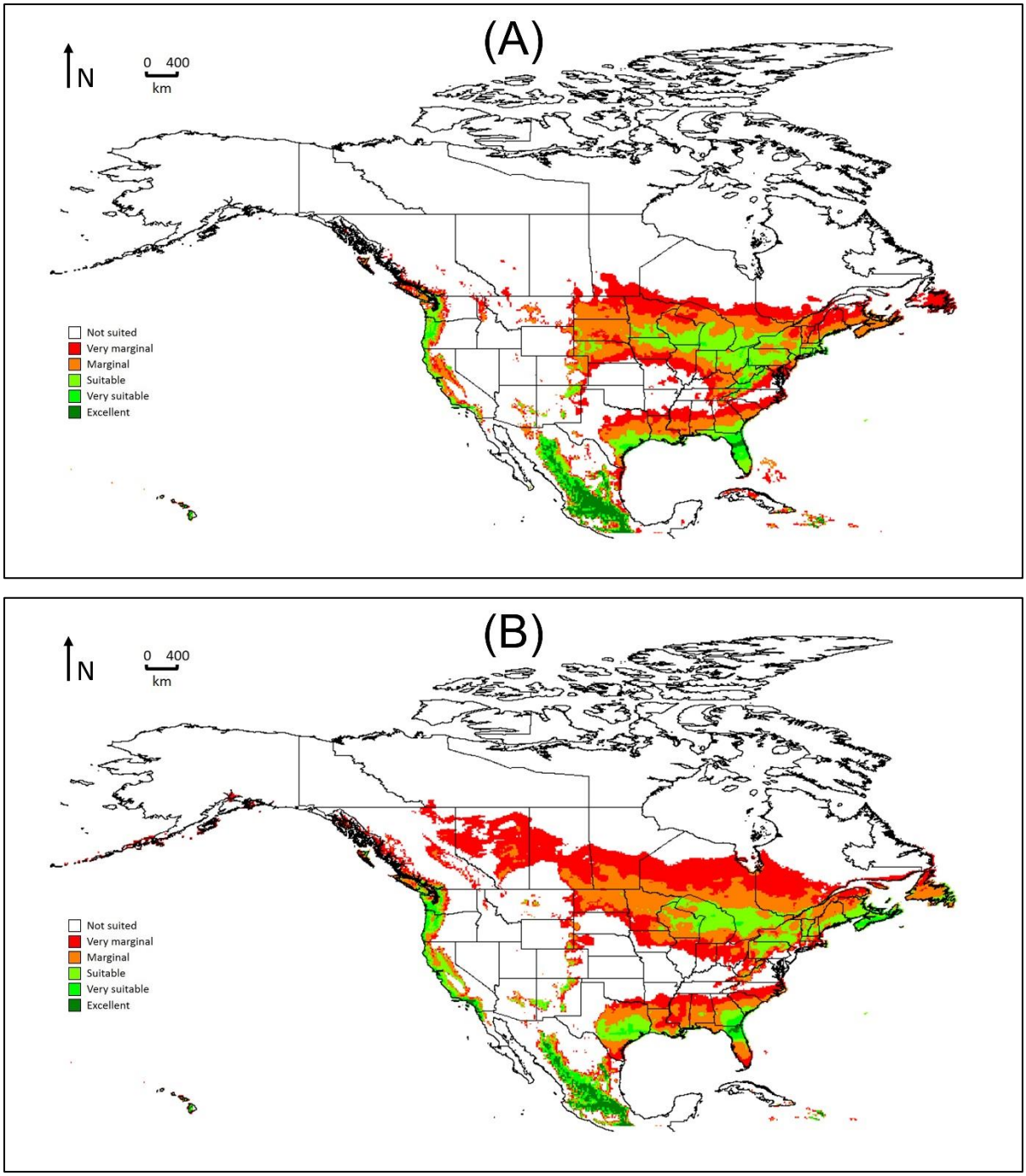
Supplementary Figure 12. Western wheatgrass (*Agropyron smithii*) suitable growing areas for baseline (A) and future (B) scenarios.



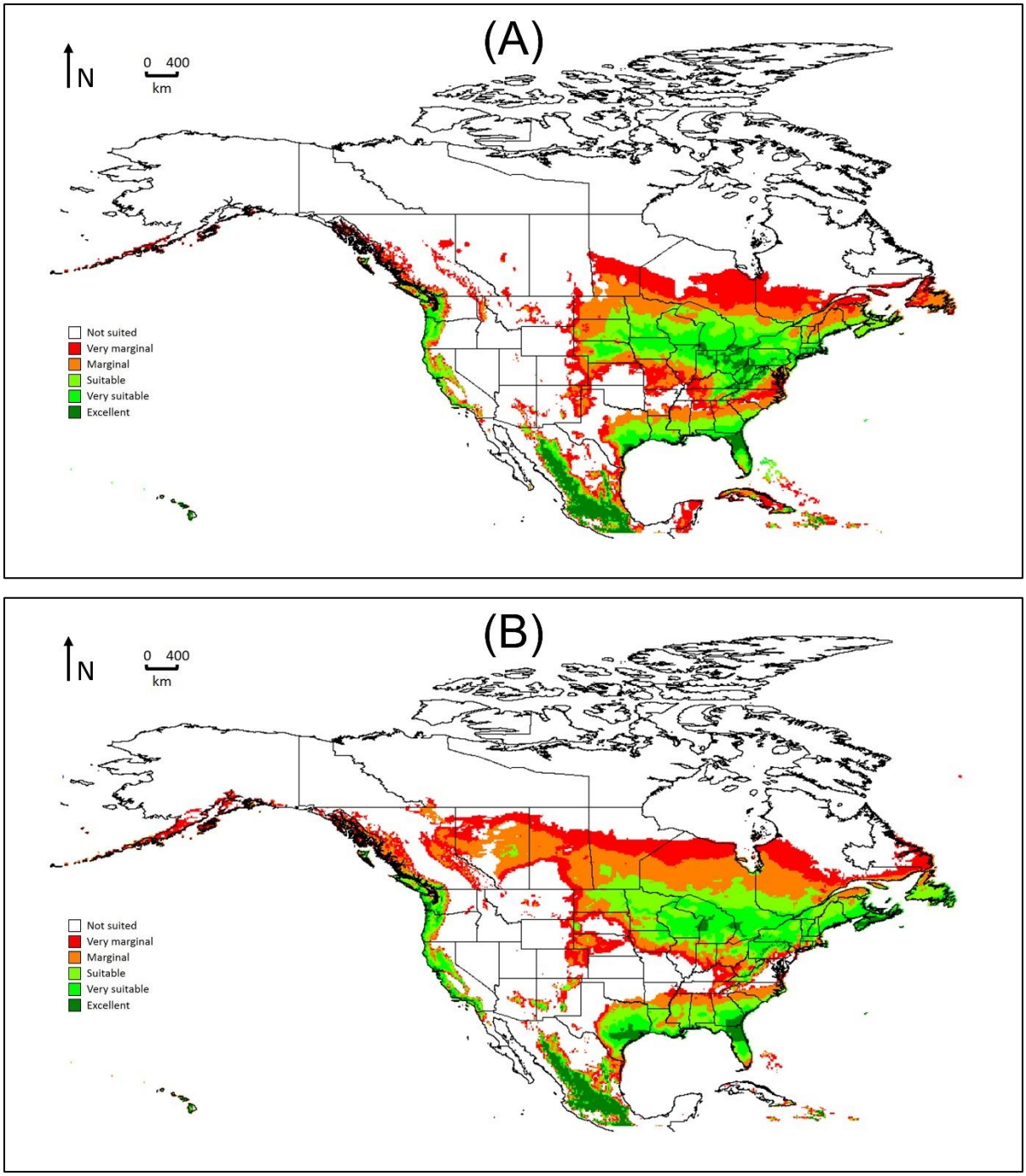
Supplementary Figure 13. Bluebunch wheatgrass (*Agropyron spicatum*) suitable growing areas for baseline (A) and future (B) scenarios.



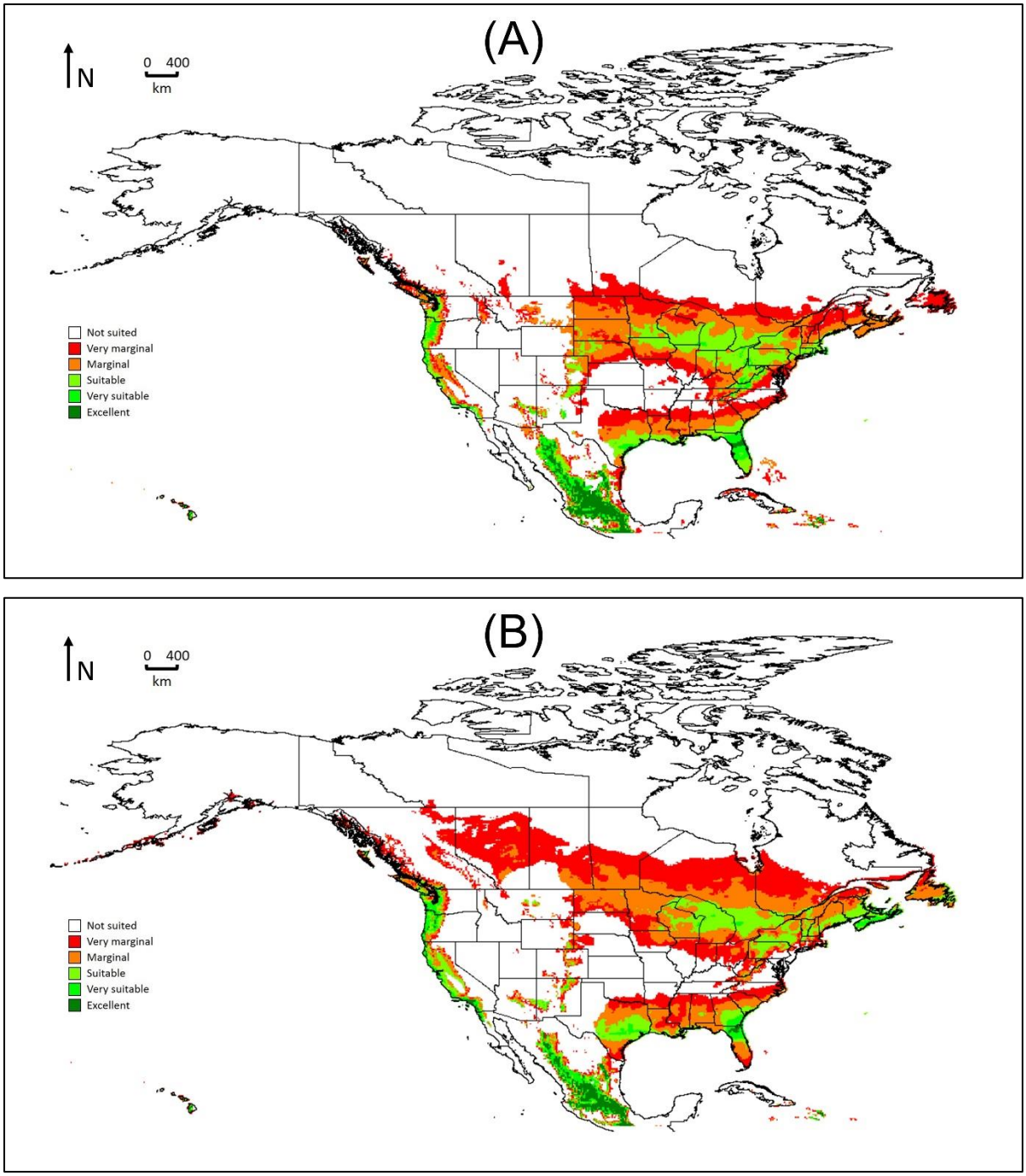
Supplementary Figure 14. Streambank wheatgrass (*Agropyron riparium*) suitable growing areas for baseline (A) and future (B) scenarios.



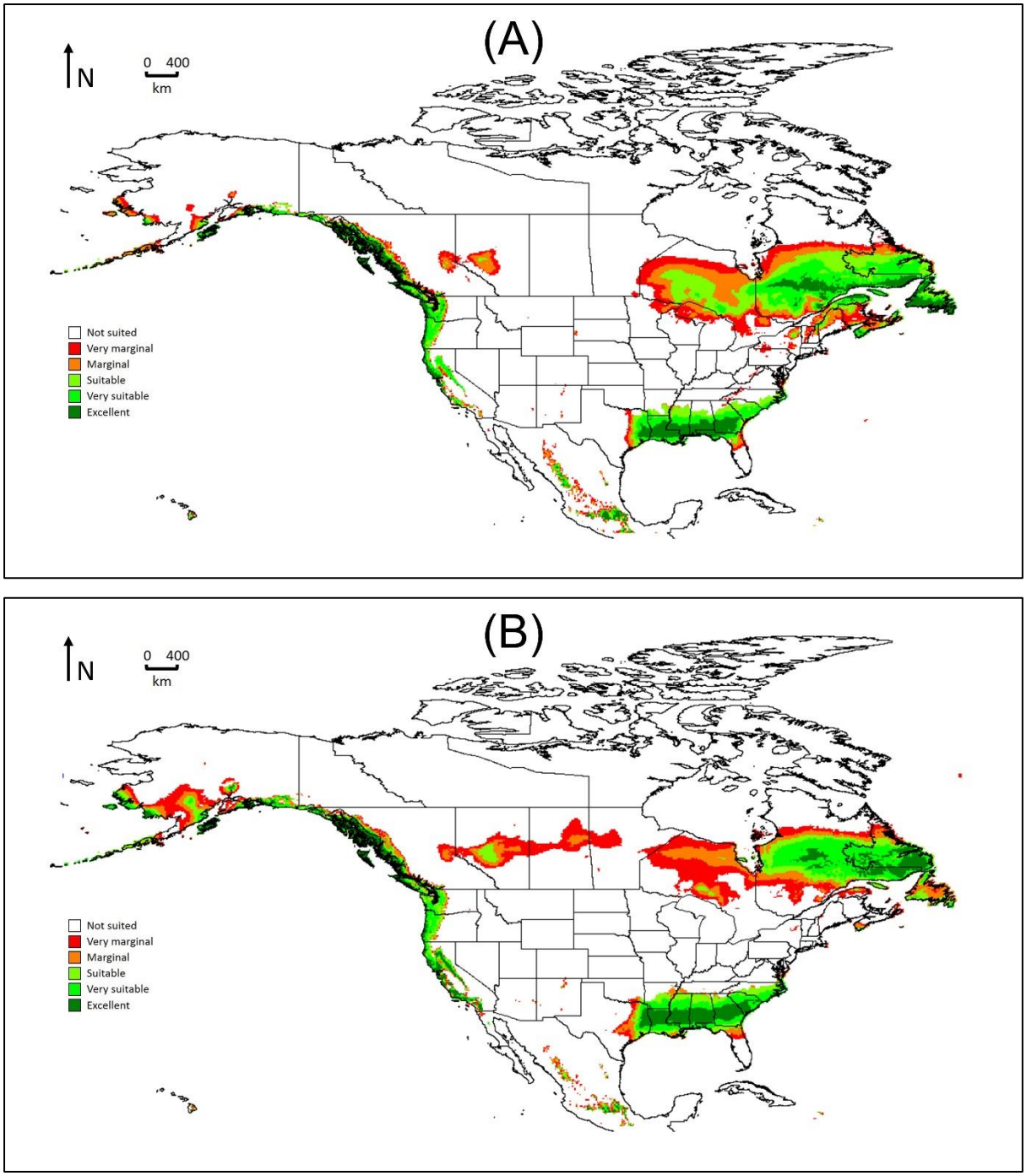
Supplementary Figure 15. *Agropyron tauri* suitable growing areas for baseline (A) and future (B) scenarios.



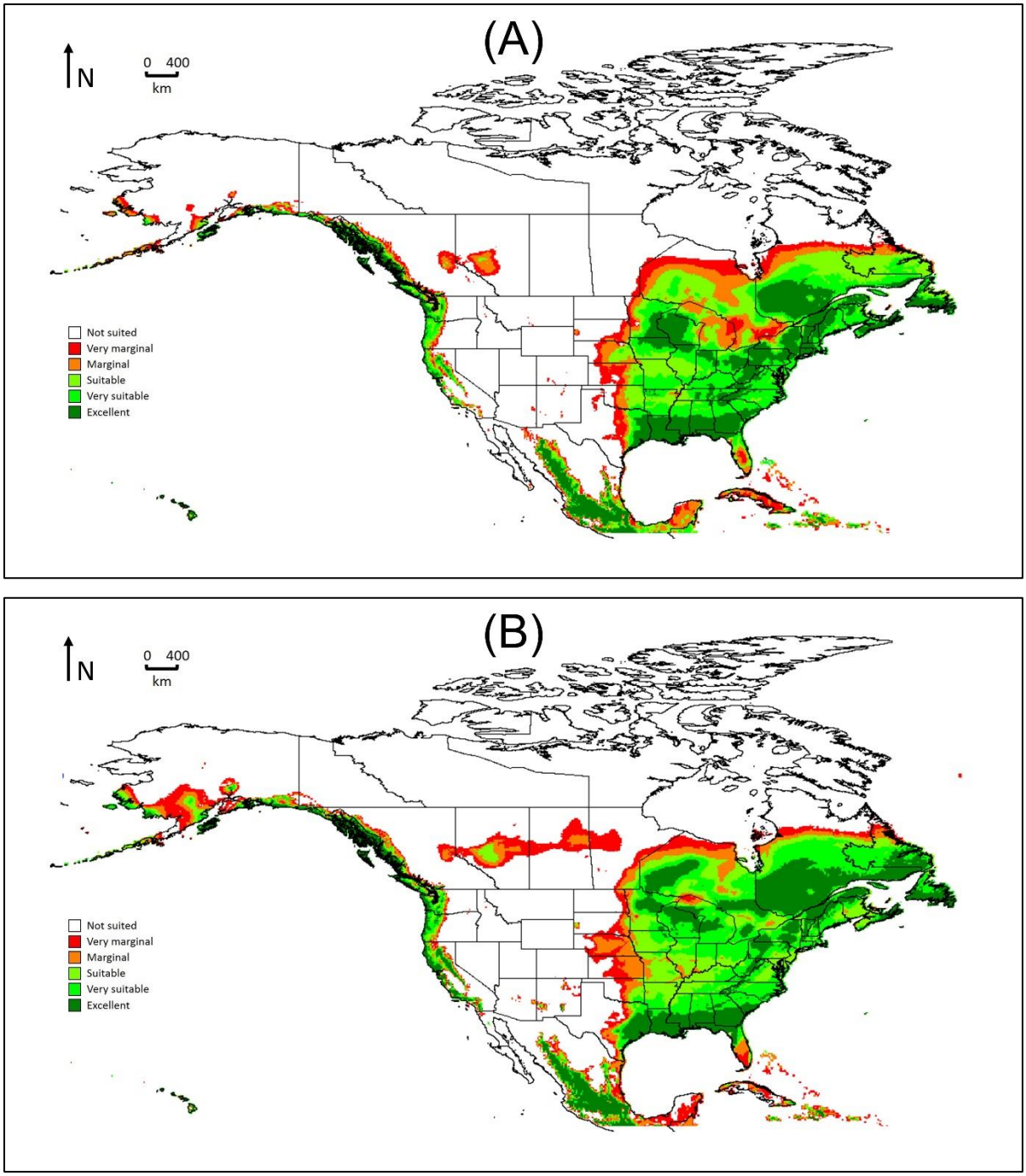
Supplementary Figure 16. Slender wheatgrass (*Agropyron trachycaulum*) suitable growing areas for baseline (A) and future (B) scenarios.



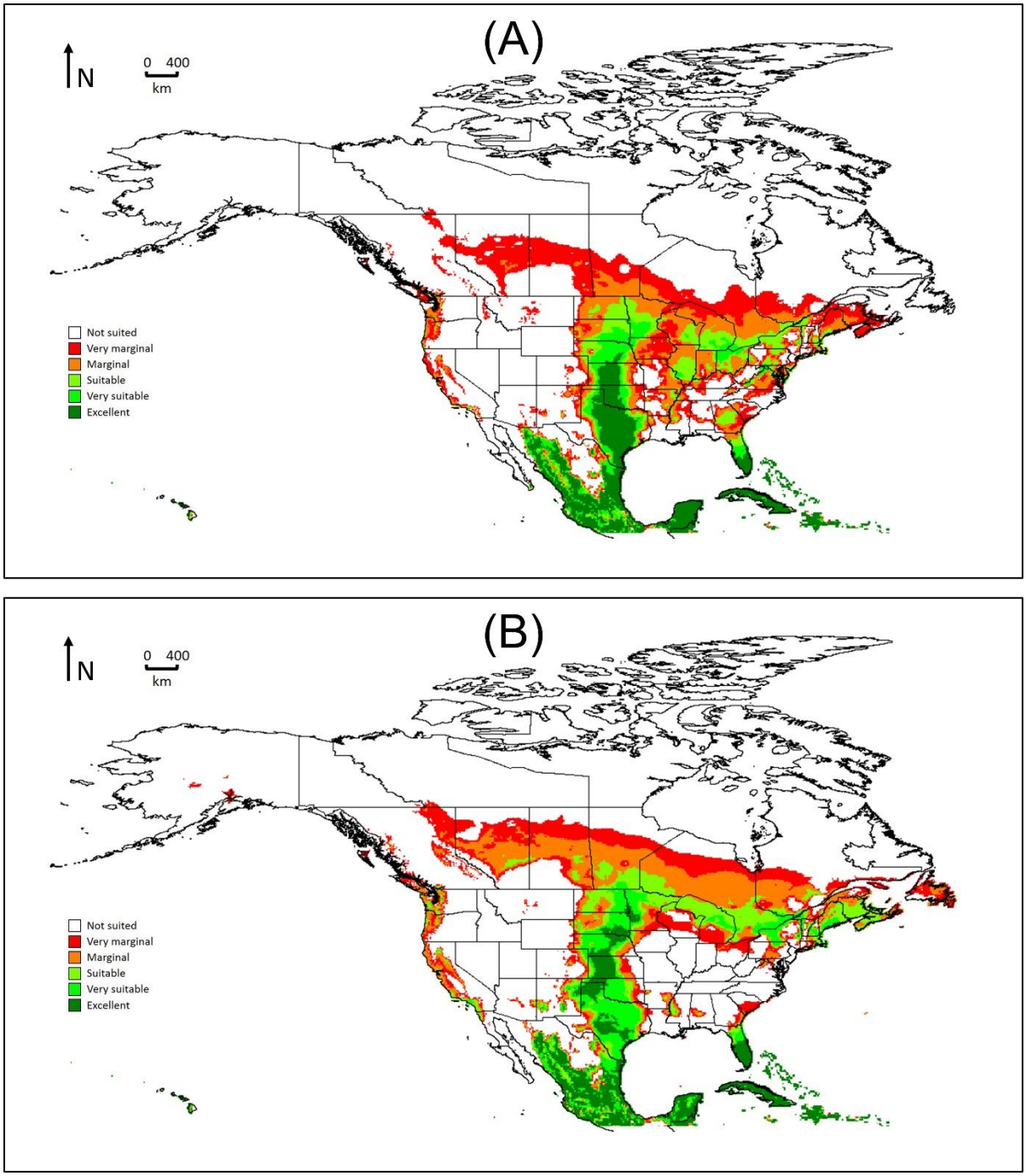
Supplementary Figure 17. Stiff hair wheatgrass (*Agropyron trichophorum*) suitable growing areas for baseline (A) and future (B) scenarios.



Supplementary Figure 18. Blue wildrye (*Elymus glaucus*) suitable growing areas for baseline (A) and future (B) scenarios.



Supplementary Figure 19. Canada wildrye (*Elymus canadensis*) suitable growing areas for baseline (A) and future (B) scenarios.



Supplementary Figure 20. *Elymus dahuricus* suitable growing areas for baseline (A) and future (B) scenarios.

Appendix 5

Reduction in nutritional quality and growing area suitability of common bean under climate change induced drought stress in Africa.

Introduction

Dietary deficiencies of micronutrients such as iron and zinc constitute major public health problems globally, particularly amongst women and children in sub-Saharan Africa (Bailey et al. 2015). While micronutrient supplementation and food fortification are important for improving delivery of micronutrients, staple food crop biofortification through breeding provides an additional route for increasing the supply of key micronutrients (iron, zinc, vitamin A) from staple crops to the diets of poorer communities in developing countries (Imdad et al. 2010; Bouis et al. 2011; Saltzman et al. 2013; Anjali 2015). The levels of micronutrients (e.g. iron, zinc) in staple crops and foods is one of the key determinants of the extent of uptake of dietary micronutrients (Sanahuja et al. 2013; Díaz-Gómez et al. 2017). However, the presence and levels of anti-nutritionals, in particular phytic acid and polyphenols, can inhibit bioavailability and hence the level of uptake of such micronutrients (Tako and Glahn 2010; Tako et al. 2014; Hart et al. 2015; Petry et al. 2015; Tako et al. 2015a; Tako et al. 2015b; Tako et al. 2016; Glahn et al. 2017). The consideration of anti-nutritionals in biofortification breeding programs is important to ensure that efforts to increase the levels of micronutrients (e.g. iron) in crops are not compromised by inadvertent increases in levels of antinutritionals (such as phytic acid and/or polyphenols) that could arise from breeding efforts (Tako et al. 2014; Hart et al. 2015; Tako et al. 2015a; Glahn et al. 2017; Hart et al. 2017) or from environmental stresses.

Previous studies have shown a negative impact of predicted mid-century elevated CO_2 levels on iron and zinc levels of C3 grain and legume crop plants (Myers et al. 2014), which is anticipated to aggravate the extent of iron deficiency in human diets globally (Smith et al. 2017). In addition to increased CO_2 levels, reduced and erratic rainfall will lead to increases in the incidence and frequency of drought in some regions, which in turn will lead to reductions in crop yields (Long and Ort 2010). For common bean, it is not known whether drought-associated reductions in crop yields will also lead to changes in the nutritional quality of beans under future climate change induced drought scenarios (Beebe et al. 2011).

In this study, we have used Ecocrop to model the impact of climate change induced changes in heat and precipitation by 2050, on the suitability for cultivation of common bean across a range of countries in southeastern Africa. In addition, we have combined the climate impact modeling with experimental field trials of common bean, under the extent of drought anticipated due to future climate change (by 2050), to determine the impact on both yield and the nutritional quality of common bean under climate-induced drought stress. Our results are

important for efforts to climate proof cultivation of the staple crop common bean, so that varieties are developed which under drought stress maintain good yields, contain high levels of the dietary micronutrients iron and zinc, while containing low levels of anti-nutritional factors such as lead and phytic acid.

Methods

EcoCrop Modelling of drought and heat impacts on common bean in East Africa

For a more detailed description of the EcoCrop model the reader is referred to Ramirez-Villegas et al. (Ramirez-Villegas et al. 2013), and Hijmans et al. (Hijmans et al. 2001). For previous assessments climate change impacts using EcoCrop for various crops (including common beans) in Africa the reader is referred to Rippke et al. (Rippke et al. 2016b) and Jarvis et al. (Jarvis et al. 2012).

EcoCrop suitability simulations for common bean were performed for a historical period (1960-1990, chosen to be a representative baseline) and then for the 2050 period (2040–2069) under the Representative Concentrations Pathway 6.0 (RCP 6.0) (Moss et al. 2010). RCP 6.0 was chosen since it is representative of a business as usual scenario. Climate data used for historical suitability simulations was derived from WorldClim (Hijmans et al. 2005), which is a global high resolution database of monthly climatological means for mean, maximum and minimum temperatures and total precipitation. For each future period, simulations were performed using 19 Global Climate Model (GCM) projections, statistically downscaled and bias-corrected (Ramirez-Villegas and Jarvis 2010a). For both historical and future simulations, we assume that the common bean crop is not viable when the overall suitability is below 43% (Rippke et al. 2016b).

EcoCrop is a relatively simple crop suitability model and as such is subject to various limitations, most notably, it does not include estimates of extreme climate impacts, soil fertility, pest and diseases, all of which can be critical to bean growth (Prasad et al. 2002; Araújo et al. 2015) it also does not provide information on crop productivity. Whilst the exclusion of these limitations likely means we underestimate climate change-related challenges for bean production, we argue that they are useful to identify key adaptation priorities for the region.

Plant material

The common bean (*Phaseolus vulgaris*) varieties used in this experiment included released varieties in Malawi, a landrace and also a range of unreleased advanced lines from the CIAT's

bean biofortification breeding program. This consisted of eight BC1F4 Andean Nutrition (NUA) bean lines (NUA 743, NUA 720, NUA 674, NUA 730, NUA 746, NUA 705, NUA 706 and NUA 740) and nine released varieties (A268, A197, CAL 143, A344, NUA 45, SUGAR 131, CAL 113, VTTT 294/4-4, NUA 59, UBR (92)25 and DRK 57) and one landrace (Nasaka).

Field trials

All twenty bean varieties were field trialed under both rainfed conditions and drought-stress conditions that are representative of conditions predicted by Ecocrop modeling to occur by 2050. The drought and rainfed trials of all bean varieties were conducted between 2013-2015 at the Kandiyani site of the International Center for Tropical Agriculture (CIAT) Chitedze Research Station, Lilongwe, Malawi. The crop trial was laid out in a Randomized Complete Block Design (RCBD) with three plots/replicates. Two seasons per year the beans were harvested, in the rainy season, plants were grown under natural, rainfed conditions between December and March. The drought trial was conducted in the winter / off-season which lasts from late August until November.

For each plot, four ridges were prepared, measuring 5m long and spaced at 0.60m. The net plot comprised of the two middle ridges with the outer ridges acting as borders. One seed was planted per planting station spaced at 0.50 cm. No fertilizer was applied in the trials. Pests and diseases were controlled using Dimethoate (Aryzta LifeScience, South Africa) and Karate (Syngenta, Basel, Switzerland) according to product label recommendations. Agronomic data collected from each plot consisted of; days to 50% flowering; days to 95% physiological maturity; plant height at harvesting; grain weight; total plot yield. Yield was determined by weight of grains.

Sampling of bean pods and grains for nutritional analysis and sample preparation

Bean samples from four harvests in the period 2013-2015 were analyzed. Bean pods were sampled when mature, from every plot 10 plants were randomly selected and of each of these plants 10 pods were taken (on different sides of the plant and not touching the soil). All procedures were followed according to HarvestPlus protocols for crop sampling to ensure samples were representative and minimize any risk of sample contamination (Stangoulis and Sison 2008).

After harvest, all samples were kept in a storage room in CIAT, Malawi at room temperature until transfer to the National University of Ireland Galway (NUI Galway). On arrival in NUI Galway, beans were removed from pods and stored at -20 °C in plastic sample

bags until further processing. Bean samples were then freeze-dried and a subsample of 20 beans (i.e. grains) were milled to a fine powder using a coffee grinder (Delonghi, KG49). Samples were vacuum-packed until ICP-MS analysis in Ionomics Lab, University of Aberdeen, Aberdeen, Scotland.

ICP-MS analysis of elemental composition of common bean grains

Briefly, bean powder of each sample was analyzed in duplicate with inductively coupled plasma-mass spectrometry (ICP-MS) as follows: trace metal grade nitric acid was added to the samples. Samples were then spiked with an internal standard plus hydrogen peroxide and left overnight to pre-digest. Following exposure to 115 °C for 4 hours, digested samples were diluted with Milli-Q water. Aliquots were transferred to a 96-well plate for analysis. A detailed description of ICP-MS analysis is provided in the Supplementary Information (available online doi: 10.1038/s41598-018-33952-4).

Protein extraction and measurement

Bean proteins were extracted in three separate stages (Jimenez et al. 2008; Oliveira et al. 2017), namely aqueous extraction (albumins) followed by saline extraction (globulins), followed by cell-lysis extraction (other, membrane-bound proteins). Microcentrifuge tubes (2 ml) were prepared with ten 1 mm glass beads (Sigma-Aldrich) inside. For aqueous extraction bean powder was suspended in distilled water (50 mg in 400 µl; 1:8), shaken in a tissue lyser (QIAGEN) at maximum speed for 30 seconds, followed by incubation at 4 °C for 30 min (samples were shaken every 10 minutes). The sample was centrifuged at 14000 rpm for 15 min. The supernatant was transferred to new tubes. Next, the water-insoluble fraction was extracted from the sediment. The sample was treated as before, but 0.5 M NaCl solution was used instead of distilled water. The supernatant was transferred to new tubes. This should result in 75% of total protein extraction. To release all membrane-bound proteins from the sediment, the sample was treated as before but Lysis buffer (50mM citric acid, pH 3.0, 1 M NaCl, 2% SDS, 0.5% Triton-X-100) was used instead of NaCl solution or distilled water. The supernatant was transferred to new tubes; this is expected to result in ≥ 91% of total protein extraction.

Samples were measured using the BCA protein assay (Thermo Fisher Scientific) on 96-well plates in triplicate. Readings were taken at 560 nm wavelength (Modulus Microplate Mulitmode Reader, Turner Biosystems). The three fractions were measured separately and total protein content is the sum of all three fractions.

Phytic acid analysis of common bean grains

Samples for phytic acid analysis were prepared using modified protocols of Harland et al. (Harland and Oberleas 1977) and Ellis et al. (Ellis and Morris 1986). We used the same milled samples as those for ICP-MS. In brief, 50 mg aliquots of powder were thoroughly mixed with 2 ml 2.4% HCl, incubated at room temperature (RT) for 1 hour followed by 3 min centrifugation at 13000 rpm (Thermo Scientific, Fresco17). 1.8 ml of supernatant was transferred into new 14 ml tubes and diluted with 8.2 ml Milli-Q water. To remove the inorganic phosphate, 10ml diluted samples were applied to the 10 ml Poly-Prep Chromatography Columns (#7311550, Bio-Rad), pre-packed with 0.3 g of AG1-X4 resin, 100-200 mesh (AG 1-X4 Resin #1401341, Bio-Rad) and equilibrated with 0.7 M NaCl, followed by elution with 10 ml 0.1 M NaCl. Phytate was eluted from the columns with 0.7 M NaCl into 15 ml Falcon tubes. Each sample was analyzed in triplicate to measure phytate with WADE reagent (0.03% FeCl3*6H2O and 0.3% sulfosalicylic acid in Milli-Q water, according to Latta et al. (Latta and Eskin 1980).

Data preparation and statistical analysis

All field trial data points, yield data and nutrient data were organized on Microsoft Excel (Microsoft, WA, USA) and all statistical analysis was performed on SPSS Statistics (version 21) (IBM, NY, USA). A detailed output of all statistical tests employed in this experiment are outlined in Supplementary Results File (available online doi: 10.1038/s41598-018-33952-4).

Results

The majority of current common bean growing areas in southeastern Africa will become unsuitable for bean cultivation by the year 2050 due to changes in temperature and precipitation

To determine climate change impacts on common bean (*Phaseolus vulgaris*) suitability for cultivation in southeastern Africa**,** the EcoCrop model was used to produce spatially-explicit simulations of potential climatic suitability for five countries in south-eastern Africa. For each spatial unit (i.e. grid cell), EcoCrop performs separate calculations for temperature-limited (heat and cold) and precipitation-limited (waterlogging and drought) suitability, and then calculates an overall suitability for the crop (Hijmans et al. 2001; Ramirez-Villegas et al. 2013).

Fig. 1 shows the spatial distribution of historical climatic suitability for common bean cultivation (Fig. 1A), as well as the spatial distribution of projected climate change impacts on bean cultivation by 2050 (Fig. 1B). The unshaded areas within the countries analysed (Malawi, Mozambique, Tanzania, Zambia and Zimbabwe) are considered unsuitable for bean cultivation (Fig. 1A). Currently, suitable areas for bean cultivation extend across most of Zambia, Zimbabwe, Tanzania, western Malawi, and northern Mozambique (Fig 1A). Our simulated historical climatic suitability for common bean cultivation agrees well with the observed distribution of bean cultivation in Africa (You et al. 2009). Our EcoCrop modelling of future climate impacts on bean cultivation suitability indicates that a significant proportion of the currently suitable areas will become unsuitable for common bean cultivation by 2050 (red areas in Fig. 1B), particularly in southern Zambia, eastern Zimbabwe, and central Tanzania. In these areas, unless appropriate climate adaptation actions (e.g. climate smart agriculture (CSA) breeding and agronomy options such as new varieties or irrigation) are put in place, it will no longer be possible to grow common beans.

Common bean cultivation suitability differs across locations within each country dependent on changes in temperature or precipitation

Our results further indicate that a reduction in the temperature-related suitability resulting from increased heat stress by 2050 is predicted as the main cause ($\Delta T < \Delta P$) for the overall reduction in climatic suitability of bean growing areas in north western Tanzania, southern Zambia, and western Zimbabwe (dark red shading in Fig 1C). Conversely, reductions in precipitation-related suitability by 2050 were found to be the major cause ($\Delta P < \Delta T$) of climatic suitability change reductions in western Malawi, northern Zambia, eastern Tanzania, and southern

Zimbabwe (orange shading in Fig. 1C). Fig. 1D demonstrates where future bean cultivation suitability by 2050 will be predominantly limited by temperature (red), precipitation (blue) or both temperature and precipitation (yellow).

Reductions in yields of common bean varieties at a climate analogue field site that is representative of predicted drought conditions by 2050

To determine the impacts of future drought scenarios by 2050 on the yields of common bean, we conducted a field trial in western Malawi of 20 common bean varieties over two growing seasons. The bean varieties used consisted of varieties commonly cultivated by smallholder farmers and also bean lines developed by the CIAT and PABRA bean breeding programs under consideration for entry into the national varietal registration and official release process. They comprise a range of different market classes (i.e. quality types) including black, red, kidney, mottled red, brown, light brown and white beans, both smaller and larger sized seed varieties. These varieties were chosen to reflect the reality of (a) what bean varieties are actually available to smallholder farmers in Malawi and (b) what bean varieties are in the regulatory pipeline in Malawi that may become accessible to smallholders within the next decade, taking account of the time-lag expected for national varietal registration, certification processes and bulking up of seed by suppliers. The suitability of the field trial site area in western Malawi for common bean cultivation is projected to fall below current suitability levels by 2050, primarily due to decreased precipitation (Fig. 1B-C).

All twenty bean varieties were field trialed under both rainfed conditions and drought-stress conditions that are representative of conditions predicted by EcoCrop modeling to occur by 2050 (Fig. S1). This corresponded to four field trial seasons in total, i.e. two trials under rainfed conditions and two under drought-stress conditions. To ensure seasons could be reliably and reproducibly grouped according to "rainfed" and "drought-stress" conditions, *k*-means cluster analysis was performed on recorded weather data for the field trial site. The identified clusters were analysed using discriminant analysis for the goodness of fit of the model, and t-statistics were applied to verify the predicted *vs*. found mean of the clusters (Supplementary Results File, available online doi: 10.1038/s41598-018-33952-4). In addition, to further verify the climate change relevance of our drought-stress growing seasons, Simulated Weather Data for Crop Modelling and Risk Assessment (MarkSim) software V.2 was employed (http://www.ccafs-climate.org/pattern_scaling/). These analyses confirmed that our drought-stress field trial conditions can be considered as a climate analogue site for bean growing seasonal weather conditions in the year 2095 in Malawi, under RCP 8.5. Average temperatures

during our drought field trials were 22 °C with extremes of 35 °C, closely resembling the predicted average temperatures for the bean growing season (March, April and May) of 24 °C and extremes of 32 °C in the year 2095. Likewise, the average rainfall during the bean grain filling period in 2015 was 54 mm, similar to the projected average rainfall of 58 mm (Fig. S2).

To determine whether the reduction in yield of the common bean varieties under drought-stress conditions at the climate analogue trial site was more influenced by genotype (variety) or weather conditions (temperature, precipitation), an F-test was performed. This revealed that the variation in yield observed at the trial site is influenced more by weather conditions than genotype (Supplementary Results File, available online doi: 10.1038/s41598-018-33952-4). To determine whether there are any statistically significant differences between the varieties under drought stress conditions, a one-way ANOVA was performed, which showed no significant yield difference between the bean genotypes (Supplementary Results File, available online doi: 10.1038/s41598-018-33952-4). Similarly, no significant yield differences were identified between the bean genotypes under rainfed conditions (Supplementary Results File, available online doi: 10.1038/s41598-018-33952-4). Overall, for all bean varieties that were field trialed at the Malawi climate analogue site, the common bean grain yield decreased by an average of 43% under drought-stress conditions, which is significantly lower (two-tailed independent samples t-test, $P < 0.001$) when compared with rainfed conditions (Fig. S3).

While iron levels in bean grains decrease, under climate-scenario relevant drought stress conditions, zinc, lead, protein and phytic acid levels increase

Many of the countries analysed in this study have moderate to high child underweight rates (Fig. 1E) (Center for International Earth Science Information Network - CIESIN - Columbia University 2005). Micronutrient deficiencies are major contributors to the global problem of maternal and child malnutrition, which causes underweight, stunting and wasting conditions in afflicted children (Black et al. 2008; Bailey et al. 2015). To determine the changes in levels of dietary micronutrients between rainfed and drought-stress conditions of each common bean variety, we used inductively coupled plasma-mass spectrometry (ICP-MS) to measure the relative concentrations of twenty-two elements (B, Na, Mg, P, S, K, Ca, Ti, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Rb, Sr, Mo, Cd and Pb), which included the important dietary micronutrients iron and zinc, and also antinutritional compounds such as lead (Table S1; Fig. S4). For each common bean variety under rainfed and drought conditions, we also determined

the protein levels and the levels of phytic acid, a major antinutritional limiting micronutrient uptake from human diets (Petry et al. 2014) (Table S1; Fig. S4).

As seen for yield, the variations in iron, zinc, lead, protein and phytic acid under drought-stress conditions at the field site are more influenced by weather conditions than genotype (variety) (Supplementary Results File, available online doi: 10.1038/s41598-018-33952-4). To further investigate the lack of variation among genotypes for iron, zinc, lead, protein and phytic acid, a one-way ANOVA was performed. When we compared across genotypes under rainfed conditions we found no significant difference between genotypes for iron, zinc, lead, protein and phytic acid content. Similarly, when we compared across genotypes under drought conditions we found no significant difference between genotypes for iron, lead, protein and phytic acid content. Under drought-stress conditions, one variety shows significantly higher ($P < 0.05$) zinc levels than the other nineteen varieties, but at the 5% significance level such a result would be expected by chance (Supplementary Results File, available online doi: 10.1038/s41598-018-33952-4). We also compared our results for concentrations of Zn and Fe for the twenty varieties tested under rainfed or drought conditions with those recorded for the 1000 accessions in the CIAT cultivated common bean core collection (Fig. S5) and find them to be representative of the range of Fe levels, and representative (or slightly enriched) for Zn. The responses of Fe and Zn in our tested varieties are therefore likely to be similar to those of other common bean genotypes.

Overall, across all of the twenty bean varieties analyzed, the levels of iron are significantly reduced under drought stress (two-tailed independent samples t-test, $P < 0.05$), from an average concentration of 59 ppm in beans from rainfed plants to 54 ppm from drought-stressed plants (Fig 2A). Conversely, the average concentration of zinc significantly increases under drought stress (two-tailed independent samples t-test, $P < 0.05$), from 35 ppm in beans from rainfed plants to 39 ppm from drought-stressed plants, (Fig. 2B). We identify a significant increase (two-tailed independent samples t-test, $P < 0.001$) in lead levels in the bean grains under drought-stress conditions, from an average concentration of 0.05 ppm to 0.22 ppm, representing a fourfold increase (Fig. 2C). This exceeds the maximum level permissible (0.1 ppm) for pulses as defined by the Codex Alimentarius[22]. The average total protein concentration also significantly increases under drought stress (two-tailed independent samples t-test, $P < 0.001$) from 326 mg/g of dry weight in beans from rainfed plants to 371 mg/g of dry weight in beans from drought-stressed plants (Fig. 2D). Similarly, there was a significant increase (two-tailed independent samples t-test, $P < 0.001$) in phytic acid levels under drought

stress, from an average level of 0.96% under rainfed conditions to 1.16% under drought stress (Fig. 2E).

Changes in precipitation and temperature correlate with changes in yield of common bean, and also with iron, lead and protein levels, but not zinc and phytic acid levels

To further investigate the link between weather conditions during the grain filling period and common bean growth, multiple linear regression analysis was performed. As the differences in rainfall between the rainfed seasons and the drought-stress seasons, were also accompanied by differences in air temperature (Fig. S1), the impact of temperature was included as well as rainfall. Over 50% of the variation in yield and iron could be explained by changes in temperature and rainfall between flowering date and harvest. Approx. 40% of the variation in lead could be explained by changes in temperature that occurred between flowering date and harvest, while rainfall that occurred during this time window did not improve the model. Approx. 30% of the variation in protein levels could be explained by changes in temperature and rainfall between flowering date and harvest. Finally, weather conditions during the grain filling period only explained a small part of the variation in zinc (4.0%) or phytic acid (17%), suggesting that other factors not included in the model influence zinc and phytic acid levels in the grain (Supplementary Results File, available online doi: 10.1038/s41598-018-33952-4).

Under climate-change induced drought scenarios, future bean servings will have lower nutritional quality

While yield is typically measured in kg/hectare, the nutritional yield can be considered as the quantity of supply of nutritionally-important compounds per unit area (Cassidy et al. 2013). To determine the supply of nutritional and anti-nutritional compounds on a per meal basis, we calculated the quantity of each dietary compound that each common bean variety would deliver per serving (50 g) under present day (rainfed) and future climate (drought-stress) scenarios (Fig. 3). This highlights that while future bean servings under climate change may become more zinc-rich, they will contain less iron and more undesirable anti-nutritionals (lead and phytic acid). Our results indicate that some varieties (NUA 59, NUA 674) may display promise for drought-proofing dietary supplies from beans under predicted future climate-change scenarios.

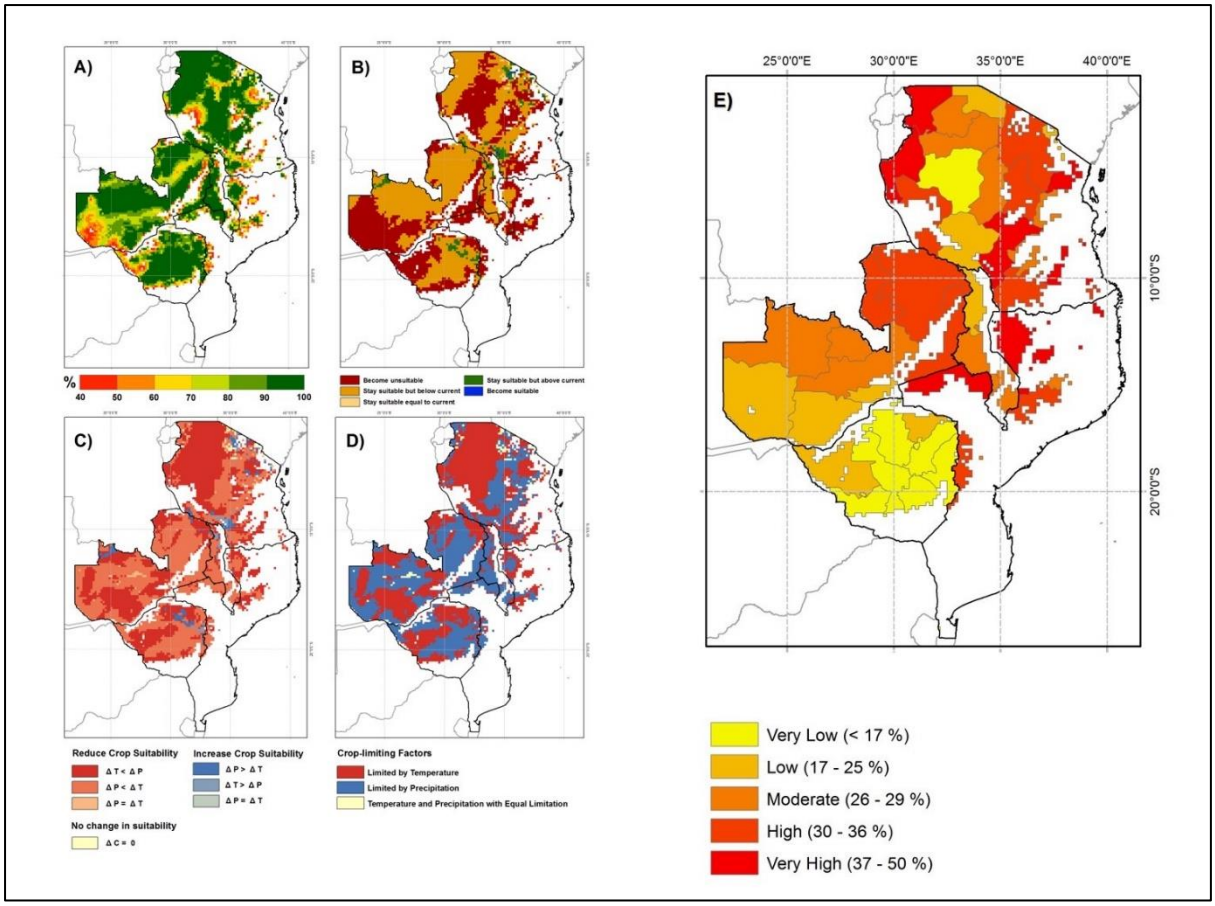


Figure 1. Historical and future (2050) common bean suitability simulations for south-east Africa and current percentage of children underweight; (A) Suitability of currently cultivated common bean for historical climate; (B) Projected impact of climate change by 2050s; (C) Driving factor of change in future climatic suitability; (D) factor most limiting to bean cultivation suitability by 2050; (E) Percentage of children, under the age of 5, who are underweight (data from CIESIN), for the period 1990–2002. Red (reduced suitability), blue (increased suitability), and beige (no change in suitability) colours are used in Fig. 1C to separate directions of change. In red areas, shades of red are used to differentiate areas where suitability reductions are due to temperature changes ($\Delta T < \Delta P$), from those where suitability reductions are due to precipitation changes ($\Delta P < \Delta T$) or where temperature- and precipitation-suitability reductions are equal ($\Delta P = \Delta T$). In blue areas, shades of blue are used to differentiate areas where suitability increases are primarily driven by precipitation ($\Delta P > \Delta T$), temperature ($\Delta T > \Delta P$), or equally driven by both ($\Delta P = \Delta T$).

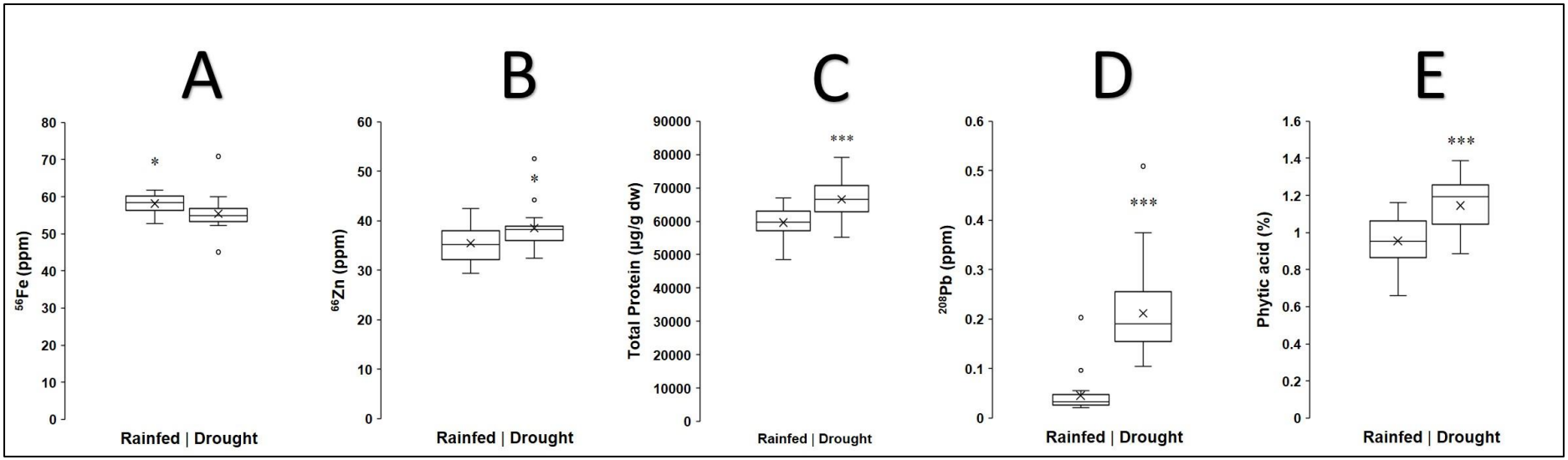


Figure 2. Box and whisker plot showing average grain iron (A), zinc (B), protein (C), lead (D), and phytic acid (E) levels of 20 common bean varieties grown under rainfed and drought-stress conditions. 'X' indicates mean value. * $P < 0.05$ *** $P < 0.001$

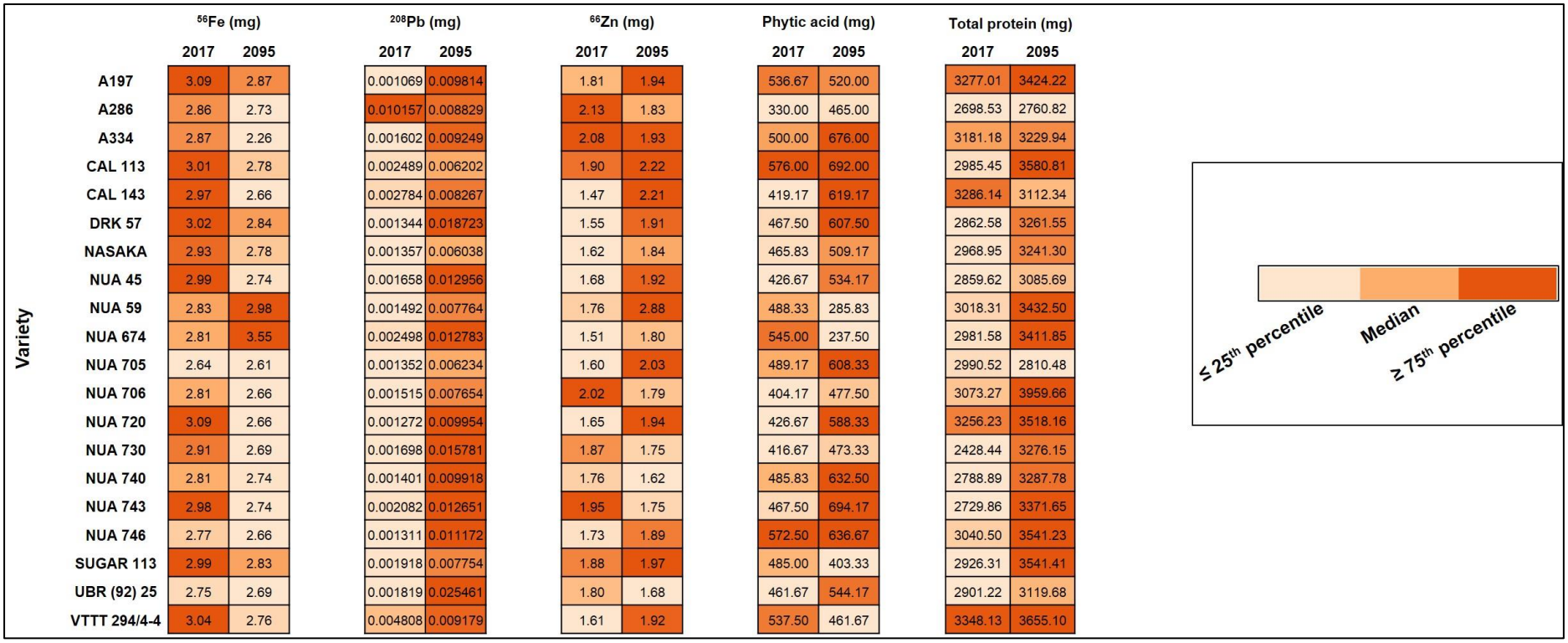


Figure 3. Heat map of nutritional quantity from one serving of beans harvested from present-day (2017) and predicted future (2095) conditions. Concentration of nutrients (^{56}Fe, ^{66}Zn, total protein) and anti-nutritionals (^{208}Pb, phytic acid) in a 50g serving of 20 common bean varieties under rainfed and drought conditions was calculated. Median, lower and upper quartile values were calculated for ^{56}Fe, ^{66}Zn, protein, ^{208}Pb and phytic acid separately.

Discussion

Climate change represents a threat to food security, particularly resulting from ongoing and anticipated negative impacts on agricultural productivity (yields/hectare) (Lobell and Field 2007b; Lobell et al. 2008; Kang et al. 2009; Challinor et al. 2014; Rosenzweig et al. 2014) While there have been a range of studies which indicate negative impacts on yields of major staple crops (Schlenker and Lobell 2010; Zhao et al. 2016; Zhao et al. 2017), there have been fewer studies that have investigated the impact of climate change stresses (e.g. rising CO_2, heat, drought) on staple crop grain quality parameters (Erda et al. 2005; Högy and Fangmeier 2008; DaMatta et al. 2010; Wang and Frei 2011; Myers et al. 2014; Zhu et al. 2018).

Common beans are the most important grain legume supporting food security and human nutrition globally, responsible for almost 15% of daily calories and 36% of daily protein in some countries in Africa and the Americas (Schmutz et al. 2014). In sub-Saharan Africa, common beans are an important staple crop for smallholder farmers and a key nutritional component in diets of poor rural communities. Smallholder farmers in sub-Saharan Africa plant a wide range of bean varieties and landraces that they access from multiple sources, including via purchasing from formal (e.g. government distributors, commercial seed companies, agro-dealers, NGO/UN) or informal (e.g. local markets, own seed stocks, neighbor) sources (Wortmann 1998; Chirwa and Aggarwal 2000; McGuire and Sperling 2016). Such smallholder farming communities are extremely vulnerable to negative impacts of climate change on their livelihoods and nutritional status, including through reductions in yields and/or nutritional quality of staple crops they both consume and trade (Challinor et al. 2007; Lloyd et al. 2011).

Climate change can change weather patterns, resulting in altered temperature and rainfall effects in different regions, which can have concomitant impacts on the suitability of crops for continued cultivation in climate-change impacted regions (Ray et al. 2015). In particular, elevated temperatures (heat) and reduced rainfall (drought) can reduce crop yields (Wheeler et al. 2000; Peng et al. 2004; Lobell et al. 2008; Schlenker and Roberts 2009). Our EcoCrop climatic suitability analyses for common bean in South Eastern Africa to 2050 indicates that predicted increases in temperature and reductions in rainfall (precipitation) will result in common bean becoming unsuitable for cultivation across the vast majority of current bean growing regions (Fig 1D). Only in specific localised zones in northern Zimbabwe, southern and northern Tanzania, and northern Malawi, are increases in climatic suitability for bean cultivation projected. Relocation of bean cultivation to different areas beyond the current range of cultivation may be possible, but in this case careful consideration should be given to

choosing varieties suited for the new area, including any change in photoperiod that occurs with changing latitude. Overall, our findings are consistent with those of previous studies where different models to EcoCrop have been used (Thornton et al. 2009; Ramirez-Villegas and Thornton 2015; Ramirez-Cabral et al. 2016). Future climate conditions will be associated with elevated atmospheric CO_2 concentrations, which may either exacerbate or alleviate the effects of increased temperature and reduced precipitation on common bean growing regions. While the EcoCrop model cannot process atmospheric CO_2 effects, it is noteworthy that C3 plants (of which common bean is one) have been shown experimentally to respond to drought by reducing photosynthesis, an effect which is not removed upon doubling CO_2 treatment (Hamim 2005). We conclude that, in the absence of implementation of significant adaptation strategies to maintain yields of common bean in southeastern Africa, it can be expected that yields of common bean will dramatically decline across the region in the period to 2050. The adaptation strategies that will be necessary to implement at scale may be incremental (e.g. breeding new bean varieties or using agronomic practices such as irrigation) or transformational (e.g. involving changing to a different protein or high-value crop species, or finding an alternative livelihood that is more climate-resilient) (Rickards and Howden 2012).

In addition to yield reductions that will negatively impact on livelihoods, there is potential for climate stresses to also impact on the nutritional quality of crops. Such climate effects may frustrate biofortification efforts to breed new biofortified varieties of staple crops that have elevated levels of essential micronutrients (Díaz-Gómez et al. 2017). Screening of over 1000 genotypes of common bean germplasm from the CIAT (Beebe et al. 2000) genebank revealed an average Fe concentration of 55 ppm, within a concentration range of 34 and 89 ppm. The average Zn concentration was 35 ppm, within a concentration range of 21 and 54ppm. Notably, all 20 genotypes investigated in our study have similar Fe and Zn concentrations. The average Fe concentration among genotypes used in our study (i.e. 59ppm) is statistically similar to the average observed across the primary genepool (i.e. 55 ppm). Likewise the average Zn concentration among the genotypes in our study (i.e. 35 ppm) is identical to the average across the genepool (i.e.35 ppm) (Fig. 2A & B, Fig. S5, Supplementary Results File, available online doi: 10.1038/s41598-018-33952-4).

While there has been a previous attempt to determine the effect of water-limiting conditions on Fe and Zn levels in common bean, using a limited number of genotypes under irrigated conditions, further studies are required to determine impacts on nutritional availability under drought in key bean growing regions (Wortmann 1998; Chirwa and Aggarwal 2000;

Schmutz et al. 2014; McGuire and Sperling 2016). In our study we have modeled the negative impacts of climate change on common bean production in southeastern Africa, which has revealed that reduced precipitation by 2050 will be the main limiting factor for bean cultivation. Furthermore, we conducted a multi-year field trial at a climate analogue site which experiences weather conditions similar to that predicted for Malawi in the year 2095. Our results are the first to demonstrate that the level of a key nutrient (i.e. iron) in common beans under climate change induced drought stress will significantly decline.

Studies conducted to date on iron bioavailability from high-Fe biofortified beans using Caco-2 cell models are congruent with findings from poultry models (Tako and Glahn 2010; Tako et al. 2011), which in turn are consistent with human feeding trials which have shown positive nutritional impacts from consumption of high-Fe biofortified beans (De Moura et al. 2014; Petry et al. 2014; Haas et al. 2016; Finkelstein et al. 2017; Glahn et al. 2017). However, antinutritional compounds in staple crop plants can negatively affect the uptake (bioavailability) of nutritional compounds (e.g. iron, zinc) (Siegenberg et al. 1991). Such antinutritional compounds include phytic acid and polyphenols, which have been shown to negatively affect the bioavailability and uptake of iron and zinc from common beans, using *in vitro* (Caco-2 cell) and animal (poultry) models (Tako and Glahn 2010; Tako et al. 2011; Petry et al. 2014; Tako et al. 2014; Hart et al. 2015; Tako et al. 2015a; Glahn et al. 2017; Hart et al. 2017). In addition, the composition of targeted diets or meal plans can affect the extent of iron uptake from high-Fe biofortified beans. For instance, some foods commonly consumed with beans (e.g. rice) can inhibit Fe bioavailability while others (e.g potato) can increase Fe bioavailability when eaten with beans (Glahn et al. 2017).

In this study we have focussed on the effect of phytic acid because of its major influence on iron bioavailability, especially in the case of consuming beans in a composite meal (Petry et al. 2014). However, we recognise that polyphenols are an additional class of antinutritionals that need to be considered in high-Fe bean biofortification efforts. For instance, studies in black beans have shown that total polyphenols inhibit iron uptake in Caco-2 cell assays (Hart et al. 2015). The overall inhibitory effect of polyphenols is combinatorial, whereby some polyphenols (catechin, 3,4-dihydroxybenzoic acid, kaempferol, and kaempferol 3-glucoside) promote iron uptake while others (myricetin, myricetin 3-glucoside, quercetin, and quercetin 3-glucoside) inhibit iron uptake (Hart et al. 2015). Because of differential potency effects between polyphenols that inhibit or promote iron uptake, it is considered that the majority of the inhibitory polyphenol compounds would need to be removed from

biofortified beats in order to substantially reduce the inhibitory effect on iron uptake (Hart et al. 2017).

There are also a number of heavy metals (e.g. lead, cadmium, arsenic), which can enter the human body via diet, that can act as toxic anti-nutritionals (depending on concentration) (Vázquez et al. 2015). In addition to drought-induced reductions in the levels of iron, our results demonstrate that drought-stressed common bean varieties also display increases in the levels of the antinutritional compounds phytic acid and lead. While we did not measure the levels of polyphenols, polyphenol levels can increase in plants in response to drought-stress (AbdElgawad et al. 2016; Varela et al. 2016), and have been shown to negatively affect iron and zinc bioavailability and uptake from dietary common bean (Tako et al. 2014; Hart et al. 2015).

The underlying physiological basis for the increases of antinutritional compounds such as phytic acid and lead under drought stress are unclear. In tropical soils, such as laterites, it has been shown that sorption values for Pb^{2+} are greater than those of other bivalent cations (Chotpantarat et al. 2011). The development of deeper rooting systems in arid soils could potentially lead to greater contact with Pb^{2+} cations adsorbed within the soil. Alternatively, the observed increase could be a secondary effect of greater investment in active cation uptake when the soluble fraction of nutrient cations is insufficient to meet plant needs, reminiscent of the increased uptake of Al^{3+} cations observed in calcicole plants under acidic conditions (Lee, 1998). However, we cannot exclude other impacts of drought and heat on root physiology (Whitmore and Whalley 2009), and direct analysis of root growth in common bean under different climatic conditions will be required to distinguish these possibilities.

The increases in phytic acid are of particular concern as it is considered the main anti-nutritional compound in legumes. The increased phytic acid accumulation likely relates to its function in limiting oxidative stress in legumes under dryer conditions (Irvine and Schell 2001; Kido et al. 2013). Indeed, phytic acid is known to accumulate in legume seeds (e.g. chickpeas) in response to drought stress (Boominathan et al. 2004; Joshi-Saha and Reddy 2015). It should be noted however that phytic acid in field peas has been found to be reduced under elevated CO_2 levels (Myers et al. 2014). If a similar response occurs in common beans then the increase in phytic acid levels we observe could be counteracted. Growth trials combining changes in climate and CO_2 simultaneously will be needed to assess interactions between possibly competing effects.

Our study indicates that ongoing efforts to develop biofortified bean varieties will need to not only develop heat- and drought-tolerant beans, but will also need to ensure that such varieties also maintain elevated iron and zinc levels, and low levels of antinutritional compounds (e.g. phytic acid, lead and specific inhibitory polyphenols) under drought or other environmental stresses. To avoid unintended consequences, our results highlight that it is critically important that biofortified crop varieties (including under abiotic stresses) do not accumulate anti-nutritionals (e.g. phytic acid, lead, arsenic, polyphenols) (Zhao and McGrath 2009). Overall, our results demonstrate that there will be a reduction in the nutritional quality of a typical bean serving, if the common bean varieties have been cultivated under the levels of drought stress predicted for southeastern Africa to 2050 and beyond.

Conclusions

Both incremental and transformational climate change adaptation (Rickards and Howden 2012; Challinor et al. 2014; Ramirez-Villegas and Thornton 2015; Rippke et al. 2016b) strategies are needed for common bean cropping systems of smallholder farmers in south-eastern Africa, whereby such farmers can have greater access to improved varieties and agronomic practices that allows their cropping systemsm to be more resilient to increased heat and drought conditions, while maintaining or improving nutritional composition of bean grains. Recent plant breeding progress to develop drought- and heat-adapted bean varieties indicates that genetics-based adaptation should be possible (Beebe et al. 2008), which can be a component of an overall portfolio of climate smart agriculture technologies and practices to ensure resilient of common bean cultivation to climate change impacts (Lipper et al. 2014). Where farmers have access to (and widely adopt) such improved bean varieties (McGuire and Sperling 2016), it may be possible to maintain yields in areas where cultivation suitability will be negatively impacted by climate change (Rippke et al. 2016b). As the breeding, testing and dissemination of new bean varieties can take a decade or more, our results highlight the need for accelerated development and seed-system dissemination of heat- and drought-tolerant common bean varieties that can maintain yields while also improving nutritional quality (e.g. via biofortification breeding) under future climate change scenarios.

References

AbdElgawad, H., G. Zinta, G.T. Beemster, I.A. Janssens, and H. Asard, 2016 Future climate CO2 levels mitigate stress impact on plants: increased defense or decreased challenge? *Frontiers in plant science* 7:556.

Anjali, V., 2015 Food fortification: a complementary strategy for improving micronutrient malnutrition (MNM) status. *Food Science Research Journal* 6 (2):381-389.

Araújo, S.S., S. Beebe, M. Crespi, B. Delbreil, E.M. González *et al.*, 2015 Abiotic stress responses in legumes: strategies used to cope with environmental challenges. *Critical Reviews in Plant Sciences* 34 (1-3):237-280.

Bailey, R.L., K.P. West Jr, and R.E. Black, 2015 The epidemiology of global micronutrient deficiencies. *Annals of Nutrition and Metabolism* 66 (Suppl. 2):22-33.

Beebe, S., A.V. Gonzalez, and J. Rengifo, 2000 Research on trace minerals in the common bean. *Food and Nutrition Bulletin* 21 (4):387-391.

Beebe, S., J. Ramirez, A. Jarvis, I.M. Rao, G. Mosquera *et al.*, 2011 Genetic improvement of common beans and the challenges of climate change. *Crop adaptation to climate change*:356-369.

Beebe, S.E., I.M. Rao, C. Cajiao, and M. Grajales, 2008 Selection for drought resistance in common bean also improves yield in phosphorus limited and favorable environments. *Crop Science* 48 (2):582-592.

Black, R.E., L.H. Allen, Z.A. Bhutta, L.E. Caulfield, M. De Onis *et al.*, 2008 Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet* 371 (9608):243-260.

Boominathan, P., R. Shukla, A. Kumar, D. Manna, D. Negi *et al.*, 2004 Long term transcript accumulation during the development of dehydration adaptation in Cicer arietinum. *Plant physiology* 135 (3):1608-1620.

Bouis, H.E., C. Hotz, B. McClafferty, J. Meenakshi, and W.H. Pfeiffer, 2011 Biofortification: a new tool to reduce micronutrient malnutrition. *Food and Nutrition Bulletin* 32 (1_suppl1):S31-S40.

Cassidy, E.S., P.C. West, J.S. Gerber, and J.A. Foley, 2013 Redefining agricultural yields: from tonnes to people nourished per hectare. *Environmental research letters* 8 (3):034015.

Center for International Earth Science Information Network - CIESIN - Columbia University, 2005 Poverty Mapping Project: Global Subnational Prevalence of Child Malnutrition. NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY.

Challinor, A., T. Wheeler, C. Garforth, P. Craufurd, and A. Kassam, 2007 Assessing the vulnerability of food crop systems in Africa to climate change. *Climatic change* 83 (3):381-399.

Challinor, A.J., J. Watson, D. Lobell, S. Howden, D. Smith *et al.*, 2014 A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change* 4 (4):287-291.

Chirwa, R.M., and V.D. Aggarwal, 2000 Bean seed dissemination systems in Malawi: A strategy. *Journal of Sustainable Agriculture* 15 (4):5-24.

Chotpantarat, S., S.K. Ong, C. Sutthirat, and K. Osathaphan, 2011 Effect of pH on transport of Pb2+, Mn2+, Zn2+ and Ni2+ through lateritic soil: column experiments and transport modeling. *Journal of Environmental Sciences* 23 (4):640-648.

DaMatta, F.M., A. Grandis, B.C. Arenque, and M.S. Buckeridge, 2010 Impacts of climate changes on crop physiology and food quality. *Food Research International* 43 (7):1814-1823.

De Moura, F.F., A.C. Palmer, J.L. Finkelstein, J.D. Haas, L.E. Murray-Kolb *et al.*, 2014 Are Biofortified Staple Food Crops Improving Vitamin A and Iron Status in Women and Children? New Evidence from Efficacy Trials–. Oxford University Press.

Díaz-Gómez, J., R.M. Twyman, C. Zhu, G. Farré, J.C. Serrano *et al.*, 2017 Biofortification of crops with nutrients: factors affecting utilization and storage. *Current opinion in biotechnology* 44:115-123.

Ellis, R., and E. Morris, 1986 Appropriate resin selection for rapid phytate analysis by ion-exchange chromatography. *Cereal chemistry* 63:58-59.

Erda, L., X. Wei, J. Hui, X. Yinlong, L. Yue *et al.*, 2005 Climate change impacts on crop yield and quality with CO2 fertilization in China. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360 (1463):2149-2154.

Finkelstein, J.L., J.D. Haas, and S. Mehta, 2017 Iron-biofortified staple food crops for improving iron status: a review of the current evidence. *Current opinion in biotechnology* 44:138-145.

Glahn, R., E. Tako, J. Hart, J. Haas, M. Lung'aho *et al.*, 2017 Iron Bioavailability Studies of the First Generation of Iron-Biofortified Beans Released in Rwanda. *Nutrients* 9 (7):787.

Haas, J.D., S.V. Luna, M.G. Lung'aho, M.J. Wenger, L.E. Murray-Kolb *et al.*, 2016 Consuming Iron Biofortified Beans Increases Iron Status in Rwandan Women after 128 Days in a Randomized Controlled Feeding Trial–3. *The Journal of nutrition* 146 (8):1586-1592.

Hamim, H., 2005 Photosynthesis of C3 and C4 Species in Response to Increased CO2 Concentration and Drought Stress. *HAYATI Journal of Biosciences* 12 (4):131.

Harland, B., and D. Oberleas, 1977 A modified method for phytate analysis using an ion-exchange procedure: application to textured vegetable proteins [Soybeans]. *Cereal chemistry* 54 (4):827-832.

Hart, J.J., E. Tako, and R.P. Glahn, 2017 Characterization of polyphenol effects on inhibition and promotion of iron uptake by Caco-2 cells. *Journal of Agricultural and Food Chemistry* 65 (16):3285-3294.

Hart, J.J., E. Tako, L.V. Kochian, and R.P. Glahn, 2015 Identification of black bean (Phaseolus vulgaris L.) polyphenols that inhibit and promote iron uptake by Caco-2 cells. *Journal of Agricultural and Food Chemistry* 63 (25):5950-5956.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones, and A. Jarvis, 2005 Very high resolution interpolated climate surfaces for global land areas. *International journal of climatology* 25 (15):1965-1978.

Hijmans, R.J., L. Guarino, M. Cruz, and E. Rojas, 2001 Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter*:15-19.

Högy, P., and A. Fangmeier, 2008 Effects of elevated atmospheric CO2 on grain quality of wheat. *Journal of Cereal Science* 48 (3):580-591.

Imdad, A., K. Herzer, E. Mayo‐Wilson, M.Y. Yakoob, and Z.A. Bhutta, 2010 Vitamin A supplementation for preventing morbidity and mortality in children from 6 months to 5 years of age. *The Cochrane Library*.

Irvine, R.F., and M.J. Schell, 2001 Back in the water: the return of the inositol phosphates. *Nature reviews Molecular cell biology* 2 (5):327.

Jarvis, A., J. Ramirez-Villegas, B.V.H. Campo, and C. Navarro-Racines, 2012 Is cassava the answer to African climate change adaptation? *Tropical Plant Biology* 5 (1):9-29.

Jimenez, O., E. Delgado, S. Rosales, S. Reveles, A. Ochoa *et al.*, 2008 Efficiency in seed protein extraction from common bean cultivars grown in Mexican Northern Highlands. *Annual report.*

Joshi-Saha, A., and K.S. Reddy, 2015 Repeat length variation in the 5′UTR of myo-inositol monophosphatase gene is related to phytic acid content and contributes to drought tolerance in chickpea (Cicer arietinum L.). *Journal of Experimental Botany* 66 (19):5683-5690.

Kang, Y., S. Khan, and X. Ma, 2009 Climate change impacts on crop yield, crop water productivity and food security–A review. *Progress in Natural Science* 19 (12):1665-1674.

Kido, E.A., J.R.F. Neto, R.L. Silva, L.C. Belarmino, J.P.B. Neto *et al.*, 2013 Expression dynamics and genome distribution of osmoprotectants in soybean: identifying important components to face abiotic stress. *BMC bioinformatics* 14 (1):S7.

Latta, M., and M. Eskin, 1980 A simple and rapid colorimetric method for phytate determination. *Journal of Agricultural and Food Chemistry* 28 (6):1313-1315.

Lipper, L., P. Thornton, B.M. Campbell, T. Baedeker, A. Braimoh *et al.*, 2014 Climate-smart agriculture for food security. *Nature Climate Change* 4 (12):1068.

Lloyd, S.J., R.S. Kovats, and Z. Chalabi, 2011 Climate change, crop yields, and undernutrition: development of a model to quantify the impact of climate scenarios on child undernutrition. *Environmental health perspectives* 119 (12):1817.

Lobell, D.B., M.B. Burke, C. Tebaldi, M.D. Mastrandrea, W.P. Falcon *et al.*, 2008 Prioritizing climate change adaptation needs for food security in 2030. *Science* 319 (5863):607-610.

Lobell, D.B., and C.B. Field, 2007 Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental research letters* 2 (1):014002.

Long, S.P., and D.R. Ort, 2010 More than taking the heat: crops and global change. *Curr Opin Plant Biol* 13 (3):241-248.

McGuire, S., and L. Sperling, 2016 Seed systems smallholder farmers use. *Food Security* 8 (1):179-195.

Moss, R.H., J.A. Edmonds, K.A. Hibbard, M.R. Manning, S.K. Rose *et al.*, 2010 The next generation of scenarios for climate change research and assessment. *Nature* 463 (7282):747.

Myers, S.S., A. Zanobetti, I. Kloog, P. Huybers, A.D. Leakey *et al.*, 2014 Increasing CO2 threatens human nutrition. *Nature* 510 (7503):139-142.

Oliveira, A.P., G.F. Andrade, B.S.O. Mateó, and J. Naozuka, 2017 Protein and Metalloprotein Distribution in Different Varieties of Beans (Phaseolus vulgaris L.): Effects of Cooking. *International journal of food science* 2017.

Peng, S., J. Huang, J.E. Sheehy, R.C. Laza, R.M. Visperas *et al.*, 2004 Rice yields decline with higher night temperature from global warming. *Proc Natl Acad Sci U S A* 101 (27):9971-9975.

Petry, N., E. Boy, J.P. Wirth, and R.F. Hurrell, 2015 Review: The potential of the common bean (Phaseolus vulgaris) as a vehicle for iron biofortification. *Nutrients* 7 (2):1144-1173.

Petry, N., I. Egli, J.B. Gahutu, P.L. Tugirimana, E. Boy *et al.*, 2014 Phytic acid concentration influences iron bioavailability from biofortified beans in Rwandese women with low iron status. *The Journal of nutrition* 144 (11):1681-1687.

Prasad, P., K.J. Boote, L.H. Allen, and J.M. Thomas, 2002 Effects of elevated temperature and carbon dioxide on seed‐set and yield of kidney bean (Phaseolus vulgaris L.). *Global Change Biology* 8 (8):710-721.

Ramirez-Cabral, N.Y.Z., L. Kumar, and S. Taylor, 2016 Crop niche modeling projects major shifts in common bean growing areas. *Agricultural and Forest Meteorology* 218:102-113.

Ramirez-Villegas, J., and A. Jarvis, 2010 Downscaling global circulation model outputs: the delta method decision and policy analysis Working Paper No. 1. *Policy Analysis* 1:1-18.

Ramirez-Villegas, J., A. Jarvis, and P. Läderach, 2013 Empirical approaches for assessing impacts of climate change on agriculture: The EcoCrop model and a case study with grain sorghum. *Agricultural and Forest Meteorology* 170:67-78.

Ramirez-Villegas, J., and P.K. Thornton, 2015 Climate change impacts on African crop production in *CCAFS Working Paper*. CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), Copenhagen, Denmark.

Ray, D.K., J.S. Gerber, G.K. MacDonald, and P.C. West, 2015 Climate variation explains a third of global crop yield variability. *Nature communications* 6.

Rickards, L., and S. Howden, 2012 Transformational adaptation: agriculture and climate change. *Crop and Pasture Science* 63 (3):240-250.

Rippke, U., J. Ramirez-Villegas, A. Jarvis, S.J. Vermeulen, L. Parker *et al.*, 2016 Timescales of transformational climate change adaptation in sub-Saharan African agriculture. *Nature Climate Change* 6 (6):605-609.

Rosenzweig, C., J. Elliott, D. Deryng, A.C. Ruane, C. Müller *et al.*, 2014 Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences* 111 (9):3268-3273.

Saltzman, A., E. Birol, H.E. Bouis, E. Boy, F.F. De Moura *et al.*, 2013 Biofortification: progress toward a more nourishing future. *Global Food Security* 2 (1):9-17.

Sanahuja, G., G. Farré, J. Berman, U. Zorrilla-López, R.M. Twyman *et al.*, 2013 A question of balance: achieving appropriate nutrient levels in biofortified staple crops. *Nutrition research reviews* 26 (2):235-245.

Schlenker, W., and D.B. Lobell, 2010 Robust negative impacts of climate change on African agriculture. *Environmental research letters* 5 (1):014010.

Schlenker, W., and M.J. Roberts, 2009 Nonlinear temperature effects indicate severe damages to US crop yields under climate change. *Proceedings of the National Academy of Sciences* 106 (37):15594-15598.

Schmutz, J., P.E. McClean, S. Mamidi, G.A. Wu, S.B. Cannon *et al.*, 2014 A reference genome for common bean and genome-wide analysis of dual domestications. *Nature genetics* 46 (7):707-713.

Siegenberg, D., R.D. Baynes, T.H. Bothwell, B.J. Macfarlane, R.D. Lamparelli *et al.*, 1991 Ascorbic acid prevents the dose-dependent inhibitory effects of polyphenols and phytates on nonheme-iron absorption. *The American journal of clinical nutrition* 53 (2):537-541.

Smith, M., C. Golden, and S. Myers, 2017 Anthropogenic carbon dioxide emissions may increase the risk of global iron deficiency. *GeoHealth*.

Stangoulis, J., and C. Sison, 2008 Crop sampling protocols for micronutrient analysis. *Harvest Plus Tech Monogr Ser* 7.

Tako, E., H. Bar, and R.P. Glahn, 2016 The combined application of the Caco-2 cell bioassay coupled with in vivo (Gallus gallus) feeding trial represents an effective approach to predicting Fe bioavailability in humans. *Nutrients* 8 (11):732.

Tako, E., S.E. Beebe, S. Reed, J.J. Hart, and R.P. Glahn, 2014 Polyphenolic compounds appear to limit the nutritional benefit of biofortified higher iron black bean (Phaseolus vulgaris L.). *Nutrition journal* 13 (1):28.

Tako, E., M.W. Blair, and R.P. Glahn, 2011 Biofortified red mottled beans (Phaseolus vulgaris L.) in a maize and bean diet provide more bioavailable iron than standard red mottled beans: Studies in poultry (Gallus gallus) and an in vitro digestion/Caco-2 model. *Nutrition journal* 10 (1):113.

Tako, E., and R.P. Glahn, 2010 White beans provide more bioavailable iron than red beans: studies in poultry (Gallus gallus) and an in vitro digestion/Caco-2 model. *International Journal for Vitamin and Nutrition Research* 80 (6):416.

Tako, E., S. Reed, A. Anandaraman, S.E. Beebe, J.J. Hart *et al.*, 2015a Studies of cream seeded carioca beans (Phaseolus vulgaris L.) from a Rwandan efficacy trial: in vitro and in vivo screening tools reflect human studies and predict beneficial results from iron biofortified beans. *PloS one* 10 (9):e0138479.

Tako, E., S.M. Reed, J. Budiman, J.J. Hart, and R.P. Glahn, 2015b Higher iron pearl millet (Pennisetum glaucum L.) provides more absorbable iron that is limited by increased polyphenolic content. *Nutrition journal* 14 (1):11.

Thornton, P.K., P.G. Jones, G. Alagarswamy, and J. Andresen, 2009 Spatial variation of crop yield response to climate change in East Africa. *Global Environmental Change* 19 (1):54-65.

Varela, M.C., I. Arslan, M.A. Reginato, A.M. Cenzano, and M.V. Luna, 2016 Phenolic compounds as indicators of drought resistance in shrubs from Patagonian shrublands (Argentina). *Plant Physiology and Biochemistry* 104:81-91.

Vázquez, M., M. Calatayud, C.J. Piedra, G. Chiocchetti, D. Vélez *et al.*, 2015 Toxic trace elements at gastrointestinal level. *Food and Chemical Toxicology* 86:163-175.

Wang, Y., and M. Frei, 2011 Stressed food–The impact of abiotic environmental stresses on crop quality. *Agriculture, Ecosystems & Environment* 141 (3-4):271-286.

Wheeler, T.R., P.Q. Craufurd, R.H. Ellis, J.R. Porter, and P.V. Prasad, 2000 Temperature variability and the yield of annual crops. *Agriculture, Ecosystems & Environment* 82 (1-3):159-167.

Whitmore, A.P., and W.R. Whalley, 2009 Physical effects of soil drying on roots and crop growth. *Journal of Experimental Botany* 60 (10):2845-2857.

Wortmann, C.S., 1998 *Atlas of common bean (Phaseolus vulgaris L.) production in Africa*: CIAT.

You, L., S. Wood, and U. Wood-Sichra, 2009 Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. *Agricultural Systems* 99 (2):126-140.

Zhao, C., B. Liu, S. Piao, X. Wang, D.B. Lobell *et al.*, 2017 Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences* 114 (35):9326-9331.

Zhao, C., S. Piao, X. Wang, Y. Huang, P. Ciais *et al.*, 2016 Plausible rice yield losses under future climate warming. *Nature plants* 3:16202.

Zhao, F.-J., and S.P. McGrath, 2009 Biofortification and phytoremediation. *Current opinion in plant biology* 12 (3):373-380.

Zhu, C., K. Kobayashi, I. Loladze, J. Zhu, Q. Jiang *et al.*, 2018 Carbon dioxide (CO2) levels this century will alter the protein, micronutrients, and vitamin content of rice grains with potential health consequences for the poorest rice-dependent countries. *Science Advances* 4 (5):eaaq1012.