NNT : 2016SACLE024

THESE DE DOCTORAT DE
L'UNIVERSITE PARIS-SACLAY
PREPAREE A
L'UNIVERSITE UNIVERSITE EVRY VAL D'ESSONNE

ÉCOLE DOCTORALE N°577
Structure et dynamique des systèmes vivants
Spécialité de doctorat : Sciences de la Vie et de la Santé

Par

Thi Vinh Ha DINH

Development and characterization of the *Brachypodium* species polyploid model

Thèse présentée et soutenue à Evry, le 27 septembre 2016

Composition du Jury :

Malika AINOUCHE	Professeur, Université de Rennes 1	Rapporteur
Etienne PAUX	Directeur de recherche, INRA Clermont-Ferrand	Rapporteur
Marie DUFRESNE	Maitre de Conférences, Université Paris Sud 11	Président
Boulos CHALHOUB	Directeur de recherche, INRA-UEVE,Evry	Directeur de thèse

Remerciements

En tout premier, je tiens à remercier profondément Dr. Boulos Chalhoub, mon *directeur de thèse, de m'avoir donné la chance et de m'avoir fait confiance pour mener ces* travaux de recherche de thèse. Je le remercie pour toute sa patie*nce et ses conseils qu'il m'a* apportés dès mes premiers jours en France ainsi que tout au long de ma thèse.

Je souhaite ensuite adresser mes remerciements les plus sincères aux membres du jury, Prof. Malika Ainouche, Dr. Etienne Paux, et Dr. Marie Dufresne *qui m'ont fait l'honneur de juger* mon travail de thèse.

Je remercie sincèrement les membres de mon comité de thèse, Dr. Joseph Jahier, Dr. Philippe Lashermes, Dr. Marie-*Angèle Grandbastien, qui m'ont offert des conseils précieux* pour ma thèse.

Je voudrais remercier vivement toutes les personnes qui m'ont apporté leurs aides et qui ont contribué au bon déroulement de cette thèse, aussi bien sur le plan scientifique que humain.

*Je tiens à remercier particulièrement l'équipe OEPG avec qui j'ai des sentim*ents *partagés dès mes premiers jours au laboratoire. Merci à Imen qui m'a donné les premières instructions pour travailler avec les Brachypodium dans la serre. Merci à Harry qui m'a* enseigné les premières techniques pour mes expériences à la paillasse, il a été à côté de moi pour me donner des conseils tout au long de mes années de thèse. Merci à Nathalie qui non *seulement m'a enseigné des connaissances scientifiques mais également des conseils académiques. Merci à Dominique qui m'a aidé pour faire les expériences d'ARN, merci à ses* opinions scientifiques pointus et son temps pour corriger les fautes de mon manuscrit de *thèse. Je remercie à tout cœur Mathieu, Jeremy, Houda pour leur bonne humeur au quotidien* et leurs encouragements.

Un grand spécial merci à *Isabelle Le Clainche qui a été toujours disponible à m'aider* pendant mes années de thèse. Merci Isabelle pour ton aide pour mes expériences, pour *m'avoir soutenir moralement à des moments les plus difficiles. Je tiens un grand merci à Aurélie qui m'a beau*coup aidé dans ma vie quotidienne pour surmonter des difficultés. Je *vous remercie toutes deux d'avoir m'aidé lorsque j'ai porté Chi Mai.*

Un grand merci à Smahane, qui m'a donné ses conseils ainsi que ses soutiens dès mes premiers jours au labo. Merci à S*mahane et Edith qui m'ont aidé pour les analyses des bases* de données RNA-Seq.

Je voudrais remercier Pascal et Florie qui m'ont aidé de s'occuper de mes plantes en serre. Sans vous, mon travail était très difficile.

Je remercie tous les collègues et amis *de l'URGV pour leur bonne humeur au* quotidien, les collègues proches du bâtiment G2: Véro, Jean-Philippe, Marie-Laure, Guillem, Philip, Jean-Luc. Je me souviens pour toujours des repars du midi avec Cécile, Zakia, Rim et Patricia. Je voudrais remercier les collègues Etienne, Stéphanie, Ludivine, Alex, Sandrine, *Mauricio, Jose, Remy… Merci à Jean, Bénédicte, Marie*-Ludivine, Ana, Marie du bâtiment G1 pour vos encourageants quotidiens.

Je tiens à remercier le service administratif *de l'URGV Mélanie, Arnaud,* Sabine, *Louisa et celui de l'Université d'Evry Marité Quintin, Florence Hamon, Véronique Fournie,* Carole Troussier, *merci pour vos aides qui m'a facilité les démarches universitaires pendent* tous mes années de thèse.

Je voudrais remercier également les collaborateurs de Genoscope Karine et Corine *pour les séquençages et les très précieux conseils sur l'analyse des NGS.*

Je remercie mes collègues à Renne : Joseph, Olivier et Virginie pour leur disponibilité, la réflexion et la discussion scientifique et leu*r travail d'analyse cytologique.*

Je tiens à remercier les collègues dans le consortium Brachypodium : Dr. John Vogel, Dr. Sean Gordon, Dr. Robert Hasterok, Dr. Pilar Catalan, merci pour vos conseils et vos coopérations.

Un grand merci à Geneviève qui m'a toujours donné ses conseils et soutiens. Merci à la famille Muon Le qui sont toujours à mon côté dans ma vie quotidienne. Merci à Trang *pour tes soutiens et tes conseils. Merci à tous mes amis qui m'ont aidé par un moyen ou un* autre pour arriver à la fin de ma thèse.

Je voudrais remercier VIED (Vietnam International Education Development -Ministry *of Education and Training) pour m'avoir fourni ma bourse pour réaliser cette thèse.*

Finalement, je tiens à remercier particulièrement les aides de ma famille, mes parents, mon mari et mes deux enfants, merci pour leurs soutiens.

Merci à toutes et à tous.

LIST OF ABBREVIATIONS

A, C, G, T	Adenine, Cytosine, Guanine, Thymine
BAC	Bacterial Artificial Chromosome
BDM	Bateson-Dobzhansky-Muller
BES	BAC End Sequence
bp	base pairs
BWA	Burrows-Wheeler Aligner
CO	Crossovers
DAPI	4',6-diamidino-2-phenylindole
DEGs	Differentially expressed genes
DNA	Deoxyribonucleic acid
DSB	Double strand break
FDR	First Division Restitution
FISH	Fluorescence in situ hybridization
GISH	Genomic in situ hybridization
HI	Hybrid incompatibility
Kb	Kilobase
LINE	Long interspersed nuclear elements
LTR	Long tandem repeat
Ma / Mya	Million years ago
Mb	Megabase
min	minute
miRNA	micro ARN interférents
mRNA	messenger RNA
NGS	Next Generation Sequencing
PCA	Principal component analysis
PCR	Polymerase Chain Reaction
Ph1	Pairing homoeologous 1
PMC	Pollen mother cell
PTGS	Post-transcriptional gene silencing
RNA	Ribonucleic Acid
RNAi	RNA interference
RPKM	Reads Per Kilobase per Million mapped read

RT	Reverse-Transcription
SDR	Second Division Restitution
siRNA	small interfering RNA
SINE	Short interspersed nuclear elements
SNP	Single Nucleotide Polymorphism
TEs	Transposable elements
WGD	Whole Genome Duplication

TABLE OF CONTENTS

GENERAL INTRODUCTION

I. POLYPLOIDY

I.1 Definition and evolutionary role

Polyploidization, or whole genome duplication (WGD), is a major evolutionary force in eukaryotes and is particularly prominent and recurrent in angiosperms (Comai 2005, Leitch and Leitch 2008, Madlung and Wendel 2013, Soltis, Visger et al. 2014). It is now well established that all plant species have experimented more than one cycle of polyploidy during their evolution (Garsmeur, Schnable et al. 2014, Jiao and Paterson 2014, Tiley, Ane et al. 2016).

Many important crop species, such as bread wheat (Triticum aestivum), rapeseed (Brassica napus) or cotton (Gossypium hirsutum), are relatively recent polyploids. Thus, polyploidy not only plays a significant role in plant speciation and evolution process but also has great impacts on agricultural resources (Renny-Byfield and Wendel 2014).

Autopolyploidy and allopolyploidy represent the two main categories of polyploidy that are distinguished based on the formation mode and the nature of genome duplication (Fig. 1) (Chen 2007, Tayale and Parisod 2013, Soltis, Visger et al. 2014).

I.1.1 Autopolyploidy

Autopolyploids contain more than two sets of homologous chromosomes in the nuclei, derived from the same species (Fig. 1).Because autopolyploids possess more than two homologous genomes, they are usually characterized by **polysomic inheritance** and multivalent formation at meiosis (Parisod, Holderegger et al. 2010).

The **polysomic inheritance** is the consequence of the multivalent formation where more than two homologous chromosomes can pair at meiosis. Such meiotic irregularities may result into gamete sterility and overall fertility reduction (Levy and Feldman 2002, Comai 2005, Hufton and Panopoulou 2009). The comparison of Arabidopsis autopolyploids with their corresponding diploids suggests that the decrease of chiasma frequency during the first meiosis may explain the decrease of fertility of autopolyploid (Lloyd and Bomblies 2016).

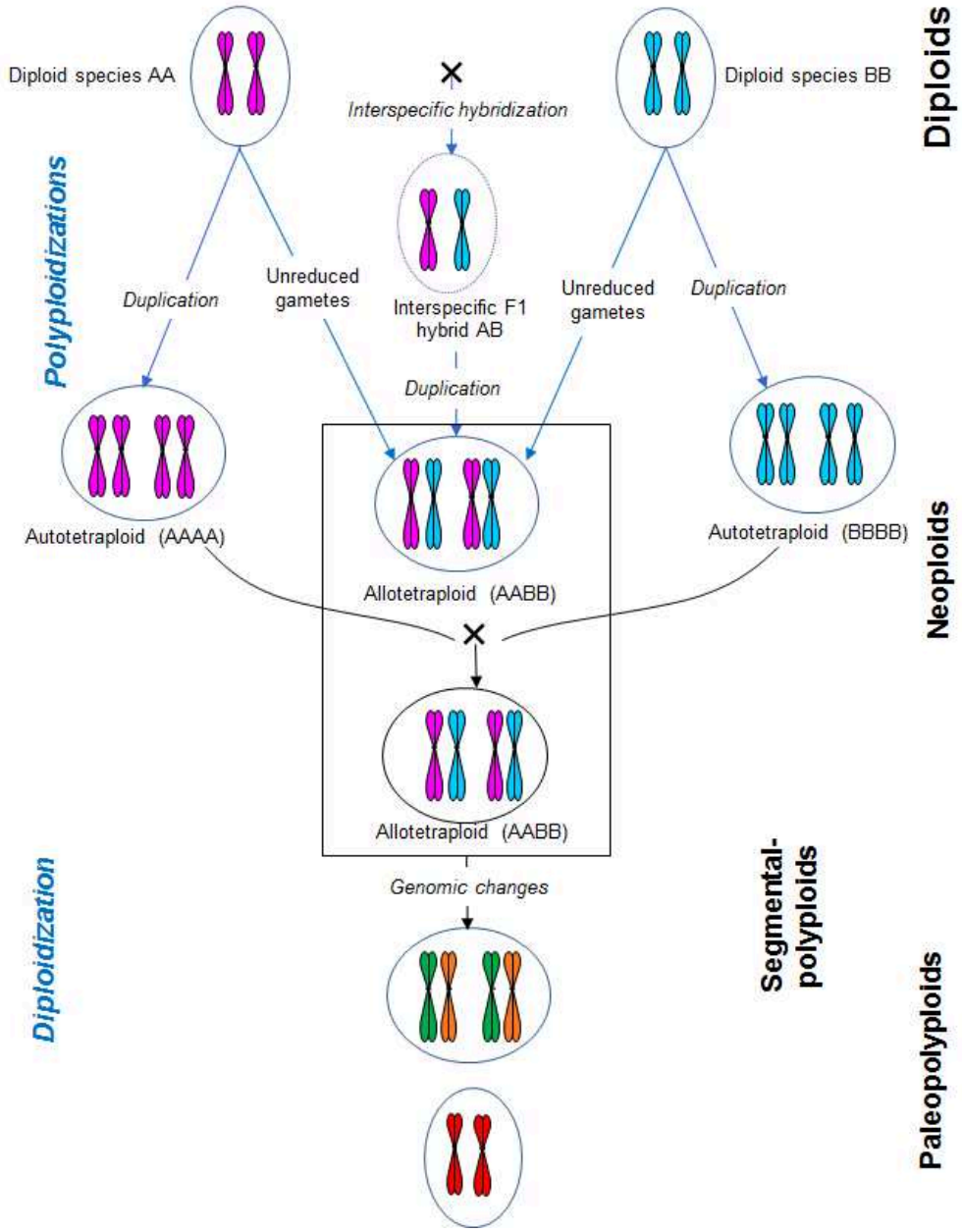


Fig. 1. Polyploid formation and Evolution. For simplicity, one pair of homologous chromosomes (pink or blue) is shown. Autopolyploid are formed by whole genome doubling. Allopolyploids are formed by interspecific hybridization; from either (i) fertilization between reduced (1n) gametes followed by genome doubling or (ii) by hybridization between two unreduced gametes (2n). The reorganization of polyploid genome leads to the formation of segmental allopolyploid (illustrated by changing of color of homologous chromosome in orange and green). Neopolyploids undergo across their evolution diploidization, leading to paleopolyploid (intense genome reorganization referred by changing colors in red of chromosomes). Adapted from Comai (2005); Chen (2007); Tayale and Parisod (2013); Soltis, Visger et al. (2014).

In the case of formation of autopolyploids by hybridization between different genotypes of the same species, a kind of heterozygosity, conferring genome flexibility, is maintained and **polysomic inheritance** could be advantageous. In this type of autotetraploids, we can expect the possibility of pairing of a given chromosome with one of its 3 remaining homologues, resulting in more possibilities of pairing between alleles at a given locus. For example, progenies of autoplolyploids heterozygotes at a given locus ($A_1A_1A_2A_2$) would result in genotypic ratios of 1/36 $A_1A_1A_1A_1$, 8/36 $A_1A_1A_1A_2$, 18/36 $A_1A_1A_2A_2$, 8/36 $A_1A_2A_2A_2$, and l/36 $A_2A_2A_2A_2$, which equates to 6% homozygotes and 94% heterozygotes. This heterozygosity would confer autopolyploids to adapt better than diploids in larger and variable environment (Soltis, Soltis et al. 2007, Parisod, Holderegger et al. 2010, Weiss-Schneeweiss, Emadzade et al. 2013).

Many important cultivated crops are autopolyploids, such as alfalfa (Medicago sativa 2n = 4x = 32) (Young, Debelle et al. 2011) and potato (Solanum tuberosum L., 2n = 4x = 48) (Consortium 2011).

Autopolyploidy is often associated with sterility and reduced seed production. Thus many autopolyploids are perennial species that also propagate vegetatively (Levy and Feldman 2002).

The importance of autopolyploidy in eukaryote evolution still needs to be appreciated. Recent investigations on autopolyploids indicate that the role of autopolyploidy in evolution may have been underappreciated in comparison to allopolyploidy (Soltis, Soltis et al. 2007, Parisod, Holderegger et al. 2010, Barker, Arrigo et al. 2016). The rate of autopolyploid formation may often be higher than those of allopolyploid formation (Ramsey and Schemske 1998), and autopolyploidy could also represent an important mechanism of speciation in plants (Soltis, Soltis et al. 2007, Ramsey and Ramsey 2014, Edger, Heidel-Fischer et al. 2015).

I.1.2 Allopolyploidy

Similar to autopolyploidy, allopolyploidy consists also in the assortment of more than two complete sets of chromosomes in the genome which, in contrast, originate from two different species, usually belonging to the same genus or closely-related genera (Fig. 1) (Stebbins 1985, Comai 2005, Soltis, Albert et al. 2009, Abbott, Albach et al. 2013, Estep,

McKain et al. 2014, del Pozo and Ramirez-Parra 2015).

Allopolyploidy is an important evolutionary pathway in plants (Soltis and Soltis 2009) and allopolyploids are found in almost plant taxa including important crops such as wheat, Brassica, and cotton. The combination of two or more different chromosome sets offers allopolyploid adaptive evolutionary advantages like hybrid vigor and genetic divergences (Ni, Kim et al. 2009, Chen 2010, Groszmann, Greaves et al. 2013). However, genome rearrangements that occur during allopolyploid evolution also cause the instabilities, imbalances and incompatibilities between allopolyploid subgenomes (Bikard, Patel et al. 2009, Burkart-Waco, Josefsson et al. 2012, Birchler and Veitia 2014).

I.1.3 Segmental auto and allopolyploidy

Segmental allopolyploidy term was first proposed by Stebbins (1947) as a polyploid which contains both homologous and homoeologous chromosome segments. Thus, they can exhibit both bivalents and multivalents during meiosis (Levy and Feldman 2002). A newly-formed allotetraploid may become a segmental allotetraploid or autopolyploid if homoeologous chromosomes, originating from parental species contain some homologous chromosomal segments (Fig. 1) (Chen 2007). This can be due to the hybrid origin of parental species genome that may have hybridize and introgress genome segments into each other (Sybenga 1996). It can be also due to crossover or non-crossover based homeologous exchanges between the two constituent homoeologous subgenomes (Chalhoub, Denoeud et al. 2014). The evidences of segmental auto- or allopolyploid were found in certain species including: maize (Zea mays) (Gaut and Doebley 1997), Hordeum. murinum subsp. leporinum (Eilam, Anikster et al. 2009), Leucaena benth (Boff and Schifino-Wittmann 2003) and more recently in B. napus (Chalhoub, Denoeud et al. 2014). It has been suggested that homoeologous exchanges between homoeologous chromosomes in the post-neolythic B. napus allopolyploid may lead to segmental autopolyploidy.

I.1.4 Aneuploidy

Aneuploidy refers to the loss or gain of individual chromosome(s), leading to change of chromosome number of the haploid chromosome set of the species genome (Torres, Williams et al. 2008, Birchler 2013). Aneuploidy has been observed in human, animals and

plant genomes (Weaver and Cleveland 2009, Hulten, Patel et al. 2010) resulting in gene dosage unbalance (Birchler 2014).

Numerous studies have indicated that chromosome combinational confusions during meiosis or mitosis cause abnormal chromosome number segregation (Compton 2011, Nagaoka, Hassold et al. 2012) leading to different types of aneuploids in offspring (Fig. 2), such as:

- Monosomy: the loss of a single chromosome, 2n-1.
- Nullisomy: the loss of a pair of homologous chromosomes in the genome, 2n-2
- Trisomy: the gain of an extra copy of a chromosome, 2n+1
- Tetrasomy: the gain of an extra pair of homologous chromosomes, 2n+2.

Aneuploidy may occur for more than one pair of homologous chromosomes, example: the gain of an extra copy of a chromosome from two different pairs will form a double trisomy (2n+1+1) (Fig. 2).

Impacts of aneuploidy on animals are usually more known than those in plants (Matzke, Mette et al. 2003). In Drosophila, aneuploidy often leads to lethality. In human, aneuploidy is often associated with malignant cell proliferation carcinogenesis or many classical syndromes, like Down's syndrome, Edward's syndrome or the formation of solid tumors (Torres, Williams et al. 2008, Weaver and Cleveland 2009, Hulten, Patel et al. 2010, Birchler 2014).

In allopolyploid plants, aneuploidy mainly results from meiotic irregularities and it is also more frequent, probably because it is more tolerated than in diploid parental species (Hufton and Panopoulou 2009). Aneuploid formation is frequent in first generations of resynthesized allopolyploids such as Brassica and wheat (Zhang, Bian et al. 2013, Zhou, Tan et al. 2016).

The incomplete chromosome set of aneuploids leads to the unbalance in gene dosage which has also effects on phenotype. The earlier research of Datura recovered a trisomic aneuploid for all 12 chromosomes of the genome each of which has a characterized phenotype that allows their distinction from the euploid progenitors (Blakeslee 1934). In a recent study on Arabidopsis, Henry, Dilkes et al. (2010) indicated that the phenotype of aneuploids was strongly associated with the dosage of specific chromosome types and that chromosomal effects can be additive.

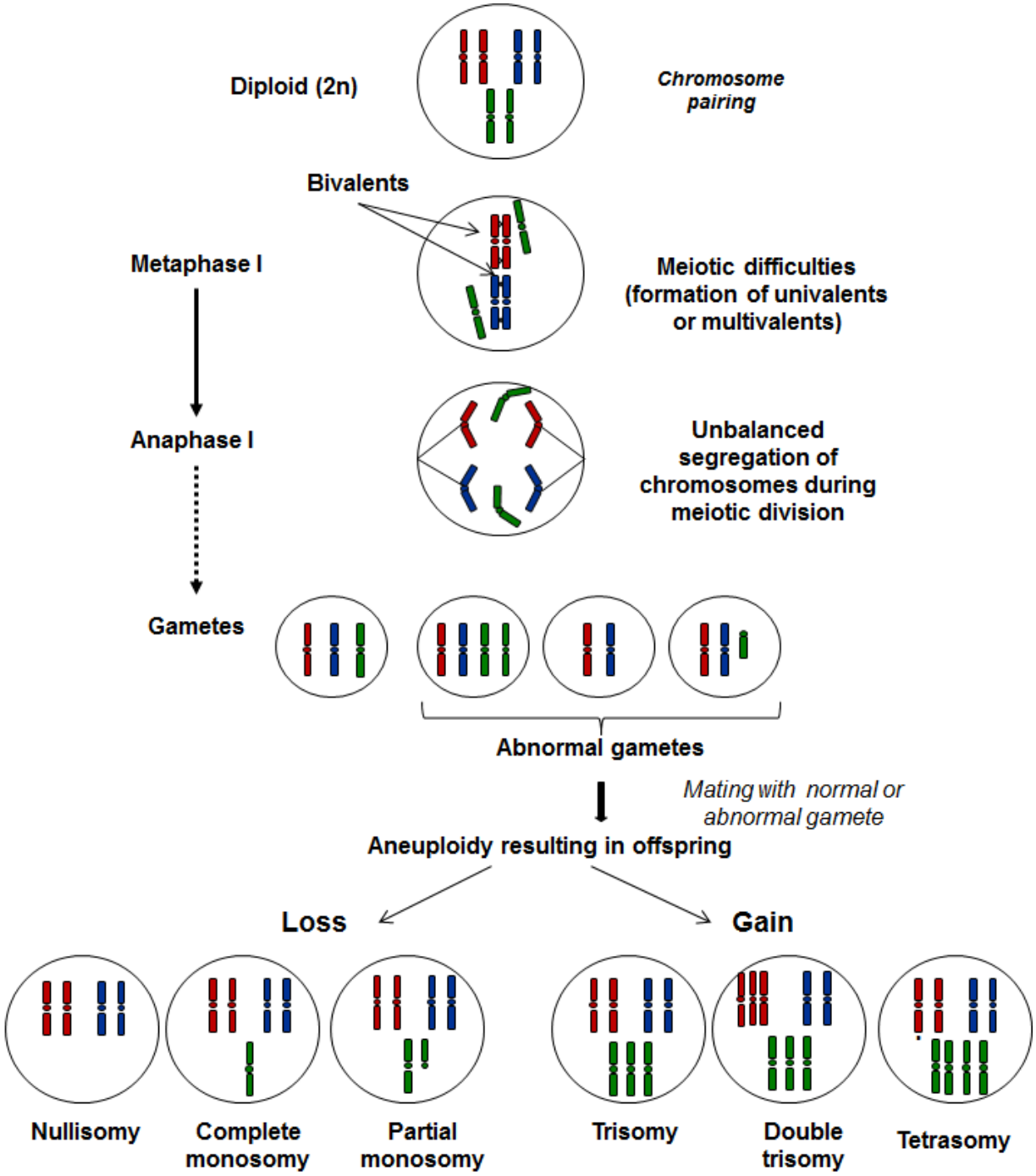


Fig. 2. The origin and types of aneuploids. Difficulties of meiosis in the diploid lead to the unbalanced segregation of homologous chromosomes during meiotic division, resulting in abnormal gametes. The combination of abnormal gametes with normal or other abnormal ones lead to the formation of aneuploids in offspring, with the loss or gain of chromosomes.

I.2 Natural and synthesis polyploidy formation and interest

I.2.1 Formation of polyploids in nature

It is clearly evidenced now that all angiosperms have undergone at least one round of polyploidy (Garsmeur, Schnable et al. 2014, Jiao and Paterson 2014, Tiley, Ane et al. 2016) implying that the formation of polyploids is frequent in nature. There are several pathways leading to polyploid formation (Ramsey and Schemske 1998, Wendel 2000, Ramsey and Schemske 2002, Tayale and Parisod 2013). The classification of polyploids into two main categories (autopolyploid and allopolyploid) is based on the way of their formations.

The spontaneous doubling of chromosome sets in somatic cells and the uniting of unreduced gametes results to polyploidy formation (Bretagnolle and Thompson 1995, Tayale and Parisod 2013).

The **formation of autopolyploid** relates to genome doubling of the same species (Ramsey and Schemske 1998). It is suggested that autopolyploids can be formed either by fertilization of unreduced gametes (2n) from the same species (Fig. 3) or by chromosome doubling of somatic cell from which a plant is generated (Bretagnolle and Thompson 1995, De Storme and Geelen 2013). The prevalence of autopolyploid species in nature indicates that autopolyploidy is probably a more common and important element of plant diversity than historic views suggest (Soltis, Soltis et al. 2007, Barker, Arrigo et al. 2016).

Hybridization between two related species from the same or related genera, (i) directly between unreduced gametes or (ii) between reduced gametes followed by genome doubling, leads to **allopolyploid formation** (Ramsey and Schemske 1998, Soltis and Soltis 2009). Interspecific hybridization can happen between species of the same or different ploidy levels. There are two main pathways for the formation of allopolyploids, both of which consist in hybridization between two different species.

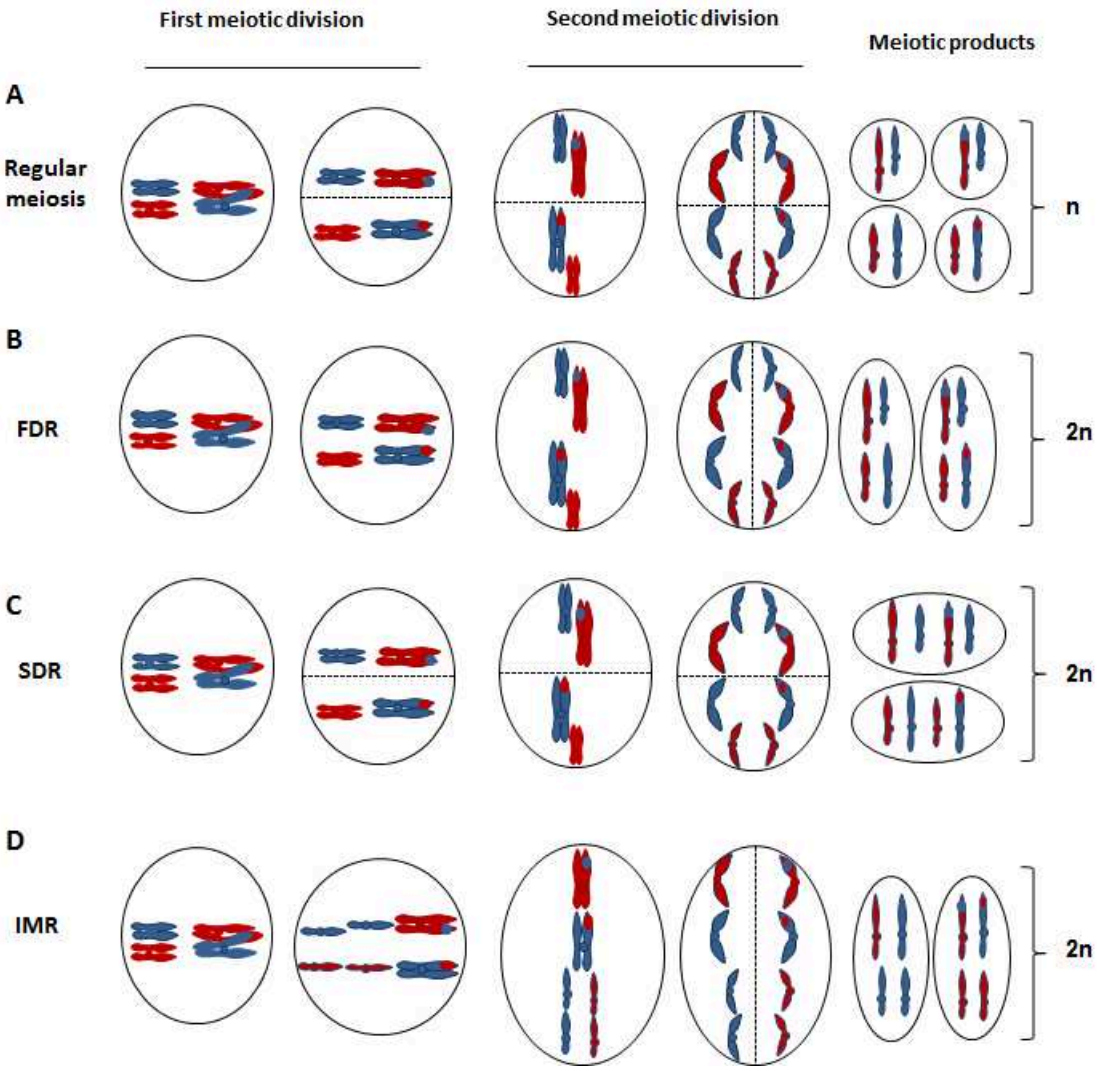


Fig. 3. Different mechanisms lead to the formation of unreduced gamete during the meiosis. (A) Regular meiosis with two meiotic division results in four n gametes in the meiotic products. (B) FDR (first division restitution,) the absence of first meiotic division results into two 2n gamete containing the non-sister chromatids. (C) SDR (second division restitution), the abnormal mechanism occurred in second meiotic division, products the sister chromatids. (D) IMR (Indeterminate meiotic restitution) relates to the univalent state in metaphase I. This figure was adapted from: Bretagnolle and Thompson (1995); Lim, Ramanna et al. (2001)

(i) The one-step model

Allopolyploids can be formed directly by hybridization between unreduced gametes (2n) of two different species. This one-step model pathway results directly in an allopolyploid formation because each of the chromosome set, from each parental species, is already doubled. This pathway of allopolyploid formation was reported to be the mode of formation of Digitalis Allium and Manihot allopolyploids (Ramsey and Schemske 1998).

(ii) The two-steps model

The two-step model of allopolyploid formation consists in a hybridization step between normal reduced gametes (1n) of two different species, resulting in F1 interspecific hybrid, which is followed by a step of chromosome doubling (Ramsey and Schemske 1998). The F1 interspecific amphihaploid hybrid contains one set of chromosomes from each parental species that are difficult to pair and cannot have normal meiosis resulting in sterility. Chromosome doubling, leading to two sets of homologous chromosomes from each parental species genome, can naturally occur at somatic or zygotic stages and restore fertility of the nascent allopolyploid as pairing at meiosis occur between the homologous chromosome pairs (Fig. 1).

The formation of bread allohexaploid wheat (Triticum aestivum, 2n = 6x = 42, AABBDD) is a classical example for this pathway of allopolyploid formation. Bread wheat was formed by two successive allopolyploidization events. The first event occurred between the diploid species T. urartu (2n = 2x = 14, AA) and a yet unidentified diploid Aegilops species of the section Sitopsis, donor of the B genome, that led to the formation of the durum or the pasta allotetraploid wheat, T. turgidum (2n = 4x = 28, AABB), less than 0.5 million year ago (Feldman 1995, Feldman 2000, Huang, Sirikhachornkit et al. 2002). The second allopolyploidization event occurred upon domestication 7,000–12,000 years ago, between the allotetraploid T. turgidum ssp. dicoccum and the goatgrass diploid species Ae. tauschii (2n = 2x = 14, DD), resulting in the allohexaploid wheat (Feldman 2000, Ozkan, Levy et al. 2001, Zhang, Zhu et al. 2014).

Formation of unreduced gametes

In plants, unreduced gametes contain two copies of each chromosome (2n) similar to the pollen mother cell, instead of the normal haploid (1n). Different mechanisms of unreduced gamete formation have been described. These include pre-meiotic genome

doubling, absence of first-division restitution (FDR), absence of indeterminate meiotic restitution (IMR), chromosome replication during the meiotic interphase, absence of second-division restitution (SDR), post-meiotic genome doubling, and apospory[1] (Fig. 3) (Bretagnolle and Thompson 1995, Lim, Ramanna et al. 2001, Brownfield and Köhler 2011, Cuenca, Aleza et al. 2015, Mason and Pires 2015).

In unreduced gametes occurring because of absence of FDR, the pairing and/or the separation of homologous chromosomes do not occur at anaphase I, resulting in the absence of the first meiotic division, whereas the second meiotic division occurs normally (Fig. 3B). Thus the FDR meiotic products possess two non-sister chromatids instead of one. This type of unreduced gamete formation has been observed in several species of the Poaceae family such as wheat (Jauhar 2007).

When unreduced gametes occurred because of abnormal SDR, the pairing and the separation of the homologous chromosomes occur normally during the first meiotic division but the sister chromatids do not separate during the second meiotic division (Fig. 3C). Therefore, the SDR meiotic products contain two sister chromatids (2n).

The IMR is an unreduced gamete formation which cannot be classified as either FDR or SDR. In this case, certain univalents divide equally (following centromere division) during the first division while bivalents disjoin normally in anaphase I. The following step in cytokinesis divides chromatids into two group, both of which contain sister and non-sister chromatid in IMR meiotic products (Fig. 3D) (Lim, Ramanna et al. 2001).

Hybridization between unreduced gametes is one notable mechanism of both autopolyploid and allopolyploid formation (Brownfield and Köhler 2011). The formation of autopolyploids may result from the combination of the 2n gametes from the same or two different genotypes of the same species (Bretagnolle and Thompson 1995) whereas the association between unreduced gametes from two different species leads to direct formation of allopolyploids.

Generation of synthetic allopolyploids using a similar one-step model pathway has reported in several studies. The 2n gamete formation is favored by primarily obtaining autopolyploids from the two diploid parental species. These produce 2n gametes and their hybridization directly results in allopolyploids. Such method has been successful in oat,

[1] . Apospory is the development of 2n gametophytes, without meiosis and spores, from vegetative, or nonreproductive, cells of the sporophyte. https://www.britannica.com/topic/plant/Reproduction-and-life-histories#ref536901

maize (Kynast, Davis et al. 2012), Brassica (Mason, Nelson et al. 2011) and Agave species (Gómez-Rodríguez, Rodríguez-Garay et al. 2012).

I.2.2 Generation of synthetic polyploids

Since the first discovery that application of colchicine leads to genome doubling (Blakeslee and Avery 1937), different procedures were used to create synthetic autopolyploids and allopolyploids, reproducing the different ways of polyploid formation in nature.

The classical procedure to create synthetic allopolyploids comprises two steps: obtaining F1 interspecific hybrids through hybridization between two related species followed by chromosome doubling through colchicine treatment of F1 dihaploid hybrids to restore fertility. This "two-step" method has been applied on many different species in order to develop synthetic allopolyploids (Song, Tang et al. 1993, Mestiri, Chagué et al. 2010). Normally, hybridization between different species in nature copes with the difficulties of reproductive isolations (pre-zygotic isolation), but forced-crossing manipulations normally overcomes these obstacles to create F1 dihaploid hybrids which contains two haploid chromosome sets of the parental species (n+n'). As they do not have a second copy of homologous chromosomes to pair and pairing between the homoeologous chromosomes (n+n') is difficult or impossible, these F1 haploid hybrids are sterile unless chromosome doubling (2n+2n') happens, spontaneously, or artificially leading to fertile allopolyploids (Ozkan, Levy et al. 2001). If chromosome doubling does not occur naturally, the doubling of chromosomes is forced through chemical treatments. Several chemical products were developed and used, the most common is colchicine which acts as super inhibitor of the mitosis, preventing the formation of mitotic spindle fibers and preventing chromosomes from segregation at the end of anaphase (Blakeslee and Avery 1937, Ranney 2006). Chromosome doubling by colchicine is usually applied on the meristematic zone stage of cell division.

Resynthesizing allopolyploids by interspecific hybridization followed by genome doubling have been done in numerous species and allowed to shed the light on polyploidy-related evolutionary processes. Studies on Brassica Song, Tang et al. (1993) reported for the first time a complete set of synthetic Brassica amphidiploids derived by reciprocal hybridization. The structural changes in synthetic polyploid genome in five generations from

F1 to F5 have been early reported (Song, Lu et al. 1995). In wheat, one model system of grass, synthetic polyploids have been created (Zhang, Sun et al. 2007, Mestiri, Chagué et al. 2010, Zhao, Zhu et al. 2011) and studied at the, structural, epigenetic and functional levels.

The recent emergence of several polyploid systems with the generation of synthetic allopolyploid have been reported in various species, such as wheat (Feldman 1995), Brassica (Song, Lu et al. 1995), cotton (Adams, Percifield et al. 2004), tobacco (Leitch, Hanson et al. 2008) or other plants as Arabidopsis (Comai 2000), Tragopogon (Soltis, Soltis et al. 2004). For other species, such as Spatina anglica, efforts for having synthetic polyploids did not get success (Aïnouche, Fortune et al. 2009). In conclusion, having synthetic polyploids represents one essential condition for a successful polyploid model.

Plant synthetic polyploids were generated not only for research purposes, but they are largely used in both crop plants, such as wheat, Brassica as well as garden plants, such as Fragaria vesca, Chrysanthemum. These man-made polyploids were created depending on human needs and commercial interests because of several advantages that polyploidization can offer.

Synthetic polyploids occupy now a significant place is plant breeding. They provide resources and new diversity for introgression and improvement of similar natural polyploids. In 1950, Stebbins designated artificial polyploidy as a tool in plant breeding (Stebbins Jr 1950). Synthetic polyploids in crop plant such as wheat, maize, rice, Brassica have been widely used in several studies (Swigonova, Lai et al. 2004, van Ginkel and Ogbonnaya 2007, Wang, Zhang et al. 2013). Synthetic polyploidy have also created new crop species such as Triticale, an artificial polyploidy between wheat and rye (Khalil, Ehdaeivand et al. 2015). It is undeniable that domesticated synthetic polyploid is increasing and becoming more popular over years (FAO document: ftp://ftp.fao.org/docrep/fao/012/i1070e/i1070e04.pdf).

Having resynthesized polyploids allows investigating immediate changes following polyploid formation through generations, which cannot be investigated in older natural polyploids (Song, Lu et al. 1995). Precise comparisons with identified progenitors allow inferring of changes induced by polyploidy. For example, it has been early found that synthetic allotetraploids of Arabidopsis and Cardaminopsis arenosa show gene silencing and changes in DNA methylation (Comai 2000, Madlung, Masuelli et al. 2002, Wang, Tian et al. 2004, Akama, Shimizu-Inatsugi et al. 2014). The data of gene expression, through mRNA sequencing, shows different patterns of homoeologous gene expression and contribution in

both resynthesized and natural Brassica napus, suggesting that changes in homoeologous gene expression regulation can occur immediately after polyploid formation and could be maintained at the long-term evolution (Higgins, Magusin et al. 2012, Chalhoub, Denoeud et al. 2014). Studying of wheat hexaploids shed the light on the prevalence of additive gene expression in highly stable polyploids (Chagué, Just et al. 2010, Chelaifa, Chagué et al. 2013).

II. GENOMIC CHANGES IN POLYPLOID GENOME

More and more studies show that polyploids are subjected to important changes at the genetic, functional and epigenetic levels, starting from the first generations of their formation and all across their evolution (Wolfe 2001, Paterson 2006, Adams 2007, Doyle, Flagel et al. 2008, Flagel and Wendel 2009, Renny-Byfield and Wendel 2014, Soltis, Visger et al. 2014, Wendel, Jackson et al. 2016).

II.1 Chromosome and DNA rearrangements

Chromosome and DNA rearrangements refer to large or small chromosome change involving deletion, duplication, inversion (paracentric and pericentric) and translocation (Fig. 4) (Chen and Ni 2006, Hufton and Panopoulou 2009). These are common pathways of allopolyploid evolution and can occur immediately in first generations that follow polyploidy formation as well as at the long-term of allopolyploid evolution (Leitch and Leitch 2008, Renny-Byfield and Wendel 2014, Wendel, Jackson et al. 2016).

Cytological and molecular evidences indicate that almost chromosomal rearrangements originated from interaction between homoeologous chromosomes, meiotic and mitotic defects which lead to the breakage and rejoining of chromosome segments (Fig. 4). Consequences of these types of rearrangements consist in chromosome segmental deletion, duplication, inversion or reciprocal translocation (Griffiths, Gelbart et al. 1999). In meiosis, the miss-recombination of homoeologous chromosomes causes chromosome mis-segregation. The formation of multivalents instead of bivalents in metaphase results in crossovers (CO) that favor deletions or duplications of chromosome segments. The intra-chromatid recombination among duplicated sequences in diploids and polyploids are also the sources of segment rearrangements. The inversion of DNA fragments can happen in two ways: paracentric, if the inversion happens outside the centromere or pericentric if the inversion spans the centromere. This type of chromosome rearrangements does not lead to immediate effects but the subsequent generations may have difficulties in recombination between homoeologous chromosomes, especially the pericentric inversion case, which can produce all types of chromosomal rearrangements because of the formation of the crossover in loop. The consequence of translocation can occur in later generations by the formation of.

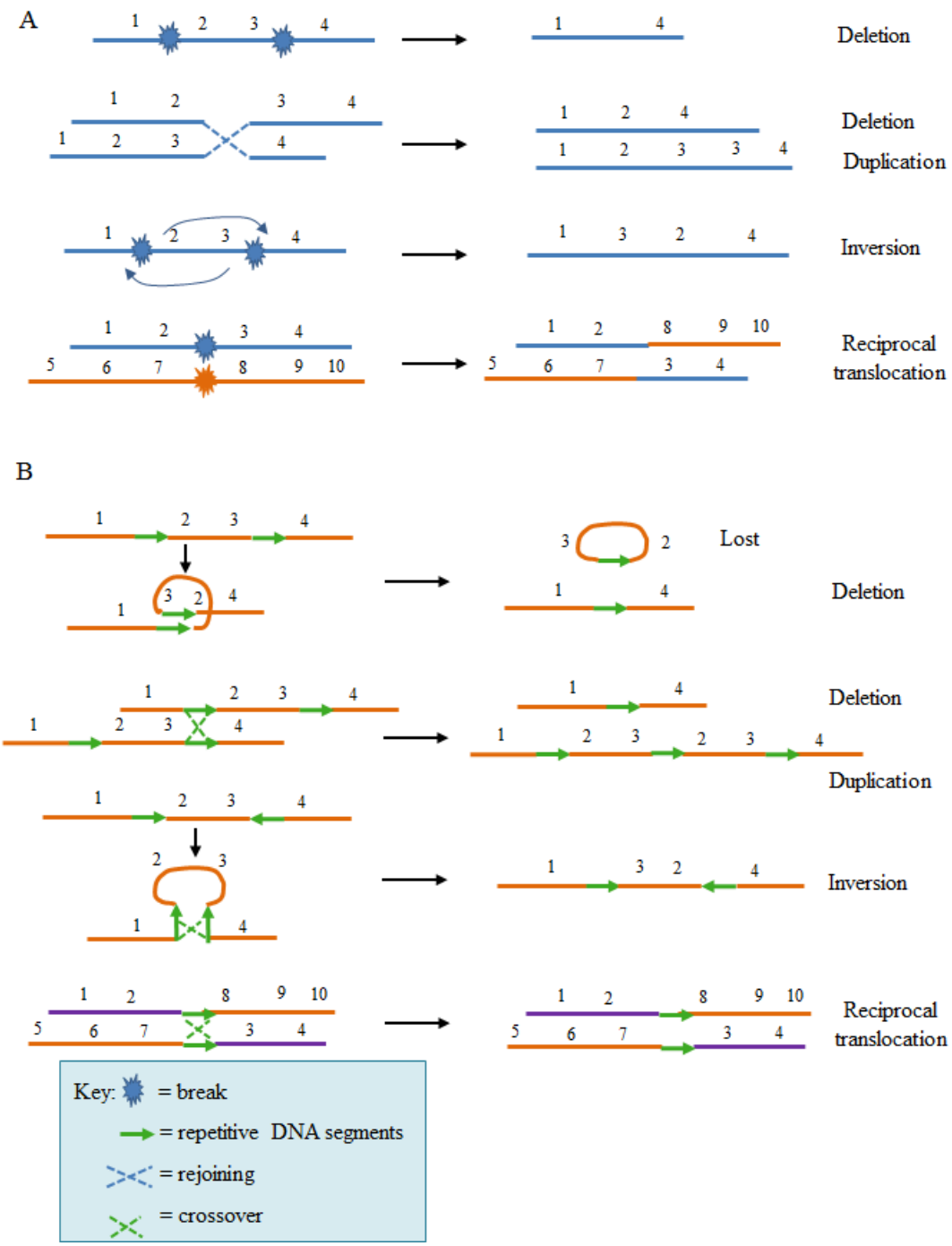


Fig. 4. Types of Chromosome rearrangements. **(A)** Homoeologous recombinations result in chromosome rearrangements via breakage and rejoining and lead to chromosome deletion, duplication, inversion or reciprocal translocation. **(B)** Chromosome rearrangements via crossover between DNA repetitive sequences result in DNA deletion, duplication, inversion or reciprocal translocation.

From Griffiths, Gelbart et al. (1999).

pairing configuration resulting into two types of segregation: adjacent 1, 2 and alternate and deriving chromosome loss and replacement or aneuploidy (Fig. 5) (Gaeta and Chris Pires 2010)

Evidences of genomic restructuring have been well documented in a number of resynthesized polyploids. In Brassica polyploids, massive genome changes, including loss or gain of parental alleles and/or appearance of novel alleles, happened immediately after allopolyploidization (Song, Lu et al. 1995, Szadkowski, Eber et al. 2010). In Arabidopsis autopolyploid, large rearrangements such as inversion, translocation, deletion or aneuploidy have been also observed (Weiss and Maluszynska 2000). Gene loss at the Hardness locus due to large genomic deletion in established accessions was found in Triticum aestivum (Chantret, Salse et al. 2005).

Genome rearrangements can lead to either expansion or contraction of polyploid genomes. There are evidences that polyploid genome size may be increased, decreased or no change occurs after polyploidizations (Renny-Byfield and Wendel 2014, Wendel, Jackson et al. 2016). Only few studies showed genome expansion after polyploid formation such as Nicotiana (Leitch, Hanson et al. 2008). In contrast, a large-scale analysis of 3,008 angiosperms revealed that the decrease of genome size is a widespread biological response to polyploidization (Leitch and Bennett 2004). Earlier studies confirmed that the rapid loss of DNA in early generations after polyploidization seems to be the cause of genome downsizing as in Tragopogon (Buggs, 2009) and wheat (Ozkan, Levy et al. 2001, Kashkush, Feldman et al. 2002, Ozkan, Tuna et al. 2003, Eilam, Anikster et al. 2009).

II.2 Dynamics of Transposable elements

Transposable elements (TEs) are mobile DNA fragments that can move (or transpose) in the genome under certain conditions. TEs are classified into two main classes: retrotransposons, which move via RNA intermediate by a ‘copy-and-paste’ mechanism, and DNA transposons, which move via DNA intermediate by either a ‘cut-and-paste’ mechanism or replication of the DNA copy (Wicker, Sabot et al. 2007, Parisod, Alix et al. 2010, Levin and Moran 2011) (Fig. 6).

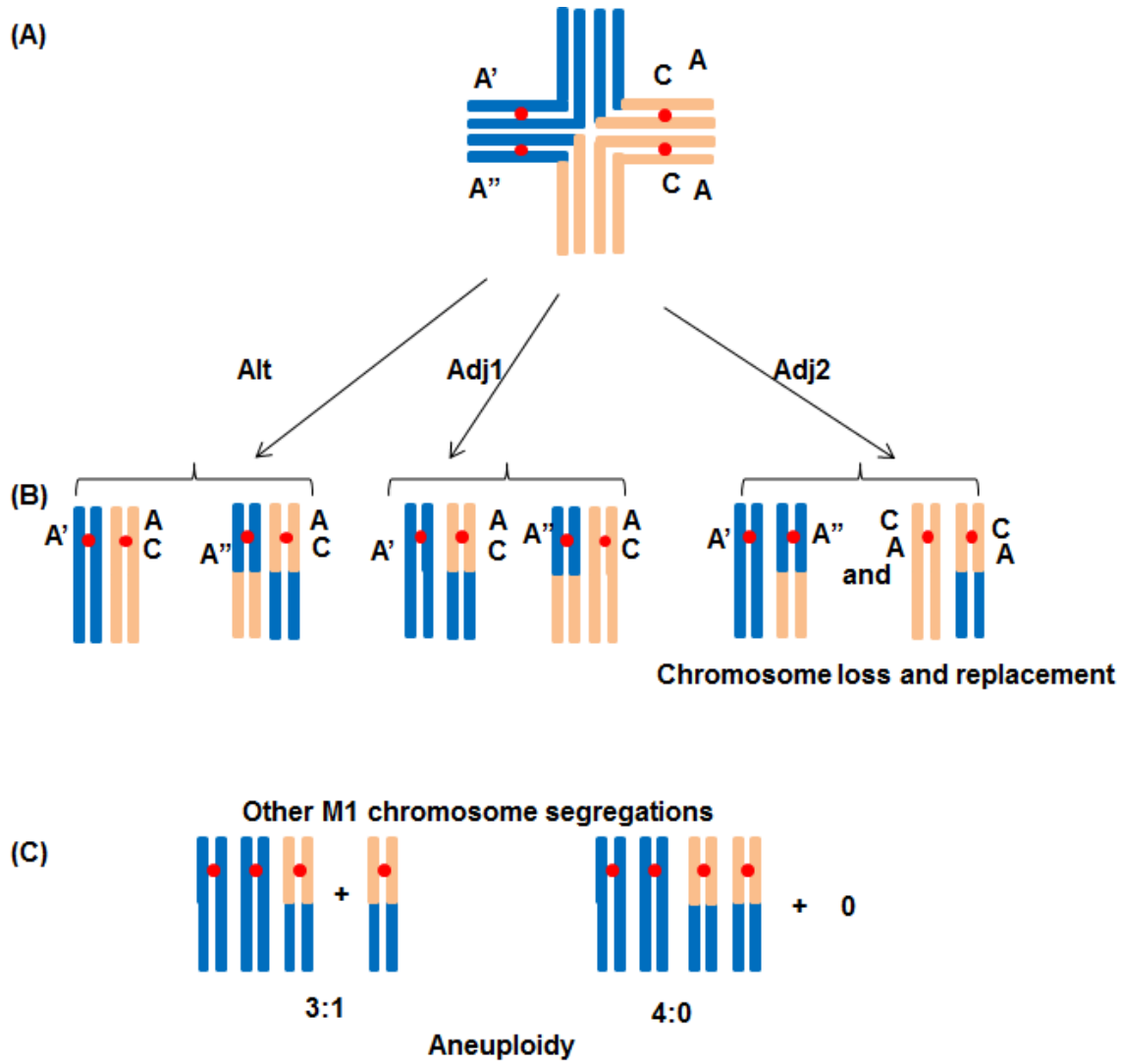


Fig. 5. Meiosis I in an allopolyploid carrying a homoeologous translocation. Homoeologous chromosomes A (blue bars) and C (orrange bars) are shown. (A) The C-homoeologs carry a terminal translocation from the distal end of the A chromosome (CA) resulting from homoeologous recombination in a previous generation. Lines carrying such translocations pair in a cross-like configuration. (B) In this example, if we assume no new recombination, chromosome segregation that is alternate (Alt) or adjacent 1 (Adj1) will generate daughter cells of the parental type. However, if adjacent 2 (Adj2) segregation occurs, daughter cells will be produced, lacking one homoeolog or the other. In this way, homoeologous chromosome loss and replacement can occur, as well as the loss of homoeologous centromeres. (C) Depending on the location and number of crossovers, meiosis I chromosome segregation could also lead to non-disjunction (3 : 1 or 4 : 0 chromosome segregations) and aneuploidy (Adapted from Gaeta, Yoo et al. (2009).

Since their first discovery by McClintock (1948), TEs were found in all prokaryote and eukaryote species. Important efforts have been realized in order to understand TEs, their behaviors and mechanisms of activation (Vicient, Jaaskelainen et al. 2001, Wicker, Sabot et al. 2007, Levin and Moran 2011, Bennetzen and Wang 2014). TEs are more abundant in plants than in animal (Lee and Kim 2014).

Based on their mechanisms of transpositions and structural features, Wicker, Sabot et al. (2007) proposed a classification into classes, families and subfamilies:

- **Class I TEs** consist of all TEs that transpose via an RNA intermediate, which is retrotranscribed and insert into the genome. These include long terminal repeat (LTR)- and non-LTR, terminal repeat retrotransposons in miniature (TRIMs) which lack coding sequences such as short and long interspersed nuclear elements (SINE, and LINE) retrotransposons (Fig. 6).
- **Class II TEs** consist in all transposons which DNA sequences can directly excise or replicate and transpose into other loci of the genome. These have terminal inverted repeats (TIRs,) with the exception of helitrons which replicate via a rolling-circle mechanism (Wicker, Sabot et al. 2007, Parisod, Alix et al. 2010) (Fig. 6).

The activation of TEs in the genome can lead to important restructuring, including deletions, duplications, inversions or reciprocal translocations (Fig. 4B). TE proliferation can lead to important genome size increase (Charles, Belcram et al. 2008, Baucom, Estill et al. 2009). TEs are also subjected to deletions by mechanisms that generate small or large deletions such as unequal homologous recombination and illegitimate recombination, resulting into genome downsizing (Devos, Brown et al. 2002, Bennetzen and Wang 2014).

In plants, TEs occupy a significant fraction of the genome and play an important role in their evolution where retrotransposons are more represented (reviewed by Bennetzen and Wang (2014)). In Brachypodium distachyon, TEs represent 28% of the genome, majority of them (27.33% of the genome) are retrotransposons (IBI 2010).

Hybridizations and polyploidizations, representing a “genetic shock” (McClintock 1984), may constitute conditions for activation or repression of TEs as well as their epigenetic control, depending on the studied model (Liu and Wendel 2000, Kashkush, Feldman et al. 2002, Petit, Guidat et al. 2010, Chalhoub, Denoeud et al. 2014). Species with big genome size have a major fraction of TEs. The numerous TEs in genome could be involved in resistance and adaptation to stress conditions (Sabot, Simon et al. 2004). Only in

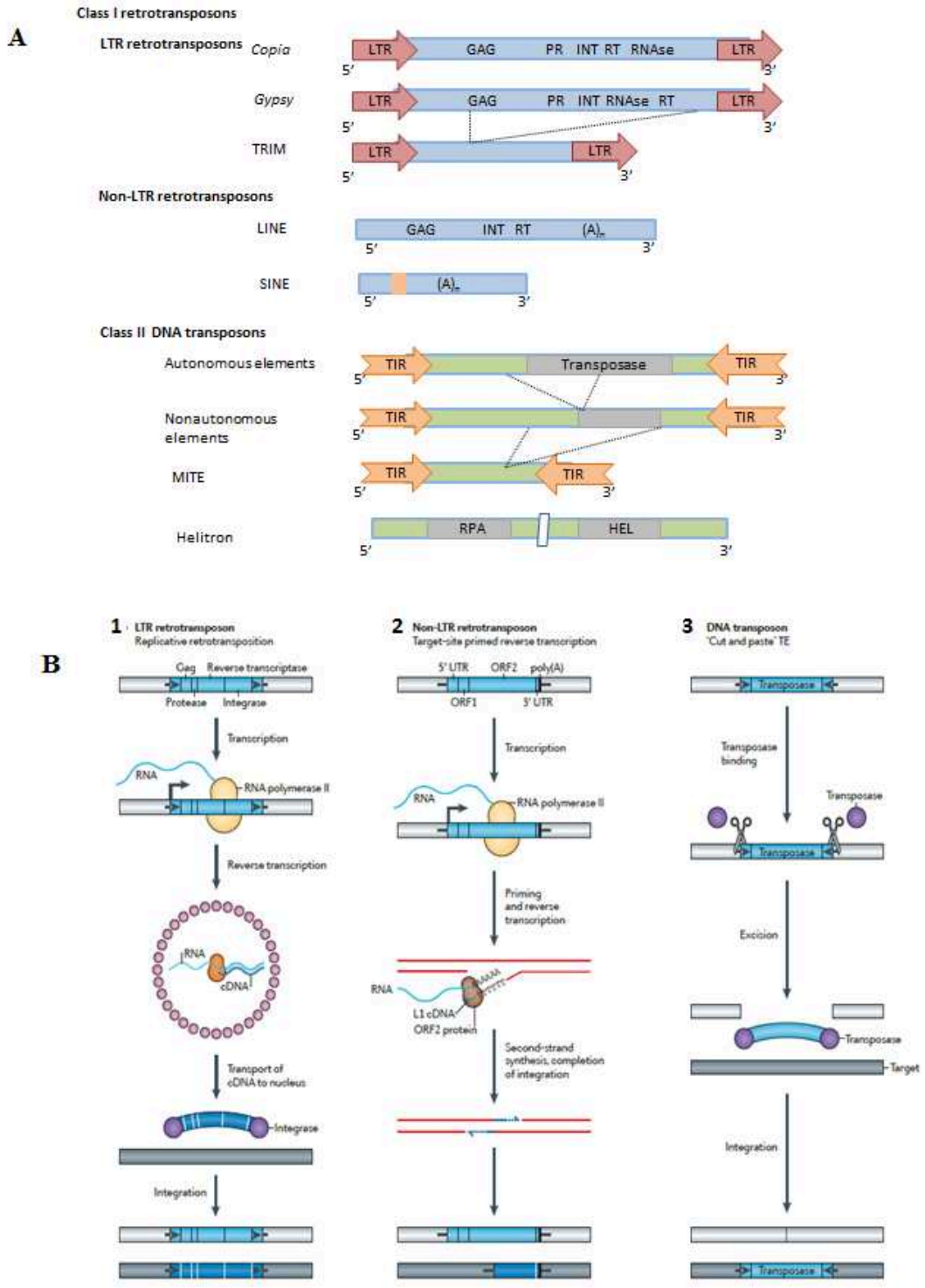


Fig. 6. Transposable elements (TEs) classification and their transposition mechanism. (A) TEs frequently observed in plant genomes. Class I TEs or restrotransposons move via RNA intermediates by a 'copy-and-paste' mechanism which is illustrated in (B1) and (B2) (the long terminal repeats (LTRs) restrotransposons and non-LTRs retrotransposons, respectively). Class II TEs or DNA transposons move via DNA intermediates, by either a 'cut-and-paste' mechanism or by replication of the DNA copy, illustrated in B3. (Adapted from Wicker, Sabot et al. (2007), Parisod, Alix et al. (2010); Levin and Moran (2011)).

wheat allopolyploids, genome restructuring such as gene loss resulting from changes involving deletion/insertion of TEs was suggested (Kashkush, Feldman et al. 2002, Hu, Hawkins et al. 2010, Kraitshtein, Yaakov et al. 2010) but this was not confirmed in other wheat allopolyploids (Mestiri, Chagué et al. 2010). In a comprehensive study of the recently formed Brassica napus allopolyploid TEs were shown to having less proliferated that in their parental species genomes (Sarilar, Palacios et al. 2013, Chalhoub, Denoeud et al. 2014).

II. 3 Epigenetic changes

Epigenetic variation refers to changes in phenotype or molecular behavior without underlying changes in its DNA sequence (Madlung 2013). Epigenetic changes have been reported in both plants and animals (Feng, Jacobsen et al. 2010, Law and Jacobsen 2010, Jones 2012, Kumar, Kumari et al. 2013, Takuno and Gaut 2013). In polyploids, epigenetic changes involve gene silencing, gene activation, genomic imprinting, maternal effects and chromatin conformation changes (Li, ZhiHui et al. 2011). In general, epigenetic changes relate to three different mechanisms: DNA methylation, chromatin modification and small and non-coding RNA changes (Bonasio, Tu et al. 2010).

II.3.1 DNA methylation

DNA methylation is a biochemical process which adds a methyl group to a cytosine or adenine base (Fig. 7A). The cytosine methylation mechanism has been well investigated. The cytosine methylation occurs almost only in the symmetric CG context in mammals (Law and Jacobsen 2010). Plant cytosine methylation occurs in all contexts: the symmetric CG, CHG contexts (H = A, T or C) and the asymmetric CHH context (Henderson and Jacobsen 2007). In Arabidopsis, CG methylation represents about 55% while 23% and 22% were found in the CHG and CHH contexts, respectively (Fig. 7B) (Lister, O'Malley et al. 2008). The cytosine methylation is also variable depending on the DNA sequence type. The gene body methylation occurs typically in CG context (in the transcriptional region of gene) and this DNA methylation context may have implications on many functional and evolutionary processes (Takuno and Gaut 2013). Around 30% of Arabidopsis genes showed CG methylation, (Fig. 7C) (Roudier, Ahmed et al. 2011).

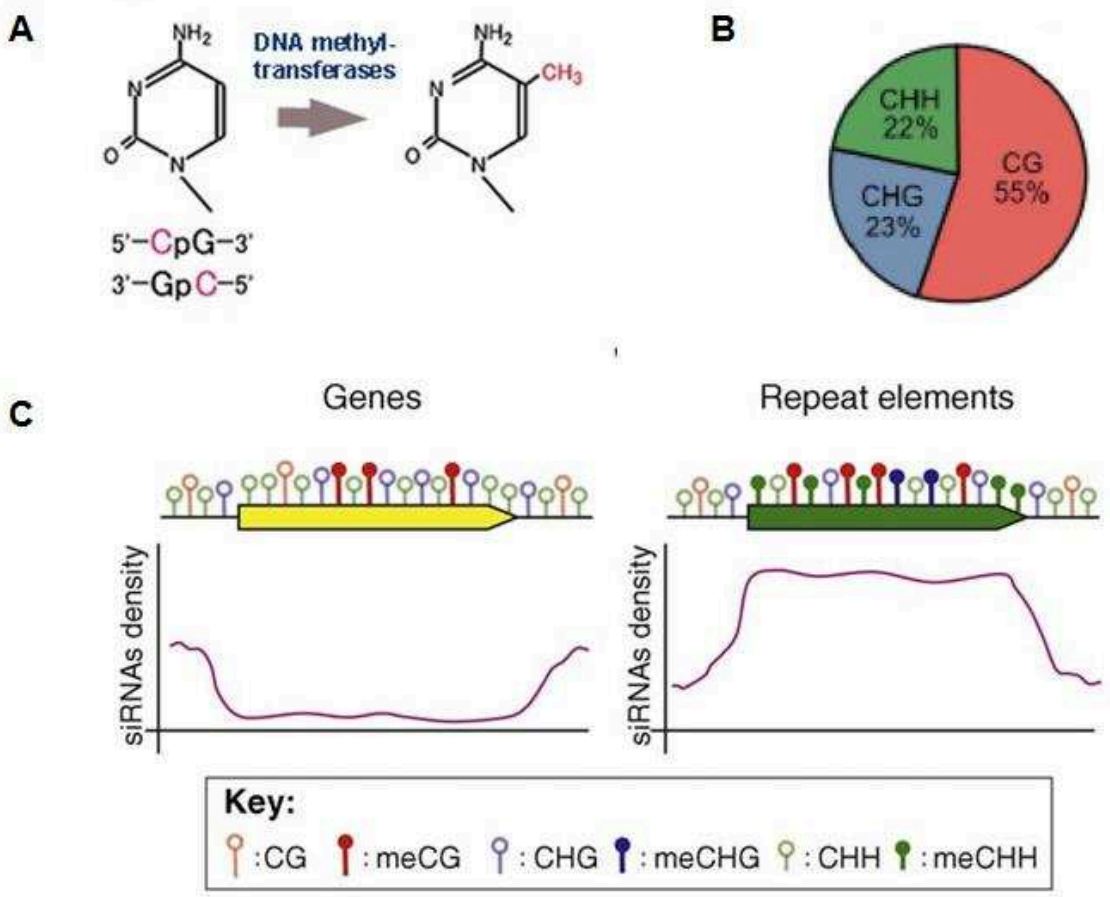


Fig. 7. Cytosine methylation and distribution of methylation context in plant genome. (A) A methyl group added to cytosine by DNA methyl-transferases. (B) The fraction of methyl-cytosines identified in each context of Arabidopsis sequence, where H = A, C, or T (from Lister, O'Malley et al. (2008)). (C) DNA methylation of genes versus repeat elements (Roudier, Ahmed et al. 2011).

In polyploids, changes in DNA methylation as compared to parental species have been reported in Arabidopsis (Madlung, Tyagi et al. 2005, Beaulieu, Jean et al. 2009), wheat (Shaked, Kashkush et al. 2001, Zhao, Zhu et al. 2011) and Brassica (Lukens, Pires et al. 2006, Xu, Zhong et al. 2009). These changes can have direct effects on gene expression as methylation. They can prevent transcription factor binding and lead to changes in chromatin structure that restrict access of transcription factors to the gene promoter (Lim and Song 2012). Changes of DNA methylation have effects on the expression regulation of homoeologous gene copies (Salmon and Ainouche 2010). In the allopolyploid B. napus differences in the cytosine methylation in the 1000 bp upstream of gene was shown to explain 35% of differences in expression between the homoeologous gene copies, suggesting that epigenetic mechanisms may play an important role in the duplicated gene functional diversification (Chalhoub, Denoeud et al. 2014).

II.3.2 Chromatin modifications

Chromatin modifications refer to covalent modifications at different amino-acid positions of histone in nucleosomes (Fig. 8A) (Dupont, Armant et al. 2009). These modification describe biochemical changes to chromatin state, such as alterations in the specific type or placement of histones, modifications of DNA or histones, or changes in the specific proteins or RNAs that associate with a genomic region (Fig. 8) (Munoz-Najar and Sedivy 2011, Eichten, Schmitz et al. 2014). Chromatin exists as a string of nucleosomes connected by DNA. Each nucleosome contains two copies of histone protein (H2A, H2B, H3 and H4) and 146 base pairs (bp) of superhelical DNA wrapped around this histone octomer. Histone proteins are relatively similar in structure and are highly conserved across evolution (McGinty and Tan 2015).

In general, most histone modifications occur at their unstructured, alkaline N-terminal tails, such as acetylation, methylation, phosphorylation, ubiquitylation, sumoylation, ADP ribosylation, proline isomerization and a variety of other post-translational modifications (Eckardt 2010, Kumar, Kumari et al. 2013). Gene activities are various between different chromatin regions (Bannister and Kouzarides 2011) and are different between euchromatin and heterochromatin (Fig. 8B). Heterochromatin regions contain genes that are differentially expressed through development then become silenced (facultative heterochromatin regions) or wholly silenced (constitutive heterochromatin regions as the centromeres and telomeres).

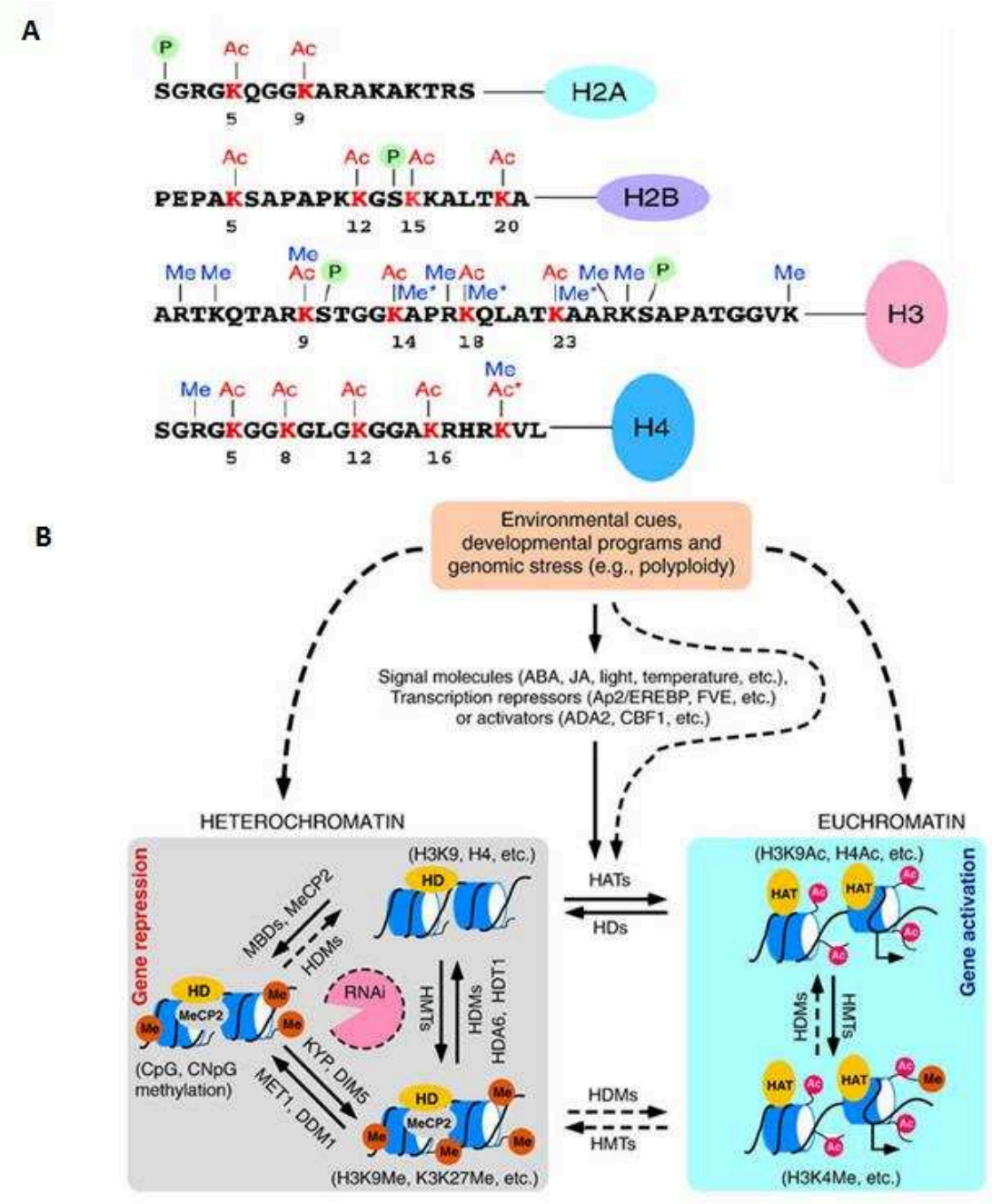


Fig. 8. Histone modifications and their effects on gene activities. (A) Different types of histone modifications, including acetylation (Ac, red), methylation (Me, blue), phosphorylation (P in green circle). In plants, ZmHAT-B has been shown to acetylate specifically lysines K 5 and 12 of H4. (B) Gene activities are effected by environmental cues, developmental programs, and genome composition (e.g., polyploidy) via histone modification and DNA methylation. Dashed arrows indicate that connections may be established with additional experimental data. ABA: abscisic acid; JA: jasmonic acid; HAT: histone acetyltransferase; HD (HDA, HDT): histone deacetylase; HDM: histone demethylase; HMT: histone methyltransferase; KYP: KRYPTONITE, a histone methyltransferase; DIM5: a SET domain protein for histone methyltransferase; MET1: DNA methyltransferase 1; DDM1: a SWI2/ SNF2 chromatin protein affecting DNA methylation; MBD: methyl-binding domain; and MeCP: methyl CpG binding protein (Chen and Tian 2007).

Meanwhile, genes express more actively in the euchromatin regions, where certain are enriched with histone modifications (Bannister and Kouzarides 2011). Comparisons of histone modifications, between heterochromatin and euchromatin, show that most activating DNA regions in euchromatin accompany with the acetylated histone modifications. In contrast, the silenced regions of heterochromatin are all together methylated histones (Fig. 8B). Thus methylations of H3K4 and H3K36 are considered as marks for open chromatin structure (Fuchs, Demidov et al. 2006). On the contrary, switch of heterochromatin from an euchromatin relates to the deacetylated in H3K9 (Fig. 8B) (Grewal and Elgin 2007). Thus histone modifications also have effects on cell divisions and chromosome recombination (Fuchs, Demidov et al. 2006, Jones 2012).

In polyploids, histone modifications and variations are considered to have large effects on gene expression during the growth and development (Fig. 8B) (Chen and Tian 2007, Madlung and Wendel 2013). In Arabidopsis polyploids, analyses of synthetic and natural A. suecica showed that histone methylation at H3K4me2 sites and histone demethylation and acetylation at H3K9 sites result in effects on transcriptional activity (Fig. 8B) (Wang, Tian et al. 2006, Ni, Kim et al. 2009). The data of histone acetylation, methylation in A. thaliana, A. arenosa and allotetraploid plants revealed that H3K9ac and H3K4me3 genes have effects on transcriptional activity and expression variability in response to developmental and environmental changes (Ha, Ng et al. 2011). This study suggests that genome-wide coordinated modifications of histone acetylation and methylation provide a general mechanism for gene expression changes within and between species (Ha, Ng et al. 2011).

II.3.3 RNA interference

RNA interference (RNAi) or post-transcriptional gene silencing (PTGS) is a biological mechanism that inhibits or activates gene expression at the stage of translation or by hindering the transcription of specific genes (Chen 2010, Lam, Chow et al. 2015). This phenomenon has been observed in many species including plants, fungi and animals (Mittal 2004, Tan and Yin 2004, Mittal, Goyal et al. 2012). In plant, RNAi involved two major groups: microRNAs (miRNAs) and small interfering RNAs (siRNAs) (Axtell 2013, Borges and Martienssen 2015, Lam, Chow et al. 2015).

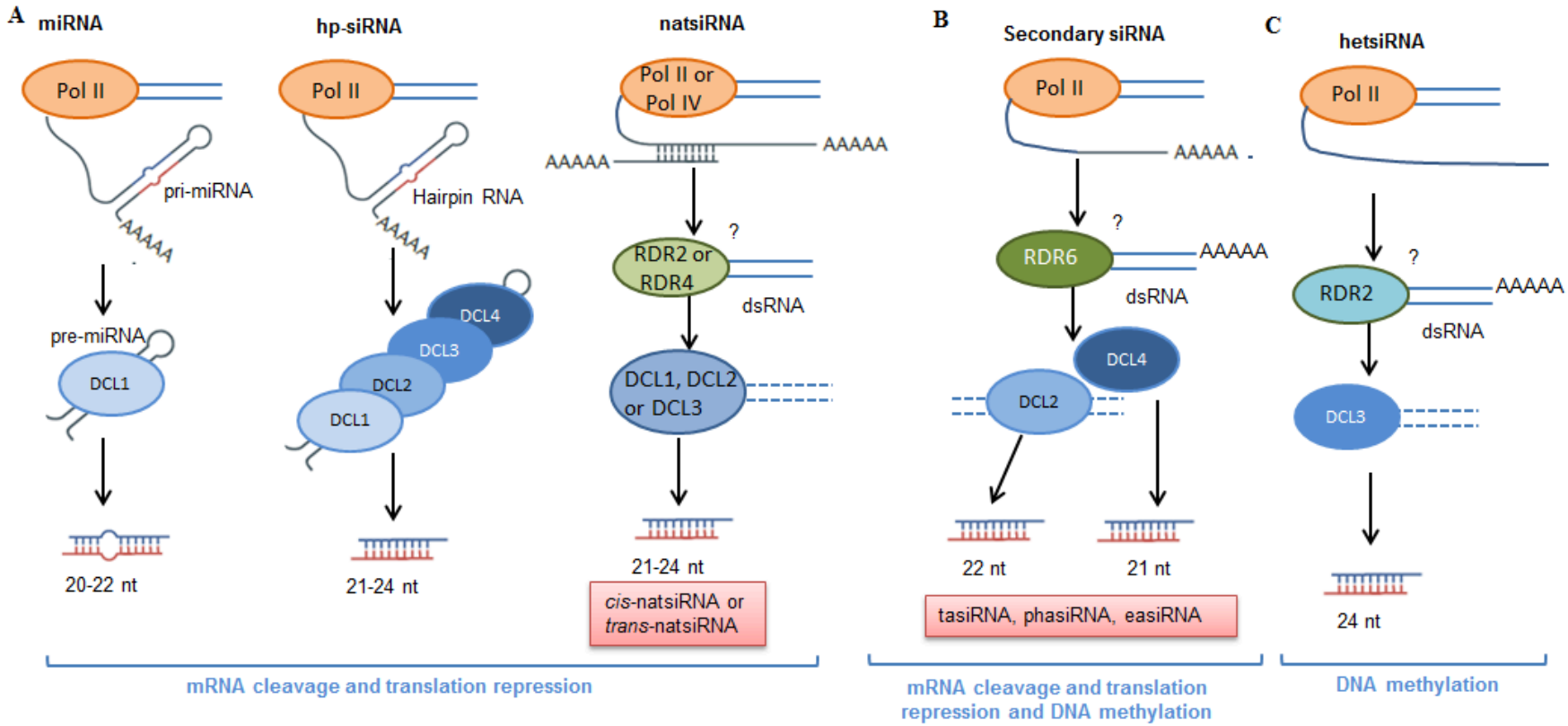


Fig. 9. Types and pathways for biogenesis of endogenous small RNAs in plants. (A) Genes encoding microRNAs (miRNAs; left) are transcribed by RNA polymerase II (Pol II) and fold into hairpin-like structures called primary miRNAs (pri-miRNAs), which are processed by DICER-LIKE 1 (DCL1) into shorter stem–loop structures called precursor miRNAs (pre-miRNAs). Pre-miRNAs are processed again by DCL1 into the mature miRNA duplex. miRNAs are involved in post-transcriptional gene silencing (PTGS) by mediating mRNA cleavage or translational repression. Hairpins-derived siRNAs (hp-siRNAs; middle), might originate from inverted repeats and are originally processed by all DCLs. Natural antisense siRNAs (natsiRNAs; right) are produced from double-stranded RNAs (dsRNAs) originating from overlapping transcription (cis-natsiRNAs), or from highly complementary transcripts originating from different loci (trans-natsiRNAs). (Continuing on next page)

Plant miRNAs comprise a large family of 20-24 bp, non-coding RNAs that have emerged as the key post-transcriptional regulators of gene expression in metazoan animals, plants and protozoa (Ghildiyal and Zamore 2009, Fabian, Sonenberg et al. 2010). Plant miRNAs are evolutionary conserved but their expression patterns often differ between species. miRNA are transcribed by RNA polymerase II into primary miRNAs transcripts that adopt a foldback structure (Fig. 9). They generally function as negative post-transcriptional regulators of gene expression through base pair interactions with mRNAs (Tian, Li et al. 2014). Depending on the degree of sequence complementarity, miRNA target the mRNA degradation or the non-degradative inhibition of translation. Plant miRNA target transcripts encoding proteins involved in diverse physiological processes, among which a set of miRNAs predominantly targets transcription factors (Chen 2010, Lauressergues, Couzigou et al. 2015).

Polyploidy effects on miRNAs have been reported in many species. In synthetic allopolyploid Arabidopsis, the rapid changes of repeat- and transposon-associated siRNAs were found in first generation, but were stably maintained in the seventh generation. Their miRNA and ta-siRNA (trans-acting small interfering RNAs) sequences were conserved, but their expression patterns were highly variable between the allotetraploids and their progenitors (Ha, Lu et al. 2009, Tian, Li et al. 2014). In wheat, the expression of miRNA in F1 interspecific hybrids did not deviate from the the average of their parental species or mid-parent value (MPVs), but a number of deviations was found in derived allopolyploids (Kenan-Eichler, Leshkowitz et al. 2011). This study also suggests that deregulation of small RNAs may stimulate TE activation in interspecific hybrids and allopolyploids (Kenan-Eichler, Leshkowitz et al. 2011).

In contrast with miRNAs, small interfering RNAs (siRNAs) may be endogenous or come from the exogenous sources such as viral infection. These siRNAs are derived from aberrant double-stranded RNAs which require the activities of RNA-dependent RNA polymerases (RDRs) resulting into two sizes of siRNAs types of 21-nt and 24-nt (Fig. 9) (Ghildiyal and Zamore 2009, Lam, Chow et al. 2015). The 24-nt endogenous siRNAs in plants are various and quite ubiquitous as demonstrated by high-throughput sequencing of small RNAs (Henderson and Jacobsen 2007, Chen 2010). The relationship between miRNAs and siRNAs is expressed through secondary siRNAs. In plants, secondary siRNAs, derived from miRNA-duplex structure, have been reported in certain studies (Bartel 2004, Pontes and Pikaard 2008, Manavella, Koenig et al. 2012). Trans-acting siRNAs regulate target gene in

Fig. 9. (Continue) The biogenesis and function of natsiRNAs is still largely unclear. (B) The precursors of secondary siRNAs are transcribed by Pol II and may originate from non-coding loci, protein-coding genes and transposable elements. These transcripts are converted into dsRNA by RNA-DEPENDENT RNA POLYMERASE 6 (RDR6) and processed by DCL2 and DCL4 to produce siRNAs of 22 or 21 nucleotides (nt) in length, respectively. Secondary siRNAs are mostly involved in PTGS, but they can also initiate RNA-directed DNA methylation (RdDM) at specific loci. They are subdivided into trans-acting siRNAs (tasiRNAs) phased siRNA (phasiRNAs) and epigenetically activated siRNAs (easiRNAs). (C) Heterochromatic siRNAs (hetsiRNAs) are derived from transposable elements and repeats that are preferentially located at pericentromeric chromatin. Their biogenesis requires Pol IV transcription and the synthesis of dsRNA by RDR2, which is subsequently processed into 24-nt-long siRNAs by DCL3. These small RNAs are involved in maintaining RdDM-mediated TGS. From Borges and Martienssen (2015).

trans like miRNAs and siRNAs do in PTGS in the same manner as miRNAs (Fig. 9). Thus they play significant roles in transitive RNA silencing and spread of RNA silencing (Vazquez and Hohn 2013). Therefore these were considered as a novel class of small RNA regulators (Chen 2010).

Similar to miRNAs, siRNAs regulate important processes including mRNA inhibition in relation to stress responses. In addition, they are also involved in the chromatin modification and transposon silencing (Pontes and Pikaard 2008, Vazquez and Hohn 2013, Lam, Chow et al. 2015). In polyploids, siRNAs seem not to contribute in non-additive gene expression between allopolyploids and their progenitors. In contrast, miRNAs and ta-siRNAs expressed variably when compared allopolyploids to their parents (Ha, Lu et al. 2009). In other polyploids such as Brassica, the ta-siRNAs were shown to be related to the self-incompatibility (Tarutani, Shiba et al. 2010), whereas the small RNA profile in B. napus has been recognized in regulating gene expression (Zhao, Wang et al. 2012).

Because of its extensive influences in many aspects of gene expression regulation, genomic changes and epigenetic variations, small RNAs become more and more important to characterize responses to polyploidy (Ha, Lu et al. 2009, Martienssen 2010, Groszmann, Greaves et al. 2011, Ng, Zhang et al. 2011).

II.4 Reprogramming of gene expression

Gene expression in polyploids can be equal to the sum or to average of parental gene expression (MPV), known as additive gene expression. It can also deviates from MPV, known as non-additive (up- or down-regulation) expression (reviewed in Soltis, Visger et al. (2014). Generally, both additive and non-additive gene expression are widely established at different proportion in diverse allopolyploids (Yoo, Liu et al. 2014).

Additive expression was found to be largely established in synthetic allohexaploid wheat where more than 93% of expressed transcripts were equal to MPVs measured from a mixture of parental RNA (Chagué, Just et al. 2010, Chelaifa, Chagué et al. 2013). In other species, the proportion of additively expressed genes varies from 65-95% in Arabidopsis allotetraploids (Wang, Tian et al. 2006), 30-70 % in cotton allotetraploids (Flagel, Udall et al. 2008), 40-73% in Senecio interspecific hybrids and allohexaploids (Hegarty, Barker et al. 2006). It is suggested that additive expression for a large number of genes may provide a

molecular basis for dosage balance and compensation of functionally redundant genes, leading to developmental stability in new allopolyploids (Jackson and Chen 2010).

Nevertheless, different allopolyploid species show a variable proportion of non-additively expressed genes. In synthetic Arabidopsis, it was shown that expression of 5.2-5.6% of genes was not additive (Wang, Tian et al. 2006). This frequency is about 1–6.1% in cotton (Adams, Percifield et al. 2004, Rapp, Udall et al. 2009). In Spartina anglica, deviation from parental additivity was most important following hybridization and was accompanied by maternal expression dominance (Chelaifa, Monnier et al. 2010).

The fate of Homoeologous duplicated genes.

The two copies of homoeologous duplicated genes can follow diverse evolutionary fates (Fig. 10).

Functional redundancy means that two or more gene copies are performing the same function after the duplication and that inactivation of one of these has little or no effect on the function or biological phenotype Fig. 10 (Nowak, Boerlijst et al. 1997).

Non-functionalization or pseudogenization consists in gene copies that became nonfunctional by random genetic drift and null mutation (Force, Lynch et al. 1999).

Sub-functionalization consists in partitioning of ancestral functions/expression patterns between duplicated genes, leading to retention of both genes because loss of either copies would be lethal (Doyle, Flagel et al. 2008). Sub-functionalization can occur by either qualitative or quantitative routes. In case of the qualitative sub-functionalization, one duplicate copy goes to fixation for a complete loss of sub-function mutation and the second locus subsequently acquires a null mutation for a different sub-function. In contrast, quantitative sub-functionalization results from the fixation of reduction of expression mutations in both duplicates (Force, Lynch et al. 1999).

There is **neo-functionalization** when a copy of the duplicated gene is selected to perform a novel function whereas the other copy maintains the ancestral function (Des Marais and Rausher 2008, Teshima and Innan 2008). This term has been proposed first time by Ohno (1970) as the hypothesis of the fate after genome duplication, one daughter gene retains the ancestral function while the other can gain novel function. Studying on the anthocyanin biosynthesis pathway gene dihydroflavonol- 4-reductase (DFR) in morning glories (Ipomea), the authors reported that this gene is constrained from improving either

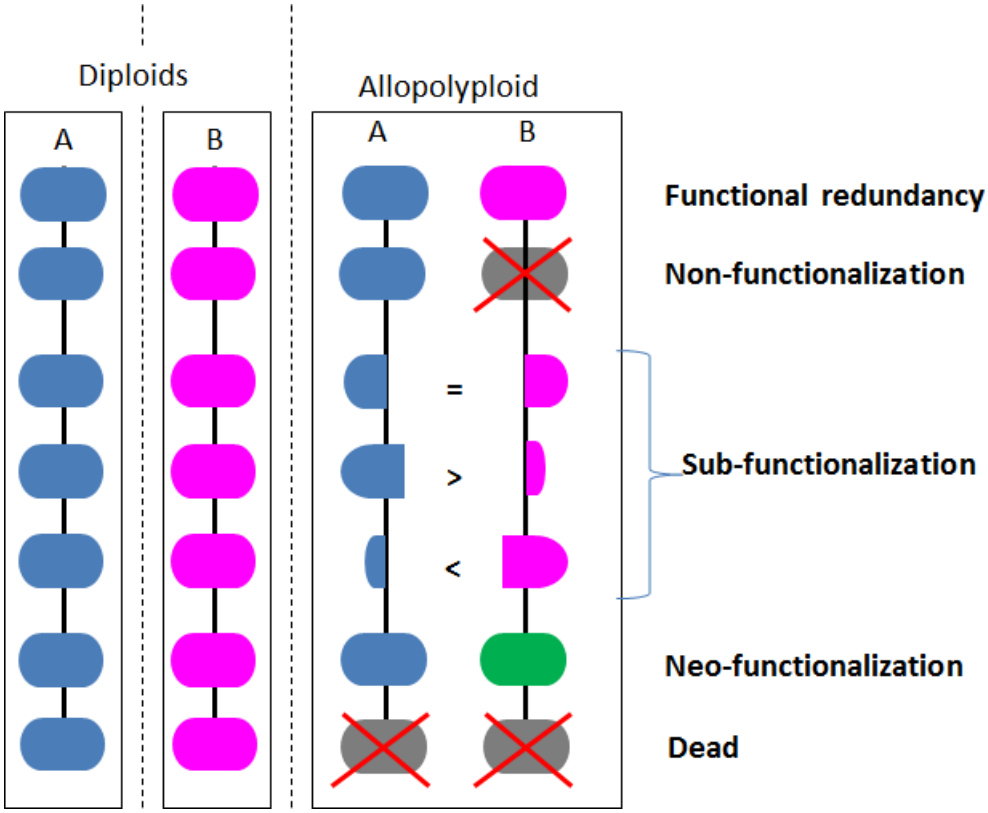


Fig. 10. **Schematic representation of the fate of duplicate genes.** Functional redundancy refers to the cases where the two gene copies are performing the same function after the duplication. Non-functionalization consists in the deletion of one copy. Sub-functionalization indicates process where two copies partition the ancestral function, either equally or bias via one of two ancestral copies. In this case the ancestral function could not be conferred completely by one single copy. Neo-functionalization occurs when a new function emerged in a duplicated copy, the original function continue to be maintain by the other copy (Ohno 1970, Force, Lynch et al. 1999, Doyle, Flagel et al. 2008, Zhang, Belcram et al. 2011, Chalabi 2014).

novel or ancestral function because of detrimental pleiotropic effects on the other function. After duplication, one copy is free to improve novel function, whereas the other is selected to improve ancestral function (Des Marais and Rausher 2008).

Recently, Barker, Baute et al. (2012) discussed that the sub-functionalization and neo-functionalization in plant genome may be classified into two different models: the duplication-degeneration-complementation (DDC) model and escape from adaptive conflict (EAC) model. The DDC model predicts that degenerative mutations in regulatory elements can increase rather than reduce the probability of duplicate gene preservation and the usual mechanism of duplicate gene preservation is the partitioning of ancestral functions rather than the evolution of new functions. Thus, the **neo-functionalization** process may occur by completely neutral process Dykhuizen-Hartl effect (Fig. 10) (Dykhuizen and Hartl 1980), an adaptive process (positive Darwinian selection) or by some combination of the two processes (Barker et al. 2012). In neutral process, a new function arises as a by-product of the accumulation of chance mutations in the cis-regulatory or protein coding regions. In fact, the neutral mutations are prevalent in the population by genetic drift; they express in the phenotypes which have not undergone the Darwinian selection. The empirical proofs were elucidated in E. coli, where the distribution of allele frequencies in natural populations is in accordance with the expectations of selective neutrality. (Dykhuizen and Hartl 1983, Kimura 1983). Neo-functionalization can occur following the adaptive process, where a new function is formed from the selection of one copy conferring a more advantage than the other copy. Several examples of neo-functionalization in plants have been reported. In Arabidopsis, based on the analysis of MEDEA, a SET- domain Polycom group protein, Spillane indicated that after duplication, MEA underwent positive Darwinian selection consistent with neo-functionalization and the parental conflict theory (Spillane, MacDougall et al. 2000).

A number of studies indicated that both neo-functionalization and sub-functionalization contribute to the evolution of new genes in plant genomes (Barker, Baute et al. 2012). Functional and phenotypic contribution of the three homoeologous copies of the important Q domestication gene in wheat have been studied in details offering a good example of the fate of duplicated genes (Zhang, Belcram et al. 2011). A single base mutation in the mi-RNA-172 target regulating site of the 5Aq copy had led its mis-regulation and consequently higher expression leading to “hyperfunctionalization” (5AQ allele). The 5Bq copy underwent pseudogenization, but remains transcriptionally active and contribution to the domestication phenotype, whereas the 5dq homoeologous copy has been sub-

functionalized. The analysis of gene expression in eight plant species for which genome sequence is available had shown important genome fractionation as well as bias in the expression of remaining duplicated gene copies (Garsmeur, Schnable et al. (2014). More interestingly the authors suggested that those paleopolyploid species showing dominance in gene fractionation and expression were derived by allopolyploidy whereas those showing equivalence were most likely derived from autopolyploidy (Garsmeur, Schnable et al. (2014).

III. CYTOLOGICAL CHARACTERISTICS, RECOMBINATION AND REPRODUCTIVE ISOLATION OF POLYPLOIDS

III.1 Regulation of cell division in polyploid plants

III.1.1 Mitotic and meiotic defects

During mitosis, the mitotic defects can result from errors in chromosomes segregation leading to the formation of aneuploids. In yeast triploids and tetraploids, chromosome loss could be observed with the syntelic[2] or merotelic[3] formation because of arising spindle irregularities specific defects (such as transient multipolarity and incomplete spindle pole separation) in the mechanism of chromosome segregation (Fig. 11A) (Storchova, Breneman et al. 2006, Hufton and Panopoulou 2009). Numerous evidences demonstrated that mitotic spindles with initial geometric defects cause high rates of chromosome mis-segregation and aneuploidy (Silkworth and Cimini 2012). In yeast, the changes of centromeric core chromatin domain and the flanking pericentric heterochromatin domain accompanied the abnormal phenotypes (Gaither, Merrett et al. 2014).

Meiotic defects have been observed in both autopolyploids and allopolyploids. Autopolyploids contain more than two homologous chromosome sets and have the polysomic inheritance. Thus, their meiosis may have obstacles because of multivalent formations which are more frequent in autopolyploids than in allopolyploids (Fig. 11B2) (Lim, Soltis et al. 2008, Hufton and Panopoulou 2009, Parisod, Holderegger et al. 2010, Lloyd and Bomblies 2016). A. thaliana and Guizotia abyssinica autopolyploids showed low frequencies of trivalents and quadrivalents formed during meiosis which result in formation of aneuploids in their progenies together with reduced fertility. The formations of multivalents in meiosis of autopolyploids also cause rearrangements of chromosomes (Weiss and Maluszynska 2000, Dagne 2001). However, the most frequent type of association was bivalent in V. corymbosum suggesting that stabilization is the general trend of this autopolyploid genome (Qu, Hancock et al. 1998). In Arabidopsis autotetraploids, several genes were shown to be involved in chromosome synapsis, cohesion, and homologous recombination and play an important role in polyploidy stabilization (Hollister et al., 2012).

[2] Describing a chromosome orientation in which both sister kinetochores attach to microtubules extending from one spindle pole
[3] Describing the attachment of one kinetochore to both mitotic spindle poles

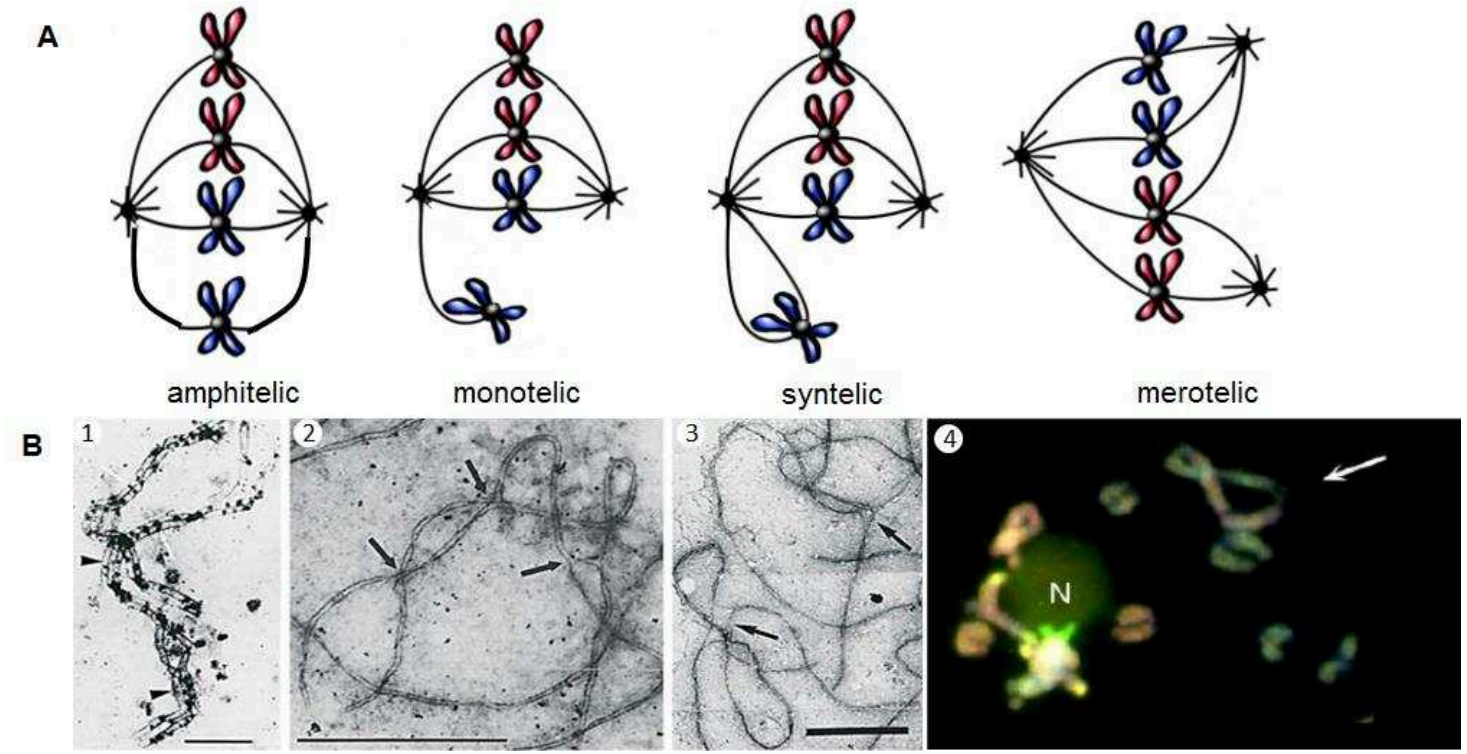


Fig. 11. Mechanisms related to mitotic and meiotic defects in polyploid plants**.** (A) The different types of attachments between chromosomes with microtubules during the early stages of mitosis. Amphitelic attachment is the only type of attachment that ensures correct chromosome segregation. Errors in chromosome segregation may occur when chromosomes establish erroneous attachments. Monotelic attachment occurs when one sister kinetochore is attached to microtubules and the other sister is unattached. Syntelic attachment advent when the two sister kinetochores bind microtubules from the same spindle pole, and merotelic attachment, in which a single kinetochore binds microtubules from both spindle poles instead of just one. (B) Meiotic defects in polyploid is the result of the synaptic and chiasmatic chromosome associations which lead to the formation of multivalents: (1) A Multiple synaptonemal complex in late-zygotene from potato male meiocytes, showing extensive triple and some quadruple synapsis (arrowheads). (2) Synaptic multivalents (arrows); surfacespread prophase I nuclei in autopolyploid Crepis capillaris and those in allotetraploid Scilla autumnalis (3). A quadrivalent with two chromosomes of Tragopogon dubius origin and two of T. porrifolius (arrow).
Adapted from Lim et al., 2008; Hufton and Panopoulou 2009; Silkworth and Cimini 2012; and Grandont et al., 2013.

The stable meiosis in allopolyploids requires a bivalent formation between homologous chromosomes and no pairing between homoeologous chromosomes. However, this ideal situation is not the case of most allopolyploids which homoeologous chromosomes can also pair during meiosis leading to aneuploid progenies with increased sterility (Ali, Lysak et al. 2004, Mestiri, Chagué et al. 2010, Grandont, Jenczewski et al. 2013). These meiotic deviations usually occur in allopolyploids through multiple associations in prophase I and metaphase I, which may happen either before or after synaptonemal complex formation. Synaptonemal complex may have a role in the adjustment during the association between homologous and homoeologous chromosomes (Fig. 11B) (Grandont, Jenczewski et al. 2013).

It has been shown in wheat and in B. napus that bivalent formation at meiosis are controlled by specific loci. In wheat allopolyploids, Ph1 (paring homoelogous 1) locus on chromosome 5B, restricts pairing to homologous chromosomes during meiosis (Griffiths, Sharp et al. 2006). The PrBn locus in Brassica napus allopolyploid has effects on recombination between homoeologous chromosomes (Jenczewski, Eber et al. 2003, Nicolas, Leflon et al. 2009).

III.1.2 Chromosome recombination during meiosis

The chromosome recombination leads to inter-chromosome exchanges. Recombination between homologous and homoeologous chromosomes which result in DNA changes such as deletions, duplications and homoelogous translocations have been reported (Song, Lu et al. 1995, Dagne 2001, Comai, Tyagi et al. 2003, Jenczewski, Eber et al. 2003, Ali, Lysak et al. 2004, Lim, Soltis et al. 2008, Chester, Leitch et al. 2010).

Mechanisms of these changes relate to chromosome recombination during meiosis such as reciprocal or nonreciprocal exchanges between homologous chromatids (Hufton and Panopoulou 2009). Normally, the recombination occurs when pairing between homologous chromosomes at prophase I, often from zygotene to diplotene, and results in equal segregation (Jones, Armstrong et al. 2003). In allopolyploid genomes, the pairing between homoeologous chromosomes may happens. Crossover or non-crossover recombination may occur under the regulation of two mechanisms: Double Strand Break Repairs (DSBR) and Synthesis Dependent Strand Annealing (SDSA) (Youds and Boulton 2011). When pairing between homoeologous chromosome arises, crossovers result in homoeologous chromosome segment exchanges (Udall, Quijada et al. 2005, Gaeta and Chris Pires 2010, Salmon, Flagel

et al. 2010, Szadkowski, Eber et al. 2010). It has been demonstrated that many changes in polyploid genomes are derived from chromosome recombination including a large set of exchanges as the sequence deletion leading to gene loss and smaller scale duplication events (Tayale and Parisod 2013) In Brassica allopolyploids, there was evidence that many genetic changes are the consequence of homoeologous recombination (Chalhoub, Denoeud et al. 2014). This processes can generate novel gene combinations and phenotypes, but may also destabilize the karyotype and lead to aberrant meiotic behavior and reduced fertility (Gaeta and Chris Pires 2010). Studies with natural polyploid population of Tragopogon miscellus Buggs, Doust et al. (2009) indicated that there were no homoeologous gene losses in the first generation, but evidences of these processes have been found in more advanced generations. Moreover, they were not fixed within natural populations and did not form a predictable pattern among populations (Buggs, Doust et al. 2009).

In conclusion, chromosome recombination contributes to genome rearrangement in polyploids. These mechanisms including DNA sequence loss, duplication or translocation have been observed during evolution of natural or synthetic polyploid populations.

III.2 Genetic control of homologous and homoeologous pairing in allopolyploids

At meiosis, homologous chromosome must recognize and pair with each other in prophase I to ensure balanced segregation. This process starts at leptotene when telomeres attach randomly along the nuclear envelope, then pass through the formation of bouquet at zygotene, during pachytene, the homologue alignments are formed along the entire length to produce the mature bivalents with fully synapsed chromosomes; the recombination between the homologous chromosomes occurs during zygotene and pachytene (Fig. 12) (Moore and Shaw 2009, Tsai and McKee 2011).

Unlike diploids, whose genome contains two homologous chromosome sets, allopolyploid genomes contain homoeologous chromosomes derived from related parental species. In order to have a corrected segregation in meiosis and maintain a stable allopolyploid genome, the homologues must be apparently distinguished from their homoeologues during pairing. Mechanism whereby the homologues and homoeologues are distinguished in polyploids is regulated by genetic factors such as PrBn (Pairing regulator in B. napus) in B. napus and Ph1 (Pairing homoeologous 1) in wheat.

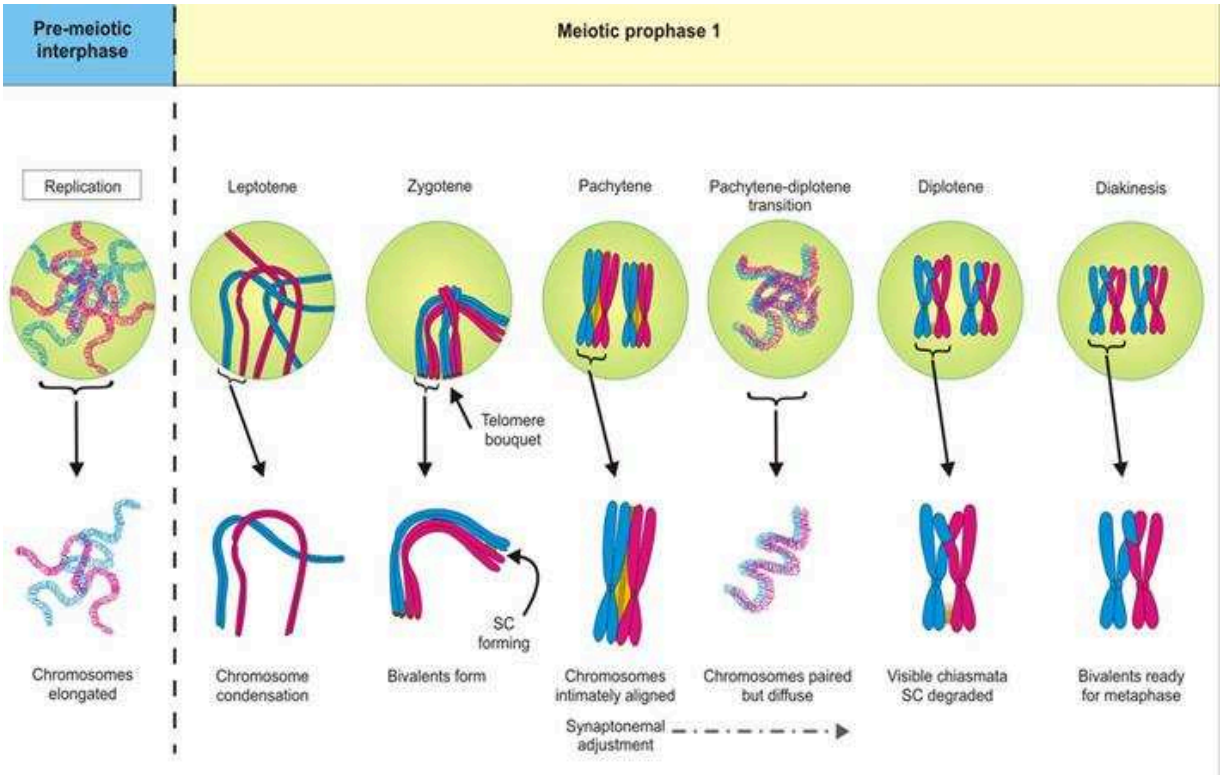


Fig. 12. Homologous chromosomes recognize and pair with each other in prophase I. Pre-meiotic interphase unpaired homologous chromosomes are distributed randomly within the nucleus, chromosomes are elongated and were replicated. In meiotic prophase I, telomeres have attached randomly along the nuclear envelope at leptotene. The telomere bouquet forms at zygotene after homologous chromosome alignment. At pachytene, high levels of homologous aligment are achieved along the entire lentgh to produce a mature bivalent with fully synapsed chromosome. Homologous recombination pairing occurs during zygotene and pachytene. The synaptonemal complex is disassembled at diplotene, when recombination is completed. Chromosomes then condense further during the diakinesis (Moore and Shaw 2009, Tsai and McKee 2011).

III.2.1. The Pairing regulator in B. napus (PrBn) gene

In B. napus (AACC, 2n = 38), a recent amphitetraploid derived from natural hybridizations between ancestors of B. rapa (AA, 2n = 20) and B. oleracea (CC, 2n = 18), a locus called PrBn has reduces meiotic homoeologous recombination by limiting the homoeologous pairing in B. napus haploids (AC, 19 chromosomes) (Jenczewski, Eber et al. 2003, Nicolas, Leflon et al. 2009). PrBn was mapped on linkage group C9 (Liu, Adamczyk et al. 2006) and the variation in crossover frequency among B. napus accessions, representing a range of genetic and geographic origins, roughly correlates with the multiple origins of B. napus and PrBn diversity (Cifuentes, Eber et al. 2010).

III.2.2 The pairing homologous gene (Ph1)

In the wheat allohexaploid, T. aestivum (AABBDD, 2n=42) each chromosome pair has two other homoeologous chromosome pairs. However, the meiosis of hexaploid wheat behaves as that of diploid species, with high rate of homologous pairing and bivalents formation (Fig. 13). It has been found that this homologous pairing in allohexaploid wheat is essentially regulated by a major locus that has been called pairing homologous gene (Ph1), located on the long arm of chromosomes 5B (Riley and Chapman 1958). By deleting the Ph1 locus, a high degree of pairing between homoeologous chromosomes and hence multivalent formation at metaphase 1 have been observed (Fig. 13) (Sears 1976).

The structure of Ph1 locus in wheat has been defined to a single wheat chromosome region that contains a cluster of cyclin-dependent kinase complex (Cdk2-like genes) related to Cdk2 of yeast humans and mouse which has been disrupted by the insertion of a segment of subtelomeric heterochromatin (Griffiths, Sharp et al. 2006). The cell cycle regulator Cdk2 is known as the factor which control meiotic progression, expression of meiotic genes, meiotic DSB formation and chromatin structure (Yousafzai, Al-Kaff et al. 2010). However, recently, Bhullar, Nagarajan et al. (2014) had claimed that Cdk2-like genes are not the candidate genes of Ph1 and the authors have identified another candidate genes (C-Ph1 gene). Having ortholog of gene At5g25610 in Arabidposis and Bradi4g33300 of Brachypodium, C-Ph1 encodes a BURP domain containing protein with a putative function in the dehydration stress response (Bhullar, Nagarajan et al. 2014). C-Ph1 may be responsible for the Ph1 effect which knockdown causes multivalent formation, homeologous pairing and chromosome

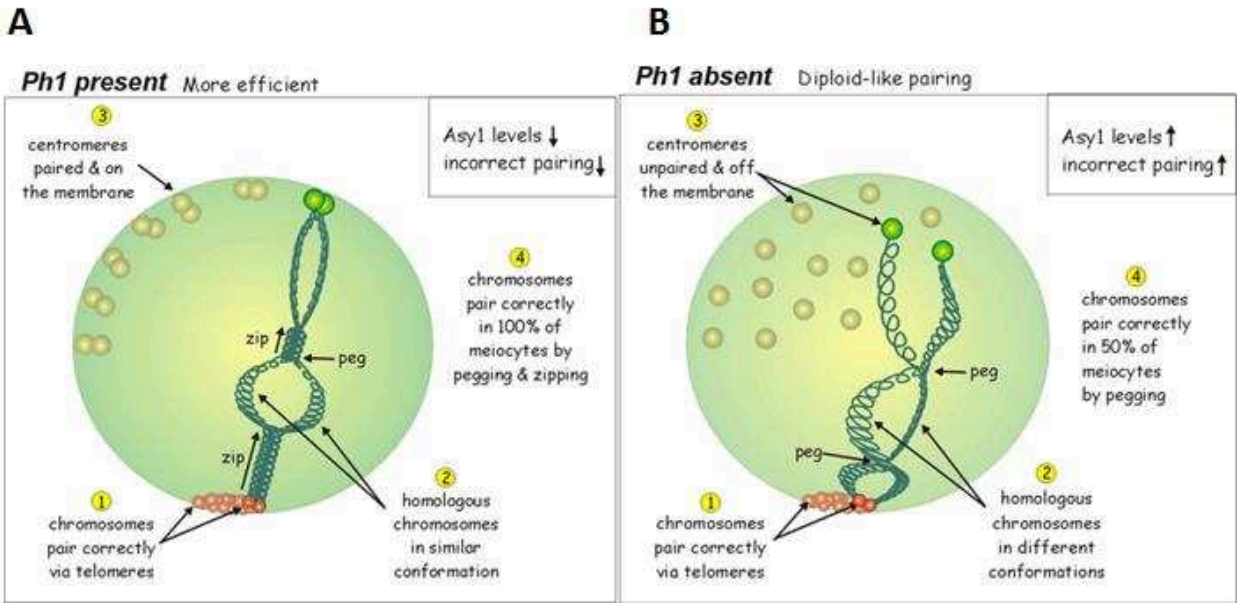


Fig. 13. Homologous pairing during wheat meiosis. (A) In the wild-type wheat, chromosomes pair correctly at the telomeres, homologs are in similar conformations and centromeres are paired and located at the nuclear envelope. Pegging and zipping mechanism are proposed to regulate chromosome pairing when Ph1 presents. (B) In the Ph1 deletion mutant, the absence of Ph1 leads to chromosomes pairing correctly in 50% of meiocytes at a few interstitial sites by pegging. The homologs are in different conformations, centromeres are unpaired and internalised from the nuclear envelope

(From Moore and Shaw (2009).

misalignment at metaphase I. This gene has different copies on chromosome 5A, 5B, and 5D. The 5A copy is truncated, while the 5D copy is expressed in the earliest stages of meiosis. The 5B copy, likely corresponding to the Ph1 allele, is expressed only in metaphase I and has an insertion relative to the other homoeologs, suggesting that it could have evolved a novel function (Bhullar, Nagarajan et al. 2014).

Evidences indicate that homologs pair via their telomeres and centromeres, and Ph1 has effects on both telomeres and centromeres. The wheat centromeres associate and Ph1 controls specificity of centromere pairing, suppressing non-homologous pairing (Martinez-Perez, Shaw et al. 2001). Moreover, a number of studies demonstrated that the telomeric regions are important for homologue recognition and synapsis in wheat (Lukaszewski 1997, Colas, Shaw et al. 2008). Effects of presence/absence of Ph1 locus have been summarized in four main points following the proposed pegging and zipping mechanism of the correctly paired chromosomes (Fig. 13). Firstly, chromosomes pair correctly via telomere in both cases Ph1 present or absent. Then homologous chromosomes combine in similar conformation in wild-type, whereas the deviation way was observed in the ph1 mutant. In case of Ph1 absence, only 50% chromosomes pair correctly by pegging (Moore and Shaw 2009).

Furthermore, Ph1 plays an important role on karyotype stability. Studies of synthetic allohexaploid wheat indicated that the multivalent formation was more frequent in the first S0 generation and decreased ion subsequent generations (Mestiri, Chagué et al. 2010). More effects of Ph1 have been considered as the effect on chromosome pairing and synapsis; the effect on chromosome recombination; the effect on chromosome organization (Martinez-Perez, Shaw et al. 2001).

III.3 Reproductive isolation in polyploids

Reproductive isolation or hybridization barriers are both an indicator of speciation and a mechanism for maintaining species identity. There are pre-zygotic and post-zygotic mechanisms for reproductive isolation. Studies on plant reproductive isolation indicated that pre-zygotic isolation is much stronger than post-zygotic (Widmer, Lexer et al. 2009). The post-zygotic isolation comprises both intrinsic and extrinsic types. The intrinsic post-zygotic isolation is related to genetic divergence and arose from the polymorphism within species which plays certain roles in speciation (Rieseberg and Willis 2007, Sweigart and Willis 2012). While the pre-zygotic barriers can be easily detected by the difference in adaptation

and ecological condition or the selection for copulation, post-zygotic barriers still remain a mystery because of the production of dead or sterile hybrids cannot be favored by natural selection.

Post-zygotic isolation is known as the hybrid unviability and sterility and the latter ecological and behavioral sterility. Studies show that the strength of post-zygotic isolation increases with increasing of genetic distance among taxa (Widmer, Lexer et al. 2009). This may indicate the importance of post-zygotic isolation in the speciation and the evolution process.

In plant, the most common form of post-zygotic is hybrid sterility. Hybrid plants can grow viably at the vegetative stage but fail to produce fertile and viable pollen or embryo-sac during reproductive development leading to reducing seed setting (Ouyang and Zhang 2013). Hybrid sterility has been found in different plants: in Arabidopsis, about 2% of intraspecific crosses give progeny that suffer from hybrid necrosis (Bomblies, Lempe et al. 2007); in Mimulus, the hybrid of the crosses between two close inbred lines of two ecological sister species M. guttatus and M. nasutus expressed high level of sterility effect (Sweigart, Fishman et al. 2006). Infertility phenotype also has been indicated in Solanum interspecific hybrids, resulting from cross between S. lycopersicum x S. habrochaites (Moyle and Nakazato 2008).

Hybrid sterility may have multiple causes, but the most common and well-studied is the hybrid incompatible loci which follow "Bateson-Dobzhansky-Muller" (BDM) model (Fig. 14). Genetic incompatibility via BDM model contributing to the sterility of hybrids has been documented in many genera such as Oryza, Arabidopsis, Solanum and Mimulus (Sweigart, Fishman et al. 2006, Bomblies, Lempe et al. 2007, Moyle and Nakazato 2008, Yamagata, Yamamoto et al. 2010)**.**

III.3.1 "Bateson-Dobzhansky-Muller" model

Bateson-Dobzhansky-Muller (BDM) is a model of the evolution of genetic incompatibility. The theory was first described by William Bateson in 1909, then independently described by Dobzhansky in 1934, and later elaborated by Herman Muller in 1942 (Bomblies, Lempe et al. 2007). BDM model suggests that the hybrid incompatibility should be the consequence of the combination of at least two mutated loci together, which function perfectly well in the context of its native genetic background, when they were

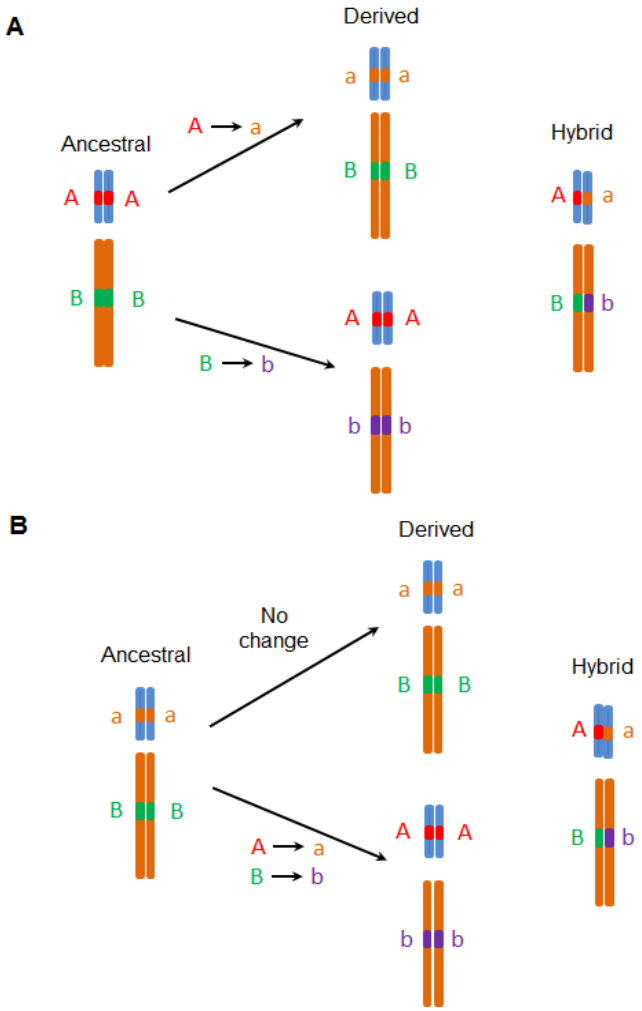


Fig. 14. The Bateson–Dobzhansky–Muller (BDM) model. The BDM model of the evolution of hybrid incompatibility (HI) posits that HI is due to changes in at least two genetic loci. (A) In the ancestral population, individuals are all AABB. The ancestral population is split into two geographically-separated populations. In the two derived populations, evolutionary pressures lead to two different results: in the top derived population, allele A changes to allele a resulting in the population being aaBB; in the bottom derived population, allele B changes to allele b resulting in the population being AAbb. If individuals from the two derived populations are crossed, their hybrids would be AaBb. HI results from the combination of the a and b alleles in the hybrids. (B) Same as above, but the ancestral population begins as aaBB. No change occurs in the top derived population, but the bottom population changes first from a to A and then from B to b. This yields derived populations that are aaBB and AAbb. The hybrids would possess the combination of the a and b alleles, and would thus have HI (Johnson 2010, Cutter 2012, Seehausen, Butlin et al. 2014).

combined in a new genetic background it might be functionally incompatible (Fig. 14) (Fishman and Willis 2001, Sweigart and Willis 2012).The reason is that the mutated allele of a single locus must be heterozygous with its original allele at the first generation that is adaptive or neutral, therefore not eliminated by natural selection. In contrast, in the new variant the mutation allele is incompatible with the original copy, and then individuals that host the mutation allele in the incompatible heterozygous format will simply not survive.

The key insight of BDM model is the interaction of two or more mutation differences between species, which may be have relationship and play crucial roles in the hybridization and polyploidization. Although the BDM model was proposed more than 60 years ago, the identification and characterization of incompatible genes was achieved recently. In plants, a number of studies reported that BDM has effects directly on hybrid sterility and then on several aspects in speciation, whereby the BDM factors may act under simple or complex genetic control (Rieseberg and Willis 2007).

III.3.2 Hybrid incompatibility genes identified in plants

BDM genes (or loci) that cause hybrid lethality or sterility seem to be more documented in yeast and animal than in plants. In yeast a number of genes (or loci) have been listed such as: AEP2, OLI1, MRS1, COX1 and AIM22. In animals, almost BDM genes in Drosophila also are well defined like: Lhr, Hmr, Zhr, OdsH, Ovd, Nup96, Nup160 and Prdm9. Plant BDM genes (or loci) have been detected quite fully in rice as DPL1/DPL2; S27/S28; SaF/SaM, and S5, while there is only one couple of genes in Arabidopsis (HPA1/HPA2) has been identified as BDM gene (Maheshwari and Barbash 2011, Ouyang and Zhang 2013).

IV THE BRACHYPODIUM MODEL

IV.1 Evolutionary relationships and taxonomic position of the Brachypodium genus

IV.1.1 Evolutionary relationships

The genus Brachypodium comprises approximately 20 taxa (18 recognized taxa and two not-yet-defined cytotypes) with different ploidy levels, genome sizes and variable basic chromosome numbers (x) of 5, 8, 9 and 10 (Fig. 15) (Robertson 1981, Catalan, López-Álvarez et al. 2016).

According to the most recent taxonomy, three species are annual (B. distachyon, B. stacei and B. hybridum) and 15 are perennials (Fig. 15). Comparative cytogenetic and molecular analyses showed that B. distachyon's large chromosomes likely originated via descending dysploidy, acting as fusions of smaller chromosomes of a putative ancestral Brachypodium species, very close to B. stacei (2n=20) or B. mexicanum (2n=40) (Betekhtin, Jenkins et al. 2014). The allotetraploid B. hybridum was derived from hybridization between B. distachyon and B. stacei (Catalan, Chalhoub et al. 2014).

Few perennial Brachypodium species show a large Eurasian (B. sylvaticum, B. rupestre, and B. pinnatum), or Mediterranean (B. retusum) distribution, and others have disjunctive distributions in South Africa, SE Asia—New Guinea and America. There are variable ploidy levels and chromosome numbers in these species (Fig. 15). Cytological evidences have indicated that B. distachyon (2n=10) has evolutionary relationships with polyploid perennial Brachypodium species probably as a parental species (Fig. 15) (Wolny and Hasterok 2009, Wolny, Lesniewska et al. 2011, Betekhtin, Jenkins et al. 2014). Having the ecological advantages, perennial grasses promise to serve as a significant source of renewable energy in the near future (Steinwand, Young et al. 2013). Thus the perennial B. sylvaticum has been proposed as a new model for perennial monocots (Steinwand, Young et al. 2013, Gordon, Liu et al. 2016).

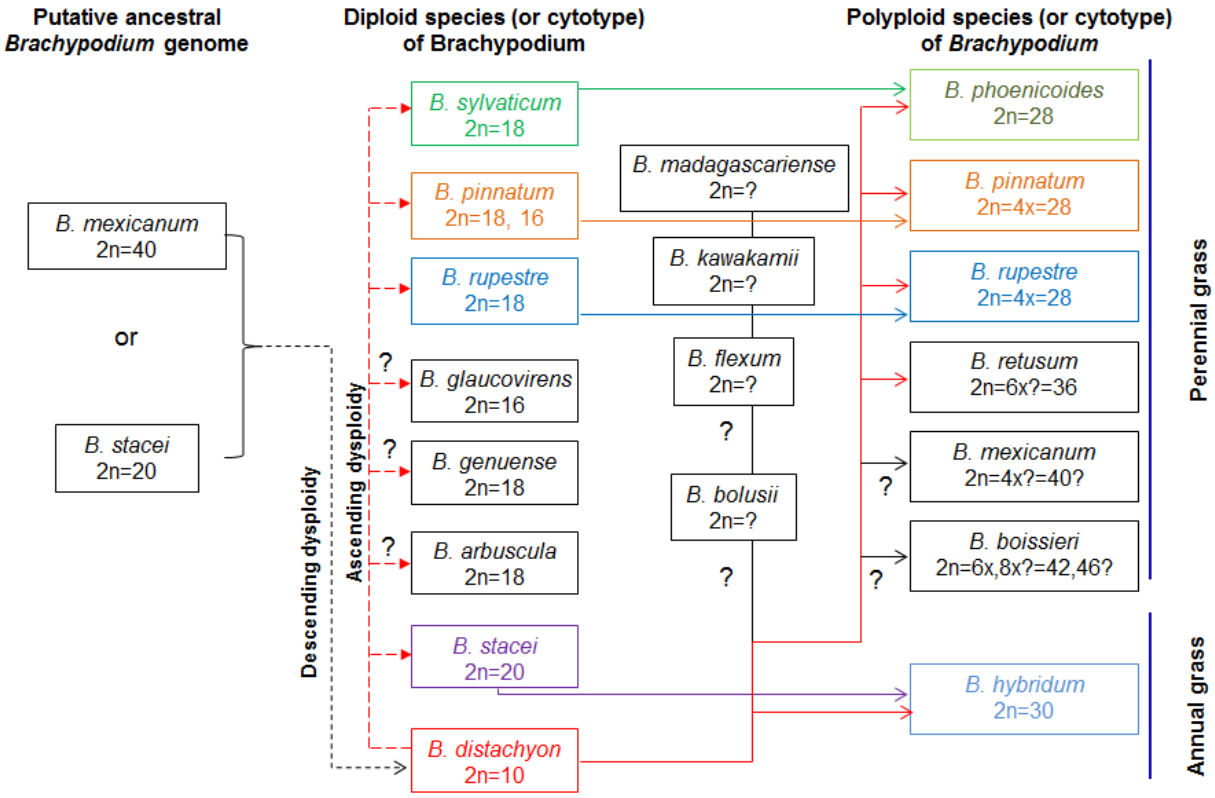


Fig. 15. The evolutionary relationship between the 20 Brachypodium species and cytotypes. B. distachyon may have resulted by descending dysploidy (dash black arrow) from a putative Brachypodium genome ancestor, very closed to B. stacei or B. mexicanum. Allopolyploids were formed by hybridization between B. distachyon (solid red arrow) and other diploids species (solid arrow, color is same as the species color). Among three annual species, the allopolyploid B. hybridum has B. distachyon and B. stacei as progenitors. Species have question mark in chromosome number have not yet well defined karyotype and ploidy level. The black arrows or lines with question marks stand for the evolutionary process which have not elucidated (Adapted from Catalan, López-Álvarez et al. (2016) and Betekhtin, Jenkins et al. (2014).

IV.1.2. Position of Brachypodium in the Poaceae family

The Poaceae is a large and nearly-ubiquitous family of monocotyledonous flowering plants and comprises more than 12,000 species, classified into 771 genera, with species having important roles in the economy and the human life, used for food production and industry as well as ornamental grasses (Soreng, Peterson et al. 2015). Brachypodium genus occupies an intermediate evolutionary position in this Poaceae family, "bridging" the gap between temperate and tropical cereals. Brachypodium genus belongs to a distinct tribe named Brachypodieae that diverged from a common ancestor of Aveneae and Triticeae about 38 MYA (Fig. 16). These tribes belong to Pooideae subfamily that has been separated from the other subfamilies such as Ehrhartoideae and Panicoideae around 55 (49-66) MYA (Chalupska, Lee et al. 2008, Bouchenak-Khelladi, Slingsby et al. 2014, Catalan, López-Álvarez et al. 2016). The Brachypodium genus has thus a close evolutionary relationship to A. savita (oats), H. vulgare (barley), S. cereal (sorghum) and Triticum spp. (wheat). These important crops are temperate grasses and possess much bigger genomes than B. distachyon. Their genomes not only are huge but also complex because of important proliferation of transposable elements (Fig. 16). Using a model plant with a smaller and more compact genome allows scientists to overcome these obstacles. Therefore, the closeness of Brachypodium to the temperate cereals promises direct applications on those important cereals. Brachyodium became a powerful model resource for genome evolution and functional studies on temperate grasses. The benefits of Brachypodium model not only permit to understand the bigger genome of other cereal, but also to detect the mechanisms in C4 plants which still have modest knowledge than those of C3 plants.

IV.2 Emergence of the Brachypodium distachyon as a model for temperate grass

Brachypodium distachyon was first proposed as a model for genomics of temperate grasses in 2001 (Draper, Mur et al. 2001). Extensive resources and studies on Brachypodium have increased rapidly since then (Brkljacic, Grotewold et al. 2011, Mur, Allainguillaume et al. 2011, Catalan, Chalhoub et al. 2014, Garvin 2016, Vogel 2016), especially after the achieving of the B. distachyon genome sequencing (IBI 2010). Nowadays, B. distachyon

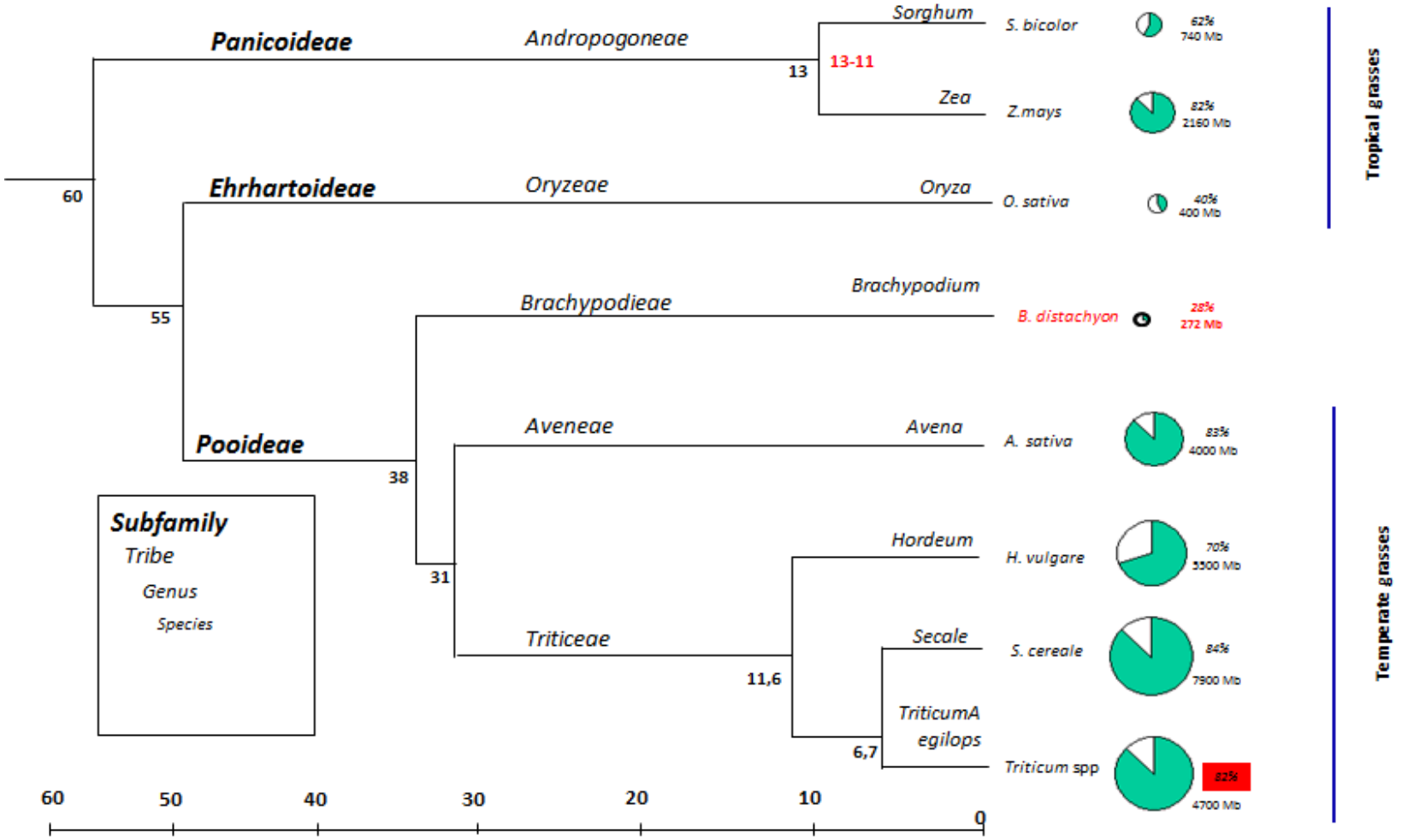


Fig. 16. Phylogenetic position of Brachypodium in the Poaceae family and divergence from other subfamilies and tribes (calculated according to gene sequence divergence from Paterson, Bowers et al. (2009) for sorghum, maize and rice, IBI (2010) for B. distachyon, Charles, Belcram et al. (2008) for wheat, Bartos, Paux et al. (2008) for rye, Bennett and Smith (1976) for oat and Wicker, Sabot et al. (2007) for barley). The sizes of genomes are indicated by circles and their transposable element proportion are illustrated by green.

represents a tractable model grass that is compatible with modern high-throughput molecular-genetic experiments (Vogel 2016).

The B. distachyon genome size of 272 Mb is intermediate between those of the two other model plants, Arabidopsis thaliana (119 Mb) and rice (430 Mb) (Initiative 2000), http://pgsb.helmholtz-muenchen.de/plant/rice/. Though A. thaliana is an extremely powerful generalized plant model but it is not adaptable for all aspects of monocots plants because of many differences in biological traits and long evolutionary distance between dicots and monocots. Oryza sativa represent a more suitable model for monocots but it is a semi-aquatic tropical grass, and has a larger plant size and difficult growth requirement conditions, less adapted to temperate grasses such as wheat, barley and oats. In this context, Brachypodium became the most suitable choice as a model of temperature grasses, most of which species possess huge genome size, rendering them less directly accessible than Brachypodium (Garvin 2016, Vogel 2016).

Evidences of the wide-spread adoption of B.distachyon as model were demonstrated in the reviews of Brkljacic, Grotewold et al. (2011), Catalan, Chalhoub et al. (2014), (Vogel 2016). The compact fully sequenced B. distachyon genome (IBI 2010) is a major reason for the success of this species as model, providing unique opportunities to study various aspects of grass genome organization and evolution. The variations in chromosome number and ploidy levels in the Brachypodium genus are also of interest for speciation and evolutionary studies.

The comparison between Brachypodium, Arabidopsis and some important crops (Table 1) showed that Brachypodium has all characters of a standard model plant. Biological traits such as small plant size, short cycle life time and important seed number per plant are particularly useful. In addition, features such as the simple growth requirement, selfing reproduction but crossing can be realized, highly-efficient transformation with Agrobacterium tumefaciens (Thole and Vain 2012, Garvin 2016) prevail Brachypodium as very suitable for generating genomic resource. On the other hand, Arabidopsis model is not suitable for answering all questions about monocots (as difference in cell wall type and photosynthesis mechanism…), whereas Brachypodium, a monocot which possesses all the essential conditions of a model plant to study the temperature cereals.

In summary, the development of B. distachyon as a model plant provides powerful tools as well as interesting resources for plant research.

Table 1. Comparison of select models and crops (Adapted from Brkljacic, 2011; Vogel, 2009)

Parameter	Arabidopsis	Brachypodium	Barley	Maize	Rice	Sorghum	Switchgrass	Wheat
Height (cm)	15–20	15–20	50–120	120–300	100	50–250	200–300	50–100
Density (plants/m2)	2000	1000	80-120	4	36	50	6	50
Generation time (weeks	8-12	8-12	10-20	8-15	12-24	13-18	26	10-20
Cell wall type	Type 1	Type 2	Type 2	Type 2	Type 2	Type 2	Type 2	Type 2
Photosynthesis	C3	C4	C4	C4	C4	C4	C4	C4
Growth requirements	Simple	Simple	Intermediate	Demanding	Demanding	Demanding	Demanding	Intermediate
Efficiently crossed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Reproduction	Selfing	Selfing	Selfing	Outcrossing/ self-compatible	Selfing	Selfing	Outcrossing	Selfing
Seeds per plant	>1000	100-1000	150-200	200-1000	>1000	>1000	>1000	50-150
Transformation	Extremely easy	Highly efficient	Efficient, but labor	Efficient, but labor intensive	Highly efficient	Inefficient	Efficient, but slow	Inefficient
Genome size (Mb)	119	272	5,500	2,300	382	700	2,400	16,000
Assembled genome sequence	Finished genome sequence	High-quality draft (finishing under way	Draft genome sequencing in progress	Draft genome	Finished genome sequence	Draft genome	Sequencing in progress	Sequencing in progress
T-DNA resources	Extensive	10,000 lines available, 40,000 more planned	None	Transposon mutants are available	Extensive	None	None	None

IV.3 Emergence of Brachypodium polyploid model

Three annual species, B. distachyon, B. stacei and B. hybridum represent an original polyploid model. The allotetraploid B. hybridum was derived from hybridization between B. distachyon and B. stacei (Catalan, Chalhoub et al. 2014). A main originality of this polyploid model is that the two parental species have a similar genome size and content. However, because of particular evolution of the former through rounds of chromosome fusions and rearrangements, B. distachyon (2n=10) has two times less chromosomes that are two times bigger than those of B. stacei (2n=20) (Fig. 17) (Betekhtin, Jenkins et al. 2014, Catalan, López-Álvarez et al. 2016). The dating analysis indicated that divergent time of B. distachyon and B. stacei is ~16 MYA (Catalan, López-Álvarez et al. 2016, Gordon, Liu et al. 2016).

One other important characteristic is that these species possess small and compact genomes. The high quality genome assemblies for all three species should be available soon. The B.distachyon genome released in 2010 showed good synteny with other grasses such as rice and Sorghum (IBI 2010). The B. stacei genome assembly and annotation and that of B. hybridum genome are progressing https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Bstacei. The analysis of distribution demonstrated that B. hybridum is more common and occupies more environmental niches than B. distachyon or B. stacei (Catalan, López-Álvarez et al. 2016), suggesting that it is more adapted to variable conditions (Bevan, Garvin et al. 2010). In addition, B. hybridum and B. stacei are easy to grow under controlled conditions, self-fertile and easy to transform (Draper, Mur et al. 2001). The evolutionary mechanism which formed the perennial polyploid grasses and the relationship between polyploid and diploid species still need to be elucidated.

The establishment of a Brachypodium polyploid model necessitates the development of tools and materials such synthetic polyploids through hybridizing the parental species that need to be characterized. These constitute an important objective of my PhD thesis.

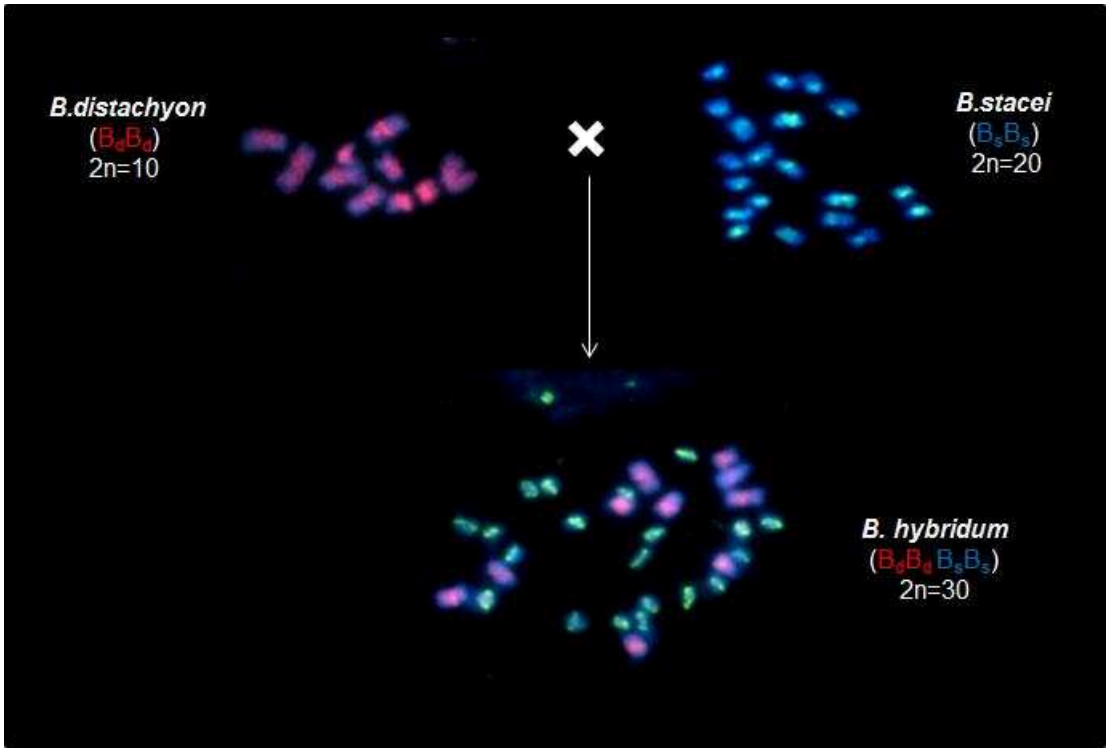


Fig. 17. Cytogenetic comparison of chromosome size and number of the three annual Brachypodium species showing that B. distachyon (2n=10) has two times less chromosomes that are two times bigger than those of B. stacei (2n=20). The allotetraploid B. hybridum (2n=30) derived from their hybridization possess an additive chromosomes number and karyotype (From Hasterok, Draper et al. (2004))

VI. CONTEXT AND OBJECTIVE OF THE THESIS

I began my PhD project in 2010, the year when the whole genome sequence of B. distachyon has been published by the International Brachypodium Initiative (IBI). The laboratory of organization and evolution of plant genomes (OEPG, led by Dr. Boulos Chalhoub) that I integrated have been very actively contributing to this consortium and the promotion of Brachypodium as a model. Since polyploidy is a main topic worked at OEPG, my supervisor, Dr. Boulos Chalhoub, suggested me to develop a Brachypodium polyploid model, in a similar way to the Brassica and the wheat models that were investigated.

The importance of polyploidy as a main plant genome evolutionary force has been investigated in several other plant models (Soltis and Soltis 2009, Beest, Le Roux et al. 2011). The evolution and the importance of polyploidy in the genus Brachypodium were just starting at the beginning of my PhD (Hasterok, Draper et al. 2004, Hasterok, Marasek et al. 2006, Idziak and Hasterok 2008, Catalan, Muller et al. 2012). Separating the so-called three cytotypes of B. distachyon into B. distachyon with n= 5, B. stacei with n=10 and the allopolyploid B. hybridum with n=15, derived from their hybridization was suggested only during my PhD preparation (Catalan, Muller et al. 2012).

This PhD subject comes also in the context of interest in Brachypodium expressed by the scientific community. After the release of the B. distachyon genome sequence, the international consortium, which my laboratory is contributing, has launched a project that consisted in sequencing a diversity collection of 98 B. distachyon inbred lines (http://jgi.doe.gov/our-science/science-programs/plant-genomics/brachypodium/), some of which sequences have been already released (http://onlinelibrary.wiley.com/doi/10.1111/tpj.12569/abstract). Importantly, the project also proposed a sequencing of B. stacei and B. hybridum which is of high interest for the development of the Brachypodium polyploid model.

There are several interests for the development of B. hybridum as a Brachypodium polyploid model (Vogel 2016). Brachypodium species grow fast, have very small genomes and are more and more becoming a model for grass species. As far as I know, this represents the first polyploid model where the parental genomes, of equal size, differ by twice in the number of chromosomes and chromosome size. Several important questions are about how these chromosomes behave when joined together in the allopolyploid genome at structural, evolutionary and meiosis levels (Betekhtin, Jenkins et al. 2014, Catalan, López-Álvarez et al.

VI. CONTEXT AND OBJECTIVE OF THE THESIS

I began my PhD project in 2010, the year when the whole genome sequence of B. distachyon has been published by the International Brachypodium Initiative (IBI). The laboratory of organization and evolution of plant genomes (OEPG, led by Dr. Boulos Chalhoub) that I integrated have been very actively contributing to this consortium and the promotion of Brachypodium as a model. Since polyploidy is a main topic worked at OEPG, my supervisor, Dr. Boulos Chalhoub, suggested me to develop a Brachypodium polyploid model, in a similar way to the Brassica and the wheat models that were investigated.

The importance of polyploidy as a main plant genome evolutionary force has been investigated in several other plant models (Soltis and Soltis 2009, Beest, Le Roux et al. 2011). The evolution and the importance of polyploidy in the genus Brachypodium were just starting at the beginning of my PhD (Hasterok, Draper et al. 2004, Hasterok, Marasek et al. 2006, Idziak and Hasterok 2008, Catalan, Muller et al. 2012). Separating the so-called three cytotypes of B. distachyon into B. distachyon with n= 5, B. stacei with n=10 and the allopolyploid B. hybridum with n=15, derived from their hybridization was suggested only during my PhD preparation (Catalan, Muller et al. 2012).

This PhD subject comes also in the context of interest in Brachypodium expressed by the scientific community. After the release of the B. distachyon genome sequence, the international consortium, which my laboratory is contributing, has launched a project that consisted in sequencing a diversity collection of 98 B. distachyon inbred lines (http://jgi.doe.gov/our-science/science-programs/plant-genomics/brachypodium/), some of which sequences have been already released (http://onlinelibrary.wiley.com/doi/10.1111/tpj.12569/abstract). Importantly, the project also proposed a sequencing of B. stacei and B. hybridum which is of high interest for the development of the Brachypodium polyploid model.

There are several interests for the development of B. hybridum as a Brachypodium polyploid model (Vogel 2016). Brachypodium species grow fast, have very small genomes and are more and more becoming a model for grass species. As far as I know, this represents the first polyploid model where the parental genomes, of equal size, differ by twice in the number of chromosomes and chromosome size. Several important questions are about how these chromosomes behave when joined together in the allopolyploid genome at structural, evolutionary and meiosis levels (Betekhtin, Jenkins et al. 2014, Catalan, López-Álvarez et al.

2016).

Starting from only seeds of these different species, I have now successfully developed the synthetic polyploids, achieving an important step towards its adoption as an important polyploidy model. My PhD work and results are completely integrated within the international consortium and efforts to characterize the Brachypodium polyploid model. My PhD program has been done in the laboratory of organization and evolution of plant genomes and I collaborated with laboratories of INRA-Rennes, Centre national de génotypage (CNG), Centre national de Séquençage (Genoscope), and other members of IBI.

RESULTS AND DISCUSSION

A key point for the development of a successful polyploid model is to obtain synthetic allopolyploids similar or different from the natural ones. During my PhD, major efforts have been put on the development and characterization of Brachypodium autopolyploids and more importantly allopolyploids similar to the natural B. hybridum.

My efforts did not consist in only development of the material but I characterized these at the cytogenetic, genetic and gene expression levels. I also generated important genomic data that will allow better characterization of this very important polyploidy model.

I am presenting here my results organized into three main chapters:

- The development and characterization of Brachypodium autopolyploids, constituting Chapter 1, are presented as a paper submitted to the journal Plant Systematic and Evolution: **Synthesis and evaluation of autotetraploids from Brachypodium distachyon and B. stacei species** (Vinh-Ha Dinh Thi, Isabelle Le Clainche, Olivier Coriton, Joseph Jahier and Boulos Chalhoub).

- The resynthesis and characterization of allopolyploids similar to the natural B. hybridum, constituting Chapter 2, are presented as a paper accepted in the journal PlosOne: **Recreating Stable Brachypodium hybridum Allopolyploids by Uniting Divergent Genomes of B. distachyon and B. stacei that Contrast in Chromosome Evolution, Number and Size** (Vinh Ha Dinh Thi, Olivier Coriton, Isabelle Le Clainche, Dominique Arnaud, Sean P. Gordon, Gabriella Linc, Pilar Catalan, Robert Hasterok, John P. Vogel, Joseph Jahier, Boulos Chalhoub).

- Finally, the experiments that I have set up to characterize polyploidy-related changes at genome structure (through DNA re-sequencing) cytosine methylation (through bisulfite sequencing) and gene expression (through RNA-Sequencing) levels, using next generation sequencing (NGS) tools, together with preliminary analysis of gene expression changes in B. distachyon autopolyploids and preceded by an updated comparison of NGS techniques, are presented as Chapter 3.

CHAPTER 1:

SYNTHESIS OF BRACHYPODIUM AUTOPOLYPLOIDS

This chapter is presented as a paper submitted to the journal Plant Systematic and Evolution:

Article 1: "Synthesis and evaluation of autotetraploids from Brachypodium distachyon and B. stacei species"

Vinh Ha Dinh Thi, Isabelle Le Clainche, Olivier Coriton, Joseph Jahier and Boulos Chalhoub

Plant Systematics and Evolution

Synthesis and evaluation of autotetraploids from Brachypodium distachyon and B. stacei species

--Manuscript Draft--

Manuscript Number:	
Full Title:	Synthesis and evaluation of autotetraploids from Brachypodium distachyon and B. stacei species
Short Title:	BRACHYPODIUM AUTOTETRAPLOID SYNTHESIS
Article Type:	S.I. : Polyploidy in Shallow and Deep evolutionary Times
Keywords:	Autopolyploidy; Brachypodium; Chromosome-doubling; cytogenetics; karyotype
Corresponding Author:	Boulos Chalhoub INRA Centre de Versailles-Grignon / Université d'Evry Val d'Essonne FRANCE
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	INRA Centre de Versailles-Grignon / Université d'Evry Val d'Essonne
Corresponding Author's Secondary Institution:	
First Author:	Vinh Ha Dinh Thi
First Author Secondary Information:	
Order of Authors:	Vinh Ha Dinh Thi
	Isabelle Le Clainche
	Olivier Coriton, PhD
	Joseph Jahier
	Boulos Chalhoub
Order of Authors Secondary Information:	
Funding Information:	ANR (ANR-11-BSV5-0015) — Dr. Boulos Chalhoub
Abstract:	Abstract Autopolyploidy is a main evolutionary and diversification mechanism of plant genomes but few studies have focus on the importance and consequences of autopolyploidy. For this purpose, we have generated and characterized in the present study synthetic autotetraploids of the emerging models species, Brachypodium distachyon (2n=10) and B. stacei (2n=20), which have similar genome content but divergent chromosome evolution. Two inbred lines of B. distachyon (Bd21 and Bd3-1) and three ones of B. stacei (ABR114, Bsta5 and TE4.3) were used. The genome doubling, through colchicine treatment, was validated by flow cytometry and karyotyping with fluorescent in situ hybridization analyses. Autopolyploids obtained from the two B. distachyon lines and lines Bsta5 and TE4.3 of B. stacei showed stability in phenotype as well as karyotype. For the third line of B. stacei (ABR114) a colchicine-treated plant gave also stable autotetraploids whereas another one showed various aneuploid progenies. Precise phenotype and quantitative comparison of inflorescences and flag leaves characters showed that both B. distachyon and B. stacei autotetraploids generally exceeded their diploid progenitors, but their fertility was reduced as illustrated by the lower number of seeds per inflorescence and lower percent of fertile florets. The generated autotetraploids provide an interesting material to study the fate of homologous duplicated genes, meiosis and various genomic consequences of autopolyploidy, comparatively between the two Brachypodium species having similar genome content but having asymmetric evolution of chromosome number and size.
Suggested Reviewers:	Jinling Meng, PhD

	Distinguished professor, Houzhang Agriculture University (HZAU) jmeng@mail.hzau.edu.cn Specialist in inter-specific hybridization and polyploidy
	Andrew Leitch Distinguished professor, Queen Mary University of London a.r.leitch@qmul.ac.uk Specialist in polyploidy and hybridization
	Ainouche Malika, PhD professor, Universite de Rennes 1 malika.ainouche@univ-rennes1.fr Specialist in polyploidy and ecology

Synthesis and evaluation of autotetraploids from *Brachypodium distachyon* and *B. stacei* species

Vinh-Ha Dinh Thi[1], Isabelle Le Clainche[1], Olivier Coriton[2], Joseph Jahier[2], Boulos Chalhoub[1]*

Runing title

BRACHYPODIUM AUTOTETRAPLOID SYNTHESIS

Affiliations

[1]Organization and evolution of plant complex genomes (OECG) Institut National de la Recherche agronomique (INRA), Université d'Evry Val d'Essonne (UEVE), 2 rue Gaston Crémieux, 91057 Evry France.

[2]UMR1349 Institut de Génétique, Environnement et Protection des Plantes (IGEPP), INRA, BP35327, 35653, Le Rheu cedex, France.

** Corresponding author:* Boulos Chalhoub

Organization and evolution of plant complex genomes (OECG), INRA/UEVE, 91057 Evry Cedex, France

Email: chalhoub@evry.inra.fr

Phone: +33 1 60 87 45 03

Fax: +33 1 60 87 45 49

Abstract

Autopolyploidy is a main evolutionary and diversification mechanism of plant genomes but few studies have focus on the importance and consequences of autopolyploidy. For this purpose, we have generated and characterized in the present study synthetic autotetraploids of the emerging models species, *Brachypodium distachyon* (2n=10) and *B. stacei* (2n=20), which have similar genome content but divergent chromosome evolution. Two inbred lines of *B. distachyon* (Bd21 and Bd3-1) and three ones of *B. stacei* (ABR114, Bsta5 and TE4.3) were used. The genome doubling, through colchicine treatment, was validated by flow cytometry and karyotyping with fluorescent *in situ* hybridization analyses. Autopolyploids obtained from the two *B. distachyon* lines and lines Bsta5 and TE4.3 of *B. stacei* showed stability in phenotype as well as karyotype. For the third line of *B. stacei* (ABR114) a colchicine-treated plant gave also stable autotetraploids whereas another one showed various aneuploid progenies. Precise phenotype and quantitative comparison of inflorescences and flag leaves characters showed that both *B. distachyon* and *B. stacei* autotetraploids generally exceeded their diploid progenitors, but their fertility was reduced as illustrated by the lower number of seeds per inflorescence and lower percent of fertile florets. The generated autotetraploids provide an interesting material to study the fate of homologous duplicated genes, meiosis and various genomic consequences of autopolyploidy, comparatively between the two *Brachypodium* species having similar genome content but having asymmetric evolution of chromosome number and size.

Key words: Autopolyploidy, *Brachypodium*, Chromosome-doubling, cytogenetics, karyotype

Introduction

Polyploidy, leading to whole genome duplication, is a major evolutionary force in eukaryotes and is particularly prominent and recurrent in angiosperm (Darlington 1937; Grant 1963; Blanc and Wolfe 2004; Schlueter et al. 2004; Adams and Wendel 2005; Cui et al. 2006; Jaillon et al. 2007; Tang et al. 2008; Soltis and Soltis 2009; Van de Peer et al. 2009; Beest et al. 2011; Barker et al. 2015). Allopolyploids combine two or more divergent homoeologous genomes, usually through interspecific or intergeneric hybridization, followed by chromosome doubling whereas autopolyploids combine less divergent genomes from the same species, resulting in more than two sets of homologous chromosomes in the nucleus.

Within the *Poaceae* family, species of the genus *Brachypodium* constitute an attractive polyploid model. The diploid plant model *B. distachyon* (2n=10) has evolved by descending dysploidy, acting as fusions of smaller chromosomes, from a putative ancestral *Brachypodium* species, very close to *B. stacei* (2n=20) (Hasterok et al. 2004; Betekhtin et al. 2014). Moreover, within this framework, these two diploid species have hybridized to give rise to the allopolyploid species, *B. hybridum.* Together, these three species comprise an excellent model to investigate the impact of polyploidy on the organization and evolution of plant genomes, because all three possess small genomes, have small plant stature, rapid generation time, significant morphometric with the originality of the high asymetric divergence between parental chromosomes (and by corollary the sub-genomes of *B. hybridium*) (Hasterok et al. 2004; Betekhtin et al. 2014; Catalan et al. 2016).

Few studies have been dedicated to the importance, the prevalence and consequences of autopolyploidy. For this purpose, we have generated and

characterized at the karyotype, phenotype a fertility levels synthetic autotetraploids of *Brachypodium distachyon* and *B. stacei,* providing an interesting material to study the fate of homologous duplicated genes, meiosis and various genomic consequences of autopolyploidy, comparatively between the two *Brachypodium* species having similar genome content but contrasting in chromosome evolution and structure.

Numerous crop species are cultivated as autopolyploids. The forage crops alfalfa (*Medicago sativa*) is an attractive system for studying autopolyploidy with several diploids and autotetraploids (Havananda et al. 2011). Potato (*Solanum tuberosum*) actually has several autopolyploidy levels, ranging from autodiploids to autohexaploids, with autotetraploids used as cultivated varieties (Consortium 2011; Cai et al. 2012). Several important fruit species are also autopolyploids such as kiwifruit (*Actinidia chinensis*) (Wu et al. 2012), apple (*Malus domestica*) (Velasco et al. 2010), and dragon fruit (*Hylocereus undatus*) (Cohen and Tel-Zur 2011; Tel-Zur et al. 2011). The important *Poaceae* family comprises many species with diploid and autopolyploid series, such as ryegrass (*Lolium perenne* L.) (Simonsen 1973; Sugiyama 2005) and switchgrass (*Panicum virgatum* L.) (Saski et al. 2011; Childs et al. 2014).

The genetic redundancy could potentially facilitate adaptation of duplicated genes, increasing the genome flexibility as a result of genome structure, functional and epigenetic changes, contributing, thus, to genome diversification (Parisod et al. 2010). Nevertheless, genetic, functional and epigenetic consequences of autopolyploidy have been modestly evaluated during the last decade (Parisod et al. 2010; Barker et al. 2015).

Containing more than two sets of homologous chromosomes, it is also expected that autopolyploids form multivalents when their chromosomes pair at meiosis, instead of bivalents observed in diploids or highly-stable allopolyploids (Ramsey and Schemske 2002). However, frequencies and types of multivalents seem to be different among the various taxa. The *Lolium perenne* exhibited 1% trivalents and 20% quadrivalents configurations at metaphase I (Simonsen 1973), while tomato autotetraploids had almost no trivalents and 19% quadrivalents (Upcott 1939). Arabidopsis autotetraploids showed higher multivalent frequencies (Santos et al. 2003). Ramsey and Schemske (2002) suggested that factors such as the non-random associations among some homologous chromosomes, and the existence of physical limitations or genetic factors may regulate bivalent and multivalent formation when pairing at meiosis. The formation of multivalents leads to fertility reduction (Levy and Feldman 2002; Comai 2005). The seedless fruits (*parthenocarpy)* are generated as unpaired autopolyploids (3x, 5x) (Varoquaux et al. 2000).

It has been found since 1937 that autopolyploids can be artificially synthesized through colchicine (N-(5,6,7,9-tetrahydro-1,2,3,10-tetra-methoxy-9-oxobenzo(a)heptalen-7-yl)acetamide) treatment (Blakeslee and Avery 1937; 1938) and numerous autopolyploids have been widely generated (Eigsti 1938; Bradley and Goodspeed 1943; Eigsti and Dustin 1955; Gupta 1981; Caperta et al. 2006). At the adequate concentration, colchicine prevents formation of mitotic spindle fiber, leading to doubling of the chromosome set (Bartels and Hilton 1973; Vaughn and Lehnen 1991). The colchicine impact is more efficient on dividing cells such as in meristem tissues. Therefore, the application of this chemical substance was often used on seeds, seedling, or meristem, cultured *in vitro* (Jahier 1992). It was not reported any deleterious effects of colchicine on chromosomes such as the

fragmentation or chromosome breakage. The common aberration following colchicine treatment and genome doubling is aneuploidy as a results of missegregation of homologous chromosomes present in more than two copies (Dermen 1940; 1945).

In this study, autopolyploids were obtained by colchicine treatment of *B. distachyon* and *B. stacei* sister species that have got asymmetric chromosome evolution since their divergence. Both species have similar genome size and content but *B. distachyon* (2n=10) has 10 larger chromosomes, that were likely to have originated via descending dysploidy, acting as fusions of twice smaller chromosomes from a putative ancestral *Brachypodium* species, very close to *B. stacei* (2n=20) (Hasterok et al. 2004; Betekhtin et al. 2014). We have generated and characterized at the phenotype and cytogenetic levels the different obtained autotetraploids. They represent an interesting material to study the fate of homologous duplicated genes, as well as genetic, functional and epigenetic consequences of autopolyploidy and regulation of meiosis**.**

Materials and methods

Plant materials

Two inbred lines from *B. distachyon* (Bd21 and Bd3-1) and four others from *B. stacei* (ABR114, Bsta5, TE4.3 and LP6.1) have been chosen based on the phylogenetic distances between accessions that have been evaluated previously (Vogel et al. 2009). The line ABR114 of *B. stacei*, characterized in previous researches (Hasterok et al. 2004; Hasterok et al. 2006), was the only available accession of this species in

our laboratory when we started autopolyploid synthesis. Through collaboration within the *Brachypodium* international consortium, we got later access to three other *B. stacei* accessions (Bsta5, TE4.3 and LP6.1) from Catalan and Vogel (Lopez-Alvarez et al. 2012).

Colchicine treatment

The protocol applied on *Brachypodium* was adapted from the method described by Jahier (1992), which was successful used on wheat (Mestiri et al. 2010). Both stages of germinated seeds, with only 1-2 cm length roots, or seedlings of 2-3 leaves (Online Resource 1) were tested. Seedlings or germinated seeds were completely immersed for three hours in an aqueous colchicine (Sigma -Aldrich Co., cat. no. C9754) solution at a concentration of 2,5mg/ml and containing 2% DMSO (dimethyl sulfoxide, Sigma -Aldrich Co., cat. no. D8418) (Online Resource 1). The colchicine-treated plants were then transferred directly into new soil pots, without rinsing by distilled water, and grown in greenhouse. The pots were watered well after treatments (Online Resource 1). Seven to 10 days after treatments we checked recovered plants with new emerging tillers with burning dots on the leaves, indicative of colchicine effect (Online Resource 1).

Flow cytometry analysis of nuclear suspensions

The flow cytometry is widely used to estimate DNA amounts and screening on leaves from colchicine treated plants and their S1 and S2 progenies as compared to non-treated diploid parents was done to primary check genome size increase.

Analyses were performed on young leaves obtained from plants grown in the greenhouse. The leaves were chopped with a sharp razor blade at room temperature in a Petri dish containing 400 µl of extraction buffer (CyStain UV Precise P Nuclei Extraction Buffer, Partec, Münster, Germany). The suspension was filtered through a 50 µM filter and nuclei were stained with 1.6 mL 4',6-diamindino-2-phenylindole (DAPI) buffer (CyStain UV Precise P Staining Buffer, Partec). The suspension was analyzed using a flow cytometer (Partec Cyflow Space, Partec) to determine the mean sample nuclei fluorescence relative to that of the internal standard (diploid control). The software used is FloMax Version 2.0 (Partec).

Chromosome counting at the metaphase stage of mitosis and fluorescence in situ hybridization (FISH) analysis

We checked chromosome number in cells at metaphase stage of mitosis. Root-tips of potential autopolyploid plants were collected as described for wheat by Jahier (1992). Freshly emerging root-tips from (2-3 cm length including the tip) were cut and put directly into α-bromonaphthalene solution (Sigma -Aldrich Co., cat. no. B73104) at 4℃ for 20-24 hours then fixed in of 3: 1 (v/v) of ethanol:glacial acetic acid solution at 4℃ for 24 hours or stored at -20℃ until use. After washing in 0.01 M enzyme buffer (citric acid-sodium citrate pH 4.5) for 15 min, the roots were digested in a solution of 5% Onozuka R-10 cellulase (Sigma) and 1% Y23 pectolyase (Sigma) at 37 ℃ for 15 min. The root tips were then washed carefully with distilled water for 30 min. A root tip was transferred to a slide and macerated with a drop of 3:1 fixation solution using a fine-pointed forceps. After air-drying, slides with good metaphase chromosome spreads were stored in -20 ℃.

For DAPI staining, the chromosomes were mounted and counterstained in Vectashield (Vector Laboratories) containing 2.5 µg/mL of DAPI. Photomicrographs were taken using a CoolSnap HQ camera (Photometrics, Tucson, Ariz) on an Axioplan 2 microscope (Zeiss, Oberkochen, Germany) and analyzed using MetaVue™ (Universal Imaging Corporation, Downington, PA).

For the FISH experiments, the ribosomal probes used in this study was 45S rDNA (pTa 71) (Gerlach W. and J. 1979) which contained a 9-kb *Eco*RI fragment of rDNA repeat unit (18S-5.8S-26S genes and spacers) isolated from *Triticum aestivum*. The pTa 71 was labelled with Alexa-488 dUTP by random priming. The centromeric DNA probe, CentBd, used for counting chromosome number contains 156 bp of long tandem repeat sequences (Wen et al. 2012) was labelled by PCR with biotin-dUTP (Roche) from *B. stacei.* Chromosome preparations were incubated in RNAse A (100ng/µL) and pepsin (0.05%) in 0.01 M HCl, post-fixed with 1% paraformaldehyde, dehydrated in an ethanol series (70%, 90% and 100%) and air-dried. The hybridization mixture consisted of 50% deionized formamide, 10% dextran sulphate, 2X SSC, 1% SDS and labelled probes (200 ng per slide) was denatured at 92℃ for 6 min, and transferred to ice. Chromosomes were denatured in a solution of 70% formamide in 2X SSC at 70℃ for 2 min. The denatured probe was placed on the slide for the night in a moist chamber at 37℃. After hybridization, the slides were washed for 5 min in 50% formamide in 0.2X SSC at 42℃ (98% stringency), followed by several washes in 4X SSC-Tween. Biotinylated probe was immunodetected by Texas Red-avidin DCS (Vector Laboratories).

Phenotype analyses

The data of morphological analysis have been recorded on the S1 and S2 confirmed autopolyploid plants in comparison to their direct diploid progenitors. Sixteen traits have been measured (Online Resource 2, 3). Five characters are directly related to fertility and include number of spikelet per inflorescences, florets number per spikelet, seed number per inflorescence, fertility ratio (percent of florets giving seeds out of total florets on the inflorescence). Other characters of inflorescence were also recorded in this experiment and include inflorescence length, spikelet length, distance between spikelets of florescence, upper glume length, upper glume width and the floret characters as floret length, lemma length (from basal floret, lemma width and awn length. The two last traits consist in flag leaf length and flag leaf width.

The S2 autopolyploid and the control diploid plants were grown in same greenhouse conditions at 22℃ with a 20 h photoperiod. Statistical analyses were performed by non-parametric Kruskal-Wallis (Kruskal and Wallis 1952).

Estimation of pollen abundance and viability

To compare pollen abundance between diploids, autopolyploids and aneuploids, the whole pollens contained in an individual anther were homogeneously spread on similar surfaces of a glass slide (size of 25 mm × 75 mm). The number of pollens was then counted under microscope (Leica DMLB) in 30 mm^2 units and averaged to represent the estimation of pollen abundance.

To estimate pollen viability, anthers were sampled the day of anthesis and were stained by acetocarmine as described by Jahier (1992). In brief, anthers were dilacerated in a drop of acetocarmine. Then the debris is removed and the liberated pollen grains are covered with a coverslip. The pollen viability was observed under

light microscopes after a few minutes. The pollen grain was considered viable if it turned into dark purple, the non-stained ones were classified as the non-viable.

Results

To produce autotetraploids we treated plants from *B. distachyon* and *B. stacei* species with colchicine to double the chromosome number. Progenies of colchicine-treated plants were checked for genome size increase and chromosome number by flow cytometry and karyotyping.

Colchicine treatment and effects

Two inbred lines of *B. distachyon* and four of *B. stacei* were treated by colchicine, either at seedling and/or germinated seeds stages to generate autopolyploids.

Out of 150 treated plants from the six accessions (105 treated at seedling stage and 45 at germinated seeds stage), 91 have survived (Table 1). Seven to 10 days after the treatment, burning or clear dots can be observed on leaves of the surviving treated plants that we designated as the zero selfed (S0) generation (Online Resource 1). The vegetative development of colchicine-treated plants was retarded in comparison to the control non-treated ones. New leaves and tillers emerged on the treated plants about 10 to 15 days after treatment (Online Resource 1). These plants grew fast and some of them were bigger than those of the non-colchicine treated

plant. Almost colchicine-treated plants flower between seven to 20 days later than the progenitor diploid plants (Table 1; Online Resource 1).

Colchicine effects on treated plants depended on the developmental stage, accessions and species of *Brachypodium* (Table 1). At the germination stage, damages caused by colchicine were very important, with 17 surviving plants obtained from 45 treated germinated seeds for *B. distachyon* (survival ratio of 37.8%), whereas those from all four accessions of *B. stacei* did not survive. Plant survival was higher when colchicine was applied at the seedling stage of the different lines and *Brachypodium* species, with 74 surviving plants out of 105 treated ones (a survival ratio of 70.5%) (Table 1). In most cases, the plant death occurred early within 10 days after the colchicine treatment. In other cases, treated plants remained longer alive but they were weak and never recover, dying before flowering. The LP6.1 accession was the most sensitive to colchicine, where all plants, treated at the two development stages, did not survive (Table 1).

Both lines of *B. distachyon* seemed more tolerant to colchicine than *B. stacei,* with 92% and 85% surviving plants for the former species and 37% and 0% for the later one, at the seedling and the germinated stages respectively (Table 1).

The surviving colchicine-treated plants exhibited between seven to 20 days delayed flowering time as compared to control non-treated ones (Table 1; Online Resource 1).

Flow cytometry (FMC) screening of S1 plants for increased DNA content

FMC allows a primary identification of plants showing a genome size increase, as compared to non-colchicine treated diploid plants.

Seeds harvested on colchicine-treated plants, designated first selfed generation (S1), were collected, sown and S1 plants were grown in greenhouse. FMC was applied on leaves of a total of 332 S1 plants, derived from 10 regenerated colchicine-treated S0 plants. Among these, 264 plants were from *B. stacei* (81 from ABR114, 19 from Bsta5, and 164 from TE4.3 lines) and 68 plants were from *B. distachyon* (41 from Bd21 and 27 from Bd3-1). Results are summarized in Table 2 and a typical example of comparison of FMC profile is shown in Fig. 1 for the diploid non-treated line ABR114 of *B. stacei* as compared to the S1 progeny (ABR114_8) of a colchicine treated plant, with almost double genome size increase. Precisely, in this S1 ABR114_8 line, the two coordinates (on the X-axis) of the two peaks of DNA content in nuclei in G1 (2C DNA) and G2/M (4C DNA) phases of mitosis, correspond to respectively those of 4C DNA and 8C of the diploid plant FMC profile, indicating thus almost double genome size increase (Fig. 1).

A total of 30 plants with almost doubled DNA size increase (named hereafter as S1 potential autopolyploid) have been screened by FMC (Table 2). Interestingly, differences in rates of plants with increased genome size were found here depending on genotypes used and individual treated plants. Among these, all the 15 S1 plants, progenies of the colchicine-treated plant ABR114_2 of line ABR114 of *B. stacei,* showed a doubled genome size profile (Table 2). This 100% genome size doubling is not usual in any documented plant species especially that, for another colchicine-treated plant of the same line (ABR114_8), we obtained only two S1 plants with doubled genome size out of 70 analyzed (doubling rate of 3.03%). No genome size increase was observed for the 51 analyzed S1 plants, progenies of the colchicine-treated plant (TE4.3_2_G2) of *B. stacei*, while 2.08% and 2.53% of plants with doubled genome size were observed among S1 progenies for the two other plants of

the same line (Table 2). The doubling rate was also different between *B. distachyon* lines and individual treated plants, with 12.5%; 17.65% and 7.41%, for Bd21_7, Bd21_8 and Bd3-1_5, respectively.

Overall, the analysis of FMC profiling, showing S1 plants with no change in genome size and others with double genome size increase as compared to non-colchicine treated plants, confirms the success of colchicine treatment. The 22 S1 plants of *B. stacei* and the eight ones of *B. distachyon* with almost doubled genome size were grown separately in greenhouse for phenotyping and chromosome counting in order to better precise whether they correspond to true autopolyploids and also to produce seeds of the second selfed (S2) generation (Table 3).

The S1 plants with doubled genome size increase derived from the ABR114_2 plant of B. stacei do not show homogeneous phenotype

No phenotypic variations were observed between the two sisters S1 plants obtained from line Bd3-1(Bd3-1_5) as well as the six ones derived from two different colchicine- treated plants (Bd21_7 and Bd21_8) of line Bd21 of *B. distachyon.* Similarly, for *B. stacei* no phenotypic differences were observed between the two S1 plants derived from line ABR114, the three ones derived from two colchicine treated plants (TE4.3_3_B6 and TE4.3_3_H10) of line TE4.3 and the two ones derived also from two treated plants (Bsta5_3_G2 and Bsta5_G11) of line Bsta5.

In contrast, there were important phenotypic variations among the 15 plants with double genome size increase, derived from the colchicine treated plant ABR114_2 of line ABR114 of *B. stacei*. These did not show a homogeneous phenotype, where we distinguished based on plant growth, spike morphology and fertility four different

phenotypes at least (Fig. 2). Twelve of these, that we called T1 type, were more or less homogeneous, with plant size smaller than the remaining three other sister plants (Fig. 2a), whereas flowering date was more similar to diploid line. They also showed one inflorescence type on their tillers (Fig. 2b), and fertility measured as the percent of florets having seeds ranges between 41 to 56% (Table 3). One other S1 plant (plant ID: ABR114_2 (3-1)), also descendant of ABR114_2, exhibited bigger plant size than T1 type sister plants, with spikes having abnormal compact structure with emergence of many secondary and small spikes at the nodes of tillers (Fig. 2b). This T2 type plant had also much reduced fertility (22%). Another S1 plant (ABR114_2 (3-5)), that we called T3, type exhibited also bigger plant size than T1 type plants and had long leaves that T1 and T2 types (Fig. 2a) whereas spikes were of similar structure to those T1 type (Fig. 2b), resembling more those of the diploid ABR114 line, with a very reduced fertility (28%). Finally, the remaining S1 sister plant (ABR114_2 (4-9)), that we call T4 type, was phenotypically similar to the T3 type but much more fertile (60%) (Fig. 2a, b).

Cytogenetic chromosome characterization and karyotyping

Cells at metaphase stage, spread on slides, were prepared from root tips of the S1 plants with doubled genome size increase as revealed by flow cytometry analysis. Counting of chromosomes on DAPI stained slides of *B. stacei* was much more technically challenging than *B. distachyon* because of their smaller size and important number (20 chromosomes in diploid *B. stacei*). Therefore, in addition to counting DAPI-stained chromosomes, we tried also to develop other types of hybridization probes to better solve and confirm chromosome number. Centromere

probe CentBd of *B. distachyon* was successfully used here to hybridize all chromosomes of *B. stacei* (Fig. 3; summarized in Table 3).The data were obtained by counting chromosomes in at least 10 cells per each plant.

In *B. distachyon* (lines Bd21 and Bd3-1), 10 chromosomes were found in the diploid progenitor, as expected, whereas we counted 20 chromosomes in cells from the autopolyploid S1 plants suggesting that each of the chromosomes was present in four copies (Fig. 3a, b). We used 45s rDNA region located on the short arm of *B. distachyon* chromosome 5 (Hasterok et al. 2004) as a probe in a typical FISH experiment. As expected, the diploid *B. distachyon* presented two hybridization signals of 45S regions, whereas four signals were observed for the autotetraploid S1 plants (Fig. 3a, b, green dots).

In *B. stacei* the S1 analyzed plants, derived from the colchicine-treated plant (ABR114_8) of accession ABR114 and of those from line Sta5 were all euploids with 40 chromosomes, expected for autopolyploids (Fig. 3). On the contrary, the analyzed S1 progenies of the plant ABR114_2 of line ABR114 showed various statuses of chromosome number variations. The S1 T1 type plants exhibited chimeric euploid and aneuploid cells with 38 to 41 chromosomes (Fig. 3c, d). Aneuploidy was also the case of the T2 type sister plant with cells of 42 and 41 chromosomes (Fig. 3e, f). Similarly, the T3 type plant exhibited chimeric cells with 40 and 41 chromosomes. Only the T4 plant showed euploid cells with 40 chromosomes, expected for an autopolyploid (Fig. 3g, h). The aneuploid of T1, T2 and T3 types showed also in S2 generation chimeric aneuploid cells and important variations in phenotype and inflorescence structure (see above).

Characterization of flower structure, abundance and pollen viability of diploids, autopolyploids and aneuploids

Flowers were compared between autopolyploids and their diploid progenitors at the same stage, about 5-6 days after the emergence of inflorescence (Fig. 4). In generally, reproductive organs of *Brachypodium* autopolyploids have a similar structure to those in diploid progenitors. Flowers of S1 autopolyploids have similar structure to those of their diploid progenitors. Inside the palea and lemma, the stamens and carpel of flowers can be seen quite clearly. Autopolyploids derived from both lines of *B. distachyon (*Bd21 and Bd3-1) have, like diploids, two stamens bearing two small plump anthers, one carpel comprising a big ovary, style and very feathery stigma in the floret (Fig. 4a). Florets of diploids and autopolyploids of *B. stacei* line Bsta5 were similar and contain three stamens with three elongated anthers and one carpel that comprises ovary, style and sigma which were more elongated than those of Bd21 (Fig. 4b). Diploid, autotetraploids and also aneuploids of the other line of *B. stacei* (ABR114) have also similar indistinguishable structure (Fig. 4c).

We compared pollen abundance and pollen viability in aneuploids (ABR114_2 (3-1), with 42 chromosome) and autopolyploids (ABR114_2 (4-9)), obtained from ABR114 line of *B. stacei*, as well as those of diploid plants (Online Resource 4). The pollen abundance was much lower in aneuploids in comparison to diploids and autopolyploids. For the aneuploid plant ABR114_2 (3-1), only ~50 pollens/30mm^2 were counted on average, while pollen abundance was more than ~225 pollens/30mm^2 and ~160 pollen/30mm^2 for diploids and autopolyploids, respectively (Online Resource 4).

The test of pollen viability was performed on the flowers the day of anthesis by coloration with the acetocarmine (Online Resource 4). A total of 78% of viable pollens were observed for the diploid line ABR114, 46% for the derived autopolyploid line ABR114_2 (4-9) and only 40% for the aneuploid line ABR114_2 (3-1).

Quantitative comparison of inflorescences, fertility and flag leaves characters between autopolyploids and diploids

The second generation S2 seeds were derived from the true autopolyploid S1 plants of Bd21, Bd3-1, ABR114 (ABR114_2, T4 type and ABR114_8) and Bsta5 lines and S2 plants were grown in greenhouse and compared with their direct diploid progenitors for 16 morphological traits (Online Resource 2, 5). Individual S2 sister plants of each of these different true autopolyploids were homogenous when compared between each other indicating a high stability with no apparent genomic changes.

B. distachyon autopolyploids showed bigger inflorescences and flowers and decreased fertility than diploids. There were significant differences between diploids and autopolyploids of Bd21 in 12 out of 16 analyzed traits (Online Resource 5). The flag leaf and traits of inflorescence and floret had greater mean values in autopolyploids than diploids. On the contrary seed number per florets per spike and number of fertile flowers were significantly lower in autopolyploid plants of Bd21 than in diploid ones. Different results were found for Bd3-1, with nine out of sixteen characters significantly different between autopolyploid and diploid plants of this line (Online Resource 2, 5). There were no differences between the flag leaf characters but, distance between two spikelets, upper glume length, upper glume width, floret

length, lemma length, and lemma width were greater in autopolyploids than in diploids of line Bd3-1.The comparison between the two *B. distachyon* autopolyploids shows that they were significantly different for ten out of sixteen traits (Online Resource 2, 5).

In *B. stacei* autopolyploids, significant differences were also observed between ABR114 and Bsta5 autopolyploids. ABR114 autopolyploids expressed more significant differences with its diploid progenitor than Bsta5 ones. In ABR114_2 T4 type autopolyploids, nine out of 16 traits were significantly different from diploids, seven of these were higher than diploid except seed number per inflorescence and percent of fertile flower were lower than diploid one (Online Resource 2, 5). Comparison between ABR114_8 autopolyploids and diploids showed that eight traits were different to diploid, six of them were higher and two ones were significant lower than diploid. Interestingly, five traits were different between S2 autopolyploid plants of ABR114_2 T4 and ABR114_8, although both of them are derived from same diploid ABR114 line (Online Resource 2, 5). Three of these (upper glume length and upper glume width and lemma width) were higher in ABR114_2 than in ABR114_8 and these were similar between ABR114_8 diploids and autopolyploids.

Only five characters were significantly higher and one was lower (the percent of fertile flower) in Bsta5 autopolyploids as compared to diploids. Both *B. distachyon* and *B. stacei* autopolyploids showed decrease in fertility in comparison with their respective diploid progenitors, illustrated by the lower number of seeds per inflorescence and lower percent of fertile florets (Online Resource 2, 5).

Discussion

In this work, we have generated, by colchicine treatment, autopolyploids from two *Brachypodium* species. We focused on *B. distachyon* and *B. stacei* that have common evolutionary history, but get since they diverged ~16 MYA, completely different chromosome evolution, leading to *B. distachyon* having two times less chromosomes (2n=10) that are two times bigger than *B. stacei* (Hasterok et al. 2004; Catalan et al. 2016). Moreover these two species naturally hybridized together, more than one time, to give the natural allopolyploid species *B. hybridum,* constituting all together an interesting polyploid model (Catalan et al. 2014; Catalan et al. 2016).

It has been shown in a number of species that plants that regenerate after colchicine treatment could show more frequently a mixture of cells with either doubled or diploid chromosomes, leading to S1 progenies either with doubled genome or that remain diploids (Blakeslee and Avery 1937; Dermen 1940; Nguyen et al. 2003; Dahanayake et al. 2010). Therefore, it was important to analyze individual S1plants, progenies of individual regenerated colchicine treated plants. While flow cytometry gave a rough and primary indication of genome size increase, more precise cytogenetic characterizations, such as karyotyping of DAPI-stained chromosomes and FISH hybridization were particularly important here to confirm autopolyploids and elucidate aneuploids. Diploid and autopolyploid S1 plants were obtained, at different frequencies, in progenies of colchicine-treated plants of the two studied lines of *B. distachyon* (Bd21 and Bd3-1) and those from lines Bsta5 and TE4.3 and the plant (ABR114_8) of line ABR114 of *B. stacei*. All these analyzed S1 autopolyploid plants and those of their S2 subsequent generation show stable phenotypes and stable chromosome structure. This was not the case of the 15 S1plants, progenies of

another colchicine-treated plant (ABR114_2) of line ABNR114 of *B. stacei*, all of which showed genome size increase, with strangely no existence of sister diploid plants. Only one of the S1 plants (ABR114_2 (4-9)), was shown to be autotetraploids, with stable phenotype and genome structure. The remaining 14 S1 plants show a mixture of cells with variable number of chromosomes ranging between 38 to 42, suggesting aneuploidy, with also phenotypes and karyotypes that continue to vary in plants of the subsequent S2 generation.

There are several explanations for the formation of aneuploids following colchicine treatments. Zhao and Davidson (1985) proposed that after three hours of colchicine treatment, there is formation of binucleate or multinucleate cells. The position of binucleate and multinucleate has effects on chromosome movement at anaphase and telophase. When binucleate divide, chromosomes may move outside the limits of the spindle region into outlying areas of cytoplasm, which lead to aneuploidy (Davidson et al. 1983).

The optimum colchicine concentration and duration the treatment is variable, depending on the species and tissue types (Müntzing and Runquist (1939). The success of treatment is normally indicated by the abnormal phenotype of the colchicine-treated regenerated plants (Müntzing and Runquist 1939). Plants of *Linum* and *Solanum* genus were more successful with the treatment of 0.025% colchicine for 2 to 4 days, whereas the potato autotetraploid was obtained by the treatment of 0.05% colchicine for 6 days. In *Salvia miltiorrhiza*, polyploids were induced by *in vitro* culture in basal MS medium containing colchicine, the optimal concentration of colchicine for this treatment is 10ppm (Gao et al. 1996). The comparison of different

colchicine concentrations, which were used on watermelon seedlings, showed that the number of autopolyploids induced is higher at 0.2% of colchicine (Jaskani et al. 2005). In *Vicia villosa* (hairy vetch) tetratraploid plants (2n=4x=28) were obtained when plants were treated at 0.005% colchicine concentration, whereas no autopolyploids were obtained at 0.002% colchicine-treated plants (Tulay and Unal 2010).

Our result showed that most of analyzed characters of *Brachypodium* autopolyploids were higher than those of their diploid progenitors. In *Hylocereus* autopolyploids, the bigger size of morphological traits was related to increase of amino acid content, tricarboxylic acid cycle intermediates, organic acids and flavonoids (Cohen et al. 2013). Similarly, the *Siraitia grosvenorii* autopolyploids were more vigourous than their diploid progenitors (Fu et al. 2012). Stomata bigger size and low density is considered as a signal to recognize ploidy level increase. The increase of stomata size was significantly accompanied by the decrease in their density in plants of higher ploidy level (Cohen et al. 2013). In *Dianthus broteri,* stomata size analysis demonstrated the relationship between morphology and ploidy level in polyploids and their progenitors (Balao et al. 2011).

Structural changes occur in early generations of newly-generated autotetraploids. Pairing between more than two homologous chromosomes at meiosis, leading to multivalent formation (Santos et al. 2003) and translocation of 45S rDNA from chromosome 4 to chromosome 3 have been reported in *Arabidopsis thaliana* (Weiss and Maluszynska 2000). Changes in gene expression and DNA methylation were also reported to occur in early generations of synthesized autopolyploids of Arabidopsis (Yu et al. 2010), *Paspalum notatum* (Martelotto et al. 2005) and *Brassica oleracea* (Albertin et al. 2005). Changes in autopolyploid have been also studied at

the gene expression level. In *Paulownia fortunei*, Zhang et al. (2014) indicated that differentially expressed transcripts related to energy metabolism and carbon fixation were enriched; most of them were upregulated in autopolyploids.

All *Brachypodium* autopolyploids studied here, especially those of *B. stacei*, have reduced seed number per inflorescence and fertile flower ratios (46%, 44% and 39% in ABR114_2 T4, ABR114_8 and Bsta5, respectively) as compared to their corresponding diploids. The reducing of fertility is one common feature observed in several other characterized autopolyploids from different species including *Arabidopsis* (Weiss and Maluszynska 2000; Hollister et al. 2012) and *Guizotia abissinica* (Dagne 2001). Most studies considered that the irregularities in meiosis of autopolyploids, which may be related to multivalent instead of bivalent formation, usually lead to fertility reduction (Ramsey and Schemske 2002; Parisod et al. 2010; Grandont et al. 2013; Lloyd and Bomblies 2016). However, other evidences suggested that irregular chromosome behaviors may not have effect on pollen viability and may not lead to the low seed production (Ortiz et al. 2011)

In *Brachypodium* genus, natural *B. hybridum* allopolyploid has been formed by hybridization between *B. distachyon* and *B. stacei*. With the objective of resynthesizing allopolyploids from hybridization between *B. distachyon* and *B. stacei*, generating of autopolyploids from these two species may be of great help.

This is the first time that synthetic autopolyploids from *Brachypodium* species are reported. These polyploids provide a starting material to study the fate of homologous duplicated genes, as well as genetic, functional and epigenetic consequences of autopolyploidy and regulation of meiosis in autopolyploids, comparatively in *B.*

distachyon and *B. stacei* species that have similar genome context but asymmetric evolution of chromosomes.

Acknowledgments

This work was supported by a French ANR (Agence Nationale de la Recherche, www.agence-nationale-recherche.fr), 2011 grant (ANR-11-BSV5-0015: Ploid-ploid-Wheat). Vinh Ha Dinh Thi was support by a PhD fellowship from the Vietnamese International Education Development (VIED) and Ministry of Education and Training Vietnam (MOET).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Legend to Electronic Supplementary Material

Online Resource 1. Colchicine treatment and effects on *Brachypodium* seedlings (adapted from Jahier (1992)).

(a) Seedlings at three leaves stage were taken out from soil and washed with water. (b) Seedlings were immersed in colchicine solution at (5g/l) concentration for 3h at room temperature. (c) Plants were then transplanted into soil pots. (d) A plant recovering 10 days after colchicine treatment. Burn dots on the leaves (red arrow) new emerging tillers (blue arrow). (e) Colchicine-treated plant (right) showing delayed flowering time in comparison to the control non-treated one (left) that is already at the flowering stage (red arrow). *B. distachyon* Bd21 is shown in in this example.

Online Resource 2. Mean values ± standard deviation for 16 morphological traits measured in different synthetic and natural *Brachypodium* autopolyploids and their parental species.

NSI: Spikelet number per inflorescence, NFS: Floret number per spikelet, NFI: Floret number per inflorescence, SI: Seed number per inflorescence, PFF: Percent of fertile florets, FLL: Flag leaf length, FLW: Flag leaf width, IL: Inflorescence length, SL: Spikelet length, DS: Distance between two spikelets of inflorescence, UGL: Upper glume length (from basal spikelet), UGW: Upper glume width (from basal spikelet), FL: Floret length (the second floret of all spikelets were taken to measure), LL: Lemma length (from basal floret), LW: Lemma width (from basal floret), AL: Awn length.

Online Resource 3. Illustration of phenotypic characters recorded in the present

study.

1. Inflorescence length; 2. Spikelet length; 3. Distance between two spikelets in florescence; 4. Upper glume from basal spikelet length; 5. Upper glume from basal spikelet width; 6. Floret length (the second floret of all spikelets were taken to measure); 7. Awn length; 8. Lemma length; 9. Lemma width; 10. Flag leaf length; 11. Flag leaf width.

Other characters considered for inflorescence or spikelet:

12. Spikelet number per inflorescence (all the spikelets in spike - 5 in this example); 13. Floret number per spikelet (i.e. the sum of all florets in each spikelet); 14. Florets number per inflorescence (i.e. the sum of all floret in all spikelet of inflorescence); 15. Seed number per inflorescence; 16. Percentage of fertile floret.

Online Resource 4. Test comparisons for pollens in anthers, pollen abondance and viability, between the diploid *B. stacei* line ABR114 and the derived autopolyploid plant (ABR114_2 (4-9), S1 generation) and the aneuploid plant ((ABR114_2 (3-1) with 42 chromosomes, S1 generation)**.**

(a) Entire anthers showing very low proportion of pollens in aneuploids as compared to diploids and autopolyploids. (b) Pollen abundance. Pollens dehisce from one anther per plant were homogenously spread on a glass slide (size of 25 mm × 75 mm). The number of pollens was then counted under microscope (Leica DMLB) in square millimeter (30mm^2) units and averaged per one mm^2. (c) Pollen viability test by acetocarmine coloration. Viable pollens are stained by dark purple and unviable ones are stained into light purple.

Online Resource 5. Morphological trait comparison between diploid lines Bd21and Bd3-1 of *B. distachyon* and the diploid lines ABR114 and Bsta5 of *B. stacei* and their derived autopolyploids (names of autopolyploid lines are indicated). Vertical bars indicate standard errors. Five biological replicates were considered.

See Online Resource 3 that shows images of the measured traits and the data in Online Resource 2.

References

Adams KL, Wendel JF (2005) Polyploidy and genome evolution in plants. Curr Opin Plant Biol 8(2):135-141.

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

Balao F, Herrera J, Talavera S (2011) Phenotypic consequences of polyploidy and genome size at the microevolutionary scale: a multivariate morphological approach. New Phytol 192(1):256-265.

Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA (2015) On the relative abundance of autopolyploids and allopolyploids. New Phytol.

Bartels GP, Hilton JL (1973) Comparison of trifluralin, Oryzalin, Pronamide, Propham and Colchicine treatment on microtubules. Pesticide Biochemistry and Physiology 3:462-472.

Beest M, Le Roux JJ, Richardson DM, Brysting AK, Suda J, Kubešová M, Pyšek P (2011) The more the better? The role of polyploidy in facilitating plant invasions. Annals of Botany 109(1):19-45.

Betekhtin A, Jenkins G, Hasterok R (2014) Reconstructing the Evolution of Brachypodium Genomes Using Comparative Chromosome Painting. PLoS One 9(12):e115108.

Blakeslee AF, Avery AG (1937) Conchicine and double diploids. Journal of Heredity 28(12):411-412.

Blakeslee AF, Avery AG (1938) Methods of inducing doubling of chromosomes in plants: Washington, D.C., Carnegie Institution of Washington.

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

Bradley MV, Goodspeed TH (1943) Colchicine-Induced Allo- and Autopolyploidy in Nicotiana. Proceedings of the National Academy of Sciences of the United States of America 29(10):295-301.

Cai D, Rodríguez F, Teng Y, Ané C, Bonierbale M, Mueller LA, Spooner DM (2012) Single copy nuclear gene analysis of polyploidy in wild potatoes (Solanum section Petota). BMC evolutionary biology 12(1):1-16.

Caperta AD, Delgado M, Ressurreicao F, Meister A, Jones RN, Viegas W, Houben A (2006) Colchicine-induced polyploidization depends on tubulin polymerization in c-metaphase cells. Protoplasma 227(2-4):147-153.

Catalan P, Chalhoub B, Chochois V, Garvin DF, Hasterok R, Manzaneda AJ, Mur LA, Pecchioni N, Rasmussen SK, Vogel JP and others (2014) Update on the genomics and basic biology of Brachypodium: International Brachypodium Initiative (IBI). Trends Plant Sci 19(7):414-418.

Catalan P, López-Álvarez D, Díaz-Pérez A, Sancho R, López-Herránz ML (2016) Phylogeny and Evolution of the Genus Brachypodium. In: Vogel PJ, editor. Genetics and Genomics of Brachypodium. Cham: Springer International Publishing. p 9-38.

Childs KL, Nandety A, Hirsch CN, Gongora-Castillo E, Schmutz J, Kaeppler SM, Casler MD, Buell CR (2014) Generation of Transcript Assemblies and Identification of Single Nucleotide Polymorphisms from Seven Lowland and Upland Cultivars of Switchgrass. Plant Genome 7(2).

Cohen H, Fait A, Tel-Zur N (2013) Morphological, cytological and metabolic consequences of autopolyploidization in Hylocereus (Cactaceae) species. BMC Plant Biology 13(1):173.

Cohen H, Tel-Zur N (2011) Morphological changes and self-incompatibility breakdown associated with autopolyploidization in Hylocereus species (Cactaceae). Euphytica 184(3):345-354.

Comai L (2005) The advantages and disadvantages of being polyploid. Nat Rev Genet 6(11):836-846.

Consortium TPGS (2011) Genome sequence and analysis of the tuber crop potato. Nature 475(7355):189-195.

Cui L, Wall PK, Leebens-Mack JH, Lindsay BG, Soltis DE, Doyle JJ, Soltis PS, Carlson JE, Arumuganathan K, Barakat A and others (2006) Widespread genome duplications throughout the history of flowering plants. Genome Res 16(6):738-749.

Dagne K (2001) Meiotic properties of induced autopolyploid Guizotia abyssinica (Lf) Cass. Journal of genetics and breadding 55(1):11-16.

Dahanayake N, Chen X-L, Zhao F-C, Yang Y-S, Wu H (2010) Separation of tetraploid and diploid plants from chimeras in in-vitro cultures of purple coneflower (Echinacea Purpurea L). Tropical Agricultural Research & Extension 13(1).

Darlington CD (1937) Recent advances in cytology. Second edition. Philadelphia. P. Blakiston's son and Co., 1937. xvi + 671 pp. The Anatomical Record 69(1):123-126.

Davidson D, Pertens E, Zhao JP (1983) Chromosome distribution between two restitution nuclei in a cell following colchicine treatment. Canadian Journal of Genetics and Cytology 25(5):437-445.

Dermen H (1940) Colchicine Polyploidy and Technique. Botanical Review 6(11):599-635.

Dermen H (1945) The Mechanism of Colchicine-Induced Cytohistological Changes in Cranberry. American Journal of Botany 32(7):387-394.

Eigsti OJ (1938) A Cytological Study of Colchicine Effects in the Induction of Polyploidy in Plants. Proceedings of the National Academy of Sciences of the United States of America 24(2):56-63.
Eigsti OJ, Dustin P (1955) Colchicine in agriculture, medicine, biology, and chemistry Ames,Iowa State College Press.
Fu W, Ma X, Tang Q, Mo C (2012) Karyotype analysis and genetic variation of a mutant in Siraitia grosvenorii. Molecular biology reports 39(2):1247-1252.
Gao SL, Zhu DN, Cai ZH, Xu DR (1996) Autotetraploid plants from colchicine-treated bud culture ofSalvia miltiorrhiza Bge. Plant Cell, Tissue and Organ Culture 47(1):73-77.
Gerlach W., J. B (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. . Nucleic Acids Res 7:1869-1885.
Grandont L, Jenczewski E, Lloyd A (2013) Meiosis and its deviations in polyploid plants. Cytogenetic and genome research 140(2-4):171-184.
Grant V (1963) The orgin of adaptations: Colombia University Press. 606 p.
Gupta PP (1981) Suppression of multivalent formation by B chromosomes in natural and artificial autopolyploids of scurvy-grass (Cochlearia L.). Theor Appl Genet 59(4):221-223.
Hasterok R, Draper J, Jenkins G (2004) Laying the cytotaxonomic foundations of a new model grass, Brachypodium distachyon (L.) Beauv. Chromosome Res 12(4):397-403.
Hasterok R, Marasek A, Donnison IS, Armstead I, Thomas A, King IP, Wolny E, Idziak D, Draper J, Jenkins G (2006) Alignment of the genomes of Brachypodium distachyon and temperate cereals and grasses using bacterial artificial chromosome landing with fluorescence in situ hybridization. Genetics 173(1):349-362.
Havananda T, Brummer EC, Doyle JJ (2011) Complex patterns of autopolyploid evolution in alfalfa and allies (Medicago sativa; Leguminosae). Am J Bot 98(10):1633-1646.
Hollister JD, Arnold BJ, Svedin E, Xue KS, Dilkes BP, Bomblies K (2012) Genetic adaptation associated with genome-doubling in autotetraploid Arabidopsis arenosa. PLoS genetics 8(12):e1003093.
Jahier J (1992) Techniques of plant cytogenetics: INRA Edition, Paris, France.
Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C and others (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449(7161):463-467.
Jaskani MJ, W.Kwon S, Kin DH (2005) Flow cytometry of DNA contents of colchicine treated watermelon as a ploidy screening method at M1 stage. Pak J Bot 37(3):685-696.
Kruskal WH, Wallis WA (1952) Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association 47(260):583-621.
Levy AA, Feldman M (2002) The impact of polyploidy on grass genome evolution. Plant Physiol 130(4):1587-1593.
Lloyd A, Bomblies K (2016) Meiosis in autopolyploid and allopolyploid Arabidopsis. Current Opinion in Plant Biology 30:116-122.
Lopez-Alvarez D, Lopez-Herranz ML, Betekhtin A, Catalan P (2012) A DNA barcoding method to discriminate between the model plant Brachypodium distachyon and its close relatives B. stacei and B. hybridum (Poaceae). PLoS One 7(12):e51058.
Martelotto LG, Ortiz JPA, Stein J, Espinoza F, Quarin CL, Pessino SC (2005) A comprehensive analysis of gene expression alterations in a newly synthesized Paspalum notatum autotetraploid. Plant Science 169(1):211-220.
Mestiri I, Chagué V, Tanguy AM, Huneau C, Huteau V, Belcram H, Coriton O, Chalhoub B, Jahier J (2010) Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity. New Phytol 186(1):86-101.
Müntzing A, Runquist E (1939) Note on some colchicine-induce polyploids. Hereditas 25(4):491-495.
Nguyen TPT, Ureshino K, Miyajima I, Ozaki Y, Okubo H (2003) Induction of tetraploids in ornamental Alocasia through colchicine and oryzalin treatments. Plant Cell, Tissue and Organ Culture 72(1):19-25.
Ortiz AM, Seijo JG, Fernández A, Lavia GI (2011) Meiotic behavior and pollen viability of tetraploid Arachis glabrata and A. nitida species (Section Rhizomatosae, Leguminosae): implications concerning their polyploid nature and seed set production. Plant systematics and evolution 292(1-2):73-83.
Parisod C, Holderegger R, Brochmann C (2010) Evolutionary consequences of autopolyploidy. New Phytol 186(1):5-17.
Ramsey J, Schemske DW (2002) Neopolyploidy in Flowering Plants. Annual Review of Ecology and Systematics 33(1):589-639.
Santos JL, Alfaro D, Sanchez-Moran E, Armstrong SJ, Franklin FC, Jones GH (2003) Partial diploidization of meiosis in autotetraploid Arabidopsis thaliana. Genetics 165(3):1533-1540.
Saski CA, Li Z, Feltus FA, Luo H (2011) New genomic resources for switchgrass: a BAC library and comparative analysis of homoeologous genomic regions harboring bioenergy traits. BMC Genomics

12(1):1-12.
Schlueter JA, Dixon P, Granger C, Grant D, Clark L, Doyle JJ, Shoemaker RC (2004) Mining EST databases to resolve evolutionary events in major crop species. Genome 47(5):868-876.
Simonsen Ø (1973) Cyto-genetic investigations in diploid and autotetraploid populations of Lolium perenne L. Hereditas 75(2):157-188.
Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. Annu Rev Plant Biol 60:561-588.
Sugiyama S (2005) Polyploidy and cellular mechanisms changing leaf size: comparison of diploid and autotetraploid populations in two species of Lolium. Ann Bot 96(5):931-938.
Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH (2008) Synteny and collinearity in plant genomes. Science 320(5875):486-488.
Tel-Zur N, Dudai M, Raveh E, Mizrahi Y (2011) In situ induction of chromosome doubling in vine cacti (Cactaceae). Scientia Horticulturae 129(4):570-576.
Tulay E, Unal M (2010) Production of colchicine induced tetraploids in Vicia villosa roth. Caryologia 63(3):292-303.
Upcott M (1939) The nature of tetraploidy inPrimula kewensis. Journal of genetics 39(1):79-100.
Van de Peer Y, Maere S, Meyer A (2009) The evolutionary significance of ancient genome duplications. Nat Rev Genet 10(10):725-732.
Varoquaux F, Blanvillain R, Delseny M, Gallois P (2000) Less is better: new approaches for seedless fruit production. Trends Biotechnol 18(6):233-242.
Vaughn KC, Lehnen LP (1991) Mitotic Disrupter Herbicides. Weed Science 39(3):450-457.
Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, Fontana P, Bhatnagar SK, Troggio M, Pruss D and others (2010) The genome of the domesticated apple (Malus [times] domestica Borkh.). Nat Genet 42(10):833-839.
Vogel JP, Tuna M, Budak H, Huo N, Gu YQ, Steinwand MA (2009) Development of SSR markers and analysis of diversity in Turkish populations of Brachypodium distachyon. BMC Plant Biol 9:88.
Weiss H, Maluszynska J (2000) Chromosomal rearrangement in autotetraploid plants of Arabidopsis thaliana. Hereditas 133(3):255-261.
Wen R, Moore G, Shaw PJ (2012) Centromeres cluster de novo at the beginning of meiosis in Brachypodium distachyon. PLoS One 7(9):e44681.
Wu J-H, Ferguson AR, Murray BG, Jia Y, Datson PM, Zhang J (2012) Induced polyploidy dramatically increases the size and alters the shape of fruit in Actinidia chinensis. Annals of Botany 109(1):169-179.
Yu Z, Haberer G, Matthes M, Rattei T, Mayer KF, Gierl A, Torres-Ruiz RA (2010) Impact of natural genetic variation on the transcriptome of autotetraploid Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America 107(41):17809-17814.
Zhang X, Deng M, Fan G (2014) Differential transcriptome analysis between Paulownia fortunei and its synthesized autopolyploid. International journal of molecular sciences 15(3):5079-5093.
Zhao JP, Davidson D (1985) Distribution of Chromosomes into Discrete Groups in Colchicine-Induced C-Metaphases of Barley. Caryologia 38(2):195-206.

Figure Captions

Fig. 1 Flow cytometry (FCM) DNA-histograms of a diploid plant of *B. stacei*, line ABR114, with two peaks corresponding to DNA content in nuclei at G1 (2C) and G2/M (4C) phases of mitosis (a), as compared to the plant (ABR114_8) from first selfed generation (S1), derived from a colchicine-treated plant of the same line (b).

Fig.2 Plant phenotype (a) and spike morphology (b) variation of first selfed generation (S1) plants, derived from the colchicine treated plant ABR114_2 and showing a double genome size increase as compared to a diploid non-treated plant (2x).

Plants ids: T1: (ABR114_2 (3-10) T2: (ABR114_2 (3-1)), T3: (ABR114_2 (3-5)), T4: (114_2 (4-9)), 2x: the diploid non-treated plant. Secondary and small spikes emerging at the nodes are shown for T2 type plant.

Fig. 3 Chromosome characterizations and counting in cells at metaphase stage of of diploid plants of *B. distachyon* and *B. stacei* as well as plants from first selfed generation (S1) derived from colchicine-treated plants and showing doubled genome size increase by flow cytometry.

(a) Ten chromosomes stained by DAPI (4',6-diamidino-2-phenylindole) are observed in the diploid plant of *B. distachyon* (line Bd21) and hybridization with 45S rDNA probe shows two 45S loci (green). (b) Twenty chromosomes are found in a S1 plant of line Bd21 and four 45S loci, demonstrating that it is autopolyploid. (c) Twenty chromosomes are found in diploid *B. stacei* lineABR114 and two 45S loci. (d) Forty chromosomes are found in the T4 type S1 plant of the same line of *B. stacei*

(ABR114_2 (4-9)) and hybridization with 45S probe exhibited four green signals of 45S loci, demonstrating that it is autopolyploid. E-H were stained by DAPI and hybridized with the centromeric sequence probe CentBd from *B. distachyon* to confirm chromosome number. (e-f): Cells with 38 and 41 chromosomes in a T1 type S1 plant of the same line of *B. stacei* (ABR114_2 (3-10)). (g) Cells with 42 chromosomes in T2 (ABR114_2 (3-1). (h): Cells wilh 41chromosomes in T3 (ABR114_2 (3-5)) S1 plants of the same line of *B. stacei*.

Fig. 4 Flower structure of diploid and autopolyploid (first selfed generation) plants of *B. distachyon* (line Bd3-1) and *B. stacei* (lines ABR114 and Bsta4) as well as a S1 aneuploid plant derived from the plant ABR114_2 of *B. stacei*.

(a): Both diploids and autopolyploid plants of line Bd3-1 of *B. distachyon* possess the same structure of flowers with a very feathery stigma and two stamens in the flower.

(b and c): Plants of both diploid lines of *B. stacei*, the autopolyploid S1 plant (Bsta5_3_GII) derived from line Bsta5, the autopolyploid S1 plant (ABR114_8) and the aneuploid S1 plant (ABR114_2 (3-1)) derived from line ABR114 show no apparent differences in flower structure that have three stamens.

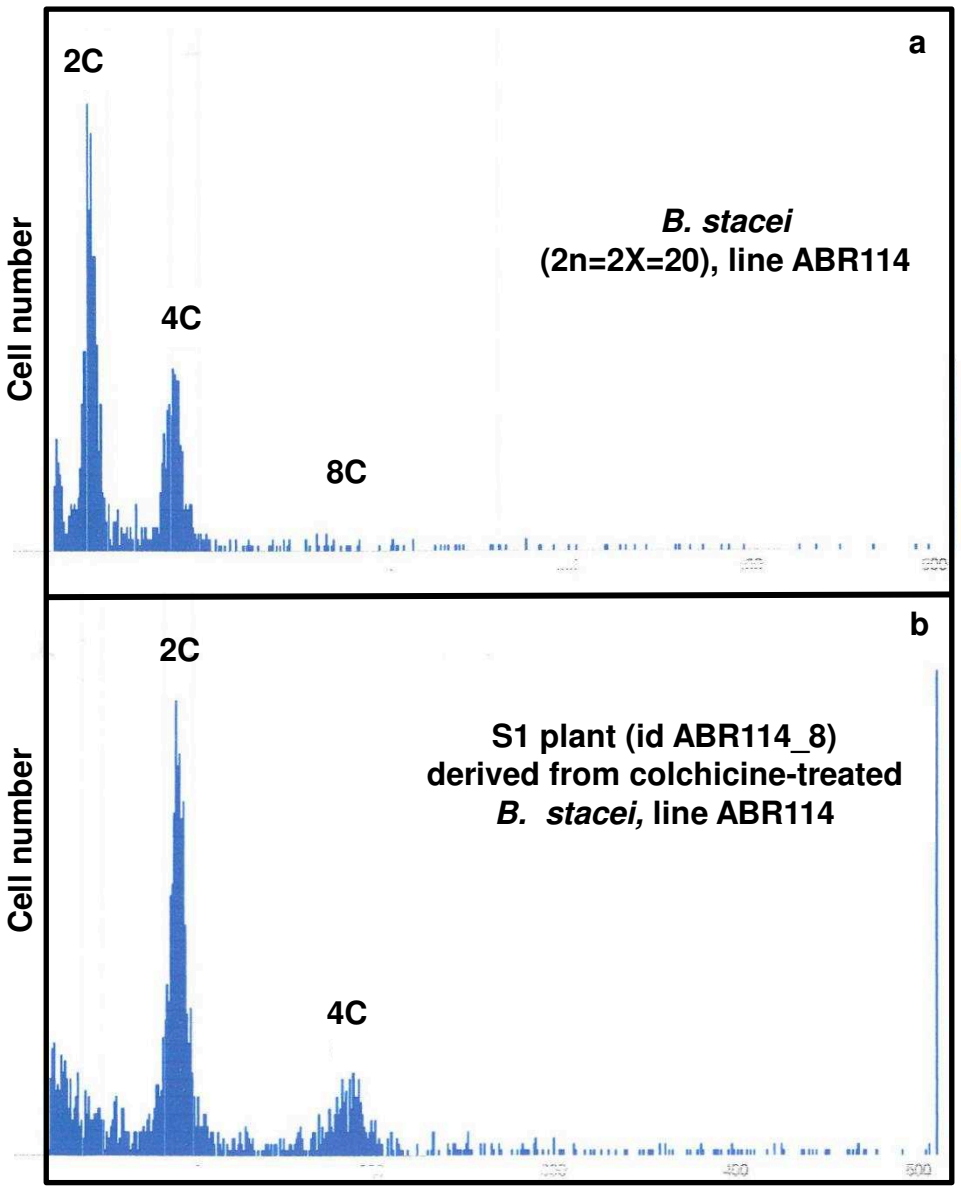

a
2C
4C
8C
Cell number
B. stacei
(2n=2X=20), line ABR114
b
2C
4C
Cell number
S1 plant (id ABR114_8)
derived from colchicine-treated
B. stacei, line ABR114
400
500
Nuclear DNA content

FIGURE 1

a

b

Fig. 2

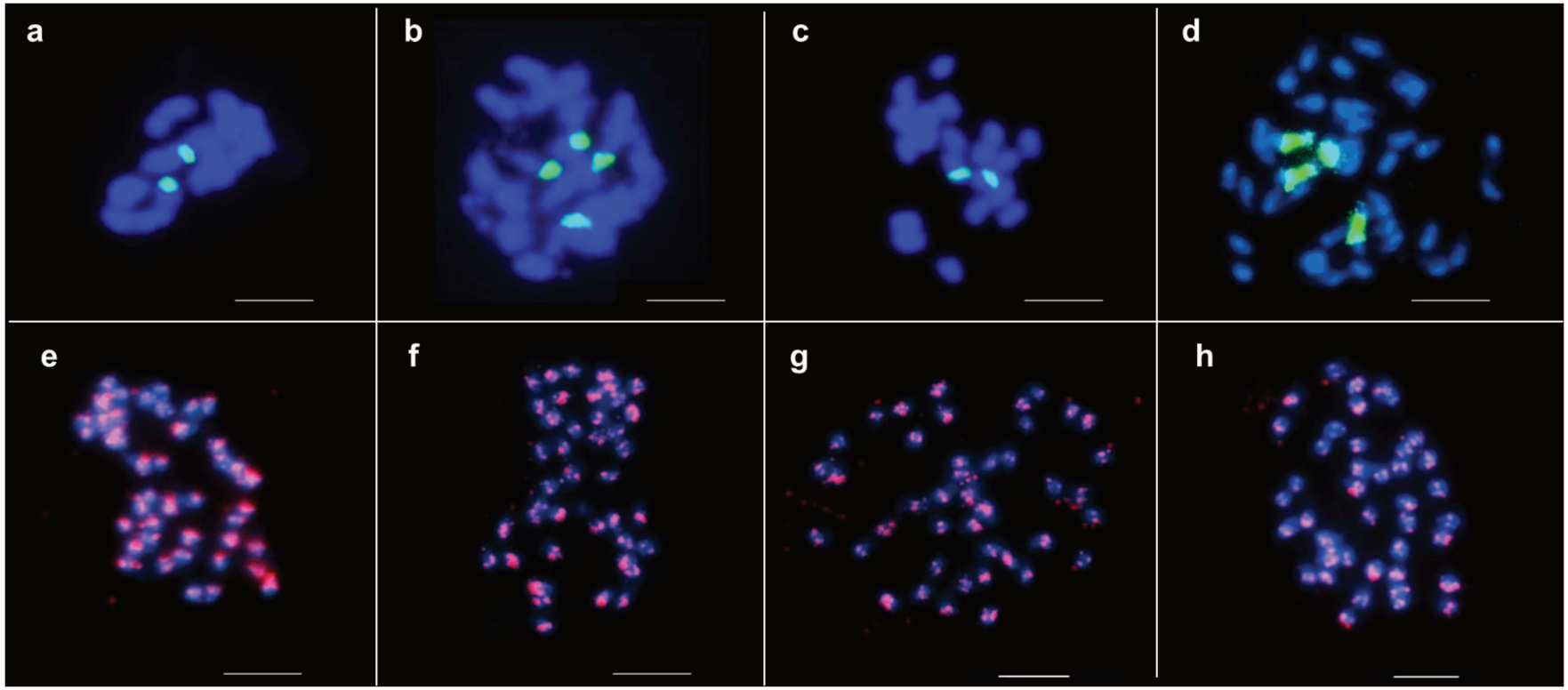


Fig. 3

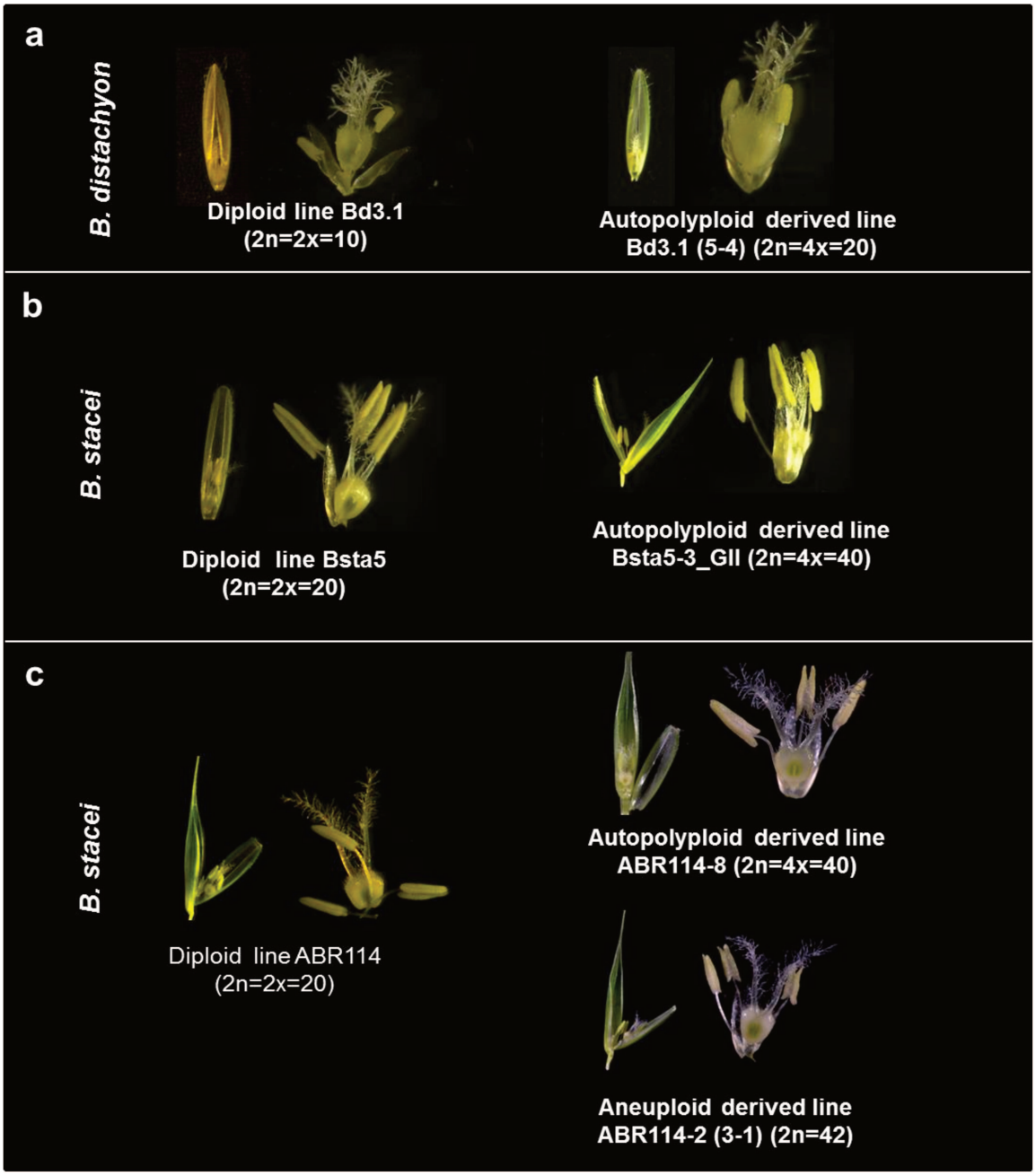

a
B. distachyon
Diploid line Bd3.1
(2n=2x=10)
Autopolyploid derived line
Bd3.1 (5-4) (2n=4x=20)
b
B. stacei
Diploid line Bsta5
(2n=2x=20)
Autopolyploid derived line
Bsta5-3_GII (2n=4x=40)
c
B. stacei
Diploid line ABR114
(2n=2x=20)
Autopolyploid derived line
ABR114-8 (2n=4x=40)
Aneuploid derived line
ABR114-2 (3-1) (2n=42)

Fig. 4

Table 1 Summary of treatment with colchicine solution (5g/l) of plants from different lines of *Brachypodium distachyon* and *B. stacei* species.

Species	Inbred line	Genome	Collection location	Treated plants at stages:		Survival plants		Delay of flowering date[a]	Seed collection
				Seedling	Germination	Seedling	Germination		
B. distachyon	Bd21	2n=2x=10	Iraq	30	10	28	8	10	yes
	Bd3-1	2n=2x=10	Iraq	30	10	27	9	20	yes
B. stacei	ABR114	2n=2x=20	Spain	30	10	14	0	7	yes
	Bsta5	2n=2x=20	Spain	5	5	2	0	13	yes
	TE4.3	2n=2x=20	Spain	5	5	3	0	17	yes
	LP6.1	2n=2x=20	Spain	5	5	0	0	-*	-
Total				105	45	74	17		

[a] as compared to flowering date of the non-treated plants of the same line used as control.
* Not available

Table 2 Flow cytometry analysis of first selfed generation (S1) plants, progenies of colchicine-treated plants of *B. distachyon* and *B. stacei.*

Species	Inbred lines	Colchicine-treated plant	Number of plants	Remaining diploids	Doubled number	Doubling rate
B. stacei	ABR114	ABR114_8	66	64	2	3.03%
		ABR114_2	15	0	15	100.00%
	Bsta5	Bsta5_3_G2	11	10	1	9.09%
		Bsta5_G11	8	7	1	12.50%
	TE4.3	TE4.3_2_G2	37	37	0	0.00%
		TE4.3_3_B6	48	47	1	2.08%
		TE4.3_3_H10	79	77	2	2.53%
B. distachyon	Bd21	Bd21_7	24	21	3	12.50%
		Bd21_8	17	14	3	17.65%
	Bd3-1	Bd3-1_5	27	25	2	7.41%
Total			332	302	30	9.04%

Table 3 Fertility, flow cytometry (FMC) and cytogenetic characterization of chosen individual autopolyploid plants (at first selfed S2 generation), of various lines of *Brachypodium distachyon* and *B. stacei*.

Species	Inbred lines	S1 plant id	Fertility[a]	FCM	Karyotype[b]	45S rDNA probe[c]
B. stacei	ABR114	ABR114_2 (3-8) (T1)	-*	Doubled	40(6), 39(1)	4
		ABR114_2 (3-10) (T1)	47.80%	Doubled	40 (4), 38(1)	4
		ABR114_2 (3-1) (T2)	22.00%	Doubled	42(6), 40(6), 41(1)	4
		ABR114_2 (3-5) (T3)	27.50%	Doubled	40(6), 41(2)	4
		ABR114_2 (4-9) (T4)	53.10%	Doubled	40 (4)	4
		ABR114_8 (2-5)	60.20%	Doubled	40 (7)	4
	Bsta5	Bsta5_3_G11	54%	Doubled	40(3)	4
B. distachyon	Bd21	Bd21_8 (2-7)	44%	Doubled	20(6)	4
	Bd21	Bd21_7 (4-9)	55%	Doubled	20(3)	4
	Bd3-1	Bd3-1 (5-4)	46%	Doubled	20(5)	4

* Not analyzed

[a] Fertility: The percent of fertile flowers (having seeds) over number of total flowers.

[b] Chromosome counting in somatic metaphase stage, stained by DAPI (4',6-diamidino-2-phenylindole) and confirmed by fluorescence *in situ* hybridization (FISH) with centromeric probe CCS1 (cereal centromeric sequences). The number of cells having the karyotype is indicated between brackets.

[c] 45S rDNA probe detecting four signals confirming autopolyploid (only two signals are observed in diploids).

Synthesis and evaluation of autotetraploids from *Brachypodium distachyon* and *B. stacei* species

Plant Systematics and Evolution

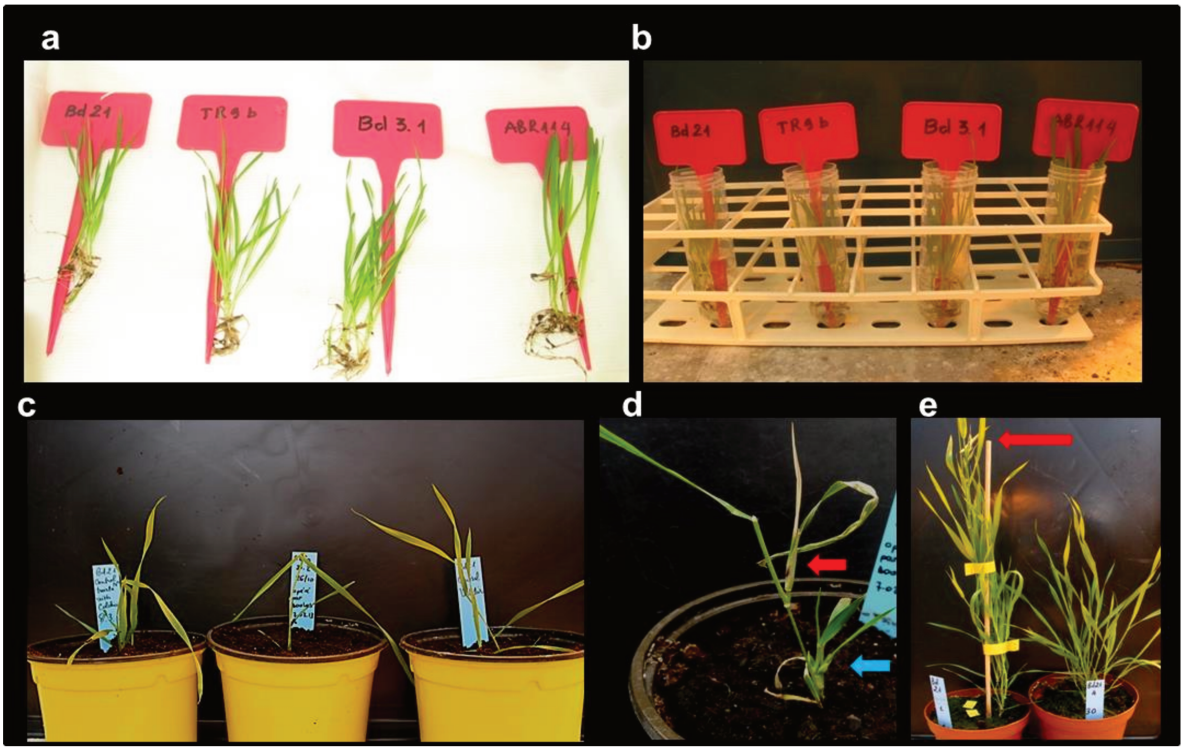
Dinh Thi Vinh-Ha[1], Le Clainche Isabelle[1], Coriton Olivier[2], Jahier Joseph[2], Chalhoub Boulos[1]*

[1]Organization and evolution of plant complex genomes (OECG) Institut National de la Recherche agronomique (INRA), Université d'Evry Val d'Essonne (UEVE), 2 rue Gaston Crémieux, 91057 Evry France.
[2]UMR1349 Institut de Génétique, Environnement et Protection des Plantes (IGEPP), INRA, BP35327, 35653, Le Rheu cedex, France.

** Corresponding author:* Boulos Chalhoub
Organization and evolution of plant complex genomes (OECG), INRA/UEV E,
91057 Evry Cedex, France
Email: chalhoub@evry.inra.fr
Phone: +33 1 60 87 45 03
Fax: +33 1 60 87 45 49

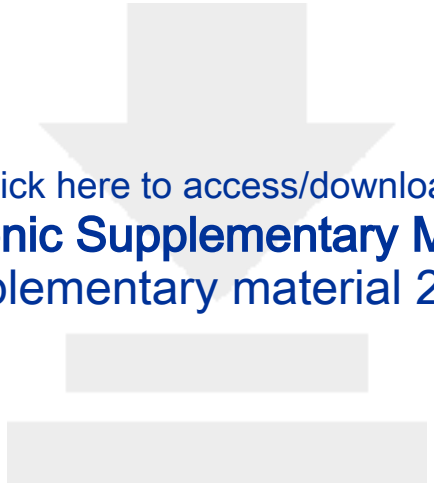
Online Resource 1. Colchicine treatment and effects on *Brachypodium* seedlings (adapted from Jahier (1992)).

(A) Seedlings at three leaves stage were taken out from soil and washed with water. (B) Seedlings were immersed in colchicine solution at (5g/l) concentration for 3h at room temperature. (C) Plants were then transplanted into soil pots. (D) A plant recovering 10 days after colchicine treatment. Burn dots on the leaves (red arrow) new emerging tillers (blue arrow). (E) Colchicine-treated plant (right) showing delayed flowering time in comparison to the control non-treated one (left) that is already at the flowering stage (red arrow). *B. distachyon* Bd21 is shown in C, D, E.



Reference:
Jahier J (1992) Techniques of plant cytogenetics: INRA Edition, Paris, France.

Electronic Supplementary Material2



Click here to access/download

Electronic Supplementary Material

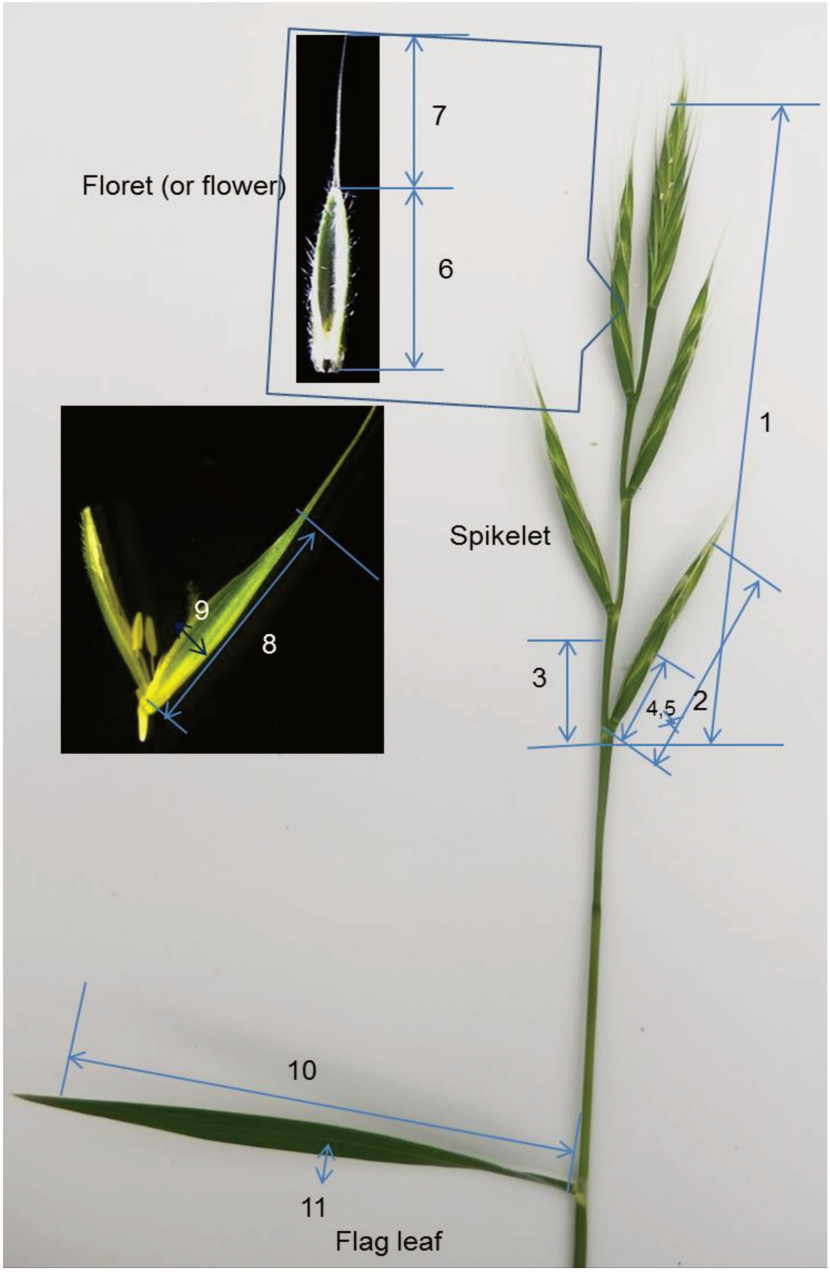
Supplementary material 2.xlsx

Online Resource 3. Illustration of phenotypic characters recorded in the present study.

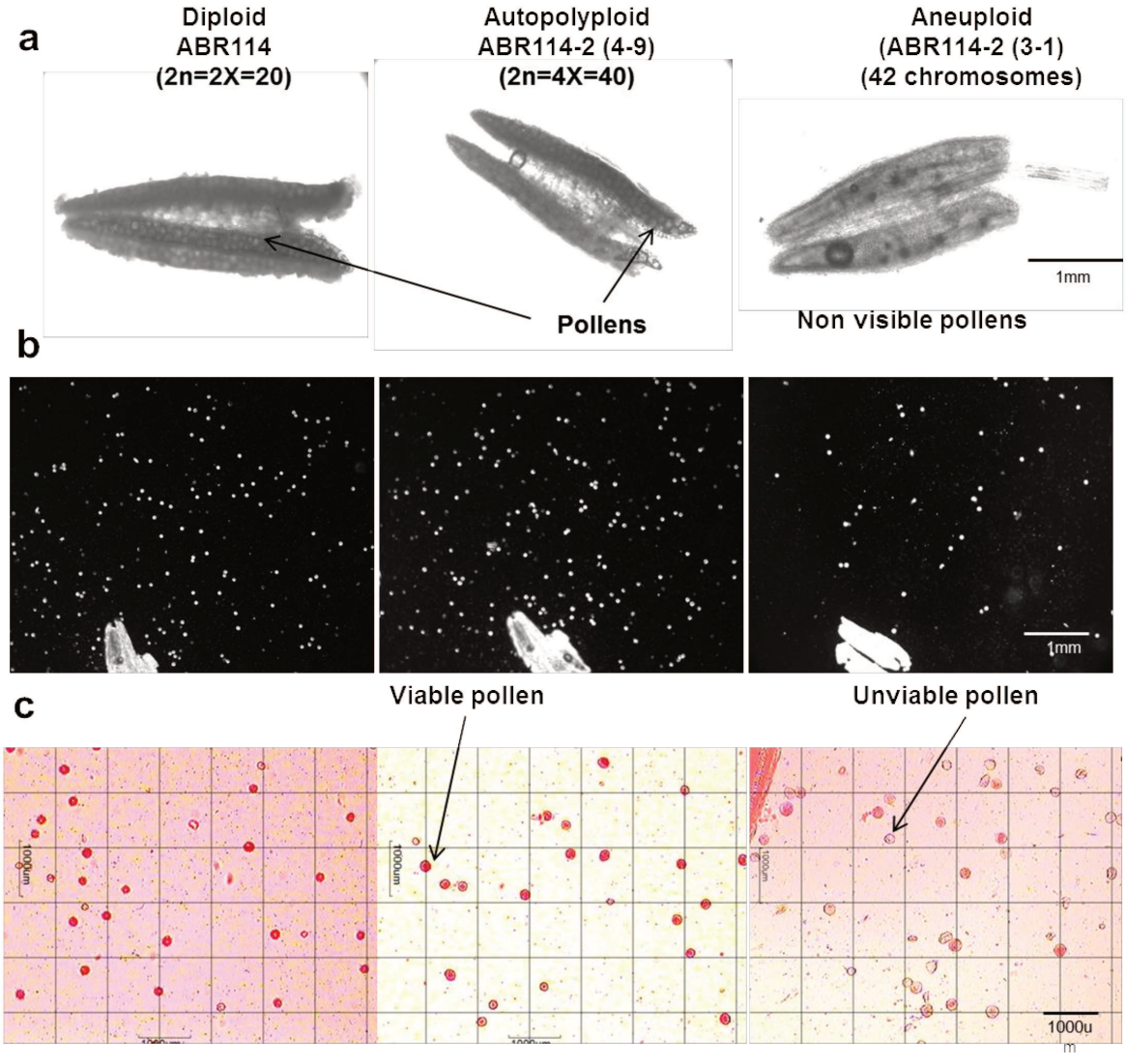
1. Inflorescence length; 2. Spikelet length; 3. Distance between two spikelets in florescence; 4. Upper glume from basal spikelet length; 5. Upper glume from basal spikelet width; 6. Floret length (the second floret of all spikelets were taken to measure); 7. Awn length; 8. Lemma length; 9. Lemma width; 10. Flag leaf length; 11. Flag leaf width.

Other characters considered for inflorescence or spikelet:

12. Spikelet number per inflorescence (all the spikelets in spike - 5 in this example); 13. Floret number per spikelet (i.e. the sum of all florets in each spikelet); 14. Florets number per inflorescence (i.e. the sum of all floret in all spikelet of inflorescence); 15. Seed number per inflorescence; 16. Percentage of fertile floret.

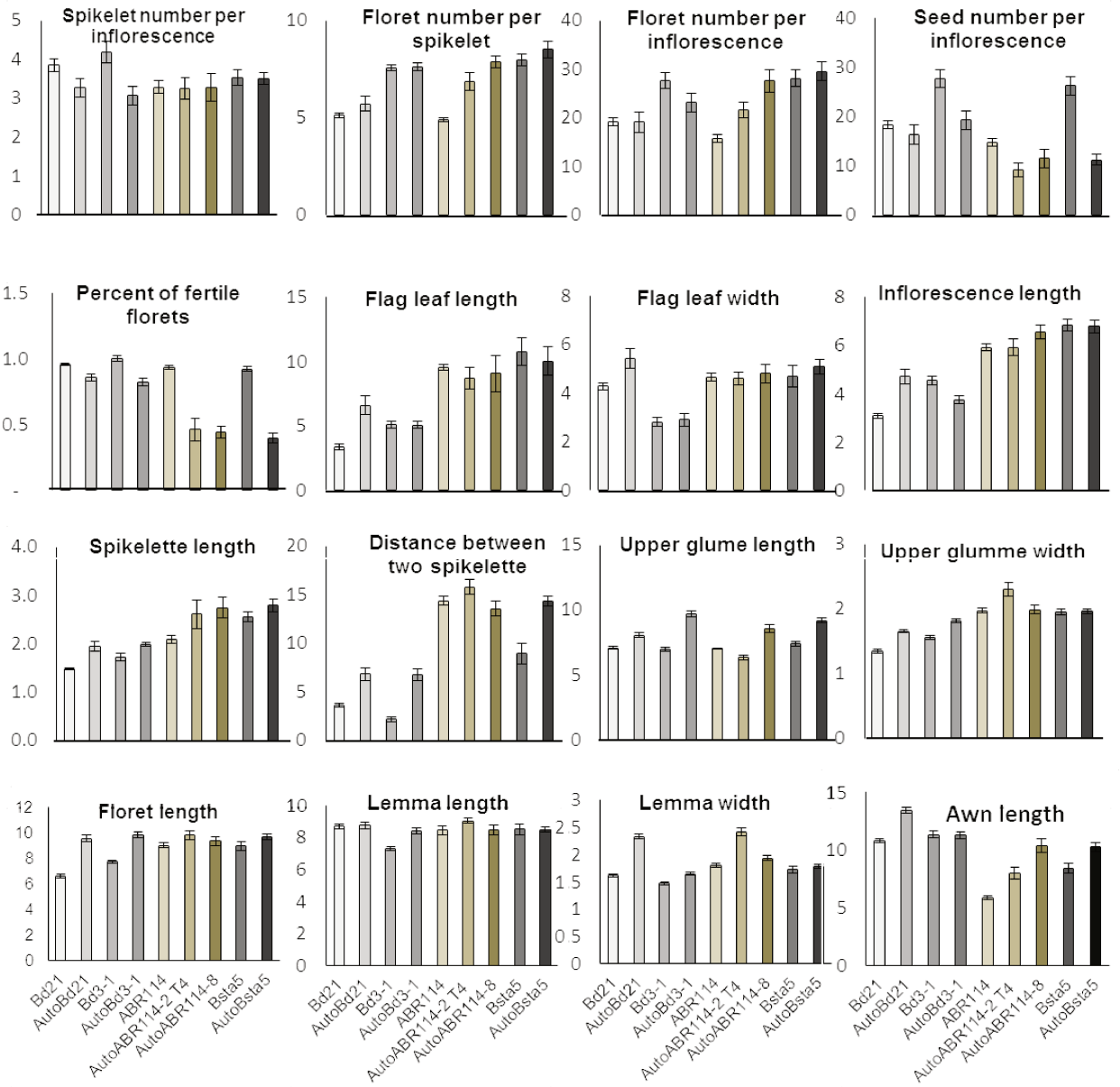


Online Resource 4. Test comparisons for pollens in anthers, pollen abondance and viability, between the diploid *B. stacei* line ABR114 and the derived autopolyploid plant (ABR114_2 (4-9), S1 generation) and the aneuploid plant ((ABR114_2 (3-1) with 42 chromosomes, S1 generation). (A) Entire anthers showing very low proportion of pollens in aneuploids as compared to diploids and autopolyploids. (B) Pollen abundance. Pollens dehisce from one anther per plant were homogenously spread on a glass slide (size of 25 mm × 75 mm). The number of pollens was then counted under microscope (Leica DMLB) in square millimeter ($30mm^2$) units and averaged per one mm^2. (C) Pollen viability test by acetocarmine coloration. Viable pollens are stained by dark purple and unviable ones are stained into light purple.



Online Resource 5. Morphological trait comparison between diploid lines Bd21and Bd3-1 of *B. distachyon* and the diploid lines ABR114 and Bsta5 of *B. stacei* and their derived autopolyploids (names of autopolyploid lines are indicated). Vertical bars indicate standard errors. Five biological replicates were considered.

See Online Resource 3 that shows images of the measured traits and the data in Online Resource 2.



Dear editor,

Please find the manuscript entitled **"Synthesis and evaluation of autotetraploids from *Brachypodium distachyon* and *B. stacei* species".** After discussing and being encouraged by the guest editor Ales Kovarik, we aim this paper for the special issue: "Polyploidy in shallow and deep evolutionary time".

Polyploidy is an important evolutionary feature widespread in eukaryotes.

During the last decades, much more important studies have focused on consequences of **allopolyploidy** (i.e. resulting from hybridization between two different species from the same or related genera) on plant genome evolution and functioning, but few ones have been dedicated to the importance and consequences of **autopolyploidy** (consisting in duplication of genome of the same species).

Species of the genus ***Brachypodium*** became, since one decade, an important monocot and polyploid model because they possess small genomes, have small plant stature, rapid generation time, significant morphometric, with the originality of the high assymetric divergence between species chromosomes and several natural polyploidy species (Hasterok et al., 2004; Betekhtin et al., 2014; Catalan et al., 2015). Among these *B. distachyon* (2n=10) has evolved by descending dysploidy, acting as fusions of smaller chromosomes of a putative ancestral *Brachypodium* species, very close to *B. stacei* (2n=20), leading to two species with similar genome content and ploidy level but contrasting in their chromosome evolution and structure. Moreover, within this framework,naturally hybridize with each others to gave rise to the the allopolyploid species, *B. hybridum.*

To further develop and characterize the *Brachypodium* polyploid model, and characterize **consequences of autopolyploidy**, we have generated in this study autotetraploids from various lines of *Brachypodium distachyon* and *B. stacei* sister species and characterized these at karyotype, phenotype and fertility levels. We showed that both *B. distachyon* and *B. stacei* autotetraploids generally exceeded their diploid progenitors in plant stature, inflorescences and flag leaves characters, but their fertility was reduced as illustrated by the lower number of seeds per inflorescence and lower percent of fertile florets. The generated autotetraploids provide an interesting

material to study the fate of homologous duplicated genes, meiosis and various genomic consequences of autopolyploidy, comparatively between the two *Brachypodium* sister species having similar genome content but contrasting in chromosome evolution and structure.

We believe that our results are appropriate for publication "**in the special issue on Polyploidy in shallow and deep evolutionary time"** of your journal. The results and material developed will be of broad interest for scientists working on *Brachypodium*, polyploidy as well genome evolution and functioning.

We stay at your disposal for any suggestion or discussion that you may need.

Best wishes

Boulos CHALHOUB

The corresponding author

CHAPTER 2:

SYNTHESIS OF BRACHYPODIUM ALLOPOLYPLOIDS

This chapter is presented as a paper accepted in the journal PLOS ONE:

Article 2: “Recreating Stable Brachypodium hybridum Allopolyploids by Uniting Divergent Genomes of B. distachyon and B. stacei that Contrast in Chromosome Evolution, Number and Size”

Vinh Ha Dinh Thi, Olivier Coriton, Isabelle Le Clainche, Dominique Arnaud, Sean P. Gordon, Gabriella Linc, Pilar Catalan, Robert Hasterok, John P. Vogel, Joseph Jahier, Boulos Chalhoub

PLOS ONE

Recreating Stable Brachypodium hybridum Allopolyploids by Uniting Divergent Genomes of B. distachyon and B. stacei that Contrast in Chromosome Evolution, Number and Size

--Manuscript Draft--

Manuscript Number:	
Article Type:	Research Article
Full Title:	Recreating Stable Brachypodium hybridum Allopolyploids by Uniting Divergent Genomes of B. distachyon and B. stacei that Contrast in Chromosome Evolution, Number and Size
Short Title:	Successful synthesis of stable Brachypodium hybridum
Corresponding Author:	Boulos Chalhoub INRA and UEVE FRANCE
Keywords:	allopolyploidy; Brachypodium; asymmetric chromosome evolution; interspecific hybridization; karyotype
Abstract:	Brachypodium hybridum (2n=30) is a natural allopolyploid with parental species, B. distachyon (2n=10) and B. stacei (2n=20), contrasting in chromosome evolution although having the same ploidy level. We recreated B. hybridum allotetraploids by hybridizing various lines from the two parental species. While sterile amphihaploid F1 interspecific hybrids were obtained at low frequencies when B. distachyon was used as the maternal parent (0.15% or 0.245% depending on the line used), no hybrids were obtained from reciprocal crosses or when autotetraploids of the parental species were crossed. Genome doubling through colchicine treatment restored fertility where doubled F1 plants produced a few S1 seed (first selfed generation) after self-pollination. S1 plants from one parental combination (Bd3-1×Bsta5) were fertile and gave rise to further generations whereas those of another parental combination (Bd21×ABR114) were sterile, illustrating the dependence of fertile allopolyploid formation based on parental genotypes. The synthetic allotetraploids were shown to be highly-stable and resembled the natural B. hybridum at the phenotypice, cytogenetic and genomic levels. The successful creation of stable synthetic B. hybridum offers the possibility to investigate allopolyploidy-related changes in genome structure and regulation at the earliest evolutionary stages of the polyploidy existence in comparison withthe parental and natural B. hybridum genomes.
Order of Authors:	Vinh Ha Dinh Thi
	Olivier Coriton
	Isabelle Leclainche
	Dominique Arnaud
	Sean P. Gordon
	Gabriella Linc
	Pilar Catalan
	Robert Hasterok
	John P. Vogel
	Joseph Jahier
	Boulos Chalhoub
Opposed Reviewers:	
Additional Information:	
Question	**Response**

Financial Disclosure Please describe all sources of funding that have supported your work. **This information is required for submission and will be published with your article, should it be accepted.** A complete funding statement should do the following: Include **grant numbers and the URLs** of any funder's website. Use the full name, not acronyms, of funding institutions, and use initials to identify authors who received the funding. **Describe the role** of any sponsors or funders in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. If the funders had **no role** in any of the above, include this sentence at the end of your statement: "*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript*." However, if the study was **unfunded**, please provide a statement that clearly indicates this, for example: "*The author(s) received no specific funding for this work*." * typeset	This project was funded by the French ANR (Agence Nationale de la Recherche, www.agence-nationale-recherche.fr) 2011 (Ploid-ploid-Wheat, ANR-11-BSV5-0015). Vinh Ha Dinh Thi was support by a PhD fellowship from the Vietnamese International Education Development (VIED) and Ministry of Education and Training Vietnam (MOET).
Competing Interests You are responsible for recognizing and disclosing on behalf of all authors any competing interest that could be perceived to bias their work, acknowledging all financial support and any other relevant financial or non-financial competing interests. Do any authors of this manuscript have competing interests (as described in the PLOS Policy on Declaration and Evaluation of Competing Interests)? **If yes**, please provide details about any and all competing interests in the box below. Your response should begin with this statement: *I have read the journal's policy and the authors of this manuscript have the following competing interests:*	The authors have declared that no competing interests exist.

If no authors have any competing interests to declare, please enter this statement in the box: "*The authors have declared that no competing interests exist.*" * typeset	
Ethics Statement You must provide an ethics statement if your study involved human participants, specimens or tissue samples, or vertebrate animals, embryos or tissues. All information entered here should **also be included in the Methods section** of your manuscript. Please write "N/A" if your study does not require an ethics statement. **Human Subject Research (involved human participants and/or tissue)** All research involving human participants must have been approved by the authors' Institutional Review Board (IRB) or an equivalent committee, and all clinical investigation must have been conducted according to the principles expressed in the Declaration of Helsinki. Informed consent, written or oral, should also have been obtained from the participants. If no consent was given, the reason must be explained (e.g. the data were analyzed anonymously) and reported. The form of consent (written/oral), or reason for lack of consent, should be indicated in the Methods section of your manuscript. Please enter the name of the IRB or Ethics Committee that approved this study in the space below. Include the approval number and/or a statement indicating approval of this research. **Animal Research (involved vertebrate animals, embryos or tissues)** All animal work must have been conducted according to relevant national and international guidelines. If your study involved non-human primates, you must provide details regarding animal welfare and steps taken to ameliorate suffering; this is in accordance with the	Not apply, our research are on plant genomes only

recommendations of the Weatherall report, "The use of non-human primates in research." The relevant guidelines followed and the committee that approved the study should be identified in the ethics statement. If anesthesia, euthanasia or any kind of animal sacrifice is part of the study, please include briefly in your statement which substances and/or methods were applied. Please enter the name of your Institutional Animal Care and Use Committee (IACUC) or other relevant ethics board, and indicate whether they approved this research or granted a formal waiver of ethical approval. Also include an approval number if one was obtained. **Field Permit** Please indicate the name of the institution or the relevant body that granted permission.	
Data Availability PLOS journals require authors to make all data underlying the findings described in their manuscript fully available, without restriction and from the time of publication, with only rare exceptions to address legal and ethical concerns (see the PLOS Data Policy and FAQ for further details). When submitting a manuscript, authors must provide a Data Availability Statement that describes where the data underlying their manuscript can be found. Your answers to the following constitute your statement about data availability and will be included with the article in the event of publication. **Please note that simply stating 'data available on request from the author' is not acceptable.** ***If, however, your data are only available upon request from the author(s), you must answer "No" to the first question below, and explain your exceptional situation in the text box provided.*** Do the authors confirm that all data underlying the findings described in their manuscript are fully available without restriction?	Yes - all data are fully available without restriction
Please describe where your data may be	Data and resorces are from the the present study whose authors may be contacted at

found, writing in full sentences. **Your answers should be entered into the box below and will be published in the form you provide them, if your manuscript is accepted.** If you are copying our sample text below, please ensure you replace any instances of XXX with the appropriate details. If your data are all contained within the paper and/or Supporting Information files, please state this in your answer below. For example, “All relevant data are within the paper and its Supporting Information files.” If your data are held or will be held in a public repository, include URLs, accession numbers or DOIs. For example, “All XXX files are available from the XXX database (accession number(s) XXX, XXX)." If this information will only be available after acceptance, please indicate this by ticking the box below. If neither of these applies but you are able to provide details of access elsewhere, with or without limitations, please do so in the box below. For example: “Data are available from the XXX Institutional Data Access / Ethics Committee for researchers who meet the criteria for access to confidential data.” “Data are from the XXX study whose authors may be contacted at XXX.” * typeset	chalhoub@evry.inra.fr
Additional data availability information:	Tick here if your circumstances are not covered by the questions above and you need the journal’s help to make your data available.

Dear editor,

We are pleased to submit our manuscript entitled **“ Recreating Stable *Brachypodium hybridum* Allopolyploids by Uniting Divergent Genomes of *B. distachyon* and *B. stacei* that Contrast in Chromosome Evolution, Number and Size“** for consideration to publication in *Plosone*

Polyploidy is an important evolutionary feature widespread in eukaryotes.
During the last decades, important studies have focused on consequences of **allopolyploidy** which results from hybridization between two different species from the same or related genera. Species of the genus ***Brachypodium*** became, since one decade, an important monocot and polyploid model because they possess small genomes, have small plant stature, rapid generation time, significant morphometric, with the originality of the high assymetric divergence between species chromosomes and several natural polyploid species (Hasterok et al., 2004; Betekhtin et al., 2014; Catalan et al., 2015). Among these *B. distachyon* (2n=10) has evolved by descending dysploidy, acting as fusions of smaller chromosomes of a putative ancestral *Brachypodium* species, very close to *B. stacei* (2n=20), leading to two species with similar genome content and ploidy level but contrasting in their chromosome evolution and structure. Moreover, within this framework, these two species naturally hybridized with each other giving rise to the allopolyploid species, *B. hybridum.*
To further develop the *Brachypodium* polyploid model and characterize **consequences of allopolyploidy**, we have generated for the first time in this study synthetic B*. hybridum* allotetraploids through interespecific hybridization between *B. distachyon* and *B. stacei* species, providing empirical evidence and established a tractable *Brachypodium* allopolyploid model. We have characterized these synthetic allotetraploids at the cytogenetic, phenotypic and fertility levels and show their high genomic stability using a battery of molecular markers in comparison to parental species and natural allotetraploids. As a consequence, an original allopolyploid model with highly-divergent subgenomes assymetric by two-folds in chromosome number and size is now emerging. These successfully generated allotetraploids provide an interesting material to further study the fate of homologous duplicated genes, meiosis and various genomic consequences of allopolyploidy at the earliest evolutionary stages of the polyploidy existence in comparison with later stages in natural *B. hybridum*.
The results and material developed will be of broad interest for scientists working on *Brachypodium*, polyploidy as well genome evolution and functioning. Therefore, we believe this study is suitable for publication in *Plosone*.

We stay at your disposal for any suggestion or discussion that you may need.

Best wishes

Boulos CHALHOUB

The corresponding author

Recreating Stable *Brachypodium hybridum* Allopolyploids by Uniting Divergent Genomes of *B. distachyon* and *B. stacei* that Contrast in Chromosome Evolution, Number and Size

Short title

Successful synthesis of stable *Brachypodium hybridum*

Vinh Ha Dinh Thi[1], Olivier Coriton[2], Isabelle Le Clainche[1], Dominique Arnaud[1], Sean P. Gordon[3], Gabriella Linc[4], Pilar Catalan[5,6], Robert Hasterok[7], John P. Vogel[3,8], Joseph Jahier[2], Boulos Chalhoub[1*]

Affiliations

[1]Organization and evolution of complex genomes, Institut national de la Recherche agronomique, Université d'Evry Val d'Essonne, Evry, France.

[2]Unité Mixte de Recherches INRA, Agrocampus Rennes - Université Rennes 1, Institut de Génétique, Environnement et Protection des Plantes, Le Rheu, France.

[3]DOE Joint Genome Institute, Walnut Creek, USA.

[4]Agricultural Institute, Centre for Agricultural Research, Hungarian Academy of Sciences,Martonvásár, Brunszvik u 2, Hungary.

[5]Department of Agriculture and Environmental Sciences, High Polytechnic School of Huesca, University of Zaragoza, Huesca, Spain.

[6]Department of Botany, Institute of Biology, Tomsk State University, Tomsk , Russia.

[7]Department of Plant Anatomy and Cytology, Faculty of Biology and Environmental Protection, University of Silesia in Katowice, Poland.

[8]Department of Plant and Microbial Biology, University of California, Berkeley, USA.

*Corresponding author: Email: chalhoub@evry.inra.fr (BC)

Abstract

Brachypodium hybridum (2n=30) is a natural allopolyploid with parental species, *B. distachyon* (2n=10) and *B. stacei* (2n=20), contrasting in chromosome evolution although having the same ploidy level. We recreated *B. hybridum* allotetraploids by hybridizing various lines from the two parental species. While sterile amphihaploid F1 interspecific hybrids were obtained at low frequencies when *B. distachyon* was used as the maternal parent (0.15% or 0.245% depending on the line used), no hybrids were obtained from reciprocal crosses or when autotetraploids of the parental species were crossed. Genome doubling through colchicine treatment restored fertility where doubled F1 plants produced a few S1 seed (first selfed generation) after self-pollination. S1 plants from one parental combination (Bd3-1×Bsta5) were fertile and gave rise to further generations whereas those of another parental combination (Bd21×ABR114) were sterile, illustrating the dependence of fertile allopolyploid formation based on parental genotypes. The synthetic allotetraploids were shown to be highly-stable and resembled the natural *B. hybridum* at the phenotypice, cytogenetic and genomic levels. The successful creation of stable synthetic *B. hybridum* offers the possibility to investigate allopolyploidy-related changes in genome structure and regulation at the earliest evolutionary stages of the polyploidy existence in comparison withthe parental and natural *B. hybridum* genomes.

Introduction

Polyploidy, leading to whole genome duplication, is a major evolutionary force in eukaryotes and is particularly prominent and recurrent in angiosperms [1-8]. Allopolyploids combine two or more divergent homoeologous genomes, usually through interspecific or intergeneric hybridization, followed by chromosome doubling. In contrast, autopolyploids combine less divergent genomes from the same species or even the same individual. Over the past few decades, numerous studies have shown that genetic, epigenetic and/or gene expression changes are common consequences of polyploidization across a wide range of species [7, 9-15]. The extent, importance, 'timing' and mechanisms by which these changes occur depend on the allopolyploid model, which in turn depends on the parental lineages crossed and the genome-doubling process (e. g., from stabilized homoploid or from sterile interspecific hybrids) [8, 16-18].

The Poaceae is one of the largest angiosperm families comprising more than 12,000 species, classified into 771 genera [19, 20]. In this family, the tribe Brachypodieae diverged from Aveneae and Triticeae about 38 million years ago (Mya), whilst tribe Ehrhartoideae (rice) diverged approximately 55 (49-66) Mya [21, 22]. The intermediate phylogenetic position of Brachypodieae between tropical cereals like *Sorghum*, *Zea* and *Oryza* of Panicoideae and Ehrhartoideae and temperate grasses such as wheat (*Triticum* and *Aegilops*) and barley (*Hordeum*) [22-25] led to establishing *Brachypodium distachyon* as a model organism for temperate grasses [26-28].

Investigations of about 20 known *Brachypodium* taxa revealed diploid species with variable basic chromosome numbers (x=5, 8, 9, 10) that have hybridized with each other, to form allopolyploid species [24, 29, 30]. This cytological diversity makes

73 *Brachypodium* an ideal model to study dynamic evolution of chromosome number
74 within a genus [20], as well as the consequences of allopolyploidy.

75 The annual *B. distachyon* (2n=10) has the lowest chromosome number and is
76 thought to have played a pivotal role in the evolution of the genus, interspecific
77 hybridizations and formation of polyploid species [24, 30, 31]. Comparative
78 cytogenetic and molecular analyses showed that *B. distachyon's* large chromosomes
79 likely originated via descending dysploidy, acting as fusions of smaller chromosomes
80 of a putative ancestral *Brachypodium* species, very close to *B. stacei* (2n=20) [31-
81 33]. Within this framework, *B. hybridum* (2n=30) was derived by allopolyploidization
82 between *B. distachyon* and *B. stacei* approximately one Mya [33] (Fig. 1). Together,
83 these three species comprise an excellent model to investigate the impact of
84 polyploidization on the organization and evolution of plant genomes, because they
85 possess small genomes, have small plant stature, rapid generation time, and
86 significant morphometric and molecular barcoding differences [33, 34], with the
87 originality of the high asymmetric divergence between parental chromosomes (and
88 by corollary the sub-genomes of *B. hybridum*) [31, 32]. Further supporting this
89 system, the genomes of all three species have been sequenced: the *B. distachyon*
90 genome sequence was first published in 2010 [28] and has recently been improved
91 to an essentially 'finished' genome sequence with the only remaining ambiguity
92 being the location of some centromeric repeats
93 (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Bdistachyon_er) ; the
94 first draft of the *B. stacei* is available
95 (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Bstacei) and a high
96 quality assemble has been created for *B. hybridum* (Vogel unpublished). In addition
97 all three species are small, self-fertile and experimentally tractable [35].

Fig. 1. The *Brachypodium* polyploidy model. (A) Evolution and origin of *B. hybridum* allotetraploid through natural hybridizations between the diploid species *B. distachyon* and *B. stacei.* (B-C) Strategies for synthesis of F1 interspecific hybrids and allopolyploids either by direct crossing of diploid accessions (B) (names indicated) or by doubling their genomes to obtain the respective autopolyploids, followed by interspecific crossing (C).

Several pathways for synthesizing allopolyploids have been proposed [36, 37]. The one-step model suggests that allopolyploids can be formed by a direct hybridization, either by the fusion of unreduced gametes, which are produced at low frequencies in diploid species, or the hybridization between two different autopolyploid species. By contrast, the two-step model requires the formation of amphihaploid interspecific hybrids from reduced gametes of two different species followed by chromosome doubling [38, 39]. Synthesizing allopolyploids by applying the two-step model has been successfully reported for various species from different genera [15, 40-43]. Allopolyploids have also been synthesized by a variant of the one step method where the genomes of the parental species were artificially doubled by colchicine treatment to form autopolyploids. These autopolyploids, each producing 2n gametes, were then hybridized to obtain allopolyploids [40, 44-46].

The aim of this study is to develop a *Brachypodium* polyploid model by synthesizing *Brachypodium* allopolyploids, through hybridization between *B. distachyon* and *B.*

stacei (Fig. 1), and characterizing their stability at the genomic, phenotypic and cytogenetic levels in comparison to parental species and the natural resembling allopolyploid *B. hybridum*.

Materials and methods

Plant material and growth

Six inbred lines of three *Brachypodium* species: *B. distachyon* (Bd21 and Bd3-1), *B. stacei* (ABR114 and Bsta5) and *B. hybridum* (ABR113 and Bhyb30) were used in this study (Fig. 1A and 1B). Seeds were obtained from USDA National Plant Germplasm System (NPGS), Aberystwyth University, UK and University of Zaragoza, Spain [33, 47]. Autopolyploid plants of *B. distachyon* lines Bd21 and Bd3-1 and *B. stacei* lines ABR114 and Bsta5 were generated in our lab (Vinh Ha Dinh Thi and Boulos Chalhoub, unpublished) and used here (Fig. 1C).

Seeds were surface sterilized after removal of the lemmas and paleas using a 10% bleach solution containing a drop of Tween-20 for three minutes. The seeds were then rinsed in sterile water three times. Germination was synchronized by incubating the seeds in Petri dishes at 4℃ for 3 days, and then at room temperature for five days. The seedlings were transferred into pots (10×7 cm) containing a mixture of equal volumes of peat moss and sand supplemented with a slow release fertilizer (2 g/L, Osmocote® Standard 14-14-14, Scotts-Sierra Horticulture, Marysville, OH, USA). Greenhouse conditions were set at day temperature of 22℃ and night temperature of 18℃, with a 16 h photoperiod.

Vegetative propagation

F1 interspecific hybrids, the colchicine-treated F1 (S0), and the S1 allopolyploids were vegetatively propagated to create large numbers of plants from these sterile or nearly sterile lines (S1 Fig.). Root development from secondary tillers was stimulated by covering the base of the plants with soil and adding solution 0.25% of the auxin indole-3-butyric Acid (IBA) [48] to the irrigation solution. Two to three weeks later, secondary tillers with roots were removed, cut and placed in new pots (S1C and S1D Fig.).

Interspecific crossing between *B. distachyon* and *B. stacei*

We tried to generate synthetic allotetraploids by interspecific hybridization between *B. distachyon* and *B. stacei*. Two different lines of each species, together with their derived autotetraploids, that we created previously (Vinh Ha Dinh Thi and Boulos Chalhoub, unpublished), were used and all possible combinations were attempted (Fig.1; Table 1). While hybridizing diploid lines from the two species would allow obtaining amphihaploid (n) F1 hybrid (Fig. 1B), crossing the autotetraploid lines would lead directly to doubled (2n) hybrids, i.e. alloteraploids (Fig. 1C). Reciprocal crossings between diploid *B. distachyon* (lines Bd21 and Bd3-1) and *B. stacei* (lines ABR114 and Bsta5) were performed over four years (2011 - 2014) in the spring and fall seasons, the best seasons for flowering and pollination in our greenhouses. Flowering time was variable among lines to be crossed. Thus, in order to ensure simultaneous flowering, multiple sowings were done for each of the lines at 15 day intervals (from January to March).

Table 1. Interspecific crosses made between plants of two lines of *B. distachyon* and two lines of *B. stacei*.

♀ \ ♂	***B. stacei*** **ABR114**	***B. stacei*** **Bsta5**	***B. distachyon*** **Bd21**	***B. distachyon*** **Bd3-1**
B. stacei **ABR114**	_*	_	2798[a] 25[b] 10[c] 0[d]	627 10 0 0
B. stacei **Bsta5**	_	_	530 6 2 0	846 13 3 0
B. distachyon **Bd21**	2664 39 16 4	565 4 0 0	_	_
B. distachyon **Bd3-1**	541 4 0 0	817 17 6 2	_	_

* No crosses made between lines of the same species

[a] number of made crosses

[b] number of obtained seeds

[c] number of germinated seeds

[d] number of true F1 interspecific hybrid (molecular markers and karyotype)

Emasculation and hybridization method were adapted from Steinwand and Vogel (http://jgi.doe.gov/our-science/science-programs/plant-genomics/brachypodium/). Emasculations were accomplished by removal of the indehiscent anthers from the female parent plants on the two or three basal florets on the spikelet in the morning (10:00 am to midday, Fig. 2A-2C). The emasculated flowers were bagged (NatureflexTM 70x130mm bags) to avoid contamination by non-selected pollen. Pollen from the selected paternal parent was collected from the male parent in the

afternoon of the same day or one day later by placing nearly ripe anthers on a glass slide for 5- 10 min. Most ripe anthers became more turgid and some of them dehisced on the slide (Fig. 2D). Pollen grains were transferred to the emasculated flower (Fig. 2E) and the pollinated spikes were bagged (Fig. 2F) to avoid pollination by stray pollen. Seed formation was recorded 5 to 6 days after pollination. The number of crosses made and the number of putative F1 amphihaploid interspecific hybrid seeds obtained are presented in Table 1.

Fig. 2. Interspecific hybridizations in *Brachypodium*. (A) Floret structure of diploid *B. distachyon* line Bd3-1 at the adequate stage of crossing having two stamens. (B) Floret structure of diploid *B. stacei* line Bsta5 at the optimal stage for crossing showing three stamens. (C) Emasculated florets of Bd3-1 (left) and Bsta5 (right). (D) Ripe anthers were collected on a microscope slide (above), anther dehisced 15-20 minutes later (below) – illustrated here anthers of Bsta5 as an example. (E) Stigma with the pollen grains after pollination - stigma of Bd3-1. (F) All other florets were removed from the spikes that were bagged to avoid unwanted cross pollination. Hybrid seed was observed after 15 days and the mature hybrid seed was collected (shown here without palea and lemma). Bars represent = 1mm.

The putative F1 amphihaploid interspecific hybrid seeds were collected at maturity (at least 4 weeks after pollination). The seeds were kept at 4℃ for three weeks and then at room temperature for 2 months. They were germinated to produce plants as

212 described above. True interspecific hybrid plants were validated by cytological
213 analysis and by PCR markers.
214 All F1 haploid interspecific hybrid plants were vegetatively-propagated as described
215 above (S1 Fig. and S1 Table). The propagated plants were split into two groups.
216 One set of plants was grown in a greenhouse without colchicine treatment to see if
217 any floral meristems with spontaneously doubled chromosomes would arise and
218 produce fertile seeds. The second group of plants was treated with colchicine to
219 double the chromosome number.
220

221 Colchicine treatment of F1 interspecific hybrids

222 We applied a colchicine treatment protocol adapted from the method described by
223 Jahier (49], which was successfully used for wheat [43].
224 Vegetatively-propagated plants at 4-5 leaf stage were completely immersed for three
225 hours in an aqueous solution of colchicine (Sigma-Aldrich Co., cat. no. C9754), at
226 concentrations of 2.5 g/l, 5g/l or 7.5 g/l, and containing 2% DMSO (dimethyl
227 sulfoxide, Sigma-Aldrich Co., cat. no. D8418). The colchicine-treated plants were
228 then transplanted into fresh soil, without rinsing, and grown in a greenhouse. By
229 seven to ten days after treatment surviving plants were recovered and producing
230 new growth. Necrotic lesions observed on the treated leaves suggested that the
231 treatment was effective.
232

233 Flow cytometry (FCM) analysis

234 FCM is a rapid and reliable method that enabled effective determination of DNA
235 content and ploidy level in a number of plant groups [50-52]. FCM was used to

236 determine the ploidy level in F1 hybrids and synthetic allopolyploids. Calibration was
237 done using the profiles and the C-values previously determined for *B. stacei* (0.564
238 pg/2C), *B. distachyon* (0.631pg/2C) and *B. hybridum* (1.265 pg/2C) [33]. Analyses
239 were performed on young leaves obtained from plants grown in the greenhouse as
240 described by Besnard, Garcia-Verdugo (51].

241

242 Chromosome counting and chromosome Fluorescent in-
243 situ hybridization (FISH)

244 Preparation of slides and hybridization using bacterial artificial chromosome -
245 fluorescent *in situ* hybridization (BAC-FISH) was carried out according to procedures
246 detailed in [53-55]. BAC clone ABR1-63-E6 which was labelled by random priming
247 with biotin-14-dUTP (Invitrogen, Life Technologies), allowed to distinguish all
248 chromosomes of *B. distachyon*, and thus effectively discriminates them from the
249 chromosomes of *B. stacei* in the synthetic allopolyploid [56]. The ribosomal DNA
250 probe used in this study was pTa 71 [57] which contained a 9-kb *Eco*RI fragment of
251 rDNA repeat unit (18S-5.8S-26S genes and spacers) isolated from *Triticum*
252 *aestivum*. pTa 71 was labelled with Alexa-488 dUTP by random priming. Biotinylated
253 probe was immunodetected by Texas Red avidin DCS (Vector Laboratories) and the
254 signal was amplified with biotinylated anti-avidin D (Vector Laboratories). The
255 chromosomes were mounted and counterstained in Vectashield (Vector
256 Laboratories) containing 2.5 µg/mL 4',6-diamidino-2-phenylindole (DAPI).
257 Fluorescence images were captured using a CoolSnap HQ camera (Photometrics,
258 Tucson, Ariz) on an Axioplan 2 microscope (Zeiss, Oberkochen, Germany) and
259 analysed using MetaVue™ (Universal Imaging Corporation,Downington, PA).

Estimation of pollen abundance and viability

To estimate pollen viability, anthers were sampled the day of anthesis and were stained by acetocarmine as described by Jahier (49]. Anthers were dilacerated in a drop of acetocarmine, the liberated pollen grains were covered with a coverslip, after removing the debris, and pollen viability was estimated after observation of pollen grains under light microscope. The pollen grain was considered viable if it turned into dark purple, the non-stained ones were classified as the non-viable.

DNA marker development and analysis

Genomic DNA was extracted from young leaves sampled as described previously [58]. Two main types of polymorphic markers were used, simple sequence repeat (SSR) and gene sequence-derived markers.

Simple sequence repeat (SSR)-derived markers

Twenty-two SSR markers were chosen from previous studies (S2 Table). Four of these (ALB165, ALB311, BdSSR330 and R2-3-ABI) were found to discriminate *B. distachyon*, *B. stacei* and *B. hybridum* [59]. The remaining 18 SSR markers were obtained from Garvin, McKenzie (60], Hammami (61] and Vogel, Tuna (62]. Polymorphic markers were also developed based on sequence comparison between *B. distachyon* and *B. stacei* orthologous genes (early release access of the *B.stacei* genome is available through Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html). Pairs of orthologous genes with 6-30

bp insertion/deletion polymorphisms, based on sequence alignments, were selected. Conserved PCR primers flanking the indels were designed using Primer3 (http://biotools.umassmed.edu/bioapps/primer3_www.cgi). A total of 149 primer pairs were designed for 134 orthologous gene regions spanning all five chromosomes of the *B. distachyon* genome (S2 Table).

Characterization of synthetic allopolyploids with the developed genetic markers

All SSR- and gene-derived markers were checked for polymorphism between *B. distachyon* and *B. stacei* lines. Only those that were polymorphic were used for characterizing the hybrids and the allopolyploids. Genomic DNA of the natural allopolyploid *B. hybridum* (lines ABR113, Bhyb30) and the mixture of *B. distachyon* and *B. stacei* parental lines (Bd21 and ABR114; Bd3-1 and Bsta5) were used as controls. For the fertile allopolyploid allo3-1×5, the S1 plant and 118 plants from S2 generation were analyzed.

As in Mestiri *et al.* (2010), a marker was considered rearranged in a synthetic allopolyploid plant if its PCR amplification pattern was different from that observed in the mixture of parental DNA and/or sister allopolyploid plants from the same generation. Polymerase chain reactions (PCR) were performed according to Charles, Belcram (58], in a 10 µl final volume with 200 µM of each deoxynucleoside triphosphate (dNTP), 500 nM of each primer, 0.2U of Taq DNA polymerase (Perkin Elmer, Norwalk, CT, USA) and 25 ng of template DNA. Migration of PCR products were performed in 3% SeaKem LE agarose gels (Lonza) staining with ethidium bromide to detect bands.

Phenotypic analysis

Fifteen morphological characters were measured and compared between synthetic allotetraploids, *B. distachyon*, *B. stacei* and *B. hybridum* (S2 Fig. and S3 Table). Three inflorescence traits that could affect fertility were recorded: number of spikelets per inflorescence (NSI), number of florets per spikelet (NFS) and number of florets per inflorescence (NFI). For synthetic polyploids, which have a high level of sterility, we also recorded percent of fertile florets (PFF) by calculating the number of seeds obtained per total number of florets in the inflorescence, seed number per inflorescence (SI) and 1,000 seed weight (P1000). Another five recorded inflorescence characters were inflorescence length (total length, without awns; IL), spikelet length (total length, without awns, averaging all spikelet lengths per inflorescence; SL), and the distance between two spikelets in the inflorescence (DS; the average of all distances in one inflorescence), upper glume length (UGL), and upper glume width (UGW). Four floral characters were also measured, floret length (FL; from the basal floret), lemma length from basal floret (LL), lemma width (LW) and awn length (longest within spikelet; AL) (S2 Fig.). At least five plants per genotype were analyzed as replicates. Statistical analysis was done using non-parametric Kruskal-wallis test [63].

Results

We assayed two approaches to synthesize allotetraploids from *B. distachyon* and *B. stacei* parental species. The first approach was to cross diploid *B. distachyon* and *B.*

stacei to produce an amphihaploid F1, followed by colchicine treatment to double the chromosomes. The second approach was to first produce *B. distachyon* and *B. stacei* autotetraploid plants and then cross them. Since these autotetraploids should have 2n gametes, the expected F1 progeny would be allotetraploid without need of further chromosome doubling.

Crossing *B. distachyon* and *B. stacei* diploids

Reciprocal interspecific crosses were performed between *B. distachyon* and *B. stacei* (Fig. 2). Two different diploid lines of *B. distachyon* (Bd21, Bd3-1) and two lines of *B. stacei* (ABR114 and Bsta5), making four genotype combinations and eight reciprocal cross types, were used to maximize the chances of success (Table 1). Flowering time was variable among lines from the two species; in order to synchronize simultaneous flowering, multiple sowings were done for each of the lines at 15 day intervals.

A total of 9,388 reciprocal crosses between the two diploid species were performed over a four years period and 122 mature seeds were obtained (Table 1). Among these, 68 were obtained from 4,587 crosses where *B. distachyon* was the maternal parent and 54 from 4,801 crosses where *B. stacei* was the female parent (Table 1). Only 38 of the 122 mature seeds (31%) germinated and produced viable plants. In comparison, the germination rates of *B. distachyon*, *B. stacei* and B*. hybridum* were usually around 96%.

To determine which of the 38 putative F1 plants were true hybrids, we first used codominant SSR markers that differentiate *B. distachyon* and *B. stacei* (see below). We also checked their genome size by flow cytometry as well as chromosome

number. This analysis identified six bona-fide F1 interspecific hybrids, four arising from the 2,664 crosses between *B. distachyon* Bd21 and *B. stacei* ABR114 (designated hereafter as F1_21×114) and two from the 817 crosses between *B. distachyon* Bd3-1 and *B. stacei* Bsta5 (designated hereafter as F1_3-1×5). The final success rate for these crosses was 0.15% and 0.245%, respectively. One F1_3-1×5 hybrid plant died before flowering. Interestingly, we failed to obtain any true F1 interspecific hybrids from the 4,801 crosses where *B. stacei* was used as the maternal parent (all four genotype combinations) as well as from crosses between the two other genotype combinations where *B. distachyon* was the female parent (Table 1).

The five remaining haploid interspecific F1 hybrids were vegetatively-propagated (S1 Fig.) and separated into two batches (S1 Table). The first batch of 99 plants were grown in greenhouse without colchicine treatment to test if spontaneous chromosome doubling would occur and lead to fertile sectors as has been observed in other systems [43, 64]. The second batch of 226 plants was treated with colchicine to induce chromosome doubling, and fertility.

Phenotypically, F1_21×114 and F1_3-1×5 amphihaploid F1 plants resembled the natural *B. hybridum* and were intermediate between *B. distachyon* and *B. stacei* for some characters, like the inflorescence and flag leaves (Fig. 3A and 3B). Floret comparisons showed that the floret hairiness and floret shape of these F1 amphihaploid hybrids were more similar to those of the *B. distachyon* maternal parent than to those of *B. stacei*, whereas the three observed stamens and the stigma structure of the dissected florets were more similar to those of the paternal *B. stacei* parent (Fig. 3C-3E).

Fig. 3. F1 interspecific haploid hybrids compared to natural *B. hybridum* and the diploid parental species, *B. distachyon* and *B. stacei*. From left to right: *B. distachyon* (Bd21), the F1 interspecific haploid hybrid F1_21×114, *B. hybridum* (ABR113) and *B. stacei* (ABR114). (A) Spike morphology; (B) Typical flag leaf; (C) Floret structure; (D) and (E) Floret dissection. Bars: 1mm.

Amphihaploid F1 interspecific hybrids were sterile:

F1 interspecific amphihaploid hybrids are normally sterile, presumably because of defective chromosome pairing at meiosis [65-67]. However, in several cases amphihaploid interspecific hybrids have been reported to produce seeds, most likely by spontaneous genome doubling prior to flowering [43, 64]. We tested this possibility with 99 vegetatively propagated plants from the five different F1 interspecific hybrids. Over a period of 2 years no seeds were produced (S1 Table). Each individual plant produced about 20-30 tillers, with two to three inflorescences per tiller and an average of 33 florets per inflorescence. Thus, about 1,320 to 2,970 florets were checked for each individual plant and a total of approximately 128,040 to 288,090 florets for all F1 plants combined. The failure to obtain seeds indicates that if spontaneous genome doubling can lead to fertile offspring it must be an extremely rare event.

Chromosome doubling of amphihaploid F1 plants and generation of allopolyploids:

One-hundred-fifty-three vegetatively propagated plants from the four original F1_21×114 interspecific hybrids and one of the F1_3-1×5 interspecific hybrids were treated with colchicine to induce chromosome doubling. The majority of plants treated with 2.5 g/l and 5 g/l colchicine solution survived (74% and 87% survival, respectively). By contrast, only 23% of the plants treated with 7.5 g/l colchicine solution survived (S1 Table).

We compared FCM profiles from leaves of colchicine-treated F1 interspecific hybrid plants with those of the non-treated plants and of the parental lines. The results revealed the expected average c-value of ~0.6 pg for the F1 interspecific hybrids, which is similar to the c-values of *B. distachyon* and *B. stacei*. The positions of G1 and G2 peaks in the F1 hybrids were also similar to their counterparts in the diploid parental species (Fig. 4A-4C). The 24 colchicine-treated F1 interspecific hybrid plants (23 plants vegetatively-multiplied from the F1_21×114 initial plant and one from the initial F1_3-1×5 plant) showed G1 and G2 peaks at similar positions to those of the natural allotetraploid *B. hybridum,* indicating that the genomes of these colchicine-treated hybrids have been partially or completely doubled (Fig. 4D and 4E; Table 2).

Fig. 4. Flow cytometry (FCM) profiles. (A) diploid *B. distachyon* Bd21 and (B) diploid *B. stacei* ABR114 showing two peaks corresponding to the G1 (2C DNA) and G2/M (4C DNA) nuclei. (C) The F1 interspecific hybrid of *B. distachyon* Bd21 and *B. stacei* ABR114, with a profile similar to that of the parental species. (D) The shift on

428 the x axis of the G1 peak to 4C position and G2 peak to 8C positions was found in *B.*
429 *hybridum* ABR113 indicating a doubling of the nuclear content of the genome. (E)
430 FCM profile of the synthetic allopolyploid allo21×114 (S0 generation) is similar to *B.*
431 *hybridum* (D) and double the genome size of the F1 (C) and diploid parental species
432 (A-B). The X and Y axes show relative DNA content per nuclei estimated by
433 fluorescence intensity and the frequency-count (cell number), respectively.

434

435 Metaphase chromosome counting of root-tip cells revealed that six plants derived
436 from F1_21×114 and one plant derived from F1_3.1×5 had a chromosome number
437 of 30 in all cell examined which is consistent with whole genome duplication. Four
438 plants derived from F1_21×114 had variable chromosome number in different cells
439 indicating that these plants were a mosaic of cells with doubled and non-doubled
440 genomes (Table 2).

441 **Table 2. Summary of cytogenetic analyses conducted on different *Brachypodium distachyon* (lines Bd21 and Bd3-1) and**

442 ***B. stacei* (lines ABR114 and Bsta5) species, their derived autopolyploids (4× followed by same line name), F1 interspecific**

443 **hybrids, synthetic polyploids (generations S0 to S2) and the natural *B. hybridum* (lines ABR113 and Bhy30).**

Species and genotypes	Genome	Flow cytometry		Mitosis metaphase stage			
		Plant number	Ploidy level	Plant number	Chromosome number (DAPI staining)	Number of 45S rDNA sites	Number of chromosomes hybridizing with the BAC ABR1-63E06
B. *distachyon*	Bd21	2	2x	1	10	2	-*
	Bd3-1	2	2x	1	10	2	-
	4×Bd21	3	4x	2	20	4	-
	4×Bd3-1	2	4x	2	20	4	-
B. *stacei*	ABR114	2	2x	1	20	2	-
	Bsta5	2	2x	1	20	2	-
	4×ABR114	15	4x	15	40	4	-
	4×Bsta5	1	4x	1	40	4	-
F1 interspecific hybrids	F1_21×114	40[a]	2x	8	15	2	5
	F1_3-1×5	1[a]	2x	1	15	2	5
Synthetic	allo21×114	23[a]	4x	10	30 (6)[b]	4(6)[b]	10 (6)[b]

allopolyploids(S0 generation)					30 and 15 (4)[c]	4 and 2(4)[c]	10 and 5(4)[c]
	allo3-1×5	1[a]	4x	1	30	4	10
Synthetic allopolyploids (S1 generation)	allo21×114	10[a]	4x	5	30	4	10
	allo3-1×5	10[a]	4x	2	30	4	10
Synthetic allopolyploids (S2 generation)	allo3-1×5	10	4x	10	30	4	10
B. hybridum	ABR113	2	4x	1	30	4	10
	Bhyb30	2	4x	1	30	4	10

444 * Not analyzed

445 [a] These plants were obtained by vegetative cuttings from one initial plant of each category

446 [b] Number of plants showing doubled genome karyotype

447 [c] Number of plants showing mixed karyotype of doubled and non-doubled genome

448

449 The seven colchicine-treated F1 plants with uniform 30 chromosomes were
450 considered to be zero-selfed (S0) generation of the allopolyploid allo21×114, derived
451 from F1_21×114, and of the allopolyploid allo3-1×5, derived from F1_3-1×5. These
452 S0 plants were maintained and also multiplied vegetatively. Only two S1 (selfed
453 generation subsequent to S0) seeds were obtained from more than 200,000 flowers
454 of 153 S0 allo21×114 plants, whereas one S1 seed was obtained from the single S0
455 plant of allo3-1×5. This indicated an overall low fertility in the first generation of the
456 synthetic allotetraploid. Only one out of the two S1 seeds of allo21×114 and the
457 single S1 seed of allo3-1×5 germinated and gave S1 mature plants. The S1 plants
458 were also vegetatively propagated to produce 161 allo21×114 and 48 allo3-1×5
459 plants.
460 All S1 plants were taller and more vigorous than the F1 hybrids and the parental
461 species. The two synthetic allotetraploids exhibited similar morphology at the first
462 stages of leaf development for the main stem and tillering (as defined by [68] (Fig.
463 5A). However, they showed differences for main stem elongation stage, where
464 allo21×114 stem tended to have longer and more internodes, leading to taller plants,
465 as compared to those of allo3-1×5 (Fig. 5B-5D). They were more similar, however, in
466 their inflorescence structure, exhibiting long inflorescences with three to five
467 spikelets (Fig. 5E). At anthesis, flowers of the two types of synthetic allopolyploids
468 were similar, both containing long hairy lemmas, three stamens and feathery stigmas
469 (Fig. 5F-5G). A summary of floret characters of F1 hybrids and S1, and the
470 generation of allopolyploids, compared with those of their progenitors and natural *B.*
471 *hybridum*, is given in S4 Table. Genome size was also assessed by FCM for the S1
472 plants of the synthetic allopolyploid plants; the positions of their G1 and G2 peaks

(Fig. 4D) were similar to those of their counterparts observed in natural *B. hybridum* (Fig. 4C).

Fig. 5. Comparison of plant morphology between the two synthesized allopolyploids at the S1 generation: allo21×114 (right) and allo3-1×5 (left). (A) Vegetatively multiplied plants; (B) Plants at the tillering stage; (C) Flowering stage; (D) Tillers; (E) Spike structure; Flower dissections of (F) allo3-1×5 and (G) allo21×114.

S1 plants of allo3-1×5 were fertile with 23% of flowers producing seed. While this is much greater than the S0 plants, it is lower than natural *B. hybridum* lines ABR113 (91%) and Bhyb30 (68%) (S3 Table). More than 100 S2 seeds of allo3-1×5 were sown and almost all of them germinated and gave plants for further cytogenetic and genomic characterization. Surprisingly, all 135 vegetatively propagated S1 allo21×114 were sterile.

Sterility of the synthetic allopolyploid allo21×114:

In order to characterize the sterility of S1 allo21×114, we used acetocarmine staining to determine pollen shape [49]. Anthers of S1 plants of allo21×114 contained few, ~15-17, normal shape viable pollen grains, whereas anthers from natural *B. hybidum* typically contain ~170-200 viable pollen grains (S3 Fig.). The near absence of normal pollen suggests that S1 allo21×114 plants have significantly reduced male fertility.

To examine the female fertility of S1 allo21×114 plants we pollinated emasculated S1 allo21×114 flowers (as well as F1_ 21×114) with pollens from the two diploid parents and two natural *B. hybridum* lines but no seeds were obtained (S5 Table). Comparatively, 10 seeds were obtained from 35 crosses done between the two *B. hy*bridum lines (data not shown).

Our results suggest that S1 plants of allo21×114 may be both male and female sterile.

Crossing autopolyploid plants of *B. distachyon* and *B. stacei*

We performed 4,384 reciprocal crosses between autotetraploids from two lines of *B. stacei* (ABR114 and Bsta5) and two lines of *B. distachyon* (Bd21 and Bd3-1), consisting in four genotype combinations and obtained 48 seeds (Table 3). However, only 11 germinated and none were true interspecific hybrid as checked by codominant molecular markers (Table 3).

Table 3. Interspecific crosses made between autopolyploids (4x) plants of two lines of *B. distachyon* and two lines of *B. stacei*.

♀ \ ♂	***B. stacei*** 4×**ABR114**	***B. stacei*** 4×Bsta5	***B. distachyon*** 4×**Bd21**	***B. distachyon*** 4×**Bd3-1**
B. stacei 4×ABR114	_*	_	748[a] 9[b] 2[c] 0[d]	550 4 1 0

B. stacei 4×Bsta5	–	–	469 6 2 0	484 5 1 0
B. distachyon 4×Bd21	608 12 4 0	540 2 0 0	–	–
B. distachyon Bd3-1	509 7 2 0	476 3 0 0	–	–

[*] No crosses made between lines of the same species

[1] number of made crosses

[b] number of obtained seeds

[c] number of germinated seeds

[d] number of true F1 interspecific hybrid (molecular markers and karyotype)

Phenotypic characterization of synthetic allopolyploids

Fifteen morphological characters related to inflorescence and floral architecture and sizes were measured and compared between synthetic S1 generation allotetraploids, their parental species *B. distachyon* and *B. stacei*, the average values of parents (mid-parent values: MPV), and *B. hybridum* natural allotetraploids (Fig. 6;S3 Table). In general, the synthetic allotetraploids were more similar to natural *B. hybridum,* exceeding generally the parental species or the average of parents. More comparisons for each of the individual traits are detailed in S1 Text.

Fig. 6. Comparison of spike and flower characters measured in synthetic allopolyploids (S1 generation), their parental species *B. distachyon* and *B. stacei*, and in the *B. hybridum* natural allopolyploid. MPV1 = Mid-parent value: average of *B. distachyon* Bd21 and *B.* stacei ABR114 parent values; MPV2: average of *B. distachyon* Bd3-1 and *B. stacei* Bsta5 parent values. See S2 Fig. for specifics of the characters scored and S3 Table for further details. There were no seeds from S1 plants of allo21×114, therefore seed number per inflorescence and percent of fertile florets were not scored.

Karyotype characterization

Metaphase chromosomal analysis was conducted in F1 interspecific amphihaploid hybrids and in S1 and S2 generations plants of the synthetic allotetraploid (Fig. 7; Table 2). The comparative analysis showed the expected 10 large chromosomes in *B. distachyon* (Fig. 7A), 20 small chromosomes in *B. stacei* (Fig. 7B) and 30 (large and small) chromosomes in the natural *B. hybridum* allopolyploid (Fig. 7C). Amphihaploid F1 interspecific hybrids contained 15 chromosomes, five derived from *B. distachyon* and 10 from *B. stacei,* (Fig. 7D). As expected, chromosomes were duplicated in the two derived S1, S2 plants of the synthetic allopolyploid allo3-1×5 (Fig. 7E and 7F) that had similar karyotypes to those of the *B. hybridum* natural allopolyploid. FISH with the 45S rDNA probe showed the expected number of two signals in *B. distachyon*, *B. stacei* and their derived F1 amphihaploid interspecific hybrids (Fig. 7A, 7B, and 7D) and four signals in their derived synthetic and S1 and S2 allopolyploid plants (Fig. 7E and 7F), similar to the natural *B. hybridum*

allopolyploid (Fig 7C). Genome-specific chromosome discrimination with the BAC ABR1-63-E6 probe demonstrated the presence of five chromosomes of *B. distachy*on in the amphihaploid F1 interspecific hybrids (Fig. 7D) and a doubled number (10) in their derived S1 and S2 synthetic allotetraploids (Fig. 7E, and 7F) and in the natural *B. hybridum* (Fig. 7C). We analyzed 53 chromosome complements from 10 plants of the S2 allo3-1×5. The results indicated that all of them were euploids, showing the expected number of 30 chromosomes, which suggested high chromosome stability (Table 2). The observed phenotypes of these plants also gave evidence that they were homogeneous and similar to each other and to their S1 parents.

Fig. 7. Cytogenetic analysis on metaphase chromosomes of natural *Brachypodium* species, F1 interspecific hybrid, and synthetic allopolyploids (Blue: DAPI staining, green: FISH with 45S rDNA probe; red: genome-specific discrimination of chromosomes with BAC clone ABR1-63-E6). (A-B) DAPI staining revealing five pairs of chromosomes in *B. distachyon* Bd21 (2n=10) and 10 smaller chromosome pairs in *B. stacei* ABR114 (2n=20) whereas FISH with 45S rDNA probe (green) reveals two sites in *B. distachyon* and two in *B. stacei*. (C) DAPI staining revealing 15 pairs of chromosomes in the *B. hybridum* ABR113 (2n=30) allopolyploid: five large chromosome pairs derived from the *B. distachyon* and 10 other smaller pairs derived from the *B. stacei* parent. FISH with 45S rDNA probe (green) reveals four sites (two on each parent-derived chromosomes) (C1) and genome-specific discrimination of chromosomes with BAC clone ABR1-63-E6 probe

582 (red) reveals specifically five large *B. distachyon*-derived chromosomes pairs and 10
583 other smaller pairs (blue) derived from the *B. stacei* parent (C2). (D) F1 Interspecific
584 amphihaploid hybrid (F1_3-1×5). DAPI staining (blue) revealing 5 *B. distachyon*-
585 derived chromosomes and 10 smaller *B. stacei*-derived chromosomes (D1). FISH
586 with 45S rDNA probe (green) reveals one site on a *B. distachyon*-derived
587 chromosome and one other on a *B. stacei*-derived smaller chromosome (D1).
588 Genome-specific discrimination of chromosomes with BAC clone ABR1-63-E6 probe
589 reveals specifically five *B. distachyon*-derived chromosomes and 10 smaller (blue) *B.*
590 *stacei*-derived chromosomes (D2). (E-F) S1 and S2 plants of the synthetic
591 allopolyploid allo3-1×5 with DAPI staining and FISH (blue) with 45S rDNA probe
592 (green) (E1, F1) and GISH-like with BAC clone ABR1-63-E6 (red, E2, F2), showing
593 similar profiles to those of natural *B. hybridum*. Bars: 5 µm.

594

595

596 Genetic characterization of synthetic allopolyploids

597 SSR- and gene-derived PCR markers were used to characterize the genetic stability
598 of synthetic allotetraploids. The single allo3-1×5 S1 plant was fertile, allowing us to
599 survey 118 individual S2 plants, whereas only a single sterile allo21×114 S1 plant
600 was analyzed. The genetic markers were classified based on the patterns of the
601 amplification and polymorphism observed between parental species, their pooled
602 DNAs, and the natural allopolyploid *B. hybridum* (Fig. 8) and then compared to those
603 observed in the synthetic allopolyploids. The classification of the marker types was
604 as recommended Mestiri, Chagué (43] and is illustrated in Fig. 8 and described in
605 more details in S2 Text.

606

607

608 **Fig. 8. PCR-based marker profiles of plants from generations F1, S1 and S2 of**
609 **synthetic *Brachypodium* allopolyploids compared with their progenitors**
610 **species and mixtures of equal amounts of parental DNAs (MPV1 and MPV2)**
611 **and natural allopolyploid *B. hybridum* (ABR113 and Bhyb30).** (A) Bd2-28 marker
612 showing no polymorphism between *B. distachyon* (Bd21 and Bd3-1) and *B. stacei*
613 (ABR114 and Bsta5) progenitor lines and no change in F1, S1 and S2 synthetic
614 allopolyploid. (B) Bd3-11 marker showing a presence⁄absence polymorphism (PAP)
615 with absence of the *B. distachyon* allele (-), all F1, S1 and S2 plants amplify the
616 same allele as *B. stacei*. (C) Bd5-14 marker showing allele size polymorphism
617 (ASP+) between *B. distachyon* and *B. stacei* diploid progenitors lines, with both
618 parental alleles amplified in MPVs, F1, S1, S2, and *B. hybridum*. (D) Bd1-26 marker
619 showing allele size polymorphism (ASP-) between *B. distachyon* (Bd21 and Bd3-1)
620 and *B. stacei* (ABR114 and Bsta5) diploid progenitors lines, but allele of *B.*
621 *distachyon* is not amplified in MPVs (red arrows) and in all plants of F1, S1 and S2
622 generations as well as in the natural *B. hybridum*. This deviation from additivity is
623 considered as most likely resulting from competition for PCR amplification between
624 progenitor alleles and not from DNA rearrangements. L: 50 bp ladder (Invitrogen,
625 Carlsbad, CA, USA). MPV1; Mixture of DNA from Bd21 and ABR114. MPV2: Mixture
626 of DNA from Bd3-1 and Bsta5.

627

628

629 A total of 151 markers (129 gene-based and 22 SSR markers) and 140 markers (123
630 gene-based and 17 SSR markers) were analyzed for allo21×114 and allo3-1×5,

631 respectively. The genetic markers obtained for F1 interspecific amphihaploid hybrids,
632 F1_21×114 and F1_3-1×5, were the same as those for their derived S0 and S1
633 synthetic allopolyploids (Table 4).
634 None of the validated markers was found rearranged in the F1 interspecific
635 amphihaploid hybrids or derived synthetic S0 and S1 of both allopolyploids that show
636 amplification patterns similar to those of mixture of parental DNA (MPV) and/or
637 parents (Fig. 8; Table 4). Similarly, none of the 118 different S2 plants available for
638 allo3-1×5 showed rearrangements of parental alleles (Fig. 8; Table 4), indicating a
639 high stability.
640 The genetic patterns in the synthetic allopolyploids were almost identical to the
641 genetic profiles observed in the natural *B. hybridum* ABR113 and Bhyb30 lines (S2
642 Table). However, we observed slightly more differences with the line Bhyb30, than
643 with line ABR113.
644
645

Table 4. Classification of gene- and simple sequence repeats (SSRs)-based PCR markers according to their amplification patterns observed on *B. distachyon* and *B. stacei* diploid species, their pooled DNAs (MPVs), their F1 interspecific hybrids and S0, S1 and S2 generations of the derived synthetic allopolyploids.

Marker classification	Genome	Chr	M	PAP	ASP [a]	Total
Gene based markers	Allo21×114	Bd1	2	14	11 (4)	**27**
		Bd2	4	12	8 (1)	**24**
		Bd3	0	14	10 (1)	**24**
		Bd4	12	7	10 (2)	**29**
		Bd5	3	14	8 (2)	**25**
		Total	**21**	**61**	**47**	**129**
	Allo3-1×5	Bd1	4	11	11(4)	**26**
		Bd2	5	10	8 (1)	**23**
		Bd3	0	9	11 (1)	**20**
		Bd4	14	5	10 (2)	**29**
		Bd5	4	13	8 (2)	**25**
		Total	**27**	**48**	**48**	**123**
SSR markers	Allo21×114	Bd1	0	3	1	**4**
		Bd2	0	6	1	**7**
		Bd3	0	1	1(1)	**3**
		Bd4	1	2	1(1)	**5**
		Bd5	0	1	1(1)	**3**
		Total	**1**	**13**	**8**	**22**
	Allo3-1×5	Bd1	1	2	1	**3**
		Bd2	0	4	1	**5**
		Bd3	0	1	2	**3**
		Bd4	1	2	1	**4**
		Bd5	0	0	1(1)	**2**
		Total	**1**	**9**	**7**	**17**

Chr: Chromosome number according to *B. distachyon* (Bd)

M: monomorphic

PAP: presence/absence polymorphic markers, amplifying one single parental allele

655 ASP: allele specific markers, amplifying both parental alleles with polymorphic allelic
656 size differences. Numbers between brackets correspond to those that amplify one
657 single parental allele in pooled DNAs from both parents.
658
659

660 Discussion

661 The origin and evolutionary relationships of the natural *B. hybridum* allotetraploid in
662 relation to its progenitor species *B. distachyon* and *B. stacei* is now clearly elucidated
663 [26, 32-34, 69]. The recreation of synthetic allotetraploid similar to natural
664 *B.hybridum* provides empirical evidence and establishes the tractable *Brachypodium*
665 allopolyploid model. This represents a unique allopolyploid model where one
666 parental genome (*B. distachyon*) has similar genome size to the other one (*B.*
667 *stacei*), but half the basic chromosome number (2n=10 and 2n=20, respectively)
668 whereas its individual chromosome size is approximately two times larger.
669 The existence of *B. hybridum* and other natural *Brachypodium* allopolyploids [24] as
670 well as our success in synthesizing similar allopolyploids, suggest that differences in
671 chromosome size and number between progenitor species do not constitute a barrier
672 to interspecific hybridization and allopolyploid formation. *B. distachyon* and its
673 derived allotetraploid *B. hybridum* hybridized with various other diploid species [70],
674 presumably leading to several additional *Brachypodium* allopolyploids including *B.*
675 *pinnatum* (2n=28), *B. phoenicoides* (2n=28), *B. phoenicoides* (2n=28) and *B.*
676 *retusum* (2n=38) [30, 31, 33, 71].
677 Moreover, the *B. hybridum*-type allotetraploids synthesized here appear highly stable
678 from the earliest generations (S1 and S2) as characterized at the phenotypic,
679 cytogenetic and genetic levels. The prominent differences in both chromosome

number and chromosome size of the two parental genomes could probably represent a factor in the chromosomal stability of both natural and synthetic *B. hybridum* as they may constitute a barrier for homoeologous pairing and rearrangements. We plan future studies of meiosis and chromosome pairing in these model polyploids.

B. hybridum allotetraploids were likely formed naturally more than once, with both *B. distachyon* or *B. stacei* as the maternal parent [34]. In our experiments surviving synthetic *Brachypodium* allotetraploids have *B. distachyon* as the maternal progenitor whereas no allopolyploids were obtained from reciprocal crosses. Moreover, the success rate of F1 amphihaploid interspecific hybrids and derived allopolyploid observed in the present study was low. Similarly, different success rates, have been previously observed, depending on the studied species and combinations of parental genotypes used [72-74].

It has been suggested that a combination of factors, including differences in flowering time, pollinator behavior and floral structure, caused by both biological and genetic factors, can cause unsuccessful or limited formation of zygotes between distantly-related species [75]. Even when pollination occurs, post-pollination barriers, such as differences in style structure and the arrest of the pollen tube growth, can inhibit the formation of zygotes between different species. These can be overcome by refining crossing methods [76]. As an example, in lilies (*Lilium candidum* L.) the pollen tubes arrest halfway down the style after interspecific pollination, a barrier that can be overcome by *in vitro* methods [77]. Post-pollination barriers have also been reported in other species, such as *Rhinanthus* and *Nicotiana* [78, 79], as the pollen tube progress at different rate in hetero-specific style or because of differences in pistil length between the crossed species. In the present study, we overcame most

barriers preventing the hybridization and zygote formation between some line combinations of *B. distachyon* and *B. stacei*, as we have obtained viable F1 interspecific hybrids. Previously, other groups have failed to obtain F1 interspecific amphihaploid hybrids between other *Brachypodium* species [70]. In comparison, by performing a high number of interspecific crosses, we obtained viable F1 interspecific hybrids in only two out of four genotype combinations. It appears that the successful combination of lines Bsta5 of *B. stacei* and Bd3-1 of *B. distachyon* that is compatible and amenable to obtain fertile allopolyploids was not tried in previous studies (G. Linc and R. Hasterok, unpublished). This illustrates the importance of the particular genotypes of the parental species. On the other hand, hybridizations between the autotetraploid lines of *B. distachyon* and *B. stacei,* did not give any successful allopolyploid. Among various possible reasons, it is important to note the low fertility in *B. stacei* and *B. distachyon* autotetraploid lines (46% and 82%, respectively), which is accompanied by reduced pollen viabilities (*data not shown*). Further characterizations of the germination of the pollen on the stigma papilla as well as the progression of the pollen tubes in the style would allow better elucidation of the potential barriers limiting zygote and interspecific hybrid formation between *B. distachyon* and *B. stacei*.

Interspecific F1 amphihaploid hybrids are normally sterile because the parental chromosomes do not pair normally during meiosis leading to unbalanced non-viable gametes [80, 81]. Doubling the genome of F1 amphihaploid plants often restores fertility and occasionally this occurs spontaneously as has been observed in a variety of plant species such as in wheat [43], *Arabidopsis* [64, 82-84], and rice [85]. In our study, there were no seeds obtained from thousands of amphihaploid F1 interspecific hybrid flowers, indicating that restoration by spontaneous genome

doubling does not occur or is exceedingly rare for the crosses we made. We were able to artificially double the genome of our two amphihaploid F1 hybrid plants leading to low fertility. Interestingly, for one cross, the next selfed generation (S1) was even more fertile and it will be very interesting to explore the changes responsible for increasing fertility. Conversely, fertility did not increased in the S1 generation of the other allopolyploid and this contrast may provide mechanistic insight. Whilst reasons of sterility of allo21×114 in comparison to the fertile allo3-1×5 allopolyploid need to be investigated, these findings suggests the existence of pre-established genetic or structural factor that influence hybridization success and stability of allopolyploid genomes, as observed for hexaploid wheat [43].

In conclusion, the successful synthesis of allotetraploids similar to the natural *B. hybridum* provides a powerful new tool to an emerging polyploid model system. When combined with the experimental resources and experimental tractability of *B. distachyon, B. stacei* and *B. hybridum*, the ability to create allotetraploids opens up exciting possibilities to study various aspects of polyploidy in grasses at genomic, cytomolecular, epigenetic and physiological levels from the very earliest stages of their evolution.

Acknowledgements

We thank Professor Andrew Leitch (Queen Mary University of London) for valuable advices and correction of the manuscript.

.

References

1. Barker MS, Arrigo N, Baniaga AE, Li Z, Levin DA. On the relative abundance of autopolyploids and allopolyploids. New Phytol. 2016;210(2):391-8. Epub 2015/10/07. doi: 10.1111/nph.13698. PubMed PMID: 26439879.
2. Blanc G, Wolfe KH. Widespread Paleopolyploidy in Model Plant Species Inferred from Age Distributions of Duplicate Genes. The Plant Cell. 2004;16(7):1667-78. doi: 10.1105/tpc.021345. PubMed PMID: PMC514152.
3. Adams KL, Wendel JF. Polyploidy and genome evolution in plants. Curr Opin Plant Biol. 2005;8(2):135-41. Epub 2005/03/09. doi: 10.1016/j.pbi.2005.01.001. PubMed PMID: 15752992.
4. Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH. Synteny and collinearity in plant genomes. Science. 2008;320(5875):486-8. Epub 2008/04/26. doi: 10.1126/science.1153917. PubMed PMID: 18436778.
5. Soltis PS, Soltis DE. The role of hybridization in plant speciation. Annu Rev Plant Biol. 2009;60:561-88. Epub 2009/07/07. doi: 10.1146/annurev.arplant.043008.092039. PubMed PMID: 19575590.
6. Van de Peer Y, Maere S, Meyer A. The evolutionary significance of ancient genome duplications. Nat Rev Genet. 2009;10(10):725-32.
7. Wendel JF, Jackson SA, Meyers BC, Wing RA. Evolution of plant genome architecture. Genome biology. 2016;17(1):37. Epub 2016/03/02. doi: 10.1186/s13059-016-0908-1. PubMed PMID: 26926526; PubMed Central PMCID: PMCPMC4772531.
8. Beest M, Le Roux JJ, Richardson DM, Brysting AK, Suda J, Kubešová M, et al. The more the better? The role of polyploidy in facilitating plant invasions. Annals of Botany. 2011;109(1):19-45. doi: 10.1093/aob/mcr277.
9. Buggs RJ, Renny-Byfield S, Chester M, Jordon-Thaden IE, Viccini LF, Chamala S, et al. Next-generation sequencing and genome evolution in allopolyploids. Am J Bot. 2012;99(2):372-82. Epub 2012/01/24. doi: 10.3732/ajb.1100395. PubMed PMID: 22268220.
10. Chalhoub B, Denoeud F, Liu S, et al. Plant genetics. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science. 2014;345(6199):950-3. doi: 10.1126/science.1253435. PubMed PMID: 25146293.
11. Ghani MA, Li J, Rao L, Raza MA, Cao L, Yu N, et al. The role of small RNAs in wide hybridisation and allopolyploidisation between *Brassica rapa* and *Brassica nigra*. BMC Plant Biol. 2014;14:272. Epub 2014/10/20. doi: 10.1186/s12870-014-0272-9. PubMed PMID: 25326708; PubMed Central PMCID: PMCPMC4209033.
12. Kashkush K, Feldman M, Levy AA. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nature Genetics. 2003;33(1):102-6.
13. Rapp RA, Udall JA, Wendel JF. Genomic expression dominance in allopolyploids. BMC biology. 2009;7:18. Epub 2009/05/05. doi: 10.1186/1741-7007-7-18. PubMed PMID: 19409075; PubMed Central PMCID: PMCPMC2684529.
14. Rieseberg LH, Sinervo B, Linder CR, Ungerer MC, Arias DM. Role of gene interactions in hybrid speciation: evidence from ancient and experimental hybrids. Science. 1996;272(5262):741-5. Epub 1996/05/03. PubMed PMID: 8662570.
15. Song K, Lu P, Tang K, Osborn TC. Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proceedings of the National Academy of Sciences of the United States of America. 1995;92(17):7719-23. Epub 1995/08/15. PubMed PMID: 7644483; PubMed Central PMCID: PMCPMC41217.
16. Stebbins GL. Polyploidy, Hybridization, and the Invasion of New Habitats. Annals of the Missouri Botanical Garden. 1985;72(4):824-32. doi: 10.2307/2399224.
17. Paun O, Bateman RM, Fay MF, Luna JA, Moat J, Hedren M, et al. Altered gene expression and ecological divergence in sibling allopolyploids of *Dactylorhiza* (*Orchidaceae*). BMC evolutionary

biology. 2011;11:113. Epub 2011/04/28. doi: 10.1186/1471-2148-11-113. PubMed PMID: 21521507; PubMed Central PMCID: PMCPMC3112086.
18. Estep MC, McKain MR, Vela Diaz D, Zhong J, Hodge JG, Hodkinson TR, et al. Allopolyploidy, diversification, and the Miocene grassland expansion. Proceedings of the National Academy of Sciences. 2014;111(42):15149-54. doi: 10.1073/pnas.1404177111.
19. Soreng RJ, Peterson PM, Romaschenko K, Davidse G, Zuloaga FO, Judziewicz EJ, et al. A worldwide phylogenetic classification of the Poaceae (Gramineae). Journal of Systematics and Evolution. 2015;53(2):117-37. doi: 10.1111/jse.12150.
20. Kellogg EA. Evolutionary history of the grasses. Plant Physiol. 2001;125(3):1198-205. PubMed PMID: 11244101; PubMed Central PMCID: PMCPMC1539375.
21. Bouchenak-Khelladi Y, Slingsby JA, Verboom GA, Bond WJ. Diversification of C4 grasses (Poaceae) does not coincide with their ecological dominance. American Journal of Botany. 2014;101(2):300-7. doi: 10.3732/ajb.1300439.
22. Chalupska D, Lee HY, Faris JD, Evrard A, Chalhoub B, Haselkorn R, et al. Acc homoeoloci and the evolution of wheat genomes. Proceedings of the National Academy of Sciences. 2008;105(28):9691-6. doi: 10.1073/pnas.0803981105.
23. Draper J, Mur LA, Jenkins G, Ghosh-Biswas GC, Bablak P, Hasterok R, et al. *Brachypodium distachyon*. A new model system for functional genomics in grasses. Plant Physiol. 2001;127(4):1539-55. PubMed PMID: 11743099; PubMed Central PMCID: PMC133562.
24. Catalan P, López-Álvarez D, Díaz-Pérez A, Sancho R, López-Herránz ML. Phylogeny and Evolution of the Genus Brachypodium. In: Vogel PJ, editor. Genetics and Genomics of Brachypodium. Cham: Springer International Publishing; 2016. p. 9-38.
25. Catalan P, Kellogg EA, Olmstead RG. Phylogeny of Poaceae subfamily Pooideae based on chloroplast ndhF gene sequences. Mol Phylogenet Evol. 1997;8(2):150-66. Epub 1997/09/23. doi: 10.1006/mpev.1997.0416. PubMed PMID: 9299221.
26. Catalan P, Chalhoub B, Chochois V, Garvin DF, Hasterok R, Manzaneda AJ, et al. Update on the genomics and basic biology of *Brachypodium*: International *Brachypodium* Initiative (IBI). Trends Plant Sci. 2014;19(7):414-8. doi: 10.1016/j.tplants.2014.05.002. PubMed PMID: 24917149.
27. Brkljacic J, Grotewold E, Scholl R, Mockler T, Garvin DF, Vain P, et al. *Brachypodium* as a model for the grasses: today and the future. Plant Physiol. 2011;157(1):3-13. Epub 2011/07/21. doi: 10.1104/pp.111.179531. PubMed PMID: 21771916; PubMed Central PMCID: PMCPMC3165879.
28. IBI T. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature. 2010;463(7282):763-8. Epub 2010/02/12. doi: 10.1038/nature08747. PubMed PMID: 20148030.
29. Robertson IH. Chromosome numbers in *Brachypodium* Beauv. (Gramineae). Genetica. 1981;56(1):55-60. doi: 10.1007/BF00126930.
30. Wolny E, Hasterok R. Comparative cytogenetic analysis of the genomes of the model grass *Brachypodium distachyon* and its close relatives. Ann Bot. 2009;104(5):873-81. Epub 2009/07/28. doi: 10.1093/aob/mcp179. PubMed PMID: 19633311; PubMed Central PMCID: PMCPMC2749528.
31. Betekhtin A, Jenkins G, Hasterok R. Reconstructing the Evolution of *Brachypodium* Genomes Using Comparative Chromosome Painting. PLoS One. 2014;9(12):e115108. doi: 10.1371/journal.pone.0115108. PubMed PMID: 25493646; PubMed Central PMCID: PMCPMC4262448.
32. Hasterok R, Draper J, Jenkins G. Laying the cytotaxonomic foundations of a new model grass, *Brachypodium distachyon* (L.) Beauv. Chromosome Res. 2004;12(4):397-403. Epub 2004/07/09. doi: 10.1023/b:chro.0000034130.35983.99. PubMed PMID: 15241018.
33. Catalan P, Muller J, Hasterok R, Jenkins G, Mur LA, Langdon T, et al. Evolution and taxonomic split of the model grass *Brachypodium distachyon*. Ann Bot. 2012;109(2):385-405. Epub 2012/01/04. doi: 10.1093/aob/mcr294. PubMed PMID: 22213013; PubMed Central PMCID: PMCPMC3268539.
34. Lopez-Alvarez D, Lopez-Herranz ML, Betekhtin A, Catalan P. A DNA barcoding method to discriminate between the model plant *Brachypodium distachyon* and its close relatives *B. stacei* and *B. hybridum* (Poaceae). PLoS One. 2012;7(12):e51058. Epub 2012/12/15. doi:

10.1371/journal.pone.0051058. PubMed PMID: 23240000; PubMed Central PMCID: PMCPMC3519806.
35. Gordon SP, Liu L, Vogel JP. The Genus Brachypodium as a Model for Perenniality and Polyploidy. In: Vogel PJ, editor. Genetics and Genomics of Brachypodium. Cham: Springer International Publishing; 2016. p. 313-25.
36. Mason AS, Pires JC. Unreduced gametes: meiotic mishap or evolutionary mechanism? Trends in Genetics. 2015;31(1):5-10. doi: 10.1016/j.tig.2014.09.011.
37. Ramsey J, Schemske DW. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual Review of Ecology and Systematics. 1998;29:467-501. doi: 10.1146/annurev.ecolsys.29.1.467. PubMed PMID: WOS:000077648200017.
38. Osabe K, Kawanabe T, Sasaki T, Ishikawa R, Okazaki K, Dennis ES, et al. Multiple mechanisms and challenges for the application of allopolyploidy in plants. International journal of molecular sciences. 2012;13(7):8696-721. Epub 2012/09/04. doi: 10.3390/ijms13078696. PubMed PMID: 22942729; PubMed Central PMCID: PMCPmc3430260.
39. Chen ZJ, Ni Z. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. BioEssays : news and reviews in molecular, cellular and developmental biology. 2006;28(3):240-52. Epub 2006/02/16. doi: 10.1002/bies.20374. PubMed PMID: 16479580; PubMed Central PMCID: PMCPmc1986666.
40. Szadkowski E, Eber F, Huteau V, Lode M, Coriton O, Jenczewski E, et al. Polyploid formation pathways have an impact on genetic rearrangements in resynthesized *Brassica napus*. New Phytol. 2011;191(3):884-94. doi: 10.1111/j.1469-8137.2011.03729.x. PubMed PMID: 21517871.
41. Madlung A, Masuelli RW, Watson B, Reynolds SH, Davison J, Comai L. Remodeling of DNA Methylation and Phenotypic and Transcriptional Changes in Synthetic *Arabidopsis* Allotetraploids. Plant Physiology. 2002;129(2):733-46. doi: 10.1104/pp.003095.
42. Lim KY, Souckova-Skalicka K, Sarasan V, Clarkson JJ, Chase MW, Kovarik A, et al. A genetic appraisal of a new synthetic *Nicotiana tabacum* (*Solanaceae*) and the Kostoff synthetic tobacco. Am J Bot. 2006;93(6):875-83. Epub 2006/06/01. doi: 10.3732/ajb.93.6.875. PubMed PMID: 21642150.
43. Mestiri I, Chagué V, Tanguy AM, Huneau C, Huteau V, Belcram H, et al. Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity. New Phytol. 2010;186(1):86-101. doi: 10.1111/j.1469-8137.2010.03186.x. PubMed PMID: 20149116.
44. Ng DW, Zhang C, Miller M, Shen Z, Briggs SP, Chen ZJ. Proteomic divergence in *Arabidopsis* autopolyploids and allopolyploids and their progenitors. Heredity (Edinb). 2012;108(4):419-30. doi: 10.1038/hdy.2011.92. PubMed PMID: 22009271; PubMed Central PMCID: PMC3313054.
45. Comai L. Genetic and epigenetic interactions in allopolyploid plants. Plant Mol Biol. 2000;43(2-3):387-99. PubMed PMID: 10999418.
46. Madlung A, Tyagi AP, Watson B, Jiang H, Kagochi T, Doerge RW, et al. Genomic changes in synthetic *Arabidopsis* polyploids. Plant J. 2005;41(2):221-30. doi: 10.1111/j.1365-313X.2004.02297.x. PubMed PMID: 15634199.
47. Vogel JP, Gu YQ, Twigg P, Lazo GR, Laudencia-Chingcuanco D, Hayden DM, et al. EST sequencing and phylogenetic analysis of the model grass *Brachypodium distachyon*. Theor Appl Genet. 2006;113(2):186-95. Epub 2006/06/23. doi: 10.1007/s00122-006-0285-3. PubMed PMID: 16791686.
48. Ludwig-Muller J, Vertocnik A, Town CD. Analysis of indole-3-butyric acid-induced adventitious root formation on *Arabidopsis* stem segments. J Exp Bot. 2005;56(418):2095-105. Epub 2005/06/16. doi: 10.1093/jxb/eri208. PubMed PMID: 15955788.
49. Jahier J. Techniques of plant cytogenetics. INRA Edition ed: INRA Edition, Paris, France; 1992.
50. Cousin A, Heel K, Cowling WA, Nelson MN. An efficient high-throughput flow cytometric method for estimating DNA ploidy level in plants. Cytometry Part A. 2009;75A(12):1015-9. doi: 10.1002/cyto.a.20816.

51. Besnard G, Garcia-Verdugo C, Rubio De Casas R, Treier UA, Galland N, Vargas P. Polyploidy in the Olive Complex (*Olea europaea*): Evidence from Flow Cytometry and Nuclear Microsatellite Analyses. Annals of Botany. 2008;101(1):25-30. doi: 10.1093/aob/mcm275.
52. Dart S, Kron P, Mable BK. Characterizing polyploidy in *Arabidopsis lyrata* using chromosome counts and flow cytometry. Canadian Journal of Botany. 2004;82(2):185-97. doi: doi:10.1139/b03-134.
53. Suay L, Zhang D, Eber F, Jouy H, Lode M, Huteau V, et al. Crossover rate between homologous chromosomes and interference are regulated by the addition of specific unpaired chromosomes in Brassica. New Phytol. 2014;201(2):645-56. Epub 2013/10/15. doi: 10.1111/nph.12534. PubMed PMID: 24117470.
54. Leflon M, Eber F, Letanneur JC, Chelysheva L, Coriton O, Huteau V, et al. Pairing and recombination at meiosis of *Brassica rapa* (AA) x *Brassica napus* (AACC) hybrids. Theoretical and Applied Genetics. 2006;113:1467-80.
55. Hasterok R, Jenkins G, Langdon T, Jones RN. The nature and destiny of translocated B-chromosome-specific satellite DNA of rye. Chromosome Res. 2002;10(1):83-6. PubMed PMID: 11863075.
56. Hasterok R, Dulawa J, Jenkins G, Leggett M, Langdon T. Multi-substrate chromosome preparations for high throughput comparative FISH. BMC biotechnology. 2006;6:20. doi: 10.1186/1472-6750-6-20. PubMed PMID: 16549000; PubMed Central PMCID: PMC1481663.
57. Gerlach WL, Bedbrook JR. Cloning and characterization of ribosomal RNA genes in wheat. Nucleic Acids Research. 1979;7:1869-85.
58. Charles M, Belcram H, Just J, Huneau C, Viollet A, Couloux A, et al. Dynamics and differential proliferation of transposable elements during the evolution of the B and A genomes of wheat. Genetics. 2008;180(2):1071-86. Epub 2008/09/11. doi: 10.1534/genetics.108.092304. PubMed PMID: 18780739; PubMed Central PMCID: PMCPMC2567357.
59. Giraldo P, Rodriguez-Quijano M, Vazquez JF, Carrillo JM, Benavente E. Validation of microsatellite markers for cytotype discrimination in the model grass *Brachypodium distachyon*. Genome. 2012;55(7):523-7. Epub 2012/07/14. doi: 10.1139/g2012-039. PubMed PMID: 22788413.
60. Garvin DF, McKenzie N, Vogel JP, Mockler TC, Blankenheim ZJ, Wright J, et al. An SSR-based genetic linkage map of the model grass *Brachypodium distachyon*. Genome. 2010;53(1):1-13. Epub 2010/02/05. doi: 10.1139/g09-079. PubMed PMID: 20130744.
61. Hammami J, Soler, Frieiro, González. Genetic diversity of SSR and ISSR markers in wild populations of *Brachypodium distachyon* and its close relatives *B. stacei* and *B. hybridum* (Poaceae). Plant Systematics and Evolution. 2014; 300:2029-40.
62. Vogel JP, Tuna M, Budak H, Huo N, Gu YQ, Steinwand MA. Development of SSR markers and analysis of diversity in Turkish populations of *Brachypodium distachyon*. BMC Plant Biol. 2009;9:88. Epub 2009/07/15. doi: 10.1186/1471-2229-9-88. PubMed PMID: 19594938; PubMed Central PMCID: PMCPMC2719641.
63. Kruskal WH, Wallis WA. Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association. 1952;47(260):583-621. doi: 10.2307/2280779.
64. Nasrallah ME, Yogeeswaran K, Snyder S, Nasrallah JB. *Arabidopsis* species hybrids in the study of species differences and evolution of amphiploidy in plants. Plant Physiol. 2000;124(4):1605-14. Epub 2000/12/15. PubMed PMID: 11115878; PubMed Central PMCID: PMCPMC59859.
65. Choudhary BR, Joshi P, Ramarao S. Interspecific hybridization between *Brassica carinata* and *Brassica rapa*. Plant Breeding. 2000;119(5):417-20. doi: 10.1046/j.1439-0523.2000.00503.x.
66. Gangadevi T, Rao PN, Rao BH, Satyanarayana KV. A study of morphology, cytology and sterility in interspecific hybrids and amphidiploids of *Nicotiana knightiana* X N. *umbratica*. Theor Appl Genet. 1985;70(3):330-2. Epub 1985/06/01. doi: 10.1007/bf00304921. PubMed PMID: 24252931.

67. Poysa V. The development of bridge lines for interspecific gene transfer between *Lycopersicon esculentum* and *L. peruvianum*. Theor Appl Genet. 1990;79(2):187-92. Epub 1990/02/01. doi: 10.1007/bf00225950. PubMed PMID: 24226217.
68. Hong SY, Park JH, Cho SH, Yang MS, Park CM. Phenological growth stages of *Brachypodium distachyon*: codification and description. Weed Research. 2011;51(6):612-20. doi: 10.1111/j.1365-3180.2011.00877.x.
69. Idziak D, Hasterok R. Cytogenetic evidence of nucleolar dominance in allotetraploid species of *Brachypodium*. Genome. 2008;51(5):387-91. Epub 2008/04/29. doi: 10.1139/g08-017. PubMed PMID: 18438442.
70. Khan MA, Stace CA. Breeding relationships in the genus *Brachypodium* (Poaceae: Pooideae). Nordic Journal of Botany1999. p. 257-68.
71. Idziak D, Hazuka I, Poliwczak B, Wiszynska A, Wolny E, Hasterok R. Insight into the karyotype evolution of *Brachypodium* species using comparative chromosome barcoding. PLoS One. 2014;9(3):e93503. Epub 2014/03/29. doi: 10.1371/journal.pone.0093503. PubMed PMID: 24675822; PubMed Central PMCID: PMCPMC3968144.
72. Yao X, Ge X, Li Z. Different fertility and meiotic regularity in allohexaploids derived from trigenomic hybrids between three cultivated *Brassica* allotetraploids and *B. maurorum*. Plant Cell Rep. 2012;31(4):781-8. Epub 2011/12/08. doi: 10.1007/s00299-011-1200-1. PubMed PMID: 22147137.
73. Tate JA, Symonds VV, Doust AN, Buggs RJ, Mavrodiev E, Majure LC, et al. Synthetic polyploids of Tragopogon miscellus and T. mirus (Asteraceae): 60 Years after Ownbey's discovery. Am J Bot. 2009;96(5):979-88. doi: 10.3732/ajb.0800299. PubMed PMID: 21628250.
74. Mejia-Jimenez A, Munoz C, Jacobsen HJ, Roca WM, Singh SP. Interspecific hybridization between common and tepary beans: increased hybrid embryo growth, fertility, and efficiency of hybridization through recurrent and congruity backcrossing. Theor Appl Genet. 1994;88(3-4):324-31. Epub 1994/06/01. doi: 10.1007/bf00223640. PubMed PMID: 24186014.
75. Presgraves DC, Glor RE. Evolutionary biology: speciation on islands. Current biology : CB. 2010;20(10):R440-2. doi: 10.1016/j.cub.2010.03.032. PubMed PMID: 20504751.
76. Sitch LA, Snape JW. Factors affecting haploid production in wheat using the *Hordeum bulbosum* system. 3. Post-fertilization effects on embryo survival. Euphytica. 1987;36(2):497-504.
77. Taylor NL, Quarles RF, Anderson MK. Methods of overcoming interspecific barriers in *Trifolium*. Euphytica. 1980;29(2):441-50.
78. Lee C, Page L, McClure B, Holtsford T. Post-pollination hybridization barriers in *Nicotiana* section Alatae. Sex Plant Reprod. 2008;21(3):183-95. doi: 10.1007/s00497-008-0077-9.
79. Natalis LC, Wesselingh RA. Post-pollination barriers and their role in asymmetric hybridization in *Rhinanthus* (Orobanchaceae). American Journal of Botany. 2012;99(11):1847-56. doi: 10.3732/ajb.1200085.
80. Scopece G, Widmer A, Cozzolino S. Evolution of postzygotic reproductive isolation in a guild of deceptive orchids. The American naturalist. 2008;171(3):315-26. Epub 2008/01/18. doi: 10.1086/527501. PubMed PMID: 18198999.
81. Ouyang Y, Zhang Q. Understanding reproductive isolation based on the rice model. Annu Rev Plant Biol. 2013;64:111-35. Epub 2013/05/04. doi: 10.1146/annurev-arplant-050312-120205. PubMed PMID: 23638826.
82. Beaulieu J, Jean M, Belzile F. The allotetraploid *Arabidopsis thaliana-Arabidopsis lyrata* subsp. *petraea* as an alternative model system for the study of polyploidy in plants. Mol Genet Genomics. 2009;281(4):421-35. doi: 10.1007/s00438-008-0421-7. PubMed PMID: 19148683.
83. Fujimoto R, Taylor JM, Sasaki T, Kawanabe T, Dennis ES. Genome wide gene expression in artificially synthesized amphidiploids of *Arabidopsis*. Plant Mol Biol. 2011;77(4-5):419-31. doi: 10.1007/s11103-011-9820-y. PubMed PMID: 21882042.
84. Akama S, Shimizu-Inatsugi R, Shimizu KK, Sese J. Genome-wide quantification of homeolog expression ratio revealed nonstochastic gene regulation in synthetic allopolyploid *Arabidopsis*.

Nucleic Acids Res. 2014;42(6):e46. doi: 10.1093/nar/gkt1376. PubMed PMID: 24423873; PubMed Central PMCID: PMC3973336.
85. Mariam AL, Zakri AH, Mahani MC, Normah MN. Interspecific hybridization of cultivated rice, *Oryza sativa* L. with the wild rice, O. *minuta* Presl. Theor Appl Genet. 1996;93(5-6):664-71. doi: 10.1007/BF00224060. PubMed PMID: 24162392.

Supporting information captions.

S1 Text. Phenotypic characterization of synthetic allopolyploids.

S2 Text. Classification and interpretation of marker polymorphism

S1 Fig. Illustration of vegetative propagation of plants from F1 interspecific hybrids, S0 and S1 generations of *Brachypodium* synthetic allopolyploids. (A) The plant has new tillers with new secondary roots emerging from node (indicated by arrows) for cutting. (B) Stimulation rooting by burying tiller node in the soil containing 0.25% indole-3-butyric Acid (IBA) and the tiller was fixed with a stick. (C) Young tillers were cut from the initial plant: (1) A tiller with enough roots to live independently, (2) A tiller with no roots. (D) The tiller (1) from Figure S1C was transferred directly into a pot to produce a new plant. (E) Rootless tiller (2) was soaked into water containing 0.25% IBA. After 7-10 days, the root emerged (indicated by blue arrows) and grew enough to transfer this tiller into a new pot.

S2 Fig. Phenotypic characters recorded in the present study. 1. Inflorescence length; 2. Spikelet length; 3. Distance between two spikelets in florescence; 4. Upper glume from basal spikelet length; 5. Upper glume from basal spikelet width; 6. floret length (the second floret of all spikelets on one inflorescence were taken to measure); 7. Awn length; 8. Lemma length; 9. Lemma width. Other characters considered for inflorescence or spikelet: 10. Spikelet number per inflorescence (all

spikelets in the spike - 5 in this example); 11. Floret number per spikelet (i.e. floret number of each spikelet); 12. Floret number per inflorescence (i.e. the sum of all floret in all spikelets in inflorescence); 13. Seed number per inflorescence; 14. Percentage of fertile floret; 15. Weight of 1,000

S3 Fig. Male sterility in the F1 interspecific hybrid F1_21×114 and S1 generation of its derived synthetic allopolyploid allo21×114, compared to diploid progenitors *B. distachyon (*Bd21) and *B. stacei (*ABR114) and the natural allopolyploid *B. hybridum* (ABR113). (A) Anthers on the day of anthesis. (B) Spontanouous release of the pollen from anthers after 15-20 minutes on the microscope slides: (B) Bd21, (C) ABR113 and (D) ABR114. This phenomenon was not observed for (E) F1_21×114 and (F) S1 generation of allo21×114; after macerating these anthers there was very little pollen compared with progenitors and natural polyploid.

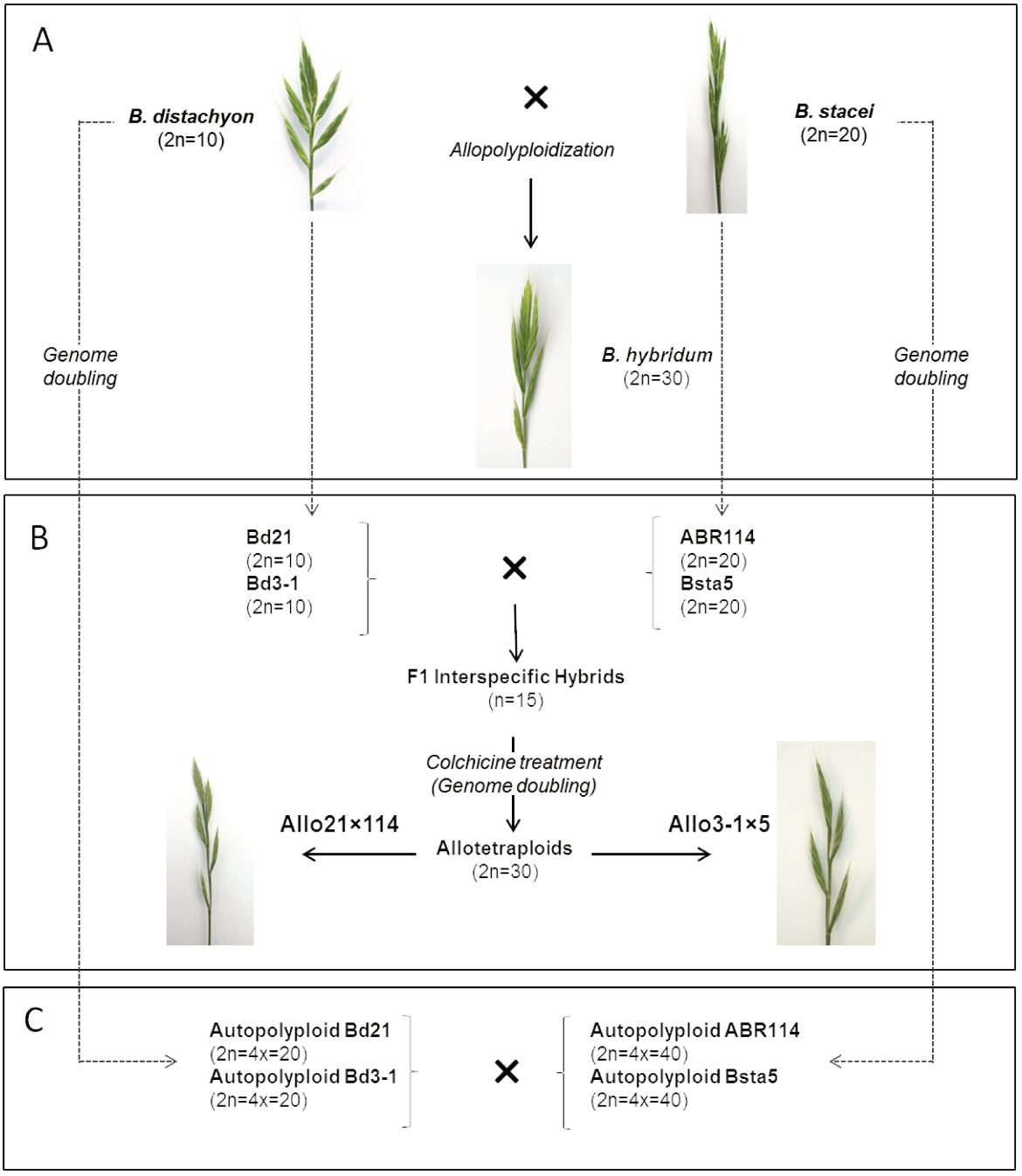
S1 Table. Results of the vegetative multiplication of haploid F1 interspecific hybrids and their treatment with varying concentrations of colchicine (2.5 g/l, 5g/l and 7.5 g/l).

S2 Table. Primer sequences and characteristics of PCR-based polymorphic markers derived from gene and SSR sequences.

S3 Table. Mean values ± standard deviation for 15 morphological traits measured in different synthetic and natural *Brachypodium* allopolyploids and their parental species.

S4 Table. Summary of comparison of some flower characters at the anthesis stage (with dissecting microscope) between *B. distachyon*, *B. stacei*, *B. hybridum,* interspecific F1 hybrids and plants of S1 generation of the synthetic allopolyploids allo21×114 and allo3-1×5.

S5 Table. Number of crosses made between the diploid parental species *B. distachyon* and *B. stacei* as well as *B. hybridum* natural allopolyploids with the interspecific hybrid F1_21×114 and the synthetic allopolyploid allo21×114 (plants of S1 generation were used here).


A
B. distachyon
(2n=10)
Allopolyploidization
B. stacei
(2n=20)
Genome doubling
B. hybridum
(2n=30)
Genome doubling
B
Bd21
(2n=10)
Bd3-1
(2n=10)
ABR114
(2n=20)
Bsta5
(2n=20)
F1 Interspecific Hybrids
(n=15)
Colchicine treatment
(Genome doubling)
Allo21×114
Allotetraploids
(2n=30)
Allo3-1×5
C
Autopolyploid Bd21
(2n=4x=20)
Autopolyploid Bd3-1
(2n=4x=20)
Autopolyploid ABR114
(2n=4x=40)
Autopolyploid Bsta5
(2n=4x=40)

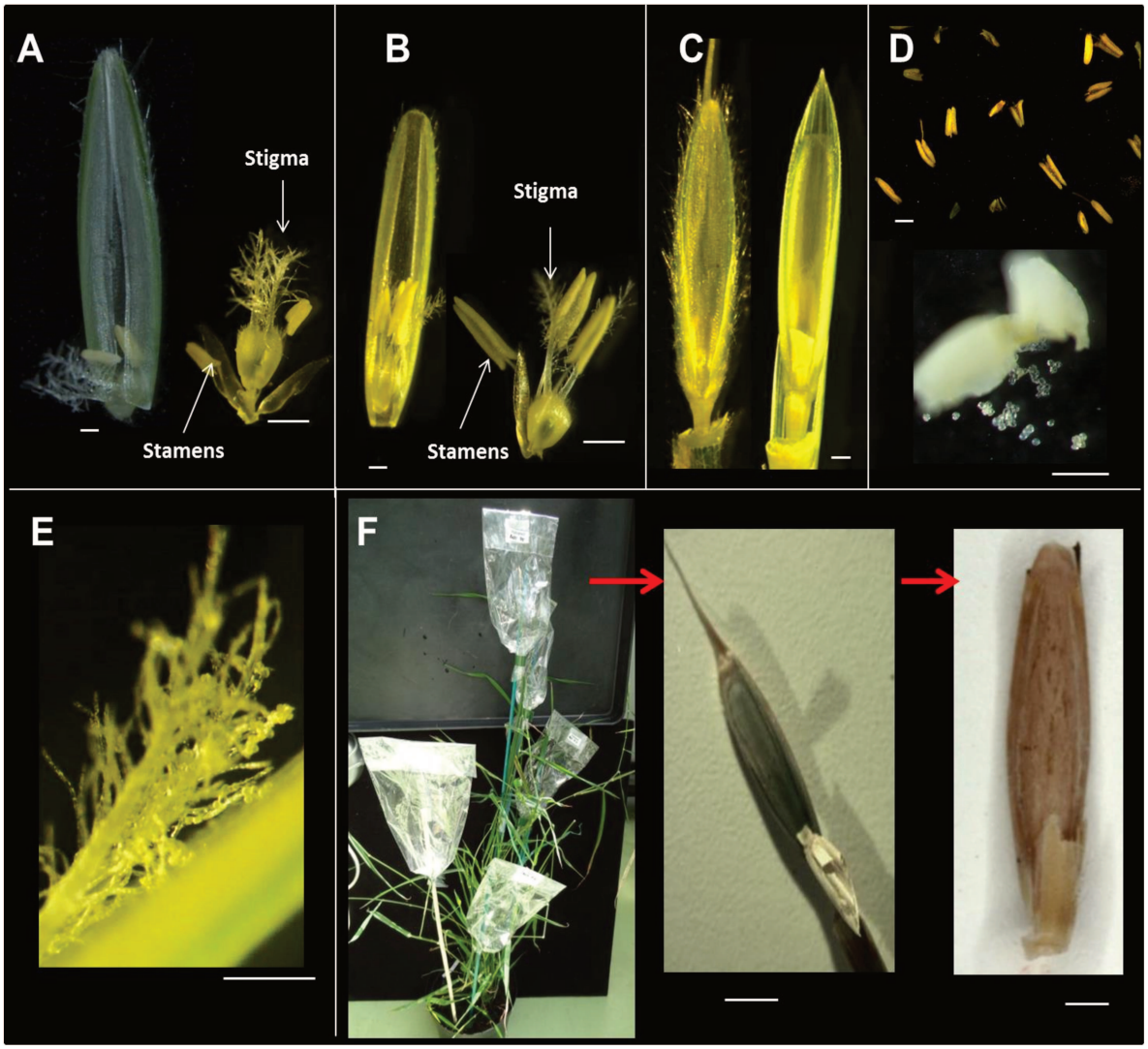


Fig. 2

Figure 3

Click here to download Figure Fig.4.tif

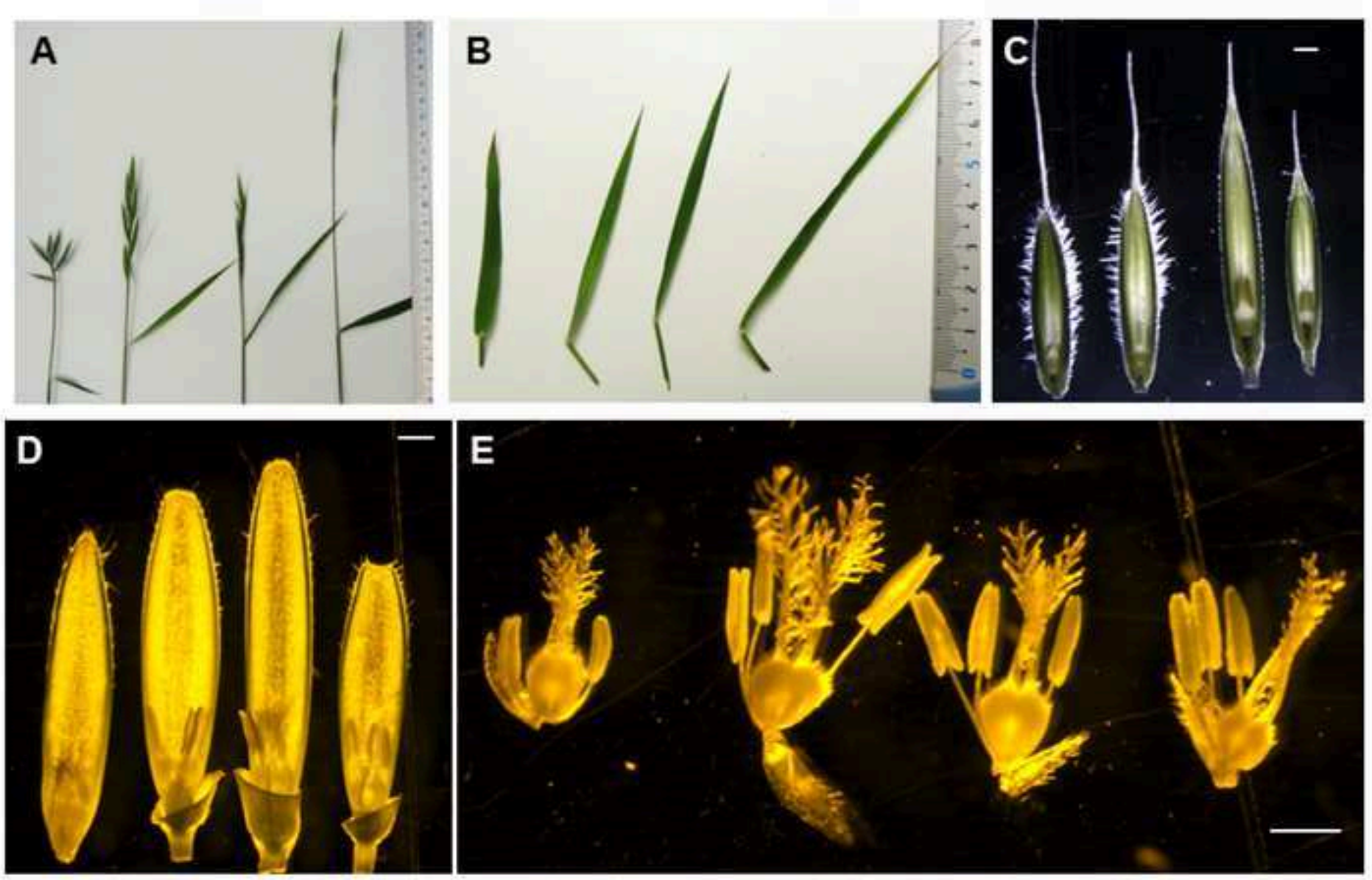
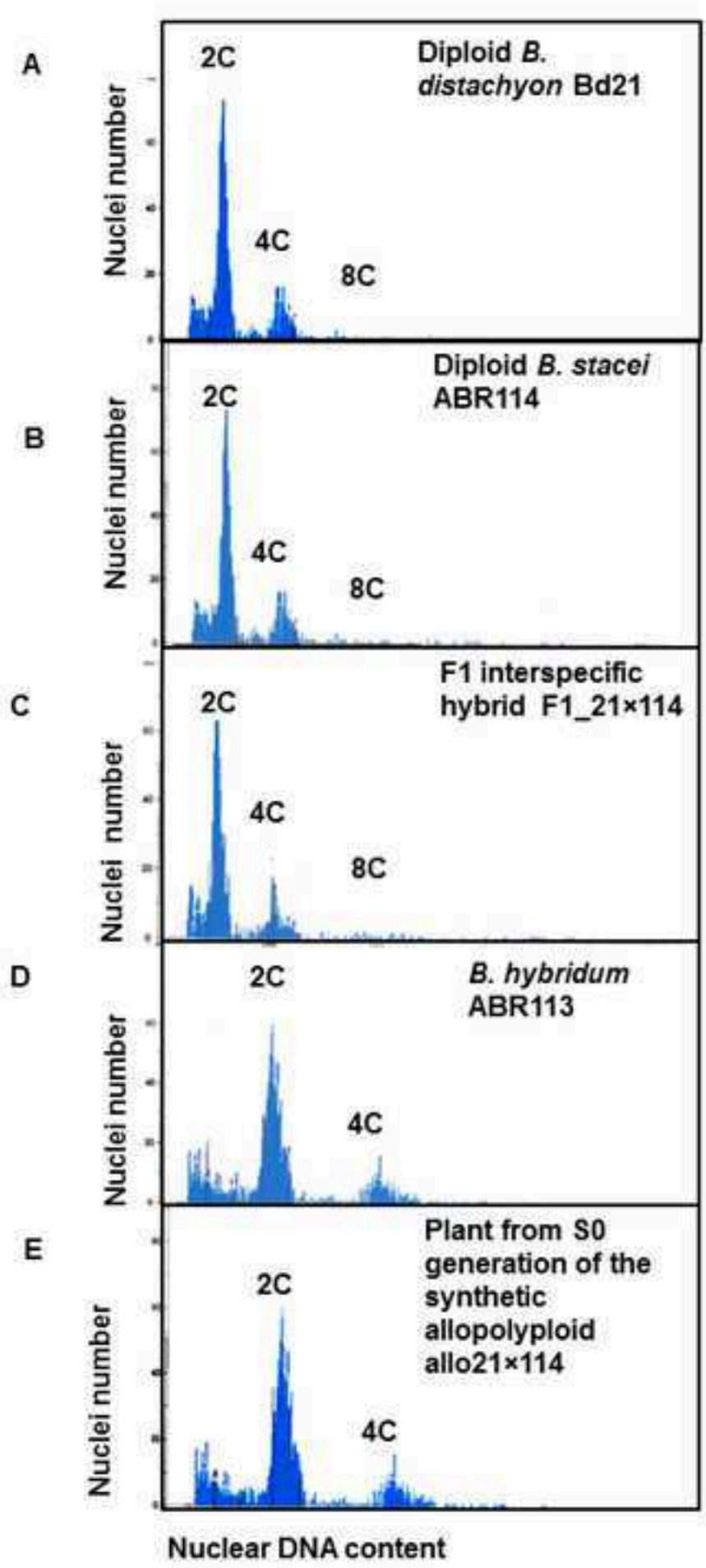


Figure 4

Click here to download Figure Fig.3.tif



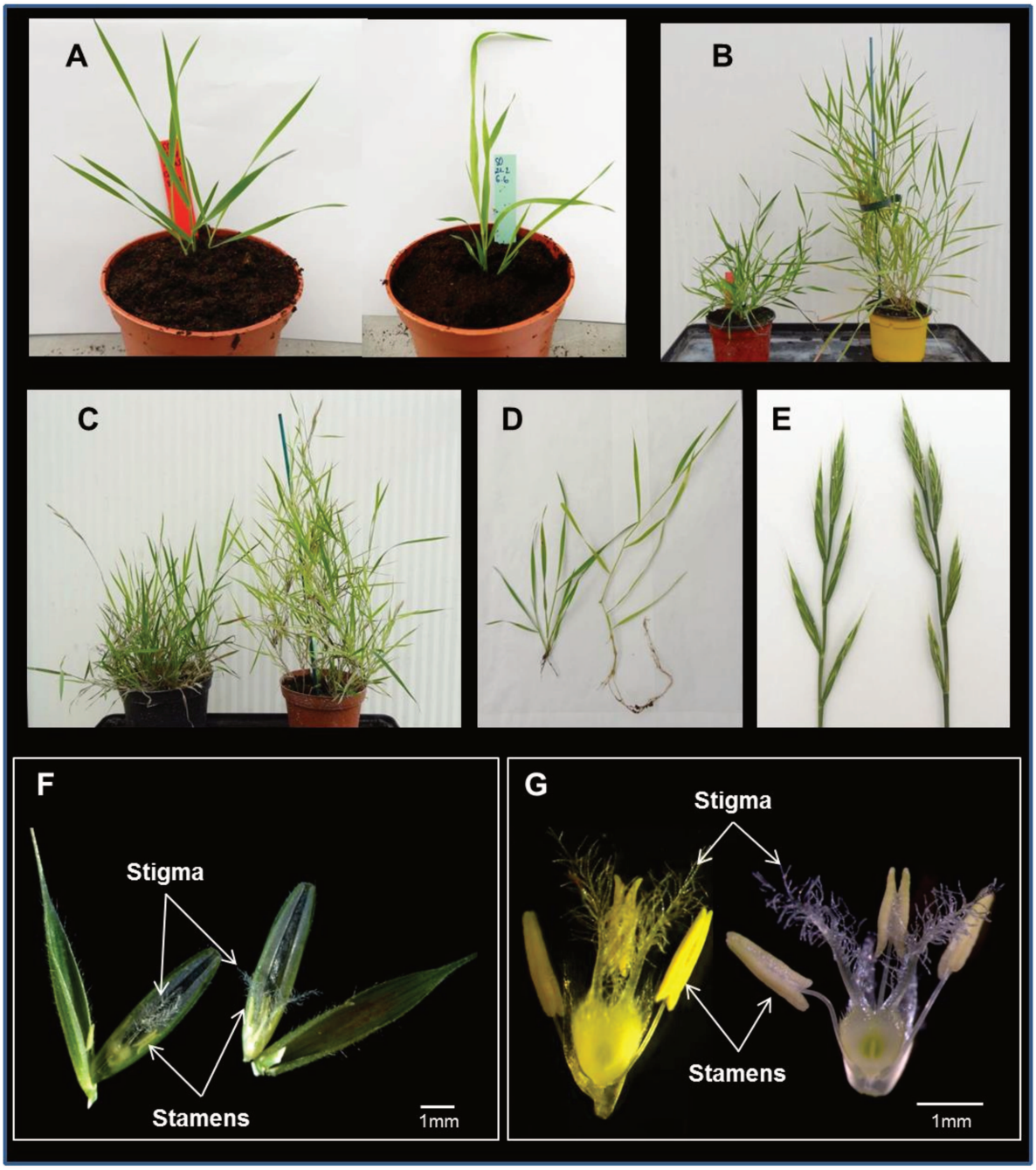


Fig. 5

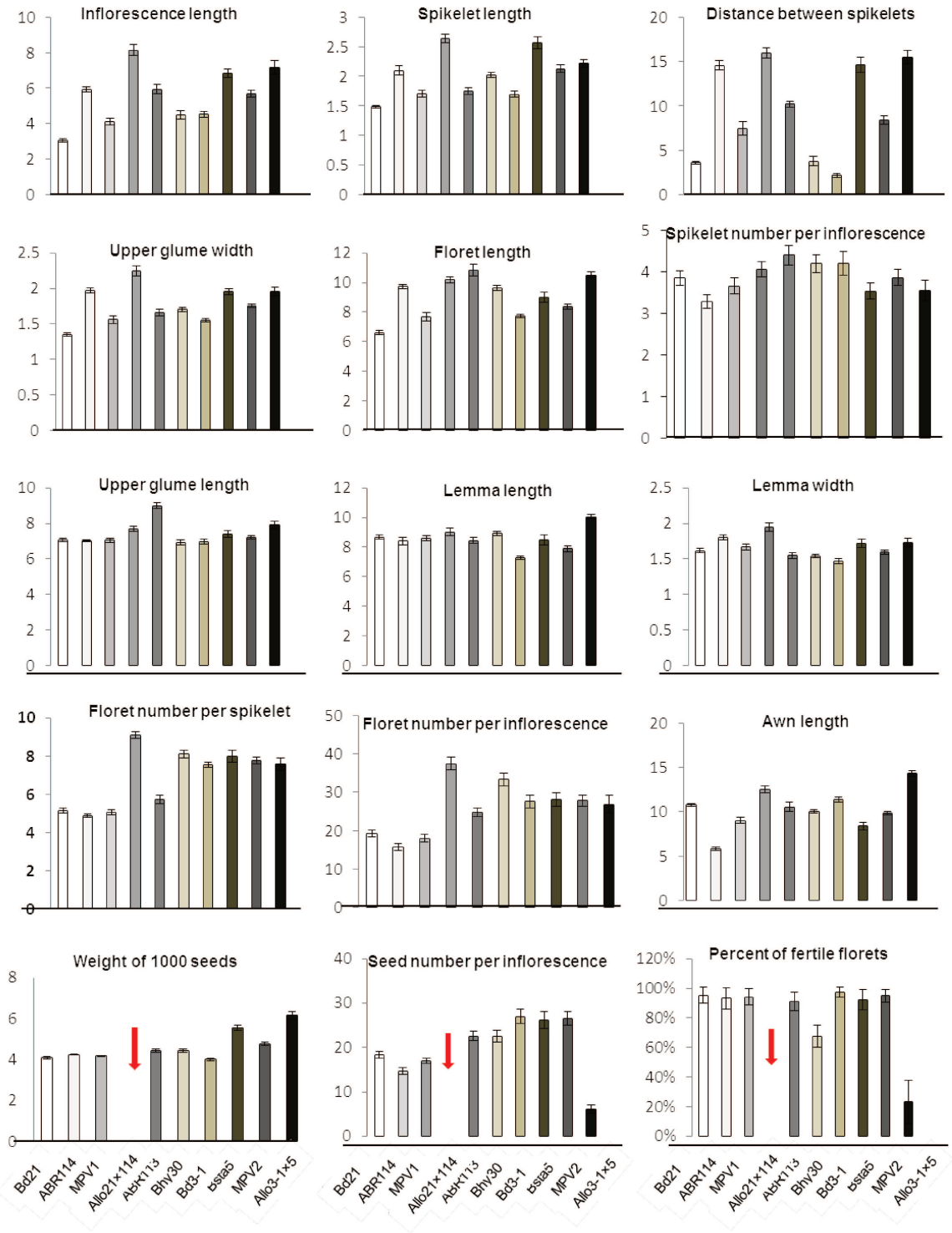
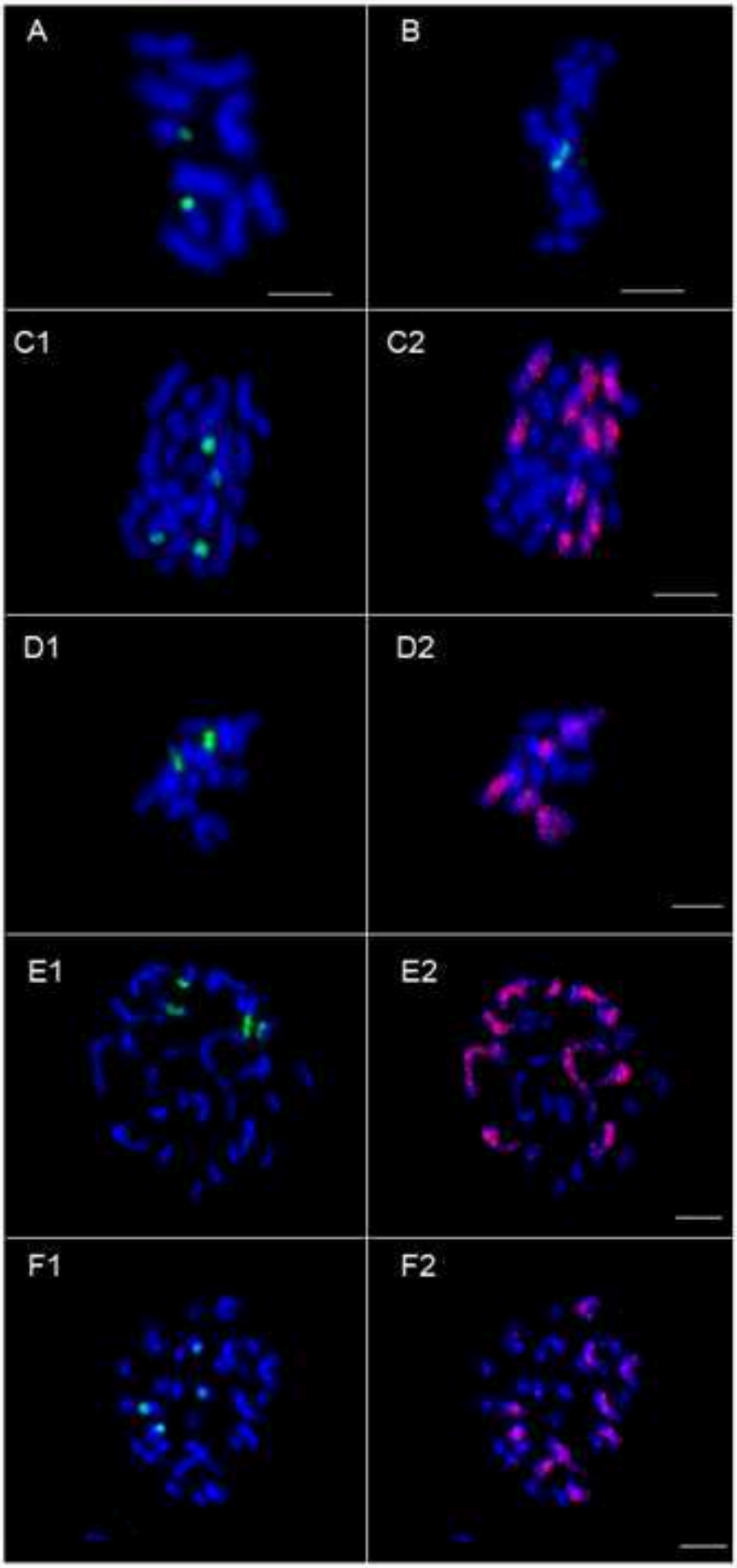

Inflorescence length
Spikelet length
Distance between spikelets
Upper glume width
Floret length
Spikelet number per inflorescence
Upper glume length
Lemma length
Lemma width
Floret number per spikelet
Floret number per inflorescence
Awn length
Weight of 1000 seeds
Seed number per inflorescence
Percent of fertile florets
Bd21
ABR114
MPV1
Allo21×114
ABR113
Bhy30
Bd3-1
Bstai5
MPV2
Allo3-1×5

Fig. 6

Figure 7

Click here to download Figure Fig.7.tif



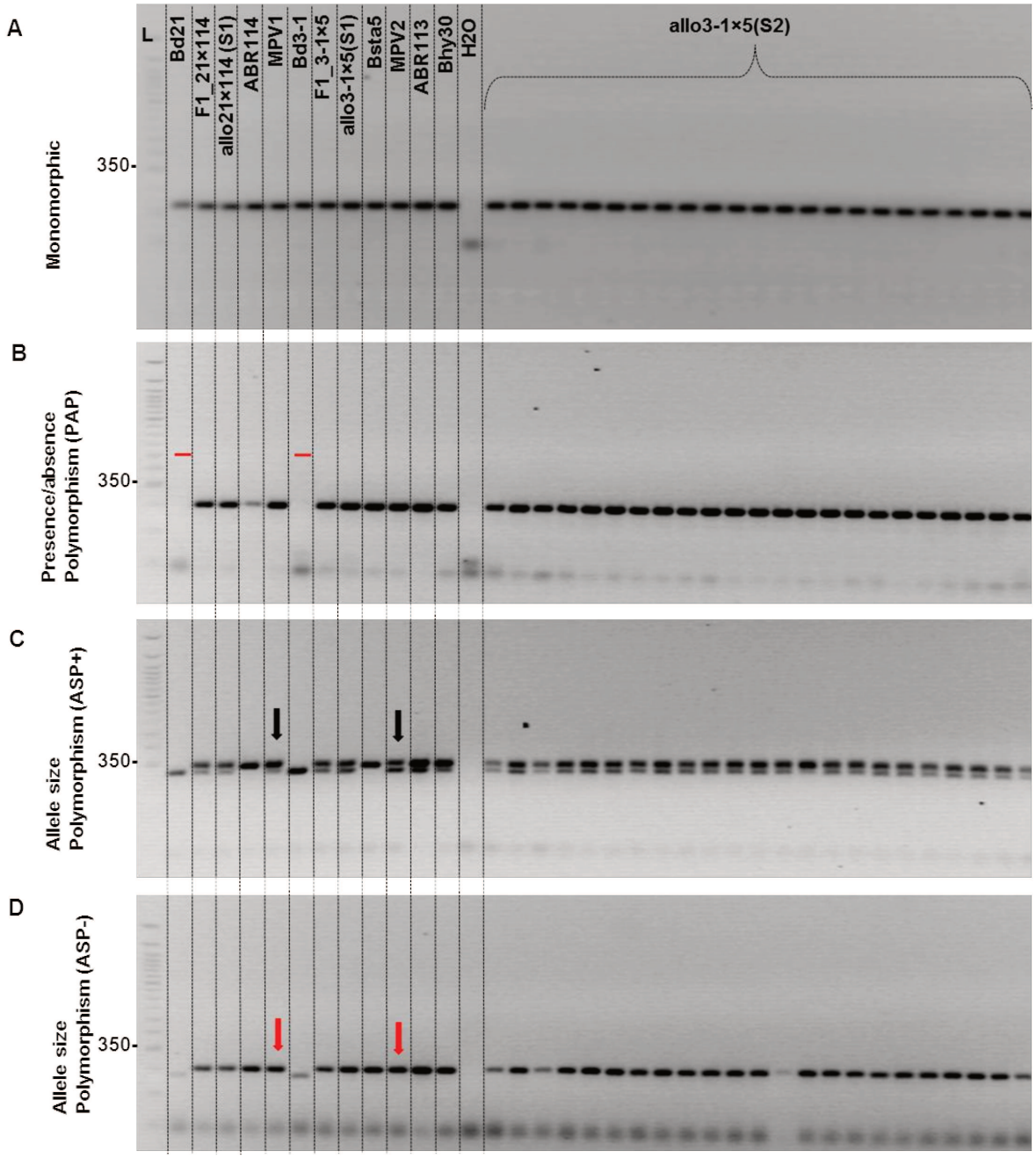
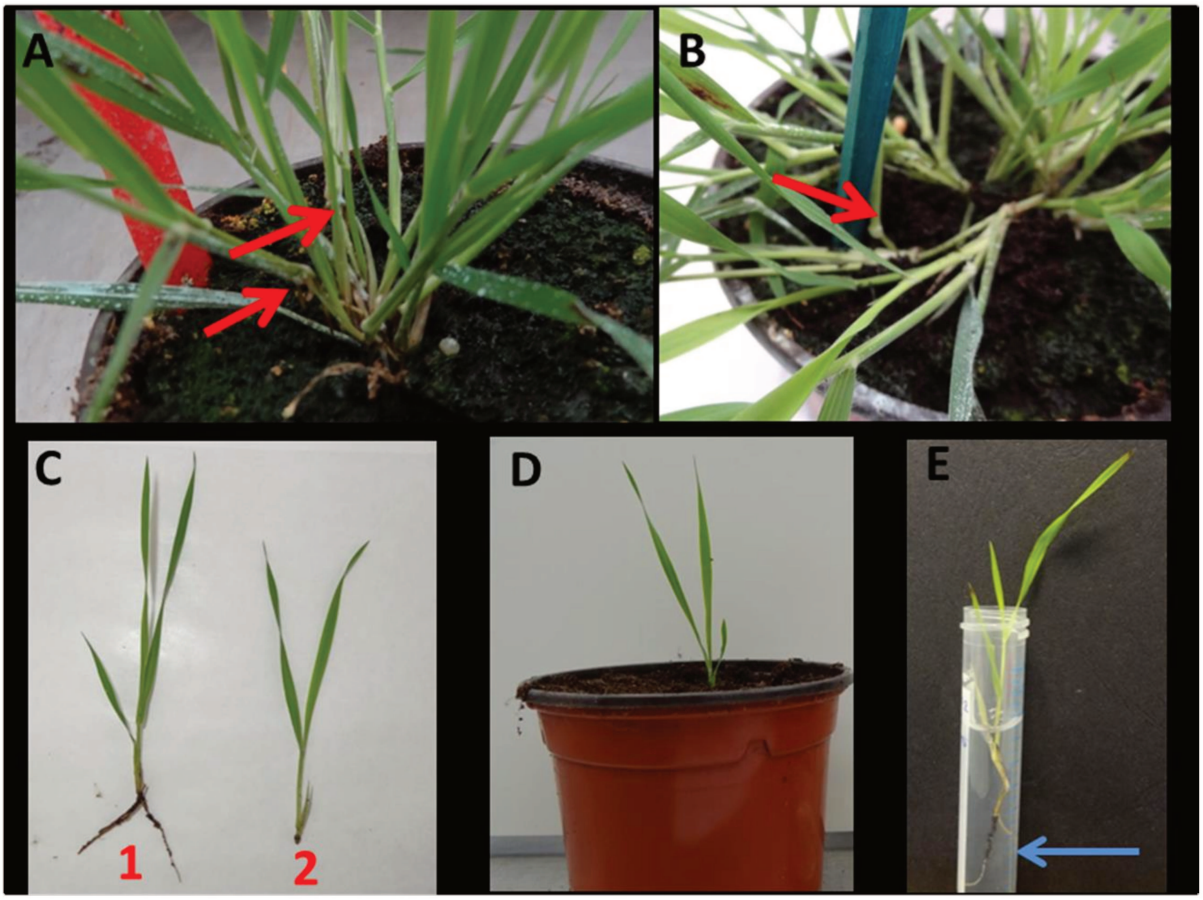
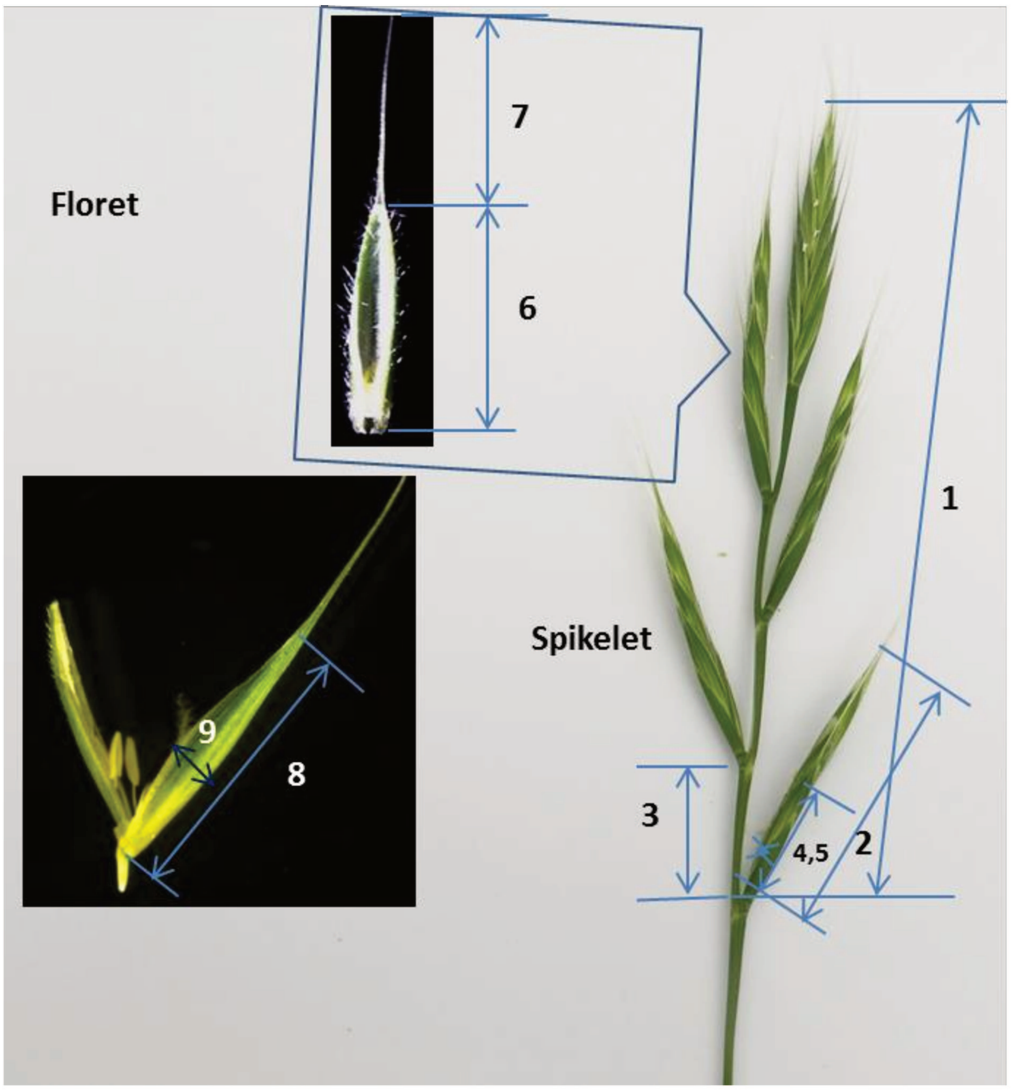
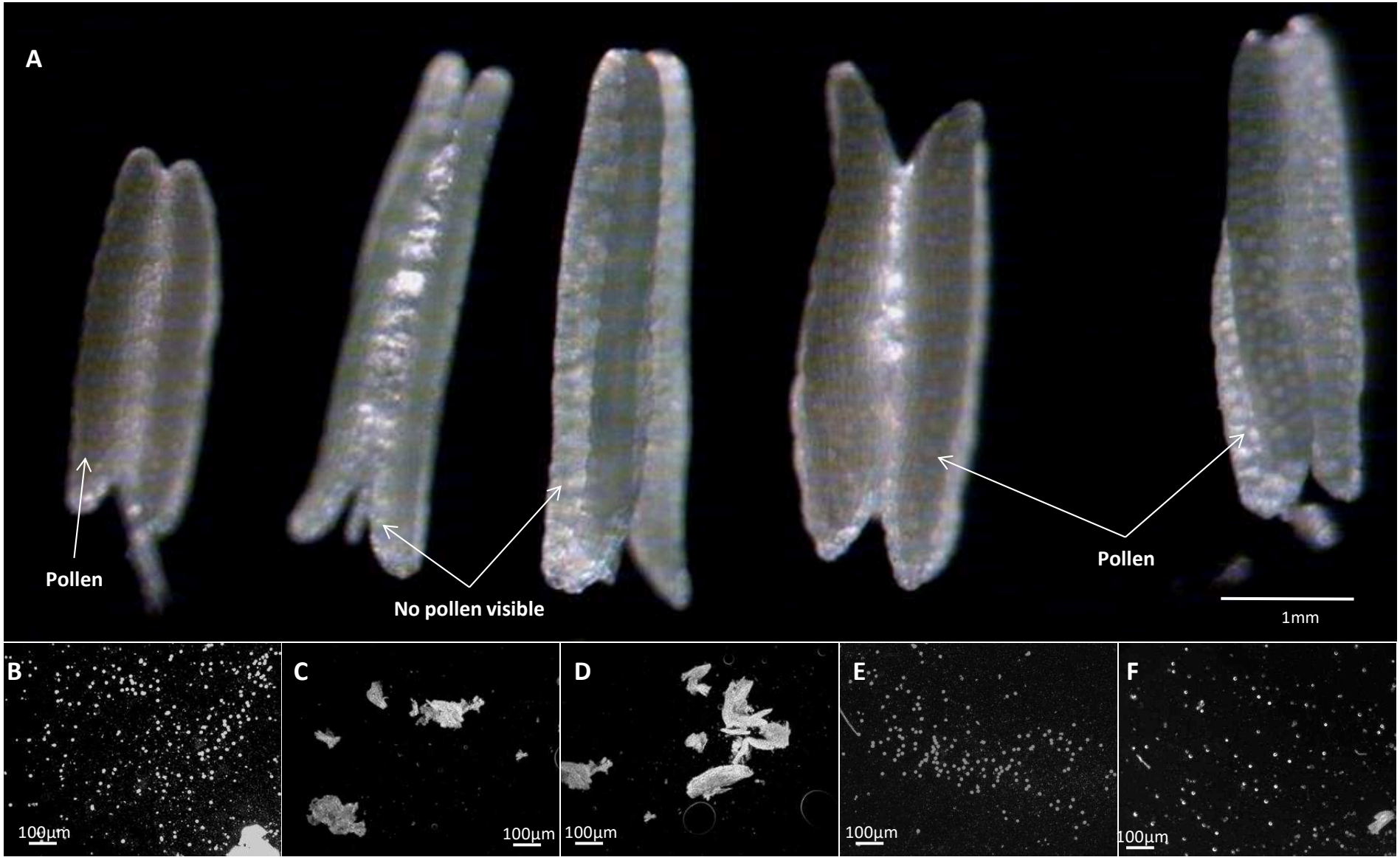

A
Monomorphic
L
Bd21
F1_21×114
allo21×114 (S1)
ABR114
MPV1
Bd3-1
F1_3-1×5
allo3-1×5(S1)
Bsta5
MPV2
ABR113
Bhy30
H2O
allo3-1×5(S2)
350-
B
Presence/absence
Polymorphism (PAP)
350-
C
Allele size
Polymorphism (ASP+)
350-
D
Allele size
Polymorphism (ASP-)
350-

Fig. 8


A
B
C
1
2
D
E


Floret
7
6
1
Spikelet
9
8
3
4,5
2

S3_Fig


A
Pollen
No pollen visible
Pollen
1mm
B
100µm
C
100µm
D
100µm
E
100µm
F
100µm

S1 Text.

Phenotypic characterization of synthetic allopolyploids

The inflorescence length was larger in *B. stacei* than in both parental lines of *B. distachyon*, whereas they were intermediate in the natural *B. hybridum* allopolyploid lines, a result that is in agreement with the findings of Catalan et al. [1]. Inflorescence length was similar in both types of synthetic allopolyploids and significantly longer than in *B. distachyon*, in averaged values of mid parents (MPVs), and in natural allopolyploids. However, compared to *B. stacei*, the inflorescence length was significantly higher in allo21×114 than in the ABR114 paternal line, whereas it was similar inallo3-1×5 and in the *B. stacei* Bsta5 parental line. The spikelet length, the distance between spikelets, the upper glume width and the floret length exhibit similar comparison trends. All four characters were higher in *B. stacei* than in the *B. distachyon* lines. They were also similar between both synthetic allopolyploids and higher than in *B. distachyon,* MPV, and both lines of the *B. hybridum* natural allopolyploid. Both synthetic allopolyploids were not significantly different from their respective *B. stacei* parent lines for all these traits. Spikelet number per inflorescence, upper glume length and lemma length and width were similar between all compared *Brachypodium* natural species and synthetic allopolyploids. The number of florets per spikelet and number of florets per inflorescence were similar in *B. distachyon* Bd21 and *B. stacei* ABR114 parental lines but significantly lower than in other parental lines, i.e. Bd3-1 and Bsta5. In both synthetic allopolyploids allo21×114 and allo3-1×5 they were significantly larger than in their respective parental lines and in the average value of parents. Awn length was the only trait significantly larger in *B. distachyon* than in *B. stacei*. Both synthetic allopolyploids were more similar to *B. distachyon*, with average values larger than those of mid-parents values and equal to those of natural allopolyploids. The average seed weight was not different between Bd21, ABR114 and natural allopolyploid, but significantly higher in Bsta5 and in allo3-1×5 synthetic allopolyploid than in natural *B. hybridum* and Bd3-1. Regarding the two remaining traits indicative of fertility, seed number per inflorescence (SI) and percent of fertile florets, allo21×114 was sterile and allo3-1×5 was fertile but its fertility rate of 23% was significantly lower than those of the natural *B. hybridum* lines ABR113 (91%) and Bhyb30 (68%).

1. **Catalan P, López-Álvarez D, Bellosta C, and Villar L** (2016) Updated taxonomic descriptions, iconography, and habitat preferences of Brachypodium distachyon, B. stacei, and B. hybridum (Poaceae). Anales del Jardín Botánico de Madrid 73.

S2_Text

S2 Text:

Classification and interpretation of marker polymorphism

Monomorphic markers (M) amplify both parental alleles at similar sizes (Fig. 8A) and allowed us to detect genomic rearrangements in the synthetic allopolyploid only if both alleles are deleted or if one of them change in size by deletion or insertions. By contrast polymorphic markers allow us to characterize the synthetic allopolyploids in different ways. The dominant type or presence/absence polymorphic (PAP) markers, amplify an allele from one parent but not from the other (Fig. 8B) and facilitate the observation of inheritance of only one-parental allele. The codominant type or allele specific polymorphic (ASP) markers, amplify both parental alleles at two different (polymorphic) sizes (Fig. 8C) and allow the characterization of both parental alleles in hybrids and allopolyploids. As recommended in Mestiri *et al.* (2010), we separate co-dominant markers that did not amplify appropriately because of technical problems (ASP-), such as competition to PCR amplification (Fig. 8D), using a mixture of parental DNA as a control.

S1 Table. Results of the vegetative multiplication of haploid F1 interspecific hybrids and their treatment with varying concentrations of colchicine (2.5 g/l, 5g/l and 7.5 g/l).

F1 interspecific dihaploid hybrids	**Number of multiplied plants**	**Number of flowering plants**	**Number of plants treated with colchicine and (surviving ones)**		
			2.5 g/l	**5g/l**	**7.5g/l**
F1_21×114.1	75	24	7 (5)	32 (28)	12 (4)
F1_21×114.2	124	46	15 (12)	46 (39)	17 (6)
F1_21×114.3	59	11	11 (7)	25 (19)	12 (2)
F1_21×114.4	64	16	10 (8)	25 (22)	13 (1)
F1_3-1×5	3	2	0 (0)	1 (1)	0 (0)
Total	**325**	**99**	**43**	**129**	**54**

S2_Table

Click here to access/download

Supporting Information

S2 Table.xlsx

S3_Table

Click here to access/download

Supporting Information

S3 Table.xlsx

S4 Table. Summary of comparison of some floral characters at the anthesis stage (with dissecting microscope) between *B. distachyon, B. stacei*, *B. hybridum,* interspecific F1 hybrids and S1 generation synthetic allopolyploids allo21×114 and allo3-1×5 plants.

Characters	***B. distachyon*** (Bd21 and Bd3-1)	***B. stacei*** (ABR114 and Bsta5)	***B. hybridum*** (ABR113 and Bhyb30)	**Interspecific hybrid** (F1_21×114 and F1_31×5)	**Synthetic allopolyploid** (allo21×114 and allo3-1×5)
Anther number	2	3	3	3	3
Anther size[a]	Small	Bigger	Bigger	Bigger	Bigger
Stigma [a]	Very feathery	Less feathery	Less feathery	Less feathery	Less feathery
Anther dehiscence of the first floret in the spike	6-7 days from the spike emergence	4-5 days from the spike emergence	6-7 days from the spike emergence	6-7 days from the spike emergence	6-7 days from the spike emergence
Paleas and lemmas	Plump oval shape, much hairy in border of paleas and outside surface of lemmas	Elongated oval shape, very few even do not have hairs in the border of paleas and outside surface of lemmas	Elongated oval shape, few hairy in border of paleas and outside surface of lemmas	Elongated oval shape, few hairy in border of paleas and outside surface of lemmas	Elongated oval shape, few hairy in border of paleas and outside surface of lemmas

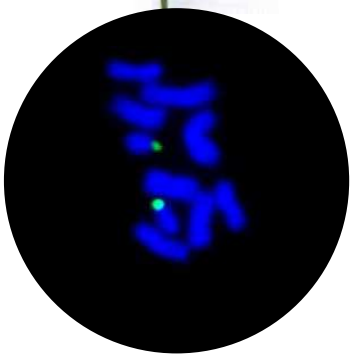
[a] See Fig. 4E for anther and stigma size and shapes

S5_Table

S5 Table. Number of crosses made between the diploid parental species *B. distachyon* and *B. stacei* as well as *B. hybridum* natural allopolyploids with the interspecific hybrid F1_21×114 and the synthetic allopolyploid allo21×114 (at the S1 generation).

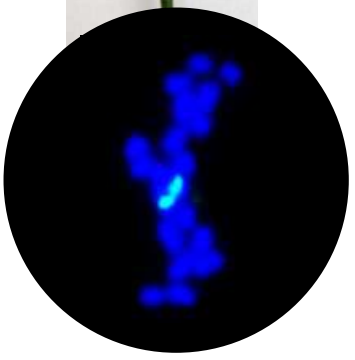
♀ ♂	***B. distachyon*** **2xBd21**	***B. stacei*** **2xABR114**	***B. hybridum*** **ABR113**	***B. hybridum*** **Bhyb30**
F1_21×114	102	95	125	130
allo21×114	105	112	136	127

Striking Image

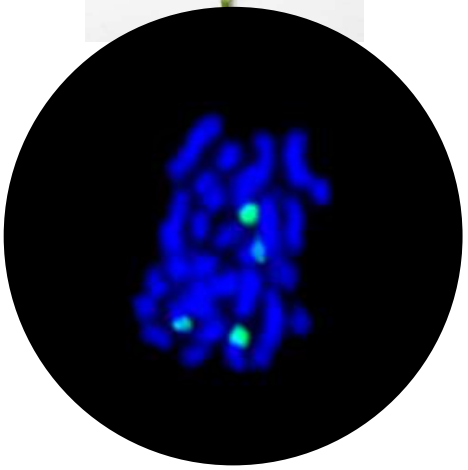


Brachypodium distachyon (2n=10)

X



B. stacei
(2n=20)



B. hybridum
(2n=30)

Cover image

Striking image caption

Cover-image caption

Brachypodium allopolyploid model: Parental species with divergent genomes having two-times chromosome-number and chromosome-size differences and still hybridizing

CHAPTER 3: CHARACTERIZATION OF THE BRACHYPODIUM AUTO- AND ALLOPOLYPLOIDS USING NEXT GENERATION SEQUENCING (NGS) TECHNOLOGIES

CONTEXT

In this chapter, I will update efforts towards genome sequencing of different Brachypodium species, followed by recent advances in next generation sequencing (NGS) technologies together with management and analysis of the huge data that they generate. I will also present the different experiments that are progressing towards characterizing Brachypodium natural and synthetic polyploids that I have constituted earlier (See Chapters I and II, (Dinh-Thi, Clainche et al. 2016, Dinh-Thi, Coriton et al. 2016)), using Illumina next generation sequencing technologies (NGS). I will finally present a pilot case of study on the gene expression changes in autopolyploids of B. distachyon that I was able to characterize.

I. PROGRESS IN GENOME SEQUENCING OF BRACHYPODIUM SPECIES POLYPLOID MODEL

The objective of the first genome sequencing of B. distachyon was initially to develop B. distachyon as a grass model, together with the development of powerful genetic and functional resources (IBI 2010, Mur, Allainguillaume et al. 2011, Catalan, Chalhoub et al. 2014). The sequence data were obtained from a whole genome shot-gun Sanger sequencing with 99.6% of the sequences anchored to five chromosomes (VanBuren and Mockler 2016). This B. distachyon genome sequence represents also the first reported high-quality genome sequence of a species of the Pooideae subfamily. Recently, the new improved assembly (version v3.1) of Bd21 line was released. In this version, the 272 Mb of sequence are arranged into five chromosomes and 22 unmapped scaffolds and cover 99.8% of the genome size (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Bdistachyon_er). The genome resequencing of six other additional inbred lines (Bd21-3, Bd3-1, Bd30-1, Bd1-1, Koz-3, and BdTR12C) of B. distachyon have been also released (Table 2) (Gordon, Priest et al. 2014). Results showed large-scale sequence variants between B. distachyon accessions and revealed more than 2400 previously unannotated genes not present in the Bd21 reference genome (Gordon, Priest et al. 2014). In a further B. distachyon sequencing project, 98 other inbred lines are now being sequenced and de novo assembled by JGI (Joint Genome institute) (Table 2). More than one decade after its suggestion as a tractable grass model (Draper, Mur et al. 2001) and six years after its genome sequencing completed (IBI 2010), B. distachyon has shown a wide range of applications with powerful resources generated for various aspects of grass biology (Brkljacic, Grotewold et al. 2011, Catalan, Chalhoub et al. 2014, Gordon, Liu et al. 2016, Vogel 2016).

The emergence of the Brachypodium polyploid model, having B. hybridum derived from interspecific hybridization between B. distachyon and B. stacei parental species promises to provide valuable resources (Gordon, Liu et al. 2016). The sequencing of B. stacei (ABR114 line) and B. hybridum (ABR113 line) genomes is in progress together with transcriptome analyses through generation of RNA-Seq data from various tissues, comparing B. hybridum with its parental species. The first sequence assembly of B. stacei is being finalized. The ~540 Mb of B. hybridum genome is nearly the sum of B. stacei and B.

distachyon genome sizes (Gordon, Liu et al. 2016), and the genome sequence of this species is also in progress.

As part of Chalhoub's lab contribution to this Brachypodium polyploid sequencing consortium, we have initially generated BAC libraries of B. stacei and B. hybridum to help sequencing the genome through BAC-end sequencing (BES). We have also generated important data using next generation sequencing (NGS) consisting of: (i) DNA resequencing, (ii) transcriptome analysis through RNA-Seq and (iii) identification of CpG methylation through bisulfite sequencing. These experiments were performed on B. hybridum and its two parental species as well as synthetic autopolyploids and allopolyploids that we have generated (Dinh-Thi, Clainche et al. 2016, Dinh-Thi, Coriton et al. 2016). The comparative analysis will allow powerful integrative analysis of genetic, transcriptomic and epigenetic changes related to auto- and allopolyploidization at the whole genome-scale.

Table 2 The list of 98 Brachypodium inbred lines in resequencing project of JGI (Joint Genome institute).

Accession	Status[a]	Accession	Status
Adi-10	de-novo assembly completed analysis in progress	Jer1	de-novo assembly completed analysis in progress
Adi-12	de-novo assembly completed analysis in progress	Arn1	de-novo assembly completed analysis in progress
Adi-2	de-novo assembly completed analysis in progress	Mur1	de-novo assembly completed analysis in progress
Bd1-1*	completed	Uni2	de-novo assembly completed analysis in progress
Bd18-1	de-novo assembly completed analysis in progress	ABR114	completed
Bd2-3	de-novo assembly completed analysis in progress	ABR113	de-novo assembly completed analysis in progress
Bd21 (control)	de-novo assembly completed analysis in progress	Alb-AL2D	sequencing in progress
Bd21-3*	completed	Alb-AL2E	sequencing in progress
Bd3-1_r*	completed	Alb-AL2F	sequencing in progress
Bd30-1*	completed	Arm-Arm3A	sequencing in progress
BdTR10C	de-novo assembly completed analysis in progress	Arm-Arm3G	sequencing in progress
BdTR11A	de-novo assembly completed analysis in progress	Geo-G30i2	sequencing in progress
BdTR11G	de-novo assembly completed analysis in progress	Geo-G31i4	sequencing in progress
BdTR11I	de-novo assembly completed analysis in progress	Geo-G34i2	sequencing in progress
BdTR12C*	completed	Geo-G34i6	sequencing in progress
BdTR13a	de-novo assembly completed analysis in progress	Ita-Sic-CSR6	sequencing in progress
BdTR13C	de-novo assembly completed analysis in progress	Ita-Sic-CSR7	sequencing in progress
BdTR1i	de-novo assembly completed analysis in progress	Ita-Sic-SLZ2	sequencing in progress
BdTR2B	de-novo assembly completed analysis in progress	Spa-Nor-S11A	sequencing in progress
BdTR2G	de-novo assembly completed analysis in progress	Spa-Nor-S11B	sequencing in progress
BdTR3C	de-novo assembly completed analysis in progress	Spa-Nor-S12B	sequencing in progress
BdTR5i	de-novo assembly completed analysis in progress	Spa-Nor-S17D-1	sequencing in progress
BdTR7A	de-novo assembly completed analysis in progress	Spa-Nor-S22B	sequencing in progress
BdTR8i	de-novo assembly completed analysis in progress	Spa-Nor-S22C	sequencing in progress
BdTR9K	de-novo assembly completed analysis in progress	Spa-Nor-S6B	sequencing in progress
Bis-1	de-novo assembly completed analysis in progress	Spa-Nor-S6D	sequencing in progress
Gaz-8	de-novo assembly completed analysis in progress	Spa-Sou-AB1_4	sequencing in progress
Kah-1	de-novo assembly completed analysis in progress	Spa-Sou-CU1_6	sequencing in progress
Kah-5	de-novo assembly completed analysis in progress	Spa-Sou-GR6_4	sequencing in progress
Koz-1	de-novo assembly completed analysis in progress	Spa-Sou-HU3_4	sequencing in progress
Koz-3*	completed	Spa-Sou-J4_3	sequencing in progress
Tek-2	de-novo assembly completed analysis in progress	Spa-Sou-J6_2	sequencing in progress
Tek-4	de-novo assembly completed analysis in progress	Spa-Sou-SG2_1	sequencing in progress
Per-1	de-novo assembly completed analysis in progress	Adi-15	sequencing in progress
Luc1	de-novo assembly completed analysis in progress	Adi-4	sequencing in progress
S8iiC	de-novo assembly completed analysis in progress	Adi-9	sequencing in progress
ABR2	de-novo assembly completed analysis in progress	BdTR10d	sequencing in progress
ABR3	de-novo assembly completed analysis in progress	BdTR11e	sequencing in progress
ABR4	de-novo assembly completed analysis in progress	BdTR12b	sequencing in progress
ABR5	de-novo assembly completed analysis in progress	BdTR13b	sequencing in progress
ABR6_r	de-novo assembly completed analysis in progress	BdTR13n	sequencing in progress
ABR7	de-novo assembly completed analysis in progress	BdTR3m	sequencing in progress
ABR8	de-novo assembly completed analysis in progress	BdTR5a	sequencing in progress
ABR9_r	de-novo assembly completed analysis in progress	BdTR9m	sequencing in progress
Foz-1	de-novo assembly completed analysis in progress	Gaz-1	sequencing in progress
Mig3	de-novo assembly completed analysis in progress	Gaz-2	sequencing in progress
Mon3	de-novo assembly completed analysis in progress	Kah-6	sequencing in progress
Sig2	de-novo assembly completed analysis in progress	Koz-5	sequencing in progress

* The initial published in http://onlinelibrary.wiley.com/doi/10.1111/tpj.12569/abstract

[a] Up to 3nd September 2015.

Data from http://jgi.doe.gov/our-science/science-programs/plant-genomics/brachypodium/

II. EVOLUTION OF SEQUENCING TECHNOLOGIES AS POWERFUL TOOLS FOR GENOME ANALYSIS

DNA sequencing is the process of determining the precise order of the four bases adenine (A), guanine (G), cytosine (C), and thymine (T) in a strand of DNA. Starting with the discovery of the structure of DNA (Watson and Crick 1953), great advances have been made in technologies aiming to decipher the DNA code. In the past 35 years, DNA sequencing technologies and applications have undergone tremendous development and provided a vast amount of genome data and subsequently broad range of research areas and multiple applications (reviewed in Liu, Li et al. (2012), Metzker (2005), Kim, Guo et al. (2016)). After the advent of first generation sequencing (Sanger or Maxim and Gilbert sequencing), more than 30 years ago, the development of high-throughput sequencing technologies known as NGS has got a brilliant era during last decades (summarized in Table 3). I will update hereafter the evolution of the DNA sequencing methods and their applications.

II.1 First generation sequencings

We refer here to the first sequencing methods, such as the Sanger (Sanger and Coulson 1975, Sanger, Nicklen et al. 1977) and Maxam - Gilbert (Maxam and Gilbert (1977)) that were developed and widely used for almost ~40 years ago.

II.1.1 Sanger sequencing

The first method for DNA sequencing described by Sanger and Coulson was called “plus and minus” when they determined two sequences of ΦX174 bacteriophage (Sanger and Coulson 1975, Sanger, Nicklen et al. 1977). This method is a mixed-mode process involving synthesis of a complementary DNA template using natural 2’-deoxynucleotides (dNTPs) and termination of synthesis using 2’,3’-dideoxynucleotides (ddNTPs) by DNA polymerase (Fig.18). Each fragment ends with a particular base that is labeled radioactively or with fluorescent dyes. Then all of the fragments are separated according to their length during electrophoresis (via acrylamide gels or capillary electrophoresis). Information regarding the last base is used to determine the original sequence (Fig.18). This method results in a read

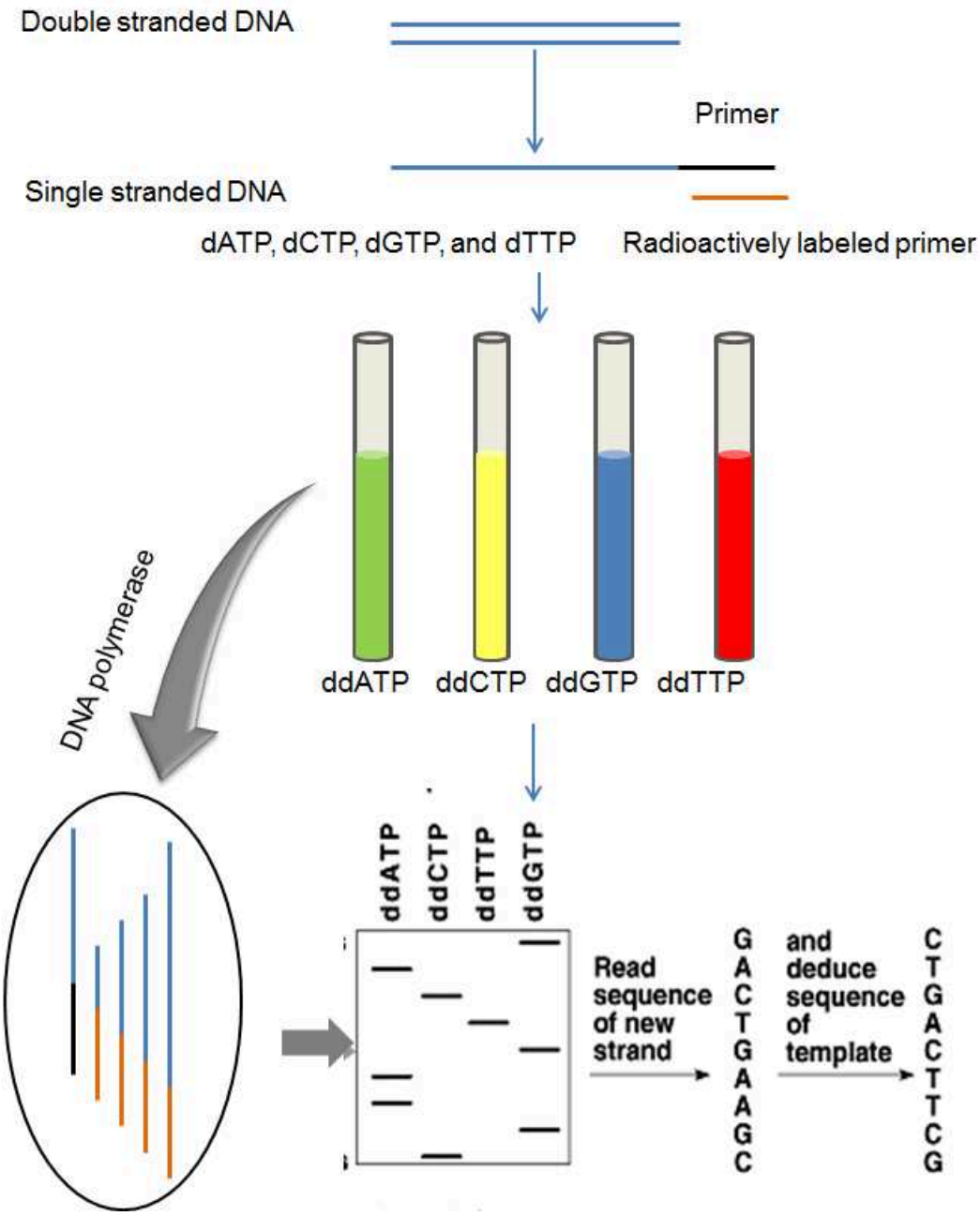


Fig 1. Sanger sequencing method. dsDNA samples separated into ssDNA, a single stranded DNA molecule is amplified and primed using DNA polymerase. This enzyme uses deoxyribonucleoside triphosphates (dNTPs) as substrates and adds them to a primer. The primer is hydrogen bonded to the 3' end of the DNA to be sequenced. The DNA with the primer is divided into four separate reaction mixtures. Each reaction mixture contains all four dNTPs and in addition, one of the four dideoxy analogs (dideoxyribonucleoside triphosphates ddNTPs) of the deoxyribonucleoside triphosphates. Because in the dideoxy sugar the 3'-hydroxyl has been replaced by a hydrogen, continued extension of the chain cannot occur. The dideoxy analog thus acts as specific chain-termination reagent. Fragments of variable length are obtained depending on the ddNTP in the mixture. The formed nucleic acid fragments are visualizes by using either a labelled (radioactive or fluorescent) primer or dNTPs.

From https://www.pearson.ch/HigherEducation/BenjaminCummings

length of ~800 bases but may be extended to above 1000 bases (Table 3)(Schadt, Turner et al. 2010).

The Sanger sequencing method has been improved from the initial radioactive dideoxynucleotide by labeling the four nucleotides reactions with different fluorophores in the same tube, and automated sequencing has been developed by fluorescence detection (Metzker 2005). The Sanger's method is very accurate with relatively and long sequences (Table 3). Its limits are the relatively low throughput with and expensive cost in comparison with NGS methods that were developed 30 years later.

As the method of choice for more than 30 years, Sanger sequencing brings to our knowledge many important data (Egan, Schlueter et al. 2012, Pérez-de-Castro, Vilanova et al. 2012, Hodkinson and Grice 2015), including the first human genome (Human Genome Sequencing 2004), Arabidopsis thaliana and Oryza sativa genome sequences (Initiative 2000, Goff, Ricke et al. 2002). Other reference genomes sequences of many species were also completed using the Sanger method for such as maize (Schnable, Ware et al. 2009), sorghum (Paterson, Bowers et al. 2009), poplar (Tuskan, Difazio et al. 2006), grapevine (Jaillon, Aury et al. 2007).

II.1.2. Maxam - Gilbert and other chemical methods

Maxam - Gilbert sequencing method is based on a chemical degradation of DNA molecule (Maxam and Gilbert (1977)). In this method, end-labelled DNA fragments are subjected to random cleavage at A, C, G or T positions (typically by a kinase reaction using gamma-^{32}P ATP) (Fig. 19). The chemical attack is based on three steps consisting in base modification, removal of the modified base from its sugar, and DNA strand breaking at that sugar position. The products of these three reactions are then separated using polyacrylamide gel electrophoresis, like the Sanger method. The gel is exposed to X-ray film for autoradiography, yielding a series of dark bands each showing the location of identical radiolabeled DNA molecules (Fig. 19).

The main advantages of the Maxam - Gilbert method compared with Sanger's chain termination reaction method are: a fragment can be sequenced from the original DNA fragment with no need of amplification step, no subcloning and no PCR reactions are required. However, this method requires the handling of toxic chemicals and radioisotopes. Thus Sanger sequencing became the prevailing DNA sequencing method and was widely used for almost 30 years

Table 3. Characteristic summary of DNA sequencing (NGS) technologies

	First generation sequencing	Next generation sequencing								
		The short-read NGS					Long-read NGS			
	Sanger 3730xl [a]	SOLiD [b]	Complete Genomic [c]	Illumina [d]	454 GS [e]	Ion Torrent [f]	Pac Bio [g]	MinION PromethION [h]	Illumina (Moleculo) [i]	10X Genomics [j]
Mechanism	Dideoxy chain termination	Sequencing by ligation	Sequencing by ligation	Sequencing by synthesis, Cyclic reversible termination	Sequencing by synthesis, Single-nucleotide addition	Sequencing by synthesis, Single-nucleotide addition	Single-molecule long-read sequencing	Single-molecule long-read sequencing	Synthetic long-read sequencing	Synthetic long-read sequencing
Read length	400-1000 bp	50 + 50 bp	35 bp	100-250 bp PE	700-1000 bp	~400 bp	10~20 kp	Up to 200 kb	Up to 100 kb	Up to 100 kb
Error rate	0.001%	0.1%	6%	1-2%	0.1%	2%	10-15% SP [k] 1% CCR [l]	12% indels 30%	1-2%	1-2%
Output (per run)	100 KB	100 GB	120 GB	1 TB	1 GB	100 GB	10 GB	1.5 GB 4TB	500 GB	500 GB

[a] https://www.thermofisher.com/order/catalog/product/3730XL, [b] https://www.thermofisher.com/us/en/home/life-science/sequencing/next-generation-sequencing/solid-next-generation-sequencing.html, [c] http://www.completegenomics.com/, [d] http://www.illumina.com/, [e] http://454.com/, [f] https://www.thermofisher.com/us/en/home/brands/ion-torrent.html, [g] http://www.pacb.com/smrt-science/smrt-sequencing/, [h] https://store.nanoporetech.com/, [i] http://www.illumina.com/technology/next-generation-sequencing/long-read-sequencing-technology.html; [j] http://www.10xgenomics.com/. [k] SP: single-pass, [l] CCR: circular consensus sequencing

II.2 The next generation sequencing (NGS) technologies

The advent of NGS almost one decade ago revolutionized genomic researches by bringing the sequencing of entire genome, which allows both DNA sequencing of individual genotypes as well as exploring cytosine methylation in addition to generation of huge transcriptome data. Having the advantages over the classical Sanger sequencing of higher and faster throughput, as well as being cost effective and smaller sample size (Table 3), the application of NGS took over in all fields of human, animal and plant genomic researches (Egan, Schlueter et al. 2012, Michael and Jackson 2013, van Dijk, Auger et al. 2014).

I will present here an overview about NGS technologies, dividing them into those generating short read sequences and those generating long read sequences.

II.2.1 The short-read NGS

Short-read sequencing technologies generate the read length of 35–700 bp. These approaches fall under two broad categories: (i) sequencing by ligation (SBL) and (ii) sequencing by synthesis (SBS) (Goodwin, McPherson et al. 2016).

Short-read NGS require DNA amplification. The DNA template is first fragmented then ligated to a common adaptor set for amplification. In most SBL and SBS approaches, DNA is clonally amplified on a solid surface. This step provides many thousands of identical copies of a DNA fragment and ensures that the signal can be distinguished from background noise. A sequencing platform can collect information from many millions of reaction simultaneously, thus sequencing many millions of DNA molecules in parallel. There are three different strategies used to generate clonal template populations: bead-based, solid-state and DNA nanoball generation (Fig. 20).

- Amplification by **bead-based**: In this amplification method, one adaptor is complementary to an oligonucleotide fragment that is immobilized on a bead (Fig. 20A). DNA template is amplified by emulsion PCR to produce as many as one million clonal DNA fragments (Fig. 20A). This approach is used in 454 (Roche), SOLiD[4] (Thermo Fisher), GeneReader (Qiagen), and Ion Torrent (Thermo Fisher) platforms (Margulies, Egholm et al. 2005, Valouev, Ichikawa et al. 2008, http://www.lifetechnologies.com).

[4] SOLiD: Sequencing by Oligonucleotide Ligation and Detection

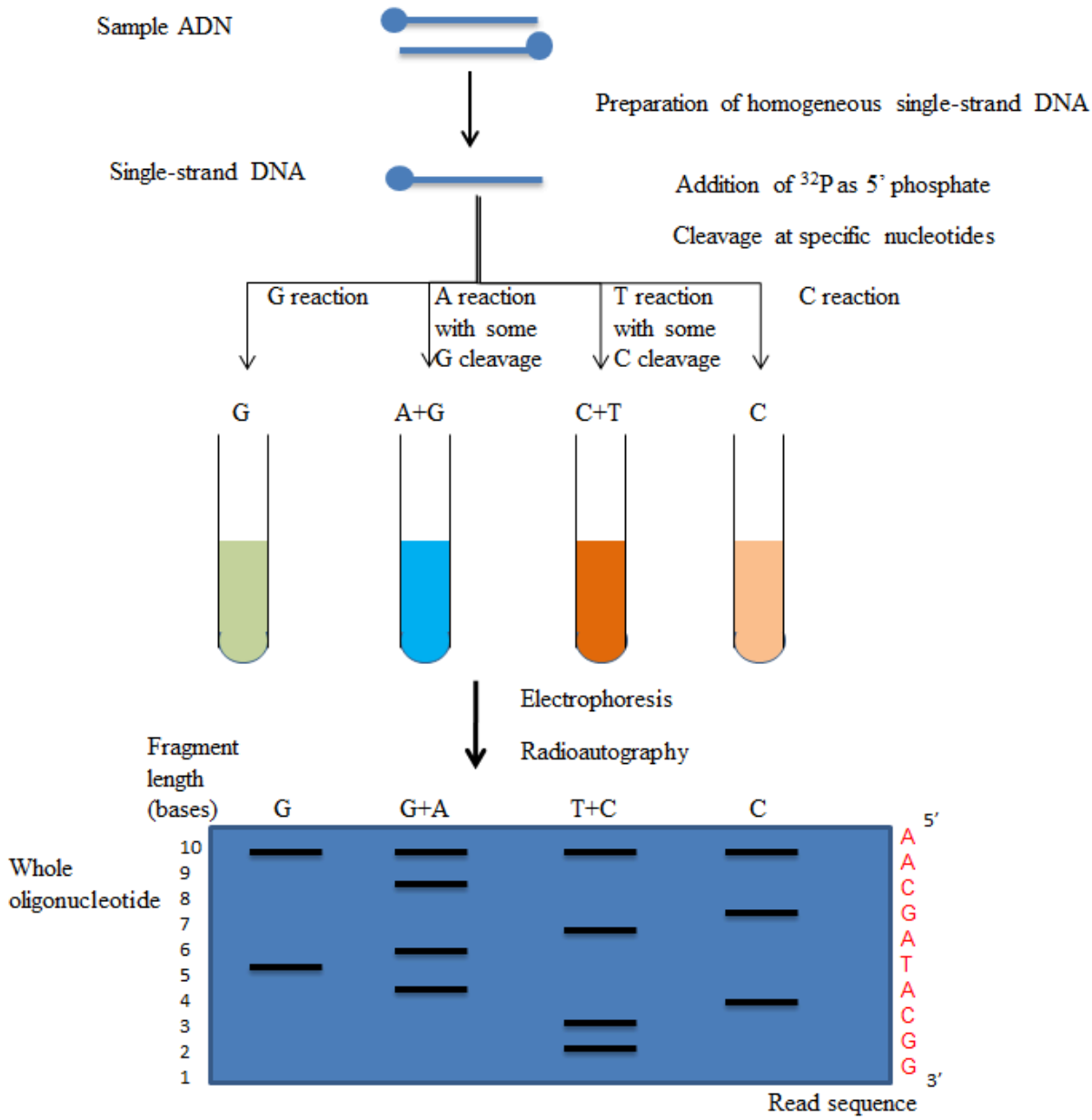


Fig. 19. Maxam-Gilbert sequencing method. Double-stranded DNA is denatured into single-stranded usually by increasing temperature. Radioactively label one 5' end of the DNA fragment to be sequenced by a kinase reaction using gamma-^{32}P, then DNA strand is cleaved at specific positions by using chemical reactions (either dimethyl sulphate selectively attacks purine (A and G) or hydrazine selectively attacks pyrimidines (C and T)). The chemical treatments cleaved at G, A+G, C and C+T. A+G means that it cleaves at A, but occasionally at G as well. Four reaction tubes contain differently sized DNA strands. Fragments are seperated by electrophoresis in high-resolution denaturing acrylamide gels. These gels are placed under X-ray film, which then yields a series of dark bands which show the location of radiolabeled DNA molecules. The fragments are ordered by size and so we can deduce the sequence of the DNA molecule.

(From Matthew and van Holde, 1990).

- **Solid-state** method amplifies directly the DNA template on a slide. It may be either bridge amplification (used in Illumina, Fig. 20B) or template walking (used in SOLiD Wildfire (Thermo Fisher), Fig. 20C). In this approach, forward and reverse primers are covalently bound to the slide surface, either randomly or on a patterned slide. These primers provide complementary ends to which single-stranded DNA (ssDNA) templates can bind (http://www.illumina.com, https://tools.thermofisher.com/content/sfs/brochures/5500-w-series-spec-sheet.pdf).

- **In-solution DNA nanoball generation** is used by Complete Genomics (BGI[5]). DNA undergoes an iterative ligation, circularization and cleavage process to create a circular template, with four distinct adaptor regions. Through the process of rolling circle amplification (RCA), up to 20 billion discrete DNA nanoballs are generated (Fig. 20D) (http://core.iddrc.org/molecular-genetics/wp-content/complete_genomics_technology.pdf).

II.2.1.1 Sequencing by ligation

In SBL, a labelled probe is hybridized and ligated to a DNA strand. The probe encodes one or two known bases (one-base-encoded or two-base-encoded probes[6]) and a series of degenerate or universal bases, driving complementary binding between the probe and template. The anchor fragment encodes a known sequence that is complementary to an adapter sequence and provides a site to initiate ligation. After ligation, the template is imaged and the known base or bases in the probe are identified. A new cycle begins after complete removal of the anchor–probe complex or through cleavage to remove the fluorophore and to regenerate the ligation site (Goodwin, McPherson et al. 2016).

The SOLiD platform (Thermo Fisher) utilizes two-base-encoded probes, in which each fluorometric signal represents a dinucleotide (Valouev, Ichikawa et al. 2008). The sequencing procedure is composed of a series of probe–anchor binding, ligation, imaging and cleavage cycles to elongate the complementary strand (Fig. 21A). The SOLiD produces up to 100 GB per run (Table 3), paired-end sequencing is available. But the read length limited to

[5] BGI: Beijing Genomics Institute
[6] Oligonucleotides that either contains a single interrogation base (one-base-encoded probes) or two adjacent interrogation bases (two-base-encoded probe) in a known position. The base corresponds to a fluorescent label on each probe. The remaining bases are either degenerate (any of the four bases) or universal (unnatural bases with nonspecific hybridization), allowing the probe to interact with many different possible template sequences.

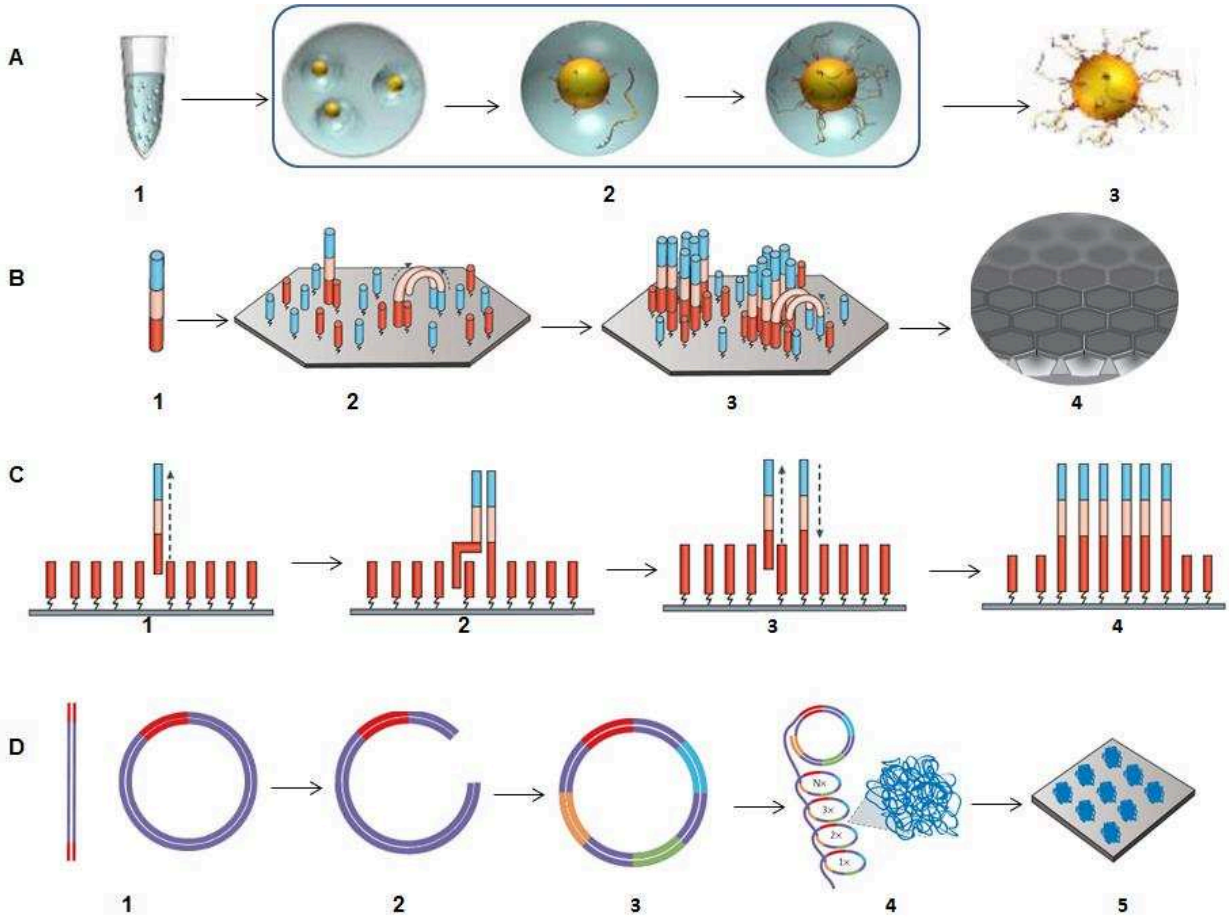


Fig 20. Template amplification strategies. (A) Emulsion PCR is used in 454 (Roche), SOLiD (Thermo Fisher), GeneReader (Qiagen), and Ion Torrent (Thermo Fisher) platforms. (1): In emulsion tube, micelle droplets are loaded with primer, template, dNTPs and polymerase. (2): On-bead amplification: templates hybridize to bead-bound primers and are amplified; after amplification, the complement strand disassociates, leaving bead-bound ssDNA templates. (3): Final product: 100–200 million beads with thousands of bound template. (B) In solid-phase bridge amplification of Illumina, free template (1) hybridize with slide-bound adapters, distal ends of hybridized templates interact with nearby primers to produce bridge amplification (2). After several rounds of amplification, 100-200 million clonal clusters are formed (3). Microwells on flow cell direct cluster generation, increasing cluster density (4). (C) Solid-phase template walking is used in SOLiD Wildfire (Thermo Fisher). (1): Free DNA templates hybridize to bound primers and the second strand is amplified. (2): Primer walking is the stage where dsDNA is partially denatured, allowing the free end to hybridize to a nearby primer. (3): Bound template is amplified to regenerate free DNA templates. (4): After several cycles of amplification, clusters on a patterned flow cell are generated. (D) In-solution DNA nanoball generation is used in Complete Genomics (BGI). (1): Adapter ligation: one set of adapters is ligated to either end of a DNA template, followed by template circularization. (2): Circular DNA templates are cleaved downstream of the adapter sequence. (3): Three additional rounds of ligation, circularization and cleavage generate a circular template with four different adapters. (4): Circular templates are amplified to generated long concatamers, called DNA nanoballs; intermolecular interactions keep the nanoballs cohesive and separate in solution. (5): DNA nanoballs are immobilized on a patterned flow cell.

Adapted from Margulies, Egholm et al. (2005), Goodwin, McPherson et al. (2016), http://www.454.com/, https://tools.thermofisher.com/content/sfs/brochures/5500-w-series-spec-sheet.pdf.

35 bp is one disadvantage of this method (http://symposcium.com/2013/07/advantages-disadvantages-roche454-hiseq-solid-sanger-sequencing/).

Complete Genomics performs DNA sequencing using combinatorial probe–anchor ligation (cPAL) (http://www.completegenomics.com/). In this approach, an anchor and a probe hybridize to a DNA nanoball at several locations (Fig. 20D). In each cycle, the hybridizing probe is a member of a pool of one-base-encoded probes, in which each probe contains a known base in a constant position and a corresponding fluorophore (Fig. 21B). After imaging, the entire probe–anchor complex is removed and a new probe–anchor combination is hybridized. Each subsequent cycle utilizes a probe set with the known base in the n + 1 position. Further cycles in the process also use adaptors of variable lengths and chemistries, allowing sequencing to occur upstream and downstream of the adaptor sequence (Fig. 21B).

II.2.1.2 Sequencing by synthesis

Sequencing by synthesis refers to approaches which use DNA polymerase and a signal (a fluorophore or a change in ionic concentration) to identify the incorporation of a nucleotide into an elongating strand (Goodwin, McPherson et al. 2016). SBS approaches are classified either as (i) cyclic reversible termination (CRT) or (ii) single-nucleotide addition (SNA).

Cyclic reversible termination approaches use of terminator molecules, in which the ribose 3′-OH group is blocked to prevent elongation. DNA template is primed by a sequence that is complementary to an adapter region, which will initiate polymerase binding to this double-stranded DNA (dsDNA) region. During each cycle, a mixture of all four individually labelled and 3′-blocked deoxynucleotides (dNTPs) are added. After the incorporation of a single dNTP to each elongating complementary strand, unbound dNTPs are removed and the surface is imaged to identify which dNTP was incorporated at each cluster. The fluorophore and blocking group can then be removed and a new cycle can begin (Fig. 22).

The Illumina CRT system is the most widely used family of NGS platforms (Fig. 22A) (Bentley, Balasubramanian et al. 2008, Quail, Kozarewa et al. 2008, Quail, Smith et al. 2012, Mardis 2013). The dNTP identification is achieved through total internal reflection

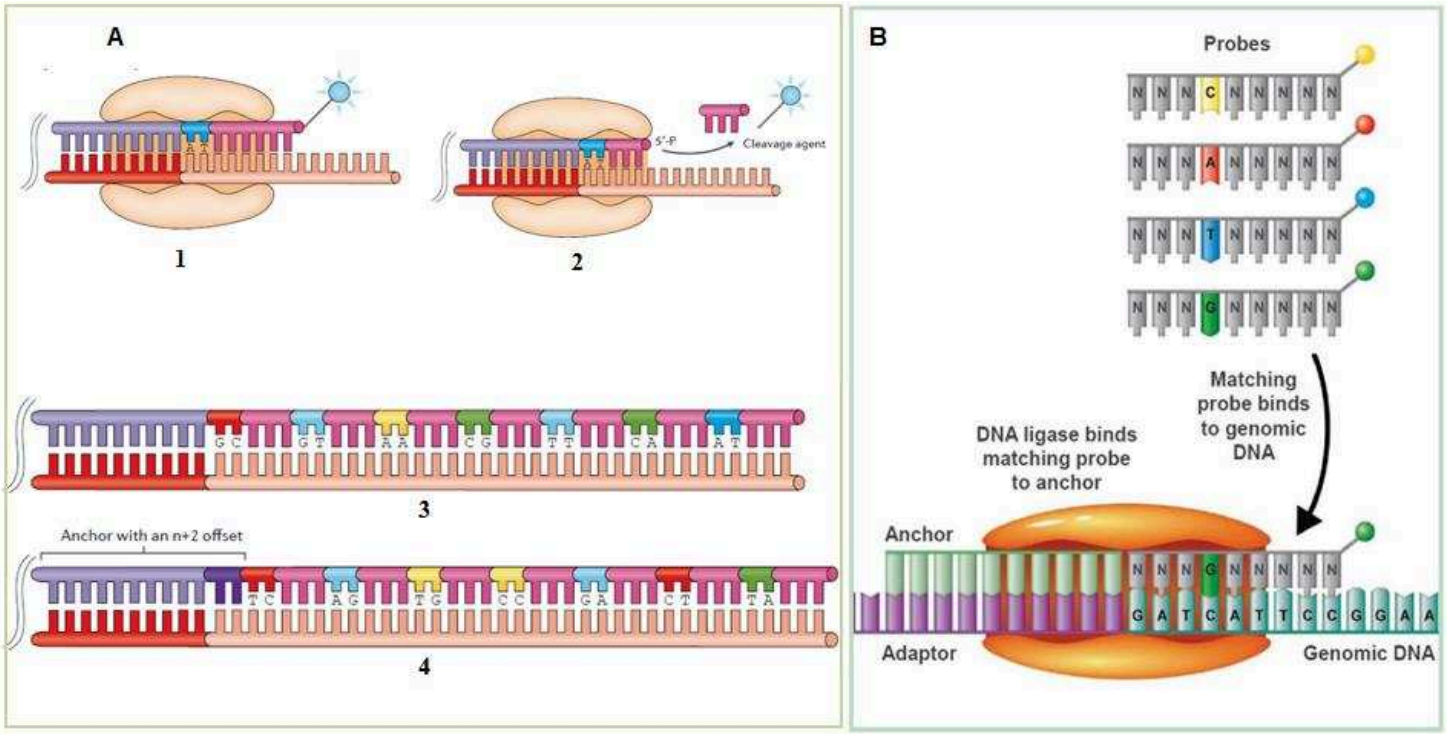


Fig. 21. Sequencing by ligation methods. (A) in SOLiD (Thermo Fisher), SOLiD color-space is interrogation probe , which consists of (3′-to-5′ direction) 2 probe-specific bases followed by 6 degenerate bases (nnnzzz) with one of 4 fluorescent labels linked to the 5′ end. The 2 probe-specific bases consist of one of 16 possible 2-base combinations. (1) Probes with two known bases followed by degenerate or universal bases hybridize to a template; ligase immobilizes the complex and the slide is imaged. (2) Fluorescence is recorded before cleavage of the last 3 degenerate probe bases. (3) Annealing and ligation of the next probe. 10 rounds of hybridization, ligation, imaging and cleavage identify 2 out of every 5 bases. (4) After a round of probe extension, all probes and anchors are removed and the cycle begins again with an offset anchor. (B) Complete Genomics. DNA is sequenced using the combinatorial probe–anchor ligation (cPAL) approach. cPAL uses pools of probes labeled with four distinct dyes (one per base) to read the positions adjacent to each adaptor. Each read position has a separate pool of probes. Complete Genomics‘ proprietary approach allows 10 contiguous bases to be read from each end of an adaptor. Ligating the matching probes with the adjacent anchors dramatically improves the full-match specificity of the probe binding as compared to hybridization without ligation. From (Goodwin, McPherson et al. 2016) and http://www.completegenomics.com/

fluorescence (TIRF) microscopy using either two or four laser channels. In most Illumina platforms, each dNTP is bound to a single fluorophore that is specific to that base type and requires four different imaging channels, whereas the NextSeq and Mini-Seq systems use a two-fluorophore system (http://www.illumina.com). Currently, the HiSeq 4000 is platform with the greatest overall output of Ilumina (650–750 Gb per run). This unsurpassed capacity means more samples can be sequenced simultaneously at a greater depth, generating richer, more important data data (http://www.illumina.com/systems/hiseq-3000-4000/system.html).

The GeneReader (Qiagen) uses virtually the same approach as that used by Illumina (Fig. 22B). However, it does not aim to ensure that each template incorporates a fluorophore-labelled dNTP. Rather, GeneReader aims to ensure that just enough labelled dNTPs are incorporated to achieve identification.

Single-nucleotide addition approaches rely on a single signal to mark the incorporation of a dNTP into an elongating strand. As a consequence, each of the four nucleotides must be added iteratively to a sequencing reaction to ensure only one dNTP is responsible for the signal (Fig. 23).

The first NGS instrument developed was the 454 pyrosequencing[7] device (Margulies, Egholm et al. 2005). This SNA system distributes template-bound beads into a PicoTiterPlate along with beads containing an enzyme cocktail. As a dNTP is incorporated into a strand, an enzymatic cascade occurs, resulting in a bioluminescence signal. Each burst of light, detected by a charge-coupled device[8] (CCD) camera, can be attributed to the incorporation of one or more identical dNTPs at a particular bead (Fig. 23A). The advantage of 454 compared to other NGS techniques is the longer read length which can be set up with de novo assembler in species lacking a genome sequence or extensive transcriptome sequences for comparison (Morozova and Marra 2008)

.

[7] The pyrosequencing approach is a sequencing-by-synthesis technique that measures the release of inorganic pyrophosphate (PPi, the by-product of the reaction when one single nucleotide is added into the a newly synthesized strand by a DNA polymerase) by chemiluminescence (Morozova et al., 2008)

[8] Charge-coupled device: A device composed of an integrated circuit that forms light-sensitive elements: pixels. When a photon interacts with the device, the light generates a charge that can be interpreted by an electronic device (Goodwin et al., 2016).

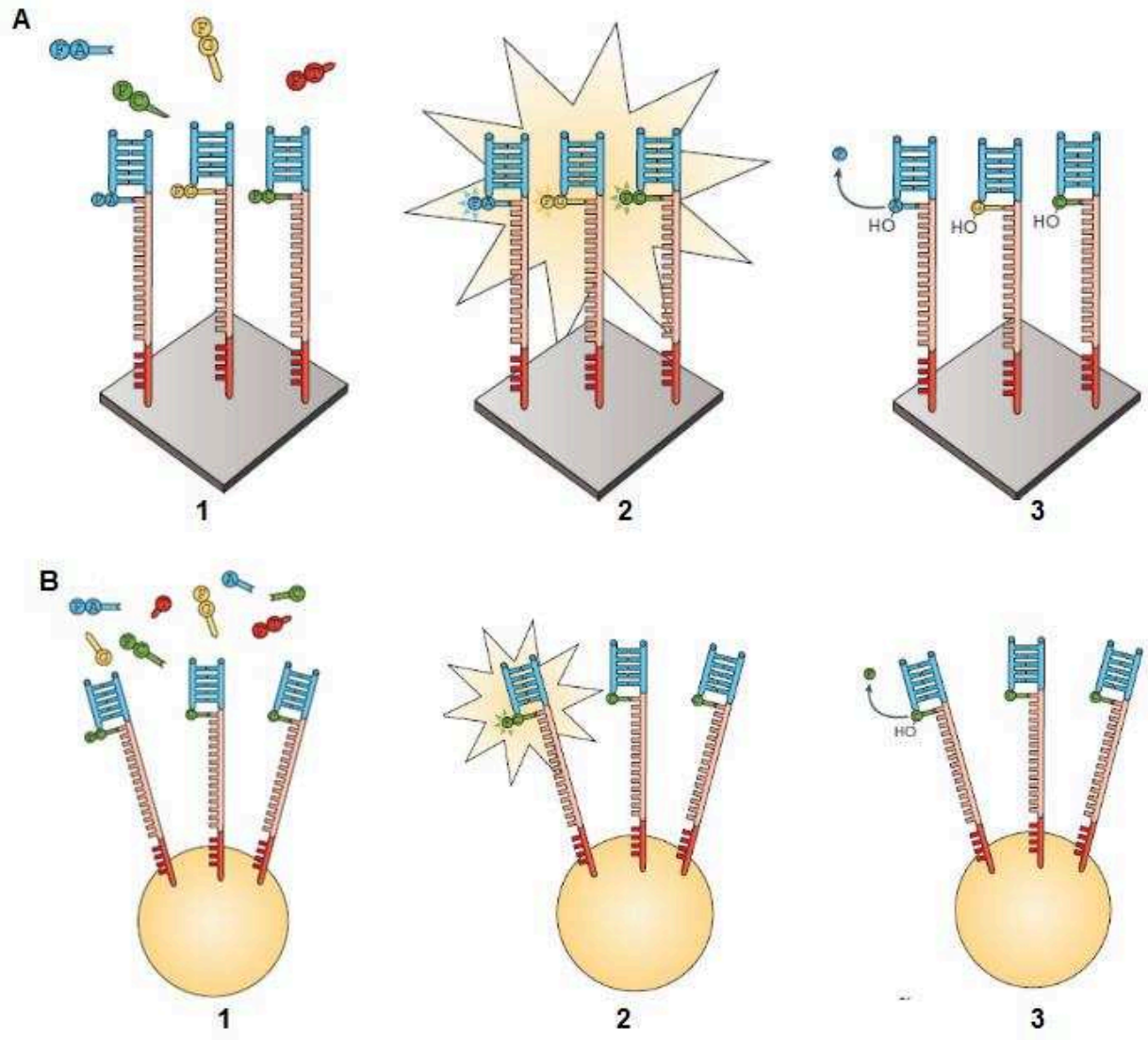


Fig. 22. Sequencing by synthesis: cyclic reversible termination approaches. (A) In Illumina approach, after solid-phase template enrichment, a mixture of primers, DNA polymerase and modified nucleotides are added to the flow cell. (1) During each cycle, fragments in each cluster will incorporate just one nucleotide as the blocked 3′ group prevents additional incorporations. (2) After base incorporation, unincorporated bases are washed away and the slide is imaged by total internal reflection fluorescence (TIRF) microscopy using either two or four laser channels; Each cluster emits a colour corresponding to the base incorporated during this cycle. (3) Fluorophores are cleaved and washed from flow cells and the 3-OH group is regenerated. A new cycle begins with the addition of new nucleotides. (B) GeneReader (Qiagen) sequencing uses bead-based template enrichment. (1) A mixture of fluorophore-labelled, terminally blocked nucleotides and unlabelled, blocked nucleotides hybridize to complementary bases. Each bead on a slide can incorporate a different base. (2) After base incorporation, unincorporated bases are washed away and the slide is imaged by TIRF using four laser channels. Fluorophores are cleaved and washed from flow cells and the 3-OH group is regenerated. A new cycle begins with the addition of new nucleotides.
From Goodwin, McPherson et al. (2016).

The Ion Torrent was the first NGS platform without optical sensing. The Ion Torrent platform detects the H+ ions that are released as each dNTP is incorporated. The change in pH is detected by an integrated complementary metal-oxide-semiconductor[9] and an ion-sensitive field-effect[10] transistor (Fig. 23B). The pH change detected by the sensor is imperfectly proportional to the number of nucleotides detected, allowing for limited accuracy in measuring homopolymer lengths. With the advantage in speed of sequencing, the lower cost, the accuracy data, and smaller instrument size, Ion Torrent sequencing provides one of most powerful NGS sequencer instrument. The "Personal Genome Machine" Ion Torrent can generate 1Gb sequence, 250 bp read length in 2 hours (http://www.lifetechnologies.com).

II.2.2 Long-read NGS sequencing

Long-read sequencing technologies can produce long reads averaging between 5,000 bp to 15,000 bp, with some reads exceeding 100,000bp (Buermans and den Dunnen 2014, Goodwin, McPherson et al. 2016, Lee, Gurtowski et al. 2016). There are two main types of long-read technologies:

(i) Single-molecule real-time sequencing approaches (Pacific Biosciences (PacBio) Single and the Oxford Nanopore Technologies (ONT) sequencing). These technologies differ from short-read approaches in that they do not rely on a clonal population of amplified DNA fragments to generate detectable signal, nor do they require chemical cycling for each dNTP added.

(ii) Synthetic long-read approaches that rely on existing short-read technologies to construct long reads in silico (the Illumina synthetic long-read sequencing platform and the 10X Genomics) (Madoui, Engelen et al. 2015, Goodwin, McPherson et al. 2016). These technologies rely on approach to library preparation that leverages barcodes to allow computational assembly of a larger fragment.

[9] Integrated complementary metal-oxide-semiconductor is an integrated circuit design that is printed on a microchip that contains different types of semiconductor transistors to create a circuit that both uses very little power and is resistant to high levels of electronic noise (Goodwin et al., 2016)

[10] Ion-sensitive field-effect transistor is a type of transistor that is sensitive to changes in ion concentration.

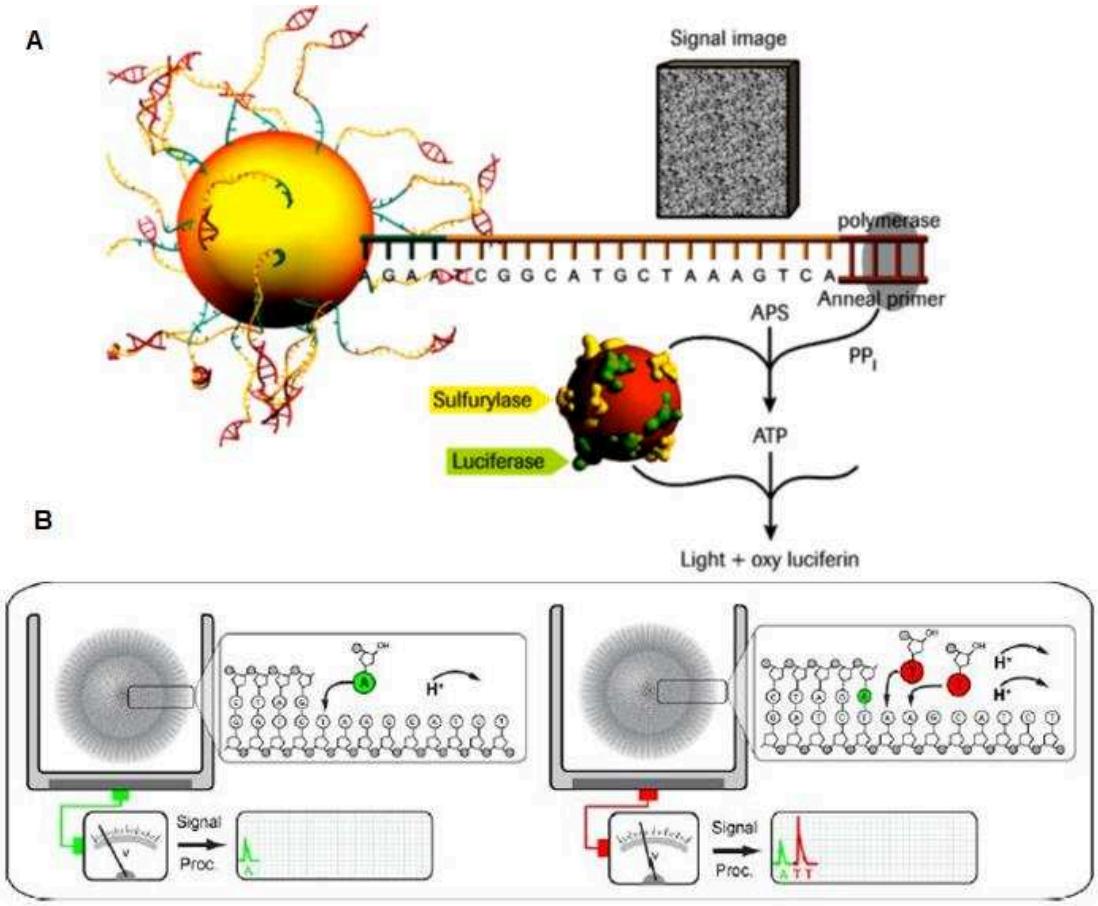


Fig. 23. Sequencing by synthesis: single-nucleotide addition approaches. (A) In 454 pyrosequencing. After bead-based template enrichment, the beads are arrayed onto a microtitre plate along with primers and different beads that contain an enzyme cocktail. During the first cycle, a single nucleotide species is added to the plate and each complementary base is incorporated into a newly synthesized strand by a DNA polymerase. The by-product of this reaction is a pyrophosphate molecule (PPi). The PPi molecule, along with ATP sulfurylase, transforms adenosine 5′ phosphosulfate (APS) into ATP. ATP, in turn, is a cofactor for the conversion of luciferin to oxyluciferin by luciferase, for which the by-product is light. Finally, apyrase is used to degrade any unincorporated bases and the next base is added to the wells. Each burst of light, detected by a charge-coupled device (CCD) camera, can be attributed to the incorporation of one or more bases at a particular bead. (B) In Ion Torrent. After bead-based template enrichment, beads are carefully arrayed into a microtitre plate where one bead occupies a single reaction well. Nucleotide species are added to the wells one at a time and a standard elongation reaction is performed. As each base is incorporated, a single H+ ion is generated as a by-product. The H+ release results in a 0.02 unit change in pH, detected by an integrated complementary metal-oxide semiconductor (CMOS) and an ion-sensitive field-effect transistor (ISFET) device. After the introduction of a single nucleotide species, the unincorporated bases are washed away and the next is added. From Margulies, Egholm et al. (2005), Rothberg, et al., 2011.

II.2.2.1 Single-molecule long-read sequencing

II.2.2.1.1 The PacBio SMRT technology

The SMRT technology is performed on SMRT cells; each patterned with 150,000 Zero-Mode Waveguide[11] (ZMW) providing the world's smallest light detection volume (Fig. 24A). Phosphate-labeled versions of all four nucleotides are present, allowing continuous polymerization of a DNA template. Base incorporation increases the residence time of the nucleotide in the ZMW, resulting in detectable fluorescent signal that is captured by a camera. A camera records the changing colours from all ZMWs; each colour change corresponds to one base (Fig. 24A).

PacBio's RS II C2 XL currently offers both one of the greatest read lengths (averaging around 4,600 bases) and the highest number of reads per run (about 47,000) (Hodkinson and Grice 2015). The disadvantage of this method is the high error rate (10-15%) in single-pass (Table 3). However, this can be overcome by the depth of sequencing since the errors are randomly distributed. Moreover, PacBio company has since incorporated circular consensus sequencing (CCS) into their system, which has greatly reduced error rates by allowing fragments to be sequenced repeatedly and thereby checked for errors (Hodkinson and Grice 2015, Lee, Gurtowski et al. 2016). PacBio is constantly improving this technology. They have recently launched the promising Sequel System which reportedly has a throughput 7× that of the existing RS II System.

The long-read of PacBio SMRT technology production is also applied in genome, transcriptome, and epigenetics research (Rhoads and Au 2015). In 2010, Flusberg, Webster et al. (2010) published a proof-of-concept study of using PacBio to discriminate between methylated and un-methylated bases, as well as between methylated adenine and methylated cytosine. As the polymerase attempts to elongate DNA containing modified bases, it pauses for longer at modified sites compared with unmodified controls, increasing a metric called the interpulse duration and thus indicating the presence of a modified base.

[11] The ZMW is a nanophotonic confinement structure that consists of a circular hole (~70 nm in diameter and ~100 nm in depth) in an aluminum cladding film deposited on a clear silica substrate.

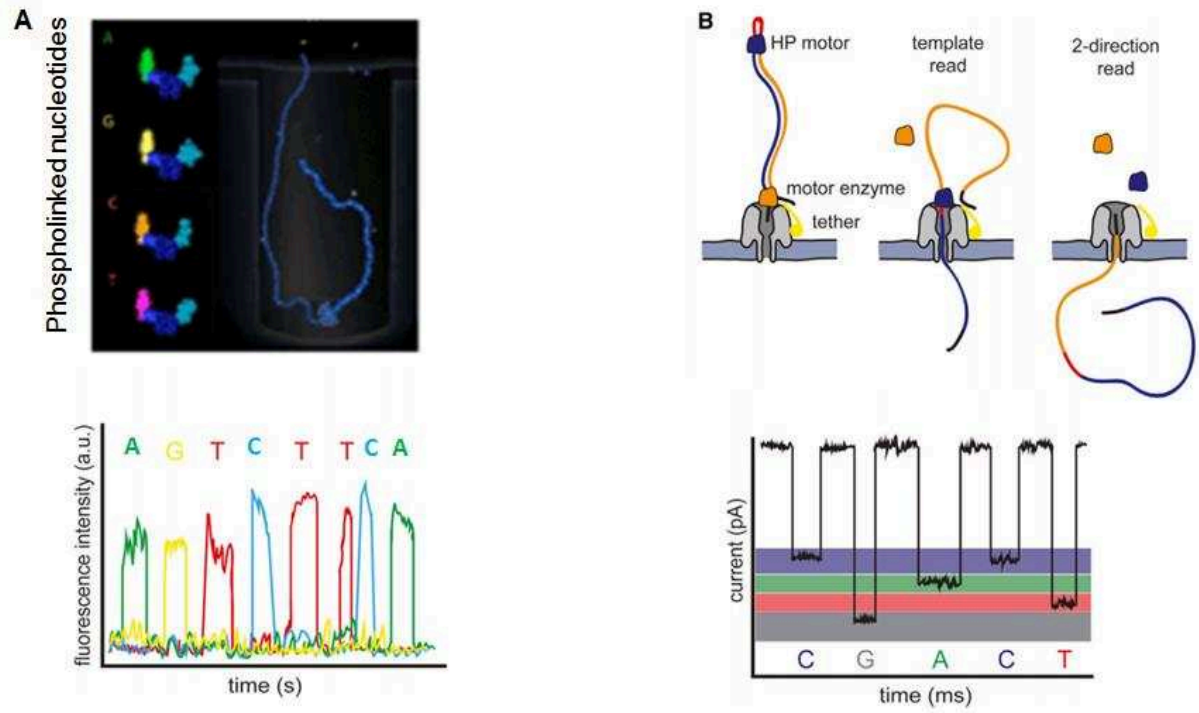


Fig. 24. Single-molecule long-read sequencing. (A) Pacific Bioscience's Single Molecule Real Time (SMRT). A single polymerase is positioned at the bottom of a Zero-Mode Waveguide (ZMW). Phosphate-labeled versions of all four nucleotides are present, allowing continuous polymerization of a DNA template. Base incorporation increases the residence time of the nucleotide in the ZMW, resulting in detectable fluorescent signal that is captured by a camera. A camera records the changing colours from all ZMWs; each colour change corresponds to one base. (B) Oxford Nanopore Sequencing. DNA templates are ligated with two adapters, the first adaptor is bound with a motor enzyme as well as a tether, whereas the second adaptor is a hairprin oligo that is bound by the HP motor protein. Changes in current that are induced as the nucleotides pass through the pore are used to discriminate bases. The library design allows sequencing of both strands of DNA from a single molecular (two-direction reads). From http://www.pacb.com/, https://nanoporetech.com/publications, and http://biochemistri.es/of-nanopores-and-isoforms.

II.2.2.1.2 Oxford Nanopore sequencing

The most recent long-read sequencing technology was released by Oxford Nanopore Technologies in 2014 (Lee, Gurtowski et al. 2016). Unlike other platforms, nanopore sequencers do not monitor incorporations or hybridizations of nucleotides guided by a template DNA strand. Nanopore sequencers directly detect the DNA composition of a native ssDNA molecule that is guided through pores of a nanopore protein membrane. The Oxford Nanopore MinION is a handheld device that sequences DNA by electronically measuring the minute disruptions to electric current as DNA molecules pass through a nanopore (Feng, Zhang et al. 2015). A nanopore is a nano scale hole that can be used to identify the target analyte (DNA, RNA or protein molecules) (Fig. 24B). DNA templates are ligated with two adapters. The first adaptor is bound with a motor enzyme as well as a tether, whereas the second adaptor is a hairprin oligo that is bound by the HP motor protein. Changes in current that are induced as the nucleotides pass through the pore are used to discriminate bases (Fig. 24B). The library design allows sequencing of both strands of DNA from a single molecula (two-direction reads or 2D read). Otherwise only the forward strand sequence is provided (called 1D read) (Madoui, Engelen et al. 2015, https://www.nanoporetech.com).

The primary studies point to a high error rate in reads reaching up to 30% from the current version of MinION that are not randomly distributed and therefore could not be corrected by read depth or coverage (Madoui, Engelen et al. 2015). Methods for correction error rates by combining with short read sequencing technologies, such as Illumina, have been proposed (Madoui, Engelen et al. 2015).

Similar to PacBio nanopore platforms also show promise for the direct detection of modified bases, as the characteristic shift in voltage across the pore is modulated by base modifications, allowing for discrimination without the need for chemical manipulations (Wescoe, Schreiber et al. 2014).

II.2.2.2 Synthetic long-read sequencing

The synthetic long-read technologies rely on a system of barcoding to associate fragments that are sequenced on existing short-read sequencers such as Illumine platforms (Goodwin, McPherson et al. 2016). There are currently two systems available for generating synthetic long-reads: the Illumina synthetic long-read sequencing (former Molecule) platform and the 10X Genomics emulsion-based system.

II.2.2.2.2 The 10x Genomic's

In 10X Genomics' use emulsion to partition DNA , with as little as 1 ng of starting material, the GemCode can partition arbitrarily large DNA fragments, up to ~100 kb, into micelles (also called 'GEMs') along with gel beads containing adapter and barcode sequences (Fig. 25B). The GEMs typically contain ~0.3× copies of the genome and 1 unique barcode out of 750,000. Within each GEM, the gel bead dissolves and smaller fragments of DNA are amplified from the original large fragments, each with a barcode identifying the source GEM. After sequencing, the reads are aligned and linked together to form a series of anchored fragments across a span of ~50 kb. The reads from a single GEM are dispersed across the original DNA fragment and the cumulative coverage is derived from multiple GEMs with dispersed — but linked — reads (Fig. 25B) (Goodwin, McPherson et al. 2016).

The development of NGS technologies is revolutionizing genome research. In the next few years, additional players seek to further democratize the field with novel sequencing solution with direct sequencing of RNA or proteins (Goodwin, McPherson et al. 2016).

II. 3. Management and analysis of NGS data

NGS produces huge amounts of raw data that need to be managed and analyzed. Here I introduce some general processes which were used for analysis of data that I have generated using Illumina technology. Similar or different methods and procedures should be adapted and used with other technologies

II.3.1 Raw data quality-control (QC) and Preprocessing

Quality control for raw reads involves the analysis of sequence quality, GC content, the presence of adaptors, overrepresented k-mers[12] and duplicated reads in order to detect sequencing errors, PCR artifacts or contaminations. Acceptable duplication, k-mer or GC content levels are experiment- and organism-specific, but these values should be homogeneous between samples for a given experiment (Conesa, Madrigal et al. 2016).

[12] k-mer: A substring within a sequence of bases of some (k) length.

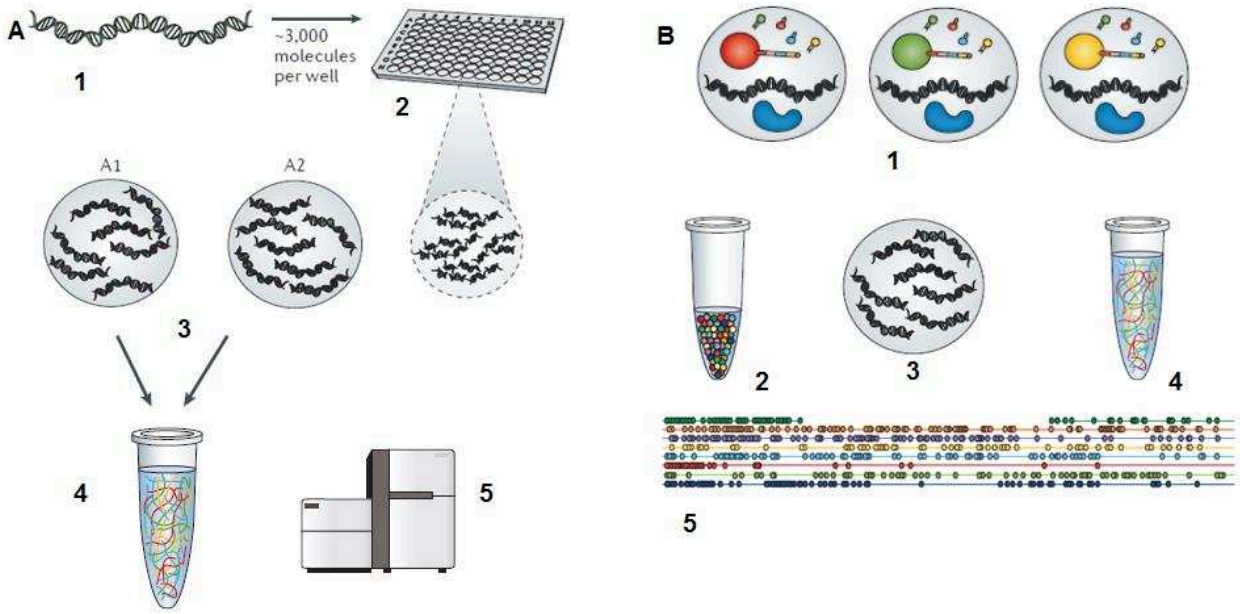


Fig. 25. Synthetic long-read sequencing approaches. (A). In Illumina synthetic long-read sequencing platform, genomic DNA templates are fragmented to 8–10 kb pieces (1). They are then partitioned into a microtitre plate (around 3,000 templates in a single well) (2). Within the plate, each fragment is sheared to around 350 bp and barcoded with a single barcode per well (3). The DNA can then be pooled (4) and sent through standard short-read pipelines (5). (B) In 10X Genomics' emulsion-based sequencing. With as little as 1 ng of starting material, the GemCode can partition arbitrarily large DNA fragments (1), up to ~100 kb, into micelles (also called 'GEMs') along with gel beads containing adapter and barcode sequences (2). The GEMs typically contain ~0.3× copies of the genome and 1 unique barcode out of 750,000. Within each GEM, the gel bead dissolves and smaller fragments of DNA are amplified from the original large fragments (3), each with a barcode identifying the source GEM (4). After sequencing, the reads are aligned and linked together to form a series of anchored fragments across a span of ~50 kb. The reads from a single GEM are dispersed across the original DNA fragment and the cumulative coverage is derived from multiple GEMs with dispersed — but linked — reads (5).

From Goodwin, McPherson et al. (2016).

Sequencing quality scores measure the probability that a base is called incorrectly. With sequencing by synthesis (SBS) technology, each base in a read is assigned a quality score (Q score) by a Phred algorithm, similar to that originally developed for Sanger sequencing experiments.

$$Q = -10\log_{10}P$$

where P is the estimated probability of the base call being wrong.

Higher Q scores indicate a smaller probability of error. Lower Q scores can result in a significant portion of the reads being unusable. They may also lead to increased false-positive variant calls, resulting in inaccurate conclusions. As shown below, a quality score of 20 represents an error rate of 1 in 100, with a corresponding call accuracy of 99%.

FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) is a java application that generates many useful data diagnosis and plots such as Phred score distribution along the reads, GC content distribution, read length distribution, and sequence duplication level (Fig. 26). It also detects over-represented sequences that may be an indication of primer or adaptor contamination. With a comprehensive raw reads QC report generated by FastQC, researchers are able to determine whether any preprocessing steps such as base trimming, read filtering, or adaptor clipping are necessary prior to alignment (Bao, Huang et al. 2014). We used this application for trimming and filtering our RNA-Seq data (see below).

II. 3.2 Sequence Alignments

Alignment is the process of mapping short nucleotide reads to a reference genome. This step is computationally intense and time consuming. It is also a critical step, as any errors in alignment to the reference genome will be carried further in the analysis (Dolled-Filhart, Lee et al. 2013).

The Sequence Alignment/Map (SAM) and Binary Alignment/Map (BAM) formats are the standard file formats for storing NGS read alignments (Li, Handsaker et al. 2009). Short reads generated from NGS may either be single-end reads (the sequencer reads a fragment from only one end to the other, generating the sequence of base pairs) or paired-end reads (the sequencer starts at one read, finishes this direction at the specified read length, and then

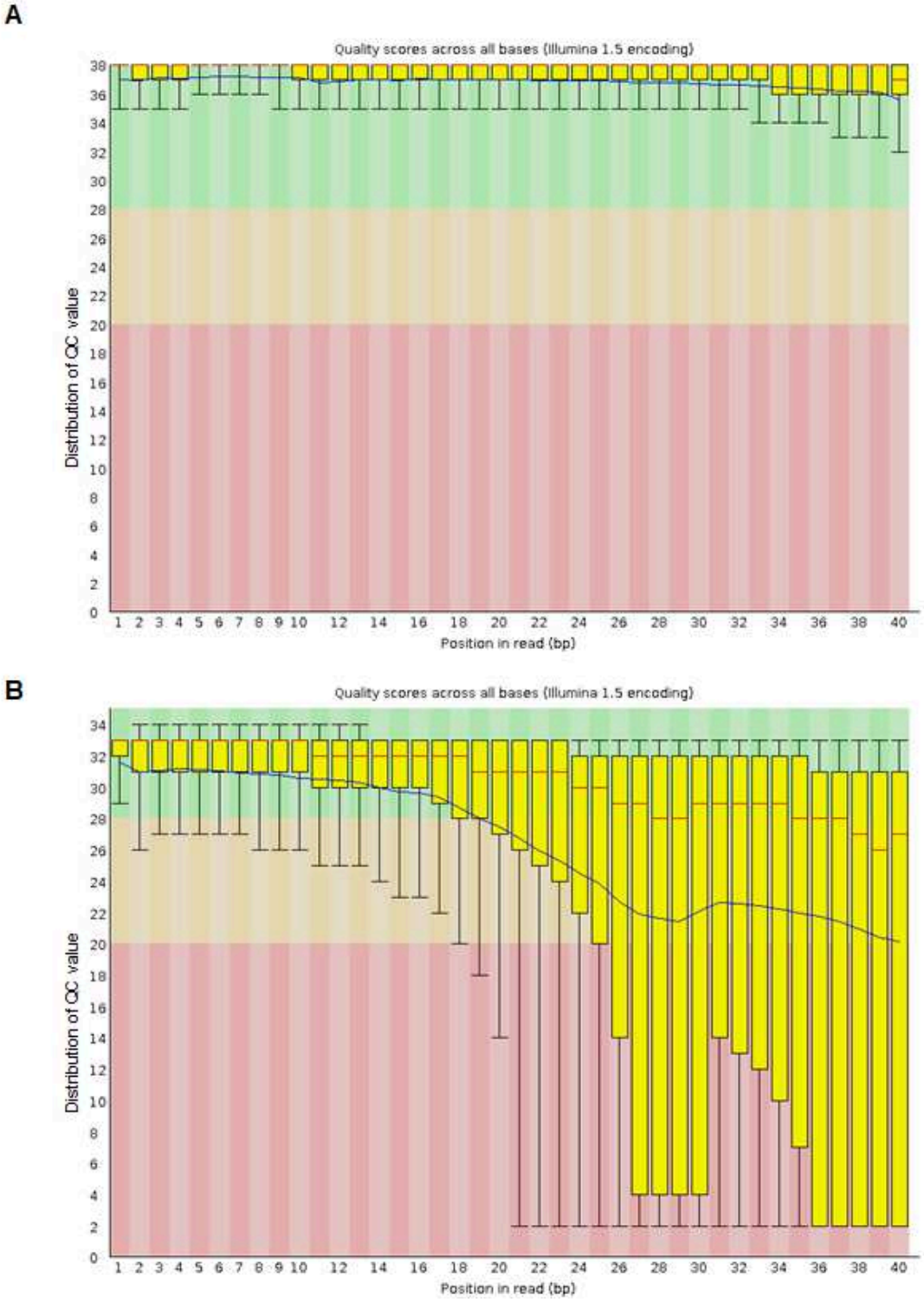


Fig. 26. Example reports of Illumina data per base sequence quality by FastQC. (A) good quality data with a homogenous of distribution of QC value, while there are the variabilities in the bad quality data (B).

From http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

starts another round of reading from the opposite end of the fragment). Their length may range from dozens to hundreds of base pairs and these reads need to be aligned correctly to their appropriate location within the reference genome.

Different tools have been developed for short reads mapping, including Bowtie/Bowtie2 (Langmead, Trapnell et al. 2009, Langmead and Salzberg 2012), BWA (Burrows-Wheeler Aligner) (Li and Durbin 2010), mrFAST (Alkan, Kidd et al. 2009), Novoalign (http://novocraft.com), SOAP (Li, Yu et al. 2009) and SSAHA2 (Ning, Cox et al. 2001). Among them, Bowtie and BWA are two well-known short reads alignment tools that implement Burrows–Wheeler Transformation (BWT) algorithm. This algorithm is tolerant to mismatches and gaps at the expense of increased computational time.

BWA is used for mapping low-divergent sequences against a large reference genome and consists of three algorithms: BWA-backtrack, BWA-SW and BWA-MEM. The first algorithm is designed for Illumina sequence reads up to 100 bp, while the remaining algorithms are for longer sequences ranged from 70 bp to 1 Mbp (Li and Durbin 2010). In the present study, we used BWA (Version: 0.6.1-r104) to map RNA-Seq data into reference genome (see below).

II. 3.3 Normalization

Normalization is a process designed to identify and minimize impacts of technical differences between samples. Massive and complex RNA-Seq datasets contain systematic variations which affect the analysis of gene expression and should be eliminated from data prior to performing statistical analysis (Dillies, Rau et al. 2012, Risso, Ngai et al. 2014, Zyprych-Walczak, Szabelska et al. 2015). Thus normalization process is a crucial step in RNA-Seq data analysis which has strong effects on the identification of differentially expressed genes (DEGs) (Soneson and Delorenzi 2013). There are some sources of variabilities in RNA-Seq data, between-samples such as library size (sequencing depth) or within-samples such as gene length GC content or unwanted variation introduced by batch effect (Zyprych-Walczak, Szabelska et al. 2015). A number of normalization methods has been developed since the emergence of RNA-Seq technology which can be classified into two groups: (i) between-lane compares a gene in different samples by scaling the data on the libraries sizes or using housekeeping genes, (ii) and the within-lane normalization compares genes in a sample by normalizing gene lengths (Dillies, Rau et al. 2012).

The choice of normalization method plays an important role in RNA-Seq analysis. In this study, I used the standard RPKM (Reads Per Kilobase per Million mapped read) normalization in R program to normalize RNA-Seq data (see below). This method facilitates comparisons between genes within a sample and combines between- and within-sample normalization. It re-scales gene counts to correct for differences due to both library sizes and gene length (Mortazavi, Williams et al. 2008, Dillies, Rau et al. 2012). Recently, a comparison of normalization methods indicated that RPKM showed the highest correlation results for RNA-Seq of a 35-nucleotide sequence (Li, Piao et al. 2015). RPKM has shown that read coverage along expressed transcripts can be non-uniform because of sequence content or RNA preparation method. RPKM method is a popular choice in many practical applications (Oshlack, Robinson et al. 2010, Dillies, Rau et al. 2012). The RPKM is calculated according to the following formula:

$$F_{ij} = \frac{N_j}{10^6} \times \frac{L_i}{10^3}$$

where

F_{ij} is the total number of mapped read in sample j (million mapped read) by the length of gene i (in kilobase),

N_j is number of mapped read in sample j,

L_i is length of gene (or the exons of gene) i.

III. INTEGRATIVE CHARACTERIZATION OF GENETIC, TRANSCRIPTOMIC AND EPIGENETIC CHANGES RELATED TO AUTO- AND ALLOPOLYPLOIDY, USING NGS

The resynthesized Brachypodium autopolyploids and allopolyploids (Dinh-Thi, Clainche et al. 2016, Dinh-Thi, Coriton et al. 2016), together with natural B. hybridum allopolyploid provide a fantastic material for integrative characterization of genetic, epigenetic and functional changes in polyploid genomes both at short and long terms of polyploid evolution. To reach this objective, I have mainly used the Illumina (Bentley, Balasubramanian et al. 2008) NGS methodology that is well developed and routinely used at the Evry Genome Institute for generation of high throughput sequence data. I have generated important data consisting of: (i) DNA resequencing, (ii) transcriptome through RNA-Seq and (iii) CpG methylation through bisulfite sequencing. The experiments were done on B. hybridum and its two parental species as well as synthetic autopolyploids and allopolyploids (Fig. 27) (Dinh-Thi, Clainche et al. 2016, Dinh-Thi, Coriton et al. 2016). Overview for the materials and tissues used in the integrative analysis is illustrated in Fig. 27 and detailed in Table 4. A total of 13 genotypes consisting in autopolyploids (derived from B. distachyon (Bd21 and Bd3-1) and from B. stacei (ABR114 and Bsta5)), synthetic allopolyploids (allo21×114, F1_21×114 and allo3-1×5), the natural diploid species B. distachyon, B. stacei and the natural allopolyploid B. hybridum (lines ABR113 and Bhy30) were chosen (Fig. 27 and Table 4).

A wide range of comparative analyses could be performed allowing mainly the elucidation: (i) the effect of homologous genome doubling when comparing diploids and autopolyploids of the same lines; (ii) the effect of homoeologous genome doubling by comparing allopolyploids to their parental species genomes (Fig. 27).

I will present hereafter the material and methods were used, the progress in the different experiments followed by planed analyses.

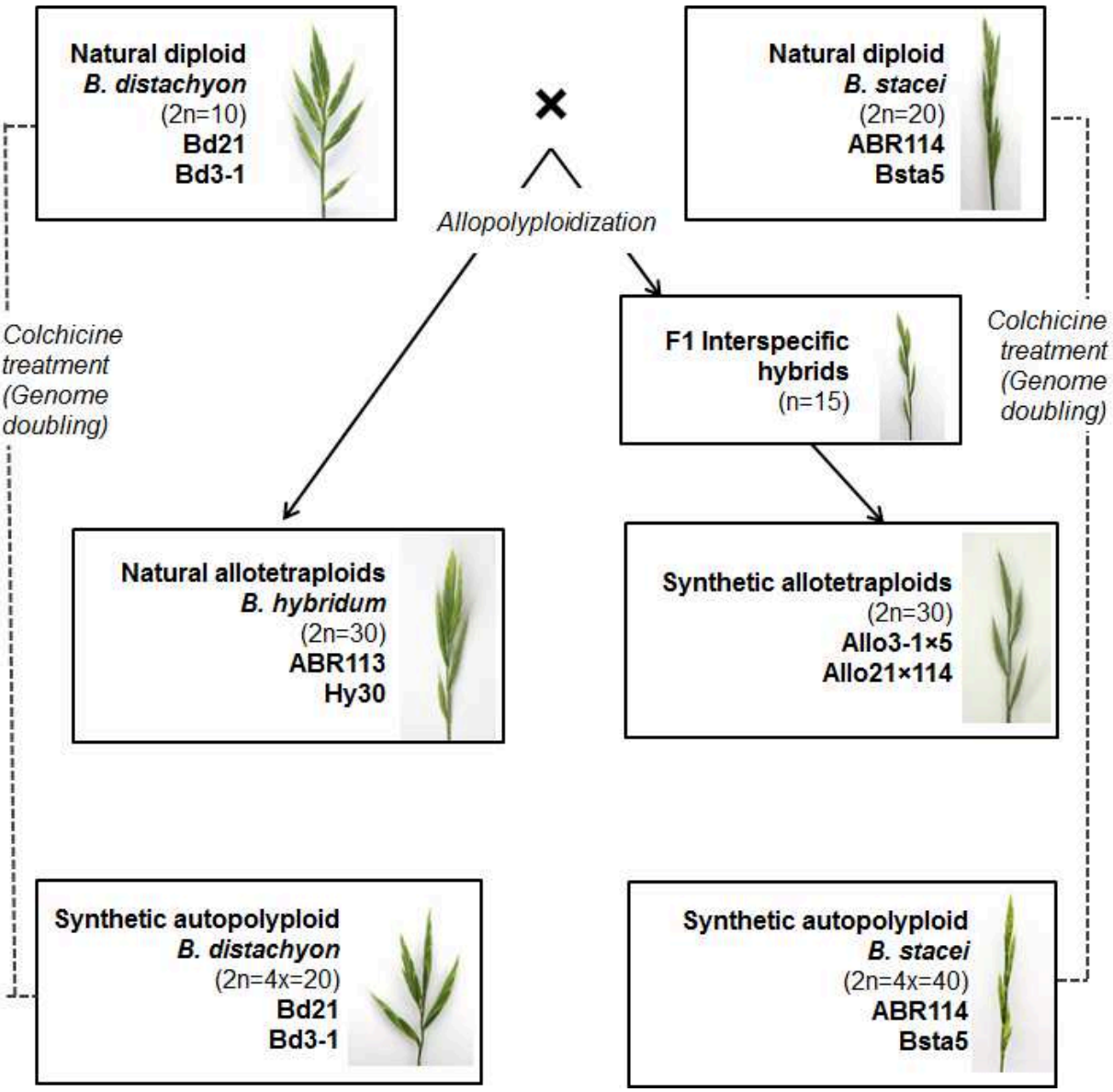


Fig. 27. Overview of natural and synthetic Brachypodium species and accessions used for DNA Illumina re-sequencing, bisulfite-treated DNA sequencing and mRNA-sequencing. Names of the lines or accessions of each species and genotypes are indicated. The photos present spike of Bd21 line and ABR114 line for diploid and autopolyploid of B. distachyon and B. stacei, spike of ABR113 line for B. hybridum and Allo21×114 for synthetic allopolyploid. For more information about the synthetic auto- and allopolyploid, see (Dinh-Thi, Clainche et al. 2016, Dinh-Thi, Coriton et al. 2016) More details on the material and samples are in Table 4

.

III.1. Revealing structural an cytosine methylation changes through DNA and DNA bisulfite-treated sequencing

High molecular weight (HMW) DNA was extracted from the same plants of the chosen 13 genotypes (Fig. 27) were used for:

- Direct resequencing in order to reveal structural and genetic changes.
- Revealing cytosine methylation where DNA was treated with bisulfite to convert methylated Cytosine into Uracil and then re-sequenced. Each genotype was replicated three times to allow quantitative evaluations of Cytosine methylation.

III.1.1 Material, methods and progress in DNA and DNA bisulfite-treated sequencing

High molecular weight (HMW) DNA extraction

High molecular weight (HMW) DNA was prepared after extraction of nuclei. Nuclei were isolated from young fresh leaves. Briefly, two gram of leaf samples was grinded in liquid nitrogen into a fine powder. 8ml of SEB1X+ β-mercaptoethanol (SEB buffer (0.01 M Tris base, 0.1 M KCl, 0.01 M EDTA, pH 9.5, 500 mM sucrose, 4 mM spermidine, 1 mM spermine tetrahydrochloride, 0.1% w/v ascorbic acid, 2% w/v PVP (MW 40,000), and 0.13% w/v sodium diethyldithiocarbamate) was added to the leaf powder. After incubation on ice for 20 minutes, the mixtures were filtered by Miracloth twice. After addition of 500µl of SEB1X-Triton, the nuclei were received by centrifugation at 600g and 8°C for15min. When the nuclei were pelleted, we proceeded to HMW DNA extraction as follows. The nuclear pellet was gently resuspended in SEB buffer. 1.5ml of lysis buffer (TRIS 1M (100mM), NaCl 5M (100mM), EDTA 0.5M (50mM), SDS 10% (2%)) was added together with 15 µl of Proteinase K and incubated for one hour at 65°C. Two ml of phenol/chloroform (24-1) was then added to each tube, mixed gently by inversion and centrifuged at 2000 rpm for 20 min in a bench top centrifuge to separate the phases. The phenol/chloroform extraction was repeated twice. HMW DNA was then precipitated by adding 0.6 volumes of isopropanol, centrifuged for 20 min at 2000 rpm and the DNA pellet rinsed 2-3x with 70% ethanol and air dried briefly on a hood. The DNA pellet was dissolved in 140 µl of TE buffer, 7 µl of RNAse stock

Table 4 The natural accessions of Brachypodium (diploid and allopolyploid) and the synthetic polyploid (autopolyploid and allopolyploid) have been used for the sequencing projects.

Species	**Sample ID**	**Genome**	**DNA Illumina re-sequencing** (replicate)	**Genome-wide bisulfite sequencing** (replicate)	**mRNA-seq sequencing** (replicate)
B. distachyon	Bd21	2n = 2x =10	1	3	6[a]
	AutoBd21	2n = 4x =20	1	3	3
	Bd3-1	2n = 2x =10	1	3	3
	AutoBd3-1	2n = 4x =20	1	3	3
B. stacei	ABR114	2n = 2x =20	1	3	6[a]
	AutoABR114	2n = 4x =40	1	3	9[b]
	Bsta5	2n = 2x =20	1	3	3
	AutoBsta5	2n = 4x =40	1	3	3
Interspecific hybrid	F1_21×114	2n = 15	1	3	3
Synthetic allopolyploid	Allo21×114	2n = 4x =30	1	3	3
	Allo3-1×5	2n = 4x =30	1	3	3
MPV[c]	Bd21+ABR114		-*	-	6[a]
	Bd3-1+Bsta5		-	-	3
B. hybridum	ABR113	2n = 4x =30	1	3	6[a]
	Bhy30	2n = 4x =30	1	3	-

* Not analysis
[a] Spike and leaf tissue have been done
[b] Three divergent phenotypes of autoABR114 have been done
[c] Mixtures of equal amounts of parental RNAs

was added and the DNA was incubated at 37°C for 1 h. The DNA was then extracted with phenol/chloroform twice, as described above. The DNA was precipitated with 0.3 M sodium acetate, and 2 volumes of cold 95% ethanol. After centrifugation the DNA pellet was rinsed with 70% ethanol, The DNA was dissolved in 100µl of 1X TE.

DNA-resequencing

DNA libraries were prepared manually following the manufacturer's protocol of NEBNext DNA Sample Prep Master Mix Set 1 (Ozyme). Briefly, samples were normalized to 5 µg DNA. 10 µl of NEBNext End Repair Reaction Buffer 10X(50mM Tris-HCl, 10mM MgCl2, 10mM DTT, 1mM ATP, 0.4mM dATP, 0.4mM dCTP, 0.4mM dGTP, 0.4mM dTTP, pH 7.5 at 25°C) and 5 µl of NEBNext End Repair Enzyme Mix (10,000 units/ml T4 PNK and 3,000 units/ml T4 DNA Polymerase) were added to the DNA solution. Then the DNAs were cleanup by AMPure XP® Beads. The dA-Tailing of end-repaired DNAs were prepared by adding 5 µl of NEBNext dA-Tailing Reaction Buffer 10X (10 mM Tris-HCl, 10 mM MgCl2, 50 mM NaCl, 1 mM DTT, 0.2 mM dATP, pH 7.9 at 25°C) and 3 µl of Klenow Fragment (3′→5′ exo–) which contains an E. coli strain containing a plasmid with a fragment of the E. coli polA (D355A, E357A) gene starting at codon 324 supplied in 25 mM Tris-HCI (pH 7.4), 0.1 mM EDTA, 1 mM DTT and 50% glycerol. After the cleanup by AMPure XP® Beads, the adaptor ligation of dA-Tailed DNA were prepared by adding 10 µl of Quick Ligation Reaction Buffer 5X (66 mM Tris-HCl, 10 mM MgCl2, 1 mM dithiothreitol, 1 mM ATP, 6% Polyethylene glycol (PEG 6000), pH 7.6 at 25°C), 10 µl of NEBNext Adaptor 15 µM (Various adaptors can be used for Illumina which can be NEB #E7335, #E7350, #E7500, #E7710, #E7730, #E7600, #E6609) and 5 µl of Quick T4 DNA Ligase (Purified from E. coli C600 pcl857 pPLc28 lig8). Then the PCR use 2-5 cycles of amplification to enrich the adaptor ligated DNA with the primers depend on the chosen adaptors. Finally, the DNA libraries quality and quantity was estimated on a BioAnalyzer using DNA1000 Chips.

The DNA libraries were sequenced by Illumina Solexa Hiseq 2000. One single replicate per each genotype was used here and DNA resequencing was performed at a depth range from 19 to 43 genome equivalents (Table 5).

Table 5 Illumina DNA resequencing results on the natural Brachypodium accessions (diploid and allopolyploid) as well as the synthetic polyploids generated in this thesis

Species	Genotype and chromosome number	Read number	Nucleotide number	Sequencing coverage (X)[a]
B. distachyon	Diploid Bd21 (2n=10)	47,897,955	9,675,386,910	35.57[b]
	Autopolyploid Bd21 (2n=20)	42,082,674	8,500,700,148	31.25
	Diploid Bd3-1	49,120,307	9,922,302,014	36.48
	Autopolyploid Bd3-1	58,439,323	11,804,743,246	43.40
B. stacei	Diploid ABR114 (2n=20)	46,376,151	9,367,982,502	40.03[c]
	Autopolyploid ABR114 (2n=40)	45,085,868	9,107,345,336	38.92
	Diploid Bsta5 (2n=20)	48,014,220	9,698,872,440	41.45
	Autopolyploid Bsta5 (2n=40)	44,447,548	8,978,404,696	38.37
Interspecific hybrid	F1_21×114 (n+n'=15)	68,341,825	13,805,048,650	25.56[d]
Synthetic allopolyploid	S0_21×114 (2n = 30)	63,384,034	12,803,574,868	23.71
	S0_3-1×5 (2n = 30)	61,131,787	12,348,620,974	22.87
B. hybridum	ABR113 (B. stacei as mother) (2n =30)	65,869,493	13,305,637,586	24.64
	Bhy30 (B. distachyon as mother) (2n =30)	51,651,879	10,433,679,558	19.32

[a] Sequencing coverage was calculated by formula: LN/G, where L is the read length, N is the number of reads and G is the reference genome size (Sims et al., 2014).

[b] B. distachyon genome is 272 Mb (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Bdistachyon_er)

[c] B. stacei genome estimates ~234 Mb (Gordon, Liu et al. 2016)

[d] B. hybridum genome estimate ~540 Mb (Gordon, Liu et al. 2016)

DNA-bisulfite treated sequencing

The high-throughput Bisulfite sequencing method is one of the appropriate methods for measuring cytosine methylation on a genome wide scale (Clark, Statham et al. 2006, Li and Tollefsbol 2011). In brief, bisulfite treatment of gDNA converts Cytosine to Uracil and has no effect on methylated Cytosine. After bisulfite treatment, unmethylated Cytosine is read as Thymine, while the methylated Cytosine is read as Cytosine. Specific sites of methylated Cytosine are detected by aligning the sequenced reads to a reference genome, and proportions of methylation can be estimated for each Cytosine (Clark, Statham et al. 2006).

In this study, genome-wide bisulfite deep sequencing was performed by an in-house protocol compatible with Illimuna's TruSeq chemistry as previous described by Chalhoub, Denoeud et al. (2014). All 13 genotypes were done in three replicates.

Briefly, three µg of HMW DNAs were fragmented with a Covaris E210 instrument and purified by Agencourt AMPure XP beads (Beckman Coulter, Pasadena, CA) to obtain DNA fragment of at least 250 bp. The purification was controlled by BioAnalyzer using High Sensitivity DNA Chips (Agilent Technologies, Santa Clara, CA). One µg of fragmented DNA was ligated to adapters compatible with the Illumina technology using an in-house developed protocol consisting in: 1) end repair, 2) 3' adenylation and 3) ligation of adaptors, followed by a purification step using Agencourt AMPure XP beads. Ligated samples were then bisulfite treated to convert methylated Cytosine into Uracil, using the EZ DNA Methylation Kit (Zymo Research, Irvine, CA) following the manufacturer's instructions. After conversion of methylated-Cytosine into Uracil, the DNA was amplified by PCR (12 cycles) using HiFi HotStart Uracil polymerase (KAPA Biosytems, Wilmington, DL). Samples were purified using a QIAquick PCR Purification Kit (Qiagen Venlo, Netherlands) and quality controlled on a BioAnalyzer using DNA1000 Chips. The libraries were size selected on an agarose gel (300 bp to 400 bp) and purified using a QIAquick Gel Extraction kit (Qiagen). The quality of DNA libraries was controlled again on a BioAnalyzer using DNA1000 Chips before 101 bp paired-end sequencing on a HiSeq2000 (Illumina, San Diego, CA).

Revealing cytosine methylation

The inference of methylated Cytosines, converted by bisulfite-treatment into Uracil, will be also revealed by mapping and comparison of the DNA bisulfite-treated resequencing data with the reference genomes of B. distachyon, B. stacei and B. hybridum, as well as on concatenated genome sequences of B. distachyon and B. stacei as described by (Fang, Martin et al. 2012, Krueger, Kreck et al. 2012, Chalhoub, Denoeud et al. 2014).

The 39 data sets of bisulfite-treated DNA from three replicates of 13 genotype samples will be analyzed following the steps bellow: FastQC v0.10.1 is executed on all read files in order to assess basic quality control metrics (base quality distribution, GC content, relative abundance of each base at each read position). Each paired-end sequence file is trimmed using an in-house Perl script that set a minimum quality threshold of Q score of 30 (Q30) and retain only reads with more than 90% of bases left after trimming. Three FastQ files are generated from this trimming: two files for reads that remained paired after trimming and one file for unpaired reads. These two new sets of sequences (one paired-end and one single end) will be aligned using Bismark v0.9.0 (Krueger and Andrews 2011) with the Bowtie2 option and with one mismatch allowed in a seed alignment. The 2 SAM files resulting from this alignment will be merged and coordinate-sorted to create a single SAM file, which is then fed into methylKit v0.5.7 (Akalin, Kormaksson et al. 2012).

Methylation calls will be generated using methylKit. Bases with too low (< 10x) or too high coverage (bases that had more than the 99.9th percentile of coverage in each sample) will be discarded. The coverage of the 39 data sets will be subsequently normalized using the median as a scaling factor across these data sets. Differentially methylated regions (DMRs) are calculated conservatively using a minimum q-value of 0.01 and methylation difference of 25%.

These planed analyses could not be done when I wrote my PhD thesis. From one side the DNA sequencing and DNA bisulfite-treated sequencing were not finished yet and from the other side we were waiting for the accomplishment of genome sequence assemblies of B. stacei and B. hybridum to map all generated DNA sequences and infer genomic as well as CpG methylation changes.

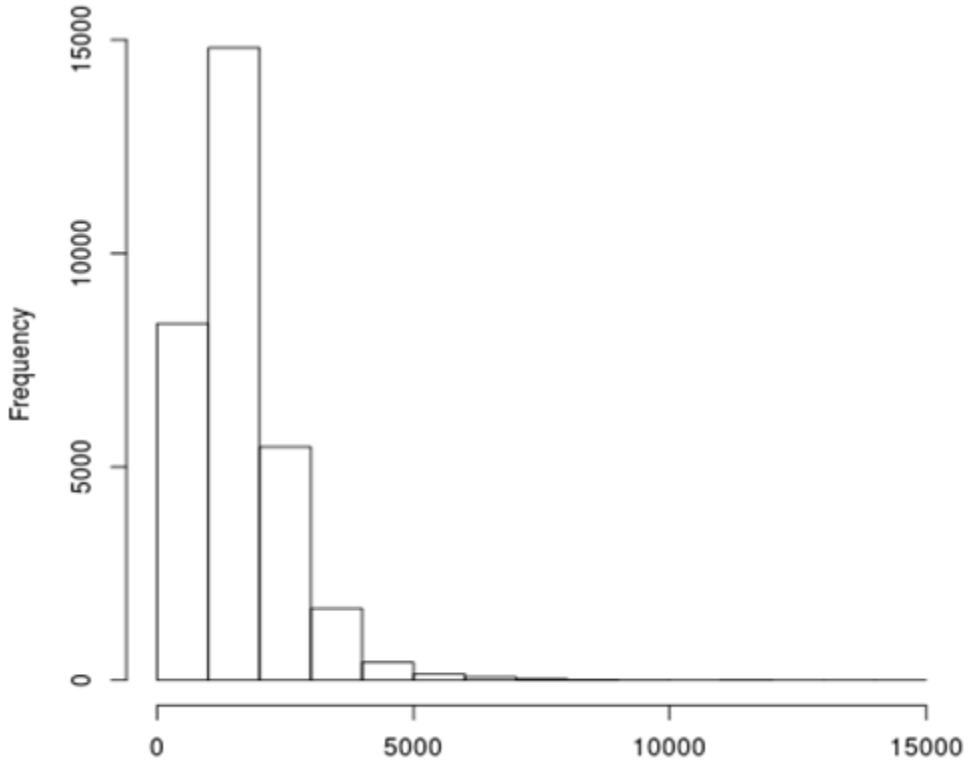


Fig. 28. Length distribution of 31,029 predicted genes of B. distachyon ranging from 34 bp to 14,968 bp with an average mean length of 1,604 bp and median of 1,429 bp.

III.2. Evaluation of changes in gene expression using RNA sequencing

RNA-Sequencing (RNA-Seq) data were generated from spikes sand leaves of the same 13 genotypes that have been used for DNA and DNA-methyl-resequencing. These include the three natural species B. distachyon, B. stacei, B. hybridum, F1 interspecific hybrid (F1_21×114), synthetic allopolyploid (allo21×114 and allo3-1×5) and autopolyploids (autoBd21, autoBd3-1, autoABR114 and autoBsta5) (Fig. 27, Table 4).

III.2.1 Material and method for analysis of gene expression using RNA sequencing

Plant material and growth condition

Plant growth condition was as described by (Dinh-Thi, Clainche et al. 2016, Dinh-Thi, Coriton et al. 2016). All genotypes were grown in a randomized design, with three biological replicates each, in a growth chamber at 22°C during day, 18°C during night and 16 h day-length. Leaf samples were collected when plants were at five-leaf stage on the main tiller and whole spikes were collected at three days after spike emergences.

RNA extraction, mRNA-Seq library construction and sequencing

Total RNA was extracted by Trizol reagents following the protocol of Invitrogen with additional washing steps to obtain salt-free RNA. After isopropanol precipitation, RNA pellets were dissolved in 500 µl RNase-free water. To further wash the RNA pellets, 1/10 volume of sodium acetate 3 M pH 5.2 and 2 volumes of ethanol were added. The mixture was incubated overnight at -70°C. RNA pellets were centrifuged at 12,000 g for 20 minutes at 4°C, washed with 75% ethanol and dissolved in RNAse-free water.

RNA concentration was measured using a NanoDrop ND-1000 Spectrophotometer. RNA quality was assessed by running 1µl of each RNA sample on an Agilent RNA 6000 Nano LabChip (Agilent Technology 2100 Bioanalyzer). Samples with an RNA Integrity Number (RIN) value higher than 8 were used for cDNA library construction by Illumina kit.

The construction of cDNA libraries from 5µg of total RNA was done by Illumina TruSeq™ RNA kit (Catalog # RS-930-20 01). Briefly, mRNAs were purified and fragmented

Table 6 Diploid and synthetic autopolyploid B. distachyon samples in leaf tissue were used for RNA-Seq experiment; the statistics of RNA-Seq read counts were obtained for each replicate from reads mapped by BWA and library size after normalization (Li and Durbin, 2010).

Inbred lines	Ploidy level	Replicate ID	Raw read number	Examination read number	Mapped read number	Mapped read with zero mismatches number	Library size after RPKM[a]
Bd21	Natural diploid 2n = 2x = 10	Bd21_L1	57707609	20095262	6794815	5235197	920228.9
		Bd21_L2	45352270	15975008	5360472	4201396	922201.8
		Bd21_L3	50365970	19129034	6340034	4890351	920046.4
	Synthetic autopolyploid 2n = 4x = 20	AutoBd21_L1	54475751	19506272	6590813	5195589	922455.4
		AutoBd21_L2	48687248	17343461	6300494	5033173	919111
		AutoBd21_L1	44540806	15135646	5504362	4460444	921539.5
Bd3-1	Natural diploid 2n = 2x = 10	Bd3-1-L1	62077017	21405632	7086849	5113838	922449.4
		Bd3-1-L2	57233560	23207849	7285704	5217827	920621.9
		Bd3-1-L3	50627187	19288641	6802850	4954586	919770.3
	Synthetic autopolyploid 2n = 4x = 20	AutoBd3-1-L1	52824829	17548568	5922125	4314941	922014.3
		AutoBd3-1-L2	53134852	17539332	6049169	4452190	922517.6
		AutoBd3-1_L3	55874124	18166356	6179348	4503886	921886.6

[a] RPKM: reads per kilobase of exon model per million reads

by magnetic bead[13] and Illumina elute-prime-fragment mix (containing oligo$(dT)_{25}$ Dynabeads, 1 M LiCl, 20 mM Tris-HCl pH 7.5, 2 mM EDTA, and SuperScript III first-strand). The cDNA was synthesized from mRNA by SuperScript II Reverse Transcriptase (ThermoFisher Scientific) for the first cDNA strand and then by Illumina second strand-master-mix (1.5 µL Blue Buffer , 1 µL dNTP mix (10 mM dATP, dCTP, dGTP, and dUTP), 0.2 µL RNase H (5 U/µL), 1 µL DNA polymerase I (10 U/µL), 1.3 µL H2O). The cDNA was purified from RNA by the Ampure XP beads. The following steps consist in repairing ends, ligating the adapters, and PCR amplification of the cDNA. The cDNA libraries were quantified and quality was controlled by running 1 µl of library on an Agilent DNA 1000 LabChip (Agilent Technology 2100 Bioanalyzer). cDNA libraries were indexed per genotype, and sequenced twice on a Illumina Solexa Hiseq2000 (100bp pair-end read).

III.2.2 Pilot analysis: comparison of gene expression between diploids and autopolyploids of B. distachyon

Analyses of gene expression using RNA sequencing require both the availability of reference genome sequences and gene annotation in order to map and quantify RNA-Seq reads on each gene. While the genome sequence of B. distachyon was available (IBI 2010), those of B. stacei and B. hybridum are now being completed. Therefore, as a pilot analysis, I present here comparison of gene expression between diploids and autopolyploids of B. distachyon.

The data analysis was adapted from previous works on wheat and B. napus (Chalabi 2014, Chalhoub, Denoeud et al. 2014). Briefly, the B. distachyon reference genome sequence that contains 31,029 predicted genes was used to align Illumina mRNA-Seq reads from leaves of diploids and autopolyploids plants derived from lines Bd21 and Bd3-1 of B. distachyon with three biological replicates for each genotype. RNA-Seq reads were mapped using BWA program with default parameters where 0 mismatch and minimum reads length of 35bp were applied (Version: 0.6.1-r104, seed 35, gap penalty 11) (described above). Mapped RNA-Seq reads were then filtered using SAMtools (Version: 0.1.12a) and only unique matches were considered. The normalization was performed by RPKM, then the read count table was computed with the method proposed in the DESeq (Differential expression analysis for sequence) package (Chalabi 2014).

[13] The poly-T oligo-attached magnetic bead was used to purify the poly-A containing mRNA molecules.

II.2.2.1 Read mapping, counting and normalization

The 31,029 predicted genes of B. distachyon (Bd21 inbred line) genome sequence (http://www.plantgdb.org/BdGDB/) were used as reference to map the RNA-Seq reads. These genes have a length ranging from 34 bp to 14,968 bp, and the average length of these genes is 1,604 bp and the median is 1,429 bp (Fig. 28).

A total of ~633 million Illumina pair-end mRNA-Seq reads, with an average of ~53 million reads (~44 to 62) per replicate per genotype have been generated on diploids and autopolyploids of Bd21 and Bd3-1 samples (Table 6). After trimming and filtering data with FastQC v0.10.1 on all read files in order to assess basic quality control metrics, a total of ~224 million reads (average ~18 million reads per replicate) were retained and were mapped onto B. distachyon reference genome (Table 6). Following our parameter an average of ~5-7 million reads per genotype per replicate were mapped onto genome totaling an average mapping rate of 34.08%, which is an acceptable result compared with the basal 30% of the total raw reads (Saminathan, Nimmakayala et al. 2015). Generally, the rate of RNA-Seq read mapping to the Bd21 reference genome was higher for diploid and autotetraploid samples of Bd21, than for Bd3-1, probably because of sequence divergence of the later.

When only mapped reads with zero mismatches were considered for differential expression analysis, 4,201,396 to 5,235,197 mapped reads per biological replicate were filtering and were used for normalization and comparison of gene expression.

The normalization was performed by RPKM (Table 6) and we focused on comparing gene expression between diploids and autopolyploids to characterize the effect of genome doubling.

A total of 28,134 expressed genes (out of 31,029 predicted ones) were considered in the comparisons (Table 7). Among these, 27,415 genes were detected as expressed in Bd21 auto- and/or diploid samples and 27,102 genes (87.34%) in those of Bd3-1, with 26,383 genes expressed in both lines, leaving 1,023 genes and 719 gene specifically expressed in Bd21 and Bd3-1, respectively.

Principal component analysis is a classical technique to resume features of RNA-Seq data (Yeung and Ruzzo 2001). The central idea of PCA is to reduce the dimensionality of the data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set

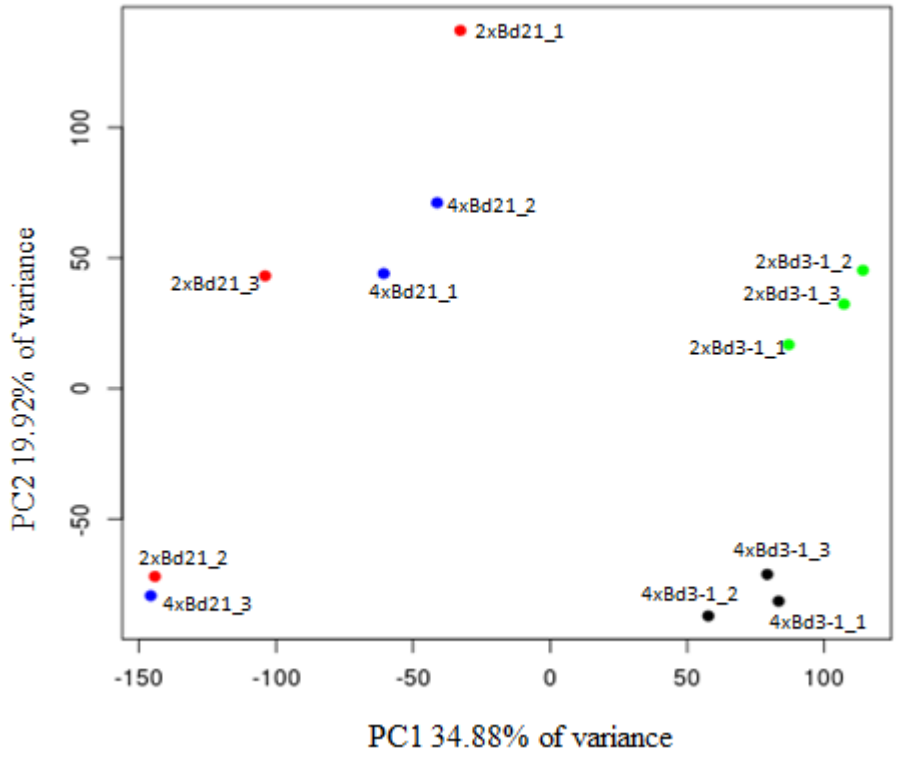


Fig. 29. The summary of the features of RNA-Seq data by principal component analysis (PCA). PCA was applied to summarize the expression data and identify spurious technical effects. The projection of the 12 samples of lines Bd21 and Bd3-1 of B. distachyon on the first two PC-axes in the sample space (two axes explaining 54.8% of the variance) indicate satisfactory reproducibility between biological replicates of Bd3-1 line. Whereas, the varieties have been found in six Bd21 samples.

2x: diploids, 4x: autotetraploids

of variables (the principal components) and summarize the feature of the data (Yeung and Ruzzo 2001, Jolliffe 2002). A PCA was applied to quickly summarize and look for spurious technical effects on RNA-Seq data (Fig. 29). The projection of the 12 samples on the first two PC-axes in the sample space shows a satisfactory reproducibility between biological replicates in six samples of Bd3-1 (the three replicates of diploid Bd3-1 and the three other replicates of autoBd3-1). However, more variability has been found in Bd21 diploid and autopolyploid samples. The three replicates of diploid Bd21 were not distributed as expected (Fig. 29). Because of these variations of Bd21 data, we did not continue analysis on the data from diploids and autopolyploids of line Bd21. Only data of diploids and autopolyploids from line Bd3-1 were used for analysis of differentially expressed genes.

II.2.2.2 Differentially expressed genes between autopolyploids and diploids of B. distachyon

We focused thus our analysis on revealing differentially expressed genes between autopolyploids and diploids of line Bd3-1 of B. distachyon.

Differential expression analysis for sequence count data (DESeq) package on R program (Anders and Huber 2010) has been widely applied for gene expression analysis (Rapaport, Khanin et al. 2013). DESeq use the related negative binomial distribution where the relation between the variance ν and mean μ is defined as $\nu = \mu + \alpha\mu$ 2 where α is the dispersion factor (Rapaport, Khanin et al. 2013).

When using DESeq and a False Discovery Rate (FDR-adjusted p-value) threshold of 0.05, we revealed 1,288 differentially expressed genes (DEGs) between autopolyploids and diploids of line Bd3-1 of B.distachyon, representing 4.75 % of total expressed genes. Among these, 235 genes (18% of total DEGs) were up-expressed genes and 1053 ones (82% of total DEGs) were down-expressed in the autopolyploids. Interestingly, there were 89 genes out of 1053 down-expressed ones that were expressed only in diploids but not detected (no RNA-seq reads that map) in autopolyploids. On the reverse, 23 out of 235 genes up-expressed in autopolyploids were not detected in diploids of Bd3-1 (Fig. 30).

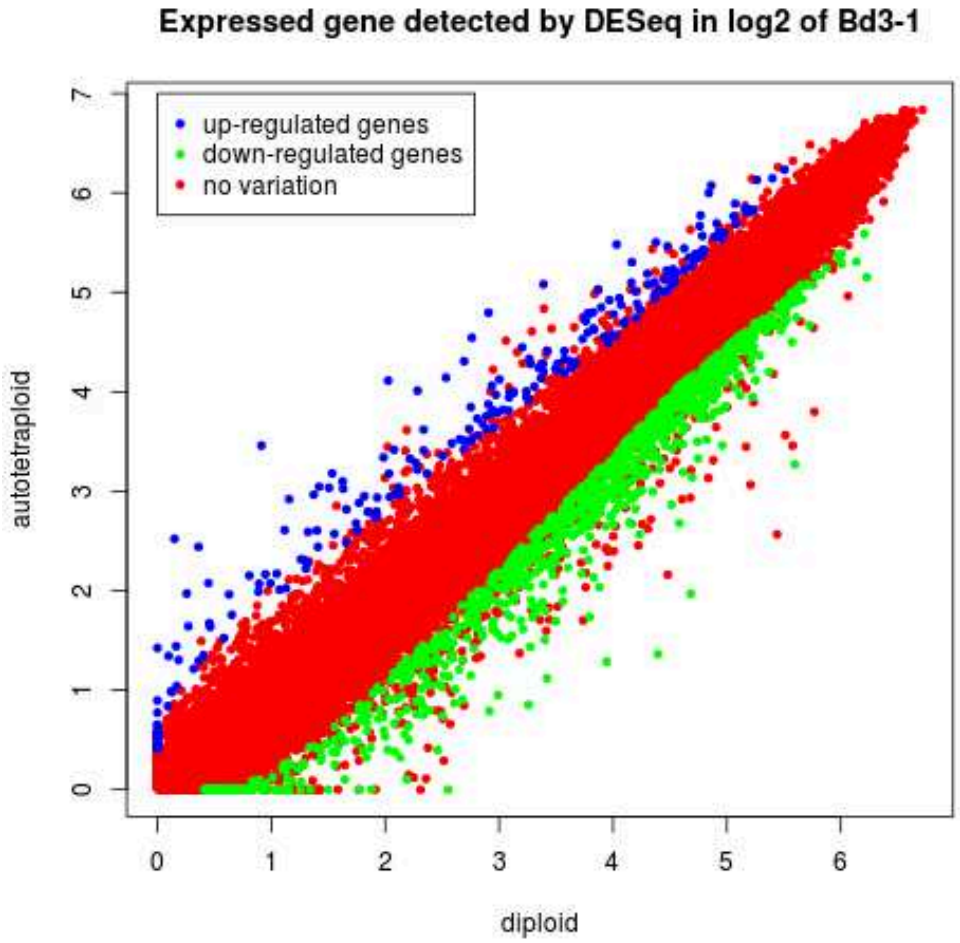


Fig. 30. Scatter plots showing 235 (blue) up- and 1053 (green) down-expressed genes in autopolyploids of as compared dipploids of line Bd3-1 of B. distachyon.
Detected by DESeq R/Bioconductor package (Anders and Huber 2010) in autopolyploid.

Functional classifications of Differentially Expressed Genes

To further analyze the putative function of the differentially expressed genes, we classified them according to their gene ontology (GO) using the public database (http://geneontology.org/). The 1,288 differentially expressed genes between autopolyploids and diploids of Bd3-1 were grouped into 25 GO terms of the three main functional categories: molecular function, biological process and cellular component (Fig. 31 and Tables 8, 9).

For the 235 up-expressed genes in autopolyploid of Bd3-1, we found enrichments in the molecular GO terms of “heme binding” and “tetrapyrrole binding”, with a fold enrichment of 4,15 and 2,53 (Tables 8, 9). The GO 'monooxygenase activity' was the most enriched GO term. This molecular function relates to the catalysis of the incorporation of one oxygen atom into a compound coupled with the reduction of the other atom of oxygen to water (Smith 2000).

Among the 1,053 down-expressed genes in the autopolyploid of Bd3-1, there were over-representations in several GO terms, including: “DNA polymerase activity”, “DNA-dependent ATPase activity” and “DNA binding” of the GO molecular function category (Tables 8, 9). These molecular functions correspond to specific biological process of cell cycle or cell division. The most enrichment GO terms of down-expressed genes were “mitotic cytokinesis”, “cytokinetic process”, and “chromosome organization involved in meiotic cell cycle” with 9.81; 9.81 and 9.18 fold as compared to expected samples (Tables 8, 9). Interestingly, there were also five GO terms of the biological process GO category that relate to the meiosis process (“meiotic cell cycle”, “meiotic cell cycle process”, “meiotic nuclear division”, “chromosome organization involved in meiotic cell cycle”, and “reciprocal meiotic recombination”), four other GO terms involve in mitotic process and many others were associated with cell cycle and DNA activities (Table 9).

Table 7. Genes expressed in autopolyploids and diploids of lines Bd21 and Bd3-1 of B. distachyon.

	Bd21	**Bd3-1**
Comparison of gene expression between inbred lines		
Total expressed genes	27415	27102
Specific expressed genes	1032	719
Total expressed genes of both two lines	28134	28134
Common to Bd21 and Bd3-1	26383	26383

II.2.2.2 Discussion of gene expression changes in autopolyploids of B. distachyon

The pilot analysis that we have conducted here, comparing gene expression between diploids and autopolyploids of B. distachyon, illustrates the strength of Illumina-based RNA-sequencing in estimating the regulation of gene expression that follows whole genome duplication by autopolyploidization. The RNA-Seq data from autopolyploids and diploids of line Bd3-1 have been mapped on the reference genome of B. distachyon line Bd21 at similar rates (~34.08%,) of those observed in other species (Saminathan, Nimmakayala et al. 2015). Thus the procedure that I have used here is reliable to perform gene expression analysis.

The important variations between biological replicates of diploid and autopolyploids of line Bd21 as estimated by PCA illustrates that comparisons could not be done and few DEGs could be detected (data not shown). Occasionally errors may have occurred in an experimental protocol and/or a sample generates improper data. A recent study had also demonstrated that "bad" replicates can have a profound distorting effect on differential expression results. Therefore, these data should be rejected (Gierliński, Cole et al. 2015).

DESeq is one of the powerful statistical approaches for the estimation of differentially expressed genes (Schurch, Schofield et al. 2016). This approach calculates the significance of the changes in gene expression between two conditions using a variation of the Fisher exact test, adopted for binomial distribution. DESeq controls false positives to select the highly significant differences in gene expression. Therefore this analysis gives reliable results (Anders and Huber 2010, Rapaport, Khanin et al. 2013). A false discovery rate (FDR)-adjusted p-value threshold of 0.05 has been applied to test selected DEGs with high significance.

There are considerations about appropriateness of using equal amount of RNA when comparing gene expression between diploids and derived autopolyploids, which have thus different genome sizes and presumably different cell sizes (that I did not estimate here), and whether it is reflecting gene expression. In our experiments, I obtained similar gene expression estimations (RPKM) when comparing 200 ng and 100ng RNA used for sequencing of line Bd21 (data not shown). This is also in agreement with similar results obtained in the laboratory with polyploid wheat (Chalabi 2014), showing similar gene expression estimations when normalizing RNA-Seq data through RPKM and reducing by half RNA quantities. Based on these experiments, I estimated that using equal amounts of RNA of diploids and autopolyploids of B. distachyon to appreciate gene expression, through Illumina RNA sequencing and RPKM normalization, is revealing true differences in gene

Table 8. Gene ontology (GO) classification of 1,288 differentially expressed genes in autopolyploids as compared to diploids of line Bd3-1 of B. distachyon (235 up- and 1053 down-expressed) as revealed by DESeq R/Bioconductor package (Anders and Huber 2010)

GO classification	Bd3-1	
	Up-expressed gene number[a]	Down-expressed gene number
Molecular function		
Translation regulator activity	2	18*
Nucleic acid binding transcription factor activity	1	20*
Binding	17*	159*
Receptor activity	6	18
Enzyme regulator activity	2	19
Structural molecule activity	2	34
Catalytic activity	62*	227*
Transporter activity	15	50
Protein binding transcription factor activity	0	4
Unmapped[b]	128	504
Biological process		
Apoptotic process	1	10
Response to stimulus	12	60
Developmental process	1	16
Cellular process	24	167*
Metabolic process	71*	303*
Biological regulation	10	74*
Cellular component organization or biogenesis	6	32
Localization	20	78*
Reproduction	0	3
Locomotion	0	1
Unmapped	90	309
Cellular component		
Membrane	4	23
Macromolecular complex	2	38
Cell part	16	98*
Organelle	8	67*
Cell junction	0	1
Extracellular region	0	5
Unmapped	205	821

[a] The number of genes that were mapped to the particular annotation data category of reference genome.

[b] Number of genes that were not mapped to the particular annotation data category of reference genome.

* GO class which contains the overrepresented terms in Table 9

expression. This is also in agreement with most of the gene expression comparison between diploids and autopolyploids where equal amount of RNA from diploids and autopolyploids were used and considered as not leading to any technical bias (Dai, Wang et al. 2015, Saminathan, Nimmakayala et al. 2015, Zhou, Kang et al. 2015).

Thus, using thus equal amounts of total RNA and similar sequencing depth in diploids and derived autotetraploids and applying DESeq, the majority (95.25%) of the 27,102 expressed genes remains equally expressed between diploids and autopolyploids implying that homologous genome doubling did not result in important changes of gene expression. This leaves 1,288 differentially expressed genes (4.75% of total genes). Surprisingly, the majority of these genes (1,053 representing 82 %) were down-expressed in Bd3-1 autopolyploids. Similar results were found in autopolyploids of cabbage (Brassica oleracea L.), where no major alterations in gene expression was observed as compared to diploid progenitors (Albertin, Brabant et al. 2005). In other species, the difference in gene expression between autopolyploid and diploid progenitor represent 2.87% of genes in Mulbery (Morus alba L.) and 6.09% in Paulownia fortunei (Zhang, Deng et al. 2014, Dai, Wang et al. 2015) and 1.8% genes in Arabidopsis (Yu, Haberer et al. 2010).

It has been revealed in several species that the differentially expressed genes between diploids and autopolyploids were related to certain pathways. In the traditional medicinal plant Chinese woad (Isatis indigotica Fort.), the up- expressed genes were mainly involved in “cell growth”, “cell wall organization”, “secondary metabolite biosynthesis”, “response to stress and photosynthetic” pathways (Zhou, Kang et al. 2015). Whereas, in Mulberry (Morus alba L.), down-expressed genes are associated with the biosynthesis, photosynthetic and signal transduction of plant hormones, including cytokinins, gibberellins, ethylene, and auxin, (Dai, Wang et al. 2015). In the present study, the up- and down-expressed genes in the autopolyploid are involved in different molecular functions and biological processes. The oxidation-reduction process was over-reprenseted in the up-expressed genes, while down-expressed genes were enriched in cell division and DNA replication processes. More precisely, the down-expressed genes in Bd3-1 autopolyploids were enriched in several GO terms that relate to cell cycle, meiosis, and mitosis. It is important to note that these genes are implicated in genome maintenance functions, including nuclear chromosome segregation, DNA repair, DNA replication, homologous recombination and transcriptional regulation. Similarly, in autotetraploid watermelon, Saminathan, Nimmakayala et al. (2015)

Table 9. Gene ontology (GO) enrichment analysis of 235 up- and 1053 down-expressed genes in autopolyploids as compared to diploids of line Bd3-1 of B. distachyon.

GO classification	Occurrence[a]	Sample expected	Fold Enrichment	+/-[a]	P value
Up-expressed genes					
GO molecular function					
Monooxygenase activity	14	2.76	5.08	+	1.02E-03
Heme binding	33	4.05	3.7	+	1.83E-02
Tetrapyrrole binding	15	4.33	4.15	+	5.27E-04
Oxidoreductase activity	18	13.03	2.53	+	1.04E-03
Unclassified	78	103.04	0.76	-	0.00E+00
GO biological process					
Oxidation-reduction process	33	13.84	2.39	+	5.46E-03
Unclassified	94	109.39	0.86	-	0.00E+00
GO cellular component					
Unclassified	127	125.64	1.01	+	0.00E+00
Down-expressed genes					
GO molecular function					
DNA polymerase activity	8	1.03	7.76	+	1.26E-02
DNA-dependent ATPase activity	13	2.77	4.69	+	6.97E-03
DNA binding	103	65.27	1.58	+	4.78E-03
Unclassified					
GO biological process					
Cytokinesis by cell plate formation	7	0.55	12.62	+	2.91E-03
Mitotic cytokinetic process	7	0.71	9.81	+	1.48E-02
Cytoskeleton-dependent cytokinesis	7	0.75	9.3	+	2.08E-02
Cell cycle process	31	5.94	5.21	+	4.34E-10
Cell cycle	32	6.82	4.69	+	2.75E-09
Mitotic cell cycle process	19	2.81	6.75	+	2.62E-07
Mitotic cell cycle	19	2.85	6.66	+	3.29E-07
Mitotic cytokinesis	7	0.71	9.81	+	1.48E-02
Cytokinetic process	7	0.71	9.81	+	1.48E-02
Chromosome organization involved in meiotic cell cycle	8	0.87	9.18	+	5.61E-03
Meiotic cell cycle	14	2.5	5.61	+	5.83E-04
Meiotic cell cycle process	14	2.38	5.89	+	3.28E-04
Regulation of mitotic cell cycle	11	1.59	6.94	+	1.35E-03
Regulation of cell cycle	14	3.41	4.11	+	1.99E-02
Nuclear chromosome segregation	10	1.51	6.64	+	6.15E-03
Regulation of cell cycle process	11	1.7	6.45	+	2.69E-03
Reciprocal meiotic recombination	9	1.43	6.31	+	2.75E-02
Reciprocal DNA recombination	9	1.43	6.31	+	2.75E-02
DNA recombination	19	3.84	4.94	+	3.78E-05
DNA metabolic process	50	16.09	3.11	+	9.04E-09
Meiotic nuclear division	12	2.22	5.41	+	5.59E-03

demonstrated that cell-cycle-related genes such as the cell division cycle protein 123-like and cell division control proteins were down-regulated. On the contrary, it was found in autopolyploids of A. arenosa, six meiosis-related genes which have functional classes corresponding to chromosome cohesion, segregation and repair show up-expression (Hollister, Arnold et al. 2012).

The down-expression of genes related to cell cycle, meiosis, and mitosis in Bd3-1 autopolyploid may explain their lower fertility, characterized by a decrease number of seeds per inflorescence and lower percent of fertile florets as compared to the diploids (Dinh-Thi, Clainche et al. 2016).

III.3. Concluding remarks and planed perspectives

Autopolyploids from the other diploid species B. stacei as well as synthetic B. hybridum allopolyploids are being also characterized together with natural diploids and allopolyploids at the genomic, DNA methylation and transcriptomic levels. While the B. distachyon genome is already available (IBI 2010) de novo and accurate sequencing of B. stacei and B. hybridum is still in progress (Catalan, López-Álvarez et al. 2016). The future availability of the genome reference sequences from these three species, together with the integrative characterization of structural, methylation and transcriptomic changes that I had already started will certainly raise Brachypodium as one of the most important polyploidy model. This will allow a better and integrative understanding of the biological responses to polyploidy and the secrets of polyploidy success

Table 9. (Continue) Gene ontology (GO) enrichment analysis of 235 up- and 1053 down-expressed genes in autopolyploids as compared to diploids of line Bd3-1 of B. distachyon.

GO classification	Occurrence[a]	Sample expected	Fold Enrichment	+/-[a]	P value
Nuclear division	17	3.01	5.64	+	3.14E-05
Organelle fission	17	3.61	4.71	+	3.88E-04
Double-strand break repair via homologous recombination	12	2.02	5.94	+	2.17E-03
Recombinational repair	12	2.02	5.94	+	2.17E-03
DNA repair	29	9.47	3.06	+	3.28E-04
Cellular response to DNA damage stimulus	30	9.87	3.04	+	2.37E-04
Double-strand break repair	16	2.85	5.61	+	8.83E-05
DNA-dependent DNA replication	14	2.66	5.27	+	1.20E-03
DNA replication	23	4.36	5.28	+	3.92E-07
Unclassified	497	492.24	1.01	+	0.00E+00
GO cellular component					
Mismatch repair complex	5	0.28	18.02	+	4.80E-03
DNA repair complex	5	0.4	12.62	+	2.59E-02
Replisome	5	0.44	11.47	+	4.04E-02
Chromosomal part	21	7.57	2.77	+	1.80E-02
Chromosome	35	10.03	3.49	+	2.15E-07
Intracellular non-membrane-bounded organelle	65	38.09	1.71	+	1.40E-02
Non-membrane-bounded organelle	65	38.09	1.71	+	1.40E-02
Condensed chromosome	9	1.51	5.98	+	1.24E-02
Kinesin complex	10	2.14	4.67	+	3.50E-02
Microtubule associated complex	12	2.89	4.15	+	2.18E-02
Microtubule cytoskeleton	15	4.48	3.35	+	2.96E-02
Nuclear chromosome	17	4.84	3.52	+	5.45E-03
Unclassified	543	565.4	0.96	-	0.00E+00
Cytoplasm	126	170.18	0.74	-	3.35E-02

[a] the number of regulated genes that map to the annotation data category.

[b] the number of genes you would expect in your list for the category, based on the reference list.

[c] over/underrepresentation

GENERAL CONCLUSSION AND PERSPECTIVES

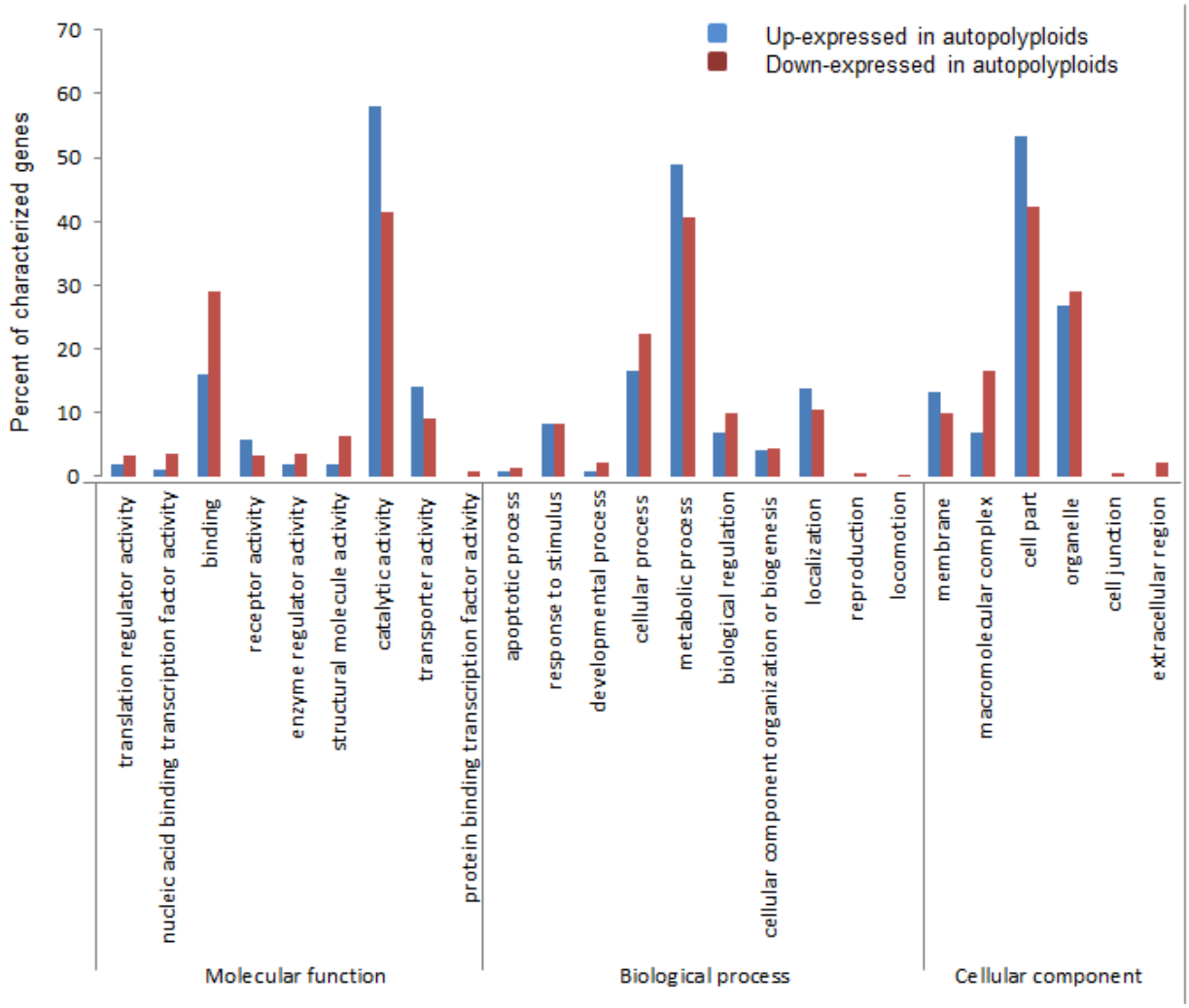


Fig. 31. Gene ontology (GO) classifications of differentially expressed genes detected by DESeq between autopolyploids and diploids of line Bd3-1 of B. distachyon

During my PhD thesis, I successfully synthesized and characterized Brachypodium autopolyploids and allopolyploids, achieving an important step towards the development of a Brachypodium polyploid model. This original model is based on B. distachyon and B. stacei sister species that have common evolutionary history, but get since they diverged ~16 MYA, completely different chromosome evolution, leading to B. distachyon having two times less chromosomes (2n=10) that are consequently two times bigger than B. stacei (Hasterok, Draper et al. 2004, Lopez-Alvarez, Lopez-Herranz et al. 2012, Catalan, López-Álvarez et al. 2015, Catalan, López-Álvarez et al. 2016). Thus, B. distachyon have originated via descending dysploidy, acting as fusions of chromosomes from a putative ancestral Brachypodium species, very close to B. stacei (Hasterok, Draper et al. 2004, Betekhtin, Jenkins et al. 2014). Moreover, the two species have naturally hybridized to give rise to the natural B. hybridum allopolyploid.

Most of my results have been discussed in the relevant chapters presented as submitted or accepted papers (See Results section). I will resume herein my main findings and conclusions, giving more attentions and highlights on the perspectives that my results offer towards deciphering polyploidy mechanisms in flowering plants.

Few studies have focused on the importance and consequences of autopolyploidy. **For this purpose, the first challenge of my PhD thesis was to create and characterize autopolyploids from two Brachyploidum species**. I have successfully generated autotetraploids from two inbred lines of B. distachyon (Bd21 and Bd3-1) and three ones of B. stacei (ABR114, Bsta5 and TE4.3), through colchicine treatments. The different autotetraploids were characterized by flow cytometry and karyotyping with fluorescent in situ hybridization analyses, as well as at the phenotype and fertility levels. All autopolyploids gave stable autotetraploids as evaluated at the cytogenetic and phenotype levels, except one derived from ABR114 of B. stacei (ABR114_2) that gave aneuploids with unstable phenotypes. Precise quantitative comparisons of inflorescences and flag leaves characters showed that both B. distachyon and B. stacei autotetraploids generally exceeded their diploid progenitors, but their fertility was reduced.

The second important challenge of my PhD thesis was to recreate Brachypodium hybridum allopolyploids by uniting, through interspecific hybridization, the divergent genomes of B. distachyon and B. stacei. B. hybridum (2n=30) is a natural allopolyploid with

parental species, B. distachyon (2n=10) and B. stacei (2n=20), contrasting in chromosome evolution, although having the same ploidy level. Several groups have previously failed in resynthesizing B. hybridum through hybridization between its parental species (G. Linc and R. Hasterok, unpublished). To maximize chances and successfully recreate B. hybridum allopolyploids, I carried out two approaches. The first one was to cross natural diploid B. distachyon and B. stacei to create an amphihaploid F1, whose spontaneous or induced chromosome doubling could lead to an allotetraploid, similar to the natural allotetraploid B. hybridum. The second approach was to directly hybridize B. distachyon and B. stacei autotetraploid plants that have been created previously. Since these autotetraploids should have 2n gametes, the expected F1 progenies would be allotetraploids without need of further chromosome doubling. Two different lines of B. distachyon (Bd21, Bd3-1) and two lines of B. stacei (ABR114 and Bsta5), representing four genotype combinations and eight reciprocal cross types, were used to maximize the chances of success. Over a period of four years, I realized a total of 9,388 reciprocal crosses between the diploid plants of the two diploid species and 4,384 crosses between their autotetraploid forms. While sterile amphihaploid F1 interspecific hybrids were obtained at low frequencies when diploid B. distachyon was used as the maternal parent (0.15% or 0.245% depending on the line used), no hybrids were obtained from reciprocal crosses or when autotetraploids of the parental species were crossed. Genome doubling through colchicine treatment restored fertility where doubled F1 plants produced a few S1 seeds after self-pollination. S1 plants from one parental combination (Bd3-1×Bsta5) were fertile and gave rise to further generations whereas those of another parental combination (Bd21×ABR114) were sterile, illustrating the dependence of fertile allopolyploid formation based on parental genotypes. The synthetic allotetraploids were shown to be highly-stable and resembled the natural B. hybridum at the phenotypic, cytogenetic and genomic levels. The recreation of synthetic allotetraploid similar to natural B. hybridum provides empirical evidence and establishes the tractable Brachypodium allopolyploid model (Hasterok, Draper et al. 2004, Idziak and Hasterok 2008, Catalan, Muller et al. 2012, Lopez-Alvarez, Lopez-Herranz et al. 2012, Catalan, Chalhoub et al. 2014). This represents a unique allopolyploid model where one parental genome (B. distachyon) has similar genome size to the other one (B. stacei), but twice lower basic chromosome number (2n=10 and 2n=20, respectively) whereas its individual chromosome size is approximately two times larger. Importantly, it offers the possibility to investigate allopolyploidy-related changes at genome structure and regulation levels at the earliest evolutionary stages of the polyploid existence in comparison to the natural B. hybridum.

The projects that I have undertook in developing and characterizing these Brachypodium autopolyploids and allopolyploids have benefited from the strong experience that the group of research that I have integrated had acquired on other polyploid models such as wheat and Brassica (Mestiri, Chagué et al. 2010) (Charles, Belcram et al. 2008, Charles, Tang et al. 2009, Chagué, Just et al. 2010, Arnaud, Chelaifa et al. 2013, Chelaifa, Chagué et al. 2013, Chalhoub, Denoeud et al. 2014).

In the scenario and based on the experience available in the lab, as soon as Brachypodium autopolyploids and allopolyploids have been developed, I was able to rapidly design and set up genomics experiments to **characterize polyploidy-related changes at genome structure** (through DNA re-sequencing) **cytosine methylation** (through bisulfite sequencing) **and gene expression** (through RNA-Sequencing) **levels, constituting the third main challenge of my PhD project.** I undertook these characterizations, before being able to analyze the data as we were waiting for the completion of reference sequencing of the genomes of B. stacei and B. hybridum, through Brachypodium international consortium to which my group contributes.

The preliminary analyses that I was able to do, comparing gene expression between autopolyploids and diploids of line Bd3-1 of B. distachyon (thanks to availability of the genome sequence), illustrates the high potential to rapidly characterize impacts of autopolyploidy and allopolyploidy. Majority (95.25%) of the 27,102 expressed genes remain equally expressed implying that homologous genome doubling did not result in important changes of gene expression. This leaves 1,288 differentially expressed genes, majority of which (1,053 representing 82 %) were surprisingly down-expressed in autopolyploids. These were enriched in gene entology related to cell cycle, meiosis, and mitosis and may explain their lower fertility of autopolyploids as compared to the diploids. This pilot experiment is very promising and illustrates the power of NGS in characterizing responses to auto and allopolyploidy as soon as the reference genomes of B. stacei and B. hybridum will be available.

In conclusion, I have developed through my PhD thesis an original Brachypodium autopolyploid and allopolyploid model and materials. The developed autotetraploids, which do not exist in nature, provide an interesting material to study the fate of homologous duplicated genes, meiosis and various genomic consequences of autopolyploidy,

comparatively between B. distachyon and B. stacei species, having similar genome content but contrasting in chromosome evolution and structure. The developed B. hybridum allopolyploid model represents also a unique allopolyploid model where parental genomes, and by corollary the sub-genomes of B. hybridum, are highly divergent in chromosome number and size. Together, these three Brachypodium species comprise an excellent model to investigate the impact of polyploidy on the organization and evolution of plant genomes, because they possess small genomes, have small plant stature, rapid generation time, and significant morphometric and molecular barcoding differences (Catalan, Muller et al. 2012, Lopez-Alvarez, Lopez-Herranz et al. 2012).

I suggest that in the near future, further exploitation of this important material and NGS data could help in:

(i) Elucidating consequences of polyploidy at the structural, epigenetic and functional genomics levels through an integrative analysis of various "omics" data that I have generated on this interesting material as soon as the genome sequences of all three species will be completed and available.

(ii) Studying the regulation of meiosis in both autopolyploids, in which four homologous chromosomes are present, as well as in allopolyploids. Several important questions should be addressed such as: pairing of the four homologous chromosomes in autopolyploids, during meiosis, and how it is regulated? Does the size difference between subgenomes represent a structural barrier for homoeologous pairing? Or reversely, does the pairing occur between homoeologous sequences.

(iii) What are the barriers for sterility or reduced fertility in autopolyploids and allopolyploids. Cytological and cytogenetic comparisons between diploids and autopolyploids as well as between the sterile synthetic allopolyploid allo21×114, the fertile allo3-1×5 and the natural B. hybridum should give insights into processes as well as possible cytological and cytogenetic barriers.

Annexes

RESUME GENERAL

La polyploïdie consiste en la duplication du génome entier et constitue une force évolutive majeure chez les eucaryotes, notamment chez les angiospermes. Les espèces du genre Brachypodium ont émergé au cours de la dernière décennie comme une modèle intéressante des monocotylédones, parce qu'elles possèdent de petits génomes, une petite taille de la plante, un cycle de vie rapide et une évolution chromosomique très divergente. Parmi ces espèces, l'allopolyploïde B. hybridum (2n = 30) a résulté de l'hybridation entre les espèces B. distachyon (2n = 10) et B. stacei (2n = 20). Les deux espèces parentales ont eu des évolutions chromosomiques assez divergentes aboutissant à ce que B. distachyon possède deux fois moins de chromosomes qui sont néanmoins deux fois plus grands. En 2010, le séquençage du génome de B. distachyon a été publié. Depuis lors, il devient la séquence du génome de référence de haute qualité appliquée dans de nombreuses études sur de nombreuses espèces de Monocote, en particulier sur la famille Poaceae. Les génomes de B. stacei et B. hybridum sont en cours de séquençage et d'assemblage. Les génomes annotés de haute qualité de progéniteurs diploïdes (B. distachyon et B. stacei) ainsi que leur allopolyploïde dérivé (B. hybridum) seraient un système utile pour déchiffrer les détails mécanistiques exacts de l'établissement, de la stabilisation et de l'évolution des polyploïdes.

Dans ce contexte, l'objectif de ma thèse est de développer un modèle de Brachypodium polyploïde en synthétisant des autopolyploïdes et les allopolyploïdes ainsi que les caractérisant aux niveaux phénotypique, cytogénétique et génomique.

Les autotétraploïdes ont été générées par traitement de colchicine de deux lignées de Brachypodium distachyon (Bd21 et Bd3-1) et trois autres de B. stacei (ABR114, Bsta5 et TE4.3). Les effets du traitement de colchicine dépendent du stade de développement et des espèces de Brachypodium. Le taux de survie assez faible au stade de germination (37,8%) et des lignées B. stacei (la moyenne de 37% et 0% au stade des plantulles et de la germination respectivement) a indiqué qu'elles sont moins tolérées à la colchicine. Le doublement du génome a été validé par les analyses de cytométrie de flux et de caryotype avec l'hybridation in situ fluorescente. Alors que la cytométrie de flux a donné une indication approximative et primaire de l'augmentation de la taille du génome, les caractérisations cytogénétiques plus précisent ont été réalisées pour confirmer les plantes autopolyploïdes et élucider les plantes aneuploïdes par les colorations DAPI des chromosomes et l'hybridation FISH. Les

autopolyploïdes obtenues à partir des deux lignées de B. distachyon et les lignées Bsta5 et TE4.3 de B. stacei ont montré la stabilité dans le phénotype ainsi que le caryotype. Pour la troisième lignée de B. stacei (ABR114) une plante traitée par colchicine a donné des autotétraploïdes stables, alors qu'une autre a donné des descendances aneuploïdes. Les comparaisons des caractères de l'inflorescence et des ligules montrent que les deux autotétraploïdes des deux espèces dépassent généralement leurs progéniteurs diploïdes, mais leurs fertilité ont été réduites comme illustré par le nombre de graines par inflorescence et le nombre de graines par le nombre total de inflorescence des fleurs plus faible.

Les allotétraploïdes synthétiques, similaire à B. hybridum, ont été générés par hybridation interspécifique entre les différentes lignées de B. distachyon et B. stacei. En outre, sur une période de quatre ans, un nombre total de 9,388 croisements réciproques ont été faites entre les plantes diploïdes des deux espèces diploïdes et de 4,384 croisements entre leurs autotétraploïdes. Alors que les hybrides interspécifiques F1 stériles ont été obtenus lorsque B. distachyon a été utilisée comme parent maternel (0,15% ou 0,245% en fonction de la lignée utilisée), aucun hybride a été obtenue à partir des croisements réciproques ou lorsque des plantes autotetraploïdes des espèces parentales ont été croisées. Le doublement du génome par traitement à la colchicine a restauré la fertilité et les plantes F1 doublées produisent par auto-polllinisation quelques graines S1 (première génération d'autofécondation). Les plantes S1_3-1×5 étaient fertiles donnants des générations suivantes alors que ceux de la S1_21×114étaient stériles, illustrant la dépendance de la formation des allopolyploïdes en fonction des génotypes parentaux. La recréation synthétique d'allotétraploïde qui est similaire à B. hybridum naturelle fournit des preuves empiriques et établit le modèle Brachypodium allopolyploïde docile. Cela représente un modèle allopolyploïde unique où un génome parental (B. distachyon) a la taille du génome similaire à l'autre (B. stacei), mais possède la moitié du nombre de chromosomes de base (2n = 10 et 2n = 20, respectivement), tandis que la taille de son chromosome individuel est environ deux fois plus grande. Ces allotétraploïdes synthétiques se sont montrées stables au niveau phenotypique, cytogénétique et génomiques, ressemblant le B. hybridum naturel. Les différences importantes dans le nombre de chromosome ainsi que la taille de chromosome des deux génomes parentaux pourraient probablement représenter un facteur dans la stabilitie des B. hybridum naturelle et synthétique car ils peuvent constituer un obstacle pour l'appariement et le réarrangement homéologue. Presque tous les traits d'inflorescence des allopolyploïdes synthétiques ont montré le dépassement à ceux des progéniteurs diploïdes,

tandis que la caractéristique fertile en F1 allo3-1 × 5 est très faibles, et inscrease chez la génération d'avantage.

Les autopolyploïdes et allopolyploïdes de Brachypodium que j'ai générés ainsi que les espèces naturelles constituent un matériel original pour la caractérisation des modifications génétiques, épigénétiques et fonctionnels liées à la polyploïdie, sur le court et le long terme de l'évolution. J'ai ainsi utilisé la méthodologie Illumina de nouvelles générations de séquençage (NGS) pour générer des données de: (i) reséquençage de l'ADN afin de caractériser les variations structurales, (ii) du transcriptome par l'ARN-Seq afin de caractériser les modifications de l'expression des ène (iii) la méthylation CpG par séquençage au bisulfite. Une large gamme de possibilités de comparaison pourrait être réalisée lorque les séquences des génomes des trois espèces naturelles de Brachypodium seraient disponible permettant ainsi d'élucider l'effet du doublement du même génome par autopolyploidie ainsi quepar allopolyploïdie.

La disponibilité de la séquence du génome de B. distachyon a permis de réaliser une analyse pilote comparant l'expression des gènes dupliqués dans les autopolyploïdes de Bd3-1 à ceux des diploïde. Dans cette étude, nous avons utilisé la technique DESeq avec la mise en place d'un seuil ajusté pour p-valeur de taux de fausses découvertes (FDR) < 0,05 pour étudier l'expression des gènes. Les plantes diploïdes et polyploïdes de B. distachyon ont montré une expression similaire pour la plupart des gènes (95,25%). Seulement 4,75% des gènes ont montrés une expression différentielle. La majeure partie de ces derniers (1053 gènes, soit 82%) étaient moins exprimés dans les autopolyploides. Des fonctions géniques tel que la ségrégation des chromosomes nucléaires, la réparation de l'ADN, la réplication de l'ADN, la recombinaison homologue, la régulation de la transcription sont enrichis et auraient un lien avec la faible fertimitée observée chez les autopolyploides. Cette étude pilote est très promettant. Les plantes autopolyploïdes des autres espèces diploïdes B. stacei ainsi que les B. hybridum allopolyploïdes synthétiques ont également été générées et caractérisées en même temps que les plantes diploïdes et allopolyploïdes naturelles au niveau du séquençage de l'ADN traité au bisulfite ainsi qu'au niveau du séquençage d'ARN.

La création des autopolyploides et allopolyploides offre la possibilité d'étudier des changements liés a la polyploïdie aux niveaux structurales, fonctionnelle et épigénétique ainsi que la régularité de la méiose.

REFERENCES

Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. J. Baird, N. Bierne, J. Boughman, A. Brelsford, C. A. Buerkle, R. Buggs, R. K. Butlin, U. Dieckmann, F. Eroukhmanoff, A. Grill, S. H. Cahan, J. S. Hermansen, G. Hewitt, A. G. Hudson, C. Jiggins, J. Jones, B. Keller, T. Marczewski, J. Mallet, P. Martinez-Rodriguez, M. Most, S. Mullen, R. Nichols, A. W. Nolte, C. Parisod, K. Pfennig, A. M. Rice, M. G. Ritchie, B. Seifert, C. M. Smadja, R. Stelkens, J. M. Szymura, R. Vainola, J. B. Wolf and D. Zinner (2013). "Hybridization and speciation." J Evol Biol **26**(2): 229-246.

Adams, K. L. (2007). "Evolution of duplicate gene expression in polyploid and hybrid plants." J Hered **98**(2): 136-141.

Adams, K. L., R. Percifield and J. F. Wendel (2004). "Organ-Specific Silencing of Duplicated Genes in a Newly Synthesized Cotton Allotetraploid." Genetics **168**(4): 2217-2226.

Aïnouche, M. L., P. Fortune, A. Salmon, C. Parisod, M.-A. Grandbastien, K. Fukunaga, M. Ricou and M.-T. Misset (2009). "Hybridization, polyploidy and invasion: lessons from Spartina (Poaceae)." Biological invasions **11**(5): 1159-1173.

Akalin, A., M. Kormaksson, S. Li, F. E. Garrett-Bakelman, M. E. Figueroa, A. Melnick and C. E. Mason (2012). "methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles." Genome Biology **13**(10): 1-9.

Akama, S., R. Shimizu-Inatsugi, K. K. Shimizu and J. Sese (2014). "Genome-wide quantification of homeolog expression ratio revealed nonstochastic gene regulation in synthetic allopolyploid *Arabidopsis*." Nucleic Acids Res **42**(6): e46.

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

Ali, H. B., M. A. Lysak and I. Schubert (2004). "Genomic in situ hybridization in plants with small genomes is feasible and elucidates the chromosomal parentage in interspecific Arabidopsis hybrids." Genome **47**(5): 954-960.

Alkan, C., J. M. Kidd, T. Marques-Bonet, G. Aksay, F. Antonacci, F. Hormozdiari, J. O. Kitzman, C. Baker, M. Malig, O. Mutlu, S. C. Sahinalp, R. A. Gibbs and E. E. Eichler (2009). "Personalized copy number and segmental duplication maps using next-generation sequencing." Nat Genet **41**(10): 1061-1067.

Anders, S. and W. Huber (2010). "Differential expression analysis for sequence count data." Genome Biology **11**(10): 1-12.

Arnaud, D., H. Chelaifa, J. Jahier and B. Chalhoub (2013). Reprogramming of Gene Expression in the Genetically Stable Bread Allohexaploid Wheat. Polyploid and Hybrid Genomics, John Wiley & Sons, Inc.**:** 195-211.

Axtell, M. J. (2013). "Classification and Comparison of Small RNAs from Plants." Annual Review of Plant Biology **64**(1): 137-159.

Bannister, A. J. and T. Kouzarides (2011). "Regulation of chromatin by histone modifications." Cell Res **21**(3): 381-395.

Bao, R., L. Huang, J. Andrade, W. Tan, W. A. Kibbe, H. Jiang and G. Feng (2014). "Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing." Cancer Inform **13**(Suppl 2): 67-82.

Barker, M. S., N. Arrigo, A. E. Baniaga, Z. Li and D. A. Levin (2016). "On the relative abundance of autopolyploids and allopolyploids." New Phytol **210**(2): 391-398.

Barker, M. S., G. J. Baute and S.-L. Liu (2012). Duplications and Turnover in Plant Genomes. Plant Genome Diversity Volume 1: Plant Genomes, their Residents, and their Evolutionary

Dynamics. F. J. Wendel, J. Greilhuber, J. Dolezel and J. I. Leitch. Vienna, Springer Vienna**:** 155-169.
Bartel, D. (2004). "MicroRNAs: genomics, biogenesis, mechanism, and function." Cell **116**(2): 281 - 297.
Bartos, J., E. Paux, R. Kofler, M. Havrankova, D. Kopecky, P. Suchankova, J. Safar, H. Simkova, C. D. Town, T. Lelley, C. Feuillet and J. Dolezel (2008). "A first survey of the rye (Secale cereale) genome composition through BAC end sequencing of the short arm of chromosome 1R." BMC Plant Biol **8**: 95.
Baucom, R. S., J. C. Estill, C. Chaparro, N. Upshaw, A. Jogi, J. M. Deragon, R. P. Westerman, P. J. Sanmiguel and J. L. Bennetzen (2009). "Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome." PLoS Genet **5**(11): e1000732.
Beaulieu, J., M. Jean and F. Belzile (2009). "The allotetraploid *Arabidopsis thaliana-Arabidopsis lyrata* subsp. *petraea* as an alternative model system for the study of polyploidy in plants." Mol Genet Genomics **281**(4): 421-435.
Beest, M., J. J. Le Roux, D. M. Richardson, A. K. Brysting, J. Suda, M. Kubešová and P. Pyšek (2011). "The more the better? The role of polyploidy in facilitating plant invasions." Annals of Botany **109**(1): 19-45.
Bennett, M. D. and J. B. Smith (1976). "Nuclear dna amounts in angiosperms." Philos Trans R Soc Lond B Biol Sci **274**(933): 227-274.
Bennetzen, J. L. and H. Wang (2014). "The contributions of transposable elements to the structure, function, and evolution of plant genomes." Annu Rev Plant Biol **65**: 505-530.
Bentley, D. R., S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes, H. R. Bignell, J. M. Boutell, J. Bryant, R. J. Carter, R. Keira Cheetham, A. J. Cox, D. J. Ellis, M. R. Flatbush, N. A. Gormley, S. J. Humphray, L. J. Irving, M. S. Karbelashvili, S. M. Kirk, H. Li, X. Liu, K. S. Maisinger, L. J. Murray, B. Obradovic, T. Ost, M. L. Parkinson, M. R. Pratt, I. M. J. Rasolonjatovo, M. T. Reed, R. Rigatti, C. Rodighiero, M. T. Ross, A. Sabot, S. V. Sankar, A. Scally, G. P. Schroth, M. E. Smith, V. P. Smith, A. Spiridou, P. E. Torrance, S. S. Tzonev, E. H. Vermaas, K. Walter, X. Wu, L. Zhang, M. D. Alam, C. Anastasi, I. C. Aniebo, D. M. D. Bailey, I. R. Bancarz, S. Banerjee, S. G. Barbour, P. A. Baybayan, V. A. Benoit, K. F. Benson, C. Bevis, P. J. Black, A. Boodhun, J. S. Brennan, J. A. Bridgham, R. C. Brown, A. A. Brown, D. H. Buermann, A. A. Bundu, J. C. Burrows, N. P. Carter, N. Castillo, M. Chiara E. Catenazzi, S. Chang, R. Neil Cooley, N. R. Crake, O. O. Dada, K. D. Diakoumakos, B. Dominguez-Fernandez, D. J. Earnshaw, U. C. Egbujor, D. W. Elmore, S. S. Etchin, M. R. Ewan, M. Fedurco, L. J. Fraser, K. V. Fuentes Fajardo, W. Scott Furey, D. George, K. J. Gietzen, C. P. Goddard, G. S. Golda, P. A. Granieri, D. E. Green, D. L. Gustafson, N. F. Hansen, K. Harnish, C. D. Haudenschild, N. I. Heyer, M. M. Hims, J. T. Ho, A. M. Horgan, K. Hoschler, S. Hurwitz, D. V. Ivanov, M. Q. Johnson, T. James, T. A. Huw Jones, G.-D. Kang, T. H. Kerelska, A. D. Kersey, I. Khrebtukova, A. P. Kindwall, Z. Kingsbury, P. I. Kokko-Gonzales, A. Kumar, M. A. Laurent, C. T. Lawley, S. E. Lee, X. Lee, A. K. Liao, J. A. Loch, M. Lok, S. Luo, R. M. Mammen, J. W. Martin, P. G. McCauley, P. McNitt, P. Mehta, K. W. Moon, J. W. Mullens, T. Newington, Z. Ning, B. Ling Ng, S. M. Novo, M. J. O'Neill, M. A. Osborne, A. Osnowski, O. Ostadan, L. L. Paraschos, L. Pickering, A. C. Pike, A. C. Pike, D. Chris Pinkard, D. P. Pliskin, J. Podhasky, V. J. Quijano, C. Raczy, V. H. Rae, S. R. Rawlings, A. Chiva Rodriguez, P. M. Roe, J. Rogers, M. C. Rogert Bacigalupo, N. Romanov, A. Romieu, R. K. Roth, N. J. Rourke, S. T. Ruediger, E. Rusman, R. M. Sanches-Kuiper, M. R. Schenker, J. M. Seoane, R. J. Shaw, M. K. Shiver, S. W. Short, N. L. Sizto, J. P. Sluis, M. A. Smith, J. Ernest Sohna Sohna, E. J. Spence, K. Stevens, N. Sutton, L. Szajkowski, C. L. Tregidgo, G. Turcatti, S. vandeVondele, Y. Verhovsky, S. M. Virk, S. Wakelin, G. C. Walcott, J. Wang, G. J. Worsley, J. Yan, L. Yau, M. Zuerlein, J. Rogers, J. C. Mullikin, M. E. Hurles, N. J. McCooke, J. S. West, F. L. Oaks, P. L. Lundberg, D. Klenerman, R. Durbin and A. J. Smith (2008). "Accurate whole human genome sequencing using reversible terminator chemistry." Nature **456**(7218): 53-59.

Betekhtin, A., G. Jenkins and R. Hasterok (2014). "Reconstructing the Evolution of *Brachypodium* Genomes Using Comparative Chromosome Painting." PLoS One **9**(12): e115108.
Bevan, M. W., D. F. Garvin and J. P. Vogel (2010). "Brachypodium distachyon genomics for sustainable food and fuel production." Current Opinion in Biotechnology **21**(2): 211-217.
Bhullar, R., R. Nagarajan, H. Bennypaul, G. K. Sidhu, G. Sidhu, S. Rustgi, D. von Wettstein and K. S. Gill (2014). "Silencing of a metaphase I-specific gene results in a phenotype similar to that of the Pairing homeologous 1 (Ph1) gene mutations." Proceedings of the National Academy of Sciences of the United States of America **111**(39): 14187-14192.
Bikard, D., D. Patel, C. Le Metté, V. Giorgi, C. Camilleri, M. J. Bennett and O. Loudet (2009). "Divergent Evolution of Duplicate Genes Leads to Genetic Incompatibilities Within A. thaliana." Science **323**(5914): 623-626.
Birchler, J. A. (2013). "Aneuploidy in plants and flies: the origin of studies of genomic imbalance." Semin Cell Dev Biol **24**(4): 315-319.
Birchler, J. A. (2014). "Facts and artifacts in studies of gene expression in aneuploids and sex chromosomes." Chromosoma **123**(5): 459-469.
Birchler, J. A. and R. A. Veitia (2014). "The Gene Balance Hypothesis: dosage effects in plants." Methods Mol Biol **1112**: 25-32.
Blakeslee, A. F. (1934). "New Jimson weeds from old chromosomes." Journal of Heredity **25**(3): 81-108.
Blakeslee, A. F. and A. G. Avery (1937). "Conchicine and double diploids." Journal of Heredity **28**(12): 411-412.
Boff, T. and M. T. Schifino-Wittmann (2003). "Segmental allopolyploidy and paleopolyploidy in species of Leucaena benth: evidence from meiotic behaviour analysis." Hereditas **138**(1): 27-35.
Bomblies, K., J. Lempe, P. Epple, N. Warthmann, C. Lanz, J. L. Dangl and D. Weigel (2007). "Autoimmune Response as a Mechanism for a Dobzhansky-Muller-Type Incompatibility Syndrome in Plants." PLoS Biol **5**(9): e236.
Bonasio, R., S. Tu and D. Reinberg (2010). "Molecular signals of epigenetic states." Science **330**(6004): 612-616.
Borges, F. and R. A. Martienssen (2015). "The expanding world of small RNAs in plants." Nat Rev Mol Cell Biol **16**(12): 727-741.
Bouchenak-Khelladi, Y., J. A. Slingsby, G. A. Verboom and W. J. Bond (2014). "Diversification of C4 grasses (Poaceae) does not coincide with their ecological dominance." American Journal of Botany **101**(2): 300-307.
Bretagnolle, F. and J. D. Thompson (1995). "Gametes with the somatic chromosome number: mechanisms of their formation and role in the evolution of autopolyploid plants." New Phytologist **129**(1): 1-22.
Brkljacic, J., E. Grotewold, R. Scholl, T. Mockler, D. F. Garvin, P. Vain, T. Brutnell, R. Sibout, M. Bevan, H. Budak, A. L. Caicedo, C. Gao, Y. Gu, S. P. Hazen, B. F. Holt, 3rd, S. Y. Hong, M. Jordan, A. J. Manzaneda, T. Mitchell-Olds, K. Mochida, L. A. Mur, C. M. Park, J. Sedbrook, M. Watt, S. J. Zheng and J. P. Vogel (2011). "*Brachypodium* as a model for the grasses: today and the future." Plant Physiol **157**(1): 3-13.
Brownfield, L. and C. Köhler (2011). "Unreduced gamete formation in plants: mechanisms and prospects." Journal of Experimental Botany **62**(5): 1659-1668.
Buermans, H. P. J. and J. T. den Dunnen (2014). "Next generation sequencing technology: Advances and applications." Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease **1842**(10): 1932-1941.
Buggs, R. J., A. N. Doust, J. A. Tate, J. Koh, K. Soltis, F. A. Feltus, A. H. Paterson, P. S. Soltis and D. E. Soltis (2009). "Gene loss and silencing in Tragopogon miscellus (Asteraceae): comparison of natural and synthetic allotetraploids." Heredity (Edinb) **103**(1): 73-81.

Burkart-Waco, D., C. Josefsson, B. Dilkes, N. Kozloff, O. Torjek, R. Meyer, T. Altmann and L. Comai (2012). "Hybrid Incompatibility in Arabidopsis Is Determined by a Multiple-Locus Genetic Network." Plant Physiology **158**(2): 801-812.
Catalan, P., B. Chalhoub, V. Chochois, D. F. Garvin, R. Hasterok, A. J. Manzaneda, L. A. Mur, N. Pecchioni, S. K. Rasmussen, J. P. Vogel and A. Voxeur (2014). "Update on the genomics and basic biology of *Brachypodium*: International *Brachypodium* Initiative (IBI)." Trends Plant Sci **19**(7): 414-418.
Catalan, P., D. López-Álvarez, C. Bellosta and L. and Villar (2016). "Updated taxonomic descriptions, iconography, and habitat preferences of Brachypodium distachyon, B. stacei, and B. hybridum (Poaceae)." Anales del Jardín Botánico de Madrid **73**(1).
Catalan, P., D. López-Álvarez, A. Díaz-Pérez, R. Sancho and M. L. López-Herránz (2015). "Phylogeny and Evolution of the Genus Brachypodium."
Catalan, P., D. López-Álvarez, A. Díaz-Pérez, R. Sancho and M. L. López-Herránz (2016). Phylogeny and Evolution of the Genus Brachypodium. Genetics and Genomics of Brachypodium. P. J. Vogel. Cham, Springer International Publishing**:** 9-38.
Catalan, P., J. Muller, R. Hasterok, G. Jenkins, L. A. Mur, T. Langdon, A. Betekhtin, D. Siwinska, M. Pimentel and D. Lopez-Alvarez (2012). "Evolution and taxonomic split of the model grass *Brachypodium distachyon*." Ann Bot **109**(2): 385-405.
Chagué, V., J. Just, I. Mestiri, S. Balzergue, A. M. Tanguy, C. Huneau, V. Huteau, H. Belcram, O. Coriton, J. Jahier and B. Chalhoub (2010). "Genome-wide gene expression changes in genetically stable synthetic and natural wheat allohexaploids." New Phytol **187**(4): 1181-1194.
Chalabi, S. (2014). Caractérisation de la reprogrammation de l'expression des gènes chez les blés allopolyploïdes, l'Université d'Evry Val d'Essonne.
Chalhoub, B., F. Denoeud, S. Liu and et al. (2014). "Plant genetics. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome." Science **345**(6199): 950-953.
Chalupska, D., H. Y. Lee, J. D. Faris, A. Evrard, B. Chalhoub, R. Haselkorn and P. Gornicki (2008). "Acc homoeoloci and the evolution of wheat genomes." Proceedings of the National Academy of Sciences **105**(28): 9691-9696.
Chantret, N., J. Salse, F. Sabot, S. Rahman, A. Bellec, B. Laubin, I. Dubois, C. Dossat, P. Sourdille, P. Joudrier, M.-F. Gautier, L. Cattolico, M. Beckert, S. Aubourg, J. Weissenbach, M. Caboche, M. Bernard, P. Leroy and B. Chalhoub (2005). "Molecular Basis of Evolutionary Events That Shaped the Hardness Locus in Diploid and Polyploid Wheat Species (Triticum and Aegilops)." The Plant Cell **17**(4): 1033-1045.
Charles, M., H. Belcram, J. Just, C. Huneau, A. Viollet, A. Couloux, B. Segurens, M. Carter, V. Huteau, O. Coriton, R. Appels, S. Samain and B. Chalhoub (2008). "Dynamics and differential proliferation of transposable elements during the evolution of the B and A genomes of wheat." Genetics **180**(2): 1071-1086.
Charles, M., H. Tang, H. Belcram, A. Paterson, P. Gornicki and B. Chalhoub (2009). "Sixty million years in evolution of soft grain trait in grasses: emergence of the softness locus in the common ancestor of Pooideae and Ehrhartoideae, after their divergence from Panicoideae." Mol Biol Evol **26**(7): 1651-1661.
Chelaifa, H., V. Chagué, S. Chalabi, I. Mestiri, D. Arnaud, D. Deffains, Y. Lu, H. Belcram, V. Huteau, J. Chiquet, O. Coriton, J. Just, J. Jahier and B. Chalhoub (2013). "Prevalence of gene expression additivity in genetically stable wheat allohexaploids." New Phytol **197**(3): 730-736.
Chelaifa, H., A. Monnier and M. Ainouche (2010). "Transcriptomic changes following recent natural hybridization and allopolyploidy in the salt marsh species Spartina x townsendii and Spartina anglica (Poaceae)." New Phytol **186**(1): 161-174.
Chen, X. (2010). "Small RNAs – secrets and surprises of the genome." The Plant Journal **61**(6): 941-958.
Chen, Z. J. (2007). "Genetic and Epigenetic Mechanisms for Gene Expression and Phenotypic Variation in Plant Polyploids." Annual Review of Plant Biology **58**(1): 377-406.

Chen, Z. J. (2010). "Molecular mechanisms of polyploidy and hybrid vigor." Trends Plant Sci **15**(2): 57-71.
Chen, Z. J. and Z. Ni (2006). "Mechanisms of genomic rearrangements and gene expression changes in plant polyploids." Bioessays **28**(3): 240-252.
Chen, Z. J. and L. Tian (2007). "Roles of dynamic and reversible histone acetylation in plant development and polyploidy." Biochim Biophys Acta **1769**(5-6): 295-307.
Chester, M., A. R. Leitch, P. S. Soltis and D. E. Soltis (2010). "Review of the Application of Modern Cytogenetic Methods (FISH/GISH) to the Study of Reticulation (Polyploidy/Hybridisation)." Genes (Basel) **1**(2): 166-192.
Cifuentes, M., F. Eber, M. O. Lucas, M. Lode, A. M. Chevre and E. Jenczewski (2010). "Repeated polyploidy drove different levels of crossover suppression between homoeologous chromosomes in Brassica napus allohaploids." Plant Cell **22**(7): 2265-2276.
Clark, S. J., A. Statham, C. Stirzaker, P. L. Molloy and M. Frommer (2006). "DNA methylation: bisulphite modification and analysis." Nature protocols **1**(5): 2353-2364.
Colas, I., P. Shaw, P. Prieto, M. Wanous, W. Spielmeyer, R. Mago and G. Moore (2008). "Effective chromosome pairing requires chromatin remodeling at the onset of meiosis." Proc Natl Acad Sci U S A **105**(16): 6075-6080.
Comai, L. (2000). "Genetic and epigenetic interactions in allopolyploid plants." Plant Mol Biol **43**(2-3): 387-399.
Comai, L. (2005). "The advantages and disadvantages of being polyploid." Nat Rev Genet **6**(11): 836-846.
Comai, L., A. P. Tyagi and M. A. Lysak (2003). "FISH analysis of meiosis in Arabidopsis allopolyploids." Chromosome Res **11**(3): 217-226.
Compton, D. A. (2011). "Mechanisms of aneuploidy." Curr Opin Cell Biol **23**(1): 109-113.
Conesa, A., P. Madrigal, S. Tarazona, D. Gomez-Cabrero, A. Cervera, A. McPherson, M. W. Szcześniak, D. J. Gaffney, L. L. Elo, X. Zhang and A. Mortazavi (2016). "A survey of best practices for RNA-seq data analysis." Genome Biology **17**(1): 1-19.
Consortium, T. P. G. S. (2011). "Genome sequence and analysis of the tuber crop potato." Nature **475**(7355): 189-195.
Cuenca, J., P. Aleza, J. Juárez, A. García-Lor, Y. Froelicher, L. Navarro and P. Ollitrault (2015). "Maximum-likelihood method identifies meiotic restitution mechanism from heterozygosity transmission of centromeric loci: application in citrus." Scientific Reports **5**: 9897.
Cutter, A. D. (2012). "The polymorphic prelude to Bateson-Dobzhansky-Muller incompatibilities." Trends Ecol Evol **27**(4): 209-218.
Dagne, K. (2001). "Meiotic properties of induced autopolyploid Guizotia abyssinica (Lf) Cass." Journal of genetics and breadding **55**(1): 11-16.
Dai, F., Z. Wang, G. Luo and C. Tang (2015). "Phenotypic and Transcriptomic Analyses of Autotetraploid and Diploid Mulberry (Morus alba L.)." Int J Mol Sci **16**(9): 22938-22956.
De Storme, N. and D. Geelen (2013). "Sexual polyploidization in plants – cytological mechanisms and molecular regulation." The New Phytologist **198**(3): 670-684.
del Pozo, J. C. and E. Ramirez-Parra (2015). "Whole genome duplications in plants: an overview from Arabidopsis." Journal of Experimental Botany.
Des Marais, D. L. and M. D. Rausher (2008). "Escape from adaptive conflict after duplication in an anthocyanin pathway gene." Nature **454**(7205): 762-765.
Devos, K. M., J. K. Brown and J. L. Bennetzen (2002). "Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis." Genome Res **12**(7): 1075-1079.
Dillies, M.-A., A. Rau, J. Aubert, C. Hennequet-Antier, M. Jeanmougin, N. Servant, C. Keime, G. Marot, D. Castel, J. Estelle, G. Guernec, B. Jagla, L. Jouneau, D. Laloë, C. Le Gall, B. Schaëffer, S. Le Crom, M. Guedj and F. Jaffrézic (2012). "A comprehensive evaluation of normalization

methods for Illumina high-throughput RNA sequencing data analysis." Briefings in Bioinformatics.
Dinh-Thi, V.-H., I. L. Clainche, O. Coriton, J. Jahier and B. Chalhoub (2016). "Synthesis and evaluation of autotetraploids from Brachypodium distachyon and B. stacei species." Plant Systematics and Evolution **Submitted**.
Dinh-Thi, V.-H., O. Coriton, I. L. Clainche, D. Arnaud, S. P. Gordon, G. Linc, P. Catalan, R. Hasterok, J. P. Vogel, J. Jahier and B. Chalhoub (2016). "Recreating Stable Brachypodium hybridum Allopolyploids by Uniting Divergent Genomes of B. distachyon and B. stacei that Contrast in Chromosome Evolution, Number and Size." PLoS One **accepted**.
Dolled-Filhart, M. P., M. Lee, C.-w. Ou-yang, R. R. Haraksingh and J. C.-H. Lin (2013). "Computational and Bioinformatics Frameworks for Next-Generation Whole Exome and Genome Sequencing." The Scientific World Journal **2013**: 730210.
Doyle, J. J., L. E. Flagel, A. H. Paterson, R. A. Rapp, D. E. Soltis, P. S. Soltis and J. F. Wendel (2008). "Evolutionary genetics of genome merger and doubling in plants." Annu Rev Genet **42**: 443-461.
Draper, J., L. A. Mur, G. Jenkins, G. C. Ghosh-Biswas, P. Bablak, R. Hasterok and A. P. Routledge (2001). "*Brachypodium distachyon*. A new model system for functional genomics in grasses." Plant Physiol **127**(4): 1539-1555.
Dupont, C., D. R. Armant and C. A. Brenner (2009). "Epigenetics: definition, mechanisms and clinical perspective." Semin Reprod Med **27**(5): 351-357.
Dykhuizen, D. and D. L. Hartl (1980). "Selective neutrality of 6PGD allozymes in E. coli and the effects of genetic background." Genetics **96**(4): 801-817.
Dykhuizen, D. E. and D. L. Hartl (1983). "Selection in chemostats." Microbiological Reviews **47**(2): 150.
Eckardt, N. A. (2010). "A Double Lock on Polyploidy-Associated Epigenetic Gene Silencing." The Plant Cell **22**(1): 3-3.
Edger, P. P., H. M. Heidel-Fischer, M. Bekaert, J. Rota, G. Glockner, A. E. Platts, D. G. Heckel, J. P. Der, E. K. Wafula, M. Tang, J. A. Hofberger, A. Smithson, J. C. Hall, M. Blanchette, T. E. Bureau, S. I. Wright, C. W. dePamphilis, M. Eric Schranz, M. S. Barker, G. C. Conant, N. Wahlberg, H. Vogel, J. C. Pires and C. W. Wheat (2015). "The butterfly plant arms-race escalated by gene and genome duplications." Proc Natl Acad Sci U S A **112**(27): 8362-8366.
Egan, A. N., J. Schlueter and D. M. Spooner (2012). "Applications of next-generation sequencing in plant biology." Am J Bot **99**(2): 175-185.
Eichten, S. R., R. J. Schmitz and N. M. Springer (2014). "Epigenetics: Beyond Chromatin Modifications and Complex Genetic Regulation." Plant Physiol **165**(3): 933-947.
Eilam, T., Y. Anikster, E. Millet, J. Manisterski and M. Feldman (2009). "Genome size in natural and synthetic autopolyploids and in a natural segmental allopolyploid of several Triticeae species." Genome **52**(3): 275-285.
Estep, M. C., M. R. McKain, D. Vela Diaz, J. Zhong, J. G. Hodge, T. R. Hodkinson, D. J. Layton, S. T. Malcomber, R. Pasquet and E. A. Kellogg (2014). "Allopolyploidy, diversification, and the Miocene grassland expansion." Proceedings of the National Academy of Sciences **111**(42): 15149-15154.
Fabian, M. R., N. Sonenberg and W. Filipowicz (2010). "Regulation of mRNA translation and stability by microRNAs." Annu Rev Biochem **79**: 351-379.
Fang, Z., J. Martin and Z. Wang (2012). "Statistical methods for identifying differentially expressed genes in RNA-Seq experiments." Cell Biosci **2**(1): 26.
Feldman, M. (1995). "Wheats: Triticum spp.(Gramineae-Triticinae)." Evolution of Crop Plants. NW Simmonds, ed.
Feldman, M. (2000). "Origin of cultivated wheat." The World Wheat Book, A history of wheat breeding.

Feng, S., S. E. Jacobsen and W. Reik (2010). "Epigenetic reprogramming in plant and animal development." Science (New York, N.Y.) **330**(6004): 622-627.
Feng, Y., Y. Zhang, C. Ying, D. Wang and C. Du (2015). "Nanopore-based Fourth-generation DNA Sequencing Technology." Genomics, Proteomics & Bioinformatics **13**(1): 4-16.
Fishman, L. and J. H. Willis (2001). "Evidence for Dobzhansky-Muller incompatibilites contributing to the sterility of hybrid between Mimulus Guttatus and M. nasutus." Evolution **55**(10): 1932-1942.
Flagel, L., J. Udall, D. Nettleton and J. Wendel (2008). "Duplicate gene expression in allopolyploid Gossypium reveals two temporally distinct phases of expression evolution." BMC Biol **6**: 16.
Flagel, L. E. and J. F. Wendel (2009). "Gene duplication and evolutionary novelty in plants." New Phytol **183**(3): 557-564.
Flusberg, B. A., D. R. Webster, J. H. Lee, K. J. Travers, E. C. Olivares, T. A. Clark, J. Korlach and S. W. Turner (2010). "Direct detection of DNA methylation during single-molecule, real-time sequencing." Nat Meth **7**(6): 461-465.
Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. L. Yan and J. Postlethwait (1999). "Preservation of duplicate genes by complementary, degenerative mutations." Genetics **151**(4): 1531-1545.
Fuchs, J., D. Demidov, A. Houben and I. Schubert (2006). "Chromosomal histone modification patterns--from conservation to diversity." Trends Plant Sci **11**(4): 199-208.
Gaeta, R. T. and J. Chris Pires (2010). "Homoeologous recombination in allopolyploids: the polyploid ratchet." New Phytol **186**(1): 18-28.
Gaeta, R. T., S. Y. Yoo, J. C. Pires, R. W. Doerge, Z. J. Chen and T. C. Osborn (2009). "Analysis of gene expression in resynthesized Brassica napus Allopolyploids using arabidopsis 70mer oligo microarrays." PLoS One **4**(3): e4760.
Gaither, T. L., S. L. Merrett, M. J. Pun and K. C. Scott (2014). "Centromeric barrier disruption leads to mitotic defects in Schizosaccharomyces pombe." G3 (Bethesda) **4**(4): 633-642.
Garsmeur, O., J. C. Schnable, A. Almeida, C. Jourda, A. D'Hont and M. Freeling (2014). "Two evolutionarily distinct classes of paleopolyploidy." Mol Biol Evol **31**(2): 448-454.
Garvin, D. F. (2016). *Brachypodium distachyon* Genetic Resources. Genetics and Genomics of Brachypodium. P. J. Vogel. Cham, Springer International Publishing**:** 183-195.
Gaut, B. S. and J. F. Doebley (1997). "DNA sequence evidence for the segmental allotetraploid origin of maize." Proc Natl Acad Sci U S A **94**(13): 6809-6814.
Ghildiyal, M. and P. D. Zamore (2009). "Small silencing RNAs: an expanding universe." Nat Rev Genet **10**(2): 94-108.
Gierliński, M., C. Cole, P. Schofield, N. J. Schurch, A. Sherstnev, V. Singh, N. Wrobel, K. Gharbi, G. Simpson, T. Owen-Hughes, M. Blaxter and G. J. Barton (2015). "Statistical models for RNA-seq data derived from a two-condition 48-replicate experiment." Bioinformatics.
Goff, S. A., D. Ricke, T. H. Lan, G. Presting, R. Wang, M. Dunn, J. Glazebrook, A. Sessions, P. Oeller, H. Varma, D. Hadley, D. Hutchison, C. Martin, F. Katagiri, B. M. Lange, T. Moughamer, Y. Xia, P. Budworth, J. Zhong, T. Miguel, U. Paszkowski, S. Zhang, M. Colbert, W. L. Sun, L. Chen, B. Cooper, S. Park, T. C. Wood, L. Mao, P. Quail, R. Wing, R. Dean, Y. Yu, A. Zharkikh, R. Shen, S. Sahasrabudhe, A. Thomas, R. Cannings, A. Gutin, D. Pruss, J. Reid, S. Tavtigian, J. Mitchell, G. Eldredge, T. Scholl, R. M. Miller, S. Bhatnagar, N. Adey, T. Rubano, N. Tusneem, R. Robinson, J. Feldhaus, T. Macalma, A. Oliphant and S. Briggs (2002). "A draft sequence of the rice genome (Oryza sativa L. ssp. japonica)." Science **296**(5565): 92-100.
Gómez-Rodríguez, V. M., B. Rodríguez-Garay and R. Barba-Gonzalez (2012). "Meiotic restitution mechanisms involved in the formation of 2n pollen in Agave tequilana Weber and Agave angustifolia Haw." SpringerPlus **1**: 17.
Goodwin, S., J. D. McPherson and W. R. McCombie (2016). "Coming of age: ten years of next-generation sequencing technologies." Nat Rev Genet **17**(6): 333-351.

Gordon, S. P., L. Liu and J. P. Vogel (2016). The Genus Brachypodium as a Model for Perenniality and Polyploidy. Genetics and Genomics of Brachypodium. P. J. Vogel. Cham, Springer International Publishing**:** 313-325.
Gordon, S. P., H. Priest, D. L. Des Marais, W. Schackwitz, M. Figueroa, J. Martin, J. N. Bragg, L. Tyler, C. R. Lee, D. Bryant, W. Wang, J. Messing, A. J. Manzaneda, K. Barry, D. F. Garvin, H. Budak, M. Tuna, T. Mitchell-Olds, W. F. Pfender, T. E. Juenger, T. C. Mockler and J. P. Vogel (2014). "Genome diversity in Brachypodium distachyon: deep sequencing of highly diverse inbred lines." Plant J **79**(3): 361-374.
Grandont, L., E. Jenczewski and A. Lloyd (2013). "Meiosis and its deviations in polyploid plants." Cytogenet Genome Res **140**(2-4): 171-184.
Grewal, S. S. and S. C. R. Elgin (2007). "Transcription and RNAi in the formation of heterochromatin." Nature **447**(7143): 399-406.
Griffiths, A. J., W. M. Gelbart, J. H. Miller and R. C. Lewontin (1999). "Modern genetic analysis."
Griffiths, S., R. Sharp, T. N. Foote, I. Bertin, M. Wanous, S. Reader, I. Colas and G. Moore (2006). "Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat." Nature **439**(7077): 749-752.
Groszmann, M., I. K. Greaves, Z. I. Albertyn, G. N. Scofield, W. J. Peacock and E. S. Dennis (2011). "Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor." Proceedings of the National Academy of Sciences **108**(6): 2617-2622.
Groszmann, M., I. K. Greaves, R. Fujimoto, W. J. Peacock and E. S. Dennis (2013). "The role of epigenetics in hybrid vigour." Trends Genet **29**(12): 684-690.
Ha, M., J. Lu, L. Tian, V. Ramachandran, K. D. Kasschau, E. J. Chapman, J. C. Carrington, X. Chen, X. J. Wang and Z. J. Chen (2009). "Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids." Proc Natl Acad Sci U S A **106**(42): 17835-17840.
Ha, M., D. W. Ng, W. H. Li and Z. J. Chen (2011). "Coordinated histone modifications are associated with gene expression variation within and between species." Genome Res **21**(4): 590-598.
Hasterok, R., J. Draper and G. Jenkins (2004). "Laying the cytotaxonomic foundations of a new model grass, *Brachypodium distachyon* (L.) Beauv." Chromosome Res **12**(4): 397-403.
Hasterok, R., A. Marasek, I. S. Donnison, I. Armstead, A. Thomas, I. P. King, E. Wolny, D. Idziak, J. Draper and G. Jenkins (2006). "Alignment of the genomes of *Brachypodium distachyon* and temperate cereals and grasses using bacterial artificial chromosome landing with fluorescence in situ hybridization." Genetics **173**(1): 349-362.
Hegarty, M. J., G. L. Barker, I. D. Wilson, R. J. Abbott, K. J. Edwards and S. J. Hiscock (2006). "Transcriptome shock after interspecific hybridization in senecio is ameliorated by genome duplication." Curr Biol **16**(16): 1652-1659.
Henderson, I. R. and S. E. Jacobsen (2007). "Epigenetic inheritance in plants." Nature **447**(7143): 418-424.
Henry, I. M., B. P. Dilkes, E. S. Miller, D. Burkart-Waco and L. Comai (2010). "Phenotypic consequences of aneuploidy in Arabidopsis thaliana." Genetics **186**(4): 1231-1245.
Higgins, J., A. Magusin, M. Trick, F. Fraser and I. Bancroft (2012). "Use of mRNA-seq to discriminate contributions to the transcriptome from the constituent genomes of the polyploid crop species Brassica napus." BMC Genomics **13**(1): 1-14.
Hodkinson, B. P. and E. A. Grice (2015). "Next-Generation Sequencing: A Review of Technologies and Tools for Wound Microbiome Research." Advances in Wound Care **4**(1): 50-58.

Hollister, J. D., B. J. Arnold, E. Svedin, K. S. Xue, B. P. Dilkes and K. Bomblies (2012). "Genetic adaptation associated with genome-doubling in autotetraploid Arabidopsis arenosa." PLoS Genet **8**(12): e1003093.
Hu, G., J. S. Hawkins, C. E. Grover and J. F. Wendel (2010). "The history and disposition of transposable elements in polyploid Gossypium." Genome **53**(8): 599-607.
Huang, S., A. Sirikhachornkit, X. Su, J. Faris, B. Gill, R. Haselkorn and P. Gornicki (2002). "Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum*/*Aegilops* complex and the evolutionary history of polyploid wheat." Proceedings of the National Academy of Sciences **99**(12): 8133-8138.
Hufton, A. L. and G. Panopoulou (2009). "Polyploidy and genome restructuring: a variety of outcomes." Current opinion in genetics & development **19**(6): 600-606.
Hulten, M. A., S. Patel, J. Jonasson and E. Iwarsson (2010). "On the origin of the maternal age effect in trisomy 21 Down syndrome: the Oocyte Mosaicism Selection model." Reproduction **139**(1): 1-9.
Human Genome Sequencing, C. (2004). "Finishing the euchromatic sequence of the human genome." Nature **431**(7011): 931-945.
IBI, T. (2010). "Genome sequencing and analysis of the model grass *Brachypodium distachyon*." Nature **463**(7282): 763-768.
Idziak, D. and R. Hasterok (2008). "Cytogenetic evidence of nucleolar dominance in allotetraploid species of *Brachypodium*." Genome **51**(5): 387-391.
Initiative, T. A. G. (2000). "Analysis of the genome sequence of the flowering plant Arabidopsis thaliana." Nature **408**(6814): 796-815.
Jackson, S. and Z. J. Chen (2010). "Genomic and expression plasticity of polyploidy." Current Opinion in Plant Biology **13**(2): 153-159.
Jaillon, O., J. M. Aury, B. Noel, A. Policriti, C. Clepet, A. Casagrande, N. Choisne, S. Aubourg, N. Vitulo, C. Jubin, A. Vezzi, F. Legeai, P. Hugueney, C. Dasilva, D. Horner, E. Mica, D. Jublot, J. Poulain, C. Bruyere, A. Billault, B. Segurens, M. Gouyvenoux, E. Ugarte, F. Cattonaro, V. Anthouard, V. Vico, C. Del Fabbro, M. Alaux, G. Di Gaspero, V. Dumas, N. Felice, S. Paillard, I. Juman, M. Moroldo, S. Scalabrin, A. Canaguier, I. Le Clainche, G. Malacrida, E. Durand, G. Pesole, V. Laucou, P. Chatelet, D. Merdinoglu, M. Delledonne, M. Pezzotti, A. Lecharny, C. Scarpelli, F. Artiguenave, M. E. Pe, G. Valle, M. Morgante, M. Caboche, A. F. Adam-Blondon, J. Weissenbach, F. Quetier and P. Wincker (2007). "The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla." Nature **449**(7161): 463-467.
Jauhar, P. P. (2007). "Meiotic restitution in wheat polyhaploids (amphihaploids): a potent evolutionary force." J Hered **98**(2): 188-193.
Jenczewski, E., F. Eber, A. Grimaud, S. Huet, M. O. Lucas, H. Monod and A. M. Chevre (2003). "PrBn, a major gene controlling homeologous pairing in oilseed rape (Brassica napus) haploids." Genetics **164**(2): 645-653.
Jiao, Y. and A. H. Paterson (2014). "Polyploidy-associated genome modifications during land plant evolution." Philos Trans R Soc Lond B Biol Sci **369**(1648).
Johnson, N. A. (2010). "Hybrid incompatibility genes: remnants of a genomic battlefield?" Trends Genet **26**(7): 317-325.
Jolliffe, I. (2002). Principal component analysis, Wiley Online Library.
Jones, G. H., S. J. Armstrong, A. P. Caryl and F. C. Franklin (2003). "Meiotic chromosome synapsis and recombination in Arabidopsis thaliana; an integration of cytological and molecular approaches." Chromosome Res **11**(3): 205-215.
Jones, P. A. (2012). "Functions of DNA methylation: islands, start sites, gene bodies and beyond." Nat Rev Genet **13**(7): 484-492.
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.

Kenan-Eichler, M., D. Leshkowitz, L. Tal, E. Noor, C. Melamed-Bessudo, M. Feldman and A. A. Levy (2011). "Wheat hybridization and polyploidization results in deregulation of small RNAs." Genetics **188**(2): 263-272.
Khalil, H. B., M. R. Ehdaeivand, Y. Xu, A. Laroche and P. J. Gulick (2015). "Identification and characterization of rye genes not expressed in allohexaploid triticale." BMC Genomics **16**: 281.
Kim, C., H. Guo, W. Kong, R. Chandnani, L.-S. Shuang and A. H. Paterson (2016). "Application of genotyping by sequencing technology to a variety of crop breeding programs." Plant Science **242**: 14-22.
Kimura, M. (1983). "Rare variant alleles in the light of the neutral theory." Molecular Biology and Evolution **1**(1): 84-93.
Kraitshtein, Z., B. Yaakov, V. Khasdan and K. Kashkush (2010). "Genetic and epigenetic dynamics of a retrotransposon after allopolyploidization of wheat." Genetics **186**(3): 801-812.
Krueger, F. and S. R. Andrews (2011). "Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications." Bioinformatics **27**(11): 1571-1572.
Krueger, F., B. Kreck, A. Franke and S. R. Andrews (2012). "DNA methylome analysis using short bisulfite sequencing data." Nat Methods **9**(2): 145-151.
Kumar, S., R. Kumari, V. Sharma and V. Sharma (2013). "Roles, and establishment, maintenance and erasing of the epigenetic cytosine methylation marks in plants." J Genet **92**(3): 629-666.
Kynast, R. G., D. W. Davis, R. L. Phillips and H. W. Rines (2012). "Gamete formation via meiotic nuclear restitution generates fertile amphiploid F1 (oatxmaize) plants." Sex Plant Reprod **25**(2): 111-122.
Lam, J. K. W., M. Y. T. Chow, Y. Zhang and S. W. S. Leung (2015). "siRNA Versus miRNA as Therapeutics for Gene Silencing." Mol Ther Nucleic Acids **4**: e252.
Langmead, B. and S. L. Salzberg (2012). "Fast gapped-read alignment with Bowtie 2." Nat Methods **9**(4): 357-359.
Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg (2009). "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." Genome Biol **10**(3): R25.
Lauressergues, D., J. M. Couzigou, H. S. Clemente, Y. Martinez, C. Dunand, G. Becard and J. P. Combier (2015). "Primary transcripts of microRNAs encode regulatory peptides." Nature **520**(7545): 90-93.
Law, J. A. and S. E. Jacobsen (2010). "Establishing, maintaining and modifying DNA methylation patterns in plants and animals." Nat Rev Genet **11**(3): 204-220.
Lee, H., J. Gurtowski, S. Yoo, M. Nattestad, S. Marcus, S. Goodwin, W. R. McCombie and M. Schatz (2016). "Third-generation sequencing and the future of genomics." bioRxiv.
Lee, S. I. and N. S. Kim (2014). "Transposable elements and genome size variations in plants." Genomics Inform **12**(3): 87-97.
Leitch, A. R. and I. J. Leitch (2008). "Genomic Plasticity and the Diversity of Polyploid Plants." Science **320**(5875): 481-483.
Leitch, I. J. and M. D. Bennett (2004). "Genome downsizing in polyploid plants." Biological Journal of the Linnean Society **82**(4): 651-663.
Leitch, I. J., L. Hanson, K. Y. Lim, A. Kovarik, M. W. Chase, J. J. Clarkson and A. R. Leitch (2008). "The Ups and Downs of Genome Size Evolution in Polyploid Species of Nicotiana (Solanaceae)." Annals of Botany **101**(6): 805-814.
Levin, H. L. and J. V. Moran (2011). "Dynamic interactions between transposable elements and their hosts." Nat Rev Genet **12**(9): 615-627.
Levy, A. A. and M. Feldman (2002). "The impact of polyploidy on grass genome evolution." Plant Physiol **130**(4): 1587-1593.
Li, ZhiHui, Lu, Xiang, Gao, Yun, Liu, ShaoJun, Tao, Min, Xiao, Heng, Qiao, YunQian, Zhang, YaPing, Luo and Jing (2011). Polyploidization and epigenetics. Heidelberg, ALLEMAGNE, Springer.

Li, H. and R. Durbin (2010). "Fast and accurate long-read alignment with Burrows-Wheeler transform." Bioinformatics **26**(5): 589-595.
Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing (2009). "The Sequence Alignment/Map format and SAMtools." Bioinformatics **25**(16): 2078-2079.
Li, P., Y. Piao, H. S. Shon and K. H. Ryu (2015). "Comparing the normalization methods for the differential analysis of Illumina high-throughput RNA-Seq data." BMC Bioinformatics **16**(1): 1-9.
Li, R., C. Yu, Y. Li, T. W. Lam, S. M. Yiu, K. Kristiansen and J. Wang (2009). "SOAP2: an improved ultrafast tool for short read alignment." Bioinformatics **25**(15): 1966-1967.
Li, Y. and T. O. Tollefsbol (2011). "DNA methylation detection: bisulfite genomic sequencing analysis." Methods Mol Biol **791**: 11-21.
Lim, K. B., M. S. Ramanna, J. H. de Jong, E. Jacobsen and J. M. van Tuyl (2001). "Indeterminate meiotic restitution (IMR): a novel type of meiotic nuclear restitution mechanism detected in interspecific lily hybrids by GISH." TAG Theoretical and Applied Genetics **103**(2-3): 219-230.
Lim, K. Y., D. E. Soltis, P. S. Soltis, J. Tate, R. Matyasek, H. Srubarova, A. Kovarik, J. C. Pires, Z. Xiong and A. R. Leitch (2008). "Rapid Chromosome Evolution in Recently Formed Polyploids in Tragopogon (Asteraceae)." PLoS ONE **3**(10): e3353.
Lim, U. and M. A. Song (2012). "Dietary and lifestyle factors of DNA methylation." Methods Mol Biol **863**: 359-376.
Lister, R., R. C. O'Malley, J. Tonti-Filippini, B. D. Gregory, C. C. Berry, A. H. Millar and J. R. Ecker (2008). "Highly integrated single-base resolution maps of the epigenome in Arabidopsis." Cell **133**(3): 523-536.
Liu, B. and J. F. Wendel (2000). "Retrotransposon activation followed by rapid repression in introgressed rice plants." Genome **43**(5): 874-880.
Liu, L., Y. Li, S. Li, N. Hu, Y. He, R. Pong, D. Lin, L. Lu and M. Law (2012). "Comparison of next-generation sequencing systems." BioMed Research International **2012**.
Liu, Z., K. Adamczyk, M. Manzanares-Dauleux, F. Eber, M. O. Lucas, R. Delourme, A. M. Chevre and E. Jenczewski (2006). "Mapping PrBn and other quantitative trait loci responsible for the control of homeologous chromosome pairing in oilseed rape (Brassica napus L.) haploids." Genetics **174**(3): 1583-1596.
Lloyd, A. and K. Bomblies (2016). "Meiosis in autopolyploid and allopolyploid Arabidopsis." Current Opinion in Plant Biology **30**: 116-122.
Lopez-Alvarez, D., M. L. Lopez-Herranz, A. Betekhtin and P. Catalan (2012). "A DNA barcoding method to discriminate between the model plant *Brachypodium distachyon* and its close relatives *B. stacei* and *B. hybridum* (Poaceae)." PLoS One **7**(12): e51058.
Lukaszewski, A. J. (1997). "The Development and Meiotic Behavior of Asymmetrical Isochromosomes in Wheat." Genetics **145**(4): 1155-1160.
Lukens, L. N., J. C. Pires, E. Leon, R. Vogelzang, L. Oslach and T. Osborn (2006). "Patterns of Sequence Loss and Cytosine Methylation within a Population of Newly Resynthesized Brassica napus Allopolyploids." Plant Physiology **140**(1): 336-348.
Madlung, A. (2013). "Polyploidy and its effect on evolutionary success: old questions revisited with new tools." Heredity (Edinb) **110**(2): 99-104.
Madlung, A., R. W. Masuelli, B. Watson, S. H. Reynolds, J. Davison and L. Comai (2002). "Remodeling of DNA Methylation and Phenotypic and Transcriptional Changes in Synthetic *Arabidopsis* Allotetraploids." Plant Physiology **129**(2): 733-746.
Madlung, A., A. P. Tyagi, B. Watson, H. Jiang, T. Kagochi, R. W. Doerge, R. Martienssen and L. Comai (2005). "Genomic changes in synthetic *Arabidopsis* polyploids." Plant J **41**(2): 221-230.
Madlung, A. and J. F. Wendel (2013). "Genetic and epigenetic aspects of polyploid evolution in plants." Cytogenet Genome Res **140**(2-4): 270-285.

Madoui, M. A., S. Engelen, C. Cruaud, C. Belser, L. Bertrand, A. Alberti, A. Lemainque, P. Wincker and J. M. Aury (2015). "Genome assembly using Nanopore-guided long and error-free DNA reads." BMC Genomics **16**: 327.
Maheshwari, S. and D. A. Barbash (2011). "The genetics of hybrid incompatibilities." Annu Rev Genet **45**: 331-355.
Manavella, P. A., D. Koenig and D. Weigel (2012). "Plant secondary siRNA production determined by microRNA-duplex structure." Proceedings of the National Academy of Sciences of the United States of America **109**(7): 2461-2466.
Mardis, E. R. (2013). "Next-Generation Sequencing Platforms." Annual Review of Analytical Chemistry **6**(1): 287-303.
Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y.-J. Chen, Z. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. I. Alenquer, T. P. Jarvie, K. B. Jirage, J.-B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. Yu, R. F. Begley and J. M. Rothberg (2005). "Genome sequencing in microfabricated high-density picolitre reactors." Nature **437**(7057): 376-380.
Martienssen, R. A. (2010). "Heterochromatin, small RNA and post-fertilization dysgenesis in allopolyploid and interploid hybrids of *Arabidopsis*." New Phytologist **186**(1): 46-53.
Martinez-Perez, E., P. Shaw and G. Moore (2001). "The Ph1 locus is needed to ensure specific somatic and meiotic centromere association." Nature **411**(6834): 204-207.
Mason, A. S., M. N. Nelson, G. Yan and W. A. Cowling (2011). "Production of viable male unreduced gametes in Brassica interspecific hybrids is genotype specific and stimulated by cold temperatures." BMC Plant Biology **11**(1): 1-13.
Mason, A. S. and J. C. Pires (2015). "Unreduced gametes: meiotic mishap or evolutionary mechanism?" Trends in Genetics **31**(1): 5-10.
Matzke, M. A., M. F. Mette, T. Kanno and A. J. Matzke (2003). "Does the intrinsic instability of aneuploid genomes have a causal role in cancer?" Trends Genet **19**(5): 253-256.
Maxam, A. M. and W. Gilbert (1977). "A new method for sequencing DNA." Proc Natl Acad Sci U S A **74**(2): 560-564.
McClintock, B. (1948). "Mutable loci in maize." Carnegie Inst Wash Year Book **47**: 155-169.
McClintock, B. (1984). "The significance of responses of the genome to challenge." Science **226**(4676): 792-801.
McGinty, R. K. and S. Tan (2015). "Nucleosome structure and function." Chem Rev **115**(6): 2255-2273.
Mestiri, I., V. Chagué, A. M. Tanguy, C. Huneau, V. Huteau, H. Belcram, O. Coriton, B. Chalhoub and J. Jahier (2010). "Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity." New Phytol **186**(1): 86-101.
Metzker, M. L. (2005). "Emerging technologies in DNA sequencing." Genome Res **15**(12): 1767-1776.
Michael, T. P. and S. Jackson (2013). "The First 50 Plant Genomes." Plant Gen. **6**(2): -.
Mittal, P., A. Goyal, R. Yadav, R. Devi and S. Sharma (2012). Phenomenal RNA Interference: From Mechanism to Application, INTECH Open Access Publisher.
Mittal, V. (2004). "Improving the efficiency of RNA interference in mammals." Nat Rev Genet **5**(5): 355-365.
Moore, G. and P. Shaw (2009). "Improving the chances of finding the right partner." Curr Opin Genet Dev **19**(2): 99-104.

Morozova, O. and M. A. Marra (2008). "Applications of next-generation sequencing technologies in functional genomics." Genomics **92**(5): 255-264.
Mortazavi, A., B. A. Williams, K. McCue, L. Schaeffer and B. Wold (2008). "Mapping and quantifying mammalian transcriptomes by RNA-Seq." Nat Meth **5**(7): 621-628.
Moyle, L. C. and T. Nakazato (2008). "Comparative genetics of hybrid incompatibility: sterility in two Solanum species crosses." Genetics **179**(3): 1437-1453.
Munoz-Najar, U. and J. M. Sedivy (2011). "Epigenetic control of aging." Antioxid Redox Signal **14**(2): 241-259.
Mur, L. A., J. Allainguillaume, P. Catalan, R. Hasterok, G. Jenkins, K. Lesniewska, I. Thomas and J. Vogel (2011). "Exploiting the Brachypodium Tool Box in cereal and grass research." New Phytol **191**(2): 334-347.
Nagaoka, S. I., T. J. Hassold and P. A. Hunt (2012). "Human aneuploidy: mechanisms and new insights into an age-old problem." Nat Rev Genet **13**(7): 493-504.
Ng, D. W., C. Zhang, M. Miller, G. Palmer, M. Whiteley, D. Tholl and Z. J. Chen (2011). "cis- and trans-Regulation of miR163 and target genes confers natural variation of secondary metabolites in two Arabidopsis species and their allopolyploids." Plant Cell **23**(5): 1729-1740.
Ni, Z., E. D. Kim, M. Ha, E. Lackey, J. Liu, Y. Zhang, Q. Sun and Z. J. Chen (2009). "Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids." Nature **457**(7227): 327-331.
Nicolas, S. D., M. Leflon, H. Monod, F. Eber, O. Coriton, V. Huteau, A.-M. Chèvre and E. Jenczewski (2009). "Genetic Regulation of Meiotic Cross-Overs between Related Genomes in Brassica napus Haploids and Hybrids." The Plant Cell **21**(2): 373-385.
Ning, Z., A. J. Cox and J. C. Mullikin (2001). "SSAHA: a fast search method for large DNA databases." Genome Res **11**(10): 1725-1729.
Nowak, M. A., M. C. Boerlijst, J. Cooke and J. M. Smith (1997). "Evolution of genetic redundancy." Nature **388**(6638): 167-171.
Ohno, S. (1970). "Evolution by Gene Duplication."
Oshlack, A., M. D. Robinson and M. D. Young (2010). "From RNA-seq reads to differential expression results." Genome Biol **11**(12): 220.
Ouyang, Y. and Q. Zhang (2013). "Understanding reproductive isolation based on the rice model." Annu Rev Plant Biol **64**: 111-135.
Ozkan, H., A. A. Levy and M. Feldman (2001). "Allopolyploidy-Induced Rapid Genome Evolution in the Wheat (Aegilops–Triticum) Group." The Plant Cell **13**(8): 1735-1748.
Ozkan, H., M. Tuna and K. Arumuganathan (2003). "Nonadditive Changes in Genome Size During Allopolyploidization in the Wheat (Aegilops-Triticum) Group." Journal of Heredity **94**(3): 260-264.
Parisod, C., K. Alix, J. Just, M. Petit, V. Sarilar, C. Mhiri, M. Ainouche, B. Chalhoub and M.-A. Grandbastien (2010). "Impact of transposable elements on the organization and function of allopolyploid genomes." New Phytologist **186**(1): 37-45.
Parisod, C., R. Holderegger and C. Brochmann (2010). "Evolutionary consequences of autopolyploidy." New Phytol **186**(1): 5-17.
Paterson, A. H. (2006). "Leafing through the genomes of our major crop plants: strategies for capturing unique information." Nat Rev Genet **7**(3): 174-184.
Paterson, A. H., J. E. Bowers, R. Bruggmann, I. Dubchak, J. Grimwood, H. Gundlach, G. Haberer, U. Hellsten, T. Mitros, A. Poliakov, J. Schmutz, M. Spannagl, H. Tang, X. Wang, T. Wicker, A. K. Bharti, J. Chapman, F. A. Feltus, U. Gowik, I. V. Grigoriev, E. Lyons, C. A. Maher, M. Martis, A. Narechania, R. P. Otillar, B. W. Penning, A. A. Salamov, Y. Wang, L. Zhang, N. C. Carpita, M. Freeling, A. R. Gingle, C. T. Hash, B. Keller, P. Klein, S. Kresovich, M. C. McCann, R. Ming, D. G. Peterson, R. Mehboob ur, D. Ware, P. Westhoff, K. F. Mayer, J. Messing and D. S. Rokhsar (2009). "The Sorghum bicolor genome and the diversification of grasses." Nature **457**(7229): 551-556.

Pérez-de-Castro, A. M., S. Vilanova, J. Cañizares, L. Pascual, J. M. Blanca, M. J. Díez, J. Prohens and B. Picó (2012). "Application of Genomic Tools in Plant Breeding." Current Genomics **13**(3): 179-195.
Petit, M., C. Guidat, J. Daniel, E. Denis, E. Montoriol, Q. T. Bui, K. Y. Lim, A. Kovarik, A. R. Leitch, M. A. Grandbastien and C. Mhiri (2010). "Mobilization of retrotransposons in synthetic allotetraploid tobacco." New Phytologist **186**(1): 135-147.
Pontes, O. and C. S. Pikaard (2008). "siRNA and miRNA processing: new functions for Cajal bodies." Curr Opin Genet Dev **18**(2): 197-203.
Qu, L., J. F. Hancock and J. H. Whallon (1998). "Evolution in an Autopolyploid Group Displaying Predominantly Bivalent Pairing at Meiosis: Genomic Similarity of Diploid Vaccinium darrowi and Autotetraploid V. Corymbosum (Ericaceae)." American Journal of Botany **85**(5): 698-703.
Quail, M. A., I. Kozarewa, F. Smith, A. Scally, P. J. Stephens, R. Durbin, H. Swerdlow and D. J. Turner (2008). "A large genome center's improvements to the Illumina sequencing system." Nature methods **5**(12): 1005-1010.
Quail, M. A., M. Smith, P. Coupland, T. D. Otto, S. R. Harris, T. R. Connor, A. Bertoni, H. P. Swerdlow and Y. Gu (2012). "A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers." BMC Genomics **13**: 341.
Ramsey, J. and T. S. Ramsey (2014). "Ecological studies of polyploidy in the 100 years following its discovery." Philosophical Transactions of the Royal Society B: Biological Sciences **369**(1648): 20130352.
Ramsey, J. and D. W. Schemske (1998). "Pathways, mechanisms, and rates of polyploid formation in flowering plants." Annual Review of Ecology and Systematics **29**: 467-501.
Ramsey, J. and D. W. Schemske (2002). "Neopolyploidy in Flowering Plants." Annual Review of Ecology and Systematics **33**(1): 589-639.
Ranney, T. G. (2006). Polyploidy: From evolution to new plant development. Combined Proceedings International Plant Propagators' Society.
Rapaport, F., R. Khanin, Y. Liang, M. Pirun, A. Krek, P. Zumbo, C. E. Mason, N. D. Socci and D. Betel (2013). "Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data." Genome Biology **14**(9): 1-13.
Rapp, R. A., J. A. Udall and J. F. Wendel (2009). "Genomic expression dominance in allopolyploids." BMC Biol **7**: 18.
Renny-Byfield, S. and J. F. Wendel (2014). "Doubling down on genomes: polyploidy and crop plants." Am J Bot **101**(10): 1711-1725.
Rhoads, A. and K. F. Au (2015). "PacBio Sequencing and Its Applications." Genomics, Proteomics & Bioinformatics **13**(5): 278-289.
Rieseberg, L. H. and J. H. Willis (2007). "Plant Speciation." Science (New York, N.Y.) **317**(5840): 910-914.
Riley, R. and V. Chapman (1958). "Genetic Control of the Cytologically Diploid Behaviour of Hexaploid Wheat." Nature **182**(4637): 713-715.
Risso, D., J. Ngai, T. P. Speed and S. Dudoit (2014). "Normalization of RNA-seq data using factor analysis of control genes or samples." Nat Biotech **32**(9): 896-902.
Robertson, I. H. (1981). "Chromosome numbers in *Brachypodium* Beauv. (Gramineae)." Genetica **56**(1): 55-60.
Roudier, F., I. Ahmed, C. Bérard, A. Sarazin, T. Mary-Huard, S. Cortijo, D. Bouyer, E. Caillieux, E. Duvernois-Berthet, L. Al-Shikhley, L. Giraut, B. Després, S. Drevensek, F. Barneche, S. Dèrozier, V. Brunaud, S. Aubourg, A. Schnittger, C. Bowler, M.-L. Martin-Magniette, S. Robin, M. Caboche and V. Colot (2011). "Integrative epigenomic mapping defines four main chromatin states in Arabidopsis." The EMBO Journal **30**(10): 1928-1938.
Sabot, F., D. Simon and M. Bernard (2004). "Plant transposable elements, with an emphasis on grass species." Euphytica **139**(3): 227-247.

Salmon, A. and M. L. Ainouche (2010). "Polyploidy and DNA methylation: new tools available." Mol Ecol **19**(2): 213-215.
Salmon, A., L. Flagel, B. Ying, J. A. Udall and J. F. Wendel (2010). "Homoeologous nonreciprocal recombination in polyploid cotton." New Phytol **186**(1): 123-134.
Saminathan, T., P. Nimmakayala, S. Manohar, S. Malkaram, A. Almeida, R. Cantrell, Y. Tomason, L. Abburi, M. A. Rahman, V. G. Vajja, A. Khachane, B. Kumar, H. K. Rajasimha, A. Levi, T. Wehner and U. K. Reddy (2015). "Differential gene expression and alternative splicing between diploid and tetraploid watermelon." Journal of Experimental Botany **66**(5): 1369-1385.
Sanger, F. and A. R. Coulson (1975). "A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase." J Mol Biol **94**(3): 441-448.
Sanger, F., S. Nicklen and A. R. Coulson (1977). "DNA sequencing with chain-terminating inhibitors." Proceedings of the National Academy of Sciences **74**(12): 5463-5467.
Sarilar, V., P. M. Palacios, A. Rousselet, C. Ridel, M. Falque, F. Eber, A. M. Chevre, J. Joets, P. Brabant and K. Alix (2013). "Allopolyploidy has a moderate impact on restructuring at three contrasting transposable element insertion sites in resynthesized Brassica napus allotetraploids." New Phytol **198**(2): 593-604.
Schadt, E. E., S. Turner and A. Kasarskis (2010). "A window into third-generation sequencing." Human Molecular Genetics **19**(R2): R227-R240.
Schnable, P. S., D. Ware, R. S. Fulton, J. C. Stein, F. Wei, S. Pasternak, C. Liang, J. Zhang, L. Fulton, T. A. Graves, P. Minx, A. D. Reily, L. Courtney, S. S. Kruchowski, C. Tomlinson, C. Strong, K. Delehaunty, C. Fronick, B. Courtney, S. M. Rock, E. Belter, F. Du, K. Kim, R. M. Abbott, M. Cotton, A. Levy, P. Marchetto, K. Ochoa, S. M. Jackson, B. Gillam, W. Chen, L. Yan, J. Higginbotham, M. Cardenas, J. Waligorski, E. Applebaum, L. Phelps, J. Falcone, K. Kanchi, T. Thane, A. Scimone, N. Thane, J. Henke, T. Wang, J. Ruppert, N. Shah, K. Rotter, J. Hodges, E. Ingenthron, M. Cordes, S. Kohlberg, J. Sgro, B. Delgado, K. Mead, A. Chinwalla, S. Leonard, K. Crouse, K. Collura, D. Kudrna, J. Currie, R. He, A. Angelova, S. Rajasekar, T. Mueller, R. Lomeli, G. Scara, A. Ko, K. Delaney, M. Wissotski, G. Lopez, D. Campos, M. Braidotti, E. Ashley, W. Golser, H. Kim, S. Lee, J. Lin, Z. Dujmic, W. Kim, J. Talag, A. Zuccolo, C. Fan, A. Sebastian, M. Kramer, L. Spiegel, L. Nascimento, T. Zutavern, B. Miller, C. Ambroise, S. Muller, W. Spooner, A. Narechania, L. Ren, S. Wei, S. Kumari, B. Faga, M. J. Levy, L. McMahan, P. Van Buren, M. W. Vaughn, K. Ying, C. T. Yeh, S. J. Emrich, Y. Jia, A. Kalyanaraman, A. P. Hsia, W. B. Barbazuk, R. S. Baucom, T. P. Brutnell, N. C. Carpita, C. Chaparro, J. M. Chia, J. M. Deragon, J. C. Estill, Y. Fu, J. A. Jeddeloh, Y. Han, H. Lee, P. Li, D. R. Lisch, S. Liu, Z. Liu, D. H. Nagel, M. C. McCann, P. SanMiguel, A. M. Myers, D. Nettleton, J. Nguyen, B. W. Penning, L. Ponnala, K. L. Schneider, D. C. Schwartz, A. Sharma, C. Soderlund, N. M. Springer, Q. Sun, H. Wang, M. Waterman, R. Westerman, T. K. Wolfgruber, L. Yang, Y. Yu, L. Zhang, S. Zhou, Q. Zhu, J. L. Bennetzen, R. K. Dawe, J. Jiang, N. Jiang, G. G. Presting, S. R. Wessler, S. Aluru, R. A. Martienssen, S. W. Clifton, W. R. McCombie, R. A. Wing and R. K. Wilson (2009). "The B73 maize genome: complexity, diversity, and dynamics." Science **326**(5956): 1112-1115.
Schurch, N. J., P. Schofield, M. Gierlinski, C. Cole, A. Sherstnev, V. Singh, N. Wrobel, K. Gharbi, G. G. Simpson, T. Owen-Hughes, M. Blaxter and G. J. Barton (2016). "How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use?" RNA **22**(6): 839-851.
Sears, E. R. (1976). "Genetic control of chromosome pairing in wheat." Annu Rev Genet **10**: 31-51.
Seehausen, O., R. K. Butlin, I. Keller, C. E. Wagner, J. W. Boughman, P. A. Hohenlohe, C. L. Peichel, G.-P. Saetre, C. Bank, A. Brannstrom, A. Brelsford, C. S. Clarkson, F. Eroukhmanoff, J. L. Feder, M. C. Fischer, A. D. Foote, P. Franchini, C. D. Jiggins, F. C. Jones, A. K. Lindholm, K. Lucek, M. E. Maan, D. A. Marques, S. H. Martin, B. Matthews, J. I. Meier, M. Most, M. W.

Nachman, E. Nonaka, D. J. Rennison, J. Schwarzer, E. T. Watson, A. M. Westram and A. Widmer (2014). "Genomics and the origin of species." Nat Rev Genet **15**(3): 176-192.
Shaked, H., K. Kashkush, H. Ozkan, M. Feldman and A. A. Levy (2001). "Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat." Plant Cell **13**(8): 1749-1759.
Silkworth, W. T. and D. Cimini (2012). "Transient defects of mitotic spindle geometry and chromosome segregation errors." Cell Division **7**(1): 1-8.
Smith, A. D. (2000). Oxford Dictionary of Biochemistry and Molecular Biology, Oxford University Press.
Soltis, D. E., V. A. Albert, J. Leebens-Mack, C. D. Bell, A. H. Paterson, C. Zheng, D. Sankoff, C. W. Depamphilis, P. K. Wall and P. S. Soltis (2009). "Polyploidy and angiosperm diversification." Am J Bot **96**(1): 336-348.
Soltis, D. E., P. S. Soltis, J. C. Pires, A. Kovarik, J. A. Tate and E. Mavrodiev (2004). "Recent and recurrent polyploidy in Tragopogon (Asteraceae): cytogenetic, genomic and genetic comparisons." Biological Journal of the Linnean Society **82**(4): 485-501.
Soltis, D. E., P. S. Soltis, D. W. Schemske, J. F. Hancock, J. N. Thompson, B. C. Husband and W. S. Judd (2007). "Autopolyploidy in Angiosperms: Have We Grossly Underestimated the Number of Species?" Taxon **56**(1): 13-30.
Soltis, D. E., C. J. Visger and P. S. Soltis (2014). "The polyploidy revolution then…and now: Stebbins revisited." American Journal of Botany.
Soltis, P. S. and D. E. Soltis (2009). "The role of hybridization in plant speciation." Annu Rev Plant Biol **60**: 561-588.
Soneson, C. and M. Delorenzi (2013). "A comparison of methods for differential expression analysis of RNA-seq data." BMC Bioinformatics **14**(1): 1-18.
Song, K., P. Lu, K. Tang and T. C. Osborn (1995). "Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution." Proc Natl Acad Sci U S A **92**(17): 7719-7723.
Song, K., K. Tang and T. C. Osborn (1993). "Development of synthetic Brassica amphidiploids by reciprocal hybridization and comparison to natural amphidiploids." Theor Appl Genet **86**(7): 811-821.
Soreng, R. J., P. M. Peterson, K. Romaschenko, G. Davidse, F. O. Zuloaga, E. J. Judziewicz, T. S. Filgueiras, J. I. Davis and O. Morrone (2015). "A worldwide phylogenetic classification of the Poaceae (Gramineae)." Journal of Systematics and Evolution **53**(2): 117-137.
Spillane, C., C. MacDougall, C. Stock, C. Kohler, J. P. Vielle-Calzada, S. M. Nunes, U. Grossniklaus and J. Goodrich (2000). "Interaction of the Arabidopsis polycomb group proteins FIE and MEA mediates their common phenotypes." Curr Biol **10**(23): 1535-1538.
Stebbins, G. (1947). "Types of polyploids: their classification and significance." Adv. Genet **1**(403429): 1939.
Stebbins, G. L. (1985). "Polyploidy, Hybridization, and the Invasion of New Habitats." Annals of the Missouri Botanical Garden **72**(4): 824-832.
Stebbins Jr, C. (1950). "Variation and evolution in plants." Variation and evolution in plants.
Steinwand, M. A., H. A. Young, J. N. Bragg, C. M. Tobias and J. P. Vogel (2013). "Brachypodium sylvaticum, a model for perennial grasses: transformation and inbred line development." PLoS One **8**(9): e75180.
Storchova, Z., A. Breneman, J. Cande, J. Dunn, K. Burbank, E. O'Toole and D. Pellman (2006). "Genome-wide genetic analysis of polyploidy in yeast." Nature **443**(7111): 541-547.
Sweigart, A. L., L. Fishman and J. H. Willis (2006). "A Simple Genetic Incompatibility Causes Hybrid Male Sterility in Mimulus." Genetics **172**(4): 2465-2479.
Sweigart, A. L. and J. H. Willis (2012). "Molecular evolution and genetics of postzygotic reproductive isolation in plants." F1000 Biology Reports **4**: 23.

Swigonova, Z., J. Lai, J. Ma, W. Ramakrishna, V. Llaca, J. L. Bennetzen and J. Messing (2004). "On the tetraploid origin of the maize genome." Comparative and functional genomics **5**(3): 281-284.
Sybenga, J. (1996). "Chromosome pairing affinity and quadrivalent formation in polyploids: do segmental allopolyploids exist?" Genome **39**: 1176-1184.
Szadkowski, E., F. Eber, V. Huteau, M. Lode, C. Huneau, H. Belcram, O. Coriton, M. J. Manzanares-Dauleux, R. Delourme, G. J. King, B. Chalhoub, E. Jenczewski and A. M. Chevre (2010). "The first meiosis of resynthesized Brassica napus, a genome blender." New Phytol **186**(1): 102-112.
Takuno, S. and B. S. Gaut (2013). "Gene body methylation is conserved between plant orthologs and is of evolutionary consequence." Proc Natl Acad Sci U S A **110**(5): 1797-1802.
Tan, F. L. and J. Q. Yin (2004). "RNAi, a new therapeutic strategy against viral infection." Cell Res **14**(6): 460-466.
Tarutani, Y., H. Shiba, M. Iwano, T. Kakizaki, G. Suzuki, M. Watanabe, A. Isogai and S. Takayama (2010). "Trans-acting small RNA determines dominance relationships in Brassica self-incompatibility." Nature **466**(7309): 983-986.
Tayale, A. and C. Parisod (2013). "Natural pathways to polyploidy in plants and consequences for genome reorganization." Cytogenet Genome Res **140**(2-4): 79-96.
Teshima, K. M. and H. Innan (2008). "Neofunctionalization of Duplicated Genes Under the Pressure of Gene Conversion." Genetics **178**(3): 1385-1398.
Thole, V. and P. Vain (2012). "Agrobacterium-mediated transformation of Brachypodium distachyon." Methods Mol Biol **847**: 137-149.
Tian, L., X. Li, M. Ha, C. Zhang and Z. J. Chen (2014). "Genetic and epigenetic changes in a genomic region containing MIR172 in Arabidopsis allopolyploids and their progenitors." Heredity **112**(2): 207-214.
Tiley, G. P., C. Ane and J. G. Burleigh (2016). "Evaluating and Characterizing Ancient Whole-Genome Duplications in Plants with Gene Count Data." Genome Biol Evol **8**(4): 1023-1037.
Torres, E. M., B. R. Williams and A. Amon (2008). "Aneuploidy: cells losing their balance." Genetics **179**(2): 737-746.
Tsai, J. H. and B. D. McKee (2011). "Homologous pairing and the role of pairing centers in meiosis." J Cell Sci **124**(Pt 12): 1955-1963.
Tuskan, G. A., S. Difazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, S. Ralph, S. Rombauts, A. Salamov, J. Schein, L. Sterck, A. Aerts, R. R. Bhalerao, R. P. Bhalerao, D. Blaudez, W. Boerjan, A. Brun, A. Brunner, V. Busov, M. Campbell, J. Carlson, M. Chalot, J. Chapman, G. L. Chen, D. Cooper, P. M. Coutinho, J. Couturier, S. Covert, Q. Cronk, R. Cunningham, J. Davis, S. Degroeve, A. Dejardin, C. Depamphilis, J. Detter, B. Dirks, I. Dubchak, S. Duplessis, J. Ehlting, B. Ellis, K. Gendler, D. Goodstein, M. Gribskov, J. Grimwood, A. Groover, L. Gunter, B. Hamberger, B. Heinze, Y. Helariutta, B. Henrissat, D. Holligan, R. Holt, W. Huang, N. Islam-Faridi, S. Jones, M. Jones-Rhoades, R. Jorgensen, C. Joshi, J. Kangasjarvi, J. Karlsson, C. Kelleher, R. Kirkpatrick, M. Kirst, A. Kohler, U. Kalluri, F. Larimer, J. Leebens-Mack, J. C. Leple, P. Locascio, Y. Lou, S. Lucas, F. Martin, B. Montanini, C. Napoli, D. R. Nelson, C. Nelson, K. Nieminen, O. Nilsson, V. Pereda, G. Peter, R. Philippe, G. Pilate, A. Poliakov, J. Razumovskaya, P. Richardson, C. Rinaldi, K. Ritland, P. Rouze, D. Ryaboy, J. Schmutz, J. Schrader, B. Segerman, H. Shin, A. Siddiqui, F. Sterky, A. Terry, C. J. Tsai, E. Uberbacher, P. Unneberg, J. Vahala, K. Wall, S. Wessler, G. Yang, T. Yin, C. Douglas, M. Marra, G. Sandberg, Y. Van de Peer and D. Rokhsar (2006). "The genome of black cottonwood, Populus trichocarpa (Torr. & Gray)." Science **313**(5793): 1596-1604.
Udall, J. A., P. A. Quijada and T. C. Osborn (2005). "Detection of chromosomal rearrangements derived from homologous recombination in four mapping populations of Brassica napus L." Genetics **169**(2): 967-979.

Valouev, A., J. Ichikawa, T. Tonthat, J. Stuart, S. Ranade, H. Peckham, K. Zeng, J. A. Malek, G. Costa, K. McKernan, A. Sidow, A. Fire and S. M. Johnson (2008). "A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning." Genome Res **18**(7): 1051-1063.
van Dijk, E. L., H. Auger, Y. Jaszczyszyn and C. Thermes (2014). "Ten years of next-generation sequencing technology." Trends Genet **30**(9): 418-426.
van Ginkel, M. and F. Ogbonnaya (2007). "Novel genetic diversity from synthetic wheats in breeding cultivars for changing production conditions." Field Crops Research **104**(1): 86-94.
VanBuren, R. and T. C. Mockler (2016). The Brachypodium distachyon Reference Genome. Genetics and Genomics of Brachypodium. P. J. Vogel. Cham, Springer International Publishing**:** 55-70.
Vazquez, F. and T. Hohn (2013). "Biogenesis and Biological Activity of Secondary siRNAs in Plants." Scientifica **2013**: 12.
Vicient, C. M., M. J. Jaaskelainen, R. Kalendar and A. H. Schulman (2001). "Active retrotransposons are a common feature of grass genomes." Plant Physiol **125**(3): 1283-1292.
Vogel, J. P. (2016). Genetics and Genomics of Brachypodium, Springer.
Vogel, J. P. (2016). The Rise of Brachypodium as a Model System. Genetics and Genomics of Brachypodium. P. J. Vogel. Cham, Springer International Publishing**:** 1-7.
Wang, A., X. Zhang, C. Yang, Z. Song, C. Du, D. Chen, Y. He and D. Cai (2013). "Development and characterization of synthetic amphiploid (AABB) between Oryza sativa and Oryza punctata." Euphytica **189**(1): 1-8.
Wang, J., L. Tian, H.-S. Lee, N. E. Wei, H. Jiang, B. Watson, A. Madlung, T. C. Osborn, R. W. Doerge, L. Comai and Z. J. Chen (2006). "Genomewide Nonadditive Gene Regulation in Arabidopsis Allotetraploids." Genetics **172**(1): 507-517.
Wang, J., L. Tian, H. S. Lee and Z. J. Chen (2006). "Nonadditive regulation of FRI and FLC loci mediates flowering-time variation in Arabidopsis allopolyploids." Genetics **173**(2): 965-974.
Wang, J., L. Tian, A. Madlung, H.-S. Lee, M. Chen, J. J. Lee, B. Watson, T. Kagochi, L. Comai and Z. J. Chen (2004). "Stochastic and Epigenetic Changes of Gene Expression in Arabidopsis Polyploids." Genetics **167**(4): 1961-1973.
Watson, J. D. and F. H. Crick (1953). "The structure of DNA." Cold Spring Harb Symp Quant Biol **18**: 123-131.
Weaver, B. A. and D. W. Cleveland (2009). "The role of aneuploidy in promoting and suppressing tumors." J Cell Biol **185**(6): 935-937.
Weiss-Schneeweiss, H., K. Emadzade, T. S. Jang and G. M. Schneeweiss (2013). "Evolutionary consequences, constraints and potential of polyploidy in plants." Cytogenet Genome Res **140**(2-4): 137-150.
Weiss, H. and J. Maluszynska (2000). "Chromosomal rearrangement in autotetraploid plants of Arabidopsis thaliana." Hereditas **133**(3): 255-261.
Wendel, J. F. (2000). "Genome evolution in polyploids." Plant Mol Biol **42**(1): 225-249.
Wendel, J. F., S. A. Jackson, B. C. Meyers and R. A. Wing (2016). "Evolution of plant genome architecture." Genome Biol **17**(1): 37.
Wescoe, Z. L., J. Schreiber and M. Akeson (2014). "Nanopores discriminate among five C5-cytosine variants in DNA." J Am Chem Soc **136**(47): 16582-16587.
Wicker, T., F. Sabot, A. Hua-Van, J. L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel and A. H. Schulman (2007). "A unified classification system for eukaryotic transposable elements." Nat Rev Genet **8**(12): 973-982.
Widmer, A., C. Lexer and S. Cozzolino (2009). "Evolution of reproductive isolation in plants." Heredity (Edinb) **102**(1): 31-38.
Wolfe, K. H. (2001). "Yesterday's polyploids and the mystery of diploidization." Nat Rev Genet **2**(5): 333-341.

Wolny, E. and R. Hasterok (2009). "Comparative cytogenetic analysis of the genomes of the model grass *Brachypodium distachyon* and its close relatives." Ann Bot **104**(5): 873-881.
Wolny, E., K. Lesniewska, R. Hasterok and T. Langdon (2011). "Compact genomes and complex evolution in the genus Brachypodium." Chromosoma **120**(2): 199-212.
Xu, Y., L. Zhong, X. Wu, X. Fang and J. Wang (2009). "Rapid alterations of gene expression and cytosine methylation in newly synthesized Brassica napus allopolyploids." Planta **229**(3): 471-483.
Yamagata, Y., E. Yamamoto, K. Aya, K. T. Win, K. Doi, Sobrizal, T. Ito, H. Kanamori, J. Wu, T. Matsumoto, M. Matsuoka, M. Ashikari and A. Yoshimura (2010). "Mitochondrial gene in the nuclear genome induces reproductive barrier in rice." Proceedings of the National Academy of Sciences **107**(4): 1494-1499.
Yeung, K. Y. and W. L. Ruzzo (2001). "Principal component analysis for clustering gene expression data." Bioinformatics **17**(9): 763-774.
Yoo, M. J., X. Liu, J. C. Pires, P. S. Soltis and D. E. Soltis (2014). "Nonadditive gene expression in polyploids." Annu Rev Genet **48**: 485-517.
Youds, J. L. and S. J. Boulton (2011). "The choice in meiosis - defining the factors that influence crossover or non-crossover formation." J Cell Sci **124**(Pt 4): 501-513.
Young, N. D., F. Debelle, G. E. D. Oldroyd, R. Geurts, S. B. Cannon, M. K. Udvardi, V. A. Benedito, K. F. X. Mayer, J. Gouzy, H. Schoof, Y. Van de Peer, S. Proost, D. R. Cook, B. C. Meyers, M. Spannagl, F. Cheung, S. De Mita, V. Krishnakumar, H. Gundlach, S. Zhou, J. Mudge, A. K. Bharti, J. D. Murray, M. A. Naoumkina, B. Rosen, K. A. T. Silverstein, H. Tang, S. Rombauts, P. X. Zhao, P. Zhou, V. Barbe, P. Bardou, M. Bechner, A. Bellec, A. Berger, H. Berges, S. Bidwell, T. Bisseling, N. Choisne, A. Couloux, R. Denny, S. Deshpande, X. Dai, J. J. Doyle, A.-M. Dudez, A. D. Farmer, S. Fouteau, C. Franken, C. Gibelin, J. Gish, S. Goldstein, A. J. Gonzalez, P. J. Green, A. Hallab, M. Hartog, A. Hua, S. J. Humphray, D.-H. Jeong, Y. Jing, A. Jocker, S. M. Kenton, D.-J. Kim, K. Klee, H. Lai, C. Lang, S. Lin, S. L. Macmil, G. Magdelenat, L. Matthews, J. McCorrison, E. L. Monaghan, J.-H. Mun, F. Z. Najar, C. Nicholson, C. Noirot, M. O/'Bleness, C. R. Paule, J. Poulain, F. Prion, B. Qin, C. Qu, E. F. Retzel, C. Riddle, E. Sallet, S. Samain, N. Samson, I. Sanders, O. Saurat, C. Scarpelli, T. Schiex, B. Segurens, A. J. Severin, D. J. Sherrier, R. Shi, S. Sims, S. R. Singer, S. Sinharoy, L. Sterck, A. Viollet, B.-B. Wang, K. Wang, M. Wang, X. Wang, J. Warfsmann, J. Weissenbach, D. D. White, J. D. White, G. B. Wiley, P. Wincker, Y. Xing, L. Yang, Z. Yao, F. Ying, J. Zhai, L. Zhou, A. Zuber, J. Denarie, R. A. Dixon, G. D. May, D. C. Schwartz, J. Rogers, F. Quetier, C. D. Town and B. A. Roe (2011). "The Medicago genome provides insight into the evolution of rhizobial symbioses." Nature **480**(7378): 520-524.
Yousafzai, F. K., N. Al-Kaff and G. Moore (2010). "Structural and functional relationship between the Ph1 locus protein 5B2 in wheat and CDK2 in mammals." Funct Integr Genomics **10**(2): 157-166.
Yu, Z., G. Haberer, M. Matthes, T. Rattei, K. F. Mayer, A. Gierl and R. A. Torres-Ruiz (2010). "Impact of natural genetic variation on the transcriptome of autotetraploid Arabidopsis thaliana." Proc Natl Acad Sci U S A **107**(41): 17809-17814.
Zhang, H., Y. Bian, X. Gou, B. Zhu, C. Xu, B. Qi, N. Li, S. Rustgi, H. Zhou, F. Han, J. Jiang, D. von Wettstein and B. Liu (2013). "Persistent whole-chromosome aneuploidy is generally associated with nascent allohexaploid wheat." Proceedings of the National Academy of Sciences of the United States of America **110**(9): 3447-3452.
Zhang, H., B. Zhu, B. Qi, X. Gou, Y. Dong, C. Xu, B. Zhang, W. Huang, C. Liu, X. Wang, C. Yang, H. Zhou, K. Kashkush, M. Feldman, J. F. Wendel and B. Liu (2014). "Evolution of the BBAA Component of Bread Wheat during Its History at the Allohexaploid Level." The Plant Cell **26**(7): 2761-2776.
Zhang, L., G. Sun, Z. Yan, Q. Chen, Z. Yuan, X. Lan, Y. Zheng and D. Liu (2007). "Comparison of newly synthetic hexaploid wheat with its donors on SSR products." J Genet Genomics **34**(10): 939-946.

Zhang, X., M. Deng and G. Fan (2014). "Differential transcriptome analysis between Paulownia fortunei and its synthesized autopolyploid." Int J Mol Sci **15**(3): 5079-5093.
Zhang, Z., H. Belcram, P. Gornicki, M. Charles, J. Just, C. Huneau, G. Magdelenat, A. Couloux, S. Samain, B. S. Gill, J. B. Rasmussen, V. Barbe, J. D. Faris and B. Chalhoub (2011). "Duplication and partitioning in evolution and function of homoeologous Q loci governing domestication characters in polyploid wheat." Proceedings of the National Academy of Sciences **108**(46): 18737-18742.
Zhao, N., B. Zhu, M. Li, L. Wang, L. Xu, H. Zhang, S. Zheng, B. Qi, F. Han and B. Liu (2011). "Extensive and heritable epigenetic remodeling and genetic stability accompany allohexaploidization of wheat." Genetics **188**(3): 499-510.
Zhao, Y. T., M. Wang, S. X. Fu, W. C. Yang, C. K. Qi and X. J. Wang (2012). "Small RNA profiling in two Brassica napus cultivars identifies microRNAs with oil production- and development-correlated expression and new small RNA classes." Plant Physiol **158**(2): 813-823.
Zhou, J., C. Tan, C. Cui, X. Ge and Z. Li (2016). "Distinct subgenome stabilities in synthesized Brassica allohexaploids." Theor Appl Genet.
Zhou, Y., L. Kang, S. Liao, Q. Pan, X. Ge and Z. Li (2015). "Transcriptomic analysis reveals differential gene expressions for cell growth and functional secondary metabolites in induced autotetraploid of Chinese woad (Isatis indigotica Fort.)." PLoS One **10**(3): e0116392.
Zyprych-Walczak, J., A. Szabelska, L. Handschuh, K. Gorczak, K. Klamecka, M. Figlerowicz and I. Siatkowski (2015). "The Impact of Normalization Methods on RNA-Seq Data Analysis." Biomed Res Int **2015**: 621690.

Résumé

Titre: ***Développement et caractérisation d'un modèle polyploïde chez les espèces de Brachypodium***

Mots clés : Brachypodium, L'expression des gènes, Polyploïdie, Allopolyploïdes, Hybridation interspécifique, autopolyploïdes

Résumé : La polyploïdie consiste en la duplication du génome entier et constitue une force évolutive majeure chez les eucaryotes, notamment chez les angiospermes. Les espèces du genre Brachypodium ont émergé comme un modèle intéressant des monocotylédones. Parmi ces espèces, l'allopolyploïde B. hybridum (2n = 30) a résulté de l'hybridation entre B. distachyon (2n = 10) et B. stacei (2n = 20). Les deux espèces parentales ont eu des évolutions chromosomiques assez divergentes aboutissant à ce que B. distachyon possède deux fois moins de chromosomes qui sont néanmoins deux fois plus grands que ceux de B. stacei.

L'objectif de ma thèse est de développer un modèle de Brachypodium polyploïde en synthétisant des autopolyploïdes et des allopolyploïdes ainsi qu'en les caractérisant aux niveaux phénotypique, cytogénétique et génomique.

Les autotétraploïdes ont été générés par traitement à la cochicine de deux lignées de B. distachyon et trois autres de B. stacei. Tous les autopolyploïdes obtenus ont été validés par cytométrtie de flux et ont montré une stabilité de caryotype et de phénotype, exceptés ceux de la lignée ABR114 de B. stacei qui ont montré des descendances autotetraploides et aneuploïdes. Les comparaisons des caractères de l'inflorescence et des ligules montrent que les autotétraploïdes des deux espèces dépassent généralement leurs progéniteurs diploïdes, mais une faible fertilité illustrée par le faible nombre de graines par inflorescence.

Les allotétraploïdes synthétiques, similaires à B. hybridum, ont été générés par hybridation interspécifique entre différentes lignées de B. distachyon et B. stacei. Alors que les hybrides interspécifiques F1 stériles ont été obtenus lorsque B. distachyon a été utilisé comme parent maternel (0,15% ou 0,245% des croisements réalisés), aucun hybride a été obtenu à partir des croisements réciproques ou lorsque des plantes autotetraploïdes des espèces parentales ont été croisées.

Le doublement du génome par traitement à la colchicine a restauré la fertilité et les plantes F1 doublées produisent par auto-polllinisation quelques graines S1 (première génération d'autofécondation). Les plantes S1 de la l'allo3-1×5 étaient fertiles donnant des générations suivantes alors que ceux de l'allo21×114 étaient stériles. Ces allotétraploïdes se sont montrés stables au niveau phénotypique, cytogénétique et génomique, ressemblant ainsi au B. hybridum naturel.

J'ai utilisé la méthodologie Illumina de nouvelle génération de séquençage pour caractériser ces différents polyploïdes aux niveaux: (i) Des variations structurales par reséquençage de l'ADN, (ii) Du transcriptome par l'ARN-Seq afin de caractériser les modifications de l'expression des gènes (iii) la méthylation CpG par séquençage de l'ADN traité au bisulfite. Une large gamme de possibilités de comparaisons pourrait être réalisée lorsque les séquences des génomes des trois espèces naturelles de Brachypodium seront disponibles.

La disponibilité de la séquence du génome de B. distachyon a permis de réaliser une analyse pilote comparant l'expression des gènes dupliqués dans les autopolyploïdes de la lignée Bd3-1 à celle des diploïdes. Une expression similaire pour 95,25% des gènes a été observée et seulement 4,75% des gènes ont montré une expression différentielle. La majeure partie de ces derniers (1053 gènes, 82%) étaient moins exprimés dans les autopolyploides. Des fonctions géniques telles que la ségrégation des chromosomes nucléaires, la réparation de l'ADN, la recombinaison homologue, la régulation de la transcription se sont révélées enrichies.

La création des autopolyploides et allopolyploides et leurs caractérisations avec les outils NGS offrent la possibilité d'investir de façon intégrée les secrets du succès des polyploïdes.

Abstract

Title: Development and characterization of the Brachypodium species polyploid model
Keywords: Brachypodium, gene expression, polyploid, allopolyploid, autopolyploid , interspecific hybrid

Abstract: Polyploidy consisting in whole genome duplication is an important evolutionary force in eukaryote and very prominent in angiosperms. Species of the Brachypodium genus emerged as an important monocot and polyploid model. Among these species, the annual allopolyploid B. hybridum (2n=30), derived from hybridizations between B. distachyon (2n=10) and B. stacei (2n=20), was shown to be polyphyletic. The two parental species have similar genome content and ploidy level but an asymmetric chromosome evolution where B. distachyon has two times less chromosomes that are two times bigger than those of B. stacei.

The objective of my PhD program consisted in developing a valuable Brachypodium polyploid model by synthesizing autopolyploids and allopolyploids and then their characterizing at the phenotype, cytogenetic and genomic levels.

Autotetraploids were generated from two inbred lines of B. distachyon and three ones of B. stacei, through colchicine treatments. The genome doubling was validated by flow cytometry and karyotyping with fluorescent in situ hybridization analyses. All autopolyploids showed stability in phenotype as well as karyotype except those of line ABR114 of B. stacei that showed various aneuploid progenies. Quantitative comparison of inflorescences and flag leaves characters showed that both B. distachyon and B. stacei autotetraploids generally exceeded their diploid progenitors, but their fertility was reduced as illustrated by the lower number of seeds per inflorescence.

Synthetic allotetraploids were generated through interspecific hybridization between various lines of B. distachyon and B. stacei species. While sterile amphihaploid F1 interspecific hybrids were obtained at low frequencies (0,15% or 0,245% of crosses) when B. distachyon was used as the maternal parent, no hybrids were obtained from reciprocal crosses or when autotetraploids of the parental species were crossed. Genome doubling through colchicine treatment restored fertility where doubled F1 plants produced a few S1 seeds (first selfed generation). S1 plants of allo3-1×5 were fertile and gave rise to further generations whereas those of allo21×114 were sterile. The synthetic allotetraploids were shown to be highly-stable and resembled the natural B. hybridum allopolyploid at the phenotypic, cytogenetic and genomic levels.

I have used the Illumina next generation sequencing (NGS) methodology to characterize the different polyploids at the levels of: (i) The DNA resequencing to reveal genetic changes, (ii) The transcriptome analysis through RNA-Sequencing and (iii) CpG methylation through bisulfite sequencing. A total of 13 genotypes containing synthetic autopolyploids, synthetic allopolyploids and the natural diploid species B. distachyon, B. stacei and the allopolyploid B. hybridum have characterized. A wide range of possibilities of comparison could be then realized, when the genomes sequences of all three species will be available.

The availability of sequence genome sequence of B. distachyon allowed a pilot gene expression comparison between diploids and autopolyploids of Bd3-1. Diploids and autotetrapolyploids of of Bd3-1 showed similar expression for most of the genes (95.25%). Only 4.75% of total genes were differentially expressed genes, the major proportion of which (1053 genes, 82%) were down-expressed in autotetraploids with important enrichment in genome maintenance functions such as nuclear chromosome segregation, DNA repair, DNA replication, homologous recombination.

The successful creation of stable autotetraploids and synthetic B. hybridum allopolyploids together with genome wide characterization using NGS offer the possibility to unravel clues of success of polyploidy in angiosperms.