GMR
The Original

CENH3 BASED HAPLOID INDUCTION SYSTEM FOR ACCELERATED CROP BREEDING

Dilbagh[1], Prastuti Bhattacharyya[2], Shivam Singh[3], Satyam Singh[4], Sagar[5], Sonu Langaya[*6]

[1]Department of Plant Breeding and Genetics, Punjab Agricultural University, Ludhiana, 141004, Punjab, India, dilbagh34437@gmail.com
[2]Department of Genetics and Plant Breeding, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, 221005, Uttar Pradesh, India, prastutibhatta@gmail.com
[3]Department of Genetics and Plant Breeding, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, 221005, Uttar Pradesh, India, sshivamm09@gmail.com
[4]Department of Genetics and Plant Breeding, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, 221005, Uttar Pradesh, India, satyamsinghp6@bhu.ac.in
[5]Department of Genetics & Plant Breeding, CCS Haryana Agricultural University, Hisar, 125004, Haryana, India, sagararora470@gmail.com
[6*]Department of Genetics and Plant Breeding, Faculty of Agricultural Sciences, GLA University, Mathura, 281406, Uttar Pradesh, India, sonulangaya@yahoo.in

ABSTRACT

Haploid induction is an efficient method to reduce the generations required for conventional inbreeding, yielding completely homozygous doubled-haploid lines that allow acceleration of the crop improvement process. Of the new haploid induction systems, a system involving manipulation of the centromere-specific histone H3 variant (CENH3), involved in the identity of centromeres and chromosome segregation, has received considerable attention. CENH3 alteration can affect the function of the centromere and cause the preferential elimination of chromosomes after fertilization leading to haploid embryo development. In this review, we discuss the development and current understanding of CENH3-mediated haploid induction with particular focus on the functional conservation of CENH3 in different plant species and the role of the mutations in both the N-terminal and histone fold in haploid induction. The effects of crossing direction on genome elimination efficiency and the reported reduction or change of inheritance of CENH3 during zygotic and early embryonic development are also discussed. The potential applications of CENH3-based haploid induction in accelerated breeding and genetic analysis are focused on, such as rapid development of homozygous lines, cytoplasmic and nuclear genome exchange, reverse breeding, and gene stacking through biparental crosses. However, variable efficiency of induction, genotype dependency, fertility constraints and the lack of understanding of the cellular mechanisms responsible for the elimination of chromosomes are still major limitations in its widespread use. The regulatory aspects and alternative approaches of efficient haploid induction systems are also discussed. Moreover, recent developments in research, such as the one developed at the Punjab Agricultural University, are also discussed to demonstrate the translational potential of CENH3-related approaches. Altogether, haploid induction by CENH3 offers a superior platform that enables the incorporation of the principles of centromere biology into current breeding strategies, and has potential to significantly reduce the time to develop genetically uniform, superior crop varieties.

KEYWORDS: CenH3, Haploid Induction, Accelerated Breeding, DNA.

INTRODUCTION

The DNA of eukaryotes is wrapped around histone proteins to form nucleosomes. The organisation and higher-order folding of chromatin helps to pack large genomes into the nucleus of a cell. This is especially relevant for flowering plants, where genome sizes vary from ~61 Mb in Genlisea tuberosa (Fleischmann et al., 2014) to 160 Gb in Paris japonica (Pellicer et al., 2010). The organization of chromatin is dynamic and changes as a function of the cell cycle and functional state of the DNA. Specific histone variants incorporated into nucleosomes and post-translational modifications of histones both affect the structural properties and accessibility of chromatin, and thus DNA packaging and gene expression (Gibcus and Dekker, 2013; Cutter and Hayes, 2015). Histone deposition and modification can be preserved after DNA replication and, in some instances, passed on from generation to generation via a variety of molecular mechanisms (Mattiroli et al., 2015). These changes affect chromatin structure and the ability of DNA to be transcribed by the transcription apparatus and are an important part of epigenetic regulation. These regulatory states can be inherited through cell divisions or across generations and may play a role in epigenetic inheritance. An important example of epigenetic regulation of chromatin organization is the specification and maintenance of centromere identity. In many eukaryotes, the centromeric DNA consists of arrays of tandem repeats, often interspersed with transposable elements, but these DNA characteristics are not always necessary for centromere formation. Not all eukaryotes have large tandem repeat arrays and some eukaryotes have different repeat compositions on different chromosomes. In addition, neocentromeres can form at unique DNA sequences that are not normally centromeric, showing that the identity of a centromere is not always associated with a particular DNA sequence (Marshall et al., 2008). Rather, in most eukaryotes examined, the identity of the centromere is mostly linked to the presence of a variant of the histone H3 protein, called CENH3, which is specifically incorporated at the centromere (Allshire and Karpen, 2008; Stimpson and Sullivan, 2010; Henikoff and Smith, 2015; Westhorpe and Straight, 2015). This protein is also called CID in Drosophila, CENP-A in mammals, and Cse4 in yeast (Talbert et al., 2002; Sekulic

and Black, 2012; Lermontova et al., 2014; Steiner and Henikoff, 2015). A notable exception is Saccharomyces cerevisiae, where the centromeric histone variant has sequence-specific DNA-binding properties and helps to specify the centromere at defined DNA sequences. In most eukaryotes, therefore, the positioning of the centromere is thought to be largely epigenetic, but not only sequence-dependent (Liu et al., 2015). CENH3 is similar to canonical histones in that it consists of an N-terminal tail that extends out from the nucleosome and is subject to post-translational modifications, and a C-terminal histone fold domain that is involved in DNA and other histone interactions. But CENH3 is a histone that evolves very quickly. It has a large N-terminal tail, which is highly variable, even between closely related species, while the histone fold domain is relatively conserved (Malik and Henikoff, 2003). Nevertheless, CENH3 has an essential role in the formation of functional centromeric chromatin and thus in the correct segregation of chromosomes during cell division. Chromosome missegregation, cell-cycle defects or lethality occur when CENH3 or its functional homologues are disrupted in several organisms. In budding yeast, mutations of CENH3 result in chromosome nondisjunction and arrest of the cell cycle at mitosis (Stoler et al., 1995). In C. elegans, the CENH3 knockout causes missegregation of chromosomes and lethality (Buchwitz et al., 1999), while in mice, CenpA knockout results in an early loss of the centromeric chromatin structure (Howman et al., 2000). In a similar manner, the down regulation of CENP-A in human cell lines leads to p53 mediated cellular senescence (Maehara et al., 2010). Down-regulation of CENH3 in Arabidopsis results in a decrease in the number of mitoses and an increase in meiotic segregation errors (Lermontova et al., 2011), and homozygous cenh3 null mutants are early embryonic lethal (Ravi et al., 2010). Taken together, these results reveal that CENH3 is essential for centromere function and genomic stability. Functional centromeres are essential for the accurate transmission of chromosomes, which forms the basis for the use of CENH3 modification in haploid induction. Incorporation and maintenance of CENH3-containing nucleosomes at the centromeres are tightly regulated processes, and disturbances of either of these processes can impair chromosome segregation. Changes in centromere function in CENH3 modified plants can lead to unequal transmission or to post-fertilization loss of chromosomes, leading to the production of haploid plants. Therefore, haploid induction by CENH3 modification is a potential approach to obtaining haploid plants which can then be chromosome doubled to obtain doubled haploids. These lines allow the production of fully homozygous breeding material in a short time and with less effort than would be needed for conventional inbreeding. CENH3 deposition and centromere maintenance have been well studied in animals and fungi (Müller and Almouzni, 2014), but the molecular mechanisms involved in these processes in plants are still poorly understood (Lermontova et al., 2013; Liu et al., 2015). Therefore, a better understanding of the function of CENH3 at the centromere and of chromosome elimination is needed to improve the haploid induction system and to make it more widely applicable in accelerated crop breeding.

THE IMPORTANCE OF HAPLOID INDUCTION

Doubled haploids- plants carrying a genome derived from a single (haploid) gamete-provide enormous timesaving in the production of true-breeding lines, literally reducing the time to acceptable levels of homozygosity by an order of magnitude. Haploid lines, and "haploid genetics" can accelerate a variety of genetic constructions and is useful in the research laboratory also (Ravi et al., 2014). The spontaneous occurrences of haploids have been reported in several species reviewed in (Dunwell, 2010; Dwivedi et al., 2015). In some species haploids have been generated through interspecies and intraspecies hybridization. (Guha and Maheshwari, 1964) showed that haploid plants could be generated from anther culture in some plant species. A variety of other protocols have been developed to generate haploids, reviewed in detail in W¸edzony et al. (2009). Though all of these techniques have proved to be useful over the last 50 years, they are limited to few crop species and/or varieties. In this review we will highlight the recent advances in HI technology via modification of CENH3. Because CENH3 is present in all plants there is promise that this technology can be translated to the majority of crop species.

CENH3 MEDIATED HAPLOID INDUCTION

In a breakthrough discovery, Ravi and Chan (2010) studying the structure/function relationship of Histone 3 variants, found that a cenh3–/– Arabidopsis null mutant, when complemented by altered version of CENH3, could induce haploids. A remarkably effective haploid inducer, termed "green fluorescent protein (GFP)-tailswap," carried GFP fused to the N-terminal tail domain of an H3 variant (At1g13370), which then replaced the N-terminal tail of CENH3 (see Figure 1). They found, surprisingly, that the CENH3 mitotic functionality was maintained in spite of the "tailswap," though meiotic function was compromised- the plants were almost entirely male sterile, with some female sterility also. On self-pollination, GFP-tailswap produces normal diploid seeds (though again, with reduced fertility). However, when crossed as a female by plants carrying wild-type CENH3, GFP-tailswap produces a remarkable 25–45% paternal haploid progeny. The maternal (GFP-tailswap) genome is lost in these haploids (Ravi and Chan, 2010). The almost total male sterility observed in GFP-tailswap plants is ascribed to its meiosis-specific loading abnormalities (Ravi et al., 2011). The authors also observed some, though less frequent, HI in plants carrying a simple N-terminal addition of GFP to CENH3 (no "tailswap"). The three independently derived lines of GFP-tailswap analyzed did not display significant differences in their level of HI frequencies (Tan et al., 2015). This suggests that HI is dependent on the modification of CENH3 rather than on expression levels- although it must be noted that a certain minimum, and perhaps maximum (Marshall et al., 2008) level of expression is selected for in these transgenic experiments, which are performed in a cenh3–/– background. The initial findings with GFP-CENH3 and GFP-tailswap suggested that addition of a bulky tag to CENH3 might interfere with CENH3 recognition, and so produce a centromere that is less competitive for the reloading of CENH3 or subsequent components of the centromere when confronted with competing wild-type centromeres. Working from this hypothesis, we will refer to these modified CENH3's, and their resulting centromeres, as "weak" or "non-competitive" vs. wild-type CENH3. Addition of GFP on its own presumably results in a slightly defective protein, and modification of the N-terminal tail (tailswap) further degrades the strength of the resulting centromere. The addition of adducts such as GFP, and changes

in the sequence of CENH3, can act additively to affect protein function. For example, as described above, GFP-CENH3, in Arabidopsis, is a haploid inducer, but otherwise fully functional. The Arabidopsis null mutant can also be complemented (mitotically and meiotically) by CENH3 from a wide range of angiosperms (Maheshwari et al., 2015), including the monocot Zea mays. However, addition of GFP to the N-terminus of this distantly related protein compromised its function to the extent that while it could localize to centromeres, it could not complement the lethality of the cenH3–/– mutant (Ravi et al., 2010; Figure 1).



FIGURE 1: Effect of various modifications of transgenic CENH3 variants in an Arabidopsis thaliana cenh3 null. Table summarizes complementation (of viability and fertility), haploid, and aneuploid induction by these lines. GFP-tailswap is the commonly used nomenclature in published literature for this construct. (Reproduced from Britt and Kuppu, 2016 under the terms of the Creative Commons Attribution License CC-BY).

FUNCTIONAL CONSERVATION OF CENH3 ACROSS DIVERSE SPECIES

In a pioneering study with human cell lines, it was demonstrated that yeast CENH3 (Cse4) was able to complement a human defect in CENH3 (CENPA) function (Wieland et al., 2004). Similarly, EYPF-tagged versions of CENH3 from monocentric plant species like, Arabidopsis lyrata, A. arenosa, Capsella bursa-pastoris, and Z. mays were able to target the A. thaliana centromere (Moraes et al., 2011). In contrast, an EYPF tagged version of CENH3 from holocentric plant Luzula nivea failed to target A. thaliana centromeres. GFP-tagged versions of CENH3 from A. aeronosa, A. lyrata, andZ. mays were able to target A. thaliana centromere. However, GFP-tagged versions of CENH3 from other kingdoms (S. cerevisiae, C. elegans and Homo sapiens) did not target to A. thaliana centromeres (Ravi et al., 2010) The study also also established that GFP tagged versions of CENH3 from Brassica rapa and Z. mays were able to localize to A. thaliana centromere, but were unable to complement A.thaliana cenh3 null function. This result was surprising given ability of the yeast protein to complement the human defect, which suggests that CENH3 function is highly conserved across kingdoms. This discrepancy was explained when Maheshwari et al. (2015), in a subsequent study employed untagged versions of CENH3 from B. rapa, Lepidium oleracuem, Vitis Vinifera, and Z. mays found that all of these untagged proteins were able to complement A. thaliana cenh3 null function. In another intriguing note, when the GFP is attached C-terminally to the AtCENH3 protein, it is able to localize to the centromere but unable to complement cenh3 null function in Arabidopsis (Ravi et al., 2010), while the N-terminally tagged protein restores viability and full fertility to this mutant. Thus GFP-tagging of CENH3 often places restrictions on its functionality. From the above studies using heterologous, untagged CENH3, it is clear that the essential functions of CENH3 are well conserved across a wide variety of species.

Table 1 | Haploid plants contain only the nuclear genome of their wild-type parent

Cross	Seeds per siliqua	Normal seed (%)	Total plants analysed	Haploids (%)	Diploids (%)	Aneuploids (%)
WT Col-0 × WT Col-0	52 ± 6 (n = 23)	99.5	224	0 (0)	224 (100)	0 (0)
GFP-tailswap × *GFP-tailswap*	0.6 (n = 1,206)	80	213	0 (0)	197 (92)	16 (8)
GFP-tailswap × WT Col-0	32 ± 9 (n = 40)	12	67	23 (34)	23 (34)	21 (32)
WT Col-0 × *GFP-tailswap*	ND	ND	116	5 (4)	99 (85)	12 (11)
GFP-tailswap × WT Ler	30 ± 4 (n = 22)	23	127	32* (25)	32 (25)	63 (50)
GFP-tailswap × WT Ws-0	23 ± 5 (n = 14)	8	22	10* (45)	7 (32)	5 (28)
GFP-tailswap × WT C24/Ler	28 ± 5 (n = 13)	30	117	34† (29)	39 (33)	44 (38)
C24/Ler ms × *GFP-tailswap*	22 ± 14 (n = 18)	63	226	12† (5)	206 (91)	8 (4)
GFP-CENH3 × *GFP-CENH3*	53 ± 4n = 21)	99	209	0 (0)	209 (100)	0 (0)
GFP-CENH3 × WT	54 ± 7 (n = 18)	67	164	8 (5)	109 (66)	47 (29)
WT × *GFP-CENH3*	48 ± 6 (n = 13)	96	112	0 (0)	108 (96)	4 (4)
Cross	**Seeds per siliqua**	**Normal seed (%)**	**Total plants analysed**	**Diploids (%)**	**Triploids (%)**	**Aneuploids (%)**
GFP-tailswap × Wa-1 (tetraploid)	21 ± 6 (n = 96)	1.8	41	11 (27)	0	30 (73)

N-TERMINAL DOMAIN MUTATIONS AND HAPLOID INDUCTION

Interspecific complementation experiments have shown that although CENH3 is rapidly evolving, its function is highly conserved. Maheshwari et al. (2015) investigated that natural variation in the sequence of CENH3 affects its critical cellular roles, given that the protein can tolerate significant changes and still be functional for mitosis and promote haploid induction. Arabidopsis cenh3-1 mutants expressing CENH3 proteins from B. rapa, L. oleraceum, V. vinifera and Z. mays were viable and fertile, indicating that CENH3 proteins from distant species can functionally compensate for the essential mitotic and meiotic functions of the native protein. But this functional equivalence was context-specific. Competitive crosses with wild-type Arabidopsis revealed that the chromosomes from the closely related species B. rapa and L.

oleraceum exhibited lower transmission efficiency, with aneuploid and haploid progeny having mostly the wild-type genome (Maheshwari et al., 2015).

The results also suggest that the CENH3 function in chromosome stability and competitive fitness are separable. The N-terminal domain (NTD) is thought to play a role in the competitive ability of CENH3, but is not essential for its core mitotic role. Plants with the *L. oleraceum* NTD but with the rest of the *Arabidopsis* genome were viable and fertile when self-pollinated, but had a high level of seed lethality and higher rates of haploid and aneuploid progeny when crossed with wild-type plants. The reciprocal construct with the NTD from Arabidopsis and the HFD from L. oleraceum did not yield any detectable haploids in the progeny tested, but some aneuploids were observed (Maheshwari et al., 2015). In this way, the competitive behaviour of centromeres during hybridization can be modified by sequence variations in the CENH3 NTD, without affecting the basic chromosome-segregation function of CENH3.

The strong sequence difference of the NTD offers a reasonable explanation for this differential centromeric competitiveness. The NTDs of L. oleraceum and A. thaliana are nearly equal in length, but diverge at 29 of 82 positions, indicating significant evolutionary divergence in this region. Interestingly, the NTD of histone H3.3, although it has only partial sequence similarity to CENH3, can substitute for the mitotic and meiotic function of CENH3, and in the absence of the GFP tag, shows weak haploid-inducing activity. These observations together imply that the NTD is not just a necessary structural part of CENH3, but that its sequence composition might affect the strength of the centromere, the competitive transmission of chromosomes and susceptibility to chromosome elimination during intergenomic competition. It is however an interesting question whether individual amino acid substitutions within the NTD are enough to regulate the strength of CENH3 and haploid induction.

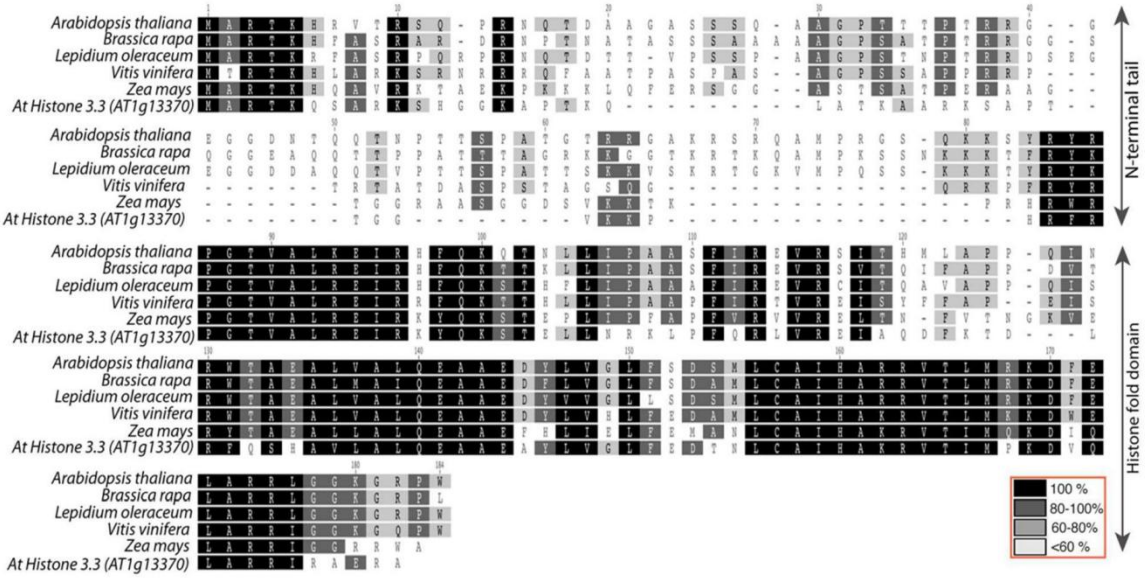


FIGURE 2: Alignment of Arabidopsis thaliana histone3.3 vs. CENH3 sequences from A. thaliana, Brassica rapa, Lepidium oleraceum, Vitis vinifera, and Zea mays. Top two rows represent N-terminal tail with less conservation and the bottom three columns represent relatively well-conserved histone fold domain. (Reproduced from Britt and Kuppu, 2016 under the terms of the Creative Commons Attribution License CC-BY).

HISTONE FOLD DOMAIN MUTATIONS AND HAPLOID INDUCTION

The tail region of CENH3 is rapidly evolving in both length and sequence whereas the histone fold domain is well conserved. This suggests that subtle changes in the HFD might have substantial effects on its activity. Recently, it has been shown that point mutations resulting in single amino acid substitutions induced uniparental genome elimination, resulting in haploidy or aneuploidy, with retention of the wild-type chromosomes (Karimi-Ashtiyani et al., 2015; Kuppu et al., 2015). Remarkably, changes in the majority of conserved residues tested resulted in plants, which were viable, fertile, and, to various degrees, haploid-inducing. None of the tested residues were essential to mitotic or meiotic function. The single amino acid substitutions P82S, G83E, A86V, A132T, L130F, and A136T (Arabidopsis CENH3) induced haploids (among 0.6–12 percent of progeny, depending on the mutation), while P102S and G173E appeared to be fully competitive with wild-type CENH3. Aneuploid progeny were also obtained from among the progenies of these plants. Because these mutations could be obtained in other crop plants by ethyl methane sulfonate (EMS) mediated mutagenesis (the sites tested were conserved among crops), this discovery suggests that non-transgenic haploid inducers could be inducedand may already exist, undetected- in EMS mutagenized crop species.

Table 1: Characterization of engineered and TILLING CENH3 point mutation lines in Arabidopsis thaliana, detailed by specific codon transitions, amino acid alterations, post-zygotic embryo abortion rates, and resulting maternal haploid induction efficiency. (Reprinted from Kuppu et al., 2015, licensed under Creative Commons Attribution 4.0 International).

Line	Codon change	Amino acid change	Aborted seeds (%)	Haploids/Total progeny (%)
WT-HFD#1	No change	No change	0	0/199 (0)
WT-HFD#10	No change	No change	0	0/243 (0)
WT-HFD#15	No change	No change	0	0/163 (0)
M1#6	CCA -> TCA	P82S	15	8/334 (2.4)
M1#8	CCA -> TCA	P82S	21	2/72 (2.7)
M1#11	CCA -> TCA	P82S	20	11/435 (2.5)
M4#16	GGA -> GAG	G83E	36	20/164 (12.2)
M4#18	GGA -> GAG	G83E	28	18/197 (9.1)
M10#6	CCG -> TCC	P102S	10	0/203 (0)
M10#19	CCG -> TCC	P102S	0	0/115 (0)
M21#2	GCT -> ACG	A132T	4	3/475 (0.63)
M21#2	GCT -> ACG	A132T	10	1/163 (0.61)
M26#4	GCA -> ACA	A136T	24	7/309 (2.26)
M47#15	GGA -> GAA	G173E	0	0/207 (0)
TILLING	GCT -> GTT	A86V	32	3/110 (2.72)
M7 # 3	GCT -> GTT	A86V	33	9/232 (3.87)



FIGURE 3: Schematic representation of transgenic CENH3 point mutant transformation and crossing. (A) Steps involved in generation of transgenic CENH3 point mutants in cenh3-1/cenh3-1 background indicated in the red box. (B) CENH3 point mutants in cenh3-1/cenh3-1 background in Col-0 ecotype with trichomes (GL1/GL1) were used as female parent and crossed by gl1-1/gl1-1 plants that carry wild-type allele for CENH3 (CENH3/CENH3) in Landsberg erecta ecotype. The possible outcomes and their phenotype are represented below. Trichomes are represented in dark green. Uniparental paternal haploids do not have trichomes a feature that was used for identification of haploids. tg CENH3* stands for transgenic CENH3 with point mutations. (Reprinted from Kuppu et al., 2015, licensed under Creative Commons Attribution 4.0 International).

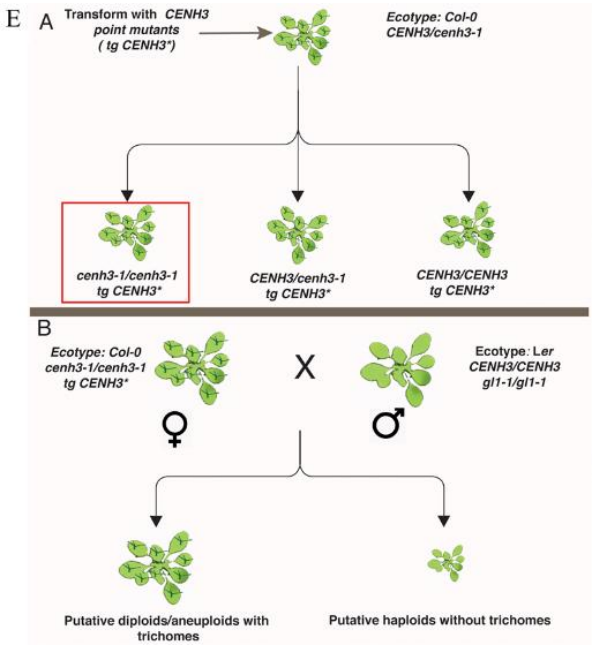


FIGURE 4: Phenotypic differences between diploid (2n) and haploid (n) Arabidopsis thaliana plants generated via CENH3 point mutant genome elimination. (A) Comparison of diploid hybrid with trichomes on the left and smaller haploid plant without trichomes on the right from haploid inducing cross. (B-C) Inset enlargements (region circled in red) show a diploid plant with trichomes on the left and a trichomeless haploid on the right. (D) Phenotype of a Ler gl1-1 haploid with undeveloped siliques on the right and a diploid Ler gl1-1 showing healthy siliques on the left. (Reprinted from Kuppu et al., 2015, licensed under Creative Commons Attribution 4.0 International).

PARENTAL DIRECTIONALITY AND THE EFFICIENCY OF GENOME ELIMINATION

Three different classes of modifications of CENH3 have been shown to produce haploids: Domain swapping with addition of fluorescent tag, complementing with CENH3 from different species, and point mutations in the highly conserved Histone Fold Domain. It is strikingly evident from all three cases that the HI works well when the CENH3 mutants are used as female parent, but not as a male. In case of GFP-tailswap the haploid inducing frequency is around 40 percent when it is used as a female parent, but drops down to 5% when used as a male (Ravi and Chan, 2010). Karimi-Ashtiyani et al. (2015) found that the point mutation L130F induced haploids only used as female parent (at the scale tested). This does not rule out the possibility of using CENH3 mediated HI system as a male, but it is clear that modification of CENH3 has a much stronger haploid-inducing effect when used as a female. The biological significance of this effect is unclear; possibly a higher level of wt CENH3 transcript in the female gametophyte helps to negate the effect of modified CENH3s on loading of additional centromeric factors. However, see below for a discussion of the phenomenology of CENH3 loading and active unloading in the gametophyte and zygote.

CENH3 DEPLETION IN ZYGOTES AND EMBRYOS

In a classical barley HI system involving crosses between Hordeum vulgare × Hordeum bulbosum, the H. bulbosum genome is lost during embryogenesis (Kasha and Kao, 1970). Sanei et al. (2011) were able to cytologically distinguish between the vulgare and bulbosum genomes via FISH, while simultaneously following deposition of CENH3 via immunolocalization (employing an antibody which does not distinguish between the CENH3s encoded by each species). In their visualization of anaphase mitoses in early embryos, the authors found that lagging chromosomes were, unsurprisingly, bulbosum-derived. More notably, these laggards lacked detectable CENH3, which is easily observed via immunolocalization in normal barley chromosomes. This indicates that there is minimally- a failure of CENH3 to reload onto the bulbosum chromosomes, rather than (or perhaps in addition to), a failure of more distal centromeric components to recognize the bulbosum-derived centromeres. It is not clear in the barley system described above whether CENH3 is lost from the bulbosum chromosomes due to dilution (with a failure to reload) after replication, or due to active unloading of CENH3 in the zygote. Interestingly, Ingouff et al. (2010) have shown- in Arabidopsis- that neither CENH3 transcripts nor CENH3-GFP (tagged at the C terminus) were detectable in the Arabidopsis egg cell. Although the authors observed localization of CENH3-GFP to sperm centromeres, this mark was rapidly lost after fertilization of wild-type eggs, prior to further cell division, in what is apparently the active removal of this mark. Localization of CENH3-GFP was visible later in embryogenesis. This very provocative observation raises the possibility that CENH3 is not present in the zygote. If this were true for CENH3, as well as for CENH3-GFP, this observation would challenge the model that centromeres are determined simply by the presence or absence of CENH3. It remains to be determined, however, whether wild-type CENH3 is similarly actively depleted from the zygote. CENH3-GFP may behave differently than CENH3; it does not, for example, complement the cenH3–/– mutation (Ravi et al., 2010). The observations with point mutants that the behavior of chromosomes in the zygote is determined by the amino acid sequence of the CENH3 protein present in the parental linesis most easily explained by assuming that significant levels of parental CENH3 are retained in both the gametes and

the zygote. However, it is also conceivable that CENH3 is truly absent, and that modification of CENH3 results in some sort of modified- and therefore "weak"- zygotic signal, or "footprint," at the centromeric locus. It is possible that CENH3 is not effectively reloaded onto these weak footprints.

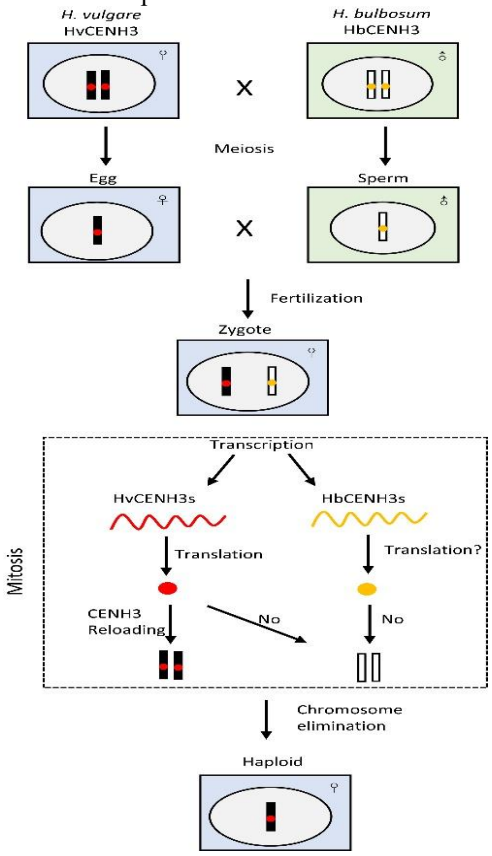
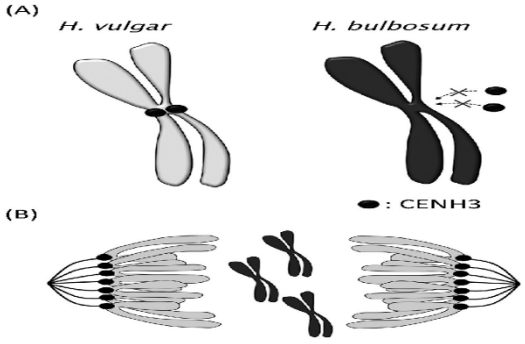


FIGURE 4: Schematic illustration of the molecular model for uniparental chromosome elimination in interspecific Hordeum vulgare × Hordeum bulbosum hybrids. The selective degradation of the paternal (H. bulbosum) genome is triggered by a post-zygotic failure in HbCENH3 protein translation or loading, resulting in inactive centromeres that fail to attach to the mitotic spindle. (Reprinted from Ren et al., 2017, licensed under Creative Commons Attribution 4.0 International).

APPLICATIONS OF HAPLOIDS IN PLANT BREEDING AND GENETIC ANALYSIS

CYTOPLASMIC AND NUCLEAR GENOME EXCHANGE



In Arabidopsis, both maternal and paternal haploids containing wild-type chromosomes and maternal cytoplasm can be generated using CENH3-mediated haploid inducers as the male or female parent. Ravi et al. (2014) developed a cenh3-1 GFPtailswap haploid inducer with Ler cytoplasm, Ler-cytoplasmic haploid inducer (HI). When pollinating Ler-cytoplasmic HI with pollen from wild type with Col-0 cytoplasm, Col-0 WT, haploids with Col-0 WT chromosomes and Ler cytoplasm are generated. This method can be used to develop any combination of cytoplasmic and nuclear genomes by transferring the male nuclear genome into a heterologous cytoplasm rapidly and conveniently. This facilitates the production of new cytoplasmic male sterile (CMS) lines for F1 hybrid seed production. If the haploid inducer line has CMS background, pollinating this haploid inducer line with different inbred lines generates paternal haploids, which carry CMS. One or a few paternal haploids need to be pollinated with pollen from the maternal inbred to produce a new diploid CMS line. Using paternal haploids for cytoplasmic conversions has three distinct advantages: (i) only two generations are needed; (ii) the new CMS line has 100% of the genomes of the inbred line; and (iii) chromosome doubling is not required (Weber, 2014). This method has been employed in maize using the ig1 system for quite a while (Evans, 2007).

REVERSE BREEDING

Hybrid seed is traditionally produced from a cross between two inbred lines. Dirks et al. (2009) proposed a novel plant breeding technology, reverse breeding, which can directly generate parental inbred lines from any hybrid. There are three steps required for reverse breeding: (i) inhibition of meiotic crossover in F1 plants to produce gametes containing combinations of nonrecombinant parental chromosomes, (ii) generation of DH lines via in vitro unfertilized ovule or anther culture and (iii) regeneration of the original hybrid through crossing DH lines with complementary sets of parental chromosomes. Reverse breeding has been tested in Arabidopsis by Wijnker et al. (2012). Firstly, they crossed Landsberg (Ler-0) and Columbia erecta (Col-0) to develop an F1 hybrid. In the hybrid, the meiosis crossover is suppressed using RNAi to knock-down the DMC1 gene, which is required for the crossover formation during meiosis. Secondly, they crossed this hybrid to the centromere mediated haploid inducer line to generate haploids which were doubled into DH lines through spontaneous doubling. Genetic analysis of 69 DH lines using SNP markers at approximately 4-Mb intervals showed absence of recombination. Lastly, they recovered the original hybrid by crossing complementing DH lines. Wijnker et al. (2014) proposed a procedure of reverse breeding in five steps: (i) the generation of DMC1:RNAi transgenic lines (achiasmatic parental lines); (ii) development of achiasmatic hybrids; (iii) haploid induction by crossing to GFP-tailswap; (iv) generation of DH lines by self-pollination of haploids; and (v) recreation of original hybrids by crossing DH lines with complementary sets of parental chromosomes. Successful reversing of breeding in Arabidopsis and availability of centromere-mediated haploid induction technology make it possible to apply this technology to other crops.

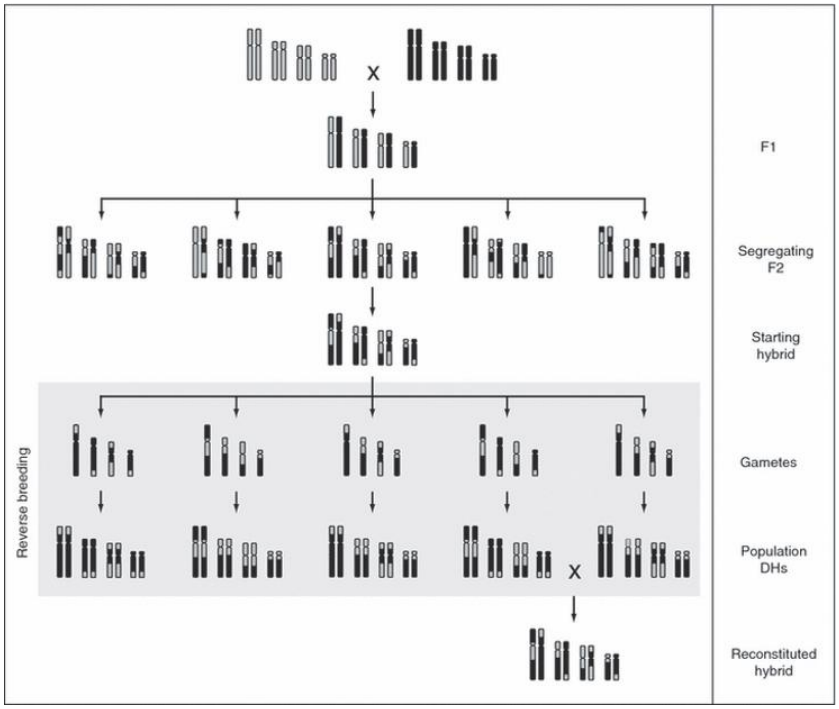


FIGURE 6: Genetic workflow of reverse breeding and the generation of chromosome substitution lines from a heterozygous hybrid. A fully heterozygous f1 starting hybrid featuring distinct parental chromosomes (light gray and black) undergoes engineered meiosis with suppressed crossover recombination (grey box). The resulting non-recombinant gametes are converted into doubled haploids (DHs), yielding completely homozygous individuals and precise chromosome substitution lines. Backcrossing BC1 these substitution lines with a homozygous parental line allows for the targeted creation of hybrids that are heterozygous for only a single specific chromosome, facilitating individual chromosome analysis and background interaction studies. (Reprinted from Dirks et al., 2009, licensed under Creative Commons Attribution 4.0 International).

GENE STACKING THROUGH BIPARENTAL CROSSES

Nowadays, introgression of one or a limited number of genes into elite inbreds by marker-assisted backcrossing is routine in plant breeding (Lubberstedt and Frei, 2012). At the end of backcross € programmes, a heterozygous plant is selfed to produce a fixed line. For single gene introgression, the expected probability of individuals with desired homozygous genotype is 1/4. The frequency of expected genotypes decreases exponentially following the formula 1/4n, where n is the number of independently segregating genes (Lubberstedt and Frei, 2012; Ravi et al., 2014; Shen et al., 2015). In contrast, haploid target genotypes are generated with a frequency of 1/2n. For example, for five loci, the frequency of the desired homozygous genotype is 1/1024 in selfed diploid progeny and 1/32 in haploid progeny. Application of doubled haploids thus significantly reduces the population size required to find desirable genotypes.

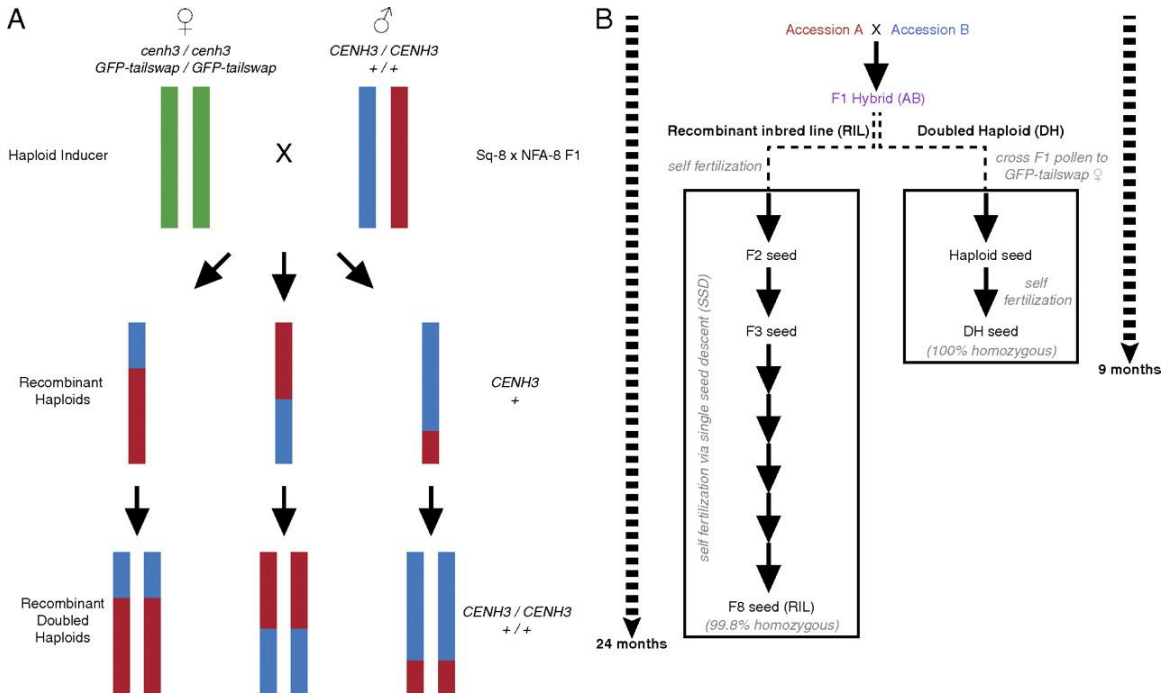


FIGURE 7: Workflow and breeding timelines of centromere-mediated genome elimination for doubled haploid production. (A) Crosses between a maternal haploid inducer line: green) and a wild-type heterozygous F1 pollen donor (blue/red) trigger paternal-specific genome elimination during early embryogenesis. Meiotic crossovers within the pollen donor yield recombinant haploids, which can subsequently be chromosome-doubled to produce fully homozygous recombinant doubled haploid (DH) lines. (B) Timeline comparison between the traditional recombinant inbred line (RIL) approach via single-seed descent, which requires approximately 24 months to reach near-homozygosity F8, and the centromere-mediated DH technology, which delivers complete homozygosity within 9 months. (Reprinted from Seymour., 2009, licensed under Creative Commons Attribution 4.0 International).

RESEARCH ADVANCES AT PUNJAB AGRICULTURAL UNIVERSITY

By Karminder et al. 2020.The OsCENH3 locus in rice, is conserved and present on chromosome 5 as a single copy gene. Comparative studies of OsCENH3, 73 wild type accessions, sequence and protein structure of this gene was studied and compared to CENH3 sequences of other important crop species revealed that rice has no effective paralogs of this gene and thus it is a single copy gene.24 accessions were sequenced completely, phylogenetic analysis revealed that CENH3 from O.rufipogon accessions cluster were closer to OsCENH3. Next step was manipulating the OsCENH3 gene, 2 TILLING mutants; C1217T and G1684A had comparable point mutants in Arabidopsis, 50% seed death observed. Genome editing by developing 4 different constructs, 20 putative transgenics were obtained from the plant transformation Facility at the University of California, Davis ,USA.

REGULATORY ISSUES AND POSSIBLE ALTERNATIVES

In a CENH3-mediated HI system, all chromosomes in the haploid product are, from a molecular point of view, non-transgenic, regardless of whether the haploid-inducing parent carried a transgenically modified CENH3 allele. From a regulatory perspective, however, it is interesting to note that HI through modification of CENH3 could be considered GMO or nonGMO based on interpretation of the law in the geographical region for which it is being considered. In some parts of the world, the genetically modified organisms are labeled based on the product- the genotype of the plant itself- whereas in other parts GMOs may be identified based on the process employed to generate that plant (Hartung and Schiemann, 2014). GFPtailswap based HI would be considered non-GMO if the product only is taken into account whereas it would be considered GMO if process of generation is an important criterion. Similarly, if CENH3 from other species were introduced through molecular tools, the resulting new varieties would be considered GMO if process of generation is considered but non-GMO if the product is considered instead. But if the same alien gene were introgressed by traditional crossing approach it would be considered non-GMO under both conditions. Defining a variety as “GMO” precipitates a costly and time-consuming approval process. Single amino acid substitutions obtained through point mutations generated by EMS is currently accepted as non-GMO under both interpretations of the law. Therefore, under existing regulations, chemically-induced point mutations of CENH3 offer a non-transgenic process for generation of haploids. Additionally, because alien CENH3s or GFP-tailswap are only functional in a cenH3–/– homozygote, identification of point-mutation-based haploid inducers may be faster than other two approaches if robust TILLING populations are already available.

CONCLUSION

Haploid induction technologies- and CENH3-based HI is only one among many-significantly reduces the time and labor required to derive new true-breeding varieties. Within the short period of its discovery, CENH3-mediated genome eliminations have also been shown to be useful in trait mapping (Seymour et al., 2012), reverse breeding (Wijnker et al., 2012), and a variety of other applications (Ravi et al., 2014), all of which can be applied to crop breeding. In spite of the great success with various applications in model plant Arabidopsis, CENH3- mediated genome elimination needs to be tested in other crop species– its application to maize, described in this volume (Kelliher et al., 2016) is very encouraging. The possible delay in implementing this technology (first published in 2010) may be due to the lack of CENH3 knockouts in other species. The recent development of efficient CRISPR-based gene targeting conveniently addresses this issue. GFP-tailswap is a simple and straightforward approach for plants with well-annotated and sequenced genomes. The alternative approaches to CENH3- mediated HI may also be employed. Non-GMO point mutants can be obtained by EMS mutagenesis, and may already be present in existing populations. The recent findings of complementation and HI by CENH3 from related species (Maheshwari et al., 2015) offers an alternative non-GMO approach for species in which EMS mutagenesis is inefficient. The mechanism, through which CENH3 modification affects competitive crosses, without affecting self-pollinations, mitosis, or meiosis, remains a mystery. The elucidation of this process will not only provide insight into CENH3 function and the mechanism of chromosome assortment, but may reveal additional targets that will, upon modification, provide the next generation of haploid inducers.

CONFLICT OF INTEREST

All authors declared that there is no conflict of interest.

REFERENCES

1. Allshire, R. C., and Karpen, G. H. (2008). Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nat. Rev. Genet. 9, 923–937. doi: 10.1038/nrg2466
2. Britt AB and Kuppu S (2016) Cenh3: An Emerging Player in Haploid Induction Technology. Front. Plant Sci. 7:357. doi: 10.3389/fpls.2016.00357
3. Buchwitz, B. J., Ahmad, K., Moore, L. L., Roth, M. B., and Henikoff, S. (1999). Cell division: a histone-H3-like protein in C. elegans. Nature 401, 547–548. doi: 10.1038/44062
4. Cutter, A. R., and Hayes, J. J. (2015). A brief review of nucleosome structure. FEBS Lett. 589, 2914–2922. doi: 10.1016/j.febslet.2015.05.016
5. Dirks, R., Van Dun, K., De Snoo, C.B., Van Den Berg, M., Lelivelt, C.L.C., Voermans, W., Woudenberg, L., De Wit, J.P.C., Reinink, K., Schut, J.W., Van Der Zeeuw, E., Vogelaar, A., Freymark, G., Gutteling, E.W., Keppel, M.N., Van Drongelen, P., Kieny, M., Ellul, P., Touraev, A., Ma, H., De Jong, H. and Wijnker, E. (2009), Reverse breeding: a novel breeding approach based on engineered meiosis. Plant Biotechnology Journal, 7: 837-845. https://doi.org/10.1111/j.1467-7652.2009.00450.x
6. Dunwell, J. M. (2010). Haploids in flowering plants: origins and exploitation. Plant Biotechnol. J. 8, 377–424. doi: 10.1111/j.1467-7652.2009.00498.x
7. Dwivedi, S. L., Britt, A. B., Tripathi, L., Sharma, S., Upadhyaya, H. D., and Ortiz, R. (2015). Haploids: constraints and opportunities in plant breeding. Biotechnol. Adv. 33, 812–829. doi: 10.1016/j.biotechadv.2015.07.001
8. Fleischmann, A., Michael, T. P., Rivadavia, F., Sousa, A., Wang, W., Temsch, E. M., et al. (2014). Evolution of genome size and chromosome number in the carnivorous plant genus Genlisea (Lentibulariaceae), with a new estimate of the minimum genome size in angiosperms. Ann. Bot. 114, 1651–1663. doi: 10.1093/aob/mcu189
9. Gibcus, J. H., and Dekker, J. (2013). The hierarchy of the 3D genome. Mol. Cell 49, 773–782. doi: 10.1016/j.molcel.2013.02.011
10. Guha, S., and Maheshwari, S. C. (1964). In vitro production of embryos from anthers of datura. Nature 204, 497–497. doi: 10.1038/204497a0
11. Hartung, F., and Schiemann, J. (2014). Precise plant breeding using new genome editing techniques: opportunities, safety and regulation in the EU. Plant J. 78, 742–752. doi: 10.1111/tpj.12413
12. Henikoff, S., and Smith, M. M. (2015). Histone variants and epigenetics. Cold Spring Harbor Perspect. Biol. 7:a019364. doi: 10.1101/cshperspect.a019364
13. Howman, E. V., Fowler, K. J., Newson, A. J., Redward, S., MacDonald, A. C., Kalitsis, P., et al. (2000). Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. Proc. Natl. Acad. Sci. U.S.A. 97, 1148–1153. doi: 10.1073/pnas.97.3.1148
14. Ingouff, M., Rademacher, S., Holec, S., Šoljic, L., Xin, N., Readshaw, A., et al. ´ (2010). Zygotic resetting of the HISTONE 3 variant repertoire participates in epigenetic reprogramming in Arabidopsis. Curr. Biol. 20, 2137–2143. doi: 10.1016/j.cub.2010.11.012
15. Karimi-Ashtiyani, R., Ishii, T., Niessen, M., Stein, N., Heckmann, S., Gurushidze, M., et al. (2015). Point mutation impairs centromeric CENH3 loading and induces haploid plants. Proc. Natl. Acad. Sci. U.S.A. 112, 11211–11216. doi: 10.1073/pnas.1504333112
16. Kasha, K. J., and Kao, K. N. (1970). High frequency haploid production in barley (Hordeum vulgare L.). Nature 225, 874–876. doi: 10.1038/225874a0
17. Kelliher, T., Starr, D., Wang, W., McCuiston, J., Zhong, H., Nuccio, M. L., et al. (2016). Maternal haploids are preferentially induced by CENH3- Tailswap transgenic complementation in maize. Front. Plant Sci. 7:414. doi: 10.3389/fpls.2016.00414

18. Kuppu, S., Tan, E. H., Nguyen, H., Rodgers, A., Comai, L., Chan, S. W. L., et al. (2015). Point mutations in centromeric histone induce post-zygotic incompatibility and uniparental inheritance. PLoS Genet. 11:e1005494. doi:10.1371/journal.pgen.1005494
19. Lermontova, I., Koroleva, O., Rutten, T., Fuchs, J., Schubert, V., Moraes, I., et al. (2011). Knockdown of CENH3 in Arabidopsis reduces mitotic divisions and causes sterility by disturbed meiotic chromosome segregation. Plant J. 68, 40–50. doi: 10.1111/j.1365-313X.2011.04664.x
20. Maheshwari, S., Tan, E. H., West, A., Franklin, F. C. H., Comai, L., and Chan, S. W. L. (2015). Naturally occurring differences in CENH3 affect chromosome segregation in zygotic mitosis of hybrids. PLoS Genet. 11:e1004970. doi: 10.1371/journal.pgen.1004970
21. Malik, H. S., and Henikoff, S. (2003). Phylogenomics of the nucleosome. Nat. Struct. Mol. Biol. 10, 882–891. doi: 10.1038/nsb996
22. Marshall, O. J., Chueh, A. C., Wong, L. H., and Choo, K. A. (2008). Neocentromeres: new insights into centromere structure, disease development, and karyotype evolution. Am. J. Hum. Genet. 82, 261–282. doi: 10.1016/j.ajhg.2007.11.009
23. Mattiroli, F., D'Arcy, S., and Luger, K. (2015). The right place at the right time: chaperoning core histone variants. EMBO Rep. 16:e201540840. doi:10.15252/embr.201540840
24. Moraes, I. R., Lermontova, I., and Schubert, I. (2011). Recognition of A. thaliana centromeres by heterologous CENH3 requires high similarity to the endogenous protein. Plant Mol. Biol. 75, 253–261. doi: 10.1007/s11103-010-9723-3
25. Müller, S., and Almouzni, G. (2014). A network of players in H3 histone variant deposition and maintenance at centromeres. Biochim. Biophys. Acta (BBA)- Gene Regul. Mech. 1839, 241–250. Pellicer, J., Fay, M. F., and Leitch, I. J. (2010). The largest eukaryotic genome of them all? Bot. J. Linn. Soc. 164, 10–15. doi: 10.1111/j.1095-8339.2010.01072.x
26. Ravi, M., and Chan, S. W. L. (2010). Haploid plants produced by centromeremediated genome elimination. Nature 464, 615–618. doi: 10.1038/nature08842
27. Ravi, M., Kwong, P. N., Menorca, R. M. G., Valencia, J. T., Ramahi, J. S., Stewart, J. L., et al. (2010). The rapidly evolving centromere-specific histone has stringent functional requirements in Arabidopsis thaliana. Genetics 186, 461–471. doi: 10.1534/genetics.110.120337
28. Ravi, M., Marimuthu, M. P. A., Tan, E. H., Maheshwari, S., Henry, I. M., MarinRodriguez, B., et al. (2014). A haploid genetics toolbox for Arabidopsis thaliana. Nat. Commun. 5:5334. doi: 10.1038/ncomms6334
29. Ravi, M., Shibata, F., Ramahi, J. S., Nagaki, K., Chen, C., Murata, M., et al. (2011). Meiosis-specific loading of the centromere-specific histone CENH3 in Arabidopsis thaliana. PLoS Genet. 7:e1002121. doi: 10.1371/journal.pgen.1002121
30. Ren, J., Wu, P., Trampe, B., Tian, X., Lübberstedt, T. and Chen, S. (2017), Novel technologies in doubled haploid line development. Plant Biotechnol J, 15: 1361-1370. https://doi.org/10.1111/pbi.12805
31. Sanei, M., Pickering, R., Kumke, K., Nasuda, S., and Houben, A. (2011). Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc. Natl. Acad. Sci. U.S.A. 108, E498–E505. doi: 10.1073/pnas.1103190108
32. Sekulic, N., and Black, B. E. (2012). Molecular underpinnings of centromere identity and maintenance. Trends Biochem. Sci. 37, 220–229. doi: 10.1016/j.tibs.2012.01.003
33. Seymour, D. K., Filiault, D. L., Henry, I. M., Monson-Miller, J., Ravi, M., Pang, A., et al. (2012). Rapid creation of Arabidopsis doubled haploid lines for quantitative trait locus mapping. Proc. Natl. Acad. Sci. U.S.A. 4227–4232. doi: 10.1073/pnas.1117277109
34. Steiner, F. A., and Henikoff, S. (2015). Diversity in the organization of centromeric chromatin. Curr. Opin. Genet. Dev. 31, 28–35. doi: 10.1016/j.gde.2015.03.010
35. Steinitz-Sears, L. M. (1963). Chromosome studies in Arabidopsis thaliana. Genetics 48, 483–490. Stimpson, K. M., and Sullivan, B. A. (2010). Epigenomics of centromere assembly and function. Curr. Opin. Cell Biol. 22, 772–780. doi: 10.1016/j.ceb.2010.07.002
36. Stoler, S., Keith, K. C., Curnick, K. E., and Fitzgerald-Hayes, M. (1995). A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. Genes Dev. 9, 573–586. doi: 10.1101/gad.9.5.573
37. Talbert, P. B., Masuelli, R., Tyagi, A. P., Comai, L., and Henikoff, S. (2002). Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. Plant Cell Online 14, 1053–1066. doi: 10.1105/tpc.010425
38. Tan, E. H., Henry, I. M., Ravi, M., Bradnam, K. R., Mandakova, T., Marimuthu, M. P., et al. (2015). Catastrophic chromosomal restructuring during genome elimination in plants. eLife 4:e06516. doi: 10.7554/eLife.06516
39. W¸edzony, M., Forster, B. P., Zur, I., Golemiec, E., Szechynska-Hebda, M., Dubas, E., et al. (2009). "Progress in doubled haploid technology in higher plants," in Advances in Haploid Production in Higher Plants, eds A. Touraev, B. Forster, and S. M. Jain (Amsterdam: Springer), 1–33. doi: 10.1007/978-1-4020-8854-4_1
40. Wang, Z., Yin, H., Lv, L., Feng, Y., Chen, S., Liang, J., et al. (2014). Unrepaired DNA damage facilitates elimination of uniparental chromosomes in interspecific hybrid cells. Cell Cycle 13, 1345–1356. doi: 10.4161/cc.28296
41. Wieland, G., Orthaus, S., Ohndorf, S., Diekmann, S., and Hemmerich, P. (2004). Functional complementation of human centromere protein A (CENP-A) by Cse4p from Saccharomyces cerevisiae. Mol. Cell. Biol. 24, 6620–6630.
42. Wijnker, E., van Dun, K., de Snoo, C. B., Lelivelt, C. L. C., Keurentjes, J. J. B., Naharudin, N. S., et al. (2012). Reverse breeding in Arabidopsis thaliana generates homozygous parental lines from a heterozygous plant. Nat. Genet. 44, 467–470. doi: 10.1128/MCB.24.15.6620-6630.2004