REVIEW

Centromere histone H3- and phospholipase-mediated haploid induction in plants

Song Wang[1], Weiwei Jin[4] and Kai Wang[1,2,3]*

Abstract

Simple and consistent production of haploid is always an appealing pursuit for both crop breeders and researchers. Although diverse strategies have been developed to produce haploids over the past decades, most of them are applicable in only a limited number of plant species. In 2010, Ravi and Chan reported that haploid *Arabidopsis thaliana* plants can be efficiently induced through the introduction of a single genetic alteration in centromere histone H3 (CENH3). Subsequent studies demonstrated that haploids can be efficiently induced either through genetic engineering of CENH3 N-terminal tail or histone fold domain or by replacing CENH3 with an ortholog. The mutation of a pollen-specific phospholipase gene, *MATRILINEAL* (*MTL*) has been revealed to trigger the haploid induction (HI) in maize, which present another promising HI approach by the editing of *MTL* in plant. Here, we review the progress of the CENH3-medialed HI and propose a revised centromere-size model by suggesting a competitive loading process between wild-type and mutant CENH3 during HI. This model can explain both the findings of HI failure when wild-type and mutant *CENH3* genes are coexpressed and the alien centromere loading of CENH3 in stable hybrids. In addition, we review the current understanding of *MTL*-mediated HI in plant. The conservation of *CENH3* and *MTL* in plants indicates wide potential application for HI. We discuss the utility and potential of these two methods in crops by comparing their mechanisms and applications to date in plants. This review will promote the study and application of both CENH3- and *MTL*-mediated haploid induction in plants.

Keywords: Haploid induction, Centromere histone H3, Phospholipase gene, Centromere-size model

Background

The ability to generate haploid plants offers tremendous benefits in plant genetics and genomics, plant breeding, and embryology. Haploid plants can undergo genome doubling to produce genetically stable lines in a single generation; this method bypasses the traditional long-term inbreeding process and can be used in crops where self-pollination is not possible. In addition, haploids will simplify heritability studies due to having only one set of chromosomes, therefore recessive mutations could be easily identified. For these reasons, great efforts have been made over the past few decades to increase the efficiency of haploid production [1]. This review emphasizes on progress of two haploid induction (HI) methods: centromere histone H3 (CENH3)- and phospholipase gene (*MATRILINEAL, MTL*)-mediated. Due to the functional conservation of the *CENH3* and *MTL* genes, both methods promise extensive applicability in plants.

Brief history of plant haploid induction

The first observation of a natural haploid in higher plants was in *Datura stramonium* by Bergner in 1922 [2]. Similar discoveries of haploids soon followed in other species, including *Nicotiana tabacum* [3] and *Triticum compactum* [4]. The potential importance of haploids in crop breeding and genetics was quickly recognized, resulting

*Correspondence: kwang@fafu.edu.cn
[1] Key Laboratory of Genetics, Breeding and Multiple Utilization of Crops, Ministry of Education, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, China
Full list of author information is available at the end of the article

Wang *et al. Plant Methods*   (2019) 15:42

Page 2 of 10

in the first boom in haploid production research. In the following decades, a wide range of haploid generation methods were found, including parthenogenesis (mainly via anther culture), pollen irradiation, seed selection with twin embryos, sparse pollination, alien cytoplasm and wide hybridization [5, 6]. Among these methods, anther culture and wide hybridization demonstrated the most promising.

Early attempts to create haploids from the male gametophyte resulted only in haploid callus tissues [7, 8]. Embryo-like haploids were first described in *Datura* by Guha and Maheshwari [9]. Although these androgenic haploids were not grown to maturity, this landmark research demonstrated the feasibility of haploid production by anther culture. Haploid plants were soon obtained from the cultured anthers of *N. sylvestris* and *N. tabacum* [10]. Thereafter, anther culture systems were successfully established in many plants. Moreover, a series of improved methods were also developed, including culture of microspores, ovaries and ovules [11]. To date, efficient protocols have been established to create haploids by anther culture or its derivatives in approximately 250 plants, including cereals, trees, and vegetables [12–14]. Thus, anther culture has become the preferred method of haploid plant production [13, 15].

Wide hybridization is another efficient way to generate haploids. In 1970, Kasha and Kao reported that the chromosomes of *Hordeum vulgare* were preferentially eliminated during embryo formation when crossing to bulb barley (*H. bulbosum*) [16]. Haploid plants were then obtained by embryo rescue, because the hybrid endosperm would often abort. Its genotype independence makes this method more promising than other haploid-producing methods in barley [17]. To date, over 60 barley cultivars have been produced around the world based on this strategy [13, 17]. Efficient HI was also found in wheat pollinated with maize, sorghum, barley, teosinte, and pearl millet [18–20]. However, this method currently has limited utility, functioning in only those few crops. Although reasons have been proposed for uniparental chromosome elimination, such as asynchronous cell cycle, formation of multipolar spindles, spatial separation of genomes during interphase, and nuclear extrusion [21–23], the actual cellular mechanism remains poorly understood. Based on detailed observation, Finch R.A. revealed that the centromeric constrictions of eliminated *Hordeum* chromosome were always either absent or much smaller than those of the retained chromosomes [24]. This finding provided the first association between centromere function and uniparental chromosome elimination in wide hybridization in *Hordeum*. Strikingly, a recent study indicated that the loss of CENH3 from *H. bulbosum* preceded uniparental chromosome elimination during the development of *H. vulgare × H. bulbosum* hybrid embryos [25]. This discovery of a key role for CENH3 in chromosome elimination suggested the possibility of producing haploids through CENH3 modification.

CENH3-mediated haploid induction

CENH3 (called CENP-A in humans) is a centromere-specific histone H3 variant that replaces histone H3 in eukaryotic centromeric nucleosomes [26, 27]. CENH3 plays an essential role in the epigenetic formation of kinetochore [28, 29] and is sufficient to determine centromere identity, at least in *Drosophila* [30]. The key role of CENH3 in centromeres indicates that any error in transcription, translation, modification, or incorporation can affect the assembly of intact kinetochore and consequently may cause centromere dysfunction or inactivation [31].

In 2010, Ravi and Chan reported the breakthrough discovery that a CENH3 mutant can induce haploids when crossed to WT *Arabidopsis thaliana* [32]. This mutated CENH3 was created by replacing the endogenous CENH3 N-terminal tail with the green fluorescent protein (GFP)-fused N-terminal tail domain of H3.3 (tailswap). This modified *GFP-tailswap CENH3* gene can complement the null mutant *cenh3-1*, in which the native *CENH3* is knocked out. Strikingly, paternal haploid progeny can be obtained at a high rate (25–45% of viable offspring) when *GFP-tailswap* plants (as female) are crossed to WT plants. Haploid progeny can also be obtained (4–5%) when many *GFP-tailswap* anthers are used to pollinate WT plant (due to *GFP-tailswap* plants are almost completely male sterile). Meanwhile, a *GFP-CENH3* mutant (simply adding a GFP at the N-terminus of CENH3) was also shown to induce haploids (5%) when crossed to WT plants [32].

The distinctive feature of this CENH3-mediated HI is that the haploids are easily generated as seeds by crosses using the inducer (*GFP-tailswap* or *GFP-CENH3* mutants) as either female or male. This advantage facilitates wide variety of applications [33], including generating mapping populations [34], chromosomal substitution lines and parental lines for reverse breeding [35] and engineering clonal reproduction through seeds [36]. In addition, it suggests that plants with appropriate modifications of CENH3 can be used as haploid inducers to easily generate haploids. Subsequent studies have revealed that point mutations in the CENH3 α-N-helix or CATD (centromere-targeting domain) can also induce haploids [37, 38], as can the replacement of *Arabidopsis* CENH3 with those related species [39, 40].

Strategies to modify CENH3 for HI
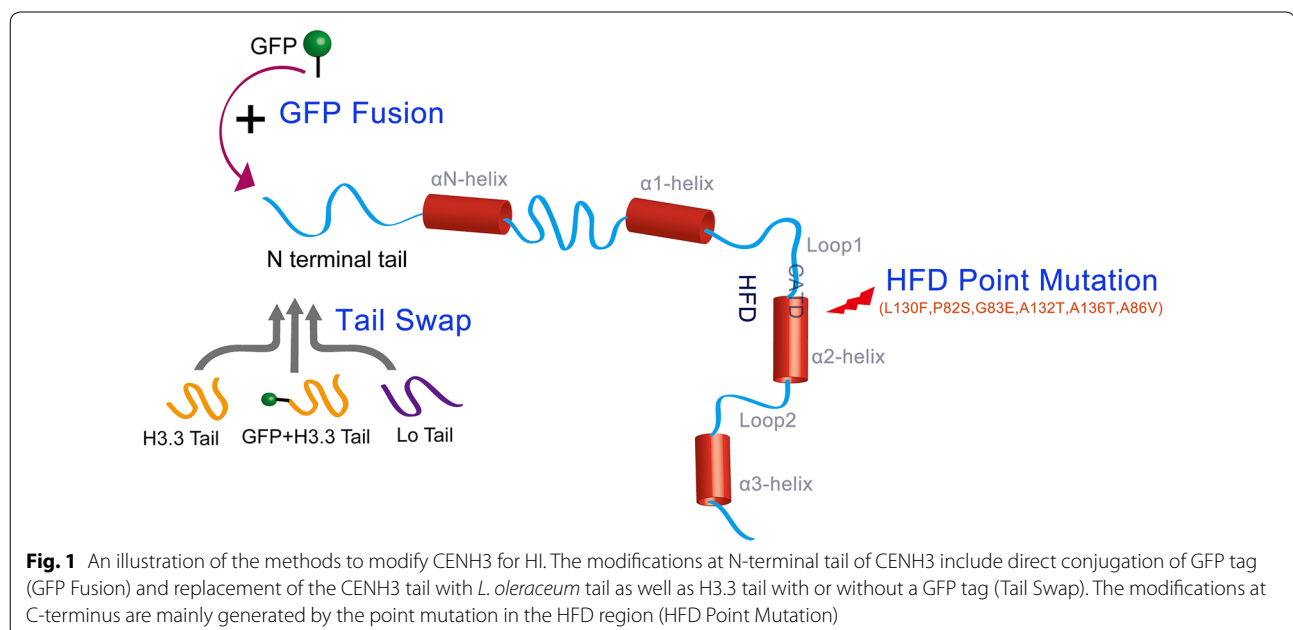
N-terminal tail editing

The initial study of Chan and colleagues suggests that any modifications of CENH3 that can weaken centromere function can be used to produce haploids [32]. Thus, the key is to determine CENH3 regions related to centromere function. It is known that eukaryotic CENH3s have a highly variable N-terminal tail and a relatively conserved C-terminal region [27]. The N-terminal tail is extremely fast-evolving [41] and differs even between CENP-As from different species [42]. N-terminal targeting assays [43] and CENH3 and H3 domain swap experiments (where the regions of CENH3 are replaced by corresponding regions from H3) [32, 44–46] have revealed that the N-terminal tail is dispensable for its mitotic centromere function.

However, exchange of *A. thaliana* N-terminal tail between less related species (*maizetailswap*) will cause severe sterility, i.e., defective centromeres [47], implying that altering the N-terminal tail composition has the potential to create haploid inducer. In addition to the simple conjugation of the GFP tag at the N-terminus (GFP Fusion, Fig. 1), replacement of the N-terminal tail with H3.3 N-terminal tails with or without a GFP tag has been shown to induce haploid formation [32, 47, 48] (Tail Swap, Fig. 1). In addition, replacement of the N-terminal tail of *A. thaliana* CENH3 with that of the mustard family species *Lepidium oleraceum* can induce haploids (Tail Swap, Fig. 1) [40]. Attempts by N-terminal tail editing have been performed in diverse species [49], and HI has been successfully achieved in maize [48], tomato

and rice [50]. Although the HI rates are relatively low (0.065–0.86% in maize, 0.2–2.3% in tomato, and 0.3–1.0% in rice), these experiments demonstrated the feasibility of HI by engineering the CENH3 N-terminal tail in monocotyledonous crop plants.

C-terminal histone fold domain editing

A portion of the C-terminal histone fold domain (HFD) that includes loop1 and the α2 helix, known as CATD (Fig. 1), is essential for centromere targeting [44, 51]. The substitution of CATD enabled the incorporation of an H3 chimera into the centromere in *Drosophila* [45]. Mutations in the CENP-A HFD loop1 residues R80 and G81 and CID (*Drosophila* CENH3) D211 lead to reduced CENP-A and CID retention in their respective centromeres [52, 53]. These results imply that modification of the HFD may create a haploid inducer. This strategy was proven to be feasible by recent studies from two labs [37, 38]. Karimi-Ashtiyani et al. created point mutations of residue L130 in *A. thaliana* CENH3 HFD. The single-insertion line, *Atcenh3* L130F induced haploids when crossed with WT plants (HFD Point Mutation, Fig. 1). Meanwhile, Kuppu et al. adopted a systematic strategy to screen potential HFD residues, determining which can result in zygotic missegregation in combination with WT CENH3. They first identified conserved amino acids from the HFDs of *A. thaliana*, *B. rapa*, *Solanum lycopersicum* and *Zea mays* and predicted functionally tolerated amino acid substitutions. After transgenic function assays, five point-mutations, P82S, G83E, A132T, A136T, and A86 V,

Fig. 1 An illustration of the methods to modify CENH3 for HI. The modifications at N-terminal tail of CENH3 include direct conjugation of GFP tag (GFP Fusion) and replacement of the CENH3 tail with *L. oleraceum* tail as well as H3.3 tail with or without a GFP tag (Tail Swap). The modifications at C-terminus are mainly generated by the point mutation in the HFD region (HFD Point Mutation)

Wang *et al. Plant Methods*   (2019) 15:42

Page 4 of 10

were shown to produce paternal haploids at the rate of 0.61–12.2%.

Unlike the tailswap-CENH3 haploid inducer, all of these point-mutant plants were viable and fully fertile when self-pollinated, without any obvious phenotypic effects [37, 38]. These fully viable and fertile inducer options will facilitate the application of this technique and statistically improve HI rate, because many more fertile gametes are generated from a single plant than from the N-terminal chimeric version of GFP-tailswap (most male sterile and ~60% female fertility) [32].

CENH3 replacement

Although highly variable in sequence, CENH3 has essential functions that are conserved across a broad evolutionary landscape in eukaryotes [40], as first proven by the finding that yeast Cse4 (CENH3) can functionally replace CENP-A in humans [54]. Studies in *Arabidopsis* have also shown functional conservation across a broad evolutionary distance, as the exogenous CENH3s from a wide range of species can target *A. thaliana* centromeres [40, 51, 55]. Interestingly, given this functional complementation, CENH3 from *B. rapa* or *L. oleraceum* were found to be 'noncompetitive' versus WT CENH3, generating haploids, aneuploids and chromosomal rearrangements [40]. This result revealed the possibility of creating a haploid inducer by completely replacing CENH3 with a CENH3 from a related species. Like the HFD point-mutant individuals, the transgenic *Arabidopsis* carrying CENH3 from *B. rapa* or *L. oleraceum* were phenotypically indistinguishable from WT and were also self-fertile, which compensate for its low HI rates (1–11%).

Mechanism of centromere-mediated haploid induction

A consistent phenomenon in the centromere-mediated HI is that the haploid plants contain only WT chromosomes. Moreover, no haploids were obtained from self-fertilized *CENH3* mutant plants and plants with coexpressed WT and mutant *CENH3* genes. These results indicate that the modified CENH3 has weakened centromere function, which leads to uniparental chromosome loss when competing with WT CENH3 during mitosis [32, 39]. Recently, a centromere-size model provided a plausible explanation for the elimination of chromosomes with defective CENH3 [56]. Briefly, defective CENH3 is expected to have a lower efficiency in recruiting some key kinetochore proteins, leading to the smaller centromere sizes in outcrosses. These smaller centromeres have less accurate metaphase plate alignment, causing inefficient recruitment of centromeric factors and leading to a high level of stochastic chromosome loss. The findings that tailswap- and point-mutant

CENH3s lead to reduced centromere loading [37, 47] provide indirect evidence to support the above assumption. Furthermore, when an inducer is self-pollinated, all centromeres are on an equal defective level and thus are competitively retained. Thus, the centromere-size model can also explain the observation that homozygous tailswap-CENH3 and other inducer lines are self-fertile and do not produce haploids. In contrast, small centromeres that can expand to match the average size of other centromeres will be retained in the progeny, producing aneuploid or diploid individuals.

This centromere-size model is derived from the concept that each centromere has a similar size within a species [57, 58]. Zhang and Dawe [57] examined ten species of grass and observed similar immunoassay signal intensities from each centromere that suggested uniform centromere size within a species. Studies through chromatin immunoprecipitation followed by sequencing (ChIP-seq) using an anti-CENH3 antibody confirmed the uniform size of centromeres within a species [59–63]. However, electron microscopy and immunoassay observation analyses revealed that centromere sizes across species are different and correlate with genome size and total centromere volume [57, 64]. Therefore, the similar size of maize and oat centromeres observed in oat-maize addition lines indicates that the alien maize centromeres have adopted a similar size to oat centromeres. This hypothesis was confirmed by the analyses of anti-CENH3 ChIP-seq in eight oat-maize addition lines [59]. Thus, the retention of the maize chromosomes in oat-maize addition lines can be attributed to expansions in centromere size allowing them to match the average size of the other centromeres, whereas failure of the smaller centromeres from one parent to adopt the required size may cause genome elimination in a hybrid.

In fact, during wide hybridizations between a large-genome species (such as oat, barley and wheat) and a small-genome species (such as maize, pearl millet, adlay millet, perennial rye grass or sorghum), chromosomes from the small-genome parent are often eliminated in early embryogenesis [65–67]. Thus, these findings not only support the centromere-size model but also suggest that the centromere-size model can explain the HI caused by wide hybridization.

Competitive CENH3 loading may influence chromosome retention or loss

An intriguing finding is that plants with coexpressed WT and mutant *CENH3* genes (*GFP-tailswap*, *GFP-CENH3* and *Hvßcenh3* L92F) do not act as haploid inducers, because no haploid or hypoploid plants are obtained [32, 37]. According to the centromere-size model, this lack of HI suggests that all the centromeres should have a

uniform size in those hybrid plants. A reasonable explanation is that WT CENH3 may competitively load to all centromeres, thus generating a uniform centromere size and producing diploid progenies. In contrast, if the WT and mutant CENH3s were constrained to load into the corresponding WT- and mutant-derived centromeres, it was expected to see haploid progeny, because the centromeres loaded with mutant CENH3 would be smaller, and they would be eliminated, producing haploid or aneuploid offspring. Therefore, the lack of HI in plants with coexpressed WT and mutant *CENH3* genes suggests the possibility that a competitive loading process between WT and mutant CENH3s occurs (Fig. 2). In oat-maize addition lines, the additional maize centromeres are consistently incorporated by oat CENH3 [68] and adopt a similar size to the oat centromeres [59]. In addition, HTR12 (*A. thaliana* CENH3) can be incorporated into the centromeres of *A. arenosa* in an allopolyploid of *A. thaliana* and *A. arenosa* [69]. All these findings support our hypothesis.

The WT and mutant-derived centromeres should retain their corresponding WT and modified CENH3s in early zygotic mitoses, because CENH3 can be sustained through mitosis after the initial steps of centromere assembly [70, 71] (Fig. 2). These defective mutant CENH3s continuously target the inducer-derived centromeres and produce smaller centromeres due to their inefficient centromere loading procedure. These smaller centromeres are noncompetitive and are later eliminated when encountering WT centromeres. However, mutant CENH3 loading must occur in a short time frame early in the zygotic mitoses. Alternatively, the WT CENH3 would competitively load into some or all of the mutant-derived centromeres, producing normal large-sized centromeres and generating normal diploid or aneuploid plants. Clearly, our revised model is consistent with and based on the centromere-size model, but it also presents a plausible explanation for both HI incapability despite co-ocurrence of WT and mutant CENH3s [32, 37] and CENH3 loading in the centromeres of alien species in stable hybrids [68, 69]. This hypothesis is testable through further centromere loading assays involving WT and mutant CENH3s during the early stage of zygotic development.
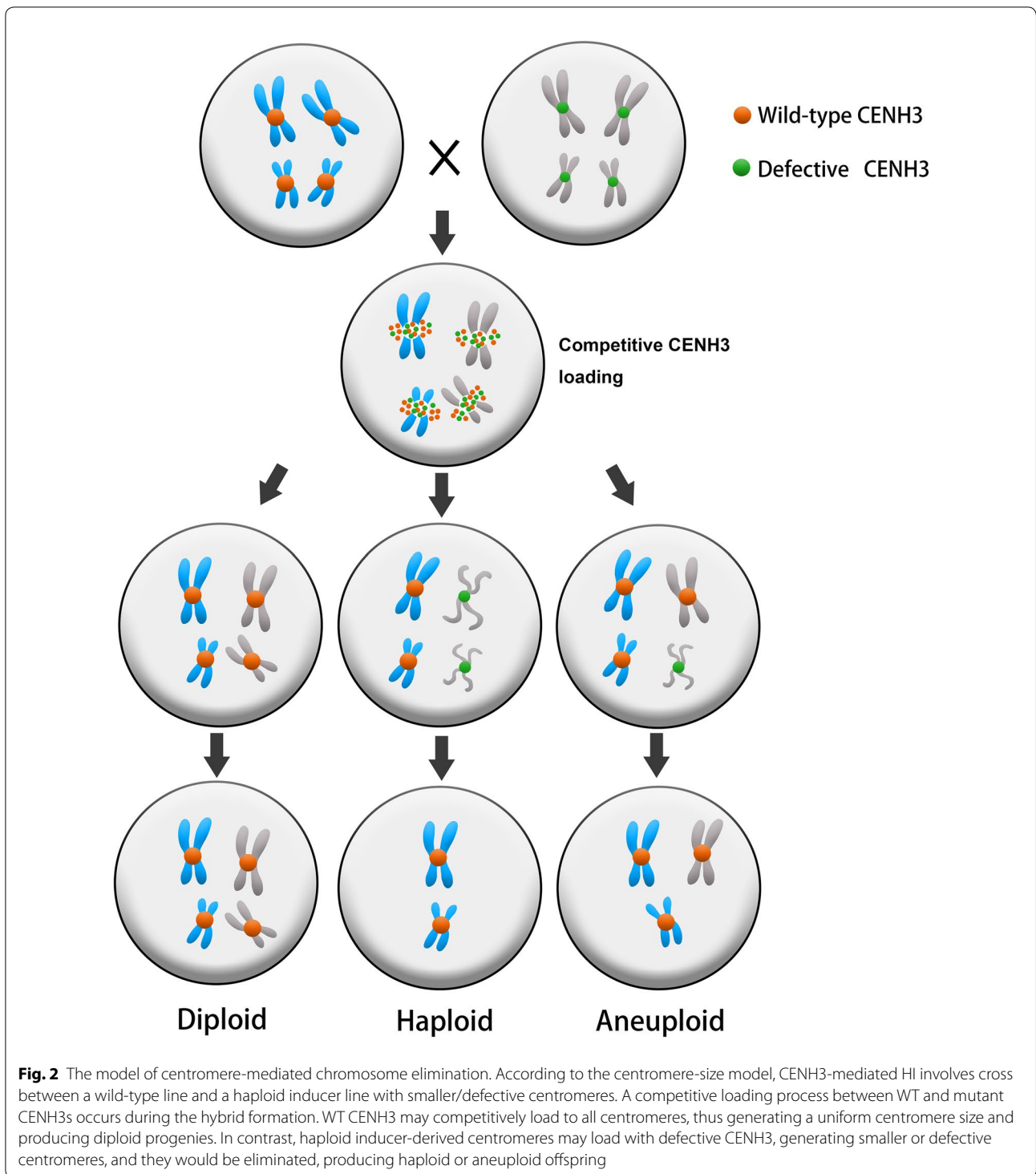
Phospholipase-triggered haploid induction in maize

In maize, the line Stock6 has been widely used to generate haploids for inbred line production. The original Stock6 inducer line was created by Coe in 1959 with HI rates of 2.52% and 1–2% when Stock6 was selfed or outcrossed as a male, respectively [72]. Selective breeding has improved the HI capability of Stock6 derivatives to 7–15% [73, 74]. Several quantitative trait loci (QTLs) associated with HI have been identified, indicating that the HI in Stock6 is controlled by multiple genes [75–79]. Among those candidate QTLs, *qhir1*, which is located in bin 1.04, was shown to explain 66% of the genetic variance and has attracted the most attention.

Researchers from Syngenta conducted the first trial to isolate the HI gene from the *qhir1* region [80]. After fine mapping, they narrowed the QTL to a 0.57 Mb region containing 7 genes. Sequence comparisons revealed that the gene *GRMZM2G471240* (*MATRILINEAL*) has a 4-bp insertion (CGAG) in the induction lines compared to the B73 genome. This 4-bp insertion in the fourth exon leads to a frame shift, causing 20 altered amino acids and a premature transcription termination that truncates the protein by 29 amino acids. After functional verification (gene knockdown and knockout), they verified that *MTL* was responsible for HI. Two other teams from China and France also achieved similar results after sequence assay and gene-editing validation (referred *GRMZM2G471240* as *ZmPLA1* and *NLD*, respectively) [81, 82]. Notably, large-scale sequence comparisons revealed that the 4-bp insertion was a distinct feature restricted to only the inducer lines, confirming its critical role in HI [81]. In addition, the absence of such a 4-bp insertion from the ancestral variety, teosinte, suggested that this mutation occurred after maize domestication [81]. Another interesting finding is that some gene-editing and knockdown events increased HI rates, suggesting that it is possible to create a high-HI-rate inducer by modifying *MTL*.

MTL encodes a patatin-like phospholipase and is expressed specifically in maize pollen [80–82]. Phospholipase alterations are associated with delayed pollen germination and pollen tube growth [83], which explain the pleiotropic phenotypes accompanying with maize HI capability [73, 80]. However, the mechanisms of HI that contribute to single fertilization [84] or postzygotic genome elimination [85, 86] remain unclear. An RNA-seq assay showed that genes associated with pollen cell endomembrane and lipid composition or Ca^{2+}-involved signaling pathways have potential effects on *mtl* HI [80]. A subcellular localization assay revealed that the MTL protein targets the sperm cell plasma membrane in both maize and *Arabidopsis*. In contrast, the truncated protein from the inducer line PK6 was absent from the plasma membrane [80]. Further in silico analysis of the MTL protein demonstrated that the absence of a lipid anchor site from the truncated protein C-terminal may contribute to its mislocalization. Collectively, these results indicate that membrane integrity or mechanisms involving signaling precursor triggered by defective MTL might be responsible for HI capacity [82].

Wang *et al. Plant Methods*   (2019) 15:42

Page 6 of 10

Fig. 2 The model of centromere-mediated chromosome elimination. According to the centromere-size model, CENH3-mediated HI involves cross between a wild-type line and a haploid inducer line with smaller/defective centromeres. A competitive loading process between WT and mutant CENH3s occurs during the hybrid formation. WT CENH3 may competitively load to all centromeres, thus generating a uniform centromere size and producing diploid progenies. In contrast, haploid inducer-derived centromeres may load with defective CENH3, generating smaller or defective centromeres, and they would be eliminated, producing haploid or aneuploid offspring

Recently, Li et al. [87] discovered that approximately 10% of kernels were aborted when pollinated with the inducer line, which has almost no detectable kernel abortion as a female parent. The proportion of viable inducer pollen is lower than that of inbred lines, suggesting that

defective pollen development may be attributed to HI. After single-nucleus sequencing of mature pollen, they demonstrated that chromosome fragmentation that begins around pollen mitosis may be the cue causing pollen abortion and HI. Given that the haploid is an abortive

Wang *et al. Plant Methods*   (2019) 15:42

Page 7 of 10

kernel with a haploid embryo, only those that initiate fragmentation after the 2nd mitosis can generate haploid nonaborted kernels [87]. However, detailed embryo development studies still need to be performed, because the possibility of single fertilization (fragmented sperm not fusing with the egg) cannot be excluded [87]. However, the results of Li et al. shed light on either the mechanism or the development of new HI superinducer. As HI is potentially caused by pollen sperm fragmentation after the 2nd mitosis, increasing the frequency of chromosome fragmentation initiated after the 2nd mitosis could potentially improve the HI rate. Alternatively, decreasing early chromosome fragmentation, before the 2nd mitosis, would be an efficient strategy to improve kernel viability.
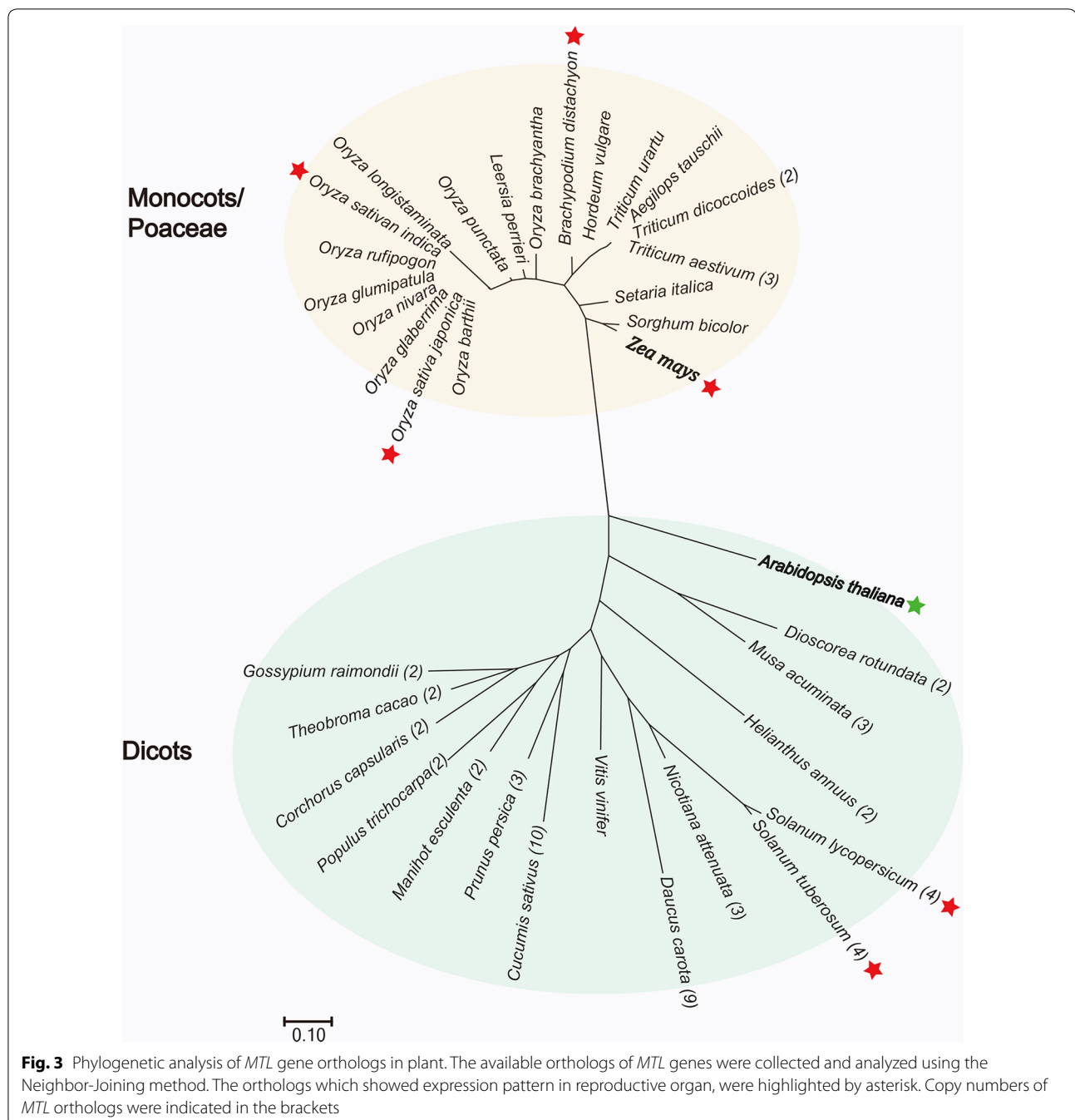
Conclusions and perspectives

The CENH3 function is highly conserved across eukaryotes. Moreover, the centromere is composed of multiple proteins, which means that other key centromere proteins, in addition to CENH3, are also expected to be potential targets for HI inducer development. Thus, CENH3-based HI has tremendous potential for successful application in any plant theoretically. To date, in addition to the model plant *Arabidopsis*, haploids have been successfully induced in maize [48] using the tailswap strategy. However, it is perplexing that the CENH3 tailswap plants in maize did not induce haploids as frequently as in *Arabidopsis* [32]. Due to the enigmatic structure of centromeres [88], functional divergence of CENH3s in other plants compared with the dicot model plant *Arabidopsis* is possible. Supporting the hypothesis, the *Arabidopsis* tailswap inducer shows severe sterility (mostly male sterile and 60–70% female fertile) [32] while that was not observed in maize [48]. It therefore implied that the HI ability might be correlated with the plant fertility. It is plausible that haploid inducer with more severe defective CENH3 will likely cause both more severe sterility and higher HI rate because the mutated CENH3 can not complement the full function of CENH3 and is less competitive when encountering with wild CENH3. In fact, the plant with the highest HI rate was also the plant with the strongest knockdown of *CENH3* in maize [48] supporting the above hypothesis. Therefore, it will be interesting to explore the correlation of inducer sterility and HI rate. Moreover, the usage of WT parent is another factor that should be considered since the HI rates can be nearly two-fold change when different WT lines were crossed with the same inducer in *Arabidopsis* [32]. Theoretically, given that the HI rated are still low using the tailswap strategy in crops, the methods of HFD modification, CENH3 replacement and genetic editing other kinetochore components provide alternative choices for the development of inducers with high HI rates. Importantly, point mutations can be achieved by non-transgenic chemical mutagenesis, and then, the results can be directly used in crop breeding.

The centromere-size model provides a basis to explain HI resulting from both wide hybridization and CENH3-mediated systems. Our CENH3 competitive loading model helps to explain the observation that inducer lines lose HI capability when both WT and mutant CENH3s are present. This hypothesis is expected to be easily tested by further WT and modified CENH3 targeting assays. Based on the centromere-size model, crosses between inducer and larger-centromere lines should produce haploids with high efficiency. Thus, this model will allow us to predict HI capacity before crossing experiments, which will greatly improve the application of HI to crop breeding.

The HI capacity of maize Stock6 is caused by at least seven potential genes (or QTLs) [75, 79], indicating that a potentially enhanced HI inducer could be created by combining multiple HI-related genes. Thus, it is worthwhile to devote effort to the identification of other major HI-associated QTLs/genes, such as *qhir8*, which was found to explain 20% of the genotypic variance [89]. In addition, alteration of the *MTL* gene renders it possible to extend this tool to other species, at least in *Poaceae* [80, 90], because of the high-level conservation of *MTL* orthologs (Fig. 3). Its successful application in rice is very encouraging [91]. However, the presence of numerous co-orthologs [82] and non-pollen-specific expression in dicots [92] (Fig. 3, Additional file 1: Table S1) raises some uncertainty about its utility in those or other plants. Thus, a careful function(s) assessment of orthologs will be required before experimental trials in other species.

Notably, the CENH3-mediated inducer can induce haploids as both female and male. This feature provides it with wide versatility, permitting diploid production from individuals with the desired genomic backgrounds. In addition, crosses using the inducer as a female will transfer the nuclear genome of the male parent to a heterologous cytoplasm, which will potentially facilitate the generation of cytoplasmic male sterile lines [32]. In contrast, the classical *MTL*-mediated Stock6 and its derivatives induce exotic-parent haploids only as male (although haploids can be induced on self-pollination), which will generate only haploids with homologous cytoplasmic and nuclear genomes. Nevertheless, the evidence from maize demonstrates that this issue can be solved by the CENH3-mediated HI system [48], which has potential utility unlimited by plant species.

Fig. 3 Phylogenetic analysis of *MTL* gene orthologs in plant. The available orthologs of *MTL* genes were collected and analyzed using the Neighbor-Joining method. The orthologs which showed expression pattern in reproductive organ, were highlighted by asterisk. Copy numbers of *MTL* orthologs were indicated in the brackets

Additional file

Additional file 1: Table S1. The information of MTL orthologs.

Abbreviations
CENH3: centromere histone H3; HI: haploid induction; *MTL*: *MATRILINEAL*; WT: wild-type; GFP: green fluorescent protein; ChIP-seq: chromatin immunoprecipitation followed by sequencing; HFD: histone fold domain; QTLs: quantitative trait loci.

Authors' contributions
SW and KW conceived this review and drafted the manuscript. WJ commented and revised the manuscript. All authors read and approved the final manuscript.

Author details
[1] Key Laboratory of Genetics, Breeding and Multiple Utilization of Crops, Ministry of Education, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Fujian Agriculture and Forestry University, Fuzhou 350002, Fujian, China. [2] National Engineering Research Center of Sugarcane, Fujian

Wang *et al. Plant Methods*   (2019) 15:42

Page 9 of 10

Agriculture and Forestry University, Fuzhou 350002, China. [3] Key Laboratory of Sugarcane Biology and Genetic Breeding, Ministry of Agriculture and Rural Affairs, Fujian Agriculture and Forestry University, Fuzhou 350002, China. [4] College of Agriculture, China Agricultural University, No. 2, Yuan Ming Yuan West Road, Haidian District, Beijing 100193, China.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Dunwell JM. Haploids in flowering plants: origins and exploitation. Plant Biotechnol J. 2010;8(4):377–424.
2. Blakeslee AF, Belling J, Farnham ME, Bergner AD. A haploid mutant in the jimson weed, "datura stramonium". Science. 1922;55(1433):646–7.
3. Clausen RE, Mann MC. Inheritance in *Nicotiana tabacum*: V. The occurrence of haploid plants in interspecific progenies. Proc Natl Acad Sci USA. 1924;10(4):121–4.
4. Gaines EF, Aase HC. A haploid wheat plant. Am J Bot. 1926;13(6):373–85.
5. Touraev A, Forster BP, Jain SM. Advances in haploid production in higher plants. New York: Springer; 2009.
6. Kasha KJ, Maluszynski M. Production of doubled haploids in crop plants An introduction. In: Maluszynski M, Kasha KJ, Forster BP, Szarejko I, editors. Doubled haploid production in crop plants: a manual. Dordrecht: Springer; 2003. p. 1–4.
7. Yamada T, Shoji T, Sinoto Y. Formation of calli and free cells in the tissue culture of *Tradescantia reflexa*. Bot Mag Tokyo. 1963;76(903):332–9.
8. Jain SM, Sopory SK, Veilleux RE. In vitro haploid production in higher plants, Volume 5—oil, ornamental and miscellaneous plants. Berlin: Springer; 1997.
9. Guha S, Maheshwari SC. In vitro production of embryos from anthers of *Datura*. Nature. 1964;204:497.
10. Bourgin J, Nitsch J. Obtention de *Nicotiana* haploides a partir determines cultivees in vitro. Ann Physiol Veg. 1967;9:377–82.
11. Khush GS, Virmani SS. Haploids in plant breeding. In: Jain SM, Sopory SK, Veilleux RE, editors. In vitro haploid production in higher plants: volume 1—fundamental aspects and methods. Dordrecht: Springer; 1996. p. 11–33.
12. Wędzony M, Forster BP, Żur I, Golemiec E, Szechyńska-Hebda M, Dubas E, et al. Progress in doubled haploid technology in higher plants. In: Touraev A, Forster BP, Jain SM, editors. Advances in haploid production in higher plants. Dordrecht: Springer; 2009. p. 1–33.
13. Maluszynski M, Kasha KJ, Szarejko I. Doubled Haploid production in crop plants: a manual. Dordrecht: Springer; 2003.
14. Germanà MA. Doubled haploid production in fruit crops. Plant Cell, Tissue Organ Cult. 2006;86(2):131–46.
15. Hazarika RR, Mishra KV, Chaturvedi R. In vitro haploid production—a fast and reliable approach for crop improvement. In: Tuteja N, Gill SS, editors. Crop improvement under adverse conditions. New York: Springer; 2013. p. 171–212.
16. Kasha KJ, Kao KN. High frequency haploid production in barley (*Hordeum vulgare L*). Nature. 1970;225:874.
17. Devaux P, Pickering R. Haploids in the improvement of poaceae. In: Don Palmer CE, Keller WA, Kasha KJ, editors. Haploids in crop improvement II. Berlin: Springer; 2005. p. 215–42.
18. Ohkawa Y, Suenaga K, Ogawa T. Production of haploid wheat plants through pollination of sorghum pollen. Jpn J Breed. 1992;42(4):891–4.
19. Riera-Lizarazu O, Mujeeb-Kazi A. Polyhaploid production in the Triticeae: wheat × Tripsacum crosses. Crop Sci. 1993;33(5):973–6.
20. Niroula RK, Bimb HP. Overview of wheat × maize system of crosses for dihaploid induction in wheat. World Appl Sci J. 2009;7(8):1037–45.
21. Gupta SB. Duration of mitotic cycle and regulation of DNA replication in *Nicotiana plumbaginifolia* and a hybrid derivative of *N. tabacum* showing chromosome instability. Can J Genet Cytol. 1969;11(1):133–42.
22. Bennett MD, Finch RA, Barclay IRJC. The time rate and mechanism of chromosome elimination in Hordeum hybrids. Chromosoma. 1976;54(2):175–200.
23. Laurie DA, Bennett MD. The timing of chromosome elimination in hexaploid wheat × maize crosses. Genome. 1989;32(6):953–61.
24. Finch RA. Tissue-specific elimination of alternative whole parental genomes in one barley hybrid. Chromosoma. 1983;88(5):386–93.
25. Sanei M, Pickering R, Kumke K, Nasuda S, Houben A. Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc Natl Acad Sci. 2011;108(33):E498–505.
26. Jiang J, Birchler JA, Parrott WA, Dawe RK. A molecular view of plant centromeres. Trends Plant Sci. 2003;8(12):570–5.
27. McKinley KL, Cheeseman IM. The molecular basis for centromere identity and function. Nat Rev Mol Cell Biol. 2016;17(1):16–29.
28. Burrack LS, Berman J. Flexibility of centromere and kinetochore structures. Trends Genet. 2012;28(5):204–12.
29. Fukagawa T, Earnshaw William C. The centromere: chromatin foundation for the kinetochore machinery. Dev Cell. 2014;30(5):496–508.
30. Mendiburo MJ, Padeken J, Fülöp S, Schepers A, Heun P. *Drosophila* CENH3 is sufficient for centromere formation. Science. 2011;334(6056):686–90.
31. Allshire RC, Karpen GH. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nat Rev Genet. 2008;9(12):923–37.
32. Ravi M, Chan SW. Haploid plants produced by centromere-mediated genome elimination. Nature. 2010;464(7288):615–8.
33. Ravi M, Marimuthu MP, Tan EH, Maheshwari S, Henry IM, Marin-Rodriguez B, et al. A haploid genetics toolbox for *Arabidopsis thaliana*. Nat Commun. 2014;5:5334.
34. Seymour DK, Filiault DL, Henry IM, Monson-Miller J, Ravi M, Pang A, et al. Rapid creation of *Arabidopsis* doubled haploid lines for quantitative trait locus mapping. Proc Natl Acad Sci USA. 2012;109(11):4227–32.
35. Wijnker E, van Dun K, de Snoo CB, Lelivelt CL, Keurentjes JJ, Naharudin NS, et al. Reverse breeding in *Arabidopsis thaliana* generates homozygous parental lines from a heterozygous plant. Nat Genet. 2012;44(4):467–70.
36. Marimuthu MP, Jolivet S, Ravi M, Pereira L, Davda JN, Cromer L, et al. Synthetic clonal reproduction through seeds. Science. 2011;331(6019):876.
37. Karimi-Ashtiyani R, Ishii T, Niessen M, Stein N, Heckmann S, Gurushidze M, et al. Point mutation impairs centromeric CENH3 loading and induces haploid plants. Proc Natl Acad Sci USA. 2015;112(36):11211–6.
38. Kuppu S, Tan EH, Nguyen H, Rodgers A, Comai L, Chan SW, et al. Point mutations in centromeric histone induce post-zygotic incompatibility and uniparental inheritance. PLoS Genet. 2015;11(9):e1005494.
39. Britt AB, Kuppu S. Cenh3: an emerging player in haploid induction technology. Front Plant Sci. 2016;7:357.
40. Maheshwari S, Tan EH, West A, Franklin FC, Comai L, Chan SW. Naturally occurring differences in CENH3 affect chromosome segregation in zygotic mitosis of hybrids. PLoS Genet. 2015;11(1):e1004970.
41. Malik HS, Henikoff S. Phylogenomics of the nucleosome. Nat Struct Biol. 2003;10:882.
42. Goutte-Gattat D, Shuaib M, Ouararhni K, Gautier T, Skoufias DA, Hamiche A, et al. Phosphorylation of the CENP-A amino-terminus in mitotic centromeric

Wang *et al. Plant Methods*   (2019) 15:42

Page 10 of 10

chromatin is required for kinetochore function. Proc Natl Acad Sci USA. 2013;110(21):8579–84.

43. Lermontova I, Schubert V, Fuchs J, Klatte S, Macas J, Schubert I. Loading of *Arabidopsis* centromeric histone CENH3 occurs mainly during G2 and requires the presence of the histone fold domain. Plant Cell. 2006;18(10):2443–51.

44. Black BE, Foltz DR, Chakravarthy S, Luger K, Woods VL Jr, et al. Structural determinants for generating centromeric chromatin. Nature. 2004;430(6999):578.

45. Vermaak D, Hayden HS, Henikoff S. Centromere targeting element within the histone fold domain of Cid. Mol Cell Biol. 2002;22(21):7553–61.

46. Sullivan KF, Hechenberger M, Masri K. Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J Cell Biol. 1994;127(3):581–92.

47. Ravi M, Shibata F, Ramahi JS, Nagaki K, Chen C, Murata M, et al. Meiosis-specific loading of the centromere-specific histone CENH3 in *Arabidopsis thaliana*. PLoS Genet. 2011;7(6):e1002121.

48. Kelliher T, Starr D, Wang W, McCuiston J, Zhong H, Nuccio ML, et al. Maternal haploids are preferentially induced by *CENH3-tailswap* transgenic complementation in maize. Front Plant Sci. 2016;7:414.

49. Tek AL, Stupar RM, Nagaki K. Modification of centromere structure: a promising approach for haploidline production in plant breeding. Turk J Agric For. 2015;39:557–62.

50. Kalinowska K, Chamas S, Unkel K, Demidov D, Lermontova I, Dresselhaus T, et al. State-of-the-art and novel developments of in vivo haploid technologies. Theor Appl Genet. 2019;132(3):593–605.

51. Ravi M, Kwong PN, Menorca RMG, Valencia JT, Ramahi JS, Stewart JL, et al. The rapidly evolving centromere-specific histone has stringent functional requirements in *Arabidopsis thaliana*. Genetics. 2010;186(2):461–71.

52. Tachiwana H, Kagawa W, Shiga T, Osakabe A, Miya Y, Saito K, et al. Crystal structure of the human centromeric nucleosome containing CENP-A. Nature. 2011;476:232.

53. Zhang W, Colmenares Serafin U, Karpen Gary H. Assembly of Drosophila centromeric nucleosomes requires CID dimerization. Mol Cell. 2012;45(2):263–9.

54. Wieland G, Orthaus S, Ohndorf S, Diekmann S, Hemmerich P. Functional complementation of human centromere protein A (CENP-A) by Cse4p from Saccharomyces cerevisiae. Mol Cell Biol. 2004;24(15):6620–30.

55. Moraes IC, Lermontova I, Schubert I. Recognition of *A. thaliana* centromeres by heterologous CENH3 requires high similarity to the endogenous protein. Plant Mol Biol. 2011;75(3):253–61.

56. Wang N, Dawe RK. Centromere size and its relationship to haploid formation in plants. Mol Plant. 2018;11(3):398–406.

57. Zhang H, Dawe RK. Total centromere size and genome size are strongly correlated in ten grass species. Chromosome Res. 2012;20(4):403–12.

58. Liu Y, Su H, Zhang J, Liu Y, Han F, Birchler JA. Dynamic epigenetic states of maize centromeres. Front Plant Sci. 2015;6:904.

59. Wang K, Wu Y, Zhang W, Dawe RK, Jiang J. Maize centromeres expand and adopt a uniform size in the genetic background of oat. Genome Res. 2014;24(1):107–16.

60. Han J, Masonbrink RE, Shan W, Song F, Zhang J, Yu W, et al. Rapid proliferation and nucleolar organizer targeting centromeric retrotransposons in cotton. Plant J. 2016;88(6):992–1005.

61. Li Y, Zuo S, Zhang Z, Li Z, Han J, Chu Z, et al. Centromeric DNA characterization in the model grass Brachypodium distachyon provides insights on the evolution of the genus. Plant J. 2018;93(6):1088–101.

62. Wolfgruber TK, Sharma A, Schneider KL, Albert PS, Koo D-H, Shi J, et al. Maize centromere structure and evolution: sequence analysis of centromeres 2 and 5 reveals dynamic loci shaped primarily by retrotransposons. PLoS Genet. 2009;5(11):e1000743.

63. Jiang J, Birchler JA. Plant centromere biology. New York: Wiley; 2013.

64. Bennett MD, Smith JB, Ward J, Jenkins G. The relationship between nuclear DNA content and centromere volume in higher plants. J Cell Sci. 1981;47(1):91–115.

65. Laurie DA, Bennett MD. The production of haploid wheat plants from wheat x maize crosses. Theor Appl Genet. 1988;76(3):393–7.

66. Ishii T, Tanaka H, Eltayeb AE, Tsujimoto H. Wide hybridization between oat and pearl millet belonging to different subfamilies of *Poaceae*. Plant Reprod. 2013;26(1):25–32.

67. Ishii T, Karimi-Ashtiyani R, Houben A. Haploidization via chromosome elimination: means and mechanisms. Annu Rev Plant Biol. 2016;67:421–38.

68. Jin W, Melo JR, Nagaki K, Talbert PB, Henikoff S, Dawe RK, et al. Maize centromeres: organization and functional adaptation in the genetic background of oat. Plant Cell. 2004;16(3):571–81.

69. Talbert PB, Masuelli R, Tyagi AP, Comai L, Henikoff S. Centromeric localization and adaptive evolution of an *Arabidopsis* histone H3 variant. Plant Cell. 2002;14(5):1053–66.

70. Hoffmann S, Dumont M, Barra V, Ly P, Nechemia-Arbely Y, McMahon MA, et al. CENP-A is dispensable for mitotic centromere function after initial centromere/kinetochore assembly. Cell Rep. 2016;17(9):2394–404.

71. Lermontova I, Rutten T, Schubert I. Deposition, turnover, and release of CENH3 at *Arabidopsis* centromeres. Chromosoma. 2011;120(6):633–40.

72. Coe EH. A line of maize with high haploid frequency. Am Nat. 1959;93(873):381–2.

73. Xu X, Li L, Dong X, Jin W, Melchinger AE, Chen S. Gametophytic and zygotic selection leads to segregation distortion through in vivo induction of a maternal haploid in maize. J Exp Bot. 2013;64(4):1083–96.

74. Eder J, Chalyk S. In vivo haploid induction in maize. Theor Appl Genet. 2002;104(4):703–8.

75. Prigge V, Xu X, Li L, Babu R, Chen S, Atlin GN, et al. New insights into the genetics of in vivo induction of maternal haploids, the backbone of doubled haploid technology in maize. Genetics. 2012;190(2):781–93.

76. Dong X, Xu X, Miao J, Li L, Zhang D, Mi X, et al. Fine mapping of *qhir1* influencing in vivo haploid induction in maize. Theor Appl Genet. 2013;126(7):1713–20.

77. Hu H, Schrag TA, Peis R, Unterseer S, Schipprack W, Chen S, et al. The genetic basis of haploid induction in maize identified with a novel genome-wide association method. Genetics. 2016;202(4):1267–76.

78. Lashermes P, Beckert M. Genetic control of maternal haploidy in maize (*Zea mays* L.) and selection of haploid inducing lines. Theor Appl Genet. 1988;76(3):405–10.

79. Barret P, Brinkmann M, Beckert M. A major locus expressed in the male gametophyte with incomplete penetrance is responsible for in situ gynogenesis in maize. Theor Appl Genet. 2008;117(4):581–94.

80. Kelliher T, Starr D, Richbourg L, Chintamanani S, Delzer B, Nuccio ML, et al. MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. Nature. 2017;542(7639):105–9.

81. Liu C, Li X, Meng D, Zhong Y, Chen C, Dong X, et al. A 4-bp insertion at *ZmPLA1* encoding a putative phospholipase A generates haploid induction in maize. Mol Plant. 2017;10(3):520–2.

82. Gilles LM, Khaled A, Laffaire JB, Chaignon S, Gendrot G, Laplaige J, et al. Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. EMBO J. 2017;36(6):707–17.

83. Kim HJ, Ok SH, Bahn SC, Jang J, Oh SA, Park SK, et al. Endoplasmic reticulum- and Golgi-localized phospholipase A2 plays critical roles in *Arabidopsis* pollen development and germination. Plant Cell. 2011;23(1):94–110.

84. Sarkar KR, Coe EH. A genetic analysis of the origin of maternal haploids in maize. Genetics. 1966;54(2):453–64.

85. Zhao X, Xu X, Xie H, Chen S, Jin W. Fertilization and uniparental chromosome elimination during crosses with maize haploid inducers. Plant Physiol. 2013;163(2):721–31.

86. Qiu F, Liang Y, Li Y, Liu Y, Wang L, Zheng Y. Morphological, cellular and molecular evidences of chromosome random elimination in vivo upon haploid induction in maize. Curr Plant Biol. 2014;1:83–90.

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

88. Cheeseman IM. The kinetochore. Cold Spring Harb Perspect Biol. 2014;6(7):a015826.

89. Liu C, Li W, Zhong Y, Dong X, Hu H, Tian X, et al. Fine mapping of *qhir8* affecting in vivo haploid induction in maize. Theor Appl Genet. 2015;128(12):2507–15.

90. Wang X. Plant phospholipases. Annu Rev Plant Physiol Plant Mol Biol. 2001;52(1):211–31.

91. Yao L, Zhang Y, Liu C, Liu Y, Wang Y, Liang D, et al. OsMATL mutation induces haploid seed formation in indica rice. Nat Plants. 2018;4(8):530–3.

92. La Camera S, Geoffroy P, Samaha H, Ndiaye A, Rahim G, Legrand M, et al. A pathogen-inducible patatin-like lipid acyl hydrolase facilitates fungal and bacterial host colonization in *Arabidopsis*. Plant J. 2005;44(5):810–25.