

Differences in homologous and heterologous cytoplasmic-nuclear interactions between cotton cytoplasmic male sterility lines (*Gossypium barbadense* L.)

Jingyi You

Guangxi University

Min Li

Guangxi University

Xiangjun Kong

Henan Institute of Science and Technology

Qiong Zhou

Guangxi University

Qingguang Hou

Hechi Institute of Agricultural Sciences, Guangxi Academy of Agricultural Sciences Hechi Branch

Hongwei Li

Shaoyang University

Bin Li

Guangxi University

Ruiyang Zhou

YOUJY0407@163.com

Guangxi University

Article

Keywords: cotton, CMS line with homologous cytoplasm and its nucleus, CMS line with heterologous cytoplasm and its nucleus, mitochondrial genome resequencing

Posted Date: May 21st, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4386176/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on April 11th, 2025. See the published version at <https://doi.org/10.1038/s41598-025-95027-5>.

Abstract

The utilization of crop hybrids plays an important role in crop breeding and production, and the innovation of the male sterile germplasm is the basis for this utilization. However, the utilization of hybrid advantage in cotton is currently dominated by cytoplasmic male sterility (CMS) lines in Harknessi cotton, which has a single cytoplasmic origin and exhibits a significant negative effect of cytoplasmic-nuclear interactions. The negative effect of cytoplasmic-nuclear interactions can only be minimized by selecting and breeding CMS lines in which the cytoplasm and nucleus originate from the same variety. However, no homologous cytoplasmic-nuclear CMS germplasm has been created, and its mechanism of occurrence has not been determined. In this study, two homologous cytoplasmic-nuclear CMS lines and two heterologous cytoplasmic-nuclear CMS lines were utilized, and the heterologous cytoplasmic-nuclear CMS lines were aborted at a relatively early stage. The physiological indexes related to reactive oxygen species (ROS) metabolism in the heterologous cytoplasmic-nuclear CMS lines were lower than those of the homologous cytoplasmic-nuclear CMS lines, including the enzyme activities of POD and CAT from the tetrad to the mature pollen grain period, and the metabolite content of malondialdehyde (MDA) was inversely correlated with the enzyme activities of the heterologous cytoplasmic-nuclear CMS lines. Resequencing analysis of four cotton mitochondrial genomes (mt genomes) revealed that the heterologous cytoplasmic-nuclear CMS lines were more complex than the homologous cytoplasmic-nuclear CMS lines, and the homologous CMS lines showed a higher degree of covariance with the maintainer lines. This indicates that heterologous cytoplasmic-nuclear interactions are more likely to lead to mtDNA structural variation. Taken together, the results showed that the cytoplasmic-nucleus homologous system was less affected by the cytoplasmic-nuclear interaction and was the best combination for the study of male sterility.

Introduction

The main characteristic of plant cytoplasmic male sterility (CMS) is the inability to produce pollen or the production of nonfunctional pollen, which is manifested at the cytological level by abnormal microspore development. Cotton CMS line failures occur at different times from the spore-making cell period to the tetrad period. Reactive oxygen species (ROS) are important signaling molecules involved in many biochemical processes during normal plant growth and development. ROS play a very important role in plant growth and development, stress and response processes, seed germination, programmed cell death (PCD) and other physiological processes¹. The sources of ROS in plants include chloroplast bodies, mitochondria, glyoxylate cyclers, cytoplasm, nucleus, peroxisomes, and membrane systems². Chloroplasts are the main source of ROS during photosynthesis in plants³⁻⁵, whereas mitochondria are the main source of ROS in processes other than photosynthesis⁶. The toxic effects of ROS are mainly manifested in oxidative damage to macromolecules and even cell death. Excessive ROS accumulation can seriously affect the stability of mitochondrial DNA (mtDNA). Shao et al⁷ studied an *Arabidopsis msh1* mutant and found that when the expression of this gene was changed, the level of ROS was reversed, the stability of mtDNA was decreased, and the recombination rate was increased. This is because mtDNA is directly exposed to ROS generated by the oxidative phosphorylation process, which is an important target molecule for free radical attack. High levels of ROS can lead to mtDNA breakage, destroying its structure and function⁸. Additionally, ROS can induce point mutations, deletions and insertions in mtDNA^{9,10}. Changes in ROS levels can lead to male sterility by preventing PCD from occurring over time in the downy mildew layer^{11,12}. Previous studies have shown that the accumulation of excess ROS leads to an oxidative burst that ultimately aborts pollen production^{13,14}. In plants, maintaining a dynamic balance of ROS is important for maintaining plant fertility.

Research on CMS at the molecular level began in 1976, when Levings and Pring's comparison of mtDNA zymography profiles of maize T-type cytoplasm and normal cytoplasm revealed significant differences¹⁵. The study of plant organelles by high-throughput sequencing dates back to the first acquisition of the genome of tobacco chloroplasts in 1986 and the mitochondrial genome of ground money in 1992¹⁶. Studies have hypothesized that the main cause of CMS in plants may be mitochondrial structural or sequence alterations¹⁷. Therefore, sequence studies and analysis of plant mitochondrial genomes are necessary to understand the nucleoplasmic interactions and the mechanism of CMS. *Arabidopsis thaliana* was the first higher plant for which mtDNA was sequenced¹⁸. To date, the mitochondrial genomes of more than five hundred plants have been sequenced¹⁹, including maize²⁰, rice²¹, tobacco²², rapeseed²³, sugar beet²⁴, kenaf²⁵, carrot²⁶, and others. With the accumulation of increasing mitochondrial sequencing data, studies in recent years have produced a growing body of evidence suggesting that CMS may be related to the rearrangement of the mitochondrial genome. Mitochondrial gene rearrangements often result in the formation of chimeras with neighboring sequences, which are cotranscribed with unmutated sequences in the form of chimeric genes, leading to CMS.

The rearrangement pattern of cotton mtDNA during the evolutionary process has been a topic of interest for scientists. Previous studies have shown that the inclusion of many repetitive sequences in mitochondria is closely related to CMS. It is hypothesized that this is due to the formation of open reading frames (ORFs) by repeated sequence-mediated gene rearrangements. ORFs generated by mitochondrial genome reorganization are usually located upstream or downstream of mitochondrial ATPase complex subunit genes, such as *atp6*, *atp8*, and *atp9*, and coexpressed with these genes. On the one hand, coexpression products interfere with normal mitochondrial function, such as toxic proteins competing with mitochondrial genes for substrates. On the other hand, cotranscription with neighboring mitochondrial protein-coding genes can affect RNA editing, leading to posttranslational protein dysfunction and microspore abortion²⁷. There are also numerous cytochrome oxidase genes in the electron transport chain that undergo rearrangements to form sterility-associated genes, such as the rearrangement of the genes *atp6* and *orf25*²⁸, the *orf256* gene upstream of *cox1* in wheat²⁹, and the *atpA* and *atp6* loci in sugar beets³⁰. Alterations in the DNA sequences of these genes may affect the activity of cytochrome oxidase, which in turn affects ATP synthesis and leads to the development of septoria.

In this study, we compared and analyzed the effects of cytoplasmic-nuclear interactions between homologous and heterologous cytoplasmic-nuclear CMS lines on microspore abortion, reactive oxygen metabolism and mtDNA rearrangement using maintainer lines as controls. We also systematically investigated the different types of cytoplasmic effects and the relationships of cytoplasmic-nuclear interactions from multiple perspectives, which provide not only theoretical guidance for the creation of optimal sterile lines and cytoplasmic-nuclear interactions but also theoretical and technological support for the creation of CMS lines for the cotton industry as a whole.

Results

Morphological and cytological observations

Flowers are the organs of sexual reproduction in plants and play an important role in the reproduction of offspring; therefore, the study of plant fertility cannot be separated from the study of flowers. Comparison of the floral organ phenotypes of homologous and heterologous cytoplasmic-nuclear CMS lines and their maintainer lines revealed that although there were no significant differences in the morphology of the floral organs and leaves between the

CMS lines, the styles of the heterologous CMS lines were relatively protruding, and the filaments and anthers were more shortened and shriveled (Fig. 1).

Comparative analysis of paraffin sections from homologous and heterologous cytoplasmic-nuclear CMS lines

Four critical periods of anther development were observed in the pollen sac cell structures of one homologous heterokaryotic cotton CMS line, two heterologous homokaryotic cotton CMS lines, and maintainer lines (Figs. 2, 3); the four periods were the pollen mother cell (PMC) stage, tetrad (Td) stage, early unicellular (early Uni) stage, and mature pollen grain (MP) stage.

Comparative analysis of paraffin sections of homologous heterokaryotic sterile lines

Comparative analysis of anther cell structures during cotton development in homologous heterokaryotic CMS lines revealed that during PMC meiosis, the PMC could be clearly observed in the center of the pollen sac lumen, with a relatively large size and a distinct nucleus, and the differences in PMC stages were not significant (Fig. 2, a1, b1, c1).

In the Td stage, the homologous cytoplasmic-nuclear CMS line 276A had tetrad structures in the anthers, while the heterologous cytoplasmic-nuclear CMS line C2-113A had no tetrad structures; both microspores were aborted, with 276A starting to abort in the Td stage and C2-113A almost completely aborted in the Ts stage (Fig. 2, a2, b2, c2).

The cells of the chorioallantoic layer of the monokaryotic maintainer line 276B were highly vesicularized and began to degrade. The callus encasing the tetrads was degraded and released microspores, while a large central vesicle began to appear that gradually pushed the nucleus toward the cell wall. The microspores of the sterile lines were completely degraded. In the two sterile lines, the chorionic layer was not degraded, and the degraded PMC structure was further vesicularized (Fig. 2, a3, b3, c3).

In the MP stage, the pollen grains of the maintainer line 276B were mature, with conspicuous punctures and complete degradation of the felted cells, whereas in both CMS lines, the pollen sacs were completely crumpled, forming a darker solid structure (Fig. 2, a4, b4, c4).

These results indicated that abortion of the heterologous cytoplasmic-nuclear CMS lines occurred earlier than that of the homologous CMS lines.

Comparative analysis of paraffin sections of homonuclear and heterologous sterile lines

Through the comparative analysis of anther cell structure during the development of cotton CMS lines, it was found that in the PMC stage of both sterile and maintainer lines, the PMC was located in the center of the pollen sac, the chorion cells were larger, and the differences in the PMC stage of the different materials were not significant (Fig. 3, a1, b1, and c1).

In the Td stage, 07-113A/B anthers had tetrad structures, while C2-113A and C4-113A anthers had no tetrad structures; the chorionic layer was not degraded; and the microspores were aborted, which indicated that the

homologous cytoplasmic-nuclear CMS lines began to abort at the Td stage, and the heterologous cytoplasmic-nuclear CMS line C2-113A was almost completely aborted in the Td stage (Fig. 3, a2, b2, c2).

The pollen sac cavity of the monokaryotic line 07-113B increased, the tetrad microspores encapsulated in the anthers separated into monokaryotic nuclei, and the chorionic layer was completely degraded; the sterile line still exhibited a tetrad structure, and the microspores had been completely degraded (Fig. 3, a3, b3, c3).

In the MP stage, the cells of the felid layer of the maintainer line 07-113B were almost completely degraded, with some cell fragments remaining, and the pollen grains were mature, with obvious punctures and pollen-dispersing anthers, whereas the pollen sacs of the two CMS lines were completely crumpled, forming a darker solid structure (Fig. 3, a4, b4, c4).

Comparative analysis of the anther cell structure during the development of homonuclear and heterologous CMS lines showed that the homologous cytoplasmic-nuclear CMS line 07-113A was abortive from the Td stage to the early monokaryotic stage, whereas the heterologous cytoplasmic-nuclear CMS lines C2-113A and C4-113A were almost completely abortive at the Td stage. These results indicated that the abortion period of the heterologous cytoplasmic-nuclear CMS lines preceded the Td stage and was earlier than that of the homologous cytoplasmic-nuclear CMS lines.

Activity of enzymes involved in reactive oxygen scavenging pathways

For the homologous cytoplasmic-nuclear CMS line 07-113A, POD activity was higher in the PMC stage and lower in the Td stage and the MP stage when compared with the sterile line 07-113B. For the heterologous cytoplasmic-nuclear CMS lines C2-113A and C4-113A, POD activity was also higher in the PMC stage and lower in the Td stage and the MP stage compared with that of the maintainer line 07-113B (Fig. 4a,b). In contrast, the POD activity of the heterologous cytoplasmic-nuclear CMS line C4-113A was consistently lower than that of the homologous cytoplasmic-nuclear CMS line 07-113A at the Td stage and the MP stage (Fig. 4b), while the POD activity of C2-113A showed a decreasing trend and was lowest at the MP stage (Fig. 4a). Similarly, comparative analysis of the homologous cytoplasmic-nuclear line 276A and its maintainer line 276B revealed that the POD activities of 276A relative to 276B showed a low-high-low trend, and the POD activities of C2-113A were consistently lower than those of the maintainer line 276B during the period of pollen development (Fig. 4c), and the heterologous cytoplasmic-nuclear CMS lines (C2-113A) had consistently lower POD activity than the corresponding homologous cytoplasmic-nuclear CMS line (276A).

The CAT activity of the homologous cytoplasmic-nuclear CMS line 07-113A was higher in the PMC stage and lower in the Td stage and the MP stage than that of the maintainer line 07-113B (Fig. 4d,e). For the heterologous cytoplasmic-nuclear CMS line C2-113A, POD activity was also higher in the PMC stage and lower in the Td stage and the MP stage compared to the maintainer line 07-113B (Fig. 4d). The activity of C4-113A was lower than that of the maintainer line 07-113B in all three periods (Fig. 4e). The CAT activity in anthers of the sterile line 276A was lower than that of the maintainer line 276B at the PMC stage; the CAT activity increased at the Td stage, and the activity of 276A was significantly higher than that of 276B at the MP stage, which may be due to the high level of ROS scavenger enzyme in 276A, resulting in a low level of ROS and affecting pollen development (Fig. 4f). The CAT activity of C2-113A was consistently lower than that of 276B during pollen development, reflecting the weaker ability of anthers in C2-113A to eliminate peroxides than that of the fertile variety 276B as microspore development occurs; the CAT activity of the heterologous cytoplasmic-nuclear CMS lines was consistently lower than that of the

homologous cytoplasmic-nuclear CMS lines (Fig. 4f). The lower enzyme activity of the heterologous CMS lines predicted a lower ability to eliminate ROS.

The MDA content reflects the degree of lipid peroxidation in organisms and may indirectly reflect the degree of cellular damage. Compared with CMS line 07-113B, the MDA content of CMS line 07-113A was higher during the PMC stage, but the MDA content of CMS line 07-113B was higher during the Td stage and lower during the maturation stage. At the onset of microspore abortion, the heterologous cytoplasmic-nuclear CMS lines (C2-113A and C4-113A) showed a low-low-high trend of MDA content compared with the maintainer line 07-113B, in which the MDA content always increased (Fig. 4g,h). The MDA content of the heterologous cytoplasmic-nuclear CMS lines was higher than that of the homologous cytoplasmic-nuclear CMS lines at the time of pollen grain maturation (Fig. 4g,h, i). The MDA content of 276A was higher than that of its maintainer line 276B in all three periods and showed a positive trend. C2-113A also followed this pattern compared with 276B, and the MDA content was higher in heterologous CMS lines than in homologous CMS lines (Fig. 4i).

Overview of cotton mtDNA sequencing data

The libraries were constructed against mtDNA and used a combination of second- and third-generation sequencing approaches, data filtering and other processing to obtain the data information statistics of each sample (Table 1). The valid data and quality distribution obtained indicate that the sequencing results are reliable.

Table 1
Sequencing data statistics of mtDNA

Sample ID	Raw data(Mb)	Clean data(Mb)	Clean data Q20(%)	Clean data Q30(%)	Clean data GC(%)
276A	5795	4337	98.12	94.1	40.11
276B	5720	5459	97.87	93.62	36.21
C2-113A	4985	4232	97.96	93.64	37.64
C4-113A	4960	3643	98.23	94.37	38.53
07-113B	5610	5322	98.46	94.88	40.88
07-113A	6864	6533	98.5	95.02	38.36

The mtDNA was assembled with the assembly software. The mt genomes of C2-113A (GenBank accession: OR906298), C4-113A (GenBank accession: OR906297), 276A (GenBank accession: OR906300) and 276B (GenBank accession: OR906299), 07-113A (GenBank accession are not yet available), 07-113B (GenBank accession are not yet available) were each assembled into a single circular molecule, and the mt genome sizes of the CMS lines were smaller than the genome size of their maintainer line (Table 2). The mitochondrial genome circle diagrams of each cotton material were plotted with the OrganellarGenomeDRAW software (Fig. 5, Fig. S1).

Table 2
Genetic composition of mitochondria of cotton material

Samples	07-113B	07-113A	C4-113A	276B	276A	C2-113A
Genome size(bp)	677,504	634,041	634,043	792,184	634,042	634,043
GC Content(%):	42.63	42.66	44.92	45.02	44.92	44.92
Gene Number:	36	36	36	36	36	36
Gene Total Length:	37,401	33,357	31,959	41,196	31,959	31,959
Intergenic region Length:	600,684	640,103	602,084	750,988	602,083	602,084

The predicted number of genes coding for proteins in the sterile and maintainer lines was 36. By analyzing the noncoding RNAs (ncRNAs) of the six materials, it was found that the number of transfer RNAs (tRNAs) was higher in all the maintainer lines than in the sterile lines (Table 3), and there were differences in the number of ribosomal RNAs (rRNAs) and tRNAs between the homologous CMS lines and heterologous CMS lines without any pattern.

Table 3
Composition of ncRNAs in plastid-nucleus homologous and plastid-nucleus heterologous CMS lines

Sample	Type	ncRNA number	Total length(bp)	Average length(bp)	ncRNA length/Genome(%)
07-113B	tRNA	33	2,543	77	0.3753
	rrn18	2	3,914	1,957	0.5777
	rrn26	2	6,750	3,375	0.9963
	rrn5	2	238	119	0.0351
07-113A	tRNA	27	2,073	76	0.327
	rrn18	2	3,914	1,957	0.6173
	rrn26	2	6,748	3,374	1.0643
	rrn5	2	238	119	0.0375
C4-13A	tRNA	28	2163	77	0.34
	rrn18	2	3914	1957	0.62
	rrn26	2	6748	3374	1.06
	rrn5	2	238	119	0.04
276B	tRNA	42	3247	77	0.41
	rrn18	3	5871	1957	0.74
	rrn26	3	10125	3375	1.28
	rrn5	3	357	119	0.05
276A	tRNA	28	2163	77	0.34
	rrn18	2	3914	1957	0.62
	rrn26	2	6748	3374	1.06
	rrn5	2	238	119	0.04
C2-113A	tRNA	28	2163	77	0.34
	rrn18	2	3914	1957	0.62
	rrn26	2	6748	3374	1.06
	rrn5	2	238	119	0.04

SV detection and analysis

SVs are large sequence and positional changes in the genome, such as duplications, insertions, inversions, heterozygotes, and deletions, which are the result of a combination of endogenous and exogenous factors. SVs have a greater impact on the genome than single-nucleotide polymorphisms (SNPs), and they can often cause birth defects, cancer, etc., in the organisms. Comparison of the SVs in the genomes of the four sterile lines showed that the genomes of the heterologous cytoplasmic-nuclear CMS lines are more complex than those of the homologous CMS lines (Table 4).

Table 4
SV types and mutations in homogeneous heteronuclear and homonuclear
heterogeneous CMS lines

Sample ID	276A	C2-113A	07-113A	C2-113A	C4-113A
Reference ID	276B	276B	07-113B	07-113B	07-113B
Translocation	7	16	8	17	8
Inversion	4	0	2	0	4
Trans + Inver	8	1	9	1	9
Deletion	3	2	1	2	2
Insertion	2	2	2	2	2
ComplexIndel	4	5	3	5	4

Analysis of mitochondrial genome covariance

By analyzing the covariance between two genomes, we can identify insertions and deletions between the genome of the target species and the reference genome and analyze the structural changes (e.g., chromosome rearrangement) of the genomes in the process of evolution. Using the maintainer line 07-113B as the reference, 07-113A, C2-113A and C4-113A were compared, and Fig. 6 shows that the covariance between 07-113A and 07-113B was better, the covariance between C2-113A and 07-113B was poorer, and more translocations occurred in C2-113A (Fig. 6a).

As shown in Fig. 6b, the two sterile lines were compared with 276B for collinearity, and the collinearity of 276A with 276B was better than that of C2-113A with 276B.

Discussion

Periods of microspore abortion in cotton CMS lines and their cytological characteristics

Due to the different sterility mechanisms of cotton CMS lines, there are great differences in the time of abortion and abortion characteristics among different sterile lines. In the land cotton CMS line J4A, the abortion period occurs from the microsporoblast stage to the Td stage (during meiosis), and abortion is characterized by the absence of degradation of the chorionic and mesophyll cells of the anthers throughout microspore development and the absence of tetrad structure formation³¹. Failure of Jin A cotton CMS lines begins during the proliferation of sporulating cells, which are unable to undergo normal mitosis and often contain multiple micronuclei, followed by cytoplasmic vesiculation of the microsporangial mother cell³². Comparative observation of cotton A and B resistance found that septicity began to occur at the same time as microspore mother cells, and microspore mother cells obviously degraded and gradually disintegrated³³. In CMS-D8, septicity started at the microsporoblast stage, and the microspores were completely degraded at the Td stage with no tetrad formation. During the septic stage, the pomatum cells are abnormally vesicularized and extruded into the drug compartment³⁴. The island cotton CMS line H276A was aborted at the Td stage, characterized by vesicularization of the nuclei of the tetrad cells, and the pomace layer was intact and did not undergo degradation during anther development^{19,32}. The

abortion period of the CMS "Anti-A" sterile lines in cotton occurs at the Td stage, and no tetrad structure is formed in the sterile lines ³⁵.

Hu et al. studied the mechanism of abortion in citrus and found that the pollen abortion of male sterile Ougan mutants occurred at the microspore Td stage ³⁶. Analysis of paraffin sections revealed that the CMS lines underwent abortion at different times, and none of the sterile lines showed degradation of the chorionic layer during microspore development. Comparing the homologous and heterologous cytoplasmic-nuclear CMS lines, we found that the abortion period of homologous CMS lines was generally later than that of heterologous CMS lines. It was hypothesized that because the mitochondrial and nuclear genomes of homologous cytoplasmic-nuclear CMS lines and their maintainer lines were almost identical, this kind of sterile line was less affected by the nucleoplasmic interaction effect, which led to late abortion and made them the best combinations for the study of male sterility.

Production and clearance of ROS in anther development

The activities of several ROS-eliminating enzymes shared the same characteristics: the highest activity was found in the maintainer lines, and the lowest activity was found in the sterile lines. From the comparative analysis of the indicators related to the metabolism of ROS in sterile and maintainer lines, combined with cytological observation in the early stage, we can speculate that starting from the PMC stage, due to the significant decrease in the activities of POD and CAT in the anthers of sterile lines, the anther's ability to eliminate ROS was gradually reduced, resulting in the peroxidation of the plasma membrane of the mitochondria, which is reflected in the extremely significant increase in the MDA content (Fig. 4i ,g ,h), disruption of the structure of the mitochondrial membrane and contraction of the plasma membrane space. The destruction of the mitochondrial membrane structure and the contraction of the plasma membrane space led to the impairment of mitochondrial function and the inability to convert energy normally, resulting in the reduction of ATP and NADH and the blockage of ascorbic acid and glutathione cycling. The inability of mitochondria to convert the energy produced during glycolysis into ATP leads to the accumulation of ROS, a byproduct of glycolysis, which creates a vicious cycle that is manifested at the cellular level by the fact that chorion cells do not undergo degradation at all times (Fig. 2, 3), resulting in the inability of microspores to obtain sufficient energy.

The analysis of ROS in homologous and heterologous cytoplasmic-nuclear CMS lines revealed that the enzyme activity of heterologous CMS lines was consistently lower than that of homologous CMS lines during the time of sporulation to mature pollen grains. This suggests that superoxide elimination from the anthers of sterile lines is weaker than that of maintainer lines and that superoxide elimination from the anthers of heterologous cytoplasmic-nuclear sterile lines is weaker than that of homologous cytoplasmic-nuclear CMS lines. The level of MDA, the end product of ROS metabolism, was opposite to the enzyme activity in the three materials and was generally lower in the maintainer lines than in the sterile lines, while the final level was higher in the heterologous CMS lines than in the homologous CMS lines. Excessive production of ROS in sterile lines, as well as significantly lower activities of various enzymes such as SOD, POD and CAT, compared with maintainer lines and insufficient elimination capacity increased the accumulation of ROS, leading to a very high accumulation of the metabolite MDA. Over time, the ROS in the sterile lines reached a level that caused toxicity to the cells and damage to the organelles, leading to abnormal cell function and ultimately to microspore failure. As plasmodial heterozygotes, CMS lines have lower anti-ROS enzyme activity and higher ROS accumulation levels, which results in higher toxicity and more pronounced negative effects. These results were confirmed by phenotypic and cytological observations.

Role of the mitochondrial genome in male sterility in cotton

Although previous studies on cotton CMS have used materials with different cytoplasmic and nuclear genetic backgrounds, few studies on materials from homologous cytoplasmic-nuclear CMS lines and heterologous cytoplasmic-nuclear CMS lines have been published. The mitochondrion is a central site for a wide range of metabolic activities and is involved in a variety of biochemical processes³⁷. Mitochondria are involved in ATP synthesis, which provides energy for normal physiological activities in plants and is also a major source of ROS production³⁸. The mitochondrial genome evolves through genomic rearrangements as well as insertions, deletions and base mutations in the sequence³⁹. In this study, we found that several cotton sterile line SNPs and insertions-deletions (InDels) were distributed in the intergenic region; all InDels occurred in the intergenic region, no insertion or deletion of small fragments occurred in the CDS region, and the InDels were mostly small fragments (1–5 bp) (Fig. S2,S3). All SNP types in the mtDNA of the sterile lines were inverted, and the vast majority of them were A-C transitions. We found that SNP mutations in the mitochondrial genome of the four CMS lines occurred relatively more often in genes for ribosome-encoded proteins, such as *rpl1*, *rpl2*, *rpl4*, *rpl5*, *rpl10*, *rpl16*, and *rps4* (Tables S1,S2). From the results of this study, it can be hypothesized that the content of H₂O₂ in the CMS strain was significantly higher than that in the maintainer strain, whereas CAT, POD and other related enzymes in the CMS strain were regulated by the related genes, which resulted in under- or overactivity and an imbalance of ROS. The heterologous cytoplasmic-nuclear CMS lines had lower anti-ROS enzyme activities and higher levels of ROS accumulation, which may be an important reason for the higher number of mutation sites and more SVs in the mitochondrial genome and poorer covariance between the maintainer lines and the heterologous CMS lines.

Conclusions

Currently, there is a single narrow source of biparental germplasm resources for hybridization, and most of the CMS materials used in previous studies are heterologous cytoplasmic-nuclear materials obtained by heterologous karyoplasmic substitution. Negative cytoplasmic effects caused by cytoplasmic-nuclear heterozygosity may interfere with progeny selection for CMS, and the overabundance of genetic information that is not related to CMS may impede the study of the CMS mechanism in cotton. Therefore, in this study, we selected homologous cytoplasmic-nuclear and heterologous cytoplasmic-nuclear CMS lines and their respective maintainer lines as experimental materials and performed a comparative analysis of the morphology, cytology, physiology and biochemistry indices and the mitochondrial genome level of the homologous and heterologous cytoplasmic-nuclear CMS lines. This study demonstrated that the homologous system is less affected by nuclear-cytoplasmic interactions, provided a theoretical basis for the creation of CMS lines optimized for cytoplasmic-nuclear interactions and laid the foundation for the analysis of cytoplasmic-nuclear homology and heterozygosity in the CMS mechanism.

Methods

Plant materials

The materials used in this study included a homologous cytoplasmic-nuclear sterile line, 07-113A; two heterologous cytoplasmic-nuclear sterile lines, C2-113A and C4-113A; and their maintainer line, 07-113B. The CMS germplasm 07-113S (referred to as the C3 cytoplasm) of *Gossypium barbadense* L. was obtained by continuous self-crossing of the 07-113 line, introduced by the Institute of cotton research of CAAS in 2007, and its cytoplasmic homologous CMS line is 07-113A. The CMS germplasm 276S (referred to as the C2 cytoplasm) of an island cotton strain, introduced from the Cotton Institute of the Chinese Academy of Agricultural Sciences, was obtained in July

2012 by transplanting the non-full-length Hcpdil5-2a gene into kenaf by the pollen tube channel method, which was saturated and backcrossed with 276B for many consecutive years to obtain a stable sterile line, which was named cytoplasmic homologous CMS line 276A. C2-113A and C4-113A are stable, sterile, cytoplasmic heterologous CMS lines obtained by multigeneration saturation backcrossing with 07-113B as the paternal plant and “C2” and “C4” cytoplasmic materials as the maternal plant. Anthers of different developmental stages and buds of different sizes were taken from cottons at full bloom. The materials were collected from the experimental field of Guangxi University, and all materials were collected under the same normal growing conditions.

Morphological and cytological observations

Several CMS lines as well as maintainer lines of cotton flowers were collected under the same conditions during the full bloom period, and the morphology of the floral organs was observed and photographed (Canon, Tokyo, Japan). Flower buds of different diameters were also collected and fixed in Canon's solution to make paraffin sections, which were measured with a Vernier caliper and classified into four grades: $2.0 \leq \text{BL} < 3$, $3.0 \leq \text{BL} < 4$, $4.0 \leq \text{BL} < 5$, and $5.0 \leq \text{BL} < 6$. The sepals and petal tissues were removed, placed in Carnot fixative, and left to be fixed in a refrigerator at 4°C for 24 h. The buds were dehydrated with a graded ethanol series ⁴⁰. Dehydrated flower buds were embedded in paraffin and sliced into 10 µm thick sections. Serial sections of flower bud tissue were placed on slides and stained with Senna-solid green. The sectioned flower buds were observed and imaged through a DMI3000B microscope (Leica, Wetzlar, Germany).

Enzyme activity and MDA content assays

The enzyme activity of CAT was determined by double-antibody sandwich assay using a plant superoxide dismutase enzyme-linked immunosorbent assay (ELISA) kit (Shanghai Enzyme-Link Bio), and the activity of POD was determined by guaiacol colorimetric assay according to the experimental method of Li Hesheng ⁴¹. The MDA content in cotton anthers was determined by a double-antibody sandwich assay using the plant hydrogen peroxide ELISA kit (Shanghai Enzyme-Link Bio).

Mitochondrial sequencing and assembly

Young plant leaves of four CMS lines and two maintainer lines of cotton were collected with no less than 10 g of each material, and the materials were subsequently treated under dark conditions for 1–2 d to cultivate yellowing seedlings. Combining density gradient and differential centrifugation, mtDNA was extracted using the modified mtDNA extraction method of Xiaofang Liao ⁴², and the extracted mtDNA was sent to Shanghai Lingen Biotechnology Co. After passing the quality control test, the library was constructed with a starting amount of 1 µg DNA, the DNA was fragmented into 300–500 bp fragments by ultrasonication with Covaris M220, the library was enriched and amplified by polymerase chain reaction (PCR), and the target bands were recovered with 2% agarose gel (Certified Low Range Ultra Agarose). The cotton mitochondrial genome was sequenced using a high-throughput sequencing platform (Illumina NovaSeq 6000). Then, Illumina sequencing data were assembled using GetOrganelle v1.7.1, and second-generation assembled sequences were aligned to Nanopore third-generation data using BWA v0.7.17 to extract the third-generation data of the target samples. Then, the extracted third-generation data were assembled by combining them with the second-generation data. Clean reads were compared to the mitochondrial genome sequence, and bases were corrected using Pilon v1.23. Finally, the start position and orientation of the

mitochondrial scaffold were determined from the reference genome to obtain the final mitochondrial genome sequence.

Analysis and annotations of the mt genomes

We used the homology matching prediction method to perform gene prediction on the sample genome. The protein sequences of the National Center for Biotechnology Information (NCBI) mitochondrial reference genome were first aligned to the sample genome using the software BLAST + 2.7.1 with the threshold parameter e-value 1e-5. Poor alignment results and redundancy were removed, and the integrity of the sample genes and the exon/intron boundaries were manually corrected to obtain the conserved gene set with a high degree of accuracy. The assembled genome sequences of the sequenced samples were combined with the predicted coding genes to obtain a circle map of the sample genome. The software used was OGDRAW (<https://chlorobox.mpimp-golm.mpg.de/OGDraw.html>).

Structural variant detection

MUMmer software was used to compare the target and reference genomes, and then LASTZ was used for region comparison to identify structural variants (SVs).

Covariance analysis

Large-scale covariance between genomes was determined using MUMmer software. Interregion comparison was performed using LASTZ to confirm the local positional alignment relationships and to find the regions of translocation (Translocation/Trans), inversion (Inversion/Inv) and translocation + inversion (Trans + Inv)⁴³.

Declarations

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the National Natural Science Foundation of China (32060465) and the Guangxi Natural Science Foundation of China (2022GXNSFBA035451). The authors are thankful for the financial support from Mr. Weng Hongwu and Weng Hongwu Original Research Fund of Peking University of China, China

Author Contribution

"YJY performed the experiments and drafted the manuscript. LM assisted during the experiment and revised the manuscript. KXJ, ZQ, HQG, LHW and LB participated in the experiments. ZRY conceived, designed, and supervised

the study. All authors read and approved the final manuscript".

Data Availability

"The mt genomes of C2-113A (GenBank accession: OR906298), C4-113A (GenBank accession: OR906297), 276A (GenBank accession: OR906300) and 276B (GenBank accession: OR906299) were assembled with the assembly software and have been submitted to NCBI (<https://www.ncbi.nlm.nih.gov/>)."

References

1. Zhou, X., Liu, Z., Ji, R. & Feng, H. Comparative transcript profiling of fertile and sterile flower buds from multiple-allele-inherited male sterility in Chinese cabbage (*Brassica campestris* L. ssp. *pekinensis*). *MGG Molecular & General Genetics* 292, 1–24 (2017).
2. Gechev, T.S., Breusegem, F.V., Stone, J.M., Denev, I. & Laloi, C. Reactive oxygen species as signals that modulate plant stress responses and programmed cell death. *BioEssays* 28(2006).
3. Asada & Kozi. Production and Scavenging of Reactive Oxygen Species in Chloroplasts and Their Functions. *Plant Physiology* (2006).
4. Dietz, K.J., Turkan, I. & Krieger-Liszkay, A. Redox- and Reactive Oxygen Species-Dependent Signaling into and out of the Photosynthesizing Chloroplast. *Plant Physiology*, 1541–1550 (2016).
5. Takagi, D., Takumi, S., Hashiguchi, M., Sejima, T. & Miyake, C. Superoxide and Singlet Oxygen Produced within the Thylakoid Membranes Both Cause Photosystem I Photoinhibition. *Plant physiology* 171, pp.00246.2016 (2016).
6. Bianka & Steffens. The role of ethylene and ROS in salinity, heavy metal, and flooding responses in rice. *Frontiers in Plant Science* 5, 685 (2014).
7. Shao, M.R., Raju, S.K.K., Laurie, J.D., Sanchez, R. & Mackenzie, S.A. Stress-responsive pathways and small RNA changes distinguish variable developmental phenotypes caused by MSH1 loss. *Bmc Plant Biology* 17(2017).
8. Bae, P.C. & Nils-G U F Ran, L. Mitochondrial DNA mutations in disease and aging. *The Journal of Experimental Medicine* (2011).
9. Kumar, M. et al. Nucleotide variations in mitochondrial DNA and supra-physiological ROS levels in cytogenetically normal cases of premature ovarian insufficiency. *Archives of Gynecology* 282, 695–705 (2010).
10. Venkatesh, S., Deecaraman, M., Kumar, R., Shamsi, M.B. & Dada, R. Role of reactive oxygen species in the pathogenesis of mitochondrial DNA (mtDNA) mutations in male infertility. *The Indian Journal of Medical Research* 129, 127–137 (2009).
11. Shi, J., Cui, M., Yang, L., Kim, Y.J. & Zhang, D. Genetic and Biochemical Mechanisms of Pollen Wall Development. *Trends in Plant Science*, 741–753 (2015).
12. Sha et al. Spatiotemporal Production of Reactive Oxygen Species by NADPH Oxidase Is Critical for Tapetal Programmed Cell Death and Pollen Development in Arabidopsis. *The Plant Cell* (2014).
13. Berghe, T.V., Linkermann, A., Jouan-Lanhuet, S., Walczak, H. & Vandenabeele, P. Regulated necrosis: the expanding network of non-apoptotic cell death pathways. *Nature Reviews Molecular Cell Biology* 15, 135–147 (2014).
14. Geng, X. et al. Identification of Proteins Involved in Carbohydrate Metabolism and Energy Metabolism Pathways and Their Regulation of Cytoplasmic Male Sterility in Wheat. *Int J Mol Sci* 19, 324 (2018).

15. Levings, C.S.I. & Pring, D.R. Restriction Endonuclease Analysis of Mitochondrial DNA from Normal and Texas Cytoplasmic Male-Sterile Maize. *Science* 193, 158–160 (1976).
16. Oda, K. et al. Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. *J Mol Biol* 223, 1–7 (1992).
17. Bentolila, M.R.H.A. Supplement: Plant Reproduction || Interactions of Mitochondrial and Nuclear Genes That Affect Male Gametophyte Development. *Plant Cell* 16, S154-S169 (2004).
18. Unseld, M., Marienfeld, J.R., Brandt, P. & Brennicke, A. The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nature Genetics* 15, 57–61 (1997).
19. Xiangjun, K. et al. Comparative Analysis of the Cytology and Transcriptomes of the Cytoplasmic Male Sterility Line H276A and Its Maintainer Line H276B of Cotton (*Gossypium barbadense* L.). *International Journal of Molecular Sciences* 18, 2240 (2017).
20. Clifton, S.W., Minx, P., Fauron, M.R., Gibson, M. & Newton, K.J. Sequence and Comparative Analysis of the Maize NB Mitochondrial Genome. *Plant Physiology* 136, 3486–3503 (2004).
21. Notsu, Y. et al. The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Molecular Genetics & Genomics* 268, 434–445 (2002).
22. Sugiyama, Y. et al. The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. *Molecular Genetics & Genomics* 272, 603–15 (2005).
23. Hirokazu & Handa. The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic acids research* (2003).
24. Tomohiko, K. et al. The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNA^{Cys}(GCA). *Nucleic acids research* 28, 2571–6 (2000).
25. Xiaofang et al. Complete sequence of kenaf (*Hibiscus cannabinus*) mitochondrial genome and comparative analysis with the mitochondrial genomes of other plants. *Scientific Reports* (2018).
26. Nakajima, Y., Yamamoto, T. & Oeda, M.K. A novel orfB-related gene of carrot mitochondrial genomes that is associated with homeotic cytoplasmic male sterility (CMS). *Plant Molecular Biology* (2001).
27. Xu, P. Huazhong Agricultural University (2010).
28. Mohr S., S.E.O.W. Mitochondrial DNA of cytoplasmic male-sterile *Triticum timopheevi*: rearrangement of upstream sequences of the atp6 and rnf25 genes. *Tag Theoretical & Applied Genetics* 86, 259–268 (1993).
29. Song, J. & Hedgcoth, C. Influence of nuclear background on transcription of a chimeric gene (orf256) and *coxI* in fertile and cytoplasmic male sterile wheats. *Genome* 37, 203–209 (1994).
30. Senda, M., Harada, T., Mikami, T., Sugiura, M. & Kinoshita, T. Genomic organization and sequence analysis of the cytochrome oxidase subunit II gene from normal and male-sterile mitochondria in sugar beet. *Current Genetics* 19, 175–181 (1991).
31. Li, M., Chen, L., Khan, A., Kong, X. & Komatsu, S. Transcriptome and MiRNAomics Analyses Identify Genes Associated with Cytoplasmic Male Sterility in Cotton (*Gossypium hirsutum* L.). *International journal of molecular sciences* 22, 4684 (2021).
32. Huang, J. & Yang, P. Ultra-structural Observations on the Microsporogenesis in Jin A CMS line of cotton. *Journal of Shanxi Agricultural University(Natural Science Edition)* 28, 4 (2008).

33. Wang, X. Zhejiang University (2006).
34. Yang, L. Huazhong Agricultural University (2017).
35. Xie, H. Zhejiang University (2006).
36. Hu et al. Abnormal microspore development leads to pollen abortion in a seedless mutant of 'Ougan' mandarin (Citrus suavissima hort. ex tanaka). J Amer Soc Hort Sci (2007).
37. Balaban, R.S., Nemoto, S. & Finkel, T. Mitochondria, Oxidants, and Aging. Cell Press (2005).
38. Murphy & P., M. How mitochondria produce reactive oxygen species? 417, 105–111 (2009).
39. Tian, X., Zheng, J. & Yu, H.J. The Rice (Oryza sativa L.) Mitochondrial Genomes and Their Variations. Plant Physiology 140, 401–410 (2006).
40. Lee, Y.P. et al. Discovery of a novel cytoplasmic male-sterility and its restorer lines in radish (Raphanus sativus L.). Theoretical and Applied Genetics 117, 905–913 (2008).
41. Li, H. *Experimental principles and techniques of plant physiology and biochemistry*, (Higher Education press, 2000).
42. Liao, X. Guangxi University (2015).
43. Ping, W. et al. Comparison of Mitochondrial Genome Between A1 Cytoplasmic Male Sterile Line and Maintainer Line of Sorghum bicolor. Biotechnology Bulletin (2019).

Figures

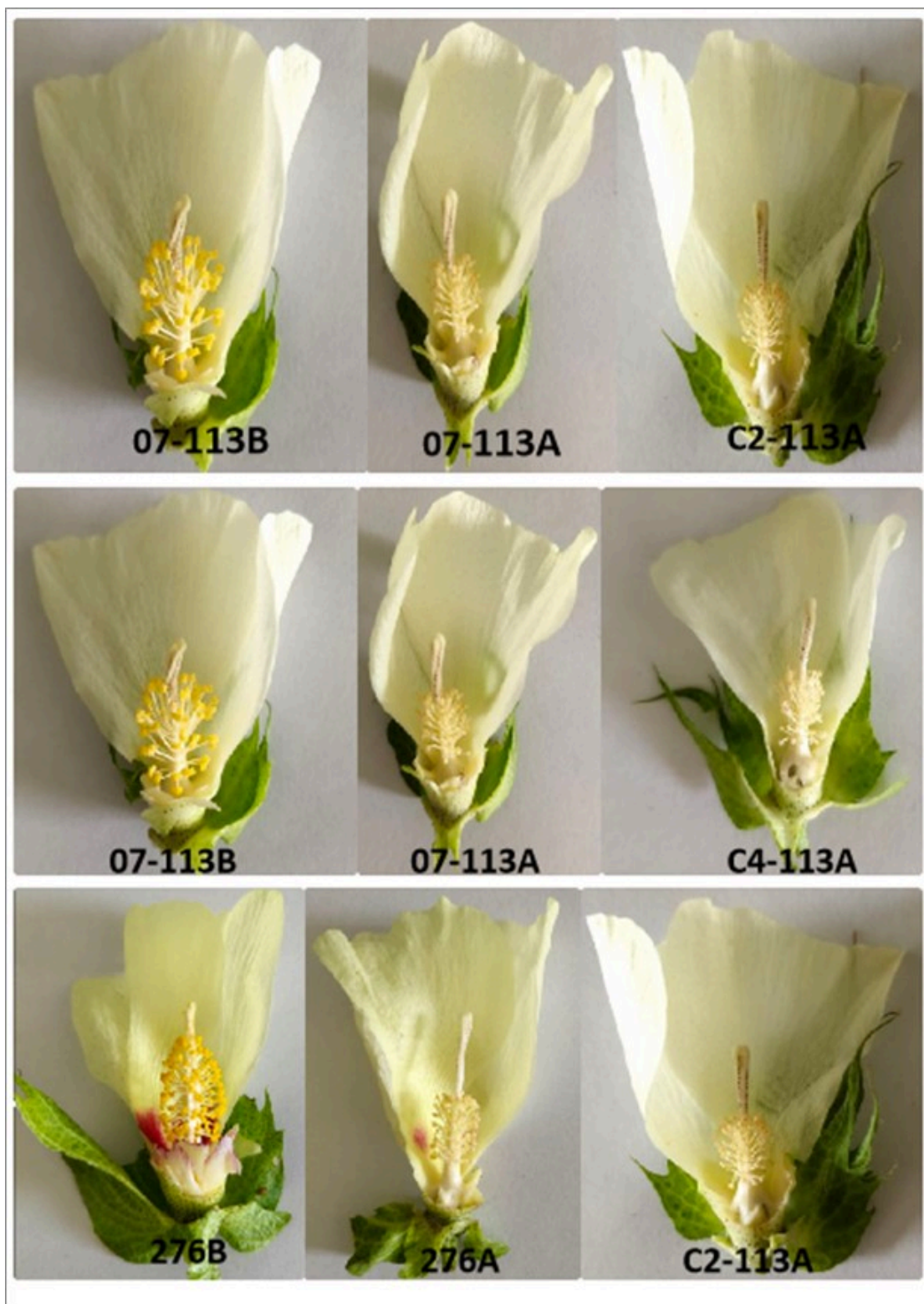


Figure 1

Morphological structure of floral organs in sterile and maintenance lines of cotton

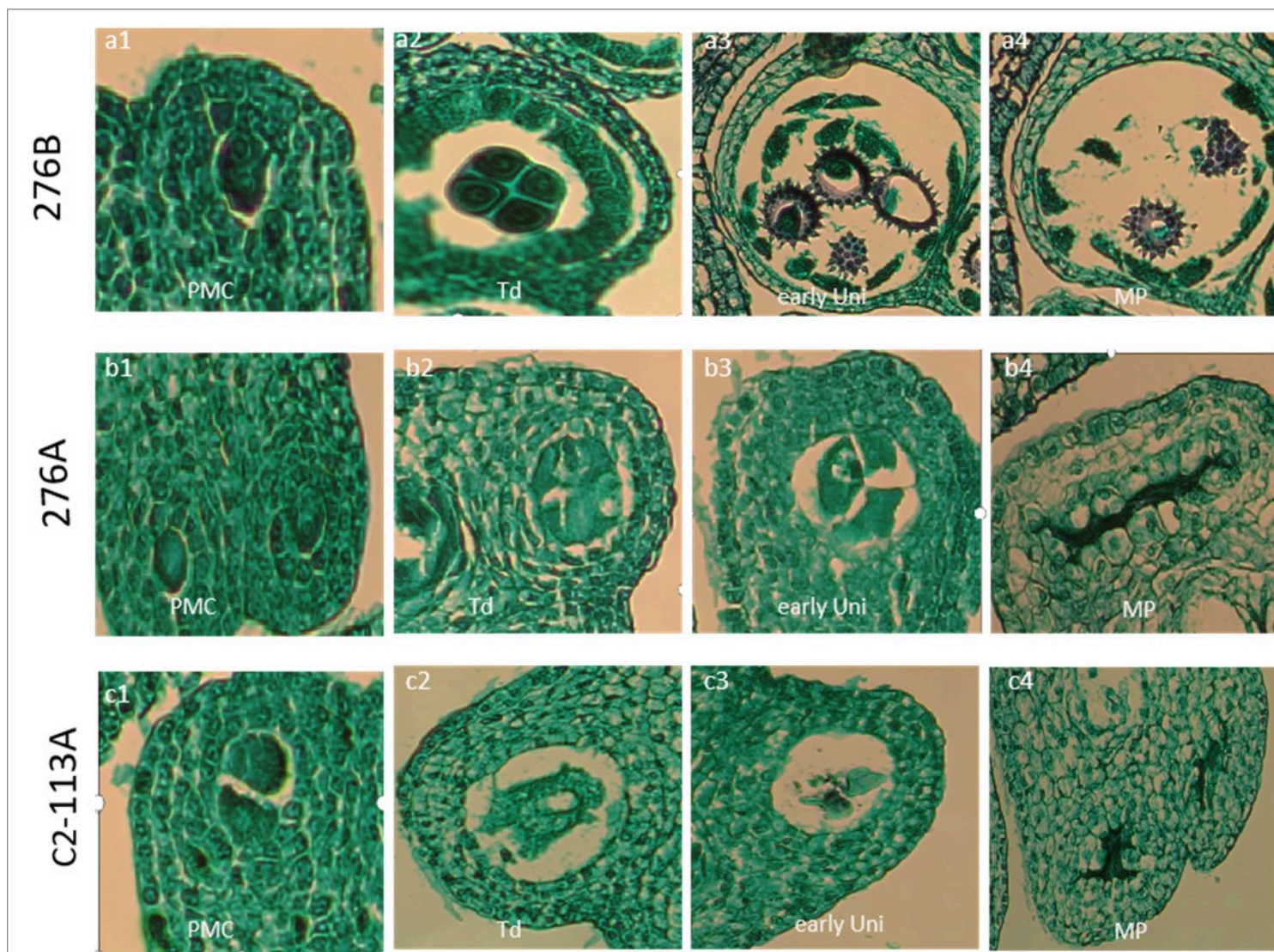


Figure 2

Paraffin section of homologous heterokaryotic CMS lineage of cotton microspore development

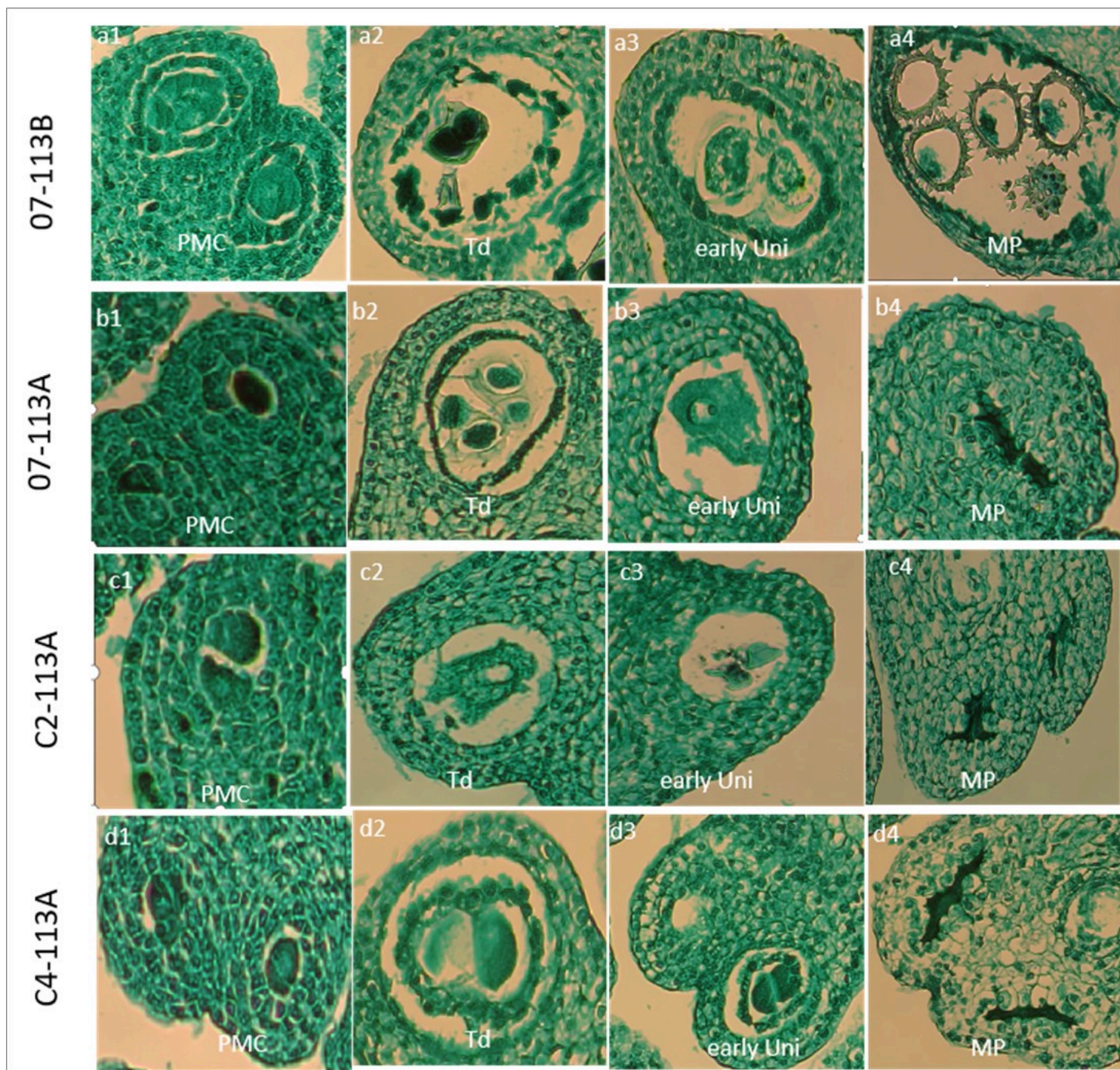


Figure 3

Paraffin section of homonuclear heterologous CMS line of cotton microspore development

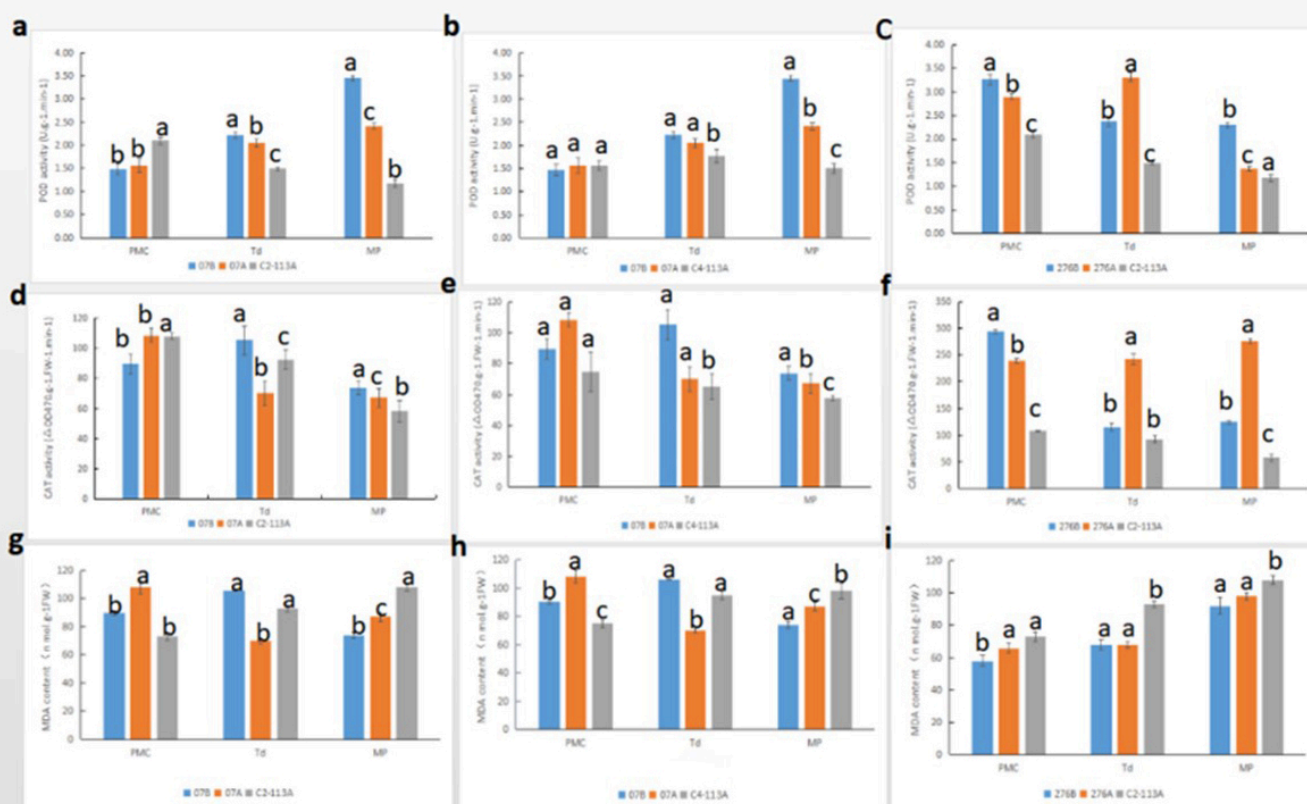


Figure 4

Active oxygen metabolism analysis. (a) POD activities of anthers in 07B, 07A, and C2-113A; (b) POD activities of anthers in 07B, 07A and C4-113A; (c) POD activities of anthers in 276B, 276A, and C2-113A; (d) CAT activities of anthers in 07B, 07A, and C2-113A; (e) CAT activities of anthers in 07B, 07A and C4-113A; (f) CAT activities of anthers in 276B, 276A, and C2-113A; (g) MDA contents in anthers of 07B, 07A and C2-113A; (h) MDA contents in anthers of 07B, 07A and C4-113A; (i) MDA contents in anthers of 276B, 276A and C2-113A;

Notes: 07A,07-113A; 07B,07-113B; PMC, Pollen mother cell; Td, Ttrad; Early Uni, Early Uninucleate; Late Uin, Late Uninucleate; MP, Mature Polle

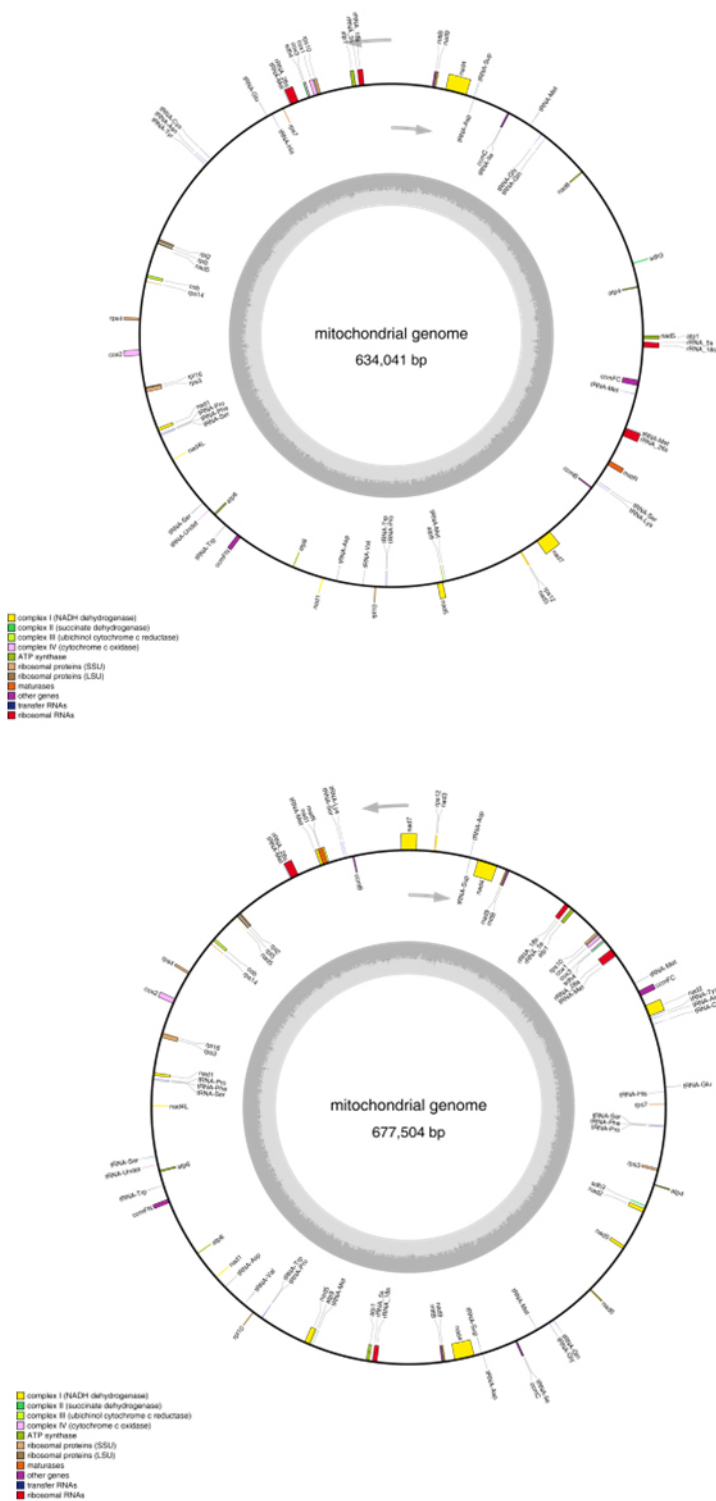


Figure 5

Mitochondrial genome circle map of 07-113A and 07-113B.

Note: The outer circle is the position coordinate of genomic components such as genes and ncRNA, with corresponding gene names; the inner circle is the content of genomic GC

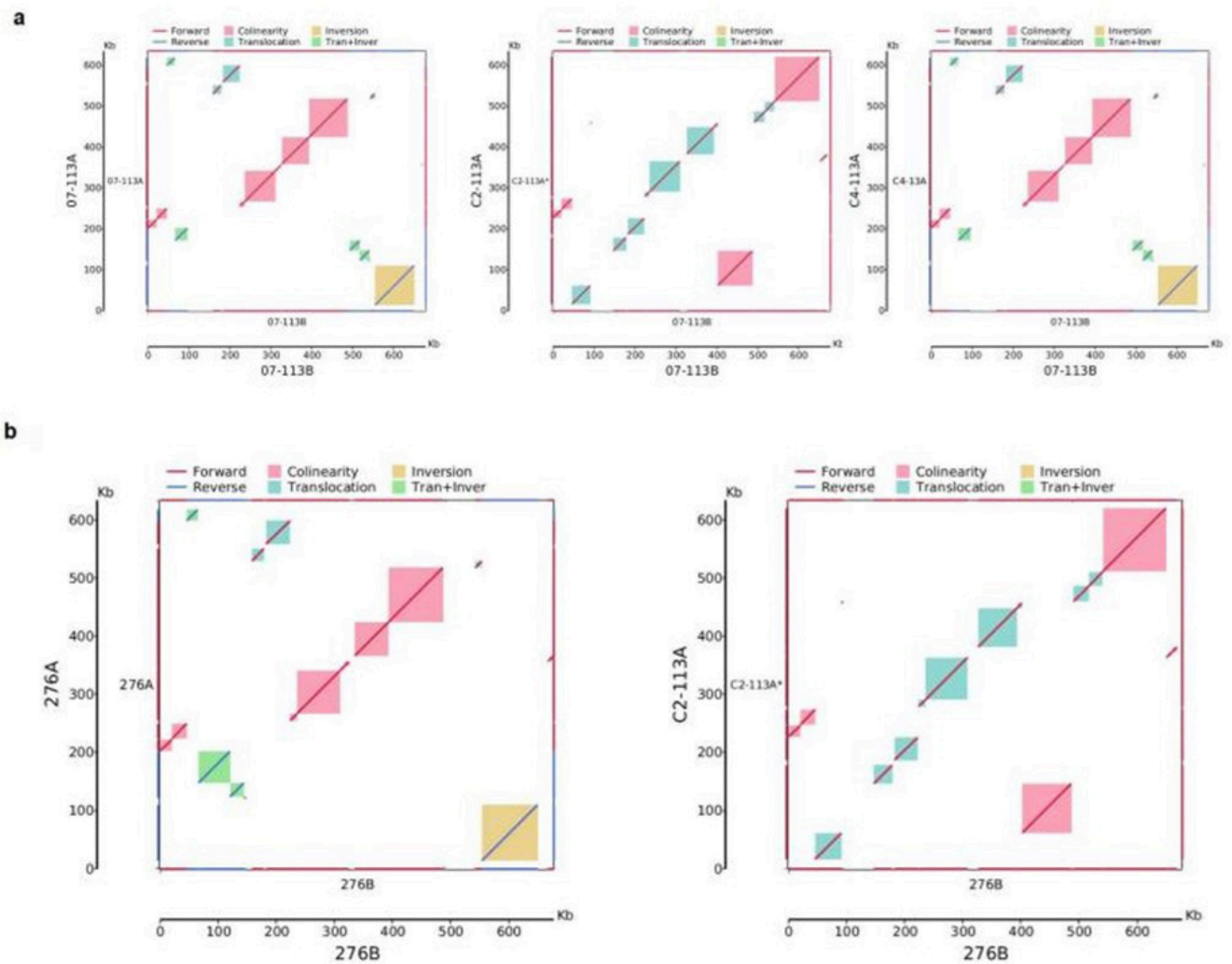


Figure 6

Two-dimensional covariance diagram of CMS lines in cotton. (a) Two-dimensional covariance diagram of homonuclear heterogeneous CMS lines in cotton;(b) Two-dimensional covariance diagram of homogeneous heteronuclear CMS lines in cotton.

Notes: The horizontal axis is the reference sequence genome, and the vertical axis is the measured genome (the "*" in the upper right corner of the name of the vertical axis sequence indicates that the negative chain of the sequence is used for comparison); The red line shows the forward alignment, and the blue shows the reverse alignment. The color of the middle square indicates the comparison type: Collinearity: homolinear comparison; Translocation: translocation comparison; Inversion: inversion comparison; Tran+Inver: translocation and inversion.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementarymaterialS.docx](#)