Chromosomal Locations of 5S and 45S rDNA in *Gossypium* Genus and Its Phylogenetic Implications Revealed by FISH

Yimei Gan[2], Fang Liu[1], Dan Chen[1], Qiong Wu[3], Qin Qin[1], Chunying Wang[1], Shaohui Li[1], Xiangdi Zhang[1], Yuhong Wang[1], Kunbo Wang[1*]

1 State Key Laboratory of Cotton Biology (China)/Institute of Cotton Research of Chinese Academy of Agricultural Sciences, Anyang, Henan, P. R. China, **2** Institute of Tropical Bioscience and Biotechnology of Chinese Academy of Tropical Agricultural Sciences/Key Laboratory of Biology and Genetic Resources of Tropical Crops, Ministry of Agriculture (China), Haikou, Hainan, P. R. China, **3** Haikou Experimental Station of Chinese Academy of Tropical Agricultural Sciences, Haikou, Hainan, P. R. China

Abstract

We investigated the locations of 5S and 45S rDNA in *Gossypium* diploid A, B, D, E, F, G genomes and tetraploid genome (AD) using multi-probe fluorescent in situ hybridization (FISH) for evolution analysis in *Gossypium* genus. The rDNA numbers and sizes, and synteny relationships between 5S and 45S were revealed using 5S and 45S as double-probe for all species, and the rDNA-bearing chromosomes were identified for A, D and AD genomes with one more probe that is single-chromosome-specific BAC clone from *G. hirsutum* (A_1D_1). Two to four 45S and one 5S loci were found in diploid-species except two 5S loci in *G. incanum* (E_4), the same as that in tetraploid species. The 45S on the 7th and 9th chromosomes and the 5S on the 9th chromosomes seemed to be conserved in A, D and AD genomes. In the species of B, E, F and G genomes, the rDNA numbers, sizes, and synteny relationships were first reported in this paper. The rDNA pattern agrees with previously reported phylogenetic history with some disagreements. Combined with the whole-genome sequencing data from *G. raimondii* (D_5) and the conserved cotton karyotype, it is suggested that the expansion, decrease and transposition of rDNA other than chromosome rearrangements might occur during the *Gossypium* evolution.

Introduction

Cotton (*Gossypium*) is an important economic fiber crop. The genus of *Gossypium* comprises about 50 species throughout tropical and subtropical regions of the world, including five tetraploid (2n=4x=52) species and about 45 diploid (2n=2x=26) species. The taxonomic and evolution study on *Gossypium* genus has been an important subject of investigation due to its economic significance. In the late 1800's or early 1900's, the taxonomy was mainly based on morphological characteristics and geographical distributions, however, it has been confusion due to un-consensus characteristics used by different taxonomists. With development of cytological methods, the diploid species have been classified genetically into seven genome types, i.e. A, B, C, D, E, F, G and K genomes [1,2]. The evolutionary history of *Gossypium* genus was reconstruct

based on geography, morphology, cytogenetics and molecular data. However, due to the continuous recombination and exchange, great differences existed between existing cotton species and their ancestors in terms of physiological feature, agronomic trait and morphology. Therefore, more interpretation about the phylogenetic and interspecific evolution in *Gossypium* genus is quite necessarily to be clarified.

Ribosomal DNA (rDNA) has highly conserved repetitive sequences in the plant genome, and the polymorphism or conservatism of their copy number and chromosomal localization are visual and comparative [3–5]. By comparing the number and distribution characteristics of rDNA sites on the chromosomes among species, interspecific phylogenetic relationships and the related mechanism of speciation and chromosomal evolution could be revealed [6]. Recently, the physical FISH location of rDNA in plants have provided much

Table 1. *Gossypium* species and their accessions used.

Species/variant	Genome	Accession/cultivar	Accession No. in nursery	Pot No. in greenhouse
G. hirsutum	A_1D_1	TM-1		
G. barbadense	A_2D_2	Pima 90-53		
G. tomentosum	A_3D_3		H0701306	H0701301
G. mustelinum	A_4D_4	A_4D_4-9	P0811807	H0804201
G. laxum	D_9		P0601001	
G. schwendimanii	D_{11}		P0602110	
G. gossypioides	D_6	D_6-2	P0814608	H0006401
G. raimondii	D_5	D_5-2	P0811506	H0006301
G. herbaceum	A_1	Hongxingcaomian		
G. herbaceum var. africanum	A_{1-a}		D2030202	H0000101
G. anomalum	B_1		P0601305	H0000201
G. capitis-viridis	B_3	B_3-1		H0004601
G. somalense	E_2	E_2-3	P0815401	H0007001
G. areysianum	E_3		P0601809	H0001901
G. incanum	E_4	E_4-4	P0815512	
G. longicalyx	F_1	F_1-3	P0815709	H0007201
G. bickii	G_1	G_1-1	P0815801	
G. nelsonii	G_3	G_3-1	P0816209	H0807601

information to the evolutional relationship of many close species and the origin of allopolyploid [7–12].

In genus of *Gossypium*, research on rDNA location in the early days was mainly focused on *G. hirsutum* (upland cotton) due to its economic importance and to the abundant genetic materials created. 5S rDNA and 18S-28S rDNA were located to chromosomes of *G. hirsutum* by FISH on chromosomes of the meiosis metaphase [13–15]. Later, the number and copy number of 5S and 18S-28S rDNA in tetraploid *G. hirsutum*, diploid species of A and D genomes, were revealed by FISH on the metaphase chromosomes of mitosis [16]. Recently, the number and copy numer of 5S and 45S rDNA, the 5S-bearing and 45S-bearing chromosomes of other tetraploid species and diploid species of A and D genomes have been revealed [17–19].

In order to further understand the cytogenetics and evolution of *Gossypium* genus, the distribution of 5S and 45S rDNA was analyzed by cocktail FISH for the four tetraploid species, as well as 13 diploid species and one variation representing diploid A, B, D, E, F and G genome. Combined with rDNA distribution in previous reports [17–19], the chromosome evolution of rDNA loci of *Gossypium* genus would be determined. Also, the phylogenetic implication based on rDNA patterns could be inferred to gain further insight into the evolutionary history of *Gossypium* genomes.

Materials and Methods

Plant materials and clones

The species and their genomes and accessions (cultivars) used in this study were shown in Table 1. The plant materials are maintained perennially in the National Wild Cotton Nursery in Sanya City, Hainan Island, sponsored and owned by the Institute of Cotton Research of Chinese Academy of Agricultural Sciences (ICR-CAAS), and at the same time, some of them are as well conserved in pots in greenhouse of ICR-CAAS at Anyang City, Henan Province, China.

For diploid D genome species and D_t subgenome of tetraploid species, four types of probes were used, including 5S rDNA, 45S rDNA, BAC clone 150D24 and some D_h chromosome-specific BAC clones. For diploid A genome and A_t subgenome of tetraploid species, three types of probes were used, including 5S rDNA, 45S rDNA and some A_h chromosome-specific BAC clones. For the diploid B, E, F and G genome species, only 5S rDNA and 45S rDNA were used. The 5S and 45S rDNA derived from *Arabidopsis thaliana* were kindly provided by Professor Yunchun Song of Wuhan University, China. The BAC clone 150D24 which contains centromere-specific repeats in D subgenome and D genome of *Gossypium* was screened from Pima 90-53 BAC library [20] to indicate centromere position. The A_h (D_h) chromosome-specific BAC clones used to identify individual chromosome were kindly provided by Professor Tianzhen Zhang of Nanjing Agricultural University, China [21].

DNA probes preparation

The probes 5S, 45S rDNA and BAC DNA were isolated using a standard alkaline extraction [22]. 45S rDNA and BAC clone 150D24 were labeled by standard Dig-nick translation reactions, whereas 5S rDNA and some A_h (D_h) subgenome chromosome-specific BAC clones [21] were labeled with Biotin-nick translation reactions, according to the manufacturer's instructions (Roche Diagnostics, USA).

Chromosome preparation and FISH

Preparation of mitotic chromosomes and the FISH procedure were conducted according to [23] with some modifications. Digoxigenin-labeled and biotin-labeled probes were detected

by anti-digoxigenin-rhodamine (red) and avidin-fluorescein (green) (Roche Diagnostics, USA), respectively. For conducting the probe-cocktail mixture, gDNA was used as block DNA instead of Cot-1 DNA. The dose of block DNA was 200 times of the chromosome-specific BAC DNA. Chromosomes were counterstained by 4', 6-diamidino-2-phenylindole (DAPI) in the antifade VECTASHIELD solutions (Vector Laboratories, Burlingame, CA). The hybridization signals were observed using a fluorescence microscope (Leica MRA2) with a charge-coupled device (CCD) camera (Zeiss) and arranged using Adobe Photoshop 7.0.

Results

The number of 5S and 45S rDNA in *Gossypium* genus

Three 45S rDNA loci and two 5S rDNA loci were detected in all three tetraploid species (Figure 1 a, b, c). Similarly, the number of 45S rDNA loci was detected three, two, four and three in D_5, D_6, D_9, D_{11}, respectively, while only one 5S rDNA locus was observed in the four D genome species (Figure 1d–1g). In A_1 and its variant A_{1-a}, three 45S loci and one 5S rDNA locus were found, respectively (Figure 1 h, 1i). In B_1, B_3, E_2, E_3, F_1 and G_3, three 45S loci and one 5S rDNA locus were observed (Figure 1j–1m, 1p), while three 45S loci and two 5S rDNA loci in E_4 (Figure 1n) as well as four 45S loci and one 5S rDNA locus in G_1 were observed (Figure 1q).

The location of 5S and 45S rDNA in *Gossypium* genus

To further identify the rDNA locations specific to individual chromosomes or even to arms in D subgenome of tetraploid and D genome cottons, the individual chromosome-specific BAC clones and a D genome centromere-specific BAC clone (150D24) as BAC-FISH probes were used in the experiments. And for D_t subgenome of tetraploid, the individual chromosome BAC clones was used as BAC-FISH probes. Figure 2a–2w showed the 45S and 5S rDNA locations specific to individual chromosomes or even to arms in the three tetraploid cottons (A_1D_1, A_3D_3, A_4D_4 and A_2D_2) and the four D genome cottons (D_9, D_{11}, D_6, D_5). In both A_1D_1 (Figure 2a–2c) and A_3D_3 (Figure 2d–2f), 45S and 5S rDNA, syntenic with BAC clones specific to chromosomes A_h09, D_h07 and D_h09 (h indicates A_1D_1), respectively, were located to the corresponding chromosomes and chromosomal arms. According to the homology within D subgenomes, chromosomes bearing with 45S locus of A_3D_3 were named as $A_{tt}09$, $D_{tt}07$ and $D_{tt}09$ (tt indicates A_3D_3), respectively. So that in these two species, three 45S loci were observed at the end of the short arm of chromosomes, whereas two 5S loci were co-localized with the 45S locus and was found interstitial on the short arm of chromosomes A_t09 and D_t09 (t indicates tetraploid species), respectively, suggesting a synteny relationship for 5S and 45S rDNA. In A_4D_4, three 45S loci, syntenic with BAC clones specific to chromosomes A_h07, A_h09 and A_h08 respectively were found located at the end of the short arm of chromosomes of A subgenome (Figure 2g–2i), while the two 5S loci, syntenic with BAC clones specific to chromosomes A_h09 and D_h09, respectively, were located to the end of the short arm of chromosomes (Figure 2h, 2j). Likewise, chromosomes bearing

45S and 5S rDNA loci were named as A_m07, A_m09, A_m08 and D_m09, respectively (m indicates A_4D_4). And so, 5S and 45S rDNA, being positioned on the chromosome A $_m09$, showed a synteny relationship, and the other 5S rDNA was positioned on the on the chromosome D $_m09$ showing no synteny relationship with any of 45S loci. In addition, 5S and 45S rDNA on the A subgenome of A_2D_2 was located to the same chromosome A $_b09$ (b indicates A_2D_2) (Figure 2k).

Figure 2l–2w showed the 45S and 5S rDNA locations specific to individual chromosomes or even to arms in the four D genome cottons (D_9, D_{11}, D_6, D_5). In D_9, 45S and 5S rDNA, syntenic with BAC clones specific to chromosomes D_h05, D_h07, D_h09 and D_h12, respectively, were located to the corresponding chromosomes and chromosomal arms (Figure 2l–2o). According to the homology between D genomes and D subgenome of A_1D_1, chromosomes bearing 45S locus of D_9 were named as D_905, D_907, D_909 and D_912, respectively. The result showed that four 45S rDNA loci were found at the end of short arm of chromosomes D_905, D_907, D_909 and D_912, with one 5S rDNA locus was found interstitial to the short arm of chromosome D $_909$. Likewise, 45S and 5S rDNA, syntenic with BAC clones specific to chromosomes D_h05, D_h07 and D_h09, were found respectively located to the corresponding chromosomes and chromosomal arms of D_{11} (Figure 2p–2r). Chromosomes bearing with 45S locus of D_{11} were named as $D_{11}05$, $D_{11}07$ and $D_{11}09$, respectively. Three 45S rDNA loci were showed at the end of short arm of chromosomes $D_{11}05$, $D_{11}07$ and $D_{11}09$, and one 5S rDNA locus was found interstitial to the short arm of chromosome D $_{11}09$. Also, 45S and 5S rDNA, syntenic with BAC clones specific to chromosomes D_h07 and D_h09, respectively, were located to the corresponding chromosomes and chromosomal arms of D_6 (Figure 2s, 2t). Chromosomes bearing with 45S locus of D_6 were named as D_607 and D_609, respectively. Therefore, two 45S rDNA loci were seen at the end of short arm of chromosomes D_607 and D_609, and one 5S rDNA locus was found interstitial to the short arm of chromosome D $_609$. In D_5, two 45S loci and one 5S rDNA, syntenic with BAC clones specific to chromosomes D_h09 and D_h11, respectively, were located to the corresponding chromosomes (Figure 2u, 2w). Since the chromosome bearing the third 45S could not be identified with BAC clones derived from *G. hirsutum* (Gan unpublished), the third 45S was therefore identified to chromosome 02 (Figure 2v) with BAC clone screened from A_2D_2 (Qinqin unpublished). According to the homology between D genome and D subgenome, chromosomes bearing with 45S locus of D_5 were named as D_509, D_511 and D_502, respectively. And, three 45S loci were shown at the end of short arm of chromosomes D_509, D_511 and D_502, while one 5S rDNA locus was found interstitial to the short arm of chromosome D $_509$.

Figure 3a–3f showed the 45S and 5S rDNA locations specific to individual chromosomes in A_1 and A_{1-a}. Three 45S and one 5S rDNA, syntenic with BAC clones specific to chromosomes A_h05, A_h07 and A_h09, were located to the corresponding chromosomes. According to the homology between A genomes and A subgenome, chromosomes bearing 45S locus were named as A_105, A_107 and A_109 (for A_1) (Figure 3a–3c), $A_{1-a}05$, $A_{1-a}07$ and $A_{1-a}09$ (for A_{1-a}) (Figure 3d–3f), respectively.
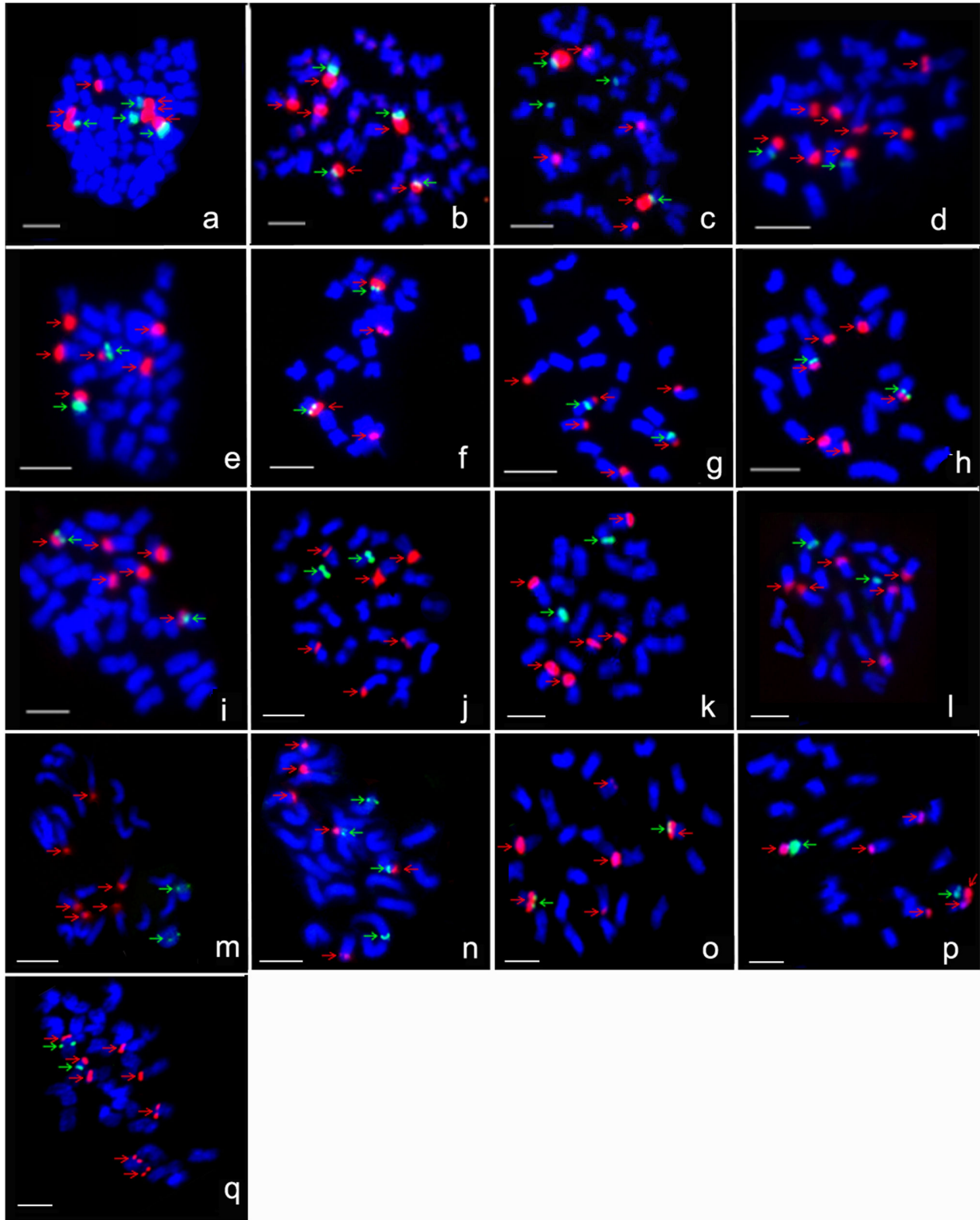
Figure 1. The number of rDNA in AD, D, A, B, E, F and G genomes. 5S rDNA: green fluorescence signals marked with green arrows; 45S rDNA: red fluorescence signals marked with red arrows. a: A_1D_1, b: A_3D_3, c: A_4D_4, d: D_5, e: D_6, f: D_9, g: D_{11}, h: A_1, i: A_{1-a}, j: B_1, k: B_3, l: E_2, m: E_3, n: E_4, o: F_1, p: G_3, q: G_1. Bar =5 μm.
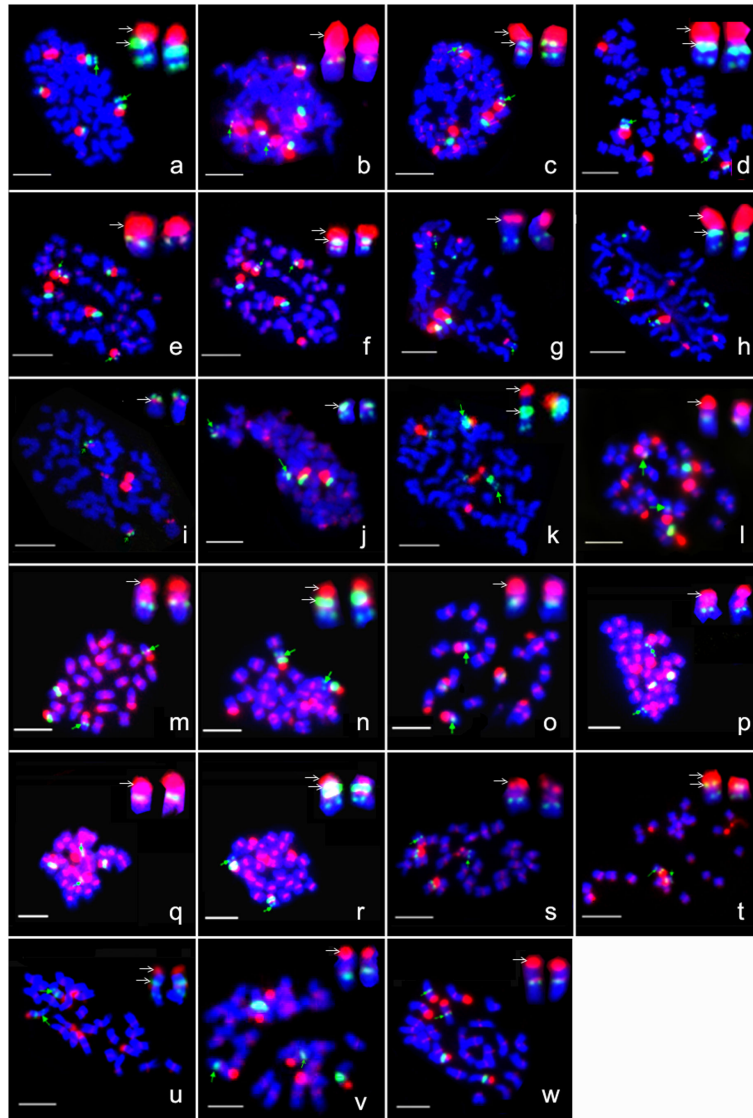
Figure 2. Locations of 5S and 45S rDNA in four tetraploid and four D-genome species. green and weak fluorescence signals with green arrow; 5S rDNA: green fluorescence signals; 45S rDNA: red fluorescence signals. For D-genome species and D-subgenome, the short arm and the long arm were distinguished by the location of 150D24 with red fluorescence signals on intercalary chromosomes. Marked chromosomes with green arrow were enlarged at the top-right corner with the short arm on the top, and the 45S or 5S signals were marked with white arrow. Bar =5μm.

a–c: FISH images with 45S on chromosomes A_h09, D_h07 and D_h09 and that with 5S on chromosomes A_h09 and D_h09 for A_1D_1, respectively.

d–f: FISH images with 45S on chromosomes $A_{tt}09$, $D_{tt}07$ and $D_{tt}09$ and that with 5S on chromosomes $A_{tt}09$ and $D_{tt}09$ for A_3D_3, respectively.

g–j: FISH images with 45S on chromosomes A_m07, A_m09 and A_m08 (g, h, i) and that with 5S on chromosomes A_m09 and D_m09 (h, j) for A_4D_4, respectively.

k: FISH images with 45S and 5S on chromosome A_b09 of A_2D_2.

l–o: FISH images with 45S and 5S on chromosomes D_905, D_907, D_909 and D_912 and that with 5S on chromosome D_909 for D_9, respectively.

p–r: FISH images with 45S and 5S on chromosomes $D_{11}05$, $D_{11}07$ and $D_{11}09$ and that with 5S on the chromosome $D_{11}09$ for D_{11}, respectively.

s, t: FISH images with 45S and 5S on chromosomes D_607 and D_609 and that with 5S on chromosome D_609 for D_6, respectively.

u–w: FISH images with 45S and 5S on chromosomes D_509, D_502 and D_511 and that with 5S on chromosome D_509 for D_5, respectively.
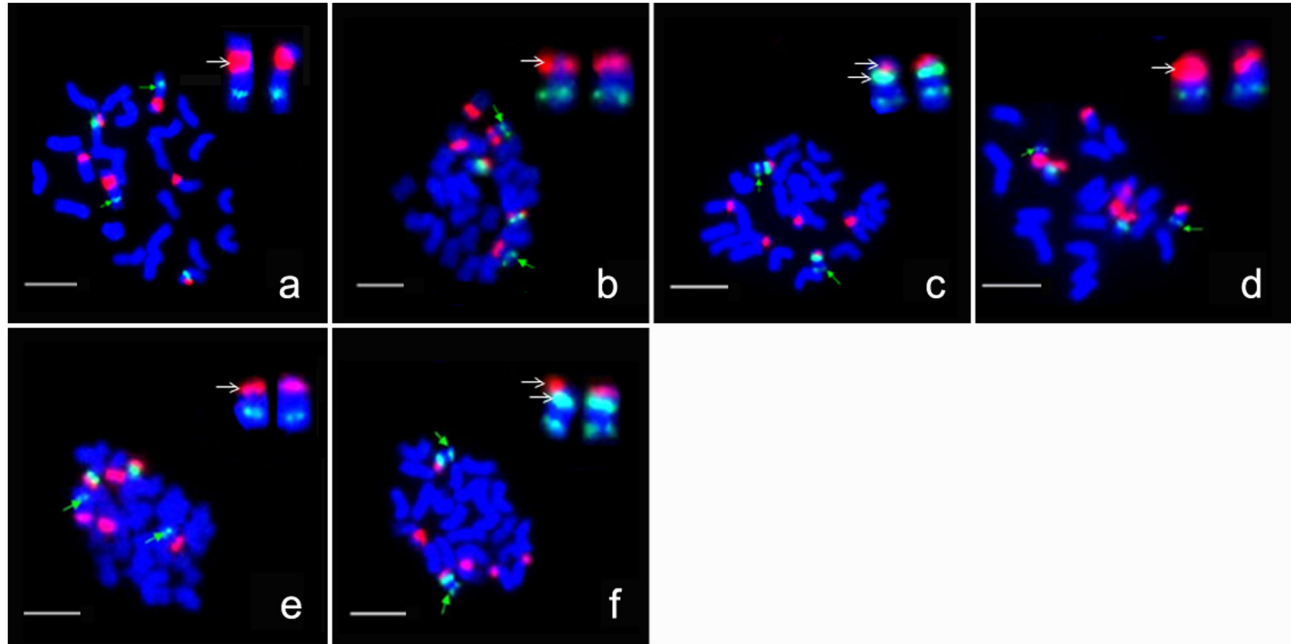
Figure 3. Locations of 5S and 45S rDNA in two A-genome species by dual-FISH. BAC DNA: green and weak fluorescence signals with green arrow; 5S rDNA: green fluorescence signals; 45S rDNA: red fluorescence signals. Marked chromosomes with green arrow were enlarged at the top-right corner with the short arm on the top, and the 45S or 5S signal were marked with white arrow. Bar =5µm.

a–c: FISH images with 45S and 5S on chromosomes A_105, A_107 and A_109 and that with 5S on chromosome A_109 for A_1, respectively.

d–f: FISH images with 45S and 5S on chromosomes $A_{1\text{-}a}05$, $A_{1\text{-}a}07$ and $A_{1\text{-}a}09$ and that with 5S on chromosome $A_{1\text{-}a}09$ for $A_{1\text{-}a}$, respectively.

Therefore, in both A_1 and $A_{1\text{-}a}$, three 45S rDNA loci were revealed at the end of short arm of chromosomes A_105 ($A_{1\text{-}a}05$), A_107 ($A_{1\text{-}a}07$) and A_109 ($A_{1\text{-}a}09$), and one 5S locus was located on chromosomes A_109 (Figure 3c) and $A_{1\text{-}a}09$ (Figure 3c), respectively.

There is limited genomic information available for species of B, E, F and G genomes, and the chromosome identification for BAC clones is unavailable either, and so the location of 5S and 45S rDNA has had been identified by using rDNA-FISH with only 5S and 45S rDNA probed. In B_1 and B_3, one 5S locus and two 45S loci were observed at the end of chromosomes, while one 45S locus was found nearby centromere, displaying the satellite-intermediate type, which is extremely rare in *Gossypium* species (Figure 1j, 1k). In E_2, E_3, E_4, 5S and 45S were both located at the end of chromosomes (Figure 1l–1n). In F_1, two 45S rDNA loci and the third 45S rDNA locus syntenic with 5S rDNA were located to the end and near centromere of chromosomes, respectively, while the 5S rDNA locus of F_1 was positioned outside the 45S rDNA and to the end of chromosome (Figure 1o). In G_1 and G_3, 45S and 5S rDNA were located at the end and near centromere of chromosomes, respectively (Figure 1p, 1q). And the relationship between 5S and 45S is nonsyntenic in B_1, B_3, E_2 and E_3, while it showed syntenic in F_1, G_1 and G_3. Among two 5S loci in E_4, one is

syteny with 45S rDNA while the other one is not synteny with 45S rDNA.

Chromosome distribution of rDNA in *Gossypium* genus

In order to compare and analyze the evolutional relationship among cotton species, the chromosomal distribution of 5S and 45S for 17 species and one variant in the present study and 8 species presented in previous papers [17–19] are summarized in Table 2 and displayed schematically in Figure 4.

The 5S rDNA loci maintained high homogeneity in the number (except E_4), distribution and the copy number among 25 species and one variant, only varied partly in copy numbers in five tetraploid species (Table 2 and Figure 4). In all diploid species except E_4, only one 5S rDNA locus was found. In tetraploid species and diploid species of A and D genome, all 5S rDNA loci were located to chromosomes 09 ($A_t09/D_t09/A_g09/D_g09$). The number of 45S rDNA is conserved in some species but is various in other species, no matter related with their genome resource. Even the same number of 45S loci was observed, they were still positioned to different chromosomes. Beside the number and chromosomes bearing 45S, the copy number (identified by signal intensity) of 45S rDNA was observed to be similar to some extend, but varied to a great extent. They varied in either different species or different

Table 2. Distribution of rDNA in *Gossypium* genus.

Genome	Species(short name)	(No.) Chr. bearing 45S	(No.) Chr. bearing 5S	Relationship between 5S and 45S	Source
AD	*G. hirsutum* (A_hD_h)	(3)A_h09, D_h07, D_h09	(2)A_h09, D_h09	synteny	This study
	G. barbadense (A_2D_2)	(3)A_b09, D_b07, D_b09	(2)A_b09, D_b09	synteny	[17]
	G. tomentosum (A_3D_3)	(3)A_{tt}09, D_{tt}07, D_{tt}09	(2)A_{to}09,D_{to}09	synteny	This study
	G. mustelinum (A_4D_4)	(3)A_m09, A_m07, A_m08	(2)A_m09,D_m09	synteny/nonsynteny	This study
	G. darwinii (A_5D_5)	(3)A_d09, D_d07, D_d09	(2)A_d09, D_d09	synteny	[18]
A	*G. herbaceum* (A_1)	(3) 09, 07, 05	(1) 09	synteny	This study
	G. herbaceum var. Africanum (A_{1-a})	(3) 09, 07, 05	(1) 09	synteny	This study
	G. arboreum (A_2)	(3) 09, 07, 05	(1)09	synteny	[19]
D	*G. thurberi* (D_1)	(4) 09, 07, 03, 11	(1) 09	synteny	[17]
	G. armourianum (D_{2-1})	(3) 09, 07, 05	(1) 09	synteny	[18]
	G. davidsonii (D_{3-d})	(4) 09, 07, 05, 12	(1) 09	synteny	[18]
	G. klotzschianum (D_{3-k})	(4) 09, 07, 05, 12	(1) 09	synteny	[18]
	G. aridum (D_4)	(3) 09, 07, 05	(1) 09	synteny	[18]
	G. raimondii (D_5)	(3) 09, 11, 02	(1) 09	synteny	This study
	G. gossypioides (D_6)	(2) 09, 07	(1) 09	synteny	This study
	G. trilobum (D_8)	(4) 09, 07, 03, 11	(1) 09	synteny	[17]
	G. laxum (D_9)	(4) 09, 07, 05, 12	(1) 09	synteny	This study
	G. schwendimanii (D_{11})	(3) 09, 07, 05	(1) 09	synteny	This study
B	*G. anomalum* (B_1)	(3) Unknown	(1) Unknown	nonsynteny	This study
	G. capitis-viridis (B_3)	(3) Unknown	(1) Unknown	nonsynteny	This study
E	*G. somalense* (E_2)	(3) Unknown	(1) Unknown	nonsynteny	This study
	G. areysianum (E_3)	(3) Unknown	(1) Unknown	nonsynteny	This study
	G. incanum (E_4)	(3) Unknown	(2) Unknown	synteny/nonsynteny	This study
F	*G. longicalyx* (F_1)	(3) Unknown	(1) Unknown	synteny	This study
G	*G. nelsonii* (G_3)	(3) Unknown	(1) Unknown	synteny	This study
	G. bickii (G_1)	(4) Unknown	(1) Unknown	synteny	This study

chromosomes of the same species. In addition, the syntenic relationship between 5S and 45S rDNA were divided into three groups. One is syntenic for A, D, F and G genome, and the second is nonsyntenic for B and two species of E genome. Then the third one is both syntenic and nonsyntenic in E_4 which has two 5S rDNA loci.

Discussion

Genomic and evolutionary researches in *Gossypium* have obtained great progress with morphological characteristics, geographical distribution, cytogenetic and molecular data. And the evolutionary history has been built into phylogenetic tree [24]. However, our understanding of evolution and chromosome structure is still extremely limited. Most of *Gossypium* species have rather small chromosomes, which in many cases are similar in shape and size, and therefore are difficult to be distinguished. In this study, we used chromosome-specific BAC clones from *G. hirsutum*, 5S and 45S rDNA as multiple probes to have FISH located precisely the 5S and 45S rDNA of tetraploid species and species of their donor genome (A and D genome). And species of other genomes are FISH located with 5S and 45S rDNA as double probes. In general, 5S and 45S rDNA were revealed for both conservation and polymorphism in the number of rDNA loci, the

number of rDNA repeats, chromosome-bearing rDNA and the synteny relationship between 5S and 45S rDNA. The rDNA pattern is generally in accordance with phylogeny with some disagreements. It is noted that, the chromosomes bearing 5S and 45S rDNA have been identified accurately for A and D genome as well as the tetraploid species using chromosome-specific BAC clones. Therefore, location of 5S and 45S rDNA in our study provided more detailed and comprehensive information for the evolution of *Gossypium* genus than the previous reports, which provided insights into molecular evolutionary changes.

Chromosomal patterns of 5S and 45S rDNA

The 5S rDNA loci among 25 cotton species (Table 2 & Figure 4) were observed to be highly conserved in this study. Only one 5S rDNA locus was found in all diploid species except E_4 and two loci in tetraploid species. The high conservation of 5S rDNA is a common phenomenon in plant genera such as *Oryza* genus [11], *Arichis* genus [8], and Quadrifaria group of *Paspalum* genus [25]. The conserved distribution of 5S rDNA might be associated with their location at pericentromeric regions [26], which rarely related to the chromosomal structure rearrangement [27]. Also, the recombination intercalary site of
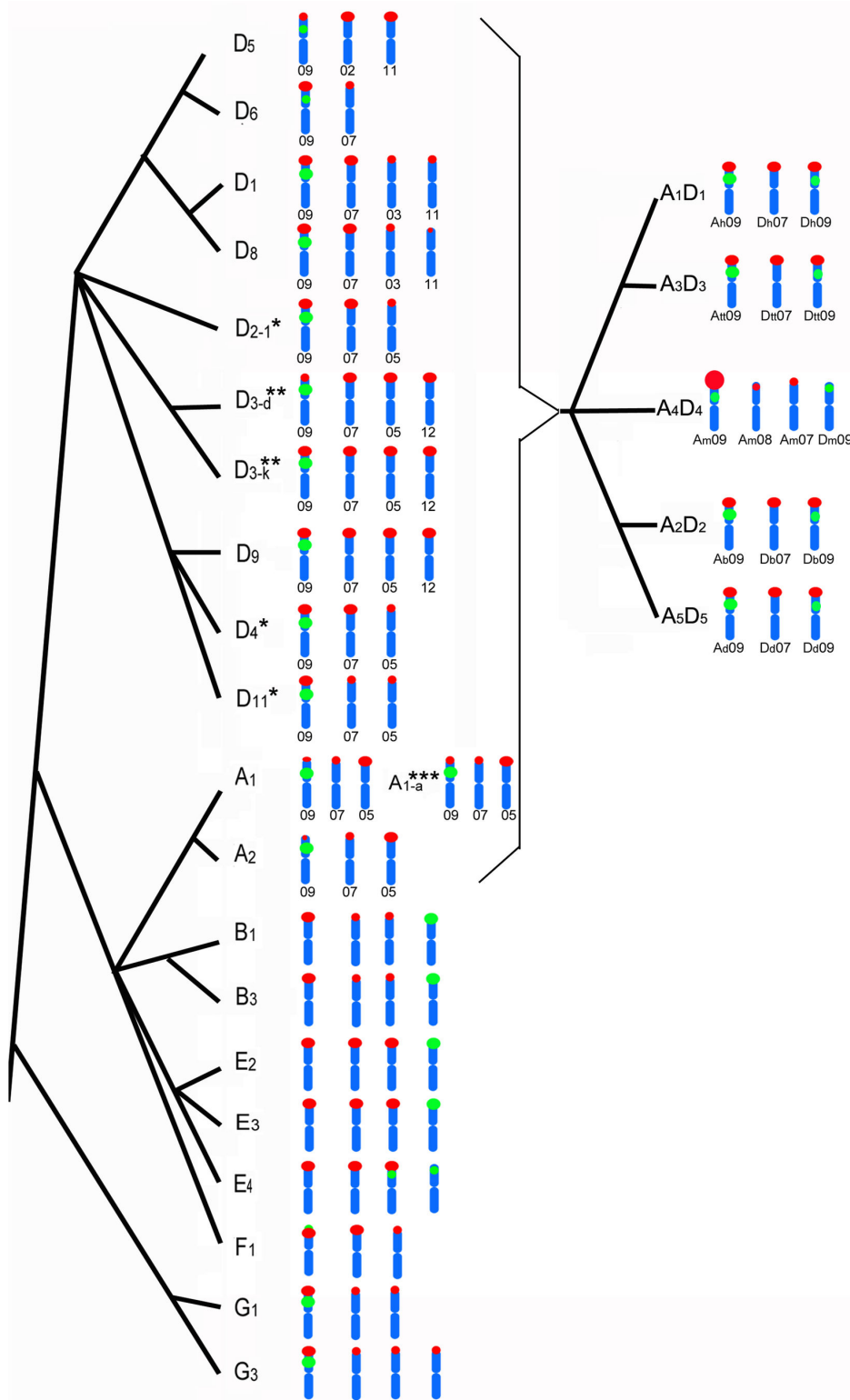
Figure 4. Simplified phylogeny of *Gossypium* genus included in idiograms chromosomes bearing 5S and 45S rDNA signals. The phylogeny is according to the reference of Wendel [24]. The words under the chromosomes indicated the chromosomes bearing the rDNA locus. 5S rDNA: green signals; 45S rDNA: red signals. * and ** indicate the differences according 45S rDNA pattern in the present study from previous report [24], respectively; *** indicates *G. herbaceum var. africanum*.

doi: 10.1371/journal.pone.0068207.g004

chromosome is much lower than that of ones at the end of chromosome in upland cotton [21].

The distribution of 45S rDNA among 25 cotton species (Table 2 & Figure 4) is both conserved and polymorphic, in accordance with other reports showing similar pattern [11,16]. The number variations of rDNA among plants of the same ploidy level have been attributed to chromosomal rearrangements, transposable events and gene silence [28]. According to the data of the whole genome sequencing of *G. raimondii* which is the smallest *Gossypium* species, a high proportion of transposable elements such as the *gypsy* and *copia*-like LTRs were found [29], suggesting that the most possible mechanisms associated with 45S rDNA variation could be transposon mobility. Taking conservative karyotypes of *Gossypium* interspecies into account, major chromosomal structural rearrangements are not frequent among species. Therefore, in *Gossypium* genus, the key mechanism facilitating diversification of 45S rDNA distribution patterns should be considered as transpositions rather than chromosome rearrangements. Moreover, the copy numbers of 45S rDNA loci was discrepant in the different chromosomes of the same species or the corresponding chromosomes of different species. The copy numbers of rDNA repeats might be amplified or decreased by unequal crossing over to the extent that these new sites can be detected by FISH.

Besides the numbers and copy numbers of rDNA, some tendencies of syntenic relationship between 5S and 45S rDNA have been as well indicated. It is syntenic for A, D, F and G genome, but nonsyntenic for B and E genome. The diploid species were divided into Australian species (C, G and K genome), American species (D genome) and Africa-Asian species (A, B, E and F genome), and the last one is considered as original species [30]. Species of B and E genome could be the most original according to the chromosomal pairing analysis in interspecies hybridization [31], electrophoretic analysis of seed protein [32] and phenotypic relationship analysis [33]. According to the evolution pattern, more advanced species evolved from the distribution center to the edge of the expansion. It is considered that Africa species (A, B, E and F genome) originated from Africa which was a distribution centre of cotton [30]. Therefore, the nonsynteny relationship in species of B and E genomes might be related the original types.

Phylogenetic implications with rDNA pattern for *Gossypium* genus

Phylogenetic implications for tetraploid species.
The variations in rDNA distribution are with phylogenetic implications, for the closeness of taxa is correlated to the similarity of their rDNA FISH patterns [6]. Theoretically, the rDNA loci of tetraploid cottons should be the sum of rDNA loci of its putative diploid ancestors (A and D genome). In the present study, the sum of 5S rDNA loci of species of A and D genome was equal to that of tetraploid cottons, while 45S rDNA loci decreased in tetraploid cottons. The nonadditive contribution of rDNA during the evolution of polyploidy species has been described in several plant genera [10,16,28,34]. Several hypotheses may explain the reason. Decreases in site

number could arise to stabilize new genomes by the formation of translocations with breakpoints proximal to the rDNA sites during the formation of polyploidy [35]. And, deletions may have eliminated loci in modern tetraploid cottons. Also, new 45S rDNA loci may have been formed by transposition of sequences containing rDNA repeats in the modern A and D-diploids.

Additionally, the copy number of 45S rDNA locus on chromosome A_t09 in tetraploid species is much higher than that of chromosome A_g09 in donor genome, but that of 5S rDNA locus of D_t09 in tetraploid species reduced obviously relative to that of D_g09 in donor genome. It is possible that the copy number of 45S rDNA locus on chromosome A_g09 in original parental species is very high, but unfortunately the species died out [35]. Or the copy number of 45S rDNA locus on chromosome A_t09 could increase after the formation of polyploidy or the subsequent evolution. Compared to parental species, 5S rDNA locus in polyploidy species according to the researches on other genus reported tended to eliminate [6,16,34–36]. Some hypotheses may explain why 5S rDNA locus disappeared or their copy number reduced following polyploidization of tetraploid cottons. Firstly, the copy number of 5S rDNA locus could be very low in the original parental species which may extinct [37]. Secondly, the copy number decreased during the course of the formation of tetraploid species. Thirdly, the copy number did not decrease in the modern tetraploid species evolved from the original ones. And the related mechanisms account for the changes could be unequal crossing over, gene exchange and transposons events, and so on [3].

Phylogenetic implication for American diploid species (D genome).
Thirteen species of D genome containing six subsections, have received considerable phylogenetic attentions [30,38–40], but evolutionary relationships among these subsections still have not enough evidence [24]. Some evolutionary evidences could be obtained from the rDNA patterns which were revealed both conserved and changeable in ten species of six subsections studied here. The evidence is that it is conserved in the chromosomal location, copy number and synteny relationship of 5S and 45S rDNA on chromosomes D_g07 and D_g09 in ten species except D_5 and D_{3-d} (Table 2 and Figure 4). D_5, divided into subsection *Austroamericana* Fryxell at the end of D genome [1], the 45S loci on chromosomes D_509, D_511 and D_502 differ greatly with other species (Table 2 and Figure 4), suggesting the greater evolutionary history than that of other D-genome species. D_6 is the sole representative of subsection *Selera* Fryxell with distinctive morphological characteristics [41]. It has only two 45S loci on chromosomes D_607 and D_609 (Table 2 and Figure 4), which is in agreement with that of D_6 and could be used as the base of D genome species [34]. D_1 and its sister species D_8, the two representative of subsection *Houzingenia* [30,39,42], both has four 45S loci on chromosomes $D_107(D_807)$, $D_109(D_809)$, $D_103(D_803)$ and $D_111(D_811)$ (Table 2 and Figure 4 [17]). The latter two 45S loci could be more unique to the two species rather than other species, which is in accordance with their clades showing different from other three clades in cladogram [24]. The rest six species, derived from the same clade but

being divided into three subsection [24], have common 45S loci on chromosomes D_g07, D_g09 and D_g05 supporting the same clade, however the variation of the last 45S locus distribution suggested two divisions among the six species as the following: On one hand, D_{3-k} and D_{3-d} of subsection *Integrifolia* Todaro and D_9 of subsection *Erioxylum* all have four 45S rDNA on the chromosomes D_g07, D_g09, D_g05 and D_g12, suggesting that they might be closer than with other three species. And copy numbers of the 45S locus on chromosome $D_{3-d}09$ is of much less than other nine species of D genome including D_{3-k} (Table 2 Figure 4 * [18]). On the other hand, the other three species, D_{2-1} of subsection *Caducibracteolata* Mauer as well as D_4 and D_{11} of subsection *Erioxylum*, all have three 45S loci on the chromosomes D_g07, D_g09 and D_g05 (Table 2 Figure 4 ** [17]), suggesting that they might be closer than their relationship previously reported [24].

Phylogenetic implication for African–Asian diploid species (A, B, E and F genome). The A, B, E and F genomes belong to the same clade, which are different from other genomes, according to the phylogenesis history of *Gossypium* genus [24]. Species of these four genomes have same numbers of 45S and 5S loci (except E_4) but varied significantly in synteny relationship. A and F genomes with the same rDNA pattern could associate with their proposed sister relationship [24]. And the 5S locus in F_1 was found near the satellite outside the 45S rather than near centromere inside the 45S unlike other species. Philips [31] proposed that F_1 should be removed from E genome and classified as F genome, as F_1 could be a new cellular type according to cytology research. F_1 with the synteny relationship of 5S and 45S is different from E genome with the nonsynteny relationship, providing more visual evidence of the further establishment of F genome. Besides, the 5S and 45S rDNA at telomere ends in F_1 are close to each other, which is rare in *Gossypium* species though it has been reported in other plants [43–45], suggesting it may relate to the stabilization of centromeric fission products [45,46].

And, the similar distribution of 5S and 45S loci was observed in B_1 and B_3. It was not extensively accepted that B_3 was classified into B genome at the early days. As far as phenotypic traits, B_3 grows likewise A_1 (with yellow crown, apetalous basis points and five-room capsule), other than any species of B genome (with ivory petal, large basis points of petals and three-room capsule). From the synteny relationship of 5S and 45S, B_3 could be confirmed to be in B genome. And the nonsynteny relationship of 5S and 45S could be considered as the classification basis of B genome. So the rDNA identification for all species of B genome has the cytogenetic evidences for the classification of *Gossypium* genus. Notably, the two 5S loci in E_4 is a great discovery in *Gossypium* genus, suggesting special evolutional implication in species of E genome.

Phylogenetic implication for Australia diploid species (G genome). The numbers of 45S rDNA were revealed three and four in G_3 and G_1, respectively, although the two G-genome species have similar morphological traits [47]. It is a certainty for G genome in terms of the taxonomy, which had been well studied [40,47]. Therefore, the origin might account for the difference in the rDNA numbers between two species. According to the cpDNA analysis, the chloroplast genome of G_1 was similar with that of *G. sturtianum*, a morphologically distant C-genome species, suggesting that G_1 could have a reticulate history with *G. sturtianum*.

In summary, the current study has clarified systematically the interrelationship of the rDNA distribution among 25 *Gossypium* species and one variant covering AD, A, B, D, E, F and G genomes. And the corresponding phylogenetic implications have been revealed for the evolution of *Gossypium* genus. Further study is needed to investigate the more precise rDNA patterns on meiotic pachytene chromosomes between with the development of cotton FISH techniques [48].

Acknowledgements

Author Contributions

Conceived and designed the experiments: YG KW FL. Performed the experiments: YG DC QQ CW YW. Analyzed the data: YG FL QW. Contributed reagents/materials/analysis tools: SL XZ KW. Wrote the manuscript: YG KW.

References

1. Endrizzi JE, Turcotte EL, Kohel RJ (1985) Genetics, cytogenetics, and evolution of *Gossypium*. Adv Genet 23: 271-375. doi:10.1016/S0065-2660(08)60515-5.

2. Stewart JMcD (1995) Potential for crop improvement with exotic germ plasm and genetic engineering. In: GA ConstableNW Forrester. Challenging the future: Proceedings of the World Cotton Research. Melbourne. pp. 313-327.

3. Leitch IJ, Heslop-Harrison JS (1992) Physical mapping of the 18S-5.8-26S rRNA genes in barley by *in situ* hybridization. Genome 35: 1013-1018. doi:10.1139/g92-155.

4. Linde-Laursen I, Ibsen E, Von Bothmer R, Giese H (1992) Physical localization of active and inactive rRNA gene loci in *Hordeum marinum* ssp. *gussoneanum* (4×) by *in situ* hybridization. Genome 35: 1032-1036. doi:10.1139/g92-158.

5. Raina SN, Mukai Y (1999) Detection of a variable number of 18S-5.8S-26S and 5S ribosomal DNA loci by fluorescent *in situ* hybridization in diploid and tetraploid *Arachis* species. Genome 42: 52-59. doi:10.1139/g98-092.

6. Liu ZL, Zhang D, Wang XQ, Ma XF, Wang XR (2003) Intragenomic and interspecific 5S rDNA sequence variation in five Asian pines. Am J Bot 90: 17-24. doi:10.3732/ajb.90.1.17. PubMed: 21659076.

7. Bogunić F, Siljak-Yakovlev S, Muratović E, Ballian D (2011) Different karyotype patterns among allopatric *Pinus nigra* (Pinaceae) populations revealed by molecular cytogenetics. Plant Biol 13: 194-200. doi:10.1111/j.1438-8677.2010.00326.x. PubMed: 21143741.

8. Robledo G, Lavia GI, Seijo G (2009) Species relations among wild *Arachis* species with the A genome as revealed by FISH mapping of rDNA loci and heterochromatin detection. Theor Appl Genet 118: 1295-1307. doi:10.1007/s00122-009-0981-x. PubMed: 19234686.

9. Hamon P, Yakovlev S, Robin O, Poncet V, Hamon S, de Kochko A (2009) Physical mapping of rDNA and heterochromatin in chromosomes of 16 *Coffea* species: A revised view of species differentiation. Chromosome Res 17: 291-304. doi:10.1007/s10577-009-9033-2. PubMed: 19333769.

10. Rosato M, Castro M, Rosselló J (2008) Relationships of the woody *Medicago* species (Section *Dendrotelis*) assessed by molecular

cytogenetic analyses. Ann Bot 102: 15-22. doi:10.1093/aob/mcn055. PubMed: 18413655.

11. Chung MC, Lee YI, Cheng YY, Chou YJ, Lu CF (2008) Chromosomal polymorphism of ribosomal genes in the genus *Oryza*. Theor Appl Genet 116: 745-753. doi:10.1007/s00122-007-0705-z. PubMed: 18214422.

12. Chacón J, Sousa A, Baeza CM, Renner SS (2012) Ribosomal DNA distribution and a genus-wide phylogeny reveal patterns of chromosomal evolution in *Alstroemeria* (Alstroemeriaceae). Am J Bot 99: 1501-1512. doi:10.3732/ajb.1200104. PubMed: 22956470.

13. Price HJ, Stelly DM, McKnight TD, Scheuring CF, Raska D et al. (1990) Molecular cytogenetic mapping of a nucleolar organizer region in cotton. J Hered 81: 365-370.

14. Crane CF, Price HJ, Stelly DM, Czeschin DG (1993) Identification of a homeologous chromosome pair by *in situ* DNA hybridization to ribosomal RNA loci in meiotic chromosomes of cotton (*Gossypium hirsutum*). Genome 36: 1015-1022. doi:10.1139/g93-135. PubMed: 18470044.

15. Ji Y, De Donato M, Crane CF, Raska WA, Islam-Faridi MN et al. (1999) New ribosomal RNA gene locations in *Gossypium hirsutum* mapped by meiotic FISH. Chromosoma 108: 200-207. doi:10.1007/s004120050369. PubMed: 10398849.

16. Hanson RE, Islam-Faridi MN, Percival EA, Crane CF et al. (1996) Distribution of 5S and 18S-28S rDNA loci in a tetraploid cotton (*Gossypium hirsutum* L.) and putative diploid ancestors. Chromosoma 105: 55-61. doi:10.1007/BF02510039. PubMed: 8662259.

17. Gan Y, Chen D, Liu F, Wang C, Li S et al. (2011) Individual chromosome assignment and chromosomal collinearity in *Gossypium thurberi*, *G. trilobum* and D subgenome of *G. barbadense* revealed by BAC-FISH. Genes Genet Syst 86: 165-174. doi:10.1266/ggs.86.165. PubMed: 21952206.

18. Gan Y, Liu F, Peng R, Wang C, Li S et al. (2012) Individual chromosome identification, chromosomal collinearity and genetic-physical integrated map in *Gossypium darwinii* and four D genome cotton species revealed by BAC-FISH. Genes Genet Syst 87: 233-241. PubMed: 23229310.

19. Wang K, Guan B, Guo W, Zhou B, Hu Y et al. (2008) Completely distinguishing individual A-genome chromosomes and their karyotyping analysis by multiple bacterial artificial chromosome-fluorescence *in situ* hybridization. Genetics 178: 1117-1122. doi:10.1534/genetics.107.083576. PubMed: 18287408.

20. Wu Q, Cheng H, Liu F, Wang S, Song G et al. (2010) Screen of FISH marker of chromosomes at *Gossypium* D genome species (in Chinese). Chinese Sci Bull (Chinese Ver.)55: 2099-2105.

21. Wang K, Guo W, Zhang T (2007) Development of one set of chromosome-specific microsatellite-containing BACs and their physical mapping in *Gossypium hirsutum* L. Theor Appl Genet 115: 675-682. doi:10.1007/s00122-007-0598-x. PubMed: 17632699.

22. Sambrook J, Russell DW (2002) Molecular cloning: a laboratory manual, 3rd ed. New York: Cold Spring Harbor Laboratory Press.

23. Wang C, Wang K, Song G, Li M et al. (2001) Protocol of cotton FISH of somatic chromosomes with rDNA as probes. J Cotton Sci 13: 75-77.

24. Wendel JF, Brubaker CL, Seelanan T (2010) The origin and evolution of *Gossypium*. In: JMcD StewartD OosterhuisJJ HeitholtJR Mauney. Physiology of Cotton. USA: Springer Verlag. pp. 5-10.

25. Vaio M, Speranza P, Valls JF, Guerra M, Mazzella C (2005) Localization of the 5S and 45S rDNA sites and cpDNA sequence analysis in species of the Quadrifaria group of *Paspalum* (Poaceae, Paniceae). Ann Bot 96: 191-200. doi:10.1093/aob/mci168. PubMed: 15911540.

26. Martins C, Galetti PMJ (2001) Two 5S rDNA arrays in Neotropical fish species: is it a general rule for fishes? Genetica 111: 439-446. doi:10.1023/A:1013799516717. PubMed: 11841188.

27. Schweizer D, Loidl J (1987) A model for heterochromatin dispersion and the evolution of C-band patterns. In: A StahlJM LucinaniAM Vagner-Capodano. Chromosomes Today vol. 9. London: Allen and Unwin. pp 61-74.

28. Moscone EA, Klein F, Lambrou M, Fuchs J, Schweizer D (1999) Quantitative karyotyping and dual-color FISH mapping of 5S and 18S-26S rDNA probes in the cultivated *Phaseolus* species (Leguminosae). Genome 42: 1224-1233. doi:10.1139/gen-42-6-1224. PubMed: 10659791.

29. Wang K, Wang Z, Li F, Ye W, Wang J et al. (2012) The draft genome of a diploid cotton *Gossypium raimondii*. Nat Genet 44: 1098-1103. doi:10.1038/ng.2371. PubMed: 22922876.

30. Wendel JF, Brubaker C, Alvarez I, Cronn R, Stewart JM (2009) Evolution and natural history of the cotton genus. In: AH Paterson. Genetics and Genomics of Cotton. USA: Springer Verlag. pp. 3-22.

31. Phillips LL (1966) The cytology and phylogenetics of the diploid species of *Gossypium*. Am J Bot 53: 328-335. doi:10.2307/2439872.

32. Johnson BL, Thein MM (1970) Assessment of evolutionary affinities in *Gossypium* by protein electrophoresis. Am J Bot 57: 1081-1092. doi:10.2307/2441274.

33. Valiček P (1978) Wild and cultivated cottons. Cot Fib Trop 33: 363-387.

34. Krishnan P, Sapra VT, Siman KM, Zipf A (2001) FISH mapping of the 5S and 18S-28S rDNA loci in different species of *Glycine*. J Hered 92: 295-300. doi:10.1093/jhered/92.3.295. PubMed: 11447252.

35. Comai L (2000) Genetic and epigenetic interactions in allopolyploid plants. Plant Mol Biol 43: 387-399. doi:10.1023/A:1026738225755. PubMed: 10999418.

36. Lavania UC, Basu S, Srivastava S, Mukai Y, Lavania S (2005) In situ chromosomal localization of rDNA sites in 'safed musli' *Chlorophytum* Ker-Gawl and their physical measurement by fiber FISH. J Hered 96: 155-160. doi:10.1093/jhered/esi018. PubMed: 15618304.

37. Wendel JF, Cronn RC (2003) Polyploidy and the evolutionary history of cotton. Adv Agron 78: 139-186. doi:10.1016/S0065-2113(02)78004-8.

38. Cronn RC, Small RL, Haselkorn T, Wendel JF (2003) Cryptic repeated genomic recombination during speciation in *Gossypium gossypioides*. Evolution 89: 707-725. PubMed: 14686525.

39. Cronn RC, Small RL, Haselkorn T, Wendel JF (2002) Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. Am J Bot 89: 707-725. doi:10.3732/ajb.89.4.707. PubMed: 21665671.

40. Wendel JF, Albert VA (1992) Phylogenetics of the cotton genus (*Gossypium* L.): Character-state weighted parsimony analysis of chloroplast DNA restriction site data and its systematic and biogeographic implications. Syst Bot 17: 115-143. doi:10.2307/2419069.

41. Fryxell PA, Koch SD (1987) New or noteworthy species of flowering plants from the Sierra Madre del Sur of Guerrero and Michoacan, Mexico. Aliso 11: 539-561.

42. Fryxell PA (1979) The Natural History of the Cotton Tribe. Texas A & M University Press.

43. Chacón J, Sousa A, Baeza CM, Renner SS (2012) Ribosomal DNA distribution and a genus-wide phylogeny reveal patterns of chromosomal evolution in *Alstroemeria* (Alstroemeriaceae). Am J Bot 99: 1501-1512. doi:10.3732/ajb.1200104. PubMed: 22956470.

44. Martínez J, Vargas P, Luceño M, Cuadrado A (2010) Evolution of *Iris* subgenus *Xiphium* based on chromosome numbers, FISH of nrDNA (5S, 45S) and *trnL-trnF* sequence analysis. Plant Syst Evol 289: 223-235. doi:10.1007/s00606-010-0345-7.

45. Chang YC, Shii CT, Chung MC (2009) Variations in ribosomal RNA gene loci in spider lily (*Lycoris* spp.). J Am Soc Hortic Sci 134: 567-573.

46. Dobigny G, Ozouf-Costaz C, Bonillo C, Volobouev V (2003) Evolution of rRNA gene clusters and telomeric repeats during explosive genome repatterning in *Taterillus* X (Rodentia, Gerbillinae). Cytogenet Genome Res 103: 94-103. doi:10.1159/000076296. PubMed: 15004471.

47. Wendel JF, Stewart JMcD, Rettig JH (1991) Molecular evidence for homoploid reticulate evolution in Australian species of *Gossypium*. Evolution 45: 694-711. doi:10.2307/2409921.

48. Peng RH, Zhang T, Liu F, Ling J, Wang CY et al. (2012) Preparations of meiotic pachytene chromosomes and extended DNA fibers from cotton suitable for fluorescence in situ hybridization. PLOS ONE 7: 33847. doi:10.1371/journal.pone.0033847. PubMed: 22442728.