Euphytica (2021)217:154
https://doi.org/10.1007/s10681-021-02771-6

# Alteration of root and shoot morphologies by interspecific replacement of individual Upland cotton chromosome or chromosome segment pairs

**Mirzakamol S. Ayubov · Tokhir M. Norov · Sukumar Saha · Te-Ming Tseng · K. Raja Reddy · Johnie N. Jenkins · Ibrokhim Y. Abdurakhmonov · David M. Stelly**

Received: 18 October 2019 / Accepted: 13 January 2021


**Abstract** Genetic approaches often lead to the most cost-effective and efficient means to improve crops, especially those grown widely. But for most crops, cotton included, genetic improvement efforts have focused far more on above-ground plant attributes than on root systems. Root system establishment is crucial to cotton seedling success, subsequent development, crop performance and sustainability. As a first step toward genetic enhancement of cotton root systems, significant heritable phenotypic variation must be found or created. The overall objective of this research was to study the effect of substituted chromosomes or chromosome segments from the donor tetraploid species *Gossypium barbadense*, *G. mustelinum*, and *G. tomentosum* on the selected traits of the stem, leaf, and especially root in CS lines. Twenty-seven chromosome substitution (CS) lines, containing different pairs or short segments of chromosomes from *G. barbadense* (CS-B lines), *G. mustelinum* (CS-M lines), and *G. tomentosum* (CS-T lines) and two parents, TM-1, parent quasi-isogenic to the CS lines and *G. barbadense* 3-79, the donor parent to all CS-B lines, were analyzed. Goals were to determine if CS lines significantly affect any of 17 morphological shoot and root traits. Indeed, significant line-based variation occurred for several root and shoot phenotypes. Comparisons of means and two-way hierarchical cluster analysis revealed several CS lines simultaneously affected multiple shoot and/or root traits, positively or negatively. Pairwise correlations of traits and the cluster analysis showed strong relationships among certain traits. The high correlation among several root traits suggests that easier-to-screen traits might be leveraged strategically to devise breeding-friendly methods for phenotypically evaluating root system morphology. Most importantly, this research identifies CS lines with prospectively novel individual trait effects and others with multi-trait effects that can be further dissected and used to improve our knowledge of cotton root systems, their development, genetic control and genetic improvement.

**Keywords** Chromosome substitution lines · *Gossypium* species · Breeding · Interspecific · WinRHIZO software

M. S. Ayubov · T. M. Norov · I. Y. Abdurakhmonov
Center of Genomics and Bioinformatics, Academy Sciences of Uzbekistan, Tashkent, Uzbekistan 111215

S. Saha (✉) · J. N. Jenkins
Genetics and Sustainable Agriculture Unit, USDA-ARS, Mississippi State, MS 39762, USA
e-mail: Sukumar.Saha@usda.gov

T.-M. Tseng · K. R. Reddy
Department of Plant and Soil Sciences, Mississippi State University, Mississippi State, MS, USA

D. M. Stelly (✉)
Department of Soil and Crop Sciences, Texas A&M University, College Station 77843-2474, TX, USA
e-mail: stelly@tamu.edu

Published online: 01 July 2021

## Introduction

Plant morphological traits such as plant height, leaf number, leaf size, root length, root size, root diameter, and other morphological features are crucial to crop performance under normal as well as biotic or abiotic stress conditions. Although roots and root systems are highly influential in overall plant composition, health and productivity, their reduced accessibility renders them far less amenable to scientific inquiry and characterization. Past improvement of crops and agricultural techniques have mainly focused on increasing shoot biomass and seed yield, commonly overlooking relevance of the root system to production (Den Herder et al. 2010). Additional information about variation in roots and root systems is central to advancing their scientific characterization, genetic dissection and subsequent genetic improvement. Prior research has shown that phenotypic differences in morphological traits result from allelic variation of the genes at multiple loci in different chromosomes (Ristova and Busch 2014). Statistically, their effects are often subject to interactions with other genes and "the environment". The discovery or creation of new genetic variation remains an essential ingredient for genetic improvement of crop plants.

The practical benefits to be gained by genetically improving root systems of Upland cotton are prospectively great, given that global cultivation of Upland cotton is extensive, e.g., Australia, Brazil, Burkina-Faso, China, India, Pakistan, Turkey, Turkmenistan, Uzbekistan, and the USA, to name a few, and often in agriculturally challenging environments. Moreover, such discoveries are demanded by increasing emphasis in cotton breeding programs and generally on enhancing sustainability and adaptation to ongoing and predicted global climate changes. While genetic and morphological variants certainly exist among cultivated Upland cotton, racestocks and other accessions of *G. hirsutum*, extant cultivars descend from a passage involving multiple genetically constrictive "bottlenecks" during evolution, domestication and/or adaptation to various production areas, most far removed from centers of origin and diversity. The opportunities for discovering beneficial genetic variation for cotton roots and root systems seem highest in wild accessions and closely related species. The discovery of additional diversity for root systems could catalyze new research efforts and bring about novel genetic improvements.

In their discussion of "Gene Pools", Harlan and de Wet (1971) characterized genetic diversity as a "two-edged sword"—with increased diversity comes increased difficulty of use in breeding. Whereas the search for simply inherited traits such as resistance to some pathogen races, or new morphological or reproductive traits can sometimes be fairly easily extended from elite germplasm to non-cultivated types and even closely related species in the primary gene pool, to search directly among non-domesticated germplasm for potentially useful variation of highly multigenic traits is typically difficult. The challenge is greater when the multigenic trait is one for which the cultivated forms are highly improved relative to non-cultivated types, because the effect of variation at an individually beneficial locus is likely to be obscured by effects of other genes, and by the multigenic "noise" of a collectively heterogenous background. An experimentally executable means of reducing such "noise" indirectly is that of hybridization and back-cross-mediated introgression of germplasm from potential donors into the elite homozygous background of one or more cultivated genotypes. Likewise, genetic variants at individual loci of a genetic network might exist quite cryptically amongst non-cultivated types, but exert significant interaction effects once introgressed or transferred into a new genetic network. Here, too, backcross-inbred extraction can serve as an incisive tool, where the donor genome is dissected as it is backcrossed into at least one common background, such that the potential effects of smaller pieces of the donor genome, along with their interactions can be evaluated in a "low-noise" genetic environment. Backcross introgression lines (Wehrhahn and Allard 1965), chromosome addition lines (O'mara 1940), chromosome substitution (CS) lines (Sears 1952) and chromosome segment substitution lines (Eshed and Zamir 1995) enable this sort of approach to be implemented. In this study, we report the use interspecific CS lines of Upland cotton to search for significant morphological effects on the root system in Upland cotton. Such lines can be especially useful for detecting the introduction of genetic factors that affect multigenic traits in the genetic milieu of a domesticated type, i.e., allowing for direct and/or epistatic effects, while also minimizing line-to-line noise from background genetic variation.

Cytogenetic analyses revealed long ago that the cultivated cotton species *G. hirsutum* and *G. barbadense* and other New World 52-chromosome species are tetraploid 2n = 4x = 52 and form 26 meiotic bivalents, whereas the other species are diploid 2n = 2x = 26 and form 13 meiotic bivalents (Skovsted 1934; Beasley 1940). All 52-chromosome *Gossypium* species are disomic tetraploids ($2n = 4x = 52$) with meiotically independent A- and D-subgenomes. The chromosomes of A subgenome and D subgenome of *G. hirsutum* were designated as chromosomes 1–13 (A subgenome), and 14–26 (D subgenome), respectively (Endrizzi et al. 1984, 1985). More recently, molecular data indicated the origin of all extant AD-genome allotetraploid cotton species to be a monophyletic from a hybridization/polyploidization event during or after hybridization of two diploid ancestors about 1–2 million years ago, one ancestor having a genome similar to extant A-genome diploid species, and the other with a genome similar to extant D-genome diploid species (Senchina et al. 2003; Wendel and Cronn 2003; Udall and Wendel 2006). The polyploidization event likely constituted a genetic bottleneck itself, but was surely followed by diversification. However, major losses of genetic variation occurred much more recently in conjunction with domestication of *G. hirsutum* (Iqbal et al. 2001), and also upon adaptation to new agricultural production areas, and yet again with modern breeding, especially the extensive reliance of crosses among closely-related elite domesticated genotypes and reselection within existing elite cultivars for high yield and superior fiber quality (Van Esbroeck et al. 1999).

Genetic improvement is usually the most cost-effective and efficient way to improve crop species. Plant morphological traits such as plant height, leaf number, leaf size, root length, root size, root diameter, and other morphological features play important roles in its performance under normal as well as biotic or abiotic stress conditions. Phenotypic differences of quantitatively inherited morphological traits is caused by allelic variation of the genes at multiple loci and chromosomes (Ristova and Busch 2014). Most of the investigations on crop morphological traits have primarily targeted the same traits that were improved during the domestication process (Ristova and Busch 2014). Deciphering the genetic mechanism associated with the complex morphological traits can benefit from methods that reduce genetic complexity and minimize background genetic "noise".

Descriptions of above-ground plant parts and their genetic characterizations are readily found for *G. hirsutum* and Upland cotton varieties, but very limited genetic information is available on specific cotton morphological traits, especially root traits (McMichael and Quisenberry 1991; Awasthi et al. 2018; Reddy et al. 2020; Singh et al. 2018). This is due in part to root traits being difficult to measure, especially in natural conditions under the soil. Growth and development of the root system reflect constant efforts by the plant to optimize its distribution in the soil, a varying and heterogeneous growth medium. Root architecture refers to the spatial configuration on the arrangement of root axes under specific soil environment (Lynch 1995). Root architecture development is vital to efficient acquisition of soil nitrogen and water, seedling establishment and survival of the plant under adverse soil conditions (Lynch 2005).

The tetraploid cotton species are very diverse in their morphological phenotypes and habitats. For example, *G. tomentosum*, the species endemic to Hawaii, is very distinct from the other tetraploid species in morphology. It has hairy, silvery-green to gray-green, palmately veined leaves, yellow corollas without petal spots; stigmas are strongly exerted, and the flowers are devoid of extra floral nectaries and was the source of the nectariless trait deployed for insect resistance. It forms 3-celled capsules containing 6–12 seeds covered with reddish brown short fibers lacking any differentiation in two layers (Meyer and Meyer 1961; Meyer and Meredith 1978; DeJoode and Wendel 1992; Percival et al. 1999). *G. mustelinum*, a wild tetraploid species native to the semi-arid regions of northeastern Brazil, is phenotypically distinct in its fruit and seed characteristics from *G. hirsutum* (Pickersgill et al. 1975; Saha et al. 2013a, b). Breeders have attempted over many years to introduce morphological characters associated with improved productivity from other species in Upland cotton, but have realized only very limited success. Interspecific introgression into Upland cotton by conventional methods typically suffers from hybrid breakdown upon inbreeding and from serious linkage drag effects that largely preclude recovery of elite Upland types containing donor-derived genes or traits. This is especially important considering that the genetic improvement of Upland cotton is constrained by its narrow genetic base.

To increase diversity, we have developed and characterized a number of traits of interspecific CS lines of Upland cotton. Each line introduces a limited amount of germplasm of donor species *G. tomentosum* (CS-T), *G. mustelinum* (CS-M) and *G. barbadense* (CS-B) into a quasi-isogenic background of Upland cotton (*G. hirsutum*) (Saha et al. 2017; Jenkins et al. 2017a, b). Each CS line is considered mostly identical to the recurrent parent, TM-1 (*G. hirsutum*) for 25 chromosome pairs, and to most other CS lines for 24 chromosome pairs, i.e., except the substituted chromosomes or chromosome segments. Isogenicity of CS lines to each other and TM-1 enhances their analytical value for phenotypic comparisons and discerning effects of substituted chromosomes or chromosome segments. The CS lines thus provide an opportunity to discover novel traits associated with important traits including root, stem, and leaf and identify their chromosomal locations (Karaca et al. 2002; Saha et al. 2015, 2017; Song et al. 2017; Jenkins et al. 2017a, b). Recently, we applied a morphometric image analyses system to genetically dissect the complex morphological traits associated with the CS lines (Awasthi et al. 2018; Reddy et al. 2020). CS-T04 and CSB08sh exhibited higher and lower low-temperature tolerances, respectively, and CS-T04 and CS-B22sh showed higher and lower drought tolerance, respectively (Awasthi et al. 2018).

The overall objective of this research was to morphologically characterize the selected traits of stem and root in CS lines comparing with the TM-1 to study the effect of the substituted chromosome or chromosome segment from the alien species on 25-day old seedlings following the methods of Awasthi et al. (2018), Singh et al. (2018). Comparative analysis of the CS lines with their almost isogenic recurrent parent the inbred 'Texas Marker-1′ (TM-1, *G. hirsutum*) has provided a method to identify and associate important traits with specific substituted chromosome or chromosome segments from the alien species (Saha et al. 2004, 2006, 2013a,b, 2018).

## Materials and methods

The evaluated germplasm included 29 interspecific CS lines of Upland cotton ($2n = 52$), each bred to be disomic for a different alien chromosome pair or chromosome short arm (sh) from *G. barbadense* (CS-B lines), *G. mustelinum* (CS-M lines), or *G. tomentosum* (CS-T lines) (Saha et al. 2004; Stelly et al. 2005), as well as two reference genotypes, the Upland inbred (*G. hirsutum*) Texas Marker (TM)-1 and the non-photoperiodic *G. barbadense* parent doubled haploid-derived line "3–79", which was the donor parent of CS-B lines (Table 2). Most of these CS lines were bred according to modified backcross-inbred development to the BC5S1, using cytological not genomic marker methods, thus other than for backcrossing, no control was exercised over inadvertent retention of unlinked segments of donor germplasm (Stelly et al. 2005). Lines CS-B02, CS-M02, CS-T02, CS-B04, CS-M04, CS-T04, CS-B06, CS-M06, CS-T06, CS-B17, CS-B18, CS-M17, CS-M18, CS-T17, CS-T18 and CS-B08sh, CS-M08sh, CS-T08sh, CS-B11sh, CS-M11sh, CS-T11sh, CS-B15sh, CS-M15sh, CS-T15sh, CS-B22sh, CS-M22sh, CS-T22sh share a common genetic background, i.e., that of the TM-1 inbred. TM-1 served as the recurrent pollen parent to create each of the isogenic hypoaneuploid *G. hirsutum* parents that was used subsequently as recurrent female parent in the modified backcrossing to create the monosomic substitution, prior to inbreeding and establishing each CS line (Stelly et al. 2005). TM-1 is a genetic and cytogenetic standard of Upland cotton (*G. hirsutum*), and line 3–79 is a *G. barbadense* genetic standard.

The overall method of measuring different traits in our experiment with the CS line was followed as per Awasthi et al. (2018). Seed from the CS lines, TM-1 and 3–79 were planted in PVC pots (15.2 cm diameter and 30.5 cm height) filled with the soil medium consisting of 3:1 sand: topsoil classified as sandy loam (87% sand, 2% clay, and 11% silt) with a 500 g of gravel at the bottom of each pot at a temperature around 95F and watered three times with automated pipeline using standard greenhouse procedure. Initially, four seeds were sown in each pot and four days after emergence; the plants were thinned to one pot$^{-1}$. Pots were arranged as a randomized complete block design with four replicates.

Each genotype was evaluated for 17 traits associated with leaf, shoot, and roots of the individual plant in a CS line. Plant heights were measured at the final harvest, 25 DAP. The leaf area of the second leaf was measured using the LI-3100 leaf area meter (LI-COR, Biosciences). Plant total dry weights (TD), including

leaves, stems, and roots were recorded after oven drying for five days at 80 °C.

The plants were carefully excavated 25 days after sowing. After harvesting, roots were separated from the base of the shoot by cutting individual plants. Roots were washed gingerly by placing on sieves and spraying gently with water and also by dipping them in clean water to remove any mud. An additional sieve of 0.2 mm was placed at the outflow of the system to make sure that no fine root material was lost. Special care was taken to avoid any damage to finer root morphology, especially finer and secondary root. The cleaned individual root systems were floated in about 5 mm of water in a 0.3- by 0.2-m Plexiglas tray and gray-scale root images were acquired as per the overall method of Singh et al. (2018). Briefly, roots were untangled and separated with a plastic paintbrush to minimize root overlap. The tray with the root was placed on a specialized dual-scan optical scanner set to the high accuracy of resolution 800 by 800 dpi equipped with a commercial software package 4.1 Win RHIZO (Regent Instruments, 2000, USA, Arsenault et al. 1995). Root images were then analyzed with a computer linked to WinRHIZO software analysis system for recording data on root morphology. The root parameters provided by the system includes root length (RL), surface area (RSA), average root diameter (RAD), number of root forks (RNF), number of root crossings (RNC), root volume (RV) and number of tips (RNT) using WinRHIZO Pro software (Singh et al. 2018; Arsenault et al. 1995). In addition, root/shoot ratio (R/S), root length/ g. wt. (RL/GW), and root surface area/ g. wt. (RSA/GW) were determined based upon measured root parameters.

### Statistical analysis

Data on 17 traits were analyzed for means and factorial analysis of variance (ANOVA) using SAS statistical package (SAS 2007) and JMP software (Statistical Institute, Inc., Cary, NC). These included 9 root-only traits—dry weight, longest root length, total root length, diameter, surface area, volume and counts of tips, forks and crossings, 6 above-ground part plant traits, and 2 combined root-shoot traits (Table 2). All data were imported from an Excel spreadsheet into SAS JMP Genomics format 3.2 (SAS). Mean comparisons among the CS lines and parents were analyzed by Tukey's $t$ test at $P \geq 0.05$ using the JMP program. Pearson's correlation coefficients were calculated between pairs of traits and tested at $P \geq 0.05$ level. For traits identified by ANOVA as significantly affected by genotypes at $p \geq 0.05$, hierarchical two-way clustering of traits and genotypes was computed and graphed using the JMP program (Statistical Institute, Inc., Cary, NC).

## Results

The 25-day old cotton CS line seedlings were used to analyze 17 phenotypes that could broadly be categorized into 3 stem, 2 leaf, 10 root, and 2 combined root and stem phenotypes. The analysis was based on the assumption that all CS lines are largely isogenic to each other and TM-1, the recurrent backcross parent of the Upland hypoaneuploid parents used subsequently for backcross-based development of the CS lines. The CS lines and TM-1 were grown in the same environment in replicated plots, so significant differences in any phenotype among the lines were considered due to the substituted chromosomes or chromosome segment from the alien species (Fig. 1). However, it should be mentioned that there is a small possibility of some differences in traits might be due to the inadvertent retention of off-target donor genes from the donor parent during backcrossing and subsequent inbreeding of CS lines; differences could also be due to interactions between donor gene(s) of the substituted chromosome with the genes on other chromosomes of the CS line. Overall, cotton lines substituted for nine different individual chromosome or chromosome segment from three donor species of *G. barbadense* (CS-B), *G. tomentosum* (CS-T), and *G mustelinum* (CS-M) respectively were subjected for genetic analysis. A total of 27 CS lines were also included to study the association of the substituted chromosome with 17 different plant morphological traits of stem, leaf, and root.

### ANOVA analysis

The CS lines exhibited substantial phenotypic variation and exemplified by representative images of each line (Fig. 2, Table 1). Graphic depiction of means and standard errors (Fig. 3) indicated fairly uniform plant-to-plant variation across most CS lines and traits. Standard errors for multiple traits of CS-B15sh, CS-

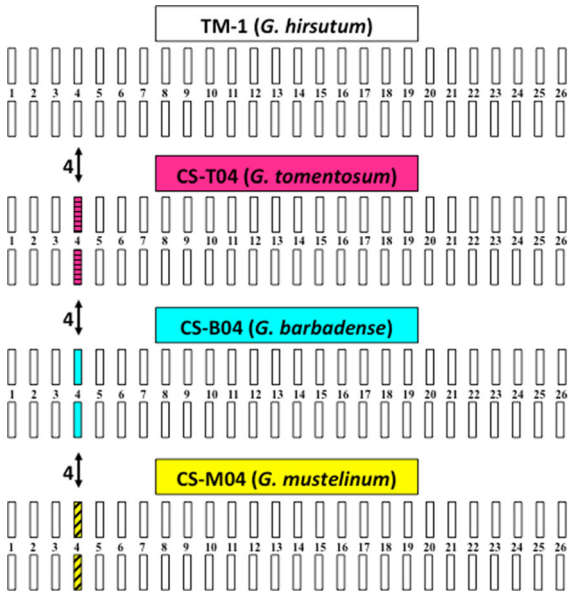


**Fig. 1** Diagrammatic illustration of isogenic, genomic and chromosomal relationships among three idealized chromosome substitution (CS) lines and the Upland cotton inbred TM-1. In each of the three interspecific CS lines, CS-T04, CS-B04 and CS-M04, the chromosome-04 pair of Upland cotton has been replaced with the homologous chromosome pair from a different donor species

M17 and perhaps CS-T11sh seemed a bit larger, but they were retained in the analyses, in spite of their potential to reduce sensitivity of comparisons. Even so, the analysis of variance (ANOVA) nonetheless indicated that differences among lines were a significant ($p < 0.05$) source of phenotypic variation for 12 out of the 17 traits, including all 9 root-only traits (Table 2).

## Average values of traits

CS lines are nearly isogenic to TM-1 (*G. hirsutum*) for all of the chromosomes except the substituted chromosome or chromosome arm segment from the alien species of *G. barbadense*, *G. tomentosum* and *G. mustelinum* (Fig. 1). Several of the morphological traits showed significant variation in the phenotypes among the CS lines, suggesting that substantial genetic variability in those traits was associated with the substituted chromosome or chromosome segment of the CS lines (Figs. 2, 3, Table 2). Several traits also differed markedly between inbreds TM-1 and 3-79, revealing differences between the CS-isogenic inbred and the species donor to CS-B lines, respectively. For example, Pima 3–79 had the highest individual root length compared to TM-1 and all other CS lines (Table 2). This suggests that considerable variation was expected among the phenotypes of the CS lines in most of the traits. For example, the number of root tips per root system was significantly higher in 3–79 than TM-1, but several CS lines had even higher numbers of root tips.

The following ranges across different traits among the 25-day-old plant CS lines and parents were observed: plant height, 11.38–23.38 cm; leaf number/plant, 3.50–7.25; stem dry weight, 0.48–1.18 g; root dry weight, 0.22–0.40 g; above-ground part weight, 1.33–3.08 g; total dry weight, 1.52–3.20 g; longest root length in a plant, 17.95–35.95 cm; total root length/plant, 1364.64–2966.24 cm; total root surface area/plant, 212.85–406.76 sq cm; root diameter, 0.40–0.55 cm; root volume, 2.82–4.63 c.c.; total number of root tips/plant, 1396.25–3086.50; number of root forks/plant, 6831.80–22,025; and, number of root crossings/plant, 454.50–1411.50 (Table 2). CS-B17 had the highest stem dry weight, root dry weight, and above-ground weight. On the other hand, CS-T06 had the highest number of leaves per plant, and CS-M15sh had highest total root length among the CS lines. CS-B11sh showed the highest root surface area.

## Correlations among traits

Correlation analysis revealed overall positive associations among almost all traits, except root/shoot ratio and root diameter, which were negatively correlated with a few traits (Table 3). The root/shoot ratio was exceptional in not being significantly correlated with any other trait except stem dry weight (− 0.28). Leaf number, leaf area, and leaf dry weight were positively correlated with all traits, except root and shoot ratio, and root diameter. Leaf number/plant, leaf dry weight, root dry weight, total dry weight, longest root length, root surface area, and root volume were positively correlated with several other root and shoot traits (Table 3). As can be seen on the left side of Table 3, most of the correlations among the various size and weight traits tended to be large ($r \geq 0.9$) and significant ($p \leq 0.05$), including plant height, leaf number, leaf area, and dry weights of leaf, shoot, root and overall plant. Plant height was highly correlated with total root dry weight ($\leq 90\%$), total root length ($\leq 80\%$), root surface area ($\leq 75\%$), root volume

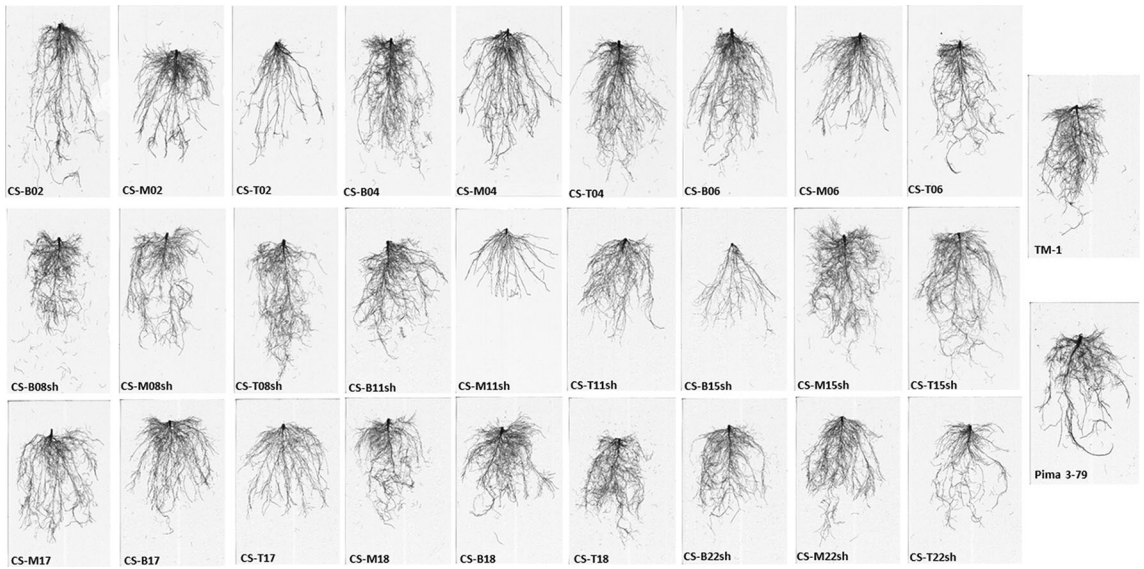


**Fig. 2** Images of representative roots from the chromosome substitution lines and two parental lines, TM-1 and Pima 3-79, taken 25 days after sowing

**Table 1** Analysis of variance across chromosome substitution lines for various morphological traits measured 25 days after sowing

| Trait | Genotype |
|---|---|
| Plant height | ** |
| Main stem leaf number | *** |
| Total leaf area | ns |
| Leaf dry weight | ns |
| Stem dry weight | ** |
| Root dry weight | * |
| Above-ground weight | ns |
| Total weight | ns |
| Root weight/Shoot weight | ns |
| Longest root length | ** |
| Total root length | * |
| Root Surface area | * |
| Root average diameter | *** |
| Root volume | ** |
| Total number of root tips | *** |
| Total number of root forks | ** |
| Total number of root crossings | *** |

#The significance levels ***, **, * and ns represent $p \leq .001$, $p \leq .01$, $p \leq .05$, and not significant (ns), respectively

($\leq 64\%$), number of root tips ($\leq 83\%$), and number of root crossings ($\leq 81\%$). Leaf number showed a high correlation value with root dry weight ($\leq 88\%$), longest root length ($\leq 59\%$), and total root length ($\leq 74\%$).

**Fig. 3** Average values of above-ground (**A**) and below-ground (**B**) morphological traits associated with different chromosome substitution lines and TM-1 ▶

## Two-way cluster analysis

Phenotype-based clustering of lines and line-based differential clustering of traits were simultaneously derived by two-way hierarchical cluster analysis. Relationships among the CS lines were plotted along one axis, while the relationships among selected traits were plotted along the other axis (Fig. 4). The centrally placed heat map shows the patterns of inter-relationships between lines and traits that underpin the clustering patterns displayed at the respective axes (Fig. 4). The results expand upon the correlation analysis discussed above (Table 3). However, only the traits that showed a significant effect by genotype based on ANOVA analysis ($p < 0.05$) were used in this analysis and interpreted. The color-encoded heat map of line-trait values provided a facile visual means to perceive line-trait patterns that indicate their

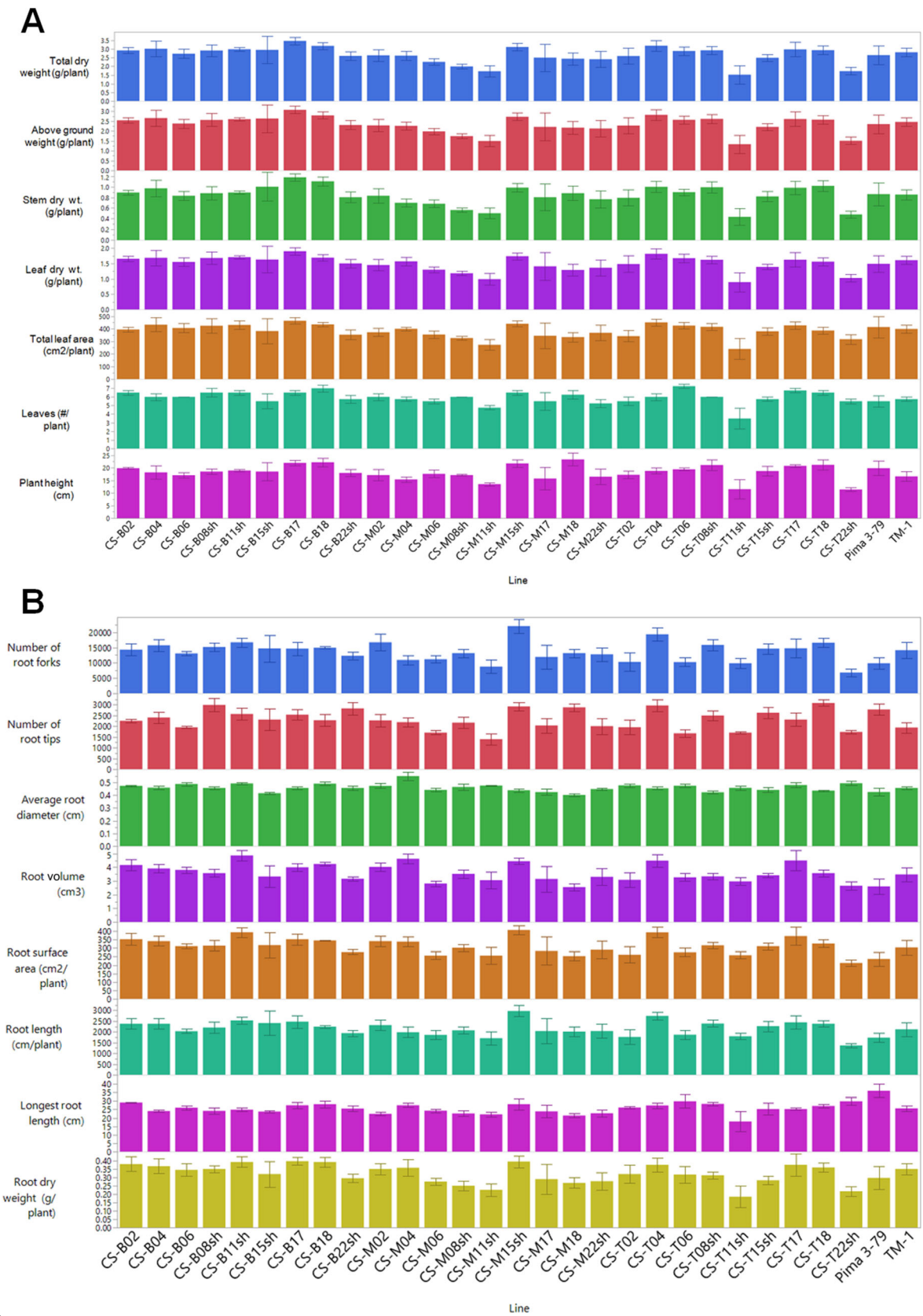
A
Total dry weight (g/plant)
Above ground weight (g/plant)
Stem dry wt. (g/plant)
Leaf dry wt. (g/plant)
Total leaf area (cm2/plant)
Leaves (#/plant)
Plant height (cm)
CS-B02
CS-B04
CS-B06
CS-B08sh
CS-B11sh
CS-B15sh
CS-B17
CS-B18
CS-B22sh
CS-M02
CS-M04
CS-M06
CS-M08sh
CS-M11sh
CS-M15sh
CS-M17
CS-M18
CS-M22sh
CS-T02
CS-T04
CS-T06
CS-T08sh
CS-T11sh
CS-T15sh
CS-T17
CS-T18
CS-T22sh
Pima 3-79
TM-1
Line
B
Number of root forks
Number of root tips
Average root diameter (cm)
Root volume (cm3)
Root surface area (cm2/plant)
Root length (cm/plant)
Longest root length (cm)
Root dry weight (g/plant)
CS-B02
CS-B04
CS-B06
CS-B08sh
CS-B11sh
CS-B15sh
CS-B17
CS-B18
CS-B22sh
CS-M02
CS-M04
CS-M06
CS-M08sh
CS-M11sh
CS-M15sh
CS-M17
CS-M18
CS-M22sh
CS-T02
CS-T04
CS-T06
CS-T08sh
CS-T11sh
CS-T15sh
CS-T17
CS-T18
CS-T22sh
Pima 3-79
TM-1
Line

**Table 2** Tukey's *t* tests for comparisons of means for various measured morphological traits of the chromosome substitution (CS) lines and the parents measured 25 days after sowing

| Name of (CS)-lines | Plant height (cm) | Leaves (number/plant) | Leaf area ($cm^2$/plant) | Leaf dry wt. (g/plant) | Stem dry dt. (g/plant) | Root dry weight (g/plant) | Above ground weight (g/plant) | Total dry weight (g/plant) | Root/shoot ratio |
|---|---|---|---|---|---|---|---|---|---|
| CS-B02 | $19.88^{a-e}$ | $6.50^{a-d}$ | $394.54^{a-d}$ | $1.66^{a-c}$ | $0.89^{a-c}$ | $0.38^{a, b}$ | $2.55^{a-c}$ | $2.93^{a-c}$ | $0.15^{a-d}$ |
| CS-B04 | $18.20^{a-f}$ | $6.00^{a-e}$ | $433.75^{a-c}$ | $1.69^{a-c}$ | $0.97^{a-c}$ | $0.37^{a-c}$ | $2.66^{a-c}$ | $3.03^{a-c}$ | $0.14^{a-e}$ |
| CS-B06 | $17.08^{b-h}$ | $6.00^{a-e}$ | $406.43^{a-c}$ | $1.55^{a-e}$ | $0.83^{a-f}$ | $0.35^{a-d}$ | $2.39^{a-d}$ | $2.73^{a-d}$ | $0.14^{a-e}$ |
| CS-B08sh | $18.50^{a-f}$ | $6.50^{a-d}$ | $424.64^{a-c}$ | $1.69^{a-c}$ | $0.88^{a-c}$ | $0.35^{a-d}$ | $2.57^{a-c}$ | $2.92^{a-c}$ | $0.14^{a-e}$ |
| CS-B11sh | $18.95^{a-f}$ | $6.50^{a-d}$ | $431.27^{a-c}$ | $1.71^{a-c}$ | $0.89^{a-c}$ | $0.39^{a, b}$ | $2.61^{a-c}$ | $3.00^{a-c}$ | $0.15^{a-d}$ |
| CS-B15sh | $18.50^{a-f}$ | $5.50^{c-e}$ | $382.01^{a-d}$ | $1.64^{a-c}$ | $1.01^{a-c}$ | $0.32^{a-e}$ | $2.64^{a-c}$ | $2.96^{a-c}$ | $0.13^{a-e}$ |
| CS-B17 | $22.00^{a-c}$ | $6.50^{a-d}$ | $463.95^{a}$ | $1.90^{a}$ | $1.18^{a}$ | $0.40^{a}$ | $3.08^{a}$ | $3.48^{a}$ | $0.13^{c-e}$ |
| CS-B18 | $22.23^{a.b}$ | $7.00^{a, b}$ | $435.47^{a-c}$ | $1.70^{a-c}$ | $1.11^{a, b}$ | $0.39^{a, b}$ | $2.80^{a, b}$ | $3.19^{a, b}$ | $0.14^{a-e}$ |
| CS-B22sh | $18.00^{a-f}$ | $5.75^{b-e}$ | $353.37^{a-e}$ | $1.51^{a-e}$ | $0.81^{b-f}$ | $0.30^{a-f}$ | $2.31^{a-d}$ | $2.61^{a-d}$ | $0.13^{c-e}$ |
| CS-M02 | $17.15^{b-g}$ | $6.00^{a-e}$ | $372.77^{a-e}$ | $1.46^{a-f}$ | $0.83^{a-f}$ | $0.35^{a-d}$ | $2.29^{a-d}$ | $2.64^{a-d}$ | $0.16^{a}$ |
| CS-M04 | $15.38^{e-h}$ | $5.75^{b-e}$ | $400.41^{a-d}$ | $1.57^{a-d}$ | $0.70^{c-g}$ | $0.36^{a-c}$ | $2.27^{a-d}$ | $2.63^{a-d}$ | $0.16^{a, b}$ |
| CS-M06 | $17.63^{b-f}$ | $5.50^{c-e}$ | $354.32^{a-e}$ | $1.30^{b-f}$ | $0.68^{c-g}$ | $0.28^{a-f}$ | $1.98^{b-e}$ | $2.26^{b-e}$ | $0.14^{a-e}$ |
| CS-M08sh | $17.13^{b-g}$ | $6.00^{a-e}$ | $326.69^{b-e}$ | $1.18^{c-f}$ | $0.56^{d-g}$ | $0.25^{c-f}$ | $1.75^{c-e}$ | $2.00^{c-e}$ | $0.14^{a-e}$ |
| CS-M11sh | $13.50^{f-h}$ | $4.75^{e.f}$ | $272.03^{d, e}$ | $0.99^{e, f}$ | $0.50^{e-g}$ | $0.23^{d-f}$ | $1.50^{d, e}$ | $1.72^{d, e}$ | $0.15^{a-c}$ |
| CS-M15sh | $21.75^{a-c}$ | $6.50^{a-d}$ | $440.95^{a-c}$ | $1.74^{a-c}$ | $0.99^{a-c}$ | $0.40^{a, b}$ | $2.73^{a, b}$ | $3.13^{a, b}$ | $0.14^{a-e}$ |
| CS-M17 | $15.75^{d-h}$ | $5.50^{c-e}$ | $345.24^{a-e}$ | $1.41^{a-f}$ | $0.81^{b-f}$ | $0.29^{a-f}$ | $2.22^{a-e}$ | $2.51^{a-e}$ | $0.13^{a-e}$ |
| CS-M18 | $23.38^{a}$ | $6.25^{a-d}$ | $333.91^{a-e}$ | $1.30^{b-f}$ | $0.88^{a-c}$ | $0.27^{b-f}$ | $2.18^{a-e}$ | $2.44^{a-e}$ | $0.13^{d, e}$ |
| CS-M22sh | $16.50^{c-h}$ | $5.25^{d-e}$ | $368.18^{a-e}$ | $1.37^{a-f}$ | $0.77^{b-g}$ | $0.28^{a-f}$ | $2.14^{b-e}$ | $2.41^{b-e}$ | $0.13^{b-e}$ |
| CS-T02 | $17.23^{b-f}$ | $5.50^{c-e}$ | $341.52^{a-e}$ | $1.49^{a-e}$ | $0.80^{b-f}$ | $0.32^{a-e}$ | $2.28^{a-d}$ | $2.60^{a-d}$ | $0.14^{a-e}$ |
| CS-T04 | $18.80^{a-f}$ | $6.00^{a-e}$ | $451.54^{a, b}$ | $1.82^{a, b}$ | $1.00^{a-c}$ | $0.38^{a-c}$ | $2.83^{a, b}$ | $3.20^{a, b}$ | $0.13^{a-e}$ |
| CS-T06 | $19.43^{a-e}$ | $7.25^{a}$ | $425.95^{a-c}$ | $1.68^{a-c}$ | $0.90^{a-d}$ | $0.32^{a-e}$ | $2.58^{a-c}$ | $2.89^{a-c}$ | $0.12^{e}$ |
| CS-T08sh | $21.13^{a-d}$ | $6.00^{a-e}$ | $415.77^{a-c}$ | $1.63^{a-c}$ | $1.00^{a-c}$ | $0.31^{a-e}$ | $2.62^{a-c}$ | $2.93^{a-c}$ | $0.12^{e}$ |
| CS-T11sh | $11.50^{g-h}$ | $3.50^{f}$ | $239.64^{e}$ | $0.90^{f}$ | $0.44^{g}$ | $0.19^{f}$ | $1.33^{e}$ | $1.52^{e}$ | $0.14^{a-e}$ |
| CS-T15sh | $18.75^{a-f}$ | $5.75^{b-e}$ | $379.91^{a-d}$ | $1.40^{a-f}$ | $0.82^{b-f}$ | $0.28^{a-f}$ | $2.22^{a-e}$ | $2.50^{a-e}$ | $0.13^{d, e}$ |
| CS-T17 | $20.80^{a-e}$ | $6.75^{a-c}$ | $427.86^{a-c}$ | $1.63^{a-c}$ | $0.99^{a-c}$ | $0.38^{a-c}$ | $2.61^{a-c}$ | $2.99^{a-c}$ | $0.14^{a-e}$ |
| CS-T18 | $21.25^{a-d}$ | $6.50^{a-d}$ | $386.07^{a-d}$ | $1.56^{a-e}$ | $1.02^{a-c}$ | $0.36^{a-c}$ | $2.58^{a-c}$ | $2.94^{a-c}$ | $0.14^{a-e}$ |
| CS-T22sh | $11.38^{h}$ | $5.50^{c-e}$ | $316.69^{c-e}$ | $1.03^{d-f}$ | $0.48^{f-g}$ | $0.22^{e, f}$ | $1.51^{d, e}$ | $1.73^{d, e}$ | $0.14^{a-e}$ |
| Pima 3–79 | $19.88^{a-e}$ | $5.50^{c-e}$ | $414.33^{a-c}$ | $1.49^{a-e}$ | $0.86^{a-c}$ | $0.30^{a-f}$ | $2.36^{a-d}$ | $2.65^{a-d}$ | $0.12^{e}$ |
| TM-1 | $16.63^{b-h}$ | $5.75^{b-e}$ | $398.77^{a-d}$ | $1.61^{a-c}$ | $0.86^{a-e}$ | $0.35^{a-d}$ | $2.47^{a-c}$ | $2.82^{a-c}$ | $0.14^{a-e}$ |

**Table 2** continued

| Name of (CS)-lines | Longest root length (cm) | Root length (cm/plant) | Root surface area ($cm^2$/plant) | Avg root diameter (cm) | Root volume ($cm^3$) | Number of root tips | Number of root forks | Number of root crossings |
|---|---|---|---|---|---|---|---|---|
| CS-B02 | 29.18$^{b}$ | 2374.43$^{a-e}$ | 352.83$^{a-d}$ | 0.47$^{b-e}$ | 4.17$^{a-e}$ | 2257.00$^{c-g}$ | 14,318.3$^{b-e}$ | 1091.75$^{b-e}$ |
| CS-B04 | 24.13$^{b-e}$ | 2376.34$^{a-e}$ | 341.53$^{a-d}$ | 0.46$^{b-g}$ | 3.91$^{a-f}$ | 2411.25$^{a-f}$ | 15,723.3$^{b-d}$ | 1243.00$^{b-e}$ |
| CS-B06 | 26.05$^{b-d}$ | 2031.39$^{b-f}$ | 311.58$^{a-f}$ | 0.49$^{b, c}$ | 3.81$^{a-g}$ | 1952.00$^{e-h}$ | 13,065.3$^{c-e}$ | 1045.00$^{b-e}$ |
| CS-B08sh | 24.23$^{b-e}$ | 2200.79$^{b-e}$ | 314.18$^{a-f}$ | 0.46$^{b-g}$ | 3.57$^{a-h}$ | 2996.00$^{a, b}$ | 15,143.5$^{b-d}$ | 1009.25$^{b-f}$ |
| CS-B11sh | 24.93$^{b-d}$ | 2534.26$^{a-c}$ | 393.66$^{a, b}$ | 0.49$^{b}$ | 4.87$^{a}$ | 2576.25$^{a-e}$ | 16,692.3$^{a-c}$ | 1023.00$^{b-e}$ |
| CS-B15sh | 23.75$^{b-e}$ | 2412.01$^{a-e}$ | 317.19$^{a-e}$ | 0.42$^{g, h}$ | 3.32$^{c-h}$ | 2315.00$^{b-g}$ | 14,693.8$^{b-e}$ | 1410.00$^{a-c}$ |
| CS-B17 | 27.55$^{b-d}$ | 2471.95$^{a-d}$ | 352.41$^{a-d}$ | 0.46$^{b-g}$ | 4.01$^{a-f}$ | 2542.50$^{a-e}$ | 14,667.8$^{b-e}$ | 1234.00$^{b-e}$ |
| CS-B18 | 28.15$^{b, c}$ | 2237.31$^{a-e}$ | 344.83$^{a-d}$ | 0.49$^{b, c}$ | 4.24$^{a-e}$ | 2277.25$^{b-g}$ | 14,964.0$^{b-d}$ | 1015.50$^{b-f}$ |
| CS-B22sh | 25.58$^{b-d}$ | 1940.85$^{c-f}$ | 276.99$^{c-f}$ | 0.46$^{b-g}$ | 3.16$^{e-h}$ | 2830.75$^{a-d}$ | 12,289.3$^{c-f}$ | 946.75$^{b-f}$ |
| CS-M02 | 22.43$^{c-e}$ | 2311.39$^{a-e}$ | 341.40$^{a-d}$ | 0.47$^{b-e}$ | 4.03$^{a-f}$ | 2273.50$^{b-g}$ | 16,721.3$^{a-c}$ | 1303.00$^{a-e}$ |
| CS-M04 | 27.55$^{b-d}$ | 1984.35$^{c-f}$ | 338.01$^{a-e}$ | 0.55$^{a}$ | 4.63$^{a, b}$ | 2194.25$^{d-g}$ | 10,925.3$^{c-f}$ | 813.00$^{d-f}$ |
| CS-M06 | 24.30$^{b-d}$ | 1856.53$^{c-f}$ | 256.14$^{d-f}$ | 0.44$^{d-h}$ | 2.82$^{f-h}$ | 1708.25$^{f-h}$ | 11,150.5$^{c-f}$ | 1072.00$^{b-e}$ |
| CS-M08sh | 22.68$^{c-e}$ | 2070.33$^{b-f}$ | 301.50$^{b-f}$ | 0.47$^{b-f}$ | 3.52$^{b-h}$ | 2166.50$^{d-g}$ | 13,142.3$^{c-e}$ | 987.00$^{b-f}$ |
| CS-M11sh | 22.15$^{c-e}$ | 1704.20$^{e, f}$ | 256.16$^{d-f}$ | 0.48$^{b-e}$ | 3.07$^{e-h}$ | 1396.25$^{h}$ | 8780.5$^{e, f}$ | 792.25$^{e, f}$ |
| CS-M15sh | 28.08$^{b, c}$ | 2966.24$^{a}$ | 406.76$^{a}$ | 0.44$^{d-h}$ | 4.45$^{a-d}$ | 2921.25$^{a-c}$ | 22,025.0$^{a}$ | 1827.25$^{a}$ |
| CS-M17 | 23.90$^{b-e}$ | 2035.59$^{b-f}$ | 283.46$^{c-f}$ | 0.43$^{f-h}$ | 3.15$^{e-h}$ | 2034.00$^{e-h}$ | 11,917.8$^{c-f}$ | 1044.00$^{b-e}$ |
| CS-M18 | 21.45$^{d, e}$ | 2011.64$^{b-f}$ | 253.60$^{d-f}$ | 0.40$^{h}$ | 2.55$^{h}$ | 2878.00$^{a-d}$ | 13,145.8$^{c-e}$ | 1212.25$^{b-e}$ |
| CS-M22sh | 22.73$^{c-e}$ | 2036.57$^{b-f}$ | 290.54$^{b-f}$ | 0.45$^{c-g}$ | 3.30$^{c-h}$ | 2001.75$^{e-h}$ | 12,750.5$^{c-f}$ | 1058.25$^{b-e}$ |
| CS-T02 | 26.13$^{b-d}$ | 1762.39$^{d-f}$ | 261.27$^{d-f}$ | 0.48$^{b-e}$ | 3.09$^{e-h}$ | 1960.25$^{e-h}$ | 10,364.8$^{d-f}$ | 845.00$^{d-f}$ |
| CS-T04 | 27.38$^{b-d}$ | 2742.40$^{a, b}$ | 392.92$^{a, b}$ | 0.45$^{b-g}$ | 4.49$^{a-c}$ | 2968.25$^{a-c}$ | 19,358.5$^{a, b}$ | 1432.25$^{a, b}$ |
| CS-T06 | 29.88$^{a, b}$ | 1863.84$^{c-f}$ | 276.48$^{c-f}$ | 0.48$^{b-e}$ | 3.27$^{d-h}$ | 1675.25$^{g, h}$ | 10,293.0$^{d-f}$ | 823.75$^{d-f}$ |
| CS-T08sh | 28.33$^{b, c}$ | 2386.46$^{a-e}$ | 316.14$^{a-f}$ | 0.42$^{f-h}$ | 3.34$^{c-h}$ | 2501.00$^{a-e}$ | 15,826.0$^{b-d}$ | 1411.50$^{a-c}$ |
| CS-T11sh | 17.95$^{e}$ | 1798.16$^{c-f}$ | 259.01$^{d-f}$ | 0.46$^{b-g}$ | 2.98$^{e-h}$ | 1701.33$^{f-h}$ | 9898.3$^{d-f}$ | 812.33$^{c-f}$ |
| CS-T15sh | 25.25$^{b-d}$ | 2266.73$^{a-e}$ | 310.88$^{a-f}$ | 0.44$^{d-h}$ | 3.41$^{b-h}$ | 2626.50$^{a-e}$ | 14,639.0$^{b-e}$ | 1196.50$^{b-e}$ |
| CS-T17 | 25.23$^{b-d}$ | 2444.09$^{a-e}$ | 371.30$^{a-c}$ | 0.48$^{b-d}$ | 4.51$^{a-c}$ | 2310.00$^{b-g}$ | 14,783.8$^{b-d}$ | 1127.75$^{b-e}$ |
| CS-T18 | 27.05$^{b-d}$ | 2381.30$^{a-e}$ | 326.85$^{a-e}$ | 0.44$^{e-h}$ | 3.57$^{b-h}$ | 3086.50$^{a}$ | 16,597.3$^{a-c}$ | 1359.50$^{a-d}$ |
| CS-T22sh | 29.90$^{a, b}$ | 1364.64$^{f}$ | 212.85$^{f}$ | 0.49$^{b}$ | 2.65$^{g, h}$ | 1733.75$^{f-h}$ | 6831.8f | 454.50$^{f-h}$ |
| Pima 3–79 | 35.95$^{a}$ | 1727.42$^{d-f}$ | 236.51$^{e-f}$ | 0.44$^{f-h}$ | 2.61$^{g, h}$ | 2782.50$^{a-d}$ | 9862.8$^{d-f}$ | 810.25$^{d-f}$ |
| TM-1 | 25.63$^{b-d}$ | 2117.97$^{b-e}$ | 303.91$^{a-f}$ | 0.46$^{b-g}$ | 3.48$^{b-h}$ | 1943.25$^{e-h}$ | 14,151.5$^{b-e}$ | 1222.50$^{b-e}$ |

Means within a column followed by the same letter are not significantly different at the 5% level as determined by Turkey's *t* test

importance to the cluster analysis results for genotypes and traits.

Cluster analysis of CS lines revealed that they could be broadly be divided into three groups based on common phenotypic trends (Fig. 4). Eight CS lines, namely CS-B02, -B11sh, -B17, -B18, -T04, -T17, -M04 and -M15sh, clustered at the top and exhibited strongly positive (red) values for multiple traits; though averages were somewhat variable for some traits, they were quite uniformly high for root dry weight, surface area and volume. Clustered at the bottom of Fig. 4 were three lines at the opposite end of the phenotypic spectrum, namely CS-M11sh, -T11sh and -T22sh. Their values were relatively uniformly low (blue) across nearly all traits, especially plant height, stem dry weight, number of leaves, and number of root tips. In spite of sharing low values for most traits, the average root diameters for these three lines were similarly intermediate to slightly high. The third (middle) group contained the most lines, including 16 CS lines, TM-1 and 3–79. However, it included 2 distinctive subgroups. One subgroup included 10 CS lines and TM-1; these lines exhibited variable but mostly intermediate values for the majority of the traits. The other subgroup, comprising 6 CS lines and *G. barbadense* 3–79, were non-uniformly low (blue) for the majority of analyzed traits. The lines in this subgroup, including CS-M06, -M17, -M22sh, -T02, -M08sh, and -T06, were strongly differentiated from the bottom-most clade by several traits—plant height, stem dry weight, number of leaves, number of root tips (Fig. 4). The heat map reveals that CS-M15sh was unusual in having high to moderately high values for all traits, except for root diameter (moderately low).

The cluster analysis among the 12 traits depicted in Fig. 4 on the horizontal axis grouped the traits into one solitary trait (root diameter) and two broad divisions (Fig. 4). A visual review of the heat map shows that the pattern of values for root diameter (rightmost trait) across the CS lines is very distinctive. All three above-ground traits and two root traits were clustered in the leftmost group, whereas the right group included only root traits. The cluster analysis and heat map suggest the potential presence of two different development patterns among the CS lines for above versus under the ground. It is interesting to note that CS-M15sh had red to far red color for the majority traits, i.e., high values, whereas CS-T22sh had blue to far blue color i.e., low values, on the same traits, a pattern indicating that these two lines differ strongly for most of the traits.

### CS line-specific effects on traits

The effect of each CS line on each trait was evaluated using pairwise *t*-test-based comparisons of line-based means (Tables 2, 4). Particular importance was placed on comparisons to TM-1, due to its quasi-isogenic relationship to the CS lines. Significantly negative deviations from TM-1 occurred for 12 traits, and involved 4 different CS lines, two from each of the donors *G. tomentosum* and *G. mustelinum*. CS-T11sh differed significantly from TM-1 on seven different traits, with reductions in longest root length, numbers of root forks, leaves, leaf area, and the five dry weight traits. For three of these same traits, the CS line disomic for the *G. mustelinum* homolog, i.e., CS-M11sh, also exhibited significant reductions relative to TM-1. CS-T22sh differed significantly from TM-1 for eight traits, including reductions in root length, numbers of root forks and crossings, and dry weights of leaves, stems, shoots, roots, and total plants. CS-M15sh, was significantly higher for four root traits—length, counts of tips, forks and crossings, perhaps underscoring its potential value. Overall, CS lines had significantly positive deviations from TM-1 for 8 traits, including 6 root traits, including length, diameter, volume, and counts of tips, forks and crossings (Tables 2, 4),. Moreover, these involved 9 different CS lines, three from each of the three donor species. These phenotypic associations suggest that multiple CS lines likely harbor genetic variation that will be useful for studying and perhaps improving cotton root systems.

## Discussion

Geneticists and breeders are continually attempting to breed plants for root, stem, and leaf traits that improve crop productivity. Several challenges, however, are hindering this breeding process, including the narrow genetic base, limited information on how important morphological traits are controlled, the polyploid nature of the genome, and the complex genome with duplicated loci in the genetic improvement of Upland cotton. The overall objective of this research was to improve our understanding of the genetic knowledge on the morphological of shoot and root traits and

**Table 3** Correlation analysis of some important traits in chromosome substitution lines

| Correlations | Plant height | Leaf number/plant | Total leaf area | Leaf dry weight | Stem dry Weight | Root dry weight | Above ground weight | Total dry weight | Root/shoot ratio |
|---|---|---|---|---|---|---|---|---|---|
| Plant height | 1 | 0.9229** | 0.9067** | 0.9115** | 0.9549** | 0.8385** | 0.9288** | 0.9215** | − 0.2414 |
| Leaf number | | 1 | 0.9218** | 0.928** | 0.9107** | 0.8863** | 0.9206** | 0.9195** | − 0.1247$^{n}$ |
| Leaf area | | | 1 | 0.986** | 0.9634** | 0.9318** | 0.9829** | 0.9842** | − 0.2366$^{n}$ |
| Leaf dry weight | | | | 1 | 0.9721** | 0.947** | 0.9939** | 0.9961** | − 0.1879$^{n}$ |
| Stem dry weight | | | | | 1 | 0.8904** | 0.9901** | 0.9858** | − 0.2794* |
| Root dry weight | | | | | | 1 | 0.9233** | 0.9342** | − 0.0482$^{n}$ |
| Above ground weight | | | | | | | 1 | 0.9993** | − 0.2419$^{n}$ |
| Total dry weight | | | | | | | | 1 | − 0.2265$^{n}$ |
| Root/shoot ratio | | | | | | | | | 1 |
| Longest Root Length | | | | | | | | | |
| Root length | | | | | | | | | |
| Root surface area | | | | | | | | | |
| Average root diameter | | | | | | | | | |
| Root Volume | | | | | | | | | |
| Number of root tips | | | | | | | | | |
| Number of root forks | | | | | | | | | |
| Number of root crossings | | | | | | | | | |

| Correlations | Longest root length | Root length | Root surface area | Average root diameter | Root volume | Number of root tips | Number of root forks | Number of root crossings |
|---|---|---|---|---|---|---|---|---|
| Plant height | 0.4668** | 0.8004** | 0.7590** | − 0.1492 | 0.6462** | 0.8392** | 0.8139** | 0.8178** |
| Leaf number | 0.5994** | 0.7404** | 0.7485** | 0.0476 | 0.6906** | 0.7726** | 0.7411** | 0.7396** |
| Leaf area | 0.4709** | 0.8241** | 0.8162** | − 0.0304 | 0.7269** | 0.8045** | 0.8309** | 0.8154** |
| Leaf dry weight | 0.4467** | 0.8569** | 0.8512** | − 0.051 | 0.7592** | 0.796** | 0.8534** | 0.8434** |
| Stem dry weight | 0.3834** | 0.8655** | 0.8223** | − 0.1697 | 0.694** | 0.8332** | 0.873** | 0.8721** |
| Root dry weight | 0.4817** | 0.8385** | 0.8836** | 0.1758 | 0.8522** | 0.7449** | 0.8162** | 0.8057** |
| Above ground weight | 0.4117** | 0.8637** | 0.8387** | − 0.1129 | 0.7253** | 0.8103** | 0.8652** | 0.8597** |
| Total dry weight | 0.4123** | 0.8661** | 0.846** | − 0.0914 | 0.737** | 0.8044** | 0.8652** | 0.859** |
| Root/shoot ratio | 0.0086$^{n}$ | − 0.2789$^{n}$ | − 0.1709$^{n}$ | 0.2257$^{n}$ | .0122$^{n}$ | − 0.4916$^{n}$ | − 0.3669$^{n}$ | − 0.3773$^{n}$ |
| Longest root length | 1 | 0.2609* | 0.3593** | 0.2436 | 0.427** | 0.4264** | 0.272* | 0.2241 |
| Root length | | 1 | 0.9548** | − 0.1511 | 0.7943** | 0.866** | 0.9883** | 0.9732** |
| Root surface area | | | 1 | 0.0315 | 0.9238** | 0.8253** | 0.9367** | 0.8911** |

**Table 3** continued

| Correlations | Longest root length | Root length | Root surface area | Average root diameter | Root volume | Number of root tips | Number of root forks | Number of root crossings |
|---|---|---|---|---|---|---|---|---|
| Average root diameter | | | | 1 | 0.2443* | − .1843$^{n}$ | − .1986$^{n}$ | − 0.2924* |
| Root volume | | | | | 1 | 0.6948** | 0.7715** | 0.7252** |
| Number of root tips | | | | | | 1 | 0.8932** | 0.8966** |
| Number of root forks | | | | | | | 1 | 0.9802** |
| Number of root crossings | | | | | | | | 1 |

The correlations were estimated by pairwise methods. The significance levels are summarized as n, * and ** represent nonsignificance, $p \leq 0.05$ and $p \leq 01$ levels of significance, respectively

improve the genetic diversity in Upland cotton by targeted introgression of useful morphological traits using CS lines from *G. barbadense*, *G. tomentosum,* and *G. mustelinum*. Knowledge of the morphology of cotton stem, leaf and root, and its developmental mechanisms will allow manipulation and exploitation of different traits to improve genetic productivity of the crop. In this study, the 25-day old cotton seedlings were used to morphologically characterize its root, stem, and leaf.

Major opportunities for improved crop productivity and sustainability are expected if root systems and root functions can be significantly improved (Lynch and Brown 2012). Though cotton plants become woody as they mature, it is extremely important for cotton to establish a uniform, vigorous, healthy stand in the early stage of the life cycle (Eissa 1983). Whereas past crop improvement efforts have focused much more on yield and shoot-related traits rather than root systems (Den Herder et al. 2010), the emphasis here was on finding novel genetic variation affecting young cotton root system morphological traits. Roots play an important role in the life of cotton against lodging, and as a major organ used for interface sensing, and interacting with, water uptake, nutrition, and interaction with biotic and abiotic factors in the soil. Previous studies indicated that multiple genes are involved in the inheritance of seedling growth (Balls 1919; Eissa et al. 1983). However, very limited information is available on the genetic mechanisms associated with the cotton root. This is primarily because of difficulties associated with studying root phenotypes under natural field conditions.

We assessed 17 shoot and root characteristics of 25-day old seedlings across 29 genotypes, 28 of which were quasi-isogenic, including 27 interspecific CS lines. Each CS line was bred previously to replace both copies of a chromosome or a large chromosome segment with the equivalent homolog or segment from one of three AD-genome 2n = 4x = 52 species in the cotton primary gene pool, namely *G. barbadense*, *G. mustelinum* and *G. tomentosum*. Analysis of data means, dispersion, correlations, two-way hierarchical clustering of selected traits and a heat-map of their inter-relationships with the genotypes indicated that the CS lines significantly affected most of the evaluated traits, including multiple root system traits. Thus, the CS lines seem likely to be useful for research and perhaps applied breeding and genetic analysis for

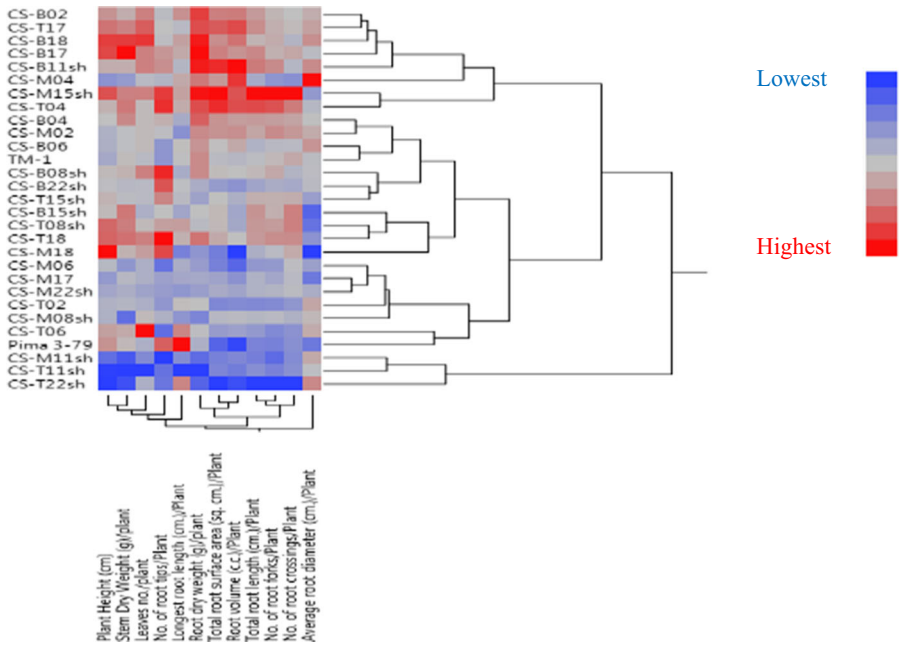


**Fig. 4** Heat map and dendrograms generated from hierarchical cluster analysis using JMP Genomics showing the relationships among the chromosome substitution lines based on the selected traits using two-way clustering method. Clustering among CS lines is depicted along the Y-axis, whereas cluster relationships among the traits are depicted along the X-axis of the heat map. Stronger red colors represent by higher positive values and stronger blue colors represent lower positive values

**Table 4** Chromosomal effect on root and shoot traits based on the mean comparison t-test with TM-1

| Trait | Chromosome substitution line |
|---|---|
| Leaf number per plant | CS-T06 |
| Total leaf area | CS-T11sh |
| Leaf dry weight per plant | CS-M11sh. CS-T11sh. CS-T22sh |
| Stem dry weight per plant | CS-T22sh. CS-T11sh |
| Root dry weight | CS-T22sh. CS-T11sh |
| Above ground plant weight | CS-T22sh. CS-M11sh. CS-T11sh |
| Total plant dry weight | CS-T22sh. CS-M11sh. CS-T11sh |
| Longest individual root length | CS-T11sh |
| Total root length | CS-T22sh. CS-M15sh |
| Root diameter | CS-M04. CS-M18 |
| Root volume | CS-B11sh |
| Number of root tips | CS-T18. CS-B08sh. CS-T04. CS-M15sh. CS-M18. CS-B22sh |
| Number of root forks | CS-M15sh. CS-M11sh. CS-T22sh |
| Number of root crossings | CS-M15sh. CS-T22sh |

The recurrent parent of the chromosome substitution lines

these traits. The discussion below highlights significant features of the findings, as well as their limitations. We surmise a small step toward genetic enhancement of cotton root systems has been achieved, but that numerous challenges remain in genetically interpreting, defining and manipulating this new diversity for research and breeding purposes, as discussed below.

Data interpretation

*CS lines and traits*

Ranges of the phenotypic ratings for the 17 traits ranged quite broadly among the 27 CS lines and two inbreds. Variance associated with line-to-line differences was significant for all 9 root-only traits and 12 of the 17 overall traits, as listed in Fig. 4. The ANOVA results were largely concordant with the previous studies on the cotton root, leaf and stem morphology with the CS and other elite cotton lines (Awasthi et al. 2018; Singh et al. 2018). Assuming that the line-to-line contributions to variance were largely of genetic basis, these CS lines should be useful breeding-relevant research and genetic improvement efforts.

*Isogenic comparisons with TM-1 to detect single- and multi-trait effects*

The comparisons of each CS line to TM1 provided a convenient vehicle for detecting major overall effects on any trait of interest, because the inbred 'TM-1' provided the genetic background common to all of the recurrent backcross parents of these CS lines, whereby it and the CS lines are quasi-isogenic. However, conducting large numbers of *t* tests, as done here, increases the experiment-wide likelihood of a Type-1 error, so guarded caution is warranted in the statistical significance levels. Among the 9 CS lines that found to be significantly superior to TM-1 for one or more traits, CS-M15sh was exceptional in that it significantly increased 4 root traits—length and counts of tips, forks and crossings. The multiplicity of trait effects by CS-M15sh traits seems to strengthen confidence in the findings for each trait, and so overall results suggest this line warrants further consideration and research as a possible source of root system modification and improvement. Moreover, the two related substitutions, CS-T15sh and CS-B15sh were quite inferior for several traits, and positioned in a different major cluster by the two-way hierarchical cluster analysis. Note, too, that this difference could indicate that the positive effects are due to the *G. mustelinum* substitution per se, rather than a loss of a deleterious TM-1 factor or segment.

CS-M18 was associated with significant increases of two traits—plant height and numbers of root tips, and also a significant decrease in root diameter. The concomitant decrease in root diameter was in keeping with the overall experiment, in which negative correlations were exhibited with plant height and numbers of root tips, albeit nonsignificantly. As above, the association of a CS line with significant changes in multiple traits seems to strengthen confidence in results for each of the individual traits. Additional evidence to this effect is provided by the observed increase in root tip count noted for CS-T18, and analogous chromosome substitution from a different donor species. Superiority of not just one but two related CS lines, i.e., CS-T18 and CS-M18, to TM-1 for this trait, could be an indication that TM-1 chromosome-18 contains one or more factors that tends to reduce the number of root tips, at least relative to chromosome-18 of the *G. tomentosum* and *G. mustelinum* donors. CS-B18 line, the third member of this chromosome-18 series, i.e. from *G. barbadense*, was intermediate and nonsignificantly different from TM-1 and CS-M18 for this trait, but significantly lower than CS-T18. The two-way hierarchical cluster analysis placed CS-B18 far away from CS-T18 and CS-M18, suggesting the possibility of different "alleles" controlling some of these traits. While CS-T18 seems likely to offer potentially useful variation in root proliferation, additional research will be needed to sort out details and relative species-specific effects. We noted that six different CS lines exhibited significantly higher numbers of root tips per plant than TM-1, more than any other trait; the relative importance of the abundance of lines exhibiting superiority is not known—it could merely reflect a spuriously low TM-1 root tip count mean, or something more profound, such as a multi-locus reduction of root proliferation in the TM-1, Upland types or *G. hirsutum* species. Some additional investigation of seedling root proliferation (tip number) in TM-1 and other *G. hirsutum* lines, seems warranted.

The chromosome-11 substitution lines, CS-T11sh and CS-M11sh, were inferior to TM-1 for multiple

traits. CS-T11sh differed significantly from TM-1 for eight traits.—longest root length, numbers of root forks and leaves, leaf area, and the five dry weight traits, and for three of these same traits. The third member of this series, CS-B11sh, behaved much differently than CS-T11sh and CS-M11sh, and it was significantly better than TM-1 for another trait—root volume. Moreover, CS-B11sh was separated widely from CS-T11sh and CS-M11sh in the two-way hierarchical cluster analysis. The observations suggest significant inferiority to TM-1 of CS-T11sh and CS-M11sh for multiple traits, but not CS-B11sh. It is noteworthy, perhaps, that chromosome 11 and its D-subgenome homeolog, chromosome 21 of the AD-genome cottons are known to contain resistance genes to various fungal and at least two nematode pathogens (Bolek et al. 2005; Zhang et al. 2015). It highly conceivable, though certainly not proven, that these differential effects have resulted from epistatic effects from species-specific changes in the profiles of R-genes that reside in chromosome 11. The relatively superior performance of CS-B11sh could involve analogous differences in chromosome-11 and/or other inadvertently substituted chromosome segments. Recent comparisons of new AD-genome assemblies indicate that in each species the chromosome 11 is well endowed with multiple R-genes, many of which differ among these species (Chen et al. 2020). Testing this hypothesis is very feasible, e.g., by genetic dissection and comparison of the respective chromosome-11 elements. Regardless, a determination of the basis and exact location of negative effects for each donor chromosome will likely increase the usefulness of that CS line for improving other traits, i.e., by enabling marker assisted selection to eliminate the locus causing undesirable effects on the root system.

Two other CS lines were significantly inferior for one or more traits relative to TM-1. CS-T22sh was associated with decreases in multiple shoot and root traits—dry weights (leaf, stem, root, shoot, and total), root length and counts of root tips, forks and crossings. Further research is needed to sort out the independence or interdependence of these effects and to localize and identify their genetic basis. One CS line, CS-M18, was associated with a significantly decrease of only one trait relative to TM-1, namely root diameter. As mentioned above, this CS line had significant associations with increased plant height and numbers of root tips. The directional discordance followed the overall patterns between these traits, both of which were nonsignificantly negatively correlated with root diameter.

### *Two-way hierarchical cluster analysis of traits and CS lines*

The relationships among CS lines defined by two-way hierarchical cluster analysis were congruent with a number of the homology and homeology relationships among the targeted chromosomes. While the significance of the patterning among clusters remains largely speculative, they are of interest and could indicate induction of similar integrated multi-trait phenotypic syndromes by homologs and/or homeologs of the same or different species. The best-performing cluster of 8 CS lines included CS-B02, T17, B18, B17, B11sh, M04, M15sh and T04, half of which involve homologs, i.e., M04 and T04, as well as CS-T17 and CS-B18; another two, CS-B02 and CS-B17, involve homeologs chromosome 2 (A genome) and 17 (D genome) of *G. barbadense* genome. The large middle cluster of the heat map included three pairs of homologs CS-T15sh and CS-M15sh, CS-T18 and CS-M18, and CS-B22sh and CS-M22sh, plus two segmental homeologs of *G. barbadense*, CS-B04 and CS-B22sh. In contrast, *G. hirsutum* TM-1 and *G. barbadense* were also clustered within the second group, but far away from each other in distal subclusters. At the bottom of the same figure, the small cluster of three inferior lines included two homologs, CS-M11sh and CS-T11sh, while the third member of the series, CS-B11sh, was mapped far away to the top-tier cluster. These patterns could reflect common effects from similar singular and/or multigenic substitutions—a possibility that might be explored by transcriptome or other molecular analysis. In contrast to the above pheno-clusters, CS-B02 and CS-T22sh were the most diverse among the tested lines, these involve two nonhomologous and non-homeologous targeted chromosomes, and were positioned at the two extreme ends of the vertical axis in the heat map. To intermate the eight best-performing CS lines and study their quasi-isogenic progenies might be a worthwhile genetic and breeding experiment.

## Interpreting the genetics

Some caution is warranted in drawing inferences from the CS line-trait associations about direct genetic effects versus epistatic interactions. The assessments in this study are based only on homozygous lines, not any derived hybrids or hybrid progenies. Thus, the two genetic modes of action, epistasis versus the individual effects of the substituted chromosome, remain inseparable. Differences between any CS line and TM-1, whether positive or negative, could have been caused by one or more gene(s) on the specific substituted chromosome or chromosome segment(s), and/or interactions between them and TM-1 genes on other chromosomes and segments. Differences might be due to loss of a function provided by a functional TM-1 allele(s), gain of function from a newly introduced donor allele(s) or gene(s), and/or some other alteration, e.g., differing expression patterns.

Caution is also warranted in terms of ascribing CS line effects to a specific location. An assumption that facilitates inferential comparisons of CS lines and TM1 is that all CS lines are quasi-isogenic to each other and to TM-1. In most situations, this assumption is almost certainly correct, and the degree isogenicity is likely sufficient in most regards. Even when not true, the phenotypic assessments and differences between lines would remain valid; however, the number and locations of factors contributing to the difference would be more complicated than if these CS lines were perfectly isogenic.

The expectation of quasi-isogenicity is based on the breeding procedures used to create them, first by developing TM-1-isogenic hypoaneuploids, then by using those to create TM-1-isogenic monosomic substitutions, followed by recovery of isogenic disomic substitutions. Prior to CS line development, hypoaneuploids were created a new in a TM-1 background or if a different genetic background, they were backcrossed as females into the TM-1 background (Stelly et al. 2005). Isogenic monosomic substitutions were created by hybridizing the donor as pollen parent to the TM-1-isogenic hypoaneuploid, isolating the hemizgyous monosomic F1 hybrid, and backcrossing repeatedly with similar steps each backcross cycle to the respective TM-1-isogenic hypoaneuploid. Maintained in hemizygous state during backcross-mediated introgression, the targeted donor chromosome remains nonrecombinant as it lacks a homologous meiotic partner; meanwhile nontargeted donor chromosomes and segments are passively lost. Once monosomic substitutions reached the $BC_5$ or higher, self-pollination allowed facile recovery of homozygous disomic substitutions, which are true-breeding and readily seed-increased and used with facility as parents for breeding. After completing backcrossing and selection of the backcross monosomic hemizygote, e.g., $BC_5F_1$, the inbreeding used to recover the homozygous disomic substitution renders homozygous about half of the heterozygous nontargeted donor segments that happen to exist in the backcross hybrid. While the default expectation is that CS lines are homozygous for the substituted chromosomes or chromosome segment from the alien species and otherwise mostly identical to TM-1 (Fig. 1), some varying number of inadvertently retained donor segments are expected in some if not most CS lines. Thus, most phenotypic differences among CS lines and between CS lines and TM-1 are likely due to the targeted substitution, but could be due to inadvertent non-targeted substitution. The presence of such inadvertent segments has little impact, because the discovery of important trait effects remains significant because subsequent analyses will nonetheless lead to similar advances, albeit involving locations different from those anticipated based on CS line identities. The discovery of such locations is relatively tractable using contemporary genotyping technologies.

## Addressing major needs

Two major long-term critical needs for Upland cotton improvement were partially addressed by this research—[1] additional genetic diversity that is beneficial, and [2] more genetic knowledge and resources for root system improvement. The CS lines of Upland cotton used in this research contain germplasm from the three donor tetraploid cotton species, and they were analyzed for key seedling morphological traits, most notably root system traits. Recently completed genome-wide sequence assemblies revealed that many novel alleles are harbored by these three alien species of Upland cotton (Chen et al. 2020). Aside from *G. barbadense*, attempts to use this diversity have been very limited; virtually none from *G. mustelinum* and *G. tomentosum* has been utilized in the improvement of Upland cotton cultivars.

One of the major constraints in the genetic improvement of Upland cotton has been the lack of effective breeding materials with desirable phenotypes, especially for valuable traits under quantitative multigenic control. Variants with qualitative effects are relatively easily recognized, whether in an elite domesticated genotypic background, wild type of a domesticated species or a related species. Examples include morphological traits, such as the *nectariless* trait of *G. tomentosum* (Meyer and Meyer 1961), certain pest/pathogen resistance traits, e.g., to reniform nematodes (Yik and Birchfield 1984). In contrast, it is especially challenging to use the more or less conventional approach of screening wild germplasm directly to identify desirable types, then introgress and use variation that might exist among wild accessions and species for complex multigenic traits, especially if that trait has been extensively improved upon during domestication and modern breeding periods.

The development and subsequent screening of backcross-inbreds (Wehrhahn and Allard 1965), chromosome substitutions (Sears 1952) or chromosome segment substitutions (Eshed and Zamir 1995) offers means to discover positive effects and interactions in the genetic context of the cultivated species, while taking advantage of its amenability to seed increase and experimentation. For an average chromosome of Upland cotton, a disomic substitution would replace about 4% of the genome, or about 3,000 genes. CS lines thus provide a powerful basis of discerning collective effects of the donor gene replacements per se, as well as their interactions with the vast majority of the elite genetic background.

In spite of their recognized importance, root systems have been insufficiently characterized even in the cultivated forms of cotton (McMichael and Quisenberry 1991; McMichael et al. 2010). Given the difficulty of root phenotyping, there has been relatively little effort put into genetically improving root systems, per se, compared to overall plant breeding investments. While domestication and adaptations leading to Upland cottons almost certainly inadvertently affected root systems, early germplasm acquisition, "breeding" and selection were focused on other factors mostly above the ground parts, such as overall plant health and productivity, boll health, ease of seedcotton harvesting and, prior to the cotton gin, ease of fiber removal (Moore 1956). Thus, the genetic foundation of Upland cottons may not have been especially conducive to good root system performance or subsequent genetic improvement. It is thus likely that addition of novel diversity for root systems could lead to significant benefits on productivity, water use and sustainability.

## Correlations with potential for root-centric selection

Our results revealed considerable variation among the CS lines for several root and shoot phenotypes (Table 2, Figs. 3, 4). Strong correlations occurred among many of the traits, suggesting that it might be possible to leverage selection for the correlated trait(s) easiest and cheapest to evaluate for simultaneous improvements in multiple highly correlated traits; this might open the door to large-scale breeding for improved root systems in cotton. For example, *G. barbadense* 3–79, a line with very delayed maturity, had the longest root length among all lines, including the recurrent parent (TM-1) of the CS lines. The previous report indicated that a larger root system is associated with delayed maturity and improved water and nutrient uptake for post-anthesis helping in grain filling in wheat (Xie et al. 2017; Pinto and Reynolds 2015). Breeding for emphasized early growth of root systems would seem likely to favor improved capture of water and nutrients later, especially in stress environments, but if that improvement leads to delayed shoot and floral development, counter-acting selection for early development of shoots and flowering may also be needed to maintain those desirable traits.

## Water deficit response

Cotton leaf, stem, and root phenotypes are seriously affected by the major abiotic factors, including water deficit, which seriously limits plant growth and crop productivity around the world (Kramer 1983). Krieg and Sung (1986) reported that water stress caused a reduction in the whole-plant leaf area by decreasing the initiation of new leaves. Pettigrew (2004) indicated that water-deficit stress led to decreasing leaf size, but mentioned that this decrease was accompanied by an increase in the specific leaf weight, a phenomenon also reported by Wilson et al. (1987). Pettigrew (2004) established that both vertical and horizontal boll load are reduced under water deficit, and account for the

majority of yield reduction. Significantly fewer nodes and lower dry weights of stems and leaves were observed for water-stressed plants compared to those of the control (Pace et al. 1999). McMichael and Quisenberry (1991) observed decreased shoot-to-root ratios of plants grown under conditions of severe water stress. Malik et al. (1979) reported that root growth seemed to be less affected by drought than shoot growth, which inferably reflects a water-stress effect, whether passive and/or proactive, that reduces plant resource allocation to shoots and proportionately increases allocations to roots as a means to increase plant water-capture relative to shoot-based demand.

The phenotypic traits for which significant genetic variation exists can potentially be genetically manipulated to help improve cotton productivity. Our previous studies reported that several of these CS lines including CS-T04, and CS-B18 had potential of heat and drought-tolerant root and shoot phenotypes (Awasthi et al. 2018; Reddy et al. 2020). The study reported here seems to validate the earlier finding in that CS-T04 and CS-B18 were among the top-performing clade of eight CS lines (Fig. 4), and seemed to perform better across multiple traits than the related CS lines for the same substituted chromosome from the other alien species such as CS-M04 and CS-B04, CS-T18 and CS-M18 as well as TM-1. We surmise that further analyses of the chromosome-04 and chromosome-18 CS lines, especially CS-T04 and CS-B18 are highly warranted.

## Feasibility of combining long root system with high root weight?

Cotton has a main tap root system that grows straight down from the stem into the soil and can grow down 2–5 cm per day and extend to depths of 3 m or more in the soil. Tap root systems divide and subdivide to produce branch root systems in the soil. Soil and air temperature influence cotton root development. The extraction of intact roots from the soil without damaging the phenotype is very difficult, laborious, time-consuming, and expensive. However, as mentioned above, an advanced method of image analysis was applied here to study early root phenotypes using CS lines (Awasthi et al. 2018; Reddy et al. 2020). Eissa et al. (1983) reported that combining long roots with high relative root weight was a difficult challenge in the development of pure breeding lines in diallel crosses with five cotton parent lines. In one cross, they found additive and additive-by-additive epistasis accounted for about one-half of the variation in root length and two-thirds of the variation for relative root weight. In another cross of the diallel analysis, residual epistasis accounted for most of the inheritance of root length and relative root weight, suggesting selection for long root with relative root weight might not be desirable for this cross. The relative amount of epistasis seems likely to depend on parental genotypes, but also seems likely to be a more prominent factor in wide-crosses. If so, CS lines and similar types of breeding materials, especially isogenic ones, might be especially effective for detecting donor genes that combine well with the background genotype of elite domesticated types.

Our results showed that long total root length was associated with CS-T15sh, and low root dry weight was related to CS-22sh and CS-T11sh. Eissa et al. (1983) suggested that fast-growing roots with high relative root weight would be useful when planting conditions are cool, and seedling diseases are a problem, because the plant should have more root tissues to slough off as diseased tissue while maintaining a viable root system necessary for growth in a healthy plant. Our results showing the association of two different chromosomes and significant positive correlation between these two traits suggests the potential of combining these two traits for genetic improvement by using two different CS lines (Fig. 3, Table 2). Eissa et al. (1983) suggested that recurrent selection with selection delayed to the $F_3$ or $F_4$ generation would be a useful strategy to enhance recombination among epistatic factors and thereby recover desired improvements in root length and relative root weight.

## Seedling root development versus maturity delay

Previous wheat research established that larger root systems are associated with delayed maturity and, in turn, extended grain filling, probably because of improved water and nutrient uptake for post-anthesis photosynthesis (Xie et al. 2017; Pinto and Reynolds 2015). Root absorption capacity is related to total root length, root surface area, and root dry weight (Fitter 1991). Root growth and development depend on the availability of carbohydrates from above ground parts, thus the reduction or increase of leaf area would

expectedly lead to similar effects on root growth (Ogbonnaya et al. 1997). Results showed that the longest root was associated with *G. barbadense* 3–79, a very late-maturing cotton line that forms large seed and leaves compared to TM-1 and several other CS lines. If this relationship is a strong one in cotton, one would infer that historical progressive selection for domestication and then earliness in Upland cottons could have inadvertently impaired some aspects of root system development. If so, racestocks of *G. hirsutum* and other species in the primary gene pool might be useful sources of diversity to regain certain root system features, e.g., using converted racestocks (McCarty et al. 2006) or the alien species donors used herein for chromosome substitution. If genes controlling these two traits are genetically distinct but linked within haplotypic blocks, then special breeding or molecular methods might be needed to create the desired genetic products (Chen et al. 2020).

### Breeding challenges and opportunities

Geneticists and breeders are constantly attempting to breed plants for root, stem, and leaf traits to improve crop productivity. Several challenges, however, are hindering this breeding process, including the narrow genetic base, limited information about the control of important morphological traits, complexity of the genome due to its tetraploid nature and preceding rounds of paleopolyploidization, which favor gene duplications and epistatic interactions. Some of these challenges are exacerbated by an extensive network of variously sized haplotypic blocks (HBs) found throughout the AD genome (Chen et al. 2020). They constitute recombinational constraints that collectively affect a large proportions of beneficial genes in the genome. In addition, many traits are associated with exceptionally low diversity globally among elite types (Chen et al. 2020). Analogous HBs have been identified in sunflower and found important as multigenic factors in development of ecotypes (Todesco et al. 2020), e.g., some perhaps functioning as "supergenes". While cotton HBs seem to constitute a doubly difficult situation for breeding—low variation and low recombination, there are indications that interspecific crosses with *G. barbadense*, *G. mustelinum* and *G. tomentosum*, i.e., the species used as donors to create the CS lines used in this research, may open the door to recombination within HBs to create new variation (Chen et al. 2020). Nonetheless, knowledge of their presence and wide distribution in the *Gossypium* A and D subgenomes expands our awareness of the challenges we face in finding, analyzing and using genetic diversity in Upland cotton improvement. CS lines will help to overcome some of these chaalenges. Our previous studies reported the potential of several of these CS lines in the genetic improvement of agronomic and fiber quality traits in TM-1 and elite Upland cultivars suggesting some of their morphological traits might be associated with the improved effects (Saha et al. 2017; Jenkins et al. 2017a, b).

### CS-RILS

Recently, we have developed chromosome-specific recombinant inbred lines (CS-RILs) using several of these CS lines (Saha et al. 2017). Each CS-RIL differs by the homozygous replacement of one or more specific segments of the previously substituted chromosome of the CS line. These CS-RILs are also quasi-isogenic to each other, the TM-1 inbred and other CS lines, such that each CS-RIL population will be a very powerful analytical tool in future investigations to reveal the locations and genetic mode of action associated with many of these traits. These CS-RILs will provide a novel approach to Upland cotton breeding program by targeted interspecific introgression of many desirable morphological traits from the alien species with reduced linkage drag effects.

## Summary

In summary, this research provides some valuable information towards the potential improvement of several morphological traits in Upland cotton seedlings, including root traits, an aspect of crop development and performance that sorely needs attention. Using novel germplasm sources, namely quasi-isogenic CS and chromosome segment substitutions lines from the donor species *G. barbadense*, *G. tomentosum* and *G. mustelinum*, specific CS lines were associated with significant changes in seedling traits. Significant positive effects on multiple shoot and root system traits were associated with CS-M15sh and CS-M18. Two-way hierarchical clustering revealed a prospectively superior clade of CS lines that included CS-

B02, -T17, -B18, -B17, -B11sh, -M04, -M15sh and -T04; these lines hold promise for research and applied crop breeding. Several additional CS lines including CS-T11sh, CS-T22sh and CS-M11sh significantly reduced multiple morphological traits, which should be mapped to implement negative MAS and recovery of nearby segments devoid of the negative effects. The possibility was raised that differential results involving chromosome 11 from different sources could reflect strong effects of their differing constellations of resistance genes. While the CS line identities suggest the potential locations of genes responsible for the respective effects, unrelated alien segments could also be present. Follow-through genetic research involving CS-RILs will be useful to genetically dissect these traits and map the responsible loci, so that they can be efficiently leveraged for breeding purposes using marker-assisted selection.

**Acknowledgements** We would like to thank Mr. David Brand of Mississippi State University and Ms. Evita Lopez-Gourley, at USDA-ARS for technical assistance and lab analysis, and Mr. Wayne A. Raska for decades of diligence in developing monosomic stocks and chromosome substitution lines. Development of the chromosome substitution lines used in this study was made possible by partial funding support from Cotton Incorporated, including agreement no. 16-349, as well as the Texas State Support Committee, Texas Food & Fiber Commission and the USDA-ARS and from the grant of USDA/ARS CRIS project 6064-21000-016.

**Funding** All authors agreed with the content and that all gave explicit consent to submit and that they obtained consent from the responsible authorities at the institute/organization where the work has been carried out, before the work is submitted. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. A submission to the journal implies that materials described in the manuscript, including all relevant raw data, will be freely available to any researcher wishing to use them for non-commercial purposes, without breaching participant confidentiality. This article does not contain any studies involving animal or human participants performed by any of the authors.

**Compliance with ethical standards**

**Conflicts of interests** All procedures performed in this research were in accordance with the ethical standards of the Committee on Publication Ethics (COPE) as per the policy of the journal. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The U.S. Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or part of an individual's income is derived from any public assistance program.

## References

Arsenault E, Bernard J-T, Carr CW, Genest-Laplante E (1995) A total energy demand model of Québec. Energy Econ 17:163–171. https://doi.org/10.1016/0140-9883(94)00003-Y

Awasthi A, Reddy KR, Saha S, Jenkins JN, Stelly DM (2018) Morph-physiological responses of cotton interspecific chromosome substitution lines to low temperature and drought stresses. Euphytica 214:218–227. https://doi.org/10.1007/S10681-018-2300-6

Balls WL (1919) The cotton plant in Egypt. MacMillan and Company Ltd., London

Beasley JO (1940) The production of polyploids in Gossypium. J Hered 31:39–48

Bolek Y, El-Zik KM, Pepper AE, Bell AA, Magill CW, Thaxton PM, Reddy OU (2005) Mapping of verticillium wilt resistance genes in cotton. Plant Sci 168:581–1590. https://doi.org/10.1016/j.plantsci.2005.02.008

Chen ZJ, Sreedasyam A, Ando A et al (2020) Genomic diversifications of five *Gossypium* allopolyploid species and their impact on cotton improvement. Nat Genet 52:525–533. https://doi.org/10.1038/s41588-020-0614-5

DeJoode DR, Wendel JF (1992) Genetic diversity and origin of the Hawaiian Islands cotton, *Gossypium tomentosum*. Am J Bot 79(1):1311–1319

Den Herder G, Van Isterdael G, Beeckman T, De Smet I (2010) The roots of a new green revolution. Trends Plant Sci 15(11):600–607

Eissa AM, Jenkins JN, Vaughan CE (1983) Inheritance of seedling root length and relative root weight in cotton. Crop Sci 23:1107–1111

Endrizzi JE, Turcotte EL, Kohel RJ (1984) Quantitative genetics, cytology and cytogenetics. In: Kohel RJ, Lewis CF (eds) Cotton. American Society of Agronomy, Madison, pp 81–129

Endrizzi JE, Turcotte EL, Kohel RJ (1985) Genetics, cytology, and evolution of *Gossypium*. Adv Genet 23:271–375

Eshed Y, Zamir D (1995) An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield associated QTL. Genetics 141:1147–1162

Fitter AH, Stickland TR (1991) Architectural analysis of plant root systems 2. Influence of nutrient supply on architecture in contrasting plant species. New Phyto L. 118(3):383–389

Harlan JR, de Wet JMJ (1971) Toward a rational classification of cultivated plants. Taxon 20:509

Iqbal MJ, Reddy OUK, El-Zik KM, Pepper AE (2001) A genetic bottleneck in the "evolution under domestication" of upland cotton *Gossypium hirsutum* L. examined using DNA fingerprinting. Theor Appl Genet 103:547–554

Jenkins JN, McCarty JC Jr, Campbell BT, Hayes RW, Wu J, Saha S, Stelly DM (2017a) Genetic effects of chromosomes 1, 4, and 18 from three tetraploid *Gossypium* species in top crosses with five elite cultivars. Crop Sci 57:1338–1346

Jenkins JN, McCarty JC Jr, Campbell BT, Hayes RW, Wu J, Saha S, Stelly DM (2017) Genotypic comparisons of chromosomes 01, 04 and 18 from three tetraploid species of *Gossypium* in top crosses with five elite cultivars of *G. hirsutum* L. Euphytica 213:107

Kramer PJ (1983) Water deficits and plant growth. In: Kramer PJ (ed) Water relations of plants. Academic Press, New York, pp 342–389

Krieg Pettigrew DR, Sung FJM (1986) Source-sink relationships as affected by water stress. In: Mauney JR, Stewart JM (eds) Cotton physiology. The Cotton Foundation, Memphis, TN, pp 73–78

Lynch JP (1995) Root architecture and plant productivity. Plant Physiol 109:7–13. https://doi.org/10.1104/pp.109.1

Lynch JP (2005) Root architecture and nutrient acquisition. In: Bassiri RH (ed) Nutrient acquisition by plants. An ecological perspective, vol 181. Springer, Berlin, pp 147–183

Lynch JP, Brown KM (2012) New roots for agriculture: exploiting the root phenome. Philos Trans R Soc Lond Ser B Biol Sci 367(1595):1598–1604. https://doi.org/10.1098/rstb.2011.0243

Malik RS, Dhankar JS, Turner NC (1979) Influence of soil water deficits on root growth of cotton seedlings. Plant Soil 53:109–115

McCarty JC, Wu J, Jenkins JN (2006) Genetic diversity for agronomic and fiber traits in day-neutral accessions derived from primitive cotton germplasm. Euphytica 148:283–293. https://doi.org/10.1007/s10681-005-9027-x

McMichael BL, Quisenberry JE (1991) Genetic variation for root-shoot relationship among cotton germplasm. Environ Exp Bot 31:461–470

McMichael BL, Oosterhuis DM, Zak JC, Beyrouty CA (2010) Growth and development of root systems. In: Stewart JM, Oosterhuis DM, Heitholt J, Mauney JR (eds) Physiology of cotton. Springer, New York, pp 57–71

Meyer JR, Meyer VG (1961) Origin and inheritance of nectariless cotton. Crop Sci 1:167–169

Meyer VG, Meredith WR Jr (1978) New germplasm for crossing upland cotton (*Gossypium hirsutum* L.) with *G. tomentosum*. J Hered 69:183–187

Moore JH (1956) Cotton Breeding in the Old South. Agric Hist 30(3):96

Ogbonnaya CI, Roy-Macauley H, Nwalozie MC, Annerose DJM (1997) Physical and histochemical properties of Kenaf (*Hibiscus cannabinus* L.) grown under water deficit on a sandy soil. Ind Crops Prod. https://doi.org/10.1016/S0926-6690(97)00034-4

O'Mara JG (1940) Cytogenetic studies on Triticale. I. A method for determining the effects of individual Secale chromosomes on Triticum. Genetics 25:401–408

Pace PF, Cralle HT, El-Halawany SHM, Cothren JT, Senseman SA (1999) Drought-induced changes in shoot and root growth of young cotton plants. J Cotton Sci 3:183–187

Percival AE, Wendel JF, Stewart JM (1999) Taxonomy and germplasm resources. In: Smith WC, Cothren JT (eds) Cotton: origin, history, technology, and production. Wiley, New York, pp 33–63

Pettigrew WT (2004) Physiological consequences of moisture deficit stress in cotton. Crop Sci 44:1265–1272

Pickersgill B, Spencer CHB, Barrett SCH, Andrade-Lima DD (1975) Wild cotton in northeast Brazil. Biotropica 7:42–54

Pinto RS, Reynolds MP (2015) Common genetic basis for canopy temperature depression under heat and drought stress associated with optimized root distribution in bread wheat. Theor Appl Genet 128:575–585. https://doi.org/10.1007/s00122-015-2453-9

Reddy KR, Bheemanahalli R, Saha S, Singh K, Lokhande SB, Gajanayake B, Read JJ, Jenkins JN, Raska DA, De Santiago LM, Hulse-Kemp AM, Vaughn RN, Stelly DM (2020) High-temperature and drought-resilience traits among interspecific chromosome substitution lines for genetic improvement of upland cotton. Plants 9(12–1747):1–22. https://doi.org/10.3390/plants9121747

Ristova D, Busch W (2014) Natural variation of root traits. Plant Physiol 166:518–527. https://doi.org/10.1104/pp.114.244749

Saha S, Jenkins JN, McCarty JC, Hayes RW, Stelly DM, Campbell BT (2018) Registration of two CS-B17-derived Upland cotton RILs with improved fiber micronaire. J Plant Reg 12(1):97–100

Saha S, Raska DA, Stelly DM (2006) Upland (*Gossypium hirsutum*) × Hawaiian cotton (*G. tomentosum*) F1 hybrid hypoaneuploid chromosome substitution series Upland (*Gossypium hirsutum* L) × Hawaiian cotton (*G. tomentosum* Nutt. ex Seem.) F1 hybrid hypoaneuploid chromosome substitution series. J Cotton Sci 10:263–272 (**ISSN: 1524-3303**)

Saha S, Raska DA, Stelly DM, Manchali S, Gutierrez OA (2013a) Hypoaneuploid chromosome substitution $F_1$ hybrids of *Gossypium hirsutum* L. × *G. mustelinum* Miers ex Watt. J Cotton Sci 17:102–114

Saha S, Stelly DM, Makamov AK, Ayubov MS, Raska D, Gutierrez OA, Manchali S, Jenkins JN, Deng D, Abdurakhmonov IY (2015) Molecular confirmation of *Gossypium hirsutum* chromosome substitution lines. Euphytica 205(2):459–473

Saha S, Wu J, Jenkins JN, McCarty JC, Campbell BT, Hayes RW, Stelly DM (2017) Tri-species shuffling of chromosomes to study the effects on fiber traits using chromosome substitution lines. Crop Sci 57:1211–1226

Saha S, Wu J, Jenkins JN, McCarty JC, Stelly DM (2013b) Interspecific chromosomal effects on agronomic traits in *Gossypium hirsutum* by AD analysis using intermated *G. barbadense* chromosome substitution lines. Theor Appl Genet 126:109–117. https://doi.org/10.1007/s00122-012-1965-9

Saha S, Wu J, Jenkins JN, McCarty JC, Gutierrez OA, Stelly DM, Percy RG, Raska DA (2004) Effect of chromosome substitutions from *Gossypium barbadense* L. 3-79 into *G. hirsutum* L. TM-1 on agronomic and fiber traits. J Cotton Sci 8:162–169 (**ISSN: 1439-0523**)

Sears ER (1952) Misdivision of univalents in common wheat. Chromosoma 4:535–550

Senchina DS, Alvarez I, Cronn RC, Liu B, Rong J, Noyes RD, Paterson AH, Wing RA, Wilkins TA, Wendel JF (2003) Rate variation among nuclear genes and the age of

polyploidy in *Gossypium*. Mol Biol Evol 20(4):633–643. https://doi.org/10.1093/molbev/msg065 (**PMID:12679546**)

Singh B, Norvell E, Wijewardana C, Wallace T, Chastain D, Reddy KR (2018) Assessing morphological characteristics of elite cotton lines from different breeding programmes for low temperature and drought tolerance. J Agron Crop Sci 204(5):467–476. https://doi.org/10.1111/jac.12276

Skovsted A (1934) Cytological studies in Cotton. II. Two interspecific hybrids between Asiatic and new world cottons. J Genet 28:407–424

Song W, Wang M, Su W, Lu Q, Xiao X, Cai J, Zhang Z, Li S, Li P, Gong J, Gong W, Shang H, Liu A, Li J, Chen T, Ge Q, Shi Y, Yuan Y (2017) Genetic and phenotypic effects of chromosome segments introgressed from *Gossypium barbadense* into *Gossypium hirsutum*. PLoS ONE 12(9):1–17 . e0184882. https://doi.org/10.1371/journal.pone.0184882

Stelly D, Saha S, Raska D, Jenkins JN, McCarty J, Gutierrez O (2005) Registration of 17 Upland (*Gossypium hirsutum*) germplasm lines disomic for different *G. barbadense* chromosome or arm substitutions. Crop Sci 45(6):2663–2665 (**ISSN 1435-0653**)

Todesco M, Owens GL, Bercovich N, Légaré JS, Soudi S, Burge DO, Huang K, Ostevik KL, Drummond EBM, Imerovski I, Lande K, Pascual-Robles MA, Nanavati M, Jahani M, Cheung W, Staton SE, Muños S, Nielsen R, Donovan LA, Burke JM, Yeaman S, Rieseberg LH (2020) Massive haplotypes underlie ecotypic differentiation in sunflowers. Nature 584(7822):602–607

Udall JA, Wendel JF (2006) Polyploidy and crop improvement. Crop Sci 46:3–14. https://doi.org/10.2135/cropsci2006.07.0489tpg

Van Esbroeck GA, Bowman DT, May OL, Calhoun DS (1999) Genetic similarity indices for ancestral cotton cultivars and their impact on genetic diversity estimates of modern cultivars. Crop Sci 39:323–328

Wehrhahn C, R.W. Allard RW, (1965) Detection and measurement of the effects of individual genes involved in the inheritance of a quantitative character in wheat. Genetics 51:109–119

Wendel JF, Cronn RC (2003) Polyploidy and the evolutionary history of cotton. Adv Agron 78:139–186

Wilson RF, Burke JJ, Quisenberry JE (1987) Plant morphological and biochemical responses to field water deficits. II. Responses of leaf glycerolipid composition in cotton. Plant Physiol 84:251–254

Xie Q, Fernando KMC, Mayes S, Sparkes DL (2017) Identifying seedling root architectural traits associated with yield and yield components in wheat. Ann Bot 119:1115–1129. https://doi.org/10.1093/aob/mcx001

Yik CP, Birchfield W (1984) Resistant germplasm in *Gossypium* species and related plants to *Rotylenchulus reniformis*. J Nematol 16:146–153

Zhang J, Yu J, Pei W, Li X, Said J, Song M, Sanogo S (2015) Genetic analysis of Verticillium wilt resistance in a backcross inbred line population and a meta-analysis of quantitative trait loci for disease resistance in cotton. BMC Genom 16:577. https://doi.org/10.1186/s12864-015-1682-2