Identification of Stable and Multiple Environment Interaction QTLs and Candidate Genes for Fiber Productive Traits Under Irrigated and Water Stress Conditions Using Intraspecific RILs of \textit{Gossypium hirsutum} var. MCU5 X TCH1218

Narayanan Manikanda Boopathi\textsuperscript{1*†}, Gopal Ji Tiwari\textsuperscript{2†}, Satya Narayan Jena\textsuperscript{2*†}, Kemparaj Nandhini\textsuperscript{3}, V. K. I. Sri Subalakhshmi\textsuperscript{3}, Pilla Shyamala\textsuperscript{1}, Babita Joshi\textsuperscript{24}, Nallathambi Premalatha\textsuperscript{3} and S. Rajeswari\textsuperscript{3}

\textsuperscript{1}Department of Plant Biotechnology, CPMB&B, Tamil Nadu Agricultural University, Coimbatore, India, \textsuperscript{2}Plant Molecular Genetics Laboratory, CSIR-National Botanical Research Institute, Lucknow, India, \textsuperscript{3}Department of Cotton, CPBG, Tamil Nadu Agricultural University, Coimbatore, India, \textsuperscript{4}Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India

Cotton productivity under water-stressed conditions is controlled by multiple quantitative trait loci (QTL). Enhancement of these productivity traits under water deficit stress is crucial for the genetic improvement of upland cotton, \textit{Gossypium hirsutum}. In the present study, we constructed a genetic map with 504 single nucleotide polymorphisms (SNPs) covering a total span length of 4,416 cM with an average inter-marker distance of 8.76 cM. A total of 181 intra-specific recombinant inbred lines (RILs) were derived from a cross between \textit{G. hirsutum} var. MCU5 and TCH1218.

In the SNP genotyping, a large number of SNPs showed either >20% missing data, duplication, or segregation distortion. However, the mapped SNPs of this study showed collinearity with the physical map of the reference genome (\textit{G. hirsutum} var. TM-1), indicating that there was no chromosomal rearrangement within the studied mapping population. RILs were evaluated under multi-environments and seasons for which the phenotypic data were acquired. A total of 53 QTL controlling plant height (PH), number of sympodial branches, boll number (BN), and boll weight (BW) were dissected by QTL analysis under irrigated and water stress conditions. Additionally, it was found that nine QTL hot spots not only co-localized for more than one investigated trait but were also stable with major QTL, i.e., with > 10% of phenotypic variation. One QTL hotspot on chromosome 22 flanked by AX-182254626–AX-182264770 with a span length of 89.4 cM co-localized with seven major and stable QTL linked to a number
INTRODUCTION

Owing to its affordability, appearance, and natural comfort in apparel, cotton (*Gossypium* spp.) is the most preferred economic crop globally. More than 80% of the top ten cotton cultivating zones are in developing countries, and the cotton export value of these zones exceeded USD 30 billion in 2017 (accessed on 24 December, 21). Thus, cotton promotes economic development and offers a key source of livelihood for millions of farmers in the world’s top three cotton exporters, North America, Central Asia, and West Africa (Jans et al., 2021).

India had a total cotton area of 13.48 million ha, accounting for 42.07% of the world’s cotton area (32.04 million ha). Particularly, the textile belt of South India alone has more than 2,000 large and small cotton mills that manufacture blended yarns and cotton with the export revenue of more than USD 3,757 million, and it is predicted that the expected flow of investment in India in various textile activities during next 5 years will be USD 14,279 million (accessed on 24 December, 21). Consequently, there is a continuous demand for cotton products in India and across the world, but fiber yield is challenged by the unpredictable climate change in all the cotton-producing countries (Niu et al., 2018). For example, more than 60% of cotton farmed in India is rainfed (Gutierrez et al., 2015), where erratic rainfall patterns severely limit lint output. Similarly, despite having one of the most efficient cotton industries in the world in the terms of water use, a recent Australian Bureau of Statistics report released on 14 May 2021 stated that climatic caprices significantly reduced Australia’s cotton area (accessed on 24 December, 21). Cotton productivity under these rainfed conditions largely depends on the timing, distribution, and quantity of monsoon rains at the different growth phases of cotton, and hence water stress limits lint productivity with different magnitudes, depending on its occurrence and severity. In general, high temperatures, below-average rainfall, and limited water availability in rainfed cotton have severely limited cotton output in recent years. There is an urgent need to develop a cotton cultivar that is more productive in water-stressed settings.

Among the six allotetraploid species (*2n = 4x, AADD*, *Gossypium hirsutum* L., contributes >90% of the worldwide cotton production (Shim et al., 2018). Conventional breeding efforts to improve the fiber yield in this species under water-limited environments have shown slow progress as these traits are more complex and governed by multiple genetic and environmental factors (Tan et al., 2018). Furthermore, the negative relationship between fiber quality and yield component characteristics impedes the simultaneous enhancement of yield and fiber quality traits in *G. hirsutum* (Abdelraheem et al., 2018). Thus, before attempting to improve these complex traits under rainfed conditions, the genetics underlying fiber production and quality attributes (such as, the pleiotropic impact and functional genes that govern trait) must be dissected by quantitative trait loci (QTL) mapping.

Among the target traits to be dissected, it is imperative to dissect the QTL for optimum plant height (PH), zero or minimum monopodial branch length, number of bolls, and boll weight, as these traits are key in introducing the mechanical harvesting, which is the need of the hour (owing to the permanent paucity of labor for boll picking). The cotton PH of 80–120 cm is strongly related to the yield under mechanical harvesting (Yan et al., 2019), and the desired form of a cotton cultivar is a compact architecture with short fruit internodes and tightly packed cotton bolls. Furthermore, such cotton designs stimulate the use of less plant growth regulators, are resistant to lodging, and adapt themselves to the dense planting and automated harvesting (Wen et al., 2021). Advances in different molecular marker tools and strategies used to dissect the complex traits genetically have enabled fast-track molecular breeding efforts in several crops even in cotton (Boopathi, 2020). A large number of QTL associated with agronomic, yield, and particular fiber quality traits were identified, and available as valuable databases on cotton QTL, such as CottonQTLdb (Said et al., 2015) and CottonGen (Yu et al., 2021). Such attempts may allow the identification of stable QTLs’ co-localization across *G. hirsutum* genetic backgrounds, validating the found QTL for future use in marker-assisted selection (MAS) for the efficient and simultaneous augmentation of yield and fiber quality traits. The use of MAS to pyramid several QTLs that affect numerous desired target characteristics enhanced the selection efficiency among breeding populations with varied genetic origins. Recent studies (Abdelraheem et al., 2021; Guo et al., 2021; Wang H. et al., 2021; Wang N. et al., 2021; Zhu et al., 2021) demonstrated the necessity of dissecting complicated fiber traits using QTL mapping and the potential of MAS in the cotton breeding program.

Keywords: drought, productive traits, genetic map, SNP, QTL, upland cotton, intra-specific cross
Though India has a long history of cotton breeding efforts, only a few reports have employed molecular markers to analyze the genetic purity of seed lots (Selvakumar et al., 2010) to examine the level of genetic diversity and linkage disequilibrium in cotton subpopulations (Jena et al., 2011) and to genetically dissect cotton fiber traits (Boopathi et al., 2015; Kumar et al., 2019; Shukla et al., 2021). An array of QTLs, genes, and gene products are specifically involved in drought response in cotton (Deeba et al., 2012), but identifying which are most useful for cotton breeding in the genetic improvement of drought tolerance in the regional breeding program has remained a significant challenge, owing to the unique characteristics of the target population of environments under rainfed conditions. As a result, anticipating the future uses of MAS in cotton, it is critical to begin the QTL mapping utilizing various parental lines and introgression of important and stable QTL for target characteristics by MAS for the genetic improvement of cotton in water-limited settings.

While selecting the donors for the QTL mapping of drought tolerance and fiber yield/quality traits, *G. arboreum* and *G. barbadense*, respectively, would be the best choice owing to their superior trait values. However, the transfer of other undesirable agronomic traits through linkage drags and problems associated with cross ability, return toward the genotype of one parent (segregation distortion), and the suppression of recombination greatly limit the progress and made the interspecific crossing program as a challenging task (Boopathi and Hoffmann, 2016). As a result, breeding initiatives are mostly focused on intraspecific *G. hirsutum* cross combinations to develop superior lint production and fiber quality attributes at a lower cost. On the other hand, the lack of polymorphism created across intraspecific *G. hirsutum* lines using breeder-friendly second-generation molecular markers, such as simple sequence repeats (SSRs) impedes the efficiency of QTL mapping (Shen et al., 2005). Through these efforts, even a consensus map derived from several intra-specific maps from different mapping populations encompassed just 31% of the cotton genome (Ulloa et al., 2005).

Alternatively, recent advancements in high throughput genotyping systems with enhanced the effectiveness in producing polymorphism among closely related individuals have made the single nucleotide polymorphism (SNP) a widely used and popular marker in plant translational quantitative genetics (Hulse-Kemp et al., 2015). With the improvement of next-generation sequencing technology, numerous methodologies have been used to identify a significant number of polymorphic SNPs in cotton, which might be useful for high-density mapping and more efficient QTL analysis. As a result, it is particularly desirable to use third-generation markers, such as SNPs to improve the detection of polymorphic loci between closely related *G. hirsutum* parental lines.

Considering the above, this present study was designed to identify QTL for various yield traits under irrigated and water-limited environments in *G. hirsutum* using SNP markers. We selected *G. hirsutum* parents, MCU5 and TCH1218, differing distinctly for drought tolerance and various other traits to develop the recombinant inbred lines (RILs) and evaluated them under irrigated and water-stressed field conditions for the identification of fiber yield related QTLs by employing 50K SNP custom array with Axiom technology (Affymetrix).

**MATERIALS AND METHODS**

**Plant Materials**

Recombinant inbred lines derived from the intra-specific cross between *G. hirsutum* var. MCUS and TCH1218 were used in this study. The F1 was made at Tamil Nadu Agricultural University (TNAU), Coimbatore, in the year 2006 and forwarded to F11 using the single seed descent method. In total, 181 RILs from the F10 generation were employed in the construction of the genetic map. Data on PH (cm), number of sympodial branches, number of bolls, and boll weight (g) obtained from the phenotypic evaluation of F7 to F11 of those same RILs were used for QTL analysis. The female parent MCU5 is a multiple cross derivative cultivated widely in South India as it possesses medium staple fiber (29 mm) and can be spun up to 70s. Besides, it has a 34% ginning outturn and produces 1,850 kg seed cotton yield per ha; however, it cannot withstand water stress at the flowering phase (Boopathi et al., 2014). On the other hand, the male parent, TCH1218, has relatively better drought tolerance but low fiber yield and quality characters than MCU5 (Boopathi et al., 2014). Our historical breeding effort has shown that MCU 5 has long-staple cotton with a good yield, while TCH 1218 is a good combiner.

**Phenotypic Data**

All the plant materials (181 RILs and two parental lines) were evaluated in randomized block design with two replications at different locations and seasons (the details of testing environments are provided in Table 1). The plant-to-plant distance was 45 cm, while the row-to-row distance was 90 cm, and 13 plants were maintained in each row. Regular crop husbandry measures were followed to ensure a healthy crop. To collect phenotypic data, three plants were selected randomly per replication on 115 days after sowing (DAS), and information was collected on different yield parameters from each experiment as detailed in Table 1. RILs were evaluated under field water stress conditions in two different water-stress environments: one with managed water stress (by withholding irrigation water after 45th day) at Maize Research Station (MRS), TNAU, Vagarai, and another under purely rainfed conditions at Cotton Research Station (CRS), TNAU, Veppanthattai.

Regular agricultural practices were used to control weeds, diseases, pests, and fertilizers were applied at sufficient levels to ensure that yield potential was not limited by any factor other than water. At MRS (T2 in Table 1), water stress was imposed by withholding water from 45th DAS, whereas the irrigated control plots received water at routine intervals. At CRS T3 and T7 (Table 1) and ARS T6 (Table 1), the RILs were grown purely under rainfed conditions, and there was no rain after 51 and 57 DAS at CRS during 2012 and 2018, respectively, and 58 DAS at ARS during 2018.
Standard descriptive statistical analysis to examine the significance of the difference in the investigated traits between two parents and among population and the estimation of Pearson’s correlation coefficient was done using Calculator Soup<sup>8,4</sup> and Minitab<sup>9</sup> respectively.

### DNA Isolation and Genotyping

Genomic DNA was extracted from the young leaves of mapping population and parents using a modified cetyltrimethylammonium bromide (CTAB) method (Shukla et al., 2021). SNP genotyping was performed using the Gene Titan Multi-channel instrument (Thermo Scientific) facility at M/s Imperial life Sciences, New Delhi, with the Cotton 50K SNP array chip. DNA samples were first processed to the DNA Isolation and Genotyping

The arrays were scanned with the Gene Titan MC and automated allele calling, and the quality assessment of called genotypes was done with Genotyping Console™ Software (GTC) with a new Axiom Genotyping Algorithm v1 (Axiom GT1). Consequently, the raw hybridization intensity data were processed for clustering and genotype calling with Affymetrix® GTC (v4.2), and those data with Dish Quality Control (DQC) value <0.82 and call rate <0.97 were excluded from the further genotyping analysis. To finish, GTC was further processed using APT (v1.19.0) and classified those SNPs into six major classes. The raw genotyping data were again filtered by selecting only those co-dominantly segregating SNPs with <20% missing genotyping data (Hulse-Kemp et al., 2015). Subsequently, the genotyping data were transformed into a mapping data format. The probe sequences (10–20 bp overlapping to SNP) were also mapped on reference genome <i>G. hirsutum</i> cultivar TM-1 (ASM98774v1), and the SNPs were assigned to the particular chromosomes. The same data were used as anchoring information while preparing the input file for the mapping software while performing chromosome assignments.

### Genetic Map Construction

A total of 522 polymorphic and segregating SNP markers were used to construct the genetic map using ‘lclmapping v4.0’ (Meng et al., 2015), and the genetic map distance (centiMorgan, cM) was calculated using the Kosambi mapping function. The threshold for logarithm of odds (LOD) score was fixed

---

<sup>8</sup>https://www.calculatorsoup.com/calculators/statistics/descriptivestatistics.php
<sup>9</sup>https://www.minitab.com/en-us/support/downloads/
at five, revealing 26 linkage groups. The exact ordering of the SNPs across chromosomes was done using RECORD, and Rippling was also executed for fine-tuning the order of markers by the sum of adjacent recombination frequencies (COUNT) with a window size of five. If there were more than three consecutive adjacent markers in the genetic map with a significance level $0.001 < p < 0.05$, it was taken as segregation distortion, and its distribution on the map was also analyzed (Coulton et al., 2020).

Quantitative Trait Loci Analysis for Production Traits

Composite interval mapping (CIM) with forward regression, in a window size of 10 cM and 5 background control markers at a walking speed of 2 cM, was employed for QTL analysis using Windows QTL Cartographer 2.5 (Wang et al., 2012). The threshold LOD of each trait QTL was determined with the trait threshold calculated with a permutation of 10,000, and multiple QTL peaks detected within 10 cM of each other were regarded as a single QTL. To calculate the effects of additive QTL in multiple environments, the MET functional module of QTL IciMapping v4.0 was employed by combined mapping analysis under multi-environment with 3.00 permutation tests’ LOD cut-off values.

Collinearity and Recombination Hotspot Analysis

All the SNP markers used to construct the genetic map reported in this study were aligned to the physical map of the upland cotton genome (TM-1 Genome NAU-NBI Assembly v1.1 and Annotation v1.1 database) through Basic Local Alignment Search Tool (BLAST). CIRCOS 0.66 with default parameters was employed to compare each investigated SNP marker’s genetic and physical positions collinearity. The recombination hotspot

| Trait | Trial* | MCU5 Mean | TCH1218 Mean | RILs | Broad sense heritability |
|-------|--------|-----------|--------------|------|-------------------------|
|       |        |           |              | Mean | SD   | CV%    | Minimum | Maximum | Skewness γ1 | Kurtosis β2 |
| Plant Height | T1      | 116.99    | 124.66       | 100.03 | 27.57 | 0.27  | 26.00   | 167.33  | −1.51     | 7.08  | 0.58 |
|         | T2      | 92.16     | 101.00       | 88.13  | 31.07 | 0.35  | 10.00   | 159.33  | −1.35     | 5.73  | 0.55 |
|         | T3      | 42.00     | 60.58        | 95.02  | 43.07 | 0.45  | 21.00   | 157.00  | −1.01     | 3.00  | 0.45 |
|         | T4      | 125.00    | 132.00       | 106.55 | 27.89 | 0.26  | 66.33   | 161.00  | −1.57     | 8.11  | 0.54 |
|         | T5      | 102.95    | 100.25       | 112.82 | 29.52 | 0.26  | 10.10   | 143.19  | −2.85     | 10.85 | 0.45 |
|         | T6      | 83.00     | 87.00        | 84.03  | 13.08 | 0.15  | 46.50   | 122.83  | −0.02     | 3.10  | 0.43 |
|         | T7      | 107.50    | 137.00       | 117.55 | 20.75 | 0.17  | 69.33   | 185.00  | 0.45      | 2.84  | 0.58 |
|         | T8      | 109.90    | 123.40       | 120.32 | 10.47 | 0.08  | 92.00   | 143.9   | −0.23     | 2.48  | 0.41 |
| Number of Sympodia | T1      | 11.87     | 10.25        | 14.23  | 6.04  | 0.42  | 2.00    | 28.5    | −0.30     | 2.72  | 0.50 |
|         | T2      | 15.49     | 39.99        | 17.30  | 7.42  | 0.42  | 6.5     | 51.33   | −0.05     | 5.22  | 0.55 |
|         | T3      | 1.83      | 3.66         | 5.21   | 2.48  | 0.47  | 1.00    | 12.00   | −0.13     | 3.21  | 0.44 |
|         | T4      | 4.00      | 3.66         | 2.23   | 0.85  | 0.38  | 1.00    | 5.00    | 0.50      | 3.25  | 0.48 |
|         | T5      | 11.40     | 11.40        | 15.32  | 8.11  | 0.52  | 1.00    | 9.05    | 0.78      | 9.93  | 0.47 |
|         | T6      | 7.5       | 10.58        | 10.57  | 1.71  | 0.16  | 6.83    | 19.24   | 0.62      | 4.93  | 0.58 |
|         | T7      | 9.33      | 19.67        | 14.34  | 3.18  | 0.22  | 5.67    | 24.00   | 0.18      | 2.76  | 0.47 |
|         | T8      | 10.90     | 12.90        | 15.21  | 1.51  | 0.09  | 11.2    | 18.8    | −0.27     | 3.09  | 0.44 |
| Number of Bolls per Plant | T1      | 10.80     | 9.10         | 6.80   | 3.70  | 0.54  | 1.33    | 25.00   | 1.37      | 6.96  | 0.55 |
|         | T2      | 2.83      | 7.33         | 4.39   | 2.44  | 0.55  | 0.66    | 11.33   | 0.26      | 2.86  | 0.58 |
|         | T3      | 2.66      | 5.16         | 13.03  | 6.62  | 0.50  | 2.50    | 27.00   | −0.48     | 2.85  | 0.41 |
|         | T4      | 16.67     | 15.33        | 17.61  | 5.99  | 0.34  | 7.67    | 48.67   | 0.18      | 7.84  | 0.48 |
|         | T5      | 24.00     | 21.00        | 21.99  | 4.60  | 0.20  | 11.0    | 30.00   | −0.52     | 2.48  | 0.49 |
|         | T6      | 11.92     | 16.25        | 19.03  | 4.67  | 0.24  | 8.67    | 43.17   | 0.64      | 5.26  | 0.45 |
|         | T7      | 17.00     | 40.33        | 23.06  | 6.70  | 0.29  | 8.33    | 45.00   | 0.64      | 3.45  | 0.58 |
|         | T8      | 13.60     | 12.60        | 15.17  | 2.95  | 0.19  | 8.80    | 22.97   | 0.19      | 2.78  | 0.57 |
| Boll Weight | T3      | 2.65      | 5.75         | 5.19   | 2.03  | 0.39  | 3.00    | 8.80    | −1.55     | 4.93  | 0.44 |
|         | T4      | 4.73      | 5.00         | 4.19   | 1.14  | 0.27  | 2.50    | 10.67   | −0.55     | 12.49 | 0.44 |
|         | T5      | 4.00      | 3.20         | 2.93   | 0.67  | 0.22  | 1.20    | 4.30    | −0.002    | 2.35  | 0.58 |
|         | T6      | 4.65      | 4.90         | 3.78   | 0.75  | 0.19  | 1.67    | 5.60    | 0.12      | 2.92  | 0.56 |
|         | T7      | 3.47      | 4.51         | 3.67   | 0.63  | 0.17  | 2.15    | 6.13    | 0.12      | 3.25  | 0.41 |
|         | T8      | 4.45      | 4.46         | 4.63   | 0.60  | 0.13  | 2.47    | 6.08    | −0.34     | 3.83  | 0.59 |

*As provided in Table 1; SD, standard deviation; CV%, coefficient of variation.
(RH) was estimated by inferring the recombination rate of investigated SNPs. If the genetic distance between adjacent SNPs was higher than 20 cm/Megabase, the genomic region between those markers was viewed as RH (Zhang J. et al., 2015).

**Major and Stable Quantitative Trait Loci and Identification of Putative Candidate Genes**

Quantitative trait loci co-localized at least in two trials with $\geq 10.0$ R² for at least one trait were depicted as major and stable QTL, and they were selected to scan for candidate genes. In general, a line with at least 10% higher variation than the parents for the target traits is considered an improved line; hence $\geq 10.0$ R² is depicted as a major QTL in this study. The sequences of SNP markers flanking the confidence intervals (CIs) of the QTL were aligned back to the physical sequence of the upland cotton genome database (Zhang J. et al., 2015). Based on the position of these flanking markers, all the genes within the target QTL were identified as candidate genes. Gene ontology (GO) grouping was employed to categorize the identified candidate genes, and pathways correlated to the candidate genes were also discovered using Kyoto Encyclopaedia of Genes and Genomes (KEGG) analysis (Charmpi and Ycart, 2015).

**RESULTS**

**Phenotypic Variation for Parents and Recombinant Inbred Lines**

The differences noticed for the investigated phenotypic traits between parents and the variation among the investigated RILs are provided in Table 2. As expected, the elite parent MCU5 performed well under the irrigated conditions, whereas TCH1218 had produced substantially good yield under both managed water stress and rainfed conditions by having higher values in the relative performance of investigated traits (significant at the 0.05 probability level) (Table 2). Transgressive segregation was observed in the RIL population for all the investigated traits under irrigated and water-stressed conditions. However, PH and boll weight have shown substantially skewed distribution as their skewness values were greater than $+1$ or lower than $-1$ (Table 2). All possible kinds of distribution, such as fairly symmetrical (skewness $-0.5$ and 0.5), moderately skewed skewness ($-1$ and $-0.5$ or between 0.5 and 1), and highly skewed (skewness less than $-1$ or greater than 1) distribution of the traits examined in this study were noticed (Table 2). As the calculated kurtosis values for all the investigated traits were greater than zero (Table 2), it can be concluded that each trait distribution has a heavier tail (leptokurtic distribution). Substantial variations (minimum and maximum values of each trait) due to genotype differences were reported for all the investigated traits among the RILs. However, the relative proportions of variance varied from one trait to another, and there was low to moderate heritability noticed for the investigated traits (Table 2). The correlation between investigated traits measured from all the trials (T1–T8) was also evaluated (Table 3). Although not in all the trials, the majority of correlations have shown a highly significant positive correlation between yield (BN and BW) and growth parameters (PH and SYM) both under irrigated and water-stressed environments, indicating that genetic improvement BN and BW would likely be accompanied by the improvement of PH and SYM. However, significant negative correlations were also observed (Table 3), which showed that the simultaneous improvement of investigated traits could not be done for all the cotton-growing regions. In general, the correlation between the investigated traits increased dramatically under water-stressed conditions (T3 and T7); however, mixed trends were noticed in the rainfed environment, T6 (Table 3). On the other hand, under managed water stress conditions (T2), a reduced level of correlation between the measured traits was noticed when compared with the traits measured from irrigated conditions (T1) (Table 3).

**Genetic Map Construction With Cotton 50K Single Nucleotide Polymorphisms Array**

Affymetrix’s Axiom custom designed Cotton 50K SNP array (unpublished) was used to the SNP typing of both parents and 181 RILs. This resulted a total of 2,457 polymorphic SNPs between the two parents and those polymorphic markers were analyzed for their segregation in the mapping population. The initial attempt of binning these 2,457 loci by excluding those SNPs with more than 20% missing data (totally there were 17 such markers) resulted in 2,440 polymorphic SNPs, which were further analyzed with $\chi^2$ test to determine if the allele frequency was deviated from the expected segregation ratio (1:1). Among the selected 2,440 polymorphic loci, 1,913 loci showed segregation distortion ($p < 0.05$), and only 527 loci followed the perfect Mendelian ratio. It was also noticed that among 1,913 distorted loci, 1,186 were in favor of TCH1218 alleles, and 727 loci were biased toward MCU5 alleles. Further, it has been found that among the 527 loci that were segregated in the RILs, and five were found to be duplicated.

Therefore, only 522 polymorphic SNPs were taken for the genetic mapping, out of which 504 were mapped on 26 AD linkage groups (LGs) or cotton chromosomes (as 18 SNPs were unlinked), and the resultant genetic map represented the genetic span length of 4,416 cM. The At sub-genome had 204 SNP markers covering 1,830.03 cM with an average genetic distance of 8.97 cM between adjacent loci, while the Dt sub-genome possessed 300 SNP markers that span 2,585.97 cM with an average of 8.62 cM between successive loci. Overall, the average marker distance among the 26 chromosomes was 8.76 cM (Table 4). The longest and most dense chromosome found in this study was chromosome 18 (395.48 cM) with 79 markers, and the shortest and most sparse was chromosome 02 (85.03 cM) with 2 markers (Table 4 and Figure 1). In the genetic map of chromosome 2, there were only two markers mapped, and 19 markers showing segregated distorted (SD) were removed from the map. Thus, there is a huge gap in chromosome 2. Besides chromosome 2, other small gaps were recorded in this genetic map due to the removing of SD markers. Therefore, the marker interval in each chromosome was ranged from 7.9 to 85.0 cM (Table 4 and Figure 1).
chromosomal rearrangements. The sequence-based reference genome of 1,027.9 Mb corresponded to our SNP based genetic map of 4,416 cM and our genetic map represented 53.13% of the total length of the sequence-based physical map (Table 5). A good collinearity was revealed by all the linkage groups with the physical reference map (Figure 2). The coverage of individual chromosome in the constructed genetic linkage map of our study ranged from 0.64 to 99.11% of the physical reference map (Table 5). Chromosome 26 was least covered whereas, chromosome 3 was comprehensively covered. Though least covered chromosomes in the genetic map would provide little information during QTL analysis, it would be useful to have such preliminary data to proceed further.

**Quantitative Trait Loci for Productive Traits Under Irrigated and Water-Stress Conditions**

As the variation found in the investigated traits was normally distributed (Table 2), it was predicted that multiple genes might control these traits. There was low to moderate heritability noticed for the investigated traits (Table 2). Thus, higher similarity between parents and offspring for the investigated trait was not expected making the QTL study more complex, and interesting for significant environmental influence. In such context, the identification of co-localized QTLs across the seasons and environments was the only means of solid proof for the presence of QTL linked to the investigated trait. A total of 53 QTLs for four investigated traits identified in this study, and among them, 21 and 32 were found to be At sub-genome- and Dt sub-genome-specific, respectively (Supplementary Table 1). The LOD values of all the identified QTLs ranged from 2.0 to 18.5. There were eight water stress-specific QTL further linked to different investigated traits. Though there are 8 QTL identified for PH (6 under irrigated conditions and two under water stress conditions), only two QTL had shown more than 10% phenotypic variation (Supplementary Table 1). Among them, qPH-c11-1 explained 38.41% phenotypic variation (PVE) and was detected under irrigated conditions; whereas qPH_S-c16-2 was detected under rainfed conditions had 19.39% PVE. A maximum number of 36 QTL were detected for the number of sympodial branches (SYM), and among them, 25 QTL were detected under irrigated conditions, and 11 were detected under managed water stress/rainfed conditions. In these, QTL detected on chromosome 18 (qSYM-c18-2) had maximum phenotypic variation under irrigated conditions (87.83%) with QTL LOD of 8.6 and a QTL observed on chromosome 21 (qSYM_S-c22-2) had accounted for a maximum of phenotypic variation under rainfed conditions (88.01%) for SYM with QTL LOD value of 3.6. Seven QTL were also identified for boll number (BN) in which all the QTL were linked to different investigated traits. Though there are 17 QTL identified for PH (6 under irrigated conditions and 11 under water stress conditions), only two QTL had shown higher LOD values (8.41 and 8.87) and explained 35.16% and 38.62% PVE, respectively (Supplementary Table 1). Among them, qPH_S-c26-2 had accounted for a maximum of 64.9% PVE under irrigated conditions, whereas qPH_S-c18-2 had accounted for a maximum of 99.11% PVE under rainfed conditions. A maximum number of 24 QTL were detected for boll weight (BW) in which all the QTL were linked to different investigated traits. Though there are 17 QTL identified for PH (6 under irrigated conditions and 11 under water stress conditions), only two QTL had shown higher LOD values (8.41 and 8.87) and explained 35.16% and 38.62% PVE, respectively (Supplementary Table 1). Among them, qPH_S-c26-2 had accounted for a maximum of 64.9% PVE under irrigated conditions, whereas qPH_S-c18-2 had accounted for a maximum of 99.11% PVE under rainfed conditions.

**High-Level Synteny, Collinearity, and Recombination Hotspot**

The quality of the genetic linkage map was evaluated using parameters, such as the segregation distortion of mapped markers, gaps in the map, and collinearity between the linkage map constructed in this study and the reference physical map of TM-1. All the 504 polymorphic SNPs found in this study were aligned to the G. hirsutum var. TM-1 reference genome using BWA in Galaxy and a high level of collinearity was shown by the SNP genetic map reported in this study with the physical map of the TM-1. It was revealed that there were no gaps in the map, and among them, 21 and 32 were found to be At sub-genome- and Dt sub-genome-specific, respectively (Supplementary Table 1). The LOD values of all the identified QTLs ranged from 2.0 to 18.5. There were eight water stress-specific QTL further linked to different investigated traits. Though there are 17 QTL identified for PH (6 under irrigated conditions and 11 under water stress conditions), only two QTL had shown higher LOD values (8.41 and 8.87) and explained 35.16% and 38.62% PVE, respectively (Supplementary Table 1). Among them, qPH_S-c26-2 had accounted for a maximum of 64.9% PVE under irrigated conditions, whereas qPH_S-c18-2 had accounted for a maximum of 99.11% PVE under rainfed conditions. A maximum number of 24 QTL were detected for boll number (BN) in which all the QTL were linked to different investigated traits. Though there are 17 QTL identified for PH (6 under irrigated conditions and 11 under water stress conditions), only two QTL had shown higher LOD values (8.41 and 8.87) and explained 35.16% and 38.62% PVE, respectively (Supplementary Table 1). Among them, qPH_S-c26-2 had accounted for a maximum of 64.9% PVE under irrigated conditions, whereas qPH_S-c18-2 had accounted for a maximum of 99.11% PVE under rainfed conditions. A maximum number of 24 QTL were detected for boll weight (BW) in which all the QTL were linked to different investigated traits. Though there are 17 QTL identified for PH (6 under irrigated conditions and 11 under water stress conditions), only two QTL had shown higher LOD values (8.41 and 8.87) and explained 35.16% and 38.62% PVE, respectively (Supplementary Table 1). Among them, qPH_S-c26-2 had accounted for a maximum of 64.9% PVE under irrigated conditions, whereas qPH_S-c18-2 had accounted for a maximum of 99.11% PVE under rainfed conditions.
Twenty QTL had a positive additive effect for the investigated traits suggesting the independent effects of alleles on the trait (Supplementary Table 1). Several QTL were found to have a dominant positive effect (Supplementary Table 1). Among the 53 QTL, 25 QTL were derived from MCU5, which exhibited a positive additive effect, and 28 QTL were derived from TCH1218, which exhibited a negative additive effect. As expected, all the QTL detected under water stress conditions were contributed by the tolerant drought parent, TCH1218 (Supplementary Table 1) except qPH_S-c16-1, qSYM_S-c18-1, and qSYM_S-c18-2, which were contributed by drought susceptible parent, MCU5.

It was found that nine QTL that were not only co-localized for more than one investigated trait but also stable (identified at least in two seasons/locations) and major QTL with >10% PVE for at least one trait that colocalized in that region (Table 6). In addition, the maximum number of seven QTL was clustered on chromosome 22 and a minimum number of two QTL was found on chromosomes 5, 21, and 24 (Table 6 and Figure 1).

Further, joint multi-environmental QTL mapping effort to estimate the impact of QTL × Environment interaction resulted in the identification of 11 additive QTL (Table 7). Eight QTL associated with SYM were identified with LOD ranging from 3.02 to 5.70. The contribution rates of interaction among eight additive and the environment QTL ranged from 0.68 to 3.33%. Two QTL associated with PH were identified with LOD ranging from 3.13 to 3.53. The contribution rates of interaction among two additves and the environment QTLs ranged from 2.16 to 4.20%. A single QTL associated with BN was identified with LOD 4.36 and the contribution rates of interaction was 4.08% (Table 7). However, no QTL for BW was detected in this joint multi-environmental QTL analysis.

Identification of Candidate Genes Within the Quantitative Trait Loci

Large array of candidate genes were identified within the QTL reported in this study (Supplementary Table 2), and they were compared and annotated with Xu et al. (2020), where meta-QTL analysis along with transcriptomic approach utilized for the identification of candidate genes related to fiber quality in upland cotton. Though there were genes specific to abiotic stress responses and productive traits, large numbers of genes identified in this study warrant the use of additional markers to fine map these QTL and identify precise genes involved for the target traits.
traits. For example, for the PH QTL under irrigated conditions, *qPH-c11-1*, there were 32 genes, whereas under water stress conditions, the QTL identified for PH, *qPH_S-c16-2*, possessed 94 candidate genes (Supplementary Table 2). Similarly, the QTL identified for a number of sympodial branches under water-stressed environment, *qSYM_S-c24-1* harbored 169 genes and QTL for BN under irrigated conditions, *qBN-c26-1* contained 733 genes, and the BW QTL under irrigated conditions, *qBW-c05-2* had 2,074 genes.

The GO analysis used all the identified candidate genes to identify potential biological functions and grouped under three main GO categories: biological process, molecular function, and cellular component (Supplementary Table 3; Sheet “GO”). Within the biological process category, there were 12 sub-categories, such as abiotic stress-specific, such as response to freezing, and response to biotic stimulus sub-category possessed the maximum number of 15 candidate genes. ADP binding with 21 genes was the main sub-category in the molecular function category with 14 principal sub-categories. Finally, in the cellular component category, there was one sub-category (vacuolar membrane) with 11 genes (Supplementary Table 3; Sheet “GO”). On the other hand, KEGG analysis resulted in only one category, namely, plant-pathogen interaction with four candidate genes (Supplementary Table 3; Sheet “KEGG”).

**DISCUSSION**

Some cotton lines have developed unique tactics to successfully handle the challenges of shifting and unexpected conditions, particularly under water stress. However, due to the intricacy of these characteristics and a lack of knowledge of the genetic processes behind these traits, introducing such drought tolerance qualities into elite cotton cultivars through conventional breeding approaches has been sluggish. We used an intraspecific linkage map generated using SNP markers to identify QTL connected to the productive attributes under water stress in this work. RILs developed from a cross between MCU5 (a well-known commercial cultivar in South India) and a good combiner, TCH1218, were tested in the field under irrigated, managed water stress, and rainfed conditions to identify QTL conferring productive traits. The RILs used in this study are useful and valuable asset for the QTL mapping of productive traits under water-limited conditions.
environments. The individuals of RILs exhibited almost all possible kinds of variations for the productive traits under water stress (Table 2).

It was noticed that water stress invariably reduced the productive trait expression in all the trials and the impact of drought included the wilting and drooping of leaves and reduced boll set and ultimately worsened the yield. Similar kinds of the impact of water stress in cotton have already been reported (Saranga et al., 2004; Zheng et al., 2016). The broad-sense heritability for the investigated traits was low to moderate (Table 2), and it indicated that the drought tolerance in cotton is greatly influenced by multiple genes and strongly affected by the environmental conditions. The impact of environmental conditions on the productive traits has already been reported in cotton (Jia et al., 2014). Though those low to moderate heritability may not accurately predict QTL, it may specify the fact that the investigated planting materials experienced relatively uniform treatments in all the trials. Furthermore, individual RILs in this work demonstrated a wide variety of reactions to the features under consideration (Table 2), and it was discovered that such responses are reasonably constant between trials. This study, in addition to having the advantage of screening the RILs under real field stress conditions with appropriate replications, also provides an example of how to carry out the experiment with less expenditure and more realistically when compared with artificial screening for drought tolerance in cotton.

This would be more useful in this genomics era, as genotyping costs are increasingly reduced, but phenotyping costs are increased tremendously when advanced and complicated infrastructure is developed for drought resistance screening. Thus, this study has shown that the replicated field screening using permanent mapping populations, such as RILs, under natural water stress conditions in multiple target environments provide valuable information for accurate QTL mapping in crop plants, such as cotton (Boopathi, 2020). Further, the positive correlation reported in this study among the examined traits specified that productive traits can be improved by selecting the agronomical traits, such as appropriate PH and the number of sympodial branches.

**Single Nucleotide Polymorphisms for Linkage Map Development**

The lack of large numbers and the segregation distortion of polymorphic markers are the main leading cause for the availability of a limited number of high-density linkage maps in cotton (Zhang T. et al., 2015). As a result, cotton genetic maps with SSR or AFLP markers often shows large gaps, poor marker density, and low marker coverage, whereas SNPs have shown their utility in constructing a high-dense genetic map as they are distributed throughout the genome (Cai et al., 2017; Liu et al., 2018). Among various strategies used in SNP genotyping (such
FIGURE 1 | Linkage map showing the relative positions of quantitative trait loci (QTL) identified in this study. QTL nomenclature is as indicated in Table 6.
as, genotyping by sequencing and whole-genome resequencing), SNP arrays have been shown to be simple and useful in the development of dense genetic maps in cotton (Li et al., 2016; Tan et al., 2018; Kumar et al., 2019). For example, in the present study 2,457 polymorphic markers between the parents of an intraspecific cross of upland cotton were identified using Cotton SNP50K array.

Progenies resulting from various cotton cross combinations frequently have unequal allelic distribution (Shao et al., 2014). Because of marker and population types, preferred fertilization, gametic combinations, genetic drift, and environmental variables, segregation distortion can vary across and within species (Shen et al., 2007). Despite the fact that a considerable number of polymorphic markers were found in this study, a large proportion of them (1,953, or 79.48% of total polymorphic loci identified between the parents) were removed due to segregation distortion. As a result, the number of loci in this study that can be adequately mapped has been drastically decreased. Cotton has also shown a
Several small effect QTL are involved in confirming the drought tolerance in cotton, and each loci represents hundreds of genes, which are the genetic basis for expressing an extensive array of water stress responses in the form of morpho-physiological traits (Abdelraheem et al., 2021). As only a few studies have focused on the QTL mapping of drought tolerance in cotton (reviewed in Mahmood et al., 2020), the molecular breeding of cotton for the genetic improvement of water stress tolerance is considered as a challenging task. Further, such QTL have large intervals (Mahmood et al., 2020; Abdelraheem et al., 2021).

This study found large intervals in the identified QTL, mainly attributed to the poor polymorphism rate detected between the two parents used in this study. Further, due to the segregation distortion, the number of mappable markers was reduced drastically from 2,457 to 504, which resulted in due to the segregation distortion. This drastic reduction of mappable markers resulted in gaps in the genetic map constructed in this study, reporting in G. hirsutum (Wang et al., 2006; Wu et al., 2009; Sun et al., 2012) and even in interspecific crosses of Gossypium spp. (He et al., 2007; Lacape et al., 2010).

Though there are different strategies (such as, genotyping by sequencing) to increase the efficiency of genetic mapping and QTL analysis through fine mapping, it would be desirable to employ high-density SNP chips as it enables speedy and automated detection of a high rate of polymorphisms across cotton accessions within a short span of time. Therefore, the development of new SNP chips by using an additional set of SNPs, detected on the water-stress responsive candidate genes identified in this study, would further enhance the efficiency of cotton molecular breeding for abiotic stress resistance.

Compared with the published reports on QTL linked to the drought tolerance and fiber yield and quality traits...
Causal genomic regions that impart at least a minimum consensus and hot spot QTL across the testing environments et al., 2005; Shen et al., 2007). It is imperative to identify the previously been demonstrated (Paterson et al., 2003; Ulloa that even the inferior parent (MCU5) might contribute to this study were under strong genotype by environment (GxE) interaction and low heritability. Despite this, nine common QTL this study were under strong genotype by environment (GxE) interaction and low heritability. Despite this, nine common QTL under water stress conditions; hyphen followed by “c” and a numeral denotes the chromosome number (chromosome 8 in the above case) and last numeral prefixed with a hyphen indicates the nth number of QTL identified for the given trait in the given chromosome. NA- not available.

**TABLE 6 | Identification of co-localized, stable, and major QTL for the investigated traits in this study.**

| Trait | Chromosome Number | QTL | Chromosome | Interval | QTL position | Additive | Dominant | Main effect | PVE% | LOD |
|-------|-------------------|-----|------------|----------|--------------|----------|----------|-------------|------|-----|
| BN 5  | gBN-c05-1         | T3  | AX-18225321-AX-182274753 | 158.11 | 12.3 | -0.7608 | 0.6414 | -0.84306 | 81.34 | 7.5 |
| BW 24 | qBW-c24-1         | T6  | AX-182258036-AX-17831424 | 81.31 | 17.94 | -0.2815 | 1.0112 | 3.592185 | 7.14 | NA |
| PH 26 | qPH-c26-3         | T3  | AX-182254129-AX-182267477 | 55.91 | 91.77 | -47.1439 | 43.6916 | -0.92677 | 0.02 | 18.5 |
| BN 26 | qBN-c26-1         | T6  | AX-182258747-AX-180652689 | 110.51 | 34.83 | -2.5734 | 8.3187 | -3.23257 | 83.62 | 7.9 |

QTL nomenclature: qPH-c08-1 can be interpreted as: first letter “q” to represent QTL; second two letters to represent trait name (PH for plant height, SYM for number of sympodial, BN for number of bolls and BW for boll weight) under irrigated conditions (if there is a “S” suffix after the trait name), it means that the given trait was identified under water stress conditions; hyphen followed by “c” and a numeral denotes the chromosome number (chromosome 8 in the above case) and last numeral prefixed with a hyphen indicates the nth number of QTL identified for the given trait in the given chromosome.

(Mahmood et al., 2020), a relatively low number of QTL were detected in this study. It indicated that the traits examined in this study were under strong genotype by environment (GxE) interaction and low heritability. Despite this, nine common QTL were detected across the seasons and locations, which pointed out that the collected phenotypic data are reliable even though there were GxE interactions. Further, a significant correlation among those traits and a substantial amount of heritability and additive effect estimates for most of the QTL identified in this study showed that the selection of investigated traits for drought tolerance improvement in cotton under filed water stress conditions would be more efficient. Similar evidence for selecting productive traits under abiotic stress environments in cotton has already been reported (Abdelraheem et al., 2018, 2021).

Surprisingly, both parents contributed to the additive effects of water stress tolerance alleles (Table 6). This demonstrated that even the inferior parent (MCU5) might contribute to the establishment of drought tolerance, a feature that has previously been demonstrated (Paterson et al., 2003; Ulloa et al., 2005; Shen et al., 2007). It is imperative to identify the consensus and hot spot QTL across the testing environments (as shown below), which will facilitate the introgression of causal genomic regions that impart at least a minimum increase in the productivity under water stress. Though this study had identified several QTL (Supplementary Table 1), only those QTL that have shown their potential in molecular breeding have been selected (Table 6) for further analysis. This study had identified a total of 11 QTL in joint multi-environmental analysis. Compared with individual analysis (Supplementary Table 1), the joint multi-environmental analysis (Table 7) identified a smaller number of QTL due to both QTL under water stress conditions; hyphen followed by “c” and a numeral denotes the chromosome number (chromosome 8 in the above case) and last numeral prefixed with a hyphen indicates the nth number of QTL identified for the given trait in the given chromosome. NA- not available.

**Hotspot Quantitative Trait Loci That Colocalized for Multiple Traits and Candidate Genes**

Further, this study attempted to identify hotspot QTL (which is defined here as a cotton genomic region), where QTL were discovered for more than one trait that was investigated across the seasons and locations expressing with more than 10% phenotypic variations (in this study) and genetic backgrounds.
TABLE 6

| Chr | Position (cm) | Left Marker | Right Marker | LOD (A/E) | LOD (A) | LOD (E) | LOD (A/E) | LOD (A) | LOD (E) | LOD (A/E) | LOD (A) | LOD (E) | LOD (A/E) | LOD (A) | LOD (E) |
|-----|---------------|-------------|--------------|-----------|---------|---------|-----------|---------|---------|-----------|---------|---------|-----------|---------|---------|
| qSYM_E-c11-1 | 11 71 | AX-182254626 | AX-182256182 | 3.0349 | 2.6598 | 0.3751 | 1.8054 | 1.3693 | 0.4361 | 1.0932 | 0.3309 | 0.8630 | 0.2186 | 0.7214 | 0.2629 |
| qSYM_E-c11-2 | 15 183 | AX-182256182 | AX-182255051 | 3.0349 | 2.6598 | 0.3751 | 1.8054 | 1.3693 | 0.4361 | 1.0932 | 0.3309 | 0.8630 | 0.2186 | 0.7214 | 0.2629 |
| qSYM_E-c15-1 | 15 183 | AX-182256182 | AX-182255051 | 3.0349 | 2.6598 | 0.3751 | 1.8054 | 1.3693 | 0.4361 | 1.0932 | 0.3309 | 0.8630 | 0.2186 | 0.7214 | 0.2629 |
| qSYM_E-c18-1 | 26 135 | AX-182256182 | AX-182255051 | 3.0349 | 2.6598 | 0.3751 | 1.8054 | 1.3693 | 0.4361 | 1.0932 | 0.3309 | 0.8630 | 0.2186 | 0.7214 | 0.2629 |
| qSYM_E-c18-2 | 24 114 | AX-182256182 | AX-182255051 | 3.0349 | 2.6598 | 0.3751 | 1.8054 | 1.3693 | 0.4361 | 1.0932 | 0.3309 | 0.8630 | 0.2186 | 0.7214 | 0.2629 |
| qSYM_E-c18-3 | 24 114 | AX-182256182 | AX-182255051 | 3.0349 | 2.6598 | 0.3751 | 1.8054 | 1.3693 | 0.4361 | 1.0932 | 0.3309 | 0.8630 | 0.2186 | 0.7214 | 0.2629 |

QTL nomenclature is that of Boopathi et al. (2018), and the alphabet E denotes joint the multi-environmental QTL mapping. LOD: Logarithm of the Odds; A by E: Additive QTL × Environment; PVE: % of phenotypic variation explained.

Identification of additive × environment interaction effect QTL by the traits investigated in this study. QTLs for Drought Tolerance in Cotton

Identification of additive × environment interaction effect QTL by the traits investigated in this study. LOD: Logarithm of the Odds; A by E: Additive QTL × Environment; PVE: % of phenotypic variation explained.

Among the three PH QTL chosen in this investigation (Table 6), a QTL on chromosome 11, qPH-c11-1, has demonstrated the most phenotypic variance (38.41%), and this region has been proven to be co-localized for a maximum number of major and stable QTL (7) linked to the number of sympodial branches both under irrigated and water-stressed conditions (Figure 1). This region was specifically detected for fruit branch node number by Hulse-Kemp et al. (2015). Another notable region was on chromosome 18 that span a length of 61.4 cM flanked by the SNPs, AX-182256182 and AX-182255051. This region was associated with five QTL linked to the number of sympodial branches both under irrigated and water-stressed conditions in this study. Similarly, a QTL cluster was found on chromosome 5 for BW and BN in this study and similar kind of correlation of this chromosomal region with several productive traits has been reported earlier (Abdelrahem et al., 2018).

Except for chromosomes 5 and 26, almost all of the hotspots discovered in this study were connected to the number of sympodial branches. Cotton bolls can form at the nodes of sympodia that develop from the monopodia, as well as at the main or secondary vegetative axis sites. Orderly development of bolls on sympodia has long been favored (McClelland, 1916), and plant design with 0% monopodial growth would be more ideal, favoring efficient mechanical harvesting and improving ultimate yield. Since only major, stable, and consistent QTL were reported in this study, it is likely that all of the QTL listed in Table 6 could be potential assets for consensus mapping across genetic backgrounds for their validation and identification of most useful QTL hotspots, map-based cloning, and MAS for productive traits in upland cotton with high predictability. Among the three PH QTL chosen in this investigation (Table 6), a QTL on chromosome 11, qPH-c11-1, has demonstrated the most phenotypic variance (38.41%), and this area has been proven to host PH QTL under irrigated (Jia et al., 2016) and salt stress environments (Abdelrahem et al., 2021). Optimum plant height in cotton is the most desirable trait in the terms of mechanical harvesting, besides its direct relation with biomass that influences the final yield (Shang et al., 2015).

It would be an elaborated list to identify candidate genes for the QTL reported in this study, as it has been estimated that ∼4,500 cM of G. hirsutum genome consisted of 70,478 predicted protein-coding genes (Zhang T. et al., 2015). Similarly, the large array of candidate genes was found in this study (Supplementary Table 2). Despite this extensive list, identifying candidate genes in the target hotspot QTL can open up new avenues in understanding the molecular basis of drought tolerance in cotton.

For example, qSYM-c22-1, a QTL on chromosome 22 flanked by AX-182254626 and AX-182258142, was found to be co-localized for the number of sympodial branches both under irrigated and water-stressed conditions, and it harbors 1,046 genes (Supplementary Table 2). Among them, the notables are GDSL esterases/lipases, involved in the regulation of plant development by the synthesis of secondary metabolites in response to the biotic and abiotic stresses (Chepyshko et al., 2012), zinc finger protein CONSTANS, regulating flowering (elsewhere) were co-localized in the same genomic segment. Nine hotspot QTL were identified in this study that were located on chromosomes 5, 7, 9, 11, 15, 18, 21, 22, 24, and 26 (Table 6). Among them, a region on chromosome 22 flanked by AX-182254626–AX-18226770 with a span length of 89.4 cM, is co-localized for a maximum number of major and stable QTL (7) linked to the number of sympodial branches both under irrigated and water-stressed conditions (Figure 1). This region was specifically detected for fruit branch node number by Hulse-Kemp et al. (2015). Another notable region was on chromosome 18 that span a length of 61.4 cM flanked by the SNPs, AX-182256182 and AX-182255051. This region was associated with five QTL linked to the number of sympodial branches both under irrigated and water-stressed conditions in this study. Similarly, a QTL cluster was found on chromosome 5 for BW and BN in this study and similar kind of correlation of this chromosomal region with several productive traits has been reported earlier (Abdelrahem et al., 2018).

Except for chromosomes 5 and 26, almost all of the hotspots discovered in this study were connected to the number of sympodial branches. Cotton bolls can form at the nodes of sympodia that develop from the monopodia, as well as at the main or secondary vegetative axis sites. Orderly development of bolls on sympodia has long been favored (McClelland, 1916), and plant design with 0% monopodial growth would be more ideal, favoring efficient mechanical harvesting and improving ultimate yield. Since only major, stable, and consistent QTL were reported in this study, it is likely that all of the QTL listed in Table 6 could be potential assets for consensus mapping across genetic backgrounds for their validation and identification of most useful QTL hotspots, map-based cloning, and MAS for productive traits in upland cotton with high predictability. Among the three PH QTL chosen in this investigation (Table 6), a QTL on chromosome 11, qPH-c11-1, has demonstrated the most phenotypic variance (38.41%), and this area has been proven to host PH QTL under irrigated (Jia et al., 2016) and salt stress environments (Abdelrahem et al., 2021). Optimum plant height in cotton is the most desirable trait in the terms of mechanical harvesting, besides its direct relation with biomass that influences the final yield (Shang et al., 2015).

It would be an elaborated list to identify candidate genes for the QTL reported in this study, as it has been estimated that ∼4,500 cM of G. hirsutum genome consisted of 70,478 predicted protein-coding genes (Zhang T. et al., 2015). Similarly, the large array of candidate genes was found in this study (Supplementary Table 2). Despite this extensive list, identifying candidate genes in the target hotspot QTL can open up new avenues in understanding the molecular basis of drought tolerance in cotton.

For example, qSYM-c22-1, a QTL on chromosome 22 flanked by AX-182254626 and AX-182258142, was found to be co-localized for the number of sympodial branches both under irrigated and water-stressed conditions, and it harbors 1,046 genes (Supplementary Table 2). Among them, the notables are GDSL esterases/lipases, involved in the regulation of plant development by the synthesis of secondary metabolites in response to the biotic and abiotic stresses (Chepyshko et al., 2012), zinc finger protein CONSTANS, regulating flowering (elsewhere) were co-localized in the same genomic segment. Nine hotspot QTL were identified in this study that were located on chromosomes 5, 7, 9, 11, 15, 18, 21, 22, 24, and 26 (Table 6).
under normal and stress conditions (Putterill et al., 1995), abscisic acid receptor PYL8, expressed in response to dark-induced leaf senescence (Lee et al., 2015), choline mono oxygenase, conferring abiotic stress tolerance by synthesizing glycine betaine (Bao et al., 2011), potassium transporter 2, copper transport protein ATX1, and peroxidase 52, contributing significantly for abiotic stress resistance (Zhang et al., 2021) and several network of reactive oxygen species (ROS) genes, those have played critical role in both abiotic stress response and fiber development (Xu et al., 2019). Similar kinds of candidate genes were identified for another hotspot QTL (Supplementary Table 2). Thus, this study provides a promising lead to precise the QTL mapping consisting of functional SNPs derived from those candidate genes identified in the target QTL by developing a new SNP chip.

CONCLUSION

Identifying major and stable QTL for productive characteristics in water-stressed conditions is a precondition for developing an effective cotton molecular breeding program. This work found such QTL using intraspecific RILs, which allowed them to be examined in multiple seasons and circumstances. Collecting replicated phenotypic data for successful QTL mapping requires intraspecific RILs, which allowed them to be evaluated in different seasons and situations. Though this study contributed preliminary information on candidate genes that unravel the molecular mechanism underlying cotton productivity under water stress, fine-mapping those QTL using additional SNPs derived from those candidate genes would be required to validate and employ them in MAS for the genetic improvement of cotton with improved productivity under water-stressed environments.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS

NB and SJ designed the experiments, prepared the documents for funding, and drafted the full manuscript. NB developed the RIL reported in this study. GT and SJ performed all the analyses of genetic and QTL mapping. NB, KN, PS, VS, NP, and SR evaluated mapping populations at different locations and years and collected phenotypic data on fiber traits. GT and BJ drafted the results sections and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

We would like to acknowledge DBT, New Delhi, for funding the national network project on cotton (BT/PR12445/AGII/106/982/2014).

ACKNOWLEDGMENTS

We are thankful to the Director, CSIR-NBRI, Lucknow, and Directors, CPMB&B, and Directors of Research, TNAU, Coimbatore for providing the facility to execute the project. We also acknowledge CSIR for partial financial support, when the DBT project was extended without additional fund. We would like to thank Imperial Life Sciences (P) Limited, New Delhi, for SNP genotyping with the 50K Axiom array.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.851504/full#supplementary-material
McClelland, C. K. (1916). On the regularity in the blooming of cotton plant. *Science* 43, 587–588. doi: 10.1126/science.43.1139.587

Meng, L., Li, H., Zhang, L., and Wang, J. (2015). QTL IciMapping; integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* 3, 269–283. doi: 10.1016/j.cj.2015.01.001

Niú, J., Zhang, S., Liu, S., Ma, H., Chen, J., Shen, Q., et al. (2018). The compensation effects of physiology and yield in cotton after drought stress. *J. Plant Physiol.* 224, 30–48. doi: 10.1016/j.jplph.2018.03.001

Paton, A. H., Saranya, M., Jiang, C. X., and Wright, R. (2003). QTL analysis of genotype × environment interactions affecting cotton fiber quality. *Theor. Appl. Genet.* 106, 384–396. doi: 10.1007/s00122-002-1025-y

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The CONSTANS gene of *Arabidopsis* promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80, 847–857. doi: 10.1016/0092-8674(95)90288-0

Said, J. J., Knapka, J. A., Song, M. Z., and Zhang, J. F. (2015). Cotton QTLdb: a cotton QTL database for QTL analysis, visualization, and comparison between *Gossypium hirsutum* and *G. barbadense* populations. *Mol. Genet. Genomics* 290, 1615–1625. doi: 10.1007/s00438-015-1021-y

Saranya, J., Jiang, C., Wright, R. J., Yakir, D., and Paterson, A. H. (2004). Genetic dissection of cotton physiological responses to arid conditions and their interrelationships with productivity. *Plant Cell Environ.* 27, 263–277. doi: 10.1111/j.1365-3040.2003.01134.x

Selvakumar, P., Ravikeshavan, R., Gopikrishnan, A., Thyigave, K., Preetha, S., and Manikanda Boopathi, N. (2010). Genetic purity analysis of cotton (*Gossypium* spp.) hybrids using SSR markers. *Seed Sci. Technol.* 38, 358–366. doi: 10.15258/sst.2010.38.2.09

Shang, L. F., Liu, W. Y., Abduselati, A., Cai, S. W., Wang, K. et al. (2015). Dynamic QTL mapping for plant height in upland cotton (*Gossypium hirsutum*). *Plant Breed.* 134, 703–712. doi: 10.1111/pbr.12316

Shao, Q. Z., Zhang, F., Tang, S., Liu, Y., Fang, X., Liu, D., et al. (2014). Identifying QTL for fiber quality traits with three upland cotton (*Gossypium hirsutum* L.) populations. *Euphytica* 198, 43–58. doi: 10.1007/s10681-014-1082-y

Shen, X., Guo, W., Lu, Q., Zuo, X., Yuan, Y., and Zhang, T. (2007). Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. *Euphytica* 155, 371–380. doi: 10.1007/s10681-007-9490-8

Shen, X., Guo, W., Zhu, X., Yuan, Y., John, Z. Y., Kohel, R. J., et al. (2005). Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. *Mol. Breed.* 15, 169–181. doi: 10.1007/s11032-004-4731-0

Shim, J., Mangat, P. K., and Aneles-Shim, R. B. (2018). Natural variation in wild *Gossypium* species as a tool to broaden the genetic base of cultivated cotton. *J. Plant Sci. Curr. Res.* 2:005.

Shukla, R. P., Tiwari, G. J., Joshi, B., Song-Beng, K., Tamta, S., Boopathi, N. M., et al. (2021). GBS-SNP and SSR based genetic mapping and QTL analysis for drought tolerance in upland cotton. *Physiol. Mol. Biol. Plants* 27, 1731–1745. doi: 10.15258/plaphy.2021.0225

Sun, F. D., Zhang, J. H., Wang, S. F., Gong, W. K., Shi, Y. Z., Liu, A. Y., et al. (2005). Chromosomal assignment of RFLP linkage groups harboring important QTLs on an intraspecific cotton (*Gossypium hirsutum* L.) germplasm. *Theor. Appl. Genet.* 110, 1397–1406. doi: 10.1007/s00122-000-0803-6

Tan, Z., Zhang, J. H., Wang, S. F., Li, Y., Yang, P., et al. (2018). Natural variation in wild *Gossypium herbaceum* L. genome: new insights from bioinformatics analysis. *Theor. Appl. Genet.* 134, 661–685. doi: 10.1007/s00122-020-05208-y

Wang, S., Basten, C. J., and Zeng, Z.-B. (2012). *QTLs for Drought Tolerance in Cotton*. Raleigh, NC: Department of Statistics, North Carolina State University.
Wen, T., Liu, C., Wang, T., Wang, M., Tang, F., and He, L. (2021). Genomic mapping and identification of candidate genes encoding nulliplex-branch trait in sea-island cotton (*Gossypium barbadense* L.) by multi-omics analysis. *Mol. Breed.* 41, 1–12.

Wu, J., Gutierrez, O. A., Jenkins, J. N., McCarty, J. C., and Zhu, J. (2009). Quantitative analysis and QTL mapping for agronomic and fiber traits in an *Gossypium tomentosum* and *Gossypium hirsutum* population of upland cotton. *Euphytica* 165, 231–245. doi: 10.1007/s10681-008-9748-8

Xu, S., Pan, Z., Yin, F., Yang, Q., Lin, Z., Wen, T., et al. (2020). Identification of candidate genes controlling fiber quality traits in upland cotton through integration of meta-QTL, significant SNP and transcriptomic data. *J. Cotton Res.* 3:34. doi: 10.1186/s42397-020-00075-z

Xu, Y., Magwanga, R. O., Cai, X., Zhou, Z., Wang, X., Wang, Y., et al. (2019). Deep transcriptome analysis reveals reactive oxygen species (ROS) network evolution, response to abiotic stress, and regulation of fiber development in cotton. *Int. J. Mol. Sci.* 20:1863. doi: 10.3390/ijms20081863

Yan, W., Du, M., Zhao, W., Li, F., Wang, X., Eneji, A. E., et al. (2019). Relationships between plant architecture traits and cotton yield within the plant height range of 80–120 cm desired for mechanical harvesting in the Yellow River Valley of China. *Agronomy* 9:587. doi: 10.3390/agronomy9100587

Yu, J., Jung, S., Cheng, C.-H., Lee, T., Zheng, P., Buble, K., et al. (2021). CottonGen: the community database for cotton genomics, genetics, and breeding research. *Plants* 10:2805. doi: 10.3390/plants10122805

Yu, J., Yu, S., Gore, M., Wu, M., Zhai, H., Li, X., et al. (2013). Identification of quantitative trait loci across interspecific F 2, F 2: 3 and testcross populations for agronomic and fiber traits in tetraploid cotton. *Euphytica* 191, 375–389. doi: 10.1007/s10681-013-0875-5

Zhang, H., Zhu, J., Gong, Z., and Zhu, J. K. (2021). Abiotic stress responses in plants. *Nat. Rev. Genet.* 23, 104–119.

Zhang, J., Zhang, Q. X., Cheng, T. R., Yang, W. R., Pan, H. T., Zhong, J. J., et al. (2015). High-density genetic map construction and identification of a locus controlling weeping trait in an ornamental woody plant (*Prunus mume* Sieb. et Zucc.). *DNA Res.* 22, 183–191. doi: 10.1093/dnares/dsv003

Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., et al. (2015). Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* 33, 531–537. doi: 10.1038/nbt.3207

Zheng, J. Y., Oluoch, G., Khan, M. K. R., Wang, X. X., and Cai, X. Y. (2016). Mapping QTLs for drought tolerance in an F 2: 3 population from an interspecific cross between *Gossypium hirsutum* and *Gossypium tomentosum*. *Genet. Mol. Res.* 15, 1–14. doi: 10.4238/gmr.15038477

Zhu, L., Andres, R. J., Zhang, K., and Kuraparthi, V. (2021). High-density linkage map construction and QTL analysis of fiber quality and lint percentage in tetraploid cotton. *Crop Sci.* 61, 3340–3360. doi: 10.1002/csc2.20519

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

**Copyright © 2022 Boopathi, Tiwari, Jena, Nandhini, Sri Subalakhshmi, Shyamala, Joshi, Premalatha and Rajeswari. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**