agronomy

Article

Evaluation of Quality Traits in Relation to Mechanical Harvesting for Screening Excellent Materials in *Gossypium barbadense* L. Germplasm Resources

Feng Lin [1,2,†], Meng Wang [2,†], Nan Zhao [3], Yubo Zhang [1,2], Weiran Wang [2], Jing Yang [2], Sumei Wan [1], Jianping Li [2], Alifu Aierxi [2], Guodong Chen [1,*] and Jie Kong [2,*]

1 College of Agriculture, Tarim University, Alar 843300, China; 10757213011@stumail.taru.edu.cn (F.L.); 10757222059@stumail.taru.edu.cn (Y.Z.); wansumei510@163.com (S.W.)
2 Institute of Cash Crops, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China; wm619715@163.com (M.W.); xaaswwr@163.com (W.W.); jingy2022@webmailhzau.edu.cn (J.Y.); ljp7786@126.com (J.L.); alip68@126.com (A.A.)
3 College of Agronomy and Biotechnology, China Agricultural University, Beijing 100083, China; nan_zhao@cau.edu.cn
* Correspondence: guodongchen@taru.edu.cn (G.C.); kongjie.258@163.com (J.K.)
† These authors contributed equally to this work.

Abstract: Sea Island cotton is renowned for its superior fiber quality. Although mechanical harvesting has the potential to significantly increase efficiency and reduce the production cost of Sea Island cotton, there is still little research in this area. In this study, we analyzed 240 Sea Island cotton germplasm resources and evaluated 19 traits related to mechanical harvesting. The coefficient of variation ranged from 5.42% to 66.96%, and the genetic diversity index spanned from 1.57 to 2.07. In most traits studied, there was a strong correlation between the height of the first fruiting branch and the defoliation rate. The 19 traits were categorized into 6 factorial groups by principal component analysis, in which the defoliation factor contributed the most (30.89%). The cluster analysis divided the 240 cotton accessions into four main groups, with the second group exhibiting favorable mechanical harvesting characteristics such as higher defoliation rate and first fruit branch height. Using stepwise regression, a model was constructed with the joint evaluation score F-value as the response variable and eight traits (X_1: PH, X_2: SNB, X_3: SBN, X_4: MBL, X_5: AFBM, X_7: MLIA, X_8: NB, and X_{13}: 15 d DR) as predictors: $Y = -7.2 + 0.01X_1 + 0.23X_2 + 0.192X_3 + 0.038X_4 + 0.007X_5 + 0.014X_7 + 0.025X_8 + 2.952X_{13}$. Selected materials suitable for machine harvesting, such as MoShi729, were identified. This study provides valuable theoretical insights into the mechanical harvesting of Sea Island cotton germplasm resources and identifies promising materials for targeted breeding and improvement programs.

Keywords: Sea Island cotton; germplasm resources; machine-harvesting traits; comprehensive evaluation; elite material screening

check for updates

Citation: Lin, F.; Wang, M.; Zhao, N.; Zhang, Y.; Wang, W.; Yang, J.; Wan, S.; Li, J.; Aierxi, A.; Chen, G.; et al. Evaluation of Quality Traits in Relation to Mechanical Harvesting for Screening Excellent Materials in *Gossypium barbadense* L. Germplasm Resources. *Agronomy* **2024**, *14*, 891. https://doi.org/10.3390/agronomy14050891

Academic Editors: Fernando Martinez-Moreno, Magdalena Ruiz, María B. Picó and María-José Díez

Received: 18 March 2024
Revised: 16 April 2024
Accepted: 23 April 2024
Published: 24 April 2024

1. Introduction

Cotton is an important economic crop that serves as the major source of natural fibers for the textile industry [1]. Historically, the genus Gossypium includes seven heterozygous tetraploid cotton species, with upland cotton and Sea Island cotton playing the dominant roles in production [2]. With the rapid development of the textile industry, the quality requirements for cotton fibers are becoming higher and higher [3]. Sea Island cotton fibers are known for their length and strength [4] and being superior to upland cotton [5] and are therefore widely used in the production of high-end cotton textiles. As one of the main producing areas of Sea Island cotton, Xinjiang region saw mechanical harvesting accounting for only about 10%, which tended to push up production costs. The Sea Island cotton varieties that are popular in Xinjiang have 'zero-branching' fruiting patterns characterized by low height of the first branch and low defoliation rate, resulting in reduced net picking rates of mechanical harvesting and higher levels of seed cotton pollution, which ultimately posed challenges to processing and impeded the development of China's Sea Island cotton industry [6]. Accordingly, enhancing the suitability of mechanical harvesting has become a key requirement for the improvement of Sea Island cotton varieties.

Mechanical cotton harvesting was used in the United States as early as 1948 and the technology was then widely adopted [7]. Mechanical harvesting was introduced in China in 1973, and by 2015 it had become standard practice for upland cotton in Xinjiang, China [8]. Mechanical cotton harvesting saves labor, reduces production costs, and expands economic benefits [9]. Due to favorable characteristics for mechanical harvesting, upland cotton has become the most widely planted variety, accounting for more than 90% of the world's cotton production [10]. Therefore, achieving mechanization is essential to increase the share of Sea Island cotton in production [11]. Previous studies have found a significant correlation between plant structures suitable for mechanical harvesting and crop yield, highlighting the effects of the first fruit branch height [12] and defoliation effectiveness [13] on mechanical harvesting efficiency. Not only that, but previous research has found that the plant structure (plant height, degree of looseness, etc.) and early maturity also have a great impact on the efficiency of mechanical picking, island cotton compared to land cotton has a higher plant height, the zero-fruit branch type of island cotton plant type is also more compact, and island cotton has a longer fertility period; all of these factors will be detrimental to the machine picking of sea island cotton [14].

Germplasm resources assessment is the basis of resource utilization and directed breeding [15]. The introduction of the term "core germplasm" [16] further emphasizes the importance of resource assessment and classification. References [17,18] conducted a comprehensive discussion on mango resource utilization and pointed out that morphological data played an indispensable role in the preliminary evaluation of resources. Similarly, the selection of cold-tolerant germplasm in tropical crops like maize allows for earlier seeding [19]. Crops such as rice [20] and peanuts [21] could also benefit from excellent germplasm ascertained through evaluation. Therefore, it is not only feasible but also significant to evaluate the suitability of mechanical harvesting of Sea Island cotton germplasm resources. The commonly used comprehensive evaluation method is based on membership function values, which have been widely used to assess the tolerance to drought [22], salt and alkali [23], and heat [24].

Different varieties of sea island cotton have different mechanical harvesting effects, and the main indexes to measure the mechanized harvesting effect of sea island cotton are not clear. This study comprehensively analyzed the global 240 varieties of Sea Island cotton, the 19 kinds of quality traits related to machine harvesting, the evaluation of Sea Island cotton machine harvesting, the screening out the reasonable plant type, and high defoliation rate, which are suitable for machine harvesting of Sea Island cotton varieties, for the new varieties of genetic improvement, and breeding to provide a material basis and theoretical reference.

2. Materials and Methods

2.1. Experimental Materials

The experimental materials used in this study were derived from a collection of Sea Island cotton germplasm resources, which have been introduced and cultivated since the 1950s by the Economic Crop Research Institute of Xinjiang Academy of Agricultural Sciences. Next, 240 Sea Island cotton germplasm accessions were carefully screened according to their genetic background and the specific investigation requirements. The sources of these materials include 14 countries and regions in Asia, Africa, and the Americas (Figure 1c). This diverse selection highlights the global representation and rich genetic diversity of the materials examined.

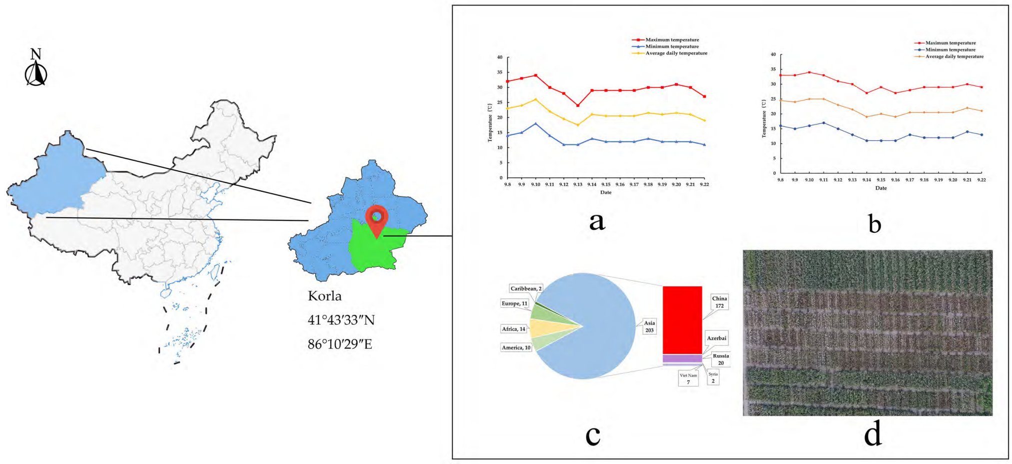

Figure 1. An overview of the test conditions and a comparison of spray effects. (**a**) Climate change in 2023; (**b**) climate change in 2022; (**c**) source of germplasm resources; (**d**) field effect of defoliant spraying.

2.2. Field Management

In 2022–2023, 240 Sea Island cotton germplasm resource materials were planted at the Korla cotton breeding base of Xinjiang Academy of Agricultural Sciences. The base has a temperate continental climate, with an average annual temperature of 3–18 °C, annual rainfall of 39.7 mm, and a loam soil. The mechanically harvested planting mode with 1 film and 4 rows (10 + 66 + 10 cm) was adopted, and the theoretical density was 222,000 plants·hm-2 with 4 rows per plot and a length of 5 m. On 18 April, artificial modulation was performed after mechanical coating, followed by artificial topdressing on 20 July. On 8 September, 12 g 360 g·L^{-1} of Thifensulfuron-180 g·L^{-1} per 667 m^2 was applied per 667 m^2 [25], and the artificial topdressing was applied. Diquat (moderately toxic, produced by Bayer, Germany) showed no rainfall within 7 days after the drug. Within 15 days, the temperature (Figure 1a,b) reached the optimal spraying conditions [26], and the defoliation effect met the experimental requirements (Figure 1d).

2.3. Data Collection

Plant structure and defoliation mainly affect the machine-harvesting effect of Sea Island cotton, and this experiment centered on the investigation of agronomic traits around these two types of factors.

Plant height (PH, cm): on 1 September, the height of Sea Island cotton was measured from the ground to the growing point using a tape measure.

Blade number (NB): on 1 September, the number of all leaves including main stem leaves and fruiting branch leaves was investigated.

Initial node height (SNB, cm): on 1 September, the height of Sea Island cotton from the ground to the first fruiting branch was measured using a tape measure.

First fruiting branch node (SBN): on 1 September, from the number of cotyledon nodes (cotyledon nodes count 0) to the first fruit branch, between the number of nodes is the number of beginning nodes.

Middle fruiting branch length (MLB, cm): on 1 September, the length of the fifth fruiting branch was measured with a tape measure.

Lower fruiting branch length (LLFB, cm): on 1 September, the length of the second fruiting branch was measured with a tape measure.

Upper leaf inclination angle (ULIA, °): YX-501 (produced by Beijing Yaxinli Instrument Technology Co., Ltd., Beijing, China) was used to measure the angle between the upper (inverted two-leaf) main stem leaves and the main stem of the cotton plant, with the vertical direction of 90° as the baseline, and, in order to minimize the effect of environmental factors on the angle of leaf inclination, the measurement time was set to be from 13:00 to 15:00

Xinjiang local time. In order to minimize the influence of environmental factors on leaf inclination angle, the measurement time was set at 13:00–15:00 Xinjiang local time.

Middle leaf inclination angle (MLIA, °): using YX-501 to measure the angle between the main stem leaves and the main stem in the middle of the cotton plant (the fifth leaf), with 90° in the vertical direction as the baseline, in order to minimize the effect of environmental factors on leaf inclination angle, set the measurement time for the local time of 13:00–15:00 in Xinjiang.

Lower leaf inclination angle (LLIA, °): use YX-501 to measure the angle between the lower part of the cotton plant (the second leaf) and the main stem of the lower part of the cotton plant (the second leaf), with the vertical direction of 90 ° as the baseline; in order to minimize the effect of environmental factors on the angle of the leaf inclination, set the measurement time for the local time of 13:00–15:00 in Xinjiang.

Middle fruiting branch angle (AFBM, °): use YX-501 to measure the angle between the middle (fifth fruiting branch) fruiting branch and the main stem, with 90° as the baseline.

Lower fruiting branch angle (ALFB, °): the angle between the lower (second fruiting branch) fruiting branch and the main stem was measured using the YX-501, with 90° as the baseline.

Fluorescence parameters: using a portable plant efficiency analyzer (Handy-PEA), we measured the fluorescence parameters of the inverted bilobed of the main stem of cotton from 11:00 to 13:00, and the leaves were dark adapted for 30 min beforehand, and maximum fluorescence (F_v), initial fluorescence (F_0), and maximum photochemical efficiency (F_v/F_m) were recorded and the potential photochemical efficiency (F_v/F_0) was counted.

2.4. Data Analysis

Excel 2021 was used for preliminary data sorting, and the genetic diversity index grading method referred to a previous study [27]. Pearson correlation analysis and cluster analysis (Euclidean distance and ward.D algorithm) were performed using R 4.3.1. The data were standardized by SPSS Statistics 26 and then subjected to principal component analysis. The composite scores were calculated according to the principal component results. The number of leaves remaining on the cotton was surveyed at 5 d, 10 d, and 15 d after defoliant spraying.

$$\text{DR (Defoliation rate, \%)} = \frac{A - B}{A} \times 100\%$$

(A: total number of leaves; B: number of remaining leaves.)

Investigate the number of cotton hanging leaves (withered but not falling, falling but hanging) 5 d, 10 d, and 15 d after spraying defoliant, respectively.

$$\text{HGY (Hanging rate, \%)} = \frac{C}{A} \times 100\%$$

(A: total number of leaves; C: number of hanging leaves.)

The relevant calculation formula [28] is as follows, genetic diversity index:

$$H' = -\sum P_i Ln P_i$$

(P_i: the fraction of germplasm resources at level I for a trait as a percentage of the total number of germplasm copies; *Ln*: natural logarithm.)

The weights of each variable in the principal components:

$$W_{ij} = \frac{\theta_j}{\sqrt{\lambda_i}}$$

(θ_j: coefficients corresponding to each variable in the component matrix; $\sqrt{\lambda_i}$: the open root value of the eigenvalue corresponds to the ith principal component).

Individual principal component scores:

$$F_i = w_{i1}X_1 + w_{i2}X_2 + \cdots + w_{in}X_n$$

Combined score value:

$$F = \alpha_1 F_1 + \alpha_2 F_2 + \cdots + \alpha_n F_n$$

(α_i: percentage of variance for the ith principal component).

3. Results and Discussion

3.1. Genetic Diversity of Machine-Harvesting Traits of Sea Island Cotton Germplasm Resources

The genetic diversity of Sea Island cotton germplasm resources was the highest in plant height (2.07) and central fruit branch angle (2.07), the lowest genetic diversity index was 1.57 for total leaf number, and the genetic diversity index of the first fruit branch node and the first fruit branch height were both 2.00. The genetic diversity index ranged from 1.92 to 2.05 for triple defoliation rates and hanging rates. Genetic diversity indices for different traits varied significantly, which helped in subsequent data analysis and material screening. The study of phenotypic genetic diversity of germplasm resources is fundamental to the utilization of germplasm resources. Genetic diversity can not only clarify the population structure and the sources of variations but also show the genetic model of the population [29,30]. The presence of high genetic diversity within a population is more favorable for the subsequent screening of specific resources. The present study found that the genetic diversity of Sea Island cotton is higher than that of land cotton, which confirms the previous conjectures [31]. The genetic diversity index of plant height (2.07), leaf inclination in the upper, middle, and lower parts of the plant, and the angle between the middle and lower parts of the fruit branches were larger, indicating that plant type, leaf type, and disease resistance of Sea Island cotton population were significantly different. Overall, the genetic diversity of Sea Island cotton was greatly abundant, which was conducive to screening materials with the targeted traits, consistent with the opinion of Pan et al. [32].

At 15 days post-dose, the variation coefficient of the hanging rate was highest (66.96%), and the maximum photochemical quantum yield of PS II was lowest (5.42%). There were seven traits in the low-variation range from 0 to 15%, and the top three traits from large to small were low fruit branch angle, defoliation rate at 10 days post-dose, and low leaf inclination angle. A total of 12 traits were in the high-variation range of 15% to 100%, and the first 3 traits from the largest to the smallest were the branching rate at 15 days post-dose, the branching rate at 10 days post-dose, and the branching rate at 5 days post-dose. The variation coefficient of defoliation rate at 5 days post-dose was the highest (24.43), which was 10.03% and 14.74% higher than that at 10 days and 15 days post-dose, respectively. The variation coefficient of the defoliation rate became smaller and smaller as the time post-dose was extended. The maximum values of the hanging rate ranged from 50% to 77%, and the difference in the extreme value became smaller and smaller with the extension of time post-dose (Table 1). The variation coefficient of defoliation rate gradually decreased at 5, 10, and 15 days after defoliant spraying, suggesting that 5 days post-dose is the key time point to identify the sensitivity of different germplasm resources to defoliant, which is consistent with the findings of Li and Wang et al. [33,34]. In this study, we found that the extreme difference in the hanging rate became smaller and the variation coefficient became larger as the time post-dose is extended. This was in line with the hanging-sensitivity characteristic of Sea Island cotton sprayed with defoliants [6]. It indicated that the difference in the hanging situation among different germplasm resources was increasingly evident. This was more favorable for selecting resource materials with a low-hanging rate. The variation coefficients of 19 machine-harvesting traits ranged from 5.42 to 66.96%, with large differences among varieties, which was consistent with the study by Yu [35].

Table 1. Analysis of the variation coefficient and genetic diversity index of mechanically harvested characters in Sea Island cotton.

Traits	Mean	Max	Min	Range	SD	CV (%)	H′
PH	86.38	123.92	47.72	76.2	15.11	17.49	2.07
SNB	18.81	38.58	9.14	29.44	5.12	27.21	2
SBN	4.8	8.6	2.7	5.9	1.1	22.87	2
MLB	12.21	30.22	4.52	25.7	6.45	52.77	1.71
AFBM	56.28	79.6	29.6	50	10.27	18.25	2.07
LLFB	12.25	26.56	4.3	22.26	5.85	47.75	1.78
ALFB	70.62	107	38.4	68.6	10.58	14.98	2.05
ULIA	80.25	103.2	51.6	51.6	8.39	10.46	2
MLIA	83.31	112.4	60	52.4	7.9	9.49	2.04
LLIA	93.99	121.2	56.4	64.8	10.89	11.59	2.04
NB	13.86	38.4	5.6	32.8	7.48	53.98	1.57
F_v/F_0	3.34	5.29	1.37	3.92	0.66	19.67	2.06
F_v/F_m	0.76	0.84	0.56	0.28	0.04	5.42	1.98
5 d DR	0.63	0.93	0.23	0.7	0.15	24.43	2.05
10 d DR	0.79	1	0.35	0.65	0.11	14.4	2.01
15 d DR	0.88	1	0.5	0.5	0.09	9.69	1.94
5 d HGY	0.29	0.77	0	0.77	0.16	55.36	1.92
10 d HGY	0.19	0.65	0	0.65	0.12	64.08	1.93
15 d HGY	0.13	0.5	0	0.5	0.09	66.96	1.94

3.2. Correlation Analysis of Machine-Harvesting Traits in Sea Island Cotton

Through the correlation analysis of machine-harvesting traits of Sea Island cotton (Figure 2), it was found that the defoliation rate at different times was negatively correlated with the hanging rate at different times, and the negative correlation coefficient between the defoliation rate at 15 days and the hanging rate at 15 days was the highest (0.969). By comparing the correlation coefficients between the defoliation rate and the hanging rate at 5, 10, and 15 days, it was found that the correlation was gradually decreasing, and thus it was inferred that the 5 days defoliation rate was the key node to identify the final defoliation effect, as hanging leaves are leaves that die on the cotton plant and cannot be shed over time. The significant negative correlation between the defoliation rate of Sea Island cotton and the rate of hanging leaves suggests that material with a high defoliation rate is sensitive to defoliants, and the leaves can be shed without difficulty and are not prone to forming hanging leaves; on the contrary, material with a low defoliation rate will reduce the quality of machine picking due to the phenomenon of hanging leaves. The defoliation rates at 10 days and 15 days were positively correlated with plant height, first fruit branch degree, and first fruit branch height. The correlation of resource pairs of zero-type fruit branches is greater than that of mixed fruiting branch types. The defoliation rate at different times was positively correlated with the leaf inclination angle in different parts of the plant, and the defoliation rate at 10 days and 15 days was significantly correlated with the leaf inclination angle in the middle part of the plant. There was a positive correlation between the defoliation rate and the total leaf number, which gradually weakened as the time post-dose lengthened. The hanging rate was negatively correlated with the length of the middle and lower fruit branches, which was mainly reflected in the mixed fruiting branch types. Correlation analysis can only distinguish the degree of closeness between variables and cannot represent the existence of causality between variables [36], but it is a prerequisite for discovering causality. Wang et al. [37] found that there was a highly significant negative correlation between the defoliation rate and flocculating rate. Our results showed the same results, except that the correlation was unstable between defoliation rate and flocculation rate in the three times, and showed a high–low–high pattern with advancing post-dose time. This indicated that the appearance of the hanging phenomenon was related to the planting environment and cultivation method in addition to the characteristics of the varieties themselves [38]. The first fruit branch node and the first fruit branch height were positively correlated with the defoliation rate, respectively, a phenomenon that needs further verification. The 19 traits of Sea Island cotton were strongly correlated in the zero-type fruit branch, and it was easier to obtain the resources containing the target traits by selecting the zero-type fruit branch as the main body.

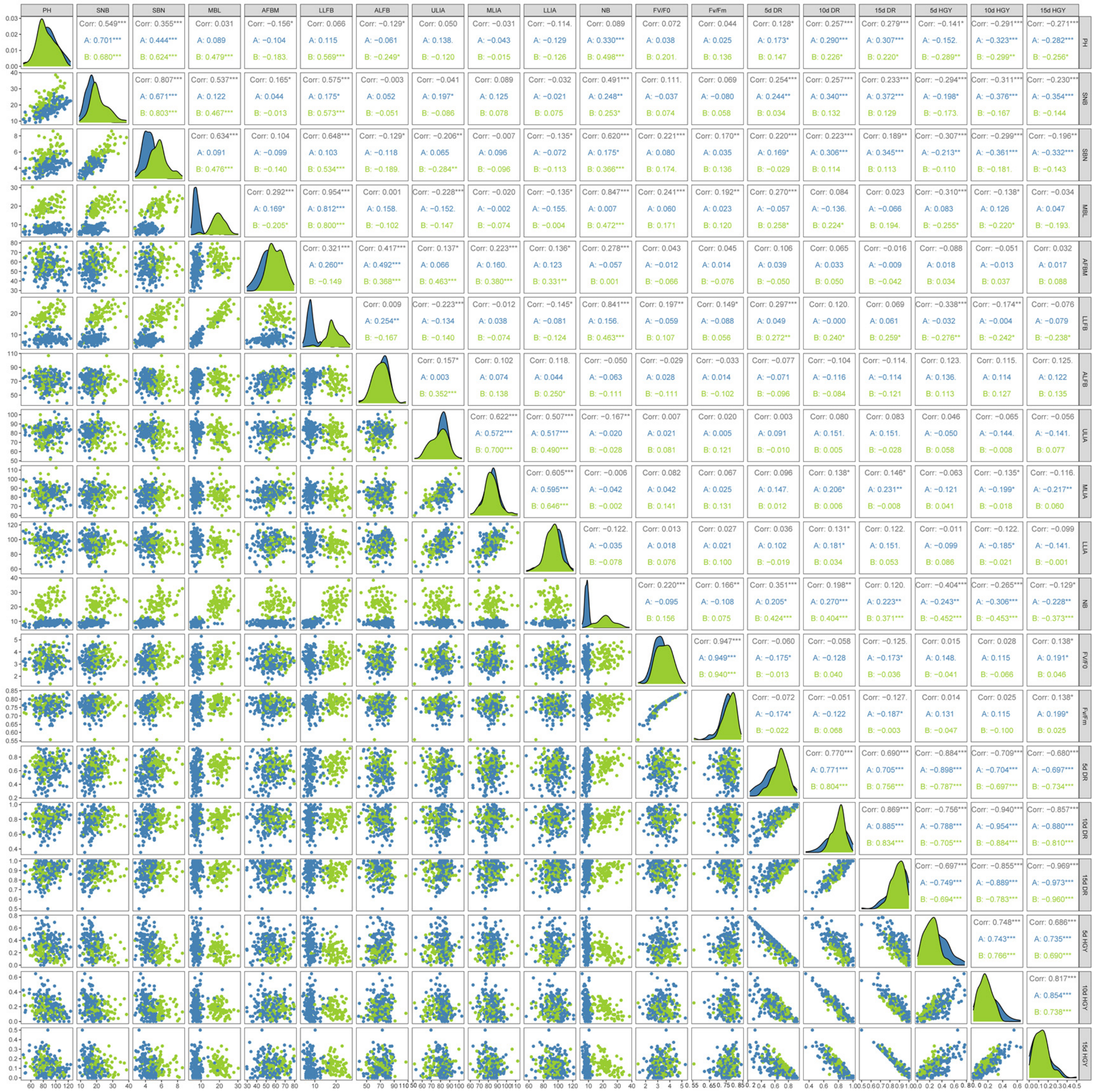

Figure 2. Correlations between machine harvested quality traits. A (blue part): zero fruiting branch type; B (green part): type I fruiting branch type. The scatterplot is at the bottom left, the data distribution is diagonal, and the correlation coefficient is at the top right; * Represents significant at the 0.1 level ($p < 0.1$), ** Represents significant at the 0.05 level ($p < 0.05$), *** Represents significant at the 0.01 level ($p < 0.01$).

3.3. *Principal Component Analysis of Machine-Harvesting Traits in Sea Island Cotton*

According to the principal component analysis of the standardized data, there were six principal components with eigenvalues greater than 1 (KMO = 0.765 > 0.6, $p = 0.000 < 0.05$), and their contribution rates were 30.89%, 19.09%, 12.24%, 9.36%, 7.5%, and 5.41% (Table 2). The cumulative contribution was 84.48%. Among them, principal component 1 was the defoliation factor, and the load of defoliation rate at different times is larger (0.33, 0.34, and 0.33). In principal component 5, loads of plant height (0.57), first fruiting branch height (0.47), first fruiting branch node (0.29), and 5 days hanging rate (0.22) were greater, indicating that principal component 5 was related to first fruit branch.

Table 2. Principal component analysis of 19 machine-harvesting traits.

Eigenvectors	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
PH	0.15	0.01	−0.09	0.26	0.57	0.28
SNB	0.25	0.23	0.04	−0.01	0.47	0.13
SBN	0.25	0.3	−0.03	0.05	0.29	−0.02
MBL	0.21	0.39	0.04	−0.16	−0.08	−0.16
AFBM	0.07	0.12	0.31	−0.36	−0.17	0.33
LLFB	0.23	0.38	0.04	−0.19	−0.05	−0.13
ALFB	−0.05	0.02	0.25	−0.31	−0.09	0.62
ULIA	−0.02	−0.19	0.47	0	0.22	−0.08
MLIA	0.05	−0.11	0.52	−0.03	0.16	−0.24
LLIA	0.01	−0.17	0.46	−0.03	0.09	−0.25
NB	0.25	0.33	0.04	−0.13	−0.08	−0.17
F_v/F_0	0.03	0.24	0.24	0.55	−0.24	0.12
F_v/F_m	0.02	0.22	0.24	0.56	−0.27	0.13
5 d DR	0.33	−0.13	−0.03	−0.06	−0.22	−0.04
10 d DR	0.34	−0.23	0	0.06	−0.09	0.08
15 d DR	0.33	−0.27	−0.04	0.05	−0.02	0.06
5 d HGY	−0.35	0.1	0.04	0.02	0.22	0.06
10 d HGY	−0.35	0.19	0	−0.07	0.05	−0.06
15 d HGY	−0.32	0.26	0.06	−0.05	0.02	−0.06
CV	5.87	3.63	2.33	1.78	1.43	1.03
CR (%)	30.89	19.09	12.24	9.36	7.5	5.41
CCR (%)	30.89	49.98	62.22	71.57	79.07	84.48

The utilization value of different fruit branch types resources was explored through dispersion mapping of the Sea Island cotton germplasm resource materials in six principal components [39] (Figure 3). In the PC1 defoliation factor, the distribution areas of the two types of fruit branch types formed a partial overlap. There are also some extreme materials with better and worse defoliation effects, where the zero-type fruit branch materials had a higher positive contribution to the defoliation factor. In PC2, the fruit branch length factor showed strong population stratification, and the zero-type fruit branch materials were mainly distributed in the negative contribution region. The principal component analysis is a method of data dimensionality reduction, which has the advantage of converting multiple data into several categories of factors that do not interfere with each other and can be described more objectively [40]. In this study, 19 traits were combined into 6 types of factors. When studying the contribution of population resources to each factor, we found that the positive contribution of the zero-type fruit branch resources was stronger for the defoliation factor, consistent with the results of the correlation analysis. The advantages and disadvantages of the two types of fruit branch resources in machine harvesting are different. The distribution of the principal components in the two-dimensional coordinate system can be chosen such that the materials are oriented in a certain factor.

The scores of the six principal components were ranked in descending order. The top 10 resource codes of mixed fruiting branch types Sea Island cotton were YueJin1, MoShi729, SYR cotton, И24−3386, LuoSaiNa, Antigua, Yue51-11, Giza80, Yunnan8040-2, and DaXuan71. The top 10 resource codes for the zero-type fruit branch Sea Island cotton were K-308, DJ9237, 16DJC01, X78, XH14, LuoSaiYa, 17-8, XH32, 572Q, and XH49 (Figure 4).

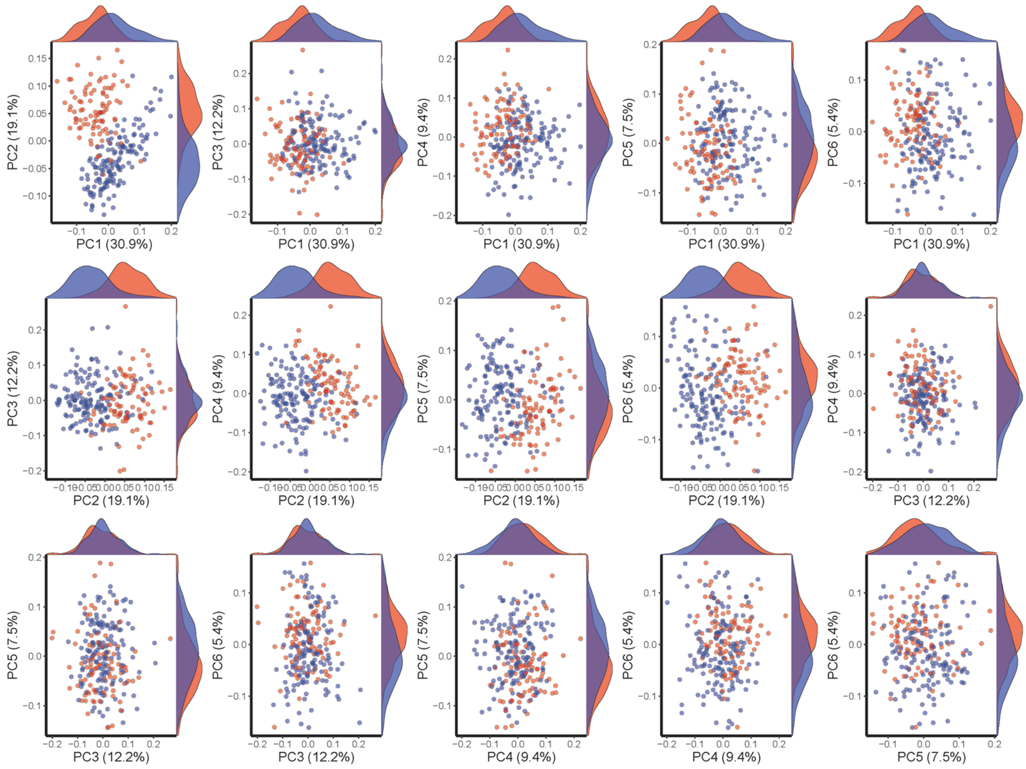

Figure 3. Principal component analysis of 240 Sea Island cotton varieties. Blue for zero fruiting branch type, red for long fruiting branch type.

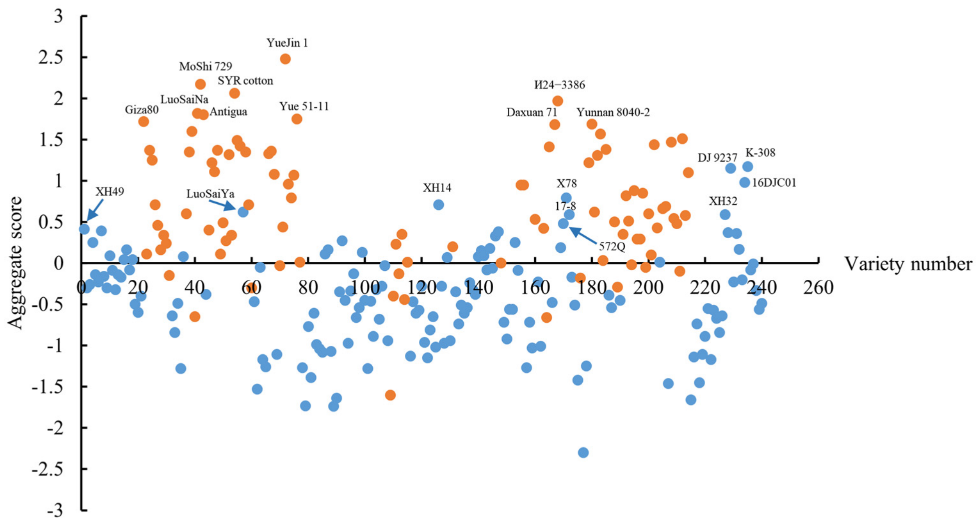

Figure 4. Comprehensive score for Sea Island cotton with different fruit branch types. Yellow for long fruiting branch types, blue for zero fruiting branch types.

3.4. Identification of Key Machine-Harvesting Traits and Construction of Regression Models

The correlation analysis of the composite score F value and 19 groups of traits was carried out to obtain 16 key traits with significant correlation (Figure 5), and then the stepwise regression equation was constructed with the F value as the independent variable and 16 groups of traits as the dependent variable, and, combined with the practical production significance, some traits were excluded, and the stepwise linear regression equation was finally obtained (VIF < 5): $Y = -7.2 + 0.01X_1 + 0.23X_2 + 0.192X_3 + 0.038X_4 + 0.007X_5 + 0.014X_7 + 0.025X_8 + 2.952X_{13}$ The eight groups of traits were X_1: PH, X_2: SNB, X_3: SBN, X_4: MBL, X_5: AFBM, X_7: MLIA, X_8: NB, X_{13}: 15 d DR. The equation was adjusted for $R^2 = 0.911$, $F = 308.114$, and the eight traits accounted for 91.1% of the variation in the composite scores. Based on the PCA results, the composite score was calculated and the regression mathematical model was constructed. The comprehensive evaluation method of resource population has been widely used in various crops, such as upland cotton [41], sorghum [42], peanut [43], etc. Chen [44] found that defoliation rate, first branch node position, and total leaf number could be used as the core indexes for evaluating the machine-harvesting traits of upland cotton, which was consistent with our results. It was found that 20 materials such as Moshi729 performed better in terms of the first fruit branch height and defoliation rate, and had the potential to be a genetically improved parent of Sea Island cotton with machine harvesting as the breeding direction. The traits used in regression model construction in this study are easy to obtain in the field, which is suitable for rapid preliminary identification of breeding materials, can greatly shorten the selection time of conventional breeding parents, and reduce the probability of missing dominant single plants in the field while avoiding the human subjective consciousness of the material characteristics.

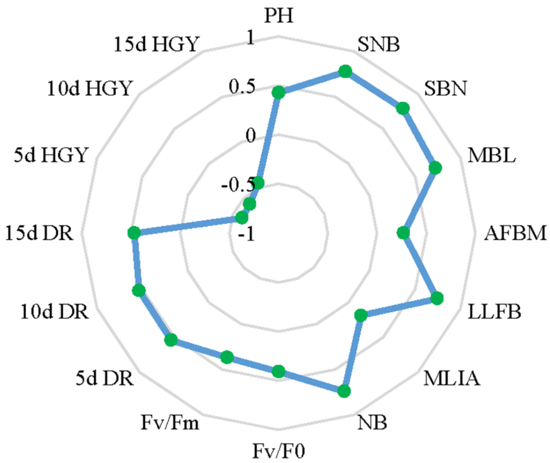

Figure 5. Correlation coefficients and significant levels ($p < 0.01$) of 16 key machine-harvesting traits.

3.5. Cluster Analysis of Sea Island Cotton Germplasm Resources

The 240 sea island cotton germplasm resources were clustered into 4 categories (Figure 6), of which 47 materials were in the first category, accounting for 19.58%. The second category contained 58 materials (24.17%). The third category included 59 materials (24.58%). The fourth category had 76 materials (31.67%).

Sea Island cotton material suitable for machine harvesting should be characterized by the high height of the first fruiting branch, moderate plant looseness, and good defoliation. By calculating the mean values of the main machine-harvesting indicators, the characteristic markers of each classified resource were derived (Table 3). The first category had low first fruiting branch height (17.36), low defoliation rate (58% at 5 days, 74% at 10 days, and 85% at 15 days), and obvious hanging fruits (37% at 5 days, 25% at 10 days, and 17% at 15 days), and the lower first fruiting branch height would affect the lower bolls' harvesting and result in yield loss, while the low defoliation rate would reduce the quality of the harvesting, and therefore belonged to the germplasm resources with poor machine-harvesting performance (SuB51, AK3836, BZ266, Ba3021, AK4154, etc.). The main characteristics of the second category were high first fruiting branch (20.62), high first fruiting branch node (5.62), increased fruiting branch node (mid-fruiting branch angle of 50.82, lower fruiting branch angle of 63.54), compact plant (upper leaf inclination of 76.39, mid-leaf inclination of 80.68, and lower leaf inclination of 88.68), and better defoliation (5-day defoliation rate of 66%, 10-day defoliation). It has a good defoliation effect (5-day defoliation rate of 66%, 10-day defoliation rate of 81%, and 15-day defoliation rate of 91%), and it is not easy to form hanging branches (5-day branching rate of 25%, 10-day branching rate of 15%, and 15-day branching rate of 11%), so the good plant structure can increase the net rate of machine harvesting, and the high defoliation rate also ensures the quality of harvesting, and therefore it belongs to the germplasm resources with strong machine-harvesting ability (Moshi 729 et al.).

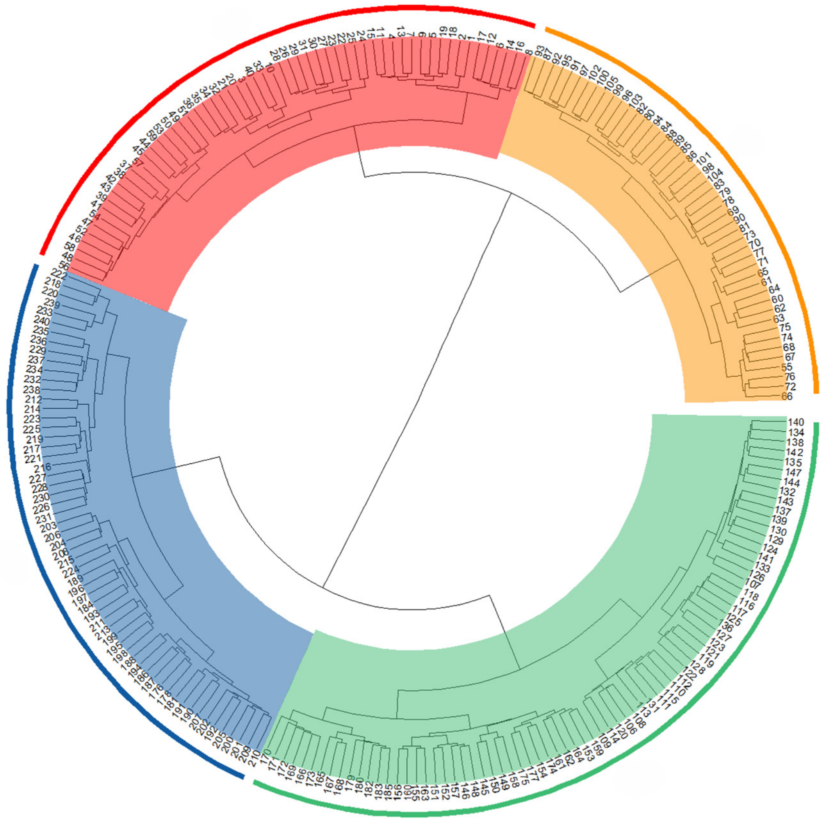

Figure 6. Cluster analysis of 240 Sea Island cotton varieties. Yellow for taxon I, red for taxon II, blue for taxon III, green for taxon IV.

Table 3. Characteristics of the four groups of phenotypic data.

Phenotype	Group I	Group II	Group III	Group IV
PH	84.95 b	95.8 a	80.41 b	84.71 b
SNB	17.36 b	20.62 a	18.3 b	18.71 b
SBN	4.47 b	5.62 a	4.77 b	4.41 b
MBL	10.60 b	14.38 a	13.34 a	10.68 b
AFBM	54.44 b	50.82 c	57.05 b	60.98 a
LLFB	10.34 c	14.32 a	12.88 ab	11.36 bc
ALFB	71.08 a	63.54 b	74.61 a	72.65 a
ULIA	81.12 a	76.39 b	80.63 a	82.35 a
MLIA	83.86 a	80.68 b	83.08 ab	85.15 a
LLIA	95.27 a	88.68 b	94.97 a	96.49 a
NB	13.20 b	16.76 a	13.48 b	12.34 b
F_v/F_0	3.51 ab	3.30 b	3.65 a	3.04 c
F_v/F_m	0.77 ab	0.76 bc	0.78 a	0.74 c
5 d DR	0.58 b	0.66 a	0.61 b	0.64 b
10 d DR	0.74 b	0.81 a	0.78 ab	0.81 a
15 d DR	0.85 b	0.91 a	0.87 b	0.90 a
5 d HGY	0.37 a	0.25 b	0.31 b	0.27 b
10 d HGY	0.25 a	0.15 c	0.21 ab	0.17 bc
15 d HGY	0.17 a	0.11 b	0.15 a	0.11 b

Note: different letters (a, b, and c) represent significant differences ($p < 0.05$).

4. Conclusions

In this study, we have found a rich diversity in plant types and defoliation effects in Sea Island cotton through the analysis of a large number of machine-harvesting traits from the germplasm resources. Combining correlation analysis, principal component analysis, and clustering analysis, key machine-harvesting traits were identified, relevant regression models were established, and excellent materials suitable for machine harvesting were selected. This study is important for the application of large-scale machine harvesting in Sea Island cotton and for the improvement of Sea Island cotton varieties suitable for machine harvesting. However, the study of phenotypic traits still has some limitations to some extent. Therefore, further in-depth research combined with molecular means is needed for better resource evaluation.

Author Contributions: Conceptualization, G.C. and J.K.; Formal Analysis, M.W. and J.Y.; Investigation, Y.Z. and F.L.; Methodology, G.C. and J.K.; Resources, J.K. and W.W.; Software, F.L.; Writing—Original Draft, F.L. and M.W.; Writing—Review and Editing, F.L., M.W., N.Z., Y.Z., J.L., S.W., A.A., G.C. and J.K. All authors have read and agreed to the published version of the manuscript.

Funding: This paper was supported by the National Key Research and Development Plan of China (No. 2022YFD1200301), China Agriculture Research System (CARS-15-09), Multi-resistant and Machine-harvested Cotton Bio-breeding Innovation Team (2023TSYCTD0002), China Agricultural University-Tarim University Joint Research Fund (ZNLH202301), and the Guiding Science and Technology Plan Projects of Xinjiang Corps (2023ZD103).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

Acknowledgments: Thanks to Yanhui Zhang and Hezhong Dong for their guidance on writing this article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Qin, Y.; Sun, M.; Li, W.; Xu, M.; Shao, L.; Liu, Y.; Zhao, G.; Liu, Z.; Xu, Z.; You, J. Single-cell RNA-seq reveals fate determination control of an individual fiber cell initiation in cotton (*Gossypium hirsutum*). *Plant Biotechnol. J.* **2022**, *20*, 2372–2388. [CrossRef]
2. Fang, L.; Zhao, T.; Hu, Y.; Si, Z.; Zhu, X.; Han, Z.; Liu, G.; Wang, S.; Ju, L.; Guo, M. Divergent improvement of two cultivated allotetraploid cotton species. *Plant Biotechnol. J.* **2021**, *19*, 1325–1336. [CrossRef]
3. Su, X.; Zhu, G.; Song, X.; Xu, H.; Li, W.; Ning, X.; Chen, Q.; Guo, W. Genome-wide association analysis reveals loci and candidate genes involved in fiber quality traits in sea island cotton (*Gossypium barbadense*). *BMC Plant Biol.* **2020**, *20*, 289. [CrossRef]
4. Yu, J.; Zhang, K.; Li, S.; Yu, S.; Zhai, H.; Wu, M.; Li, X.; Fan, S.; Song, M.; Yang, D. Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum* × *Gossypium barbadense* backcross inbred line population. *Theor. Appl. Genet.* **2013**, *126*, 275–287. [CrossRef]
5. Ma, L.; Zhao, Y.; Wang, Y.; Shang, L.; Hua, J. QTLs Analysis and Validation for Fiber Quality Traits Using Maternal Backcross Population in Upland Cotton. *Front. Plant Sci.* **2017**, *8*, 2168. [CrossRef]
6. Hong, M.; Zhang, L.; Chen, J.; Lu, Z.; Fan, A.; Tursun, T.; Ma, Z.; Xie, L. Practice and Thinking on Mechanical Harvest of Long Staple Cotton in Xinjiang. *China Cotton* **2020**, *47*, 8–10. [CrossRef]
7. Heinicke, C.; Grove, W.A. "Machinery Has Completely Taken Over": The Diffusion of the Mechanical Cotton Picker, 1949–1964. *J. Interdiscip. Hist.* **2008**, *39*, 65–96. [CrossRef]
8. Wu, J.; Chen, X. Present situation, problems and countermeasures of cotton production mechanization development in Xinjiang Production and Construction Corps. *J. Agric. Eng.* **2015**, *31*, 6.
9. Su, J.; Ma, Q.; Li, M.; Hao, F.; Wang, C. Multi-Locus Genome-Wide Association Studies of Fiber-Quality Related Traits in Chinese Early-Maturity Upland Cotton. *Front. Plant Sci.* **2018**, *9*, 1169. [CrossRef]
10. Zhang, T.; Hu, Y.; Jiang, W.; Fang, L.; Guan, X.; Chen, J.; Zhang, J.; Saski, C.A.; Scheffler, B.E.; Stelly, D.M. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nat. Biotechnol.* **2015**, *33*, 531–537. [CrossRef]
11. Kong, Q. The analysis of comparative advantage of production situation for sea island cotton (*G. barbadense* L.) in China. *China Cotton* **2002**, *29*, 19–23. [CrossRef]

12. Qi, H.; Wang, N.; Qiao, W.; Xu, Q.; Zhou, H.; Shi, J.; Yan, G.; Huang, Q. Construction of a high-density genetic map using genotyping by sequencing (GBS) for quantitative trait loci (QTL) analysis of three plant morphological traits in upland cotton (*Gossypium hirsutum* L.). *Euphytica* **2017**, *213*, 83. [CrossRef]
13. Shi, J.; Wang, N.; Zhou, H.; Xu, Q.; Yan, G. Transcriptome analyses provide insights into the homeostatic regulation of axillary buds in upland cotton (*G. hirsutum* L.). *BMC Plant Biol.* **2020**, *20*, 228. [CrossRef] [PubMed]
14. Chandel, R.; Sharma, K. Multiple Attributed Parametric Review Study on Mechanical Cotton (*Gossypium hirsutum* L.) Harvesters. *J. Agric. Sci.* **2022**, *142*, 122. [CrossRef]
15. Cheng, X.; Feng, Y.; Chen, D.; Luo, C.; Yu, X.; Huang, C. Evaluation of Rosa germplasm resources and analysis of floral fragrance components in *R. rugosa*. *Front. Plant Sci.* **2022**, *13*, 1026763. [CrossRef] [PubMed]
16. Egan, L.M.; Conaty, W.C.; Stiller, W.N. Core Collections: Is There Any Value for Cotton Breeding? *Front. Plant Sci.* **2022**, *13*, 895155. [CrossRef]
17. Ferguson, A.R. The need for characterization and evaluation of germplasm: Kiwifruit as an example. *Euphytica* **2007**, *154*, 371–382. [CrossRef]
18. Khan, A.S.; Ali, S.; Khan, I.A. Morphological and molecular characterization and evaluation of mango germplasm: An overview. *Sci. Hortic.* **2015**, *194*, 353–366. [CrossRef]
19. Rodríguez, V.M.; Romay, M.C.; Ordás, A.; Revilla, P. Evaluation of European maize (*Zea mays* L.) germplasm under cold conditions. *Genet. Resour. Crop Evol.* **2010**, *57*, 329–335. [CrossRef]
20. Yan, W.; Rutger, J.N.; Bryant, R.J.; Bockelman, H.E.; Fjellstrom, R.G.; Chen, M.H.; Tai, T.H.; McClung, A.M. Development and Evaluation of a Core Subset of the USDA Rice Germplasm Collection. *Crop Sci.* **2007**, *47*, 869–876. [CrossRef]
21. Upadhyaya, H.D.; Swamy, B.P.M.; Goudar, P.V.K.; Kullaiswamy, B.Y.; Singh, S. Identification of diverse groundnut germplasm through multienvironment evaluation of a core collection for Asia. *Field Crop Res.* **2005**, *93*, 293–299. [CrossRef]
22. Song, Q.; Liu, C.; Bachir, D.G.; Chen, L.; Hu, Y. Drought resistance of new synthetic hexaploid wheat accessions evaluated by multiple traits and antioxidant enzyme activity. *Field Crop Res.* **2017**, *210*, 91–103. [CrossRef]
23. Huang, K.; Dai, X.; Xu, Y.; Dang, S.; Shi, T.; Sun, J.; Wang, K. Relation between level of autumn dormancy and salt tolerance in lucerne (*Medicago sativa*). *Crop Pasture Sci.* **2018**, *69*, 194. [CrossRef]
24. Tian, Z.; Yang, Y.; Wang, F. A comprehensive evaluation of heat tolerance in nine cultivars of marigold. *Hortic. Environ. Biotechnol.* **2015**, *56*, 749–755. [CrossRef]
25. Li, X.; Ma, H.; Wang, L.; Aimerjiang, A.; Guo, R.; Zhong, J.; Ou, H.; Dai, L.; Zhang, Y.; Xi, Y. Effects of spraying different times and dosages of defoliants on defoliation and ripening of long staple cotton in Xinjiang. *China Cotton* **2022**, *49*, 6–9. [CrossRef]
26. Chen, J.; Zhang, L.; Hong, M.; Fan, A.; Tursun, T.; Ma, Z.; Xie, L. The Suitable for Integrated Mechanically Harvested Cultivation Technical of Long-fiber Cotton in Xinjiang. *Cotton Sci.* **2021**, *43*, 64–68.
27. Zhang, L.; Fan, A.; Hong, M.; Ma, Z.; Chen, J.; Zhao, S.; Zheng, K.; Telson, T. Genetic diversity analysis of 647 sea island cotton germplasm resources. *J. Plant Genet. Resour.* **2023**, *24*, 307–324. [CrossRef]
28. Dai, P.; Sun, J.; He, S.; Wang, L.; Jia, Y.; Pan, Z.; Pang, B.; Du, X.; Wang, Q. Comprehensive evaluation and genetic diversity analysis of phenotypic traits of core collection in upland cotton. *Sci. Agric. Sin.* **2016**, *49*, 3694–3708. [CrossRef]
29. Martínez-Castillo, J.; Colunga-GarcíaMarín, P.; Zizumbo-Villarreal, D. Genetic erosion and in situ conservation of Lima bean (*Phaseolus lunatus* L.) landraces in its Mesoamerican diversity center. *Genet. Resour. Crop Evol.* **2008**, *55*, 1065–1077. [CrossRef]
30. Pfeiffer, V.W.; Ford, B.M.; Housset, J.; McCombs, A.; Blanco Pastor, J.L.; Gouin, N.; Manel, S.; Bertin, A. Partitioning genetic and species diversity refines our understanding of species–genetic diversity relationships. *Ecol. Evol.* **2018**, *8*, 12351–12364. [CrossRef] [PubMed]
31. Rahman, M.U.; Shaheen, T.; Tabbasam, N.; Iqbal, M.A.; Ashraf, M.; Zafar, Y.; Paterson, A.H. Cotton genetic resources. A review. *Agron. Sustain. Dev.* **2012**, *32*, 419–432. [CrossRef]
32. Pan, Z.; He, S.; Jia, Y.; Podolnaya, L.P.; Sun, J.; Wang, L.; Du, X. Genetic Diversity Analysis of the Sea Island Cotton (*Gossypium barbadense* L.) Mostly from Former USSR Using SSR Markers. *J. Plant Genet. Resour.* **2014**, *15*, 399–404. [CrossRef]
33. Li, J. Loci Associated with Defoliant Sensitivity of Upland Cotton were Identified by Association Analysis. Master's Thesis, Tarim University, Alar, China, 2021.
34. Wang, T.; Zhou, Y.; Liu, C.; Li, J.; He, L. Screening of Upland Cotton Germplasm Resources Which Are Sensitive to Defoliant in Southern Xinjiang. *Xinjiang Agric. Sci.* **2019**, *56*, 2189–2198. [CrossRef]
35. Yu, H. Genetic Diversity and Genome-Wide Association Analysis of Fiber Quality and Yield Traits of Self-Bred Island Cotton Cultivars in Xinjiang. Master's Thesis, Zhejiang University, Hangzhou, China, 2019.
36. Baldovin, M.; Cecconi, F.; Vulpiani, A. Understanding causation via correlations and linear response theory. *Phys. Rev. Res.* **2020**, *2*, 043436. [CrossRef]
37. Wang, J.; Zhang, Z.; Zhang, N.; Liang, Y.; Gong, Z.; Wang, J.; Ditta, A.; Sang, Z.; Li, X.; Zheng, J. The Correlation of Machine-Picked Cotton Defoliant in Different *Gossypium hirsutum* Varieties. *Agronomy* **2023**, *13*, 2151. [CrossRef]
38. Li, L. Effects of Row Spacing on Canopy Microenvironment and Yield Formation of Cotton of Different Plant Types. Master's Thesis, Xinjiang Agricultural University, Urumqi, China, 2021.
39. Molania, R.; Foroutan, M.; Gagnon-Bartsch, J.A.; Gandolfo, L.C.; Jain, A.; Sinha, A.; Olshansky, G.; Dobrovic, A.; Papenfuss, A.T.; Speed, T.P. Removing unwanted variation from large-scale RNA sequencing data with PRPS. *Nat. Biotechnol.* **2022**, *41*, 82–95. [CrossRef] [PubMed]

40. Wold, S.; Esbensen, K.; Geladi, P. Principal component analysis. *Chemom. Intell. Lab.* **1987**, *2*, 37–52. [CrossRef]
41. Wang, X.; Xing, A.; Yang, R.; He, S.; Jia, Y.; Pan, Z.; Wang, L.; Du, X.; Song, X. Comprehensive Evaluation of Phenotypic Characters of Nature Population in Upland Cotton. *Sci. Agric. Sin.* **2022**, *55*, 1082–1094. [CrossRef]
42. Zhao, N.; Wang, W.; Jiang, K.; Grover, C.E.; Cheng, C.; Pan, Z.; Zhao, C.; Zhu, J.; Li, D.; Wang, M. A Calmodulin-Like Gene (GbCML7) for Fiber Strength and Yield Improvement Identified by Resequencing Core Accessions of a Pedigree in *Gossypium barbadense*. *Front. Plant Sci.* **2022**, *12*, 815648. [CrossRef]
43. Zhang, H.; Jiang, C.; Yin, D.; Dong, J.; Ren, Q.; Zhao, X.; Zhong, C.; Wang, X.; Yu, H. Establishment of comprehensive evaluation system for cold tolerance and screening of cold-tolerance germplasm in peanut. *Acta Agron. Sin.* **2021**, *47*, 1753–1767. [CrossRef]
44. Chen, X. Study on Comprehensive Evaluation of Mechanical Performance of Different Cotton Varieties in Northern Xinjiang. Master's Thesis, Xinjiang Agricultural University, Urumqi, China, 2020.