

RESEARCH

Open Access



# Evaluation of the genetic diversity of fibre quality traits in upland cotton (*Gossypium hirsutum* L.) inferred from phenotypic variations

SUN Zhengwen, WANG Xingfen, LIU Zhengwen, GU Qishen, ZHANG Yan, LI Zhikun, KE Huifeng, YANG Jun, WU Jinhua, WU Liqiang, ZHANG Guiyin and MA Zhiying\*

## Abstract

**Background:** Evaluating phenotypic traits is very important for the selection of elite lines in *Gossypium hirsutum* L. Cotton breeders are interested in using diverse genotypes in hybridization that can segregate for traits of interest with the possibility of selection and genetic gain. Information on phenotypic and molecular diversity helps the breeders for parental selection.

**Methods:** In this study, 719 global collections of *G. hirsutum* L. were evaluated for five fibre-related traits during two consecutive years in eight different environments. A series of phenotypic data for fibre quality traits were obtained and the elite accessions were further screened using principal component analysis and phylogenetic tree construction based on single nucleotide polymorphism markers.

**Results:** We found that fibre quality traits showed a wide range of variation among the *G. hirsutum* accessions over 2 years. In general, accessions from outside China tended to have higher fibre length (FL) and fibre strength (FS) than did Chinese accessions. Among different regional accessions in China, North/Northwest accessions tended to have the highest FL, FS and best fibre micronaire. By assessing five fibre quality traits over 2 years with genotypic data, 31 elite germplasms reaching double-thirty quality values ( $FL \geq 30$  mm and  $FS \geq 30$  cN·tex<sup>-1</sup>) were selected.

**Conclusions:** This study provided a detailed phenotypic diversity description of a population representing a wide range of upland cotton germplasm. Our findings provide useful information about possible elite fibre quality parents for cotton breeding programs.

**Keywords:** *Gossypium hirsutum* L., Fibre quality, Morphological traits, Genetic diversity, Double-thirty

## Background

Cotton is one of the major fibre crops worldwide, and has extensive phenotypic diversity among the 50 representative species (Wendel and Cronn 2003). Among these species, upland cotton (*Gossypium hirsutum* L.) is the leading fibre production crop and grown in more than 80 countries/regions of the world. Currently, *G. hirsutum* is responsible for 95% of the annual cotton production in the world (Chen et al. 2007). Because of

its economic importance, such as high-yield and environmental suitability, *G. hirsutum* has attracted considerable scientific interest of plant breeders and agricultural scientists.

However, domesticated upland cotton genotypes have narrow genetic base (Abdurakhmonov et al. 2008; Tyagi et al. 2014). To broaden the genetic base through hybrid breeding programs, the genetic divergence among available germplasm is a prerequisite. Genetic variation among different cotton genotypes for morphological and fibre quality traits was studied for their improvement. Cultivated cotton was domesticated in Mesoamerica and has a low level of genetic

\* Correspondence: [mzhy@hebau.edu.cn](mailto:mzhy@hebau.edu.cn)

North China Key Laboratory for Crop Germplasm Resources of Ministry of Education/Key Laboratory for Crop Germplasm Resources of Hebei Province, Hebei Agricultural University, Baoding 071001, China



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

polymorphism (Rungis et al. 2005; Ai et al. 2017). The narrow genetic base of upland cotton has become a serious concern since limited genetic diversity translates to limited allelic availability for continued genetic gain. So, it is crucial to explore novel germplasm resources for potential natural genetic diversity and develop innovative genomics tools to efficiently mobilize these useful genetic variations to breeding germplasm, which should help to overcome existing and potential problems of cotton production associated with narrow genetic base of the cultivar germplasm (Ulloa et al. 2013; Hinze et al. 2017).

Variation in germplasm collections has been utilized for identifying desirable genotypes to enhance yield improvement. To identify the desired genotype, various morphological traits are being employed such as fibre quality, yield component factors, resistance, seed quality (Abdurakhmonov et al. 2008; Zhao et al. 2014; Badiganavar and Myers 2015; Huang et al. 2017; Sun et al. 2017). These distinct qualitative traits, called morphological markers, are often reliable for germplasm characterization and trait associated selections; such qualitative traits are stable in expression across various environments. Thus, characterization of germplasm plays a vital role in crop improvement.

Evaluation of germplasm and quantification of genetic diversity are indispensable for a pragmatic use of plant genetic resources, also for determining evolutionary relationships (Amezrou et al. 2017). Studies of the variation present in germplasm collections have been carried out by employing plant morphological attributes as characterization tools (Zhang et al. 2012; Salazar et al. 2016). For cotton geneticists and breeders, the precise evaluation of the genetic diversity of excellent *G. hirsutum* germplasm will provide a guide for choosing parents and predicting the degree of inheritance, variation, and level of heterosis, which are essential for realizing the breeding goals (Hinze et al. 2015).

To better understand and effectively utilize *G. hirsutum* germplasm in China, it is necessary to evaluate global collections of the species for phenotypic traits and to study their genetic variation in different environments. The objectives of the present study were: (i) to assess the phenotypic and genotypic variation of five fibre quality traits in 719 global collections of *G. hirsutum*; (ii) to compare variation of phenotypic traits of accessions from the different regions; (iii) to detect and analyze correlations among the investigated traits; and (iv) to group the accessions using principal component analysis (PCA) and phylogenetic tree based on single nucleotide polymorphism (SNP) markers. Our findings could provide useful information for selecting elite parental materials for breeding programs of upland cotton.

## Methods

### Plant materials and field trials

A collection of 719 upland cotton germplasms was evaluated in this study. These accessions were from different countries/regions, 588 were collected from China and 131 were from other countries/regions (Additional file 1: Table S1). The accessions were grown in eight different environments in a randomized complete block at Baoding (115°47'N, 38°87'E), Hejian (116°13'N, 38°42'E), Xinji (115°12'N, 37°54'E), and Qingxian (116°91'N, 38°65'E) in Hebei Province and Yacheng (109°20'N, 18°38'E) in Hainan Province in 2014, denoted 14BD, 14HJ, 14XJ, 14QX and 14HN, and Xinji and Qingxian in Hebei Province and Yacheng in Hainan Province in 2015, denoted 15XJ, 15QX and 15HN, respectively. Two replicates were performed for each accession in five locations in 2014, and three replicates were performed for the three locations in 2015. Briefly, one row of each accession was planted for each replicate, with 20–22 plants per row, 30–35 cm between plants within rows and 80 cm between rows.

### Measurement of fibre quality traits and data collection

When mature, 20 naturally open bolls from the central part of the plants from each accession were hand-harvested at each location and ginned by machine. Fibre samples were sent to the Supervision and Testing Center of Cotton Quality, Ministry of Agriculture of China in Anyang, Henan Province for fibre property determination. Fibre quality traits, including the upper-half mean fibre length (FL, mm), fibre strength (FS, cN·tex<sup>-1</sup>), fibre micronaire (FM), fibre uniformity (FU, %) and fibre elongation (FE, %), were measured using a high volume instrument (HVI).

### Statistical analysis

The data for the phenotypic characters were analyzed by determining the mean, minimum, maximum, standard deviation (SD) and frequency distribution. The relationships among traits were calculated using Pearson's correlation analysis for the accession means using SPSS 22.0 software.

### Cluster based on phenotypes and genotypes

The accessions were classified by principal component analysis (PCA) on mean values of all five-quality traits among different environments. PCA was performed using the OmisShare tools ([www.omicshare.com/tools](http://www.omicshare.com/tools)). Phylogenetic tree among individual SNPs (Sun et al. 2017) was constructed by calculating Nei's genetic distance using PowerMarker version 3.25 (Liu and Muse 2005) and was visualized by iTOL (<http://itol.embl.de/>).

## Results

### Characterization of five fibre quality traits

We found a wide range of variation in the five quality traits among the evaluated accessions (Fig. 1). For example, FL varied from 23.68 mm to 33.74 mm; FS varied from 25.02 cN·tex<sup>-1</sup> to 33.94 cN·tex<sup>-1</sup>; FM varied from 3.65 to 5.97; FU varied from 81.69% to 86.65%; FE varied from 6.21% to 7.13%.

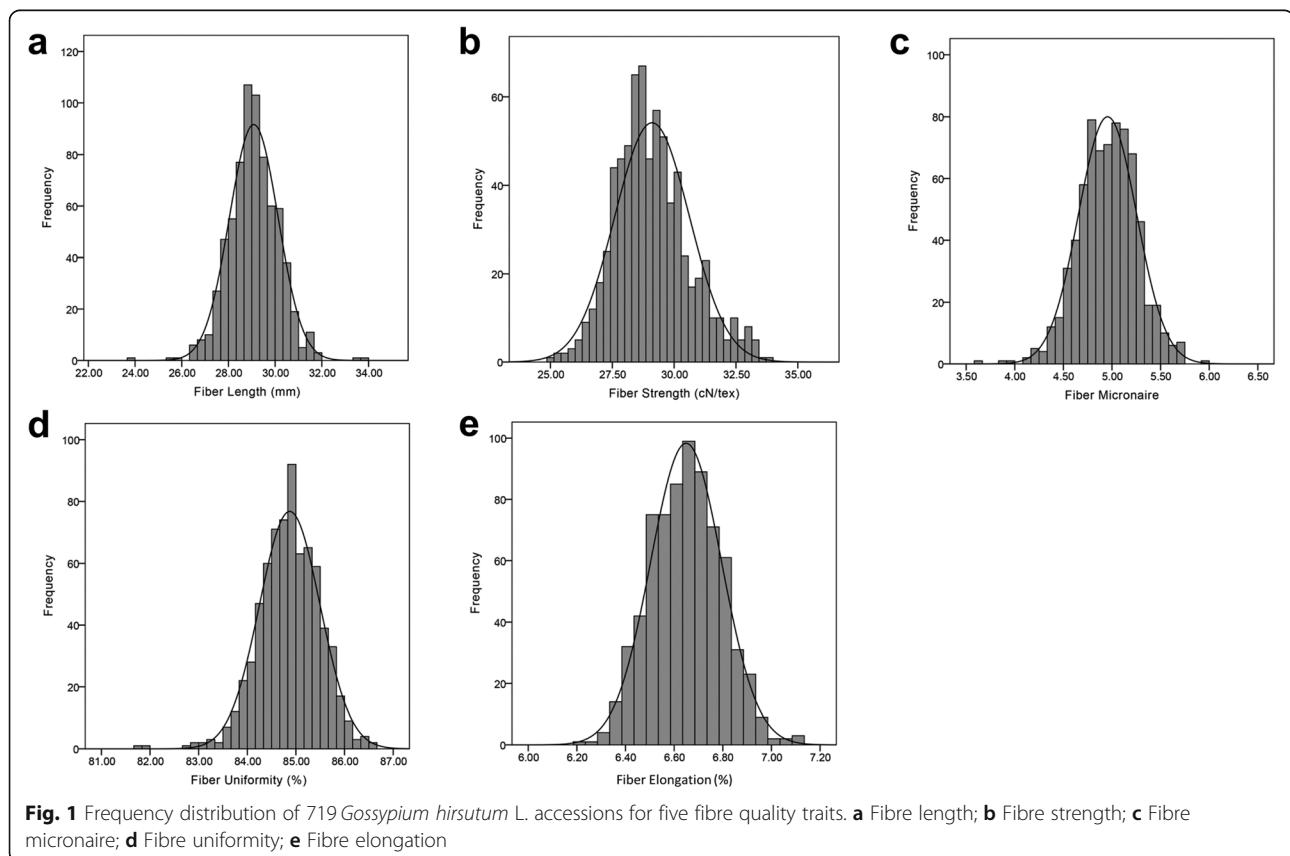
In this study, all cotton accessions were assigned four different germplasm types, namely the Yellow River (406), the Yangtze River (123), North/Northwest China (59) and Abroad (131). The means and variation ranges of the five quality traits of accessions from different regions across the two years are presented in Table 1. Through comparison analysis, we found that the abroad accessions tended to have the highest FL (29.51 mm and 28.66 mm) in 2 years, the highest FS (30.05 cN·tex<sup>-1</sup>), highest FM (4.99) and highest FU (85.11%) in 2014, compared with the other three germplasm types accessions (Table 1). Additionally, our results demonstrated that the North/Northwest accessions in China tended to have higher FL, higher FS, lower FM and lower FE than did the other ecotype accessions in 2 years, such as those from the Yellow River and the Yangtze River.

### Correlation between phenotypic traits

Correlation coefficients ( $r$ ) were highly significant ( $P < 0.01$ ) in 540 of the 780 trait combinations, where  $r$  ranged from  $-0.398$  to  $1$  (Fig. 2). The correlation coefficient and  $P$  values among the traits under investigation are presented in Additional file 1: Table S2. The objective was to explore which traits are well associated and meaningful for breeding. The correlations of each trait were largely consistent between different locations in 2 years. Overall, FL had high positive correlation coefficients with FS, FU and FE, whereas FS also had high positive correlations with FU and FE. We also found negative correlations between FM and FL, FS, FU. Of the above correlation coefficients, FL<sub>14BD</sub> was positively and significantly correlated with FS<sub>14BD</sub> ( $r = 0.784^{**}$ ), and FL<sub>14HJ</sub> had a significantly positive correlation with FE<sub>14HJ</sub> ( $r = 0.734^{**}$ ). There were negative and significant correlations between FM<sub>14HN</sub> and FL<sub>14HN</sub> ( $r = -0.398^{**}$ ) as well as between FM<sub>14BD</sub> and FS<sub>14BD</sub> ( $r = -0.280^{**}$ ). The results showed that the higher FL and FS, the lower FM.

### Mining germplasm resources with excellent quality traits

Through evaluating the consistency of the data between the two years, analysis of the means of 719 *G. hirsutum*

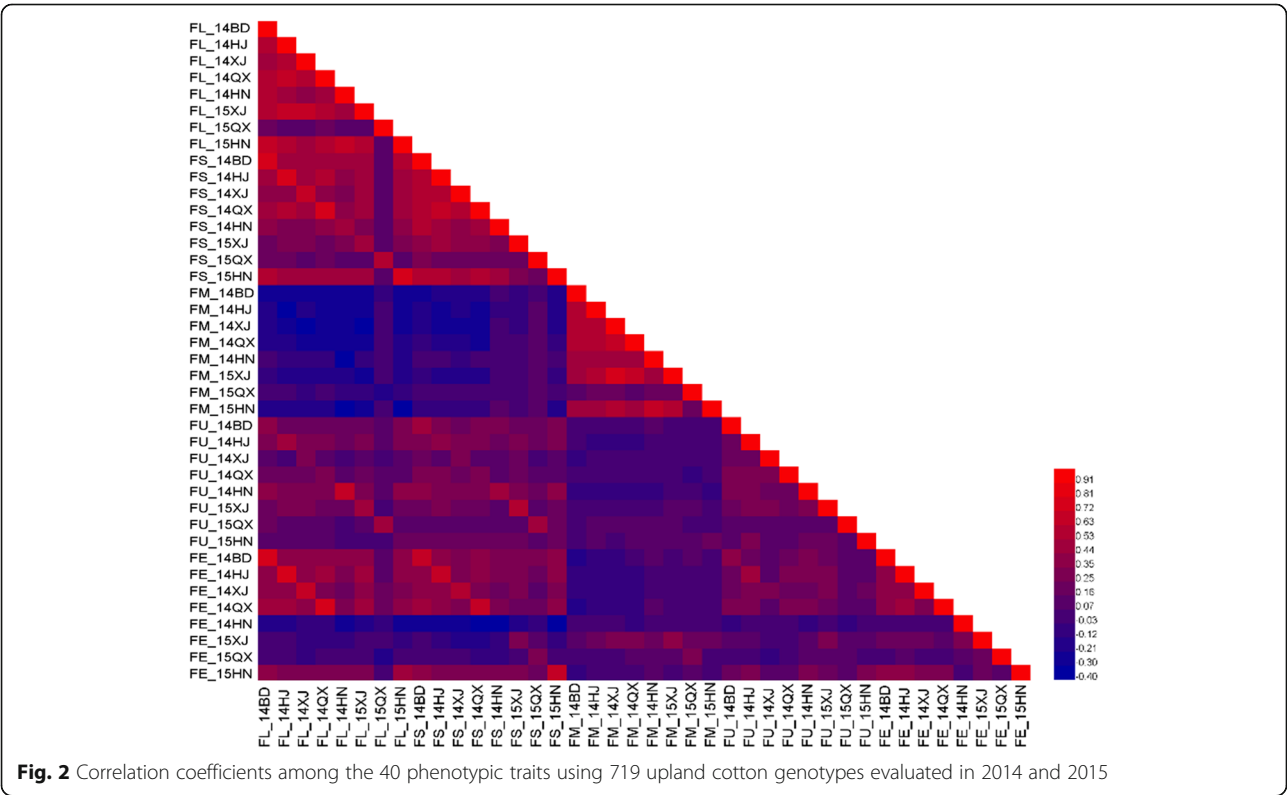


**Table 1** Descriptive statistics of the five measured traits in 719 upland cotton accessions from different regions

Trait	Year	The Yellow River region		The Yangtze River region		North/Northwest China		Abroad	
		Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range
FL	2014	29.36 ± 1.528	22.07–34.65	29.31 ± 1.534	24.50–33.39	29.46 ± 1.561	26.09–33.43	29.51 ± 1.535	25.14–35.56
	2015	28.53 ± 1.974	22.74–34.82	28.63 ± 1.975	23.17–32.85	28.64 ± 1.999	22.10–34.16	28.66 ± 1.975	22.00–36.14
FS	2014	29.37 ± 2.365	23.10–36.10	29.43 ± 2.365	23.30–36.30	30.01 ± 2.379	23.40–36.30	30.05 ± 2.373	23.30–36.80
	2015	28.30 ± 3.005	22.04–38.02	28.19 ± 2.996	22.81–36.18	28.62 ± 3.028	20.82–37.03	28.59 ± 3.019	18.79–36.25
FM	2014	4.98 ± 0.495	3.28–6.34	4.97 ± 0.488	3.39–5.93	4.94 ± 0.494	3.75–6.18	4.99 ± 0.491	3.14–6.12
	2015	4.93 ± 0.491	3.16–6.26	4.91 ± 0.493	3.24–6.01	4.87 ± 0.501	3.33–6.00	4.90 ± 0.491	3.33–6.30
FU	2014	85.04 ± 1.170	78.90–88.70	84.96 ± 1.175	81.40–88.30	84.99 ± 1.179	81.70–88.10	85.11 ± 1.174	81.50–89.00
	2015	84.63 ± 1.738	77.80–88.59	84.62 ± 1.740	78.39–88.07	84.58 ± 1.744	78.34–88.10	84.49 ± 1.743	71.15–88.13
FE	2014	6.60 ± 0.524	3.55–8.65	6.65 ± 0.528	4.00–8.55	6.52 ± 0.546	4.05–7.30	6.57 ± 0.526	3.95–8.20
	2015	6.75 ± 0.437	5.52–7.95	6.73 ± 0.441	5.73–8.29	6.71 ± 0.443	5.81–7.69	6.72 ± 0.439	5.72–8.25

accessions showed that 31 accessions not only had FL values higher than 30.00 mm but also had FS values higher than 30.00 cN·tex<sup>-1</sup> (double-thirty quality, Table 2). Eight accessions were from abroad, 15 were in the Yellow River, four from the Yangtze River and four from North/Northwest China. Meanwhile, these accessions reached double-thirty quality values in at least six environments and had FM values between 3.5 and 4.9. The accession W82-1 had both FL and FS values more than 30.00 in eight environments.

The accession MSCO-12 had highest mean value of FL (33.74 mm); the accession J02-508 had highest mean value of FS (33.94 cN·tex<sup>-1</sup>). In addition, the accessions MSCO-12, W82-1, Xinluzao17, Zhong078 and Jinmian12 were the top five for mean value of FL; the accessions J02-508, Nongda13, SuBR6202Bt, W82-1 and Zhong078 were the top five for mean value of FS; three accessions Zhong078, Shiwu107 and Luwu16 reached A level (3.7–4.2) for FM value in addition to double-thirty quality values.



**Table 2** The 31 elite upland cotton germplasms screened based on phenotypic traits

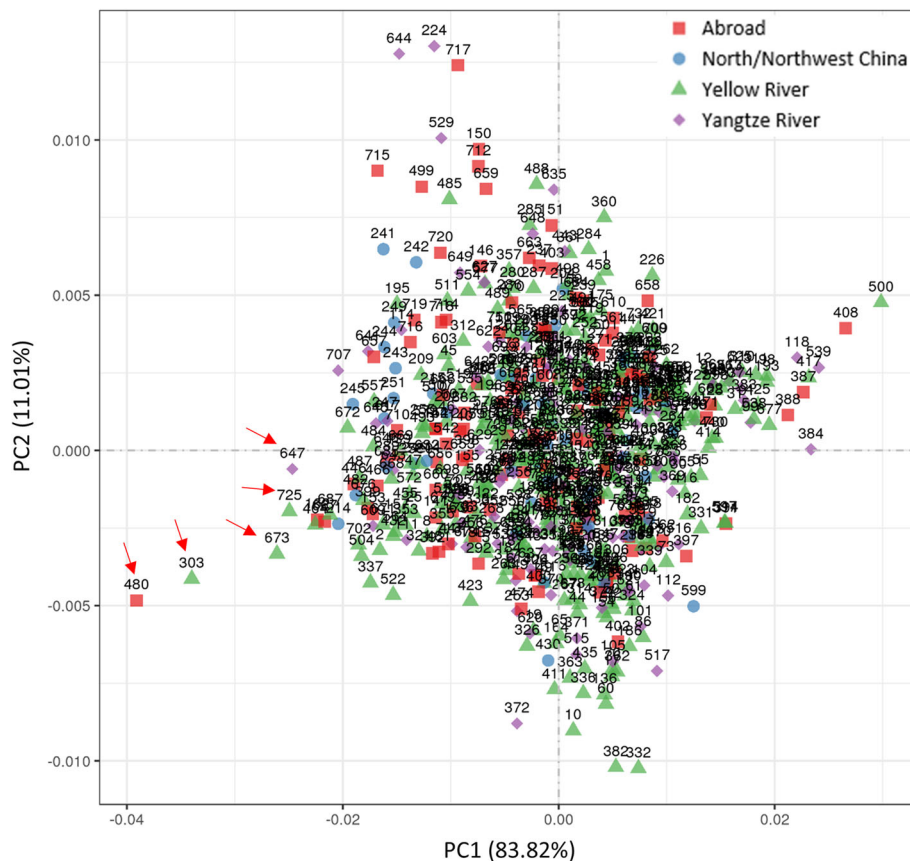
Accessions	Source	FL	FS	FM	FU	FE
MSCO-12	Abroad	33.74	32.93	3.65	83.91	6.52
W82-1	The Yellow River	33.38	33.26	4.34	85.28	6.41
Zhong078	The Yellow River	31.79	33.25	4.09	85.66	6.90
Ji4025	The Yellow River	31.59	30.93	4.75	84.49	6.44
AcalaSJ-3	Abroad	31.58	32.68	4.73	85.64	6.70
Zhongyuan9115	The Yellow River	31.56	31.17	4.42	85.18	6.56
Jiwu2046	The Yellow River	31.43	31.17	4.48	85.47	6.69
Arcot-1	Abroad	31.42	32.62	4.61	85.79	6.62
Nongda13	The Yellow River	31.40	33.65	4.70	85.69	6.75
ZhongR014121	The Yellow River	31.40	32.96	4.48	85.99	6.68
CCRI49	The Yellow River	31.25	32.59	4.67	85.52	6.70
J02-508	The Yangtze River	31.06	33.94	4.68	85.76	6.62
Zhong152	The Yellow River	31.06	31.02	4.89	85.24	6.72
Zhong2220	The Yellow River	30.95	32.39	4.86	85.83	6.51
Ekanmian3	The Yangtze River	30.92	30.98	4.58	85.14	6.90
Xinluzao38	North/Northwest China	30.87	32.48	4.50	86.07	6.49
F280	Abroad	30.85	31.36	4.75	84.60	6.40
ZhongArc-185	The Yellow River	30.84	30.91	4.54	84.86	6.56
PD3264	Abroad	30.75	32.20	4.79	85.12	6.49
M2Bt4133	Abroad	30.69	32.44	4.59	86.28	6.95
American8123	Abroad	30.56	31.74	4.73	85.07	6.47
Han8901	The Yellow River	30.56	33.03	4.67	85.81	6.51
Emian9	The Yangtze River	30.50	30.23	4.54	85.15	6.63
Hewu237	The Yellow River	30.46	30.97	4.59	84.87	6.74
Sumian9108	The Yangtze River	30.36	32.06	4.63	85.57	6.52
Shuofeng165	North/Northwest China	30.36	31.79	4.71	85.15	6.49
Kang17	North/Northwest China	30.34	32.43	4.76	86.09	6.63
Yu284	The Yellow River	30.15	32.52	4.56	85.58	6.37
ISA4ne	Abroad	30.15	30.76	4.89	85.81	6.73
Kang4	North/Northwest China	30.12	32.68	4.70	85.88	6.50
Zhong521	The Yellow River	30.12	30.59	4.62	85.31	6.51

### PCA and genetic diversity

Hierarchical cluster analysis among 719 *G. hirsutum* accessions from the different regions were obtained using principal component analysis (PCA) with the five fibre quality traits (Fig. 3). The first two PCA components explained 94.83% of the total variation. The PC1 was the most important and explained 83.82% of the total variance, and 11.01% of the variation among accessions attributed to the PC2. The Yellow River accessions distributed in the lower part of the plot; Most of North/Northwest China accessions distributed in the middle part of the plot; and most of the Yangtze River and abroad accessions distributed in the upper part of the plot. However, most of the accessions distributed in the middle of the plot and there was no clear

boundary among the four regions. Notably, elite quality germplasms, 480 for MSCO-12, 303 for W82-1, 673 for Zhong078, 725 for Nongda13, and 647 for J02-508, were clearly distinguished from the other accessions in the plot. Additionally, 31 accessions with excellent fibre quality are shown in the phylogenetic tree based on the SNPs (Fig. 4). The tree provides information on the genetic relationships between these elite accessions. The result show that there is a relatively long distance among these germplasms except for some individual accessions. Although the elite lines are scattered through the tree, elite Chinese accessions are grouped closely with abroad accessions, emphasizing that genetic relationships within the Chinese cultivars are more important than the





**Fig. 3** PCA for 719 upland cotton accessions based on the mean values of five fibre quality traits

regions where they currently grown for selecting superior parents in cotton quality breeding.

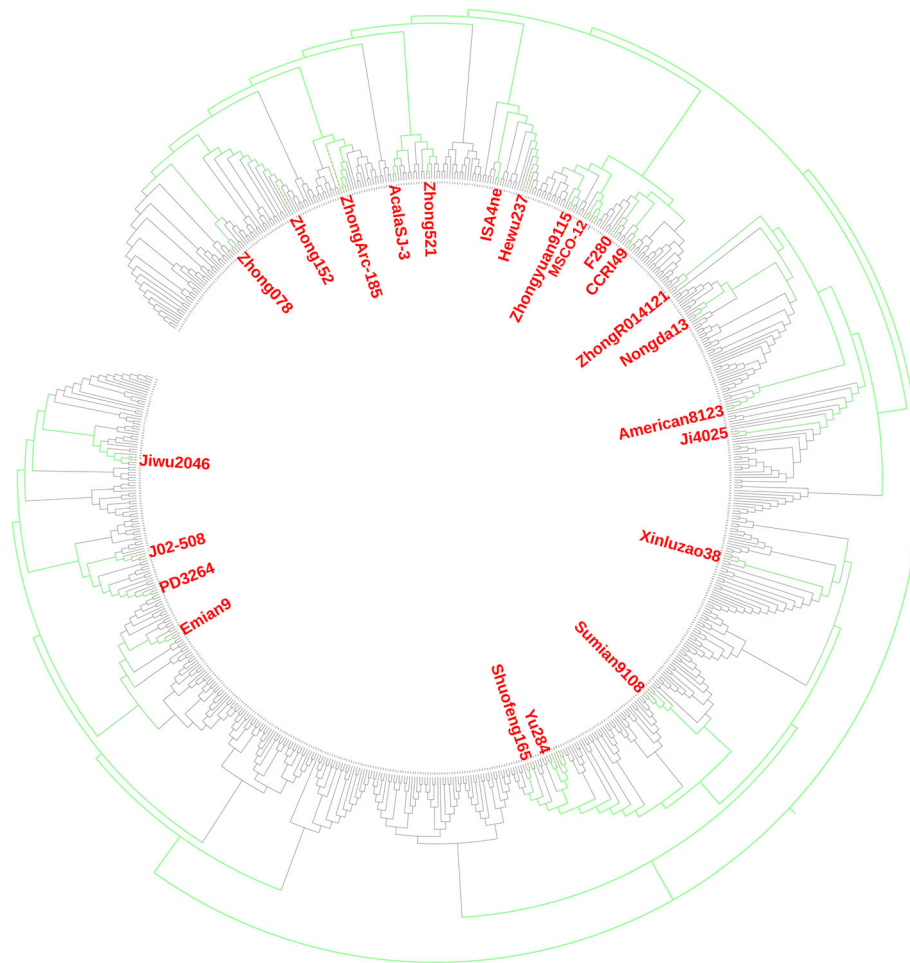
### Discussion

Morphological and molecular markers have been used extensively to describe the variability of different crops (Li et al. 2008; Huang et al. 2016; Lei et al. 2017). Most of these investigations showed a high variation of measured traits among accessions of various geographical regions. Improving fibre quality is one of the most important challenges in *G. hirsutum* breeding. However, the phenotypic variation of these traits is continuous and influenced by different aspects (Abdurakhmonov et al. 2008). In this study, we found relatively high variation of the five fibre quality traits among the accessions. These results, especially the identification of excellent germplasm, will provide important support for improving fibre quality in upland cotton.

Our results showed that the five fibre traits have high levels of variation in different regions. However, the average FL and FS of the Chinese accessions was lower than abroad ones. The reason was closely related to the history of Chinese cotton breeding. It was known that

Chinese cotton breeders mainly focused on improving yield in several decades, and this goal was indeed achieved (Dai et al. 2017). More recently, the goal of cotton breeding in China has turned to emphasis on fibre quality improvement, especially breeding for double-thirty cultivars in fibre quality to benefit cotton financial markets (Fang et al. 2017; Ma et al. 2018). All of the five traits demonstrated highly significant correlations between the two years, indicating substantial reproducibility of the results. The correlations among different fibre quality traits suggest that the FL and FS could be over other traits for breeding aims because of their positive relationship. These results are consistent with previously reported results (Nie et al. 2016; Sun et al. 2017).

By comparing the means of the fibre quality related traits over 2 years, we screened elite germplasms with longer fibre length ( $> 30$  mm), higher FS ( $> 30$  cN·tex<sup>-1</sup>) and better FM (3.7–4.2), while some genetic materials with shorter fibre length ( $< 26$  mm) and fewer FS ( $< 26$  cN·tex<sup>-1</sup>) was also observed. We identified a wide range of variations in these traits and screened elite germplasms that may be used as excellent parents for upland cotton breeding in China. Simultaneously, we preserved



**Fig. 4** Phylogenetic relationship of 719 upland cotton genotypes using SNP markers

rich genetic resources by broadening possible phenotypic variations for future genetic research on the various traits.

In this study, PCA showed that there was no obvious distinction among accessions from different growing regions (Fig. 3). Chinese accessions and abroad accessions could not be clearly distinguished, even among Chinese accessions from different geographic regions. This was most likely due to the breeding history of the cotton germplasm. Current Mexico-Guatemala region is considered the site of original domestication and the primary center of *G. hirsutum* diversity (Brubaker et al. 1999). Early Chinese accessions came from abroad and China breeders have since created distinct cotton cultivars from these imported accessions by using various breeding methods. Our results illuminated the complex intermingling of imported genetic stocks in modern Chinese accessions, indicating that distinct, region-specific cultivars have not emerged. For optimizing fibre quality, the results indicate that elite and high-performing

lines bred in China or abroad will also have high performing in diverse growing areas within China. However, breeding focused on these lines will further restrict the already limited diversity in cultivated upland cotton varieties.

## Conclusions

This study provided a detailed description of a population that represents a wide range of upland cotton diversity germplasm. The results of the fibre quality traits evaluations of the *G. hirsutum* accessions showed a wide range of variation over 2 years. In general, abroad accessions tended to have higher FL and FS than did Chinese accessions. Among different geographic regions accessions in China, North/Northwest accessions tended to have the highest FL, FS and best FM. Through evaluating the five fibre quality traits over 2 years, we selected 31 elite germplasms reaching double-thirty quality values. PCA based on

phenotypes revealed no clear boundary among the germplasm materials along the first two principal coordinates of different geographic origins and different regions. This study will enable breeders to make useful information about possible parents for cotton breeding programs.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s42397-019-0041-2>.

**Additional file 1: Table S1.** List of 719 upland accessions used in this study. **Table S2.** Correlation coefficients among the 40 phenotypic traits using 719 upland cotton.

## Abbreviations

FE: Fibre elongation; FL: Fibre length; FM: Fibre micronaire; FS: Fibre strength; FU: Fibre uniformity; PCA: Principal component analysis; SD: Standard deviation; SNP: Single nucleotide polymorphism

## Acknowledgments

Not applicable.

## Authors' contributions

MZY conceived and designed the research; SZW, WXF, LZW, GQS, ZY, LZK, KHF, YJ, WJH, WLQ and ZGY performed experiments; SZW, LZW and WXF analyzed data; SZW and WXF wrote the manuscript; MZY revised the manuscript. All authors read and approved the manuscript.

## Funding

This work was supported by grants from the National Key Research and Development Program of China (2016YFD0101006 and 2016YFD0101405) and the Top Talent Fund of Hebei Province.

## Availability of data and materials

No other data related to this study is available at this time.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

Received: 19 July 2019 Accepted: 22 November 2019

Published online: 23 December 2019

## References

- Abdurakhmonov IY, Kohel RJ, Yu JZ, et al. Molecular diversity and association mapping of fibre quality traits in exotic *G. hirsutum* L. germplasm. *Genomics*. 2008;92:478–87. <https://doi.org/10.1016/j.ygeno.2008.07.013>.
- Ai XT, Liang YJ, Wang JD, et al. Genetic diversity and structure of elite cotton germplasm (*Gossypium hirsutum* L.) using genome-wide SNP data. *Genetica*. 2017;145:409–16. <https://doi.org/10.1007/s10709-017-9976-8>.
- Amezrou R, Gyawali S, Belqadi L, et al. Molecular and phenotypic diversity of ICARDA spring barley (*Hordeum vulgare* L.) collection. *Genet Resour Crop Evol*. 2017;65:255–69. <https://doi.org/10.1007/s10722-017-0527-z>.
- Badigannavar A, Myers GO. Genetic diversity, population structure and marker trait associations for seed quality traits in cotton (*Gossypium hirsutum*). *J Genet*. 2015;94:87–94. <https://doi.org/10.1007/s12041-015-0489-x>.
- Brubaker CL, Paterson AH, Wendel JF. Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. *Genome*. 1999;42:184–203. <https://doi.org/10.1139/gen-42-2-184>.
- Chen ZJ, Scheffler BE, Dennis E, et al. Toward sequencing cotton (*Gossypium*) genomes. *Plant Physiol*. 2007;145:1303–10. <https://doi.org/10.1104/pp.107.107672>.
- Dai J, Kong X, Zhang D, et al. Technologies and theoretical basis of light and simplified cotton cultivation in China. *Field Crop Res*. 2017;214:142–8. <https://doi.org/10.1016/j.fcr.2017.09.005>.
- Fang L, Wang Q, Hu Y, et al. Genomic analyses in cotton identify signatures of selection and loci associated with fiber quality and yield traits. *Nat Genet*. 2017;49:1089–98. <https://doi.org/10.1038/ng.3887>.
- Hinze LL, Fang DD, Gore MA, et al. Molecular characterization of the *Gossypium* diversity reference set of the US national cotton germplasm collection. *Theor Appl Genet*. 2015;128:313–27. <https://doi.org/10.1007/s00122-014-2431-7>.
- Hinze LL, Hulse-Kemp AM, Wilson IW, et al. Diversity analysis of cotton (*Gossypium hirsutum* L.) germplasm using the CottonSNP63K array. *BMC Plant Biol*. 2017;17:37. <https://doi.org/10.1186/s12870-017-0981-y>.
- Huang C, Nie XH, Shen C, et al. Population structure and genetic basis of the agronomic traits of upland cotton in China revealed by a genome-wide association study using high-density SNPs. *Plant Biotechnol J*. 2017;15:1374–86. <https://doi.org/10.1111/pbi.12722>.
- Huang J, Guo N, Li YH, et al. Phenotypic evaluation and genetic dissection of resistance to *Phytophthora sojae* in the Chinese soybean mini core collection. *BMC Genet*. 2016;17:85. <https://doi.org/10.1186/s12863-016-0383-4>.
- Lei QY, Zhou JJ, Zhang WH, et al. Morphological diversity of panicle traits in Kam fragrant glutinous rice (*Oryza sativa*). *Genet Resour Crop Evol*. 2017;65:775–86. <https://doi.org/10.1007/s10722-017-0570-9>.
- Li ZK, Wang XF, Zhang Y, et al. Assessment of genetic diversity in glandless cotton germplasm resources by using agronomic traits and molecular markers. *Front Agric China*. 2008;2:245–52. <https://doi.org/10.1007/s11703-008-0063-x>.
- Liu K, Muse SV. PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics*. 2005;21:128–9. <https://doi.org/10.1093/bioinformatics/bti282>.
- Ma ZY, He SP, Wang XF, et al. Resequencing a core collection of upland cotton identifies genomic variation and loci influencing fibre quality and yield. *Nat Genet*. 2018;50:803–13. <https://doi.org/10.1038/s41588-018-0119-7>.
- Nie XH, Huang C, You CY, et al. Genome-wide SSR-based association mapping for fibre quality in nation-wide upland cotton inbred cultivars in China. *BMC Genomics*. 2016;17:352. <https://doi.org/10.1186/s12864-016-2662-x>.
- Rungis D, Llewellyn D, Dennis ES, et al. Simple sequence repeat (SSR) markers reveal low levels of polymorphism between cotton (*Gossypium hirsutum* L.) cultivars. *Aust J Agric Res*. 2005;56:301–7. <https://doi.org/10.1071/AR04190>.
- Salazar E, Correa J, Araya MJ, et al. Phenotypic diversity and relationships among Chilean Choclero maize (*Zea mays* L. mays) landraces. *Plant Genet Resour*. 2016;15:461–73. <https://doi.org/10.1017/S1479262116000137>.
- Sun ZW, Wang XF, Liu ZW, et al. Genome-wide association study discovered genetic variation and candidate genes of fibre quality traits in *Gossypium hirsutum* L. *Plant Biotechnol J*. 2017;15:982–96. <https://doi.org/10.1111/pbi.12693>.
- Tyagi P, Gore MA, Bowman DT, et al. Genetic diversity and population structure in the US upland cotton (*Gossypium hirsutum* L.). *Theor Appl Genet*. 2014;127:283–95. <https://doi.org/10.1007/s00122-013-2217-3>.
- Ulloa M, Abdurakhmonov IY, Perez-M C, et al. Genetic diversity and population structure of cotton (*Gossypium* spp.) of the New World assessed by SSR markers. *Botany*. 2013;91:251–9. <https://doi.org/10.1139/cjb-2012-0192>.
- Wendel JF, Cronn RC. Polyploidy and the evolutionary history of cotton. *Adv Agron*. 2003;78:139–86. [https://doi.org/10.1016/S0065-2113\(02\)78004-8](https://doi.org/10.1016/S0065-2113(02)78004-8).
- Zhang YX, Zhang XR, Che Z, et al. Genetic diversity assessment of sesame core collection in China by phenotype and molecular markers and extraction of a mini-core collection. *BMC Genet*. 2012;13:102. <https://doi.org/10.1186/1471-2156-13-102>.
- Zhao YL, Wang HM, Chen W, et al. Genetic structure, linkage disequilibrium and association mapping of Verticillium wilt resistance in elite cotton (*Gossypium hirsutum* L.) germplasm population. *PLoS One*. 2014;9:e86308. <https://doi.org/10.1371/journal.pone.0086308>.