



Molecular diversity and association mapping of fiber quality traits in exotic *G. hirsutum* L. germplasm

I.Y. Abdurakhmonov^{a,*}, R.J. Kohel^{b,1}, J.Z. Yu^{b,1}, A.E. Pepper^{c,2}, A.A. Abdullaev^a, F.N. Kushanov^a, I.B. Salakhutdinov^a, Z.T. Buriev^a, S. Saha^{d,3}, B.E. Scheffler^e, J.N. Jenkins^d, A. Abdukarimov^a

^a Center of Genomic Technologies, Institute of Genetics and Plant Experimental Biology, Academy of Sciences of Uzbekistan, Yuqori Yuz, Qibray region Tashkent district, 702151, Uzbekistan

^b USDA-ARS, Crop Germplasm Research Unit, 2881 F&B Road College Station, TX 77845, USA

^c Department of Biology, Texas A&M University, College Station, Texas 77843, USA

^d USDA-ARS, Crop Science Research Laboratory, Genetics and Precision Agriculture, P. O. Box 5367, 8120 Highway 12E, Mississippi State, MS 39762, USA

^e USDA-ARS, 141 Experiment Station Road, Stoneville, MS 38776, USA

ARTICLE INFO

Article history:

Received 29 January 2008

Accepted 29 July 2008

Available online 5 October 2008

Keywords:

Cotton germplasm

Genetic diversity

Fiber quality

Linkage disequilibrium (LD)

Simple sequence repeats (SSR) markers

LD-based association mapping

ABSTRACT

The narrow genetic base of cultivated cotton germplasm is hindering the cotton productivity worldwide. Although potential genetic diversity exists in *Gossypium* genus, it is largely 'underutilized' due to photoperiodism and the lack of innovative tools to overcome such challenges. The application of linkage disequilibrium (LD)-based association mapping is an alternative powerful molecular tool to dissect and exploit the natural genetic diversity conserved within cotton germplasm collections, greatly accelerating still 'lagging' cotton marker-assisted selection (MAS) programs. However, the extent of genome-wide linkage disequilibrium (LD) has not been determined in cotton. We report the extent of genome-wide LD and association mapping of fiber quality traits by using a 95 core set of microsatellite markers in a total of 285 exotic *Gossypium hirsutum* accessions, comprising of 208 landrace stocks and 77 photoperiodic variety accessions. We demonstrated the existence of useful genetic diversity within exotic cotton germplasm. In this germplasm set, 11–12% of SSR loci pairs revealed a significant LD. At the significance threshold ($r^2 \geq 0.1$), a genome-wide average of LD declines within the genetic distance at < 10 cM in the landrace stocks germplasm and > 30 cM in variety germplasm. Genome wide LD at $r^2 \geq 0.2$ was reduced on average to ~ 1 – 2 cM in the landrace stock germplasm and 6–8 cM in variety germplasm, providing evidence of the potential for association mapping of agronomically important traits in cotton. We observed significant population structure and relatedness in assayed germplasm. Consequently, the application of the mixed liner model (MLM), considering both kinship (K) and population structure (Q) detected between 6% and 13% of SSR markers associated with the main fiber quality traits in cotton. Our results highlight for the first time the feasibility and potential of association mapping, with consideration of the population structure and stratification existing in cotton germplasm resources. The number of SSR markers associated with fiber quality traits in diverse cotton germplasm, which broadly covered many historical meiotic events, should be useful to effectively exploit potentially new genetic variation by using MAS programs.

© 2008 Elsevier Inc. All rights reserved.

Introduction

The genus *Gossypium* includes approximately 45 diploid and 5 allotetraploid species distributed mostly in tropical and subtropical regions of the world [1,2]. It provides most of the world's natural textile fiber, sources of oil, and cottonseed meal as feed products [3,4]. Diploid cottons are classified into eight (A–G to K) cytogenetically defined genome groups [5]. Hybridization between A-genome

and D-genome diploids and subsequent polyploidization about 1.5 million years ago created the five AD allotetraploid lineages that are indigenous to the Americas and Hawaii [6–8]. These New World tetraploid cottons include the commercially important species, *Gossypium hirsutum* and *G. barbadense*, where *G. hirsutum* is the most widely cultivated (90%) industrial cotton among all *Gossypium* species [9,10]. The origin of *G. hirsutum* is Guatemala, but its large indigenous range encompassed most of Mesoamerica and Caribbean. According to archaeobotanical findings, *G. hirsutum* probably was domesticated originally within the Southern end of Mesoamerican gene pool [11,12]. Mexico–Guatemala is considered the site of original domestication and primary center of diversity. *G. hirsutum* exhibits a diversity of types including a single wild race, 'yucatanense', six primitive domesticated forms, 'richmondi', 'punctatum', 'latifolium', 'palmeri',

* Corresponding author. Fax: +7 998 712 642 390.

E-mail address: genomics@uzsci.net (I.Y. Abdurakhmonov).

¹ Phone: +1 979 260 9311; fax: +1 979 260 9333.

² Phone: +1 979 845 2518; fax: +1 979 862 4790.

³ Phone: +1 662 320 7386; fax: +1 662 320 7528.

Table 1
Summary of fiber quality traits from Mexican environment

Traits	Number	Mean	Minimum	1 st Q	median	3 rd Q	Maximum
Micronaire (MIC)	203	4.73	3	4.25	4.7	5.15	6.6
Fiber length (UHM)	203	1.09	0.84	1.04	1.1	1.15	1.28
Uniformity (UI)	203	83.84	79	82.45	83.8	85.35	88
Strength (STR)	203	27.54	21.3	25.55	27.4	29.3	35.5
Elongation (ELO)	203	6.52	3.8	5.8	6.5	7.2	10.1
Reflectance (RD)	203	72.11	0	74.65	76.2	77.9	81.5

Note: 1st Q–25%-ile; 3rd Q–75%-ile.

'morilli', 'marie-galante' stages [13,14], and a number of domesticated variety accessions, 'euhirsutum' [15], grown in 80 different countries worldwide [3]. The accessions of wild and primitive races of *G. hirsutum* are the 'golden' reservoirs of genetic diversity, but they are largely 'underutilized' in breeding programs due to their short-day photoperiodic flowering [16,17]. The 'euhirsutum' accessions are the main germplasm resources of worldwide cotton breeding programs to improve cultivars, and world cotton production greatly relies on these variety germplasm. Consequently, these variety germplasm of *G. hirsutum* make the major portion (at least 50–80%) of cotton germplasm collections worldwide [10]. Currently, cotton researchers and producers are concerned with the narrow genetic base of cultivated cottons that has experienced recent cotton yield and quality declines [18,19]. The yield and quality stagnation could be directly attributed to many factors, perhaps, including rapid genetic erosion of genetically uniform elite cultivars and paucity of information on the complex nature of the cotton genome especially in relation to economically important traits. These indicate the critical need to 1) explore novel germplasm resources for potential natural genetic diversity and 2) develop innovative genomics tools to efficiently mobilize these useful genetic variations to breeding germplasm. This should help to overcome existing and potential problems of worldwide cotton production associated with narrow genetic base of the cultivar germplasm.

During past years, the international cotton research community has developed extensive genomic resources [20,21]. A large collection of robust, portable, and PCR-based Simple Sequence Repeat (SSR) marker resources were developed in multiple laboratories and made available to the cotton research community through the cotton marker database (CMD) [22]. Molecular marker technology was successfully used to assess genetic diversity of germplasm resources, create genetic linkage maps, and map important agronomic QTLs in bi-parental mapping populations [20,21], which are imperative for acceleration of marker-assisted-selection (MAS) programs in cotton. Alternatively, turning the gene-tagging efforts from bi-parental crosses to germplasm collections, and from traditional linkage mapping to linkage disequilibrium (LD)-based association study promises the most effective utilization of *ex situ* conserved natural genetic diversity of worldwide cotton germplasm resources. LD refers to a historically reduced (non-equilibrium) level of the recombination of specific alleles at different loci controlling particular genetic variation in a population. This LD can be measured statistically, and has been widely applied to map and eventually clone a number of genes underlying complex genetic traits in humans [23,24].

The advantages of population-based association study over conventional QTL mapping in bi-parental crosses primarily are due to 1) availability of broader genetic variations with wider background for marker-trait correlations (i.e. multiple alleles evaluated simultaneously), 2) likelihood of a higher resolution mapping because of the utilization of majority recombination events in the germplasm's developmental history, 3) possibility of exploiting historically measured trait data for associations, and 4) no need for the development of expensive and tedious bi-parental populations that makes the approach timesaving and cost-effective [25–27]. Conversely, traditional QTL mapping is 1) very costly [27], 2) has poor resolution with

the evaluation of only a few alleles [28], and 3) it requires a longer research time period. Additionally, association mapping in plants, compared to human populations, has more power with the opportunity to create mapping populations with required amount of LD and diversity [29].

The measurement of the LD patterns for genomic regions and specificity of LD extent among different populations or groups of the 'target' organisms is the important starting point to design and conduct association mapping [30,31]. LD has been quantified in several plant species [32,33] including the model organism *Arabidopsis* and now extended to crops such as maize, barley, durum wheat, spring wheat, sorghum, soybean, sugarcane, sugar beet and grapevine, as well as in trees such as European aspen, and loblolly pine [28,32,33]. These studies revealed that the genome-wide extent of LD varied across genomes and between species with the examples of longer stretches of LD in some local populations. Moreover, LD-based association mapping was successfully used in plant germplasm resources that highlighted serious influence of population structure and relatedness of individuals in conducting association mapping [32,33].

The application of LD-based association mapping for cotton will not only accelerate still lagging MAS programs in cotton, but also adds to our knowledge and understanding of the complex cotton genome and its evolution. However, the extent of genome-wide LD has not been determined in cotton. In this study, we utilized largely 'unexploited' cotton germplasm resource from the Republic of Uzbekistan to search for new genetic variation in fiber quality traits in cotton and mapping the main fiber quality traits using LD-based association mapping strategy. Here we report the extent and distribution of molecular diversity, population structure, kinship, and an average extent of genome-wide LD for exotic germplasm measured using SSR markers. LD-based association mapping for fiber quality traits was applied using mixed linear model approach that found several SSR markers associated with the main fiber quality traits of exotic cotton germplasm. The results provide preliminary insight into genome-wide averages for extent of LD in cotton and are very useful as a framework for future 'association studies' in cotton that will accelerate development of superior cotton cultivars through MAS programs.

Results

Fiber quality properties of selected accessions in Mexican environment

The cotton accessions including 208 landrace stocks and 77 photoperiodic variety accessions from Mexico and Africa (Table S1) revealed a wide range of phenotypic variation among fiber quality traits including fiber strength (STR), fiber length (upper half mean-UHM), uniformity (UI), elongation (ELO), micronaire (MIC), and reflectance (Rd) (Table 1). We observed significant trait correlations among different fiber traits in the Mexican environment (Tables 2). The negative correlations were observed between MIC and UHM, MIC and STR, and STR and ELO. Positive correlations were observed

Table 2
Correlation of fiber quality traits from Mexican environment

TRAITS	MIC	UHM	UI	STR	ELO	RD
MIC	1					
UHM	-0.49****	1				
UI	-0.09	0.44****	1			
STR	-0.32****	0.69****	0.29****	1		
ELO	-0.02	-0.11	0.09	-0.29****	1	
RD	0.09	0.28****	0.22**	0.32****	-0.03	1

MIC–Micronaire; UHM–fiber length; UI–uniformity; STR–fiber strength; ELO–elongation; RD–reflectance. **, ****, $p \leq 0.01$, 0.001 , respectively.

Download English Version:

<https://daneshyari.com/en/article/2821364>

Download Persian Version:

<https://daneshyari.com/article/2821364>

[Daneshyari.com](https://daneshyari.com)