Feasibility assessment of phenotyping cotton fiber maturity using infrared spectroscopy and algorithms for genotyping analyses

KIM Hee Jin1*, LIU Yongliang2, FANG David D.1 and DELHOM Christopher D.2

Abstract

Background: Cotton fiber maturity is an important property that partially determines the processing and performance of cotton. Due to difficulties of obtaining fiber maturity values accurately from every plant of a genetic population, cotton geneticists often use micronaire (MIC) and/or lint percentage for classifying immature phenotypes from mature fiber phenotypes although they are complex fiber traits. The recent development of an algorithm for determining cotton fiber maturity ($M_{IR}$) from Fourier transform infrared (FT-IR) spectra explores a novel way to measure fiber maturity efficiently and accurately. However, the algorithm has not been tested with a genetic population consisting of a large number of progeny plants.

Results: The merits and limits of the MIC- or lint percentage-based phenotyping method were demonstrated by comparing the observed phenotypes with the predicted phenotypes based on their DNA marker genotypes in a genetic population consisting of 708 F2 plants with various fiber maturity. The observed MIC-based fiber phenotypes matched to the predicted phenotypes better than the observed lint percentage-based fiber phenotypes. The lint percentage was obtained from each of F2 plants, whereas the MIC values were unable to be obtained from the entire population since certain F2 plants produced insufficient fiber mass for their measurements. To test the feasibility of cotton fiber infrared maturity ($M_{IR}$) as a viable phenotyping tool for genetic analyses, we measured FT-IR spectra from the second population composed of 80 F2 plants with various fiber maturities, determined $M_{IR}$ values using the algorithms, and compared them with their genotypes in addition to other fiber phenotypes. The results showed that $M_{IR}$ values were successfully obtained from each of the F2 plants, and the observed $M_{IR}$-based phenotypes fit well to the predicted phenotypes based on their DNA marker genotypes as well as the observed phenotypes based on a combination of MIC and lint percentage.

Conclusions: The $M_{IR}$ value obtained from FT-IR spectra of cotton fibers is able to accurately assess fiber maturity of all plants of a population in a quantitative way. The technique provides an option for cotton geneticists to determine fiber maturity rapidly and efficiently.

Keywords: Attenuated total reflection Fourier transform infrared spectroscopy, Cotton fiber maturity, Crystallinity, Immature fiber (im) mutant
Background
Cotton fiber maturity is an important physical property that affects both yield and fiber quality (Peirce and Lord 1939). It is directly correlated to dye uptake of yarn and fabric products as well as fiber breakage and entanglement during mechanical processes (Kelly et al. 2015). Cotton fiber maturity may be referred to as circularity (θ) that is defined as the ratio of the cross-sectional cell wall area to the area of a circle having the same perimeter. In lieu of θ, maturity ratio (MR = θ/ 0.577) is frequently used by cotton breeders and the textile industry (Gordon and Rodgers 2017). The maturity values can be directly determined by image analysis microscopy (IAM) by measuring average cell wall area and perimeters from 300–500 cross-sectioned fibers for each cotton sample (Hequet et al. 2006; Thibodeaux and Evans 1986). The IAM method has rarely been used for classifying cotton materials in genetic studies due to its lengthy and laborious process. MR values can also be indirectly measured by Advanced Fiber Information System (Kelly et al. 2012) or Cottonscope® (Rodgers et al. 2011). For a quick and automated assessment of fiber maturity, the cotton community has depended on the High Volume Instrument (HVI) that is a standardized instrument for measuring cotton fiber properties including Micronaire (MIC) as recognized by the International Cotton Advisory Committee and other organizations (ASTM D5867-12e1 2012). MIC represents a combination of fiber maturity and fineness by measuring air-flow resistance through a plug of cotton fibers of a given weight which has been compressed to a known volume (Frydrych and Thibodeaux 2010).

A cotton fiber mutant that produces immature fibers was originally identified from an upland cotton variety Acala 4–42 (Kohel et al. 1974) and later named as immature fiber (im) mutant (Kohel and McMichael 1990). By backcrossing the original im mutant several times with the wild type (WT) Texas Marker-1 (TM-1), a pair of near isogenic lines differing in fiber maturity was developed (Kohel and McMichael 1990). The MIC values of im fibers are significantly lower than that of TM-1 fibers. The MIC value difference was originally suggested as a way to classify the im plant from the WT plant. However, cotton geneticists faced difficulties obtaining MIC values from every plant in a segregating F2 population from a cross between the im and WT cotton plants due to insufficient fiber mass required by HVI or Fibroniare. In an attempt to find a way to identify the im phenotypes from the F2 plants producing insufficient fiber mass for the MIC measurements, cotton geneticists have been primarily using lint percentage since Kohel and his colleagues reported the 40% dry weight difference between the im and WT fibers (Kohel et al. 1974). Lint percentage has been closely associated with yield improvements of commercial cultivars (Bridge et al. 1971; Meredith Jr and Bridge 1973; Meredith Jr 1984) and is significantly and positively correlated with MIC readings of cotton fibers in general (Meredith Jr 1984; Wan et al. 2007). Unlike the MIC value, the lint percentage was successfully obtained from every F2 plant of the segregating populations (Kim et al. 2013a; Kohel and McMichael 1990; Thyssen et al. 2016; Wang et al. 2013). For identifying the im locus by mapping-by-sequencing that required quantitative fiber trait data from 2,837 F2 plants, a combination of the lint percentage observed from all F2 plants with the MIC data observed from a portion of the population was used to distinguish im phenotype from WT phenotype (Thyssen et al. 2016).

Plant biologists have been using Fourier transform infrared (FT-IR) spectroscopy to classify secondary cell wall (SCW) cellulose from primary cell wall (PCW) cellulose of model plants in a rapid and non-invasive way (McCann et al. 1992). FT-IR spectroscopy has been recently used to monitor cotton fiber wall composition (Abidi et al. 2008), SCW cellulose development (Abidi et al. 2010a; Islam et al. 2016; Kim et al. 2018), sugar composition (Abidi et al. 2010b) and crystallinity (Abidi et al. 2014; Abidi and Manike 2018; Liu et al. 2012) from a few cotton species or several upland cotton cultivars. Based on attenuated total reflection (ATR) FT-IR spectral differences between immature and mature seed cotton fibers, simple algorithms that utilized the intensity ratios of three IR vibrations at 1 500, 1 032, and 956 cm⁻¹ (Liu et al. 2011) and another three IR vibrations at 1,500, 1,032, and 956 cm⁻¹ (Liu et al. 2012) were proposed to determine cotton fiber infrared maturity (MIR) and crystallinity (CIr), respectively. The observed MIR values of cotton fibers harvested from im mutant and WT parents were able to distinguish the two phenotypes successfully (Kim et al. 2017; Liu and Kim 2019), monitor the development of cotton fiber grown in planta and in tissue culture (Liu and Kim 2015), and validated with the fiber maturity that was measured from developing and developed fibers by a cross-sectional image analysis (Liu et al. 2019). However, the technique has not been tested with segregating populations for genetic analyses despite the advantages including that (1) a simple and direct ATR FT-IR measurement of cotton fibers avoids the need to perform any preparation or pretreatment of cotton samples, (2) the technique requires a small amount of fibers (as little as 0.5 mg) as compared to the minimum fiber mass (> 10.0 g) for HVI measurement, and (3) a short time (less than 2 min) for sample loading, spectral acquisition, and subsequent result reporting.

In the present research, we used two different sets of cotton materials. The first set consisting of 708 F2 plants was used to find merits and limits of conventional fiber maturity phenotyping methods including MIC and lint percentage for genetic analyses. The second set consisting...
of 80 F₂ plants was used to compare the genotypes with the phenotypes based on MIC, lint percentage, and $M_{IR}$ values. The results showed that the $M_{IR}$ value obtained from FT-IR spectra was significantly correlated with the MIC and successfully classified the im phenotype from WT phenotype. Unlike the MIC values that were unable to be obtained from all 80 F₂ plants, the $M_{IR}$ values were observed quantitatively from each of F₂ plants.

Results and discussion

The first set of cotton materials with various MIC values

For this study, we used 708 F₂ plants derived from a cross between the WT cotton line MD52ne and the im mutant. The MD52ne produces fluffy cotton bolls, whereas the im mutant generates non-fluffy cotton bolls (Fig. 1a). Comparisons of cross-sectioned fibers between the MD52ne and im mutant showed visible differences of the cell wall area (Fig. 1a, inset).

In our previous research (Thyssen et al. 2016), the phenotype of each F₂ plant was obtained by calculating the lint percentage and by measuring HVI MIC values where possible, since the lint percentage were acquired from all F₂ plants and the HVI was unable to measure the MIC values from some F₂ plants that produced less than 10.0 g of fiber. Therefore, the genotypes of all F₂ plants were compared with the phenotypes determined by lint percentage, but not by MIC values. In this study, we measured additional MIC value from the F₂ progeny plants by using the Fibronaire Instrument which measures MIC values on 3.24 g of fiber mass. These 708 F₂ progeny were designated as the first set of cotton materials and used to compare the strength and weakness of the conventional MIC- and lint percentage-based phenotypes of the F₂ population.

Genotyping analysis of the first set of cotton materials for classifying the immature fiber phenotype

Based on the single recessive im gene (Kim et al. 2013a; Kohel and McMichael 1990; Thyssen et al. 2016), the 1:2:1 ratio of genotypes including WT homozygosity (WT-homo), WT heterozygosity (WT-hetro), and im homozygosity (im-homo) is expected among the 708 F₂ progeny plants. To determine the three types of genotypes, we extracted genomic DNAs from young leaves of each F₂ plant, and analyzed them with the previously developed markers located at the coding sequence of the im gene (CFBid0001) or in its flanking regions (DPL1071 and SHIN-1511) (Kim et al. 2013a; Thyssen et al. 2016). Of the 708 F₂ plants, the WT-homo, WT-hetro, and im-homo genotypes were observed in 218 (30.8%), 335 (47.3%), and 155 (21.9%) progeny plants, respectively (Fig. 1b). The observed WT-homo genotype (30.8%) was numerically greater than the expected one (25%), and the observed WT-hetro genotype (47.3%) was less than the expected one (50%). The observed im-homo genotype (21.9%) was less than the expected one (25.0%). The calculated chi-square ($\chi^2$, 0.524) and probability ($P$, 0.469) confirmed that the observed genotype ratio fit the expected segregation ratio (1: 2: 1). Based on the observed genotype ratio (30.8, 47.3%: 21.9%), the expected phenotype ratio between WT and im mutant should be 78.1% (blue): 21.9% (red) (Fig. 1b).

MIC: accurate, but limited to perform quantitative genetic analysis for entire F₂ plants

Despite the striking fiber phenotypic differences between the im mutant and WT plants (Fig. 1a), it has been a challenge to distinguish the field grown im mutant from WT plants (Kim et al. 2013a; Kim et al. 2013b; Kohel et al. 2013a). The MIC values obtained from FT-IR spectra showed a significant correlation with the MIC, and successfully classified the im phenotype from WT phenotype. Unlike the MIC values that were unable to be obtained from all 80 F₂ plants, the $M_{IR}$ values were observed quantitatively from each of F₂ plants. However, the MIC values were unable to be obtained from some F₂ plants that produced less than 10.0 g of fiber. Therefore, the genotypes of all F₂ plants were compared with the phenotypes determined by lint percentage, but not by MIC values. In this study, we measured additional MIC value from the F₂ progeny plants by using the Fibronaire Instrument which measures MIC values on 3.24 g of fiber mass. These 708 F₂ progeny were designated as the first set of cotton materials and used to compare the strength and weakness of the conventional MIC- and lint percentage-based phenotypes of the F₂ population.
The non-fluffy cotton boll phenotype is not unique to the \textit{im} mutant. In field conditions with biotic and abiotic stress, a WT cotton cultivar may also produce an \textit{im} mutant-like phenotype, referring to tight lock bolls. Previous reports showed that the observed MIC values by the HVI measurement were able to classify the \textit{im} phenotype from the WT phenotype (Kim et al. 2014; Kim et al. 2017; Kothari et al. 2007).

The MIC value from individual \(F_2\) progeny plants can be measured with HVI when each plant produces more than 10.0 g of fibers, or Fibronaire if more than 3.24 g of fibers but less than 10.0 g. Among the 708 \(F_2\) cotton plants, MIC values were measured by either HVI or Fibronaire from 547 WT phenotype plants (77.8%) including WT-homo (217 plants) and WT-hetero (330 plants) genotypes as well as 52 \textit{im} phenotype plants with the \textit{im}-homo genotype (Fig. 2a). Using both instruments, we were able to measure the MIC values from 599 \(F_2\) progeny plants (84.6%), but were unable to obtain MIC values from 109 \(F_2\) plants (15.4%) due to production of less than 3.24 g of cotton fibers (Fig. 2a). We assigned the 109 plants as unmeasurable (U.M.) samples for being distinguished from the measurable (M) 599 samples.

A frequency distribution curve of the MIC values from the measurable 599 \(F_2\) progeny plants showed two distinct peaks (Fig. 2b). The greater peak was observed from a normal range of MIC values (3.65 – 5.41) for WT phenotype. In contrast, a low range of MIC values (2.03 – 3.60) was noted for the mostly \(F_2\) plants with the \textit{im} genotype. The MIC peak of the \textit{im} curve occurred around 2.30, and MIC values below 2.0 are not measurable due to the detection limits of the HVI.

We compared the MIC values with the three different genotype data (Fig. 2c). Among the unmeasurable 109 \(F_2\)....

---

**Fig. 2** Phenotyping analysis based on MIC values. **a** Diagram representing ratios between WT and \textit{im} phenotypes. MIC values were measured from the 599 plants of the 708 \(F_2\) progeny. **b** Frequency distribution curve based on MIC values. **c** Scatter dot plot of comparing genotypes with the MIC-based phenotypes. Genotypes representing homozygous wild type (WT-homo), heterozygous wild type (WT-hetero), and homozygous \textit{im} type (\textit{im}-homo) were determined from DNA markers. T, threshold value; N, sample number; M, measureable; U.M, unmeasurable.
progeny plants, 103 plants belonged to the im-homo genotype. The other six WT plants (one WT-homo plant and five WT-hetro plants) also produced insufficient fiber mass (<3.24 g) for MIC measurement (U.M. in Fig. 2c). Average MIC values of 4.85 and 4.77 were observed for the measurable WT plants (217 WT-homo and 330 WT-hetro genotypes) with a range from 3.65 to 5.51 (Fig. 2c). Average MIC value of the measurable 52 of the 103 im plants was 2.64. Based on the minimum MIC value of the WT phenotype, the MIC values ranging from 3.50 to 3.60 appeared to be a threshold for classifying the im phenotype from WT phenotypes despite the four outliers of the im mutant (Fig. 2c). We suspect the outliers might be recombinants as discussed in Thyssen et al. (2016). Using the results of the Fig. 2b and c, we arbitrarily classified the F₂ progeny plants (MIC< 3.60) and im phenotypes (MIC<3.60). Due to the lack of MIC values from the entire F₂ population, we were unable to compare the observed phenotype with the predicted phenotypes with the calculated chi-square and probability.

Based on the observation that 109 F₂ plants of the total 708 F₂ progeny plants produced sufficient and unmeasurable amount of fiber samples (<3.24 g) for MIC measurement, we were aware that MIC value alone is not sufficient to meet genetic analysis of the F₂ population. To perform genetic analysis with quantitative phenotype from mature phenotype, the lint percentage determined by the lint percentage was lower than the observed im genotype (21.9%), but much lower than the expected phenotype (25.0%). Since we have already determined the im genotype, we compared the known im genotype with the observed im phenotype based on the lint percentage. The calculated chi-square (χ², 0.669) and probability (P, 0.413) suggested that the observed im phenotype fit the expected im phenotype determined by the im genotype. However, the observed segregation ratio of the im phenotype by the lint percentage did not meet the expected 3:1 segregation ratio according to the calculated chi-square (χ², 7.239) and probability (P, 0.071). As a result, we concluded that lint percentage in conjunction with MIC data can be used for classifying im phenotype from the WT phenotype. However, the lint percentage alone is not sufficient to classify fiber maturity for genetic analysis.

**FT-IR spectral characteristics of cotton fibers with various MIC values**

To illustrate the difference of IR spectra of cotton fibers having various fiber maturity values, we compared the IR spectra of three typical F₂ progeny plants including the WT-homo (MIC, 5.13), WT-hetro (MIC, 4.65), and im-homo (MIC, 2.09) in the first set cotton materials. Figure 4 showed that apparent ATR FT-IR spectral intensity increased or decreased in the region from 1 100 cm⁻¹ to 650 cm⁻¹. As the MIC values increased from 2.09, 4.65 to 5.13, intensities of the vibrations at 1 055 cm⁻¹ and 1 028 cm⁻¹ due to C-O stretching mode (Liu et al. 2019) decreased, while those in the region between 850 cm⁻¹ and 700 cm⁻¹
increased. Spectral intensity variations of those vibrations have been well characterized in earlier studies (Abidi et al. 2014; Liu and Kim 2015). The intensities of the vibration region between 1100 cm$^{-1}$ and 900 cm$^{-1}$ originate from the stretching modes of C–O and C–C vibrations. The vibrations in the region between 800 cm$^{-1}$ and 700 cm$^{-1}$ are likely due to crystal I$\beta$ form of cotton cellulose (Liu and Kim 2019). The depth of IR light penetration is approximately 1.8 μm–3.3 μm into a fiber bundle sample, and the variations of the IR spectra were detected from the three cotton fibers with different MIC values (Fig. 4).

Thus, the algorithm for determining cotton fiber infrared maturity ($M_{IR}$) utilized the intensity ratios of three IR vibrations at 1500, 1032, and 956 cm$^{-1}$ (Liu et al. 2011). Recently, the $M_{IR}$ values of two distinct fiber sets have been shown a consistency and equivalency of the fiber maturity values that were directly measured by image analysis of cross-sectioned cotton fibers (Liu et al. 2019).

Selection of the second set cotton materials for comparing IR maturity ($M_{IR}$) with MIC and lint percentage

To test if the IR maturity value is compatible to the conventional MIC or lint percentage value, we used the
second set of cotton materials that were also previously constructed by crossing the \( im \) parent plant with multiple WT upland cotton cultivars including Texas Marker-1 (Kohel et al. 1970), Sure-Grow 747 (Lege 1999), Deltapine Acala 90, UA-48 (Bourland 2013), and MDS2ne (Meredith Jr 2005) for studying the relationships of fiber maturity with single fiber breaking force and strength (Kim et al. 2019). This set of cotton materials was composed of 20, 40, and 20 individual \( F_2 \) progeny plants from WT-homo, WT-hetro, and \( im \) homo genotypes, respectively, for an idealized 3:1 ratio of the \( F_2 \) segregation. We measured IR spectra, calculated \( M_{IR} \) values, and compared them with the MIC values for the second set of cotton materials (Fig. 5a). We were able to observe MIC values from 76 \( F_2 \) plants. Four \( F_2 \) plants produced insufficient fiber mass (< 3.24g) for Fibronaire measurement. The MIC range of the 76 \( F_2 \) plants was from 2.09 to 5.52. In contrast, the \( M_{IR} \) value was obtained from all 80 \( F_2 \) plants, and it ranged from 0.39 to 0.93. Based on the algorithm for determining IR maturity (Liu et al. 2011) on the 76 \( F_2 \) plants, the \( M_{IR} \) value was positively \( (r=0.890) \) and significantly \( (P < 0.0001) \) correlated with the MIC values. Similarly, comparison of the \( C_{IR} \) values with their corresponding MIC values showed a positive \( (r=0.675) \) and significant association with the MIC values (Fig. 5b) according to the algorithm for estimating IR crystallinity (Liu et al. 2012). However, the \( r \) and \( R^2 \) values between the \( C_{IR} \) and MIC values were less than those between the \( M_{IR} \) and MIC values (Fig. 5a and b).

**Comparisons of genotypes with three phenotypes including MIC, lint percentage, and IR maturity**

We compared the genotypes of the second set of cotton materials with their phenotypes that were classified by MIC, lint percentage, or \( M_{IR} \) value. Figure 6a is a scatter dot plot that compared the genotypes with the observed MIC phenotypes from the 76 \( F_2 \) plants consisting of WT-homo (20 plants), WT-hetro (40 plants), and \( im \)-homo genotypes (16 plants). The minimum MIC value from the WT phenotypes was 3.74, and the maximum MIC value of the \( im \) phenotype was 3.80. Using the observed MIC values, we were able to identify threshold MIC values (3.74–3.80) and classify the phenotypes of the 74 plants of these 76 \( F_2 \) plants. The observed phenotypes of the 74 plants showed consistency with the expected phenotypes based on their DNA marker genotypes (Fig. 6a). We were unable to classify the phenotypes from the two plants found at the threshold region (3.74–3.80) based on the MIC values alone, the genotype data showed that they consisted of a WT and an \( im \) mutant. The genotype data also predicted the four plants that produced insufficient fiber mass for the MIC measurements were all \( im \) mutant (Fig. 6a).

Unlike the MIC phenotypes in which the threshold value was chosen in the narrow overlapping range between the two phenotypes, the lint percentage phenotypes showed a broad overlapping range between the minimum value of the WT phenotypes (27.4%) and the maximum value of the \( im \) phenotype (30.4%) as shown in Fig. 6b. There were 16 plants in the overlapping region. The other 64 plants showed consistency between the observed and predicted phenotypes. Due to the broad range of the overlapping lint percentages (27.4%–30.4%) between the observed WT and \( im \) phenotypes, it was a challenge to determine the lint percentage threshold. This difficulty exists despite the fact that the lint percentage was obtained from all 80 plants of the second set. When we chose 27.4% as a lint percentage threshold, the observed phenotype ratio between the WT and \( im \) plants was 65:15. The calculated Chi square (1.667) and \( P \) value (0.197) suggested that the observed segregation ratio determined by the lint percentage with the MIC data still fit to the expected ratio by the genotypes.

The \( M_{IR} \) value of the 80 \( F_2 \) plants were compared with their genotypes (Fig. 6c). Interestingly, the WT and \( im \) phenotypes were clearly distinguished by a threshold \( M_{IR} \) value of 0.74 without an obvious overlapping range.
when we excluded one obvious outlier. The observed phenotype ratio (61: 19) between the WT and im plants determined by the \( M_{IR} \) value was very similar to the expected phenotype ratio (60: 20) based on the genotype data. Statistical analysis with the calculated Chi square (0.667) and \( P \) value (0.796) also suggested that the observed segregation ratio by the \( M_{IR} \) value was consistent with the expected ratio by the genotypes.

With the second set cotton materials that were grown in a different field and year, we found the threshold MIC (3.74-3.80) and lint percentage (27.4%) were noticeably greater than the threshold MIC (3.60) and lint percentage (24.0%) of the first set cotton materials that were grown in an ARS field located at Stoneville, MS. In addition, the average MIC (4.98) and lint percentage (34.0%) of the WT-homo genotype in the second set materials were also greater than the average MIC (4.85) and lint percentage (33.5%) of the WT-homo genotype in the first set materials. Similarly, the average MIC (2.98) and lint percentage (23.6%) of the im-homo genotype in the second set materials were also greater than the average MIC (2.64) and lint percentage (15.8%) of the im-homo genotype in the first set materials. Since the MIC, lint percentage, and fiber maturity are greatly affected by environmental conditions (Bradow and Davidonis 2000; Kim et al. 2013b; Kohel and McMichael 1990), we interpreted that the growth and environmental conditions of the second set materials were more favorable than those of the first set materials.

Classifications of the immature fiber phenotypes by a combination of \( M_{IR} \) value with others fiber traits

To overcome the limits of MIC or lint percentage for classifying the im phenotype from the WT phenotype accurately and quantitatively, cotton geneticists have been using a combination of MIC and lint percentage (Kim et al. 2013a; Thyssen et al. 2016). As predicted, the combination of MIC and lint percentage values improved the separation of the im phenotype from the WT phenotype (Fig. 7a). The MIC values were obtained from 76 plants of the total 80 plants, and helped identify the correct threshold lint percentage for classifying im from WT phenotype. When the lint percentage alone was used for phenotyping the second set materials, we chose 27.4% as a threshold, and the observed segregation ratio (65: 15 between WT and im phenotype). With both MIC and lint percentage, the observed ratio (59: 21) between WT and im phenotype almost perfectly fit the expected ratio (60: 20) as shown in Fig. 7a.

We tested if and how a combination of the \( M_{IR} \) value with other fiber traits could improve the classification of the two phenotypes from the second set of cotton materials. The combination of \( M_{IR} \) and MIC clearly distinguished the im phenotype from the WT phenotype (Fig. 7b). The observed phenotype ratio (59: 21) between WT and im determined by \( M_{IR} \) and MIC values was similar to the expected ratio (60: 20). Figure 7c also showed that the combination of \( M_{IR} \) and lint percentage...
can be used to distinguish the \( im \) from WT phenotype clearly. The observed phenotype ratio (59: 20) among WT and \( im \) phenotypes determined by \( M_{IR} \) and lint percentage values was closest to the expected ratio (60: 20) despite one outlier. Another algorithm using different IR spectral vibrations of cotton fibers enabled to determine the IR crystallinity (\( CI_{IR} \)) in addition to the \( M_{IR} \) value. Using the combination of \( M_{IR} \) and \( CI_{IR} \) also enabled to distinguish the \( im \) from WT phenotype (Fig. 7d). The observed phenotype ratio (61: 19) between WT and \( im \) determined by \( M_{IR} \) and \( CI_{IR} \) values was similar to the expected ratio (60: 20) in spite of the difficulty with identifying a few \( im \) plants located at the overlapping area in Fig. 7d. Thus, the FT-IR spectroscopy method alone may provide two fiber traits that can distinguish the \( im \) and WT phenotypes without \( MIC \) or lint percentage measurement.

**Conclusion**

To determine the threshold phenotype accurately, cotton geneticists often measure additional fiber traits for their genetic analyses. Those multiple phenotypic analyses of several thousand \( F_2 \) progeny plants for a fine mapping analysis can be an expensive, laborious and time-consuming process. In order to identify economical, efficient, and expeditious methods for measuring fiber maturity in a quantitative way for genetic analysis, we determined threshold phenotypes between \( im \) and WT phenotypes using a combination of multiple fiber traits determined by HVI, lint percentage, and ATR FT-IR spectra, and compared merits and weaknesses of the methods. Our results showed that the IR maturity (\( M_{IR} \)) index can be used to determine the threshold value for classifying the immature fiber phenotype from the wild type phenotype. The \( M_{IR} \) value can be also used with a combination of \( MIC \), lint percentage, and IR crystallinity for further quantitative genetic analyses.

**Methods**

**Cotton fiber materials and population construction**

We used two sets of upland cotton populations composed of various fiber \( MIC \) and maturity. Both populations were previously constructed by crossing the immature fiber (\( im \)) mutant with \( G. \) hirsutum normal cultivars (wild type) described at Fig. 1a. The first set was constructed by a cross between the \( im \) mutant and a \( G. \) hirsutum improved fiber quality germplasm, MD52ne (Meredith Jr 2005). The \( F_1 \) plants were self-pollinated to obtain \( F_2 \) seeds and the \( F_2 \) population along with the parents were grown in a field located at Stoneville, Mississippi in 2014. The soil type in Stoneville, MS was Bosket fine sandy loam. Standard conventional field practices were applied during planting season. Leaf samples were collected from the individual \( F_2 \) plants as well as parents for DNA isolation. The first set consisting of the \( F_2 \) population of 708 plants was previously used for mapping-by-sequencing to identify the \( im \) gene (Thyssen et al. 2016). The second \( F_2 \) population was produced by crossing the \( im \) mutant with four different WT upland cotton cultivars, Texas Marker-1 (PI 607172; Kohel et al. 1970), Sure-Grow 747 (PVP 9800118; Lege 1999), Deltapine Acala 90 (PI 564767), and UA-48 (PI 660508; Bourland 2013).
The segregating F₂ plants along with parents were grown side by side in the same field in New Orleans, Louisiana from 2011 to 2015. The soil type of the cotton field was Aquents dredged over alluvium in an elevated location to provide adequate drainage. The second set consisting of the F₂ population of the 80 plants had been used for studying relationships of fiber maturity with single fiber strength (Kim et al. 2019).

**Fiber property measurements**
Cotton bolls were manually harvested from the individual F₂ plants as well as parents. Cotton fibers were collected by ginning with a laboratory roller gin. The ginned fibers were conditioned at 21 ± 1°C and 65 ± 2% relative humidity for 48 h before testing (ASTM D1776 / D1776M-16 2016). Phenotypes of the segregating F₂ population were evaluated for lint percentage, fiber properties determined by HVI, Fibronaire, and microscopic measurements. Lint percentage was measured by dividing lint weight with the cottonseed weight, and multiplying by 100. HVI 1000 (Uster Technologies Inc., Knoxville, TN) was used to measure fiber properties from the individual F₂ plants producing more than 10 g. Average HVI values were obtained from five replicates. The Fibronaire instrument (Motion Control Inc., Dallas, TX) was used to measure MIC values for the F₂ progenies that produced insufficient lint mass (3.3~10.0 g) for HVI measurement. Both instruments were properly calibrated according to the manufacturers’ instructions and standard procedure (ASTM D5867-12e1 2012).

**ATR-FTIR spectral collection and data analysis**
All spectra from the second set of cotton materials were collected with an FTS 3000MX FTIR spectrometer (Varian Instruments, Randolph, MA) equipped with a ceramic source, KBr beam splitter, and deuterated triglycine sulfate (DTGS) detector. The ATR sampling device utilized a DuraSampIR single-pass diamond-coated internal reflection accessory (Smiths Detection, Danbury, CT), and a consistent contact pressure was applied by way of a stainless-steel rod and an electronic load display. At least six measurements at different locations for individual samples were collected over the range of 4000–600 cm⁻¹ at 4 cm⁻¹ and 16 coadded scans. All spectra were given in absorbance units and no ATR correction was applied. Following the import to GRAMS IQ application in Grams/AI (Version 9.1, Thermo Fisher Scientific, Waltham, MA), the spectra were smoothed with a Savitzky–Golay function (polynomial = 2 and points = 11). Then, the spectral set was loaded into Microsoft Excel 2007 to assess cotton Fiber maturity M₀ from IR measurement by using a previously proposed algorithm analysis (Liu and Kim 2015; Liu et al. 2011).

**Genotyping of the F₂ plants by SSR and InDel markers**
For genotyping of the F₂ plants, InDel and SNP (Thysen et al. 2016) and SSR (Kim et al. 2013a) markers linked to the im gene on chromosome 3 were used. The forward primers were fluorescent-labeled at 5’ end with 6-FAM (6-carboxyfluorescein), or HEX (4, 7, 2’, 4’, 5, 7-hexachloro-carboxyfluorescein). Primers were purchased from Sigma Genosys (Woodlands, TX). PCR amplification was according to the method that was previously described in Fang et al. (2010). Amplified PCR products were separated and measured on an automated capillary electrophoresis system ABI 3730 XL (Applied Biosystems Inc. Foster City, CA). GeneScan-400 ROX (Applied Biosystems Inc. Foster City, CA) was used as an internal DNA size standard.

**Statistical analyses**
Statistical analyses and construction of graphs were performed using correlation, linear regression, frequency distribution from Prism version 7 software (Graph-Pad Software, Inc., San Diego, CA). The correlation coefficient value (r) was determined by Pearson’s method (Pearson 1895). The P value cutoff for significance was 0.05. Samples in individual fiber sets were fitted to an exponential function by the use of Microsoft Excel 2007.

**Additional files**

**Additional file 1:** Comparisons of the observed infrared maturity (M₀, IM) and crystallinity (CI₀, IC) with the predicted phenotypes based on their DNA marker genotypes. ATR FT-IR spectra were measured from the second set composed of 80 F₂ plants with various MIC values. M₀ and CI₀ values were determined by algorithms and compared with the predicted phenotypes based on their DNA marker genotypes as well as the observed MIC and lint percentage. (XLSX 61 kb)

**Additional file 2:** Comparisons of the observed infrared maturity (M₀) and crystallinity (CI₀) with the predicted phenotypes based on their DNA marker genotypes. ATR FT-IR spectra were measured from the second set composed of 80 F₂ plants with various MIC values. M₀ and CI₀ values were determined by algorithms and compared with the predicted phenotypes based on their DNA marker genotypes as well as the observed MIC and lint percentage. (XLSX 15 kb)

**Abbreviations**
ATR FTIR: Attenuated total reflection fourier transform infrared; CI₀: Cotton fiber infrared crystallinity; F₁: First filial generation; F₂: Second filial generation; HVI: High volume instrument; IAM: Image analysis microscopy; im: Immature fiber; im-homo: Homozygosity for the im type; M₀: Micronaire; Mₚ: Cotton fiber infrared maturity index; MR: Maturity ratio; PCW: Primary cell wall; r: Correlation coefficient value; SCW: Secondary cell wall; SNP: Single nucleotide polymorphism; SSR: Simple sequence repeats; U.M.: Unmeasurable; WT: Wild type; WT-hetero: Heterozygosity for the wild type; WT-homo: Homozygosity for the wild type; R: Circularity; χ²: Chi-square.

**Acknowledgements**
Authors acknowledge Zeng LH of USDA-ARS at Stoneville, Mississippi for growing the first set cotton materials. We also thank Moiseyev R and King...
H for phenotyping, Li P for genotyping, and Buttsam W and Stevenson K for preparing cotton fields for the second set cotton materials. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U. S. Department of Agriculture which is an equal opportunity provider and employer.

Authors’ contributions
Kim HJ and Liu Y conceived the research. Kim HJ constructed the cotton populations. Liu Y, Fang DD, Kim HJ, and Delhom CD were responsible for FT-IR, genotype, lint percentage, and MIC measurements, respectively. Kim HJ analyzed the data and wrote the paper. All authors read, revised, and approved the manuscript.

Funding
This research was supported by the USDA-ARS Research Project # 6054–21000-017-00D, and Cotton Incorporated-sponsored project #19–858.

Availability of data and materials
All supporting data can be found within the manuscript and its Additional files 1 and 2.

Ethics approval and consent to participate
No materials of human or animals were used. Authors followed the experimental research guidelines of USDA ARS and Louisiana Department of Agriculture & Forestry on cotton plant growth and material collections.

Consent for publication
All data were collected by the authors who consented for submission of the manuscript according to the USDA ARS guidance.

Competing interests
The research described in this manuscript was in part funded by a grant from Cotton Incorporated. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This does not alter the authors’ adherence to all polices on sharing data and materials.

Author details
1USDA-ARS, Southern Regional Research Center, Cotton Fiber Bioscience Research Unit, 1100 Robert E. Lee Blvd, New Orleans, LA 70124, USA.
2USDA-ARS, Southern Regional Research Center, Cotton Structure and Quality Research Unit, 1100 Robert E. Lee Blvd, New Orleans, LA 70124, USA.

Received: 29 April 2019 Accepted: 17 June 2019
Published online: 04 July 2019

References
Abidi N, Cabrales L, Haigler CH. Changes in the cell wall and cellulose content of developing cotton fibers investigated by FTIR spectroscopy. Carbohydr Polym. 2014;100(9):1–16. https://doi.org/10.1016/j.carbpol.2013.01.074.

Abidi N, Cabrales L, Haigler CH. Changes in the cell wall and cellulose content of developing cotton fibers investigated by FTIR spectroscopy. Carbohydr Polym. 2014;100(9):1–16. https://doi.org/10.1016/j.carbpol.2013.01.074.

Kim HJ, Delhom CD, Rodgers J, et al. Effect of fiber maturity on bundle and single-fiber strength of upland cotton. Crop Sci. 2019;59:115–24. https://doi.org/10.2136/cropsci2018.05.0324.

Abidi N, Hequet E, Cabrales L, et al. Evaluating cell wall structure and composition of developing cotton fibers using Fourier transform infrared spectroscopy and thermogravimetric analysis. J Appl Polym Sci. 2008;107(1):476–86. https://doi.org/10.1002/app.27100.

Abidi N, Manike M. X-ray diffraction and FTIR investigations of cellulose deposition during cotton fiber development. Text Res J. 2018;88(7):719–30. https://doi.org/10.1177/0040517517768654.

ASTM D1776 / D1776M-16. Standard practice for conditioning and testing textiles. West Conshohocken: ASTM International; 2016. www.astm.org. https://doi.org/10.1520/D1776_D1776M-16.

ASTM D5867-12e1. Standard test methods for measurement of physical properties of raw cotton by cotton classification instruments. West Conshohocken: ASTM International; 2012. www.astm.org. https://doi.org/10.1520/D5867-12e1.

Bourland F. Cotton cultivar UA-48. United States patent US 8492618; 2013. [2012-04-12].

Bradov JM, Davidson GH. Quantitation of fiber quality and the cotton production–processing interface: a physiologist’s perspective. J Cotton Sci. 2002;434–64.

Bridge RR, Meredith WR Jr, Chism JF. Comparative performance of obsolete varieties and current varieties of upland cotton. Crop Sci. 1971;11:29–32.

Fang DD, Xiao J, Canci PC, et al. A new SNP haplotype associated with blue disease resistance gene in cotton (Gossypium hirsutum L.). Theor Appl Genet. 2010;120(5):943–53. https://doi.org/10.1007/s00122-010-1223-y.

Frydrych I, Thibodeaux DP. Fiber quality evaluation: current and future trends/ intrinsic value of fiber quality in cotton. In: Wakeley PJ, Chaudhry MR, editors. Cotton: technology for the 21st century. Washington DC: International Cotton Advisory Committee; 2010. p. 251–96.

Gordon S, Rodgers J. Cotton fibre cross-section properties. In: Gordon S, Abidi N, editors. Cotton fibres, characteristics, uses and performance. New York: Nova Science publishers, Inc; 2017. p. 65–86.

Hequet EF, Wyatt B, Abidi N, et al. Creation of a set of reference material for cotton fiber maturity measurements. Text Res J. 2006;76(7):756–86. https://doi.org/10.1177/004051750601600109.

Islam MS, Fang DD, Thysen GN, et al. Comparative fiber property and transcriptome analyses reveal key genes potentially related to high fiber strength in cotton (Gossypium hirsutum L.) line MD52ne. BMC Plant Biol. 2016;16(1):36. https://doi.org/10.1186/h12870-016-0272-2.

Kelly B, Abidi N, Ethridge D, et al. Fiber to fabric. In: Fang DD, Percy RG, editors. Cotton. 2nd ed. Madison: American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America; 2015. p. 665–744. https://doi.org/10.2135/cropsy2016.04.015.

Kohel RJ, Quisenberry JE, Benedict CR. Fiber elongation and dry weight changes in mutant lines of cotton. Crop Sci. 1974;14:471-476.

Kohler R, Richmond T, Lewis C. Texas marker-1. Description of a genetic standard for Gossypium hirsutum L. Crop Sci. 1990;30(2):419–24.

Kohler RJ, McMichael SC. Immature fiber mutant of upland cotton. Crop Sci. 1990;30(2):419–21.

Kohler RJ, Quisenberry JE, Benedict CR. Fiber elongation and dry weight changes in mutant lines of cotton. Crop Sci. 1974;14:471–4.

Kothari N, Abidi N, Hequet E, et al. Fiber quality variability within a plant. In: World cotton research conference-4, Lubbock; 2007.

Lege KE. Sure-grow 747, a new early-maturing picker variety. In: Proceedings of the beltwide cotton conferences, vol. 1; 1999. p. 69–71.

Liu Y, Kim HJ. Use of attenuated total reflection Fourier transform infrared (ATR FT-IR) spectroscopy in direct, non-destructive, and rapid assessment of developmental cotton fibers grown in planta and in culture. Appl Spectrosc. 2015;69:1004–10. https://doi.org/10.1366/15-07876.

Liu Y, Kim HJ. Comparative investigation of secondary cell wall development in cotton fiber near isogenic lines using attenuated total reflection Fourier
transform infrared spectroscopy (ATR FT-IR), Appl Spectrosc. 2019;73(3):329–36. https://doi.org/10.1177/0003702818818171.

Liu Y, Kim HJ, Delhom CD, et al. Investigation of fiber maturity measurement by cross-sectional image analysis and Fourier transform infrared spectroscopy on developing and developed upland cottons. Cellulose. 2019; Accepted. https://doi.org/10.1007/s10570-019-02502-8.

Liu Y, Thibodeaux D, Gamble G. Development of Fourier transform infrared spectroscopy in direct, non-destructive, and rapid determination of cotton fiber maturity. Text Res J. 2011;81:1559–67. https://doi.org/10.1177/0040517511410107.

Liu Y, Thibodeaux D, Gamble G, et al. Comparative investigation of Fourier transform infrared spectroscopy and X-ray diffraction (XRD) in the determination of cotton fiber crystallinity. Appl Spectrosc. 2012;66(8):983–6. https://doi.org/10.1366/12-06611.

McCann MC, Hammouri M, Wilson R, et al. Fourier transform infrared spectroscopy is a new way to look at plant cell walls. Plant Physiol. 1992;100(4):1940–7.

Meredith WR Jr. Quantitative genetics. In: Kohel RJ, Lewis CL, editors. Cotton agronomy monograph 24, 1984. Madison: American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America; 1984. p. 131–50.

Meredith WR Jr. Registration of a high fiber quality germplasm, MD52ne and its recurrent parent, MD90ne. Crop Sci. 2005;45:807–8.

Meredith WR Jr, Bridge RR. Recurrent selection for lint percent within a cultivar of cotton (Gossypium hirsutum L.). Crop Sci. 1973;13(6):698–701.

Pearson K. Contributions to the mathematical theory of evolution. III. Regression, heredity, and panmixia. Proc R Soc Lond. 1895;59(353–358):69–71.

Peirce F, Lord E. The fineness and maturity of cotton. J Text Inst. 1939;30(12):T173–210.

Rodgers J, Delhom C, Fortier C, et al. Rapid measurement of cotton fiber maturity and fineness by image analysis microscopy using the Cottonscope. Text Res J. 2011;81(9):259–71. https://doi.org/10.1177/0040517511410107.

Thibodeaux DP, Evans JP. Cotton fiber maturity by image analysis. Text Res J. 1986;56(2):130–9.

Thyssen GN, Fang DD, Zeng L, et al. The immature fiber mutant phenotype of cotton (Gossypium hirsutum) is linked to a 22-bp frame-shift deletion in a mitochondrial targeted pentatricopeptide repeat gene. G3. 2016;6(6):1627–33. https://doi.org/10.1534/g3.116.027649.

Wan Q, Zhang Z, Hu M, et al. T locus in cotton is the candidate gene affecting lint percentage, fiber quality and spiny bollworm (Earias spp.) resistance. Euphytica. 2007;158(1–2):241–7. https://doi.org/10.1007/s10681-007-9446-y.

Wang C, Zhang T, Guo W. The im mutant gene negatively affects many aspects of fiber quality traits and lint percentage in cotton. Crop Sci. 2013;53(1):27–37. https://doi.org/10.2135/cropsci2012.05.0304.