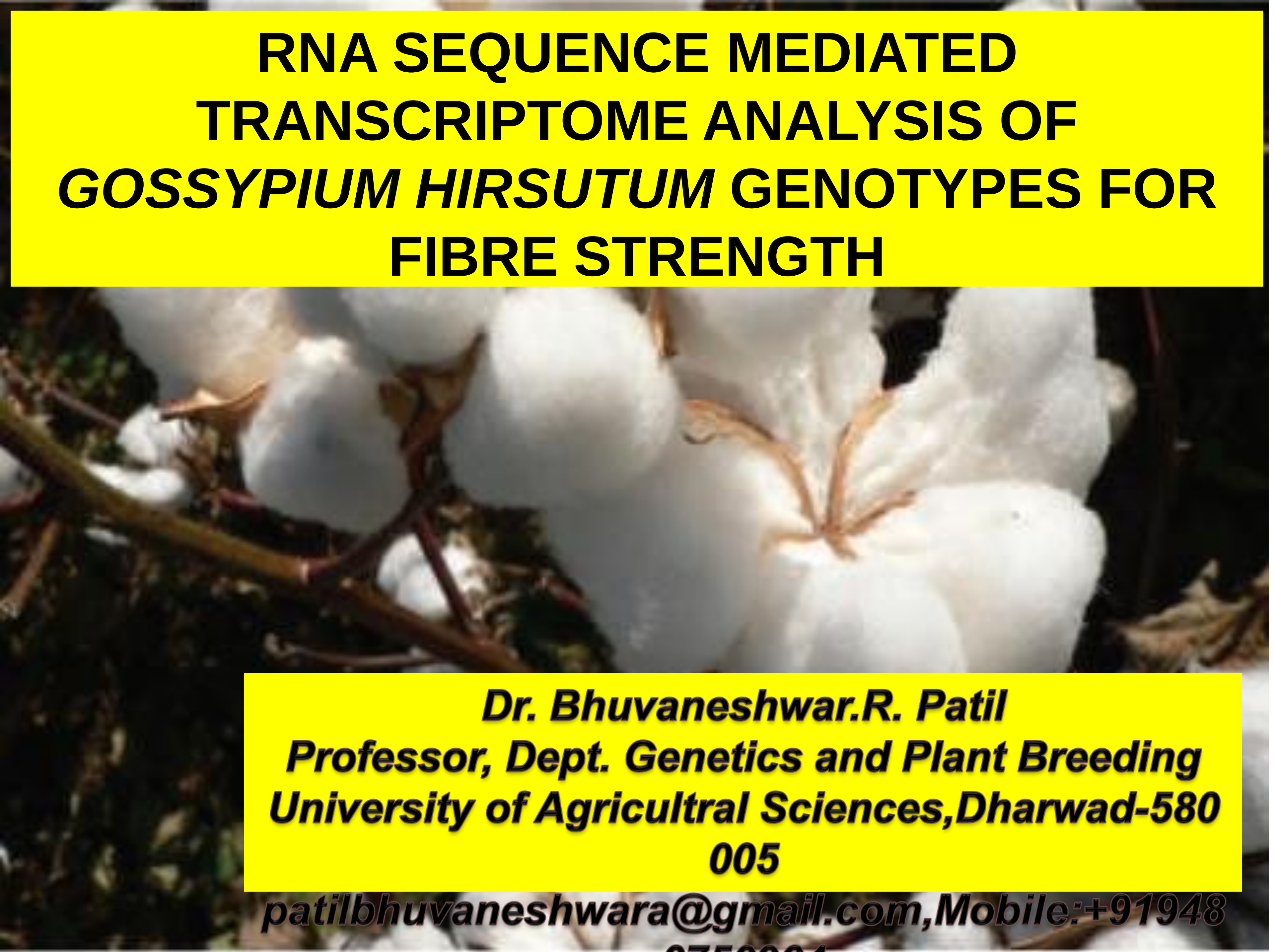


RNA SEQUENCE MEDIATED TRANSCRIPTOME ANALYSIS OF *GOSSYPIUM HIRSUTUM* GENOTYPES FOR FIBRE STRENGTH

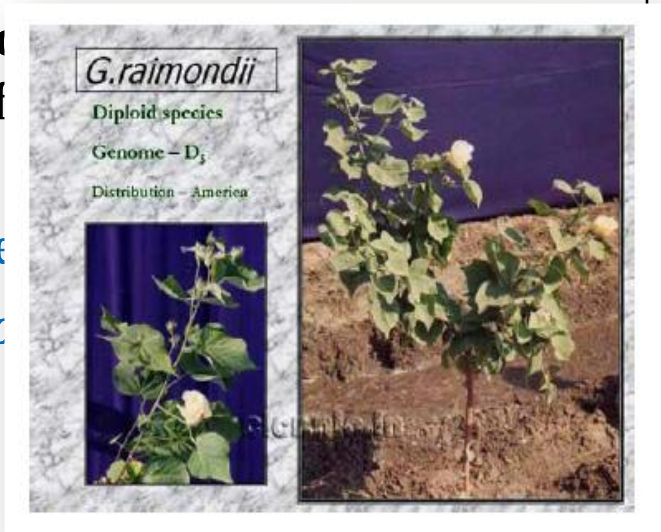
Dr. Bhuvaneshwar.R. Patil
Professor, Dept. Genetics and Plant Breeding
University of Agricultral Sciences,Dharwad-580 005
patilbhuvaneshwara@gmail.com,Mobile:+91948

Introduction

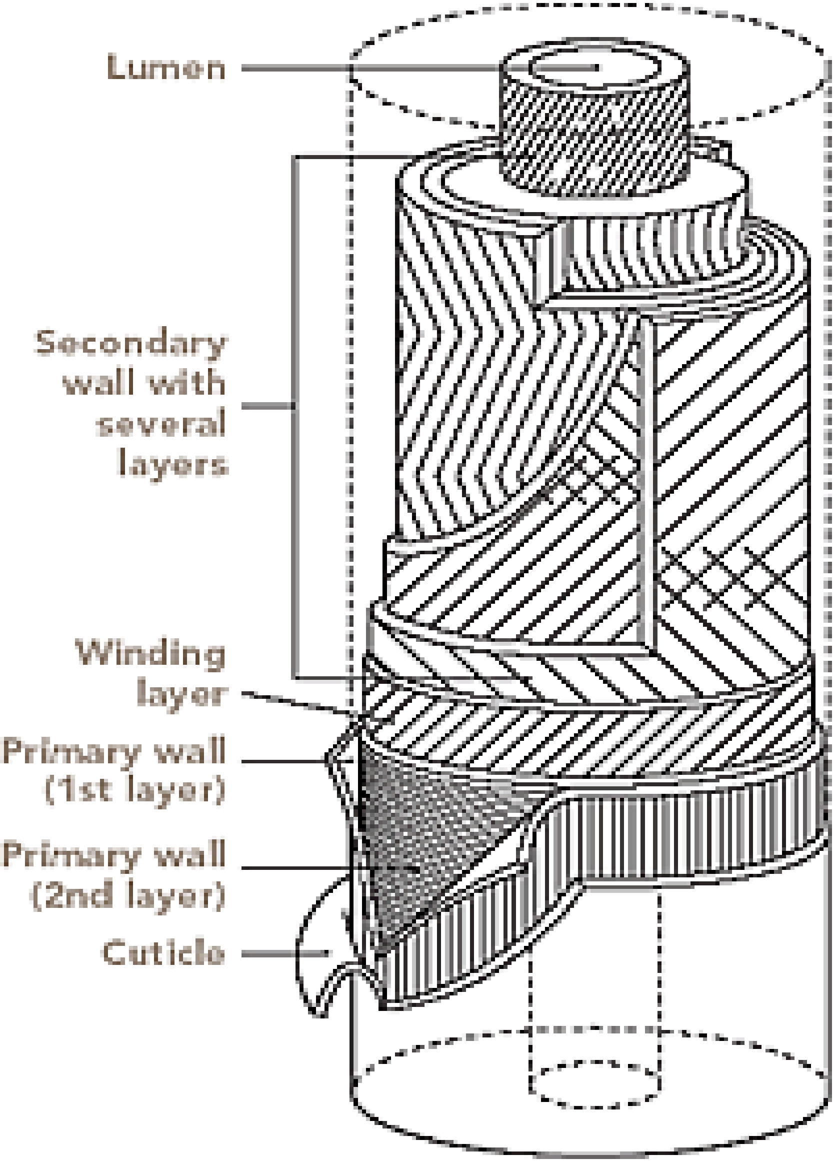
- **Fiber strength in cotton is major quality trait with textile sector importance.**
- **A large proportion of the cotton genome expresses during both fibre elongation and subsequent secondary cell wall thickening.**
- **Recently there has been considerable investment in investigating the global molecular and biochemical processes underlying fiber development through both expressed sequence tag (EST) analysis and microarray-based differential expression analyses(Udall et al. 2006)**
- **In recent years, cotton functional genomics has gained much new insight into fiber development, particularly in fiber expansion, elongation and secondary cell wall synthesis.**
- **Genome-wide quantification of mRNA transcripts is highly informative about cellular states and has helped to elucidate regulatory networks .**
- **Transcriptomes have been profiled by large-scale sequencing of the expressed sequence tags (EST) from elongating cotton**

Cotton Genome Sequencing: an update

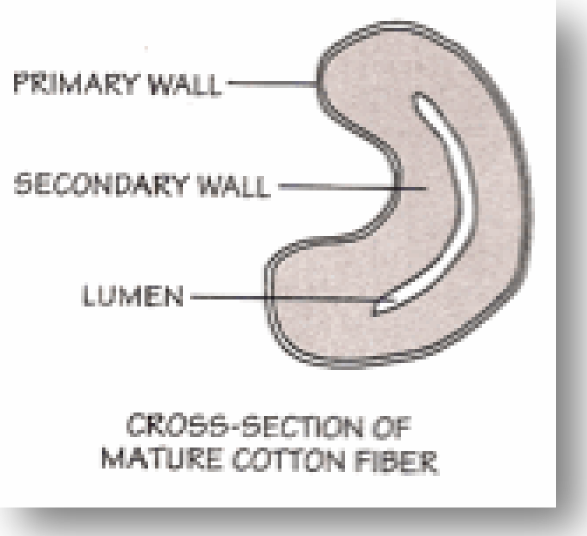
- **Wang, K. *et al.* 2012. The draft genome of a diploid cotton *Gossypium raimondii*. *Nat. Genet.* 44, 1098–1103.**
- **Paterson, A.H. *et al.* 2012. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. *Nature* 492, 423–427.**
- **Li, F.G. *et al.* 2015.Genome sequence of the cultivated cotton *Gossypium arboreum*. *Nat. Genet.* 46, 567–572.**
- **Li, F.G.et al 2015.Genome sequence of cultivated Upland cotton (*Gossypium hirsutum*TM-1) provides insights into genome evolution *Nat. Biotech.*Volume:33,Pages: 524–530.**
- **Zhang,T et al 2015. Sequencing of allotetraploid c L. acc. TM-1) provides a resource for f *Biotech.*Volume: 33,Pages: 531–537.**
- **Hendrix, B. & Stewart, J.M. 2005. Estimation of the nuclear DNA content of *Gossypium* species. *Ann. Bo* 789–797.**

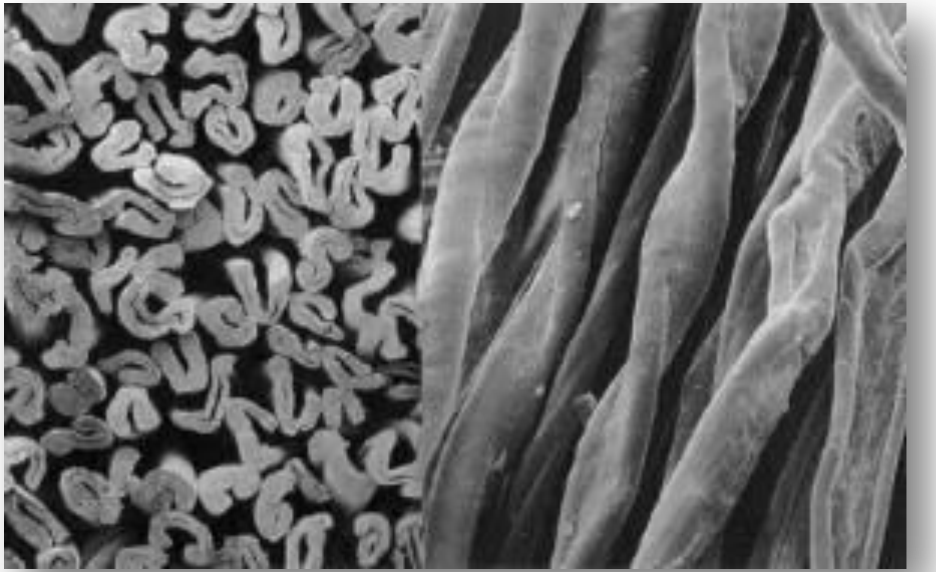


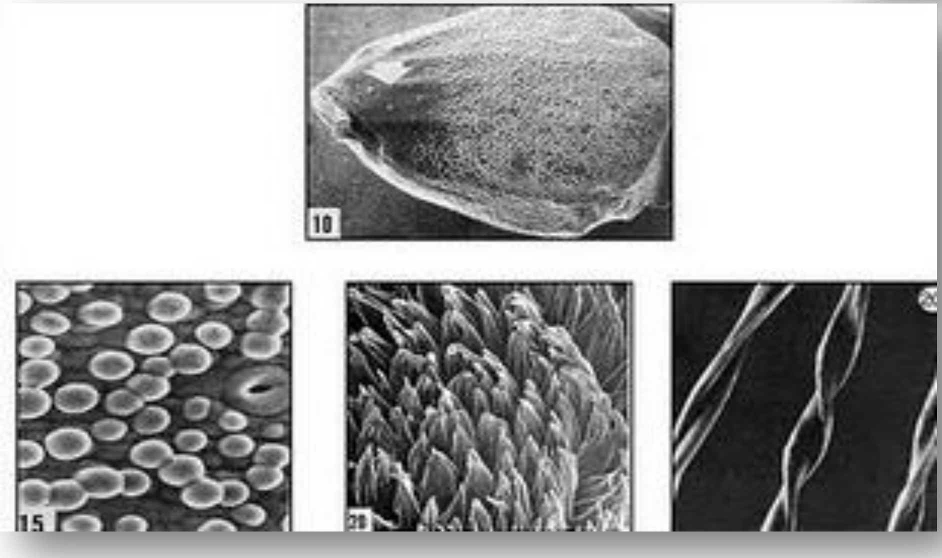
Structure of fiber…



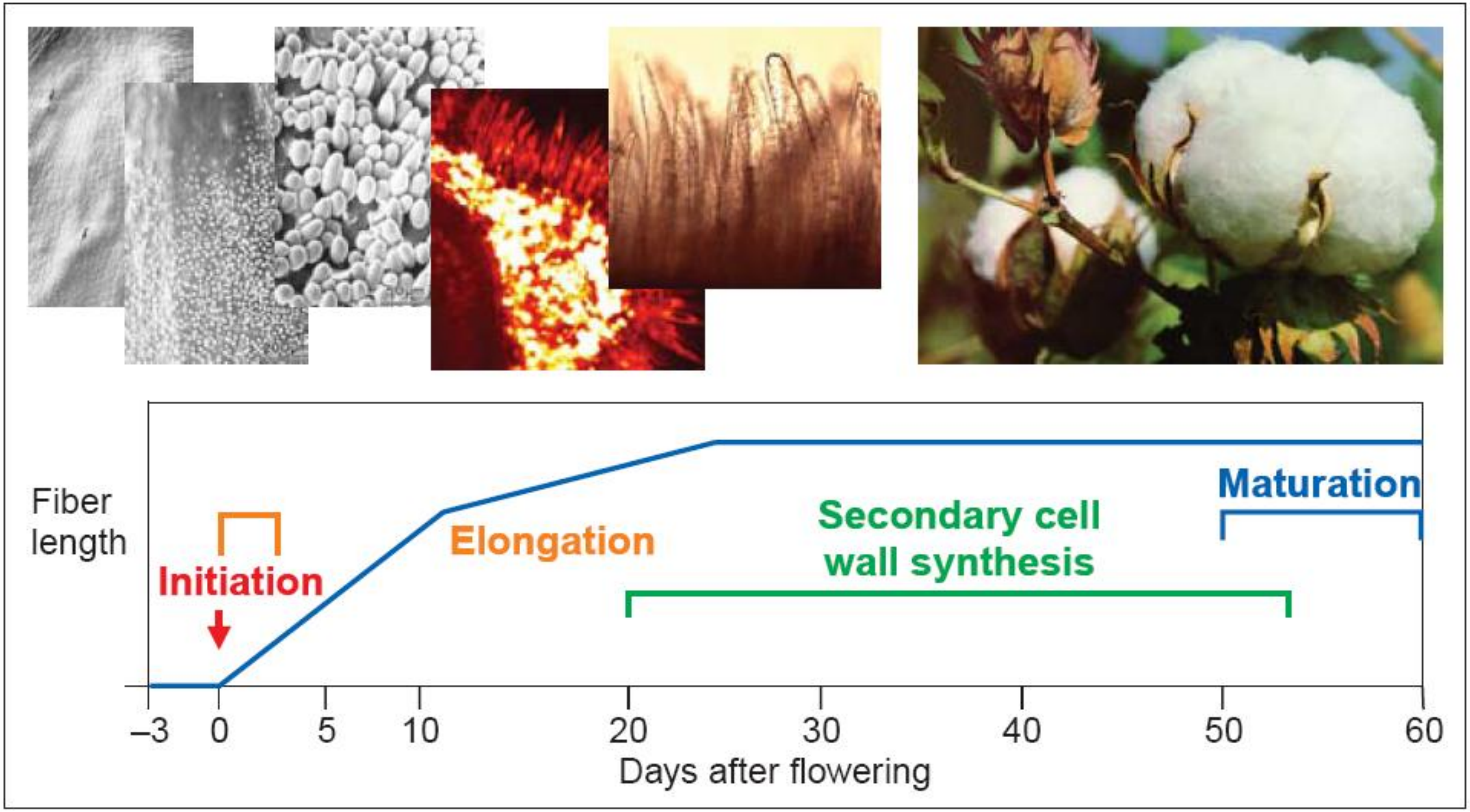
Microscopic structure of fiber…







Developmental stages of fiber…



Arioli (2005)

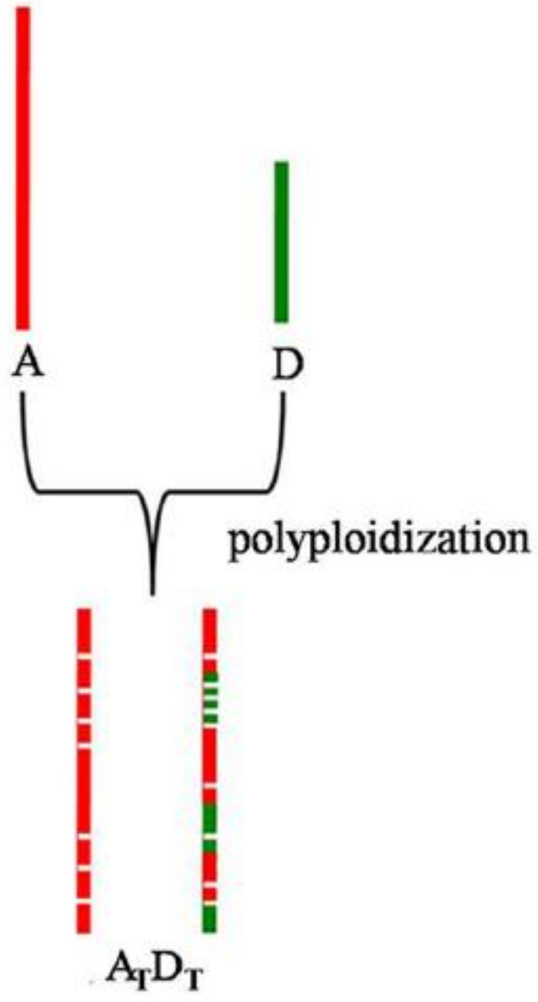


Figure 5. Schematic drawing that illustrates the evolution of most chromosomes. This figure is available in black and white in print and in colour at *DNA Research* online.

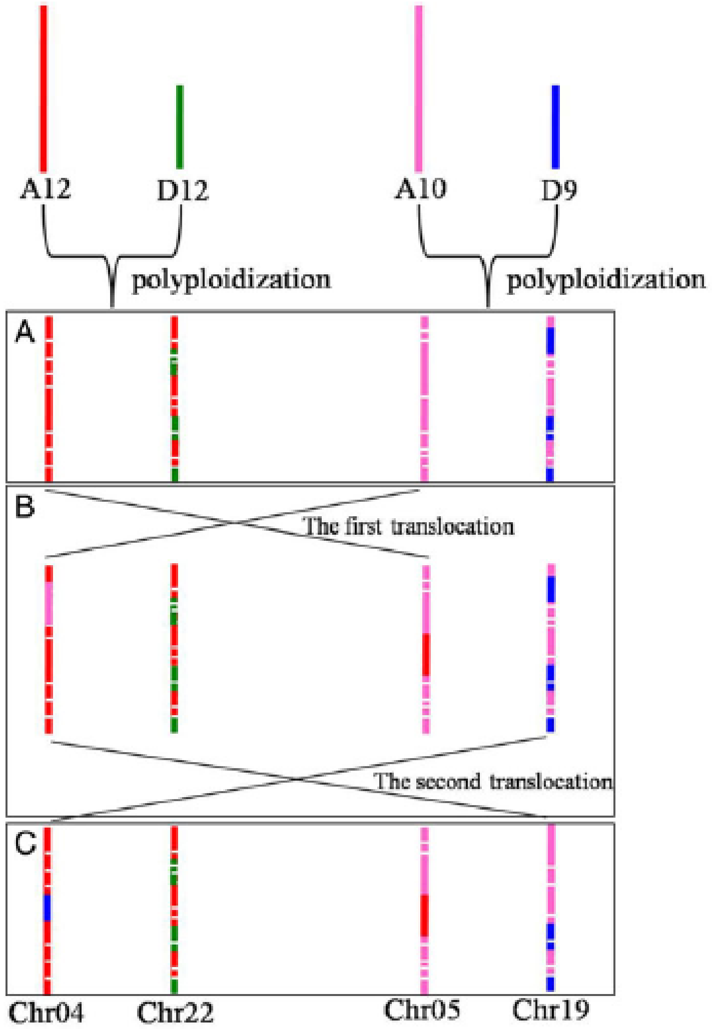


Figure 6. Schematic drawing that illustrates the evolution of Chr04/Chr22 and Chr05/Chr19. This figure is available in black and white in print and in colour at *DNA Research* online.

- **Molecular evidence suggests that the tetraploid cottons arose on the Amerian continent during the pleistocene period through a unique interspecific hybridization event which explains the monophyletic origin of tetraploid cotton involving an A genome and D genome species (Wendel and Cronn,2003).**

- **The A genome of the AD tetraploid species would derive from a type close to G.herbaceum, which would have migrated to the Meso-American region via the Pacific,thanks to the capacity of lint bearing seeds to float and withstand long periods of imbibitions in sea water without loss of power of germination (Fryxell,1976).**

- **The cytoplasm of tetraloid species would originate from the A-genome progenitor (Wendel and Albert, 1992) ,while the D genome would originate from a D-genome species nowadays extinct, but close to that of G.raimondii. Following this unique hybridization event, three main branches (G.mustelinum; G. Hirsutum-G.tomentosum; G.barbadense-G.darwinii) would have diverged (Wendel *et al* 1994).**

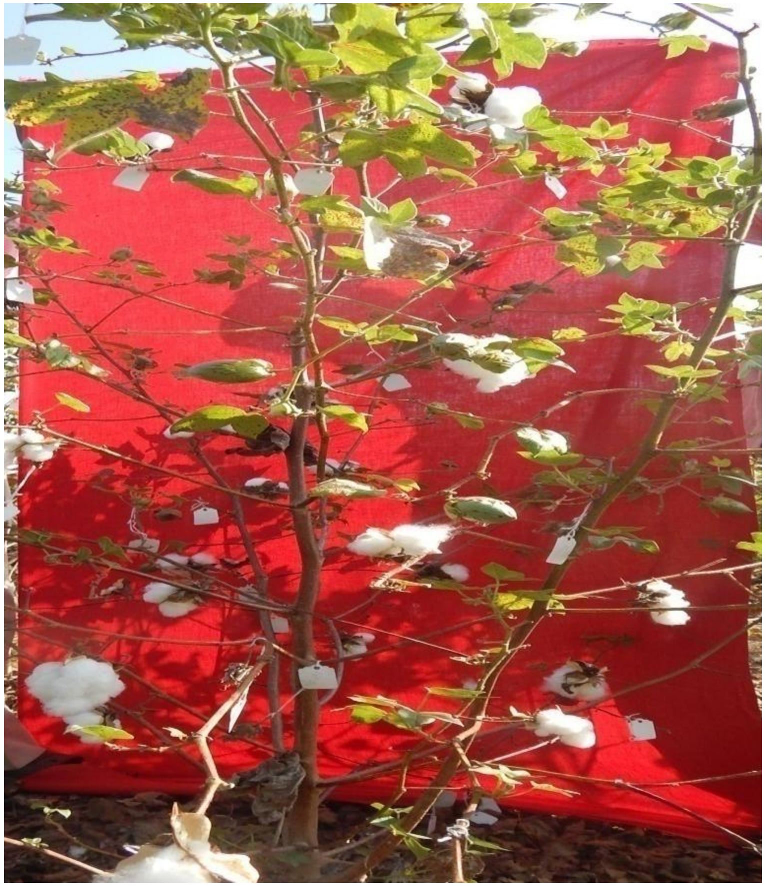
- **After an allopolyploidization event between the A-genome and the D-genome, the formed tetraploid upland cotton markedly changed its plant morphology to economically significant properties, suggesting that the D-subgenome may also contribute specific to fiber quality traits. (Sun *et al*., 2012 ,Zhang *et al*., 2015)**

- **Although the D-genome species of *Gossypium* donot produce spinnable fiber and the genome size of the D-genome is only half of that of the A-genome, it is estimated that both have same total numbers of genes [Wang *et al*., 2012, Paterson *et al*., 2012, Li *et al*., 2014].**

Fibre quality parameters of *G hirsutum* cotton genotypes used in sequencing

HBS144

HBS187

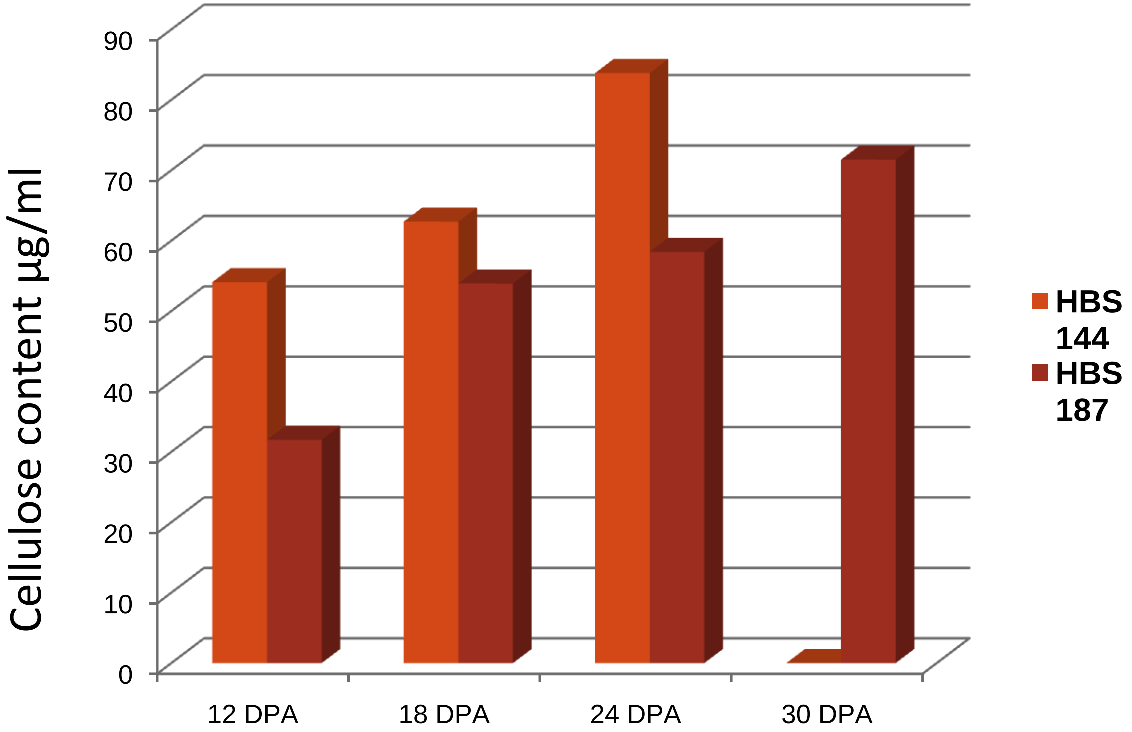


Sl. No.	Character	HBS-187	HBS-144
1	Fibre strength (g/tex)	20.00	28.00
2	Micronaire index (µg/inch)	3.7	3.5
3	Uniformity ratio (%)	82.2	83.6

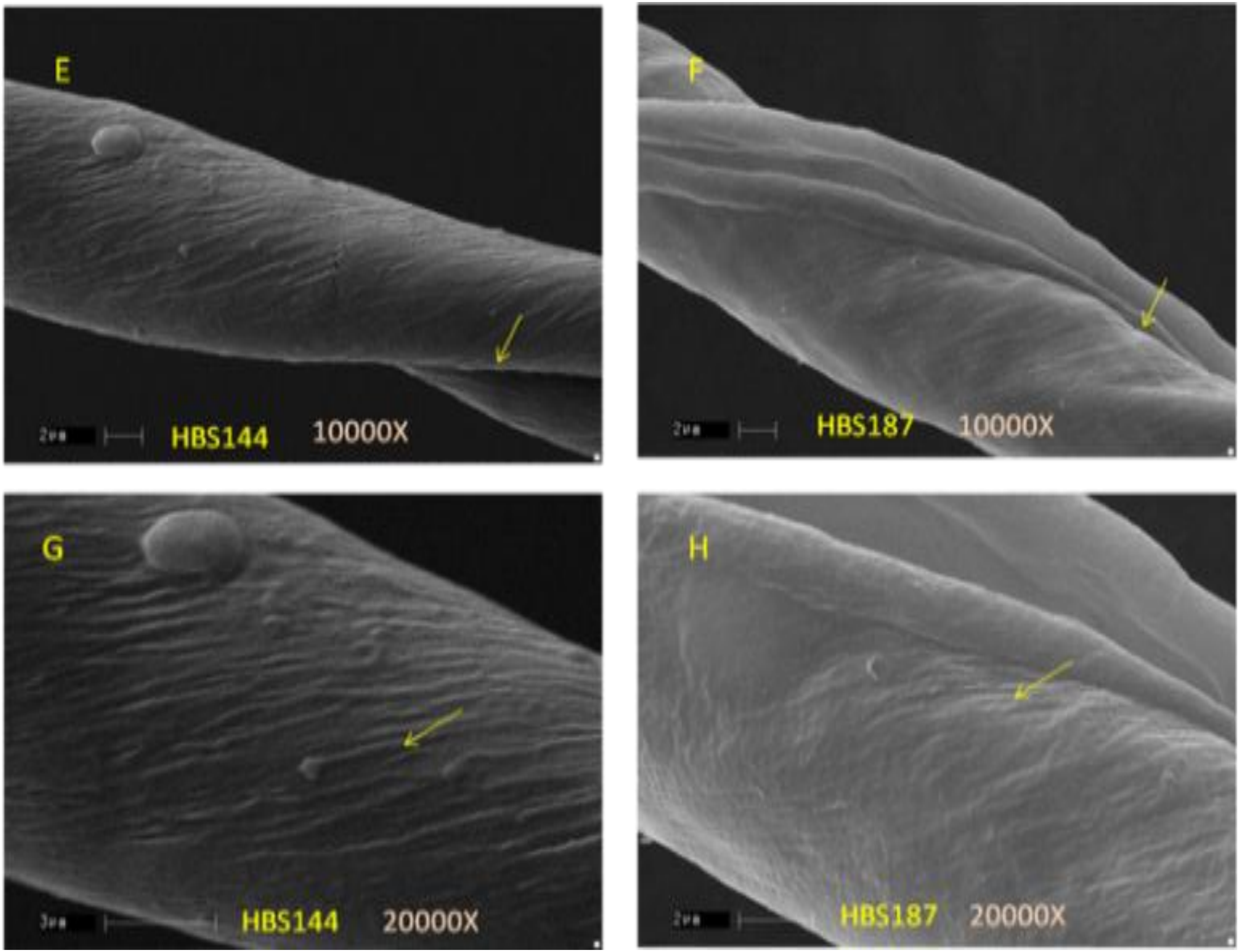
Cellulose Estimation: Updegroff DM (1969) protocol.

S.No	Genotypes	Cellulose content (µg/ml)			
		12DPA	18DPA	24DPA	30DPA
1	HBS 144	54.1	62.7	83.8	110.5
2	HBS 187	31.7	53.9	58.4	71.5

Eppendorf Biospectrometer (Kinetic)



HBS 144:28,
HBS 187:20

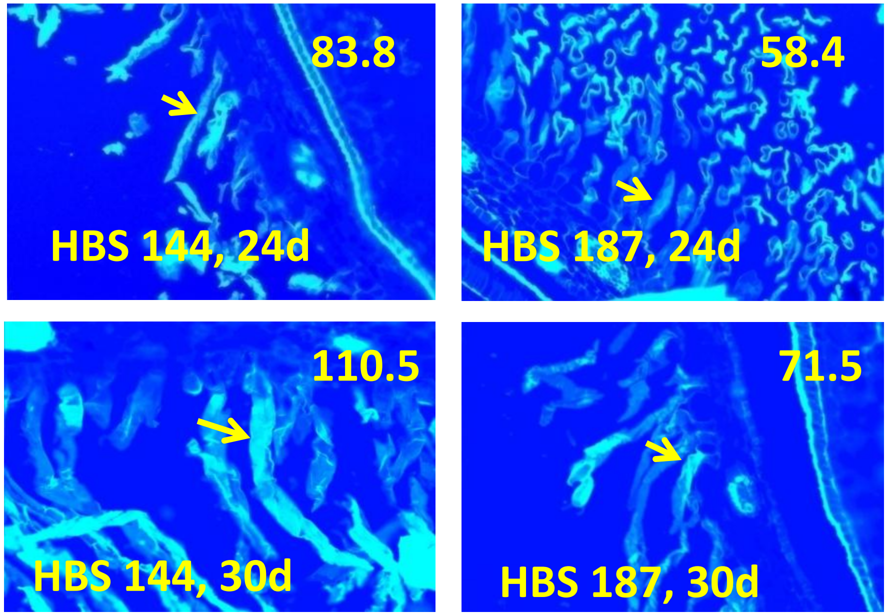

E
2μm
HBS144
10000X
F
2μm
HBS187
10000X
G
3μm
HBS144
20000X
H
2μm
HBS187
20000X

Thickness of the cotton fibre

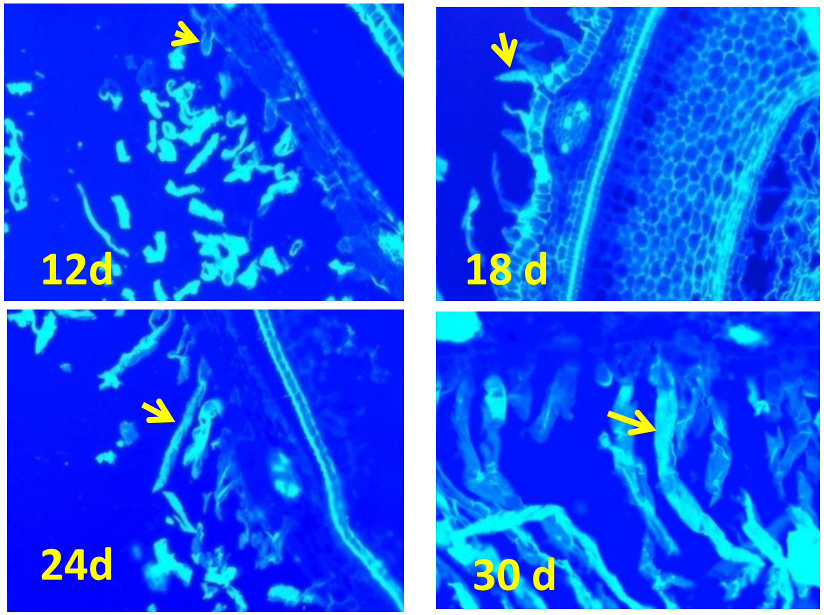
Sl. No.	Genotype	Fibre Diameter (µm)	fibril Diameter (µm)
1	HBS 144	17	0.2
2	HBS 187	10	0.1

These fibres contain rough thick fibrillar coils and formed the folds on fibre surfaces and their inclination is related to the mechanism of cellulose deposition as spiral structures.

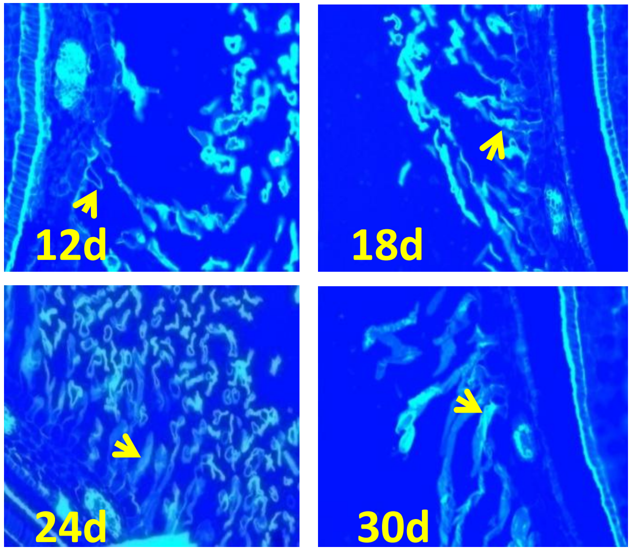
CELLULOSE,ARABINOGALACTAN AND GALACTAN CONTENT



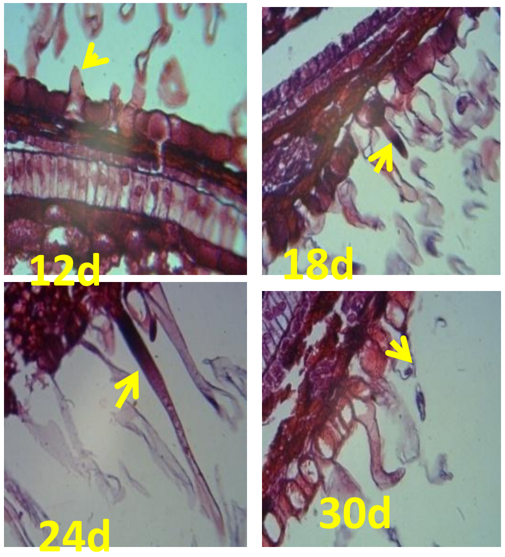
HBS 144 and HBS 187 - The seed sections stained with Calcofluor white stain



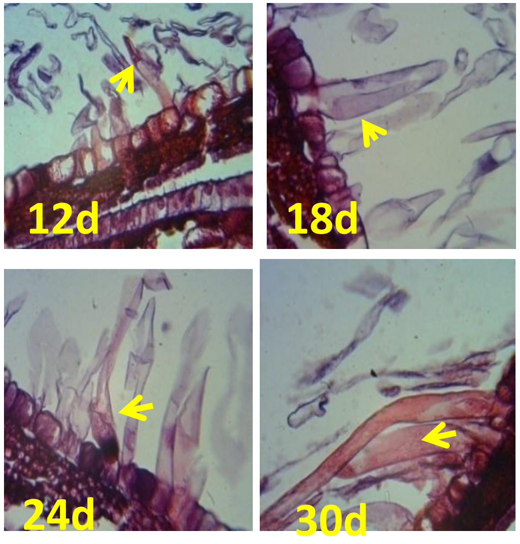
HBS 144- The seed sections stained with Calcofluor white stain



HBS 187- The seed sections stained with Calcofluor white stain



HBS 144- The seed sections stained with safranin

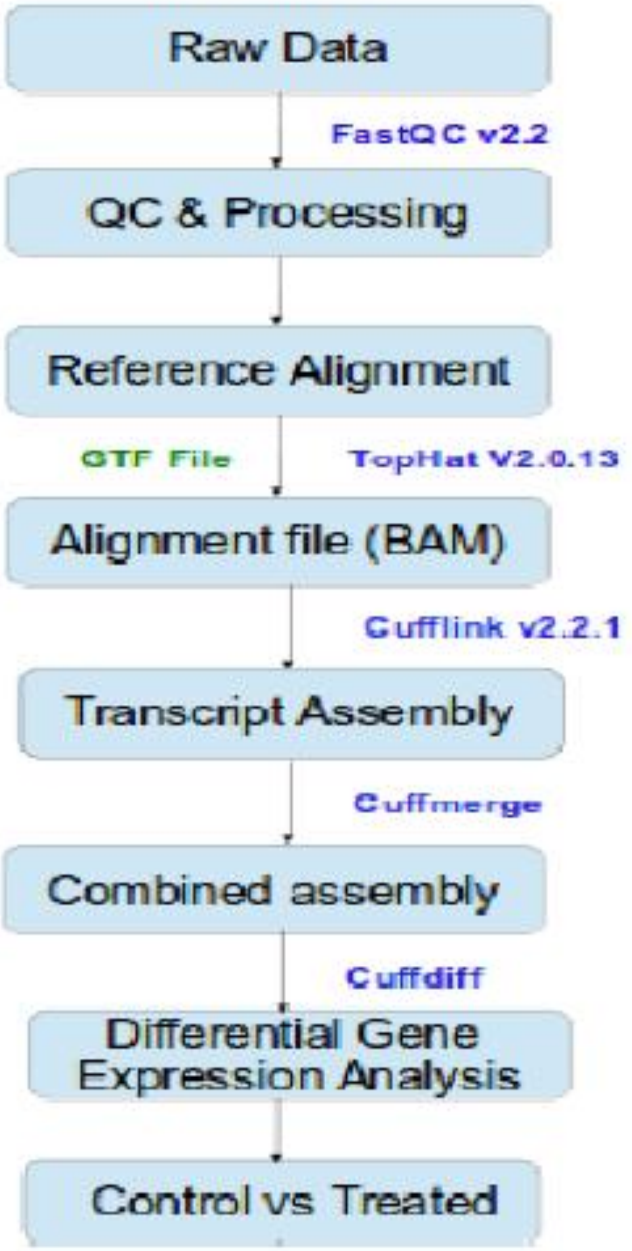


HBS 187- The seed sections stained with safranin

Materials and methods…

- mRNA was isolated from HBS-144 (28.00 g/tex) and HBS187 (20.00 g/tex) *G. hirsutum* lines. at 12, 18, 24 and 30 DPA.
- mRNA converted to cDNA used for sequencing.
- cDNA library was generated for transcriptome sequencing.

Bioinformatics pipeline



Raw Data

FastQC v2.2

QC & Processing

Reference Alignment

GTF File

TopHat V2.0.13

Alignment file (BAM)

Cufflink v2.2.1

Transcript Assembly

Cuffmerge

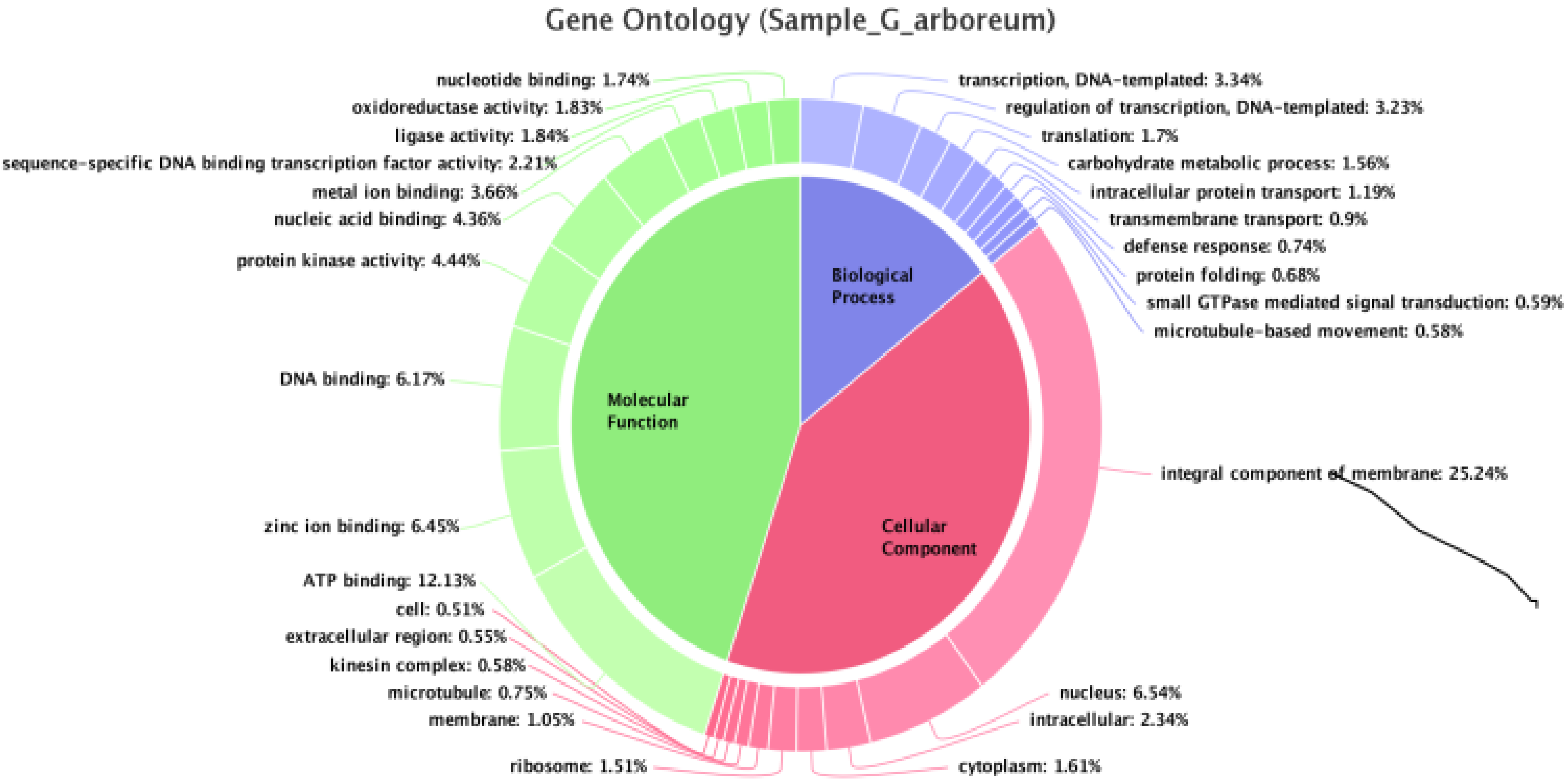
Combined assembly

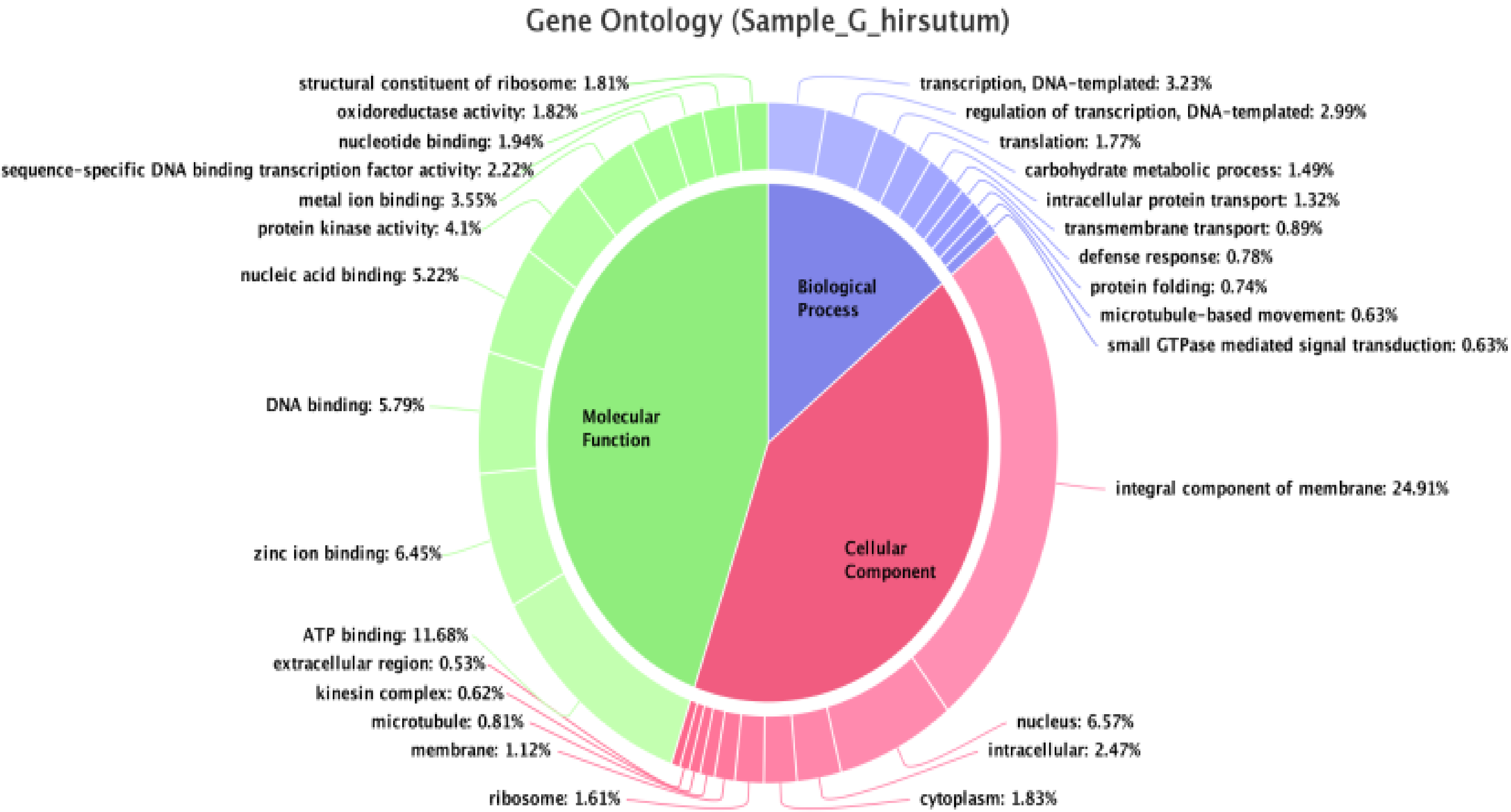
Cuffdiff

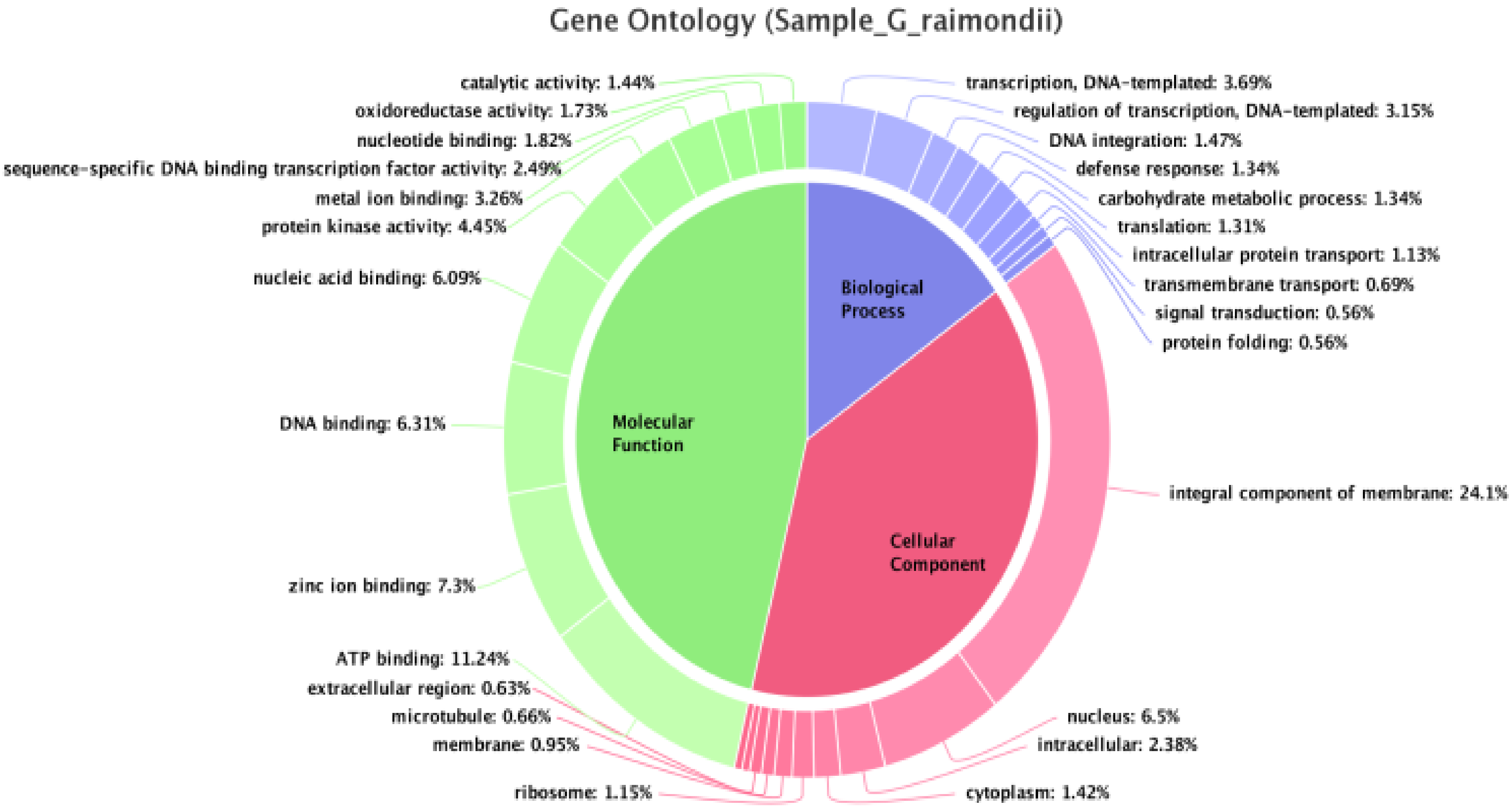
Differential Gene Expression Analysis

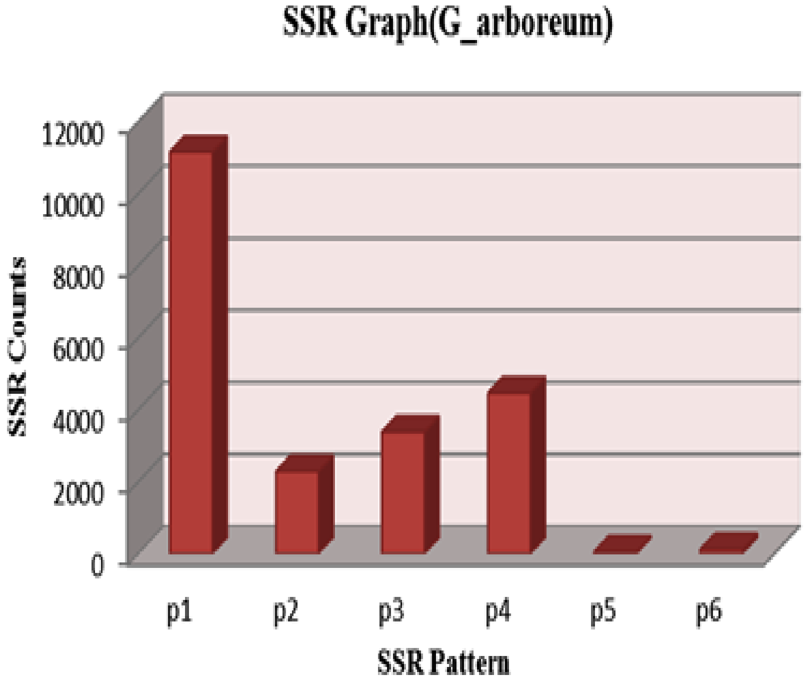
Control vs Treated

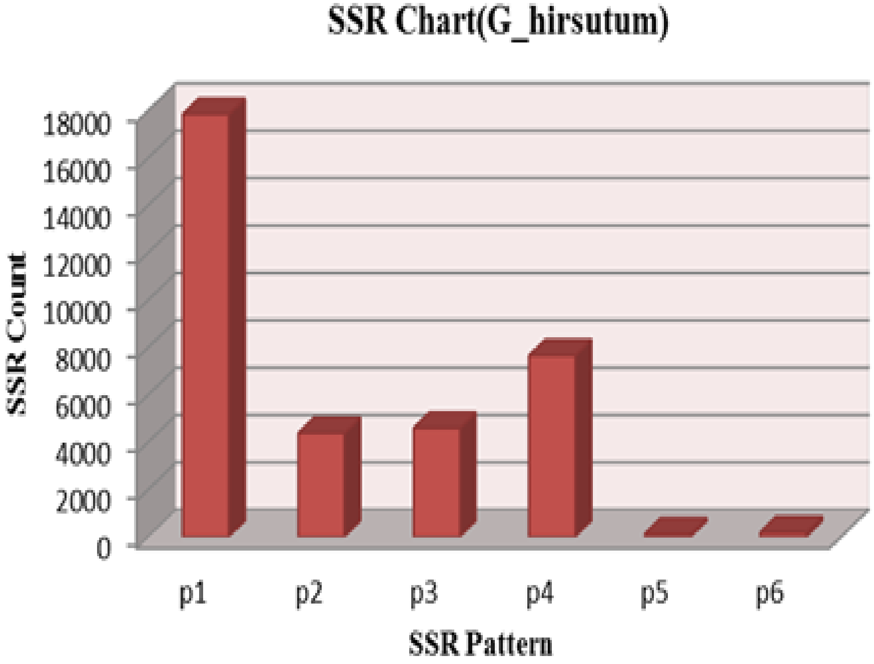
Gene ontology

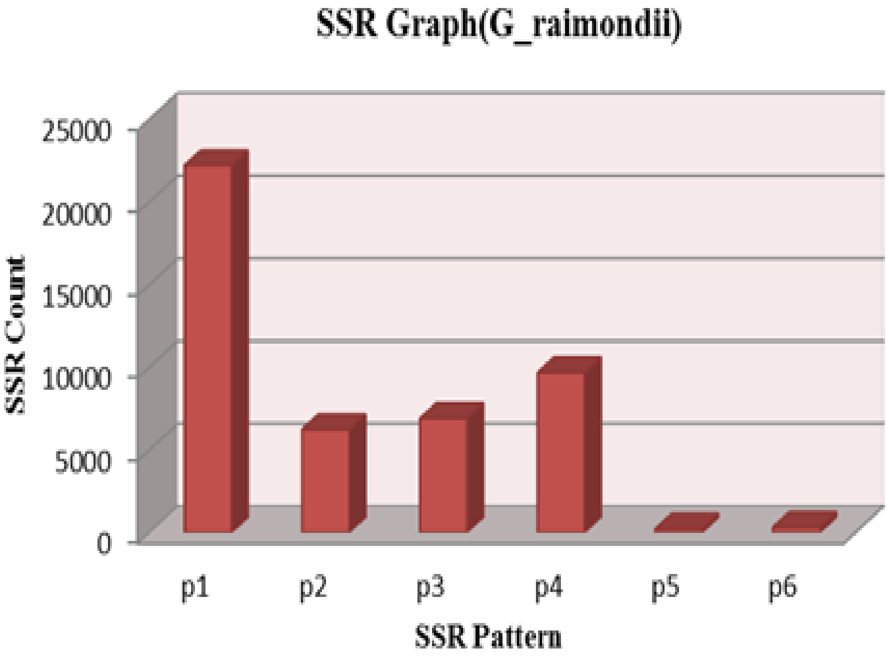



Gene Ontology (Sample_G_hirsutum)
structural constituent of ribosome: 1.81%
oxidoreductase activity: 1.82%
nucleotide binding: 1.94%
sequence-specific DNA binding transcription factor activity: 2.22%
metal ion binding: 3.55%
protein kinase activity: 4.1%
nucleic acid binding: 5.22%
DNA binding: 5.79%
zinc ion binding: 6.45%
ATP binding: 11.68%
extracellular region: 0.53%
kinesin complex: 0.62%
microtubule: 0.81%
membrane: 1.12%
ribosome: 1.61%
transcription, DNA-templated: 3.23%
regulation of transcription, DNA-templated: 2.99%
translation: 1.77%
carbohydrate metabolic process: 1.49%
intracellular protein transport: 1.32%
transmembrane transport: 0.89%
defense response: 0.78%
protein folding: 0.74%
microtubule-based movement: 0.63%
small GTPase mediated signal transduction: 0.63%
integral component of membrane: 24.91%
nucleus: 6.57%
intracellular: 2.47%
cytoplasm: 1.83%
Biological Process
Molecular Function
Cellular Component


Gene Ontology (Sample_G_raimondii)
Molecular Function
Biological Process
Cellular Component
catalytic activity: 1.44%
oxidoreductase activity: 1.73%
nucleotide binding: 1.82%
sequence-specific DNA binding transcription factor activity: 2.49%
metal ion binding: 3.26%
protein kinase activity: 4.45%
nucleic acid binding: 6.09%
DNA binding: 6.31%
zinc ion binding: 7.3%
ATP binding: 11.24%
transcription, DNA-templated: 3.69%
regulation of transcription, DNA-templated: 3.15%
DNA integration: 1.47%
defense response: 1.34%
carbohydrate metabolic process: 1.34%
translation: 1.31%
intracellular protein transport: 1.13%
transmembrane transport: 0.69%
signal transduction: 0.56%
protein folding: 0.56%
integral component of membrane: 24.1%
nucleus: 6.5%
intracellular: 2.38%
cytoplasm: 1.42%
ribosome: 1.15%
membrane: 0.95%
microtubule: 0.66%
extracellular region: 0.63%


SSR Graph(G_arboreum)


SSR Chart(G_hirsutum)


SSR Graph(G_raimondii)

P1 – Mononucleotide repeats
P2 – Dinucleotide repeats
P3 – Trinucleotide repeats
P4 – Tetranucleotide repeats
P5 – Pentanucleotide repeats
P6 – Hexanucleotide repeats

VARIANT ANLYSIS IN HBS144 AND HBS 187

SNP's	**HBS 144**	**HBS 187**
G. arboreum	**343855**	**307180**
G. hirsutum	**56287**	**46078**
G. raimondii	**352907**	**280861**

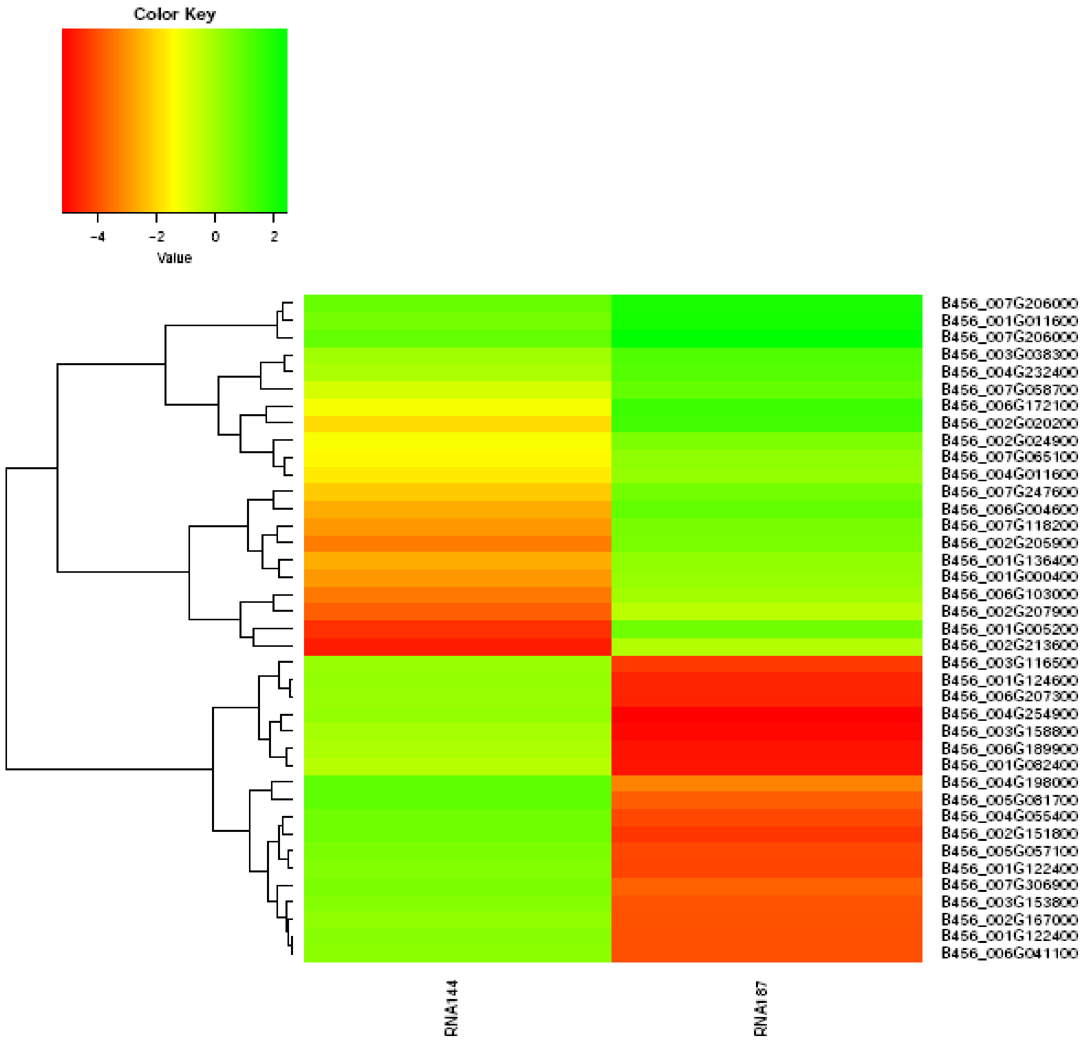
INDELs	**HBS 144**	**HBS 187**
G. arboreum	**15459**	**14733**
G. hirsutum	**2884**	**2400**
G. raimondii	**15323**	**11396**

DIFFERENTIAL GENE EXPRESSION

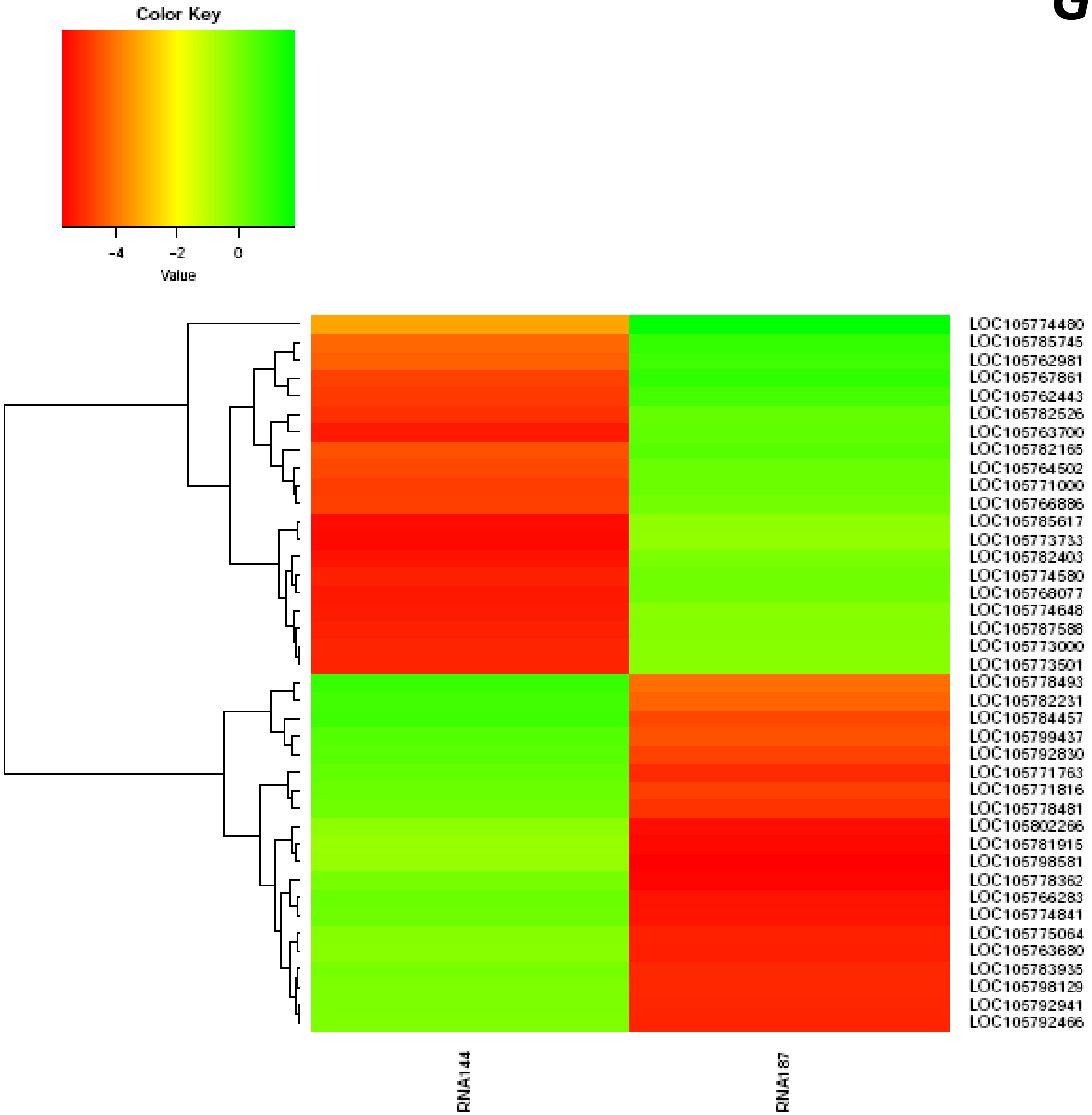
	G. arboreum	G. hirsutum	G. raimondii
Up Regulated:	4499	24885	11299
Down Regulated:	17979	29377	12379
Neutrally Regulated:	50765	80926	29708

Heat map

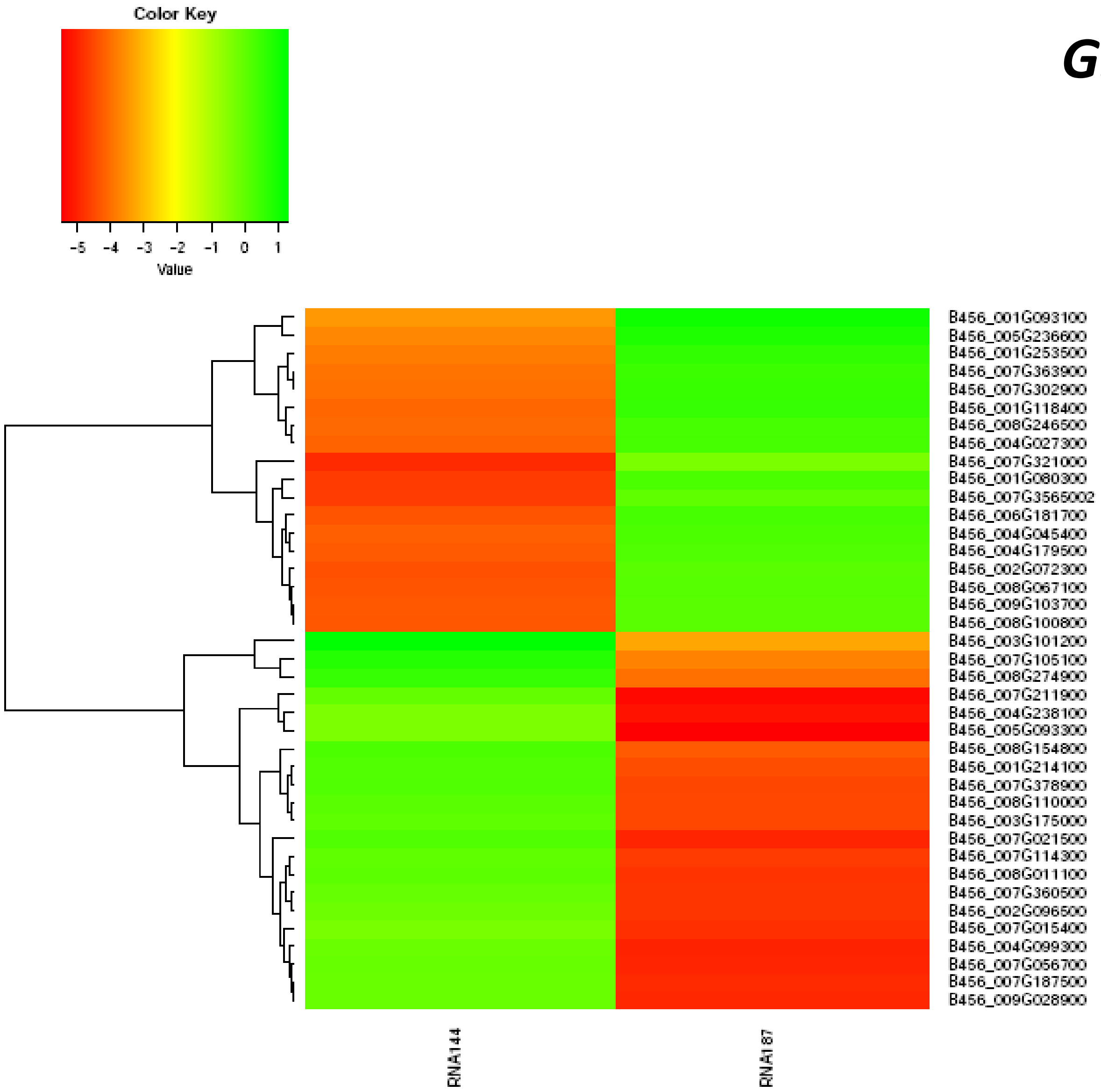
G. arboreum



G. ramondii


Color Key
-4
-2
0
Value
LOC105774480
LOC105785745
LOC105762981
LOC105767861
LOC105762443
LOC105782526
LOC105763700
LOC105782165
LOC105764502
LOC105771000
LOC105766886
LOC105785617
LOC105773733
LOC105782403
LOC105774580
LOC105768077
LOC105774648
LOC105787588
LOC105773000
LOC105773501
LOC105778493
LOC105782231
LOC105784457
LOC105799437
LOC105792830
LOC105771763
LOC105771816
LOC105778481
LOC105802266
LOC105781915
LOC105798581
LOC105778362
LOC105766283
LOC105774841
LOC105775064
LOC105763680
LOC105783935
LOC105798129
LOC105792941
LOC105792466
RNA144
RNA187

G. hirsutum


Color Key
-5
-4
-3
-2
-1
0
1
Value
B456_001G093100
B456_005G236600
B456_001G253500
B456_007G363900
B456_007G302900
B456_001G118400
B456_008G246500
B456_004G027300
B456_007G321000
B456_001G080300
B456_007G3565002
B456_006G181700
B456_004G045400
B456_004G179500
B456_002G072300
B456_008G067100
B456_009G103700
B456_008G100800
B456_003G101200
B456_007G105100
B456_008G274900
B456_007G211900
B456_004G238100
B456_005G093300
B456_008G154800
B456_001G214100
B456_007G378900
B456_008G110000
B456_003G175000
B456_007G021500
B456_007G114300
B456_008G011100
B456_007G360500
B456_002G096500
B456_007G015400
B456_004G099300
B456_007G056700
B456_007G187500
B456_009G028900
RNA144
RNA187

WAY FORWARD

- **Construction of genetic linkage map in F_2 population for fiber strength using polymorphic markers**

  &

  Analysis of QTLs and detection of their association with molecular markers.

Populations developed

- F_2 population of cross

  RHAP 12 X RHAP 24

  RHAP 7 X RHAP 12

- RIL population of cross

  RHAP 15 X RHAP 12

Characters of parents

- **RHAP 12**
- Thin leaf
- Early maturity
- Less pubescent
- Susceptible to sucking pest
- High fibre strength and short fibres
- **RHAP 15**
- Thick leaf
- Late maturity
- High pubescence
- Resistant to sucking pest
- Stay green
- Low fibre strength short

- **RHAP 24**
- Thin leaf
- Medium maturity
- Less pubescent
- Susceptible to sucking pest
- Brittle branch
- High fibre strength and long fibres
- **RHAP 7**
- Thick leaf
- Late maturity
- Resistant to sucking pest
- Low fibre strength short fibres

RHAP 15

RHAP 24

RHAP 7

RHAP 12

To study gene expression level of selected genes for fibre strength in selected lines of *Gossypium herbaceum* L. cotton genotypes

***G. herb 2,* 23.75 g/tex**

***G. herb 100,* 17.58 g/tex**

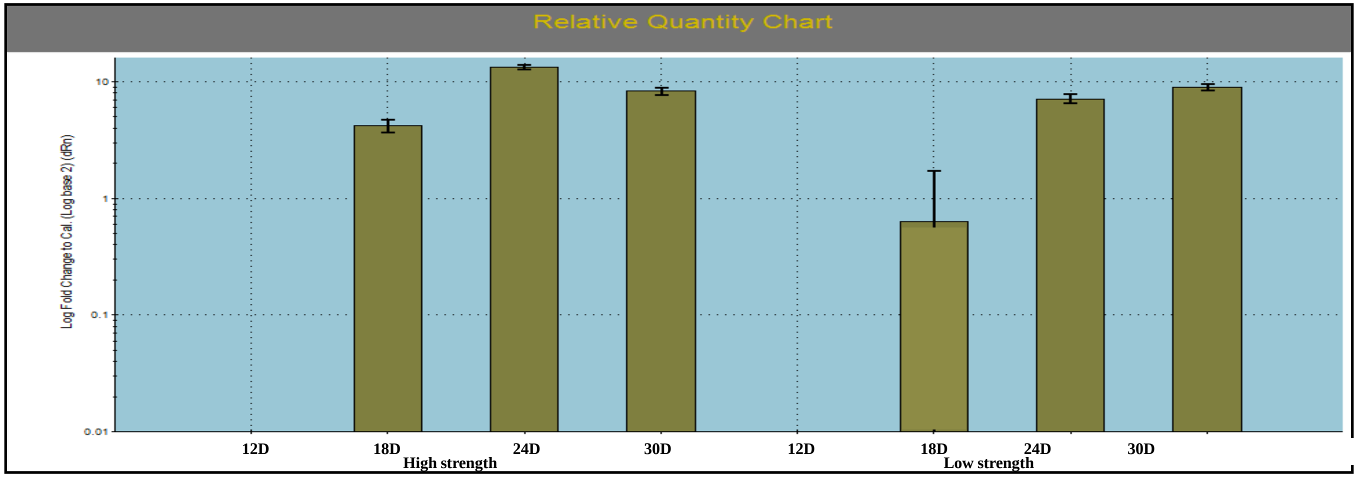


Fig. 5: Relative fold change in expression of *GhMT1* gene in comparison with calibrator value as zero

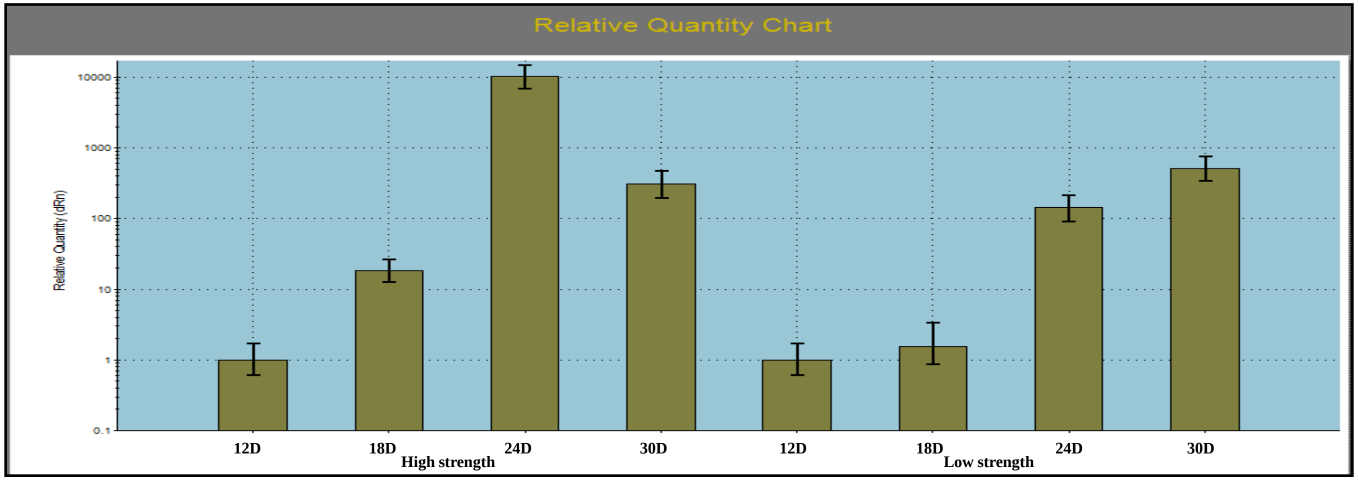


Fig. 6: Relative fold change in expression of *GhMT1* gene in comparison with calibrator value as one

Role of target genes in different stages of cotton fiber development in herbaceum lines

Target gene	Cotton fiber development stage
CesA8	**Secondary cell wall synthesis**
GhMT1	**Secondary cell wall synthesis**
Gene32671	**Maturity stage**
Gene43908	**Secondary cell wall synthesis and Maturity stage**

TO VALIDATE THE AD GENOME SPECIFIC GENE/S IMPLICATED IN FIBRE STRENGTH IN *GOSSYPIUM ARBOREUM* LINES.

QTL-6

G
arb-120

Sl. No.	Character	QTL.6	*G.arb* -120
1	Fiber strength (g/tex)	34.37	8.50
2	Mean length (mm)	27.9	24.1
3	Micronaire index (µg/inch)	3.3	4.5
4	Uniformity ratio (%)	84.5	78.5

UP REGULATION OF P SIGNIFICANT GENES AT VARIOUS STAGES OF FIBER DEVELOPMENT IN G. ARBOREUM L. GENOTYPES(GENE EXPRESSION STAGES)

Sl. No	EST,s/Gene	High fiber strength genotype (QTL-6)	Low fiber strength genotype(G arb-120)
1	CesA 8	Secondary cell wall synthesis Fiber maturation	Fiber elongation Fiber maturation
2	Gene 1272	Secondary cell wall synthesis Fiber maturation	Nil
3	Gene 47473	Fiber elongation Secondary cell wall synthesis	Nil
4	Gene 32671	Fiber maturation Fiber elongation	Fiber maturation
5	Gene 43908	Fiber maturation	Fiber elongation Secondary cell wall synthesis Fiber maturation
6	GhMT1	Fiber maturation	Fiber elongation

Two F_2 populations of used for current study and their parents characteristics

1. **QTP 5 x *G.herbaceum* 90**
2. **QTP 16 x *G. arboreum* 151**

QTP 5 x *G.herbaceum* 90

	QTP 5	G.herbaceum 90
Fiber strength g/tex	31.17	18.20

QTP 16 x *G. arboreum* 151

	QTP 16	G.arboreum 151
Fiber strength g/tex	23.08	33.13

***G. herbaceum* - A1**
***G. arboreum* – A2**

FIBER QUALITY PARAMETERS OF SELECTED HIGH AND LOW FIBER STRENGTH COTTON GENOTYPES IN F_2 (QTP 5 X *G.HERBACEUM* 90) FOR GENE EXPRESSION STUDY

Sl. No.	Character	High fiber strength genotype	Low fiber strength genotype
1	Fiber strength (g/tex)	28.2	22.80
2	Mean length (mm)	27.40	22.80
3	Micronaire index (µg/inch)	5.05	5.24
4	Uniformity ratio (%)	85.20	

Role of target genes in different stages of cotton fiber development

S.no	Target gene	Cotton fiber development stage
1.	Gene 1272	Secondary cell wall synthesis
2.	Gene 32671	Secondary cell wall synthesis
3.	Gene 55808	Fiber Maturity
4.	Gene 59335	Fiber Maturity
5.	Gene 42183	Fiber elongation
6.	Gene 24332	Differential expression

Acknowledgement

- Dr I.S.Katageri, UAS Dharwad
- Dr K.P.Raghvendra,CICR,Nagpur
- Dr G. Balasubramani,CICR,Nagpur
- Dr J. Amudha

Thank you…