Cloning and expression of draught- & salt-tolerant genes on cotton

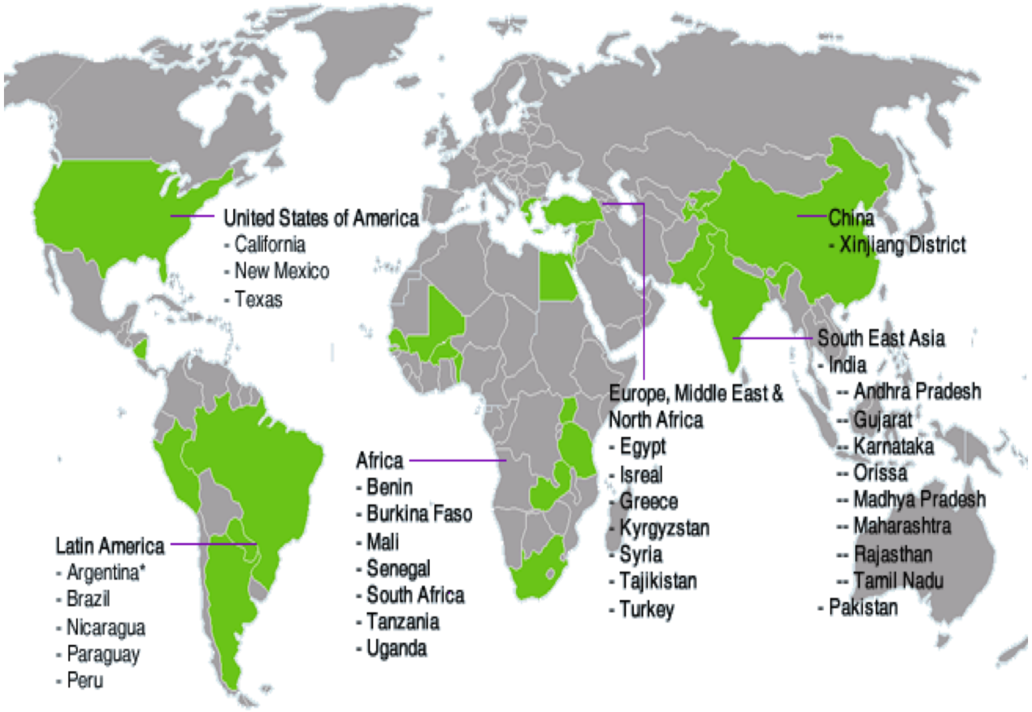
Wuwei Ye, PHD

Yew158@163.com

State Key Laboratory of Cotton Biology,
China Cotton Research Institute, CAAS,
Anyang, China

- **Cotton is one of the most important economic crops**

75 cotton-producing countries located between 32° south and 47° north latitude on the globe, about 33 million ha or 5% of the world’s arable land is used for cotton planting annually.

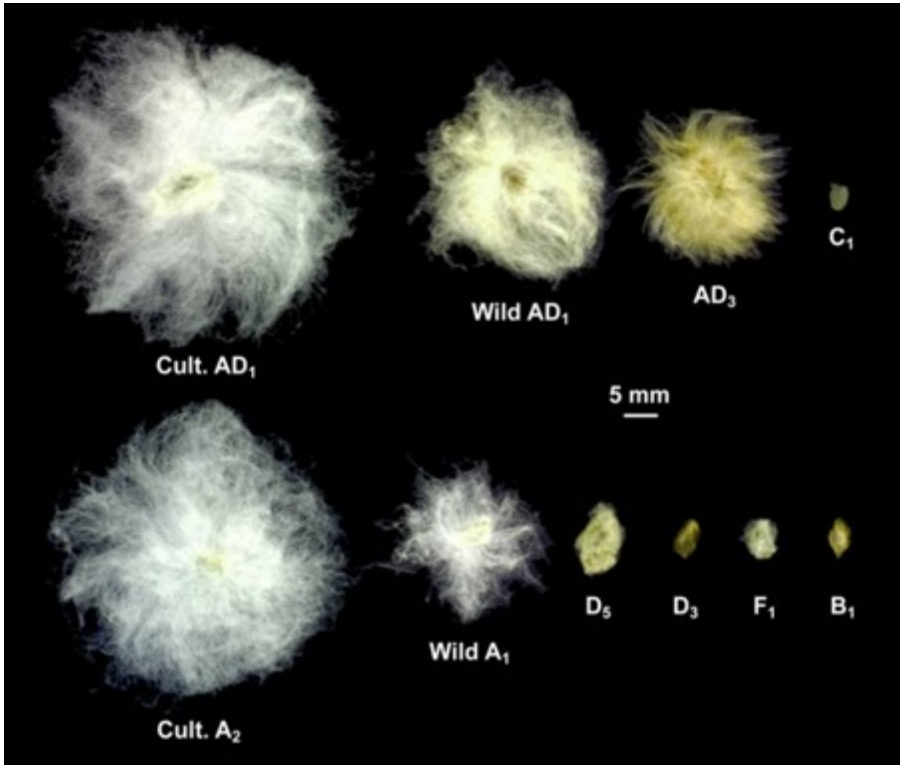


From http://farmhub.textileexchange.org/learning-zone/growing-regions

- **Cotton is an excellent model system for studying salinity tolerance and draught tolerance.**

Variation in seed trichome (fiber) morphology in wild and domesticated cottons.

Gossypium seeds exhibits remarkable variation among the ~50 wild and domesticated species.

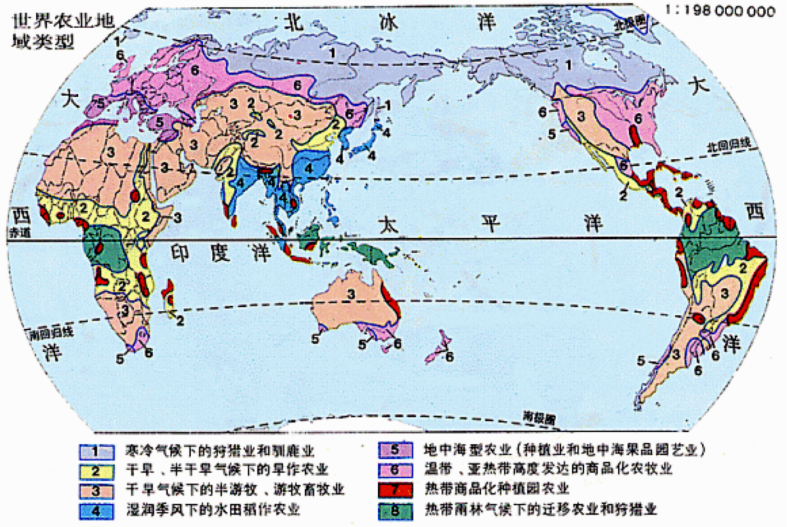


http://cottonevolution.info/

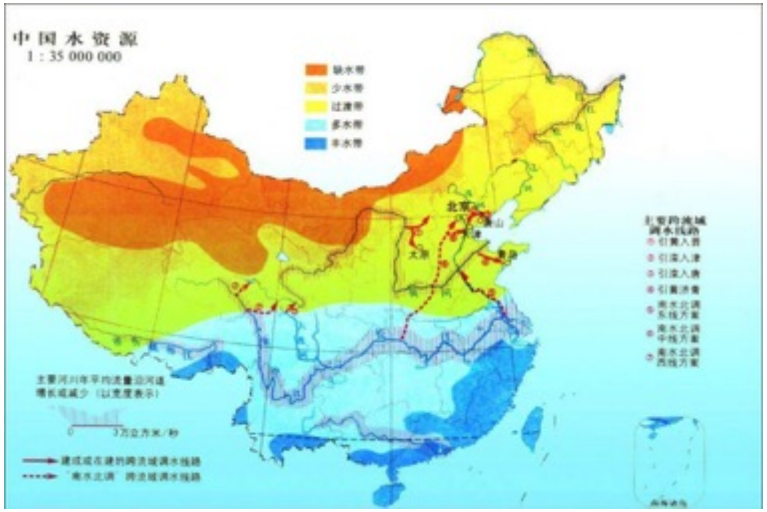
Outline

1. Background
2. Germplasm identification
3. Mechanism of salinity tolerance
4. Salinity-tolerant genes isolated

Drought and semi-drought regions



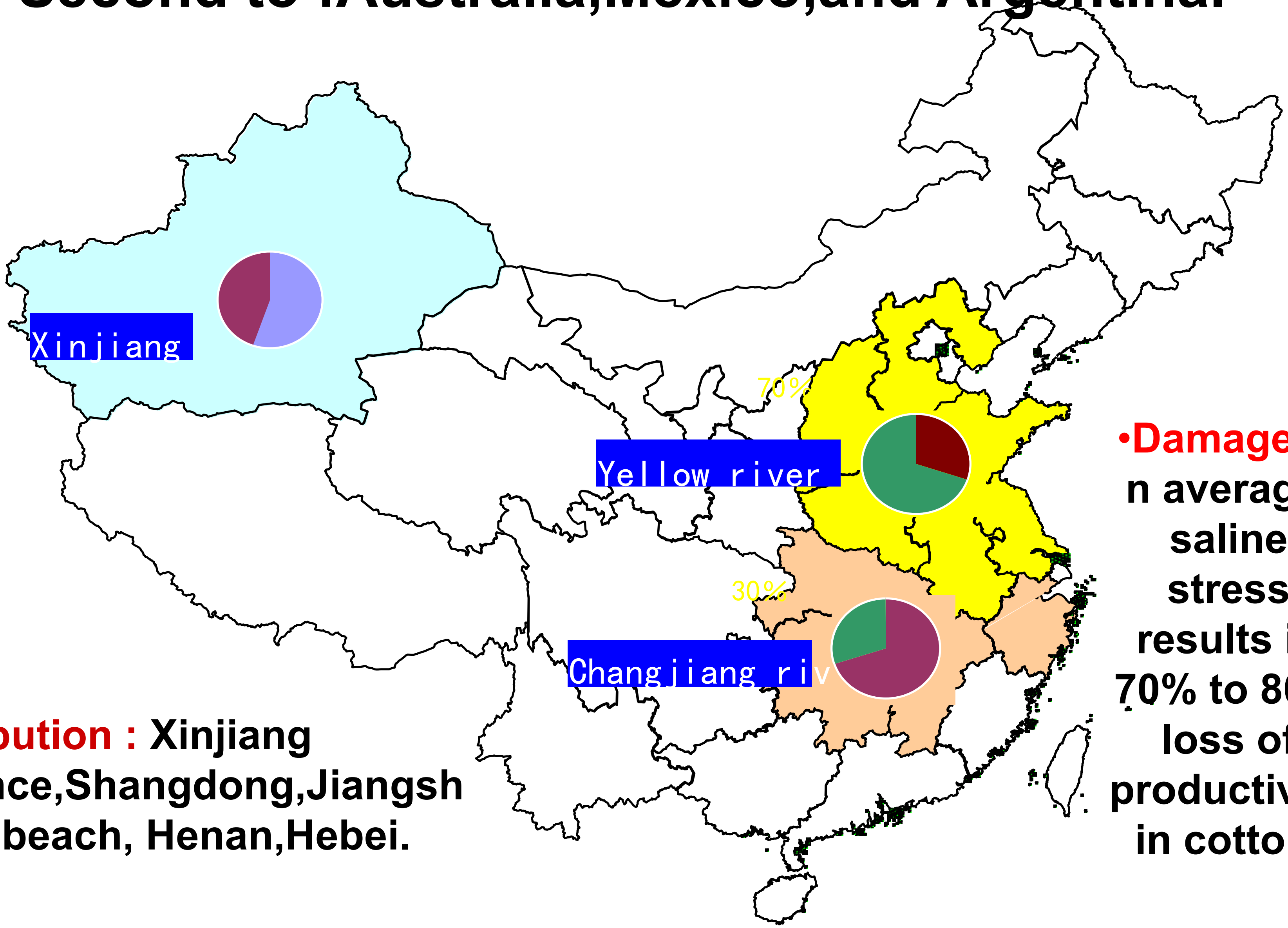
Drought areas in China



Silk Route

Saline areas in China

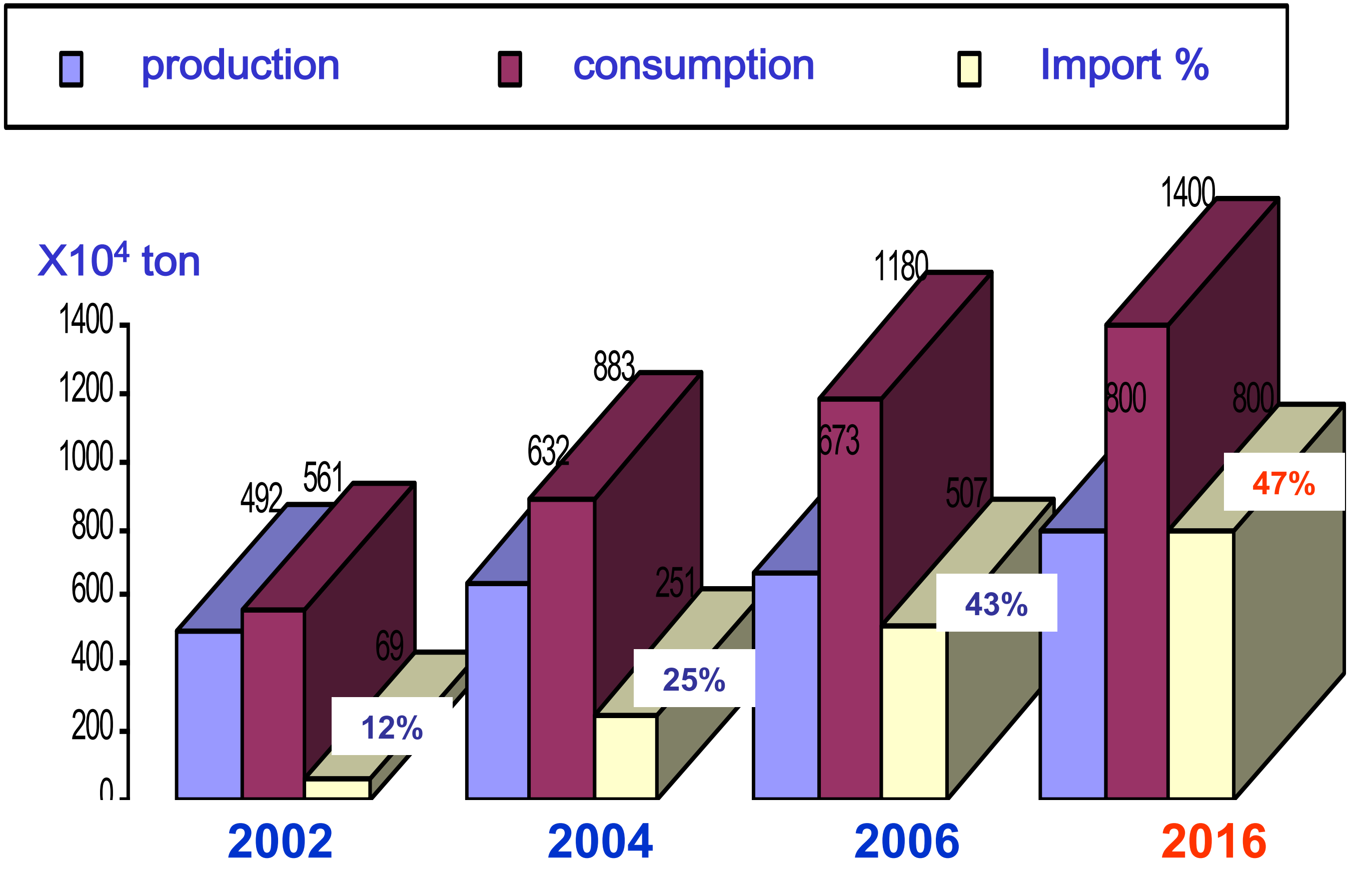
China: 3.5×10^{7} ha saline land of the total land, Second to :Australia,Mexico,and Argentina.



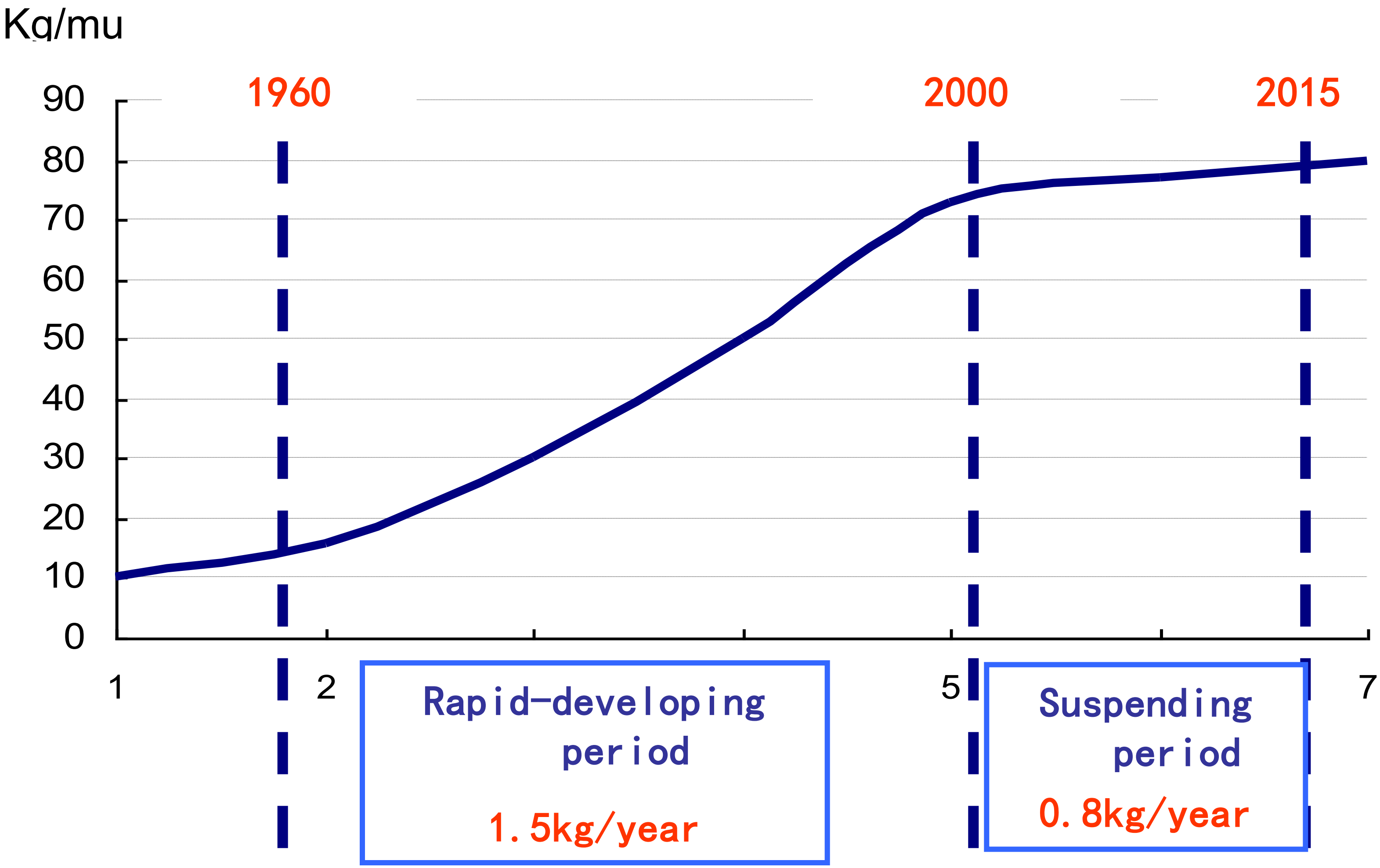
- **Damage**:On average, saline stress results in 70% to 80% loss of productivity in cotton.

Distribution : Xinjiang province,Shangdong,Jiangshu,Sea beach, Henan,Hebei.

Cotton production and consumption in China



Annual cotton yield in China



Rapid-developing period

1.5kg/year

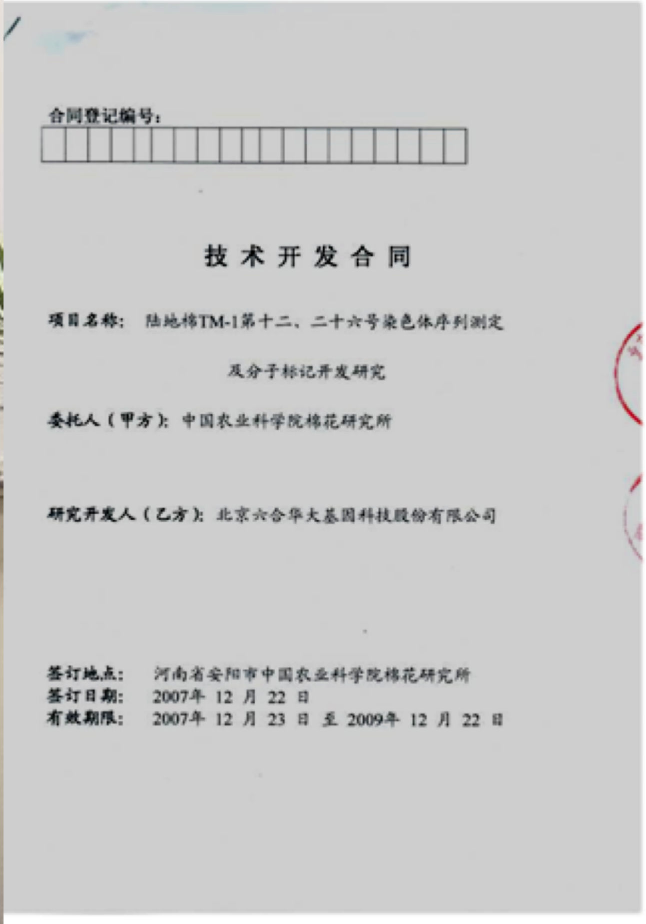
Suspending period

0.8kg/year

Cotton Genome Projecct

Sino-US, In 2007

华大基因
BGI
SEQUENCING



合同登记编号：□□□□□□□□□□□□□□□□□□□

技术开发合同

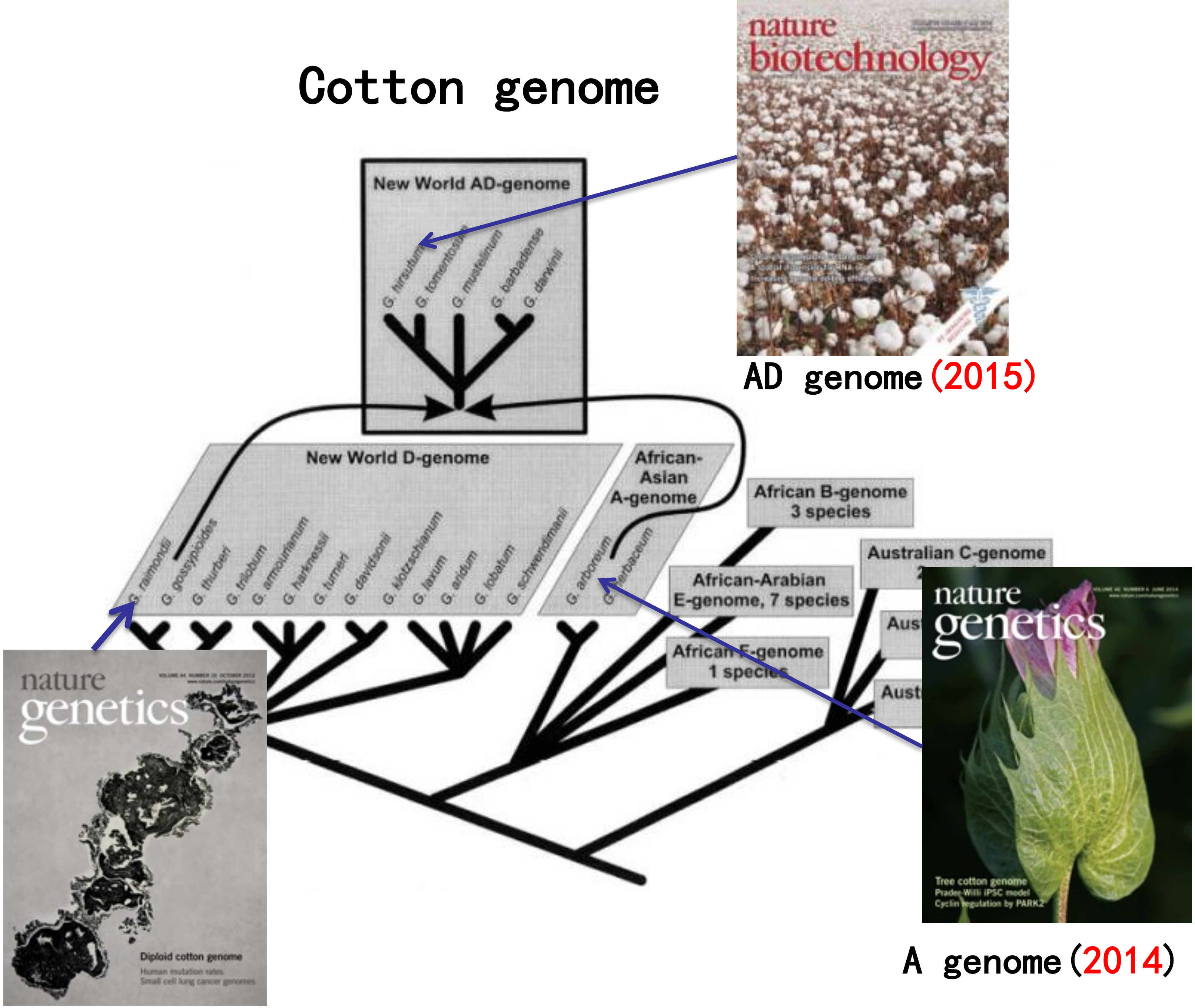
项目名称：　陆地棉TM-1第十二、二十六号染色体序列测定
及分子标记开发研究

委托人（甲方）：中国农业科学院棉花研究所

研究开发人（乙方）：北京六合华大基因科技股份有限公司

签订地点：　河南省安阳市中国农业科学院棉花研究所
签订日期：　2007年 12 月 22 日
有效期限：　2007年 12 月 23 日 至 2009年 12 月 22 日

Cotton genome

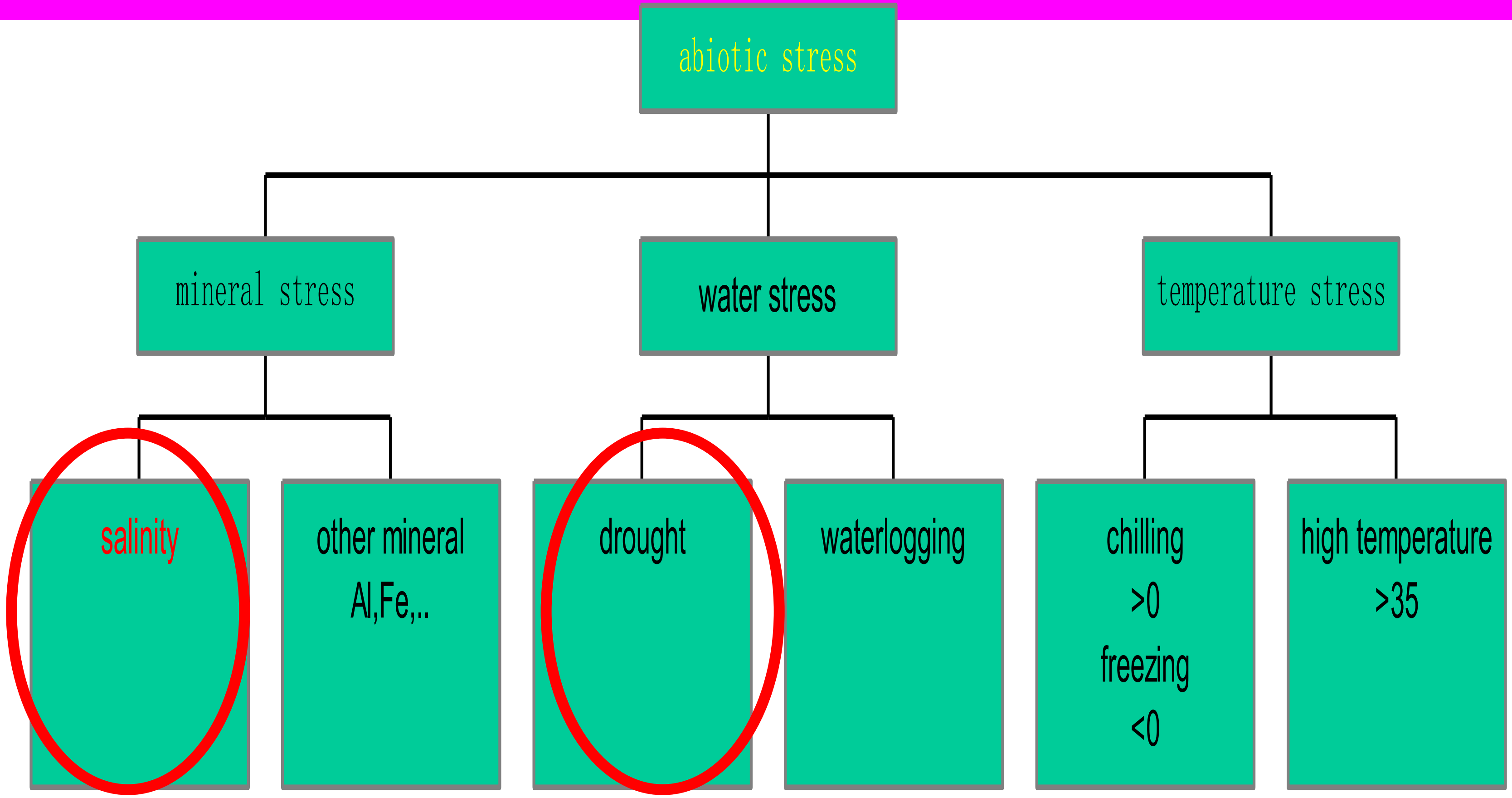


AD genome (2015)

A genome (2014)

D genome (2012)

Abiotic stress


abiotic stress

mineral stress

water stress

temperature stress

salinity

other mineral

Al,Fe,..

drought

waterlogging

chilling

>0

freezing

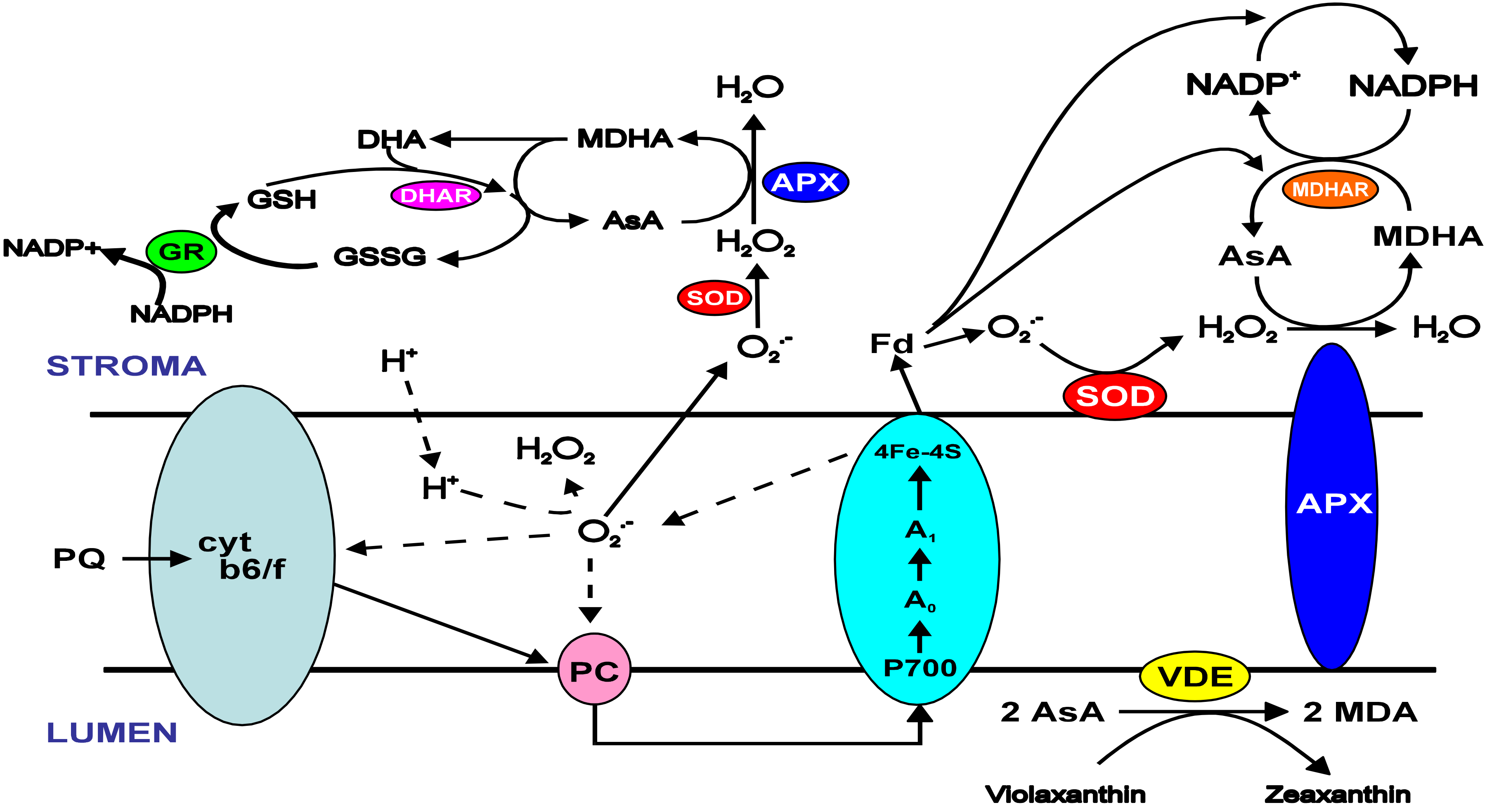
<0

high temperature

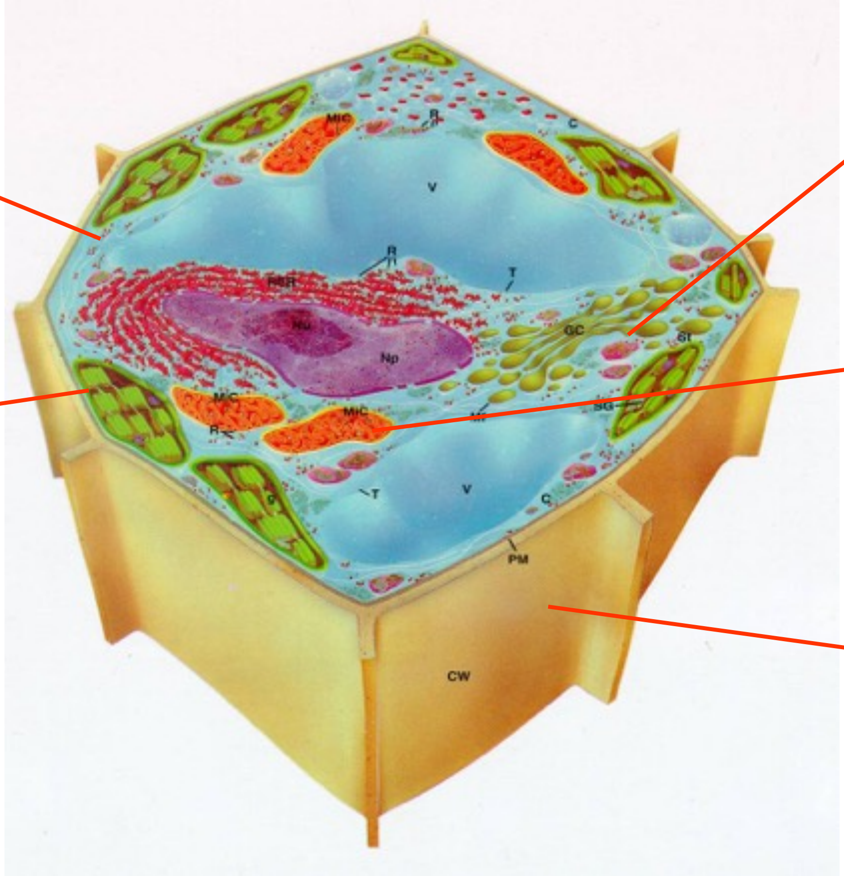
>35

- **Damage**: On average, saline or draught stress results in 30% to 40% loss of productivity in cotton.

Main genetic loci related to salinity stress



Salt-tolerance related enzymes



Cytosol

Cu/Zn SOD
APX
GPX
GST

Chloroplast

Cu/Zn SOD
Fe SOD
stromal APX
thylakoid APX

Peroxisome

Catalase
APX

Mitochondria

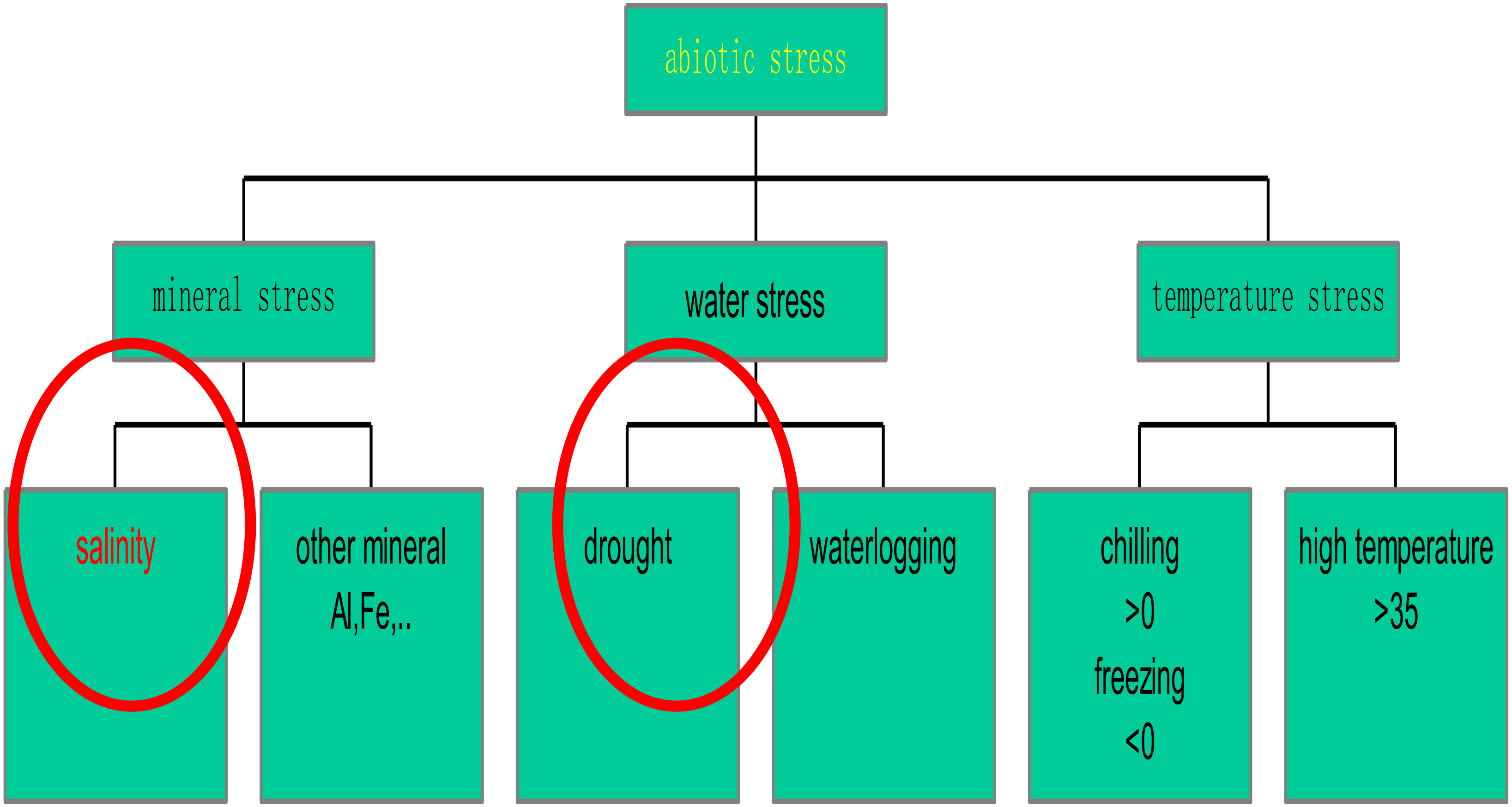
Mn SOD
Catalase
APX

Cell Wall

Cu/Zn SOD
Peroxidases
APX

Germplasm identification

Abiotic stress



abiotic stress

mineral stress

water stress

temperature stress

salinity

other mineral
Al,Fe,..

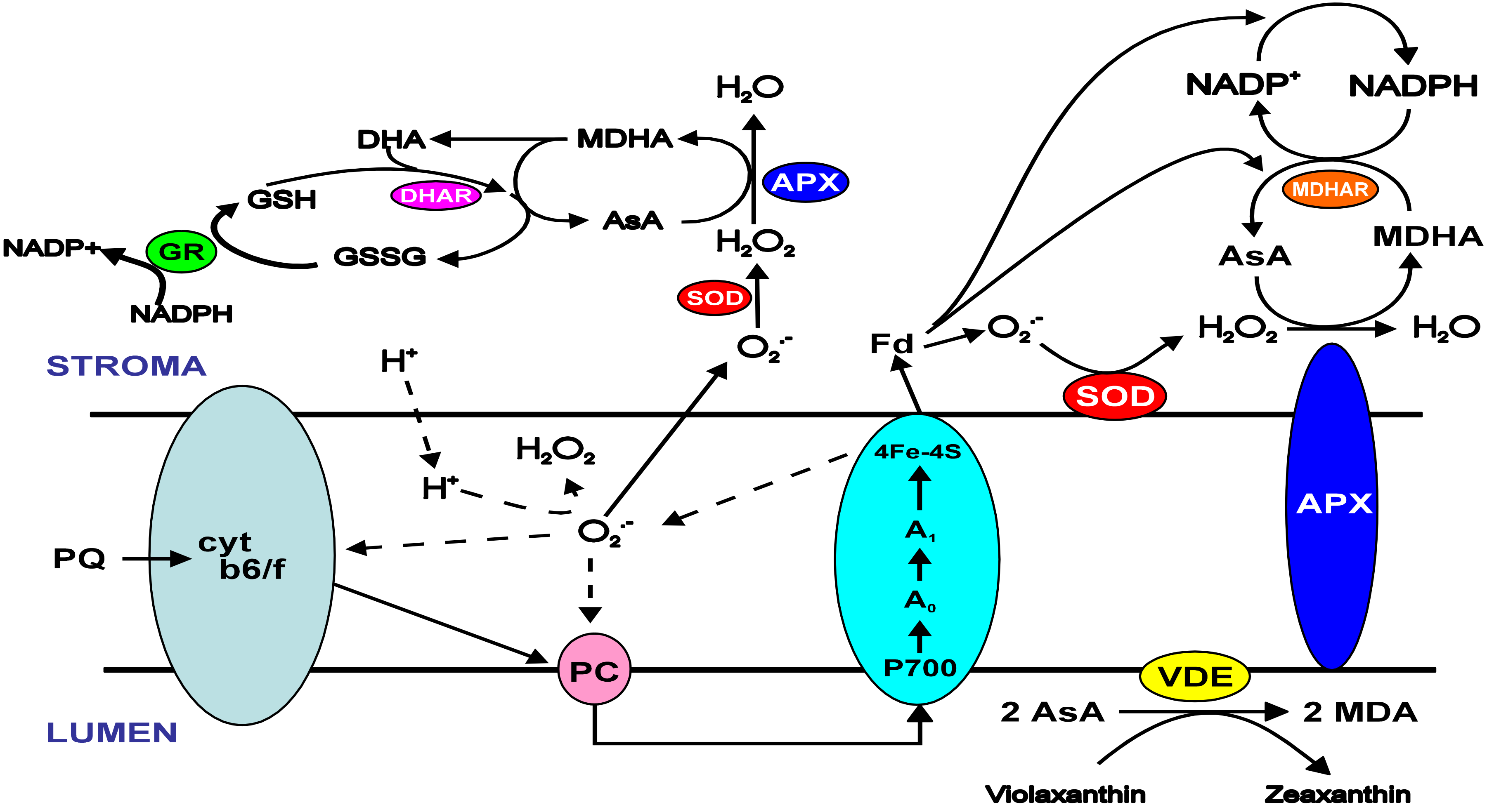
drought

waterlogging

chilling
>0
freezing
<0

high temperature
>35

Main genetic loci related to salt- & draught-stress



Cotton germplasm (US cv China)

Species	US	China
G.hirsutum L.	3400	7522
G.barbadense L.	1600	565
G.herbaceum L.	190	17
G.arboreum L.	1700	378
G.hirsutum L.*landrances*	2200	350
others	250	41
total	9340	8873

Screening systems for salt- /drought-stress

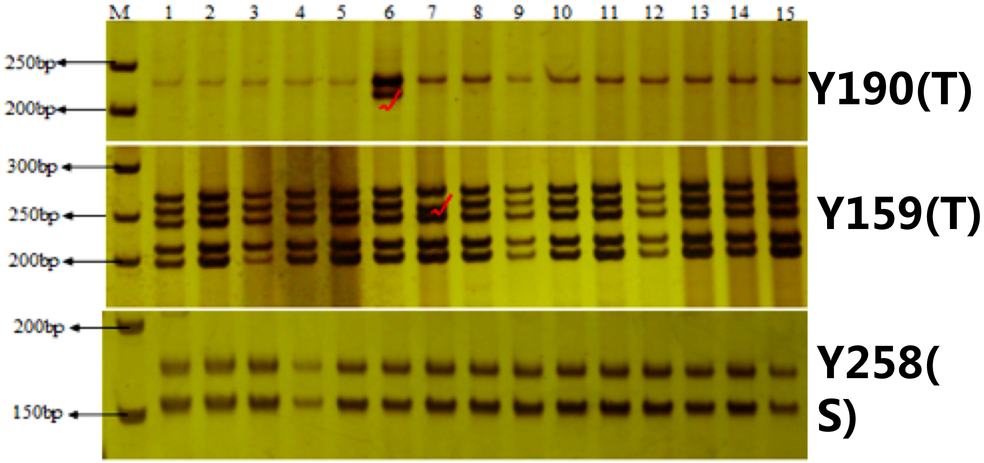
Cotton germplasm and salinity resistance in China

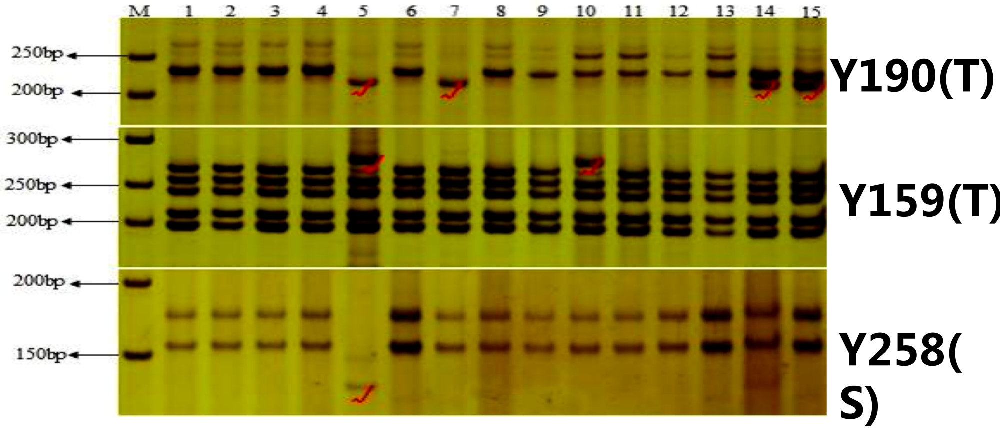
1. Germplasm: 8873.
2. Few of them are resistant to salinity, as well as to drought, chilling.
3. Distribution of salinity-resisted:

G.herbceum,	33.33％
G.barbadense,	3.86％
G.hirsutum,	0.09％
G.arboreum,	none

SSR markers to identify the salinity tolerance in cotton germplasm

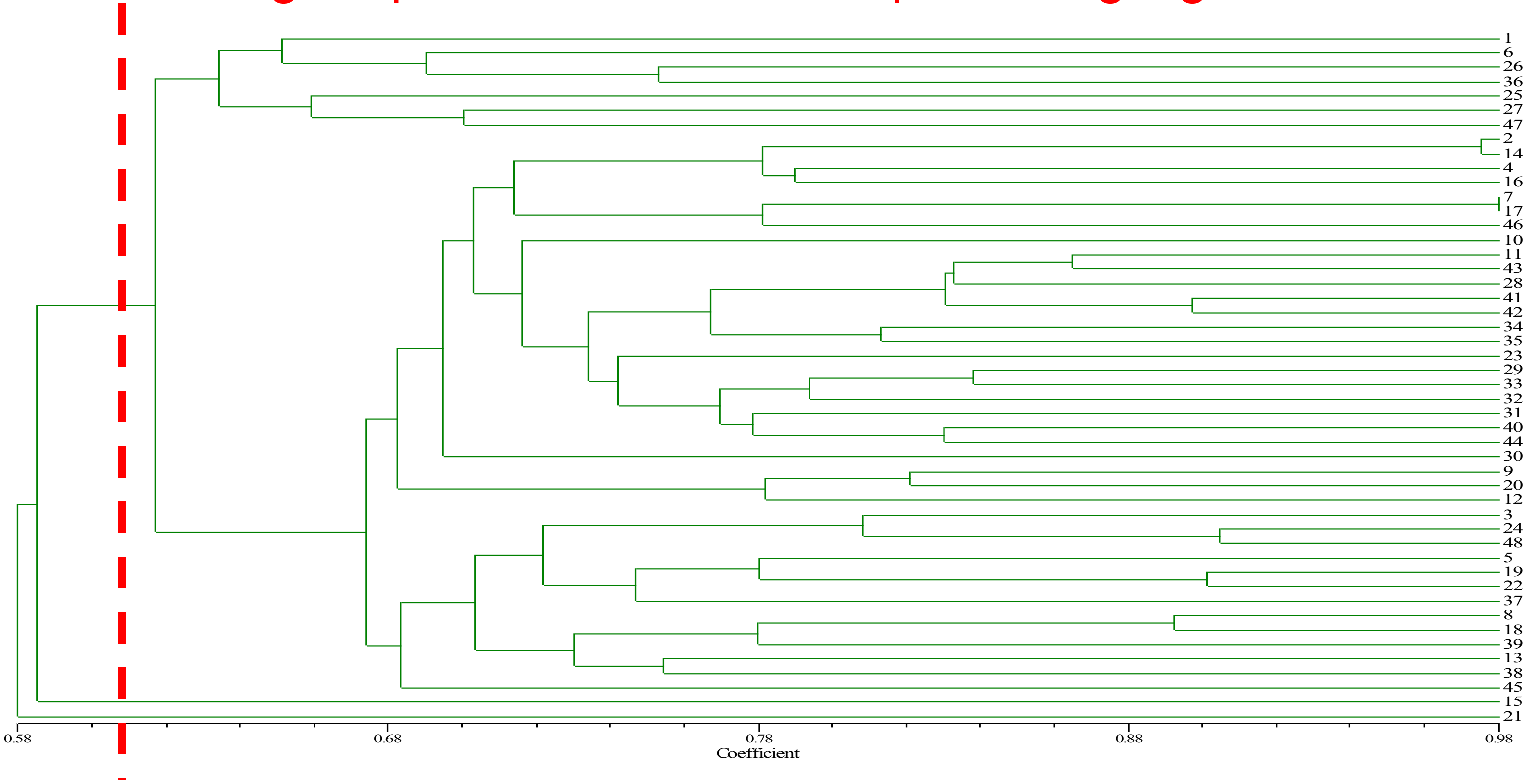
- 26 of 5053 pairs of SSR core primers were screened.
- A new SSR multi-markers salinity-identification method, was initially established to identify salinity-tolerance of cotton.
- The combination of 3 markers are Y190, Y159 and Y258.
- It showed the coincidence of 90.91% in consistence with the identification result of 0.4%NaCl identification method.





The genealogy of cotton salinity-tolerant varieties by SSR markers

- The genetic diversity of cotton germlasm was analyzed among 48 cotton salinity-tolerance relevant accessions using 95 SSR molecular markers, which showed that most of germplasm had closer genetic relationship.
- The genealogy of the 60 cotton salinity-tolerant varieties were clustered into four basic germplasms such as Deltapine, King,Uganda and Russia.



Mechanism of salinity-tolerance

1. DNA methylation analysis under different types of salt

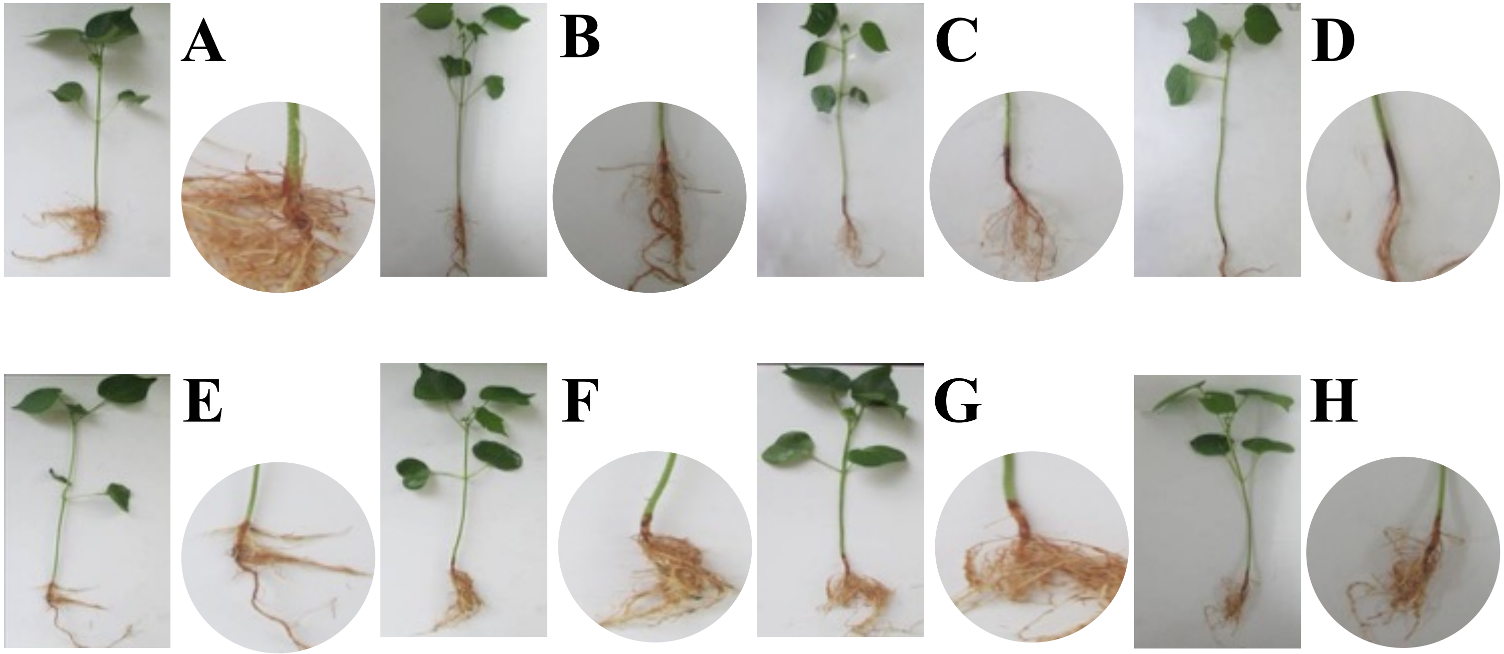


Fig. 2-1 Cotton seedlings under different salt stresses

A) Zhong07–dH_2O; **B)** Zhong07–NaCl; **C)** Zhong07–$NaHCO_3$; **D)** Zhong07–Na_2CO_3; **E)** ZhongS9612–dH_2O; **F)** ZhongS9612–NaCl; **G)** ZhongS9612–$NaHCO_3$; **H)** ZhongS9612–Na_2CO_3

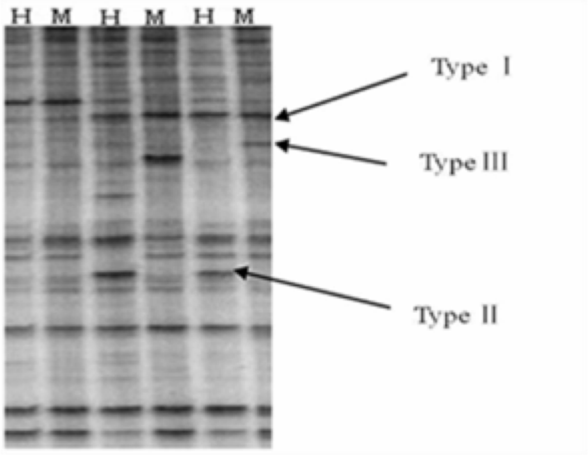


Fig. 2-2 MSAP amplification map of genome DNA of leaves of Zhong07 under NaCl stress

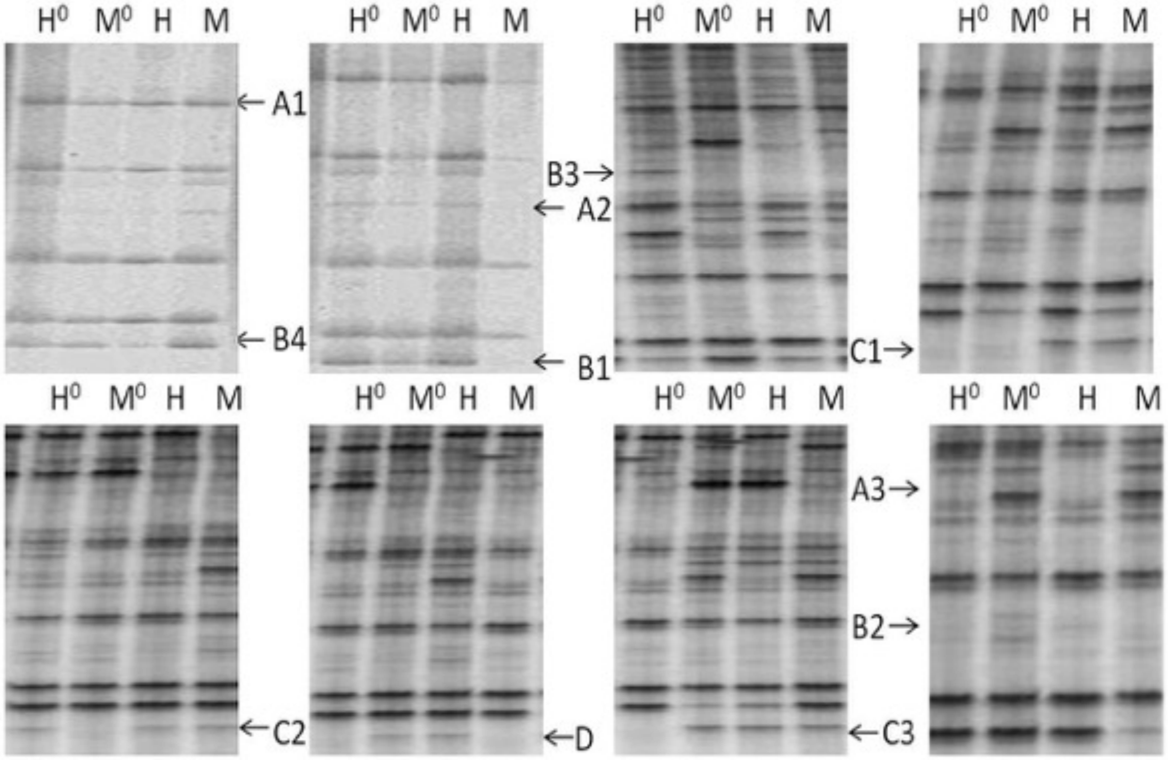


Fig. 2-3 MSAP amplification results between the salt stress and the control

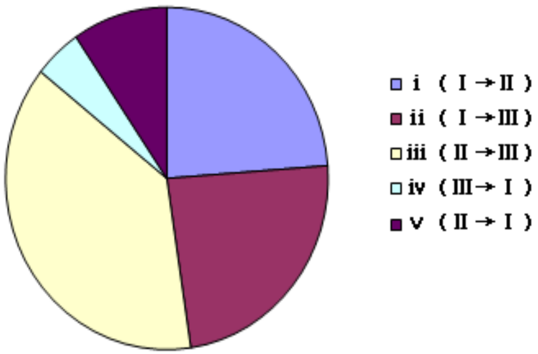
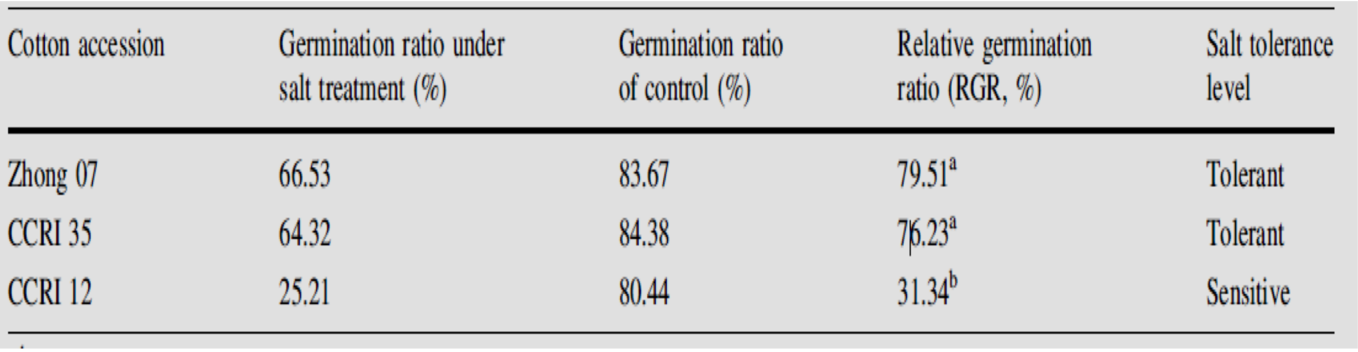


Fig. 2-4 The pattern of DNA methylation transformation and its percentages

- **Different salt stresses exerted distinct effects on cotton seedling growth.**
- **Neutral salt NaCl<alkalescent salt $NaHCO_3$ < alkaline salt Na_2CO_3**

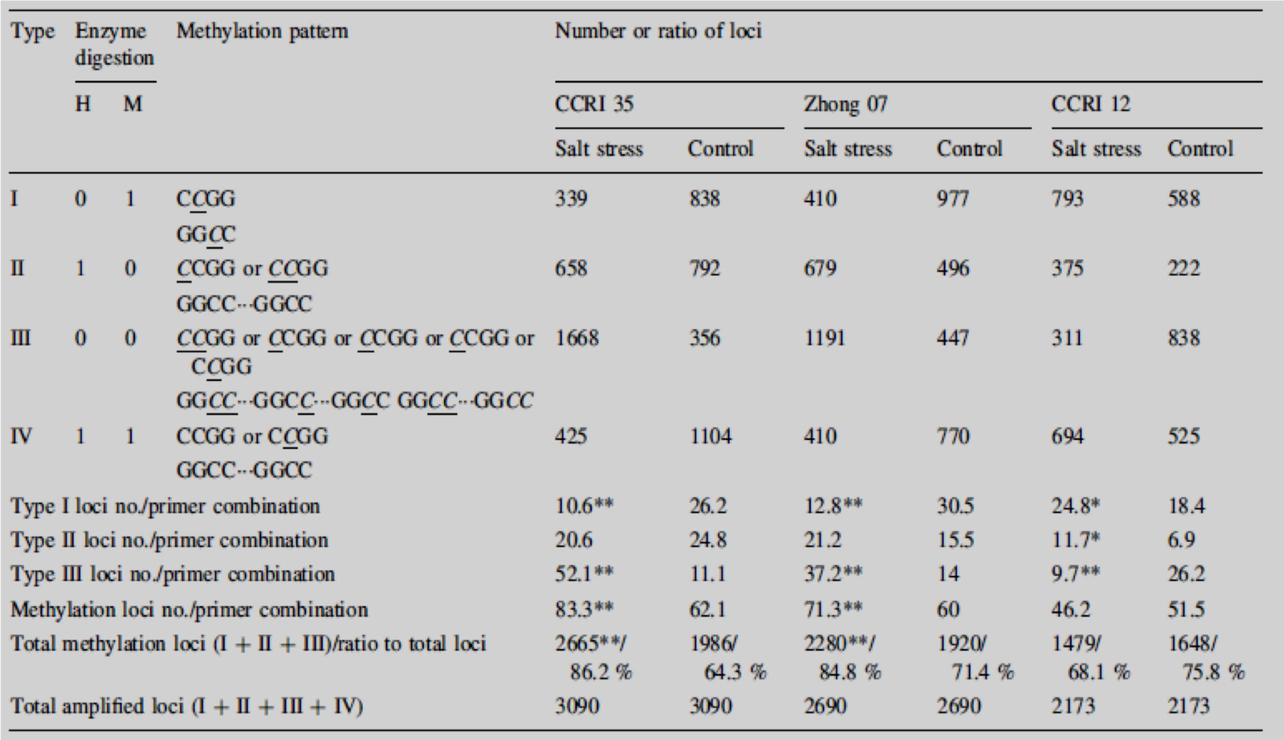
2. DNA methylation analysis in different cotton accessions under salt stress

Table 1 Salt-tolerance level of different cotton accessions



Cotton accession	Germination ratio under salt treatment (%)	Germination ratio of control (%)	Relative germination ratio (RGR, %)	Salt tolerance level
Zhong 07	66.53	83.67	79.51^{a}	Tolerant
CCRI 35	64.32	84.38	76.23^{a}	Tolerant
CCRI 12	25.21	80.44	31.34^{b}	Sensitive

Table 2 DNA methylation patterns of different cotton accessions detected by MSAP



Type	Enzyme digestion		Methylation pattern	Number or ratio of loci					
				CCRI 35		Zhong 07		CCRI 12	
	H	M		Salt stress	Control	Salt stress	Control	Salt stress	Control
I	0	1	C*C*GG GG*C*C	339	838	410	977	793	588
II	1	0	*C*CGG or *CC*GG GGCC···GGCC	658	792	679	496	375	222
III	0	0	*CC*GG or *C*CGG or *C*CGG or *C*CGG or C*C*GG GG*CC*···GGC*C*···GG*C*C GG*CC*···GG*CC*	1668	356	1191	447	311	838
IV	1	1	CCGG or C*C*GG GGCC···GGCC	425	1104	410	770	694	525
Type I loci no./primer combination				10.6**	26.2	12.8**	30.5	24.8*	18.4
Type II loci no./primer combination				20.6	24.8	21.2	15.5	11.7*	6.9
Type III loci no./primer combination				52.1**	11.1	37.2**	14	9.7**	26.2
Methylation loci no./primer combination				83.3**	62.1	71.3**	60	46.2	51.5
Total methylation loci (I + II + III)/ratio to total loci				2665**/ 86.2 %	1986/ 64.3 %	2280**/ 84.8 %	1920/ 71.4 %	1479/ 68.1 %	1648/ 75.8 %
Total amplified loci (I + II + III + IV)				3090	3090	2690	2690	2173	2173

- **The higher DNA methylation diversity (CHG methylation being more diverse than CG methylation) in cotton genotypes suggest epigenetic regulation may be important for cotton.**
- **The change in DNA methylation between genetypes indicates epigenetically developed in various *spp*.**

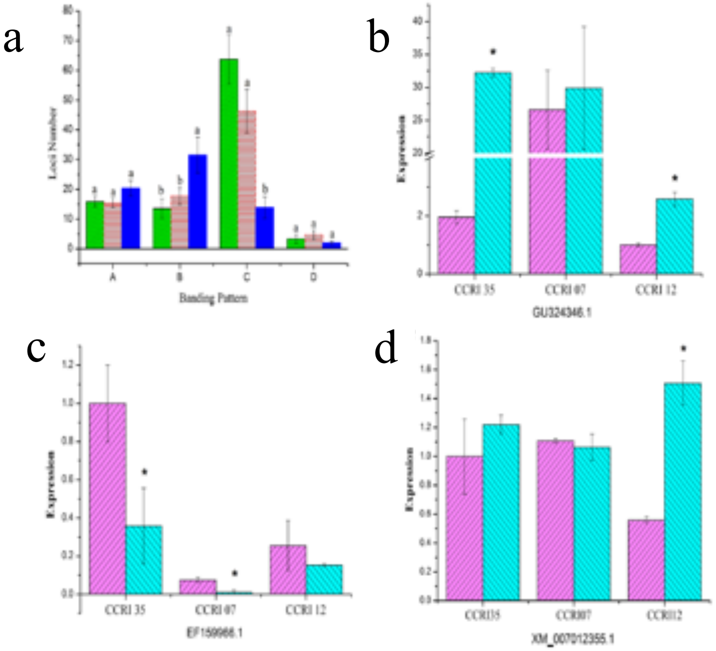
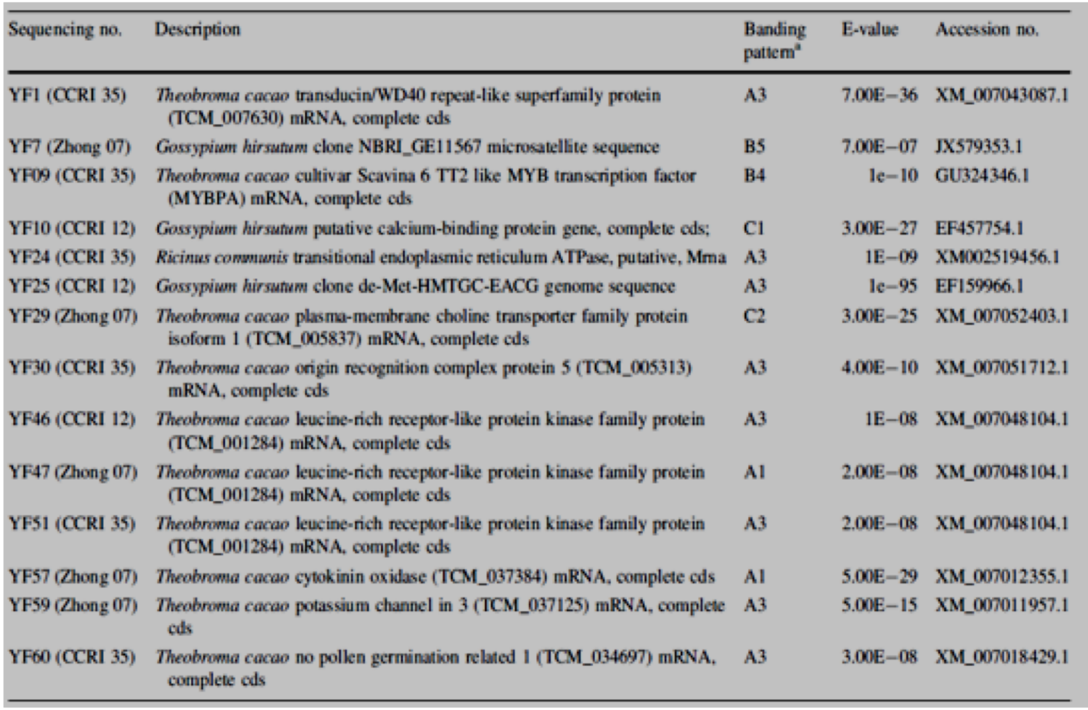


Fig. 2-5 Methylation patterns induced by salt



Sequencing no.	Description	Banding pattern[a]	E-value	Accession no.
YF1 (CCRI 35)	*Theobroma cacao* transducin/WD40 repeat-like superfamily protein (TCM_007630) mRNA, complete cds	A3	7.00E−36	XM_007043087.1
YF7 (Zhong 07)	*Gossypium hirsutum* clone NBRI_GE11567 microsatellite sequence	B5	7.00E−07	JX579353.1
YF09 (CCRI 35)	*Theobroma cacao* cultivar Scavina 6 TT2 like MYB transcription factor (MYBPA) mRNA, complete cds	B4	1e−10	GU324346.1
YF10 (CCRI 12)	*Gossypium hirsutum* putative calcium-binding protein gene, complete cds;	C1	3.00E−27	EF457754.1
YF24 (CCRI 35)	*Ricinus communis* transitional endoplasmic reticulum ATPase, putative, Mrna	A3	1E−09	XM002519456.1
YF25 (CCRI 12)	*Gossypium hirsutum* clone de-Met-HMTGC-EACG genome sequence	A3	1e−95	EF159966.1
YF29 (Zhong 07)	*Theobroma cacao* plasma-membrane choline transporter family protein isoform 1 (TCM_005837) mRNA, complete cds	C2	3.00E−25	XM_007052403.1
YF30 (CCRI 35)	*Theobroma cacao* origin recognition complex protein 5 (TCM_005313) mRNA, complete cds	A3	4.00E−10	XM_007051712.1
YF46 (CCRI 12)	*Theobroma cacao* leucine-rich receptor-like protein kinase family protein (TCM_001284) mRNA, complete cds	A3	1E−08	XM_007048104.1
YF47 (Zhong 07)	*Theobroma cacao* leucine-rich receptor-like protein kinase family protein (TCM_001284) mRNA, complete cds	A1	2.00E−08	XM_007048104.1
YF51 (CCRI 35)	*Theobroma cacao* leucine-rich receptor-like protein kinase family protein (TCM_001284) mRNA, complete cds	A3	2.00E−08	XM_007048104.1
YF57 (Zhong 07)	*Theobroma cacao* cytokinin oxidase (TCM_037384) mRNA, complete cds	A1	5.00E−29	XM_007012355.1
YF59 (Zhong 07)	*Theobroma cacao* potassium channel in 3 (TCM_037125) mRNA, complete cds	A3	5.00E−15	XM_007011957.1
YF60 (CCRI 35)	*Theobroma cacao* no pollen germination related 1 (TCM_034697) mRNA, complete cds	A3	3.00E−08	XM_007018429.1

Table 3 Expression analysis of homologous sequences

3. Epigenetic mechanisms of salt Tolerance and heterosis in upland cotton

Table 4 Salt-tolerance level of CCRI29 and its two parents

Cotton material	Salt-tolerance index (%)*	Salt-tolerance level
CCRI29	60.84a	Tolerant
P1	54.50b	Tolerant
RP4	42.39c	Sensitive
Mid-parent heterosis of salt-tolerance level	25.59	

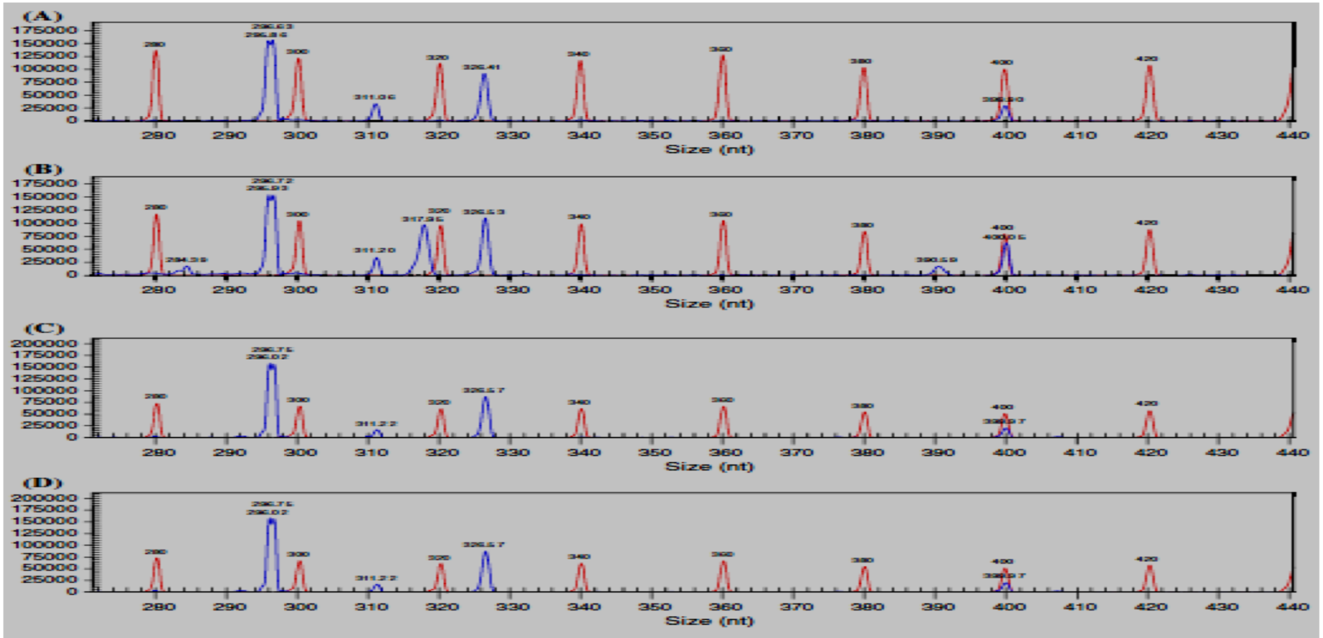
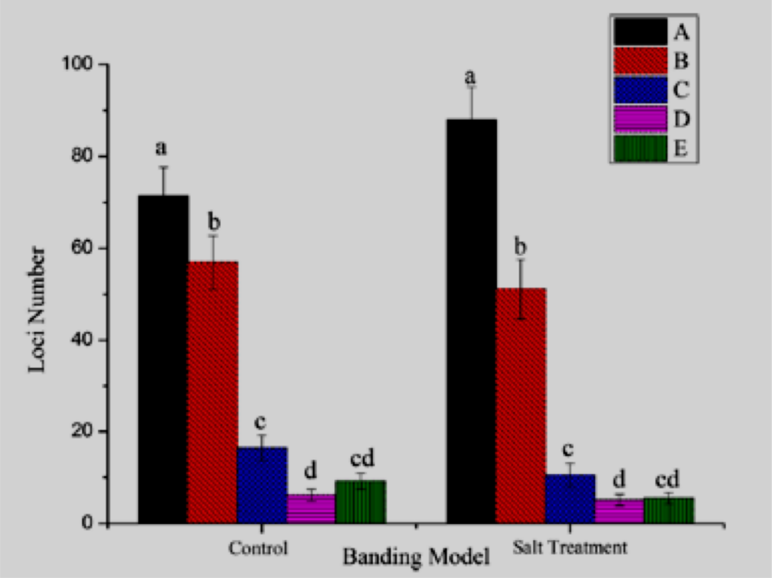
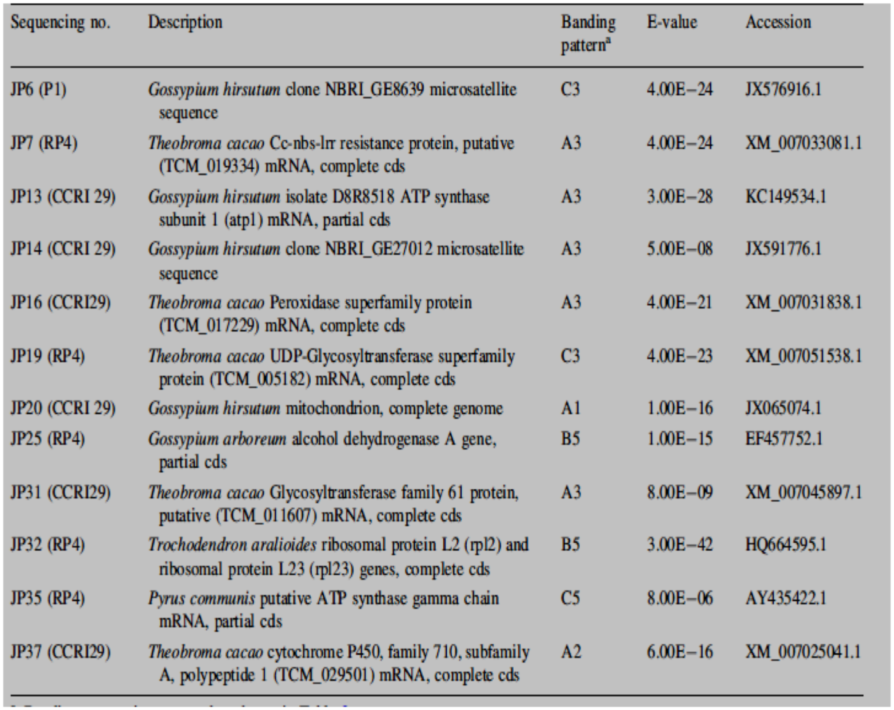


Fig.2-6 DNA methylation patterns of CCRI 29 detected in CE with the primer combination

- **Most cytosine methylation sites in hybrid CCRI 29 shared the same status as that of at least one of the parents**
- **The number of parents hypomethylation is significantly higher than that of hypermethylation in CCRI 29 under both control and salt stress.**
- **Demethylation could be the mechanism to explain heterosis in cotton hybrid.**

Table 4 Variation of DNA methylation patterns between control and salt treatment



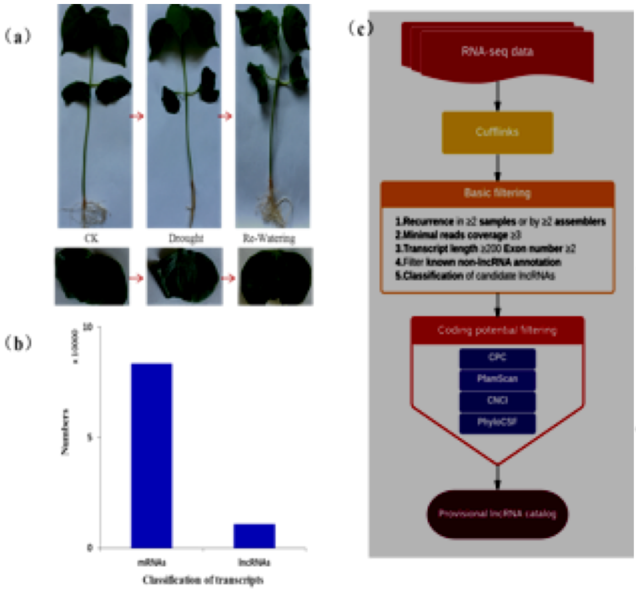


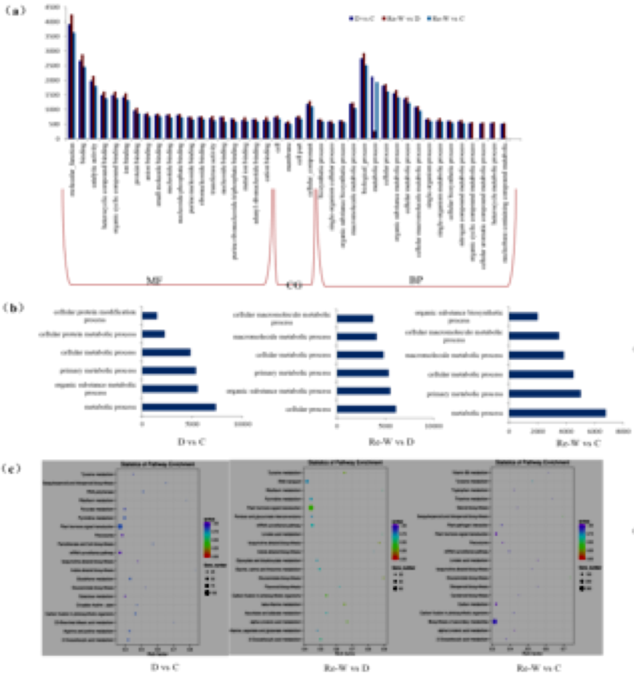
Sequencing no.	Description	Banding pattern[a]	E-value	Accession
JP6 (P1)	*Gossypium hirsutum* clone NBRI_GE8639 microsatellite sequence	C3	4.00E−24	JX576916.1
JP7 (RP4)	*Theobroma cacao* Cc-nbs-lrr resistance protein, putative (TCM_019334) mRNA, complete cds	A3	4.00E−24	XM_007033081.1
JP13 (CCRI 29)	*Gossypium hirsutum* isolate D8R8518 ATP synthase subunit 1 (atp1) mRNA, partial cds	A3	3.00E−28	KC149534.1
JP14 (CCRI 29)	*Gossypium hirsutum* clone NBRI_GE27012 microsatellite sequence	A3	5.00E−08	JX591776.1
JP16 (CCRI29)	*Theobroma cacao* Peroxidase superfamily protein (TCM_017229) mRNA, complete cds	A3	4.00E−21	XM_007031838.1
JP19 (RP4)	*Theobroma cacao* UDP-Glycosyltransferase superfamily protein (TCM_005182) mRNA, complete cds	C3	4.00E−23	XM_007051538.1
JP20 (CCRI 29)	*Gossypium hirsutum* mitochondrion, complete genome	A1	1.00E−16	JX065074.1
JP25 (RP4)	*Gossypium arboreum* alcohol dehydrogenase A gene, partial cds	B5	1.00E−15	EF457752.1
JP31 (CCRI29)	*Theobroma cacao* Glycosyltransferase family 61 protein, putative (TCM_011607) mRNA, complete cds	A3	8.00E−09	XM_007045897.1
JP32 (RP4)	*Trochodendron aralioides* ribosomal protein L2 (rpl2) and ribosomal protein L23 (rpl23) genes, complete cds	B5	3.00E−42	HQ664595.1
JP35 (RP4)	*Pyrus communis* putative ATP synthase gamma chain mRNA, partial cds	C5	8.00E−06	AY435422.1
JP37 (CCRI29)	*Theobroma cacao* cytochrome P450, family 710, subfamily A, polypeptide 1 (TCM_029501) mRNA, complete cds	A2	6.00E−16	XM_007025041.1

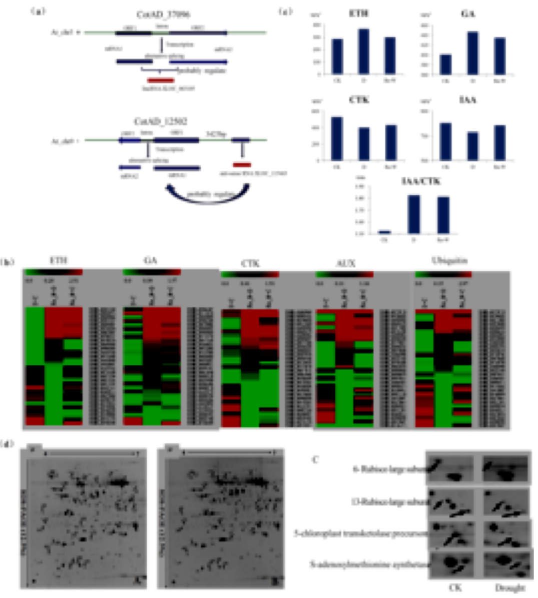
Table 5 Results of BLASTn for part of methylated sequences

4. lncRNAs responding to drought

- 10,820 lncRNAs of high-confidence were observed, Small RNAs revealed that 196 lncRNAs may be the precursors to small RNAs, most of which (35.7%, 70) were miRNAs.
- LncRNAs may be likely to involve in regulating plant hormones pathway in response to drought stress.

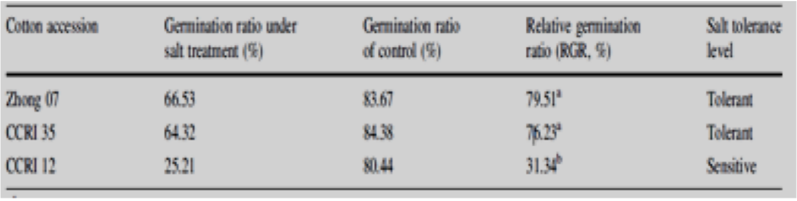




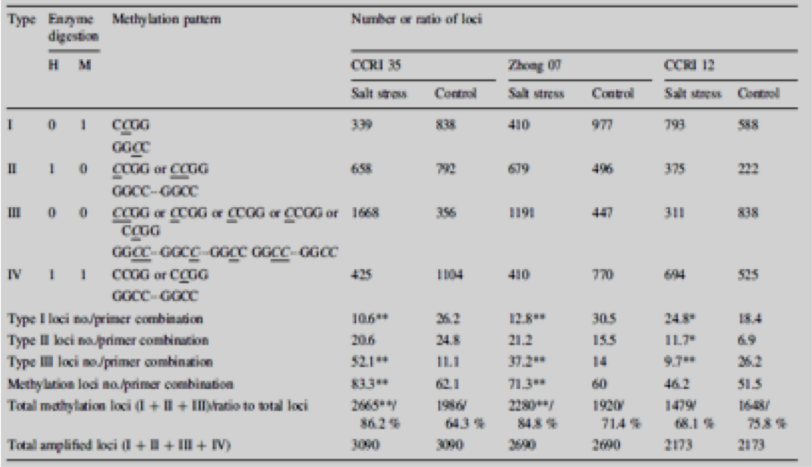


5. DNA methylation helps to understand the responding to salt stress

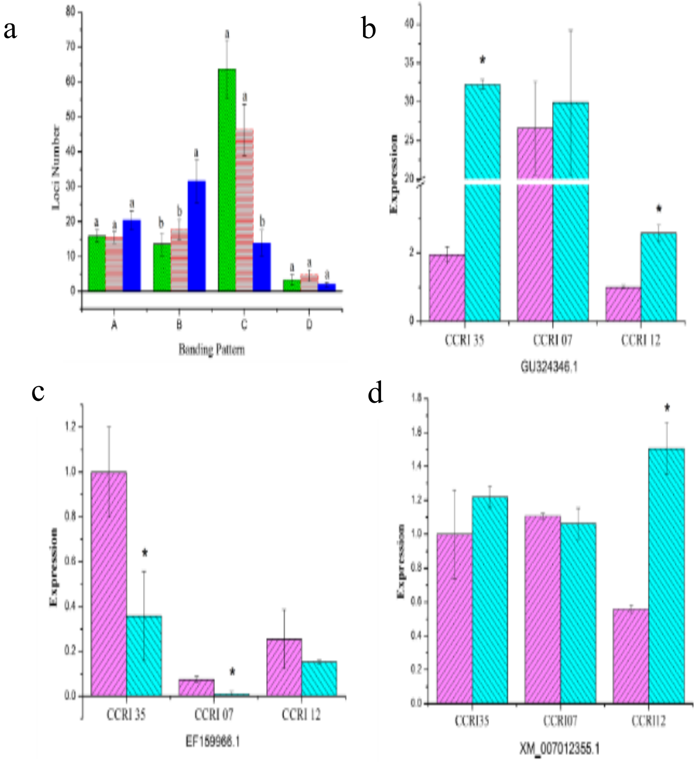
- Salt-tolerant cotton might have a mechanism of increasing the methylation level when responding to salt stress;
- Increasing level of DNA methylation and different methylation patterns might play important roles in active responding to salt stress in cotton.



Cotton accession	Germination ratio under salt treatment (%)	Germination ratio of control (%)	Relative germination ratio (RGR, %)	Salt tolerance level
Zhong 07	66.53	83.67	79.51^a	Tolerant
CCRI 35	64.32	84.38	76.23^a	Tolerant
CCRI 12	25.21	80.44	31.34^b	Sensitive

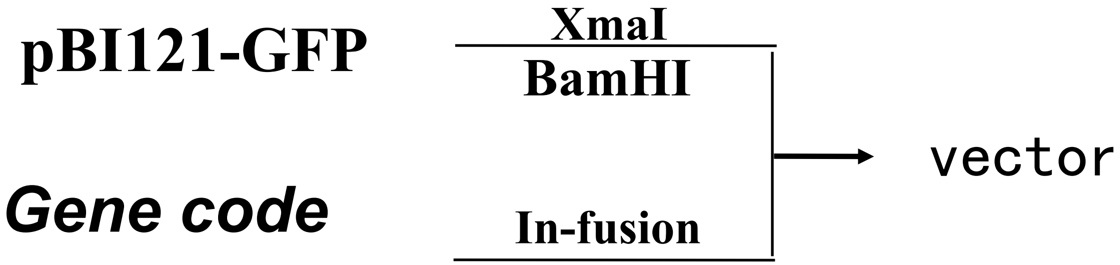


Type	Enzyme digestion		Methylation pattern	Number or ratio of loci					
				CCRI 35		Zhong 07		CCRI 12	
	H	M		Salt stress	Control	Salt stress	Control	Salt stress	Control
I	0	1	CCGG GGCC	339	838	410	977	793	588
II	1	0	CCGG or CCGG GGCC–GGCC	658	792	679	496	375	222
III	0	0	CCGG or CCGG or CCGG or CCGG or CCGG GGCC–GGCC–GGCC GGCC–GGCC	1668	356	1191	447	311	838
IV	1	1	CCGG or CCGG GGCC–GGCC	425	1104	410	770	694	525
Type I loci no./primer combination				10.6**	26.2	12.8**	30.5	24.8*	18.4
Type II loci no./primer combination				20.6	24.8	21.2	15.5	11.7*	6.9
Type III loci no./primer combination				52.1**	11.1	37.2**	14	9.7**	26.2
Methylation loci no./primer combination				83.3**	62.1	71.3**	60	46.2	51.5
Total methylation loci (I + II + III)/ratio to total loci				2665**/ 86.2 %	1986/ 64.3 %	2280**/ 84.8 %	1920/ 71.4 %	1479/ 68.1 %	1648/ 75.8 %
Total amplified loci (I + II + III + IV)				3090	3090	2690	2690	2173	2173



Genes of salinity-tolerance

pBI121-GFP vector



pBI121-GFP

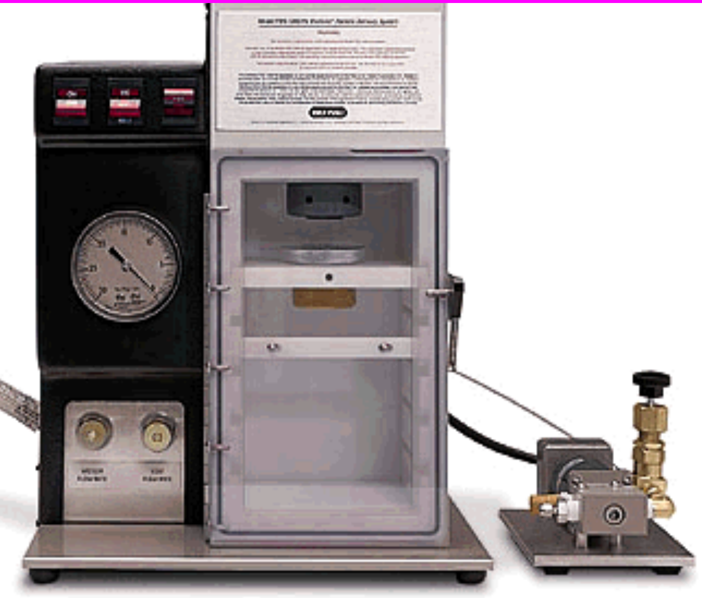
XmaI
BamHI

Gene code

In-fusion

vector

Transgenetic method: gene gun



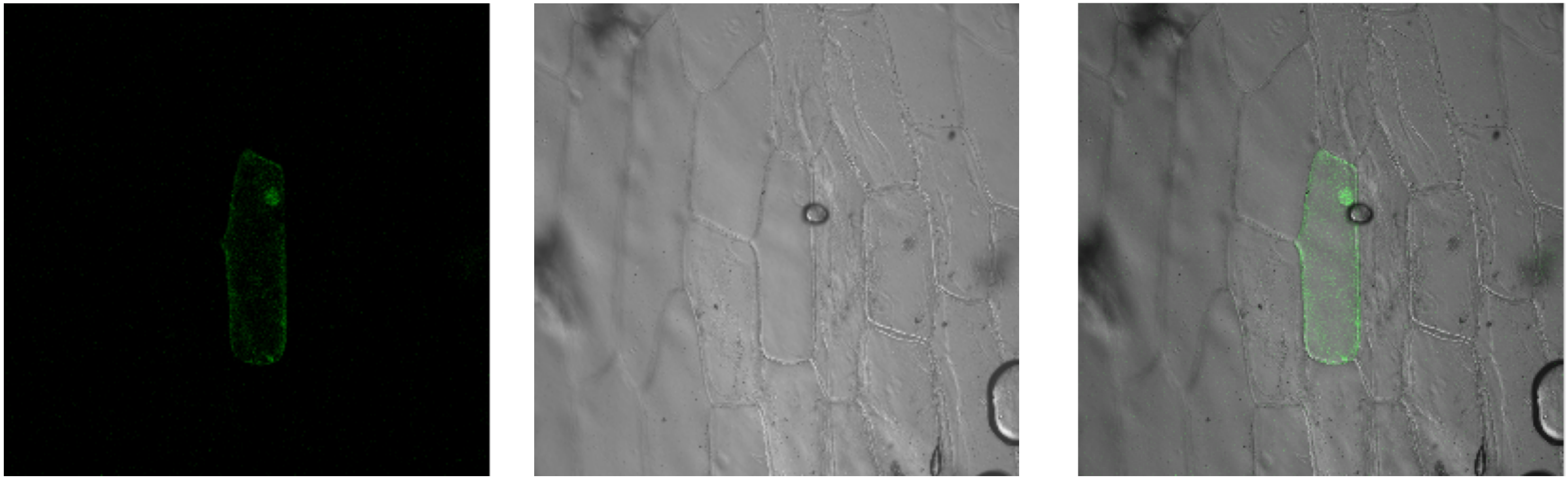
G-I：Biorad PDS-1000

G-2：Biorad helios

G-3: Wealtec GDS-80

Onion epidermis

50 psi/3 cm (1 μg：10kb/0.6 mg)

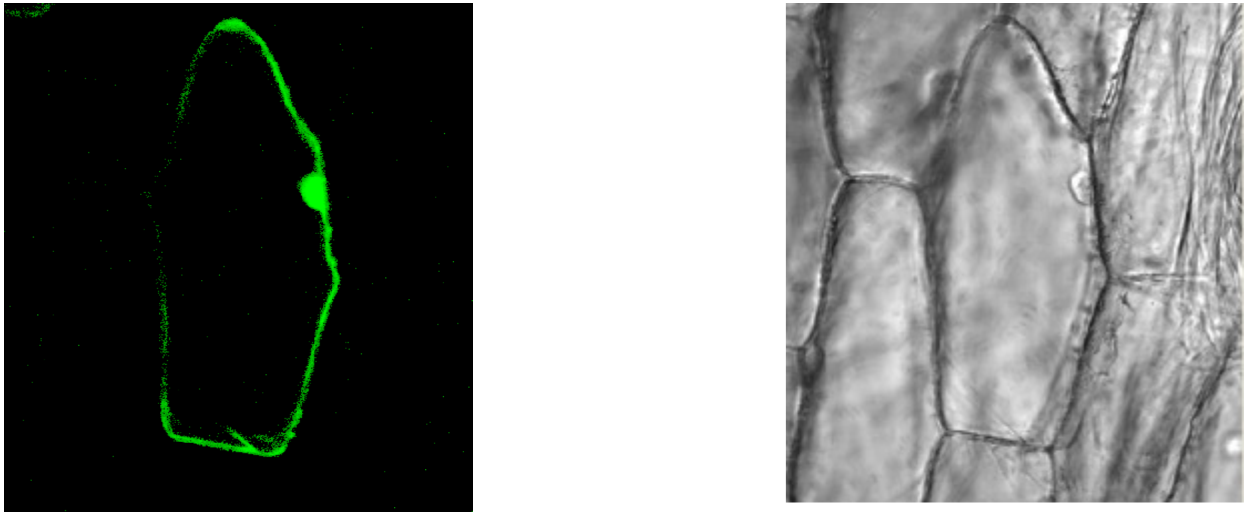


1. Subcellular Localization of *GhVP* and *GhSAMS*

onion

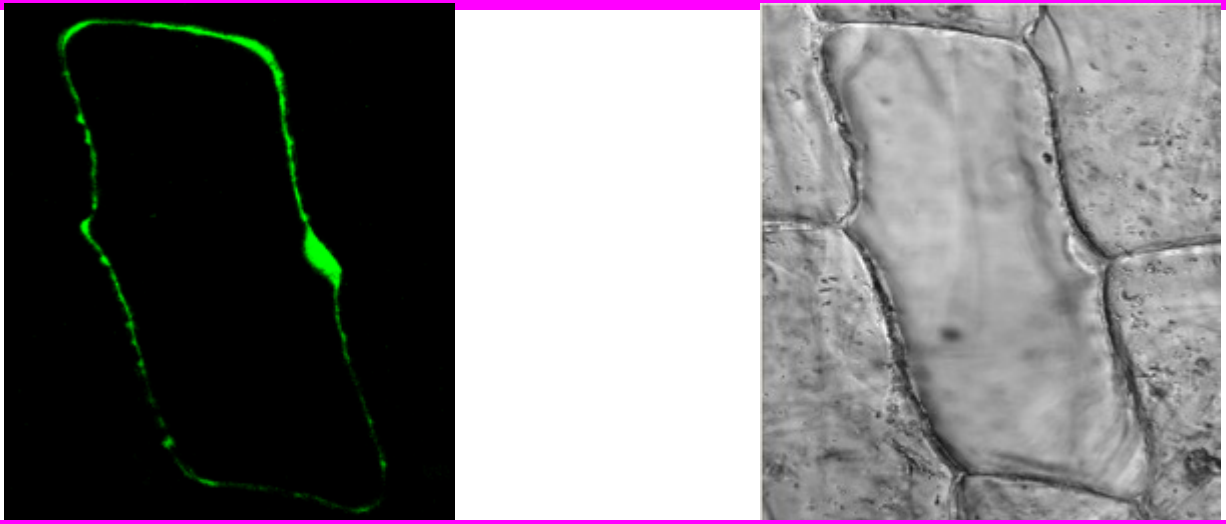
cotton

GhVP



Subcellular location: membrane of the plasma membrane and nucleus membrane---onion epidermis

GhSAMS



Subcellular location: membrane of the plasma membrane and nucleus membrane---onion epidermis

Cotton Varieties: ccri9835,ccri2067
Genes: GhVp and GhSAMS

Sperm

Top of the growing point

50 psi/3 cm (1 µg：10kb/0.6 mg)

Transgenetic seeds (2016)

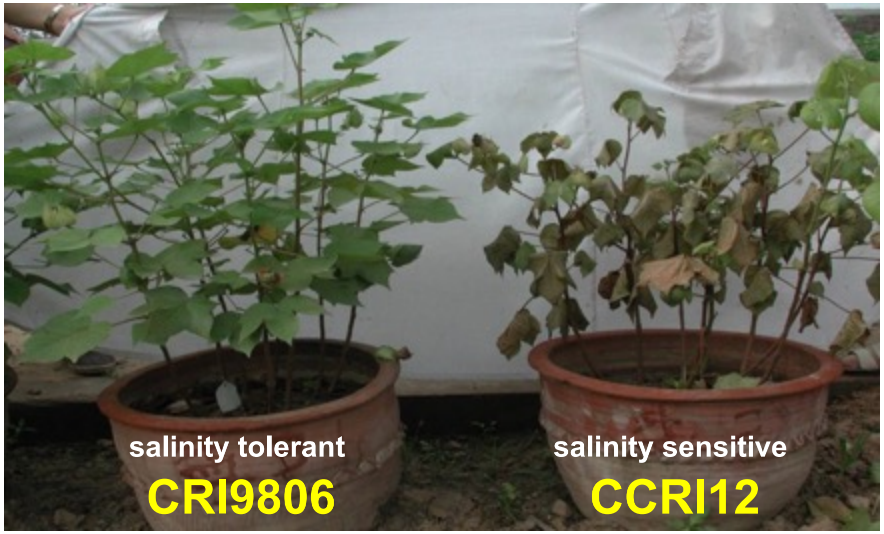
No.	receptor	genes	No.of seeds
1	ccri45	*GhVP*	180
2	ccri63	*GhVP*	60
3	ymz21	*GhVP*	96
4	sGK958	*GhVP*	87
5	ccri45	*GhSAMS*	99
6	ccri63	*GhSAMS*	120
7	ymz21	*GhSAMS*	78
8	sGK958	*GhSAMS*	56

Transgenetic seeds screened under the salinity stress(0.4%NaCl)

Some salinity toeralant races

Race No.	Salinity resistance index/%	Salinity resistance level	Race No.	Salinity resistance index/%	Salinity resistance level
ccri9806	76.5	resistant	804129	58.2	tolerant
806081	74.9	tolerant	805021	55.3	tolerant
810151	72.3	tolerant	810011	50.6	tolerant
807071	72.3	tolerant	804137	64.3	tolerant
807017	71.2	tolerant	805133	63.2	tolerant
810121	65.3	tolerant	806011	58.9	tolerant
807067	55.2	tolerant	ccri9612	21.5	sensitive

CRI9806 grown in the saline(0.4%NaCl) land.



1	**The draft genome of a diploid cotton Gossypium raimondii**	**Nature Genetics**	**2012,44(10):1098-1103**
2	**Genome sequence of the cultivated cotton Gossypium arboreium**	**Nature Genetics**	**2014,46, 567-574**
3	**Genome sequence of cultivated Upland cotton (Gossypium hirsutum TM-1) provides insights into genome evolution**	**Nature biotechnology**	**2015,33(5):524-530**
4	**Genome-wide Identification and Structural Analysis of Pyrophosphatase Gene Family in Cotton**	**Crop Science**	**2016,56:1-10**
5	**Genome-Wide Analysis of Long Noncoding RNAs and Their Responses to Drought Stress in Cotton (G. hirsutum L.)**	**PLOS ONE**	**2016,11(6):e0156723**
6	**Mining and Analysis of SNP in Response to Salinity Stress in Upland Cotton (Gossypiumhirsutum L.)**	**PLOS ONE**	**2016,11(6):e0158142**
7	**Genome-wide Identification and analysis of the stress-resistance function of the TPS (Trehalose-6-Phosphate Synthase) gene family in cotton**	**BMC Genetics**	**2016,17:54**
8	**In vitro regeneration protocol for synthetic seed production in upland cotton (Gossypium hirsutum L.)**	**Plant Cell, Tissue & Organ Culture**	**2016,123:673-679**
9	**Genome-wide identification and expression analysis of CIPK genes in diploid cottons**	**Genetics & Mol. Research**	**2016,15(4)gmr15048852**
10	**Epigenetic mechanisms of salt tolerance and heterosis in Upland cotton (Gossypium hirsutum L.) revealed by methylation-sensitive amplified polymorphism analysis**	**Euphytica**	**2016, 208:477–491**
11	**Cloning of SjCA gene and its expression analysis on upland cottons**	**Journal of Biomedical Engineering and Informatics**	**2016, 2(2):150-162**

Thanks!!