

作物智能设计技术与育种应用

谷晓峰

中国农业科学院生物技术研究所

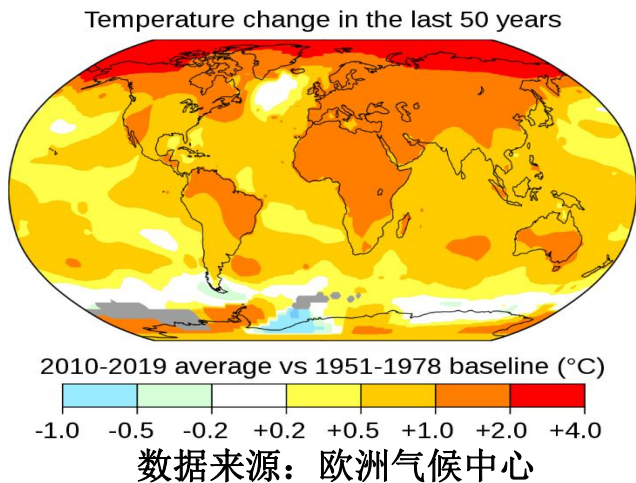
中国农业科学院生物技术研究所

Biotechnology Research Institute, CAAS

中国农业科学院生物技术研究所
Biotechnology Research Institute, CAAS

全球气候变暖，极端天气频发严重影响粮食产量

- 矮化育种和杂种优势技术两次变革，分别使水稻等作物产量提高了15%以上
- 实现增产、抗逆、高效、品质等性状提升目标，有赖于基础理论和新一轮技术突破和应用



生长发育迟缓，结实率下降

育种技术创新

驯化选育
育种1.0



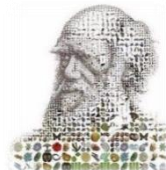
1900年
孟德尔遗传学



1953年
DNA双螺旋
分子生物学



1990年
组学和系统生物学等



21世纪
计算生物学和合成生物学等

科学理论突破



杂交品种
1930s
产量增益
70%



高肥水高产品种
1960s, 水稻和小麦半
矮秆基因育种
农业绿色革命

种质资源



转基因品种
1990s
抗虫耐除草剂
生产方式变革

基因资源

杂交育种
育种2.0

分子育种
育种3.0



高产、优质、固氮等品种
2000s, 基因编辑、合成生物
人工智能等颠覆性技术应用

智能设计育种
育种4.0

智能设计推动生物育种技术变革

- 新一轮生物育种技术突破成为驱动种业核心竞争力提升的战略需求

- 多维大数据、基因位点和性状调控网络为智能设计技术提供支撑

生物技术和信息技术融合，推动智能设计育种技术创新



表观遗传是加速智能设计育种的关键和新方向

性状决定要素：遗传 + **表观遗传** + 环境 + 遗传 X **表观遗传** X 环境

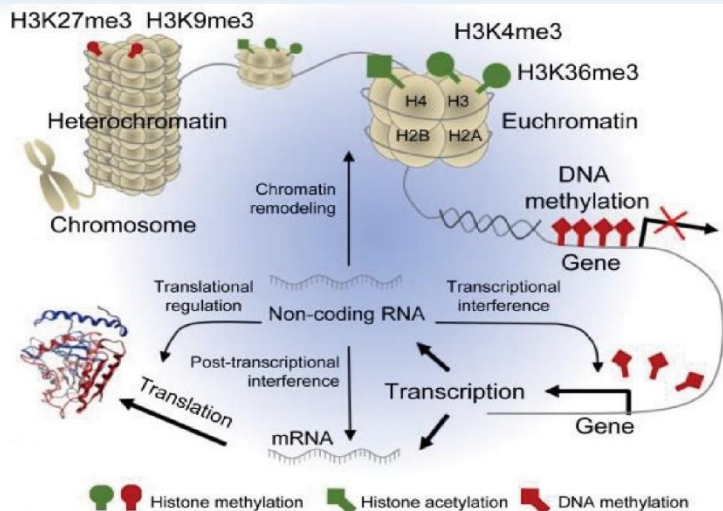


表观遗传在遗传和环境互作中起着重要作用，精准控制表观遗传与二者之间的协同调控是培育高产稳产品种关键

表观遗传参与调控生命发育全过程

● **表观遗传**是不依赖DNA序列改变调控基因表达的方式

● **参与生物发育的每一个阶段**，起着关键的调控作用



Jiang et al., JIPB, 2022

修饰类型：

DNA 甲基化

RNA 甲基化

组蛋白修饰

nc RNA

染色质重塑

.....

精准控制复杂性状形成，决定物种多样性

表观智能设计推进基础理论和资源创新

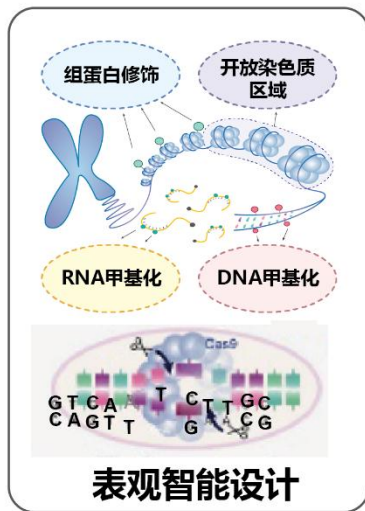
● **基础理论**：大数据联合分析、智能预测和解析修饰位点和调控机制

● **表观设计**：融合合成生物、基因编辑、人工智能等前沿技术

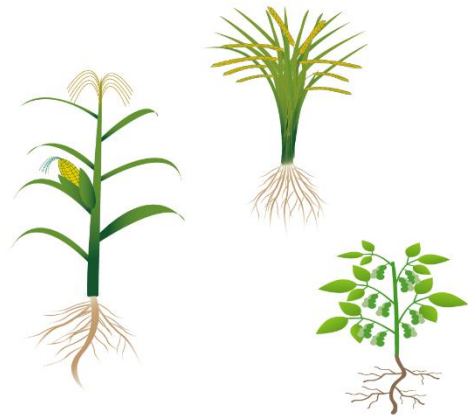
● **优异种质**：加速种质资源多性状的鉴定和优异新种质的定向创制



.....
种质资源性状多样性



高产
优质
多抗
高效
...



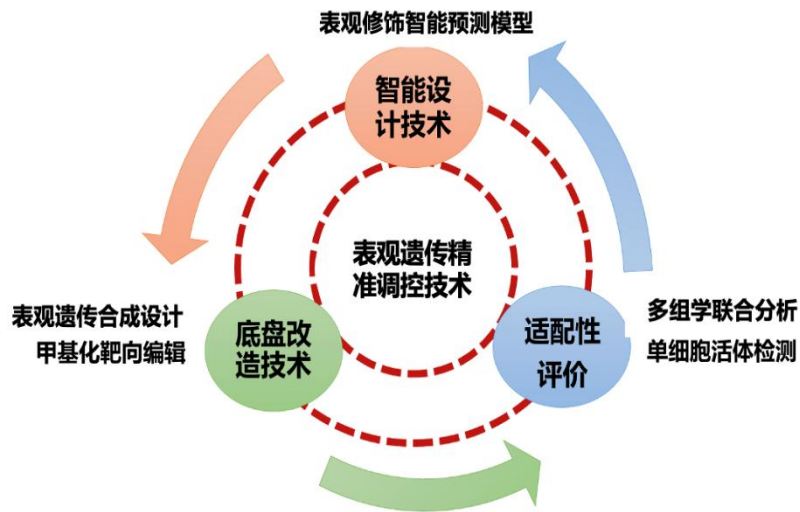
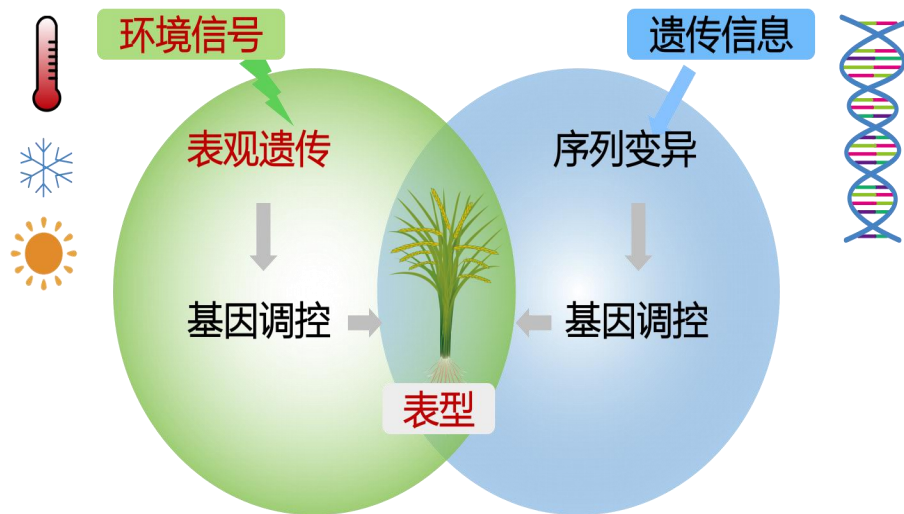
.....
优异种质

作物表观遗传智能设计技术体系亟待建立

● 多物种新型表观遗传修饰网络了解较少

● 缺少多物种表观智能预测模型和设计技术

● 仅在少数物种中实现定向调控修饰水平



一、多组学绘制表观遗传图谱和鉴定关键因子

1、表观遗传修饰标记鉴定

- **DNA腺嘌呤甲基化(6mA)、RNA胞嘧啶甲基化 (m⁵C)**等新表观遗传修饰类型，参与调控发育、产量等重要农艺性状
- **确定6mA等是高温胁迫快速响应的表观标记**，鉴定*DDM1*等6个关键基因和育种元件，为抗逆性改良提供资源
- **DRW1、SFL1**等调控组蛋白修饰的动态变化

Developmental Cell, 2018

Molecular Plant, 2017, 2018, 2020

Trends in Plant Science, 2020

Plant Biotech J

Plant Cell, 2021

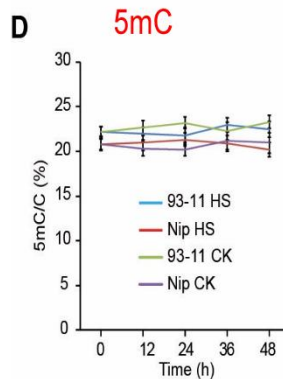
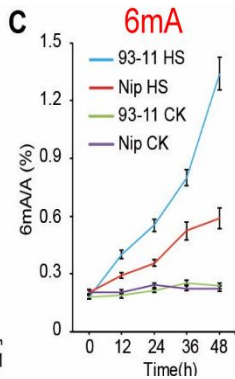
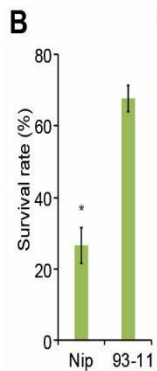
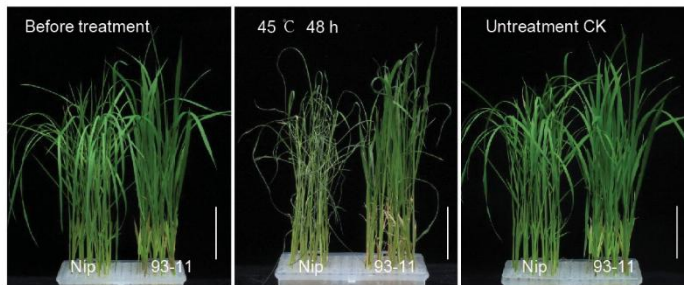
获得国内发明专利6项

国际PCT 专利1项

籼粳间DNA 6mA 修饰水平差异影响耐逆性能

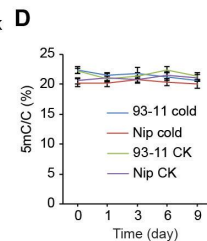
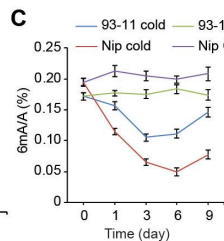
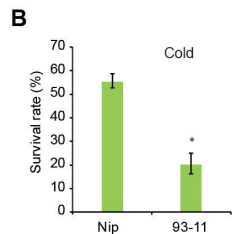
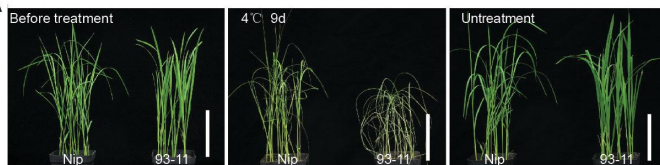
高温

A



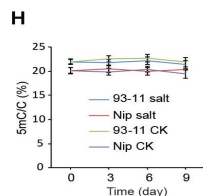
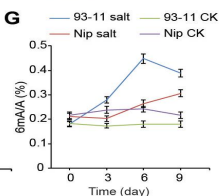
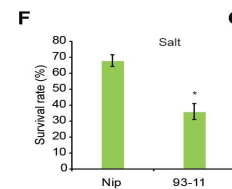
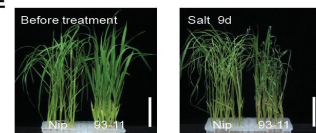
低温

A

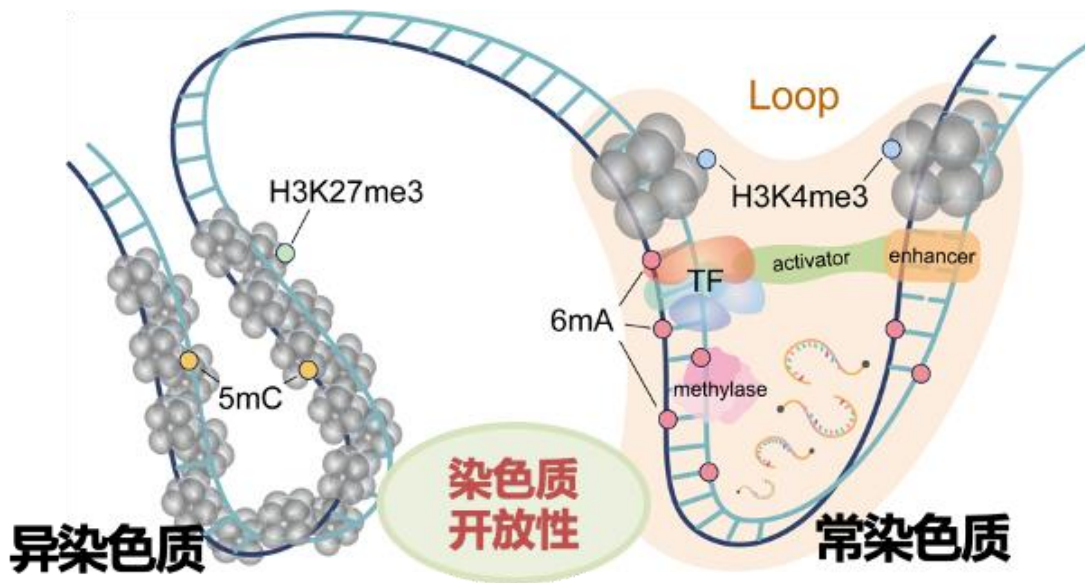


盐碱

E

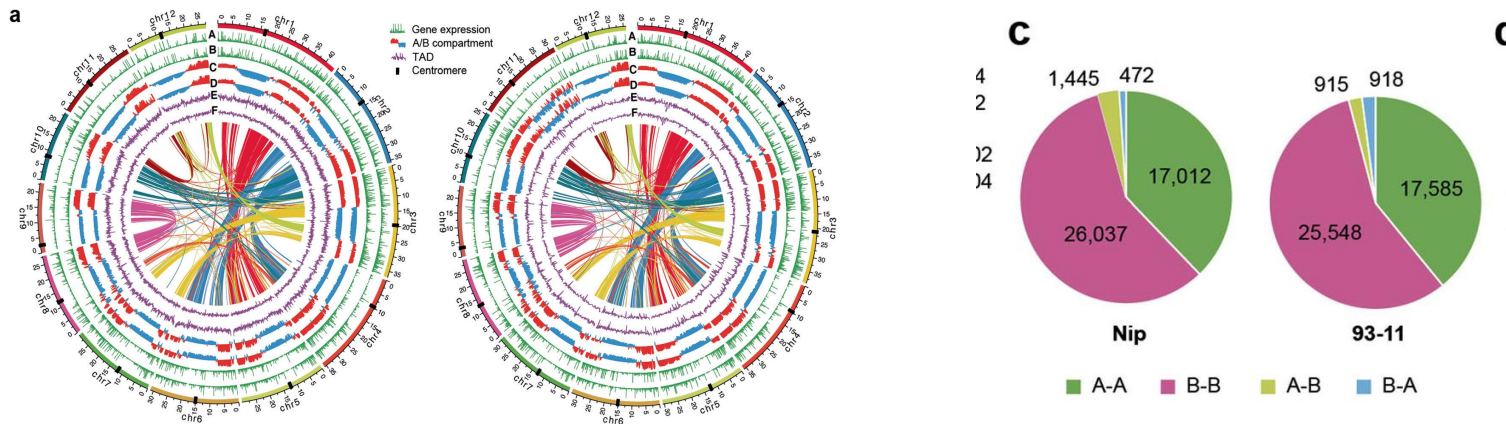


2、籼粳间染色质3D结构影响耐高温性能



- **表观遗传调控的染色质开放性**对三维空间结构变化起核心作用
- 染色质三维空间结构随细胞分化过程呈现高度动态变化

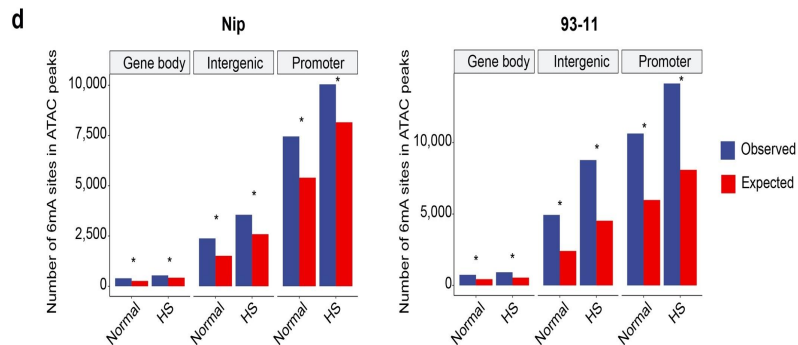
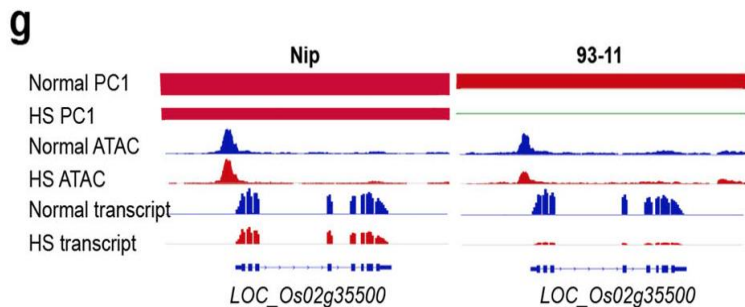
2、籼粳间染色质3D结构影响耐高温性能



HiC-seq、ATAC-seq、RNA-seq多组学联合分析

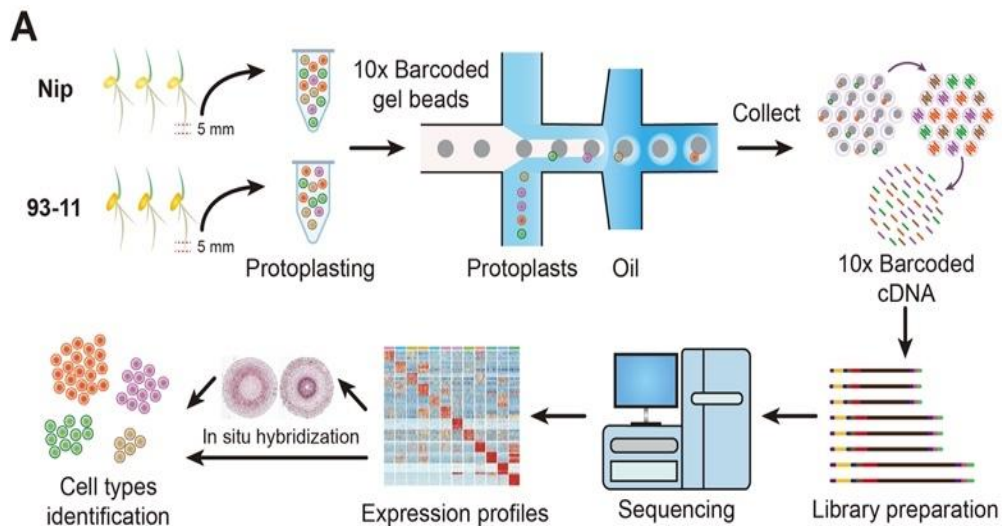
- 高温胁迫条件下，籼稻NIP和粳稻93-11染色质三维空间结构不同层级结构单元出现明显变化
- 93-11比NIP 表现出更多的**基因动态表达和染色质可及性变化**，这与“93-11”具备更强的高温耐受性一致。

2、籼粳间染色质3D结构影响耐高温性能



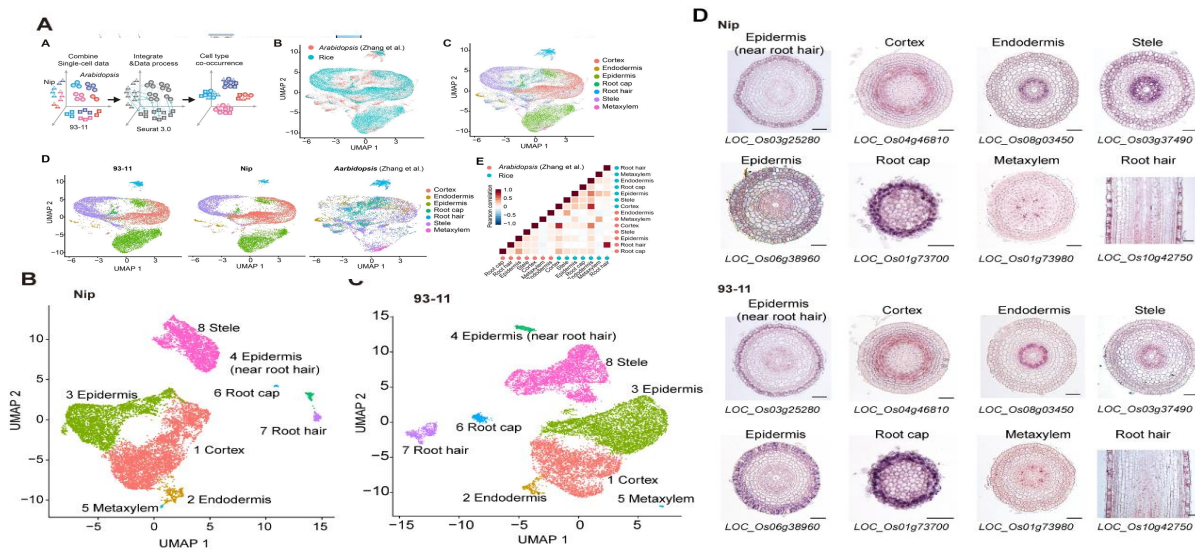
- 全基因组ATAC信号显著与DNA 腺嘌呤甲基化6mA修饰位点重叠
- 染色质可及性变化可能和DNA 6mA协同调控水稻的高温胁迫响应

3、籼粳稻单细胞转录组和染色质可及性图谱



- 建立高效的水稻单细胞分离和分析体系

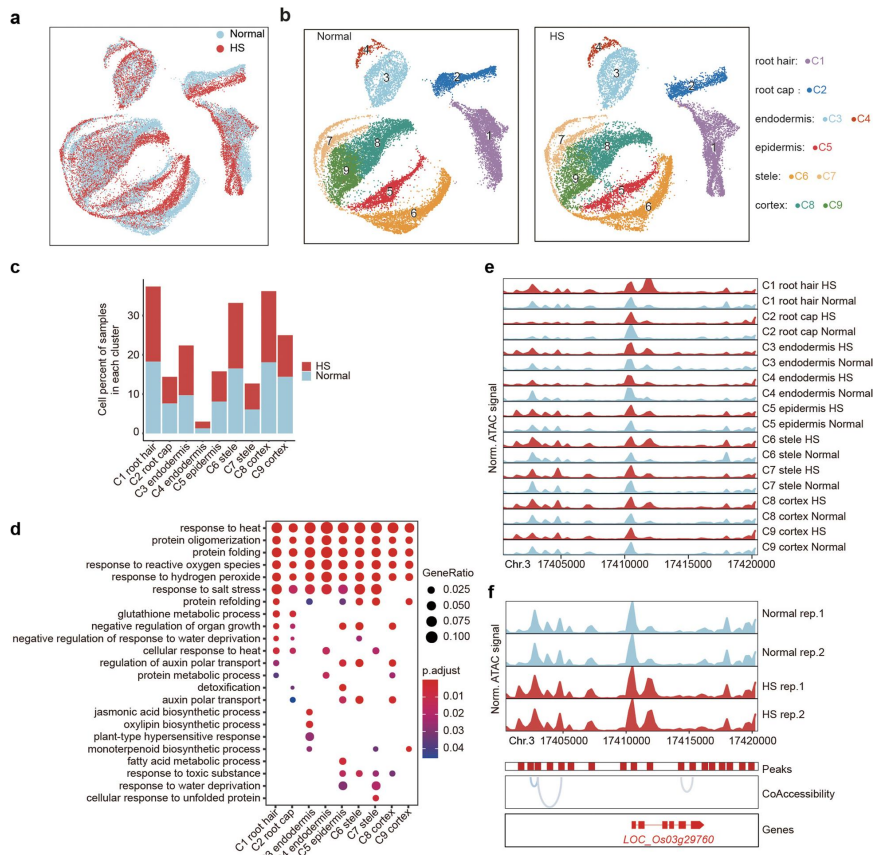
(1) 籼粳稻单细胞图谱绘制



- 鉴定了多个特异性细胞调控基因和胁迫响应基因
- 为研究植物细胞类型特化、功能验证提供了新途径和方法


Molecular Plant, 2021
获得软件著作权1项

(2) 水稻单细胞染色质可及性图谱和数据库



- **首个水根单细胞染色质可及性图谱**
- **揭示了不同细胞类型的特异性染色质可及性差异，重点阐述了单细胞水平染色质开放区对环境信号热胁迫下的变化**
- **发现了热胁迫下细胞类型特异性的染色质可及性及相关基因的表达变化**

(2) 水稻单细胞染色质可及性图谱和数据库



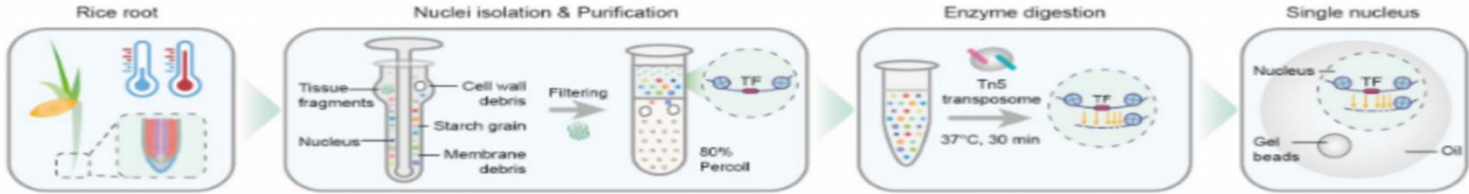
水稻根单细胞染色质可及性数据库SCAR

[Home](#) [How to Cite](#) [Contact](#)

[HOME](#) [Cluster search](#) [Locus search](#) [Genome JBrowse](#) [Help](#) [Download](#)

Chromatin accessibility map of rice root tips at single-cell resolution

Here, we applied single-cell ATAC-seq to 46,758 cells from rice root tips under normal and heat stress conditions to comprehensively delineate chromatin accessibility during root development of an important crop. In this platform, we developed a single-cell chromatin accessibility database for describing chromatin responses to heat, and identifying cell type specific accessibility changes to this key environmental stimulus.



The flowchart illustrates the experimental workflow for single-cell ATAC-seq on rice root tips. It consists of four main stages: 1. **Rice root**: Starting with a rice root sample. 2. **Nuclei isolation & Purification**: Tissue fragments are processed to isolate nuclei, removing cell wall debris, starch grains, and membrane debris. The nuclei are then filtered and purified using 80% Percoll. 3. **Enzyme digestion**: The purified nuclei are treated with a Tn5 transposome at 37°C for 30 minutes to generate chromatin accessibility. 4. **Single nucleus**: The final product is a single nucleus containing a nucleus, gel beads, and oil.

Welcome to eLab in CAAS
eLab located at Biotechnology Research Institute, Chinese Academy of Agricultural sciences, 12 Zhongguancun South Street, Beijing, China.

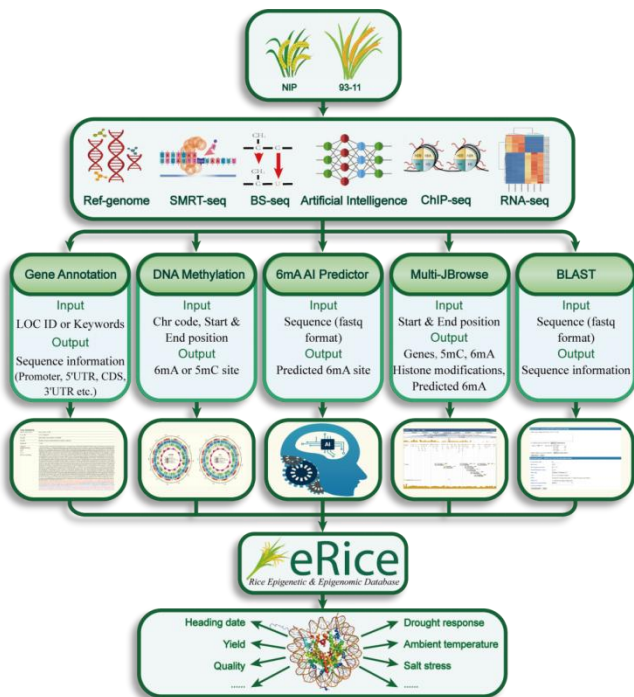
表观遗传修饰位点鉴定和应用

- 表观遗传具有细胞、组织、发育等时空动态变化特点，检测技术和高通量测序的发展推动了表观组学发展
- 受到检测技术、实验成本、取材组织以及表观修饰动态可逆的调控特性，**仍有大量表观修饰位点没有得到发掘和研究**

如何利用现有表观组学数据深入探索潜在的表观修饰位点是表观遗传研究面临的重要问题

二、表观遗传智能预测和设计

1、构建了水稻表观遗传智能数据库



测序品种：粳稻日本晴和籼稻9311

数据种类：基因组

DNA 腺嘌呤甲基化

DNA 胞嘧啶甲基化

人工智能全基因组预测

组蛋白修饰

转录数据

数据格式：输入和输出格式

Plant Biotechnology J , 2020

获得软件著作权1项

数据库界面及栏目

eRice
Rice Epigenetic & Epigenomic Database

About Site Map How to Cite Help Contact

Search in eRice Search

HOME Gene Annotation BLAST DNA Methylation Components 6mA AI Predictor Multi-JBrowse Download

Welcome to eRice

News

November 1, 2020
Optimized webpage.
Added components column.

January 9, 2020
eRice database is published on *Plant Biotechnology Journal*.

November 1, 2019
eRice v1.1 is online.

August 3, 2019
Improved eRice website design.

May 1, 2019
Improved Genome Annotation.

March 25, 2019
Updated 6mA AI Predictor.

January 29, 2019
eRice v1.0 is online.

November 17, 2018
A paper describing DNA 6mA methylation in rice and its association with gene expression, plant development and stress responses has been published on *Molecular Plant*.

Rice Epigenomic & Genomic Annotation

Epigenetic marks played an essential role in various pathways, including transcription regulation, DNA replication and repair, development and differentiation as well as response to environmental factors. Thanks to single-molecule real-time (SMRT) sequencing, recent discovery of DNA N6-methyladenosine (6mA) methylation provided new insights into altering the activities and functions of developmental signals and stress responses.

In rice (*Oryza sativa*), genome-wide 6mA distribution at single-nucleotide resolution has been reported in two main and parental cultivars (Nipponbare, *japonica* group; 93-11, *indica* group), which were regarded as reference genomes of *japonica* and *indica* groups. Here, we have developed a species-specific epigenomic database, eRice (an Epigenomic & Genomic Annotation Database for Rice), to facilitate efficient epigenomic studies for both *japonica* and *indica* groups.

6mA AI-predictor

6mA SMRT-seq

Related to sites

NCBI

Rice Genome Annotation Project

基因注释

基因和表观修饰检索

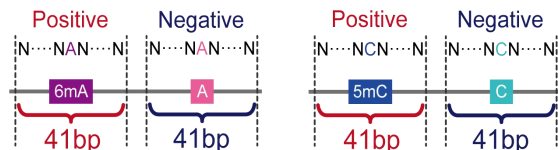
DNA甲基化

表观修饰元器件

DNA 甲基化人工智能预测

多组学可视化窗口

DNA甲基化 AI 预测界面



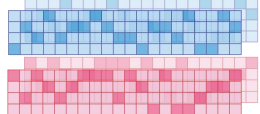
41 x 4



(20~800) x 4



Convolutions



Training data

Example:

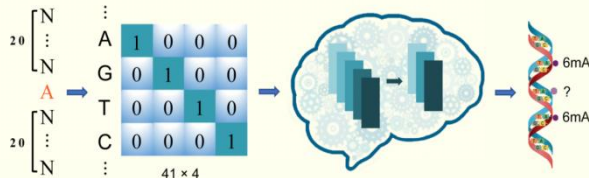
```

>test
CCGGAGAGACACGGGATCCAGGGCCGACGGAGGGATCCGGGATCTGCCGCCACGACCCGAGTCATCCTTGGATCCACACDITGGCCACTAGAGAAAAGGA
GGGGAGAGGCGAGGGCAACCGAGAGAGATAGTTGAGGGAGAGGGGACAGCCGAACTCTCCGATGGGGATGATTTTTATTATAGACTAGAGATGATTATTGGAT
GAAGGAGAGAGAGTAGAGAGGAGCTAGAGAGATCCCTCTCACTCCATCTCAACCTACCCCTCCCAAGATTTTTTAGATATGTATATCTGTADCGCTCAGT
TGTACTGCTGGCTACCCCATTTGCAACGCTATGCAAGATCTAGCTGCTGTCAATTCAGATGGAAGAACAGGGCTCAGTTTGCATCTGGCCATATCTGGA
AAAGAAAGGATCGAGATTTAGTTTTAGTTTCCCTTACTCTGAGGCTACTAGAGGATAGGATCTTATATATGAGATGGAGGAGGACTACTAGATCTGTA
TCTTTAATTTTGGAAAGGAGGAGTAGATTAATTTGTTATTTGGAAAGGAGGAGTAGTATTAATTTGTTGTTTTAGAGGTAATGAAGACTGAGATTAG
ATATTGCAATAGATGTTGAGAAACCATATCCATTCACTAGACAGGCTGCTAGACTCTGCGATGAGATTCAGCGGATTCGATCAGCAGACTGGCAATCCGGTGA
ATTCGATCGAGAGAGCTTTTTAGCTGAGATCCATTAAGTTTTGATCCAGAGGGTAGAGAGGTTTTGGCTGGCCACACTCAGTTGAGAAAAAATTTTAGAGAGS
TATAAAATCACATCTTCATAAAGTCAGGAGCCACATCATCTCCAGGGCTGTTTTCCCTCAGCAAAAAGAACAGGTCGAG
    
```

example results

Type/Paste the gene sequence with the Fasta format below:

Predictor Reset



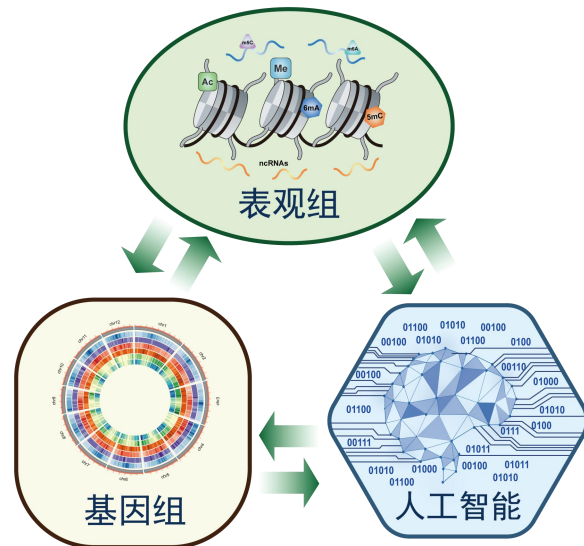
6mA AI 预测工具

6mA AI Predictor				
No	Sequence	Probability to be 6mA	Predicted	Result
1	CCGGAGAGACACGGGATCCAGGGCCGACGGAGGGATCCGG	0.953	6mA	No
2	GACAAAGGGATCCAGGGCCGACGGAGGGATCCGGATCTGC	0.018	No	No
3	ACGGGATCCAGGGCCGACGGATCCGGGATCTGCCGCC	0.014	No	No
4	GATCCAGGGCCGACGGATCCGGGATCTGCCGCCACG	0.030	No	No
5	GGCCAGGACGGATCCGGATCCGCCGCCAGCCGAG	0.391	No	No
6	GGATCCGGGATCTGCCGCCACGACCCGAGTATCTGTTGA	0.001	No	No
7	TCCGGGATCTGCCGCCACGACCCGAGTATCTGGATCC	0.023	No	No
8	GATCTGCCGCCACGACCCGAGTATCTGGATCCACCC	0.890	6mA	No
9	TGCCGCCACCCGAGTATCTGGATCCACCCGTTG	0.001	No	No
10	ACGACCCGAGTATCTGGATCCACCCGTTGCCACTAGA	0.585	6mA	No
11	CCCCGATCTGGGATCCACGCTGCCACTAGAGAAA	0.003	No	No
12	GAGTCACTGTTGATCCACCTGTTGCCACTAGAGAAAAG	0.005	No	No
13	GTTGGATCCACCTGTTGCCACTAGAGAAAAGGAGGCGAG	0.010	No	No
14	GGATCCACCTGTTGCCACTAGAGAAAAGGAGGCGAGCG	0.519	6mA	No
15	ATCCACCTGTTGCCACTAGAGAAAAGGAGGCGAGAGCGA	0.014	No	No
16	CCACCTGTTGCCACTAGAGAAAAGGAGGCGAGGCGAGG	0.002	No	No
17	CACCACTGTTGCCACTAGAGAAAAGGAGGCGAGGCGAGG	0.004	No	No
18	ACCACTGTTGCCACTAGAGAAAAGGAGGCGAGGCGAGG	0.004	No	No
19	CCACTGTTGCCACTAGAGAAAAGGAGGCGAGGCGAGGCA	0.003	No	No
20	CACGTTGCCACTAGAGAAAAGGAGGCGAGGCGAGGCGAA	0.018	No	No
21	GTTGCCACTAGAGAAAAGGAGGCGAGGCGAGGCGAAAAC	0.953	6mA	No
22	CACTAGAGAAAAGGAGGCGAGGCGAGGCGAAAACGGAAG	0.576	6mA	No
23	CTAGAGAAAAGGAGGCGAGGCGAGGCGAAAACGGAAGAA	0.136	No	No
24	AGAAAAGGAGGCGAGGCGAGGCGAAAACGGAAGATAAG	0.314	No	No
25	AAGGAGGCGAGGCGAGGCGAAAACGGAAGATAAGTTGAG	0.003	No	No
26	AGGAGGCGAGGCGAGGCGAAAACGGAAGATAAGTTGAGG	0.004	No	No
27	GGAGGCGAGGCGAGGCGAAAACGGAAGATAAGTTGAGGG	0.000	No	No
28	GAGGCGAGGCGAGGCGAAAACGGAAGATAAGTTGAGGGA	0.005	No	No
29	CGAGGCGAGGCGAAAACGGAAGATAAGTTGAGGGAGAGG	0.043	No	No
30	GAGGCGAGGCGAAAACGGAAGATAAGTTGAGGGAGAGG	0.001	No	No

甲基化位点预测
(位置、序列、修饰水平)

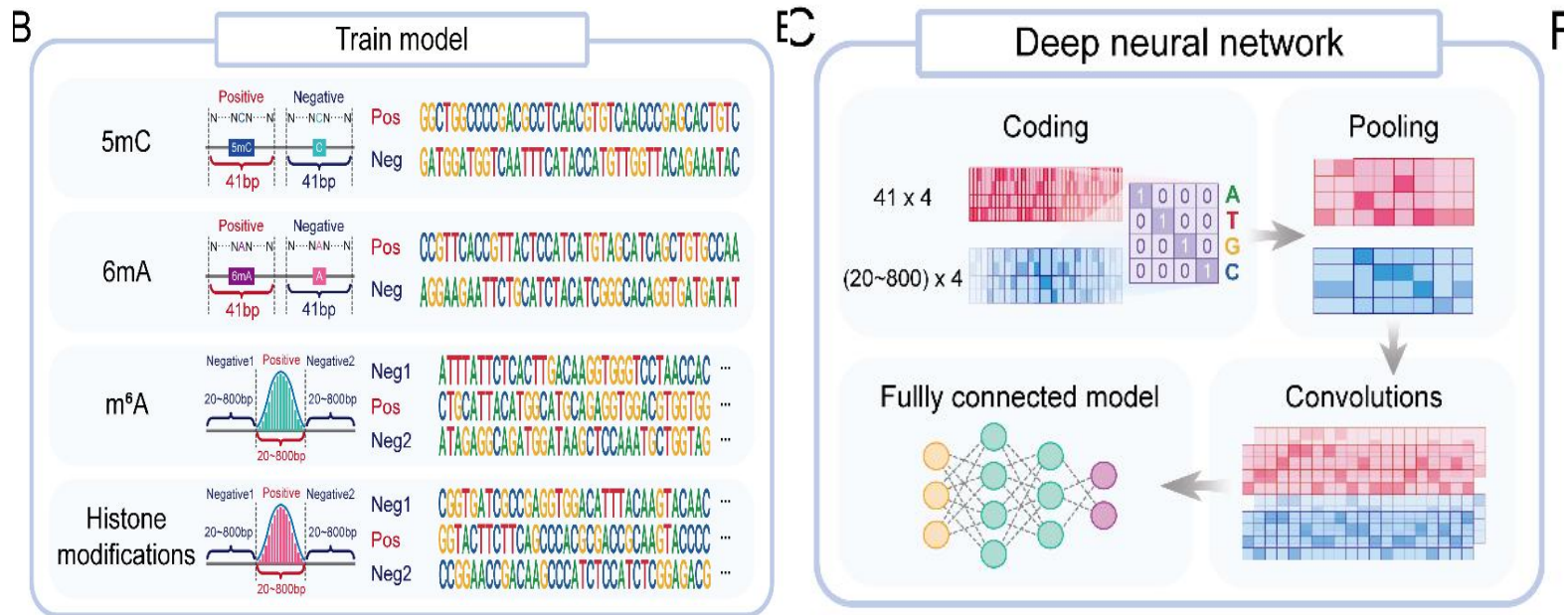
多组学和人工智能交叉融合

Method	Sn(%)	Sp(%)	ACC(%)	MCC(%)
PseDNC	63.52	65.57	64.55	0.29
i6mA-Pred	82.95	83.30	83.13	0.66
MM-6mAPred	89.32	90.11	89.72	0.79
iDNA6mA	86.70	86.59	86.64	0.73
SDM6A	85.20	90.90	88.10	0.76
iDNA6mA-rice	83.86	83.41	83.63	0.67
SNNRice6mA	92.16	94.32	93.24	0.87
SMEP	90.48	90.58	90.48	0.81



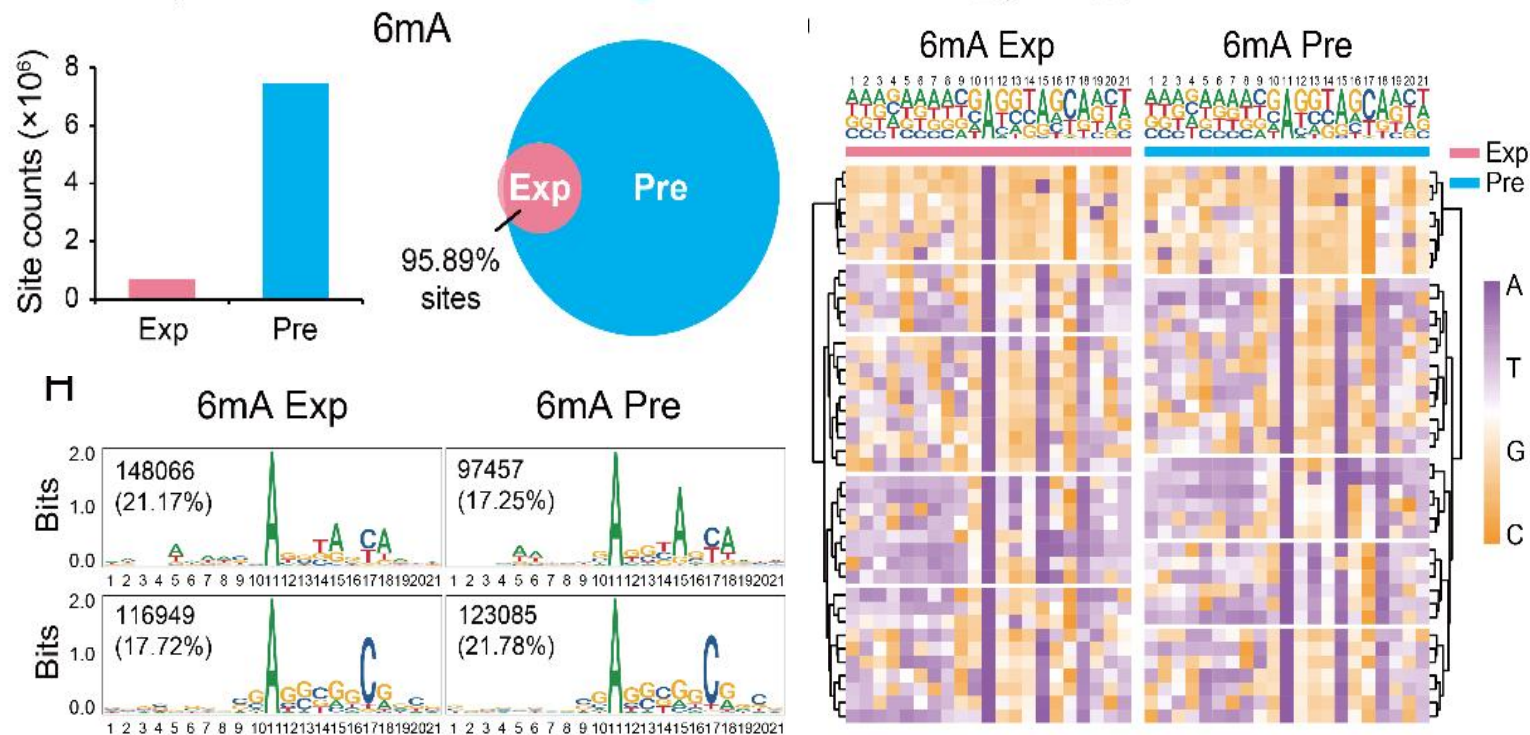
- 预测准确率达到95%
- 优于已发表的人工智能预测模型
- 表观修饰大数据集上准确率最高的预测模型

2、表观遗传智能预测模型SMEP

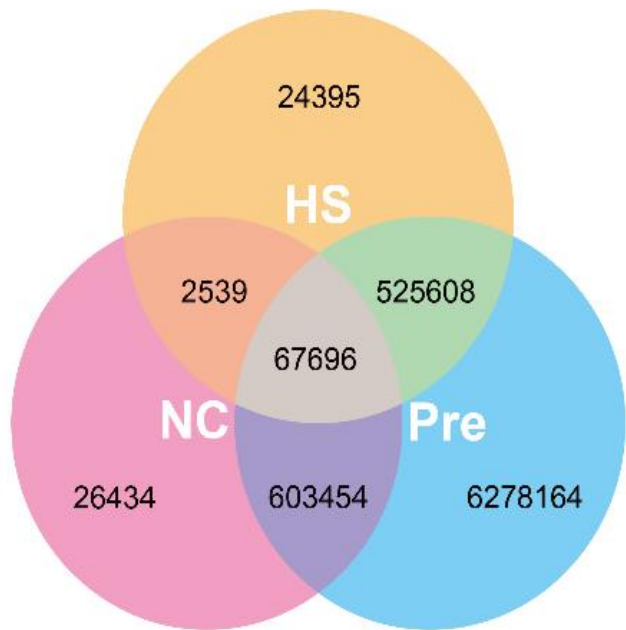


表观遗传修饰深度模型训练

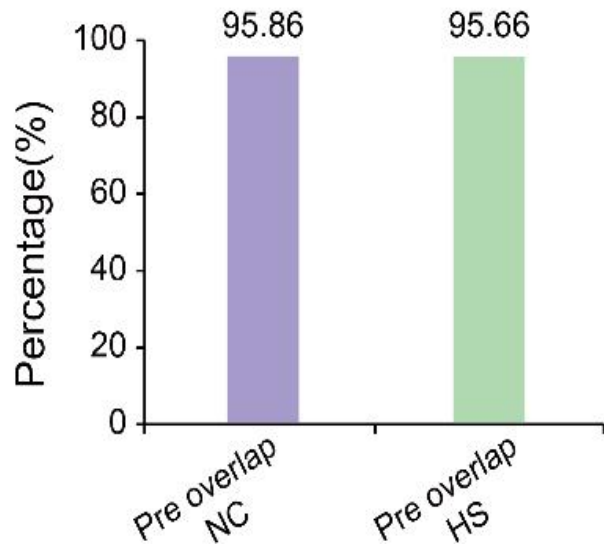
DNA 腺嘌呤预测模型



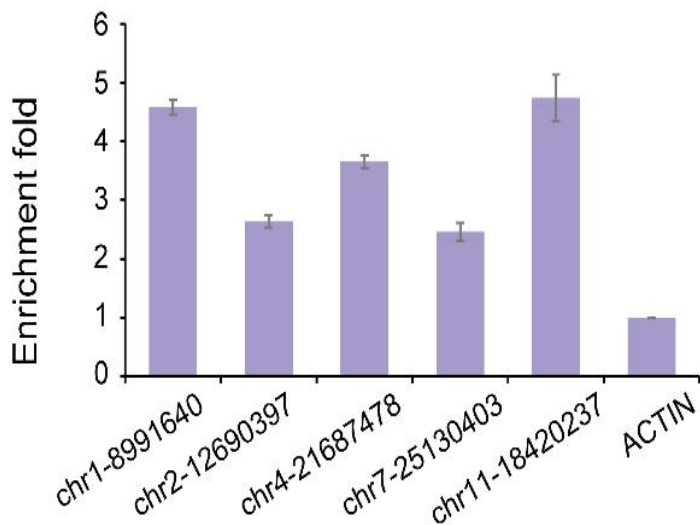
AI 预测位点和试验数据高度吻合



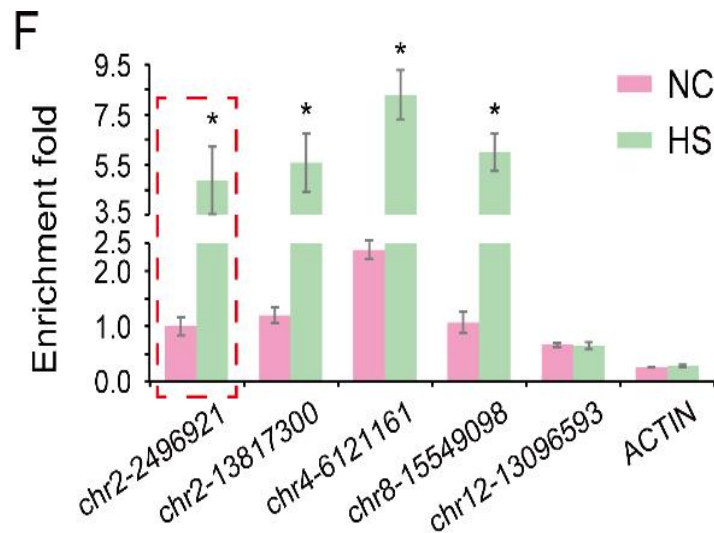
正常条件和热激条件



AI 预测位点的验证

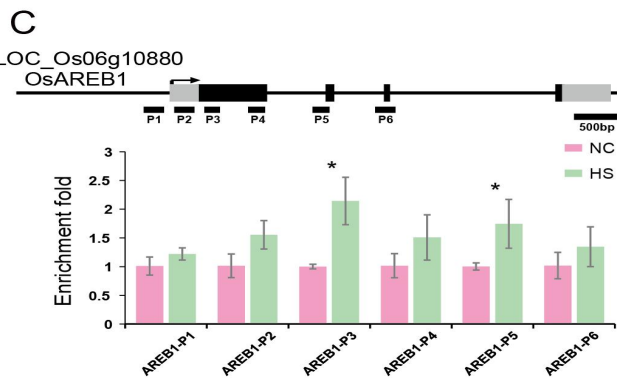
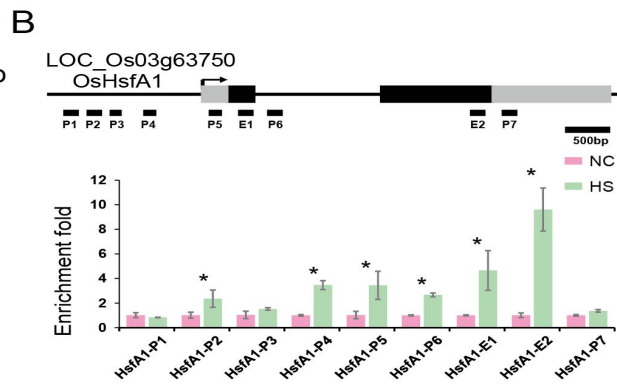
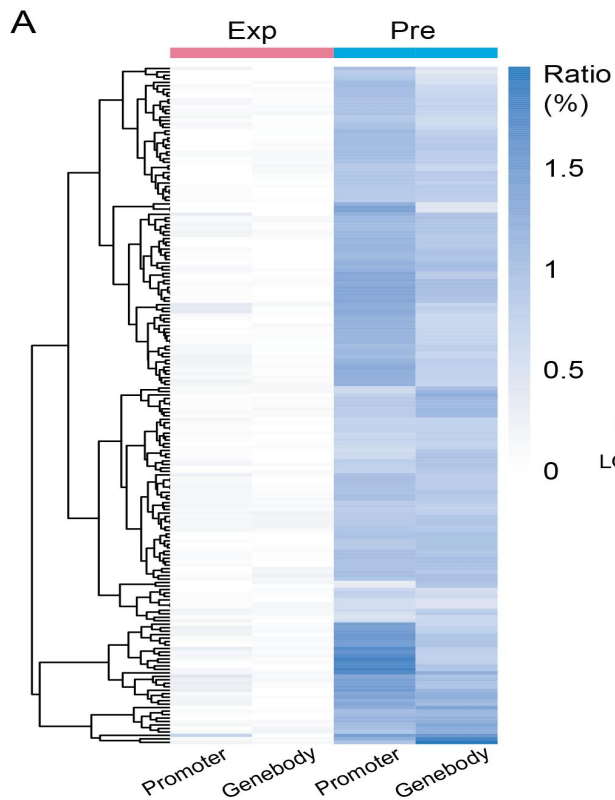


正常条件下预测位点验证



热激条件下6mA修饰含量增加

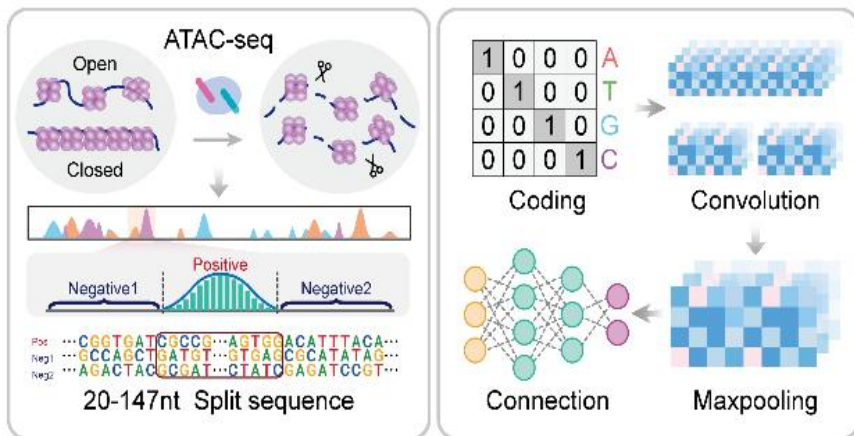
热激相关基因分析



3、作物三维空间结构预测模型SMOC

A

Feature data extraction and model training

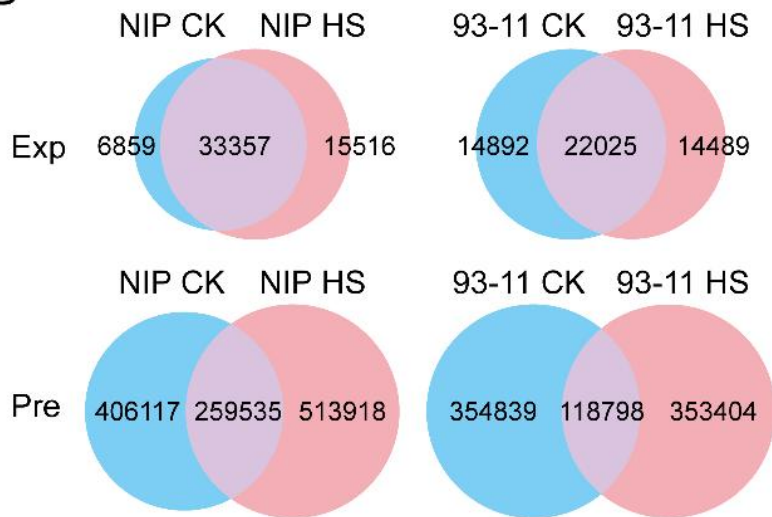


B^A

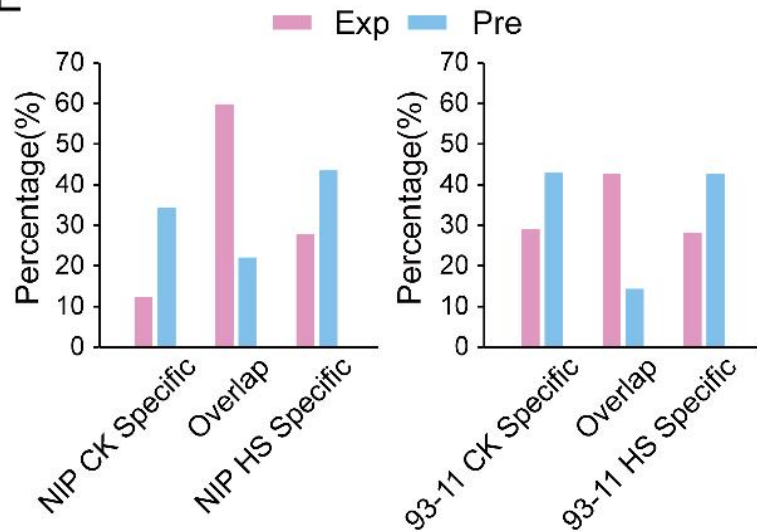
The screenshot shows the SMOC website interface. The header includes 'HOME', 'Japonica group cultivar Nip', 'Indica group cultivar 93-11', 'JBrowse', and 'Download'. The main content area is titled 'A Smart Model for Open Chromatin Region Prediction in Rice Genomes'. Below the title is a paragraph describing the model's purpose and architecture. At the bottom, there is a diagram comparing 'Normal Condition' and 'Heat Shock' for two rice cultivars, 'NIP' and '93-11'. The diagram shows that under heat shock, chromatin transitions from an 'Open' state to a 'Closed' state, which is predicted by the SMOC model.

预测了在不同环境下三维空间结构的变化

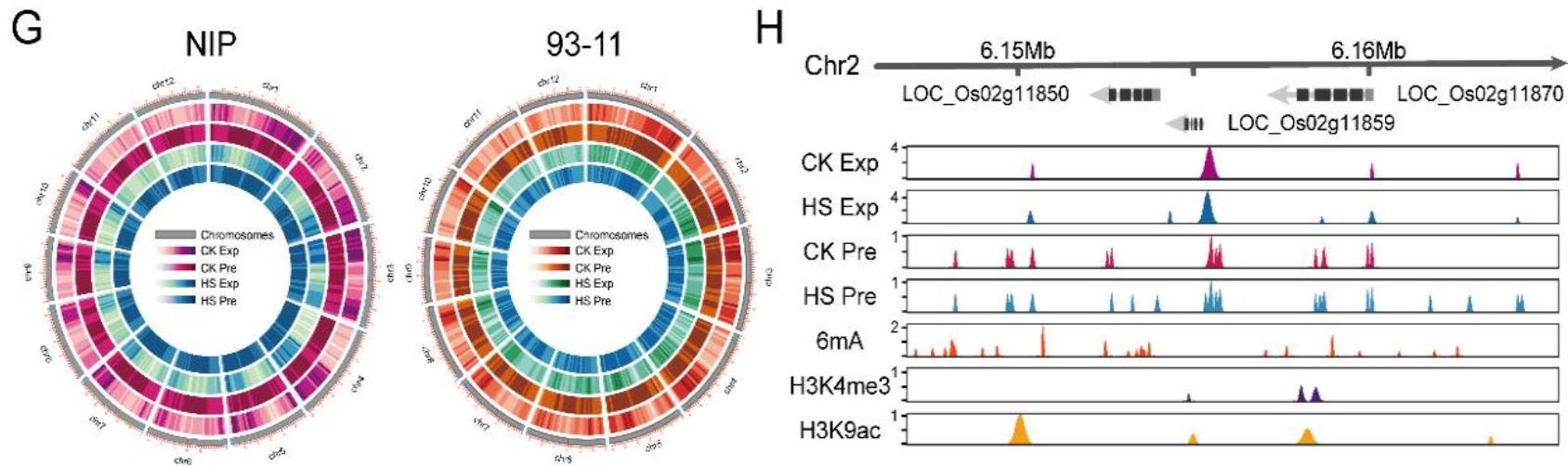
D



E

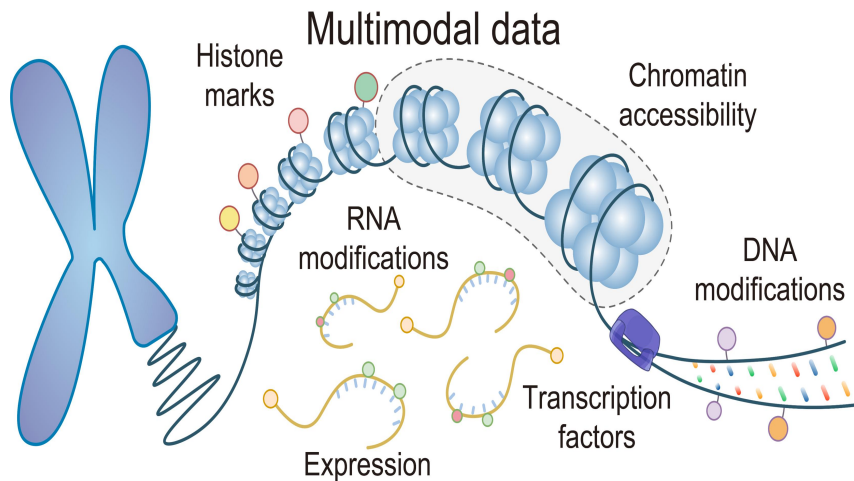


三维空间结构和表观遗传修饰关系

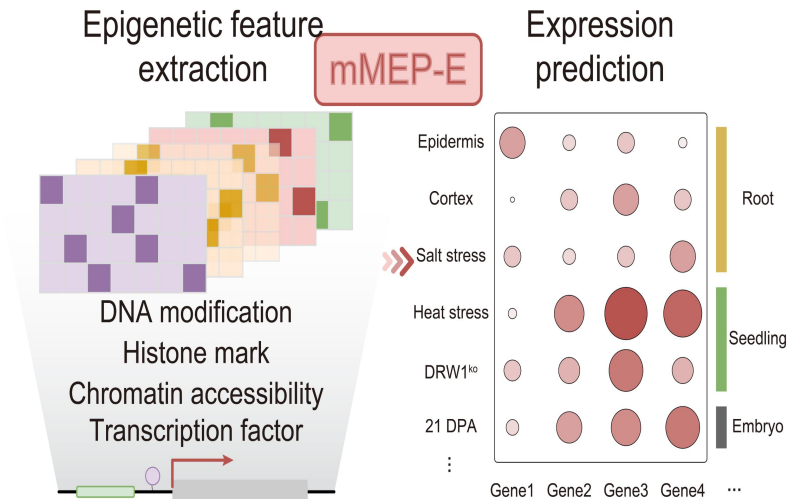


4、整合表观遗传修饰和三维空间结构的模型

mMEP多模态数据集



mMEP整合遗传与表观信息

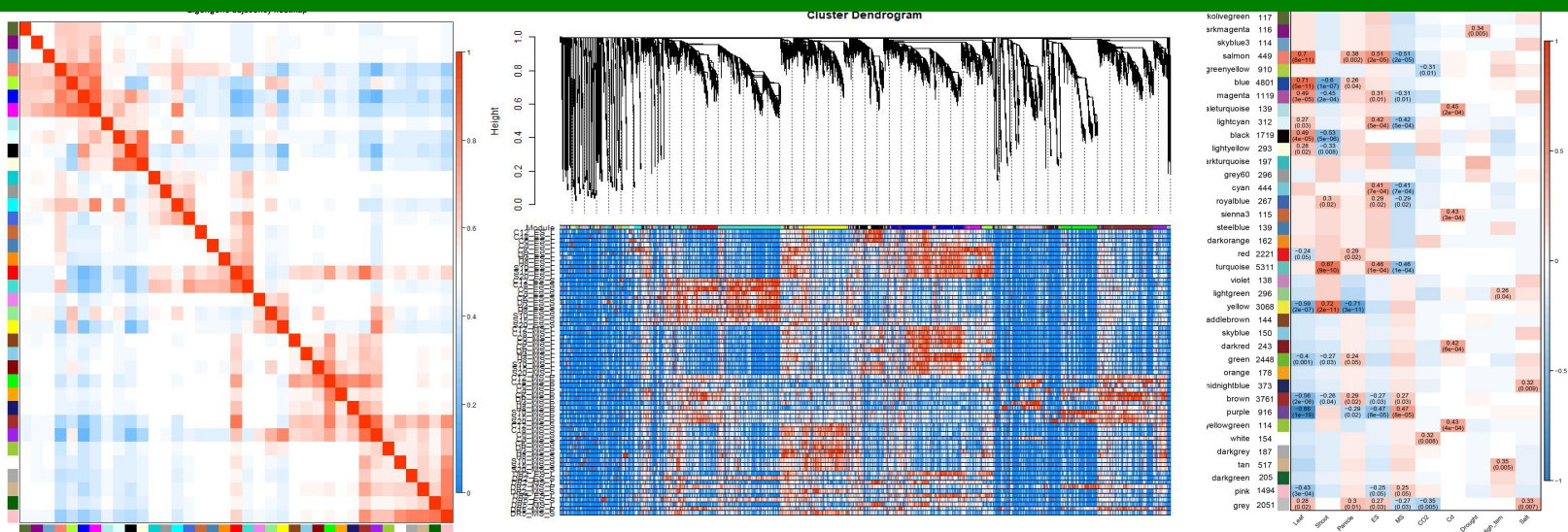


- 不同组织、细胞类型、胁迫处理、生长发育、突变体的多维表观组学数据

- DNA修饰、组蛋白修饰、转录因子、染色质可及性等多个维度的信息，进行精准预测和设计

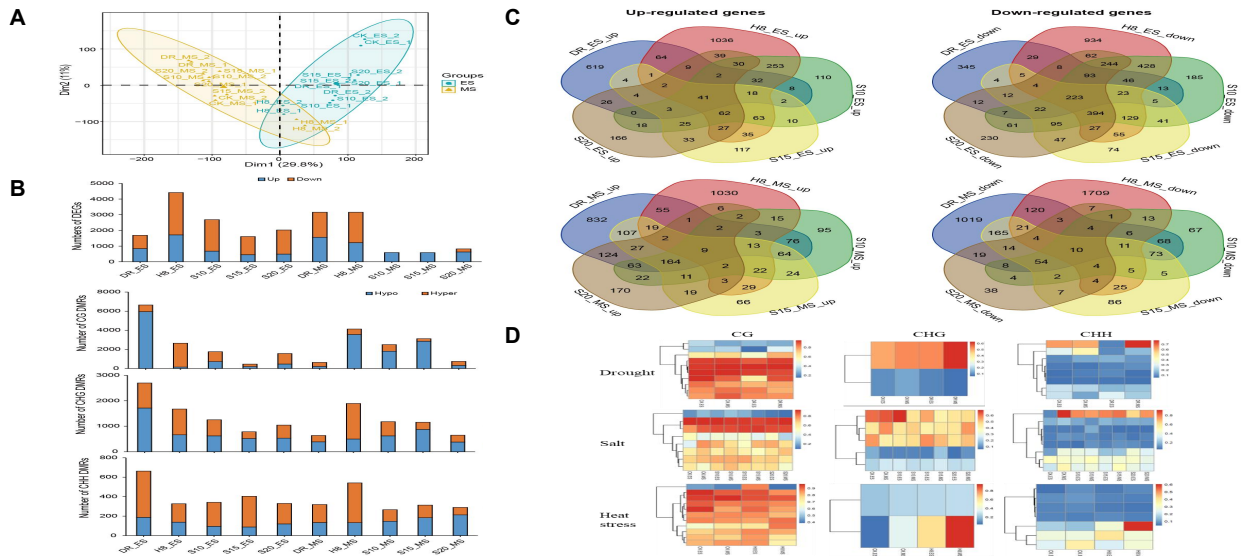
Unpublished data

水稻多元环境条件下基因共表达网络



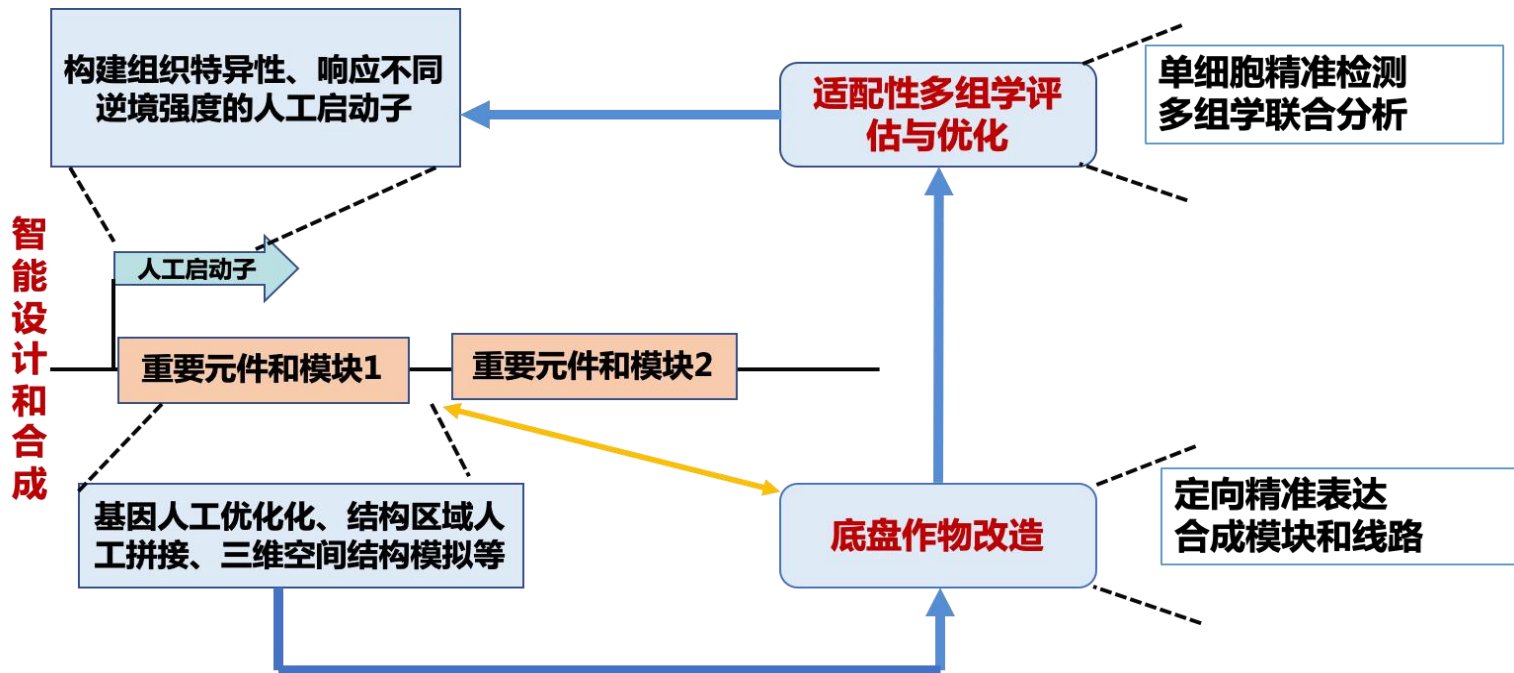
- 构建了具有33448 Nodes、28788847 Edges和38 Modules的基因共表达网络
- 筛选了具有时期特异性和胁迫表型特异性的Modules，根据特征重要性确定了系列基因集合

水稻多元环境条件下DNA甲基化动态调控调基因表达网络



- 构建了60439 Nodes、215682117 Edges和50 Modules多组学整合网络
- 确定了影响关键表型变化的DNA 5mC--基因表达网络。

三、水稻表观遗传回路设计和合成



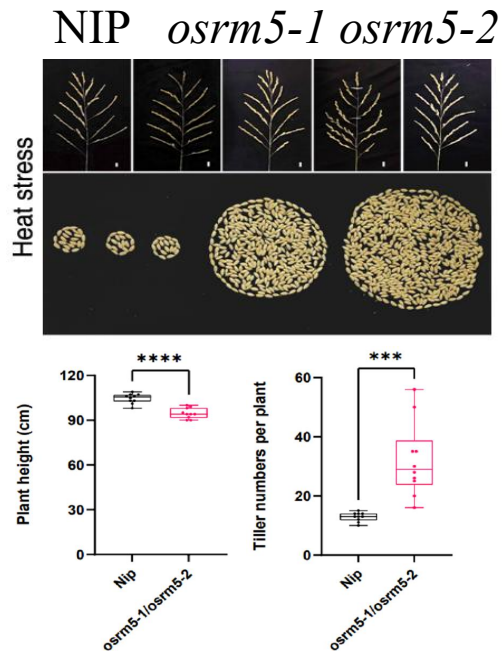
(1) 重要育种价值基因聚合创制协同改良新种质

● 创制产量和耐高温显著提升的协同改良新种质



NIP *ospon1* *ospon2*

PON2 定点编辑新材料单产增加15%以上，耐热性增强2.5倍

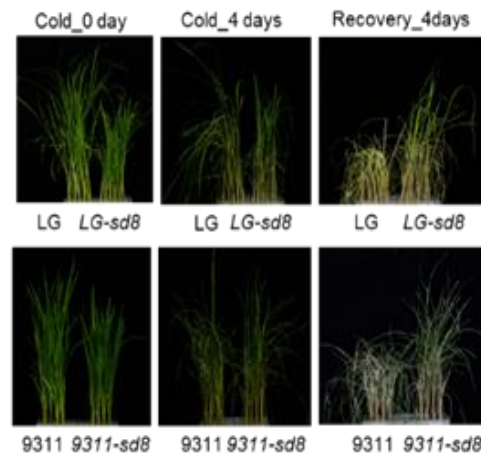
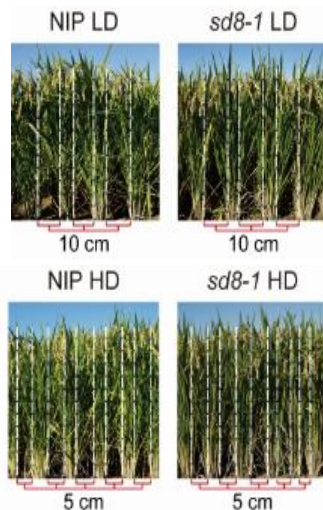
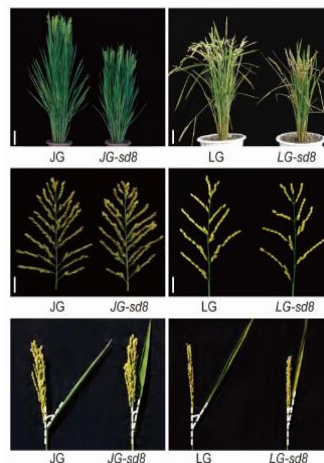
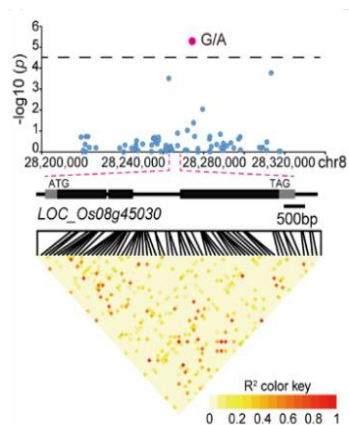


基因聚合新材料高温与产量协同改良，分蘖增加约40%

Unpublished data

(1) 重要育种价值基因聚合创制协同改良新种质

- **耐密植、耐低温与产量**协同改良的SD8籼稻和粳稻新材料，籼稻和粳稻主栽品种小区密植产量提升**10%以上**



全基因组关联分析鉴定到SD8新位点

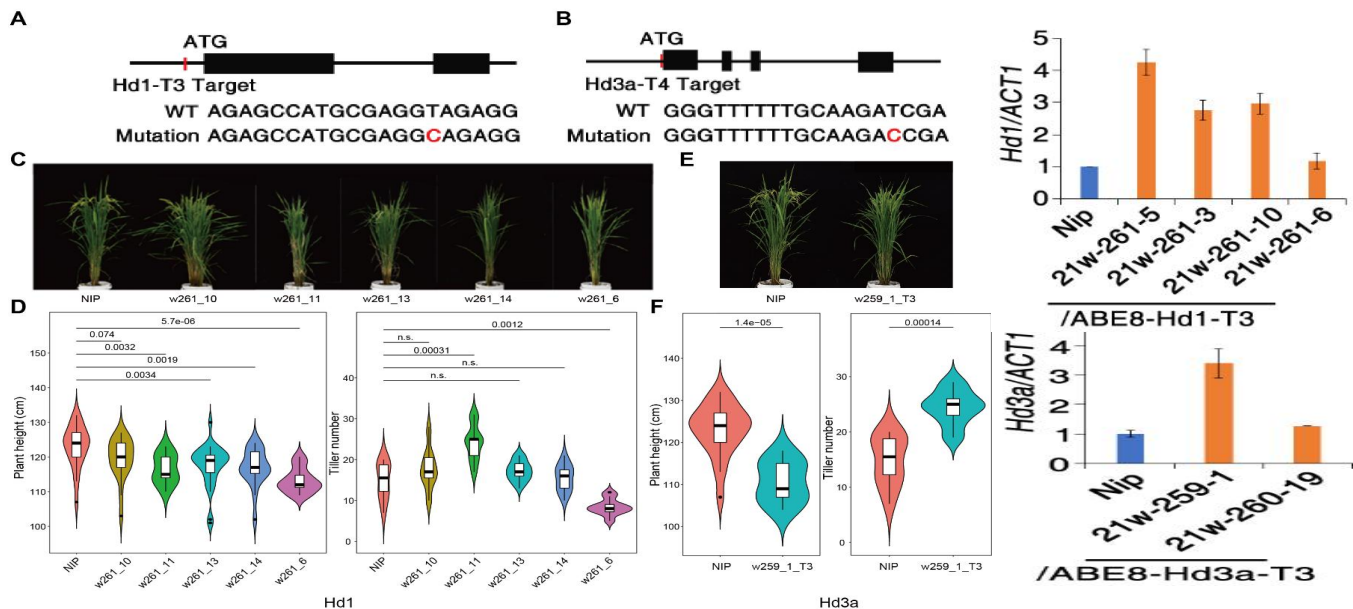
6个籼粳主栽品种株高显著降低，株型改良

密植条件下，新材料产量提升**10%以上**

主栽品种背景下具有显著耐冷性表型

(2) 精准设计表观修饰位点创制新种质

- 靶向设计合成DNA 6mA位点，创制了16份表观编辑水稻新材料
- *Hd1*和*Hd3a*启动子上6mA位点数量增加促使基因表达上调



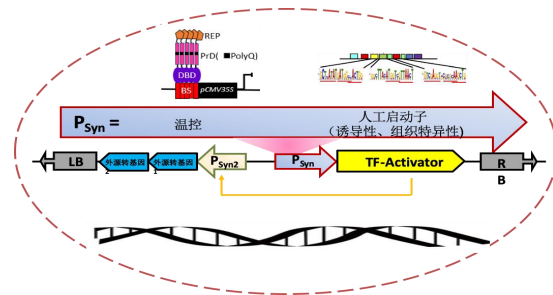
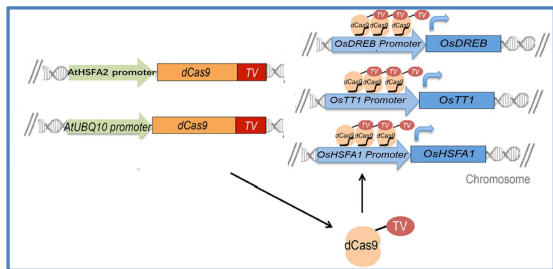
精准编辑DNA甲基化修饰位点调控株高和分蘖数

Unpublished data

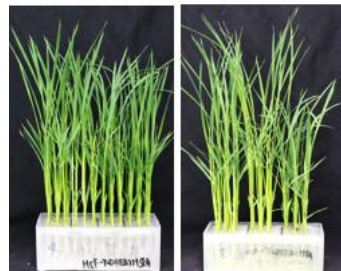
(3) 设计合成回路创制新材料

转录激活 (热激基因): *HSFA1*、*DREB2A*等

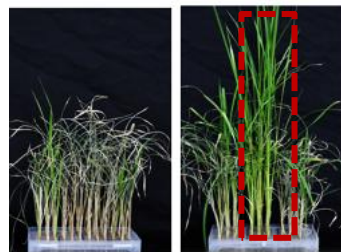
抗逆合成回路: *dosH*、*dwhy*、*IrrE*等耐逆基因



高温胁迫
处理前



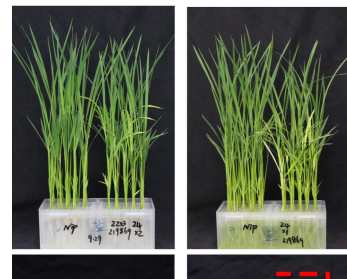
处理后恢
复7d



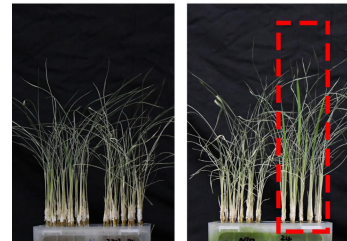
Nip HSF-T2DREB-T1HSFA

Nip HSF-T2DREB-T1HSFA

盐胁迫
处理前



处理后恢
复培养



Nip 3-1dosH-1

Nip 3-1dosH-1

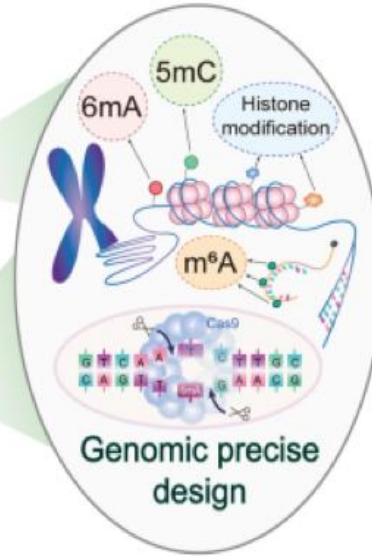
150 mm NaCl treatment

Unpublished data

Application in plants



Smart crop



Hybrid vigor

Yield & quality

Biotic and abiotic stresses

Photosynthesis architectures

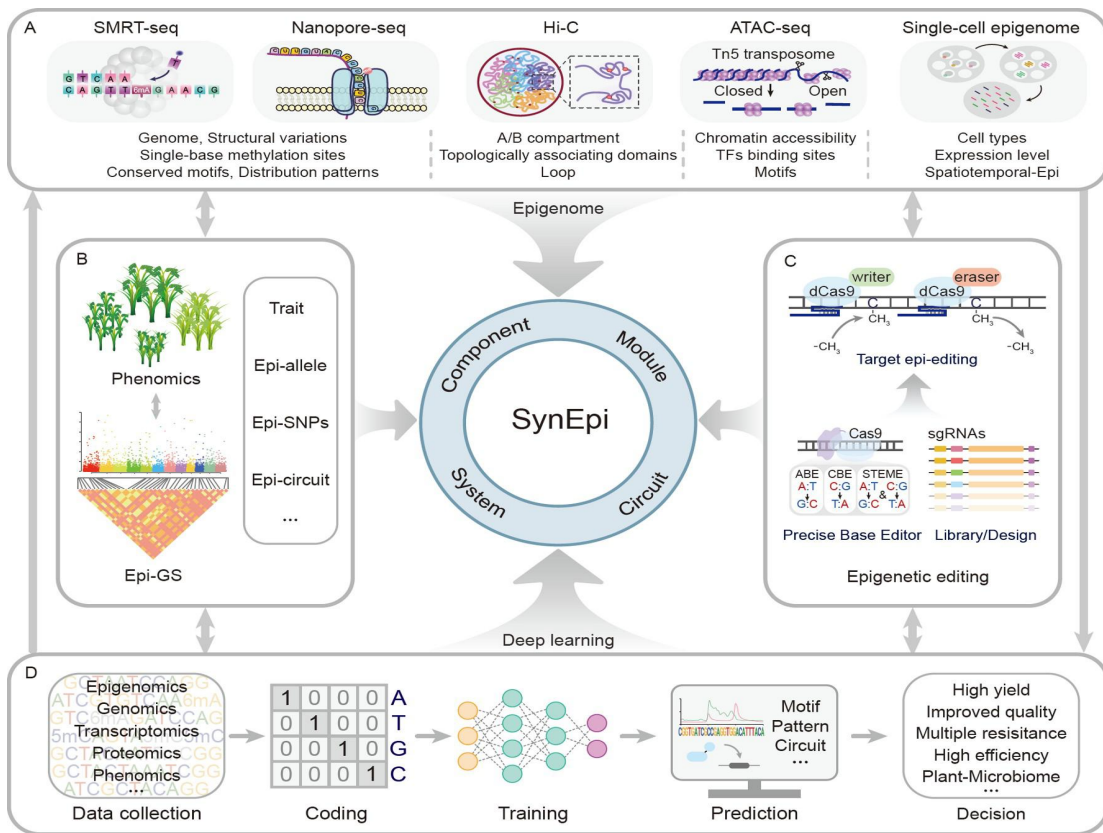
Plant-microbiome interactions

.....

SMART Rice

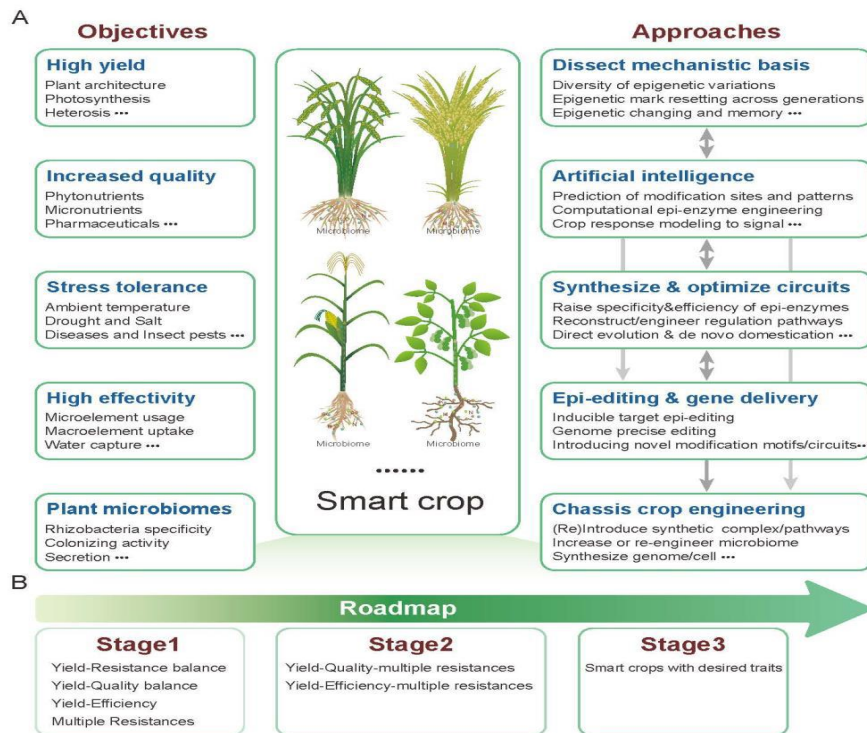
Self-Monitoring, Adapted and Responsive Technology

SMRT Crop 创制途径和路线图



SMRT Crop 创制途径和路线图

Plant synthetic epigenomic engineering for crop improvement



致谢

Lab members

Qin Liu

Dan Feng

Qian Zhang

Pingxian Zhang

Yifan Wang

Xiulan Li

Zhuoying Du

Guihua Hu

Collaborators

Dr. Jan U. Lohmann, Heidelberg University

Dr Zhe Liang , BRI, CAAS

Dr. Li Pu, BRI, CAAS

Dr. Jian Tian , BRI, CAAS

Dr. Tiegang Lu BRI, CAAS

Dr. Yuehui He PKU

Dr. Yu Hao NUS



中华人民共和国科学技术部
Ministry of Science and Technology of the People's Republic of China



谢谢!

