National key basic research program (973 project)  
(2010CB126600, 2010-2014. 5.09M USD)

The Basic Research of Cassava  
Important Tropical Crop for the World

Ming Peng  
Institute of Tropical Bioscience & Biotechnology, Chinese Academy of Tropical Agricultural Sciences  
Jan 22, 2016  Nanning, China
Cassava is an important crop in the world

- The sixth important food crop, known as “the king of starch crop”, staple food for over 700 million people.

- Resources for feed, bioenergy and industrial materials
  (Market share in China: industry starch 10%, fuel ethanol 40%)

Total production of the top six food crops in 2014

- Wheat: 80 M Ton
- Rice: 60 M Ton
- Maize: 60 M Ton
- Potato: 40 M Ton
- Cassava: 260 M Ton
- Soybean: 20 M Ton

FAO Stat 2014
Cassava in Tropic China

Area, 0.48 M km²
8 provinces
200 counties
36 nationalities
170 Million population

The bottleneck of local tropical plantation: interval drought, barren soil under the huge yield potential
Biological features of Cassava —— to be a tropical model plant

- High efficiency of photosynthesis: C3-C4, high up to 45-50 µmolCO\textsubscript{2}s\textsuperscript{-1}·cm\textsuperscript{-2} (Pn)
- High biomass productivity: Starch yielding to 22.5T/ha, Harvest index: 0.50-0.55.
- Drought tolerance, High Water-Use-Efficiency (WUE).
- Tolerance to barren or Nitrogen/Phosphates deficient soil.
- Grown in tropics: prefer to high solar radiation, high temperature, 28-35 °C, and rainy environment.
- Starch characters: Storage root with 30-32% starch content, average amylose 22.4%, amylopectin 77.6%.
Challenges of breeding technology and field production of cassava

- High heterozygosity, genetic loads, and low genetic advantage.
- High yield, higher starch content, tolerance to barren soil, low temperature, drought, PPD, and satisfy starch quality.
- Food, feed, bioenergy and advanced bio-materials.
Two major scientific concerns refined

- Determination of the main pathways for starch efficient accumulation and yield.
- Elucidation of the molecular mechanism of tolerance to abiostress/biostress and the relationship with biomass production.
Research program

Starch biosynthesis and yield regulation  
**Project 1**

Mechanism of abiotic stress tolerance  
**Project 2**

Comparative genomics in Eurphorbiaceae  
**Project 3**

GenBank

Genome annotation and integration  
**Project 4**

Genome sequencing: KU50, CAS36, W14

Integration of technology and germplasm

Validation of key gene function and application  
**Project 5**

Integrated breeding and germplasm enhancement  
**Project 6**

Output of theory and technologies for improvement of high yield and abiotic stress of tropical crops
Whole genome sequencing and comparative annotation of wild ancestor and cultivated varieties

KU50 (Manihot esculenta Crantz) successful variety

W14 (Manihot esculenta ssp. flabellifolia) ancestor sub species
PEPC/RuBP activities of cultivated varieties are higher than those of the wild ancestor.
Chromosome (2n=36) and repeats

Evolution from wild ancestor to cultivars

Mining of millions of SNV/Indel

Comparison of genomes
Description of cassava draft genomes

- Over 96% gene region coverage.
- 34,000 genes for W14 and 38,000 genes for KU50 were annotated.
- Genome heterozygosity self SNV: W14 3.9%, KU50 3.5%, AM560 1.44%.
- Up to 65% sequences are estimated as repeated sequences, including 34% annotated, and about 30% are unassembled;
- Non-coding RNA annotated: 143 miRNA, 861 tRNA and 337 rRNA.
The unique genes (PAV) and the extremely selected genes (Ka/Ks) from wild ancestor to the cultivar W14 KU50+AM560

Gene Models

Highly selected genes (3254)
Major biological processes involved in PAV and CNV

Present & Absent Variation (PAV)

Copy Number Variation (CNV)
GO ways of genes have been highly selected (W14 vs KU50/AM560)

- **a** Cellular process
- **b** Metabolic process
- **c** Cell & Cell part
- **d** Developmental process
- **e** Biological regulation
- **d** Response to stimulus
Summary: there were 9 major important biological processes took place predominantly during the evolution and domestication of cassava

- Cellular process (PAV, Ka/Ks)
- Metabolic process (PAV, Ka/Ks)
- Cell & Cell part (PAV, CNV, Ka/Ks)
- Catalytic activity (PAV, CNV)
- Binding (PAV, CNV)
- Developmental process (Ka/Ks)
- Biological regulation (Ka/Ks)
- Response to stimulus (Ka/Ks)
- Transferase activity (CNV)
RNA Sequencing and the comparative transcriptomic annotation

Genotypes: KU50, Arg7, SC124 and W14

Materials: Leaves, stem and storage roots in the developed stage of 90d, 150d and 240d at AM 10:00 or under treatments of phloem transport block and drought.

RNASeq: 64 samples; smallRNA Seq: 6 samples.
Comparative transcriptome revealed genes predominantly expressed in leaves of cultivars (KU50 vs W14)
Expression of genes for photosynthesis and carbon assimilation increased.
Enrichment of genes respond to multiple environmental factors in domesticated leaf

18/131: HSPs (HSP20, HSP21, HSP70B)
Enrichment of genes for ABA metabolism in storage root of domesticated cassava

4/11: ZEP; nine-cis-epoxycarotenoid dioxygenase; abscisic aldehyde oxidase; NAD(P)-binding Rossmann-fold superfamily protein
Enrichment of genes for cell and cellparts in storage root of domesticated cassava

140 (51.1%)
The gene network involved in starch synthesis in the storage root of cassava
Expression of genes for cell wall synthesis and the second metabolism predominantly downregulated.

(Arg7/KU50 vs W14)
miRNAs play a role in regulation of starch biosynthesis of cassava

The targets of miRNAs were found in promoter region of SuSy
A carbon flux and starch efficient accumulation model in cassava

Cassava genome from wild ancestor to cultivated varieties, Nature Comm. 2014
An apoplastic phloem sucrose loading model in leaves of cassava

- Sucrose is the main transport form of carbohydrates in phloem.
- A sucrose apoplastic loading model in leaves of cassava is drawn on the results of C\(^{14}\) tracing and SUTs expression analysis.
- It is similar to that of grasses and other short life cycle plants, which is adaptable to flexible leaves dropping under drought stresses.
MeSUT1–MeSUT4 structure

MeSUT1x
Structure no. 1
Segments included: 1 2 3 4 5 6 7 8 9 10 11 12

MeSUT2
Structure no. 1
Segments included: 1 2 3 5 6 7 8 9 10 11 12

MeSUT4
Structure no. 1
Segments included: 1 2 3 4 5 6 7 8 9 10 11 12
A symplasmic carbohydrate unloading model is correlated with efficient starch accumulation in cassava

1. Sucrose unloading from phloem is a symplasmic model rely on plasmodesmata.

2. There is a shift from apoplastic to symplasmic at the stage of storage root forming and starch accumulation beginning.

3. This structurized unloading mode will benefit the efficient starch accumulation.

Kun Pan et al., A carbohydrate symplasmic unloading model correlated to efficient starch accumulation in cassava (*Manihot esculenta* Crantz), (Scientific Reports, 2016, under review)
The plasmodesmatas in phloem of roots

<table>
<thead>
<tr>
<th></th>
<th>SE-CC</th>
<th>SE-CC</th>
<th>PP-PP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Num. / μm)</td>
<td>CC/PP</td>
<td>CC/PP</td>
<td>CC/PP</td>
</tr>
<tr>
<td>初生须根</td>
<td>0.013</td>
<td>0.007</td>
<td>0.030</td>
</tr>
<tr>
<td>次生须根</td>
<td>0.008</td>
<td>0.001</td>
<td>0.049</td>
</tr>
<tr>
<td>伸长期块根</td>
<td>0.033</td>
<td>0.022</td>
<td>0.147</td>
</tr>
<tr>
<td>膨大期块根</td>
<td>0.040</td>
<td>0.027</td>
<td>0.095</td>
</tr>
<tr>
<td>成熟期块根</td>
<td>0.032</td>
<td>0.021</td>
<td>0.100</td>
</tr>
</tbody>
</table>

SE, sieve element; CC, companion cell; PP, phloem parenchyma
Cassava response to drought treatments

Light: 40000-50000LX
Temperature: 30-36°C
Moisture in soil: 70%-75%
Two kinds of phenotype in response to drought

1) The growth slow down (SC124);
2) The leaves abscission rapidly (Arg7);
Proline-polyamine-ROS-ethylene mechanism regulates leaf abscission under drought.
A glutathione (GSH) mediated drought tolerance mechanism

Glutaredoxin (Grx)

Expression of Grx785 and Grx058 are induced under drought
Grx785 could be self activated in yeast cell nucleus, but Grx058 not
The ALWL-motif determines function difference between Grx058 and Grx785.

ProADH1::NLS::GAL4 DNA-BD::Grx785-P65 mutant::TerADH1 (P65→L65)
ProADH1::NLS::GAL4 DNA-BD::Grx785-G75 mutant::TerADH1 (G75→D75)
GM cassava plants with Grx785-RNAi and Grx058-OE appeared better tolerance to drought.

Grx785 over expressed in cassava sensitive to ABA and other osmotic stresses, and the plant growth was depressed.

Grx785-RNAi, and Grx058-OE or Grx058 RNAi in cassava are insensitive to ABA and other osmotic stresses.
Both Grx785 and Grx058 can interact with three TGA transcription factors in cassava

Grx785 interacted with TGA by double fluorescence molecular verification

<table>
<thead>
<tr>
<th>Bait</th>
<th>Prey</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>BD: Grx785 mP65</td>
<td>AD: TGA074</td>
<td>+</td>
</tr>
<tr>
<td>BD: Grx785 mP65</td>
<td>AD: TGA304</td>
<td>+</td>
</tr>
<tr>
<td>BD: Grx785 mP65</td>
<td>AD: TGA351</td>
<td>+</td>
</tr>
<tr>
<td>BD: Grx058</td>
<td>AD: TGA074</td>
<td>+</td>
</tr>
<tr>
<td>BD: Grx058</td>
<td>AD: TGA304</td>
<td>+</td>
</tr>
<tr>
<td>BD: Grx058</td>
<td>AD: TGA351</td>
<td>+</td>
</tr>
<tr>
<td>BD: Grx058</td>
<td>AD: TGA813</td>
<td>-</td>
</tr>
</tbody>
</table>

Interaction between Grx058 and TGA verified by double fluorescence molecular
TGA351 has no transcription activity in yeast
Grx785 is necessary for the transcription function of TGA351 in yeast.
Competitive binding of Grx785/Grx058 with TGA regulate a water deficient response action

A: When Grx058 instead of Grx785 binds to TGA, TGA will be inactive state, the plant growth will continue.

B: While Grx785 predominantly binds to TGA which will be active state, the plant growth to be ceased.
GSH mediated a drought signal transduction pathway via interaction between Grx785 and TGA transcription factor.

- Drought induced expression of Grx785.
- Grx785 protein associated with GSH activated TGA in nucleus.
# Identification of drought responsive miRNAs

## Drought and cold related miRNAs in cassava

<table>
<thead>
<tr>
<th>miRNAs</th>
<th>Drought</th>
<th>Cold</th>
<th>Drought and Cold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novel31, novel18, miR164,</td>
<td>miR319, miR477b, nove16</td>
<td>Novel4, novel42, miR52, miR395</td>
<td></td>
</tr>
<tr>
<td>miR169, miR393, miR394</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Putative Target mRNAs of miRNA novel 4

<table>
<thead>
<tr>
<th>miRNA</th>
<th>target ID</th>
<th>score</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>novel-4</td>
<td>cassava4.1_029179m</td>
<td>4.5</td>
<td>GATA type zinc finger transcription factor family protein light-responsive</td>
</tr>
<tr>
<td>novel-4</td>
<td>cassava4.1_016357m</td>
<td>4.5</td>
<td>Peptidase M20/M25/M40 family protein kinase metabolic process, hydrolase</td>
</tr>
<tr>
<td>novel-4</td>
<td>cassava4.1_007251m</td>
<td>5</td>
<td>Peptidase M20/M25/M40 family protein kinase metabolic process, hydrolase</td>
</tr>
<tr>
<td>novel-4</td>
<td>cassava4.1_013753m</td>
<td>5</td>
<td>5CKB1 casein kinase II beta chain 1 protein kinase metabolic process</td>
</tr>
<tr>
<td>novel-4</td>
<td>cassava4.1_000724m</td>
<td>5</td>
<td>Di-glucose binding protein with Kinesin motor domain long-range intracellular transport</td>
</tr>
<tr>
<td>novel-4</td>
<td>cassava4.1_011709m</td>
<td>5</td>
<td>5BEH4 BES1/BZR1 homolog 4 Brassinosteroid signalling lipid trafficking ABC</td>
</tr>
<tr>
<td>novel-4</td>
<td>cassava4.1_013303m</td>
<td>5</td>
<td>5TGD2 trigalactosyldiacylglycerol2 lipid trafficking</td>
</tr>
<tr>
<td>novel-4</td>
<td>cassava4.1_009577m</td>
<td>5</td>
<td>5TGD2 trigalactosyldiacylglycerol2 ABC</td>
</tr>
</tbody>
</table>
The expression level of miRNA Novel 4 is negatively related to that of its target mRNA.
Trehalose get involved in osmotic regulation response to drought stress in cassava

Higher trehalose has been measured in leaves of all 39 cassava genotypes range from 0.23 to 1.3 mg.g\(^{-1}\)FW, that is beyond MeTPS1-3 OE tobacco and Arabidopsis lines (0.01-0.03mg.g\(^{-1}\)FW).

Trehalose content positively correlated to water holding ability in cassava leaves.
Osmotic stress induced expression of genes in trehalose synthesis in cassava (SC124)

Trehalose content could be induced up to 230-550% in cassava.

MeTPS1-3 OE tobacco lines exhibited stronger drought tolerance 1: Wild type; 2, 3: GM lines

Fu et al. Crop Science, 2014
Improved transgenic protocol used for gene function validation and GM breeding in cassava

Embryogenic Suspension → Somatic Embryogenesis → Shoot Organogenesis

Yielding rate above 20%
Transgenic events of MeSOD and proteomics unlocked new mechanism of PPD in cassava

Transgenic cassava with MeSOD (Cu/Zn) and MeAPX2 obtained tolerance to PPD

Ca\(^{2+}\)-CaM join the controlling of metabolic PPD
Waxy mutants created by transgenic

A series of transgenetic lines with MeGBSSii RNAi, the amylose content in storage root range from 0-12%.

(Zhao et al., Biotechnology & Bioengineering, 2011, 108: 1925-1935)
Database: Cassava-genome.cn

>> Browse cassava genome
Input format: (1) track:start-end, like scf20036:75000-196000. (2) scf20036:start+length, like scf20036:75000+121000. (3) scfId, like scf20036.

>> BLAST against cassava database
Here: you can input query nucleic acid sequences to find similar sequences in the cassava database.

>> Browse cassava database
In cassava database, the latest data is available, including nucleic acid sequences, genome annotation etc.

>> Plant microRNA function database
miRFDB is an online database that aims at exploring the regulatory function of plant miRNAs. A large number of expression datasets of both miRNAs and their target genes under various experimental conditions, gene function annotations and pathways are integrated into miRFDB.

>> Information
- About the database
- 073 Project Website
Germplasm enhancement and integrated breeding

- New accessions: Wild ancestor (W14), Dwarf species, Sugary mutant, Waxy, tolerance to PPD, cold tolerant CH12, and H1426.
Sugary mutant CAS36

High yield, resistant to wind, H1426

Cold tolerant F224

Tolerant to PPD, OZ1
Released new varieties: SC12 and NS 048

SC12: fresh root yield 40.05t/ha, starch content of storage root, 31.11%, HCN, 42.20mg/Kg.
General achievements of the 973 program (2010-2014)

- Publications: total of 197 articles, including over 100 SCI papers, published in such as Nature Biotechnology, Nature Comuni. NAR, Plant Physiology, Scientific Reports, J EXP. B., BMC plant biology, BMC genomics.

- Patents, authorized 12

- Scientific and Technology Awards: 7 items

- National cassava team: about 100 for basic research, other 100 for R&D.

- Training: 2 outstanding young scientists, 5 professors, 12 associated professors, 44 got Ph.D, 94 got Master degree and 10 completed Post doc.
International collaboration

- NSFC-CG joint foundations (4):
  - Whole genome association analysis modelling genomic based breeding in cassava (Wenquan Wang, 2013-2017, 2.39 M RMB)
  - Molecular evaluation of cassava germplasm by transcriptomic and proteomic methods (Songbi Cheng, 2014-2018, 2.3M RMB)
  - Molecular breeding of CBSV resistant cassava by gene editing with eIF4E (Zhixi Liu, 2015-2019, 2.3M RMB)
  - The molecular regulation of Cold tolerance in cassava (Ming Peng, 2016-2020, 2.6M RMB)

- International cooperation projects involved in cassava (2008-2016, 5 items, there are about 12 M RMB)
Acknowledgements

- Institute of Tropical Biosciences & Biotechnology, CATAS
- Tropical Crop Genetic Resources Institute, CATAS
- Shanghai Academy of Life Sciences, CAS
- Beijing Genome Institute, CAS
- South China Botany Garden, CAS
- Fudan University
- Huzhong Agriculture University
- Guangxi University
木薯碳硫分配及光合产物高效转运与淀粉积累模型

栽培/野生种基因组、大量转录组数据比较及重要基因的生物学验证，提出：

 木薯光合产物装载与卸载方式；
 碳硫分配与淀粉积累的模型。