Front. Agr. Sci. Eng. 2016, **3**(4): 308–320
DOI: 10.15302/J-FASE-2016113
Available online at http://engineering.cae.cn

RESEARCH ARTICLE

Annotation and validation of genes involved in photosynthesis and starch synthesis from a *Manihot* full-length cDNA library

Yang ZHANG[1*], Xin CHEN[1*], Haiyan WANG[1*], Zhiqiang XIA[1], Peng LING[2], Wenquan WANG (✉)[1]

1 Institute of Tropical Bioscience and Biotechnology, Chinese Academy of Tropical Agricultural Sciences, Haikou 571101, China
2 Department of Horticultural Sciences, Citrus Research and Education Center, University of Florida, Lake Alfred, FL 33850, USA

Abstract A full-length cDNA library from leaf and root tissues of cassava (*Manihot esculenta*) Arg7 and one accession of its wild ancestor W14 (*M. esculenta* ssp. *flabellifolia*) has been constructed. The library is comprised of four sub-libraries, containing 32640 recombinant clones, 6028 cDNA clones from their 5′ ends, and 128 clones from the 3′ ends were sequenced. In total, 5013 high-quality expressed sequence tags (ESTs) and 1259 unigenes were obtained. Of these, 746 unigenes were identified by their sequence homologies to ESTs from model plants, and 323 unigenes were mapped onto 114 different KEGG pathways. From these, 24 differentially expressed genes involved in starch metabolism and photosynthesis were identified and five of them were selected to compare their expression level between Arg7 and W14. Notably, Arg7 has a higher net photosynthesis rate in leaves, higher ribulose-1,5-bisphosphate carboxylase oxygenase activities in leaves, and higher AGPase activity in roots. This resource is the first EST collection from wild cassava and should be of value for gene discovery, genome annotation and studies of *Manihot* evolution.

Keywords *Manihot esculenta*, expressed sequence tag, unigene, pathway, expression pattern

1 Introduction

Cassava (*Manihot esculenta* Crantz), an important tropical crop and model plant in the Euphorbiaceae, is an important source of carbohydrates for more than 700 million people, particularly in many developing countries in Africa, Asian and South America[1]. Its storage roots are rich in starch[2], adapted for survival in barren soil and under drought conditions, and relatively protected from herbivory owing to the presence of cyanogens[3,4]. With the current global crisis related to sustainable energy and climate change, interest in cassava has increased, both as a potential resource for bio-fuel to replace oil reserves[5,6] and for its productivity even on land poorly suited to agriculture. Cassava has an unusually high rate of photosynthetic carbon assimilation (CO_2, 43 $\mu mol \cdot m^{-2} \cdot s^{-1}$), as well as a high temperature optimum (45°C) for photosynthesis[7–9]. It is reported to have one of the highest rates of CO_2 assimilation into sucrose of any plant measured. Although classified as a C_3 plant, the leaves exhibit C_3–C_4 intermediate behavior under certain conditions, and also exhibit high activity of the C_4 enzyme phosphoenolpyruvate carboxylase (PEPC)[10]. The approximate composition of cassava starch on a dry weight basis is 0.24% ash, 0.13% fat, 0.49% protein, 0.15% crude fiber and 98.4% starch[11]. The molecular mechanisms that control growth and development of cassava, and the complex molecular networks that regulate photosynthesis and starch accumulation in the species, are still far from clear. Therefore, analysis of the cassava transcriptome, especially the expression of the genes related to starch metabolism and other desirable traits, is sorely needed. Large-scale analysis technology, such as the construction of full-length cDNA libraries and the genome-wide sequencing of EST collections, are increasingly being applied to identify related genes and/or pathways. Three full-length cDNA libraries that focus on cassava genes related to tolerance and resistance have already been constructed[12–14]. Recently, a cDNA library prepared from root tissue of cassava cv. Huanan 124 at the root bulking stage was constructed to identify genes that regulate starch metabolism[15]. Single-pass sequencing of 9600 cDNA clones from this library established a catalog of expressed sequence tags (ESTs) that were assembled into 2878 putative unigenes. All of these EST resources will help to clarify the mechanism of starch accumulation, gene discovery and genome annotation in cassava.

Received September 8, 2016; accepted October 8, 2016

Correspondence: wangwenquan@itbb.org.cn
*These authors contribute equally to the work

Several genes that regulate photosynthesis, carbon assimilation, carbon allocation, and starch synthesis have been identified in model plants. The first step in photosynthetic CO_2 assimilation and photorespiratory carbon oxidation is catalyzed by ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCO)[16], and there are high rates of consumption of NADP and NADH during photosynthesis and photorespiration, respectively[17]. The carboxylation of phosphoenolpyruvate to yield oxaloacetate and inorganic phosphate is catalyzed by PEPC (EC4.1. 1.31)[10,18,19]. Soluble acid invertase (SAI, EC3.2.1.26) has been suggested to be a key regulator of the accumulation of sucrose, and UDPglucose pyrophosphorylase (UGPase, EC2.7.7.9) converts UDPglucose to glucose-1-phosphate in storage roots[20,21]. There are two possible pathways for cleaving cytosolic sucrose. The first involves the conversion of sucrose to glucose and fructose by SAI with the glucose then being phosphorylated by hexokinase. The second is the conversion of sucrose to UDPglucose and fructose, with the UDPglucose subsequently converted to glucose-1-phosphate by UGPase for use in starch synthesis[22]. The main enzyme responsible for the synthesis of ADP-glucose, AGPase (EC2.7.7.27), is considered to catalyze the first and key step in starch biosynthesis[23,24]. Analysis of transgenic crops using techniques such as quantitative PCR will be essential for identifying and understanding the functions of the full complement of genes involved in these pathways[25].

In this study, we constructed a library of full-length cDNAs prepared from two cassava genotypes. The first, Arg7, is a cultivar of *M. esculenta* characterized by the high starch content (~32%) of its tuberous roots. The second, W14, is a selection of *M. esculenta* subsp. *flabellifolia*, a wild ancestor of *M. esculenta*, which has tuberous roots with low starch content (~5%). We used the SMART technology to obtain more full-length transcripts during library construction[26]. With the goal of establishing the molecular genetic bases of the substantial differences in root-starch contents between Arg7 and W14, we constructed cDNA libraries from leaf and root tissues of both varieties. We then used deep sequencing to identify genes that are differentially expressed in the four samples. Although we focused on genes likely to be important in photosynthesis and starch accumulation, availability of the library and the expression information obtained is likely to be of value for improving other aspects of cassava growth and development.

2 Materials and methods

2.1 Isolation of total RNA from *Manihot* root and leaf tissues

Two *Manihot* species were used in this study. The first was Arg7, a cultivar of *M. esculenta* with a high starch content (32%–34%) originating from Argentina. The second was W14 a selection of *M. esculenta* subsp. *flabellifolia*, the ancestral relative of *M. esculenta* with a low starch content (4%–5%) originating from Brazil. Both species were grown under field conditions using standard cultivation methods. Tissues of both species including leaves and roots were collected at 60, 120 and 270 days after planting, corresponding to storage root formation, storage root bulking and starch maturity stages. At each sampling three replicate samples were taken from separate individual plants. For leaves, the tip and tenth leaves down the top were sampled, and the petioles were removed before use. For each species, all root-derived samples and all leaf-derived samples were pooled separately. We used the RNAplant Reagent kit (TIANGEN, Beijing, China) to isolate total RNAs from both *Manihot* roots and leaves, using the extraction procedure recommended by the manufacturer. The quantity and quality of total RNA extracts were assessed using spectrophotometry and electrophoresis through 1% agarose gels.

2.2 Full-length cDNA synthesis and size fractionation

Before synthesizing first-strand cDNA, total RNAs were treated with DNase to remove contaminating genomic DNA. First-strand cDNA was synthesized from approximately 300 ng purified total RNA using a Creator SMART cDNA Library Construction Kit (DB Clontech, CA, USA) according to the manufacturer's protocol. We used the SMART IV oligonucleotide [5′-AAGCAGTGGTATCAACGCAGAGTGGCCATACGGCCGGG-3′] containing the *Sfi*I A site, and a CDS III/3′PCR primer [5′-ATTCTAGAGGCCGCCTCGGCCGACATG-d(T)30N1N3] containing the *Sfi*I B site. The first-strand cDNA product was used as template to synthesize double-stranded cDNA using long-distance PCR. Each PCR reaction was performed using 20 pmol of 5′ PCR primer [5′-AAGCAGTGGTATCAACGCAGAGT-3′], and 20 pmol of CDS III/3′PCR Primer. Each 50 μL PCR mixture contained 2 μL of first-strand cDNA product as template, 5 μL of 10 × Advantage 2 PCR Buffer, 1 μL of 50 × dNTP, 1 μL of 50 × Advantage 2 polymerase mix, 39 μL of deionized H_2O, and 20 μL of mineral oil overlaying the PCR mixture. The PCR reaction was performed using a Biometra thermal cycler (Biometra, Horsham, Germany) with the following program: denaturation at 95°C for 1 min, followed by 22 cycles of 95°C for 15 s, 68°C for 6 min, and incubation at 10°C at the end of cycling. We set up four 50 μL mixtures for each RNA sample. The 200 μL double-stranded cDNA mixture was extracted with phenol and chlorophenol, and precipitated by adding 20 μL of 3 $mol \cdot L^{-1}$ sodium acetate, 2.6 μL of glycogen (20 $mg \cdot \mu L^{-1}$) and 520 μL of 95% ethanol. The pellet comprising double-strand cDNA was washed twice with 75% ethanol, dried and suspended in 79 μL of deionized H_2O.

Double-strand cDNA was incubated with SfiI enzyme at 50°C for 2 h, and the reaction mixture contained 79 μL of double-strand cDNA, 10 μL of 10 × Sfi Buffer, 1 μL of BSA (100 μg·mL^{-1}) and 10 μL of *Sfi*I enzyme (200 U), in a total reaction volume of 100 μL. The digested double-stranded cDNA was fractionated in a 1% agarose gel by electrophoresis (6 V·cm^{-1}) and the size fraction corresponding to fragments between 100 bp and ~6.0 kb was excised. The excised gel slice was further divided into four pieces, ranging in size from ~500 bp to ~4.0 kb. The digested double-strand cDNA was then extracted and purified from the gel slice by using EZNA gel extraction kit (Omega, CA, USA). The cDNA of each slice was eluted in 50 μL of elution buffer, dried, and then resuspended in 8 μL of deionized H_2O.

2.3 cDNA library construction

The *Sfi*I-digested double-stranded cDNA fragments were ligated to 30 ng pDNR-Lib vector (DB Clontech, CA, USA) with T4 ligase (0.7 μL = 280 U, New England Biolabs, USA) in a 16°C water incubator overnight or for 16 h. The 10 μL ligation product was washed with 75% ethanol twice to remove salt ions, and dried in a super clean-bench, then suspended in 5 μL deionized H_2O. The purified ligation product was transformed into competent DH10B cell using electroporation. Transformed DH10B cells were allowed to recover by incubation in 400 μL of SOC medium for 45–60 min with gentle shaking, and then spread onto LB plates containing chloramphenicol (12.5 μg·mL^{-1}), and incubated at 37°C for 20 h. For the specific pDNR-Lib vector, only clones containing cDNA inserts could grow on the plates. Therefore, all white colonies were regarded as recombinant clones. These colonies were picked and manually transferred into separate wells of 384-well micro-titer plates. Each well of every 384-well plate contained 75 μL of freezing storage medium [360 mmol·L^{-1} K_2HPO4, 17 mmol·L^{-1} Na citrate, 4 mmol·L^{-1} $MgSO_4$, 68 mmol·L^{-1} $(NH_4)_2SO_4$, 44% (v/v) glycerol, 12.5 μg·mL^{-1} of chloramphenicol, LB]. Colonies were incubated at 37°C overnight, and were stored at −70°C.

2.4 cDNA library evaluation and clone sequence analysis

Forty-eight recombinant clones were randomly selected from each 384-well plate, and grown separately in the 2 mL wells of 96-well plates each containing 1.0 mL LB and 12.5 μg·mL^{-1} chloramphenicol. Plates were incubated overnight at 37°C with shaking at 200 r·min^{-1}. We then used colony PCR or digestion of purified plasmid DNA with *Hin*d III and *Eco*R I to evaluate the recombinant ratio (i.e., the number of PCR reaction clones with inserts divided by total number of PRC reaction clones) and the size of plasmid inserts. We randomly selected 6167 cDNA clones for sequencing analysis. Prior to sequencing, all plasmids were isolated from bacterial clones by cellular lysis and purified in 96-well plates. Single-pass sequencing using the M13 universal primer was used to characterize cDNA inserts from their 5′-ends.

2.5 EST sequence data processing

Original sequence data, as well as trace files, were processed using the Phred base-caller[26–28]. Sequences were trimmed using SeqClean software (TIGR Gene Indices Sequence Cleaning and Validation script, available at http://compbio.dfci.harvard.edu/tgi/software) to remove low-quality sequences and sequences derived from rRNA, poly-A tails, bacteria and vectors. All the resulting EST sequences with a sequence length longer than 100 nucleotides were initially clustered using the Megablast program[29], and assembled into contigs by CAP3 assembler with parameters set to identify any match of over 100 nucleotides and identify those spanning at least 95% of the sequences length[30,31]. All of the assembled ESTs were reported as either a singleton with one EST sequence or a contig with at least two overlapping EST sequences.

2.6 Functional annotation of the sequences

To compare the homologies of ESTs to known sequences, we performed online Blastn, using a cutoff value of 1E ⩽ −5 to compare the EST collections with entries in the integrative PlantGDB database (http://www.plantgdb.org). Gene annotation was performed using both online Blastx analysis with a cutoff value of 1E ⩽ −5 and comparison with entries in the GenBank protein database (http://www.ncbi.nlm.nih.gov). The top match was selected as the most likely potential homolog for the query EST sequence. Functional categorization and annotation were conducted as Gene Ontology (GO) annotation[30,31] (http://geneontology.org). The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database (http://www.genome.jp/kegg/) was used to assign genes to metabolism pathways using Blast2GO software[32] (https://www.blast2go.com).

2.7 Quantitative RT-PCR analysis

Total RNA used for quantitative RT-PCR was extracted from leaf and root samples using the method described above. For leaf tissue, we sampled unfolded leaves from the top two branches of 200 d old plants at six sampling times (once every 2 h, from 8 am to 6 pm). We took samples from three Arg7 individuals and two W14 individuals at each time point, and samples taken at the same time point from different individuals were pooled. We extracted RNA from 12 leaf samples altogether. We sampled storage roots from Arg7 plants that were 120, 160 and 200 d old. Roots of three individuals at the same time point were pooled. All fibrous root samples were also

pooled. Owing to the low frequency of storage roots in W14, we obtained only one storage-root sample from a 200 d old W14 plant. Altogether, we extracted RNA from five different root samples.

All qRT-PCR reactions were performed using a 6200HRM Corbett Life Science Rotor-Gene 6000 instrument (Corbett Research, Germany) with SYBR Premix Ex Taq^{TM} (TaKaRa Biotechnology Co., Ltd, Dalian, China), according to the manufacturer's instructions. We used Primer Premier 5.0 and Vector NIT 11.0 software (Applied Biosystems) to design primers based on the sequences of key genes of interest identified in our library. Whenever unigene sequences differed between the two species, the region with the highest homology was used for primer design. The average size of PCR products was approximately 170 bp. The cassava actin gene was used as an internal control. The sequences of primers used in this study are provided in Table 1.

Real-time PCR was performed following a standard SYBR Premix Ex TaqTM kit (TaKaRa) protocol. The reactions (1.0 μL first-strand cDNA product, 1 × Universal PCR Master Mix, 1.0 mmol·L^{-1} primers) were set up in 0.1 mL tubes of the Rotor-gene 6000 machine (QIANGEN, Hilden, Germany). The procedure for thermal cycling involved incubation at 95°C for 10 s, followed by 40 cycles of a program involving incubation at 95°C for 5 s, then 56°C for 15 s and finally, 72°C for 15 s. The procedure ended with a melt-curve ramping from 60 to 95°C, with temperature increases of 0.5°C at each step. Real-time PCR reactions were repeated six times for leaf samples and nine times for root samples. The relative mRNA level was calculated as $2^{-\Delta C_T}$. Correlation analysis and significance analysis were carried out using SPSS 17.0 software.

2.8 Evaluation of light intensity and net photosynthetic rate

The intensity of sunlight was determined using a HT4-GLZ-B photosynthetically active radiation meter HT4-GLZ-B (Zhongxi, Inc, Shengyang, Liaoning, China). The net photosynthetic rate (Pn) was determined using an LI-6400XT instrument (LI-6400: LI-COR, Inc, Lincoln, NE, USA).

3 Results

3.1 Construction and validation of the cDNA library in cassava

A full-length cDNA library was constructed from RNA extracted from leaves and roots of two *Manihot* species, cv. Arg7 (tuber starch content 32%) and its wild ancestral relative W14 (tuber starch content 4%–5%). The library comprises 32640 recombinant clones, with cDNA inserts having an average length of approximately 0.6 kb, ranging from 0.4 to 1.1 kb. In this library, only 60% of the sequences include a poly-A tail, although single-pass sequencing was only initiated from the 5′ ends of the cDNA inserts. The recombinant ratio of the cDNA library was 92%.

We sequenced 6028 randomly selected clones from their 5′ ends and 128 clones from their 3′ ends, generating 5593 reads in total, in which 3235 reads were derived from Arg7

Table 1 Genes of interest, primers and size of amplification products

Gene cellular activity	Abbreviation	EST sequence source	Forward/Reverse primer	Amplicon size/bp
RuBisco (Ribulose-1,5-bisphosphate carboxylase oxygenase)	*Me*rbcS	Arg7	5′-TGGCAGCTTCCACATTGTC-3′ 5′-CTCATACCTTCCACCACTCTCA-3′	175
Ferredoxin:NADP^{+} oxidoreductase	*Me*FNR	Arg7	5′-CGCTCTCTAGTGCCATGTTTAG-3′ 5′-GCACTCTATTTCCACTTCACCC-3′	185
NADH dehydrogenase (quinone reductase)	*Me*ndh	Arg7	5′-ATTCCAGCCAAGATGACGG-3′ 5′-ACAGAATTTAGTGAGCGGCG-3′	151
PEPC (Phophoenolpyruvate carboxylase)	*Me*ppc	Arg7/W14	5′-CGGATCAACGGGAAGCAAGA-3′ 5′-GCAAGATGAGTGGGTCCTCC-3′	196
AGPase (adenosine 5′-diphosphate glucose pyrophosphorylase)	*Me*LSU	W14	5′- AATCCAGATAATCCATTCGG-3′ 5′- TAGCAGCAACTCAAACGCCT-3′	165
UGPase (UDPglucose pyrophosphorylase)	*Me*UGP	*Jatropha curcas* (GenBank: GT973508.1)	5′- GGCTTGTAGAATCCGACGCA-3′ 5′- AGTTGCCTTCACTGGAAGGA-3′	170
SAI (soluble acid inverse)	*Me*Inv	Provided by Dr. Peng Zhang[#]	5′-ATCTACGACACCGAGGACTTCAT-3′ 5′-CTGTTCAATTCATCGTAAGTTCCA-3′	151
Constitutive control gene				
Actin	*Me*Act		5′-CGATGGTCGTACAACTGGTAT-3′ 5′-ATCCTCCAATCCAGACACTGT-3′	450

Note: [#], Shanghai Center for Cassava Biotechnology, Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, China.

Table 2 Summary of library properties and assembly results after sequencing the clones and their distributions among cDNA sub-libraries

Category	Library	Sub-library			
		Arg7 leaf	Arg7 root	W14 leaf	W14 root
No. of clones sequenced	6176	1082	2413	615	2066
EST	5593	1038	2197	438	1920
Analyzed EST > 100 bp	5067	964	1995	391	1717
Cluster	1153	195	746	157	322
Contig	476	62	256	47	143
Singlet	783	145	538	114	213
Unigene	1259	207	749	161	356
ORF prediction	1093	188	658	133	336

and 2358 reads were derived from W14 (Table 2). The resulting sequences were processed using the base-calling program 'Phred' (http://www.phrap.com/phred/) with satisfactory quality standards (100 nt of Q16, corresponding to an approximate error probability of 0.01)[33]. This generated 5013 high-quality EST sequences (more than 100 bp of EST sequence after removal of vector sequences, NCBI access numbers JK733886 to JK738898), of which 2671 included poly-A tail sequences. The lengths of ESTs ranged from 100 to 1003 bp, with an average length of 556.4 bp. The 5013 high-quality EST sequences clustered into 1153 clusters, in accordance with UniGene (http://www.ncbi.nlm.nih.gov/unigene) classification, and were further assembled into 476 contigs and 783 singletons using CAP3. The contigs and singletons represented 1259 putative unigenes with an average length of 563.2 bp. The longest unigene was 2380 bp. The percentage redundancy for each library varied between 79% (W14 root) and 58% (W14 leaf). The recombinant ratio of the whole library was 92%.

The information in the CAP3 assembly that we used to build a cluster profile represents the number of ESTs per assembled unigene (Fig. 1). Most (62.16%) of the unigenes had only one EST sequence. The analysis of the 5013 ESTs collection from this work revealed an average GC-content of 44.26%, ranging from 13.6% to 73.8%.

3.2 Functional categorization and reconstruction of metabolic pathways

GO analysis based on sequence similarity, enabled us to assign the 1259 unigenes to KEGG metabolic pathways[34]. We matched 746 (59.3%) of the unigenes to ESTs from model plants (considered as known genes), whereas 513 (49.7%) of the unigenes had no function assigned. Overall 323 (25.76%) of the unigenes were assigned to 114 different pathways, with some unigenes assigned to more than one pathway. As shown in Table 3, unigenes were assigned to all five major categories of the KEGG pathway: metabolism (176 unigenes), genetic information processing (48 unigenes), processing of environmental information (25 unigenes), cell processing (25 unigenes) and human diseases (58 unigenes).

The number of unigenes mapped to KEGG pathways differed substantially between the four sub-libraries (Appendix A, Table S1). For the carbohydrate metabolism and energy metabolism pathways we focused on in our study, we found significant differences in pathways related to photosynthesis, carbon fixation in photosynthetic organisms, glycolysis/gluconeogenesis, and oxidative phosphorylation (Fig. 2). Only six unigenes mapped to the starch and sucrose metabolism pathway, of which three, the genes encoding AGPase, sucrose synthase and starch branching enzyme, are important in starch synthesis.

Of the 103 unigenes predicted to function in carbohydrate and energy metabolism, 23 were only found in one of the four sub-libraries. Of these 23 unigenes, 14 were unique to Arg7 (with seven unigenes expressed in leaf tissue and seven in roots), and nine were unique to W14 (with two unigenes expressed in leaf tissue and seven in roots). Six unigenes that encode enzymes involved in carbohydrate metabolism were found in the Arg7 root sub-library, compared with only four in the W14 root sub-library. However, the AGPase unigene was only found in the W14 root sub-library. It is interesting that three unigenes encoding important photosynthetic enzymes (NADH dehydrogenase, ferredoxin: NADP + oxidoreductase and RuBisCO) were only found in the Arg7 leaf sub-library. The unigenes that encode enzymes involved in specialized metabolic pathways, such as acetohydroxyacid synthase (butanoate metabolism), phospholipase C (inositol phosphate metabolism) and donor: H_2O_2-oxidoreductase (methane metabolism), were found mainly in the sub-library prepared from W14 root tissue.

3.3 Differential expression of genes involved in photosynthesis and starch metabolism

To further understand the differential expression of genes involved in photosynthesis and starch metabolism in Arg7

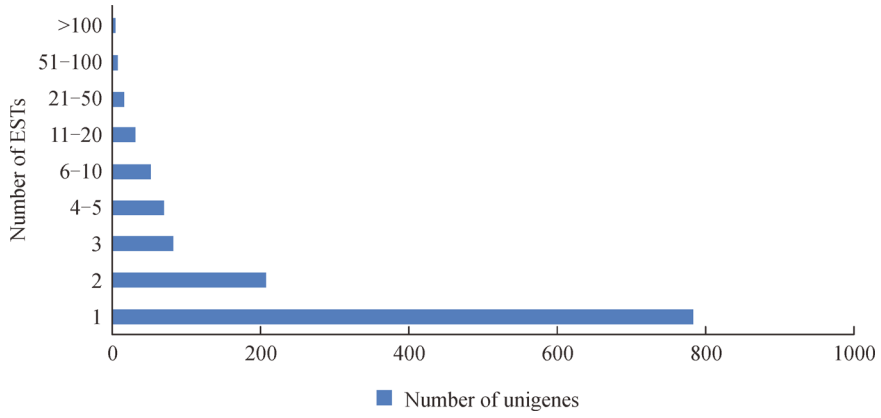

Fig. 1 Analysis of cassava ESTs assembled in terms of the number of ESTs per unigene. A total of 783 unigenes were represented by a single EST.

Table 3 The 323 unigenes mapped to KEGG pathways

KEGG pathway	No. of pathways mapped	No. of genes	Total genes mapped/%
Metabolism	69	176	54.49
Amino acid metabolism	8	12	3.72
Biosynthesis of secondary metabolites	6	8	2.48
Carbohydrate metabolism	15	51	15.79
Energy metabolism	8	52	16.10
Glycan biosynthesis and metabolism	5	6	1.86
Lipid metabolism	8	13	4.02
Metabolism of cofactors and vitamins	7	11	3.41
Metabolism of other amino acids	3	4	1.24
Nucleotide metabolism	2	11	3.41
Xenobiotics biodegradation and metabolism	7	8	2.48
Genetic information processing	8	48	14.86
Folding, sorting and degradation	4	8	2.48
Replication and repair	2	2	0.62
Transcription	4	38	11.76
Environmental information processing	8	16	4.95
Membrane transport	2	2	0.62
Signal transduction	6	14	4.33
Cellular processes	13	25	7.74
Cell communication	2	3	0.93
Cell growth and death	3	7	2.17
Cell motility	1	3	0.93
Endocrine system	4	9	2.79
Nervous system	2	2	0.62
Sensory system	1	1	0.31
Human diseases	14	58	17.96
Cancers	6	8	2.48
Immune disorders	1	9	2.79
Infectious diseases	2	8	2.48
Metabolic disorders	1	4	1.24
Neurodegenerative diseases	4	29	8.98

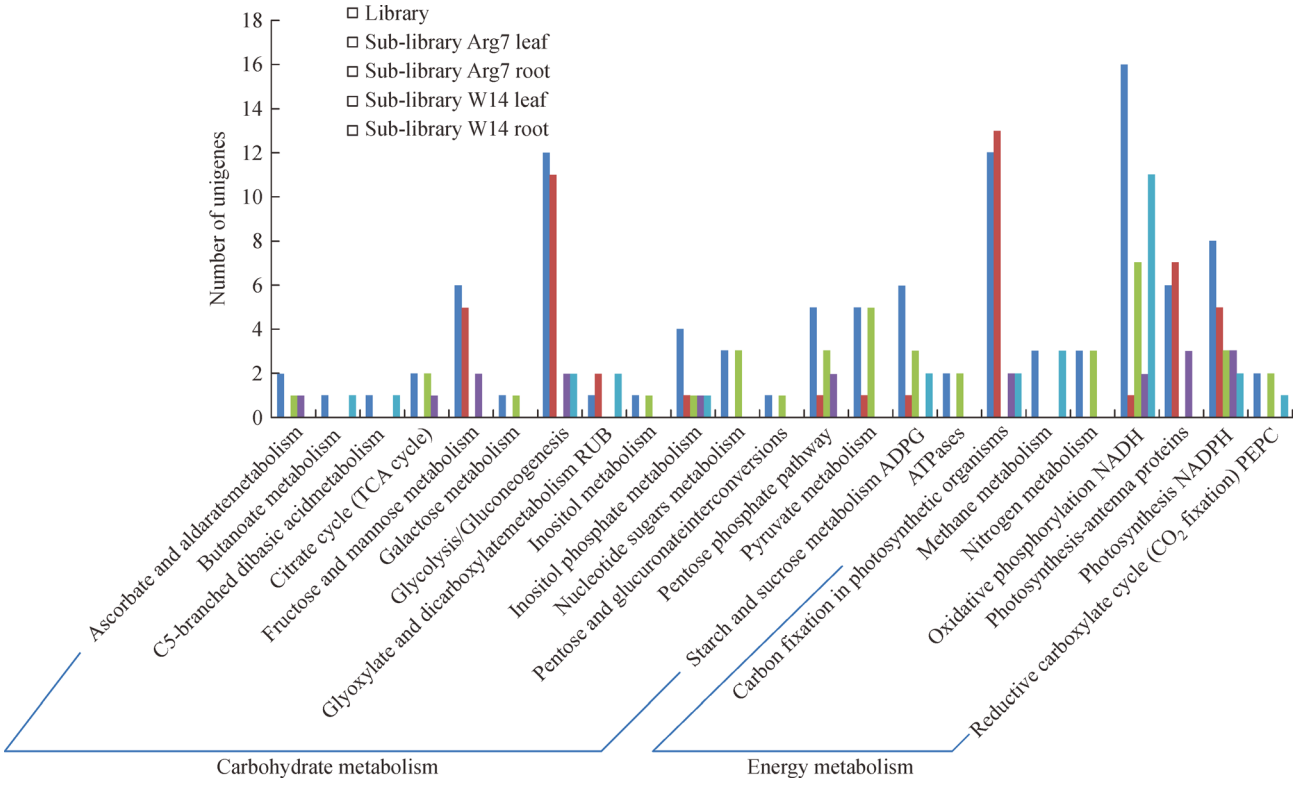

Fig. 2 The differential distribution of annotated unigenes in different pathways related to carbohydrate and energy metabolism among the four sub-libraries.

and W14, we selected five key genes, which were found in only one of the four sub-libraries and compared their general expression between Arg7 and W14. Of these four genes, those encoding PEPC, NADH dehydrogenase (NADH:ubiquinone reductase, EC1.6.5.3), and ferredoxin-NADP + oxidoreductase (EC1.18.1.2) were from the Arg7 leaf and root sub-libraries. The remaining gene, which encodes AGPase, was from the W14 root sub-library. Other key regulators of sucrose accumulation that we studied were SAI and UGPase. The genes encoding these enzymes (and abbreviations used for this in this paper) are listed in Table 1.

3.4 Daily expression pattern of genes related to photosynthesis in leaves

Compared with its wild ancestor, W14, Arg7 has both a higher Pn (Fig. 3) and higher starch content in storage roots (32%–34% compared with 4%–5% in W14). Figure 4 compares the expression patterns of five genes at six time points during the day. Expression of the *Me*FNR gene was higher and better correlated ($P < 0.0.5$) with ambient light intensity in Arg7 plants than in W14 plants (Fig. 4a; Appendix A, Table S1). Expression of *Me*FNR was highest at 10 am in both Arg7 and W14, and significantly lower at the five other time points examined (Fig. 4b). The expression levels at noon and 2 pm were higher than at the other four times in Arg7 and higher than at any time in

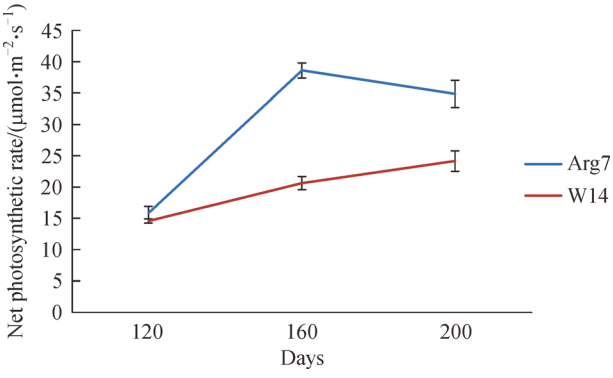

Fig. 3 Net photosynthetic rates in the leaves of Arg7 and W14 during the growing period 120, 160, and 200 d after planting. All measurements were made at 10:00 am. Each data point represents the mean±SD of three replicates.

W14, but its expression levels in W14 were significantly higher than those in Arg7 at 10 am and 6 pm (Fig. 4c). The expression of *Me*FNR was significantly correlated with light intensity ($P < 0.05$). Given the high rates of consumption of NADP during photosynthesis and NADH during photorespiration[17], levels of expression of *Me*FNR in leaves are likely to influence the catalytic activity of RuBisCO, which has both carboxylase and oxygenase activities. In this regard, it is notable that the

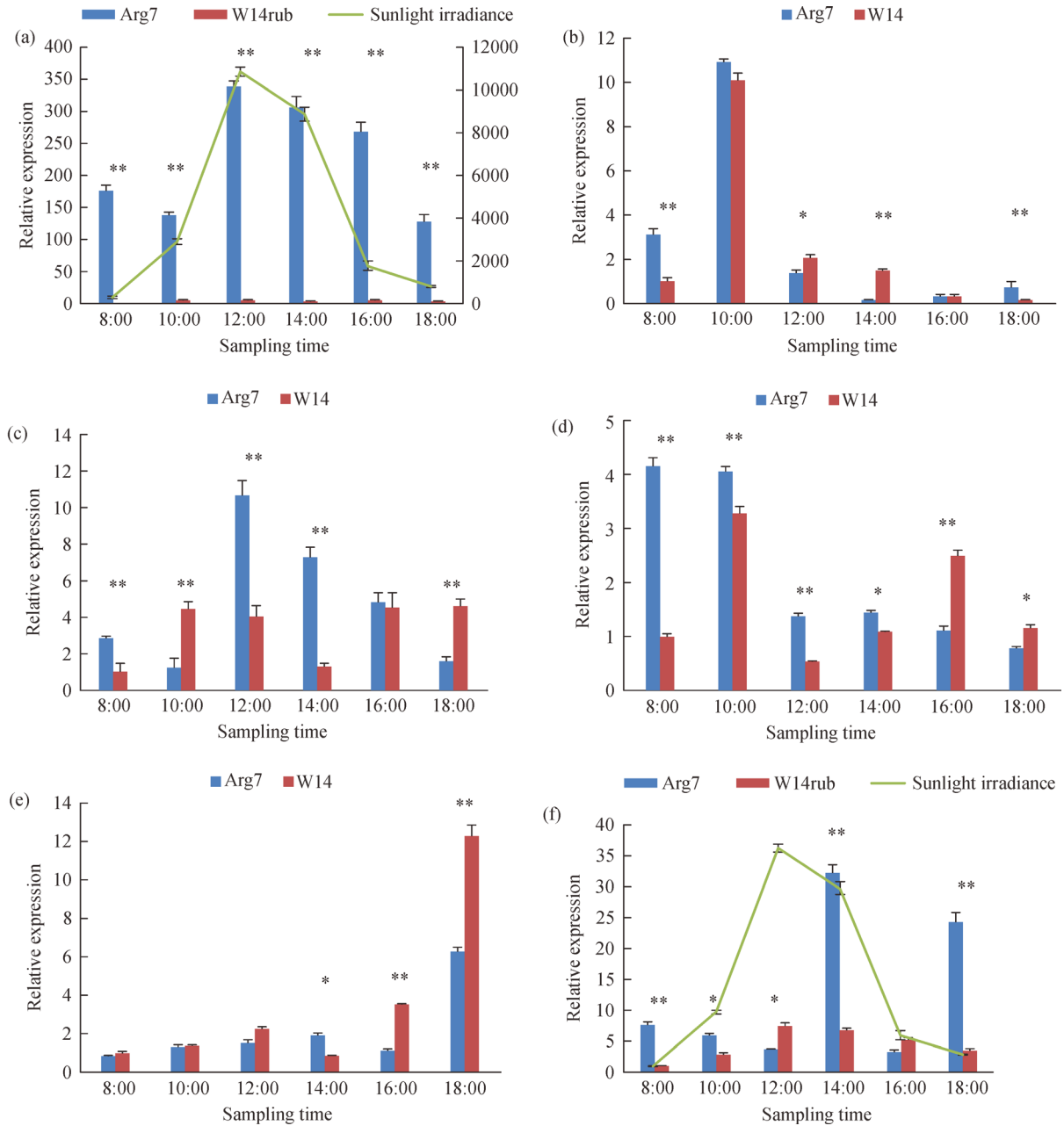

Fig. 4 Diurnal fluctuations in the expression of selected genes in cassava leaves measured using real-time PCR, with variations in transcript levels associated with changes in light intensity between 8:00 h and 18:00 h. (a) *Me*rbcS; (b) *Me*FNR; (c) *Me*ndh; (d) *Me*ppc; (e) *Me*LSU; (f) *Me*Inv. Each data point represents the mean±SD of nine replicates, with values indicating fold-changes in relative gene expression following normalization relative to the actin cassava gene. Asterisks above each bar indicate the degree of significance differences in the expression levels between the Arg7 cultivar of *Manihot esculenta* and the W14 variety of *Manihot flabellifolia* (*, $P < 0.05$; **, $P < 0.01$).

changes in the expression levels of *Me*rbcS and were significantly correlated ($P < 0.01$).

Real-time PCR analysis indicated that the expression of *Me*PPC was higher in Arg7 than in W14 from 8 am to 2 pm, but lower at 4 and 6 pm (Fig. 4d). This result is consistent with higher levels of PEPC activity in leaves of Arg7 than in those of W14, and the observation of distinct, green bundle sheath cells in leaves of Arg7 but not in leaves of W14 (unpublished data). These observations suggest that PEPC fixes more CO_2 in Arg7 than in W14. The high rates of photosynthesis in individual leaves and existence of distinct green bundle sheath cells in leaves of certain *Manihot* species suggest that cassava leaves possess some novel photosynthesis characteristics[33,35]. In both varieties, expression of *Me*LSU increased from 8 am to 6 pm, and at 4 pm in Arg7 (Fig. 4e).

The enzyme responsible for the synthesis of ADP-glucose, AGPase, is considered the first and key step in starch biosynthesis. Starch synthesized during the day in leaves is temporarily stored in chloroplasts, and then degraded to sucrose at night for transport to roots[36,37]. Levels of expression of *Me*INV were fairly consistent in Arg7 over most of the time points measured, but with peaks in transcript at 2 and 6 pm. For W14, trends in the expression of *Me*INV transcripts were significantly related to light intensities, although levels were generally lower than those observed in Arg7 (Fig. 4f). SAI has been suggested to be a key regulator of sucrose accumulation. In leaves, reducing sugars produced from photosynthetically generated sucrose support leaf growth and are partly transported to sink organs for storage. Accordingly, high acid invertase activities always follow high rates of plant growth[38]. Sucrose export from leaves is higher during the day than the night[39]. Near the end of the day (6 pm), the content of sucrose in leaves was nearly at its highest level[40], increased expression of *Me*INV and *Me*LSU transcripts was likely attributable to their regulation by cellular osmotic pressure. Moreover, the efficiency of photosynthesis was greatest at midday under the highest light intensity. This likely increased sucrose levels, which in turn likely accounts for the upregulated expression of *Me*INV at 2 pm in Arg7.

3.5 Expression of genes related to starch synthesis in roots

We selected six genes (*MeFNR*, *Mendh*, *Meppc*, *MeLSU*, *MeUGP* and *MeInv*) that encode key enzymes in the pathways of photosynthesis and starch accumulation to assess their expression in roots. All five genes were upregulated with the enlargement of storage roots in Arg7. The oxidized and reduced forms of the cofactors NAD (including NAD^+ and NADH) and NADP (including $NADP^+$ and NADPH) are fundamental mediators of various biological processes, including energy metabolism, mitochondrial functions, calcium homeostasis, the regulation of cellular redox potential, gene expression, pathogen defenses, aging and cell death[41]. The increased expression of MeFNR is consistent with increases in the rate of cellular metabolism from the early periods of root formation to the rapid enlargement of storage roots (Fig. 5a, Fig. 5b). The PEPC enzyme is important in regulating the metabolism of phosphorus and organic acids in roots[42,43], and cassava AGPase is activated by 3-phosphoglycerate and inhibited by up to 90% in the presence of inorganic phosphate[44]. Given that rates of expression of both *Me*ppc and *Me*LSU were upregulated during the process of storage-root enlargement (Fig. 5c, Fig. 5d), we speculated that there may be a synergistic relationship between *Me*ppc and *Me*LSU. The expression of *Me*UGP appeared to reach a maximum when plants were around 200 d old, whereas the expression of *Me*Inv was slightly lower at 200 d than that at 160 d in Arg7 storage roots (Fig. 5e, Fig. 5f). These observations may be attributable to the increase in the ratio of sucrose allocation to starch synthesis during the course of storage-root enlargement in Arg7. The observation that the highest expression level of *Me*Inv was measured in Arg7 fibrous roots suggests that more sucrose was converted to glucose and fructose in fibrous roots than stored as starch.

Overall, photosynthetic efficiency was higher in Arg7 than in W14, and the proportion of photosynthesis products that were stored in leaf as temporary starch was less in Arg7 than in W14. All of the selected genes involved in starch synthesis were expressed at higher levels in Arg7 storage roots than in W14 storage roots. This might explain, at least in part, the higher starch content in Arg7 storage roots compared to W14 storage roots.

4 Discussion

4.1 Value of the full-length cDNA library in cassava

We constructed a library comprising 32640 full-length recombinant cDNA clones from cassava, and sequenced 6176 clones to generate 5067 EST sequences representing 1259 unigenes, which have been deposited in EST databases. Identification of cassava genes that encode either key enzymes in photosynthesis or starch metabolism, or regulators of these pathways, should greatly facilitate improvement of this important crop. Of the unigenes represented in the library, 746 unigenes were homologous to ESTs from model plants represented in the KEGG database, and 323 unigenes were mapped to 114 different pathways. In particular, 103 of the unigenes represented in our library are implicated in carbohydrate and energy metabolism, including the unigenes that encode RuBisCO, PEPC, pyruvate kinase, AGPase, starch synthase, trehalose-6-phosphate synthase, and ATP synthase. It has been reported that overexpression of AGPase, pyrophosphatase and NADH-glutamate synthase leads to increased root biomass in potato and cassava[39,45]. Some of the unigenes showing no similarity with other proteins in the database may be specific to cassava. It is possible that a few of them may simply represent untranslated regions or be too short to have domains that would present significant matches with other proteins, although we suspect that the majority of these genes define proteins specific to cassava or its close relatives. In summary, these gene resources in our library will contribute to the genetic improvement of cassava not only in terms of some critically important metabolic pathways but also to bolster ongoing efforts to assemble a complete gene catalog for this species.

It is noteworthy that some unigenes were mapped to KEGG metabolic pathways parallel to those in humans (Table 3). It is possible that some of the pathways have

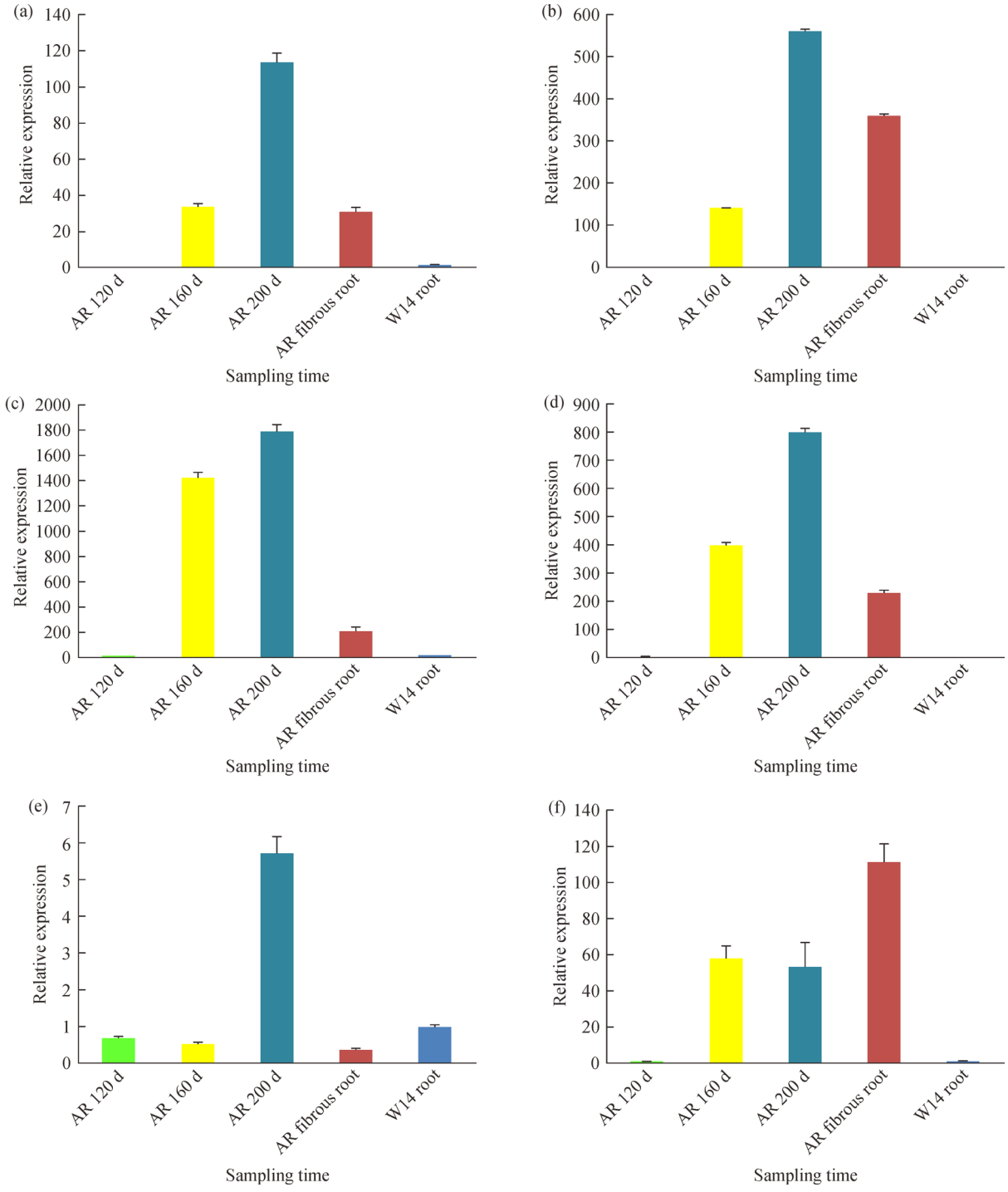

Fig. 5 Expression levels of selected genes implicated in photosynthesis and carbohydrate metabolism in cassava, using real-time PCR. (a) *Me*FNR; (b) *Me*ndh; (c) *Me*ppc; (d) *Me*LSU; (e) *Me*UGP; (f) *Me*Inv. Each data point represents the mean±SD of nine replicates, with values indicating fold-changes in relative gene expression following normalization relative to the actin cassava gene. AR 120 d, AR 160 d and AR 200 d are levels in Arg7 storage roots 120, 160, and 200 d after planting, respectively.

common ancient origins dating to before the evolutionary divergence of plants and animals[46]. Recently, it was reported that the tomato 14-3-3 protein 7, a homolog of human protein 14-3-3e, regulates immunity-associated programmed cell death in tomato, and that both proteins share similar signaling transduction pathways dependent on phosphorylation[47]. Therefore, it will be valuable to investigate whether cassava has signaling pathways that parallel the functions of human signaling cascades.

Our full-length cDNA library was assembled from four

sub-libraries in order to derive functional information through parallel comparisons of the compositions of the sub-libraries. There was no evidence of differential expression between W14 and Arg7 for 5067 of the ESTs in our library. These insights were enabled primarily by next generation sequencing technology, which can generate millions of reads in a relatively short time and at low cost, and is regarded as a powerful method to analyze the transcriptomes of eukaryotic genomes[48]. Despite the limited number of clones sequenced, we still obtained some valuable insights, related especially to carbohydrate and energy metabolism (Table 3). Therefore, in contrast to the stress-related full-length cDNA library for cassava published by BMC genomics[14], our library was valuable in discovering the genes related to starch accumulation in cassava. At the same time, our library is the first EST collection of *M. esculenta* subsp. *flabellifolia*, an ancestral relative of cultivated cassava and will be useful to study evolution of the genus *Manihot*.

4.2 Understanding of higher starch content in Arg7 storage roots

Photosynthesis products are partly exported from chloroplasts as sucrose and partly turned into starch, a temporary energy store in chloroplasts, during the day[36]. Plastidic sucrose is partly transported into the vacuole and then cleaved into glucose and fructose, with the remainder transported to storage roots, which act as sink organs. There are two pathways for the transport of sucrose from leaves to sink organ cells. One is the pathway forming storage starch in amyloplasts, whereas the other is transport into the vacuole for cleavage into glucose and fructose for energy storage[49].

Whereas cultivated cassava (*M. esculenta*) has several characteristics typical of C_4 photosynthesis that are not usual for C_3 plants[9,10,50], wild *Manihot* species are incapable of C_4 photosynthesis owing to the low activity of PEPC, the key enzyme of the C_4 pathway[51]. In this study, the expression level of *Me*PPC was higher than that in W14 and was at significant or very significant levels between 8 am and 2 pm. Moreover, Pn was higher in Arg7 than in W14 in 120, 160 and 200 d old plants (Fig. 3). In summary, photosynthetic efficiency is higher in Arg7 than in W14.

Activation of AGPase stimulates starch synthesis and decreases levels of glycolytic intermediates in leaves and storage roots[36,52]. In our study, the expression level of *Me*LSU was higher in leaves of W14 than in those of Arg7 throughout most of the day, especially from 4 to 6 pm (Fig. 5e). Nonetheless, in 200 d old roots, the level of expression of *Me*LSU in Arg7 was nearly 800 times more than that in W14 (Fig. 5d). These data suggest that W14 temporarily stores more photosynthate as starch in leaves than Arg7, and the latter has a greater ability to synthesize starch in storage roots than W14.

4.3 Starch accumulation in cassava

The differential expression of the aforementioned genes that encode key enzymes involved in photosynthesis and starch synthesis possibly partly explains the differences in the efficiencies with which the two species accumulate starch in storage roots, although we cannot dismiss the possibility of regulation at the post-transcriptional level. Analysis of the promoters of these genes and the transcription factors that control their expression will be the focus of our future studies. For instance, a better understanding of the molecular basis that underlies the differential expression of *Me*INV and *Me*UGP in storage roots and fibrous roots may provide insights into the mechanism of sucrose allocation between organs with different functions.

Further research related to the pathway of starch accumulation metabolism in cassava should concentrate on both the regulation of the genes that encode key enzymes, the distribution of photosynthetic products throughout different parts of the plant, and the relationship between their transport and the accumulation of starch in roots. A first step toward this goal would involve evaluation of the expression profiles in different organs and under various conditions of the homologous genes detected within our library. This could be done through the construction of microarrays, for which once again, the availability of our full-length cDNA library would be an invaluable resource[53].

5 Conclusions

Comparable full-length cDNA libraries of the leaves and storage roots of cultivar Arg7 and wild ancestor W14 were constructed with a total of 32640 recombinant clones. Primary annotation and expression validation revealed that Arg7 has a higher net photosynthesis rate and RuBPase activities than W14 in leaves, and has absolute higher starch accumulation efficiency than W14 in storage roots. This resource will provide bases for gene discovery, genome annotation and evolution in cassava.

Acknowledgements This work was supported by a National Nonprofit Institute Research Grant of CATAS-ITBB, and the earmarked fund for China Agricultural Research System (CARS-12). We thank Prof. Peng Zhang of Chinese Academy of Sciences for his constructive suggestions on this research.

Supplementary materials The online version of this article at http://dx.doi.org/10.15302/J-FASE-2016113 contains supplementary material (Appendix A).

Compliance with ethics guidelines The authors Yang Zhang, Xin Chen, Haiyan Wang, Zhiqiang Xia, Peng Ling, and Wenquan Wang declare that they have no conflicts of interest or financial conflicts to disclose.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

1. Taylor N, Chavarriaga P, Raemakers K, Siritunga D, Zhang P. Development and application of transgenic technologies in cassava. *Plant Molecular Biology*, 2004, **56**(4): 671–688
2. de Souza C R B, Carvalho L J C B, de Mattos Cascardo J C. Comparative gene expression study to identify genes possibly related to storage root formation in cassava. *Protein and Peptide Letters*, 2004, **11**(6): 577–582
3. Koch B, Sibbesen O, Swain E, Kahn R, Liangcheng D, Bak S, Halkier B, Møller B. Possible use of a biotechnological approach to optimize and regulate the content and distribution of cyanogenic glycosides in cassava to increase food safety. *Acta Horticulturae*, 1994, (375): 45–60
4. Siritunga D, Sayre R T. Domestication of cassava: generation of cyanogen-free cassava. *Planta*, 2003, **217**(3): 367–373
5. Jansson C, Westerbergh A, Zhang J, Hu X W, Sun C. Cassava, a potential biofuel crop in the People's Republic of China. *Applied Energy*, 2009, **86**(S1): S95–S99
6. Ziska L H, Runion G B, Tomecek M, Prior S A, Torbet H A, Sicher R. An evaluation of cassava, sweet potato and field corn as potential carbohydrate sources for bio-ethanol production in Alabama and Maryland. *Biomass and Bioenergy*, 2009, **33**(11): 1503–1508
7. Hunt C A, Wholey D M, Cock J H. Growth physiology of cassava (*Manihot esculenta* Crantz). *Field Crop Abstracts*, 1977, **30**: 77–89
8. Edwards G E, Sheta E, Moore B D, Dai Z, Fransceschi V R, Cheng S H, Lin C H, Ku M S B. Photosynthetic characteristics of cassava (*Manihot esculenta*), a C_3 species with chlorenchymatous bundle sheath cells. *Plant & Cell Physiology*, 1990, **31**(8): 1199–1206
9. Angelov M N, Sun J, Byrd G T, Brown R H, Black C C. Novel characteristics of cassava, *Manihot esculenta* Crantz, a reputed C_3–C_4 intermediate photosynthesis species. *Photosynthesis Research*, 1993, **38**(1): 61–72
10. EI-Sharkawy M A. International research on cassava photosynthesis, productivity, eco-physiology, and responses to environmental stresses in the tropics. *Photosynthetica*, 2006, **44**(4): 481–512
11. Osunsami A T, Akingbala J O, Oguntimein G B. Effect of storage on starch content and modification of cassava starch. *Starch*, 1989, **41**(2): 54–57
12. Lopez C, Jorge V, Piegu B, Mba C, Cortes D, Restrepo S, Soto M, Laudie M, Berger C, Cooke R, Delseny M, Tohme J, Verdier V. An unigene catalogue of 5700 expressed genes in cassava. *Plant Molecular Biology*, 2004, **56**(4): 541–554
13. Lokko Y, Anderson J V, Rudd S, Raji A, Horvath D, Mikel M A, Kim R, Liu L, Hernandez A, Dixon A G, Ingelbrecht I L. Characterization of an 18166 EST dataset for cassava (*Manihot esculenta* Crantz) enriched for drought-responsive genes. *Plant Cell Reports*, 2007, **26**(9): 1605–1618
14. Sakurai T, Plata G, Rodriguez-Zapata F, Seki M, Salcedo A, Toyoda A, Ishiwata A, Tohme J, Sakaki Y, Shinozaki K, Ishitani M. Sequencing analysis of 20,000 full-length cDNA clones from cassava reveals lineage specific expansions in gene families related to stress response. *BMC Plant Biology*, 2007, **7**: 66
15. Li Y Z, Pan Y H, Sun C B, Dong H T, Luo X L, Wang Z Q, Tang J L, Chen B. An ordered EST catalogue and gene expression profiles of cassava (*Manihot esculenta*) at key growth stages. *Plant Molecular Biology*, 2010, **74**(6): 573–590
16. Robert J S, Michael E S. RUBISCO: structure, regulatory interactions, and possibilities for a better enzyme. *Annual Review of Plant Biology*, 2002, **53**(4): 449–475
17. Allen J F. Oxygen reduction and optimum production of ATP in photosynthesis. *Nature*, 1975, **256**(5518): 599–600
18. Bandyopadhyay A, Datta K, Zhang J, Yang W, Raychaudhuri S, Miyao M, Datta S K. Enhanced photosynthesis rate in genetically engineered *indica* rice expressing *pepc* gene cloned from maize. *Plant Science*, 2007, **172**(6): 1204–1209
19. Chollet R, Vidal J, O'Leary M H. Phosphoenolpyruvate carboxylase: a ubiquitous, highly regulated enzyme in plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, 1996, **47**(1): 273–298
20. Yoshihara N, Imayama T, Fukuchi-Mizutani M, Okuhara H, Tanaka Y, Ino I, Yabuya T. cDNA cloning and characterization of UDP-glucose: anthocyanidin 3-O-glucosyltransferase in *Iris hollandica*. *Plant Science*, 2005, **169**(3): 496–501
21. Tomlinson P T, Duke E R, Nolte K D, Koch K E. Sucrose synthase and invertase in isolated vascular bundles. *Plant Physiology*, 1991, **97**(3): 1249–1252
22. Fernie A R, Willmitzer L, Trethewey R N. Sucrose to starch: a transition in molecular plant physiology. *Trends in Plant Science*, 2002, **7**(1): 35–41
23. Kleczkowski L A. Is leaf ADP-glucose pyrophosphorylase an allosteric enzyme? *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*, 2000, **1476**(1): 103–108
24. Kleczkowski L A. A new player in the starch field. *Plant Physiology and Biochemistry*, 2001, **39**(9): 759–761
25. Stam M, Mol J N M, Kooter J M. The silence of genes in transgenic plants. *Annals of Botany*, 1997, **79**(1): 3–12
26. Zhu Y Y, Machleder E M, Chenchik A, Li R, Siebert P D. Reverse transcriptase template switching: a SMART^{TM} approach for full-length cDNA library construction. *BioTechniques*, 2001, **30**(4): 892–897
27. Ewing B, Green P. Base-calling of automated sequencer traces using Phred. II. Error probabilities. *Genome Research*, 1998, **8**(3): 186–194
28. Ewing B, Hillier L D, Wendl M C, Green P. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome Research*, 1998, **8**(3): 175–185
29. Zhang Z, Schwartz S, Wagner L, Miller W.A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, 2000, **7**(1–2): 203–214
30. Huang X, Madan A. CAP3: a DNA sequence assembly program. *Genome Research*, 1999, **9**(9): 868–877
31. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B, Tsai J, Quackenbush J. TIGR gene indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics*, 2003, **19**(5): 651–652
32. Camon E, Barrell D, Lee V, Dimmer E, Apweiler R. The gene ontology annotation (GOA) database-an integrated resource of GO annotations to the UniProt knowledgebase. *In Silico Biology*, 2004, **4**(4): 5–6
33. El-Sharkawy M A, Cock J H, de Cadena G. Influence of differences

in leaf anatomy on net photosynthetic rates of some cultivars of cassava. *Photosynthesis Research*, 1984, **5**(3): 235–242

34. Mullikin J C, McMurray A A. Sequencing the genome, fast. *Science*, 1999, **283**(5409): 1867–1868
35. Angelov M N, Sun J, Byrd G T, Brown R H, Black C C. Novel characteristics of cassava, *Manihot esculenta* Crantz, a reputed C_3–C_4 intermediate photosynthesis species. *Photosynthesis Research*, 1993, **38**(1): 61–72
36. Zeeman S C, Smith S M, Smith A M. The diurnal metabolism of leaf starch. *Biochemical Journal*, 2007, **401**(1): 13–28
37. Zeeman S C, Kossmann J, Smith A M. Starch: its metabolism, evolution, and biotechnological modification in plants. *Annual Review of Plant Biology*, 2010, **61**(1): 209–234
38. Estruch J J, Beltrán J P. Change in invertase activities precede ovary growth induced by gibberellic acid in *Pisum sativum. Physiologia Plantarum*, 1991, **81**(3): 319–326
39. Geigenberger P, Stitt M. Diurnal changes in sucrose, nucleotides, starch synthesis and AGPs transcript in growing potato tubers that are suppressed by decreased expression of sucrose phosphate synthase. *Plant Journal*, 2000, **23**(6): 795–806
40. Gerhardt R, Stitt M, Heldt H W. Subcellular metabolite levels in spinach leaves. *Plant Physiology*, 1987, **83**(2): 399–407
41. Ying W H. NAD^+/NADH and $NADP^+$/NADPH in cellular functions and cell death: regulation and biological consequences. *Antioxidants & Redox Signalling*, 2008, **10**(2): 179–206
42. Schulze J, Shi L F, Blumenthal J, Samac D A, Gantt J S, Vance C P. Inhibition of alfalfa root nodule phosphoenolpyruvate carboxylase through an antisense strategy impacts nitrogen fixation and plant growth. *Phytochemistry*, 1998, **49**(2): 341–346
43. Begum H H, Osaki M, Shinano T, Miyatake H, Wasaki J, Yamamura T, Watanabe T. The function of a maize-derived phosphoenol pyruvate carboxylase (PEPC) in phosphorus-deficient transgenic rice. *Soil Science and Plant Nutrition*, 2005, **51**(4): 497–506
44. Munyikwa T, Kreuze J, Fregene M, Suurs L, Jacobsen E, Visser R. Isolation and characterisation of cDNAs encoding the large and small subunits of ADP-glucose pyrophosphorylase from cassava (*Manihot esculenta* Crantz). *Euphytica*, 2001, **120**(1): 71–83
45. Ihemere U, Arias-Garzon D, Lawrence S, Sayre R. Genetic modification of cassava for enhanced starch production. *Plant Biotechnology*, 2006, **4**(4): 453–465
46. Ausubel F M. Are innate immune signaling pathways in plants and animals conserved? *Nature Immunology*, 2005, **6**(10): 973–979
47. Oh C S, Pedley K F, Martin G B. Tomato 14–3-3 protein 7 positively regulates immunity-associated programmed cell death by enhancing protein abundance and signaling ability of MAPKKK α. *Plant Cell*, 2010, **22**(1): 260–272
48. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 2009, **10**(1): 57–63
49. Azcón-Bieto J, Lambers H, Day D A. Effect of photosynthesis and carbohydrate status on respiratory rates and the involvement of the alternative pathway in leaf respiration. *Plant Physiology*, 1983, **72** (3): 598–603
50. El-Sharkawy M A, Cock J H, Held A A. Photosynthetic responses of cassava cultivars (*Manihot esculenta* Crantz) from different habitats to temperature. *Photosynthesis Research*, 1984, **5**(3): 243–250
51. Tiessen A, Hendriks J H M, Stitt M, Branscheid A, Gibon Y, Farré E M, Geigenberger P. Starch synthesis in potato tubers is regulated by post-translational redox modification of ADP–Glucose pyrophosphorylase: a novel regulatory mechanism linking starch synthesis to the sucrose supply. *Plant Cell*, 2002, **14**(9): 2191–2213
52. Calatayud P A, Barón C H, Velásquez H, Arroyave J A, Lamaze T. Wild *Manihot* species do not possess C_4 photosynthesis. *Annals of Botany*, 2002, **89**(1): 125–127
53. Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y, Shinozaki K. Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. *Plant Cell*, 2001, **13**(1): 61–72