

# Characterization of Carotenoid-protein Complexes and Gene Expression Analysis Associated with Carotenoid Sequestration in Pigmented Cassava (*Manihot Esculenta* Crantz) Storage Root

Luiz Joaquim Castelo Branco Carvalho<sup>\*1</sup>, John Lippolis<sup>2</sup> Songbi Chen<sup>3</sup>, Claudia Regina Batista de Souza<sup>4</sup>, Eduardo A. Vieira<sup>5</sup>, James V. Anderson<sup>6</sup>

<sup>1</sup>EMBRAPA Genetic Resources and Biotechnology, Brazil

<sup>2</sup>USDA-ARS, National Animal Disease Center, Ames-IA, USA

<sup>3</sup>CATAS Tropical Crops Genetic Resources Institute, Dangsou, China

<sup>4</sup>Department of Genetics – UFPA, Brazil

<sup>5</sup>EMBRAPA Cerrados, Brazil

<sup>6</sup>USDA-ARS, Sunflower and Plant Biology Research Unit, Fargo, ND, USA

**Abstract:** Carotenoid-protein complex (CPC) was isolated from chromoplast-enriched suspensions of cassava storage root (CSR) using size exclusion chromatography and characterized. Peptide sequences (LC\_MS/MS spectrum) obtained from CPC and their corresponding proteins were obtained using publically available databases. Small Heat Shock Proteins (sHSPs) were the most abundant proteins identified in the CPC. Western blot analysis showed that Fibrillin and Or-protein were present in chromoplast-enriched suspensions of yellow root but not in the complex or white root. Results from qRT-PCR helped identify an isoform of HSP21 possessing four single point mutations in the intense yellow CSR that may be responsible for increased sequestration of b-carotene.

**Keywords:** Cassava root, carotenoid accumulation, carotenoid-protein complex, heat shock proteins.

## INTRODUCTION

Plastid lipid-associated proteins (PAPs) and Carotenoid Associated Proteins (CAP) are general terms conferred to two classes of proteins accumulated at high levels in chromoplasts of plant cells [1,2] that are involved in the sequestration of hydrophobic compounds including lipids and carotenoids. These proteins are being investigated by several groups to establish the role of their parallel accumulation with carotenoids [3] and lipids of different types [4,5]. These investigations also include examining chromoplast formation [2,3,6], which specially focus on the Fibrillar type that forms a carotenoid-protein complex (CPC) within supramolecular structures where different carotenoid types are sequestered [2,7,8]. Most of this information was obtained from green tissues where carotenoid function is well documented in dissipating excess excitation energy by participating in non-photochemical quenching, which is essential in protecting the chloroplasts from photo-oxidative damage [9] and lipid changes that occur during leaf senescence due to

interconversion of chloroplasts into chromoplast [5]. Here we provide information on the characterization of carotenoid-protein complex isolated in cassava storage root that is a non green tissue.

Analysis of sequence data generated by LC-MS/MS is one of the classic PROTEOMICs technologies widely adopted for sequencing complex protein mixtures and predicting individual proteins [10]. Although erroneous protein identification is appreciated [11], the identification of individual proteins by LC-MS/MS is largely supported by statistical scoring systems [12, <http://systemsbiology.org/research/software.html>], which increase our confidence about protein sequence identification and homology matches to peptide sequences deduced from species specific EST databases, including the EMBRAPA cassava genome project database [<http://genoma.embrapa.br/genoma/>], ESTIMA database [[http://titan.biotech.uiuc.edu/cgi-bin/ESTWebsite/-estima\\_start?seqSet=cassava](http://titan.biotech.uiuc.edu/cgi-bin/ESTWebsite/-estima_start?seqSet=cassava)], CIAT data base [<http://web-app.ciat.cgiar.org/biotechnology/bioinformatics.htm>], Riken full-length cDNA database [<http://www.brc.riken.go.jp/lab/epd/Eng/catalog/pDNA.shtml>] and the *Arabidopsis* counterparts (ATG\_ORTHOLOGUE) (<http://www.arabidopsis.org/>). These resources and other public databases facilitate identification of proteins and their corresponding gene se-

\*Address correspondence to this author at the EMBRAPA Genetic Resources and Biotechnology, Laboratory of Biochemistry and Biophysics, Caixa Postal 02372, 70 770 916 Brasilia, DF, Brazil; Tel: + 55 61 3448 4664, + 55 61 3448 4937, + 55 61 9557 0838; Fax: + 55 61 3340 3658; E-mail: [luiz.castelo@embrapa.br](mailto:luiz.castelo@embrapa.br)

quences using validated peptide sequences derived from shotgun PROTEOMIC approaches. Here, we report results from combining chromoplast-enriched suspension preparation, protein-carotenoid complex separation in conventional size exclusion chromatography (SEC), protein separation by SDS-PAGE, and shotgun PROTEOMIC sequence techniques to identify proteins from carotenoid-protein complex present in cassava storage root.

## MATERIALS AND METHODOLOGY

### Plant Material and Tissue Preparation

Cassava Storage Root (CSR) from field-grown plants at EMBRAPA Genetic Resources and Biotechnology Research Unit (Brasilia, DF, Brazil) were used in this study. For protein extraction, quantification and protein blot analysis we used landrace BGM019 (intense yellow root) and the cv. IAC12.829 (white root). Protein sequence from carotenoid-protein complex was obtained from landrace BGM019. For gene expression analysis we used landrace BGM019 and cv. IAC12-829. Fresh storage roots (cylinders of 30-40 cm length x 4-6 cm diameters) were manually sliced and immediately frozen in liquid nitrogen, freeze-dried and stored at  $-80^{\circ}\text{C}$  until used. For the chromoplast-enriched suspension studies, fresh intact storage roots were peeled off and processed immediately after harvest.

### Protein Extraction, Quantification and Blot Analysis

One gram of freeze-dried CSR tissue was rehydrated and added with 2 ml of extraction buffer (pH 6.8, 5% SDS, 10% glycerol, 80 mM Tris, and 25 mM DTT), and 20 ml acetone and incubated at  $-20^{\circ}\text{C}$  for more than 1h to obtain a protein precipitate free of pigments. After centrifugation (30000 rpm @  $4^{\circ}\text{C}$  for 20 min) the supernatant was drained, the pellet dried by blowing  $\text{N}_2$  air and suspended with 5 ml extraction buffer (EB). The re-suspended solution was centrifuged again (30000 rpm @  $4^{\circ}\text{C}$  for 10 min) to collect 1ml aliquot of the supernatant that was stored at  $-80^{\circ}\text{C}$  until use. Aliquots (100  $\mu\text{l}$ ) were treated with DOC (0.15%) and TCA (12%) to precipitate soluble proteins. After centrifugation (13000 rpm @  $4^{\circ}\text{C}$  for 20 min) the pellet was suspended in 100  $\mu\text{l}$  of AB, and 50  $\mu\text{l}$  was used for tissue protein estimation by the BCA Micro Kit method (PIERCE) according to the manufacturer protocol. Proteins from crude extract (CE), peak1 and 2 (30  $\mu\text{g}$  each) were separated by 12% SDS-PAGE and blotted onto nitrocellulose as previously described [13]. Transblots were blocked with nonfat dry milk and incubated with primary antibody to Or-protein or Fibrillin-protein at a 1:200 dilution and secondary AP-conjugated goat-anti-rabbit IgG antibody (Bio-Rad) at 1:3000 dilution. Western Blue Stabilized Substrate (Promega) was used for chromogenic detection.

### Preparation of Chromoplast Enriched Suspensions

Fresh CSR was harvested, washed with tap water, sliced, and homogenized with Homogenate Buffer (HB) adjusted to pH 8.2 (100 mM Tris-HCl, 8 mM EDTA, 10 mM KCl, 2 mM  $\text{MgCl}_2$ , 400 mM sucrose, and 1 mM PMSF) using a household blender to obtain a paste that was filtered through

three layers of cheesecloth. The filtrate was centrifuged (600 rpm @  $4^{\circ}\text{C}$  for 30 min) and the resulting supernatant was centrifuged (2000 rpm @  $4^{\circ}\text{C}$  for 30 min) a second time. The resulting pellet was suspended by brushing the surface with 50 mL HB, washed once with HB, followed by centrifugation (2000 rpm @  $4^{\circ}\text{C}$  for 30 min). The final pellet was collected by surface brushing with 50 mL of suspension buffer (SB) adjusted to pH 6.8 (100 mM Tris-HCl, and 250 mM NaCl), named Non-denatured Chromoplast-Enriched Suspension (NDCES), and used for separation of carotenoid-protein complex and non-carotenoid associated proteins by Size Exclusion Chromatography (SEC).

### Size Exclusion Chromatography of NDCES

Chromoplast-enriched suspensions were sonicated for 10 min and 5 mL was loaded onto a column (1 m length x 1.5 cm internal diameter) filled with a resin-bed of Sepharose CL 6B-200 and equilibrated with SB. Fractions (1 mL) were collected at a flow rate of 1.5 mL  $\text{min}^{-1}$  and OD readings @ 280 nm and 461 nm were recorded using a spectrophotometer (Spectra max Plus\_384, Molecular Devices) using UV transparent microplate. Fractions from two observed peaks were pooled separately, named Peak 1 (fractions 33-49, carotenoid-protein complex - CPC) and Peak 2 (fractions 108-122, non carotenoid-protein complex - NCPC), concentrated, and stored at  $-80^{\circ}\text{C}$  until further analyzes.

### HPLC Analysis of Peak 1 Carotenoids

Pooled fractions 33-49 from SEC were treated with acetone to isolate the carotenoid constituent, which was further separated on a C30 column according to [14] to identify the composition of carotenoids present in Peak 1.

### Protein Preparation for LC\_MS/MS Analysis

Pooled protein fractions from peaks 1 and 2 were separated by one-dimensional SDS\_PAGE as previously described [10,15,16] and gels zones were excised and used for in-gel protein digestion with trypsin. Equivalent protein amount (100 mg) from both peak 1 and 2 were individually transferred to a siliconized eppendorf tube and dried in a speed-vac devise. The dried pellet was re-suspended in 200 mL of TCEP solution (10 mM) prepared fresh in ambicarb and incubated at  $56^{\circ}\text{C}$  for one hour and cooled to room temperature. Proteins were then precipitated with 4 volume of acetone after incubation at  $-80^{\circ}\text{C}$  for one hour. After centrifugation at maximum speed for 30 min at  $4^{\circ}\text{C}$ , the pellet was collected and added to 200 mL of Iodacetamide (55 mM) in ambicarb, and incubated for one hour at room temperature in the dark. The sample was incubated at  $-80^{\circ}\text{C}$  for one hour with 4 volumes acetone and then centrifuged (maximum speed @  $4^{\circ}\text{C}$  for 30 min). The resulting pellet was incubated with 100 mL of ambicarb by vortexing and sonicating for one hour. A trypsin solution (20 mg  $\text{mL}^{-1}$  in ambicarb) was added to have a protease:protein ration of 1:50, incubated at  $37^{\circ}\text{C}$  for the first two hours, after which an additional 100 mL of the same trypsin solution was added for incubation over night. After cooling to room temperature the sample was used for mass spectrometry analysis.

## LC-MS/MS Analysis

After initial protein digestion, peptides were ionized, detected, isolated, fragmented and sequenced in a qTOF MALDI MS/MS. The operating software generated MS and MS/MS spectra, and the peptide sequences were determined using the PROTEOMIC facilities and software available at USDA/ARS-National Animal Disease Center in Ames, Iowa.

## Bioinformatics Analysis of Peptide and Protein Sequence Data

Peptide sequences were established using a MASCOT server database ([http://www.matrixscience.com/cgi/search\\_form.pl?FORMVER=2&SEARCH=PMF](http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=PMF)) and statistical validation as well as protein sequence matching identification. Peptide sequences that produced no protein hits were further analyzed using *tblastx* and translation tools available in ExPASy at Proteomics Server (<http://www.expasy.org/>) using four cassava EST databases located at CIAT (<http://webapp.ciat.cgiar.org/biotechnology/bioinformatics.htm>), ESTIMA ([http://titan.biotech.uiuc.edu/cgi-bin/ESTWebsite/estima\\_start?seqSet=cassava](http://titan.biotech.uiuc.edu/cgi-bin/ESTWebsite/estima_start?seqSet=cassava)), EMBRAPA (<http://www.genoma.embrapa.br/>), and RIKEN (<http://www.brc.riken.go.jp/lab/epd/Eng/species/manihot.shtml>). Finally, the protein fragment sequences derived from the EST matching results were mapped to an Arabidopsis protein database [<http://www.arabidopsis.org>] and to cassava PROTEOME and GENOME databases [<http://www.phytozome.net/cassava>]. Information on cassava proteome loci, for identified protein, and genome scaffold (for their corresponding gene) were obtained to characterize their location and distribution of identified proteins in the carotenoid-protein complex.

## RNA Isolation, Quantification, and cDNA Synthesis

Total RNA was extracted twice using the phenol chloroform procedure described previously [17] and treated with RNase free DNase. Total RNA was quantified using the Quant-It RiboGreen RNA Kit according to the manufacturer (Molecular Probe). Total RNA (400 ng) was reverse transcribed in a 20  $\mu$ L reaction volume using a SuperScript III Platinum Two-Step qRT-PCR Kit (Invitrogen). The reactions were terminated by heat inactivation and the cDNA products were treated with RNase H and stored at  $-20^{\circ}\text{C}$ .

## Gene-Specific Primers, Certified Standard and Housekeeping Gene Primers

Fluorogenic primers (FAM labeled LUX primer) and corresponding unlabeled primers were designed using the LUX Designer-Desktop version (Invitrogen) for each sequence of the cDNA fragment coding for genes as shown in Table S1 (see supplementary data). Standard certified primers for 18S ribosomal (Invitrogen Cat. 115HM-02) and Gus (Invitrogen Cat. 112H-02), and qPCR plasmid Standards (Invitrogen Cat. 11741-100) with Gus ORF were used as internal controls and quantitative standards to generate standard curves respectively. All primers were synthesized and purchased from Invitrogen.

## Quantitative Real Time PCR Analysis

Gene sequences coding for proteins identified from the carotenoid-protein complexes were used to quantify transcript abundance in white and intense yellow CSR using (qRT-PCR). The qRT-PCR assays were performed in triplicate for each extraction using a Bio-Rad system (BioRad model iCycler). Ribosomal RNA (18S) was used as an internal control and PCR efficiency was evaluated using GUS certified primers to obtain a quantitative standard curve for absolute expression analysis of target transcript level in samples. Each experiment was replicated at least three times. Alternatively, a relative expression assay using MEC 1 (gene coding for Pt2L4 protein, 17) was used as an internal control. qRT-PCR products were sequenced at the EMBRAPA Genetic Resources and Biotechnology Genome facilities.

## Data Analysis

To access the intrinsic property for the level of each transcript corresponding to the target gene in all samples, qRT-PCR amplification procedures consisted of absolute quantification expression using total RNA as normalizer. These methods provided an absolute quantitative standard curve for a dilution series based on the qPCR plasmid standards kit protocol (Invitrogen). Preliminary experiments with two unknown samples were performed to set up qRT-PCR amplification optimization conditions, including reliable exponential phase of amplification, qRT-PCR efficiency ranging from 90% to 105%, Ct threshold and baseline, and evaluation of each primer sets for target gene within an unknown sample. After setting up the optimized conditions, Ct values for all unknown samples and replications were interpolated from the standard curve estimating the level of transcript in unknown samples. In addition, a five point curve made with each amplicon was performed to obtain the efficiency of each primer (needed for the  $\Delta\Delta\text{Ct}$  formula) and to know the range of amplification. After setting up the optimized conditions, Ct values for all unknown samples and replications were collected and normalized with housekeeping genes (internal controls) using the  $\Delta\Delta\text{Ct}$  formula. Statistical parameters, such as mean and mean standard deviation, were performed in Excel.

## RESULTS

### Carotenoid-protein Complex Isolation and Characterization

Proteins from non-denatured chromoplast-enriched suspension (CES) present in white and intense yellow cassava, containing chromoplast-associated proteins (CAP) were separated by SEC Fig. (1A and 1B), respectively.

Size exclusion chromatography profiles monitored with Abs<sub>461</sub> (carotenoids) and Abs<sub>280</sub> (proteins) indicated the presence of two major peaks. Peak 1 (fractions 33-49) eluted within the void volume of the column, as determined with blue dextran (MW 2000 kDa) and overlap between the carotenoid and protein profiles indicates that Peak 1 represents a carotenoid-protein complex (CPC) with an estimated size close to 400 kDa. Peak 2 (fractions 108-122) eluted later in



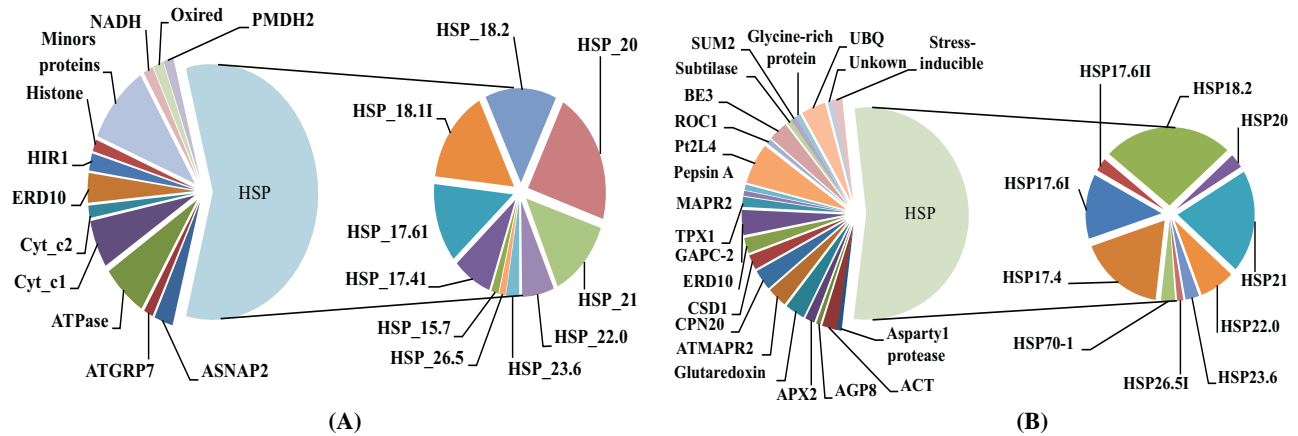


Fig. (3). Distribution of proteins present in the carotenoid-protein complex (A) and non-carotenoid associated protein (B).

This differential distribution of HSPs between these two fractions may indicate (1) differential properties of these HSPs to complex with carotenoids, (2) differential solubility in the SDS-PAGE gel separation procedure used in the present work, or (3) aggregation dynamics of a particular HSP *in vitro* or *in vivo* as reported by others [21].

**Occurrence of Fibrillin- and Or-protein in White and Intense Yellow Cassava**

Protein-blot analysis was used to test for the presence of Fibrillin- and Or-proteins as markers for fibrillar structure of chromoplast occurrence in CSR as suggested in others plant systems [7,22], as well as chromoplast formation and differentiation [23,24] respectively. Fibrillin- (Fig. 4A) and Or-protein (Fig. 4B) are present in CES and NCPC of intense yellow cassava, but not in CPC isolated fractions or CES of the white cassava. These results suggest that the type of chromoplast structure accumulating carotenoid in cassava is of fibrillar-type. Since Fibrillin- and Or-proteins were not detected in the CES or in CPC of white cassava, these results also indicate that intense yellow root accumulate more chromoplasts than white roots of cassava. At the same time, these results suggest that further microscopic studies will be needed to confirm the form, number and the behavior of chromoplast in CSR, which was not addressed in the present work.

**Assign Assessed CAP Sequences to Cassava PROTEOME and GENOME**

List of peptide and identified proteins sequences present in CPC are showed in Supplementary data in Table S2 and Supplementary data in Table S3 respectively. List of peptide and identified proteins sequences present in NCPC are showed in Supplementary data in Table S4 and Supplementary data in Table S5 respectively. The identified proteins were assigned to cassava PROTEOME and GENOME databases and results are summarized in Tables 1 to 4.

It was observed that 30 identified proteins in CPC were distributed in 28 cassava-protein loci and 19 families while 33 proteins in NCPC are distributed in 44 loci and 25 families. The HSP group individuals for CPC are located in 9 loci

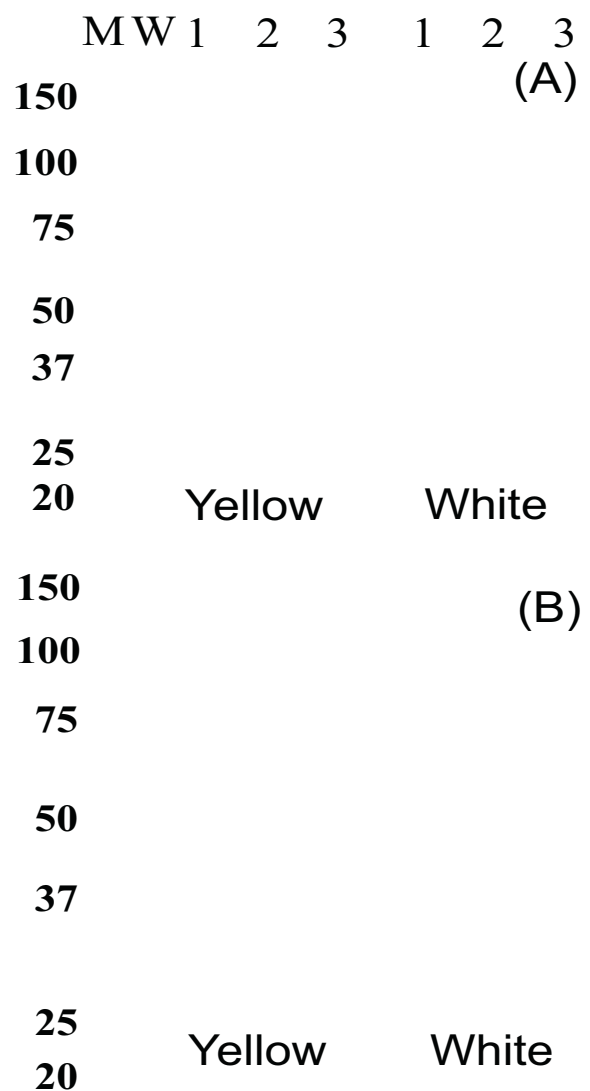


Fig. (4). Western blot analysis of proteins (35µg/lane) extracts using antiserum raised against Fibrillin- (A) and Or-protein (B). Line 1 is chromoplast enriched suspension. Line 2 is total protein fraction from CPC. Line 3 is total protein fraction from NCAP identified in the profile of size exclusion chromatography as shown in Fig. (1).

**Table 1. Identified Proteins in the Carotenoid-Protein Complex and their Distribution in the Cassava Genome**

Cassava Genome Scaffold	Score	E_Value	Protein Name Abbreviation
scaffold06948	971.5	0	ACT7
scaffold09150	836.2	0	Actin
scaffold10689	291.6	4.9e-77	Alpha_SNAP2
scaffold04718	513.4	6.9e-144	ATGRP7
scaffold03604	1004.0	0	ATPase
scaffold04355	921.0	0	ATPase
scaffold03264	73.4	2.7e-11	ATPsyn
scaffold04355	161.8	4.5e-38	ATPsyn
scaffold04355	248.3	1.1e-63	ATRZ-1A
scaffold06708	629.8	3.4e-179	BE
scaffold08729	132.9	1.8e-29	COX2
scaffold06028	499.0	1.8e-139	Cyto_c1
scaffold06697	502.6	1.8e-140	Cyto_c1
scaffold05875	122.1	6.8e-26	Dihy_acet
scaffold06708	747.9	0	ERD10
scaffold08053	302.4	2.2e-80	HIR1
scaffold12794	261.0	6.6e-68	HIR1
scaffold06656	244.7	4.4e-63	Histone
scaffold05875	731.7	0	HSP_17.4I
scaffold06512	1292.5	0	HSP_17.6I
scaffold06908	949.9	0	HSP_17.6I
scaffold06512	868.7	0	HSP_18.1I
scaffold06908	1323.2	0	HSP_18.1I
scaffold06512	868.7	0	HSP_18.2
scaffold06908	1323.2	0	HSP_18.2
scaffold08265	1545.0	0	HSP_18.2
scaffold06512	868.7	0	HSP_20
scaffold07543	1041.8	0	HSP_21
scaffold09151	1204.1	0	HSP_22.0
scaffold06701	1162.7	0	HSP_23.6
scaffold04151	881.3	0	HSP_26.5
scaffold04355	161.8	4.5e-38	HSP17.6
scaffold06512	358.4	4.0e-97	HSP20
scaffold06908	949.9	0	HSP20
scaffold08265	250.2	3.0e-65	HSP20
scaffold00069	978.7	0	HSP21

Table 1. Contd.....

Cassava Genome Scaffold	Score	E_Value	Protein Name Abbreviation
scaffold06701	1162.7	0	HSP23.6
scaffold02264	149.2	2.3e-34	MAB1
scaffold00069	1038.2	0	NADH
scaffold10097	468.4	3.2e-130	OEP16
scaffold07528	1402.5	0	Oxireductase
scaffold07851	834.4	0	PATL1
scaffold06782	573.0	2.2e-161	PMDH2
scaffold06089	719.0	0	RBP31
scaffold00325	1031.0	0	RRM_1
scaffold07851	383.6	8.2e-105	SEC14
scaffold02892	798.4	0	snRNP
scaffold12794	1123.0	0	snRNP

Table 2. Identified Proteins in the Carotenoid-Protein Complex Assigned to Cassava Proteome Locus and their Distribution in Protein Families Assigned by *Pfam* Search

Pfam	Manihot Locus	Protein Name Abbreviation	Score	E_Value
PF00022	cassava4.1_033108m	ACT7	312.0	4.8e-85
PF00022	cassava4.1_033108m	Actin	350.5	1.1e-96
PF00515	cassava4.1_013135m	ASNAP2	367.5	2.8e-102
PF10559	cassava4.1_006793m	ATPase	194.5	3.5e-50
PF00306	cassava4.1_004726m	ATPsyn	228.0	3.2e-60
PF02806	cassava4.1_001595m	BE	1726.8	0
PF02167	cassava4.1_012354m	Cyt_c1	425.6	1.1e-119
PF00198	cassava4.1_004864m	Dihy_acet	440.3	5.2e-124
PF00257	cassava4.1_015875m	ERD10	116.7	7.3e-27
PF04716	cassava4.1_017840m	RBP31	336.7	1.9e-92
PF01145	cassava4.1_013221m	HIR1	252.7	6.6e-68
PF00125	cassava4.1_019911m	Histone	167.5	2.1e-42
PF00011	cassava4.1_019659m	HSP_15.7	225.7	1.6e-59
PF00011	cassava4.1_018038m	HSP_17.4I	164.1	1.5e-61
PF00011	cassava4.1_018127m	HSP_17.6I	261.2	3.8e-70
PF00011	cassava4.1_017871m	HSP_18.1I	291.2	3.8e-79
PF04969	cassava4.1_017974m	HSP_18.2	266.9	9.5e-72
PF00011	cassava4.1_017871m	HSP_20	284.3	2.7e-77
PF00011	cassava4.1_015256m	HSP_21	396.4	1.1e-110
PF04969	cassava4.1_032436m	HSP_22.0	352.8	9.6e-98
PF00011	cassava4.1_016267m	HSP_23.6	382.9	1.3e-106

Table 2. Contd.....

Pfam	Manihot Locus	Protein Name Abbreviation	Score	E_Value
PF00011	cassava4.1_015518m	HSP_26.5	305.8	2.3e-83
PF01423	cassava4.1_019986m	snRNP	58.9	2.8e-9
PF02779	cassava4.1_010415m	MAB1	237.3	1.6e-63
PF02466	cassava4.1_018513m	OEP16	256.1	9.4e-69
PF00107	cassava4.1_031433m	Oxireductase	377.5	4.6e-105
PF02866	cassava4.1_010611m	PMDH2	612.1	0
PF00076	cassava4.1_012182m	RNP	304.7	6.5e-83
PF00076	cassava4.1_019294m	RRM_1	161.4	4.5e-40
PF03765	cassava4.1_003845m	SEC14	309.3	6.2e-85

Table 3. Identified Proteins in the Non-Carotenoid-Protein Complex and their Distribution in the Cassava Genome

Cassava Scaffold	Total Score	E_Value	Protein Name Abbreviation
scaffold09150	839.9	0	ACT7
scaffold03602	206.9	2.2e-51	ACT8
scaffold03942	917.4	0	AGP8
scaffold06708	482.8	1.0e-134	Amylase
scaffold00506	398.0	8.5e-109	APX2
scaffold06844	1067.1	0	APX2
scaffold00069	978.7	0	Aspartyl protease
scaffold00847	906.6	0	ATMAPR2
scaffold06708	482.8	1.0e-134	BE
scaffold00506	398.0	8.5e-109	CPN20
scaffold10114	462.9	2.3e-128	CPN20
scaffold04285	621.6	2.8e-176	CSD1
scaffold06705	1104.9	0	CSD1
scaffold06844	1067.1	0	ERD10
scaffold04953	627.1	1.3e-177	GAPC
scaffold06948	720.8	0	GAPC
scaffold07543	122.1	6.2e-26	GAPC
scaffold03245	145.6	6.2e-33	GAPC
scaffold07660	356.5	1.3e-96	Glutaredoxin
scaffold07762	1103.1	0	Glycine-rich protein
scaffold05255	919.2	0	HSP17.4
scaffold03651	1106.8	0	HSP17.4
scaffold01127	1005.8	0	HSP17.4
scaffold07859	1440.4	0	HSP17.4
scaffold06908	502.3	4.3e-141	HSP17.4



Table 3. Contd.....

Cassava Scaffold	Total Score	E_Value	Protein Name Abbreviation
scaffold10469	1287.1	0	HSP17.4
scaffold12439	982.3	0	HSP17.4
scaffold09151	1204.1	0	HSP17.4
scaffold06512	1224.0	0	HSP17.6I
scaffold05255	919.2	0	HSP17.6I
scaffold10469	1287.1	0	HSP17.6II
scaffold07859	1440.4	0	HSP17.6II
scaffold06908	1285.3	0	HSP18.2
scaffold05875	731.7	0	HSP18.2
scaffold05255	919.2	0	HSP18.2
scaffold12439	1561.2	0	HSP18.2
scaffold08265	1545.0	0	HSP18.2
scaffold06908	1285.3	0	HSP18.2
scaffold06512	392.6	2.7e-107	HSP20
scaffold08265	250.2	3.0e-65	HSP20
scaffold00069	1038.2	0	HSP21
scaffold03241	872.3	0	HSP21
scaffold00467	596.4	1.1e-168	HSP21
scaffold02817	1178.9	0	HSP21
scaffold06908	502.3	4.3e-141	HSP22.0
scaffold09151	1204.1	0	HSP22.0
scaffold00069	978.7	0	HSP23.6
scaffold06512	868.7	0	HSP23.6
scaffold04151	881.3	0	HSP26.5I
scaffold08265	571.2	3.8e-161	HSP70-1
scaffold08265	250.2	3.0e-65	MAPR2
scaffold04175	419.7	1.0e-115	NAD binding
scaffold06598	378.2	3.4e-103	Pepsin A
scaffold06844	1067.1	0	Pt2L4
scaffold07859	1440.4	0	Pt2L4
scaffold11638	1393.5	0	ROC1
scaffold08559	903.0	0	Stress-inducible
scaffold00467	520.7	6.1e-146	Subtilase
scaffold11661	414.3	3.4e-114	SUM2
scaffold00847	780.3	0	TPX1
scaffold10669	699.2	0	UBQ10
scaffold12498	558.5	2.3e-157	UBQ2

Table 3. Contd.....

Cassava Scaffold	Total Score	E_Value	Protein Name Abbreviation
scaffold03055	1043.6	0	UBQ2
scaffold10669	699.2	0	UBQ3
scaffold00069	978.7	0	UBQ6
scaffold07591	749.7	0	Unkown

Table 4. Identified Proteins in Non-carotenoid-protein Complex Assigned to Cassava Proteome Locus and their Distribution in Protein Families Assigned by Pfam Search

Pfam	Cassava Locus	Protein Name Abbreviation	Total Score	E_Value
PF00022	cassava4.1_033108m	ACT7	371.3	3.4e-103
PF00022	cassava4.1_033108m	ACT8	448.7	1.5e-126
PF02806	cassava4.1_001595m	AMYLASE	255.8	6.2e-69
PF00141	cassava4.1_014643m	APX2	484.6	3.8e-137
PF03489	cassava4.1_005735m	Aspartyl protease	517.7	4.6e-147
PF00173	cassava4.1_019974m	ATMAPR2	203.8	6.6e-53
PF00166	cassava4.1_014439m	CPN20	461.5	3.2e-130
PF00080	cassava4.1_018289m	CSD1	278.9	1.8e-75
PF00257	cassava4.1_015875m	ERD10	188.7	3.7e-48
PF02469	cassava4.1_031631m	Fasciclin	323.2	4.5e-89
PF00044	cassava4.1_011175m	GAPC	402.5	1.6e-112
PF00044	cassava4.1_011176m	GAPC	287.0	6.3e-78
PF00044	cassava4.1_011326m	GAPC	394.4	2.9e-110
PF00044	cassava4.1_011366m	GAPC	634.4	0
PF00462	cassava4.1_019777m	glutaredoxin	203.0	7.9e-53
PF01277	cassava4.1_016619m	glycine-rich protein	243.4	7.0e-65
PF00011	cassava4.1_015256m	HSP17.4	403.3	9.5e-113
PF00011	cassava4.1_016281m	HSP17.4	302.8	1.5e-82
PF04969	cassava4.1_017267m	HSP17.4	295.0	9.6e-80
PF00011	cassava4.1_017871m	HSP17.4	291.2	3.8e-79
PF00011	cassava4.1_018066m	HSP17.4	241.1	2.4e-64
PF00011	cassava4.1_018093m	HSP17.4	281.6	3.8e-76
PF00011	cassava4.1_018121m	HSP17.4	305.4	1.4e-83
PF00011	cassava4.1_018127m	HSP17.4	261.2	4.0e-70
PF00011	cassava4.1_018158m	HSP17.4	267.7	5.5e-72
PF00011	cassava4.1_018376m	HSP17.4	295.8	1.0e-80
PF00011	cassava4.1_017871m	HSP17.6I	291.2	3.8e-79
PF00011	cassava4.1_018127m	HSP17.6I	261.2	3.8e-70
PF00011	cassava4.1_018131m	HSP17.6I	167.2	2.1e-56

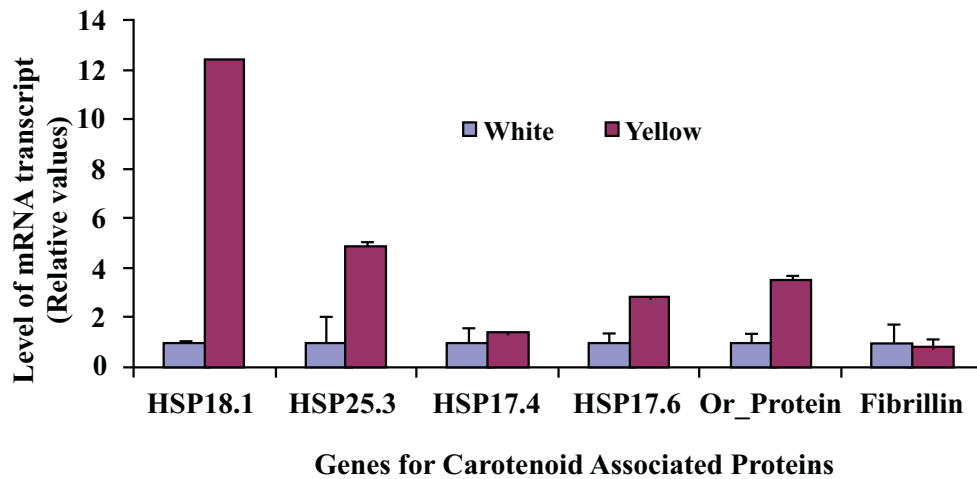
Table 4. Contd.....

Pfam	Cassava Locus	Protein Name Abbreviation	Total Score	E_Value
PF00011	cassava4.1_018158m	HSP17.6II	267.7	5.5e-72
PF00011	cassava4.1_017871m	HSP18.2	291.2	3.8e-79
PF00011	cassava4.1_018038m	HSP18.2	164.1	1.5e-61
PF00011	cassava4.1_018066m	HSP18.2	241.1	2.4e-64
PF00011	cassava4.1_017871m	HSP20	291.2	3.8e-79
PF00011	cassava4.1_018131m	HSP20	235.0	3.7e-62
PF00011	cassava4.1_015256m	HSP21	396.4	1.1e-110
PF00011	cassava4.1_018066m	HSP21	241.1	2.4e-64
PF00011	cassava4.1_017871m	HSP22.0	291.2	3.8e-79
PF00011	cassava4.1_018066m	HSP22.0	241.1	2.4e-64
PF00011	cassava4.1_015256m	HSP23.6	396.4	1.1e-110
PF00011	cassava4.1_015518m	HSP23.6	305.8	2.3e-83
PF04969	cassava4.1_017267m	HSP23.6	295.0	9.6e-80
PF04969	cassava4.1_017974m	HSP23.6	266.9	9.5e-72
PF04969	cassava4.1_022774m	HSP23.6	226.9	1.1e-59
PF04969	cassava4.1_032436m	HSP23.6	352.8	9.6e-98
PF00012	cassava4.1_003331m	HSP70-1	194.5	9.2e-59
PF00012	cassava4.1_003343m	HSP70-1	379.0	3.0e-105
PF05193	cassava4.1_005898m	pepsin A	274.2	1.3e-73
UNKOWN	cassava4.1_017439m	Pt2L4	57.0	1.8e-8
PF00160	cassava4.1_017662m	ROC1	323.9	6.3e-89
PF00515	cassava4.1_004441m	stress-inducible	305.8	1.3e-83
PF05922	cassava4.1_019197m	subtilase	221.5	2.9e-58
PF00240	cassava4.1_019468m	SUM2	117.1	2.1e-27
PF00578	cassava4.1_017973m	TPX1	322.0	1.9e-88
PF00240	cassava4.1_019139m	UBQ10	247.3	4.5e-66
PF07708	cassava4.1_009650m	UBQ2	313.9	2.6e-160
PF11976	cassava4.1_018181m	UBQ2	266.5	6.5e-72
PF00240	cassava4.1_019139m	UBQ3	247.3	4.6e-66
UNKOWN	cassava4.1_015690m	Unkown	287.0	1.1e-77

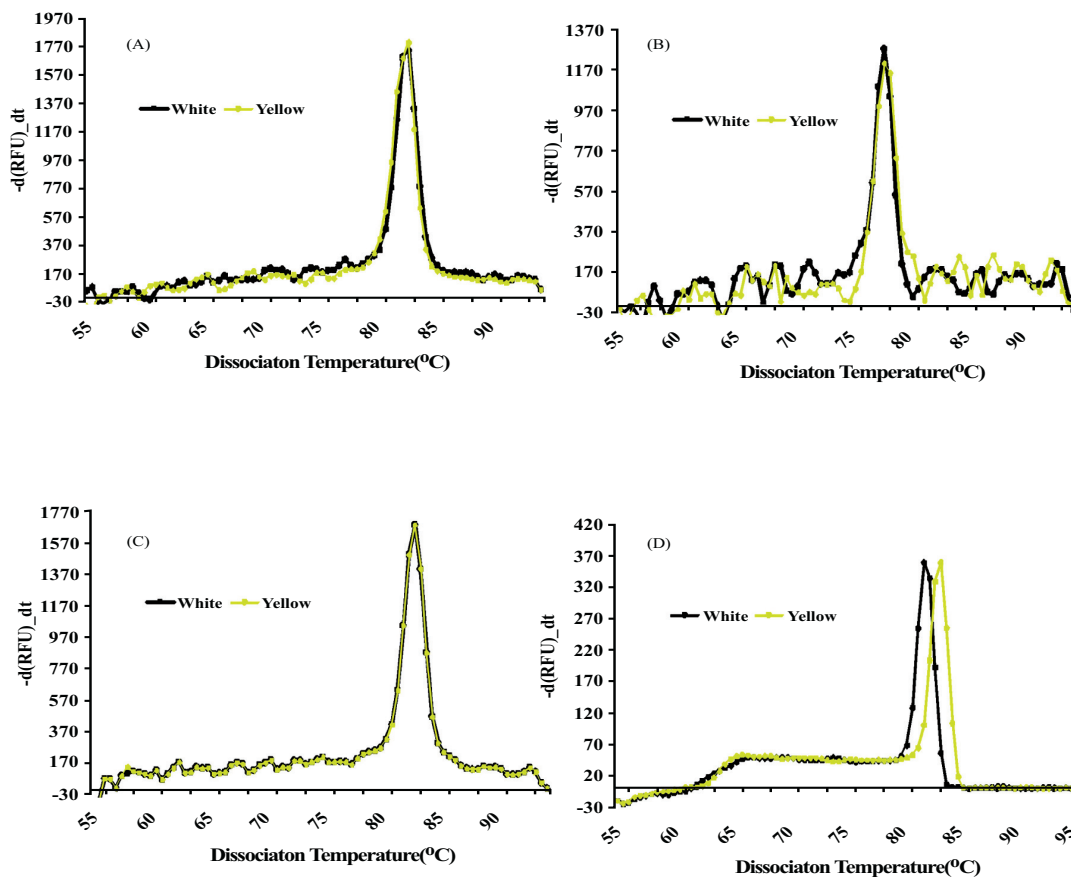
and distributed in 2 protein families. The gene coding for those proteins are distributed in 30 scaffolds for the CPC and 43 for the NCPC in the cassava GENOME. For the case of HSP, the genes coding for 10 HSP identified in CPC and in NCPC showed different distribution pattern, being 11 scaffolds for the CPC group and 15 scaffolds for the NCPC group. These results contribute to the overall observation about the functional complexity of HSPs in plants as well as the distribution and arrangement of their genes in the cassava genome.

### Gene Expression Analysis

Expression of two genes, one related to chromoplast structure definition (Fibrillin gene) and another related to chromoplast development (Or-gene coding for Or-protein), were analyzed to gain insights on abundance and formation of chromoplast in CSR. Four identified gene sequences coding for proteins in the CPC (HSP18.1, HSP17.6, HSP17.4 and HSP21) were used to search for their possible roles in differential sequestration of  $\beta$ -carotene in white and intense



**Fig. (5).** Differential levels of transcript for genes coding for a set of proteins related to carotenoid accumulation in white (A) and intense yellow (B). Relative values expressed in relation to white roots.



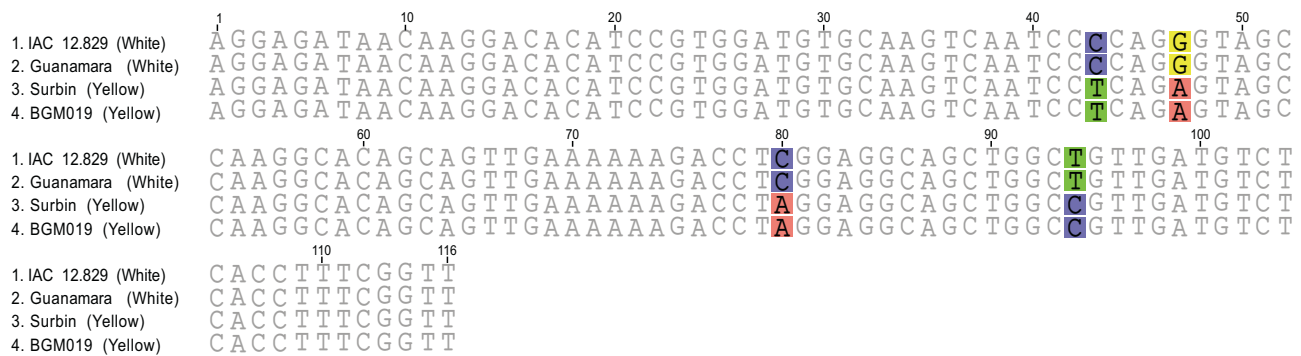
**Fig. (6).** Dissociation curve derived from qRT-PCR amplification for genes coding for Fibrillin (A), Or-protein (B), CAP1 (C), CAP2 (D), present in CPC of white and intense yellow cassava storage root.

yellow genotypes. Transcript levels for the above listed genes were more abundant in the intense yellow than in white CSR (Fig. 5) and followed the pattern of carotenoid accumulation.

The dissociation temperature curve patterns shown in Fig. (6) of the qRT-PCR amplification of the four genes listed above indicates a change in nucleotide coding sequence for HSP21 (Fig. 6D), which is associated with high amounts of carotenoid accumulation in the intense yellow

landrace. The overall observation about the functional complexity of HSPs in plants as well as their distribution and arrangements of their genes in the cassava genome. Indeed, the sequence of the corresponding amplified fragment (Fig. 7) revealed four single point mutations in the gene coding for this protein that may favor carotenoid sequestration in the CSR of intense yellow phenotype.

In addition, these results corroborate those observed by protein blot analysis where Fibrillin- and Or-protein support



**Fig. (7).** Sequence comparison for the amplicon resulting from qRT-PCR of HSP21 from white and intense yellow cassava storage root.

a role for enhancement of protein carotenoid sink capacity (i.e. chromoplast formation and number) as well as a possible specific role for HSP21 in the carotenoid sequestration mechanism.

## DISCUSSION

Carotenoid-protein complex (CPC) formation in specialized supramolecular structures inside the chromoplast has been associated with several molecular functional interactions [1,24], including sequestration of specific carotenoids that aid in driving the carotenoid synthesis pathway towards completion. The characterization of CPC protein composition in chromoplast of cassava storage root (CSR) identified HSPs as the major type of proteins present in the complex. Further, specific markers for proteins (Fibrillin) used in this study indicate chromoplasts of intense yellow cassava storage root are of fibrillar type capable of chromoplast differentiation [25-27] and carotenoid accumulation [25]. Results from this study also suggest that a specific haplotype of HSP21 could explain the differential carotenoid accumulation in intense yellow cassava landraces.

HSPs are classes of proteins that are unique, abundant and diverse in plants, especially in non-green tissue [28,29]. Higher plants have at least 20 small Heat Shock Proteins (sHSPs) and some species may have up to 40 different sHSPs [30]. The sHSPs are involved in response to environmental stress such as heat, cold, drought and salinity stress [31,32]. In the absence of environmental stress, sHSPs expression in plants are restricted to certain stages of development such as embryogenesis, seed germination, pollen development, and fruit maturation [33,34]. In the case of CSR, we observed that the proteins forming the CPC in root chromoplast are mainly composed of the sHSP family with variable levels of expression, depending on the type of carotenoid accumulating genotype. For instance, transcripts for HSP18.1 and HSP21 are more abundant in intense yellow cassava storage root where values of  $\beta$ -carotene accumulation were 100x more than in white root. While HSP18.1 may be responsible for the enhancement of total carotenoid accumulation, HSP21 may also play a role in the specificity to  $\beta$ -carotene accumulation throughout the sequestration mechanism as observed in tomato [35]. In addition, a new heteroduplex double strand cDNA sequence for the gene coding for HSP21 was observed, suggesting that an isoform of this protein is coded by a different gene haplotype possessing four single point mutations in the sequence of the amplicon (Fig. 7) that may be responsible for the sequestration

of high  $\beta$ -carotene in intense yellow CSR. In other plant systems, the presence of a specific HSP21 protein, has been speculated to act as an antioxidant under oxidative stress [36] or as an enhancer of carotenoid accumulation in tomato fruit of a transgenic plant [35]. More in-depth research on the genotypes enriched with special carotenoids from other cassava landraces, will have a larger impact on carotenoid genetic improvement in new varieties and strengthen our knowledge of protein content enhancement and carotenoid regulation in non-green tissue.

Fibrillin has been found in a number of plants and plays a structural role in carotenoid-storing fibrillar type chromoplasts [37]. This prompted us to examine the presence of fibrillin in CSR, which could help define the type of ultramolecular structure of chromoplast present in CSR. Here we demonstrated that fibrillin antibody recognized a putative cassava fibrillin protein in CES of intense yellow but not in white cassava storage root. This evidence supports the presence of carotenoid-fibril type in CSR, rather than other type of structures observed in roots such as carrot [18] that also store high amounts of carotenoids. Based on the discovery of Or-protein and its cellular function in inducing formation of chromoplasts [25,38], we used Or-protein antibody to recognize this protein in CSR as a possible indicator of the differential number of chromoplasts between intense yellow and white CSR. Indeed, Or-protein was observed in CES of intense yellow but not in white root. This evidence indicates the potential for more chromoplasts present in intense yellow CSR creating a metabolic sink for carotenoid accumulation, similar to that observed in transgenic potato tuber overexpressing the Or-gene that accumulate greater levels carotenoids [39].

## CONCLUSION

In conclusion, this study provides evidence that the carotenoid-protein complex, in cassava storage root, is enriched with HSPs that interact with different carotenoids, depending on the genetic background of cassava landrace. The presence of Fibrillin- and Or-protein may confirm the fibrillar type of chromoplast and intense chromoplast formation in storage root of intense yellow compared to white cassava CSR. The novel allele of HSP21, identified in the present work, may be an important tool for genetic manipulation of protein content in cassava storage root contributing to a natural sink for protein and carotenoid accumulation in intense yellow CSR. Further elucidation of the roles of HSP21 in binding specifi-

cally with  $\beta$ -carotene or in inducing accumulation of carotenoid in CSR would provide a significant contribution to our understanding of the molecular machinery controlling carotenoid accumulation in non-green tissues.

## ACKNOWLEDGEMENTS

Financial support from EMBRAPA/LABEX-US, in collaboration with the USDA-ARS, Red River Valley Agricultural Research Center in Fargo, ND, is acknowledged. Special acknowledges are also extended for the financial support provided by the Conselho Nacional de Desenvolvimento Científico e Tecnológico-CNPq (grant number 680.410/01-1), the National Biotechnology Program of EMBRAPA (grant number 06.03.02.058), International Atomic Energy Agency (grant number BRA\_13188), State Administration of Foreign Experts Affairs, China (grant number Y20100-326001) as well as the KLUGRTC, MA/CATAS\_TCGRI from China (grant number KFKT\_2010\_01).

## CONFLICTS OF INTEREST

There are no conflicts of interest and the authors agree not to withdraw their manuscript at any stage prior to publication.

## SYMBOLS AND ABBREVIATIONS

PAP	=	plastid lipid-associated proteins
CAP	=	carotenoid associated proteins
EB	=	extraction buffer
SB	=	Suspension Buffer
CES	=	Chromoplast enriched suspension
SEC	=	Size exclusion chromatography
CPC	=	carotenoid-protein complex
NDCES	=	Non-denatured Chromoplast-Enriched Suspension
NCPC	=	non carotenoid-protein complex
CSR	=	cassava storage root

## URL ADDRESSES OF SUPPORT ANALYTICAL SOFTWARE AND DATA BASE

MASCOT server	=	<a href="http://www.matrixscienc.com/cgi/search_form.pl?FORMER=2&amp;SEARCH=PMF">http://www.matrixscienc.com/cgi/search_form.pl?FORMER=2&amp;SEARCH=PMF</a>
ExpASy	=	<a href="http://www.expasy.org/">http://www.expasy.org/</a>
NCBI	=	<a href="http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucst&amp;cmd=search&amp;term=carvalho%20cassava">http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucst&amp;cmd=search&amp;term=carvalho%20cassava</a>
CIAT data base	=	<a href="http://webapp.ciat.cgiar.org/biotechnology/bioinformatics.htm">http://webapp.ciat.cgiar.org/biotechnology/bioinformatics.htm</a>
ESTIMA data base	=	<a href="http://titan.biotech.uiuc.edu/cgi-bin/ESTWebsite/estima_start?seqSet=cassava">http://titan.biotech.uiuc.edu/cgi-</a>

		<a href="http://titan.biotech.uiuc.edu/cgi-bin/ESTWebsite/estima_start?seqSet=cassava">bin/ESTWebsite/estima_start?seqSet=cassava</a>
EMBRAPA data base	=	<a href="http://www.genoma.embrapa.br/">http://www.genoma.embrapa.br/</a>
RIKEN data base	=	<a href="http://www.brc.riken.go.jp/lab/epd/Eng/species/manihoto.shtml">http://www.brc.riken.go.jp/lab/epd/Eng/species/manihoto.shtml</a>
Cassava genome data base	=	<a href="http://www.phytozome.net/cassava">http://www.phytozome.net/cassava</a>
Arabidopsis data base	=	<a href="http://www.arabidopsis.org/">http://www.arabidopsis.org/</a>
Pfam Server	=	<a href="http://pfam.sanger.ac.uk/">http://pfam.sanger.ac.uk/</a>

## REFERENCES

- Camara, B.; Huguency, P.; Bouvier F.; Kuntz M.; Moneger, R. Biochemistry and molecular Biology of chromoplast development. *Int. Rev. Cytol.*, **1995**, *163*, 175-235.
- Deruere, J.; Romer, S.; D'Harlingue, A.; Backhouse, R.A.; Kuntz, M.; Camara, B. Fibril assembly and carotenoid over accumulation in chromoplasts: a model for supramolecular lipoprotein structures. *Plant Cell*, **1994**, *6*, 119-133.
- Smirra, I.; Halevy, A.H.; Vainstein, A. Isolation and Characterization of a Chromoplast-Specific Carotenoid-Associated Protein *Cucumis sativus* Corollas. *Plant Physiol.*, **1993**, *102*, 491-496.
- Mishra, S.; Sanwal, G.G. Effect of *Cuscuta* infection on chloroplast lipid composition of Brassica leaves. *Eur. J. Plant Pathol.* **1994**, *100*, 61-70.
- Mishra, S.; Tyagil, A.; Singhl, I.V.; Sangwan, R.S. Changes in lipid profile during growth and senescence of *Catharanthus roseus* leaf. *Braz. J. Plant Physiol.*, **2006**, *18*(4), 447-454.
- Libal-Weksler, Y.; Vishnevetsky, M.; Ovadis, M.; Vainstein, A. Isolation and regulation of accumulation of a minor chromoplast-specific protein from cucumber corollas. *Plant Physiol.*, **1997**, *113*, 59-63.
- Vishnevetsky, M.; Ovadis, M.; Zuker, A.; Vainstein, A. Molecular mechanisms underlying carotenogenesis in the chromoplast: multi-level regulation of carotenoid-associated genes. *Plant J.*, **1999**, *20*, 423-431.
- Vishnevetsky, M.; Ovadis, M.; Vainstein, A. Carotenoid sequestration in plants: the role of carotenoid-associated proteins. *Trends Plant Sci.*, **1999**, *4*, 232-235.
- Monte, E.; Ludevid, D.; Prat, S. Leaf 40.4: A carotenoid-associated protein involved in the modulation of photosynthetic efficiency? *Plant J.*, **1999**, *19*(4), 399-410.
- Lippolis, J.D.; Reinhardt, T.A. Proteomic survey of bovine neutrophils. *Vet. Immunol. Immunopathol.*, **2005**, *103*, 53-65.
- Resing, K.A.; Ahn, N.G. Proteomics strategies for protein identification. *FEBS Lett.*, **2005**, *579*, 885-889.
- Nesvizhskii, A.I.; Keller, A.; Kolker, E.; Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.*, **2003**, *75*, 4646-4658.
- Carvalho, L.J.C.B. *Maze and invertases: Purification, characterization, localization and function*. PhD Thesis, Cornell University, Ithaca (NY), August 2000.
- Kimura, M.; Kobori, C.N.; Rodriguez-Amaya, D.B.; Nestel, P. Screening and HPLC methods for carotenoids in sweet potato, cassava and maize for plant breeding trials. *Food Chem.*, **2005**, *100*, 1734-1746.
- Reinhardt, T.A.; Lippolis, J.D. Bovine milk fat globule membrane proteome. *J. Dairy Res.*, **2006**, *73*, 406-416.
- Lippolis, J.D.; Reinhardt, T.A. CENTENNIAL PAPER: Proteomics in animal science. *J. Anim Sci.*, **2008**, *86*, 2430-2441.
- de Souza, C.R.B.; Carvalho, L.J.C.B.; De Almeida, E.R.P.; Gander, E.S. Identification of cassava root protein genes. *Plant Foods Hum. Nutr.*, **2004**, *57*, 353-363.
- Bryant, J.D.; McCord, J.D.; Unlu, L.K.; Erdman, J.W. Isolation and partial characterization of  $\alpha$  and  $\beta$ -Carotene-containing carotene-protein from carrot (*Daucus carota* L.) root chromoplast. *J. Agric. Food Chem.*, **1992**, *40*, 545-549.

- [19] Zhou, J.R.; Gugger, T.E.; Jr J.W.E. Isolation and partial characterization of an 18 kDa carotenoid-protein complex from carrot roots. *J. Agric. Food Chem.*, **1996**, *42*, 2386-2390.
- [20] Milicus, J.C.G.; Juarros, L.J.; De las Rivas, J.; Ibarrondo, J.; Gomez, R. Isolation of a yellow carotenoprotein from carrot. *Phytochemistry*, **1991**, *30*, 1535.
- [21] Osteryoung, K.W.; Vierling, E. Dynamics of small heat shock protein distribution within the chloroplasts of higher plants. *J. Biol. Chem.*, **1994**, *18*, 28676-28682.
- [22] Vishnevetsky, M.; Ovadis, M.; Vainstein, A. Carotenoid sequestration in plants: the role of carotenoid-associated proteins. *Trends Plant Sci.*, **1999**, *4*, 232-235.
- [23] Vishnevetsky, M.; Ovadis, M.; Itzhaki, H.; Vainstein, A. CHRC, encoding a chromoplast-specific carotenoid-associated protein, is an early gibberellic acid-responsive gene. *J. Biol. Chem.*, **1997**, *272*(40), 24747-50.
- [24] Bartley, G.; Scolnik, P. Plant carotenoids: pigments for photoprotection, visual attraction, and human health. *Plant Cell*, **1995**, *7*, 1027-1038.
- [25] Li, L.; Paolillo, D.J.; Parthasarathy, M.V.; DiMuzio, E.M.; Garvin, D.F. A novel gene mutation that confers abnormal patterns of  $\beta$ -carotene accumulation in cauliflower (*Brassica oleracea* var. botrytis). *Plant J.*, **2001**, *26*, 59-67.
- [26] Lopez, A.B.; Eck, J.V.; Conlin, B.J.; Paolillo, D.J.; O'Neill, J.; Li, L. Effect of the cauliflower Or transgene on carotenoid accumulation and chromoplast formation in transgenic potato tubers. *J. Exp. Bot.*, **2008**, *59* (2), 213-223.
- [27] Lu, S.; Eck, J.V.; Zhou, X.; Lopez, A.B.; O'Halloran, D.M.; Cosman, K.M.; Conlin, B.J.; Paolillo, D.J.; Garvin, D.F.; Vrebalov, J.; Kochian, L.V.; Kupper, H.; Earle, H.D.; Cao, H.; Li, L. The cauliflower or gene encodes a dnaj cysteine-rich domain-containing protein that mediates high levels of  $\beta$ -carotene accumulation. *Plant Cell*, **2006**, *18*, 3594-3605.
- [28] Waters, E.R.; Lee, G.J.; Vierling, E. Evolution, structure and function of the proteins in plants. *J. Exp. Bot.*, **1996**, *47*(296), 325-338.
- [29] Mansfield, M.A.; Key, J.L. Synthesis of the low molecular weight heat shock proteins in plants. *Plant Physiol.*, **1987**, *84*(4), 1007-1017.
- [30] Vierling, E. The roles of heat shock proteins in plants. *Ann. Rev. Plant Physiol. Plant Mol. Biol.*, **1991**, *42*, 579-620.
- [31] Sun, W.; Van Montagu, M.; Verbruggen, N. Small heat shock proteins and stress tolerance in plants. *Biochim. Biophys. Acta*, **2002**, *1577*, 1-9.
- [32] Wang, W.; Vinocur, B.; Shoseyov, O.; Altman, A. Role of plant heat-shock proteins and molecular chaperones in abiotic stress response. *Trends Plant Sci.*, **2004**, *9*, 1360-1385.
- [33] Wehmeyer, N.; Hernandez, L.D.; Finkelstein, R.R.; Vierling, E. Synthesis of small heat-shock proteins is part of the developmental program of late seed maturation. *Plant Physiol.*, **1996**, *112*, 747-757.
- [34] Hopf, N.; Plesofsky-Vig, N.; Brambl, R. The heat shock response of pollen and other tissues of maize. *Plant Mol. Biol.*, **1992**, *19*, 623-30.
- [35] Neta-Sharir, I.; Isaacson, T.; Lurie, S.; Weiss, D. Dual Role for Tomato Heat Shock Protein 21: Protecting photosystem II from oxidative stress and promoting color changes during fruit maturation. *Plant Cell*, **2005**, *17*, 1829-1838.
- [36] Levine, R.L.; Mosoni, L.; Berlett, B.S.; Stadtman, E.R. Methionine residues as endogenous antioxidants in proteins. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 15036-15040.
- [37] Laizet, Y.; Pontier, D.; Mache, R.; Kuntz, M. Subfamily organization and phylogenetic origin of genes encoding plastid lipid-associated proteins of the fibrillin type. *J. Genome Sci. Technol.*, **2004**, *3*(1), 19-28.
- [38] Paolillo, D.J.; Garvin, D.E.; Parthasarathy, M.V. The chromoplast of Or mutants of cauliflower (*Brassica oleracea* L. var. botrytis). *Protoplasma*, **2004**, *224*(3-4), 245-253.
- [39] Zhou, X.; Van Eck, J.; Li, L. Use of the cauliflower Or gene encodes a DNaj cysteine-rich domain-containing protein that mediates high levels of  $\beta$ -carotene accumulation. *Plant Cell*, **2006**, *18*(12), 3594-3605.

---

Received: March 30, 2012

Revised: May 24, 2012

Accepted: May 24, 2012

© Carvalho et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.