



Journal of
Plant Sciences

ISSN 1816-4951



Academic
Journals Inc.

www.academicjournals.com

Isolation of Genomic Clones Encoding Granule-bound Starch Synthase (GBSS I) in Cassava (*Manihot esculenta* Crantz)

²J.T. Opatode, ¹O.O. Oyelakin, ²O.A. Akinyemiju and ¹I.L. Ingelbrecht

¹Central Biotechnology Laboratory, International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria; c/o Lambourn Ltd, Carolyn House, 26 Dingwall Road, Croydon CR9 3EE, UK

²Department of Crop Production and Protection, Obafemi Awolowo University, Ile-Ife, Nigeria

Corresponding Author: Jelili T. Opatode, Department of Crop Production and Protection, Obafemi Awolowo University, Ile-Ife, Nigeria

ABSTRACT

Cassava (*Manihot esculenta*) is a major source of starch in tropical and subtropical countries. The genomic organization of the granule-bound starch synthase I (GBSS I) in cassava was examined to increase our knowledge of starch biosynthesis and facilitate the production of modified starches in cassava. Three partial genomic clones were generated and the sequence analysis revealed that they were identical at both nucleotide and amino acid levels. The insert in one of the clones, pOYE303-1 (accession no. HM038439), was 731 nucleotides long and encoded a partial protein of 124 amino acids with predicted molecular weight of 13.65 kDa and calculated 10.01 isoelectric point. HM038439 contained the first conserved domains of GBSS including the KTTGG motif responsible for ADP-glucose binding site. Phylogenetic analysis revealed that cassava GBSS I belongs to dicot subgroup and is closely related to GBSS I of potato, sweet potato and buckthorn. *In silico* identification by nucleotide sequence alignment between genomic and cDNA sequences revealed three putative introns in GBSS I. The longest intron was 121 bp long and was located between nucleotide 595 and 716. The implications of these findings on starch biosynthesis and modification are discussed.

Key words: Cassava, granule-bound starch synthase I, *Manihot esculenta*, amylose, starch

INTRODUCTION

Cassava is a root crop and constitutes an important source of energy in the diet of 600 million people in tropical and subtropical countries (Defloor *et al.*, 1998). The starch content of cassava roots ranges from 65 -91% of its total root dry weight depending on the cultivar (Snchez *et al.*, 2009). Like starches from other plants, cassava starch can be fractionated into two types of polymers namely: amylose and amylopectin. The proportion of amylose to amylopectin in starch determines the physical, chemical and functional properties of cassava starch (Hoover, 2001). Low-amylose (<5%) starches have better physicochemical properties compared with those containing amylose making it more useful for industrial applications. For instance, low-amylose potato starch showed a higher granule melting temperature, less retrogradation and better adhesive properties than that containing higher amylose levels (Visser *et al.*, 1991). Similarly, the use of waxy (low-amylose) wheat has been reported to extend the shelf-life of baked, refrigerated and frozen wheat-based products (Chakraborty *et al.*, 2004). There is considerable interest in generating cassava plants with altered starch characteristics which would offer new industrial

applications for this important root crop. For example, creation of cassava variety that produces low-amylose (waxy) starch would take advantage of existing huge market in global food industry for low-amylose starches which is estimated to be US\$10 million (Blennow, 2003).

Conventional breeding of cassava is challenging due to the highly heterozygous nature of the crop preventing a backcross scheme, exacerbated by poor flowering and limited seed set of many varieties (Ceballos *et al.*, 2004). Many genes are involved in starch biosynthesis and there is limited natural variation with regards to starch metabolism in cassava compare to other crops such as maize. To date, only two starch mutants have been reported in cassava: A natural mutation in a GBSS gene resulting in production of amylose-free starch and a gamma irradiation-induced mutation in isoamylase gene resulting in high-amylose starch (Ceballos *et al.*, 2007, 2008). Because most mutations affecting starch metabolism are recessive, moving useful mutations into the appropriate genetic background via conventional breeding is arduous.

In the field, cassava is typically propagated clonally by stem cuttings. This propagation strategy is ideal for a transgenic approach to crop improvement as gene segregation through outcrossing is limited (Taylor *et al.*, 2004). Transgenic cassava producing amylose-free starch has been reported using antisense inhibition of granule-bound starch synthase I (Raemakers *et al.*, 2005). In addition, transgenic cassava plants with an enhanced quantity of starch have been produced (Ihemere *et al.*, 2006). Further successful manipulation of starch characteristics in cassava requires a more thorough characterization of the various cassava starch biosynthesis genes.

It has been established that granule-bound starch synthase I (GBSS I) is responsible for the biosynthesis of amylose. Cloning and characterization of the complementary DNA (cDNA) encoding granule-bound starch synthase I (GBSS I) in cassava have been reported. The full-length cDNA was isolated from a tuber-specific library. Sequencing of the cloned cDNA showed that it has 74% identity with potato GBSS and 60-72% identity with GBSS from other plant species. The cDNA encodes a 608 amino acid protein of which 78 amino acids form a chloroplast/amyloplast transit peptide of 8.37 kDa. The mature protein has a predicted molecular mass of 58.61 kDa (530 amino acids). Comparison of the GBSS proteins of various plant species and glycogen synthase of bacteria showed extensive identity among the mature form of plant GBSS proteins, in which the monocots and dicots form two separate branches in the evolutionary tree. Nucleotide sequences of cDNA provide no insight on the non-expressed portion of the gene (intron) and recent studies have revealed the importance and regulatory roles of intron in gene functions. From the analysis of the genomic DNA of an allotetraploid cassava, it was shown that GBSS I is encoded by a low-copy-number gene. The GBSS I transcript was synthesized in a number of different organs such as stem, petiole and leaves, but most abundantly in tubers (Salehuzzaman *et al.*, 1993; Munyikwa *et al.*, 1997). To date, the genomic organization of GBSS I is not known in cassava. The identification of the location of intron in GBSS I sequence will have impact on the production of amylose-free starch in cassava and other genetic engineering applications in other crop. This prompted us to conduct the study. The objectives of this study were to generate genomic clone, describe the genomic organization and identify introns *in silico* of GBSS I, the enzyme responsible for biosynthesis of amylose in cassava.

MATERIALS AND METHODS

Plant materials and DNA extraction: DNA was extracted from young leaves (0.5-1.0 g) of field-grown cassava genotype TMS 4 (2) 1425 as described by Dellaporta *et al.* (1983). DNA was

resuspended at a concentration of 500 ng μL^{-1} . The quality of the DNA was verified by running 2 μL of the DNA alongside a molecular weight marker $\lambda\text{Pst I}$ on 0.8% agarose gel electrophoresis in 1 x TAE (Tris Acetate EDTA) buffer at 500 volts for 1 h.

Primer design: A cassava granule-bound starch synthase I gene-specific primer pair was designed from cDNA sequences (accession X14760) earlier deposited in GenBank database from nucleotide +99 to +530. The downloaded sequences were used as template to design primer specific for the genes using Lasergene sequence analysis software (DNASTAR Inc, Madison, USA). The composition of the primer pairs designed and used for the study were as follows: Forward 5'-TGGACCCAAACTATCACTC -3' and Reverse 5'-AGTTTTGCCCCATACCTTC -3'. Synthesis of the primers was done by Integrated DNA Technologies Incorporation (Iowa, USA).

PCR amplification, cloning and DNA sequencing: Amplification was carried out in 50 μL reaction volume, which composed of 1 μL of 10x buffer, 0.5 μL of MgCl_2 (25 mM), 1 μL each of primer F and R (1 μM), 0.5 μL of dNTPs (2.5 mM), 1 μL template DNA (500 ng), 4.8 μL of H_2O and 0.2 U of Taq DNA polymerase (Bioline, USA). The PCR amplification profile consisted of initial denaturation at 94°C for one min and 30 cycles of amplification (94°C for 30 sec, 50°C for 30 sec, 72°C for 45 sec) with a final cycle of 5 min at 72°C. The PCR amplification was carried out in a Peltier thermal cycler (PTC 2000, MJ Research, India). The PCR fragments were purified and cloned into pDRIVE vector (QIAGEN, California, USA). The presence of the insert in the recombinant plasmid was confirmed by restriction digestion. Both strands of the DNA inserts were sequenced and any sequence ambiguities were resolved by re-sequencing. The three independent PCR clones were sequenced from each PCR reaction. The sequences were manually edited and vector sequences removed. DNA sequencing was performed by Iowa State University, USA.

Sequence analyses: The genomic sequence of GBSS I has been submitted to GenBank under the accession numbers HM038439. Putative introns were identified *in silico* by (1) alignments between genomic and cDNA sequences using CLC DNA WORKBENCH software version 6.1 and (2) and confirmed by gene prediction method using GENSCAN web server (Burge and Karlin, 1998). The genomic sequence of GBSS I obtained in this study was aligned with available cDNA sequence (accession X14760) of the gene in the GenBank to identify the locations of the intron.

Phylogenetic analysis: Sequence searches of the non-redundant and unfinished genome databases at NCBI were conducted with the isolated GBSS I ORFs as query sequence. The ClustalX package was used to create an alignment of the sequences that was then submitted to a neighbour-joining analysis to generate a branching pattern. The phylogenetic tree was displayed using the CLC TREEVIEW program. Nineteen GBSS I sequences of higher plants (and their accession numbers) included in the study were: *Frangula anus* (AF500430), *Ipomoea batatas* (AB524726), *Ipomoea trifida* (EU192912), *Nelumbo nicifera* (FJ602702), *Solanum tuberosum* (X58453), *Musa acuminata* (HQ646360), *Gossypium hirsutum* (FJ713477), *Oryza sativa* (AB066094), *Hordeum vulgare* (FN179380), *Saccharum officinarum* (AF446084), *Secale cereale* (FJ491377), *Glycine max* (EU586117), *Perilla frutescens* (AF210699), *Vigna unguiculata* (EF472252), *Phaseolus vulgaris* (AB029546), *Pisum sativum* (AJ345045), *Zea mays* (EU970982), *Manihot esculenta* (HM038439) and *Triticum aestivum* (AF286320).

RESULTS

Amplification, isolation and characterization of cassava GBSS I clones: Out of the three annealing temperatures (45, 50 and 55°C) and two template DNA concentrations (100 and 500 ng μL^{-1}) tested for amplification of GBSS I, the band was distinct at annealing temperature of 50°C using DNA template concentration of 500 ng μL^{-1} . These conditions were used for amplification of GBSS I in 50 μL reaction volume for fragment purification and cloning purposes (Fig. 1). The observed product size was 731 bp against the expected of 323 bp. Three independent clones obtained with the primer pair designed to amplify +99 to +530 portion of *Manihot esculenta* GBSS I (accession X14760) gene were sequenced and their nucleotide sequences were all identical. Therefore, only clone pOYE303-1 was subjected to further analysis. The insert in clone pOYE303-1 (accession no. HM038439) was 731 nucleotides long and encoded a partial protein of 124 amino acids with predicted molecular weight of 13.65 kDa and calculated 10.01 isoelectric point (Fig. 2).

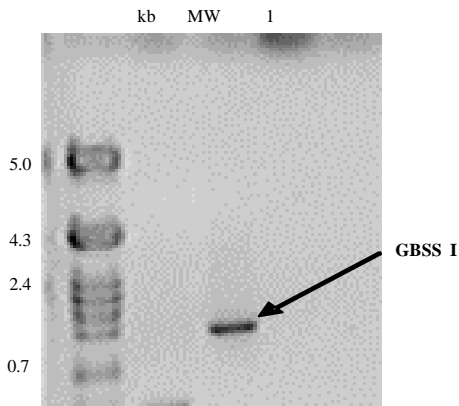


Fig. 1: Amplification of GBSS I by PCR. MW: DNA marker, 1 = 500 ng of DNA at 55°C, Molecular weight markers are at the left in kb

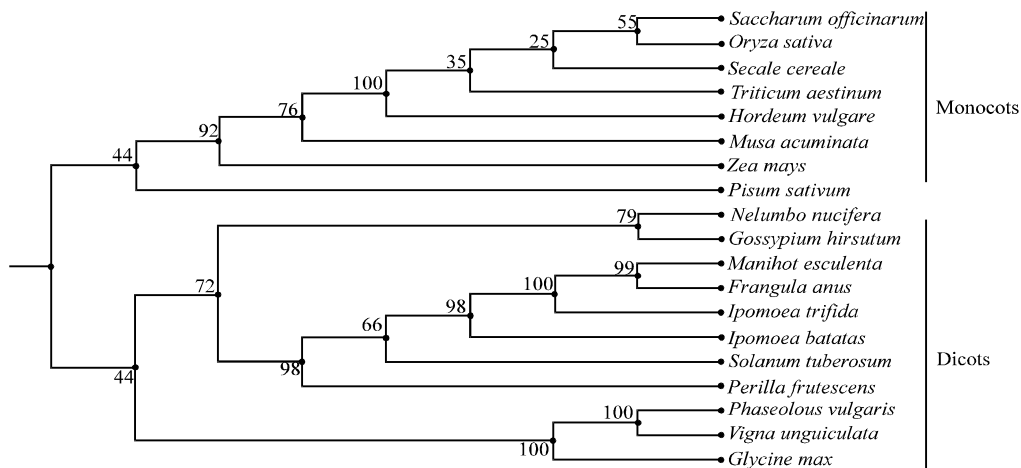


Fig. 2: Phylogenetic relationship of GBSS I sequences of higher plants. The tree was constructed on the basis of available GBSS I sequences in the GenBank and from this study (cassava) using neighbour-joining method from bootstrapped data sets. The number of bootstrap replicates is indicated next to each branch. The scale indicates the average

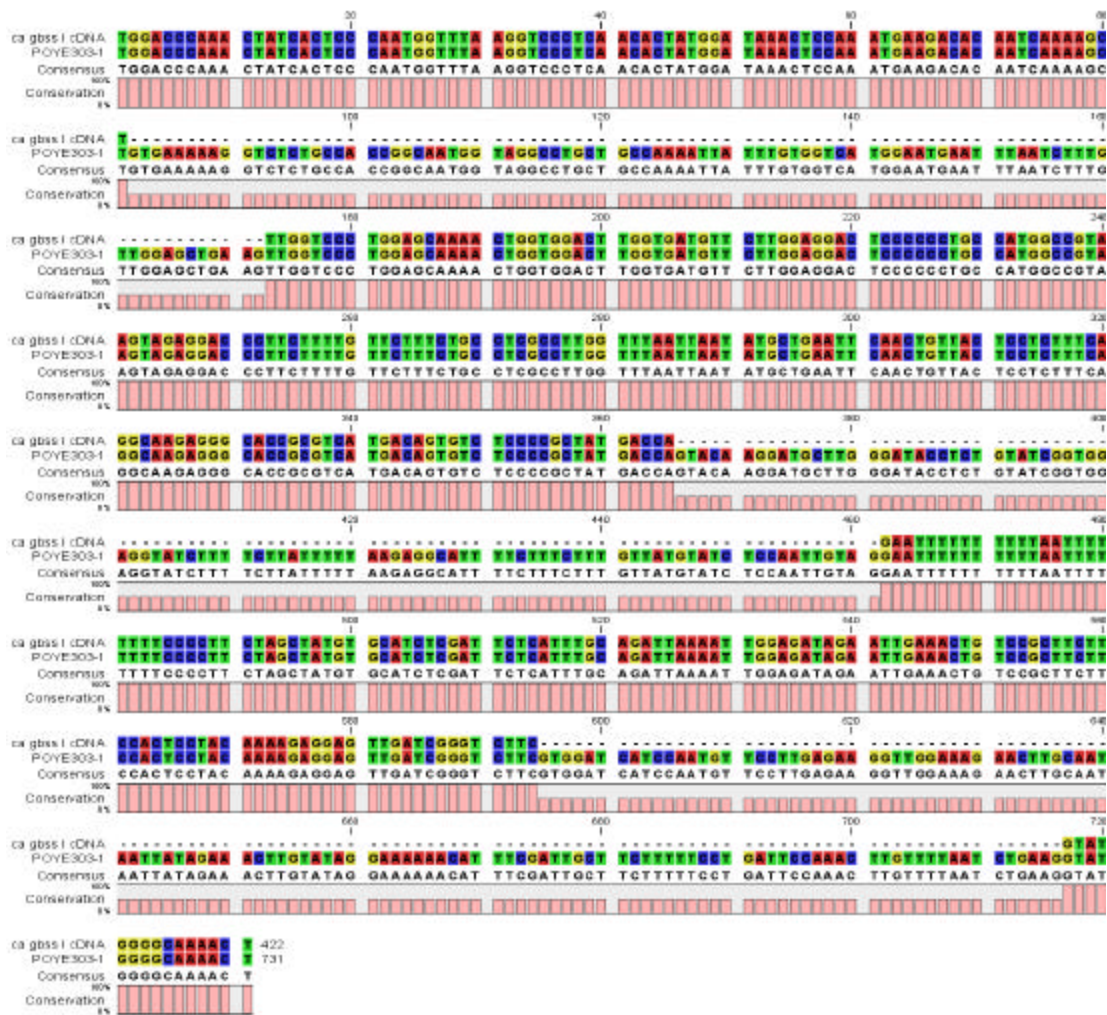


Fig. 3: Alignment of cassava GBSS I genomic and cDNA sequences. Ca GBSS I -cassava GBSS I cDNA (X14760) sequence, POYE303-1- cassava GBSS I genomic (HM038439) sequence of clone POYE303-1

HM038439 was rich in glycine (11.29%), valine (10.48%) and lysine (9.85). The BLAST search showed that GBSS I cDNA of cassava (accession no. X74160) was the most significant hit (100% similarity) with $1e-120$ E-value, as expected. The clone contained the consensus BOX 1 of starch synthase gene. Box I (16 amino acids) was a glycine-rich region containing KTGGL, the substrate-binding site for ADP-glucose (Furukawa *et al.*, 1990). The identity of HM038439 to GBSS I in sweet potato and potato is 66 and 67%, respectively. The predicted secondary structure of HM038439 consist of 9.68% alpha helix, 33.06% extended strand and 57.26% random coil.

Phylogenetic analysis: The GBSS I proteins in higher plants diverged into two groups represented by monocots and dicots (Fig. 2). Cassava GBSS I protein belongs to the dicot group; dicot GBSS I protein further divided into three subgroups. Cassava GBSS I protein is found in the

subgroup of potato and sweet potato. However, it is closely related to buckthorn GBSS I than potato and sweet potato GBSS I.

***In silico* identification of introns:** The locations of three putative introns identified by sequence alignment between the GBSS I isolated genomic sequence (accession No. HM038439) and GBSS I cDNA sequence (accession X14760) deposited in the GenBank are in Fig. 3. The putative intron locations were further confirmed by gene prediction method using GENSCAN web server. The first intron was 91 bp long and located between nucleotide 82 and 172. The first intron was 12.5% of the entire genomic sequence and made up of 30.0%T, 25.3%A, 16.5%C and 27.5%G. The second intron was 96 bp long and about 13.1% of HM038439. The second intron was located between 366 and 461 and made up of 19.8%A, 21.9%G, 43.8%T and 12.5%C. The third intron was 121 bp long and about 16.6% of HM038439. It was located between 595 and 716 nucleotide and 32.2%A, 18.2%G, 35.6%T and 14.9%C rich. When the intron sequences were removed, the nucleotide sequences of HM038439 and GBSS I cDNA were identical. All the introns of GBSS I sequence contain the splice site sequences consistent with the consensus 5'-GT-AG-3' (Hanley and Schuler, 1988).

DISCUSSION

The successful PCR amplification of GBSS I gene indicated that the composition of oligonucleotide primer pairs used was adequate and specific for the gene. Three partial genomic clones of GBSS I was isolated and sequenced in this study. They were 100% identical at both nucleotide and amino acid levels, suggesting that the GBSS I protein in cassava has no isoform, in agreement with previous work reported in cassava (Salehuzzaman *et al.*, 1993) and sweet potato (Wang *et al.*, 1999). GBSSI protein is known to contain three conserved domains known as BOX I, BOX II and BOX III. One of the conserved domains, BOX I, was identified in the insert of each of the three partial *gbss* I genomic clones. The BOX I contains 16 amino acids among which is KTGGL motif that is highly conserved among plants and bacteria starch synthases. The motif is the ADP-glucose substrate-binding site for GBSS I. The KTGGL motif has been reported in sweet potato, maize, sorghum, pea and potato (Wang *et al.*, 1999; Denyer *et al.*, 2001; Edwards *et al.*, 2002).

The phylogenetic study of higher plant GBSS I previously reported (Salehuzzaman *et al.*, 1993) was expanded by including newly identified sequences from higher plants. The divergence of GBSS I into monocot and dicot groups have been documented. However the newly released GBSS I sequence of buckthorn is closely related to cassava GBSS I than potato and sweet potato GBSS I. The reason for these is not understood. Cassava belongs to family Euphorbiaceae and produce tubers while buckthorn is in Rhamnaceae and not producing tubers.

The observed amplification product size (731 bp) was larger than the expected amplification product size (323 bp) of the starch gene studied. Since genomic DNA was used as template DNA for PCR amplification, the difference in predicted and observed PCR product size could be attributable to the presence of introns in genomic DNA. An intron is any nucleotide sequence within a gene that is removed by RNA splicing to generate the final mature RNA product of a gene (Rose, 2008). A total of three putative introns were identified in the genomic sequence of GBSS I cloned. This is the first identification of intron sequence in cassava GBSS I gene, which is expected to have wide applications on gene expression studies and starch modification. The inclusion of introns in plant transformation constructs increased transformation efficiency (Gonzalez *et al.*,

1998; Hankoua *et al.*, 2006). The intron sequences can be used in the design of plant gene constructs to silence or over express starch genes in cassava or other plant genes for enhanced starch functionality (Chen *et al.*, 2003). The introns can also be used for expression studies of other genes in other plants as the significance of introns in gene expression has recently been established (Rose, 2008).

CONCLUSION

The study produced and described, for the first time, partial genomic organization of GBSS I in cassava. The study revealed that cassava genome contains only one isoform of GBSS I. The study also identified three introns in the genomic arrangement of GBSS I. The introns will find applications in genetic modification of cassava for enhanced starch quality and in genetic manipulation of other crops.

ACKNOWLEDGMENT

The Visiting Research Fellowship opportunity given to the first author by International Institute of Tropical Agriculture (IITA) is appreciated. The financial support of Obafemi Awolowo University Research Committee is acknowledged.

REFERENCES

- Blennow, A., 2003. Towards predictable functionalization of starch. *Starch*, 50: 58-64.
- Burge, C.B. and S. Karlin, 1998. Finding the genes in genomic DNA. *Curr. Opin. Struct. Biol.*, 8: 346-354.
- Ceballos, H., C.A. Iglesias, J.C. Perezze and A.G.O. Dixon, 2004. Cassava breeding: Opportunities and challenges. *Plant Mol. Bio.*, 56: 503-516.
- Ceballos, H., T. Sanchez, N. Morante, M. Fregene and D. Dufour *et al.*, 2007. Discovery of an amylose-free starch mutant in cassava (*Manihot esculenta* Crantz). *J. Agric. Food Chem.*, 55: 7469-7476.
- Ceballos, H., T. Sanchez, K. Denyer, A.P. Tofino and E.A. Rosero *et al.*, 2008. Induction and identification of a small-granule, high-amylose mutant in cassava (*Manihot esculenta* Crantz). *J. Agric. Food Chem.*, 56: 7215-7222.
- Chakraborty, M., K. Matkovic, D.G. Greir, E.L. Jarabek, W.A. Berzonsky, M.S. McMullen and D.C. Doehlert, 2004. Physicochemical and functional properties of tetraploid and hexaploid waxy wheat starch. *Starch*, 56: 339-347.
- Chen, S., D. Hofius, U. Sonnewald and F. Bornke, 2003. Temporary and spatial control of gene silencing in transgenic plants by inducible expression of double-stranded RNA. *Plant J.*, 36: 731-740.
- Defloor, I., I. Dehing and J.A. Delcour, 1998. Physico-chemical properties of cassava starch. *Trop. Sci.*, 31: 189-207.
- Dellaporta, S.L., J. Wood and J.B. Hicks, 1983. A plant DNA mini preparation: Version II. *Plant Mol. Biol. Rep.*, 1: 19-21.
- Denyer, K., P. Johnson, S. Zeeman and A.M. Smith, 2001. The control of amylose synthesis. *J. Plant Physiol.*, 158: 479-487.
- Edwards A., J.P. Vincken, C.J.M. Suurs, R.G.F. Visser, S. Zeeman, A. Smith and C. Martin, 2002. Discrete forms of amylose are synthesized by isoforms of GBSS I in pea. *Plant Cell*, 14: 1767-1785.

- Furukawa, K., M. Tagaya, M. Inouyes, J. Preiss and Y. Fukui, 1990. Identification of lysine 15 at the active site in *E. coli* glycogen synthase. *J. Biol. Chem.*, 265: 2085-2090.
- Gonzalez, A.E., C. Schopke, N.J. Taylor, R.N. Beachy and C.M. Fauquet 1998. Regeneration of transgenic cassava plants (*Manihot esculenta* Crantz) through *Agrobacterium*-mediated transformation of embryogenic suspension cultures. *Plant Cell Rep.*, 17: 827-831.
- Hankoua, B.B., N.J. Taylor, S.Y.C. Ng, I. Fawole and J. Puonti-Kaerlas *et al.*, 2006. Production of the first transgenic cassava in Africa via direct shoot organogenesis from friable embryogenic calli and germination of maturing somatic embryos. *Afr. J. Biotech.*, 5: 1700-1712.
- Hanley, B.A. and M.A. Schuler, 1988. Plant intron sequences: Evidence for distinct groups of introns. *Nuc. Acid Res.*, 16: 7159 -7174.
- Hoover, R., 2001. Composition, molecular structure and physicochemical properties of tuber and root starches: A review. *Carbohydr. Polym.*, 45: 253-267.
- Ihemere, U., D. Ariaz-Garzon, S. Lawrence and R. Sayre, 2006. Genetic modification of cassava for enhanced starch production. *Plant Biotech. J.*, 4: 453-465.
- Munyikwa, T.R.I., S. Langeveld, S.N.I.M. Salehuzzaman, E. Jacobsen and R.G.F. Visser, 1997. Cassava starch biosynthesis: New avenues for modifying starch quantity and quality. *Euphytica*, 96: 65-75.
- Raemakers, K., M. Schreuder, L. Suurs, J.P. Vincken, E. Jacobsen and R.G.F. Visser, 2005. Improved cassava starch by antisense inhibition of granule-bound starch synthase I. *Mol. Breeding*, 16: 163-172.
- Rose, A.B., 2008. Intron-mediated regulation of gene expression. *Curr. Top. Microbio. Immun.*, 326: 277-290.
- Salehuzzaman, S.N.I.M., E. Jacobsen and R.G.F. Visser, 1993. Isolation and characterization of cDNA encoding granule-bound starch synthase in cassava (*Manihot esculenta* Crantz) and its antisense expression in potato. *Plant Mol. Biol.*, 23: 947-962.
- Snchez, T., E. Salcedo, H. Ceballos, D. Dufour and G. Mafla *et al.*, 2009. Screening of starch quality traits in cassava (*Manihot esculenta* Crantz). *Starch/Starke*, 61: 310-310.
- Taylor, N., P. Chavarriaga, K. Raemakers, D. Siritunga and P. Zhang, 2004. Development and application of transgenic technologies in cassava. *Plant Mol. Biol.*, 56: 671-688.
- Visser, R.G., I. Somhorst, G.J. Kuipers, N.J. Ruys, W.J. Feenstra and E. Jacobsen, 1991. Inhibition of the expression of the gene for granule-bound starch synthase in potato by antisense constructs. *Mol. Gen. Gent.*, 225: 289-296.
- Wang, S.J., K.W. Yeh and C.Y. Tsai, 1999. Molecular characterization and expression of starch granule-bound starch synthase in the sink and source tissues of sweet potato. *Physiol. Plant*, 106: 253-261.