High-quality pathway of cassava starch biosynthesis in non-photosynthetic tissues inferred by multiple plant templates

Treenut Saithong\textsuperscript{a,b,d}, Orathai Rongsirikul\textsuperscript{a}, Saowalak Kalapanulak\textsuperscript{a,b,d}, Supatcharee Netrphan\textsuperscript{a}, Malinee Suksangpanomrung\textsuperscript{a}, Asawin Meechai\textsuperscript{a,d}, Supapon Cheevadhanarak\textsuperscript{b}

\textsuperscript{a}Bioinformatics and systems biology program, School of Bioresources and Technology, King Monkut’s University of Technology Thonburi, Bangkok, 49 Teintale 25, Bangkuntein-Chaitale Road, Thakam, Bangkuntein, Bangkok, 10150, Thailand.
\textsuperscript{b}School of Bioresources and Technology, King Monkut’s University of Technology Thonburi, Bangkok, 49 Teintale 25, Bangkuntein-Chaitale Road, Thakam, Bangkuntein, Bangkok, 10150, Thailand.
\textsuperscript{c}Department of Chemical Engineering, King Monkut’s University of Technology Thonburi, Bangkok, 126 Pracha-u-tid Road, Toongkhu, Bangnmod, Bangkok, 10140, Thailand.
\textsuperscript{d}Biological modeling lab, Systems Biology and Bioinformatics research group, King Monkut’s University of Technology Thonburi, Bangkok, 49 Teintale 25, Bangkuntein-Chaitale Road, Thakam, Bangkuntein, Bangkok, 10150, Thailand.
\textsuperscript{e}National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency, Thailand Science Park, 113 Pahonyothin Road, Pathumthani 12120, Thailand.

Abstract

Food security and green energy are becoming a popular research, because the world starts to concern about the unbalance of the dramatically increased human population and the current capacity of the natural resources. Cassava, which is a decade ago recognised as a cheap staple crop served as food for developing countries, is at the moment known as a crop of hope in both food and energy aspects. A niche of cassava underlying its significance lies on the capability to produce and accumulate huge amount of starch in the underground roots with respect to the other crops. Despite being a high starch-content crop, the research on cassava mostly aims at the improvement of the starch yield in storage roots. One strategy is to understand the starch biosynthesis pathway, a mechanism of starch production inside plant cells. Taking the advantage of having the recently released cassava genome data, we employed an intensive alignment-based study to infer the high quality pathway of starch biosynthesis in cassava, starting from transported sucrose to starch formation. The cassava genomic data was passed through reciprocal sequence alignment process using five plant templates (i.e. rice, maize, Arabidopsis, potato, castor bean). All identified cassava orthologs were given a confidence score based upon (1) the quality of sequences matching in the annotation process and (2) the consistency of the annotation results among the multiple templates. Furthermore, we performed literature validation for our established pathway by using the experimentally identified sequences from Genebank database. At the end, the informative visualisation of the reconstructed network was provided to integrate all results into a single map. In addition to the reconstructed map, we also discussed the information implied by the results, for instance the high conservation of the starch biosynthesis pathway across the plant species and the advantages of the employed method over the others. The overall results of the study, including a well-defined network, could increase the understanding in the biology of starch biosynthesis process in a cassava plant and is expected to facilitate the next step of the study in unwinding the dynamic regulation of the process.

Keywords: starch biosynthesis, cassava, metabolic pathway reconstruction

Introduction and Objective

Cassava is an important crop for the world population. It is in the fourth rank after rice, maize and sugar cane in terms of the important food sources. Moreover, at the moment, the potential of cassava for being a green energy source is becoming well-established, at least in the bio-diesel production.

At the early stage, the cassava research was aimed to identify the optimal growth condition. Later, it has been found that yield of starch relies very much on the genetics (1), so that breeding technique was introduced and played an important role in finding the better cassava strains. In 2009, the release of cassava genome moved the cassava research a step forward. It accelerated the modern biological study on a process governing the starch biosynthesis.
The starch biosynthesis pathway in cassava was firstly introduced by (2), in which the network was reconstructed from the library of full-length cDNA EST. Then, Rongsirikul et al. (2010) (3) exploited the recently released cassava genome sequences to acquire more complete pathway of starch production in cassava. Although there are already two versions of the pathway, neither of them is informative and claimed as a high quality network. We, therefore, employed an intensive alignment-based study to infer the high quality pathway of starch biosynthesis in cassava, starting from sucrose conversion to starch formation.

Materials and Methods

2.1 Data resources

The nucleotide sequences of the five plant templates, including Arabidopsis, rice, maize, caster bean and potato, were mainly retrieved from two public databases: KEGG (http://www.genome.jp/kegg/) and PMN (http://www.plantcyc.org). For cassava, the data was relied on the sequencing data from the Cassava Genome Project 2009 deployed by Phytozome database (http://www.phytozome.net).

2.2 Sequence alignment

Due to the limitation of the cassava genome annotation at the time of study, Reciprocal Best Hit (RBH) of nucleotide BLAST was employed to identify the putative orthologs of starch biosynthesis genes in cassava. The settings for similarity-based sequence alignment were as follows: E-value < 1E-10, Identity score > 60%, Coverage score > 20%.

Results and Discussion

Through the sequence similarity based study, approximately 70 genes in cassava genome were annotated to function in the biochemical process involved in converting sucrose to storage starch. The pathway comprised of these annotated genes is demonstrated in Figure 1, in which the source and quality of annotation are denoted by the color of circles and boxes, respectively. In addition to the information indicating the confidence of the gene annotation, the quality of the reconstructed pathway is also certified by the literature validation using experimental data in Genbank database (marked as underlined in Figure 1).

Figure 1. The starch biosynthesis pathway of cassava reconstructed through the intensive alignment-based study

In conclusion, we presented the high quality pathway of starch biosynthesis in cassava which could contribute in acquiring more understanding in the biology of such process in a cassava plant.

References