Cassava (Manihot esculenta subsp. esculenta Crantz), is the fourth more important basic product after rice, wheat and maize. In addition to being a basic food in the human diet, cassava constitutes one of the more common raw materials for the starch production. The yield, the dry matter content (indirectly starch content) and planting materials are affected by Cassava Bacterial Blight (CBB). One tool which holds a lot of promise in unraveling the complexities in gene expression is Expresssed Sequence Tags (ESTs). In this study, we have targeted 2 economically important characters, starch content and CBB resistance to generate several types of libraries including subtracted. A high collection of EST was generated. We report here the single pass sequencing of 18202 cDNA clones from the 5'end.

INTRODUCTION

MATERIALS AND METHODS

RESULTS

Characterization of a unigene set

After trimming a total of 13043 sequences was obtained. These were assembled in a unigene set of 5700 sequences, including 1875 Tentative Contigs (TC) and 3825 singletons

Comparing the cassava unigene set with other plant species

Conceptual translation of 16% of the sequences did not show any similarity of available EST of other plant species. They might represent new cassava-specific genes, although some of them might be 5' or 3' untranslatable sequences.

Genes involved in defense

Some ESTs showed similarity to known genes related to defense.

- 39 Protein kinases
- 9 Xa21-like protein
- 13 similarity to R-gene
- 10 Cytochrome P450
- 3 peroxidases
- Transcription factor WRKY
- EDS1
- Chitinase
- 56% Non similarity

Comparing the specific or common sequences among the different cassava cultivars we observed that most of them were specific to each cultivar-Xam interaction (either susceptible or resistant reactions)

Functional categories

Functional categories were defined using the Gene Ontology (GO) classification scheme (http://www.geneontology.org). A total of 63% of unigenes did not have a significant match. Thus, 37% of the unigene set was assigned a putative function using this method (Fig below).

Characterization of EST associated to starch biosynthesis

Putative genes involved in the starch biosynthesis were identified in the cassava EST collection

- 3 Isoformes alpha-1,4 glucan phosphorylase
- 2 Tubulin-specific and sucrose-responsive element binding factor
- 2 Isoforms Starch phosphorylase
- Glycogen [starch] synthase, chloroplast precursor (GBSIII)
- Granule bound starch synthase II precursor
- Soluble glycogen [starch] synthase, chloroplast precursor (SS III)
- Protein induced upon tuberization

Comparing in silico gene expression

We performed a differential analysis using the method reported by Stekel el al. (2000) to study differences in gene expression between cultivars CM523-7, and MPer183. Results are presented below:

CONCLUSIONS AND PERSPECTIVES

- The cassava EST data presented here is the first effort in the large scale sequencing of the cassava expressed genome and also in cataloguing cassava genes.
- A unigene set of 5700 sequences was identified and a putative function assigned to 37% of unigenes.
- A number of ESTs were found to be present only in the Xam challenged libraries.
- The EST resource will increase the density of gene markers on the cassava genetic map.
- Information obtained here will be used to develop microarray technology for further cassava gene expression studies.
- Based on this, developing new cassava varieties having high dry matter content and durable resistance to CBB will be the next challenge.