Seventh framework programme
Food, Agriculture and Fisheries, and Biotechnology

Specific International Co-operation Actions
Small or medium scale focused research project

Sweet Fuel

Sweet Sorghum
an alternative energy Crop

Grant Agreement n° 227422

WP4
Deliverable 4.6:

Drought regulated genes identified *in silico*

Composition of the consortium

CIRAD
ICRISAT
EMBRAPA
KWS
IFEU
UniBO
UCSC
ARC-GCI
UANL
WIP

In order to evaluate different Sorghum genotypes for their tolerance against drought stress it is important to study the genes involved in the complex mechanism of plant response to drought. Some of these genes and some QTLs that control the process are described in scientific publications but a lot of work is still necessary to increase our understanding on drought response, considering also different environmental conditions.

The main purpose of the WP4 is to identify useful traits in sweet sorghums for bio-ethanol production, characterizing their physiological and genetic basis. Carbohydrate accumulation is directly related with biotic and abiotic stresses in particular with drought. Within this WP we (UCSC) will investigate the molecular genetics of drought tolerance with the objective to develop useful molecular markers for MAS (task 4.5), on the basis of candidate genes or gene sequences. For this purpose differential gene expression screening using microarrays and association mapping using SNPs markers will be exploited.

Task 4.5.1. plans the identification of drought-related genes by computational analysis of sequence information with the forthcoming Arabidopsis Information Resource (http://www.arabidopsis.org) and with the Gramene database, particularly with the maize and sorghum EST indexes (http://www.gramene.org). This is an important step to proceed with the task 4.5.2., analysis of differentially gene expression of two different sorghum genotypes (Keller and IS-19453) under drought and irrigated conditions, by using microarray technology. Drought-related sequences contained in public databases provide useful information on genes involved in the drought response mechanisms, furthermore they can be used to realize a more specific chip, including genes discovered in species different from sorghum and putatively important for understanding sorghum drought response.

We identified several drought-related genes in the public databases, in particular in the Plant Gene Index at http://compbio.dfci.harvard.edu/tgi/plant.html. This database resulted, at the moment of the database creation, the most complete and updated. We selected the tentative consensus (TC) sequences listed in the section “Functional annotation and analysis” – “Gene Ontology”, then “Response to stimulus” (as biological process) and “response to stress”. Each TC annotation is identified by a TC number and by a description of its molecular function, obtained as the best hit after BLAST operations.

The resulting drought-database is composed by all the available TC sequences of six crop species (Sorghum, Arabidopsis, Rice, Maize, Barley, Wheat) representing the genes involved in response to drought and other abiotic stresses. Since several genes are induced in more than one abiotic stress we decided to include other type of stresses in addition to those more closely involved with water deficiency. Beside sorghum we chose the above mentioned five species for which several genomic information are available, furthermore for maize, rice and Arabidopsis the complete genome has been sequenced and maize and rice have high genomic similarity with sorghum. The aim of these approach is to test some drought related genes present in other crop species and not annotated in sorghum and verify their putative expression in stressed and unstressed sorghum samples.

In detail the database contains 9938 TC sequences of which 509 of Sorghum, 482 of Arabidopsis, 375 of Rice, 1241 of Barley, 3429 of Maize and 3839 of Wheat.

All genes identified by TC sequences are divided in the following stress groups:

response to osmotic stress
hypotonic response

hypotonic salinity response
hyperosmotic response
hyperosmotic salinity response
response to salt stress
regulation of translation initiation in response to osmotic stress
response to water deprivation
response to desiccation
cellular response to water deprivation
response to heat
response to heat acclimation
heat shock-mediated polytene chromosome puffing
response to cold
response to cold acclimation
response to vernalization
response to freezing

Within each stress group, the genes have been divided on the basis of their biological function in the following functional classes:

- Cell rescue, defence and virulence (heat shock protein, chaperonins, enzymes and proteins involved in oxidation-reduction activity, etc.)
- Cellular communication/signal transduction mechanism (kinases, phosphatases)
- Nucleic acid binding (RNA binding protein, transcriptional factors)
- Sugar metabolism (Glyceraldehyde-3-phosphate dehydrogenase, Galactinol synthase 3, Enolase 1, etc.)
- Nucleotide binding (GTP and ATP binding protein)
- Energy (Ribulose bisphosphate carboxylase/oxygenase activase A, Phosphosulfolactate synthase-related proteins, etc.)
- Interaction with the environment (Low temperature and salt responsive proteinS, Temperature-induced lipocalin-2, Neoxanthin cleavage enzyme, etc.)
- Systemic interaction with the environment (Allene oxide cyclase and other enzymes involved in hormones biosynthesis, hormones receptors and responsive elements)
- Aminoacids metabolism (Pyrroline-5-carboxylate reductase, S-adenosylmethionine synthetase 2, Methionine aminopeptidase, etc.)
- Cellular transport (Aquaporin and other transport membrane protein)
- Protein fate (Cysteine proteinase, ubiquitins, Peptidyl-prolyl cis-trans isomerases, etc.)
- Fatty acid metabolism (Fatty acid desaturase DES1, Lipoxygenases)
- Cell cycle and DNA processing (histones)
- Protein synthesis (Ribosomal proteins, Elongation factor 1-gamma, etc.)
- Cell fate
- Storage protein
- Subcellular localization
- Unknown

Database can be consulted selecting for species, type of stress, biological process and molecular function. Unknown category represents a group of expressed sequences (EST) during stresses but not yet characterized and annotated.

We picked out some genes to test via PCR their presence and profile in the Keller and IS-19453 genotypes. We chose some dehydrins, Late Embryogenesis Abundant protein, P5CS. From literature these genes are known to be involved in response to drought. The objective is to check for the presence of polymorphic genes between the two genotypes and use them as candidate genes for the future mapping of drought-involved QTLs, exploiting F_2 and F_3 populations from the cross Keller X Is-19453. Moreover through qRT-PCR is possible to quantify expression of each gene at different stress levels and evaluate the type and the intensity of the response in each genotype. This part of work is currently in progress.