Comparaison des génomes de trois Arecaceae d'importance, co-évolution du sorgho et du striga et étude de l'association chez le sorgho

Comparison of three Arecacea genomes: oil palm (Elaeis guineensis), date palm (Phoenix dactylifera) and coconut tree (Cocos nucifera); co-evolution of sorghum and striga, association study in sorhum and pearl millet

ABSTRACT

1 - Comparison of three Arecacea genomes: oil palm (Elaeis guineensis), date palm (Phoenix dactylifera) and coconut tree (Cocos nucifera)

The objective of the project is to compare five genomic regions in three Arecacea species: oil palm (Elaeis guineensis), date palm (Phoenix dactylifera) and coconut tree (Cocos nucifera) and with rice.

For the project, we used two BAC libraries own by Cirad: one for oil palm and one for coconut tree. The date-palm fosmid library was own by J. Bennetzen laboratory. The sequencing of the clones was obtained through a Genoscope project.

To select the BAC clones, we hybridized the BAC high-density filters with probes corresponding to a few genes:
- The ADH gene (alcool deshydrogenase); this loci has already been largely studied in comparative studies. One probe has been derived from the sequence of the ADH gene in sugarcane and specific primers for ADH 1 and ADH 2 were defined from the alignment of several ADH genes.
- The genes EgDEF1 (MADS-box sub-family deficiens) et EgGLO2 (MADS-box sub-family Globosa) were selected since they are involved in the somaclonal variation "mantled", which induce a conversion of male floral component in female one with an important impact on production of palm oil. Specific probes were derived from the EgDEF1 and EgGLO2 cDNAs.
- The other regions were selected from the partial assembly of the date-palm sequence (http://qatar-weill.cornell.edu/research/datepalmGenome/download.htm). The scaffolds were annotated automatically with Fgenesh and Eugène. This annotation allowed identifying gene-rich scaffolds. Two gene rich scaffolds were chosen corresponding to the loci PEPC and FBP (phosphoenolpyruvate carboxylase and Fructose-bi-phosphate). Probes were developed from the genes on these scaffolds and PCR primers were designed for BAC validation.

ADH probes allowed identifying 18 oil palm BACs and 13 coconut tree BACs. BAC fingerprint and PCR analysis allowed distinguishing the BAC corresponding to ADH1 and ADH2 loci. Two oil palm BACs Eg033E10 (ADH1) and Eg060I13 (ADH2) and one coconut BAC comprising the ADH1 and ADH2 genes were selected for sequencing.

Probes corresponding to EgDEF1 and EgGLO2 allowed identifying 4 and 2 BAC respectively but did not allow identifying coconut tree BACs. Since these two genes correspond to multigenic family, we had to design very specific probes based on date palm. This may explained the absence of hybridization on coconut trees. The oil palm BAC Eg133H20 (EgDEF1) and Eg033E06 (EgGLO2) were selected for sequencing.

Probe corresponding to the loci PEPC identified 4 oil palm and 5 coconuts BAC clones. PCR analysis and fingerprint allowed selecting BAC Eg132P05 (oil palm) and BAC Cn227M10 (coconut) for sequencing.

Probe corresponding to the loci FBP identified 5 oil palm and 3 coconuts BAC clones. PCR analysis and

fingerprint allowed selecting BAC Eg172L05 (oil palm) and BAC Cn142B10 (coconut) for sequencing.

Selection of date-palm fosmids was performed in J. Bennetzen laboratory. Eight fosmids were selected Pd007D03_D6 and Pd009A12 for the loci ADH; Pd005B81_4D2 and Pd005B81_5H2 for the loci EgGLO2; Pd009B12_D2 and Pd009H8_H8 for the loci FBP; Pd003C11_1C9 and Pd006A61_G9 for the loci PEPC.

We encountered an important problem of delay in the sequencing phase that took much more time than anticipated and lead to important delay in the progress of the project.

In addition, some of the fosmids sent by J. Bennetzen lab for sequencing to Genoscope appeared to be redundant: 4 fosmids (Pd007D03_D6, Pd005B81_4D2, Pd005B81_5H2 and Pd003C11_1C9) had the same sequence.

Finally, in total 6 oil-palm BAC, 3 coconut-tree BAC and 8 date-palm fosmids were sequenced

2. Co-evolution of sorghum and striga
J. Bennetzen was interested to study sorghum/pearl millet/striga co evolution. Striga is a parasite plant very prejudicial in Africa. We provide him contact for his mission in Mali to collect samples.

3. Association analyses in sorghum and pearl millet

K. Devos was interested to analyse gene involved in height variation through association study in sorghum and/or pearl millet.

Association analyses is done most effectively when a panel of varieties is shared between different research groups. In discussion with IRD, she adopted their association panel for pearl millet, and expanded it with varieties from the United States. her trait of interest is plant height. Three genes had been identified in her lab as underlying height variation in pearl millet. In first instance, we needed to isolate the pearl millet orthologs of those genes. Using grass comparative information, primers were designed against conserved regions in the candidate genes. Pearl millet amplification products were cloned and sequenced. Unfortunately, most of the cloned products appeared to be retroelements and were thus the result from non-specific amplification.

4. Characterization of translocation break points in wheat

K. Devos has been working on genomics in Triticeae for many years and more recently focused on the analysis of translocations. She wanted to use our expertise in molecular cytogenetics to be trained in new methods in molecular cytogenetic to continue this work in Montpellier.

The objective of the study was to determine the physical size of the region that spanned the 4L/5L translocation breakpoint in wheat. Using comparative information from rice, four markers (two for the group four and two for group five) had been identified that spanned the breakpoint. A partial contig for the breakpoint region had indicated that the region likely spanned a few hundred kilobases and consisted mainly of repetitive DNA. To determine the physical size of the region, in situ hybridization needed to be carried out with the markers flanking the breakpoint on both chromosomes 4 and 5. Visualization of in situ hybridization signals requires probe lengths of minimum 2 to 3 kb. The available probes were less than 500 bp long. To generate additional probes for the four loci, primers were constructed using the rice genes as template. Alignment of the rice genes with wheat ESTs allowed the design of primers that would amplify the orthologous sequences in wheat. The amplified fragments were cloned and sequenced to confirm their identity. This brought the total available probe length to 2 - 3.8 kb for each of the four loci. In order to increase the resolution with which two probes could be differentiated, in situ hybridization was carried out on flow sorted chromosomes, provided by J. Dolezel (IEB, Czech Republic) and we tested the protocol of J. Dolezel to stretch the chromosome before hybridization. A series of labeling and hybridization procedures were tested to enhance the sensitivity with which hybridization signals could be detected. But the conclusion was that the probe length needs to be further increased to be able to obtain successful hybridization and detection at both two loci that flank a breakpoint.

Year : 2008
Project number : 0800-001
Type of funding : PC
Project type : SP
Research units in the network : DIADE LGDP
Start date : 2008-05-01
End date : 2010-06-30
Flagship project : no

Project leader : Angélique d'hont
Project leader's institution : CIRAD
Project leader's RU : AGAP

Budget allocated : 93706.6 €
Total budget allocated (including co-financing) : 93706.6 €
Funding : RTRA

PERSPECTIVES

Comparison of three Arecacea genomes : we are currently annotating all the sequences then the comparison will be performed between the tree Arecacea species and with rice.
A publication on these results will be written jointly by the two research teams