POSTER - DNA AND GENOMICS

LTR-RETROTRANSPOSON CHARACTERIZATION IN RHYNCHOSPORA HOLOCENTRIC GENOMES

Letícia Maria Parteka (leticia.maria.parteka@uel.br)

Rafael De Assis (rafaeldeassiis@gmail.com)

Guiherme De Carvalho Borges (guilherme.borges@uel.br)

André Marques (amarques@mpipz.mpg.de)

Andre Luis Laforga Vanzela (andrevanzela@uel.br)

Long terminal repeat-retrotransposons (LTR-RTs) comprise a large portion of plant genomes, highly repeated and distributed along chromosomes. There are autonomous elements, which have all the genes responsible for transposition intact and functional, as well as non-autonomous presenting degraded sequences. In plants, there are two large superfamilies of LTR-RTs: Copia and Gypsy, based on the internal organisation of protein domains in the polygenic chain. Each superfamily is further subclassified into lineages based on similarities of the coding region. Non-autonomous LTR-RTs can be classified according to their sizes and with the structure remaining in them, i.e. Terminal-Repeat Retrotransposons in Miniature (TRIM), Large Retrotransposon Derivatives (LARD) and Terminal Repeat Retrotransposons with GAG Domain (TR_GAG). Rhynchospora species (sedges) have holocentric chromosomes, which exhibit kinetochore distributed along chromatids. This research was conduced to analyze the set of autonomous and non-autonomous LTR-RTs in the fully sequenced genomes of seven Rhynchospora species (R. pubera, R.

breviuscula, R. tenuis, R. colorata, R. tenerrima, R. cephalotes, and R. alba). The objective was to understand whether the diversity of LTR-RTs in holocentic genomes is similar to monocentrics, which represent the majority of organisms. The strategy consisted of filtering the LTR-RT fraction, by analysing the frequency with which they occur in each genome, in search of centromeric signatures. LTR-RTs were filtered from high coverage genomic sequencing (PacBio) using the EDTA tool and sorted with the Inpactor program. For the autonomous elements, reverse transcriptase sequences had been isolated and aligned, along with RT database of GypsyDB. The non-autonomous were just aligned without any references. The phylogenetic trees of the elements were created in the FastTree tool and then analyzed in the iTOL platform. The autonomous LTR-RTs of the Copia superfamily was predominant in all genomes, ranging from 40% (R. pubera) to 80% (R. alba). The Gypsy superfamily varied from less than 10% (R. alba) to 35% (R. pubera). The main Copia lineages were Bianca, Oryco, Retrofit, Sire and Tork, 95% of which were found in all genomes. Among those, Retrofit and Tork were the most numerous. The main Gypsy elements were Athila/TAT, CRM, DEL, Reina and Galadriel, but the latter was found only in R. alba. Athila and Reina members were the most abundant of Gypsy. The non-autonomous LTR-RTs elements were grouped according to each species, but their respective lineages could not be identified. Sequences from the different clusters were selected for annotation, and in some of them small satellite DNA repeats were observed. A satellite named Tyba related to centromeric Gypsy autonomous elements (CRM) were found in R. pubera. These data will be used for design probes in order to produce physical mapping of LTR-RTs in these holocentric chromosomes.