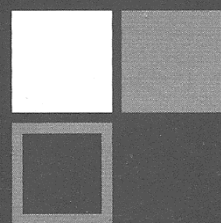


Plant GEMS

5.

**Plant
Genomics
European
Meetings**

Venice, 2006



P 2.1

NEW UTILITY TOOLS FOR TANDEM REPEAT DATA MINING

Bilgen M. (bilgen@akdeniz.edu.tr), İnce A.G., Onus A.N., Karaca M.

Department of Field Crops, Faculty of Agriculture, Akdeniz University, Antalya II 07059, Turkey

With the advent of new biotechnological tools number of DNA sequences being deposited in GenBank are rapidly proliferating. Such as 26,858,818 ESTs deposited in GenBank in May 13, 2005, the database increased to 36,205,655 in May 5, 2006 for a total of 1,087 organism. Here we report on the development of two utility tools that make easier and faster to extract, collect and sort the DNA sequence data to be used in Tandem Repeats Analyzer 1.5 (TRA, Bilgen et al., 2004) and Exact Tandem Repeats Analyzer (e-TRA, Karaca et al., 2005) or some other data mining programs including Tandem Repeats Finder, (TRF, Benson, 1999), the Repeats Computer (REPuter, Kurtz et al., 2001). Simple Sequence Repeat Identification Tool (SSRIT, Kantety et al., 2002), Simple Sequence Repeat Finder, (SSRF, Sreenu et al., 2003), Search for Tandem Repeats In Genomes, (STRING, Parisi et al., 2003), and Microsatellite Search, (MISA, Thiel et al., 2003). We also demonstrated the applications of the two utilities on 243,334 ESTs sequences of *Gossypium* and *Capsicum* that spliced from downloaded GenBank data using 36,205,655 ESTs. ESTs from the two plant species were mined for organism, organ, tissue and developmental stages. Tandem Repeat Analyzer program was then implemented for development of tissue, organ specific EST primer pairs.