Proceedings of the XLIX Italian Society of Agricultural Genetics Annual Congress
Potenza, Italy – 12/15 September, 2005
ISBN **88-900622-6-6**

Oral Communication Abstract - 5.03

COMPLETE SEQUENCE AND ANNOTATION OF THE CHLOROPLAST GENOME OF GRAPE (*VITIS VINIFERA* L.) CV PINOT NOIR

M. MOROLDO*,**, M. BARBARO*, S. SCALABRIN*, R. VELASCO**, M. MORGANTE*

*) Dipartimento di Scienze Agrarie e Ambientali, Università degli Studi di Udine, Via delle Scienze 208, 33100 Udine
**) Istituto Agrario San Michele all'Adige, Via Mach 1, 38010 San Michele (TN)

chloroplast, Vitis vinifera, annotation, microsatellites

Chloroplast genomes of higher plants exhibit a remarkable level of similarity in sequence, organisation and gene order. They have small sizes (about 150 kbp) and contain about 120 genes; several microsatellites are found as well. Moreover, chloroplasts are mainly uniparentally unherited and their molecular lineages are not perturbed by recombination. For these reasons, they are a unique tool to study population genetics, to unveil the origin of polyploids and hybrids and to infer phylogenetic relationships.

In this work we present the first sequence of the chloroplast genome of grape (*Vitis vinifera* L.), together with its annotation for genes and microsatellites.

As a first step, a BAC clone containing the whole chloroplast genome was identified by analyzing through Blast search 27000 BAC ends from a BAC library of grape cv. Pinot Noir. The DNA extracted from this clone was then sheared by nebulization. Fragments in the size range of 1500-2500 bp were produced and then ligated into pCR-Script. The ligation mix was then used to transform *E. coli* strain DH10β and about 3600 clones were picked up; about 6800 paired-ended sequences were obtained in this way. Further 500 sequences were obtained by Blast search on two genomic libraries of grape cv. Pinot Noir, in which cpDNA represents a contamination of nuclear DNA.

All the about 7300 paired-ended sequences were then assembled using Arachne2 (Whitehead Institute, MA) and 38 contigs were obtained. These contigs were furtherly processed using Cap3 (Huang and Madan, Genome Research 1999, 9(9): 868-877), reducing their number to 28. After this assembly the contigs were annotated using the web-based software DOGMA (http://bugmaster.jgi-psf.org/dogma/), which identifies genes by performing a Blast search against a custom database of genes from 16 complete chloroplast genomes of green plants.

Taking advantage of the high level of colinearity between chloroplast genomes and using *Nicotiana tabacum* as a reference, this first annotation step made it possible to know the relative order of the contigs between them. Moreover, this led to determine the position of the gaps, which were then closed by PCR amplification and sequencing.

Once the sequence was complete, it was annotated again for genes using DOGMA. 112 genes were identified, a number which is very close to the one usually found in green plants, and the gene order appeared to be the same known for plants like *Nicotiana tabacum*. A modified version of the software Sputnik (Morgante *et al.*, Nature Genetics 2002, 30(2): 194-200) was then run and 57 SSRs were found. Subsequently, a phylogenetic study was undertaken to better establish the position of *V. vinifera* among green plants, especially among the dycotiledonous.

To elucidate the patterns of evolution of microsatellites in cpDNA, the SSRs found in the plastid of grape were aligned with the syntenic loci of other species, allowing a direct comparison between the corresponding sequences to make inferences on mutations and homoplasy.