

BAC library of *Lupinus angustifolius* nuclear genome

The BAC library of nuclear genome of narrow-leaved lupin cv. Sonet (Kasprzak et al. 2006) contains 55 296 clones of mean insert size of 100 ± 5 kbp, which theoretically is equivalent to the six-fold coverage of the *L. angustifolius* genome. The library is characterized by a very low level of contamination with DNA fragments coming from the chloroplasts and mitochondria, which expressed in percentages accounts for 0.018% for mtDNA and 0.045% for cpDNA. The probability of finding a genome sequence in the library was estimated using the Clarke-Carbon formula at 99.7%. The BAC clones are freeze-stored (-80°C) on LB medium with an addition of 20% glycerol and 12.5 $\mu\text{g/ml}$ chloramphenicol, placed in 144 plates of 384 wells each. For the purpose of searching the BAC library using molecular probes, DNA isolated from clones included in the library was placed in nylon filters, actually being DNA macrotemplates of high density. For searching the BAC genomic library of narrow-leaved lupin molecular probes are used, designed on the basis of the following DNA sequences:

- Fragments of ENOD40 and SymRK genes, responsible for the formation of root nodules,
- fragments of the chalcone synthase gene from *L. albus*, one of the genes of the phenylpropanoid metabolic pathway,
- AntjM1 and AntjM2 markers, conjugated with resistance of *L. angustifolius* to anthracnose, caused by *Colletotrichum lupini*,
- Ph258M2 and PhtjM2 markers, connected with resistance of *L. angustifolius* stem brown rot, caused by *Diaporthe toxica*,
- RustM1 marker, conjugated with resistance of *L. angustifolius* to rot, caused by a pathogen *Uromyces lupinicolus*.

DNA isolated from BAC clones was used as a template in PCR with primers specific for sequences of applied molecular probes. Having templates with a high DNA concentration, ends of inserts of isolated BAC clones were sequenced (BES). A total of 485 BESs were obtained, with mean length of approx. 620 bp. These sequences were functionally annotated, targeting the identification of sequences of known genes and proteins, based on the similarity of the sequence of occurrence for individual nucleotides or amino acids. The COBALT software and implemented data bases of sequences repeated in the genome and available BLAST algorithms, mainly BLASTX

and BLASTN were used for this purpose. Genes and proteins were identified, belonging to different classes and sometimes coming from considerably evolutionarily distant species of plants and animals. Identified genes are characterized by a considerable functional differentiation – they are both structural proteins and catalytic, transport, or regulatory proteins. Based on BESs exhibiting affinity to identified genes primers were designed for PCR, which were applied to generate STS genetic markers and to localize BAC clones on the genetic map of narrow-leafed lupin. BAC clones were also subjected to physical mapping, using restriction fingerprinting with the use of enzymes recognizing six-nucleotide sequences. On this basis several contigs of BAC clones were generated. Verification of clone position in these contigs was performed in several stages, both using genetic mapping and at the application of fluorescent in-situ hybridization (BAC-FISH).