NEXT EUKARYOTIC GENOME REVEALED WITH COOPERATION OF THE POLISH LABORATORIES.

POTATO GENOME SEQUENCED

Summary

In mid-July 2011 Nature published the paper „Genome sequence and analysis of the tuber crop potato” by the Potato Genome Sequencing Consortium – PGSC.

This international consortium consisted of 32 teams from 14 countries that are significant potato producers are active in potato breeding programs. The Polish team, representing the PAS Institute of Biochemistry and Biophysics, co-authored the paper, which was signed by 94 consortium members from 25 leading institutions. The work, begun in 2007, was funded by the participating countries’ governments – including Poland’s Ministry of Science and Higher Education – within the 47/PGS/2006/01.

After the yeast and paramecium genomes this is the third large genome sequencing in which IBB teams have participated.

Consortium sequenced two strains of potato DM1-3 516 R44 and RH 89-039-16. The DM strain was sequenced by the consortium using the “Shotgun method” on genome sequencing platforms, one of which was created in Warsaw (Roche GS FLX Titanium 454 sequencer).The second strain sequenced was a laboratory *S. tuberosum* heterodiploid RH 89-039-16.

Ab initio predictions of genes and their functions were verified by RNA transcriptome analysis done for various tissues, development stages and under stress. This allowed for the identification of 39031 gene-coding proteins. 25,3% of these genes produce RNA that undergoes splicing, coding for an average of 2–3 different proteins. It therefore seems that the potato genome codes for approximately 100 000 different proteins.

Comparisons of the DM and RH sequences show that the potato exhibits high heterozygosity. Single-nucleotiste polymorphisms (SNP) are encountered on average every 40 nucleotides, while insertions or deletions (so-called indel) of an average length of 12.8 nucleotides are encountered on average every 394 base pairs.

This data clearly shows that gene damage in the potato genome is a frequent occurrence, which accounts for the easy degeneration of industrial strains.