

## **Program SEED: Zpracování a analýza houbových sekvencí snadno a rychle**

TOMÁŠ VĚTROVSKÝ\*, PETR BALDRIAN

Mikrobiologický ústav AV ČR, v.v.i., Vídeňská 1083, 142 20 Praha 4

\*vetrovsky@biomed.cas.cz

Je mnoho zavedených programů na zpracování amplikonových dat z NGS, z nichž naprostá většina je založena na práci v příkazovém řádku, a proto je hůře stravitelná pro biology bez základních znalostí bioinformatiky. Program SEED (<http://www.biomed.cas.cz/mbu/lbwrf/seed/>) je volně dostupnou “okenní” aplikací pro zpracování amplikonových dat, vyvinutou pro 64-bitovou platformu Windows (1). Program byl vytvořen s cílem poskytnout intuitivní prostředí pro rychlou bioinformatickou analýzu amplikonových sekvencí odvozených od různých taxonomických markerů (houbových ITS oblastí, bakteriálních 16S rDNA, aj.) na stolním počítači. Může však být použit jako multifunkční nástroj pro práci s nukleovými sekvencemi obecně. SEED verze 2.0 je plně optimalizován pro zpracování amplikonových dat ze sekvenátoru Illumina a umožňuje pohodlnou práci s daty o velikosti okolo 8 milionů sekvencí (~4 GB) na běžném počítači s 8 GB RAM. Nová verze 2.0 má široké spektrum funkcí zahrnujících editaci sekvencí a jejich názvů, spojování “pair-end” sekvencí, filtrování podle kvality, třídění a seskupování sekvencí, “alignování” a klastrování sekvencí, tvorbu OTU tabulek, vyhledávání pomocí algoritmu BLAST, taxonomickou identifikaci, konstrukci a editaci fylogenetických stromů a mnoho dalších.

- (1) Větrovský, T. and P. Baldrian (2013). "Analysis of soil fungal communities by amplicon pyrosequencing: current approaches to data analysis and the introduction of the pipeline SEED." *Biology and Fertility of Soils* 49: 1027–1037.