Vizualizace a identifikace organismů v metagenomických studiích

Karel Sedlář
Jiří Těthal
Tomáš Reigl

PA055 Vizualizace komplexních dat
projekt

Popis dat

- amplikonové sekvenování 16S rRNA vzorků vod

- pomocí DNA barcodingu kvantitativně určeno složení mikroflóry

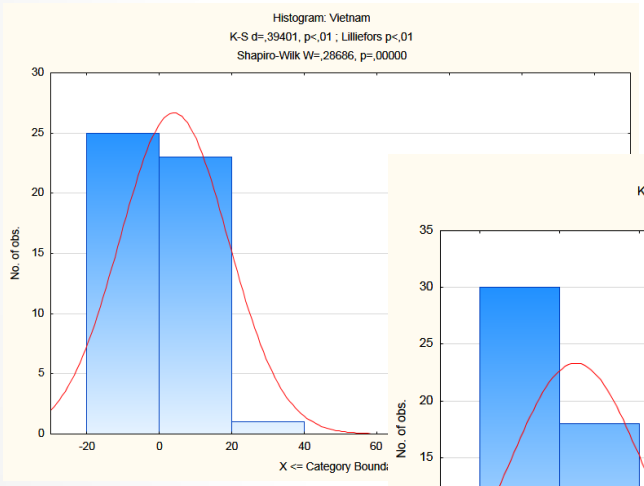
- qPCR genů antibiotické rezistence z rybích vzorků

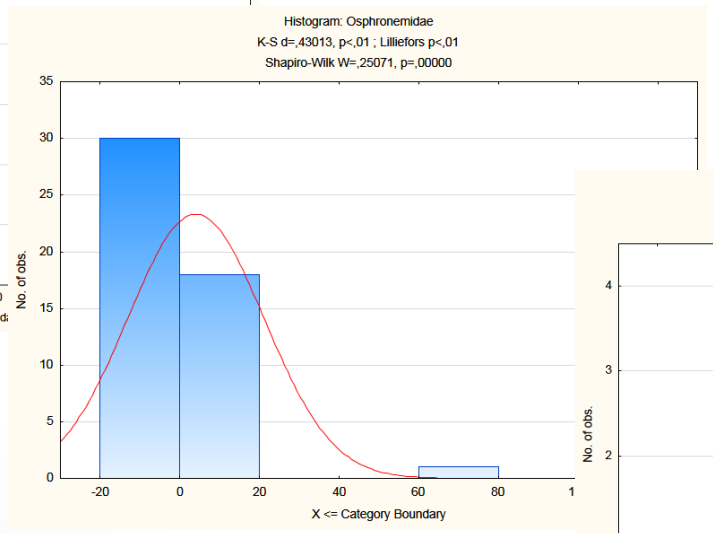
Důvod studie

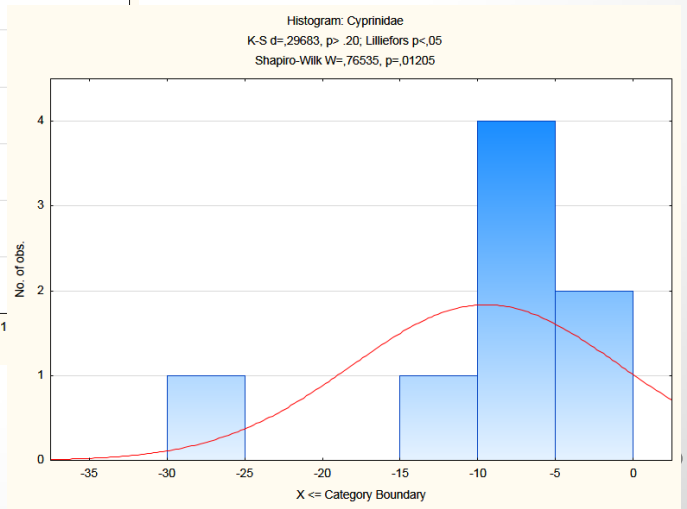
- celosvětový obchod s rybami z dálného východu je velice rozšířený
- antibiotika v chovu nejsou používána pouze jako léčba, ale masivně také jako prevence před transportem
- otázka vzniku antibiotické rezistence
- ve vodě je masivní horizontální přenos genetické informace

Normalita dat

- Použity testy Komolgorov-Smirnov a Shapiro-Wilks
- Rozložení dat není normální - použití neparametrických statistických technik

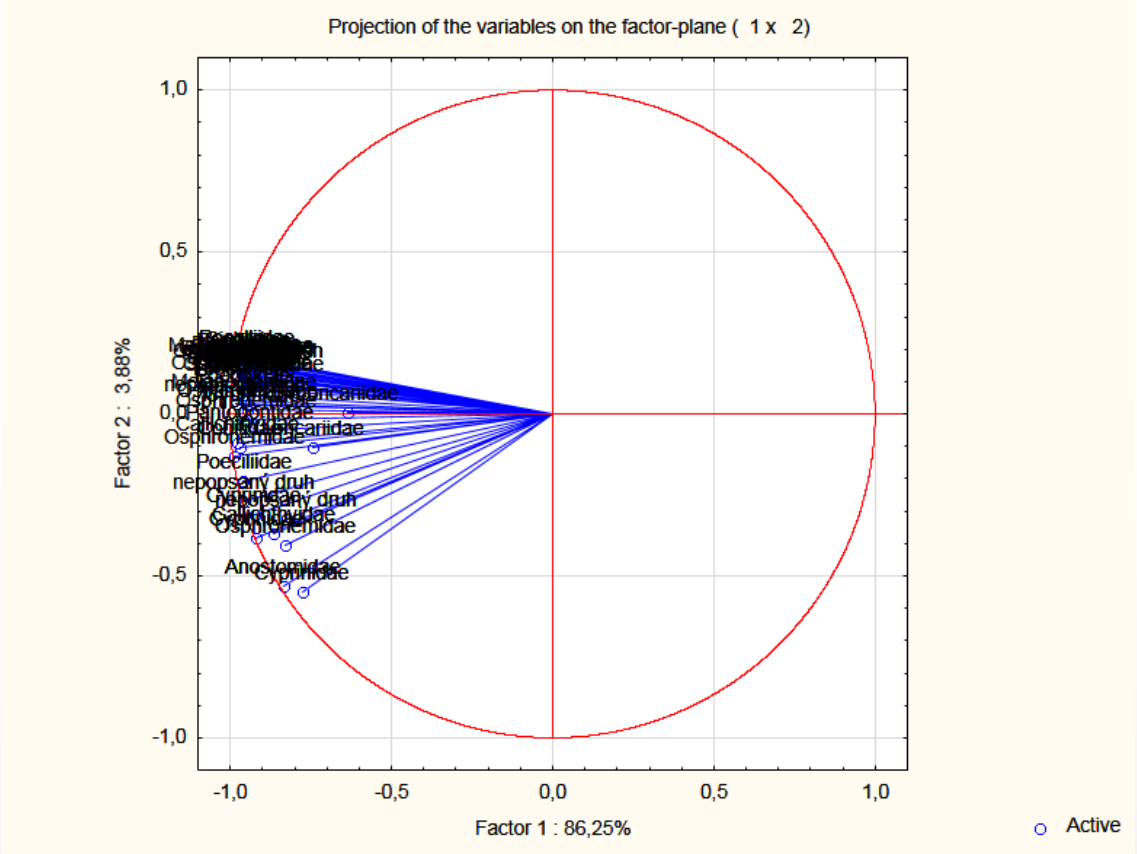






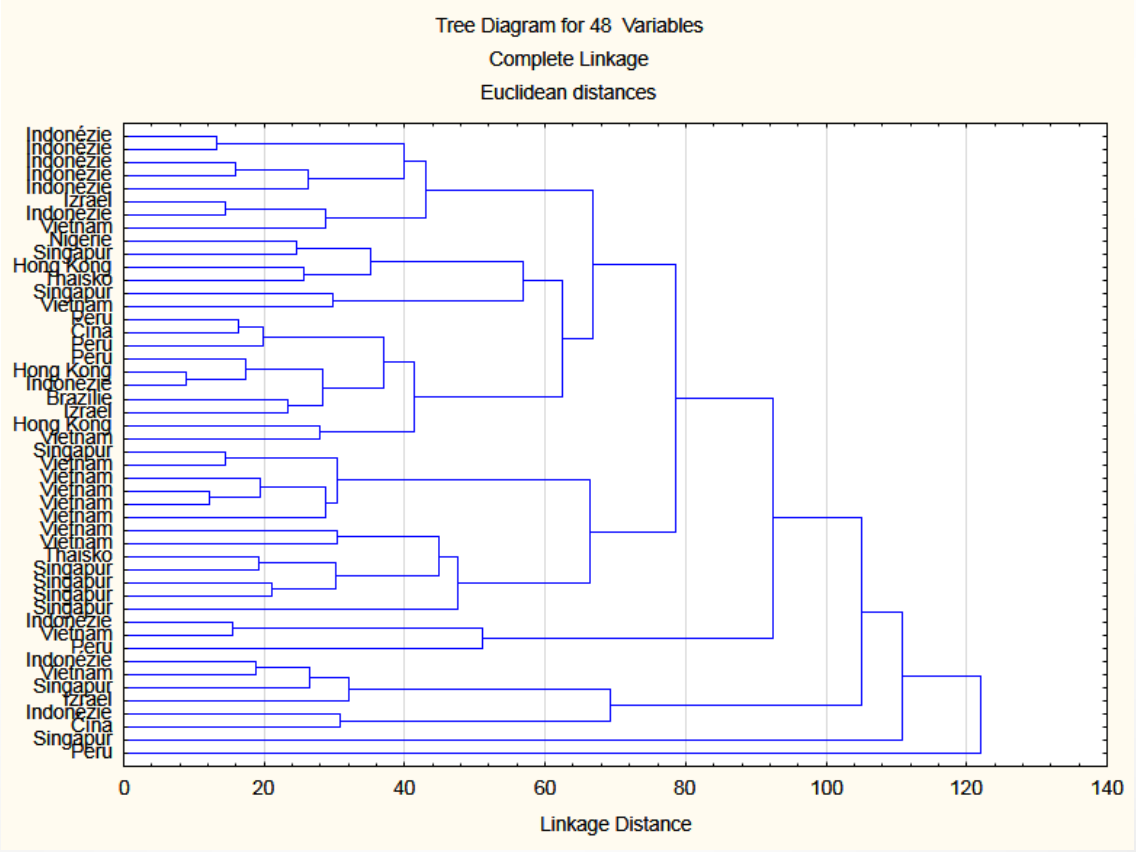
Analýza hlavních komponent (PCA)

- 86,25% variability vyčerpává pouze jediná faktorová osa
- velká korelace jednotlivých skupin



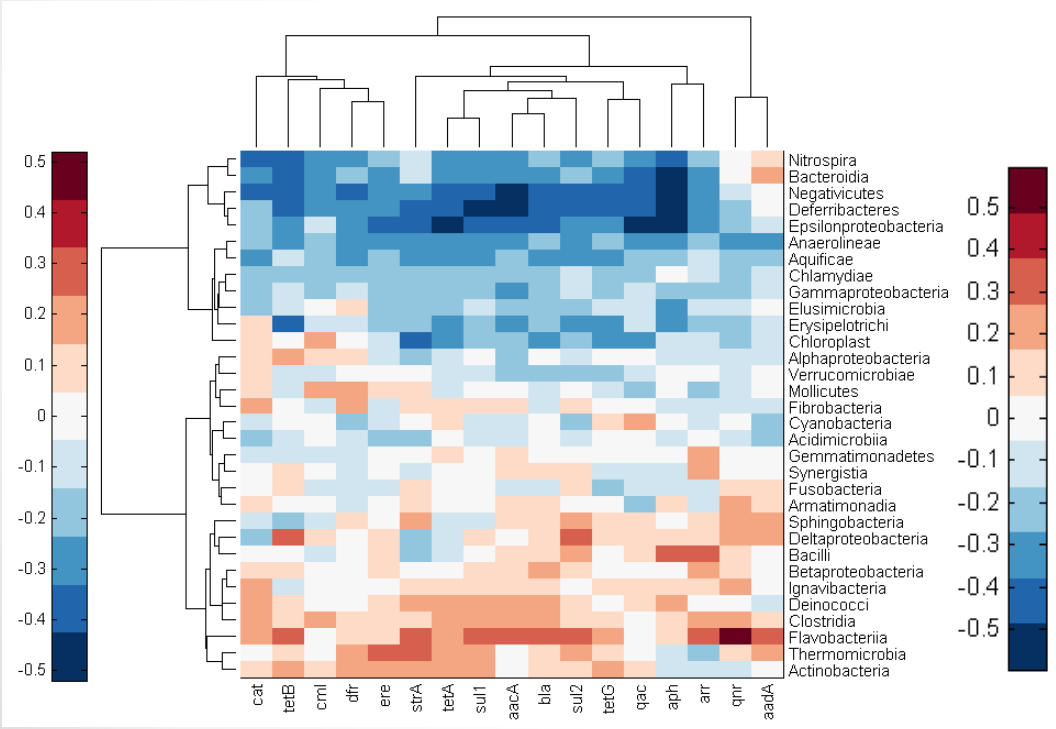
Shluková analýza

- hierarchické aglomerativní shlukování metodou nejvzdálenějšího souseda (complete linkage) na euklidovské metrice

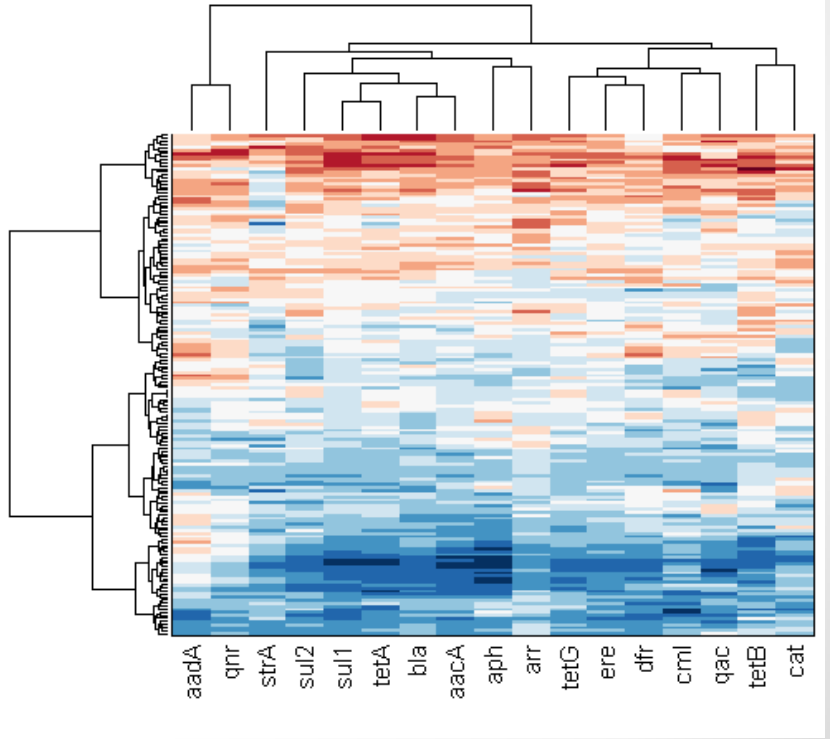


Korelační analýza

- taxon-gen



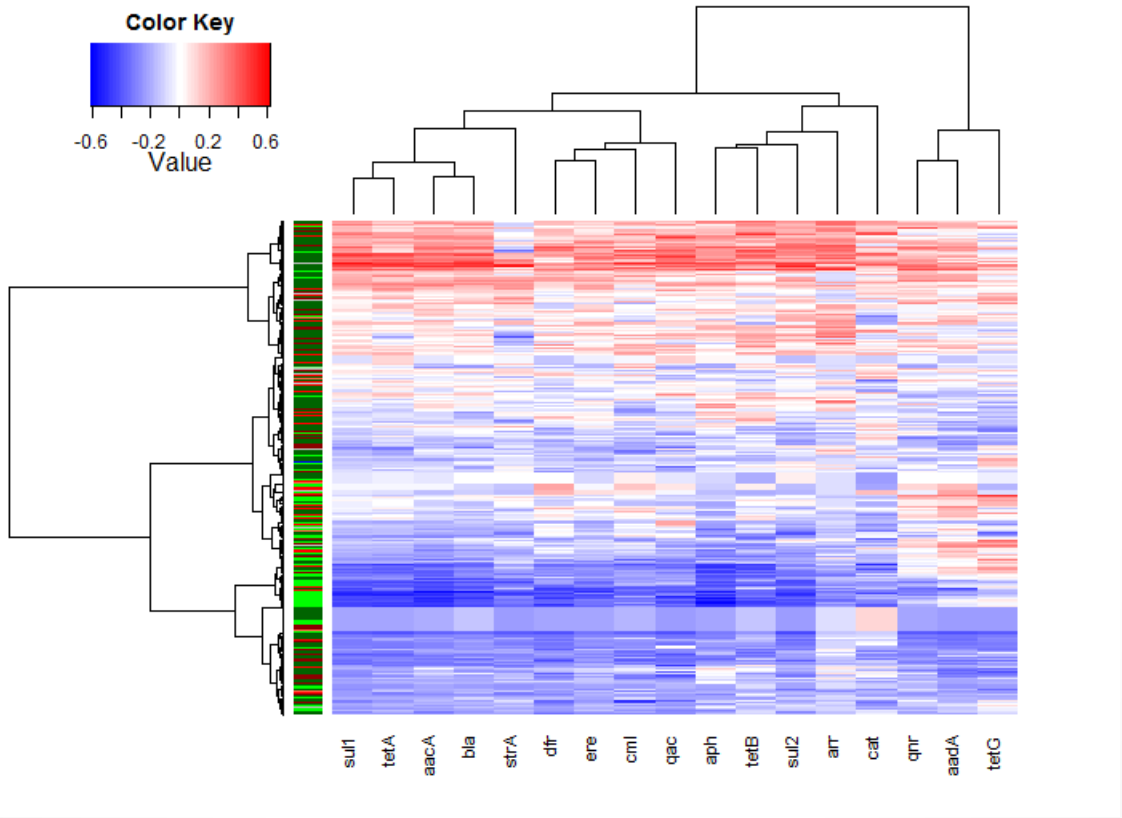
třída



čeleď

Korelační analýza

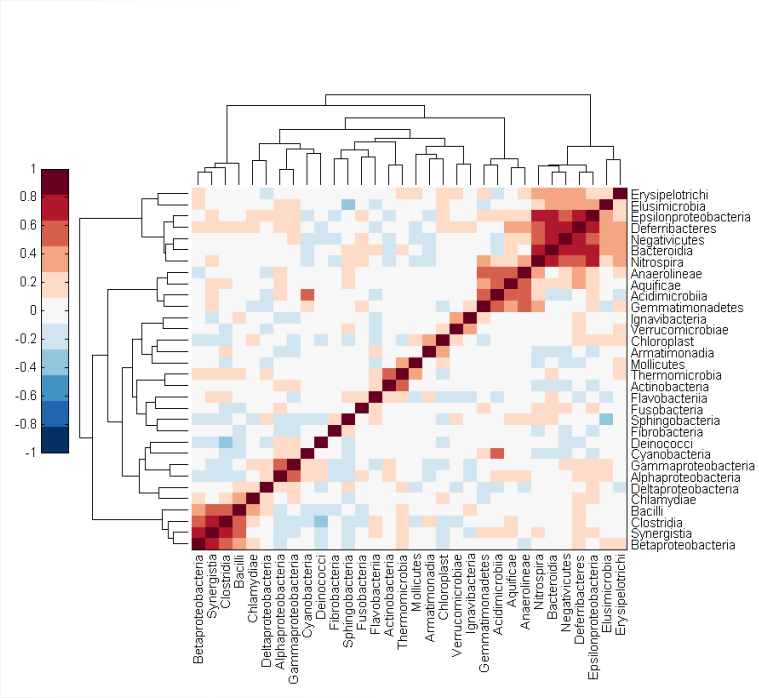
- taxon-gen



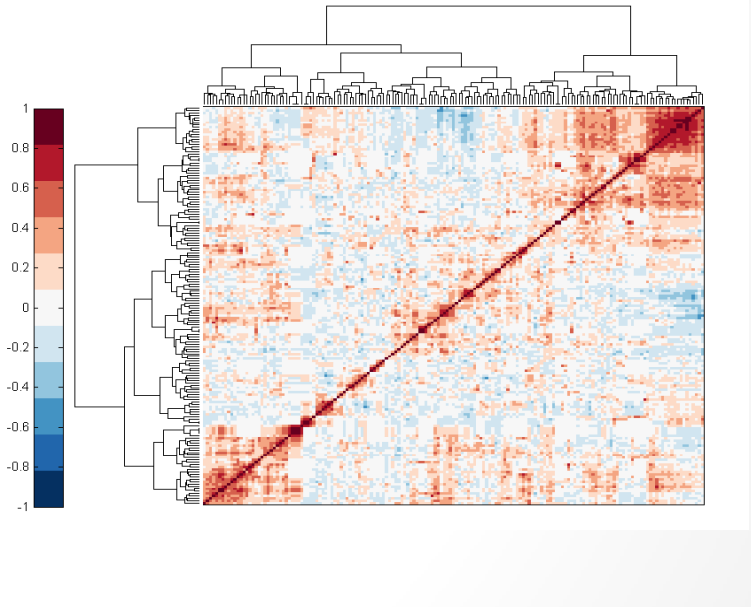
rod

Korelační analýza

- taxon-taxon



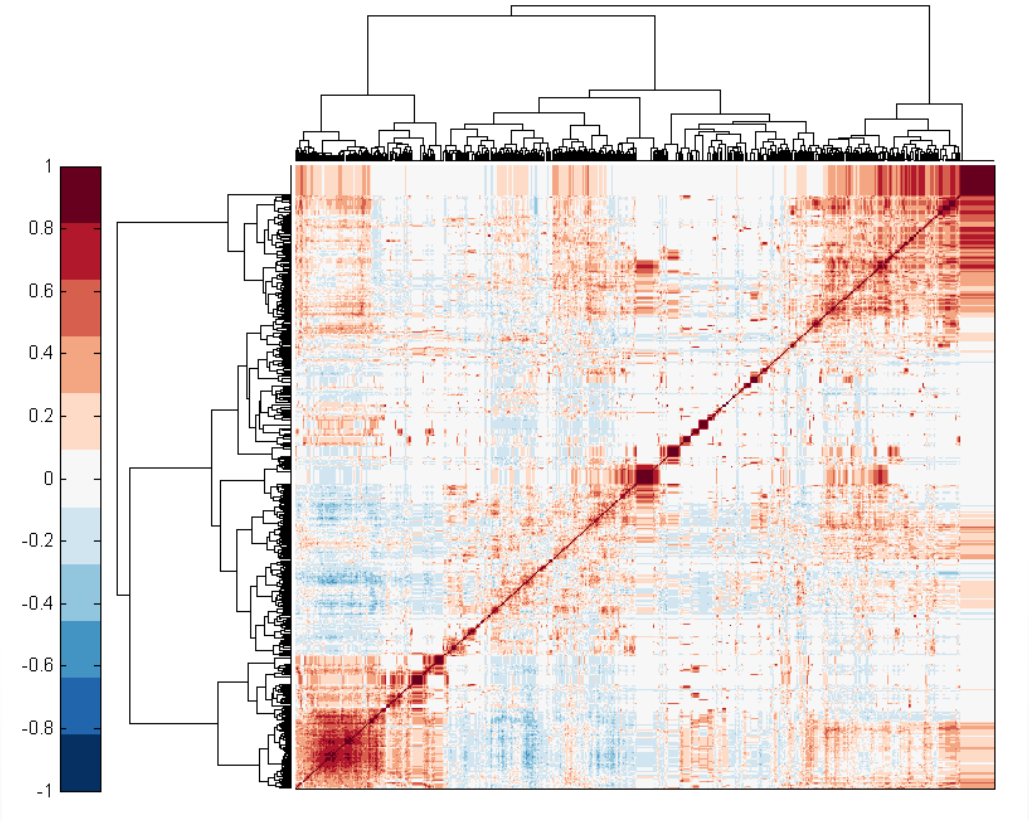
třída



čeleď

Korelační analýza

- taxon-taxon



rod

Závěr

- na základě popisné statistiky byla navrhnuta vizualizační strategie, která doposud pro zdrojová data není typická
- poskytuje přehledné znázornění vztahů v mikroflóře
- dokazuje možnost použití heatmap i na jiná než microarray data

Děkujeme za pozornost