

BIOINFORMATICKÉ SPRACOVANIE PROTEOMICKÝCH DÁT

M. Menkyna¹, M. Marcin¹, M. Šulíková¹, M. Kacířová², D. Tóth³, M. Lenárt⁴, J. Sabo¹

¹Ústav lekárskej a klinickej biofyziky UPJŠ LF, Trieda SNP 1, 040 11 Košice, SR

²CKPV Medipark, UPJŠ LF, Trieda SNP 1, 040 11 Košice, SR

³Gynekologicko-pôrodná klinika, UPJŠ LF, Trieda SNP 1, 040 11 Košice, SR

⁴1. chirurgická klinika, UPJŠ LF, Trieda SNP 1, 040 11 Košice, SR

Úvod

Proteomické merania generujú obrovské množstvo dát, ale cesta od tých prvotných (hmotnostných spektier) až k výsledkom, ktoré by popisovali biologický rozdiel medzi dvoma alebo viacerými skupinami, vie byť krkolomná. Cieľom tejto práce je spraviť prehľad bioinformatického spracovania proteomických dát z hmotnostného spektrometra pri štúdiu zmeny expresie T-lymfocytov s CD4⁺ a CD8⁺ znakom pri karcinóme prsníka.

Materiál a metódy

Pri predloženej práci boli použité dáta z merania krvných vzoriek pacientok s karcinómom prsníka pomocou nanoUHPLC (Ultimate 3000 RSLC Nano) a hmotnostného spektrometra (amaZon speed EDT ion trap). Tieto dáta sú pôvodne v „raw“ formáte, takže poskytujú informáciu iba o MS/MS² spektrách. Na ich spracovanie je potrebný celý rad programov, počnúc DataAnalysis a ProteinScape (Bruker Daltonics Inc.) cez Proline Software (ProfiProteomics) až po samotný softvér na nájdenie a vyhodnotenie relevancie signifikantných signálnych dráh pomocou Gene Set Enrichment Analysis (Broad Institute MIT/Harvard). Tento postup, ktorý je u nás používaný a jednotlivé zmeny dát v ňom budú detailne popísané.

Výsledky

Prezentované budú výsledky z predstavenej štúdie (porovnanie CD4⁺ a CD8⁺), pričom hlavný dôraz bude kladený na postupnú transformáciu dát, ktoré boli namerané. Takisto sa dostaneme k prezentácii výsledných signálnych dráh ako aj štatistické podmienky, ktoré musia spĺňať. Ako jedna z takýchto signálnych dráh sa ukazuje byť glykolyza, ktorej konkrétne charakteristiky sú takisto ukázané.

Záver

V dnešnej dobe sa dá našťastie pri najmodernejších hmotnostných spektrometroch spraviť analýza nameraných vzoriek od začiatku až po koniec v jednom programe, avšak z viacerých dôvodov môže byť tento stav niekedy ťažko dosiahnuteľný, respektíve neponúka možnosť nahliadnuť do vnútorných procesov v jednotlivých krokoch. Prezentácia súčasnej situácie pri práci na starších spektrometroch je dobrá ilustrácia tohto procesu na pochopenie celej logickej postupnosti.

Táto práca bola vypracovaná s finančnou podporou projektov APVV-19-0476 a ITMS2014+: 313011V44.