

Plán standardizace – Program rozvoje metrologie na rok 2017

Zpráva pro závěrečnou oponenturu

Číslo úkolu: PRM VII/14/17

Název: Metrologické postupy v metabolomickém vyšetřování potravin a krmiv.

Zadavatel: Česká republika – Úřad pro technickou normalizaci, metrologii a státní zkušebnictví, organizační složka státu

Řešitel: Vysoká škola chemicko-technologická v Praze, Technická 5, 16628 Praha 6

Vypracoval: prof. Ing. Jana Hajšlová, CSc.....

Schválil:

Datum: 13. 11. 2017

Rozdělovník: 1x ÚNMZ, řešitel, oponenti

Výtisk č.....

OBSAH

1. ÚVOD (anotace).....	2
2. POPIS ŘEŠENÍ ÚKOLU.....	2
3. DOSAŽENÉ VÝSLEDKY.....	4
4. ZÁVĚRY.....	25
5. PŘÍLOHY.....	26

1. ÚVOD (anotace)

Cílem je vytvořit návrh schématu metrologické návaznosti metod založených na metabolomickém fingerprintu a dále politiku (strategii) standardizace těchto inovativních vyšetřovacích postupů od zpracování vzorku přes získání a vyhodnocení spektrometrických záznamů až po chemometrické zpracování a interpretaci výsledků. Úkol byl zadán jako součást Programu rozvoje metrologie na rok 2017. Závěrečná zpráva zahrnuje a doplňuje i dílčí zprávu ze srpna 2017.

2. POPIS ŘEŠENÍ ÚKOLU

Dílčí cíl A:

Získání a vyhodnocení metabolomických fingerprintů / profilů nejméně 3 potravin různého druhu (duben – červenec 2017)

Pro měření metabolomických fingerprintů /profilů bude primárně využita technika ultraúčinné kapalinové chromatografie U-HPLC, případně superkritická fluidní chromatografie (SFC), ve spojení s hmotnostním detektorem typu kvadrupól-analyzátor doby letu QqTOF. Jako metabolom označujeme soubor všech nízkomolekulárních (cca do 1500 Da) sloučenin (složek vzorku) detekovaných v rámci necílového screeningu, jako metabolomický profil pak soubor vybraných sloučenin s určitým společným charakteristických znakem (např. strukturní podobnost).

Čokoláda: problémem bývá zejména neshoda složení s deklarovanou kategorií výrobku, např. snížený obsah kakaového másla, přídavek rostlinných tuků nad hranici apod.

Rostlinné oleje: panenské, za studena lisované: u tohoto typu výrobku se lze setkat s přídavkem levnějších rostlinných olejů, záměna suroviny. Kritické může být i kvalita oleje (žluklost)

Koření: v úvahu, zejména u dražších druhů (např. šafrán, pepř, chilli paprika), připadá přídavek či záměna levnějším rostlinným materiálem, přibarvování apod.

Dílčí cíl B:

Návrh pilotní zkušební metody „Stanovení látkového profilu / metabolomu metodou hmotnostní spektrometrie“ (červenec – listopad 2017)

Na základě experimentů provedených při řešení dílčího cíle A bude vypracován obecný vzor pro standardizaci analytických metod necílového stanovení a zpracování látkových profilů / metabolomu potravin s cílem provádět jejich autentikaci, resp. prokázat falšování vybraných komodit. Metoda bude implementována v laboratoři řešitelského pracoviště a zahrnuta do rozsahu akreditace podle ČSN EN ISO/IEC 17025.

Dílčí cíl C:

Zpracování a validace schématu metrologické návaznosti - (září – listopad 2017)

Ve spolupráci s EURACHEM ČR bude posouzena validita primárních dat, získaných při měření v tomto úkolu a měřicích postupů. Primární data budou potom statisticky zpracována uni- a multivariačními metodami s cílem vhodné interpretace získaných dat i s ohledem na evropskou a českou legislativu. Bude navrženo modelové schéma návaznosti pro metody metabolomického vyšetření.

3. DOSAŽENÉ VÝSLEDKY

3.1 Metabolomika čokolády

Úvod

U čokolád je nejčastějším problémem neshoda s deklarovaným složením výrobku, např. snížený obsah kakaového másla, snížený obsah celkové kakaové sušiny či přídavek nedeklarovaných či nepovolených rostlinných tuků. V některých případech tak může docházet i k nedodržení legislativních požadavků pro čokoládu a tím i k nesprávnému označení druhu výrobku (viz např. Vyhláška č. 76/2003 Sb. v platném znění). K dispozici byl soubor 100 vzorků, který zahrnoval hořké čokolády v maximálním rozmezí obsahu celkové kakaové sušiny (35 – 100 %), mléčné čokolády a další kakaové výrobky (polevy, pochoutky).

Vzhledem k velmi komplexnímu složení těchto výrobků byla pro analýzu zvolena technika ultra-účinné superkritické fluidní chromatografie ve spojení s vysokorozlišovací tandemovou hmotnostní spektrometrií (UHPSFC-HRMS/MS) s ionizací elektrosprejem. Tato moderní metoda nabízí unikátní selektivitu a rychlost separace a je velmi vhodná pro nepolární látky, jakými jsou acylglyceroly, které jsou dominantní složkou čokolád.

Vstupní experimenty

V úvodní fázi projektu byla provedena optimalizace extrakčního postupu a optimalizace chromatografické separace.

Vzorky byly před extrakcí jemně namlety a uchovávány v chladu. Pro extrakci byla do 15ml kyvety navážena 0,1 g vzorku, protřepána s 5 ml extrakční směsí methanol/ethylacetát 1:2 (v/v) po dobu 1 min, odstředěna (5 min, 10000 min⁻¹, 20 °C) a přefiltrována přes mikrofiltr (0,22 µl). Získaný extrakt byl před analýzou 10 x naředěn.

Extrakční rozpouštědlo bylo voleno s cílem zachytit v extraktu co nejširší spektrum látek přítomných ve vzorku. Důraz byl kladen na lipidy, které jsou důležité pro ověření autenticity kakaového másla. Další významné látky, které souvisí s kvalitou čokolád, jsou alkaloidy theobromin a kofein charakterizující tukuprostou kakaovou sušinu. Při výběru extrakčního rozpouštědla byla proto pozornost soustředěna nejen na celkový počet detekovaných látek (features), ale také na intenzitu konkrétních látek, které jsou pro čokoládu charakteristické. Jako účinnější se obvykle pro extrakci osvědčují, díky nižší selektivitě, směsi dvou až tří rozpouštědel. Kromě čistých rozpouštědel (methanol; isopropanol) byly proto testovány také směsi o různé polaritě: (i) hexan/isopropanol 1:3 (v/v), (ii) isopropanol/ethylacetát 2:1 (v/v), (iii) isopropanol/methanol 2:1 (v/v), (iv) methanol/ethylacetát 1:2 (v/v), a (v) methanol/voda 19:1 (v/v). Vzhledem k tomu, že triacylglyceroly se ve vzorcích nacházejí ve vysokých koncentracích, bylo snahou podpořit extrakci především více polárních látek. Při použití nejpolarnější směsi (v) ovšem již docházelo k významnému potlačení extrahovatelnosti triacylglycerolů a jejich derivátů a celkový počet features v pozitivním módu ionizace byl zhruba třetinový (cca 1400) v porovnání s ostatními směsmi. Směsi se střední polaritou poskytovaly pro extrahovatelnost látek širokého rozmezí polaritě dobrý kompromis; jako nevhodnější byla s ohledem na intenzitu významných látek zvolena směs (iv) s celkovým počtem features okolo 4400. S použitím negativního módu ionizace bylo detekováno významně nižší množství features, asi 120 (především mastné kyseliny a diacylglyceroly). Chromatogramy vybraných extraktů jsou znázorněny v **Příloze 6.1.1**.

Přípravná měření

Výhodou separací pomocí SFC je možnost využití jak polárních, tak nepolárních stacionárních fází s mobilní fází o jednom složení. V rámci projektu bylo testováno několik typů stacionárních fází od společnosti Waters dostupných v naší laboratoři: Acquity UPC² HSS C18 SB (3,0 x 100 mm, 1,8 μ m); Acquity UPC² BEH (3,0 x 100 mm, 1,7 μ m); Acquity UPC² Torus 1-AA (3,0 x 100 mm, 1,7 μ m); a Acquity UPC² Torus 2-PIC (3,0 x 100 mm, 1,7 μ m). Jako modifikátor byla použita směs rozpouštědel methanol/acetonitril 4:1 (v/v) s 5 % vody a 15mM mravenčanem amonným, která se osvědčila již u předchozích experimentů. Zatímco „klasická“ nepolární kolona typu C18 se ukázala jako velmi vhodná pro analýzu triacylglycerolů, pro získání komplexního profilu vzorku byla jako nejvhodnější vybrána polární kolona BEH. S jejím využitím dochází nejen k dobré separaci jednotlivých tříd lipidů (triacylglyceroly, diacylglyceroly, fosfolipidy aj.), ale je také vhodná pro analýzu polárnějších látek, jako jsou cukry. Příklady chromatogramů získaných z analýzy vzorku čokolády na uvedených kolonách s využitím pozitivního módu ionizace jsou uvedeny v **Příloze 6.1.2.**

V dalším kroku byla provedena finální optimalizace složení a průtoku mobilní fáze a dalších podmínek separace, jako je teplota kolony a automatická regulace zpětného tlaku (ABPR). Výsledné podmínky použité pro analýzy jsou uvedeny v **Příloze 6.1.3.**

Vyhodnocení výsledků - primární data

K primárnímu zpracování dat byl použit software MarkerLynx, který je součástí programu MassLynx používaného pro ovládání přístroje. Po nastavení základních parametrů, jako je limitní hodnota signál/šum a velikost hmotnostního okna, byla provedena spektrální dekonvoluce a odstranění izotopů. V jednotlivých vzorcích byly nalezeny všechny eluované látky charakterizované hodnotou m/z, retenčním časem a intenzitou. Intenzity jednotlivých iontů jsou v rámci každého vzorku normalizovány vzhledem k celkové intenzitě všech iontů, vnitřní standard nebyl zatím používán.

Následně byla extrahovaná data přenesena do statistického programu SIMCA (Umetrics AB, Švédsko). Pro zjištění primární struktury dat byla provedena analýza hlavních komponent (PCA). Data získaná v obou ionizačních módech byla dobře strukturována a došlo k dobrému oddělení čokolád od ostatních alternativních výrobků. Pro zdůraznění rozdílů mezi kategoriemi vzorků byla použita diskriminační analýza částečných nejmenších čtverců (PLS-DA); v tomto případě vykazoval model rozpoznávací schopnost mezi čokoládami a ostatními výrobky 100 %. Mezi významnými markery čokolád byly identifikovány kromě triacylglycerolů 50:1 a 52:1, oxidovaných triacylglycerolů 52:1+O, 50:1+O, 54:1+O či diacylglycerolů 36:1 a 34:1 také theobromin (m/z 181,0729, t_r 6,4 min) a kofein (m/z 195,0883, t_r 4,79 min).

Vyzkoušeno bylo také vyloučení triacylglycerolů, jejichž intenzity mohou být zkreslené saturací detektoru, z tohoto modelu. Jak je patrné z obr. P3-B, i v takovém případě je oddělení skupin vzorků při PCA dostatečné. Při použití PCA pro zpracování dat z negativního módu ionizace jsou rozdíly mezi skupinami vzorků méně významné než u dat z pozitivního módu ionizace, PLS-DA však opět poskytuje rozpoznávací schopnost 100 % (viz **Příloha 6.1.4.**). Mezi hlavní markery zodpovědné za separaci vzorků se řadí mastné kyseliny palmitová (m/z 255,2317), stearová (m/z 283,2632) a olejová (m/z 281,2473), dále fosfatidylinositol 34:1 (m/z 835,5336) a fosfatidylglyceroly 42:1 (m/z 859,6653) a 44:1 (m/z 885,6655).

Data získaná analýzou čokolád v pozitivním módu ionizace byla dále zpracována PLS-DA s rozdělením vzorků do tří skupin podle obsahu celkové kakaové sušiny (35 - 59 %, 60 – 79 %, 80 – 100 %). Protože přechod mezi těmito skupinami je plynulý, nelze očekávat jejich ostré oddělení. Navíc se celková kakaová sušina skládá ze dvou různých složek, a to obsahu kakaového másla a tukuprosté kakaové sušiny, jejichž poměr nemusí být vždy stejný. Přesto jsou z modelu (viz obrázek P3-D) patrné rozdíly mezi čokoládami s vysokým a nízkým obsahem celkové kakaové sušiny, a to zejména v relativní intenzitě triacylglycerolů 54:0, 54:1, 54:5, 54:6, 50:1, theobrominu, kofeinu či sacharosy (m/z 365,1064, t_r 8,4 min). Vzorky s obsahem mléka se oddělily na základě obsahu triacylglycerolů s krátkým řetězcem, a to 34:0, 36:0, 38:0 aj.

Závěr

Použitý postup umožňuje dobrou charakterizaci jednotlivých vzorků. Získány jsou informace o přítomném tuku (tri-, di- a monoacylglyceroly, mastné kyseliny, fosfolipidy, sfingolipidy, amidy mastných kyselin), tukuprosté kakaové sušině (theobromin, kofein) i dalších přítomných složkách (např. cukru). S využitím PLS-DA je možné rozlišení vzorků čokolád a ostatních kakaových výrobků, dále je možné zjistit přítomnost mléčného tuku a v případě čokolád také orientační zařazení podle obsahu celkové kakaové sušiny.

3.2 Metabolomika olivového oleje

Úvod

Olej získávaný z plodů olivovníku evropského je používán po tisíce let a jeho spotřeba dlouhodobě stoupá. Olivové oleje nejvyšší kvality, převážně extra panenské olivové oleje (EVOO), bývají vzhledem k vysoké nutriční a sensorické hodnotě a ceně velmi často falšovány. Jednak může docházet k nesprávné deklaraci olivových olejů nižší kategorie za oleje extra panenské, k nahrazování části či celého výrobku jiným rostlinným olejem a v neposlední řadě také ke špatné deklaraci země původu. Tato problematika je reflektována evropskou komisí, která v nedávné době přijala přísnější opatření s cílem falšování zamezit či ho alespoň podstatně ztížit.

Cílem této studie je komplexní analýza autentického souboru olivových olejů s využitím techniky kapalinové chromatografie ve spojení s hmotnostně-spektrickou detekcí za účelem budoucí klasifikace vzorků dle geografického původu. Analyzovaný soubor vzorků zahrnoval 154 olivových olejů z Itálie a 58 vzorků ze Španělska. Veškeré italské olivové oleje byly nejvyšší kvality - extra panenské (EVOO). V rámci této studie byla optimalizována extrakční technika i podmínky separace a detekce. Získaná data byla procesována a využita pro multivariační statistickou analýzu za účelem tvorby statistických modelů a identifikaci diskriminativních markerů.

Vstupní experimenty

Prvním krokem analýzy byla optimalizace extrakčního kroku, která by obecně ve studiu metabolomu neměla výrazně diskriminovat některé ze složek vzorku. Cílem extrakce by měla být izolace co největšího počtu komponent vzorku, protože i minoritně zastoupené sloučeniny mohou mít vysokou diskriminační schopnost. Volba strategie izolačního kroku či již výběr konkrétních rozpouštědel značně závisí na povaze zkoumaného vzorku. V rámci této studie byla optimalizace extrakce provedena na menším souboru 50 vzorků autentických EVOO pocházejících z Evropy ($n = 36$) a mimoevropských ($n = 14$).

Vzhledem k majoritnímu zastoupení nepolárního podílu tvořeného především acylglyceroly, jehož složení bývá v olivových olejích obdobné, byla pozornost zaměřena především na obsáhlý komplex fenolických komponent olivového oleje polárního charakteru, které jsou v něm zastoupeny minoritně. Na základě předchozích zkušeností s analýzou matric podobného typu byla testována dvě extrakční činidla - prvním byla směs methanol:voda (1:1, v/v), kterým by měly být izolovány především polární fenolické komponenty olivového oleje a druhým pak okyselený methanol (0,2 % kys. mravenčí), který má i přes relativně vysokou polaritu vyšší schopnost izolace nepolárních složek olivového oleje (**Příloha 6.2.1.** a **6.2.2.**).

Obecně by hlavním kritériem výběru extrakčního rozpouštědla měl být počet proměnných, tj. potenciálních markerů, charakterizovaných hmotou iontu (m/z) a retenčním časem (RT, min) získaných po zpracování primárních dat. Rozpouštědlo, poskytující nejvyšší počet těchto proměnných, by pak mělo být zvoleno jako ideální. Vzhledem k nastavení této studie a dostupnosti dalšího setu autentických vzorků olivového oleje byla zvolena strategie popsána výše a jako vhodnější extrakční rozpouštědlo bylo vybráno takové, které vedlo k získání statistických modelů s lepšími charakteristikami, konkrétně rozpoznávací (R²Y) a predikční schopností (Q²). Statistická rozdílnost v rámci souboru vzorků dle geografického původu jednoznačně vypovídala pro variantu směsi methanol:voda (R²Y > 0,9 a Q² > 0,6), zatímco pro okyselený methanol byly hodnoty R²Y a Q² dokonce i záporné.

Volba zaměření se na „nepolární frakci“ by mohla být vhodnější s využitím metaboliky realizované pomocí instrumentace GC–MS, zatímco pro LC–MS s testovanými extrakčními činidly se jeví jako vhodnější zaměření na fenolický komplex olivového oleje.

Přípravná měření

Pro vlastní analýzu byla využita ultra-účinná chromatografie ve spojení s vysokorozlišovací tandemovou hmotnostně-spektrální detekcí (UHPLC–H²MS/MS; UltiMate 3000 RSLD - Q-Exactive Plus; Thermo Scientific). Přehled podmínek ionizace, separace a detekce je uveden v **Příloze 6.2.3. – 6.2.6.** Pro chromatografickou separaci byla zvolena analytická kolona s reverzní fází (Acquity BEH C18, 100 x 2,1, 1,7 μ m; Waters), která je obecně vhodná pro širokou škálu analytů a vyznačuje se značnou stabilitou. Jako mobilní fáze byl zvolen methanol a voda, obě fáze s přídavkem mravenčanu amonného (5 mM) a kyseliny mravenčí (0,1 %), které zajišťují dostatečnou stabilitu pH mobilní fáze a tudíž reprodukovatelné výsledky v rámci delší analýzy či sérií analýz. V rámci této studie byl optimalizován eluční gradient, jehož cílem byla maximální distribuce separovaných složek, což by mělo být obecně stejné pro metabolickou analýzu s využitím chromatografie. Jeden z prvních testovaných a finální gradient jsou uvedeny v **Příloze 6.2.7.** společně s 3D iontovými mapami vyjadřujícími distribuci separovaných komponent dle m/z v čase.

HRMS/MS detekce byla realizována s využitím MS experimentu *fullMS-data dependentMS2*, v celo-spektrálním módu (fullMS) v hmotnostním rozsahu m/z 100 - 1200 s fragmentací 20 iontů s nejvyšším signálem v rámci předchozího *fullMS* spektra. Nastavení fragmentace je nutno zvolit tak, aby bylo možné v úzkém časovém intervalu kontinuálně sbírat *fullMS* spektra v širokém rozsahu hmot m/z nutných pro konstrukci chromatografických píků. Proto je nutné brát v potaz i průměrnou šířku chromatografického píku související s analytickou kolonou a nastavenou rozlišovací schopností detektoru (70 000 FWHM). Data po fragmentaci nachází využití až v posledním kroku po multivariační analýze pro identifikaci diskriminativních sloučenin - markerů. Pro fragmentaci bylo zvoleno generické nastavení tří hladin kolizní energie po sobě (28, 35 a 42 %), kterou lze získat fragmentační spektra méně i více stabilních látek.

Vyhodnocení výsledků - primární data

Získaná data byla procesována pomocí softwaru CompoundDiscoverer 2.2 (Thermo Scientific), který po zadání klasifikačních kritérií zpracuje primární data a výstupem je rozsáhlá matrice intenzit signálů pro danou proměnnou charakterizovanou hmotou m/z a retenčním čase ($m/z@RT$) v daném vzorku, která jsou dále zpracovávána pomocí nástrojů multivariační analýzy. Vlastní procesování dat zahrnuje extrakci veškerých hmot m/z a jejich signálu ze všech vzorků, z nichž jsou dále sestaveny chromatogramy. Vzhledem k tomu, že jsou v MS spektrech zastoupeny různé adukty a také izotopové příspěvky té samé sloučeniny, je nutné tyto redundantní signály nejprve vyloučit. V následném kroku je nutné provést tzv. zarovnání retenčních časů vzhledem k drobným posunům, drobný vliv může mít i míra vyhlazení píku, díky čemuž může dojít k posunu vrcholku píku o jednotky setin minuty. Dále je nutné nastavit limit pro signál, nižší hodnoty jsou vyloučeny a v dalším kroku poté odstranit proměnné ($m/z@RT$) s velmi nízkým zastoupením v souboru analyzovaných vzorků. V případě analýzy olivového oleje, kdy jsou klasifikačním kritériem 2 obsáhlé skupiny, je spíše na škodu pracovat s proměnnými se zastoupením v řádu jednotek výskytů v souboru desítek až stovek vzorků.

Tento předfinální set dat je nakonec vhodné doplnit o signály s nižší intenzitou, než byl stanovený limit, pro proměnné, u kterých by v případě některých vzorků mohly zůstat prázdná pole, což by mohlo vést ke zkreslení míry diskriminativnosti dané proměnné. Tento krok bývá nazýván často jako „vyplnění mezer“.

Vzhledem k variabilitě intenzity signálů s využitím LC–MS instrumentace je nezbytné v posledním kroku zpracování primárních dat provést normalizaci. Jednotlivé intenzity signálů daného vzorku jsou tak vyjádřeny jako podíl vůči sumě veškerých signálů. Obecně lze po zpracování LC–HRMS dat očekávat přibližně stovky až jednotky tisíc proměnných. Využití vnitřních standardů pro normalizaci nelze pro LC–HRMS data jednoznačně doporučit, neboť signál jednotlivých komponent se nemění stejně v případě všech sledovaných iontů, a proto ani několik vnitřních standardů eluujících se v různých retenčních časech nelze označit jako ideální řešení. Využití více vnitřních standardů by mělo smysl při zavádění metabolického fingerprintingu v dlouhodobém měřítku, kdy by mohlo docházet k drobnému systematickému posunu retenčních časů v důsledku vlivu analytické kolony (např. jiná šarže), servisních zásahů do fluidiky kapalinového chromatogramu (výměna komponent - kapilár, nástřikové smyčky a stříkačky, ap.). Dále je nutno zmínit obtížnost jednoduchého transferu zavedené metodiky mezi různými instrumenty, neboť jejich citlivost vůči různým sloučeninám může být značně variabilní, např. z důvodu jiné konstrukce stroje (ionizační zdroj, ap.).

Řada specializovaných softwarů poskytovaných výrobcí umožňuje provedení výše popsaných kroků genericky, avšak vždy je nutné brát v potaz aspekty související s využitou instrumentací, analytickou kolonou a typem analyzované matrice. Mimo těchto SW jsou pro některé typy dat dostupné i freeware SW (MZmine, MetAlign, ad.), který může najít využití společně se SW pro konverzi primárních dat (Raw Converter, ReAdW, ad.) v případě, kdy není dostupný SW přímo od výrobce.

Chemometrické zhodnocení primární dat

Získaná data byla zpracovávána s využitím SW SIMCA v. 13 (Umetrics AB, Švédsko), která umožňuje provedení multivariační analýzy, nejprve analýzou hlavních komponent (principal component analysis, PCA), která má za účel snížit dimenzionalitu dat a zároveň zachovat maximální možné množství informací pro klasifikaci vzorků do požadovaných kategorií. Původní rozsáhlá matice dat je transformována do tzv. hlavních komponent, které jsou dále využity pro diskriminační analýzu. Zde je teoreticky možné vyloučit z datového souboru statisticky odlehle vzorky, které by mohly mít negativní vliv na kvalitu statistických modelů. Pokud by však nešlo o vzorky významně odlehle, je vhodné je zachovat, neboť zvyšují robustnost finálních statistických modelů.

V rámci studie metabolomiky olivového oleje žádný takový vzorek vyloučen nebyl. Hlavními kvalitativními parametry statistických modelů jsou rozpoznávací (R²Y) a predikční (Q²) schopnost. Rozpoznávací schopnost modelu popisuje, jak by byly klasifikovány vzorky vyřazené z daného souboru a poté testovány jeden po druhém. Predikční schopnost naopak popisuje pravděpodobnost klasifikace neznámých vzorků po vyloučení náhodné 1/7 vzorků za předpokladu otestování celého souboru vzorků. Vybraný počet hlavních komponent (PC) byl dále zpracován metodou ortogonální lineární diskriminační analýzy částečných nejmenších čtverců (orthogonal partial least square discriminant analysis, OPLS-DA). S využitím SW Simca je vhodný počet PC pro OPLS-DA automaticky vypočten, při zahrnutí většího počtu PC hrozí získání přetřénovaného modelu, tj. ideálního pro zkoumaný trénovací soubor vzorků s velmi dobrými charakteristikami, který by ale nebyl nepoužitelný pro následné analýzy.

Q2 vypovídá o kvalitě modelu lépe, vždy vychází hůře nežli R2Y (vždy v intervalu 0 - 1), přičemž za kvalitní model je v případě LC-MS dat považován model s $Q2 > 0,5$ a perfektní s $Q2 > 0,9$, avšak vždy záleží na vstupním souboru vzorků a kvalitě zpracování primárních dat. V každém případě by neměl být rozdíl mezi těmito dvěma hodnotami příliš velký, což ukazuje na nepříliš kvalitní výsledek pro daný účel, který může být způsoben např. významnou variabilitou mezi zkoumanými vzorky či jejich nedostatečným počtem. V ideálním případě by součástí trénovacího souboru měly být desítky až stovky vzorků v závislosti na předem definovaném účelu.

V rámci naší studie byla z pohledu statistiky zjištěna významně kvalitnější data získaná při ionizaci ESI+, která jsou dále diskutována. Přehled grafických výstupů multivariační analýzy společně s charakteristikami statistických modelů jsou uvedeny v **Příloze 6.2.8**. Klasifikačním kritériem byla země původu EVOO (Itálie, Španělsko), pro které byly zjištěny velmi dobré charakteristiky OPLS-DA modelu $R2Y = 0,74$ a $Q2 = 0,70$. V posledním kroku byla pozornost zaměřena na proměnné, které mají nejvyšší vliv na klasifikaci vzorků dle specifikovaných kritérií. Tyto markery lze graficky znázornit ve formě tzv. S-křivky (S-plot), kde jsou zobrazeny potenciální markery v protějších rozích grafu, naopak proměnné podobné pro klasifikované soubory dat jsou orientovány blíže středu. Alternativně lze diskriminativnost proměnných vyjádřit pomocí tzv. VIP křivky (variable importance in projection, VIP). Sloučeniny s nejvyššími hodnotami VIP jsou nejvíce diskriminativní. S využitím procesovacího SW, online databází a odborné literatury lze takovéto látky identifikovat. V některých případech je to ale velmi komplikované, např. v olivovém oleji obsažené terpenické látky se vyskytují ve formě mnoha desítek izomerů a i s využitím LC-HRMS je jejich jednoznačná identifikace v mnoha případech nedosažitelná, proto bývá využíváno termínu „předběžně identifikováno“. V rámci studie zaměřené na geografický původ EVOO byly jako nejvíce diskriminativní sloučeniny pro skupinu italských EVOO identifikovány fenolické látky (kyselina elenolová, aglykon ligstrosidu a jeho deriváty, tyrosol, ad.), zatímco pro španělské EVOO byly diskriminativní spíše nepolární sloučeniny (acylglyceroly, triterpenoidy, ad.), viz. **Příloha 6.2.9**. Obecným závěrem může být vyšší zastoupení hodnotných fenolických sloučenin EVOO pocházejících z Itálie.

Závěr

Metabolický fingerprinting představuje zajímavou strategii pro účely autentikace potravin a jejich surovin, které bývají falšované vzhledem k jejich nedostupnosti či vysoké ceně. Dobře realizovatelným cílem by mohla být zejména identifikace vzorků, jejichž složení může být alespoň částečně odlišné vzhledem k použité výrobní technologii či rozdílné složce vzorku. Pro standardizaci databází fingerprintů lze obecně doporučit analýzu rozsáhlých souborů vzorků po delší časový úsek, aby byla co nejvíce minimalizována variabilita získaných dat. Dále je žádoucí disponovat referenčními vzorky, které by mohly být analyzovány v průběhu času za účelem zlepšení dlouhodobé opakovatelnosti analýz.

3.3 Metabolomika koření- černý pepř

Úvod

Pro pilotní studii metabolického zpracování dat byl vybrán černý pepř, přesněji sušené plody rostliny *piper nigrum*, která patří z hlediska objemu mezi nejvýznamnější druhy koření. Falšování je často realizováno náhradou části nebo celého objemu suroviny levnějším materiálem různého původu. Bylo popsáno například použití semen papayi, jalovce, případně jiných druhů pepře (*P. attenuatum*, *P. galeatum*). V případě mletého pepře lze k výše uvedenému seznamu zástupných surovin přidat spent (reziduální materiál po extrakci oleoresinu), slupky černého pepře, různé druhy škrobu a další blíže nedefinované mleté materiály.

Vstupní experimenty

Vzhledem k intenzivnímu fingerprintu těkavých látek, které mletý pepř emituje, byla pro autentikaci vybrána technika mikroextrakce na tuhou fázi ve variantě headspace (HS-SPME), která slouží k zachycení těkavých látek v prostoru nad vzorkem, spojená s plynovou chromatografií a hmotnostně spektrometrickou detekcí (GC-MS). Z hlediska MS detekce byl použit průletový hmotnostní analyzátor (TOF), zejména z důvodu možnosti provedení dekonvoluce v získaných datech.

Podmínky mikroextrakce na tuhou fázi v provedení „Headspace“ (HS-SPME) byly optimalizovány s cílem dosáhnout reprezentativního vzorkování těkavých látek emitovaných vzorky černého pepře a případně jeho zástupných surovin. V tomto ohledu byla testována velikost navážky vzorku, přísadek nasyceného roztoku NaCl, teploty a doby inkubace vzorku a sorpce těkavých látek na SPME vlákno. Pro extrakci bylo vybráno osvědčené SPME vlákno pro automatizovanou extrakci se stacionární fází PDMS/CAR/DVB.

Navážka 100 mg suchého homogenizovaného vzorku se ukázala být dostatečně reprezentativní a postačovala i pro dosažení vyhovující opakovatelnosti naměřených dat. Přísadek nasyceného roztoku NaCl se neosvědčil, jelikož nebylo možno vytvořit suspenzi.

Z důvodu celkově značné intenzity získaného fingerprintu, který při běžném základním nastavení SPME metody (10 min inkubace, 10 min sorpce, vše při 40 °C) prakticky neumožňoval dostatečnou chromatografickou separaci, byla testována různá doba extrakce těkavých sloučenin na vlákno a split poměr při nástřiku do plynového chromatografu. Výsledné podmínky měření jsou shrnuty v **Příloze 6.3.1**.

Přípravná měření

Optimalizovanou metodou využívající HS-SPME vzorkování těkavých látek následované plynově chromatografickou separací (Agilent 6890; Agilent; USA) a hmotnostně spektrometrickou detekcí (TruTOF; LECO; USA) bylo proměřeno 15 autentických vzorků (**Příloha 6.3.2**). Tyto autentické vzorky byly získány v rámci spolupráce s průmyslovým partnerem Gurmeko s.r.o. a představují vzorky pepře, nakoupené jako nemleté od důvěryhodných zdrojů a byly homogenizovány ve firmě Gurmeko. Vedle těchto vzorků bylo proměřeno i několik vzorků zástupných surovin – potenciálních adulterantů (spent-zbytek po extrakci oleoresinu a slupky černého pepře). V neposlední řadě byly získány i fingerprinty těkavých látek maloobchodních i velkoobchodních vzorků, které byly označeny jako podezřelé z hlediska nízké ceny nebo senzorických vlastností.

Vyhodnocení výsledků - primární data

Primární zpracování získaných dat spočívalo v automatickém vyhledání chromatografických píků a dekonvoluci hmotnostně spektrometrických záznamů pomocí přístrojového software ChromaTOF (LECO). Dále byly uvažovány jen sloučeniny s přednastaveným poměrem signálu k šumu (S/N) 20, 50. V doplňku software ChromaTOF – Statistical Compare byl proveden automatický „peak alignment“ spočívající v přiřazení signálů ve skupině vzorků na základě shody retenčního času a současně shody spektra větší než 60 %. Pro zpětné vyhledávání byl parametr S/N snižen na 25 (resp. 15). Finální nastavení podmínek „peak alignment“ ukazuje **Příloha 6.3.1**. Setříděná data byla zpracovávána v programu MS Excel. Chybějící hodnoty, tj případy kdy sloučenina nebyla detekována, resp. parametr S/N nepřesáhl nastavenou mez, byly nahrazeny poloviční hodnotou intenzity šumu. Dále byly odstraněny sloučeniny, jejichž odezva (plocha kvantifikačního iontu) byla větší než $8,5 \times 10^7$, jelikož u těchto sloučenin byl překročen lineární rozsah detektoru a tedy jejich intenzita neodpovídala skutečnosti. Takto upravená data byla transformována metodou normalizace na „konstantní řádek“ – plocha vybraného kvantifikačního iontu dané sloučeniny (automaticky vybraná unikátní m/z v retenčním čase) je dělena sumou ploch všech kvantifikačních iontů v daném vzorku a vynásobená 100. Schéma matice zpracovaných dat ukazující relativní odezvy je uvedeno v **Příloze 6.3.3**. V rámci experimentů nebyl používán vnitřní standard.

Primární identifikace těkavých látek byla provedena automatickým srovnáním naměřených hmotnostních spekter s těmi uloženými v knihovně NIST 2.0 (2014). Tato identifikace však není dostatečná pro isomerní sloučeniny, protože v hmotnostním spektru poskytují stejné fragmenty o podobných intenzitách. Proto pro další identifikaci byla u významných sloučenin použita navíc metoda porovnání lineárních retenčních indexů.

Chemometrické zhodnocení primární dat

Takto upravená data byla doplněna o dostupné deskriptory skupin vzorků a použita jako vstupní matice pro statistické zpracování v programu SIMCA (Umetrics, Sweden), kde bylo použito rozdělení PCA (analýza hlavních komponent) a následně OPLS-DA (ortogonální diskriminační analýza nejmenších čtverců). Pro prvotní PCA rozdělení byly skupiny vzorků označeny jako autentické (doručené od důvěryhodných dodavatelů a namleté ve firmě Gurmeko) a jako černý pepř pro vzorky získané z maloobchodu i velkoobchodu. Výsledky v **příloze 6.3.5. 2** ukazují, že zatímco vzorky autentického pepře tvoří relativně těsný cluster, jeden vzorek označený jako pepř je umístěn mimo 95% elipsu a další 3 vzorky jsou významně odlehlé. PCA, které obsahuje i vzorky zástupných surovin a využívá barevné zvýraznění vzorků podezřelých (**příloha 6.3.6.**) ukazuje, že všechny odlehlé vzorky byli současně označeny za podezřelé na základě jejich ceny nebo sensorických vlastností. Z tohoto rozdělení byly odstraněny vzorky, u kterých nebyly zaručeny důvěryhodné deklarace původu a z takto redukované matice dat byl vytvořen a validován OPLS-DA model. Parametry modelu byly následující: $R^2X=80\%$, $R^2Y=96\%$, $Q^2 = 90\%$. R^2 ukazuje variaci popsanou všemi komponenty modelu, zatímco Q^2 indikuje, jak přesně může model předpovídat příslušnost vzorkům ke skupinám (třídám). Dále byl model podroben manuální validaci, která probíhala tak, že z modelu byl odstraněn vždy jeden vzorek a na něm bylo testováno, zda bude zařazen do správné třídy. Rozpoznávací a predikční schopnost tohoto modelu zjištěná tímto způsobem byla 100 %. Takto validovaný model byl použit pro otestování autenticity vzorků, jejichž původ nebyl důvěryhodně doložen. Všechny velkoobchodní i maloobchodní vzorky označené za podezřelé byly přiřazeny do skupin slupky, spent nebo směsné koření. Vedle toho ve skupině maloobchodních vzorků byly zjištěny dva další vzorky, které model přiřadil do skupiny slupek (**Příloha 6.3.7.**).

OPLS-DA model, respektive jeho graf vlivů (S-plot) byl použit pro identifikaci sloučenin nejvýznamnějších pro separaci zástupných surovin od autentických vzorků. Tímto způsobem bylo zjištěno, že pro skupinu autentických vzorků je nejvýznamnější sloučenina β -myrcen. Residuální organická rozpouštědla ethylacetát a aceton, používané pro extrakci oleoresinu byly spolu s kyselinou octovou důležitými markery směsného koření. Aceton a dekan byly významné pro spent a vyšší relativní zastoupení kopaenu a karyofylenu bylo významné pro slupky černého pepře. Z hlediska přítomnosti organických rozpouštědel nutno poznamenat, že jejich přítomnost nebyla prokázána v ostatních vzorcích připravovaných ve stejný čas ve stejné laboratoři ani ve slepých stanoveních provedených ve stejné sekvenci, což dokazuje, že přítomnost těchto látek je spíše spojena s výrobou daného produktu než kontaminací způsobenou v laboratoři.

OPLS-DA model, respektive jeho graf vlivů (S-plot) byl použit pro identifikaci sloučenin nejvíce významných pro separaci zástupných surovin od autentických vzorků. Tímto způsobem bylo zjištěno, že pro skupinu autentických vzorků je nejvýznamnější sloučenina β -myrcen. Residuální organická rozpouštědla ethylacetát a aceton, používané pro extrakci oleoresinu byly spolu s kyselinou octovou důležitými markery směsného koření. Aceton a dekan byly významné pro spent a vyšší relativní zastoupení kopaenu a karyofylenu bylo významné pro slupky černého pepře. Z hlediska přítomnosti organických rozpouštědel nutno poznamenat, že jejich přítomnost nebyla prokázána v ostatních vzorcích připravovaných ve stejný čas ve stejné laboratoři ani ve slepých stanoveních provedených ve stejné sekvenci, což dokazuje, že přítomnost těchto látek je spojena s výrobou daného produktu nikoli sekundární kontaminací v laboratoři.

Závěr

Metabolomický fingerprinting těkavých látek černého pepře technikou HS-SPME-GC-TOFMS se jeví jako efektivní nástroj pro rozpoznání autentického pepře od směsného koření, obsahujícího vedle pepře také spent a oleoresin a také pro odlišení těchto zástupných surovin. Jelikož, vzhledem k obtížnosti získání autentických vzorků, byl uvedený model vytvořen pouze na základě 15 autentických vzorků, bylo by žádoucí tento soubor vzorků doplnit a dále ověřit funkci modelu. Z důvodu malého množství vzorků nebylo možno dále provádět autentikaci původu vzorku nebo roku sklizně.

3.4 Metoda „Stanovení látkového profilu / metabolomu technikami MS“

Obecný procesní diagram metabolomické analýzy lze pro danou oblast měření znázornit následovně:

1. **Výběr SOP a přezkoumání postupu: „fit-to-purpose“**

2. **Výběr a zpracování testovacího souboru (vzorků)**

3. **Optimalizace podmínek měření a verifikace měř. systému**

4. **Měření primárních dat souboru vzorků vč. QC**

5. **Zpracování dat (filtrace, normalizace, redukce,...)**

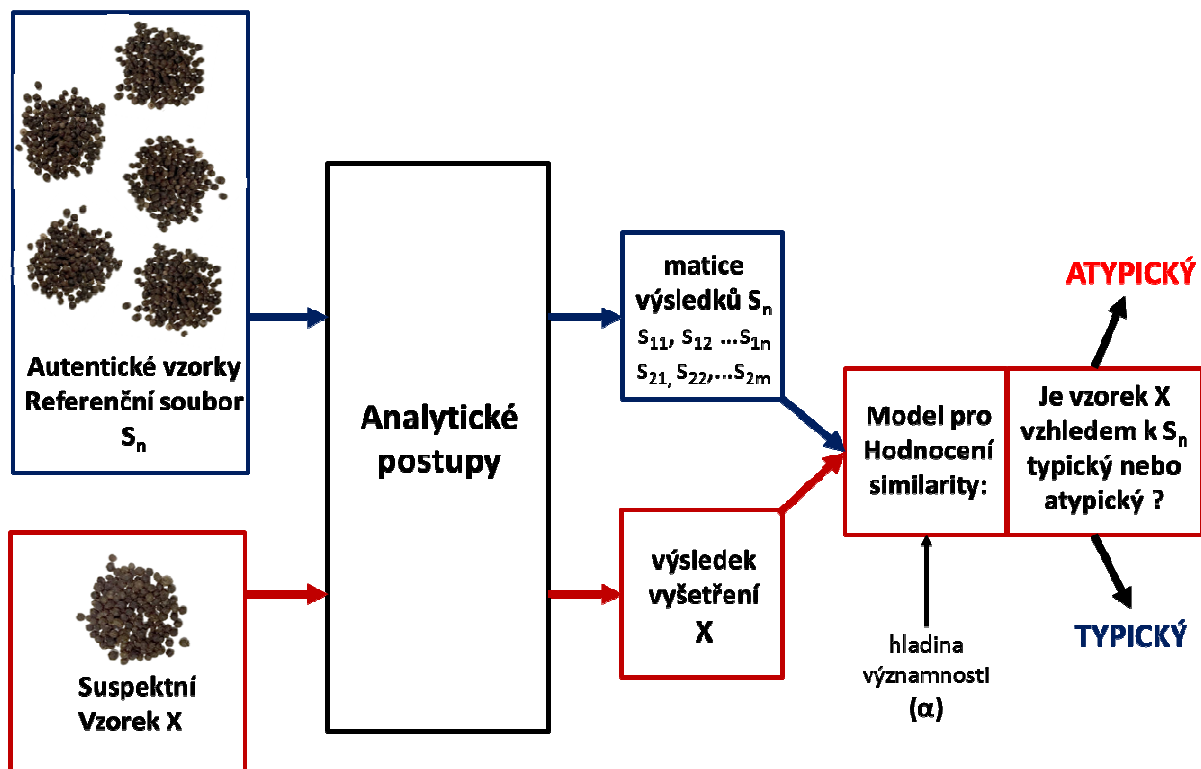
6. **Multivariační analýza dat a interpretace**

V tomto smyslu byl pak také vypracován obecný vzor pro standardizaci analytických metod necílového stanovení a zpracování látkových profilů / metabolomu potravin s cílem provádět jejich autentikaci, resp. prokázat falšování vybraných komodit. Konkrétní validované metody (postupy, podmínky, parametry) byly do tohoto vzoru zpracovány a úspěšně implementovány v akreditované Metrologické a zkušební laboratoři řešitelského pracoviště.

Na základě výsledků předchozího dílčího cíle A byly jako případová studie zpracovány a validovány vzorové SOP pro dvě optimalizované a standardizované pilotní analytické metody nazvané „**Stanovení látkového profilu / metabolomu [potraviny] metodou hmotnostní spektrometrie**“. Metoda **SOP X1** je určena pro měřicí postupy založené na plynové chromatografii s MS detekcí a je zpracována pro metabolom koření, zatímco metoda **SOP X2** je vypracována pro kapalinovou chromatografii ve spojení s MS detekcí a je zpracována pro olivové oleje. **Obě tyto SOP jsou samostatnými přílohami této zprávy.**

Protože však tato Závěrečná zpráva obsahuje i výsledky z Průběžné zprávy, které byly pro navržené SOP následně využity, jsou některé grafy a tabulky uváděny jak v této zprávě, tak i v SOP, neboť SOP budou dále využitelné v praxi již samostatně.

Genericky pak bude možné podle těchto dvou vzorů / schémat navrhovány další SOP pro prakticky libovolné další druhy (skupiny) potravin, případně i krmiv a nutraceutik. Pro navrhování dalších specificky zaměřených SOP (podle výše uvedeného procesního diagramu) byla experimentálně dokumentována platnost vstupní hypotézy, tj. v úvodu navrženého schématu pro získávání a zpracování dat s interpretační návazností na referenční soubory autentických vzorků:



3.5 Schéma metrologické návaznosti pro metody metabolického vyšetření

Úvod

Určování autenticity vzorků můžeme zařadit do kvalitativní analýzy, která využívá kvantitativních měřicích metod. Měřené signály určitým „filtrem“ na kvalitativní veličiny, které jsou základem rozmanitých klasifikací.

Kvalitativní analýza je úzce spojena s řešením čtyř hlavních otázek:

- přítomnost či nepřítomnost určité substance (analytu); tento problém je obvykle řešen jednoduchou měřicí metodou, při které je signál (odezva) transformována na binární veličinu – pozitivní či negativní hodnota
- testovaný vzorek srovnáváme pomocí souboru naměřených údajů se „vzory“, u kterých je známo zařazení do jednotlivých kvalitativních tříd. Tento typ kvalitativní analýzy je znám pod pojmem „analýza obrazu“ (pattern recognition).
- tzv. limitní test (threshold test), při kterém je hodnota měřené veličiny srovnávána s předem určenou hraniční, limitní hodnotou (threshold value)
- přiřazení molekulové struktury či určité substance testovanému analytu

Výše naznačené problémy kvalitativní analýzy nám dovolují formulovat obecnou definici kvalitativní analýzy: Kvalitativní analýza se zabývá klasifikací objektů podle specifických kritérií za účelem splnění dohodnutých požadavků.

Specifická kritéria jsou velice rozmanitá, např. shoda barvy objektu s testovací škálou, shoda spektrálního profilu píku s posicí a intenzitou, expertní názor, překročení kvantitativního limitu, atd. S klasifikací v kvalitativní analýze jsou rovněž spojeny různé statistické techniky, které závisí na charakteru dat či údajů. Data mohou mít charakter binární (ano/ne), mohou to být kategorizovaná data (červený, zelený, modrý; nemocný, zdravý) nebo mohou být vyjádřeny pořadově (0,1,..10). Statistické testování kvalitativních metod je neobvyklé a obtížné z toho důvodu, že se nejedná o normální rozdělení a musíme tedy používat jiných, vhodných testů. Důležitý rozdíl mezi kvantitativní a kvalitativní analýzou je i tzv. výkonových charakteristikách (performance characteristics). Kvantitativní metody analýzy jsou charakterizovány takovými parametry, jako pravdivost, správnost, přesnost, robustnost, analytická selektivita, analytická citlivost, mez detekce (LOD), linearita, pracovní rozsah. Kvalitativní analýza používá úplně jiné výkonové charakteristiky. Nejdůležitější charakteristikou kvalitativní analýzy je **indikace falešně pozitivních a negativních odezev**. Této charakteristice neodpovídá žádný z parametrů, charakterizujících kvantitativní analýzu. Důležitým aspektem kvalitativních metod je skutečnost, že vždy měříme analytický signál (kvantitativní veličina) a konstruujeme „filtr“ pro převod kvantitativních údajů na kvalitativní, vyjádřené různými způsoby.

Nejistota v kvalitativní analýze je charakterizována jiným způsobem, než je tomu v kvantitativní analýze.

Následující kontingenční tabulka a obrázek znázorňují všechny situace, které mohou v kvalitativní analýze nastat.

KONTINGENČNÍ TABULKA PRO KVALITATIVNÍ MĚŘICÍ METODU

<u>Reálná situace:</u>	<i>analyt přítomen</i>	<i>analyt nepřítomen</i>	<i>total</i>
Důkaz:			
<i>pozitivní</i>	TP	FP	n_{pos,měř}
<i>negativní</i>	FN	TN	n_{neg,měř}
<i>total</i>	n_{přít}	n_{nepřít}	

Citlivost (sensitivity) v kvalitativní analýze, SEN, je pravděpodobnost, že vzorek, který obsahuje analyt, je testovací metodou identifikován jako pozitivní

$$SEN = \frac{TP}{TP + FN}$$

Specifická (specificity) v kvalitativní analýze, SP, je pravděpodobnost, že vzorek, který neobsahuje analyt, je testovací metodou identifikován jako negativní

$$SP = \frac{TN}{TN + FP}$$

Falešně pozitivní poměr (false positive rate) je definován pomocí pravděpodobností vztahem:

$$FPR = \frac{FP}{TN + FP}$$

Je to vlastně poměr nekorektních výsledků, získaných na vzorcích, které neobsahují daný analyt (viz předchozí tabulka).

V literatuře najdeme i jinou definici falešně pozitivního poměru, definovaného na základě počtu falešně pozitivních výsledků experimentu:

$$\% FP = \frac{\text{počet}_{\text{ falešně}_{\text{ pozitivních}_{\text{ výsledků}}}}}{\text{počet}_{\text{ všech}_{\text{ výsledků}}}}$$

Spolehlivost (v%) se dá vyjádřit jako

$$\text{Spolehlivost (\%)} = 100 - FP (\%) - FN (\%)$$

Návaznost je základní vlastnost výsledků z metrologického hlediska. Metrologická definice se váže k výsledkům získaných kvantitativní analýzou, dá se však aplikovat i do oblasti kvalitativní analýzy s použitím referenčních materiálů a kalibrací instrumentů, poskytující kvantitativní signál, rovněž tak s použitím množiny tzv. trénovačích vzorků, u kterých známe správnou klasifikaci. Standardy (referenční materiály, množina trénovačích vzorků) charakterizuje vhodnou instrumentální technikou (např. spojením chromatografických technik s hmotnostní spektrometrií). Naměřený soubor signálů různého typu potom slouží k analýze neznámých vzorků. Většinou se to děje automaticky s použitím sofistikovaných statistických metod. Pyramida návaznosti se potom skládá z analytických metod, vhodných příznaků množiny standardů, chemometrických metod a klasifikace neznámého vzorku. Nejistota klasifikace je dána příslušnou kontingenční tabulkou (čtyřpolní pro binární rozhodnutí, vícepolní pro klasifikaci do více skupin (tříd).

Přehled statistických metod pro klasifikaci

Organizace datové matice pro N vzorků, D tříd a k příznaků

Vzorky	Třída	Příznaky			
1	A	X11	X12	X1k
2	A	X21	X22	X2k
3	B	X31	X32	X3k
.....
N	D	XN1	XN2	XNk

Příznaky mohou být kvantitativní i kvalitativní (např. v senzorické analýze). Kvantitativní data mohou být rovněž diskretizována (kategorizována) na kvalitativní. Např. relativní intenzita píku může být diskretizována na různé intervaly (0-0,2;0,2-0,4; atd). Pokud máme rozsáhlý soubor dat, kategorizaci provádíme automaticky pomocí software..

Významnost měřených signálů

Moderní analytické přístroje poskytují velký počet signálů pro jeden vzorek. Pro věrohodnost správné klasifikace je třeba selekce významných signálů na množině standardních (trénovačích) vzorků. Pokud máme větší počet vzorů pro jednotlivé třídy, je třeba vybrat takový soubor signálů, který s určitou pravděpodobností je spolehlivě popisuje. Většinou volíme pravděpodobnosti 0,90 nebo 0,95, což vede k tzv. **cut-off** hodnotě $p=0,1$ nebo $p=0,05$.

Pro testování významnosti používáme neparametrických metod testování, neboť v těchto případech klasifikace nezáleží na typu rozdělení pravděpodobností. Dále se nebudeme zabývat podrobnými statistickými postupy.

Mann Whitneův (dále MW) test používáme pro testování v případech, kdy máme rozlišit dvě třídy (např.vzorek je ze dvou pěstebních oblastí nebo je falešný či autentický).

MW test posuzuje rozdíl mezi dvojicemi měřených signálů pro dvě třídy. Nulová hypotéza, H_0 , je založena na statistickém posouzení průměrných diferencí pro jednotlivé

třídy. H_0 je splněna, pokud rozdíl je statisticky nevýznamný. Výsledek MW může být např. ve tvaru tabulky (zde pro tři parametry)

Proměnná	p*
A	0.0003
B	0.199
C	0.049

*p je risk odmítnutí nulové hypotézy

Interpretace této tabulky je taková, že parametry A a C jsou významně odlišné, parametr B vykazuje shodnost a tedy nerozlišení dvou tříd.

Pro případ, že máme rozlišit více tříd (vzorů), používáme **Kruskal-Wallisův** test.

Vhodným testem, který je rovnocenný předchozím testům, je **Moodův** test. Moodův test využívá pro testování rozdílů mezi třídami medián jednotlivých parametrů (signálů) pro jednotlivé třídy. Výhodou mediánového testu je nezávislost na vychýlených hodnotách.

Je třeba si uvědomit, že testy významnosti významně snižují počet parametrů a eliminují ty, které nemají na rozlišení žádný vliv. Je to důležité i z hlediska věrohodnosti získaných výsledků klasifikace, neboť důležitou podmínkou je co nejvyšší poměr počtu signálů (parametrů) a počtu vzorků v trénovací množině.

Klasifikační metody

Klastrová analýza

V případech, kdy neznáme počet tříd, provedeme předběžné zkoumání datového souboru. K tomu volíme dvě metody: metodu nejbližšího souseda (k-nearest neighbour) a aglomerativní shlukování..

Metoda nejbližšího souseda je iterativní metoda, ve které nulté přiblížení klastrů (tříd, skupin) se rovná počtu objektů. Následuje redefinice klastrů pro jednotlivé objekty podle zvoleného klasifikačního kritéria. Nejběžnější je tzv. Wilks lambda kritérium, což je poměr vnitroklastrové kovarianční matice k celkové kovarianční matici. Testovací kritérium je F-kritérium.

Agglomerativní shlukování je založeno na shlukování objektů podle euklidovské nebo Mahalanobisovy vzdálenosti. Kritérium shlukování je např. chi-kvadrát test. Výsledkem je grafické zobrazení dendogramem.

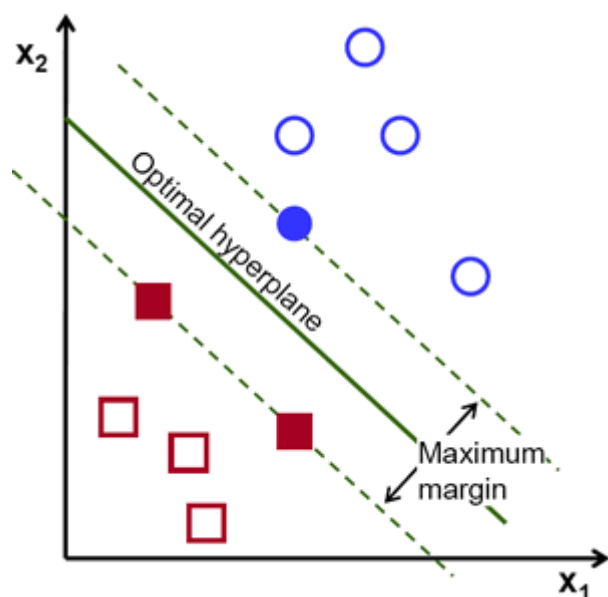
Metoda nejbližšího souseda a aglomerativní shlukování jsou komplementární. Při k-nearest neighbour metodě se při iteracích může měnit zařazení objektu do různých klastrů, což není možné při aglomerativním shlukování. Toto předběžné zkoumání je nutné v těch případech, kdy není jasné apriorní rozdělení do tříd. To je např. v případě binárních rozhodování (autentický a falešný vzorek).

Pro binární data (dvě třídy), což je v našem případě nejběžnější, využíváme několik statistických metod. Budeme se zabývat těmi metodami, které jsou schopné poskytnout kontingenční tabulku zahrnující metrologická kritéria pro nejistotu v kvalitativní analýze (selektivita a sensitivita).

Support vector machine (SVM)

Metoda je zřejmá z přiloženého (převzatého) obrázku pro dva parametry (signály). Iterativním způsobem se „konstruuje“ optimální hyperploha, charakterizovaná souborem

vektorů ve vícerozměrném prostoru. Neznámý vzorek je potom automaticky přiřazen na danou stranu hyperplochy, tedy do jedné ze dvou tříd.



Jako příklad uvádíme použití této metody na vzorky italských a španělských olejů.

Konfuzní matice (kontingenční tabulka) trénovacího souboru 239 vzorků.

from \ to	ITA	SPA	Total	% correct
ITA	154	0	154	100.00%
SPA	0	85	85	100.00%
Total	154	85	239	100.00%

Počet příznaků (signálů): 253

Počet vektorů: 416

Logistická regrese

Jedné skupině přiřadíme kódové označení **0**, druhé skupině **1**. Výsledkem je regresní rovnice pro pravděpodobnost, že vzorek se zařadí do skupiny 0 nebo 1. Dalším výsledkem je kontingenční tabulka. Logistická regrese je použitelná i v případě více skupin. Nevýhodou je nemožnost použití metody v případě velkého množství příznaků a malého počtu vzorků.

Lineární diskriminační analýza (LDA)

Slouží k *diskriminaci* (rozlišení) objektů pocházejících z konečného počtu tříd (kategorií) na základě objektů z jisté podmnožiny všech objektů, tzv. *trénovací množiny*, sestavením *rozhodovacího pravidla* a k následné *klasifikaci*.

Jako příklad uvádíme výsledek LDA pro stejný soubor olejů

Kontingenční tabulka

from \ to	ITA	SPA	Total	% correct
ITA	154	0	154	100.00%
SPA	0	85	85	100.00%
Total	154	85	239	100.00%

Doprovodným výsledkem může být i simulovaná závislost citlivost na specifičnosti, tzv. ROC křivka.

Možným použitím je i kombinace **LDA s metodou PLS** (Partial Least Square), známá v literatuře jako SIMCA. Metoda dává podobné výsledky jako LDA. Obě metody rovněž poskytují přehled o parametrech, které významně ovlivňují rozdělení. Na základě statistického testování můžeme rovněž podstatně redukovat počet příznaků.

Pro zpracování datové matice o více třídách využíváme metody Logistické regrese a LDA, popř. SIMCA. Výsledkem Logistické regrese a LDA je rovněž kontingenční tabulka.

Kvalita primárních dat

Jakékoli měření, tedy i rozsáhlý soubor primárních dat, musí vykazovat znaky zajištění kvality. Budeme se zabývat měřicími metodami, které jsou používány v metabolomice. Kromě NMR jsou to převážně metody vysokoúčinné chromatografie spojené s hmotnostní spektrometrií. Nedílnou součástí zajištění kvality dat je vysoká kvalifikace personálu provádějícího měření a rozhodování. Ke každé matici se musí přistupovat individuálně, jak z hlediska měření, tak i zpracování a interpretaci. Měřicí metody musí být validovány, rovněž přístrojový software, určený k předzpracování a úpravě naměřených dat, by měl být validován. Je to ostatně i požadavek ISO 17025.

Matrice řešené v tomto projektu

Jedná se o čokolády, rostlinné oleje a koření. Vzorky jsou obecně charakterizovány a zvoleny způsoby separací a měření.

Čokoláda

Problémem je neshoda s deklarovaným složením. Bylo k dispozici 100 vzorků různých čokolád. Vzhledem ke komplexnímu složení čokolád byla jako měřicí metoda zvolena technika ultra účinné superkritické fluidní chromatografie ve spojení s vysokorozlišovací tandemovou hmotnostní spektrometrií (UHPSFC-HRMS/MS).

Olivový olej

Analyzovaný soubor obsahoval 154 olivových olejů z Itálie a 58 vzorků ze Španělska. Jako měřicí metoda byla zvolena vysokoúčinná kapalinová chromatografie ve spojení s vysokorozlišovací tandemovou hmotnostní spektrometrií (U-HPLC-HRMS/MS).

Černý pepř

Analyzovaný soubor obsahoval 15 autentických vzorků černého pepře a několik vzorků zástupných surovin, popř. podezřelých vzorků z maloobchodu a velkoobchodu. Jako měřicí metoda byla zvolena technika mikroextrakce na tuhou fázi ve variantě head space (HS

SPME) ve spojení s plynovou chromatografií a hmotnostně spektrometrickou detekcí (HS-SPME-GC-HR-TOF-MS).

Zobecněná optimalizace měřicí metody a prvotní zpracování primárních dat

Nedílnou součástí metrologické paramidy je používání optimalizovaných metod měření. V tom je podle mého názoru hlavní překážka přenosu metodiky na jiná pracoviště a vytvoření obecné databáze standardních vzorků popř. datové matice trénovacího souboru. Optimalizace měřicího postupu je časově i finančně nákladná záležitost a je obchodním tajemstvím organizace. Nedílnou součástí, jak již bylo uvedeno, je vysoce kvalifikovaný personál, jehož výchova je také velmi drahá a časově náročná.

Optimalizace měřicího postupu

Optimalizaci můžeme rozložit do několika kroků:

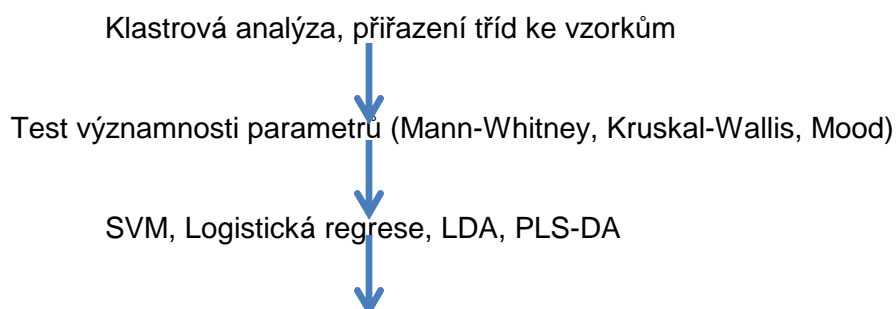
1. Neexperimentální úvaha o matici a možných sloučenin, které mohou být předmětem měření
2. Výběr vhodných separačních technik – rovněž neexperimentální úvaha
3. Počáteční experimenty na známé, autentické matici a jejich optimalizace
4. Úvodní měření na vybrané technice a optimalizace instrumentálních parametrů
5. Případná identifikace hlavních sloučenin, popř. skupin chemických látek a jejich výběr pro další zpracování

Získání primárních dat trénovacího souboru

1. Výběr časově stabilních vzorků – kontrola krátkodobé stability
2. Měření a získání datové matice, procesně upravené (např. pomocí speciálních software)
3. Výsledek: rozsáhlá matice intenzit signálů pro daný poměr m/z a retenční čas (m/z@RT). Matice může být normalizována k nejvyšší intenzitě, popř. tzv. zarovnání retenčních časů (drobné posuny vrcholů píků)

Statistické zpracování dat trénovacího souboru

Neznáme rozdělení do tříd nebo dělení není jednoznačné



- Výsledek:** Konfuzní matice /kontingenční tabulka
 Citlivost, specifita, správnost
 Regresní rovnice v logistické regresi a pravděpodobnost zařazení do tříd
 Rozhodovací pravidlo (klasifikační rovnice) LDA nebo PLS-DA

Známe rozdělení do tříd
 Odpadá první krok

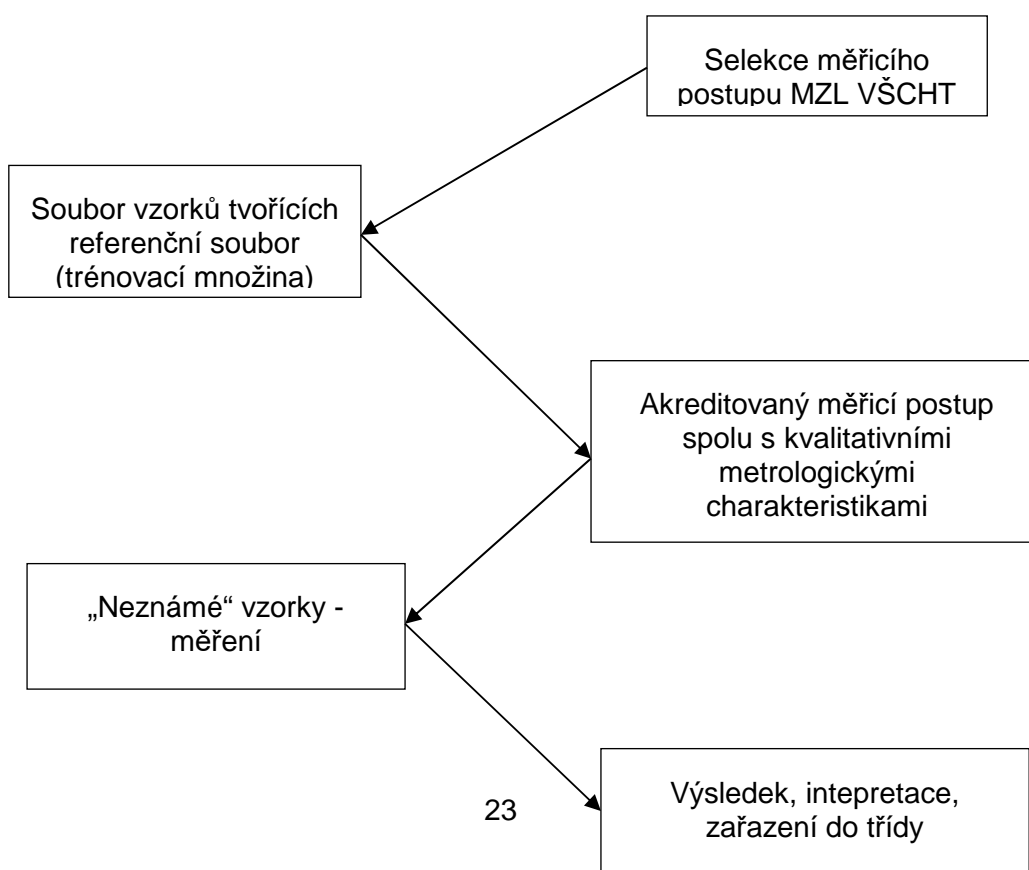
Schéma návaznosti metabolických měření

Mezinárodní metrologická organizace ani VIM prozatím nedefinovaly návaznost kvalitativních měření. Náš návrh je tedy prvním pokusem to učinit. Návrh je v částečném souladu s ISO 17511. Tato norma pojednává o návaznosti v medicíně. Je navrženo několik schémat podle typu matrice, metody měření a analytu. Jedno ze schémat se nazývá

„Calibration hierarchy and metrological traceability to the manufacturer’s selected measurement procedure that is not primary“

Podobně bylo vytvořeno schéma návaznosti v tomto projektu. Je třeba zdůraznit, že se vždy jedná o návaznost k trénovací množině vzorků, na jejíž kvalitě závisí i kvalita schématu návaznosti. Rovněž, jak bylo již zdůrazněno, je třeba striktně dodržovat SOP dané metody a personální kvalifikaci účastníků měření a interpretace. MZL VŠCHT je akreditována Českým institutem pro akreditaci a splňuje tedy všechny náročné podmínky.

Schéma návaznosti může být vyjádřeno jednoduchým schématem:



Metrologické postupy v metabolomickém vyšetřování potravin a krmiv

Ve schématu není prozatím krok, kterým by bylo vyjádřeno měření jiným účastníkem, než je MZL VŠCHT. To může být předmětem další diskuse s ČMI.

4. ZÁVĚRY

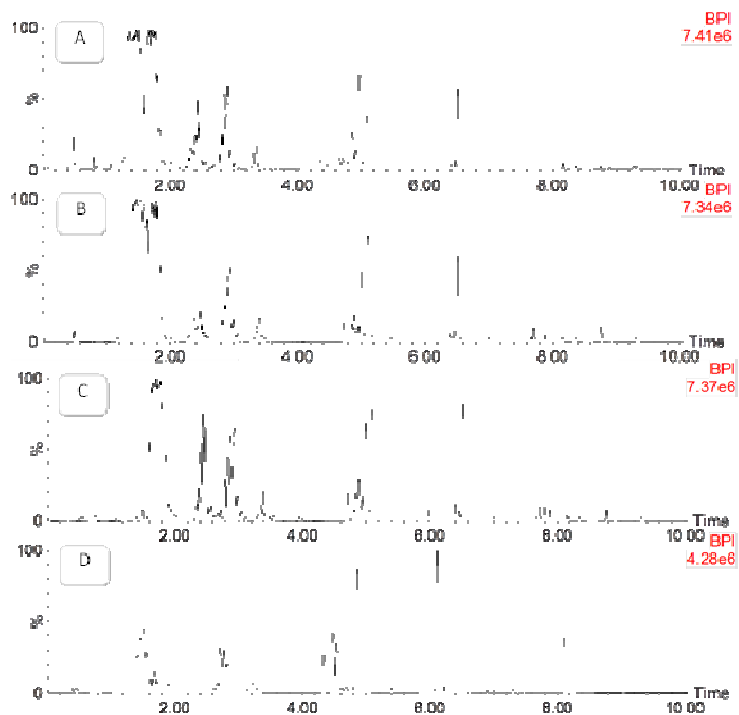
1. Pro všechny tři typy testovaných matric byla provedena optimalizace izolačního resp. extrakčního postupu. Byla navržena a ověřena analýza s využitím vysoce účinných separačních metod a s detekcí HR-MS. Kromě separace chromatografické byla realizována také separace pomocí spektrální dekonvoluce, což je výhodné v případě koalujících se sloučenin. Ve všech případech se podařilo matici velmi komplexně charakterizovat a to jak z pohledu zastoupených sloučenin, tak i relativní intenzity jejich signálů. Vhodné vnitřní standardy nebyly dosud navrženy.
2. Primární data byla zpracována multidimenzionální analýzou. Metodou PCA bylo dosaženo zjednodušení popisu skupiny korelovaných znaků. Klasifikace je realizována diskriminační analýzou (DA) klasifikačního modelu získaného pomocí metody částečných nejmenších čtverců (Partial Least Square, PLS) či s využitím její varianty ortogonální-PLS (OPLS), jež zahrnují mnohonásobné lineárně regresní postupy založené na PCA. Pro vybrané komodity byly nalezeny vhodné markery autenticity.
3. Podle navržených schémat byly do standardizované formy zpracovány a validovány dvě optimalizované pilotní analytické metody s názvem „Stanovení látkového profilu / metabolomu [potravina] metodou hmotnostní spektrometrie“. Postupy (SOP) budou zařazeny do rozsahu akreditace Metrologické a zkušební laboratoře VŠCHT Praha (audit ČIA stanoven na únor 2018) a dále využity pro navrhování dalších generický SOP.
4. Ve spolupráci s EURACHEM ČR byla potvrzena validita získaných primárních dat a postupů jejich zpracování, včetně interpretace. Bylo navrženo obecné schéma metrologické návaznosti pro metody metabolomického vyšetření potravin a krmiv.
5. Řešení proběhlo podle časového a věcného plánu.

Hlavní řešitelkou projektu je prof. Ing. Jana Hajšlová, CSc., která je vedoucí Metrologické a zkušební laboratoře VŠCHT Praha, akreditované Českým institutem pro akreditaci, o.p.s. podle ČSN EN ISO/IEC 17025:2005. Na řešení projektu a dílčí zprávě se podíleli zejména prof. Ing. Vladimír Kocourek, CSc., Ing. Zbyněk Džuman, Ing. Jaromír Hradecký, Ing. Eliška Kludská a Ing. Michaela Rektorisová.

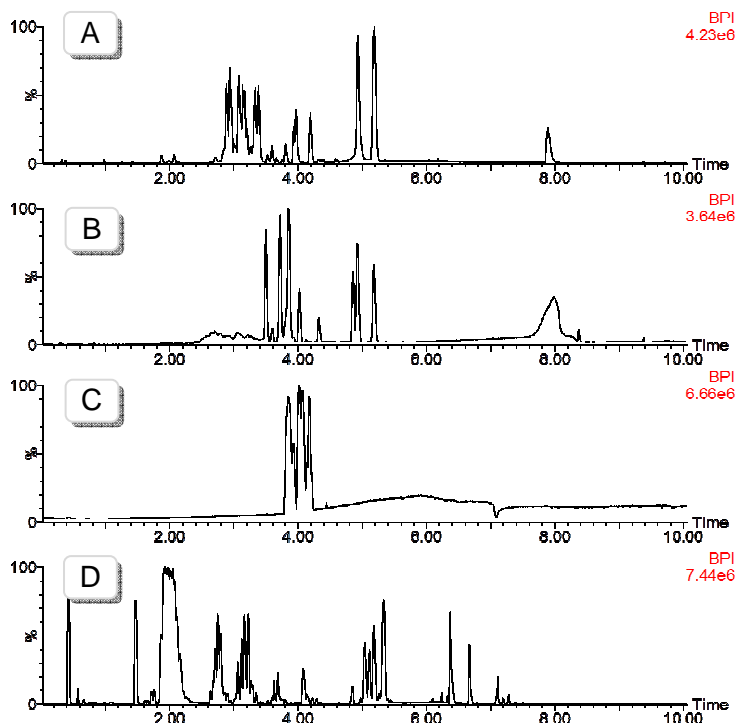
5. PŘÍLOHY

5.1. Přílohy- metabolomika čokolády

Příloha 6.1.1.: Chromatogramy různých typů extraktů vzorku čokolády získané separací na koloně BEH s využitím pozitivního módu ionizace. A – isopropanol; B – isopropanol/methanol 2:1 (v/v); C – methanol/ethylacetát 1:2 (v/v); D – methanol/voda 19:1 (v/v).



Příloha 5.1.2.: Příklady chromatogramů získaných z analýzy vzorku čokolády s využitím pozitivního módu ionizace. Separace na kolonách: A - HSS C18 SB; B – Torus 2-PIC; C – Torus 1-AA; D - BE



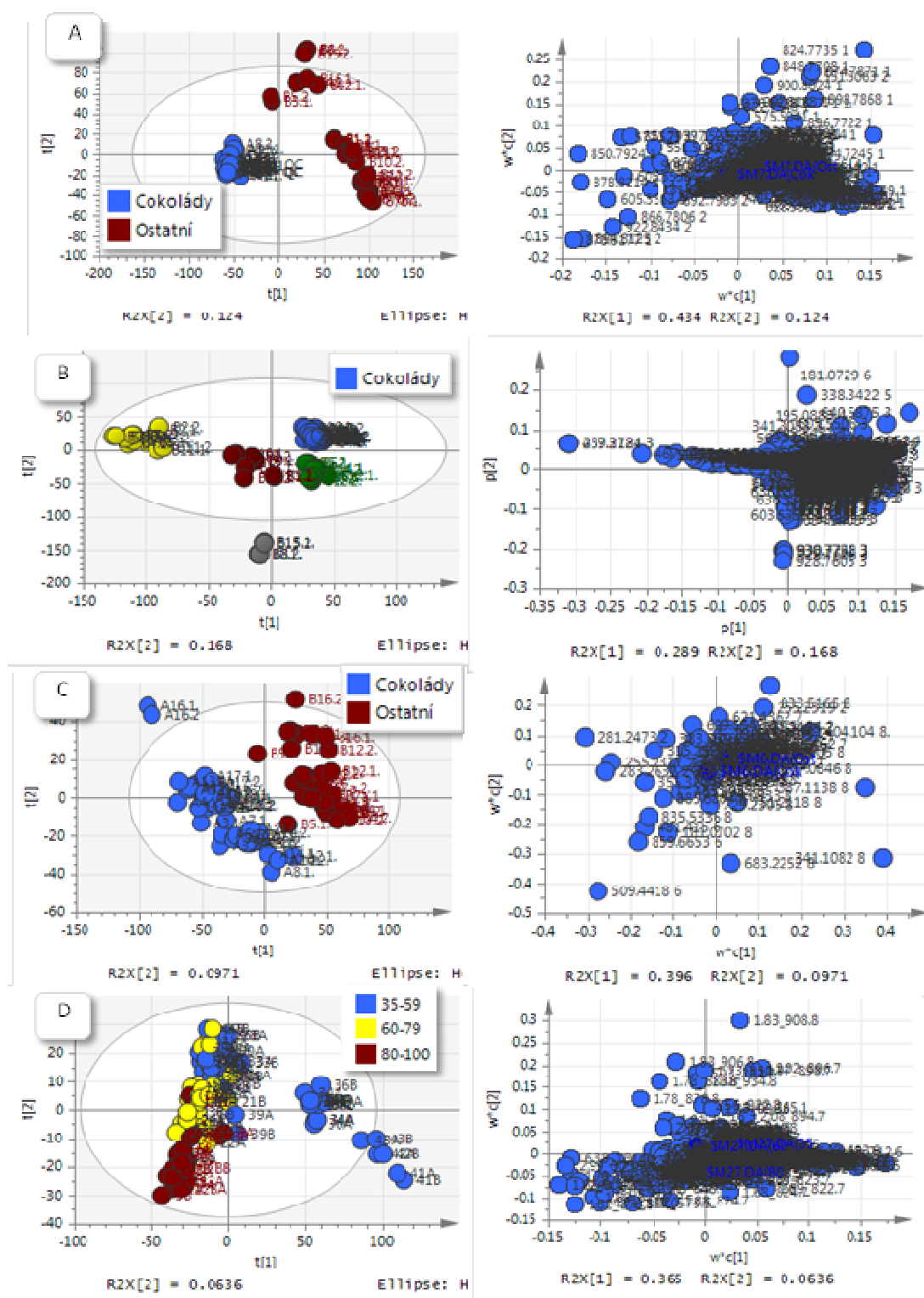
Příloha 5.1.3.: Tabulka optimalizovaných podmínek měření UHPSFC-HRMS/MS

UPC ²	
Složení mobilní fáze	A: CO ₂ B: MeOH/MeCN (80/20, v/v), 5 % H ₂ O, 15mM HCOONH ₄
Průtok mobilní fáze	1,8 ml/min
Kolona	Acquity UPC ² BEH (3,0 x 100 mm, 1,7 μm)
Teplota kolony	60 °C
Objem nástřiku	1 μl
Zpětný tlak (ABPR)	2000 psi (13,8 MPa)
Make-up	MeOH s 0,1% HCOOH; 0,25 ml/min
ESI-QTOFMS	
Napětí sprejovací kapiláry	1 kV / -0,6 kV
Napětí na kóně	35 V
Napětí na extrakční kóně	4 V
Teplota zdroje	120 °C
Desolvatační teplota	350 °C
Průtok desolvatačního plynu	800 l/h
Průtok kónového plynu	30 l/h
Akviziční rychlost	10 spekter/s
Hmotnostní rozsah	50 – 1200 Da

Rozlišení	25000 FWHM
-----------	------------

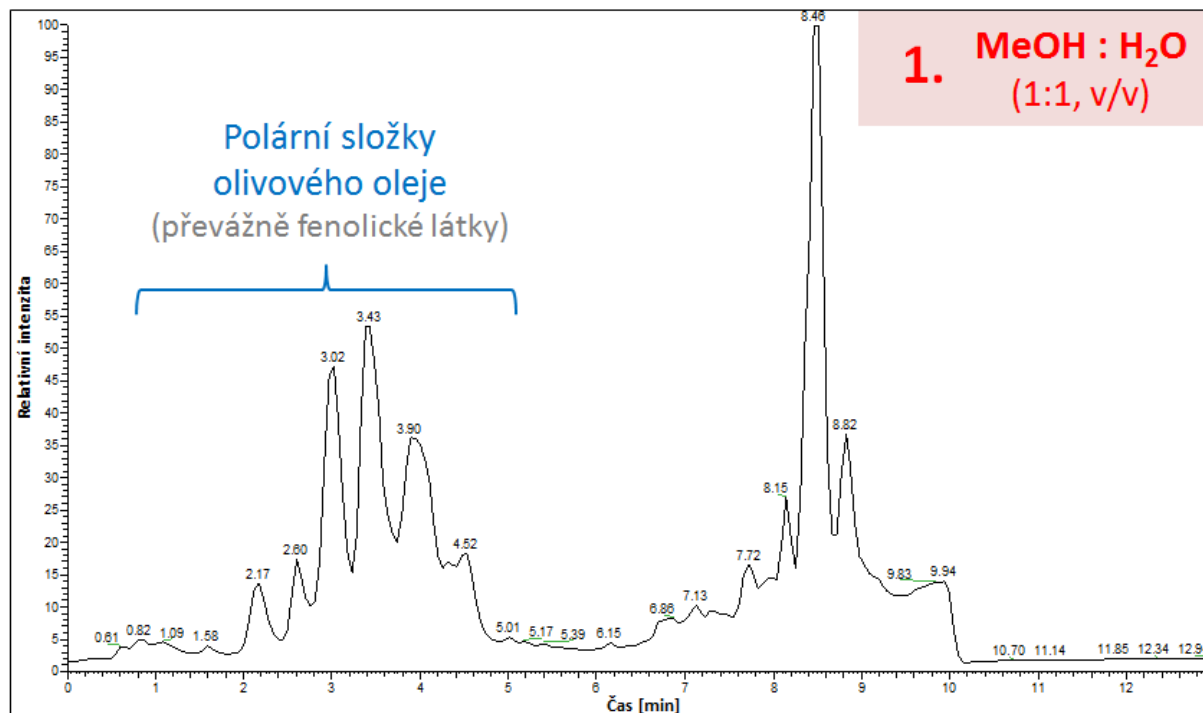
Příloha 5.1.4.: Diagramy komponentního skóre znázorňující rozdělení jednotlivých vzorků a diagramy komponentních vah zobrazující příslušné markery zodpovědné za toto rozdělení. A – PLS-DA, data z pozitivní ionizace; B – PCA, data z pozitivní ionizace s retenčním časem vyšším než 3 minuty; C – PLS-DA, data z negativní ionizace; D – PLS-DA, data z pozitivní ionizace, pouze vzorky čokolád.

Metrologické postupy v metabolickém vyšetřování potravin a krmiv

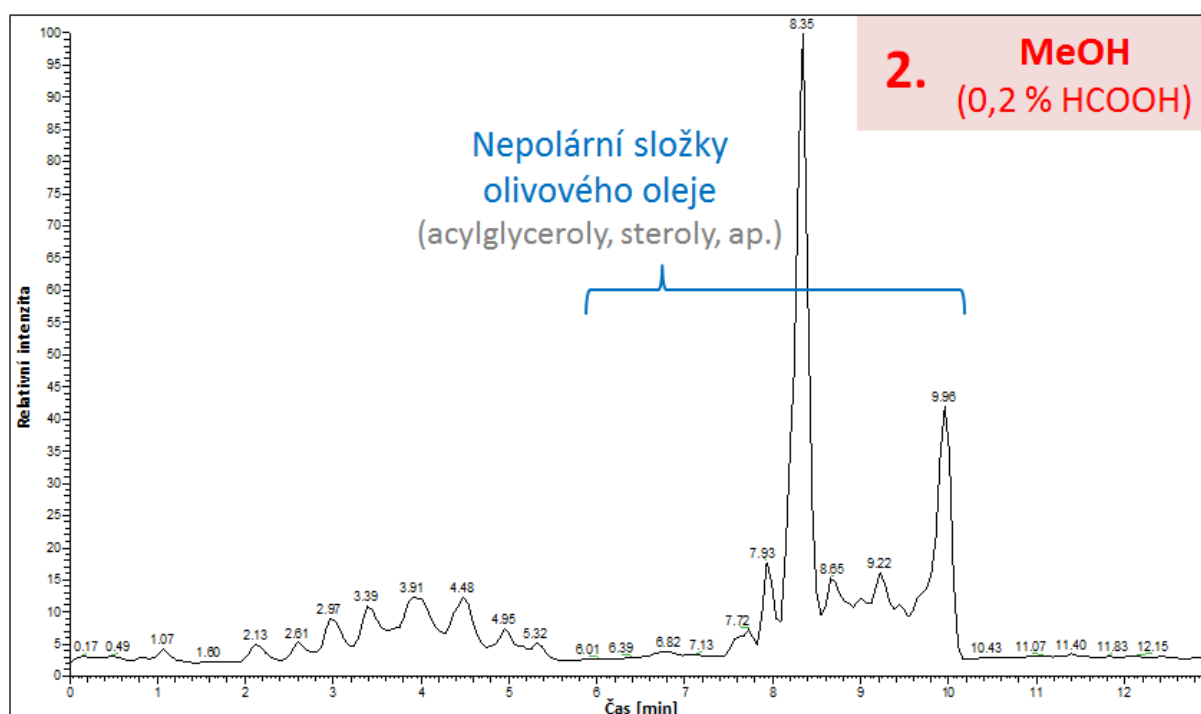


5.2. Přílohy- metabolika olivového oleje

Příloha 6.2.1.: Celkový iontový chromatogram extraktu extra panenského olivového oleje se zaměřením na polární složky vzorku



Příloha 5.2.2.: Celkový iontový chromatogram extraktu extra panenského olivového oleje se zaměřením na polární složky



Příloha 5.2.3. Gradient mobilní fáze pro U-HPLC–HRMS/MS

t [min]	průtok	A [%]	B [%]
0	0,3	80	20
1,5	0,3	50	50
8,0	0,3	0	100
10,0	0,4	0	100
13,0	0,3	80	20

Příloha 5.2.4. Podmínky ionizace (ESI)

Ionizace	ESI+/-
Nosný/pomocný plyn	32/7 arb. j.
Teplota kapiláry	300 °C
Vypařovací teplota	220 °C
Napětí na elektrospreji	+/- 3,3 kV
S-lens hodnota	60

Příloha 5.2.5. Podmínky akvizičního módu *fullMS*

Rozlišení	70 000 FWHM
Akviziční rychlost	1,5 Hz
Hmotnostní rozsah	<i>m/z</i> 100 -
Automatická regulace citlivosti (AGC)	3e ⁶
Maximální doba nástřiku (max IT)	100 ms

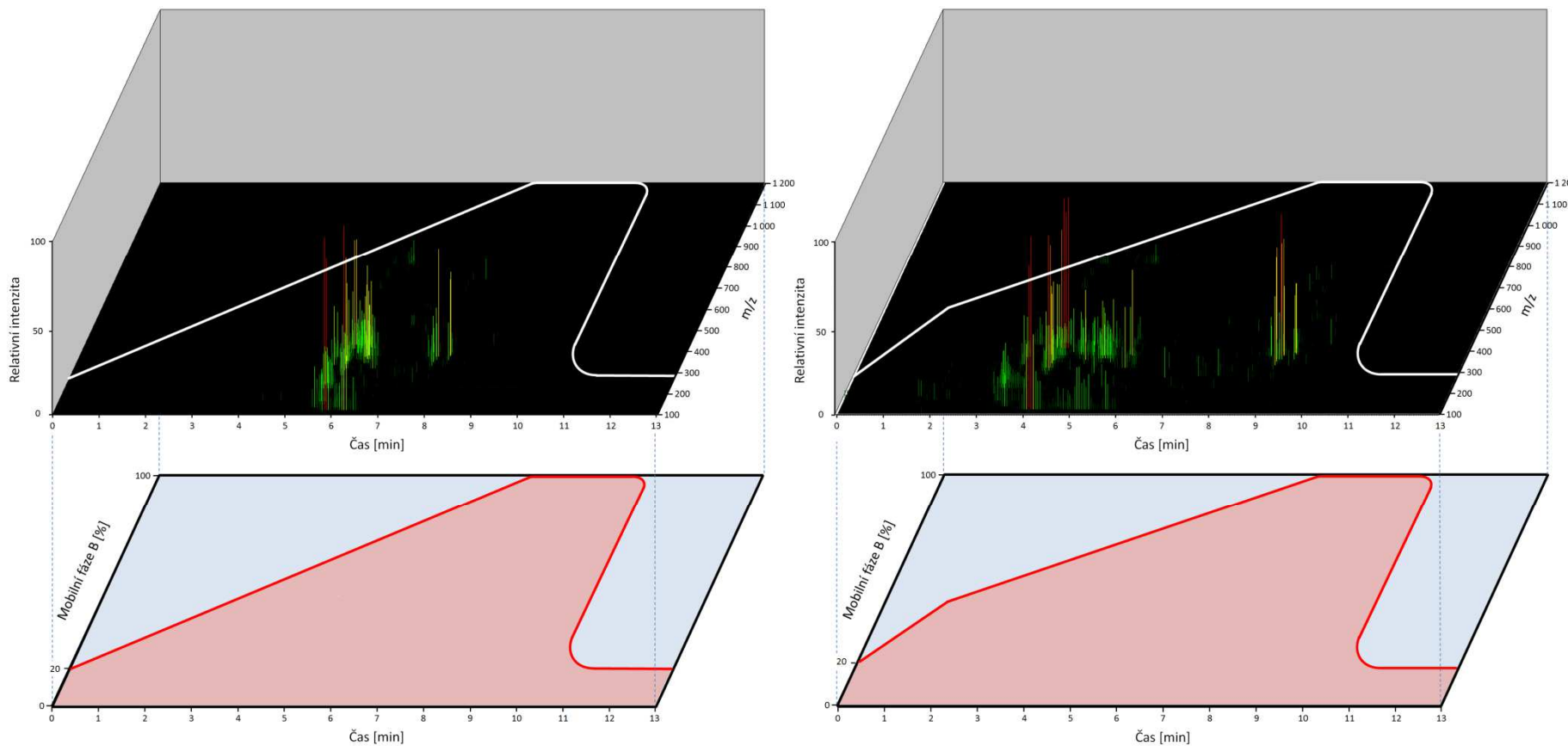
Příloha 5.2.6. Podmínky akvizičního módu *ddMS2 (data dependent MS2)*

Rozlišení	17 500 FWHM
Akviziční rychlost	12 Hz
Hmotnostní rozsah	<i>m/z</i> 50 - <i>m/z</i> fragmentovaného analytu

Metrologické postupy v metabolomickém vyšetřování potravin a krmiv

Šířka izolačního okna	1 m/z
Automatická regulace citlivosti (AGC)	$1e^5$
Maximální doba nástřiku (max IT)	50 ms
Normalizovaná kolizní energie (NCE)	25 %, 30 %, 35 %
Dynamické vyloučení analytu	5 s
Mez intenzity pro fragmentaci	$2e^4$

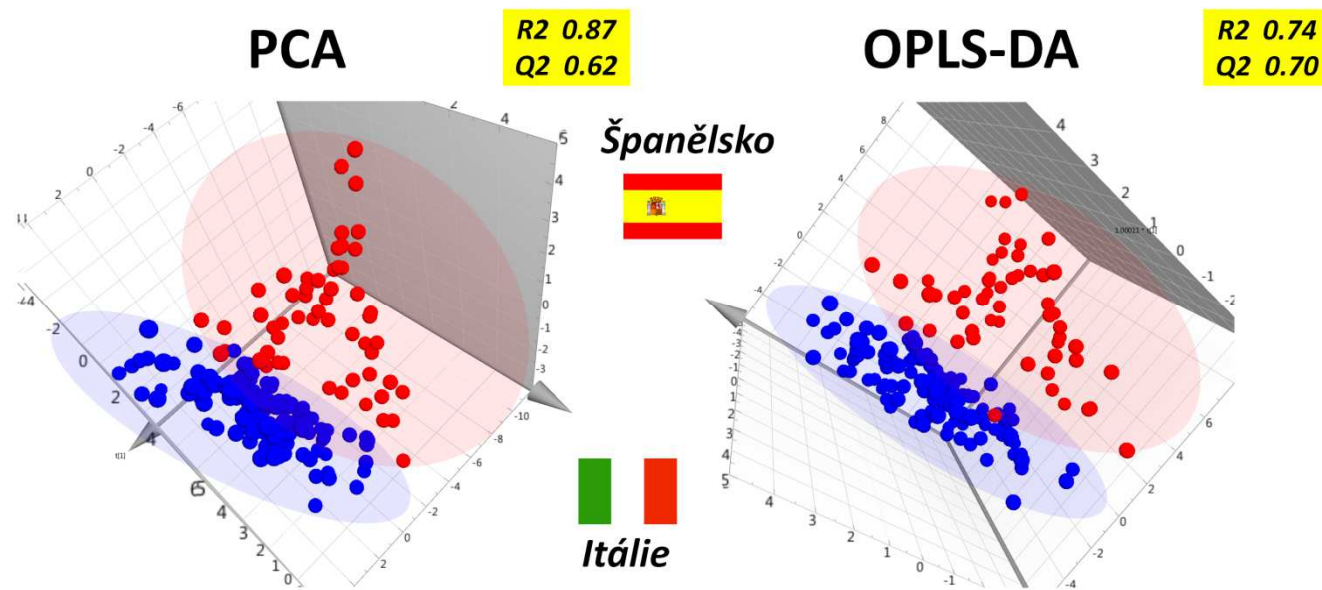
Příloha 5.2.7. Optimalizace chromatografického gradientu demonstrováno jako celkový iontový tok (TIC) ve formě 3D mapy (13 min, m/z 100 - 1200); před (vlevo) a po (vpravo) optimalizaci



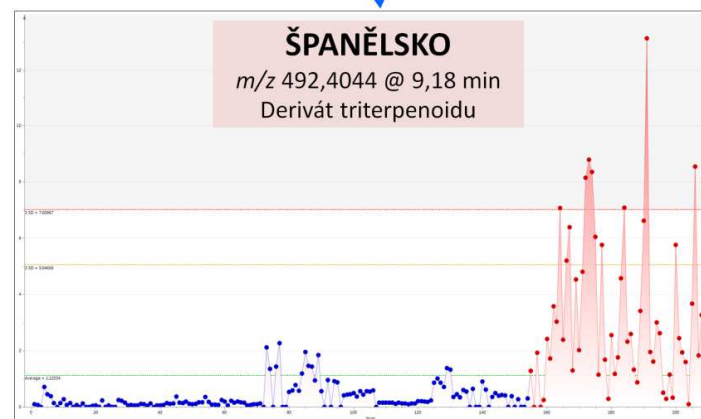
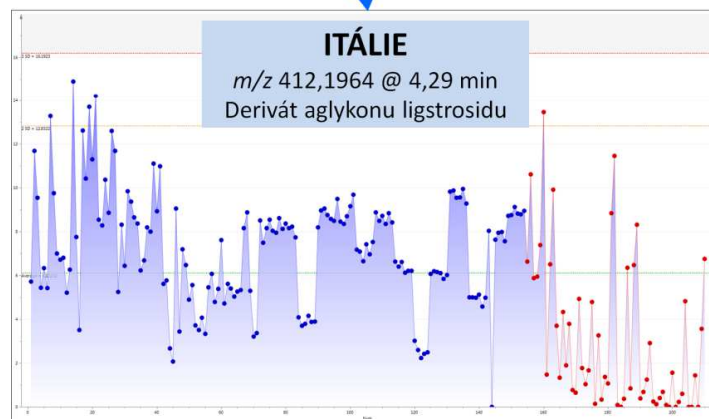
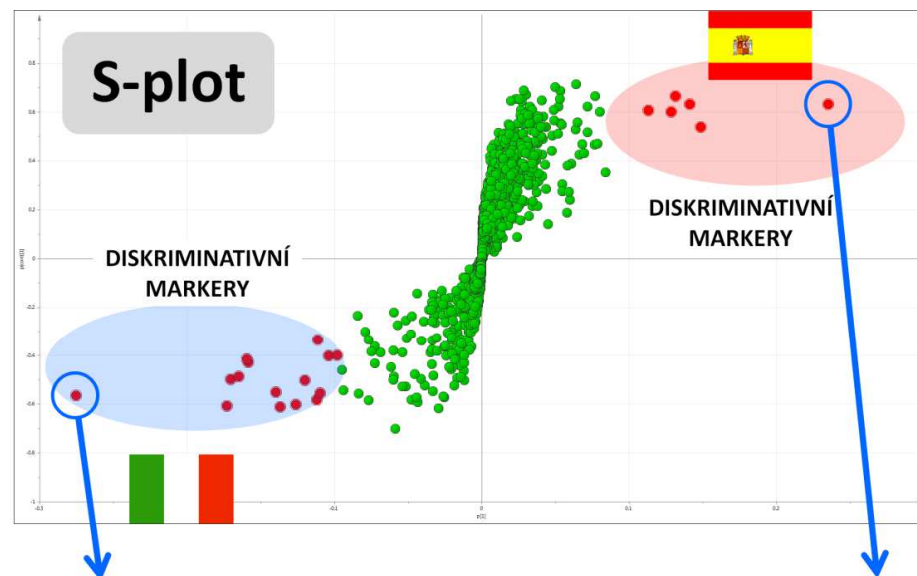
Příloha 5.2.8. Grafický výstup multivariační analýzy (PCA, OPLS-DA) pro autentikaci 154 italských a 58 španělských olivových olejů dle geografického původu

Geografický původ

- Itálie (n = 154)
- Španělsko (n = 58)



Příloha 5.2.9. Vizualizace identifikace diskriminativních markerů pro klasifikaci EVOO dle geografického původu



5.3. Přílohy- metabolomika černého pepře

Příloha 5.3.1.: Parametry HS- SPME-GC-HR-TOF-MS metody pro získání fingerprintů černého pepře

<u>Příprava vzorku</u>	
Naváženo 100 mg vzorku do 10 ml vialky pro SPME	
<u>HS-SPME nastavení</u>	
SPME vlákno	PDMS/CAR/DVB
Inkubace	5 min / 50 °C
Sorpce	1 min / 50 °C
Desorpce	10 min / 270 °C
<u>GC separace</u>	
Chromatografická kolona	HP-Innowax (30 m x 0.25 mm x 0.25 µm)
Nástřik vzorku split/splitless	270 °C (split poměr 100:1)
Průtok	1 ml/min (constantní průtok)
Teplotní program pece	40 °C 1min, dále gradient 10 °C/min do 170 °C a dále gradient 30 °C/min do 260 (zádrž 3,5min)
Nosný plyn	Helium
<u>MS detekce (TOF)</u>	
Rozsah hmot	10-510 m/z
Akviziční rychlost	10 Hz
Teplota iontového zdroje	250 °C
Ionizační energie (EI mód)	70 eV
Napětí detektoru	2700-2850 V
Kalibrační látka (PFTBA)	kalibrováno před analýzou
<u>Statistical compare - nastavení</u>	
Spektrální similarita	625
S/N	20/15
Odchylka retenčního času	45 s

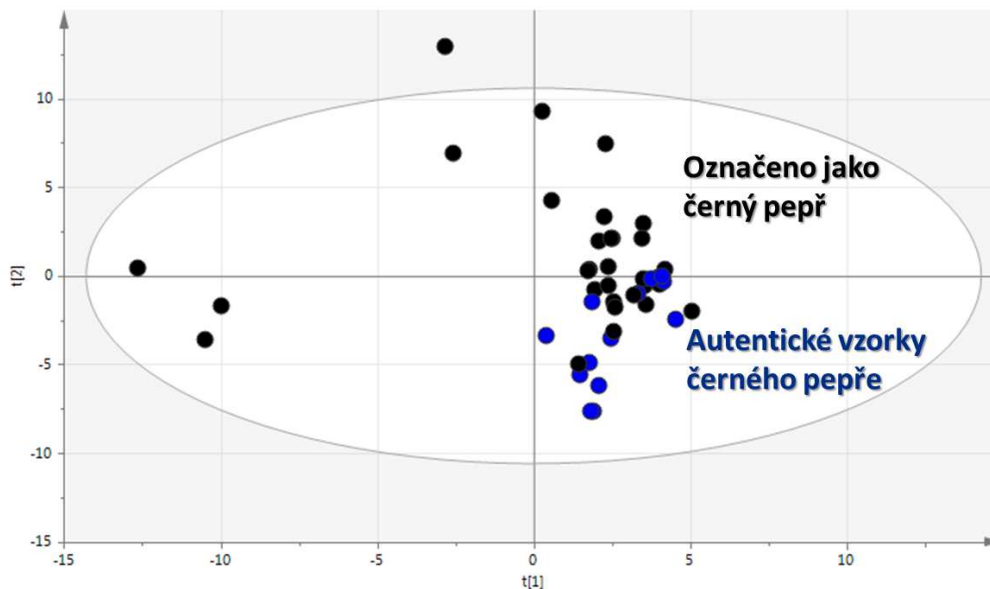
Příloha 5.3.2.: Přehled vzorků černého pepře

Typ vzorku	n	
Autentický	15	
Maloobchodní	18	
<i>Z toho. 'podezřelých'</i>	1	
Velkoobchodní	12	
<i>Z toho. 'podezřelých'</i>	5	
Adulterant	spent	2
	slupky	2
jiné	Směsné koření	3

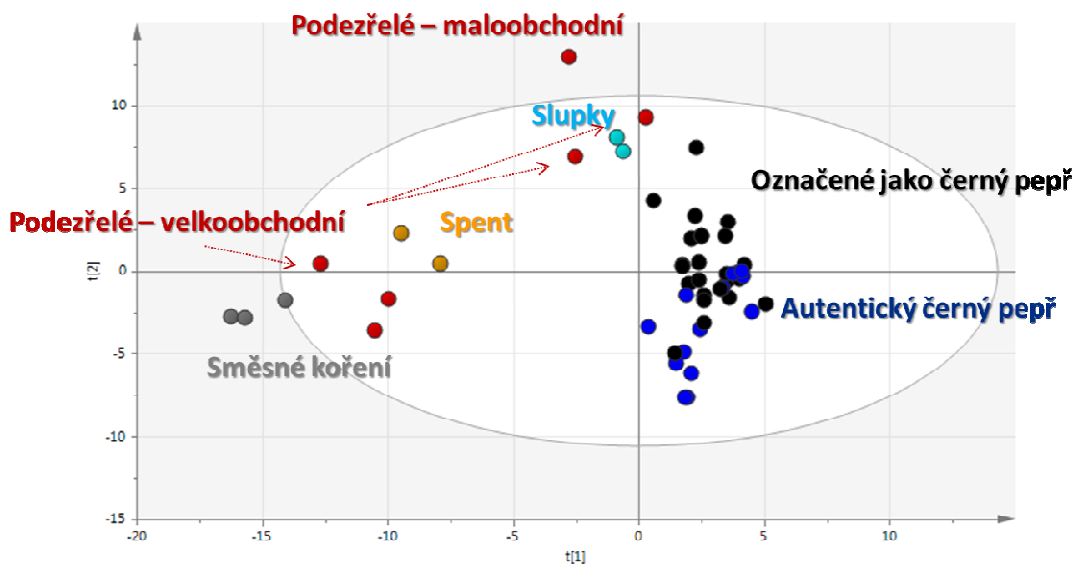
Příloha 5.3.3.: Schéma matice dat

číslo sloučeniny	1	2	3	m	suma relativních ploch (rA); (%)
m/z	m/z_1	m/z_2	m/z_3	m/z_m	-
retenční čas (s)	RT_1	RT_2	RT_3	RT_m	-
vzorek 1	rA_1	rA_2	rA_3	rA_m	100
vzorek 2	rA_1	rA_2	rA_3	rA_m	100
vzorek 3	rA_1	rA_2	rA_3	rA_m	100
.....	100
vzorek n	rA_1	rA_2	rA_3	rA_m	100

Příloha 5.3.5.: Grafické výstupy statistického zpracování dat



Příloha 5.3.6.: PCA rozdělení vzorků autentického pepře a vzorků označených jako černý pepř



Příloha 5.3.7. Výsledky klasifikace vzorků černého pepře:

Typ vzorku		klasifikace			
		autentické	slupky	spent	směsné koření
Maloobchodní	18	15	3	-	
	'podezřelé'	1	1	-	-
Velkoobchodní	12	7	2	2	1
	'podezřelé'	5	2	2	1

Příloha 5.4. Příklad části rozsáhlé LC–HRMS databáze detekovaných sloučených specifikovaných přesnou hmotou m/z a retenčním časem (min) v analyzovaných vzorcích olivových olejů po normalizaci

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
Sample ID	Country	Harvest	Quality grade	Blending	Country/Quality	121.06497@5.231	121.06498@4.159	105.05456@3.42	187.12628@4.449	211.05959@3.384	225.07599@3.402	243.0862@3.378	319.15379@4.15	363.14349@5.233	377.1592@4.313	377.1592@4.306	379.13855@4.741	380.17005@5.242	412.15638@4.289	440.22757@5.289	446.22756@4.304	476.40926@9.253	
1	ITA	2013	EVOO	single	ITA-EVOO	4.128	6.095	1.488	0.261	1.459	2.211	3.255	6.660	4.773	2.667	0.413	2.042	2.156	5.711	0.611	1.303	0.073	
2	ITA	2013-2	ITA	2013	EVOO	3.156	5.026	0.994	0.095	0.908	1.697	1.718	2.489	3.068	3.573	1.945	1.885	1.700	11.681	0.873	1.422	0.030	
3	ITA	2013-3	ITA	2013	EVOO	3.438	5.965	1.313	0.238	1.075	2.071	2.979	6.221	4.732	0.001	1.632	1.966	1.618	9.535	1.344	1.655	0.001	
4	ITA	2013-4	ITA	2013	EVOO	6.591	5.615	1.526	0.523	1.007	2.054	2.509	6.551	2.567	2.511	1.187	1.896	1.954	3.439	1.810	1.218	0.001	
5	ITA	2013-5	ITA	2013	EVOO	4.610	6.167	1.561	0.349	1.029	2.049	2.173	6.645	4.455	0.001	0.001	1.417	0.745	6.338	2.363	1.407	0.215	
6	ITA	2013-6	ITA	2013	EVOO	6.038	7.702	1.345	0.303	1.070	2.050	1.905	4.850	4.489	2.189	0.957	0.001	0.001	2.375	5.417	1.844	1.671	0.230
7	ITA	2013-7	ITA	2013	EVOO	0.001	2.962	1.709	0.527	2.525	3.977	5.910	0.441	1.803	0.001	2.526	2.263	1.796	13.281	2.038	3.773	0.108	
8	ITA	2013-8	ITA	2013	EVOO	3.275	6.662	1.174	0.288	0.934	1.880	1.631	5.107	3.564	2.822	1.614	1.223	0.001	9.763	0.847	1.710	0.077	
9	ITA	2013-9	ITA	2013	EVOO	4.101	6.632	1.321	0.137	0.966	1.584	2.122	6.725	5.523	3.422	0.001	2.355	2.549	6.992	1.101	1.340	0.117	
10	ITA	2013-10	ITA	2013	EVOO	1.819	6.609	1.289	0.001	0.745	1.456	1.932	7.598	4.784	6.335	1.088	1.867	0.001	1.419	1.124	0.107	0.107	
11	ITA	2013-11	ITA	2013	EVOO	4.675	4.567	1.111	0.179	0.877	1.754	1.693	3.844	6.861	3.334	1.234	2.750	2.960	6.815	0.472	1.008	0.085	
12	ITA	2013-12	ITA	2013	EVOO	4.188	5.475	1.085	0.001	0.877	1.634	2.018	6.727	4.647	2.760	1.289	2.234	2.619	5.210	1.269	1.036	0.084	
13	ITA	2013-13	ITA	2013	EVOO	7.180	7.372	1.246	0.415	1.241	2.579	2.421	5.188	5.604	2.460	1.101	0.001	3.304	6.258	1.973	1.376	0.053	
14	ITA	2013-14	ITA	2013	EVOO	0.001	4.119	1.628	0.097	1.509	2.803	4.042	1.221	1.898	5.365	4.495	4.238	1.659	2.774	14.898	0.576	0.001	0.059
15	ITA	2013-15	ITA	2013	EVOO	4.395	2.998	3.691	0.519	3.247	6.959	6.464	2.884	3.886	0.001	1.255	0.001	1.481	7.753	1.507	1.208	0.199	
16	ITA	2013-16	ITA	2013	EVOO	3.374	1.053	1.998	0.001	2.275	3.349	5.416	0.342	5.463	3.714	0.591	0.944	3.372	1.911	1.784	1.377	0.100	
17	ITA	2013-17	ITA	2013	EVOO	3.040	2.958	1.945	0.001	1.837	3.207	3.004	6.640	3.864	3.645	2.156	1.112	1.679	12.612	0.751	1.283	0.034	
18	ITA	2013-18	ITA	2013	EVOO	3.686	5.500	1.834	0.381	1.870	2.855	2.687	0.001	2.744	0.001	1.773	1.893	1.773	10.416	1.576	2.177	0.001	
19	ITA	2013-19	ITA	2013	EVOO	2.816	3.109	1.930	0.142	1.687	3.201	3.471	1.087	3.973	3.969	2.128	2.440	1.915	13.697	0.533	1.922	0.038	
20	ITA	2013-20	ITA	2013	EVOO	3.569	4.037	2.294	0.126	2.191	4.099	5.592	1.802	4.964	4.567	2.863	2.110	2.622	11.289	1.446	1.638	0.047	
21	ITA	2013-21	ITA	2013	EVOO	0.001	2.976	2.351	0.152	2.150	4.312	6.126	5.525	4.502	4.804	3.576	1.285	3.134	14.219	1.121	1.873	0.007	
22	ITA	2013-22	ITA	2013	EVOO	7.101	2.521	1.019	0.987	1.714	1.724	0.649	6.212	1.447	3.702	0.001	3.891	3.940	6.891	3.940	2.987	0.164	
23	ITA	2013-23	ITA	2013	EVOO	3.837	9.058	1.132	0.421	0.869	1.747	1.781	9.114	3.525	0.001	1.706	0.978	1.560	8.286	1.008	2.377	0.001	
24	ITA	2013-24	ITA	2013	EVOO	5.126	4.623	1.910	0.296	1.597	2.798	3.317	4.672	6.252	0.001	0.001	1.940	5.307	10.371	0.768	1.516	0.051	
25	ITA	2013-25	ITA	2013	EVOO	5.071	4.071	1.279	0.258	1.790	2.009	2.743	2.784	5.805	6.802	2.068	2.830	0.001	8.860	0.607	1.641	0.001	
26	ITA	2013-26	ITA	2013	EVOO	8.550	4.715	1.047	0.195	0.798	2.588	1.622	3.629	5.794	3.299	1.791	2.111	3.809	12.597	2.072	2.237	0.001	
27	ITA	2013-27	ITA	2013	EVOO	3.244	2.567	1.792	0.662	2.062	3.614	4.308	0.469	3.597	1.540	1.832	0.001	1.982	11.681	1.868	2.712	0.116	
28	ITA	2013-28	ITA	2013	EVOO	5.342	3.960	1.215	0.610	1.242	2.145	2.546	4.304	3.594	2.104	0.748	0.001	2.764	5.251	1.462	1.510	0.177	
29	ITA	2013-29	ITA	2013	EVOO	0.001	3.072	1.404	0.927	1.484	1.610	0.712	1.860	1.617	0.001	3.479	0.001	1.356	8.308	0.001	3.706	0.001	
30	ITA	2013-30	ITA	2013	EVOO	5.303	5.067	1.121	0.207	0.945	1.734	1.553	4.005	4.548	3.110	1.136	0.001	3.238	6.444	0.579	0.996	0.001	
31	ITA	2013-31	ITA	2013	EVOO	3.661	4.019	1.123	0.113	1.032	1.738	1.959	4.516	3.871	5.151	2.332	3.662	2.974	9.840	1.338	1.489	0.065	
32	ITA	2013-32	ITA	2013	EVOO	4.964	3.781	1.333	0.123	1.344	2.507	2.544	3.614	0.001	5.167	2.311	2.878	3.214	9.375	1.564	1.508	0.039	
33	ITA	2013-33	ITA	2013	EVOO	4.797	4.742	0.990	0.144	0.990	1.348	1.915	3.904	3.909	2.441	2.139	2.843	6.949	1.234	1.291	0.011		
34	ITA	2013-34	ITA	2013	EVOO	3.797	5.054	1.387	0.207	0.821	1.704	1.834	6.212	3.625	0.001	1.397	0.001	8.370	5.370	1.534	1.508	0.039	
35	ITA	2013-35	ITA	2013	EVOO	5.511	5.951	1.362	0.229	1.174	2.076	1.911	4.744	3.971	1.771	1.627	1.842	1.957	6.234	1.060	1.030	0.001	
36	ITA	2013-36	ITA	2013	EVOO	6.602	4.635	1.155	0.153	0.923	1.829	2.033	1.174	4.927	3.770	1.707	2.417	0.001	6.685	1.361	1.129	0.001	
37	ITA	2013-37	ITA	2013	EVOO	4.120	4.584	0.900	0.141	0.759	1.442	1.336	3.769	4.099	5.123	2.022	2.950	2.318	8.197	0.422	1.013	0.001	
38	ITA	2013-38	ITA	2013	EVOO	5.090	3.953	1.371	0.123	1.407	2.660	2.342	3.896	4.442	4.967	1.453	2.902	0.001	7.994	1.212	1.322	0.001	
39	ITA	2013-39	ITA	2013	EVOO	2.534	4.265	0.967	0.100	0.746	1.453	1.358	3.941	3.730	1.899	3.343	1.711	1.112	11.112	0.481	1.001	0.001	
40	ITA	2013-40	ITA	2013	EVOO	4.050	4.532	1.183	0.130	0.858	1.780	2.194	3.333	5.016	4.957	2.202	2.820	0.001	8.926	0.421	1.119	0.001	
41	ITA	2013-41	ITA	2013	EVOO	2.801	4.214	1.024	0.099	0.803	1.492	1.757	2.214	3.662	3.095	2.028	2.880	1.681	10.991	0.476	1.089	0.049	
42	ITA	2013-42	ITA	2013	EVOO	7.040	4.505	0.957	0.182	0.634	1.621	1.236	5.555	5.334	7.537	0.895	2.774	0.001	5.610	0.419	1.146	0.111	
43	ITA	2013-43	ITA	2013	EVOO	5.798	4.362	1.005	0.100	0.880	1.599	1.880	4.380	3.886	3.426	0.792	3.823	3.153	5.770	0.352	0.711	0.001	
44	ITA	2013-44	ITA	2013	EVOO	6.888	5.384	0.866	0.210	0.635	1.209	0.964	6.964	6.262	5.463	0.001	2.660	0.001	6.445	0.642	0.990	0.001	
45	ITA	2013-45	ITA	2013	EVOO	1.116	6.983	1.140	0.343	0.762	1.641	1.747	7.919	0.982	1.235	0.310	0.001	1.717	2.075	1.340	0.495	0.100	
46	ITA	2013-46	ITA	2013	EVOO	3.842	5.766	0.825	0.215	0.597	1.246	1.046	4.019	4.723	2.250	1.408	1.837	2.229	9.052	0.534	1.318	0.097	
47	ITA	2013-47	ITA	2013	EVOO	7.523	5.575	1.216	0.228	0.908	1.824	1.613	6.318	5.366	7.239	0.589	2.963	0.001	3.447	0.566	0.734	0.070	
48	ITA	2013-48	ITA	2013	EVOO	3.722	4.587	1.115	0.165	0.977	1.863	2.548	4.470	4.990	4.226	1.010	2.795	2.004	7.202	0.461	0.933	0.001	
49	ITA	2013-49	ITA	2013	EVOO	1.489	5.009	1.032	0.198	0.795	1.520	1.816	4.703	5.043	5.285	1.112	2.591	0.001	6.473	0.464	0.986	0.106	
50	ITA	2013-50	ITA	2013	EVOO	6.233	6.233	1.101	0.199	0.818	1.667	1.455	5.064	4.336	0.001	0.001	3.443	1.621	4.890	1.136	0.895	0.001	
51	ITA	2013-51	ITA	2013	EVOO	5.874	6.045	1.072	0.194	0.804	1.771	1.611	6.480	6.292	5.634	1.233	0.001	0.001	5.564	0.339	1.090	0.072	
52	ITA	2013-52	ITA	2013	EVOO	4.402	4.846	1.690	0.505	1.192	2.505	3.599	4.794	3.668	0.001	0.644	2.056	2.187	3.716	1.810	1.079	0.001	
53	ITA	2013-53	ITA	2013	EVOO	4.907	6.750	1.309	0.378	0.951	2.038	2.061	8.295										

