



**VYSOKÁ ŠKOLA
CHEMICKO-TECHNOLOGICKÁ
V PRAZE**

**Fakulta potravinářské a biochemické technologie
Ústav analýzy potravin a výživy**

CERTIFIKOVANÁ METODIKA

Metody metabolomického profilování a stanovení analytických markerů autenticity jablek

**Kamila Bechyňská, Michaela Rektorisová, Petra Vacková, Vladimír Kocourek,
Jana Hajšlová**

**VŠCHT Praha
2023**

© Kamila Bechyňská, Michaela Rektorisová, Petra Vacková, Vladimír Kocourek,
Jana Hajšlová, VŠCHT Praha, 2023

Publikace neprošla jazykovou ani redakční kontrolou, za obsah příspěvků odpovídají autoři.

Kontakt na vedoucí autorského kolektivu: jana.hajslova@vscht.cz

ISBN 978-80-7592-230-4

OBSAH

1	Cíl metodiky	4
2	Vlastní popis metodiky	5
2.1	Použitelnost zvolené analytické strategie.....	5
2.2	Princip analytické metody.....	5
2.3	Chemikálie a spotřební materiál	5
2.4	Přístroje a zařízení.....	6
2.5	Použitý software.....	6
2.6	Pracovní postup.....	7
2.6.1	Příprava vzorku	7
2.7	Podmínky pořízení metabolomických „fingerprintů“ pomocí U-HPLC-HRMS.....	7
2.8	Detekce a konfirmace specifických markerů	8
2.8.1	U-HPLC-HRMS analýza.....	8
2.8.2	Zpracování dat	9
2.8.3	Statistické zpracování dat a tvorba modelů.....	10
2.8.4	Identifikace charakteristických markerů	22
2.9	Pracovní charakteristiky metody.....	30
3	Srovnání novosti postupů	31
4	Popis uplatnění metodiky	32
5	Ekonomické aspekty.....	33
6	Seznam použité související literatury	34
7	Seznam publikací, které předcházely metodice.....	35

1 Cíl metodiky

Jedním z dílčích cílů projektu QK1910104 (Výzkum metabolomických metod pro laboratorní ověření geografické autenticity jablek) podporovaného Ministerstvem zemědělství ČR, bylo navrhnout analytickou metodu pro získání metabolomických fingerprintů vzorků jablek a vytvořit matematicko-statistický model, podle kterého bude možno plody jabloně uváděné na trh v ČR zařadit do skupin podle odrůdy a geografického původu. Pro splnění tohoto cíle jsou dále na základě pokročilého statistického zpracování dat generovaných pomocí necílového metabolického fingerprintingu (autentických) plodů definovaného původu vybrány tzv. chemické markery autenticity. Celá tato metodika má pak poskytnout nástroj pro screeningovou laboratorní podporu ověřování autenticity významných tržních odrůd jablek v kontrolní a inspekční praxi.

2 Vlastní popis metodiky

2.1 Použitelnost zvolené analytické strategie

Použitá analytická strategie necílového metabolomického screeningu představuje současný trend aplikovaný při autentikaci potravin. Pro tuto studii byla využita moderní analytická technika ultra-vysokoúčinné kapalinové chromatografie ve spojení s vysokorozlišovací tandemovou hmotnostní spektrometrií, která umožňuje přesné a citlivé měření širokého spektra malých molekul (metabolitů) izolovaných daným postupem ze vzorku. Komplexní data generovaná při vyšetření vzorků je nutné zpracovat a interpretovat pomocí jednorozměrných a vícerozměrných statistických metod, pomocí nichž mohou být získány klasifikační modely a identifikovány diagnostické proměnné.

Představená strategie navrhuje aplikaci metody metabolomického fingerprintingu pro klasifikaci odrůdy a geografického původu jablek na základě analýzy extraktů složek povrchové, epikutikulární vrstvy plodů. Pro vytvoření statistických modelů bylo použito 274 autentických vzorků jablek (137 vzorků ve 2 biologických opakováních) odrůdy Gala, Golden Delicious, Idared a Jonagold, které pocházely z České republiky a z Polska z let 2020-2022.

2.2 Princip analytické metody

Pro izolaci metabolitů z epikutikulární vrstvy jablek byla použita extrakční směs rozpouštědel dichlormethan:methanol (1:1, v:v). Získané extrakty byly po zakoncentrování analyzovány technikou ultra-vysokoúčinné kapalinové chromatografie ve spojení s vysokorozlišovací tandemovou hmotnostní spektrometrií (U-HPLC-HRMS/MS) s hmotnostním analyzátozem typu kvadrupól-analyzátor doby letu (Q-ToF). Získané „chemické otisky“ (fingerprinty) izolovaných sekundárních metabolitů byly následně zpracovány a s využitím statistických metod byly vytvořeny vícerozměrné modely umožňující rozlišit geografický původ jablek. Statistickým hodnocením byly vytipované nejvíce diagnostické metabolity resp. jejich ionty (m/z) a byla rovněž navržena jejich identifikace pomocí přesné hmoty, isotopového profilu a fragmentačních spekter.

2.3 Chemikálie a spotřební materiál

- Methanol, gradient grade pro LC (Merck KGaA, Německo)
- Dichlormethan, čistota 99,9%, (Merck spol. s.r.o, USA)
- Deionizovaná voda (zařízení Milli-Q Millipore, Německo)
- Mravenčan amonný, čistota 99,9 %, (Sigma-Aldrich, Německo)

- Kyselina mravenčí, čistota pro MS - 98 %, (Sigma-Aldrich, Německo)
- Kalibrační roztok APCI Calibration Solution (SCIEX, Kanada)
- Membránové filtry LUT Syringe, PTFE, 25 mm, 0,22 μm (HPST, ČR)
- Špičky – Finntip, 1–10 ml (Analytika, Česká republika)
- Vialky, septa, víčka a další příslušenství (Labicom s.r.o., ČR)
- Běžné laboratorní sklo a vybavení

2.4 Přístroje a zařízení

Příprava vzorku:

- Rotační vakuová odparka s vodní lázní, Büchi Vacuum Pump V-100, Büchi Interface I-100, Büchi Rotavapor R-100, Büchi Heating Bath B-100 (Büchi, Švýcarsko)
- Třepačka HS 250 basic (IKA Labortechnik, IKA Werk, Německo)
- Analytické váhy Balance AND, model GF-1200 (A&D Instruments LTD., Japonsko)
- Automatické pipety (TraffLab-Transferpette[®], Švýcarsko)
- Přístroj na přípravu deionizované vody Milli-Q (Merck Millipore, Německo)

Instrumentální systém U-HPLC-HRMS:

- Kapalinový chromatograf U-HPLC Dionex UltiMate 3000 (Thermo Fisher Scientific, USA)
- Hmotnostní spektrometr typu QTOF – TripleTOF[™] 6600 (SCIEX, Kanada)
- Analytická kolona BEH C18 (2,1 mm $\times$ 100 mm; 1,7 μm) (Waters, USA)

2.5 Použitý software

- Analyst[®] verze 1.6 TF (Sciex, Kanada)
- SCIEX OS, verze 1.5 (Sciex, Kanada)
- MS Dials, verze 5.0.6
- MS Finder, verze 3.52
- SIMCA[™] verze 13.0 (Umetrics, Švédsko)
- Rstudio, verze 1.3.1073
- Microsoft Excel (Microsoft Corporation, USA)

Poznámka: Kromě přístroje a zařízení uvedených dodavatelů, které byly použity pro tuto metodiku, lze použít instrumentaci od alternativních dodavatelů, pokud jsou pro daný účel vhodné. Vhodnost musí být dokumentována při implementaci metody v konkrétních podmínkách.

2.6 Pracovní postup

2.6.1 Příprava vzorku

Vzorky jablek o známé hmotnosti byly vloženy do kádinky a převrstveny směsí rozpouštědel dichlormethan:methanol (1:1,v:v) tak, aby byl celý plod ponořený. Kádinka byla po přikrytí víčkem umístěna na třepačku a oplach povrchu jablka probíhal po dobu 10 minut s frekvencí třepání 120 min⁻¹. Extrakt byl převeden do odpařovací baňky a zakonzentrován pomocí vakuové odpary. Zbytkové rozpouštědlo bylo odfoukáno proudem dusíku. Suchý extrakt (odparek) byl zvážen a následně rozpuštěn ve směsi rozpouštědel dichlormethan:methanol (1:1,v:v) na výslednou koncentraci matrice 33,33 mg/ml. Rozpuštěný extrakt byl přefiltrován přes stříkačkový filtr s velikostí pórů 0,22 µm a převeden do skleněné vialky. Vzorek pro kontrolu kvality (QC) měření dat byl připraven jako pool alikvotů všech připravených vzorků.

2.7 Podmínky pořízení metabolomických „fingerprintů“ pomocí U-HPLC-HRMS

Pro analýzu vzorků byla využita technika ultra účinné kapalinové chromatografie ve spojení s vysokorozlišovací tandemovou hmotnostní spektrometrií (UHPLC-HRMS/MS) s hmotnostním analyzátozem typu kvadrupól-analyzátor doby letu (Q-ToF).

Podmínky měření extraktů jsou uvedeny v **Tabulkách 1-3**.

Tabulka 1: Podmínky UHPLC separace

U-HPLC Dionex UltiMate 3000	
Kolona	BEH C18 (2,1x100 mm; 1,7 µm)
Teplota kolony	60 °C
Objem nástřiku	1 µl
Teplota autosampleru	5 °C
Mobilní fáze	A: 5 mM mravenčan amonný ve směsi voda:methanol (95:5, v:v) + 0,1 % kyselina mravenčí
	B: 5 mM mravenčan amonný ve směsi isopropanol, methanol a voda (65:30:5, v:v:v) + 0,1 % kyselina mravenčí

Tabulka 2: Gradient mobilní fáze

Čas [min]	Průtok [ml/min]	A [%]	B [%]
0	0,4	70	30
2	0,4	50	50
7	0,4	20	80
13	0,4	0	100
20	0,4	0	100
20,1	0,4	70	30
22	0,4	70	30

Tabulka 3: Podmínky MS detekce

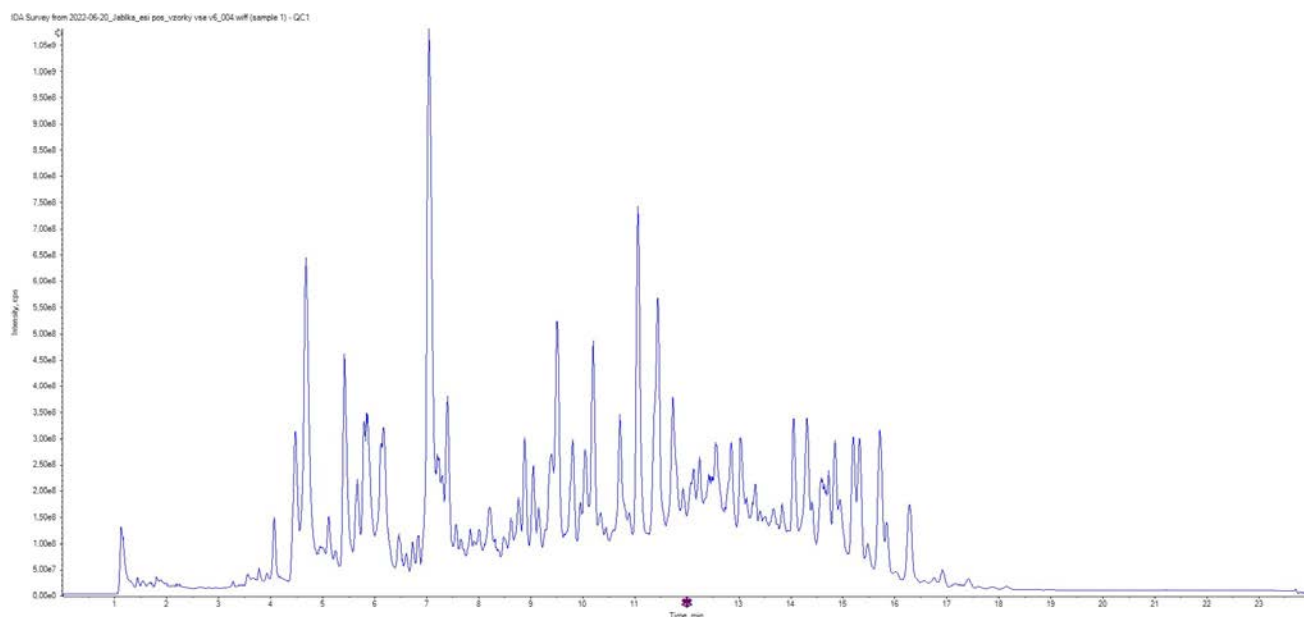
QTOF-TripleTOF™ 6600	
Iontový zdroj	Duo Spray™ s oddělenými ESI a APCI iontovými zdroji
Ionizační technika	ESI +/ESI-
Akviziční mód	MS1 + MS2 (DDA)
Teplota iontového zdroje	500 °C
Napětí na kapiláře	+4500 V/-4000 V
Rozsah snímaných hmot (<i>m/z</i>)	100–1200
Kolizní energie	35 ± 15 eV
Tlak nepulzačního plynu	55 psi (379 kPa)
Tlak sušícího plynu	55 psi (379 kPa)
Dosažené rozlišení	FWHM > 40 000 (<i>m/z</i> 829,5393 pro ESI+ a <i>m/z</i> 933,637 pro ESI-)

Automatická kalibrace přesnosti hmoty se provádí pomocí kalibračního roztoku po každých 10 nástricích. Pro kontrolu kvality měřených dat byl použit QC vzorek, který byl nastříknut v 8 opakováních před začátkem sekvence, dále po každých 17 vzorcích v průběhu sekvence a v 6 opakováních na konci sekvence. Pro zhodnocení kvality měření byla porovnána konzistence intenzity signálů, stabilita retenčních časů a tlakových záznamů mobilní fáze.

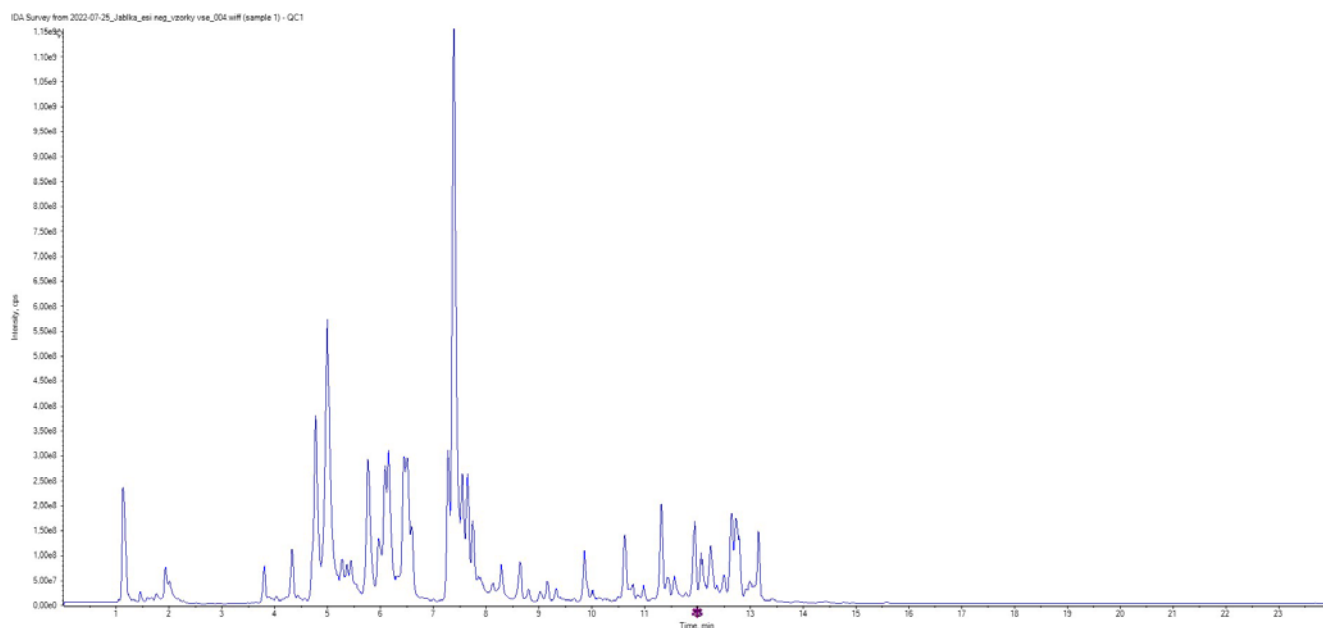
2.8 Detekce a konfirmace specifických markerů

2.8.1 U-HPLC-HRMS analýza

Extrakty metabolitů z epikutikulární vrstvy jablek byly analyzovány technikou UHPLC-HRMS/MS v pozitivním (ESI+) i negativním (ESI-) módu ionizace. Získané „fingerprinty“ jsou v obou polaritách poměrně bohaté, ukázka chromatogramů je uvedena na **Obrázcích 1 a 2**.



Obrázek 1: Chromatogram celkové iontové intezity QC vzorku z pozitivního módu ionizace ESI+



Obrázek 2: Chromatogram celkové iontové intezity QC vzorku z negativního módu ionizace ESI-

2.8.2 Zpracování dat

Získaná data byla po konverzi do formátu *.abf zpracována v programu MS-Dial, který umožňuje extrakci proměnných, deizotopizaci a alignment. Parametry zpracování dat jsou uvedeny v **Tabulce 4**.

Tabulka 4: Parametry zpracování dat v programu MS Dial

Tolerance přesné hmoty (m/z)	MS1 (0,01 Da), MS2 (0,025 Da)
Minimální výška píku	10 000
Typ aduktů	[M+H] ⁺ , [M+NH ₄] ⁺ , [M+Na] ⁺ (ESI ⁺) [M-H] ⁻ , [M+HCOO] ⁻ ; (ESI ⁻)
Tolerance retenčního času pro alignment dat	0,05 min (ESI ⁺), 0,3 min; (ESI ⁻)
Tolerance přesné hmoty (m/z) pro zarovnání dat	0,015 Da (ESI ⁺ i ESI ⁻)

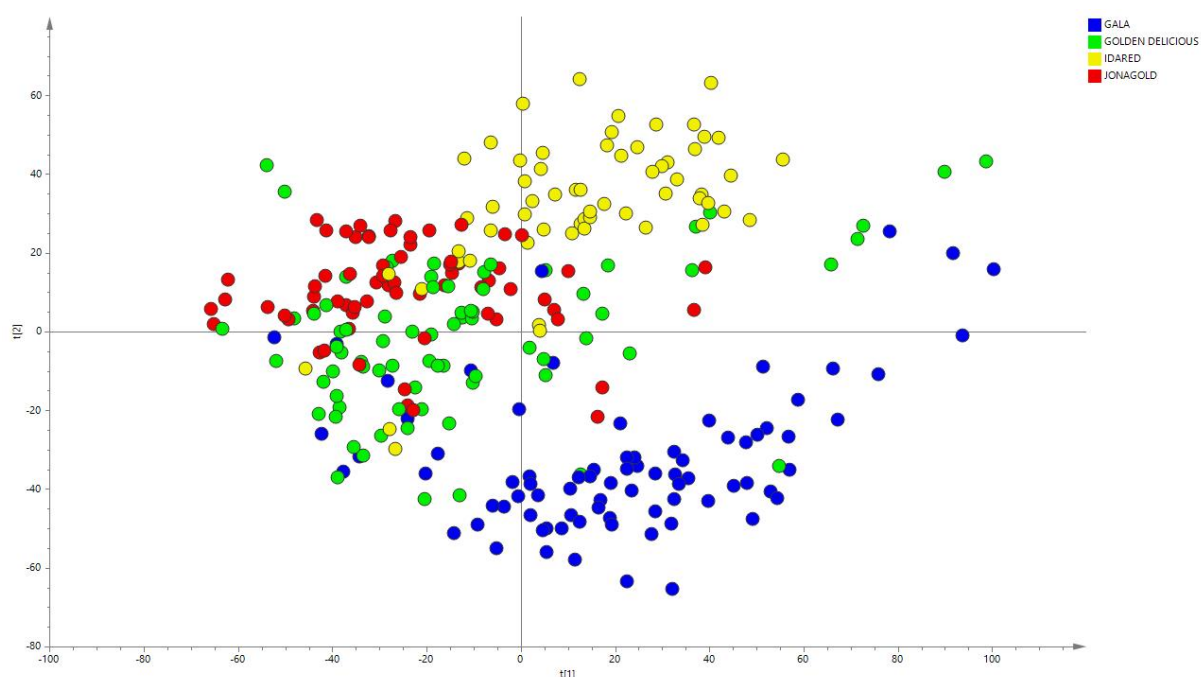
Výsledkem zpracování dat byly datové matice, které čítaly v pozitivním módu ionizace 96072 proměnných a v negativním módu ionizace 21040 proměnných. Proměnné jsou charakterizovány naměřenou přesnou hmotou iontů a retenčním časem. Následovala filtrace proměnných podle relativní směrodatné odchylky (RSD) intenzity signálů (plochy píků) v QC vzorcích. Dále byly odfiltrovány všechny proměnné, jejichž plocha ve vzorcích nebyla alespoň 3x vyšší než v rozpouštědlovém blanku. Pro následné statistické zpracování bylo použito 16044 proměnných detekovaných v ESI⁺ a 2132 proměnných detekovaných v ESI⁻.

2.8.3 Statistické zpracování dat a tvorba modelů

Získané datové matice byly před každým statistickým zpracováním normalizovány na sumu, logaritmicky transformovány a Pareto škálovány. Na data byly následně aplikovány jednorozměrné a vícerozměrné statistické metody s cílem vytvoření modelů, které budou na základě diagnostických proměnných „markerů“ rozlišovat odrůdu a geografický původ jablek. Pro náhled na vícerozměrná data byla využita analýza hlavních komponent (*Principal Component Analysis*, PCA). Diagnostické proměnné byly hledány pomocí metod t-testu, změny střední hodnoty a Receiver Operating Characteristics (ROC), kde každá z nich posuzuje významnost proměnné jiným způsobem a umožňuje tak výběr relevantních markerů. Na základě vybraných proměnných byly vytvořeny klasifikační modely diskriminační analýzy parciálně nejmenších čtverců (*Partial least square Discriminant Analysis*, PLS-DA) a diskriminační analýzy orthogonálních parciálních nejmenších čtverců (*Orthogonal Partial least square Discriminant Analysis*, OPLS-DA). Klasifikační modely byly validovány pomocí sedminásobné křížové validace a charakterizovány popsáním (R^2Y) a predikovaným (Q^2Y) rozptylem. Jako validní byl klasifikován model, jehož hodnota Q^2Y nabývala minimálně hodnoty 0,5 [1].

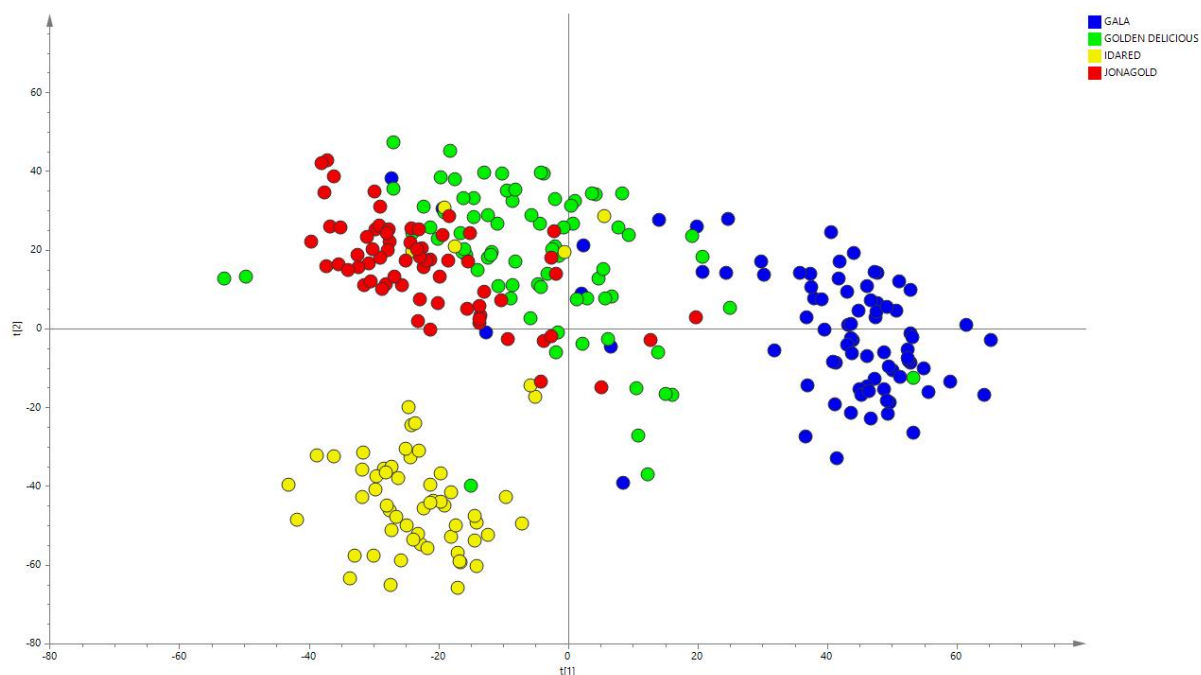
2.8.3.1 Klasifikace odrůdy na základě metabolického fingerprintu získaného v ESI+

Na metabolické fingerprinty zahrnující všech 16044 proměnných pro celkem 284 vzorků jablek 4 odrůd (Golden Delicious, Jonagold, Idared, Gala) byla aplikována analýza hlavních komponent (**Obrázek 3**). Na grafu komponentního skóre je patrné oddělení vzorků jablek odrůdy Gala (modrá) a Idared (žlutá). U vzorků odrůd Golden Delicious (zelená) a Jonagold (červená) k průkazné separaci nedošlo.



Obrázek 3: PCA graf komponentního skóre metabolických fingerprintů extraktů vosků jablek odrůdy Gala (modrá), Golden Delicious (zelená), Idared (žlutá) a Jonagold (červená) v ESI+

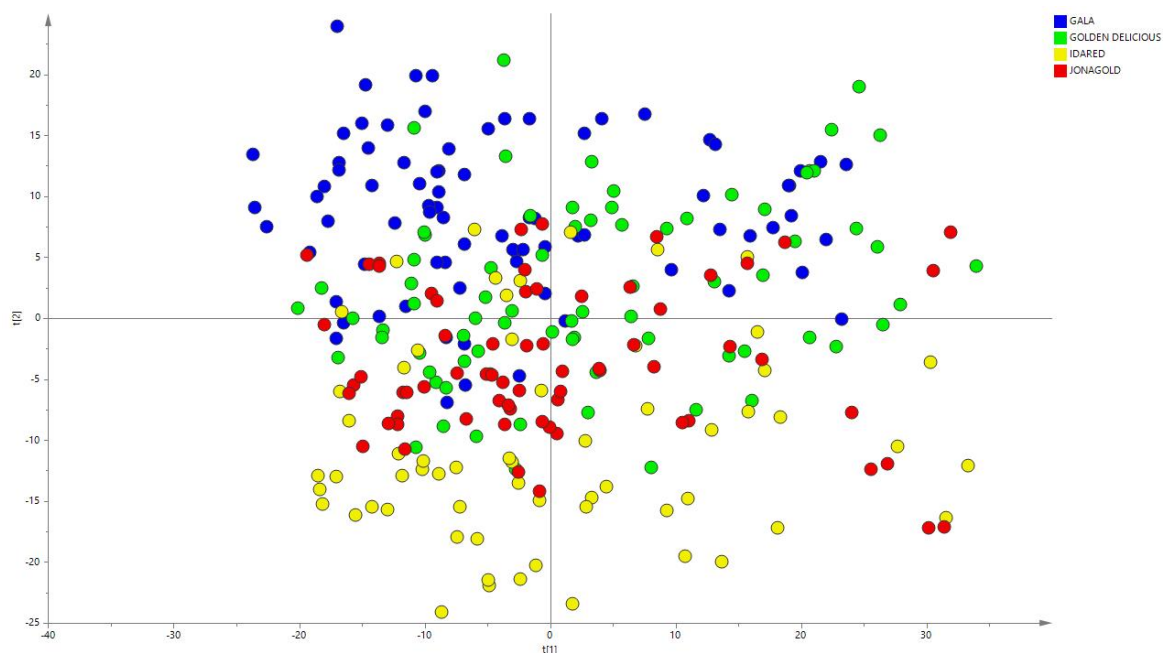
Na proměnné byla aplikována analýza rozptylu (*Analysis of Variance*, ANOVA), která vybrala 14551 proměnných, které se na hladině významnosti $\alpha=0,05$ statisticky významně lišily alespoň mezi jednou dvojicí odrůd. Následně byl na této podmnožině proměnných vytvořen PLS-DA model (**Obrázek 4**). Validita modelu byla ověřena na základě sedminásobné křížové cross validace a parametr R^2Y dosáhl hodnoty 0,790 a Q^2Y hodnoty 0,756, což splňuje požadavky na validní model. Významné proměnné byly vybrány na základě parametru VIP skóre $> 1,5$ a u těchto proměnných byla snaha o jejich identifikaci.



Obrázek 4: PLS-DA graf komponentního skóre metabolomických fingerprintů extraktů vosků jablek odrůdy Gala (modrá), Golden Delicious (zelená), Idared (žlutá) a Jonagold (červená) v ESI+

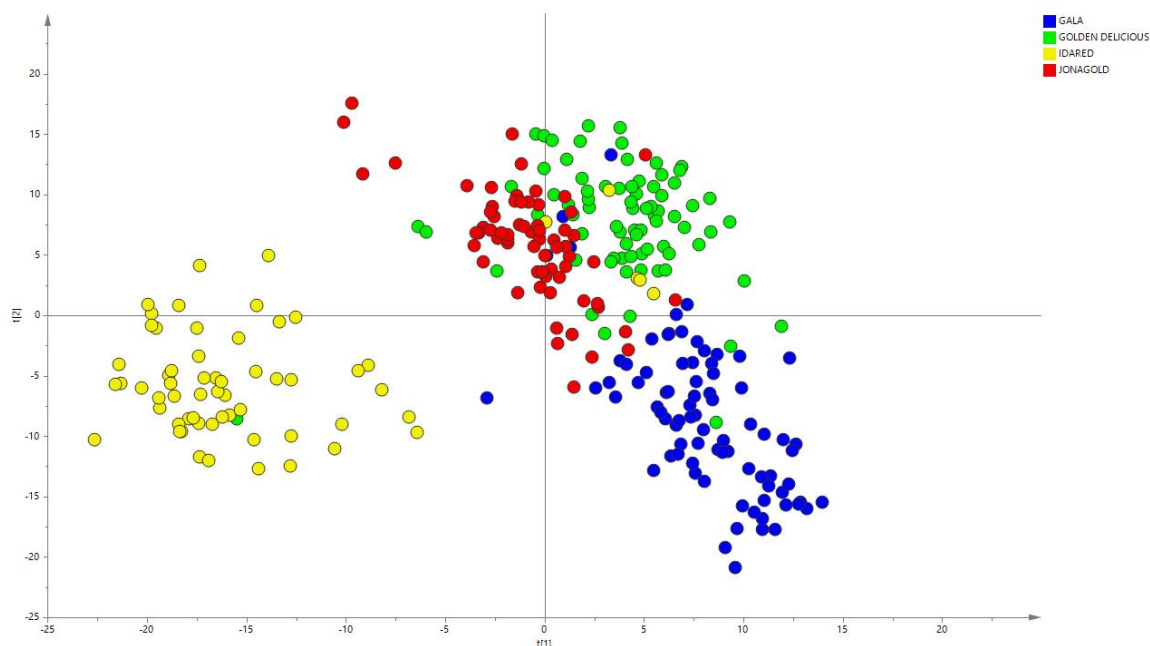
2.8.3.2 Klasifikace odrůdy na základě metabolomického fingerprintu získaného v ESI-

Na metabolomické fingerprinty zahrnující všech 2132 proměnných pro celkem 284 vzorků jablek 4 odrůd (Golden Delicious, Jonagold, Idared, Gala) byla aplikována analýza hlavních komponent (**Obrázek 5**). Na grafu komponentního skóre je patrný náznak oddělení vzorků jablek odrůdy Gala (modrá) a Golden Delicious (zelená) od jablek odrůdy Idared (žlutá) a Jonagold (červená).



Obrázek 5: PCA graf komponentního skóre metabolomických fingerprintů extraktů vosků jablek odrůdy Gala (modrá), Golden Delicious (zelená), Idared (žlutá) a Jonagold (červená) v ESI-

Na proměnné byla aplikována ANOVA, která vybrala 1678 proměnných, které se na hladině významnosti $\alpha=0,05$ statisticky významně lišily alespoň mezi jednou dvojicí odrůd. Následně byl na této podmnožině proměnných vytvořen PLS-DA model (**Obrázek 6**). Validita modelu byla ověřena na základě sedminásobné křížové cross validace a parametr R^2Y dosáhl hodnoty 0,770 a Q^2Y hodnoty 0,756, což splňuje požadavky na validní model. Významné proměnné pro ESI- byly vybrány na základě parametru VIP skóre, který byl u 122 proměnných vyšší než 1,5 a u těchto proměnných byla snaha o jejich identifikaci.



Obrázek 6: PLS-DA graf komponentního skóre metabolomických fingerprintů extraktů vosků jablek odrůdy Gala (modrá), Golden Delicious (zelená), Idared (žlutá) a Jonagold (červená) v ESI-

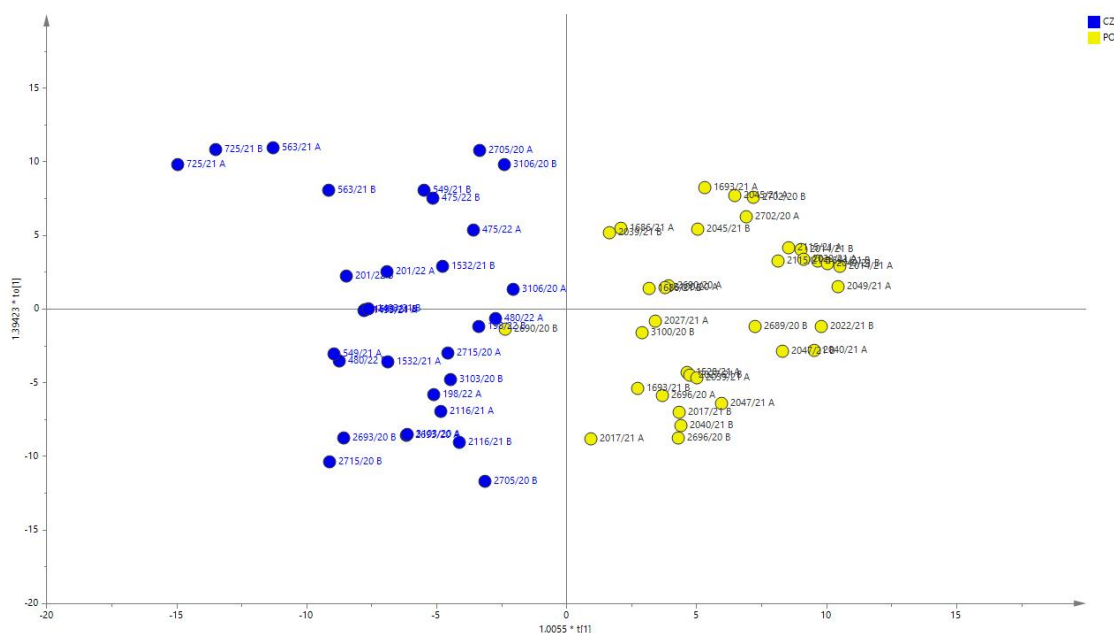
2.8.3.3 Klasifikace země původu na základě metabolomického fingerprintu získaného v ESI+

Odrůda jablka má významnější vliv na metabolom epikutikulární vrstvy než geografický původ plodu, proto je nutné analyzovat vliv geograficity samostatně pro každou odrůdu. Výchozím souborem byla datová matice s 16 044 proměnnými, na data byl následně aplikován t-test a ROC a byla spočítána změna střední hodnoty. Z proměnných, které splnily podmínku t-testu o statistické významnosti na hladině $\alpha=0,05$ a hodnota area under curve (AUC) byla vyšší než 0,75, byl vytvořen OPLS-DA model. Validace modelu byla provedena pomocí sedminásobné křížové validace a byly získány parametry RMSEE, R^2Y a Q^2Y . Provedeny byly rovněž permutační testy (100 permutací), jejichž výsledek je vyjádřen jako pravděpodobnost, že parametry R^2Y a Q^2Y vytvořeného modelu se shodují s parametry získaných z modelů při permutačních testů. Dále byla ověřena validita modelu v čase a tím i relevance použitých proměnných, kdy do OPLS-DA modelu vytvořeného ze vzorků z roku 2021 byly vloženy vzorky z roku 2020 a 2022 a byla spočítána predikční schopnost modelu správně určit geografický původ.

Gala

Metabolomické fingerprinty extraktů vosků jablek odrůdy Gala zahrnovaly 35 vzorků, z nichž 16 pocházelo z České republiky (5 vzorků z roku 2020, 8 vzorků z roku 2021, 3 vzorky z roku 2022) a 19 pocházelo z Polska (5 vzorků z roku 2020, 13 vzorků z roku 2021, 1 vzorek z roku 2022). Všechny vzorky byly ve 2 biologických repikátech. Z původních 16 044 proměnných bylo na základě t-testu na hladině významnosti $\alpha=0,05$ vybráno 1 706

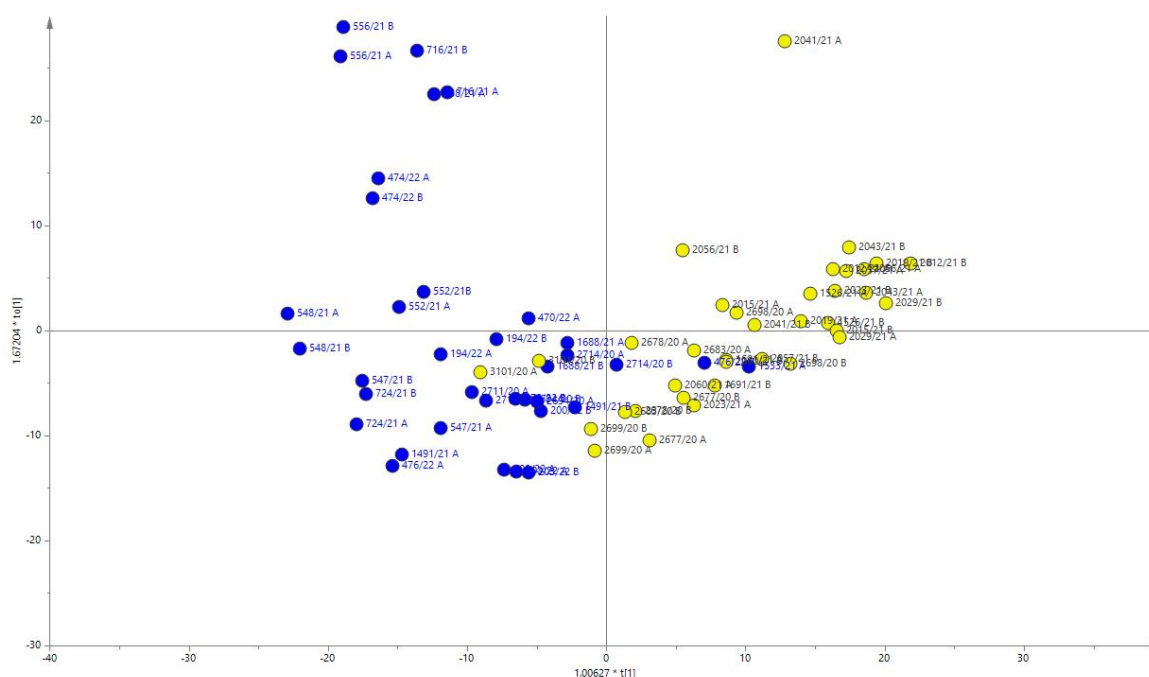
proměnných, které se statisticky významně lišily mezi vzorky z České republiky a z Polska. Z nich bylo vybráno pouze 506 proměnných, které splňovaly hodnotu $AUC > 0,75$. Vytvořený OPLS-DA model (**Obrázek 7**) dosáhl parametrů RMSEE 0,265, R^2Y 0,735 a Q^2Y 0,624. Výsledek permutačních testů byl pro oba parametry R^2Y a Q^2Y roven pravděpodobnosti menší než 1%. Predikční schopnost OPLS-DA modelu vytvořeného ze vzorků z roku 2020 a 2021, do kterého byly vloženy vzorky z roku 2022 dosáhla hodnoty 82 %.



Obrázek 7: OPLS-DA graf komponentního skóre vytvořený na základě 506 markerů získaných z metabolomických fingerprintů extraktů vosků jablek odrůdy Gala pocházejících z České republiky (modrá) a z Polska (žlutá) v ESI+

Golden Delicious

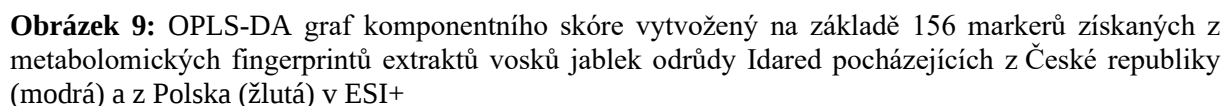
Metabolomické fingerprinty extraktů vosků jablek odrůdy Golden Delicious zahrnovaly 37 vzorků, z nichž 19 pocházelo z České republiky (3 vzorky z roku 2020, 10 vzorků z roku 2021, 6 vzorků z roku 2022) a 18 pocházelo z Polska (6 vzorků z roku 2020, 11 vzorků z roku 2021, 1 vzorek z roku 2022). Všechny vzorky byly ve 2 biologických replikátech. Z původních 16 044 proměnných bylo na základě t-testu na hladině významnosti $\alpha = 0,05$ vybráno 5 792 proměnných, které se statisticky významně lišily mezi skupinou vzorků z České republiky a z Polska. Z nich bylo vybráno pouze 1 048 proměnných, které splňovaly hodnotu $AUC > 0,75$. Vytvořený OPLS-DA model (**Obrázek 8**) dosáhl parametrů RMSEE = 0,309, R^2Y = 0,635 a Q^2Y = 0,554. Výsledek permutačních testů byl pro oba parametry R^2Y a Q^2Y roven pravděpodobnosti menší než 1%. Predikční schopnost OPLS-DA modelu vytvořeného ze vzorků z roku 2021, do nějž byly vloženy vzorky z roku 2020 a 2022 dosáhla hodnoty 65 %.



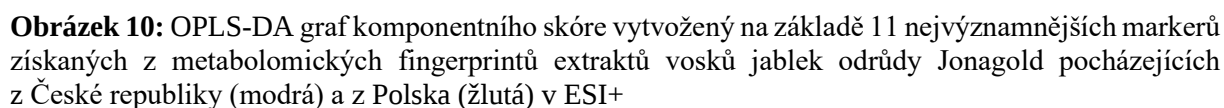
Obrázek 8: OPLS-DA graf komponentního skóre vytvořený na základě 1048 markerů získaných z metabolických fingerprintů extraktů vosků jablek odrůdy Golden Delicious pocházejících z České republiky (modrá) a z Polska (žlutá) v ESI+

Idared

Metabolické fingerprinty extraktů vosků jablek odrůdy Idared zahrnovaly 32 vzorků, z nichž 17 pocházelo z České republiky (5 vzorků z roku 2020, 7 vzorků z roku 2021, 5 vzorků z roku 2022) a 15 pocházelo z Polska (5 vzorků z roku 2020, 10 vzorků z roku 2021). Všechny vzorky byly ve 2 biologických repikátech. Z původních 16 044 proměnných bylo na základě t-testu na hladině významnosti $\alpha=0,05$ vybráno 438 proměnných, které se statisticky významně lišily mezi skupinou vzorků z České republiky a Polska. Z nich bylo vybráno pouze 156 proměnných, které splňovaly hodnotu $AUC > 0,75$. Vytvořený OPLS-DA model (**Obrázek 9**) dosáhl parametrů $RMSEE=0,175$, R^2Y 0,886 a Q^2Y 0,809. Výsledek permutačních testů byl pro oba parametry R^2Y a Q^2Y roven pravděpodobnosti menší než 1%. Predikční schopnost OPLS-DA modelu vytvořeného ze vzorků z roku 2021, do nějž byly vloženy vzorky z roku 2020 a 2022 dosáhla hodnoty 85 %.



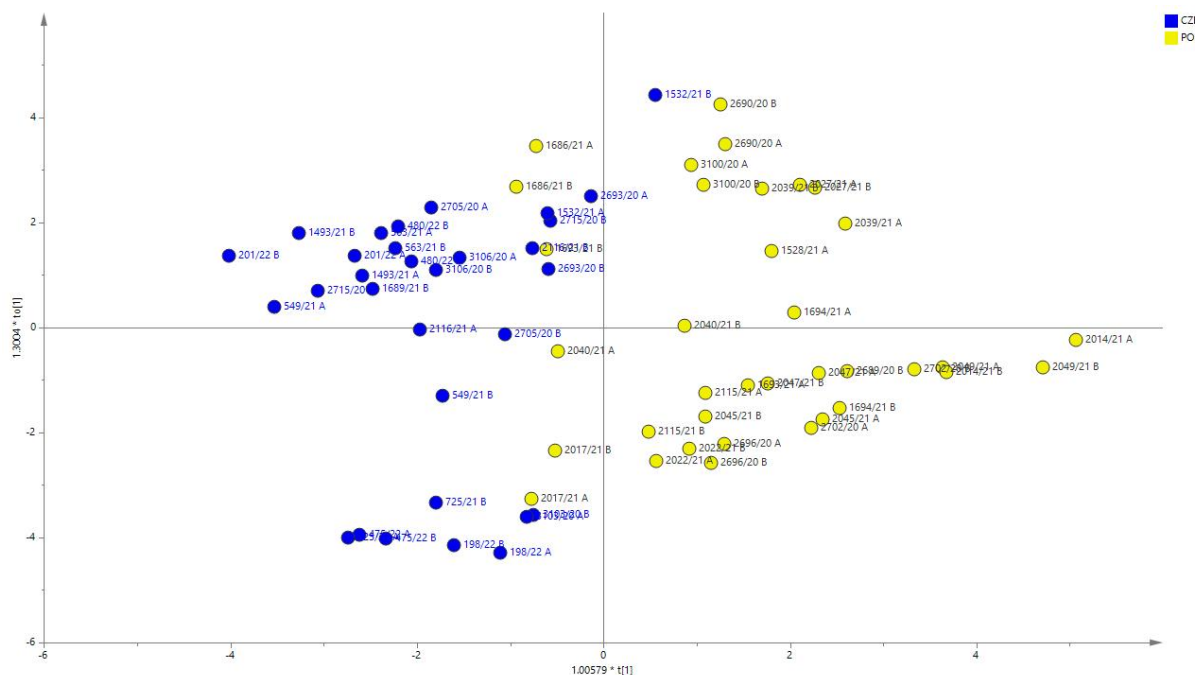
Metabolomické fingerprinty extraktů vosků jablek odrůdy Jonagold zahrnovaly 33 vzorků, z nichž 13 pocházelo z České republiky (6 vzorků z roku 2020, 3 vzorky z roku 2021 a 4 vzorky z roku 2022) a 20 pocházelo z Polska (8 vzorků z roku 2020, 12 vzorků z roku 2021). Všechny vzorky byly ve 2 biologických repikátech. Z původních 16044 proměnných bylo na základě t-testu na hladině významnosti $\alpha=0,05$ vybráno 15 proměnných, které se statisticky významně lišily mezi skupinou vzorků z České republiky a Polska. Z nich bylo vybráno pouze 12 proměnných, které splňovaly hodnotu $AUC > 0,75$. Po kontrole tvaru chromatografických píků bylo vybráno všech výsledných 11 markerů, z nichž byl vytvořen OPLS-DA model (**Obrázek 10**). Parametr RMSEE dosahoval hodnoty 0,335, parametr R^2Y dosahoval hodnoty 0,561 a parametr Q^2Y dosahoval hodnoty 0,501. Výsledek permutačních testů byl pro oba parametry R^2Y a Q^2Y roven pravděpodobnosti menší než 1%. Predikční schopnost OPLS-DA modelu vytvořeného ze vzorků z roku 2020 a 2021, do kterého byly vloženy vzorky z roku 2022, dosáhla hodnoty 88 %.



Stejný postup byl aplikován i u dat z negativního módu ionizace. Výchozím souborem byla datová matice s 2132 proměnnými, na data byl následně aplikován t-test a ROC a byla spočítána změna střední hodnoty. U proměnných, které splnily podmínku t-testu o statistické významnosti na hladině $\alpha=0,05$ a AUC byla vyšší než 0,75, byl zkontrolován tvar odpovídajícího chromatografického píku. Z takto vybraných proměnných „markerů“ byl vytvořen OPLS-DA model, který byl validován na základě sedmi-kolové křížové validace a popsán pomocí parametrů R^2Y a Q^2Y , dále byla ověřena validita modelu v čase, kdy do OPLS-DA modelu vytvořeného ze vzorků z roku 2021 byly vloženy vzorky z roku 2020 a 2022 a byla spočítána predikční schopnost.

Metabolomické fingerprinty extraktů vosků jablek odrůdy Gala zahrnovaly 35 vzorků, z nichž 16 pocházelo z České republiky (5 vzorků z roku 2020, 8 vzorků z roku 2021, 3 vzorky z roku 2022) a 19 pocházelo z Polska (5 vzorků z roku 2020, 13 vzorků z roku 2021, 1 vzorek z roku 2022). Všechny vzorky byly ve 2 biologických repikátech. Z původních 2 132 proměnných bylo na základě t-testu na hladině významnosti $\alpha=0,05$ vybráno 52 proměnných, které se statisticky významně lišily mezi skupinou vzorků z České republiky a Polska. Z nich bylo vybráno pouze 17 proměnných, které splňovaly hodnotu $AUC > 0,75$. Po kontrole tvaru chromatografických píků bylo vybráno výsledných 13 markerů, z nichž byl vytvořen OPLS-

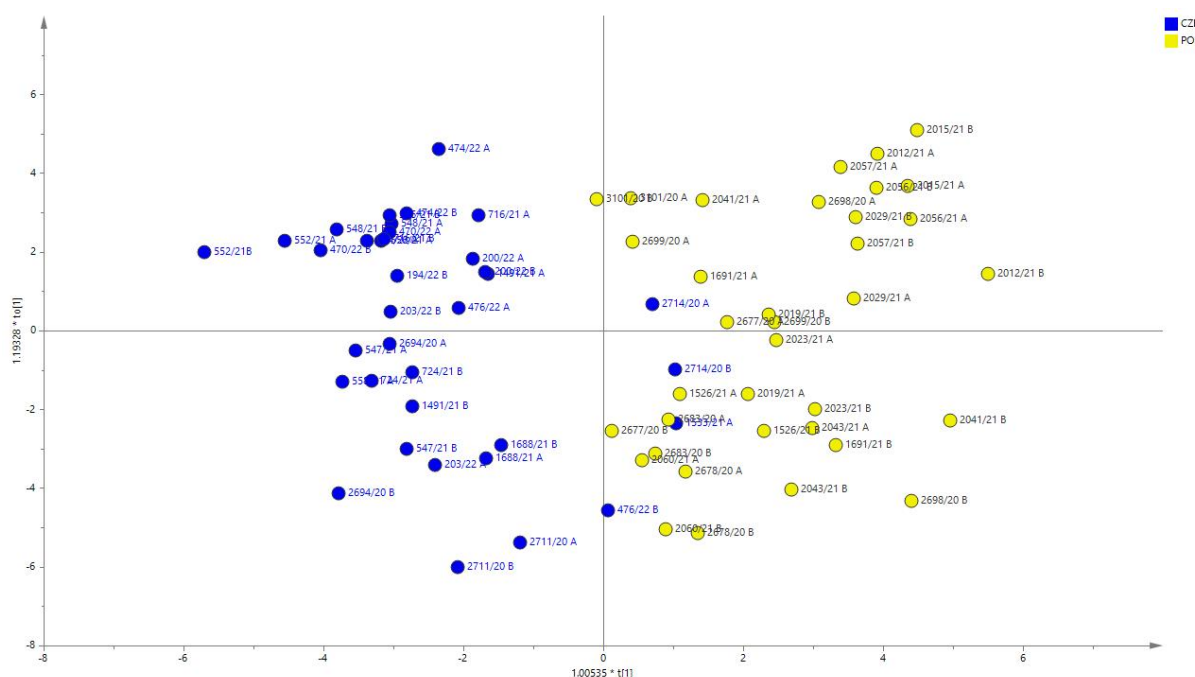
DA model (**Obrázek 11**). Parametr RMSEE dosahoval 0,307, parametr R^2Y dosahoval hodnoty 0,639 a parametr Q^2Y dosahoval hodnoty 0,543. Výsledek permutačních testů byl pro oba parametry R^2Y a Q^2Y roven pravděpodobnosti menší než 1%. Predikční schopnost OPLS-DA modelu vytvořeného ze vzorků z roku 2021, do kterého byly vloženy vzorky z roku 2020 a 2022 dosáhla hodnoty 78 %.



Obrázek 11: OPLS-DA graf komponentního skóre vytvořený na základě 13 nejvýznamnějších markerů získaných z metabolomických fingerprintů extraktů vosků jablek odrůdy Gala pocházejících z České republiky (modrá) a z Polska (žlutá) v ESI-

Golden Delicious

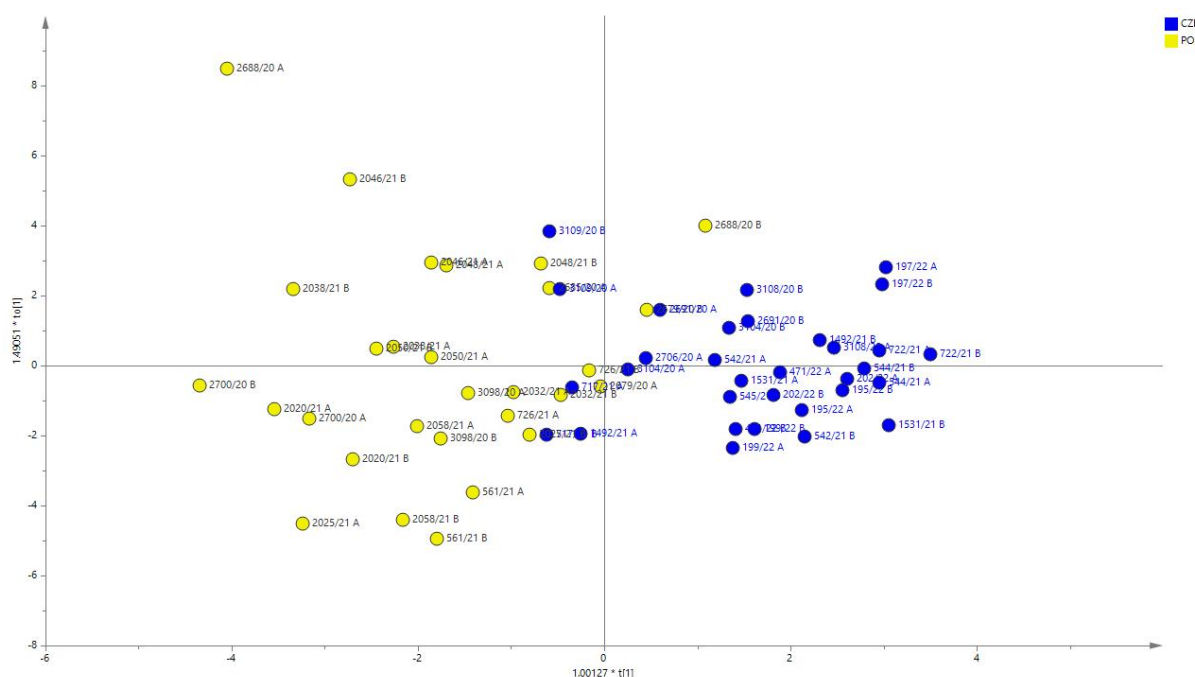
Metabolomické fingerprinty extraktů vosků jablek odrůdy Golden Delicious zahrnovaly 37 vzorků, z nichž 19 pocházelo z České republiky (3 vzorky z roku 2020, 10 vzorků z roku 2021, 6 vzorků z roku 2022) a 18 pocházelo z Polska (6 vzorků z roku 2020, 11 vzorků z roku 2021, 1 vzorek z roku 2022). Všechny vzorky byly ve 2 biologických repikátech. Z původních 2132 proměnných bylo na základě t-testu na hladině významnosti $\alpha=0,05$ vybráno 260 proměnných, které se statisticky významně lišily mezi skupinou vzorků z České republiky a Polska. Z nich bylo vybráno pouze 57 proměnných, které splňovaly hodnotu $AUC > 0,75$. Po kontrole tvaru chromatografických pík bylo vybráno výsledných 44 markerů, z nichž byl vytvořen OPLS-DA model (**Obrázek 12**). Parametr RMSEE dosahoval 0,261, parametr R^2Y dosahoval hodnoty 0,738 a parametr Q^2Y dosahoval hodnoty 0,686. Výsledek permutačních testů byl pro oba parametry R^2Y a Q^2Y roven pravděpodobnosti menší než 1%. Predikční schopnost OPLS-DA modelu vytvořeného ze vzorků z roku 2021, než byly vloženy vzorky z roku 2020 a 2022 dosáhla hodnoty 77 %.



Obrázek 12: OPLS-DA graf komponentního skóre vytvořený na základě 44 nejvýznamnějších markerů získaných z metabolomických fingerprintů extraktů vosků jablek odrůdy Golden Delicious pocházejících z České republiky (modrá) a z Polska (žlutá) v ESI-

Idared

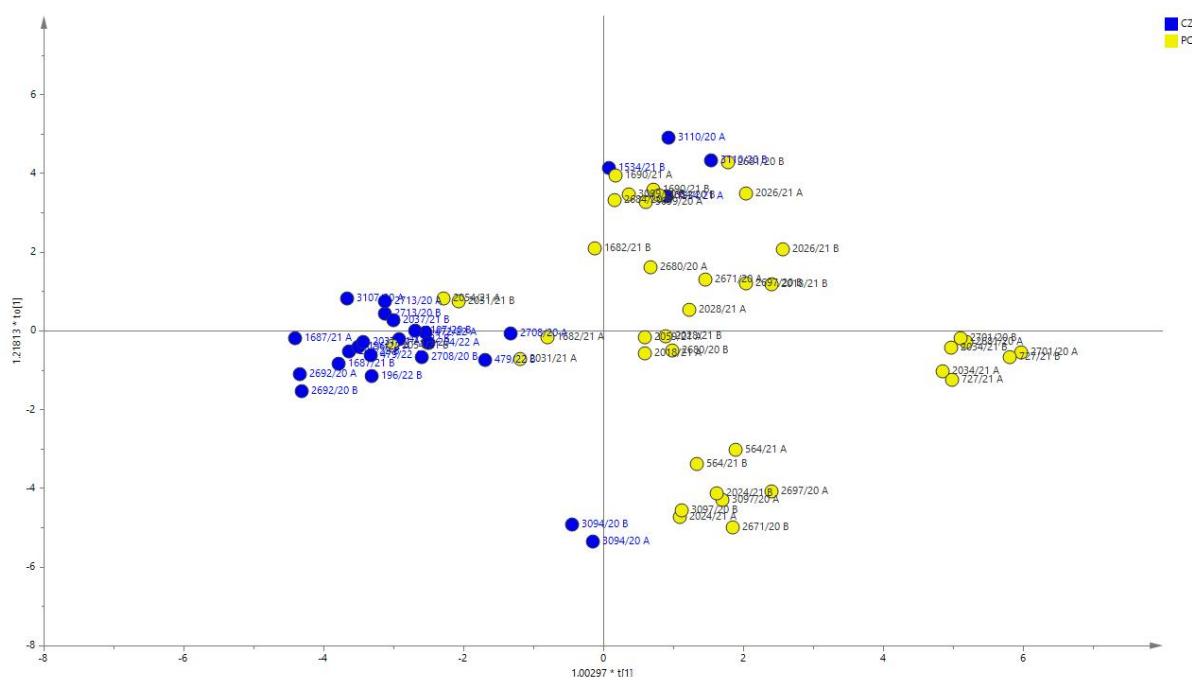
Metabolomické fingerprinty extraktů vosků jablek odrůdy Idared zahrnovaly 32 vzorků, z nichž 17 pocházelo z České republiky (5 vzorků z roku 2020, 7 vzorků z roku 2021, 5 vzorků z roku 2022) a 15 pocházelo z Polska (5 vzorků z roku 2020, 10 vzorků z roku 2021). Všechny vzorky byly ve 2 biologických repikátech. Z původních 2132 proměnných bylo na základě t-testu na hladině významnosti $\alpha=0,05$ vybráno 88 proměnných, které se statisticky významně lišily mezi skupinou vzorků z České republiky a Polska. Z nich bylo vybráno pouze 31 proměnných, které splňovaly hodnotu $AUC > 0,75$. Po kontrole tvaru chromatografických píků bylo vybráno výsledných 24 markerů, z nichž byl vytvořen OPLS-DA model (**Obrázek 13**). Parametr RMSEE dosahoval 0,308, parametr R^2Y dosahoval hodnoty 0,646 a parametr Q^2Y dosahoval hodnoty 0,574. Výsledek permutačních testů byl pro oba parametry R^2Y a Q^2Y roven pravděpodobnosti menší než 1%. Predikční schopnost OPLS-DA modelu vytvořeného ze vzorků z roku 2021, než byly vloženy vzorky z roku 2020 a 2022 dosáhla hodnoty 78 %.



Obrázek 13: OPLS-DA graf komponentního skóre vytvořený na základě 24 nejvýznamnějších markerů získaných z metabolomických fingerprintů extraktů vosků jablek odrůdy Idared pocházejících z České republiky (modrá) a z Polska (žlutá) v ESI-

Jonagold

Metabolomické fingerprinty extraktů vosků jablek odrůdy Jonagold zahrnovaly 33 vzorků, z nichž 13 pocházelo z České republiky (6 vzorků z roku 2020, 3 vzorků z roku 2021, 4 vzorků z roku 2022) a 20 pocházelo z Polska (8 vzorků z roku 2020, 12 vzorků z roku 2021). Všechny vzorky byly ve 2 biologických repikátech. Z původních 2 132 proměnných bylo na základě t-testu na hladině významnosti $\alpha=0,05$ vybráno 16 proměnných, které se statisticky významně lišily mezi skupinou vzorků z České republiky a Polska. Z nich bylo vybráno pouze 9 proměnných, které splňovaly hodnotu $AUC > 0,75$. Po kontrole tvaru chromatografických pík bylo vybráno všech výsledných 9 markerů, z nichž byl vytvořen OPLS-DA model (**Obrázek 14**). Parametr RMSEE dosahoval 0,362, parametr R^2Y dosahoval hodnoty 0,480 a parametr Q^2Y dosahoval hodnoty 0,436. Výsledek permutačních testů byl pro oba parametry R^2Y a Q^2Y roven pravděpodobnosti menší než 1%. Predikční schopnost OPLS-DA modelu vytvořeného ze vzorků z roku 2021, do kterého byly vloženy vzorky z roku 2020 a 2022 dosáhla hodnoty 88 %.



Obrázek 14: OPLS-DA graf komponentního skóre vytvořený na základě 9 nejvýznamnějších markerů získaných z metabolomických fingerprintů extraktů vosků jablek odrůdy Jonagold pocházejících z České republiky (modrá) a z Polska (žlutá) v ESI-

2.8.4 Identifikace charakteristických markerů

Proměnné, které byly použity pro tvorbu klasifikačních modelů, jsou popsány pomocí charakteristik jednorozměrných i vícerozměrných statistických metod. Návrh identifikace proměnných byl uskutečněn v programu SIRIUS 4 [2], který na základě přesné hmoty detekovaného iontu a isotopového profilu počítá sumární vzorec. Dále pracuje s naměřeným fragmentačním spektrem a hledá jeho shodu s online databázemi nebo predikuje možnou strukturu pomocí nástroje CANOPUS [3]. V Tabulkách 5-12 jsou pro identifikované proměnné uvedeny naměřené hodnoty m/z a retenční čas, navržený sumární vzorec, typ aduktu, návrh identifikace, chyba hmoty v ppm pro navrženou identifikaci a VIP skóre z odpovídajícího modelu (v případě modelů predikujících odrůdu) nebo parametry jednorozměrných metod t-test, změna střední hodnoty a ROC (v případě modelů predikujících geografický původ). Pro markery je udán stupeň spolehlivosti jejich identifikace na základě stupnice navržené v metabolomických publikacích [4-6]. Znalost identity diagnostických markerů umožní jednodušší převedení analytické metody do rutinní kontrolní praxe, kdy místo měření fingerprintů bude omezen rozsah sledovaných látek pouze na vybrané látky.

Tabulka 5: Návrh identifikace metabolitů použitých pro klasifikaci odrůdy (ESI+)

Experimentální m/z	Experimentální retenční čas [min]	Typ aduktu	Sumární vzorec	Chyba hmoty [ppm]	Návrh identifikace	PLSDA VIP skóre	Spolehlivost identifikace
701.7138	14.03	[M+H] ⁺	C ₄₈ H ₉₂ O ₂	-5.4	Ester mastné kyseliny a alkoholu (18:1/30:1)	3.1	2
317.064	2.06	[M+H] ⁺	C ₁₆ H ₁₂ O ₇	-6.7	Isorhamnetin	2.9	3
673.6829	13.69	[M+H] ⁺	C ₄₆ H ₈₈ O ₂	-5	Ester mastné kyseliny a alkoholu (18:1/28:1)	2.9	2
461.1111	2.14	[M-H] ⁻	C ₂₂ H ₂₂ O ₁₁	5.9	Isorhamnetin rhamnosid	2.8	3
671.6652	13.43	[M+H] ⁺	C ₄₆ H ₈₆ O ₂	-8	Ester mastné kyseliny a alkoholu (46:3)	2.8	3
699.691	14.64	[M+Na] ⁺	C ₄₆ H ₉₂ O ₂	-12.1	Ester mastné kyseliny a alkoholu (46:0)	2.7	3
979.8971	14.81	[M+Na] ⁺	C ₆₃ H ₁₂₀ O ₅	-6.4	TAG (60:2)	2.4	3
509.4234	12.24	[M+H] ⁺	C ₃₁ H ₅₆ O ₅	5.6	DAG (28:2)	2.1	3

Tabulka 6: Návrh identifikace metabolitů použitých pro klasifikaci odrůdy (ESI-)

Experimentální m/z	Experimentální retenční čas [min]	Typ aduktu	Sumární vzorec	Chyba hmoty [ppm]	Návrh identifikace	PLSDA VIP skóre	Spolehlivost identifikace
533.0917	1.33	[M-H] ⁻	C ₂₄ H ₂₂ O ₁₄	2.6	Luteolin-O-malonylglukosid	2.1	3
663.3906	5.99	[M+HC OO] ⁻	C ₃₉ H ₅₄ O ₆	1.3	Caffeoylbutynic acid	2	3
535.4747	12.87	[M-H] ⁻	C ₃₄ H ₆₄ O ₄	3.8	FAHFA (18:1/16:0)	1.9	2
549.3436	3.7	[M+HC OO] ⁻	C ₃₀ H ₄₈ O ₆	1.6	Triterpenová kyselina	1.6	3
749.6105	13.21	[M-H] ⁻	C ₄₉ H ₈₂ O ₅	2.8	DAG (46:7)	1.6	3

Tabulka 7: Návrh identifikace metabolitů použitých pro klasifikaci geografického původu jablek odrůdy Gala (ESI+)

Experimentální m/z	Experimentální retenční čas [min]	Typ aduktu	Sumární vzorec	Chyba a hmoty [ppm]	Návrh identifikace	Parametry jednorozměrné statistiky			Spolehlivost identifikace
						t-test FDR p-hodnota	FC	AUC ROC	
635.6317	13.92	[M+H] ⁺	C ₄₂ H ₈₂ O ₃	-3.9	Ester hydroxy mastné kyseliny a alkoholu (18:1-O/24:0)	1.90E-08	0.78	0.91	2
661.6446	14.14	[M+H] ⁺	C ₄₄ H ₈₄ O ₃	-8	Ester hydroxy mastné kyseliny a alkoholu (18:2-O/26:0)	1.08E-05	0.61	0.9	2
619.634	13.48	[M+H] ⁺	C ₄₂ H ₈₂ O ₂	-8.6	Ester mastné kyseliny a alkoholu (42:1)	3.03E-08	0.86	0.86	3
607.5995	13.51	[M+H] ⁺	C ₄₀ H ₇₈ O ₃	-5.6	Ester hydroxy mastné kyseliny a alkoholu (40:1)	6.40E-07	0.73	0.85	3
637.6466	13.78	[M+H] ⁺	C ₄₂ H ₈₄ O ₃	-5.1	Ester hydroxy mastné kyseliny a alkoholu (16:0-O/26:0)	4.70E-07	1.09	0.84	2
663.6599	14.33	[M+H] ⁺	C ₄₄ H ₈₆ O ₃	-8.5	Ester hydroxy mastné kyseliny a alkoholu (18:1-O/26:0)	8.07E-05	0.6	0.84	2
609.6143	13.39	[M+H] ⁺	C ₄₀ H ₈₀ O ₃	-7	Ester hydroxy mastné kyseliny a alkoholu (16:0-O/24:0)	1.13E-06	1.02	0.83	2
192.0759	1.8	[M+H] ⁺	C ₉ H ₉ N ₃ O ₂	-7.1	Karbendazim	1.10E-06	-8.36	0.82	3
521.524	14.33	[M+H] ⁺	C ₃₅ H ₆₈ O ₂	-11	Ester mastné kyseliny a alkoholu (9:1/26:0)	3.53E-05	0.53	0.82	2
659.6268	13.44	[M+H] ⁺	C ₄₄ H ₈₂ O ₃	-11.2	Ester hydroxy mastné kyseliny a alkoholu (18:3-O/26:0)	1.12E-06	0.68	0.82	2
619.6351	13.77	[M+H] ⁺	C ₄₂ H ₈₂ O ₂	-6.8	Ester mastné kyseliny a alkoholu (16:1/26:0)	2.26E-06	1.04	0.81	2
691.6917	14.73	[M+H] ⁺	C ₄₆ H ₉₀ O ₃	-7.5	Ester hydroxy mastné kyseliny a alkoholu (18:1-O/28:0)	3.51E-04	0.5	0.81	2
631.5949	13.02	[M+H] ⁺	C ₄₂ H ₇₈ O ₃	-12.7	Ester hydroxy mastné kyseliny a alkoholu (18:3-O/24:0)	4.68E-05	0.73	0.81	2
455.3456	5.17	[M+H] ⁺	C ₃₀ H ₄₆ O ₃	-15.2	Oxo ursenová kyselina	6.76E-06	0.42	0.81	3
615.5967	13.74	[M+H] ⁺	C ₄₂ H ₇₈ O ₂	-18.4	Ester mastné kyseliny a alkoholu (18:3/24:0)	1.90E-05	0.74	0.8	2
661.6455	12.91	[M+H] ⁺	C ₄₄ H ₈₄ O ₃	-6.7	Ester hydroxy mastné kyseliny a alkoholu (18:2-O/24:0)	8.11E-06	0.84	0.8	2
265.2493	9.67	[M+H] ⁺	C ₁₈ H ₃₂ O	-14.5	Linolenyl alkohol	8.69E-06	-1.12	0.79	3
287.2342	6.71	[M+H] ⁺	C ₂₀ H ₃₀ O	-11.5	Vitamin A	2.43E-04	-0.89	0.79	3
471.3458	5.61	[M+H] ⁺	C ₃₀ H ₄₆ O ₄	-3.5	Oxohydroxy ursenová kyselina	1.18E-04	0.44	0.79	3

673.6816	14.74	[M+H] +	C ₄₆ H ₈₈ O ₂	-6.9	Ester mastné kyseliny a alkoholu (18:2/28:0)	1.19E-04	0.43	0.79	2
457.3627	6.54	[M+H] +	C ₃₀ H ₄₈ O ₃	-11.9	Monohydroxy ursenoic acid	4.60E-04	0.27	0.79	3
549.5176	11.68	[M+H] +	C ₃₆ H ₆₈ O ₃	-12.9	Ester hydroxy mastné kyseliny a alkoholu (36:2)	7.75E-05	0.41	0.78	3
263.2368	9.13	[M+H] +	C ₁₈ H ₃₀ O	-2.6	Farnesylacetol	9.36E-05	-1.02	0.77	3
318.2987	3.84	[M+H] +	C ₁₈ H ₃₉ NO ₃	-6.7	Phytosphingosine	3.41E-05	0.75	0.77	3
453.3358	5.61	[M+H] +	C ₃₀ H ₄₄ O ₃	-2.4	Oxodehydroursolová kyselina	2.13E-04	0.42	0.77	3
469.3625	6.73	[M+H] +	C ₃₁ H ₄₈ O ₃	-12.1	Methyl oxooleanenová kyselina	6.82E-04	0.34	0.77	3
615.6071	13.19	[M+H] +	C ₄₂ H ₇₈ O ₂	-1.5	Ester mastné kyseliny a alkoholu (18:3/24:0)	9.98E-05	0.87	0.77	2
671.6644	13.95	[M+H] +	C ₄₆ H ₈₆ O ₂	-9.2	Ester mastné kyseliny a alkoholu (18:3/28:0)	5.31E-05	0.71	0.77	2
673.6794	14.02	[M+H] +	C ₄₆ H ₈₈ O ₂	-10.1	Ester mastné kyseliny a alkoholu (18:2/28:0)	4.17E-04	0.56	0.77	2
703.6281	15.83	[M+H] +	C ₄₅ H ₈₂ O ₅	5.7	DAG (42:3)	6.02E-05	1.13	0.77	3
379.0965	1.34	[M+Na] +	C ₁₆ H ₂₀ O ₉	-10.6	Feruloyl glucose	1.02E-04	1.11	0.76	3
475.1182	1.59	[M+Na] +	C ₂₁ H ₂₄ O ₁₁	-7.2	Sieboldin	8.92E-04	0.6	0.76	3
525.3529	6.44	[M+Na] +	C ₃₁ H ₅₀ O ₅	-5.1	Methyl ester euscaphové kyseliny	9.20E-05	0.85	0.76	3
641.6169	15.19	[M+Na] +	C ₄₂ H ₈₂ O ₂	-6.7	Ester mastné kyseliny a alkoholu (42:1)	4.38E-05	0.35	0.76	3
669.6464	13.82	[M+H] +	C ₄₆ H ₈₄ O ₂	-12.7	Ester mastné kyseliny a alkoholu (18:4/28:0)	3.92E-04	0.62	0.76	2
683.6628	15.97	[M+Na] +	C ₄₅ H ₈₈ O ₂	-7.9	Ester mastné kyseliny a alkoholu (45:1)	2.78E-04	0.76	0.76	3
471.3437	6.56	[M+H] +	C ₃₀ H ₄₆ O ₄	-7.9	Oxohydroxy ursenová kyselina	1.75E-04	0.25	0.75	3
453.3329	3.93	[M+H] +	C ₃₀ H ₄₄ O ₃	-8.8	Oxodehydroursolová kyselina	3.66E-04	0.63	0.75	3
469.3327	5.91	[M+H] +	C ₃₀ H ₄₄ O ₄	1.9	Dioxohydroxy ursenová kyselina	4.58E-04	0.75	0.75	3
633.6124	13.74	[M+H] +	C ₄₂ H ₈₀ O ₃	-9.7	Ester hydroxy mastné kyseliny a alkoholu (18:2-O/24:0)	2.92E-05	0.79	0.75	3

Tabulka 8: Návrh identifikace metabolitů použitých pro klasifikaci geografického původu jablek odrůdy Golden Delicious (ESI+)

Experimentální m/z	Experimentální retenční čas [min]	Typ aduktu	Sumární vzorec	Chyba hmoty [ppm]	Návrh identifikace	Parametry jednorozměrné statistiky			Spolehlivost identifikace
						t-test FDR p-hodnota	FC	AUC ROC	
671.6644	13.95	[M+H] ⁺	C ₄₆ H ₈₆ O ₂	-9.2	Ester mastné kyseliny a alkoholu (18:3/28:0)	1.30E-11	0.78	0.91	2
661.6455	12.91	[M+H] ⁺	C ₄₄ H ₈₄ O ₃	-6.7	Ester hydroxy mastné kyseliny a alkoholu (18:2-O/24:0)	5.98E-10	1.08	0.9	2
661.6455	13.58	[M+H] ⁺	C ₄₄ H ₈₄ O ₃	-6.6	Ester hydroxy mastné kyseliny a alkoholu (18:2-O/24:0)	2.78E-10	1.1	0.9	2
659.6295	12.78	[M+H] ⁺	C ₄₄ H ₈₂ O ₃	-7.1	Ester hydroxy mastné kyseliny a alkoholu (18:3-O/26:0)	4.99E-08	0.94	0.88	2
691.6893	14.19	[M+H] ⁺	C ₄₆ H ₈₉ O ₃	-10.8	Ester hydroxy mastné kyseliny a alkoholu (18:1-O/28:0)	2.92E-10	0.81	0.88	2
631.6307	14.14	[M+H] ⁺	C ₄₃ H ₈₂ O ₂	-13.7	Ester mastné kyseliny a alkoholu (43:2)	1.50E-08	0.68	0.87	3
643.6345	13.58	[M+H] ⁺	C ₄₄ H ₈₂ O ₂	-7.5	Ester mastné kyseliny a alkoholu (18:3/26:0)	4.11E-09	1.01	0.87	3
459.3801	5.89	[M+H] ⁺	C ₃₀ H ₅₀ O ₃	-8	Hydroxybetulin	1.91E-07	0.95	0.86	3
703.6281	15.83	[M+H] ⁺	C ₄₅ H ₈₂ O ₅	5.7	DAG (42:3)	7.21E-09	0.97	0.86	3
669.5425	12.12	[M+Na] ⁺	C ₄₁ H ₇₄ O ₅	-1.4	DAG (38:3)	2.96E-08	1.3	0.86	3
615.6071	13.19	[M+H] ⁺	C ₄₂ H ₇₈ O ₂	-1.4	Ester mastné kyseliny a alkoholu (18:3/24:0)	1.72E-07	1.24	0.85	2
647.557	12.13	[M+Na] ⁺	C ₃₉ H ₇₆ O ₅	-3.1	DAG (36:0)	1.68E-06	1.26	0.85	3
163.0745	10.88	[M+H] ⁺	C ₁₀ H ₁₀ O ₂	-8.9	Methyl cinnamate	7.18E-08	-1.63	0.84	3
183.0862	1.15	[M+H] ⁺	C ₆ H ₁₄ O ₆	-3.6	Sorbitol	9.54E-08	1.19	0.84	3
605.5795	11.99	[M+H] ⁺	C ₄₀ H ₇₆ O ₃	-12.8	Ester hydroxy mastné kyseliny a alkoholu (18:2-O/22:0)	5.83E-06	0.94	0.84	2
133.0628	12.05	[M+H] ⁺	C ₉ H ₈ O	-19.3	Cinnamaldehyde	1.06E-06	0.67	0.83	3
425.3765	12.47	[M+H] ⁺	C ₃₀ H ₄₈ O	-4.4	Lupeone	1.90E-07	-1.13	0.83	3
451.4453	14.44	[M+H] ⁺	C ₃₀ H ₅₈ O ₂	-13.9	Triacantenová kyselina	2.38E-06	-1.73	0.83	3
487.4499	12.11	[M+H] ⁺	C ₃₃ H ₅₈ O ₂	-3.3	Ester mastné kyseliny a alkoholu (33:4)	1.12E-06	-0.96	0.83	3
165.0757	1.15	[M+H] ⁺	C ₆ H ₁₂ O ₅	-3.9	Fukóza	6.22E-08	1.27	0.83	3
659.6288	13.76	[M+H] ⁺	C ₄₄ H ₈₂ O ₃	-8.2	Ester hydroxy mastné kyseliny a	2.26E-07	0.55	0.83	2

					alkoholu (18:3-O/26:0)				
663.6599	14.33	[M+H] ⁺	C44H86O3	-8.5	Ester hydroxy mastné kyseliny a alkoholu (18:1-O/26:0)	6.63E-07	0.57	0.83	2
665.6752	14.16	[M+H] ⁺	C44H88O3	-9	Ester hydroxy mastné kyseliny a alkoholu (44:0)	4.72E-07	0.64	0.83	3
691.6917	14.73	[M+H] ⁺	C46H89O3	-7.5	Ester hydroxy mastné kyseliny a alkoholu (18:1-O/28:0)	2.32E-06	0.27	0.83	2
711.585	11.77	[M+Na] ⁺	C44H80O5	-7.6	DAG (41:3)	5.56E-07	-1.38	0.83	3
445.3968	13.46	[M+H] ⁺	C30H52O2	-17.4	Ester mastné kyseliny a alkoholu (30:4)	1.43E-06	-1.57	0.82	3
659.6268	13.44	[M+H] ⁺	C44H82O3	-11.2	Ester hydroxy mastné kyseliny a alkoholu (18:3-O/26:0)	1.87E-06	0.44	0.82	2
265.2524	11.32	[M+H] ⁺	C18H32O	-2.6	Linolenyl alkohol	3.45E-06	-1.17	0.81	3
457.3661	7.28	[M+H] ⁺	C30H48O3	-4.6	Monohydroxy ursenoic acid	1.51E-06	0.1	0.81	3
635.6317	13.92	[M+H] ⁺	C42H82O3	-3.9	Ester hydroxy mastné kyseliny a alkoholu (18:1-O/24:0)	5.44E-06	0.28	0.81	2
673.6816	14.74	[M+H] ⁺	C46H88O2	-6.9	Ester mastné kyseliny a alkoholu (18:2/28:0)	7.60E-06	0.2	0.81	2
675.5843	12.6	[M+Na] ⁺	C41H80O5	-9	DAG (38:0)	8.74E-07	1.14	0.81	3
853.7221	15.53	[M+Na] ⁺	C53H98O6	-4.8	TAG (40:2)	3.75E-07	-1.41	0.81	3
875.7018	14.04	[M+Na] ⁺	C55H96O6	-9.9	TAG (52:5)	2.86E-05	0.66	0.81	3
877.7224	15.24	[M+Na] ⁺	C55H98O6	-4.2	TAG (52:4)	1.48E-06	-1.65	0.81	3
263.2368	9.13	[M+H] ⁺	C18H30O	-2.8	Farnesylacetol	1.04E-04	-1.31	0.8	3
133.0642	10.42	[M+H] ⁺	C9H8O	-8.9	Cinnamaldehyde	2.39E-05	-1.27	0.8	3
193.1605	9.13	[M+H] ⁺	C13H20O	6.6	Damascone	2.07E-06	-1.37	0.8	3
443.3858	13.44	[M+H] ⁺	C30H50O2	-7.1	Uvaol	7.36E-06	-1.27	0.79	3
471.3458	5.61	[M+H] ⁺	C30H46O4	-3.5	Hydroxyoxooleane nová kyselina	4.97E-05	0.37	0.79	3
643.6299	14.14	[M+H] ⁺	C44H82O2	-14.6	Ester mastné kyseliny a alkoholu (18:3/26:0)	1.34E-05	0.18	0.79	2
427.3927	7.83	[M+H] ⁺	C30H50O	-3.1	Amyrin	2.32E-05	0.44	0.78	3
445.3971	10.15	[M+H] ⁺	C30H52O2	-16.8	Ester mastné kyseliny a alkoholu (30:4)	9.40E-03	-2.15	0.78	3
445.3997	13.72	[M+H] ⁺	C30H52O2	-10.9	Ester mastné kyseliny a alkoholu (30:4)	2.48E-05	-1.24	0.78	3
453.3358	5.61	[M+H] ⁺	C30H44O3	-2.4	Oxodehydroursolov á kyselina	8.40E-05	0.37	0.78	3
593.5082	10.9	[M+H] ⁺	C37H68O5	-10.6	DAG (34:2)	2.11E-05	-1.63	0.78	3
631.5949	13.02	[M+H] ⁺	C42H78O3	-12.7	Ester hydroxy mastné kyseliny a alkoholu (18:3-O/24:0)	3.94E-05	0.65	0.78	2

661.6446	14.14	[M+H] ⁺	C44H84O3	-8	Ester hydroxy mastné kyseliny a alkoholu (18:2- O/26:0)	3.26E-05	0.19	0.78	2
635.4605	10.35	[M+Na] ⁺	C39H64O5	-7.4	DAG (36:6)	3.96E-05	1.86	0.78	3
685.5684	13.44	[M+Na] ⁺	C42H78O5	-9.2	DAG (38:2)	9.87E-05	-1.23	0.78	3
275.0891	1.8	[M+H] ⁺	C15H14O5	-10.4	Phloretin	3.23E-05	0.43	0.77	3
443.3849	7.64	[M+H] ⁺	C30H50O2	-9.1	Ester mastné kyseliny a alkoholu (30:5)	3.72E-05	0.32	0.77	3
543.4717	12.4	[M+H] ⁺	C36H62O3	-11	Ester hydroxy mastné kyseliny a alkoholu (36:5)	2.23E-05	-1.17	0.77	3
609.6143	13.39	[M+H] ⁺	C40H80O3	-7	Ester hydroxy mastné kyseliny a alkoholu (16:0-O- 24:0)	6.61E-05	0.55	0.77	3
619.6351	13.77	[M+H] ⁺	C42H82O2	-6.7	Ester mastné kyseliny a alkoholu (16:1/26:0)	1.40E-05	0.59	0.77	2
755.6085	12.28	[M+H] ⁺	C43H83N2 O6P	2.4	SM (20:2)	2.02E-05	-1	0.77	3
615.4914	11.35	[M+Na] ⁺	C37H68O5	-8.2	DAG (34:2)	5.68E-05	0.63	0.77	3
699.5872	12.84	[M+Na] ⁺	C43H80O5	-4.6	DAG (40:2)	2.73E-04	0.7	0.77	3
447.4142	10.22	[M+H] ⁺	C30H54O2	-13.5	Ester mastné kyseliny a alkoholu (30:3)	7.10E-05	-1.29	0.76	3
455.3494	5.84	[M+H] ⁺	C30H46O3	-6.9	Oxooleanolová kyselina	2.98E-05	0.04	0.76	3
663.5863	12.23	[M+H] ⁺	C42H78O5	-9.7	DAG (38:2)	1.47E-04	-1.35	0.76	3
495.3431	5.41	[M+Na] ⁺	C30H48O4	-3.9	Hydroxyursolová kyselina	4.54E-05	0.24	0.76	3
650.5706	11.22	[M+NH4]] ⁺	C40H72O5	-2.7	DAG (36:3)	2.52E-05	-1.36	0.76	3
198.0964	1.15	[M+NH4]] ⁺	C6H12O6	-6.7	Manóza	2.32E-04	1.1	0.76	3
631.5981	13.29	[M+H] ⁺	C42H78O3	-7.7	Ester hydroxy mastné kyseliny a alkoholu (18:3- O/24:0)	2.59E-03	0.21	0.75	2
613.4818	8.72	[M+Na] ⁺	C37H66O5	1.7	DAG (34:3)	2.75E-04	0.81	0.75	3
685.5693	10.25	[M+Na] ⁺	C42H78O5	-7.9	DAG (38:2)	1.04E-04	-1.12	0.75	3
851.7066	13.8	[M+Na] ⁺	C53H96O6	-4.5	TAG (40:3)	3.96E-05	-0.93	0.75	3

Tabulka 9: Návrh identifikace metabolitů použitých pro klasifikaci geografického původu jablek odrůdy Idared (ESI⁺)

Experimentální m/z	Experimentální retenční čas [min]	Typ aduktu	Sumární vzorec	Chyba hmoty [ppm]	Návrh identifikace	Parametry jednorozměrné statistiky			Spolehlivost identifikace
						t-test FDR p- hodnota	FC	AUC ROC	
192.0759	1.8	[M+H] ⁺	C9H9N3O2	-7.3	Karbendazim	4.65E-08	-4.45	0.92	3
167.0709	1.74	[M+H] ⁺	C9H10O3	0.5	Ethyl salicylát	1.37E-05	-2.82	0.82	3
561.1335	1.45	[M+H] ⁺	C30H24O11	-11	Biflavonoid	1.31E-04	1.69	0.78	3

Tabulka 10: Návrh identifikace metabolitů použitých pro klasifikaci geografického původu jablek odrůdy Gala (ESI-)

Experimentální m/z	Experimentální retenční čas [min]	Typ aduktu	Sumární vzorec	Chyba hmoty [ppm]	Návrh identifikace	Parametry jednorozměrné statistiky			Spolehlivost identifikace
						t-test FDR p-hodnota	FC	AUC ROC	
135.0301	1.22	[M-H]-	C ₄ H ₈ O ₅	5.6	Trihydroxybutanová kyselina	1.05E-08	-2.83	0.87	3

Tabulka 11: Návrh identifikace metabolitů použitých pro klasifikaci geografického původu jablek odrůdy Golden Delicious (ESI-)

Experimentální m/z	Experimentální retenční čas [min]	Typ aduktu	Sumární vzorec	Chyba hmoty [ppm]	Návrh identifikace	Parametry jednorozměrné statistiky			Spolehlivost identifikace
						t-test FDR p-hodnota	FC	AUC ROC	
131.0465	1.16	[M-H]-	C ₄ H ₈ N ₂ O ₃	6.3	Asparagin	8.20E-08	3.85	0.84	3
181.0744	1.15	[M-H]-	C ₆ H ₁₄ O ₆	17.6	Galakticol	5.16E-07	0.83	0.81	3
133.0139	1.25	[M-H]-	C ₄ H ₆ O ₅	1.5	Kyselina jablečná	3.65E-05	2.11	0.77	3
209.0666	1.2	[M-H]-	C ₇ H ₁₄ O ₇	2.2	Heptulose	5.58E-05	1.32	0.77	3

Tabulka 12: Návrh identifikace metabolitů použitých pro klasifikaci geografického původu jablek odrůdy Idared (ESI-)

Experimentální m/z	Experimentální retenční čas [min]	Typ aduktu	Sumární vzorec	Chyba hmoty [ppm]	Návrh identifikace	Parametry jednorozměrné statistiky			Spolehlivost identifikace
						t-test FDR p-hodnota	FC	AUC ROC	
225.0587	1.15	[M+HCOO]-	C ₆ H ₁₂ O ₆	10.4	Manóza	1.04E-05	1.25	0.82	3
209.0666	1.2	[M-H]-	C ₇ H ₁₄ O ₇	2.2	Heptulóza	4.21E-05	1.18	0.78	3
353.0853	1.48	[M-H]-	C ₁₆ H ₁₈ O ₉	5.6	Chlorogenová kyselina	3.18E-04	1.06	0.76	3
131.0465	1.16	[M-H]-	C ₄ H ₈ N ₂ O ₃	6.2	Asparagin	4.54E-04	1.15	0.75	3

2.9 Pracovní charakteristiky metody

Všechny klasifikační modely byly validovány pomocí sedminásobné křížové validace a jsou charakterizovány popsáním (R^2Y) a predikovaným (Q^2Y) rozptylem. Dosažené hodnoty R^2Y a Q^2Y jsou uvedeny a komentovány u jednotlivých modelů. *Za validní model je považován takový, jehož hodnota Q^2Y nabývá minimálně hodnoty 0,5.*

Chyby hmoty v ppm pro navrženou identifikaci jednotlivých metabolitů použitých pro klasifikaci jsou uváděny v tabulkách 6 – 13.

Pro kontrolu stability a opakovatelnosti měření byl využit 'quality control' (QC) vzorek, který byl nástříknut před začátkem sekvence v 8 opakováních, na konci sekvence v 6 opakováních a dále byl nastřikován v pravidelných odstupech každých 17 vzorků v průběhu celé sekvence. Jedním z prvních kritérií při filtraci proměnných byla použita i hodnota relativní směrodatné odchylky (RSD) proměnných v QC vzorcích z celé sekvence, čímž byly odstraněny nekonzistentní proměnné.

Výběr stabilních signálů, které tvoří výsledné klasifikační modely, byl zaručen díky zmíněné filtraci na RSD proměnných v QC vzorcích a dále použitím kombinace jednorozměrných statistických metod (Receiver Operating Characteristics, t-test, změna střední hodnoty).

Robustnost metody je zaručena normalizací extrahovaných složek povrchové epikutikulární vrstvy jablek na uniformní koncentraci.

Robustnost klasifikačních modelů pro určení země původu byla ověřena na základě vložení vzorků získaných v roce 2021 do modelu vytvořeného na základě vzorků z let 2020 a 2022. Všechny testované modely vykazaly predikční schopnost vyšší než 65 %.

3 Srovnání novosti postupů

Předkládaná metodika vychází z metabolomické analýzy resp. strategie metabolomického fingerprintingu, tedy využití necílové analýzy, která umožňuje provést rychlé vyšetření vzorků pro autentikaci geografického původu a odrůdy. Vyvinutá metodika přináší nový způsob analýzy pomocí „fingerprintingu“ celkového metabolomu s využitím ultra-účinné kapalinové chromatografie spojené s vysokorozlišovací hmotnostní spektrometrií (U-HPLC–HRMS) s ionizací elektrosprejem. Tento přístup je zcela unikátní a nebyl doposud publikován. Je ale třeba zdůraznit potřebu dalšího zvyšování robustnosti chemometrických modelů pomocí rozšiřování počtu analyzovaných autentických vzorků jablek používaných k jejich konstrukci.

Předností komplexního necílového fingerprintingu frakcí extraktů metodou UHPLC-HRMS je především množství informací obsažených v hmotnostních spektrech detekovaných metabolitů (veškeré ionizovatelné sloučeniny v oblasti molekulových hmot od 100 do 1200 Da). Metodika poskytuje spolehlivou kvalitativní i kvantitativní informaci a je robustní. Tato inovativní metodika analýzy metabolomu jablek umožňuje také retrospektivní prohledávání dat a dodatečnou identifikaci původně necílené sloučeniny, což konvenční rutinní metody cílové analýzy neumožňují. Experimentální data byla zpracována pomocí námi nově vypracovaných skriptů v programovacím jazyku R. Tato strategie umožnila zkvalitnit interpretaci experimentálních dat a tedy i využitelnost získaných výsledků. V rámci rutinní autentikace by ale mohla být analýza jednodušší, stačilo by cíleně vyhledávat jen zjištěné markery.

Jedinečnost a originalita řešení spočívá zejména ve vytvoření unikátní databáze metabolomických 'fingerprintů' jablek různého geografického původu, která může být použita k ověření jejich autenticity. Opakované vstupy v různých letech sklizně umožnily podchycení variability metabolomu v důsledku meziročního kolísání klimatických podmínek, agrotechniky i podmínek skladování. Pomocí validovaných statistických modelů bylo možno identifikovat některé analytické markery pro určení země původu. Vývoj, standardizace a implementace diagnostických postupů, umožňujících rozlišit falešně deklarované dodávky jablek na trhu z hlediska původu, je důležitým předpokladem pro zvýšení účinnosti státního dozoru a zlepšení tržního prostředí ve vztahu k českým producentům. V dostupné odborné literatuře jsme dosud nenašli publikaci týkající se laboratorního prokazování falešné deklarace zeměpisného původu jablek.

4 Popis uplatnění metodiky

Parametry a pracovní charakteristiky uvedené analytické metody, která je založená na principu U-HPLC-HRMS, jsou vhodné pro získání rychlé kvalitativní informace o autenticitě jablek, tj. ověření jejich deklarovaného geografického původu a odrůdové pravosti.

Metodika byla vyvinuta v rámci řešení projektu **QK1910104 „Výzkum metabolomických metod pro laboratorní ověření geografické autenticity jablek“**. Aplikace metody navržené v rámci tohoto projektu významně přispěje ke screeningu či potvrzování původu a odrůdy jablek uváděných na trh.

Metodika je určena zejména pro laboratoře úřední kontroly potravin a zemědělských produktů, vědeckovýzkumné organizace nebo privátní kontrolní laboratoře zabývající se prokazováním autenticity produktů, v obecném pojetí je pak metodika využitelná širokou odbornou veřejností. Mohou ji využít zejména instituce jako Ústřední kontrolní a zkušební ústav zemědělský (ÚKZÚZ) nebo Státní zemědělská a potravinářská inspekce (SZPI) či komerční laboratoře.

5 Ekonomické aspekty

Ekonomicky významným přínosem je zvýšení účinnosti a spolehlivosti kontroly dodržování nařízení (EU) č. 1169/2011, nařízení (ES) č. 178/2002 ale zejména pak nařízení (EU) č. 2023/2429, které vyžaduje uvádět u čerstvých jablek údaj o zemi původu a odrůdě. Dalším přínosem je zvýšení důvěry veřejnosti k informacím o potravinách uváděných na český trh. To pak má nepochybně velký význam pro podporu českých producentů jablek, respektive zvýšení jejich konkurenceschopnosti. Ekonomické přínosy očekávané v této oblasti nelze však přesně kvantifikovat.

Jednou z výhod výše popsané metodiky analýzy jablek pro účely získání komplexních informací o jejich metabolomu je využití ultra-vysokoúčinné kapalinové chromatografie, která má oproti konvenční HPLC vyšší separační potenciál, umožňuje významné zkrácení doby analýzy, ale i snížení spotřeby rozpouštědel. To je podstatné nejen pro snížení nákladů na rozpouštědla ale také z hlediska ochrany životního prostředí. Samotná příprava vzorků k měření je zde velmi jednoduchá, rychlá a levná.

Při bilanci ekonomických aspektů bylo předpokládáno, že laboratoř, která chce uvedenou metodu zavést, již má zkušenosti s analýzou v systémech U-HPLC-HRMS a vlastní již také kapalinový chromatograf s vysokorozlišovacím hmotnostním spektrometrem. Implementace metody zahrnuje zjištění aktuálních pracovních charakteristik, zejména zjištění citlivosti a opakovatelnosti metody. Za předpokladu, že by bylo provedeno 20 verifikačních a kontrolních analýz, přičemž provozní náklady na jednu analýzu by byly 1500 Kč, celková výše prostředků na implementaci lze odhadnout na asi 30 tis. Kč.

6 Seznam použité související literatury

- [1] Triba, M. N., Le Moyec, L., Amathieu, R., Goossens, C., Bouchemal, N., Nahon, P., Rutledge, D. N., & Savarin, P. (2015). PLS/OPLS models in metabolomics: The impact of permutation of dataset rows on the K-fold cross-validation quality parameters. *Molecular BioSystems*, 11(1), 13–19. [<https://doi.org/10.1039/c4mb00414k>]
- [2] Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., Dorrestein, P. C., Rousu, J., & Böcker, S. (2019). SIRIUS 4: A rapid tool for turning tandem mass spectra into metabolite structure information. *Nature Methods*, 16(4), 4. [<https://doi.org/10.1038/s41592-019-0344-8>]
- [3] Kim, H. W., Wang, M., Leber, C. A., Nothias, L.-F., Reher, R., Kang, K. B., van der Hooft, J. J. J., Dorrestein, P. C., Gerwick, W. H., & Cottrell, G. W. (2021). NPClassifier: A Deep Neural Network-Based Structural Classification Tool for Natural Products. *Journal of Natural Products*, 84(11), 2795–2807. [<https://doi.org/10.1021/acs.jnatprod.1c00399>]
- [4] Blaženović, I., Kind, T., Ji, J., & Fiehn, O. (2018). Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics. *Metabolites*, 8(2), 31. [[doi:10.3390/metabo8020031](https://doi.org/10.3390/metabo8020031)]
- [5] Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W.-M., Fiehn, O., Goodacre, R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J. C., Marriott, P., Nicholls, A. W., ... Viant, M. R. (2007). Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics: Official Journal of the Metabolomic Society*, 3(3), 211–221. [<https://doi.org/10.1007/s11306-007-0082-2>]
- [6] Reisdorph, N. A., Walmsley, S., & Reisdorph, R. (2019). A Perspective and Framework for Developing Sample Type Specific Databases for LC/MS-Based Clinical Metabolomics. *Metabolites*, 10(1), 8. [<https://doi.org/10.3390/metabo10010008>]

7 Seznam publikací, které předcházely metodice

- Hajšlová J., Kocourek V., Sedlák J., Vacková P., Navrátilová K., Hůrková K., Hrbek V.: Identifikace země původu jablek pomocí chemických analýz. Vědecké práce ovocnářské 27(1): 25-38 (2021). ISSN 2695-1347 (online).
- Rektorisová M., Hajšlová J., Kocourek V., Sedlák J., Vacková P., Navrátilová K., Hůrková K., Hrbek V.: Využití analýzy metabolomických profilů fenolických látek jablek pro jejich autentikaci. Vědecké práce ovocnářské 27(2): 25-37 (2021). ISSN 2695-1347 (online).
- Bechynska K., Rektorisova M., Navratilova K., Sedlak J., Uttl L., Kosek V., Kocourek V., Hajšlova, J.: Metabolomics on apples (*Malus domestica*) cuticle - search for authenticity markers. *Manuscript připraven pro Journal of Food Composition and Analysis*.

Autoři metodiky:

Ing. Kamila Bechyňská	podíl 20 %
Ing. Michaela Rektorisová, Ph.D.	podíl 10 %
Ing. Petra Vacková	podíl 10 %
prof. Ing. Vladimír Kocourek, CSc.	podíl 30 %
prof. Ing. Jana Hajšlová, CSc.	podíl 30 %

Vysoká škola chemicko-technologická v Praze
Ústav analýzy potravin a výživy
Technická 1905/5
166 28 Praha 6 - Dejvice

Jména oponentů:

Prof. Ing. Lenka Kouřimská, Ph.D.

Fakulta agrobiologie, potravinových a přírodních zdrojů, Česká zemědělská univerzita, Praha

Ing. Šárka Hanzelková, Ph.D.

Státní zemědělská a potravinářská inspekce, Brno

Dedikace:

Metodika vznikla za finanční podpory Ministerstva zemědělství České republiky a je výstupem řešení projektu NAZV QK1910104 s názvem „**Výzkum metabolomických metod pro laboratorní ověření geografické autenticity jablek**“.

Metody metabolomického profilování a stanovení analytických markerů autenticity jablek

Vydavatel:	Vysoká škola chemicko-technologická v Praze, Technická 5, 166 28 Praha 6
Oponenti:	Prof. Ing. Lenka Kouřimská, Ph.D., Ing. Šárka Hanzelková, Ph.D.
Rok vydání:	2023
Počet stran:	36
Vydání:	první