

TECHNIKY METABOLOMIKY V BIOMEDICÍNĚ

**PETR WOJTOWICZ, HANA JANEČKOVÁ,
DAVID FRIEDECKÝ a TOMÁŠ ADAM**

*Laboratoř dědičných metabolických poruch a Ústav molekulární a translační medicíny, LF UP v Olomouci,
I. P. Pavlova 6, 775 20 Olomouc
petr.wojtowicz@gmail.com, janeckovah@gmail.com,
david.friedecky@gmail.com, tomasadam@gmail.com*

Došlo 2.4.12, přijato 18.6.12.

Klíčová slova: metabolomika, plynová chromatografie, kapalinová chromatografie, kapilární elektroforéza, hmotnostní spektrometrie, zpracování dat

Obsah

1. Úvod
2. Příprava vzorků pro metabolomickou analýzu
3. Metody metabolomické analýzy
 - 3.1. Plynová chromatografie
 - 3.2. Kapalinová chromatografie
 - 3.3. Kapilární elektroforéza
4. Analýza metabolomických dat
 - 4.1. Předúprava dat
 - 4.2. Statistická analýza
5. Závěr

1. Úvod

Metabolomika je termín používaný pro analýzu velkého souboru nízkomolekulárních látek, které se v daném čase za daného stavu nacházejí v biologickém vzorku (buňka, tkáň, organismus), tzv. metabolomu. Toto odvětví je analogií dalších „omik“ – genomiky, transkriptomiky a proteomiky, ovšem od předešlých se v mnoha ohledech liší. V případě člověka zahrnuje metabolom mnohem menší počet jednotlivých chemických individuí (cca 2500 metabolitů vs. cca 32 000 genů vs. cca 1 000 000 proteinů¹), které však patří do velkého počtu chemických tříd, pro jejich analýzu je tedy nutno použít více různých technik. Ačkoliv metabolom závisí na genotypu, je velice významně ovlivněn aktuálním stavem systému a je dynamický. Informace, které poskytuje, jsou k ostatním „omikám“ komplementární. Například aktuální metabolický stav člověka je ovlivněn celou řadou vnitřních (genotyp, věk, zdravotní stav, denní cyklus apod.) i vnějších (fyzická aktivita, výživa, léky, mikrobiom aj.) faktorů, které však

mají na jeho genovou expresi malý vliv¹. Jinými slovy lze s jistotou nadsázkou říci, že genom říká, co by se v organismu mohlo stát a metabolom co se právě děje.

Metabolomika jako směr vědeckého bádání zahrnuje poměrně širokou oblast ke zkoumání, vyplývající z biologické rozmanitosti naší planety – od experimentů mikrobiologických přes rostlinné a živočišné až po analýzy lidských metabolitů. Z hlediska pracovního postupu pak zahrnuje návrh experimentu, odběr a přípravu vzorku, analytickou část, zpracování dat a jejich interpretaci, popřípadě archivaci a revizi.

Existuje několik přístupů ke zkoumání metabolomu a tyto varianty mají každá svou vlastní metodiku práce. Jejich přehled je uveden v tab. I.

V současnosti je metabolomika velmi dynamicky se rozvíjející odvětví s širokým uplatněním v mnoha chemických, biologických a lékařských odvětvích. Rostoucí význam metabolomiky v systému přírodních věd podnítilo vydávání odborného časopisu *Metabolomics* (Springer Science+Business Media, Inc.). První číslo vyšlo v lednu roku 2005 (IF₂₀₁₀= 3,608). V tomtéž roce byla ustanovena Iniciativa MSI (Metabolomics Standards Initiative), která v rámci pěti pracovních skupin sdružuje osobnosti aktivní v příslušné oblasti³: I. Biologický kontext metadat (nezbytné informace, kterými je nutno opatřit biologický vzorek), II. Chemická analýza⁴, III. Procesování dat⁵, IV. Ontologie⁶, V. Výměna dat⁷ (transfer, sdílení, archivace). Jejím hlavním úkolem je standardizovat pravidla pro provedení metabolomického experimentu s cílem umožnit přesné zopakování práce a kompatibilitu výsledků mezi jednotlivými pracovišti. Také používaná terminologie stále ještě není ustálená a pro zavedené anglické termíny navíc mnohdy neexistuje ekvivalentní český překlad, autoři v tomto případě volí původní anglické termíny.

V tomto referátu jsou diskutovány jednotlivé části metabolomického experimentu s důrazem na biomedicínské aplikace a v analytické části je popsáno spojení separačních technik s hmotnostní spektrometrií.

2. Příprava vzorku pro metabolomickou analýzu

Významným faktorem úspěchu každého analytického procesu je kvalita přípravy vzorku, která by měla být co nejvíce jednoduchá a univerzální, ale především minimálně destruktivní k analytům⁸. Zkoumané analyty se extrahují z nejrůznějších komplexních biologických matic, jako je např. sérum, plazma, plná krev, mozkomíšni mok, moč, sliny, tkáňové homogenáty, buněčné pelety či buněčná média. Při extrakci je nutno převést biologický materiál na formu, která bude kompatibilní s analytickou technikou a zároveň dojde k odstranění matričních složek, které by v analýze interferovaly (solí, peptidy či proteiny).

Tabulka I
Definice strategií používaných v metabolomických experimentech^a

Pojem	Vysvětlení
Metabolomika	Identifikace a kvantifikace všech metabolitů v biologickém systému (metabolomu). V praxi je tento cíl v současnosti nedosažitelný, neboť neexistuje analytická technika, která by dokázala tento úkol splnit. Nutností je užití vhodné kombinace vysoce selektivních a senzitivních analytických technik.
Metabonomika	Kvantitativní měření dynamické metabolické reakce živého organismu na podnět (fyziologická stimulace, genetická modifikace).
Targetovaná (cílená) metabolomika	Stanovení jednoho či více konkrétních metabolitů souvisejících se specifickou metabolickou drahou. Součástí bývá mnohdy složitá příprava vzorku z biologické matrice a užití separační techniky s citlivou detekcí.
Metabolické profilování	Identifikace a kvantifikace metabolitů majících podobnou chemickou strukturu (aminokyseliny) nebo patřících k metabolické dráze (citrátový cyklus). Obvykle zahrnuje vysoceúčinnou separační techniku.
Metabolický fingerprinting	Komplexní analýza vzorku nebo extraktu bez identifikace a kvantifikace jednotlivých metabolitů. Minimální příprava vzorku. Může sloužit například ke klasifikaci vzorků nebo ke screeningovým účelům.
Metabolický footprinting	Komplexní analýza extracelulárních metabolitů (vylučovaných do růstového média). Součástí je biotransformace komponent média.

^a Modifikováno z cit.²

Důležitým biologickým materiálem pro výzkum a diagnostiku jsou kultivované buňky. Metabolom biotekutiny (krev, moč) je ovlivňován mnoha faktory, jako jsou věk, nemoc, strava aj., tento vliv je minimalizován na buňky v kultuře, které jsou pěstovány v kontrolovaném a přesně definovaném prostředí. Metabolity jsou obecně labilní, chemicky velmi různorodé látky, které se navíc vyskytují ve velkém rozsahu koncentrací. Dalším úskalím je fakt, že metabolické reakce probíhají v řádu sekund, tedy mnohem rychleji, než je doba syntézy či degradace velkých molekul. Velmi důležitou součástí analýzy buněčného materiálu je tedy velmi rychlé zastavení metabolismu – tzv. metabolické zhášení (quenching). Je nežádoucí, aby buňka na podnět vedoucí k zastavení metabolismu zareagovala změnou koncentrace metabolitů nebo je ztrácela narušením buněčné stěny (metabolite leakage). Pro různé typy vzorků byly vyvinuty různé metody metabolického zhášení^{9–11}.

Po metabolickém zhášení následuje extrakce metabolitů z biologického materiálu. Výběr extrakčního činidla je závislý na chemické povaze metabolitů. Postup, jak zcela extrahovat všechny typy látek v jednom extrakčním kroku je, z důvodu již zmiňované rozmanitosti, v současnosti neznámý. Avšak existují mnohá kompromisní řešení. Pro analýzu intracelulárního obsahu je nutno nejprve narušit buněčné stěny, čehož lze dosáhnout chemicky nebo fyzikálně (tepelně, mechanicky). Vzorky jsou extrahovány organickými rozpouštědly a jejich směsmi (např. methanol nebo chloroform). Srovnání extrakčních procedur uvádí Detmer a spol.¹² nebo Sellick a spol.¹³. Všechny tyto postupy, přestože funkční a validované pro různé typy buněk,

jsou vždy limitovány tím, že popisují metabolismus buňky jako celku a neberou v potaz subcelulární kompartmentizaci metabolických procesů, tj. jejich lokalizaci v buněčných organelách.

3. Metody metabolické analýzy

Biologické vzorky obecně mají velmi komplexní charakter – obsažené látky patří k mnoha chemickým typům a mají několikařádrový koncentrační rozsah – např. v lidské krvi je možno nalézt stopové koncentrace některých hormonů vedle řádově milimolární hladiny glukosy. Dalším charakteristickým znakem (významným zejména pro lidskou metabolomiku) je omezené množství a dostupnost vzorků, zejména pro analýzy buněk v kulturách. Tyto skutečnosti dělají z metabolomiky jako analýzy všech nízkomolekulárních látek značně obtížnou disciplínu.

Značná část metabolomických experimentů se neobejde bez účinné separační části – užití chromatografických a elektroforetických metod. Nejčastěji používaným detektorem je v tomto případě hmotnostní spektrometr^{8,14}, jehož výhodou je zejména možnost současné identifikace i citlivé kvantifikace metabolitů. Odlišnou platformu komplexního měření pak zastupují techniky nukleární magnetické rezonance. Srovnání těchto technik z hlediska vhodnosti pro metabolomickou analýzu uvádí např. Musilová a Glatz¹⁵. Mezi méně používané platformy pak patří např. metody přímého nástřiku do hmotnostního spektrometru nebo infračervená a Ramanova spektrometrie.

3.1. Plynová chromatografie

Plynová chromatografie je metoda instrumentální chemické analýzy pro separaci a stanovení především plyných a těkavých kapalných látek. Princip metody a obecné aplikace byly popsány v mnoha monografiích^{16–18}. Ač jsou metody GC používány již přes 40 let, v současnosti podléhají vývojovým trendům. Použití MS jako detektoru pro GC se stalo dnes již rutinním. Svou renesanci dnes zažívá, díky rychlosti sběru dat a pulznímu dávkování iontů, které nezkrusuje MS spektra, time-of-flight analyzátor. V poslední době bylo toto uspořádání použito mimo jiné pro profilování látek v játrech pro charakterizaci hyperlipidemií¹⁹. Jinou užívanou kombinací je pak tandemová MS používaná pro kvantifikaci s velmi nízkým limitem detekce. Kvítang a spol.²⁰ vyvinuli vysoce senzitivní a časově nenáročnou metodu pro kvantifikaci 67 metabolitů v moči a séru derivatizovaných deuterovaným methyl-chlorformiátem. Dále je možno uvést např. metodu stanovení ethanolamidů dlouhých mastných kyselin v plazmě²¹, acetaminofenu²², steroidních hormonů²³ nebo kanabinoidů²⁴.

Vzhledem ke komplexnosti vzorku se do popředí dostává kompletní dvojrozměrná plynová chromatografie (comprehensive two-dimensional gas chromatography, GC×GC), technika využívající dvě kolony o různé selektivitě zapojené v sérii pomocí tzv. modulátoru. Modulátor (nejčastěji na základě kombinace studených a horkých pulzů plynu) shromáždí eluent z první kolony, který reinjekuje na druhou kolonu. Díky rozdílům v polaritě kolon jsou analyty podrobeny dvěma rozdílným separačním mechanismům. V porovnání s jednorozměrnou GC přináší tato technika zvýšení separační účinnosti, díky kryofokusaci zlepšení chromatografického rozlišení a snížení limitu detekce. Technické aspekty GC×GC byly popsány v několika přehledných článcích^{25–28}. GC×GC našla od svého uvedení v roce 1991 uplatnění v mnoha rozdílných oblastech, zejména v petrochemii, environmentální analýze, analýze potravin, ale také při metabolomickém studiu biologických vzorků²⁹. Mítrevski a spol.³⁰ studovali možnosti analýzy 27 látek steroidní povahy v moči. Popsány jsou extrakční a derivatizační účinnosti a možnosti identifikace. Li a spol.³¹ popisují uplatnění GC×GC pro hledání biomarkerů diabetu. Plazma diabetických pacientů byla deproteinována methanolem. Extrakt byl lyofilizován a silylován. Statistické analýze bylo podrobeno 692 metabolitů. Jako diferencující mezi pacienty s diabetem a kontrolami byly kromě glukosy označeny např. 2-hydroxyisobutyryát, fosfát, kyselina linoleová nebo palmitová. V práci Pasikantiho a spol.³² je představen metabolický obtisk (footprinting) potenciálně rakovinných buněk. Odebrané kultivační médium bylo deproteinováno methanolem a toluenem a následně byl extrakt odpařený v proudu dusíku silylován. Celkem 268 metabolitů bylo statisticky zpracováno a bylo objeveno 20 markerů, mezi nimi např. aminocyklopropankarboxylová kyselina, benzylalkohol nebo glukuronát. Orešić a spol.³³ se pokusili na vzorcích silylovaného séra definovat “metabolom“ schi-

zofrenie. Při vyhodnocování 201 identifikovaných metabolitů byly v tomto případě zaznamenány zvýšené hladiny aminokyselin – isoleucinu, fenylalaninu, prolinu a kyseliny glutamové. Technika byla použita také pro diagnostiku metabolických poruch z moči^{34,35}.

3.2. Kapalinová chromatografie

Vysokoučinná kapalinová chromatografie je v současnosti nejvíce používanou separační technikou. Spojením s hmotnostní spektrometrií se pak stává účinným nástrojem, který dovoluje separaci a charakterizaci mnoha metabolitů. HPLC může být použita k separaci široké škály sloučenin, a to látek jak hydrofilní, tak i hydrofobní povahy, solí, kyselin, bází apod., není totiž limitována pouze jedním mechanismem separace. HPLC může být zaměřena na separaci specifické třídy sloučenin, přičemž se k tomu využívají různé systémy např. na reverzní fázi (RP), normální fázi (NP) či hydrofilní interakční chromatografie (HILIC), nebo se využívá jejich kombinace. Lehmann a spol.³⁶ využili separaci na RP pro zjištění metabolitů, které se účastní procesu adaptace organismu po jeho zatížení. Bylo zjištěno zvýšení hladin středně dlouhých acylovaných karnitinů v plazmě. Cubbon a spol.³⁷ uplatnili HILIC separaci v analýze moči. Výsledky byly statisticky porovnány s výsledky RP separace. Bylo zjištěno, že oba systémy mají pro metabolomiku srovnatelné analytické parametry, z důvodu umožnění separace i velmi polárních látek je pak HILIC vhodnou metodou pro zvýšení pokrytí metabolomu. Lewis a spol.³⁸ použili, s cílem rozlišení většího počtu analytů, souběžně systém tří kolon s ortogonálními separačními charakteristikami a to pro metabolické profilování krve pacientů s infarktem myokardu.

Nevýhodou konvenční HPLC je relativně nízká účinnost, z toho důvodu byla vyvíjena ultraúčinná kapalinová chromatografie (UHPLC), která umožňuje separaci látek při vysokých tlacích na koloně obsahující částice menší než 2 μm. Dosud bylo publikováno několik klinických studií, v nichž byla UHPLC uplatněna. Příkladem může být Ma a spol.³⁹, kteří ji využili k metabolomické analýze moči k porovnání pacientů s rakovinou tlustého střeva.

Využití LC-MS v metabolomice již bylo několikrát přehledně shrnuto^{40–43}. Nejvíce je využívána ionizace elektrosprejem (ESI), chemická ionizace za atmosférického tlaku (APCI) a fotoionizace za atmosférického tlaku (APPI). V důsledku rozličných chemických vlastností metabolitů se často vyžaduje analyzovat biologický vzorek v pozitivních i negativních módech k docílení pokrytí celého metabolomu.

Ionizace elektrosprejem je nejvíce vhodná pro semi-polární a polární látky, tato ionizace se tedy užívá ve většině metabolomických studií. APCI a APPI se používají především pro látky nepolární, jako jsou lipidy. Ceglarek a spol.⁴⁴ využili APCI ionizaci při analýze steroidních hormonů v séru za účelem diagnostiky a monitorování endokrinních onemocnění. V dnešní době se stává trendem kombinace dvou ionizačních technik, především se jedná o kombinace ESI a APCI nebo ESI a APPI, což má opět za

následek zvýšení pokrytí metabolomu.

Hmotnostní analyzátoři mohou být rozděleny do několika typů, mezi něž patří kvadrupól (Q), iontová past (IT), průletový analyzátor (TOF), orbitrap a iontová cyklotronová rezonance s Fourierovou transformací (FT-ICR). Hybridní či tandemové hmotnostní spektrometry jsou pak kombinací dvou nebo více analyzátorů.

LC-MS umožňuje cílenou i necílenou analýzu metabolomu. Při cílené analýze se využívají hlavně IT a trojitě kvadrupóly (QqQ), kde druhý kvadrupól slouží jako kolizní cela. Tento systém je zde uplatňován hlavně díky svým možnostem kvantifikace, nejvíce se využívá analýz v tzv. MRM (multiple reaction monitoring) módu. Příkladem může být práce Wanga a spol.⁴⁵, kteří se zaměřili na hodnocení methylace argininu a kardiovaskulárních rizik.

Necílený přístup lze provádět především pomocí technik TOF, orbitrap a FT-ICR. Všechny tyto techniky charakterizuje vysoká rozlišovací schopnost a také vysoká přesnost určení hmoty (v současnosti i nižší než 3 ppm). Tento přístup nám umožňuje identifikovat markery onemocnění, např. různých typů rakoviny⁴⁶, diabetu⁴⁷, či ischemických chorob srdečních⁴⁸. Dalším příkladem může být i využití v diagnostice dědičných metabolických poruch. Wikoff a spol.⁴⁹ ukázali užitečnost necílené metabolomiky pro rozlišení pacientů s methylmalonovou a propionovou acidemií. Touto studií potvrdili, že společným markerem těchto onemocnění je propionylkarnitin.

Ritchie a spol.⁵⁰ uplatnili oba zmiňované přístupy za účelem identifikace sérových biomarkerů rakoviny tlustého střeva, necílenou analýzu prostřednictvím FT-ICR a cílenou pomocí Q-TOF a QqQ. Zjistili, že novými specifickými markery tohoto onemocnění mohou být snížené hladiny hydroxylovaných mastných kyselin s velmi dlouhým řetězcem.

I když se ve většině metabolických analýz využívá před hmotnostní detekcí chromatografická separace, byly též publikovány studie, kdy byla hmotnostní spektrometrie použita samostatně. Příkladem mohou být studie zabývající se dědičnými metabolickými poruchami^{51,52} či kardiovaskulárními chorobami⁵³.

3.3. Kapilární elektroforéza

Kapilární elektroforéza je analytická metoda používaná pro analýzu látek nesoucích náboj. Má tedy velký potenciál pro analýzu velké části metabolomu, většímu rozšíření metody však brání zejména problémy se spojením s MS detekcí. Ta je žádoucí zejména z důvodů detekce látek, které neabsorbují v UV/VIS oblasti, nebo identifikace neznámých analytů. Nejvíce komplikací působí on-line zapojení kapiláry do iontového zdroje (vzájemné ovlivňování napětí CE a iontového zdroje), jiný problém je způsoben kompatibilitou pufrů používaných v CE. Tyto nedostatky mají za následek pokles robustnosti analýz⁵⁴. Nejpoužívanějším iontovým zdrojem ve spojení s CE je ESI a to s použitím pomocné kapaliny (sheath liquid), která kromě vodivého spojení vytváří dostatečný průtok pro vytvoření

spreje. Z nastíněných důvodů byla CE dosud použita pouze v několika případech, např. Ramautar a spol.⁵⁵ popisují profilovou analýzu až 600 kationtů a aniontů v krysí moči s CE kapilárou pokrytou vrstvami polybren/dextranulfát/polybren a TOF MS detekcí. Poslední vývoje CE-MS v metabolomické analýze přehledně zpracovali Ramatour a spol.⁵⁶.

4. Analýza metabolomických dat

Metabolomické experimenty generují velké množství dat^{57–59}, které je nutno efektivně zpracovat, interpretovat, prezentovat a archivovat. Tato část je obvykle časově velmi náročná a vyžaduje specializované postupy vyplývající z charakteru získaných dat. Pro vyvození správných závěrů je však klíčová a je nutno jí věnovat stejnou pozornost jako plánování experimentu. Pro větší či menší části procesu analýzy jsou dostupné mnohé softwary jak volně přístupné (např. R, XCMS, MZmine), tak placené komerční balíky (např. Statistica, SIMCA-P, Minitab). Některé úvodní části procesu pak mohou být součástí samotného ovládacího softwaru přístroje. V metabolomice jsou často sbírána data mající dynamický (časově závislý) charakter (např. popis účinků léku na metabolickou dráhu v průběhu času), k čemuž je nutno také při vyhodnocení přihlížet⁶⁰.

4.1. Předúprava dat

Surová data metabolomického experimentu mají nejčastěji strukturu chromatografického (elektroforetického) záznamu, v němž se jako další dimenze vyskytují příslušná hmotnostní nebo NMR spektra. To vše je znásobeno příslušným počtem provedených měření. Pro další analýzu je nejprve nutno data vhodně upravit, což zahrnuje několik níže popsaných kroků, jejichž cílem je redukovat systematický a náhodný analytický šum. Konečným výsledkem pro následovné statistické zpracování je pak matice s obvyklým tvarem: $\mathbf{X}(m \times n)$, která má m řádků (vzorky) a n sloupců (plochy nebo koncentrace jednotlivých metabolitů). V dalším textu je míněna matice s touto strukturou.

Nejprve bývá provedena korekce podle základní linie (baseline correction). Prakticky to znamená její identifikaci (např. metodou Savitzky-Golay nebo Wavelet Transform) a odečtení od dat. Poté je nutno vyhledat části záznamu, ve kterých se nachází píky (peak-picking, peak-finding, peak detection). Lze použít několik strategií pro rozlišení signálu od šumu, nejjednodušeji pouze na základě prahového limitu poměru signál/šum.

Protože separace komplexního vzorku nebývá dokonalá (existuje možnost, že detegovaný pík obsahuje více analytů) následuje dekonvoluce signálu tj. matematické rozlišení překrývajících se signálů, např. GC chromatogramů. Tento proces s využitím druhé dimenze dat (obvykle MS) rozliší a seskupí ionty příslušné jednotlivým látkám a provádí se porovnáváním časového průběhu jednotlivých extrahovaných záznamů.

Do této doby jsou jednotlivé záznamy hodnoceny

samostatně, pro jejich vzájemné porovnání je nutno zajistit synchronizaci retenčních časů i m/z (alignment). V případě m/z dimenze je řešení jednoduché – správná hmotová kalibrace, změny v retenci jsou však problematictější. Metody vyvinuté pro popsání procesu zahrnují např. korekce na základě spektrální podobnosti ve zvoleném časovém intervalu nebo tzv. wrapping metody založené na protahování a zkracování jednotlivých chromatografických segmentů⁶¹.

Dalším krokem předúpravy dat může být jejich normalizace, tedy korekce řádkových profilů tak, aby byly kvantitativně porovnatelné. Příkladem může být normalizace analýzy moče na její zředění podle hladiny kreatininu nebo podle přídatku vhodného interního standardu, dále je třeba zahrnout také extrakční nebo ionizační účinnosti.

Pro většinu statistických přístupů je žádoucí, aby data pocházela z normálního rozdělení, proto jsou v tomto případě na místě transformace – logaritmická, mocnná, Box-Coxova aj. Mezi dále zmíněné běžné úpravy zajišťující objektivnější porovnatelnost patří centrování (tj. odečtení příslušného průměru eventuálně mediánu, obvykle sloupcové) a škálování tj. vydělení příslušnou směrodatnou odchylkou (unit-variace scaling), rozptylem (Paretovo škálování) nebo rozsahem. Nutno si také uvědomit, že zmíněné operace nejsou asociativní a záleží, v jakém pořadí jsou provedeny.

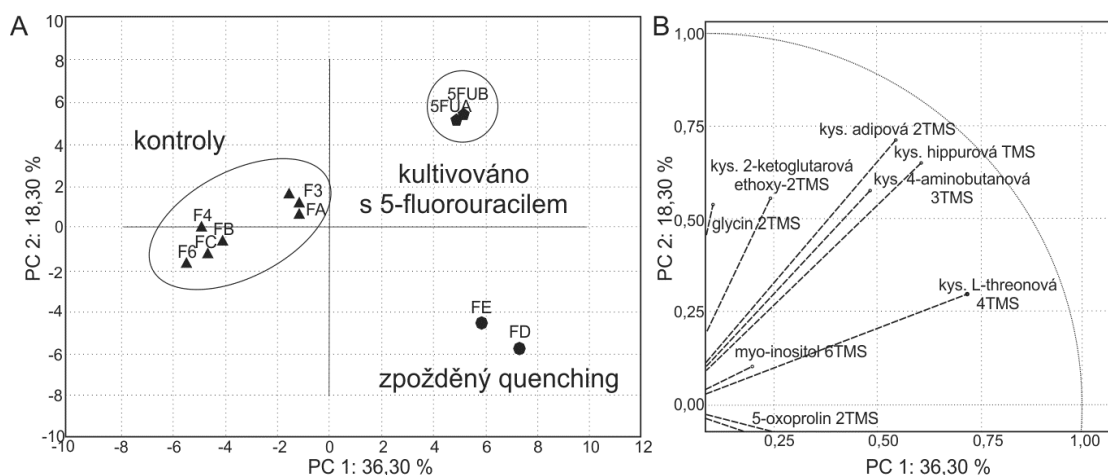
Nezanedbatelnou součástí předúpravy by mělo být také zhodnocení chybějících hodnot. Tzv. „nuly“ jsou pro statistiku problémem vedoucím k deformované struktuře dat – je nutno identifikovat příčinu vzniku těchto nul a odstranit je např. nahrazením minimální hodnotou (šum). Podobný problém mohou způsobit také odlehle body – mělo by být zhodnoceno, zdali je jejich výskyt způsoben biologickou variabilitou nebo nějakou hrubou chybou (např. při přípravě vzorku).

4.2. Statistická analýza

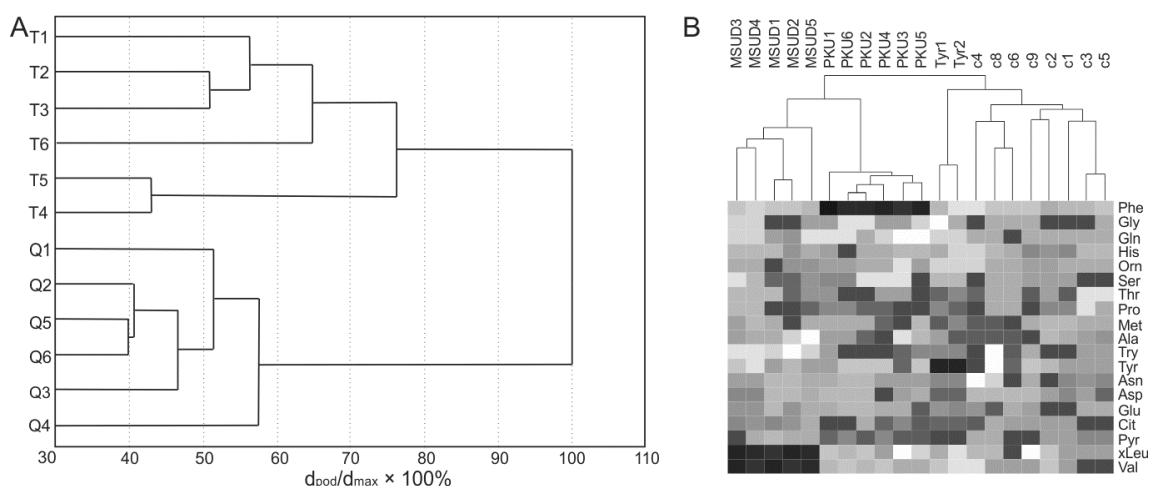
Metabolomický experiment v zásadě generuje multidimenzionální data, použití jednorozměrné statistiky je tedy omezeno. Lze ji použít např. pro testování hypotéz o významnosti jednotlivých metabolitů mezi dvěma skupinami vzorků. Jedná se zejména o parametrické testy – t -test, ANOVA (analýza rozptylu), nebo neparametrické testy pro data, která nepocházejí z normálního rozdělení (Kruskal-Wallis).

Dominantní pro metabolomiku jsou ovšem multidimenzionální statistické postupy. Jejich hlavními úkoly jsou systematizace dat, nalezení vyšších struktur, trendů a souvislostí. Používané postupy lze rozdělit na dvě skupiny – „unsupervised“ a „supervised“. První skupina zahrnuje zejména metody redukce dimenze dat a metody shlukovací. Typickou metodou redukující dimenzi dat je analýza hlavních komponent (Principal Component Analysis, PCA). V tomto případě se jedná o transformaci původních dat do nové soustavy malého počtu proměnných (hlavní komponenty, principal components, PC), které však vystihují velké množství původní variability a navzájem nekorelují, a tudíž poskytují nezávislé informace o systému. Grafickými výstupy pro diagnostiku PCA jsou rozptylový diagram komponentních skóre (score plot, zobrazuje průmět objektů do roviny zvolených PC, obr. 1A) a graf komponentních zátěží (loadings plot, vyjadřuje vztah původních proměnných k novým PC, obr. 1B). Jinou možností je tzv. biplot kombinující oba předchozí. PCA je v mnoha metabolomických studiích používána jako první průzkumný krok před dalšími analýzami a také pro srozumitelné zobrazení multidimenzionálních dat do roviny.

Shluková analýza (Cluster Analysis, CA) je postup, který vyhledává podobnosti mezi jednotlivými objekty nebo proměnnými na základě jejich definované vzdálenos-



Obr. 1. **Analýza hlavních komponent;** (A) Projekce případů do komponentní roviny pro porovnání 78 látek metabolomu fibroblastů, odděleny jsou kontroly od vzorků s nesprávným postupem přípravy a od vzorků kultivovaných s 5-fluorouracilem; (B) část projekce zátěží do komponentní roviny – ukazuje, které látky metabolomu fibroblastů jsou v přítomnosti 5-fluorouracilu nejvíce ovlivněny; upraveno dle cit.⁷²



Obr. 2. **Shluková analýza;** (A) dendrogram porovnávající dva způsoby sklizení fibroblastů – trypsinizaci (T1–T6) a methanický quenching (Q1–Q6); (B) „heat-map“ zobrazující metabolický vzor vybraných dědičných onemocnění (MSUD – leucinoza, PKU – fenylketurie, Tyr – tyrosinémie, c – kontroly), data jsou normalizována a podrobena $\log(2)$ transformaci; upraveno dle cit.⁷³

ti v multidimenzionálním prostoru. Na výsledek shlukování mají zásadní vliv dva parametry – metrika vzdálenosti (jak jsou shluky vlastně daleko – Eukleidovská, Mahalanobisova, Manhattanovská vzdálenost, korelace) a metoda shlukování (které dva shluky lze považovat za jeden – centroidní, mediánová, Wardova, metoda nejbližšího souseda a další). Zobrazením CA je tzv. dendrogram (obr. 2A), vývojový strom zobrazující průběh shlukování. S CA úzce souvisí i tzv. teplotní mapa (heat-map, obr. 2B). Jedná se o klastrovaná data pro mnoho vzorků (a)nebo analytů. Kvantitativní stránka datového souboru je zde vyjádřena pomocí barevné škály, nejčastěji dvoubarevné, vyjadřující zvýšení a snížení. Toto uspořádání umožňuje jednoduše vizualizovat i rozsáhlé datové soubory. Uplatňuje se například pro klinická data k rozpoznávání vzorů (pattern recognition)⁶².

„Supervised“ postupy pracují jak s **X** maticí naměřených dat, tak s **Y** maticí tzv. prediktorů – tedy předem známých informací (pro lidskou metaboliku se může jednat např. o věk, krevní tlak, ale i o kategorické proměnné – zdravý/nemocný). Úkol těchto metod zahrnuje zejména problém regrese (numerický výstup) a klasifikace (identifikace tříd). Metody částečných nejmenších čtverců (PLS, Partial Least Square) a její varianta ortogonální-PLS (OPLS) zahrnují mnohonásobně lineárně regresní postupy založené na PCA, které kvantifikují vztah mezi **X** a **Y**. V případě klasifikace také existují příslušné metody diskriminační analýzy (DA, Discriminant Analysis) – PLS-DA, OPLS-DA. Dále lze zmínit např. SVM (Support Vector Machine), CVA (Canonical Variates Analysis), SIMCA (Soft Independent Modeling of Class Analogy) nebo jednu z variant neuronové sítě tzv. Kohonenovy mapy (též Self-Organizing Maps, SOM). Jejich použití je vhodné v případech, kdy je nutné prohloubit rozdělení jednotli-

vých skupin pozorování na maximum.

Jiná skupina „supervised“ postupů je, i když se stejnými cíli, založena na poněkud odlišných principech. Jsou to například pro klasifikaci sloužící rozhodovací stromy (Decision Trees) nebo množství umělých neuronových sítí (Artificial Neural Networks, ANN), které mohou sloužit jak pro klasifikaci, tak pro regresi nebo predikci časových řad. Základem neuronových sítí je neuron, jednotka propojená s ostatními do hierarchické struktury podobně jako v lidském mozku. Každý neuron může mít mnoho vstupů s různými vahami, které vlastním algoritmem vyhodnocuje a předává jinému neuronu. Mezi vstupní (**X**) a výstupní (**Y**) data je pak vloženo několik vrstev neuronů. Neuronová síť je schopna se na modelovém setu dat „naučit“ řešit daný problém a na tomto základě pak řešit podobné. Je třeba zdůraznit, že výčet statistických metod není zdaleka vyčerpávající a použití dané metody pro určitou aplikaci je nutno pečlivě zvážit.

Potřeba integrace dat dává vzniknout databázím, z nichž řada je dostupná volně na internetu a jejichž přehled je uveden v tab. II.

5. Závěr

Metabolika je rychle se rozvíjející vědou, která spojuje biochemii a medicínu, přičemž využívá pokročilé analytické techniky a sofistikované statistické metody a stojí v centru zájmu systémové biologie. Hlavní směry jejího současného vývoje zahrnují jak zlepšení limitu detekce přístrojů, tak rozšíření univerzálnosti a komplexnosti používaných metod, což vyžaduje nové přístupy nejen v instrumentaci, ale i v následném zpracování naměřených dat. Závěrem lze říci, že ač jsou metabolické postupy

Tabulka II
Přehled některých databází s metabolickou tematikou

Název	Poznámky	Odkaz
AraCyc ^a	zaměřená se <i>Arabidopsis thaliana</i>	http://www.arabidopsis.org/biocyc/index.jsp
EcoCyc ^{a63}	zaměřená na <i>Escherichia coli</i> K-12 MG1655	http://ecocyc.org/
HumanCyc ^a	encyklopedie lidských genů a metabolismu	http://humancyc.org/
MetaCyc ^{a64}	metabolické dráhy více než 2000 organismů	http://metacyc.org/index.shtml
YeastCyc ^a	zaměřená na <i>Saccharomyces cerevisiae</i>	http://biocyc.org/YEAST/organism-summary?object=YEAST
KEGG ⁶⁵	Kyoto Encyclopedia of Genes and Genomes; komplexní databáze obsahující informace o mikrobiálním, rostlinném i živočišném metabolomu (metabolické dráhy, reakce, enzymy)	http://www.genome.jp/kegg/ligand.html
HMDB ⁶⁶	Human Metabolome Database; databáze lidských metabolitů, obsahuje chemická, klinická a molekulárně biologická data	http://www.hmdb.ca/
Reactome ⁶⁷	databáze primárně lidských biochemických drah	http://www.reactome.org/
MeMo ⁶⁸	metabolomické modelování	
PUMA2 ⁶⁹	vývojová analýza metabolomu	http://compbio.mcs.anl.gov/puma2
BRENDA ⁷⁰	systém informací o enzymech	http://www.brenda-enzymes.org/

^a Součástí kolekce databází BioCyc⁷¹

stále ve vývoji, jsou již v mnoha ohledech dobře zavedené. Množství dat, které lze takto generovat, analyzovat a proměnit v biologicky relevantní znalosti, je obrovské a lze předpokládat, že bude dále narůstat.

Tato práce vznikla za podpory Studentské grantové soutěže Univerzity Palackého v Olomouci, č. projektu LF_2012_016 a grantu IGA MZCR NT12218. Infrastrukturní část projektu (Ústav molekulární a translační medicíny LF UP) je financována z projektů Operačního programu věda a výzkum pro inovace (projekt BIOMEDREG, CZ.1.05/2.1.00/01.0030).

LITERATURA

- Goodacre R.: *J. Nutr.* 137, 259S (2007).
- Dunn W. B., Bailey N. J. C., Johnson H. E.: *Analyst* 130, 606 (2005).
- Fiehn O., Robertson D., Griffin J., van der Werf M., Nikolau B., Morrison N., Sumner L. W., Goodacre R., Hardy N. W., Taylor C., Fostel J., Kristal B., Kaddurah-Daouk R., Mendes P., van Ommen B., Lindon J. C., Sansone S.: *Metabolomics* 3, 157 (2007).
- 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., Reily M. D., Thaden J. J., Viant M. R.: *Metabolomics* 3, 211 (2007).
- Goodacre R., Broadhurst D., Smilde A. K., Kristal B. S., Baker J. D., Beger R., Bessant C., Connor S., Capuani G., Craig A., Ebbels T., Kell D. B., Manetti C., Newton J., Paternostro G., Somorjai R., Sjöström M., Trygg J., Wulfert F.: *Metabolomics* 3, 231 (2007).
- Sansone S. A., Schober D., Atherton H. J., Fiehn O., Jenkins H., Rocca-Serra P., Rubtsov D. V., Spasic I., Soldatova L., Taylor C., Tseng A., Viant M. R.: *Metabolomics* 3, 249 (2007).
- Hardy N. W., Taylor C. F.: *Metabolomics* 3, 243 (2007).
- Dettmer K., Aronov P. A., Hammock B. D.: *Mass Spectrom. Rev.* 26, 51 (2007).
- Lorenz M. A., Burant C. F., Kennedy R. T.: *Anal. Chem.* 83, 3406 (2011).
- Teng Q., Huang W., Collette T. W., Ekman D. R., Tan C.: *Metabolomics* 5, 199 (2008).
- Villas-Boas S. G., Bruheim P.: *Anal. Biochem.* 370, 87 (2007).
- Dettmer K., Nurnberger N., Kaspar H., Gruber M. A., Almstetter M. F., Oefner P. J.: *Anal. Bioanal. Chem.* 399, 1127 (2011).
- Sellick C. A., Knight D., Croxford A. S., Maqsood A. R., Stephens G. M., Goodacre R., Dickson A. J.: *Metabolomics* 6, 427 (2010).
- Villas-Bôas S. G., Mas S., Åkesson M., Smedsgaard J., Nielsen J.: *Mass Spectrom. Rev.* 24, 613 (2005).
- Musilová J., Glatz Z.: *Chem. Listy* 105, 745 (2011).
- Jennings W., Mittlefehldt E., Stremple P.: *Analytical*

- Gas Chromatography*. Academic Press, San Diego 1997.
17. Grob R. L., Barry E. F. (ed.): *Modern Practice of Gas Chromatography*. J. Wiley, Hoboken 2004.
 18. Sparkman O. D., Penton Z. E., Kitson F. G.: *Gas Chromatography and Mass Spectrometry: A Practical Guide*. Academic Press, San Diego 2011.
 19. Gu S. H., Jiye A., Wang G. J., Zha W. B., Yan B., Zhang Y., Ren H. C., Cao B., Liu L. S.: *Biomed. Chrom.* 24, 245 (2010).
 20. Kvitvang H. F., Andreassen T., Adam T., Villas-Boas S. G., Bruheim P.: *Anal. Chem.* 83, 2705 (2011).
 21. Zoerner A. A., Gutzki F. M., Suchy M. T., Beckmann B., Engeli S., Jordan J., Tsikas D.: *J. Chromatogr., B* 877, 2909 (2009).
 22. Trettin A., Zoerner A. A., Bohmer A., Gutzki F. M., Stichtenoth D. O., Jordan J., Tsikas D.: *J. Chromatogr., B* 879, 2274 (2011).
 23. Hansen M., Jacobsen N. W., Nielsen F. K., Bjorklund E., Styrihave B., Halling-Sorensen B.: *Anal. Bioanal. Chem.* 400, 3409 (2011).
 24. Emidio E. S., Prata V. D., Dorea H. S.: *Anal. Chim. Acta* 670, 63 (2010).
 25. Gorecki T., Harynuk J., Panic O.: *J. Sep. Sci.* 27, 359 (2004).
 26. Mondello L., Tranchida P. Q., Dugo P., Dugo G.: *Mass Spectrom. Rev.* 27, 101 (2008).
 27. Adahchour M., Beens J., Vreuls R. J. J., Brinkman U. A. T.: *TrAC, Trends Anal. Chem.* 25, 438 (2006).
 28. Adahchour M., Beens J., Vreuls R. J. J., Brinkman U. A. T.: *TrAC, Trends Anal. Chem.* 25, 540 (2006).
 29. Adahchour M., Beens J., Brinkman U. A.: *J. Chromatogr., A* 1186, 67 (2008).
 30. Mitrevski B. S., Brenna J. T., Zhang Y., Marriott P. J.: *J. Chromatogr., A* 1214, 134 (2008).
 31. Li X., Xu Z., Lu X., Yang X., Yin P., Kong H., Yu Y., Xu G.: *Anal. Chim. Acta* 633, 257 (2009).
 32. Pasikanti K. K., Norasmaria J., Cai S., Mahendran R., Esuvaranathan K., Ho P. C., Chan E. C.: *Anal. Bioanal. Chem.* 398, 12853 (2010).
 33. Oresic M., Tang J., Seppanen-Laakso T., Mattila I., Saarni S. E., Saarni S. I., Lonqvist J., Sysi-Aho M., Hyotylainen T., Perala J., Suvisaari J.: *Genome Med.* 3, 19 (2011).
 34. Kouremenos K. A., Pitt J., Marriott P. J.: *J. Chromatogr., A* 1217, 104 (2010).
 35. Wojtowicz P., Zrostlikova J., Kovaleczuk T., Schurek J., Adam T.: *J. Chromatogr., A* 1217, 8054 (2010).
 36. Lehmann R., Zhao X., Weigert C., Simon P., Fehrenbach E., Fritsche J., Machann J., Schick F., Wang J., Hoene M., Schleicher E. D., Haring H. U., Xu G., Niess A. M.: *PLoS One* 5, e11519 (2010).
 37. Cubbon S., Bradbury T., Wilson J., Thomas-Oates J.: *Anal. Chem.* 79, 8911 (2007).
 38. Lewis G. D., Wei R., Liu E., Yang E., Shi X., Martinovic M., Farrell L., Asnani A., Cyrille M., Ramathanathan A., Shaham O., Berriz G., Lowry P. A., Palacios I. F., Tasan M., Roth F. P., Min J., Baumgartner C., Keshishian H., Addona T., Mootha V. K., Rosenzweig A., Carr S. A., Fifer M. A., Sabatine M. S., Gerszten R. E.: *J. Clin. Invest.* 118, 3503 (2008).
 39. Ma Y. L., Qin H. L., Liu W. J., Peng J. Y., Huang L., Zhao X. P., Cheng Y. Y.: *Dig. Dis. Sci.* 54, 2655 (2009).
 40. Juo C. G., Chiu D. T., Shiao M. S.: *Biofactors* 34, 159 (2008).
 41. Zhou B., Xiao J. F., Tuli L., Resson H. W.: *Mol. Biosyst.* 8, 470 (2012).
 42. Roux A., Lison D., Junot C., Heilier J. F.: *Clin. Biochem.* 44, 119 (2011).
 43. Becker S., Kortz L., Helmschrodt C., Thiery J., Ceglarek U.: *J. Chromatogr., B* 883–884, 68 (2012).
 44. Ceglarek U., Kortz L., Leichtle A., Fiedler G. M., Kratzsch J., Thiery J.: *Clin. Chim. Acta* 401, 114 (2009).
 45. Wang Z., Tang W. H., Cho L., Brennan D. M., Hazen S. L.: *Arterioscler. Thromb. Vasc. Biol.* 29, 1383 (2009).
 46. Chen J., Zhang X. Y., Cao R., Lu X., Zhao S. M., Fekete A., Huang Q., Schmitt-Kopplin P., Wang Y. S., Xu Z. L., Wan X. P., Wu X. H., Zhao N. Q., Xu C. J., Xu G. W.: *J. Proteome Res.* 10, 2625 (2011).
 47. Kim K., Aronov P., Zakharkin S. O., Anderson D., Perroud B., Thompson I. M., Weiss R. H.: *Mol. Cell. Proteomics* 8, 558 (2009).
 48. Liu G. W., Snapp H. M., Ji Q. C., Arnold M. E.: *Anal. Chem.* 81, 9225 (2009).
 49. Wikoff W. R., Gangooi J. A., Barshop B. A., Siuzdak G.: *Clin. Chem.* 53, 2169 (2007).
 50. Ritchie S. A., Ahiahonu P. W. K., Jayasinghe D., Heath D., Liu J., Lu Y. S., Jin W., Kavianpour A., Yamazaki Y., Khan A. M., Hossain M., Su-Myat K. K., Wood P. L., Krenitsky K., Takemasa I., Miyake M., Sekimoto M., Monden M., Matsubara H., Nomura F., Goodenowe D. B.: *BMC Med.* 8 (2010).
 51. Janeckova H., Hron K., Wojtowicz P., Hlidkova E., Baresova A., Friedecky D., Zidkova L., Hornik P., Behulova D., Prochazkova D., Vinohradska H., Peskova K., Bruheim P., Smolka V., Stastna S., Adam T.: *J. Chromatogr., A* 1226, 11 (2012).
 52. Jansson J., Willing B., Lucio M., Fekete A., Dicksved J., Halfvarson J., Tysk C., Schmitt-Kopplin P.: *PLoS One* 4, e6386 (2009).
 53. Turer A. T., Stevens R. D., Bain J. R., Muehlbauer M. J., van der Westhuizen J., Mathew J. P., Schwinn D. A., Glower D. D., Newgard C. B., Podgoreanu M. V.: *Circulation* 119, 1736 (2009).
 54. Klampfl C. W.: *Electrophoresis* 27, 3 (2006).
 55. Ramautar R., Torano J. S., Somsen G. W., de Jong G. J.: *Electrophoresis* 31, 2319 (2010).
 56. Ramautar R., Mayboroda O. A., Somsen G. W., de Jong G. J.: *Electrophoresis* 32, 52 (2011).
 57. Goodacre R., Broadhurst D., Smilde A. K., Kristal B. S., Baker J. D., Beger R., Bessant C., Connor S., Capuani G., Craig A., Ebels T., Kell D. B., Manetti C., Newton J., Paternostro G., Somorjai J., Sjöström M.,

- Trygg J., Wulfert F.: *Metabolomics* 3, 231 (2007).
58. Trygg J., Holmes E., Lundstedt T.: *J. Proteome Res.* 6, 469 (2007).
59. G. Boccard J., Veuthey J., Rudaz S.: *J. Sep. Sci.* 33, 290 (2010).
60. Smilde A. K., Westerhuis J. A., Hoefsloot H. C. J., Biljsma S., Rubingh C. M., Vis D. J., Jellema R. H., Pijl H., Roelfsema F., van der Greef J.: *Metabolomics* 6, 3 (2010).
61. Skov T., van den Berg F., Tomasi G., Bro R.: *J. Chemometr.* 20, 484 (2007).
62. Moon J. Y., Jung H. J., Moon M. H., Chung B. C., Choi M. H.: *J. Am. Soc. Mass Spectrom.* 20, 1626 (2009).
63. Keseler I. M., Bonavides-Martinez C., Collado-Vides J., Gama-Castro S., Gunsalus R. P., Johnson D. A., Krummenacker M., Nolan L. M., Paley S., Paulsen I. T., Peralta-Gil M., Santos-Zavaleta A., Shearer A. G., Karp P. D.: *Nucleic Acids Res.* 37, D464 (2009).
64. Caspi R., Altman T., Dale J. M., Dreher K., Fulcher C. A., Gilham F., Kaipa P., Karthikeyan A. S., Kothari A., Krummenacker M., Latendresse M., Mueller L. A., Paley S., Popescu L., Pujar A., Shearer A. G., Zhang P., Karp P. D.: *Nucleic Acids Res.* 38, D473 (2010).
65. Kanehisa M., Goto S., Furumichi S., Tanabe M., Hirakawa M.: *Nucleic Acids Res.* 38, D335 (2010).
66. Wishart D. S., Knox C., Guo A. C., Eisner R., Young N., Gautam B., Hau D. D., Psychogios N., Dong E., Bouatra S., Mandal R., Sinelnikov I., Xia J., Jia L., Cruz J. A., Lim E., Sobsey C.A., Shrivastava S., Huang P., Liu P., Fang L., Peng J., Fradette R., Cheng D., Tzur D., Clement M., Lewis A., De Souza A., Zuniga A., Dawe M., Xiong Y., Clive D., Greiner R., Nazyrova A., Shaykhtudinov R., Li L., Vogel H. J., Forsythe I.: *Nucleic Acids Res.* 38, D603 (2010).
67. Croft D., O'Kelly G., Wu G., Haw R., Gillespie M., Matthews L., Caudy M., Garapati P., Gopinath G., Jassal B., Jupe S., Kalatskaya I., Mahajan S., May B., Ndegwa N., Schmidt E., Shamovsky V., Yung C., Birney E., Hermjakob H., D'Eustachio P., Stein L.: *Nucleic Acids Res.* 39, D691 (2011).
68. Spasić I., Dunn W. B., Velarde G., Tseng A., Jenkins H., Hardy N., Oliver S. G., Kell D. B.: *BMC Bioinf.* 7, 281 (2006).
69. Maltsev N., Glass E., Sulakhe D., Rodriguez A., Syed M. H., Bompada T., Zhang Y., D'Souza M.: *Nucleic Acids Res.* 34, D369 (2006).
70. Scheer M., Grote A., Chang A., Schomburg I., Munnaretto C., Rother M., Söhngen C., Stelzer M., Thiele J., Schomburg D.: *Nucleic Acids Res.* 38, 1 (2010).
71. Karp P. D., Ouzounis C. A., Moore-Kochlacs C., Goldovsky L., Kaipa P., Ahren D., Tsoka S., Darzentas N., Kunin V., Lopez-Bigas N.: *Nucleic Acids Res.* 33, 6083 (2005).
72. Wojtowicz P., Dostálová E., Adam T.: *Klin. Biochem. Metab.* 20, 38 (2012).
73. Wojtowicz P., Zrostlikova J., Šťastná V., Dostálová E., Židková L., Bruheim P., Adam T., v knize: *Gas Chromatography - Biochemicals, Narcotics and Essential Oils* (Salih B., ed.), kap. 2, str. 42, 46. InTech, Rijeka 2012.

P. Wojtowicz, H. Janečková, D. Friedecký, and T. Adam (*Laboratory for Inherited Metabolic Disorders and Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University, Olomouc*): **Metabolomic Techniques in Biomedicine**

Global metabolite profiling, often called metabolomics, is an expanding research field. Together with genomics, transcriptomics and proteomics, it provides additional information on specific changes occurring in biological systems, allowing to better understand metabolic pathways and their pathologies. This review highlights the current metabolomic challenges in biomedical research using chromatographic and electrophoretic separation techniques coupled mostly with mass spectrometry. The most frequently used procedures for obtaining bioinformatics data are introduced as well.