Možnosti využití celogenomové sekvenace (WGS) k analýze izolátů *Streptococcus pneumoniae*

Vohrnová S.[1, 2], **Kozáková J.**[1]

[1]Národní referenční laboratoř pro streptokokové nákazy, Oddělení bakteriálních vzdušných nákaz, Centrum epidemiologie a mikrobiologie, Státní zdravotní ústav, Praha
[2]3. lékařská fakulta Univerzity Karlovy, Praha

SOUHRN

Streptococcus pneumoniae (pneumokok) je grampozitivní kok vyvolávající jak neinvazivní, tak invazivní infekční onemocnění. Onemocnění vyvolaná pneumokokem mohou být preventabilní očkováním. Invazivní pneumokoková onemocnění (IPO) musí splňovat mezinárodní definici případu, jsou hlášena na národní i mezinárodní úrovni a v řadě zemí, včetně České republiky, jsou sledována v programu surveillance. Důležitou součástí surveillance IPO je sledování vyskytujících se sérotypů, hodnocení četnosti jednotlivých sérotypů v čase a v relaci k probíhajícím vakcinačním programům.

Ve světě i v České republice je u pneumokoků stále častěji prováděna metoda celogenomové sekvenace (whole genome sequencing, WGS), která umožňuje určování sérotypu pneumokoků ze sekvenačních dat, dále přesnou analýzu jejich genetických vztahů a studium genů obsažených v jejich genomu.

Celogenomová sekvenace umožňuje získávat spolehlivá a mezinárodně srovnatelná sekvenační data, která lze snadno sdílet. Sekvenační data jsou analyzována pomocí bioinformatických nástrojů, které vyžadují znalosti z oblasti přírodních věd s důrazem na genetiku a znalosti z oblasti bioinformatiky.

V publikaci jsou představeny některé možnosti analýzy pneumokoka, kterými jsou sérotypizace, multilokusová sekvenační typizace (MLST), ribozomální MLST (rMLST), core genome MLST (cgMLST), whole genome MLST (wgMLST), single nucleotide polymorphism (SNP) analýza, určení Global Pneumococcal Sequence Cluster (GPSC), stanovení genů virulence a genů antibiotické rezistence.

U metody WGS je prezentována strategie její aplikace v Evropě a realizace v praxi.

Analýza pneumokoků metodou WGS představuje zlepšení v provádění surveillance IPO, kdy je sérotyp určován molekulárně geneticky, jsou prováděny další podrobnější typizace, získaná data lze snadno mezinárodně porovnávat a lze lépe hodnotit účinnost vakcinačních programů.

KLÍČOVÁ SLOVA

Streptococcus pneumoniae – celogemová sekvenace– sérotypitace – molekulární surveillance – invazivní pneumokokové onemocnění

ABSTRACT

Vohrnová S., Kozáková J.: Posibilities for use of whole genome sequencing (WGS) for the analysis of *Streptococcus pneumoniae* isolates

Streptococcus pneumoniae (pneumococcus) is a Gram-positive coccus causing both non-invasive and invasive infectious diseases. Pneumococcal diseases are vaccine preventable. Invasive pneumococcal diseases (IPD) meeting the international case definition are reported nationally and internationally and are subject to surveillance programmes in many countries, including the Czech Republic. An important part of IPD surveillance is the monitoring of causative serotypes and their frequency over time and in relation to ongoing vaccination programmes.

In the world and in the Czech Republic, whole genome sequencing (WGS) is increasingly used for pneumococci, which allows for serotyping from sequencing data, precise analysis of their genetic relationships, and the study of genes present in their genome. Whole-genome sequencing enables the generation of reliable and internationally comparable data that can be easily shared. Sequencing data are analysed using bioinformatics tools that require knowledge in the field of natural sciences with an emphasis on genetics and expertise in bioinformatics.

This publication presents some options for pneumococcal analysis, i.e., serotyping, multilocus sequence typing (MLST), ribosomal MLST (rMLST), core genome MLST (cgMLST), whole genome MLST (wgMLST), single nucleotide polymorphism (SNP) analysis, assignment to Global Pneumococcal Sequence Cluster (GPSC), and identification of virulence genes and antibiotic resistance genes. The WGS strategies and applications for Europe and WGS implementation in practice are presented.

WGS analysis of pneumococci allows for improved IPD surveillance, thanks to molecular serotyping, more detailed typing, generation of internationally comparable data, and improved evaluation of the effectiveness of vaccination programmes.

KEYWORDS

Streptococcus pneumoniae – whole genome sequencing – serotyping – molecular surveillance – invazive pneumococcal disease

Epidemiol Mikrobiol Imunol, 2024; 73(1): 30–36
https://doi.org/10.61568/emi/11-6254/20240123/136240

proLékaře.cz | 30.11.2026

ÚVOD

Streptococcus pneumoniae (pneumokok) je grampozitivní, opouzdřený, kataláza negativní kok. Identifikace pneumokoků je prováděna pomocí alfa hemolýzy na krevním agaru, testu rozpustnosti v 10% roztoku deoxycholátu sodného (test rozpustnosti ve žluči – kapkové či zkumavkové provedení), testováním citlivosti k optochinu, latexovou aglutinací s polysacharidovým pouzdrem pneumokoka, molekulárně genetickým testováním genů *cps*A, *lyt*A, *ply*, *sod*A a sekvenováním 16S rRNA [1]. Nejspolehlivější metodou identifikace pneumokoka je zkumavkový test rozpustnosti ve žluči, následován testováním citlivosti k optochinu [2, 3].

Pneumokok je obalen mukózním pouzdrem, které je tvořeno provázanými řetězci polysacharidů. Na základě specifické charakteristiky pouzdra rozlišujeme u pneumokoků sérotyp, který je definován jako kmen pneumokoka produkující polysacharid pouzdra o unikátním složení a imunochemických vlastnostech [4]. Sérotyp je určován pomocí vazby protilátek obsažených v typovém séru, které je speciálně vytvořeno proti imunochemické charakteristice konkrétního sérotypu, s pouzdrem vyšetřovaného kmene pneumokoka [5]. Aktuálně rozeznáváme přes 100 sérotypů [6, 7]. Imunitní systém hostitele tvoří proti polysacharidovému pouzdru specifické protilátky a po prodělání infekčního onemocnění vyvolaném pneumokokem určitého sérotypu nalezneme následně v krvi pacienta protilátky proti tomuto vyvolávajícímu sérotypu. Vytvořené protilátky poskytují ochranu před opětovným nakažením stejným sérotypem.

Pneumokok je lidským patogenem způsobujícím širokou škálu infekčních onemocnění od méně závažných infekcí horních cest dýchacích a sinusitid, přes akutní hnisavé záněty středního ucha, až po závažná invazivní pneumokoková onemocnění (IPO), jako jsou pneumonie, sepse a meningitida.

Pneumokok kromě svého patogenního působení také běžně kolonizuje lidské sliznice horních cest dýchacích, především nazofaryngu a orofaryngu. Nosičství pneumokoka je častější v rozvojových než ve vyspělých zemích, dále se četnost nosičství liší dle věkových skupin – je častější u dětí než u dospělé populace. U dětí navštěvujících mateřské školy bylo zaznamenáno nosičství pneumokoka u 3–60 % dětí [8, 9, 10]. V české studii z roku 2010 bylo zjištěno nosičství pneumokoka až u 60 % dětí ve věku 4 až 5 let navštěvujících mateřské školy [11]. U dospělých bývá nosičství pneumokoka prokazováno až u 8 % dospělé populace [12, 13]. Nosičství pneumokoka bývá bezpříznakové, ale může se rozvinout infekční onemocnění. Rozvoj pneumokokového onemocnění je podmíněn předchozím nosičstvím pneumokoka [14, 15].

Ochranu proti infekčním onemocněním vyvolaným pneumokoky představuje očkování. Aktuálně je dostupných pět vakcín proti pneumokokům. Pneumokoková polysacharidová vakcína PPV-23, která obsahuje polysacharidové antigeny 23 sérotypů pneumokoků (obsahuje sérotypy 1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19A, 19F, 20B, 22F, 23F, 33F). Dále jsou dostupné pneumokokové konjugované vakcíny (PCV), a to 10valentní (obsahuje sérotypy 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F, 23F), 13valentní (sérotypy 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 23F), 15valentní (sérotypy 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 22F, 23F, 33F) a 20valentní (sérotypy 1, 3, 4, 5, 6A, 6B, 7F, 8, 9V, 10A, 11A, 12F, 14, 15B, 18C, 19A, 19F, 22F, 23F, 33F). V budoucnosti lze očekávat uvedení nových 21valentních a 24valentních konjugovaných vakcín, u kterých v současnosti probíhají klinické testy [16, 17].

V současnosti je klíčovým momentem při navrhování, zavádění a hodnocení vakcinační strategie znalost vyskytujících se sérotypů pneumokoka v průběhu času v dané populaci – ideálně již před zavedením vakcinace, dále při zavedení vakcinace a v časovém odstupu po zavedení vakcinace. Vakcinační strategie je v průběhu času upravována s ohledem na aktuálně dostupné vakcíny a výskyt sérotypů v populaci, která má být vakcínou očkována.

Určování sérotypů pneumokoka je založeno na různých přístupech. Metody lze rozdělit na fenotypové, které využívají imunologické vlastnosti pouzdra pneumokoka, a genotypové, které testují genový podklad tvorby pouzdra [4, 18].

Historicky nejdéle a dosud nejčastěji je sérotyp pneumokoka určován pomocí tzv. Quellung neboli Neufeldovy reakce, která je známa a využívána od roku 1902 [19]. Jedná se o sérologickou reakci mezi typovým pouzderným polysacharidovým antigenem pneumokoka a typově specifickou protilátkou v hyperimunním králičím antiséru. V případě pozitivní reakce dojde během několika minut k nabobtnání pouzdra, které lze pozorovat mikroskopem s fázovým kontrastem. Tato metoda je uváděna jako zlatý standard sérotypizace [18, 20]. Z imunologických testů je dále využívána například metoda latexové aglutinace [18].

Kromě imunologických testů k určování sérotypu jsou využívány i molekulárně genetické metody. Syntéza polysacharidového pouzdra je kódována v genovém *cps* (capsule polysaccharide synthesis) operonu [5]. Geny obsažené v *cps* operonu se liší dle jednotlivých sérotypů/séroskupin. Sérotypy, které jsou geneticky velmi příbuzné, řadíme do séroskupin (např. 6A, 6B, 6C a 6D). V rámci jedné séroskupiny se sérotypy liší pořadím obsažených genů a diskrétnějšími rozdíly na úrovni jednotlivých nukleotidů. Sérotyp můžeme molekulárně geneticky vyšetřit pomocí multiplexové polymerázové řetězové reakce (mPCR) či PCR v reálném čase (RT-PCR), využitím microarray metody a dále metodou celogenomové sekvenace (whole genome sequencing, WGS).

V České republice je sérotypizace pneumokoků prováděna v Národní referenční laboratoři pro streptokokové nákazy (NRL/STR). Vyšetření v případě dostupnosti kmene

proLékaře.cz | 30.11.2026

pneumokoka spočívá v kombinaci mPCR a Quellung reakce [21]. V případě typizace z klinického materiálu bez možnosti vyšetřit bakteriální kulturu se neprovádí Quellung reakce, ale pouze mPCR.

Protokol zahrnující Quellung reakci vyžaduje dostupnost čisté bakteriální kultury, zato ale umožňuje určit mnohem širší spektrum sérotypů ve srovnání s mPCR reakcí. Reakcí Quellung nelze ozřejmit pouze neobalené a tedy „netypovatelné" pneumokoky. Při typování pomocí mPCR není nutné mít k dispozici čistou kulturu, ale je výrazně limitován počet sérotypů, které lze takto určit, protože je omezený počet dostupných primerů k jednotlivým sérotypům. Molekulárně genetické metody umožňují typovat i neobalené pneumokoky, které sice nevytváří polysacharidové pouzdro, ale mohou v genomu mít geny pro tvorbu pouzdra.

STRATEGIE WGS V EVROPĚ

V roce 2008 vydalo European Centre for Disease Prevention and Control (ECDC) dokument Surveillance of communicable diseases in the European Union – A long-term strategy: 2008–2013, v němž je mimo jiné zdůrazněna důležitost zajištění kvalitních laboratorních metod, které jsou nezbytné k získání porovnatelných laboratorních dat [22]. Dále je v dokumentu uvedeno, že má být podporována standardizace a zavádění molekulárně genetické subtypizace a její kombinace s epidemiologickými a klinickými daty. Následoval dokument Surveillance of communicable diseases in Europe – a concept to integrate molecular typing data into EU-level surveillance z roku 2011 zdůrazňující a podporující zavádění molekulární typizace do surveillance infekčních onemocnění v Evropě [23].

V roce 2013 vyšel dokument Roadmap for integration of molecular typing into European-level surveillance and epidemic preparedness, který uvádí celogenomovou sekvenaci jako nástroj k charakterizaci patogenních mikroorganismů a k hodnocení jejich genetické různorodosti a krátkodobé evoluce [24, 25]. V dokumentu Expert opinion on whole genome sequencing for public health surveillance z roku 2016 je celogenomová sekvenace považována za referenční metodu typizace vzorků při outbreacích nahrazující klasické metody typizace [26]. Další ECDC dokumenty z následujících let již uvádí celogenomovou sekvenaci jako standardní nástroj k vyšetřování epidemií a k typizaci a zabývají se podporou zavádění metody v členských státech EU/EEA [27].

V dokumentu ECDC strategic framework for the integration of molecular and genomic typing into European surveillance and multi-country outbreak investigations – 2019–2021 jsou rozebrány prioritní patogeny, u nichž by měla být prováděna analýza pomocí celogenomové sekvenace [28]. Jednou z priorit je studium bakterie *Streptococcus pneumoniae* metodou celogenomové sekvenace. ECDC zde klade důraz na sentinelovou surveillance izolátů z invazivního pneumokokového onemocnění (IPO), na molekulární typizaci a hodnocení dopadu zaváděných vakcín proti pneumokokům na výskyt vakcinačních a nevakcinačních sérotypů, dále na identifikaci a monitorování úspěšných klonů, identifikaci změn v syntéze pouzdra a na zjišťování faktorů virulence spojených se závažným průběhem IPO.

REALIZACE WGS V PRAXI

Data získaná celogenomovou sekvenací prochází celou řadou bioinformatických analýz. Hodnotí se kvalita získaných dat, data se používají k sestavení genomu agens, následně se určuje rod a druh sekvenovaného agens, analyzuje se příbuznost, hodnotí se přítomnost jednotlivých genů a vyhledávají se genové mutace.

Prvotní formát dat získaný ze sekvenátoru se nazývá raw data neboli soubor s příponou fastq, kdy q na konci říká, že fastq soubor nese nejen sekvence daného genomu, ale též informaci o kvalitě sekvenování = quality. Z raw dat se následně skládá genom do formátu fasta souboru neboli „assembled" genom (volně přeloženo poskládaný genom). Programy na analýzu sekvenačních dat mohou při analýze vyžadovat vstupní data ve formátu fastq, fasta či lze využít oba formáty.

V případě pneumokoka je prioritou určit ze sekvenačních dat sérotyp, provést multilokusovou sekvenační typizaci (MLST), ribozomální MLST (rMLST), core genome MLST (cgMLST), whole genome MLST (wgMLST), single nucleotide polymorphism (SNP) analýzu, určení Global Pneumococcal Sequence Clusteru (GPSC), dále lze zkoumat přítomné geny virulence a geny antibiotické rezistence, přítomnost plasmidů atd.

GENOM PNEUMOKOKA

Pneumokoky mají genom dlouhý přibližně 1,9–2,1 Mb, tedy zhruba 2 miliony párů nukleotidů. Obsah bází cytosinu a guaninu (CG-content) je zhruba 39% [29].

Pneumokoky jsou stejně jako ostatní streptokoky obdařeny vysokou mírou plasticity genomu. Například při zavádění vakcín je na populaci pneumokoků vytvářen selekční tlak, na který reagují obměnou sérotypů (serotype replacement), obměnou pouzdra (capsular switch) či horizontálním genovým transportem [30].

Při setkání kmenů pneumokoka s dalšími streptokoky či jinými bakteriemi může dojít ke konjugaci a následně k horizontálnímu genovému transportu, kdy je prostřednictvím integrativních konjugativních elementů sdílena genetická informace mezi bakteriemi. Pneumokoky tak mohou získávat nové geny antibiotické rezistence, geny virulence, či může docházet k modifikaci sérotypu [31, 32].

proLékaře.cz | 30.11.2026

Další zdroj nové genetické informace pro pneumokoka, jakož i pro ostatní bakterie představují bakteriofágy, kdy se profág či jeho část může začlenit do genomu bakterie a pozměnit její vlastnosti, jak bylo popsáno v případě multidrug rezistentní linie *Streptococcus pneumoniae* Spain 23F/ST81 [33, 34].

SÉROTYPIZACE PNEUMOKOKA Z WGS DAT

Sérotypizace (přesněji řečeno genotypizace, ale v rámci zachování jednoty nomenklatury v článku bude nadále používán termín sérotypizace) je jednou z hlavních analýz, kterým jsou sekvenační data pneumokoků podrobována. První práce využívající sérotypizaci ze sekvenačních dat využívaly mapování sekvencí genů v *cps* operonu s dříve publikovanými referenčními sekvencemi [32, 35]. V roce 2014 byl publikován návod k sérotypizaci přímo ze sekvenačních dat [36]. O dva roky později byl publikován první bioinformatický automatický sérotypizační nástroj PneumoCaT (Pneumococcal Capsule Typing), který je dostupný volně ke stažení [37]. Proces určení sérotypu má dva kroky, kdy v prvním kroku je porovnána „přečtená" sekvence pneumokoka s 92 referenčními sekvencemi sérotypů, pokud je sekvence stoprocentně shodná s referencí, proces sérotypizace je ukončen a sérotyp je úspěšně určen. Pokud není shoda s referencí úplná, ve 2. kroku jsou hodnoceny SNPs a je rozlišováno mezi sérotypy v rámci jedné séroskupiny. Takto je možno určit 89 z 94 tehdy známých sérotypů a určit kmen na úroveň séroskupiny pro séroskupiny 24 a 32. Zdokonalení sérotypizace přinesl nástroj SeroBa, který umožňuje určit vyšší počet sérotypů, a to v kratším čase a s využitím menší kapacity počítače [38]. Nástroje PneumoCaT i SeroBA předpokládají, že byla sekvenována čistá kultura pneumokoka. Nově jsou dostupné i nástroje k typizaci vzorků, kdy lze očekávat více sérotypů pneumokoka v jednom vzorku, například při testování vzorků z nosičství. Takovými volně dostupnými programy jsou SeroCall a PneumoKITy (Pneumococcal K-mer Integrated Typing), přičemž PneumoKITy je první program umožňující sérotypizaci jak z raw dat, tak z fasta souboru [39, 40].

Výše zmíněné nástroje jsou primárně určeny pro sekvenační data získaná prostřednictvím sekvenační platformy Illumina, kromě programu PneumoKITy, který umí zpracovat i fasta soubory. Základní algoritmy všech výše zmíněných programů se do jité míry liší, všechny ale využívají stejný postup porovnávající sekvenci *cps* lokusu vzorku a referenční databázi *cps* lokusů pro různé sérotypy.

Nevýhodou výše uvedených nástrojů je nutnost získat kvalitní sekvenační data s vysokou coverage. Coverage udává, kolikrát byl který úsek genomu sekvenován, u bakteriálního genomu je cílem coverage vyšší než 20, v ideálním případě je dobré získávat coverage 30 a více – čím vyšší coverage, tím spolehlivější jsou získaná data. Programem řešícím situaci sekvenačních dat s nedostatečnou coverage je PfaSTer [41], který využívá jako vstupní data fasta soubor. Predikce sérotypu pak probíhá na základě sekvenačních dat a strojového učení. Program PfaSTer potřebuje mít k dispozici kvalitně sekvenované genomy jednotlivých sérotypů, jeho limitací je, že nabídka kvalitních sekvencí u vzácných sérotypů je stále omezená, a proto lze pomocí programu spolehlivě určit 65 nejčastějších sérotypů.

MLST

MLST je molekulárně genetická metoda umožňující relativně snadné hodnocení příbuznosti izolátů jednoho druhu na lokální i na mezinárodní úrovni [42]. Hodnocení příbuznosti je prováděno na základě srovnání krátkých fragmentů sedmi alel tzv. „housekeeping" genů. U izolátů je identifikována sada těchto sedmi alel a dle alelického profilu je jim přidělen sekvenční typ (ST). Zavedení MLST u pneumokoků umožnilo spolehlivější srovnávání příbuznosti izolátů než do té doby používané metody PFGE (Pulsed field gel electrophoresis) či MLEE (multilocus enzyme electrophoresis) [43]. U pneumokoků lze pozorovat, že v rámci jednotlivých ST či v clusterech příbuzných ST je často sdílen stejný sérotyp, nejedná se však o pravidlo bez výjimek, což může být dáno nedostatečnou diskriminační schopností MLST.

rMLST

Základ systematiky a fylogenetické analýzy bakterií pomocí ribozomálních genů byl položen v 80. letech, kdy začala být používána nukleotidová sekvence malé podjednotky (16S) ribozomální RNA neboli 16S rRNA [44]. Sekvence genu pro 16S rRNA je univerzálně přítomna v bakteriálním genomu, je dostatečně konzervovaná v čase a je snadno sekvenovatelná. Taxonomické a evoluční studie 16S rRNA hrály klíčovou roli při zavedení tří domén života – archea, bakterie, eukaryota [45].

Se zavedením celogenomové sekvenace došlo k rozšíření počtu využívaných ribozomálních genů. V rámci ribozomálního MLST (rMLST) je stanovováno 53 *rps* genů [46]. Geny *rps* mají dostatečně konzervované sekvence, což umožňuje taxonomické rozlišení na úroveň rodu a druhu, a zároveň jsou dostatečně variabilní, aby bylo možné typizační rozlišení izolátů stejného druhu. Ribozomální geny jsou rozmístěny v různých částech bakteriálního chromozomu a jedná se o protein-kódující geny.

Profil alelových variant ribozomálních genů izolátu poskytne ribozomální sekvenční typ (rST) a umožní přiřadit izolát k rodu a druhu. Analýza je prováděna v programu Bacterial Isolate Genome Sequence Database (BIGSdb) na platformě PubMLST (https://pubmlst.org/) [47, 48].

Praktické využití rMLST se ukazuje v případě pochybností o správné identifikaci rozporuplně fenotypicky reagujících pneumokoků a viridujících streptokoků [49].

proLékaře.cz | 30.11.2026

cgMLST

Při určování core genome MLST (cgMLST) probíhá srovnání alel všech tzv. core genů (obvykle stovky či tisíce genů), které jsou v čase konzervované a společné všem či téměř všem izolátům jednoho druhu [50]. cgMLST poskytuje významné rozšíření původního sedmilokusového MLST. cgMLST lze využít k taxonomickému určování izolátů, k vyšetřování nozokomiálních nákaz, infekčních ohnisek a clusterů či k analýze příbuznosti a původu izolátů. Takto byl například zkoumán původ a šíření pneumokoka sérotypu 19A/ST416 v České republice a v zemích Evropy [51]. cgMLST pneumokoka zahrnuje celkem 1 222 tzv. core genů. Určování cgMLST je možné prostřednictvím platformy PubMLST (https://pubmlst.org/) [48].

wgMLST

Celogenomovou sekvenací a přiřazením složeného genomu izolátu k referenční sekvenci zjistíme přítomnost nejen core genů, ale určíme i tzv. accessory geny. Určený počet genů se tak značně rozšiřuje a pro podrobnější analýzy příbuznosti mezi izoláty lze využít přístup porovnávání gen-za-genem (gene-by-gene) neboli MLST celého genomu (whole genome MLST, wgMLST) [52].

Příkladem využití této strategie hodnocení příbuznosti je srovnání sekvenačních dat skupiny invazivních pneumokoků o vybraných sérotypech na základě všech v té době dostupných genů v PubMLST databázi [53].

SNP

Jiný pohled na příbuznost izolátů jednoho druhu nabízí metoda single nucleotide polymorphism (SNP) analýzy [54]. SNP je jednonukleotidová mutace a analýza těchto mutací benefituje mimo jiné z toho, že jsou hodnoceny nejen protein-kódující části genomu, ale i proteiny-nekódující oblasti. Provádí se mapováním sekvenačních dat izolátu proti referenčnímu genomu. SNP analýza může sloužit k identifikaci bakterií, k hodnocení genetické diverzity, fylogenetickým analýzám, k účelům surveillance a vyšetřování ohnisek infekce, k podrobným analýzám na úrovni genových operonů a jednotlivých genů [55, 56].

GPSC

V roce 2009 vznikly první návrhy na projekt, jenž by umožňoval genomickou surveillance pneumokoků na mezinárodní úrovni. Postupně vznikl Global Pneumococcal Sequencing project (GPS). Prvotním cílem projektu mělo být porozumění mechanismům antibiotické rezistence, postupně bylo těžiště zkoumání přesunuto ke strategiím úniku pneumokoků před vakcinací. Vznikla databáze >20 000 genomů pneumokoků z celého světa. V rámci projektu byla zavedena nová typizační metoda založená na definování pneumokokových kmenů pomocí variací v genomu. Pomocí typizační pipeline je určován Global Pneumococcal Sequence Cluster (GPSC).

Typizace probíhá pomocí programu PopPUNK (Population Partitioning Using Nucleotide K-mers), což je software umožňující populační analýzy izolátů a jejich rozdělení do clusterů [57]. Hodnocení příbuznosti či odlišnosti izolátů je prováděno ve sdílených úsecích genomu a v obsažených genech, a to s použitím porovnávání k-merů o různé délce. V každé jedné linii GPSC jsou obsaženy jak vakcinační, tak nevakcinační sérotypy.

Analýzou GPSC databáze bylo zjištěno, že existují linie pneumokoků, které jsou více rezistentní k antibiotikům, jiné linie mají vyšší potenciál vyvolávat invazivní onemocnění, a to bez ohledu na sérotyp [58]. Zavedení GPSC umožnilo clusterování pneumokoků a napomohlo k jejich srovnávání na mezinárodní úrovni.

Geny virulence

V genomu pneumokoka jsou přítomny geny virulence a ostrovy patogenicity (pathogenicity islands), které zodpovídají za interakci s hostitelem a za různou míru invazivity jednotlivých izolátů [59]. Mezi nejvýznamnější faktory virulence patří pouzdro na povrchu bakterie. Dále jsou to pily umožňujících úchyt bakterie na sliznici hostitele a potencující imunitní odpověď hostitele [60].

Vyhledávání genů virulence pneumokoků je možné pomocí databází referenčních genů virulence – například ve Virulence Factor Database (VFDB, http://www.mgc.ac.cn/VFs/) či v online bioinformatické platformě Bacterial and Viral Bioinformatics Resource Center (BV-BRC, https://www.bv-brc.org/), která sdružuje databáze VFDB, Victor a spravuje svou vlastní databázi genů virulence [61, 62, 63].

Geny antibiotické rezistence

Monitorování antibiotické rezistence představuje důležitou součást surveillance IPO. Od 70. let se začaly objevovat studie dokumentující šíření pneumokoků, které vykazovaly sníženou citlivost a rezistenci k penicilinu, tetracyklinu, erythromycinu, klindamycinu, rifampicinu a chloramfenikolu [64].

K analýze sekvenačních dat a k vyhledávání genů antibiotické rezistence lze využít online databázi Comprehensive Antibiotic Resistance Database (CARD, https://card.mcmaster.ca/home), dále platformu BV-BRC, která umí vyhledávat geny antibiotické rezistence obsažené v databázi CARD a v databázi PATRIC [65, 62, 66]. Platforma BV-BRC kromě stanovení přítomnosti genů antibiotické rezistence předpovídá i fenotypický projev testovaného pneumokoka. Dalším online nástrojem je PathogenWatch (https://pathogen.watch/), který kromě jiných funkcí stanovuje přítomnost genů antibiotické rezistence pneumokoků a předpovídá jejich vliv na fenotypický projev daného pneumokoka – tedy jeho fenotypickou antibiotickou rezistenci [67].

proLékaře.cz | 30.11.2026

ZÁVĚR

Celogenomová sekvenace (WGS) rozšiřuje hranice poznání mikroorganismů. Získaná data lze snadno mezinárodně sdílet a pomocí bioinformatických nástrojů analyzovat.

V článku popsané metody představují pouze základní kroky analýzy sekvenačních dat pneumokoků. Možnosti analýzy jsou vzhledem k objemu celogenomových dat a rozšíření WGS ve světě rozsáhlé a vždy záleží na zkoumané problematice, jaké další metody a postupy je třeba zvolit.

Analýza pneumokoků metodou WGS představuje zlepšení v provádění surveillance IPO, kdy je sérotyp určován molekulárně geneticky, jsou prováděny další podrobnější typizace, získaná data lze snadno mezinárodně porovnávat a lze lépe hodnotit účinnost vakcinačních programů.

V některých státech probíhá surveillance IPO z velké části pomocí celogenomové sekvenace a lze očekávat rutinní celogenomové sekvenování v reálném čase ve stále širším spektru mikrobiologických laboratoří.

LITERATURA

1. Arbique JC, Poyart C, Trieu-Cuot P, et al. Accuracy of phenotypic and genotypic testing for identification of Streptococcus pneumoniae and description of Streptococcus pseudopneumoniae sp. nov. *J Clin Microbiol.*, 2004;42(10):4686–4696. doi: 10.1128/JCM.42.10.4686-4696.2004. PMID: 15472328; PMCID: PMC522306.
2. Wasilauskas BL, Hampton KD. An analysis of Streptococcus pneumoniae identification using biochemical and serological procedures. *Diagn Microbiol Infect Dis*., 1984;2(4):301–307. doi: 10.1016/0732-8893(84)90061-0. PMID: 6488747.
3. Kellogg JA, Bankert DA, Elder CJ, et al. Identification of Streptococcus pneumoniae revisited. *J Clin Microbiol*., 2001;39(9):3373–3375. doi: 10.1128/JCM.39.9.3373-3375.2001. PMID: 11526182; PMCID: PMC88350.
4. Geno KA, Gilbert GL, Song JY, et al. Pneumococcal Capsules and Their Types: Past, Present, and Future. *Clin Microbiol Rev*., 2015;28(3):871–899. doi: 10.1128/CMR.00024-15. Erratum in: *Clin Microbiol Rev*., 2020; 34(2): PMID: 26085553; PMCID: PMC4475641.
5. Bentley SD, Aanensen DM, Mavroidi A, et al. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet*., 2006;2(3):e31. doi: 10.1371/journal.pgen.0020031. Epub 2006 Mar 10. PMID: 16532061; PMCID: PMC1391919.
6. Ganaie F, Saad JS, McGee L, et al. A New Pneumococcal Capsule Type, 10D, is the 100th Serotype and Has a Large cps Fragment from an Oral Streptococcus. *mBio*., 2020;11(3):e00937–20. doi: 10.1128/mBio.00937-20. PMID: 32430472; PMCID: PMC7240158.
7. Global Pneumococcal Sequencing Project (GPS) | Serotypes. [online]. [cit. 2023-08-10]. Dostupné na www: https://www.pneumogen.net/gps/serotypes.html.
8. De Lencastre H, Kristinsson KG, Brito-Avô A, et al. Carriage of respiratory tract pathogens and molecular epidemiology of Streptococcus pneumoniae colonization in healthy children attending day care centers in Lisbon, Portugal. *Microb Drug Resist*., 1999;5(1):19–29. doi: 10.1089/mdr.1999.5.19. PMID: 10332718.
9. Kellner JD, Ford-Jones EL. Streptococcus pneumoniae carriage in children attending 59 Canadian child care centers. Toronto Child Care Centre Study Group. *Arch Pediatr Adolesc Med*., 1999;153(5):495–502. doi: 10.1001/archpedi.153.5.495. PMID: 10323630.
10. Tvedskov ESF, Hovmand N, Benfield T, et al. Pneumococcal carriage among children in low and lower-middle-income countries: A systematic review. *Int J Infect Dis*., 2022;115:1–7. doi: 10.1016/j.ijid.2021.11.021. Epub 2021 Nov 18. PMID: 34800691.
11. Vančíková Z, Trojánek M, Zemličková H, et al. Pneumococcal urinary antigen positivity in healthy colonized children: is it age dependent? *Wien Klin Wochenschr*., 2013;125(17–18):495–500. doi: 10.1007/s00508-013-0405-4. Epub 2013 Aug 9. PMID: 23928934.
12. Hussain M, Melegaro A, Pebody RG, et al. A longitudinal household study of Streptococcus pneumoniae nasopharyngeal carriage in a UK setting. *Epidemiol Infect*., 2005;133(5):891–898. doi: 10.1017/S0950268805004012. PMID: 16181510; PMCID: PMC2870321.
13. Arguedas A, Trzciński K, O'Brien KL, et al. Upper respiratory tract colonization with Streptococcus pneumoniae in adults. *Expert Rev Vaccines*., 2020;19(4):353–366. doi: 10.1080/14760584.2020.1750378. Epub 2020 Apr 17. PMID: 32237926.
14. Bogaert D, De Groot R, Hermans PW. Streptococcus pneumoniae colonisation: the key to pneumococcal disease. *Lancet Infect Dis*., 2004;4(3):144–154. doi: 10.1016/S1473-3099(04)00938-7. PMID: 14998500.
15. Simell B, Auranen K, Käyhty H, et al. The fundamental link between pneumococcal carriage and disease. *Expert Rev Vaccines*, 2012;11(7):841–855. doi: 10.1586/erv.12.53. PMID: 22913260.
16. Platt H, Omole T, Cardona J, et al. Safety, tolerability, and immunogenicity of a 21-valent pneumococcal conjugate vaccine, V116, in healthy adults: phase 1/2, randomised, double-blind, active comparator-controlled, multicentre, US-based trial. *Lancet Infect Dis*., 2023;23(2):233–246. doi: 10.1016/S1473-3099(22)00526-6. Epub 2022 Sep 15. PMID: 36116461.
17. Chichili GR, Smulders R, Santos V, et al. Phase 1/2 study of a novel 24-valent pneumococcal vaccine in healthy adults aged 18 to 64 years and in older adults aged 65 to 85 years. *Vaccine*, 2022;40(31):4190–4198. doi: 10.1016/j.vaccine.2022.05.079. PMID: 35690500.
18. Jauneikaite E, Tocheva AS, Jefferies JM, et al. Current methods for capsular typing of Streptococcus pneumoniae. *J Microbiol Methods*, 2015;113:41–49. doi: 10.1016/j.mimet.2015.03.006. Epub 2015 Mar 25. PMID: 25819558.
19. Neufeld F. Ueber die Agglutination der Pneumokokken und uber die Theorien der Agglutination. *Zeitschrift für Hygiene und Infektionskrankheiten*, 1902;40:54–72. doi:10.1007/bf02140530. S2CID 1143320.
20. Merrill CW, Gwaltney JM Jr., Hendley JW, et al. Rapid identification of pneumococci. Gram stain vs. the quellung reaction. *N Engl J Med*., 1973;288:510–512.
21. Vacková Z, Klímová M, Kozáková J. Nová molekulární metoda a schéma typizace Streptococcus pneumoniae v České republice. *Zprávy Epidemiologie a Mikrobiologie*, 2013; 22(1):16–18.
22. European Centre for Disease Prevention and Control. Surveillance of communicable diseases in the European Union – A long-term strategy: 2008–2013. Stockholm: ECDC; 2008.
23. European Centre for Disease Prevention and Control. Surveillance of communicable diseases in Europe – a concept to integrate molecular typing data into EU-level surveillance. Stockholm: ECDC; 2013.
24. European Centre for Disease Prevention and Control. Roadmap for integration of molecular typing into European – level surveillance and epidemic preparedness – Version 1.2, 2013. Stockholm: ECDC; 2016.
25. European Centre for Disease Prevention and Control. ECDC roadmap for integration of molecular and genomic typing into European-level surveillance and epidemic preparedness – Version 2.1, 2016-19. Stockholm: ECDC; 2016.
26. European Centre for Disease Prevention and Control. Expert opinion on whole genome sequencing for public health surveillance. Stockholm: ECDC; 2016.
27. ECDC public health microbiology strategy 2018–2022.
28. European Centre for Disease Prevention and Control. ECDC strategic framework for the integration of molecular and genomic typing into European surveillance and multi-country outbreak investigations – 2019–2021. Stockholm: ECDC; 2019.

proLékaře.cz | 30.11.2026

29. Tagini F, Greub G. Bacterial genome sequencing in clinical microbiology: a pathogen-oriented review. *Eur J Clin Microbiol Infect Dis*., 2017;36(11):2007–2020. doi: 10.1007/s10096-017-3024-6. PMID: 28639162; PMCID: PMC5653721.
30. Andam CP, Hanage WP. Mechanisms of genome evolution of Streptococcus. *Infect Genet Evol*., 2015;33:334–342. doi: 10.1016/j.meegid.2014.11.007. PMID: 25461843; PMCID: PMC4430445.
31. Coffey TJ, Enright MC, Daniels M, et al. Recombinational exchanges at the capsular polysaccharide biosynthetic locus lead to frequent serotype changes among natural isolates of Streptococcus pneumoniae. *Mol Microbiol*., 1998;27(1):73–83. doi: 10.1046/j.1365-2958.1998.00658.x. PMID: 9466257.
32. Croucher NJ, Harris SR, Fraser C, et al. Rapid pneumococcal evolution in response to clinical interventions. *Science*, 2011;331(6016):430–434. doi: 10.1126/science.1198545. PMID: 21273480; PMCID: PMC3648787.
33. Gindreau E, López R, García P. MM1, a temperate bacteriophage of the type 23F Spanish/USA multiresistant epidemic clone of Streptococcus pneumoniae: structural analysis of the site-specific integration system. *J Virol*., 2000;74(17):7803–7813. doi: 10.1128/jvi.74.17.7803-7813.2000. PMID: 10933687; PMCID: PMC112310.
34. Croucher NJ, Walker D, Romero P, et al. Role of conjugative elements in the evolution of the multidrug-resistant pandemic clone Streptococcus pneumoniae Spain 23F ST81. *J Bacteriol*., 2009;191(5):1480–1489. doi: 10.1128/JB.01343-08. PMID: 19114491; PMCID: PMC2648205.
35. Everett DB, Cornick J, Denis B, et al. Genetic characterisation of Malawian pneumococci prior to the roll-out of the PCV13 vaccine using a high-throughput whole genome sequencing approach. *PLoS One*, 2012;7(9):e44250. doi: 10.1371/journal.pone.0044250. PMID: 22970189; PMCID: PMC3438182.
36. Liyanapathirana V, Ang I, Tsang D, et al. Application of a target enrichment-based next-generation sequencing protocol for identification and sequence-based prediction of pneumococcal serotypes. *BMC Microbiol*., 2014;14:60. doi: 10.1186/1471-2180-14-60. PMID: 24612771; PMCID: PMC3984734.
37. Kapatai G, Sheppard CL, Al-Shahib A, et al. Whole genome sequencing of Streptococcus pneumoniae: development, evaluation and verification of targets for serogroup and serotype prediction using an automated pipeline. *PeerJ*., 2016;4:e2477. doi: 10.7717/peerj.2477. PMID: 27672516; PMCID: PMC5028725.
38. Epping L, van Tonder AJ, Gladstone RA, et al. SeroBA: rapid high-throughput serotyping of Streptococcus pneumoniae from whole genome sequence data. *Microb Genom*., 2018;4(7):e000186. doi: 10.1099/mgen.0.000186. Erratum in: *Microb Genom*., 2018;4(8): PMID: 29870330; PMCID: PMC6113868.
39. Knight JR, Dunne EM, Mulholland EK, et al. Determining the serotype composition of mixed samples of pneumococcus using whole-genome sequencing. *Microb Genom*. 2021;7(1):mgen000494. doi: 10.1099/mgen.0.000494. PMID: 33355528; PMCID: PMC8115901.
40. Sheppard CL, Manna S, Groves N, et al. PneumoKITy: A fast, flexible, specific, and sensitive tool for Streptococcus pneumoniae serotype screening and mixed serotype detection from genome sequence data. *Microb Genom*., 2022;8(12):mgen000904. doi: 10.1099/mgen.0.000904. PMID: 36748701; PMCID: PMC9837567.
41. Lee JT, Li X, Hyde C, et al. PfaSTer: a machine learning-powered serotype caller for Streptococcus pneumoniae genomes. *Microb Genom*., 2023;9(6). doi: 10.1099/mgen.0.001033. PMID: 37279053.
42. Maiden MCJ, Bygraves JA, Feil E, et al. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A*, 1998;95(6):3140–3145. doi: 10.1073/pnas.95.6.3140. PMID: 9501229.
43. Enright MC, Spratt BG. A multilocus sequence typing scheme for Streptococcus pneumoniae: identification of clones associated with serious invasive disease. *Microbiology (Reading)*, 1998;144 (Pt 11):3049–3060. doi: 10.1099/00221287-144-11-3049. PMID: 9846740.
44. Woese CR. Bacterial evolution. *Microbiol Rev*., 1987;51(2):221–271. doi: 10.1128/mr.51.2.221-271.1987. PMID: 2439888; PMCID: PMC373105.
45. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A*, 1990;87(12):4576–4579. doi: 10.1073/pnas.87.12.4576. PMID: 2112744; PMCID: PMC54159.
46. Jolley KA, Bliss CM, Bennett JS, et al. Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. *Microbiology*, 2012;158(4):1005–1015. doi: 10.1099/mic.0.055459-0. PMID: 22282518.
47. Jolley KA, Maiden MC. BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics*, 2010;11:595. doi: 10.1186/1471-2105-11-595. PMID: 21143983; PMCID: PMC3004885.

Poděkování
Podpořeno z programového projektu Ministerstva zdravotnictví ČR s registračním číslem NU22-09-00433.

Do redakce došlo 11. 8. 2023.

Adresa pro korespondenci:
MUDr. Sandra Vohrnová
Národní referenční laboratoř pro streptokokové nákazy,
Oddělení bakteriálních vzdušných nákaz,
Centrum epidemiologie a mikrobiologie
Státní zdravotní ústav
Šrobárova 49/48
100 00 Praha 10
e-mail: sandra.vohrnova@szu.cz