

Univerzita Pardubice
Fakulta chemicko-technologická

Analýza struktury a funkce genomu patogenních bakterií

2025

Darja Vymazalová

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Darja Vymazalová**
Osobní číslo: **C22272**
Studijní program: **B0914P360019 Laboratorní diagnostika ve zdravotnictví**
Téma práce: **Analýza struktury a funkce genomu patogenních bakterií**
Téma práce anglicky: **Structure and Function Analysis of Pathogenic Bacteria Genome**
Zadávací katedra: **Katedra biologických a biochemických věd**

Zásady pro vypracování

1. Vypracujte literární rešerši o analýze struktury a funkci genomu patogenních bakterií.
2. V práci se zaměřte na popis patogenních bakterií a jejich vybraných zástupců, definujte, co je to genom, jakou má strukturu a funkci a v poslední části práce se zaměřte na analýzu struktury a funkci genomu u vybraných patogenních bakterií.
3. Ke studiu tématiky použijte hlavně zahraniční odbornou literaturu dostupnou v elektronických vědeckých databázích, jako jsou např. *NCBI Pubmed*, *ScienceDirect*, *Web of Science*, apod. Informace přehledně zpracujte podle pokynů a doporučení školitele.

Rozsah pracovní zprávy: **25 s.**
Rozsah grafických prací: **dle potřeby**
Forma zpracování bakalářské práce: **tištěná**

Seznam doporučené literatury:

Podle pokynů vedoucího bakalářské práce.

Vedoucí bakalářské práce: **Mgr. Barbora Jankovičová, Ph.D.**
Katedra biologických a biochemických věd
Konzultant bakalářské práce: **Ing. Lucie Michalcová**
Katedra biologických a biochemických věd

Datum zadání bakalářské práce: **20. prosince 2024**

Termín odevzdání bakalářské práce: **1. července 2025**

L.S.

prof. Ing. Petr Němec, Ph.D. v.r.
děkan

prof. RNDr. Tomáš Roušar, Ph.D. v.r.
vedoucí katedry

V Pardubicích dne 28. února 2025

Prohlášení

Prohlašuji: Práci s názvem Analýza struktury a funkce genomu patogenních bakterií jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Beru na vědomí, že v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů, a směrnicí Univerzity Pardubice č. 7/2019 Pravidla pro odevzdávání, zveřejňování a formální úpravu závěrečných prací, ve znění pozdějších dodatků, bude práce zveřejněna prostřednictvím Digitální knihovny Univerzity Pardubice.

V Pardubicích dne: 26.06.2025

Darja Vymazalová v.r.

PODĚKOVÁNÍ:

Tímto bych chtěla poděkovat Mgr. Barboře Jankovičové, Ph.D. za cenné rady, připomínky, ochotu a trpělivost při zpracování bakalářské práce. Poděkování také patří Ing. Lucii Michalcové, která moji bakalářskou práci začala vést.

ANOTACE

Tato bakalářská práce se věnuje analýze struktury a funkce genomu u vybraných patogenních bakterií. V úvodní části jsou popsány základní vlastnosti těchto mikroorganismů. Následně jsou shrnuty vybrané metody sekvenování a jejich využití při studiu genomu bakterií.

KLÍČOVÁ SLOVA

genom, patogenní bakterie, sekvenování, Illumina, Ion Torrent

TITLE

Analysis of the structure and function of pathogenic bacteria genomes

ANNOTATION

This bachelor thesis aims to analyze structure and function of genome within chosen pathogenic bacteria. In the preamble basic properties of these microorganisms are described. Subsequently, selected sequencing methods and their application in the study of bacteria genomes are summarized.

KEYWORDS

genome, pathogenic bacteria, sequencing, Illumina, Ion Torrent

Obsah

Seznam ilustrací	9
Seznam zkratk a značek	10
Úvod	12
1. Bakterie	13
1.1 Rod <i>Staphylococcus</i>	14
1.1.1 <i>Staphylococcus aureus</i>	14
1.1.2 <i>Staphylococcus epidermidis</i>	15
1.2 Rod <i>Streptococcus</i>	15
1.2.1 <i>Streptococcus pyogenes</i>	16
1.2.2 <i>Streptococcus pneumoniae</i>	16
1.3 Rod <i>Pseudomonas</i>	17
1.3.1 <i>Pseudomonas aeruginosa</i>	17
2. Genom	18
2.1 Struktura genomu	18
2.2 Funkce genomu	19
3. Analýza struktury a funkce genomu	20
3.1 Metody sekvenování	20
3.2 Sekvenování první generace	21
3.2.1 Sekvenování dle Sangera	22
3.2.2 Sekvenování dle Maxama-Gilberta	23
3.3 Sekvenování druhé generace	24
3.3.1 Pyrosekvenování	24
3.3.2 Illumina	26
3.3.3 Ion Torrent	27
3.4 Sekvenování třetí generace	29
3.4.1 SMRT	30
3.5 Sekvenování čtvrté generace	31
3.6 Shluková analýza	31
Závěr	36
Seznam informačních zdrojů	37

Seznam ilustrací

Obr. 1: Trojrozměrná organizace genomu.....	19
Obr. 2: Technologie sekvenování DNA první generace.....	23
Obr. 3: Postup pyrosekvenování.....	25
Obr. 4: Proces sekvenování pomocí metody Illumina.....	27
Obr. 5: Schéma mikročipu a zpracování signálu technikou Ion Torrent.....	28
Obr. 6: Replikační proces sekvenování technikou SMRT.....	30
Obr. 7: Dendrogram homologních chromozomů genomů <i>S. aureus</i>	33
Obr. 8: Dendrogram izolátu z populace streptokoků.....	34
Obr. 9: Dendrogram izolátu z populace pseudomonád	35

Seznam zkratek a značek

A: adenin

ASLO: antistreptolysin O

ATP: adenosintrifosfát

C: cytosin

CCD: detektor citlivý na elektrický náboj (z angl. charge-coupled device)

ddNTP: dideoxynukleotidtrifosfáty (z angl. dideoxynucleoside triphosphates)

DNA: deoxyribonukleová kyselina, zřídka i DNK (z angl. deoxyribonucleic acid)

dNTP: deoxynukleotidtrifosfáty (z angl. deoxynucleotide triphosphates)

G: guanin

MRSA: methicilin- rezistentní *Staphylococcus aureus*

NGS: sekvenování nové generace (z angl. next-generation sequencing)

PCR: polymerázová řetězová reakce (z angl. polymerase chain reaction)

PGM: osobní genomový přístroj (z angl. personal genome machine)

RNA: ribonukleová kyselina (z ang. ribonucleic acid)

S. aureus: *Staphylococcus aureus*

S. epideirmidis: *Staphylococcus epidermidis*

S. pneumoniae: *Streptococcus pneumoniae*

S. pyogenes: *Streptococcus pyogenes*

SAB: bakteriémie *Staphylococcus aureus* (z angl. *Staphylococcus aureus* bacteremia)

SMRT: sekvenování jednotlivých molekul v reálném čase (z angl. single molecule real- time sequencing)

SMS: sekvenování jednotlivých molekul (z angl. single-molekule sequencing)

T: thymin

TAD: topologicky asociované domény

ZMW: vlnovody nulového módu (z angl. zero-mode waveguides)

Úvod

Patogenní bakterie představují závažný zdravotnický problém, neboť jsou schopné vyvolávat široké spektrum infekčních onemocnění u lidí i zvířat. Schopnost těchto mikroorganismů způsobovat onemocnění (patogenita) je výsledkem specifických genetických vlastností, které jsou zakódovány v jejich genomu. Studium bakteriálních genomů je proto klíčové pro pochopení mechanismů infekce, přenosu a rezistence vůči antibiotikům.

Tato práce se zaměřuje na studium genomu u vybraných patogenních bakterií, konkrétně rodů *Staphylococcus*, *Streptococcus* a *Pseudomonas*. Tyto bakterie patří mezi klinicky významné patogeny, které mohou způsobovat jak běžné komunitní infekce, tak závažné nozokomiální nákazy. Znalost jejich genetické výbavy je nezbytná nejen pro vývoj nových diagnostických a terapeutických přístupů, ale i pro sledování epidemiologických trendů a efektivní kontrolu šíření infekcí.

V úvodní části práce jsou popsány základní vlastnosti uvedených mikroorganismů a jejich význam v medicíně. Dále jsou shrnuty metody sekvenování DNA, které umožňují detailní analýzu genomu těchto patogenů. Práce se věnuje jak tradičním metodám sekvenování, tak i moderním přístupům, včetně technologií třetí a čtvrté generace, které přinášejí nové možnosti v oblasti mikrobiální genomiky.

Sekvenování bakteriálních genomů dnes představuje nedílnou součást mikrobiologického výzkumu i klinické praxe. Tato práce si klade za cíl přiblížit význam a praktické využití těchto poznatků a metod v kontextu současné biomedicíny.

1. Bakterie

Bakterie jsou jednobuněčné prokaryotické organismy s poměrně jednoduchou strukturou. Jejich tvar je kokovitý (kulatý) nebo tyčinkovitý. Podle složení buněčné stěny je můžeme rozdělit na grampozitivní a gramnegativní. Grampozitivní bakterie mají buněčnou stěnu složenou ze silné vrstvy peptidoglykanů, gramnegativní bakterie mají buněčnou stěnu složenou z tenké vrstvy peptidoglykanů a tu překrytou vnější membránou. Bakterie jsou v přírodě nenahraditelné, ve významu koloběhu látek, symbiózy s člověkem či pozitivním vlivem na ekosystém. Vedle těchto prospěšných bakterií jsou některé bakterie spojovány s nemocemi neboli patogenitou. Na základě toho lze bakterie obecně rozdělit na nepatogenní, oportunně patogenní a primárně patogenní [1].

Primární patogeny napadají zdravého jedince, zatímco oportunní patogeny vyvolají infekci jen pokud je poškozen obranný mechanismus jedince, s čímž samozřejmě souvisí oslabená imunita. Nepatogenní mikroorganismy jsou nekonfliktní, nezpůsobují onemocnění ani jiné škodlivé reakce a můžeme sem zařadit mikroorganismy vzdušné flóry nebo normální mikroflóry (například na kůži, sliznicích či ve střevě). Z oportunně patogenních bakterií se v této práci zaměřím na *Pseudomonas aeruginosa*. Primárně patogenních bakterií není mnoho a onemocnění jimi způsobeným se snažíme zabránit očkováním, v této práci se více zaměřím na rod *Staphylococcus* a rod *Streptococcus* [2].

Patogenita bakterií je schopnost vyvolat onemocnění u hostitele. Mírou patogenity jsou faktory virulence, které vyjadřují, do jaké míry dokáže mikroorganismus hostitele poškodit, jak dobře dokáže unikat imunitní obraně hostitele a jak rychle se dokáže replikovat. Mezi virulentní faktory řadíme vylučované látky mikroorganismu, jako jsou toxiny a enzymy, ale také struktury buněčného povrchu jako jsou kapsle nebo lipopolysacharidy. Další faktory virulence jsou adheziny, které umožňují přichycení k hostiteli, a invaziny zajišťující schopnost vstoupit do organismu. Přítomnost těchto faktorů bývá podmíněna specifickými geny, které jsou často součástí tzv. ostrovů virulence [3].

Zkoumání bakterií výrazně rozvinuli Robert Koch a Louis Pasteur. Louis Pasteur dokázal v roce 1859, že kvašení způsobují bakterie, a také zastával názor, že bakterie jsou příčinou nemocí. Toto tvrzení potvrdil Robert Koch, který studoval původce cholery

a tuberkulózy a následně prokázal, že bakterie opravdu způsobují tuberkulózu. Za tento objev v roce 1905 obdržel Nobelovu cenu [1].

1.1 Rod *Staphylococcus*

Stafylokoky řadíme mezi grampozitivní bakterie, kulovitého tvaru. Pokud je budeme pozorovat pod mikroskopem, uvidíme shluky ve tvaru hroznů. Nejčastěji se vyskytují v nose a na kůži člověka. Nejrozšířenější druhy, na které se zaměřím v této práci, jsou *Staphylococcus aureus* (*S. aureus*) a *Staphylococcus epidermidis* (*S. epidermidis*) [4].

1.1.1 *Staphylococcus aureus*

Tuto bakterii najdeme nejčastěji v nosní dutině nebo na kůži. *S. aureus* řadíme mezi fakultativně anaerobní bakterie. Při kultivaci na bohatém živném médiu tvoří *S. aureus* poměrně velké žluté kolonie. Při kultivaci na krevním agaru způsobuje často hemolýzu [4].

Projevy patogenity tohoto mikroorganismu jsou velmi široké. Zahrnují rezistenci k antibiotikům, produkci velkého množství extracelulárních proteinů (enzymů i toxinů) a také tvorbu biofilmu. Právě tvorba biofilmu umožňuje přežít na plastových materiálech a způsobovat tak nozokomiální nákazy. *S. aureus* způsobuje různě závažné infekce a onemocnění, od povrchových infekcí kůže až po těžké systémové stavy. Mezi typické projevy patří enterotoxikózy (způsobené produkovaným enterotoxinem), a syndrom opařené kůže neboli Ritterův syndrom (způsobený exfoliatinem, což je epidermolytický toxin). Z dalších produkovaných toxinů lze ještě zmínit hemolyziny (alfa, beta, gama, delta), které způsobují lýzu erytrocytů [1, 5].

Souvislost s nozokomiálními nákazami má také MRSA, což je methicilin-rezistentní kmen *Staphylococcus aureus*, který v poslední době představuje velké problémy. Pro obecnou populaci není příliš nebezpečný ale pro pacienty s oslabenou imunitou či s otevřenými rány už nebezpečný je. V současné době se k léčbě využívá vankomycin, ovšem vzhledem k neustále se zvyšující minimální účinné koncentraci, vzniká obava z vybudování rezistence i pro toto antibiotikum. Proto se k léčbě antibiotiky přistupuje jen ve vážných případech jako je nekrotizující pneumonie, infekční endokarditida, apod. [5, 6].

Nejzávažnější projev patogenity této bakterie je tzv. bakteriémie *Staphylococcus aureus* (SAB). Bakteriémie obecně znamená, že v krvi jedince koluje daná bakterie ale ne její toxiny. Tato konkrétní bakteriémie je nejčastější infekcí krevního řečiště s vysokou incidencí (tj. počet nemocných za určitý časový úsek neboli pravděpodobnost onemocnění s určitou diagnózou) a letalitou (tj. smrtelnost, poměr počtu zemřelých na danou chorobu vůči celkovému počtu živých osob s danou chorobou) [4].

1.1.2 *Staphylococcus epidermidis*

Jedná se o bakterii nejčastěji se vyskytující na kůži, jak lze odvodit z názvu. Původně byla pojmenována *Staphylococcus albus* díky tomu, že při kultivaci na bohatém živném médiu tvoří malé bílé a kulovité kolonie. Při kultivaci na krevním agaru nezpůsobuje hemolýzu, což je jednou z odlišností od *S. aureus*. Při běžném výskytu na pokožce není *S. epidermidis* patogenem, ovšem v nemocničním prostředí patogenem být může. Stejně jako *S. aureus* produkuje biofilm, který ho chrání před působením antibiotik a také obrannými mechanismy hostitele. Díky tvorbě biofilmu poté ulpívá na katetrech a dalších přístrojích a náčiních. Většina onemocnění, které způsobuje, je chronického charakteru, například infekce intravaskulárního katetru nebo protetického kloubu a jejich komplikace [1, 7].

1.2 Rod *Streptococcus*

Jde o rod kokovitých grampozitivních bakterií, charakteristického řetízkovitého uspořádání. Najdeme zde zastoupení bakterií přirozeně se vyskytujících u lidí i zvířat, bakterie využívající se v potravinářském průmyslu, ale i patogenní bakterie. Patogenní bakterie tohoto rodu způsobují infekce především horních a dolních dýchacích cest. Tento rod můžeme rozdělit na tři skupiny, a to skupinu A, B a pneumokoky, podle přítomných antigenů. Do skupiny A řadíme streptokoky, jejichž rozpoznávací antigen je M-protein, který inhibuje fagocytózu a je zodpovědný za jejich virulenci. Streptokoky, které řadíme do skupiny B, mají na svém povrchu antigen kapsulární polysacharid, který je důležitý pro jejich virulenci a rozpoznání. Tyto streptokoky jsou rizikové především pro novorozence, těhotné ženy a starší jedince, konkrétně jsou původcem např. meningitid a sepsí. Pro pneumokoky je typický pouzdrový polysacharidový antigen, který je klíčový v jejich virulenci [1, 2].

1.2.1 *Streptococcus pyogenes*

Streptococcus pyogenes (*S. pyogenes*) je grampozitivní bakterie, která se kultivuje na krevním agaru, kde roste v drobných koloniích se zónou úplné β -hemolýzy. Řadí se mezi streptokoky skupiny A, je to výhradně lidský bakteriální patogen, způsobující lehké infekce kůže a sliznic, až hnisavé toxicko-invazivní onemocnění, která často bývají smrtelná. Virulentní potenciál je u tohoto druhu opravdu vysoký. *S. pyogenes* se dokáže vyrovnat se všemi úrovněmi imunitní obrany hostitele, tudíž ho považujeme za velmi nebezpečný patogen [1, 8].

Mezi významná onemocnění spojená se *S. pyogenes* můžeme zařadit postsreptokokovou glomerulonefritidu, což je autoimunitní onemocnění většinou vyvolané opakovanými infekcemi *S. pyogenes*. Dále akutní revmatickou horečku a revmatické onemocnění srdce [9].

Revmatická horečka je zánětlivé systémové onemocnění pojivových tkání, které se vyznačuje horečkou, postižením kloubů, kůže a srdce. V Evropě se skoro nevyskytuje, je spíše problémem rozvojových zemí. Ovšem revmatická horečka může přejít až v revmatické onemocnění srdce, což je velmi vážný stav. Název je odvozen na základě podobnosti příznaků s revmatem. Při těžších případech může dojít až k selhání srdce. Problém však představuje i samotné zjizvení tkáně chlopní. Diagnostika se pojí s anamnézou prodělané streptokokové infekce a detekcí protilátek ASLO, které jsou pro tuto infekci typické [10].

Na tyto onemocnění ročně umře půl milionu lidí, v celosvětovém měřítku. Účinnými látkami, kterými se zatím *S. pyogenes* léčí, jsou peniciliny a cefalosporiny, ovšem rezistence vůči všem ostatním antibiotikům neustále stoupá [9].

1.2.2 *Streptococcus pneumoniae*

Streptococcus pneumoniae (*S. pneumoniae*) patří také mezi grampozitivní bakterie, při kultivaci na krevním agaru je alfa hemolytický, netvoří spory a je nepohyblivý. *S. pneumoniae* je součástí přirozené mikroflóry horních cest dýchacích. Pokud však hostitel má slabou imunitu, kolonizace *S. pneumoniae* se šíří do dolních cest dýchacích a dalších orgánů, tkání a stává se tak patogenním. Obzvláště nebezpečný je pro malé děti, seniory a imunokompromitované osoby. *S. pneumoniae* způsobuje pneumokoková onemocnění jako například: pneumonii

(obvykle lobárního typu), paranasální sinusitidu, meningitidu (ta je většinou až sekundární infekcí tohoto patogenu), septickou artritidu nebo mozkové abscesy [9, 11].

1.3 Rod *Pseudomonas*

Rod *Pseudomonas* byl popsán v roce 1894 a je jedním z nejrozmanitějších a nejrozšířenějších bakteriálních rodů, jehož druhy byly izolovány po celém světě ve všech typech prostředí od Antarktidy po tropy. Zástupci rodu *Pseudomonas* jsou většinou pohyblivé gramnegativní tyčinky. Rozsáhlé rozšíření tohoto rodu je způsobeno nenáročností na prostředí a nutriční všestranností. Jsou schopny využívat velké množství organických sloučenin jako zdroj uhlíku a dusíku. Některé druhy jsou dokonce schopny žít v destilované vodě díky přítomnosti stopových živin. Jsou odolné vůči většině používaných antibiotik. Naštěstí jsou nákazy tímto rodem spíše oportunní neboli omezené na pacienty se špatnou imunitou či jinými vážnými stavy [1].

1.3.1 *Pseudomonas aeruginosa*

Pseudomonas aeruginosa je gramnegativní bakterie, která je známá především kvůli své přirozeně vysoké rezistenci vůči antibiotikům. Dále má schopnost udržet plazmidy s informací o rezistenci vůči dalším antibiotikům, tudíž je obávaným a nebezpečným patogenem. Také má tendenci kolonizovat povrchy ve formě biofilmu, kvůli kterému jsou její buňky neprostupné vůči terapeutickým koncentracím antibiotik. Konkrétně způsobuje různé druhy infekcí, ať už se jedná o infekce močových cest nebo respirační infekce, vzhledem k rezistenci na antibiotika jsou tyto infekce těžce léčitelné a problematické. Nákaza touto bakterií je nejčastější u pacientů s cystickou fibrózou. Kolonizace *P. aeruginosa* se zvyšuje s věkem pacienta a také tím, že kolonizace jiných mikroorganismů je často nahrazována právě tímto patogenem. Dle nových studií jsou časté kolonizace způsobené hlavně tím, že se patogen nevyskytuje převážně v plicích ale kolonizace přichází z dutiny ústní. Pacienti s cystickou fibrózou si navíc často tento patogen předávají mezi sebou. Účinných antibiotik je opravdu málo např. gentamicin. Bohužel vůči některým kmenům nejsou účinná žádná antibiotika, takže v některých případech je infekce nevléčitelná a často smrtelná [12].

2. Genom

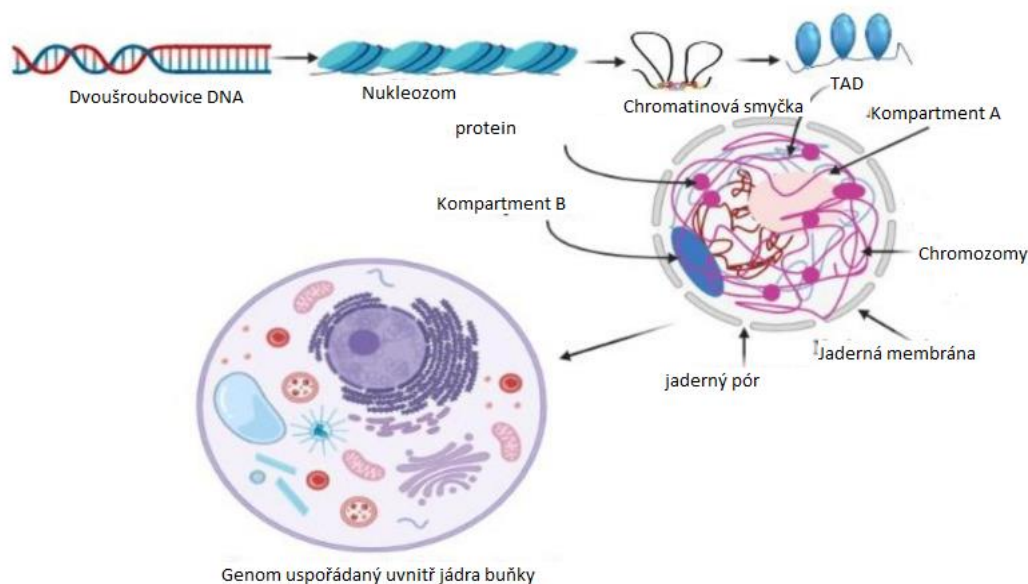
Genom je úplný soubor všech genů uložených v DNA (v případě některých virů i v RNA) obsažených v buňkách organismu. Sekvence genomu představuje kompletní pořadí všech nukleotidů v DNA. Nukleotidy se skládají z dusíkaté báze (adenin, guanin, thymin, cytosin, uracil), pětiuhlíkatého monosacharidu (ribóza nebo deoxyribóza) a jednoho nebo více zbytků kyseliny fosforečné. Struktura genomu a genetická sekvence jsou hlavními rozlišovacími znaky čeledi, druhu a kmene daného organismu. Jednotlivé kmeny mikroorganismů tak mohou být odlišeny nejen na základě mikrobiologické kultivace, ale také pomocí sekvenování [1].

2.1 Struktura genomu

Genom je nejdůležitější funkční součástí buňky. Obsahuje miliony nukleotidových bází uspořádaných do specifického rámce, které tvoří kompaktní trojrozměrné struktury tzv. chromozomy. Genom zahrnuje jak kódující, tak nekódující sekvence DNA a představuje celkovou genetickou výbavu organismu. Rekonstrukce trojrozměrné struktury chromatinu (komplexu DNA a proteinů) z jednorozměrné lineární sekvence DNA představuje vzhledem k fyzikálním omezením prostorového uspořádání v buněčném jádře poněkud nesnadný úkol. Vizualizace této trojrozměrné struktury genomu je uvedena na Obr. 1. Genom prokaryotních bakterií je zpravidla tvořen pouze jedním kruhovým chromozomem, ale faktory patogenity mohou být lokalizovány i na plasmidech, ostrovech virulence nebo být součástí mobilních genetických elementů, jako jsou profágy a transpozony. Tyto faktory zahrnují například geny pro toxiny, adheziny, invaziny, enzymy nebo sekreční systémy, které umožňují patogenu kolonizaci hostitele a vyvolání onemocnění [13].

Chromozom má v jádře své konkrétní místo, které je dále rozděleno na chromozomové kompartmenty, topologicky asociované domény (TAD) a chromatinové smyčky. Skládání chromatinu hraje zásadní roli v genové regulaci, přičemž transkripční kontrola je zprostředkována fyzickými kontakty mezi cílovými geny a příslušnými enhancery - oblastmi DNA, na které se vážou regulační proteiny. Funkční smyčka mezi genem a regulační doménou obvykle probíhá v rámci jedné TAD. Genetická informace a funkce genomu jsou ovlivněny nejen epigenetickými markery v jednorozměrných lineárních sekvencích DNA, ale také jejím

nenáhodným prostorovým uspořádáním. Organizace trojrozměrné struktury chromatinu tak přímo souvisí s funkčností genomu [13].



Obr.1: Trojrozměrná organizace genomu [13].

2.2 Funkce genomu

Hlavní funkcí genomu je udávání specifity dané buňky. Jednotlivé typy buněk jsou definovány určitými programy genové exprese a liší se také vzorci globální organizace genomu, například rozsahem a umístěním heterochromatinových domén. Jedním z předpokladů vyplývajících z funkce genomu jako hnací síly jeho prostorové organizace je, že celkový morfologický vzhled genomu by se měl lišit mezi jednotlivými buněčnými typy a také s rostoucí evoluční vzdáleností buněčných linií. Struktura genomu jasně souvisí s jeho funkcí, jak ukazují rozsáhlé korelace mezi stavem kondenzace chromatinu a transkripční aktivitou genů. Stejně tak je však zřejmé, že samotná struktura neurčuje funkční stav genu. K dalším funkcím genomu patří replikace a obnova, přičemž obě tyto funkce závisejí na struktuře chromatinu. Dále pak transkripce a translace, které jsou ovlivňovány regulačními oblastmi genomu [14].

3. Analýza struktury a funkce genomu

Od roku 1995, kdy byly zveřejněny první dvě kompletní sekvence bakteriálních genomů, se věda o bakteriích dramaticky změnila. Pomocí sekvenování DNA třetí generace je dnes možné kompletně sekvenovat bakteriální genom během několika hodin a také identifikovat některé typy metylačních míst podél genomu. Sekvenování bakteriálních genomů je nyní standardním postupem a informace získané z desítek tisíc bakteriálních genomů mají zásadní vliv na náš pohled na svět bakterií. V současnosti jsou k dispozici sekvence více než 50-ti tisíc genomů různých bakteriálních kmenů. Jejich rozložení je však značně nevyvážené ve prospěch několika kmenů, které zahrnují modelové organismy. Rozsah dostupných dat se však stále rozšiřuje, přičemž probíhající projekty se zaměřují na doplňování méně charakterizovaných taxonomických skupin. Nutno zmínit, že na tento pokrok měla a má velký vliv metagenomika zabývající se izolací DNA a následnou sekvenací genomu z jakéhokoli vzorku přímo z prostředí (např. vzorky půdy, vody nebo obsah žaludku). Právě díky této metodě bylo objeveno velké množství nových druhů mikroorganismů [15].

Důvodem, proč je sekvenování genomů tak důležité je, že umožňuje mnoho praktických aplikací, například modelování metabolismu v měřítku celého genomu, biodohled, bioforenzní analýzu, aj. [15].

3. 1 Metody sekvenování

Sekvenováním rozumíme určování pořadí nukleotidů v molekulách nukleových kyselin. Metody sekvenování můžeme rozdělit na klasické metody a metody nové generace. Mezi klasické metody řadíme metodu dle Sangera (1977) a metodu Barnes (1978). Tyto metody byly schopné zjistit jen přibližné pořadí bází u krátkých fragmentů DNA. Metody nové generace jsou založené na masivní paralelní syntéze pomocí DNA polymerázy s cyklickým začleňováním bází [16].

Hlavní rozdíl mezi tradičními metodami a sekvenováním nové generace spočívá v objemu dat, který lze analyzovat. Zatímco tradiční metody umožňují sekvenovat pouze jeden fragment DNA v čase, NGS technologie dokáže ve jednom běhu zpracovat velké množství fragmentů současně. To přináší výhody nejen z hlediska času, ale i nákladů. Tudiž pro sekvenování většího objemu dat je lepší NGS ale pokud se zaměřujeme na nízký počet cílů

bude vhodnější tradiční sekvenování. NGS má nevýhodu oproti tradičnímu sekvenování především ve složitější analýze dat a ve složitějším pracovním postupu [17].

3.2 Sekvenování první generace

James Watson a Francis Crick v roce 1953 na základě krystalografických dat Rosalindy Franklinové a Maurice Wilkinse vyřešili trojrozměrnou strukturu DNA, což přispělo k vytvoření koncepčního rámce pro replikaci DNA i kódování proteinů v nukleových kyselinách. Schopnost "číst" neboli sekvenovat DNA však nenásledovala ihned po objasnění její struktury. Ukázalo se, že strategie vyvinuté k určování sekvence proteinových řetězců se nedají snadno použít pro výzkum nukleových kyselin. Molekuly DNA byly mnohem delší a skládaly se z menšího počtu jednotek, které si byly navzájem podobné, takže bylo obtížnější je rozlišit. Počáteční snahy se soustředily na sekvenování nejsnáze dostupných populací relativně čistých druhů RNA, jako jsou mikrobiální ribozomální nebo transferová RNA nebo genomy bakteriofágů s jednořetězcovou RNA. Výhodou bylo, že nejsou komplikovány komplementárním vláknem a jsou často podstatně kratší než molekuly eukaryotické DNA. Navíc již byly známy a k dispozici enzymy RNázy schopné štěpit řetězce RNA na specifických místech. Navzdory těmto výhodám zůstával pokrok pomalý, dosavadní technika dokázala analyzovat pouze složení nukleotidů, nikoliv jejich pořadí. V roce 1965 se Robertu Holleymu a jeho kolegům podařilo vytvořit první sekvenci celé nukleové kyseliny. Souběžně s tím Fred Sanger a jeho kolegové vyvinuli příbuznou techniku založenou na detekci částečně štěpených radioaktivně značených fragmentů po dvourozměrné frakcionaci, která výzkumníkům umožnila postupně doplňovat rostoucí knihovnu sekvencí ribozomální a transferové RNA. Také díky této metodě dvourozměrné frakcionace se v laboratoři Waltera Firse podařilo v roce 1972 získat první kompletní sekvenci genu kódujícího bílkoviny, o čtyři roky později pak kompletní genom [18].

Metody sekvenování první generace spadají do kategorie tradičních sekvenačních technik. V následujících kapitolách budou popsány dvě nejvýznamnější z nich, a to sekvenování podle Sangera a podle Maxama-Gilberta.

3.2.1 Sekvenování dle Sangera

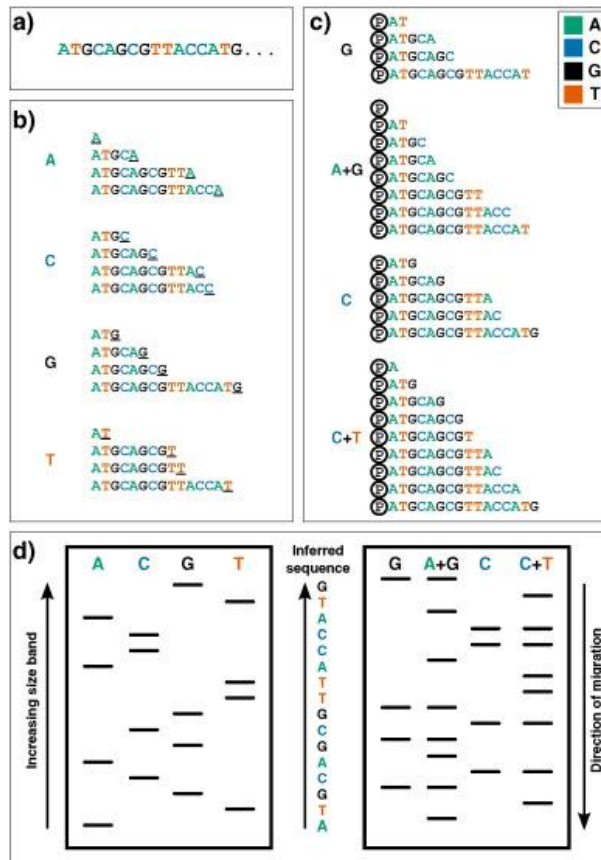
Další praktickou změnou, která měla velký dopad, bylo nahrazení 2-D frakcionace (která často zahrnovala elektroforézu i chromatografii) jedinou separací podle délky polynukleotidu pomocí elektroforézy v polyakrylamidovém gelu, která poskytla mnohem větší rozlišovací schopnost. Tuto techniku využili Alan Coulson a Frederic Sanger v roce 1975. Technika využívala DNA polymerázu k syntéze DNA z primeru, přičemž byly začleňovány radioaktivně značené nukleotidy. Následně byly provedeny dvě polymerační reakce, které umožnily detekci vzniklých fragmentů: "plusová" reakce, ve které je přítomen pouze jeden typ nukleotidu, takže všechna rozšíření končí touto bází, a "minusová" reakce, ve které se používají tři typy nukleotidů, které vytváří sekvence končící těsně před chybějícím čtvrtým nukleotidem. Právě touto technikou Sanger a jeho kolegové sekvenovali první DNA genom, který má dodnes v mnoha sekvenačních laboratořích své místo jako pozitivní kontrolní genom [18].

Zásadní průlom, který navždy změnil vývoj technologie sekvenování DNA, však přišel v roce 1977, kdy Sanger vyvinul techniku tzv. řetězového zakončení. Tato technika využívá chemické analogy deoxynukleotidtrifosfátů (dNTP), konkrétně dideoxynukleotidtrifosfáty (ddNTP), které slouží jako terminátory syntézy DNA řetězce. Dideoxynukleotidtrifosfáty (ddNTP) postrádají 3' hydroxylovou skupinu, která je nutná pro prodlužování řetězců DNA pomocí vazby s 5' fosfátem dalšího dNTP. Z tohoto důvodu jejich začlenění do syntetizovaného řetězce DNA vede k jeho ukončení. Přimícháním malého množství radioaktivně značených ddNTP do reakční směsi obsahující běžné dNTP dochází při polymeraci k náhodnému zařazení ddNTP, což vede ke vzniku řetězců DNA různé délky – vždy zakončených specifickým ddNTP. Provedením čtyř paralelních reakcí – každé s jiným typem ddNTP (ddATP, ddTTP, ddCTP, ddGTP) – a následnou elektroforézou v polyakrylamidovém gelu, kde každý pruh odpovídá jedné bázi, lze pomocí autoradiografie odvodit sekvenci nukleotidů. Na příslušné pozici gelu se zobrazí radioaktivní pás, který označuje zakončení řetězce konkrétní bází. Sekvenování dle Sangera je zobrazeno na Obrázku 2-b,d [18].

Přestože tato metoda, známá jako Sangerovo sekvenování, pracuje na stejném principu jako ostatní techniky (generování řetězců různých délek a značení konečného nukleotidu), vyniká svou přesností a relativní jednoduchostí. Díky těmto vlastnostem se stala na dlouhá léta nejběžnější technologií používanou k sekvenování DNA [18].

3.2.2 Sekvenování dle Maxama-Gilberta

Jedná se o metodu založenou na chemické fragmentaci DNA a jejím následném zobrazení pomocí elektroforézy. Tato metoda byla představena koncem 70. let 20. století. Prvním krokem je převedení vzorku DNA na jedno vlákno. Pomocí alkalické fosfatázy se odstraní fosfátová skupina v sekvenci DNA a na 5' konec jednořetězcové DNA se přidá radioaktivní fosfátová skupina (P^{32}). Následuje amplifikace pomocí polymerázové řetězové reakce (PCR), výsledkem jsou miliony kopií templátové DNA. Pomocí chemických látek, specifických pro každou bázi, se tyto řetězce rozpadnou na čtyři skupiny, každá skupina je specifická pro jednu bázi. Výsledkem tedy je, že reakce probíhá ve čtyřech zkumavkách, každá ze zkumavek obsahuje látku specifickou pro jednu bázi a radioaktivně značenou templátovou DNA. Následuje vizualizace pomocí elektroforézy [19], viz. Obr. 2-c,d.



Obr. 2: Technologie sekvenování DNA první generace: a) příklad DNA, která má být sekvenována, b) sekvenování dle Sangera, c) sekvenování dle Maxama-Gilberta, d) vizualizace obou metod pomocí elektroforézy [18].

3.3 Sekvenování druhé generace

Sekvenování nové generace (NGS) dalo možnost používat sekvenování pro rutinní vyšetřování v klinických laboratořích. Například lidský genom zvládne sekvenovat už za jeden den oproti projektu lidského genomu, který trval dvanáct let. V klinice je velmi důležité ve screeningu genetických onemocnění, rizika i existence rakoviny, atd. Hlavní výhodou je možnost porovnání pacientova genomu s referenčním genomem [20].

Technologie NGS prošly vývojem, v jehož rámci se rozlišují různé generace (druhá, třetí, čtvrtá) podle technických principů a úrovně výkonnosti. Druhá generace sekvenování, označovaná jako klasické NGS, umožňuje hromadné paralelní sekvenování milionů fragmentů DNA současně. Typickými zástupci jsou technologie pyrosekvenování, Illumina, Ion Torrent, popsané níže. Proces zahrnuje fragmentaci DNA, její amplifikaci a následné sekvenování. Hlavními výhodami jsou vysoká propustnost, rychlost a nízká cena za bázi. Nevýhodou je nutnost amplifikace a relativně krátká délka čtení (zpravidla 100–300 bp), což může komplikovat analýzu složitých genomů [20].

Sekvenování druhé generace, se často využívá v metagenomice (studium genetického materiálu získaného přímo ze životního prostředí). Pyrosekvenování se využívá v projektu lidského mikrobiomu, který má za cíl studovat provázanost mezi složením lidského mikrobiomu a vznikem různých onemocnění [21].

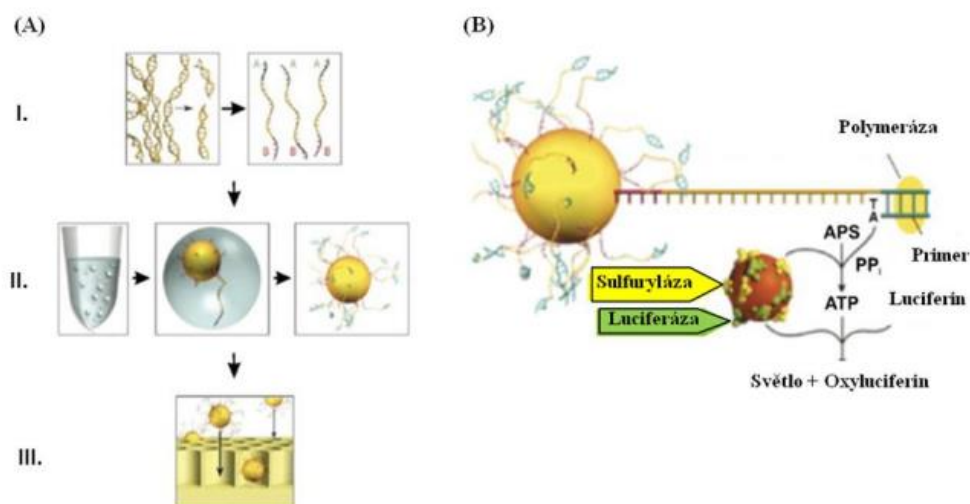
3.3.1 Pyrosekvenování

Současně s rozvojem rozsáhlého Sangerova sekvenování se objevila další technika, která připravila půdu pro první vlnu sekvenátorů DNA nové generace. Tato metoda se výrazně lišila od stávajících metod tím, že neodvozovala identitu nukleotidů pomocí radioaktivně nebo fluorescenčně značených dNTP nebo oligonukleotidů před vizualizací pomocí elektroforézy. Místo toho vědci využili nedávno objevenou luminiscenční metodu pro měření syntézy pyrofosfátu. Ta spočívala ve dvouenzymovém procesu, při němž se ATP-sulfuryláza používá k přeměně pyrofosfátu na adenosintrifosfát (ATP), což je hlavní zdroj energie v buňce, který se pak používá jako substrát pro luciferázu, čímž vzniká světlo úměrné množství pyrofosfátu. Tento přístup byl použit k odvození sekvence měřením produkce pyrofosfátu, odpovídající intenzitě luminescence, při postupném promývání každého nukleotidu systémem přes

templátovou DNA imobilizovanou na pevné fázi. Postup pyrosekvenování je zobrazen na Obrázku 3 [18].

Tato technika pyrosekvenování měla řadu vlastností, které byly považovány za přínosné. Bylo možné ji provádět s použitím přirozených nukleotidů (namísto silně modifikovaných dNTP používaných v protokolech ukončujících řetězce) a umožňovala sledovat reakci v reálném čase, čímž odpadla potřeba zdlouhavé elektroforézy. Hlavním problémem této techniky však bylo určení počtu po sobě jdoucích stejných nukleotidů (tzv. homopolymerů). Intenzita uvolněného světla sice odpovídala délce homopolymeru, ale šum signálu způsobil nelineární odečet u sekvencí s více než čtyřmi nebo pěti stejnými nukleotidy [18].

Pyrosekvenování bylo později licencováno biotechnologickou společností 454 Life Sciences, založenou Jonathanem Rothbergem, kde se vyvinulo v první významnou a úspěšnou komerční technologii sekvenování nové generace (NGS) [18].



Obr. 3: Postup pyrosekvenování [29]: A) zobrazuje I. fragmentaci DNA a přípravu DNA knihovny, II. kuličky s navázaným fragmentem se vloží do zkumavky, kde dochází k amplifikaci pomocí PCR a v bodu III. se kuličky vloží do pikotitrační destičky. Část B) zobrazuje samotnou sekvenační reakci [22].

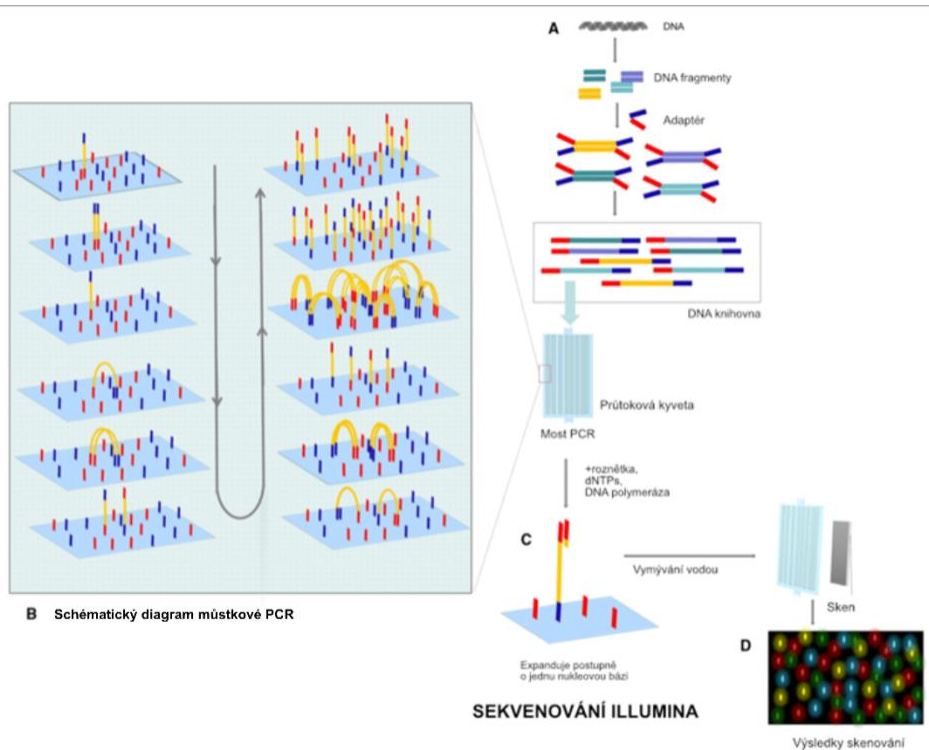
3.3.2 Illumina

Metoda se začala používat v roce 2006 a jejím principem je reverzibilní terminace řetězce. Illumina je velmi rozšířenou platformou v genomických studiích a zahrnuje šest různých typů sekvenátorů [16].

Sekvenátory využívají technologii sekvenování syntézou. DNA knihovna s fixními adaptéry je nejprve denaturována na jednotlivá vlákna a navázána na povrch průtokové buňky, kde následně dochází k můstkové amplifikaci. Tento proces vede ke vzniku klastrů, které obsahují klonální fragmenty DNA. Před samotným sekvenováním se knihovna pomocí linearizačního enzymu rozštěpí, čímž vznikají jednořetězcové templáty. Během sekvenační reakce jsou k templátu postupně připojovány čtyři druhy nukleotidů, z nichž každý je opatřen fluorescenčním barvivem a dočasně blokující skupinou, která brání přidání další báze. Po každém cyklu je zaznamenán fluorescenční signál pomocí CCD (charge-coupled device) detektoru, který je citlivý na elektrický náboj a blokující skupina je následně odstraněna, což umožňuje pokračování syntézy [23].

Ze sekvenátoru Illuminy je možné získat od 1,5 – 600 Gb dat na jeden běh dle výběru konkrétního sekvenátoru (záleží na typu). Chybovost je do jednoho procenta a kvalita fragmentů klesá s nárůstem počtu bází. Illumina má přesné vyhodnocení v homopolymerních úsecích, ale v úsecích obsahujících C+G její spolehlivost klesá [16].

Hlavní výhodou této metody je její využitelnost v rámci studia různých onemocnění. Zaměřuje se na sekvenování celého exomu, který tvoří méně než 2 % lidského genomu, ale zahrnuje přibližně 85 % známých variant souvisejících s chorobami [16]. Proces sekvenování pomocí metody Illumina je zobrazený na Obrázku 4.



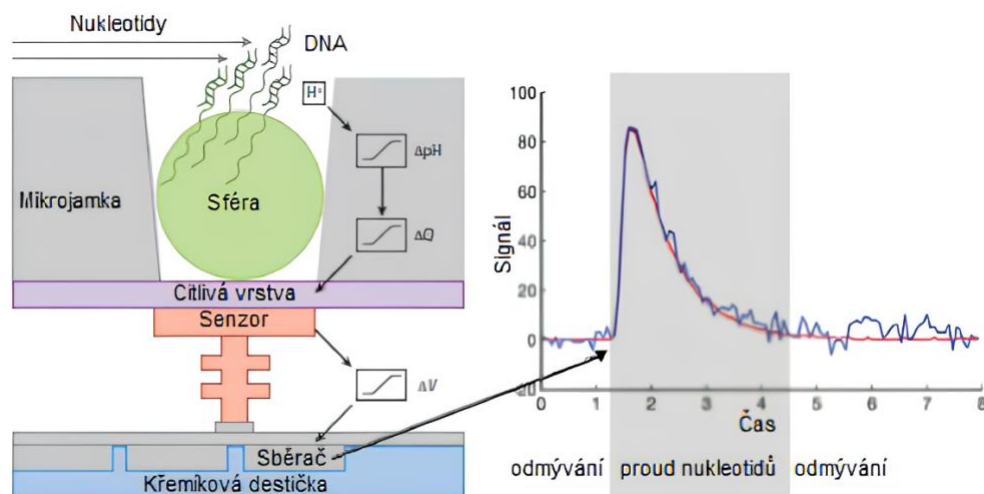
Obr. 4: Proces sekvenování pomocí metody Illumina: A) fragmentace DNA a vytvoření knihovny DNA, B) můstková amplifikace, ve které vznikají shluky tvořené mnoha klonálními fragmenty DNA, C) do průtokové kyvety jsou přidány primery, specificky fluorescenčně označené dNTPs a DNA polymeráza, D) vizualizace fluorescenčních výsledků sekvenování.

Převzato a upraveno podle [24].

3.3.3 Ion Torrent

Ion PGM (Personal Genome Machine) byl uveden na trh společností Ion Torrent koncem roku 2010. Představuje jednu z posledních technik druhé generace sekvenování. PGM využívá polovodičovou sekvenační technologii, která detekuje změny pH související se změnami napětí vznikajícími při inkorporaci nukleotidů. Pokud je přidán správný nukleotid, dojde k uvolnění vodíkového iontu, což způsobí měřitelnou změnu pH, a tudíž i změnu napětí, která je detekována senzorem. Při přidání nesprávného nukleotidu k žádné reakci nedochází a změna pH se nedetekuje. Pokud jsou přidány 2 identické nukleotidy za sebou, je detekováno dvojnásobné napětí [18].

Při přípravě knihovny fragmentů se využívá párové uspořádání. Nejprve se štěpí DNA a ligují se adaptéry na oba konce fragmentů. Po imobilizaci na nosičové sféry následuje amplifikace pomocí emulzní PCR. Posledním krokem přípravy knihovny je nanesení sfér centrifugací na destičku s mikrojamkami. Následuje samotné sekvenování, které probíhá na mikročipu, který obsahuje vrstvu mikrojamek, pod nimiž je vodivá vrstva s integrovaným detektorem, přičemž celý systém je umístěn na křemíkové destičce. Schéma mikročipu je uvedeno na Obr. 5. Mikročip je překrytý mikrofluidní vrstvou, ve které postupně proudí jednotlivé báze a polymerázy. Sekvenování probíhá tak, že báze vstupují do systému v přesně definovaném pořadí. První čtyři odpovídají čtyřem nukleotidům na adaptéru pod primerem. Při každém zařazení správné báze během syntézy DNA se uvolní vodíkový kation, což se projeví jako změna pH v příslušné mikrojamce. Tato změna je zaznamenána jako elektrický signál díky vodivé vrstvě, která je kovovým sloupkem spojená s detektorem. Zaznamenaný signál umožňuje čipu určit, který nukleotid byl právě připojen, a tak postupně rekonstruovat sekvenci templátové DNA [16].



Obr. 5: Schéma mikročipu a zpracování signálu technikou Ion Torrent [16].

Příkladem použití sekvenátoru Ion Torrent PGM je identifikace mikrobiálních patogenů. V Německu proběhla v roce 2011 velká nákaza patogenní *Escherichia coli*, díky použití této technologie byly poměrně rychle nalezeny fungující antibiotika. Z čehož vyplývá, že Ion Torrent má potenciál rychlého, ale výkonově omezeného sekvenátoru v případě výskytu nového onemocnění [18].

PGM je první komerční sekvenátor, který nevyužívá fluorescenční značení a skenování kamerou, což vede k vyšší rychlosti a menší velikosti přístroje. Velkou výhodou je i jednodušší pracovní postup a dostupnější cena. Nevýhoda spočívá v obtížném čtení homopolymerních extenzí a opakování. Metoda je ceněná hlavně pro svoji rychlost a je určena spíše pro sekvenování menších genomů [18, 29].

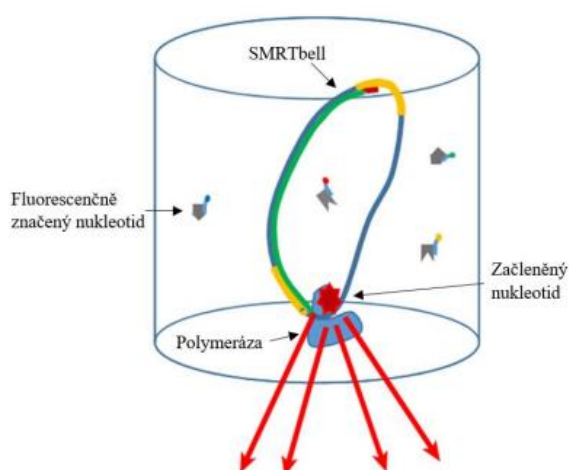
3.4 Sekvenování třetí generace

O tom, co definuje různé generace technologie sekvenování DNA, se vedou rozsáhlé diskuse. Objevují se argumenty, že definujícími znaky třetí generace by mělo být sekvenování jednotlivých molekul (z angl. single molecule sequencing, SMS), sekvenování v reálném čase a prostý odklon od předchozích technologií. Zde jsou považovány za technologie třetí generace ty, které jsou schopny sekvenovat jednotlivé molekuly, čímž se eliminuje požadavek na amplifikaci DNA, který byl společný všem předchozím technologiím. Hlavní výhodou sekvenování třetí generace jsou extrémně dlouhá čtení (desítky kilobází), což usnadňuje analýzu komplexních genomů a strukturních variant. Technologie navíc umožňuje detekci epigenetických modifikací. Nevýhodou bývá vyšší chybovost a náročnější bioinformatické zpracování [24]. První technologie SMS byla vyvinuta v laboratoři Stephena Quakea, později ji komercializovala společnost Helicos BioSciences a fungovala v podstatě stejným způsobem jako technologie Illumina, ale bez amplifikačního můstku. Templáty DNA se připojí k planárnímu povrchu a poté se vlastní fluorescenční reverzibilní terminátory dNTP (tzv. "virtuální terminátory") promyjí vždy po jedné bázi a zobrazí se, načež se odštěpí a cyklus pokračuje na další bázi. Ačkoli je tato technologie relativně pomalá a drahá, byla to první technologie, která umožnila sekvenování neamplifikované DNA, čímž se vyhnula všem souvisejícím zkreslením a chybám [18].

3.4.1 SMRT

Jedna z nejčastěji využívaných technologií třetí generace je platforma SMRT (z angl. single molecule real time) od společnosti Pacific Biosciences, která je k dispozici na přístrojích řady PacBio. Během analýzy SMRT probíhá polymerizace DNA v soustavách mikrofabrikovaných nanostruktur zvaných zero-mode waveguides (ZMW), což jsou v podstatě malé otvory v kovové fólii pokrývající čip. Tyto ZMW využívají vlastnosti světla procházejícího otvory o průměru menším, než je jeho vlnová délka, což způsobuje jeho exponenciální rozklad a osvětlení výhradně samotného dna jamek. To umožňuje vizualizaci jednotlivých fluoroforních molekul v blízkosti dna ZMW, protože zóna laserové excitace je tak malá, a to i na pozadí sousedních molekul v roztoku. Během replikačního procesu dochází při začlenění fluorescenčně značených nukleotidů k uvolnění fluorescenčního signálu (Obr. 6) a tyto záblesky jsou zaznamenávány kamerovým systémem v reálném čase. Tímto postupem lze sekvenovat jednotlivé molekuly ve velmi krátkém čase [18, 25].

Přínosem této metody, je i možnost sledování jiných parametrů studované molekuly DNA. Například detekci methylací, methylenové báze může snadno detekovat, protože přímo ovlivňují kinetiku polymerázy při sekvenování. Methylace je jedna z nejčastějších modifikací DNA bakterií, kterým slouží pro odlišení vlastní DNA od cizí. Před zavedením technologie SMRT, byly všechny metody využívané pro studium methylace velmi časově i finančně náročné [21].



Obr. 6: Replikační proces sekvenování technikou SMRT [25].

3.5 Sekvenování čtvrté generace

Čtvrtá generace sekvenování zatím není formálně standardizovaná, ale zahrnuje nové přístupy, které kombinují nebo rozšiřují předchozí technologie. Tyto přístupy uchovávají prostorové souřadnice sekvencí RNA a DNA až na úrovni subcelulárního rozlišení, a umožňují tak zpětné mapování sekvenačních čtení do původního histologického kontextu.

Takto získané informace jsou aktuálně využívány například ve výzkumu mozku, kde pomáhají lépe chápat funkční organizaci tkání. Také ve výzkumu rakoviny má sekvenování čtvrté generace potenciál zásadně změnit možnosti mapování molekulárního mikroprostředí nádorů. Patří sem například sekvenování *in situ*, tedy sekvenování nukleových kyselin přímo ve fixovaných buňkách, a také integrace s umělou inteligencí. Cílem je dosáhnout ještě rychlejší, levnější a přístupnější sekvenování v reálném čase, s důrazem na personalizovanou medicínu, klinické využití a možnost mobilního nasazení [26].

3.6 Shluková analýza

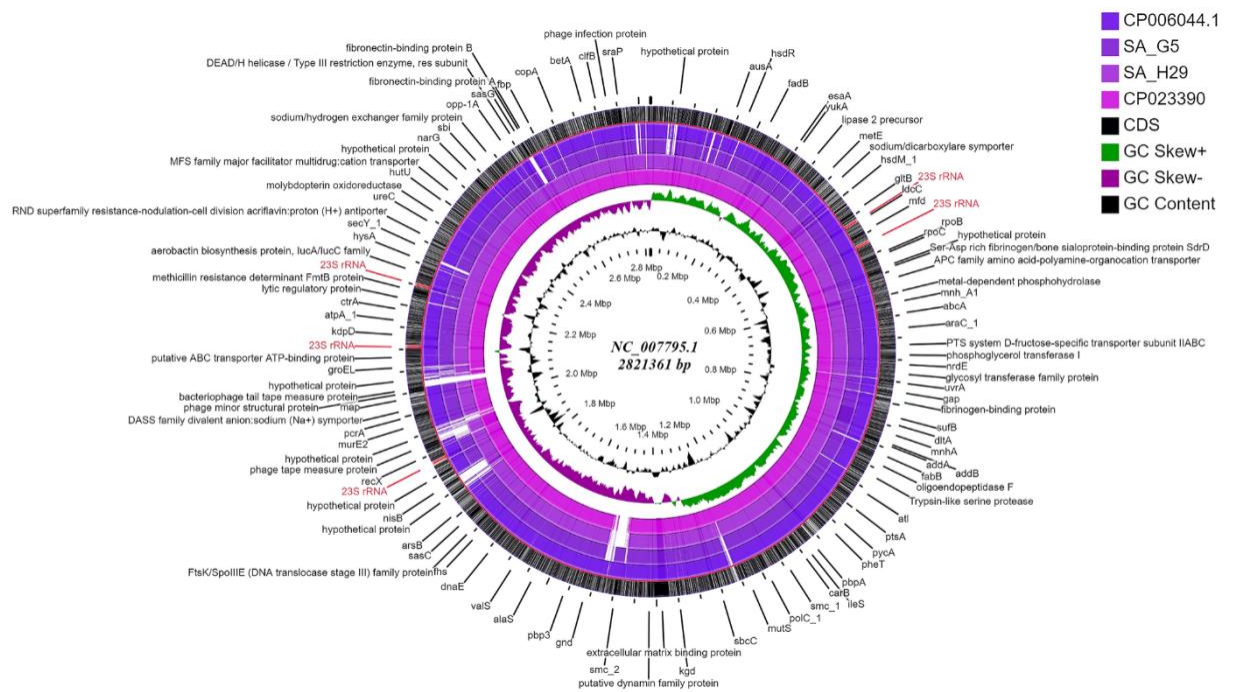
Shluková analýza (clustering) je statistická metoda používaná k seskupování objektů na základě jejich podobnosti. Použití je vhodné zejména tam, kde objekty mají již své třídy (kmeny, druhy), takže tyto objekty už samy tvoří shluky [27].

Obecně uznávané jsou čtyři základní využití techniky shlukové analýzy a to: vytváření klasifikací homogenních skupin, objevování nových vztahů anebo zkoumání koncepčních schémat, testování hypotéz a potvrzující analýza dříve identifikovaných klasifikací [18].

V genomice patogenních bakterií se shluková analýza uplatňuje při analýze sekvenačních dat, kdy umožňuje identifikovat skupiny genů, které vykazují podobné vlastnosti – např. podobnou expresi, přítomnost v určitém kmeni, nebo strukturální motivy. V souvislosti se strukturou genomu pomáhá shluková analýza rozlišit konzervované genové bloky, mobilní genetické elementy nebo oblasti s vysokou variabilitou (např. ostrovy patogenity). Z hlediska funkce je klíčová pro odhalení funkčně souvisejících genů – např. genů zapojených do virulence, rezistence vůči antibiotikům nebo metabolických drah [29].

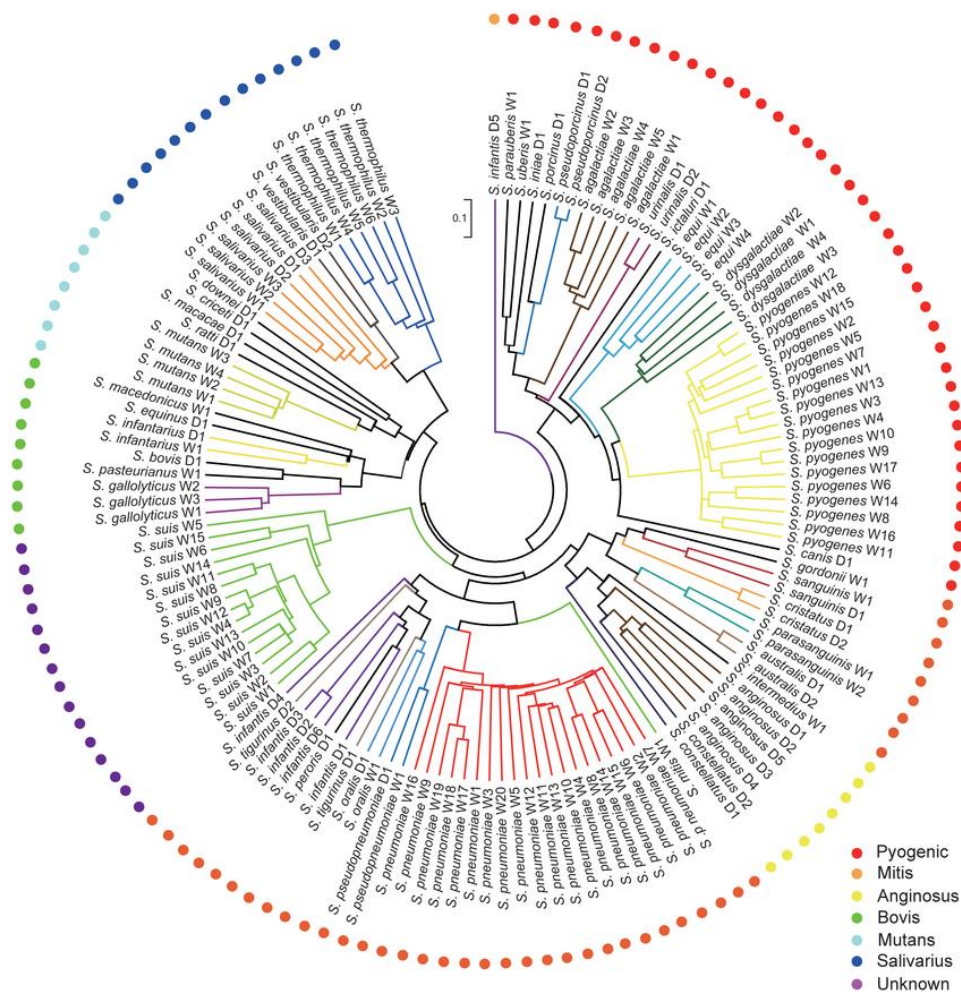
Díky integraci s dalšími bioinformatickými metodami lze pomocí shlukové analýzy odhalit regulační sítě a funkční moduly, které jsou klíčové pro adaptaci a patogenitu bakterií. Tato metoda tak výrazně přispívá k porozumění evolučních strategií bakteriálních patogenů a může podpořit vývoj nových diagnostických či terapeutických nástrojů.

Výsledkem shlukové analýzy jsou často také tzv. dendrogramy, což jsou „genealogické stromy“ studovaných sekvencí. Jedná se vlastně o grafy, které znázorňují vztahy mezi operačními taxonomickými jednotkami a jejich hypotetickými předky. Z tohoto grafu můžeme vyčíst také evoluční vzdálenost, což představuje odhad počtu mutací, které oddělují jednotlivé generace. Kruhová forma dendrogramu není vhodná k dalším výpočtům, pokud tedy chceme s daty nadále pracovat analyticky, je lepší pracovat se sadou standardně zapsaných znaků. Na následujících obrázcích (Obr. 7-9) jsou zobrazené dendrogramy vybraných druhů mikroorganismů (druhu *Staphylococcus aureus*, rodů streptokoků a rodu *Pseudomonas*) [29].



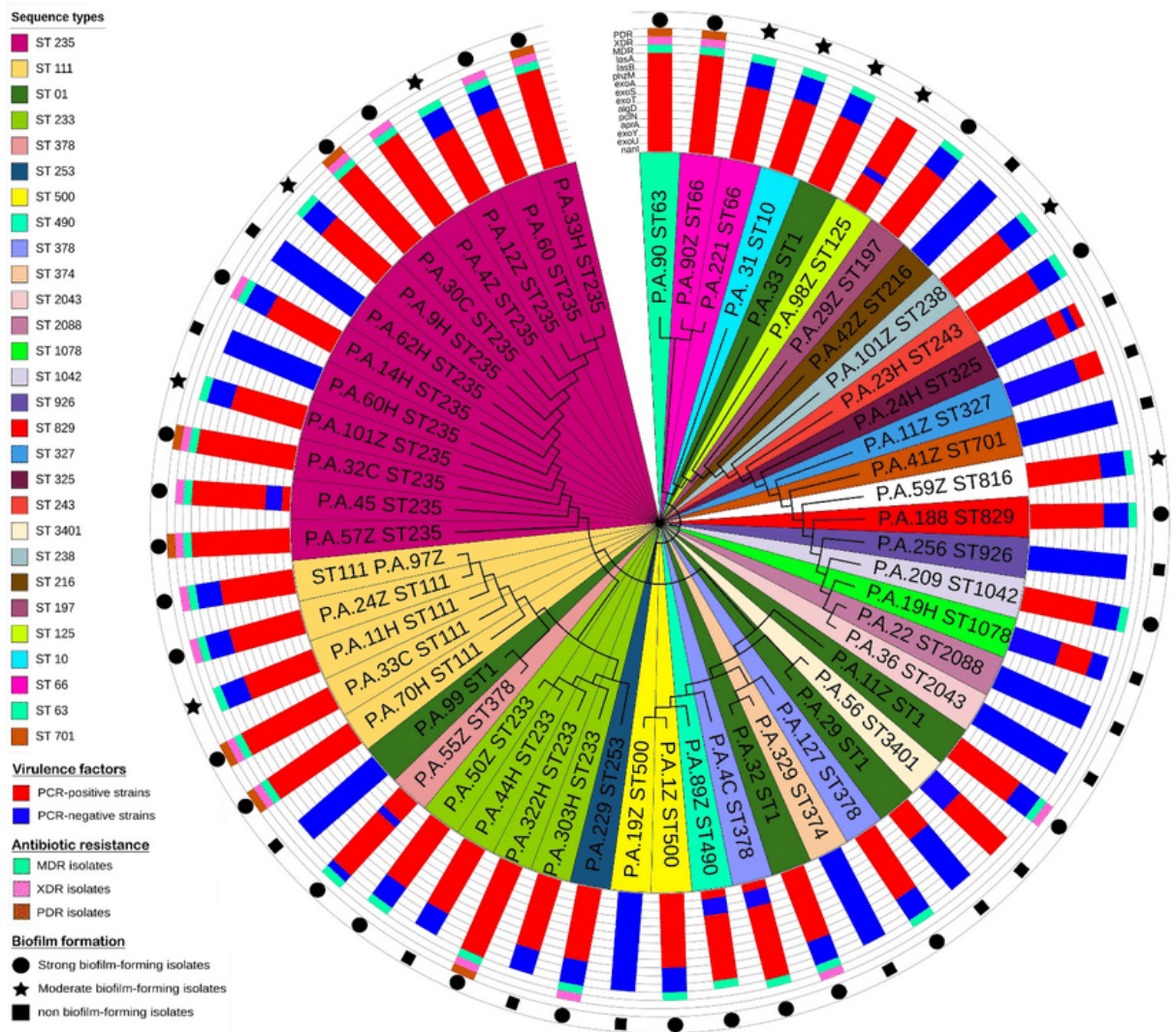
Obr. 7: Dendrogram homologních chromozomů genomů *S. aureus* proti genomu *S. aureus* subsp. *aureus* NCTC 8325 (NC_007795.1) [30].

Genom kmene *S. aureus* H29 má nejvyšší počet unikátních (singletových) genů, které kódují 37 proteinů spojených s virulencí, rezistencí, a mobilními genetickými elementy. Dendrogram na Obr. 7 znázorňuje srovnání dvou genomů (SA G5 a SA H29) a jejich příbuzných kmenů proti referenčnímu genomu *Staphylococcus aureus* subsp. *aureus* NCTC 8325 (NC_007795.1), kde každý genom je v kruhovém diagramu reprezentován jinou barvou. Tmavší oblasti v kruhovém genomu označují 100% sekvenční podobnost s referenčním genomem, světlejší oblasti pak přibližně 70% podobnost. Z těchto výsledků vyplývá, že kmeny SA G5 a SA H29 jsou si geneticky velmi podobné [30].



Obr. 8: Dendrogram izolátu z populace streptokoků [31].

Tento dendrogram zobrazuje izolát rodu *Streptococcus*. Dendrogram byl zkonstruován hierarchickým shlukováním na základě rozdílů v obsahu genů mezi 138 kmeny *Streptococcus* s použitím párových Jaccardových vzdáleností, které se pohybují od 0 do 1. Různé barevně odlišené větve označují různé druhy [31].



Obr. 9: Dendrogram izolátu z populace *Pseudomonas aeruginosa* [32].

Tento dendrogram zobrazuje kruhové shlukování 55 izolátů *P. aeruginosa* na základě profilů typu sekvence. Sekvence ve stejném klonálním komplexu jsou stínovány stejnou barvou [32].

Závěr

V této práci jsem se zaměřila na analýzu struktury a funkce genomu patogenních bakterií, konkrétně rodů *Staphylococcus*, *Streptococcus* a *Pseudomonas*. Uvedla jsem základní vlastnosti těchto mikroorganismů a popsala metody sekvenování, které se používají pro studium jejich genomu.

Momentálně nejvyužívanější jsou v klinické praxi i ve výzkumu metody sekvenování druhé a třetí generace. Nicméně prostor začíná postupně dostávat i sekvenování čtvrté generace, které ještě podléhá technologickému vývoji a ladění. Do budoucna má však velký smysl hlavně pro studium mikroprostředí karcinomů ale i pro sekvenování genomu bakterií, jelikož umožňuje sekvenaci DNA přímo ve fixovaných buňkách.

Sekvenování DNA umožňuje detailní analýzu genetické informace a je klíčovým nástrojem pro pochopení patogenity, mechanismů rezistence vůči antibiotikům a pro vývoj nových terapeutických strategií. Práce shrnuje současné poznatky o struktuře a funkci genomu těchto bakterií a zdůrazňuje význam sekvenování pro výzkum a medicínu. Detailní analýza genomu patogenních bakterií přispívá k lepšímu pochopení jejich biologie a mechanismů, které využívají k infikování hostitele a vyhýbání se imunitnímu systému. Tyto poznatky jsou zásadní pro vývoj nových léčebných postupů a prevenci infekčních onemocnění.

Seznam informačních zdrojů

1. MURRAY, Patric R., Ken S. ROSENTHAL a Michael A. PFALLER. *Medical Microbiology* [online]. 1. Canada: Elsevier, 2016 [cit. 2021-12-20]. ISBN 978-0-323-29956-5. Dostupné z: <https://iums.ac.ir/files/microb/files/Murray.pdf>
2. BEDNÁŘ, Marek et al. *Lékařská mikrobiologie: Bakteriologie, virologie, parazitologie*. 1. Praha: Marvil, 1996 [cit. 2021-12-20]. ISBN 8023802976.
3. LEITÃO, Jorge H., 2020. Microbial Virulence Factors. *International journey of microbial sciences*. [online] No. 21(15). [cit. 2025-06-09]. Dostupné z: <https://www.mdpi.com/1422-0067/21/15/5320>.
4. TONG, Steven Y. C., Joshua S. DAVIS, Emily EICHENBERGER, Thomas L. HOLLAND a Vance G. FOWLER. *Staphylococcus aureus* Infections: Epidemiology, Pathophysiology, Clinical Manifestations, and Management. *Clinical Microbiology Reviews* [online]. 2015, **28**(3), 603-661 [cit. 2022-03-17]. ISSN 0893-8512. Dostupné z: [doi:10.1128/CMR.00134-14](https://doi.org/10.1128/CMR.00134-14)
5. BŘEČKOVÁ, Adéla, 2022. *Expres toxinu u Staphylococcus aureus během infekce terapeutickými bakteriofágy*. [online], diplomová práce. Brno: Masarykova univerzita Přírodovědecká fakulta Ústav experimentální biologie [cit. 2025-05-27]. Dostupné z: https://is.muni.cz/th/lco8h/Expres_toxinu_u_Staphylococcus_aureus_behem_infekce_terapeutickymi_bakteriofagy.pdf.
6. LIU, George Y., 2009. Molecular Pathogenesis of *Staphylococcus aureus* Infection. [online] *Pediatric Research*. Č. 65, s. 71-77 [cit. 2025-05-28]. Dostupné z: <https://doi.org/https://doi.org/10.1203/PDR.0b013e31819dc44d>.
7. OTTO, Michael, 2011. Molecular basis of *Staphylococcus epidermidis* infections. [online] *Seminars in Immunopathology*. Vol. 34, s. 201-214 [cit. 2025-05-28]. Dostupné z: <https://link.springer.com/article/10.1007/s00281-011-0296-2>.
8. FIEDLER, Tomas, Thomas KÖLLER a Bernd KREIKEMEYER. *Streptococcus pyogenes* biofilms-formation, biology, and clinical relevance [online]. 2015 [cit. 2022-03-24]. Dostupné z: [doi:10.3389/fcimb.2015.00015](https://doi.org/10.3389/fcimb.2015.00015)

9. WALKER, Mark J., Timothy C. BARNETT, Jason D. MCARTHUR, et al. Disease Manifestations and Pathogenic Mechanisms of Group A *Streptococcus*. *Clinical Microbiology Reviews* [online]. 2014, **27**(2), 264-301 [cit. 2022-03-24]. ISSN 0893-8512. Dostupné z: doi:10.1128/CMR.00101-13
10. MRÁZ, Tomáš, 2025. *Revmatická horečka*. [online], [cit. 2025-05-30]. Dostupné z: <https://www.homolka.cz/nase-oddeleni/11635-kardiovaskularni-program/11635-kardiologie-kar/informacni-portal-pro-pacienty/17-revmaticka-horecka>.
11. BROOKS, Lavidia R. K. a George I. MIAS. *Streptococcus pneumoniae*'s Virulence and Host Immunity: Aging, Diagnostics, and Prevention. *Frontiers in Immunology* [online]. 2018, **9** [cit. 2022-03-31]. ISSN 1664-3224. Dostupné z: doi:10.3389/fimmu.2018.01366
12. PARKINS, Michael D., Ranjani SOMAYAJI a Valerie J. WATERS. Epidemiology, Biology, and Impact of Clonal *Pseudomonas aeruginosa* Infections in Cystic Fibrosis. *Clinical Microbiology Reviews* [online]. 2018, **31**(4), e00019-18 [cit. 2022-03-31]. ISSN 0893-8512. Dostupné z: doi:10.1128/CMR.00019-18
13. MOHANTA, Tapan Kumar, Awdhesh Kumar MISHRA a Ahmed AL-HARRASI. *The 3D Genome: From Structure to Function* [online]. 2021 [cit. 2022-06-13]. Dostupné z: doi:10.3390/ijms222111585
14. MISTELI, Tom. *The self-organizing genome: Principles of genome architecture and function* [online]. 2020, 28-45 [cit. 2022-06-13]. Dostupné z: doi:10.1016/j.cell.2020.09.014
15. LAND, Miriam; HAUSER, Loren; JUN, Se-Ran a NOKAEW, Intawat, 2015. Insights from 20 years of bacterial genome sequencing. [online] *Funct Integr Genomics*. 27.2.2015, vol. 15, s. 141-161. [cit. 2022-06-12]. Dostupné z: <https://doi.org/10.1007/s10142-015-0433-4>.
16. KOVAŘOVIC, Vojtěch, 2015. *Aplikace sekvenování nové generace pro fylogenetickou analýzu bakteriálních genomů a metagenomiku*. [online], Bakalářská práce. Brno: Masarykova univerzita. [cit. 2022-04-01]. Dostupné z: file:///C:/Users/HP%20-%20PC/Downloads/BP_-_Vojtech_Kovarovic_Archive.pdf.
17. SLABÝ, Ondřej, 2018. *Technologie sekvenování nové generace: celogenomové, celoexomové a cílené – hotspot – sekvenování*. [online], [cit. 2025-06-03]. Dostupné z: <https://www.homolka.cz/nase-oddeleni/11635-kardiovaskularni-program/11635-kardiologie-kar/informacni-portal-pro-pacienty/17-revmaticka-horecka>.

- z: <https://www.prolekare.cz/tema/precizni-medicina/detail/technologie-sekvenovani-nove-generace-celogenomove-celoexomove-a-cilene-hotspot-sekvenovani-105628>.
18. HEATHER, James M. a Benjamin CHAIN. The sequence of sequencers: The history of sequencing DNA. *Genomics* [online]. 2016, 1-8 [cit. 2022-06-06]. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S0888754315300410>
 19. EREN, Kübra; TAKTAKOĞLU, Nursema a PIRIM, Ibrahim, 2022. DNA Sequencing Methods: From Past to Present. [online], *Eurasian Journal of Medicin.* [cit. 2025-06-08] Dostupné z: <https://doi.org/10.5152/eurasianjmed.2022.22280>.
 20. MUZZEY, Dale; EVANS, Eric A. a LIEBER, Caroline, 2015. Understanding the Basics of NGS: From Mechanism to Variant Calling. [online], *Genetic Counseling and Clinical Testing*. Vol. 3 [cit. 2025-06-08] Dostupné z: <https://link.springer.com/article/10.1007/s40142-015-0076-8>.
 21. PAVLÍKOVÁ, Magdaléna, 2013. *Využití metod paralelního sekvenování v mikrobiologii*. [online], Bakalářská práce. Praha: Univerzita Karlova Přírodovědecká fakulta. [cit. 2025-06-23]. Dostupné z: https://dspace.cuni.cz/bitstream/handle/20.500.11956/56386/BPTX_2012_1_11310_0_319328_0_130268.pdf?sequence=1&isAllowed=y.
 22. BAČOVÁ, Zuzana, 2015. *Transkriptomové sekvenování u akutní lymfoblastické leukémie*. [online], diplomová práce. Olomouc: Katedra buněčné biologie a genetiky, PřF UP Olomouc. [cit. 2025-06-08]. Dostupné z: <https://theses.cz/id/1k5cp3/15449316>.
 23. LIU, LIN, YINHU LI, SILIANG LI, et al. Comparison of Next-Generation Sequencing Systems. *Journal of Biomedicine and Biotechnology*. [online], 2012, 1-11 [cit. 2022-06-07]. ISSN 1110-7243. Dostupné z: [doi:10.1155/2012/251364](https://doi.org/10.1155/2012/251364)
 24. ZHANG, Lu; CHEN, FengXin; ZENG, Zhan a XU, Mengjiao, 2021. Advances in Metagenomics and Its Application in Environmental Microorganisms. [online], *Frontiers in Microbiology*. Vol. 12. [cit. 2025-06-08]. Dostupné z: <https://doi.org/10.3389/fmicb.2021.766364>.
 25. MAJZLÍKOVÁ, Gabriela, 2020. *Analýza struktury genomu a genetické variability triploidních banánovníků - plantainů (genom AAB)*. [online], Diplomová práce. Olomouc: Univerzita Palackého v Olomouci. [cit. 2022-04-03]. Dostupné z: <https://theses.cz/id/nddzhr/Majzlikova.Gabriela.2020.pdf>.

26. KE, Rongqin; MIGNARDI, Marco; HAULING, Thomas a NILSSON, Mats, 2016. Fourth Generation of Next-Generation Sequencing Technologies: Promise and Consequences. [online], *Next Generation Sequencing*. Vol. 37. [cit. 2025-06-24]. Dostupné z: <https://onlinelibrary.wiley.com/doi/full/10.1002/humu.23051>.
27. *Učební text ke kurzu APLBIO2* [online]. 1-2 [cit. 2022-06-13]. Dostupné z: https://is.muni.cz/www/98951/41610771/43823411/43823458/Analyza_a_hodnoc/44563155/00_Shlukova_analyza_Hierarchicka_FINAL.pdf
28. DUNN, Heather a Laurie QUINN. *Cluster Analysis in Nursing Research: An Introduction, Historical Perspective, and Future Directions* [online], 2017, 10-11 [cit. 2022-06-13]. Dostupné z: doi:10.1177/0193945917707705
29. ALLENDES OSORIO, Rodolfo S., Lokesh P. TRIPATHI a Kenji MIZUGUCHI. *CLINE: a web-tool for the comparison of biological dendrogram structures*. [online], 2019, 20 [cit. 2022-06-13]. Dostupné z: doi:10.1186/s12859-019-3149-y
30. NAOREM, Romen Singh; URBAN, Peter; GOSWAMI, Gunajit a FEKETE, Csaba, 2020. *Characterization of methicillin-resistant Staphylococcus aureus through genomics approach*. [online], In: Springer nature link. [cit. 2025-06-08]. Dostupné z: <https://link.springer.com/article/10.1007/s13205-020-02387-y/figures/4>.
31. GAO, Xiao-Yang; ZHI, Xiao-Yang; LI, Hong-Wei a KLENK, Hans-Peter. Comparative Genomics of the Bacterial Genus Streptococcus Illuminates Evolutionary Implications of Species Groups. [online], [cit. 2025-06-20]. Dostupné z: <https://doi.org/10.1371/journal.pone.0101229>.
32. TAHMASEBI, Hamed; DEHBASHI, Sanaz; NASAJ, Mona a ARABESTANI, Mohammad Reza, 2022. Molecular epidemiology and collaboration of siderophore-based iron acquisition with surface adhesion in hypervirulent *Pseudomonas aeruginosa* isolates from wound infections. [online], *Scientific reports*. Č. 12. [cit. 2025-06-20]. Dostupné z: <https://doi.org/10.1038/s41598-022-11984-1>.