



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA CHEMICKÁ

FACULTY OF CHEMISTRY

ÚSTAV FYZIKÁLNÍ A SPOTŘEBNÍ CHEMIE

INSTITUTE OF PHYSICAL AND APPLIED CHEMISTRY

**VYTIPOVÁNÍ A SLEDOVÁNÍ EXPRESE GENŮ
OVLIVŇUJÍCÍCH SYNTÉZU KYSELINY HYALURONOVÉ
VE STREPTOCOCCUS EQUI SUBSP. ZOOEPIDEMICUS
POMOCÍ TECHNOLOGIE DNA ČIPŮ A REAL TIME PCR**

STUDYING OF GENE EXPRESSION INVOLVED IN HYALURONIC ACID SYNTHESIS IN STREPTOCOCCUS
EQUI SUBSP. ZOOEPIDEMICUS USING DNA MICROARRAYS AND REAL-TIME PCR

DIZERTAČNÍ PRÁCE

DOCTORAL THESIS

AUTOR PRÁCE

AUTHOR

Ing. Radka Hrudíková

ŠKOLITEL

SUPERVISOR

doc. RNDr. Vladimír Velebný, CSc.

BRNO 2020

Zadání dizertační práce

Ústav: Ústav fyzikální a spotřební chemie Akademický rok: 2019/20
Studentka: **Ing. Radka Hrudíková**
Studijní program: Fyzikální chemie
Studijní obor: Fyzikální chemie
Vedoucí práce: **doc. RNDr. Vladimír Velebný, CSc.**

Název dizertační práce:

Vytipování a sledování exprese genů ovlivňujících syntézu kyseliny hyaluronové ve streptococcus equi subsp. zooepidemicus pomocí technologie dna čipů a real time PCR

Zadání dizertační práce:

Charakterizace genomu bakteriálního kmene Streptococcus equi subsp. zooepidemicus sekvenací chromozomální DNA.

Sledování exprese genů u bakteriálních kmenů rodu Streptococcus za využití metody DNA čipů a real time PCR.

Identifikace genů zapojených přímo, nebo nepřímo do metabolismu kyseliny hyaluronové v bakteriálním kmenu Streptococcus equi subsp. zooepidemicus za různých kultivačních podmínek (suplementace kultivačního média N-acetylglukózaminem).

Termín odevzdání dizertační práce: 10.6.2020:

Ing. Radka Hrudíková
student(ka)

doc. RNDr. Vladimír Velebný, CSc.
vedoucí práce

prof. Ing. Miloslav Pekař, CSc.
vedoucí ústavu

V Brně dne 1.9.2019

prof. Ing. Martin Weiter, Ph.D.
děkan

ABSTRAKT

Kyselina hyaluronová (HA) je biologicky významnou látkou, která je využívána v potravinářském, farmaceutickém a především kosmetickém průmyslu. Jde o glykosaminoglykan běžně se vyskytující v lidském těle. Jako jeden z virulentních faktorů ji některé bakteriální kmeny produkují ve formě kapsule. Obaluje bakteriální buňku za účelem ochrany před imunitním systémem napadeného organismu. Jedním z hlavních producentů je *Streptococcus equi* subsp. *zooepidemicus*. Tato Gram-pozitivní bakterie je využívána ve společnosti Contipro a.s. k výrobě HA. Produkční kmen CO4A byl získán náhodnou fyzikální mutagenezí. Cílem této práce bylo prostudovat změny v genomu a transkriptomu kmene, které vedly k významnému navýšení produkce HA. Genom kmene CO4A byl osekvenován a porovnán s referenčním kmenem ATCC35246 [1]. Velikost genomu je 2 167 251 bp a bylo v něm identifikováno 83 relevantních variant (59 SNV a 34 indelů). Varianty v kódujících oblastech byly anotovány a byly také určeny změny aminokyselinové sekvence. U SNV mutací došlo ke změně aminokyselinové sekvence v 45 případech, u indel mutací byla změna aminokyselinové sekvence identifikována vždy. U obou kmenů byla sledována míra exprese vybraných skupin genů metodou DNA čipů. U kmene CO4A byla pozorována kaskáda zvýšené míry exprese genů metabolismu aminocukrů, vedoucí k syntéze UDP-*N*-acetylglukosaminu (navýšení exprese u těchto genů v porovnání s ATCC35246 bylo v průměru o 28 %). Následně byla míra exprese vybraných genů ověřována real-time PCR (qPCR). Zde nebyl identifikován signifikantní rozdíl míry exprese u genů *has* operonu u obou kmenů. Metodou qPCR byl sledován také vliv suplementace kultivačního média *N*-acetylglukosaminem (GlcNAc), který je jedním z prekurzorů syntézy HA. Byl zaznamenán pozitivní vliv suplementace kultivačního média externím GlcNAc na expresi vybraných skupin genů a také na výtěžek HA z média po ukončení kultivace u kmene CO4A (nárůst výtěžku v průměru o 17 %). U kmene ATCC35246 byl sledován pozitivní vliv na výtěžek HA z kultivací (nárůst výtěžku v průměru o 9 %), ale nebyla zjištěna signifikantní změna míry exprese u vybraných skupin genů.

KLÍČOVÁ SLOVA

kyselina hyaluronová, *Streptococcus zooepidemicus*, genová exprese, celogenomová sekvenace, DNA čipy, real-time PCR, referenční gen, bioinformatika, regulace exprese genů

ABSTRACT

Hyaluronic acid (HA) is an important substance, which is mostly used in pharmaceutical and cosmetic industry. This substance is commonly found in the human body. HA is one of the factors contributing to virulence of microorganisms. Some bacterial strains produce hyaluronic acid in the form of a mucoid capsule that encapsulates the cell to protect bacteria against the immune system of the host organism. One of the main producers is the bacterial strain *Streptococcus equi* subsp. *zooepidemicus*. Contipro a.s. uses the strain CO4A to produce hyaluronic acid in large scale. The production strain was obtained by random mutagenesis by UV light. The aim of the work was to study changes in the genome, which led to a significant increase in hyaluronic acid production, using DNA microarray and real-time PCR (qPCR). The genome of the strain CO4A was sequenced and compared to reference ATCC35246 [1]. The size of the genome is 2,167,251 bp and 83 relevant variants (59 SNV and 34 indels) have been identified. Variants in coding regions were annotated and amino acid sequence changes were determined. In SNV mutations there was a change in the amino acid sequence in 45 cases. The change was identified in every case of indel mutations. The expression level of selected groups of genes was monitored in both strains by the method of DNA microarrays. A cascade of increased expression level of amino sugar metabolism genes leading to the synthesis of UDP-*N*-acetyl glucosamine was observed in strain CO4A (the increase in expression level of these genes compared to ATCC35246 was on average 28 %). Subsequently, the expression of selected genes was verified by qPCR. There was no significant difference in the expression level of the *has* operon genes of both strains. The effect of supplementation of the culture medium with *N*-acetylglucosamine (GlcNAc), which is one of the precursors of HA synthesis, was also studied by qPCR. A positive effect of the supplementation of the culture medium with external GlcNAc in the CO4A strain has been recorded. Also, the supplementation has positive effect on the yield of HA from the medium (increase in yield was on average by 17 %). GlcNAc has been shown to have a positive effect on the yield of HA in ATCC35246 strain as well (increase in yield was 9 % on average), but no significant changes in the expression levels were found in selected groups of genes in ATCC35246.

KEY WORDS

hyaluronic acid, *Streptococcus equi* subsp. *zooepidemicus*, gene expression, whole genome sequencing, DNA microarrays, real-time PCR, housekeeping gene, bioinformatics, regulation of gene expression

HRUDÍKOVÁ, R.. *Vytipování a sledování exprese genů ovlivňujících syntézu kyseliny hyaluronové ve Streptococcus equi subsp. zooepidemicus pomocí technologie dna čipů a real-time PCR*. Brno: Vysoké učení technické v Brně, Fakulta chemická, 2020. 121s. Vedoucí dizertační práce: doc. RNDr. Vladimír Velebný, CSc.

Prohlášení:

Prohlašuji, že jsem dizertační práci vypracovala samostatně a že všechny použité literární zdroje jsem správně a úplně citovala. Dizertační práce je z hlediska obsahu majetkem Fakulty chemické Vysokého učení technického v Brně a může být využita ke komerčním účelům jen se souhlasem vedoucího dizertační práce a děkana FCH VUT.

.....

podpis studenta

Poděkování:

Zde bych ráda poděkovala mému školiteli doc. RNDr. Vladimíru Velebnému, CSc. za příležitost provádět výzkum ve společnosti Contipro a.s. a seznámit se s prací v laboratoři na špičkové úrovni. Dále bych chtěla poděkovat Mgr. Janě Jilkové za vedení a cenné rady při směřování mé práce a Mgr. Stanislavu Pepeliaevovi Ph.D. za supervizi v průběhu její experimentální části. Velké díky také patří Mgr. Matoušovi Čihákovi, který se mnou prací velmi intenzivně konzultoval a pomáhal mi ji směřovat až do její konečné podoby.

Dále bych ráda poděkovala své rodině a zejména partnerovi za podporu, protože vím, že to se mnou v době vypracovávání práce nebylo vůbec jednoduché.

OBSAH

1	ÚVOD.....	9
2	TEORETICKÁ ČÁST.....	10
2.1	Kyselina hyaluronová (HA)	10
2.1.1	<i>Historie objevu a produkce HA</i>	11
2.1.2	<i>Mikrobiální produkce</i>	11
2.2	Biosyntéza HA ve <i>S. zooepidemicus</i>	12
2.2.1	<i>has operon</i>	14
2.2.2	<i>HA syntáza</i>	14
2.3	Sekvenace celogenomové dna <i>S. zooepidemicus</i>	17
2.4	Studium genové exprese <i>S. zooepidemicus</i> – DNA čipy	20
2.5	Studium genové exprese <i>S. zooepidemicus</i> – qPCR	22
3	CÍL PRÁCE.....	24
3.1	Charakterizace bakteriálního kmene <i>S. zooepidemicus</i> CO4A	24
3.2	Sledování exprese genů u vybraných kmenů <i>S. zooepidemicus</i> CO4A	24
4	MATERIÁL A METODY.....	25
4.1	Bakteriální kmeny a jejich kultivace.....	25
4.1.1	<i>Kultivace bakteriálního kmene pro sekvenaci celogenomové DNA</i>	25
4.1.2	<i>Kultivace bakteriálních kmenů pro analýzu na DNA čípech</i>	25
4.1.3	<i>Kultivace bakteriálních kmenů pro analýzu qPCR</i>	26
	Kultivace s N-acetylglukosaminem (GlcNAc) a bez jeho suplementace	26
	Zpracování biomasy po kultivacích srážecí metodou	26
4.2	Izolace chromozomální DNA.....	26
4.3	Izolace RNA a zjištění její kvality	26
4.4	Přepis RNA do cDNA pomocí reverzní transkripce	27
4.5	Sekvenace celogenomové DNA.....	27
4.6	Analýza genomu SEZ CO4A	28
4.7	Sledování exprese genů – DNA čipy	28
4.7.1	<i>Skenování čipu a zpracování dat</i>	28
4.7.2	<i>Výpočet poměru celkové exprese sledovaného genu v genomu CO4A/ATCC</i>	29
4.7.3	<i>Výpočet poměru celkové exprese genu v genomu CO4A/ATCC normalizováno vůči expresi gyrázy A (gyrA)</i>	29
4.8	Sledování exprese genů – qPCR	29

4.8.1	Provedení analýzy qPCR.....	29
	Návrh primerů.....	29
	Normalizace koncentrace templátu.....	29
	Negativní kontrola qPCR.....	30
	SYBR Green fluorescenční značení.....	30
4.8.2	Stanovení referenčního genu pro qPCR.....	30
4.8.3	Analýza C_t hodnot genů <i>has operonu</i>	31
4.8.4	Analýza C_t hodnot vybraných skupin genů metodou $2^{(-\Delta\Delta Ct)}$	31
5	VÝSLEDKY A DISKUZE.....	32
5.1	Sekvence chromozomální DNA <i>Streptococcus equi</i> subsp. <i>zooepidemicus</i> CO4A.....	32
5.2	Analýza genomu <i>S. zooepidemicus</i> CO4A.....	33
5.3	Sledování exprese genů pomocí DNA čipů.....	49
5.3.1	<i>Geny energetického metabolismu</i>	49
	Glykolýza a glukoneogeneze.....	49
	Metabolické dráhy glutaminu a glutamátu.....	55
	Metabolické dráhy aminocukrů a nukleotidů.....	56
5.3.2	<i>Geny metabolických drah syntézy buněčné stěny</i>	62
	Peptidoglykan.....	62
	Glycerolipidy.....	66
5.3.3	<i>Geny replikace DNA a degradace RNA</i>	69
	Replikace DNA.....	69
	Degradace RNA.....	71
5.3.4	<i>Geny has operonu a vybraných metabolických drah</i>	73
	<i>has operon</i>	73
	Vybrané metabolické dráhy.....	75
5.4	Sledování exprese vybraných genů pomocí metody qPCR.....	78
5.4.1	Stanovení referenčního genu.....	78
5.4.2	Ověření exprese genů <i>has operonu</i>	85
5.4.3	Vliv <i>N-acetylglukosaminu</i> na expresi vybraných skupin genů.....	88
	Analýza změn relativní genové exprese metodou $2^{(-\Delta\Delta Ct)}$	90
5.4.4	Výtěžek HA z provedených kultivací.....	99
6	ZÁVĚR.....	102
7	PŘEHLED LITERATURY.....	105

8 SEZNAM POUŽITÝCH ZKRATEK	118
9 STRUČNÝ ŽIVOTOPIS AUTORA	119
10 SEZNAM PUBLIKACÍ AUTORA.....	120

1 ÚVOD

Česká biotechnologická společnost Contipro a.s. se zabývá produkcí a využitím kyseliny hyaluronové (HA) pro kosmetický a farmaceutický průmysl. HA má široké využití především proto, že je přirozenou součástí mezibuněčné hmoty živých organismů a hlavně člověka. Navíc má hydrofilní vlastnosti a biologické účinky (regenerace pokožky, hojení ran aj.), díky kterým je právě využívána ve farmaceutických a kosmetických přípravcích. Společnost k její produkci využívá bakteriálního kmene *Streptococcus equi* subsp. *zooepidemicus*, který je jejím přirozeným producentem. HA u tohoto bakteriálního kmene představuje jeden z nejdůležitějších virulentních faktorů. Kmen produkuje HA ve formě kapsule, která obaluje celý povrch bakterie a tím ji skrývá před buňkami imunitního systému napadeného organismu. Bakteriální kmen, který využívá k produkci společnost Contipro, byl získán náhodnou fyzikální mutagenézí a je nepatogenní (nehemolytický). Cílem této práce je především zmapování změn, ke kterým v genetické výbavě kmene došlo. Hlavním a nejvýraznějším důsledkem mutace kmene byla právě velmi vysoká produkce HA. Proto jsme se pokusili najít změny v genomu a spojit je s metabolickými drahami syntézy HA, aby bylo možné produkci ovlivňovat či cíleně upravovat. Ovlivňování její produkce přes cílené úpravy se stalo i nástrojem k prohloubení znalostí mechanismu tvorby tohoto virulentního faktoru a jeho významu pro růst buněk v laboratorních podmínkách [2, 3, 4, 5].

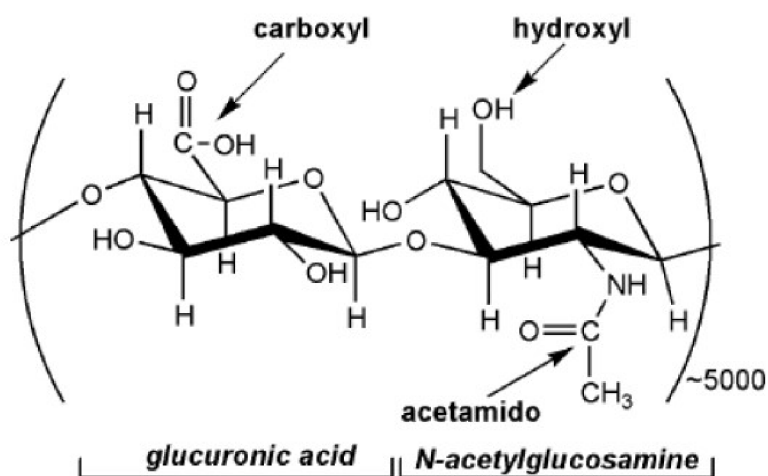
Aby byla problematika komplexně prostudována, byla velká část teorie věnována metabolickým drahám syntézy HA v bakteriálním druhu *S. zooepidemicus* a jejich regulačním mechanismům. Další kapitoly byly věnovány metodám, kterými byly tyto dráhy studovány na základě genové exprese. Prvním krokem bylo zjištění sekvence celého genomu kmene studovaného a referenčního metodou celogenomové sekvenace. Následovala analýza získaných dat s vyhledáváním variant. Dále byly provedeny kontrolní kultivace a odběry pro izolaci RNA, která byla použita pro transkriptomickou analýzu metodou DNA čipů a qPCR.

Poslední část práce byla zaměřena na sledování změn exprese genů u metabolických drah, které jsou přímo či nepřímo zapojeny v produkci kyseliny hyaluronové, nebo ji mohou výrazně ovlivnit. Jednalo se především o dráhy syntézy buněčné stěny, peptidoglykanu a glykolýzy. Na základě výsledků publikovaných v roce 2013 v Austrálii [6] jsme se rozhodli inspirovat se jejich experimentem a ověřit vliv suplementace kultivačního média *N*-acetylglukosaminem (GlcNAc) na biosyntézu HA u kmene ATCC35246. GlcNAc je jedním z prekurzorů syntézy HA a jeho suplementace měla dle výše uvedené studie pozitivní dopad na její produkci. Výsledky získané molekulárně-biologickými metodami jsou doplněné analýzou výtěžku HA z experimentálních kultivací.

2 TEORETICKÁ ČÁST

2.1 Kyselina hyaluronová (HA)

HA je přirozeně se vyskytující glykosaminoglykan, který nabízí mnoho jedinečných výhod jako stavební jednotka pro biomateriály. V lidském těle se vyskytuje ve formě hyaluronátu sodného. Při fyziologickém pH existuje v náhodném klubkovitém uspořádání [7]. Nejvyšší koncentrace HA v lidském těle můžeme zaznamenat v kůži, v pupeční šňůře či ve sklivci očí. V mikrobiální říši se s HA nejčastěji setkáváme jako s extracelulární kapsulí některých druhů bakterií rodu *Streptococcus* (např. *S. zooepidemicus*, *S. pyogenes*, *S. equisimilis*) [8]. Je to neimunogenní enzymaticky degradovatelný lineární polysacharid o vysoké molekulové hmotnosti (4 000-8 000 000 Da) [9], skládající se z disacharidických jednotek, které jsou tvořeny D-glukuronovou kyselinou (GlcUA) a *N*-acetylglukosaminem (GlcNAc) spojenými β -1-3 a β -1-4 glykosidickou vazbou (obr. 1) [10]. Do skupiny glykosaminoglykanů patří dále například chondroitin sulfát, dermatan sulfát a heparan sulfát. Kyselina hyaluronová neobsahuje na rozdíl od ostatních ze skupiny sulfátovou skupinu a interaguje nekovalentně s komplexy proteinů. Stále více se dozvídáme o biologických funkcích HA, které jsou zprostředkovány specifickými receptory na povrchu buněk, jako jsou CD44 a RHAMM [11]. Vysoká koncentrace HA je zaznamenávána v oblastech proliferace a migrace buněk, dále je důležitá pro morfogenezi a regeneraci tkání [10].



Obrázek 1: Struktura kyseliny hyaluronové [10].

HA má významnou roli v mnoha procesech zahrnujících například angiogenezi, homeostázy, hojení ran a podobně. Ve vysoké koncentraci je přítomna v extracelulární matrix. Svými vlastnostmi si našla široké uplatnění v kosmetickém průmyslu, biomedicině, farmacii a potravinářském průmyslu. Komerční produkce je zaměřena na široké spektrum použití jak HA, tak jejích derivátů (př. oční operace) [7]. V současnosti se výzkum HA rozšiřuje i do oborů jako jsou nanotechnologie, nanomedicína, tkáňové inženýrství, metabolické inženýrství a podobně [12]. Tradičním způsobem získávání HA byla extrakce z kohoutích hřebínků, teď se vyrábí především mikrobiální produkcí s nízkou produkční cenou a menším dopadem na životní prostředí [13]. Je úspěšně produkována v průmyslovém měřítku zejména vhodnými

bakteriálními kmeny rodu *Streptococcus*, které mají vysokou výtěžnost a nízké náklady na kultivační podmínky.

2.1.1 Historie objevu a produkce HA

V roce 1934 Karl Meyer a John Palmer popsali nový polysacharid izolovaný z bovinního sklivce očí. Našli látku obsahující uronovou kyselinu a aminosacharid, kterou pojmenovali hyaluronová kyselina jako spojení slov hyaloid (sklivce) a uronová kyselina. Dále také popsali její chemickou strukturu [14].

Termín hyaluronan byl však zaveden až roku 1986. V letech 1930–1940 byla HA izolována z mnoha různých zdrojů, od sklivce a kohoutích hřebínků po bakterie rodu *Streptococcus*. Fyzikálně chemická charakterizace kyseliny hyaluronové byla provedena v období 1950–1960. Výzkumníci zjistili extrémně vysokou viskozitu při velmi nízké koncentraci a výjimečné rheologické vlastnosti. Díky tomu HA může sloužit jako lubrikant, stabilizátor struktur, regulátor obsahu vody a může tvořit hydrogely. Viskoelastické vlastnosti jsou kontrolovány koncentrací a molekulovou hmotností použité kyseliny hyaluronové [15, 16]. Původní HA jako produkt použitelný v medicíně a farmacii byl připraven Endre A. Balazsem, který připravil první nezánětlivou a vysoce čistou vysokomolekulární HA z pupečních šňůr a kohoutích hřebínků [17]. Začátkem osmdesátých let byla HA poprvé použita k vytvoření nitroočních čoček pro implantaci. Po tomto úspěchu se stala hlavním materiálem používaným k očním operacím [17].

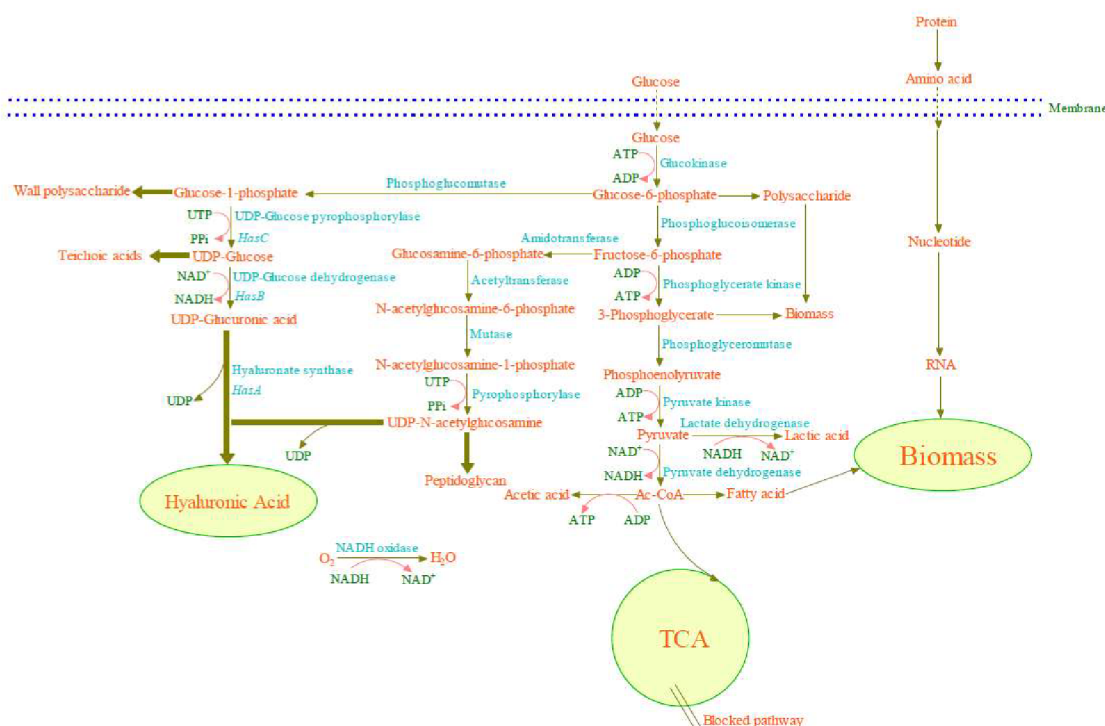
Extrakce kyseliny hyaluronové z kohoutích hřebínků postupem času přestala stačit poptávce na trhu a především požadavkům nových biomedicínských a farmaceutických aplikací. Jako alternativní produkce kyseliny hyaluronové byla zavedena mikrobiální fermentace [4, 18]. První fermentačně připravená HA pro komerční účely byla produkována bakteriálním druhem *S. equi* subsp. *zooeidemicus*, který zůstává nejpoužívanějším v industriální produkci. Po kmenech rodu *Streptococcus* lze HA produkovat i v jiných bakteriích, jak přirozenou cestou s nižší účinností, tak za pomoci metod metabolického a genového inženýrství. Můžeme uvést jak Gram negativní bakterie (*Agrobacterium sp.*, *Escherichia coli*), tak i Gram pozitivní (*Bacillus sp.*, *Lactococcus lactis*) [19, 20, 21, 22, 23].

2.1.2 Mikrobiální produkce

Kyselina hyaluronová získávaná extrakcí z kohoutích hřebínků nebo bovinního sklivce tvoří komplex s proteoglykany. Je proto složité izolovat vysokomolekulární HA, jelikož několikanásobná extrakce způsobuje nežádoucí redukci molekulové hmotnosti. Velmi důležitým faktorem je také využívání živočišného produktu k léčbě lidí a s tím spojená rizika imunitní odpovědi, virové infekce a podobně. Proto se průmysl přiklonil k mikrobiální produkci. Bakteriální kmeny produkují HA jako extracelulární polysacharid, který je nekovalentně vázaný k buněčné stěně a může být vlivem vnějšího prostředí uvolňován do živného média. Množství tohoto polymeru produkovaného bakteriální fermentací může být teoreticky neomezené. Prakticky je to už komplikovanější. S přibývajícím produktem roste viskozita média, mění se pH kultivačního prostředí a dochází i k dalším změnám v průběhu fermentace, které vedou ke zpomalení či zastavení syntézy [24].

Bakteriální kmeny produkují HA nejčastěji ve formě mukoidní kapsule pokrývající celou buňku. Kapsule je biokompatibilní s hostitelským organismem a je jedním z faktorů virulence [25, 26]. Chrání bakteriální buňku před působením buněk imunitního systému a zvyšuje její

přilnavost ke sliznicím hostitelského organismu [27]. Ve kmenech rodu *Streptococcus* je HA produkována jako sekundární metabolit a není tedy přímo zapojená v procesu růstu, vývoje, nebo rozmnožování buňky). Její produkce je podmíněna geneticky a může být ovlivněna nutričními i dalšími faktory. Běžně používaný *S. zooepidemicus* (přirozený producent) může vyprodukovat až 6 g/l HA za vhodných kultivačních podmínek a v závislosti na vybraném kmenu. Vysoká viskozita roztoku i s nízkou koncentrací HA způsobuje komplikace spojené s mícháním a vzdušením fermentačních reaktorů. Z hlediska metabolismu zde soupeří o společné prekurzory (UDP-GlcNAc a UDP-GlcUA) metabolické dráhy syntézy HA a buněčného růstu (energetický metabolismus, biosyntéza peptidoglykanu a polysacharidů buněčné stěny). Posledním významným faktorem je syntéza kyseliny mléčné, která je produktem metabolismu cukrů. Její vysoká koncentrace (až 65 g/l [28]) vede k silné inhibici buněčného růstu a produkce HA, tím že snižuje hodnotu pH, čímž deprimuje metabolismus. Kontinuální úprava pH zase zvyšuje koncentraci solí v médiu a mění osmotické prostředí buňky [8]. Souhrnný přehled metabolických drah, propojených s metabolickou dráhou produkce kyseliny hyaluronové je uveden na obr. 2.



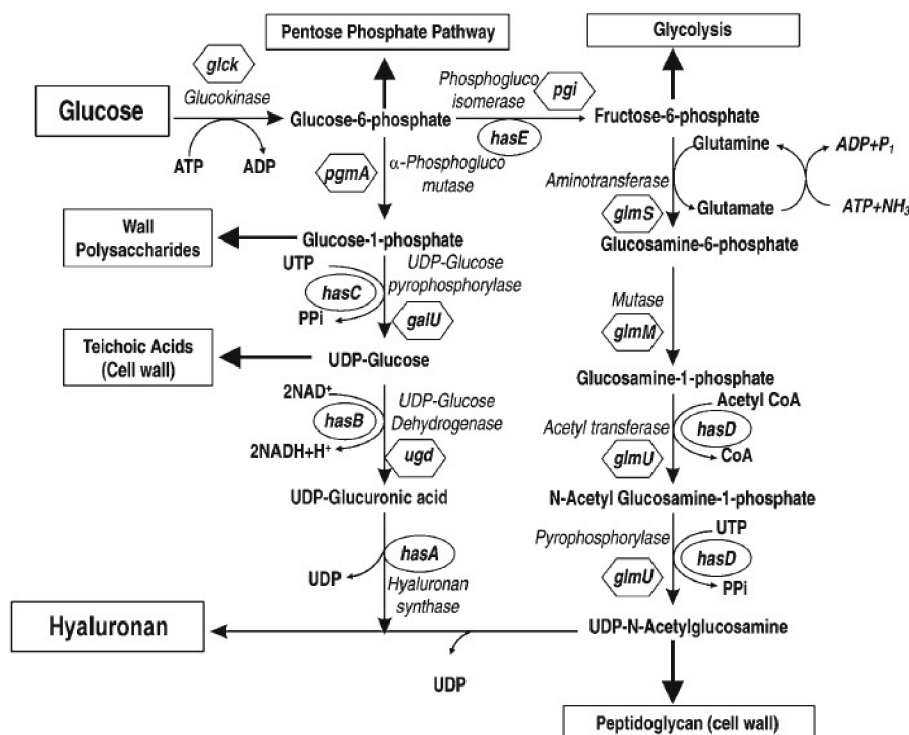
Obrázek 2: Souhrnný přehled propojení metabolických drah zapojených v produkci HA [8].

2.2 Biosyntéza HA ve *S. zooepidemicus*

Kyselina hyaluronová je derivátem monosacharidických jednotek GlcNAc a GlcUA. Syntéza HA je rozdělena do tří fází. V první fázi se syntetizuje UDP-glukuronová kyselina (UDP-GlcUA). Glukóza-6-fosfát je konvertován α -fosfoglukomutázou (*PgmA*) na glukóza-1-fosfát reverzibilní reakcí. UDP-glukóza pyrofosforyláza (*GalU*) následně katalyzuje reakci UTP a glukózy-1-fosfátu za vzniku UDP-glukózy. UDP-glukuronová kyselina je získána specifickou oxidací primární alkoholové skupiny na UDP-glukóze katalyzované UDP-glukóza dehydrogenázou (*Ugd*) [10].

Druhá fáze je zodpovědná za syntézu UDP-*N*-acetylglukosaminu (UDP-GlcNAc) a začíná reakcí fruktózy-6-fosfátu. Aminokupina z glutaminu je přenesena na fruktózu-6-fosfát za katalýzy amidotransferázou (*GlmS*) a za vzniku glukosamino-6-fosfátu. Následující krok je velmi energeticky náročný. Fosfátové skupiny jsou přeskupeny za působení mutázy (*GlmM*), což vede ke vzniku *N*-acetyl-glukosamin-1-fosfátu. Následně je acetylová skupina přesunuta na glukosamin-6-fosfát za pomoci acetyltransferázy a za vzniku *N*-acetyl glukosaminu-6-fosfátu. Nakonec za katalýzy pyrofosforylázou (bifunkční enzym *GlmU*) UTP reaguje s *N*-acetyl-glukosamin-1-fosfátem a vzniká UDP-GlcNAc.

Ve třetí fázi jsou nasyntetizované aktivované prekurzory spojovány HA syntázou (*HasA*) do konečné struktury kyseliny hyaluronové (obr. 3) [10].



Obrázek 3: Biosyntéza HA ve *S. zooepidemicus* [21].

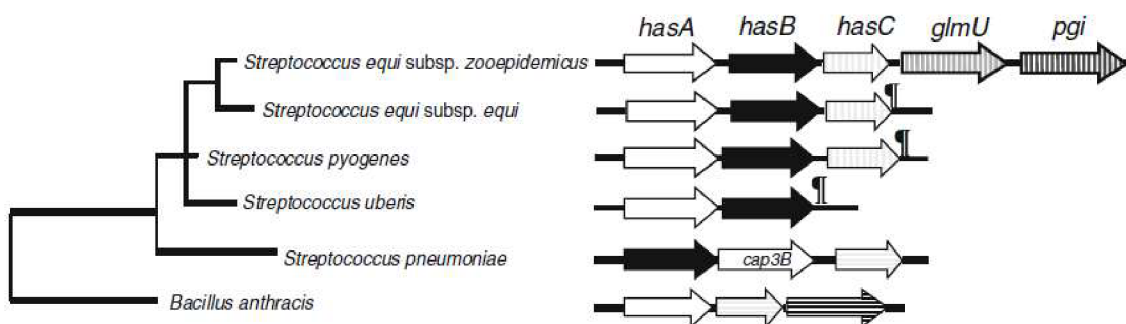
Celkové množství ATP spotřebované na přípravu 1 molu disacharidických jednotek HA jsou 4 moly, z toho 2 moly ATP jsou spotřebovány ve dvou glukokinázových reakcích a 2 moly ATP jsou spotřebovány na regeneraci donorových UTP. Oxidační reakce katalyzovaná UDP-glukóza dehydrogenázou uvolňuje 2 moly NADH na 1 mol nasyntetizovaných HA disacharidických jednotek [10].

Tři klíčové enzymy jsou zapojeny do první části dráhy – konverze glukózy-6-fosfátu na UDP-glukuronovou kyselinu. Jsou to fosfoglukomutáza (gen *pgmA*), UDP-glukóza pyrofosforyláza (geny *hasC*, nebo *galU*) a v neposlední řadě UDP-glukóza dehydrogenáza (geny *hasB*, nebo *ugd*). Kritickými enzymy pro syntézu druhého prekurzoru – UDP-*N*-acetylglukosaminu – jsou fosfoglukoizomeráza (geny *hasE*, nebo *pgi*) a bifunkční enzym glukosamin-1-fosfát acetyltransferáza/*N*-acetylglukosamin-1-fosfát uridylyltransferáza (geny *hasD*, nebo *glmU*). Iniciační molekuly (prekurzory) pro syntézu HA jsou také intermediáty tvorby buněčné stěny a mohou být rovněž zpracovány glykolýzou či pentózovým cyklem [29].

2.2.1 *has operon*

Enzymy, zapojené do metabolické dráhy kyseliny hyaluronové, jsou u streptokoků kódovány *has* operonem. Počet genů, z nichž se *has* operon skládá, se může lišit od 3 do 5 u různých kmenů [30]. U přirozeného producenta kyseliny hyaluronové *S. zooepidemicus* ATCC35246 a příbuzných kmenů (*S. zooepidemicus* H70, *S. zooepidemicus* MGCS10565) je všech pět genů (*hasA*, *hasB*, *hasC*, *hasD*, *hasE*) lokalizováno na stejném lokusu genomu v *has* operonu. Transkripce těchto genů je pod kontrolou místa počátku transkripce a σ^{70} promotoru umístěného 50 bp před *hasA* [31, 32, 10]. Tyto kmeny mají ve svém genomu také jednu kopii *hasC* genu navíc. Kopie *hasC* byla nalezena v *gps* operonu obsahujícím pouze dva geny: *hasC* a *gpsA* gen, který kóduje NAD(P)⁺ dependentní glycerol-3-fosfát dehydrogenázu. Dodatečná kopie *hasC* má za následek stabilizaci a zlepšení produkce HA v buňkách [30]. Bylo prokázáno, že biosyntéza HA je v buňkách *Streptococcus sp.* kontrolována na úrovni transkripce silou promotoru *has* operonu a regulována dvousložkovým regulačním systémem *csrS/csrR* (respektive *arls/R*, nebo *covR/S*) [33, 34, 5].

Byly charakterizovány *has* operony dvou patogenních zástupců *S. zooepidemicus* a *S. equi*, které jsou využívány ke komerční produkci HA [31]. Dále byla zkoumána skladba genů *has* operonu u těchto kmenů i s dalšími kmeny rodu *Streptococcus* a některými kmeny rodu *Bacillus*, který přirozeně formuje kapsuli z poly-D-gamma-glutamové kyseliny. Příbuznost operonů je naznačena na obrázku 4 evolučním distančním dendrogramem (Kimurův dvouparametrový model), založeným na analýze 16S rRNA. Dále je na schématu naznačeno uspořádání genů *has* operonu u *S. equi* subsp. *zooepidemicus* [31].



Obrázek 4: Skladba *has* operonu u *S. equi* subsp. *zooepidemicus* a dalších kmenů rodu *Streptococcus*, navíc se skladbou genů identifikovaných u *Bacillus anthracis* [31].

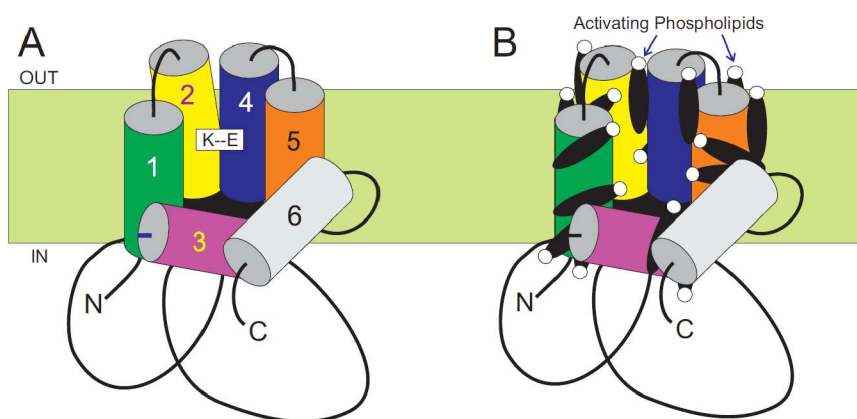
Gen *gpi* kóduje fosfoglukoizomerázu, která je pro růst bakterií na sacharidech jako uhlíkovém zdroji esenciální. Není ovšem esenciální pro produkci HA a proto je vhodná samostatná regulace transkripce *has* operonu a *gpi*. V té by mohl hrát roli silně konzervovaný σ^{70} promotor před *gpi*, který naznačuje, že tento gen je pod kontrolou alespoň dvou promotorů [31].

2.2.2 *HA syntáza*

Hyaluronan syntáza (HAS) je enzym kódovaný *hasA* genem umístěným na začátku *has* operonu a je esenciální pro produkci kyseliny hyaluronové. Existují dvě nepříbuzné skupiny HA syntáz [35]. Syntázy, které nacházíme u obratlovců a mimo jiné u bakterií rodu *Streptococcus*, patří do skupiny HAS I. HA syntázy, izolované z bakterie *Pasteurella multocida*, jsou řazeny do skupiny HAS II. Skupina HAS I se dále dělí na dvě podskupiny podle

směru polymerizace: podskupinu HAS I-R u níž probíhá polymerizace od redukujícího konce (streptokokální, myší a lidská HA) a podskupinu HAS I-N, u níž probíhá polymerizace od neredukujícího konce, patří sem například HAS z *Xenopus leavis* [36]. Syntázy třídy HAS I patří do GT2 skupiny glykozyl transferáz, která zahrnuje i β -glykozyl transferázy jako jsou celulóza syntáza a chitin syntáza [36]. Skupina HAS II je poněkud odlišná od ostatních syntáz. Polymerizace probíhá sice od neredukujícího konce (stejně jako u HAS I-N skupiny), ale vyskytují se především u Gram-negativních bakterií [36].

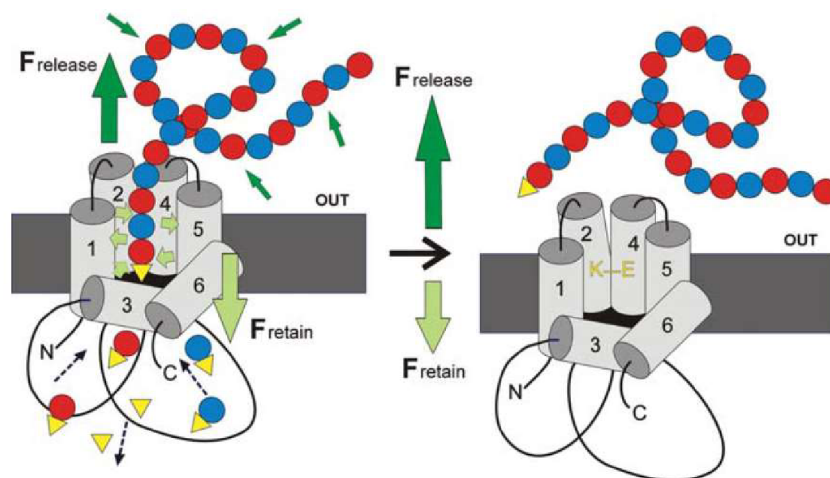
HA je u streptokoků syntetizována na cytoplasmatické membráně pomocí membránově asociované procesivní syntázy (obr. 5) se čtyřmi transmembránovými a dvěma k membráně asociovanými doménami [37, 38, 39, 40]. HAS je lipid-dependentní integrální protein a pro jeho aktivitu jsou vyžadovány specifické fosfolipidy tzv. kardiolipiny. Funkční katalytickou jednotkou je HAS monomer tvořící komplex se šestnácti kardiolipinovými molekulami [36, 41].



Obrázek 5: (A) Navrhované uspořádání transmembránových a membránově asociovaných domén HAS u *S. equi subsp. zooepidemicus* ve své neaktivní formě bez zakomplexování s kardiolipiny. Domény 1 a 2 a domény 4 a 5 jsou spojeny krátkými extracelulárními smyčkami, tudíž jsou asociovány jako sousední domény [38]. Dvojice polárních reziduí označenými v bílém rámečku K^{48} a E^{327} zprostředkovává interakce mezi těmito sousedními doménami. Domény 3 a 6 mohou být rozmístěny tak, jak je uvedeno na obrázku, nebo si mohou své pozice vyměnit [42]. (B) Navrhované uspořádání transmembránových a membránově asociovaných domén HAS u *S. equi subsp. zooepidemicus* ve své aktivní formě – tvoří komplex s fosfolipidy (černě vyznačeny s bílou „hlavičkou“). Fosfolipidy se aktivně zapojují do tvorby póru, kterým je zřejmě HA translokována přes membránu ven z buňky [43] [převzato z 36].

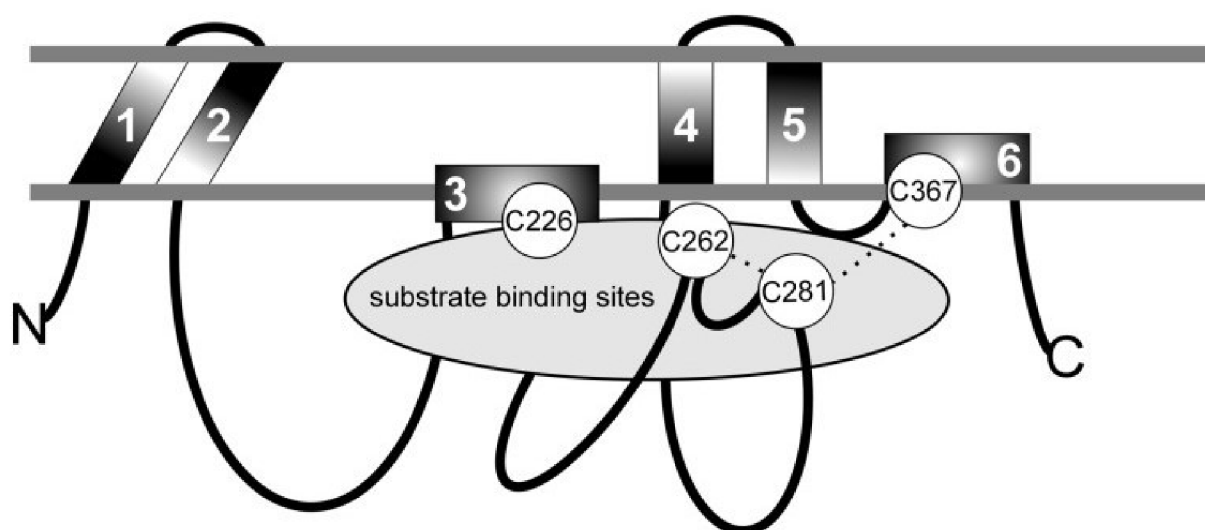
Syntéza HA probíhá *de-novo* prodlužováním polymeru z jeho redukujícího konce [44]. Nejsou tvořeny žádné kovalentně napojené intermediáty. Mechanismus translokace řetězce polymeru přes plazmatickou membránu, stejně jako zakončení polymerizace, není ještě zcela objasněn. Žádný přímý důkaz o zapojení HAS do translokace HA řetězce neexistuje, nicméně zapojení ABC transportéru bylo vyloučeno [10]. Zatím byla jen vyslovena hypotéza, že HA řetězec je extrudován skrze pór formovaný samotnou HA syntázou a terminace nastává stříhovou silou, která řetězec HA tlačí pryč od hyaluronan syntázy [43, 45]. Rostoucí řetězec zůstává připojený k HAS na buněčné membráně a zároveň je extrudován do extracelulárního prostoru (obr. 6). Vyslovená teorie může být odpovědí na otázku, jak HAS reguluje velikost

HA a jak může být komplex HAS-UDP-HA (UDP – uridin difosfát) stabilní až několik hodin [40].



Obrázek 6: HA – HAS retain/release model kontroly velikosti produktu u hyaluronan syntáz skupiny I. Levé schéma zobrazuje šest membránových domén HAS (šedé válce) tvořící pór, který obklopuje a váže rostoucí HA řetězec [46]. HA řetězec je tvořen střídavě GlcUA (červený kruh) a GlcNAc (modrý kruh), ke kterým je vázán UDP (žlutý trojúhelník) z redukujícího konce. Malé černé šipky ukazují zásobování aktivního místa substrátem UDP-GlcUA a UDP-GlcNAc, přičemž UDP reakční produkt je z místa odváděn. Malé světle zelené šipky naznačují působení sil, které se podílejí na udržení HA-HAS komplexu pohromadě (interakce s fosfolipidy, s cysteinovými rezidui a jiné). Působení těchto sil je znázorněno velkou světle zelenou šipkou. Malé tmavě zelené šipky zastupují interakce (vazba k externím proteinům, či matrix) a síly (střihová síla, rotační pohyb buňky, Brownův pohyb), které odlučují HA řetězec od HAS. Jejich působení je znázorněno jako velká tmavě zelená šipka nahoru. Síly udržující komplex HA – HAS pohromadě (F_{retain}) zůstávají po dobu syntézy HA více méně konstantní, zatímco síly odtrhávající HA řetězec od komplexu ($F_{release}$) přirozeně narůstají se zvětšující se velikostí řetězce. Na začátku polymerizace jsou F_{retain} mnohem větší než $F_{release}$, postupně se síly vyrovnávají, až v jistém momentu v průběhu elongace řetězce se stává $F_{release}$ větší než F_{retain} . Dochází k odtržení řetězce HA od HAS a k uvolnění do extracelulárního prostoru (schéma vpravo). Odtržený řetězec HA je zobrazen s navázaným UDP, protože není jasné, zda je proces uvolnění řetězce asociován se ztrátou UDP. V pravé části obrázku jsou naznačena rezidua K^{48} v doméně 2 a E^{327} v doméně 4 (žlutá písmena). Rezidua se podílejí na silách udržujících komplex HA – HAS navázán [převzato ze 40].

Ke stabilitě komplexu HAS-UDP-HA a regulaci velikosti HA přispívá i přítomnost cysteinových reziduí na HA syntáze. Membránově vázaná HAS má 4 cysteinová rezidua na pozicích 226, 262, 281 a 367. Cysteinová rezidua Cys²⁶², Cys²⁸¹ a Cys³⁶⁷ jsou umístěna velmi blízko sebe a společně s Cys²²⁶ jsou lokalizována na vnitřní straně membrány. Všechna tato rezidua jsou taktéž lokalizována velmi blízko UDP-prekurzor vazebného místa (obr. 7) [36].



Obrázek 7: Navrhované rozložení cysteinových reziduí a membránových domén kolem aktivního místa HAS [převzato ze 40].

Ani jedno cysteinové reziduum není vyžadováno pro aktivitu HAS [40]. V předchozích letech byla tato rezidua studována na molekulární úrovni. Později byla snaha nahradit je serinovými a alaninovými rezidui na různých místech a sledovat vliv na velikost produkované HA [41, 47].

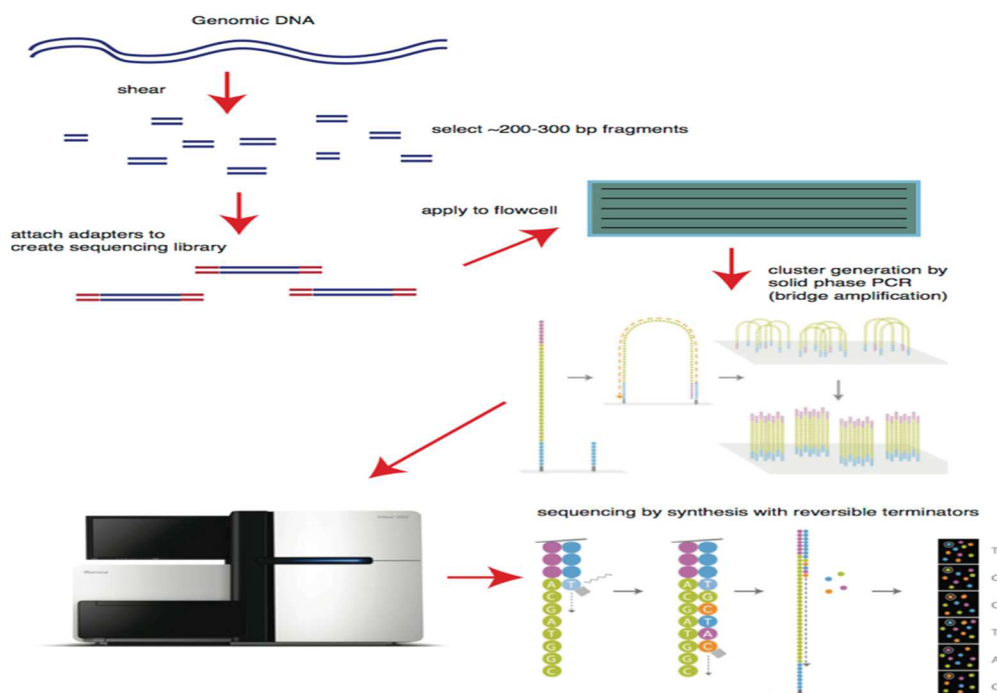
2.3 Sekvence celogenomové dna *S. zooepidemicus*

V posledních desetiletích je stále se rozvíjející celogenomová sekvenace (WGS – Whole-Genome Sequencing) jednou z nejslibnějších technik v mikrobiologii [48, 49]. Od doby osekvenování prvního bakteriálního genomu v roce 1995 prošla technika rozsáhlým vývojem [50, 51] Využití našla také v diagnostice, biotechnologiích, forenzní biologii, molekulární biologii a v dalších oblastech vědeckého zájmu. Zavedení metody významně a v pozitivním smyslu ovlivnilo biologický výzkum [52]. My této metody využíváme u dosud neosekvenovaného kmene *S. zooepidemicus* CO4A, který byl upraven fyzikální mutagenézí (UV světlo).

U streptokokálního genomu jsou využívány různé metody sekvenace, podle kritérií jednotlivých experimentů. Kompletní genom kmene *S. agalactiae* GD201008-001 byl osekvenován za použití Roche 454 GS FLX DNA sekvencer a systému Illumina Solexa GA IIx sekvencer v módu paired-end [53]. Platformy Roche 454 GS FLX sekvencer bylo například použito i v případě sekvenace dvou izolátů kmene *S. pneumoniae* SV35-T23 a SV36-T3 [54]. U *S. pyogenes* (sérologická skupina A -GAS) bylo WGS provedeno na přístroji HiSeq 2500 (Illumina) a sekvenováno v paired-end módu [55]. Platforma MiSeq (Illumina) byla použita například u WGS *S. pyogenes* kmene JMUB1235 [56]. Illumina je velmi přesná, ale nabízí kratší čtení (ready do 300 bp). Při skládání nového genomu je tak vhodné použít kombinaci dvou sekvenačních platform např. Illumina pro vysokou přesnost čtení a PacBio pro generování delších readů [57]. Zvýší se tím přesnost a správnost sestaveného genomu. Vždy je však nutné se řídit potřebami jednotlivých analýz, počtem readů, módem sekvenace, kapacitou sekvenátoru atd. Důležitým faktorem je i výběr softwaru, kterým budou data vyhodnocována.

Sekvenace na platformě Illumina (Solexa) má svůj původ v práci týmu doktora Turcattiho [58]. Knihovna DNA fragmentů (cca 300–1500 bp) může být vytvořena metodami PCR (generuje směs fragmentů o délce do několika stovek párů bází s připojenou sekvencí adaptoru), fyzikálními metodami (sonikátor), nebo enzymaticky (např. Nextera kit od Illuminy). V metodě se uplatňuje pro klastrování průtokové komory (angl. flow cell) můstková amplifikace nafragmentované DNA. Průtoková komora může nést až několik milionů klastrů umístěných na specifických místech. Díky indexování vzorků je navíc možné otestovat několik nezávislých DNA knihoven paralelně v průběhu jednoho runu [59]. Vše záleží na kapacitě sekvenátoru, velikosti knihoven aj. Po skončení generace klastrů jsou amplikony linearizovány a sekvenční primer je hybridizován k univerzální sekvenci adaptoru na regionu našeho zájmu. Detekce fluorescenčních záblesků nově nasedajících nukleotidů probíhá během čtení. U Solexa systému dochází k sekvenaci syntézou [59].

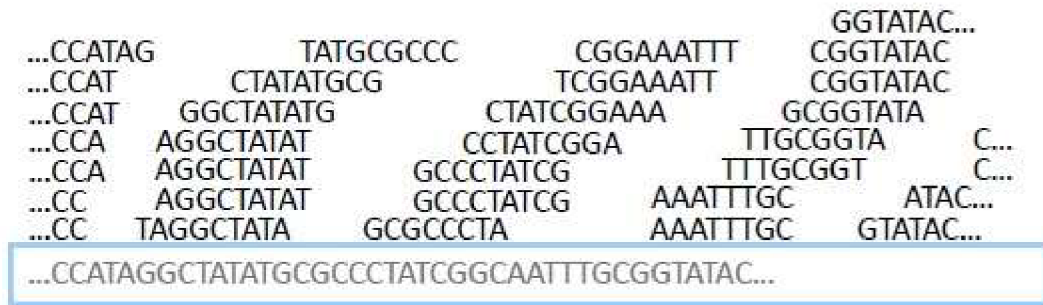
Technologie sekvenace syntézou používá čtyři fluorescenčně značených nukleotidů k sekvenaci desítek milionů klastrů současně na povrchu průtokové cely. V průběhu každého cyklu jsou jednotlivě značené deoxynukleotid trifosfáty (dNTPs) přidávány ke vznikajícímu řetězci DNA. Značený dNTP po inkorporaci do řetězce DNA vyzáří odpovídající fluorescenční signál, podle něhož je identifikována příslušná báze. Vznikající signál je v přímé úměře s měřeným signálem v průběhu cyklu [60]. Schéma procesu sekvenace syntézou s můstkovou amplifikací DNA je na obr. 8.



Obrázek 8: Schéma procesu WGS metodou sekvenace syntézou při použití fluorescenčně značených nukleotidů [61].

Výsledkem sekvenace jsou krátké úseky DNA ve formátu FASTQ (Illumina), u kterých známe sekvenci a kvalitu čtení. Těchto úseků se při celogenomové sekvenaci získá i několik set tisíc. Záleží na módu sekvenace (single-end vs. paired-end), délce čtení (75 bp, 150 bp, u sekvenací Illumina je to maximálně 300 bp), hustotě naklastrování průtokové komory aj. Sekvenace se provádí v takovém nastavení, aby bylo získáno dostatečné množství readů

a současně se úseky překrývaly větší částí své sekvence, aby mohly být sestaveny do větších úseků, tzv. contigů (obr. 9).



Obrázek 9: Sestavování delších sekvencí DNA (contigů) z krátkých úseků, které byly získány sekvenací. Modrým rámečkem je vyznačená sekvence contigu získaného z překrývajících se sekvencí [62].

Contig je delší sekvence vzniklá ze souboru DNA kratších sekvencí, které měly homologní konce a byly tak sloučeny do tohoto většího celku. Contigů z celogenomové sekvenace získáme řádově desítky. Jedná se o DNA sekvence nejčastěji ve formátu textového souboru FASTA. Sekvence contigu může být dlouhá několik stovek až několik desítek tisíc bází. Mají překrývající se sekvence, podle nichž jsou následně spojeny do jediné sekvence genomové DNA [63].

Sestavování genomové DNA z krátkých fragmentů pouze ručně by bylo velmi časově náročné. Proto existují porovnávací softwary a různé algoritmy, které se pro zpracování dat ze sekvenace dají použít. Pokud sekvenaci neprovádíme sami, ale provede ji externí laboratoř, jak tomu bylo i v našem případě, dostanou se k nám výsledky sekvenace ve formě contigů, které přeskupíme podle překrývajících se sekvencí (angl. *de-novo* assembling), nebo sestavíme genom podle osekvenovaného referenčního genomu mateřského kmene (angl. assemble to reference) [63].

Systémů ke zpracování sekvenačních dat existuje velké množství. Většinu jich lze použít pro různé sekvenační platformy a lze s nimi pracovat v různých operačních systémech (Linux, macOS, Windows) [63]. Vstupní formát sekvenačních dat je zásadní předpoklad k výběru správného nástroje pro NGS-pipeline. Tyto nástroje se velmi dynamicky rozvíjejí a vybrat kvalitní, nízkonákladový a uživatelsky příjemný software je obtížné. Některý je zdarma přístupný, ale ne vždy dlouhodobě podporovaný ze strany vývojářů. Komerční software (např. CLC Genomics Workbench [64]) je kontinuálně vyvíjen s ohledem na nejnovější trendy, zachovává uživatelsky příjemné prostředí a zajišťuje podporu pro koncové uživatele. Kvalitě však odpovídá i cena softwaru.

Analýzu dat z celogenomové sekvenace je možné provádět za použití různých softwarových nástrojů. Často se uvádí nástroje jako MAUVE [65, 66], DEP3 [67, 68], GLIMMER3 [53], CLC Genomic Workbench [64] a další. Software, jenž umožní *de-novo* sestavování genomu z krátkých fragmentů, je například systém ABySS (Assembly By Short Sequences) [69], který je navrhnout pro skládání velmi krátkých fragmentů (do 50 bp délky) do větších celků.

Dále pak systém Edena (Exact De Novo Assembler) [70] schopný zpracovat milióny krátkých fragmentů získaných platformou Illumina Genome Analyzer [71]. Systém Geneious

[72] je vhodný pro sestavování genomů z krátkých fragmentů DNA po sekvenaci a je jedním ze systémů, jež byly použity pro sestavování genomu SEZ [59]. Geneious je kompatibilní s mnohými platformami (Sanger sequencing, 454 pyrosequencing, Illumina, Oxford Nanopore aj.). Volně dostupné programy pro *de-novo* sestavení sekvence jsou např. A5 pipeline a Velvet [70, 73, 55]. Příkladem on-line nástroje, který můžeme využít k porovnávání a vyhledávání sekvencí, je BLAST (z angl. Basic Local Alignment Search Tool) [74]. Z porovnávacích systémů (algoritmů) vytvořených přímo pro platformu Illumina (Solexa) je to například ELAND (z angl. Efficient Large-Scale Alignment of Nucleotide Databases) [63].

Součástí přípravy sekvence genomu bývá také jeho anotování – přiřazování vlastností k specifickým DNA sekvencím. Ruční zadávání anotací je časově náročné a proto bylo pro anotaci genomu zavedeno rovněž velké množství systémů, které lze použít. Systém EagleView je schopen zobrazit a zahrnout velké množství informací o sekvenci a byl vyvinut na univerzitě v Bostonu [63]. Systém SAM (Sequance Assembly Manager) je nástroj k vizualizaci a upravování genomu, který byl vyvinut v kanadském *Michael Smith Genome Science* centru pro Linux [63].

2.4 Studium genové exprese *S. zooepidemicus* – DNA čipy

Vzrůstající frekvence používání sekvenačních metod, a tudíž i stále větší znalosti sekvencí DNA různých organismů a mikroorganismů, vyvolala také potřebu zavedení techniky, kterou by bylo možné sledovat expresi jednotlivých genů, skupin genů, operonů, nebo celých genomů. Dosavadními metodami bylo možné testovat jen aktivitu několika málo genů. Průlom v tomto oboru přinesla technologie DNA čipů. Umožnila totiž získat kvantitativní informaci o kompletním transkripčním profilu buňky a stala se tak cenným zdrojem poznání.

Technologie DNA čipů propojuje mnoho sekcí biologie, ale i matematiky a statistického zpracování dat. Potřeba takovéto metody vycházela právě z masivního nárůstu DNA sekvencí, ke kterým byly časem doplněny jejich reálné, nebo hypotetické funkce. Schopnost simultánního sledování exprese velkého množství genů je zcela klíčová a dnes je základem funkční genomiky. Technologie DNA čipů byla poprvé zmíněna v roce 1995 a následně nastal její rychlý rozvoj [75, 76, 77].

U prokaryotických systémů můžeme zmínit několik základních aplikací DNA čipů. Zásadní aplikací je studium celogenomové transkripční aktivity v buňce. Všechny buněčné funkce startují s expresí jednotlivých genů. Metodou DNA čipů lze sledovat transkripční aktivity genů, genových regionů za standardních nebo specifických růstových podmínek. Byly použity například u *E. coli* [78], *Caulobacter crescentus* [79, 80], nebo *Bacillus subtilis* [81]. Další aplikací může být studium regulonu. Regulon je skupina genů, jejichž transkripce je pod kontrolou regulátoru. Příkladem může být *NtrC* regulon u *E. coli* [82], sigma faktor u *B. subtilis* [83, 84] a u zástupce streptokoků *S. pneumoniae* [85, 86, 87]. Dále lze identifikovat operon a sledovat jeho strukturu, studovat neznámé regiony DNA, studovat interakci mezi DNA a proteiny [88]. V neposlední řadě mohou být DNA čipy použity v komparativní genomice a genotypingu např. *S. pneumoniae* [89, 90].

Pokud bychom měli zmínit specifické aplikace pro užití metody DNA čipů v mikrobiologii, byla by to především identifikace virulentních faktorů patogenů [90]. Další aplikací by bylo studium buněčné odpovědi hostitelského organismu na patogen, stanovení expresního profilu při aplikaci léčiv, inhibitorů a toxických látek a analýza mikrobiální evoluce a epidemiologie.

Metoda může být úspěšně použita i k diagnostickým účelům, metabolickému inženýrství a optimalizaci mikrobiálních procesů [90].

Specifické značení mRNA pro analýzu exprese genů metodou DNA čipů je u prokaryot výzvou [91]. U eukaryotické mRNA se využívá polyA řetězce k nabohacení molekul. U prokaryot bylo použití první specificky značené mRNA pro monitoring genové exprese na vysokodenzitním čipu publikováno v roce 1998 [92].

Technologie pracuje na principu hybridizace. Fluorescenčně značená RNA nebo DNA se ve formě roztoku hybridizuje s oligonukleotidy uchycenými na specifickém místě na povrchu čipu. Povrch může být z různého materiálu (porézní membrány, sklo, porcelán), na kterém jsou fragmenty DNA uchyceny. Jako čip mohou rovněž posloužit například mikrotitrační destičky. DNA fragmenty se získávají nejčastěji z cDNA, genomických nebo plazmidových knihoven, ale můžeme je připravit i PCR reakcí. Obecně jsou k dostání čipy nízkou a vysokodenzitní. Rozdílem u nich je počet sekvencí, které můžeme analyzovat, a kvalita signálu. Dnes jsou k dostání také čipy, u kterých lze sledovat paralelně až 1 milion sond [93, 94, 95].

DNA čipy jsou komerčně k dostání a je výhodné si je nechat připravit odborníky, pokud na pracovišti není příslušné přístrojové vybavení. Výrobci a prodejci také často nabízejí na svých webových stránkách službu automatického návrhu sekvence sond. Pro účely daného experimentu volíme také materiál, na který se sondy budou vázat a techniku, kterou budou sondy na povrch vázány [96]. Obecně se používají dvě metody nanášení a uchycení sond na čip. První metodou jsou sondy syntetizovány *in-situ*, tedy přímo na čipu [96]. Druhá metoda je založena na robotickém rozmístění a natiskování vybraných sond syntetizovaných mimo povrch čipu na zvolený nosič. Vše se provádí automatizovaně metodou *ink-jet*. Méně používané jsou vysokodenzitní kuličkové čipy, elektronické čipy a suspenze kuličkových čipů [96].

In-situ syntetizované čipy mají extrémně vysokou denzitu oligonukleotidových sond. Nejznámější výrobce takových čipů je Affymetrix. Oligonukleotidové sondy jsou syntetizovány přímo na povrch čipu, který je běžně velký 1,2 cm². Protože *in-situ* syntetizované sondy jsou typicky 20-25 nukleotidů dlouhé, navrhuje se několik různých sond pro jednu targetovou sekvenci, aby se zvýšila senzitivita, specifita a statistická významnost. Běžně se používá a navrhuje 11 sond pro úsek zájmu, který má délku 600 bp [97]. Set sond pro jeden úsek obsahuje jednu sondu, která přesně sedí na daném úseku a jednu sondu, která nese změnu v 1 bp uprostřed sondy. Sonda nesoucí změnu slouží jako negativní kontrola k identifikaci možných nespecifických hybridizací [96].

Affymetrix GeneChips mají více než 10⁶ sond na čipu, odvíjí se to od vzdáleností mezi sondami [97]. Sondy jsou syntetizovány polovodičovou fotochemickou metodou. Na povrch čipu jsou syntetizovány linkery, které jsou modifikované ochrannou skupinou citlivou na světlo [98]. Povrch čipu je tak chráněn před navázáním nukleotidů, dokud není čip vystaven působení UV záření. Pokud je povrch čipu vystaven působení UV záření, ochranná skupina bude degradována a řetězec může být prodloužen. K zacílení specifických nukleotidů na přesné místo se používají fotolitografické masky. Každá maska má svůj definovaný vzor – počet rámečků, které slouží jako filtr pro UV záření a ohraničují specifické části čipu [96]. Použití litografických masek zásadně limituje možnost změn výběru sond. Proto společnost Affymetrix produkuje čipy vhodné pouze pro určité mikroorganismy (*Saccharomyces cerevisiae*, *E. coli*, *B. subtilis*, *Pseudomonas aeruginosa*, *Salmonella typhimurium*) [99].

Další vysokodenzitní čipy jsou produkovány společností Roche NimbleGen (Madison, WI) a společností Agilent Technologies (Palo Alto, CA). Obě platformy používají delší oligonukleotidové sondy (60 až 100 bp). NimbleGen používá pro syntézu sond bezmaskovou technologii a Agilent Technologies využívá *ink-jet* technologii. Zatímco u Affymetrix GeneChips je limitováno značení vzorků na jednu barvu, u ostatních technologie umožňuje vícebarevnou hybridizaci. Roche NimbleGen technologie *in-situ* syntézy je podobná jako u Affymetrix GeneChips, jen fotolitografickou masku nahradili maskou virtuální. Tyto technologie umožňují flexibilitu v navrhování sond a je prakticky možné nechat nasyntetizovat sondy pro jakýkoliv mikroorganismus [99]. Každý NimbleGen čip může obsahovat více než 10^6 sond a mohou být připraveny v následujících formátech sond: $1 \times 2\ 100\ 000$, $3 \times 720\ 000$, $1 \times 385\ 000$, $4 \times 72\ 000$ a $12 \times 135\ 000$. Agilent Technologies čipy jsou syntetizovány na skleněné podložky a využívají metody *ink-jet* tisknutí, která eliminuje potřebu jakékoliv masky. Agilent čipy jsou dostupné ve formátech sond: $1 \times 244\ 000$, $2 \times 105\ 000$, $4 \times 44\ 000$, $8 \times 15\ 000$, $8 \times 60\ 000$, $4 \times 160\ 000$ a v mnohých dalších [96].

Velkou výhodou DNA čipů je možnost studia rozsáhlého množství genů v jedné analýze. Nízkodenzitní čipy nesou desítky až stovky specifických oligonukleotidových sond. Vysokodenzitní čipy mohou nést tisíce až miliony sond [100]. Nevýhodou je poměrně častý výskyt falešně pozitivních výsledků. Dochází k tomu hlavně, pokud má gen více homologních sekvencí v genomu, pokud má organismus vysoce variabilní genom apod. [100]. K chybám detekce vede i saturace čipu (při velkém množství stejných hybridizujících sekvencí), nebo omezení hybridizační kinetiky [100]. Čipy jsou založeny na principu nepřímého měření relativní koncentrace – signál by měl odpovídat množství hybridizované DNA. Měřený signál ovšem nebývá vždy lineární, a to může vést ke zkreslení výsledků. Cena komerčních DNA čipů bývá také poměrně vysoká [100].

2.5 Studium genové exprese *S. zooepidemicus* – qPCR

Metoda real-time polymerázové řetězové reakce (qPCR) je založena, stejně jako sekvenční techniky, na revoluční metodě PCR zavedené Kary Mullisem v osmdesátých letech minulého století. Metoda umožňovala amplifikovat specifické úseky DNA více než bilionkrát. Tento objev měl revoluční vliv na metody molekulární biologie, klonování a genové manipulace. Velkým přínosem byl i pro projekt sekvenace lidského genomu (Human Genome Project) [101, 102, 103, 104].

Od objevu PCR a zavedení do praxe bylo vyvinuto mnoho protokolů a přístrojů umožňujících efektivnější, rychlejší a především plně automatizovanou analýzu. qPCR může být použita pro řadu účelů, nejčastěji se využívá ke kvantifikaci nukleových kyselin a genotypingu. Můžeme ji ovšem využít i k detekci mutací, analýze chimerismů a dalším výzkumným aplikacím. Stejně jako i v této práci bývá metoda qPCR využívána ke studiu genové exprese ve spojení s analýzou na DNA čipech. Našla své místo i v klinické diagnostice. Zásadní rozdíl qPCR od původní PCR spočívá v tom, že vznik a amplifikace produktu jsou sledovány v reálném čase za použití fluorescenčních technik [105, 106, 107]. Hlavní výhodou je extrémně velký dynamický rozsah a signifikantně vyšší spolehlivost [108]. Fluorescenční signál je přímo úměrný vznikajícímu PCR produktu. Je generován použitím fluorescenčních barviv, které jsou specifické pro dvouvláknovou DNA (dsDNA specifická barviva: YO-PRO-1, YOYO-1, SYBR Green aj.) [109, 110], nebo použitím fluorescenčně značených sond sekvenčně specifických pro PCR produkt (např. TaqMan sondy) [109, 110].

SYBR Green je nejfrekventovaněji používaná fluorescenční barva v qPCR analýzách [111]. Jde o kyaninové interkalační barvivo strukturně podobné dalším dsDNA specifickým barvivům YOYO-1 a TOTO-1 [112, 113]. Nese dva pozitivní náboje, díky nimž má vysokou afinitu k negativně nabitě dsDNA. Vazebná účinnost je stokrát vyšší než u dříve používaného ethidium bromidu. Fluorescence navázaného SYBR Greenu je tisíckrát vyšší než nenavázaného, tudíž se dá velmi dobře využít k monitorování akumulace produktu qPCR. SYBR Green může být excitován pomocí modrého světla o vlnové délce 480 nm. Jeho emisní spektrum je srovnatelné s emisním spektrem fluoresceinu (FAM) s maximem při 520 nm [112]. Může docházet k tzv. „self-quenchingu“, neboli samozhášení, když je barvivo vázáno k dsDNA. Samozhášení není zdrojem chyb [111]. Často se také uvádí, že nespecifická detekce jakékoliv dsDNA je nevýhodou SYBR Greenu v porovnání se sekvenčně specifickými druhy detekce. Každý protokol použitý pro analýzu by měl být optimalizován, tudíž při použití optimalizovaného protokolu by měla být vysoká specifita PCR produktu (nevznikají nespecifické produkty, které by mohly zkreslovat výsledek), stejně jako efektivita reakce (vysoká účinnost reakce). Optimalizací se zamezí nespecifickým interakcím s náhodnou dsDNA a pokud by i tak docházelo ke zkreslení výsledků, bude se jednat o systematickou chybu, ať už bude nebo nebude docházet ke vzniku nespecifických produktů [111].

Obecně je metoda qPCR rozšířenou laboratorní technikou. Její hlavní výhodou je použití pro široké množství aplikací stejných jako u DNA čipů s výrazně vyšší efektivitou a při nižším prahu detekce. Nevýhodou je podstatně menší kapacita testovaný vzorků. Proto se velmi často obě metody (DNA čipy a qPCR) využívají souběžně [114, 115]. qPCR metoda byla použita při mnoha studiích k potvrzení výsledků získaných z hybridizačních reakcí na DNA čipech [114, 115, 116, 117].

3 CÍL PRÁCE

3.1 Charakterizace bakteriálního kmene *S. zooepidemicus* CO4A

Bakteriální kmen *S. zooepidemicus* CO4A využívaný společností Contipro a.s. k průmyslové výrobě HA byl získán náhodnou mutagenezí UV zářením. Prvním cílem mé práce bylo u mutanta charakterizovat sekvenci chromozomální DNA a porovnat ji se sbírkovým kmenem ATCC35246, z něhož vznikl (kap. 4.2, 4.5 a 4.6). Bioinformatickou analýzou byly nalezeny varianty v sekvenovaném vzorku chromozomální DNA (kap. 5.1 a 5.2). Varianty byly následně anotovány a filtrovány z hlediska jejich biologické relevantnosti.

3.2 Sledování exprese genů u vybraných kmenů *S. zooepidemicus* CO4A

Druhým cílem mé práce bylo sledovat úroveň exprese genů produkčního kmene CO4A a kmene sbírkového ATCC35246 v různých časových úsecích kultivace za daných podmínek (kap. 4.1). Sledování exprese bylo provedeno metodami DNA čipů a qPCR (kap. 4.7 a 4.8). Nutnou součástí byla optimalizace metod potřebných k dosažení relevantních výsledků (např. stanovení referenčního genu pro qPCR, kap. 4.8.2). Byl také testován vliv suplementace média *N*-acetylglukosaminem na expresi genů spojených s produkcí HA (kap. 5.4.3). Dalším krokem bylo určení genů se signifikantním nárůstem a poklesem exprese (kap. 5.3 a 5.4). V případě diferenciální exprese anotovaných genů bez variant byla provedena analýza metabolických drah a zjištěno zapojení genu do metabolismu HA. Identifikované geny zapojené přímo či nepřímo do metabolismu HA byly dále diskutovány.

4 MATERIÁL A METODY

4.1 Bakteriální kmeny a jejich kultivace

Sbírkový kmen *Streptococcus equi* subsp. *zooepidemicus* ATCC35246 [1] a interní kmen CO4A (připravený UV ozařováním kmene ATCC35246) byly skladovány v -80 °C ve formě glycerinových konzerv (20 % glycerolu).

4.1.1 Kultivace bakteriálního kmenepro sekvenaci celogenomové DNA

Z kryokonzervy kmene CO4A byla zaočkována Petriho miska s pevným živným médiem THY (Todd-Hewitt Broth; OXOID) s kvasničným autolyzátem [118]. Miska byla uložena v inkubátoru při 37 °C po dobu 24 h. Z takto na misce narostlé kultury, byla zaočkována 300ml Erlenmeyerova baňka se 100 ml tekutého živného média THY. Baňka byla inkubována na třepačce (Incubator Hood TH 15, Edmund Bühler GmbH; 150 rpm) 24 h při 37 °C a 150 rpm. Připravená kultura byla použita pro izolaci chromozomální DNA (kap. 4.2). Složení použitých médií je uvedeno v tabulce 1.

Tabulka 1: Složení živných médií použitých ke kultivaci pro izolaci celogenomové DNA [upraveno dle 118].

Navážka na 100 ml média	Tekuté živné médium	Pevné živné médium (agar)
THY (Todd Hewitt Broth)	3,7 g	3,7 g
Kvasničný autolyzát	0,3 g	0,3 g
Agar	-	1,6–2,0 g

4.1.2 Kultivace bakteriálních kmenů pro analýzu na DNA čípech

Příprava bakteriální kultury pro izolaci celkové RNA k hybridizaci na čip proběhla následovně. Z kryokonzervy *S. zooepidemicus* kmene CO4A a ATCC35246 byly zaočkovány 3 Petriho misky s THY agarem [118] pro každý ze dvou kmenů (tři biologické replikáty). Misky byly uloženy v inkubátoru 24 h při 37 °C. Z narostlé kultury byly zaočkovány 300ml Erlenmeyerovy baňky se 100 ml tekutého média THY. Baňky byly inkubovány na třepačce (Incubator Hood TH 15, Edmund Bühler GmbH) 24 h při 37 °C a 150 rpm. Připravené kultury byly použity pro inokulaci bioreaktorů (Multifors) od společnosti INFORS o objemu 500 ml s kontrolovaným monitorováním kultivačních parametrů.

Kultivace trvala 16 h. Fermentory byly očkované 20 ml kultury přes noc narostlé na $OD_{600} = 0,5$ (4% inokulum). Kultivace probíhaly v médiu THY a byly provedeny ve třech biologických replikátech. Vzorky byly odebírány a zpracovávány ve 4., 8., 12., a 16. hodině, abychom získali profil genové exprese v průběhu kultivace. Ve fermentorech byly kultivace provedeny při 37 °C, pH 7 (automatické dávkování 20% NaOH), vzdušnění 0,5 vvm a míchání 300 rpm (Rushtonova turbína). Odebrané vzorky byly centrifugovány při $6800 \times g$ (biomasa se usadila, ale buňky nebyly rozrušeny) a ze získané biomasy byla izolována celková RNA (kap. 4.3).

4.1.3 Kultivace bakteriálních kmenů pro analýzu qPCR

Kultivace s N-acetylglukosaminem (GlcNAc) a bez jeho suplementace

Kultivace probíhaly podobně jako kultivace pro analýzu na DNA čípech (kap. 4.1.2). Rozdílné bylo složení kultivačního média THY s kvasničným autolyzátem, které bylo obohaceno o GlcNAc (0,5 g/l) [6]. Médium kontrolní kultivace bylo bez GlcNAc. Vzorby byly odebrány a zpracovávány v 6., 9., 12., a 16. hodině kultivace. První vzorky byly odebrány později, protože analýza byla zaměřena na produkci HA, která probíhá v pozdějších fázích kultivace. Odebrané vzorky byly centrifugovány při $6800 \times g$ (biomasa se usadila, ale buňky nebyly rozrušeny) a ze získané biomasy byla izolována celková RNA (kap. 4.3).

Zpracování biomasy po kultivacích srážecí metodou

Biomasa byla zpracována ihned po skončení fermentací. Do kádinky bylo odváženo 250 g kultury v poměru 1:3 s vodou (250 g bakteriální suspenze + 750 ml demi H₂O). Před samotnou filtrací bylo přidáno do suspenze 15 g NaCl a poté byla suspenze přefiltrována přes navlhčenou filtrační desku HS800 (PALL) pod tlakem 2 bary. Filtrační zařízení bylo vyrobeno v laboratoři přímo pro tuto aplikaci. pH filtrátu bylo upraveno pomocí kyseliny octové na hodnotu 5. Srážení HA ve vzorku bylo provedeno za intenzivního míchání na hřídelové míchačce pomalým přiléváním 1,5 l regenerovaného (86 %) a 0,5 l absolutního izopropanolu (IPA). Vysrážený produkt byl sedimentován a opatrně dekantován. Následně dvakrát promyt 300 ml regenerovaného IPA a dvakrát 300 ml absolutního IPA. Vždy se sediment HA promíchal, homogenizoval (T 25 Digital Ultra-Turrax®, IKA) a znovu nechal usadit. Roztok byl po poslední sedimentaci dekantován.

Promytý produkt byl sušen na Petriho misce 16 h v sušárně při 60 °C. Množství HA bylo určeno gravimetricky vážením (navážka 100 mg HA) s přesností na dvě desetinná místa. Sušina produktu byla stanovena automatickými halogenovými sušícími váhami HR73 (Mettler Toledo) při teplotě sušení 120 °C. Sušení probíhalo až do zastavení úbytku hmotnosti sušeného vzorku. Výtěžek HA byl vyjádřen v jednotkách g/l po přepočtu na 100 % sušiny.

4.2 Izolace chromozomální DNA

Izolace celogenomové DNA k její sekvenaci byla provedena za použití komerčního kitu UltraClean® Microbial DNA Isolation Kit (MO-BIO Laboratories, Inc). Kit je určen pro izolaci genomové DNA z různých mikroorganismů. Izolace byla provedena podle protokolu výrobce. Kvalita izolované genomové DNA byla ověřena gelovou elektroforézou (TSA pufr). Na 1 % agarózový gel byly nanášeny 2 µl vzorku DNA smíšeného s nanášecím pufrem (Gel Loading Dye, Purple (6×), no SDS, NEB) a DNA barvičkou (MIDORI Green Direct, Elisabeth® Pharmacon). Koncentrace genomové DNA byla ověřena měřením na spektrofotometru DENOVIX DS-11 FX⁺. Čistota DNA byla vyhodnocena dle poměru absorbancí $A_{260/280}$ a $A_{260/230}$.

4.3 Izolace RNA a zjištění její kvality

K izolaci bakteriální RNA byl použit kit UltraClean™ Microbial RNA Isolation Kit od společnosti MOBIO Laboratories, Inc). Při izolaci bylo postupováno podle protokolu výrobce. Kvalita a kvantita izolované RNA pro analýzu na DNA čípech byla ověřena za použití Experion™ RNA StdSens Analysis Kit od společnosti BIO-RAD. Kit byl použit s přístrojovým vybavením Experion systému pro analýzu RNA. Systém zahrnoval automatickou

elektroforetickou stanicí, primingovou stanicí, vortexovou stanicí a software k analýze a normalizaci dat. Čistota a koncentrace izolované RNA byla také ověřena měřením na spektrofotometru DENOVIK DS-11 FX⁺. Čistota RNA byla odečtena z poměrů absorbancí $A_{260/280}$ a $A_{260/230}$. K analýze byly přijaty vzorky s hodnotami $A_{260/280}$ a $A_{260/230}$ většími než 1,9.

4.4 Přepis RNA do cDNA pomocí reverzní transkripce

K přepisu bakteriální RNA ze vzorků byl použit High-Capacity cDNA Reverse Transcription Kit od společnosti Applied Biosystems (RNA-dependentní DNA polymerázou je zde MultiScribeTM Reverse Transcriptase). Při přípravě směsi bylo postupováno podle protokolu výrobce kitu. Reakce proběhla v termocykleru od společnosti BioRad za podmínek uvedených v následující tabulce 2.

Tabulka 2: Nastavení programu termocykleru při přepisu RNA do cDNA

	Krok 1	Krok 2	Krok 3	Krok 4
Teplota	25 °C	37 °C	85 °C	4 °C
Čas	10 min	120 min	5 min	∞

Připravená cDNA byla uchovávána při -20 °C a byla použita jako templát pro následnou analýzu qPCR. V tomto kroku byla sjednocena koncentrace všech vzorků, aby bylo možné je porovnat. Koncentrace každého vzorku RNA (cDNA) byla unifikována na 150 ng/μl. Předpokladem byla absolutní efektivita přepisu RNA do cDNA.

4.5 Sekvenace celogenomové DNA

Připravený vzorek gDNA kmene CO4A byl odeslán k sekvenaci do externí laboratoře (EdgeBio, USA). Sekvenace byla provedena na přístroji MiSeq (Illumina) s délkou čtení paired-end 150 bp. Data byla získána ve formátu FASTQ. Sekvenační data byla dodána ve formě contigů a dále byla zpracována manuálně i pomocí softwaru. Analýza *de-novo* sestavení genomu byla provedena za pomoci několika druhů softwaru (tab. 3) a byly porovnány počty vzniklých contigů, délka nejdelšího contigu a konečná délka sekvence celého genomu. K anotaci a sestavování contigů do konečné podoby genomu produkčního kmene CO4A bylo použito několik různých softwarů a online dostupných databází (tab. 4).

Tabulka 3: Softwary použité k *de-novo* sestavení genomu.

Software	Typ	Operační systém	Výrobce/Vydavatel	Reference
Geneious 7.0.3	komerční	Windows, Linux, MacOS	Biomatters, Inc.	http://www.geneious.com
CLC Genomics Workbench	komerční	Windows, MacOS	QIAGEN	http://www.clcbio.com/products/clc-genomics-workbench/
Velvet	zdarma	Linux	EMBL-EBI	https://www.ebi.ac.uk/~zerbino/velvet/
Abyss	zdarma	Linux	BC Cancer Research Centre	https://www.bcgsc.ca/resources/software/abyss
Minia	zdarma	Linux	Salikhov K., Sacomoto G., a Kucherov G. [119]	http://minia.genouest.org

Tabulka 4: Softwary a databáze k anotaci sekvence a sestavení contigů do konečné sekvence genomu CO4A.

Software	Typ	Výrobce/Vydavatel	Reference
BLAST	Online, zdarma	National Center for Biotechnology Information	https://blast.ncbi.nlm.nih.gov/Blast.cgi
MAUVE	Windows, zdarma	The University of Technology Sydney	http://darlinglab.org/mauve/mauve.html
CLC Genomics Workbench	windows, MacOS, komerční	QIAGEN	http://www.clcbio.com/products/clc-genomics-workbench/
ApE	Windows, zdarma	The University of Utah	https://jorgensen.biology.utah.edu/wayned/apE/

4.6 Analýza genomu SEZ CO4A

K analýze genomu byl použit software CLC Genomics Workbench 20 [64, 120], jehož vydavatelem je společnost Qiagen. Původní sekvenační data SEZ CO4A byla srovnána s daty referenčního genomu ATCC35246 [1] použitím modulu Basic Frequency Variant Detection. Jako relevantní byly filtrovány varianty s frekvencí 90 % a vyšší. Pro zvýšení kvality mapování a eliminaci počtu falešně pozitivních variant byla původní sekvenační data podle kvality zkrácena o cca 20 bází na obou koncích 150 bp dlouhých sekvencích na základě Q-score (Phred Quality Score) [120]. Získaná data měla Q-score ≥ 30 . Následně byly upravené sekvence v CLC Genomics Workbench použity pro hledání variant v modulu Basic Frequency Variant Detection. Z analýzy byly vybrány varianty s frekvencí větší než 90 %. Takto získaná data byla porovnána s výsledky předchozí analýzy. Vyhledávány byly SNV (jednonukleotidový polymorfismus) a indely (inserce a delece). Následně byla provedena *in-silico* analýza vlivu změn nukleotidové sekvence na sekvenci aminokyselinovou.

4.7 Sledování exprese genů – DNA čipy

Sledování celogenomové exprese na DNA čipech bylo provedeno na *in-situ* syntetizovaných oligonukleotidových čipech SurePrint G3 Gene Expression Microarray (Agilent Technologies). Byly použity čipy o velikosti 4×180 000 sond. Délka sond byla 60 bp. Sondy byly navrženy softwarově přes webové stránky firmy Agilent Technologies. Průběh kultivace je popsán v kapitole 4.1.2. Průběh izolace RNA a ověření její kvality a kvantity je popsán v kapitole 4.3.

Poté, co byla potvrzena koncentrace a kvalita izolované RNA, byla přepsána do formy cDNA metodou reverzní transkripce (kap. 4.4). Pomocí kitu FairPlay III Microarray Labeling kit (Agilent Technologies) byly do jednovláknové cDNA inkorporovány amino allyl-dNTP, které jsou nezbytné k následnému ukotvení fluorescenční značky na řetězec. Vzorky cDNA byly naznačeny dle protokolu výrobce. Značená cDNA byla použita pro hybridizaci na DNA čipech s navrženými sondami. Hybridizace na DNA čipy byla provedena pomocí Gene Expression Hybridization kitu (Agilent Technologies) podle protokolu výrobce.

4.7.1 Skenování čipu a zpracování dat

Skenování čipu bylo provedeno na microarray detektoru Molecular Device Axon GenePix 4000B (Molecular Devices) a ke zpracování obrazu byl použit software GenePix Pro 7.1 (Molecular Devices). Hodnoty vygenerované softwarem ihned po konverzi z obrazového formátu byly dále zpracovávány a vyhodnocovány v programu Microsoft Excel. Jako referenční gen byla použita gyráza A (*gyrA*) a to na základě získaných výsledků při stanovení referenčního genu *S. zooepidemicus* CO4A a ATCC35246 (kap. 4.8.2 a 5.4.1).

4.7.2 Výpočet poměru celkové exprese sledovaného genu v genomu CO4A/ATCC

Celková exprese sledovaného genu (CEG) byla vypočtena jako součet všech hodnot exprese daného genu (E_g) v průběhu celé kultivace (resp. ve 4., 8., 12. a 16. odběrové hodině). Výpočet byl proveden dle rovnice 1:

$$CEG = E_g(4h) + E_g(8h) + E_g(12h) + E_g(16h)$$

Rovnice 1: Výpočet celkové exprese genu.

4.7.3 Výpočet poměru celkové exprese genu v genomu CO4A/ATCC normalizováno vůči expresi gyrázy A (gyrA)

Celková exprese každého sledovaného genu (CEG_g) v genomu byla vztažena k celkové expresi referenčního genu (CEG_{gyrA}) *gyrA* a tím byla získána hodnota celkové exprese genu vůči *gyrA* ($CEG_{g/gyrA}$) pro jednotlivé kmeny. Výpočet uveden jako rovnice 2:

$$CEG_{g/gyrA} = \frac{CEG_g}{CEG_{gyrA}}$$

Rovnice 2: Výpočet použitý pro normalizaci dat vůči referenčnímu genu gyráze A.

Po vypočtení celkové exprese genů vůči *gyrA* pro jednotlivé kmeny ($CEG_{g/gyrA}(ATCC)$ a $CEG_{g/gyrA}(CO4A)$) byl odečten poměr p těchto hodnot (rovnice 3). Hodnota poměru byla ukazatelem toho, kolikrát byla síla exprese u produkčního kmene vyšší než u kmene referenčního.

$$p = \frac{CEG_{g/gyrA}(CO4A)}{CEG_{g/gyrA}(ATCC)}$$

Rovnice 3: Výpočet poměru celkových expresí genů ve sledovaných kmenech vztažených k referenčnímu genu gyrA.

4.8 Sledování exprese genů – qPCR

4.8.1 Provedení analýzy qPCR

Návrh primerů

Primery pro qPCR analýzu byly navrženy pomocí Primer-BLAST [74] podle těchto pravidel:

- Délka primerů 20 bází (± 1 báze).
- Teplota tání (T_m) primerů mezi 59-61 °C. Preference byla dána na teplotu 60 °C.
- Délka vznikajícího amplikonu v rozmezí 50-150 bází.
- Procentuální obsah GC párův rozmezí 30-80 %, ideálně kolem 50 %.

Dále byl sledován obsah G a C bází v posledních 5 bázích na 3' konci primerů. Jejich obsah ovlivňuje teplotu tání primerů a může zkreslovat výslednou sílu signálu (čím vyšší obsah G a C bází, tím vyšší bude teplota tání). Požadovanými vlastnostmi u primerů byly také žádné vzájemné překryvy, vnitřní struktury, dlouhé homopolymery a žádné vlásenky a dimery [121].

Normalizace koncentrace templátu

Templátem pro analýzu qPCR byly vzorky RNA přešpané do cDNA. Koncentrace vzorků byly sjednoceny na 150 ng/ μ l v objemu 20 μ l. Tato koncentrace je pro analýzu qPCR příliš

vysoká. Pokud by byl použit templát o tak vysoké koncentraci, nebyla by dobře znatelná počáteční fáze amplifikační křivky, podle které je nastavována baseline v softwaru. Proto byl vzorek doplněn do objemu 1 ml na konečnou koncentraci 3 ng/μl.

Negativní kontrola qPCR

Jako negativní kontrola pro analýzy qPCR byla zvolena cDNA připravená z celkové RNA bakterie *Escherichia coli DH5α* a k ní byly vybrány primery pro gen kódující hyaluronan syntázu (*hasA*). Tento gen bakterie za normálních okolností neexprimuje, tudíž její hodnoty by měly být nedetekovatelné. Negativní kontrola byla připravena pro každou analýzu.

SYBR Green fluorescenční značení

K provedení analýzy bylo použito sekvenčně nespecifické fluorescenční barvivo SYBR Green v podobě komerčně dodávaného kitu Power SYBR[®]Green PCR Master Mix (Applied Biosystems). Program nastavený na přístroji je uveden v tab. 5. Rovněž při přípravě PCR směsi bylo postupováno podle protokolu výrobce.

Tabulka 5: Nastavení přístroje pro provedení analýzy qPCR.

	AmpliTaQ Gold[®] Enzymová aktivace	Denaturace	Nasedání primerů Prodlužování řetězce	Závěrečná fáze	Počet cyklů
Teplota	95 °C	95 °C	60 °C	4 °C	40
Čas	10 min	15 s	1 min	∞	

4.8.2 Stanovení referenčního genu pro qPCR

Na základě informací získaných z provedené analýzy na DNA čípech bylo vybráno 10 genů (tab. 6), které mohly být potenciálně referenčními geny pro qPCR.

Tabulka 6: Potenciální referenční geny a jejich funkce.

Gen	Anotace genu
Sesec_00096	NUDIX hydrolase
Sesec_00097	ispE, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
Sesec_00549	lacB, galactose-6-phosphate isomerase subunit LacB
Sesec_00725	ileS, isoleucyl-tRNA synthetase
Sesec_00767	transcriptional repressor
Sesec_00768	hypothetical protein
Sesec_02327	mscS, mechanosensitive ion channel protein
Sesec_02508	mviM, oxidoreductase
Sesec_01348	gyrA, DNA gyrase subunit A
Sesec_01041	gyrB, DNA gyrase subunit B

Geny byly vybírány podle míry a stability exprese. Sledována byla hodnota C_t^1 (z anglického „cycle of threshold“) u každého genu, průběh exprese genu v čase (4 časové body odběrů vzorků) a nakonec byla provedena analýza softwarovým doplňkem programu Microsoft Excel – GeNorm ke stanovení stability exprese genů [122]. Software GeNorm vyhodnocuje hodnotu stability exprese genu M [122]. Čím je tato hodnota nižší, tím je exprese genu stabilnější. Limitní kritická hodnota M je většinou nastavena na 1,5. Tudiž $M < 1,5$ lze považovat za dostatečně stabilní expresi pro referenční gen [122]. Z genů byl vybrán jeden jako referenční (*gyrA*; viz 5.4.1) a byl použit pro všechny následné analýzy.

4.8.3 Analýza C_t hodnot genů *has operonu*

Hodnoty C_t genů *has operonu* každého biologického replikátu obou sledovaných kmenů (ATCC35246, CO4A) byly normalizovány k referenčnímu genu *gyrA* výpočtem poměru p dle rovnice 4.

$$p = \frac{C_t gyrA}{C_t genu}$$

Rovnice 4: Výpočet poměru C_t referenčního genu *gyrA* k C_t genu.

Po normalizaci k referenčnímu genu byla vypočtena průměrná hodnota C_t genu ze všech tří biologických replikátů. Průměrné hodnoty C_t pro geny obou kmenů byly vyneseny do grafů. Hodnoty chybových úseček byly získány výpočtem směrodatné odchylky z dat biologických replikátů.

4.8.4 Analýza C_t hodnot vybraných skupin genů metodou $2^{-\Delta\Delta C_t}$

Vybrané skupiny genů zapojené v biosyntéze HA byly testovány u kultivací se suplementací *N*-acetylglukosaminem a bez jeho suplementace. Byly porovnávány hodnoty C_t genů z kontrolní kultivace u kmene CO4A a referenčního kmene ATCC35246 a hodnoty C_t genů z kultivace se suplementací média GlcNAc u týchž kmenů. Dále byly hodnoty normalizovány vůči referenčnímu genu *gyrA*. Genová exprese byla vyhodnocena pomocí metody $2^{-\Delta\Delta C_t}$ [123]. Metodou bylo možné vyjádřit relativní míru exprese (fold change) vybraných genů zapojených v metabolismu HA při kultivaci se suplementací GlcNAc. Současně mohla být data normalizována vůči *gyrA* jako referenčnímu genu a vůči kontrolní kultivaci bez suplementace GlcNAc. Díky biologickým replikátům mohly být stanoveny směrodatné odchylky pro takto vyjádřené relativní míry exprese. Výpočet $\Delta\Delta C_t$ a fold change $2^{-\Delta\Delta C_t}$ je uveden rovnicemi 5 a 6.

$$(C_{t,gen} - C_{t,gyrA})_{GlcNAc} - (C_{t,gen} - C_{t,gyrA})_{kontrola} = \Delta\Delta C_t$$

Rovnice 5: Výpočet $\Delta\Delta C_t$ sledovaného genu při normalizaci na referenční gen *gyrA* a kontrolní kultivaci bez suplementace GlcNAc.

$$fold\ change = 2^{-\Delta\Delta C_t}$$

Rovnice 6: Výpočet relativní míry exprese (fold change) sledovaného genu.

¹ Hodnota C_t je číslem cyklu, ve kterém se signifikantně projeví nárůst fluorescenčního signálu. Fluorescenční signál v průběhu qPCR analýzy sílí až do svého maxima, tedy do stavu nasycení reakce. Vývoj křivky na jejím počátku je exponenciální, a právě v této fázi je hodnota odečítána. Hodnota C_t tedy určuje míru exprese genu v daném vzorku. Pokud je na počátku reakce vyšší množství templátu pro reakci, jeho fluorescenční signál exponenciálně narůstá a o to dříve se objeví signál silnější, než je pozadí. Tedy vyšší míra exprese genu se vyznačuje nižší hodnotou C_t [124].

5 VÝSLEDKY A DISKUZE

V první části této kapitoly jsou diskutovány výsledky celogenomové sekvenace mutantního kmene *S. zooepidemicus* CO4A, který vykazuje zvýšenou produkci HA oproti kmeni mateřskému ATCC35246. Analýzou genomu CO4A a srovnáním s referenčním kmenem byly identifikovány relevantní mutace typu SNV a indelů. Každá z mutací v kódujících sekvencích byla *in-silico* vyhodnocena z hlediska vlivu na aminokyselinovou sekvenci příslušných proteinů. V další části byla s využitím metodiky DNA-čipů vyhodnocována celková míra exprese genů v genomu CO4A a to v porovnání s genomem kmene ATCC35246. V závěrečné části je zkoumán vliv suplementace kultivačního média *N*-acetylglukosaminem na produkci HA a expresi vybraných skupin genů u obou studovaných kmenů.

5.1 Sekvenace chromozomální DNA *Streptococcus equi* subsp. *zooepidemicus* CO4A

Z kultury kmene CO4A byla izolována chromozomální DNA (kap. 4.2). Minimální doporučené množství DNA pro celogenomovou sekvenaci malých genomů (≤ 5 Mbp) je 250 ng. Množství našeho vzorku bylo stanoveno na 1 155 ng (23,11 ng/ μ l, celkový objem 50 μ l). Výsledné hodnoty relevantních poměrů absorpací při 230, 260 a 280 nm byly následující: $A_{260/280} = 1,85$ a $A_{260/230} = 2,09$.

DNA byla úspěšně osekvenována a data byla analyzována za pomoci softwaru (kap. 4.5). Sekvenace byla provedena v módu paired-end o délce čtení 150 bp. Délka sekvencí se pohybovala v rozmezí 35-151 bp (90 % sekvencí bylo delších než 140 bp). Množství sekvencí v páru bylo 7 632 658, tedy $2 \times 3 816 329$. Více než 90 % přečtených bází vykazovalo Q-score >30 , tedy správnost čtení vyšší než 99,9 %. Konečné výsledky byly dodány ve formě contigů o velikosti řádově od stovek párů bází po statisíce párů bází. Počet contigů pro kmen CO4A se lišil s použitým softwarem pro *de-novo* sestavení sekvence (tab. 7). Contigy byly zapisovány do formátu FASTQ, který je kompatibilní s vyhledávacím a porovnávacím softwarem, který může být použit k sestavení celého genomu (kap. 4.5).

Tabulka 7: Rozdíly v počtu a délce contigů s použitým softwarem (kmen CO4A)

Software	Počet contigů	Nejdelší contig [bp]	Délka genomu [bp]
CLC Genomics Workbench	77	118733	2103015
Geneious	512	83112	2111453
Velvet	7382	68315	2205676
ABYSS	1434	68315	2149762
Minia	275	68315	2062678

Contigy měly překrývající se komplementární konce. Byla provedena analýza ručním vyhledáváním komplementárních konců a následné spojování contigů v programu ApE (A plasmid Editor, ApE 2020) [125] a doplněna softwarovou analýzou za použití volně dostupné webové aplikace BLAST (Basic Local Alignment Search Tool) [74]. Pomocí BLAST lze porovnávat genové sekvence s on-line databází a nacházet tak homologní sekvence v různých bakteriálních kmenech. Contigy byly srovnány s nejpříbuznějším osekvenovaným kmenem a byla zjištěna jejich pravděpodobná pozice v genomu v porovnání s referenčním genomem. K bioinformatické analýze byla použita také volně dostupná verze programu Mauve [66], který sestaví contigy vůči referenčnímu genomu a graficky je zobrazí. Výsledkem byl

kompletní genom kmene CO4A o velikosti 2 167 251 párů bází. Bylo identifikováno 1 999 kódujících oblastí a obsah bází GC byl 41,7 %.

Referencí byl genom kmene ATCC35246, který je kmeni CO4A nejbližším příbuzným genomem (vykazuje 99,5% identitu). Jeho sekvence je také dostupná v databázi NCBI [74]. Sekvenace byla provedena v roce 2011 týmem čínských vědců za použití platformy Roche 454 genome sequencer FLX systému a bylo získáno celkem 649 contigů [1]. Kompletní genom je složen z 2 167 264 párů bází s četností nukleotidů GC 41,7 %. Obsahuje 1 984 sekvencí kódujících protein.

5.2 Analýza genomu *S. zooepidemicus* CO4A

V dalším kroku byla provedena analýza genomu kmene *S. zooepidemicus* CO4A (kap. 4.6). Detekováno bylo 1453 variant, nicméně pro analýzu byly směrodatné především varianty s frekvencí 90 % a vyšší. Varianty s nižší frekvencí mohly vzniknout např. při přípravě sekvenační knihovny, v průběhu sekvenace, nebo při kombinaci obou dějů [126]. Po odfiltrování variant s frekvencí nižší než 90 % jsme identifikovali 80 relevantních variant. Z toho se jedná v 59 případech o SNV (tab. 8), v 7 případech o delecii (tab. 9) a ve 14 případech o inserci (tab. 11). Po provedení rozšířené analýzy (kap. 4.6) bylo získáno 83 variant, z toho bylo 59 případů SNV dále 15 případů inserce a 9 případů delecce. Navíc tedy byla identifikována pouze jedna inserce a dvě delecce (tab. 13). V ostatních případech se výsledky obou analýz shodovaly.

Identifikované varianty v jednotlivých nukleotidech u SNV, mohly způsobit změnu v aminokyselinovém řetězci. Případné změny AMK sekvence byly zaneseny do tab. 8 a dále diskutovány. Při bližším pohledu na referenční sekvenci genomu ATCC35246 zjistíme, že byl letos (2020) nově anotován [127], ale zůstala původní sekvenační data získaná na platformě Roche 454 genome sequencer FLX systému v roce 2011 [1]. Pyrosekvenace na platformě Roche byla v průběhu let optimalizována, nicméně má své limity použití (správnost čtení, délka readů aj.) [128] a společnost Roche v roce 2016 ukončila své aktivity v tomto odvětví. Je tedy nutné se ptát, zda byly všechny detekované SNV mutace a indely zapříčiněny působením UV záření. Genom CO4A byl osekvenován na platformě Illumina (kap. 5.1) a později byla sekvenace ověřena ještě na platformě PacBio. V současnosti je to zřejmě nejspolehlivější kombinace sekvenačních technik k sekvenaci bakteriálního genomu [129]. Máme důvod se domnívat, že se v původním genomu ATCC35246 mohly vyskytnout změny způsobené chybnou sekvenací, popř. nedostatečnou délkou pokrytí genomu z důvodu menšího množství sekvenačních dat. Tím mohlo dojít i k chybné interpretaci některých změn AMK sekvencí. Zvláště pak u genů, které se zdály být u kmene ATCC35246 podle sekvence zcela nefunkční (například nesmyslná přítomnost velkého množství stop kodónů v sekvenci) a domnělou náhodnou mutací byly v genomu kmene CO4A opraveny. Možnost chybné sekvenace referenčního genomu ATCC35246 nelze vyloučit a proto je na ni dále v textu upozorňováno a je také u genů s variantami diskutována.

Tabulka 8: V tabulce jsou uvedeny pozice 59 identifikovaných SNV a v jaké oblasti genu kmene ATCC35246 se nachází. Dále je uveden význam genu, jeho označení (dle databáze NCBI [74] a KEGG [130] – pokud jsou dohledatelné) a jakým způsobem ovlivní změna nukleotidu aminokyselinovou sekvenci proteinu. Červeně jsou označeny SNV vyvolávající změny AK sekvence, které mohly ovlivnit růst buňky, transkripci, translaci, nebo produkci HA. CDS – protein kódující sekvence (z anglického CoDing Sequence), NCS – nekódující sekvence (Non Coding Sequence)

Region ATCC 35246	Typ varianty	Báze v ATCC 35246	Báze v CO4A	AMK v ATCC 35246	AMK v CO4A	Délka proteinu	Shoda	Frekvence [%]	Gen	Oblast	Funkce genu
186212	SNV	T	C	Ala (GCT)	Ala (GCC)	477	477	100,0	Sesec_00200	CDS	polA, DNA polymerase I [EC:2.7.7.7], [K02319]
221149	SNV	T	C	Ser (AGT)	Gly (GGT)	553	553	100,0	Sesec_00247	CDS	hypothetical protein
277654	SNV	A	G	Lys (AAG)	Arg (AGG)	359	359	100,0	Sesec_00319	CDS	rpsL, 30S ribosomal protein S12 [K02950]
300278	SNV	C	T	Gly (GGA)	Glu (GAA)	537	537	100,0	Sesec_00348	CDS	rpsN, 30S ribosomal protein S14 [K02954]
306976	SNV	T	C	Cys (TGC)	Arg (CGC)	525	525	100,0	Sesec_00359	CDS	two-component system, OmpR family, sensor histidine kinase SaeS [EC:2.7.13.3], [K10681]
370386	SNV	G	A	Arg (CGT)	His (CAT)	462	462	100,0	Sesec_00434	CDS	alr, alanine racemase [EC:5.1.1.1], [K01775]
393853	SNV	C	T	Thr (ACG)	Met (ATG)	484	484	100,0	Sesec_00454	CDS	codY, GTP-sensing pleiotropic transcriptional regulator CodY (repressor) [K03706]
434492	SNV	G	A	Ala (GCG)	Thr (ACG)	450	450	100,0	Sesec_00506	CDS	polyisoprenyl-teichoic acid-peptidoglycan teichoic acid transferase [EC:2.7.8.-], [K01005]
649924	SNV	C	T	Lys (AAG)	Lys (AAA)	456	456	100,0	Sesec_00785	CDS	tRNA1(Val) (adenine(37)-N6)-methyltransferase
696395	SNV	T	C	Leu (CTT)	Pro (CCT)	485	485	100,0	Sesec_00836	CDS	uidA, beta-glucuronidase [EC:3.2.1.31], [K01195]
824533	SNV	G	A	Thr (ACG)	Thr (ACA)	391	391	100,0	Sesec_00996	CDS	YvcK family protein
932808	SNV	C	T	Pro (CCG)	Ser (TCG)	489	489	100,0	Sesec_01110	CDS	aminotransferase
1016814	SNV	T	C	Ile (ATC)	Thr (ACC)	314	314	100,0	Sesec_01207	CDS	pstB, phosphate ABC transporter ATP-binding protein [EC:7.3.2.1], [K02036]
1129125	SNV	C	T	Arg (CGA)	Gln (CAA)	416	416	100,0	Sesec_01341	CDS	murM, UDP-N-acetylmuramoylpentapeptide-lysine N(6)-alanyltransferase [EC:2.3.2.10], [K05363]
1203154	SNV	C	A	Arg (AGG)	Met (ATG)	521	521	100,0	Sesec_01429	CDS	ABC transporter ATP-binding cassette, subfamily B, tetracycline resistant protein [K18216]
1333515	SNV	G	A	Ser (TCA)	Leu (TTA)	494	494	100,0	Sesec_01587	CDS	ABC transporter utative ABC transport system ATP-binding protein [K05833]
1351663	SNV	A	T	Ala (GCT)	Ala (GCA)	466	466	100,0	Sesec_01607	CDS	SDR family oxidoreductase [K07124]
1358472	SNV	C	A	Arg (CGG)	Leu (CTG)	445	445	100,0	Sesec_01616	CDS	glgP, glycogen/starch/alpha-glucan family phosphorylase [EC:2.4.1.1], [K00688]

1516920	SNV	G	A	Pro (CCG)	Leu (CTG)	462	462	100,0	Sesec_01822	CDS	ccpA, catabolite control protein A, LacI family transcriptional regulator [K02529]
1519276	SNV	T	C	Asn (AAC)	Ser (AGC)	272	272	100,0	Sesec_01825	CDS	ABC transporter ATP-binding/permease protein
1861115	SNV	G	A	Asp (GAC)	Asp (GAT)	550	550	100,0	Sesec_02260	CDS	OppA, oligopeptide ABC transporter periplasmic oligopeptide-binding protein, oligopeptide ABC transporter substrate-binding protein [K15580]
1996819	SNV	C	T	Arg (CGT)	His (CAT)	539	539	100,0	Sesec_02414	CDS	relA, GTP pyrophosphokinase, GTP diphosphokinase / guanosine-3',5'-bis(diphosphate) 3'-diphosphatase [EC:2.7.6.5,EC:3.1.7.2], [K01139]
2107642	SNV	G	A	Leu (CTC)	Leu (CTT)	646	646	100,0	Sesec_02545	CDS	DUF4065 conserved hypothetical protein
772980	SNV	G	A	Lys (AAG)	Lys (AAA)	633	634	99,8	Sesec_00926	CDS	PcfB family protein (DUF3801, LPD29, Phage_integrase – hypotetical protein)
66680	SNV	A	G	His (CAT)	Arg (CGT)	609	610	99,8	Sesec_00065	CDS	rplV, 50S ribosomal protein L22 [K02890]
2057183	SNV	G	A	Leu (TTG)	Leu (TTA)	579	580	99,8	Sesec_02478	CDS	dkp, deoxynucleoside kinase protein
299568	SNV	G	A	Arg (AGA)	Lys (AAA)	550	551	99,8	Sesec_00347	CDS	DUF218, YdcF family protein
618107	SNV	G	A	Glu (GAA)	Lys (AAA)	547	548	99,8	Sesec_00745	CDS	xseA, exodeoxyribonuclease VII large subunit [EC:3.1.11.6], [K03601]
353141	SNV	A	G	Thr (ACC)	Ala (GCC)	536	537	99,8	Sesec_00418	CDS	conserved hypothetical protein
408995	SNV	A	G	Ile (ATT)	Val (GTT)	518	519	99,8	Sesec_00473	CDS	dnaK, chaperone protein DnaK (heat shock protein 70) [K04043]
1751733	SNV	T	C	Ile (ATA)	Met (ATG)	506	507	99,8	Sesec_02125	CDS	mraY, phospho-N-acetylmuramoyl-pentapeptide-transferase [EC:2.7.8.13], [K01000]
541853	SNV	A	C	Gln (CAA)	His (CAC)	503	504	99,8	Sesec_00647	CDS	FtsK/SpoIIIE family protein putative secretion system component EssC/YukA, DNA segregation ATPase FtsK/SpoIIIE, S-DNA-T family [K03466]
194169	SNV	C	T	Asp (GAC)	Asp (GAT)	475	476	99,8	Sesec_00206	CDS	YSIRK-type signal peptide-containing protein
1739978	SNV	G	A	Ala (GCC)	Ala (GCT)	450	451	99,8	Sesec_02108	CDS	nadE, ammonia-dependent NAD(+) synthetase [EC:6.3.1.5], [K01916]
10100	SNV	C	T	Cys (TGC)	Cys (TGT)	446	447	99,8	Sesec_RS00040	CDS	oligosaccharide flippase family protein
210788	SNV	C	T	Asn (AAC)	Asn (AAT)	427	428	99,8	Sesec_00234	CDS	glmU, bifunctional UDP-N-acetylglucosamine diphosphorylase/glucosamine-1-phosphate N-acetyltransferase GlmU [EC:2.3.1.157;EC:2.7.7.23], [K11528]
39978	SNV	T	C	Cys (TGT)	Cys (TGC)	417	418	99,8	Sesec_00461	CDS	EamA family transporter, membrane protein
1397893	SNV	A	G	Cys (TGT)	Arg (CGT)	409	410	99,8	Sesec_01663	CDS	RsnA, ribosomal small subunit pseudouridine synthase A, 16S rRNA pseudouridine516 synthase [EC:5.4.99.19], [K06183]
69527	SNV	C	T	Pro (CCA)	Ser (TCA)	621	623	99,7	Sesec_00072	CDS	rplE, 50S ribosomal protein L5, large subunit ribosomal protein L5 [K02931]
2071716	SNV	G	A	Ala (GCC)	Val (GTC)	570	572	99,7	Sesec_02503	CDS	hypothetical protein (DUF2371)
1477303	SNV	G	A	Asp (GAC)	Asp (GAT)	548	550	99,6	Sesec_01775	CDS	hypothetical protein (DUF5126, UPF0020)
134534	SNV	C	T	Ala (GCC)	Val (GTC)	447	449	99,6	Sesec_00138	CDS	PepO, endopeptidase O, putative endopeptidase [EC:3.4.24.-], [K07386]

494127	SNV	G	A	Gly (GGG)	Glu (GAG)	440	442	99,5	Sesec_00592	CDS	Streptolysin S biosynthesis protein SagD, YcaO-like family protein
934041	SNV	C	T	Ala (GCG)	Val (GTG)	438	440	99,5	Sesec_01112	CDS	amino acid ABC transporter, amino acid-binding/permease protein, polar amino acid transport system substrate-binding protein [K02030]
2135214	SNV	C	T	Val (GTC)	Ile (ATC)	631	634	99,5	Sesec_02576	CDS	gidA, glucose inhibited division protein A, tRNA uridine 5-carboxymethylaminomethyl modification enzyme [K03495]
121433	SNV	G	A	Pro (CCC)	Ser (TCC)	405	407	99,5	Sesec_01436	CDS	CCA tRNA nucleotidyltransferase [EC:2.7.7.72], [K07558]
373867	SNV	G	A	Gly (GGT)	Ser (AGT)	403	405	99,5	Sesec_00437	CDS	NEAT domain-containing protein, LRR 8, putative Fe ³⁺ -siderophore transport protein, leucine-rich repeat protein
1818371	SNV	G	A	Pro (CCT)	Ser (TCT)	403	405	99,5	Sesec_02214	CDS	SWF/SNF family helicase, Helicase conserved C-terminal domain
363697	SNV	G	A	Ala (GCA)	Thr (ACA)	400	402	99,5	Sesec_00429	CDS	LRR 8, endo-beta-N-acetylglucosaminidase F2 precursor,
1990184	SNV	G	A	His (CAC)	His (CAT)	553	556	99,5	Sesec_02406	CDS	dexB, glucan 1,6-alpha-glucosidase [EC:3.2.1.70], [K01215]
236344	SNV	G	A	Ala (GCG)	Thr (ACG)	334	336	99,4	Sesec_00266	CDS	cynT, carbonic anhydrase [EC:4.2.1.1], [K01673]
2097462	SNV	G	A	Val (GTA)	Ile (ATA)	499	502	99,4	Sesec_02532	CDS	argS, arginyl-tRNA synthetase [EC:6.1.1.19], [K01887]
1703037	SNV	A	T	Lys (AAA)	Asn (AAT)	161	162	99,4	Sesec_02068	CDS	ISL3-like element ISSeq1 family transposase [K07485]
558939	SNV	G	A	Ala (GCT)	Thr (ACT)	609	613	99,3	Sesec_00669	CDS	CRISPR-associated protein Csd2 [K19118]
1282926	SNV	C	T	NC	NCS	551	555	99,3		NCS	Mezigenová sekvence (NCS – non coding sequence)
225027	SNV	C	T	Ala (GCT)	Thr (ACT)	401	405	99,0	Sesec_00255	CDS	galU; UTP--glucose-1-phosphate uridylyltransferase HasC.2 [KO:K00963] [EC:2.7.7.9]
1185071	SNV	T	A	Asn (AAT)	Lys (AAA)	83	84	98,8	Sesec_01410	CDS	ISL3-like element ISSeq1 family transposase [K07485]
1458684	SNV	G	A	Ala (GCC)	Val (GTC)	410	417	98,3	Sesec_01755	CDS	Tex, transcription accessory protein [K06959]
501250	SNV	T	C	Arg (AGA)	Gly (GGA)	251	260	96,5	Sesec_RS02460	CDS	thioester-forming surface-anchored protein

V tabulce 8 jsou uvedeny identifikované substituce jednoho nukleotidu v sekvenci (SNV). Některé mutace vedly ke změnám aminokyselinové sekvence vznikajícího proteinu a některé k tzv. tiché mutaci (beze změny v aminokyselinové sekvenci). Tiché mutace byly v posledních letech hojně studovány a byl vyvrácen názor, že takovéto mutace nemají vliv na míru exprese genu. Naopak bylo zjištěno, že mohou mít významný vliv na míru exprese genu [131]. Změny v nukleotidové sekvenci mohou mít vliv na tvorbu sekundárních struktur mRNA a tím významně snížit efektivitu exprese aniž by došlo ke změně v sekvenci aminokyselinového řetězce [131]. V diskusi musí být zohledněna i možnost chybovosti v sekvenci referenčního genomu ATCC35246. Jak bylo dříve zmíněno, sekvence genomu kmene ATCC35246 [1] byla získána ne zcela spolehlivou metodou (pyrosekvenace) a mohly se v ní vyskytnout chyby.

Geny s SNV byly zařazeny podle své funkce do skupin, aby bylo možné posoudit, zda mutace mohou mít vliv na buněčné procesy. První větší skupinou byly geny zapojené v životně nepostradatelných procesech buňky. Bylo identifikováno několik SNV v genech kódujících ribozomální proteiny. Ribozom je komplexní struktura, na které probíhá translace nukleotidové sekvence do sekvence proteinu. Změny AMK řetězce proteinů, které jsou součástí ribozomu, mohly mít vliv na účinnost translace, případně její chybovost. U všech nalezených SNV ribozomálních proteinů došlo ke změně aminokyselinové sekvence (Sesec_00319, Lys→Arg; Sesec_00348, Gly→Glu; Sesec_00065, His→Arg; Sesec_01663, Cys→Arg; Sesec_00072, Pro→Ser). Mutace identifikovaná v genu Sesec_00473 (Ile→Val), který kóduje chaperon DnaK. DnaK chaperon systém byl charakterizován u několika patogenních bakteriích a existuje hypotéza, že je zapojen v důležitých buněčných reakcích i u streptokoků. Změna v tomto genu by mohla mít vliv na patogenicitu a stresovou rezistenci buněk [132].

Rychlost nárůstu biomasy mohla být ovlivněna mutací v genech zapojených v replikaci DNA. Gen Sesec_00200 kóduje DNA polymerázu I. Došlo u ní ke změně v nukleotidové sekvenci, ale AMK řetězec zůstal zachován (Ala→Ala). Na replikaci DNA se podílí i helikázy, které napomáhají k rozplétání vlákna DNA. Gen Sesec_02214 kóduje C-terminální doménu enzymu patřícího do vysoce konzervované rodiny helikáz. SNV mutace měla za následek změnu aminokyselinové sekvence (Pro→Ser). Gen Sesec_00745 kóduje enzym odpovědný za opravy chyb v sekvenci DNA. U něj nastala změna AMK sekvence (Glu→Lys), která mohla vést například ke snížení účinnosti oprav DNA a destabilizaci genomu. Nicméně z našich pozorování bakteriální kultury v průběhu fermentace můžeme vyvodit, že v důsledku výše uvedených změn nedošlo k zásadnímu ovlivnění životaschopnosti buněk. Bakterie *S. zooepidemicus* CO4A nevykazovaly odchylky v růstu či buněčném dělení oproti kmeni ATCC35246.

Nejvíce nás však zajímaly geny, ve kterých došlo k mutaci a byly spojené s produkcí kyseliny hyaluronové, nebo s blízkými metabolickými drahami. U genů přímo zapojených v produkci HA byly identifikovány dvě mutace. Gen Sesec_00234 (*glmU*) obsahoval tzv. tichou mutaci, při které nedochází ke změně v AMK sekvenci (Asn→Asn). Naopak změna proteinové sekvence (Ala→Thr) nastala u druhé kopie genu UTP-glukóza-1-fosfát-uridylyltransferázy kódované Sesec_00255 (*galU*). Geny Sesec_00255 (*galU*) a Sesec_01989 (*glmU*) jsou v genomu homologními sekvencemi genů Sesec_00234 (*glmU*) a Sesec_00233 (*galU*). Slouží ke stabilizaci produkce HA tím, že zajišťují dostatečné množství klíčových enzymů pro energetický metabolismus buněk [31].

Vliv na produkci HA mohly mít i mutace genů souvisejících metabolismů. Syntéza buněčné stěny zahrnuje kompetující metabolické dráhy k produkci HA (obr. 2, str. 7). U drah syntézy

peptidoglykanu byly identifikovány dva geny s SNV. U sekvence proteinu MurM kódovaného genem Sese_01341 (*murM*) byla změněna jedna aminokyselina (Arg→Gln). Enzym MraY kódovaný Sese_02125 (*mraY*) měl také v důsledku SNV pozměněnou aminokyselinu (Ile→Met). Expresi obou genů byla studována v souvislosti s produkcí HA i za pomoci metody DNA čipů (kap. 5.3.2). Dráha syntézy kyseliny lipoteichoové zahrnovala gen Sese_00506, u kterého byla zjištěna SNV mutace (Ala→Thr) a mohla mít vliv na expresi daného genu [133]. Mutace byly nalezeny také u dvou genů okrajových drah energetického metabolismu. U genu Sese_01616 (*glgP*) došlo ke změně Arg→Leu a u genu Sese_00836 (*uidA*) došlo k záměně Leu→Pro. Gen *glgP* je u prokaryot zapojen v drahách degradace glykogenu [134].

Větší množství změn aminokyselinové sekvence bylo zaznamenáno u genů souvisejících s ABC transportním systémem. Jde o následující geny a změny: Sese_01207 (Ile→Thr), Sese_01429 (Arg→Met), Sese_01587 (Ser→Leu), Sese_01825 (Asn→Ser), Sese_02260 (Asp→Asp), Sese_01112 (Ala→Val). Jako ABC transportní systém se označuje velké množství transmembránových proteinů, které jsou schopné aktivně převádět různé substráty přes membránu dovnitř a ven z buňky. Studie přenosu HA přes plazmatickou membránu střídavě potvrzovaly [135] nebo odmítaly [136] zapojení ABC transporteru. Zdá se však, že zapojení ABC transportního systému bylo v nejnovější studii vyvráceno [10].

Za zmínku také stojí identifikované SNV v genech transkripčních faktorů, které by mohly mít vliv na míru exprese některých genů. Jsou to geny Sese_00454 (*codY*, Thr→Met) a Sese_01822 (*ccpA*, Pro→Leu). Gen *codY* je regulátor zapojen především v metabolismu aminokyselin, ale bylo studováno i jeho zapojení do regulace virulence u *S. pneumoniae* [137]. Gen *ccpA* je zapojen v regulaci genové exprese v závislosti na uhlíkatém zdroji [138].

Delece (tab. 9) způsobily ve většině případů posun čtecího rámce, tudíž část řetězce aminokyselin mohla mít pozměněnou sekvenci či délku. Původní a nově vzniklé AMK sekvence jsou uvedeny v tabulce 10.

Tabulka 9: Bylo identifikováno 7 delecí v genomu CO4A v porovnání s referenčním genomem ATCC35246. V tabulce je uvedeno o jaké delece se jedná a na jakých pozicích je můžeme najít. Dále je uvedeno, zda se nachází v kódující oblasti genu a jaká je funkce daného genu. Červeně jsou označeny geny, ve kterých je změna, které by mohla ovlivnit růst buňky, transkripce, translaci, nebo produkci HA. CDS – protein kódující sekvence (z anglického CoDing Sequence)

Region ATCC 35246	Typ varianty	Báze v ATCC 35246	Báze v CO4A	AMK v ATCC 35246	AMK v CO4A	Délka proteinu	Shoda	Frekvence [%]	Gen	Oblast	Funkce genu
1993285	Deletion	T	-	Arg (AGA)	Asp (GAC)	503	510	98,6	Sesec_RS09440	CDS	PucR family transcriptional regulator [K09684]
1558030	Deletion	T	-	Lys (AAA)	Asn (AAC)	699	711	98,3	Sesec_RS07350	CDS	hypothetical protein
1447299	Deletion	A	-	Phe (TTT)	Leu (TTA)	487	496	98,2	Sesec_RS06850	CDS	DHA2 family multidrug efflux MFS transporter permease subunit
842641	Deletion	C	-	Pro (CCG)	Arg (CGT)	401	409	98,0	Sesec_01019	CDS	chorismate mutase [EC:5.4.99.5], [K04516]
501514	Deletion	T	-	Arg (AGG)	Gly (GGG)	67	69	97,1	Sesec_RS02460	CDS	thioester-forming surface-anchored protein
2001049	Deletion	T	-	Leu (CTT)	Leu (CTT)	574	592	97,0	Sesec_02416	CDS	surface antigen transcriptional activator SezV, AraC family regulatory protein
724137	Deletion	A	-	Lys (AAA)	Lys (AAA)	581	641	90,6	Sesec_00868	CDS	DUF4135 Lanthionine synthetase C family protein

Delece identifikované analýzou způsobují posun čtecího rámce, čímž zapříčiní změnu všech aminokyselin od místa odstranění daného nukleotidu. Změny v sekvencích AMK jsou uvedeny v tabulce 10. Gen Sesec_01019 je zapojen v biosyntéze sekundárních metabolitů a aminokyselin. Delece byly zjištěny i u dvou genů kódující transkripční regulátory Sesec_02416 a Sesec_RS09440. Změna nastala i u thioester-formujícího proteinu (Sesec_RS02460), u kterého se vyskytla i SNV mutace (tab. 8). U genu Sesec_00868 může identifikovaná varianta v případě homopolymerů vzniknout i v důsledku chyby DNA polymerázy během přípravy sekvenační knihovny či v průběhu sekvenace [139].

Tabulka 10: Identifikované delece způsobují posun v nukleotidové sekvenci genu, tudíž nutně musí mít vliv i na sekvenci aminokyselin ve vznikajícím proteinu. V tabulce jsou uvedeny sekvence AMK genu původního (ATCC35246) a také potenciální změněná sekvence AMK v CO4A. Červeně podbarvená aminokyselina je methionin, která je kódována start kodónem. Určuje začátek překladu sekvence mRNA do sekvence proteinu. Žlutě podbarvené části znázorňují sekvenci, která byla mutací změněna.

Gen	Region ATCC 35246	Typ varianty	báze v ATCC	báze v CO4A	původní AMK sekvence ATCC35246	změněná AMK sekvence CO4A
Sesec_RS09440	1993285	Deletion	T	-	M ALKDYFPDMQEGAFPLADQDWISVEVGEQYLHFPKSSLSER EQLLGLGRQHKEQSKDQESPWYHYLIQGGQEPEGLKGCQFIYIN HQSPLAGDLIELLMGIIEGVEAVLPISQTRTAFLVCHHSRSDLFRLL ADLLPTIESDFGLALTIFINGWYQVSGNALR RLF*RRKCCVYLSL *ALCWSLADLFRGHALVTLSS**WFASSIKAF*SAAFG**RGYGS GAC HVAGAWQPADSPALVYSS*FAAI*AG*ICSTVRTPFKATR*LGCL SIYTEAL	M ALKDYFPDMQEGAFPLADQDWISVEVGEQYLHFPKSSLSEREQL LLGLGRQHKEQSKDQESPWYHYLIQGGQEPEGLKGCQFIYINHQSP AGDLIELLMGIIEGVEAVLPISQTRTAFLVCHHSRSDLFRLLADLLPTIES DFGLALTIFINGWYQVSGNALR DYFKEENA VFTAYLAEHSAGRLL TF SEVMLWSLLANSGLPALSKHFSQLLL ANKEVMDLVLAMWREHGNLV QTAORLYIHRNSLQYKLDKFAQQSGLHLKQLDDLAFAYLFILKH*
Sesec_RS07350	1558030	Deletion	T	-	M KLFKKRTQHQDSFKRLIHRLSGMSESDLHKVNQLLDVVFDA TSQTNRDEKMSPELPIEPILDETEAKNTLKTQELEKRIEQFRQAK KPKKRITHQKWCVIRFLCYNESN*	M KLFKKRTQHQDSFKRLIHRLSGMSESDLHKVNQLLDVVFDA T SQ TNRDEKMSPELPIEPILDETEAKNTLKTQELEKRIEQFRQAK PK *RTKNGALSVFCVIMKVT
Sesec_RS06850	1447299	Deletion	A	-	M ETYDIHGKPYHRTLMIIILLVGTAFVNLNQTSLGTAIPTLMA NFKVPLSTTQQATTWFLLANGIMIPVSAFLATRFSTKRLYMTAYAI LVGGLLLAALAPT AHWGVFFSSA YRSGLCGRYHAFDAGGDDQR FSCGAAWCSNGNQWLGGWASPGCWTLNLRDLDFEKGFSSVRYLFR L AGHFYLAFSDFAAAGFKLISFHAPRCYPSAGC*A*FNISRLIYLRAG CFSLGLYKCCSLWLARYGVRDFTYFSRYWDDLPLFOASATIREAL FRY*SVSQQAIIYLDNNSNCLNDDGYDWC*DDAAA VSAKQRLISI KFRNGFIAGCFDYGADQSSFWPHL*QDRCSLSGVIRLYHFSCGDSTI CISK*KYLTLFDYCFVCH*NVWHINDFNAFNRKCYECLAIA*IGAW NGS**YCTPDLDFYCCSCVIECDTKHYQQQTA*A*NAPIHTP*AIC*A ASRSQFKWLSCLFFPWLYVCSGRLYCFAFPA*GKGYREI	M ETYDIHGKPYHRTLMIIILLVGTAFVNLNQTSLGTAIPTLMA N FK VPLSTTQQATTWFLLANGIMIPVSAFLATRFSTKRLYMTAYAILV G GLL LAALAPT AHWGVFL LAARIVQACAVGITMPLMQVVMINVFPAEQRGA A MGINGLVVGLAPAVGPTLSGWILKDFSVLGIYLGWRAIFILPLVILLV LSLSPFMLRDVIPQAVKLDSSISLGLSILGLGAFLWGFTNVAAYGWLD MVYVILPILVIGGIYLFKSKRQLQLEKPFDIRVFRNKQFTLTIAIALTM MAMIGVEMMLPLYLQNVRLSPLNSGMVLLPGALIMGLISPVSGRIYD KIGARYLALLGFTILAAATLPPVFLNENTSRLITVLYAIRMFGISMILMP LTASAMNALPSHESAHGTAANNTVRISSSIVAVLSSVTQNISSKQPP KMLLSIHPEQYAKQLLEANLNGFHASFFLGFMAVAVGFIVVFLHKGK VIEK*
Sesec_01019	842641	Deletion	C	-	M GLETIRQDINA VDRQLIALLEKRMALVDQVAAYKQVHDVPV LDQEREAEVLRSVSQIRHKAFEPAILLELFKAIMAASR PLSG*ANS G WFK***SL L	M GLETIRQDINA VDRQLIALLEKRMALVDQVAAYKQVHDVPVLD QEREAEVLRSVSQIRHKAFEPAILLELFKAIMAASR RYQAEQI AVGSND DEAY*

Sesec_RS02460	501514	Deletion	T	-	<p>MRKTMKKMLAASTLCIIMSGSFISGSARVLAEQYYGYDDGTQKGGSDPAFLYVSPKDNPNMSAIVYCFNKENTPPNPWEKHIWKKPEELQYPLPLYNKYSGTDNLFESNASTKLRKKEVTNITATLVA VLSHGYPHGKSHFCSEYCRDENKARKVTQLALWYFTDGDIDENTLKTNYSLSEDDMKGLQYLIKVGKEAKVGSQTLDIFYIAQENSSVNRKYQNLLGSTPILKPETPNNHCCQCTKIYIKDDGENGYWLYVY YDTNGNNIHDNNEHILQOKYIKNGKD GKQGRIGFPGLPGPKGPKGDRGEK GDRGETGPRGERGPAGEPGKPGKGDGTGPQGRGDKGETGEKGPQGNPGKD GKPGERGERGPKGDTGPQGRGDKGETGERGKDGEPGVPGPQGPGRDGDGKGEKGERGEQGPQGERGEQGPQGERGEQGRASR*TRGAR*TSR*TRRARASR*TRRARASRRTRGARASR*TRGARTSR*TRRARASR*TRRARASRRTRGARASR*TRGARTSR*TRRARASR*TRRARASR*TRGTRTSGRKSYSPRANATTNA*PYT*ANGS*TRIART*TYTSTRDKTST*TCASVRS*ASKTNAQAINHY*PI*WQITAKNK*HRQLEHSIGNRFTLSSRIRLPN*TQKKA</p>	<p>MRKTMKKMLAASTLCIIMSGSFISGSARVLAEQYYGYDDGTQKGGSDPAFLYVSPKDNPNMSAIVYCFNKENTPPNPWEKHIWKKPEELQYPLPLYNKYSGTDNLFESNASTKLRKKEVTNITATLVA VLSHGYPHGKSHFCSEYCRDENKARKVTQLALWYFTDGDIDENTLKTNYSLSEDDMKGLQYLIKVGKEAKVGSQTLDIFYIAQENSSVNRKYQNLLGSTPILKPETPNNHCCQCTKIYIKDDGENGYWLYVY YDTNGNNIHDNNEHILQOKYIKNGKD GKQGRIGFPGLPGPKGPKGDRGEK GDRGETGPRGERGPAGEPGKPGKGDGTGPQGRGDKGETGEKGPQGNPGKD GKPGERGERGPKGDTGPQGRGDKGETGERGKDGEPGVPGPQGPGRDGDGKGEKGERGEQGPQGERGEQGPQGERGEQGRASR*TRGAR*TSR*TRRARASR*TRRARASRRTRGARASR*TRGARTSR*TRRARASR*TRRARASR*TRGTRTSGRKSYSPRANATTNA*PYT*ANGS*TRIART*TYTSTRDKTST*TCASVRS*ASKTNAQAINHY*PI*WQITAKNK*HRQLEHSIGNRFTLSSRIRLPN*TQKKA</p>
Sesec_02416	2001049	Deletion	T	-	<p>MDTMSLVKSLQACLKLSITVFDKDHQVTEFVNDDTFEFYYRFQVVLQMDKEQATNLFQGSYNELFLYRYQDHYILIGPWRCNVLDKTFKVA VADKDLTKEEEEILYQALKHLPVYPLNAIRDLLIVVHYFFSGKIEDLLSPPLRHYIRDFSDNIEVQKANMFVKQESKVCLYNYEKNLSSIVQEGDVHKLKEIIFSLRNTFIPVVS GDTLRSEKNYSIAFDRLAQIAIQSGMDIADAYHSREEFMRDSELATSPMDVLKIRDASLVFYTQK VSELRDNRWKHVSPVNAIIQYISVNLHKSVTTAELAQHFN MSESRLRQIFKRDTGLTIYQYTSNMKIEEAKTLLRSAYAINHISDVLGFTDPSHFSKFFKHTGESPKTYQNRKH*</p>	<p>MDTMSLVKSLQACLKLSITVFDKDHQVTEFVNDDTFEFYYRFQVVLQMDKEQATNLFQGSYNELFLYRYQDHYILIGPWRCNVLDKTFKVA VADKDLTKEEEEILYQALKHLPVYPLNAIRDLLIVVHYFFSGKIEDLLSPPLRHYIRDFSDNIEVQKANMFVKQESKVCLYNYEKNLSSIVQEGDVHKLKEIIFSLRNTFIPVVS GDTLRSEKNYSIAFDRLAQIAIQSGMDIADAYHSREEFMRDSELATSPMDVLKIRDASLVFYTQK VSELRDNRWKHVSPVNAIIQYISVNLHKSVTTAELAQHFN MSESRLRQIFKRDTGLTIYQYTSNMKIEEAKTLLRSAYAINHISDVLGFTDPSHFSKFFKHTGESPKTYQNRKH*</p>
Sesec_00868	724137	Deletion	A	-	<p>MKTLYSKCYKEKLDYDIDVTFEQWLGLKSSLPDKIVIKNLKELFGSEYQANESIKEIYLQEQESIDIFNLLDERLVGDFNDFTKVLSKENSQIKEFLNNAKGQEQLTIENIAPFIFYSLYKEKKLLANLKRCHIDTFVELITKGVL SIAIKNHILDMNHNEKYSFYKEKSTY TEFIRNEVLNFDYERFYISNPMLLSRCFSKTIDLISYFKEVVHNISIDFDLQISKMNFSEGDTHSGGKTVVKFTSGRKNYYKPRNLFIEKVFYDVLDFISIQHVENSENKYYESYSIAAEVKYLPVNDNEIKDFYYTLGKIQFVLYLGGSDIHYENLIACGKLPVIIDLETLFQINYDHIRYGTIEVKKRLKNQLDLLQNIIGLDLYLPGE GEGMNVSA LF GSKESK SCSKRDF*</p>	<p>MKTLYSKCYKEKLDYDIDVTFEQWLGLKSSLPDKIVIKNLKELFGSEYQANESIKEIYLQEQESIDIFNLLDERLVGDFNDFTKVLSKENSQIKEFLNNAKGQEQLTIENIAPFIFYSLYKEKKLLANLKRCHIDTFVELITKGVL SIAIKNHILDMNHNEKYSFYKEKSTY TEFIRNEVLNFDYERFYISNPMLLSRCFSKTIDLISYFKEVVHNISIDFDLQISKMNFSEGDTHSGGKTVVKFTSGRKNYYKPRNLFIEKVFYDVLDFISIQHVENSENKYYESYSIAAEVKYLPVNDNEIKDFYYTLGKIQFVLYLGGSDIHYENLIACGKLPVIIDLETLFQINYDHIRYGTIEVKKRLKNQLDLLQNIIGLDLYLPGE GEGMNVSA LF GSKESK VVQKEIF</p>

Změny v nukleotidové sekvenci a sekvenci AMK mohly mít vliv například na funkci proteinu, translaci ze sekvence mRNA, případně degradaci. U čtyř genů (Sesec_RS02460, Sesec_RS09440, Sesec_RS06850, Sesec_01019) se zdá, že došlo ke stabilizaci jejich sekvence a k obnovení aktivity genu. Ale pravděpodobněji se jedná právě o chybu v sekvenci genomu kmene ATCC35246, který byl sekvenován před 9 lety metodou pyrosekvence. Možnost mutace v jednom nukleotidu právě v místě, kdy dojde k obnově aktivity enzymu, je velmi nepravděpodobná. U genů Sesec_RS07350 a Sesec_02416 došlo delecí ke zkrácení AMK sekvence a mohlo dojít k omezení transkripční aktivity genů. Možností je i to, že zjištěná delece byla způsobena chybnou sekvenací u homopolymerů anebo chybou při přípravě sekvenační knihovny [139] U posledního genu

Sesec_00868 zřejmě došlo naopak k prodloužení AMK sekvence, což mohlo vést také ke změně transkripční aktivity genu. Ale identifikovaná chyba mohla být také způsobena i chybou sekvenace či přípravou sekvenační knihovny [139].

Posledním druhem zjišťovaných mutací jsou inserce nukleotidů (tab. 11). Stejně jako delece i inserce má za následek posun nukleotidů v sekvenci (posun čtecího rámce) a změnu AMK řetězce. Původní AMK řetězec a potenciální nová AMK sekvence jsou uvedeny v tabulce 12.

Tabulka 11: V tabulce jsou uvedeny identifikované inserce v genomu kmene CO4A, jejich pozice a to, zda se nacházejí v kódující oblasti genu. Uvedena je i funkce genu a ve kterém regionu genomu kmene ATCC35246 se nachází. Červeně jsou označeny geny, ve kterých je změna, které by mohla ovlivnit růst buňky, transkripci, translaci, nebo produkci HA. CDS – protein kódující sekvence (z anglického CoDing Sequence)

Region ATCC 35246	Typ varianty	Báze v ATCC35246	Báze v CO4A	AMK v ATCC 35246	AMK v CO4A	Délka proteinu	Shoda	Frekvence [%]	Gen	Oblast	Funkce genu
258129^258130	Insertion	-	TA	Val (G-TA)	Val (GTA)	17	17	100,0	Sesec_00296	CDS	nanE, N-acylglucosamine-6-phosphate 2-epimerase [EC:5.1.3.9], [K01788]
1581568^1581569	Insertion	-	A	Asn (-AAT)	Lys (AAA)	190	192	98,9	Sesec_01901	CDS	ISL3-like element ISSeq1 family transposase [K07485]
377828^377829	Insertion	-	A	Gln (C-AA)	Gln (CAA)	408	414	98,5	Sesec_RS01855	CDS	IsdE, heme ABC transporter substrate-binding protein IsdE
2164689^2164690	Insertion	-	T	Asp (-GAC)	Stop (TGA)	588	599	98,2	Sesec_RS10210	CDS	rlmH, 23S rRNA (pseudouridine(1915)-N(3))-methyltransferase RlmH
1454257^1454258	Insertion	-	T	Ile (A-TT)	Ile (ATT)	622	641	97,0	Sesec_RS06885	CDS	DUF3270 family protein
1383191^1383192	Insertion	-	A	Thr (AC-A)	Thr (ACA)	473	488	96,9	Sesec_RS06560	CDS	ABC transporter permease subunit
222567^222568	Insertion	-	T	Lys (AA-G)	Asn (AAT)	515	532	96,8	Sesec_00251	CDS	LPXTG cell wall anchor domain-containing protein, collagen-like surface protein SclZ.5
943998^943999	Insertion	-	A	Leu (TT-A)	Leu (TTA)	458	474	96,6	Sesec_RS04460	CDS	ABC transporter, ATP-binding cassette domain-containing protein
1720652^1720653	Insertion	-	T	Met (A-TG)	Ile (ATT)	448	464	96,5	Sesec_RS08130	CDS	pknB, Stk1 family PASTA domain-containing Ser/Thr kinase
2028892^2028893	Insertion	-	T	Lys (AA-G)	Asn (AAT)	479	497	96,4	Sesec_02448	CDS	LPXTG cell wall anchor domain-containing protein
2164639^2164640	Insertion	-	T	Asn (AA-C)	Asn (AAT)	571	595	96,0	Sesec_RS10210	CDS	rlmH, 23S rRNA (pseudouridine(1915)-N(3))-methyltransferase RlmH
2164570^2164571	Insertion	-	T	Lys (AA-G)	Asn (AAT)	551	575	95,8	Sesec_RS10210	CDS	rlmH, 23S rRNA (pseudouridine(1915)-N(3))-methyltransferase RlmH
9549^9550	Insertion	-	T	Val (G-TT)	Val (GTT)	676	707	95,6	Sesec_RS00040	CDS	oligosaccharide flippase family protein
2083511^2083512	Insertion	-	T	Asn (AA-T)	Asn (AAT)	634	672	94,3	Sesec_RS09845	CDS	recA, recombinase RecA

Ve skupině identifikovaných insercí se také vyskytly geny, které je vhodné zmínit v souvislosti s produkcí HA. Jedním zásadním genem je Sesec_00296 (*nanE*), který je klíčovým genem v drahách energetického metabolismu buňky. U CO4A došlo u *nanE* vlivem inserce TA v sekvenci kódující predikovanou N-acetylmannosamine-6-fosfát-2-epimerázu k posunu čtecího rámce, čímž mohla být ovlivněna funkce tohoto enzymu, který je klíčový také pro biosyntézu neuraminové kyseliny u bakterií [140, 141, 142]. Neuraminová kyselina se vyskytuje na povrchu buněk, přispívá k jejich adhezi a je zapojena i do signálních mechanismů [141]. Patogenní bakterie ji produkují jako virulentní faktor a má podobnou funkci v buňce jako HA (ochrana před imunitním systémem hostitelského organismu) [141].

Inserce nastaly i u genů kódující proteiny spojeny s ABC transportním systémem. Jde o geny s označením Sesec_RS01855, Sesec_RS06560 a Sesec_RS04460. Forma zápisu označení genu, se kterou se setkáme v NCBI databázi je novější [74]. V databázi KEGG [130] nebyly geny pod tímto označením nalezeny, proto u nich nejsou uvedena referenční čísla pro vyhledávání. Zdá se, že byly anotovány u *S. zooepidemicus* ATCC35246 poměrně nedávno.

Zajímavými geny jsou Sesec_00251 a Sesec_02448. Kódují ve své sekvenci doménu LPXTG, jejíž funkce je blíže popsána jako protein-uchycující se k buněčné stěně. Uvedená doména se objevuje nejčastěji na C-terminálním konci proteinu a váže se kovalentní vazbou k pentaglycinovému vazebnému místu peptidoglykanu přes threonin vazebné místo domény LPXTG [133]. Je jedním ze základních virulentních faktorů u streptokoků, stejně jako kyselina hyaluronová. Jeho přítomnost na povrchu buněk může mít vliv na translokaci HA ven z buněk a tvorbu HA kapsule.

Posledním genem, který chci uvést je Sesec_RS10210 (*rlmH*). Je ribozomálním proteinem a mutace v něm byla identifikována hned třikrát. Jeho sekvence byla pozměněna a v tabulce 12 je uvedena sekvence se všemi změnami, ačkoliv v tab. 11 jsou inserce v genu uvedeny zvlášť.

Tabulka 12: Inserce stejně jako delece způsobuje posun v nukleotidové sekvenci genu, tudíž mohou způsobit změnu v kompletní části sekvence proteinu. V tabulce jsou uvedeny původní AMK sekvence dle referenčního genomu ATCC35246 a potenciální změněné sekvence AMK u CO4A. Červeně podbarvená aminokyselina je Methionin, která je kódována start kodónem. Žlutě podbarvené části znázorňují sekvenci, která byla mutací změněna.

Gen	Region ATCC 35246	Typ varianty	báze v ATCC	báze v CO4A	původní AMK sekvence ATCC35246	změněná AMK sekvence CO4A
Sesec_00296	258129^ 258130	Insertion	-	TA	<p>MPEKPSKAELLSVLKDGLIVSCQALPGEPLYSEAGGIMPLMAK AAEEAGAVGIRANGVRDIKEIQAVTDLPIIGIHKKEYPPQSPFITPTM AEVDALVALDVAVIALDCTQRRERHDGLAIAADFIRAIKEKYPDQLFM ADISTLEEGLAAEAGIDFVGTTLSTGYTPYSRQGTGPDMEIGALC EAGVSVIAEGKHSPEEARAVQALGVAGIVVGGAITRPKEIAERFVA ALTSSLSIKKY*</p>	<p>MPEKPSKAELLSVLKDGLIVSCQALPGEPLYSEAGGIMPLMAKA AEEAGAVGIRANGVRDIKEIQAVTDLPIIGIHKKEYPPQSPFITPTMAEV DALVALDVAVIALDCTQRRERHDGLAIAADFIRAIKEKYPDQLFMADIST LEEGLAAEAGIDFVGTTLSTGYTPYSRQGTGPDMEIGALCEAGVYQ **LRARFTVQRRLLGLFRLWEWLVLWLVVLLPDQRKLQSAWRL*HR HCL*KNI</p>

Sesec_01901	1581569	Insertion	-	A	<p>MEQSHVITQLLGKDPHITFSKEIHDMMKTHKELKAVLDYDAPP CPNCQSQMGKYDFQRESKVPFLDCAGYKTLIRLKKRRFKCQFCGK ITVAETSLVPKNHQIPTIVKQKVAQLLIEKVSMTAIADRLSISTSTVM RKLNEFTFKSHLTYLPEHMSWDEYAFKKGKMSFIAQDFDTNNIIAI LDGRTKAVIRNHFLRYPRQVRNDVKLITMDMFTPYYNLAKMLFPN AQIVLDRFHIVQHLGRAMNRIRTOIMNSFDRKSHEYKALKRYWKL IQDSSKLSDKRFYRPTFRMHLTNKEVVERLLSYDELKRHYDLY QLLLFHFQEKQADHFFGLIEEQIDSSNPLFQTVFKTFLKDRDKIENA LDLPYSNAKLEATNNLIKVIKRNAFGFRNFENFKRRILIAINMKKEK TNWSSPDVSLSITHYS*</p>	<p>MEQSHVITQLLGKDPHITFSKEIHDMMKTHKELKAVLDYDAPCP NCQSQMGKYDFQRESKVPFLDCAGYKTLIRLKKRRFKCQFCGKITV AETSLVPKNHQIPTIVKQKVAQLLIEKVSMTAIADRLSISTSTVMRKL NEFTFKSHLTYLPEHMSWDEYAFKKGKMSFIAQDFDTNNIIAILDGR KAVIRNHFLRYPRQVRNDVKLITMDMFTPYYNLAKMLFPNAQIVLDR RFHIVQHLGRAMNRIRTOIMNSFDRKSHEYKALKRYWKLIIQDSSKLS SDKRFYRPTFRMHLTNKEVVERLLSYDELKRHYDLYQLLLFHFQEK QADHFFGLIEEQIDSSNPLFQTVFKTFLKDRDKIENALDLPYSNAKLE ATNNLIKVIKRNAFGFRNFENFKRRILIAINMKKEKKLVL SRC*LVN YPLQL</p>
Sesec_RS01855	377829	Insertion	-	A	<p>MKKVIKLIGLAVLSLLSACVDQHPQAEKKADKERIVATSVAV VDICDRNLNLDLIGVCDLKYQLPKRYDAVKRIGLPMNPDMEIASL KPTWILSPNSLQNDLEPKYQLDTEYGFLNLTSEIGMYQSIDDLGD LFDKKEAKRLRADYEEFYQAFKKQHKQKSLRFLF*WGFLAAI* LLQTNPMLAICLSWLVVRFMFIILRTRNFCLPIQKTCFSSSLS*FCVLP MLFQTRLRTRCLRNLKMFISLKLSSLSRFMI*TISFLA*VLSLTTK RL*IICHKSFMERKST</p>	<p>MKKVIKLIGLAVLSLLSACVDQHPQAEKKADKERIVATSVAVV DICDRNLNLDLIGVCDLKYQLPKRYDAVKRIGLPMNPDMEIASLKP TWILSPNSLQNDLEPKYQLDTEYGFLNLTSEIGMYQSIDDLGDLFDR KKEAKRLRADYEEFYQAFKKQHKQKKPKVLILMGLPGSYLIATNK SYVGNLLELAGGENVYHSEDKEFLSANPEDMLQQPELILRTAHAIP DKVKDMFDKEFAENDIWKHFEAVKSQQVYDLNLFMGSAKFNKY EALNHLSQLIFYGAKEVN*</p>
Sesec_RS10210	2164690	Insertion	-	T	<p>MKIKLICVGLKEAYLRDGLAIEYQKDCLAFANVISLN*QMKNT **G*PC*KTADYGG*QDOEKLQORDFVIALAIEGKQLASEQFSL LSEVTVKGYSDIAFVIGGSLGLDQSIKRNANLLMSFGLLTLPHQLM RLVLIEQVYRAFMIQQGSPYHK*</p>	<p>MKIKLICVGLKEAYLRDGLAIEYQKRLSRFCQCDIIELADEKTPD KASHAEKQMAKEADRIKKKLQORDFVIALAIEGKQLASEQFSL SEVTVKGYSDIAFVIGGSLGLDQSIKRNANLLMSFGLLTLPHQLMRL VLIEQVYRAFMIQQGSPYHK*</p>
Sesec_RS06885	1454258	Insertion	-	T	<p>MTTISFYGKMGVYLTGKERTMPTPLKQRLKQEQEQVSOQK FQEQDLNGHSSKLELIFLPVLLFLA*SPSCLHLSFLL*I*HQFGPFS LLALSA*PSLFFQRPSPRY*CH</p>	<p>MTTISFYGKMGVYLTGKERTMPTPLKQRLKQEQEQVSOQKF QEQDLNGHSSKLELIFARIAVFSLITVLFVLLVMNLAPVWAFLI ASAISLAITSFLSKAITLLMS*</p>
Sesec_RS06560	1383192	Insertion	-	A	<p>MKNTLKA WMLAMLSVILMTAGIASAETIDIVSDTAYAPFEFKD SDQVYKIDVDIINEVAKRKGWNINMTFPGFDAAVNAIQSGQADA LMAGTTITAARQKVHFSDPYDTKIIVATKKTLLRSIAS*KEKQS VLKMEQLLKPS*INTKNRMVILLKHLTQVT*CTTVLQOALLTLLWT TSLSFNMPLTKSRILRST*RERLLEALALLSKKAAAMIISLTTSMKL* AL*KLMEPIKLS*QNGLAAMTRSQQARLLAMPLRPLL*SRFIKSW PTPPLLPLNFKLAENMSVLIWTSRQLSSKALPLRSAILALTL* MLFKLVRLTLLQVCLSHLHVVRFLIFQIPTTHQIFALGSNPEALS HTRTLRVK*SVPKKEQPPMISLKA TRIFMVTV*KPLMRPPOCTIA*IL VLLTGLWMMKLF*GTLFSEATSLIFLSKGSLLVSMALLKKEPILSS SKCLTMVWLLKHOENTTTPSLADTLIQRHLSKQPMSQPFMD* LITTSNY*LGLEQPLV*HSSPLPLLWLVLFLE*WLYPHQRTCAWPLS LLMWCVVSP*LLPLSSGECQTSKLA*LATNHRML*LLLLHFR* MVEPILPKLFVGVSRLCRQVKWRLVAAWACHITQQCEGLSFLRPL S*CCQTSISLSFP*RIRP*SQRSV*WSSSKQKGLSLEITINPSACMLS* QLST*L*LHS*QGLQNV*KRGLT</p>	<p>MKNTLKA WMLAMLSVILMTAGIASAETIDIVSDTAYAPFEFKDS DQVYKIDVDIINEVAKRKGWNINMTFPGFDAAVNAIQSGQADALM AGTTITAARQKVHFSDPYDTKIIVATKKTNPICKYSELKGTGTVGV KNGTAAQAFLDKYKKSYYGTVKTFDTGDLMSNLSLATGSDIDAVMDD EPVIQYAINQKQDLAINMKGEAIGSFVAVKKGSGYDYLIDDFNEAL GAMKADGAYQAIMTKWLGSDNDKVTAGKATGDASAKATPLKPVYKI VADSSAFPEFQIGSGEYVVIDMDLIKAIKQGGFTIEISNPGFDAALN AVQAGQADAVIAGMSITPARREIFDFSDSYTSNIRLVKSGSSIKSY KDLKGVIVGAKKGTASYDFLESNKDIYGYRLKAFDEASSMYDSLNS GSIDGLMDDEAILRYAIQRGHKLDIPIKGEPSGEYGFVKKGANPELI EMFNNGLAALKASGKYDAISR YLDSEASVKTADESTIYGLLANNYK QLLAGLGTTLGLTLISFAIALVIGIVFGMMAVSPSKNLRMASTIFVDV VRGIPLMIVA AFIFWVGNLIESMTGHQSPINDFVAATIALSLNGGAY AEIVRGGIEAVPAGQMEASRSLGLSYHTTMRRIIPQAIKMLPNFIN QFVISLKDTTIVSAIGLVELFQGTGKIIARNYQSFMYAAILAVIYLIMITL LTRLAKRLEKRLN*</p>
Sesec_00251	222568	Insertion	-	T	<p>MRYTLNKNKRARNYGLCSAAIALAAVASLGTAGSAKAEAAK TLDQLPNPEEQILYNYWDTNDQKAREYLGDLYGFLKSYLGELEK LKSLDLLDLKVKKEKKEKVVIRVSKALRGRVTKALRGRVTKA KKVKRATKAKKARRVTKAKKARKANAVSKVKKVCRVTKVSKAK</p>	<p>MRYTLNKNKRARNYGLCSAAIALAAVASLGTAGSAKAEAAKTL DQLPNPEEQILYNYWDTNDQKAREYLGDLYGFLKSYLGELEKKAKE PGPAGPKGEKGEKGDGPGQPKGEKGDQGPKEKGDQGGQKGE KGDQSQKGEKGDQGGQKGEKGORGEQQKGEKGDQGIQQGEGHGH</p>

					ASMDIRVTMEKMLSHQHLKQILNQOQKLWRQQLKRAHYQLOQR QITHSSPLQPLVLSLAPAC*H*KARKT	KGHHGKDAKPSAPKADSKSTAKAMAPAAQKGTLPATGETNHPFFTL AALGVIASAGMLTLKGKN*
Sesec_RS04460	943999	Insertion	-	A	M ANKEIAVTVDHVSKSFKLPTESTKSFRTALVNRFGIKGYTE QHVLDRDISFDVYKGDFFGIVGRNGSGKSTLLKIIISQIYVPEKGQVTV EGKMVSFIELGVGFNPELTGRENVMNGAMLGFTTEEVDMMYDD IVDFAEHLHDFMNOQLKIIQACRFVLPFLSLSRHREMC*FLMRC*LL VMRPFSENAMTTSWSARRAARLLY**PMIWWLSKNTAIAQFSSKM ALSRLTANPLMSLTNTVLIILN*KMLSKELLSLCLIWQAS*R*D*PL SQALVQMSPPFLRFPIMS*RTSLPMLPFR*LILTVIFGYTMTIHRISRQ VVLATRIYPTS VSSH*SISN*SLR*LFAIKTGKCCSFQQLIILR*SCF NETILLQMTCLLWTLVPVAYIKEMVAGSLIN	M ANKEIAVTVDHVSKSFKLPTESTKSFRTALVNRFGIKGYTEQ HYVLDRDISFDVYKGDFFGIVGRNGSGKSTLLKIIISQIYVPEKGQVTV EGKMVSFIELGVGFNPELTGRENVMNGAMLGFTTEEVDMMYDDIVDF AELHDFMNOQYKKNYSSGMQVRLAFVVAIKAQGDVLLILDEVLA VGDVAFQKCNDFMERKESGKTTLVTHDMGAVKCYCNRAVLIENGLV KAYGEPFDVANQYSVDNTELKNAEQGAVAEPVSDLASQLEVR LTSKPSLSPDEPISFEISYHVLKDEPTYVAFSLTDIDRNIWVYNDNSQDQPTSG PGHKNISYQCQLSQLNDIKLLEVTVRDQNGQMLLFSTADHSPLIVL QRNDIAPDDLSALDSASGLYQRNGSWLINQ*
Sesec_RS08130	1720653	Insertion	-	T	M IQIGKLFAGRYRILKSIGRGGMADVLLANDLILDNEDVAIKV LRTNYQTDQVAVARFQREARAMELNHPNIVAIRDIGEEDGQQFL VMEYVDGSDLKKYIQDHAPLSNNEVVRIMEEVL SAMTLAHQKGI VHRDLKPNILLTKSGVVKVTDGFI AVAF AETS LTQTNSMLGSVHYL SPEARQSKATI QSDIYAMGIMLFEMLTGHIPYDGD SAVTIALQHF QKPLPSIDENPNVQALENVVIRATAKKLSDRYRSTFEMSRDLMT ALSYNRSRERKVFDDVESTKPLPKVTPAPITAPKVSAAASQASGAA DTRPEQTSQMDTLQKPKKKKMGAF*ERS*KSYFPFSLSVWFLLLI WS*PSQLRLRFLM*QALVSMPLSRSWRRQD*RSVRSIRLKA TQSVRIRWLRQILKLAQLSEKAPALTFTRSEIKALRWKIIRDRIIRM**PV* L RPTVYQSQKSRLSVS*QMNTLKTPLLASHQAREKNLTQLVNPRLRL V*QWVIRSSCQW*LSTAMQMR*IH*QP*ELTPQGSRPTLQAQALRR ALFR*APQAQRRLSVGRHRIMVPRLAFLKRVILASTSTLRRHVHQA AHRALIPQLLIHQLEIQQTQAVIPN*VKMTLL*HR	M IQIGKLFAGRYRILKSIGRGGMADVLLANDLILDNEDVAIKVLR TNYQTDQVAVARFQREARAMELNHPNIVAIRDIGEEDGQQFLVME YVDGSDLKKYIQDHAPLSNNEVVRIMEEVL SAMTLAHQKGI VHRDLKPNILLTKSGVVKVTDGFI AVAF AETS LTQTNSMLGSVHYL SPEARQSKATI QSDIYAMGIMLFEMLTGHIPYDGD SAVTIALQHF QKPLPSIDENPNVQALENVVIRATAKKLSDRYRSTFEMSRDLMT ALSYNRSRERKVFDDVESTKPLPKVTPAPITAPKVSAAASQASGAA DTRPEQTSQMDTLQKPKKKKNGRFLGTLKILFSFFVGVV ALFAYLVLTTPKPTTKV PDVTGSSLNAAKQVLEEAGLKVGVKHQIESDTVSEDQVVKTNPQAG SAKRKGSVDLYLSIGNKGFEMENYKQDYQDVMTSLTETYGV PKS KIKIERIVTNEYPTVVISQSPSAGEKFNPA GKSKITLSVAVGDTVIMP MVTEYSYADAVNTLTALGIDAARI KAYTPSSSPTGFVQVSSPSKAI VSGOTPYGTQISLSEK GDISLYLYPEETRASSSSSSSDTASNP SAGDTTSSSHTELQONDAVITP*
Sesec_02448	2028893	Insertion	-	T	M TNKTKRTGLVRKYGACSAAIALAALASLGAGKAVKADQPA ALKYPEPRDYFLHTREGDVIYDEDIKRYFEDLEAYLTARLGGIDKK VEEAAQKPGIPGPMGLAGPQGPKGDKGDPGAPGERGPAGPKGDT GEAGPRGEQGPAGQAGERGPKGDTGAPGPKGEKGDTGAVGPKGE KGDGTGATGPKGDKGERGEKGEQQRGEKGEQQRGEKGEQQR GEKGEQQRGEKGEQQRGEKGEQKPKGDQKDKTSPAPKAPK APAPKAPKASEQSSNPKAPKAPKSAPNKSAAPTGOKATLPA TGETNH PFFTLAALSVIASAGVLT LTKGKTKLVIVSFPYDTNFNS*	M TNKTKRTGLVRKYGACSAAIALAALASLGAGKAVKADQPA LKYPEPRDYFLHTREGDVIYDEDIKRYFEDLEAYLTARLGGIDKKVE EAAQKPGIPGPMGLAGPQGPKGDKGDPGAPGERGPAGPKGDTG EAGPRGEQGPAGQAGERGPKGDTGAPGPKGEKGDTGAVGPKGE KGDGTGATGPKGDKGERGEKGEQQRGEKGEQQRGEKGEQQR GEKGEQQRGEKGEQQRGEKGEQKPKGDQKDKTSPAPKAPK APAPKAPKASEQSSNPKAPKAPKSAPNKSAAPTGOKATLPA TGETNH PFFTLAALSVIASAGVLT LTKGKID*ASYCIPV*H*F*QL
Sesec_RS00040	9550	Insertion	-	T	VRLNLEKLPKHTKQSSLLIVISGFMSKLLAASYRIPYQNLV GDRGFYAYQIYPLLGIISALGLTALPNVIASMAQKKKQLQ LAALFKLQ CCTSFLLSGILVLSHKALASW MGAQQLAPSIITAMVLLTIPFISFYR GLAQADMNMAPTALSQVLEQIRV AIIIAALCYSVFGWTVYVTA NVAAFGNL VASLVILAYLKRHSA YSLKSFFSEDTV AIRDLTSLGLPT LVFYFFQFTCLFFNW*ILFW*KIF WLIQAYQS*QLK*LRVFTIGD SPC CNLVSFFRQPYLQPIYRI*Q PYFMLKGRSIKNRVNSLSFILV* R*LLALLVFFI**IEPYLR TIRAGWP**FT*L*QCQV* FNSFIKNASLKIT SSSLCLFC CLAFY*N*A*HLY*LTITLL LEAPYQL*SPCWSFYSYILC WLILILKR**MLNMGLL** SCSCVCLQVSVICHYQVGLVR* LVFWC HLLL VYPLF**LAKS* ISLRRNCGLFYHLSKKN	VRLNLEKLPKHTKQSSLLIVISGFMSKLLAASYRIPYQNLV GDRGFYAYQIYPLLGIISALGLTALPNVIASMAQKKKQLQ LAALFKLQ CCTSFLLSGILVLSHKALASW MGAQQLAPSIITAMVLLTIPFISFYR GLAQADMNMAPTALSQVLEQIRV AIIIAALCYSVFGWTVYVTA NVAAFGNL VASLVILAYLKRHSA YSLKSFFSEDTV AIRDLTSLGLPT LVFYFFQFTCLFFNW*ILFW*KIF WLIQAYQS*QLK*LRVFTIGD SPC CNLVSFFRQPYLQPIYRI*Q PYFMLKGRSIKNRVNSLSFILV* R*LLALLVFFI**IEPYLR TIRAGWP**FT*L*QCQV* FNSFIKNASLKIT SSSLCLFC CLAFY*N*A*HLY*LTITLL LEAPYQL*SPCWSFYSYILC WLILILKR**MLNMGLL** SCSCVCLQVSVICHYQVGLVR* LVFWC KKLDFEKEKLSFLPFVKEK*

Sesec_RS09845	2083512	Insertion	-	T	<p>IGKKS*KNEEITKKFGDERRKALDDALKNIEKDFGKGA VMRLG ERAEQKVQVMSSGSLALDIALGAGGYPKGRIIEYGPESGKTTVA LHAVAQAQKEGGIAAFIDAEHALDPAYAAALGVNIDELLSQPDS GEQGLEIAGKLIDSGAVDLVVVD SVAALVPRAEIDGDIIGNHVGL QARMMSQAMRKLASINKTKTIAIFINQLREKVGVMFGNPETTPGG RALKFYASVRLDVRGTTQIKGTGDKDSSIGKETKIKVVKNKVAPP FKVAEVEIMYGEISRTGELIKIASDLDIQKAGAWFSYNGEKIGQG SENAKRYLADHPELFEIDHKVRVKFGLLEDTEESAAVDPAAAKA DELVLELDDAIEIED*</p>	<p>IGKKS*KK*RNHQKIW**TS*GT**CVKEH*KGFW*GCGYAPW*A CRAKGSDELRSQCFRHCASWRWLSQRAYY*NLWTRVFW*NNSCP ACSSAGSKRRWYCSLH*CGACLGSCCLCGARC*Y**AAFVTAGFW* ART*DSR*TD*FWCC*FGCCRLCCSFSASC*D*W*YW**PRWSAGSYD ESGNA*AFSLNQ*NQDNCDLY*PAA*KGRGYVW*SRDDTRWSCFEIL CLCPGCSWNNTNKRNRWSKRQ*YW*GNQD*GC*E*GCSAI*GG*G* NHVWRRHLTYR*AD*NCFRFRHYSKGWCLVLL*R*KNWSGL*KCQE IFG*SPRAV**D*P*GAC*IWLA*R YRGKCSCRSSCSQSR*AGFRARRC H*N*GL</p>
---------------	---------	-----------	---	---	--	---

U většiny uvedených genů zasažených mutací se zdá, že došlo ke stabilizaci sekvence proteinu a pravděpodobně i k aktivaci jeho funkce (Sesec_RS01855, Sesec_RS06885, Sesec_RS06560, Sesec_00251, Sesec_RS04460). Zde se však pravděpodobně jedná o chybu při sekvenaci referenčního genomu ATCC35246. Pravděpodobnost mutace s takovým důsledkem na sekvenci proteinu je velmi nízká.

Když se podíváme na změnu u genu *nanE* (Sesec_00296), kterou mutace vyvolala v sekvenci aminokyselinového řetězce proteinu, zjistíme, že koncová část proteinu zřejmě nebude zcela funkční. Vznik mnoha nových stopkodónových sekvencí pravděpodobně zastaví translaci předčasně. Musíme ovšem vzít v úvahu i možnost chybné sekvence v genomu kmene ATCC35246. Analýzou pomocí metody DNA čipů (kap. 5.3.1) a qPCR (kap. 5.4) bylo zjištěno signifikantní navýšení (nejvyšší jaké bylo v práci zaznamenáno) míry exprese u tohoto genu v porovnání s genem referenčního kmene ATCC35246. Gen *nanE* má v genomu homologní sekvenci (Sesec_00291).

Ve dvou dalších případech (Sesec_01901, Sesec_02448) byla sekvence díky mutaci zkrácena (zařazení stopkodónu) o několik aminokyselin. Zkrácení nemuselo mít negativní vliv na funkci proteinu. U proteinu kódovaného genem Sesec_RS09845 byla sekvence natolik destabilizována, že došlo pravděpodobně k inaktivaci jeho funkce. Nicméně spolu s genem Sesec_RS00040 jsou v NCBI [74] databázi popsány jako pseudogeny. Jejich AMK sekvence odpovídá sekvenci v databázi. Startkodónem u některých prokaryot může být i kodon pro Leucin (UUG) nebo pro Valin (GUG) [143]. V obou případech se v nukleotidové sekvenci vyskytují opakující se sekvence několika T za sebou. Oblast proto mohla být osekvenována nesprávně, nebo řada T nukleotidů mohla být vpravena chybou DNA polymerázy při přípravě knihovny či samotném sekvenování [139].

Rozšířenou analýzou byly nalezeny ještě tři mutace mimo hlavní analýzu. Jde o jednu inserci (Sesec_00735) a dvě delece (Sesec_00791, Sesec_01092). Geny a jejich charakterizace jsou uvedeny v tabulce 13.

Tabulka 13: V této tabulce jsou uvedeny dvě významné delece a jedna inserce, které byly identifikovány rozšířenou srovnávací analýzou u kmene CO4A a jejich frekvence byla větší než 90 %. Jsou zde jejich pozice, rozsah, ve kterém genu byly nalezeny a zda jsou v kódující oblasti genu. Červeně jsou označeny geny, ve kterých je změna, které by mohla ovlivnit růst buňky, transkripci, translaci, nebo produkci HA.

Region ATCC 35246	Typ varianty	Báze v ATCC35246	Báze v CO4A	AMK v ATCC 35246	AMK v CO4A	Délka proteinu	Shoda	Frekvence [%]	Gen	Oblast	Funkce genu
611156	Insertion	-	T	Lys(A-AG)	Ile (ATA)	59	59	100	SeseC_00735	CDS	collagen-like protein SclZ.7
656982	Deletion	AGCATGGAAGCAT GGA	-	Glu (GAA)	Glu (GAA)	86	86	100	SeseC_00791	CDS	hypothetical protein, Transglutaminase-like superfamily
910673	Deletion	GCCAAT	-	Glu (GAG)	Glu (GAG)	31	34	91	SeseC_01092	CDS	LPXTG cell wall anchor domain-containing protein

Získané doplňující informace rozšířily významné mutace v genomu kmene CO4A o tři další. Gen SeseC_01092 kóduje protein obsahující doménu LPXTG, která byla identifikována jako jeden z virulentních faktorů u streptokoků [133]. V této sekvenci došlo k deleci šesti nukleotidů v pozici, kdy nedochází k posunu celé sekvence. Důsledkem byla delece dvou aminokyselin ze sekvence proteinu. Odstranění dvou aminokyselin nemusí, ale může vést k zásadním změnám funkce na úrovni proteinu a k odchylkám u exprese mutovaného genu. U gen SeseC_00735 kódujícího protein kolagenového typu, který je pravděpodobně uchycen na stěně buňky, byla identifikována inserce jednoho nukleotidu. Povrchové buněčné proteiny mohou mít vliv na transportu HA ven z buňky a na formaci kapsule.

V tabulce 14 jsou uvedeny změny v sekvenci aminokyselin na základě změn v nukleotidovém řetězci.

Tabulka 14: Identifikované delece, při druhé podrobnější analýze genomu CO4A, byly podstatně delší. V této tabulce jsou uvedeny původní AMK sekvence (vycházející z genomu kmene ATCC35246) a AMK sekvence potenciálně změněné (kmen CO4A). Červeně podbarvená aminokyselina je Methionin, která je kódována start kodónem. Žlutě podbarvené části znázorňují sekvenci, která byla mutací změněna.

Gen	Region ATCC 35246	Typ varianty	báze v ATCC	báze v CO4A	původní AMK sekvence ATCC35246	změněná AMK sekvence CO4A
Sesec_00735	611156	Insertion	-	T	<p>MRKPIRKYSVCSTVIALATMASLGTAGNVRAETKLDY PGVQSGIPYKYGTVDQDFGVRKYLSELEKFFKDYLDLEQQA QKGGPQGPVGPAGPRGERGEAGPQGPAGQTGPMGPRGETG PAGPAGPIGPKGEAGQPGEPSKGEQGPAGPTGPIGPRGDKG ETGEAGPKGEQGPAGPAGPKGEAGPKGDKGEPGTPGER GPKGDRGERGEKGEQGORGEKGEKGQRGEKGEKGQRGEK GEKANAAKKASKANAAKKVNKANAVKKAKKANAAKRAK KANVVKKAKKANAAKKVNKANVVKKASNSQVSKERIPSH QLQKRLKRLHQRLLQKNQLSLRHORLSQHQLKRARYQLQ GRQITHSSPLQLASSLAQAS*A*RRKKA</p>	<p>MRKPIRKYSVCSTVIALATMASLGTAGNVRAETKLD YPGVQSGIPYKYGTVDQDFGVRKYLSELEKFFKDYLDLEQ QAQKGGPQGPVGPAGPRGERGEAGPQGPAGQTGPMGPRG ETGPAGPAGPIGPKGEAGQPGEPSKGEQGPAGPTGPIGPR GDKGETGEAGPKGEQGPAGPAGPKGEAGPKGDKGEP GTPGERGPKGDRGERGEKGEQGORGEKGEKGQRGEKGEK GQRGEKGEKGEQQRGEKGEQQRGEKGEQQRGEKGEKGO RGEKGEKGEQQRGEKGEKGEQQRGEKGEQQRGEKGEQPKG EQGKDTKPSAPKAPEKAPTPEKSTQPKAPAPKSAAPAA QKSTLPATGETNHPFFTLAALSVIASAGVLSLKKKKS*</p>
Sesec_00791	656982	Deletion	AGCATGGAAGC ATGGA	-	<p>MEAWKHGSMEAWKHGRTLCKGLLLGLACGAISMSLVV KSTTVVADDVRLSEQSTVILDNTVSNLISEVKDELAKNHQDA LEIYVLGEGSPKYSVSPDRLDGYKDYKPVRLNGKVIHGRA FFMMKKEDQLVSTKEEYDRIIAERFSNLEKFDYPTIHYTFDA ELKDKYPYLNILKANGLSLSDSISLDSILGINGVTIHTQHLRI DNQFAYRYTKDDYKQDSHITFERFRNDDKLELIVTSGAD REGLTDAQRVKAYVLYMANHVDYDWEAARYNPDYLFICQ ASDLFALTERNKAMCIGFSTAAARGLNLLGIPSYVTVGRNRN GAAHAIRVFDYDKKWRDLVDTGVGRGKYTEEHYDNLNTH DYKIFDLTKETDVLDRGYMEIKDFEDWILSHNTKELLFIN QDANIRNKVAKDDFKYIQESEKQSLIDKLLKFFVKVENISSY FKKEHFKKLVIGLGEDAKNELAKLTLSPDLEDDYKKSCTV LEWLNIIYSQCIREGDPDLIRKDGELKEMDS*</p>	<p>MEAWKHGSMEGLCVKDYD*D*HVVLYLCL*SSRVQL *LRM*GLASSRRSF*ITLCQI*LVLRLMS*PKIIRTR*KSMS WEREVQNIVECLPIVLVTT*KIISLGC*MGR*YMEEPFS**K RKISWYQLKKNMTA**QKGFQI*KSLIILLYTIPLMQS*RINI LI*T*Y*KRTAFHSLIR*A*TQY*V*MVLLFIRSICVLIINLLIG ILKTIISKTAI*RLNVLRGMMTN*RN*L*HLELIERG*QMLN E*KPMSCIWQTMLIMIGRQQDIILII*YFVRLICLHLQSEIRP CV*DSVQRLQED*IY*GYHLM*LMVEIEMAQHMRRYVFFT IKSGEH*MLQELVEGNTQKNIMTISIRTIIRFLI*LRKQTLI*I VAIWSIRILKIGSSHTIRKSYYLLIKMRIYVIRLQKMILSTY KSLRSNR**IS*RNLSRLKISVAILRKSISKS*LLD*VRTLK MS*QN*P*VLILKRTTTKSPKLC*SG*FIIRSVLEKTEIL*LE KMES*KRWIV</p>
Sesec_01092	910673	Deletion	GCCAAT	-	<p>MLTNQTYRHRLSKQKLPISMVSCLLGMGVFISPTSISAD SPPTTIKSAADKDISSNLPKIPLSNAKPEEKLEDTQPLNEKG YGNDIRDLGKPEDLPKLPDEKGYGNDIRKPEKTPELNEKGGQ GDDLKKEEDLKKQRKELTEKLRGLKTVEEVLEFLKMPGP VTDVDWQFLMAFVEKGTANKTFEKDKLTLDDIYSTSSGHMI WLGYLVEKPEPMPEPKPEPKMPKPEPMPKMPKPEPKPM PEPKMPPEPKMPKPEPKMPKPEPKMPPEPKMPPEPKPM PKPEPKMPKPEPKPEPKMPKPEPKMPKPEPKMPKPEPKM KAMEHKLPSTGEATAPFFTAALAVLASLGVAVVSRKRKK H*</p>	<p>MLTNQTYRHRLSKQKLPISMVSCLLGMGVFISPTSIS ADSPPTTIKSAADKDISSNLPKIPLSNAKPEEKLEDTQPLN EKGYGNDIRDLGKPEDLPKLPDEKGYGNDIRKPEKTPEL NEKGGDDLKKEEDLKKQRKELTEKLRGLKTVEEVLEFL KMPGPVTDVDWQFLMAFVEKGTANKTFEKDKLTLDDIY STSSGHMIWLGYLVEKPEPMPEPKPEPKMPKPEPKMPK PEPKMPPEPKMPKPEPKMPKPEPKMPKPEPKMPPEPKM MPEPKMPKPEPKPEPKPEPKMPKPEPKMPKPEPKMPKPE QKGDKPAVKAMEHKLPSTGEATAPFFTAALAVLASLGV VAVVSRKRKKH*</p>

U prvního genu *Sesec_00735* se zdá, že došlo vlivem delece k úpravě AMK sekvence, která vedla k její stabilizaci. V tomto případě spíše půjde o chybu v původních sekvenačních datech referenčního genomu ATCC35246, uvedená oprava sekvence je nepravděpodobná. U genu *Sesec_00791* se zdá, že došlo naopak k inaktivaci sekvence proteinu. Ovšem deletovaná sekvence (AGCATGGAAGCATGGA) je problematická z důvodu repetitivních úseků. Může se jednat o chybu při sekvenaci genomu CO4A, protože takováto změna UV mutagenézí je nepravděpodobná. Původní data ze sekvenace genomu ATCC35246 se v tomto případě zdají být v pořádku. U posledního genu *Sesec_01092* došlo k deleci dvou AMK z řetězce. Zde se může také jednat o chybu sekvenace genomu CO4A z důvodu repetitivní oblasti, stejně tak jako chybu v sekvenci referenčního genomu ATCC35246. Gen kóduje protein s LPXTG doménou, která je virulentním faktorem streptokoků [133].

Porovnáním osekvenovaného genomu produkčního kmene *S. zooepidemicus* CO4A s kmenem referenčním (sbírkovým) *S. zooepidemicus* ATCC35246 [1] byly identifikovány změny v nukleotidové sekvenci. Celkem bylo popsáno 59 SNV a 24 indel mutací. V důsledku mutací byly v některých případech pozmeněny sekvence vznikajících proteinů a mohla být ovlivněna jejich funkce, nebo míra exprese genů, kterými jsou kódovány. Byla diskutována i možnost zkrácení výsledků kvůli chybným sekvenačním datům referenčního genomu kmene ATCC35246.

5.3 Sledování exprese genů pomocí DNA čipů

Syntéza kyseliny hyaluronové může být ovlivněna intenzitou exprese genů u různých metabolických drah, které spolu kompetují (biosyntéza peptidoglykanu, metabolismus aminocukrů a nukleotidů aj.) nebo se doplňují (metabolismus zpracování sacharózy aj.). Sledování celogenomové exprese metodou DNA čipů u bakteriálních kmenů *S. equi* subsp. *zooepidemicus* (interní kmen s označením CO4A a kmen referenční ATCC35246) bylo provedeno za kultivačních podmínek uvedených v kap. 4.1.2. Vyizolovaná RNA (kap. 4.3) byla reverzní transkripcí přepsána do cDNA (kap. 4.4). Značení cDNA, hybridizace na čip a zpracování dat jsou popsány v kapitole 4.7. Analýza byla zaměřena na srovnání míry exprese u kmene referenčního ATCC35246 a kmene produkčního CO4A v souvislosti se zařazením k určité metabolické skupině. Pokud byl rozdíl exprese mezi kmeny nižší než 10 %, byla exprese považována za signifikantně neměnnou.

5.3.1 Geny energetického metabolismu

Glykolýza a glukoneogeneze

První velkou genovou skupinou byly geny energetického metabolismu a s ním související dráhy. Patří sem geny zapojené v glykolýze a glukoneogenezi (tab. 15).

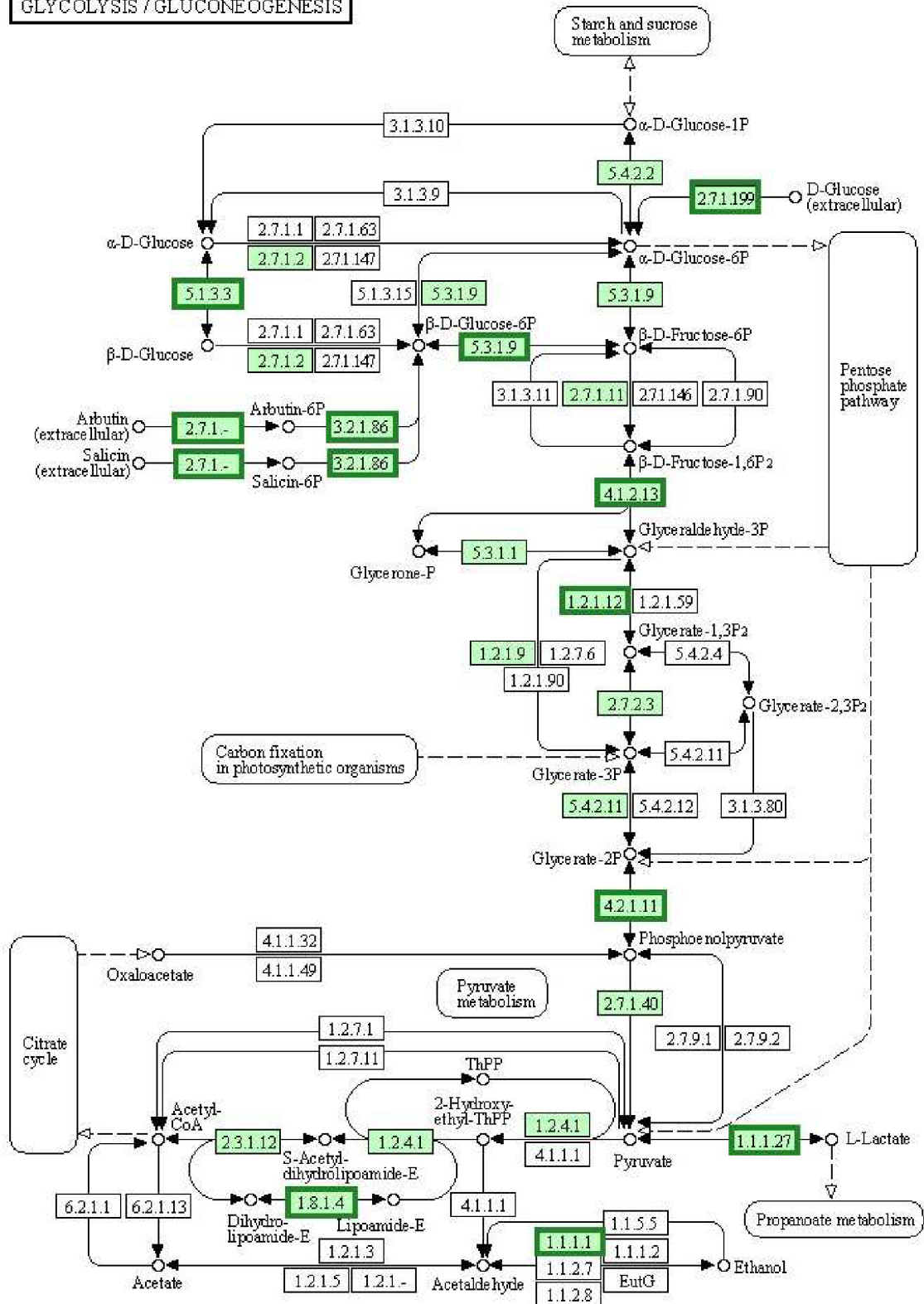
Tabulka 15: V tabulce jsou uvedeny geny zapojené v energetickém metabolismu kmene CO4A vztahující se k metabolickým drahám glykolýzy a glukoneogeneze. Uvedeno je označení genu, jeho funkce, zkratka genu a číselné označení k vyhledávání v databázi KEGG [130]. Jsou zde uvedeny poměry exprese genů u obou kmenů normalizované k referenčnímu genu *gyrA*. Na obrázku 10 jsou geny vykresleny v kontextu dalších drah. ¹⁾ Poměr celkové exprese genu v ATCC35246 k celkové expresi *gyrA* (ATCC35246), ²⁾ Poměr celkové exprese genu v CO4A k celkové expresi *gyrA* (CO4A), ³⁾ Poměr celkových exprese obou kmenů normalizovaných ke *gyrA*: $CEG_{gyrA}(CO4A) / CEG_{gyrA}(ATCC)$

Gen	CEG_{gyrA} (ATCC) ¹⁾	CEG_{gyrA} (CO4A) ²⁾	Poměr ³⁾	Změna exprese	Funkce genu
Sesec_00051	1,29	2,05	1,59	↑59%	adhE; aldehyde-alcohol dehydrogenase 2 [KO:K04072] [EC:1.2.1.10 1.1.1.1]
Sesec_00054	1,74	1,98	1,14	↑14%	adhP; alcohol dehydrogenase [KO:K13953] [EC:1.1.1.1]
Sesec_00235	1,83	2,21	1,21	↑21%	gpi; glucose-6-phosphate isomerase [KO:K01810] [EC:5.3.1.9]
Sesec_00292	0,96	1,12	1,17	↑17%	ptsG; sugar phosphotransferase system (PTS), IIBC component [KO:K02778 K02779] [EC:2.7.1.199]
Sesec_00293	0,86	0,97	1,13	↑13%	PTS system, maltose and glucose-specific IIA-B-C component [KO:K02777] [EC:2.7.1.-]
Sesec_00324	4,59	6,38	1,39	↑39%	gapA; glyceraldehyde-3-phosphate dehydrogenase [KO:K00134] [EC:1.2.1.12]
Sesec_00326	1,87	1,85	0,99	↓1%	pgk; phosphoglycerate kinase [KO:K00927] [EC:2.7.2.3]
Sesec_00709	1,82	1,83	1,00	0%	glk; glucokinase [KO:K00845] [EC:2.7.1.2]
Sesec_00766	1,27	1,36	1,07	↑7%	gpmA; 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase GpmA [KO:K01834] [EC:5.4.2.11]
Sesec_01035	1,87	1,97	1,05	↑5%	gapN; NADP-dependent glyceraldehyde-3-phosphate dehydrogenase [KO:K00131] [EC:1.2.1.9]
Sesec_01044	2,47	3,02	1,22	↑22%	eno; phosphopyruvate hydratase [KO:K01689] [EC:4.2.1.11]
Sesec_01167	1,72	1,62	0,94	↓6%	pfk; 6-phosphofructokinase [KO:K00850] [EC:2.7.1.11]
Sesec_01168	1,53	1,67	1,09	↑9%	pyk; pyruvate kinase [KO:K00873] [EC:2.7.1.40]
Sesec_01237	1,19	1,22	1,02	↑2%	manB; phosphomannomutase PgmA [KO:K01835] [EC:5.4.2.2]
Sesec_01349	1,48	1,98	1,33	↑33%	ldh; L-lactate dehydrogenase [KO:K00016] [EC:1.1.1.27]
Sesec_01576	3,17	3,48	1,10	↑10%	lpd; dihydrolipoyl dehydrogenase LpdA [KO:K00382] [EC:1.8.1.4]
Sesec_01577	1,62	1,50	0,93	↓7%	aceF; branched-chain alpha-keto acid dehydrogenase subunit E2 [KO:K00627] [EC:2.3.1.12]
Sesec_01578	1,49	1,45	0,97	↓3%	pdhB; pyruvate dehydrogenase E1 component, beta subunit [KO:K00162] [EC:1.2.4.1]
Sesec_01579	1,82	1,71	0,94	↓6%	pdhA; pyruvate dehydrogenase E1 component alpha subunit PdhA [KO:K00161] [EC:1.2.4.1]
Sesec_01655	0,70	0,90	1,29	↑29%	bglB; beta-glucosidase [KO:K01223] [EC:3.2.1.86]
Sesec_01706	1,16	1,19	1,03	↑3%	tpiA; triosephosphate isomerase [KO:K01803] [EC:5.3.1.1]

GLYKOLÝZA A GLUKONEOGENEZE

Sesec_01767	0,86	1,03	1,19	↑19%	6-phospho-beta-glucosidase [KO:K01223] [EC:3.2.1.86]
Sesec_02318	1,61	1,92	1,20	↑20%	fbaA; fructose-bisphosphate aldolase [KO:K01624] [EC:4.1.2.13]
Sesec_02351	0,50	0,61	1,21	↑21%	galM; aldose 1-epimerase GalM [KO:K01785] [EC:5.1.3.3]
Sesec_02421	1,57	1,96	1,25	↑25%	gamP; glucose-specific phosphotransferase system (PTS), IIABC component [KO:K02777 K20107 K20108] [EC:2.7.1.- 2.7.1.208]
Sesec_02449	0,71	0,68	0,97	↓3%	phosphoglycerate mutase family protein [KO:K01834] [EC:5.4.2.11]
Sesec_00425	0,59	0,63	1,05	↑5%	sacA, sucrose-6-phosphate hydrolase, beta-fructofuranosidase [EC:3.2.1.26], [K01193]
Sesec_00426	3,11	3,31	1,06	↑6%	scrA, sucrose-specific phosphotransferase system (PTS), IIABC component [EC:2.7.1.211], [K02810]

GLYCOLYSIS / GLUCONEOGENESIS



Obrázek 10: Na obrázku jsou vyznačeny zeleným orámováním sledované geny metabolických drah části energetického metabolismu *S. zooeconomicus* se změnou míry exprese vyšší než 10 % (poměr exprese genu v kmeni CO4A vůči kmeni ATCC35246). Zeleně podbarvené geny se vyskytují u *S. zooeconomicus*. Bíle podbarvené geny nejsou u *S. zooeconomicus* přítomny. [upraveno dle 130]

Metabolické dráhy glykolýzy a glukoneogeneze jsou centrálními drahami energetického metabolismu prokaryotických buněk. Byly identifikovány geny, které dosahovaly 10 % změny míry exprese v porovnání kmene CO4A s kmenem referenčním (obr. 10). Na schématu jsou ukázány tři extracelulární substráty (D-glukóza, arbutin a salicin), které jsou těmito drahami zpracovávány. Geny, zapojeny v jejich využití v buňce (Sesec_00292, EC:2.7.1.199; Sesec_01655, EC:3.2.1.86; Sesec_00293, Sesec_02421, EC:2.7.1.-, skupina genů PTS systému), měly vyšší než 10 % nárůst exprese v porovnání s kmenem ATCC35246. Zmíněné extracelulární substráty nebyly v kultivaci použity (arbutin a salicin jsou rostlinného původu) proto zřejmě navýšení exprese u jmenovaných genů nesouvisí se zpracováním arbutinu a salicinu. Pro naše kultivace byla jako uhlíkatý zdroj použita sacharóza, která je štěpena na fruktózu a glukózu-6-P. V souvislosti s tím byla hodnocena i míra exprese u genů zpracovávajících extracelulární sacharózu. Jsou to geny Sesec_00426 (*scrA*), který je odpovědný za transport sacharózy do buňky (je součástí PTS transportního systému) a gen Sesec_00425 (*sacA*), který je odpovědný za štěpení sacharózy na glukózu-6-P a fruktózu. Jejich zvýšená exprese by mohla dodat prekurzory do drah biosyntézy HA. Tato hypotéza se nepotvrdila. Byl zjištěn jen minimální nárůst exprese v porovnání s kmenem referenčním (5 % u genu *scrA* a 6 % u genu *sacA*).

Následující skupina genů zahrnuje enzymy centrální linie metabolismu sacharidů. Patří sem gen Sesec_00235 kódující enzym glukózu-6-P izomerázu (*gpi*, EC:5.3.1.9), která je centrálním enzymem a katalyzuje změnu β -D-glukózy-6-P na β -D-fruktózu-6-P. Její exprese byla o 21 % vyšší než u kmene ATCC35246. Enzym je součástí *has* operonu, ale jeho exprese je řízena samostatným promotorem umístěným v sekvenci před genem [31]. Jeho exprese tedy není spojena s expresí ostatních genů *has* operonu. Navýšení míry exprese v drahách glykolýzy mohlo mít potenciálně vliv na dráhy syntézy HA v pozdějších hodinách kultivace. Dalším genem hlavní linie je gen Sesec_02351, který kóduje aldózu-1-epimerázu (*galM*, EC:5.1.3.3). Katalyzuje změnu α -D-glukózy na β -D-glukózu. Zde bylo zaznamenáno navýšení míry exprese o 21 %. Gen pro fruktózu-bisfosfát aldolázu a označením Sesec_02318 (*fbaA*, EC:4.1.2.13) kóduje enzym měnící β -D-fruktózu-1,6-bisfosfát na glycerinaldehyd-3-fosfát. Do tohoto kroku vstupuje i glycerinaldehyd-3-P z pentózového cyklu. Navýšení exprese pro tento gen bylo stanoveno na 120 % v porovnání s referenčním kmenem. Dalším studovaným genem centrální linie byl gen Sesec_00324 kódující glycerinaldehyd-3-P dehydrogenázu (*gapA*, EC:1.2.1.12). Enzym navazuje na předešlý krok metabolismu a katalyzuje změnu glycerinaldehyd-3-P na glycerát-1,3-bisfosfát. Jeho expresní aktivita byla o 39 % vyšší než u kmene ATCC35246. Tak výrazné navýšení míry exprese mohlo být způsobeno přísunem substrátu z drah pentózového cyklu. Posledním sledovaným genem centrální linie byl Sesec_01044 kódující fosfopyruvát hydrolázu (*eno*, EC:4.2.1.11). Enzym mění glycerát-2-P na fosfoenolpyruvát (vstupní substrát pro metabolismus pyruvátu). Zde bylo zaznamenáno 22 % navýšení míry exprese. Všechny sledované geny centrální linie měly v porovnání s referenčním kmenem navýšení míry exprese o cca 20 %. Gen Sesec_00324 (*gapA*) měl navýšení exprese o téměř 40 %.

U vedlejší linie drah metabolismu sacharidů (metabolismus pyruvátu) byl sledován gen Sesec_00051 kódující aldehyd-alkohol dehydrogenázu 2 (*adhE*, EC:1.2.1.10; EC:1.1.1.1). Jeho míra exprese dosahovala nejvyšších hodnot z této skupiny genů, a to navýšení o 59 %. Gen je zapojen v celé řadě metabolických drah (metabolismus pyruvátu, degradace mastných kyselin, biosyntéza sekundárních metabolitů, metabolismus aminokyselin aj. [130]). Katalyzuje reakci přeměny etanolu na acetaldehyd. Navýšení míry exprese může mít více příčin, ale je zde jen

malá pravděpodobnost, že by tento gen výrazně ovlivňoval produkci HA. Gen *Sesec_01349* kóduje L-laktát dehydrogenázu (*ldh*, EC:1.1.1.27). Enzym katalyzuje změnu pyruvátu na L-laktát (vstupuje do metabolismu propanoátu). Bylo zde zjištěno navýšení míry exprese o 33 %. Posledním sledovaným genem v této skupině byl *Sesec_01576* kódující dihydrolipoyl dehydrogenázu (*lpdA*, E.C:1.8.1.4). Enzym katalyzuje změnu dihydro-lipoamidu-E na lipoamid-E. Z této části dráhy vystupuje acetylCoA. U genu bylo zaznamenáno navýšení míry exprese o 10 %. Je to nejmenší nárůst exprese ze všech zmiňovaných genů.

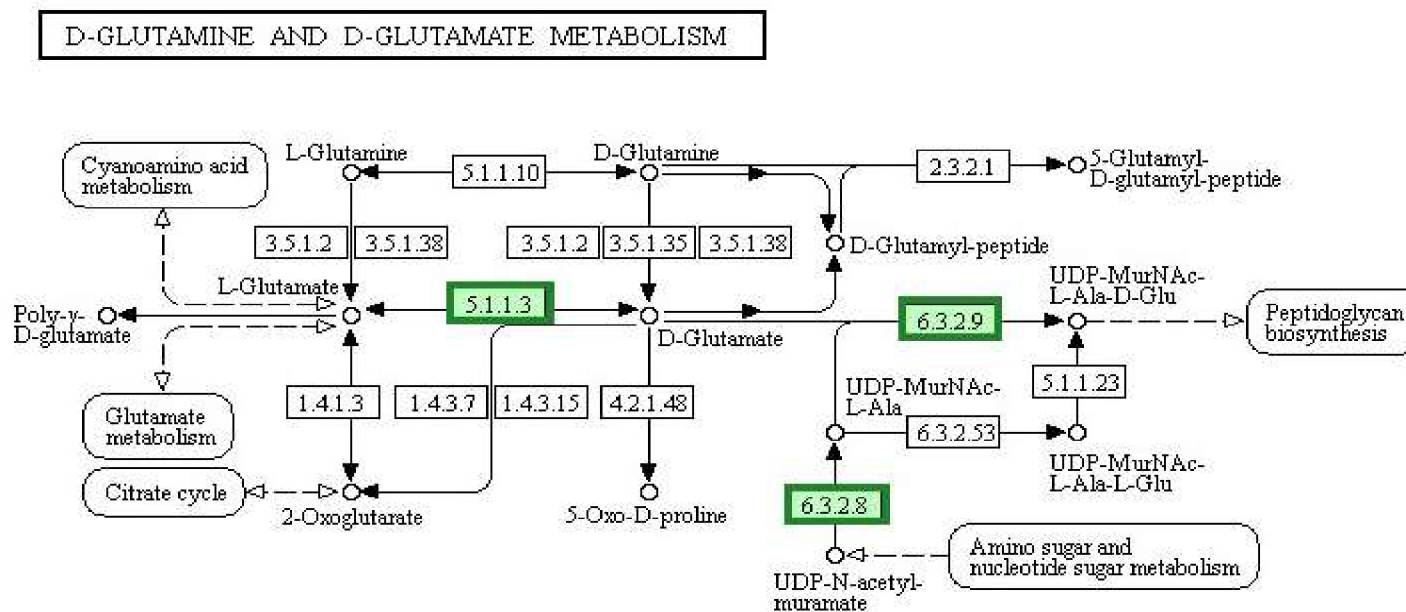
V uvedeném komplexu metabolických drah nelze identifikovat jednoznačnou kaskádu genů s navýšenou mírou exprese, která by měla vliv na biosyntézu HA. Pozorujeme navýšení míry exprese u genů, bezprostředně zpracovávajících metabolity vycházející z pentózafosfátového cyklu. Zejména byla exprese navýšena u reakcí přeměny pyruvátu na laktát coby konečný produkt metabolismu.

Metabolické dráhy glutaminu a glutamátu

Druhou skupinou sledovaných genů byly ty zapojené v metabolismu glutaminu a glutamátu (tab. 16). Metabolické dráhy glutaminu a glutamátu souvisejí s centrálními drahami energetického metabolismu pouze okrajově (obr. 11). Z této skupiny byly sledovány geny, které zpracovávají UDP-*N*-acetylmuramát přicházející z drah metabolismu aminocukrů a nukleotidů. Prvním sledovaným genem byl SeseC_02212, který kóduje UDP-*N*-acetylmuramát-L-alanin ligázu (*murC*, EC:6.3.2.8). Enzym katalyzuje připojení L-alaninu na UDP-*N*-acetylmuramát. Navýšení exprese tohoto genu bylo o 12 % v porovnání s referenčním kmenem. Hned dalším genem pokračujícím v kaskádě připojuje D-glutamát k UDP-*N*-acetylmuramát-L-alaninu. Jedná se o gen s označením SeseC_00714 kódující UDP-*N*-acetylmuramoyl-L-alanyl-D-glutamát syntázu (*murD*, EC:6.3.2.9). Míra exprese genu byla o 13 % vyšší než u kmene ATCC35246. Posledním studovaným genem byl gen s označením SeseC_02200, který kóduje enzym glutamát racemázu (*murI*, EC:5.1.1.3). Enzym konvertuje L-glutamát na D-glutamát vstupující do reakce s dříve zmíněnými substráty. Zjištěné navýšení míry exprese u tohoto genu bylo o 19 %. Studované dráhy jsou okrajovým metabolismem, který vstupuje do drah biosyntézy peptidoglykanu. Dráha může potenciálně kompetovat s dráhou biosyntézy HA ve smyslu zpracování UDP-*N*-acetylglukosaminu, který je prekurzorem pro UDP-*N*-acetylmuramát. Teoreticky čím větší množství UDP-GlcNAc odejde v podobě UDP-*N*-acetylmuramátu mimo dráhu syntézy HA, tím více se sníží produkce HA.

Tabulka 16: V tabulce jsou uvedeny geny zapojené v energetickém metabolismu kmene CO4A vztahující se k metabolickým drahám glutaminu a glutamátu. Uvedeno je označení genu, jeho funkce, zkratka a číselné označení k vyhledávání v databázi KEGG [130]. Jsou zde uvedeny poměry exprese genů u obou kmenů (CO4A a ATCC35246) normalizované k referenčnímu genu *gyrA*. Na obrázku 11 jsou geny vykresleny v kontextu dalších drah. ¹⁾ Poměr celkové exprese genu v ATCC35246 k celkové expresi *gyrA* (ATCC35246), ²⁾ Poměr celkové exprese genu v CO4A k celkové expresi *gyrA* (CO4A), ³⁾ Poměr celkových exprese obou kmenů normalizovaných ke *gyrA*: $CEG_{gyrA} (CO4A) / CEG_{gyrA} (ATCC)$

	Gen	CEG_{gyrA} (ATCC) ¹⁾	CEG_{gyrA} (CO4A) ²⁾	Poměr ³⁾	Změna exprese	Funkce genu
GLUTAMIN GLUTAMÁT	SeseC_00714	1,41	1,59	1,13	↑13%	<i>murD</i> ; UDP- <i>N</i> -acetylmuramoyl-L-alanyl-D-glutamate synthetase [KO:K01925] [EC:6.3.2.9]
	SeseC_02200	1,55	1,84	1,19	↑19%	<i>murI</i> ; glutamate racemase [KO:K01776] [EC:5.1.1.3]
	SeseC_02212	1,02	1,14	1,12	↑12%	<i>murC</i> ; UDP- <i>N</i> -acetylmuramate-L-alanine ligase [KO:K01924] [EC:6.3.2.8]



Obrázek 11: Na obrázku jsou vyznačeny sledované geny metabolických drah části metabolismu glutaminu a glutamátu v kontextu s dalšími drahami metabolismu (zeleně orámované). Vyznačeny jsou geny kmene CO4A se změnou exprese vyšší než 10 % v porovnání s referenčním kmenem ATCC35246. Zeleně podbarvené geny se vyskytují u *S. zooepidemicus* ATCC352246. Bíle podbarvené geny nejsou u *S. zooepidemicus* přítomny. [upraveno dle 130].

Metabolické dráhy aminocukrů a nukleotidů

Třetí velkou skupinou sledovaných genů energetického metabolismu jsou ty, zapojené do drah zpracování aminocukrů a nukleotidů (tab. 17).

Tabulka 17: V tabulce jsou uvedeny geny zapojené v energetickém metabolismu kmene CO4A vztahující se k drahám aminocukrů a nukleotidů. Uvedeno je označení genu, jeho funkce, zkratka a číselné označení k vyhledávání v databázi KEGG [130]. Jsou zde uvedeny poměry exprese genů u obou kmenů normalizovány k referenčnímu genu *gyrA*. Na obrázku 12 jsou geny vykresleny v kontextu dalších drah. ¹⁾ Poměr celkové exprese genu v ATCC35246 k celkové expresi *gyrA* (ATCC35246). ²⁾ Poměr celkové exprese genu v CO4A k celkové expresi *gyrA* (CO4A). ³⁾ Poměr celkových exprese obou kmenů normalizovaných ke *gyrA*: $CEG_{gyrA} (CO4A) / CEG_{gyrA} (ATCC)$

Gen	CEG_{gyrA} (ATCC) ¹⁾	CEG_{gyrA} (CO4A) ²⁾	Poměr ³⁾	Změna exprese	Funkce genu
Sesec_00232	1,20	1,35	1,13	↑13%	ugd; UDP-glucose 6-dehydrogenase HasB [KO:K00012] [EC:1.1.1.22]
Sesec_00233	2,35	3,14	1,33	↑33%	galU; UTP--glucose-1-phosphate uridylyltransferase HasC.2 [KO:K00963] [EC:2.7.7.9]
Sesec_00234	1,35	1,73	1,28	↑28%	glmU; Bifunctional protein glmU [KO:K04042] [EC:2.7.7.23 2.3.1.157]
Sesec_00235	1,83	2,21	1,21	↑21%	gpi; glucose-6-phosphate isomerase [KO:K01810] [EC:5.3.1.9]
Sesec_00255	2,55	2,66	1,04	↑4%	galU; UTP--glucose-1-phosphate uridylyltransferase HasC.2 [KO:K00963] [EC:2.7.7.9]
Sesec_00291	1,69	2,40	1,42	↑42%	nanE; N-acetylmannosamine-6-phosphate 2-epimerase [KO:K01788] [EC:5.1.3.9]
Sesec_00292	0,96	1,12	1,17	↑17%	ptsG; sugar phosphotransferase system (PTS), IIBC component [KO:K02778 K02779] [EC:2.7.1.199]
Sesec_00293	0,86	0,97	1,13	↑13%	PTS system, maltose and glucose-specific IIA-B-C component [KO:K02777] [EC:2.7.1.-]
Sesec_00296	3,25	5,98	1,84	↑84%	nanE; N-acetylmannosamine-6-phosphate 2-epimerase [KO:K01788] [EC:5.1.3.9]
Sesec_00302	0,87	1,09	1,25	↑25%	dapA; N-acetylneuraminic lyase [KO:K01639] [EC:4.1.3.3]
Sesec_00430	1,54	2,05	1,33	↑33%	scrK; fructokinase ScrK [KO:K00847] [EC:2.7.1.4]
Sesec_00431	1,06	1,30	1,23	↑23%	manA; mannose-6-phosphate isomerase Pmi [KO:K01809] [EC:5.3.1.8]
Sesec_00494	1,73	1,83	1,06	↑6%	manZ; mannose-Specific Phosphotransferase System Component IID ManZ [KO:K02796] [EC:2.7.1.191]
Sesec_00495	1,67	1,75	1,05	↑5%	manY; mannose-specific phosphotransferase system (PTS), IIC component [KO:K02795] [EC:2.7.1.191]
Sesec_00497	0,84	0,93	1,11	↑11%	manX; PTS system mannose-specific EIIAB component ManX [KO:K02793 K02794] [EC:2.7.1.191 2.7.1.191]
Sesec_00575	1,18	1,22	1,03	↑3%	nagA; N-acetylglucosamine-6-phosphate deacetylase [KO:K01443] [EC:3.5.1.25]
Sesec_00709	1,82	1,83	1,00	0%	glk; glucokinase [KO:K00845] [EC:2.7.1.2]
Sesec_00795	0,88	0,96	1,09	↑9%	nagB; glucosamine-6-phosphate isomerase [KO:K02564] [EC:3.5.99.6]
Sesec_00830	0,97	1,17	1,21	↑21%	murA; UDP-N-acetylglucosamine 1-carboxyvinyltransferase [KO:K00790] [EC:2.5.1.7]
Sesec_00848	0,77	0,89	1,16	↑16%	nagZ; glycosyl hydrolase family protein [KO:K01207] [EC:3.2.1.52]
Sesec_01008	1,15	1,48	1,29	↑29%	PTS system mannose/fructose permease IIC component [KO:K02795] [EC:2.7.1.191]

Sesec_01009	1,24	1,64	1,32	↑32%	sugar phosphotransferase system (PTS), mannose/fructose/sorbose family, IID component [KO:K02796] [EC:2.7.1.191]
Sesec_01011	0,65	0,75	1,14	↑14%	sugar phosphotransferase system (PTS), fructose family, IIA component [KO:K02793] [EC:2.7.1.191]
Sesec_01059	1,16	1,24	1,07	↑7%	glgC; glucose-1-phosphate adenylyltransferase [KO:K00975] [EC:2.7.7.27]
Sesec_01060	0,86	1,09	1,27	↑27%	glgC; glycogen biosynthesis protein GlgD [KO:K00975] [EC:2.7.7.27]
Sesec_01072	1,02	0,99	0,97	↓3%	murA; UDP-N-acetylglucosamine 1-carboxyvinyltransferase [KO:K00790] [EC:2.5.1.7]
Sesec_01134	1,27	1,36	1,07	↑7%	galE; UDP-glucose 4-epimerase [KO:K01784] [EC:5.1.3.2]
Sesec_01170	1,73	2,37	1,37	↑37%	glmS; glucosamine--fructose-6-phosphate aminotransferase [KO:K00820] [EC:2.6.1.16]
Sesec_01237	1,19	1,22	1,02	↑2%	manB; phosphomannomutase PgmA [KO:K01835] [EC:5.4.2.2]
Sesec_01392	1,21	1,39	1,15	↑15%	murB; UDP-N-acetylenolpyruvoylglucosamine reductase [KO:K00075] [EC:1.3.1.98]
Sesec_01545	1,40	1,51	1,08	↑8%	sugar phosphotransferase system (PTS), mannose/fructose family, IID component [KO:K02796] [EC:2.7.1.191]
Sesec_01547	1,14	1,19	1,05	↑5%	agaC; mannose permease IIC component ManY [KO:K02795] [EC:2.7.1.191]
Sesec_01549	0,58	0,62	1,07	↑7%	PTS system mannose/fructose family IIA component [KO:K02793] [EC:2.7.1.191]
Sesec_01569	2,31	2,77	1,20	↑20%	glmM; phosphoglucosamine mutase GlmM [KO:K03431] [EC:5.4.2.10]
Sesec_01989	2,15	2,79	1,30	↑30%	glmU; Bifunctional protein glmU [KO:K04042] [EC:2.7.7.23 2.3.1.157]
Sesec_02348	0,95	1,19	1,25	↑25%	galK; galactokinase [KO:K00849] [EC:2.7.1.6]
Sesec_02349	0,68	0,77	1,13	↑13%	galT; galactose-1-phosphate uridylyltransferase [KO:K00965] [EC:2.7.7.12]
Sesec_02350	0,57	0,69	1,21	↑21%	galE; UDP-glucose 4-epimerase [KO:K01784] [EC:5.1.3.2]
Sesec_02421	1,57	1,96	1,25	↑25%	gamP; glucose-specific phosphotransferase system (PTS), IIABC component [KO:K02777 K20107 K20108] [EC:2.7.1.- 2.7.1.208]

Výběr sledovaných genů metabolických drah aminocukrů a nukleotidů je zaměřen na metabolismus UDP cukrů (*N*-acetylglukosaminu, GlcNAc) jako prekurzoru biosyntézy kyseliny hyaluronové. *N*-acetylglukosamin může být získáván extracelulárně anebo může být produktem metabolismu. Gen SeseC_00848 (*nagZ*, EC:3.2.1.52) kóduje β -*N*-acetylhexosaminidázu, která katalyzuje hydrolýzu β -*N*-acetylglukosaminu (NAG) z neredukujícího konce různých glukokonjugátů a přivádí GlcNAc do metabolismu [144]. U tohoto genu bylo sledováno navýšení exprese o 16 %. Gen SeseC_00430 kóduje enzym fruktokinázu (*scrK*, EC:2.7.1.4), který mění fruktózu na fruktózu-6-P. Fruktóza přichází z drah metabolismu fruktózy, manózy a zejména sacharózy. Fruktóza-6-P může dále odcházet ke zpracování do kaskád glykolýzy a syntézy HA. Zde bylo pozorováno navýšení míry exprese o 33 % v porovnání s referenčním kmenem. Fruktóza-6-P může být jako meziprodukt konvertována na glukózu-6-P. Konverzi katalyzuje enzym glukóza-6-fosfát izomeráza (*gpi*, EC:5.3.1.9), která je kódována genem SeseC_00235. Navýšení jeho relativní míry exprese v porovnání s kmenem ATCC35246 bylo stanoveno na 21 %. Gen je součástí jak drah glykolýzy, tak biosyntézy HA. Fruktóza-6-P může být odvedena jako substrát i do drah metabolismu manózy. Změnu na manózu-6-P katalyzuje enzym manóza-6-fosfát izomeráza, která je kódována genem SeseC_00431 (*manA*, EC:5.3.1.8). Navýšení míry exprese bylo o 23 % vyšší než u referenčního kmene. Glukózu-6-P může buňka získat hydrolýzou sacharózy, nebo jako extracelulární substrát a k tomu je využíván gen SeseC_00292 kódující část fosfotransferázového systému PTS (*ptsG*, EC:2.7.1.199). U tohoto genu bylo sledováno navýšení exprese o 17 %.

V metabolismus aminocukrů a nukleotidů je zajímavý gen *nanE* (SeseC_00296), který katalyzuje obousměrnou reakci měnící GlcNAc-6-P na ManNAc-6-P a může tak odvádět prekurzor syntézy HA do drah metabolismu manózy. Gen SeseC_00296 (*nanE*) má v genomu homologní sekvenci SeseC_00291. Obě kopie vykazovaly vysokou míru exprese v porovnání s referenčním kmenem ATCC35246 (SeseC_00291 měl nárůst exprese o 42 % a SeseC_00296 dokonce o 84 %). V sekvenci SeseC_00296 byla identifikována inserce (kap. 5.2 tab. 11), která posunula a především zkrátila čtecí rámec. Důsledkem mohl být pozorovaný nárůst exprese genů. Gen *nanE* je klíčovým enzymem také u biosyntézy *N*-acetylneuraminové kyseliny (sialové kyseliny) [140, 141, 142]. Geny SeseC_00291 (*nanE*), SeseC_00292 (*ptsG*), SeseC_00293 (PTS systém), SeseC_00296 (*nanE*) a SeseC_00302 (*dapA*) jsou zřejmě součástí operonu metabolické dráhy biosyntézy kyseliny neuraminové, kde se jako iniciační molekula uplatňuje *N*-acetylglukosamin [141]. Z dat je patrné, že se zvýšila jejich exprese oproti ATCC35246 v rozmezí 13-84 %. Může to naznačovat právě nedostatek enzymu *N*-acetylmannosamin-6-fosfát-2-epimerázy (popř. omezení jeho funkčnosti v porovnání s ATCC35246), potažmo depresi této metabolické dráhy, která spotřebová u *S. zooepidemicus* CO4A prekurzor syntézy HA – GlcNAc. Pokud tedy jedna z cest zpracování GlcNAc není funkční, může to vést k posílení zbylých drah a to by mohlo vysvětlovat i zvýšenou expresi genů, které tento cukr zpracovávají.

Pokud se podíváme na geny více zapojené do metabolických drah UDP-GlcNAc, můžeme začít u velmi běžného metabolitu fruktózy-6-P. Fruktóza-6-P přichází z cyklu glykolýzy a glukoneogeneze. Pomocí enzymu glukózamino-fruktóza-6-fosfát aminotransferáza (*glmS*, EC:2.6.1.16), který je kódován genem SeseC_01170, je konvertována na glukózamin-6-P. Míra exprese byla o 37 % vyšší než u kmene ATCC35246. Kaskáda pokračuje genem SeseC_01569, který kóduje enzym fosfoglukózamino mutázu (*glmM*, EC:5.4.2.10). Mění glukózamin-6-P na

D-glukózamin-1-P. Míra exprese byla stanovena na hodnotu o 20 % vyšší. D-glukózamin-1-P je následně konvertován na *N*-acetylglukózamin-1-P za katalýzy C-terminální domény bifunkčního enzymu glukosamin-1-fosfát acetyltransferázy (*glmU*, EC:2.3.1.157; EC:2.7.7.23), který je kódován genem *Sesec_01989* (*Sesec_00234*). Míra jeho exprese byla navýšena o 30 % v porovnání s referenčním kmenem. Tento enzym (*glmU*) spolu s předchozími dvěma enzymy (*glmS*, *glmM*) se přímo podílí na syntéze prekurzoru pro produkci HA. N-terminální doména stejného enzymu (*glmU*) katalyzuje změnu *N*-acetylglukosaminu-1-P na aktivovanou formu UDP-*N*-acetylglukosamin [145]. Zde bylo zaznamenáno navýšení míry exprese o 28 %.

Geny zapojené přímo v metabolických reakcích vedoucích ke vzniku *N*-acetylglukosaminu měly vyšší změnu míry exprese, než v drahách okrajových, nebo tam kde byl metabolit odváděn do jiných cyklů. Můžeme říci, že u kmene CO4A měly klíčové geny signifikantně vyšší změnu exprese a tudíž je pravděpodobné, že se projevila i na úrovni proteomu. Můžeme pozorovat kaskádu genů s vyšší mírou exprese směřující od glukózy-6-P a fruktózy (pocházející ze sacharózy, jako uhlíkového zdroje v kultivaci) až k UDP-*N*-acetylglukosaminu (prekurzoru biosyntézy HA). Zajímavým genem byl i gen *nanE*, který je klíčovým jak pro energetický metabolismus buňky, tak pro biosyntézu neuraminové kyseliny [140].

5.3.2 Geny metabolických drah syntézy buněčné stěny

Další skupinou metabolických drah, které mohou souviset i se syntézou kyseliny hyaluronové, jsou geny drah syntézy buněčné stěny. Byly sem zařazeny geny metabolismu syntézy peptidoglykanu (tab. 18) a glycerolipidů (tab. 19).

Peptidoglykan

Tabulka 18: V tabulce jsou uvedeny geny zapojené do metabolismu syntézy buněčné stěny kmene CO4A vztahující se k metabolickým drahám syntézy peptidoglykanu. Uvedeno je označení genu, jeho funkce, zkratka a číselné označení k vyhledávání v databázi KEGG [130]. Jsou zde uvedeny poměry exprese genů u obou kmenů normalizované k referenčnímu genu *gyrA*. Na obrázku 13 jsou geny vykresleny v kontextu dalších metabolických drah. ¹⁾ Poměr celkové exprese genu v ATCC35246 k celkové expresi *gyrA* (ATCC35246). ²⁾ Poměr celkové exprese genu v CO4A k celkové expresi *gyrA* (CO4A). ³⁾ Poměr celkových exprese obou kmenů normalizovaných ke *gyrA*: $CEG_{gyrA} (CO4A) / CEG_{gyrA} (ATCC)$.

Gen	CEG_{gyrA} (ATCC) ¹⁾	CEG_{gyrA} (CO4A) ²⁾	Poměr ³⁾	Změna exprese	Funkce genu
Sesec_00102	0,94	1,19	1,27	↑27%	pbp; multimodular transpeptidase-transglycosylase [KO:K03693] [ClassA PBP]
Sesec_00175	1,77	2,07	1,17	↑17%	pbp2A; multimodular transpeptidase-transglycosylase [KO:K12555] [EC:2.4.1.129 3.4.16.4]
Sesec_00468	1,25	1,45	1,16	↑16%	vanY; D-alanyl-D-alanine carboxypeptidase [KO:K07260] [EC:3.4.17.14]
Sesec_00714	1,41	1,59	1,13	↑13%	murD; UDP-N-acetylmuramoyl-L-alanyl-D-glutamate synthetase [KO:K01925] [EC:6.3.2.9]
Sesec_00715	1,14	1,26	1,10	↑10%	murG; undecaprenyldiphospho-muramoylpentapeptide beta-N-acetylglucosaminyltransferase [KO:K02563] [EC:2.4.1.227]
Sesec_00773	0,77	0,92	1,20	↑20%	pbp2B; penicillin-binding protein 2b [KO:K00687]
Sesec_00775	1,04	1,11	1,07	↑7%	ddl; D-alanine--D-alanine ligase Ddl [KO:K01921] [EC:6.3.2.4]
Sesec_00776	0,78	0,91	1,15	↑15%	murF; UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase [KO:K01929] [EC:6.3.2.10]
Sesec_00830	0,97	1,17	1,21	↑21%	murA; UDP-N-acetylglucosamine 1-carboxyvinyltransferase [KO:K00790] [EC:2.5.1.7]
Sesec_01072	1,02	0,99	0,97	↓3%	murA; UDP-N-acetylglucosamine 1-carboxyvinyltransferase [KO:K00790] [EC:2.5.1.7]
Sesec_01341	0,90	1,08	1,20	↑20%	UDP-N-acetylmuramoylpentapeptide-lysine N(6)-alanyltransferase MurM [KO:K05363] [EC:2.3.2.10]
Sesec_01392	1,21	1,39	1,15	↑15%	murB; UDP-N-acetylenolpyruvoylglucosamine reductase [KO:K00075] [EC:1.3.1.98]
Sesec_01572	0,78	0,85	1,09	↑9%	murE; Mur ligase family protein [KO:K23393] [EC:6.3.5.13]
Sesec_01573	1,32	1,49	1,13	↑13%	conserved hypothetical protein [KO:K07009] [EC:6.3.5.13]
Sesec_01704	0,68	0,75	1,10	↑10%	peptidoglycan pentaglycine interpeptide biosynthesis protein [KO:K05363] [EC:2.3.2.10]

Sesec_01705	0,70	0,75	1,06	↑6%	peptidoglycan branched peptide synthesis protein [KO:K12554] [EC:2.3.2.-]
Sesec_01943	1,21	1,35	1,12	↑12%	dgkA; diacylglycerol kinase DgkA [KO:K00887] [EC:2.7.1.66]
Sesec_02105	1,25	1,29	1,04	↑4%	mrcA; penicillin-binding protein 1A [KO:K05366] [EC:2.4.1.129 3.4.16.4]
Sesec_02125	2,26	2,94	1,30	↑30%	mraY; phospho-N-acetylmuramoyl-pentapeptide-transferase [KO:K01000] [EC:2.7.8.13]
Sesec_02126	0,79	0,86	1,09	↑9%	ftsI; cell division protein FtsI [KO:K12556]
Sesec_02166	0,60	0,69	1,15	↑15%	murE; UDP-N-acetylmuramoylalanyl-D-glutamate--L-lysine ligase [KO:K05362] [EC:6.3.2.7]
Sesec_02180	1,48	1,78	1,20	↑20%	PAP2 superfamily protein [KO:K19302] [EC:3.6.1.27]
Sesec_02212	1,02	1,14	1,12	↑12%	murC; UDP-N-acetylmuramate--L-alanine ligase [KO:K01924] [EC:6.3.2.8]
Sesec_02261	0,70	0,76	1,09	↑9%	dacC; D-alanyl-D-alanine carboxypeptidase precursor DacA.1 [KO:K07258] [EC:3.4.16.4]
Sesec_02262	0,51	0,61	1,20	↑20%	dacC; D-alanyl-D-alanine carboxypeptidase [KO:K07258] [EC:3.4.16.4]
Sesec_02273	1,32	1,70	1,29	↑29%	bacA; undecaprenol kinase (bacitracin resistance protein) [KO:K06153] [EC:3.6.1.27]

Metabolická dráha syntézy peptidoglykanu je přímo kompetující k syntéze HA. Jeho produkce vychází z aktivovaného prekurzoru syntézy HA UDP-*N*-acetylglukosaminu (UDP-GlcNAc). UDP-GlcNAc pochází z drah metabolismu aminocukrů a nukleotidů. Syntéza peptidoglykanu začíná konverzí UDP-GlcNAc na UDP-GlcNAc-enolpyruvát, která je katalyzována enzymem UDP-*N*-acetylglukosamin-1-carboxyvinyltransferázou, která je kódována genem *Sesec_00830* (*murA*, EC:2.5.1.7). Navýšení jeho exprese v kmeni CO4A bylo stanoveno na 21 % v porovnání s kmenem ATCC35246. Dále byl předchozí metabolit změněn na UDP-*N*-acetylmuramát za působení enzymu UDP-*N*-acetylenolpyruvoylglukosamin reductázy (*murB*, EC:1.3.1.98), která je kódována genem *Sesec_01392*. Míra jeho exprese byla o 15 % vyšší než u referenčního kmene.

Dalšími kroky metabolismu se postupně dostaneme až k peptidoglykanu. Jednotlivé enzymy a geny po sobě jdoucích metabolických reakcí s jejich změnami míry exprese jsou uvedeny ve zkrácené formě zde. Gen *Sesec_02212* kóduje *murC* (EC:6.3.2), exprese byla vyšší o 12 %. Gen *Sesec_00714* kóduje *murD* (EC:6.3.2.9), exprese byla vyšší o 13 %. Gen *Sesec_02166* kóduje *murE* (EC:6.3.2.7), exprese byla vyšší o 15 %. Gen *Sesec_00776* kóduje *murF* (EC:6.3.2.10), exprese byla vyšší o 15 %. Gen *Sesec_02125* kóduje *mraY* (EC:2.7.8.13), exprese byla vyšší o 30 %. Gen *Sesec_00715* kóduje *murG* (EC:2.4.1.227), exprese byla vyšší o 10 %. Gen *Sesec_01341* kóduje *murM* (EC:2.3.2.10), exprese byla vyšší o 20 %. A enzym končící celou konverzi na peptidoglykan kóduje gen *Sesec_00468 VanY* (EC:3.4.17.14), jeho exprese byla vyšší o 16 % v porovnání s referenčním kmenem ATCC35246.

Gen *Sesec_02125* kódující *mraY* (EC:2.7.8.13) měl nejvyšší změřenou změnu míry exprese (30 %). U ostatních genů se jejich změna exprese pohybovala mezi 10-20 %. Takto vysoká změna exprese mohla být způsobena zapojením do dráhy zpracování undecaprenolu, kde byly zapojeny i poslední dva sledované geny *Sesec_01943* kódující *dgkA* (EC:2.7.1.66) a *Sesec_02180* kódující *PAP2* (EC:3.6.1.27). Změny jejich exprese byly 12 % (*dgkA*) a 20 % (*PAP2*).

Glycerolipidy

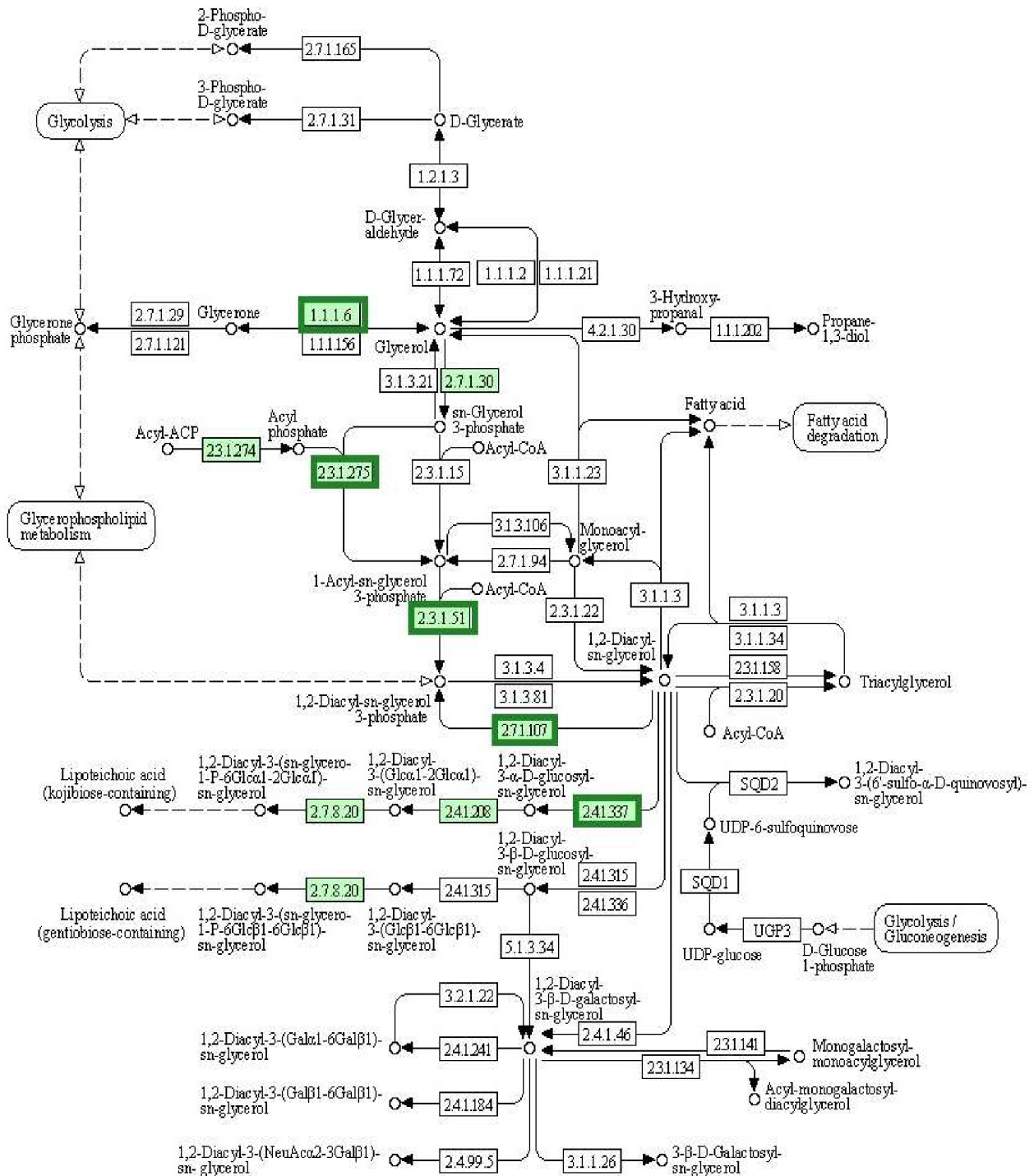
U této skupiny vykazovalo 5 genů změnu míry exprese vyšší než 10 % v porovnání s referenčním kmenem ATCC35246 (obr. 14). Jde o kaskádu genů vedoucích k syntéze lipoteichoové kyseliny, která je součástí buněčné stěny. Kyselina lipoteichoová (LTA) je amfifilní molekula² vyskytující se u Gram-pozitivních bakterií a regulátor autolytických enzymů buněčné stěny (muramidáz). Patří také k virulentním faktorům streptokoků [146].

Tabulka 19: V tabulce jsou uvedeny geny zapojené v metabolismu syntézy buněčné stěny kmene CO4A vztahující se k metabolickým drahám glycerolipidů. Uvedeno je označení genu, jeho funkce, zkratka a číselné označení k vyhledávání v databázi KEGG [130]. Jsou zde uvedeny poměry exprese genů u obou kmenů normalizované k referenčnímu genu gyrA. Na obrázku 14 jsou geny vykresleny v kontextu dalších metabolických drah.
¹⁾ Poměr celkové exprese genu v ATCC35246 k celkové expresi gyrA (ATCC35246), ²⁾ Poměr celkové exprese genu v CO4A k celkové expresi gyrA (CO4A), ³⁾ Poměr celkových exprese obou kmenů normalizovaných ke gyrA: $CEG_{gyrA} (CO4A) / CEG_{gyrA} (ATCC)$

Gen	CEG_{gyrA} (ATCC) ¹⁾	CEG_{gyrA} (CO4A) ²⁾	Poměr ³⁾	Změna exprese	Funkce genu
Sesec_00026	1,45	1,34	0,92	↓8%	plsX; fatty acid/phospholipid synthesis protein [KO:K03621] [EC:2.3.1.274]
Sesec_00786	0,90	1,09	1,20	↑20%	plsC; acyltransferase [KO:K00655] [EC:2.3.1.51]
Sesec_01054	0,70	0,94	1,33	↑33%	conserved hypothetical protein [KO:K07029] [EC:2.7.1.107]
Sesec_01147	0,78	0,78	1,00	0%	sulfatase [KO:K19005] [EC:2.7.8.20]
Sesec_01505	0,83	1,02	1,23	↑23%	plsY; putative glycerol-3-phosphate acyltransferase PlsY [KO:K08591] [EC:2.3.1.275]
Sesec_01819	1,65	2,09	1,27	↑27%	ugtP; glycosyltransferase [KO:K19002] [EC:2.4.1.337]
Sesec_01820	1,32	1,44	1,09	↑9%	glycosyltransferase [KO:K13677] [EC:2.4.1.208]
Sesec_02145	0,59	0,64	1,08	↑8%	glpK; glycerol kinase [KO:K00864] [EC:2.7.1.30]
Sesec_02453	0,75	0,93	1,24	↑24%	gldA; glycerol dehydrogenase [KO:K00005] [EC:1.1.1.6]

² Amfifilní molekuly obsahující nepolární hydrofobní oblast a polární hydrofilní oblast. Ve vodném roztoku mohou tvořit různé struktury jako jsou micely, vesikuly a tubuly [147].

GLYCEROLIPID METABOLISM



Obrázek 14: Metabolické dráhy glycerolipidů jsou spojeny s energetickým metabolismem především přes dráhy glykolýzy. Byly sledovány geny centrálních drah transformace glycerolu. Výstupem těchto drah jsou například triacylglyceroly a další cukerné alkoholy, ale také kyselina lipoteichoová. Kyselina teichoová je součástí buněčné stěny, ale ne v takové míře jako předcházející peptidoglykan. Dráha není výrazně kompetitivní s dráhou syntézy HA. Mohla by ovlivňovat energetický metabolismus prostřednictvím drah degradace lipidů (nebylo testováno). Vyznačeny jsou geny (zeleně orámováni) se změnou exprese vyšší než 10 %. Zeleně podbarvené geny se vyskytují u *S. zooepidemicus* ATCC35246. Bíle podbarvené geny nejsou u *S. zooepidemicus* přítomny. [upraveno dle 130]

Prvním sledovaným genem byl *Sesec_02453*, který kóduje enzym glycerol dehydrogenázu (*gldA*, EC:1.1.1.6). Katalyzuje přeměnu glyceronu na glycerol. Glyceron může také odcházet do metabolických drah glykolýzy. Míra změny exprese byla o 24 % vyšší v porovnání s referencí. Dalším genem byl *Sesec_01505* kódující enzym acyltransferázu (*plsY*, EC:2.3.1.275), kdy byla připojena ke glycerolu-3-P acylová skupina a vznikl 1-acylglycerol-3-P. Míra změny jeho exprese byla 23 %. Další acyltransferázu (*plsC*, EC:2.3.1.51) kóduje gen *Sesec_00786* a z předchozího metabolitu vzniká 1,2-diacylglycerol-3-P. Změna míry exprese byla 20 %. Jedním z enzymů konečné fáze vzniku lipoteichoové kyseliny je glycosyltransferáza (*ugtP*, EC:2.4.1.337), která je kódována genem *Sesec_01819*. Katalyzuje první krok ke vzniku lipoteichoové kyseliny z 1,2-diacylglycerolu. Míra exprese tohoto genu byla o 27 % vyšší než u referenčního kmene. Jako poslední v řadě byl sledován gen *Sesec_01054* (EC:2.7.1.107), který je zařazen jako katalyzátor změny 1,2-diacylglycerolu zpět na 1,2-diacylglycerol-3-P. Tato zpětná reakce měla míru exprese o 33 % vyšší než reference.

5.3.3 Geny replikace DNA a degradace RNA

Do třetí části byly zařazeny metabolické dráhy související s buněčným dělením. Jsou to dráhy replikace DNA (tab. 20) a metabolické dráhy degradace RNA (tab. 21). U drah replikace DNA můžeme předpokládat, že rychlost replikace je přímo úměrná dělení buněk. Je tedy teoreticky možné, ji vztáhnou k rychlosti nárůstu biomasy.

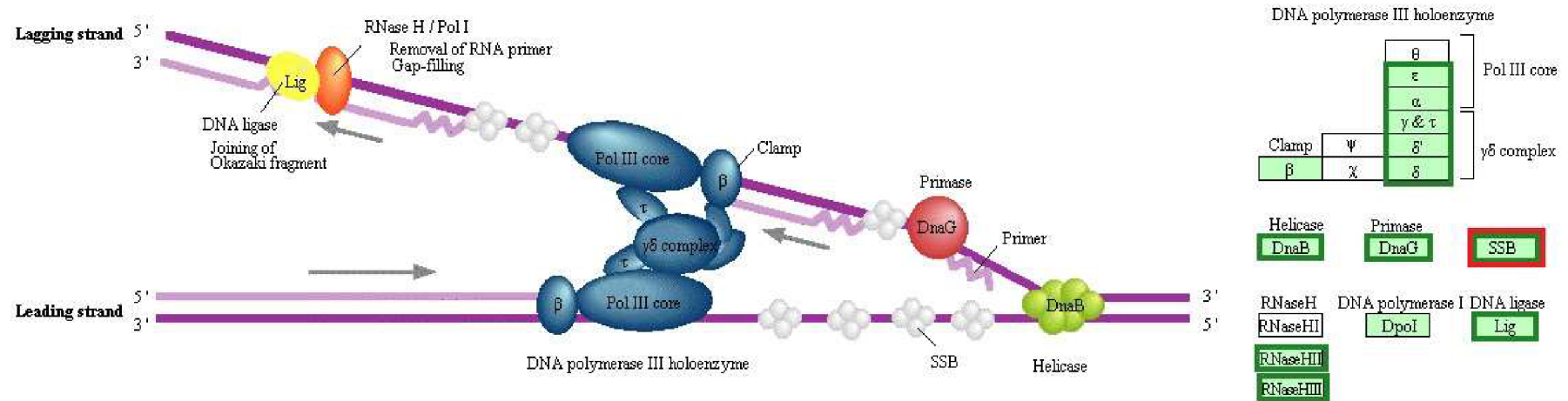
Replikace DNA

Tabulka 20: V tabulce jsou uvedeny geny zapojené v metabolických drahách replikace DNA u kmene CO4A. Uvedeno je označení genu, jeho funkce, zkratka a číselné označení k vyhledávání v databázi KEGG [130]. Na obrázku 15 jsou geny uvedené ve schématu replikace DNA. ¹⁾ Poměr celkové exprese genu v ATCC35246 k celkové expresi gyrA (ATCC35246). ²⁾ Poměr celkové exprese genu v CO4A k celkové expresi gyrA (CO4A). ³⁾ Poměr celkových exprese obou kmenů normalizovaných ke gyrA: $CEG_{gyrA} (CO4A) / CEG_{gyrA} (ATCC)$

Gen	CEG _{gyrA} (ATCC) ³⁾	CEG _{gyrA} (CO4A) ⁴⁾	Poměr ⁵⁾	Změna exprese	Funkce genu
Sesec_00003	1,17	1,22	1,04	↑4%	dnaN; DNA polymerase III subunit beta [KO:K02338] [EC:2.7.7.7]
Sesec_00200	1,06	1,08	1,02	↑2%	polA; DNA polymerase I [KO:K02335] [EC:2.7.7.7]
Sesec_00352	0,59	0,77	1,29	↑29%	DNA polymerase III subunit epsilon [KO:K02342] [EC:2.7.7.7]
Sesec_00390	1,68	1,87	1,11	↑11%	rnhC; ribonuclease HIII [KO:K03471] [EC:3.1.26.4]
Sesec_00406	2,29	1,84	0,80	↓20%	ssb; single strand binding protein (SSB) [KO:K03111]
Sesec_00789	0,61	0,70	1,15	↑15%	holA; DNA polymerase III delta subunit [KO:K02340] [EC:2.7.7.7]
Sesec_00817	0,70	0,80	1,14	↑14%	dnaX; DNA polymerase III subunits gamma and tau [KO:K02343] [EC:2.7.7.7]
Sesec_01052	1,46	1,97	1,35	↑35%	ligA; NAD-dependent DNA ligase LigA [KO:K01972] [EC:6.5.1.2]
Sesec_01116	1,08	1,39	1,29	↑29%	dnaG; DNA primase [KO:K02316] [EC:2.7.7.101]
Sesec_01166	1,32	1,46	1,10	↑10%	dnaE; DNA polymerase III alpha subunit [KO:K02337] [EC:2.7.7.7]
Sesec_01273	0,51	0,58	1,13	↑13%	rnhB; ribonuclease HII [KO:K03470] [EC:3.1.26.4]
Sesec_02156	0,60	0,67	1,12	↑12%	holB; DNA polymerase III, delta' subunit [KO:K02341] [EC:2.7.7.7]
Sesec_02310	0,52	0,71	1,36	↑36%	single-strand binding protein ssB [KO:K03111]
Sesec_02395	1,06	1,32	1,24	↑24%	polC; DNA polymerase III PolC-type [KO:K03763] [EC:2.7.7.7]
Sesec_02489	1,51	2,11	1,39	↑39%	ssb; single stranded DNA-binding protein [KO:K03111]
Sesec_02572	1,61	1,99	1,24	↑24%	dnaB; replicative DNA helicase [KO:K02314] [EC:3.6.4.12]

DNA REPLICATION

Replication complex (Bacteria)



70

Obrázek 15: Schéma komplexu replikace DNA v bakteriální buňce. Zeleně podbarvené geny se vyskytují u *S. zooepidemicus*. Bíle podbarvené geny nejsou u *S. zooepidemicus* přítomny. Sledované geny jsou označený zeleným rámečkem. [převzato ze 124]

DNA polymeráza III je složena z několika podjednotek a u některých byla analyzována jejich exprese (obr. 15). Sledovány byly i podjednotky RNázy H, DNA polymerázy I, DNA ligázy, helicázy a dalších zapojených enzymů. Na schématu jsou vyznačeny části, u kterých byl zjištěn vyšší než 10 % nárůst, nebo pokles míry exprese v porovnání s referenčním kmenem ATCC35246. Zde jsme narazili na gen, který měl o 20 % nižší míru exprese než je tomu u reference (Sesec_00406, *SSB* – single strand binding protein), zároveň dva geny stejné skupiny vazebných proteinů *SSB* mají míru exprese výrazně vyšší (gen Sesec_02310 má míru exprese vyšší o 36 %, gen Sesec_02489 má míru exprese vyšší o 39 %). *SSB* proteiny mají významnou roli při replikaci DNA a jejich opravách [148]. Využívají se také v biochemických aplikacích, jako je například zvýšení specifity a efektivity PCR reakcí [148]. Gen Sesec_00352 kódující DNA polymerázu III podjednotku epsilon, měl míru exprese o 29 % vyšší. Gen Sesec_01052 kódující NAD-dependentní DNA ligázu (*ligA*), měl míru exprese o 35 % vyšší atd. Tato skupina genů je nezbytná pro přežití a množení prokaryotických buněk. Jejich exprese v buňce by měla probíhat stabilně a v dostatečné míře. Navýšení míry exprese by mohlo znamenat kratší cyklus buněčného dělení a rychlejší nárůst biomasy při kultivaci u kmene CO4A. Na množství produkované HA neměla replikace DNA

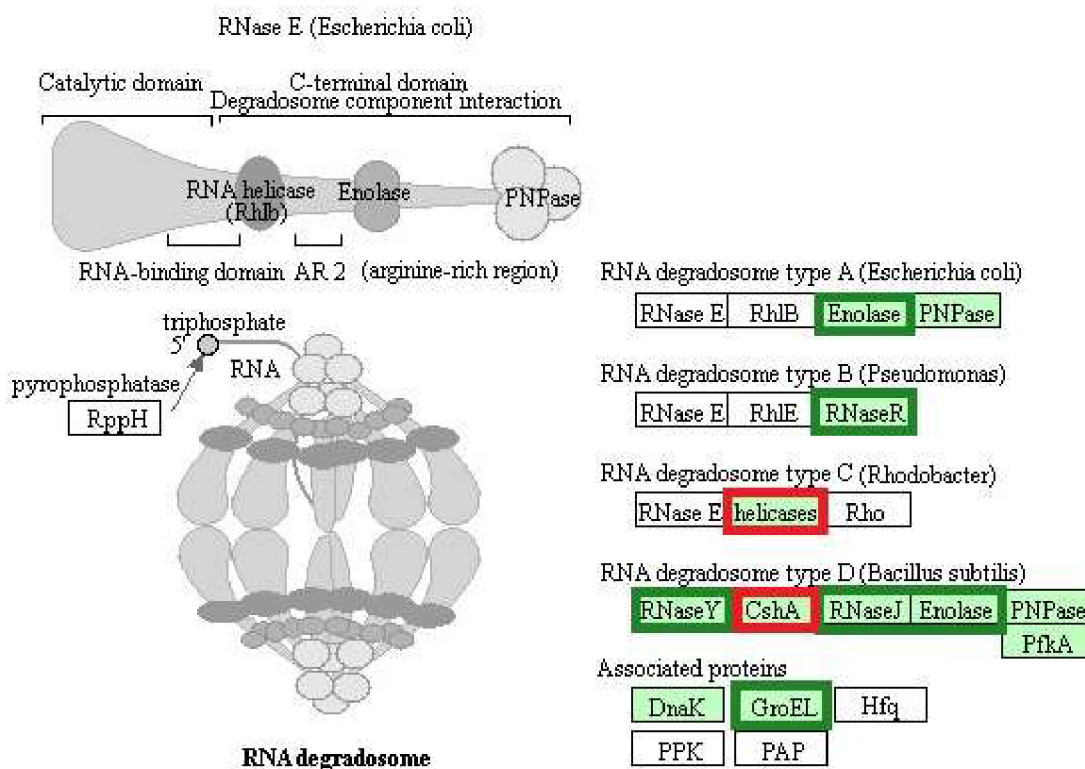
zřejmě signifikantní vliv. V genech kódujících tuto dráhu nebyla nalezena žádná mutace. Mutace v těchto genech by nejspíše měla fatální následky na život buněk.

Degradace RNA

Tabulka 21: V tabulce jsou uvedeny geny zapojené v metabolických drahách degradace RNA kmene CO4A. Uvedeno je označení genu, jeho funkce, zkratka a číselné označení k vyhledávání v databázi KEGG [130]. Jsou zde uvedeny poměry exprese genů u obou kmenů normalizované k referenčnímu genu *gyrA*. Na obrázku 16 jsou geny vykresleny v kontextu jednotky degradosomu RNA. ¹⁾ Poměr celkové exprese genu v ATCC35246 k celkové expresi *gyrA* (ATCC35246). ²⁾ Poměr celkové exprese genu v CO4A k celkové expresi *gyrA* (CO4A). ³⁾ Poměr celkových exprese obou kmenů normalizovaných ke *gyrA*: $CEG_{gyrA} (CO4A) / CEG_{gyrA} (ATCC)$

Gen	CEG_{gyrA} (ATCC) ¹⁾	CEG_{gyrA} (CO4A) ²⁾	Poměr ³⁾	Změna exprese	Funkce genu
Sesec_00156	1,54	1,71	1,12	↑12%	groEL; 60 kDa chaperonin [KO:K04077]
Sesec_00340	1,08	0,98	0,90	↓10%	mj; metallo-beta-lactamase family protein, Zn-dependent hydrolase [KO:K12574] [EC:3.1.-.-]
Sesec_00473	2,36	2,29	0,97	↓3%	dnaK; chaperone protein DnaK (heat shock protein 70) [KO:K04043]
Sesec_00780	1,29	1,11	0,86	↓14%	deaD; DEAD box helicase family protein [KO:K05592] [EC:3.6.4.13]
Sesec_01044	2,47	3,02	1,22	↑22%	eno; phosphopyruvate hydratase [KO:K01689] [EC:4.2.1.11]
Sesec_01167	1,72	1,62	0,94	↓6%	pfk; 6-phosphofruktokinase [KO:K00850] [EC:2.7.1.11]
Sesec_01585	1,07	1,25	1,17	↑17%	metallo-beta-lactamase superfamily protein [KO:K12574] [EC:3.1.-.-]
Sesec_01836	1,31	1,54	1,18	↑18%	rrr; exoribonuclease R [KO:K12573] [EC:3.1.13.1]
Sesec_02095	1,25	1,27	1,02	↑2%	phosphohydrolase [KO:K18682] [EC:3.1.-.-]
Sesec_02384	1,08	1,14	1,06	↑6%	pnp; polyribonucleotide nucleotidyltransferase [KO:K00962] [EC:2.7.7.8]

Bacterial RNA degradation



Obrázek 16: Schéma znázorňuje jednotku RNA degradosomu, kde dochází k rozkladu nepotřebných, chybných, nebo nebezpečných sekvencí RNA. V obrázku jsou vyznačeny geny se změnou míry exprese vyšší než 10 % (zelené orámování). Zeleně podbarvené geny se vyskytují u *S. zooepidemicus*. Bile podbarvené geny nejsou u *S. zooepidemicus* přítomny. [upraveno dle 130]

Degradace RNA slouží v buňce jako regulační a kontrolní mechanismus. Odstraňuje se poškozená nebo nepotřebná RNA a slouží i jako obrana proti RNA virům. U bakteriálních buněk probíhá degradace na degradozomu [149]. Ten obsahuje exo- a endonukleázy a fosforylázy, které společně štěpí RNA. Signálem k degradaci RNA je polyadenylace 3' konce [149].

V této skupině se vyskytly geny, které měly změnu míry exprese o více než 10 % menší než u referenčního kmene. Gen Sese_00780 kódující helikázu (*deaD*, EC:3.6.4.13) měl o 14 % nižší míru exprese. Helikáza napomáhá k rozplétání vlákna RNA při její degradaci. Nižší míru exprese o 10 % měl gen Sese_00340 kódující Zn-dependentní hydrolázu (EC:3.1.-.-). Další enzymy podílející se na štěpení RNA byly fosfopyruvát hydroláza (Sese_01044, *eno*, EC:4.2.1.11) a exoribonukleáza R (Sese_01836, *mr*, EC:3.1.13.1). Hydroláza měla o 22 % vyšší míru exprese a exonukleáza o 18 % v porovnání s referenčním kmenem. Fosfopyruvát hydroláza (*eno*) je přímo zapojena i do metabolických drah glykolýzy jako jeden z centrálních enzymů. Z dat genové exprese u metabolických drah degradace RNA nelze jednoznačně určit, jaký vliv dráha může mít na produkci HA.

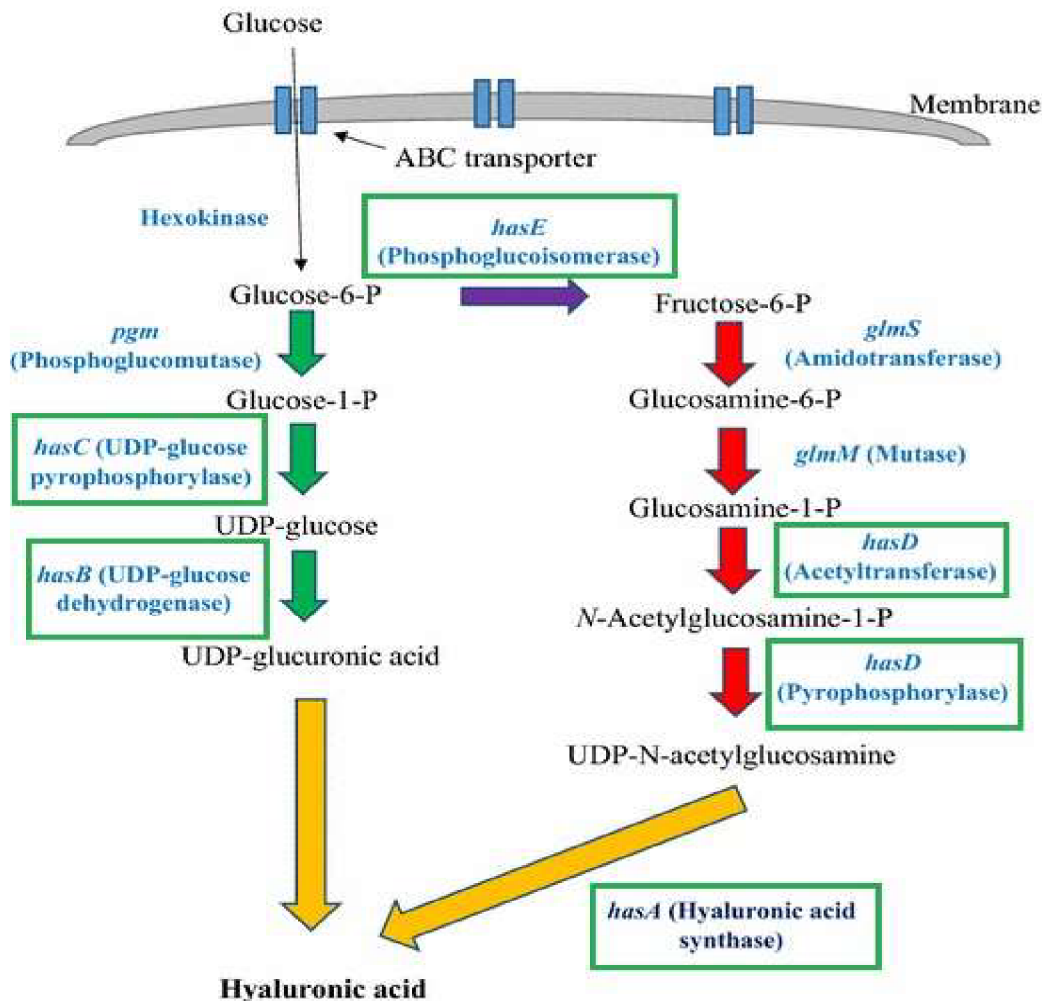
5.3.4 Geny *has* operonu a vybraných metabolických drah

Do posledního oddílu byly zařazeny geny metabolických drah, které souvisí přímo, nebo nepřímo s produkcí HA, nebo ji mohou svou aktivitou ovlivňovat (tab. 23). Byly sem zařazeny geny *has* operonu (tab. 22), tedy přímo dráhy syntézy HA.

has operon

Tabulka 22: V tabulce jsou uvedeny geny zapojené v metabolických drahách produkce kyseliny hyaluronové (geny *has* operonu). Uvedeno je označení genu, jeho funkce, zkratka a číselné označení k vyhledávání v databázi KEGG [130]. Jsou zde uvedeny poměry exprese genů u obou kmenů normalizované k referenčnímu genu *gyrA*. Schéma genů *has* operonu v kontextu biosyntézy HA je uvedeno na obrázku 17. ¹⁾ Poměr celkové exprese genu v ATCC35246 k celkové expresi *gyrA* (ATCC35246). ²⁾ Poměr celkové exprese genu v CO4A k celkové expresi *gyrA* (CO4A). ³⁾ Poměr celkových exprese obou kmenů normalizovaných ke *gyrA*: $CEG_{gyrA} (CO4A) / CEG_{gyrA} (ATCC)$

	Gen	CEG_{gyrA} (ATCC) ¹⁾	CEG_{gyrA} (CO4A) ²⁾	Poměr ³⁾	Změna exprese	Funkce genu
HAS OPERON	Sesec_00231hasA	2,56	4,77	1,86	↑86%	hasA; hyaluronansynthase HasA [KO:K00752][EC:2.4.1.212]
	Sesec_00232ugd	1,20	1,35	1,13	↑13%	ugd; UDP-glucose 6-dehydrogenase HasB [KO:K00012] [EC:1.1.1.22]
	Sesec_00233galU	2,35	3,14	1,33	↑33%	galU; UTP--glucose-1-phosphate uridylyltransferase HasC.2 [KO:K00963] [EC:2.7.7.9]
	Sesec_00234glmU	1,35	1,73	1,28	↑28%	glmU; Bifunctional protein glmU [KO:K04042] [EC:2.7.7.23 2.3.1.157]
	Sesec_00235gpi	1,83	2,21	1,21	↑21%	gpi; glucose-6-phosphate isomerase [KO:K01810] [EC:5.3.1.9]



Obrázek 17: Analyzované geny *has* operonu se změnou míry exprese vyšší, než 10 % jsou zeleně orámované. [upraveno dle 150]

has operon kóduje centrální enzymy biosyntézy kyseliny hyaluronové a je regulován silným promotorem. Očekávali jsme, že vyšší výtěžek HA u kmene CO4A v porovnání s referencí by mohl mít původ ve zvýšené expresi genů *has* operonu. Naše výsledky skutečně ukázaly, že exprese *has* operonu u CO4A je vyšší než u ATCC35246. Nejvyšší změna genové exprese byla zaznamenána u genu SeseC_00231 kódujícího hyaluronan syntázu (*hasA*, EC:2.4.1.212). Enzym je klíčový pro spojení dvou prekurzorů a vzniku HA. Navýšení bylo o 86 % v porovnání s referenčním kmenem. U ostatních genů nebylo zaznamenáno tak výrazné navýšení míry exprese (SeseC_00232 o 13 %; SeseC_00233 o 33 %; SeseC_00234 o 28 %; SeseC_00235 o 21 %). Síla exprese genů v jednom operonu klesá se vzdáleností od promotorové sekvence [31]. Zde se tento jev úplně nepotvrdil. Je však nutné zohlednit i to, že gen *gpi* je regulován samostatným promotorem (σ^{70}) [31]. Podílí se, jako jeden s klíčových genů, na drahách metabolismu cukrů, a proto je pro buňku nezbytná samostatná regulace jeho exprese.

U SeseC_00234 (*glmU*) byla, jako u jediného genu *has* operonu, identifikována bodová mutace (kap. 5.2, tab. 8). Jde se o tzv. tichou mutaci, kdy nedochází ke změně v sekvenci aminokyselin vzniklého proteinu. Mutace nezpůsobila změnu aminokyselinové sekvence, nicméně stejný polypeptid s jinou nukleotidovou sekvencí by mohl být translatován s rozdílnou efektivitou [131].

Vybrané metabolické dráhy

Tabulka 23: V tabulce jsou uvedeny geny zapojené ve vybraných metabolických drahách souvisejících s produkcí HA. Uvedeno je označení genu, jeho funkce, zkratka a číselné označení k vyhledávání v databázi KEGG [130]. Jsou zde uvedeny poměry exprese genů u obou kmenů normalizované k referenčnímu genu *gyrA*. Geny jsou vykresleny v kontextu biosyntézy HA na obrázku 18. ¹⁾ Poměr celkové exprese genu v ATCC35246 k celkové expresi *gyrA* (ATCC35246). ²⁾ Poměr celkové exprese genu v CO4A k celkové expresi *gyrA* (CO4A). ³⁾ Poměr celkových exprese obou kmenů normalizovaných ke *gyrA*: $CEG_{gyrA} (CO4A) / CEG_{gyrA} (ATCC)$

Gen	CEG_{gyrA} (ATCC) ¹⁾	CEG_{gyrA} (CO4A) ²⁾	Poměr ³⁾	Změna exprese	Funkce genu
Sesec_00293	0,86	0,97	1,13	↑13%	PTS system, maltose and glucose-specific IIA-B-C component [KO:K02777] [EC:2.7.1.-]
Sesec_00296nanE	3,25	5,98	1,84	↑84%	nanE; N-acetylmannosamine-6-phosphate 2-epimerase [KO:K01788] [EC:5.1.3.9]
Sesec_00467pgm	1,73	1,91	1,10	↑10%	gpm; phosphoglycerate mutase family protein
Sesec_00575nagA	1,18	1,22	1,03	↑3%	nagA; N-acetylglucosamine-6-phosphate deacetylase [KO:K01443] [EC:3.5.1.25]
Sesec_00795nagB	0,88	0,96	1,09	↑9%	nagB; glucosamine-6-phosphate isomerase [KO:K02564] [EC:3.5.99.6]
Sesec_00830murA	0,97	1,17	1,21	↑21%	murA; UDP-N-acetylglucosamine 1-carboxyvinyltransferase [KO:K00790] [EC:2.5.1.7]
Sesec_01044eno	2,47	3,02	1,22	↑22%	eno; phosphopyruvate hydratase [KO:K01689] [EC:4.2.1.11]
Sesec_01147	0,78	0,78	1,00	0%	sulfatase [KO:K19005] [EC:2.7.8.20]
Sesec_01167pfk	1,72	1,62	0,94	↓6%	pfk; 6-phosphofructokinase [KO:K00850] [EC:2.7.1.11]
Sesec_01168pyk	1,53	1,67	1,09	↑9%	pyk; pyruvate kinase [KO:K00873] [EC:2.7.1.40]
Sesec_01170glmS	1,73	2,37	1,37	↑37%	glmS; glucosamine--fructose-6-phosphate aminotransferase [KO:K00820] [EC:2.6.1.16]
Sesec_01349ldh	1,48	1,98	1,33	↑33%	ldh; L-lactate dehydrogenase [KO:K00016] [EC:1.1.1.27]
Sesec_01392murB	1,21	1,39	1,15	↑15%	murB; UDP-N-acetylenolpyruvoylglucosamine reductase [KO:K00075] [EC:1.3.1.98]
Sesec_01569glmM	2,31	2,77	1,20	↑20%	glmM; phosphoglucosamine mutase GlmM [KO:K03431] [EC:5.4.2.10]
Sesec_01819ugtP	1,65	2,09	1,27	↑27%	ugtP; glycosyltransferase [KO:K19002] [EC:2.4.1.337]
Sesec_01820	1,32	1,44	1,09	↑9%	glycosyltransferase [KO:K13677] [EC:2.4.1.208]

uchycení na sliznicích hostitelského organismu [151, 146]. Směrem z metabolismu syntézy HA byly sledovány dva geny (Sesec_01044, Sesec_01349). U genu Sesec_01044 (*eno*, EC:4.2.1.11) bylo vypočteno navýšení exprese o 22 % v porovnání s referencí. U genu Sesec_01349 (*ldh*, EC:1.1.1.27) bylo zjištěno navýšení o 33 %. Změna exprese genů směrem do drah glykolýzy byla vyšší než u drah syntézy peptidoglykanu. Buňka by pravděpodobně preferovala energetický metabolismus před navýšením syntézy komponent buněčné stěny. Samozřejmě bude záležet na stavu kultivace a jejích podmínkách. Fruktóza-6-P je konvertována i na glukosamin-6-P, který vede k syntéze UDP-*N*-acetylglukosaminu. Konverze je katalyzována aminotransferázou, která je kódována genem Sesec_01170 (*glmS*, EC:2.6.1.16). Jeho změna exprese byla o 37 % vyšší než u kmene ATCC35246. Pokud porovnáme dva směry, kterými může být fruktóza-6-P zpracována, intenzita exprese genů je porovnatelná.

Jedním z faktorů, které by mohly pozitivně ovlivnit produkci HA, je zpracování *N*-acetylglukosaminu, který byl dodán jako extracelulární substrát. Enzym ke konverzi GlcNAc na *N*-acetylglukosamin-6-P kóduje gen Sesec_00293 (PTS transportní systém, EC:2.7.1.-), jeho exprese byla o 13 % vyšší než u reference. Následně se mění glukosamin-6-P na glukosamin-1-P za katalýzy enzymem, který je kódován genem Sesec_01569 (*glmM*, EC:5.4.2.10). Exprese genu byla o 20 % vyšší než u kmene ATCC35246. *N*-acetylglukosamin-6-P nemusí být směřován přímo do dráhy syntézy HA. Gen Sesec_00296 (*nanE*, EC:5.1.3.9) odvádí *N*-acetylglukosamin-6-P konverzí na *N*-acetylmanosamin-6-P do drah metabolismu manózy (tj. také do drah biosyntézy kyseliny *N*-acetylneuraminové [140, 141, 142]). Exprese genu byla o 84 % vyšší než u kmene ATCC35246. Otázkou je, jaký vliv bude mít dodání externího GlcNAc do kultivačního prostředí při cíleném experimentu. Vliv GlcNAc na produkci HA byl studován pomocí metody qPCR (kap. 5.4.3).

Využití externě dodané sacharózy, která je štěpena na glukózu-6-P a fruktózu, může mít také potenciálně pozitivní vliv na produkci HA. Tento směr dráhy byl sledován prostřednictvím genu Sesec_00467 (*gpm*, EC:5.4.2.2), který kóduje mutázu měnící glukózu-6-P na glukózu-1-P. Míra exprese genu byla o 10 % vyšší než u reference. Navíc glukóza-6-P může vstupovat do drah pentózového cyklu.

5.4 Sledování exprese vybraných genů pomocí metody qPCR

Metoda qPCR umožňuje sledování a analýzu míry exprese genů v reálném čase a je schopná detekovat velmi malé koncentrace analyzovaného substrátu. Klíčové pro hodnocení a interpretaci výsledků získaných touto metodou bylo stanovení referenčního genu. Následně byly analyzovány geny zapojené do metabolických drah syntézy HA a do metabolických drah potenciálně participujících na produkci HA. Byly srovnány standardní kultivace kmenů CO4A a ATCC35246 s kultivacemi suplementovanými GlcNAc, jedním z prekurzorů syntézy HA. Byl testován jeho vliv na expresi uvedených genů i na celkový výtěžek HA. U všech analýz byla jako negativní kontrola použita cDNA z bakteriálního kmene *Escherichia coli DH5α* společně s primery pro gen kódující hyaluronan syntázu (*hasA*). Gen *hasA* není v *E. coli* přítomen, hodnoty byly proto metodou qPCR nedetekovatelné, nebo zcela mimo standardní rozsah detekce.

5.4.1 Stanovení referenčního genu

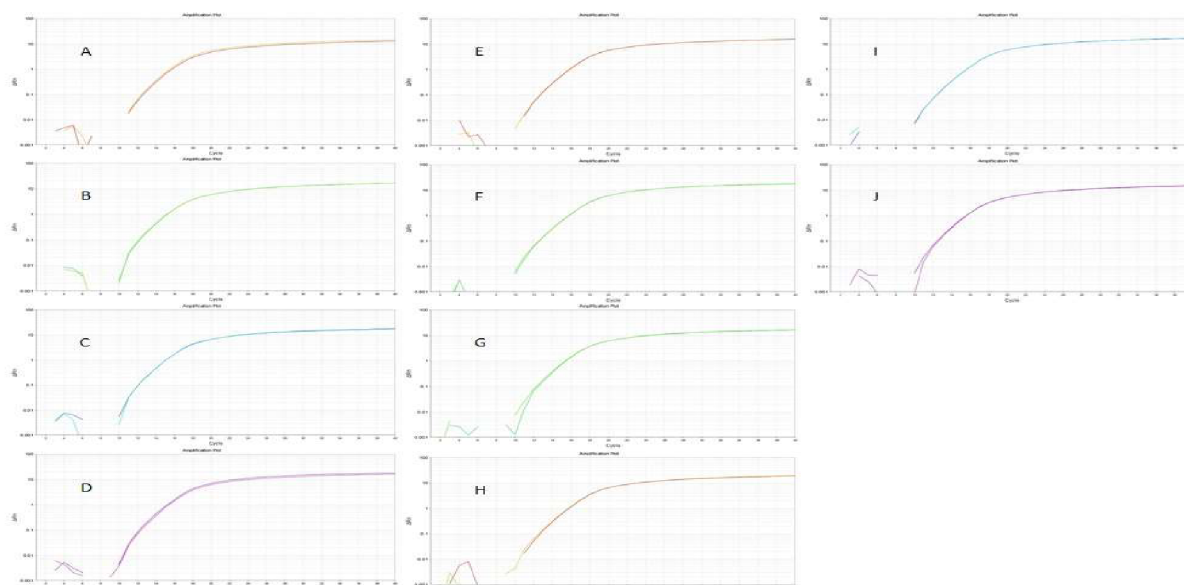
Z výsledků hybridizace na sondách DNA čipů bylo vytipováno 10 potenciálních referenčních genů podle míry a stability jejich exprese v buňkách. Stabilita exprese byla určena vývojem exprese genů ve vzorcích odebíraných v daných časových intervalech. Referenční gen byl stanovován s cílem použít ho pro analýzy u obou zmiňovaných kmenů, aby byly výsledky qPCR analýzy obou kmenů mezi sebou po normalizaci porovnatelné.

Následně byly pro tyto geny navrženy primery (kap. 4.8.1). Přehledný seznam vybraných genů s jejich charakteristikou a navrženými primery je uveden v tabulce 24.

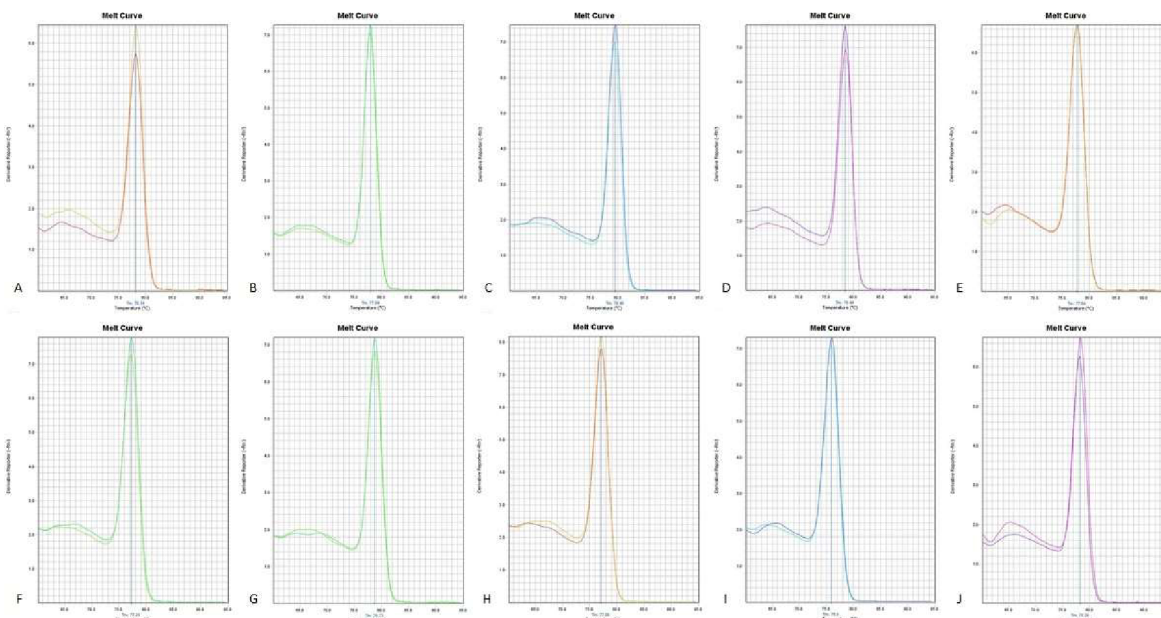
Tabulka 24: V tabulce je uveden přehled vybraných potenciálních referenčních genů s jejich charakterizací podle databáze NCBI [74]. Ke každému genu byla navržena dvojice primerů, ohraničující amplifikovaný úsek o velikosti mezi 50–150 bp. Primery byly následně nasyntetizovány společností GenScript®.

Označení genu	Primery		Charakteristika genu
Sesec_00096	Forward	TGGTTCACGCTGTGTCTTT	NUDIX hydrolase
	Reverse	GCTGGGCCAACCCTCTTTAT	
Sesec_00097	Forward	GCTTGTCTCTTGGGTGGTT	ispE, 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
	Reverse	ACCTGCCTCAATGGCTGAAA	
Sesec_00549	Forward	AGGTTTCTGGCATTTCGCTCA	lacB, galactose-6-phosphate isomerase subunit LacB
	Reverse	CTCGCCTGTGATTTTACCGC	
Sesec_00725	Forward	CGGTAACAGCTTCTCGTGGT	ileS, isoleucyl-tRNA synthetase
	Reverse	CAGTCCTTCTGCCACCACAT	
Sesec_00767	Forward	GAGCAGACAAGGCAGGAGTT	transcriptional repressor
	Reverse	AGGGTGCTCTACCTCCACAT	
Sesec_00768	Forward	GCAGGGCTTGGTGCTGTAATA	hypothetical protein
	Reverse	AGCTCAATACCAGCCCTCCT	
Sesec_02327	Forward	ACTCAAGTGCAGGCTTTGA	mscS, mechanosensitive ion channel protein
	Reverse	GCCGCGTGATTTGTTGCTAA	
Sesec_02508	Forward	CCCTCGATAAGCTCGCAAGG	mviM, oxidoreductase
	Reverse	CTACAGTCAGGTAGCAGGCG	
Sesec_01348	Forward	GGTTTCAATCTCGGTGCGTG	gyrA, DNA gyrase subunit A
	Reverse	CCTTGGTGATGGGACGTTCA	
Sesec_01041	Forward	TACTGCTATGGGACAGGCT	gyrB, DNA gyrase subunit B
	Reverse	CCCCATCGACATCTGCATCA	

Po konečné optimalizaci qPCR byla provedena analýza exprese potenciálně referenčních genů. Amplifikační křivky a křivky tání deseti vybraných genů kmene ATCC35246 (vzorky ze 4té hodiny kultivace) jsou uvedeny na obrázku 19 a 20.



Obrázek 19: Zde jsou uvedeny amplifikační křivky potenciálních referenčních genů kmene ATCC35246 ze 4té hodiny kultivace. Každý vzorek byl připraven ve dvou technických replikátech, u kterých se křivky překrývají. A – Sesece_00096; B – Sesece_00097ispE; C – Sesece_00549lacB; D – Sesece_00725ileS; E – Sesece_00767; F – Sesece_00768; G – Sesece_01041gyrB; H – Sesece_01348gyrA; I – Sesece_02327mscS; J – Sesece_02508mviM



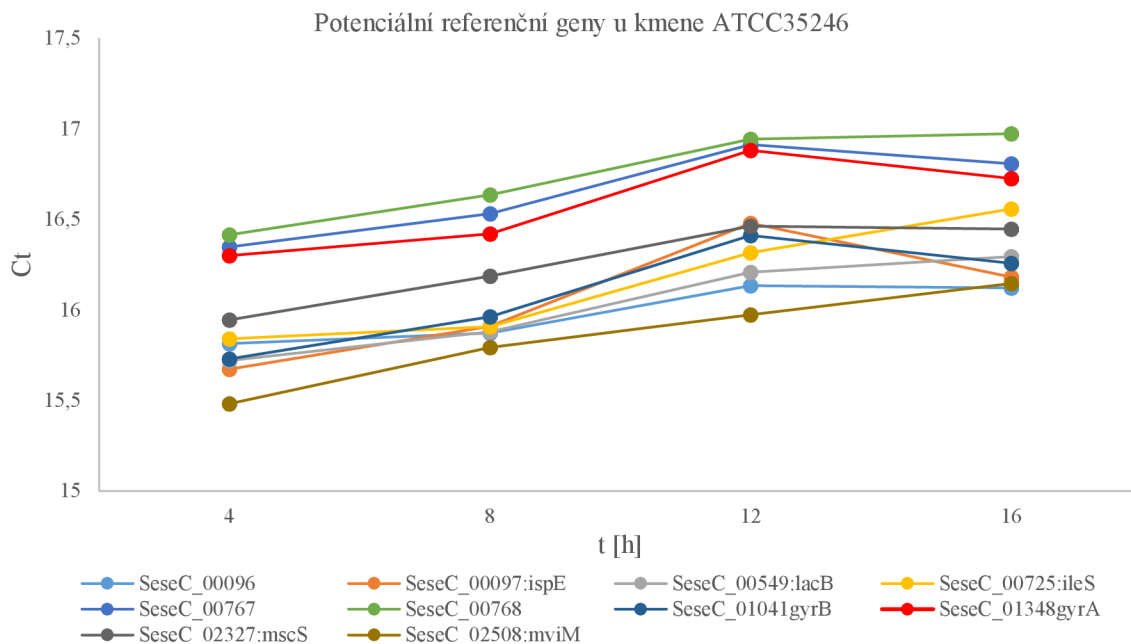
Obrázek 20: Zde jsou uvedeny křivky tání potenciálních referenčních genů kmene ATCC35246 ze 4té hodiny kultivace. Každý vzorek byl připraven ve dvou technických replikátech, u kterých se křivky překrývají. Z obrázků je patrná přítomnost jednoho výrazného piku, který je jasně vymezen. Pokud vznikal nespecifický produkt, tak pouze v minimálním množství, které neovlivnilo konečný výsledek analýzy. A – Sesece_00096; B – Sesece_00097ispE; C – Sesece_00549lacB; D – Sesece_00725ileS; E – Sesece_00767; F – Sesece_00768; G – Sesece_01041gyrB; H – Sesece_01348gyrA; I – Sesece_02327mscS; J – Sesece_02508mviM

Stejná analýza byla provedena pro vybrané geny u interního kmene CO4A. Grafická podoba byla exportována do excelového souboru a byly generovány hodnoty C_t , které mají vypovídající charakter o míře exprese genů. Po zpracování dat v programu Excel s doplňkem GeNorm [122] bylo možné posoudit, který z genů by byl nejvhodnějším referenčním genem. Míra exprese testovaných genů vyjádřena jako C_t (kap. 4.8.2) je uvedena v následující tabulce 25.

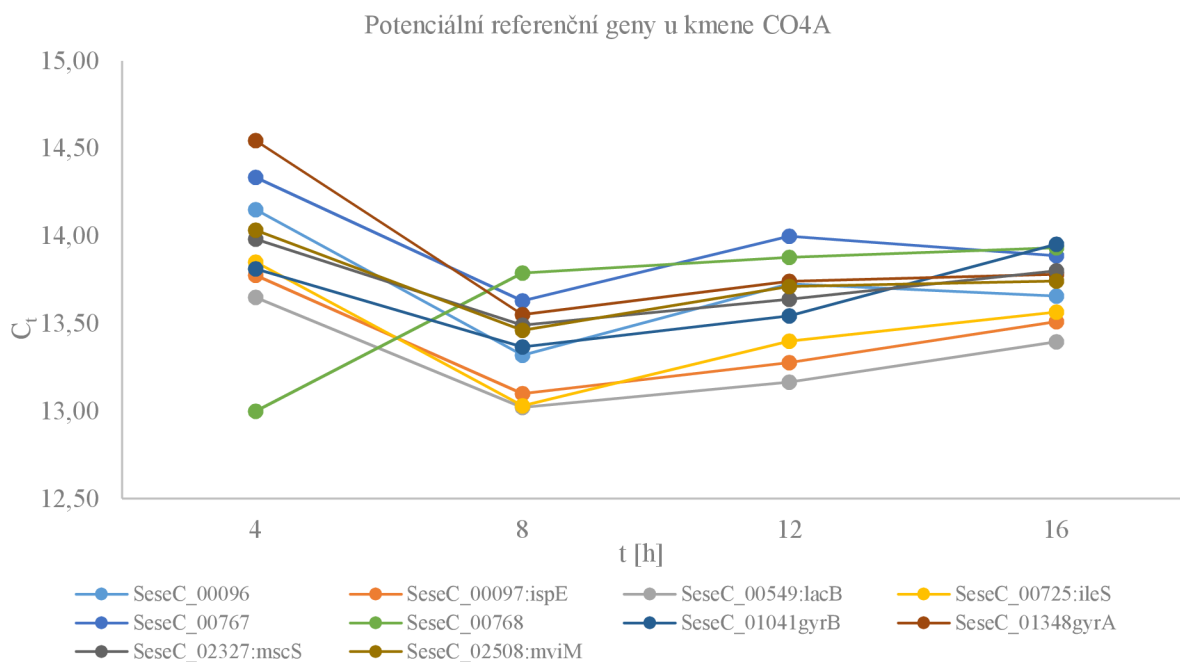
Tabulka 25: V tabulce jsou uvedeny výsledné hodnoty C_t pro první biologický replikát kmene ATCC35246 a kmene CO4A.

C_t	CO4A				ATCC35246				
	Gen	4 h	8 h	12 h	16 h	4 h	8 h	12 h	16 h
SeseC_00096		14,15	13,32	13,73	13,66	15,81	15,87	16,13	16,12
SeseC_00097:ispE		13,78	13,10	13,28	13,51	15,67	15,91	16,48	16,18
SeseC_00549:lacB		13,65	13,02	13,17	13,40	15,72	15,88	16,21	16,29
SeseC_00725:ileS		13,85	13,03	13,40	13,57	15,84	15,91	16,32	16,56
SeseC_00767		14,33	13,63	14,00	13,89	16,35	16,53	16,91	16,81
SeseC_00768		13,00	13,79	13,88	13,93	16,42	16,64	16,94	16,97
SeseC_01041gyrB		13,81	13,37	13,54	13,95	15,73	15,96	16,41	16,26
SeseC_01348gyrA		14,54	13,55	13,74	13,78	16,30	16,42	16,88	16,73
SeseC_02327:mscS		13,98	13,49	13,64	13,80	15,95	16,19	16,46	16,45
SeseC_02508:mviM		14,03	13,46	13,71	13,74	15,48	15,79	15,97	16,15

Výše uvedené hodnoty C_t pro každý gen jsou vyneseny v grafu 1 pro referenční kmen ATCC35246 a v grafu 2 pro produkční kmen CO4A.

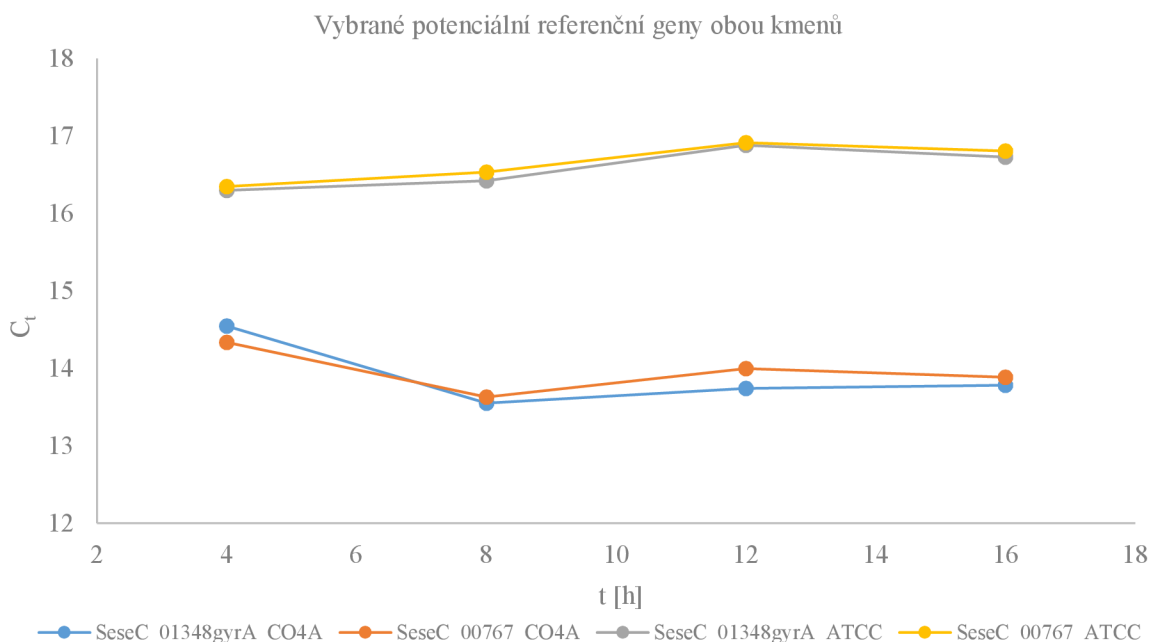


Graf 1: V grafu jsou uvedeny výsledné hodnoty C_t , které byly získány analýzou qPCR. Výsledky jsou uvedeny pro první biologický replikát kmene ATCC35246. Byly testovány vzorky ze 4., 8., 12. a 16. hodiny kultivace. Z grafu jsou patrné průběhy exprese vybraných genů. Z těchto dat byl hodnocen průběh křivky genu ve srovnání s ostatními potenciálními referenčními geny.



Graf 2: V grafu jsou uvedeny výsledné hodnoty C_t , pro první biologický replikát kmene CO4A. Z grafu jsou patrné průběhy exprese vybraných genů. Z těchto dat byl hodnocen průběh křivky genu ve srovnání s ostatními potenciálními referenčními geny.

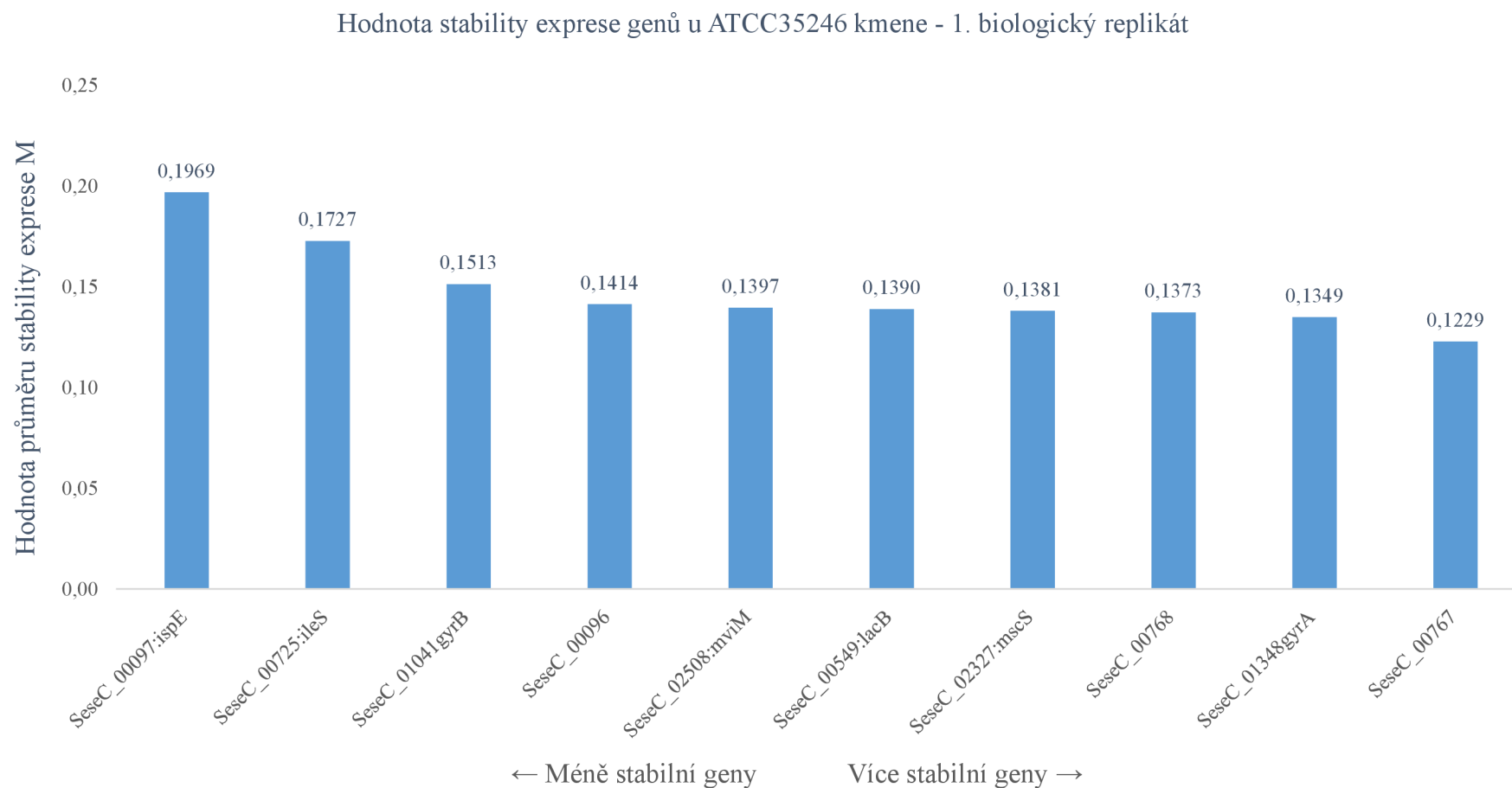
Z předchozího vyhodnocení dat a vývoje exprese genů v průběhu celé kultivace byly vybrány dva geny (Sesec_01348 – *gyrA* a Sesec_00767), které by mohly být použity jako geny referenční. Tyto geny se vyznačují tím, že mají stabilní míru exprese ve všech odebíraných časech u obou sledovaných kmenů. Průběh exprese u vybraných genů je zobrazen v grafu 3.



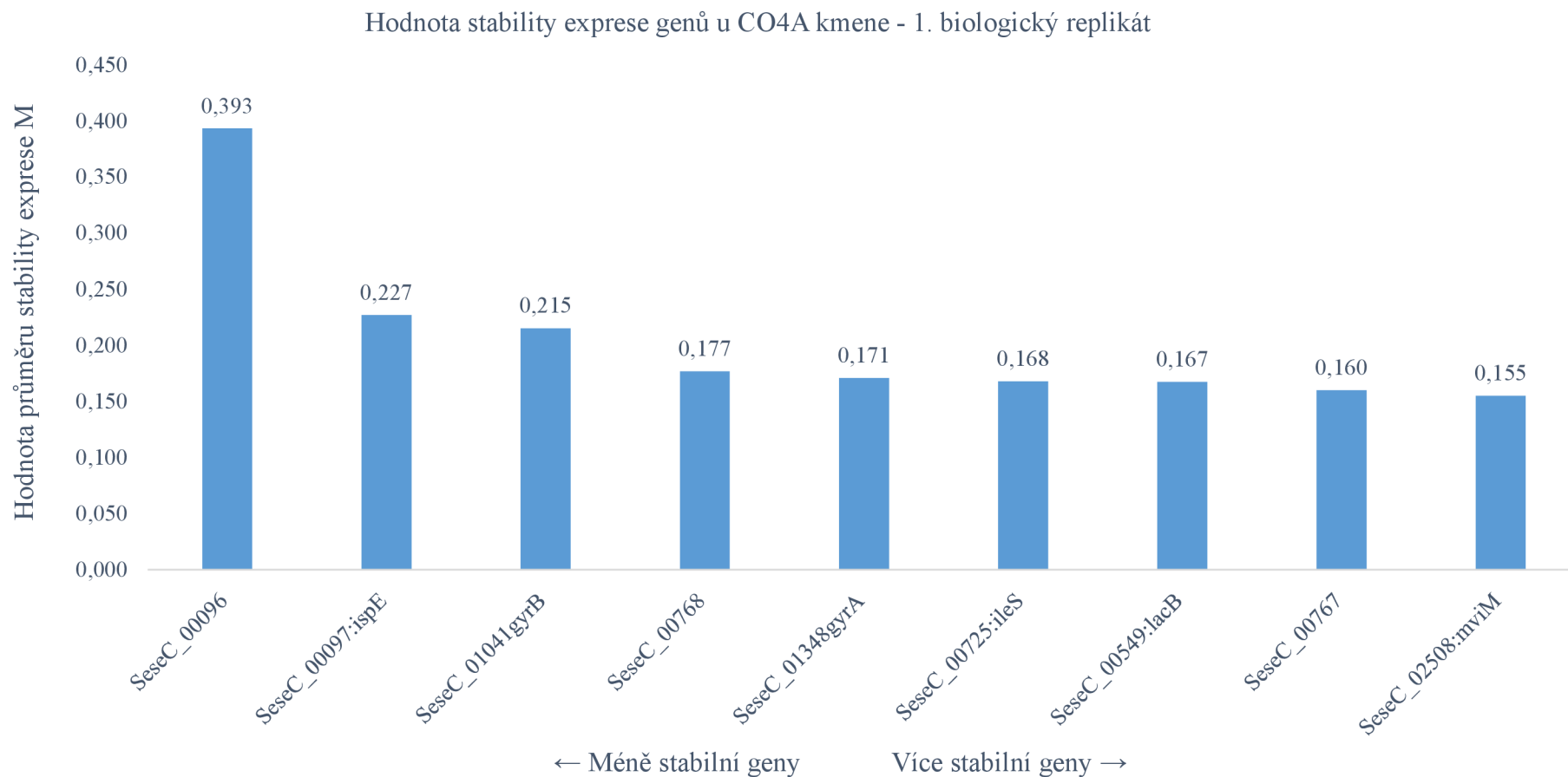
Graf 3: Z předchozích analýz dat získaných metodou qPCR pro vybrané potenciální referenční geny byly vybrány dva geny nejvíce vyhovující (Sesec_01348gyrA a Sesec_00767). Na grafu jsou uvedeny průběhy míry jejich exprese vyjádřené hodnotou C_t po dobu kultivace v bioreaktoru.

Pro potvrzení našich výsledků byla navíc provedena analýza softwarovým doplňkem programu Microsoft Excel – GeNorm [122], který je určen přímo ke stanovení stability exprese genů, tudíž ke stanovení nejvhodnějšího referenčního genu.

V grafu 4 jsou uvedeny výsledky analýzy GeNorm doplňkem pro první biologický replikát sbírkového kmene ATCC35246 a všech 10 vybraných potenciálních referenčních genů. V grafu 5 jsou též data pro kmen CO4A. GeNorm software využívá k analýze hodnoty M , která představuje průměr stability exprese genu. Přičemž lze nastavit kritický limit, za kterým hodnoty M nevyhovují požadavkům pro referenční gen. Limit je obvykle stanovován na $M = 1,5$ [122].



Graf 4: Hodnoty průměrné stability exprese M pro první biologický replikát sbírkového kmene ATCC35246. Hodnoty M pro jednotlivé vybrané geny nepřesahují hodnotu 0,2, z čehož můžeme usuzovat, že jako referenční geny mohou být teoreticky použity všechny výše vybrané geny.



Graf 5: Hodnoty průměrné stability exprese M pro první biologický replikát produkčního kmene CO4A. Hodnoty M pro jednotlivé vybrané geny nepřesahují hodnotu 0,4. Jako referenční geny mohou být teoreticky použity všechny výše vybrané geny.

Častým referenčním genem u prokaryotických organismů je v literatuře uváděn gen *gyrA* kódující podjednotku A DNA gyrázy [152, 153, 154]. Na základě našich výsledků byl tento závěr potvrzen a k dalším analýzám byl jako referenční gen zvolen *gyrA*. Hodnotami z analýzy Genorm doplňkem byla potvrzena správnost výběru potenciálních referenčních genů na základě dat získaných sledováním celogenomové exprese metodou DNA čipů (kap 5.3). Nepřímo byla tímto výsledkem potvrzena i správnost výsledků analýzy exprese genů metodou DNA čipů.

5.4.2 Ověření exprese genů *has* operonu

Geny *has* operonu jsou odpovědné za produkci kyseliny hyaluronové. Kyselina hyaluronová jako jeden z virulentních faktorů je pro přežití streptokoka v hostitelském organismu velmi důležitá. Proto bylo zajímavé sledovat expresi těchto genů u kmene CO4A, který vykazuje vysokou míru produkce HA v porovnání se sbírkovým kmenem ATCC35246. Geny a navržené primery jsou uvedeny v tabulce 26.

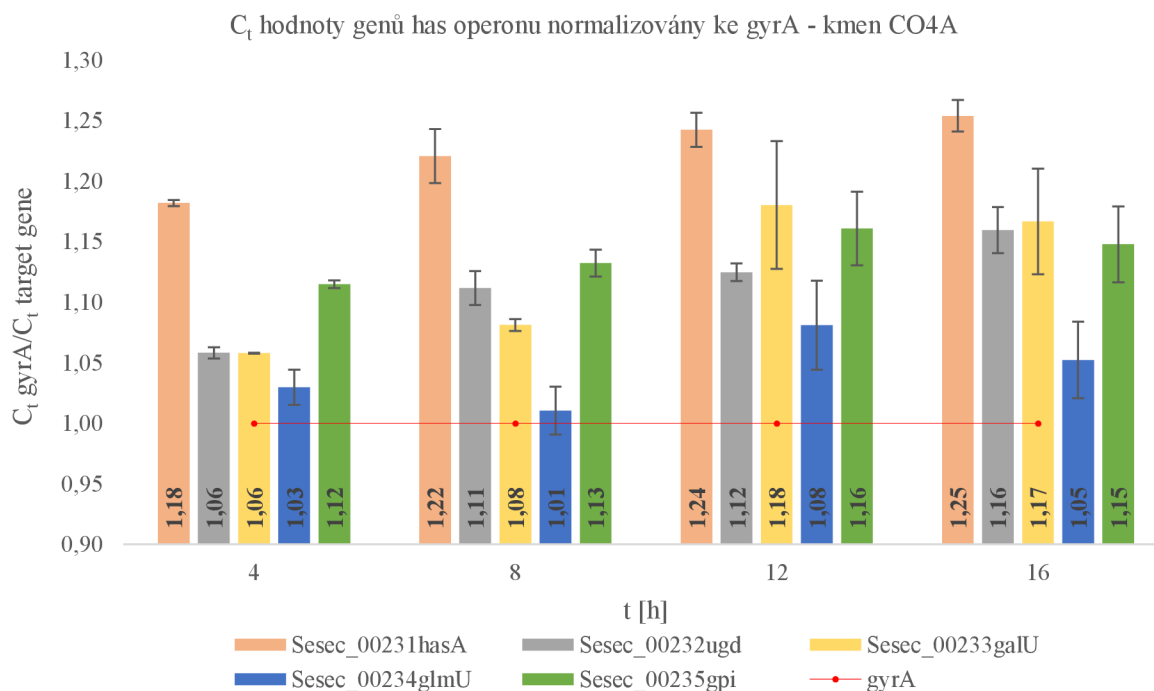
Tabulka 26: Geny *has* operonu s navrženými primery pro analýzu qPCR

Označení genu	Primery		Produkt genu (protein)
Sesec_00231	Forward	TTAAGCACCCGCTGTCCTTC	hasA, hyaluronan synthase HasA
	Reverse	TTACGTGTTCCCCAGTCAGC	
Sesec_00232	Forward	TTTATGGGCTCACAGGAGGC	ugd, UDP-glucose 6-dehydrogenase HasB
	Reverse	AGACACCCGCATAGCCAAAA	
Sesec_00233	Forward	CTCCAAGCCAAGGGCAAAAAC	galU, UTP-glucose-1-phosphate uridylyltransferase (HasC)
	Reverse	AGTCCTCTAGGGTGGCTCTG	
Sesec_00234	Forward	GCCAGCCAAGAAAAGGTTGG	glmU, Bifunctional protein glmU
	Reverse	TTGTTGATCCGCTCCTGCAT	
Sesec_00235	Forward	TTTCTTGGGCTGGCTTGAGT	gpi, glucose-6-phosphate isomerase
	Reverse	AGCACCAAGGTAAGAGCCAC	

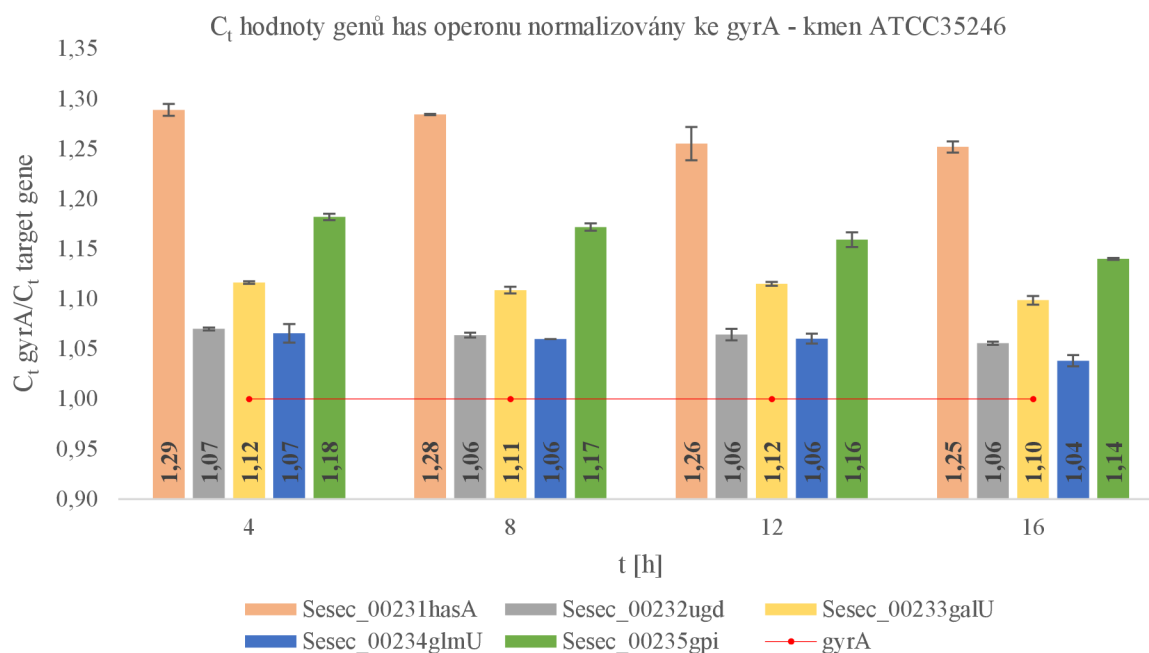
Míra exprese genů *has* operonu byla sledována metodou qPCR (kap. 4.8.1). Získané hodnoty C_t byly normalizovány k referenčnímu genu *gyrA* a byl vypočten průměr ze tří biologických replikátů experimentu (kap. 4.8.1). Data jsou uvedena v tabulce 27 a byla vynesena do grafů pro oba kmeny (graf 6 a 7). Geny jsou uvedeny v pořadí, v jakém jsou přítomny v *has* operonu. Směrem od promotoru míra exprese klesá. Výjimkou je gen *gpi* kódující glukózu-6-fosfoizomerázu, který je řízen samostatným, silně konzervovaným σ^{70} promotorem [31].

Tabulka 27: Průměrné hodnoty C_t biologických replikátů obou kmenů v poměru k C_t referenčnímu genu *gyrA* ($C_{tgyrA}/C_{t genu}$).

<i>has</i> operon	ATCC35246				CO4A			
Označení genu	4 h	8 h	12 h	16 h	4 h	8 h	12 h	16 h
SeseC_00231:hasA	1,29	1,28	1,26	1,25	1,18	1,22	1,24	1,25
SeseC_00232:ugd	1,07	1,06	1,06	1,06	1,06	1,11	1,12	1,16
SeseC_00233:galU	1,12	1,11	1,12	1,10	1,06	1,08	1,18	1,17
SeseC_00234:glmU	1,07	1,06	1,06	1,04	1,03	1,01	1,08	1,05
SeseC_00235:gpi	1,18	1,17	1,16	1,14	1,12	1,13	1,16	1,15
SeseC_01348:gyrA	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00



Graf 6: Normalizované C_t hodnoty genů has operonu kmene CO4A. has operon je regulován dvěma konzervovanými promotory. Míra exprese odpovídá předpokladu. Gen kódující hasA měl nejvyšší expresi ze všech sledovaných genů. Míra exprese genů klesá se vzdáleností od promotorové oblasti, tak jak ukazují data v grafu. Výjimku tvoří gen Sesec_00235 (gpi), který kóduje enzym klíčový pro dráhy energetického metabolismu. Je tedy pravděpodobná možnost samostatné regulace a proto je před genem umístěna druhá promotorová sekvence regulující expresi pouze gpi genu [31].



Graf 7: U kmene ATCC35246 byl vývoj exprese genů stejný jako u kmene CO4A. Dokonce u ATCC35246 byla relativní míra exprese hasA ve 4. hodině kultivace vyšší než u CO4A. Ostatní geny měly míru exprese srovnatelnou s CO4A.

V této kapitole výsledků byla zhodnocena exprese genů *has* operonu u kmene produkčního CO4A a referenčního ATCC35246. Expresse genů je vyšší než u referenčního genu *gyrA*. Signifikantně vyšší pak v případě genů *hasA* a *gpi*. Geny *has* operonu jsou pod kontrolou dvou nezávislých promotorů [31]. Pod kontrolou druhého silně konzervovaného promotoru σ^{70} je pouze gen *gpi*. Jeho exprese je nižší než exprese genu *hasA*, nicméně výrazně vyšší než exprese zbylých genů *has* operonu. Gen *hasA* kódující hyaluronan syntázu je exprimován s nejvyšší intenzitou. Je zařazen na první místo hned po sekvenci prvního promotoru (expresse genů klesá se vzdáleností od sekvence promotorové části [31]). Co se týká změny míry exprese v čase, nepozorovali jsme žádný výrazný trend u kmene CO4A, ani u kmene ATCC35246. Celkově to vypadá, že exprese genů *has* operonu je u obou sledovaných kmenů srovnatelná. Avšak u analýzy metodou DNA čipů (kap. 5.3.4) byl zaznamenán signifikantní nárůst míry exprese genů, *has* operonu, u kmene CO4A v porovnání s referenčním kmenem ATCC35246. Mohlo to být způsobeno rozdílným stavem buněk při kultivaci. Počet buněk dodaných v inokulu by mělo být velmi podobné (ověřeno spektrofotometricky), nicméně kvalita kultury mohla být odlišná (u kmene ATCC35246 horší) a tím mohla být kultivace zpožděná. Metoda DNA čipů také může ve vyšší míře vykazovat falešně pozitivní výsledky, právě proto je hojně využívána v kombinaci s metodou qPCR. Bakteriální buňky kmene ATCC35246 však produkují výrazně méně HA než buňky produkčního kmene CO4A (kap. 5.4.4). Protože u studovaných kmenů metodou qPCR nebyla zjištěna signifikantní změna exprese genů *has* operonu, je možné konstatovat, že biosyntéza kyseliny hyaluronové je v nich regulována i jiným způsobem.

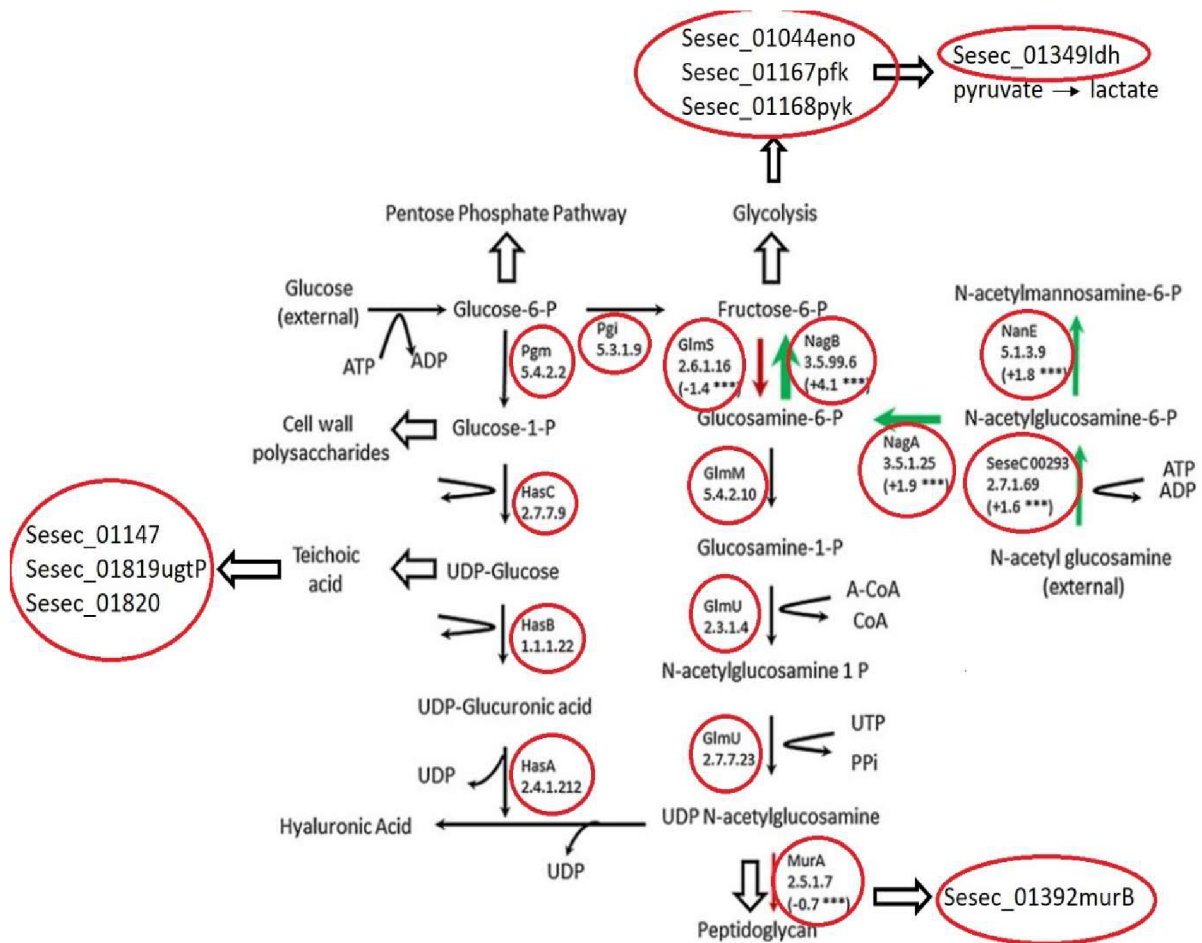
5.4.3 Vliv *N*-acetylglukosaminu na expresi vybraných skupin genů

Již dříve byl v literatuře publikován pozitivní vliv suplementace média GlcNAc na produkci HA u bakterií rodu *Streptococcus* [6]. Ke studiu exprese metodou qPCR bylo vybráno 21 genů. Část byla vybrána na základě výše uvedeného článku a část byly námi zvolené geny (tab. 28).

Tabulka 28: V tabulce jsou uvedeny vybrané geny, jejichž exprese byla sledována při kultivaci bez suplementace a se suplementací GlcNAc v kultivačním médiu.

Označení genu	Primery		Charakteristika genu
Sesec_00231	Forward	TTAAGCACCCGCTGTCCTTC	hasA, hyaluronan synthase HasA
	Reverse	TTACGTGTTCCCCAGTCAGC	
Sesec_00232	Forward	TTTATGGGCTCACAGGAGGC	Ugd, UDP-glucose 6-dehydrogenase HasB
	Reverse	AGACACCCGCATAGCCAAAA	
Sesec_00233	Forward	CTCCAAGCCAAGGGCAAAAC	galU, UTP-glucose-1-phosphate uridylyltransferase HasC
	Reverse	AGTCCTCTAGGGTGGCTCTG	
Sesec_00234	Forward	GCCAGCCAAGAAAAGGTTGG	glmU, Bifunctional protein glmU
	Reverse	TTGTTGATCCGCTCCTGCAT	
Sesec_00235	Forward	TTTCTTGGGCTGGCTTGAGT	gpi, glucose-6-phosphate isomerase
	Reverse	AGCACCAAGGTAAGAGCCAC	
Sesec_00293	Forward	GACGGTTTTTCCGACCAAGC	PTS system, maltose and glucose-specific IIA-B-C component
	Reverse	CCAGCGACTACCTTTGACC	
Sesec_00467	Forward	TGCTTTGGTAGCCTGGATGG	pgm, phosphoglycerate mutase family protein
	Reverse	AGGCACATCACGCAGATTGT	
Sesec_00575	Forward	TACAATGGGATGCGTGGCTT	nagA, <i>N</i> -acetylglucosamine-6-phosphate deacetylase
	Reverse	ATCGCAAGCAACAGGTGAGA	
Sesec_00795	Forward	CTTGGCAGAAGGGGCTACAA	nagB, glucosamine-6-phosphate isomerase
	Reverse	TAATCGGTGTGCTGCCTGTT	
Sesec_00830	Forward	GGCGTAAGGATGACGGTTGA	murA, UDP- <i>N</i> -acetylglucosamine 1-carboxyvinyltransferase
	Reverse	GCCCTTAAAAGCAGAGGGGT	
Sesec_01147	Forward	GGAAAAGCCCTCTTCACTGCT	sulfatase
	Reverse	CGTCAGCAACCACTACCCTT	
Sesec_01167	Forward	TGTTGAAGTGATGGGACGCA	pfk, 6-phosphofructokinase
	Reverse	GCAATACCTGCCCAAAGTGC	
Sesec_01168	Forward	ACTTGGTCTGCGGTTGAAA	pyk, pyruvate kinase
	Reverse	AACTGACCAGCCCAATAGCC	
Sesec_01170	Forward	CAACGCATGGTCAAGCAACA	glmS, glucosamine-fructose-6-phosphate aminotransferase
	Reverse	CGTGTGTCCAGCAAGTACT	
Sesec_01349	Forward	GATGCACGTTCTGTCCATGC	ldh, <i>L</i> -lactate dehydrogenase
	Reverse	TCGTACAATCCAACCCCTGC	
Sesec_01392	Forward	CAAAGGCCATCTCTTGGCT	murB, UDP- <i>N</i> -acetylenolpyruvoylglucosamine reductase
	Reverse	TGCAGGTGCTTATGGTGGAG	
Sesec_01569	Forward	TGCCATCACACGCAACAATG	glmM, phosphoglucosamine mutase GlmM
	Reverse	CTATGGAGGTGCCAGCGATT	
Sesec_01819	Forward	TCTCTAGTGCAATGCCGGTC	ugtP, glycosyltransferase
	Reverse	CCCCTGCTTCGTGGGTATTT	
Sesec_01820	Forward	AAATAACGCCGAAAGGCAGC	glycosyltransferase
	Reverse	GGACGCCATTATCACACCGA	
Sesec_00296	Forward	GAAGCAGGAGCAGTAGGCAT	nanE, <i>N</i> -acetylmannosamine-6-phosphate 2-epimerase
	Reverse	GTTGGCGTGATAAAGGGGGA	
Sesec_01044	Forward	AAGCTGGCTACACTGCTGTT	eno, phosphopyruvate hydratase
	Reverse	GTTTGTGCCACTGCGATGT	

Na následujícím schématu (obr. 21) jsou tyto geny vykresleny v souvislosti s metabolickými drahami HA. U nepřímo zapojených metabolických drah se jedná především o enzymy zapojené v syntéze buněčné stěny, peptidoglykanu a do energetického metabolismu buňky.



Obrázek 21: Geny označeny červeným kroužkem byly vybrány jako potenciálně zajímavé pro studium jejich exprese. [upraveno dle 6]

Byly provedeny kultivace bakteriálních kmenů CO4A a ATCC35246 v kultivačním médiu suplementovaném *N*-acetylglukosaminem (0,5 g/l) a v kultivačním médiu bez suplementace (kap. 4.1.3). qPCR byla nastavena dle parametrů uvedených v tabulce 5 (kap. 4.8.1). Samozřejmě bylo zařazení negativní kontroly, pro kterou byla použita jako templát cDNA získaná přepisem RNA z bakterie *E. coli DH5α* a primery pro gen *hasA*, který se u *E. coli* nevyskytuje. Získaná data byla dále analyzována.

Analyza změn relativní genové exprese metodou $2^{(-\Delta\Delta C_t)}$

Získané C_t hodnoty byly normalizovány k referenčnímu genu *gyrA* a byla provedena analýza relativních změn genové exprese metodou $2^{(-\Delta\Delta C_t)}$ [123] (kap. 4.8.4). Výsledné hodnoty pro 1. biologický replikát obou kmenů jsou uvedeny v tab. 29.

Tabulka 29: Hodnoty relativní změny exprese vybraných skupin genů získané metodou $2^{(-\Delta\Delta C_t)}$

Gen	Fold change 1. biologického replikátu							
	CO4A				ATCC			
	6	9	12	16	6	9	12	16
<i>gyrA</i>	1	1	1	1	1	1	1	1
Sesec_00231 <i>hasA</i>	1,11	1,23	0,44	0,46	1,15	1,05	0,83	0,74
Sesec_00232 <i>ugd</i>	1,18	1,18	0,71	0,48	0,69	0,78	1,42	1,07
Sesec_00233 <i>galU</i>	1,15	1,25	0,22	0,26	1,23	0,80	0,73	0,67
Sesec_00234 <i>glmU</i>	1,17	1,81	0,24	0,40	0,96	0,72	0,72	0,74
Sesec_00235 <i>gpi</i>	1,04	1,08	0,43	0,55	1,01	1,07	1,04	1,03
Sesec_00293	1,00	1,33	0,44	0,65	1,22	0,72	1,09	0,91
Sesec_00296 <i>nanE</i>	1,17	1,57	0,78	0,99	0,74	0,91	1,02	0,74
Sesec_00467 <i>pgm</i>	1,11	1,30	0,63	0,56	1,02	0,80	0,89	0,70
Sesec_00575 <i>nagA</i>	0,91	1,19	0,65	0,80	1,23	0,71	1,19	1,03
Sesec_00795 <i>nagB</i>	1,18	1,23	0,42	0,32	0,67	0,58	0,70	0,53
Sesec_00830 <i>murA</i>	1,14	1,70	0,65	0,95	1,12	0,91	1,28	1,14
Sesec_01044 <i>eno</i>	1,09	1,25	0,53	0,54	1,35	1,11	0,98	0,93
Sesec_01147	1,25	1,33	0,49	0,47	1,11	0,70	1,22	1,01
Sesec_01167 <i>pfk</i>	1,22	1,40	0,33	0,24	1,03	0,90	1,14	0,97
Sesec_01168 <i>pyk</i>	1,01	1,03	0,47	0,39	0,85	0,88	1,18	0,99
Sesec_01170 <i>glmS</i>	1,16	1,81	0,52	0,87	0,79	0,80	0,75	0,69
Sesec_01349 <i>ldh</i>	0,89	1,19	0,75	0,83	0,84	0,85	1,19	1,06
Sesec_01392 <i>murB</i>	1,03	0,75	0,55	0,60	0,52	0,45	0,97	0,95
Sesec_01569 <i>glmM</i>	0,76	0,94	0,68	0,78	1,03	0,73	0,79	0,75
Sesec_01819 <i>ugtP</i>	0,70	1,27	0,41	0,51	0,93	0,97	1,09	0,75
Sesec_01820	1,07	1,38	0,77	0,92	0,65	0,64	0,84	0,89

Hodnoty získané metodou $2^{(-\Delta\Delta C_t)}$ byly pro jednotlivé geny vyneseny do grafů 8 až 28. Definují relativní změny exprese genu ve vztahu k referenčnímu genu a kontrolní kultivaci [123]. Chybové úsečky byly získány z dat obou biologických replikátů daného kmene.

Hlavní studovanou metabolickou dráhou byla biosyntéza HA. Geny kódující enzymy v ní zapojené, se nacházejí v sekvenci za sebou v jednom operonu. *has* operon je pojmenován podle produkce HA. První čtyři geny jsou regulovány jedním promotorem, a proto můžeme očekávat obdobnou reakci na změnu kultivačního prostředí (obohacení média GlcNAc) [31]. Skutečně u zmíněných genů kmene CO4A sledujeme v prvních hodinách (6. a 9. hodině) kultivace navýšení míry exprese v porovnání s kontrolní kultivací (u Sesec_00231 *hasA* o 11 % a 23 %, u Sesec_00232 *ugd* o 18 % v obou případech, u Sesec_00233 *galU* o 15 % a 25 %, u Sesec_00234 *glmU* o 17 % a 81 %). U kmene ATCC35246 tento efekt nepozorujeme zřejmě proto, že GlcNAc je využíván v pozdější fázi kultivace. Gen Sesec_00234 (*glmU*) v CO4A má

ve své nukleotidové sekvenci bodovou mutací (kap. 5.2, tab. 8), která ovšem nezpůsobuje změnu na proteinové sekvenci. Vliv této mutace na expresi *glmU* bude pravděpodobně minimální, nicméně ji nelze zcela vyloučit [131]. U genu kódujícího *glmU* byl u CO4A pozorován výrazný nárůst exprese v 9. hodině kultivace (81 %). Geny kódující *glmU* a *galU* mají v sekvenci genomu i svou druhou kopii (homologní sekvenci). Ta může souviset s výrazným nárůstem exprese *glmU* v CO4A a mohla by vést k vyšší koncentraci enzymu syntetizujícího prekurzor (GlcNAc) syntézy HA. V případě genu *Sesec_00233 (galU)* byla v jeho kopii identifikována SNV mutace, kdy došlo i ke změně AMK sekvence (Ala→Thr). Především u kmene ATCC35246 ve 4. hodině kultivace byl sledován nárůst u *galU* o 23 %. Oproti tomu přidavek GlcNAc do média neměl vliv na expresi *Sesec_00235 (gpi)*. Ke konci kultivace (ve 12. a 16. hodině) pozorujeme celkový pokles exprese. Kultura se dostává do stacionární fáze růstu, kultivační podmínky se zhoršují a může docházet k intenzivnější degradaci RNA. Zhoršení kultivačních podmínek spočívá především ve změně pH (například vysoká produkce laktátu [28]), jeho vyrovnáváním pomocí NaOH a s tím spjatými změnami v osmotickém prostředí buněk. Dále dochází v důsledku vysoké viskozity média na konci kultivace ke zhoršení toku živin kvůli ztíženým podmínkám míchání a vzdušnění v reaktoru. Stres buněk může mít také vliv na míru exprese jednotlivých genů a na adaptaci organismu na jiné prostředí, kde kapsule již ztrácí svou úlohu.

U kmene ATCC35246 není soudržnost v míře exprese genů *has* operonu tak výrazná. Zejména u genu *Sesec_00232 (ugd)* data poněkud překvapivě ukazují nižší expresi oproti kontrole bez GlcNAc v prvních hodinách kultivace a naopak výrazný nárůst exprese ve 12. hodině experimentu. Pokud by se pokles exprese genu *Sesec_00232* u ATCC35246 projevil i na proteinové úrovni mohl by zabrzdit navazující procesy, což by v konečném důsledku mohlo mít za následek snížení produkce HA. Vzhledem k tomu, že jeho substrátem je UDP-glukóza, tak jeho nízká exprese v počátku kultivace by mohla mít vliv na navýšení syntézy struktur buněčné stěny. Je pravděpodobné, že kmen ATCC35246 v první polovině kultivace upřednostňuje syntézu buněčné stěny před zvýšenou produkcí HA, která může být zpožděná. Tento kmen tak za daných kultivačních podmínek produkuje méně HA než kmen CO4A.

U dalších genů kmene CO4A zapojených přímo v drahách syntézy HA pozorujeme stejný expresní vývoj (*Sesec_00467 (pgm)*, *Sesec_01167 (pfk)*). U genu *Sesec_001167 (pfk)* pozorujeme v 9. hodině kultivace nárůst o 40 % v porovnání s kontrolní kultivací. Enzym je zapojen také v drahách glykolýzy. Metabolitem, který je odváděn tím směrem a postupně zpracováván enzymem *pfk* společně s geny *Sesec_01168 (pyk)* a *Sesec_01044 (eno)* je fruktóza-6-P. U dalších dvou genů *Sesec_01168 (pyk)* a *Sesec_01044 (eno)* nepozorujeme tak výrazné navýšení exprese jako u *Sesec_01167 (pfk)*. Enzym *pfk* je centrálním enzymem glykolýzy. Enzym *eno* mění glycerát na fosfoenolpyruvát a na tuto reakci navazuje enzym *pyk*, který katalyzuje změnu fosfoenolpyruvátu na pyruvát. Zmíněná enzymatická posloupnost odvádí metabolity z hlavních drah až k odbourávání pyruvátu na laktát za katalýzy enzymu *ldh* (*Sesec_01349*), který má z uvedených genů nejnižší expresi. Útlum genové exprese v konečných hodinách kultivace (12. a 16. hodině) je u kmene CO4A společným jevem, důvodem mohou být již zmíněné nepříznivé kultivační podmínky. U kmene ATCC35246 vykazovaly geny *Sesec_01167 (pfk)* a *Sesec_01168 (pyk)* signifikantní nárůst exprese jen ve 12. hodině kultivace (14 % *pfk* a 18 % *pyk*), nejspíše bude kultivace a s ní produkce HA zpožděná. Dále byl zaznamenán nárůst exprese u genu *Sesec_01044 (eno)* o 35 % v 4. hodině kultivace. Gen *Sesec_01349 (ldh)* měl výraznější nárůst exprese ve 12. hodině kultivace. Zde

můžeme pozorovat snižující se míru navýšení exprese u kaskády po sobě jdoucích enzymatických reakcí s časem kultivace: *eno* (35 % ve 4. hodině) → *pyk* (18 % ve 12. hodině) → *ldh* (19 % ve 12. hodině).

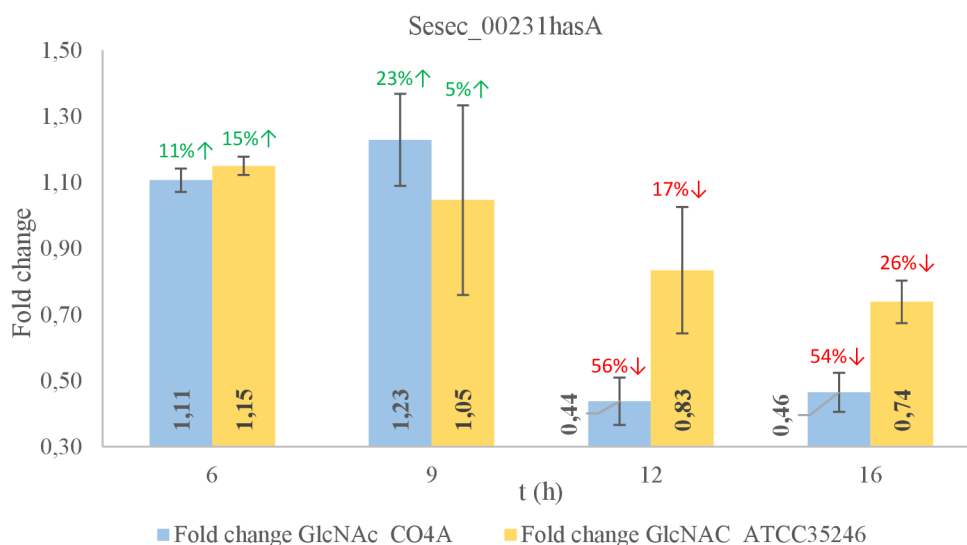
Ústředním aspektem tohoto experimentu bylo sledování vlivu externě dodaného *N*-acetylglukosaminu na změny exprese vybraných genů. Zpracování externě dodaného GlcNAc je zajištěno mimo jiné i geny *Sesec_00293*, *Sesec_00575* (*nagA*) a *Sesec_00795* (*nagB*). Vysoký nárůst exprese v porovnání s kontrolní kultivací byl zaznamenán u genu *Sesec_00293* v 9. hodině kultivace (33 % navýšení). Tento gen zajišťuje první zpracování externího GlcNAc v buňkách. Navazuje na něj gen kódující *nagA* enzym, který přivádí GlcNAc do linie syntézy jednoho z prekurzorů produkce HA. U *nagA* bylo v 9. hodině kultivace zjištěno navýšení exprese o 19 %. Navazuje na něj gen kódující *nagB* enzym, který odvádí metabolit přes fruktózu-6-P směrem do drah glykolýzy. Navýšení jeho exprese bylo v první polovině kultivace signifikantní (23 %). Enzym *glmS* (*Sesec_001170*) kompetuje s enzymem *nagB* a vrací fruktózu-6-P zpět do dráhy syntézy HA. V 9. hodině kultivace dosahovala jeho míra exprese navýšení o 81 %. Z toho vyplývá, že i když aktivita genu vedoucí k odchodu důležitého metabolitu mimo dráhy syntézy HA (*Sesec_00795*, *nagB*) byla vysoká (23 %), exprese genu, kódujícího enzym *glmS*, katalyzujícího opačnou reakci byla vyšší (81 %) a metabolit byl zřejmě intenzivněji přiváděn do drah syntézy HA. Aktivita genů spojených s externím GlcNAc se zvyšovala přibližně do poloviny kultivace, až mezi 9. a 12. hodinou kultivace pravděpodobně dosahovala svého vrcholu. Od 12. hodiny exprese genů zpravidla klesala pod úroveň kontrolní kultivace. U kmene ATCC35246 byl pozorován nárůst exprese genu *Sesec_00293* už ve 4. hodině experimentu. Stejně tak byl zjištěn nárůst ve 4. hodině u genu *Sesec_00575* (*nagA*). Co by naznačovalo, že ke zpracování GlcNAc mohlo docházet dříve než u kmene CO4A. Nicméně pak nastává výrazný propad míry exprese v porovnání s kontrolní kultivací. Gen *Sesec_00795*, (*nagB*) má expresi velmi nízkou po celou dobu kultivace, stejně tak jako gen *Sesec_001170* (*glmS*). Podle všech indicií je velmi pravděpodobné, že externě dodaný GlcNAc u kmene ATCC35246 odchází v podobném množství jak do drah glykolýzy, tak biosyntézy HA.

Z počátku experimentu buňky kmene CO4A zřejmě upřednostňovaly metabolismus aminocukrů a nukleotidů, kam byl GlcNAc směřován genem *nanE* (*Sesec_00296*). *nanE* měl v 9. hodině kultivace navýšení exprese o 57 % v porovnání s kontrolní kultivací. Tento gen má v genomu homologní sekvenci *Sesec_00291*. U analýzy metodou DNA čipů (kap. 5.3.1) byla prokázána zvýšená exprese u obou kopií genu (*Sesec_00296* měl nárůst exprese o 84 % a *Sesec_00291* měl nárůst exprese o 42 %), nicméně to bylo při kultivaci bez suplementace GlcNAc a ve srovnání kmene CO4A vůči kmenu ATCC35246. Navíc v sekvenci *Sesec_00296* byla identifikována mutace (kap. 5.2, tab. 8), která posunula čtecí rámec a pozměnila posledních 55 AMK. Změna mohla mít vliv na expresi genu a v konečném důsledku na funkčnost proteinu. Jeho míra exprese v porovnání s ATCC35246 je velmi vysoká, což mohlo být způsobeno vyšší potřebou enzymu u kmene CO4A.

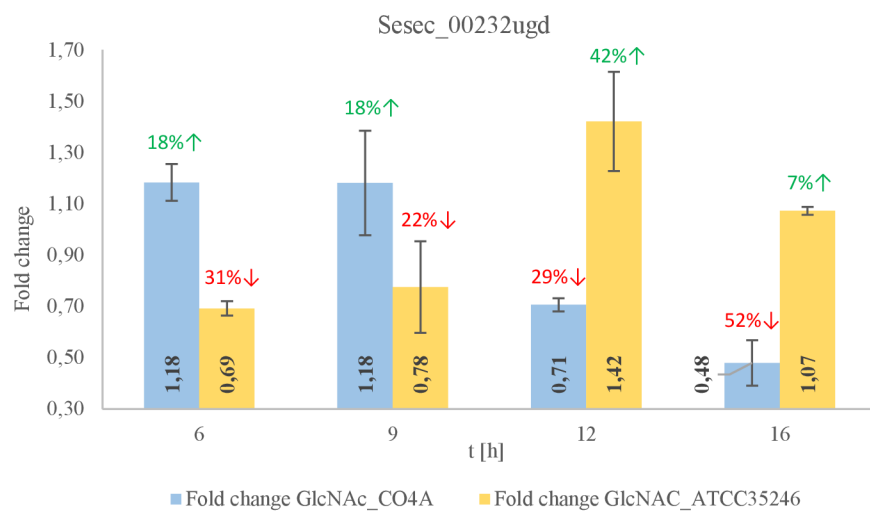
Významný nárůst, v porovnání s kontrolní kultivací, byl pozorován i u genu *Sesec_00830*, který kóduje enzym *murA* odvádějící *N*-acetylglukosamin do drah syntézy peptidoglykanu. V první polovině kultivace (v 9. hodině) byl nárůst exprese tohoto genu u kmene CO4A o 70 %. Peptidoglykan jako stavební jednotka buněčné stěny bude v počátcích kultivace silně upřednostňován i vzhledem k tomu, že gen *Sesec_00231* (*hasA*) měl v 9. hodině kultivace nárůst exprese jen 23 %. Oproti tomu u kmene ATCC35246 dosahoval gen *Sesec_00830*

(*murA*) nárůstu exprese (28 %) v porovnání s kontrolní kultivací až ke konci kultivace (12. hodině). Ve 4. hodině byl jeho nárůst exprese (12 %) porovnatelný s genem kódujícím enzym *hasA* (15 %). Produkce HA může být ovlivněna i odlivem meziproductů do drah syntézy kyseliny lipoteichoové. V tomto směru byly sledovány geny *Sesec_01820*, který měl v 9. hodině kultivace nárůst exprese 38 % a *Sesec_01819 (ugtP)*, který měl v tomtéž čase nárůst exprese 27 %. Lipoteichoová kyselina je stejně jako kyselina hyaluronová jedním z virulentních faktorů a povrchový protein [133, 146].

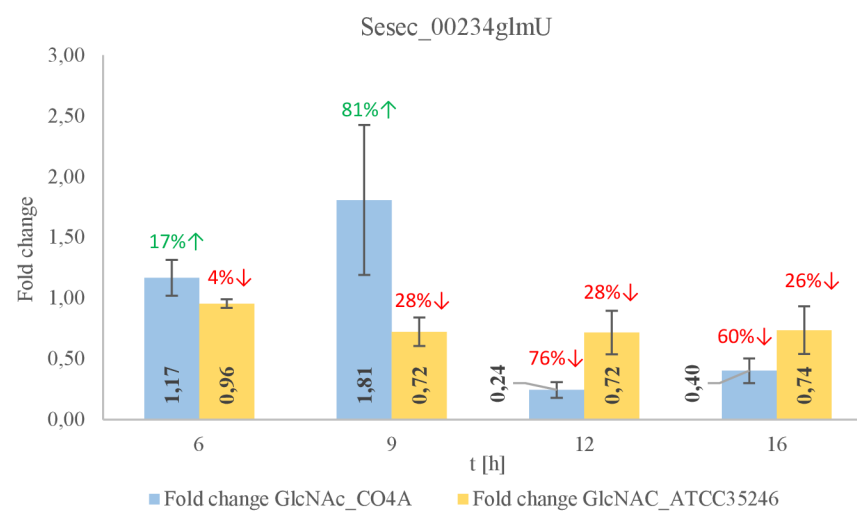
U kmene ATCC35246 měly sledované geny většinou nižší míru exprese ve srovnání s kontrolní kultivací než geny kmene CO4A. Můžeme konstatovat, že u kmene ATCC35246 nebyl pozorován jednoznačný pozitivní vliv dodaného GlcNAc na míru exprese vybraných genů. U kmene CO4A bylo pozorováno navýšení exprese genů *has* operonu v počátcích kultivace. GlcNAc jako prekurzor syntézy HA se po suplementaci do média mohl inkorporovat přímo do HA a navyšovat tak výtěžek v porovnání s kontrolní kultivací bez GlcNAc. Nicméně schopnost inkorporace může být u obou kmenů odlišná. Zatímco u kmene CO4A by mohla větší část suplementovaného GlcNAc přecházet do HA, kmen ATCC35246 ho nebude schopný využít s takovou efektivitou a spíše ho začlení do drah energetického metabolismu. Odlišnost genomická a transkriptomická, kterou v práci studuji u obou kmenů, bude mít vliv na flux GlcNAc. Utilizace tohoto prekurzoru bude u obou kmenů odlišná, předpokladem je vždy určitá genetická predispozice, míra transkripce katabolických a anabolických genů v drahách zpracování GlcNAc a další úrovně regulace. Svůj význam určitě mají podmínky kultivace, povrchové proteiny, plasticita a složení membrány, transportní systémy, rychlost zpracování aj.



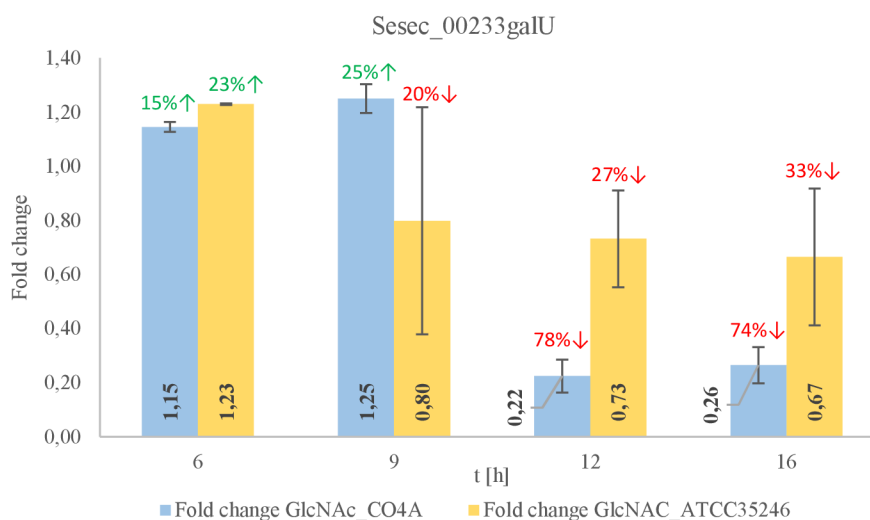
Graf 8: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



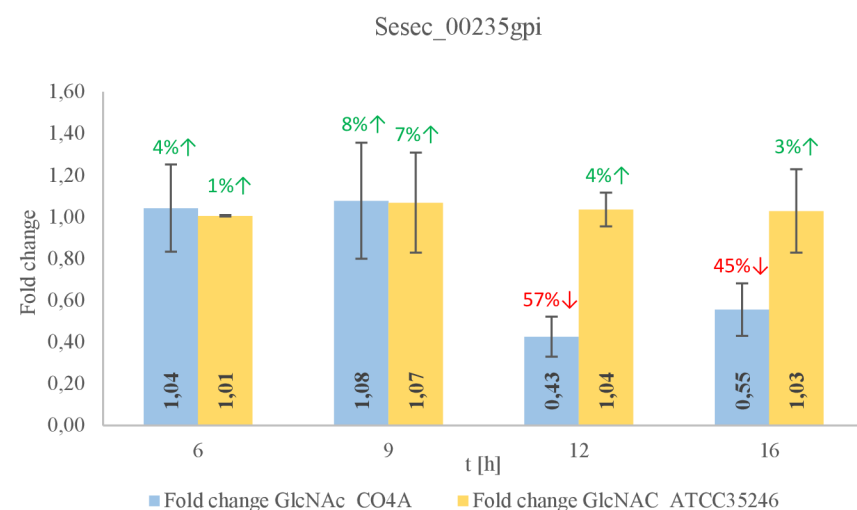
Graf 9: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



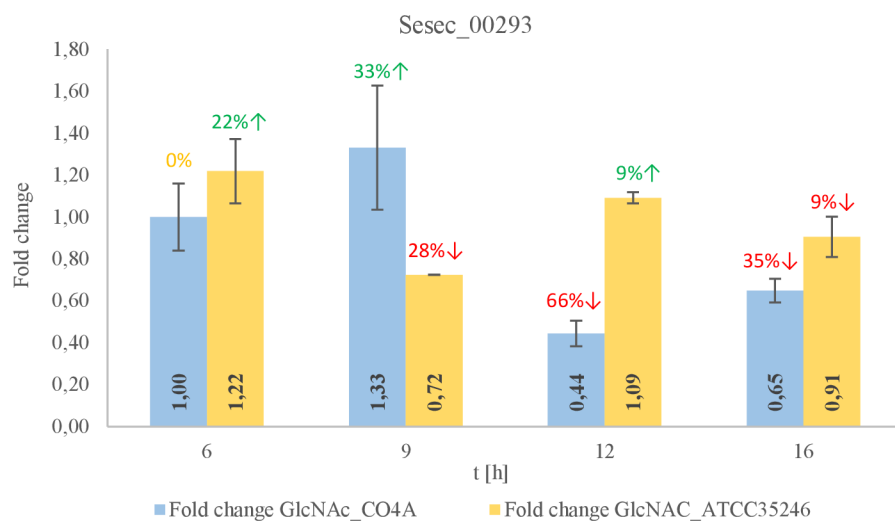
Graf 11: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



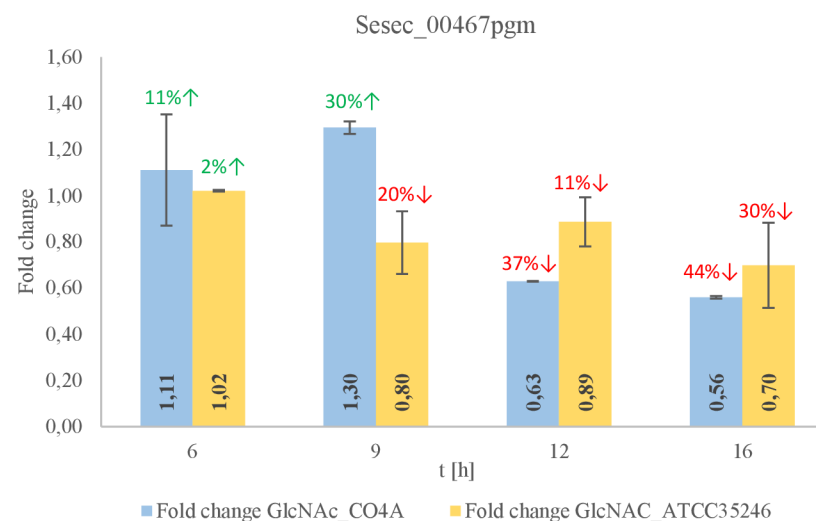
Graf 10: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



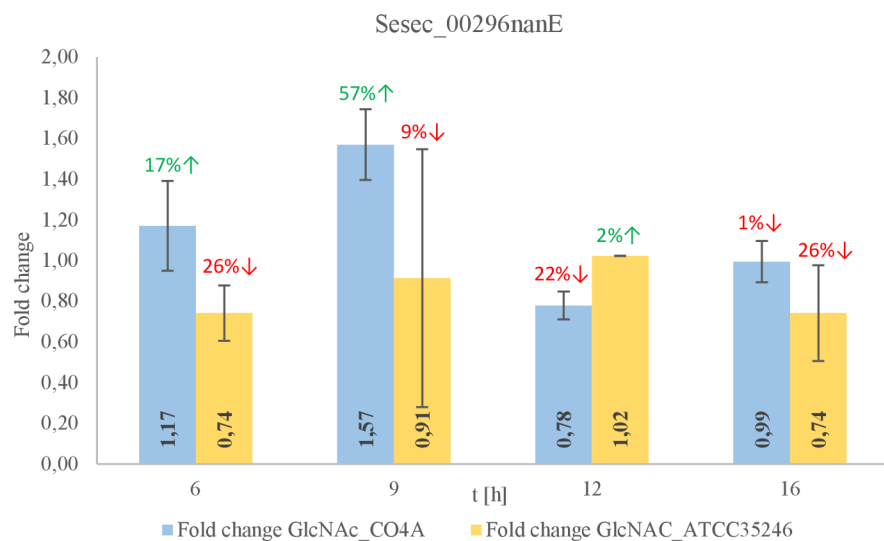
Graf 12: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



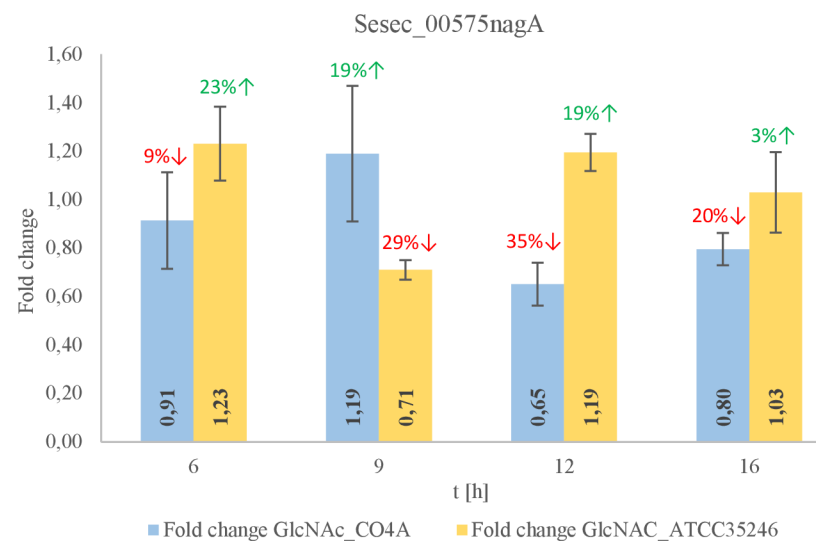
Graf 13: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



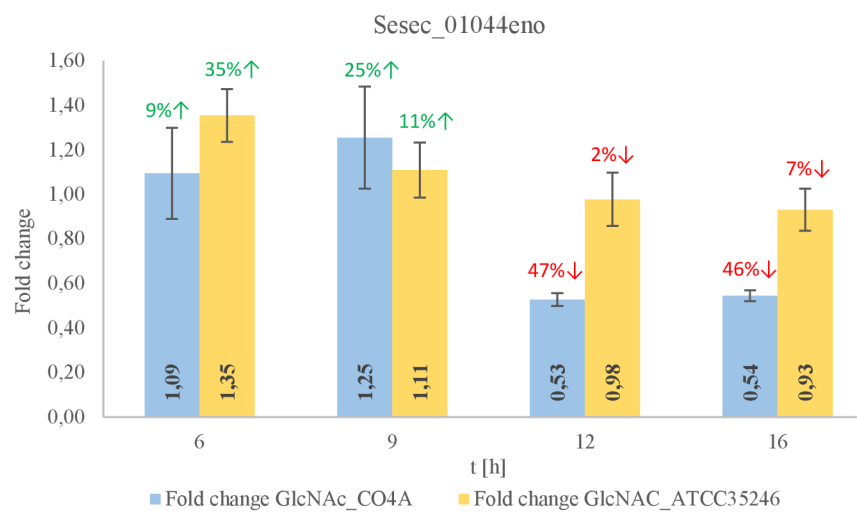
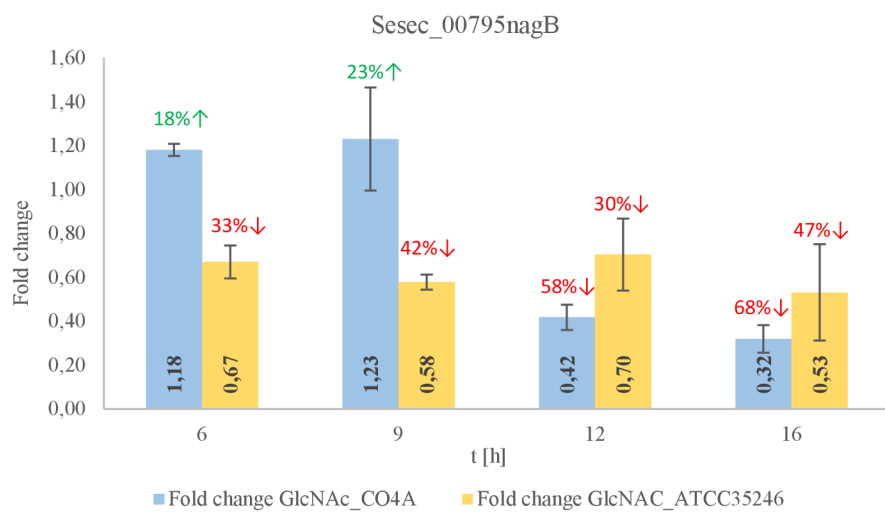
Graf 15: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



Graf 14: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



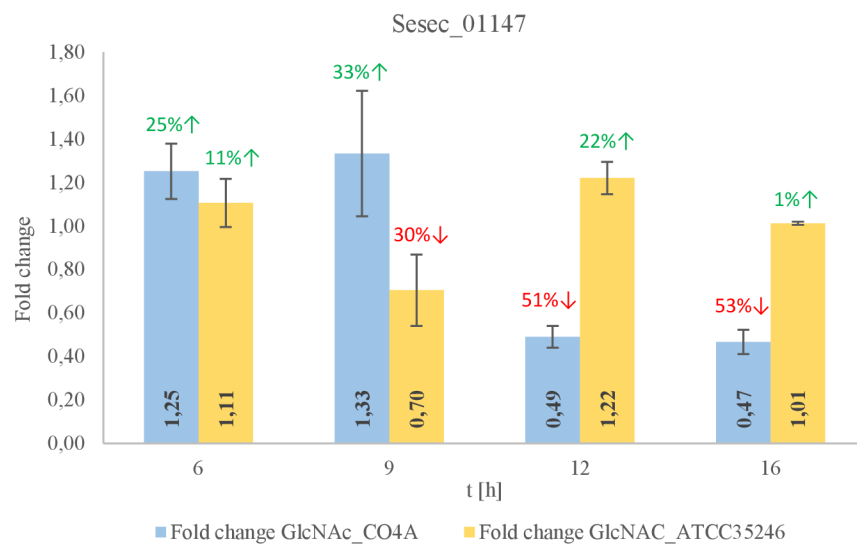
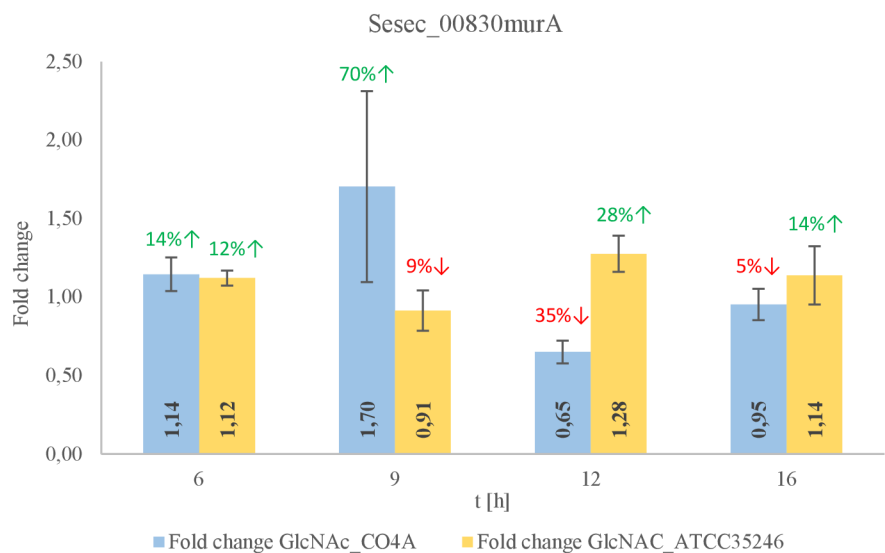
Graf 16: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



Graf 17: Změna relativní míry exprese genu u kmene CO4A a ATCC35246

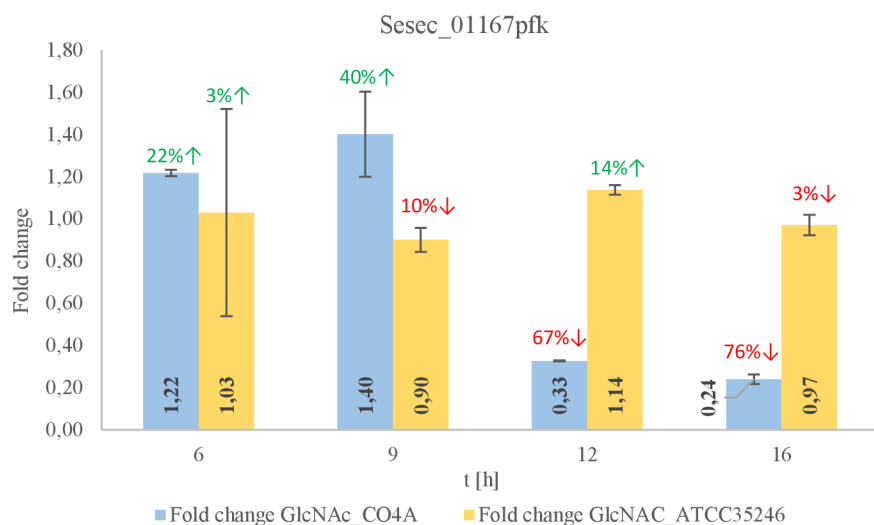
Graf 19: Změna relativní míry exprese genu u kmene CO4A a ATCC35246

96

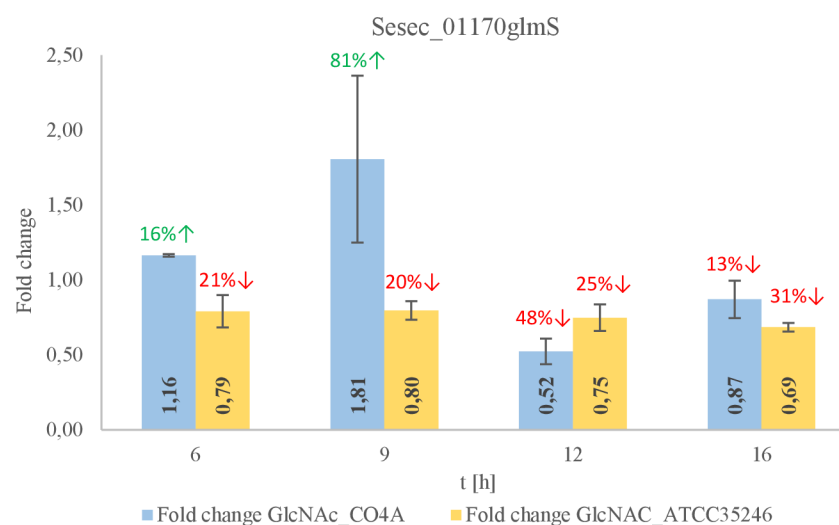


Graf 18: Změna relativní míry exprese genu u kmene CO4A a ATCC35246

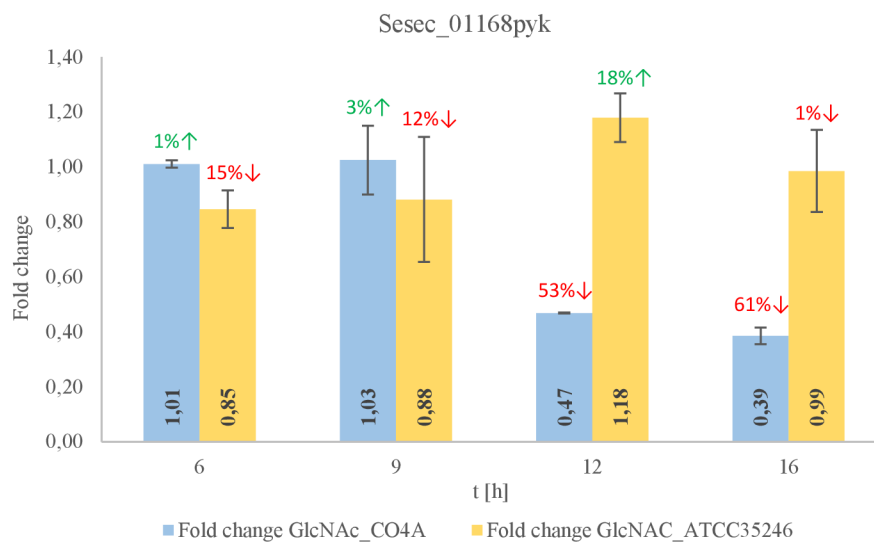
Graf 20: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



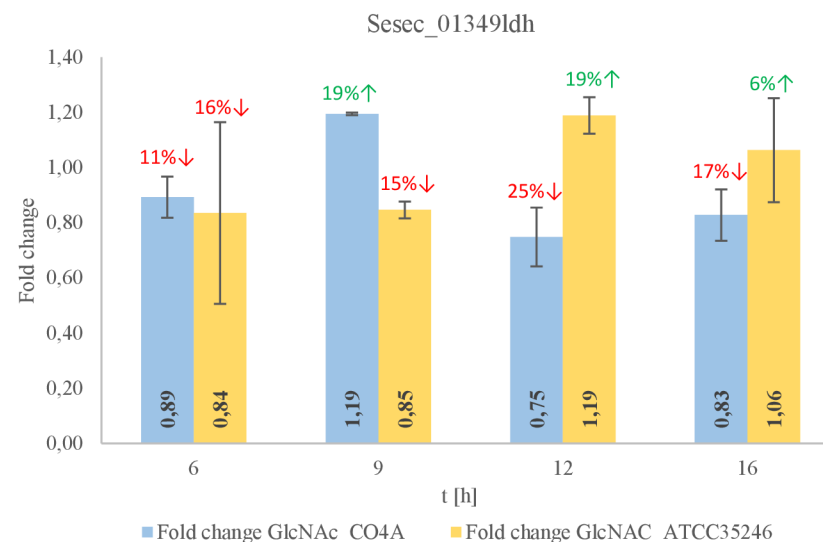
Graf 21: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



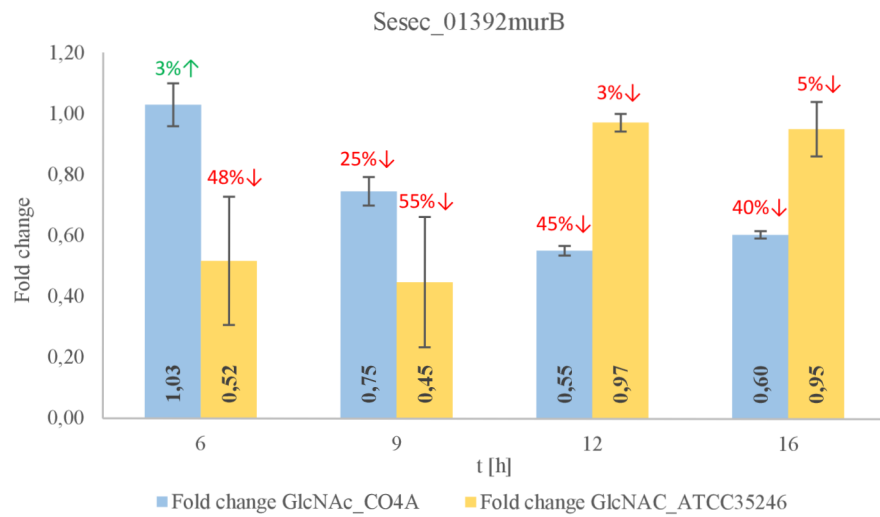
Graf 23: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



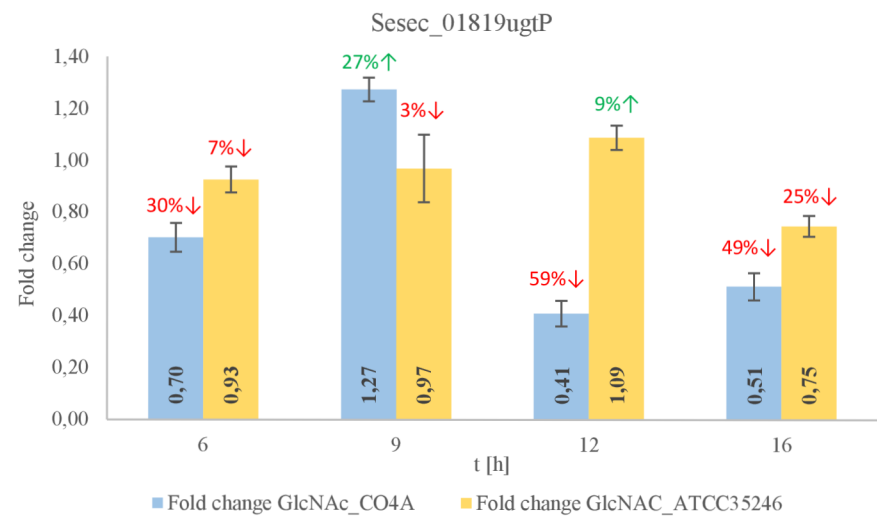
Graf 22: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



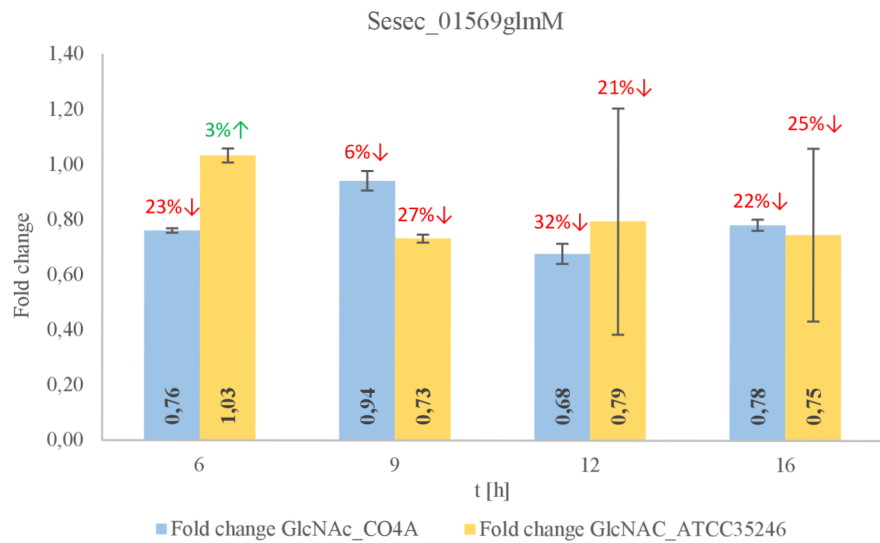
Graf 24: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



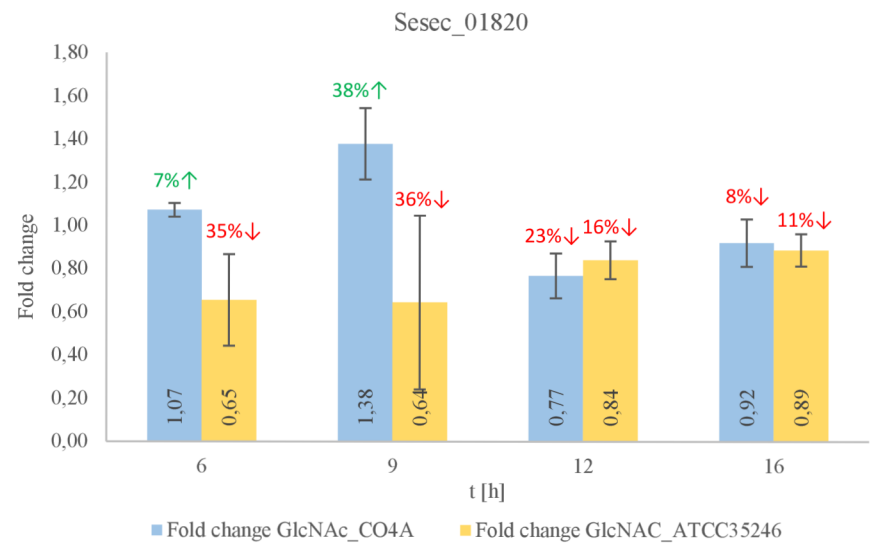
Graf 25: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



Graf 27: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



Graf 26: Změna relativní míry exprese genu u kmene CO4A a ATCC35246



Graf 28: Změna relativní míry exprese genu u kmene CO4A a ATCC35246

5.4.4 Výtěžek HA z provedených kultivací

Po ukončení kultivací bylo zjištěno, jaké množství kyseliny hyaluronové bylo vyprodukováno. K získání výtěžku HA byla použita metoda srážení (kap. 4.1.3). Výchozí množství kultury po fermentaci bylo 250 ml. Výtěžek z objemu byl následně přepočítán na litr média, bylo stanoveno množství sušiny a procentuální nárůst výtěžku HA z kultivace s *N*-acetylglukosaminem v porovnání s kontrolní kultivací. Hodnoty výtěžků kultivací pro kmen ATCC35246 jsou uvedeny v tabulce 30.

Tabulka 30: Výtěžky kyseliny hyaluronové z provedených kultivací u kmene ATCC35246.

Kmen ATCC35246	Kultivace	Výtěžek HA [g/250 ml]	Výtěžek HA [g/l]	Sušina [%]	Výtěžek HA na 100 % sušiny [g]
Fermentor 1	GlcNAc	1,01	4,04	93,28	3,77 (+0,27 g)
Fermentor 2	Kontrola	0,94	3,76	93,16	3,50
Fermentor 3	GlcNAc	1,07	4,28	91,74	3,93 (+0,39 g)
Fermentor 4	Kontrola	0,94	3,76	94,22	3,54
Fermentor 5	GlcNAc	1,02	4,08	94,07	3,84 (+0,33 g)
Fermentor 6	Kontrola	0,94	3,76	93,46	3,51

Z předchozích výsledků qPCR reakcí, kde byla sledována míra a vývoj exprese vybraných genů metabolických drah syntézy HA a drah k nim přidružených, nebyl patrný žádný specifický trend míry exprese. Nemůžeme proto říci, že by *N*-acetylglukosamin měl u kmene ATCC35246 dle hladin exprese jednoznačně pozitivní vliv na produkci HA. Nicméně z výsledků srážecích reakcí je patrný nárůst výtěžku HA na litr média. V případě prvního biologického replikátu³ (fermentor 1 a 2) byl nárůst výtěžku o cca 7 %, u druhého biologického replikátu (fermentor 3 a 4) byl nárůst výtěžku cca o 13 % a v případě třetího biologického replikátu (fermentor 5 a 6) byl nárůst výtěžku v porovnání s kontrolní kultivací cca o 8 %.

Také po ukončení kultivací s kmenem CO4A bylo zjišťováno, jaké množství kyseliny hyaluronové bylo vyprodukováno. Hodnoty výtěžků kultivací jsou uvedeny v tab. 31.

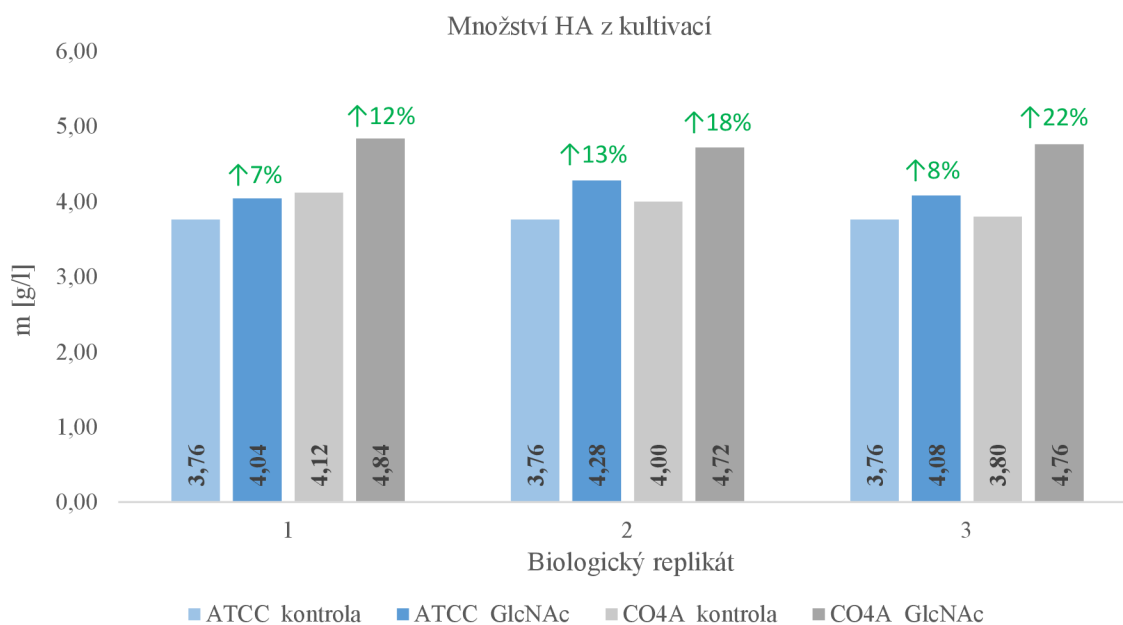
Tabulka 31: V tabulce jsou uvedeny výtěžky kyseliny hyaluronové z provedených kultivací se suplementací *N*-acetylglukosaminem (GlcNAc) v porovnání s kontrolní kultivací bez GlcNAc u kmene CO4A.

Kmen CO4A	Kultivace	Výtěžek HA [g/250 ml]	Výtěžek HA [g/l]	Sušina [%]	Výtěžek na 100 % sušiny [g]
Fermentor 1	GlcNAc	1,21	4,84	90,35	4,37 (+0,51 g)
Fermentor 2	Kontrola	1,03	4,12	93,75	3,86
Fermentor 3	GlcNAc	1,18	4,72	93,85	4,43 (+0,78 g)
Fermentor 4	Kontrola	1,00	4,00	91,26	3,65
Fermentor 5	GlcNAc	1,19	4,76	92,24	4,39 (+0,96 g)
Fermentor 6	Kontrola	0,95	3,80	90,39	3,43

³ Fermentory 1 a 2 byly inokulovány z jedné prekultury, aby byl stav v počátku experimentu srovnatelný jak u kultivace s GlcNAc, tak i u kontrolní kultivace. Stejně bylo inokulováno i v případě dalších fermentorů u kmene ATCC35246 a i u kmene CO4A.

Z předchozích výsledků analýzy metodou qPCR můžeme soudit, že suplementace kultivačního média *N*-acetylglukosaminem měla u produkčního kmene CO4A pozitivní vliv na expresi genů zapojených v produkci HA. Ke stejným závěrům můžeme dojít i z výsledků srážecích reakcí, kde je patrný nárůst výtěžku HA na litr média. V případě prvního biologického replikátu (fermentor 1 a 2) byl nárůst výtěžku o cca 12 %, u druhého biologického replikátu (fermentor 3 a 4) byl nárůst výtěžku cca o 18 % a v případě třetího biologického replikátu (fermentor 5 a 6) byl nárůst výtěžku v porovnání s kontrolní kultivací cca o 22 %.

Srovnání obou kultivací je uvedeno v grafu 29.



Graf 29: Z kultivací byla získána HA srážecí metodou. Z grafu je viditelné, že množství získané kyseliny hyaluronové u kultivací se suplementovaným *N*-acetylglukosaminem byla v porovnání s kontrolními kultivacemi vyšší.

U našeho experimentu s kultivací s přidavkem GlcNAc jsme se inspirovali v článku z roku 2013 [6], kde se zaměřovali na vliv extracelulárních prekurzorů HA na její molekulovou hmotnost. Pracovali mimo jiné i se sbírkovým kmenem *S. zooepidemicus* ATCC35246, který ve své práci používám jako referenční kmen. Zjistili, že přidavek GlcNAc do kultivačního média má pozitivní vliv na množství produkované HA, ale molekulová hmotnost klesá [6]. Nebylo zjištěno signifikantní navýšení exprese u genů zapojených v drahách biosyntézy HA [6]. K jejich seznamu vybraných metabolických drah byly doplněny námi vybrané geny zájmu (na základě dat z analýzy na DNA čípech), aby bylo pokryto co nejvíce možností, které by mohly ovlivnit produkci HA. Došli jsme k závěru, že suplementace kultivace *N*-acetylglukosaminem (prekurzor syntézy HA) může pozitivně ovlivňovat expresi genů ve sledovaných metabolických drahách u kmene CO4A. V porovnání s kontrolní kultivací, kam nebyl dodán *N*-acetylglukosamin, byly hodnoty exprese u produkčního kmene nižší než u kultivací se suplementací GlcNAc. Výtěžek HA z těchto kultivací zaznamenal ve všech třech opakování signifikantní nárůst, a to o 12 %, 18 % a 22 %. Rozdíly ve výtěžku napříč kultivacemi jsou výrazné hlavně v případě první a třetí kultivace. Důvodem byl pravděpodobně stav buněk na počátku kultivace. Inokulováno bylo stejné množství buněk u všech kultivací (určeno

spektrofotometricky), ale vitálnost a celková kvalita kultury se mohly lišit. To mohlo prodloužit lag fázi kultivace a tím zpozdit celou kultivaci. Následkem mohl být nižší výtěžek HA.

U kmene referenčního ATCC35246 jsme jednoznačný trend ve vývoji míry exprese v průběhu kultivace neidentifikovali. I přes to byl zaznamenán nárůst výtěžku HA v porovnání s kontrolními kultivacemi (o 7 %, 13 % a 8 %). V případě druhého replikátu byl nárůst výtěžku HA také signifikantní (13 %). Předpokladem našich experimentů byl vysoce efektivní proces přepisu sekvence mRNA do sekvence proteinu. V reálném systému míru exprese ovlivňují z velké části i procesy předcházející translaci, dále může docházet k degradaci mRNA, nebo může být míra přepisu sekvence do proteinu méně efektivní z důvodu přítomnosti mutací atd. Nicméně na základě získaných dat můžeme konstatovat, že při kultivaci CO4A v médiu s GlcNAc byl zjištěn signifikantní nárůst exprese genů zapojených v metabolismu HA, jenž oproti ATCC35246 mohl zapříčinit vyšší výtěžek HA.

Koncentrace suplementovaného GlcNAc v kultivačním médiu byla 0,5 g/l. Kultivace probíhaly v objemu 0,5 l, tudíž reálné množství extracelulárního GlcNAc, které bylo buňkami využitelné, odpovídá 0,25 g. Pokud budeme uvažovat, že zcela dochází k inkorporaci dodaného GlcNAc do řetězce HA, došlo by ke zvýšení výtěžku HA o 0,5 g (0,25 g GlcNAc + 0,5 g GlcUA). Kmen ATCC35246 vykazoval navýšení produkce HA v průměru o cca 0,33 g, což by znamenalo využití 0,165 g GlcNAc (66 %) z 0,25 g dostupného GlcNAc. Jsme tedy v rozmezí konverze 0-100 % a navýšení produkce HA mohlo být způsobeno pouze dodáním extracelulárního GlcNAc do média, Zbylé množství suplementu může být zároveň zpracováváno i jinými metabolickými drahami, např. glykolýzou.

U kmene CO4A však došlo v průměru k navýšení produkce HA o cca 0,75 g, což by znamenalo využití 0,375 g GlcNAc z 0,25 g dostupného GlcNAc. Zde by se tedy konverze dodaného GlcNAc pohybovala kolem 150 %. V případě kmene CO4A proto můžeme předpokládat, že se kromě konverze GlcNAc také uplatňuje jeho další regulační efekt na dráhy produkce HA či flux většího množství prekurzorů. Toto by mohlo být u CO4A způsobeno právě depresí metabolické dráhy syntézy kyseliny *N*-acetylneuraminové [140, 141, 142], z důvodu vnesené inzerce TA v sekvenci kódující predikovanou *N*-acetylmannosamin-6-fosfát-2-epimerázu (*nanE* gen). V této dráze je iniciační molekulou GlcNAc [141]. Pokud tedy jedna z cest zpracování GlcNAc není funkční, může to vést k posílení zbylých drah a to by mohlo vysvětlovat i zvýšenou expresi genů, které tento cukr zpracovávají. Navíc vidíme, že suplementace GlcNAc má u sledovaných genů kmene CO4A pozitivní efekt na nárůst exprese, což u původního kmene ATCC35246 nepozorujeme a zvyšuje se i exprese genu *nanE* (Sesec_00296).

6 ZÁVĚR

Práce se zabývala studiem exprese genů za použití fyzikálně chemických metod (WGS, NGS, qPCR, DNA čipy) na poli molekulární biologie. Studován byl produkční kmen *S. zooepidemicus* CO4A, který využívá společnost Contipro a.s. při komerční přípravě kyseliny hyaluronové pro kosmetický a farmaceutický průmysl. Jako referenční kmen byl použit sbírkový kmen *S. zooepidemicus* ATCC35246 [1].

Genom *S. zooepidemicus* CO4A byl úspěšně sekvenován a za použití specifického softwaru byl sestaven do kompletní sekvence. Výsledkem byl genom o velikosti 2 167 251 párů bází a v něm bylo identifikováno 1 999 kódujících oblastí. Obsah bází GC byl 41,7 %. Analýzou získaného genomu bylo detekováno 83 relevantních variant s frekvencí vyšší než 90 %. Z toho se jednalo v 59 případech o SNV mutace a ve 24 případech o indel mutace. U SNV mutací došlo ke změně aminokyselinové sekvence v 45 případech, u indel mutací byla změna aminokyselinové sekvence identifikována vždy spolu s posunem čtecího rámce. Při analýze genomu jsme se zaměřili na geny *has* operonu, kde byla nalezena mutace v kódující oblasti genu *Sesec_00234 (glmU)*. Byla identifikována také SNV mutace v kódující oblasti genu *Sesec_00255 (galU)*, který je druhou kopií genu *Sesec_00233 (galU)* nacházející se v *has* operonu.

Při studiu exprese genů celého genomu metodou DNA čipů jsme se zaměřili na skupiny specifických metabolických drah. Byly to dráhy energetického metabolismu (glykolýza, glukoneogeneze, dráhy zpracování glutaminu a glutamátu a metabolické dráhy aminocukrů a nukleotidů), dráhy syntézy buněčné stěny (syntéza peptidoglykanu a dráhy glycerolipidů), dráhy replikace DNA a degradace RNA, dráhy produkce HA (*has* operon) a vybrané geny potenciálně ovlivňující její produkci.

U drah energetického metabolismu buňky (dráhy metabolismu aminocukrů) byla zjištěna zajímavá kaskáda genů s vysokou mírou exprese v porovnání s referenčním kmenem, která vedla k syntéze UDP-GlcNAc. Kaskáda zpracovává substrát fruktózu-6-P, která přichází z glykolýzy, nebo je získávána také z extracelulární sacharózy jako uhlíkatého zdroje pro kultivaci. Posloupnost genů a jejich navýšení míry exprese bylo následující: *Sesec_01170 (glmS)*, navýšení o 37 % → *Sesec_01569 (glmM)*, navýšení o 20 % → *Sesec_01989/00234 (glmU)*, navýšení o 30 % → *Sesec_01989/00234 (glmU)*, navýšení o 28 %. Gen *glmU* kóduje bifunkční enzym a katalýza probíhá za účasti C-terminální domény, nebo za účasti N-terminální domény enzymu, podle toho jakou reakci katalyzuje. Výsledkem byla aktivovaná forma UDP-*N*-acetylglukosamin [145], který je prekurzorem syntézy HA a jeho vyšší koncentrace pozitivně ovlivňovala její produkci. Metabolická dráha syntézy peptidoglykanu je přímo kompetující dráha k syntéze HA a odebírá UDP-GlcNAc pocházející z drah metabolismu aminocukrů a nukleotidů. Výběr drah, ve kterých se bude prekurzor utilizovat záleží na kultivačních podmínkách a délce kultivace, regulačních faktorech aj.

V drahách metabolismu glycerolipidů byla identifikována zajímavá kaskáda genů se zvýšenou expresí vedoucí k syntéze lipoteichoové kyseliny. Kyselina lipoteichoová (LTA) je součástí buněčné stěny, ale jen ve výrazně menší míře než peptidoglykan. Je to amfifilní molekula vyskytující se u Gram-pozitivních bakterií a regulátor autolytických enzymů buněčné stěny (muramidáz). Také je jedním z virulentních faktorů a povrchových proteinů u streptokoků [146]. HA se dostává ven z buněk přes membránu a formuje kapsuli obalující buňku. Změny

povrchových proteinů mohou mít následky například v podobě nepřilnutí HA k povrchu buněk, případně mohou mít pozitivní či negativní efekt při průchodu HA skrz buněčnou membránu.

Metodou DNA čipů lze sledovat expresi velkého množství genů současně. Nevýhodou však je její nedostatečná citlivost a vyskytující se falešně pozitivní výsledky. Proto byly geny, které se podílejí na syntéze HA, sledovány i za pomoci metody qPCR. Prvním krokem bylo stanovení referenčního genu (*gyrA*), ke kterému se vztahovaly výsledné hodnoty C_t . U genů *has* operonu kmene CO4A a kmene ATCC35246 nebyla zjištěna signifikantní změna exprese genů. Celkově můžeme říct, že exprese genů *has* operonu u obou sledovaných kmenů je srovnatelná. Bakteriální buňky kmene ATCC35246 však produkují méně HA než buňky produkčního kmene CO4A (kap. 5.4.4). Biosyntéza kyseliny hyaluronové byla zřejmě regulována i jiným mechanismem než jen genovou expresí.

Jako další byl sledován vliv suplementace GlcNAc do kultivačního média. U kmene ATCC35246 měly sledované geny většinou nižší, nebo srovnatelnou míru exprese ve srovnání s kontrolní kultivací než geny kmene CO4A. Můžeme konstatovat, že u kmene ATCC35246 nebyl pozorován jednoznačný pozitivní vliv dodaného GlcNAc na míru exprese vybraných genů. I přes to byl pozorován nárůst výtěžku HA v porovnání s kontrolními kultivacemi, a to o 7 %, 13 % a 8 %. U kmene CO4A docházelo k navýšení míry exprese genů *has* operonu a tudíž můžeme říci, že suplementace měla pozitivní dopad na vývoj jejich exprese. U výtěžku HA z těchto kultivací byl zaznamenán signifikantní nárůst ve všech třech opakování, a to o 12 %, 18 % a 22 %. Utilizace tohoto prekurzoru bude u obou kmenů odlišná, předpokladem je vždy určitá genetická predispozice, míra transkripce katabolických a anabolických genů v drahách zpracování GlcNAc a další úrovně regulace. Svůj význam určitě mají podmínky kultivace, povrchové proteiny, plasticita a složení membrány, transportní systémy, rychlost zpracování aj. Experiment by se v budoucnu mohl rozšířit i o druhý prekurzor GlcUA.

U kmene CO4A měl GlcNAc zřejmě i signální a regulační funkci. Vzhledem k tomu, že koncentrace GlcNAc byla 0,5 g/l a kultivace probíhala v objemu 0,5 l, bude celkové množství dostupného GlcNAc 0,25 g. V průměru došlo u kultivací CO4A k nárůstu výtěžku HA o cca 0,75 g, což by znamenalo využití 0,375 g GlcNAc z 0,25 g dostupného GlcNAc (150 %). Konverze GlcNAc je vyšší než 100 %. Můžeme se domnívat, že byl konvertován všechny externě dodaný GlcNAc a navíc tato suplementace vyvolala další navýšení produkce HA. Uplatňuje se tudíž také jeho další regulační efekt GlcNAc na dráhy produkce HA či flux většího množství prekurzorů. Toto by mohlo být u CO4A způsobeno právě depresí metabolické dráhy syntézy kyseliny *N*-acetylneuraminové [140, 141, 142], z důvodu vnesené inzerce TA v sekvenci kódující gen *nanE*. Pokud tedy jedna z cest zpracování GlcNAc není funkční, může to vést k posílení zbylých drah a to by mohlo vysvětlovat i zvýšenou expresi genů, které tento cukr zpracovávají. U kmene ATCC35246 došlo v průměru k navýšení produkce HA o cca 0,33 g, což by znamenalo využití 0,165 g GlcNAc z 0,25 g dostupného GlcNAc (66 %). Jsme tedy v rozmezí konverze 0-100 %, což značí, že navýšení produkce HA bylo pravděpodobně způsobeno pouze konverzí extracelulárního GlcNAc z média.

V průběhu provádění experimentů bylo získáno velké množství dat, ze kterých lze vycházet při dalších experimentech. Je možné se zaměřit přímo na regulační a signální dráhy zapojené v produkci HA a v přidružených drahách. Může být prostudována skupina povrchových proteinů buněk, které mohou mít vliv na uchycení kapsule na buňkách, nebo na transport HA ven z buněk. Metodou DNA čipů byla sledována celogenomová exprese. Ze získaných dat je

také možné vytipovat další kandidáty z řad genů, které by mohly ovlivňovat syntézu HA. Například pro bližší studium by byl zajímavý gen *nanE*. Měl velmi vysokou míru exprese (až 84 %) u všech analýz a provedených experimentů. Je jedním z centrálních genů metabolismu aminocukrů a biosyntézy kyseliny neuraminové. Byla u něj identifikována mutace a má svou druhou kopii v genomu (Sesec_00296, Sesec_00291). Dále by byl zajímavý gen Sesec_00255 (*galU*), který je druhou kopií genu Sesec_00233 v *has* operonu. Byla v něm identifikována také mutace SNV v kódující oblasti (kap. 5.2, tab. 8). Další kopie genu měla zřejmě zajistit stabilní množství enzymů nutných k produkci HA [31].

V práci bylo dosaženo stanovených cílů. Byla zjištěna sekvence genomu a byly identifikovány změny, které vznikly působením fyzikálního mutagenu (UV záření). Byl odhalen pozitivní vliv suplementace jedním z prekurzorů HA (GlcNAc) a byl potvrzen i jeho pozitivní vliv na výtěžek HA. Na práci lze navázat a čerpat ze získaných dat při dalších analýzách a experimentech.

7 PŘEHLED LITERATURY

- [1] MA, Z., J. GENG, H. ZHANG, et al. Complete Genome Sequence of *Streptococcus equi* subsp. *zooepidemicus* Strain ATCC 35246. *Journal of Bacteriology* [online]. 2011, **193**(19), 5583-5584 [cit. 2020-02-18]. DOI: 10.1128/JB.05700-11. ISSN 0021-9193. Dostupné z: <http://jb.asm.org/cgi/doi/10.1128/JB.05700-11>
- [2] KIM, Jeong-Hyun, Seung-Jong YOO, Deok-Kun OH, Young-Gi KWEON, Dong-Woo PARK, Chul-Hoon LEE a Gwang-Hoon GIL. Selection of a *Streptococcus equi* mutant and optimization of culture conditions for the production of high molecular weight hyaluronic acid. *Enzyme and Microbial Technology* [online]. 1996, **19**(6), 440-445 [cit. 2020-05-26]. DOI: 10.1016/S0141-0229(96)00019-1. ISSN 01410229. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S0141022996000191>
- [3] YU, Huimin a Gregory STEPHANOPOULOS. Metabolic engineering of *Escherichia coli* for biosynthesis of hyaluronic acid. *Metabolic Engineering* [online]. 2008, **10**(1), 24-32 [cit. 2020-05-26]. DOI: 10.1016/j.ymben.2007.09.001. ISSN 10967176. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1096717607000493>
- [4] KRAHULEC, Ján a Jana KRAHULCOVÁ. Increase in hyaluronic acid production by *Streptococcus equi* subsp. *zooepidemicus* strain deficient in β -glucuronidase in laboratory conditions. *Applied Microbiology and Biotechnology*. 2005, vol. 71, issue 4, s. 415-422. DOI: 10.1007/s00253-005-0173-9.
- [5] KRAHULEC, Ján, Marcela TLUSTÁ, Stanislav STUHLÍK a Ján TURŇA. Structure of the has Operon Promoter and the Effect of Mutations on the has Promoter Strength in *Streptococcus equi* subsp. *zooepidemicus*. *Molecular Biotechnology*. 2011, vol. 49, issue 2, s. 166-175. DOI: 10.1007/s12033-011-9388-4. Dostupné z: <http://link.springer.com/10.1007/s12033-011-9388-4>
- [6] CHEN, Wendy Yiting, Esteban MARCELLIN, Jennifer A. STEEN a Lars Keld NIELSEN. The Role of Hyaluronic Acid Precursor Concentrations in Molecular Weight Control in *Streptococcus zooepidemicus*. *Molecular Biotechnology* [online]. 2014, **56**(2), 147-156 [cit. 2018-07-01]. DOI: 10.1007/s12033-013-9690-4. ISSN 1073-6085. Dostupné z: <http://link.springer.com/10.1007/s12033-013-9690-4>
- [7] SARANRAJ, P a M.A NAIDU. Hyaluronic acid production and its applications: a review. *International Journal of Pharmaceutical and Biological Archives*. 2013, 4 (5), s. 853-859.
- [8] LIU, Long, Yanfeng LIU, Jianghua LI, Guocheng DU a Jian CHEN. Microbial production of hyaluronic acid: current state, challenges, and perspectives. *Microbial Cell Factories*. 2011, vol. 10, issue 1. DOI: 10.1186/1475-2859-10-99.
- [9] Armstrong DC, Johns MR. Culture Conditions Affect the Molecular Weight Properties of Hyaluronic Acid Produced by *Streptococcus zooepidemicus*. *Appl Environ Microbiol*. 1997 Jul;63(7):2759-64. PMID: 16535649; PMCID: PMC1389204.
- [10] CHONG, Barrie Fong, Lars M. BLANK, Richard MCLAUGHLIN a Lars K. NIELSEN. Microbial hyaluronic acid production. *Applied Microbiology and Biotechnology*. 2005, vol. 66, issue 4, s. 341-351. DOI: 10.1007/s00253-004-1774-4. Dostupné z: <http://link.springer.com/10.1007/s00253-004-1774-4>
- [11] VASVANI, Shyam, Pratik KULKARNI a Deepak RAWTANI. Hyaluronic acid: A review on its biology, aspects of drug delivery, route of administrations and a special emphasis on its approved marketed products and recent clinical studies. *International Journal of Biological Macromolecules* [online]. 2020, **151**, 1012-1029 [cit. 2020-05-25]. DOI: 10.1016/j.ijbiomac.2019.11.066. ISSN 01418130. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S014181301936547X>
- [12] KOGAN, Grigorij, Ladislav ŠOLTÉS, Robert STERN, Peter GEMEINER, J. SCHILLER, N. VOLPI, E. HRABÁROVÁ a L. ŠOLTÉS. Hyaluronic acid: a natural biopolymer with a broad range of biomedical and industrial applications. *Biotechnology Letters*. 2006, vol. 29, issue 1, s. 1-34. DOI: 10.1002/9781118164792.ch1.

- [13] LIU, L., M. WANG, G. DU a J. CHEN. Enhanced hyaluronic acid production of *Streptococcus zooepidemicus* by an intermittent alkaline-stress strategy. *Letters in Applied Microbiology*. 2008, vol. 46, issue 3, s. 383-388. DOI: 10.1111/j.1472-765x.2008.02325.x.
- [14] Meyer, K. and Palmer, J.W. The polysaccharide of the vitreous humor. *Journal of Biological Chemistry*. 1934, 107, 629-634.
- [15] LARSEN, Nancy E., Edward A. LESHCHINER, Edward G. PARENT, Julie HENDRIKSON-AHO, Endre A. BALAZS a Sadek K. HILAL. Hylan gel composition for percutaneous embolization. *Journal of Biomedical Materials Research*. 1991, vol. 25, issue 6, s. 699-710. DOI: 10.1002/jbm.820250602. Dostupné z: <http://doi.wiley.com/10.1002/jbm.820250602>
- [16] LAURENT, Torvard C, Ulla BG LAURENT a J Robert E FRASER. The structure and function of hyaluronan: An overview. *Immunology and Cell Biology*. 1996, vol. 74, issue 2, A1-A7. DOI: 10.1038/icb.1996.32.
- [17] BALAZS, E. A.; SWEENEY, D. B.; MCPHERSON, A. New and controversial aspects of retinal detachment. by McPherson, A., Harper and Row, New York, 1968, 371.
- [18] CHONG, B a L NIELSEN. Amplifying the cellular reduction potential of. *Journal of Biotechnology*. 2003, vol. 100, issue 1, s. 33-41. DOI: 10.1016/s0168-1656(02)00239-0.
- [19] CHIEN, L.-J. a C.-K. LEE. Enhanced hyaluronic acid production in *Bacillus subtilis* by coexpressing bacterial hemoglobin. *Biotechnology Progress*. 2007, vol. 0, issue 0. DOI: 10.1021/bp070036w.
- [20] WIDNER, R. BEHR, S. Von DOLLEN, M. TANG, T. HEU, A. SLOMA, D. STERNBERG, P. L. DEANGELIS, P. H. WEIGEL a S. BROWN. Hyaluronic Acid Production in *Bacillus subtilis*. *Applied and Environmental Microbiology*. 2005, vol. 71, issue 7, s. 3747-3752. DOI: 10.1128/aem.71.7.3747-3752.2005.
- [21] CHIEN, Liang-Jung a Cheng-Kang LEE. Hyaluronic acid production by recombinant *Lactococcus lactis*. *Applied Microbiology and Biotechnology*. 2007, vol. 77, issue 2, s. 339-346. DOI: 10.1007/s00253-007-1153-z.
- [22] MAO, Z. a R.R. CHEN. Recombinant Synthesis of Hyaluronan by *Agrobacterium* sp. *Biotechnology Progress*. 2007, vol. 0, issue 0. DOI: 10.1021/bp070113n.
- [23] YU, Huimin, Gregory STEPHANOPOULOS a Rainer M. FIGGE. Metabolic engineering of *Escherichia coli* for biosynthesis of hyaluronic acid. *Metabolic Engineering*. 2008, vol. 10, issue 1, s. 163-193. DOI: 10.1007/7171_2006_059.
- [24] Kitchen JR, Cysyk RL. Synthesis and release of hyaluronic acid by Swiss 3T3 fibroblasts. *Biochemical Journal*. 1995;309(Pt 2):649-656.
- [25] HARRINGTON, Dean J, Iain C SUTCLIFFE a Neil CHANTER. The molecular basis of *Streptococcus equi* infection and disease. *Microbes and Infection* [online]. 2002, 4(4), 501-510 [cit. 2020-05-03]. DOI: 10.1016/S1286-4579(02)01565-4. ISSN 12864579. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1286457902015654>
- [26] SEGURA, Mariela. Extracellular virulence factors of streptococci associated with animal diseases. *Frontiers in Bioscience* [online]. 2004, 9(1-3) [cit. 2020-05-03]. DOI: 10.2741/1287. ISSN 10939946. Dostupné z: <http://www.bioscience.org/2004/v9/af/1287/list.htm>
- [27] CLEARY, P. PATRICK; LARKIN, ANNICE. Hyaluronic acid capsule: strategy for oxygen resistance in group A streptococci. *Journal of bacteriology*, 1979, 140.3: 1090-1097.
- [28] ZHANG, Jinyu, Ning HAO, Guo-Qiang CHEN a A.-P. TAIRI. Effect of expressing polyhydroxybutyrate synthesis genes (phbCAB) in *Streptococcus zooepidemicus* on production of lactic acid and hyaluronic acid: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Applied Microbiology and Biotechnology*. 2006, vol. 71, issue 2, s. 222-227. DOI: 10.1007/s00253-005-0164-x. Dostupné z: <http://link.springer.com/10.1007/s00253-005-0164-x>

- [29] PRASAD, Shashi Bala, K. B. RAMACHANDRAN a Guhan JAYARAMAN. Transcription analysis of hyaluronan biosynthesis genes in *Streptococcus zooepidemicus* and metabolically engineered *Lactococcus lactis*. *Applied Microbiology and Biotechnology*. 2012, vol. 94, issue 6, s. 1593-1607. DOI: 10.1007/s00253-012-3944-0. Dostupné z: <http://link.springer.com/10.1007/s00253-012-3944-0>
- [30] WIDNER, B., R. BEHR, S. Von DOLLEN, M. TANG, T. HEU, A. SLOMA, D. STERNBERG, P. L. DEANGELIS, P. H. WEIGEL a S. BROWN. Hyaluronic Acid Production in *Bacillus subtilis*. *Applied and Environmental Microbiology*. 2005, vol. 71, issue 7, s. 3747-3752. DOI: 10.1128/aem.71.7.3747-3752.2005.
- [31] BLANK, Lars M., Philip HUGENHOLTZ a Lars K. NIELSEN. Evolution of the Hyaluronic Acid Synthesis (has) Operon in *Streptococcus zooepidemicus* and Other Pathogenic Streptococci. *Journal of Molecular Evolution*. 2008, vol. 67, issue 1, s. 13-22. DOI: 10.1007/s00239-008-9117-1. Dostupné z: <http://link.springer.com/10.1007/s00239-008-9117-1>
- [32] CRATER, D. L., I. van de RIJN, Sebastian ALBERTI a Michael R. WESSELS. Hyaluronic Acid Synthesis Operon (has) Expression in Group A Streptococci. *Journal of Biological Chemistry*. 1995, vol. 270, issue 31, s. 975-978. DOI: 10.1007/978-1-4899-1825-3_229.
- [33] ALBERTI, Sebastian, Cameron D. ASHBAUGH, Michael R. WESSELS, Sebastian ALBERTI a Michael R. WESSELS. Structure of the has operon promoter and regulation of hyaluronic acid capsule expression in group A *Streptococcus*. *Molecular Microbiology*. 1998, vol. 28, issue 2, s. 975-978. DOI: 10.1007/978-1-4899-1825-3_229.
- [34] LEVIN, James C. a Michael R. WESSELS. Identification of *csrR/csrS*, a genetic locus that regulates hyaluronic acid capsule synthesis in group A *Streptococcus*. *Molecular Microbiology*. 1998, vol. 30, issue 1, s. 209-219. DOI: 10.1046/j.1365-2958.1998.01057.x.
- [35] WEIGEL, Paul H., David CHIPMAN, Zéev BARAK, Stanislav ENGEL, Sharon MENDEL a Maria VYAZMENSKY. Functional Characteristics and Catalytic Mechanisms of the Bacterial Hyaluronan Synthases. *IUBMB Life (International Union of Biochemistry and Molecular Biology: Life)*. 2002, vol. 54, issue 4, s. 201-211. DOI: 10.1201/9780203913420.ch14.
- [36] WEIGEL, P. H. a P. L. DEANGELIS. Hyaluronan Synthases: A Decade-plus of Novel Glycosyltransferases. *Journal of Biological Chemistry*. 2007-12-14, vol. 282, issue 51, s. 36777-36781. DOI: 10.1074/jbc.R700036200. Dostupné z: <http://www.jbc.org/cgi/doi/10.1074/jbc.R700036200>
- [37] DEANGELIS, Paul L. a Paul H. WEIGEL. Immunochemical Confirmation of the Primary Structure of Streptococcal Hyaluronan Synthase and Synthesis of High Molecular Weight Product by the Recombinant Enzyme. *Biochemistry*. 1994, vol. 33, issue 31, s. 9033-9039. DOI: 10.1021/bi00197a001.
- [38] HELDERMON, C., P. L. DEANGELIS a P. H. WEIGEL. Topological Organization of the Hyaluronan Synthase from *Streptococcus pyogenes*. *Journal of Biological Chemistry*. 2001-01-19, vol. 276, issue 3, s. 2037-2046. DOI: 10.1074/jbc.M002276200. Dostupné z: <http://www.jbc.org/cgi/doi/10.1074/jbc.M002276200>
- [39] KUMARI, K. a P. H. WEIGEL. Molecular Cloning, Expression, and Characterization of the Authentic Hyaluronan Synthase from Group C *Streptococcus equisimilis*. *Journal of Biological Chemistry*. 1997, vol. 272, issue 51, s. 32539-32546. DOI: 10.1074/jbc.272.51.32539.
- [40] WEIGEL, P. H. a B. A. BAGGENSTOSS. Hyaluronan synthase polymerizing activity and control of product size are discrete enzyme functions that can be uncoupled by mutagenesis of conserved cysteines. *Glycobiology*. 2012-08-22, vol. 22, issue 10, s. 1302-1310. DOI: 10.1093/glycob/cws102. Dostupné z: <http://www.glycob.oxfordjournals.org/cgi/doi/10.1093/glycob/cws102>
- [41] TLAPAK-SIMMONS, V. L. Purification and Lipid Dependence of the Recombinant Hyaluronan Synthases from *Streptococcus pyogenes* and *Streptococcus equisimilis*. *Journal of Biological Chemistry*. 1999, vol. 274, issue 7, s. 4239-4245. DOI: 10.1074/jbc.274.7.4239. Dostupné z: <http://www.jbc.org/cgi/doi/10.1074/jbc.274.7.4239>

- [42] KUMARI, K., B. A. BAGGENSTOSS, A. L. PARKER a P. H. WEIGEL. Mutation of Two Intramembrane Polar Residues Conserved within the Hyaluronan Synthase Family Alters Hyaluronan Product Size. *Journal of Biological Chemistry*. 2006-04-21, vol. 281, issue 17, s. 11755-11760. DOI: 10.1074/jbc.M600727200. Dostupné z: <http://www.jbc.org/cgi/doi/10.1074/jbc.M600727200>
- [43] TLAPAK-SIMMONS, V. L., E. S. KEMPNER, B. A. BAGGENSTOSS a P. H. WEIGEL. The Active Streptococcal Hyaluronan Synthases (HASs) Contain a Single HAS Monomer and Multiple Cardiolipin Molecules. *Journal of Biological Chemistry*. 1998-10-02, vol. 273, issue 40, s. 26100-26109. DOI: 10.1074/jbc.273.40.26100. Dostupné z: <http://www.jbc.org/cgi/doi/10.1074/jbc.273.40.26100>
- [44] WEIGEL, Paul H. Hyaluronan Synthase: The Mechanism of Initiation at the Reducing End and a Pendulum Model for Polysaccharide Translocation to the Cell Exterior. *International Journal of Cell Biology* [online]. 2015, 2015, 1-15 [cit. 2020-05-29]. DOI: 10.1155/2015/367579. ISSN 1687-8876. Dostupné z: <https://www.hindawi.com/journals/ijcb/2015/367579/>
- [45] WEIGEL, P. H., V. C. HASCALL a M. TAMMI. Hyaluronan Synthases. *Journal of Biological Chemistry*. 1997-05-30, vol. 272, issue 22, s. 13997-14000. DOI: 10.1074/jbc.272.22.13997. Dostupné z: <http://www.jbc.org/cgi/doi/10.1074/jbc.272.22.13997>
- [46] MEDINA, Andria P, Jialing LIN a Paul H WEIGEL. Hyaluronan synthase mediates dye translocation across liposomal membranes. *BMC Biochemistry*. 2012, vol. 13, issue 1. DOI: 10.1186/1471-2091-13-2.
- [47] WEIGEL, P. H. a B. A. BAGGENSTOSS. Hyaluronan synthase polymerizing activity and control of product size are discrete enzyme functions that can be uncoupled by mutagenesis of conserved cysteines. *Glycobiology* [online]. 2012, 22(10), 1302-1310 [cit. 2020-05-29]. DOI: 10.1093/glycob/cws102. ISSN 0959-6658. Dostupné z: <https://academic.oup.com/glycob/article-lookup/doi/10.1093/glycob/cws102>
- [48] DIDELOT, Xavier, Rory BOWDEN, Daniel J. WILSON, Tim E. A. PETO a Derrick W. CROOK. Transforming clinical microbiology with bacterial genome sequencing. *Nature Reviews Genetics* [online]. 2012, 13(9), 601-612 [cit. 2020-05-26]. DOI: 10.1038/nrg3226. ISSN 1471-0056. Dostupné z: <http://www.nature.com/articles/nrg3226>
- [49] BERTELLI, C. a G. GREUB. Rapid bacterial genome sequencing: methods and applications in clinical microbiology. *Clinical Microbiology and Infection* [online]. 2013, 19(9), 803-813 [cit. 2020-05-26]. DOI: 10.1111/1469-0691.12217. ISSN 1198743X. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1198743X14631979>
- [50] FRASER, C. M., J. D. GOCAYNE, O. WHITE, et al. The Minimal Gene Complement of *Mycoplasma genitalium*. *Science* [online]. 1995, 270(5235), 397-404 [cit. 2020-05-26]. DOI: 10.1126/science.270.5235.397. ISSN 0036-8075. Dostupné z: <https://www.sciencemag.org/lookup/doi/10.1126/science.270.5235.397>
- [51] FLEISCHMANN, R., M. ADAMS, O. WHITE, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* [online]. 1995, 269(5223), 496-512 [cit. 2020-05-26]. DOI: 10.1126/science.7542800. ISSN 0036-8075. Dostupné z: <https://www.sciencemag.org/lookup/doi/10.1126/science.7542800>
- [52] HUSON, D. *Bioinformatics I: DNA sequencing*. Germany, 2010. Dostupné z: <http://ab.inf.uni-tuebingen.de/teaching/ws09/bioinformatics-i/10-sequencing.pdf>. Studijní materiály. University of Tuebingen.
- [53] LIU, Guangjin, Wei ZHANG a Chengping LU. Complete Genome Sequence of *Streptococcus agalactiae* GD201008-001, Isolated in China from Tilapia with Meningoencephalitis. *Journal of Bacteriology* [online]. 2012, 194(23), 6653-6653 [cit. 2020-05-01]. DOI: 10.1128/JB.01788-12. ISSN 0021-9193. Dostupné z: <http://jb.asm.org/lookup/doi/10.1128/JB.01788-12>
- [54] HU, Fen Z., Rory EUTSEY, Azad AHMED, et al. In Vivo Capsular Switch in *Streptococcus pneumoniae* – Analysis by Whole Genome Sequencing. *PLoS ONE* [online]. 2012, 7(11) [cit. 2020-05-01]. DOI: 10.1371/journal.pone.0047983. ISSN 1932-6203. Dostupné z: <https://dx.plos.org/10.1371/journal.pone.0047983>

- [55] ATHEY, T. B. T., S. TEATERO, A. LI, A. MARCHAND-AUSTIN, B. W. BEALL a N. FITTIPALDI. Deriving Group A Streptococcus Typing Information from Short-Read Whole-Genome Sequencing Data. *Journal of Clinical Microbiology* [online]. 2014, **52**(6), 1871-1876 [cit. 2020-05-01]. DOI: 10.1128/JCM.00029-14. ISSN 0095-1137. Dostupné z: <http://jcm.asm.org/cgi/doi/10.1128/JCM.00029-14>
- [56] WATANABE, Shinya, Teppei SASAHARA, Naoshi ARAI, Kazumasa SASAKI, Yoshifumi AIBA, Yusuke SATO a Longzhu CUI. Complete Genome Sequence of Streptococcus pyogenes Strain JMUB1235 Isolated from an Acute Phlegmonous Gastritis Patient. *Genome Announcements* [online]. 2016, **4**(5), e01133-16 [cit. 2020-05-01]. DOI: 10.1128/genomeA.01133-16. ISSN 2169-8287. Dostupné z: <http://genomea.asm.org/lookup/doi/10.1128/genomeA.01133-16>
- [57] RHOADS, Anthony a Kin Fai AU. PacBio Sequencing and Its Applications. *Genomics, Proteomics & Bioinformatics* [online]. 2015, **13**(5), 278-289 [cit. 2020-05-25]. DOI: 10.1016/j.gpb.2015.08.002. ISSN 16720229. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1672022915001345>
- [58] TURCATTI, G., A. ROMIEU, M. FEDURCO a A.-P. TAIRI. A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Nucleic Acids Research*. 2007-12-20, vol. 36, issue 4, e25-e25. DOI: 10.1093/nar/gkn021. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/gkn021>
- [59] SHENDURE, Jay a Hanlee JI. Next-generation DNA sequencing. *Nature Biotechnology*. 2008, vol. 26, issue 10, s. 1135-1145. DOI: 10.1038/nbt1486. Dostupné z: <http://www.nature.com/doi/10.1038/nbt1486>
- [60] Illumina.com. ILLUMINA. *Illumina* [online]. 2015 [cit. 2015-03-05]. Dostupné z: <http://www.illumina.com/neoprep-hp-takeover.html>
- [61] Sequencing-by-Synthesis: Explaining the Illumina Sequencing Technology. In: *BiteSizeBio* [online]. August 30, 2012 [cit. 2020-05-04]. Dostupné z: <https://bitesizebio.com/13546/sequencing-by-synthesis-explaining-the-illumina-sequencing-technology/>
- [62] GREGORY, Simon G. Contig Assembly. *Encyclopedia of Life Sciences*. Chichester, UK: John Wiley, 2001-05-30. DOI: 10.1038/npg.els.0005365. Dostupné z: <http://doi.wiley.com/10.1038/npg.els.0005365>
- [63] A reasonably thorough table of next-gen-seq software available in the commercial and public domain. In: *SEQanswers: The Next Generation Sequencing Community* [online]. Australia, 2008 [cit. 2015-03-23]. Dostupné z: <http://seqanswers.com/forums/showthread.php?t=43>
- [64] LIU, Chia-Hsin a Y. Peter DI. Analysis of RNA Sequencing Data Using CLC Genomics Workbench. KEOHAVONG, Phouthone, Kamaleshwar P. SINGH a Weimin GAO, ed. *Molecular Toxicology Protocols* [online]. New York, NY: Springer US, 2020, 2020-01-28, s. 61-113 [cit. 2020-05-26]. Methods in Molecular Biology. DOI: 10.1007/978-1-0716-0223-2_4. ISBN 978-1-0716-0222-5. Dostupné z: http://link.springer.com/10.1007/978-1-0716-0223-2_4
- [65] HILLER, N. Luisa, Azad AHMED, Evan POWELL, et al. Generation of Genic Diversity among Streptococcus pneumoniae Strains via Horizontal Gene Transfer during a Chronic Polyclonal Pediatric Infection. *PLoS Pathogens* [online]. 2010, **6**(9) [cit. 2020-05-26]. DOI: 10.1371/journal.ppat.1001108. ISSN 1553-7374. Dostupné z: <https://dx.plos.org/10.1371/journal.ppat.1001108>
- [66] DARLING, A. C.E. Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. *Genome Research* [online]. 2004, **14**(7), 1394-1403 [cit. 2020-05-26]. DOI: 10.1101/gr.2289704. ISSN 1088-9051. Dostupné z: <http://www.genome.org/cgi/doi/10.1101/gr.2289704>
- [67] EHRLICH, Garth D., Azad AHMED, Josh EARL, et al. The distributed genome hypothesis as a rubric for understanding evolution in situ during chronic bacterial biofilm infectious processes. *FEMS Immunology & Medical Microbiology* [online]. 2010, **59**(3), 269-279 [cit. 2020-05-26]. DOI: 10.1111/j.1574-695X.2010.00704.x. ISSN 0928-8244. Dostupné z: <https://academic.oup.com/femspd/article-lookup/doi/10.1111/j.1574-695X.2010.00704.x>

- [68] EHRLICH, Garth D, Fen Ze HU, Kai SHEN, Paul STOODLEY a J Christopher POST. Bacterial Plurality as a General Mechanism Driving Persistence in Chronic Infections. *Clinical Orthopaedics and Related Research* [online]. 2005, &NA;(437), 20-24 [cit. 2020-05-26]. DOI: 10.1097/00003086-200508000-00005. ISSN 0009-921X. Dostupné z: <http://gateway.ovid.com/ovidweb.cgi?T=JS&PAGE=crossref&AN=00003086-200508000-00005>
- [69] BIROL, I., S. D. JACKMAN, C. B. NIELSEN, et al. De novo transcriptome assembly with ABySS. *Bioinformatics* [online]. 2009, 25(21), 2872-2877 [cit. 2020-05-25]. DOI: 10.1093/bioinformatics/btp367. ISSN 1367-4803. Dostupné z: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp367>
- [70] SCHMIDT, B., R. SINHA, B. BERESFORD-SMITH a S. J. PUGLISI. A fast hybrid short read fragment assembly algorithm. *Bioinformatics* [online]. 2009, 25(17), 2279-2280 [cit. 2020-05-25]. DOI: 10.1093/bioinformatics/btp374. ISSN 1367-4803. Dostupné z: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp374>
- [71] QUAIL, Michael A., Harold SWERDLOW a Daniel J. TURNER. Improved Protocols for the Illumina Genome Analyzer Sequencing System. *Current Protocols in Human Genetics* [online]. 2009, 62(1) [cit. 2020-06-09]. DOI: 10.1002/0471142905.hg1802s62. ISSN 1934-8266. Dostupné z: <https://onlinelibrary.wiley.com/doi/abs/10.1002/0471142905.hg1802s62>
- [72] AMIN, Shorash, Peter J PRENTIS, Edward K GILDING a Ana PAVASOVIC. Assembly and annotation of a non-model gastropod (*Nerita melanotragus*) transcriptome: a comparison of De novo assemblers. *BMC Research Notes* [online]. 2014, 7(1) [cit. 2020-05-25]. DOI: 10.1186/1756-0500-7-488. ISSN 1756-0500. Dostupné z: <http://bmcresearchnotes.biomedcentral.com/articles/10.1186/1756-0500-7-488>
- [73] ZERBINO, D. R. a E. BIRNEY. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* [online]. 2008, 18(5), 821-829 [cit. 2020-05-01]. DOI: 10.1101/gr.074492.107. ISSN 1088-9051. Dostupné z: <http://www.genome.org/cgi/doi/10.1101/gr.074492.107>
- [74] *National Center for Biotechnology Information: Basic Local Alignment Search Tool* [online]. 8600 Rockville Pike, Bethesda MD, 20894 USA: U.S. National Library of Medicine, 2018 [cit. 2018-07-01]. Dostupné z: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>
- [75] SCHENA, M., D. SHALON, R. W. DAVIS a P. O. BROWN. Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science*. 1995, vol. 270, issue 5235, s. 467-470. DOI: 10.1126/science.270.5235.467
- [76] SOUTHERN, E.M., S.C. CASE-GREEN, J.K. EIDER, M. JOHNSON, K.U. MIR, L. WANG a J.C. WILLIAMS. Arrays of complementary oligonucleotides for analysing the hybridisation behaviour of nucleic acids. *Nucleic Acids Research*. 1994, vol. 22, issue 8, s. 1368-1373. DOI: 10.1093/nar/22.8.1368. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/22.8.1368>
- [77] FODOR, Stephen P. A., Richard P. RAVA, Xiaohua C. HUANG, Ann C. PEASE, Christopher P. HOLMES a Cynthia L. ADAMS. Multiplexed biochemical assays with biological chips. *Nature*. 1993-8-5, vol. 364, issue 6437, s. 555-556. DOI: 10.1038/364555a0. Dostupné z: <http://www.nature.com/doi/10.1038/364555a0>
- [78] SELINGER, Douglas W., Kevin J. CHEUNG, Rui MEI, Erik M. JOHANSSON, Craig S. RICHMOND, Frederick R. BLATTNER, David J. LOCKHART a George M. CHURCH. RNA expression analysis using a 30 base pair resolution *Escherichia coli* genome array. *Nature Biotechnology* [online]. 2000, 18(12), 1262-1268 [cit. 2020-05-26]. DOI: 10.1038/82367. ISSN 1087-0156. Dostupné z: http://www.nature.com/articles/nbt1200_1262
- [79] STEPHENS, Craig. Bacterial cell cycle: Seeing the big picture with microarrays. *Current Biology* [online]. 2001, 11(6), R222-R225 [cit. 2020-05-26]. DOI: 10.1016/S0960-9822(01)00110-5. ISSN 09609822. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S0960982201001105>

- [80] LAUB, M. T., H. H. MCADAMS, T. FELDBLYUM, C. M. FRASER a L. SHAPIRO. Global Analysis of the Genetic Network Controlling a Bacterial Cell Cycle. *Science* [online]. 2000, **290**(5499), 2144-2148 [cit. 2020-05-26]. DOI: 10.1126/science.290.5499.2144. ISSN 0036-8075. Dostupné z: <https://www.sciencemag.org/lookup/doi/10.1126/science.290.5499.2144>
- [81] YE, Rick W., Wang TAO, Laura BEDZYK, Thomas YOUNG, Mario CHEN a Liao LI. Global Gene Expression Profiles of *Bacillus subtilis* Grown under Anaerobic Conditions. *Journal of Bacteriology* [online]. 2000, **182**(16), 4458-4465 [cit. 2020-05-26]. DOI: 10.1128/JB.182.16.4458-4465.2000. ISSN 1098-5530. Dostupné z: <https://JB.asm.org/content/182/16/4458>
- [82] ZIMMER, D. P., E. SOUPENE, H. L. LEE, V. F. WENDISCH, A. B. KHODURSKY, B. J. PETER, R. A. BENDER a S. KUSTU. Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: Scavenging as a defense against nitrogen limitation. *Proceedings of the National Academy of Sciences* [online]. 2000, **97**(26), 14674-14679 [cit. 2020-05-26]. DOI: 10.1073/pnas.97.26.14674. ISSN 0027-8424. Dostupné z: <http://www.pnas.org/cgi/doi/10.1073/pnas.97.26.14674>
- [83] FAWCETT, P., P. EICHENBERGER, R. LOSICK a P. YOUNGMAN. The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences* [online]. 2000, **97**(14), 8063-8068 [cit. 2020-05-26]. DOI: 10.1073/pnas.140209597. ISSN 0027-8424. Dostupné z: <http://www.pnas.org/cgi/doi/10.1073/pnas.140209597>
- [84] HECKER, Michael a Susanne ENGELMANN. Proteomics, DNA arrays and the analysis of still unknown regulons and unknown proteins of *Bacillus subtilis* and pathogenic Gram-positive bacteria. *International Journal of Medical Microbiology* [online]. 2000, **290**(2), 123-134 [cit. 2020-05-26]. DOI: 10.1016/S1438-4221(00)80080-6. ISSN 14384221. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1438422100800806>
- [85] DE SAIZIEU, Antoine, Christophe GARDÈS, Nicholas FLINT, et al. Microarray-Based Identification of a Novel *Streptococcus pneumoniae* Regulon Controlled by an Autoinduced Peptide. *Journal of Bacteriology* [online]. 2000, **182**(17), 4696-4703 [cit. 2020-05-26]. DOI: 10.1128/JB.182.17.4696-4703.2000. ISSN 1098-5530. Dostupné z: <https://JB.asm.org/content/182/17/4696>
- [86] RIMINI, Rebecca, Birger JANSSON, Georg FEGER, et al. Global analysis of transcription kinetics during competence development in *Streptococcus pneumoniae* using high density DNA arrays. *Molecular Microbiology* [online]. 2000, **36**(6), 1279-1292 [cit. 2020-05-26]. DOI: 10.1046/j.1365-2958.2000.01931.x. ISSN 0950382X. Dostupné z: <http://doi.wiley.com/10.1046/j.1365-2958.2000.01931.x>
- [87] PETERSON, Scott, Robin T. CLINE, Hervé TETTELIN, Vasily SHAROV a Donald A. MORRISON. Gene Expression Analysis of the *Streptococcus pneumoniae* Competence Regulons by Use of DNA Microarrays. *Journal of Bacteriology* [online]. 2000, **182**(21), 6192-6202 [cit. 2020-05-26]. DOI: 10.1128/JB.182.21.6192-6202.2000. ISSN 1098-5530. Dostupné z: <https://JB.asm.org/content/182/21/6192>
- [88] REN, B. Genome-Wide Location and Function of DNA Binding Proteins. *Science* [online]. **290**(5500), 2306-2309 [cit. 2020-05-26]. DOI: 10.1126/science.290.5500.2306. ISSN 00368075. Dostupné z: <https://www.sciencemag.org/lookup/doi/10.1126/science.290.5500.2306>
- [89] TUOMANEN, E. I., Regine HAKENBECK, Nadège BALMELLE, Beate WEBER, Christophe GARDÈS, Wolfgang KECK a Antoine DE SAIZIEU. Mosaic Genes and Mosaic Chromosomes: Intra- and Interspecies Genomic Variation of *Streptococcus pneumoniae*. *Infection and Immunity* [online]. 2001, **69**(4), 2477-2486 [cit. 2020-05-26]. DOI: 10.1128/IAI.69.4.2477-2486.2001. ISSN 1098-5522. Dostupné z: <https://IAI.asm.org/content/69/4/2477>
- [90] YE, Rick W, Tao WANG, Laura BEDZYK a Kevin M CROKER. Applications of DNA microarrays in microbial systems. *Journal of Microbiological Methods* [online]. 2001, **47**(3), 257-272 [cit. 2020-05-02]. DOI: 10.1016/S0167-7012(01)00308-6. ISSN 01677012. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S0167701201003086>

- [91] MILLER, Melissa B. a Yi-Wei TANG. Basic Concepts of Microarrays and Potential Applications in Clinical Microbiology. *Clinical Microbiology Reviews* [online]. 2009, 22(4), 611-633 [cit. 2020-05-26]. DOI: 10.1128/CMR.00019-09. ISSN 0893-8512. Dostupné z: <https://CMR.asm.org/content/22/4/611>
- [92] SAIZIEU, Antoine de, Ulrich CERTA, Janet WARRINGTON, Christopher GRAY, Wolfgang KECK a Jan MOUS. Bacterial transcript imaging by hybridization of total RNA to oligonucleotide arrays. *Nature Biotechnology* [online]. 1998, 16(1), 45-48 [cit. 2020-05-26]. DOI: 10.1038/nbt0198-45. ISSN 1087-0156. Dostupné z: <http://www.nature.com/articles/nbt0198-45>
- [93] PEASE, A. Caviani, et al. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proceedings of the National Academy of Sciences*, 1994, 91.11: 5022-5026.
- [94] LIPSHUTZ, Robert J., Stephen P.A. FODOR, Thomas R. GINGERAS a David J. LOCKHART. High density synthetic oligonucleotide arrays. *Nature Genetics*. vol. 21. DOI: 10.1038/4447.
- [95] LOCKHART, David J. a Elizabeth A. WINZELER. *Nature*. vol. 405, issue 6788, s. 827-836. DOI: 10.1038/35015701. Dostupné z: <http://www.nature.com/doi/10.1038/35015701>
- [96] MILLER, M. B., Y.-W. TANG, M. FEDURCO a A.-P. TAIRI. Basic Concepts of Microarrays and Potential Applications in Clinical Microbiology: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Clinical Microbiology Reviews*. 2009-10-12, vol. 22, issue 4, s. 611-633. DOI: 10.1128/CMR.00019-09. Dostupné z: <http://cmr.asm.org/cgi/doi/10.1128/CMR.00019-09>
- [97] DALMA-WEISZHAUSZ, Dennise D., Janet WARRINGTON, Eugene Y. TANIMOTO a C. Garrett MIYADA. [1] The Affymetrix GeneChip® Platform: An Overview. *Clinical Microbiology Reviews*. 2009-10-12, vol. 22, issue 4, s. 3. DOI: 10.1016/S0076-6879(06)10001-4. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S0076687906100014>
- [98] FODOR, S., J. READ, M. PIRRUNG, L STRYER, A. LU a D SOLAS. Light-directed, spatially addressable parallel chemical synthesis: An Overview. *Science*. 1991-02-15, vol. 251, issue 4995, s. 767-773. DOI: 10.1126/science.1990438. Dostupné z: <http://www.sciencemag.org/cgi/doi/10.1126/science.1990438>
- [99] EHRENREICH, Armin. DNA microarray technology for the microbiologist: an overview. *Applied Microbiology and Biotechnology* [online]. 2006, 73(2), 255-273 [cit. 2020-05-02]. DOI: 10.1007/s00253-006-0584-2. ISSN 0175-7598. Dostupné z: <http://link.springer.com/10.1007/s00253-006-0584-2>
- [100] AUSUBEL, Frederick M., Roger BRENT, Robert E. KINGSTON, David D. MOORE, J.G. SEIDMAN, John A. SMITH a Kevin STRUHL, ed. *Current Protocols in Molecular Biology* [online]. Hoboken, NJ, USA: John Wiley & Sons, 2001 [cit. 2020-05-26]. ISBN 0471142727.
- [101] VALASEK, M. A. The power of real-time PCR. *AJP: Advances in Physiology Education*. 2005-09-01, vol. 29, issue 3, s. 151-159. DOI: 10.1152/advan.00019.2005. Dostupné z: <http://ajpadvan.physiology.org/cgi/doi/10.1152/advan.00019.2005>
- [102] MULLIS, Kary B., Tock HING, Y. V, Duska DELIC, Patricia HERNANDEZ-RODRIGUEZ a Arlen GOMEZ. The Unusual Origin of the Polymerase Chain Reaction: Types, Utilities and Limitations. *Scientific American*. 1990, vol. 262, issue 4, s. 56-65. DOI: 10.5772/37450.
- [103] MULLIS, Kary B. a Fred A. FALOONA. [21] Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. s. 335. DOI: 10.1016/0076-6879(87)55023-6. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/0076687987550236>
- [104] OLSON, M, L HOOD, C CANTOR a D BOTSTEIN. A common language for physical mapping of the human genome. *Science*. 1989-09-29, vol. 245, issue 4925, s. 1434-1435. DOI: 10.1126/science.2781285. Dostupné z: <http://www.sciencemag.org/cgi/doi/10.1126/science.2781285>

- [105] HEID, C A, J STEVENS, K J LIVAK a P M WILLIAMS. Real time quantitative PCR. *Genome Research*. 1996-10-01, vol. 6, issue 10, s. 986-994. DOI: 10.1101/gr.6.10.986. Dostupné z: <http://www.genome.org/cgi/doi/10.1101/gr.6.10.986>
- [106] HIGUCHI, Russell, Carita FOCKLER, Gavin DOLLINGER a Robert WATSON. Kinetic PCR Analysis: Real-time Monitoring of DNA Amplification Reactions. *Bio/Technology*. 1993, vol. 11, issue 9, s. 1026-1030. DOI: 10.1038/nbt0993-1026. Dostupné z: <http://www.nature.com/doi/doi/10.1038/nbt0993-1026>
- [107] NAZARENKO, I. A., S. K. BHATNAGAR a R. J. HOHMAN. A closed tube format for amplification and detection of DNA based on energy transfer. *Nucleic Acids Research*. 1997-06-01, vol. 25, issue 12, s. 2516-2521. DOI: 10.1093/nar/25.12.2516. Dostupné z: <http://nar.oxfordjournals.org/lookup/doi/10.1093/nar/25.12.2516>
- [108] RIRIE, Kirk M., Randy P. RASMUSSEN a Carl T. WITWER. Product Differentiation by Analysis of DNA Melting Curves during the Polymerase Chain Reaction. *Analytical Biochemistry*. 1997, vol. 245, issue 2, s. 154-160. DOI: 10.1006/abio.1996.9916. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S0003269796999169>
- [109] WITWER, Carl T., et al. Continuous fluorescence monitoring of rapid cycle DNA amplification. *Biotechniques*, 1997, 22.1: 130-139.
- [110] TSENG, Susan Y., Dan MACOOL, Viki ELLIOTT, George TICE, Ray JACKSON, Mark BARBOUR a Douglas AMORESE. An Homogeneous Fluorescence Polymerase Chain Reaction Assay to Identify Salmonella. *Analytical Biochemistry*. 1997, vol. 245, issue 2, s. 207-212. DOI: 10.1006/abio.1996.9970. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S0003269796999704>
- [111] WILHELM, Jochen a Alfred PINGOUD. Real-Time Polymerase Chain Reaction. *ChemBioChem*. 2003-11-07, vol. 4, issue 11, s. 1120-1128. DOI: 10.1002/cbic.200300662. Dostupné z: <http://doi.wiley.com/10.1002/cbic.200300662>
- [112] JIN, X., et al. Sybr Green (Tm)-1-a new fluorescent dye optimized for detection of picogram amounts of DNA in gels. In: *Biophysical Journal*. 9650 ROCKVILLE PIKE, BETHESDA, MD 20814-3998: BIOPHYSICAL SOCIETY, 1994. p. A159-A159.
- [113] SINGER, V. L., et al. SYBR™ Green dyes: ultrasensitive strains for detection of DNA and RNA in electrophoretic gels. *Biomedical Products*, 1994, 19: 68-72.
- [114] SHEMESH, Moshe., Avshalom. TAM a Doron. STEINBERG. Differential gene expression profiling of Streptococcus mutans cultured under biofilm and planktonic conditions. *Microbiology* [online]. 2007, **153**(5), 1307-1317 [cit. 2020-05-03]. DOI: 10.1099/mic.0.2006/002030-0. ISSN 1350-0872. Dostupné z: <https://www.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.2006/002030-0>
- [115] SHEMESH, Moshe, Avshalom TAM a Doron STEINBERG. Expression of biofilm-associated genes of Streptococcus mutans in response to glucose and sucrose. *Journal of Medical Microbiology* [online]. 2007, **56**(11), 1528-1535 [cit. 2020-05-03]. DOI: 10.1099/jmm.0.47146-0. ISSN 0022-2615. Dostupné z: <https://www.microbiologyresearch.org/content/journal/jmm/10.1099/jmm.0.47146-0>
- [116] SMOOT, L. M., J. C. SMOOT, M. R. GRAHAM, G. A. SOMERVILLE, D. E. STURDEVANT, C. A. L. MIGLIACCIO, G. L. SYLVA a J. M. MUSSER. Global differential gene expression in response to growth temperature alteration in group A Streptococcus. *Proceedings of the National Academy of Sciences* [online]. 2001, **98**(18), 10416-10421 [cit. 2020-05-03]. DOI: 10.1073/pnas.191267598. ISSN 0027-8424. Dostupné z: <http://www.pnas.org/cgi/doi/10.1073/pnas.191267598>
- [117] DARGAHI, Narges, Joshua JOHNSON, Vasso APOSTOLOPOULOS a Hossam M ASHOUR. Streptococcus thermophilus alters the expression of genes associated with innate and adaptive immunity in human peripheral blood mononuclear cells. *PLOS ONE* [online]. 2020, **15**(2) [cit. 2020-05-03]. DOI: 10.1371/journal.pone.0228531. ISSN 1932-6203. Dostupné z: <https://dx.plos.org/10.1371/journal.pone.0228531>

- [118] TODD, E. W. a L. F. HEWITT. A new culture medium for the production of antigenic streptococcal hæmolysin. *The Journal of Pathology and Bacteriology*. 1932, **35**(6): 973-974. DOI: 10.1002/path.1700350614. ISSN 0368-3494. Dostupné také z: <http://doi.wiley.com/10.1002/path.1700350614>
- [119] SALIKHOV, Kamil, Gustavo SACOMOTO a Gregory KUCHEROV. Using cascading Bloom filters to improve the memory usage for de Bruijn graphs. *Algorithms for Molecular Biology* [online]. 2014, **9**(1) [cit. 2020-06-01]. DOI: 10.1186/1748-7188-9-2. ISSN 1748-7188. Dostupné z: <http://almob.biomedcentral.com/articles/10.1186/1748-7188-9-2>
- [120] *CLG Genome Workbench 20* [online]. QIAGEN Digital Insights [cit. 2020-02-18]. Dostupné z: <https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-clc-genomics-workbench/>
- [121] APTE, A. a S. DANIEL. PCR Primer Design. *Cold Spring Harbor Protocols* [online]. 2009, **2009**(3), pdb.ip65-pdb.ip65 [cit. 2020-05-25]. DOI: 10.1101/pdb.ip65. ISSN 1559-6095. Dostupné z: <http://www.cshprotocols.org/cgi/doi/10.1101/pdb.ip65>
- [122] SCHLOTTER, Yvette M., Eveline Z. VEENHOF, Bas BRINKHOF, Victor P.M.G. RUTTEN, Bart SPEE, Ton WILLEMSE a Louis C. PENNING. A GeNorm algorithm-based selection of reference genes for quantitative real-time PCR in skin biopsies of healthy dogs and dogs with atopic dermatitis. *Veterinary Immunology and Immunopathology*. 2009, **129**(1-2), 115-118. DOI: 10.1016/j.vetimm.2008.12.004. ISSN 01652427. Dostupné také z: <https://linkinghub.elsevier.com/retrieve/pii/S0165242708007769>
- [123] LIVAK, Kenneth J. a Thomas D. SCHMITTGEN. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* [online]. 2001, **25**(4), 402-408 [cit. 2020-05-26]. DOI: 10.1006/meth.2001.1262. ISSN 10462023. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1046202301912629>
- [124] GONI, Ramon, Patricia GARCÍA a Sylvain FOISSAC. The qPCR data statistical analysis. In: *Integromics White Paper* [online]. Spain, 2009, September 2009, s. 1-9 [cit. 2020-06-08]. Dostupné z: <https://www.gene-quantification.com/integromics-qpcr-statistics-white-paper.pdf>
- [125] *ApE - A plasmid Editor* [online]. University of Utah Biology Department [cit. 2019-03-31]. Dostupné z: <http://jorgensen.biology.utah.edu/wayned/apE/>
- [126] VAN DIJK, Erwin L., Yan JASZCZYŹYŹYN a Claude THERMES. Library preparation methods for next-generation sequencing: Tone down the bias. *Experimental Cell Research* [online]. 2014, **322**(1), 12-20 [cit. 2020-05-14]. DOI: 10.1016/j.yexcr.2014.01.008. ISSN 00144827. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S0014482714000160>
- [127] *Streptococcus equi subsp. zoepidemicus ATCC 35246, complete sequence: The annotation was added by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP)*. NCBI [online]. NCBI RefSeq. 2020, 2 February 2020 [cit. 2020-06-09]. Dostupné z: https://www.ncbi.nlm.nih.gov/nuccore/NC_017582.1?report=genbank&to=2167264
- [128] AMEND, ANTHONY S., KEITH A. SEIFERT a THOMAS D. BRUNS. Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Molecular Ecology* [online]. 2010, **19**(24), 5555-5565 [cit. 2020-05-29]. DOI: 10.1111/j.1365-294X.2010.04898.x. ISSN 09621083. Dostupné z: <http://doi.wiley.com/10.1111/j.1365-294X.2010.04898.x>
- [129] GARZETTI, Debora, Sandrine BRUGIROUX, Boyke BUNK, Rüdiger PUKALL, Kathy D. MCCOY, Andrew J. MACPHERSON a Bärbel STECHER. High-Quality Whole-Genome Sequences of the Oligo-Mouse-Microbiota Bacterial Community. *Genome Announcements* [online]. 2017, **5**(42), e00758-17 [cit. 2020-06-08]. DOI: 10.1128/genomeA.00758-17. ISSN 2169-8287. Dostupné z: <https://MRA.asm.org/content/5/42/e00758-17>
- [130] *Kyoto Encyclopedia of Genes and Genomes: KEGG* [online]. [cit. 2020-05-26]. Dostupné z: <https://www.kegg.jp/kegg/kegg2.html>

- [131] ANGOV, Evelina. Codon usage: Nature's roadmap to expression and folding of proteins. *Biotechnology Journal* [online]. 2011, **6**(6), 650-659 [cit. 2020-05-25]. DOI: 10.1002/biot.201000332. ISSN 18606768. Dostupné z: <http://doi.wiley.com/10.1002/biot.201000332>
- [132] TOMOYASU, Toshifumi, Atsushi TABATA, Hidenori IMAKI, Keigo TSURUNO, Aya MIYAZAKI, Kenji SONOMOTO, Robert Alan WHILEY a Hideaki NAGAMUNE. Role of Streptococcus intermedius DnaK chaperone system in stress tolerance and pathogenicity. *Cell Stress and Chaperones* [online]. 2012, **17**(1), 41-55 [cit. 2020-06-03]. DOI: 10.1007/s12192-011-0284-4. ISSN 1355-8145. Dostupné z: <http://link.springer.com/10.1007/s12192-011-0284-4>
- [133] MITCHELL, A.M. a T.J. MITCHELL. Streptococcus pneumoniae: virulence factors and variation. *Clinical Microbiology and Infection* [online]. 2010, **16**(5), 411-418 [cit. 2020-05-11]. DOI: 10.1111/j.1469-0691.2010.03183.x. ISSN 1198743X. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1198743X14633115>
- [134] ARORA, Gunjan, Andaleeb SAJID a Vipin Chandra KALIA, ed. *Drug Resistance in Bacteria, Fungi, Malaria, and Cancer* [online]. Cham: Springer International Publishing, 2017 [cit. 2020-06-08]. DOI: 10.1007/978-3-319-48683-3. ISBN 978-3-319-48682-6.
- [135] OUSKOVA, G. Hyaluronan release from Streptococcus pyogenes: export by an ABC transporter. *Glycobiology* [online]. 2004, **14**(10), 931-938 [cit. 2020-05-11]. DOI: 10.1093/glycob/cwh115. ISSN 1460-2423. Dostupné z: <https://academic.oup.com/glycob/article-lookup/doi/10.1093/glycob/cwh115>
- [136] THOMAS, Natalie K. a Tracey J. BROWN. ABC transporters do not contribute to extracellular translocation of hyaluronan in human breast cancer in vitro. *Experimental Cell Research* [online]. 2010, **316**(7), 1241-1253 [cit. 2020-05-11]. DOI: 10.1016/j.yexcr.2010.01.004. ISSN 00144827. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S0014482710000054>
- [137] HENDRIKSEN, Wouter T., Hester J. BOOTSMA, Silvia ESTEVÃO, Theo HOOGENBOEZEM, Anne DE JONG, Ronald DE GROOT, Oscar P. KUIPERS a Peter W. M. HERMANS. CodY of Streptococcus pneumoniae: Link between Nutritional Gene Regulation and Colonization. *Journal of Bacteriology* [online]. 2008, **190**(2), 590-601 [cit. 2020-06-03]. DOI: 10.1128/JB.00917-07. ISSN 0021-9193. Dostupné z: <https://JB.asm.org/content/190/2/590>
- [138] ABRANCHES, Jacqueline, Marcelle M. NASCIMENTO, Lin ZENG, Christopher M. BROWNGARDT, Zezhang T. WEN, Mercedes F. RIVERA a Robert A. BURNE. CcpA Regulates Central Metabolism and Virulence Gene Expression in Streptococcus mutans. *Journal of Bacteriology* [online]. 2008, **190**(7), 2340-2349 [cit. 2020-06-03]. DOI: 10.1128/JB.01237-07. ISSN 0021-9193. Dostupné z: <https://JB.asm.org/content/190/7/2340>
- [139] LAEHNEMANN, David, Arndt BORKHARDT a Alice Carolyn MCHARDY. Denoising DNA deep sequencing data—high-throughput sequencing errors and their correction. *Briefings in Bioinformatics* [online]. 2016, **17**(1), 154-179 [cit. 2020-06-06]. DOI: 10.1093/bib/bbv029. ISSN 1467-5463. Dostupné z: <https://academic.oup.com/bib/article-lookup/doi/10.1093/bib/bbv029>
- [140] VIMR, Eric R., Kathryn A. KALIVODA, Eric L. DESZO a Susan M. STEENBERGEN. Diversity of Microbial Sialic Acid Metabolism. *Microbiology and Molecular Biology Reviews* [online]. 2004, **68**(1), 132-153 [cit. 2020-06-08]. DOI: 10.1128/MMBR.68.1.132-153.2004. ISSN 1092-2172. Dostupné z: <https://MMBR.asm.org/content/68/1/132>
- [141] SEVERI, Emmanuele, Derek W. HOOD a Gavin H. THOMAS. Sialic acid utilization by bacterial pathogens. *Microbiology* [online]. 2007, **153**(9), 2817-2822 [cit. 2020-06-08]. DOI: 10.1099/mic.0.2007/009480-0. ISSN 1350-0872. Dostupné z: <https://www.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.2007/009480-0>
- [142] AFZAL, Muhammad, Sulman SHAFEEQ, Hifza AHMED, Oscar P. KUIPERS a M. KIVISAAR. Sialic Acid-Mediated Gene Expression in Streptococcus pneumoniae and Role of NanR as a Transcriptional Activator of the nan Gene

Cluster. *Applied and Environmental Microbiology* [online]. 2015, **81**(9), 3121-3131 [cit. 2020-06-08]. DOI: 10.1128/AEM.00499-15. ISSN 0099-2240. Dostupné z: <http://aem.asm.org/lookup/doi/10.1128/AEM.00499-15>

[143] BLATTNER, F. R. The Complete Genome Sequence of Escherichia coli K-12. *Science* [online]. **277**(5331), 1453-1462 [cit. 2020-06-08]. DOI: 10.1126/science.277.5331.1453. ISSN 00368075. Dostupné z: <https://www.sciencemag.org/lookup/doi/10.1126/science.277.5331.1453>

[144] JIANG, Yong-Liang, Wei-Li YU, Jun-Wei ZHANG, Cecile FROLET, Anne-Marie DI GUILMI, Cong-Zhao ZHOU, Thierry VERNET a Yuxing CHEN. Structural Basis for the Substrate Specificity of a Novel β -N-Acetylhexosaminidase StrH Protein from Streptococcus pneumoniae R6. *Journal of Biological Chemistry* [online]. 2011, **286**(50), 43004-43012 [cit. 2020-05-20]. DOI: 10.1074/jbc.M111.256578. ISSN 0021-9258. Dostupné z: <http://www.jbc.org/lookup/doi/10.1074/jbc.M111.256578>

[145] MENGIN-LECREULX, D a J VAN HEIJENOORT. Copurification of glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridyltransferase activities of Escherichia coli: characterization of the glmU gene product as a bifunctional enzyme catalyzing two subsequent steps in the pathway for UDP-N-acetylglucosamine synthesis. *Journal of Bacteriology* [online]. 1994, **176**(18), 5788-5795 [cit. 2020-05-08]. DOI: 10.1128/JB.176.18.5788-5795.1994. ISSN 0021-9193. Dostupné z: <https://JB.asm.org/content/176/18/5788>

[146] GINSBURG, Isaac. Role of lipoteichoic acid in infection and inflammation. *The Lancet Infectious Diseases* [online]. 2002, **2**(3), 171-179 [cit. 2020-05-25]. DOI: 10.1016/S1473-3099(02)00226-8. ISSN 14733099. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1473309902002268>

[147] SADO, J.F. a J. CHARVOLIN. Frustration in bilayers and topologies of liquid crystals of amphiphilic molecules. *Journal de Physique* [online]. 1986, **47**(4), 683-691 [cit. 2020-06-08]. DOI: 10.1051/jphys:01986004704068300. ISSN 0302-0738. Dostupné z: <http://www.edpsciences.org/10.1051/jphys:01986004704068300>

[148] NAUFER, M. Nabuan, Michael MORSE, Gudfridur B. MOLLER, James MCISAAC, Ioulia ROUZINA, Penny J. BEUNING a Mark C. WILLIAMS. E. coli Single Stranded Binding Protein (SSB) Self-Regulates Wrapping of ssDNA through Competitive Binding. *Biophysical Journal* [online]. 2020, **118**(3) [cit. 2020-05-25]. DOI: 10.1016/j.bpj.2019.11.559. ISSN 00063495. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S0006349519314882>

[149] CARPOUSIS, Agamemnon J. The RNA Degradosome of Escherichia coli: An mRNA-Degrading Machine Assembled on RNase E. *Annual Review of Microbiology* [online]. 2007, **61**(1), 71-87 [cit. 2020-05-09]. DOI: 10.1146/annurev.micro.61.080706.093440. ISSN 0066-4227. Dostupné z: <http://www.annualreviews.org/doi/10.1146/annurev.micro.61.080706.093440>

[150] SZE, Jun Hui, Jeremy C. BROWNLIE a Christopher A. LOVE. Biotechnological production of hyaluronic acid: a mini review. *3 Biotech* [online]. 2016, **6**(1) [cit. 2020-05-07]. DOI: 10.1007/s13205-016-0379-9. ISSN 2190-572X. Dostupné z: <http://link.springer.com/10.1007/s13205-016-0379-9>

[151] SRIVASTAVA, Satish K. a Donald A. BARNUM. The role of lipoteichoic acids on the adherence of Streptococcus equi to epithelial cells. *Veterinary Microbiology* [online]. 1983, **8**(5), 485-492 [cit. 2020-05-09]. DOI: 10.1016/0378-1135(83)90042-1. ISSN 03781135. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/0378113583900421>

[152] YASIR, Muhammad, Muhammad FARMAN, Muhammad W. SHAH, et al. Genomic and antimicrobial resistance genes diversity in multidrug-resistant CTX-M-positive isolates of Escherichia coli at a health care facility in Jeddah. *Journal of Infection and Public Health* [online]. 2020, **13**(1), 94-100 [cit. 2020-05-25]. DOI: 10.1016/j.jiph.2019.06.011. ISSN 18760341. Dostupné z: <https://linkinghub.elsevier.com/retrieve/pii/S1876034119301996>

[153] DIÉGUEZ, Ana L., Sabela BALBOA a Jesús L. ROMALDE. Halomonas borealis sp. nov. and Halomonas niordiana sp. nov., two new species isolated from seawater. *Systematic and Applied Microbiology* [online]. 2020, **43**(1) [cit. 2020-05-25].

DOI: 10.1016/j.syapm.2019.126040. ISSN 07232020. Dostupné z:
<https://linkinghub.elsevier.com/retrieve/pii/S0723202019303352>

[154] WANG, Yu, Xiawei JIANG, Zemin XU, Chaoqun YING, Wei YU a Yonghong XIAO. Identification of Raoultella terrigena as a Rare Causative Agent of Subungual Abscess Based on 16S rRNA and Housekeeping Gene Sequencing. *Canadian Journal of Infectious Diseases and Medical Microbiology* [online]. 2016, **2016**, 1-4 [cit. 2020-05-25]. DOI: 10.1155/2016/3879635. ISSN 1712-9532. Dostupné z: <http://www.hindawi.com/journals/cjidmm/2016/3879635/>

8 SEZNAM POUŽITÝCH ZKRATEK

- HA – hyaluronová kyselina / Hyaluronic Acid
- GlcUA – D-glukuronová kyselina
- GlcNAc – N-acetylglukosamin
- UTP – uridin trifosfát
- UDP – uridin difosfát
- ATP – adenosin trifosfát
- HAS – hyaluronan syntáza
- mRNA – mediátorová ribonukleová kyselina
- RNA – ribonukleová kyselina
- DNA – deoxyribonukleová kyselina
- PCR – polymerázová řetězová reakce / Polymerase Chain Reaction
- qPCR – polymerázová řetězová reakce v reálném čase / real-time PCR
- bp – páry bází / base pair
- BLAST – Basic Local Alignment Search Tool
- cDNA – komplementární DNA / complementary DNA
- UV – ultrafialové záření
- dsDNA – dvouřetězová deoxyribonukleová kyselina
- THY – Todd Hewitt Broth médium
- OD – optická denzita
- NCBI – National Center for Biotechnology Information
- WGS – celogenomová sekvenace / Whole Genome Sequencing
- NGS – sekvenace nové generace / New Generation Sequencing
- dNTP – deoxyribonukleotid trifosfát
- IPA – izopropanol
- gDNA – genomová DNA
- SNV – bodová mutace / Single Nucleotide Variant
- CDS – kódující sekvence / CoDing Sequence
- NCS – nekódující sekvence / Non CoDing sequence
- KEGG – Kyoto Encyclopedia of Genes and Genomes
- LTA – kyselina lipoteichoová
- AMK – aminokyselina / aminokyselinový

9 STRUČNÝ ŽIVOTOPIS AUTORA

Osobní informace

Jméno: Radka Hrudíková
Trvalé bydliště: Sokolská 439, Holešov 769 01
Datum narození: 30.01.1988

Vzdělání

2012 – současnost Vysoké učení technické v Brně, Fakulta chemická, program Fyzikální chemie, obor Fyzikální chemie
2010 – 2012 Vysoké učení technické v Brně, Fakulta chemická, program Chemie a technologie potravin, obor Potravinářská chemie a biotechnologie, dosažený titul Ing., téma diplomové práce: Využití magnetických mikročástic pro izolaci bakteriální DNA
2007 – 2010 Vysoké učení technické v Brně, Fakulta chemická, program Chemie a technologie potravin, obor Biotechnologie, dosažený titul Bc., téma bakalářské práce: Mikrofauna aktivovaného kalu biologických čistíren odpadních vod na Kroměřížsku
2003 – 2007 Gymnázium L. Jaroše v Holešově, maturitní zkouška: Český jazyk, Anglický jazyk, Chemie, Biologie

Pracovní zkušenosti

8/2012 – 7/2013 Biotechnologická a farmaceutická společnost Contipro a.s., odborný pracovník v laboratoři
8/2014 – 2/2016 Biotechnologická a farmaceutická společnost Contipro a.s., odborný pracovník v laboratoři
1/2013 – 6/2013 Mendelova univerzita v Brně, Ústav chemie a biochemie, externí lektor
6/2016 – souč. AVČR, Ústav molekulární genetiky, výzkumný pracovník ve skupině buněčné motility

Zahraniční stáž

8/2013 – 6/2014 Norwegian University of Science and Technology (NTNU), Department of Biotechnology, Trondheim, Norway, Aim of work: New insights into Pseudomonas fluorescens alginate biosynthesis relevant for the establishment of an efficient production process for microbial alginates

Pedagogická činnost

1/2013 – 6/2013 Mendelova univerzita v Brně, Ústav chemie a biochemie, externí lektor pro praktikum z biochemie

10 SEZNAM PUBLIKACÍ AUTORA

- KONFERENCE

HRUDÍKOVÁ, R. PEPELIAEV, S. *Determination of the reference genes in Streptococcus equi subsp. zooepidemicus for real-time PCR.* CEITEC PhD Retreat, Book of Abstracts. Brno: Masaryk University, 2015. s. 49-49. ISBN: 978-80-210-7825-3.

Radka Hrudíková, Stanislav Pepeliaev, Vladimír Velebný. *Determination of the Reference Genes in Streptococcus equi subsp. zooepidemicus for Real-Time PCR.* Sborník abstraktů. 1. Brno: Vysoké Učení Technické v Brně, Fakulta chemická, Purkyňova 464/118, Brno 61200, 2015. s. 65-65. ISBN: 978-80-214-5290-9.

Paprskářová A.; **Hrudíková R.**; Švadlák D.; Zápotocký V.; Velebný V.; Suchý. *Susceptibility of selected antibiotics, used individually and in their combinations, to the most common causative agents of nosocomial infections, in cardiovascular surgery.* Sborník abstraktů, 65. česko-slovenské farmakologické dny. Praha: Farmakologický ústav 1. LF UK a VFN v Praze, 2015. s. 35-35. ISBN: 978-80-260-8027-5. <http://www.quickcomp.cz/Conference/files/Sbornik.pdf>

Stanislav Pepeliaev, Jana Jilkova, **Radka Hrudikova**, Dzianis Smirnou, Lukas Franke. *Mutant hyaluronan lyase sphyl lacking a-domain possesses catalytic activity.* Proceedings of the 11th International Conference on Polysaccharides- Glycoscience. 1. Prague, Czech Republic: Czech Chemical Society, Prague, 2015. s. 147-150. ISBN: 978-80-86238-85-2. ISSN: 2336-6796.

Susan Maleki, Mali Mærk, **Radka Hrudikova**, Svein Valla, Helga Ertesvåg. *IN PSEUDOMONAS FLUORESCENS ALGINATE PRODUCTION LEVELS ARE DETERMINED BY THE AVAILABLE FRUCTOSE 6-PHOSPHATE.* Book of Abstracts. 1. Řím: SAPIENZA universita di Roma, 2015.

- IMPAKTOVANÉ ČLÁNKY

PEPELIAEV, Stanislav, **Radka HRUDÍKOVÁ**, Jana JÍLKOVÁ, Jaroslav PAVLÍK, Dzianis SMIRNOU, Zbyněk ČERNÝ a Lukáš FRANKE. *Colorimetric enzyme-coupled assay for hyaluronic acid determination in complex samples.* *European Polymer Journal* [online]. 2017, **94**, 460-470 [cit. 2018-07-02]. DOI: 10.1016/j.eurpolymj.2017.07.036. ISSN 00143057. Dostupné z: <http://linkinghub.elsevier.com/retrieve/pii/S0014305717300551>

MALEKI, Susan, Mali MÆRK, **Radka HRUDIKOVA**, Svein VALLA a Helga ERTESVÅG. *New insights into Pseudomonas fluorescens alginate biosynthesis relevant for the establishment of an efficient production process for microbial alginates.* *New Biotechnology*. 2016, DOI: 10.1016/j.nbt.2016.08.005. ISSN 18716784. Dostupné také z: <http://linkinghub.elsevier.com/retrieve/pii/S1871678416323925>

ŠTĚPÁNKOVÁ, Veronika, Daniel ŠVADLÁK, **Radka HRUDÍKOVÁ**, Alice PAPRSKÁŘOVÁ, Vojtěch ZÁPOTOCKÝ a Vladimír VELEBNÝ. *The minimal inhibition concentration obtained by E-test as a criterion for selection of mutual combination of antimicrobial agents.* *New Biotechnology*. 2016, **33**, S80-. DOI: 10.1016/j.nbt.2016.06.994. ISSN 18716784. Dostupné také z: <http://linkinghub.elsevier.com/retrieve/pii/S1871678416317770>

FRANKE, Lukáš, Dzianis SMIRNOU, Stanislav PEPELIAEV, Zbyněk ČERNÝ, **Radka HRUDÍKOVÁ** a Vladimír VELEBNÝ. *N-Acetylglucosamine as a signalling molecule triggering hyaluronic acid increase in Streptococcus zooepidemicus.* *New Biotechnology*. 2016, **33**, S152-. DOI: 10.1016/j.nbt.2016.06.1248. ISSN 18716784. Dostupné také z: <http://linkinghub.elsevier.com/retrieve/pii/S1871678416320313>

- PATENTY

CONTIPRO A.S. *Method of determination of hyaluronic acid*. Inventors: PEPELIAEV Stanislav, **HRUDIKOVA Radka**, JILKOVA Jana, PAVLIK Jaroslav, VELEBNY Vladimir. Czech Republic. Application number: WO2018166545A1. Date of application: 2018-09-20. online: https://worldwide.espacenet.com/publicationDetails/biblio?FT=D&date=20180920&DB=EPODOC&locale=en_EP&CC=WO&NR=2018166545A1&KC=A1&ND=5#