

UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

ANALÝZA VÍNNEHO MIKROBIÓMU  
BAKALÁRSKA PRÁCA

2023  
MAREK KAJAN



UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

ANALÝZA VÍNNEHO MIKROBIÓMU  
BAKALÁRSKA PRÁCA

Študijný program: Bioinformatika  
Študijný odbor: Informatika a Biológia  
Školiace pracovisko: Katedra informatiky  
Školiteľ: Mgr. Jaroslav Budiš, PhD.

Bratislava, 2023  
Marek Kajan





Univerzita Komenského v Bratislave  
Fakulta matematiky, fyziky a informatiky

---

## ZADANIE ZÁVEREČNEJ PRÁCE

**Meno a priezvisko študenta:** Marek Kajan  
**Študijný program:** bioinformatika (Medziodborové štúdium, bakalársky I. st., denná forma)  
**Študijné odbory:** informatika  
biológia  
**Typ záverečnej práce:** bakalárska  
**Jazyk záverečnej práce:** slovenský  
**Sekundárny jazyk:** anglický

**Názov:** Analýza vínneho mikrobiómu  
*Wine Microbiome Analysis*

**Anotácia:** Mikroorganizmy zohrávajú pri výrobe vína kľúčovú úlohu. Množstvo mikroorganizmov, ich taxonomické zloženie a s tým súvisiaca aktivita sa líšia v rôznych fázach výroby vína. Analýza vplyvu pridaných štartovacích kultúr na vývoj mikrobiálnych komunit nám môže pomôcť pochopiť vnútorné procesy pri kvasení, a tým skvalitniť kontrolovanú výrobu vína.

Vstupnými údajmi pre bioinformatické analýzy sú zloženia mikrobiálnych komunit z rôznych fáz výroby vína. Študent zanalyzuje rozdiely v mikrobiálnych komunitách kvasiniek a baktérií pri rôznych podmienkach, ako zmena štartovacích kultúr kvasiniek, iné ročníky a rozdiely medzi meraním molekúl typu DNA a RNA. Opíše systematické vzory, ako aj špecifiká pre jednotlivé skupiny.

**Vedúci:** Mgr. Jaroslav Budiš, PhD.  
**Katedra:** FMFI.KI - Katedra informatiky  
**Vedúci katedry:** prof. RNDr. Martin Škoviera, PhD.  
**Dátum zadania:** 22.10.2022

**Dátum schválenia:** 22.10.2022

doc. Mgr. Bronislava Brejová, PhD.  
garant študijného programu

.....  
študent

.....  
vedúci práce

**Pod'akovanie:** Chcem sa poďakovať môjmu školiteľovi Jaroslavovi Budišovi, jeho kolegom Michalovi Zemanovi a Miroslavovi Böhmerovi za skvelú pomoc v každom štádiu tejto bakalárskej práce.

## Abstrakt

V tejto práci sme sa pozreli na vývoj vínneho mikrobiómu počas fáz kvasenia vina. Pozerali sme sa na zloženie fungálnych a aj bakteriálnych komunít. Taktiež sme porovnávali vplyv rôznych štartérov na vínny mikrobióm. Dáta boli zberané v rokoch 2018 a 2019, medzi ktorými bol aj veľký rozdiel v počasí počas rastu hrozna. Fungálne komunity boli dominované kvasinkami z rodu *Saccharomyces* vo všetkých fázach kvasenie, v oboch rokoch a v každom štartéry. Vo vzorkách, kde bol použitý štartér obsahujúci kvasinky *Lachancea thermotolerans* sme zaznamenali mierne zvýšenie metrík alfa diverzity, oproti iným štartérom. Fungálne populácie boli bohatšie a vyrovnanejšie v týchto vzorkách. Stále však dominovali kvasinky rodu *Saccharomyces*. V bakteriálnych populáciách nebolo vidno žiaden vplyv rôznych štartérov na zloženie populácie baktérií. Rozdiel však bolo vidno na základe rôzneho počasia v rokoch odberu vzoriek. V oboch rokoch sa vo vzorkách nachádzalo veľa baktérií, produkujúcich kyselinu octovú ako napríklad *Gluconobacter* a *Komagataeibacter*. Taktiež sme našli baktérie zodpovedné za mliečne kvasenie ako napríklad *Lactobacillus*. Veľký rozdiel medzi rokmi bol hlavne pri RNA aktivite v roku 2018. Túto aktivitu dominovali baktérie rodu *Escherichia-Shigella*. V roku 2019 takáto dominancia nebola.

**Kľúčové slová:** víno, mikrobióm, amplicónové sekvenovanie, kvasenie

## Abstract

In this paper we looked at the evolution of wine microbiome during different stages of the fermentation process. We looked at the composition of fungal and bacterial communities. We also looked at how different starter cultures would impact the wine microbiome. Data was gathered in the years 2018 and 2019. There was a big difference in weather conditions between these 2 years during the growth of grapes. Fungal communities were dominated by *Saccharomyces* yeast in all stages of fermentation in both years, regardless of starter culture. In samples where yeast *Lachancea thermotolerans* were used in the starter culture, we noticed a slight increase in alpha diversity metrics in comparison to other starter cultures. The fungal population in these samples was a little richer and a little more even. But there was still a clear domination of *Saccharomyces* yeast. In bacterial communities, we did not notice any differences based on different starter cultures. We observed some differences based on the year the samples were gathered. In both years, we observed a lot of acetic acid producing bacteria such as *Gluconobacter* and *Komagataeibacter*. We also observed bacteria responsible for malolactic fermentation such as *Lactobacillus*. There was a big difference in RNA activity between the years. In 2018, the RNA activity was dominated by *Escherichia-Shigella* bacteria. In 2019, there was no such dominance.

**Keywords:** wine, microbiome, amplicon sequencing, fermentation





# Obsah

<b>Úvod</b>	<b>1</b>
<b>1 Teoretický základ</b>	<b>3</b>
1.1 Mikrobióm . . . . .	3
1.2 Výroba vína . . . . .	6
1.2.1 Štartér . . . . .	6
1.2.2 Fázy kvasenia . . . . .	6
1.3 Next Generation Sequencing . . . . .	7
1.3.1 Amplicon Sequence Variant . . . . .	7
<b>2 Metodika práce</b>	<b>9</b>
2.1 Použité dátové zdroje . . . . .	9
2.1.1 Laboratórne spracovanie . . . . .	10
2.1.2 Určenie zastúpenia taxonomických druhov . . . . .	10
2.2 Spracovanie taxonomických počtov . . . . .	12
2.2.1 Vývoj mikrobiómu pri kvasení vína . . . . .	12
2.3 Diferenčná analýza . . . . .	14
2.3.1 DESeq2 . . . . .	15
2.4 Alfa diverzita . . . . .	15
2.4.1 Shannonov index . . . . .	15
2.4.2 Rarefaction . . . . .	16
2.4.3 Porovnanie alfa diverzít . . . . .	17
2.5 Beta diverzita . . . . .	17
2.5.1 Bray-Curtis metóda . . . . .	18
2.5.2 Analýza hlavných súradníc . . . . .	18
<b>3 Výsledky</b>	<b>21</b>
3.1 Určenie zastúpenia taxonomických druhov . . . . .	21
3.1.1 Fungálne 28S dáta . . . . .	21
3.1.2 Bakteriálne 16S dáta . . . . .	25
3.1.3 Prežívanie experimentálnych štartérov . . . . .	29

3.2	Diferenčná analýza . . . . .	31
3.2.1	Fungálne 28S dáta . . . . .	32
3.2.2	Bakteriálne 16S dáta . . . . .	33
3.3	Alfa diverzita . . . . .	34
3.3.1	Rarefaction . . . . .	34
3.3.2	Porovnanie alfa diverzít . . . . .	35
3.4	Beta diverzita . . . . .	39
3.4.1	Analýza hlavných súradníc pre fungálne 28S dáta . . . . .	39
3.4.2	Analýza hlavných súradníc pre bakteriálne 16S dáta . . . . .	44
<b>4</b>	<b>Diskusia</b>	<b>49</b>
	<b>Záver</b>	<b>51</b>
	<b>Príloha A</b>	<b>59</b>

# Zoznam obrázkov

1.1	Ilustračný fylogenetický strom kvasiniek . . . . .	5
2.1	Diagram získavanie vzoriek . . . . .	10
2.2	Diagram laboratórneho spracovania vzoriek . . . . .	11
2.3	Ukážka stacked bar chart grafu . . . . .	13
2.4	Ukážka čiarového grafu . . . . .	14
2.5	Ukážka rarefaciton grafu . . . . .	17
2.6	Ukážka 3D projekcie PCoA . . . . .	19
2.7	Ukážka porovnania komponentov PCoA navzájom . . . . .	20
3.1	Výsledok taxonomickej analýzy pre fungálne vzorky DNA roku 2018 . .	22
3.2	Výsledok taxonomickej analýzy pre fungálne vzorky RNA roku 2018 . .	23
3.3	Výsledok taxonomickej analýzy pre fungálne vzorky DNA roku 2019 . .	24
3.4	Výsledok taxonomickej analýzy pre fungálne vzorky RNA roku 2019 . .	25
3.5	Výsledok taxonomickej analýzy pre bakteriálne vzorky DNA roku 2018	26
3.6	Výsledok taxonomickej analýzy pre bakteriálne vzorky RNA roku 2018	27
3.7	Výsledok taxonomickej analýzy pre bakteriálne vzorky DNA roku 2019	28
3.8	Výsledok taxonomickej analýzy pre bakteriálne vzorky RNA roku 2019	29
3.9	Prežívanie kvasiniek štartéru lasa . . . . .	30
3.10	Prežívanie kvasiniek štartéru mesa . . . . .	31
3.11	Výsledok rarefaction grafov . . . . .	35
3.12	Výsledok porovnania bohatostí pre fungálne dáta . . . . .	36
3.13	Výsledok porovnania rovnováhy pre fungálne dáta . . . . .	37
3.14	Výsledok porovnania bohatostí pre bakteriálne dáta . . . . .	38
3.15	Výsledok porovnania rovnováhy pre bakteriálne dáta . . . . .	39
3.16	3D projekcia PCoA pre fungálne DNA dáta . . . . .	40
3.17	PCoA pre fungálne DNA dáta po jednotlivých komponentoch . . . . .	41
3.18	3D projekcia PCoA pre fungálne RNA dáta . . . . .	42
3.19	PCoA pre fungálne RNA dáta po jednotlivých komponentoch . . . . .	43
3.20	3D projekcia PCoA pre bakteriálne DNA dáta . . . . .	44
3.21	PCoA pre bakteriálne DNA dáta po jednotlivých komponentoch . . . . .	45

3.22 3D projekcia PCoA pre bakteriálne RNA dáta . . . . .	46
3.23 PCoA pre bakteriálne RNA dáta po jednotlivých komponentoch . . . .	47

# Úvod

Vínny mikrobióm je zaujímavý koncept z pohľadu potravinárskeho priemyslu. Mikroorganizmy, ktoré sú prítomné na povrchu hrozna, v pôde a v prostredí viniča ovplyvňujú procesy fermentácie a zrenia vína. Tieto mikroorganizmy zahŕňajú baktérie, kvasinky a huby.

Výskum vínneho mikrobiómu nám pomáha lepšie porozumieť interakciám medzi týmito mikroorganizmami a ich vplyvu na kvalitu a charakter vína. Rôzne druhy mikroorganizmov prispievajú k rôznym aspektom fermentačného procesu. Kvasinky napríklad premenia cukry na alkohol a oxid uhličitý, čím dochádza k alkoholickému kvaseniu. Baktérie môžu ovplyvniť chuť a aroma vína, a to buď priamo prostredníctvom svojich metabolických aktivít alebo nepriamo prostredníctvom interakcií s kvasinkami.

Pri procese výroby vína sa kvasenie vína môže uskutočniť spontánne, teda len za prítomnosti kvasiniek a baktérií nachádzajúcich sa na hrozne alebo kontrolovane za pomoci rôznych štartérov. Jeden z bežne používaných je štartér obsahujúci kvasinky *Saccharomyces cerevisiae*. Štartér sú mikroorganizmy, ktoré sa vložia do vínneho muštu pred začatím procesu kvasenia, čím menia kvalitu, arómu a charakter výsledného vína. V tejto práci sa pozrieme aj na vplyv rôznych štartovacích kultúr na celkový mikrobióm vína. Získané poznatky môžu prispieť k vylepšeniu procesu kontrolovaného kvasenia vína, aby sa docielil požadovaný charakter vína.

V prvej kapitole práce opíšeme proces výroby vína a zavedieme niektoré odborné pojmy používané v tejto práci. V druhej kapitole sa zameriame na praktickú časť práce, opíšeme dátové zdroje a použité výpočtové metodiky na ich analýzu. V tretej kapitole sa zoberieme výsledky dátových analýz.



# Kapitola 1

## Teoretický základ

### 1.1 Mikrobióm

Pojem mikrobióm opisuje všetky mikroorganizmy, ich metagenóm a podmienky okoli-  
tého prostredia v ktorom žijú [25]. Mikrobióm sa dá vzhľadom na typ biochemického  
zdroja charakterizovať pomocou rôznych techník, ako je napríklad metagenomika, me-  
tatranskriptomika, metaproteomika v kombinácii s metadátami o prostredí [25].

Mikroorganizmy, ktoré žijú v jednom prostredí spolu interagujú a tým dynamicky  
menia štruktúru populácie mikrobiómu ako aj prostredia, v ktorom žijú. Interakcie  
medzi mikroorganizmami sa dajú rozdeliť na priame a nepriame. Priame interakcie  
nastávajú pri fyzickom kontakte. Patrí sem napríklad parazitizmus, symbióza alebo  
predácia. Pri parazitizme jeden organizmus získava výhodu na úkor iného organizmu.  
Pri symbióze oba organizmy benefitujú zo vzájomnej interakcie. Pri predácii jeden z  
organizmov je potrava pre druhý organizmus.

Nepriame interakcie sú spôsobené prítomnosťou mimobunkových metabolitov. Me-  
dzi tento druh interakcie patrí neutralizmus, mutualizmus, komenzalizmus, alelopátia  
a kompetícia [21]. Pri neutralizme interakcia medzi organizmami neprináša výhodu ani  
újmu ani jednému z organizmov. Pri mutualizme oba organizmy získavajú výhodu z  
danej interakcie. Pri komenzalizme jeden z organizmov získa výhodu pričom to nemá  
žiadny efekt na druhý organizmus. Pri alelopácii mimobunkové metabolity jedného or-  
ganizmu prinášajú újmu druhému organizmu. A pri kompetícii mikroorganizmy bojujú  
o limitovaný zdroj potravy.

Štúdium týchto interakcií je jednoduchšie, ak ide o 2 organizmy. Avšak keď sa na ne  
pozeráme z pohľadu celého mikrobiómu, tak interakcie medzi rôznorodými populáciami  
mikroorganizmov sa stávajú zložité.

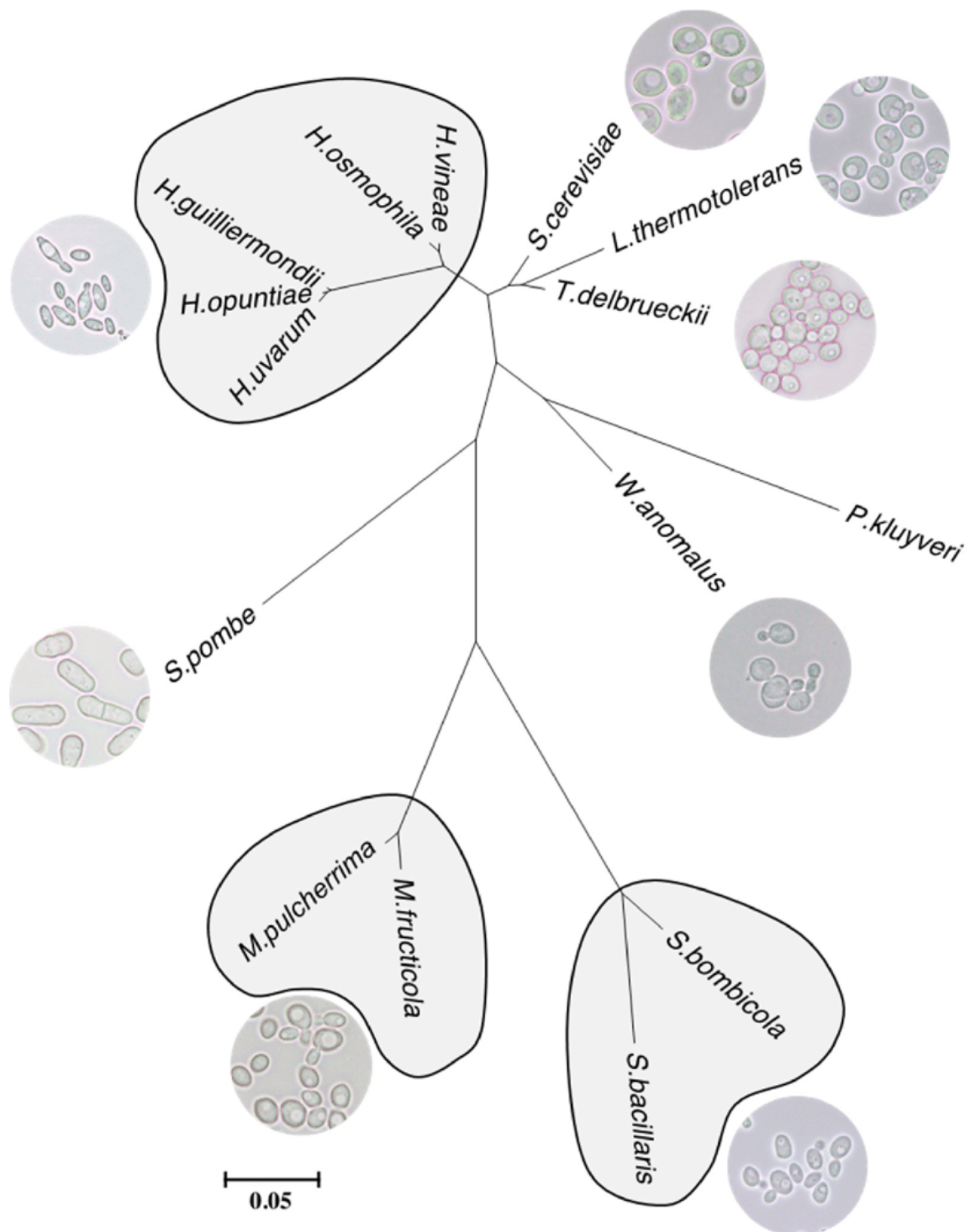
Víno je zaujímavé prostredie pre štúdiu týchto interakcií a vývoja populácií mikro-  
organizmov. Mikrobiálne populácie sa dajú študovať priebežne počas procesu výroby  
vína. Veľká biodiverzita kvasiniek a baktérií bola pozorovaná na povrchu hrozna, v



mušte a vo víne. Jej aktivita mikrobiómu vo víne je ovplyvňuje nutričné a aromatické kvality vína [22].

Vôňa a chuť vína je určená zložením prchavých a neprchavých aromatických látok. Tieto látky sa nachádzajú v mušte vína ešte predtým ako začne fermentácia a ďalšie sú metabolizované počas procesu kvasenia a dozrievania vína [43]. Tieto aromatické látky sa dajú rozdeliť na primárne aromatické látky, sekundárne (fermentačné) látky a terciárne (postfermentačné) látky [2]. Primárne aromatické látky majú pôvod v hrozne, tvoria odrodový charakter vína [2]. Sekundárne (fermentačné) látky vytvárajú jednak kvasinky počas alkoholového kvasenia, ako aj mliečne baktérie počas mliečneho kvasenia [2]. Terciárne (postfermentačné) látky zahŕňajú zlúčeniny vytvárané počas vyzrievania vína [2]. Medzi hlavné aromatické látky ovplyvňujúce charakter vína patria ethylfenoly, vinylfenoly, mastné kyseliny, alkoholy, estery, terpenoly, pyrazíny a mnohé ďalšie [43][2]. Štúdiá mikrobiómu vína nám dáva lepšie porozumenie tomu, aké mikroorganizmy vytvárajú ktoré aromatické látky. Prípadne vznik niektorej z aromatických látok môže byť dôsledok interakcie medzi špecifickými mikroorganizmami v populácii. Takéto poznanie môže vinárom pomôcť pri výrobe kvalitnejších vín alebo vín so špecifickým charakterom.

Zloženie mikrobiómu na povrchu hrozna je veľmi variabilné. Hlavným dôvodom tejto variability sú vonkajšie faktory prostredia ako sú geografická lokalita a spôsoby, akými sa vinári starajú o samotné rastliny pri raste hrozna. Vďaka tomuto mikrobiómu hrozno dokáže aj spontánne kvasiť (nie je potrebné pridať kvasinky do vínneho muštu). Štúdiá ukázala, že hlavné druhy nachádzajúce sa v prvotnej fáze kvasenia hrozna sú z rodov *Saccharomyces*, *Metschnikowia*, *Lachancea*, *Pichia* a *Hanseniaspora* [30]. Na obrázku 1.1 môžeme vidieť ich fylogenetické zaradenie. Avšak v neskorších fázach kvasenia sa ukázalo, že vďaka prítomnosti etanolu dominujú kvasinky z rodu *Saccharomyces* [30][26].



Obr. 1.1: Fylogenetický strom príbuznosti kvasiniek a hub, ktoré sa vyskytujú na povrchu hrozna. Ohraničené oblasti označujú rovnaké rodové zaradenie. Vzdialenosti boli určené metódou maximálnej vierohodnosti. Mierka zobrazuje pravdepodobnosť substitúcie bázy na pozícii v zarovnaných génoch. Získané zo zdroja [26]

## 1.2 Výroba vína

Pri výrobe vína hrozno prechádza tzv. vinifikačným procesom. Tento proces začína zberom hrozna, ktoré sa následne lisuje aby sa získal mušt. Tento mušt sa môže nechať spontánne kvasiť alebo prípadne môžeme do muštu pridať štartér pre kontrolované kvasenie.

### 1.2.1 Štartér

Pod pojmom štartér sa vo vinárstve označuje zmes mikrobiálnych kultúr, ktoré sa inokulujú do muštu vína. Inokulácia v mikrobiológii znamená, pridanie mikroorganizmov do kultúrneho média. V našom prípade je teda kultúrne médium mušt vína [1]. Pri úspešnej inokulácii môže štartér ovplyvniť mikrobióm vína počas všetkých fáz kvasenia a teda aj výsledný produkt. V dnešnej dobe sa pri výrobe vína typicky používajú komerčné štartéry. Niektoré z týchto štartérov obsahujú ako hlavnú zložku kvasinky *Saccharomyces cerevisiae*. Takýto štartér sa v našej analýze využíva ako kontrolná skupina (v texte ho označujem ako kontrolný štartér, v grafoch a tabuľkách ako ctrl).

V tejto práci tento kontrolný štartér budeme porovnávať s experimentálnymi štartérmi. Máme 2 takéto štartéry. Prvý z týchto štartérov pozostáva z kvasiniek *Metschnikowia pulcherrima* a *Saccharomyces cerevisiae*. Pre zvýšenie čitateľnosti bude v dátových analýzach a v grafoch tento štartér označený pod skratkou mesa. Druhý štartér, ktorý bol použitý pozostáva z kvasiniek *Lachancea thermotolerans* a *Saccharomyces cerevisiae*. Tento je označený skratkou lasa.

### 1.2.2 Fázy kvasenia

Hrozno prechádza pri výrobe vína rôznymi fázami kvasenia. Pri zbere dát, použitých v tejto práci vinári definovali tri hlavné fázy kvasenia.

Prvou fázou kvasenia je iniciačný mušt. Táto fáza označuje vínny mušt, ktorý bol inokulovaný štartérom. V dátach a grafoch bude táto fáza označená ako m1.

Druhou fázou je fermentovaný mušt. V tejto fáze v mušte vína aktívne prebieha fermentácia. Najmä sa jedná o alkoholovú fermentáciu. Pri tejto je cukor z hrozna metabolizovaný na etanol. Pri tomto procese vznikajú aj niektoré aromatické látky [12]. V dáta a grafoch je označená ako m2.

Tretia fáza je označená ako mladé víno pre filtráciou. V tejto fáze by víno malo byť už na konci fermentačného procesu. V tejto fáze je už väčšina cukru spotrebovaná pri alkoholovej fermentácii. Môže tu ešte prebiehať mliečne kvasenie. Pri mliečnom kvasení dochádza ku konverzii kyseliny jablčnej na kyseliny mliečnu. Následne sa z kyseliny mliečnej produkujú aromatické látky a dochádza k poklesu kyslosti [35]. V dátach a v grafoch je označená ako m3.

## 1.3 Next Generation Sequencing

Zloženie mikrobiómu sa dá študovať mnohými metódami. V súčasnosti jedna z obľúbených metód je Next Generation Sequencing. Next Generation Sequencing (ďalej len NGS) je charakteristické tým, že sa paralelne sekvenuje veľké množstvo (môže byť aj milióny) malých fragmentov polynukleotidov [6]. Toto sa nazýva aj masívne paralelne sekvenovanie. Výhodou tejto metódy sekvenovania je, že sa môže osekvenovať veľké množstvo dát za krátky čas a pomerne nízku cenu.

Táto vlastnosť NGS sa využíva pri amplikónovom sekvenovaní [42]. Pri amplikónovom sekvenovaní sa digitalizuje iba časť markerového génu. Markerový gén môže byť hociktorý gén, ktorý je v genóme každého skúmaného organizmu. Napríklad v baktériách sa ako markerový gén využíva gén pre 16S ribozomálnu RNA. Tento gén sa používa, pretože je vysoko konzervovaný a je potrebný pre každú baktériu. Navyše tento gén má 9 hypervariabilných oblastí, ktoré sa líšia medzi druhmi a rodmi baktérií [42]. Zacielenie sekvenovania na tento gén je docielené pomocou PCR. Pred sekvenovaním sa DNA a RNA vo vzorkách amplifikuje pomocou PCR. PCR primery sú navrhnuté tak, aby sa viazali na konzervovanú oblasť génu a amplifikovali hypervariabilnú oblasť. Takéto produkty PCR sa upravujú vo viacerých krokoch. Napríklad jedným z týchto krokov je odstránenie sekvencií primerov. Následne sa tieto upravené produkty PCR digitalizujú pomocou NGS [42]. Digitalizované sekvencie získané z PCR sa ďalej analyzujú. Najčastejšie sa jedná o organizáciu čítaní do skupín taxónov. Jednou z týchto analýz je Amplicon Sequence Variant [42].

### 1.3.1 Amplicon Sequence Variant

Amplicon Sequence Variant (ďalej len ASV) je metóda klastrovania čítaní. Táto metóda najskôr odstráni čítania, pri ktorých nie je štatisticky isté, že nedošlo ku chybe pri amplifikácii alebo sekvenovaní. Toto sa deje za predpokladu, že biologicky správne sekvencie sa budú vo vzorke nachádzať častejšie ako tie s chybami [11]. Rovnaké sekvencie sa klastrujú dokopy. Klastre sa navzájom môžu líšiť aj na základe jedného nukleotidu [11]. Každému klastru je priradená veľkosť na základe počtu čítaní v ňom. Tieto počty odrážajú zastúpenie jednotlivých taxónov vo vzorke. Táto informácia o počtoch je hlavným vstupným dátovým zdrojom v tejto práci. Každému klastru je následne priradené taxonomické zaradenie na základe referenčných databáz [42]. Toto však znamená, že taxonomické zaradenie je závislé od kvality referenčnej databázy [42]. Taktiež sa môže stať, že čítania pochádzajúce z rovnakého taxónu sú rozdelené do viacerých klastrov, pretože organizmy môžu mať aj viacej ako jednu kópiu markerového génu. Tieto kópie sa môžu líšiť v sekvencii nukleotidov [42].



# Kapitola 2

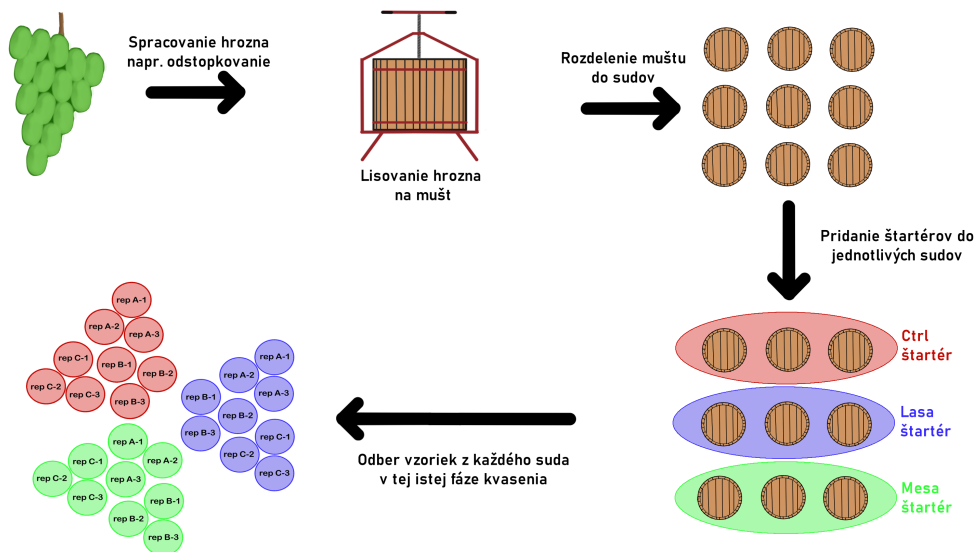
## Metodika práce

### 2.1 Použité dátové zdroje

Tento dátový zdroj pochádza zo vzoriek hrozna odrody Rulandské biele, ktoré bolo zozbierané v obci Modra, ktorá sa nachádza v Malých Karpatách počas rokov 2018 a 2019.

Zozbierané strapce hrozien z týchto rokov boli odstopkované a následné lisované. V roku 2018 sa vylisovaný mušt rozdelil rovnomerne do 5 sudov. Každý sud pre nás predstavuje jednu biologickú repliku. Následne do troch sudov bol pridaný kontrolný štartér pozostávajúci z komerčne dostupných kvasiniek. Do ďalšieho suda bol pridaný experimentálny štartér lasa obsahujúci *L. thermotolerans* a do ďalšieho experimentálny štartér mesa obsahujúci *M. pulcherrima*. Z roku 2019 bolo pripravených 9 sudov. Do troch sudov bol vložený kontrolný štartér, a ostávajúcich 6 boli vložené experimentálne štartéry, vždy jeden do troch sudov.

Následne všetky sudy prešli rovnakým procesom výroby vína v oboch rokoch. Počas tohto procesu sa vo fázach kvasenia definovaných v 1.2.2 Fázy kvasenia (m1 - mušt, m2 - fermentovaný mušt, m3 - mladé víno pred filtráciou) odobrali z každejologickej repliky tri vzorky (technické repliky). Na obrázku 2.1 môžeme vidieť diagram, ktorý znázorňuje akým spôsobom boli získané jednotlivé vzorky v tomto dátovom zdroji.



Obr. 2.1: Diagram, ktorý zobrazuje akým spôsobom boli získané jednotlivé vzorky v dátovom zdroji použitom v tejto práci. Vzorky boli získavané v rokoch 2018 a 2019. V oboch týchto rokoch bolo zozbierané hrozno, ktoré bolo následne odstopkované a vylišované na mušt. Tento mušt bol rozdelený do niekoľko sudov. Následne sa do každého suda pridal niektorý zo štartérov. V roku 2018 sa do 3 sudov pridal kontrolný štartér, do 1 suda štartér lasa a do 1 suda štartér mesa. V roku 2019 sa pridal každý z týchto štartérov do 3 sudov. Následne sa v jednotlivých fázach kvasenia odobrali vzorky.

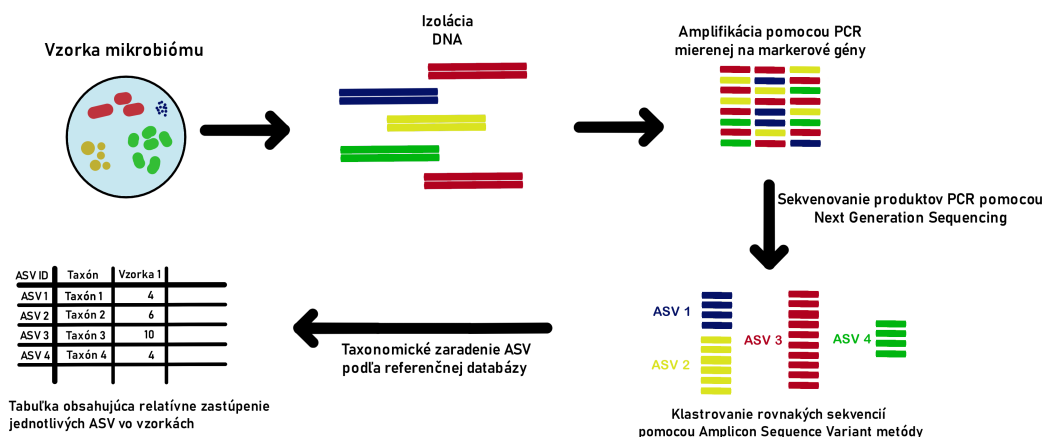
### 2.1.1 Laboratórne spracovanie

Z odobratých vzoriek bola izolovaná DNA a RNA. DNA obsahuje informáciu o prítomnosti jednotlivých mikroorganizmov vo vzorke a RNA o tom, ako veľmi aktívne sú jednotlivé mikroorganizmy. Tento materiál bol následne amplifikovaný pomocou metódy PCR. Primery pre amplifikáciu cielili na markerové gény bakteriálnej 16S rRNA a rDNA a fungálnej 28S rRNA a rDNA. Následne vytvorená knižnica bola podrobená cieľnému amplikónovému masívnemu paralelnému sekvenovaniu systémom MiSeq od firmy Illumina Inc. Výstupom tohto sekvenovania sú čítania, ktoré poskytujú informácie o genomickom zložení amplifikovaných častí markerových génov mikroorganizmov vo vzorkách.

### 2.1.2 Určenie zastúpenia taxonomických druhov

Sekvenačné dáta boli následne podrobené bioinformatickej analýze. Čítania získané zo sekvenovania boli klastrované metódou ASV vysvetlenou v sekcii 1.3.1 Amplicon Sequence Variant. Pri taxonomickej klasifikácii jednotlivých ASV klastrov bola využitá

databáza SILVA v138.1 [31] pre 16S rRNA a rDNA čítania a databáza RefSeq v220113 [28] pre 28S rRNA a rDNA čítania. Podrobnejšie informácie o spracovaní vzoriek vína a sekvenčných dát sa nachádzajú v dizertačnej práci Modelové aplikácie charakterizácie mikrobiálneho zloženia pomocou cieleného amplikónového masívne paralelného sekvenovania vo vinárskom výskume od Mgr. Miroslav Böhmer, PhD [10]. Na obrázku 2.2 môžeme vidieť diagram laboratórneho spracovania vzoriek.



Obr. 2.2: Diagram, ktorý zobrazuje akým spôsobom boli z jednotlivých vzoriek získané tabuľky relatívneho zastúpenia ASV, ktoré sú vstupné dáta pre túto prácu. Zo vzoriek mikrobiómu je izolovaná DNA (alebo RNA). Táto DNA (alebo RNA) je následne amplifikovaná pomocou PCR, ktorej primery sú cielené na markerové gény. V našom prípade to bolo 16S rRNA gén pre baktérie a 28S rRNA gén pre huby. Následne sa produkty PCR sekvenujú pomocou Next Generation Sequencing. Čítania sa klastrujú dokopy metódou ASV popisovanie v sekcii 1.3.1 Amplicon Sequence Variant a následne je každému klastru priradené taxonomické zaradenie podľa referenčnej databázy. Informácia o vzorke je zapísaná do tabuľky relatívnych zastúpení, kde stĺpce sú jednotlivé vzorky v riadku sú jednotlivé ASV klastre s taxonomickým zaradením a relatívnym zastúpením vo vzorke.

Vstupné dáta do tejto práce získané spôsobom popísaným vyššie boli dostupné v tabuľkovom formáte, V prvom stĺpci je unikátny identifikátor ASV klastru. V druhom stĺpci je sekvencia markerového génu, špecifická pre daný ASV klastr. V treťom stĺpci je celé taxonomické zaradenie. Všetky nasledujúce stĺpce obsahujú merania z analyzovaných vzoriek, teda počet čítaní prislúchajúci danému ASV klastru. Riadky reprezentujú jednotlivé ASV klastre. Máme dve takéto tabuľky, jednu pre bakteriálne 16S dáta a druhú pre fungálne 28S dáta.

V týchto tabuľkách sa mohlo stať, že niektorému taxonomickému zaradeniu patrilo viacej riadkov. Toto sa deje z dôvodu, že čítania pochádzajúce z rovnakého taxónu môžu byť rozdelené do viacerých ASV klastrov, pretože organizmy môžu mať aj viacej ako jednu kópiu markerového génu a tieto kópie sa môžu líšiť v sekvencii nukleotidov. Pri



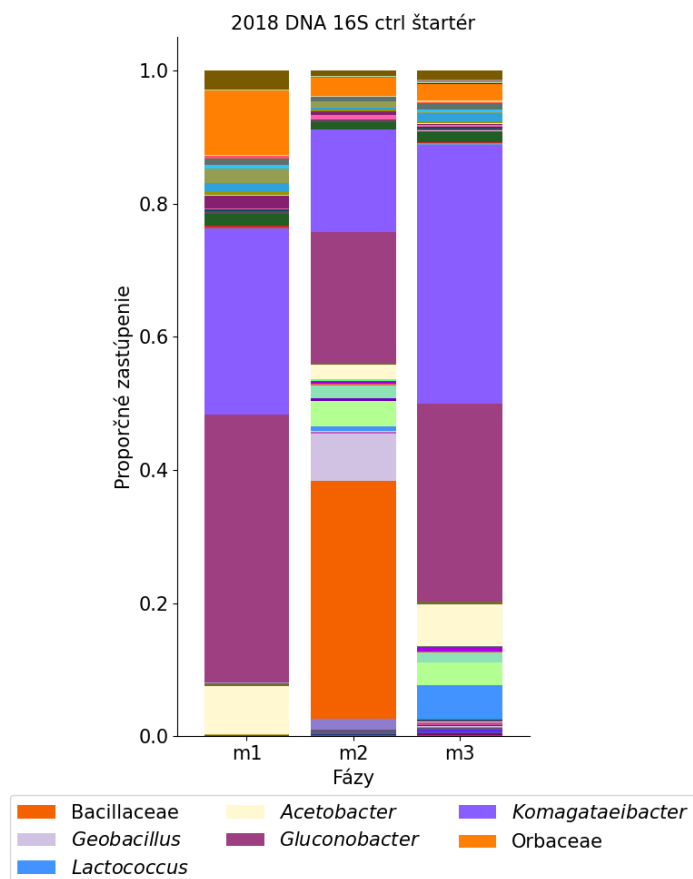
niektorých našich analýzach by sa mohlo stať, že rovnaký taxón uvidíme vo výslednom grafe alebo tabuľke viackrát. Tomuto sa chceme vyhnúť, pretože by to robilo interpretáciu výsledkov ťažšiu. Riadky sme sa snažili spájať podľa rodového taxonomického zaradenia. Ak riadok nemal rodové zaradenie, tak sme ho spájali podľa čeľade. Ak nemal ani zaradenie do čeľade, tak sme ho spájali podľa ešte vyššieho taxonomického zaradenia a tak ďalej, dokým neboli pospájané všetky riadky s rovnakým taxonomickým zaradením. Na takto upravených tabuľkách boli potom vykonané ďalšie analýzy opísané nižšie v tejto kapitole.

## 2.2 Spracovanie taxonomických počtov

### 2.2.1 Vývoj mikrobiómu pri kvasení vína

Pri tejto analýze nás zaujímalo percentuálne zastúpenie taxónov v jednotlivých vzorkách počas troch sledovaných fáz kvasenia. Chceli sme hlavne vedieť, ktoré sú majoritné taxóny, teda ktoré majú najväčšie zastúpenie vo vzorkách.

Pri identifikácii majoritných taxónov sme sa rozhodli na vizualizáciu použiť stacked bar chart z Python knižnice Matplotlib [18]. Tieto grafy sme zvolili z toho dôvodu, že je ľahko vidno, ktoré taxóny majú najväčšie zastúpenie v určitej fáze kvasenia. Ukážku tohto grafu môžeme vidieť na obrázku 2.3. Na osi x máme jednotlivé fázy kvasenia a na osi y je percentuálne zastúpenie taxónu vo vzorke. Keďže v jednej vzorke mohlo byť namerané veľké množstvo rôznych taxónov, do finálnej legendy sme sa rozhodli uviesť len také, ktoré mali percentuálne zastúpenie aspoň 5%. Takto sme vizuálne poukázali, na ktoré taxóny sa oplatí pozrieť bližšie, pretože tieto majú najväčší podiel na procese kvasenia vína.



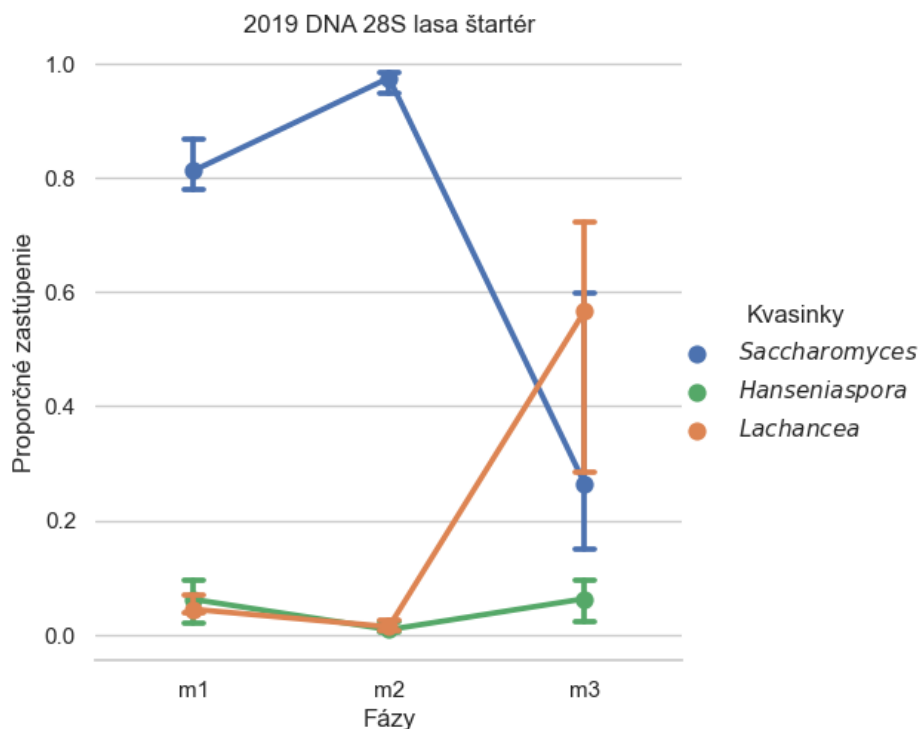
Obr. 2.3: Ukážka stacked bar chart grafu, ktoré sme použili na vizualizáciu taxonomickeho zastúpenia mikrobiómu v rôznych skupinách vzoriek počas fáz kvasenia. Každý stĺpec predstavuje jednu fázu kvasenia. Na osi y je proporčné zastúpenie taxónov. Nadpis hovorí o aké typy vzoriek ide. V legende sa nachádzajú iba taxóny, ktoré v niektorej fáze kvasenia mali zastúpenie aspoň 5%.

### Prežívanie kvasiniek z experimentálnych štartérov

Pri fungálnych vzorkách sme identifikovali, že väčšinové zastúpenie majú kvasinky z rodu *Saccharomyces* a *Hanseniaspora*. Chceli sme teda porovnať, ako sa pri procese kvasenia darí kvasinkám, ktoré sme pridali do vzoriek v podobe našich dvoch experimentálnych štartérov v porovnaní s týmito dvoma dominantnými kvasinkami. Chceli sme vidieť, aký je medián proporčného zastúpenia a aká je maximálna a minimálna hodnota proporčného zastúpenia jednotlivých kvasiniek vo fázach kvasenia. Rozhodli sme sa to vizualizovať cez čiarové grafy.

Pre vizualizáciu sme použili Python knižnicu seaborn [41]. Ukážku grafu môžeme vidieť na obrázku 2.4. Na x osi sú fázy kvasenia a na y osi je percentuálne zastúpenie taxónov. Môžeme vidieť, že pri každej fáze je na grafe ukazovateľ rozsahu dát. Tento ukazovateľ v tomto prípade ukazuje maximálne a minimálne zastúpenie taxónu vo

všetkých vzorkách v určitej fáze pre ten konkrétny štartér v konkrétnom roku. Analýzy sme robili zvlášť pre DNA a RNA.



Obr. 2.4: Ukážka čiarových grafu, ktoré sme použili, aby sme zistili, či kvasinky z experimentálnych štartérov prežívajú v jednotlivých fázach kvasenia. Na osi x sú jednotlivé fázy kvasenia a na osi y je proporčné zastúpenie kvasinky vo fáze. Body predstavujú medián proporčného zastúpenia kvasinky vo vzorkách a ukazovatele rozsahu ukazujú maximálne a minimálne proporčné zastúpenie kvasinky vo vzorkách. Nadpis grafu popisuje vzorky použité ako vstupné dáta do grafu.

## 2.3 Diferenčná analýza

Pri diferenčnej analýze sme sa chceli pozrieť na to, aké sú rozdiely v taxonomickom zložení medzi kontrolným štartérom a experimentálnymi štartérmi. Túto analýzu sme sa rozhodli vykonať pomocou knižnice v jazyku R, DESeq2 [32][23].

Diferenčná analýza sa robila vždy pre vzorky z rovnakého roku, z rovnakej fázy kvasenia a z rovnakého genetického materiálu, DNA alebo RNA. Taktiež aby malo zmysel robiť takúto analýzu, každá porovnávaná skupina musela mať aspoň 3 vzorky. Takéto vzorky sa potom vložili do tabuľky, kde v prvom stĺpci sa nachádzajú mená vzoriek. V prvom riadku ostatných stĺpcov sa nachádzajú mená jednotlivých taxónov. Tabuľka je vyplnená nameranými hodnotami ASV. Okrem tejto tabuľky je potrebná aj tabuľka obsahujúca údaje o jednotlivých vzorkách. V našom prípade v tejto tabuľke

v prvom stĺpci sú mená vzoriek a v druhom stĺpci označenie, či sa jedná o kontrolnú skupinu alebo štartér.

### 2.3.1 DESeq2

V DESeq2 sa počty modelujú aby zodpovedali negatívnemu binomickému rozdeleniu [23]. Negatívne binomické rozdelenie má 2 parametre. Prvým je priemer, ktorý je vypočítaný, na základe kvantít  $q_{ij}$  proporčnej ku množstvu zastúpenia taxónu vo vzorke. Druhý parameter je disperzia, ktorá je vypočítaná na základe log-normálnej distribúcie závislej od priemeru normalizovaných počtov zastúpenia taxónu vo vzorkách [23].

DESeq2 vypočíta veľa parametrov, avšak my sme sa zamerali len na logarithmic fold change (LFC) a upravenú P hodnotu. LFC nám hovorí o zmenilo zastúpenie organizmu vo vzorke v kontrolnej skupine oproti experimentálnej skupine [23]. P hodnota je následne vypočítaná na základe LFC pomocou Wald testu [23]. Nulová hypotéza pri tomto teste je, že nie je rozdiel v zastúpení organizmov v kontrolnej a experimentálnej skupine [23]. Táto P hodnota je ešte upravená Benjamini-Hochberg metódy, pretože pri organizmoch s nízkym počtom môže dôjsť ku chybnému výpočtu P hodnoty [23].

## 2.4 Alfa diverzita

Alfa diverzita vyjadruje bohatosť komunity organizmov v rámci jedného ekosystému, v našom prípade v rámci jednej analyzovanej vzorky. Alfa diverzitu je možné kvantifikovať rôznymi spôsobmi. Tieto metriky typicky vyjadrujú buď počet nameraných druhov, rovnomernosť zastúpenia týchto druhov, prípade kombináciu týchto metrík. Pre túto prácu sme vybrali 2 metriky. Jedna je ako bohatá je populácia a druhá je, ako rovnomerná je populácia. Ako bohatá je populácia sme jednoducho vyjadrili pomocou počtu rôznych ASV pozorovaných vo vzorkách. Ako rovnomerne sú jednotlivé taxóny zastúpené v populácii sme vyjadrili pomocou Shannonovho indexu. Na výpočet týchto metrík sme použili Python knižnicu scikit-bio [34].

### 2.4.1 Shannonov index

Shannonov index je bežne používaná metrika na vyjadrenie rovnováhy populácie vo vzorke. Táto metrika sa počíta pomocou nasledujúceho vzorca:

$$H = \frac{-\sum P_i \log P_i}{\log N}$$

$P_i$  je koľkokrát sa vo vzorke konkrétny druh nachádza a  $N$  je koľko rôznych druhov sa vo vzorke nachádza [29]. Ide teda o kombináciu bohatosti a rovnomernosti. Čím vyššia je táto hodnota, tým vyrovnanejšia je populácia vo vzorke. Hodnota 0 bude

znamenat', že v populácii sa nachádza len jeden druh organizmu. Maximálna hodnota tohto indexu môže byť  $\log N$ . Treba podotknúť, že hodnota vypočítaná týmto indexom je len relatívna informácia [29]. Dá sa porovnávať len s hodnotami Shannonovho indexu, ktoré sú vypočítané v rámci tejto práce. Nemôžeme teda porovnávať naše hodnoty Shannonovho indexu s hodnotami vypočítanými v iných prácach.

### 2.4.2 Rarefaction

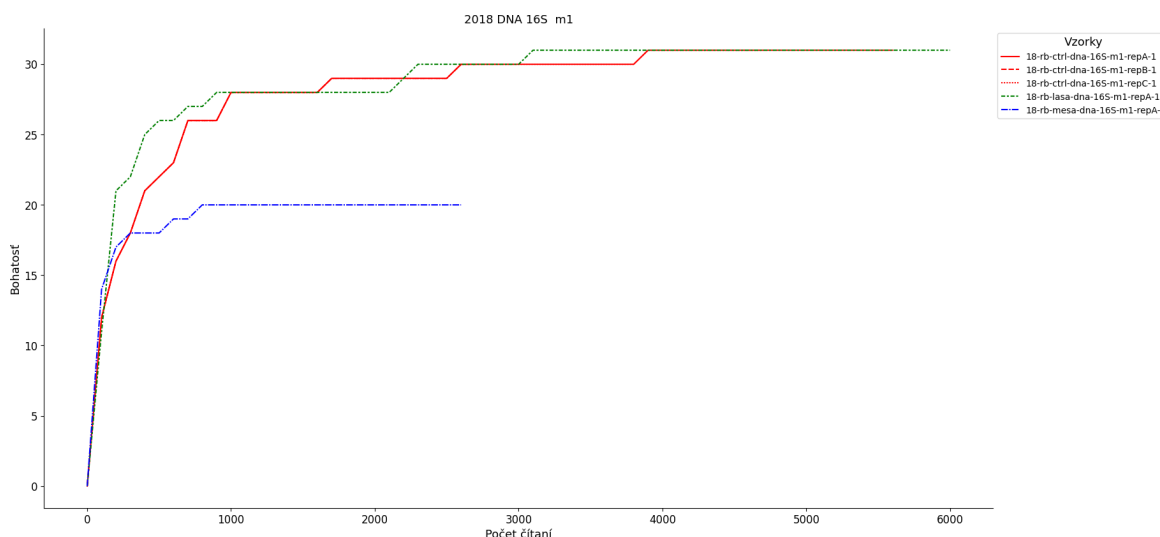
Cieľom techniky rarefaction je korigovať nevyvážené veľkosti vzoriek, ktoré sú jedným z hlavných problémov pri porovnávaní diverzít [15]. Vezmime si 2 vzorky ako príklad. Z prvej vzorky spravíme 5000 čítaní a z druhej vzorky spravíme 20000 čítaní. Pri menšom počte čítaní máme menšiu šancu nájsť málo zastúpené organizmy aj pri odbere z rovnakého prostredia. Teda v prvej vzorke je menšia pravdepodobnosť, že nameriame všetky organizmy nachádzajúce sa vo vzorke. Toto by sa mohlo premietnuť do inej hodnoty niektorej z metrick alfa diverzity ako by bola, keby sme vzorku osekvenovali hlbšie.

Výstupom tejto techniky je grafická vizualizácia v podobe rarefaction krivky, na základe ktorej je možné posúdiť, či vzorky boli dostatočne hlboko osekvenované, teda či sme namerali všetky organizmy vo vzorke.

Pri vytváraní takýchto kriviek sme zvolili nasledovný postup. Zo všetkých čítaní pre konkrétnu vzorku sme postupne náhodne vyberali 50 čítaní. Pre týchto 50 čítaní sme sa pozreli, koľko organizmov sme namerali. Toto sme opakovali desaťkrát a z vypočítaných hodnôt sme zobrali priemernú hodnotu. Následne sme toto opakovali pre 100 čítaní, 150 čítaní a tak ďalej, kým sme nedosiahli počet čítaní vo vzorke.

Ak sa dlhodobo nezvyšuje počet identifikovaných druhov so zvyšujúcim sa počtom čítaní, tak predpokladáme, že sme už identifikovali všetky rozumne zastúpené organizmy. Toto potom vidíme na rarefaction krivke, ktorá sa ustáli na jednej hodnote. Nejedná sa však o presnú metriku. Rarefaction krivky sú len vizuálne zhodnotenie, či sú vzorky dostatočne hlboko osekvenované.

Na osi x je počet čítaní a na y osi je metrika alfa diverzity. Každá krivka predstavuje jednu vzorku. V legendu sú aj názvy konkrétnych vzoriek. Nadpis opisuje typ vzoriek, ktoré sú vizualizované na grafe.



Obr. 2.5: Ukážka rarefaction grafu na vizuálne zhodnotenie, či sú vzorky dostatočne hlboko osekvenované. Na osi x je počet čítaní a na osi y je jedna z metrik alfa diverzity. V tomto konkrétnom prípade to je bohatosť vzoriek. Nadpis popisuje aké vzorky boli použité ako vstupné dáta do grafu. V legende sú aj názvy konkrétnych vzoriek.

### 2.4.3 Porovnanie alfa diverzít

Porovnanie jednotlivých alfa diverzít sme vykonali nasledovne. Najskôr sme sa pozreli, ako sa líšia metriky diverzity v jednotlivých fázach kvasenia na pri rôznych štartéroch štartéru. Vzorky sme rozdelili podľa rokov a podľa toho, či sa jedná o DNA alebo RNA. Následne sme tieto dáta vizualizovali pomocou boxplotov.

## 2.5 Beta diverzita

Hlavnou myšlienkou beta diverzity je zmerať ako sa mení štruktúra populácie medzi vzorkami navzájom [4]. Vzdialenosť nám umožní kvantifikovať rozdiely medzi vzorkami napríklad na časovom gradiente alebo územnom gradiente. Naš časový gradient sú buď jednotlivé fázy kvasenia alebo rôzne roky odberu vzoriek. Naš ďalší gradient sú rôzne štartéry, ktoré boli použité. Tak ako pri alfa diverzite, tak aj beta diverzita sa dá popísať rôznymi metrikami. Pre našu analýzu sme si vybrali Bray-Curtis metóda [5]. Táto metrika sa bežne používa a je aj vhodná pre naše dáta, pretože pracuje len s tabuľkami počtov, čo sú presne aj naše ASV tabuľky. Pre výpočet tejto metriky sme použili opäť Python knižnicu scikit-bio [34].

### 2.5.1 Bray-Curtis metóda

Pri tejto metóde sa ako vstupné dáta používa zastúpenie druhov vo vzorkách, ktorú v našej dátovej sade predstavujú počty čítaní v ASV klastroch. Následne sa táto metóda snaží vypočítať maticu vzdialeností medzi každou dvojicou vzoriek [5]. Hodnoty matice sú počítané nasledujúcim vzorcom:

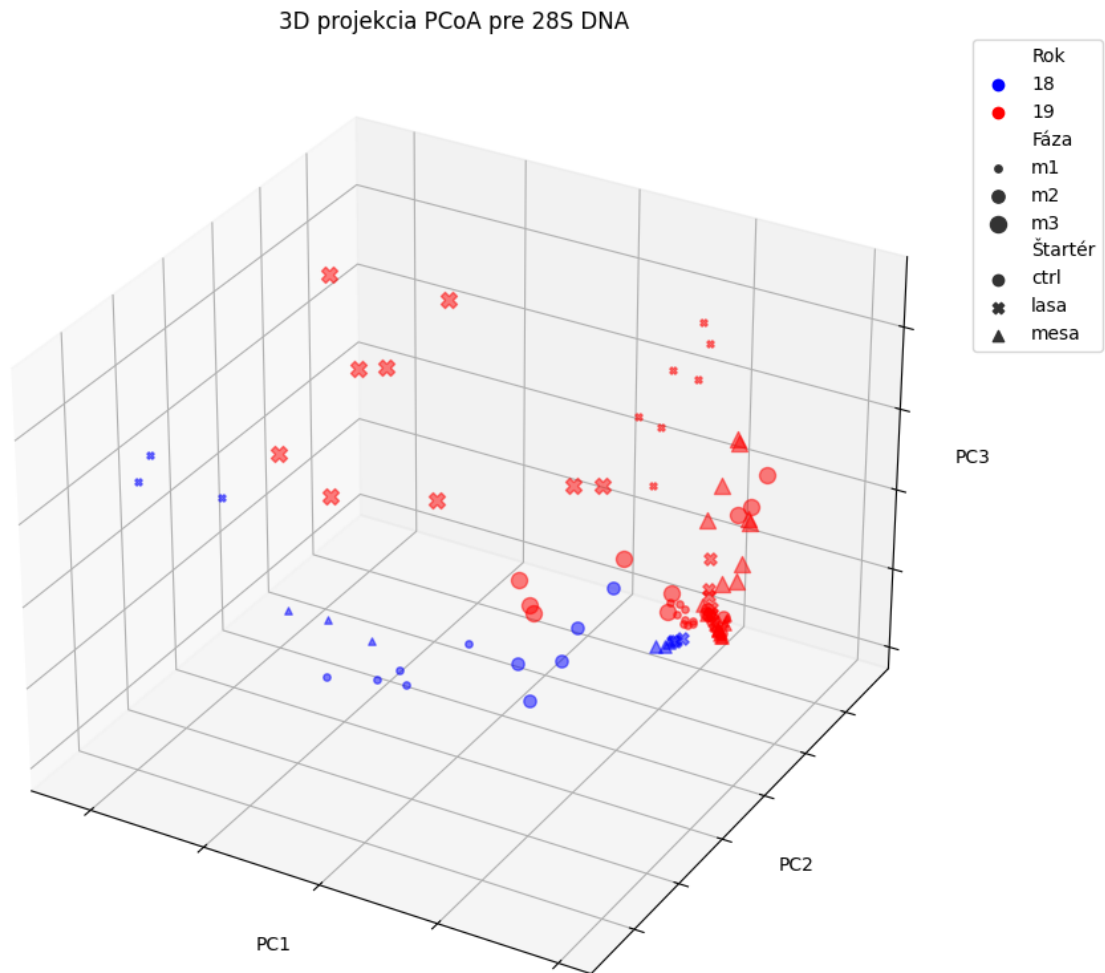
$$BC_{UV} = \frac{\sum_{j=1}^S |x_{Uj} - x_{Vj}|}{\sum_{j=1}^S (x_{Uj} + x_{Vj})}$$

$S$  je celkový počet druhov v porovnávaných vzorkách a  $x_{Uj}$  a  $x_{Vj}$  sú hodnoty zastúpenia druhu vo vzorkách [33].

Pri výpočte sme vzorky rozdelili podľa toho, či sa merala DNA alebo RNA. Následne pre každú túto vzorku sme vypočítali maticu vzdialeností.

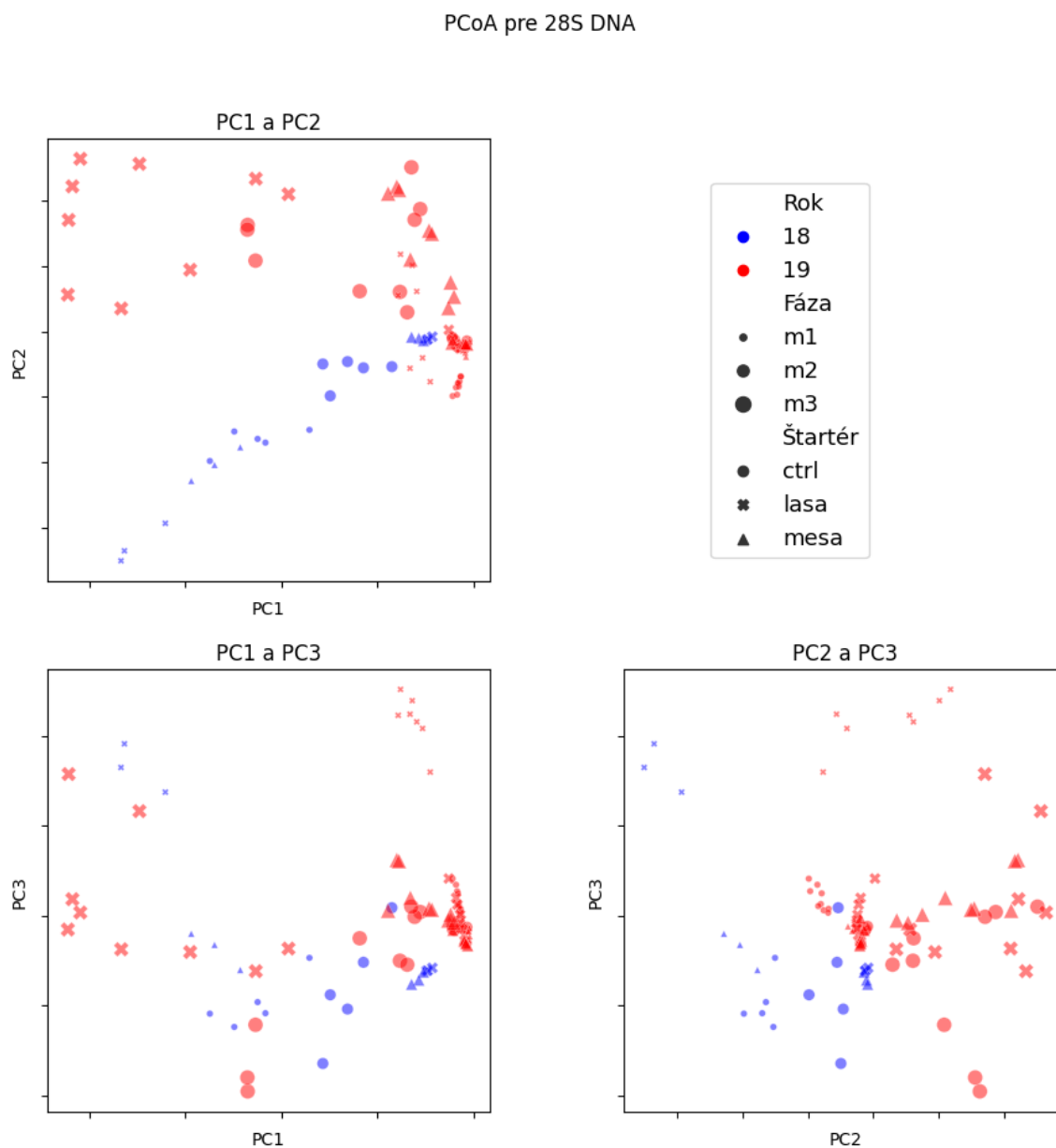
### 2.5.2 Analýza hlavných súradníc

Z matíc vzdialeností sme vypočítaných metódou vyššie sme spravili analýzu hlavných súradníc (ďalej len PCoA). Opäť knižnica scikit-bio [34] nám priamo poskytla už implementovanú funkciu. Pomocou PCoA vieme previesť našu maticu relatívnych vzdialeností beta diverzity do karteziánskej sústavy súradníc. Takýmto spôsobom následne môže vizualizovať beta diverzity našich vzoriek. Vzorky blízko pri sebe na grafe PCoA budú blízko pri sebe aj čo sa týka beta diverzity. PCoA sme spravili pre tri komponenty, teda naša PCoA sa dá zobrazíť v 3D priestore alebo môžeme porovnávať jednotlivé komponenty navzájom. Na obrázku 2.6 môžeme vidieť príklad 3D projekcie a na obrázku 2.7 môžeme vidieť príklad porovnanie jednotlivých komponentov navzájom.



Obr. 2.6: Ukážka 3D projekcie PCoA grafu. Na osiach sú jednotlivé komponenty PCoA. Body sú označené podľa roku, fázy kvasenia a štartéru. Body z roku 2018 sú označené modrou a z roku 2019 označené červenou. Fázy sú znázornené rozdielnymi veľkosťami bodov. Štartéry sú znázornené podľa rôzneho tvaru bodov.





Obr. 2.7: Ukážka grafov porovnanie komponentov PCoA navzájom. Na osiach sú jednotlivé komponenty PCoA. Body sú označené podľa roku, fázy kvasenia a štartéru. Body z roku 2018 sú označené modrou a z roku 2019 označené červenou. Fázy sú znázornené rozdielnymi veľkosťami bodov. Štartéry sú znázornené podľa rôzneho tvaru bodov.

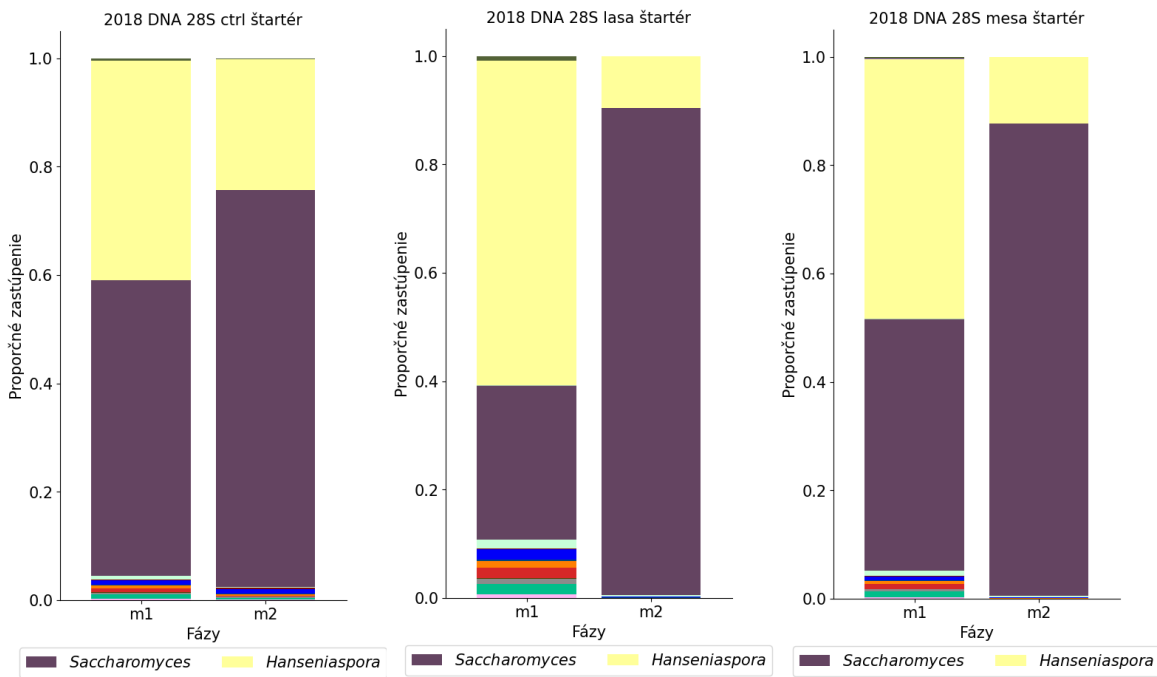
# Kapitola 3

## Výsledky

### 3.1 Určenie zastúpenia taxonomických druhov

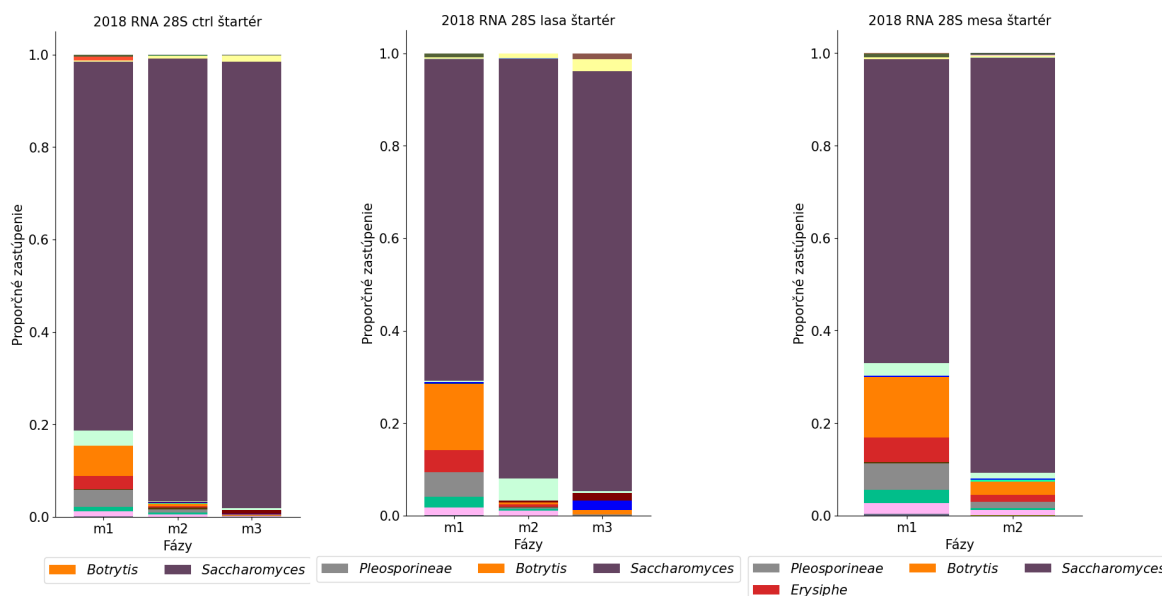
#### 3.1.1 Fungálne 28S dáta

Pre vzorky DNA z roku 2018, bez ohľadu na štartér dominujú kvasinky z rodu *Saccharomyces* vo fázach m1 a m2. Dáta pre vzorky m3 nám chýbajú. Vo väčšom zastúpení sa v týchto vzorkách taktiež nachádzajú kvasinky z rodu *Hanseniaspora*. Vo fáze m1 sú zastúpené na podobnej úrovni ako kvasinky z rodu *Saccharomyces*, ale vo fáze m2 ich zastúpenie značne klesne. V biologickej replike inokulovanej štartérom lasa vidíme prítomnosť týchto kvasiniek z rodu *Lachancea*, avšak ich percentuálna prítomnosť nie je ani 5%. Najväčšie zastúpenie má v m1 fáze a vo fáze m2 jej prítomnosť z pohľadu DNA klesá. V biologickej replike inokulovanej štartérom mesa nevidíme žiadnu prítomnosť DNA kvasiniek z rodu *Metschnikowia*. Toto naznačuje tomu, že inokulácia týmto štartérom bola neúspešná. Tieto výsledky môžeme vidieť na obrázku 3.1.



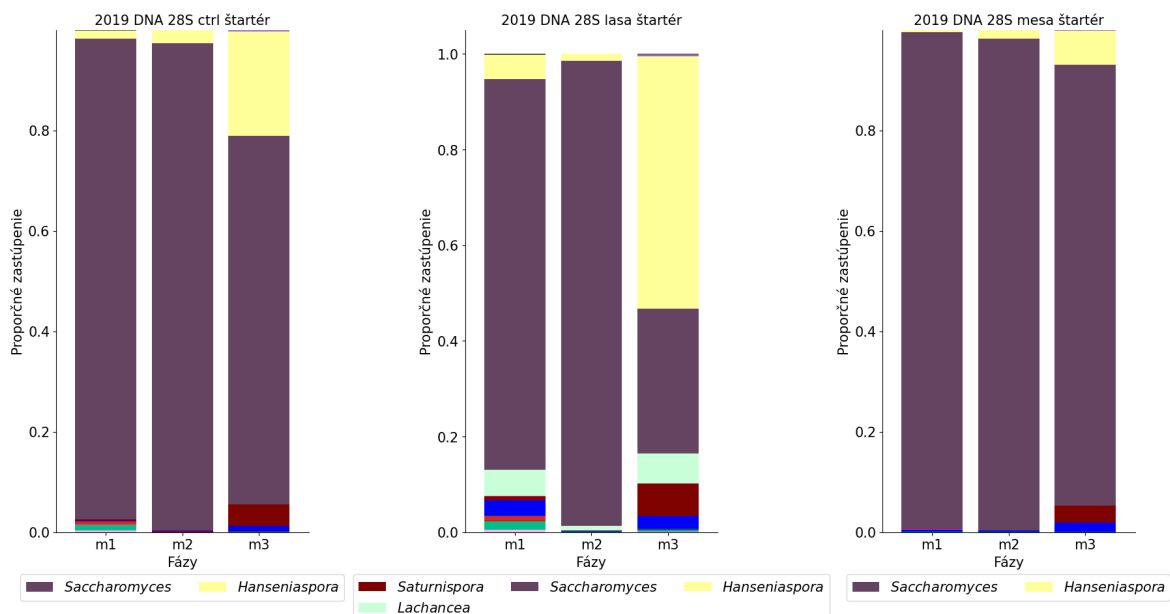
Obr. 3.1: Výsledok taxonomickej analýzy pre fungálne vzorky DNA roku 2018. Na obrázku sú tri grafy. Každý graf je pre jeden zo štartérov použitých v dátach. Na osi x sú fázy kvasenia. Na osi y je proporčné zastúpenie taxónov. Nadpis hovorí o aké typy vzoriek ide. V legende sa nachádzajú iba taxóny, ktoré v niektorej fáze kvasenia mali zastúpenie aspoň 5%.

Pri pohľade na vzorky RNA z roku 2018 vo všetkých vzorkách bez ohľadu na štartér a fázu dominujú kvasinky rodu *Saccharomyces*. Nevidíme veľkú aktivitu kvasiniek z rodu *Hanseniaspora* na úrovni RNA. Vo vzorkách pre štartér lasa bola nameraná RNA aktivita kvasiniek rodu *Lachancea* ale podobne ako pri DNA, nebolo to ani na úrovni 5%. Rozdiel oproti DNA bolo to, že najväčšia aktivita bola nameraná vo fáze m2. Pre štartér mesa je výsledok rovnaký ako pri DNA. Nevidíme žiadnu aktivitu kvasiniek rodu *Metschnikowia*. Pri meraní RNA môžeme vidieť aj nové taxóny, ktoré sme nezaregistrovali pri meraní DNA. Bez ohľadu na štartér bola vo vzorkách nameraná RNA aktivita *Botrytis*. Plesne tohto rodu infikujú povrch hrozna. Bežne sa považuje za rastlinný patogén, ale za určitých podmienok tieto plesne môžu spôsobiť zlepšenie kvality hrozna [39]. Tieto plesne sú aktívne hlavne v m1 fáze kvasenia. Ďalšie taxóny, ktorého RNA aktivitu sme zaregistrovali sú huby radu *Pleosporineae* a rodu *Erysiphe*. Tieto huby sú rastlinné patogény. Rad *Pleosporineae* spôsobuje odumieranie rôznych častí rastlín. Rod *Erysiphe* spôsobuje chorobu múčnatka viniča [14]. Aj pri týchto taxónoch môžeme pozorovať, že ich aktivita je najväčšia vo fáze m1 a postupne klesá, podobne ako pri plesni rodu *Botrytis*. Tieto výsledky môžeme vidieť na obrázku 3.2.



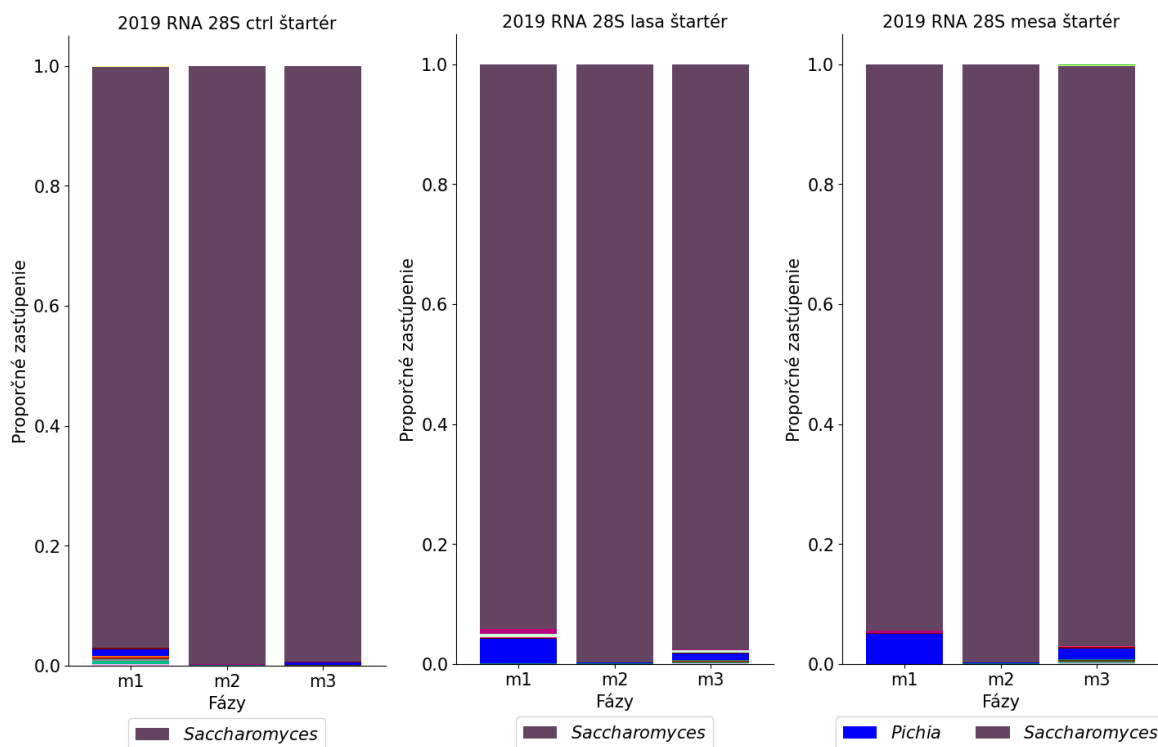
Obr. 3.2: Výsledok taxonomickej analýzy pre fungálne vzorky RNA roku 2018. Na obrázku sú tri grafy. Každý graf je pre jeden zo štartérov použitých v dátach. Na osi x sú fázy kvasenia. Na osi y je proporčné zastúpenie taxónov. Nadpis hovorí o aké typy vzoriek ide. V legende sa nachádzajú iba taxóny, ktoré v niektorej fáze kvasenia mali zastúpenie aspoň 5%.

Vzorky DNA z roku 2019 ukazujú ešte väčšiu dominanciu kvasiniek rodu *Saccharomyces*. Je vidieť aj prítomnosť DNA kvasiniek rodu *Hanseniaspora*. Pri každom štartéry prítomnosť týchto kvasiniek prudko stúpa vo fáze m3. Najväčšiu diverzitu DNA môžeme vidieť vo vzorkách, kde bol použitý štartér lasa. Podobne ako v roku 2018 je vidno prítomnosť kvasiniek rodu *Lachancea* v tomto štartéry ale na rozdiel od roku 2018 zastúpenie týchto kvasiniek bolo aspoň 5%. Okrem týchto sa vo vzorkách pre štartér lasa nachádzali aj kvasinky z rodu *Saturnispora*. O týchto kvasinkách sa mi však nepodarilo nájsť informácie o ich vplyve na proces kvasenia vína. Avšak vo vzorkách sú prítomné aj kvasinky rodu *Pichia*. *Pichia* a *Saturnispora* sú obe z čeľade *Pichiaceae*. Rod *Pichia* je bežne pozorovaný na povrchu hrozna a môže počas kvasenia ovplyvniť rôzne kvality vína. Je možné, že rod *Saturnispora* zohráva podobnú úlohu ako *Pichia* pri fázach kvasenia. Tieto výsledky môžeme vidieť na obrázku 3.3.



Obr. 3.3: Výsledok taxonomickej analýzy pre fungálne vzorky DNA roku 2019. Na obrázku sú tri grafy. Každý graf je pre jeden zo štartérov použitých v dátach. Na osi x sú fázy kvasenia. Na osi y je proporčné zastúpenie taxónov. Nadpis hovorí o aké typy vzoriek ide. V legende sa nachádzajú iba taxóny, ktoré v niektorej fáze kvasenia mali zastúpenie aspoň 5%.

Nakoniec vzorky RNA z roku 2019 ukazujú jasnú dominanciu rodu kvasiniek *Saccharomyces*. Vo všetkých fázach bez ohľadu na použitý štartér takmer úplne vytlačila všetky ostatné taxóny. Jedine pri štartéry mesa sme detekovali podiel RNA nad 5% pre kvasinky rodu *Pichia* vo fáze m1. Tieto výsledky môžeme vidieť na obrázku 3.4.

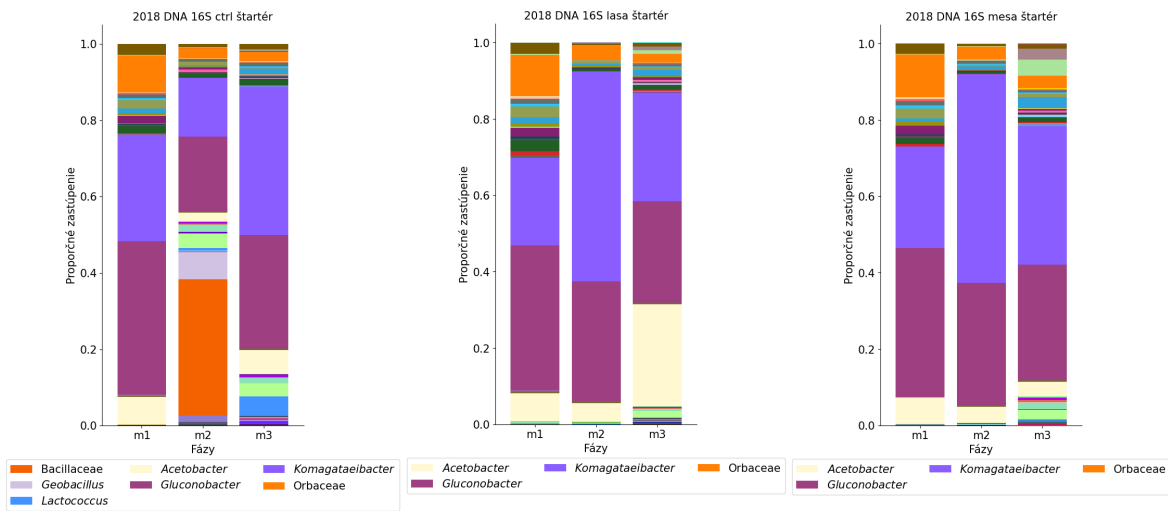


Obr. 3.4: Výsledok taxonomickej analýzy pre fungálne vzorky RNA roku 2019. Na obrázku sú tri grafy. Každý graf je pre jeden zo štartérov použitých v dátach. Na osi x sú fázy kvasenia. Na osi y je proporčné zastúpenie taxónov. Nadpis hovorí o aké typy vzoriek ide. V legende sa nachádzajú iba taxóny, ktoré v niektorej fáze kvasenia mali zastúpenie aspoň 5%.

### 3.1.2 Bakteriálne 16S dáta

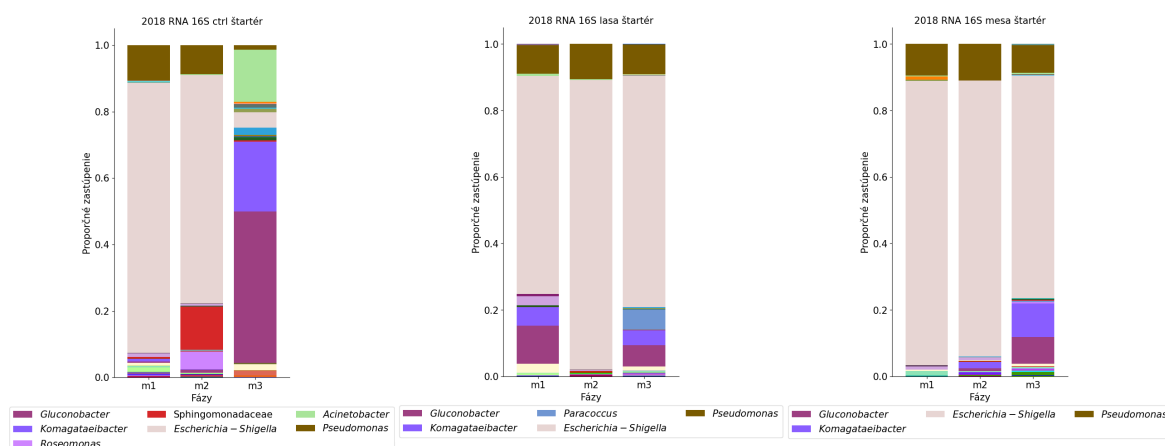
Bakteriálne komunity vo vzorkách sú rôznorodejšie než kvasinkové. Vzorky DNA z roku 2018 bez ohľadu na štartér obsahujú obsahujú baktérie z rodov *Gluconobacter*, *Acetobacter* a *Komagataeibacter*. Baktérie z týchto rodov sa bežne nachádzajú na povrchu hrozna a sú známe kazením vína. Taktiež sa nachádzajú aj v iných kvasených nápojoch, ako je napríklad kvas alebo kombucha a iné [38]. Tieto baktérie dokážu metabolizovať cukry a alkoholy na organické kyseliny. Napríklad vo víne metabolizujú etanol na kyselinu octovú [38]. Ďalej sa vo vzorkách bez ohľadu na štartér nachádzala čeľaď *Orbaceae*. Nie je známe, aký dopad na proces kvasenia by mohli mať tieto baktérie. Hlavne sú asociované s črevným traktom hmyzu [20]. Toto by mohlo vysvetľovať, ako sa dostali na povrch hrozna. Vo vzorkách, kde bol použitý kontrolný štartér boli vo fáze m2 namerané baktérie z čeľade *Bacillaceae* a taktiež rod tejto čeľade *Geobacillus*. Je zaujímavé, že tieto baktérie boli pod limitom detekcie vo vzorkách z fázy m1 a aj v neskoršej fáze m3. Je možné, že tieto baktérie boli v štádiu spory a teda nebolo možné izolovať ich DNA počas tohto štádia. Tieto baktérie sú schopné fermentácie [37], teda

ich prítomnosť a prežívanie vo víne nie je až tak prekvapujúca. Zvláštne je, že neboli namerané vo vzorkách iných štartérov a boli namerané len vo fáze m2. Tieto výsledky môžeme vidieť na obrázku 3.5.



Obr. 3.5: Výsledok taxonomickej analýzy pre bakteriálne vzorky DNA roku 2018. Na obrázku sú tri grafy. Každý graf je pre jeden zo štartérov použitých v dátach. Na osi x sú fázy kvasenia. Na osi y je proporčné zastúpenie taxónov. Nadpis hovorí o aké typy vzoriek ide. V legende sa nachádzajú iba taxóny, ktoré v niektorej fáze kvasenia mali zastúpenie aspoň 5%.

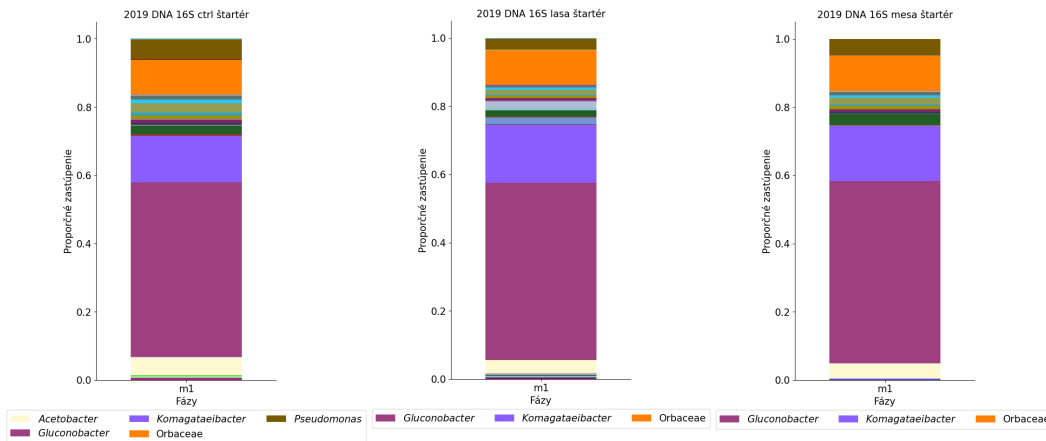
Vo vzorkách RNA z roku 2018 opäť vidíme zastúpenie baktérií z rodov *Gluconobacter*, *Acetobacter* a *Komagataeibacter* bez ohľadu na štartér. Ich aktivita stúpa vo fáze m3, pravdepodobne z dôvodu, že premieňajú alkohol prítomný vo víne na kyselinu octovú. Avšak bez ohľadu na štartér, vo vzorkách dominuje RNA aktivita baktérií z rodu *Escherichia-Shigella* a *Pseudomonas*. Akú úlohu pri procese kvasenia hrajú tieto baktérie nie je známe. Je však známe, že baktérie rodu *Escherichia-Shigella* môžu mať negatívny vplyv na zdravie človeka [24]. Tieto baktérie sa dajú nájsť aj v črevách ľudí a zvierat. Väčšina druhov je neškodných, ale niektoré môžu spôsobovať hnačku, vracanie alebo horúčku [27]. Tieto výsledky môžeme vidieť na obrázku 3.6.



Obr. 3.6: Výsledok taxonomickej analýzy pre bakteriálne vzorky RNA roku 2018. Na obrázku sú tri grafy. Každý graf je pre jeden zo štartérov použitých v dátach. Na osi x sú fázy kvasenia. Na osi y je proporčné zastúpenie taxónov. Nadpis hovorí o aké typy vzoriek ide. V legende sa nachádzajú iba taxóny, ktoré v niektorej fáze kvasenia mali zastúpenie aspoň 5%.

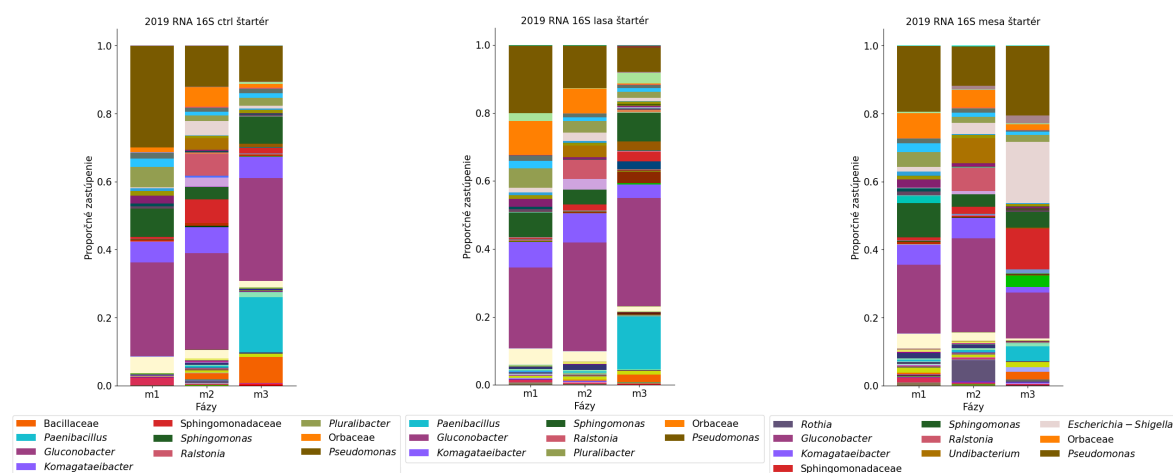
Pre vzorky z roku 2019 chýbajú dáta z merania DNA pre fázy m2 a m3. Teda môžeme vyhodnotiť len dáta z fázy m1. Fáza m1 vyzerá podobne ako v roku 2018. Opäť boli detekované hlavne baktérie z rodov *Gluconobacter*, *Acetobacter* a *Komagataeibacter*. Tieto výsledky môžeme vidieť na obrázku 3.7.





Obr. 3.7: Výsledok taxonomickej analýzy pre bakteriálne vzorky DNA roku 2019. Na obrázku sú tri grafy. Každý graf je pre jeden zo štartérov použitých v dátach. Na osi x sú fázy kvasenia. Na osi y je proporčné zastúpenie taxónov. Nadpis hovorí o aké typy vzoriek ide. V legende sa nachádzajú iba taxóny, ktoré v niektorej fáze kvasenia mali zastúpenie aspoň 5%.

Mikrobiálne populácie vzoriek RNA z roku 2019 vyzerajú oveľa vyváženejšie ako tie z roku 2018. Nie je tu taká jasná dominancia baktérií rodu *Escherichia-Shigella*, avšak stále sú tieto baktérie prítomné. V tomto roku je ich prítomnosť vo väčšine vzoriek pod 5%. Taktiež stále vidíme RNA aktivitu baktérií z rodov *Gluconobacter*, *Acetobacter* a *Komagataeibacter*. Ich aktivita na rozdiel od roku 2018, kde sme videli zvýšenú aktivitu vo fáze m3, vyzerá byť veľmi podobná vo všetkých fázach kvasenia. Oproti roku 2018 tu môžeme vidieť baktérie z rodu *Ralstonia*, *Sphingomonas*, *Pluralibacter* a *Paenibacillus*. Baktérie rodu *Ralstonia* sú patogén hlavne u rastlín, ale boli zaznamenané aj ako patogén u človeka. Preto by sa ich prítomnosť vo víne mala inhibovať [24]. Rod *Sphingomonas* je rastlinný a niekedy aj ľudský patogén, avšak tieto baktérie majú pozitívny vplyv na charakteristiky vína [8]. Baktérie z rodu *Paenibacillus* dokážu podporiť rast rastlín. V pôde dokážu premieňať dusík, fosfor a železo do formy, ktorú vedľa rastliny spracovať [17]. Taktiež dokážu chrániť rastliny pred rastlinnými patogénmi a bylinožravým hmyzom. Toto zabezpečuje vyvolávaním obranných mechanizmov rastliny alebo produkciou biocídnych látok (látky, ktoré ničia a odpudzujú škodlivé organizmy) [17]. Ich enzýmy majú využitie aj v medicíne a v priemysle [17]. Tieto výsledky môžeme vidieť na obrázku 3.8.

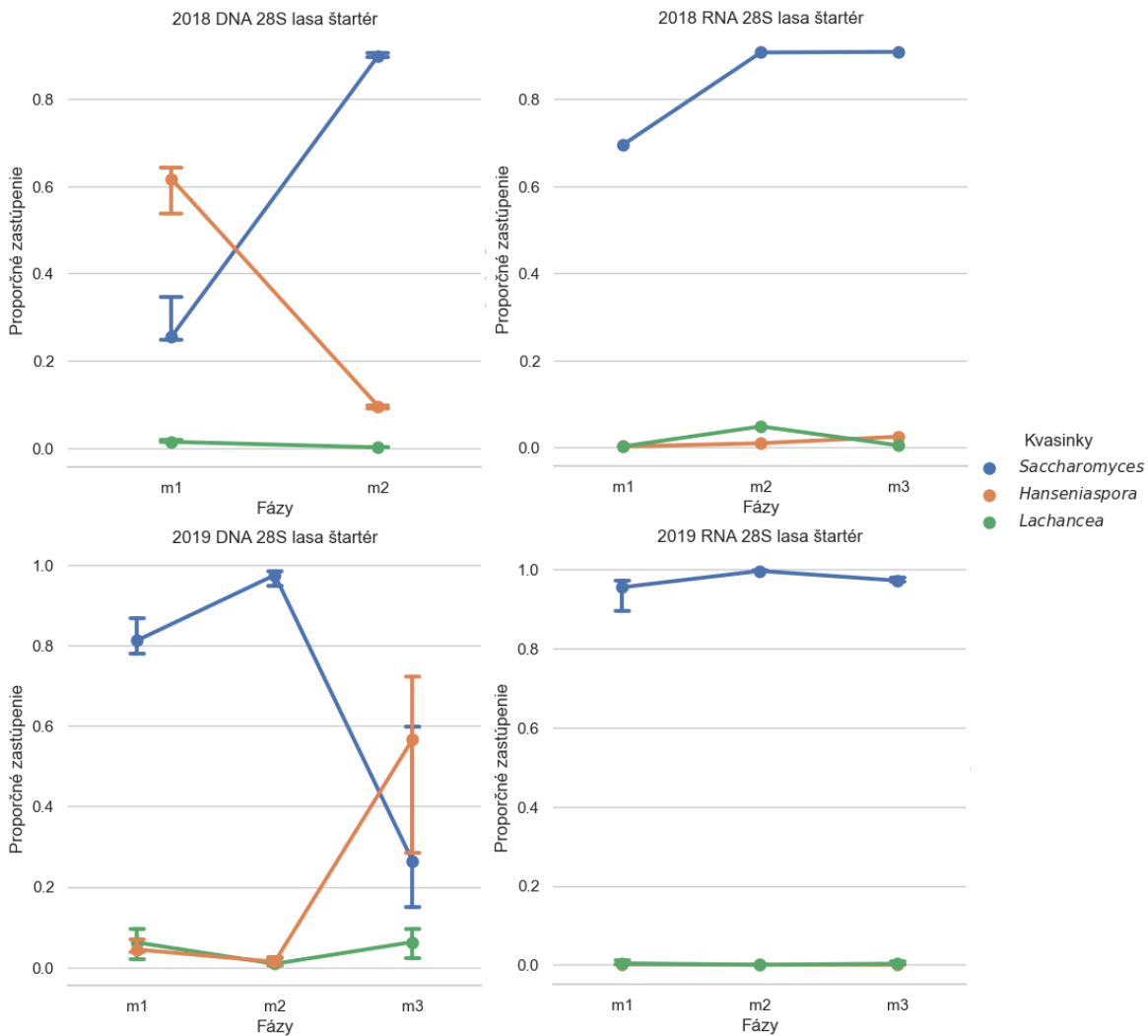


Obr. 3.8: Výsledok taxonomickej analýzy pre bakteriálne vzorky RNA roku 2019. Na obrázku sú tri grafy. Každý graf je pre jeden zo štartérov použitých v dátach. Na osi x sú fázy kvasenia. Na osi y je proporčné zastúpenie taxónov. Nadpis hovorí o aké typy vzoriek ide. V legende sa nachádzajú iba taxóny, ktoré v niektorej fáze kvasenia mali zastúpenie aspoň 5%.

### 3.1.3 Prežívanie experimentálnych štartérov

#### Štartér lasa

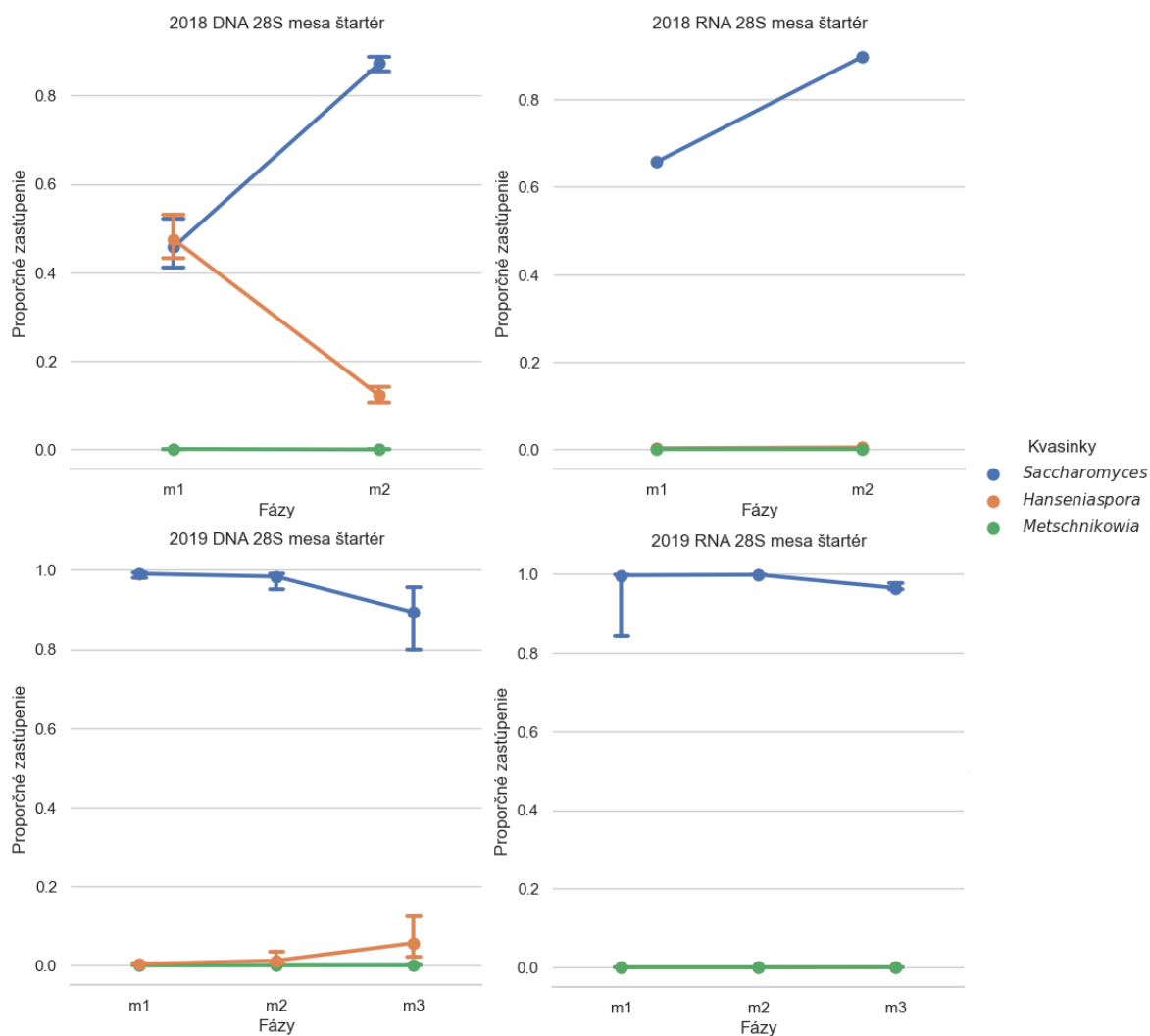
DNA týchto kvasiniek bola nameraná vo fáze m1 vo vzorkách z roku 2018 aj z roku 2019. Toto naznačuje, že inokulácia týmto štartérom bola úspešná. Taktiež v neskorších fázach m2 a m3 boli namerané. Avšak oproti kvasinkám rodu *Saccharomyces* je to zastúpenie kvasiniek štartéru veľmi nízke. Keď sa pozrieme na RNA, tak nameraná aktivita bola veľmi nízka, v niektorých fázach až nulová. Tieto výsledky vidíme na obrázku 3.9



Obr. 3.9: Prežívanie kvasiniek štartéru lasa. Na osi x sú jednotlivé fázy kvasenia a na osi y je proporčné zastúpenie kvasinky vo fáze. Body predstavujú medián proporčného zastúpenia kvasinky vo vzorkách a ukazovatele rozsahu ukazujú maximálne a minimálne proporčné zastúpenie kvasinky vo vzorkách. Nadpis grafu popisuje vzorky použité ako vstupné dáta do grafu.

### Štartér mesa

Nameraná DNA kvasiniek štartéru mesa bola nulová vo všetkých fázach v roku 2018 aj 2019. Taktiež RNA aktivita nebola nameraná v žiadnej fáze ani v roku 2018 ani 2019. Dôvodom tohto môže byť, že inokulácia týmto štartérom nebola úspešná ani v jednom roku. Tieto výsledky vidíme na obrázku 3.10



Obr. 3.10: Prežívanie kvasiniek štartéru mesa. Na osi x sú jednotlivé fázy kvasenia a na osi y je proporčné zastúpenie kvasinky vo fáze. Body predstavujú medián proporčného zastúpenia kvasinky vo vzorkách a ukazovatele rozsahu ukazujú maximálne a minimálne proporčné zastúpenie kvasinky vo vzorkách. Nadpis grafu popisuje vzorky použité ako vstupné dáta do grafu.

## 3.2 Diferenčná analýza

Výsledkom diferenčnej analýzy boli 3 tabuľky pre 28S dáta a 3 tabuľky pre 16S dáta. Každá tabuľka má na začiatku stĺpce, ktoré popisujú na akých dátach sa robila diferenčná analýza. V prvom stĺpci je rok. V druhom stĺpci je fáza kvasenia. V treťom stĺpci je, či sa jednalo o vzorky DNA lebo RNA. V štvrtom stĺpci je z ktorých biologických replík boli vzorky odobrané. V piatom a šiestom stĺpci je, ktoré dva štartéry boli porovnávané. zbytok tabuľky je vyplnený niektorou z hodnôt, vypočítanou pomocou DESeq2. V prvej tabuľke sú hodnoty korigovanej P-hodnoty. V druhej tabuľke je  $\log_2$

fold change. Táto hodnota hovorí o tom, ako sa líšilo zastúpenie taxónu vo vzorke medzi kontrolnou a experimentálnou skupinou. Ak bola táto hodnota záporná, tak väčšie zastúpenie bolo v kontrolnej skupine. Ak bola táto hodnota kladná, tak väčšie zastúpenie bolo v experimentálnej skupine. V tretej tabuľke sú pomocou x označené bunky tabuľky, ktoré mali v prvej tabuľke korigovanú P-hodnotu nižšiu ako 0.01. Takto sme si uľahčili analýzu tabuliek, lebo nám to umožňuje zamerať sa len na štatisticky významné výsledky. Podľa tejto poslednej tabuľky boli utriedené aj stĺpce všetkých 3 tabuliek. Stĺpec s najväčším počtom x v stĺpci je prvý v poradí. Takéto 3 tabuľky sme spravili aj pre fungálne 28S dáta aj pre bakteriálne 16S dát.

### 3.2.1 Fungálne 28S dáta

V roku 2018 boli všetky porovnania na úrovni DNA. V RNA dátach nebolo dostatok vzoriek, aby boli zmysluplné robiť tieto porovnania. Keďže sa jedná o štatistickú metódu, potrebujeme aj dostatok meraní na správne určenie rozdielov medzi vzorkami. Pri porovnávaní kontrolného štartéru so štartérom lasa boli ako štatisticky významné rozdiely označené kvasinky *Saccharomyces* vo fáze m1 s vyšším zastúpením vo vzorkách s kontrolným štartérom. Vo fáze m2 to boli *Saccharomyces*, *Lachancea* a *Filobasidium*. Iba rod *Filobasidium* mal vyššie zastúpenie vo vzorkách s kontrolným štartérom. Vo fáze m3 nebolo dostatok dát na vykonanie diferenčnej analýzy. Kvasinky rodu *Filobasidium* boli nájdené na podlahe, stenách a novom vybavení vinárni [3]. Produkujú enzým lipáza, ktorý by mohol byť využitý pri produkcii biopalív [9].

Pri porovnaní kontrolného štartéru a štartéru mesa boli ako štatisticky významné rozdiely označené kvasinky *Saccharomyces*, *Lachancea* a *Filobasidium* vo fáze m2. Opäť iba rod *Filobasidium* mal väčšie zastúpenie vo vzorkách s kontrolným štartérom. Vo fáze m1 neboli žiadne štatisticky významné rozdiely a vo fáze m3 nebolo dostatok dát na vykonanie diferenčnej analýzy.

Keď sme porovnali štartéry mesa a lasa, tak jediný štatisticky významný rozdiel bol vo fáze m1 v zastúpení kvasiniek rodu *Saccharomyces*. Väčšie zastúpenie mali vo vzorkách so štartérom mesa.

V roku 2019 máme porovnania aj z DNA aj z RNA dát. Pozrieme sa najskôr na DNA porovnania. Porovnanie kontrolného štartéru a lasa štartéru ukázalo ako štatisticky významné vo fáze m1 rody *Saccharomyces*, *Lachancea*, *Hanseniaspora*, *Pichia* a *Saturniaspora*. Všetky okrem rodu *Saccharomyces* boli vo väčšom zastúpení vo vzorkách so štartérom lasa. Vo fáze m2 bol označený iba rod *Lachancea*, ktorý mal väčšie zastúpenie vo vzorkách so štartérom lasa. Vo fáze m3 kvasinky z rodu *Saccharomyces* a *Lachancea*. *Saccharomyces* mali väčšie zastúpenie vo vzorkách s kontrolným štartérom a *Lachancea* vo vzorkách so štartérom lasa.

Pri porovnávaní kontrolného štartéru a štartéru mesa vo fáze m1 vidíme ako štatis-

ticky významné rody *Saccharomyces*, *Hanseniaspora*, *Pichia*, *Sacchotheciaceae*, *Nakaseomyces*, *Cladosporiaceae* a *Aspergillaceae*. Rody *Sacchotheciaceae* a *Cladosporiaceae* mali väčšie zastúpenie vo vzorkách s kontrolným štartérom. Ostatné mali väčšie zastúpenie vo vzorkách so štartérom mesa. Vo fáze m2 sa nenašli žiadne štatisticky významné rozdiely medzi mikrobiómami týchto dvoch štartérov. Vo fáze m3 boli označené rody *Saccharomyces*, *Hanseniaspora* a *Pichia*. Z týchto len *Hanseniaspora* bola vo väčšom zastúpení vo vzorkách s kontrolným štartérom.

Porovnanie štartérov lasa a mesa vo fáze m1 ukázal štatisticky významné rozdiely v zastúpení rodov *Saccharomyces*, *Lachancea*, *Nakaseomyces*, *Cladosporiaceae* a *Aspergillaceae*. Vo vzorkách so štartérom lasa mali väčšie zastúpenie kvasinky rodov *Lachancea* a *Cladosporiaceae*. Vo fáze m2 sa ako štatisticky významné ukázali rody *Lachancea* a *Pichia*. Rod *Lachancea* mal väčšie zastúpenie vo vzorkách so štartérom lasa a rod *Pichia* vo vzorkách so štartérom mesa. Vo fáze m3 boli označené za štatisticky významné rody *Saccharomyces*, *Lachancea*, *Hanseniaspora*, *Pichia* a *Sacchotheciaceae*. Vo vzorkách so štartérom lasa boli vo väčšom zastúpení rody *Lachancea*, *Hanseniaspora* a *Sacchotheciaceae*.

Z RNA dát máme v tomto roku porovnanie len z fáz m1 a m3. Vo fáze m2 nebolo dostatok dát na to, aby sme spravili diferenčné analýzy. Pri porovnaní kontrolného štartéru so štartérom lasa a kontrolného štartéru so štartérom mesa vo fáze m1 sa ukázali ako štatisticky významné taxóny *Filobasidium*, *Sacchotheciaceae*, *Nakaseomyces*, *Cladosporiaceae*, *Erysiphe*, *Pleosporineae*, *Borytis* a *Kazachstania*. Väčšie zastúpenie vo vzorkách s experimentálnym štartérom mali *Nakaseomyces* a *Kazachstania*. Ostatné mali väčšie zastúpenie vo vzorkách s kontrolným štartérom. Pri porovnaní kontrolného štartéru a štartéru lasa a pri porovnaní experimentálnych štartérov v tejto fáze bol označený aj rod *Lachancea* a mal väčšie zastúpenie vo vzorkách so štartérom lasa. Vo fáze m3 je štatisticky významný len rod *Lachancea* pri porovnaní kontrolného a lasa štartéru s väčším zastúpením vo vzorkách so štartérom lasa.

### 3.2.2 Bakteriálne 16S dáta

V roku 2018 máme porovnanie len pre vzorky DNA z fázy m3. V tejto fáze pri porovnaní kontrolného a lasa štartéru vidíme, že štatisticky významný rozdiel je v baktériách *Komagataeibacter* a *Acetobacter*. *Komagataeibacter* mal väčšie zastúpenie vo vzorkách s kontrolným štartérom a *Acetobacter* vo vzorkách so štartérom lasa.

Pri porovnaní kontrolného a mesa štartéru sú ako štatisticky významné označené baktérie *Komagataeibacter*, *Acinetobacter* a *Acidovorax*. Iba *Komagataeibacter* mal väčšie zastúpenie vo vzorkách s kontrolným štartérom.

Pri porovnaní lasa a mesa štartérov vidíme ako štatisticky významné s väčším zastúpením vo vzorkách s mesa štartérom baktérie *Acinetobacter*. Vo vzorkách s lasa

štartérom baktérie *Acetobacter*.

V roku 2019 pre vzorky DNA máme diferenčné analýzy len pre fázu m1. Pri porovnaní kontrolného a lasa štartéru vidíme ako štatisticky významné baktérie *Komagataeibacter*, *Corynebacterium*, *Gluconobacter*, *Paracoccus*, *Parapusillimonas*, *Alicycliphilus* a *Diaphorobacter*. Všetky mali väčšie zastúpenie vo vzorkách so štartérom lasa.

Porovnanie kontrolného a mesa štartéru ukázalo ako štatisticky významné populácie baktérií *Komagataeibacter* a *Gluconobacter*. Väčšie zastúpenie mali vo vzorkách so štartérom mesa.

Pri porovnaní štartérov lasa a mesa boli označené ako štatisticky významné a mali väčšie populácie v štartéry lasa baktérie *Komagataeibacter*, *Paracoccus*, *Parapusillimonas*, *Alicycliphilus* a *Diaphorobacter*.

Vo vzorkách kde sa merala RNA máme porovnania zo všetkých troch fáz. Pri pohľade na porovnania kontrolného a lasa štartéru vo fáze m1 sú štatisticky významné baktérie *Enterococcus*, *Lactobacillus*, *Paenibacillus* a *Escherichia-Shigella*. Všetky boli vo väčšom zastúpení vo vzorkách so štartérom lasa. Vo fáze m2 sa ukázali ako štatisticky významné baktérie *Streptococcus* a *Xylophilus*. *Streptococcus* mali väčšie zastúpenie vo vzorkách so štartérom lasa a *Xylophilus* vo vzorkách s kontrolným štartérom. Vo fáze m3 boli štatisticky významné baktérie *Caulobacter* a *Bradyrhizobium*, obe s väčším zastúpením vo vzorkách so štartérom lasa.

Porovnávanie kontrolného a mesa štartéru vo fáze m1 označilo ako štatisticky významné baktérie *Streptococcus*, *Enterococcus*, *Lactobacillus*, *Paenibacillus* a *Escherichia-Shigella*. Vždy mali väčšie zastúpenie vo vzorkách so štartérom mesa. Vo fáze m2 sa našli ako štatisticky významné len baktérie *Xylophilus* s väčším zastúpením vo vzorkách s kontrolným štartérom. Fáza m3 bola na tom podobne, rozdiel je, že v tomto prípade sa jednalo o baktérie *Corynebacterium*.

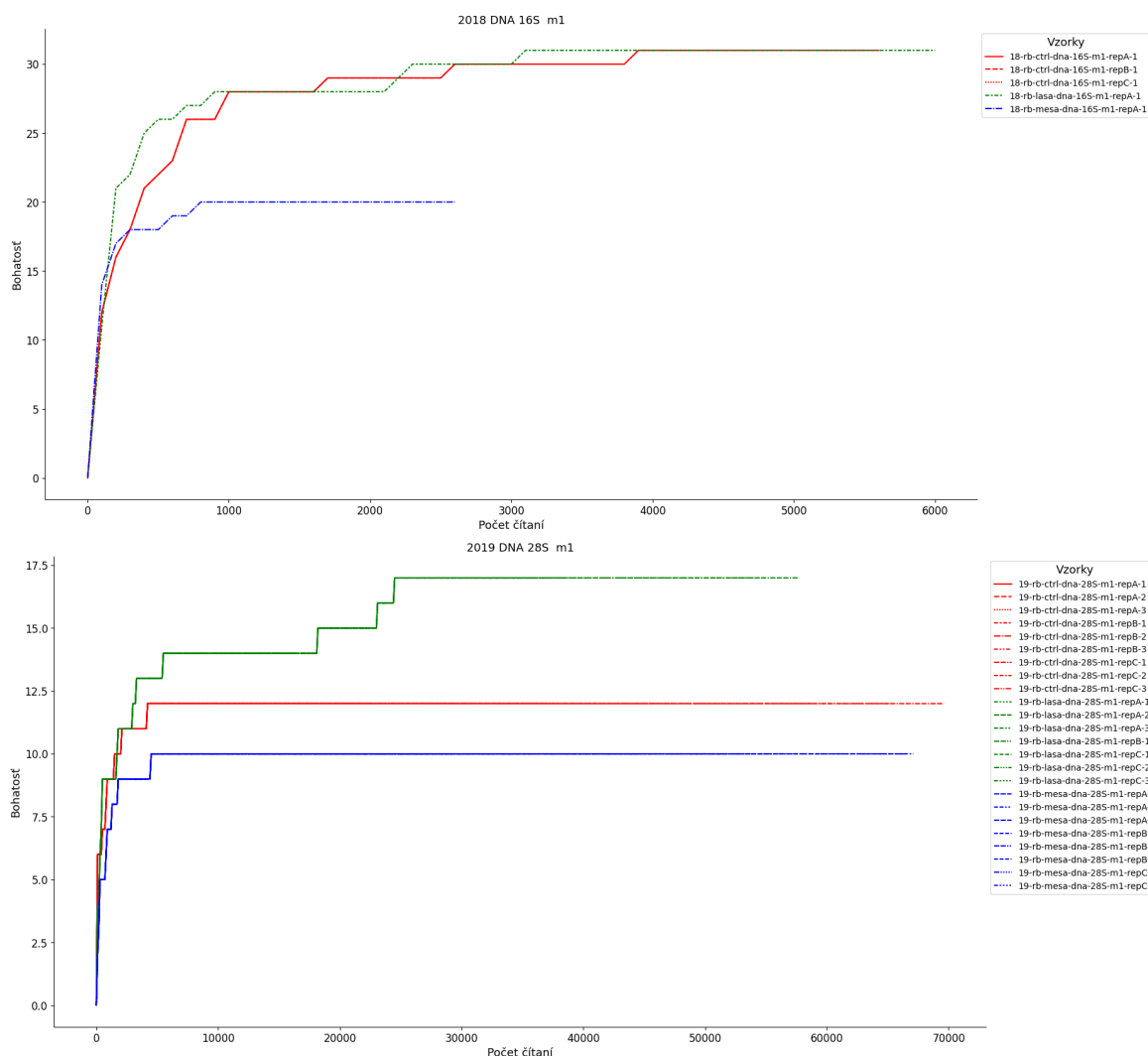
Pri porovnávaní experimentálnych štartérov vidíme štatisticky významné rozdiely len vo fáze m3. Baktérie *Caulobacter* a *Bradyrhizobium* mali väčšie zastúpenie vo vzorkách so štartérom lasa a baktérie *Corynebacterium* mali väčšie zastúpenie vo vzorkách so štartérom mesa.

## 3.3 Alfa diverzita

### 3.3.1 Rarefaction

Na základe vygenerovaných rarefaction kriviek usudzujeme, že jednotlivé vzorky boli dostatočne hlboko osekvenované. Nie je vylúčené, že počet detekovaných taxónov by mohol pri hlbšom sekvenovaní ďalej rásť. Je ale nepravdepodobné, že objavíme nejakú novú skupinu významne zastúpených baktérií alebo húb. Tieto výsledky môžeme vidieť na obrázku 3.11, kde sú však vybrané 2 grafy na ilustráciu. Všetky grafy z tejto

analýzy sa nachádzajú v elektronickej prílohe.



Obr. 3.11: Výsledok rarefaction grafov. V legende sa nachádzajú názvy jednotlivých vzoriek. Na osi x je počet čítaní a na osi y je nameraná bohatosť vzoriek. Nadpis popisuje typ vzoriek, ktoré boli použité ako vstupné dáta. V legende sa nachádzajú aj názvy konkrétnych vzoriek. Tieto 2 grafy sú vybrané ilustračne. Všetky grafy, ktoré boli výsledkom tejto analýzy sa nachádzajú v elektronickej prílohe.

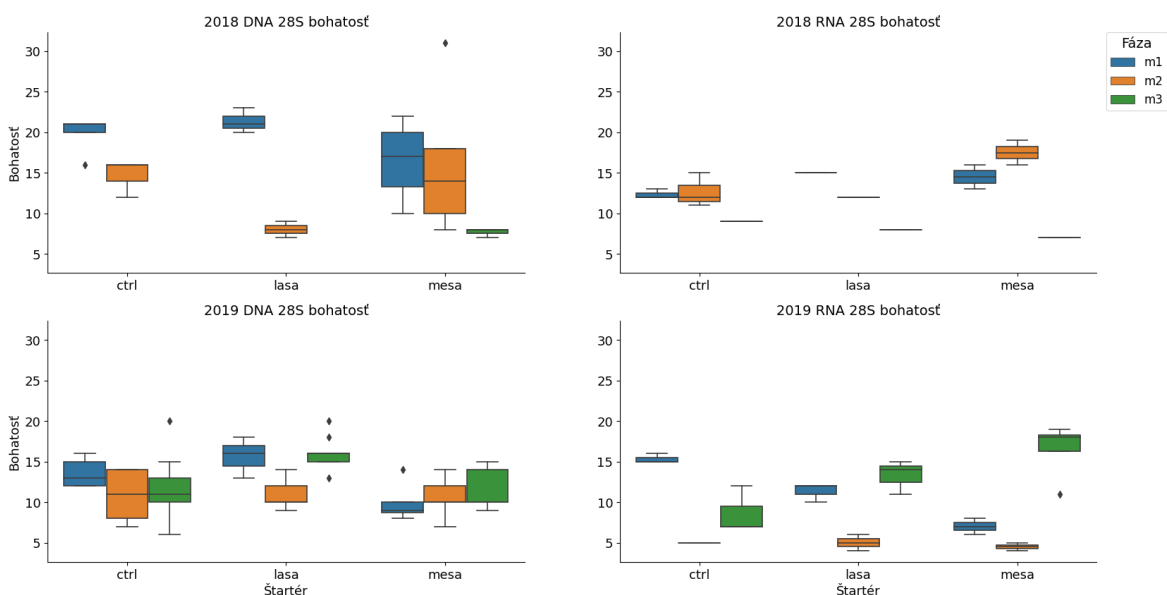
### 3.3.2 Porovnanie alfa diverzít

Čo sa týka bohatosti pre fungálne 28S dáta, môžeme vidieť na obrázku 3.12, že aj v roku 2018 aj 2019 bola najväčšia bohatosť vo fáze m1. V tejto fáze by vo vzorkách mali byť pozorovateľné takmer všetky kmene, ktoré sa nachádzali aj na hrozne. Následne vidíme prudký pokles vo fáze m2. V tejto fáze prebieha najbúrlivejšie kvasenie, pri ktorom je dominantná kvasinka *Saccharomyces*. Pri tejto fáze kvasenia aj zahynie veľa iných kvasiniek a húb, ktoré nevedia prežiť v prítomnosti alkoholu. V roku 2018 vo fáze



m3 toho nevieme veľa povedať o bohatosti populácií. Máme dáta len zo štartéru mesa, pre ostatné štartéry tieto dáta chýbajú. Ale pri štartéry mesa vidíme ďalšie znižovanie bohatosti populácie kvasiniek a húb. V roku 2019 je táto fáza zaujímavejšia. Pre kontrolný štartér bohatosť ostala zhruba rovnaká ako vo fáze m2. Avšak pri štartéroch mesa a lasa to vyzerá, že bohatosť populácie narástla oproti fáze m2. Je možné, že viac kvasiniek a húb prežilo za prítomnosti týchto štartérov ako pri kontrolnom štartéry. Keďže v tejto fáze už je dokončené búrlivé kvasenie, tak už nedominuje *Saccharomyces* a môžu sa prejaviť aj iné kvasinky a huby.

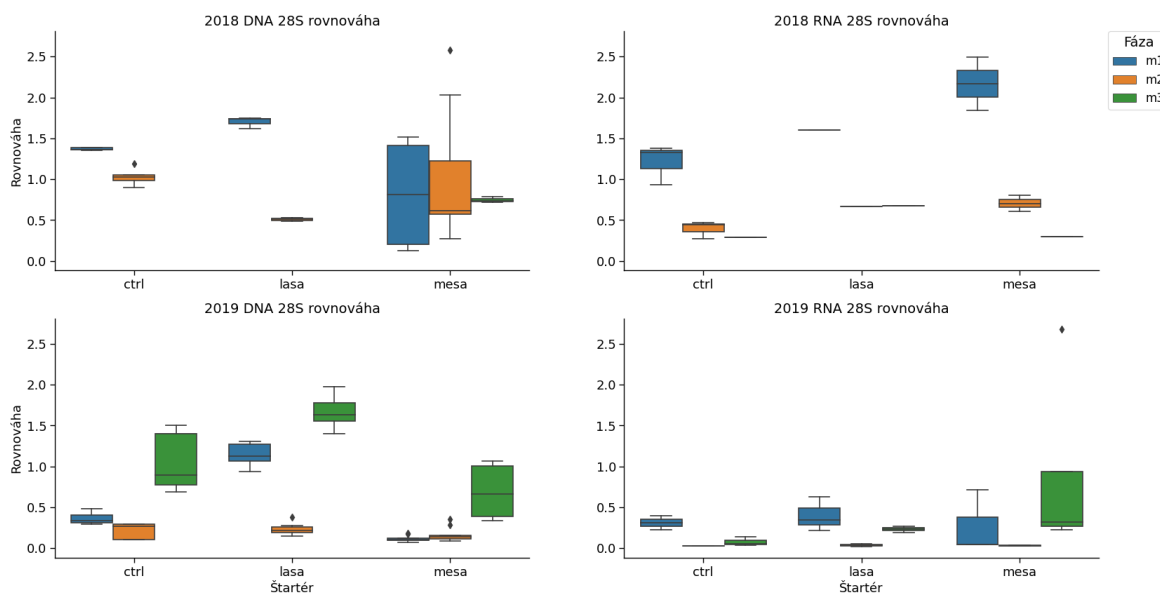
Pri RNA by sme očakávali, že bohatosť bude rovnaká alebo nižšia v jednotlivých fázach ako pri DNA. V roku 2018 to tak je, s výnimkou fázy m2 pri štartéry lasa, kde bola nameraná RNA aktivita vyššia ako prítomnosť DNA. V roku 2019 je tiež takáto výnimka, konkrétne fáza m3 pre štartér mesa. Vidíme, že najmenej bohaté fázy, čo sa týka RNA, v roku 2018 boli m3 pre všetky štartéry a v roku 2019 to bola fáza m2. Očakávali sme skôr fázu m2, že bude najmenej bohatá čo sa týka RNA, keďže túto fázu dominuje kvasinka *Saccharomyces*. Tieto RNA výsledky sú podobné ako im korešpondujúce DNA výsledky. Tieto výsledky vidíme na obrázku 3.12



Obr. 3.12: Výsledok porovnania bohatostí pre fungálne dáta. Na obrázku sú 4 grafy, každý patrí inému typu vzoriek. Typy vzoriek sú špecifikované v názvoch grafov. Na osi x sú štartéry. Pre každý štartér sú na grafoch 3 boxy, každý box pre 1 fázu kvasenia. Boxy sú vždy v poradí m1, m2, m3. Na osi y je vypočítaná hodnota bohatosti mikrobiómu vo vzorkách.

Keď sa pozrieme na obrázok 3.13 vidíme, že rovnováha týchto populácií, v roku 2018 je najvyrovnanejšia vo fáze m1 čo sa týka DNA aj RNA. Najmenej vyrovnaná je vo fáze m2. Toto platí pre všetky štartéry. Toto by sme očakávali, keďže vo fáze m1 by mala byť najväčšia diverzita a vo fáze m2 by mala dominovať hlavne kvasinka *Saccharomyces*.

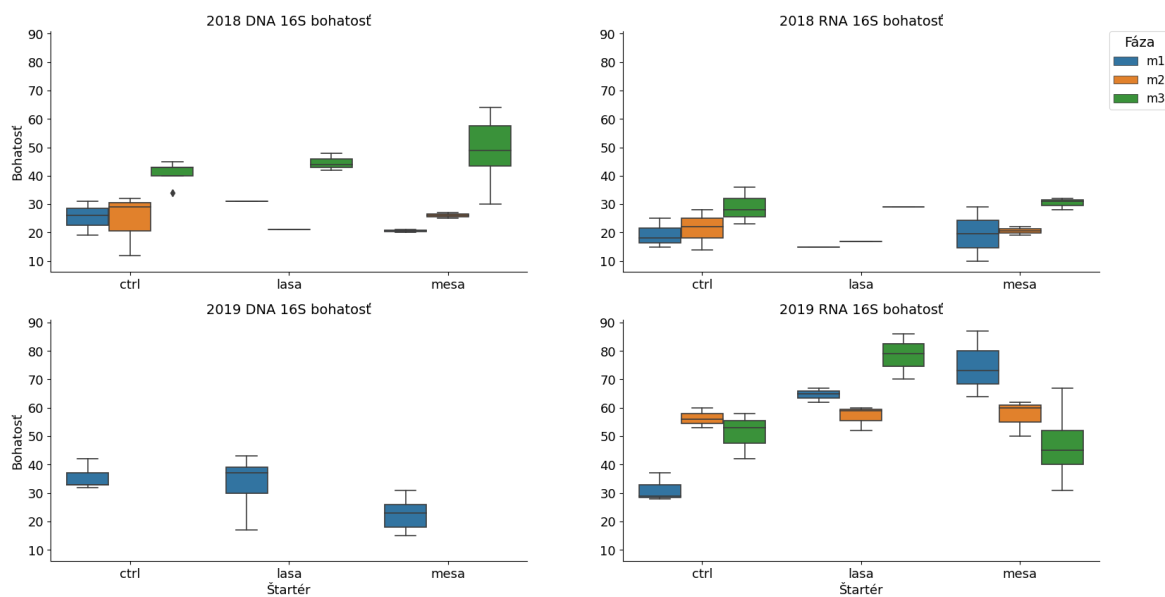
V roku 2019 je situácia trochu zaujímavejšia. Fázy m1 a m2 sú podobné ako v roku 2018, avšak vo fáze m3 je populácia omnoho vyrovnannejšia ako v ostatných fázach pri všetkých štartéroch v DNA dátach. Pri RNA aktivite tomu už tak nie je. Fáza m3 je pri kontrolnom štartéry podobne vyrovnaná ako m2, pri lasa štartéry trochu vyrovnannejšia a pri mesa štartéry je vyrovnannejšia aj ako fáza m1.



Obr. 3.13: Výsledok porovnania rovnováhy pre fungálne dáta. Na obrázku sú 4 grafy, každý patrí inému typu vzoriek. Typy vzoriek sú špecifikované v názvoch grafov. Na osi x sú štartéry. Pre každý štartér sú na grafoch 3 boxy, každý box pre 1 fázu kvasenia. Boxy sú vždy v poradí m1, m2, m3. Na osi y je vypočítaná hodnota rovnováhy mikrobiómu vo vzorkách.

Keď sa pozrieme na obrázok 3.14 na bakteriálne 16S dáta, meranie DNA v roku 2018 ukazuje, že najbohatšie boli vzorky vo fáze m3 vo všetkých štartéroch. Kontrolný štartér a mesa štartér majú rovnaký priebeh, kde bohatosť vzorky stúpa v každej fáze kvasenie. Pri štartéry lasa je najnižšia bohatosť vo fáze m2. Pri RNA aktivite v roku 2018 opäť vidíme tento trend, kde bohatosť stúpa s v každej fáze kvasenia, tentoraz vo všetkých štartéroch.

V roku 2019 nám chýbajú dáta o meraniach DNA vo fáze m2 a m3 pre všetky štartéry. Vo fáze m1 kontrolný štartér a lasa štartér sú podobne bohaté, mesa štartér je menej bohatý ako tieto dva. Pri pohľade na RNA aktivitu však nevidno nejaký trend medzi štartérmi. Pri kontrolnom štartéry v m1 fáza je najmenej bohatá, potom m2 je najbohatšia a m3 je o trochu menej bohatá ako m2. Pri lasa štartéry m1 je o trochu bohatšia ako m2 a m3 je oveľa bohatšia ako m2. Pri mesa štartéry je m1 najbohatšia a postupne to klesá v ďalších fázach.



Obr. 3.14: Výsledok porovnania bohatostí pre bakteriálne dáta. Na obrázku sú 4 grafy, každý patrí inému typu vzoriek. Typy vzoriek sú špecifikované v názvoch grafov. Na osi x sú štartéry. Pre každý štartér sú na grafoch 3 boxy, každý box pre 1 fázu kvasenia. Boxy sú vždy v poradí m1, m2, m3. Na osi y je vypočítaná hodnota bohatosti mikrobiómu vo vzorkách.

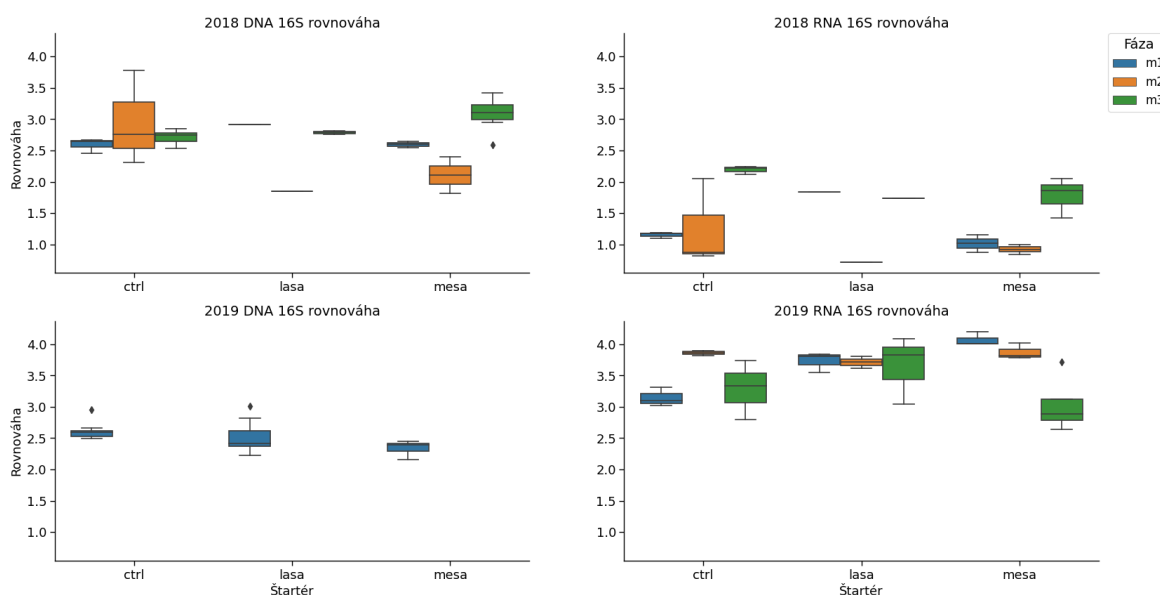
Keď sa pozrieme na obrázok 3.15 na rovnováhu týchto vzoriek, tak v roku 2018 vidíme, že pri meraní DNA je kontrolný štartér zhruba rovnako vyrovnaný vo všetkých fázach. Štartér lasa je rovnako vyrovnaný vo fázach m1 a m3, vo fáze m2 je menej vyrovnaný ako v týchto dvoch. Tento pokles rovnováhy je spôsobený tým, že baktérie z rodu *Komagataeibacter* začali prevažovať v populácii oproti ostatným baktériám. Následne vo fáze m3 jeho pomerové zastúpenie kleslo a populácia sa vyrovnala. Štartér mesa je vo fáze m1 vyrovnanejší ako vo fáze m2, následne vo fáze m3 je najvyrovnanejší. Spôsobené to je z rovnakého dôvodu ako pri štartéry lasa. Tieto zmeny môžeme sledovať na obrázku 3.5

Vyrovnanosť RNA aktivity v roku 2018 je na nižšej úrovni ako pri meraní DNA. Pri kontrolnom štartéry vo fáze m1 dominovali baktérie rodu *Escherichia-Shigella* a postupne jej dominancia klesla a začali byť aktívne aj iné baktérie. Toto vidíme aj na grafe (obrázok 3.6), kde rovnováha populácia sa postupne zvyšuje vo fázach. Pri štartéry lasa taktiež dominovali baktérie z rodu *Escherichia-Shigella*. Avšak tu to najviac dominovala vo fáze m2, pri fáze m1 a m3 to bolo na podobnej úrovni. Toto vidíme aj na grafe, kde rovnováha populácie je najnižšia vo fáze m2 a vo fázach m1 a m3 je na podobnej úrovni. Pri štartéry mesa je situácia rovnaká ako pri lasa štartéry.

V roku 2019 chýbajú dáta pre DNA vzorky vo fázach m2 a m3 vo všetkých štartéroch. Ale fáza m1 je zhruba rovnako vyrovnaná vo všetkých štartéroch.

RNA aktivita v roku 2019 sa pohybuje okolo rovnakej úrovne vyrovnanosti vo všet-

kých štartéroch vo všetkých fázach. Pri kontrolnom štartéry vidíme nárast rovnováhy vo fáze m2. Toto korešponduje s tým, že vo vzorkách narástlo zastúpenie baktérií z rodov *Ralstonia*, *Pluralibacter* a čeľade *Sphingomonadaceae* a iných. Tieto populácie sa vo fáze m3 vrátili do úrovne, v ktorej boli vo fáze m1. Pri štartéry lasa je RNA aktivita zhruba rovnako vyrovnaná vo všetkých fázach. Pri štartéry mesa vidíme, že rovnováha vzorky postupne klesala. Vo fáze m1 sa vo vzorkách nachádzalo mnoho rôznych baktérií, ale s pomerne malým zastúpením. Vo fáze m2 zanikli niektoré tieto malé populácie. Nakoniec vo fáze m3 vidíme prudký nárast populácií baktérií z rodu *Pseudomonas*, *Escherichia-Shigella* a z čeľade *Sphingomonadaceae*. Toto môže za pokles rovnováhy vzoriek vo fáze m3. Tieto zmeny vidno na grafoch na obrázku 3.8.



Obr. 3.15: Výsledok porovnania rovnováhy pre bakteriálne dáta. Na obrázku sú 4 grafy, každý patrí inému typu vzoriek. Typy vzoriek sú špecifikované v názvoch grafov. Na osi x sú štartéry. Pre každý štartér sú na grafoch 3 boxy, každý box pre 1 fázu kvasenia. Boxy sú vždy v poradí m1, m2, m3. Na osi y je vypočítaná hodnota bohatosti mikrobiómu vo vzorkách.

## 3.4 Beta diverzita

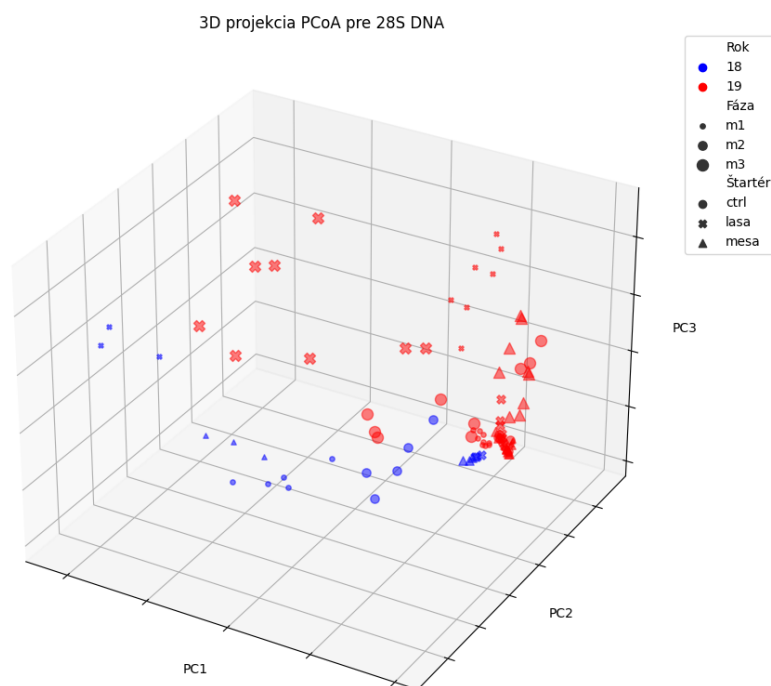
### 3.4.1 Analýza hlavných súradníc pre fungálne 28S dáta

PCoA pre DNA vzorky (obrázky 3.16 a 3.17) nám ukazuje, že roky 2018 a 2019 sú dostatočne odlišné. Graf s PC1 oproti PC2 nám toto ilustruje najlepšie. Táto odlišnosť v rokoch sa dá vysvetliť tým, že v roku 2018 a 2019 bolo veľmi odlišné počasie v období kedy rastie a dozrieva hrozno. Údaje zo SHMÚ ukazujú, že v roku 2018 bolo v tomto

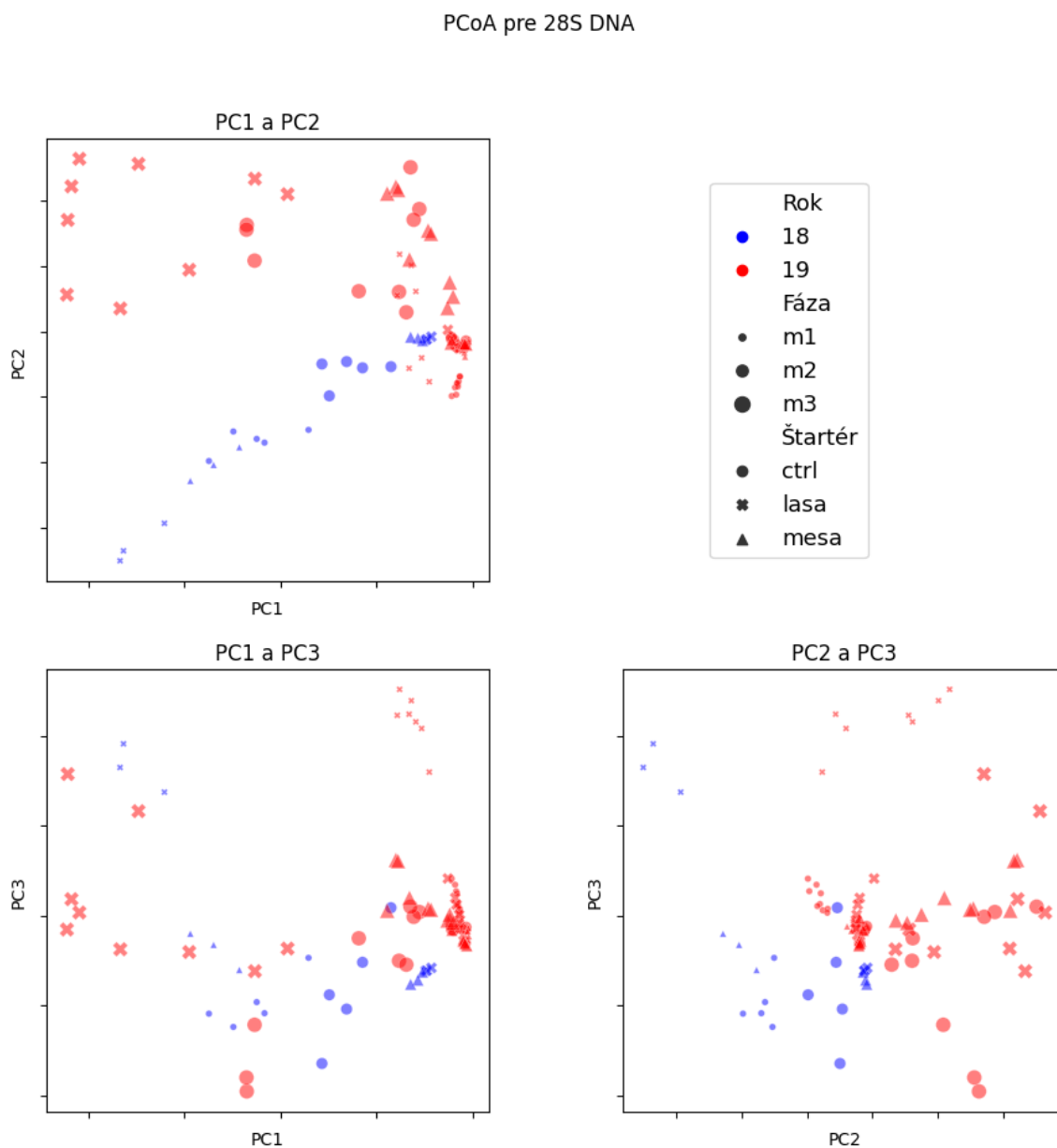
období omnoho viacej zrážok ako v roku 2019 [44][45]. Taktiež v roku 2019 priemerná teplota dosahovala vyššie hodnoty častejšie ako tomu bolo v roku 2018.

V grafe s PC2 oproti PC3 by sa dal rozdeliť na oblasti podľa fázy. Keď si predstavíme diagonálu z ľavého dolného rohu k pravému hornému rohu, tak vzorky z m1 fázy sú väčšinou nad diagonálou. Vzorky z fázy m2 sa nachádzajú hlavne v centre grafu. Vzorky fázy m3 sú všetky na pravej strane grafu. Avšak z fázy m3 máme len vzorky z roku 2019. Vidíme aj, že vzorky obsahujúce štartér lasa sa vo fáze m1 sa na grafoch držia ďalej od väčšiny vzoriek. Taktiež vzorky štartéru lasa vo fáze m3 z roku 2019 na grafe s PC1 oproti PC2 tvoria vlastnú skupinu. Toto nám ukazuje, že štartér lasa má vplyv na kvasinkové komunity vo fázach m1 a m3.

Vzorky fázy m2 z rokov 2018 aj 2019 sa zhlukujú blízko seba. Taktiež vzorky obsahujúce štartér mesa a vzorky obsahujúce kontrolný štartér sa vo všetkých fázach držia blízko seba. Toto nám podporuje našu hypotézu, že inokulácia štartérom mesa nebola úspešná ani v jednom roku.

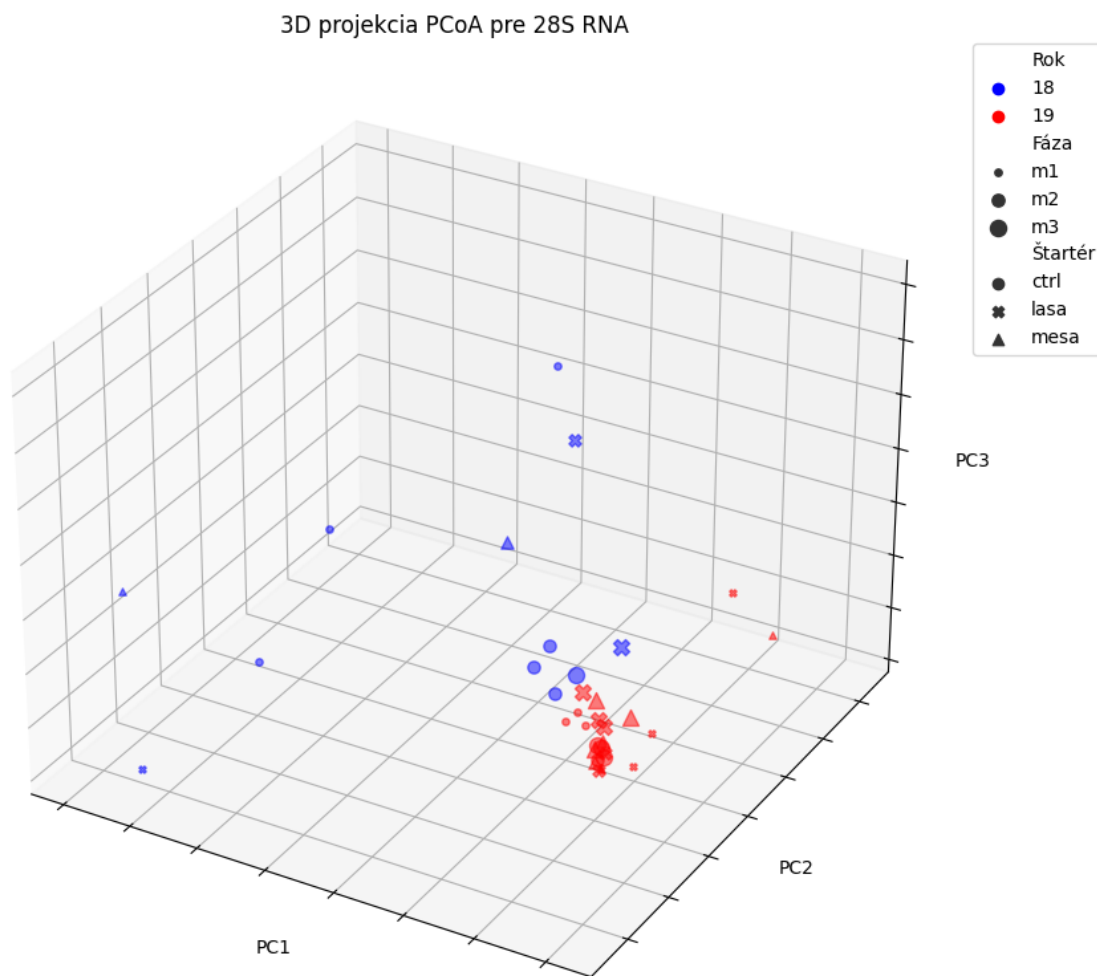


Obr. 3.16: 3D projekcia beta diverzity na základe PCoA pre fungálne DNA dáta. Na osiach sú jednotlivé komponenty PCoA. Body sú označené podľa roku, fázy kvasenia a štartéru. Body z roku 2018 sú označené modrou a z roku 2019 označené červenou. Fázy sú znázornené rozdielnymi veľkosťami bodov. Štartéry sú znázornené podľa rôzneho tvaru bodov.

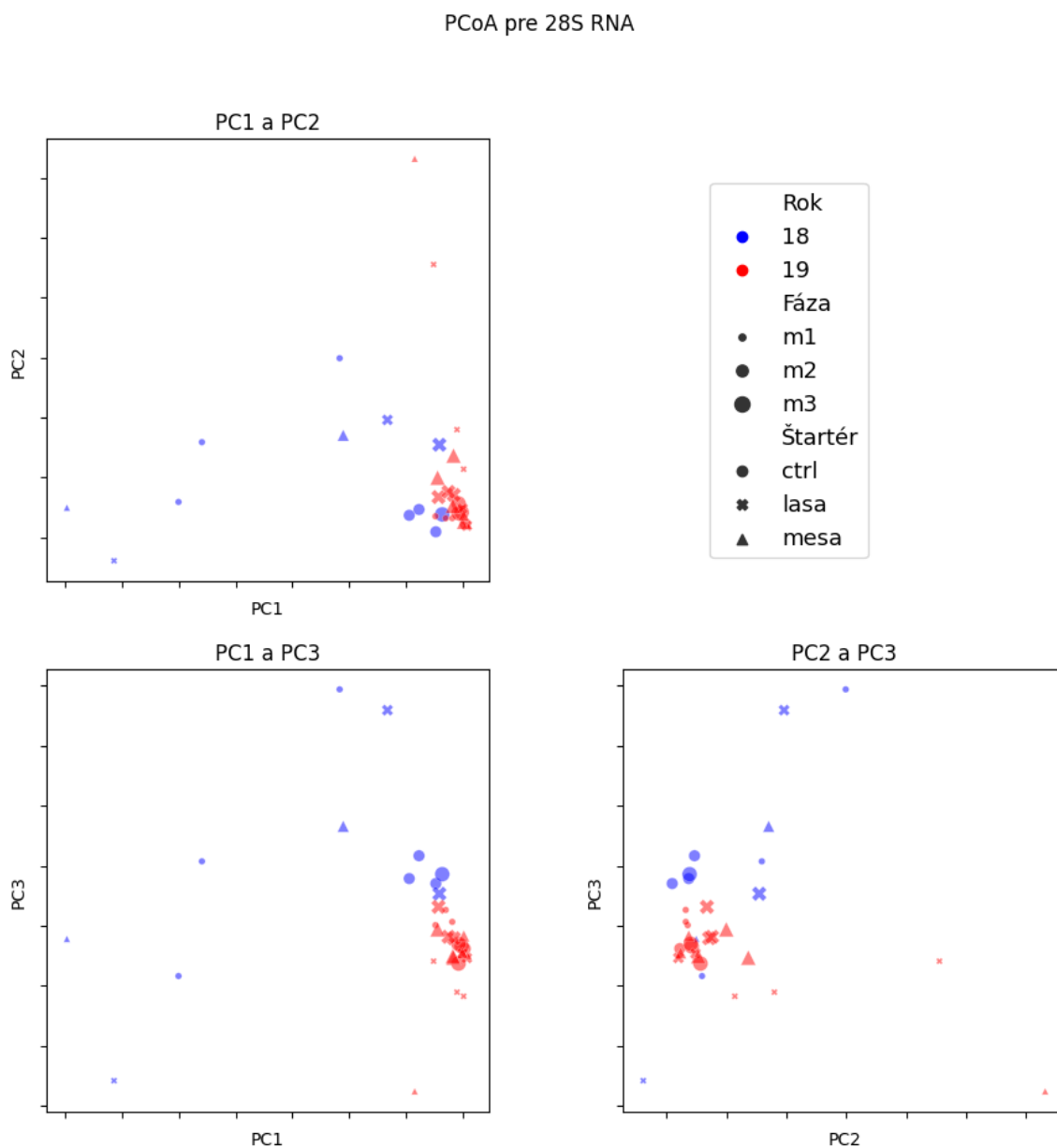


Obr. 3.17: PCoA pre fungálne DNA dáta po jednotlivých komponentoch. Na osiach sú jednotlivé komponenty PCoA. Body sú označené podľa roku, fázy kvasenia a štartéru. Body z roku 2018 sú označené modrou a z roku 2019 označené červenou. Fázy sú znázornené rozdielnymi veľkosťami bodov. Štartéry sú znázornené podľa rôzneho tvaru bodov.

PCoA pre vzorky RNA (obrázky 3.18 a 3.19) opäť ukazujú, že medzi rokmi boli rozdiely, hlavne vo fáze m1. Vo nasledujúcich fázach sa vzorky z oboch rokov zhlukujú veľmi blízko seba. Pre štartér lasa sa v m1 fáze dajú najst vzorky, ktoré sú oddialené od ostatných v tejto fáze, ale nie je to také jednoznačné ako sme videli pri DNA. Hlavne vzorky z 2019 sa držia blízko hlavného zhluku vzoriek. Toto sme aj očakávali, keďže ako sme už videli, tak RNA aktivita bola dominovaná prevažne kvasinkami rodu *Saccharomyces*.



Obr. 3.18: 3D projekcia beta diverzity na základe PCoA pre fungálne RNA dáta. Na osiach sú jednotlivé komponenty PCoA. Body sú označené podľa roku, fázy kvasenia a štartéru. Body z roku 2018 sú označené modrou a z roku 2019 označené červenou. Fázy sú znázornené rozdielnymi veľkosťami bodov. Štartéry sú znázornené podľa rôzneho tvaru bodov.



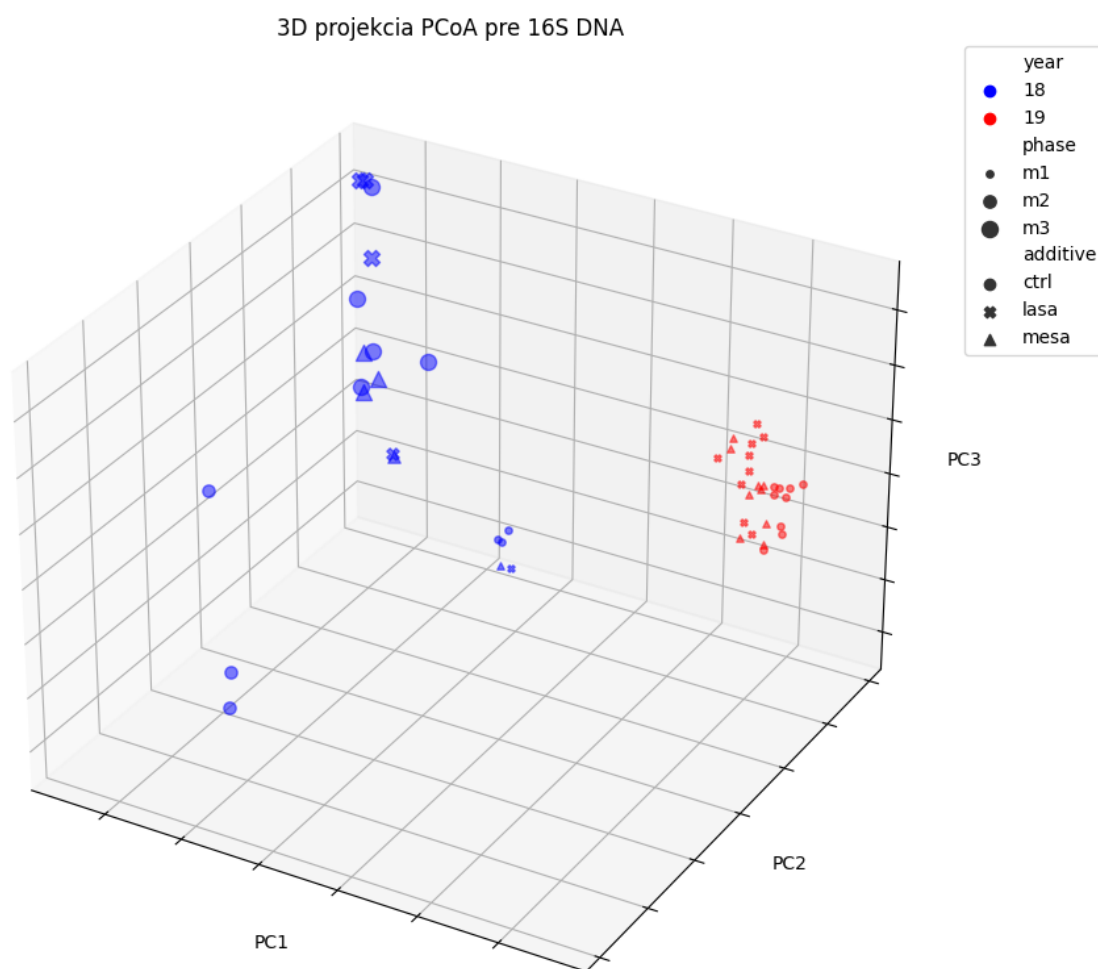
Obr. 3.19: PCoA pre fungálne RNA dáta po jednotlivých komponentoch. Na osiach sú jednotlivé komponenty PCoA. Body sú označené podľa roku, fázy kvasenia a štartéru. Body z roku 2018 sú označené modrou a z roku 2019 označené červenou. Fázy sú znázornené rozdielnymi veľkosťami bodov. Štartéry sú znázornené podľa rôzneho tvaru bodov.



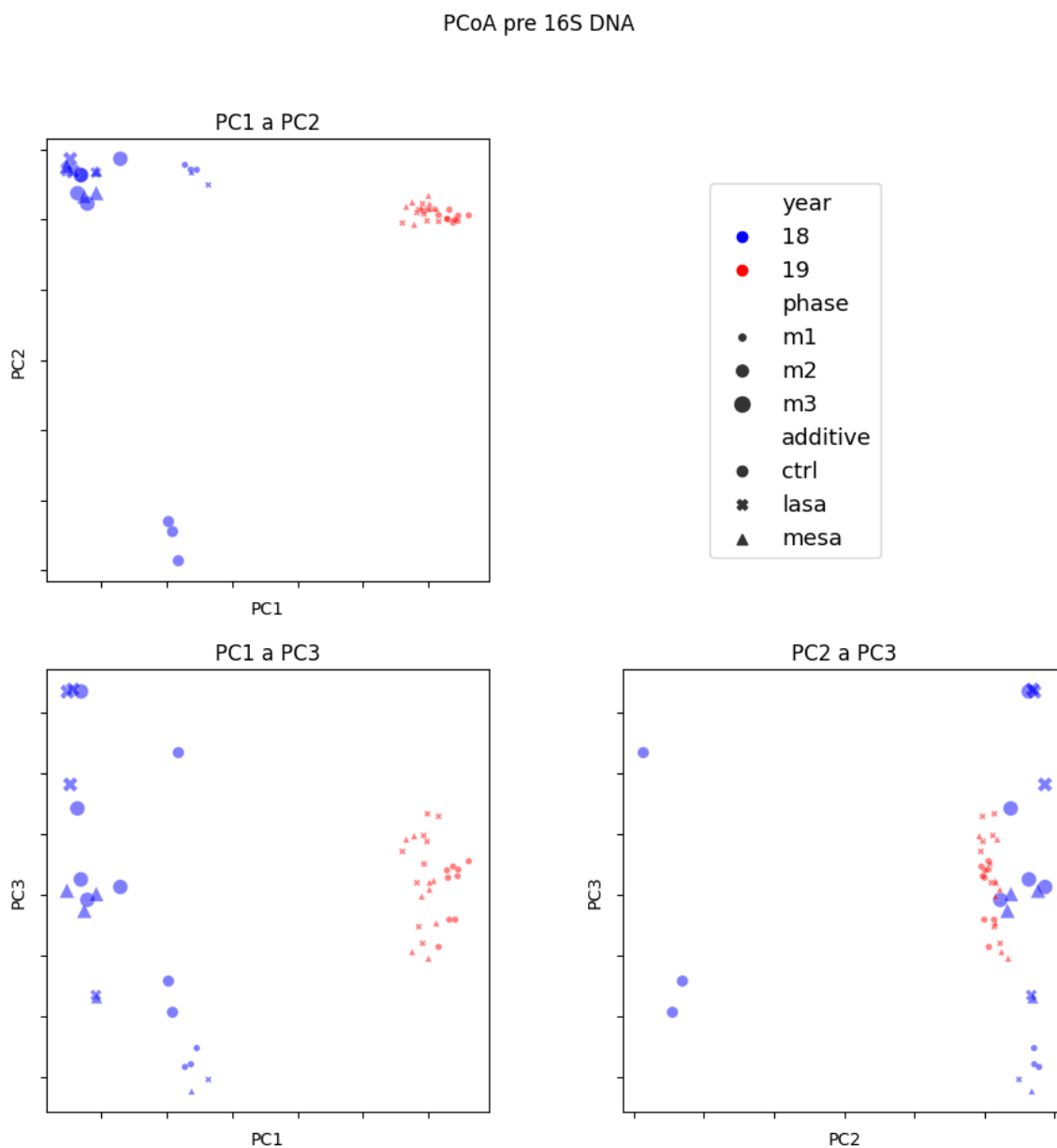
### 3.4.2 Analýza hlavných súradníc pre bakteriálne 16S dáta

Na grafoch pre DNA vzorky (obrázky 3.20 a 3.21) vieme jasne odlíšiť roky 2018 a 2019. V týchto grafoch jasne vidieť odlišnosť bakteriálnych komunit v rokoch 2018 a 2019. Domievame sa, že bakteriálne komunity boli najviac zasiahnuté odlišným počasím v rokoch 2018 a 2019.

V roku 2019 máme len dáta pre fázu m1, ale nevyzerá to, že beta diverzita by bola ovplyvnená rôznym štartérom, keďže netvorí nejaké zhluky na základe štartéru. V roku 2018 vieme na grafe s PC1 oproti PC2 odlíšiť jednotlivé fázy, ale vplyv štartérov nie je vidieť.

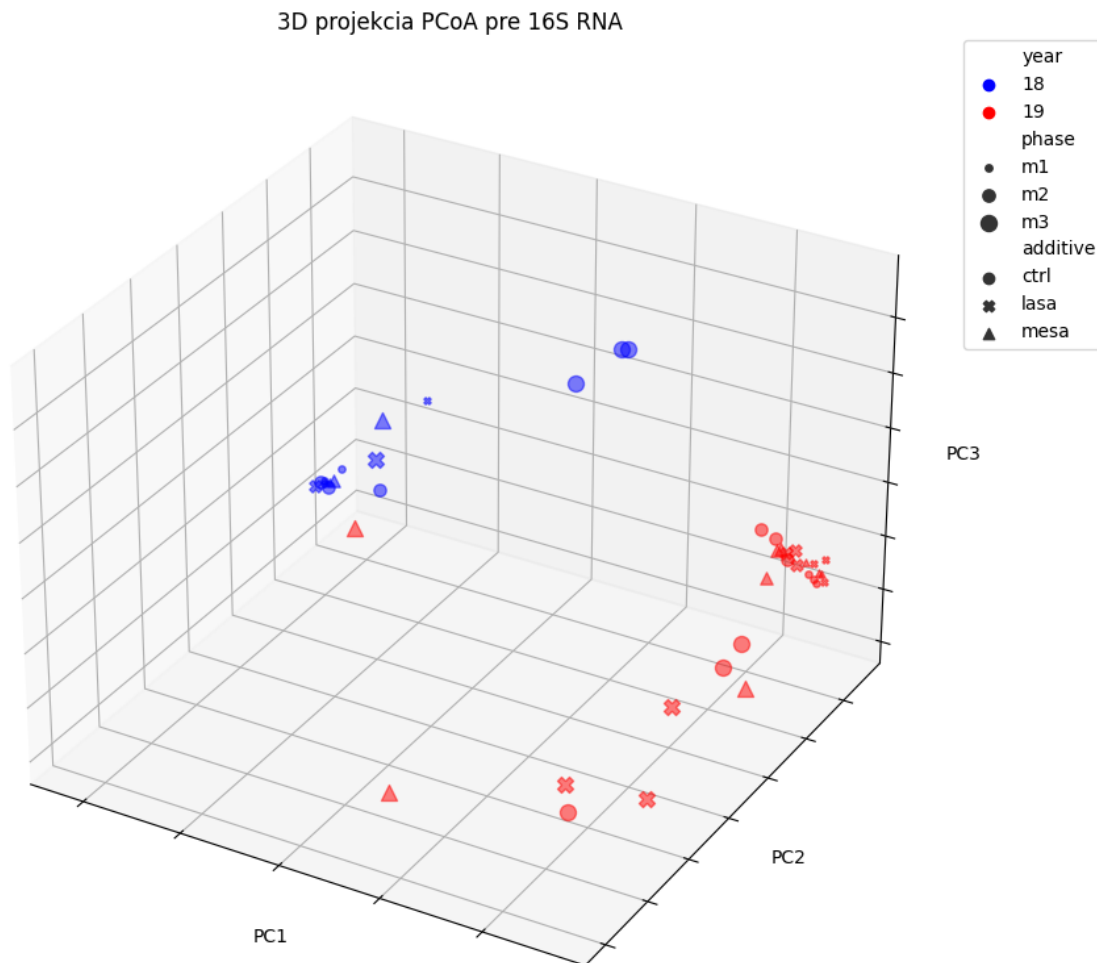


Obr. 3.20: 3D projekcia beta diverzity na základe PCoA pre bakteriálne DNA dáta. Na osiach sú jednotlivé komponenty PCoA. Body sú označené podľa roku, fázy kvasenia a štartéru. Body z roku 2018 sú označené modrou a z roku 2019 označené červenou. Fázy sú znázornené rozdielnymi veľkosťami bodov. Štartéry sú znázornené podľa rôzneho tvaru bodov.

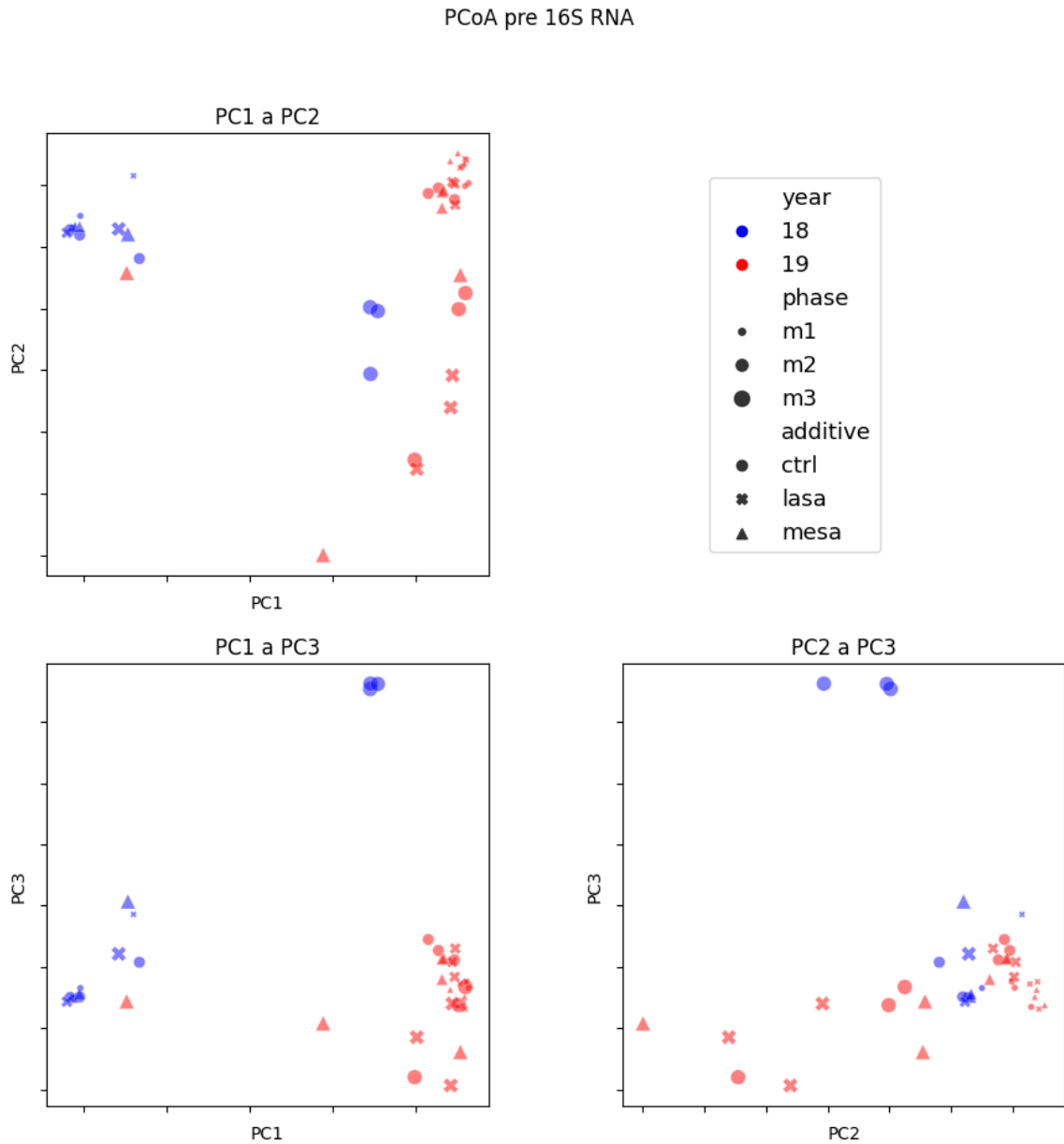


Obr. 3.21: PCoA pre bakteriálne DNA dáta po jednotlivých komponentoch. Na osiach sú jednotlivé komponenty PCoA. Body sú označené podľa roku, fázy kvasenia a štartéru. Body z roku 2018 sú označené modrou a z roku 2019 označené červenou. Fázy sú znázornené rozdielnymi veľkosťami bodov. Štartéry sú znázornené podľa rôzneho tvaru bodov.

Pri grafoch pre RNA vzorky (obrázky 3.22 a 3.23) opäť vidíme, že vzorky sú rozličné v závislosti od rokov. V oboch rokoch si môžeme všimnúť, že fázy m1 a m2 sú v oboch rokoch zhlukované dokopy a vzorky fázy m3 sa od tohto zhluku vzdiaľujú. Opäť nevidíme nejaký výrazný rozdiel na základe štartéru. Je tu však jedna zaujímavá skupina vzoriek, a to sú vzorky z roku 2018 vo fáze m3 pre kontrolný štartér, ktoré sú vzdialené od všetkých ostatných vzoriek. V týchto vzorkách oproti iným z roku 2018 výrazne klesla aktivita baktérie *Escherichia-Shigella*, zatiaľ čo v ostatných bola dominantná baktéria. Toto vysvetľuje, prečo sú práve tieto vzorky tak vzdialené od ostatných.



Obr. 3.22: 3D projekcia beta diverzity na základe PCoA pre bakteriálne RNA dáta. Na osiach sú jednotlivé komponenty PCoA. Body sú označené podľa roku, fázy kvasenia a štartéru. Body z roku 2018 sú označené modrou a z roku 2019 označené červenou. Fázy sú znázornené rozdielnymi veľkosťami bodov. Štartéry sú znázornené podľa rôzneho tvaru bodov.



Obr. 3.23: PCoA pre bakteriálne RNA dáta po jednotlivých komponentoch. Na osiach sú jednotlivé komponenty PCoA. Body sú označené podľa roku, fázy kvasenia a štartéru. Body z roku 2018 sú označené modrou a z roku 2019 označené červenou. Fázy sú znázornené rozdielnymi veľkosťami bodov. Štartéry sú znázornené podľa rôzneho tvaru bodov.



# Kapitola 4

## Diskusia

Analýza jednotlivých taxónov a alfa diverzity pre fungálne 28S dáta ukázala, že najväčšia diverzita tejto populácie je vo fáze m1. V tejto fáze kvasenia sa v mušte nachádzajú mnohé nesacharomycétne kvasinky, ktoré zohrávajú dôležitú úlohu v týchto skorých fázach kvasenia. Produkujú enzýmy a aromatické látky, ktoré dodávajú vínu špecifický charakter [10]. Následne táto diverzita klesá vo fáze m2, kedy sa zvyšuje koncentrácia alkoholu vo víne a dominujú kvasinky rodu *Saccharomyces*. Toto sa ukázalo aj v iných štúdiách [30][26][10]. Vo fáze m3 môžeme vidieť opätovný nárast kvasiniek *Hanseniaspora*. Takýto nárast v neskorších fázach kvasenia bol zaznamenaný aj v iných štúdiách [13][10]. Na úrovni RNA môžeme v roku 2018 vidieť vo fáze m1 pomerne veľkú aktivitu plesne rodu *Botrytis*, ktorú nie je vidno v roku 2019. Toto môže byť spôsobené tým, že v roku 2018 bolo počasie omnoho vlhkejšie ako v roku 2019, čo podporuje rast tejto plesne [19][39]. V roku 2018 v mesiacoch dozrievania a zberu hrozna bol úhrn zrážok v priemere okolo 30-40 mm zrážok [44]. V roku 2019 bol tento priemer zhruba okolo 10mm zrážok [45]. Vo fázach m2 a m3 sa RNA aktivita na základe beta diverzity nelíši medzi rokmi. Toto je spôsobené už spomínanou dominanciou kvasiniek rodu *Saccharomyces*.

Pri pohľade na vplyv jednotlivých štartérov na fungálne populácie sme si všimli, že štartér lasa má pri DNA bohatšie a vyrovnanejšie populácie vo fázach m1 a m3 ako kontrolný štartér. Je teda možné, že vyššia prítomnosť kvasiniek rodu *Lachancea* môže za tento nárast. Domnievame sa, že to môže byť spôsobené kompetíciou medzi rodmi *Lachancea* a *Saccharomyces* [16]. Taktiež jedna štúdia ukázala, že za dostatočnej prítomnosti kyslíka má kvasinka *Lachancea thermotolerans* potenciál aj dominovať fermentáciu proti *Saccharomyces* [36]. Vidíme aj pri analýze beta diverzity, že fungálne populácie vo fázach m1 a m3 pre štartér lasa tvoria samostatné klastre.

Pri analýze prežívania štartéru mesa sa ukázalo, že pravdepodobne sa nepodarila inokulácia vínneho muštu týmto štartérom, keďže sme nezaznamenali prítomnosť DNA kvasiniek *Metschnikowia* ani ich RNA aktivitu v žiadnej zo vzoriek. Tým pádom

vzorky týchto mesa štartéru sú veľmi podobné so vzorkami kontrolného štartéru v oboch rokoch. Toto sme videli aj analýze beta diverzity, kde vzorky kontrolného a mesa štartéru sa zoskupujú spolu. Nevieme teda usúdiť, aký vplyv má štartér mesa na celkové zloženie mikrobiómu.

Bakteriálne 16S vzorky DNA máme dáta len pre fázu m1. V tejto fáze medzi rokmi na prvý pohľad vyzerajú veľmi podobne. Avšak analýza beta diverzity ukázala, že tieto vzorky sa veľmi líšia medzi rokmi. V jednotlivých rokoch však nevidíme rozdiely medzi vzorkami s kontrolným štartérom a vzorkami s lasa štartérom. Teda môžeme posúdiť, že štartér lasa nemá vplyv na zloženie bakteriálnej populácie. Keďže inokulácia štartérom mesa sa nepodarila, tak neviem usúdiť aký vplyv by tento štartér mal na bakteriálne komunity.

Vzorky RNA sa taktiež veľmi líšili medzi rokmi. V roku 2018 väčšinu fáz v každom štartéry dominovali baktérie rodu *Escherichia-Shigella*. Výnimkou bola fáza m3 v kontrolnom štartéry. Pravdepodobne sa ale jedná len o technický artefakt, keďže vzorky so štartérom mesa mali túto dominanciu vo všetkých fázach. V roku 2019 táto dominancia nebola. RNA aktivita *Escherichia-Shigella* v jednotlivých fázach a štartéroch je na veľmi nízkej úrovni. Tento rozdiel sa opäť dá vysvetliť veľmi rozdielnym počasím v roku 2018 a 2019, kde v roku 2018 bolo prostredie vlhkejšie. *Escherichia-Shigella* lepšie prežíva vo vlhkom prostredí ako v suchom [7].

Celkovo sa bakteriálne populácia v oboch rokoch skladali hlavne z proteobaktérií (*Acetobacter*, *Gluconobacter*, *Komagataeibacter*, *Pseudomonas*). Tieto baktérie sa označujú aj skratkou AAB (acetic acid producing bacteria), pretože premieňajú etanol na kyselinu octovú[38]. Pri produkcii bieleho vína to v malom množstve nie je problém, pretože biele vína sa obecné vyžadujú kyslejšie. Vo veľkom množstve sú však nežiadúce, pretože kyselina octová môže pokaziť kvalitu vína [38]. Taktiež sme pozorovali prítomnosť zo skupiny LAB (lactic acid producing bacteria), ktoré sú zodpovedné za mliečne kvasenie. Pri mliečnom kvasení dochádza ku konverzii kyseliny jablčnej na kyseliny mliečnu a pri tomto procese vznikajú aj aromatické látky [35]. Najčastejšie sa vo vínach z tejto skupiny nachádza baktéria *Oenococcus oeni* [40]. Vo vzorkách z dátového zdroja sa však nenachádza. Z tejto skupiny baktérií sa vo vzorkách nachádzajú hlavne baktérie z rodov *Lactobacillus*, *Lactococcus* a *Streptococcus*. Ďalšou väčšou skupinou baktérií boli baktérie z čeľade *Bacillaceae*. Tieto baktérie dokážu kvasiť, teda môžu byť prospešné pri procese kvasenia [37].

# Záver

V tejto práci sme sa pozreli na zloženie vínného mikrobiómu. Pri fungálnych vzorkách sme videli dominanciu kvasiniek rodu *Saccharomyces* čo sa týka prítomnosti DNA aj RNA aktivity. Pridanie kvasiniek *Lachancea thermotolerans* v podobe štartéru malo mierne pozitívny efekt na alfa diverzitu fungálnych vzoriek. Avšak malo to veľmi malý alebo žiadny efekt na bakteriálne vzorky. Toto sa premietlo aj do beta diverzity, kde pri fungálnych vzorkách sa dali odlíšiť klastre štartéru lasa od klastru kontrolného a mesa štartéru. Pri bakteriálnych vzorkách boli všetky štartéry v spoločných zhlukoch.

Bakteriálne vzorky vyzerali byť ovplyvnené len rokom ich pôvodu. Veľmi rozdielne počasie v rokoch 2018 a 2019, kedy boli vzorky odoberané malo najväčší vplyv na zloženie bakteriálnych komunit. Videli sme, že vďaka vlhkému počasiu v roku 2018 baktérie rodu *Escherichia-Shigella* boli schopné dominovať RNA aktivitu vo všetkých fázach. V roku 2019 za suchšieho počasia bola RNA aktivita viac vyrovnaná. Počasie malo vplyv aj na fungálne vzorky kde sme v roku 2018 videli značnú RNA aktivitu plesne rodu *Botrytis*, pravdepodobne vďaka vlhkému počasiu. V roku 2019 sme takúto RNA aktivitu nevideli. V tomto roku takmer všetká nameraná RNA aktivita patrila rodu *Saccharomyces*.





# Literatúra

- [1] Inoculation. [Citované 2023-05-5] Dostupné z <https://www.biologyonline.com/dictionary/inoculation>.
- [2] Ján Peťka a Pavel Farkaš. Prchavé látky vo víne – zloženie a ich vplyv na arómu i. primárne aromatické látky, 2001.
- [3] Hany Abdo, Claudia Rita Catacchio, Mario Ventura, Pietro D’Addabbo, Hervé Alexandre, Michèle Guilloux-Bénatier, and Sandrine Rousseaux. The establishment of a fungal consortium in a new winery. *Scientific Reports*, 10(1):7962, May 2020.
- [4] Marti J. Anderson, Thomas O. Crist, Jonathan M. Chase, Mark Vellend, Brian D. Inouye, Amy L. Freestone, Nathan J. Sanders, Howard V. Cornell, Liza S. Comita, Kendi F. Davies, Susan P. Harrison, Nathan J. B. Kraft, James C. Stegen, and Nathan G. Swenson. Navigating the multiple meanings of  $\beta$  diversity: a roadmap for the practicing ecologist. *Ecology Letters*, 14(1):19–28, 2011.
- [5] Edward W. Beals. Bray-curtis ordination: An effective strategy for analysis of multivariate ecological data. volume 14 of *Advances in Ecological Research*, pages 1–55. Academic Press, 1984.
- [6] Sam Behjati and Patrick S Tarpey. What is next generation sequencing? *Archives of Disease in Childhood - Education and Practice*, 98(6):236–238, 2013.
- [7] Alexandra M Belias, Adrian Sbodio, Pilar Truchado, Daniel Weller, Janneth Pinzon, Mariya Skots, Ana Allende, Daniel Munther, Trevor Suslow, Martin Wiedmann, and Renata Ivanek. Effect of weather on the Die-Off of escherichia coli and attenuated salmonella enterica serovar typhimurium on preharvest leafy greens following irrigation with contaminated water. *Appl Environ Microbiol*, 86(17), August 2020.
- [8] Nicholas A. Bokulich, Thomas S. Collins, Chad Masarweh, Greg Allen, Hildegard Heymann, Susan E. Ebeler, and David A. Mills. Associations among wine grape

microbiome, metabolome, and fermentation behavior suggest microbial contribution to regional wine characteristics. *mBio*, 7(3):e00631–16, 2016.

- [9] Roberta Bussamara, Alexandre Meneghello Fuentefria, Eder Silva de Oliveira, Leonardo Broetto, Michaela Simcikova, Patrícia Valente, Augusto Schrank, and Marilene Henning Vainstein. Isolation of a lipase-secreting yeast for enzyme production in a pilot-plant scale batch fermentation. *Bioresource Technology*, 101(1):268–275, 2010.
- [10] Miroslav Böhmer. *Modelové aplikácie charakterizácie mikrobiálneho zloženia pomocou cieleného amplikónového masívne paralelného sekvenovania vo vinárskom výskume*. Dizertačná práca, Univerzita komenského v Bratislava, Prírodovedecká fakulta, 2021.
- [11] Benjamin J Callahan, Paul J McMurdie, and Susan P Holmes. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis, Jul 2017. [Citované 2023-01-26] Dostupné z <https://www.nature.com/articles/ismej2017119>.
- [12] M. Ciani, F. Comitini, and I. Mannazzu. Fermentation. In Sven Erik Jørgensen and Brian D. Fath, editors, *Encyclopedia of Ecology*, pages 1548–1557. Academic Press, Oxford, 2008.
- [13] Vasileios Englezos, Kalliopi Rantsiou, Fabrizio Torchio, Luca Rolle, Vincenzo Gerbi, and Luca Cocolin. Exploitation of the non-saccharomyces yeast *starmerella bacillaris* (synonym *candida zemplinina*) in wine fermentation: physiological and molecular characterizations. *Int J Food Microbiol*, 199:33–40, January 2015.
- [14] DAVID M. GADOURY, LANCE CADLE-DAVIDSON, WAYNE F. WILCOX, IAN B. DRY, ROBERT C. SEEM, and MICHAEL G. MILGROOM. Grapevine powdery mildew (*erysiphe necator*): a fascinating system for the study of the biology, ecology and epidemiology of an obligate biotroph. *Molecular Plant Pathology*, 13(1):1–16, 2012.
- [15] John J. Gart, Andrew F. Siegel, and Rebecca Z. German. Rarefaction and taxonomic diversity. *Biometrics*, 38(1):235–241, 1982.
- [16] Mirko Gobbi, Francesca Comitini, Paola Domizio, Cristina Romani, Livio Lencioni, Ilaria Mannazzu, and Maurizio Ciani. *Lachancea thermotolerans* and *saccharomyces cerevisiae* in simultaneous and sequential co-fermentation: A strategy to enhance acidity and improve the overall quality of wine. *Food Microbiology*, 33(2):271–281, 2013.

- [17] Elliot Nicholas Grady, Jacqueline MacDonald, Linda Liu, Alex Richman, and Ze-Chun Yuan. Current knowledge and perspectives of paenibacillus: a review. *Microbial Cell Factories*, 15(1):203, Dec 2016.
- [18] J. D. Hunter. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, 9(3):90–95, 2007.
- [19] Anton M. Kofranek. 1 - cut chrysanthemums. In Roy A. Larson, editor, *Introduction to Floriculture (Second Edition)*, pages 3–42. Academic Press, New York, second edition edition, 1992.
- [20] Waldan K. Kwong and Nancy A. Moran. Cultivation and characterization of the gut symbionts of honey bees and bumble bees: description of *snodgrassella alvi* gen. nov., sp. nov., a member of the family neisseriaceae of the betaproteobacteria, and *gilliamella apicola* gen. nov., sp. nov., a member of orbaceae fam. nov., orbales ord. nov., a sister taxon to the order ‘enterobacteriales’ of the gammaproteobacteria. *International Journal of Systematic and Evolutionary Microbiology*, 63(Pt 6), 2013.
- [21] Youzhong Liu, Sandrine Rousseaux, Raphaëlle Tourdot-Maréchal, Mohand Soudou, Régis Gougeon, Philippe Schmitt-Kopplin, and Hervé Alexandre. Wine microbiome: A dynamic world of microbial interactions. *Critical Reviews in Food Science and Nutrition*, 57(4):856–873, 2017. PMID: 26066835.
- [22] Youzhong Liu, Sandrine Rousseaux, Raphaëlle Tourdot-Maréchal, Mohand Soudou, Régis Gougeon, Philippe Schmitt-Kopplin, and Hervé Alexandre. Wine microbiome: A dynamic world of microbial interactions. *Critical Reviews in Food Science and Nutrition*, 57(4):856–873, 2017. PMID: 26066835.
- [23] Michael I. Love, Wolfgang Huber, and Simon Anders. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biology*, 15:550, 2014.
- [24] Yao Lu, Xiaolei Guan, Rongbin Li, Jie Wang, Yaqiong Liu, Yanli Ma, Jiawei Lv, Suwen Wang, and Jianlou Mu. Comparative study of microbial communities and volatile profiles during the inoculated and spontaneous fermentation of persimmon wine. *Process Biochemistry*, 100:49–58, 2021.
- [25] Julian R Marchesi and Jacques Ravel. The vocabulary of microbiome research: a proposal. *Microbiome*, 3:31, July 2015.
- [26] Antonio Morata, Carlos Escott, Maria Bañuelos, Iris Loira, Juan Fresno, Carmen Gonzalez, and Suárez-Lepe. Contribution of non-saccharomyces yeasts to wine freshness. a review. *Biomolecules*, 10:34, 12 2019.

- [27] Centers of Disease Control and Prevention. E. coli and food safety, April 2023. [Citované 2023-05-05] Dostupné z <https://www.cdc.gov/foodsafety/communication/ecoli-and-food-safety.html>.
- [28] N.A. O’Leary, M.W. Wright, Ciufu Brister, JR, Haddad S, McVeigh D, Rajput R, Robbertse B, Smith-White B, Ako-Adjei B, Astashyn D, Badretdin A, Bao A, Blinkova Y, Brover O, Chetvernin V, Choi V, Cox J, Ermolaeva E, Farrell O, Goldfarb CM, Gupta T, Haft T, Hatcher D, Hlavina E, Joardar W, Kodali VS, Li VK, Maglott W, Masterson D, McGarvey P, Murphy KM, O’Neill MR, Pujar K, Rangwala S, Rausch SH, Riddick D, Schoch LD, Shkeda C, Storz A, Sun SS, Thibaud-Nissen H, Tolstoy F, Tully I, Vatsan RE, Wallin AR, Webb C, Wu D, Landrum W, Kimchi MJ, Tatusova A, DiCuccio T, Kitts M, Murphy P, Pruitt TD, and K.D. Reference sequence (refseq) database at ncbi: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* Jan 4;44(D1):D733-45 PubMed PubMedCentral.
- [29] Robert K. Peet. Relative diversity indices. *Ecology*, 56(2):496–498, 1975.
- [30] Cátia Pinto, Diogo Pinho, Remy Cardoso, Valéria Custódio, Joana Fernandes, Susana Sousa, Miguel Pinheiro, Conceição Egas, and Ana C. Gomes. Wine fermentation microbiome: a landscape from different portuguese wine appellations. *Frontiers in Microbiology*, 6, 2015.
- [31] Christian Quast, Elmar Pruesse, Pelin Yilmaz, Jan Gerken, Timmy Schweer, Pablo Yarza, Jörg Peplies, and Frank Oliver Glöckner. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*, 41(Database issue):D590–6, November 2012.
- [32] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021.
- [33] C. Ricotta and J. Podani. On some properties of the bray-curtis dissimilarity and their ecological meaning. *Ecological Complexity*, 31:201–205, 2017.
- [34] The scikit-bio development team. scikit-bio: A bioinformatics library for data scientists, students, and developers, 2020.
- [35] G.R. Scollary. Wines: Wine production. In Benjamin Caballero, Paul M. Finglas, and Fidel Toldrá, editors, *Encyclopedia of Food and Health*, pages 571–576. Academic Press, Oxford, 2016.
- [36] Kirti Shekhawat, Tristan Jade Porter, Florian F. Bauer, and Mathabatha E. Setati. Employing oxygen pulses to modulate lachancea thermotolerans–saccharomyces

- cerevisiae chardonnay fermentations. *Annals of Microbiology*, 68(2):93–102, Feb 2018.
- [37] Wenhua Tong, Ping He, Ying Yang, Zongwei Qiao, Dan Huang, Huibo Luo, and Xinjun Feng. Occurrence, diversity, and character of bacillaceae in the solid fermentation process of strong aromatic liquors, Dec 2021.
- [38] Janja Trček, Aleksander Mahnič, and Maja Rupnik. Diversity of the microbiota involved in wine and organic apple cider submerged vinegar production as revealed by dhpc analysis and next-generation sequencing. *International Journal of Food Microbiology*, 223:57–62, 2016.
- [39] Andrea Vannini and Gabriele Chilosi. Botrytis infection: Grey mould and noble rot - wiley online library, Apr 2013.
- [40] Carla Viridis, Krista Sumbly, Eveline Bartowsky, and Vladimir Jiranek. Lactic acid bacteria in wine: Technological advances and evaluation of their functional role. *Frontiers in Microbiology*, 11, 2021.
- [41] Michael L. Waskom. seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60):3021, 2021.
- [42] Caroline R Wensel, Jennifer L Pluznick, Steven L Salzberg, and Cynthia L Sears. Next-generation sequencing: insights to advance clinical investigations of the microbiome. *J Clin Invest*, 132(7), April 2022.
- [43] Fengmei Zhu, Bin Du, and Jun Li. Aroma compounds in wine. In Antonio Morata and Iris Loira, editors, *Grape and Wine Biotechnology*, chapter 12. IntechOpen, Rijeka, 2016.
- [44] Slovenský Hydrometeorologický Ústav. Klimagramy, atmosférické zrážky 2018. [Citované 2023-05-05] Dostupné z [https://www.shmu.sk/sk/?page=1&id=klimat\\_operativneudaje2&identif=11816&rok=2018&obdobie=1981-2010](https://www.shmu.sk/sk/?page=1&id=klimat_operativneudaje2&identif=11816&rok=2018&obdobie=1981-2010).
- [45] Slovenský Hydrometeorologický Ústav. Klimagramy, atmosférické zrážky 2019. [Citované 2023-05-05] Dostupné z [https://www.shmu.sk/sk/?page=1&id=klimat\\_operativneudaje2&identif=11816&rok=2019&obdobie=1981-2010](https://www.shmu.sk/sk/?page=1&id=klimat_operativneudaje2&identif=11816&rok=2019&obdobie=1981-2010).



# Príloha A: obsah elektronickej prílohy

V elektronickej prílohe sa nachádza zdrojový kód v jazyku Python pre vykonanie jednotlivých analýz. Taktiež sa tu nachádzajú všetky tabuľky, ktoré boli výstupom diferenčných analýz DESeq2, pretože boli moc rozsiahle na to, aby boli uvedené priamo v texte. Taktiež sú tam všetky rarefaction grafy.