

**Univerzita Karlova v Praze, Přírodovědecká fakulta**

**Katedra filosofie a dějin přírodních věd**

**Charles University in Prague, Faculty of Science**

**Department of Philosophy and History of Science**

Doktorský studijní program: Teoretická a evoluční biologie

Ph.D. study program: Theoretical and Evolutionary Biology

Autoreferát disertační práce

Summary of the Ph.D. Thesis



**Metagenomické profilování mikrobiálních společenstev**

**Metagenomic profiling of microbial consortia**

**Jakub Rídl**

Školitel/Supervisor: Mgr. Jan Pačes, Ph.D.

Praha, 2016

## Obsah

Obsah.....	i
Abstrakt .....	ii
Abstract.....	iii
1 Úvod.....	1
2 Cíle práce.....	2
3 Materiál a metody.....	3
4 Výsledky a diskuze.....	6
5 Závěry.....	13
1 Introduction .....	15
2 Aims of the study .....	16
3 Material and methods .....	17
4 Results and discussion.....	20
5 Conclusions .....	26
Použitá literatura / References.....	28
Curriculum vitae.....	32
Seznam publikací / Selected publications .....	34

## Abstrakt

Metody molekulární biologie umožňují studium mikrobiální diverzity a analýzu genů kódujících procesy a biochemické dráhy jednotlivých mikroorganismů a celých mikrobiálních společenstev. K tomu zásadně přispěl vývoj technologií masivně paralelního sekvenování DNA. Tyto metody rozšířily možnosti zkoumání diverzity od studia jednotlivých genomů modelových a v laboratoři kultivovatelných mikroorganismů přes jednoduché komunity v extrémním prostředí až k výzkumům komplexních mikrobiálních konsorcií. Tento experimentální přístup je založen na analýze celého metagenomu.

Pozornost je věnována ekosystémům negativně ovlivněným lidskou aktivitou, kde mikroorganismy dokáží nejen přežít, ale také adaptovat metabolismus k využívání a odbourávání látek toxických pro vyšší organismy. Příkladem je bakterie *Achromobacter xylosoxidans* A8 izolovaná z půdy kontaminované toxickými chlorbenzoáty. Sekvence a analýza genomu *Achromobacter xylosoxidans* A8 umožnila studium genů kódujících enzymy zapojené do degradace chlorbenzoátů v kontextu kompletní genetické informace.

V extrémně kyselém prostředí bývalého dolu ve Zlatých Horách (Česká republika) vznikají zajímavé útvary bakteriálního biofilmu, želatinové stalaktity. Ty jsou tvořené jednou z taxonomicky nejjednodušších komunit s majoritním zastoupením dvou bakterií rodu *Ferrovum* a *Acidithiobacillus*. Sekvence DNA izolované přímo ze vzorku biofilmu a bionformatická analýza získaných dat nám umožnila sestavení téměř kompletních genomů bez nutnosti pěstování těchto těžko kultivovatelných bakterií. Analýzou RNA jsme identifikovali aktivně transkribované geny konsorcia. Rekonstruované genomy představují unikátní kombinace genů vyvinuté adaptací na konkrétní ekosystém důlní vody ve Zlatých Horách.

Mnohem komplexnější systém představují společenstva půdních mikroorganismů. Sekvenováním celkové „půdní DNA“, amplifikací genů pro 16S rRNA a metagenomickým profilováním jsme odhalili vliv několika druhů rostlin na taxonomické složení a funkční potenciál mikrobiálních společenstev schopných biodegradace polutantů.

## Abstract

Methods of molecular biology enable studies on microbial diversity based on analysis of genes encoding processes and biochemical pathways of individual microorganisms and also complete microbial consortia. For this a crucial step was elaboration of new technologies of high-throughput DNA sequencing. These methods made it possible to advance studies of diversity from analysis of genomes of model microorganisms easily cultivated in laboratories to simple communities living in extreme environments and further to complex microbial consortia. This experimental approach is based on metagenomic analyses.

Important are studies on ecosystems negatively affected by human activity where microorganisms not only survive but they can convert their metabolism to degrade compounds toxic for higher organisms. An example is bacterium *Achromobacter xylosoxidans* A8 isolated from soils contaminated by toxic chlorobenzoates. Sequencing and analysis of *Achromobacter xylosoxidans* A8 genome made it possible to study genes coding for enzymes that are involved in chlorobenzoates degradation in the context of the complete genetic background.

An interesting microbial biofilm – gelatinous stalactites – was discovered in an extremely acidic environment of the abandoned mine in Zlaté Hory (the Czech Republic). It is formed by a simple consortium with predominantly present bacterial genus *Ferroplasma* and genus *Acidithiobacillus*. DNA sequencing of the biofilm sample and bioinformatic analysis of the obtained data enabled us to reconstruct two nearly complete genomes without growing individual bacterial strains that are difficult to cultivate. By RNA analysis expressed genes of the consortium were identified. Thus, metabolic processes of the present bacterial strains can be described. Comparative analysis shows unique properties of individual members of the analyzed consortium that are important in the particular natural conditions.

Much more complex systems are found in soil. We sequenced „soil DNA“, amplified 16S rRNA genes and constructed metagenomic profiles of the samples. Using this approach we discovered effect of different plant species on taxonomic composition and functional potential of microbial communities that can degrade pollutants.

# 1 Úvod

Mikroorganismy obývají veškeré části povrchu planety Země a jejich činnost zásadním způsobem ovlivňuje ekosystémy, ve kterých žijí. Podílejí se na biochemických procesech, jsou zodpovědné za koloběh organických sloučenin a globální koloběh živin, díky čemuž veškeré organismy v biosféře závisejí na mikrobiálních aktivitách. Rozvoj metod molekulární biologie a především vývoj technologií masivně paralelního sekvenování DNA umožnil nejen studium nepřeberné míry mikrobiální diverzity v různých prostředích, ale také analýzu genů zodpovědných za kódování funkčních procesů mikroorganismů a celých mikrobiálních společenstev. Tato oblast označovaná metagenomika vychází ze dvou jevů: (1) jen méně než procento mikroorganismů dokážeme kultivovat a následně analyzovat v laboratorních podmínkách a (2) na jednotlivých funkčních procesech v ekosystému se podílí celá řada druhů tvořících mikrobiální konsorcia (Handelsman et al., 1998).

V této práci používáme spektrum sekvenačních technologií a bioinformatických nástrojů v bakteriálních genomových projektech a metagenomických analýzách. Práce provádí čtenáře příběhem od rozklíčování jednoho bakteriálního genomu, přes studium nejjednodušších bakteriálních konsorcií, kde pospolu žije jen několik organismů, až po analýzu komplexních mikrobiálních společenstev.

Pozornost je věnována ekosystémům negativně ovlivněným lidskou aktivitou, kde mikroorganismy dokáží nejen přežít, ale také adaptovat metabolismus k využívání a odbourávání látek toxických pro vyšší organismy. Prvním příkladem je bakterie *Achromobacter xylosoxidans* A8 využívající toxické chlorbenzoáty jako zdroj uhlíku a energie, důležitý člen konsorcia mikroorganismů schopného degradovat polutanty v kontaminované půdě. Sekvenační analýza genomu *A. xylosoxidans* A8 poskytuje první kompletní bakteriální genom stanovený pouze sekvenační platformou nové generace v české laboratoři a umožňuje další studium biodegradačního potenciálu bakterie.

V extrémně kyselém ekosystému důlní vody ve Zlatých Horách (Česká Republika) jsou zase acidofilní bakterie schopny tvořit makroskopické nárůsty biofilmu, například dlouhé želatinové stalaktity. Předchozí studie ukázala přítomnost taxonomicky velmi jednoduché komunity ve vzorku krápníkového biofilmu tvořené převážně dvěma bakteriemi rodu *Ferrovum* a *Acidithiobacillus* (Falteisek & Čepička, 2012). S využitím masivně paralelního sekvenování DNA a RNA jsme se rozhodli analyzovat geny zapojené do funkčních a metabolických procesů vyvinutých adaptacemi na konkrétní extrémní podmínky prostředí s hodnotou pH < 3.

Zcela odlišný systém představují společenstva půdních mikroorganismů obecně vykazující nejvyšší míru biodiverzity na Zemi. Převážně dynamické prostředí vzniká v těsném okolí kořenů rostlin. Předchozí výzkumy dokládají vliv výsadby některých druhů rostlin na snížení koncentrace toxických polychlorovaných bifenylnů (PCB) v kontaminované půdě (Ionescu et al., 2009; Kurzawová et al., 2012). Otázkou však zůstávají samotné mechanismy působení rostlin – zda fungují pouze prostřednictvím vlastní absorpce a transformace polutantů, či se zásadním způsobem podílejí na utváření a podporování půdních mikrobiálních společenstev schopných biodegradace bifenylnů. Sekvenační a bioinformatickou analýzou celkové „půdní DNA“ jsme hodnotili vliv kořenů křenu (*Armoracia rusticana*), lilku (*Solanum nigrum*) a tabáku (*Nicotiana tobacum*) na taxonomické složení a funkční potenciál mikrobiálních společenstev v půdě kontaminované PCB.

## 2 Cíle práce

- Sekvenační analýzou genomové DNA s využitím GS FLX přístroje (Roche) stanovit sekvenci genomu bakterie *Achromobacter xylosoxidans* A8, která je schopná využívat chlorbenzoáty jako zdroj uhlíku a energie. Na jejím základě predikovat oblasti kódující geny, vytvořit jejich funkční anotaci a poskytnout základ pro analýzu degradační dráhy chlorbenzoátů v kontextu kompletní genetické informace.
- Aplikovat přístupy ověřené na základě analýzy genomu *A. xylosoxidans* A8 pro výzkum taxonomicky jednoduché komunity tvořící krápníkové biofilmy v prostředí extrémně kyselé důlní vody bývalého dolu ve Zlatých Horách (Česká Republika). Rovněž amplifikací a sekvenací 16S rRNA genů určit složení daného mikrobiálního konsorcia. Sekvenovat celkovou metagenomovou DNA ze vzorku biofilmu, sestavit souvislé sekvence pokrývající genomy majoritně zastoupených mikroorganismů a anotovat predikované geny. Na základě anotace rekonstruovat fyziologické a metabolické vlastnosti mikroorganismů vyvinuté v konkrétním ekosystému s extrémně nízkou hodnotou pH. Prostřednictvím sekvenování cDNA zpětně přepsané z mRNA izolované z biofilmu identifikovat exprimované geny aktivně využívané zastoupenými organismy.

- Aplikací ampliconového a shotgunového sekvenování a adaptací bioinformatických metod analyzovat komplexní společenstva mikroorganismů v půdě kontaminované polychlorovanými bifenyly. Sekvenovat amplicony 16S rRNA genů a celkovou metagenomovou DNA, anotovat čtení a sestavit taxonomické a funkční profily zkoumaných vzorků. Provést komparativní analýzu profilů vzorků ovlivněných přítomností křenu (*Armoracia rusticana*), lilku (*Solanum nigrum*) a tabáku (*Nicotiana tobacum*) a aplikací minerálního hnojiva.

### 3 Materiál a metody

#### 3.1 Analýza genomu *Achromobacter xylosoxidans* A8

Kultura buněk *A. xylosoxidans* A8 byla pěstována při teplotě 28°C v ABC minerálním mediu, celkovou DNA jsme izolovali za použití Gentra Puregene Yeast/Bact. A kitu (Qiagen). Pro 454 pyrosekvenování jsme připravili shotgunovou knihovnu krátkých fragmentů DNA přibližně o délce 700 bp sekvenovaných z jednoho směru, a párové knihovny s délkou úseků (inzertů) 3 kb a 8 kb, kdy jsou následně sekvenovány oba konce, což má význam při následném sestavení čtení a rekonstrukci genomové sekvence. Sekvenování jsme provedli na přístroji GS FLX verze Titanium (Roche).

První sestavení (assembly) bylo provedeno programem Newbler verze 2.3 (Roche), mezery jsme doplnili individuálně lokálním sestavením sekvencí. Predikci genů kódujících proteiny (*coding sequences*, CDS) jsme provedli paralelně třemi programy: Critica (Badger & Olsen, 1999), Glimmer (Delcher et al., 1999) a Prodigal (Hyatt et al., 2010). Software tRNAscan (Lowe & Eddy, 1997) a Aragorn (Laslett & Canback, 2004) jsme použili pro nalezení genů pro transferové RNA (tRNA) a transferové-mediátorové RNA (tmRNA); ribosomální geny (rRNA) jsme identifikovali programem RNAmmer (Lagesen et al., 2007). Funkční anotaci jsme přiřadili predikovaným CDS na základě vyhledání podobnosti k sekvencím v databázích UniRef90 (Tatusov et al., 1999), NCBI-NR (Benson et al., 2010) a KEGG (Kanehisa et al., 2008) programem BLASTP ( $E\text{-value} < 10^{-10}$ ). Zároveň jsme využili pro anotaci automatické online nástroje RAST Annotation Server (Aziz et al., 2008) a IGS Annotation Engine (<http://manatee.sourceforge.net>). Výsledky anotace byly ověřeny v programu Artemis (Carver et al., 2008). Začátek replikace chromosomu *A. xylosoxidans* A8 jsme identifikovali podle průběhu rozdílného obsahu G a C bazí mezi vedoucím a zpoždujícím se vláknem replikace (tzv. GC skew analýza) (Lobry, 1998).

### **3.2 Metagenomika a metatranskriptomika bakteriální komunity z extrémně kyselého ekosystému důlní vody ve Zlatých Horách**

Několik centimetrů želatinového bakteriálního stalaktitu (vzorek označen ZH7) bylo odebráno ze stropu šachty bývalého dolu ve Zlatých Horách v hloubce přibližně 190 m pod zemí. Fyzikálně-chemické vlastnosti vody, která jím protékala, ukazují na typickou kyselou důlní vodu s pH hodnotou 2,9, teplotou 8,7 °C, s velmi nízkou koncentrací organického uhlíku a celkovou mineralizací přibližně 5,8 g/l, tvořenou převážně sírany, Fe, Al, Mg a Cu (Falteisek & Čepička, 2012). DNA byla izolována s využitím sady ZR Soil Microbe DNA Kit (Zymo Research), RNA jsme izolovali pomocí Qiagen RNeasy Mini Kit s přechodí homogenizací přístrojem TissueLyser (Qiagen). K odstranění ribosomální RNA jsme použili sadu Ribo-Zero rRNA Kit (Epicentre / Illumina). 454 shotgunové a párové DNA knihovny (s inzertem 3 – 8,8 kb) byly připraveny dle protokolů firmy Roche a sekvenovány na přístroji GS FLX verze Titanium (Roche). Knihovna krátkých fragmentů byla také připravena sadou TruSeq DNA Sample Prep Kit (Illumina) pro sekvenování instrumentem MiSeq (Illumina) v režimu 250 bp z obou konců (2x250 *paired-end*). RNA sloužila k přepisu do cDNA a následné tvorby sekvenační knihovny pomocí NEXTflex RNA-seq Kit (BIOO Scientific). Sekvence byla provedena přístrojem MiSeq (2x150 *paired-end*).

Pro detailní taxonomickou analýzu vzorku jsme prostřednictvím PCR primerů (f8-27 a r372-357) amplifikovali úsek 16S rRNA genu a produkty (amplikony) sekvenovali přístrojem GS FLX Titanium. Zpracování sekvencí bylo provedeno programem Mothur verze 1.28 (Schloss et al., 2009), který umožňuje kontrolu a korekci kvality čtení, alignment sekvencí a tvorbu operačních taxonomických jednotek (OTU) se sekvenční podobností alespoň 97 %, které zhruba odpovídají přítomným bakteriálním druhům.

Sestavení genomových 454 čtení jsme provedli programem Newbler verze 2.8 (Roche), MiSeq-DNA a MiSeq-RNA čtení jsme použili pro zaplnění mezer programem GapCloser (Lua et al., 2012). MiSeq-RNA sekvence byly sestaveny do transkriptů pomocí programu Trinity (Grabherr et al., 2011).

Pro predikci CDS v sestavených genomových sekvencích jsme využili programy Critica (Badger & Olsen, 1999), Glimmer (Delcher et al., 1999) a Prodigal (Hyatt et al., 2010). Pro predikované CDS jsme vyhledali podobné sekvence v databázi NCBI-NR programem BLASTP ( $E\text{-value} < 10^{-10}$ ). Rovněž jsme využili anotační server RAST (Aziz et al., 2008) a KAAS – KEGG Automatic Annotation Server (Moriya et al., 2007). Sestavené sekvence jsme



roztřídili a přiřadili genomům dvou majoritně zastoupených bakterií na základě vlastní pipeline zahrnující informace o kompozici sekvencí, výsledných parametrů assembly a anotace CDS.

Pro identifikaci vysoce exprimovaných genů jsme mapovali MiSeq-RNA čtení na anotované genomové sekvence programem STAR (Dobin et al., 2013) a na základě počtu RNA čtení mapujících do oblastí CDS jsme stanovili relativní transkripční aktivitu genů.

### ***3.3 Metagenomické profilování mikrobiálních společenstev v kontaminované půdě ovlivněné kořeny rostlin***

Půda kontaminovaná PCB odebraná ze skládky u obce Lhenice v Jižních Čechách byla použita v laboratoři pro pěstování křenu (*Armoracia rusticana*, vzorek Hor), lilku (*Solanum nigrum*, vzorek Nig) a tabáku (*Nicotiana tabacum*, vzorek Tob) po dobu 6 měsíců. Polovina od každých vzorků s rostlinami (vzorky HorF, NigF a TobF) byla hnojena minerálním hnojivem Univerzal KH (Nohel Garden, Česká Republika), stejným způsobem byly připraveny také kontrolní vzorky bez výsadby rostlin (vzorky Ctrl a CtrlF). DNA z půdy jsme izolovali sadou PowerMax Soil DNA Isolation Kit (MoBio Laboratories Inc., USA). Shotgunová sekvenační knihovna byla připravena podle protokolu firmy Roche, pro přípravu amplikonové knihovny jsme amplifikovali úsek genu pro 16S rRNA s využitím primerů f563-577 a r1406-1392. Knihovny byly sekvenovány na přístroji GS FLX verze Titanium resp. Titanium+ (Roche).

Shotgunová čtení byla anotována prostřednictvím serveru MG-RAST (Meyer et al., 2008), na základě anotovaných čtení jsme sestavili taxonomické a funkční profily pro jednotlivé vzorky. Amplikonová čtení jsme zpracovali v programu Mothur verze 1.28 a výsledky uspořádali rovněž do taxonomických profilů. Nemetrické mnohorozměrné škálování (nMDS) jsme provedli v balíku vegan (Oksanen et al., 2013) pro statistický software R (R Development Core Team, 2009).

Pro hodnocení rozdílně zastoupených taxonů a funkčních genů v profilech vzorků jsme adaptovali bioinformatické metody z balíku DESeq2 (Love et al., 2014) ve statistickém programu R (R Developmental Core Team, 2009) původně navržené pro analýzu rozdílně exprimovaných genů na základě RNA-seq dat. Statisticky významné rozdíly byly vizualizovány formou heatmap grafů v programu R.

## 4 Výsledky a diskuze

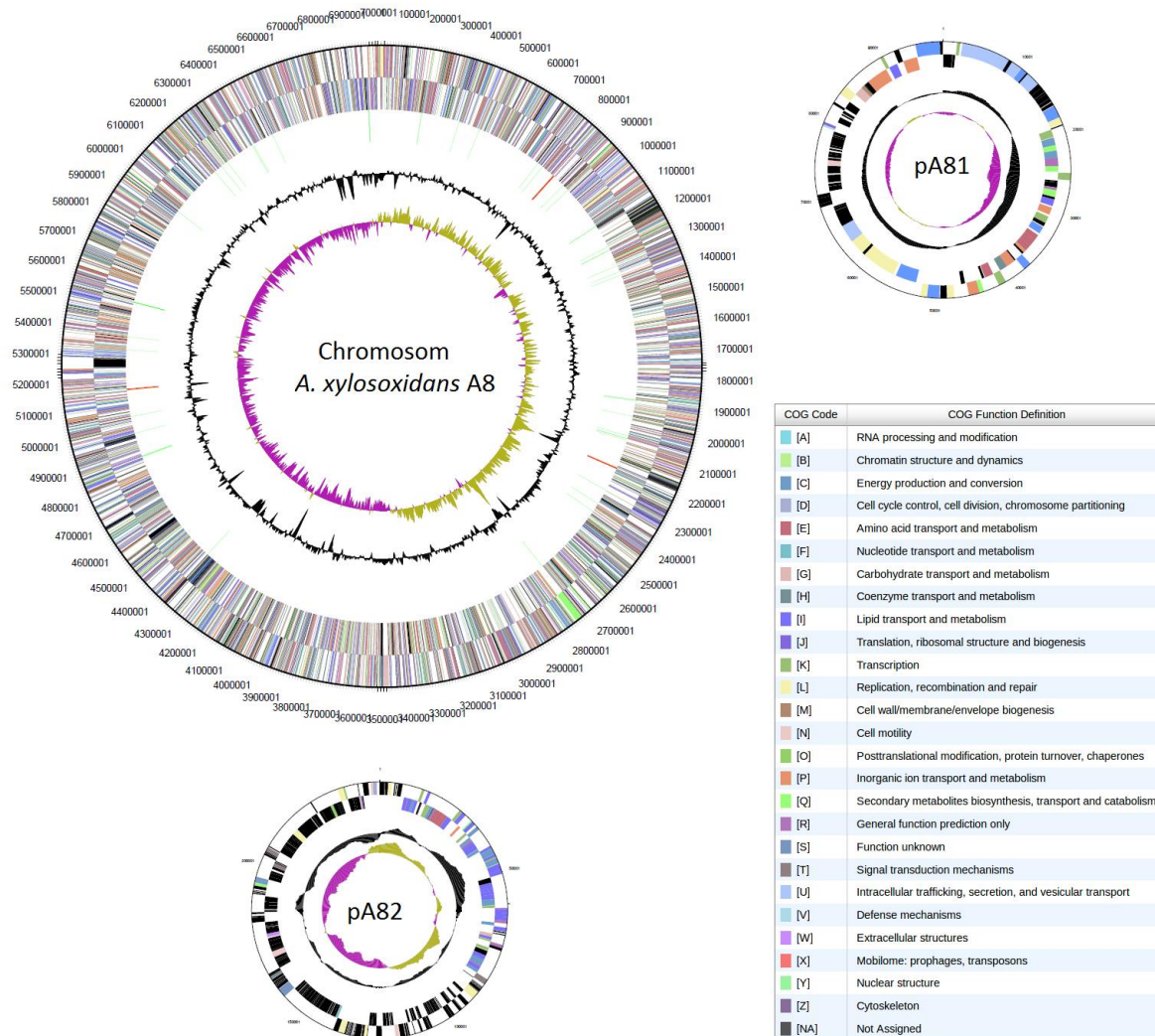
### 4.1 Genomová sekvence *Achromobacter xylosoxidans* A8

Sestavili jsme kompletní genomovou sekvenci *A. xylosoxidans* A8 sestávající z chromosomu o délce 7 013 095 bp a dvou plasmidů: pA81 (98 156 bp) a pA82 (247 895 bp) (obrázek 1). Chromosom vykazuje vysoký obsah GC párů (66 %), plasmidy v průměru obsahují 62 % GC párů. Začátek replikace chromosomu jsme identifikovali v blízkosti *dnaA* genu.

Identifikovali jsme 6 459 kódujících sekvencí (CDS) na chromosomu *A. xylosoxidans*, 104 CDS v rámci plasmidu pA81 a 252 CDS v plasmidu pA82. Chromosom dále obsahuje 3 rRNA operony, 60 tRNA pro přepis všech 20 aminokyselin a selenocysteinu a jeden gen pro tmRNA. Kódující oblasti pokrývají 91,3% genomu *A. xylosoxidans*. Na základě anotace byla přiřazena funkce 5 620 nalezeným CDS (82,5%), funkci u 783 (11,5%) CDS se prohledáním databází přiřadit nepodařilo. Ostatních 412 predikovaných CDS nevykazovalo podobnost s žádným záznamem v databázích.

Identifikovali jsme operon *ohbRAB* kódující a regulující enzym dioxygenázu iniciující degradaci chlorbenzoátů na chlorkatechol, kompletní klastr genů *mocpRABCD* pro katabolismus chlorkatecholů a *hyb* operon umožňující degradaci salicylátu, čímž jsme potvrdili jejich umístění na pA81 plasmidu (Jenčová et al., 2008). Dále jsme identifikovali přibližně 70 genů kódujících další dioxygenázy, které mohou být rovněž asociovány s bioremediačními schopnostmi bakterie.

Kontaminace prostředí aromatickými polutanty je často doprovázena také vysokým obsahem těžkých kovů, které ovlivňují složení bakteriálních konsorcií (Roane et al., 2001; Gremion et al., 2004). Několik predikovaných genů *A. xylosoxidans* A8 bylo asociováno s rezistencí vůči  $Pb^{2+}$ ,  $Cd^{2+}$  a  $Zn^{2+}$  iontům prostřednictvím aktivního transportu efluxními pumpami, ATPázami typu P. Patří sem skupina *pbt* genů kódovaná rovněž na plasmidu pA81, jejíž funkce u *A. xylosoxidans* A8 byla následně potvrzena skupinou dr. Kotrby expresí v bakterii *E. coli* (Hložková et al., 2013). Ztráta plasmidu pA81 však vede pouze k poklesu, nikoli ztrátě rezistence u *A. xylosoxidans* A8 (Hložková et al., 2013), což ukazuje na roli dalších ATPázových pump kódovaných geny predikovanými v sekvenci chromosomu.



**Obrázek 1:** Grafické znázornění chromosomu *A. xylosoxidans* A8 a dvou plasmidů pA81 a pA82. Cirkulární obrázky znázorňují (od periferie ke středu): (1) geny v (+) orientaci (barevné kódy podle funkčních systémů viz legenda vpravo), (2) geny v (-) orientaci (barevné kódy podle funkčních systémů viz legenda vpravo), (3) GC obsah lišící se od průměru (počítaný v okně 1000 bp), (4) GC skew (počítáno v okně 200 bp).

Genom rovněž obsahuje 28 genů spojených s patogenními vlastnostmi a 33 genů asociovaných s produkcí toxinů a rezistencí vůči antibiotikům. Bakteriální druh *Achromobacter xylosoxidans* byl původně izolován z pacientů a popsán jako oportunistický patogen u jedinců s oslabenou imunitou (Holmes et al., 1977). Li et al. (2013) provedli komparativní analýzu založenou na naší publikované sekvenci kompletního genomu společně s pěti dalšími dostupnými nekompletními sekvencemi genomů rodu *Achromobacter*, která naznačuje fylogenetickou příbuznost s lidskými patogeny rodu *Bordetella*.

## 4.2 Taxonomické složení a funkční procesy acidofilních bakterií tvořících

### krápníkový biofilm ve Zlatých Horách

Výsledky amplikonového sekvenování úseků 16S rRNA genu (tabulka 1) ukazují majoritní zastoupení dvou bakterií rodu *Ferrovum* a *Acidithiobacillus* doprovázené třetím dosud nepopsaným bakteriálním druhem, jehož 16S rRNA sekvence se shoduje v databázi NCBI-NT s 16S rRNA genem nekultivované bakterie amplifikovaným z environmentálního vzorku. Řádově nižší zastoupení vykazuje bakterie rodu *Acidiphilium*. Acidofilní heterotrofní zástupci tohoto rodu byli dříve popsáni v konsorciích tvořených majoritně rodem *Ferrovum* (Johnson et al., 2014; Ullrich et al., 2016). Obdobně jsou zastoupeny další tři dosud nepopsané bakteriální druhy, doprovázené ostatními sporadicky zastoupenými OTU. Reprezentativní sekvence OTU1 vykazovala 100% sekvenční shodu s druhem *Ferrovum myxofaciens* P3G; pro potřeby následujícího textu budeme pro námi analyzovanou bakterii používat označení „*Ferrovum myxofaciens*“ ZH7. Bakterie rodu *Acidithiobacillus*, která se vyskytuje v našem vzorku (OTU2), nese 16S rRNA gen s 99% podobností k druhu *Acidithiobacillus ferrooxidans* a nadále o ní budeme referovat jako o „*Acidithiobacillus*“ linie ZH7B.

**Tabulka 1:** Přehled počtu amplikonových 16S rRNA čtení náležejících jednotlivým OTU. BLASTN vs. NCBI-NT a podob. [%] udává klasifikaci na základě vyhledání podobností proti NCBI-NT databázi a míru podobnosti se záznamem v NCBI-NT. RDP klasifikace uvádí taxonomickou klasifikaci reprezentativních sekvencí OTU na základě podobnosti s RDP databází, procenta udávají bootstrapový odhad spolehlivosti.

OTU	# čtení	BLASTN vs. NCBI-NT podob. [%]	RDP klasifikace	
OTU1	1736	<i>Ferrovum</i>	100	Betaproteobacteria;100%; Ferroviales;100%; Ferroviales;100%; Ferrovum;100%
OTU2	1938	<i>Acidithiobacillus</i>	99	Gammaproteobacteria;100%; Acidithiobacillales;100%; Acidithiobacillaceae;100%; Acidithiobacillus;100%
OTU3	95	nekultivovaná bakterie	100	Acidobacteria_Gp1;100%; Granulicella;83%
OTU4	103	<i>Acidiphilium</i>	99	Alphaproteobacteria;100%; Rhodospirillales;100%; Acetobacteraceae;100%; Acidiphilium;75%
OTU5	12	<i>Acidobacteria</i>	99	Acidobacteria_Gp1;100%; Gp1;81%
OTU6	44	<i>Acidocella</i>	100	Alphaproteobacteria;100%; Rhodospirillales;100%; Acetobacteraceae;100%; Acidocella;92%
OTU7	36	<i>Alicyclobacillus</i>	98	Bacilli;70%; Bacillales;62%; Bacillaceae 1;9%; Domibacillus;8%
OTU8	4	<i>Leptospirillum</i>	100	Nitrospira;100%; Nitrospirales;100%; Nitrospiraceae;100%; Leptospirillum;100%
OTU9	3	<i>Sphingomonas</i>	99	Alphaproteobacteria;100%; Sphingomonadales;100%; Sphingomonadaceae;100%; Sphingomonas;97%
OTU10	28	<i>Thiomonas</i>	100	Betaproteobacteria;100%; Burkholderiales;100%; Burkholderiales incertae sedis;100%; Thiomonas;100%
OTU11	1020	nekultivovaná bakterie	100	Betaproteobacteria;99%; Gallionellales;72%; Gallionellaceae;72%; Sideroxydans;72%
OTU12	107	NA	100	Gammaproteobacteria;85%; Xanthomonadales;85%; Xanthomonadaceae;85%; Dokdonella;52%
OTU13	65	NA	99	Acidobacteria_Gp1;100%; Terriglobus;76%

Na základě sekvenace metagenomové DNA z biofilmu, sestavení sekvencí a jejich rozřídění jsme byli schopni rekonstruovat téměř kompletní genomové sekvence majoritně zastoupených bakterií „*Ferrovum myxofaciens*“ ZH7 a „*Acidithiobacillus*“ linie ZH7B bez nutnosti jejich kultivace. Funkční anotace sekvencí umožňuje rekonstruovat fyziologické a metabolické procesy. Z výsledků vyplývá, že obě bakterie jsou plně vybaveny pro život v daných extrémních podmínkách, jsou chemolitoautotrofní a jejich metabolické dráhy jsou

zcela oddělené (nejsou syntrofní). Obě získávají uhlík fixací CO<sub>2</sub> a energii oxidací železa. „*Acidithiobacillus*“ linie ZH7B je navíc schopen oxidovat redukovanou síru a jako zdroj dusíku dokáže metabolizovat močovinu, což umožňuje dvěma popsaným bakteriím využívat z části oddělené niky.

V porovnání s dalšími zástupci rodu *Ferrovum* je námi popsaná bakterie „*Ferrovum myxofaciens*“ ZH7 schopna fixovat molekulární dusík, je vybavena bakteriálním imunitním systémem CRISPR/Cas a její buňky jsou „obrněny“ kapsulí, což reflektuje nároky daného ekosystému.

Rekonstruovaný genom „*Acidithiobacillus*“ linie ZH7B se vyznačuje směsicí znaků vyskytujících se u dříve popsaných druhů *At. ferrooxidans* a *At. ferrivorans*. 16S rRNA gen vykazuje nejvyšší míru podobnosti s 16S rRNA sekvencí genu *At. ferrooxidans*. „*Acidithiobacillus*“ linie ZH7B však využívá geny zapojené do dráhy oxidace železa, která byly navržena u linií druhu *At. ferrivorans* (Hallberg et al., 2010; Amouric et al., 2011). Zajímavé také je, že *At. ferrivorans* byl popsán jako vysoce pohyblivá bakterie (Hallberg et al., 2010), zatímco rekonstruovaný genom „*Acidithiobacillus*“ linie ZH7B strukturní geny pro bakteriální bičík postrádá. Obsahuje však geny pro CRISPR/Cas systém, které v rámci rodu *Acidithiobacillus* lze nalézt pouze v publikovaném genomu *At. ferrooxidans* ATCC 23270. Tato data dokládají variabilitu uvnitř rodu *Acidithiobacillus*. Námi popsaná bakterie představuje unikátní linii, která se vyvinula prostřednictvím specifických adaptací v konkrétním prostředí kyselé důlní vody ve Zlatých Horách.

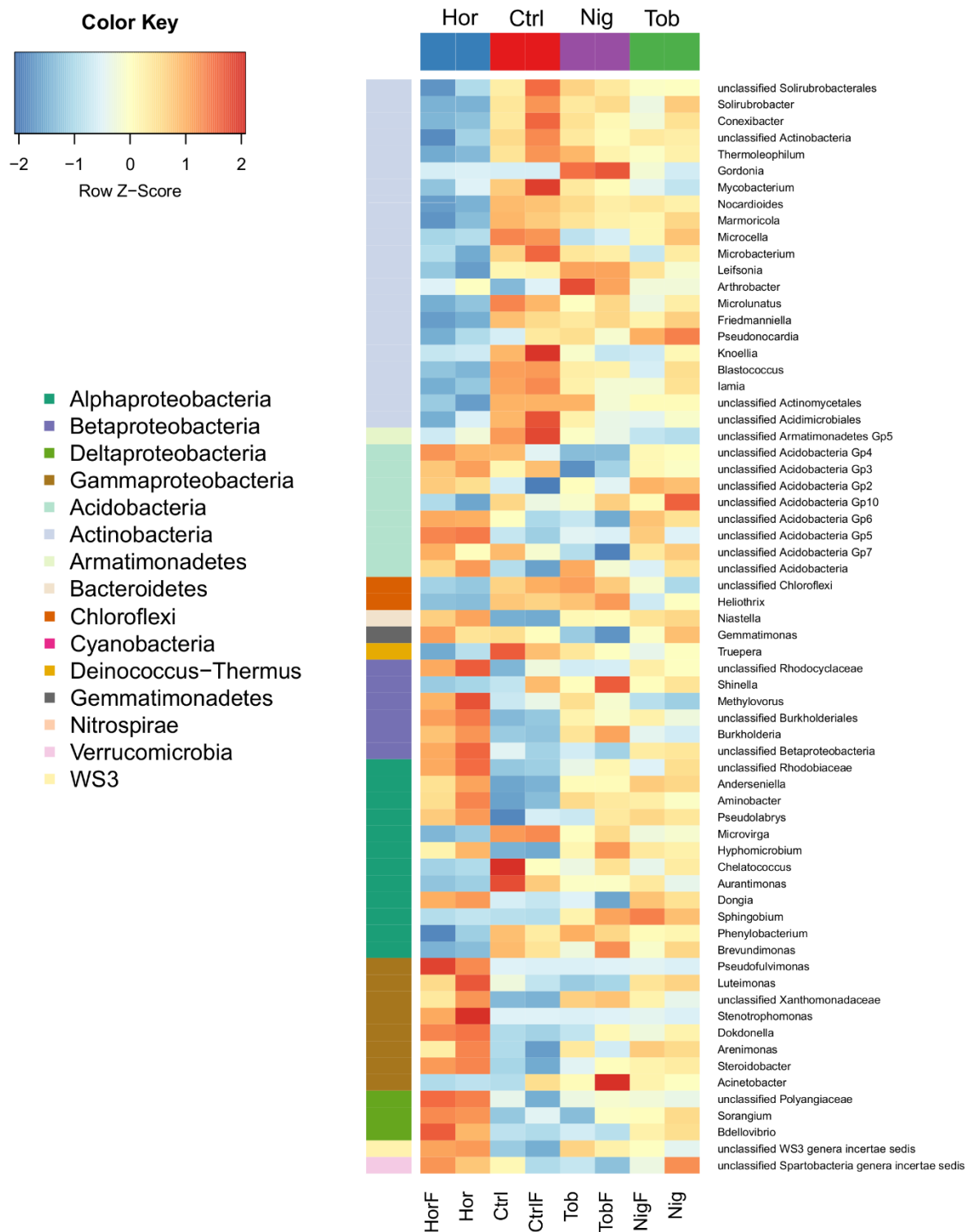
Analýza RNA ze vzorku biofilmu nám umožnila identifikovat exprimované geny. Nejvíce exprimované geny bakterií „*Ferrovum myxofaciens*“ ZH7 a „*Acidithiobacillus*“ linie ZH7B se účastní dráhy oxidace železa a dále adaptací na extrémní podmínky důlního prostředí, kde společným vlivem nízké hodnoty pH a oxidačním stresem hrozí nestabilita buněčných struktur a poškození DNA a proteinů. Zajímavostí také je, že vysoce exprimované geny ostatních členů bakteriálního společenstva jsou z velké části nepopsané, nebo je na základě porovnání s databázemi nelze anotovat vůbec. To vypovídá o výskytu dalších, dosud neznámých bakterií a funkčních procesů, které jsou v daném ekosystému aktivní.

### 4.3 Vliv kořenů rostlin na taxonomické složení a funkční potenciál

#### *mikrobiálních společenstev schopných biodegradace PCB v půdě*

Biodegradační potenciál konsorcií ovlivněných přítomností rostlin demonstruje vyšší zastoupení bakteriálních kmenů, které předchozí studie identifikovaly jako PCB degradéry (obrázek 2). Zemina z okolí kořenů křenu je nabohacena o rod *Burkholderia*, jehož zástupci byli popsáni jako velmi efektivní degradéři PCB (Mukerjee-Dhar et al., 1998; Tillman et al., 2005; Chain et al., 2006; Uhlík et al., 2013), rod *Stenotrophomonas*, který byl dříve izolován ze stejné skládkové zeminy z kultury pěstované na PCB (Uhlík et al., 2013) či *Methylovorus*, jehož schopnost metabolismu bifenyly byla přímo detekována metodou SIP ze vzorků rhizosféry křenu (Uhlík et al., 2009). Naše data rovněž dokládají, že zemina s pěstovaným křenem obsahuje mikrobiální konsorcium, které se nejvíce odlišuje od všech ostatních vzorků. *Arthrobacter* a *Gordonia*, bakteriální rody často asociované s biodegradací PCB (Gilbert & Crowley, 1997; Abraham et al., 2005; Leigh et al., 2007; Koubek et al., 2012), jsou naproti tomu nabohacené u vzorku s pěstovaným tabákem.

Aplikace hnojiva má velký význam nejen v zemědělství, ale je mu také věnována pozornost v ekologii. Předmětem studia jsou vlivy chemického hnojení na konsorcia půdních mikroorganismů (Geisseler & Scow, 2014), zvláště pak z hlediska formování funkčních systémů v zemině a jejich vztažení ke specifickým populacím mikroorganismů (Su et al., 2015). Řada studií poukazuje na změny mikrobiálních komunit v důsledku aplikace chemického hnojiva (Leff et al., 2015; Su et al., 2015), zatímco jiné výzkumy dokládají výrazně vyšší účinek samotné vegetace než použitého hnojiva (Benizri & Amiaud, 2005; Liliensiek et al., 2012). V některých studiích dokonce nebyl prokázán signifikantní efekt hnojení na půdní mikroorganismy (Marschner et al., 2001; Liliensiek et al., 2012). V tomto ohledu přispívají do diskuze výsledky našeho výzkumu: neprokázali jsme signifikantní posun v mikrobiálních konsorciích následkem aplikace chemického, minerálního hnojiva a to jak na úrovni taxonomického složení společenstev tak jejich funkčního potenciálu. Naproti tomu výsadba vybraných rostlin měla za následek změny mikrobiálních populací ve sledovaných vzorcích (obrázek 2 a 3).

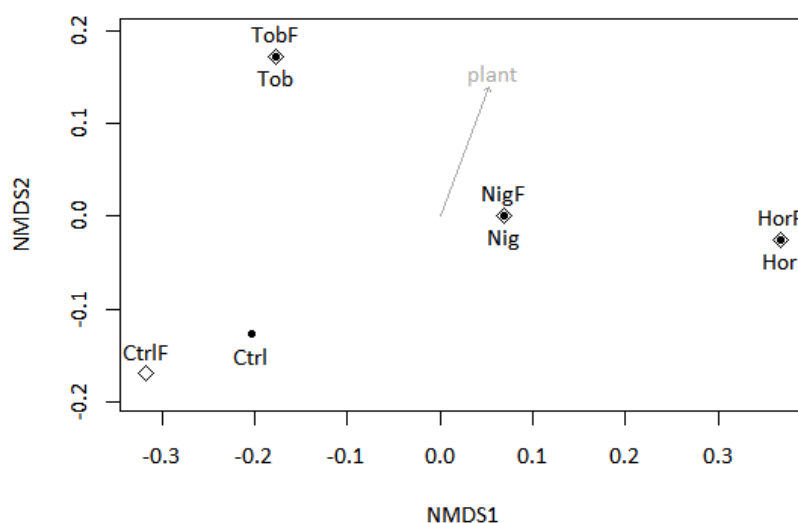


**Obrázek 2:** Statisticky významné rozdíly v zastoupení taxonů ve vzorcích na základě amplikonového sekvenování 16S rRNA genu. Heatmapa ukazuje zastoupení rodů, červená barva značí vyšší zastoupení, modrá nižší. Seskupení na úrovni kmenu je barevně zvýrazněno na levé straně grafu. Nehnojené vzorky: Hor (křen), Nig (lilek), Tob (tabák), Ctrl (kontrola). Hnojené vzorky: HorF (křen), NigF (lilek), TobF (tabák), CtrlF (kontrola).

Mikrobiální konsorcia v zemině bývají selektivně nabohaceny o kopiotrofní či naopak oligotrofní organismy v závislosti na množství přítomného organického uhlíku (Fierer et al.,



2007). Prostředí přímo ovlivněné rhizodepozicí skrze kořeny rostlin obecně upřednostňuje výskyt kopiotrofních mikroorganismů (Dennis et al., 2010). Přestože typ růstové strategie nelze zcela jednoznačně vztáhnout k taxonomické příslušnosti organismu na úrovni kategorií kmene či třídy, členové Actinobacteria, Bacteroidetes, Alphaproteobacteria, Betaproteobacteria a Gammaproteobacteria bývají považováni spíše za kopiotrofní, zatímco kmene Acidobacteria a Planctomycetes za oligotrofní (Fierer et al., 2007; Prober et al., 2015; Leewis et al., 2016). Zvýšená abundance zástupců Alpha-, Beta-, Gammaproteobacteria a kmene Bacteroidetes (obrázek 2) ve vzorcích s pěstovanými rostlinami podporuje hypotézu, že oblast kořenů rostlin favorizuje výskyt kopiotrofních organismů. Výskyt oligotrofních kmenů bakterií Acidobacteria a Planctomycetes nesleduje tak jednoznačný trend. Zatímco kmen Planctomycetes byl málo četný napříč všemi vzorky, zástupci kmene Acidobacteria vykazují signifikantně odlišnou míru zastoupení ve vzorcích, přičemž největší zastoupení můžeme sledovat v prostředí kořenů křenu. Již dříve byl sledován rozdílný výskyt určitých linií Acidobacteria v rhizosféře specifické vegetace (Nunes de la Rocha et al., 2013) a naše výsledky dokládají pozitivní vliv křenu na vznik podmínek preferovaných pro růst členů třídy Acidobacteria.



**Obrázek 3:** Nemetrické mnohorozměrné škálování (NMS, stress < 0.001) na základě 16S rRNA amplikonových dat s proložením vektoru odpovídajícímu vlivu rostlin (*plant*) ( $P$ -value < 0.05, 40 320 permutací). Symboly: vyplněné body – nehnojené vzorky, nevyplněné kosočtverce – hnojené vzorky. Nehnojené vzorky: Hor (křen), Nig (lilek), Tob (tabák), Ctrl (kontrola). Hnojené vzorky: HorF (křen), NigF (lilek), TobF (tabák), CtrlF (kontrola).

Důležitá je v tomto ohledu rovněž skutečnost, že některé funkční kategorie jsou připisovány specificky kopiotrofním či oligotrofním růstovým strategiím. Funkční geny spadající do kategorií, jako jsou přenos signálu (*signal transduction mechanisms*), obranné



mechanismy (*defense mechanisms*) či transport a metabolismus aminokyselin (*amino acid transport and metabolism*), bývají nabohacené u kopiotrofních organismů (Lauro et al., 2009; Leff et al., 2015). Jejich výskyt se rovněž signifikantně odlišuje mezi shotgunovými daty z našich vzorků – můžeme sledovat nižší abundanci v kontrolní zemině ve srovnání s ostatními vzorky ovlivněnými přítomností vegetace.

## 5 Závěry

- Kombinace shotgunových a párových knihoven, jejich sekvenace s využitím GS FLX Titanium (Roche) a assembly doprovázené manuálním zaplněním mezer umožnily přečíst a rekonstruovat kompletní genom bakterie *Achromobacter xylosoxidans* A8. Genom sestává z jednoho chromosomu dlouhého 7 013 095 bp a dvou plasmidů pA81 (98 156 bp) a pA82 (247 895 bp).
- Bioinformatickými metodami jsme identifikovali celkem 6 815 oblastí kódujících geny, 5 620 z nich jsme anotovali na základě podobnosti ke genům ve veřejných databázích. Potvrdili jsme lokalizaci genů pro degradaci chlorbenzoátů na plasmidu pA81. Zároveň jsme identifikovali další geny s potenciálním využitím v biodegradaci polutantů.
- Několik identifikovaných genů poskytuje bakterii *A. xylosoxidans* A8 rezistenci vůči těžkým kovům, což z něho činí vhodného kandidáta pro tvorbu bioremediačního nástroje. Vysoký obsah těžkých kovů často doprovází kontaminaci aromatickými polutanty a negativně ovlivňuje bioremediační potenciál mikrobiálních konsorcií.
- Testovaný postup zahrnující sekvenování shotgunových a párových knihoven jsme dále aplikovali při analýze mikrobiálního společenstva tvořícího krápníkové biofilmy v extrémně kyselém prostředí dolu ve Zlatých Horách. Prvotním 454 amplikonovým sekvenováním úseků 16S rRNA genu jsme stanovili složení této velmi jednoduché bakteriální komunity s dominantním zastoupením dvou bakterií rodu *Ferrovum* a *Acidithiobacillus*.
- Podařilo se nám získat téměř kompletní genomové sekvence bakterií označených jako „*Ferrovum myxofaciens*“ ZH7 a „*Acidithiobacillus*“ linie ZH7B ze smíšeného metagenomického vzorku bez nutnosti předcházející kultivace. Anotace predikovaných genů umožňuje rekonstruovat fyziologické a metabolické vlastnosti jednotlivých organismů. Genomy bakterií „*Ferrovum myxofaciens*“ ZH7 a „*Acidithiobacillus*“ linie ZH7B představují unikátní kombinace genů vyvinuté adaptací na konkrétní ekosystém extrémně kyselých důlních vod ve Zlatých Horách.
- S využitím dat získaných sekvenováním mRNA jsme identifikovali exprimované geny. Mezi nejvíce exprimovanými jsou geny související s oxidací redukovaného železa jako

zdroje energie a geny spojené s adaptací na nízké pH a oxidační stres, kterému jsou bakterie v důlní vodě vystaveny.

- Data velkou měrou připívají k poznatkům o některých prozatím málo zkoumaných acidofilních bakteriích, převážně obtížně kultivovatelných linií rodu *Ferrovum*. Představují důležitý základ pro další analýzy dosud neodhalených metabolických drah, například navržené dráhy oxidace železa.
- Amplikonové a shotgunové sekvenování jsme využili rovněž při výzkumu komplexních mikrobiálních společenstev v půdě kontaminované polychlorovanými bifenyly. Na základě sekvencí 16S rRNA genů a anotovaných shotgunových čtení jsme sestavili taxonomické a funkční profily mikrobiálních komunit v kontaminované zemině ovlivněné výsadbou vybraných druhů rostlin a aplikací minerálního hnojiva.
- Pro bioinformatickou komparativní analýzu rozdílně zastoupených taxonů a funkčních systémů jsme adaptovali metody původně navržené pro analýzu rozdílné míry exprese z RNA-seq dat. To umožnilo odhalit statisticky významný vliv kořenů rostlin na složení a funkční potenciál mikrobiálních společenstev. Od kontrolního vzorku se nejvíce odlišovala populace z oblasti kořenů křenu. Aplikace hnojiva používaná pro podporu růstu rostlin neměla významný vliv na mikrobiální komunity v půdě.
- Tyto informace přispívají k pochopení vztahu rostlin a půdních mikroorganismů a mohou být dále využity pro bioremediační výzkumy.

# 1 Introduction

Physiological and metabolic processes of microorganisms inhabiting every part of this planet greatly affect both biotic and abiotic components of each ecosystem. Methods of molecular biology enable studies on microbial diversity based on analysis of genes encoding processes and biochemical pathways of individual microorganisms and also complete microbial consortia. For this a crucial step was elaboration of new technologies of high-throughput DNA sequencing. These methods made it possible to advance studies of diversity from analysis of genomes of model microorganisms easily cultivated in laboratories to simple communities living in extreme environments and further to complex microbial consortia. Metagenomics, the study of theoretically every genome of given community (Handelsmann et al., 1998), is based on two phenomena: (1) less than 1% of microorganisms could be cultivated under the laboratory conditions, and (2) more than one microorganism is often involved in individual functional pathway in the ecosystem.

The aim of this thesis is to describe the modern sequencing methods and bioinformatic tools and to discuss their applications in bacterial genomics and metagenomics using three of our experimental studies as examples. The presented studies spawn from deciphering one bacterial genome to analyses of taxonomically simple community formed by only a few bacterial members and further to metagenomic profiling of complex soil microbial consortia.

Important are studies on ecosystems negatively affected by human activity where microorganisms not only survive but they can adapt their metabolism to degrade compounds toxic for higher organisms. The first example is a bacterium *Achromobacter xylosoxidans* strain A8, which is able to use toxic chlorobenzoates as sole source of carbon and energy. Sequencing of the genomic DNA of *A. xylosoxidans* A8 resulted in the first complete bacterial genome assembled in the Czech laboratory using only the next-generation platform, which makes it possible to further study the genes involved in biodegradation of chlorobenzoates in the context of the complete genetic background.

An interesting microbial biofilm – macroscopic gelatinous stalactites – was discovered in an extremely acidic environment of the abandoned mine in Zlaté Hory (the Czech Republic). Previous study indicated that a sample of the biofilm was formed by the taxonomically simplest community consisted mostly of two bacteria of genus *Ferroplasma* and *Acidithiobacillus* (Falteisek & Čepička, 2012). Here we present the analysis of the physiological and metabolic characteristics of the predominating members of the acidophilic community based on the deep high-throughput DNA and RNA sequencing.

Soil on the other hand contains numerous microsites and gradients which harbor enormous microbial diversity. Rhizosphere, the thin layer of soil directly influenced by the plant roots, is a particularly active spot for microbial-plant interactions. In polluted soils, rhizosphere interactions are of particular importance and much research have been devoted to rhizodegradation. Previous results have shown that presence of some plant species attenuated polychlorinated biphenyl (PCB) concentrations in legacy contaminated soil (Ionescu et al., 2009; Kurzawová et al., 2012). Based on these studies, we hypothesized that not only do the plants absorb and transform PCBs but also that they alter soil community structure and enrich or stimulate degradative populations. Therefore, we used shotgun metagenomics and 16S rRNA gene amplicon pyrosequencing to determine the effect of cultivated horseradish (*Armoracia rusticana*), black nightshade (*Solanum nigrum*) and tobacco (*Nicotiana tabacum*) on the structure and functional potential of the microbial consortia in legacy contaminated soil.

## 2 Aims of the study

- Sequence and assemble the complete genome of *Achromobacter xylosoxidans* A8 using the GS FLX instrument (Roche). Predict coding sequences (CDS), make a functional annotation of the CDS, identify genes involved in biodegradation of xenobiotics and provide the complete genetic background for the study toward a bioremediation tool.
- Use the tested approach from the step one as a basis for the analysis of the stalactite biofilm sample from Zlaté Hory. In addition, use amplicon sequencing of 16S rRNA for the in depth taxonomical analysis of the biofilm bacterial community. Using the GS FLX (Roche) and MiSeq (Illumina) instruments, sequence the total DNA from the biofilm sample and assemble the continuous sequences covering the genomes of the predominating bacteria. Predict and annotate the CDS and reconstruct the physiological and metabolic processes of the bacteria that evolve in the extremely acidic ecosystem. Sequence mRNA reverse transcribed to cDNA using the MiSeq instrument (Illumina) and identify the expressed genes of the biofilm community from Zlaté Hory.

- Analyze the PCB degrading microbial consortia in legacy contaminated soil using the 16S rRNA gene amplicon sequencing and shotgun sequencing. Using an adopted RNA-seq bioinformatic approach of data analysis, determine the significant differences in the composition and functional potential of the samples influenced by the cultivation of horseradish (*Armoracia rusticana*), black nightshade (*Solanum nigrum*) and tobacco (*Nicotiana tabacum*) and by the application of chemical fertilization.

### 3 Material and methods

#### 3.1 Sequencing of the *Achromobacter xylosoxidans* A8 genome

*Achromobacter xylosoxidans* A8 was grown in the ABC mineral medium at 28 °C, total DNA was isolated with Gentra Puregene Yeast/Bact. A Kit (Qiagen). 454 pyrosequencing shotgun library of approx. 700bp fragments (sequenced from one direction) and paired-end libraries with insert size 3kb and 8kb (both ends sequenced, important for subsequent assembly) were prepared according to the Roche protocols. The libraries were sequenced using the GS FLX Titanium (Roche).

The sequenced reads were assembled in Newbler 2.3 software (Roche), gaps were closed by local assembly of selected reads. The complete sequence was analyzed with Critica (Badger & Olsen, 1999), Glimmer (Delcher et al., 1999), and Prodigal (Hyatt et al., 2010) for the protein-coding genes, tRNAscan (Lowe & Eddy, 1997) and Aragorn (Laslett & Canback, 2004) for the tRNA and tmRNA genes, and RNAmmer (Lagessen et al., 2007) for the rRNA genes. The functions of the predicted protein-coding genes were annotated by comparing our annotation – based on similarity search with BLASTP ( $E\text{-value} < 10^{-10}$ ) against the UniRef90 (Tatusov et al., 1999), NCBI-NR (Benson et al., 2010), and KEGG (Kanehisa et al., 2008) databases – with annotation provided by RAST Annotation Server (Aziz et al., 2008) and IGS Annotation Engine (<http://manatee.sourceforge.net>). The annotation results were verified using Artemis (Carver et al., 2008). Start of the replication was identified by the GC skew analysis (Lobry, 1998).

### ***3.2 Metagenomics and metatranscriptomics of the bacterial community from the extremely acidic mine ecosystem from Zlaté Hory***

Sample of the gelatinous stalactite (named ZH7) was collected in the abandoned mine in Zlaté Hory from the depth of 190m under the ground. Previous physicochemical analysis of the surrounding water showed the typical acid mine drainage ecosystem with pH 2.9, temperature of 8.7 °C, very low amount of organic carbon and total mineralization of 5.8 g/L consisting mainly of sulfates, Fe, Al, Mg, and Cu (Falteisek & Čepička, 2012). DNA was isolated with ZR Soil Microbe DNA Kit (Zymo Research), for RNA isolation we homogenized the biofilm sample using TissueLyser instrument (Qiagen) and extracted the total RNA with RNeasy Mini Kit (Qiagen). To enrich the mRNA, we removed rRNA molecules using the RiboZero rRNA Kit (Epicentre/Illumina).

454 shotgun and paired-end (insert size 3 – 8.8kb) libraries were prepared according to protocols from Roche and sequenced using the GS FLX Titanium (Roche). Sequencing library of short DNA fragments was also prepared with TruSeq DNA Sample Prep Kit (Illumina) and sequenced on the MiSeq instrument (Illumina) from both ends (2x250bp). mRNA was reverse transcribed to cDNA sequencing library with NEXTflex RNA-seq Kit (BIOO Scientific) and sequenced on the MiSeq instrument (2x150bp).

For in depth taxonomical analysis, we amplified 16S rRNA gene from total DNA using the primers f8-27 and r372-357 and sequenced the resulting amplicons on the GS FLX Titanium (Roche). Amplicon reads were processed with Mothur 1.28 software (Schloss et al., 2009) allowing the quality control and correction of the reads and grouping the sequences into operational taxonomic units (OTUs) – with 97% similarity cutoff, the OTUs roughly correspond to the present bacterial species.

Assembly of 454 reads were performed using the Newbler 2.8 software (Roche), MiSeq-DNA and MiSeq-RNA reads were used to further filled some of the gaps in GapCloser software (Lua et al., 2012). MiSeq-RNA reads were assembled to transcripts corresponding contigs with software Trinity (Grabherr et al., 2011).

CDS prediction was carried out using Critica (Badger & Olsen, 1999), Glimmer (Delcher et al., 1999), and Prodigal (Hyatt et al., 2010) software, functional annotation was based on the similarity search against NCBI-NR database using BLASTP ( $E\text{-value} < 10^{-10}$ ). We also used annotation server RAST (Aziz et al., 2008) and KAAS – KEGG Automatic Annotation Server (Moriya et al., 2007).

For the binning of the assembled sequences we used our own pipeline including the measurements of sequence composition, assembly characteristics and CDS annotation. The two of the resulting bins were associated with the genomes of the two predominating bacteria.

To analyze gene expression, we mapped the MiSeq-RNA reads to the annotated assembled genomic sequences and count the relative transcriptional activity to identify the most expressed genes.

### ***3.3 Metagenomic profiling of the PCB degrading consortia from legacy contaminated soil influenced by plants and fertilizations***

Soil samples were collected from a landfill of legacy contaminated soil in Lhenice, south Bohemia, Czech Republic and used for in laboratory 6-months cultivation of horseradish (*Armoracia rusticana*, sample id Hor), black nightshade (*Solanum nigrum*, sample Nig) and tobacco (*Nicotiana tabacum*, sample Tob). We prepared non-fertilized and fertilized samples with each plant species. For the fertilized samples (HorF, NigF, TobF), we used Univerzal KH fertilizer (Nohel Garden, Czech Republic). Control soil samples without a cultivated plant were prepared in the same way (samples Ctrl and CtrlF). DNA was isolated with PowerMax Soil DNA Isolation Kit (MoBio Laboratories Inc., USA). Shotgun libraries for pyrosequencing were prepared according to the Roche manual. Amplicon library were prepared using PCR primers f563-577 and r1406-1392 for 16S rRNA gene amplification. Shotgun and amplicon libraries were sequenced using the GS FLX Titanium and Titanium+, respectively (Roche).

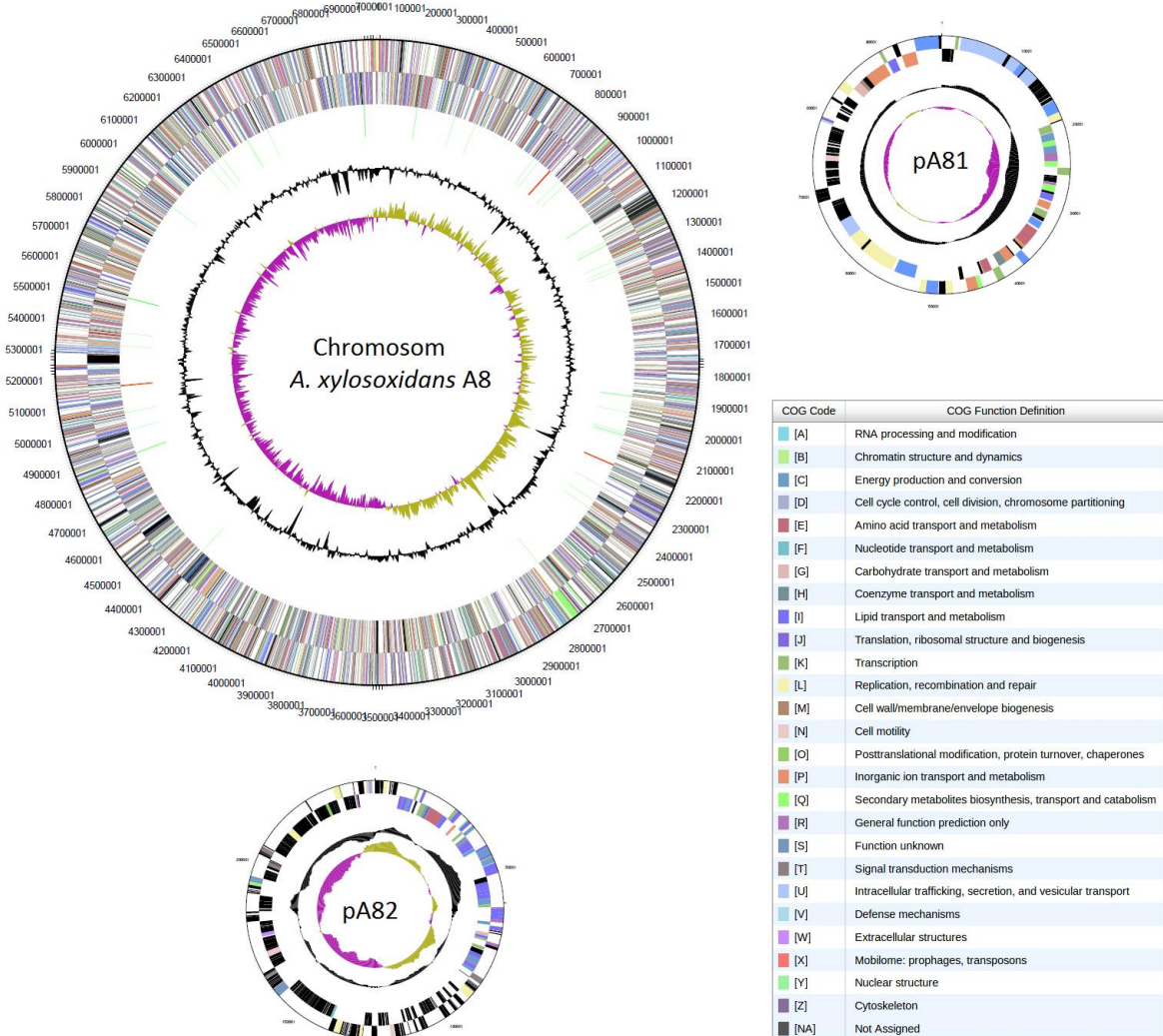
Shotgun reads were annotated using the MG-RAST server (Meyer et al., 2008) and taxonomic and functional abundance profiles were constructed from the list of annotated genes for each sample. Amplicon reads were processed with Mothur 1.28 software (Schloss et al., 2009) and the taxonomic profiles were also constructed based on the counts of classified amplicons. Non-metric multidimensional scaling and vector fitting were performed in vegan package (Oksanen et al., 2013) in R statistical software (R Development Core Team, 2009).

To analyze differences between the samples, we adopted an RNA-seq approach using the DESeq2 package (Love et al., 2014) in R statistical software (R Development Core Team, 2009). Statistically significant differential abundances were plotted in a heatmap graphs in software R.

## 4 Results and discussion

### 4.1 Complete genome sequence of *Achromobacter xylosoxidans* A8

We assembled the complete genome sequence of *A. xylosoxidans* A8 consisting of one chromosome comprising 7,013,095bp and two plasmids: pA81 (98,156bp) and pA82 (247,895bp) (Fig. 1). The chromosome has relatively high GC content (66%), the plasmids have GC contents of 62%. Based on the GC skew analysis, the origin of replication has been localized to the vicinity of the *dnaA* gene.



**Fig. 1:** Graphical representation of the chromosome *A. xylosoxidans* A8 and two plasmids pA81 and pA82. Circular graphs represents (from periphery to the center): (1) genes in (+) orientation (for color codes of functional systems see the legend in the figure), (2) genes in (-) orientation (for color codes of functional systems see the legend in the figure), (3) GC content varying from the mean (calculated in 1000bp window), (4) GC skew (calculated in 200bp window).



We found 6,459 coding sequences (CDS) in the chromosome, 104 CDS in plasmid pA81, and 252 CDS in plasmid pA82. There are 3 rRNA operons, 60 tRNA genes for all 20 amino acids and selenocystein, and one tmRNA gene, all of them located on the chromosome. The coding density of the *A. xylosoxidans* genome is 91.3%. Functions were assigned to 5,620 CDS (82.5%). Seven hundred eighty-three CDS (11.5%) represent genes with some similarity to hypothetical genes in databases. The remaining 412 CDS have no homologues in the databases (*E-value*  $10^{-10}$ ).

We identified the complete *mocpRABCD* gene cluster and the dioxygenase gene for chlorobenzoate degradation and *hyb* operon for salicylate 5-hydroxylase degradation both localized on the plasmid pA81 (Jenčová et al., 2008). We also identified about 70 dioxygenase genes, some of which may be associated with the bioremediation ability of this bacterial strain. Eight genes are probably responsible for the resistance to heavy metals (metal efflux P-type ATPase). High concentration of heavy metals often accompany the contamination with organic compounds and may negatively affect the microbial communities responsible for bioremediation (Roane et al., 2001; Gremion et al., 2004). The function of heavy metal resistance genes predicted in the plasmid pA81 sequence was subsequently confirmed by the expression in *E. coli* by other research group (Hložková et al., 2013). Since the loss of the plasmid does not result in the loss of heavy metal resistance of *A. xylosoxidans* A8 (Hložková et al., 2013), other genes predicted in the chromosome sequence are also involved.

There are 28 genes associated with pathogenesis and 33 genes associated with toxin production and resistance. Actually, the bacterial species *Achromobacter xylosoxidans* was initially isolated from patients with compromised immunity (Holmes et al., 1977). Based on our complete genome sequence of the strain A8, Li et al. (2013) performed a comparative analysis including some other draft genomes of *Achromobacter* strains showing the phylogenetic relationship with the pathogenic bacterial genus *Bordetella*.

## ***4.2 Taxonomical composition and functional processes of the acidophilic***

### ***bacteria forming the stalactite biofilm in Zlaté Hory***

Results of the amplicon 16S rRNA gene sequencing shows the predominating bacteria of genus *Ferrovum* and *Acidithiobacillus* in the biofilm sample from acidic mine in Zlaté Hory (Tab. 1). Relatively more abundant is also an unknown bacterial species (OTU11), followed by a bacteria of genus *Acidiphillium* and other unknown bacterial strains. Acidophilic heterotrophs from the genus *Acidiphillium* were also described in the acid mine drainage waters (Johnson et

al., 2014; Ullrich et al., 2016). Representative sequence of the OTU1 shows 100% similarity with the 16S rRNA gene of *Ferrovum myxofaciens* P3G. In the following text, we will refer to this bacterium associated with OTU1 as to “*Ferrovum myxofaciens*” ZH7. Representative sequence of the OTU2 is 99% similar to the 16S rRNA gene of *Acidithiobacillus ferrooxidans* and hereby we will use the designation “*Acidithiobacillus*” strain ZH7B for this bacterium from Zlaté Hory.

**Tab. 1:** Number of amplicon 16S rRNA reads grouped to OTUs. BLASTN vs. NCBI-NT and simil. [%] represents the best hit in the NCBI-NT and its similarity based on the BLASTN search. RDP classification is based on the similarity search against the RDP database with the bootstrap confidence level (percentages).

OTU	# reads	BLASTN vs. NCBI-NT	simil. [%]	RDP classification
OTU1	1736	<i>Ferrovum</i>	100	Betaproteobacteria;100%; Ferroales;100%; Ferroaceae;100%; <i>Ferrovum</i> ;100%
OTU2	1938	<i>Acidithiobacillus</i>	99	Gammaproteobacteria;100%; Acidithiobacillales;100%; Acidithiobacillaceae;100%; <i>Acidithiobacillus</i> ;100%
OTU3	95	nekultivovaná bakterie	100	Acidobacteria_Gp1;100%; Granulicella;83%
OTU4	103	<i>Acidiphilium</i>	99	Alphaproteobacteria;100%; Rhodospirillales;100%; Acetobacteraceae;100%; <i>Acidiphilium</i> ;75%
OTU5	12	<i>Acidobacteria</i>	99	Acidobacteria_Gp1;100%; Gp1;81%
OTU6	44	<i>Acidocella</i>	100	Alphaproteobacteria;100%; Rhodospirillales;100%; Acetobacteraceae;100%; <i>Acidocella</i> ;92%
OTU7	36	<i>Alicyclobacillus</i>	98	Bacilli;70%; Bacillales;62%; Bacillaceae 1;9%; <i>Domibacillus</i> ;8%
OTU8	4	<i>Leptospirillum</i>	100	Nitrospira;100%; Nitrospirales;100%; Nitrospiraceae;100%; <i>Leptospirillum</i> ;100%
OTU9	3	<i>Sphingomonas</i>	99	Alphaproteobacteria;100%; Sphingomonadales;100%; Sphingomonadaceae;100%; <i>Sphingomonas</i> ;97%
OTU10	28	<i>Thiomonas</i>	100	Betaproteobacteria;100%; Burkholderiales;100%; Burkholderiales_incertae_sedis;100%; <i>Thiomonas</i> ;100%
OTU11	1020	nekultivovaná bakterie	100	Betaproteobacteria;99%; Gallionellales;72%; Gallionellaceae;72%; Sideroxydans;72%
OTU12	107	NA	100	Gammaproteobacteria;85%; Xanthomonadales;85%; Xanthomonadaceae;85%; <i>Dokdonella</i> ;52%
OTU13	65	NA	99	Acidobacteria_Gp1;100%; <i>Terriglobus</i> ;76%

Our sequencing strategy of the metagenomic DNA and the assembly of the reads resulted in nearly complete genomic sequences of the predominating bacteria “*Ferrovum myxofaciens*” ZH7 and “*Acidithiobacillus*” strain ZH7B without prior cultivation. It allowed us to analyze the physiological and metabolic processes. Both bacteria are fully adapted to the acid mine ecosystem, they are chemolithoautotrophic with separated metabolic pathways that are not involved in any form of syntrophy. They both use CO<sub>2</sub> fixation to obtain carbon and oxidation of Fe<sup>2+</sup> as a source of energy. In addition, “*Acidithiobacillus*” strain ZH7B is also able to oxidase sulfur as a source of energy and can use urea as an alternative source of nitrogen. These traits probably enable the predominating bacteria to keep partially separated niches.

In contrast with a few other members of the genus *Ferrovum* based on the published genome sequences, bacterium “*Ferrovum myxofaciens*” ZH7 is able to fix molecular nitrogen, possesses a bacterial immunity system CRISPR/Cas and its cells are “armored” with capsule. All of this features reflects the unique conditions of the given ecosystem.

Genome of the “*Acidithiobacillus*” strain ZH7B shows a mix of features found in bacterial species *At. ferrooxidans* and *At. ferrivorans*. 16S rRNA gene sequence is similar to *At. ferrooxidans*. On the other hand, “*Acidithiobacillus*” strain ZH7B uses genes for Fe<sup>2+</sup>

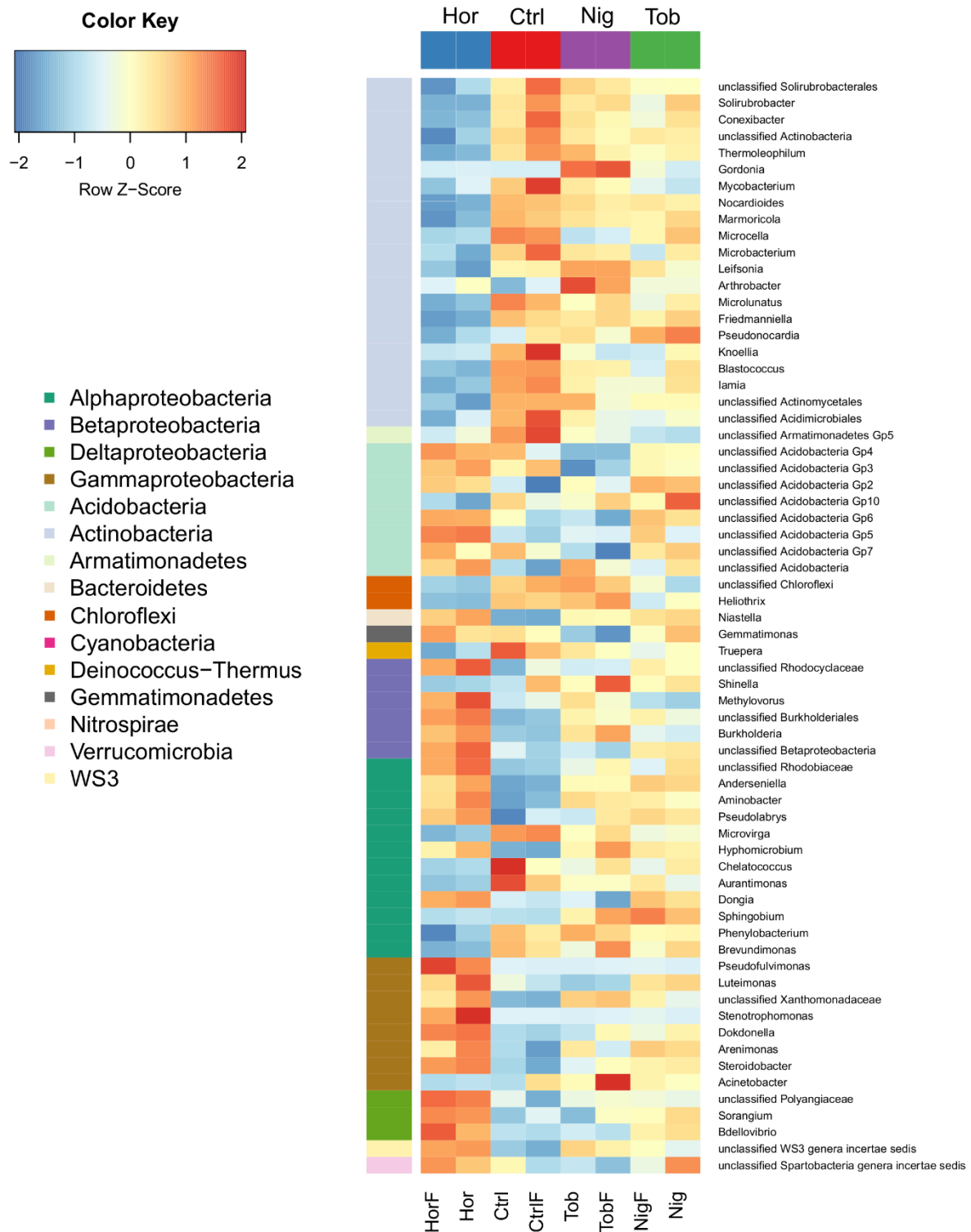
oxidation pathway previously described in *At. ferrivorans* (Hallberg et al., 2010; Amouric et al., 2011). Interestingly, *At. ferrivorans* has been described as highly motile (Hallberg et al., 2010), whereas the genome of “*Acidithiobacillus*” strain ZH7B completely lacks the genes for flagellar assembly. Our data demonstrates the variability inside the bacterial genus *Acidithiobacillus*. Bacterium “*Acidithiobacillus*” strain ZH7B represents a unique strain that had evolved through the adaptations to the given mine environment in Zlaté Hory.

The most expressed genes of the both “*Ferrovum myxofaciens*” ZH7 and “*Acidithiobacillus*” strain ZH7B are involved in Fe<sup>2+</sup> oxidation. Also highly expressed are genes associated with adaptations to the low pH conditions and oxidative stress, which put the bacteria under the constant risk of cellular structures instability and DNA/protein damage. Interestingly, most expressed genes of the other members of the biofilm community (not associated with the two reconstructed predominating genomes) are hypothetical or unknown. This shows the presence of the unknown bacterial taxa and functional processes that are also active in the biofilm.

### ***4.3 Plants rather than mineral fertilization shape microbial community***

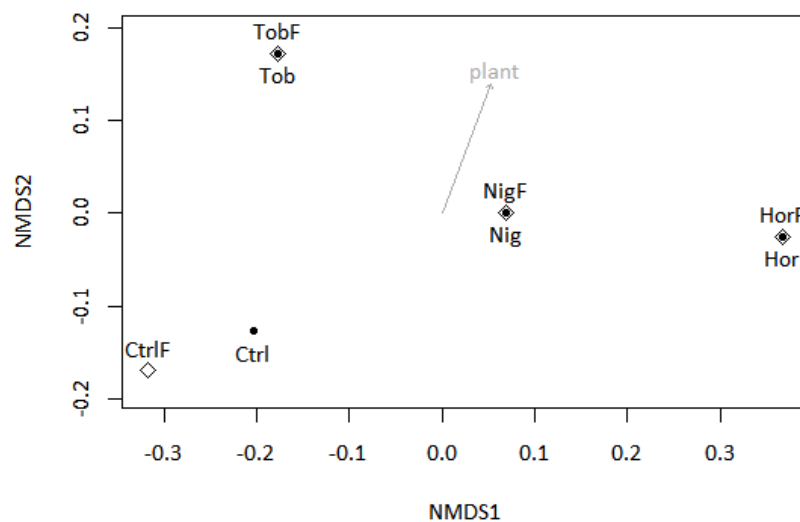
#### ***structure and functional potential in legacy contaminated soil***

We found that reads affiliated with some genera of previously described PCB-degraders were more abundant in vegetated soils versus bulk (control) soil, which may indicate the potential for more efficient degradation (Fig. 2). For instance, planting with horseradish resulted in an increased abundance of *Burkholderia*, which have thoroughly been described as very efficient degraders of PCBs (Mukerjee-Dhar et al., 1998; Tillmann et al., 2005; Chain et al., 2006; Uhlík et al., 2013), *Stenotrophomonas*, which has been isolated from the same soil by growth on biphenyl (Uhlík et al., 2013), or *Methylovorus*, which has been directly implicated in biphenyl metabolism through stable isotope probing in the horseradish rhizosphere (Uhlík et al., 2009). We also determined that horseradish-vegetated soil had the most unique soil microbial community when compared to the control and other two plant samples (Fig. 2). *Arthrobacter* (Gilbert & Crowley, 1997; Abraham et al., 2005; Leigh et al., 2007) and *Gordonia* (Koubek et al., 2012), also often associated with PCB-degradation, were significantly enriched in soils vegetated by tobacco.



**Fig. 2:** Statistically significant differences in the abundance of bacterial taxa based on the amplicon 16S rRNA sequencing. Heatmap shows the abundances at the level of genus (red indicates higher abundance, blue indicates lower abundance). Phylum level is represented in the color bar at the left side of the figure. Non-fertilized samples: Hor (horseradish), Nig (nightshade), Tob (tobacco), Ctrl (control). Fertilized samples: HorF (horseradish), NigF (nightshade), TobF (tobacco), CtrlF (control).

Due to the importance of fertilization for agricultural and environmental sciences, assessment of the effects of chemical fertilization is an active field of study (Geisseler and Scow, 2014), especially in connection with attribution of soil functions to specific microbial populations (Su et al., 2015). Many studies have demonstrated that there are shifts in soil microbial community structure associated with the use of chemical fertilizers (Leff et al., 2015; Su et al., 2015) but the populations are less strongly influenced by the fertilization than plant species (Benizri & Amiaud, 2005; Liliensiek et al., 2012), sometimes indicating the fertilizer does not have a significant effect (Marschner et al., 2001; Liliensiek et al., 2012). Our results also indicate that the response of soil microbial populations to fertilization is not significant, however the presence of plants is (Fig. 2, and 3).



**Fig. 3:** Non-metric multidimensional scaling ordination analysis (NMDS, stress < 0.001) of soil 16S rRNA genes with subsequent fitting of environmental vectors of treatment (plant and fertilizer) onto the ordination ( $P$ -value < 0.05, 40,320 permutations). Symbols: full points – non-fertilized treatments, open diamonds – fertilized treatments. Non-fertilized samples: Hor (horseradish), Nig (nightshade), Tob (tobacco), Ctrl (control). Fertilized samples: HorF (horseradish), NigF (nightshade), TobF (tobacco), CtrlF (control).

Soils tend to select for copiotrophic or oligotrophic microbial populations based on the amount of available organic carbon (Fierer et al., 2007). Therefore, soils directly under the influence of plants tend to favor copiotrophic microorganisms in comparison to populations in bulk soil (Dennis et al., 2010). Although the ecological grow strategies cannot be completely generalized on the level of phylum or class, bacteria affiliated with Actinobacteria, Bacteroidetes, Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria are commonly considered copiotrophic whereas Acidobacteria or Planctomycetes oligotrophic (Fierer et al., 2007; Prober et al., 2015; Lewis et al., 2016b). Increases in the relative

abundances of Alpha-, Beta-, Gammaproteobacteria, or Bacteroidetes (Fig. 2) in vegetated soils also supports the hypothesis that the vegetated soil was enriched for copiotrophic organisms. The trend with Acidobacteria and Planctomycetes, commonly considered oligotrophic, was not that unambiguous. Whereas Planctomycetes were of low abundance throughout the treatments, Acidobacteria were dramatically differently represented, and were enriched especially in the root zone of horseradish. Previously, plant-specific colonization of rhizosphere was proposed for certain acidobacterial lineages (Nunes da Rocha et al., 2013). Our results suggest that horseradish creates favorable environment for Acidobacteria.

Importantly, several functional categories have also been attributed to either copiotrophic or oligotrophic strategies. We observed that some genetic determinants associated with signal transduction mechanisms, defense mechanisms or amino acid transport and metabolism, categories which tend to be overrepresented in copiotrophic organisms (Lauro et al., 2009; Leff et al., 2015), significantly differed among treatments, with significantly lower abundances in the bulk soil compared to vegetated soils.

## 5 Conclusions

- Combination of shotgun and paired-end sequencing using the GS FLX (Roche), and assembly followed by individual gap filling based on the local assembly allowed us to reconstruct the complete genomic sequence of *Achromobacter xylosoxidans* A8. Genome consists of one chromosome of 7,013,095bp and two plasmids: pA81 (98,156bp) and pA82 (247,895bp).
- Using the bioinformatic tools, we identified 6,815 coding sequences (CDS), 5,620 CDS were successfully annotated. We confirmed the localization of the genes involved in chlorobenzoates degradation in the pA81 plasmid, and identified other genes of bioremediation potential.
- Some of the identified genes are involved in the heavy metal resistance, which also makes the bacterium a good starting microorganism for further development toward a bioremediation tool.
- The tested sequencing approach was further used to analyze the acidophilic bacterial community forming the macroscopic biofilm in the acid mine ecosystem

in Zlaté Hory. In addition, the amplicon 16S rRNA gene sequencing revealed the taxonomic composition of this simple community consisting predominantly of two bacteria of genus *Ferrovum* and *Acidithiobacillus*.

- We reconstructed nearly complete genome sequences of the two bacteria named “*Ferrovum myxofaciens*” ZH7 and “*Acidithiobacillus*” strain ZH7B from metagenomic biofilm sample without prior cultivation. Annotation of the sequences allowed us to identify the physiological and metabolic processes. Both bacteria possess a unique combination of features that had evolved in the given extreme ecosystem with  $\text{pH} < 3$ .
- Sequencing of mRNA isolated from the biofilm sample was used to measure gene expression. The most expressed genes of “*Ferrovum myxofaciens*” ZH7 and “*Acidithiobacillus*” strain ZH7B are involved in the oxidation of  $\text{Fe}^{2+}$  as a source of energy and in the adaptations to low pH and oxidative stress.
- Since these acidophilic genera, especially the genus *Ferrovum*, are not yet well studied, our data represents a considerable contribution to the field and can be further used in the study of unknown pathways, for example the hypothesized pathway of  $\text{Fe}^{2+}$  oxidation.
- Amplicon sequencing of 16S rRNA gene and shotgun sequencing of the total DNA was also used to decipher the effects of plants and fertilizer on the microbial consortia in contaminated soil. Adopting the bioinformatical methods previously designed for RNA-seq analyses, we were able to identify statistically significant differences in the soil samples with cultivated horseradish (*Armoracia rusticana*), black nightshade (*Solanum nigrum*) and tobacco (*Nicotiana tabacum*).
- Our results shows significant effect of the cultivated plant species on the composition and functional potential of the microbial community which is able to degrade polychlorinated biphenyls in soil. On the other hand, the fertilization showed no significant effect. This information can help us to better understand the plant-microbe relationships in contaminated environments, therefore allowing us to better understand the complicated dynamics potentially associated with the phyto/rhizoremediation of contaminated areas.

## Použitá literatura / References

- Abraham, WR, Wenderoth, DF, Glasser, W (2005) Diversity of biphenyl degraders in a chlorobenzene polluted aquifer. *Chemosphere* 58, 529-533.
- Amouric A, Brochier-Armanet C, Johnson DB, Bonnefoy V, Hallberg KB (2011) Phylogenetic and genetic variation among Fe(II)-oxidizing acidithiobacilli supports the view that these comprise multiple species with different ferrous iron oxidation pathways. *Microbiology* 157, 111-22.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9, 75.
- Badger JH, Olsen GJ (1999) CRITICA: coding region identification tool invoking comparative analysis. *Molecular Biology and Evolution* 16, 512-24.
- Carver T, Berriman M, Tivey A, Patel C, Böhme U, Barrell BG, Parkhill J, Rajandream MA (2008) Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* 24, 2672-6.
- Delcher AL, Harmon D, Kasif S, White O, Salzberg SL (1999) Improved microbial gene identification with GLIMMER. *Nucleic Acids Research* 27, 4636-41.
- Dennis, PG, Miller, AJ, Hirsch, PR (2010) Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? *FEMS Microbiology Ecology* 72, 313-327.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15-21.
- Faltesek L, Cepička I (2012) Microbiology of diverse acidic and non-acidic microhabitats within a sulfidic ore mine. *Extremophiles* 16, 911-22.
- Fierer, N, Bradford, MA, Jackson, RB (2007) Toward an ecological classification of soil bacteria. *Ecology* 88, 1354-1364.
- Geisseler, D, Scow, KM (2014) Long-term effects of mineral fertilizers on soil microorganisms - A review. *Soil Biology & Biochemistry* 75, 54-63.
- Gilbert, ES, Crowley, DE (1997) Plant compounds that induce polychlorinated biphenyl biodegradation by *Arthrobacter* sp. strain B1B. *Applied and Environmental Microbiology* 63, 1933-1938.
- Gremion F, Chatzinotas A, Kaufmann K, Von Sigler W, Harms H (2004) Impacts of heavy metal contamination and phytoremediation on a microbial community during a twelve-month microcosm experiment. *FEMS Microbiology Ecology* 48, 273-83.
- Hallberg KB, González-Toril E, Johnson DB (2010) *Acidithiobacillus ferrivorans*, sp. nov.; facultatively anaerobic, psychrotolerant iron-, and sulfur-oxidizing acidophiles isolated from metal mine-impacted environments. *Extremophiles* 14, 9-19.
- Hložková K, Šuman J, Strnad H, Ruml T, Pačes V, Kotrba P (2013) Characterization of pbt genes conferring increased Pb<sup>2+</sup> and Cd<sup>2+</sup> tolerance upon *Achromobacter xylosoxidans* A8. *Research in Microbiology* 164, 1009-18.



Holmes B, Snell JJS, Lapage SP (1977) Strains of *Achromobacter xylosoxidans* from clinical material. *Journal of Clinical Pathology* 30, 595-601.

Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11, 119.

Chain, PS, Deneff, VJ, Konstantinidis, KT, Vergez, LM, Agullo, L, Reyes, VL, Hauser, L, Cordova, M, Gomez, L, Gonzalez, M, Land, M, Lao, V, Larimer, F, Lipuma, JJ, Mahenthiralingam, E, Malfatti, SA, Marx, CJ, Parnell, JJ, Ramette, A, Richardson, P, Seeger, M, Smith, D, Spilker, T, Sul, WJ, Tsoi, TV, Ulrich, LE, Zhulin, IB, Tiedje, JM (2006) *Burkholderia xenovorans* LB400 harbors a multi-replicon, 9.73-Mbp genome shaped for versatility. *Proceedings of the National Academy of Sciences of the United States of America* 103, 15280-15287.

Ionescu, M, Beranová, K, Dudková, V, Kochánková, L, Demnerová, K, Macek, T, Macková, M (2009) Isolation and characterization of different plant associated bacteria and their potential to degrade polychlorinated biphenyls. *International Biodeterioration & Biodegradation* 63, 667-672.

Jenčová V, Strnad H, Chodora Z, Ulbrich P, Vlček C, Hickey WJ, Pačes V (2008) Nucleotide sequence, organization and characterization of the (halo)aromatic acid catabolic plasmid pA81 from *Achromobacter xylosoxidans* A8. *Research in Microbiology* 159, 118-27.

Johnson DB, Hallberg KB, Hedrich S (2014) Uncovering a microbial enigma: isolation and characterization of the streamer-generating, iron-oxidizing, acidophilic bacterium "*Ferroplasma myxofaciens*". *Applied and Environmental Microbiology* 80, 672-680.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Research* 36, D480-4.

Koubek, J, Uhlík, O, Ječná, K, Junková, P, Vrkoslavová, J, Lipov, J, Kurzawová, V, Macek, T, Macková, M (2012) Whole-cell MALDI-TOF: Rapid screening method in environmental microbiology. *International Biodeterioration & Biodegradation* 69, 82-86.

Kurzawová, V, Štursa, P, Uhlík, O, Norková, K, Strohalm, M, Lipov, J, Kochánková, L, Macková, M (2012) Plant-microorganism interactions in bioremediation of polychlorinated biphenyl-contaminated soil. *New Biotechnology* 30, 15-22.

Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW (2007) RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research* 35, 3100-8.

Laslett D, Canback B (2004) ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research* 32, 11-6.

Leewis, M-C, Uhlík, O, Leigh, MB (2016b) Synergistic Processing of Biphenyl and Benzoate: Carbon Flow Through the Bacterial Community in Polychlorinated-Biphenyl-Contaminated Soil. *Scientific Reports* 6, 22145.

Leff, JW, Jones, SE, Prober, SM, Barberán, A, Borer, ET, Firn, JL, Harpole, WS, Hobbie, SE, Hofmockel, KS, Knops, JMH, McCulley, RL, La Pierre, K, Risch, AC, Seabloom, EW, Schütz, M, Steenbock, C, Stevens, CJ, Fierer, N (2015) Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. *Proceedings of the National Academy of Sciences* 112, 10967-10972.

Leigh, MB, Pellizari, VH, Uhlík, O, Sutka, R, Rodrigues, J, Ostrom, NE, Zhou, J, Tiedje, JM (2007) Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs) *ISME Journal* 1, 134-148.

- Li X, Hu Y, Gong J, Zhang L, Wang G (2013) Comparative genome characterization of *Achromobacter* members reveals potential genetic determinants facilitating the adaptation to a pathogenic lifestyle. *Applied Microbiology and Biotechnology* 97, 6413-25.
- Liliensiek, A-K, Thakuria, D, Clipson, N (2012) Influences of Plant Species Composition, Fertilisation and *Lolium perenne* Ingression on Soil Microbial Community Structure in Three Irish Grasslands. *Microbial Ecology* 63, 509-521.
- Lobry JR (1996) A simple vectorial representation of DNA sequences for the detection of replication origins in bacteria. *Biochimie* 78, 323-6.
- Love, M, Huber, W, Anders, S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15, 550.
- Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* 25, 955-64.
- Marschner, P, Yang, CH, Lieberei, R, Crowley, DE (2001) Soil and plant specific effects on bacterial community composition in the rhizosphere. *Soil Biology and Biochemistry* 33, 1437-1445.
- Meyer, F, Paarmann, D, D'souza, M, Olson, R, Glass, E, Kubal, M, Paczian, T, Rodriguez, A, Stevens, R, Wilke, A, Wilkening, J, Edwards, R (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9, 386.
- Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research* 35, W182-5.
- Mukherjee, S, Juottonen, H, Siivonen, P, Lloret Quesada, C, Tuomi, P, Pulkkinen, P, Yrjala, K (2014) Spatial patterns of microbial diversity and activity in an aged creosote-contaminated site. *The ISME Journal* 8, 2131-2142.
- Nunes Da Rocha, U, Plugge, CM, George, I, Van Elsas, JD, Van Overbeek, LS (2013) The Rhizosphere Selects for Particular Groups of Acidobacteria and Verrucomicrobia. *PLoS ONE* 8, e82443.
- Oksanen, J, Blanchet, FG, Kindt, R, Legendre, P, Minchin, PR, O'hara, RB, Simpson, GL, Solymos, P, Stevens, MHH, Wagner, H (2013) *vegan: Community Ecology Package* [Online]. Available: <http://CRAN.R-project.org/package=vegan>.
- R Development Core Team (2009) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Roane TM, Josephson KL, Pepper IL (2001) Dual-bioaugmentation strategy to enhance remediation of cocontaminated soil. *Applied and Environmental Microbiology* 67, 3208-15.
- Tatusov RL, Koonin EV, Lipman DJ (1999) A genomic perspective on protein families. *Science* 278, 631-7.
- Su, J-Q, Ding, L-J, Xue, K, Yao, H-Y, Quensen, J, Bai, S-J, Wei, W-X, Wu, J-S, Zhou, J, Tiedje, JM, Zhu, Y-G (2015) Long-term balanced fertilization increases the soil microbial functional diversity in a phosphorus-limited paddy soil. *Molecular Ecology* 24, 136-150.
- Tillmann, S, Strompl, C, Timmis, KN, Abraham, WR (2005) Stable isotope probing reveals the dominant role of Burkholderia species in aerobic degradation of PCBs. *FEMS Microbiology Ecology* 52, 207-217.

Uhlík, O, Ječná, K, Macková, M, Vlček, C, Hroudová, M, Demnerová, K, Pačes, V, Macek, T (2009) Biphenyl-metabolizing bacteria in the rhizosphere of horseradish and bulk soil contaminated by polychlorinated biphenyls as revealed by stable isotope probing. *Applied and Environmental Microbiology* 75, 6471-6477.

Uhlík, O, Musilová, L, Rídl, J, Hroudová, M, Vlček, C, Koubek, J, Holečková, M, Macková, M, Macek, T (2013) Plant secondary metabolite-induced shifts in bacterial community structure and degradative ability in contaminated soil. *Applied microbiology and biotechnology* 97, 9245-9256.

Ullrich SR, Poehlein A, Tischler JS, González C, Ossandon FJ, Daniel R, Holmes DS, Schlömann M, Mühling M (2016) Genome Analysis of the Biotechnologically Relevant Acidophilic Iron Oxidising Strain JA12 Indicates Phylogenetic and Metabolic Diversity within the Novel Genus "Ferrovum". *PLoS One* 11, e0146832.

## Curriculum vitae

Narozen: 10. 2. 1982, Plzeň, Česká republika  
Afilie: Oddělení genomiky a bioinformatiky, Ústav molekulární genetiky AV ČR, v. v. i., Vídeňská 1083, 142 20 Praha 4  
Telefon: +420 241 063 412  
E-mail: jakub.ridl@img.cas.cz

### Vzdělání

2008–dosud doktorské studium, Katedra filosofie a dějin přírodních věd, Univerzita Karlova v Praze, Přírodovědecká fakulta; disertace: *Metagenomické profilování mikrobiálních společenstev*.  
2004–2006 Mgr., Katedra antropologie a genetiky člověka, Univerzita Karlova v Praze, Přírodovědecká fakulta; diplomová práce: *Diversita mitochondriální DNA jemenských populací: příspěvek regionálního vzorkování k interpretaci genetické struktury jihu Arabského poloostrova*.  
2001–2004 Bc., Sociální a kulturní antropologie, Západočeská univerzita, Filozofická fakulta; bakalářská práce: *Využití extrakce aDNA v antropologických výzkumech*.

### Zaměstnání

2007–dosud Výzkumný asistent, Oddělení genomiky a bioinformatiky, Ústav molekulární genetiky AV ČR, v. v. i., Vídeňská 1083, 142 20 Praha 4

### Zahraniční stáže

2013, 2014 GeneCore, EMBL, Heidelberg, Německo

### Zahraniční kurzy

2014 Computational Genomics and Personalized Medicine. Lipari school on bioinformatics and computational biology, Lipari, Itálie  
2010 GS FLX Software Winter School, 2010, Bielefeld, Německo

2009 GS FLX Titanium training, Penzberg, Německo

### **Pedagogická činnost**

2011–2013 přednáška Genetické metody v antropologii v rámci kurzu Metody kosterní biologie Katedry Antropologie, Západočeské Univerzity v Plzni

2007–2010 přednáška Úvod do forenzní genetiky v rámci kurzu Forenzní antropologie Katedry Antropologie, Západočeské Univerzity v Plzni

## Seznam publikací / Selected publications

### *Seznam publikací v impaktovaných časopisech*

**Rídl J**, Kolář M, Strejček M, Strnad H, Štursa P, Pačes J, Macek T, Uhlík O (2016) Plants rather than mineral fertilization shape microbial community. *Front Microbiol* – přijato do tisku  
IF = 3,989 (2015)

Strnad H\*, **Rídl J**\*, Pačes J, Kolář M, Vlček C, Pačes V (2011) Complete genome sequence of the haloaromatic acid-degrading bacterium *Achromobacter xylosoxidans* A8. *J Bacteriol* 193, 791-2.  
\* autoři přispěli k publikaci rovným dílem  
IF = 3,726

Strejček M, Wang Q, **Rídl J**, Uhlík O (2015) Hunting Down Frame Shifts: Ecological Analysis of Diverse Functional Gene Sequences. *Front Microbiol* 6, 1267.  
IF = 3,989

Uhlík O, Strejček M, Vondráček J, Musilová L, **Rídl J**, Lovecká P, Macek T (2014) Bacterial acquisition of hexachlorobenzene-derived carbon in contaminated soil. *Chemosphere* 113, 141-5.  
IF = 3,499

Uhlík O, Wald J, Strejček M, Musilová L, **Rídl J**, Hroudová M, Vlček C, Cardenas E, Macková M, Macek T (2012) Identification of bacteria utilizing biphenyl, benzoate, and naphthalene in long-term contaminated soil. *PLoS One* 7, e40653.  
IF = 4,092

Uhlík O, Musilová L, **Rídl J**, Hroudová M, Vlček C, Koubek J, Holečková M, Macková M, Macek T (2013) Plant secondary metabolite-induced shifts in bacterial community structure and degradative ability in contaminated soil. *Appl Microbiol Biotechnol* 97, 9245-56.  
IF = 3,689

Zubáčová Z, Novák L, Bublíková J, Vacek V, Fousek J, **Rídl J**, Tachezy J, Doležal P, Vlček C, Hampl V (2013) The mitochondrion-like organelle of *Trimastix pyriformis* contains the complete glycine cleavage system. *PLoS One* 8, e55417.  
IF = 3,730

Zákostelská Z, Kverka M, Klimešová K, Rossmann P, Mrázek J, Kopečný J, Hornová M, Srutková D, Hudcovic T, **Rídl J**, Tlaskalová-Hogenová H (2011) Lysate of probiotic *Lactobacillus casei* DN-114 001 ameliorates colitis by strengthening the gut barrier function and changing the gut microenvironment. *PLoS One* 6, e27961.  
IF = 4,411

Klimešová K, Kverka M, Zákostelská Z, Hudcovic T, Hrnčíř T, Štěpánková R, Rossmann P, **Rídl J**, Kostovčík M, Mrázek J, Kopečný J, Kobayashi KS, Tlaskalová-Hogenová H (2013) Altered gut microbiota promotes colitis-associated cancer in IL-1 receptor-associated kinase M-deficient mice. *Inflamm Bowel Dis* 19, 1266-77.

IF = 5,119

Baldrian P, Kolařík M, Stursová M, Kopecký J, Valášková V, Větrovský T, Zifčáková L, Snajdr J, **Rídl J**, Vlček C, Voříšková J (2012) Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. *ISME J* 6, 248-58.  
IF = 3,375

Mořkovský L, Pačes J, **Rídl J**, Reifová R (2015) Scrimmer: designing primers from transcriptome data. *Mol Ecol Resour* 15, 1415-20.  
IF = 3,712

Hroudová M, Vojta P, Strnad H, Krejčík Z, **Rídl J**, Pačes J, Vlček C, Pačes V (2012) Diversity, phylogeny and expression patterns of Pou and Six homeodomain transcription factors in hydrozoan jellyfish *Craspedacusta sowerbyi*. *PLoS One* 7, e36420.  
IF = 4,092

Nývltová E, Stairs CW, Hrdý I, **Rídl J**, Mach J, Pačes J, Roger AJ, Tachezy J (2015) Lateral gene transfer and gene duplication played a key role in the evolution of *Mastigamoeba balamuthi* hydrogenosomes. *Mol Biol Evol.* 2015 Apr;32(4):1039-55.  
IF = 9,105

Pačes J, Huang YT, Pačes V, **Rídl J**, Chang CM (2013) New insight into transcription of human endogenous retroviral elements. *N Biotechnol* 30, 314-8.  
IF = 1,706

## ***Seznam dalších výsledků***

### **Kapitoly v knize**

Rídl J (2012) Sekvence DNA. In: Zvarova J, Mazura I (ed), *Biomedicínská informatika V. Metody molekulární biologie a bioinformatiky*. Nakladatelství Karolinum. Pp: 189-210.

Rídl J, Edens CM, Černý V (2009) Mitochondrial DNA structure of Yemeni population: regional differences and the implications for different migratory contributions. In: Michael D. Petraglia & Jeffrey I. Rose (ed), *The Evolution of Human Populations in Arabia. Paleoenvironments, Prehistory and Genetics. Series: Vertebrate Paleobiology and Paleoanthropology*. Springer. Pp: 69-78.

### **Přednášky na konferencích a plakátová sdělení**

Rídl J, Kolář M, Strejček M, Strnad H, Štursa P, Pačes J, Macek T, Uhlík O (2016) Analýza taxonomicky jednoduché bakteriální komunity z prostředí hlubinného dolu s využitím metagenomiky a bioinformatiky. ENBIK (Národní bioinformatická konference), plakátové sdělení, 3. místo v soutěži o nejlepší plakátové sdělení.

Rídl J, Falteisek L, Hroudová M, Pačes J, Strnad H, Vlček C (2014) Analýza taxonomicky jednoduché bakteriální komunity z prostředí hlubinného dolu s využitím metagenomiky a bioinformatiky. ENBIK (Národní bioinformatická konference), plakátové sdělení.

Rídl J, Pačes J, Stavělová M, Kosinová H, Hroudová M, Fousek J, Vlček C (2011) Metagenomic profiling of bacterial consortia inhabiting environments contaminated by chlorinated ethylenes. 5th European Bioremediation Conference, Crete, plakátové sdělení.

Rídl J, Stavělová M, Kosinová H, Pačes J (2010) Metagenomika mikrobiálních konzorcií v lokalitách kontaminovaných chlorovanými ethyleny. Inovativní sanační technologie ve výzkumu a praxi, Beroun, přednáška.