

**Univerzita Hradec Králové**

**Přírodovědecká fakulta**

**Katedra biologie**

**Struktura a evoluce 5S rRNA genů u ryb**

**Bakalářská práce**

Autor: Kateřina Endrlová  
Studijní program: B1501 / Biologie  
Studijní obor: 1501R015 / Systematická biologie a ekologie  
Vedoucí práce: Mgr. Radka Symonová, Ph.D

Hradec Králové

srpen 2021



## Zadání bakalářské práce

<b>Autor:</b>	<b>Kateřina Endrlová</b>
Studium:	S17BI002BP
Studijní program:	B1501 Biologie
Studijní obor:	Systematická biologie a ekologie
<b>Název bakalářské práce:</b>	<b>Struktura a evoluce 5S rRNA genů u ryb</b>
Název bakalářské práce AJ:	Structure and evolution of 5s rRNA genes in fish

### **Cíl, metody, literatura, předpoklady:**

VŠKP bude spočívat v práci s DNA a RNA sekvencemi ribozomálního genu pro 5S podjednotku. Konkrétně půjde o alignmenty 5S rDNA sekvencí z osekvenovaných druhů ryb, které jsou v současné době k dispozici v online databázi [www.ensembl.org](http://www.ensembl.org). Alignment se bude opírat o data sekundární struktury 5S rRNA. Predikované sekundární struktury budou sloužit ke zmapování evoluce sekvence 5S rDNA a 5S rRNA napříč fylogenetickým stromem ryb.

Vierna J, Wehner S, Höner zu Siederdisen C, Martínez-Lage A, Marz M. 2013. Systematic analysis and evolution of 5S ribosomal DNA in metazoans. *Heredity (Edinb)*. 2013 Nov;111(5):410-21. doi: 10.1038/hdy.2013.63

Castro SI, Hleap JS, Cárdenas H, Blouin C. 2016. Molecular organization of the 5S rDNA gene type II in elasmobranchs. *RNA Biol*. 2016;13(4):391-9. doi: 10.1080/15476286.2015.1100796

Garantující pracoviště:	Katedra biologie, Přírodovědecká fakulta
Vedoucí práce:	Mgr. Radka Symonová, Ph.D.
Datum zadání závěrečné práce:	6.1.2018

Prohlášení:

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně a že jsem v seznamu použité literatury uvedla všechny prameny, ze kterých jsem vycházela.

V Hradci Králové dne

Kateřina Endrlová

## Poděkování

Ráda bych poděkovala Mgr. Radce Symonové, Ph.D. za vedení mé bakalářské práce, odborný dohled, konzultace a za její nesmírnou trpělivost, ochotu a čas, který mi věnovala. Děkuji také RNDr. Aleši Kovaříkovi, CSc. za jeho ochotu, čas a cenné připomínky. Rovněž děkuji svým blízkým za jejich neutuchající podporu.

## **Anotace**

ENDRLOVÁ, K. *Struktura a evoluce 5S rRNA genů u ryb*. Hradec Králové, 2021. Bakalářská práce na Přírodovědecké fakultě Univerzity Hradec Králové. Vedoucí bakalářské práce Mgr. Radka Symonová, Ph.D, 45 s.

Geny pro ribozomální RNA (rRNA) jsou důležitou součástí genomů eukaryot, protože jejich produkty tvoří ribozomy, které jsou zodpovědné za proteosyntézu. 5S rDNA je zhruba 120 bp dlouhá, vysoce konzervovaná sekvence kódující 5S rRNA, jež je součástí velké ribozomální podjednotky eukaryot. Cílem této práce je přinést stručný úvod do problematiky rDNA a systematicky zanalyzovat molekulární strukturu 5S rDNA co nejvyššího počtu druhů ryb, dostupných v současné době ve veřejně přístupných databázích a popsat sekundární strukturu s ní korespondujících 5S rRNA.

### **Klíčová slova**

rDNA, 5S rRNA, sekundární struktura, Actinopterygii

## **Annotation**

ENDRLOVÁ, K. *Structure and evolution of 5S rRNA genes in fish*. Hradec Králové, 2021. Bachelor Thesis at Faculty of Science University of Hradec Králové. Thesis Supervisor Mgr. Radka Symonová, Ph.D, 45 p.

Ribosomal RNA genes are an important part of the genomes of eukaryotes because their products form ribosomes responsible for protein synthesis. The 5S rDNA is an approximately 120 bp long, highly conserved sequence encoding 5S rRNA, which is part of the large ribosomal subunit of eukaryotes. The aim of this work is to provide a brief introduction to the problematics of rDNA, to systematically analyze the molecular structure of 5S rDNA of as many fish species currently available in the publicly accessible databases as possible and to describe the secondary structure of corresponding 5S rRNA.

### **Keywords**

rDNA, 5S rRNA, secondary structure, Actinopterygii

# Obsah

Úvod .....	8
1 Charakteristika rDNA.....	9
1.1 Jaderná ribozomální DNA.....	9
1.2 Mitochondriální rDNA .....	10
1.3 Pseudogeny a orphony .....	10
1.4 Počet kopií rDNA.....	11
1.5 Organizace 45S rDNA.....	13
1.6 Primární sekvence 5S rDNA .....	14
1.7 Sekundární struktura 5S rRNA.....	15
1.8 Interakce 5S rRNA s dalšími molekulami.....	18
1.9 Typy 5S rDNA.....	20
1.10 Evoluce rDNA.....	21
1.11 Metody detekce a způsoby využití rDNA .....	21
2 Materiál a Metody.....	23
2.1 NCBI .....	23
2.2 Ensembl.....	23
2.3 5SRNadb .....	24
2.4 Ensembl BioMart .....	24
2.5 AliView.....	25
2.6 Vienna RNA Websuite.....	25
3 Výsledky.....	26
3.1 Primární sekvence.....	26
3.2 Sekundární struktura.....	27
4 Diskuze .....	32
4.1 Primární sekvence.....	32
4.2 Sekundární struktura.....	32
Závěr.....	35
Seznam použité literatury.....	36
Přílohy .....	44

## Úvod

Geny pro ribozomální RNA (rDNA) jsou důležitou součástí genomů živých organismů, protože jejich produkty tvoří spolu s ribozomálními proteiny ribozomy zodpovědné za proteosyntézu. Živočišná jaderná rDNA je multigenová rodina, skládající se ze dvou hlavních, tandemově uspořádaných jednotek - 5S a 45S rDNA. Nukleolární (jadéřková) 45S rDNA kóduje 18S, 5.8S a 28S rRNA, zatímco extranukleolární (mimojadéřková) 5S rDNA kóduje 5S rRNA (Sochorová et al. 2018). Mitochondriální genom živočichů pak kóduje ještě 12S a 16S rRNA, jež se podílí na tvorbě mitochondriálních ribozomů (Van Etten et al. 1980).

Nedávno bylo zjištěno, že rDNA se vyskytuje ve dvou typech sekvencí. Prvním jsou sekvence přepisované do rRNA, jež jsou vysoce konzervované napříč celým fylogenetickým stromem, což odpovídá úloze, kterou v organismu hrají. Potřeba funkčních ribozomů je všem organismům společná a mutace v této části rDNA mohou být fatální. Druhým typem jsou sekvence nepřepisované do rRNA. Ty nepodléhají tak vysokému selekčnímu tlaku a mohou tedy hromadit mutace, čímž zvyšují variabilitu rDNA. Konzervované úseky se díky snadné rozpoznatelnosti používají k identifikaci pozice klastrů rDNA na chromozomech a ke zjišťování fylogenetických vztahů mezi příbuznými organismy, např. na úrovni druhu. Variabilní úseky se používají ke zjišťování vztahů mezi fylogeneticky vzdálenějšími organismy (Lim et al. 2012).

S ohledem na výše uvedené je důležitým faktorem správného fungování organismu počet kopií rDNA, protože kvůli vysoké potřebě rRNA při výrobě ribozomů buňky potřebují jejich velké množství (Porokhovnik a Gerton 2019). Počet kopií rDNA/rRNA hraje roli v efektivitě proteosyntézy (Hay a Gurdon 1967), v regulaci odpovědi na buněčný stres (Grummt 2013), ve stabilizaci genomu a ochraně před působením mutagenů (Ide et al. 2010). Může ovlivňovat genovou expresi a fenotyp jednice (Paredes et al. 2011), je zdrojem rychlých a vratných změn, jež mohou organismu pomoci se přizpůsobit náhlým změnám prostředí (Gibbons et al. 2014), ovlivňuje buněčnou senescenci (Kobayashi 2014) a jeho změny jsou spjaté s výskytem některých chorob (Xu et al. 2017, Warmerdam a Wolthuis 2019).

Výzkumem 5S rDNA ryb se zabývá mnoho zejména cytogenetických studií. Často však pojednávají o spíše úzkém spektru druhů (např. Symonová et al. 2017; Campo et al. 2009; Nakajima et al. 2012). Systematicky molekulární strukturu 5S rDNA ryb dosud řešilo jen malé množství studií, navíc na malém vzorku druhů ryb a paryb (Castro et al. 2016; Vierna et al. 2013). Tato bakalářská práce má za cíl zpracovat sekvence co nejvyššího počtu druhů ryb, dostupných v současné době v databázích NCBI, Ensembl a 5SRNAdb (NCBI Resource Coordinators et al. 2018; Howe et al. 2021; Szymanski et al. 2016).



# 1 Charakteristika rDNA

## 1.1 Jaderná ribozomální DNA

Jaderná ribozomální DNA (rDNA) je část genomu sestávající z genů kódujících ribozomální RNA (rRNA). Je přítomna v jádrech všech buněk a obsahuje jak konzervované úseky, které se v cytogenetice využívají k detekci těchto oblastí na chromozomech, tak variabilní úseky, díky kterým můžeme zjišťovat fylogenetické vztahy. (Lim et al. 2012)

Názvy jednotlivých rRNA jsou odvozeny od jejich velikosti. Ta je stanovena pomocí sedimentačního koeficientu a udávána ve Svedbergových jednotkách (S). 23S rRNA tedy sedimentuje rychlostí 23S a je větší než 16S rRNA, která sedimentuje rychlostí 16S. (Röhl 1983 cit. dle Sochorové, 2019, s.10)

Živočišná rDNA obsahuje dvě tandemově uspořádané jednotky, které mají proměnlivé množství kopií (od desítek po tisíce). První jednotkou je nukleolární 45S rDNA, která kóduje 18S, 5.8S a 28S rRNA. Druhou jednotkou je extranukleolární 5S rDNA. Díky tomu, že jsou obě tyto jednotky organizovány do opakujících se klastrů, bývají na chromozomech snadno cytogeneticky zobrazitelné. (Sochorová et al. 2018)

U člověka se 45S rDNA nachází na akrocentrických chromozomech 13, 14, 15, 21 a 22, kde tvoří většinu krátkého raménka, zatímco 5S rDNA se nachází pouze na chromozomu 1. Tyto jednotky nejsou fyzicky propojené, nenesou ani žádné homologní segmenty a jsou přepisovány jinými RNA polymerázami – 45S je přepisována RNA polymerázou I, zatímco 5S je přepisována RNA polymerázou III. (Gibbons et al. 2015)

45S rDNA se podílí na vzniku nukleolu (jádérka). Ten se v telofázi formuje okolo NORs („nucleolar organizer regions“), které jsou tvořeny tandemově uspořádanými, opakujícími se úseky 45S rDNA (McStay 2016). V průběhu mitózy nukleolus mizí a NORs můžeme pozorovat jen jako sekundární, často peritelomerická vlákna („cord“) v kondenzovaných chromozomech (Kupriyanova 2000). U některých eukaryot se NORs nachází v částech chromozomů, které v interfázi tvoří perinukleolární konstituční heterochromatin a díky tomu jsou izolovány od genů kódujících proteiny (McStay 2016).

Nukleolus je největší částí buněčného jádra; dochází zde k syntéze a zpracování rRNA a jejímu spojení s ribozomálními proteiny (Németh a Längst 2011). Dále se také podílí na sledování stresových signálů. Pokud je buňka vystavena podmínkám vyvolávajícím stresovou reakci, dojde k utlumení syntézy rRNA a biogeneze ribozomů. Protože se jedná o vysoce energeticky náročné procesy, utlumení transkripce rRNA genů je efektivní způsob, jak může buňka ušetřit energii, kterou

potřebuje k udržení homeostázy. Z toho vyplývá, že nukleolus (tedy i 45S rRNA) hraje klíčovou roli při koordinaci buněčné stresové reakce. (Grummt 2013)

## 1.2 Mitochondriální rDNA

V genomu živočichů se dále nachází mitochondriální 12S a 16S rDNA. 12S rRNA společně s proteiny utváří malou podjednotku, zatímco 16S rRNA se podílí na tvorbě velké podjednotky mitochondriálního ribozomu. (Van Etten et al. 1980)

Uspořádání mitochondriálních rRNA (mt rRNA) genů u myší se nápadně podobá prokaryotickým rRNA operonům a signalizuje, že mt rRNA živočichů nepochází z vlastního jaderného, ale spíše z cizího prokaryotického genomu, což potvrzuje teorii endosymbiózy (Martin et al. 2015). Úseky kódující mt rRNA jsou uspořádány velmi úsporně, blízkost rRNA a tRNA genů naznačuje, že jsou přepisovány jednou RNA polymerázou do polycistronického transkriptu. Z něj je zralá rRNA a tRNA produkována jednoduchým endonukleolytickým štěpením na 5' a 3' hranicích tRNA. (Van Etten et al. 1980)

I členovci mají charakteristickou strukturu mt rRNA, jež může souviset s redukcí mitochondriálního genomu. V tom se navíc vyskytuje daleko větší množství mutací než v jaderném genomu, což je nejspíš způsobeno nedostatečnou rekombinací a korekturou mutací („proofreading“), společně s vysoce oxidativním prostředím. Díky tomu se mitochondriální genom z evolučního hlediska vyvíjí daleko rychleji než jaderný. (Gillespie et al. 2006)

## 1.3 Pseudogeny a orphony

V rámci jaderného genomu lze narazit na další typy DNA – pseudogeny a orphony.

Pseudogeny jsou DNA sekvence, jejichž struktura je podobná funkčním genům, ale které obsahují určité defekty, jako například ztrátu start kodonu, přítomnost dodatečných stop kodónů nebo různé změny v přilehlých kontrolních oblastech. Tyto defekty ve většině případů zabraňují tvorbě náležitých produktů a mají tak za následek ztrátu dané funkce pseudogenu. Ten však v průběhu evoluce může získat funkci novou. (D'Errico 2004)

Pseudogeny se vyskytují jak v prokaryotních, tak v eukaryotních organismech, i když jejich počet i pozice se u různých živočichů značně liší. Důvody těchto odlišností zatím nejsou moc dobře prozkoumané, ale zřejmě svědčí o způsobu, jakým se různé druhy organismů adaptovaly na přežití, a jsou tedy důležitým zdrojem informací pro evoluční i komparativní genomiku. (D'Errico 2004)

Rozlišujeme pseudogeny procesované, neprocesované a mitochondriální (D'Errico 2004). Neprocesované pseudogeny se většinou nachází na stejných chromozomech, uvnitř klastrů nezměněných sekvencí a mohou tedy mít introny a přilehlé kontrolní oblasti stejně jako funkční geny (Harris et al. 1984). Nejčastěji vznikají při

duplikaci genu jako přebytečná kopie, na které se díky její nepotřebnosti mohou hromadit mutace, aniž by došlo k poškození organismu; také však mohou vznikat v důsledku nerovnoměrného dělení při crossing-overu (D'Errico 2004). Typicky obsahují předčasné stop kodony, posuny čtecího rámce, změny v regulačních oblastech a místech střihu; většinou tedy dojde ke ztrátě funkce pseudogenu, jeho delecí, nebo ztrátě veškeré jeho podobnosti s původním genem (Lodish 2000). Duplikovaný pseudogen však stále může být přepisován a pokud získá novou funkci či způsob exprese, může dojít k jeho fixaci v populaci (D'Errico 2004). Neprocesované pseudogeny tedy mohou být jedním z mnoha zdrojů fylogenetické diverzity (Madan Babu 2003).

Procesované pseudogeny vznikají při reverzní transkripci zralé mRNA do DNA, která je následně vložena do jiného chromozomu (Vanin 1985). Nemají tedy protisměrné („upstream“) regulační oblasti ani introny, končí poly-A oblastí na 3' konci a jsou obklopeny přímými repetitivními; mohou to být kompletní i nekompletní přepisy kódujících sekvencí, protože jsou však přepisovány RNA polymerázou II a ne III a neměly by obsahovat promotor, pravděpodobně nebudou schopny aktivace ani v případě, že je jejich kódující úsek neporušen (D'Errico 2004). Díky nedostatku selekčního tlaku mohou obsahovat další vložené sekvence, jako například dlouhé rozptýlené repetice (LINE) nebo *ALU* sekvence (Foord et al. 1996).

Mitochondriální pseudogeny jsou jaderné inserce mtDNA, které neprodukují funkční proteiny kvůli odlišnostem v mitochondriálním genetickém kódu; nejspíše vznikly neúspěšným přesunem některých úseků DNA z mitochondriálního do jaderného genomu (D'Errico 2004). Tento přesun může probíhat buď přímým přenosem DNA, nebo zprostředkovaně prostřednictvím RNA, přičemž přímý přenos se zdá být pravděpodobnější (Woischnik a Moraes 2002). Mitochondriální pseudogeny jsou hojně přítomné v jaderných genomech ptáků a savců, ale v genomech ryb se jejich výskyt zdá být velmi sporý (Bensasson 2001).

Orphony jsou individuálně rozptýlené sekvence, které jsou odvozené od tandemově uspořádaných multigenových rodin (jak těch, které kódují proteiny, tak proteiny nekódujících) a jež se nachází na pro daný druh netypickém místě. Většina orphonů obsahuje pouze jeden kódující úsek a soubor orphonů v genomech různých jedinců v populaci se zdá být odlišný. (Childs et al. 1981)

## 1.4 Počet kopií rDNA

Eukaryotická buňka obsahuje od 1,5 do 3 milionů ribozomů; v průběhu jednoho buněčného cyklu je jich syntetizováno *de novo* až 10 milionů a kvůli vysoké potřebě rRNA při jejich výrobě tedy buňky potřebují velké množství kopií rDNA (Porokhovnik a Gerton 2019). Počet kopií rDNA je mnohdy tak rozsáhlý, že v podstatě tvoří vlastní subgenom, neboli rDNAom (Symonová 2019). Např. gen pro 5S rRNA měl v genomu dánie pruhovaného (*Danio rerio*) 9 081 kopií (Locati et al. 2017) a v genomu štiky obecné (*Esox lucius*) dosahoval přinejmenším 20 000 kopií

(Symonová et al. 2017). Gong et al. (2013) zjistili, že kopie rDNA můžou u některých druhů, např. vířenky (*Vorticella*), dosáhnout počtu i přes 300 000. Poměr rDNAomu ke zbytku genomu ovlivňuje proteosyntetickou aktivitu buněk a tím i fitness jedince; proto je někdy rDNA během diferenciacce buněk (tedy v době intenzivní proteosyntézy) amplifikována, z čehož vyplývá, že její množství není konstantní, ale mění se v průběhu života jedince (Hay a Gurdon 1967). Replikace 5S rDNA může být řízena nezávisle na zbytku genomu; během oogeneze dochází u některých živočichů k její extrémní amplifikaci, zatímco v dalším stádiu počet jejích kopií výrazně klesne (De Fátima Chasko Ribeiro a Aparecida Fernandez 2004).

Rámcový počet kopií rDNA bývá pro daný druh charakteristický. Přestože mnoho těchto kopií je navíc (nejsou přepisovány do rRNA), jsou esenciální pro stabilitu genomu. Ten se totiž po jejich odstranění stává náchylnější k poškození DNA v důsledku působení mutagenů. (Ide et al. 2010)

Počet kopií rDNA se však může i v rámci druhu značně odlišovat. Důsledky proměnlivosti počtu kopií („copy number variation“, CNV) ale ještě nejsou úplně prozkoumány. Zdá se, že např. u octomilky (*Drosophila melanogaster*) ovlivňuje CNV rDNA na chromozomu Y různých jedinců euchromatin i heterochromatin, a je zdrojem rozdílů v genové expresi, jež ovlivňují celý genom. Indukované delece rDNA na chromozomu Y vedly k odlišné expresi stovek až tisíců jiných genů, což mohlo přispět k fenotypovým změnám daných jedinců. Počet kopií rDNA je tedy důležitým determinantem diverzity genové exprese v přirozených populacích a jeho změny mohou pomoci objasnit mechanismy epigenetické dědičnosti a způsob, jakým přispívá k dědivosti komplexních znaků. (Paredes et al. 2011)

rDNA je jednou z nejstarších součástí eukaryotických genomů. Její všudypřítomnost a univerzálnost, patrná v průběhu evoluce eukaryot, naznačuje, že má mnohem více funkcí, než „jen“ syntézu ribozomů (Symonová 2019). CNV rDNA, může být jedním z mechanismů napomáhajících buněčné homeostáze na molekulární úrovni a může představovat také zdroj rychlých a vratných změn genomu, které přispívají k adaptaci jak samotných buněk, tak celých organismů (Gibbons et al. 2014).

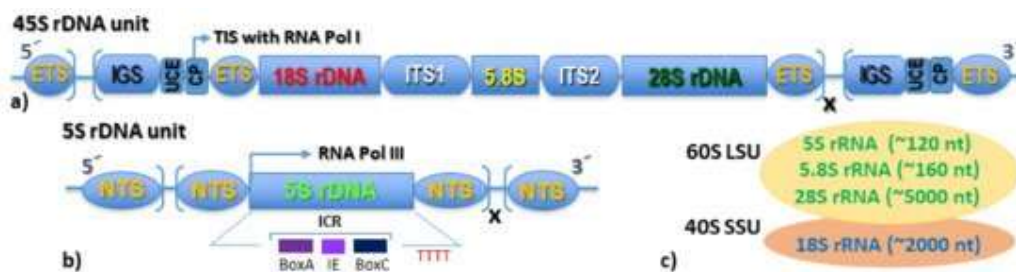
V rámci genomu má rDNA vysoce repetitivní tandemové uspořádání, a to ji činí evolučně dynamickou, ale také náchylnou ke ztrátě kopií v důsledku homologních rekombinací (pokud je jedna kopie poškozena, pokus o její opravu rekombinací s jinou kopií vede ke ztrátě kopií jak v místě poškození, tak v templátovém místě). Počet těchto kopií je sice následně obnoven díky jedinečnému systému amplifikace rRNA genů, avšak tato nestálost může hrát významnou roli nejen v regulaci stability celého genomu, ale může ovlivňovat i životaschopnost buněk. (Kobayashi 2014)

Nestabilita v počtu kopií u člověka souvisí s předčasným stárnutím a neurologickými poruchami u syndromu Louis-Barové a Bloomova syndromu (Warmerdam a Wolhuis 2019). Dřívější výzkumy také poukázaly na menší množství kopií a zvýšenou variabilitu v rDNA genomu buněk postižených rakovinou (Xu et al.

2017). Ukazuje se však, že tyto výsledky mohly být zkresleny neúplností tehdy dostupných dat (Hall et al. 2021).

## 1.5 Organizace 45S rDNA

45S rDNA je v jadérku přepisována RNA polymerázou I na prekurzor pre-rRNA (Obr. 1a, c), skládající se z 18S, 5.8S a 28S rRNA oddělených vnitřními přepisovanými mezerníky (ITS1 a 2), které jsou v průběhu maturace rRNA odstraněny. Tato jednotka je na koncích ohraničena vnějšími přepisovanými mezerníky (ETS). Každá 45S rDNA je od té předchozí oddělena intergenickým mezerníkem (IGS). (Henras et al. 2015)



**Obr. 1:** (a) Organizace 45S rDNA genového klastru (45S rRNA transkripční jednotky); (b) Organizace 5S rDNA jednotky; (c) 80S eukaryotický ribozom složený z velké (60 LSU) a malé (40 SSU) podjednotky. CP – jádrový promotor („core promoter“), ETS – vnější přepisovaný mezerník („external transcribed spacer“), ICR – vnitřní kontrolní oblast („internal control region“), IE – interní element („internal element“), IGS – intergenový mezerník („intergenic spacer“), ITS1, ITS2 – vnitřní přepisovaný mezerník 1 a 2 („internal transcribed spacer 1 and 2“), RNA Pol I a III – RNA polymeráza I a III, LSU – velká podjednotka, SSU – malá podjednotka, nt – nukleotidy, NTS – nepřepisovaný mezerník („non-transcribed spacer“), TIS – iniciační místo transkripce („transcription initiation site“), TTTT – polyT terminační místo transkripce („polyT transcription termination site“), UCE – protisměrný kontrolní element („upstream control element“) (Symonová 2019).

V průběhu evoluce dochází k prodlužování úseků 18S a 28S rDNA a přepisovaných mezerníků, ale obecná struktura přepisované oblasti zůstává stejná. Například savčí 28S rRNA je zhruba o 1,5 kb delší než 28S rRNA kvasinek a skládá se z alternujících konzervovaných a variabilních oblastí. Rozložení těchto oblastí je spojeno se způsobem, jakým se rRNA skládá v ribozomu. Úseky tvořící dvouvláknové útvary sekundární struktury jsou konzervované, zatímco jednovláknové úseky jsou variabilní. Mezi nejvíce konzervované úseky dále patří ty, které interagují s 5S a 5.8S rRNA. (Kupriyanova 2000)

28S rRNA obratlovců obecně obsahuje GC bohaté úseky, které formují stabilní vlásečkové struktury. Jejich mezerníky i variabilní úseky 28S rDNA obsahují jednoduché repetice a mikrosatelity (např. u krysy to jsou motivy: CG, GA, CT, GT, CCT, GGT, GCC, TCG, ACCC, TCCC, GCCC a GAGAC). (Kupriyanova 2000)

Překvapující konzervativnost sekvencí (a univerzálnost sekundárních struktur) některých oblastí 18S a 28S rRNA je způsobena obdobným uspořádáním ribozomů eukaryot – sekundární struktura rRNA patří k nejdůležitějším faktorům normálního fungování ribozomu. Tomu nasvědčuje i fakt, že substituce na jednom z vláken dvouvláknového konzervovaného úseku je ve většině případů kompenzována substitucí na druhém vlákně („compensatory base change“). Dalším důvodem konzervativnosti těchto úseků je jejich pravděpodobná spolupráce v průběhu maturace pre-rRNA. Oblasti kódující zralou rRNA jsou tedy vysoce konzervované ve všech organismech, což reflektuje konzervativnost celého translačního systému. (Kupriyanova 2000)

Sekvence vlastních rRNA genů je u eukaryot tedy vysoce konzervovaná. Naproti tomu IGS se mohou lišit délkou i sekvencí, vždy však obsahují společné znaky. U *Saccharomyces cerevisiae* se jedná o počátek replikace, místo blokující replikační vidlici a promotor řídící syntézu nekódujících transkriptů (Kobayashi 2014). U 45S rDNA savců IGS obsahují regulační oblasti („upstream control element“, UCE), jadrový promotor („core promoter“, CP) a místo ukončení transkripce (Russell a Zomerdijs 2005). UCE je uložen 150 bp před místem počátku transkripce a jeho vzdálenost od CP silně ovlivňuje její účinnost (Kupriyanova 2000). IGS také hrají důležitou roli v umlčování rDNA („rDNA silencing“) (Mayer et al. 2006). V kvasince *S. cerevisiae* (jejíž genom je velmi dobře prozkoumán) je v rámci IGS 45S rDNA obsažena i jednotka 5S rDNA (odděluje IGS1 a IGS2) (James et al. 2009).

Sekvence IGS je velmi variabilní i mezi příbuznými organismy, což naznačuje vyšší evoluční rychlost v porovnání s kódujícími oblastmi. U eukaryot je IGS většinou delší než přepisovaná oblast a v průběhu evoluce se zvětšuje (kvasinky mají 2kb dlouhé IGS, zatímco savci 30kb). Jedinečné oblasti však tvoří jen malou část IGS, za většinu jeho délky (a variability) odpovídají repetice. Ty jsou většinou tvořeny řadou jednotek, skládajících se z úseků od 2-6 bp, tvořících mikrosatelity, po úseky dlouhé několik tisíc bp, tvořících bloky dlouhých repetic. Počet některých repetic umístěných před promotorem ovlivňuje expresi rRNA genů a tím i buněčnou kapacitu proteosyntézy a celkový stav organismu. (Kupriyanova 2000)

## 1.6 Primární sekvence 5S rDNA

5S rDNA je kratší a méně komplexní než 45S rDNA (Obr. 1b, c). Jedná se o vysoce konzervovanou, zhruba 120 bp dlouhou sekvenci, která kóduje 5S rRNA, jež je součástí velké ribozomální podjednotky všech organismů a chybí pouze v mitochondriálních ribozomech hub a živočichů (Smirnov et al. 2008). Přepisovaná část 5S rDNA obsahuje vnitřní kontrolní oblast (ICR), která se skládá ze tří funkčních

elementů: boxu A, vnitřního elementu (IE) a boxu C, je zakončena polyT terminačním místem a z obou stran ohraničena nepřepisovanými mezerníky (NTS); v rámci ICR obsahují IE a box A více mutací než box C (Castro et al. 2016). ICR slouží jako promotor transkripce; box A je vazebným místem RNA polymerázy III a ostatní elementy interagují s TFIIIA (Korn a Brown 1978).

NTS mají vyšší mutační rychlost než kódující oblast (Long a Dawid 1980). Jejich délka i sekvence se u různých druhů liší. Tyto mezidruhové rozdíly jsou způsobeny především insercí, delecí a nukleotidovou substitucí (Campo et al. 2009; Wasko et al. 2001), ale roli hrají i různá uspořádání mikrosatelitů a jejich komplexita (Castro et al. 2016). Obecně jsou ale NTS delší než přepisované oblasti (De Fátima Chasko Ribeiro a Aparecida Fernandez 2004). Živočišná 5S rDNA dále vykazuje vysokou variabilitu v protisměrných („upstream“), vnitřních i posměrných („downstream“) regulačních oblastech a také vzdálenost mezi kódující oblastí a terminačním signálem je variabilní; navíc zde existují minimálně tři typy terminačních signálů (Vierna et al. 2013).

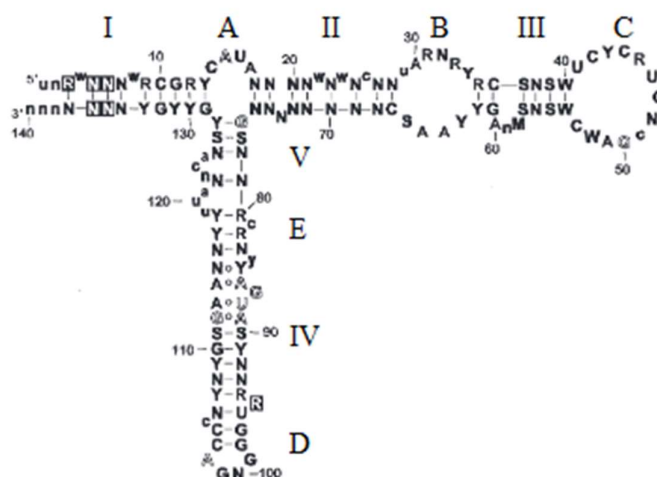
5S rRNA obsahuje velké množství nekanonických párů bází a přestože se její primární nukleotidová sekvence může lišit i v rámci jednoho druhu, má vysoce konzervovanou sekundární a terciární strukturu (Smirnov et al. 2008). Zobrazení její konsenzuální sekundární struktury může být využito k rozlišení mezi rRNA geny a pseudogeny, jež se v eukaryotních genomech často vyskytují (Barciszewska et al. 2001).

## 1.7 Sekundární struktura 5S rRNA

Sekundární struktura rRNA označuje lokální geometrické uspořádání jejího nukleotidového řetězce do různých druhů smyček a helixů, vytvořené vznikem vodíkových můstků (Lodish 2000).

Sekundární struktura 5S rRNA (Obr. 2) se skládá z pěti helixů (I – V), dvou vlásenkových smyček (C, D), dvou vnitřních smyček (B, E) a kloubového regionu A (též smyčky A), uspořádaných do tří helixových ramen (Barciszewska et al. 2001).

Tyto části následně tvoří tři domény. Helix I tvoří doménu  $\alpha$ , helixy II a III se smyčkami B a C tvoří doménu  $\beta$  a helixy IV a V se smyčkami D a E tvoří doménu  $\gamma$ . Individuální domény jsou více či méně nezávislé strukturální útvary, které jsou do jednotného celku spojeny pomocí smyčky A. (Smirnov et al. 2008).



**Obr. 2:** Model obecné eukaryotní sekundární struktury 5S rRNA, I až V – helixy, A – kloubový region, B až D – smyčky. Nukleotidy jsou označeny podle kódu IUPAC (Tab. 1). Malá písmena značí inserce a písmena v boxech značí delece. Upraveno dle Barciszewska et al. (2001).

IUPAC kód	Báze
A	Adenin
C	Cytosin
G	Guanin
T (U)	Thymin (Uracil)
R	A nebo G
Y	C nebo T
S	G nebo C
W	A nebo T
K	G nebo T
M	A nebo C
B	C nebo G nebo T
D	A nebo G nebo T
H	A nebo C nebo T
V	A nebo C nebo G
N	Jakákoli báze
. nebo -	mezera

**Tab. 1:** Tabulka mezinárodního nukleotidového kódu IUPAC, upraveno dle Cornish-Bowden (1985).

Helix I je většinou velmi stabilní, a to především díky vyššímu obsahu G-C párů, také je specifický zvýšeným počtem G-U párů v porovnání s jinými částmi molekuly (Smirnov et al. 2008). G-U patří mezi tzv. kolísavé páry bází („Wobble base pair“), tj. páry dvou nukleotidů, které se neřídí pravidly Watson-Crick párování, podle



kterých se guanin (G) páruje výhradně s cytosinem (C) a adenin (A) s thyminem (T) v DNA a uracilem (U) v RNA (Reece a Campbell 2011).

Na jedné straně helixu I se nachází 5' a 3' konce 5S rRNA, které jsou často nespárované, případně se liší délkou (5' konec bývá uložen hlouběji, resp. 3' konec přesahuje). Na druhé straně na helix I navazuje smyčka A. Díky lokálním deformacím helixu (které jsou výsledkem neefektivního stohování kvůli přítomnosti nekompensovaných G-U párů) je tento úsek pravděpodobně velmi důležitý pro RNA-protein rozpoznávání. Dále je jeho specifická sekundární a terciární struktura (tj. prostorové uspořádání podmíněné interakcemi mezi vedlejšími skupinami řetězce) zcela zásadní pro stabilizaci 5S rRNA a pro její začlenění do ribozomu. Protože stabilita Helixu I ovlivňuje strukturu smyčky A, která se přímo účastní interakce s transkripčním faktorem TFIIIA, je neporušenost této domény velmi důležitá. (Smirnov et al. 2008)

Helix II obsahuje pouze kanonické Watson-Crick páry; nukleotidová sekvence v této oblasti je velmi variabilní, ale sekundární struktura zůstává beze změn (Smirnov et al. 2008). Helix II je u většiny eukaryot dlouhý 8 bp a obsahuje drobnou smyčku tvořenou zbytkovým nespárovaným nukleotidem (Delihás a Andersen 1982). Ta podporuje lokální změnu konformace, která má za následek odkrytí vazebného místa pro proteiny. (Smirnov et al. 2008)

Helix III má velmi konzervativní primární sekvenci i sekundární strukturu. To souvisí s jeho úlohou v interakci s L5 proteinem. Vznik vazebného místa zde usnadňuje přítomnost dvou (případně tří) nespárovaných nukleotidů, které uprostřed tohoto helixu tvoří krátkou smyčku. Tato oblast má též zvýšenou schopnost konformační mobility (tj. schopnost změny tvaru molekuly, často indukovaná vnějšími faktory). To je důležité pro její integraci do ribozomu, protože reorganizace struktury helixu III jsou také spojeny s vazbou na protein L5. (Smirnov et al. 2008)

Helix IV je charakteristický zvýšeným počtem výskytu nekanonických párů bází. Má zvýšenou konformační mobilitu a spolu se smyčkou D je zodpovědný za integraci 5S rRNA do velké ribozomální podjednotky. (Smirnov et al. 2008)

Jednovláknové oblasti 5S rRNA jsou velmi kompaktní a dobře strukturované prvky se specifickou konformací a stabilitou, která je zajišťována nekanonickými interakcemi (které charakterizují především smyčky A, B a C). Smyčka A hraje centrální roli v terciárním uspořádání 5S rRNA. Její nukleotidová sekvence je vysoce konzervativní a mutace v této oblasti většinou vedou k vážným změnám v prostorové (terciární) struktuře 5S rRNA a k poruchám jejích interakcí s proteiny. Smyčka B pomáhá udržovat nenarušené stohování bází, čímž snižuje vnitřní mobilitu domény  $\beta$ . Zdá se však, že nemá žádnou jinou než čistě strukturní funkci. I když je jedním z úseků chráněných proteinem L5 (po jeho navázání), nemá žádnou úlohu ve formování jeho vazebného místa. Naopak smyčka C hraje v interakci

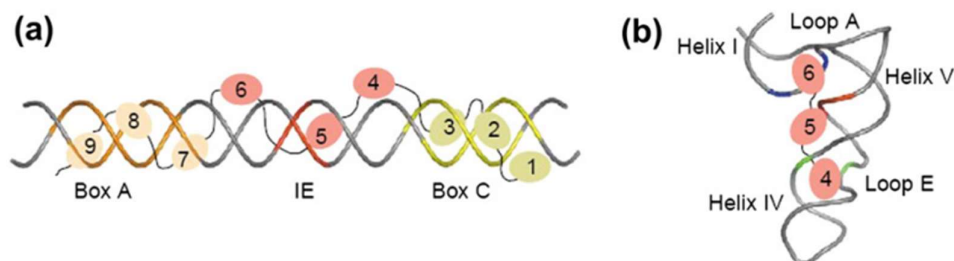
s tímto proteinem významnou úlohu. Smyčka E interaguje s TFIIIA. (Smirnov et al. 2008)

Castro et al. (2016) ovšem kromě této typické sekundární struktury 5S rRNA, kterou klasifikovali jako variantu  $\alpha$ , uvádí ještě další dva typy sekundárních struktur 5S rRNA ( $\beta$  a  $\gamma$ ), které mají nikoli 3, ale 4 ramena.

## 1.8 Interakce 5S rRNA s dalšími molekulami

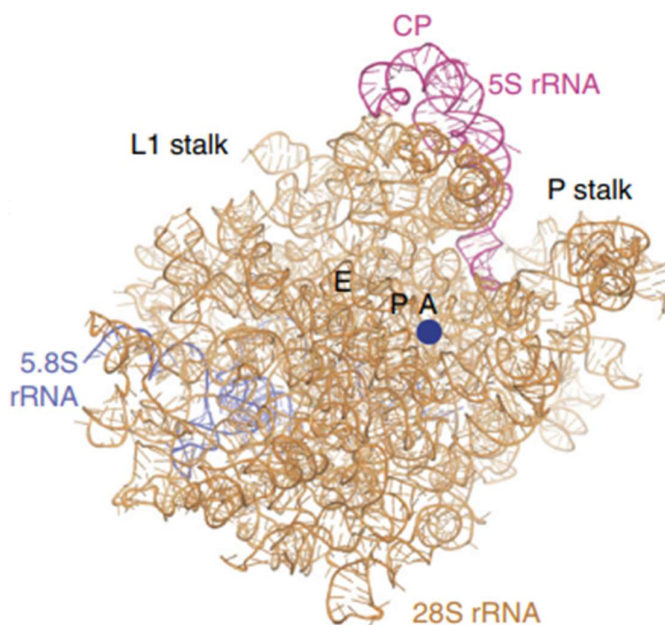
Podobně jako ostatní buněčné RNA, 5S rRNA je vždy vázaná na některý z proteinových faktorů. V případě eukaryotní 5S rRNA se jedná o ribozomální protein L5 nebo transkripční faktor TFIIIA. Proteinová složka komplexu rozhoduje nejen o stabilitě, ale i o směru přenosu 5S rRNA v rámci buňky. Její transportní cesta je velmi složitá, jelikož její transkripce neprobíhá v nukleolu a musí být tedy na místo ribozomální konstrukce dopravena. (Smirnov et al. 2008).

Jak již bylo zmíněno, 5S rRNA je přepisována pomocí RNA polymerázy III, spolu s ní se však transkripce účastní i faktory TFIIIA (Obr. 3a), TFIIIB a TFIIIC; nově syntetizovaná 5S rRNA se poté naváže na TFIIIA, který ji stabilizuje do doby než je potřebná k syntéze ribozomu (Smirnov et al. 2008). Tři centrální zinkové prsty (ZIF 4-6) z celkových devíti prstů TFIIIA se vážou na fragment 55 nukleotidů složený ze smyček A a E a helixů IV a V (Obr. 3b); ZIF 7, který výrazně přispívá k afinitě a specificitě interakce mezi těmito dvěma molekulami, se váže na helix II (Neely et al. 1999). TFIIIA dále chrání smyčky D a E a části helixů I a II před štěpením nukleázami. Chrání tedy 5S rRNA před degradací, ale zároveň nenarušuje její rozpoznávání proteinem L5, protože jeho vazebné místo se nachází na helixu III a smyčce C, jež jsou nutné k integraci do ribozomu (Smirnov et al. 2008).



**Obr. 3:** (a) TFIIIA se váže na 5S rDNA pomocí zinkových prstů 1-3 (rozeznávají box C; vyznačeny žlutě), prstu 5 (rozeznává interní element IE; vyznačen červeně) a prstů 7-9 (rozeznávají box A; vyznačeny oranžově). Zinkové prsty 4 a 6 (vyznačeny červeně) mají roli mezerníků nevázaných se na DNA a pomáhají rozlišit výše zmíněné elementy. (b) Na 5S rRNA se TFIIIA váže hlavně díky zinkovým prstům 4-6. Prst 4 se váže na smyčku E (zeleně), prst 5 se váže na helix V (červeně) a prst 6 se váže na smyčku A (modře). (Hall 2005).

Protein L5 je poměrně špatně strukturovaná molekula, po navázání na 5S rRNA však změní konformaci a dojde k jeho stabilizaci (ke stejné reakci dojde i u 5S rRNA, ovlivnění je tudíž vzájemné); obě makromolekuly jsou tedy částečně nestrukturovanými chaperony zodpovědnými za vzájemné správné prostorové uspořádání a výslednou funkčně aktivní konformaci (Tompá a Csermely 2004). Protein L5 pomáhá transportovat 5S rRNA zpět do jádra a integrovat ji do velké podjednotky ribozomu (Obr. 4), kde tento komplex tvoří většinu centrálního výčnělku (Smirnov et al. 2008).



**Obr. 4:** Pozice 5S (fialově), 5.8S (modřě) a 28S (hnědě) rRNA v rámci velké ribozomální podjednotky; Aminoacylové místo (A), je místem, na které se naváže tRNA nesoucí aminokyselinu. Aminokyseliny se spojují v řetěz polypeptidů v centru peptidyltransferázy (modrý bod). tRNA, na které je navázán polypeptidový řetězec, se posune na peptidylové místo (P), zatímco deacetylovaná tRNA se posune na místo výstupu (E), odkud je následně z ribozomu uvolněna. CP značí centrální výčnělek velké ribozomální podjednotky. (Lafontaine 2015).

Eukaryotní ribozom (80S) je makromolekulární útvar zodpovědný za proteosyntézu, vyskytující se ve všech eukaryotních organismech. Tvoří ho dvě podjednotky (Obr. 1c), složené z ribozomálních RNA a proteinů. Malá podjednotka (40S, jejíž součástí je 18S rRNA) je místem kodon-antikodon interakcí mezi mRNA a tRNA, zatímco velká podjednotka (60S, jejíž součástí jsou 5S, 5.8S a 28S rRNA) katalyzuje vznik peptidových vazeb. (Barciszewska et al. 2001)

Díky mnoha RNA-protein a RNA-RNA interakcím slouží 5S rRNA jako spojovací článek mezi funkčními centry ribozomu a pomáhá koordinovat jejich činnost (Kiparisov et al. 2005). 5S rRNA není s těmito centry v přímém kontaktu, nachází se

však na zvláštní pozici, která jí umožňuje roli prostředníka (Dinman 2005). Mezi tato centra patří A- a P- místo, centrum peptidyltransferázy a GTPázy; dále pomáhá propojit velkou a malou podjednotku (Smirnov et al. 2008). Navíc 5S rRNA stabilizuje ribozom a zvyšuje aktivitu peptidyltransferázy (Barciszewska et al. 2001).

Kiparisov et al. (2005) také po analýze uměle zanesených mutací a alelických variant 5S rRNA zjistili, že variabilita této RNA může sama o sobě sloužit jako regulační mechanismus genové exprese. Různé alely 5S rRNA jsou zodpovědné za různou přesnost translace, což může být podle autorů buňkou využito pro drobné změny v syntetizovaném proteinu.

## 1.9 Typy 5S rDNA

U nižších obratlovců se vyskytují dva typy 5S rDNA specifické pro určité tkáně a stupně individuálního vývoje. Jedná se o somatický a oocytový typ. Oocytový typ tvoří 120bp dlouhá 5S rDNA obklopená několik stovek bp dlouhou sekvencí bohatou na AT, zatímco somatický typ tvoří podobná, stejně dlouhá 5S rDNA, která je však obklopena sekvencí bohatou na GC. (Tomaszewski 1997)

Oocytový typ je během embryogeneze deaktivován kvůli nestabilitě jeho transkripčního komplexu (Wolffe a Brown 1988). U ryb dochází během vývinu k úplné ztrátě tohoto typu, zatímco u žab zůstává částečně zachován (Locati et al. 2017; Peterson et al. 1980).

U ryb a paryb se dále vyskytují jiné dva typy 5S rDNA (typ I a II), které spolu koexistují ve všech tkáních a liší se délkou a sekvencí NTS oblastí (Castro et al. 2016). U lososa obecného (*Salmo salar*) se tyto dva typy lišily nejen délkou NTS, ale i třemi bázemi v úseku kódujícím 5S rRNA (Pendas et al. 1994). U síhů (*Coregonus sp.*) byly rozdíly v délce dvou typů NTS způsobeny přítomností 130 bp dlouhé přímé repetice; rozdíly v úsecích kódujících 5S rRNA byly způsobeny dvěma typy sekvencí boxu C; první typ (A) obsahoval nukleotidy 83A a 85G, druhý typ (B) obsahoval nukleotidy 83G a 85A (Sajdak et al. 1998). U štikozubců (*Merluccius sp.*) byly detekovány dva typy 5S rDNA, nazvané typ A a M, které se lišily přítomností či absencí jednoduché repetitivní sekvence v rámci NTS, některé druhy obsahovaly dvě varianty typu M (S a L), které se lišily délkou; úseky kódující 5S rRNA se lišily na pozicích 3 a 25 (Campo et al. 2009). Všechny výše zmíněné práce podporují hypotézu duálního systému 5S rRNA u ryb.

Savčí genom rovněž obsahuje dva typy 5S rDNA. První typ je vysoce konzervativní, vyskytuje se ve všech savčích genomech a neobsahuje téměř žádné mutace, zatímco druhý typ je vysoce variabilní. (Vierna et al. 2013)

## 1.10 Evoluce rDNA

Evoluce rDNA je ovlivňována dvěma evolučními procesy na molekulární úrovni: koncertovanou evolucí („concerted evolution“) a evolucí birth-and-death. Dlouhodobá evoluce rDNA je charakteristická vysokým selekčním tlakem na konzervativní části sekvencí, zatímco genová diverzifikace flexibilních částí může poskytovat adaptivní potenciál pro případ změn v životním prostředí organismů. Na jedné straně tedy silná purifikační selekce a další mechanismy koncertované evoluce zachovávají integritu konzervativních sekvencí, zatímco birth-and-death procesy pomáhají generovat nové variace genů díky duplikaci rDNA. (Vierna et al. 2013)

Evoluční model birth-and-death tedy uvádí jako zdroj variability rRNA genů jejich duplikaci, po které následně některé kopie zůstávají funkční, zatímco z ostatních se postupně díky nahromaděným mutacím stávají pseudogeny. Tento model také zdůrazňuje purifikační selekci, která napomáhá konzervativnosti kódující sekvence. Jejím výsledkem je malá variabilita kódující oblasti genů i u nepříbuzných druhů, nepřepisované oblasti jsou však velmi variabilní i u druhů příbuzných (díky chybějícímu selekčnímu tlaku). (Nei a Rooney 2005)

Koncertovaná evoluce funguje na bázi genové konverze a asymetrického crossing-overu. Podle tohoto modelu se geny v jedné genové rodině vyvíjejí současně (nikoli každá jednotka zvlášť). Vysvětluje především proces homogenizace DNA, jelikož variabilita způsobená mutacemi se díky jejím procesům buď rozšíří a fixuje, nebo zaniká. (Nei a Rooney 2005)

Genová konverze je lokální jednosměrný přenos úseku DNA z donorového do recipientního genu. Tato interakce nezpůsobuje změny v přilehlých oblastech DNA. Jejím prostřednictvím vznikají dva produkty, z nichž jeden je identický s původní molekulou, zatímco druhý se od ní odlišuje (jedna z alel se přemění podle vzoru druhé alely). Tento mechanismus může též sloužit k opravě genů. (Kupriyanova 2000)

Při asymetrickém crossing-overu vznikají z úseku DNA s mutací dva typy nových úseků DNA. Jeden, jež danou mutaci obsahuje dvakrát a druhý, ve kterém daná mutace chybí. Po mnoha asymetrických crossing-overech se mutace buď rozšíří, nebo ze všech úseků vymizí. (Eickbush a Eickbush 2007)

## 1.11 Metody detekce a způsoby využití rDNA

Genomová organizace rDNA je snadno identifikovatelná díky štěpení pomocí restričních enzymů, a to především díky své repetitivní struktuře. Restriční fragment lze identifikovat například molekulární hybridizací. (Kobayashi 2014)

Lokusy rDNA na chromosomech mohou být zjištěny pomocí *in situ* hybridizace s použitím fluorescenčních barviv, nebo radioaktivního značení. Označená sonda se

přichytí na komplementární sekvenci a následně může být zobrazena. Při radioaktivním značení se používá nejčastěji tritium, při fluorescenčním značení se používá buď molekula reportéru s navázanou fluorescenční molekulou, nebo fluorochrom. (Phillips a Reed 1996)

45S rDNA ve formě NORs aktivních v předchozí interfázi, je možno detekovat pomocí barvení stříbrem (Howell a Black 1980). U nižších obratlovců je také možno využít GC-specifické fluorochromy, jež dokáží označit i neaktivní NORs díky jejich obohacení o GC nukleotidy na GC homogenním a chudším pozadí (Schmid 1980).

rDNA lze také zobrazit *in silico* přístupem vizualizací genomové polohy jejích lokusů pomocí nástroje BioMart, díky kterému lze zmapovat rDNA na vazebných skupinách. Jedná se o velmi citlivou metodu, je možné detekovat i jedinou molekulu rDNA, ale její přesnost závisí na kvalitě genomových dat, jež jsou pro daný druh k dispozici. (Symonová 2019)

rDNA je všudypřítomná a její sekvence jsou poměrně dobře konzervované, využívá se tedy ve fylogenetických studiích k určení míry odlišnosti druhů (např. Stock a Whitt 1992). Dále se využívá v cytogenetických studiích (např. Nakajima et al. 2012). 5S rRNA se díky své malé velikosti a bohatosti na nekanonické párování bází využívá jako model pro zkoumání RNA struktury, konformací a RNA-protein interakcí (Barciszewska et al. 2001).

## 2 Materiál a Metody

Sekvence 5S rDNA využitá ke zpracování této bakalářské práce pochází z následujících volně přístupných online databází.

### 2.1 NCBI

National Center for Biotechnology Information (NCBI), neboli Národní centrum pro biotechnologické informace, je součástí Národní lékařské knihovny USA (NLM) spadající pod Národní zdravotní instituty (NIH). Bylo založeno v roce 1988 za účelem vývoje informačních systémů s využitím v molekulární biologii. Spravuje velké množství databází z šesti okruhů – literatura, zdraví, genomy, geny, proteiny a chemické látky. Pro každou z těchto kategorií dále poskytuje relevantní nástroje, umožňující snadnější manipulaci s daty obsaženými v jednotlivých databázích. Data získává NCBI díky mezinárodní spolupráci s dalšími poskytovateli databází (např. s DNA Data Bank of Japan, DDBJ) a také díky příspěvkům vědecké komunity. Sekvence DNA a RNA jsou uloženy konkrétně v databázi „Nucleotide“ (dostupné z: <https://www.ncbi.nlm.nih.gov/nucleotide>). (NCBI Resource Coordinators et al. 2018)

NCBI také spravuje systém Entrez, který poskytuje možnost vyhledávat a získávat data z více než 30 databází, a to prostřednictvím webového rozhraní, i pomocí rozhraní pro programování aplikací („Application Programming Interface“, API). Entrez byl navržen tak, aby zjednodušoval propojení mezi nejrůznějšími typy biologických dat a usnadnil tak výzkum. (Schuler et al. 1996)

### 2.2 Ensembl

Projekt Ensembl (dostupný z: <http://www.ensembl.org/index.html>) byl započat v roce 1999 a spadá pod Evropský institut bioinformatiky při Evropské laboratoři pro molekulární biologii (EMBL-EBI) sídlící ve Velké Británii. Jeho úkolem je automatická anotace genomů (snaha o lokalizaci významných oblastí DNA jako např. genů a míst jejich interakcí s proteiny v rámci genomu), integrace získaných anotací s dalšími dostupnými biologickými daty a zveřejňování výsledků prostřednictvím webových stránek. Od jejich spuštění v roce 2000 již bylo do této databáze přidáno více než 50 000 genomů. DNA sekvence jsou Ensemblu poskytovány projekty z celého světa. (Howe et al. 2021)

Ensembl také poskytuje nástroje usnadňující zpracování genomických dat, jako například BLAST („Basic Local Alignment Search Tool“) (Altschul et al. 1990), VEP („Variant Effect Predictor“) (McLaren et al. 2016) a mnohé další, a to jak prostřednictvím webového rozhraní, tak API.

## 2.3 5SRNAdb

5SRNAdb (dostupná z: <http://combio.pl/rrna/>) je databáze zaměřená na 5S rDNA a 5S rRNA. Obsahuje více než 11 000 sekvencí pro více než 7 000 druhů. Sekvence pro eukaryota autoři získali z genomových assembly, které jsou k dispozici na NCBI, jedná se tedy o sekundární databázi. Její výhodou je přehlednost a snadná dostupnost již upravených sekvencí. 5SRNAdb umožňuje snadno vyhledávat a stahovat jednotlivé sekvence i jejich alignmenty. U každé sekvence je také k dispozici interaktivní sekundární struktura. Ta však zobrazuje nukleotidové sekvence v kontextu obecné sekundární struktury pro danou doménu (např. eukaryota). Přizpůsobuje tedy (pomocí mezer) primární sekvenci sekundární struktury. (Szymanski et al. 2016)

## 2.4 Ensembl BioMart

BioMart je jedním z webových nástrojů poskytovaných projektem Ensembl. Umožňuje snadné získávání dat bez nutnosti programovacích zkušeností a bez potřeby učit se základní strukturu databáze, která může být poněkud matoucí. Ensembl také poskytuje BioMart Perl API verzi, která umožňuje přímou integraci funkcí tohoto nástroje do skriptů psaných v programovacím jazyce Perl. (Kinsella et al. 2011)

V této práci bylo nicméně využité webové rozhraní, jehož ovládání je velmi přehledné. Nejprve je potřeba zvolit databázi. Sekvence zpracované v této bakalářské práci pochází z databáze Ensembl Genes 104. Verze databází se pravidelně aktualizují, nicméně předešlé verze jsou dostupné prostřednictvím Ensembl Archives. Po zvolení databáze následuje výběr datasetu, tj. konkrétního druhu, s jehož daty chceme pracovat. Seznam druhů zpracovaných pro tuto bakalářskou práci je uveden v příloze (Příloha 1).

BioMart následně nabídne možnosti zvolit si filtry, atributy (tedy výstupy, které se zobrazí) a dodatečný dataset. Pro potřeby této práce byl použit filtr „Gene type“ s hodnotou „rRNA“. Zvolenými atributy byly „Exon sequences“ a „Exon stable ID“. Žádný dodatečný dataset zvolen nebyl.

Všechny výsledky je možno jednoduše stáhnout ve formátu FASTA. Tento formát se používá při manipulaci se sekvencemi DNA, RNA i proteinů. Data ukládá ve formě dvou řádků, první je „header line“, která začíná znakem „>“ a obsahuje popis sekvence (například v našem případě „Exon stable ID“). Ve druhém řádku je uložena vlastní sekvence. Před stažením výsledků nabízí BioMart možnost „Unique results only“, která v tomto případě byla zvolena. Také je možné prostřednictvím tlačítka „Count“ zobrazit počet sekvencí odpovídajících hodnotám filtru z celkového počtu sekvencí v konkrétním genomu, avšak tato hodnota neodpovídá výslednému počtu použitelných 5S rDNA sekvencí, protože se zde vyskytují sekvence nekompletní, i sekvence jiných rDNA (nejčastěji 5.8 S rDNA).



## 2.5 AliView

AliView je program, umožňující prohlížet a editovat alignmenty (seřazení a zarovnání) sekvencí. Byl navržen s důrazem na použitelnost a intuitivní ovládání. Je schopný velmi rychle zpracovat i rozsáhlé alignmenty a to s minimálním ovlivněním paměti zařízení. Dokáže přečíst neomezeně velké soubory ve formátech FASTA, Phylip, Nexus, Clustal a MSF. V závislosti na velikost paměti, kterou má konkrétní počítač k dispozici, může buď načíst celý alignment do paměti najednou, nebo ponechat určitou část souboru nenačtenou ve složce, čímž umožňuje práci s alignmenty i na starých počítačích a dalších zařízeních s omezenou kapacitou paměti. Jeho intuitivní grafické rozhraní umožňuje jednoduše kontrolovat, třídit, mazat, spojovat a přeskupovat sekvence v rámci manuální úpravy dat. Také umožňuje získat snadný přehled o datasetu díky možnosti zobrazit konsensuální a nekonsensuální prvky. (Larsson 2014)

Program AliView byl využit k úpravě datasetů sekvencí stažených z výše zmíněných databází. Bylo nutné vyloučit sekvence jiného typu než 5S rDNA, stejně jako sekvence nekompletní. Také umožnil získání základního přehledu o variabilitě sekvencí v rámci druhu.

## 2.6 Vienna RNA Websuite

Vienna RNA Websuite je kolekci volně dostupných nástrojů pro skládání („folding“), design a analýzu RNA sekvencí, spravovaná Institutem teoretické chemie při Vídeňské univerzitě. Poskytuje webové rozhraní pro nejpoužívanější nástroje Vienna RNA balíku (který obsahuje knihovnu pro programovací jazyk C a několik samostatných programů k predikci a porovnání sekundárních struktur RNA). Tyto nástroje umožňují skládání jednotlivých i alignovaných sekvencí, predikci RNA-RNA interakcí a design sekvencí s předem danou strukturou. (Gruber et al. 2008)

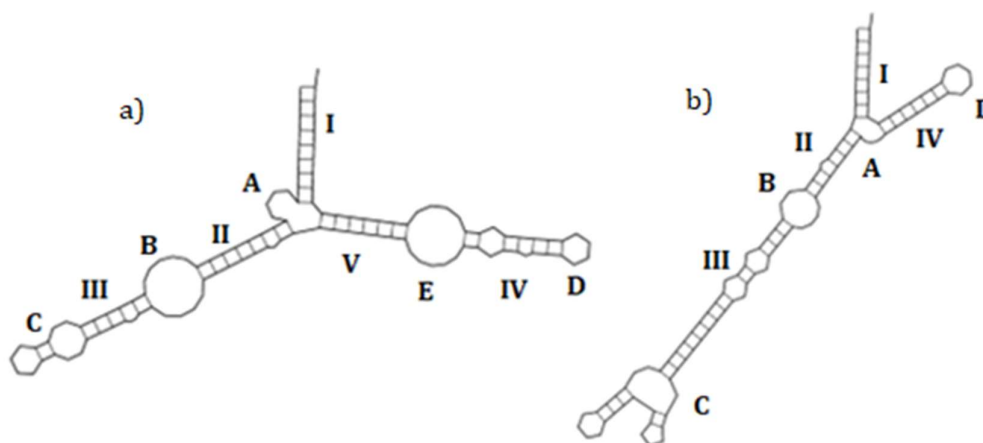
Dostupné nástroje mají jednotný vzhled. Na úvodní stránce se nalézají krátký popis funkce nástroje spolu s upozorněním na limit velikosti vstupních dat (ten je zde z důvodu omezené kapacity serveru, na kterém nástroje pracují). Dále jsou zde pole pro vložení vstupních dat a nastavení některých parametrů. Ty se dělí na základní a pokročilá nastavení. Každý parametr je zde krátce vysvětlen. Výchozí nastavení parametrů však bylo pečlivě nastaveno tak, aby jejich přizpůsobení nebylo ve většině případů nezbytné k dosažení optimálních výsledků. (Gruber et al. 2008)

V této práci byly využity nástroje RNAfold a RNAalifold. RNAfold predikuje sekundární strukturu s minimální volnou energií (MFE) pro jednu sekvenci, zatímco RNAalifold predikuje totéž pro více alignovaných sekvencí, maximálně však 300. V případě RNAfoldu mohou být vstupní data vložena jako prostý text nebo FASTA formát, RNAalifold vyžaduje data ve formátu FASTA nebo CLUSTAL W. (Gruber et al. 2008)

Nastavení parametrů obou nástrojů bylo ponecháno na výchozích hodnotách. Kvůli omezené kapacitě RNAalifoldu nemohly být u některých druhů použity kompletní alignmenty získané z databáze Ensembl, v takovém případě bylo použito 300 náhodně vybraných sekvencí daného druhu. Kromě predikovaných sekundárních struktur byly z výstupů těchto nástrojů využity i rRNA sekvence (v případě RNAalifoldu se jednalo o konsenzuální sekvenci).

### 3 Výsledky

Pro potřeby této bakalářské práce byly zpracovány 5S rDNA sekvence 92 druhů ryb z 34 řádů. U 24 řádů byla detekována sekundární struktura odpovídající konsenzuální sekundární struktuře a sekundárním strukturám pro eukaryotní 5S rRNA navrhovaným dalšími autory (dále jen „odpovídající“ sekundární struktura; Obr. 5a, 7, 8). Pro dva řády (*Carangiformes* a *Labriformes*) byla navržena kandidátní sekundární struktura odlišná od konsenzuální struktury (Obr. 5b). U 8 řádů se nepodařilo detekovat ani jednu z výše zmíněných struktur (Příloha 1).

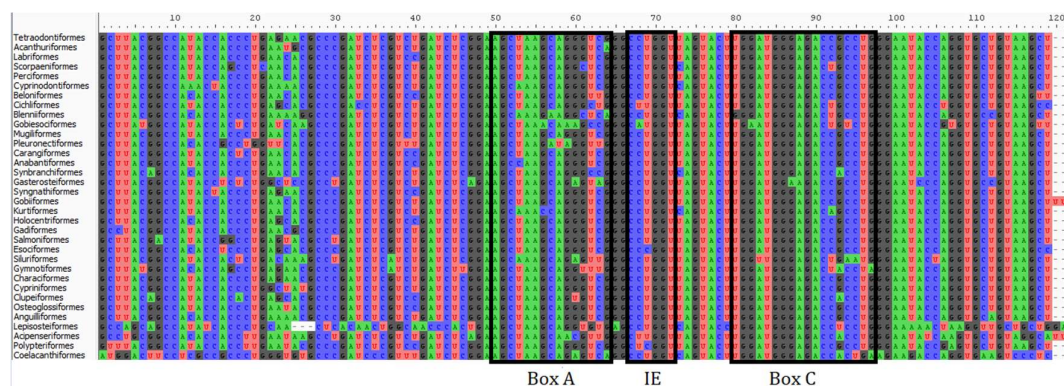


**Obr. 5:** (a) „Odpovídající“ sekundární struktura predikovaná pro druh *Anabas testudineus* (*Anabantiformes*), shodná s celkovou konsenzuální strukturou. (b) Kandidátní sekundární struktura pro druh *Echeneis naucrates* (*Carangiformes*). I až V – helixy, A až E – smyčky. Obě struktury byly vygenerovány pomocí RNAalifoldu (Gruber et al. 2008).

#### 3.1 Primární sekvence

Pro zachování efektivity transkripce musí všechny geny pro 5S rRNA obsahovat ICR („Internal control region“). Ten se skládá z boxu A, jež se nachází na pozicích 50-64; interního elementu IE, jež se nachází na pozicích 67-72 a boxu C na pozicích 80-97 (Castro et al. 2016; Martins a Wasko 2004). Tyto elementy byly detekovány v konsenzuálních sekvencích všech řádů zpracovaných v této práci (Obr. 6, ve větším měřítku Příloha 2). U většiny řádů se shodovala i pozice těchto elementů, jen

u řádu *Lepisosteiformes* byly posunuty o tři nukleotidy blíže k 5' konci (tj. box A na pozicích 47-61, IE na pozicích 64-69 a box C na pozicích 77-94).



**Obr. 6:** Vyznačení pozice ICR v alignmentu konsenzuálních sekvencí řádů. První ohraničení označuje box A, druhé interní element IE, třetí box C. Pro názornost byly do sekvence řádu *Lepisosteiformes* vloženy tři mezery. Větší verze obrázku je k dispozici v Příloze 2. Řády jsou seřazeny dle Hughes et al. (2018) od fylogeneticky nejmladších po fylogeneticky nejstarší.

### 3.2 Sekundární struktura

Detekovaná „odpovídající“ sekundární struktura (Obr. 5a, 7, 8) má klasické uspořádání 5S rRNA, tedy 3 helixová ramena spojená smyčkou A do vidlice. Sekundární struktury jednotlivých řádů (Tab. 2) se odlišovaly především velikostí smyček, popř. počtem smyček v levém rameni.

U řádů *Anguilliformes*, *Gadiformes* a *Osteoglossiformes* byla nalezena smyčka i v helixu I. Řád *Gadiformes* měl dále stejné uspořádání smyčky B jako řád *Belontiiformes*, tedy v helixu III místo smyčky dvou volných nukleotidů „klasickou“ smyčku (strukturou tak odpovídali Vierna et al. (2013), více v diskuzi).

U řádů *Acipenseriformes* a *Gobiesociformes* se vyskytla smyčka v helixu II. řád *Gobiesociformes* pak měl také výrazně zvětšené smyčky E a D a výrazně zkrácený helix IV.

U řádu *Blenniiformes* byla smyčka E o 7 nukleotidů menší, zatímco helixy IV a V se prodloužily. Volný nukleotid v helixu V byl „nahrazen“ smyčkou, vzniklou dvěma nespárovanými nukleotidy. Stejné uspořádání této oblasti platilo i pro řád *Clupeiformes*, který měl navíc místo volných nukleotidů v helixech II a III „klasické“ smyčky. Dále měl smyčku B o 3 nukleotidy menší a na začátku helixu I 4 nespárované nukleotidy (helix I byl tedy o 2 bp kratší).

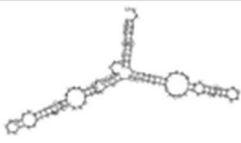
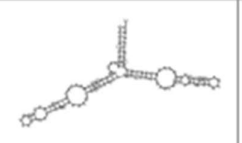
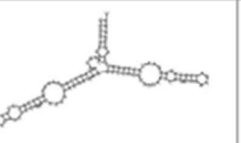
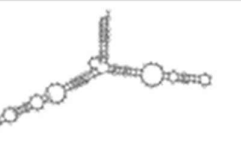
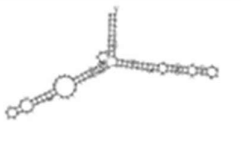
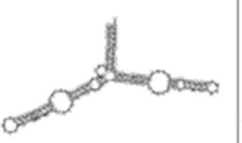
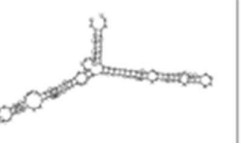
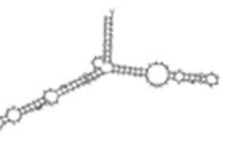
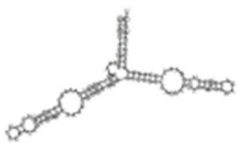
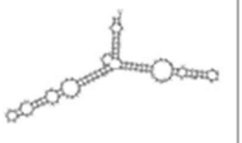
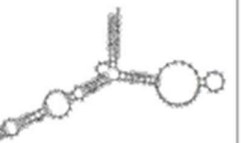
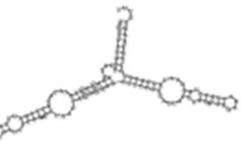
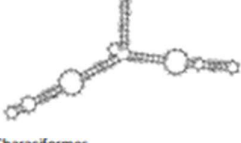
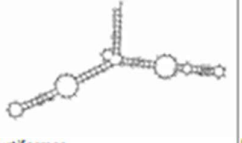
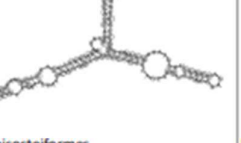
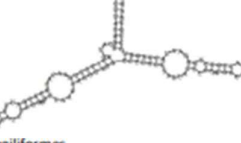
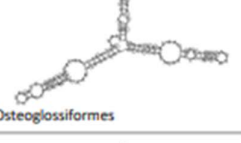
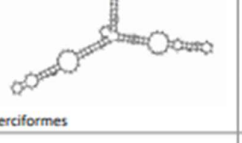
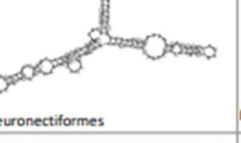
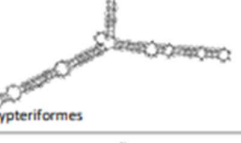
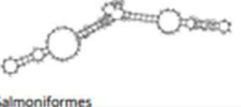

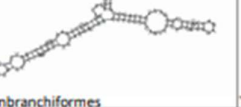
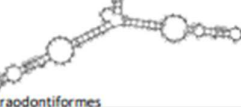
Řády *Cichliformes* a *Kurtiformes* měly „koncovou část“ rozdělené smyčky C o 2 nukleotidy větší, zatímco „vnitřní část“ této smyčky byla redukována na 2 volné

nukleotidy. Řád *Cichliformes* měl navíc v helixu II místo volného nukleotidu „klasickou“ smyčku (stejně jako řád *Clupeiformes*).

Řády *Polypteriformes* a *Synbranchiformes* měly o 5 nukleotidů menší smyčku B, a o 2 bp delší helix III kde se také nacházel další volný nukleotid (C). Podobné uspořádání měl i řád *Cypriniformes*, kde však došlo k prodloužení a k výskytu volného nukleotidu (G) v helixu II.

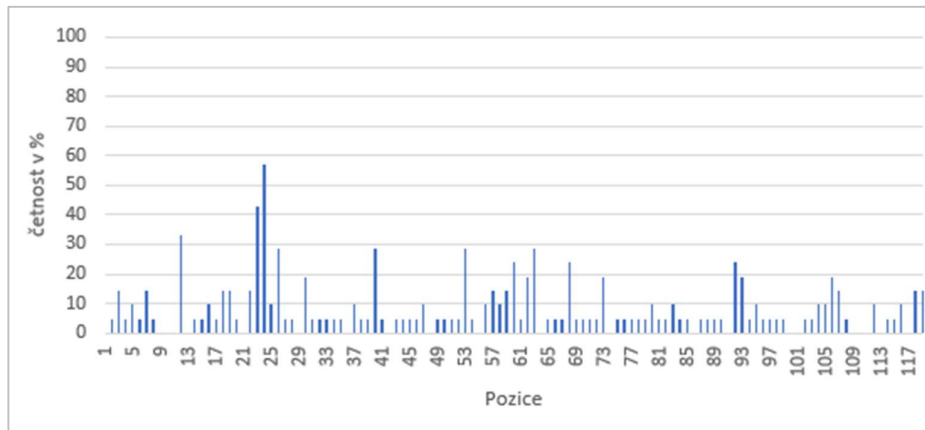
Řády *Salmoniformes* a *Siluriformes* měly „zaškrčení“ smyčky C o 2bp delší, „koncovou část“ smyčky C o 2 nukleotidy větší a „vnitřní část“ o 1 nukleotid menší. Také měly zvětšenou smyčku B a zkrácený helix III. U řádu *Salmoniformes* dále došlo k posunu volného nukleotidu v helixu II blíže směrem ke smyčce B. U řádu *Siluriformes* pak došlo také ke zkrácení helixu II. „Zaškrčení“ smyčky C měl delší (o 1 bp) i řád *Lepisosteiformes*. Ten měl také menší smyčku B, delší helix II a nespárované první dva nukleotidy v helixu I.

U řádu *Pleuronectiformes* byla detekována „vyhřezlá“ smyčka B, což naznačuje, že by tato struktura mohla mít narušenou funkci. Větší počet volných nukleotidů na 3' konci u řádů *Acipenseriformes* a *Gobiiformes* je způsoben zahrnutím části poly-T konce v použitých sekvencích.

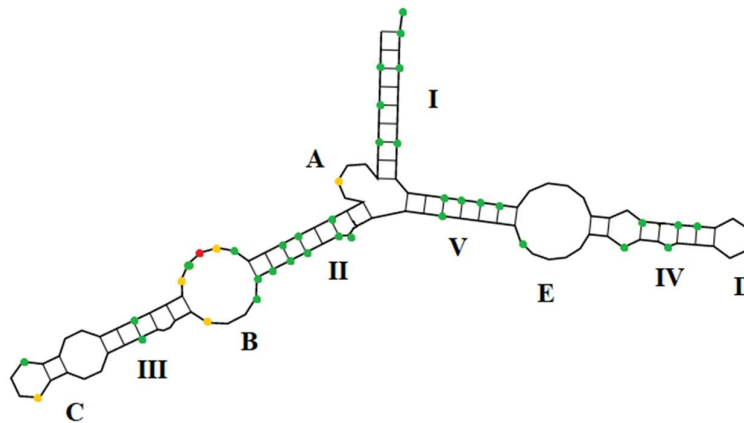
 Acipenseriformes	 Anabantiformes	 Anguilliformes	 Belontiiformes
 Blenniiformes	 Cichliformes	 Clupeiformes	 Cypriniformes
 Cyprinodontiformes	 Gadiformes	 Gobiiesociformes	 Gobiiformes
 Characiformes	 Kurtiformes	 Lepisosteiformes	 Mugiliformes
 Osteoglossiformes	 Perciformes	 Pleuronectiformes	 Polypteriformes
 Salmoniformes	 Siluriformes	 Synbranchiformes	 Tetraodontiformes

**Tab. 2:** Tabulka „odpovídajících“ konsenzuálních sekundárních struktur jednotlivých řádů. Řády jsou řazeny abecedně. Všechny struktury byly vygenerovány pomocí RNAalifoldu (Gruber et al. 2008).

Mutace v primárních sekvencích (Graf 1.) řádů s „odpovídající“ sekundární strukturou, se vyskytovaly především v helixech I, II, IV, V a ve smyčkách B a C, jak je znázorněno na Obr. 7. Jsou zde vyznačeny pouze mutace, jejichž relativní četnost byla vyšší než 5 %. Tato hladina odpovídá mutaci, jež se objevila v 1 z 24 sekvencí a byla zvolena kvůli odlišnosti sekvence druhu *Lepisosteus oculatus* (kostlín skvrnitý), jež je jediným zástupcem řádu *Lepisosteiformes* (který nepatří mezi koncoústé ryby) zahrnutým v této bakalářské práci. Jeho sekvence se lišila téměř na každé pozici (přestože obsahovala shodné motivy – zejména všechny části ICR, jen posunutá o tři nukleotidy směrem k 5' konci).

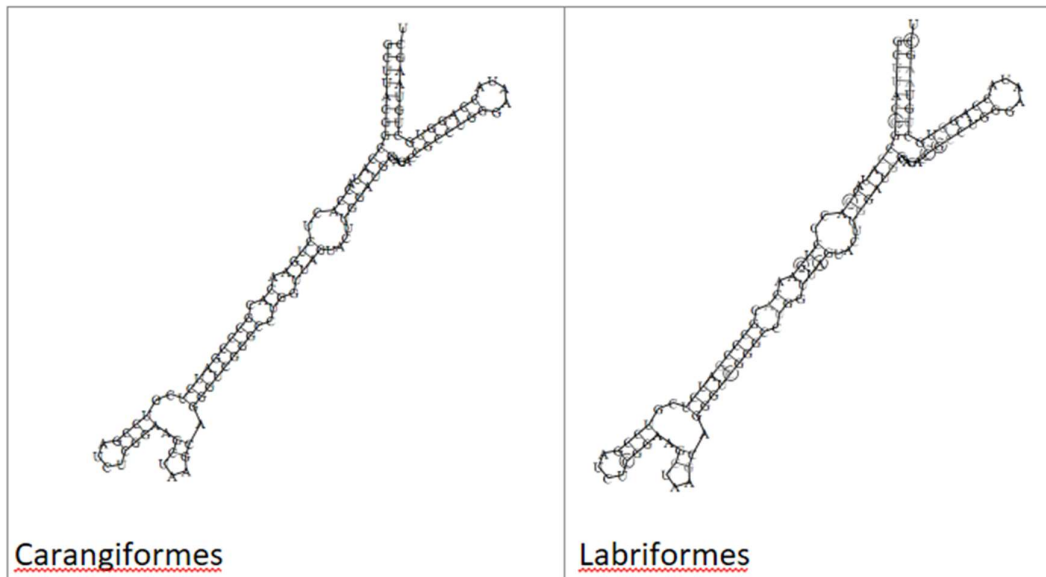


**Graf 1:** Graf relativních četností mutací na konkrétních pozicích v primárních sekvencích „odpovídajících“ sekundárních struktur (při porovnání konsenzuálních sekvencí řádů).



**Obr. 7:** Vyznačení pozic, na kterých se vyskytly mutace v primárních sekvencích „odpovídajících“ sekundárních struktur. Pozice, na kterých bylo detekováno 6-25% mutací jsou vyznačeny zeleně, 26-50% mutací žlutě a více než 50% mutací červeně. I až V – helixy, A až E – smyčky. Struktura byla vygenerována pomocí RNAalifoldu (Gruber et al. 2008).

Kandidátní sekundární struktura navržená pro řády *Carangiformes* a *Labriformes* (Obr. 5b, Tab. 3) se výrazně odlišuje od konsenzuální struktury, ale jednotlivé domény se částečně dají rozpoznat. Helix I je dlouhý 8 bp (u konsenzuální struktury je to 9 bp), smyčka A má menší „vyboulení“ na opačné straně struktury, volný nukleotid v helixu II se nachází blíže smyčce B (také na opačné straně struktury) a tento helix je dlouhý pouhých 7 bp (místo 8 bp, jako u konsenzuální struktury). Helix III je více než dvakrát delší a obsahuje dvě menší smyčky. Smyčka C je výrazně deformovaná. Helix V a smyčka E úplně chybí.



**Tab. 3:** Tabulka kandidátních sekundárních struktur jednotlivých řádů. Řády jsou řazeny abecedně. Obě struktury byly vygenerovány pomocí RNAalifoldu (Gruber et al. 2008).

## 4 Diskuze

### 4.1 Primární sekvence

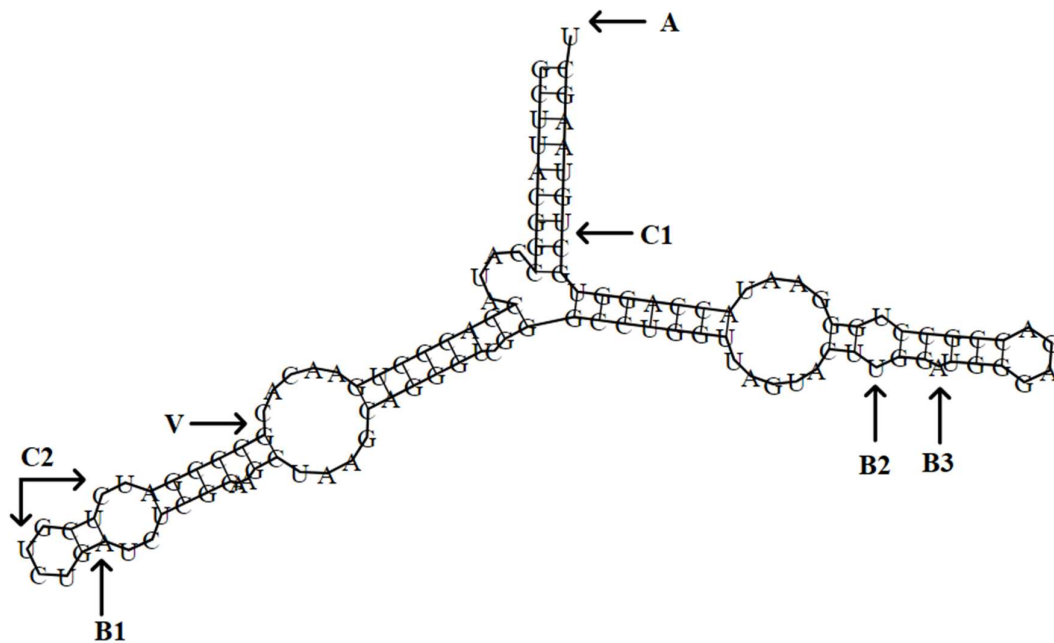
Martins a Wasko (2004) uvádí alignment sekvencí 5S rDNA pro vybrané druhy ryb, obojživelníků, plazů, ptáků a savců, v nichž jako box A označili sekvenci nukleotidů „AGCTAAGCAGGGTCG“, jako IE sekvenci „CCTGGT“ a jako box C sekvenci „TGGATGGGAGACCGCCTG“. Tyto sekvence jsou shodné s těmi, jež byly detekovány v rámci této bakalářské práce. Box C odpovídá typu A 5S rDNA, který detekovali Sajdak et al. (1998). Castro et al. (2016), kteří se zabývali 5S rDNA paryb (konkrétně typem II) uvádí jako motiv boxu A sekvenci „AGCTAAGCAGGCTCA“, na pozici 61 tedy detekovali cytosin místo guaninu a na pozici 64 adenin místo guaninu. Může to být znak specifický jak pro paryby tak pro 5S rDNA typu II, box A se však celkově zdá být nejvíce variabilním z motivů ICR. V obou zmíněných pracích (Martins a Wasko 2004; Castro et al. 2016) obsahoval nejvíce mutací právě box A, zatímco box C byl nejkonzervovanější. I v této bakalářské práci byla detekována nejvyšší pravděpodobnost výskytu mutace v rámci ICR právě v boxu A (4,8 %). Pro IE byla pravděpodobnost výskytu mutace 3,3 % a pro box C 3 %.

### 4.2 Sekundární struktura

Celková predikovaná konsenzuální sekundární struktura rybí 5S rRNA byla shodná se sekundární strukturou 5S rRNA popsanou dalšími autory (Obr. 8), a to i přes to, že jednotlivé sekvence (potažmo řády) měly v mnoha případech predikovanou sekundární strukturu výrazně odlišnou. Dle práce Castro et al. (2016) může být tato nižší přesnost predikce sekundárních struktur způsobena tím, že pro správné uspořádání potřebuje 5S rRNA pravděpodobně interagovat s dalšími molekulami a programy generující predikci sekundárních struktur zatím tuto interakci nedokáží zohlednit.

Barciszewska et al. (2000) predikovali obecný model sekundární struktury 5S rRNA pro všechna eukaryota, Castro et al. (2016) predikovali sekundární strukturu 5S rRNA pro paryby a Vierna et al. (2013) predikovali konsenzuální strukturu pro metazoa. Na rozdíl od predikovaných struktur všech výše zmíněných autorů končí níže zobrazená konsenzuální sekundární struktura (Obr. 8) na 3' konci pouze jedním volným nukleotidem (nikoli 2). To je způsobeno tím, že sekvence stažené z Ensemblu prostřednictvím BioMartu zpracovávané v této práci byly dlouhé 119, nikoli 120 bp, došlo tedy o zkrácení sekvencí o poslední nukleotid.





**Obr. 8:** Odchylky výsledné konsenzuální struktury od struktur predikovaných dalšími autory. A – odchýlení od všech dále zmíněných autorů, B1-3 – odchýlení od struktury predikované v Barciszewska et al. (2000), C1 a 2 – odchýlení od struktury predikované v Castro et al. (2016), V – odchýlení od struktury predikované ve Vierna et al. (2013). Konsenzuální struktura byla vygenerována pomocí RNAalifoldu (Gruber et al. 2008).

Odchylky (Obr. 8) od struktury predikované prací Barciszewska et al. (2000) spočívaly: 1) ve zmenšení smyčky C (resp. jejímu rozdělení na dvě menší smyčky) prostřednictvím párů 35U:42A a 36C:41G, 2) pozice 80U a 96U netvoří pár, ale smyčku a 3) v posunu drobné smyčky tvořené volným nukleotidem z pozice 84 na pozici 83.

Odchylky (Obr. 8) od struktury predikované prací Castro et al. (2016) spočívaly: 1) ve vytvoření páru 7G:112U, čímž z této pozice zmizela smyčka a 2) v posunu párů bází redukujících smyčku C ze 34C:41G a 33U:42A na 35U:42A a 36C:41G, čímž došlo ke zmenšení „koncové části“ smyčky C o 2 nukleotidy a ke zvětšení „vnitřní části“ smyčky C o 2 nukleotidy.

Odchylka (Obr. 8) od struktury predikované prací Vierna et al. (2013) spočívala v rozpojení bází na pozicích 25A, 52C a 26C, 51G, místo nichž vznikly páry 27G:52C a 28C:51G, čímž došlo ke zvětšení smyčky B o 2 nukleotidy a redukci sousedící smyčky o 2 nukleotidy.

Výrazná odlišnost kandidátní struktury navržené pro řády *Carangiformes* a *Labriformes* od konsenzuální struktury (a potažmo obecné struktury všech eukaryotních organismů) naznačuje, že by se mohlo jednat o pseudogen. Přestože primární sekvence 5S rDNA může být do značné míry variabilní, její sekundární

a terciární struktura (nezbytná pro správnou funkci v rámci ribozomu) je vysoce konzervovaná (Smirnov et al. 2008). Na druhou stranu jsou však pseudogeny díky nízkému selekčnímu tlaku typické hromaděním mutací (Lodish 2000, D'Errico 2004), což může vést ke zvyšování fylogenetické diverzity (Madan Babu 2003). Tato predikovaná kandidátní sekundární struktura však byla u všech druhů na úrovni primární sekvence vysoce konzervovaná, což naznačuje, že pokud se opravdu jedná o pseudogen, mohl v průběhu evoluce získat novou funkci (D'Errico 2004). Jak však poukázali Castro et al. (2016), predikce sekundárních struktur 5S rRNA je v porovnání s ostatními rRNA méně přesná a může se tedy jednat o funkční gen, jež potřebuje ke správnému prostorovému uspořádání další interakce, které v algoritmu, jež predikuje sekundární strukturu, nejsou zahrnuty. To otevírá možnost dalšího výzkumu.

Kvalitu predikovaných konsensuálních sekvencí a struktur však mohl ovlivnit i počet kopií 5S rRNA genů v rámci druhu, stejně jako počet druhů v rámci řádu, pro které byla dostupná data. Tyto rozdíly jsou patrné v tabulce zpracovaných druhů (Příloha 1). Zatímco však počet dostupných druhů mohl silně ovlivnit, detekci různých typů sekundárních struktur, zdá se, že počet sekvencí neměl vliv na kvalitu výsledné konsensuální struktury (pro srovnání např. rozdíly mezi *Esociformes* a *Gadiformes*).

## Závěr

V rámci této bakalářské práce byly zpracovány sekvence 5S rDNA 92 druhů ryb z 34 řádů, pocházející z volně dostupných databází Ensembl, NCBI a 5SRNAdb. Tyto sekvence se daly rozdělit do tří skupin: 1) Na ty, které tvořily sekundární strukturu odpovídající obecné struktuře 5S rRNA eukaryot; 2) ty, které tvořily predikovanou kandidátní sekundární strukturu a 3) ty, jejichž sekundární struktury byly příliš variabilní (a pravděpodobně se tedy jednalo o pseudogeny).

Přesto, že se sekundární struktury u různých řádů částečně lišily, konsenzuální sekundární struktura všech řádů odpovídala doposud predikovaným sekundárním strukturám 5S rRNA eukaryotních organismů.

Predikovaná kandidátní sekundární struktura byla od konsenzuální sekundární struktury eukaryotních organismů výrazně odlišná, což by nasvědčovalo tomu, že se ve skutečnosti jednalo o pseudogen, avšak konzervativnost jejích primárních sekvencí u různých řádů ryb naznačuje, že pokud se skutečně jednalo o pseudogen, mohl by to být pravděpodobně pseudogen neprocesovaný, který získal novou funkci a podařilo se mu tak udržet se v populaci.

U osmi řádů se nepodařilo identifikovat ani jeden z těchto typů sekundárních struktur (patřily tedy do třetí skupiny). To bylo částečně způsobeno jejich nedostatečnou reprezentací – od každého z těchto řádů byly získány sekvence pouze jednoho druhu. Počet osekvenovaných druhů však neustále stoupá a genomové anotace se zpřesňují. Do budoucna tedy bude možnost tuto práci o tyto druhy rozšířit.

## Seznam použité literatury

ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS a D. J. LIPMAN, 1990. Basic local alignment search tool. *Journal of Molecular Biology* [online]. 215(3), 403–410. ISSN 00222836. Dostupné z: doi:10.1016/S0022-2836(05)80360-2

BARCISZEWSKA, M. Z., M. SZYMAŃSKI, V. A. ERDMANN a J. BARCISZEWSKI, 2001. Structure and functions of 5S rRNA. *Acta Biochimica Polonica* [online]. 48(1), 191–198. ISSN 1734-154X, 0001-527X. Dostupné z: doi:10.18388/abp.2001\_5126

BARCISZEWSKA, M. Z., M. SZYMAŃSKI, V. A. ERDMANN a J. BARCISZEWSKI, 2000. 5S Ribosomal RNA. *Biomacromolecules* [online]. 1(3), 297–302. ISSN 1525-7797, 1526-4602. Dostupné z: doi:10.1021/bm000293o

BENSASSON, D., 2001. Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology & Evolution* [online]. 16(6), 314–321. ISSN 01695347. Dostupné z: doi:10.1016/S0169-5347(01)02151-6

CAMPO, D., G. MACHADO-SCHIAFFINO, J. L. HORREO a E. GARCIA-VAZQUEZ, 2009. Molecular Organization and Evolution of 5S rDNA in the Genus *Merluccius* and Their Phylogenetic Implications. *Journal of Molecular Evolution* [online]. 68(3), 208–216. ISSN 0022-2844, 1432-1432. Dostupné z: doi:10.1007/s00239-009-9207-8

CASTRO, S. I., J. S. HLEAP, H. CÁRDENAS a C. BLOUIN, 2016. Molecular organization of the 5S rDNA gene type II in elasmobranchs. *RNA Biology* [online]. 13(4), 391–399. ISSN 1547-6286, 1555-8584. Dostupné z: doi:10.1080/15476286.2015.1100796

CORNISH-BOWDEN, A., 1985. Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. *Nucleic Acids Research* [online]. 13(9), 3021–3030. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/13.9.3021

DE FÁTIMA CHASKO RIBEIRO, L. a M. APARECIDA FERNANDEZ, 2004. Molecular characterization of the 5S ribosomal gene of the *Bradysia hygida* (Diptera: Sciaridae). *Genetica* [online]. 122(3), 253–260. ISSN 0016-6707, 1573-6857. Dostupné z: doi:10.1007/s10709-004-1704-5

DELIHAS, N. a J. ANDERSEN, 1982. Generalized structures of the 5S ribosomal RNAs. *Nucleic Acids Research* [online]. 10(22), 7323–7344. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/10.22.7323

D'ERRICO, I., 2004. Pseudogenes in metazoa: Origin and features. *Briefings in Functional Genomics and Proteomics* [online]. 3(2), 157–167. ISSN 1473-9550, 1477-4062. Dostupné z: doi:10.1093/bfgp/3.2.157

DINMAN, J. D., 2005. 5S rRNA: Structure and Function from Head to Toe. *International journal of biomedical science: IJBS*. 1(1), 2–7. ISSN 1550-9702.

- EICKBUSH, T. H. a D. G. EICKBUSH, 2007. Finely Orchestrated Movements: Evolution of the Ribosomal RNA Genes. *Genetics* [online]. 175(2), 477–485. ISSN 1943-2631. Dostupné z: doi:10.1534/genetics.107.071399
- FOORD, S. M., B. MARKS, M. STOLZ, E. BUFFLIER, N. J. FRASER a M. G. LEE, 1996. The Structure of the Prostaglandin EP4 Receptor Gene and Related Pseudogenes. *Genomics* [online]. 35(1), 182–188. ISSN 08887543. Dostupné z: doi:10.1006/geno.1996.0337
- GIBBONS, J. G., A. T. BRANCO, S. A. GODINHO, S. YU a B. LEMOS, 2015. Concerted copy number variation balances ribosomal DNA dosage in human and mouse genomes. *Proceedings of the National Academy of Sciences* [online]. 112(8), 2485–2490. ISSN 0027-8424, 1091-6490. Dostupné z: doi:10.1073/pnas.1416878112
- GIBBONS, J. G., A. T. BRANCO, S. YU a B. LEMOS, 2014. Ribosomal DNA copy number is coupled with gene expression variation and mitochondrial abundance in humans. *Nature Communications* [online]. 5(1), 4850. ISSN 2041-1723. Dostupné z: doi:10.1038/ncomms5850
- GILLESPIE, J. J., J. S. JOHNSTON, J. J. CANNONE a R. R. GUTELL, 2006. Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta: Hymenoptera): structure, organization, and retrotransposable elements. *Insect Molecular Biology* [online]. 15(5), 657–686. ISSN 0962-1075, 1365-2583. Dostupné z: doi:10.1111/j.1365-2583.2006.00689.x
- GONG, J., J. DONG, X. LIU a R. MASSANA, 2013. Extremely High Copy Numbers and Polymorphisms of the rDNA Operon Estimated from Single Cell Analysis of Oligotrich and Peritrich Ciliates. *Protist* [online]. 164(3), 369–379. ISSN 14344610. Dostupné z: doi:10.1016/j.protis.2012.11.006
- GRUBER, A. R., R. LORENZ, S. H. BERNHART, R. NEUBOCK a I. L. HOFACKER, 2008. The Vienna RNA Websuite. *Nucleic Acids Research* [online]. 36(Web Server), W70–W74. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/gkn188
- GRUMMT, I., 2013. The nucleolus—guardian of cellular homeostasis and genome integrity. *Chromosoma* [online]. 122(6), 487–497. ISSN 0009-5915, 1432-0886. Dostupné z: doi:10.1007/s00412-013-0430-0
- HALL, A. N., T. N. TURNER a Ch. QUEITSCH, 2021. Thousands of high-quality sequencing samples fail to show meaningful correlation between 5S and 45S ribosomal DNA arrays in humans. *Scientific Reports* [online]. 11(1), 449. ISSN 2045-2322. Dostupné z: doi:10.1038/s41598-020-80049-y
- HALL, T. M. T., 2005. Multiple modes of RNA recognition by zinc finger proteins. *Current Opinion in Structural Biology* [online]. 15(3), 367–373. ISSN 0959440X. Dostupné z: doi:10.1016/j.sbi.2005.04.004

HARRIS, S., P. A. BARRIE, M. L. WEISS a A. J. JEFFREYS, 1984. The primate  $\psi\beta 1$  gene. *Journal of Molecular Biology* [online]. 180(4), 785–801. ISSN 00222836. Dostupné z: doi:10.1016/0022-2836(84)90257-2

HAY, E. D. a J. B. GURDON, 1967. Fine Structure of the Nucleolus in Normal and Mutant *Xenopus* Embryos. *Journal of Cell Science* [online]. 2(2), 151–162. ISSN 1477-9137, 0021-9533. Dostupné z: doi:10.1242/jcs.2.2.151

HENRAS, A. K., C. PLISSON-CHASTANG, M.-F. O'DONOHUE, A.CHAKRABORTY a P.-E. GLEIZES, 2015. An overview of pre-ribosomal RNA processing in eukaryotes: Pre-ribosomal RNA processing in eukaryotes. *Wiley Interdisciplinary Reviews: RNA* [online]. 6(2), 225–242. ISSN 17577004. Dostupné z: doi:10.1002/wrna.1269

HOWE, K. L, P. ACHUTHAN, J. ALLEN et al., 2021. Ensembl 2021. *Nucleic Acids Research* [online]. 49(D1), D884–D891. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/gkaa942

HOWELL, W. M. a D. A. BLACK, 1980. Controlled silver-staining of nucleolus organizer regions with a protective colloidal developer: a 1-step method. *Experientia* [online]. 36(8), 1014–1015. ISSN 0014-4754, 1420-9071. Dostupné z: doi:10.1007/BF01953855

HUGHES, L. C., G.ORTÍ, Y. HUANG, Y. SUN, C. C. BALDWIN, A. W. THOMPSON, D. ARCILA, R. BETANCUR-R., Ch. LI, L. BECKER, N. BELLORA, X. ZHAO, X. LI, M. WANG, Ch. FANG, B. XIE, Z. ZHOU, H. HUANG, S. CHEN, B. VENKATESH a Q. SHI, 2018. Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proceedings of the National Academy of Sciences* [online]. 115(24), 6249–6254. ISSN 0027-8424, 1091-6490. Dostupné z: doi:10.1073/pnas.1719358115

CHILDS, G., R. MAXSON, R. H. COHN a L. KEDES, 1981. Orphans: Dispersed genetic elements derived from tandem repetitive genes of eucaryotes. *Cell* [online]. 23(3), 651–663. ISSN 00928674. Dostupné z: doi:10.1016/0092-8674(81)90428-1

IDE, S., T. MIYAZAKI, H. MAKI a T. KOBAYASHI, 2010. Abundance of Ribosomal RNA Gene Copies Maintains Genome Integrity. *Science* [online]. 327(5966), 693–696. ISSN 0036-8075, 1095-9203. Dostupné z: doi:10.1126/science.1179044

JAMES, S. A., M. J.T. O'KELLY, D. M. CARTER, R. P. DAVEY, A. VAN OUDENAARDEN a I. N. ROBERTS, 2009. Repetitive sequence variation and dynamics in the ribosomal DNA array of *Saccharomyces cerevisiae* as revealed by whole-genome resequencing. *Genome Research* [online]. 19(4), 626–635. ISSN 1088-9051. Dostupné z: doi:10.1101/gr.084517.108

KINSELLA, R. J., A. KAHARI, S. HAIDER, J. ZAMORA, G. PROCTOR, G. SPUDICH, J. ALMEIDA-KING, D. STAINES, P. DERWENT, A. KERHORNOU, P. KERSEY a P. FLICEK,

2011. Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database [online]. 2011(0), bar030–bar030. ISSN 1758-0463. Dostupné z: doi:10.1093/database/bar030
- KIPARISOV, S., A. PETROV, A. MESKAUSKAS, P. V. SERGIEV, O. A. DONTSOVA aj. D. DINMAN, 2005. Structural and functional analysis of 5S rRNA in *Saccharomyces cerevisiae*. Molecular Genetics and Genomics [online]. 274(3), 235–247. ISSN 1617-4615, 1617-4623. Dostupné z: doi:10.1007/s00438-005-0020-9
- KOBAYASHI, T., 2014. Ribosomal RNA gene repeats, their stability and cellular senescence. Proceedings of the Japan Academy, Series B [online]. 90(4), 119–129. ISSN 0386-2208, 1349-2896. Dostupné z: doi:10.2183/pjab.90.119
- KORN, L. J. a D. D. BROWN, 1978. Nucleotide sequence of xenopus borealis oocyte 5S DNA: Comparison of sequences that flank several related eucaryotic genes. Cell [online]. 15(4), 1145–1156. ISSN 00928674. Dostupné z: doi:10.1016/0092-8674(78)90042-9
- KUPRIYANOVA, N. S., 2000. Conservation and variation of ribosomal DNA in eukaryotes. Molecular Biology [online]. 34(5), 637–647. ISSN 0026-8933, 1608-3245. Dostupné z: doi:10.1007/BF02759600
- LAFONTAINE, D. L. J., 2015. Noncoding RNAs in eukaryotic ribosome biogenesis and function. Nature Structural & Molecular Biology [online]. 22(1), 11–19. ISSN 1545-9993, 1545-9985. Dostupné z: doi:10.1038/nsmb.2939
- LARSSON, A., 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics [online]. 30(22), 3276–3278. ISSN 1460-2059, 1367-4803. Dostupné z: doi:10.1093/bioinformatics/btu531
- LIM, K., Y. FURUTA a I. KOBAYASHI, 2012. Large Variations in Bacterial Ribosomal RNA Genes. Molecular Biology and Evolution [online]. 29(10), 2937–2948. ISSN 0737-4038, 1537-1719. Dostupné z: doi:10.1093/molbev/mss101
- LOCATI, M. D., J. F.B. PAGANO, W. A. ENSINK, M. VAN OLST, S. VAN LEEUWEN, U. NEHRDICH, K. ZHU, H. P. SPAINK, G. GIRARD, H. RAUWERDA, M. J. JONKER, R. J. DEKKER a T. M. BREIT, 2017. Linking maternal and somatic 5S rRNA types with different sequence-specific non-LTR retrotransposons. RNA [online]. 23(4), 446–456. ISSN 1355-8382, 1469-9001. Dostupné z: doi:10.1261/rna.059642.116
- LODISH, H. F., ed., 2000. Molecular cell biology. 4th ed. New York: W.H. Freeman. ISBN 978-0-7167-3136-8.
- LONG, E. O. a I. B. DAWID, 1980. Repeated Genes in Eukaryotes. Annual Review of Biochemistry [online]. 49(1), 727–764. ISSN 0066-4154, 1545-4509. Dostupné z: doi:10.1146/annurev.bi.49.070180.003455

- MADAN BABU, M, 2003. Did the loss of sigma factors initiate pseudogene accumulation in *M. leprae*? Trends in Microbiology [online]. 11(2), 59–61. ISSN 0966842X. Dostupné z: doi:10.1016/S0966-842X(02)00031-8
- MARTIN, W. F., S. GARG a V. ZIMORSKI, 2015. Endosymbiotic theories for eukaryote origin. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences [online]. 370(1678), 20140330. ISSN 1471-2970. Dostupné z: doi:10.1098/rstb.2014.0330
- MARTINS, C. a A. P. WASKO, 2004. Organization And Evolution of 5S Ribosomal DNA In The Fish Genome. In: Clyde R. WILLIAMS, ed. Focus on Genome Research. New York: Nova Science Publishers, Inc., s. 335–363. ISBN 978-1-59033-960-2.
- MAYER, Ch., K.-M. SCHMITZ, J. LI, I. GRUMMT a R. SANTORO, 2006. Intergenic Transcripts Regulate the Epigenetic State of rRNA Genes. Molecular Cell [online]. 22(3), 351–361. ISSN 10972765. Dostupné z: doi:10.1016/j.molcel.2006.03.028
- MCLAREN, W., L. GIL, S. E. HUNT, H. S. RIAT, G. R. S. RITCHIE, A. THORMANN, P. FLICEK a F. CUNNINGHAM, 2016. The Ensembl Variant Effect Predictor. Genome Biology [online]. 17(1), 122. ISSN 1474-760X. Dostupné z: doi:10.1186/s13059-016-0974-4
- MCSTAY, B., 2016. Nucleolar organizer regions: genomic ‘dark matter’ requiring illumination. Genes & Development [online]. 30(14), 1598–1610. ISSN 0890-9369, 1549-5477. Dostupné z: doi:10.1101/gad.283838.116
- NAKAJIMA, R. T, D.C CABRAL-DE-MELLO, G.T VALENTE, P. C VENERE a C. MARTINS, 2012. Evolutionary dynamics of rRNA gene clusters in cichlid fish. BMC Evolutionary Biology [online]. 12(1), 198. ISSN 1471-2148. Dostupné z: doi:10.1186/1471-2148-12-198
- NCBI RESOURCE COORDINATORS, R. AGARWALA, T. BARRETT et al., 2018. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research [online]. 46(D1), D8–D13. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/gkx1095
- NEELY, L. S., B. M. LEE, J. XU, P. E. WRIGHT a J. M. GOTTFELD, 1999. Identification of a minimal domain of 5 S ribosomal RNA sufficient for high affinity interactions with the RNA-specific zinc fingers of transcription factor IIIA 1 Edited by D. E. Draper. Journal of Molecular Biology [online]. 291(3), 549–560. ISSN 00222836. Dostupné z: doi:10.1006/jmbi.1999.2985
- NEI, M. a A. P. ROONEY, 2005. Concerted and Birth-and-Death Evolution of Multigene Families. Annual Review of Genetics [online]. 39(1), 121–152. ISSN 0066-4197, 1545-2948. Dostupné z: doi:10.1146/annurev.genet.39.073003.112240



- NÉMETH, A. a G. LÄNGST, 2011. Genome organization in and around the nucleolus. *Trends in Genetics* [online]. 27(4), 149–156. ISSN 01689525. Dostupné z: doi:10.1016/j.tig.2011.01.002
- PAREDES, S., A. T. BRANCO, D. L. HARTL, K. A. MAGGERT a B. LEMOS, 2011. Ribosomal DNA Deletions Modulate Genome-Wide Gene Expression: “rDNA-Sensitive” Genes and Natural Variation. *PLoS Genetics* [online]. 7(4), e1001376. ISSN 1553-7404. Dostupné z: doi:10.1371/journal.pgen.1001376
- PENDAS, A.M., P. MORAN, J.P. FREIJE a E. GARCIA-VAZQUEZ, 1994. Chromosomal mapping and nucleotide sequence of two tandem repeats of Atlantic salmon 5S rDNA. *Cytogenetic and Genome Research* [online]. 67(1), 31–36. ISSN 1424-859X, 1424-8581. Dostupné z: doi:10.1159/000133792
- PETERSON, R. C., J. L. DOERING a D. D. BROWN, 1980. Characterization of two xenopus somatic 5S DNAs and one minor oocyte-specific 5S DNA. *Cell* [online]. 20(1), 131–141. ISSN 00928674. Dostupné z: doi:10.1016/0092-8674(80)90241-X
- PHILLIPS, R. B. a K. M. REED, 1996. Application of fluorescence in situ hybridization (FISH) techniques to fish genetics: a review. *Aquaculture* [online]. 140(3), 197–216. ISSN 00448486. Dostupné z: doi:10.1016/0044-8486(95)01140-4
- POROKHOVNIK, L. aj. L. GERTON, 2019. Ribosomal DNA-connecting ribosome biogenesis and chromosome biology. *Chromosome Research* [online]. 27(1–2), 1–3. ISSN 0967-3849, 1573-6849. Dostupné z: doi:10.1007/s10577-018-9601-4
- REECE, J. B. a N. A. CAMPBELL, ed., 2011. *Campbell biology*. 9th ed. Boston: Benjamin Cummings / Pearson. ISBN 978-0-321-55823-7.
- RUSSELL, J. aj. C.B.M. ZOMERDIJK, 2005. RNA-polymerase-I-directed rDNA transcription, life and works. *Trends in Biochemical Sciences* [online]. 30(2), 87–96. ISSN 09680004. Dostupné z: doi:10.1016/j.tibs.2004.12.008
- SAJDAK, S. L., K. M. REED a R. B. PHILLIPS, 1998. Intraindividual and Interspecies Variation in the 5S rDNA of Coregonid Fish. *Journal of Molecular Evolution* [online]. 46(6), 680–688. ISSN 0022-2844. Dostupné z: doi:10.1007/PL00006348
- SCHMID, M., 1980. Chromosome banding in amphibia: IV. Differentiation of GC- and AT-rich chromosome regions in Anura. *Chromosoma* [online]. 77(1), 83–103. ISSN 0009-5915, 1432-0886. Dostupné z: doi:10.1007/BF00292043
- SCHULER, G. D., J. A. EPSTEIN, H. OHKAWA aj. A. KANS, 1996. Entrez: Molecular biology database and retrieval system. In: *Methods in Enzymology* [online]. B.m.: Elsevier, s. 141–162 [vid. 2021-07-17]. ISBN 978-0-12-182167-8. Dostupné z: doi:10.1016/S0076-6879(96)66012-1

- SMIRNOV, A. V., N. S. ENTELIS, I. A. KRASHENINNIKOV, R. MARTIN a I. A. TARASSOV, 2008. Specific features of 5S rRNA structure — Its interactions with macromolecules and possible functions. *Biochemistry (Moscow)* [online]. 73(13), 1418–1437. ISSN 0006-2979, 1608-3040. Dostupné z: doi:10.1134/S000629790813004X
- SOCHOROVÁ, J. *Databáze živočišné rDNA*. Brno, 2019. Disertační práce na Přírodovědecké fakultě Masarykovy univerzity. Školitel RNDr. A. Kovařík, CSc., 92s.
- SOCHOROVÁ, J., S. GARCIA, F. GÁLVEZ, R. SYMONOVÁ a A. KOVAŘÍK, 2018. Evolutionary trends in animal ribosomal DNA loci: introduction to a new online database. *Chromosoma* [online]. 127(1), 141–150. ISSN 0009-5915, 1432-0886. Dostupné z: doi:10.1007/s00412-017-0651-8
- STOCK, D. a G. WHITT, 1992. Evidence from 18S ribosomal RNA sequences that lampreys and hagfishes form a natural group. *Science* [online]. 257(5071), 787–789. ISSN 0036-8075, 1095-9203. Dostupné z: doi:10.1126/science.1496398
- SYMONOVÁ, R., 2019. Integrative rDNAomics—Importance of the Oldest Repetitive Fraction of the Eukaryote Genome. *Genes* [online]. 10(5), 345. ISSN 2073-4425. Dostupné z: doi:10.3390/genes10050345
- SYMONOVÁ, R., K. OCALEWICZ, L. KIRTIKLIS, G. B.DELMASTRO, Š. PELIKÁNOVÁ, S. GARCIA a A. KOVAŘÍK, 2017. Higher-order organisation of extremely amplified, potentially functional and massively methylated 5S rDNA in European pikes (*Esox* sp.). *BMC Genomics* [online]. 18(1), 391. ISSN 1471-2164. Dostupné z: doi:10.1186/s12864-017-3774-7
- SZYMANSKI, M., A. ZIELEZINSKI, J. BARCISZEWSKI, V. A. ERDMANN a W. M. KARLOWSKI, 2016. 5SRNadb: an information resource for 5S ribosomal RNAs. *Nucleic Acids Research* [online]. 44(D1), D180–D183. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/gkv1081
- TOMASZEWSKI, R., 1997. The AT-rich flanks of the oocyte-type 5S RNA gene of *Xenopus laevis* act as a strong local signal for histone H1-mediated chromatin reorganization in vitro. *Nucleic Acids Research* [online]. 25(3), 458–466. ISSN 13624962. Dostupné z: doi:10.1093/nar/25.3.458
- TOMPA, P. a P. CSERMELY, 2004. The role of structural disorder in the function of RNA and protein chaperones. *The FASEB Journal* [online]. 18(11), 1169–1175. ISSN 0892-6638, 1530-6860. Dostupné z: doi:10.1096/fj.04-1584rev
- VAN ETTEN, R. A., M. W. WALBERG a D. A. CLAYTON, 1980. Precise localization and nucleotide sequence of the two mouse mitochondrial rRNA genes and three immediately adjacent novel tRNA genes. *Cell* [online]. 22(1), 157–170. ISSN 00928674. Dostupné z: doi:10.1016/0092-8674(80)90164-6

VANIN, E. F., 1985. Processed Pseudogenes: Characteristics and Evolution. *Annual Review of Genetics* [online]. 19(1), 253–272. ISSN 0066-4197, 1545-2948. Dostupné z: doi:10.1146/annurev.ge.19.120185.001345

VIERNA, J., S. WEHNER, C. HÖNER ZU SIEDERDISSEN, A. MARTÍNEZ-LAGE a M. MARZ, 2013. Systematic analysis and evolution of 5S ribosomal DNA in metazoans. *Heredity* [online]. 111(5), 410–421. ISSN 0018-067X, 1365-2540. Dostupné z: doi:10.1038/hdy.2013.63

WARMERDAM, D. O. a R. M. F. WOLTHUIS, 2019. Keeping ribosomal DNA intact: a repeating challenge. *Chromosome Research* [online]. 27(1–2), 57–72. ISSN 0967-3849, 1573-6849. Dostupné z: doi:10.1007/s10577-018-9594-z

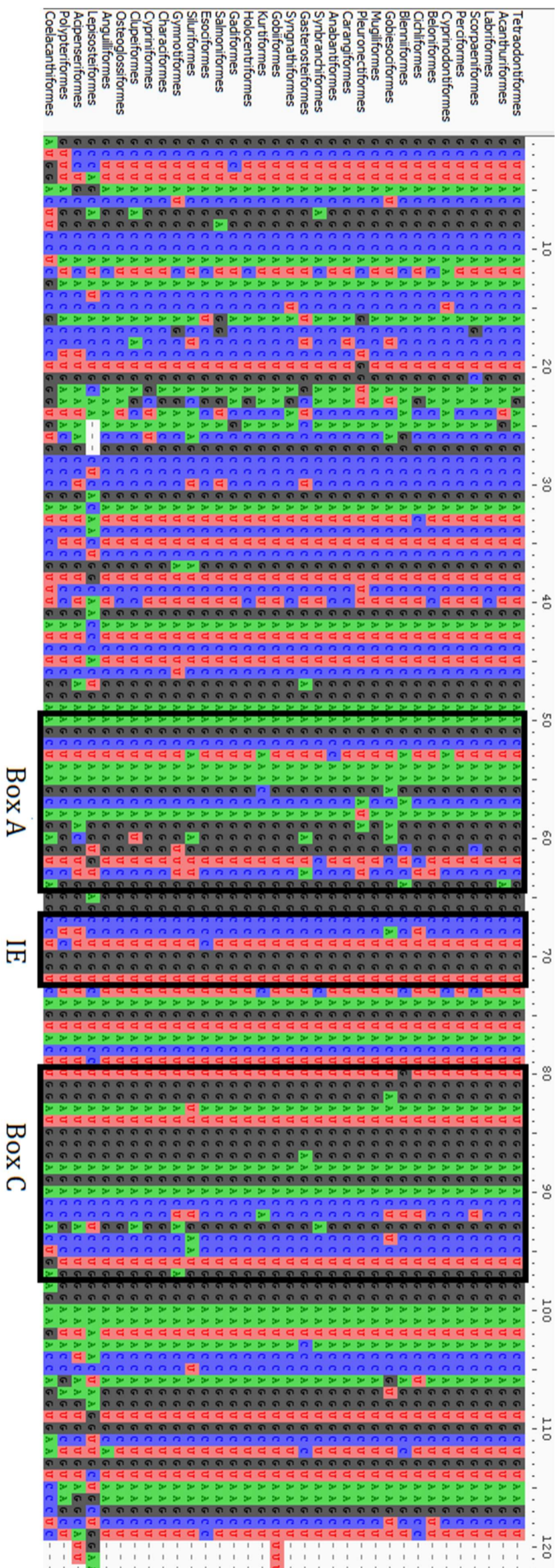
WASKO, A. P., C. MARTINS, J. M. WRIGHT a P. M. GALETTI, 2001. Molecular organization of 5S rDNA in fishes of the genus *Brycon*. *Genome*. 44(5), 893–902. ISSN 0831-2796.

WOISCHNIK, M. a C. T. MORAES, 2002. Pattern of Organization of Human Mitochondrial Pseudogenes in the Nuclear Genome. *Genome Research* [online]. 12(6), 885–893. ISSN 1088-9051. Dostupné z: doi:10.1101/gr.227202

WOLFFE, A. a D. BROWN, 1988. Developmental regulation of two 5S ribosomal RNA genes. *Science* [online]. 241(4873), 1626–1632. ISSN 0036-8075, 1095-9203. Dostupné z: doi:10.1126/science.3420414

XU, B., H. LI, J. M. PERRY, V. Pratap SINGH, J. UNRUH, Z. YU, M. ZAKARI, W. MCDOWELL, L. LI a J. L. GERTON, 2017. Ribosomal DNA copy number loss and sequence variation in cancer. *PLOS Genetics* [online]. 13(6), e1006771. ISSN 1553-7404. Dostupné z: doi:10.1371/journal.pgen.1006771





**Příloha 2:** Vyznačení pozice ICR v alignmentu konsenzuálních sekvencí řádů. První ohraničení označuje box A, druhé interní element IE, třetí box C. Pro názornost byly do sekvence řádu *Lepisosteiformes* vloženy tři mezery. Řády jsou seřazeny dle Hughes et al. (2018) od fylogeneticky nejmladších po fylogeneticky nejstarší.