

BAKALÁŘSKÁ PRÁCE

Michal Hercík

Webový plugin pro vizualizaci sady sekundárních struktur RNA

Katedra softwarového inženýrství

Vedoucí bakalářské práce: doc. RNDr. David Hoksza, Ph.D.

Studijní program: Informatika

Studijní obor: Programování a vývoj software

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů. Tato práce nebyla využita k získání jiného nebo stejného titulu.
Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.
V dne
Podpis autora

TODO: podekovani

Název práce: Webový plugin pro vizualizaci sady sekundárních struktur RNA

Autor: Michal Hercík

Katedra: Katedra softwarového inženýrství

Vedoucí bakalářské práce: doc. RNDr. David Hoksza, Ph.D., Katedra softwaro-

vého inženýrství

Abstrakt: TODO

Klíčová slova: bioinformatika RNA sekundární struktura web plugin

Title: Web plugin for multiple RNA secondary structure visualization

Author: Michal Hercík

Department: Department of Software Engineering

Supervisor: doc. RNDr. David Hoksza, Ph.D., Department of Software Enginee-

ring

Abstract: TODO

Keywords: bioinformatics RNA secondary structure web plugin

Obsah

Ú٦	vod		3
1	Úvo	od do problematiky	4
	1.1	Seznámení s biologickými pojmy	4
		1.1.1 RNA	4
	1.2	Vizualizace sekundárních RNA struktur	4
	1.3	Podobné projekty	6
		1.3.1 VARNA	6
		1.3.2 RNAStructViz	7
		1.3.3 Forna	8
		1.3.4 R-chie	9
		1.3.5 Shrnutí existujících nástrojů	9
	1.4	Kreslení grafů na základě šablony	10
		1.4.1 TRAVeLer	10
		1.4.2 R2DT	10
2	Met	tody vizualizace a porovnání	11
	2.1	Překládání struktur	11
	2.2	Transformace na vzor	14
	2.3	Mapovací čáry	14
	2.4	Demonstrace metod	16
3	Dok	kumentace	17
	3.1	Vstupní data	17
	3.2	Objektový návrh	19
		3.2.1 Obecný pohled na třídy	19
	3.3	Uživatelská dokumentace	22
		3.3.1 Instalace	22
		3.3.2 Rozhraní: IAnimation	$\frac{-}{22}$
		3.3.3 Příklady	$\frac{-}{22}$
	3.4	Vývojová dokumentace	$\frac{-}{22}$
	0.1	3.4.1 Volba technologií	22
		3.4.2 Verzování	23
		3.4.3 Rozbor implementace	23
Zá	ivěr		24
		Y	
Se	znan	n použité literatury	25
Se	znan	n obrázků	26
Se	znan	n tabulek	27
\mathbf{Se}	znan	n použitých zkratek	28

\mathbf{A}	A Přílohy														2	9																	
	A.1	První příloha																														2	C

$\mathbf{\acute{U}vod}$

TODO: uvod

1. Úvod do problematiky

1.1 Seznámení s biologickými pojmy

1.1.1 RNA

Rozdělit do více kapitol. Jedna z kapitol struktura RNA.

RNA (zkratka z anglického ribonucleic acid) je biomolekula, která hraje klíčovou roli v procesu přenosu genetické informace u všech živých organismů. RNA se skládá z řetězce nukleotidů, které obsahují cukr ribózu, fosfátovou skupinu a jednu z pěti dusíkatých bází (adenin, guanin, cytosin, uracil nebo inosin). Existují různé typy RNA, jako jsou messenger RNA (mRNA), ribozomální RNA (rRNA) a transfer RNA (tRNA), které mají každý svou specifickou funkci v buňce.

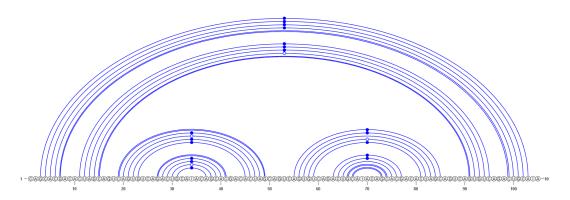
RNA sekundární struktura se týká způsobu, jakým se molekula RNA skládá na sebe díky vzniku bázových párů mezi komplementárními nukleotidy. Bázové párování se děje mezi dusíkatými bázemi RNA nukleotidů, přičemž adenin (A) se páruje s uracilem (U) a guanin (G) se páruje s cytosinem (C).

RNA sekundární struktura je důležitá, protože může ovlivnit to, jak RNA molekula funguje. Například stem-loop struktura v mRNA molekule může ovlivnit přístupnost mRNA k ribozomům, což je buněčný mechanismus zodpovědný za překládání mRNA na proteiny.

1.2 Vizualizace sekundárních RNA struktur

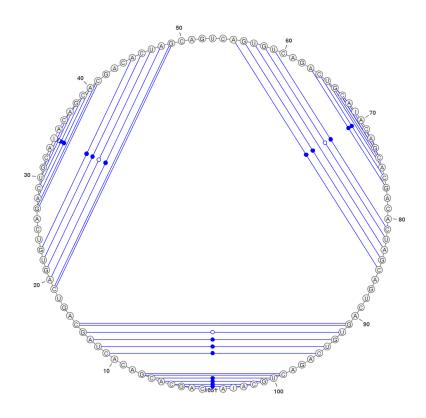
Pro reprezentaci sekundární RNA struktury se používají jak textové, tak grafické způsoby. Pro nás jsou nejzajímavějsí ty grafické, ze kterých v této části představíme tři nejpoužívanější - arc diagram, circular diagram a radiate diagram. Obrázky ukázek diagramu v této části jsou získané za pomoci nástroje VARNA[2].

V arc diagramu jsou nukleotidy zobrazeny na rovné čáře ve stejném pořadí jako v sekvenci a bázové páry nukleotidů jsou spojeny obloukem.



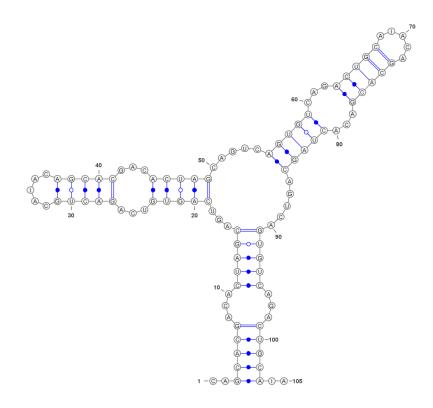
Obrázek 1.1: Ukázka arc diagramu

Circular diagram je velmi podobný. Nukleotidy neleží na rovné čáře, ale po obvodu kruhu. Bázové páry jsou spojeny buď čárou nebo obloukem.



Obrázek 1.2: Ukázka circular diagramu

Obě tyto reprezentace postrádají schopnost zachytit motivy sekundární struktury, a proto se radiate diagram používá tam, kde je potřeba detailní vizuální analýza motivů sekundární RNA struktury a její interakce. V radiate diagramu jsou pozice nukleotidů voleny tak, aby bylo možné rozeznat motivy sekundární struktury, jako jsou hairpins, bulges nebo vícevětvené smyčky.



Obrázek 1.3: Ukázka radiate diagramu

1.3 Podobné projekty

chybi motivacni uvod, ktery by rekl, co chceme delat, aby bylo mozne pochopit podobne k cemu

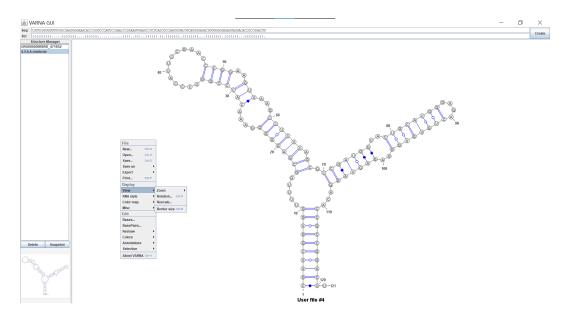
Rádi bychom čtenáře seznámili s některými nástroji, které jsou používáné pro vizualizaci sekundárních RNA struktur. Většina z nich jsou programy s uživatelským rozhraním a mohlo by se proto zdát zbytečné je zmiňovat nebo porovnávat s naší knihovnou. Nicméně u níže zmíněných programů není duležité řešení samotného uživatelského rozhraní, jako především druh zvolených metod pro vizualizaci a následné porovnávání.

Z velkého množství existujících nástrojů, byla snaha vybrat takové, které mají rozdílné přístupy a nabízí nejširší paletu funkcí.

1.3.1 VARNA

VARNA (Visualization Applet for RNA) je nástroj pro automatické kreslení, vizualizaci a anotaci sekundárních RNA struktur, navržený jako doprovodný software pro webové servery a databáze.

VARNA implementuje algoritmy pro vykreslení všech tří výše zmíněných diagramů, podporuje různé textové formáty pro vstup i výstup a je schopný exportovat kresbu do rastrových nebo vektorových formátů. Umožňuje ruční úpravy a strukturální anotace výsledku kresby a je považován za standard pro práci se sekundárními strukturami RNA.

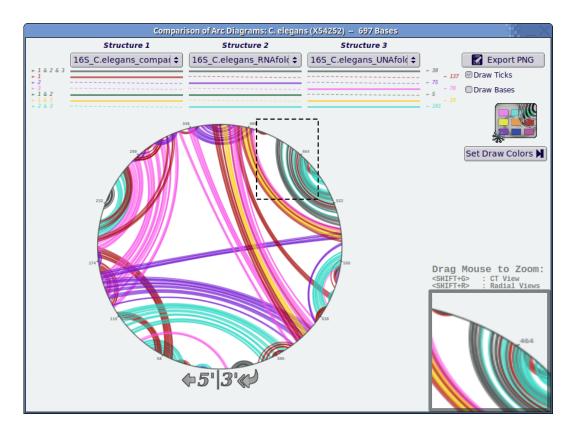


Obrázek 1.4: Snimek nástroje Varna. Zobrazená struktura je d.5.b.A.madurae.

1.3.2 RNAStructViz

RNAStructViz[1] je grafický nástroj pro analýzu sekundárních RNA struktur. Jeho předností je vizuální porovnání tří konfigurací v circular arc diagramu. Doplněné zabudovaným prohlížečem CT-style¹ souboru a prohlížečem radial diagramu podstruktury, která je přímo propojená s arc diagram oknem skrze nástroj pro výběr zoom. Mezi další funkce patří vypočítání číselných informací a možnost exportu obrázků a dat pro pozdější použití.

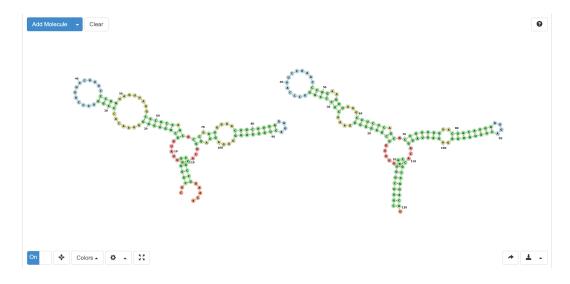
 $[\]overline{^{1}\mathrm{CT}}$ formát souboru slouží k ukládání informace o sekvenci a bázových párů.



Obrázek 1.5: Snimek nástroje rnaStructViz, zobrazující tři struktury RNA.

1.3.3 Forna

Forna[4] (force-directed rna) nabízí webové rozhraní a server, který umožňuje uživateli vložit sekundární RNA strukturu ve formátu dot-bracket a zobrazí ji jako force-directed graf². Uživatel může následně upravit pozice přetažením myší a lze i upravovat přímo strukturu.



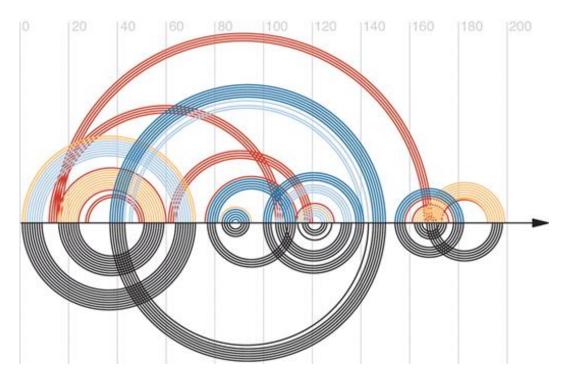
Obrázek 1.6: Snimek nástroje Forna. Nalevo odvozená sekundární RNA struktura URS00000B9D9D_471852 od struktury d.5.b.A.madurae napravo.

²https://cs.brown.edu/people/rtamassi/gdhandbook/chapters/force-directed.pdf

1.3.4 R-chie

R-chie [5] je web server, který umí vygenerovat šest různých typů arc diagramu. Vývoj tohoto nástroje byl se zaměřením především na složitější struktury, které nelze hezky nakreslit v radial diagramu. R-chie umí vygenerovat diagram pro porovnávání dvou sekundárních RNA struktur. Důležitým cílem byla možnost generovat diagramy pro velké množství dat, proto také nenabízí grafické rozhraní a s ním spojenou interakci se strukturami.

Projekt také nabízí balíček napsaný v jazyce R³ zvaný R4RNA, který umožňuje spuštění programu lokálně a napříč operačním systémům.



Obrázek 1.7: Výsledný arc diagram nástroje R-chie, zobrazující dvě sturktury. První struktura je nad horizontální čárou a druhá pod ní.

1.3.5 Shrnutí existujících nástrojů

Nástroje představené v této kapitole se soustředí především na prácí s circular diagramem nebo arc diagramem, a právě pouze pro tyto diagramy nabízí nějaké metody pro porovnávání omezeného množství sekundárních struktur RNA. Forna podporuje pouze radial diagram, ale porovnávání dvou struktur, které sice jdou zobrazit vedle sebe, už nijak neusnadňuje.

Varna Podporuje všechny tři zmíněné diagrami, ale nelze ani zobrazit dvě sekundární rna struktury vedle sebe. Velkou výhodou nástroje VARNA by byla možnost použití na webu, ale k tomu používá Java Applets ⁴, které jsou od roku 2017 považované za zastaralé ⁵.

Ze zmíněných projektů je nejpodobnější tomu našemu R-chie, který se snaží usnadnit porovnávání sekundárních RNA struktur a nabízí i knihovnu napsanou

³https://www.r-project.org/

⁴https://docs.oracle.com/javase/tutorial/deployment/applet/index.html

⁵https://www.oracle.com/java/technologies/javase/9-deprecated-features.html

v jazyce R. Liší se pak v samotném přístupu, protože jejich rozhraní generuje pouze statické circular nebo arc diagramy.

1.4 Kreslení grafů na základě šablony

Níže jsou zmíněné dva projekty, které úzce souvisí s naší knihovnou, protože produkují data ve formátu, se kterým pracuje naše knihovna a metody použité ke generovaní takových dat jsou klíčové pro naší knihovnu.

1.4.1 TRAVeLer

Traveler[3] je nástroj pro vizualizaci cílové sekundární struktury, využívající existující rozložení dostatečně podobné RNA struktury jako vzor. Traveler je založený na algoritmu, který konvertuje cílovou a vzorovou strukturu do odpovídající stromové reprezentace a využije stromovou editační vzdálenost společně s modifikací rozložení k přetvoření vzorové struktury do cílové. Traveler přijme na vstupu sekundární strukturu a vzor rozložení a na výstupu dá rozložení cílové struktury. Je to tedy command-line open source nástroj schopný rychle generovat rozložení i pro největší RNA struktury za poskytnutí dostatečně podobného rozložení.

Do vzniku Traveleru neexistoval žádný nástroj, který by dokázal velké struktury vizualizovat ve standardni notaci, se kterou jsou biologové naučení pracovat a porovnávat struktury napříč druhům.

1.4.2 R2DT

R2DT[6] je metoda pro predikci a vizualizaci široké škály sekundárních RNA struktur ve radial diagramu. R2DT je postaveno na knihovně se 3 647 vzory reprezentujícími většinu známých RNA struktur. R2DT se používá na ncRNA⁶ (non-coding RNA) sekvencích z RNAcentral⁷ databáze a vytvořila více než 27 miliónů diagramů⁸, čímž tvoří největší světovou sadu dat s 2D RNA strukturami. Pro vizualizaci neboli 2D rozložení používá R2DT právě výše zmíněný nástroj Traveler.

⁶RNA, která se nepřekládá do proteinů

⁷https://rnacentral.org/

⁸Číslo je aktuální k datu 11.4. 2023

2. Metody vizualizace a porovnání

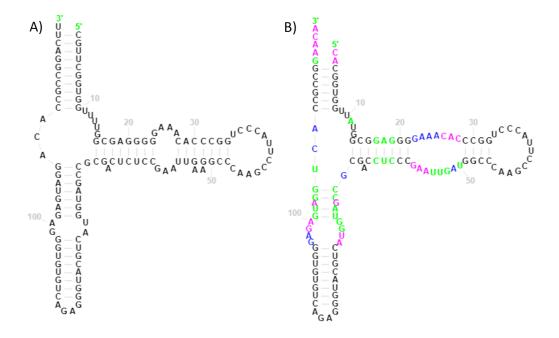
Cílem této knihovny není pouze vizualizovat sekundární strukturu RNA, ale také usnadnit analýzu rozdílů a podobností mezi více strukturami RNA. Proto jsme se zaměřili na práci s radiálními diagramy.

Kromě toho jsme rozpoznali potenciál generování distribuce na základě vzorové struktury, jak to dělá nástroj Traveler. Výstupem Traveleru je soubor ve formátu JSON, který obsahuje informace o vzoru každého nukleotidu a provedených úpravách.

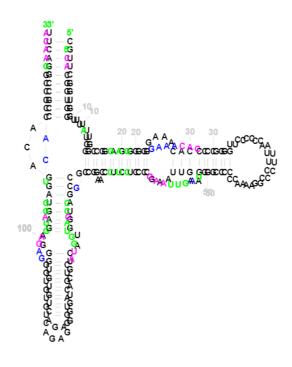
 ${\bf V}$ důsledku toho jsme se rozhodli používat v našich metodách zmíněné mapování na referenční strukturu. Náš nástroj je tudíž schopný pracovat s N strukturami, mezi nimiž je referenční struktura, ze které jsou generovány všechny ostatní struktury.

2.1 Překládání struktur

Vzhledem k tomu, že struktury odvozené od stejného vzoru jsou obvykle velmi podobné, má smysl umožnit jejich překrytí, aby se spojily společné části a zvýraznily se rozdíly. Pouhým přeložením podobných struktur přes sebe však získáme výsledek, který neposkytuje příliš zajímavé informace a je málo přehledný.



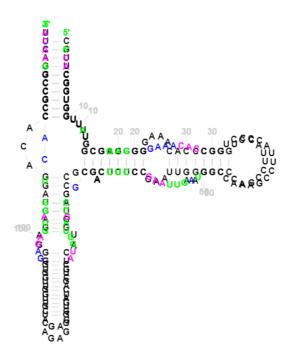
Obrázek 2.1: Struktury vedle sebe.



Obrázek 2.2: Struktury přeložené přes sebe.

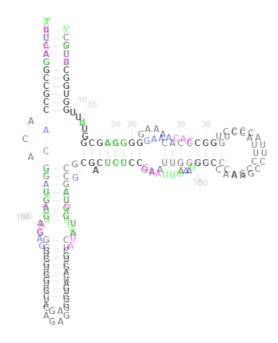
Je potřeba najít způsob, jak řešit problém přesunu a zarovnání struktury, protože manuální manipulace pomocí přetažení myší nebo zadávání pozice může být zbytečně obtížná, zejména pokud se snažíme dosáhnout přesného zarovnání. Proto je velmi užitečné umožnit zarovnání sekundární RNA struktury na konkrétní nukleotid nebo skupinu nukleotidů ze vzorové struktury.

Pokud je vybrán vzorový nukleotid v pro určitý nukleotid n z ostatních struktur, je celá struktura přesunuta tak, aby se nukleotidy v a n překrývaly. Pokud se snažíme najít větší skupinu nukleotidů, na které chceme strukturu zarovnat, koukáme se na posunutí pro konkrétní nukleotidy a následně vybereme takové posunutí které zarovná nejvíc nukleotidů.



Obrázek 2.3: Struktury přeložené přes sebe a zarovnané.

Pouhým přeložením struktur nelze snadno rozeznat, které nukleotidy jsou společné a překrývají se a které nejsou. Přidáním průhlednosti je možné tuto situaci rozlišit, protože překrývající se nukleotidy budou mít sytější barvu než ty, které se nepřekrývají.



Obrázek 2.4: Struktury přeložené přes sebe, zarovnané a s průhledností.

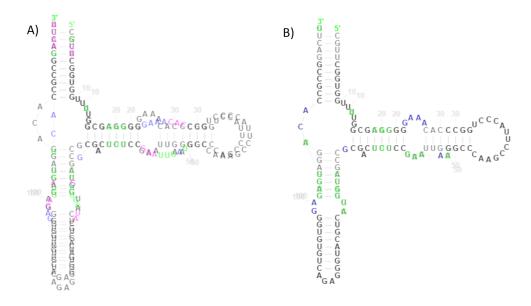
Zarovnávání struktur bohužel neřeší všechny výzvy. Výsledné obrázky se mohou zdát rozmazané. To je způsobeno tím, že ačkoli má nukleotid vzorový nukleotid, který je stejný, jeho pozice se může v rozložení mírně lišit v důsledku metody

generování dat. Tento fakt může způsobit, že popisky nukleotidů vypadají rozmazaně.

Jako přímočaré řešení se může zdát posunout jednotlivé nukleotidy, které jsou blízko sebe, aby dokonale překrývaly jejich vzory. Věříme, že by to vyřešilo zmíněný problém. Nicméně naše knihovna tuto funkci přímo neimplementuje.

2.2 Transformace na vzor

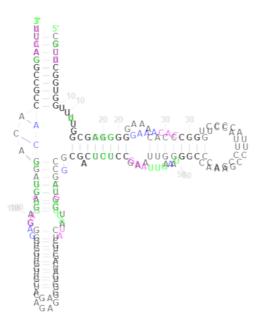
Užitečnou metodou je transformace mezi vzorovou a cílovou strukturou. Každý nukleotid, který má svůj vzorový nukleotid, se přemístí na pozici vzorového nukleotidu a nukleotidy, které ve vzoru nejsou, jsou skryté. Tato metoda je velmi užitečná pro práci s dvěma strukturami, které jsou si podobné, nebo pro získání počátečního přehledu o tom, co je na co namapováno. Slabou stránkou této metody je její použití při práci s více než dvěma strukturami nebo strukturami, které jsou velmi odlišné. V takových situacích se na obrazovce děje mnoho věcí a je obtížné se soustředit a vypozorovat něco užitečného.



Obrázek 2.5: Vzorová a odvozená struktura přeložené přes sebe před transformací (A) a po transformaci (B).

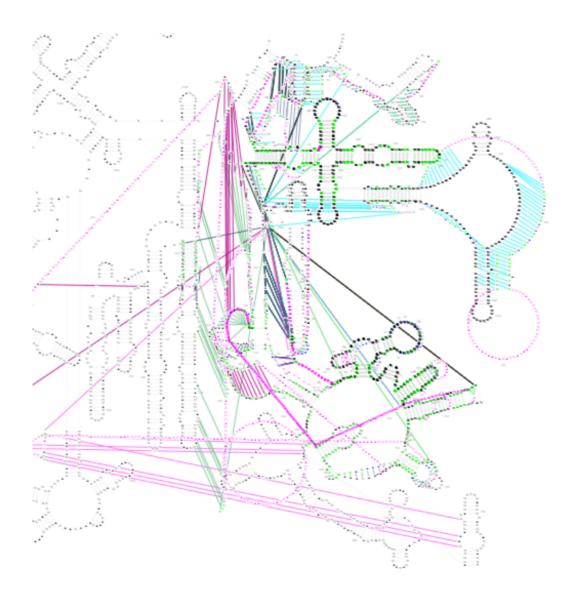
2.3 Mapovací čáry

Vědomost o tom, který nukleotid se na co mapuje, může být významná pro odhalení rozdílů a podobností mezi strukturami. V našem úsilí zprostředkovat tuto informaci již před transformací na vzorovou sturkturu jsme přišli s dalším způsobem - s čárami, které spojují každý nukleotid s jeho vzorovým nukleotidem.



Obrázek 2.6: Vzorová a odvozená struktura přeložené přes sebe s mapovacíma čárama.

Bohužel tento způsob se zvětšující se velikostí struktury stáva velmi nepřehledným, přesto si myslíme že můžou být užitečné a naše knihovna je podporuje.



Obrázek 2.7: Výřez z mnoha struktur přeložených přes sebe s mapovacími čárami. Každá struktura má vlastní barvu mapovacích čar.

2.4 Demonstrace metod

V rámci naší knihovny jsme vytvořili také webovou aplikaci¹, která demonstruje možnosti naší knihovny. Tato aplikace umožňuje uživatelům pracovat s jednotlivými metodami pro usnadnění porovnání zmíněnými v této kapitole nebo souběžně se všemi.

 $^{^{1}}$ https://michalhercik.github.io/rna-visualizer/

3. Dokumentace

3.1 Vstupní data

Jak jsme již zmiňovali, naše knihovna využívá výstupní data nástroje Traveler jako vstupní data. Jedná se o data ve formátu JSON, obsahující všechny potřebné informace o rozložení nukleotidů, jejich párování, velikostech popisků, barvách a tlouštkách čar. Kromě informací o rozložení obsahuje také informace o potřebných editacích vzorové sekundární struktury.

V rámci R2DT projektu vzníká i JSON schéma¹, které by mělo popisovat strukturu vstupních dat. Schéma je stále ve vývoji, proto aktuální výstupy R2DT nebo Traveleru neodpovídájí schématu a je dost možné, že se jejich výstupy budou v budoucnu měnit a naše knihovna se jim bude přizpůsobovat, TODO: přepsatprotože RNAcentral, využívající R2DT, je největší databází s 2D RNA strukturama.

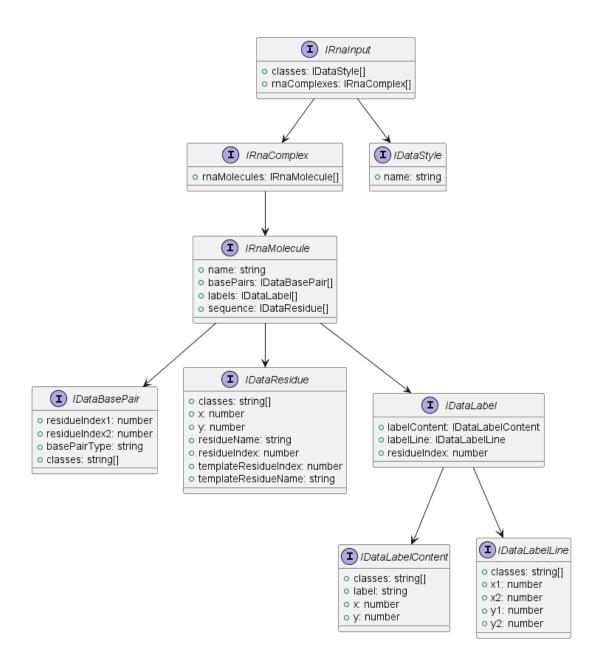
Samotná struktura dat není složitá, ale popíšeme zde pouze tu část, kterou aktuálně využíváme, kromě toho, že ostatní data pro nás nejsou duležitá, tak jak již bylo zmíněno samotná struktura dat není pevně daná a může se měnit.

Jedná se o objekt, který má dvě položky - classes, což je pole objektů popisující třídy říkající způsob zobrazení struktury, podobně jako to kaskádové styly (CSS) diktují pro webové stránky a rnaComplexes.

rnaComplexes je pole polí sloužící pro popis celých skupin RNA struktur. Naše knihovna pracuje vždy pouze s nultým prvkem. Neviděli jsme důvod to dělat jinak, a pokud by se nějaký důvod našel v budoucnu, neměl by být problém naší knihovnu přizpůsobit situaci (např. rozšířením o novou metodu pro zachování zpětné kompatibility).

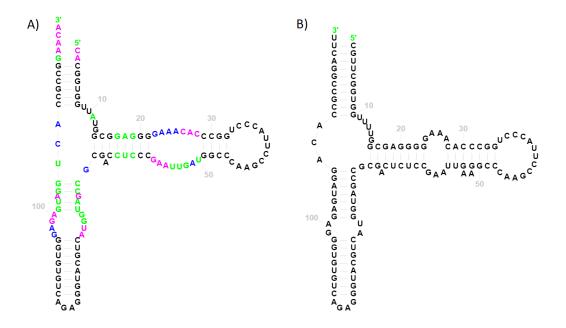
V rámci naší knihovny jsme vytvořili interface, který vstupní data musí splňovat. Struktura zbytku dat by měla být jasně viditelná z následujícího diagramu těchto interfaceů.

¹https://github.com/LDWLab/RNA2D-data-schema



Obrázek 3.1: Interface pro vstupní data

Při vykreslení dat si můžeme všimnout různého obarvení jednotlivých residue. Tyto barvy slouží k lepšímu zorientování ve struktuře vzhledem ke vzorové struktuře. Černá barva značí, že residue leží na poloze vzorového residue se stejným názvem. Zelenou barvou jsou označený ty residue jejiž vzorový residue bylo třeba přejmenovat. Modrou barvou jsou vyznačený posunutý residue. A poslední růžovou barvu mají nově přidaný residue.



Obrázek 3.2: A) Odvozená sekundární RNA struktura URS00000B9D9D_471852, B) Vzorová sekundární RNA struktura d.5.b.A.madurae

3.2 Objektový návrh

Udělat dobrý objektový návrh je pro knihovnu, která má usnadňovat práci velmi důležité, zároveň je tento úkol velmi těžký, speciálně v případě kdy není předem jasné, co všechno má knihovna umět. To v našem případě nebylo, a proto se některá rozhodnutí mohou zdát zpětně zvláštní, při nejmnenším ne ideální.

TODO: přepsatV následující kapitole se pokusíme čtenáře seznámit s objektovým návrhem naší knihovny včetně s myšlenkama, které nás k takovému návrhu dovedli. Nejdříve dáme čtenáři obecný pohled na strukturu a následně rozebere jednotlivé třídy podrobněji.

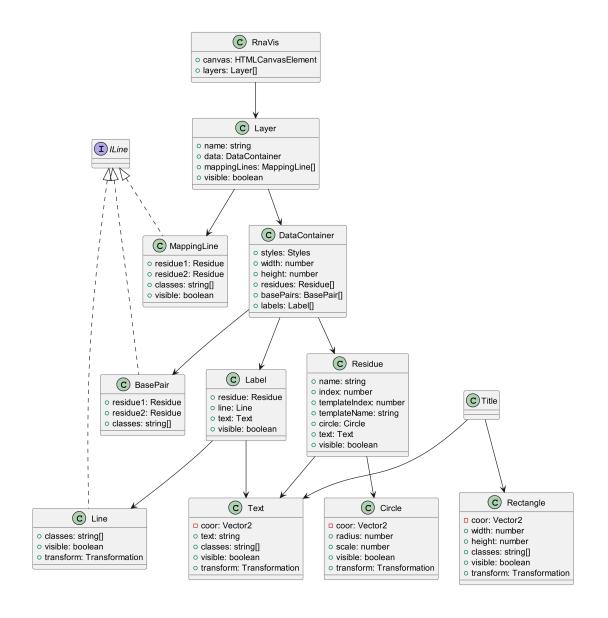
3.2.1 Obecný pohled na třídy

Srdcem celé knihovny je třída RnaVis, která vykresluje struktury na canvas a nastavuje se přes ní zoom/panning. Kromě toho má v sobě také uložené vrstvy realizované třídou Layer, představující jednotlivé struktury.

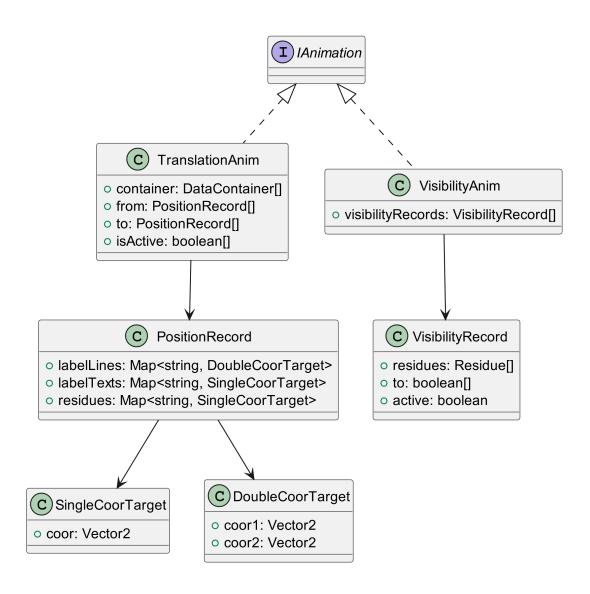
Pro určení vykreslovacích parametrů (např. font, barva, velikost) objektů máme třídu Styles. Před každým vykreslení se zeptáme této třídy na vykreslovací parametry pro daný objekt.

Poslední důležitou velkou částí jsou animace. Tuto funkcionalitu zpřístupňují dvě hlavní třídy - TranslationAnim, VisibilityAnim. Obě implementují interface IAnimation.

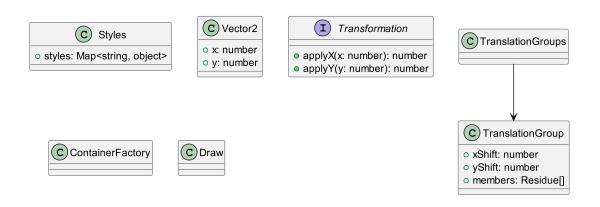
Níže jsou tři diagramy tříd, které dohromady obsahují každou třídu. Diagramama se snažíme vyjádřit obecnou strukturu, tím pádem pro přehlednost neobsahují všechny informace - všechny metody, některé privátní vlastnosti a některé závislosti.



Obrázek 3.3: Diagram tříd



Obrázek 3.4: Diagram tříd



Obrázek 3.5: Diagram tříd

3.3 Uživatelská dokumentace

3.3.1 Instalace

3.3.2 Rozhraní: IAnimation

Rozhraní pro definici animace.

Implementováno

- TranslationAnim
- VisibilityAnim

Metody

animate

animate(rna, duration, after): void Provede animaci.

Parametry

Název	Тур	Popis
rna	RnaVis	Objekt RnaVis, na kterém se má animace
		provést
duration	number	Délka animace
after	AfterFn	Funkce, která se má zavolat po dokončení
		animace

3.3.3 Příklady

Vykreslení struktur

Transformace

3.4 Vývojová dokumentace

3.4.1 Volba technologií

Programovací jazyk

Volba programovacího jazyka byla poměrně přímočará. Chtěli jsme napsat knihovnu, která se bude používat na webu. Javascript² je v tomto případě jasnou volbou, protože to je v podstatě to jediné, co se používá. Přesto Javascript není jazyk, ve kterým je naše knihovna psaná, protože se jedná o dynamicky typovaný jazyk, což s sebou nese určité výhody pro jednoduché a rychlé psaní kódu, ale u větších projektů se to stává nevýhodou. Naše knihovna je napsaná v Typescriptu³, což je nadmnožina Javascriptu, snažící se řešit jeho slabiny a navíc ho

²https://developer.mozilla.org/en-US/docs/Web/JavaScript

³https://www.typescriptlang.org/

lze snadno přeložit do Javasrciptu a v této formě distribuovat. Při tomto překladu je možné zachovat i informaci o konkrétních typech, tím pádem projekty, využívající naší knihovnu, psané v Typescriptu nejsou ochuzené o typovou kontrolu, kterou typescript nabízí.

3.4.2 Verzování

Chtěli jsme mít možnost se vracet ke starším verzím projektu v případě, že něco rozbijeme, nebo by nás pouze zajímali implementace, které už nepoužíváme. Pro verzování používáme Git⁴, který lze jednoduše používat s GitHubem⁵. Díky kterému, je náš projekt přístupný odkukoliv.

Jediným důvodem pro tato rozhodnutí je, že obě technologie známe, nemáme s nimi problém a neviděli jsme tedy důvod hledat jiné možnosti.

Knihovna D3.js

D3.js⁶ je knihovna v jayzce Javascript pomáhající přivést data k životu využívající především SVG⁷ formát, se kterým se dá v HTML⁸ pohodlně pracovat. Její důraz na webové standardy dává uživateli možnost využívat moderní prohlížeče naplno bez dalších frameworků. S knihovnou není nejjednodušší se naučit pracovat, ale její velkou předností je rychlost.

SVG nebo canvas

V počátcích jsme chtěli k zobrazování používat SVG. Jedná se o webový standard, který lze skvěle kombinovat s ostatními standardy jako je CSS⁹, DOM¹⁰, Javascript. V kombinací s D3.js knihovnou je pak dělání animací nebo aktualizaci stavu SVG objektů jednoduché.

Nebyl důvod SVG nevyužívat, ale později jsme na vlastní kůži pocítili slabinu SVG. SVG se při vykreslování tisícovek objektů stává velmi pomalé. Takového počtu objektů můžeme dosáhnout pouze s jednou sekundární RNA strukturou, my navíc chceme zobrazit více takových struktur a ještě s nimi dynamicky pracovat.

Tím se pro nás stalo SVG nepoužitelné. Další možností bylo využití canvasu¹¹, který slibuje výrazně lepší výkon a lze ho stále jednoduše používat.

Některé pro nás klíčové funkce D3.js knihovny lze využít i pro práci s canvasem. Konkrétně se jedná o zoom, panning a animace.

U canvasu jsme nakonec i zůstali, přestože při práci s desítkama velkých struktur vykreslování není plynulé.

3.4.3 Rozbor implementace

 $^{^4}$ https://git-scm.com/

⁵https://github.com/

⁶https://d3js.org/

⁷https://developer.mozilla.org/en-US/docs/Web/SVG

⁸https://developer.mozilla.org/en-US/docs/Learn/Getting_started_with_the_web/HTML_basics

⁹https://developer.mozilla.org/en-US/docs/Web/CSS

 $^{^{10} \}rm https://developer.mozilla.org/en-US/docs/Web/API/Document_Object_Model$

¹¹https://developer.mozilla.org/en-US/docs/Web/API/Canvas_API

Závěr

Seznam použité literatury

- [1] CHENNEY, S., HEITSCH, C., MIZE, C., SWENSON, S., SCHMIDT, M. D., KIRKPATRICK, A. a YOON, I. (2019). Rnastructviz. URL https://github.com/gtDMMB/RNAStructViz/wiki. Accessed on March 29, 2023.
- [2] Darty, K., Denise, A. a Ponty, Y. (2009). VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**(15), 1974–1975.
- [3] ELIAS, R. a HOKSZA, D. (2017). Traveler: a tool for template-based rna secondary structure visualization. *BMC Bioinformatics*, **18**(1), 487. ISSN 1471-2105. doi: 10.1186/s12859-017-1885-4. URL https://doi.org/10.1186/s12859-017-1885-4.
- [4] KERPEDJIEV, P., HAMMER, S. a HOFACKER, I. L. (2015). Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams. *Bioinformatics*, **31**(20), 3377–3379. ISSN 1367-4803. doi: 10.1093/bioinformatics/btv372. URL https://doi.org/10.1093/bioinformatics/btv372.
- [5] LAI, D., PROCTOR, J. R., ZHU, J. Y. A. a MEYER, I. M. (2012). R-chie: a web server and R package for visualizing RNA secondary structures. *Nucleic Acids Research*, 40(12), e95–e95. ISSN 0305-1048. doi: 10.1093/nar/gks241. URL https://doi.org/10.1093/nar/gks241.
- [6] SWEENEY, B. A., HOKSZA, D., NAWROCKI, E. P., RIBAS, C. E., MADE-IRA, F., CANNONE, J. J., GUTELL, R., MADDALA, A., MEADE, C. D., WILLIAMS, L. D., PETROV, A. S., CHAN, P. P., LOWE, T. M., FINN, R. D. a PETROV, A. I. (2021). R2dt is a framework for predicting and visualising rna secondary structure using templates. Nature Communications, 12(1), 3494. ISSN 2041-1723. doi: 10.1038/s41467-021-23555-5. URL https://doi.org/10.1038/s41467-021-23555-5.

Seznam obrázků

1.1	Ukázka arc diagramu	4
1.2	Ukázka circular diagramu	5
1.3	Ukázka radiate diagramu	6
1.4	Snimek nástroje Varna. Zobrazená struktura je d.5.b.A.madurae	7
1.5	Snimek nástroje rnaStructViz, zobrazující tři struktury RNA	8
1.6	Snimek nástroje Forna. Nalevo odvozená sekundární RNA struktura URS00000B9D9D_471852 od struktury d.5.b.A.madurae na-	
	pravo	8
1.7	Výsledný arc diagram nástroje R-chie, zobrazující dvě sturktury.	
	První struktura je nad horizontální čárou a druhá pod ní	9
2.1	Struktury vedle sebe	11
2.2	Struktury přeložené přes sebe	12
2.3	Struktury přeložené přes sebe a zarovnané	13
2.4	Struktury přeložené přes sebe, zarovnané a s průhledností	13
2.5	Vzorová a odvozená struktura přeložené přes sebe před transfor-	
	mací (A) a po transformaci (B)	14
2.6	Vzorová a odvozená struktura přeložené přes sebe s mapovacíma	
	čárama	15
2.7	Výřez z mnoha struktur přeložených přes sebe s mapovacími čá-	
	rami. Každá struktura má vlastní barvu mapovacích čar	16
3.1	Interface pro vstupní data	18
3.2	A) Odvozená sekundární RNA struktura URS00000B9D9D_471852,	
	B) Vzorová sekundární RNA struktura d.5.b.A.madurae	19
3.3	Diagram tříd	20
3.4	Diagram tříd	21
3.5	Diagram tříd	21

Seznam tabulek

Seznam použitých zkratek

A. Přílohy

A.1 První příloha