



Agenda

- Zarys
- Zakres pracy
- Założenia i ograniczenia
- Diagram użycia aplikacji
- Technologie
- Architektura
- Algorytm
- Eksperyment obliczeniowy

Zarys

- ▢ Wprowadzenie DNA
- ▢ Mutageneza

Minusy:

Losowość miejsca modyfikacji DNA

Liczne kopie niektórych genów

Modyfikacje w komórkach zarodkowych

Nieodwracalność zmian, niewrażliwość
na zmianę warunków

- ▢ „Wyciszenie” genów

„Wyciszanie” genów

Krótki fragment RNA

- Przyłączony do mRNA powstałego genu, stanowi przeszkodę lub kieruje do degradacji

Zalety:

możliwość pracy z org. dojrzałymi
działa na wszystkie kopie
utrzymanie org. przy życiu
wyciszenie trwałe lub przejściowe

Wady:

precyzja

Krótkie RNA

- MicroRNA
 - 20-25 elementowe odcinki RNA
 - U roślin zdefiniowano ponad 4500:
- Rola
 - Różnicowanie organów, morfologia liścia, tożsamość organów, własna biogeneza, ...

artificial microRNA- amiRNA

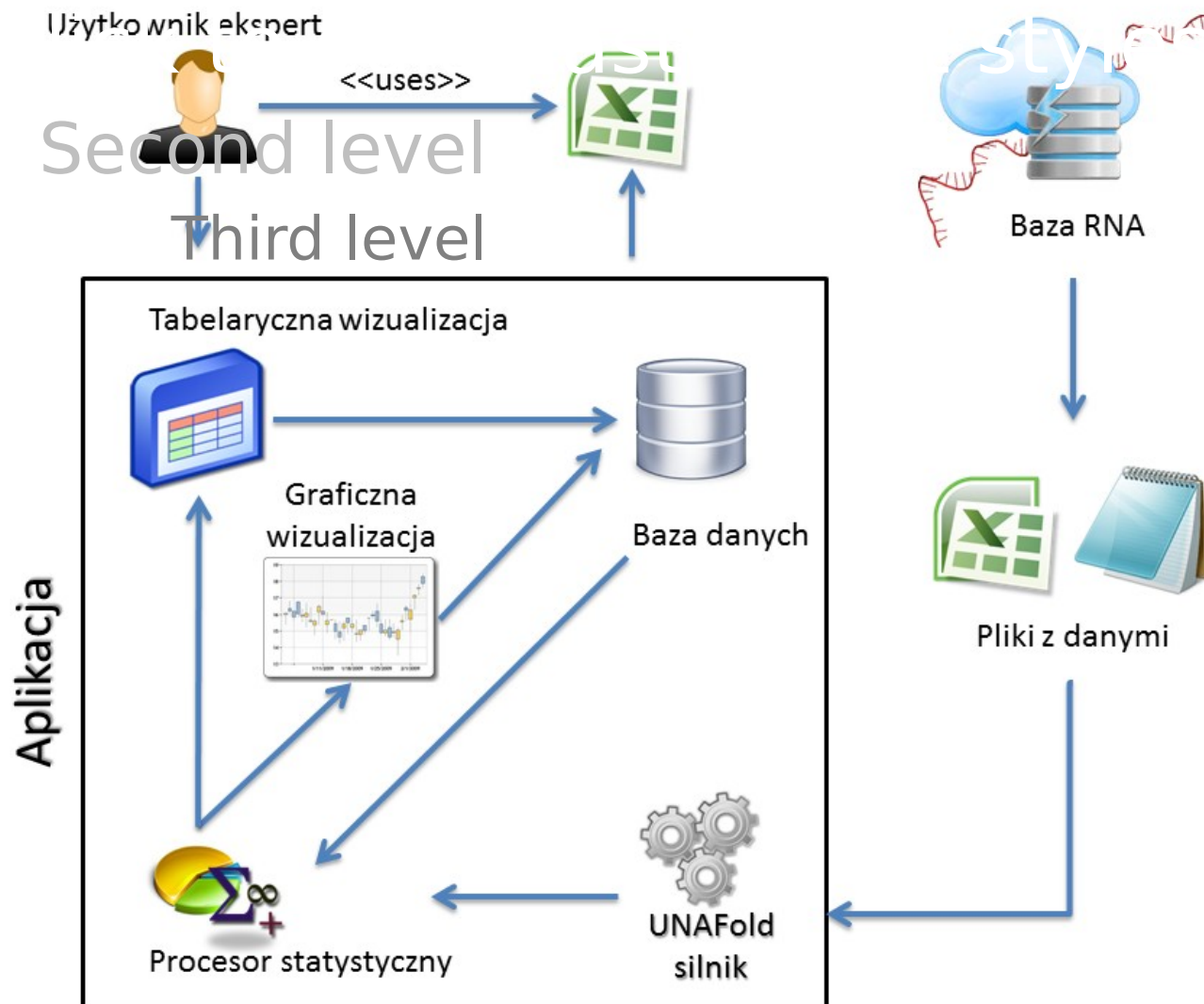
Zakres pracy

- Opracowanie koncepcji algorytmu
umożliwiającego projektowanie sztucznych
cząsteczek małych regulatorowych RNA
- Implementacja aplikacji wykonującej
algorytm
- Eksperyment obliczeniowy
- Przygotowanie bazy danych

Założenia/ograniczenia



Diagram użycia aplikacji



Technologie

- ▢ Aplikacja desktopowa

- ▢ .NET 4.0

- ▢ WPF

- ▢ Perl, ActivePerl

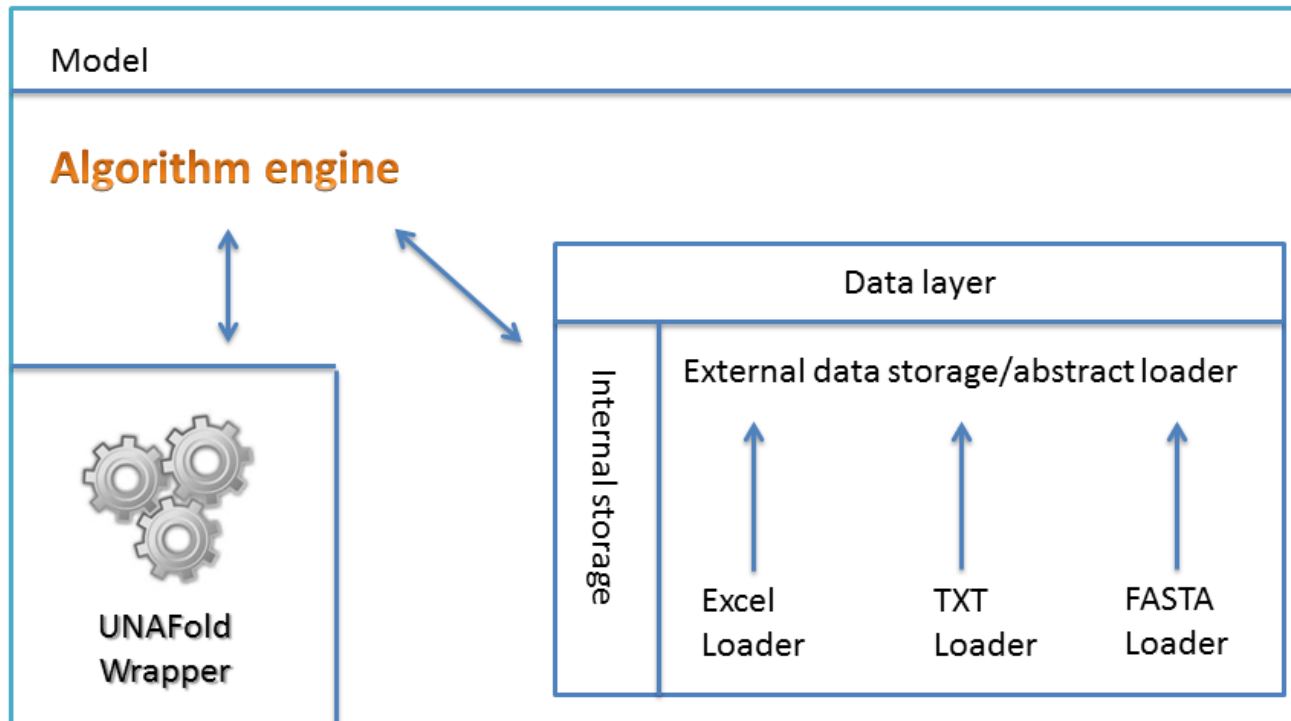
Architektura - MVVM



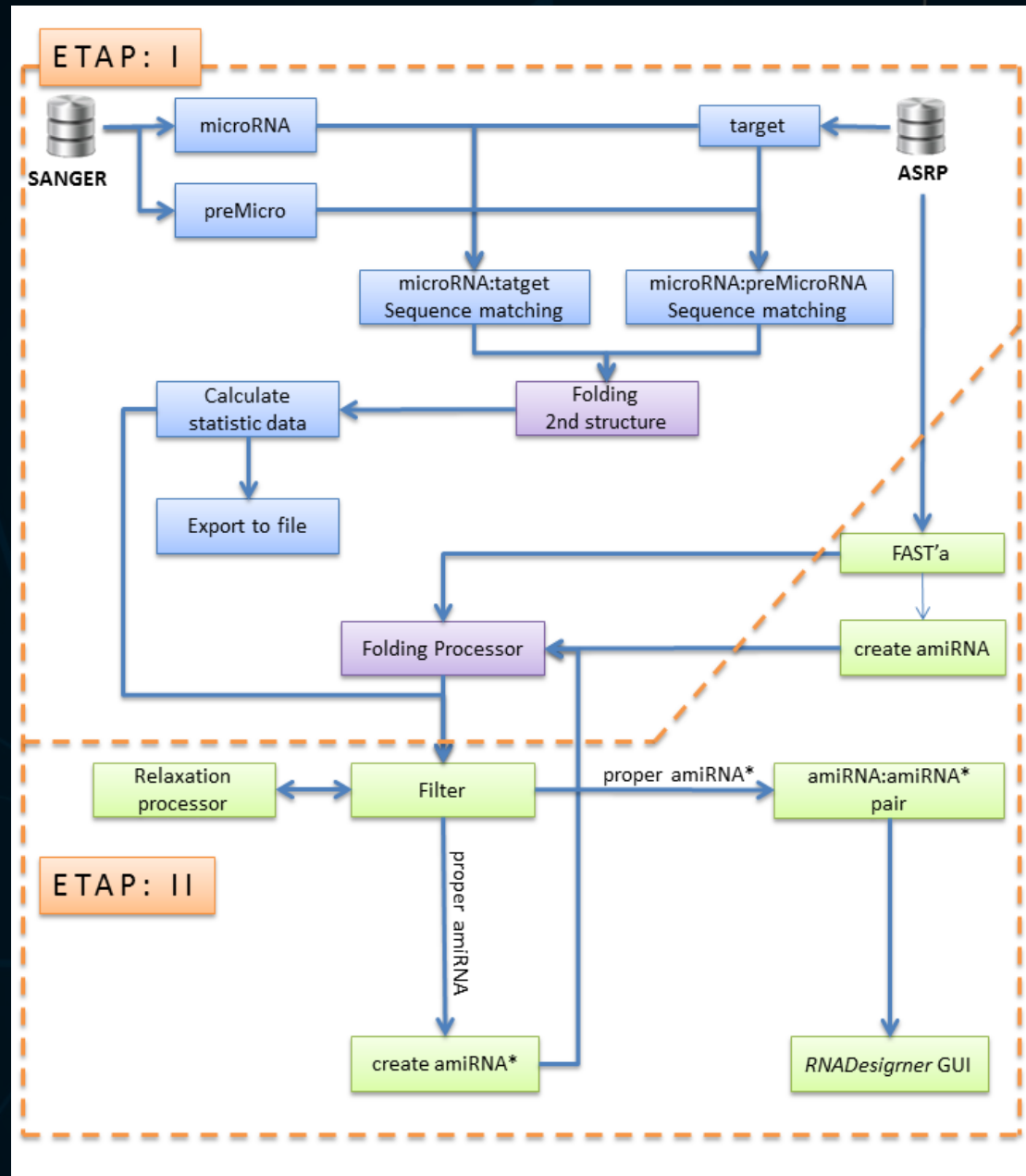
Second level



Third level



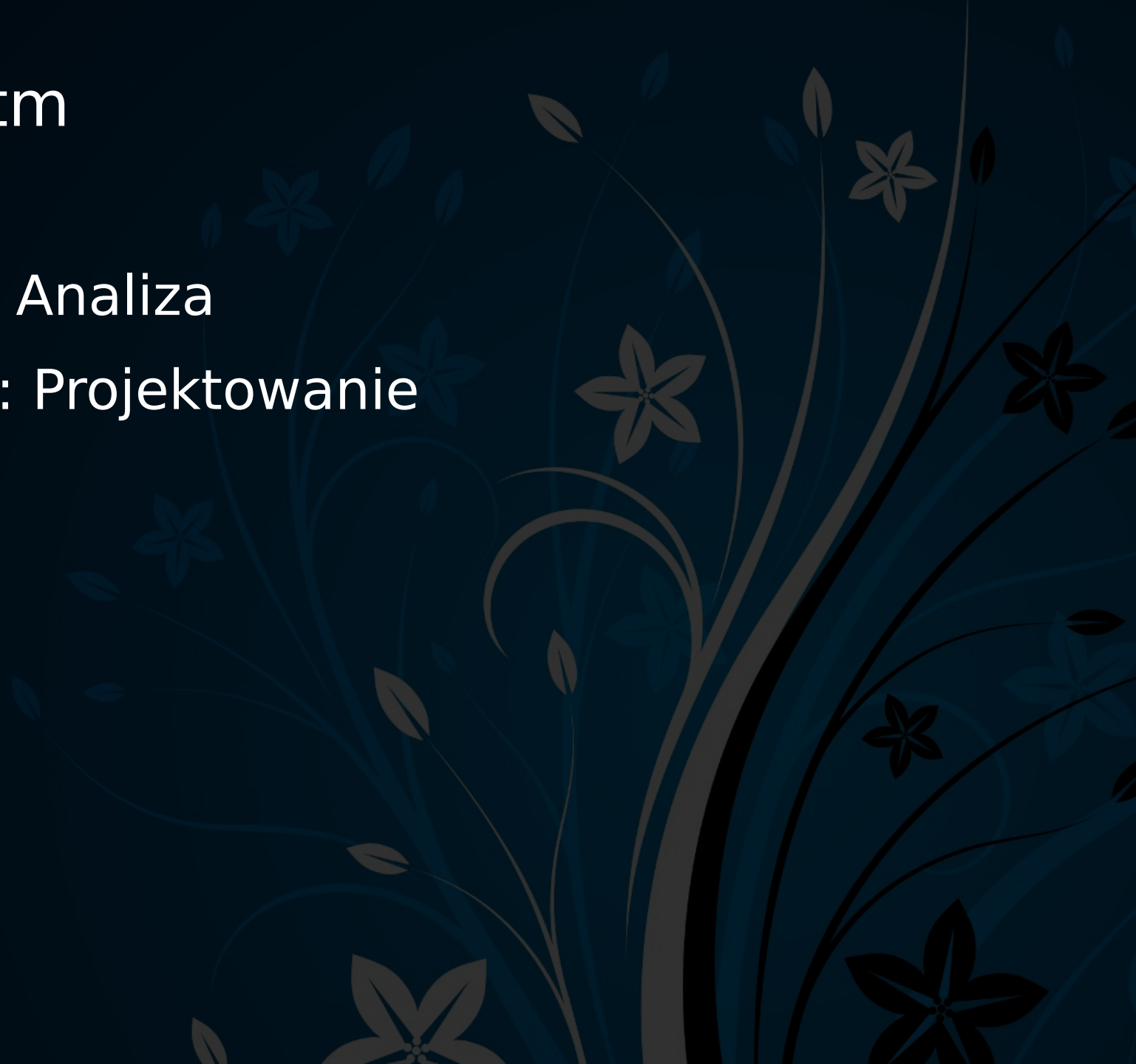
Algorytm



Algorytm

Etap I : Analiza

Etap II : Projektowanie



Algorytm

Etap I : Analiza

- Wczytanie danych
- Obliczenie danych termodynamicznych (UNAFold)
- Parsowanie wyniku
- Uśrednienie temperatur
- Eksport do pliku

Etap II : Projektowanie

Algorytm

Etap I : Analiza

- Wczytanie danych
- Obliczenie danych termodynamicznych (UNAFold)
- Parsowanie wyniku
- Uśrednienie temperatur
- Eksport do pliku

Etap II : Projektowanie

Algorytm

Etap I : Analiza

- Wczytanie danych
- Obliczenie danych termodynamicznych (UNAFold)
 - zapis sekwencji do pliku
 - uruchomienie zestawu skryptów
 - pobranie pliku tekstowego z wynikiem
- Parsowanie wyniku
- Uśrednienie temperatur
- Eksport do pliku

Etap II : Projektowanie

Algorytm

Etap I : Analiza

- Wczytanie danych
- Obliczenie danych termodynamicznych (UNAFold)
- Parsowanie wyniku
 - konwersja pliku tekstowego na obiekty biznesowe
- Uśrednienie temperatur
- Eksport do pliku

Etap II : Projektowanie

Algorytm



Algorytm

Etap I : Analiza

- Wczytanie danych
- Obliczenie danych termodynamicznych (UNAFold)
- Parsowanie wyniku
- Uśrednienie temperatur
- Eksport do pliku

Etap II : Projektowanie

Algorytm

Etap I : Analiza

Etap II : Projektowanie

- Wczytanie danych z pliku FASTA
- Generowanie sekwencji targetu
- Walidacja sekwencji targetu
- Kalkulacja statystyk
- Projektowanie amiRNA
- Projektowanie amiRNA*

Algorytm

Etap I : Analiza

Etap II : Projektowanie

- Wczytanie danych z pliku FASTA
- Generowanie sekwencji targetu
- Walidacja sekwencji targetu
- Kalkulacja statystyk
- Projektowanie amiRNA
- Projektowanie amiRNA*

Algorytm

Etap I : Analiza

Etap II : Projektowanie

- Wczytanie danych z pliku FASTA
- Generowanie sekwencji targetu

For i = 1 to i = FASTA.Length - 20

Begin

Take 21 nt;

End

- Walidacja sekwencji targetu
- Kalkulacja statystyk
- Projektowanie amiRNA
- Projektowanie amiRNA*

Algorytm

Etap I : Analiza

Etap II : Projektowanie

- Wczytanie danych z pliku FASTA
- Generowanie sekwencji targetu
- Walidacja sekwencji targetu
 - na 21. pozycji musi znajdować się nukleotyd A
 - na 12. pozycji musi znajdować się nukleotyd U
 - na 2. pozycji nie może znajdować się nukleotyd C
 - na 9. pozycji nie może znajdować się nukleotyd G
- Kalkulacja statystyk
- Projektowanie amiRNA
- Projektowanie amiRNA*

Algorytm

Etap I : Analiza

Etap II : Projektowanie

- Wczytanie danych z pliku FASTA
- Generowanie sekwencji targetu
- Walidacja sekwencji targetu
- Kalkulacja statystyk
 - temperatura minimalna
 - 1.kwartyl
 - mediana
 - 3.kwartyl
 - temperatura maksymalna
- Projektowanie amiRNA

Projektowanie amiRNA

- Utworzenie nici komplementarnej
- Wyznaczenie danych termodynamicznych
- Parsowanie wyniku
- Uśrednienie temperatur
- Walidacja
- Obsługa rozluźnień

Projektowanie amiRNA

- Utworzenie nici komplementarnej
 - A na U
 - U na A
 - G na C
 - C na G
- Wyznaczenie danych termodynamicznych
- Parsowanie wyniku
- Uśrednienie temperatur
- Walidacja
- Obsługa rozluźnień

Projektowanie amiRNA

- Utworzenie nici komplementarnej
- Wyznaczenie danych termodynamicznych
- Parsowanie wyniku
- Uśrednienie temperatur
- Walidacja
- Obsługa rozluźnień

Projektowanie amiRNA

- Utworzenie nici komplementarnej
- Wyznaczenie danych termodynamicznych
- Parsowanie wyniku
- Uśrednienie temperatur
- Walidacja
 - temperatura na każdej z 21 pozycji nie może przekroczyć 3.kwartyla
 - temperatura nie może przekroczyć wartości mediany przy jednoczesnym spadku mediany
- Obsługa rozluźnień

Projektowanie amiRNA

- Utworzenie nici komplementarnej
- Wyznaczenie danych termodynamicznych
- Parsowanie wyniku
- Uśrednienie temperatur
- Walidacja
- Obsługa rozluźnień
 - temperatura niższa od 1.kwartyla przy jednoczesnym rosnącym trendzie mediany

Obsługa rozluźnień

- Wprowadzane przy zbyt silnych wiązaniach
- Wstępna analiza możliwości rozluźnienia
- Generowanie wszystkich możliwych kombinacji spełniających podane kryteria
- Odrzucenie kombinacji niespełniających warunku odstępu
- Ponowne reprocesowanie

Obsługa rozluźnień

- Wprowadzane przy zbyt silnych wiązaniach
 - zmiana dokonywana na sekwencji amiRNA
- Wstępna analiza możliwości rozluźnienia
- Generowanie wszystkich możliwych kombinacji spełniających podane kryteria
- Odrzucenie kombinacji niespełniających warunku odstępu
- Ponowne reprocesowanie

Obsługa rozluźnień

- Wprowadzane przy zbyt silnych wiązaniach
- Wstępna analiza możliwości rozluźnienia
 - nukleotyd można rozluźnić (C na U, A na G)
 - nukleotyd będący prawym sąsiadem można rozluźnić (reguła prawego sąsiada)
- Generowanie wszystkich możliwych kombinacji spełniających podane kryteria
- Odrzucenie kombinacji niespełniających warunku odstępu
- Ponowne reprocesowanie

Obsługa rozluźnień

- Wprowadzane przy zbyt silnych wiązaniach
- Wstępna analiza możliwości rozluźnienia
- Generowanie wszystkich możliwych kombinacji spełniających podane kryteria
 - maksymalna liczba rozluźnień w sekwencji = 3
- Odrzucenie kombinacji niespełniających warunku odstępu
- Ponowne reprocesowanie

Obsługa rozluźnień

- Wprowadzane przy zbyt silnych wiązaniach
- Wstępna analiza możliwości rozluźnienia
- Generowanie wszystkich możliwych kombinacji spełniających podane kryteria
- Odrzucenie kombinacji niespełniających warunku odstępu
 - minimalny odstęp pomiędzy rozluźnieniami = 5
- Ponowne reprocesowanie

Obsługa rozluźnień

- Wprowadzane przy zbyt silnych wiązaniach
- Wstępna analiza możliwości rozluźnienia
- Generowanie wszystkich możliwych kombinacji spełniających podane kryteria
- Odrzucenie kombinacji niespełniających warunku odstępu
- Ponowne reprocesowanie

Projektowanie amiRNA*

- Utworzenie nici komplementarnej
- Wyznaczenie danych termodynamicznych
- Parsowanie wyniku
- Uśrednienie temperatur
- Walidacja
- Obsługa rozluźnień

Eksperyment obliczeniowy

Etap I :

- Dane wejściowe :
 - plik zawierający sekwencje microRNA (187 sekwencji)
 - plik zawierający sekwencje targetu (226 sekwencji)
 - plik zawierający dane preMicro (158 pozycji)

Eksperyment obliczeniowy

Etap I :

- Dane wejściowe :
 - 210 par microRNA:target
 - 157 sekwencji preMicro

Eksperyment obliczeniowy

Etap I :

- Fragment przetworzonych danych microRNA:target :

Gene	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
	-0,06	-2,06	-2,62	-2,34	-1,73	-1,85	-2,17	-1,96	-1,63	-1,29	-1,36	-1,69	-1,79	-1,78	-1,7	-1,93	-2,16	-2,19	-2,27	-2,08	-2,17	-2,01	-1,68	-1,09	-1,56	-2,35	-0,77	0,54
miR157a-"AT1G27360	0	-2,07	-2,73	-2,36	-1,8	-2,08	-2,51	-2,51	-2,23	-2,35	-2	-2	-2	-1,35	-0,44	-0,16	-1,35	-2,27	-2,55	-2,27	-2,91	-2,71	-2,88	-2,33	-2,9	-2,83	-0,77	0,54
miR157a-"AT1G53160	0	-2,07	-1,82	-0,8	-0,23	-1,43	-2,51	-2,51	-2,23	-2,35	-2	-2	-2	-1,35	-0,44	-0,16	-1,35	-2,27	-2,55	-2,27	-2,91	-2,71	-1,79	-0,23	-0,8	-1,82	-0,77	0,54
miR157a-"AT1G69170	0	-2,07	-2,88	-2,33	-1,76	-1,9	-2,51	-2,51	-2,23	-2,35	-2	-2	-2	-1,35	-0,44	-0,16	-1,35	-2,27	-2,55	-2,27	-2,91	-2,71	-2,88	-2,33	-2,9	-2,83	-0,77	0,54
miR157a-"AT5G50570	0	-2,07	-1,82	-0,8	-0,23	-1,43	-2,51	-2,51	-2,23	-2,35	-2	-2	-2	-1,35	-0,44	-0,16	-1,35	-2,27	-2,55	-2,27	-2,91	-2,71	-2,88	-2,33	-2,9	-2,83	-0,77	0,54
miR159a-"AT3G60460	0	-2,07	-2,73	-2,36	-1,45	-1,53	-2,25	-2,8	-2,44	-1,73	-1,19	-1,76	-1,75	-1,08	0,11	-0,61	-1,81	-2,72	-2,91	-2,63	-2,91	-2,41	-2,07	-1,79	-2,29	-2,73	-0,77	0,54
miR159a-"AT2G32460	0	-2,07	-2,73	-2,36	-1,45	-1,53	-2,25	-1,6	-0,61	0,11	-0,88	-2,08	-2,08	-2	-2,17	-2,31	-2,05	-1,6	-2,38	-2,63	-2,91	-2,27	-2,09	-1,81	-2,45	-2,73	-0,77	0,54
miR159b-"AT2G32460	0	-2,07	-2,73	-2,36	-1,45	-1,53	-2,25	-1,6	-0,61	0,11	-0,88	-2,08	-2,08	-2	-2,17	-2,31	-2,05	-1,6	-2,38	-2,63	-2,91	-2,27	-1,35	-0,16	-0,8	-1,82	-0,77	0,54
miR159a-"AT2G26950	0	-2,07	-2,88	-2,33	-1,42	-1,35	-2,25	-2,8	-1,81	-0,44	-0,23	-1,43	-2,08	-2	-2,17	-2,9	-3,18	-2,72	-2,91	-2,63	-2,91	-2,41	-2,07	-1,79	-2,29	-2,73	-0,77	0,54
miR159b-"AT2G26950	0	-2,07	-2,88	-2,33	-1,42	-1,35	-2,25	-2,8	-1,81	-0,44	-0,23	-1,43	-2,08	-2	-2,17	-2,9	-3,18	-2,72	-2,91	-2,63	-2,91	-2,41	-1,5	-0,3	-0,8	-1,82	-0,77	0,54

Eksperyment obliczeniowy

Etap II :

- Dane wejściowe :
 - plik w formacie FASTA (1642 nukleotydy)

Eksperyment obliczeniowy

Etap II :

- Dane wejściowe :
 - 1622 sekwencje targetu

Eksperyment obliczeniowy

Etap II :

- Sekwencje po walidacji :
 1. na 21. pozycji musi znajdować się nukleotyd A
 2. na 12. pozycji musi znajdować się nukleotyd U
 3. na 2. pozycji nie może znajdować się nukleotyd C
 4. na 9. pozycji nie może znajdować się nukleotyd G

Reguły	targety zgodne z regułami (%)
1,2,3,4	65 sekwencji (4,01%)
1,2,3	85 sekwencji (5,24%)
1,2	109 sekwencji (6,72%)
1	392 sekwencje (24,17%)
2	483 sekwencje (29,78%)
2,3	392 sekwencje (24,17%)
2,3,4	294 sekwencje (18,13%)

Eksperyment obliczeniowy

Etap II :

- Projektowanie dla reguł :
 - na 21. pozycji musi znajdować się nukleotyd A (392 sekwencje)
- Rezultat :
 - 50 sekwencji amiRNA
 - 5 sekwencji amiRNA*

Eksperyment obliczeniowy

Etap II :

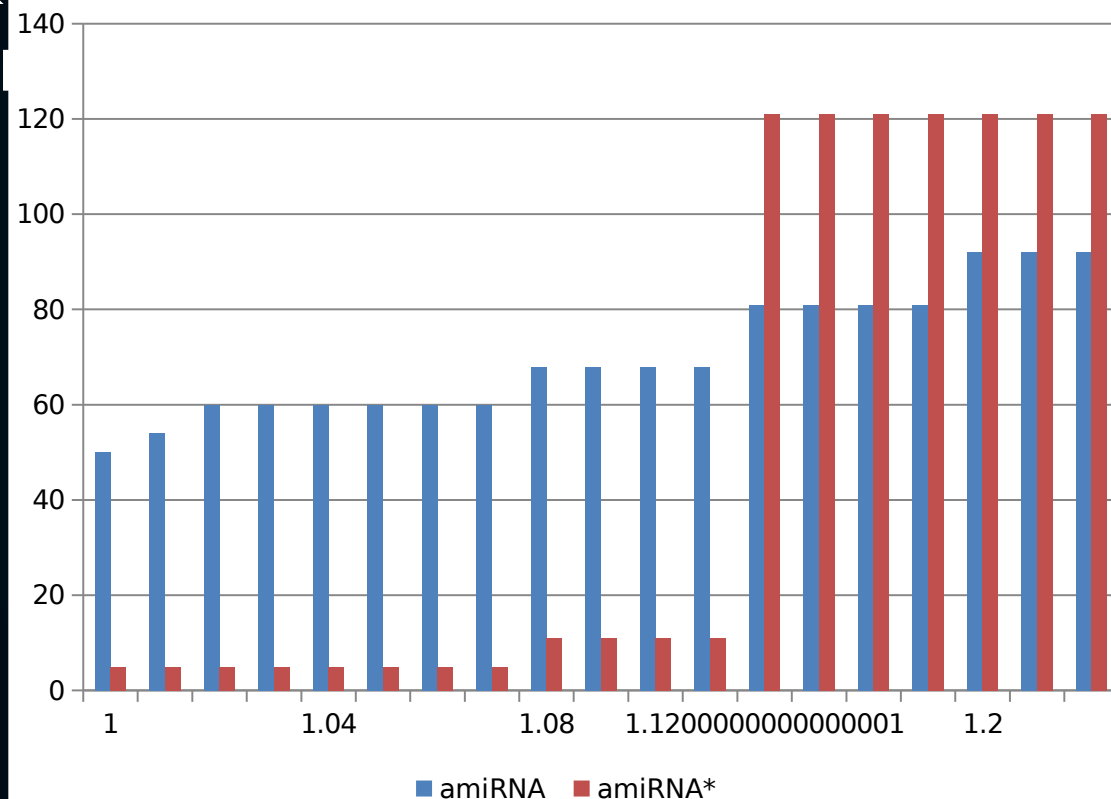
- **Rezultat :**
 - 5 sekwencji amiRNA
 - 5 sekwencji amiRNA*

amiRNA ID	amiRNA	amiRNA* ID	amiRNA*
amiR_FAS_968_20G	UCUCCCUCGUGACCAUGAAG G	antyamiR_FAS_968_20G_15 G_11U	CCUUCAUGGUUACGGGGGA GA
amiR_FAS_968_20G	UCUCCCUCGUGACCAUGAAG G	antyamiR_FAS_968_20G_15 G	CCUUCAUGGUCACGGGGGA GA
amiR_FAS_968_20G	UCUCCCUCGUGACCAUGAAG G	antyamiR_FAS_968_20G_11 U	CCUUCAUGGUUACGAGGGAG A
amiR_FAS_1564_19G	UUGAGUAGCGUGUUGAAGG GG	antyamiR_FAS_1564_19G_11 G	CCCCUUAACGCGCUACUCA A
amiR_FAS_1564_20G	UUGAGUAGCGUGUUGAAGG GG	antyamiR_FAS_1564_20G_11 G	CCCCUUAACGCGCUACUCA A

Eksperyment obliczeniowy

Etap II :

- 3.kwartyl – testy dla zwiększonych wartości



Podsumowanie

Dziękujemy za uwagę ^^