

53 Olimpiada Biologiczna

Pracownia 201/A – arkusz zadań

Czas trwania: 90 minut

Liczba punktów: 30

Rozwiązując zadania, możesz korzystać z wymienionego poniżej oprogramowania:

- składniki pakietu biurowego LibreOffice,
- środowisko obliczeniowe R,
- ApE,
- Chimera,
- ClustalX,
- domyślne składniki systemu operacyjnego, np. kalkulator, notatnik.

W każdym zadaniu samodzielnie wybierz odpowiedni program lub programy. Metoda rozwiązania zadania nie będzie oceniana, tzn. liczy się wyłącznie końcowy wynik.

Odpowiedzi udziel – w zależności od rodzaju polecenia – na karcie odpowiedzi, zamalowując odpowiednie pola albo zapisując pliki z sekwencjami w formacie FASTA. Pliki FASTA z udzielonymi odpowiedziami **zapisz w podkatalogu „odpowiedzi”**.

W **załączniku 1** umieszczono listę wszystkich plików z sekwencjami nukleotydowymi, które znajdują się w podkatalogu „sekwencje”.

Wstęp do zadania 1.

U ludzi występuje pięć funkcjonalnych hialuronidaz kodowanych przez następujące geny: *HYAL1*, *HYAL2*, *HYAL3*, *HYAL4* i *HYAL5* oraz pseudogen *HYAL6*. Geny *HYAL1–3* są zlokalizowane na chromosomie 3, podczas gdy geny *HYAL4–5* oraz pseudogen *HYAL6* są zlokalizowane na chromosomie 7. Białka *HYAL1* i *HYAL2* są głównymi hialuronidazami w większości tkanek.

W poniższej tabeli przedstawiono podstawowe informacje dotyczące pięciu różnych dojrzałych mRNA genu *HYAL3*, które powstają w wyniku alternatywnego splicingu (składania eksonów).

dojrzały mRNA genu <i>HYAL3</i>	nazwa pliku z sekwencją nukleotydową dojrzałego mRNA	m.cz. kodowanego białka	aktywność enzymatyczna
wariant 1.	NM_003549.fasta	46,5 kDa	aktywna
wariant 2.	NM_001200030.fasta	43,6 kDa	nieaktywna
wariant 3.	NM_001200031.fasta	18,8 kDa	nieaktywna
wariant 4.	NM_001200032.fasta	15,9 kDa	nieaktywna
wariant 5.	NM_001200029.fasta	46,5 kDa	aktywna

Na potrzeby tej pracowni przyjmij następujące definicje eksonu i intronu:

- ekson – ta część pre-mRNA (pierwotnego transkryptu), która jest obecna także w dojrzałym mRNA
- intron – ta część pre-mRNA (pierwotnego transkryptu), która jest wycinana w trakcie dojrzewania mRNA i nie jest w nim obecna.

Plik „NC_000003.fasta” zawiera fragment chromosomu 3, w obrębie którego to fragmentu leży gen *HYAL3*.

W tym fragmencie leży jeszcze inny gen kodujący podjednostkę katalityczną *N*- α -acetylotransferazę 80 (NAA80). Plik „NM_012191.fasta” zawiera sekwencję nukleotydową dojrzałego mRNA powstającego w wyniku transkrypcji genu *NAA80*.

Pliki z pojedynczymi sekwencjami nukleotydowymi możesz otworzyć zarówno w Notatniku, jak i w programie ApE. Jednak w programie ApE możesz wykonać przyrównanie dwóch lub większej liczby sekwencji. W programie ApE możesz także znaleźć otwarte ramki odczytu oraz przeprowadzić translację *in silico*.

Skrócona instrukcja do programu ApE

Przygotowanie przyrównania sekwencji:

1. Otwórz pliki FASTA w programie ApE.
2. Kliknij **Tools** → **Align Sequences**.
3. Wybierz sekwencję referencyjną, np. w postaci wariantu transkrypcyjnego 1. w menu **Reference sequence**.
4. Wybierz inne sekwencje do przyrównania w części **Align to Windows**. Jeżeli chcesz wybrać więcej niż jedną sekwencję, to przytrzymaj klawisz *Shift* na klawiaturze i kliknij po kolei w każdą nazwę sekwencji.

Uwaga: Wykonaj przyrównanie, stosując domyślne ustawienia programu ApE.

Znajdywanie otwartej ramki odczytu:

1. Otwórz plik FASTA w programie ApE.
2. Kliknij **ORFs** → **Find Next**.
3. W przypadku dojrzałych mRNA przyjmij, że to właśnie najdłuższa otwarta ramka odczytu spośród wszystkich znalezionych ulega translacji.

Translacja *in silico*:

1. Zaznacz myszką część sekwencji wyświetlonej w programie ApE, która ma być poddana translacji *in silico*.
2. Kliknij **ORFs** → **Translate**.
3. Jeśli chcesz uzyskać tylko sekwencję aminokwasową (bez numerów reszt aminokwasowych oraz sekwencji nukleotydowej), wybierz **None** w odpowiednio opcjach **Line Numbers** oraz **DNA**.

Do wykonania przyrównania sekwencji nukleotydowych albo aminokwasowych możesz wykorzystać program ClustalX.

Zadanie 1.1. (2 pkt.)

Określ liczbę intronów, które zostały usunięte podczas składania wariantu 1. dojrzałego mRNA genu *HYAL3*.

- A. dwa
- B. trzy
- C. cztery
- D. pięć
- E. sześć

Zadanie 1.2. (1 pkt.)

Określ, czy w eksonie 1. wariantu 2. dojrzałego mRNA genu *HYAL3* znajduje się sekwencja kodująca białko.

- T. tak
- N. nie

Zadanie 1.3. (2 pkt.)

Określ długość 5' UTR (rejon nieulegający translacji) w wariancie 3. dojrzałego mRNA genu *HYAL3*.

- A. 191 nukleotydów
- B. 194 nukleotydy
- C. 414 nukleotydy
- D. 695 nukleotydów
- E. 698 nukleotydów

Zadanie 1.4. (1 pkt)

Określ kodon stop w otwartej ramce odczytu wariantu 4. dojrzałego mRNA genu *HYAL3*.

- A. UAA (zapisany w sekwencji jako TAA)
- B. UAG (zapisany w sekwencji jako TAG)
- C. UGA (zapisany w sekwencji jako TGA)

Zadanie 1.5. (2 pkt.)

Określ, długość sekwencji nukleotydowej wspólnej w wariantach 1. i 5. dojrzałych mRNA genu *HYAL3* znajdującej się w 5' UTR od jego strony 3'.

- A. 16 nukleotydów
- B. 17 nukleotydów
- C. 57 nukleotydów
- D. 153 nukleotydy
- E. 193 nukleotydy

Zadanie 1.6. (2 pkt.)

Określ, czy następujące stwierdzenia są prawdziwe. Zaznacz P, jeśli stwierdzenie jest prawdziwe, albo F – jeśli jest fałszywe. Pozycje reszt aminokwasowych odnoszą się do polipeptydu powstającego w wyniku translacji wariantu 1. dojrzałego mRNA genu *HYAL3*.

Stwierdzenie	Prawda czy fałsz?
1. Występowanie w białku <i>HYAL3</i> reszt 1–250 jest wystarczające do tego, aby białko to było aktywne enzymatycznie.	<input type="checkbox"/> prawda / <input type="checkbox"/> fałsz
2. Występowanie w białku <i>HYAL3</i> reszt 300–329 jest wystarczające do tego, aby to białko było aktywne enzymatycznie.	<input type="checkbox"/> prawda / <input type="checkbox"/> fałsz

Zadanie 1.7. (2 pkt.)

Określ, czy następujące stwierdzenia są prawdziwe. Zaznacz P, jeśli stwierdzenie jest prawdziwe, albo F – jeśli jest fałszywe.

Stwierdzenie	Prawda czy fałsz?
1. Sekwencje 5' UTR <u>wariantu 1.</u> dojrzałego mRNA genu <i>HYAL3</i> oraz dojrzałego mRNA genu <i>NAA80</i> są takie same.	<input type="checkbox"/> prawda / <input type="checkbox"/> fałsz
2. Sekwencja kodująca białko <i>NAA80</i> <u>w całości</u> znajduje się w jednym z intronów genu <i>HYAL3</i> .	<input type="checkbox"/> prawda / <input type="checkbox"/> fałsz

Wstęp do zadania 2.

Do tej pory opisano w bazie Protein Data Bank (PDB) tylko sześć struktur hialuronidaz. Jedną z nich – rozwiązaną dzięki zastosowaniu krystalografii rentgenowskiej – jest hialuronidaza występująca w jadzie pszczelim pszczoły miodnej. Ta struktura jest zarejestrowana w PDB pod numerem identyfikacyjnym 1FCQ. Plik o nazwie „1fcq.pdb” ze strukturą 1FCQ zapisano w podkatalogu „struktury”.

Struktura hialuronidazy człowieka HYAL3 do tej pory nie została rozwiązana. Jednak współczesne narzędzia bioinformatyczne oparte na uczeniu maszynowym – takie jak AlphaFold – pozwalają z dużą dozą prawdopodobieństwa przewidzieć strukturę przestrzenną białka. Plik „AF-O43820-F1-model_v4.pdb” zawiera przewidzianą algorytmem AlphaFold strukturę białka HYAL3.

W podkatalogu sekwencje znajduje się plik „NM_001011619.fasta” zawierający sekwencję dojrzałego mRNA genu kodującego hialuronidazę u pszczoły miodnej. Najdłuższa otwarta ramka odczytu koduje sekwencję aminokwasową tego enzymu.

Wykonaj następujące kroki w programie Chimera i przeprowadź analizę strukturalną hialuronidaz człowieka i pszczoły miodnej.

1. Otwórz pliki w formacie PDB („1fcq.pdb” i „AF-O43820-F1-model_v4.pdb”) w programie Chimera. Kliknij **File** → **Open**. Wciśnij klawisz *shift* na klawiaturze i kliknij obie nazwy plików w formacie PDB. Obie struktury zostaną wyświetlone w programie Chimera.
2. Wykonaj przyrównanie strukturalne. Kliknij **Tools** → **Structure Comparison** → **MatchMaker**. W części **Reference structure** zaznacz „1fcq.pdb”, a w części **Structure(s) to match** – „AF-O43820-F1-model_v4.pdb”. Kliknij **OK**. Obie struktury zostaną przyrównane tak, aby korespondujące ze sobą atomy były jak najmniej oddalone.
3. Wyświetl przyrównanie sekwencji aminokwasowych obu białek. Kliknij **Tools** → **Structure Comparison** → **Match** → **Align**. Upewnij się, że obie nazwy są zaznaczone i kliknij **OK**.

W programie Chimera można obracać wyświetloną strukturę przez kliknięcie lewego przycisku myszy i przesuwanie kursora po ekranie. Przyciśnięcie prawego przycisku myszy umożliwia powiększenie lub zmniejszenie wyświetlonej struktury. Pozostawienie kursora przez ok. 0,5 sekundy nad wybranym elementem wyświetlonej struktury powoduje wyświetlenie informacji o danej reszcie aminokwasowej, np. „#0 MET 1.A”. Ta informacja oznacza, że kursor leży nad strukturą przyrównaną (oznaczenie #0 ma sekwencja, która była otwarta w programie Chimera jako pierwsza) i znajduje się tam metionina w pozycji 1. łańcucha polipeptydowego. Oznaczenie literowe „A” odnosi się do pierwszego łańcucha polipeptydowego w danym pliku PDB. Ponieważ oba pliki zawierają po jednym łańcuchu polipeptydowym, można pominąć to oznaczenie.

Zadanie 2.1. (2 pkt.)

Do rozwiązania struktury przestrzennej hialuronidazy pszczoły miodnej wykorzystano rekombinowane białko o niepełnej długości.

Określ pozycję reszty aminokwasowej oznaczonej w strukturze przestrzennej jako „CYS 189.A” w strukturze pierwszorzędowej pełnej długości hialuronidazy pszczoły miodnej.

- A. pozycja 189.
- B. pozycja 205.
- C. pozycja 221.
- D. pozycja 233.
- E. pozycja 345.

Zadanie 2.2. (4 pkt.)

Określ, czy następujące stwierdzenia są prawdziwe. Zaznacz P, jeśli stwierdzenie jest prawdziwe, albo F – jeśli jest fałszywe.

Stwierdzenie	Prawda czy fałsz?
1. Algorytm AlphaFold przewidział strukturę przestrzenną całego białka HYAL3 występującego u człowieka.	<input type="checkbox"/> prawda / <input type="checkbox"/> fałsz
2. Pierwszym elementem struktury drugorzędowej od końca aminowego w hialuronidazie jest α -helisa.	<input type="checkbox"/> prawda / <input type="checkbox"/> fałsz
3. Strukturze α -helisy obecnej w hialuronidazie z pszczoły miodnej o sekwencji „LTKHLQVFRDHLINQ” odpowiadają strukturalnie pozycje 101–115 z funkcjonalnej hialuronidazy człowieka.	<input type="checkbox"/> prawda / <input type="checkbox"/> fałsz
4. W hialuronidazie pszczoły miodnej występują tylko dwie β -kartki.	<input type="checkbox"/> prawda / <input type="checkbox"/> fałsz

Zadanie 2.3. (2 pkt.)

Określ, która z poniższych reszt cysteiny tworzy mostek disiarczkowy z cysteiną oznaczoną jako „CYS 189.A” w hialuronidazie z pszczoły miodnej.

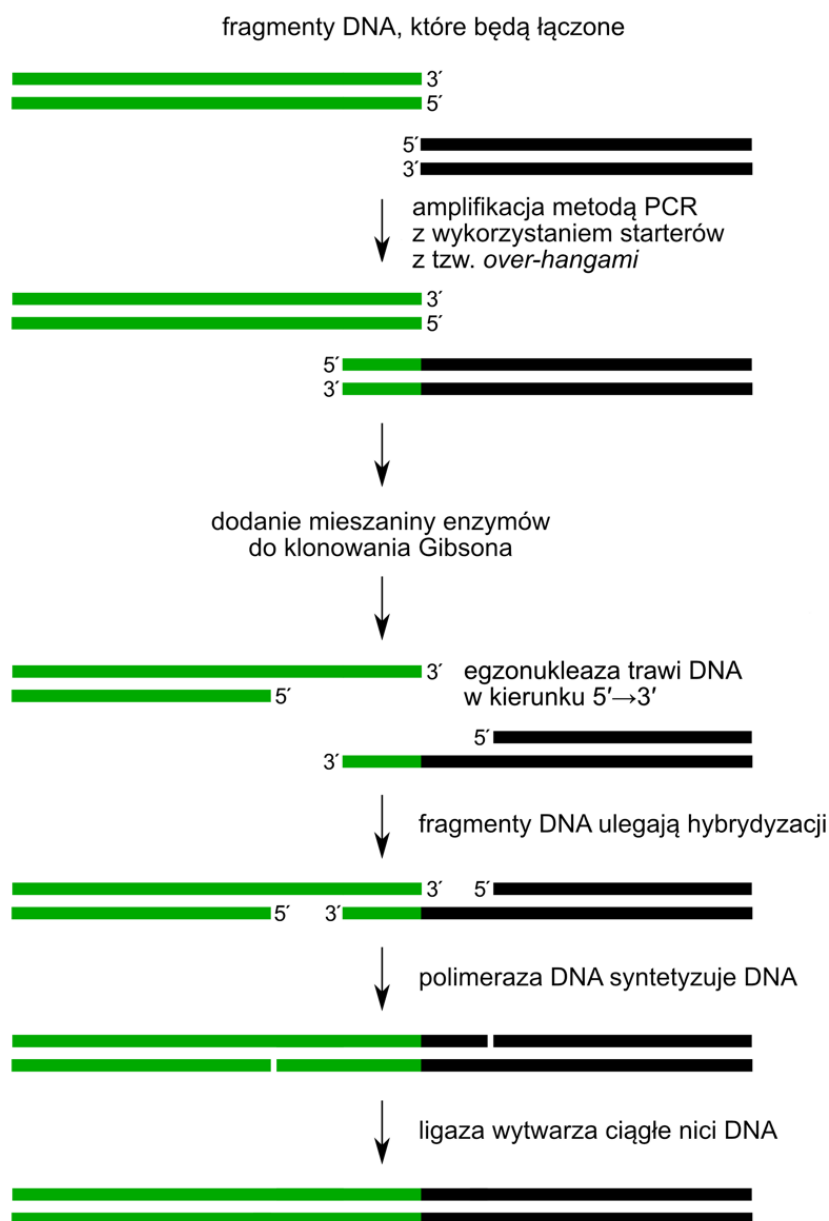
- A. CYS 16.A z człowieka
- B. CYS 205.A z człowieka
- C. CYS 22.A z pszczoły miodnej
- D. CYS 201.A z pszczoły miodnej
- E. CYS 313.A z pszczoły miodnej

Wstęp do zadania 3.

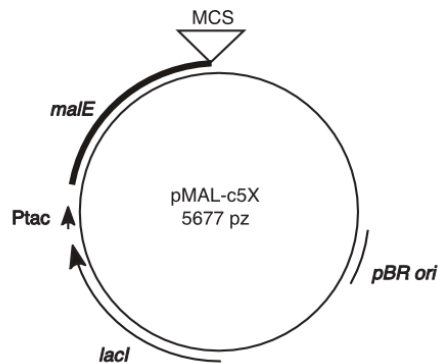
Klonowanie Gibsona (ang. *Gibson assembly*) to metoda pozwalająca na połączenie ze sobą nawet kilku fragmentów DNA w jednej reakcji zachodzącej w stałej temperaturze. W tej metodzie kluczowe znaczenie mają nachodzące na siebie, komplementarne fragmenty o długości 20–40 par zasad. Następnie fragmenty ulegają połączeniu dzięki trzem enzymom obecnym w mieszaninie reakcyjnej:

- Egzonukleaza – hydrolizuje pojedynczą nić DNA od końca 5' i tworzy „lepkie końce”;
- Polimeraza DNA – wypełnia utworzone przez egzonukleazę przerwy po hybrydyzacji komplementarnych nici z sąsiadujących fragmentów DNA;
- Ligaza DNA – łączy kowalencyjnie sąsiadujące ze sobą fragmenty.

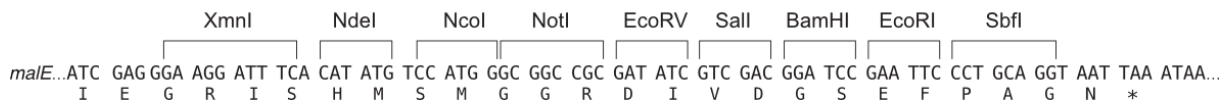
Na poniższym schemacie przedstawiono kolejne etapy klonowania Gibsona.



Zaprojektuj startery do PCR, które pozwolą namnożyć sekwencję kodującą białko HYAL3 człowieka i połączyć ją z wektorem ekspresyjnym pMAL-c5X metodą klonowania Gibsona.

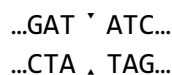


pMAL-c5X MCS



Na powyższym schemacie przedstawiono budowę wektora pMAL-c5X. Cechą charakterystyczną tego wektora jest to, że pod kontrolą promotora Ptac występuje gen *malE* kodujący białko wiążące maltozę, który nie kończy się kodonem STOP, ponieważ bezpośrednio za nim znajduje się miejsce ułatwiające klonowanie (ang. *multiple cloning site*; MCS). Translacja tej ramki odczytu skutkuje powstaniem białka MalE z wydłużonym końcem C o łącznej m.cz. 44,7 kDa.

Pod sekwencją nukleotydową obszaru MCS umieszczono sekwencję aminokwasową polipeptydu powstającego w tej samej ramce odczytu, co białko MalE. W MCS występuje miejsce dla enzymu restrykcyjnego EcoRV, który przecina rozpoznawaną sekwencję i tworzy tępe końce:



Do przecięcia kolistego wektora wykorzystaj enzym restrykcyjny EcoRV. Jako matrycę do przeprowadzenia reakcji PCR wykorzystaj cDNA zsyntetyzowany na matrycy wariantu 1. dojrzałego mRNA genu *HYAL3*.

Zaprojektowane startery muszą:

- umożliwić powstanie fuzyjnego białka MalE-HYAL3 (nie usuwaj pierwszej metioniny z białka HYAL3)
- zapewnić brak dodatkowych reszt aminokwasowych na końcu C rekombinowanego białka MalE-HYAL3 (wykorzystaj kodon stop znajdujący się w otwartej ramce odczytu genu HYAL3).

i składać się z dwóch części: (1) hybrydującej z sekwencją nukleotydową otwartej ramki odczytu kodującej białko HYAL3 o długości 20 nukleotydów oraz (2) tzw. *over-hang* o długości 30 nukleotydów, nachodzącego na sekwencję wektora pMAL-c5X – oba startery (tzw. lewy i prawy) muszą zawierać 50 reszt nukleotydowych.

Zadanie 3.1. (3 pkt.)

Zaprojektuj tzw. lewy (ang. *forward*) starter spełniający wymagania przedstawione we wstępie do zadania. Sekwencję zapisz w formacie FASTA. Jako identyfikator sekwencji zapisz numer uczestnika bez znaków diakrytycznych wraz z oznaczeniem kierunku (*_fwd*), np. „>ZOLT01_fwd”. Plik nazwij wg takej samej konwencji, ale usuń znak „>” i dodaj rozszerzenie „.fasta”, np. „ZOLT01_fwd.fasta”, oraz zapisz w podkatalogu „odpowiedzi”.

Zadanie 3.2. (3 pkt.)

Zaprojektuj tzw. prawy (ang. *reverse*) starter spełniający wymagania przedstawione we wstępie do zadania. Sekwencję zapisz w formacie FASTA. Jako identyfikator sekwencji zapisz numer uczestnika bez znaków diakrytycznych wraz z oznaczeniem kierunku (*_rev*), np. „>ZOLT01_rev”. Plik nazwij wg takej samej konwencji, ale usuń znak „>” i dodaj rozszerzenie „.fasta”, np. „ZOLT01_rev.fasta”, oraz zapisz w podkatalogu „odpowiedzi”.

Zadanie 3.3. (2 pkt.)

Określ przewidywaną masę cząsteczkową rekombinowanego białka MaIE-HYAL3.

- A. 44,7 kDa
- B. 46,5 kDa
- C. 90,1 kDa
- D. 90,8 kDa
- E. 91,1 kDa

Zadanie 3.4. (2 pkt.)

Określ, która kombinacja enzymu restrykcyjnego oraz wielkości fragmentów powstałych w wyniku przeprowadzenia analizy restrykcyjnej wskazuje na prawidłowe wykonanie rekombinowanego wektora pMAL-c5X-HYAL3.

Odpowiedź	Enzym restrykcyjny	Długość fragmentów DNA (pz)
A.	ApaI	1961, 1223, 1163, 830, 500
B.	EcoRI	5677
C.	NcoI	5832, 595, 504
D.	NdeI	5677
E.	PvuII	3908, 1676, 93

Załącznik 1. – wykaz plików z sekwencjami nukleotydowymi

nazwa pliku	format	krótka charakterystyka
NC_000003.fasta	FASTA	Fragment chromosomu 3 człowieka – w tym rejonie występuje gen <i>HYAL3</i> oraz gen <i>NAA80</i>
NM_003549.fasta	FASTA	Wariant 1. dojrzałego transkryptu genu <i>HYAL3</i>
NM_012191.fasta	FASTA	Dojrzały transkrypt genu <i>NAA80</i> człowieka
NM_001011619.fasta	FASTA	Dojrzały mRNA hialuronidazy pszczoły miodnej
NM_001200029.fasta	FASTA	Wariant 5. dojrzałego transkryptu genu <i>HYAL3</i>
NM_001200030.fasta	FASTA	Wariant 2. dojrzałego transkryptu genu <i>HYAL3</i>
NM_001200031.fasta	FASTA	Wariant 3. dojrzałego transkryptu genu <i>HYAL3</i>
NM_001200032.fasta	FASTA	Wariant 4. dojrzałego transkryptu genu <i>HYAL3</i>
pMAL-c5x.apc	ApE	Wektor ekspresyjny pMAL-c5X – jest ustawiony jako cząsteczka kolista (ang. <i>circular</i>)