

### Szczegółowy opis przedmiotu zamówienia

„Wykonanie usługi transkryptomiki i metabolomiki na potrzeby projektu ORCHIDOMICS realizowanego na Wydziale Biologii Uniwersytetu Gdańskiego”

#### Wprowadzenie

Innowacyjność projektu ORCHIDOMICS wynika z równoczesnego zastosowania najnowocześniejszych technologii analitycznych o wysokiej przepustowości, a mianowicie transkryptomiki i metabolomiki do analizy dużej liczby prób roślin zbieranych w naturalnych populacjach. Połączenie tych technologii umożliwia uzyskanie kompleksowego, o ile nie wyczerpującego obrazu fizjologii badanych roślin. Do tej pory metody te były stosowane niezależnie od siebie i dla prób pochodzących z organizmów modelowych, hodowanych w kontrolowanych warunkach.

Przed projektem ORCHIDOMICS stoi więc kilka wyzwań:

- **Stosowanie dwóch metod omicznych** do nietypowych próbek, których **dostępność jest ograniczona**, co w efekcie nie pozwala na wielokrotne próbkowanie w celu optymalizacji metod.
- **Integracja wyników analiz omicznych** – jednoczesne zastosowanie transkryptomiki i metabolomiki dostarczy wyniki, które żadna z powyższych metod stosowana osobno nie byłaby w stanie wygenerować, jednak jedynie pod warunkiem poprawnej ich integracji. Ten proces jest jednak bardzo skomplikowany zwłaszcza, że obiektem badań są organizmy niemodelowe a próby pobierane są w warunkach naturalnych czyli w niekontrolowanych warunkach wzrostu. Wymaga to wykonania badań z zastosowaniem obu metod na dokładnie tych samych próbach (**dokładność i precyzja pracy na obu zestawach prób**) oraz w rygorystycznie kontrolowanych warunkach analiz tak aby próby badane były w jednakowych warunkach, co pozwoli uniknąć lub znacznie zredukować dodatkowy szum informacyjny, który może poważnie wpłynąć na jakość integracji danych.
- **Rygor, dokładność i jednorodność analizy** w projekcie ORCHIDOMICS musi być utrzymane w stosunku do stosunkowo dużej liczby prób (288). Próby te obejmują korzenie, liście oraz protokormy (kielekujące nasiona) różnych gatunków storczyków (*Epipogium aphyllum*, *Neottia nidus-avis*, *Epipactis helleborine*, *Epipactis atrorubens*, *Cephalanthera damasonium*, *Epipactis palustris*, *Listera ovata*, *Platanthera bifolia* and *Dactylorhiza majalis*).

#### I. Próby dostarczone przez Zamawiającego

##### Pkt 1. Transkryptomika:

- 288 prób oczyszczonego RNA. Co najmniej cztery mikrogramy całkowitego RNA wyekstrahowanego z zastosowaniem protokołu NucleoZol (Macherey-Nagel) i traktowanych DNazą, RNA rozpuszczone w wodzie wolnej od RNaz, w próbkach wolnych od Rnaz w stężeniu nie mniejszym niż 100 ng/ul. Wartość A<sub>260</sub> / A<sub>280</sub> nm będzie wynosić nie mniej niż 1.7. Kontrola jakości RNA powinna być przeprowadzona przez wykonawcę po otrzymaniu prób.

##### Pkt 2. Metabolomika:

- 288 próbek dostarczonych w porcjach 10 mg liofilizowanego proszku.

##### Pkt 3. Sposób i czas dostarczenia prób:

- a) Łączna liczba próbek wynosi 288 dla każdej z analiz omicznych (transkryptomika i metabolomika). Probki będą przesłane Wykonawcy do 2 tygodni, licząc od daty zawarcia umowy.
- b) Próby będą dostarczone osobiście przez zamawiającego lub jeżeli to niemożliwe przez dedykowanego kuriera w warunkach, które nie wpłyną na utratę ich właściwości (wyzolowane RNA będzie dostarczone w ciekłym azocie lub suchym lodzie). Wykonawca zostanie powiadomiony o terminie dostarczenia przesyłki. Za potwierdzenie odbioru przesłanych materiałów uznaje się podpis na stosownym dokumencie

(np. liście przewozowym) przedstawiciela Wykonawcy uprawnionego do odbioru prób zgodnie z zapisami w umowie.

- c) Kompletność przesłanych próbek z załączoną listą powinna być niezwłocznie sprawdzona przez przedstawiciela Wykonawcy. Zamawiający powinien być niezwłocznie powiadomiony o wynikach odbioru. Wykonawca jest zobowiązany bezzwłocznie po otrzymaniu prób zapewnić, do czasu wykonania analiz, przechowywanie w odpowiednich warunkach (w przypadku prób RNA w temperaturze min. -80°C)

## **II. Rozliczenia z Wykonawca**

Wykonawca wykona analizy łącznie dla 288 próbek transkryptomocnych i 288 próbek metabolomicznych, przy czym koszt każdej z analiz powinien figurować jako osobna pozycja. Wykonawca wykona i wyceni usługę proporcjonalnie do dostarczonej liczby próbek. Od ceny końcowej zostaną odliczone próby, dla których analizy nie wyszły (powyżej akceptowanego przez Zamawiającego błędu na poziomie 3%, mimo odpowiedniej jakości dostarczonego materiału).

## **III. Rodzaj analizy do wykonania**

### **Pkt 1. Transkryptomika:**

Projekt ma na celu analizę zmian jądrowego oraz organellowego (plastydy i mitochondria) transkryptomu w celu zbadania, które geny ulegają ekspresji na określonych etapach rozwoju roślin i w określonych warunkach. Nie zajmuje się jednak badaniem małych cząsteczek RNA. Z uwagi na fakt, że przedmiotem badań są organizmy niemodelowe, brak dla nich odpowiednich genomów referencyjnych. Uzyskane dane transkryptomowe posłużą więc początkowo do złożenia (assemblingu) referencyjnego transkryptomu metodą *de novo* dla każdego z badanych gatunków storczyków. Zamawiający wymaga zatem, aby sekwencjonowanie transkryptomu wykonane zostało z użyciem dwóch bibliotek o różnych długościach odczytów, z założeniem, że jedna z bibliotek wykonana jest w technologii paired-end, co czyni proces składania transkryptomu dużo dokładniejszy i mniej wymagający pod kątem obliczeniowym. Wymagane jest zastosowanie wysokoprzepustowej technologii sekwencjonowania transkryptomu (RNA-seq) o wysokiej dokładności (niski poziom błędów,  $\geq 75\%$  zsekwencjonowanych zasad o jakości powyżej Q30) i czułości (możliwość wygenerowania do 5 miliardów odczytów na jeden bieg instrumentu sekwencjonującego).

### **Pkt 2. Metabolomika:**

Projekt ma na celu analizę składu metabolitów, a zwłaszcza metabolitów podstawowych w próbkach tkanek badanych storczyków. Analiza powinna zostać wykonana z zastosowaniem technologii chromatografii sprzęgniętej ze spektrometrią mas. Technologia GC-MS (chromatografia gazowa sprzęgnięta ze spektrometrią mas) jest preferowana, gdyż dostarcza wystarczającą ilość danych dla tego typu analiz, jednakże zastosowanie bardziej czułych metod jest również dopuszczalne.

## **IV. Sposób wykonania analiz:**

### **Pkt 1. Transkryptomika:**

- a) 40 wskazanych próbek powinno być zsekwencjonowane techniką paired-end 2 x 150 bp lub dłuższych o pokryciu (coverage) co najmniej 35 mln par odczytów na próbkę po przycięciu odczytów pod kątem jakości (Quality trimming) i usunięciu odczytów rRNA.
- b) pozostałe 248 próbek powinno być zsekwencjonowane w technologii single-end 50 bp lub dłuższym o pokryciu co najmniej 35 mln odczytów na próbkę po przycięciu i usunięciu odczytów rRNA.

### **Pkt 2. Metabolomika:**

288 próbek analizowanych metodą chromatografii gazowej sprzęgniętej ze spektrometrią mas (GC-MS) lub inną (pkt. III ppkt 2).

## **V. Jakość analizy:**

### Pkt 1. Transkryptomika:

Otrzymane sekwencje powinny być w formie plików fastq z kontrolowaną jakością i przycięte, o minimalnej długości odczytów powyżej 30 bp ze wskaźnikiem jakości Phred >20, nie powinny zawierać sekwencji adaptorowych, sekwencje odpowiadające rybosomalnemu RNA powinny być odfiltrowane i załączone w osobnym pliku. Zamawiający zaakceptuje możliwość nieudanych analiz na poziomie nieprzekraczającym 3% całkowitej liczby analizowanych prób z powodów innych niż jakość RNA.

### Pkt 2. Metabolomika:

Jakość każdego wstrzyknięcia należy oceniać za pomocą kontroli wewnętrznej (np. z użyciem rybitolu). Wyniki nie powinny się różnić od wartości spodziewanych więcej niż 30%. Liczbę i ilość poszczególnych metabolitów należy podawać na podstawie powierzchni pików w oparciu o referencyjną bazę danych Fiehn lub inną zawierającą minimum 500 metabolitów i po ręcznej korekcie profili. (tj. ręcznej integracji wysokości piku, jego powierzchni oraz czasu retencji dla danego metabolitu). Zamawiający, na podstawie własnych wstępnych badań, załącza listę 120 metabolitów i wymaga aby minimum 80% z nich zostało przedstawione w wynikach analiz dla każdej próby. Zamawiający zaakceptuje możliwość nieudanych analiz na poziomie nieprzekraczającym 3% całkowitej liczby analizowanych prób. Otrzymane wyniki powinny być w formie plików Excel, bądź równoważnym formacie tabelarycznym (np. csv, tsv).

## **VI. Wyniki:**

Wyniki powinny być przekazane zamawiającemu w terminie wykonania zamówienia określonym w SIWZ, tj. nie później niż w ciągu 40 tygodni od dnia odbioru próbek przez Wykonawcę, potwierdzonego zgodnie z zapisem § 7 ust. 1 pkt 2 projektu umowy - *z uwzględnieniem zapisów rozdziału XIV SIWZ – termin ten jest jednym z kryteriów oceny ofert*. Wyniki analiz będą udostępnione na bezpiecznym serwerze FTP przez okres minimum 3 tygodni.

## **VII. Inne wymagania:**

Pkt 1. Ze względu na opisane we wstępie wyzwania stojące przed projektem jak i specyfikę prób opisaną w części wstępnej, przygotowanie bibliotek do RNA-seq oraz przygotowanie prób do analizy metabolomu są kluczowym krokiem w każdej z analiz gdyż na tym etapie może zdarzyć się najwięcej błędów technicznych. W konsekwencji Zamawiający wymaga aby przynajmniej te analizy były wykonane w całości przez Wykonawcę.

Pkt 2. Wykonawca jest zobowiązany przyjąć dwie osoby na staż naukowy odbywany w czasie zleconych analiz i w miejscu ich wykonywania. Celem stażu będzie przeszkolenie w zakresie stosowanych do badań technologii. Komunikacja w języku polskim i/lub angielskim. Koszty podróży i zakwaterowania tych osób pokrywa Zamawiający i nie powinny być wliczone cenę oferty.

Przeszkolenie będzie obejmowało następujące zagadnienia:

- prezentacja i omówienie protokołów przeprowadzanych analiz,
- prezentacja urządzeń analitycznych i szkolenie w ich użytkowaniu,
- obserwacji i omówienie poszczególnych czynności podczas analiz wykonywanych przez specjalistę ze strony Wykonawcy,
- udział w dyskusji na temat wyników analiz.

LISTA 120 METABOLITÓW

- sucrose
- D-glucose 1
- phosphoric acid
- citric acid
- L-asparagine 2
- L-glutamic acid 3 (dehydrated)
- D-malic acid
- L-glutamic acid 2
- aspartic acid 2
- L-glutamine 3
- L-tryptophan 2
- fructose 1
- pyruvic acid
- fructose 2
- L-serine 2
- D-glucose 2
- succinic acid
- glyceric acid
- L-threonine 2
- L-serine 1
- myo-inositol
- aspartic acid 1
- L-valine 2
- ethanolamine
- glyoxylic acid
- L-histidine 3
- shikimic acid
- threonic acid
- palmitic acid
- trans-4-hydroxy-L-proline 2
- norvaline 1
- DL-isoleucine 2
- L-tyrosine 2
- beta-cyano-L-alanine
- D-mannose 1
- glycine
- xanthotoxin 2
- D-(+) trehalose
- phenylalanine
- L-leucine 2
- L-threonine 1
- gamma-aminobutyric acid (GABA)
- levoglucosan
- 2-isopropylmalic acid
- galactonic acid 2
- 3-hydroxy-3-methylglutaric acid
- D-mannitol
- L-lysine 2
- glycerol 1-phosphate
- L-proline 2
- putrescine
- linoleic acid
- L-ornithine 1
- threonic acid-1,4-lactone B
- pipercolic acid 1
- maleic acid
- caffeic acid
- phytosphingosine 2
- D-threitol
- 2-hydroxypyridine
- L-methionine 2
- fumaric acid
- glycerol
- citramalic acid
- gluconic acid 2
- alpha ketoglutaric acid
- L-glutamine 1
- pipercolic acid 2
- citrulline 2
- acetohydroxamic acid
- L-(+) lactic acid
- benzoic acid
- quinic acid
- D-glucose-6-phosphate 1
- maltose 1
- allantoin 3
- dehydroascorbic acid 3
- neohesperidin
- 4-hydroxycinnamic acid
- lactose 1
- 4-hydroxybenzoic acid
- trans-aconitic acid
- O-phosphocolamine
- nicotinic acid
- Phenylalanine 1
- 1,6-anhydro-glucose
- allantoin 1
- arabinose
- xylitol
- glycolic acid
- 1,3-diaminopropane 1
- xylose 2
- urea
- arbutin
- 4-hydroxy-3-methoxybenzyl alcohol
- beta-sitosterol
- lauric acid
- tartaric acid
- L-alanine 1
- L-glutamine 2
- rhamnose 1
- ribose
- acetyl-L-serine 2
- 4-hydroxybenzaldehyde 1
- 4-hydroxy-3-methoxybenzoic acid
- Beta- alanine 1
- uracil
- alpha tocophereol
- galacturonic acid 1
- N-methylglutamic acid 4
- 4-hydroxypyridine
- ferulic acid
- L-methionine sulfoxide 3
- galacturonic acid 2
- N-acetyl-D-glucosamine 1
- suberic acid
- beta-gentiobiose 1
- D-glucose-6-phosphate 2
- 1-hexadecanol
- 3-phosphoglycerate