



## Biblioteki metagenomowe jako źródło genów przydatnych w biotechnologii

Edyta Deja-Sikora, Marcin Sikora, Marcin Gołębiowski,  
Andrzej Tretyn

Zakład Biotechnologii, Wydział Biologii i Nauk o Ziemi, Uniwersytet  
Mikołaja Kopernika, Toruń

### Metagenomic libraries as sources of genes useful for biotechnology

#### Summary

The vast majority of microorganisms cannot be cultured under laboratory conditions. It was estimated that over 99% of microbial genetic information are inaccessible due to the inability to isolate and culture bacteria. It is also widely known that among those uncultured organisms there are such that bear the genes which are interesting from the biotechnological point of view, i.e. coding for novel enzymes (lipases, amylases, cellulases, polymerases etc.), responsible for resistance to various chemical substances (heavy metals, aromatic compounds, pesticides, antibiotics), or the genes encoding the elements of biosynthetic pathways (for instance producing novel antibiotics). Recently, the methods have been developed that allow (i) isolation and purification of environmental DNA, (ii) construction of random fragment libraries of such DNA, and (iii) effective screening of those libraries in search for interesting genes. These methods are collectively known as 'metagenomics' or 'environmental genomics'.

We aim to review the metagenomic methods, which will be done in Part I of our paper, and to present the up – to – date achievements and future perspectives for obtaining biotechnologically important genes from environmental samples in Part II. The main attention will be paid to soil metagenomics, as this kind of environment seems to be the most promising in terms of microbial biodiversity and the spectrum of biochemical reactions performed by inhabiting bacteria.

We will treat the perspectives for isolation of novel, useful genes such as those coding for biosynthesis of antibiotics, organic compounds degrading pathways, and heavy metal resistance and prospects for their biotechnological application. Assessing the microbial biodiversity through metagenomic methods will also be covered.

#### Key words:

metagenomics, environmental DNA, metagenomic libraries, screening, polluted environments, bioremediation, biodiversity.

#### Adres do korespondencji

Marcin Gołębiowski,  
Zakład Biotechnologii,  
Wydział Biologii i Nauk  
o Ziemi,  
Uniwersytet Mikołaja  
Kopernika,  
ul. Gagarina 9,  
87-100 Toruń;  
e-mail:  
mgoleb@biol.uni.torun.pl

## 1. Wstęp

Wszystkie ekosystemy funkcjonujące w przyrodzie zdominowane są przez jednokomórkowe organizmy prokariotyczne. Główną grupę wśród nich stanowią eubakterie. Dzięki swojej różnorodności fizjologicznej, plastyczności, zdolności szybkiego wzrostu i adaptacji do zmieniających się warunków środowiskowych, bakterie opanowały nisze ekologiczne niedostępne dla organizmów eukariotycznych. Ważną rolę w zasiedlaniu przez nie wszelkich możliwych środowisk odgrywa ogromny potencjał biochemiczny komórek, które przystosowały się do wykorzystywania bardzo złożonych źródeł energii, dostępnych w środowisku. Aktywności metaboliczne poszczególnych grup mikroorganizmów oraz zachodzące pomiędzy nimi interakcje regulują funkcjonowanie całych ekosystemów poprzez rozkład i przekształcanie materii organicznej i nieorganicznej, co jest podstawą obiegu pierwiastków w przyrodzie.

Potencjał biochemiczny mikroorganizmów ma swoje źródło w olbrzymiej różnorodności informacji genetycznej zawartej w ich komórkach. W genomach bakterii znajdują się użyteczne geny mogące kodować np. nie znane do tej pory enzymy o znaczeniu przemysłowym, oporność na antybiotyki, metale ciężkie czy pestycydy, szlaki degradacji trudno rozkładalnych związków organicznych (węglowodory aromatyczne i ich halogeno-pochodne, pestycydy itp.), czy wreszcie szlaki biosyntezy użytecznych związków (np. nowych rodzajów antybiotyków).

### 1.1. Bioróżnorodność mikroorganizmów

Na podstawie wyników ostatnio przeprowadzanych badań w celu oszacowania liczebności organizmów prokariotycznych sugeruje się, że ogólna liczba komórek bakteryjnych, żyjących na Ziemi, zbliżona jest do wartości  $4-6 \times 10^{30}$ , z czego około  $2,6 \times 10^{29}$  komórek egzystuje w glebie (1). Identyfikacja i porównywanie ze sobą genów kodujących 16S rRNA małej podjednostki rybosomalnej (SSU, ang. *small subunit*), ze względu na bardzo niską zmienność sekwencji, dostarcza informacji o obecności poszczególnych grup bakteryjnych w danym środowisku. Pozwala również na wykrycie nowych i nie scharakteryzowanych do tej pory rodzin (2).

Szacuje się, że w jednym gramie gleby znajduje się około  $10^9$  organizmów prokariotycznych i więcej niż dwa tysiące różnych typów genomów. Średnia reprezentacja jednego rodzaju genomu wynosi zatem mniej niż 0,05% (3). Analizując długości poszczególnych typów genomów mikroorganizmów glebowych, zauważono, że ich łączny rozmiar przewyższa długość genomu *Escherichia coli* około 6-10 tysięcy razy w glebach nie zanieczyszczonych oraz 350-1500 razy w glebach wysoko zanieczyszczonych metalami ciężkimi (4). Ponadto podczas badania różnorodności glebowych organizmów prokariotycznych, wykazano, że tylko 0,1-1,0% bakterii może być pozyskanych ze środowiska za pomocą tradycyjnych metod mikrobiologicznych i następ-

nie hodowanych w warunkach laboratoryjnych (4,5). Wynik ten wskazuje, że pozostałe 99% glebowej populacji bakterii pozostaje niezbadane i może być źródłem nieznanymi genów.

## 1.2. Czym jest metagenomika?

Ogromna liczebność, różnorodność form i szeroka gama przeprowadzanych przez bakterie procesów biochemicznych nakierowały uwagę badaczy na ogromne bogactwo puli genetycznej tych organizmów (6). Badania nad dostępem do „genów środowiskowych” zaowocowały wyodrębnieniem nowej gałęzi genomiki, która nazwana została metagenomiką. Określenie to wywodzi się z połączenia dwóch terminów – metaanalizy i genomiki. Metaanaliza oznacza proces statystycznego łączenia oddzielnych analiz lub wtórnego odkrywania wiedzy metodą uogólniania informacji zawartych w publikacjach naukowych lub źródłach pierwotnych, natomiast genomika obejmuje kompleksową analizę materiału genetycznego organizmów (6). Jej pole zainteresowania obejmuje budowę i działanie genomów, poszczególnych genów i operonów, a także badanie produktów ich ekspresji (6). Na tej podstawie dąży się do poznania: różnorodności organizmów prokariotycznych, zależności pomiędzy budową ich genomów a funkcją środowiskową oraz do zidentyfikowania nowych genów kodujących białka o pożądanym funkcjach (7,8). Od strony metodologicznej metagenomika wykorzystuje podstawowe techniki biologii molekularnej takie jak: ekstrakcja kwasów nukleinowych ze środowiska, ich amplifikacja za pomocą reakcji PCR oraz bezpośrednio klonowanie DNA środowiskowego w celu konstrukcji tzw. bibliotek metagenomowych.

Sformułowanie „metagenom” po raz pierwszy pojawiło się w literaturze naukowej w 1998 r. i zostało użyte jako określenie obejmujące genomy wszystkich mikroorganizmów występujących w przyrodzie (9). Do tej pory analizie metagenomicznej poddano mikroorganizmy bytujące w kilku środowiskach, m. in. w wodach morskich (10) i osadach czynnych (11). Bardzo dużym zainteresowaniem cieszy się również gleba (12).

## 2. Izolacja DNA metagenomowego

Wyzolowanie środowiskowego DNA jest pierwszym etapem każdego doświadczenia metagenomicznego. Opracowanie prostej i wydajnej techniki izolacji jest trudne i wciąż pozostaje przedmiotem badań. Dobra metoda pozyskiwania środowiskowego DNA powinna (13):

- zapobiegać zbytnej fragmentacji materiału genetycznego na skutek działania czynników fizycznych,
- uniemożliwiać degradację DNA, będącą wynikiem działania nukleaz,

– zapewniać uzyskanie materiału genetycznego o wysokiej jakości i niskim stopniu zanieczyszczenia substancjami hamującymi reakcje enzymatyczne.

Wyróżniamy dwa typy izolacji DNA metagenomowego: metodę bezpośrednią i pośrednią (14). Metoda bezpośrednia polega na przeprowadzeniu lizy komórek bakteryjnych, znajdujących się w swoim naturalnym środowisku (np. macierzy glebowej) (15). Uwolniony materiał genetyczny kontaktuje się zatem ze związkami obecnymi w środowisku, np. w przypadku gleby z kwasami humusowymi, stanowiącymi trudne do usunięcia zanieczyszczenia hamujące reakcje enzymatyczne. Konieczne jest zatem doczyszczanie DNA w celu przygotowania go do tych reakcji. Główną zaletą metod bezpośrednich jest uzyskanie dużej ilości materiału genetycznego, dobrze reprezentującego różnorodność metagenomu. Pominięcie etapu izolacji bakterii ze środowiska nie dyskryminuje mikroorganizmów zaadsorbowanych na powierzchni cząstek stałych (14). Największymi wadami tych metod są: zanieczyszczenie wyizolowanego DNA inhibitorami reakcji enzymatycznych (16) oraz DNA eukariotycznym, a także znaczna fragmentacja DNA (17). Metody bezpośrednie stosowane są najczęściej do izolacji DNA służącego jako matryca w reakcji PCR bądź do konstrukcji bibliotek o małych wstawkach (18).

W metodzie pośredniej bakterie wyodrębnione ze środowiska, np. na drodze wiorowania w gradiencie gęstości, mogą być poddane standardowej lizie alkalicznej (14,19), enzymatycznej (20,21) lub łagodnej lizie w żelu agarozowym i elektroforezie (22). Metoda pośrednia została opracowana aby zmniejszyć ilość zanieczyszczeń chemicznych oraz zredukować mechaniczną i enzymatyczną degradację materiału genetycznego, co pozwala uzyskać fragmenty DNA o długości powyżej 300 kpz (22). Ostateczny wybór techniki izolacji uwarunkowany jest właściwościami środowiska i późniejszym zastosowaniem DNA (14).

Omówienie metod izolacji i oczyszczania metagenomowego DNA dotyczy głównie gleb, jako środowiska najbogatszego w mikroorganizmy.

## 2.1. Metody bezpośrednie

Na metody te składają się dwa krytyczne etapy: pierwszy polega na zniszczeniu ścian i błon komórkowych mikroorganizmów w celu uwolnienia kwasów nukleinowych do buforu ekstrakcyjnego, drugi to oddzielenie DNA od pozostałych wyekstrahowanych substancji.

Wyróżnia się trzy rodzaje lizy komórek: mechaniczną (fizyczną), chemiczną i enzymatyczną. Sposoby te stosuje się oddzielnie lub w kombinacjach.

Metody fizyczne niszczą struktury komórkowe i powodują rozdrobnienie agregatów glebowych, co umożliwia dostęp do komórek znajdujących się w mikroporach macierzy glebowej. Destrukcji ulegają zarówno postaci wegetatywne, jak i spory. Pozwalają one uzyskać największe ilości DNA, jednak jest ono mocno pofragmentowane. Średnia wielkość fragmentów przy zastosowaniu lizy mechanicznej wy-

nosi od 600 pz do 25 kpz, co umożliwia wykorzystanie ich do konstrukcji bibliotek plazmidowych, fagowych lub kosmidowych i przeprowadzenie reakcji PCR. Do metod fizycznych zalicza się: zamrażanie-rozmarzanie oraz zamrażanie-gotowanie (23), rozdrabnianie homogenizatorem ostrzowym (ang. *warring blender*) (24), rozdrabnianie w tak zwanym młynie ziarnowym (ang. *bead-beater*) (25), rozdrabnianie w moździerzu w obecności ciekłego azotu (26), ultrasonifikacja i mikrofalowy szok termiczny (23).

Metody chemiczne są łagodniejsze niż metody fizyczne. Wymagają wstępnego rozdrobnienia materiału co umożliwia dostęp buforu lizującego do komórek bytujących wewnątrz agregatów glebowych. Do lizy używa się najczęściej SDS z dodatkiem CTAB, który działa jako detergent i w obecności wysokiego stężenia NaCl wytrąca z roztworu związki polifenolowe (27). W niektórych protokołach stosuje się związki denaturujące, takie jak izotiocyanian guanidyny (28), czy ditiotreitol (DTT) (19). Zapobiegają one degradacji DNA dezaktywując nukleazy poprzez redukcję mostków dwusiarczkowych.

Metody enzymatyczne są najłagodniejsze dla DNA. Opierają się na enzymatycznym trawieniu ścian komórkowych, co jest szczególnie użyteczne w przypadku bakterii gramdodatnich, odpornych na metody mechaniczne i chemiczne. Metody enzymatyczne stosuje się zwłaszcza, gdy wielkość DNA ma podstawowe znaczenie (na przykład w konstrukcji bibliotek BAC). Enzymów używa się również do niszczenia nukleaz degradujących DNA i do usuwania zanieczyszczającego RNA. Najczęściej stosowane enzymy to lizozym (E.C. 3.2.1.17) (20,21), achromopeptydaza (E.C. 3.4.21.50; niszczy ściany komórkowe, skuteczna w przypadku bakterii opornych na lizozym) (29), proteinaza K (E.C. 3.4.21.64) (27) i RNAza A (rybonukleaza A, EC 3.1.27.5).

## 2.2. Metody pośrednie

W metodach pośrednich pierwszym i najważniejszym etapem jest dyspersja macierzy glebowej i wyizolowanie jak największej ilości nienaruszonych komórek bakteryjnych, co gwarantuje otrzymanie nieuszkodzonego DNA. Oddzielone komórki muszą reprezentować możliwie najpełniej różnorodność mikroflory próbki gleby. Drugi etap to liza komórek i izolacja DNA. Ostatnim etapem jest oczyszczenie kwasów nukleinowych.

Do rozdrabniania gleby stosuje się metody chemiczne i fizyczne.

Najpowszechniej stosowane metody fizyczne to homogenizacja homogenizatorem ostrzowym, sonifikacja i łagodne rozpraszanie przez wytrząsanie i wirujący tłuczek. Najskuteczniejsze w rozpraszaniu cząsteczek gleby są homogenizatory ostrzowe i wirujące tłuczki (30), których zastosowanie jest możliwe dla małych i dużych próbek.

Metody chemiczne stosowane są najczęściej w połączeniu z fizycznymi. Efektywne okazały się żywice kationowymienne (31), stosowano również związki takie jak

cholinian sodu i deoksycholinian sodu (32) oddziałujące na bakteryjne lipopolisacharydy oraz glikol polietylenowy (PEG) i SDS, które rozpuszczają materiał hydrofobowy (28). Poważną wadą metod chemicznych jest uszkodzanie komórek bakteryjnych.

Innym sposobem oddzielenia komórek bakterii od macierzy glebowej jest wirowanie oparte na różnicach w sedymentacji między poszczególnymi komponentami próbki (32). Metoda składa się z dwóch następujących po sobie wirowań: przy niskim przyspieszeniu, służącego do usunięcia dużych fragmentów gleby i plech grzybowych oraz przy dużym przyspieszeniu, podczas którego wirując supernatant uzyskuje się frakcję bakteryjną w postaci osadu. Jeden cykl wirowań pozwala na oddzielenie około 10% zawartych w próbce gleby bakterii i według autorów metody, ilość taka reprezentuje całą różnorodność biologiczną próbki. Kolejne cykle zwiększają tylko ilość otrzymanego materiału genetycznego (24).

Kolejną z metod jest wirowanie w gradiencie gęstości. Jako medium gradientowe testowano m.in. Metryzamid, Percoll (33), Nycodenz (30) i sacharozę (34). Wydajność tej metody oddzielania bakterii od gleby to od 6 do 50% ogólnej liczby komórek bakterii zawartych w próbce gleby. Skuteczność zależy głównie od składu gleby. Najtrudniejsze do obróbki są gleby z dużą zawartością gliny. W porównaniu z metodą sedymentacyjną wirowanie w gradiencie Nycodenzu pozwala na otrzymanie mniej zanieczyszczonych komórek bakteryjnych (24).

Po oddzieleniu komórek od gleby następuje izolacja i doczyszczanie DNA. Procedury izolacji są podobne jak w przypadku metod bezpośrednich. W celu usunięcia niektórych zanieczyszczeń glebowych opracowano dodatkowe protokoły. Delikatna liza komórek połączona z wirowaniem w gradiencie chlorku cezu jest metodą pozwalającą uzyskać DNA o wysokiej czystości i rozmiarach przekraczających 100 kbp, jest zatem niezwykle przydatna przy konstruowaniu bibliotek z użyciem BAC (ang. *Bacterial Artificial Chromosome*) (19). Skompilowana metoda uwzględniająca izolowanie bakterii przez wirowanie w gradiencie, lizę komórek w korkach agarozowych i elektroforezę pulsacyjną, pozwoliła na uzyskanie fragmentów o długości powyżej 300 kbp i czystości odpowiedniej dla klonowania i wytworzenia biblioteki fosmidowej lub z użyciem BAC (22).

### **3. Oczyszczanie metagenomowego DNA po izolacji**

DNA po izolacji ze środowiska, a zwłaszcza po ekstrakcji z gleby, musi zostać poddane oczyszczeniu. Surowy ekstrakt zawiera najczęściej zanieczyszczenia uniemożliwiające zastosowanie technik biologii molekularnej takich jak na przykład PCR czy hybrydyzacja oraz powodujące inaktywację enzymów restrykcyjnych i ligazy. Najpoważniejszym zanieczyszczeniem preparatów DNA lub RNA izolowanych z gleby są kwasy humusowe. Ich usunięcie umożliwia przeprowadzenie reakcji PCR, odwrotnej transkrypcji, trawienia czy ligacji. Techniki oczyszczania obejmują:



- zastosowanie rozpuszczalników organicznych takich jak fenol i chloroform,
- wirowanie w gradiencie chlorku cezu (CsCl) lub chlorku cezu-bromku etydyiny (CsCl-EtBr) (35),
- oczyszczanie na kolumnach hydroksypatytowych z użyciem mocznika (25),
- elektroforeza w żelu agarozowym,
- molekularne sączenie żelowe (26),
- chromatografia jonowymienna z użyciem komercyjnych zestawów do oczyszczania DNA (36).

#### 4. Charakterystyka bibliotek metagenomowych

Biblioteka genomowa to kolekcja klonów zawierających losowe fragmenty heterologicznego DNA, reprezentujące cały genom określonego organizmu lub grupy organizmów (37). Reprezentatywność biblioteki (prawdopodobieństwo znalezienia w niej klonu zawierającego określoną sekwencję) związana jest z całkowitym rozmiarem klonowanego DNA, średnią długością wstawki i liczbą klonów. Ze względu na duży łączny rozmiar DNA metagenomowego skonstruowanie reprezentatywnej biblioteki wymaga zgromadzenia dużej liczby klonów. Teoretycznie bank DNA reprezentujący cały metagenom wyizolowany z próby glebowej, powinien zawierać w ok.  $10^6$  klonów z insertami o długości 100 kpz (38). Kalkulacja ta oparta jest na założeniu, że wszystkie typy genomów posiadają równe reprezentacje w wyizolowanym DNA. W rzeczywistości w glebie występuje także frakcja bardzo rzadkich genomów, co sprawia, że liczba klonów potrzebna do odzyskania całej informacji genetycznej powinna być dużo większa niż  $10^6$  (12). Jeżeli biblioteka przeszukiwana będzie pod kątem funkcji, prawdopodobieństwo zidentyfikowania genu zależy dodatkowo od spodziewanego typu ekspresji tego genu w komórce gospodarza (39).

Biblioteki metagenomowe reprezentują zbiór wszystkich genomów prokariotycznych, występujących w danym środowisku. Konstruowane są według standardowej procedury, wymagają jednak modyfikacji protokołów stosowanych na poszczególnych etapach. Najistotniejsze w konstrukcji biblioteki jest dobranie odpowiedniej pary gospodarz-vektor.

Gospodarzem najczęściej wykorzystywanym do klonowania są specjalnie skonstruowane szczepy *Escherichia coli* (12,37,38). Spośród laboratoryjnych szczepów *E. coli* do konstrukcji bibliotek metagenomowych najczęściej wykorzystywane są DH10B i DH5 $\alpha$ , zdolne do efektywnego pobierania i stabilnego utrzymywania fragmentów DNA o długości większej niż 100 kpz (38). Wykazano, że w komórkach *E. coli* zachodzi ekspresja heterologicznego DNA, pochodzącego od *Bacillus cereus* (2). Jednak zastosowanie *E. coli* jako jedyne gospodarza ogranicza różnorodność genów ulegających ekspresji, dlatego powstała koncepcja wykorzystania innych bakterii, np. *Streptomyces lividans* i *Pseudomonas putida* (40).

*S. lividans* to bakterie należące do rzędu *Actinomycetales*, który znany jest z olbrzymiego potencjału metabolicznego i produkcji licznych metabolitów wtórnych (np. antybiotyków, poliketonów) oraz enzymów (19,41). Jako gospodarza do ekspresji bibliotek metagenomowych stosuje się szczep zmodyfikowany genetycznie. Z jego chromosomu usunięto operony *red* i *act*, odpowiedzialne za syntezę barwników i antybiotyków – brak zatem endogennych aktywności, które mogą interferować z detekcją heterologicznych genów (40). Ekspresja metagenomowego DNA w komórkach *S. lividans* możliwa jest dzięki konstrukcji wektorów bifunkcyjnych (wahadłowych, ang. *shuttle vectors*). Zawierają one sekwencję *oriT*, umożliwiającą transfer wektora z *E. coli* do *S. lividans* na drodze mobilizacji przez plazmidy IncP oraz kasetę *attP-int*, pochodzącą od faga  $\phi$ C31 umożliwiającą integrację z miejscem *attB* na chromosomie (40). Ponadto wektor wyposażony jest w marker oporności – np. na tiosrepton (Th). Tak skonstruowany system zapewnia stabilne utrzymywanie obcego DNA w formie zintegrowanej z chromosomem *S. lividans* (42).

Innym typem bakteryjnego gospodarza do ekspresji metagenomowego DNA jest niepatogenny, glebowy szczep *Pseudomonas putida* KT 2240 (43). Modyfikacja tego szczepu, umożliwiającą integrację wektorów wahadłowych, polega na wprowadzeniu do chromosomu *Pseudomonas* sekwencji *attB* (40).

Wykazano, że wektory wahadłowe na bazie BAC, mogą przenosić i utrzymywać fragmenty DNA o wielkości 120 kbp w gospodarzu innym niż *E. coli* (42). Wysoka wydajność transferu koniugacyjnego sprawia, że ekspresja i przeszukiwanie bibliotek o dużych insertach w gospodarzach takich jak *Streptomyces* i *Pseudomonas* jest znacznie łatwiejsze (40).

Wybór wektora do tworzenia biblioteki jest bardzo istotny i decyduje o właściwościach otrzymanego banku genów. Biblioteki metagenomowe klasyfikuje się na podstawie typu wykorzystanego wektora. Można wyróżnić biblioteki oparte na: małych plazmidach bakteryjnych i wektorach bakteriofagowych (18), kosmidach i fosmidach (13,44), sztucznych chromosomach bakteryjnych (BAC) (45).

#### **4.1. Biblioteki oparte na małych plazmidach o wysokiej liczbie kopii i fagach**

Wektory plazmidowe i bakteriofagowe stosowane są do konstrukcji bibliotek metagenomowych, zawierających małe fragmenty heterologicznego DNA. Ich pojemność jest bardzo ograniczona, ponieważ maksymalny rozmiar przyjmowanego przez nie insertu wynosi od 10 kbp (dla plazmidów) do 20 kbp (dla wektorów fagowych). Oba typy wektorów posiadają na ogół elementy, które umożliwiają ekspresję obcych genów w komórkach gospodarza (46). Wśród plazmidów bardzo popularne są wektory pUC18/19.

Wektory fagowe znacznie różnią się od plazmidowych i najczęściej oparte są na genomie bakteriofaga lambda. Ligacja heterologicznego materiału genetycznego



z ramionami faga daje długie, liniowe cząsteczki DNA, które następnie są fragmentowane i pakowane do białkowego kapsydu (46). Zaletą stosowania wektorów przenoszących małe fragmenty DNA jest łatwość ich izolacji oraz proste i dobrze opracowane metody transformacji i dalszej manipulacji materiałem genetycznym. Wadą jest mała pojemność, która umożliwia odnajdywanie pojedynczych genów lub małych operonów.

#### 4.2. Biblioteki kosmidowe i fosmidowe

Kosmidy i fosmidy są wektorami do klonowania fragmentów DNA o długości 30-45 kpz. Kosmidy łączą w sobie cechy wektorów plazmidowych i fagowych. Niosą plazmidową sekwencję *ori*, gen oporności na antybiotyk oraz sekwencję *cos* (umożliwiającą pakowanie do kapsydów fagowych). Wektory kosmidowe wprowadzane są do komórki gospodarza jako fagi, natomiast wewnątrz tych komórek namnażają się jak typowe plazmidy. Fosmidy to odmiana wektorów kosmidowych, opartych na replikonie plazmidu *F. coli*, zdolnych do przenoszenia fragmentów DNA o długości do 45 kpz. Po transformacji komórka bakteryjna utrzymuje tylko jedną kopię fosmidu, dzięki czemu materiał genetyczny klonowany w wektorze jest bardziej stabilny (45).

#### 4.3. Biblioteki oparte na sztucznych chromosomach bakteryjnych (BAC)

Spośród wszystkich wektorów dostępnych do konstrukcji bibliotek metagenomowych, sztuczne chromosomy bakteryjne (BAC) wykorzystywane są najczęściej. Wektory te odznaczają się bardzo wysoką pojemnością, pozwalającą na utrzymywanie w komórce insertów DNA o wielkości do 350 kpz oraz szeregiem cech, które czynią je dogodnym narzędziem do klonowania obcych fragmentów DNA (17). Sztuczne chromosomy bakteryjne skonstruowane zostały w oparciu na replikonie plazmidu *F. coli* (45). Sztuczne chromosomy bakteryjne dają możliwość identyfikacji całego zespołu genów tworzących szlak metaboliczny (39), oraz wykazują dużą stabilność materiału genetycznego, co zapobiega jego rearanżacjom (38). Wadą wektorów BAC jest trudność w izolacji (ze względu na ekstremalnie niską liczbę kopii) oraz transformacji nimi komórek gospodarza.

### 5. Ocena bioróżnorodności mikroorganizmów w środowisku

Izolacja DNA metagenomowego w celu prowadzenia badań filogenetycznych, opartych na analizie sekwencji genów 16S rRNA, stanowi obecnie jedną z najdogodniejszych metod charakteryzowania mikroorganizmów glebowych (47). Identyfikacja i porównywanie ze sobą tych genów dostarcza informacji o obecności poszczegól-

gólnych grup bakteryjnych w metagenomie, oraz pozwala na wykrycie nowych i nie scharakteryzowanych do tej pory rodzin i gatunków (2). Na podstawie analizy sekwencji genów 16S rRNA, wyodrębnionych z ekosystemów wysoko zanieczyszczonych, zasugerowano, że systemy te w ogromnej przewadze składają się z nie znanych jeszcze organizmów prokariotycznych, które mogą być zaangażowane w procesy bioremediacji (48).

Typowy schemat doświadczenia badającego bioróżnorodność mikroorganizmów w danym środowisku wygląda następująco: z badanej próbki izolowane jest DNA metagenomowe, które służy jako matryca do namnożenia genów 16S rRNA za pomocą zdegenerowanych primerów, specyficznych dla poszczególnych grup bakterii. Namnożone fragmenty są klonowane w wektor o wysokiej liczbie kopii, a powstała biblioteka jest sekwencjonowana. Po ustaleniu sekwencji wystarczającej liczby klonów porównuje się je z obecnymi w bazach danych, np. w Ribosomal Database Project (RDP-II) i konstruuje drzewa filogenetyczne (49).

## 6. Przeszukiwanie bibliotek metagenomowych

Biblioteki metagenomowe można przeszukiwać pod kątem identyfikacji klonów wykazujących pożądaną funkcję lub przenoszących określoną sekwencję.

Wykrywanie aktywności klonów w bibliotece zachodzi może poprzez: wizualną identyfikację charakterystycznego dla nich fenotypu (np. zabarwienia) lub bezpośrednią selekcję w warunkach wzrostowych, w których zachodzi preferowany wzrost komórek bakteryjnych, posiadających pożądaną aktywność (np. zdolność wykorzystywania danego substratu jako jedyne źródła węgla lub azotu) (43,50).

Zaletą przeszukiwania bibliotek pod kątem określonej aktywności jest możliwość detekcji środowiskowych genów kodujących nowe typy i klasy enzymów oraz całych szlaków syntezy nie znanych dotąd metabolitów wtórnych. Techniki te nie wymagają znajomości sekwencji genów i są wysoce selektywne wobec tych genów, których sekwencja jest pełna i ekspresja daje funkcjonalny produkt (12). Ponadto są one proste w wykonaniu i pozwalają na jednoczesne przetestowanie tysięcy klonów tworzących bibliotekę (50). Jednym z najtrudniejszych do rozwiązania problemów jest niewielka ilość aktywnych klonów w bibliotece, co wynika z braku ekspresji heterologicznego DNA w komórkach zastępczego gospodarza (39). Ekspresja obcego materiału genetycznego wymaga bowiem rozpoznania charakterystycznych dla innych mikroorganizmów sekwencji regulatorowych w obrębie DNA oraz obecności specyficznych dla nich czynników transkrypcyjnych. Dodatkowo formowanie się aktywnego produktu białkowego uwarunkowane jest zajściem efektywnej translacji oraz poprawnego fałdowania białka, a także obecnością odpowiednich chaperonów i kofaktorów, regulujących jego aktywność. Często wymagana jest również obecność specyficznych enzymów, uczestniczących w potranslacyjnych modyfikacjach białka oraz wydajny proces jego sekrecji (37,39).

W celu ominięcia tych ograniczeń skonstruowano nowe systemy do ekspresji heterologicznego DNA (np. wektory wahadłowe oraz zmodyfikowane szczepy *E. coli* wykazujące podwyższoną ekspresję obcych genów), a także zaprojektowano nowe metody przeszukiwania bibliotek, wykorzystujące podłoża płynne, które zapewniają swobodną dyfuzję substratów i produktów pomiędzy komórkami i dzięki temu pozwalają na identyfikację klonów uzupełniających się metabolicznie (należących do tzw. konsorcjów metabolicznych – ang. *metabolic consortium*) (44).

Na podstawie analizy dokonanej pod kątem funkcji nie można odnaleźć wszystkich potencjalnie użytecznych genów ze względu na ograniczenia opisanej ekspresji. Identyfikację takich genów pominiętych przy poszukiwaniu funkcjonalnych klonów może zapewnić przeszukiwanie pod kątem sekwencji, poprzez zastosowanie reakcji łańcuchowej polimeryzacji (PCR) (15,16), sond molekularnych (12), sekwencjonowania (10,46), mikromacierzy (51) i analiz bioinformatycznych. Detekcja poszukiwanej sekwencji za pomocą reakcji PCR lub sond molekularnych wymaga zaprojektowania wysoce specyficznych oligonukleotydów hybrydujących z konserwowanymi regionami poszukiwanych genów. Informacje na temat konserwowanych sekwencji uzyskuje się poprzez analizy genów zgromadzonych w bazach danych (12). Do sekwencjonowania najczęściej używa się bibliotek o niewielkich wstawkach w wysokokopijowych plazmidach. Ustalenie sekwencji dużej liczby klonów pozwala na złożenie genomów najbardziej reprezentatywnych organizmów i ich późniejszą analizę (52). Na podstawie porównania uzyskanych sekwencji z homologami z baz danych można wnioskować o potencjalnej użyteczności genów. Z poskładanych sekwencji genomowych uzyskuje się dodatkowe informacje, np. o kontekście w jakim występują dane geny, czy o pozycji taksonomicznej organizmu z którego pochodzą.

Mikromacierze umożliwiają jednoczesną i szybką analizę wielu tysięcy fragmentów metagenomowego DNA (51). Ich wykorzystanie w celach badawczych polega na hybrydowaniu znakowanych fluorescencyjnie cząsteczek DNA z odpowiednimi polami mikromacierzy, w obrębie których znajdują się sekwencje komplementarne do badanych prób. Dzięki temu możliwe jest ustalenie sekwencji DNA lub wzorca ekspresji genów w komórce (53).

## 7. Zastosowanie i perspektywy rozwoju metagenomiki

Największym do tej pory projektem metagenomicznym było sekwencjonowanie DNA populacji mikroorganizmów z Morza Sargassowego. Badania te prowadzono zgodnie z klasycznym schematem (izolacja, klonowanie, sekwencjonowanie). W pierwszym etapie wyizolowano z wody morskiej mikroorganizmy (filtracja, zebrano osad z filtrów o porach od 0,1 do 3  $\mu\text{m}$ ), następnie wyekstrahowano z nich całkowite DNA, które pofragmentowano na odcinki o długości od 2 do 6 kbp i sklonowano w wektor o dużej liczbie kopii. Wygenerowano w sumie 993 tysiące klonów, których inserty zsekwencjonowano z obu końców, co dało 1,63 Gbp sekwencji (52).

Na podstawie uzyskanych wyników istnieje możliwość złożenia prawie kompletnych sekwencji genomowych z losowych klonów DNA metagenomowego. Otrzymano m. in. genom organizmu zbliżonego do *Burkholderia*, dwa genomy bakterii spokrewnionych z *Shewanella oneidensis*, zestaw kontigów reprezentujących SAR86, *Prochlorococcus* i nie znaną do tej pory archebakterię. Ponadto, zidentyfikowano m. in. 1800 nie znanych do tej pory genów bakteriorodopsyny, co uwiadcza bogactwo informacji genetycznej nawet w tak oligotroficznym środowisku, jakim jest woda morska (52).

Oprócz znaczenia poznawczego, dostęp do metagenomów środowiskowych, a zwłaszcza glebowych, ma znaczenie biotechnologiczne. Gleba już od dawna uważana jest za jedno z ważniejszych źródeł substancji użytecznych w procesach produkcyjnych. Ponieważ liczba gatunków bakterii znajdujących się w jednej tonie gleby, może wynosić nawet 4 miliony (54), możliwości odkrycia nie znanych jeszcze związków aktywnych biologicznie są ogromne. Zróżnicowanie budowy i właściwości metabolicznych mikroorganizmów oraz opracowanie technik pozwalających wyizolować geny ze środowiska, przyczyniło się do zainteresowania metagenomem jako źródłem genów kodujących:

- białka odpowiedzialne za syntezę związków biologicznie czynnych takich jak antybiotyki, substancje antynowotworowe i immunosupresyjne (55),
- biokatalizatory zdolne do przeprowadzania procesów biochemicznych w warunkach przemysłowych lub katalizujące nowe typy przemian (37),
- szlaki degradacji ksenobiotyków (53) oraz odporności, np. na metale ciężkie (56).

Obecnie większość antybiotyków odkrywana jest na drodze izolacji z czystych kultur mikroorganizmów hodowanych laboratoryjnie, jednak liczba antybiotyków identyfikowanych metodami tradycyjnymi wciąż spada (4). Alternatywnym i znacznie bogatszym ich źródłem staje się metagenom glebowy, w którym możliwa jest zarówno identyfikacja nowych produktów jak i wykrywanie mechanizmów oporności na antybiotyki (4). Wyizolowano m.in. geny odpowiadające za syntezę nowego typu antybiotyków (turbomycyn A i B) (57).

Biblioteki metagenomowe umożliwiają również uzyskanie wielu rodzajów enzymów. Izolowano m.in. lipazy (58), esterazy (59), endoglukanazy (60), pektynazy (61) proteazy (62), dehydrogenazy (63) i oksigenazy (64). Spośród enzymów o dużym znaczeniu biotechnologicznym w metagenomie szczególnie poszukiwane są lipazy i esterazy, ponieważ pozostają one wysoce aktywne w rozpuszczalnikach organicznych, wykazują wysoką stereoselektywność i nie wymagają obecności kofaktorów do funkcjonowania (41). W bibliotece skonstruowanej z DNA wyizolowanego z próby gleby pobranej na pustyni alkalicznej zidentyfikowano 120 nowych genów należących do klasy lipaz i esteraz, uwiadczeniając tym samym olbrzymią różnorodność biokatalizatorów kodowanych przez metagenomowe DNA (65). Równie poszukiwane i szeroko stosowane w procesach katalizy są oksydoreduktazy wykorzystywane głównie do syntezy związków, których produkcja na drodze chemicznej jest bardzo

trudna i kosztowna, np. hydroksy- i aminokwasów oraz chiralnych alkoholi (66). Solbak (61) i Rhee (59) w swoich pracach dobrze pokazują możliwości wyizolowania z próbek środowiskowych biokatalizatorów przydatnych w przemyśle. W pierwszej z nich autorzy donoszą o izolacji nowych genów pektynaz mających zastosowanie w przemyśle tekstylnym, natomiast w drugiej opisali identyfikację termofilnych lipaz i esteraz. Szczególnie istotny w tego typu badaniach jest dobór środowiska o określonych właściwościach selekcyjnych, w którym egzystują specyficzne populacje bakteryjne. Przykładem może być poszukiwanie genów kodujących pektynazy w metagenomie wyizolowanym z gleby lasu tropikalnego bogatej w rozkładającą się materię organiczną (61).

W związku z coraz bardziej obniżającymi się kosztami sekwencjonowania i doskonaleniem narzędzi bioinformatycznych umożliwiających składanie i analizę sekwencji metagenomowych, przyszłość, jak się wydaje, leży w masowym sekwencjonowaniu bibliotek o niewielkich insertach.

## Literatura

1. Torsvik V., Øvreås L., (2002), *Curr. Opin. Microbiol.*, 5, 240-245.
2. Osoegawa K., Woon P. Y., Zhao B., Frengen E., Tateno M., Catanese J. J., de Jong P. J., (1998), *Genomics*, 52, 1-8.
3. Stein J. L., Marsh T. L., Wu K. Y., Shizuya H., DeLong E. F., (1996), *J. Bacteriol.*, 178, 591-599.
4. Steele H. L., Streit W. R., (2005), *FEMS Microbiol. Lett.*, 247, 105-111.
5. Hugenholtz P., Goebel B. M., Pace N. R., (1998), *J. Bacteriol.*, 180, 4765-4774.
6. Cowan D. A., (2000), *Trends Biotechnol.*, 18, 14-16.
7. Radajewski S., McDonald I. R., Murrell J. C., (2003), *Curr. Opin. Biotechnol.*, 14, 296-302.
8. Daniel R., (2002), *Construction of environmental libraries for functional screening of enzyme activity*, in: *Directed Molecular Evolution of Proteins*, Eds. Brakmann S., Weinheim J. K., Wiley-VCH Verlag GmbH, Weinheim.
9. Handelsman J., Rondon M. R., Brady S. F., Clardy J., Goodman R. M., (1998), *Chem. Biol.*, 5, R245-249.
10. Beja O., Suzuki M. T., Koonin E. V., Aravind L., Hadd A., Nguyen L. P., Villacorta R., Amjadi M., Garrigues C., Jovanovich S. B., Feldman R. A., DeLong E. F., (2000), *Environ. Microbiol.*, 2, 516-529.
11. Abraham W. R., Nogales B., Golyshin P. N., Pieper D. H., Timmis K. N., (2002), *Curr. Opin. Microbiol.*, 5, 246-253.
12. Daniel R., (2004), *Curr. Opin. Biotechnol.*, 15, 199-204.
13. Kim U., Birren B. W., Slepak T., Mancino V., Boysen C., Kang H. L., Simon M. I., Shizuya H., (1996), *Genomics*, 34, 213-218.
14. Bertrand H., Poly F., van Tran V., Lombard N., Nalin R., Vogel T. M., Simonet P., (2005), *J. Microbiol. Methods*, 62, 1-11.
15. Hallin S., Lindgren P., (1999), *Appl. Environ. Microbiol.*, 65, 1652-1657.
16. Tsai Y.-L., Olson B. H., (1992), *Appl. Environ. Microbiol.*, 58, 2292-2295.
17. Rondon M. R., Goodman R. M., Handelsman J., (1999), *Trends Biotechnol.*, 17, 403-409.
18. Hanahan D., (1985), *Techniques for transformation of Escherichia coli*, in: *DNA cloning, a practical approach*, Eds. Glover D. M., Homes B. D., Oxford.
19. Berry A. E., Chiocchini C., Selby T., Sosio M., Wellington E. M. H., (2003), *FEMS Microbiol. Lett.*, 223, 15-20.
20. Tsai Y.-L., Olson B. H., (1991), *Appl. Environ. Microbiol.*, 57, 1070-1074.

21. LaMontagne M. G., Michel Jr. F. C., Holden P. A., Reddy C. A., (2002), *J. Microbiol. Methods*, 49, 255-264.
22. Robe R. P., (2002), United States Patent 6989249.
23. Picard C., Ponsonnet C., Paget E., Nesme X., Simonet P., (1992), *Appl. Environ. Microbiol.*, 58, 2717-2722.
24. Holben W. E., Jansson J. K., Chelm B. K., Tiedje J. M., (1987), *Appl. Environ. Microbiol.*, 54, 703-711.
25. Bürgmann H., Pesaro M., Widmer F., Zeyer J., (2001), *J. Microbiol. Methods*, 45, 7-20.
26. Maarit-Niemi R., Heiskanen I., Wallenius K., Lindstrom K., (2001), *J. Microbiol. Methods*, 45, 155-165.
27. Turner P. C., McLennan A. G., Bates A. D., White M. R. H., (1999), *Biologia molekularna*, PWN, Warszawa.
28. Steffan R. J., Goksoyr J., Bej A. K., Atlas R. M., (1988), *Appl. Environ. Microbiol.*, 54, 2908-2915.
29. Ezaki T., Suzuki S., (1982), *J. Clinical. Microbiol.*, 16, 844-846.
30. Lindahl V., Bakken L. R., (1995), *Microbiol. Ecol.*, 16, 135-142.
31. Jacobsen C. S., Rasmussen O. F., (1992), *Appl. Environ. Microbiol.*, 58, 2458-2462.
32. Faegri A., Torsvik V. L., Goksoyr J., (1977), *Soil Biol. Biochem.*, 9, 105-112.
33. Bakken L. R., (1985), *Appl. Environ. Microbiol.*, 49, 1482-1487.
34. Mayr C., Winding A., Hendriksen N. B., (1999), *J. Microbiol. Methods*, 36, 29-33.
35. Ogram A., Saylor G. S., Barkay T., (1987), *J. Microbiol. Methods*, 7, 57-66.
36. Miller D. N., Bryant J. E., Madsen E. L., Ghiorse W. C., (1999), *Appl. Environ. Microbiol.*, 65, 4715-4724.
37. Shizuya H., Kouros-Mehr H., (2001), *Keio J. Med.*, 50, 26-30.
38. Sambrook J., Fritsch E. F., Maniatis T., (1989), *Molecular Cloning: a Laboratory Manual*, 2<sup>nd</sup> ed., Cold Spring Harbor Laboratory Press., New York.
39. Gabor E. M., de Vries E. J., Janssen D. B., (2003), *FEMS Microbiol. Ecol.*, 44, 153-163.
40. Liles M. R., Manske B. F., Bintrim S. B., Handelsman J., Goodman R. M., (2003), *Appl. Environ. Microbiol.*, 69, 2684-2691.
41. Jaeger K. E., Eggert T., (2002), *Curr. Opin. Biotechnol.*, 13, 390-397.
42. Sosio M., Giusino F., Cappellano C., Bossi E., Puglia A. M., Donadio S., (2000), *Nat. Biotechnol.*, 8, 43-345.
43. Schloss P. D., Handelsman J., (2003), *Curr. Opin. Biotechnol.*, 14, 303-310.
44. Entcheva P., Liebl W., Johann A., Hartsch T., Streit W. R., (2001), *Appl. Environ. Microbiol.*, 67, 89-99.
45. Beja O., (2004), *Curr. Opin. Biotechnol.*, 15, 187-190.
46. Rondon M. R., August P. R., Bettermann A. D., Brady S. F., Grossman T. H., Liles M. R., Loiacono K. A., Lynch B. A., MacNeil L. A., Minor C., Tiong C. L., Gilman M., Osburne M. S., Clardy J., Handelsman J., Goodman R. M., (2000), *Appl. Environ. Microbiol.*, 66, 2541-2547.
47. Abed R. M., Safi N. M. D., Köster J., de Beer D., El-Nahhal Y., Rullkötter Y., Garcia-Pichel F., (2002), *Appl. Environ. Microbiol.* 68, 1674-1683.
48. Cole J. R., Chai B., Marsh T. L., Farris R. J., Wang Q., Kulam S. A., Chandra S., McGarrell D. M., Schmidt T. M., Garrity G. M., Tiedje J. M., (2003), *Nucleic Acids Res.*, 31, 442-443.
49. Cole J. R., Chai B., Farris R. J., Wang Q., Kulam S. A., McGarrell D. M., Garrity G. M., Tiedje J. M., (2005), *Nucleic Acids Res.*, 33, D294-D296.
50. Gabor E. M., (2004), *Harvesting novel biocatalysts from the metagenome*, PhD thesis, Groningen.
51. Zhou J., Thompson D. K., (2002), *Curr. Opin. Biotechnol.*, 13, 204-207.
52. Venter J. C., Remington K., Heidelberg J. F., Halpern A. L., Rusch D., Eisen J. A., Wu D., Paulsen I., Nelson K. E., Nelson W., (2004), *Science*, 304, 66-74.
53. Eyers L., George I., Schuler L., Stenuit B., (2004), *Appl. Microbiol. Biotechnol.*, 66, 123-130.
54. Curtis T. P., Sloan W. T., Scannell J. W., (2002), *Proc. Natl. Acad. Sci. USA*, 99, 10494-10499.
55. Brady S. F., Chao C. J., Clardy J., (2002), *J. Am. Chem. Soc.*, 124, 9968-9969.
56. Strohl W. R., (2000), *Drug Discov. Today*, 5, 39-41.
57. Gillespie D. E., Brady S. F., Bettermann A. D., Cianciotto N. P., Liles M. R., Rondon M. R., Clardy J., Goodman R. M., Handelsman J., (2002), *Appl. Environ. Microbiol.*, 68, 4301-4306.



58. Henne A., Schmitz R. A., Bomeke M., Gottschalk G., Daniel R., (2000), *Appl. Environ. Microbiol.*, 66, 3113-3116.
59. Rhee J.-K., Ahn D.-G., Kim Y.-G., Oh1 J.-W., (2005), *Appl. Environ. Microbiol.*, 71, 817-825.
60. Yun J., Kang S., Park S., Yoon H., Kim M. J., Heu S., Ryu S., (2004), *Appl. Environ. Microbiol.*, 70, 7229-7235.
61. Solbak A. I., Richardson T. H., McCann R. T., Kline K. A., Bartnek F., Tomlinson G., Tan X., Parra-Gessert L., Frey G. J., Podar M., Luginbuhl P., Gray K. A., Mathur E. J., Robertson D. E., Burk M. J., Hazlewood G. P., Short J. M., Kerovuo J., (2005), *J. Biol. Chem.*, 280, 9431-9438.
62. Santosa D. A., (2001), *Mol. Biotechnol.*, 17, 59-64.
63. Henne A., Daniel R., Schmitz R. A., Gottschalk G., (1999), *Appl. Environ. Microbiol.*, 65, 3901-3907.
64. Erwin D. P., Erickson I. K., Delwiche M. E., Colwell F. S., Strap J. L., Crawford R. L., (2005), *Appl. Environ. Microbiol.*, 71, 2016-2025.
65. Miller C. A., (2000), *Biotechnol. Reports*, 11, 489-495.
66. Davis B. G., Boyer V., (2001), *Nat. Prod. Rep.*, 18, 618-640.