

Jacek Kozdrój^{1*}

¹Katedra Mikrobiologii, Uniwersytet Rolniczy im. Hugona Kołłątaja, al. Mickiewicza 24/28, 30-052 Kraków

Wpłynęło w lutym 2013 r.

1. Wstęp. 2. Problemy z poznaniem mikrobiomu oraz próby ich przewycięzania. 3. Metagenomika a specyficzny charakter gleby. 4. Analiza genetyczna mikrobiomu. 5. Analiza funkcjonalna metagenomu. 6. Strukturalna różnorodność mikroorganizmów w różnych środowiskach glebowych. 7. Podsumowanie

Metagenome – a new source of information about soil microorganisms

Abstract: Soil environment, due to its high heterogeneity, is considered as a major reservoir of microbial genetic and metabolic diversity in the biosphere. The knowledge on this diversity is limited, because most of the soil microorganisms cannot be cultured under the usual laboratory conditions. During the last two decades, development of methods to isolate nucleic acids from soil has opened a window to a previously unknown microbial world. In consequence, a new metagenomic approach based on the analyses of total microbial DNA has appeared in soil studies. Total microbial DNA extracted from soil by direct or indirect methods is mostly used for amplification of marker genes (e.g. SSU rRNA) which is further differentiated by fingerprinting (e.g. DGGE, T-RFLP) or sequenced directly. Until recently, sequencing was mainly performed after first cloning PCR products to produce a clone library of amplicons. Lately, another approach has been introduced to reduce costs and labour; it is commonly known as 454-pyrosequencing, the method that does not require cloning. These methods as well as DNA microarrays have demonstrated an unanticipated level of microbial diversity, especially in the newly discovered world of the biosphere. Thousands or even several hundred thousands of different bacterial phylotypes can be present in a gram of soil. They belong to dozens of phyla. The molecular approach changed the picture of structural diversity of soil microbiome, also indicating that bacteria, archaea, fungi and even viruses are diverse both globally and locally. Moreover, soil metagenomics, allows for a comprehensive search for gene expression and metabolic activity within microbiome.

1. Introduction. 2. Difficulties with microbiome analysis and attempts to overcome them. 3. Metagenomics vs. distinct features of soil. 4. Genetic analysis of microbiome. 5. Functional analysis of microbiome. 6. Microbial structural diversity in different soil environments. 7. Summary

Słowa kluczowe: bioróżnorodność, metagenomika, mikrobiom glebowy

Key words: biodiversity, metagenomics, soil microbiome

1. Wstęp

Mikroorganizmy stanowią najbardziej istotną a jednocześnie zróżnicowaną grupę organizmów wchodzących w skład biosfery ziemskiej [93]. Większość bioróżnorodności, spotykana w obrębie trzech domen: bakterii, archeonów oraz eukariotów, jest reprezentowana przez mikroorganizmy [97], które szeroko rozprzestrzeniły się we wszystkich typach środowisk, nawet tych ekstremalnych, w ciągu blisko czterech miliardów lat swojej ewolucji. Tak długi czas umożliwił wykształcenie się najróżnorodniejszych mechanizmów metabolicznych, które pozwoliły im zaadaptować się do prawie każdej możliwej niszy ekologicznej i wykorzystać do swego rozwoju różnorodne warunki środowiskowe [59].

Środowisko glebowe jest uważane za najbogatszy rezerwar różnorodnych mikroorganizmów [90], co nie jest trudne do zaakceptowania, zważywszy na wysoce heterogeniczny charakter warunków fizykochemicznych panujących w tym środowisku, a także uwzględniając zróżnicowanie typów gleb powstałych na Ziemi. Jedno-

ześnie, specyficzny charakter gleby powoduje, że jest ona traktowana jak swego rodzaju „czarna skrzynka”: pod względem składu mikrobiomu, aktywności i procesów przeprowadzanych przez występujące w niej mikroorganizmy. Nie jest pozbawiona racji myśl sformułowana przez Leonarda da Vinci już w 1510 roku, że być może więcej wiemy o ruchach ciał astralnych, aniżeli o tym, co dzieje się w ziemi pod naszymi stopami. Kilka wieków później równie sceptyczny był Selman Waksman, który w 1925 roku stwierdził, że: „brak jest stosownych metod w badaniach mikrobiologicznych gleby, a te z metod stosowanych obecnie są dalekie od tego, aby w adekwatny sposób dać nam wiedzę o tym, co rzeczywiście dzieje się w glebie” [46]. Co ciekawe, ten sam mikrobiolog w 1931 roku sformułował przeciwną opinię, że: „zgromadzona już olbrzymia baza informacji pozwala przedstawić jasny obraz mikroskopowej populacji mikroorganizmów glebowych” [32]. Tak diametralnie odmienne opinie mogą budzić zdziwienie i pytanie o stabilność naukową kryteriów badań mikrobiologicznych. Zgodnie z ustaleniem przyjętym

* Autor korespondencyjny: Katedra Mikrobiologii, Uniwersytet Rolniczy im. Hugona Kołłątaja, al. Mickiewicza 24/28, 30-052 Kraków; tel.: 12 662 40 96; e-mail: j.kozdroj@ur.krakow.pl

w 1923 roku przez *Bergey's Manual*, podstawą identyfikacji nowego gatunku mikroorganizmu jest konieczność sklasyfikowania jego szeregu cech morfologicznych, biochemicznych, fizjologicznych i genetycznych po wcześniejszym jego wyhodowaniu na odpowiednim podłożu [32]. Jak wielkie stanowi to wyzwanie, niech świadczy fakt znajomości tylko ponad 9000 gatunków bakterii, wobec potencjalnych 10^3 – 10^7 różnych drobnoustrojów występujących w gramie gleby [27]. Ta rozbieżność między liczbą mikroorganizmów dających się wyhodować, a tymi, o istnieniu których świadczą wyniki analiz mikroskopowych i molekularnych, jest powodem ustawicznej frustracji mikrobiologów.

2. Problemy z poznaniem mikrobiomu oraz próby ich przezwyciężania

Przez blisko sto lat rozwoju mikrobiologii zdobywana wiedza o świecie mikroorganizmów dotyczyła tych, które hodowano na różnych podłożach. Podstawowe właściwości poszczególnych gatunków (szczepów) tych organizmów określano w oparciu o wyselekcjonowane monokultury wyrosłe na podłożu hodowlanym. To podejście, będące konsekwencją postulatów K o c h a w badaniach nad drobnoustrojami chorobotwórczymi, zaważyło na bagatelizowaniu faktu obecności w środowisku form żywych, których nie można było wyhodować. Dominujący model badawczy w mikrobiologii kazał traktować je jako artefakty, a nie jako istotną grupę poszerzającą spektrum różnorodności mikroorganizmów, jak postulowali będący w mniejszości zwolennicy istnienia tych organizmów. Swoiste przesilenie, zmieniające utarte poglądy w mikrobiologii, nastąpiło w połowie lat osiemdziesiątych ubiegłego wieku, kiedy oficjalnie potwierdzono istnienie wyraźnej rozbieżności między liczebnością mikroorganizmów określoną na podłożu hodowlanym oraz w preparacie mikroskopowym, zwaną wielką anomalią liczenia metodą płytkową [78]. Jak się okazało, jedynie 0,1 do 1% bakterii glebowych może wyrosnąć na tradycyjnych podłożach hodowlanych w standardowych warunkach laboratoryjnych [44, 84]. Innymi słowy, ponad 99% różnorodnych bakterii w gramie gleby jest poza zasięgiem poznawczym i kryje w sobie rzeszę organizmów o być może istotnym znaczeniu dla funkcjonowania ekosystemu.

Wśród tego typu mikroorganizmów można wyróżnić grupę, którą da się wyhodować na specyficznych podłożach, spełniających wymogi żywieniowe organizmów, i umożliwiających ich wzrost w odpowiednich warunkach fizycznych i chemicznych. Drugą grupę natomiast stanowią mikroorganizmy, których stan fizjologiczny stanowi przeszkodę dla wzrostu na jakimkolwiek podłożu hodowlanym i dlatego są znane tylko w postaci sekwencji nukleotydowych rRNA (ryboty-

pów), jak np. WS3, OP11, OD1 lub BRC-1. Tutaj należą też powszechne w glebie mikroorganizmy, występujące w stadium spoczynkowym o ograniczonej aktywności komponentów komórkowych warunkujących wzrost – w przypadku bakterii są to formy drobne (ultra mikrobakterie). Te drzemiące formy drobnoustrojów w sprzyjających warunkach podlegają wybudzeniu, a więc teoretycznie mogą być też wyhodowane na odpowiednim podłożu. Jednakże są też takie mikroorganizmy glebowe widoczne pod mikroskopem, które można określić jako trwale niezdolne do wzrostu (nonviable) o nieodwracalnych zmianach komórkowych, lub będące w fazie zamierania [52].

Manipulacje składem podłoża hodowlanego w kierunku obniżenia jego troficzności, dobór odpowiedniego pH i temperatury, wydłużanie czasu inkubacji, stosowanie polimerowego substratu wzrostowego (np. ksylanu) oraz żelowych mikrogranul do immobilizacji komórek izolowanych ze środowiska przed umieszczeniem ich w podłożu hodowlanym, pozwalają znacząco poszerzyć spektrum możliwych do wyhodowania mikroorganizmów glebowych [10, 40, 41, 101]. Wolno rosnące bakterie glebowe z typów: *Acidobacteria*, *Chloroflexi*, *Actinobacteria* (podklasa *Rubrobacteridae*), *Planctomycetes*, *Bacteroidetes*, *Firmicutes* oraz *Proteobacteria*, niewykrywalne według tradycyjnej procedury, mogą być dostrzeżone w postaci mini kolonii (25–200 μm) po 12 tygodniowej hodowli na podłożu zawierającym ksylan jako substrat pokarmowy oraz gellan (polisacharyd produkowany przez *Pseudomonas elodea*) jako środek zestalający zamiast agaru [16]. Ten postęp w metodach hodowania pospolitych w glebie bakterii, wykrywanych jedynie na podstawie sekwencji nukleotydowej genu 16S rRNA, daje nadzieje na uzyskiwanie monokultur tych organizmów, które wymykały się spod tradycyjnej praktyki laboratoryjnej mikrobiologów. Dzięki temu możliwe jest ich pełne identyfikowanie do gatunku i bliższe poznanie ich właściwości fizjologicznych i potencjalnie zajmowanej niszy ekologicznej w glebie.

Przytłaczająca jest jednakże dysproporcja między liczbą mikroorganizmów możliwych do wyhodowania, a tymi reprezentowanymi tylko przez fragmenty genów kwasów rybonukleinowych małej podjednostki rybosomów (SSU rRNA), które *W o e s e i F o x* [96] zaproponowali jako chronometry ewolucyjne organizmów. To odkrycie dało podstawę do sformułowania paradygmatu biologicznego o podziale świata żywego na trzy domeny: *Bacteria*, *Archaea* oraz *Eucarya* [97]. *Norman P a c e* wraz ze współpracownikami, jako pierwszy wykorzystał markery komórkowe, geny 5S oraz 16S rRNA, do charakteryzowania różnorodności mikroorganizmów prokariotycznych w środowisku bez potrzeby ich hodowania [60, 77]. Początkowo taka procedura wymagała bezpośredniego sekwencjonowania RNA lub jego odwrotnych transkryptów w postaci

kopii DNA (cDNA), co było technicznie kłopotliwe. Znaczący postęp dokonał się wraz z rozwojem technologii PCR i skonstruowaniem starterów, umożliwiających amplifikację prawie kompletnego genu rRNA [29]. To zintensyfikowało wykrywanie coraz to nowych taksonów wraz z badaniem różnych biotopów ziemskich przy użyciu tej nowej techniki [3, 32, 72]. Gromadzone sekwencje nukleotydowe SSU rRNA w bazie Ribosomal Database Project (RDP) początkowo w liczbie kilkuset pozycji, szybko uległy zwielokrotnieniu do 10 tysięcy w roku 1998, 72 626 w 2003, czy 2 639 157 w grudniu 2012 (RDP-10.31). W tej ostatniej wersji bazy RDP bakterie są reprezentowane przez 1 186 838 sekwencji nukleotydowych, natomiast archeony obejmują jedynie 22 615 pozycji. Wśród sklasyfikowanych sekwencji dominują *Firmicutes* (407 377), *Proteobacteria* (322 871), *Actinobacteria* (176 308) oraz *Bacteroidetes* (135 261).

Rozmiar dysproporcji, między światem mikroorganizmów potencjalnie istniejących, a tymi, które da się wyhodować, został uwidoczniiony po zastosowaniu metody pomiaru tempa reasocjacji jednoniciowego DNA, powstałego po denaturacji całkowitego DNA w wysokiej temperaturze, wyekstrahowanego z mikroorganizmów glebowych [83]. Okazało się, że w gramie różnych gleb może występować od 1000 do 40 000 genomów mikroorganizmów, co przekracza przeszło 200-krotnie liczbę organizmów otrzymanych w hodowli [46]. Dalsze postępy w analizie reasocjowanego DNA pozwoliły oszacować różnorodność mikroorganizmów prokariotycznych w gramie gleby na poziomie od 500 tysięcy do 10 milionów genomów [27]. W tym oceanie różnych genomów identyfikacja do konkretnych taksonów jest wysoce utrudniona. Jediną pomocą są bazy sekwencji nukleotydowych genów SSU rRNA i opracowywane na ich podstawie znakowane sondy filogenetyczne, które umożliwiają także ilościowe oznaczenia występowania różnych taksonów (gatunków), analizując bezpośrednio próbkę pobraną ze środowiska za pomocą metody fluorescencyjnej hybrydyzacji *in situ* [28].

Pomimo znaczącego postępu w badaniach mikroorganizmów, oraz aby uniknąć konsekwencji jedynie katalogowania różnorodnych sekwencji nukleotydowych genów SSU rRNA znajdujących w środowisku, mikrobiolodzy podjęli się opracowania metod, umożliwiających poznanie genetyki, fizjologii i ekologii drobnoustrojów nie poddających się hodowli.

3. Metagenomika a specyficzny charakter gleby

Pełna charakterystyka organizmów wymaga poznania sekwencji nukleotydowych ich genomów. Obecnie w pełni znanych jest 2809 genomów wirusów, 2096 genomów bakterii, 153 genomów archeonów oraz 37 genomów eukariotów (wg IMG/M, Integrated Microbial

Genomes with Microbiome Samples). Liczba ta jest znikoma, aby wykorzystać ją, jako podstawę do analizy różnorodności mikroorganizmów w próbce środowiskowej, w oparciu o poszukiwanie sekwencji DNA odpowiadających znanym genomom. W dodatku taka procedura byłaby bardzo żmudna, kosztochłonna, i co najważniejsze, w niewielkim stopniu uzupełniałaby tradycyjną analizę mikrobiologiczną. Sytuacja zmienia się, jeśli zamiast sekwencjonowania poszczególnych genomów, analizie podda się zbiorczo cały DNA izolowany z próbki środowiskowej, reprezentujący wszystkie mikroorganizmy.

Ideę analizy zbioru podobnych, ale nie identycznych jednostek (genomów), stanowiących razem metagenom, wprowadziła, pod postacią metagenomiki, H a n d e l s m a n i wsp. [33], jako sposób na poznanie nieznanych mikroorganizmów glebowych i ich zdolności syntezy nowych produktów (np. antybiotyków). Klasyczna analiza metagenomiczna jest oparta na: izolacji DNA z próbki środowiskowej, klonowaniu fragmentów DNA w odpowiednim wektorze, transformowaniu otrzymanych rekombinantów do stosownego gospodarza bakteryjnego, a następnie analizy przesiewowej uzyskanych transformantów bakterii. Skonstruowana w ten sposób biblioteka klonów DNA jest dalej analizowana pod kątem zawartości sekwencji nukleotydowych różnych genomów, wykorzystując filogenetyczne markery (najczęściej gen 16S rRNA), lub ekspresji funkcjonalnych genów kodujących różne enzymy (np. lipazy, amylazy czy proteazy), genów odpowiedzialnych za biosyntezę nowych antybiotyków (np. turbomycyny A i B), albo też DNA klonów poddawane jest losowemu sekwencjonowaniu [32, 14]. Pionierskie prace, oparte na idei klonowania środowiskowego DNA sformułowanej przez P a c e i wsp. [60], dotyczyły: (a) analizy społeczności pikoplanktonu morskiego przy użyciu wektora fagowego [72]; (b) konstrukcji pierwszej biblioteki metagenomowej konsorcjum mikroorganizmów, w tym o aktywności celulolitycznej, namnożonych na wysuszonej trawie w laboratorium [35]; (c) konstrukcji pierwszej biblioteki otrzymanej bezpośrednio ze społeczności prokariotów morskich, i stwierdzenia w genomach przedstawicieli *Gamma-proteobacteria* obecności genu kodującego bakteriorodopsynę, wcześniej znanego tylko u archeonów [79].

Gleba stanowi szczególne wyzwanie dla analizy metagenomicznej. Jest to związane z jej budową strukturalną, składem chemicznym i niejednorodnym rozmieszczeniem mikroorganizmów, zależnym z kolei od tych pierwszych czynników. Różnie rozwinięta, w zależności od typu gleby, struktura agregatowa oparta na składnikach mineralno-organicznych stanowi skomplikowaną mozaikę mikrobiotopów o zmiennych czasowo i przestrzennie warunkach fizykochemicznych. W tych mikrobiotopach, o rozmiarach od 2 do 20 μm lub mniejszych niż 2 μm , drobnoustroje występują na powierzchni cząstek piasku lub kompleksów iłowo-organicznych, silnie do

nich przylegając w postaci mikrokolonii lub zasiedlają mieszczące się w nich mikropory, preferując te o średnicy 2 μm . Łączna powierzchnia, jaką zajmują mikroorganizmy, jest znikoma i stanowi jedynie od tysięcznych do milionowych części procenta ogółu dostępnej powierzchni, biorąc pod uwagę fakt, że powierzchnia frakcji ilastej może zajmować do 10^3 m^2 w gramie gleby [52]. Mikroagregaty, połączone egzopolisacharydami produkowanymi przez drobnoustroje, a także strzępkami grzybów, tworzą makroagregaty, wewnątrz i między którymi występują makropory o średnicy powyżej 6 μm . Środowiskiem wyraźnie stabilniejszym dla mikroorganizmów pod względem: zasobności w wodę, zawartości pokarmu, zaadsorbowanych makrocząstek (białka, DNA) i toksycznych gazów (CO_2 , NO_x), a także zagrożenia ze strony drapieżnych pierwotniaków, wazonkowców, skoczogonków lub roztoczy, są mikroagregaty. Dlatego też strukturalna różnorodność drobnoustrojów, czyli mozaika różnych populacji składających się na daną społeczność, jest znacząco odmienna w mikro i makroagregatach glebowych [51].

Przy badaniu tak złożonego środowiska pojawia się problem, jaką strategię poboru próbek, jeżeli chodzi o analizowany obszar, ich liczbę oraz wielkość, zastosować, aby uzyskać reprezentatywny dostęp do jak największej liczby mikroorganizmów występujących w danej glebie. Niedocenianie tego problemu, zwłaszcza w przypadku stosowania metod molekularnych w odniesieniu do bezpośrednich ekstraktów DNA/RNA, może skutkować zafałszowanym obrazem strukturalnej różnorodności i aktywności metabolicznej mikrobiomu. Analizy mikrobiologiczne gleby muszą uwzględniać zmienność bioróżnorodności w czasie [95] mierzony w skali wieloletniej, rocznej (efekt pór roku), dziennej czy nawet godzinowej (pora dnia). Drugim faktem, który należy uwzględnić, jest zmienność bioróżnorodności w przestrzeni w skali jednego pola lub większej (regionu, kontynentu), gdzie duże znaczenie mają zmieniające się warunki edaficzne, głównie pH [5, 51, 70]. Ta zmienność implikuje dokonywanie losowego poboru wielu próbek (5–10), na obszarze kilkudziesięciu metrów kwadratowych danego pola (terenu), które mogą być analizowane osobno lub razem po zmieszaniu. W tym ostatnim przypadku obserwuje się mniejszą zmienność warunków fizycznych gleby oraz strukturalnej różnorodności mikrobiomu [2]. Przy porównywaniu różnorodności mikroorganizmów, występujących w różnych miejscach, musi być uwzględniony stopień zmienności tej różnorodności w obrębie pojedynczego miejsca, a więc bardzo ważna jest liczba wykonanych powtórzeń [64]. Pojedyncze próbki, dające reprezentatywny obraz mikrobiomu glebowego, mają minimalną wielkość między 0,125 a 0,250 g [43, 65], jakkolwiek niekiedy lepiej jest zastosować większe 10-g próbki zamiast 1-g lub 0,1-g, aby zmniejszyć stopień zmienności struk-

tury społeczności mikroorganizmów obserwowany między próbkami [57]. Znalezienie złotego środka w tym przypadku opiera się na: rodzaju badanych populacji mikroorganizmów – dominujące lub rzadkie (tu lepiej zastosować większe próbki), a także samej glebie – efektywność ekstrakcji DNA zmienia się w zależności od typu gleby [51]. W mikroskali glebowej drobnoustroje charakteryzują się wybitnie nierównomiernym rozmieszczeniem między makro i mikroagregatami [54]. Wiele rzadkich populacji może być przeoczonych, jeśli badania nie będą prowadzone na poziomie tych agregatów, z uwagi na maskujący efekt dominujących populacji widoczny podczas analizy w makroskali tradycyjnych próbek glebowych [43].

Wysoce heterogeniczny charakter gleby, pod względem struktury oraz rozmieszczenia mikroorganizmów, stanowi olbrzymie utrudnienie dla efektywnej ekstrakcji DNA lub RNA metagenomu glebowego. Zagadnienie to zostało szerzej opisane we wcześniejszej pracy przeglądowej [45]. Najogólniej, kwasy nukleinowe ekstrahuje się z gleby, stosując podejście bezpośrednie lub pośrednie, kiedy DNA lub RNA izoluje się z wcześniej wyodrębnionej frakcji komórkowej drobnoustrojów glebowych. Otrzymany metagenomowy DNA oczywiście nie jest jednakowy pod względem ilościowym, i co ważniejsze, jakościowym (czystość preparatu, stopień pofragmentowania cząsteczki – a więc liczba i wielkość fragmentów DNA). Bezpośrednia ekstrakcja daje od 10 do 100 razy więcej DNA, który może być bardziej zanieczyszczony związkami humusowymi, zewnątrzkomórkowym DNA oraz silniej pofragmentowany na krótsze odcinki (0,5–20 kpz), niż w przypadku izolacji DNA z frakcji komórkowej (10–50 kpz, a nawet 150–1000 kpz). Ta procedura jest jednakże najczęściej stosowana w analizie metagenomu glebowego, chociaż rzadko można uzyskać więcej niż 60% ogólnej puli DNA mikroorganizmów obecnych w danej glebie, który jeszcze nawet w 30% może być tracony podczas etapów oczyszczania [51]. Niekiedy korzystniejsze bywa stosowanie procedury pośredniej ekstrakcji DNA, ale i w tym przypadku należy liczyć się z tym, że trudno jest uzyskać więcej niż 20–50% bakteryjnego DNA, z uwagi na silną adsorpcję komórek do cząstek glebowych [67]. Szereg specyficznych przeszkód charakterystycznych dla każdej ze strategii izolacji metagenomowego DNA powoduje, że uzyskiwany obraz różnorodności mikrobiomu glebowego jest mimo wszystko fragmentaryczny i zależy w dużej mierze od rodzaju mikroorganizmów i ich podatności na lizę, rozmieszczenia oraz stopnia dostępności komórek i uwolnionego DNA w strukturze gleby. W takiej sytuacji, pewnym rozwiązaniem, umożliwiającym uzyskanie w miarę wiarygodnego obrazu bioróżnorodności, może być przeprowadzenie sukcesywnego kilkakrotnego ekstrahowania DNA lub izolowania komórek z tej samej próbki glebowej [23].

4. Analiza genetyczna mikrobiomu

Różnorodność mikroorganizmów w glebie może być rozpatrywana pod kątem ogólnej oceny struktury mikrobiomu i porównania między różnymi próbkami, lub też może dotyczyć dogłębnej identyfikacji składu społeczności drobnoustrojów zasiedlającej daną glebę. W związku z tym należy zastosować różne strategie do sporządzania próbek oraz ekstrakcji DNA, jak wyżej opisano. Co więcej, odmienny charakter będzie też mieć analiza metagenomiczna w obu przypadkach. Jeśli celem jest poznanie ogólnej struktury społeczności mikroorganizmów i ewentualnych zachodzących w niej zmian, wystarczające jest wykorzystanie znanych metod „fingerprintowych”, jak np. DGGE/TGGE czy T-RFLP [46]. Ponieważ większość z tych metod opiera się na wcześniejszej amplifikacji filogenetycznych markerów przez PCR, należy być świadomym pewnych problemów specyficznych dla reakcji PCR. Przykładem istotnego problemu jest preferencyjne amplifikowanie fragmentów genowych, pochodzących od populacji będących w mniejszości, jeśli takie fragmenty charakteryzują się ponad przeciętną efektywnością przyłączania starterów i polimerazy DNA [91]. W konsekwencji otrzymane profile DNA nie będą rzetelnie odzwierciedlać składu analizowanego mikrobiomu. Charakterystyczną cechą metod „fingerprintowych” jest także to, że przedstawiony przez nie obraz mikrobiomu obejmuje ograniczoną liczbę dominujących populacji. Te populacje mogą być identyfikowane poprzez klonowanie, a następnie sekwencjonowanie poszczególnych fragmentów DNA pochodzących od różnych mikroorganizmów. Nowa generacja metod „fingerprintowych” umożliwiła identyfikację bakterii i archeonów glebowych w oparciu o mikromacierze genowe, które zawierają od 500 000 do 110 000 025-merowych sond oligonukleotydowych hybrydujących z metagenomowymi fragmentami genów SSU rRNA po ich amplifikacji PCR, lub bezpośrednio z rRNA, albo dwuniciowym komplementarnym DNA (dscDNA). Daje to możliwość wykrycia od około 9 000 (PhyloChip G2) do ponad 50 000 (PhyloChip G3) OTU (operational taxonomic unit) drobnoustrojów potencjalnie obecnych w metagenomie glebowym [1, 11, 17, 74].

Szczegółowe poznanie składu mikrobiomu glebowego wymaga zastosowania innego podejścia analitycznego w stosunku do wyekstrahowanego z gleby metagenomowego DNA. Jednym ze sposobów, jednocześnie najczęściej stosowanym, jest wykorzystanie filogenetycznego markera w postaci genu SSU rRNA (16S u prokariotów lub 18S u eukariotów), który po amplifikacji PCR podlega klonowaniu celem otrzymania biblioteki reprezentującej wszystkie mikroorganizmy obecne w badanej glebie. Następnie poszczególne sklonowane fragmenty DNA są sekwencjonowane i poddawane bioinformatycznej analizie. Drugim sposobem, dzięki postępowi technologicznemu, jest analizowanie,

na podobnej zasadzie, losowych fragmentów DNA lub całych metagenomów. W tym przypadku pofragmentowany materiał genetyczny jest swoistą mieszaniną fragmentów DNA, w której szereg sekwencji nukleotydowych częściowo się pokrywa. W związku z tym, po sekwencjonowaniu wystarczającej liczby małych odcinków DNA istnieje możliwość odtworzenia dużych fragmentów DNA, wystarczających do identyfikacji taksonów, lub będących wręcz rekonstrukcją kompletnego genomu [99]. Oczywiście jest, że w tej procedurze będą wykrywane, w przeważającej mierze, najliczniej reprezentowane organizmy, dlatego tak ważne jest posługiwanie się większymi próbkami gleby w powtórzeniach, co jest często trudne do spełnienia, aby wychwycić także sekwencje DNA pochodzące od rzadkich taksonów. Pewną pomocą jest przypadkowa natura sekwencjonowania typu „shotgun”, która zapewnia, że wiele tych organizmów jest reprezentowanych przynajmniej przez jakieś małe fragmenty sekwencji nukleotydowych.

Niezależnie od strategii zastosowanej do analizy metagenomu, w klasycznej postaci niezbędny etap klonowania może dotyczyć małych lub dużych fragmentów DNA, które też są wstawiane w różne wektory. Biblioteka małych wstawek dotyczy fragmentów DNA o wielkości do 20 kbp wprowadzonych do plazmidów, natomiast duże fragmenty o wielkości do 40 kbp wstawiane są do fosmidów lub kosmidów, a te o wielkości jeszcze większej, przekraczającej nawet 100 kbp, znajdują się w sztucznym chromosomie bakteryjnym BAC. Wspomniane wektory posiadają zalety i wady oraz różnią się przeznaczeniem (Tab. I), a ich użycie zależy od: jakości izolowanego glebowego DNA, wymagań odnośnie przeciętnej wielkości wstawek w bibliotece, wymagań w stosunku do liczby kopii wektora w komórce gospodarza, rodzaju samego gospodarza, oraz strategii użytej do przeszukiwania klonów [14]. Niemal wszystkie biblioteki klonów w początkowym etapie tworzenia, a także w celu przechowywania sekwencji DNA metagenomu, wykorzystują jako gospodarza bakterię *Escherichia coli*. W przypadku klonowania fragmentów DNA mikroorganizmów glebowych, zwłaszcza w celu późniejszej analizy funkcjonalnej genów, jako gospodarzy wybiera się występujące w glebie bakterie z rodzaju *Pseudomonas* lub *Streptomyces* [12, 92]. Komórki *E. coli* często nie mają odpowiedniej maszyny transkrypcyjno-translacyjnej (np. efektywne promotory, czynniki sigma czy białka regulatorowe), która umożliwiłaby ekspresję genów lub operonów pochodzących od członków mikrobiomu glebowego często różniących się znacznie zawartością zasad G + C w genomie oraz systemem wydzielniczym białek, co skutkuje niską ilością (niekiedy mniej niż 0,01%) uzyskanych pozytywnych klonów podczas jednej rundy przeszukiwania biblioteki [13, 30]. Konstruowane biblioteki metagenomu glebowego są z reguły bardzo obszernymi kolekcjami fragmentów

Tabela I
Porównanie użyteczności wektorów niosących różne wstawki DNA w konstrukcji bibliotek metagenomowych

Małe wstawki DNA (plazmidy)	Duże wstawki DNA (fosmidy, kosmidy, BAC)
<p>Zalety</p> <ul style="list-style-type: none"> • Duża liczba kopii umożliwiająca wykrywanie genów o słabej ekspresji • Klonowanie pojedynczych genów lub małych operonów (np. geny enzymów lub antybiotykooporności) • Możliwa jest ekspresja genów spod promotora wektora • Możliwe jest klonowanie pofragmentowanego lub zanieczyszczonego DNA • Technicznie proste 	<ul style="list-style-type: none"> • Duże fragmenty DNA (nawet > 100 kz) mogą zawierać całe operony dla funkcji metabolicznych (np. różnych biosyntezy, enzymów katabolicznych) • Mniejsza liczba klonów konieczna do przeszukiwania w celu wykrycia nosicieli genu(ów) • Modyfikowane BAC dają możliwość przełączania między zachowaniem w pojedynczej i wielu kopiach co przynosi korzyści dla ilości otrzymywanego DNA i poziomu ekspresji genów • Odpowiednie dla częściowej charakterystyki genomowej niehodowanych mikroorganizmów glebowych
<p>Wady</p> <ul style="list-style-type: none"> • Małe fragmenty DNA (< 20 kz) niosą ograniczoną informację genetyczną • Możliwa eliminacja części informacji filogenetycznej z uwagi na ekspresję toksycznych dla gospodarza genów • Konieczność selekcji właściwej procedury ekstrakcji DNA oraz klonowania dla uzyskania dużego udziału DNA pożądanym mikroorganizmów • Konieczność regulacji poziomu ekspresji klonowanych genów przez dobór wektorów z indukowalną kontrolą ekspresji genu wstawki lub liczby kopii plazmidu • Olbrzymia liczba klonów konieczna do przeszukiwania w celu znalezienia nosicieli genu • Niewłaściwe do badania aktywności oraz szlaków biochemicznych kodowanych przez większe ugrupowania genowe 	<ul style="list-style-type: none"> • Małe fragmenty DNA metagenomu są tracone, a więc pewne taksony mogą być zbyt rzadko reprezentowane • Mała liczba kopii uniemożliwia wykrywanie genów o słabej ekspresji • Ograniczona ekspresja genów spod promotora wektora • Możliwa eliminacja części informacji filogenetycznej z uwagi na ekspresję toksycznych dla gospodarza genów, stąd konieczność zachowania pojedynczych kopii w bibliotece do momentu skryningu funkcji • Niezbędny wysoko-molekularny DNA o znacznej czystości do konstrukcji biblioteki • Technicznie trudne

DNA. Oznaczając liczbę różnych genomów prokariotycznych obecnych w gramie gleby obliczono, że niezbędne jest w tym celu posłużenie się ponad 10^7 plazmidowych klonów, zawierających 15 kpz wstawki DNA, albo 10^6 klonów BAC, niosących wstawki o wielkości 100 kpz [33]. Te obliczenia są oparte na założeniu, że wszystkie gatunki są równocześnie reprezentowane, co w warunkach naturalnej gleby jest niespotykane. Co więcej, ażeby uzyskać rzeczywistą reprezentację genomów, uwzględniając także rzadko występujące (poniżej 1%), należałoby skonstruować biblioteki zawierające ponad 10 000 Gpz glebowego DNA, co odpowiadałoby 10^{11} klonów BAC [66]. Zakładając poprawność tych obliczeń, genetyczna zawartość opublikowanych do tej pory bibliotek glebowych (często 100–200 Mpz, maksymalnie kilka Gpz, dla danej gleby) jest nikła, jeśli chodzi o pokrycie całego metagenomu glebowego [32, 76]. Ponadto, biblioteka BAC genów 16S rRNA, zawartych we fragmentach otrzymanych po trawieniu enzymem restrykcyjnym (np. *Hind*III) wysokomolekularnego DNA glebowego, może reprezentować odmienny skład populacji bakterii (mniejszy udział *Bacillus*, *Alphaproteobacteria*, grupy *Cytophaga-Flavobacterium-Bacteroides*, *Gemmimonas*; brak *Chloroflexi* oraz Grupy Termitowej), aniżeli biblioteka klonów genów 16S rRNA otrzymanych z wyekstrahowanego DNA i poddanego bezpośredniej amplifikacji PCR (mniejszy udział *Betaproteobacteria*, *Gammaproteobacteria*, *Acidobacterium*, WS3; brak

OP10, *Planctomycetes* oraz *Actinomycetes*). Ten przykład z badań L i l e s i wsp. [50] jest wymownym świadectwem wpływu sposobu konstruowania biblioteki metagenomu na końcowy obraz strukturalnej różnorodności mikrobiomu w analizowanej glebie. W sumie, wiedza jaką dysponujemy obecnie obejmuje populacje dominujących mikroorganizmów. Fakt ten nie powinien dziwić, jeśli uświadomimy sobie ograniczenia metodyczne związane z ekstrakcją DNA z gleby, jego klonowaniem, a także minimalnym pokryciem metagenomu glebowego (przy wymaganym ponad 100, a nawet 1000 krotnie większym, aby uwzględnić rzadkie mikroorganizmy) zapewnionym przez konstruowane biblioteki. Dodatkowo należy pamiętać, że rzadkie mikroorganizmy nie będą reprezentowane w bibliotece bez zastosowania wcześniejszej procedury ich wzbogacania w glebie [42].

Zarówno małe, jak i duże wstawki DNA metagenomowego, zgromadzone w odpowiednich bibliotekach, od początku były sekwencjonowane, podobnie jak w przypadku analizy pojedynczych genomów, w oparciu o metodę Sangera opierającą się na włączaniu dideoksynukleotydów i syntezie komplementarnej nici DNA do jednoniciowej matrycy. W analizie sekwencyjnej genomów niezbędne było zastosowanie znaczników fluorescencyjnych dla nukleotydów i zautomatyzowanie procesu z wykorzystaniem 96 kanałowych kapilarnych sekwenatorów [4]. Dla przykładu, oparty na metodzie Sangera system 3730 xl firmy Applied Biosystems

umożliwia sekwencjonowanie w ciągu dnia od ponad 1000 do ponad 3800 próbek DNA o wielkości od 500 do ponad 800 pz, dając łączny odczyt sekwencji DNA od 1 do 2 Mz. Dalszy postęp przyniosło wprowadzenie do badań w 1996 roku sekwencjonowania nowej generacji, które jest ponad 50-krotnie wydajniejsze i zamiast metody dideoksy wykorzystuje technikę niezależną od elektroforezy, ani innego sposobu rozdzielania fragmentów DNA, a opartą na pirosekwencjonowaniu. W tej nowej metodzie komplementarna do matrycy nic DNA jest syntetyzowana bez udziału dideoksynukleotydów, a wstawianym przez polimerazę DNA nukleotydowi towarzyszy uwolnienie cząsteczki pirofosforanu i, w obecności sulfurylasy ATP oraz lucyferazy, dochodzi do emisji światła. Niewłączone nukleotydy i nadmiarowe ATP są usuwane przez apirazę, a sekwencja nukleotydów jest rejestrowana w postaci pirogramu, w którym poszczególne piki o różnej wielkości odpowiadają odpowiedniej ilości rejestrowanych sygnałów światła proporcjonalnych do liczby wstawianych w danym momencie nukleotydów. Ogromna ilość uzyskanej informacji, w postaci ponad 2 500 000 sekwencji DNA (powyżej 350 pz), przekracza łącznie miliard zasad na dzień, przy zdecydowanie mniejszych kosztach w porównaniu z sekwencjonowaniem kapilarnym [56].

Sekwencjonowanie nowej generacji opiera się obecnie na trzech konkurencyjnych systemach: 454 Life Sciences, Solexa/Illumina oraz SOLiD [56], które różnią się wielkością i ilością fragmentów DNA uzyskanych po sekwencjonowaniu, wydajnością syntezy na cykl oraz dzień (Tab. II). Ta odmiana sekwencjonowania nie wymaga wcześniejszego klonowania fragmentów DNA w wektorach, które też nie dla każdego z nich musi być skuteczne. Najnowsza wersja platformy Solexa/Illumina, oparta na sekwenatorze krótkich sekwencji SOLiD4, generuje DNA składający się z 50-z fragmentów o łącznej długości przeszło 100 Gpz, w ciągu jednego cyklu. Niemniej do analizy metagenomicznej najwygodniejsze jest stosowanie 454-pirosekwencjonowania, zwłaszcza nowszej platformy GS Titanium, która zapewnia syn-

tezę większych fragmentów DNA (400–700 pz) i ponad pięciokrotnie większą wydajność (400–600 Mpz), w stosunku do pierwszej wersji GS20. Dłuższe sekwencje ułatwiają dalszą analizę bioinformatyczną, ponieważ zapewniają poprawne poszukiwanie homologii za pomocą BLAST w bankach sekwencji DNA i składanie długich kontigów. W konsekwencji ułatwia to filogenetyczne powiązanie genów z ich macierzystymi organizmami oraz poprawia identyfikowanie gatunków z dostępnych fragmentów pochodzących z genu 16S rRNA [98].

Sekwencjonowanie metagenomu może dotyczyć całkowitego DNA wyizolowanego z gleby, i jest przeprowadzane według strategii typu „shotgun”, albo też jest przeprowadzane na frakcji tego DNA wzbogaconej przez PCR [7]. Pierwsza strategia prowadzi do uzyskania maksymalnej informacji o sekwencjach nukleotydowych członków społeczności mikroorganizmów i wszystkich funkcjonalnych genach obecnych w mikrobiomie. Ponieważ mikrobom niesie olbrzymią ilość informacji genetycznej, dostęp do niej wiąże się z koniecznością przeprowadzenia bardzo wielu sekwencjonowań, co jest kosztowne i wymaga znacznego zaangażowania mocy obliczeniowych oraz zaawansowanych metod bioinformatycznych. Jednakże nawet wtedy należy liczyć się z możliwością przeoczenia populacji rzadko reprezentowanych w społeczności mikroorganizmów. Wzbogacanie tylko części DNA przez PCR, najczęściej w postaci genów SSU rRNA, ale także innych konserwowanych ewolucyjnie (np. *recA*), wymaga: użycia mniejszych próbek źródłowych (wystarczy nawet ng DNA), mniejszej liczby sekwencjonowań w celu wykrycia również gatunków rzadkich oraz zastosowania prostszej analizy bioinformatycznej. Chociaż 454-pirosekwencjonowanie generuje długie sekwencje DNA, to jednak są one wyraźnie krótsze, niż np. gen 16S rRNA (1542 pz), co mogłoby budzić zastrzeżenia w kwestii przydatności tej metody do analizy filogenetycznej. Jednakże okazało się, że amplifikacja fragmentów obejmujących region V6 oraz V7 (V6 + V7) genu 16S rRNA przy użyciu uniwersalnych starterów pozwala na oszacowanie porównywalnego

Tabela II

Porównanie wydajności sekwencjonowania techniką elektroforezy kapilarnej oraz trzech odmian sekwencjonowania nowej generacji

System	Wydajność (milion pz/dzień)	Wydajność (milion pz/cykl)	Wydajność (sekwencje/dzień)	Sekwencje (długość, pz)
Dideoksy 3730xl	1 – 2	0,8	8000	> 800
454 Titanium ¹	1000	450 – 700	1–2,5 × 10 ⁶	600 – 1000
SOLiD ²	7000 – 20000	90000 – 300000	> 1,3 × 10 ⁸	50 – 75
Solexa/Illumina ³	55000 – 106000	120000 – 600000	5,5 – 10 × 10 ⁸	2 × 100 – 2 × 150

¹ Podane wartości dotyczą systemu GS FLX XL+ oraz GS FLX XLR70, których cykl pracy wynosi odpowiednio 23 lub 10 godz.;

² podane wartości dotyczą systemu 5500 oraz 5500xl;

³ dostępne są cztery główne odmiany sekwenatorów. Podane wartości dotyczą typu HiSeq2500 pracującego w wariantcie szybkim (cykl ok. 27 godz.) lub wysokiej wydajności (cykl ok. 11 dni)

bogactwa gatunków drobnoustrojów, jak w przypadku stosowania kompletnych sekwencji genu [100]. Również inne regiony zmienne genu 16S rRNA mogą być pomyślnie wykorzystywane do analizy metagenomicznej, a generowane podczas pirosekwencjonowania fragmenty genu określa się mianem pirotagów [18, 75]. Ponadto, pirotagi pochodzące z różnych próbek (np. gleb) mogą być równocześnie syntetyzowane i następnie rozpoznawane dzięki specyficznej konstrukcji starterów, które między odcinkami adaptorów A i B, niezbędnych dla procesu pirosekwencjonowania, a właściwymi regionami starterów, zawierają unikalne 5-z lub dłuższe odcinki kodowe charakterystyczne dla każdej próbki [7]. W ten sposób, nawet ponad 100 różnych próbek DNA amplifikowanych przez PCR może być sekwencjonowane podczas jednego cyklu w sekwenatorze 454 Titanium [31]. W większości badań kolekcja około 25 000–50 000 pirotagów jest wystarczająca dla adekwatnego zobrazowania strukturalnej różnorodności mikrobiomu [76].

Pirosekwencjonowanie metagenomu glebowego może generować setki tysięcy krótkich (300–800 z), jeśli porównać z metodą dideoksy, sekwencji DNA lub pirotagów podczas jednego cyklu, a ich końcowa liczba zależy od stopnia złożoności metagenomu. Taka olbrzymia ilość sekwencji nukleotydowych daje możliwość wychwycić także rzadkie populacje mikroorganizmów, przy założeniu, że każdy fragment DNA (pirotag) identyfikuje odpowiedniego członka mikrobiomu. Analiza złożonego metagenomu glebowego, w oparciu o tradycyjne sekwencjonowanie (kapilarna elektroforeza, CAE) ograniczonej liczby klonów, nie daje tej możliwości i skutkuje rozpoznaniem jedynie dominujących populacji, reprezentowanych przez odpowiednie filotypy wyróżnione na bazie ponad 1000 pełnych sekwencji genu 16S rRNA [20, 76]. Jednakże ta przewaga pirosekwencjonowania może być okupiona niebezpieczeństwem przeszacowania liczby OTU w postaci stwierdzenia ich nadmiernej lub zbyt małej liczby. Jak się okazało, pewne regiony zmienne (V1 + V2, V6) genu 16S rRNA dają podczas analizy przesadnie dużo OTU, inne wysoce zmienne regiony (V3, V7, V7 + V8) generują zbyt małą liczbę OTU, natomiast regiony V4, V5 + V6 oraz V6 + V7 są źródłem pirotagów pozwalających oszacować podobną liczbę OTU, jak w przypadku sekwencjonowania prawie kompletnego genu [100]. Aby uniknąć nadmiaru OTU, zwłaszcza wśród rzadkich populacji, przyjmuje się też 0,2% dla błędu pirosekwencjonowania oraz 97% podobieństwo sekwencji DNA, jako progowe wartości dla definiowania wiarygodnych OTU (filotypów) do gatunków i wyznaczania ich bogactwa w danym mikrobiomie [47]. Liczba stwierdzanych OTU rośnie wraz ze wzrostem złożoności metagenomu i ilością sekwencjonowanych fragmentów DNA, a zależność ta jest obrazowana w postaci krzywej rozgęszczenia filotypów, którą można wyznaczać przy poziomach

dystansu genetycznego 3%, 5% lub 20%, charakterystycznych odpowiednio dla gatunku, rodzaju lub typu mikroorganizmu. W bogatej w bakterie glebie stan nasycenia krzywej daje się zaobserwować tylko na poziomie typu, natomiast w przypadku niższych taksonów, nawet znaczna liczba sekwencji DNA otrzymanych przez pirosekwencjonowanie nie zapewnia pokrycia wszystkich organizmów zasiedlających glebę [55].

Oznaczanie liczby filotypów, wchodzących w skład mikrobiomu, i ich identyfikowanie na podstawie analizy sekwencji nukleotydowych sklonowanych w wektorach, wymaga stosunkowo prostej procedury bioinformatycznej. Bogate bazy sekwencji nukleotydowych genu 16S rRNA mogą być przeszukiwane za pomocą programu BLAST pod kątem największego podobieństwa w stosunku do badanych próbek. Olbrzymia ilość sekwencji DNA generowanych podczas pirosekwencjonowania wymaga o wiele większej obróbki surowych danych, wśród których sekwencje (pirotagi) o właściwej długości, stosownej do analizy filogenetycznej, muszą być wyselekcjonowane i pogrupowane przez tworzenie kontigów lub klastrów podobnych fragmentów DNA, celem redukcji liczby danych. Z reguły im dłuższe są sekwencje DNA, tym niosą bardziej przydatną informację do filogenetycznego lub taksonomicznego dopasowania przez porównanie ze znanymi sekwencjami DNA, choć platforma MG-RAST zadowala się nawet 75-z fragmentami [81]. Należy pamiętać, że liczba pirotagów powstających podczas jednego cyklu pracy sekwenatora stanowi poważne wyzwanie dla mocy obliczeniowych związanych z dopasowaniem sekwencji DNA i działaniem algorytmów konstrukcji drzew pokrewieństw. Znaczący procent sekwencji DNA z metagenomu glebowego nie znajduje istotnych homologów BLAST w dostępnych bibliotekach [62]. O możliwych rozmiarach tego procenta niech świadczy fakt, że nawet w przypadku kompletnie poznanych genomów czystych kultur bakteryjnych od 35 do 45% przewidzianych otwartych ramek odczytu (ORF) nie ma istotnych homologów w bazie sekwencji nukleotydowych GenBank [73]. Nie ma programu, który byłby w stanie obliczyć dopasowanie dla setek tysięcy pirotagów, co wymagałoby stworzenia matrycy odległości o pojemności przekraczającej terabajt. Stosując narzędzia relacyjnej bazy danych, można próbować sobie poradzić z tym problemem, ale i wtedy można dla przykładu dopasować maksymalnie 37 000 unikalnych sekwencji DNA spośród całości wynoszącej 725 000 pirotagów [76].

5. Analiza funkcjonalna metagenomu

Podczas analizy metagenomicznej, nie tylko chodzi o poznanie różnorodności mikroorganizmów tworzących mikrobiom glebowy, ale równie ważne jest poszu-

kiwanie funkcjonalnych genów, a więc dowodów na występowanie określonych aktywności metabolicznych wśród członków mikrobiomu. Biblioteki klonów, zawierające na wektorach małe bądź duże wstawki metagenomowego DNA, są przeszukiwane pod kątem ekspresji genu określonej nowej cechy fenotypowej w bakteryjnym gospodarzu (np. *E. coli*), który nie ma tej cechy zakodowanej w swoim genomie. Wymaga to zastosowania dokładnego przeszukiwania wszystkich klonów wyrostłych na podłożu różnicującym, i często można wykryć obecność badanej cechy u zaledwie kilku klonów na setki tysięcy czy kilka milionów analizowanych [42]. Sprawa jest łatwiejsza, jeśli poszukiwana cecha może być jednocześnie markerem selekcyjnym (np. oporność na antybiotyki czy metal ciężki), co pozwala otrzymać na podłożu selekcyjnym wzrost klonów niosących stosowne geny warunkujące metalooporność [53, 61]. Inny sposób polega na wykorzystaniu, jako gospodarzy, szczepów lub ich mutantów, które wymagają heterologicznej komplementacji ze strony genu pochodzącego z metagenomu, aby rosnąć na podłożu selekcyjnym. W tym przypadku obserwuje się jedynie wzrost zrekombinowanych klonów, które posiadają stosowny gen i wytwarzają aktywny produkt tego genu, np. dioksygenazę naftalenową czy racemazę lizyny [9, 58]. Interesującą metodą przeszukiwania

biblioteki klonów jest wzbudzanie ekspresji genu z metagenomu w klonach, które wyrastają na podłożu wzbożonym w stosowny substrat dla enzymu kodowanego przez badany gen (operon). W ten sposób są wykrywane przykładowo geny: esterazy lub lipazy przez powstawanie przezroczystego halo wokół kolonii na podłożu wskaźnikowym, dioksygenazy ekstradiolowej przez wytwarzanie żółtego produktu oraz geny związane z rozkładem 4-hydroksymaślanu na podłożu z dodatkiem chlorku tetrazolowego [21, 32, 36, 37, 80].

Wspomniane metody przeszukiwania biblioteki klonalnej metagenomu są potężnym narzędziem służącym identyfikowaniu genów kodujących naturalne produkty lub odpowiedzialnych za określoną aktywność metaboliczną wśród wszystkich członków mikrobiomu, a nie tylko tych, którzy poddają się hodowli (Tab. III). Jednakże również tutaj pojawia się szereg ograniczeń, które są trudne do pokonania. Wykrywanie rekombinowanych klonów, w których wprowadzony gen (operon) ulega ekspresji, wybitnie zależy od jego skutecznej transkrypcji, translacji, składania białka oraz efektywnego wydzielania produktu przez komórkę bakteryjnego gospodarza. Dlatego coraz częściej *E. coli* jest zastępowana przez bakterie, które występują w glebie (np. *Pseudomonas*, *Streptomyces*, *Sinorhizobium*, *Rhizobium*,

Tabela III

Przykłady analizy funkcjonalnej bibliotek klonalnych metagenomów różnych gleb [14, 82, 89]

Gleba	Wektor	Klony	Wstawka (kz)	Metagenom (Gz)	Geny
Nieuprawna	BAC	12000	37	0,42	substancji antybakteryjnych (indyrubina i pokrewne)
Alkaliczna lessowa	plazmid	100000	8–12	1,0	proteaz
Glina piaszczysto-ilasta	fosmid	100000	30–40	3,5	syntazy poliketydowej
Leśna	fosmid	33700	35	1,18	lipaz
Glina pylasta	plazmid	200000 58000 250000 650000	4,1 2,7 3,5 3,5	4,2	antybiotykooporności ¹ (Km ^r , Tc ^r , Am ^r , Bs ^r , Tm ^r , Sm ^r)
Łąkowa Pole buraków cukrowych	plazmid	1500000	5–8	7,8	rozkładu 4-hydroksymaślanu lipaz antyporterowe Na ⁺ /H ⁺
Pole uprawne	plazmid	80000	5,2	0,42	amidaz
Łąka Pole buraków cukrowych Pole zbóż	plazmid	583000 360000 324000	4,4 3,8 3,5	4,05	produkcji karbonylu
Przy strumieniu pokryta mchem	plazmid	30000	3,5	0,11	amylaz
Leśna	fosmid	113700	> 30	3,4	syntaz poliketydowych
Trawiasta	fosmid	15000	35–45	1,0	biosyntezy leinamycyny, mupirocyny i pimarycyny syntazy poliketydowej
Ilasta Zwirowo-ilasta sucha	fagemid ²	55000 495000	6,5–7	3,5	antybiotykooporności ³ (Ap ^r , Gm ^r , Cm ^r , Tmp ^r)

¹ Nowe geny oporności na kanamycynę, tetracyklinę, apramycynę, butyrozynę, tobramycynę oraz sisomycynę;² Konstrukt oparty na fagu lambda;³ Nowe geny oporności na ampicylinę, gentamycynę, chloramfenikol oraz trimetoprim

Burkholderia, *Bacillus*, *Sphingomonas*), a więc dysponują stosowną maszyną regulacyjną dla ekspresji genów wniesionych na wektorze [21]. Równocześnie dobierane są odpowiednie wektory o szerokim spektrum gospodarzy lub wektory wahadłowe (np. BAC), które zawierają więcej niż jedno miejsce startu replikacji lub mogą integrować się do chromosomu. Wektory wahadłowe umożliwiają na przykład przejęcie wstawki DNA metagenomowego z biblioteki wyjściowej skonstruowanej na *E. coli* przy użyciu stabilnego wektora fosmidowego i wprowadzenie jej, w celu ekspresji, do nowego gospodarza, np. *Pseudomonas* [8].

Wykrywanie funkcjonalnych genów, w tym także nowych, w metagenomie glebowym odbywa się również poprzez sekwencjonowanie zawartości biblioteki klonów lub bezpośrednio przez pirosekwencjonowanie fragmentów wyekstrahowanego DNA mikrobiomu. Podczas sekwencjonowania typu „shotgun” interesujących genów, występujące w regionach flankujących markery filogenetyczne (np. gen 16S rRNA) pozwalają na identyfikację taksonomiczną aktywnych metabolicznie mikroorganizmów [32]. Funkcjonalną analizę można także przeprowadzić, wykorzystując mikromacierze zawierające geny zaangażowane w kluczowe procesy biogeochemiczne. Przykładem jest obszerna mikromacierz GeoChip oparta na 28 000 50-merowych oligonukleotydach obejmujących 57 000 sekwencji DNA z 292 kategorii genów, które odpowiadają za przemiany C, N, S, P, metalooporność, redukcję metali lub degradację ksenobiotyków [34]. W ten sposób można monitorować obecność lub ekspresję tysięcy genów, uwzględniając zarówno aspekty jakościowe, jak i ilościowe, w jednym eksperymencie. W dodatku ta technika pozwala dokonywać analizy porównawczej wielu różnych próbek równocześnie. Jednakże efektywność tej techniki w dużej mierze zależy od doboru i konstrukcji stosownych sond genowych, jakości i ilości izolowanego metagenomowego DNA oraz właściwego przeprowadzenia procesu PCR. Ponadto, należy pamiętać, że metoda mikromacierzy DNA bazuje na sondach opracowanych w oparciu o znane sekwencje nukleotydowe genów i dlatego może przeoczyć nowe geny obecne w metagenomie, a których brak w mikromacierzy. Stanowi to poważny problem, ponieważ metoda ta nie wykrywa wielu funkcji obecnych wśród populacji bakterii glebowych, które nie zostały opisane, a mogą prawdopodobnie reprezentować znaczącą frakcję mikrobiomu glebowego [13, 19].

6. Strukturalna różnorodność mikroorganizmów w różnych środowiskach glebowych

Podejście metagenomowe znacząco zmieniło nasze wyobrażenie o strukturalnej różnorodności mikrobiomu glebowego oraz aktywności metabolicznej, i w kon-

sekwencji roli ekologicznej poszczególnych członków tej społeczności. Nasze możliwości poznawcze uległy poszerzeniu o świat mikroorganizmów stanowiących do niedawna milczącą większość. Pierwszą publikacją, przedstawiającą klonalną bibliotekę genu 16S rRNA drobnoustrojów glebowych, była praca Liesack i Stackebrandt z 1992 roku [49]. Wkrótce potem takie biblioteki zaczęto otrzymywać w odniesieniu do szerokiego zakresu różnych gleb, a taki sposób identyfikacji mikroorganizmów, zwłaszcza wraz z rozwojem metod „fingerprintowych”, na trwale zadomowił się w mikrobiologii środowiskowej. Pierwszy metagenom glebowy o wielkości przekraczającej 1 Gpb, w postaci biblioteki klonalnej konstruowanej na bazie BAC, wykazał obecność w glebie bakterii z taksonów *Acidobacterium*, *Cytophagales* oraz *Proteobacteria* [69]. Na podstawie meta-analizy dostępnych informacji z klonalnych bibliotek wynika, że w glebie może występować 32 typy bakterii, spośród których 9 dominuje [39]. Do tych ostatnich należą: *Proteobacteria* (39%, zakres 10–77%), *Acidobacteria* (20%, zakres 5–46%), *Actinobacteria* (13%, zakres 0–34%), *Verrucomicrobia* (7%, zakres 0–21%), *Bacteroidetes* (5%, zakres 0–18%), *Chloroflexi* (3%, zakres 0–16%), *Planctomycetes* (2,5%, zakres 0–8%) oraz *Gemmatimonadetes* (2%, zakres 0–4%). Zwracająca uwagę jest mała reprezentacja bakterii z typu *Firmicutes* (2%, zakres 0–8%) zawierających takie typowe bakterie glebowe, jak *Bacillus* oraz *Clostridium*, co może mieć jednak związek z trudnością ekstrakcji DNA z tych Gram-dodatnich i przetrwalnikujących bakterii.

Przeciętne biblioteki metagenomu glebowego są małe i obejmują między 300 a 500 klonów, wyjątkiem są biblioteki sporządzone w odniesieniu do gleb z Minnesoty oraz Alaski, które zawierają odpowiednio 1700 oraz 1033 klonów [71, 86]. Efektem tego jest niepełne pokrycie mikrobiomu glebowego i ograniczenie strukturalnej różnorodności do dominujących populacji. Pozostała reszta mikroorganizmów, występująca w małych liczebnie populacjach, skupia jednakże w sobie większość genetycznej różnorodności mikrobiomu i stanowi „rozrzedzoną biosferę” [75]. By móc dotrzeć do tych drobnoustrojów, Elshahed i wsp. [22] opracowali bibliotekę zawierającą 13 001 klonów prawie kompletnej sekwencji genu 16S rRNA bakterii występujących w glebie periowej. Analiza filogenetyczna umożliwiła im wyróżnienie 34 typów bakterii, wśród których 15 ma znanych przedstawicieli hodowlanych, 14 jest określanych jako typy o statusie kandydatów, natomiast 5 (KFS1-KFS5) stanowi całkiem nowe typy. Wśród najliczniej reprezentowanych w bibliotece są typy: *Actinobacteria* (24%), *Acidobacteria* (20%), *Alphaproteobacteria* (10%), *Deltaproteobacteria* (8%), *Chloroflexi* (7%), *Verrucomicrobia* (6%), *Bacteroidetes* (5%), *Planctomycetes* (4%), *Gemmatimonadetes* (3%). Mały udział, poniżej 1%, ma aż 24 typów bakterii (np. TM7, OP11,

WS3, OD2, *Cyanobacteria*, *Chlorobi*), a 14 z nich jest reprezentowanych tylko przez kilka klonów (np. *Fibrobacter*, *Chlamydia*).

Gleba jest przykładem środowiska, w którym występuje znacząca liczba rzadkich gatunków bakterii, które mogą stanowić od 18 do 37% wszystkich OTU_{0,03} reprezentowanych w bibliotece klonów [22]. Najwięcej rzadkich gatunków należy do typu *Planctomycetes* (34–78%), a najmniej do *Acidobacteria* (12–25%). Ponadto, wśród tych rzadkich gatunków można wyróżnić dwie grupy, z których pierwsza, reprezentowana przez 84–92% członków, wykazuje bliskie pokrewieństwo z przedstawicielami licznie zasiedlających glebę typów bakterii, natomiast druga grupa, w ilości 8–16%, obejmuje gatunki wybitnie unikalne filogenetycznie, lub należące do typów znanych z innych środowisk (np. *Chlorobi*, *Caldithrix*, BRC-1). O roli ekologicznej tych rzadkich mikroorganizmów, zwłaszcza należących do drugiej grupy, prawie nic nie wiadomo. Można przypuszczać, że przejmują one określone funkcje biogeochemiczne w sytuacji drastycznych zakłóceń fizykochemicznych zachodzących w glebie (np. podczas suszy, zanieczyszczenia ksenobiotykami), kiedy ich populacje mogą stanowić istotny składnik zmienionej struktury mikrobiomu [22]. Można też sądzić, że szereg unikalnych gatunków wśród tej „rozrzedzonej biosfery” spełnia specyficzne, nieznanne jeszcze funkcje w ekosystemie glebowym, lub też jest pozostałością ewolucji mikroorganizmów – niemniej stanowiącą istotny bank informacji genetycznej. Te swego rodzaju relikty ewolucji, jakkolwiek współcześnie przytłoczone konkurencją taksonów dominujących, mogą charakteryzować się wyjątkową zdolnością przeżycia oraz unikania procesu wymierania, stanowiąc jakby pamięć genetyczną ekosystemu wykorzystywaną do odtworzenia jego typowej struktury i funkcji po katastrofie ekologicznej.

Zaletą analizy metagenomicznej jest możliwość dokonania porównania różnorodności bakterii, archeonów, grzybów oraz wirusów występujących w skrajnie różnych środowiskach glebowych [25]. Okazuje się, że niezależnie od typu gleby liczba unikalnych OTU_{0,03} (gatunków) archeonów i grzybów może być równa, a nawet przewyższać liczbę unikalnych OTU_{0,03} bakterii, co jest nowym spostrzeżeniem, gdyż przyjęło się uważać bakterie za najbardziej zróżnicowaną grupę mikroorganizmów glebowych [6]. Bakterie jednakże mogą charakteryzować się podobnym bogactwem gatunków (ok. 10⁴ OTU_{0,03}) w różnych glebach, podczas gdy archeony licznie są reprezentowane w glebie pustynnej (10⁷ OTU_{0,03}), grzyby w glebie preriowej (10⁹ OTU_{0,03}), a wirusy w glebie lasu deszczowego (10⁸ OTU_{0,02}). Badania F i e r e r i i wsp. [25] dobitnie wykazały, że w przypadku każdej z grup mikroorganizmów występuje minimalne taksonomiczne podobieństwo między glebami z różnych typów ekosystemów, co wyraźnie sugeruje, że

bakterie, archeony, grzyby oraz wirusy są zróżnicowane zarówno w skali globalnej, jak i lokalnej.

Na podstawie szacunków, że w glebie można się spodziewać od 2000 do 8,3 mln gatunków drobnoustrojów, G a n s i i wsp. [27] wskazali na niepraktyczność stosowania amplifikacji i sekwencjonowania genu 16S rRNA w celu weryfikacji i identyfikacji liczby gatunków bakterii w glebie. W przypadku tej procedury, liczba gatunków jest ograniczona tylko do kilku tysięcy w gramie gleby. Pirosekwencjonowanie daje rzetelnniejsze wyniki. R o e s c h i i wsp. [68], wykorzystując tę metodę w odniesieniu do metagenomów czterech różnych gleb, otrzymali ponad 25 000 pirotagów dla każdej z gleb, co pozwoliło oszacować maksymalną liczbę OTU w glebie, przy najwyższym poziomie rozdzielczości filogenetycznej (0% różnic), na około 52 000. Przy tym poziomie rozdzielczości, bezpośrednio obserwowana liczba OTU wynosi od 3000 do 7000, a po ekstrapolacji do postaci krzywej rozgęszczenia, odczytana liczba OTU osiąga wartości między 11 000 a 21 000 w gramie gleby. Niezależnie więc od sposobu określania liczby gatunków w glebie, pirosekwencjonowanie pokazuje, że jest ich ponad 100 razy mniej niż sugeruje G a n s i i wsp. [27]. Analiza porównawcza pozwala stwierdzić, że gleby uprawne mogą charakteryzować się bogactwem różnych gatunków bakterii oraz archeonów z typu *Crenarchaeota* (4–12%), ale są ubogie na poziomie typów. Natomiast gleba leśna może obfitować w różne typy bakterii przy znikomym udziale (0,009%) archeonów [68]. Najczęściej reprezentowanymi spośród ogółu 30 typów występujących w tego typu glebach są *Proteobacteria* (40%), następnie *Bacteroidetes* (15–25%), *Acidobacteria* (3–12%), *Actinobacteria* (2–9%), *Firmicutes* (2–4%) oraz bakterie niesklasyfikowane (6–12%). Wśród pozostałych typów zwraca uwagę większa obecność *Gemmatimonadetes* (ok. 4%) i *Thermotogae* (ok. 2,5%) w glebie borealnej, a *Nitrospira* (2–3%) oraz TM 7 (ok. 4%) w glebach uprawnych.

F u l t h o r p e i i wsp. [26] stwierdzili, że na poziomie rozdzielczości filogenetycznej przyjętym dla gatunku (0 oraz 3% różnic) różne gleby mogą wykazywać nieznaczne podobieństwo (odpowiednio 1,5 oraz 4,1%), czyli w drastycznie wysokim stopniu charakteryzują się unikalnym składem gatunkowym mikrobiomów. Wśród 10 najliczniej reprezentowanych rodzajów, które obejmują od 31 do 37% wszystkich sekwencji pirotagów, tylko dwa, *Chitinophaga* (*Bacteroidetes*) oraz *Acidobacterium* (*Acidobacteria*), występują w glebach uprawnych oraz borealnej w znacznie większych ilościach (odpowiednio 7,5–13,8% oraz 2,4–9,3%). Inny często obecny rodzaj, *Acidovorax* (1,8–7,5%), występuje tylko w glebach uprawnych. Zwracający uwagę jest fakt, że spośród 10 uznanych za najistotniejsze w glebie rodzaje bakterii (*Agrobacterium*, *Alcaligenes*, *Arthrobacter*, *Bacillus*, *Flavobacterium*, *Micromonospora*, *Nocardia*, *Pseudomonas* oraz *Streptomyces*) tylko trzy są potwierdzone

przez wyniki pirosekwencjonowania metagenomowego DNA, a mianowicie *Bacillus*, *Flavobacterium* oraz *Pseudomonas*. Natomiast pojawiają się w różnych glebach nieoczekiwane takie rodzaje, jak *Chitinophaga*, *Thiobacter*, *Gemmatimonas*, *Chondromyces*, *Nevskia*, *Aquabacterium* czy *Dyadobacter* [26].

Jeszcze bardziej holistyczny obraz mikrobiomu glebowego można uzyskać dzięki bezpośrednio izolowanemu metagenomowemu RNA z dużych (LSU) i małych podjednostek (SSU) rybosomów, który poddany odwrotnej transkrypcji generuje produkty cDNA do procesu 454-pirosekwencjonowania [87]. Taka procedura pozwala uzyskać ponad 250 000 wysokiej jakości sekwencji RNA-tagów o wielkości około 98 pz, spośród których 74,8% jest rybotagami (rozpoznawalne taksonomicznie), w równej proporcji SSU oraz LSU. Prowadzi to do równoczesnego identyfikowania taksonów należące do bakterii, archeonów oraz eukariotów, a także tych mikroorganizmów, dla których brak jest zestawów starterów do amplifikacji PCR. W glebie piaszczystej najliczniej reprezentowane oraz najbardziej zróżnicowane okazują się bakterie w liczbie 19 spośród 24 znanych typów, oraz 20 o statusie kandydatów stanowiących jedynie 2% wszystkich rybotagów. Najliczniej spotykane są *Actinobacteria*, *Proteobacteria*, *Firmicutes*, *Planctomycetes* oraz *Acidobacteria*. Grzyby reprezentowane są głównie przez *Ascomycota* obejmujące dwie trzecie rybotagów, zaś najmniej liczną grupę zarówno wśród SSU (1,5%), jak i LSU (1,4%) rybotagów stanowią archeony (Grupa I.1b należąca do typu *Crenarchaeota*) zaangażowane w proces utleniania amoniaku w glebie [48]. Takie podejście metatranskryptomomiczne, zaprezentowane przez U r i c h i wsp. [87], daje także możliwość równoczesnego określenia aktywności fizjologicznej mikrobiomu. Wśród otrzymanych RNA-tagów ponad 21 000 zostało określone jako sekwencje mRNA odpowiadające transkryptomu całej społeczności mikroorganizmów glebowych.

Analiza metagenomiczna pozwala w nowym świetle spojrzeć na: (a) efekty nawożenia mineralnego lub organicznego na strukturalną różnorodność i aktywność fizjologiczną mikrobiomu glebowego [63], (b) wpływ rodzaju gospodarowania rolniczego lub leśnego na skład społeczności mikroorganizmów w glebie [55], (c) rozmieszczenie populacji mikroorganizmów w profilu gleby [94] oraz różnych jej frakcjach [15], (d) strukturę mikrobiomu ryzosferowego [38, 88].

Nawożenie gleby prowadzi jedynie do nieznacznej zmiany w strukturze mikrobiomu, reprezentowanego w 90–94% przez pirotagi pochodzące z bakterii, wynikającej głównie z większej obecności *Firmicutes* oraz archeonów (głównie *Crenarchaeota*) w glebie nawożonej oraz *Cyanobacteria* w glebie nienawożonej. Natomiast zmniejszenie liczebności *Rhizobiales* dają się zauważyć w glebie nawożonej [63].

Wśród ponad 590 000 pirotagów (255 pz) bakterii, pochodzących z gleb terenów podlegających różnemu rodzajowi gospodarowania rolniczego lub leśnego, sklasyfikowano 79,3% do 17 typów, które w większej ilości występują w glebach trawiastych niż leśnych. Jako dominujące grupy bakterii da się wyróżnić *Acidobacteria*, *Alphaproteobacteria*, *Actinobacteria*, *Betaproteobacteria*, *Deltaproteobacteria* oraz *Firmicutes*. Więcej pirotagów należących do *Actinobacteria*, *Firmicutes*, *Verrucomicrobia*, *Cyanobacteria* oraz *Gemmatimonadetes* występuje w glebach pastwiskowych, natomiast *Alphaproteobacteria* przeważają w glebach leśnych [55]. Pirosekwencjonowanie pozwala także dostrzec znaczące różnice w strukturze mikrobiomów na poziomie rodzajów. Jakkolwiek *Mycobacterium* dominuje w obu typach gospodarowania glebami, to w przypadku *Phenylobacter*, *Polyangium*, *Azospirillum* i *Dyella* można stwierdzić większą reprezentację w glebach leśnych, natomiast *Bacillus*, *Kribbella*, *Agromyces*, *Defluviicoccus*, *Rubrobacter*, *Streptomyces* i *Bradyrhizobium* przeważają w glebach trawiastych. Na poziomie przypuszczalnych gatunków (OTU_{0,03}) średnio 3219 OTU może występować w glebach leśnych, natomiast 2611 OTU w glebach trawiastych [55].

Mikroorganizmy występują nie tylko w powierzchniowej warstwie gleby, gdzie są najczęściej badane, ale także głębiej, gdzie są mniej poznane. Jak się okazuje, większa różnorodność bakterii jest w warstwie próchnicznej (A), gdzie znacząca liczba spośród wykrytych 750 000 pirotagów reprezentuje *Actinobacteria*, *Bacteroidetes*, *Cyanobacteria*, *Fibrobacteres*, *Firmicutes*, *Spirochaetes*, *Verrucomicrobia* oraz *Proteobacteria*. Dla odmiany większość sekwencji DNA należących do *Acidobacteria*, *Chloroflexi*, *Gemmatimonadetes*, *Nitrospira*, TM7 oraz WS3 można wyizolować z metagenomu poziomu wzbogacania [94].

Gleba ma swoistą budowę strukturalną, co umożliwia wyróżnienie trzech głównych mikrośrodków: makroagregatowe (powyżej 250 μm), mikroagregatowe (53–250 μm) oraz ilasto-pylaste (poniżej 53 μm). W tych mikrośrodkach społeczności mikroorganizmów różnią się składem oraz liczebnością w odpowiedzi na odmienne warunki chemiczne i fizyczne. Porównując mikrośrodkówiska, okazuje się że, makroagregaty zawierają więcej przedstawicieli *Alphaproteobacteria*, *Gammaproteobacteria*, *Actinobacteria* oraz *Bacteroidetes*, natomiast we frakcji mikroagregatowej gleby występuje więcej *Rubrobacteriales* z typu *Actinobacteria*, oraz *Chloroflexi*. W trzecim mikrośrodku, ilasto-pylastym, liczniej reprezentowane są *Deltaproteobacteria*, *Betaproteobacteria*, *Gemmatimonadetes*, WS3, *Acidobacteria*, OP10 oraz *Nitrospirae*. Analiza pirotagów wykazuje, że środowisko mikroagregatów charakteryzuje się mniejszą różnorodnością gatunkową (ok. 2500 OTU_{0,03}), aniżeli frakcje makroagregatowa oraz ilasto-pylasta, gdzie stwierdza się ponad 3000 OTU_{0,03} [15].

Istnienie efektu ryzosferowego pod postacią selektywnego wpływu rośliny na skład mikrobiomu w ryzosferze wyraźnie pokazują wyniki pirosekwencjonowania fragmentów metagenomowego DNA. Na przykład, w ryzosferze dębu przeważającą liczbę otrzymywanych pirotagów stanowią przedstawiciele *Betaproteobacteria*, *Gammaproteobacteria* oraz niesklasyfikowanych proteobakterii. Bardziej dogłębna analiza pozwoli zaobserwować, większą reprezentację rodziny *Burkholderiaceae* z rodzajami *Burkholderia* i *Collimonas* w ryzosferze, zaś w glebie przeważanie przedstawicieli rodziny *Hyphomicrobiaceae* z rodzajem *Rhodoplanes* oraz bakterie z rzędu *Acidimicrobiales* [88]. W ryzosferze roślin zielnych, np. ziemniaka, także wyraźnie daje się dostrzec silny wpływ rośliny na występujące w niej mikroorganizmy. Wyróżnić tu można 25 typów bakterii znacząco (8–50%) reprezentowanych przez *Actinobacteria* oraz *Alphaproteobacteria* oraz w mniejszym stopniu (1–5%) przez *Gammaproteobacteria*, *Betaproteobacteria*, *Acidobacteria*, *Gemmatimonadetes*, *Firmicutes*, *Verrucomicrobia*, *Bacteroidetes* oraz TM7. Nowością jest odkrycie, że najliczniej reprezentowanym rodzajem w ryzosferze ziemniaka jest *Rhodanobacter* stanowiący od 8 do 26% wszystkich sklasyfikowanych sekwencji DNA. Na poziomie gatunku (OTU_{0,03}) można wykryć ponad 7300 różnych OTU w gramie gleby ryzosferowej, co jest jednakże liczbą wyraźnie mniejszą niż w przypadku gleby poza ryzosferowej (ponad 9000 OTU) [38]. Analiza metagenomiczna uwidacznia wyraźnie selekcjonujący wpływ rośliny na różnorodność i strukturę mikrobiomu ryzosferowego, która zależna jest od gatunku rośliny i jej stadium rozwojowego.

7. Podsumowanie

Analiza metagenomiczna jawi się jako poszukiwany od dziesięcioleci Graal mikrobiologii środowisk, który umożliwi dokonania holistycznego spojrzenia na strukturalną różnorodność mikrobiomu oraz jego funkcjonowanie w czasie i przestrzeni. Złożoność problemu, jeśli chodzi o oddziaływania między środowiskiem a mikrobiomem, jest tak wielka, zwłaszcza w przypadku gleby, że nawet przy zastosowaniu najnowszej generacji dogłębnych technologii analizy metagenomu możemy dostrzec zaledwie część (przeciętnie od 30 do 60%) świata mikroorganizmów, a o aktywności fizjologicznej i roli ekologicznej poszczególnych taksonów (gatunków) wiemy jeszcze mniej. Metagenomika rozbudziła olbrzymie nadzieje na dotarcie do mikroorganizmów niehodowanych, a wśród nich także do tych, które wchodzi w skład „rozrzedzonej biosfery”. Jakie jest wśród nich bogactwo poszczególnych taksonów idąc od typów, przez klasy, rzędy, rodziny czy wreszcie rodzaje i gatunki? Które z nich i w jakich warunkach są liczniej

reprezentowane, które zawsze stanowią tło, a które i pod wpływem jakich czynników abiotycznych i biotycznych środowiska podlegają ograniczeniu lub eliminacji? Jak rozmieszczone są poszczególne populacje i jakim podlegają zmianom w skali mikroagregatów glebowych, jak przekłada się to na strukturę mozaiki mikrobiotopów w glebie? Które z tych mikroorganizmów są kosmopolityczne, a które wykazują charakter endemiczny, a więc czy jest możliwe rzetelne poznawanie biogeografii drobnoustrojów i porównywanie mikrobiomów glebowych, nie tylko w skali ekosystemu czy regionu, ale także kontynentu czy globalnie? Wielkie nadzieje wzbudza metagenomika w obszarze poszukiwania genów nowych substancji (metabolitów) o znanych funkcjach i przeznaczeniu, jak i metabolitów całkiem nowych. Które mikroorganizmy i w jakich warunkach są producentami tych substancji? Kiedy i pod wpływem jakich czynników dochodzi do ekspresji poszukiwanych genów w metagenomie glebowym? Czy ekspresja, np. genów katabolicznych, zachodzi w sposób skoordynowany u wszystkich właścicieli a oni współpracują w rozkładaniu substratów, czy też występuje konkurencja i formowanie się swoistej hierarchii pod względem możliwości ich wykorzystania?

W kontekście tych nadziei możliwości metagenomiki nie są pozbawione ograniczeń. Do badań przede wszystkim wymagany jest DNA/RNA o odpowiednim stopniu czystości, możliwie jak najwyższym, o wielkości efektywnej dla poprawnej analizy filogenetycznej i funkcjonalnej. Spełnienie tych warunków jest olbrzymim wyzwaniem podczas badania metagenomu glebowego. Każdy z następujących etapów jest źródłem kolejnych możliwych błędów, jak np. brak adekwatnych dla różnych mikroorganizmów starterów dla PCR czy oligonukleotydów dla konstruowania sond hybrydacyjnych, preferencyjne amplifikowanie pewnych fragmentów DNA (często rzadko reprezentowanych w ogólnej puli), problem z doбором wektora i gospodarza w procesie klonowania, zbyt mało liczne biblioteki klonów w sekwencjonowaniu dideoksy, czy występowanie olbrzymiej ilości krótkich fragmentów DNA podczas pirosekwencjonowania wymagającej rozbudowanej analizy bioinformatycznej dla uzyskania efektywnej informacji filogenetycznej. W konsekwencji otrzymywane wyniki, jakkolwiek obejmują mikroorganizmy niezależnie od ich podatności na wyhodowanie, dotyczą tych członków mikrobiomu glebowego, z których uzyskano DNA lub RNA, zachowujące swoją reprezentatywność od ekstrakcji przez kolejne etapy przetwarzania. W związku z tym można przyjąć, że swój wyraz najczęściej znajdują mikroorganizmy dominujące i najłatwiej dostępne w glebie. Różnica między metodami polega na tym, jak daleko są w stanie poszerzyć tę grupę o kolejne populacje ze zakresu „rozrzedzonej biosfery”.

Postęp technologiczny i malejące koszty sekwencjonowania oraz rozbudowywane bazy mikromacierzy DNA

dają asumpt do coraz powszechniejszego stosowania analizy metagenomicznej. Budowane bazy metagenomów gromadzą dane dostarczane przez zespoły badawcze, które służą też dla analiz porównawczych dla nowych projektów. Baza IMG/M (Integrated Microbial Genomes with Microbiome Samples – <http://www.img.jgi.doe.gov/cgi-bin/m/main.cgi>) zgromadziła 385 metagenomów ze środowiska, z czego 64 to metagenomy glebowe. Inna baza, GOLD 4.0 (DOE Joint Genome Institute), zawiera 370 metagenomy, a wśród nich 45 to metagenomy glebowe, 11 pochodzi z ryzoplenu, a 1 z fylosfery.

Analiza metagenomiczna znacząco poszerza obraz mikrobiomu glebowego, ale w dalszym ciągu jest on nieostry i wygląda tak, jakbyśmy chcieli dostrzec detale przy niewielkim powiększeniu. Miliardy drobnoustrojów, zawierające 2,142 km DNA niosący 4×10^{12} genów, zamieszkuje każdy gram gleby, aktywnie wpływając na jej właściwości [85]. Wiele z nich przeżywa, wiele wymiera, niezliczone ilości podlegały szybkiemu procesowi specjacji. Mikrobiom glebowy jest zatem dynamiczną jednostką podlegającą ustawicznej ewolucji i należy mieć nadzieję, że choć trochę bardziej będziemy go rozumieć dzięki badaniom na poziomie metagenomu.

Piśmiennictwo

- Andersen G.L., He Z., DeSantis T.Z., Brodie E.L., Zhou J.: The use of microarrays in microbial ecology (w) *Environmental Molecular Microbiology*, red. Liu W.-T., Jansson J.K., Caister Academic Press, Norfolk, 2010, s. 87–110
- Baker K.L., Langenheder S., Nicol G.W., Ricketts D., Killham K., Campbell C.D., Prosser J.I.: Environmental and spatial characterisation of bacterial community composition in soil to inform sampling strategies. *Soil Biol. Biochem.* **41**, 2292–2298 (2009)
- Barns S.M., Fundyga R.E., Jeffries M.W., Pace N.R.: Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proc. Natl. Acad. Sci. USA*, **91**, 1609–1613 (1994)
- Brown T.A.: *Genomy*. Wyd. Naukowe PWN, Warszawa, 2009
- Bru D., Ramette A., Saby N.P., Dequiedt S., Ranjard L., Jolivet C., Arrouays D., Philippot L.: Determinants of the distribution of nitrogen-cycling microbial communities at the landscape scale. *ISME J.* **5**, 532–542 (2011)
- Buckley, D., Schmidt T.: *Exploring the biodiversity of soil – a microbial rain forest* (w) *Biodiversity of Microbial Life*, red. Staley J., Reysenbach A., John Wiley & Sons, New York, 2002, s. 183–208
- Caffrey S.M.: Which microbial communities are present? Sequence-based metagenomics (w) *Applied Microbiology and Molecular Biology in Oilfield Systems*, red. Whitby C., Skovhus T.L., Springer Science+Business Media B.V., Dordrecht, 2011, s. 63–76
- Charles T.C.: The potential for investigation of plant-microbe interactions using metagenomics methods (w) *Metagenomics. Theory, Methods and Applications*, red. Marco D., Caister Academic Press, Norfolk, 2010, s. 107–118
- Chen I.C., Lin W.D., Hsu S.K., Thiruvengadam V., Hsu W.H.: Isolation and characterization of a novel lysine racemase from a soil metagenomic library. *Appl. Environ. Microbiol.* **75**, 5161–5166 (2009)
- Connon S.A., Giovannoni S.J.: High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl. Environ. Microbiol.* **68**, 3878–3885 (2002)
- Cooper M., La Duc M.T., Probst A., Vaishampayan P., Stam Ch., Bernardini J.N., Piceno Y.M., Andersen G.L., Venkateswaran K.: Comparison of innovative molecular approaches and standard spore assays for assessment of surface cleanliness. *Appl. Environ. Microbiol.* **77**, 5438–5444 (2011)
- Courtois, S., Pernodet J.L. i wsp.: Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Appl. Environ. Microbiol.* **69**, 49–55 (2003)
- Cubillos-Ruiz A., Junca H., Baena S., Venegas I., Zambrano M.M.: Beyond metagenomics: integration of complementary approaches for the study of microbial communities (w) *Metagenomics. Theory, Methods and Applications*, red. Marco D., Caister Academic Press, Norfolk, 2010, s. 15–38
- Daniel R.: The metagenomics of soil. *Nat. Rev. Microbiol.* **3**, 470–478 (2005)
- Davinic M., Fultz L.M., Acosta-Martinez V., Calderón F.J., Cox S.B., Dowd S.E., Allen V.G., Zak J.C., Moore-Kucera J.: Pyrosequencing and mid-infrared spectroscopy reveal distinct aggregate stratification of soil bacterial communities and organic matter composition. *Soil Biol. Biochem.* **46**, 63–72 (2012)
- Davis K.E.R., Sangwan P., Janssen P.H.: *Acidobacteria, Rubrobacteridae and Chloroflexi* are abundant among very slow-growing and mini-colony-forming soil bacteria. *Environ. Microbiol.* **13**, 798–805 (2011)
- DeAngelis K.M., Pace N.R. i wsp. Phylogenetic stains: ribosomal RNA-based probes for the identification of single cells. *Science*, **243**, 1360–1363 (1989)
- Dowd S.E., Callaway T.R., Wolcott R.D., Sun Y., McKeenhan T., Hagevoort R.G., Edrington T.S.: Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiol.* **8**, 125 (2008)
- Dugat-Bony E., Peyretailade E., Parisot N., Biderre-Petit C., Jaziri F., Hill D., Rimour S., Peyret P.: Detecting unknown sequences with DNA microarrays: explorative probe design strategies. *Environ. Microbiol.* **14**, 356–371 (2012)
- Dunbar J., Barns S.M., Ticknor L.O., Kuske C.R.: Empirical and theoretical bacterial diversity in four Arizona soils. *Appl. Environ. Microbiol.* **68**, 3035–3045 (2002)
- Ekkers D.M., Cretoiu M.S., Kielak A.M., van Elsland J.D.: The great screen anomaly – a new frontier in product discovery through functional metagenomics. *Appl. Microbiol. Biotechnol.* **93**, 1005–1020 (2012)
- Elend C., Schmeisser C., Leggewie C., Babiak P., Carballeira J.D., Steele H.L., Reymond J.L., Jaeger K.E., Streit W.R.: Isolation and biochemical characterization of two novel metagenome-derived esterases. *Appl. Environ. Microbiol.* **72**, 3637–3645 (2006)
- Elshahed M.S., Krumholz L.R. i wsp.: Novelty and uniqueness patterns of rare members of the soil biosphere. *Appl. Environ. Microbiol.* **74**, 5422–5428 (2008)
- Feinstein L.M., Sul W.J., Blackwood C.B.: Assessment of bias associated with incomplete extraction of microbial DNA from soil. *Appl. Environ. Microb.* **75**, 5428–5433 (2009)
- Fierer N., Jackson R.B. i wsp.: Metagenomic and small-subunit rRNA analyses reveal the genetic diversity of *Bacteria*, *Archaea*, *Fungi*, and viruses in soil. *Appl. Environ. Microbiol.* **73**, 7059–7066 (2007)
- Fulthorpe R.R., Roesch L.F.W., Riva A., Triplett E.W.: Distantly sampled soils carry few species in common. *ISME J.* **2**, 901–910 (2008)

27. Gans J., Wolinsky M., Dunbar J.: Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science*, **309**, 1387–1390 (2005)
28. Giovannoni S.J., DeLong E.F., Olsen G.J., Pace N.R.: Phylogenetic group-specific oligodeoxynucleotide probes for identification of single microbial cells. *J. Bacteriol.* **170**, 720–726 (1988)
29. Giovannoni S.J., Britschgi T.B., Moyer C.L., Field K.G.: Genetic diversity in Sargasso Sea bacterioplankton. *Nature*, **345**, 60–63 (1990)
30. Guazzaroni M.E., Golyshin P.N., Ferrer M.: Analysis of complex microbial communities through metagenomic survey (w) Metagenomics. Theory, Methods and Applications, red. Marco D., Caister Academic Press, Norfolk, 2010, s. 55–78
31. Hamady M., Walker J.J., Harris J.K., Gold N.J., Knight R.: Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat. Methods*, **5**, 235–237 (2008)
32. Handelsman J.: Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* **68**, 669–685 (2004)
33. Handelsman, J., Rondon M.R., Brady S.F., Clardy J., Goodman R.M.: Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* **5**, R245–R249 (1998)
34. He Z., Zhou J. i wsp.: GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. *ISME J.* **4**, 1167–1179 (2010)
35. Healy F.G., Ray R.M., Aldrich H.C., Wilkie A.C., Ingram L.O., Shanmugam K.T.: Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulose. *Appl. Microbiol. Biotechnol.* **43**, 667–674 (1995)
36. Henne, A., Daniel, R., Schmitz, R.A., Gottschalk, G.: Construction of environmental DNA libraries in *Escherichia coli* and screening for the presence of genes conferring utilization of 4-hydroxybutyrate. *Appl. Environ. Microbiol.* **65**, 3901–3907 (1999)
37. Henne, A., Schmitz R.A., Bomeke M., Gottschalk G., Daniel R.: Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. *Appl. Environ. Microbiol.* **66**, 3113–3116 (2000)
38. Inceoğlu Ö., Al-Soud W.A., Falcão Salles J., Semenov A.V., van Elsas J.D.: Comparative analysis of bacterial communities in a potato field as determined by pyrosequencing. *PLoS ONE* **6**(8): e23321. doi:10.1371/journal.pone.0023321
39. Janssen P.H.: Identifying the dominant soil bacteria taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl. Environ. Microbiol.* **72**, 1719–1728 (2006)
40. Janssen P.H., Yates P.S., Grinton B.E., Taylor P.M., Sait M.: Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions *Acidobacteria*, *Actinobacteria*, *Proteobacteria* and *Verrucomicrobia*. *Appl. Environ. Microbiol.* **68**, 2391–2396 (2002)
41. Joseph S.J., Hugenholtz P., Sangwan P., Osborne C.A., Janssen P.H.: Laboratory cultivation of widespread and previously uncultured soil bacteria. *Appl. Environ. Microbiol.* **69**, 7210–7215 (2003)
42. Kakirde K.S., Parsley L.C., Liles M.R.: Size does matter: application-driven approaches for soil metagenomics. *Soil Biol. Biochem.* **42**, 1911–1923 (2010)
43. Kang S., Mills A.L.: The effect of sample size in studies of soil microbial community structure. *J. Microbiol. Methods*, **66**, 242–250 (2006)
44. Kellenberger E.: Exploring the unknown: the silent revolution of microbiology. *EMBO Rep.* **2**, 5–7 (2001)
45. Kozdrój J.: Izolacja kwasów nukleinowych ze środowiska – pierwszy krok w analizie metagenomu. *Kosmos*, **59**, 141–150 (2010)
46. Kozdrój J.: Różnorodność mikroorganizmów glebowych w świetle badań molekularnych. *Post. Mikrobiol.* **43**, 375–398 (2004)
47. Kunin V., Engelbrektson A., Ochman H., Hugenholtz Ph.: Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ. Microbiol.* **12**, 118–123 (2010)
48. Leininger S., Urich T., Schloter M., Schwark L., Qi J., Nicol G.W., Prosser J.I., Schuster S.C., Schleper C.: *Archaea* predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**, 806–809 (2006)
49. Liesack W., Stackebrandt E.: Occurrence of novel groups of the domain *Bacteria* as revealed by analysis of genetic material isolated from an Australian terrestrial environment. *J. Bacteriol.* **174**, 5072–5078 (1992)
50. Liles M.R., Manske B.F., Bintrim S.B., Handelsman J., Goodman R.M.: A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Appl. Environ. Microbiol.* **69**, 2684–2691 (2003)
51. Lombard N., Prestat E., Van Elsas J.D., Simonet P.: Soil-specific limitations for access and analysis of soil microbial communities by metagenomics. *FEMS Microbiol. Ecol.* **78**, 31–49 (2011)
52. Madsen E.L.: Environmental Microbiology. From genomes to biogeochemistry. Blackwell Publishing, Oxford, 2008
53. Mirete, S., de Figueras, C.G., Gonzalez-Pastor, J.E.: Novel nickel resistance genes from the rhizosphere metagenome of plants adapted to acid mine drainage. *Appl. Environ. Microbiol.* **73**, 6001–6011 (2007)
54. Mummey D., Holben W., Six J., Stahl P.: Spatial stratification of soil bacterial populations in aggregates of diverse soils. *Microb. Ecol.* **51**, 404–411 (2006)
55. Nacke H., Thürmer A., Wollherr A., Will Ch., Hodac L., Herold N., Schöning I., Schrumpf M., Daniel R.: Pyrosequencing-based assessment of bacterial community structure along different management types in German forest and grassland soils. *PLoS ONE* **6**(2): e17000. doi:10.1371/journal.pone.0017000 (2011)
56. Nelson K.E., Bryan P.A., White B.A.: Genomics and metagenomics: history and progress. (w) Metagenomics. Theory, Methods and Applications red. Marco D., Caister Academic Press, Norfolk, 2010, s. 21–36
57. Nicol G.W., Glover L.A., Prosser J.I.: Spatial analysis of archaeal community structure in grassland soil. *Appl. Environ. Microbiol.* **69**, 7420–7429 (2003)
58. Ono A., Miyazaki R., Sota M., Ohtsubo Y., Nagata Y., Tsuda M.: Isolation and characterization of naphthalene-catabolic genes and plasmids from oil contaminated soil by using two cultivation-independent approaches. *Appl. Microbiol. Biotechnol.* **74**, 501–510 (2007)
59. Øvreås L.: Population and community level approaches for analysing microbial diversity in natural environments. *Ecol. Lett.* **3**, 236–251 (2000)
60. Pace N.R., Stahl D.A., Lane D.J., Olsen G.J.: The analysis of natural microbial populations by ribosomal RNA sequences. *Adv. Microb. Ecol.* **9**, 1–55 (1986)
61. Parsley L.C., Consuegra E.J., Kakirde K.S., Land A.M., Harper Jr. W.F., Liles M.R.: Identification of diverse antimicrobial resistance determinants carried on bacterial, plasmid, or viral metagenomes from an activated sludge microbial assemblage. *Appl. Environ. Microbiol.* **76**, 3753–3757 (2010)
62. Pignatelli M., Aparicio G., Blanquer I., Hernandez V., Moya A., Tamames J.: Metagenomics reveals our incomplete knowledge of global diversity. *Bioinformatics* **24**, 2124–2125 (2008)
63. Poulsen P.H.B., Abu Al-Soud W., Bergmark L., Magid J., Hansen L.H., Sørensen S.J.: Effects of fertilization with urban and agricultural organic wastes in a field trial – Prokaryotic

- diversity investigated by pyrosequencing. *Soil Biol. Biochem.* <http://dx.doi.org/10.1016/j.soilbio.2011.12.023> (2012)
64. Prosser J.I.: Replicate or lie. *Environ. Microbiol.* **12**, 1806–1810 (2010)
 65. Ranjard L., Lejon D.P., Mougel C., Schehrer L., Merdinoglu D., Chaussod R.: Sampling strategy in molecular microbial ecology: influence of soil sample size on DNA fingerprinting analysis of fungal and bacterial communities. *Environ. Microbiol.* **5**, 1111–1120 (2003)
 66. Riesenfeld C.S., Goodman R.M., Handelsman J.: Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environ. Microbiol.* **6**, 981–989 (2004)
 67. Robe P., Nalin R., Capellano C., Vogel T.M., Simonet P.: Extraction of DNA from soil. *Eur. J. Soil Biol.* **39**, 183–190 (2003)
 68. Roesch L.F.W., Fulthorps R.R., Riva A., Casella G., Hadwin A.K.M., Kent A.D., Daroub S.M., Camargo F.A.O., Farmerie W.G., Triplett E.W.: Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J.* **1**, 283–290 (2007)
 69. Rondon M.R., Goodman R.M. i wsp.: Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* **66**, 2541–2547 (2000)
 70. Rousk J., Brookes P.C., Bååth E.: Investigating the mechanisms for the opposing pH relationships of fungal and bacterial growth in soil. *Soil Biol. Biochem.* **42**, 926–934 (2010)
 71. Schloss P.D., Handelsman J.: Toward a census of bacteria in soil. *PLoS Comput. Biol.* **2**, 786–793 (2006)
 72. Schmidt T.M., DeLong E.F., Pace N.R.: Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J. Bacteriol.* **173**, 4371–4378 (1991)
 73. Schwartz I.: Microbial genomics, from sequence to function. *Emerg. Infect. Dis.* **6**, 493–495 (2000)
 74. Singer M.E., Tom L.M., Andersen G.L.: PCR amplification-independent methods for detection of microbial communities by the high-density microarray PhyloChip. *Appl. Environ. Microbiol.* **77**, 6313–6322 (2011)
 75. Sogin M.L., Morrison H.G., Huber J.A., Welch M.D., Huse S.M., Neal P.R., Arrieta J.M., Herndl G.J.: Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc. Natl. Acad. Sci. USA*, **103**, 12115–12120 (2006)
 76. Sogin M.L.: Characterizing microbial population structures through massively parallel sequencing (w) Microbiology Monographs Vol. 10, red. Epstein S.S., Springer-Verlag Berlin, Heidelberg, 2009, s. 19–33
 77. Stahl D.A., Lane D.J., Olsen G.J., Pace N.R.: Characterization of a Yellowstone hot spring microbial community by 5S rRNA sequences. *Appl. Environ. Microbiol.* **49**, 1379–1384 (1985)
 78. Staley J.T., Konopka A.: Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu. Rev. Microbiol.* **39**, 321–346 (1985)
 79. Stein J.L., T.L. Marsh K.Y., Shizuya W.H., DeLong E.F.: Characterization of uncultivated prokaryotes: isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. *J. Bacteriol.* **178**, 591–599 (1996)
 80. Suenaga H., Ohnuki T., Miyazaki K.: Functional screening of a metagenomic library for genes involved in microbial degradation of aromatic compounds. *Environ. Microbiol.* **9**, 2289–2297 (2007)
 81. Thomas T., Gilbert J., Meyer F.: Metagenomics – a guide from sampling to data analysis. *BMC Microb. Inf. Exp.* **2**, 3 (2012)
 82. Torres-Cortés G., Millán V., Ramírez-Saad H.C., Nisa-Martínez R., Toro N., Martínez-Abarca F.: Characterization of novel antibiotic resistance genes identified by functional metagenomics on soil samples. *Environ. Microbiol.* **13**, 1101–1114 (2011)
 83. Torsvik V., Goksoyr J., Daae F.L.: High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* **56**, 782–787 (1990)
 84. Torsvik V., Øvreås L.: Microbial diversity and function in soil: from genes to ecosystems. *Curr. Opin. Microbiol.* **5**, 240–245 (2002)
 85. Trevors J.T.: One gram of soil: a microbial biochemical gene library. *Antonie van Leeuwenhoek*, **97**, 99–106 (2010)
 86. Tringe S.G., Rubin E.M. i wsp.: Comparative metagenomics of microbial communities. *Science*, **308**, 554–557, 2005.
 87. Urich T., Lanzén A., Qi J., Huson D.H., Schleper Ch., Schuster S.C.: Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. *PLoS ONE*, **3**(6), e2527, doi:10.1371/journal.pone.0002527 (2008)
 88. Uroz S., Buée M., Murat C., Frey-Klett P., Martin F.: Pyrosequencing reveals a contrasted bacterial diversity between oak rhizosphere and surrounding soil. *Environ. Microbiol. Rep.* **2**, 281–288 (2010)
 89. Van Elsas J.D., Costa R., Jansson J., Sjoling S., Bailey M., Nalin R., Vogel T.M., Van Overbeek L.: The metagenomics of disease-suppressive soils – experiences form the METACONTROL project. *Trends Biotechnol.* **26**, 591–601 (2008)
 90. Van Elsas J.D., Jansson J.K., Trevors J.T.: Modern Soil Microbiology II. CRC Press-Taylor and Francis, Boca Raton, 2007
 91. Von Wintzingerode F., Gobel U.B., Stackebrandt E.: Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol. Rev.* **21**, 213–229 (1997)
 92. Wang G.Y., Davies J. i wsp.: Novel natural products from soil DNA libraries in a streptomycete host. *Org. Lett.* **2**, 2401–2404 (2000)
 93. Whitman W.B., Coleman D.C., Wiebe W.J.: Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. USA*, **95**, 6578–6583 (1998)
 94. Will Ch., Daniel R. i wsp.: Horizon-specific bacterial community composition of German grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes. *Appl. Environ. Microbiol.* **76**, 6751–6759 (2010)
 95. Williams M.A., Rice C.W.: Seven years of enhanced water availability influences the physiological, structural, and functional attributes of a soil microbial community. *Appl. Soil Ecol.* **35**, 535–545 (2007)
 96. Woese C.R., Fox G.E.: Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. USA*, **74**, 5088–5090 (1977)
 97. Woese C.R., Kandler O., Whellis M.L.: Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc. Natl. Acad. Sci. USA*, **87**, 4576–4579 (1990)
 98. Wommack K.E., Bhavsar J., Ravel J.: Metagenomics: read length matters. *Appl. Environ. Microbiol.* **74**, 1453–1463 (2008)
 99. Xu J.: Metagenomics and ecosystems biology: conceptual frameworks, tools and methods (w) Metagenomics. Theory, Methods and Applications, red. Marco D., Caister Academic Press, Norfolk, 2010, s. 1–14
 100. Youssef N., Sheik C., Krumholz L., Najjar F., Roe B., Elshahed M.: A Comparative study of species richness estimates obtained using nearly complete fragments and simulated pyrosequencing-generated fragments in 16S rRNA gene-based environmental surveys. *Appl. Environ. Microbiol.* **75**, 5227–5236 (2009)
 101. Zengler K., Toledo G., Rappe M., Elkins J., Mathur E.J., Short J.M., Keller M.: Cultivating the uncultured. *Proc. Natl. Acad. Sci. USA*, **99**, 15681–15686 (2002)