

Wykorzystanie potencjału mikrobiomu środowiskowego w kryminalistyce

Agata Jagiełło¹, Anna Woźniak¹, Błażej Szczerba², Rafał Płoski³, Małgorzata Rydzanicz³, Agnieszka Pollak³, Alina Frolova^{4,5}, Michał Kowalski⁵, Paweł Łabaj⁵, Andrzej Ossowski⁶, Wojciech Branicki^{1,5}, Kinga Herda⁵, Kamila Marszałek⁵, Renata Zbieć-Piekarska^{1*}

¹ Centralne Laboratorium Kryminalistyczne Policji

² Ardigen SA

³ Warszawski Uniwersytet Medyczny

⁴ Instytut Biologii Molekularnej i Genetyki Narodowej Akademii Nauk Ukrainy

⁵ Uniwersytet Jagielloński

⁶ Pomorski Uniwersytet Medyczny w Szczecinie

* autor korespondencyjny: renata.zbiec@policja.gov.pl

Streszczenie

Konsorcjum naukowe pod przewodnictwem Centralnego Laboratorium Kryminalistycznego Policji podjęło się opracowania metody analizy DNA mikrobiomu gleby, która znajdzie zastosowanie w badaniach kryminalistycznych. Celem projektu o akronimie SMAFT (Soil Microbiome Analysis Forensic Tool, <http://smaft.eu/>), finansowanego przez Narodowe Centrum Badań i Rozwoju (DOB-BIO10/03/01/2019), jest stworzenie nowego narzędzia umożliwiającego powiązanie śladu w postaci próbki gleby z określoną lokalizacją geograficzną. W pierwszej części artykułu przybliżono pojęcie mikrobiomu oraz przedstawiono możliwości wykorzystania analiz DNA mikrobiomu w kryminalistyce. W jego drugiej części szczegółowo opisano etapy realizowanego projektu, począwszy od zbierania próbek gleby z różnych miejsc Polski w czterech porach roku i izolacji z nich DNA mikrobiomów, poprzez oparte na technologii MPS (ang. *Massively Parallel Sequencing*) sekwencjonowanie izolatów oraz opracowanie testu genetycznego zawierającego zestaw markerów metagenomicznych pozwalających na skuteczną indywidualizację próbek gleby, aż po stworzenie systemu informatycznego umożliwiającego analizę i interpretację otrzymanych wyników, który obejmuje bazę danych profili DNA mikrobiomów gleb pochodzących z różnych miejsc Polski.

Słowa kluczowe: mikrobiom, gleba, kryminalistyka, metagenomika, analizy metataksonomiczne, sekwencjonowanie MPS, bioinformatyka, projekt NCBiR

Wstęp

Centralne Laboratorium Kryminalistyczne Policji jako instytut badawczy prowadzi badania naukowe z dziedziny kryminalistyki. Realizowane tutaj projekty mają na celu opracowywanie metod i technologii wspierających zapobieganie, wykrywanie oraz zwalczanie przestępczości. Jednym z głównych zagadnień pozostających od kilku lat w obszarze zainteresowania genetyki sądowej jest wykorzystanie genomiki różnic indywidualnych, epigenomiki oraz metagenomiki i analizy metataksonomicznej do projektowania i rozwijania innowacyjnych narzędzi przewidzianych do stosowania w policyjnych laboratoriach kryminalistycznych. Obrany kierunek badań naukowych wpisuje się w priorytety wyznaczone przez państwa Sojuszu Północnoatlantyckiego (NATO) z Sojuszniczym Dowództwem Transformacji i Organizacją ds. Nauki i Technologii Sojuszu oraz Europejską Agencją Obrony na czele, uznające genomikę za jedną z dwudziestu technologii przełomowych, których

rozwijanie ma kluczowe znaczenie dla zapewnienia bezpieczeństwa.

W przeciwieństwie do genetyki, która bada głównie geny i mechanizmy ich dziedziczenia, genomika jest stosunkowo nową dziedziną biologii zajmującą się analizą genomu, czyli pełnej informacji genetycznej danego organizmu. Powstała dzięki intensywnemu rozwojowi informatyki i technik biologii molekularnej, a przede wszystkim nowoczesnych technologii sekwencjonowania, zwłaszcza masowego, równoległego sekwencjonowania – MPS (ang. *Massively Parallel Sequencing*) stanowiącego jedno z najważniejszych osiągnięć nauk biologicznych ostatnich dwóch dekad (Heather, Chain, 2017). Genomika strukturalna obejmuje określanie sekwencji genomu danego organizmu i jego organizacji oraz identyfikowanie obszarów odpowiadających poszczególnym genom. Zadaniem genomiki funkcjonalnej wraz z epigenomiką (modyfikacje DNA), transkryptomiką (mRNA), proteomiką (białka)

i metabolomiką (metabolity) jest określanie funkcji genów i regionów niekodujących, a także mechanizmów ich regulacji, badanie interakcji pomiędzy genami oraz między genami a środowiskiem, w tym analiza wpływu produktów ekspresji genów na funkcjonowanie komórek, tkanek, narządów i całych organizmów (Khodadadian i in., 2020). Poszukiwaniem ogólnych praw rządzących genami zajmuje się genomika teoretyczna, ewolucją genów – genomika porównawcza, a osobnicze zróżnicowanie genomów, najbardziej interesujące z kryminalistycznego punktu widzenia, jest przedmiotem badań prowadzonych w ramach genomiki różnic indywidualnych (Ślósarek, 2012).

Zastosowanie w badaniach kryminalistycznych genomiki w połączeniu z technologią MPS pozwala na prowadzenie analiz, które jeszcze do niedawna były zbyt kosztowne lub wręcz niewykonalne z powodu ograniczeń technicznych. Obecnie możliwe jest jednoczesne badanie markerów STR (ang. *Short Tandem Repeat*) i SNP (ang. *Single Nucleotide Polymorphism*) autosomalnego DNA, markerów chromosomów pci X i Y oraz mitochondrialnego DNA (mtDNA), pozwalające na pozyskanie oprócz „standardowych” danych identyfikacyjnych osoby również informacji dotyczących jej cech fenotypowych oraz pochodzenia biogeograficznego, nawet w przypadku małych ilości lub słabej jakości DNA znalezionej na miejscu przestępstwa (Kayser, 2015; Ambers i in., 2016; Bruijns, Tiggelaar, Gardeniers, 2018). Dodatkowo, korzystając z markerów epigenetycznych, można analizować profile metylacyjne ludzkich genomów i na ich podstawie względnie dokładnie określać wiek nieznaney osoby, rozróżniać bliźnięta monozygotyczne oraz identyfikować tkanki i płyny ustrojowe (Zbieć-Piekarska i in., 2015; Vidaki, Kayser, 2018).

Mikrobiom i jego potencjał w kryminalistyce

Pomimo ciągłego udoskonalania metod służących identyfikacji osobniczej zdarzają się sytuacje, w których zastosowanie jedynie genomiki nie pozwala rozwiązać wątpliwości pojawiających się w toku prowadzonych dochodzeń. Publikowane na przestrzeni ostatnich lat wyniki badań wskazują, że w takich sytuacjach odpowiedzi na niektóre pytania można znaleźć dzięki metagenomice, dziedzinie nauki zajmującej się badaniem DNA mikrobiomów, tj. DNA pozyskiwanego bezpośrednio ze środowiska, bez konieczności prowadzenia hodowli laboratoryjnych zamieszkujących je społeczności.

Chociaż naukowcy zajmują się badaniem mikrobiomów już od kilku dziesięcioleci, jeszcze do niedawna brakowało jasnej, uniwersalnej definicji mikrobiomu. Z tego powodu w marcu 2019 r. w ramach projektu MicrobiomeSupport (<https://www.microbiomesupport.eu/>) odbyło się połączone z warsztatami spotkanie dużej międzynarodowej grupy ekspertów zajmujących się badaniem mikrobiomów, którzy dyskutowali nad zdefiniowaniem tego terminu. Rok później Berg i in. (2020)

w opublikowanym obszernym komentarzu, będącym podsumowaniem spotkania, przedstawili wypracowaną wówczas propozycję. Opiera się ona na definicji z 1988 r. (Whipps, Lewis, Cooke, 1988), która opisuje mikrobiom jako społeczność mikroorganizmów (mikrobiota) zasiedlających ściśle określone środowisko o charakterystycznych cechach fizykochemicznych. Jak zauważają Berg i in. (2020), definicja ta oprócz społeczności drobnoustrojów o określonych właściwościach i funkcjach uwzględnia również jej interakcje ze środowiskiem prowadzące do tworzenia specyficznych nisz ekologicznych. Mikrobiom stanowiący dynamiczny i interaktywny mikroekosystem, który podlega zmianom w czasie i przestrzeni, pozostaje zintegrowany z makroekosystemami, np. eukariotycznym gospodarzem, wpływając na jego funkcjonowanie i zdrowie. Należy podkreślić, że w skład mikrobiomu wchodzi mikrobiota (tj. żywe mikroorganizmy należące do różnych taksonów) oraz miejsce ich aktywności (ang. *theatre of activity*) obejmujące całe spektrum cząsteczek wytwarzanych przez mikroorganizmy, w tym ich elementy strukturalne (kwasy nukleinowe, białka, lipidy, polisacharydy), metabolity (cząsteczki sygnałowe, toksyny, cząsteczki organiczne i nieorganiczne) oraz cząsteczki powstające w wyniku warunków panujących w otaczającym środowisku, w tym produkowane przez współistniejących gospodarzy, a także wszystkie ruchome elementy genetyczne, np. fagi, wirusy, plazmidy, transpozony, integrony oraz pozakomórkowe DNA, w tym reliktywne DNA pochodzące z martwych komórek (Berg i in., 2020).

Każdy człowiek, podobnie jak każdy inny żywy organizm, każde konkretne miejsce, np. łąka i pochodząca z niej gleba czy jezioro i pochodząca z niego woda, są unikatowymi środowiskami determinującymi w dużym stopniu, jakie mikroorganizmy i w jakich proporcjach ilościowych w nich żyją. Sekwencjonowanie i/lub analiza metataksonomiczna DNA mikrobiomu danego miejsca pozwala na określenie jego unikatowego profilu, który może być porównywany z profilami mikrobiomów pochodzących z innych siedlisk. Wyniki badań prac prowadzonych na przestrzeni ostatnich kilkudziesięciu lat wskazują, że w przyszłości możliwe będzie wykorzystywanie w celach kryminalistycznych DNA ludzkiego mikrobiomu pochodzącego z różnych miejsc ciała człowieka (Fierer i in., 2008; Ravel i in., 2011; Tridico i in., 2014; Schmedes i in., 2018). Człowiek pozostawia swój mikrobiologiczny „ślad” w miejscach, które odwiedza, np. na miejscu zbrodni (Hampton-Marcell i in., 2020), jak również na dotykanych przez siebie przedmiotach (Lax i in., 2014), takich jak telefony komórkowe (Fierer i in., 2010), garderoba (Lax i in., 2015) i tkaniny (Lee i in., 2016), oraz na osobach, z którymi ma kontakt (Neckovic i in., 2020). Dowiedziono też, że ludziom towarzyszą specyficzne „chmury” drobnoustrojów, których skład może stanowić podstawę ich identyfikacji osobniczej (Meadow i in., 2015). Dodatkowo, jak zauważają Clarke i in.

(2017), zsekwencjonowanie mikrobiomu konkretnego człowieka może dostarczyć danych wystarczających nie tylko do jego identyfikacji, ale także do uzyskania informacji o jego płci, stanie zdrowia i stylu życia, jakże ważnych z punktu widzenia organów ścigania.

Równoległe do badań prowadzonych nad ludzkim mikrobiomem w kontekście kryminalistycznym w wielu ośrodkach naukowych opracowywane są metody i narzędzia umożliwiające wykorzystanie mikrobiomu gleby do celów dochodzeniowo-śledczych i wykrywczych (ang. *forensic DNA intelligence*). Na przykład w latach 2013–2015 w ramach działań Unii Europejskiej na rzecz bezpieczeństwa w Europie realizowano projekt o akronimie MiSAFE (Microbial Soil Analysis), którego celem było opracowanie narzędzi i procedur pozwalających na rutynowe badanie próbek gleby w laboratoriach kryminalistycznych (<https://forensicmisafe.wixsite.com/misafe/project>). Powszechnie wiadomo, że wszechobecność, zróżnicowanie, a także możliwość łatwego przenoszenia na butach, pojazdach, narzędziach czy odzieży sprawia, iż gleba jest szczególnie cennym materiałem dowodowym, o którego istnieniu przestępcy często zapominają, zacierając pozostawione przez siebie ślady (Young, Austin, Weyrich, 2017). W związku z tym analiza próbki gleby może dostarczyć informacji pozwalających na powiązanie podejrzanego, ofiary lub przedmiotu z miejscem przestępstwa (Johll, 2009; Dawson, Hillier, 2010; Concheri i in., 2011), jak i na temat prawdopodobnej lokalizacji miejsca pochodzenia próbki gleby (Pirrie, Dawson, Graham, 2017). Gleby są złożonymi materiałami składającymi się z minerałów, materii organicznej, w tym organizmów żywych, gazów oraz wody (Needelman, 2013). Obecnie kryminalistyczne badania gleby opierają się na cechach fizycznych i składzie chemicznym. Analizowane są kolor gleby, jej tekstura, wielkość cząstek, pH, skład pierwiastkowy i zawartość minerałów, a niekiedy także związki organiczne, np. woski roślinne (Habtom i in., 2017; Murray, 2012; Woods i in., 2016), które dostarczają istotnych dla prowadzonego śledztwa informacji (Fitzpatrick, Raven, Self, 2017; Petraco, Kubic, Petraco, 2008). Jednak w szczególnych przypadkach, kiedy badane próbki gleb pochodzą z jednorodnego geologicznie obszaru, z miejsc zlokalizowanych bardzo blisko siebie lub charakteryzują się niską zawartością składników nieorganicznych (np. gleby torfowe), niemożliwe jest ich rozróżnienie za pomocą powyższych analiz (Giampaoli i in., 2014; Young, Austin, Weyrich, 2017; Young, Higgins, Austin, 2019). Ograniczenia związane z rozróżnieniem gleb w skali lokalnej można pokonać dzięki analizom materiału biologicznego. Szacuje się, że w 1 gramie suchej gleby znajduje się średnio: 10 miliardów wirusów, 10 miliardów bakterii i archeonów (w tym 100 milionów promieniowców), po 1 milionie grzybów i glonów, 100 tysięcy pierwotniaków i 100 nicieni (Trevors, 2010). Dodatkowo gleba może zawierać także fragmenty roślin (np. korzenie, pyłki, zarodniki, nasiona, liście) oraz inne niż nicienie

bezkęgowce (Young, Austin, Weyrich, 2017) i poza-komórkowe DNA. Ponieważ wszystkie organizmy żyjące w glebie charakteryzują się określonymi wymogami siedliskowymi, warunki środowiskowe (np. rodzaj i tekstura gleby, pH, wilgotność, temperatura, ilość węgla organicznego) odgrywają ważną rolę w kształtowaniu składu mikrobiomu, tj. społeczności mikroorganizmów żyjących w danym miejscu (Maron, Mougel, Ranjard, 2011; Pasternak i in., 2013). Gleba ze względu na swoją zmienność struktury przestrzennej nie zawiera „typowego” mikrobiomu (Fierer, 2017). Badania względnych liczebności głównych taksonów bakteryjnych i archeonów w próbkach gleby wykazały różnice nie tylko w odniesieniu do różnych typów gleby, lecz także w przypadku gleb pobieranych z miejsc oddalonych zaledwie o kilkanaście czy nawet kilka centymetrów (Habtom i in., 2019; O'Brien i in., 2016; Pasternak i in., 2013; Sensabaugh, 2009). Analiza DNA mikrobiomu glebowego może dostarczyć danych pozwalających na jej skuteczną indywidualizację i w przyszłości okazać się równie przydatna do porównywania śladów gleby lub określania miejsca jej pochodzenia jak profile ludzkiego DNA do ustalania związku między śladami biologicznymi a podejrzanym (Damaso i in., 2018).

Należy podkreślić, że chociaż coraz częściej sugeruje się możliwość wykorzystania analiz DNA mikrobiomów jako dowodów lub uzupełnienia tradycyjnych kryminalistycznych metod badawczych, dziś są one prawie nieobecne zarówno w dochodzeniach, jak i w sądach (Robinson i in., 2021). Aby potencjał mikrobiomu mógł być skutecznie wykorzystywany przez organy ścigania, konieczne jest prowadzenie dalszych badań idących przede wszystkim w kierunku wykazania, że wnioskowanie statystyczne oparte na tego typu analizach jest wystarczająco rygorystyczne, aby mogło być uznane przez sądy za dowód mający wystarczające podstawy naukowe, np. znany i akceptowany poziom błędów (Velsko, 2020; Robinson i in., 2021). Powiększanie puli próbek poddawanych analizie w prowadzonych badaniach, tworzenie baz danych mikrobiomów pochodzących z różnych środowisk na podstawie jasno określonych i dobrze udokumentowanych procedur, doskonalenie narzędzi bioinformatycznych, w tym technik uczenia maszynowego służących do interpretacji wyników, czy badanie dynamiki zmian czasowych i przestrzennych mikrobiomów to tylko przykłady celów badawczych, których realizacja może przybliżyć dzień włączenia analiz mikrobiomu do zestawu narzędzi wykorzystywanych w laboratoriach kryminalistycznych (Robinson i in., 2021).

Projekt SMAFT

Włączając się w realizację wyzwań przedstawionych powyżej, Centralne Laboratorium Kryminalistyczne Policji (CLKP) jako lider reprezentujący konsorcjum w składzie: Warszawski Uniwersytet Medyczny (WUM), Uniwersytet Jagielloński (UJ), Pomorski Uniwersytet

Medyczny (PUM) oraz firma ARDIGEN, uzyskało z Narodowego Centrum Badań i Rozwoju dofinansowanie na realizację projektu o akronimie **SMAFT (Soil Microbiome Analysis Forensic Tool, <http://smaft.eu/>)**, którego zadaniem jest wykorzystanie potencjału mikrobiomu gleby w kryminalistyce (DOB-BIO10/03/01/2019). Głównym celem projektu jest skonstruowanie predykcyjnego narzędzia do kryminalistycznej analizy DNA mikrobiomu gleby, które umożliwi geograficzną lokalizację próbki gleby o nieznanym pochodzeniu na terytorium Polski. Innymi słowy, opracowywany system sprawdzi, czy istnieje powiązanie śladu w postaci próbki gleby pobranej z obuwia, opony lub łopaty z określoną lokalizacją geograficzną. Dostarczenie tego typu dowodów może uprawdopodobniać obecność podejrzanych osób w określonych miejscach lub pozwalać na prześledzenie drogi, którą przemieszczali się przestępcy. W przyszłości uzyskane dzięki zastosowaniu systemu SMAFT informacje będą mogły ukierunkowywać, a tym samym przyspieszać śledztwa zarówno w sprawach kryminalnych, jak i związanych z atakami terrorystycznymi. Opracowany system będzie mógł być również wykorzystywany do ścigania przestępstw ekologicznych, a także w szerszym wymiarze w badaniach bioróżnorodności.

Badania zaplanowane w ramach projektu SMAFT podzielone są na kilka etapów, z których pierwszym jest pobranie próbek gleby. Projekt zakłada zebranie blisko 1000 próbek gleby, po 250 w okresie jesiennym, zimowym, wiosennym i letnim z 80 różnych lokalizacji na terenie Polski. Według autorów projektu takie podejście umożliwia wychwycenie ewentualnych różnic w składzie mikrobiomów wierzchniej warstwy gleby niezależnych od pory roku. Wyboru miejsc pobrania próbek gleby dokonano na podstawie analizy informatycznej danych z pomiarów wykonanych we wszystkich stacjach hydrologiczno-meteorologicznych zlokalizowanych na terenie całej Polski, obejmujących okres ostatnich 20 lat. Mapa Polski została podzielona na pięć obszarów różniących się warunkami klimatycznymi. Ostatecznie wybrano dwanaście głównych punktów pobierania próbek, a w okolicach każdego z nich wyznaczono po kilka (od pięciu do ośmiu) miejsc pobrania. Ustalono również, że z każdej konkretnej lokalizacji zebrane zostaną trzy próbki gleby. Dodatkowo wybierając miejsca pobrania próbek, brano pod uwagę cechy fizykochemiczne gleb występujących na terenie Polski, korzystając z danych zebranych w ramach unijnego projektu LUCAS (Land Use/Cover Area frame statistical Survey) (<https://ec.europa.eu/eurostat/web/lucas/>). Etap ten obejmuje również opracowanie szczegółowej metodyki pobierania próbek gleby, ich zabezpieczania i opisywania, wyznaczania miejsc pobrania i ich dokumentowania oraz transportu do laboratorium w odpowiednich warunkach. W kolejnym etapie projektu z zebranych próbek wyizolowane zostanie DNA mikrobiomów. Ze względu na fakt, że gleba jest trudnym, niejednorodnym materiałem, zawierającym dużo

substancji potencjalnie hamujących enzymy wykorzystywane w kolejnych zadaniach realizowanego projektu, znalezienie metody izolacji pozwalającej na otrzymanie DNA mikrobiomu gleby w odpowiedniej ilości, o odpowiedniej czystości i jakości warunkuje sukces kolejnych etapów. Izolacja DNA z bakterii gram dodatnich, gram ujemnych, tworzących przetrwalniki czy wytwarzających otoczki przebiega z różną wydajnością w zależności od zastosowanej procedury, w związku z tym pozyskanie z gleby DNA reprezentatywnego dla całej społeczności bakterii danego mikrobiomu nie jest łatwym zadaniem. W trzecim etapie projektu z izolatów DNA zostaną przygotowane biblioteki zawierające fragmenty DNA mikroorganizmów wyizolowanych z próbek gleby. W celu uzyskania jak najlepszej jakości bibliotek, o pożądanej długości fragmentów, planowana jest optymalizacja procedury tworzenia bibliotek, zarówno na etapie fragmentacji DNA, jak i amplifikacji. W procesie przygotowywania bibliotek wszystkie fragmenty DNA pochodzące z każdej próbki otrzymają znacznik (ang. *barcode*) pozwalający – po uzyskaniu danych z sekwencjonowania – na ich identyfikację i przypisanie do konkretnej próbki gleby. Długości fragmentów przygotowanych bibliotek zostaną zweryfikowane za pomocą elektroforezy kapilarnej, a stężenie DNA w bibliotekach – oznaczone metodą fluorymetryczną. Po wyliczeniu stężenia molarnego i normalizacji wszystkich bibliotek zostaną one połączone z uwzględnieniem unikatowej kombinacji indeksów. Przygotowane biblioteki w następnym – czwartym etapie zostaną zsekwencjonowane z zastosowaniem technologii SBS firmy Illumina® wykorzystującej sekwenator najnowszej generacji NovaSeq 6000 Illumina®. Zaplanowano uzyskanie od 80 do 100 milionów par odczytów 150 nukleotydowych fragmentów przypadających na pojedynczą próbkę gleby. Otrzymane surowe dane zostaną przekonwertowane do formatu pozwalającego na ich analizę bioinformatyczną. Analiza danych, prowadzona w celu wyłonienia optymalnego zestawu markerów, który umożliwi ocenę składu mikrobiomu w próbce nieznanego pochodzenia i pozwoli na przyporządkowanie jej do określonej lokalizacji, będzie stanowiła piąty etap projektu. Autorzy projektu planują stworzyć panel przeznaczony do identyfikacji DNA gleby, zawierający unikatowy zestaw wysoce informatywnych markerów umożliwiający porównywanie próbek gleby. Szóstym etapem jest opracowanie, optymalizacja i walidacja celowanej metody NGS umożliwiającej analizę mikrobiomu gleby, a ściślej opracowanie testu genetycznego opartego na celowanej metodzie sekwencjonowania następnej generacji (NGS, ang. *Next Generation Sequencing*). Testowanie próbek będzie się opierać na wyselekcjonowanych sekwencjach genomowych zdefiniowanych na etapie piątym. Sprawdzona zostanie także kompatybilność wyników sekwencjonowania próbek gleby otrzymanych na etapie czwartym (głębokie sekwencjonowanie) z wynikami sekwencjonowania uzyskanymi dla opracowanego na tym etapie testu

(sekwencjonowanie celowane) przeprowadzonego za pomocą kilku metod sekwencjonowania o średniej przepustowości, co w efekcie doprowadzi do wyłonienia optymalnej technologii. Wybrana metoda zostanie poddana walidacji, ze szczególnym uwzględnieniem specyfiki i ograniczeń występujących w analizach kryminalistycznych. W kolejnym (siódmym) etapie autorzy projektu zaplanowali stworzenie systemu informatycznego do analizy i interpretacji wyników uzyskiwanych za pomocą testu genetycznego i wybranej technologii NGS. Dane uzyskane z sekwencjonowania DNA mikrobiomów gleby zostaną umieszczone w bazie danych będącej elementem tworzonego systemu, stanowiąc „mapę” mikrobiomów gleb Polski. Dodatkowo w powstającym systemie zostanie zaimplementowane narzędzie do efektywnego przeszukiwania zasobów bazy i interpretacji wyników analiz. Uzyskiwane za pomocą testu wyniki sekwencjonowania DNA wyizolowanego z próbki gleby o nieznanym pochodzeniu dzięki stworzonemu oprogramowaniu będą porównywane z bazą i przypisywane do najbardziej prawdopodobnej lokalizacji na mapie Polski. Ostatecznie powstanie kompletny system predykcyjny zawierający test identyfikujący społeczności bakteryjne mikrobiomu gleby i oprogramowanie pozwalające na interpretację otrzymanych za pomocą testu danych. Ostatni etap projektu będzie obejmował przetestowanie skuteczności działania stworzonego programu przy zachowaniu warunków zbliżonych do rzeczywistych oraz przygotowanie procedur SOP (ang. Standard Operating Procedure) umożliwiających korzystanie z opracowanego systemu predykcyjnego. Przeprowadzona zostanie także ocena parametrów i poprawności działania programu. Dodatkowo opracowane zostaną wytyczne, procedury i instrukcje niezbędne do prowadzenia badań predykcyjnych i analiz porównawczych próbek DNA wyizolowanych z gleby, a wykonywanych dzięki użyciu stworzonego w projekcie SMAFT systemu.

Produktem końcowym projektu SMAFT będzie kompletny system predykcyjny przeznaczony do identyfikacji i ustalania miejsca pochodzenia próbek gleby na podstawie składu ich mikrobiomu.

Należy podkreślić, że według najlepszej wiedzy autorów w żadnej z opublikowanych dotąd prac naukowych dotyczących badań nad mikrobiomem gleby nie zastosowano tak głębokiego sekwencjonowania tak wielu izolatów DNA pochodzących z różnych próbek gleby. Przeprowadzone badania przyczynią się także do wzbogacenia wiedzy dotyczącej bioróżnorodności gleb w różnych częściach Polski.

Bibliografia

1. Ambers, A.D., Churchill, J.D., King, J.L., Stoljarova, M., Gill-King, H., Assidi, M., Abu-Elmagd, M., Buhmeida, A., Al-Qahtani, M., Budowle, B. (2016). More comprehensive forensic genetic marker analyses for accurate human remains identification using massively parallel DNA sequencing. *BMC Genomics*, 17, doi: 10.1186/s12864-016-3087-2.
2. Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G.H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J.A., Maguin, E., Mauchline, T., McClure, R., Mitter, B., Ryan, M., Sarand, I., Smidt, H., Schelkle, B., Roume, H., Kiran, G.S., Selvin, J., Souza, R.S.C., van Overbeek, L., Singh, B.K., Wagner, M., Walsh, A., Sessitsch, A., Schlöter, M. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, 8(1), doi:10.1186/s40168-020-00875-0.
3. Bruijns, B., Tiggelaar, R., Gardeniers, H. (2018). Massively parallel sequencing techniques for forensics: A review. *Electrophoresis*, 39, doi: doi.org/10.1002/elps.201800082.
4. Clarke, T.H., Gomez, A., Singh, H., Nelson, K.E., Brinkac, L.M. (2017). Integrating the microbiome as a resource in the forensics toolkit. *Forensic Science International: Genetics*, 30, doi: 10.1016/j.fsigen.2017.06.008.
5. Concheri, G., Bertoldi, D., Polone, E., Otto, S., Larcher, R., Squartini, A. (2011). Chemical elemental distribution and soil DNA fingerprints provide the critical evidence in murder case investigation. *PLOS ONE*, 6(6), doi: 10.1371/journal.pone.0020222.
6. Damaso, N., Mendel, J., Mendoza, M., von Wettberg, E.J., Narasimhan, G., Mills, D. (2018). Bioinformatics approach to assess the biogeographical patterns of soil communities: The utility for soil provenance. *Journal of Forensic Sciences*, 63(4), doi: 10.1111/1556-4029.13741.
7. Dawson, L.A., Hillier, S. (2010). Measurement of soil characteristics for forensic applications. *Surface and Interface Analysis*, 42(5), doi: 10.1002/sia.3315.
8. Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, 15, doi: 10.1038/nrmicro.2017.87.
9. Fierer, N., Hamady, M., Lauber, C.L., Knight, R. (2008). The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proceedings of the National Academy of Sciences*, 105(46), doi: 10.1073/pnas.0807920105.
10. Fierer, N., Lauber, C.L., Zhou, N., McDonald, D., Costello, E.K., Knight, R. (2010). Forensic identification using skin bacterial communities. *Proceedings of the National Academy of Sciences*, 107, doi: 10.1073/pnas.1000162107.
11. Fitzpatrick, R.W., Raven, M.D., Forrester, S.T. (2009). A systematic approach to soil forensics: Criminal case studies involving transference from crime scene to forensic evidence. W: K. Ritz, L. Dawson, D. Miller (red.), *Criminal and Environmental Soil Forensics*. Dordrecht: Springer.

12. Fitzpatrick, R., Raven, M., Self, P. (2017). The role of pedology and mineralogy in providing evidence for 5 crime investigations involving a wide range of earth materials. *Episodes*, 40, doi: 10.18814/epiugs/2017/v40i2/017017.
13. Giampaoli, S., Berti, A., Di Maggio, R.M., Pilli, E., Valentini, A., Valeriani, F., Gianfranceschi, G., Barni, F., Ripani, L., Romano Spica, V. (2014). The environmental biological signature: NGS profiling for forensic comparison of soils. *Forensic Science International*, 240, doi: 10.1016/j.forsciint.2014.02.028.
14. Habtom, H., Demanèche, S., Dawson, L., Azulay, C., Matan, O., Robe, P., Gafny, R., Simonet, P., Jurkevitch, E., Pasternak, Z. (2017). Soil characterisation by bacterial community analysis for forensic applications: A quantitative comparison of environmental technologies. *Forensic Science International: Genetics*, 26, doi: 10.1016/j.fsigen.2016.10.005.
15. Habtom, H., Pasternak, Z., Matan, O., Azulay, C., Gafny, R., Jurkevitch, E. (2019). Applying microbial biogeography in soil forensics. *Forensic Science International: Genetics*, 38, doi:10.1016/j.fsigen.2018.11.010.
16. Hampton-Marcell, J.T., Larsen, P., Anton, T., Cralle, L., Sangwan, N., Lax, S., Gottel, N., Salas-Garcia, M., Young, C., Duncan, G., Lopez, J.V., Gilbert, J.A., (2020). Detecting personal microbiota signatures at artificial crime scenes. *Forensic Science International*, 313, doi 10.1016/j.forsciint.2020.110351.
17. Heather, J.M., Chain, B. (2017). The sequence of sequencers: The history of sequencing DNA. *Genomics*, 107(1), doi: 10.1016/j.ygeno.2015.11.003.
18. Jøhll, M.E. (2009). *Investigating Chemistry. A Forensic Science Perspective*, wyd. 2. New York: W.H. Freeman and Company.
19. Kayser, M. (2015). Forensic DNA phenotyping: Predicting human appearance from crime scene material for investigative purposes. *Forensic Science International: Genetics*, 18, doi: 10.1016/j.fsigen.2015.02.003.
20. Khodadadian, A., Darzi, S., Haghi-Daredeh, S., Sadat Eshaghi, F., Babakhanzadeh, E., Mirabutalebi, S.H., Nazari, M. (2020). Genomics and transcriptomics: The powerful technologies in precision medicine. *International Journal of General Medicine*, 13, doi: 10.2147/IJGM.S249970.
21. Lax, S., Smith, D.P., Hampton-Marcell, J., Owens, S.M., Handley, K.M., Scott, N.M., Gibbons, S.M., Larsen, P., Shogan, B.D., Weiss, S., Metcalf, J.L., Ursell, L.K., Vázquez-Baeza, Y., Van Treuren, W., Hasan, N.A., Gibson, M.K., Colwell, R., Dantas, G., Knight, R., Gilbert, J.A. (2014). Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science*, 345(6200), doi: 10.1126/science.1254529.
22. Lax, S., Hampton-Marcell, J.T., Gibbons, S.M., Colares, G.B., Smith, D., Eisen, J.A., Gilbert, J.A. (2015). Forensic analysis of the microbiome of phones and shoes. *Microbiome*, 3, doi: 10.1186/s40168-015-0082-9.
23. Lee, S.Y., Woo, S.K., Lee, S.M., Eom, Y.B. (2016). Forensic analysis using microbial community between skin bacteria and fabrics. *Toxicology and Environmental Health Sciences*, 8(3), doi: 10.1007/s13530-016-0284-y.
24. Maron, P.A., Mougél, C., Ranjard L. (2011). Soil microbial diversity: Methodological strategy, spatial overview and functional interest. *Comptes Rendus Biologies*, 334(5–6), doi: 10.1016/j.crv.2010.12.003.
25. Meadow, J.F., Altrichter, A.E., Bateman, A.C., Stenson, J., Brown, G.Z., Green, J.L., Bohannan, B.J.M. (2015). Humans differ in their personal microbial cloud. *PeerJ Life & Environment*, 3, doi: 10.7717/peerj.1258.
26. Murray, R.C. (2004). Forensic geology: Yesterday, today and tomorrow. *Geological Society, Special Publications*, 232(1), doi: 10.1144/GSL.SP.2004.232.01.02.
27. Murray, R.C. (2012). Forensic examination of soils. W: L. Kobilinsky (red.), *Forensic Chemistry Handbook*. Hoboken, New Jersey: John Wiley & Sons, doi: 10.1002/9781118062241.ch4.
28. Neckovic, A., van Oorschot, R.A.H., Szkuta, B., Durdle, A. (2020). Investigation of direct and indirect transfer of microbiomes between individuals. *Forensic Science International: Genetics*, 45, doi: 10.1016/j.fsigen.2019.102212.
29. Needelman, B.A. (2013) What are soils? *Natural Education Knowledge*, 4(3).
30. O'Brien, S.L., Gibbons, S.M., Owens, S.M., Hampton-Marcell, J., Johnston, E.R., Jastrow, J.D., Gilbert, J.A., Meyer, F., Antonopoulos, D.A. (2016). Spatial scale drives patterns in soil bacterial diversity. *Environmental Microbiology*, 18(6), doi: 10.1111/1462-2920.13231.
31. Pasternak, Z., Al-Ashhab, A., Gatica, J., Gafny, R., Avraham, S., Minz, D., Gillor, O., Jurkevitch, E. (2013). Spatial and temporal biogeography of soil microbial communities in arid and semiarid regions. *PLOS ONE*, 8(7), doi: 10.1371/journal.pone.0069705.
32. Petraco, N., Kubic, T.A., Petraco, N.D.K. (2008). Case studies in forensic soil examinations. *Forensic Science International*, 178(2–3), doi: 10.1016/j.forsciint.2008.03.008.
33. Pirrie, D., Dawson, L., Graham, G. (2017). Predictive geolocation: Forensic soil analysis for provenance determination. *Episodes*, 40(2), doi: 10.18814/epiugs/2017/v40i2/017016.
34. Pye, K. (2004). Forensic geology. W: R.C. Selley, L.R.M. Cocks, I.R. Plimer (red.), *Encyclopedia of Geology*, t. 2. Amsterdam: Elsevier.
35. Ravel, J., Gajer, P., Abdo, Z., Schneider M.G., Koenig, S.S.K., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O., Brotman, R.M., Davis, C.C., Ault, K., Peralta, L., Forney, L.J. (2011). Vaginal microbiome of reproductive-age women.

- Proceedings of the National Academy of Sciences*, 108 (Supplement 1), doi: 10.1073/pnas.1002611107.
36. Robinson, J.M., Pasternak, Z., Mason, C.E., Elhaik, E. (2021). Forensic applications of microbiomics: A review. *Frontiers in Microbiology*, 11, doi: org/10.3389/fmicb.2020.608101.
37. Ruffell, A., McKinley, J. (2005). Forensic geoscience: Applications of geology, geomorphology and geophysics to criminal investigations. *Earth-Science Reviews*, 69, doi: 10.1016/j.earscirev.2004.08.002.
38. Schmedes, S.E., Woerner, A.E., Novroski, N.M.M., Wendt, F.R., King, J.L., Stephens, K.M., Budowle, B. (2018). Targeted sequencing of clade-specific markers from skin microbiomes for forensic human identification. *Forensic Science International: Genetics*, 32, doi: 10.1016/j.fsigen.2017.10.004.
39. Sensabaugh, G.F. (2009). Microbial community profiling for the characterisation of soil evidence: Forensic considerations. W: K. Ritz, L. Dawson, D. Miller (red.), *Criminal and Environmental Soil Forensics*. Dordrecht: Springer, doi: 10.1007/978-1-4020-9204-6_4.
40. Ślósarek, G. (2012). Współczesna rewolucja naukowa na pograniczu fizyki i biologii. *Zagadnienia Filozoficzne w Nauce (Philosophical Problems in Science)*, 51.
41. Trevors, J.T. (2010). One gram of soil: A microbial biochemical gene library. *Antonie van Leeuwenhoek*, 97, doi: 10.1007/s10482-009-9397-5.
42. Tridico, S.R., Murray, D.C., Addison, J., Kirkbride, K.P., Bunce, M. (2014). Metagenomic analyses of bacteria on human hairs: A qualitative assessment for applications in forensic science. *Investigative Genetics*, 5(1), doi: 10.1186/s13323-014-0016-5.
43. Velsko, S.P. (2020). Chapter 24 – Inferential validation and evidence interpretation. W: B. Budowle, S.E. Schutzer, S.A. Morse (red.), *Microbial Forensics*, wyd. 3. San Diego: Academic Press, doi: 10.1016/B978-0-12-815379-6.00024-6.
44. Vidaki, A., Kayser, M. (2018). Recent progress, methods and perspectives in forensic epigenetics. *Forensic Science International: Genetics*, 37, doi: 10.1016/j.fsigen.2018.08.008.
45. Whipps, J.M., Lewis, K., Cooke, R.C. (1988). Myco-parasitism and plant disease control. W: M.N. Burge (red.), *Fungi in Biological Control Systems*. Manchester: Manchester University Press.
46. Woods, B., Lennard, C., Kirkbride, K.P., Robertson, J. (2016). Soil examination for a forensic trace evidence laboratory – Part 3: A proposed protocol for the effective triage and management of soil examinations. *Forensic Science International*, 262, doi: 10.1016/j.forsciint.2016.02.034.
47. Young, J.M., Austin, J.J., Weyrich, L.S. (2017). Soil DNA metabarcoding and high-throughput sequencing as a forensic tool: Considerations, potential limitations and recommendations, *FEMS Microbiology Ecology*, 93(2), doi: 10.1093/femsec/fiw207.
48. Young, J.M., Higgins, D., Austin, J.J. (2019). Soil DNA: Advances in DNA technology offer a powerful new tool for forensic science. W: R.W. Fitzpatrick, L.J. Donnelly (red.), *Forensic Soil Science and Geology*. London: Geological Society, Special Publications, doi: 10.1144/SP492-2017-351.
49. Zbieć-Piekarska, R., Spólnicka, M., Kupiec, T., Parys-Proszek, A., Makowska, Ż., Pałeczka, A., Kucharczyk, K., Płoski, R., Branicki, W. (2015). Development of a forensically useful age prediction method based on DNA methylation analysis. *Forensic Science International: Genetics*, 17, doi: 10.1016/j.fsigen.2015.05.001.

Harnessing the potential of the environmental microbiome in forensic science

Agata Jagiełło¹, Anna Woźniak¹, Błażej Szczerba², Rafał Płoski³, Małgorzata Rydzanicz³, Agnieszka Pollak³, Alina Frolova^{4,5}, Michał Kowalski⁵, Paweł Łabaj⁵, Andrzej Ossowski⁶, Wojciech Branicki^{1,5}, Kinga Herda⁵, Kamila Marszałek⁵, Renata Zbieć-Piekarska^{1*}

¹ Central Forensic Laboratory of the Police

² Ardigen SA

³ Medical University of Warsaw

⁴ Institute of Molecular Biology and Genetics of National Academy of Sciences of Ukraine, Kyiv, Ukraine

⁵ Jagiellonian University

⁶ Pomeranian Medical University in Szczecin

* Corresponding author: renata.zbiec@policja.gov.pl

Summary

A scientific consortium led by the Central Forensic Laboratory of the Police has undertaken to develop a method for DNA analysis of the soil microbiome to be used in forensic investigations. The aim of the project entitled Soil Microbiome Analysis Forensic Tool – SMAFT (<http://smaft.eu/>), financed by the National Center for Research and Development (DOB-BIO10/03/01/2019), is to develop a new tool that enables the association of a trace in the form of a soil sample with a specific geographical location. The first part of the paper introduces the concept of the microbiome and presents the possibilities of using microbiome DNA analysis in forensic science. In the second part, the stages of the SMAFT project are described in detail, beginning from the collection of soil samples from different sites in Poland across all seasons and isolation of microbiome DNA through massively parallel sequencing (MPS) technology-based analysis of isolates and the development of a genetic test containing a set of metagenomic markers allowing for effective individualization of soil samples, up to the creation of an IT system enabling analysis and interpretation of the obtained results, which includes a database of soil microbiome DNA profiles from various locations in Poland.

Key words: microbiome, soil, forensics, metagenomics, metataxonomic analysis, MPS sequencing, bioinformatics, NCBiR project

Introduction

The Central Forensic Laboratory of the Police (CFLP), as a research institute, conducts scientific research in the field of forensic science. The projects carried out at the CFLP are aimed at developing methods and technologies to support the prevention, detection and combating of crime. One of the major issues of interest to forensic genetics in recent years is the use of possibilities offered by forensic genomics (genomics of individual differences), epigenomics, metagenomics, and metataxonomic analysis to design and develop innovative tools intended for use in forensic police laboratories. This is in line with the priorities set by the member states of the North Atlantic Treaty Organization (NATO), particularly by the Allied Command Transformation, the NATO Science and Technology Organization and the European Defence Agency, recognizing genomics as one of twenty breakthrough technologies whose development is critical to ensuring security.

Unlike genetics, which mainly involves the study of genes and the mechanisms of their inheritance, genomics is a relatively new field of biology that addresses the analysis of the genome (complete genetic information of an organism). Genomics has emerged due to the intensive development of computer science and molecular biology techniques, especially modern sequencing technologies such as massively parallel sequencing (MPS), which represents one of the most important advances in biological sciences in the last two decades (Heather, Chain, 2017). Structural genomics involves determining the sequence of an organism's genome and its organization, whilst identifying the regions corresponding to individual genes. The role of functional genomics, together with epigenomics (DNA modifications), transcriptomics (mRNA), proteomics (proteins), and metabolomics (metabolites), is to determine the functions of genes and non-coding regions, the mechanisms of their regulation, and to study the gene-gene and gene-environment

interactions, including the analysis of the effects of gene expression products on the functioning of cells, tissues, organs, and whole organisms (Khodadadian et al., 2020). Theoretical genomics, comparative genomics and genomics of individual differences analyze general rules applicable to genes, the evolution of genes and the individual variation of genomes, respectively (Ślósarek, 2012). The last is the most interesting from a forensic point of view.

The use of genomics combined with MPS technology in forensic investigations allows analyses that until recently were too expensive or even unfeasible due to technical limitations. Nowadays it is possible to simultaneously test short tandem repeat (STR) and single nucleotide polymorphism (SNP) markers of autosomal DNA, X and Y sex chromosome markers as well as mitochondrial DNA (mtDNA). These methods, in addition to the forensic identification data of a person, allow to obtain information on phenotypic traits and biogeographical origin, even in the case of small amounts or poor quality DNA found at the crime scene (Kayser, 2015; Ambers et al., 2016; Bruijns, Tiggelaar, Gardeniers, 2018). Additionally, epigenetic markers can be used in the analysis of methylation profiles of human genomes and may contribute to a relatively accurate determination of the age of an unknown individual, distinguish between monozygotic twins, and identify tissues and body fluids (Zbieć-Piekarska et al., 2015; Vidaki, Kayser, 2018).

The potential of the microbiome in forensic science

Regardless of ongoing improvements in the area of identification methods, a stand-alone application of genomics is not always sufficient to solve a crime. A review of recent research papers indicates that in such situations certain questions can be answered with the aid of metagenomics, a field of science that involves the study of microbiome DNA, i.e. DNA recovered directly from environmental samples, without the need for establishing laboratory cultures.

Although scientists have been studying microbiomes for several decades, until recently no clear, universal definition of the microbiome was available. For this reason, in March 2019, the Microbiome Support project (<https://www.microbiomesupport.eu/>) hosted a combined workshop meeting of a large international group of microbiome experts to discuss redefinition of this term. In the following year, Berg et al. (2020) presented a proposal developed at that time in a comprehensive commentary published as a summary of the meeting. It is based on the definition from 1988 (Whipps, Lewis, Cooke, 1988), which describes the microbiome as a community of microorganisms (microbiota) inhabiting a well-defined environment with characteristic physicochemical features. As noted by Berg et al. (2020), this definition takes into account, in addition to a microbial community with distinct characteristic and functions, its interactions with the environment

resulting in the formation of specific ecological niches. The microbiome, which is a dynamic and interactive microecosystem that undergoes changes in time and space, remains integrated with macroecosystems, such as the eukaryotic host, affecting its functioning and condition. It should be emphasized that the microbiome comprises the microbiota (i.e. living microorganisms of different taxa) and their numerous activities (their “theatre of activity”) encompassing the whole spectrum of molecules produced by the microorganisms, including their structural elements (nucleic acids, proteins, lipids, polysaccharides), metabolites (signaling molecules, toxins, organic and inorganic molecules) and molecules formed by various environmental conditions in the surrounding environment, including those produced by coexisting hosts, as well as any mobile genetic elements, e.g. phages, viruses, plasmids, transposons, integrons and extracellular DNA, including relic DNA from dead cells (Berg et al., 2020).

Every human being, like every other living organism or every particular place, such as a meadow and its soil or a lake and its water, represent unique habitats determining to a large extent the diversity and quantitative proportions of the indigenous microorganisms. Sequencing and/or metataxonomic analysis of the microbiome DNA from a specific site allows the identification of its unique profile, which can be compared to microbiome profiles from other habitats. The results of research performed over the past few decades indicate that in the future the use of human microbiome DNA from various parts of the human body will be possible for forensic purposes (Fierer et al., 2008; Ravel et al., 2011; Tridico et al., 2014; Schmedes et al., 2018). Humans leave their microbial “footprint” in places they visit, such as crime scenes (Hampton-Marcell et al., 2020), on objects they touch (Lax et al., 2014), such as cell phones (Fierer et al., 2010), clothes (Lax et al., 2015), and fabrics (Lee et al., 2016), and on other individuals they come in contact with (Neckovic et al., 2020). It has also been proven that humans are accompanied by specific microbial clouds whose composition can be potentially utilized for the purpose of identification (Meadow et al., 2015). Furthermore, as noted by Clarke et al. (2017), sequencing the microbiome of an individual can provide data sufficient not only for identification, but also to obtain information about the individual’s gender, health, and lifestyle, which is very important from a law enforcement perspective.

In parallel to the research conducted on the human microbiome in a forensic context, many research centers are developing methods and tools that enable the use of the soil microbiome for investigative and forensic DNA intelligence purposes. For example, between 2013 and 2015, a Microbial Soil Analysis (MiSAFE) project was implemented within the framework of the European Union’s action for security in Europe, dedicated to the development of tools and procedures for routine testing of soil samples in forensic laboratories

(<https://forensicmisafe.wixsite.com/misafe/project>). It is well known that soil represents particularly valuable evidence due to its omnipresence in the environment, its diversity as well as the ability to adhere to shoes, tires, tools or clothes. It is also often overlooked by suspects attempting to obliterate their traces (Young, Austin, Weyrich, 2017). Therefore, the analysis of soil samples can provide sufficient information to link a suspect, victim, or object to a crime scene (Johll, 2009; Dawson, Hillier, 2010; Concheri et al., 2011), as well as to infer the likely geographic location of origin of the soil sample. (Pirrie, Dawson, Graham, 2017). Soil is a complex mixture of minerals, organic matter including living organisms, gases, and water (Needelman, 2013). Currently, forensic soil analysis is based on determining physical characteristics and chemical composition. The analysis includes soil color, texture, particle size, pH, elemental composition, mineral content, and sometimes organic compounds such as plant waxes (Habtom et al., 2017; Murray, 2012; Woods et al., 2016), which provide information relevant to the investigation (Fitzpatrick, Raven, Self, 2017; Petraco, Kubic, Petraco, 2008). However, in particular cases, e.g. when the soil samples come from a geologically homogeneous area, from adjacent sites, or have low inorganic content (e.g., peat soils), routine analysis is unfeasible for discrimination between samples (Giampaoli et al., 2014; Young, Austin, Weyrich, 2017; Young, Higgins, Austin, 2019). The limitations of differentiating soil samples at the local scale can be overcome by performing analyses of biological material. It is estimated that 1 gram of dry soil contains on the average: 10^{10} viruses, 10^{10} bacteria and archaeons (including 10^8 actinomycetes), 10^6 each of fungi and algae, 10^5 protozoa, and 10^2 nematodes (Trevors, 2010). Additionally, soil may also contain plant fragments (e.g. roots, pollen, spores, seeds, leaves) and invertebrates other than nematodes (Young, Austin, Weyrich, 2017) as well as extracellular DNA. Because all soil-dwelling organisms have specific habitat requirements, environmental conditions, such as soil type and texture, pH, moisture, temperature and organic carbon levels significantly influence the composition of the microbiome, i.e. the community of microorganisms living at a specific site (Maron, Mougel, Ranjard, 2011; Pasternak et al., 2013). Due to a variable spatial structure of the soil, it does not contain a "typical", uniform microbiome (Fierer, 2017). Studies of the relative abundances of major bacterial taxa and archaeons in soil samples have demonstrated diversity not only for different soil types, but also for soils collected from sites only several or even a few centimeters away (Habtom et al., 2019; O'Brien et al., 2016; Pasternak et al., 2013; Sensabaugh, 2009). DNA analysis of the soil microbiome may provide data to enable its effective individualization and, in the future, prove as useful for comparing soil traces or determining the soil origin, as human DNA profiles for establishing a link between biological traces and the offender (Damaso et al., 2018).

It is important to note that currently the analysis of microbiome DNA is practically absent in the investigation, although the potential of this type of data as evidence or aid in conventional forensic testing methods has been increasingly recommended (Robinson et al., 2021). In order for the potential of the microbiome to be utilized effectively by law enforcement, further research is needed to demonstrate that statistical inference based on such analyses is stringent enough to be considered by courts as scientific evidence as being characterized, for instance, by a known and accepted error rate (Velsko, 2020; Robinson et al., 2021). Expanding the set of samples to be analyzed in ongoing studies, creating databases of microbiomes from diverse environments, based on clearly defined and well-documented procedures, improving bioinformatics tools, including machine learning techniques for interpreting results, or studying the dynamics of temporal and spatial changes in the microbiomes, are examples of research goals that, if achieved, may contribute to including microbiome analysis in the toolbox used in forensic laboratories (Robinson et al., 2021).

SMAFT project

Bearing in mind the challenges outlined above, the Central Forensic Laboratory of the Police (CFLP) as a leader representing a consortium consisting of: Medical University of Warsaw, Jagiellonian University, Pomeranian Medical University and ARDIGEN company, has received funding from the National Center for Research and Development to implement the **Soil Microbiome Analysis Forensic Tool – SMAFT** (<http://smaft.eu/>) project aimed at using the potential of the soil microbiome in forensic science (DOB-BIO10/03/01/2019). The main objective of the project is to develop a predictive tool for forensic DNA analysis of the soil microbiome, which will allow profiling the geographical location of soil samples of unknown origin in Poland. In other words, the system under development will seek to determine the possibility of establishing the link between a soil trace recovered from a shoe, tire or shovel and a specific geographic location. This type of evidence is likely to link the suspect with a particular location with a higher degree of probability, or it can allow to trace the movement of the criminal offender. In the future, the information from the SMAFT system will aid to direct and speed up both criminal and terrorist investigations. The system under development also has the potential to be used to prosecute environmental crimes or, more broadly, to conduct biodiversity research.

The research planned under the SMAFT project is divided into several stages, beginning with the collection of soil samples. The project involves collecting close to 1,000 soil samples, 250 each in the fall, winter, spring and summer, from 80 different locations across Poland. According to the authors, such an approach allows to detect possible seasonally independent differences in

topsoil microbiome composition. The selection of soil sampling sites was based on computer analysis of data from measurements taken at all hydrological and meteorological stations located throughout Poland, covering the period of the last 20 years. The map of Poland has been divided into five areas with different climatic conditions. Twelve main sampling locations were eventually selected, and several (five to eight) sampling sites were designated around each location. Three soil samples would be collected from each specific sampling site. In addition, the physicochemical characteristics of soils in Poland, using data collected under the EU Land Use/Cover Area frame statistical Survey – LUCAS project (<https://ec.europa.eu/eurostat/web/lucas/>) were taken into account while selecting the sampling sites. This stage also includes developing a detailed methodology for collecting, preserving and labeling soil samples, designating and documenting the collection sites, as well as transporting samples to the laboratory under appropriate conditions. In the next stage of the project, microbiome DNA will be extracted from the collected samples. As soil is a difficult, heterogeneous material, containing a lot of substances potentially inhibiting the enzymes used in subsequent project activities, finding an isolation method that will yield soil microbiome DNA of sufficient quantity, purity and quality will determine the success of subsequent stages. The efficiency of DNA extraction from gram-positive, gram-negative, spore-forming, or envelope-producing bacteria varies depending on the procedure used, hence obtaining DNA from soil that is representative of the entire bacterial community of a given microbiome is not straightforward. In the third stage of the project, the obtained DNA isolates will be used to prepare libraries containing DNA fragments of microorganisms extracted from soil samples. In order to obtain the best quality libraries with the desired fragment lengths, optimization of the library construction procedure is planned, at both the DNA fragmentation and amplification stages. In the process of preparing the libraries, all DNA fragments from each sample will be assigned a unique barcode allowing – after obtaining sequencing data – their identification and assignment to a specific soil sample. The fragment lengths of the prepared libraries will be verified by capillary electrophoresis and the concentration of DNA within the libraries will be determined by a fluorometric assay. After calculating the molar concentration all the libraries will be normalized and pooled, taking into account unique index combinations. The prepared library pools will be sequenced in the next (fourth) stage using Illumina® SBS technology and the latest generation Illumina® NovaSeq 6000 sequencer. Between 80–100 million 150 bp paired-end reads per single soil sample will be generated. The resulting raw data will be converted to a format that allows bioinformatics analysis. The fifth stage will involve data analysis, in order to identify the optimal set of markers to evaluate

the microbiome composition of a sample of unknown origin and assign the sample to a specific location. The authors plan to design a soil DNA identification panel, containing a unique set of highly informative markers that enable comparison of soil samples. The sixth stage of the project provides for the development, optimization, and validation of a targeted NGS method to analyze the soil microbiome or, more specifically, the development of a genetic test based on next-generation sequencing (NGS). Sample testing will be performed using the selected genomic sequences defined in stage five. The compatibility of the soil sample sequencing results obtained in stage four (deep sequencing) with the sequencing results of the genetic test developed in stage five (targeted sequencing) will also be determined using several medium throughput sequencing methods, ultimately leading to selection of the optimal technology. The selected method will be validated, with particular requirements and limitations of forensic analyses. In the next (seventh) stage, the authors aim to create an IT system for analysis and interpretation of the results obtained by genetic tests and selected NGS technology. The data gathered from DNA sequencing of soil microbiomes will be uploaded to a database included in the IT system under development, thus creating a “map” of the soil microbiomes in Poland. Additionally, a tool for efficient database searching and interpreting analysis results will be implemented as part of the system. The results of sequencing DNA isolated from a soil sample of unknown origin obtained by the test will be compared with the database and assigned to the most probable location on the map of Poland. Eventually, a complete predictive system will be developed that includes a test for identification of bacterial communities of the soil microbiome and the software for interpretation of test data. The last stage of the project will involve testing the effectiveness of the predictive system in conditions mimicking real-life situations and preparation of Standard Operating Procedures (SOPs) enabling its implementation into practice. An evaluation of the system’s parameters and performance will also be conducted. Additionally, guidelines, procedures, and instructions necessary to conduct predictive testing and comparative analysis of DNA isolated from soil samples, performed using the system created under the SMAFT project will be developed.

As a result of the SMAFT project a complete predictive system designed to identify and determine the site of origin of soil samples based on the composition of the microbiome will be created.

It should be noted that, to the best of the authors’ knowledge, no scientific paper published to date on soil microbiome research has reported such deep sequencing of so many DNA isolates from different soil samples. The conducted research will additionally contribute to broadening the knowledge on soil biodiversity in various regions of Poland.

Bibliography

1. Ambers, A.D., Churchill, J.D., King, J.L., Stoljarova, M., Gill-King, H., Assidi, M., Abu-Elmagd, M., Buhmeida, A., Al-Qahtani, M., Budowle, B. (2016). More comprehensive forensic genetic marker analyses for accurate human remains identification using massively parallel DNA sequencing. *BMC Genomics*, *17*, doi: 10.1186/s12864-016-3087-2.
2. Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G.H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J.A., Maguin, E., Mauchline, T., McClure, R., Mitter, B., Ryan, M., Sarand, I., Smidt, H., Schelkle, B., Roume, H., Kiran, G.S., Selvin, J., Souza, R.S.C., van Overbeek, L., Singh, B.K., Wagner, M., Walsh, A., Sessitsch, A., Schloter, M. (2020). Microbiome definition re-visited: old concepts and new challenges. *Microbiome*, *8*(1), doi:10.1186/s40168-020-00875-0.
3. Bruijns, B., Tiggelaar, R., Gardeniers, H. (2018). Massively parallel sequencing techniques for forensics: A review. *Electrophoresis*, *39*, doi: doi.org/10.1002/elps.201800082.
4. Clarke, T.H., Gomez, A., Singh, H., Nelson, K.E., Brinkac, L.M. (2017). Integrating the microbiome as a resource in the forensics toolkit. *Forensic Science International: Genetics*, *30*, doi: 10.1016/j.fsigen.2017.06.008.
5. Concheri, G., Bertoldi, D., Polone, E., Otto, S., Larcher, R., Squartini, A. (2011). Chemical elemental distribution and soil DNA fingerprints provide the critical evidence in murder case investigation. *PLOS ONE*, *6*(6), doi: 10.1371/journal.pone.0020222.
6. Damaso, N., Mendel, J., Mendoza, M., von Wettberg, E.J., Narasimhan, G., Mills, D. (2018). Bioinformatics approach to assess the biogeographical patterns of soil communities: The utility for soil provenance. *Journal of Forensic Sciences*, *63*(4), doi: 10.1111/1556-4029.13741.
7. Dawson, L.A., Hillier, S. (2010). Measurement of soil characteristics for forensic applications. *Surface and Interface Analysis*, *42*(5), doi: 10.1002/sia.3315.
8. Fierer, N. (2017). Embracing the unknown: disentangling the complexities of the soil microbiome. *Nature Reviews Microbiology*, *15*, doi: 10.1038/nrmicro.2017.87.
9. Fierer, N., Hamady, M., Lauber, C.L., Knight, R. (2008). The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proceedings of the National Academy of Sciences*, *105*(46), doi: 10.1073/pnas.0807920105.
10. Fierer, N., Lauber, C.L., Zhou, N., McDonald, D., Costello, E.K., Knight, R. (2010). Forensic identification using skin bacterial communities. *Proceedings of the National Academy of Sciences*, *107*, doi: 10.1073/pnas.1000162107.
11. Fitzpatrick, R.W., Raven, M.D., Forrester, S.T. (2009). A systematic approach to soil forensics: Criminal case studies involving transference from crime scene to forensic evidence. In: K. Ritz, L. Dawson, D. Miller (ed.), *Criminal and Environmental Soil Forensics*. Dordrecht: Springer.
12. Fitzpatrick, R., Raven, M., Self, P. (2017). The role of pedology and mineralogy in providing evidence for 5 crime investigations involving a wide range of earth materials. *Episodes*, *40*, doi: 10.18814/epiugs/2017/v40i2/017017.
13. Giampaoli, S., Berti, A., Di Maggio, R.M., Pilli, E., Valentini, A., Valeriani, F., Gianfranceschi, G., Barni, F., Ripani, L., Romano Spica, V. (2014). The environmental biological signature: NGS profiling for forensic comparison of soils. *Forensic Science International*, *240*, doi: 10.1016/j.forsciint.2014.02.028.
14. Habtom, H., Demanèche, S., Dawson, L., Azulay, C., Matan, O., Robe, P., Gafny, R., Simonet, P., Jurkevitch, E., Pasternak, Z. (2017). Soil characterisation by bacterial community analysis for forensic applications: A quantitative comparison of environmental technologies. *Forensic Science International: Genetics*, *26*, doi: 10.1016/j.fsigen.2016.10.005.
15. Habtom, H., Pasternak, Z., Matan, O., Azulay, C., Gafny, R., Jurkevitch, E. (2019). Applying microbial biogeography in soil forensics. *Forensic Science International: Genetics*, *38*, doi:10.1016/j.fsigen.2018.11.010.
16. Hampton-Marcell, J.T., Larsen, P., Anton, T., Cralle, L., Sangwan, N., Lax, S., Gottel, N., Salas-Garcia, M., Young, C., Duncan, G., Lopez, J.V., Gilbert, J.A., (2020). Detecting personal microbiota signatures at artificial crime scenes. *Forensic Science International*, *313*, doi 10.1016/j.forsciint.2020.110351.
17. Heather, J.M., Chain, B. (2017). The sequence of sequencers: The history of sequencing DNA. *Genomics*, *107*(1), doi: 10.1016/j.ygeno.2015.11.003.
18. Jøhl, M.E. (2009). *Investigating Chemistry. A Forensic Science Perspective*, ed. 2. New York: W.H. Freeman and Company.
19. Kayser, M. (2015). Forensic DNA phenotyping: Predicting human appearance from crime scene material for investigative purposes. *Forensic Science International: Genetics*, *18*, doi: 10.1016/j.fsigen.2015.02.003.
20. Khodadadian, A., Darzi, S., Haghi-Daredeh, S., Sadat Eshaghi, F., Babakhanzadeh, E., Mirabutalebi, S.H., Nazari, M. (2020). Genomics and transcriptomics: The powerful technologies in precision medicine. *International Journal of General Medicine*, *13*, doi: 10.2147/IJGM.S249970.
21. Lax, S., Smith, D.P., Hampton-Marcell, J., Owens, S.M., Handley, K.M., Scott, N.M., Gibbons, S.M., Larsen, P., Shogan, B.D., Weiss, S., Metcalf, J.L., Ursell, L.K., Vázquez-Baeza, Y., Van Treuren, W., Hasan, N.A., Gibson, M.K., Colwell, R., Dantas, G., Knight, R., Gilbert, J.A. (2014). Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science*, *345*(6200), doi: 10.1126/science.1254529.

22. Lax, S., Hampton-Marcell, J.T., Gibbons, S.M., Colares, G.B., Smith, D., Eisen, J.A., Gilbert, J.A. (2015). Forensic analysis of the microbiome of phones and shoes. *Microbiome*, 3, doi: 10.1186/s40168-015-0082-9.
23. Lee, S.Y., Woo, S.K., Lee, S.M., Eom, Y.B. (2016). Forensic analysis using microbial community between skin bacteria and fabrics. *Toxicology and Environmental Health Sciences*, 8(3), doi: 10.1007/s13530-016-0284-y.
24. Maron, P.A., Mougél, C., Ranjard L. (2011). Soil microbial diversity: Methodological strategy, spatial overview and functional interest. *Comptes Rendus Biologies*, 334(5–6), doi: 10.1016/j.crv.2010.12.003.
25. Meadow, J.F., Altrichter, A.E., Bateman, A.C., Stenson, J., Brown, G.Z., Green, J.L., Bohannon, B.J.M. (2015). Humans differ in their personal microbial cloud. *PeerJ Life & Environment*, 3, doi: 10.7717/peerj.1258.
26. Murray, R.C. (2004). Forensic geology: Yesterday, today and tomorrow. *Geological Society, Special Publications*, 232(1), doi: 10.1144/GSL.SP.2004.232.01.02.
27. Murray, R.C. (2012). Forensic examination of soils. In: L. Kobilinsky (ed.), *Forensic Chemistry Handbook*. Hoboken, New Jersey: John Wiley & Sons, doi: 10.1002/9781118062241.ch4.
28. Neckovic, A., van Oorschot, R.A.H., Szkuta, B., Durdle, A. (2020). Investigation of direct and indirect transfer of microbiomes between individuals. *Forensic Science International: Genetics*, 45, doi: 10.1016/j.fsigen.2019.102212.
29. Needelman, B.A. (2013) What are soils? *Natural Education Knowledge*, 4(3).
30. O'Brien, S.L., Gibbons, S.M., Owens, S.M., Hampton-Marcell, J., Johnston, E.R., Jastrow, J.D., Gilbert, J.A., Meyer, F., Antonopoulos, D.A. (2016). Spatial scale drives patterns in soil bacterial diversity. *Environmental Microbiology*, 18(6), doi: 10.1111/1462-2920.13231.
31. Pasternak, Z., Al-Ashhab, A., Gatica, J., Gafny, R., Avraham, S., Minz, D., Gillor, O., Jurkevitch, E. (2013). Spatial and temporal biogeography of soil microbial communities in arid and semiarid regions. *PLOS ONE*, 8(7), doi: 10.1371/journal.pone.0069705.
32. Petraco, N., Kubic, T.A., Petraco, N.D.K. (2008). Case studies in forensic soil examinations. *Forensic Science International*, 178(2–3), doi: 10.1016/j.forsciint.2008.03.008.
33. Pirrie, D., Dawson, L., Graham, G. (2017). Predictive geolocation: Forensic soil analysis for provenance determination. *Episodes*, 40(2), doi: 10.18814/epiugs/2017/v40i2/017016.
34. Pye, K. (2004). Forensic geology. In: R.C. Selley, L.R.M. Cocks, I.R. Plimer (ed.), *Encyclopedia of Geology*, vol. 2. Amsterdam: Elsevier.
35. Ravel, J., Gajer, P., Abdo, Z., Schneider M.G., Koenig, S.S.K., McCulle, S.L., Karlebach, S., Gorle, R., Russell, J., Tacket, C.O., Brotman, R.M., Davis, C.C., Ault, K., Peralta, L., Forney, L.J. (2011). Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences*, 108 (Supplement 1), doi: 10.1073/pnas.1002611107.
36. Robinson, J.M., Pasternak, Z., Mason, C.E., Elhaik, E. (2021). Forensic applications of microbiomics: A review. *Frontiers in Microbiology*, 11, doi.org/10.3389/fmicb.2020.608101.
37. Ruffell, A., McKinley, J. (2005). Forensic geoscience: Applications of geology, geomorphology and geophysics to criminal investigations. *Earth-Science Reviews*, 69, doi: 10.1016/j.earscirev.2004.08.002.
38. Schmedes, S.E., Woerner, A.E., Novroski, N.M.M., Wendt, F.R., King, J.L., Stephens, K.M., Budowle, B. (2018). Targeted sequencing of clade-specific markers from skin microbiomes for forensic human identification. *Forensic Science International: Genetics*, 32, doi: 10.1016/j.fsigen.2017.10.004.
39. Sensabaugh, G.F. (2009). Microbial community profiling for the characterisation of soil evidence: Forensic considerations. In: K. Ritz, L. Dawson, D. Miller (ed.), *Criminal and Environmental Soil Forensics*. Dordrecht: Springer, doi: 10.1007/978-1-4020-9204-6_4.
40. Ślósarek, G. (2012). Współczesna rewolucja naukowa na pograniczu fizyki i biologii. *Zagadnienia Filozoficzne w Nauce*, 51.
41. Trevors, J.T. (2010). One gram of soil: A microbial biochemical gene library. *Antonie van Leeuwenhoek*, 97, doi: 10.1007/s10482-009-9397-5.
42. Tridico, S.R., Murray, D.C., Addison, J., Kirkbride, K.P., Bunce, M. (2014). Metagenomic analyses of bacteria on human hairs: A qualitative assessment for applications in forensic science. *Investigative Genetics*, 5(1), doi: 10.1186/s13323-014-0016-5.
43. Velsko, S.P. (2020). Chapter 24 – Inferential validation and evidence interpretation. In: B. Budowle, S.E. Schutzer, S.A. Morse (ed.), *Microbial Forensics*, ed. 3. San Diego: Academic Press, doi: 10.1016/B978-0-12-815379-6.00024-6.
44. Vidaki, A., Kayser, M. (2018). Recent progress, methods and perspectives in forensic epigenetics. *Forensic Science International: Genetics*, 37, doi: 10.1016/j.fsigen.2018.08.008.
45. Whipps, J.M., Lewis, K., Cooke, R.C. (1988). Mycoparasitism and plant disease control. In: M.N. Burge (red.), *Fungi in Biological Control Systems*. Manchester: Manchester University Press.
46. Woods, B., Lennard, C., Kirkbride, K.P., Robertson, J. (2016). Soil examination for a forensic trace evidence laboratory – Part 3: A proposed protocol for the effective triage and management of soil examinations. *Forensic Science International*, 262, doi: 10.1016/j.forsciint.2016.02.034.
47. Young, J.M., Austin, J.J., Weyrich, L.S. (2017). Soil DNA metabarcoding and high-throughput sequencing as a forensic tool: Considerations, potential limi-

- tations and recommendations, *FEMS Microbiology Ecology*, 93(2), doi: 10.1093/femsec/fiw207.
48. Young, J.M., Higgins, D., Austin, J.J. (2019). Soil DNA: Advances in DNA technology offer a powerful new tool for forensic science. In: R.W. Fitzpatrick, L.J. Donnelly (red.), *Forensic Soil Science and Geology*. London: Geological Society, Special Publications, doi: 10.1144/SP492-2017-351.
49. Zbieć-Piekarska, R., Spólnicka, M., Kupiec, T., Parys-Proszek, A., Makowska, Ż., Pałeczka, A., Kucharczyk, K., Płoski, R., Branicki, W. (2015). Development of a forensically useful age prediction method based on DNA methylation analysis. *Forensic Science International: Genetics*, 17, doi: 10.1016/j.fsigen.2015.05.001.

Translation Hanna Wierzchośławska, Ewa Bartnik