

Bazy DNA pozwalają zidentyfikować 60% białych w USA

15 października 2018

Ujęcie seryjnego mordercy nazwanego przez prasę Golden State Killer pokazało, jak przydatnymi narzędziami w rękach policji mogą być ogólnodostępne bazy danych genealogicznych, w których obywatele dobrowolnie pozostawiają próbki DNA. Teraz analizy przeprowadzone przez naukowców wykazały, że już za kilka lat – przynajmniej w USA – przed śledczymi nie ukryje się nikt, kto pozostawił na miejscu przestępstwa swój ślad genetyczny.



Jeśli jesteś biały, mieszkasz w USA i twój daleki krewny korzysta z serwisów genealogicznych, to istnieje olbrzymie prawdopodobieństwo, że policja zidentyfikuje cię dzięki kodowi genetycznemu. Badania przeprowadzone właśnie przez naukowców z Columbia University wykazały, że wykorzystując próbkę DNA daleko spokrewnionej osoby i z grubsza szacując wiek poszukiwanego przestępcy, policja już teraz jest w stanie zawęzić krąg podejrzanych do mniej niż 20 osób. A wszystko to dzięki bazie danych zawierających kod genetyczny zaledwie 1,3 miliona osób.

Wykorzystanie takiej metody pozwala w tej chwili na identyfikację około 60% białych Amerykanów, nawet jeśli oni sami nigdy nie udostępnili swojego DNA. „Za kilka lat będzie tak można zidentyfikować każdego”, mówi Yaniv Erlich, genetyk z Columbia University.

Naukowcy postanowili bliżej przyjrzeć się tej metodzie po tym, jak policja wykorzystała ją do złapania seryjnego mordercy, który działał w latach 70. i 80. DNA pozostawione przez niego na miejscach zbrodni zostało porównane z bazą danych GEDmatch. To bezpłatny serwis, zawierający dane o mniej niż milionie osób, do którego każdy może przesłać próbkę, a firmy takie jak 23andMe przeanalizują DNA i określą krąg spokrewnionych osób, które również przesłały próbki.

W całej tej sprawie najbardziej niezwykły jest fakt, że schwytany morderca był bardzo daleko spokrewniony z osobami, których DNA znajdowało się w bazie. Na tyle daleko, że najprawdopodobniej się nie znali. Otóż Golden State Killer i osoby, których DNA było w przeszukanej bazie mieli tych samych prapradziadków. Pradziadkowie byli inni, dziadkowie byli inni, rodzice też byli inni. Do tych informacji wystarczyło dodać szacowany wiek mordercy i miejsca popełnienia przestępstw, by znacząco zawęzić krąg poszukiwanych.

Naukowcy szybko zaczęli zastanawiać się, jak wiele osób o nieznanym DNA można zidentyfikować za pomocą takich serwisów. W tym celu Erlich i jego zespół spojrzeli się na bazę MyHeritage, która zawiera 1,28 miliona profili DNA. Analiza wykazała, że jeśli mieszkasz w USA, a twoi przodkowie pochodzą z Europy, to istnieje 60% szans, że osoba, z którą masz wspólnych prapradziadków lub jesteście bliżej spokrewnieni, dostarczyła swój profil genetyczny do MyHeritage. Dla osoby, której przodkowie pochodzą z Afryki Subsaharyjskiej odsetek ten spada do 40%.

Erlich postanowił też sprawdzić, jaka jest szansa, że w takiej sytuacji policja zidentyfikuje cię podobnie, jak Golden State

Killera. Okazało się, że dodając do informacji genetycznej takie dane jak szacunkowa ocena wieku i prawdopodobna lokalizacja geograficzna można zmniejszyć początkową pulę 850 osób do zaledwie 17 podejrzanych.

GEDmatch, dzięki której policja znalazła mordercę, zawiera DNA około 0,5% dorosłej populacji USA. Gdy odsetek ten wzrośnie do 2% dzięki tej bazie będzie można zidentyfikować ponad 90% osób mających europejskie korzenie. „To zaskakujące, jak niewielka musi być ta baza” – mówi genetyk Noah Rosenberg z Uniwersytetu Stanforda.

Co więcej, połączenie baz policyjnych z ogólnodostępnymi może dać jeszcze lepsze wyniki. Jak donosi najnowszy numer Cell, analizy wykazały, że ponad 30% osób, których profil DNA znajduje się w policyjnej bazie danych, można skojarzyć z ich krewnymi z komercyjnych baz danych. Jakby jeszcze tego było mało, dodatkowe łączne analizy danych zawartych w obu typach baz pozwalają ujawnić szczegóły wyglądu podejrzanego lub jego krewnych, jak kolor oczu czy przebyte choroby. „Z tymi bazami można zrobić więcej, niż się powszechnie uważa” – mówi Rosenberg.

Uczni zwracają uwagę, że takie analizy, które są pocieszające z punktu widzenia pracy policji, mogą jednak stanowić zagrożenie dla prywatności. Rozwiązaniem może być np. przesyłanie serwisom genealogicznym swoich danych DNA w formie zaszyfrowanej. Wówczas policja mogłaby z nich korzystać tylko we współpracy z danym serwisem, a zatem za zgodą sądu. Erlich uważa też, że USA powinny zmienić prawo dotyczące prywatności osób zgłaszających się na ochotnika do badań naukowych.

Teoretycznie osoby takie powinny być trudne do zidentyfikowania. Jednak Erlich wykorzystał bazę GEDmatch do zidentyfikowania anonimowej kobiety, która dostarczyła swoją próbkę naukowcom. Jej dane zostały zanonimizowane przez instytucję badawczą, jednak kobieta korzystała też z serwisów genealogicznych, a tam dane nie są anonimowe. Zatem ktoś, kto

działa w złej wierze i ma dostęp do próbek w instytucjach naukowych, może pokusić się o identyfikację osób biorących udział w badaniach.

Być może rozwiązaniem byłoby przyjęcie przez Kongres przepisów ograniczających policji możliwość korzystania z danych genealogicznych. Użycie takich baz byłoby uzasadnione w przypadkach poważnych przestępstw jak morderstwo, ale już nie w drobnych sprawach.

Autor: Mariusz Błoński

Zdjęcie: [qimono](#) (CC0)

Na podstawie: ScienceMag.org

Źródło: [KopalniaWiedzy.pl](#)