

Човешкият геномен проект – значение и постижения

Десислава Нешева

„Проектът за човешкия геном“ (**Human Genome Project – HGP**) бе едно от най-амбициозните начинания от края на миналия век, а днес вече може да се смята за едно от най-големите постижения в историята на науката. Стартиран през 1990 г., или 37 години след откриването на структурата на ДНК, този проект спомага за разгадаването на голяма част от механизмите на наследствеността. Знаковото му постижение е генерирането на първата последователност от човешкия геном. Той предостави фундаментална информация за човешката генетична структура, ускори изучаването на човешката биология и подобри медицинската практика. Разбира се, изследването на човешката ДНК продължава – с откриване на нови, неразкрити функционално и структурно райони от нея.

А ето накратко и историята на този проект. В края на XX век международна мрежа от генетици се обединява от идеята да секвенира цялата човешка ДНК. Това е дръзко начинание, не само защото е пионерско дело, а като се има предвид и какъв колосален обем генетичен материал е трябвало да се разчете. Тогава е било известно, че двоищата верига на ДНК в нашите клетки съдържа приблизително три милиарда двойки бази, чиято последователност, ако се разгледа като текст, ще запълни стотици книги. В началото на своята работа екипът е можел да прочита ДНК

последователност от няколко бази, след което е трябвало да сглобяват фрагментите – подобно на пъзел, като се търсели еднакви краища, които се припокриват. Нищо чудно, че тази задача е отнела години!

През 2000 г. вече се обявява, че е завършен първият етап от разчитането на човешкия геном, но все още дълги участъци от веригата оставали неизвестни. А самото сглобяване на поредицата от малки парчета се оказало наистина сложно. Част от нашите гени съществуват като множество почти идентични копия, които понякога изпълняват раз-

лични функции. Други копия – известни като псевдогени – са деактивирани от мутации. А гените съставляват само малък процент от генома. Останалото може да бъде още по-объркващо. Голяма част от генома се състои от вирусоподобни участъци от ДНК, които съществуват до голяма степен, за да направят нови копия за новите вирусни частици или за да се вмъкнат отново някъде в генома.

Все пак работата напредвала. Учените постоянно усъвършенствали методите си на работа и успявали да секвенират все по-голяма част от генома. Подобрявал се както софтуерът, така и мощта на компютрите, с които се сглобявали все по-големи фрагменти. Периодично се обновявал *референтният човешки геном*, който представлява актуален вариант на разчетената до момента част. Благодарение на този референтен геном, сравнявайки пробите на пациенти, молекулярните генетици могат да класифицират все по-прецизно мутации, водещи до различни патологии и заболявания.



Десислава Нешева е възпитаник на Биологическия факултет на СУ „Св. Климент Охридски“. Бакалавър по молекулярна биология (2006) и магистър по генетика (2008). Работи в Катедра по Медицинска генетика към Медицински университет – София. Доктор по генетика от 2016 г. Специализирала е 2 години популационна, еволюционна и антропогенетика във Флорентинския университет, Италия. Участва в множество проекти и има интерес в области, свързани с изследване на онкологични заболявания, вродени малформации, проблеми на дълголетието и социално значими болести.



Логото на Човешкия геномен проект

Колаж: National Human Genome Research Institute (NHGRI)



Първата международна стратегическа среща за секвениране на човешкия геном на Бермудските острови през 1996 г. Нобелистът Джеймс Уотсън (отпред вляво с шапка) изигра важна роля в стартирането на проекта за човешкия геном и насърчаване на споделянето на геномни данни от проекта

Най-скорошният референтен геном е обновен през 2013 г. и е много по-пълнен от предходните, но все още има много неразчетени и неясни участъци, които генетиците се опитват да разучат.

Амбициите на проекта за човешкия геном не се ограничават само до *Homo Sapiens*. Още през 1988 г. замисълът е включвал сравнителни изследвания и на други геноми – списъкът с организми включва бактериите *E. coli*, хлебната мая, плодовата мушица *Drosophilla*

melanogaster, нематоди и мишки. Така се поставя началото на нова ера в биомедицинските изследвания. Проектът за човешкия геном в крайна сметка нахвърли първоначално заложените цели още до 2003 г., две години преди планираното завършване (2005 г.). Много от постиженията на проекта тогава са отвъд рамките, които учените са смятали за постижими през 1988 г.

Постигнат е значителен прогрес в методиката на работа. В началото

е използван един конкретен метод за секвениране на ДНК, наречен „Сангер“. Той е създаден още през 1977г. от биохимика Фредерик Сангер (Frederick Sanger) – единственият учен с две нобелови награди. Методът му обаче е сложен и бавен, като в хода на работата по **HGP** претърпява значителен напредък чрез серия от технически иновации.

Тук трябва да подчертаем, че последователността на човешкия геном, генерирана в проекта **HGP** не произхожда от един конкретен човек. Тя отразява смесица от множество хора, чиято самоличност е анонимна. По-голямата част от последователността идва от кръводарители от Ню Йорк.

А огромната научна работа е дело на „Международен консорциум за секвениране на човешки геном“. В него са се трудили хиляди високкоквалифицирани учени от 20 отделни университета и изследователски центрове в Съединените щати, Обединеното кралство, Франция, Германия, Япония и Китай.

Въпреки огромната работа не може да се каже, че този проект приключва с пълно разгадаване на човешкия геном. По време на проекта изследователите непрекъснато подобряват методите за секвениране на ДНК, но са ограничени в способностите си да определят последователността на някои участъци от човешка ДНК. През юни 2000 г.



Д-р Колинс анализира авторадиограма, показваща резултатите от експеримент за секвениране на ДНК по метода на Сангер, като този, използван в ранните години на проекта за човешкия геном. (Снимков архив на NHGRI)

WANTED
20 Volunteers
 to participate in the
Human Genome Project
 a very large international scientific research effort.

The goal is to decode the human hereditary information (human blueprint) that determines all individual traits inherited from parents. The outcome of the project will have tremendous impact on future progress of medical science and lead to improved diagnosis and treatment of hereditary diseases.

Volunteers will receive information about the project from the Clinical Genetics Service at Roswell Park, and sign a consent form before participating.

No personal information will be maintained or transferred.

Volunteers will provide a one-time donation of a small blood specimen. A small monetary reimbursement will be provided to the participants for their time and effort.

Individuals must be at least 18 years of age.
 Persons who have undergone chemotherapy are not eligible.


 For more information please contact the
 Clinical Genetics Service
 845-5720 (9:00 am - 3:00 pm)
 March 24 - 26, 1997

Реклама във вестник от 1997 г. в Бъфало, Ню Йорк, набираща доброволци за предоставяне на кръвни проби за изолиране на ДНК за проекта за човешкия геном. (Архив на програмата за история на геномиката на NHGRI)



Изследовател от Вашингтонския университет в Сейнт Луис държи замразени проби от човешка ДНК, които се използват за целите на проекта. (Снимков архив на NHGRI)

Международният консорциум обяви, че е изготвил проект на последователност, който представлява 90% от човешкия геном. В него обаче има повече от 150 000 района, в които ДНК последователността е неизяснена, тъй като не може да бъде разчетена точно. През 2003 г. се генерира последователност, която е значително подобрена, представляваща 92% от човешкия геном.

Резултатите са добре известни, защото проектът се развива в условията на абсолютна публичност. Учените публикуват всяка част от получените резултати на всеки етап и след поява на нови данни. Участниците в консорциума са подписали споразумение, наречено „Бермудски принципи“, което въвежда по-голяма осведоменост и споделяне на данни в биомедицинските изследвания. Това също е едно от най-важните постижения на проекта за човешкия геном.

В хода на работата се поставят широк кръг етични и социални въпроси, свързани с придобиването и използването на геномната информация. Стават ясни рисковете и ползите от включването на нови геномни данни в изследванията и медицината, както и възможности-

те за потенциална злоупотреба с тях. За справянето с тези проблеми се създава **Програмата за изследване на етични, правни и социални последици**, която се превръща в модел за спазване на биоетиката в световен мащаб.

Проектът е и успешен пример за събирането на интердисциплинарни групи, включващи експерти от областта на техниката, биологията и компютърните науки, наред с други области. Тези вид мащабни научни начинания станали по-често срещани и добре приети, отчасти като следствие от успеха на проекта за човешкия геном. Макар официално той да е завършен през 2003 г., са оставали все още неразчетени 8 процента от генома. Затова има и продължение.

През 2019 г. стартира нов проект, като консорциумът **Telomere-to-Telomere (T2T)** обявява, че ще запълни липсите и ще създаде първата наистина пълна последователност на човешкия геном. Основатели на консорциума са Адам Филипи (Adam Philippi), изчислителен биолог в Националния институт за изследване на човешкия геном, и Карън Мига (Karen Miga), генетик от Калифорнийския университет в Санта Круз.

Д-р Филипи признава, че липсващите участъци са го гразнели лично и това го е мотивирало за работа: „Вземате красив пейзаж, изваждате сто части и го гледате – това е много досадно за един перфекционист“, обяснява

той. Заедно с д-р Мига те отправят призив към учени от цял свят да довършат пъзела. Привличат 99 специалисти, работещи директно върху секвенирането на човешкия геном и десетки други, които се включват в обработката на данните.

Консорциумът използва нови секвенатори, които могат да разчетат участъци от ДНК, достигащи дължина от десетки хиляди бази. Създаден е и софтуер за установяване позицията на някои от повтарящите се последователности в генома. Като цяло учените добавят повече от 200 милиона базови двойки в референтния геном и могат да кажат с увереност, че човешкият геном е с дължина 3,05 милиарда базови двойки. В тези нови последователности на ДНК



Карън Мига и Адам Филипи, основатели на консорциума „От теломер до теломер“. Двамата обединяват и ръководят международен екип от учени в търсенето на липсващите части от човешкия геном

учените откриха повече от 2000 нови гена. Повечето съдържат мутации и са неактивни, но 115 от тях представляват белтък кодиращи гени, чиято функция тепърва ще се изучава. До момента се изчислява, че човешкият геном съдържа 19 969 гени, кодиращи протеини.

Със сглобения пълен геном изследователите могат да разгледат по-добре вариациите в ДНК на различните индивиди. Те откриват повече от два милиона нови точки в генома, където хората се различават. Някои лекари започват да използват референтния геном в своите изследвания. Примерно – на редки детски болести, при които, сравнявайки ДНК от пациентите си спрямо референтната последователност, могат да се търсят мутации, асоциирани със заболяванията.

Последната секвенирана последователност на човешкия геном от 2022 г. разкрива нови генетични варианти, свързани с развитието и прогресията на различни генетични заболявания. Откриват се кандигат-гени, асоциирани с появата, развитието и прогресията на различни заболявания като мускулна дистрофия, някои форми на рак, както и такива, свързани с дълголетието и стареенето. Новоразчетените райони, според екипа на проекта T2T, ще доведат до иновативни лечения и по-добро разбиране на човешката вариабилност.

Сред новата разчетена последователност са две ключови области на нашите хромозоми – теломерите и центромерите. Теломерите са участъци, намиращи се в края на нашите хромозоми. Те се скъсяват при всяко клетъчно делене и с напредването на възрастта. Някои изследвания показват, че хората, които имат по-къси теломери, отколкото е нормално за тяхната възраст, са изложени на повишен риск от болести на

старостта и умират по-рано от хората с по-дълги. Проучванията показват, че удължаването на теломерите е един възможен начин за забавяне на стареенето.

Центромерите имат структурна функция в хромозомата. Тази област е важна по време на клетъчното делене, тъй като именно на това място хромозомите се залавят за делителното вретено и могат да бъдат правилно разделени в дъщерните клетки. Промениите, които настъпват в тези райони, се откриват при пациенти с някои видове рак, при инфертилитет и оказват влияние и в процесите на стареене. Новият геном, наречен T2T-CHM13, разкрива близо 2000 кандигат-гени, които изследователите вече могат да изучават допълнително. Някои от идентифицираните райони разкриват гени, свързани с имунния отговор и адаптацията, както и протекцията при някои инфекции. Други райони се асоциират с фармакогенетичните процеси и реакцията на отделните организми към лекарствени препарати и форми на лечение. Гени, асоциирани с еволюционните процеси в човешкия мозък, показват как той е еволюирал при *Homo sapiens*, което ни прави уникални и ни отличава от други примати. В рамките на новоразчетените региони на генома се идентифицират „стотици хиляди варианти, които обещаваат възможност за еволюционни и биомедицински открития“, пишат изследователите.

Това стига и до практиката. Генетични варианти вече се използват за подпомагане на лечение – например за специализирани терапии на някои видове рак на гърдата. Смята се, че новата карта на генома ще извади наяве повече генетично свързани заболявания. В перспектива е и възможността за *инвитро редакция*

на някои гени, още при формирането на зиготата и ембриона. Това ще спомогне за намаляване и дори за елиминирание на тежки наследствени състояния, синдроми и форми на рак. Отваря се вратата към нова ера на високопроизводителна цифрова биология.

За разлика от предходните изследвания, в които се използва ДНК, изолирана от доброволци, в това изследване се използват специфични, изкуствено моделирани и култивирани *инвитро* клетки, носещи генетичната информация единствено от сперматозоиди. Те се получават чрез оплождане с нормален сперматозоид на яйцеклетка с отстранено ядро.

Изследователите вече разполагат и с две нови технологии за секвениране на ДНК, които им помагат да осъществят този проект: методът на Oxford Nanopore, който може да прочита до 1 милион ДНК бази наведнъж, но с някои грешки и методът на PacBio HiFi, който може да разчете 20 000 букви с 99,9% точност.

Методът Oxford Nanopore позволява секвениране на ДНК или РНК чрез прокачане на веригата на нуклеиновата кисе-

лина през нанопора (пори с размери в диапазона 1 – 100 нм). В този процес всяка база, която преминава през нанопората, може да бъде идентифицирана в реално време. Секвенирането чрез нанопори позволява директен анализ на къси до ултрадълги фрагменти от ДНК/РНК. Така учените получават бърз достъп до важна информация за изключително кратко времето, имат възможност за генериране на информация за много проби и по-голям контрол върху целия процес.

Методът на секвениране PacBio HiFi осигурява разделителна способност на разчетените бази с точност на четене на една молекула от порядъка на 99,9%. Този метод може да се използва широко за секвениране – на целия геном и неговото сглобяване, за цялостно откриване на варианти, за епигенетично характеризиране, както и на РНК и груги.

Секвенирането на ДНК най-общо напомня сглобяване на пъзел. С по-дългите фрагменти от новите методи пъзелът се нарежда много по-лесно и бързо. Стандартните човешки клетки съдържат гва комплекта ДНК – майчино копие и бащино копие, но този екип използва ДНК от група клетки, наречена пълна



Изображение, представящо преминаването на четящата се верига през нанопората. Източник: <https://nanoporetech.com/support/how-it-works>

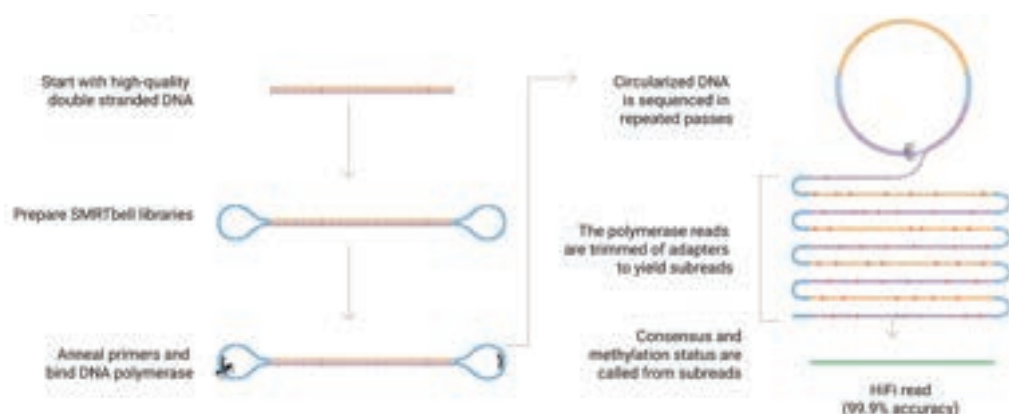


Схема на работа на метода PacBio HiFi, източник: <https://www.pacb.com/technology/hi-fi-sequencing/Tbx0>

хидатидиформна бенка, която съдържа дубликат на бащиния набор от ДНК. Пълната хидатидиформна бенка е рядко усложнение на бременността, причинено от анормален растеж на клетки, които произхождат от плацентата. Този подход опростява генома така, че учените се нуждаят от последователност само на един комплект, а не на два комплекта ДНК. Учените се надяват, че през следващите 10 години секвенирането на геномите на индивидите може да се превърне в рутинен медицински тест, който ще струва много по-евтино и ще е по-достъпен. Това ще доближава все повече науката до персонализираната медицина.

Ако се върнем към целия процес на разчитането на човешкия геном, трябва да напомним, че дълго време нито един индивидуален човешки геном не е бил разчетен в неговата цялост. Референтните последователности всъщност са съглобени от парчета ДНК на различни хора. Това бе най-доброто, което можеше да се направи по това време, но то бе свързано с големи пропуски и грешки. Едва днес технологията напредна достатъчно, че е възможно да се разчете целия човешки геном с минимални грешки.

И вече е представен целият геном на един реален човек – той се казва Леонид Миронович Пешкин (Leonid Peshkin). Роден е през 1970 г. в Москва, учи в Харвард, където започва научна кариера като биолог с публикации по проблемите на изкуствения интелект и стареенето. Световна известност обаче добива с това, че щатският проект „Геном в бутилка“ (**Genome in a Bottle**) избира него и родителите му като световен стандарт и през 2022 е обявено, че консорциумът Telomere-to-Telomere (T2T) е създал изцяло пълен човешки геном, като е използвана Y-хромозомата на Пешкин.

Този пълен, единствен човешки геном е монументално техническо постижение. Само 70 години, след като структурата на двойната спирала на ДНК беше разкрита за първи път, вече имаме способността да прочетем цялата генетична последователност, която прави човек уникален. Разбира се, изследванията не свършват дотук. Това е началото на нов етап – секвениране на геномите на хора от цял свят, за да се изгради истинска картина на генетичното разнообразие на нашия вид. Секвенирането на генома от край до край в клиниките ще стане рутинна процедура с цел по-добра

диагностика, прогнозиране и терапия на много заболявания.

Човешкият геном обещава много неща, като това да открием функцията на всички гени, особено на тези, които участват в детерминирането и прогресията на различни наследствени заболявания. Това би позволило създаване на персонализирана медицина, в която получаваме лечение, оформено според генетичния профил. Геномът също така може да даде отговор на въпросите, свързани с еволюционния ни произход и примерно как точно се различаваме от най-близките ни живи роднини – шимпанзетата.

Със сигурност днес знаем много повече за функциите на много гени и тяхната роля при заболявания, вариращи от рак на гърдата до шизофрения. На практика обаче се оказва, че повечето заболявания се влияят от стотици гени, така че геномната медицина има още много път да извърви. Малко наследствени заболявания са причинени от един дефектен ген и използването на генетичен скрининг за откриване на носители на редки заболявания се използва до голяма степен само за тези, които се считат за най-рискови.

Консорциумът Telomere-to-Telomere не е просто самоцелен грандиозен проект, а масово усилие, което е провокирано от пандемията Ковид 19. Първият голям успех на T2T е през юли 2020 г., когато публикуваха пълната последователност на човешки X хромозом. На следващата година те представят пълната последователност на хромозома 8. През 2021 г. те публикуват материал, озаглавен „Пълната последователност на човешкия геном“, в който попълват голяма част от липсите. Повечето от клетките в телата ни съдържат две копия на всяка хромозома: едно получено от май-

ката, а другото – от бащата. Това прави по-трудно повторното съгласяване на части от последователността, тъй като двете копия са различни. За преодоляване на тази трудност T2T използват необичайни клетки, които имат две почти идентични копия на ДНК на бащата. Клетките идват от яйцеклетка без ядро, която губи своето ДНК и след това се опложда от сперматозоид. След това клетката губи ДНК от сперматозоида. Секвенцията на T2T добавя над 200 милиона бази към човешкия геном и 2000 допълнителни гена.

Наличието на пълен геном означава, че е възможно да се разбере каква функция имат повтарящи се дълги участъци от ДНК. Повтарящата се ДНК включва и много последователности, които могат да се движат в генома, наречени „мобилен ДНК елементи“. Много еволюционни скокове, включително загубата на опашката, както и някои мозъчни функции, могат да се дължат на някои видове мобилна ДНК. Други важни последователности са „сегментните губливания“, при които дълги участъци от ДНК, които могат да включват множество гени, се губират многократно. Те водят до появата на нови гени, които са специфични за човека и допринасят непропорционално за различията, които ни правят хора. Докато човешкият геном и геномът на шимпанзето са 99% идентични, сегментните губливания са един от начините, по които могат да възникнат важни разлики, а първоначалната последователност на човешкия геном ги е пропуснала до голяма степен.

Голямата оставаща празнина до този момент е Y-хромозомата, която се среща и предава само при мъже. Сперматозоидите обикновено носят само една полова хромозома – или X, или Y. Тъй като ДНК на изкуствено създадената оп-



Лео Пешкин, учен и изследовател, човекът с най-детайлно секвенирания геном. Източник <https://kirschner.hms.harvard.edu/people/leon-peshkin>

логена куха яйцеклетка, използвана от T2T, идва от сперматозоид, който носи X хромозома, У хромозомата остава неразчетена. С цел да се запълнят липсите по последователността на У хромозомата, екипът се обръща към Леон Пешкин.

ДНК-то на Пешкин е използвано от проекта Genome in a Bottle (GIAB), който има за цел да разучи геномите на клетъчни линии, които могат да се отглеждат неограничено дълго време в лабораторията, което улеснява изучаването на ефектите от мутациите. Геномът на Пешкин е важен, защото той представя генома на родителите си, давайки на GIAB триото майка-баща-син. През декември 2022 г. T2T публикува статия, описваща последователността от край до край на У хромозомата на Пешкин. Поради наличието на много повторения и специфики, повече от половината от хромозомата липсва в предходните референтни геноми. Новата последователност добавя над 30 милиона нуклеотида в разчетената последователност, включително десетки гени. През последните месеци екипът работи върху целия геном на Пешкин, включително и двете копия на всяка хромозома. Полученият геном е „пълн и без пропуски“, а дублиращите се хромозоми са напълно

разделени. В края на 2022 г. и началото на 2023 г., Национален институт за изследване на човешкия геном (NHGRI), част от Националния здравен институт, САЩ публикува данни за първия човешки пангеном, който представлява смесица от разчетени геноми на 47 различни етнически и расови групи хора. Той отразява най-добре характерното генетично разнообразие на човешкия вид. Разработват и софтуер, наречен Verkko, означаващо „мрежа“ на фински, който прави процеса на съгласяване на пълни геномни последователности по-бърз и е достъпен. Желанието им е този метод за секвениране на генома да бъде включен в рутинните изследвания и да има голяма достъпност. Все пак трябва да напомним, че няма единен човешки геном. ДНК на всеки един човек е различна и тези разлики имат значение. Няма да разберем наистина генома, докато не установим как той варира в отделните човешки популации.

Ето защо много членове на консорциума T2T се включват и в груп проект: „Референтен консорциум за човешкия пангеном“. Целта му е да се секвенират стотици хора от цялата планета, за да може базата данни да бъде непрекъснато подновявана. По този начин ще се постигне и крайната цел – да има референтен геном, който да се използва за всички пациенти и този метод да влезе в практиката и диагностиката като рутинен метод.

Учените смятат, че имаме още много да научим от нашите геноми, тъй като новите методи разкриват повече от тайните на геномите и позволяват повече да бъдат секвенирани. За тях предизвикателството продължава. Според изследователите „Докато има хора, проектът за човешкия геном ще продължи“. ■

Първи случай на заболяване на човек от растителна гъба

В началото на 2023 г. научните списания съобщиха тревожна новина: установен е първи случай на заболяване на човек от растителна патогенна гъба, поразила белите му дробове! Касае се за 61-годишен мъж от Индия, който бил диагностициран с остра кашлица, липса на апетит, умора, остър фарингит и гр. признаци, които продължили почти 3 месеца преди отиването му в медицинска клиника. Той работил до момента на заболяването в лаборатория за изучаване на гъбите като миколог и изследвал материали от различни видове, предимно стари и разлагащи се гъби. Този първи в медицинската практика в света случай на



Външен вид на дървесна гъба *Chondrostereum purpureum*

заболял човек от растителна гъба е описан по-подробно в списанието *Medical Mycology Case Reports* (Март, 2023 г.).

Първите медицински изследвания на заболялелия чрез компютърна томография установили наличие на възел от дясната страна на дихателната му тръба, в който открили хифи на неизвестна гъба в процес на разрастване в човешките тъкани. Тъй като непознатата гъба не е било възможно да се идентифицира само по откритите растящи хифи, материал от нея била изпратен и в Центъра за гъбни заболявания към Световната здравна организация (WHO) в Индия. След генетичен анализ там е уста-

новено, че причинителят се отнася към вида *Chondrostereum purpureum*, причинител на заболяването „Сребърен лист“ (Silver leaf) по растенията. След успешно сложно лечение пациентът е изписан и вече се чувства здрав, по данни от клиниката.

Случаят представлява значителен интерес от медицинско гледище като първи случай в медицинската практика. Гъбата причинител на заболяването е широко разпространен патоген по растенията, включително и по плодни растения като ябълки, череши, декоративни растения и гр. и възможността за заразяване на човек с нея са доста големи. Характерно за причинява-

ното от нея заболяване по растенията е появата на т.н. „сребърни листа“, пронизани от хифите на гъбата и отличаващи се със сребриста лъскавина (гланц) от здравите листа и растения. Гъбата образува понякога и едри плодни тела по стволите на дървесните видове. Този случай, заключават авторите на изследването, „показва неочаквани възможности за проникването и на гъбни растителни патогени в човек, контактен с тях... А това изисква още много усилия да бъде изяснена тяхната патогенност и възможности за контрол на тези инфекции“.

По Live Science