

Геноміка — медицині XXI століття



Юрій Бажора
доктор мед. наук,
проректор
Одеського
медичного
університету,
м. Одеса

Досягнення молекулярної біології другої половини XX століття надали вченим змогу розпочати реалізацію проекту щодо розшифровки генома людини (*Human Genome Organization — HUGO*). Його було завершено, в цілому, 2001 р.. В процесі виконання цієї програми виникли нові напрями молекулярної біології — геноміка і біоінформатика.

Геноміка — наука, що вивчає структуру та функцію генів. Вона інвентаризує гени, створюючи геномні карти живих істот. Під час роботи над розшифруванням генома, комп'ютерний аналіз мільйонів послідовностей ДНК трансформувалася в самостійну наукову галузь — біоінформатику.

Розвиток цих двох наукових напрямів дозволив наприкінці першого етапу проекту визначити послідовність близько 3 млрд. пар нуклеотидів у геномі людини. Якщо перший мільярд було прочитано за чотири роки, то наступні — за чотири місяці.

Завершення програми HUGO, незважаючи на її масш-

табність, призвело, певною мірою, до деякого розчарування її результатами як в наукових колах, так і в соціальних структурах.

Однак слід підкреслити основний її результат — вона сформувала **геноміку** як науку. Остання почала свій інтенсивний розвиток, створюючи такі наукові напрями як **структурна геноміка, порівняльна геноміка та функціональна геноміка** (рис. 1).

Структурна геноміка вивчає послідовність або сіквенс (від *sequence*, англ. — *послідовність*) нуклеотидів з першого до останнього в кожній хромосомі. В цьому напрямі вченим доведеться ще багато попрацювати для виявлення послідовностей в геномах багатьох прокариот, рослин та тварин.

Результати, отримані структурною геномікою, є відправною точкою для порівняльної геноміки, яка порівнює зміст та організацію геномів різних організмів. Методичні підходи, котрі розроблялися під час секвенування генома людини, дозволили значно прискорити розшифрування геномів багатьох прокариот і еукариот.

Першим представником прокариот, геном якого було повністю картовано, була бактерія *H. influenza* (1995 р.), в наступному розшифрували геноми понад двадцяти інших мікроорганізмів, збудників важких захворювань людини (туберкульозу, сифілісу та ін.).

1996 р. було секвеновано геном еукаріотичної клітини — дріжджів, а згодом низки більш високоорганізованих багатоклітинних організмів: *нематоди C. elegans*, *мухи дрозофіли*, *рослини арабидопсис*, *щура, мавп*.

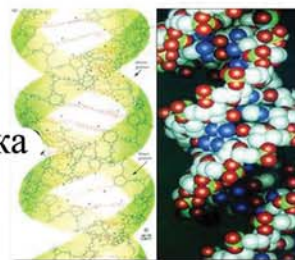
Картування геномів представників видів, які перебувають на різних рівнях організації, дає можливість порівняти їхні повні геноми. При цьому визначаються спільні гени і виділяються "особисті" гени.

Розшифровка геному людини

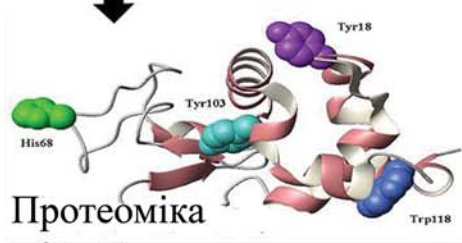
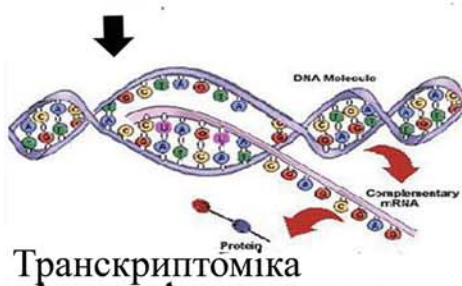


Біоінформатика

Геноміка



- * Структурна геноміка
- * Порівняльна геноміка
- * Функціональна геноміка



Функція



Рис. 1.

Скажімо, багато генів нематоди, котрих немає у одноклітинних дріжджів, мабуть, пов'язані з міжклітинними взаємодіями у багатоклітинного організму. *У людини кількість генів тільки в 4 – 5 разів перевищує їх чисельність у нематоди. Виходить, що багато з них спільні з нематодами і дріжджами. Ці результати дозволяють, по-перше, значно полегшити пошук "своїх" генів людини, а, по-друге, визначаючи спільні гени людини і черв'яка або дріжджів, виключати їх у нижче розташованих організмів і з'ясувати функцію продуктів цих генів (структурних білків та ферментів).*

Порівняльна геноміка надасть можливість з'ясувати тонкі механізми еволюції живого світу. Результати порівняння структури геномів дозволили зробити висновок, що еволюція ссавців проходила не стільки через помноження різноманітності генів, скільки шляхом поступового копіювання, модифікації та комбінації вже наявних генів та шляхом формування механізмів їх регуляції. Встановлено також, що людина відрізняється від шимпанзе приблизно 2% своїх генів і дещо більше — від горили. Гени гомеобоксу, які керують формою тіла у людини та інших ссавців, дуже схожі з аналогічними генами у простіших істот, які виникли 500—600 млн. років тому.

Порівняння геномів різних організмів дозволило встановити у них співвідношення кодуєчих та некодуєчих інформацію ділянок. Скажімо, в одноклітинних дріжджів відносний вміст екзонів, які кодують інформацію, дуже великий; у нематоди *C. elegans* співвідношення екзонів (27%) та інтронів (26%) приблизно однакове, а некодуєчі ділянки складають 47%; у людини кодуєчі ділянки геному становлять близько 3%.

Таким чином, *в процесі еволюції відносна кількість інформації про білки та РНК на довжину ДНК зменшується. Значний об'єм інформації ДНК залишається невідомим. У свій час Ф. Крік назвав її "егоїстичною". Тепер зрозуміло, що під час еволюції вона надає більш прогресивним видам певні переваги.*

Функціональна геноміка вивчає функції нещодавно відкритих генів та так званих "генних мереж", що визначають розвиток тканин і органів в нормі та їх порушення при різних захворюваннях. Внаслідок розшифровки генома людини ідентифіковано понад 30 тис. генів і встановлено функцію приблизно 8 тис. генів. Функція решти картованих і некартованих генів залишається невідомою. З'ясування їх — основне завдання програми "Функціональна геноміка". Зараз вона реалізується швидкими темпами і особлива увага приділяється *поліморфізму поодиноких нуклеотидів (single nucleotide polymorphism — SNP або сніп). Вважають, що за весь час існування людини як біологічного виду накопичена достатня кількість SNPs (приблизно одна мутація на тисячу нуклеотидів).*

Заміна нуклеотидів, що виникає в регуляторних ділянках ДНК, наприклад, у промоторах структурних генів, суттєво впливає на експресію цих генів, що підвищує ризик виникнення у людини різних хвороб (злоякісні пухлини, діабет тощо). Ці мутації успадковують наступні покоління. SNPs, пов'язані з генами, які містяться поряд, служать генетичними маркерами, за допомогою яких можна точно встановити локалізацію генів в хромосомах. Це дуже складна та кропітка робота, яка в багатьох лабораторіях світу виконується з середини 1990-х рр., проте вона є одним із шляхів до **"персоналізованої медицини"**.

На початку роботи над проектом щодо складання зведеної карти сніпів його автори вважали, що дос-

татньо буде встановити 150 тис. SNPs, але потім необхідність в них зросла до 500 тис.

З'ясовано зв'язок SNPs в генах, які містять спадкову інформацію про ферменти, що метаболізують лікарські засоби в організмі людини, та в генах, які мають відношення до транспорту цих ферментів, у генах, які пов'язані з рецепторами клітин, з терапевтичною активністю, побічними ефектами та токсичністю лікарських препаратів. Вивчення точкових варіацій у зазначених генах та їхня роль у долі ліків в організмі людини — основне завдання фармакогеноміки. Швидке визначення SNPs в ключових генах ферментів транспорту та метаболізму ліків дозволяє індивідуалізувати дозування ліків, ухилитися від побічних дій і зменшити вартість лікування хворих. Пов'язана з SNPs—поліморфізмом різноманітність ізоформ цитохрому P450 проявляється в повільному, нормальному або швидкому метаболізмі ліків.

Наступний напрям у вивченні генної експресії — **транскриптоміка**, до завдання якої належить виявлення в клітині інформаційних РНК (i-РНК) та розшифровка їх функцій, що вважається одним із перспективних шляхів у вивченні роботи генів. Водночас клітина виробляє велику кількість різних i-РНК. Розшифрування цих i-РНК дозволить виявити всі гени, які працюють цієї миті. Вилловлюють і пізнають i-РНК за допомогою ДНК-мікрочіпів. Певна i-РНК взаємодіє (гібридується) зі своїм геном або його фрагментом в ДНК. В різних клітинах експресуються різні гени, тому картини гібридизації будуть різними між цими клітинами. Картина гібридизації генів, які експресуються в клітині, називається **паттерном (візерунком, схемою) гібридизації**, а дані про одночасну експресію певного гена в різних клітинах — **профілем експресії гена**. Картина експресії генів може бути критерієм оцінки функціонального стану клітини.

На підставі геноміки виникла **протеоміка**, яка вивчає будову і функцію білків, систематизує їх. Протеоміка застосовує високотехнологічні методи вивчення білків. За допомогою їх можна визначити кількість того чи іншого білка в зразках, ідентифікувати його, уточнити первинну структуру, а також післятрансляційні модифікації.

Протеомний аналіз здійснюють за три етапи. Спочатку виконують двомірний електрофорез нормальної і патологічно зміненої тканини, що дозволяє виявити до 10 тис. різних білків. Аналіз електрофореграми дає можливість встановити, які білки і наскільки збільшились чи зменшились у змінній тканині. Останні можуть стати маркерами для конкретного захворювання і потенційними мішенями для лікарських засобів, що розробляються. Зазначені білки в наступному вивчаються на мас-спектрометрі (для визначення молекулярної маси) і на секвенаторі (для визначення первинної структури). Досліджувати білки можна в суміші, не розділяючи їх. Щоправда, робота ускладнюється тим, що за допомогою електрофорезу важко розділити гідрофобні білки, а вони якраз і є найцікавішими для медицини, оскільки до них належать білки-рецептори — головні мішені лікарських засобів. Крім того, молекулярна маса і первинна структура не дають повної уяви про роботу білка, бо його функція, значною мірою, визначається його просторовою структурою. Остання дуже чутлива до дії оточуючих білкову молекулу факторів.

Складність досліджень у протеоміці полягає в тому, що сукупність генів у кожній клітині однакова, а набір

білків у різних клітинах досить різноманітний, і різниця може обчислюватися порядками. Протеом кожної клітини, на відміну від геному, змінюється залежно від її функціонального стану. І все ж, інтенсивний розвиток протеоміки триває.

Одним з перших її вагомих здобутків є дослідження *протеомного статусу жінки-донора*. Він вміщував інформацію про 150 тис. білків (їх отримали із 150 біоптатів), які кодувалися дванадцятьма тисячами генів.

Зараз йде мова про створення *"каталогу протеїнів"*. Такі проекти вже започатковано, але вони не стали успішними. Лише після розшифрування геному людини стали можливими масштабні дослідження протеому клітини людини. Для цього в 2001 р. було створено міжнародну організацію, еквівалентну до HUGO — HUPO (Human Proteome Organization). Одним з головних її завдань є створення атласу білків людини.

З геномікою та протеомікою пов'язаний ще один науковий напрям, що інтенсивно розвивається — *метаболоміка*, яка вивчає метаболічні шляхи, котрі мають місце в клітині у їх взаємозв'язку з активністю генів і вмістом відповідних білків-ферментів, які контролюють певні етапи кожного метаболічного шляху. При вивченні інтегральних показників у метаболоміці широко використовують комп'ютерні методи.

В клінічній практиці рідко буває так, щоб одна речовина була надійним маркером хвороби. Та й цінність його для діагностики може бути сумнівною. Наприклад, рівень холестерину в крові не завжди відповідає клінічному стану хворого. На жаль, багато з біохімічних показників при тій чи іншій хворобі недостатні для точної діагностики. Для повної характеристики хвороби необхідні комплекси показників.

Ще на початку проекту HUGO було зрозумілим, що без системного комп'ютерного аналізу сотень мільйонів послідовностей ДНК розшифрувати геном у зазначений термін неможливо. При виконанні цієї програми такий комп'ютерний аналіз перетворився в самостійну науку — *біоінформатику*.

Для дешифрування інформації, яка записана в ДНК, було створено спеціальні програми, наприклад, програма для статистичного аналізу розподілу нуклеотидів в ДНК. З їхньою допомогою вирішено чимало біологічних задач. Було з'ясовано, що сполучення нуклеотидів в кодуючих ділянках ДНК підкоряється певним правилам, а в міжгенних ділянках сполучення нуклеотидів близьке до випадкового.

Статистичний аналіз послідовностей нуклеотидів дозволив установити ділянки геному з визначеними властивостями. Наприклад, у бактерій більшість генів розташовано в "острівках патогенності", в яких є певна частота наявності пар нуклеотидів А-Т і Г-Ц. Відповідне сполучення нуклеотидів вказує на екзон чи інтрон гена, а їх межі зазначені спеціальними нуклеотидними сполученнями. Після секвенування ДНК пошук гена займає хвилини при наявності добрих алгоритмів пошуку.

Методи біоінформатики використовують не тільки для пошуку генів, але й для виявлення регуляторних ділянок в ДНК, передбачення структури та функції білка, його локалізації в клітині для з'ясування шляхів метаболізму. Реконструкція взаємопов'язаних реакцій метаболізму — головне спрямування наступних після розшифрування генетичних досліджень.

Таким чином, *біоінформатика*, що виникла як наукова дисципліна на підставі геноміки, супроводжує її поступовий розвиток.

Біоінформатика, фактично, стала однією зі сполучених ланок у комплексі єдиного ланцюга: *геноміка — транскриптоміка — протеоміка — метаболоміка* (рис. 1). На цьому рисунку показано не тільки взаємозв'язок галузей молекулярної біології, що динамічно розвиваються, але й інтеграцію їх в єдину систему для вивчення основ життя.

Розшифрування структури геномів різних за рівнем організації живих істот показало, що в них є великий запас функціональної інформації, про яку вченим мало що відомо. Крім того, стало зрозумілим, що механічне об'єднання складників не дає цілісної уяви про функціональні процеси в клітині, тим більш у складному багатоклітинному організмі. В цих процесах задіяна велика кількість проміжних складників. Вони, взаємодіючи один із одним, формують багатомірні просторово-часові мережі. Наявні методи досліджень таких складних систем не дозволяють вивчити їх на відповідному рівні.

Із зазначеного випливає необхідність системного підходу до вивчення живої клітини, яка є структурно-функціональною одиницею живого. Складність у тому, що неймовірна кількість клітин складного організму відрізняється між собою за функціональним станом їхнього геному і, відповідно, структурою та функцією білків та інших речовин, які містяться в них. Крім того, організми одного виду неоднакові, оскільки відрізняються між собою за генотипом і фенотипом.

На нинішньому етапі розвитку молекулярної біології, де все ще панує редуccionізм у вивченні живих систем, продовжується накопичення інформації про структуру і функцію генів та їхніх продуктів (РНК, білків). Паралельно розробляються та вдосконалюються технології, за допомогою яких можна досліджувати клітину в тривимірному просторі і у динаміці. Великі надії покладають на біоінформатику, котра дозволяє систематизувати та інтегрувати факти стосовно будови і функції молекулярних та субклітинних утворень та показати, як вони працюють разом. Таким чином, передбачено вирішити ще багато питань, які поставлені геномікою, протеомікою та метаболомікою, щоб вийти на якісно новий рівень розуміння основ живого.

Слід підкреслити, що геноміка виникла як наука в процесі розшифрування геному людини. Мабуть, у зв'язку з цим, її досягнення мають безпосереднє практичне втілення в фундаментальні та прикладні напрямки медицини. Наочними прикладами можуть бути численні спадкові захворювання, причиною яких є дефекти у відповідних генах. ДНК-діагностика батьків, плода та новонароджених у родинах зі схильністю до того чи іншого захворювання дозволяє або завчасно (за рішенням батьків) перервати вагітність, або своєчасно розпочати лікування дитини.

На цей час ідентифіковано багато генів, з дефектами яких пов'язані спадкові хвороби (муковіцидоз, м'язова дистрофія Дюшенна, хорея Гентінгтона, хвороба Альцгеймера, талассемія та інші).

Найоптимальнішим рішенням цих проблем була б заміна дефектних генів. Проте вона поки що залишається невирішеною. Необхідно знати точну локалізацію гена в геномі. Введений ген має проникнути у більшість клітин організму. Крім того, не відомо, як введений ген буде поводити себе в своїй "генній сітці" в процесі реалізації кінцевої структури або функції. Незважаючи на зазначені труднощі, дослідження в цьому напрямку проводяться й, імовірно, зазначені перепони буде подолано.

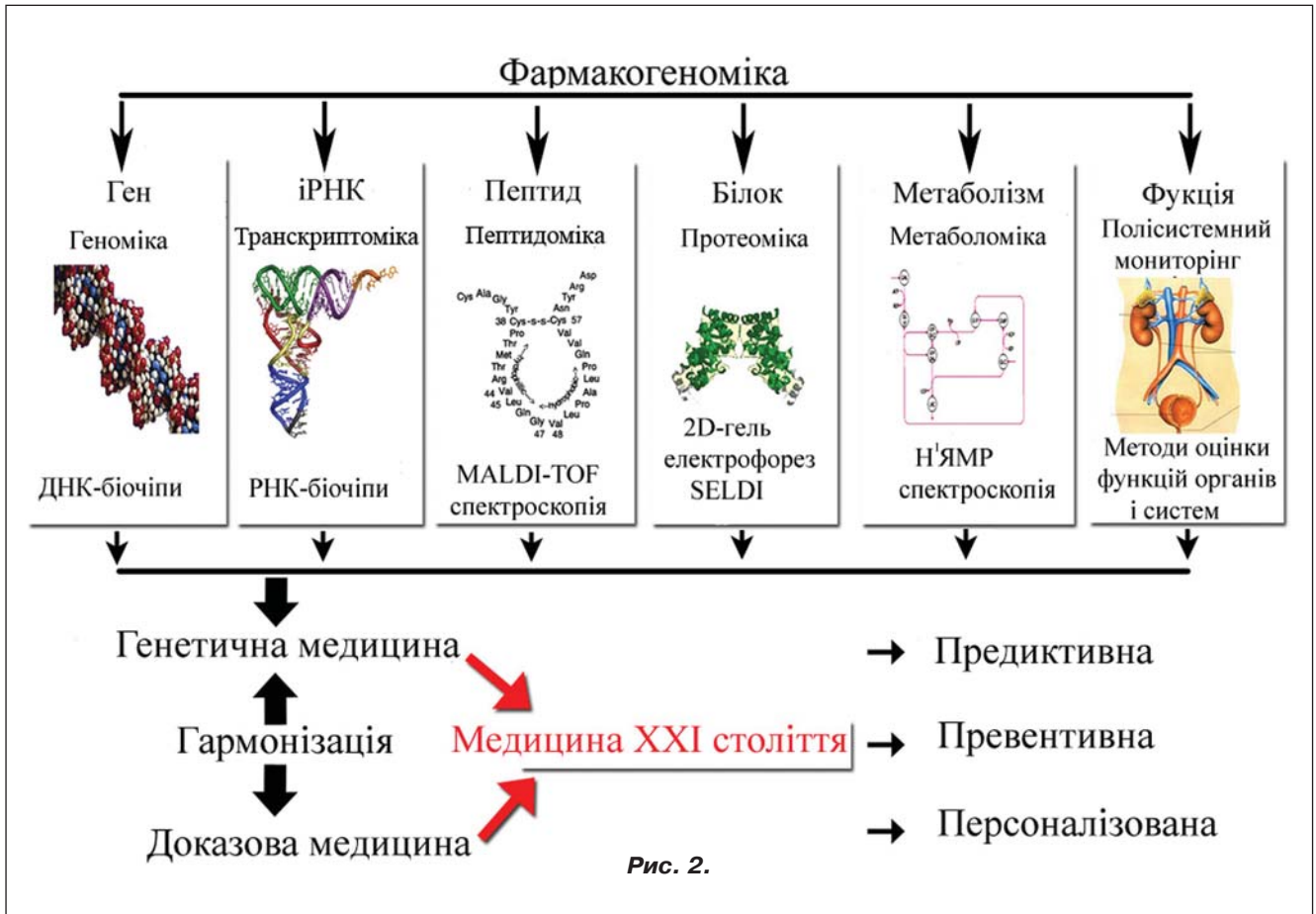


Рис. 2.

Спадкові захворювання, як багато їх не було б, становлять приблизно 2% всієї патології у людства. Тобто, вони не є соціально значущими. Решта 98% хвороб, які досить розповсюджені, а багато з них посідають провідні місця стосовно смертності та інвалідності сотень мільйонів людей, залишаються за межами медичної генетики.

Розвиток геноміки та інших напрямків молекулярної біології, розробка новітніх надшвидкісних автоматизованих апаратних комплексів для аналізу генів та їхніх продуктів, дозволили відкрити молекулярні механізми багатьох захворювань, пов'язаних з генетичною схильністю, яка, в сполученні з несприятливими навколишніми факторами, є основою розвитку соматичних, онкологічних, інфекційних захворювань.

Геноміка спричинила трансформацію фармакогенетики, яка спиралася на феноменологію різної чутливості хворих до лікарських засобів, у **фармакогеноміку**, що вивчає дію особливостей будови генів, які визначають метаболізм відповідних ліків. Фармакогеноміка акумулює також досягнення транскриптоміки, протеоміки, метаболоміки (рис. 2).

Наприклад, з протеомікою пов'язане створення ліків спрямованої дії. Знання точної будови активного центру ферменту, дефект якого викликає захворювання і який потрібно послабити, дозволить моделювати для цього відповідну хімічну сполуку. Отримані ліки будуть ефективнішими і менш токсичними. На підґрунті протеоміки розробляються, в першу чергу, ліки для лікування онкологічних захворювань.

Відомо, що пухлинні клітини мають на своїй поверхні білки-маркери, котрі відсутні у здорових клітин. Причому, ці білки специфічні для кожного виду пухлин. Створюються компоненти лікарських засобів, що

впізнають відповідний білок, приєднуються до нього і, таким чином, полегшується проникнення ліків в уражену клітину. Крім того, стан білків у клітинах відповідних органів має велике значення для діагностики. Захворювання може ініціюватися зовнішніми чинниками або помилкою, що закодована в гені. Проте, наслідки у всіх випадках проявляються на рівні білків.

Важливе значення для вдосконалення діагностичних методів у сучасній медицині мають досягнення **метаболоміки**. Як вже зазначалося, метаболоміка вивчає метаболічні шляхи. Метаболізм порушується практично при всіх захворюваннях.

За допомогою сучасних методів метаболоміки зсуви в обміні речовин можна виявити ще до розвитку клінічних симптомів і навіть до змін показників рутинних методів клінічної лабораторії. Під метаболічним аналізом маються на увазі дослідження паттернів метаболізму організму в цілому. Для цього використовують комп'ютерні програми розпізнавання образів при дослідженні інтегральних показників. Такі показники становлять основу полісистемного моніторингу оцінки функціонального стану як окремих систем органів, так і організму в цілому.

Таким чином, на цей час на основі молекулярної біології починає формуватися **генетична медицина** (рис. 2). Вона гармоніює з фундаментальними принципами доказової медицини, які вимагають незалежних доказів діагностичної цінності тих чи інших методів і терапевтичної ефективності призначених ліків.

Можна передбачувати, що ці два напрями будуть локомотивами руху медицини XXI століття, яка має стати справді предиктивною, превентивною і персоналізованою.