ГЕНОМ ЛЮДИНИ ТА МЕТОДИ ЙОГО ВИВЧЕННЯ
ЛЕКЦІЯ 1

Тема: Геном людини

План:
1. Загальна характеристика геному людини
2. Тандемно повторювані послідовності
3. Диспергувані повторювані послідовності
4. Картування генів людини

1. Загальна характеристика геному людини

Проєкт розшифрування геному людини (Human Genome Project) був одним із найамбітніших наукових проєктів XX століття. Він тривав 13 років і завершився у 2003 р. Проєкт мав на меті побудувати достатньо точну генетичну карту (з кроком 2–5 сМ), отримати послідовність 3 млрд пар основ ДНК, ідентифікувати генні послідовності, визначити геномні варіації. У результаті виконання проєкту побудовано генетичну карту з кроком у 1 сМ, отримано послідовності 99,9 % еухроматинових ділянок геному з точністю 99,99 %, отримано 15 тис. повних кДНК та ідентифіковано близько 3 млн нуклеотидів, що варіюють у геномах людини (SNP — single nucleotide polymorphism).

Традиційно геном людини представляють як сукупність послідовностей ДНК 22 аутосом і двох статевих хромосом X та Y. Незважаючи на великий розмір геному (приблизно 3,2 млрд пар основ), верхня оцінка кількості білкових генів у геномі людини не перевищує 25 тис (ідентифікація всіх генів РНК є досить далекою від остаточного завершення). Останні версії каталогу білкових генів людини включають близько 21 тис. білкових генів і 4 тис. генів РНК (остання оцінка, напевно, занижена). Загальна довжина кодуючих ділянок генів становить приблизно 34 млн пар основ, тобто 1,2 % геному. Складнішим є визначення зв'язку між генним локусом, продуктом і фенотипом, який він може зумовлювати.

Щодо мітохондріальної ДНК людини, геном мітохондрій представлений кільцевою молекулою ДНК розміром 16 тис. 569 пар основ і містить 37 генів: гени білків, які беруть участь в окисному фосфорилюванні, 2 гени рРНК і 22 гени тРНК.

Ядерна геномна ДНК людини організована у хромосоми. Нормальний хромосомний набір людини складається із 46 хромосом: 22 пар аутосом і пари статевих хромосом (XX у жінок та XY у чоловіків). Рівномірно пофарбовані метафазні хромосоми людини за розміром і морфологією поділяються на 7 груп (позначаються латинськими літерами від A до G). Для більш точної ідентифікації хромосом та їхніх специфічних ділянок застосовують спеціальні методи диференційного забарвлення хромосом, які базуються на використанні певних барвників і процедур попередньої обробки препаратів.

ГЕНОМ ЛЮДИНИ ТА МЕТОДИ ЙОГО ВИВЧЕННЯ

Найпоширенішим методом диференційного пофарбування є метод G-забарвлення, який дозволяє отримати хромосоми, нерівномірно пофарбовані вздовж їхньої довжини. Малюнок чергування темних і світлих смуг є унікальним для кожної пари гомологічних хромосом. Цікаво, що темні та світлі смуги асоційовані з ділянками, збагаченими на АТ- і GC-пари відповідно (тобто з геномними зонами, збагаченими на мобільні елементи та кодуючі послідовності). Практично таку саму картину чергування смуг дає метод Q-забарвлення (заснований на використанні флуоресцентних барвників). Малюнок сегментів при R-забарвленні є протилежним до такого, що спостерігається при G- і Q-забарвленнях: R-позитивні смуги відповідають GC-збагаченим ділянкам геному.

Існують також методи диференційного забарвлення, що дозволяють, наприклад, виявляти центромери та перицентромерний гетерохроматин (С-забарвлення) або ядерцеві організатори — місця локалізації кластерів генів рРНК - (ЯОР-забарвлення), які знаходяться у людини в ділянках вторинної перетяжки коротких плечей акроцентричних хромосом (хромосоми 13–15, 21 і 22). Результати аналізу геному людини свідчать про надзвичайно високу подібність послідовностей ДНК у різних індивідуумів: відповідно до останніх оцінок, будь-які дві людини є ідентичними за нуклеотидними послідовностями на 99,5 %, тобто вся сукупність різноманітних фенотипів у людини зумовлюється варіаціями лише 0,5 % геному. Першими, ще задовго до появи методів секвенування ДНК, були описані варіанти геному людини, що характеризуються зміною кількості та структури хромосом.

Анеуплоїдії та хромосомні перебудови часто асоційовані з різними захворюваннями, але гетероморфізм гомологічних хромосом (наявність у них сегментів, які відрізняються за розміром, морфологією або особливостями забарвлення) не обов'язково пов'язаний з патологічними змінами фенотипу й достатньо часто зустрічається в людській популяції (хромосомний поліморфізм). Такі великі геномні зміни, що торкаються послідовностей ДНК розміром понад 3 млн пар основ, визначають як мікроскопічні структурні варіанти.

У процесі аналізу нуклеотидних послідовностей людини було виявлено численні "дрібномасштабні" варіанти геномних послідовностей, розмір яких не перевищує 1 тис. пар основ. Приблизно 80 % таких ДНК-варіантів представлено у вигляді поодиноких нуклеотидних замін (SNP). За приблизними оцінками в людській популяції одна нуклеотидна заміна зустрічається на кожні 300 нуклеотидних пар. Крім SNP до "дрібномасштабних" варіантів геному відносять невеликі делеції, дуплікації та інверсії, варіації кількості мікро- та мінісателітних повторів. SNP і дрібні перебудови також спостерігаються в мітохондріальному геномі.

На початку XXI сторіччя були розроблені нові підходи до ефективного аналізу структури геному, які дозволили виявити варіації нуклеотидних послідовностей із проміжною довжиною між 1 тис. та 3 млн пар основ. Такі геномні варіанти отримали назву субмікроскопічних структурних варіантів. Субмікроскопічні структурні варіанти в основному представлені інверсіями та

CNVs (Copy-Number Variants) - варіантами за кількістю копій елементів послідовності порівняно з "еталонним геномом". Нині описано декілька сотень субмікроскопічних структурних варіантів.

Під час дослідження індивідуальних геномних варіацій основну увагу приділяють SNP та іншим "дрібномасштабним" геномним варіаціям. Вважається, що саме вони зумовлюють різноманітність фенотипових варіантів, а також можуть бути відповідальними за схильність до деяких хвороб за дією несприятливих факторів середовища. Очевидно, що більша частина "дрібномасштабних" варіацій спостерігається в некодуючих ділянках геному (яких просто значно більше), і такі варіації не впливають на структуру білків. Тому аналіз здебільшого зосереджують на екзонних варіантах, особливо на несинонімічних нуклеотидних замінах. Проте накопичуються дані про роль синонімічних замін і замін у некодуючих ділянках у деяких фенотипових проявах.

Крім установлення зв'язку поліморфізм - фенотип, "дрібномасштабні" варіації геному використовуються для потреб медичної генетики (при виявленні носіїв рецесивних мутантних генів), популяційної генетики людини (визначення спорідненості між популяціями людей) і судово-медичної експертизи при встановлені особистості або батьківства (у цьому випадку широко використовують поліморфізм за міні- та мікросателітними локусами). Міні- та мікросателіти є гіперваріабельними тандемними повторами. Їхня різноманітність настільки висока, що кожна людина (крім монозиготних близнюків) має індивідуальний набір варіантів цих послідовностей. За цими наборами одну людину можна відрізнити від іншої так само легко, як і за відбитками пальців, тому методика ідентифікації особистості за міні- та мікросателітними повторами отримала назву ДНК-фінгерпринтингу.

Тож, повторювальні послідовності у геномі людини можна об'єднати в кілька груп:

- Високочастотні повтори (сателітна ДНК)
- Помірні частотні повтори (малокопійні повтори)
- Інвертовані повтори
- Міні та мікросателітні повтори

Сателітна ДНК – це послідовність з 5-5000 п.н. яка повторюється мільйони разів і утворює кластери переважно у центромерних та теломерних ділянках хромосом.

Помірні частотні повтори не утворюють кластерів і чергуються з унікальними послідовностями.

Довгі диспергувані повтори (5000-7000 п.н.) обмежені з обох боків прямими повторами (300-600 п.н.).

Короткі диспергувані повтори активно транскрибуються як компоненти інтронів або як самостійні елементи.

Інвертовані повтори складаються із двох ідентичних копій довжиною 300 п.н., розташованих у зворотній послідовності стосовна одна одної на відстані від нуля до десяти тисяч п.н. Вони складають до 5% геному, і

щонайменше третина з них є паліндромами, що входять до складу регуляторних елементів геному.

Міні- та мікросателітні послідовності – це розсіяні по геному короткі тандемні повтори.

2. Тандемно повторювані послідовності

Основним класом тандемно повторюваних послідовностей є центромерна ДНК. Найпоширеніший тип центромерної ДНК — альфоїдна ДНК. ЇЇ сегменти мають довжину приблизно 170 п.н. Ці одиниці утворюють ряди, довжина яких від 250 т.п.н до 5 млн п.н. і вони складають приблизно 3% генома. Всередині одного ряду повтори неідентичні, вони трохи відрізняються, і між хромосомами ці відмінності ще суттєвіші. Центромерна ДНК утворює центромери — складні утворення, що, окрім ДНК, містять білки кількох типів, до яких під час поділу приєднуються нитки веретена поділу.

Тандемно повторювані послідовності також знайдені в теломерах, що розташовуються на кінцях хромосом. У людини теломерними послідовностями є GGGTTA, в різних хромосомах вони повторюються від 250 до 1500 разів. Останнім часом теломерна ДНК привертає увагу через знайденим зв'язком між скороченням теломер та старінням. У соматичних клітин та у культурі клітини мають обмежену здатність до розмноження через, у тому числі, і скорочення теломер. Однак у зародкових (статевих) клітинах, і стволових клітинах та у деяких ракових клітинах є фермент теломераза, здатний відновлювати теломерні послідовності, які за його відсутності скорочуються під час кожної реплікації ДНК. Теломераза це незвичайний фермент який складається з двох компонентів: РНК та білок. Саме РНК слугує матрицею для відновлення втрачених теломерних повторів.

Деякі інші класи тандемно повторюваних послідовностей — це гени некодуючих РНК. Найбільш відомими з них є гени рибосомальної РНК. У геномі людини є 5 груп таких генів, кожна з яких містить близько 60 копій. Вони розташовані в коротких плечах акроцентричних хромосом 13, 14, 15, 21 та 22. Ці кластери для рРНК разом із кількома додатковими ділянками ДНК отримали назву ядерцевих організаторів, оскільки кожен із них може утворити ядерце, розташоване в ядрі. Існує понад 80 типів рибосомальних білків. Вони синтезуються в цитоплазмі та мігрують у ядро, де зв'язуються з рРНК, що утворюється в ядрі. Крім того, у зборці рибосом беруть участь невеличкі некодуючі РНК кількох типів. В хромосомі 1 розташований ще один кластер тандемних повторів для іншого важливого компонента рибосом – 5S-рРНК.

3. Диспергувані повторювані послідовності

Диспергувані повторювані послідовності зазвичай розташовуються в геномі окремо, а не в кластерах. За розмірами їх поділяють на дві групи: довгі (LINE) та короткі (SINE). Обидва класи це рухомі (мобільні) генетичні елементи, що мають назву ретротранспозони. Функціонуючий

ретротранспозон здатен розмножувати себе або подібні послідовності. Ретротранспозони можуть виникати як ретровіруси.

Одним з найважливіших класів диспергованих повторювальних послідовностей є LINE-1 або група L1, яка представлена у геномі людини у кількості до 500000 копій та складає приблизно 15% всього геному. Більшість елементів L1 — це скорочені копії повних одиниць, довжина яких становить близько 5000 п.н., проте кілька тисяч елементів L1 мають повну довжину. Тільки 40–50 з них функціонують; вони кодують кілька білків, які здатні викликати транспозицію або самого L1, або деяких інших мобільних елементів. Елементи L1 містять два гени, що мають назву відкритих рамок зчитування (ORF). ORF1 кодує білок, який зв'язується з нуклеїновими кислотами; ORF2 кодує і зворотну транскриптазу (цей фермент використовує мРНК як матрицю для синтезу одноланцюгової ДНК, кДНК) та ендонуклеазу, яка робить розрізи геномної ДНК, куди може вбудуватися кДНК.

Коли мРНК з елемента L1 транслюється, білки, що утворилися, можуть безпосередньо зв'язуватися зі своєю мРНК. Такий комплекс білка з РНК переміщується в ядро, де ендонуклеаза розрізає один із ланцюгів ДНК, утворюючи вільний кінець. Зворотна транскриптаза використовує цю ділянку для синтезу затравки та утворення ДНК-копії мРНК елемента L1. Врешті утворюється другий ланцюжок кДНК і дволанцюгова молекула вбудовується в хромосому на місці одноніткового розриву.

Вважається, що зворотня транскріптаза відповідає також за утворення процессованих псевдогенів, що є ДНК-копіями мРНК та вбудовуються у місця негомологічні вихідному гену. Ці псевдогени не містять інтронів і зазвичай не експресуються у вигляді поліпептидів. Це відбувається через те, що вони або не містять регуляторних послідовностей, або мають мутації. Псевдогени доволі розповсюджені та становлять 0,5–1% геному людини.

Найбільший клас елементів SINE складається з послідовностей Alu. В геномі людини міститься біля мільйону послідовностей Alu, що дорівнює приблизно 10–12% всієї ДНК. Довжина основної одиниці становить приблизно 300 п.н., але в класі Alu існують і інші послідовності. Вони переважно містяться між генами та всередині інтронів, але інколи можуть входити до складу мРНК. Послідовності Alu не кодують білків і тому не здатні самі переміщуватись. Однак багато послідовностей Alu транскрибуються та мають на кінцях короткі послідовності подібні до РНК у елементів L1. І саме тому поки що існує непідтверджена думка, що ферменти, які продукують елементи L1, беруть участь у ретротранспозиції послідовностей Alu.

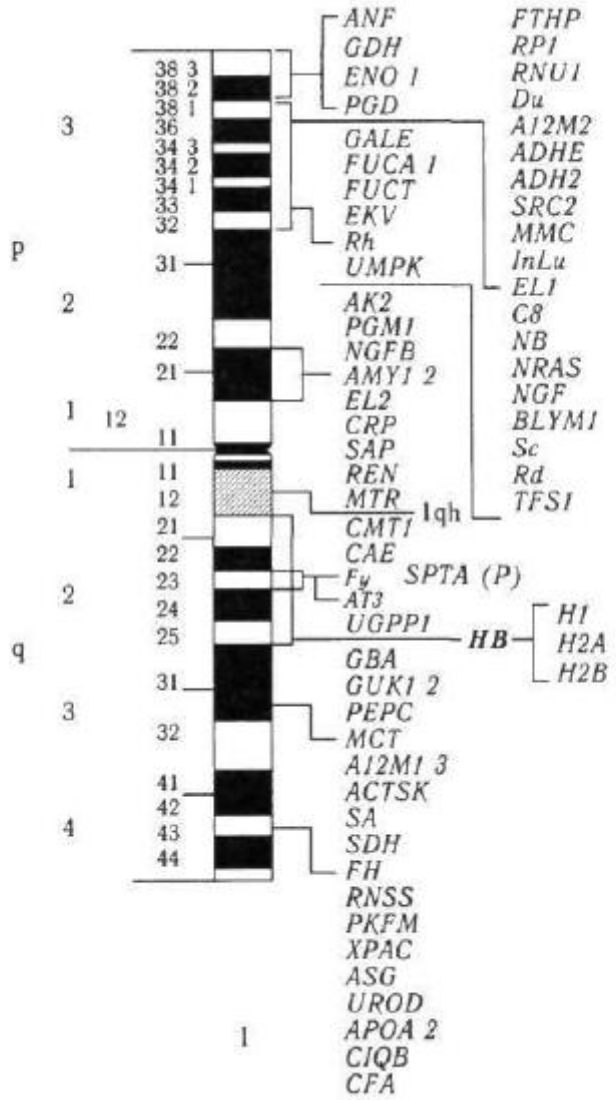
Вбудова нового мобільного елемента в ДНК потенційно здатна порушити функцію гену. У людини відомо близько 30 захворювань, спричинених ретротранспозицією. Більш того, наявність великої кількості споріднених послідовностей може призвести до втрати або дуплікації генетичного матеріалу під час мейозу та викликати появу аномальних фенотипів. Повторювальні послідовності в ДНК надають геному еволюційної пластичності та є важливим джерелом його змін.

ГЕНОМ ЛЮДИНИ ТА МЕТОДИ ЙОГО ВИВЧЕННЯ

4. Картування генів людини

Картування генів людини – це надзвичайно складне завдання, яке, проте, трохи полегшується тим, що деякі гени зібрані в кластери. Для картування генів людини використовують різні методи, проте особливо ефективними виявилися методи генетики соматичних клітин. Дуже часто їх поєднують з іншими підходами – вивченням родоводів, використанням мічених зондів, гібридизацією нуклеїнових кислот та іншими.

Сьогодні генетична карта хромосоми 1 є найдетальнішою.



Цитологічна карта генів хромосоми 1 людини
(*Ф. Фогель, А. Мотульскі*)

Продукти деяких генів хромосоми 1 легко визначити у окремої людини, і тому, їх можна використовувати для генетичного аналізу родин.

Величина геному людини у сантиморганах становить 3000, тобто на одну хромосому припадає приблизно 130 сМ. Сучасні методи забарвлення

дозволили виявити до 1000 смуг на усіх хромосомах. Тож одна смуга, виявлена цитогенетично, відповідає приблизно 3 сМ.

Генно-інженерні методи мають вищу роздільну здатність. Застосовуючи ці методи можна картувати послідовності ДНК довжиною від декількох до 100000 п.н., що складає приблизно 0,1 сМ.

Досить успішно відбувається картування генів Х-хромосоми. Ген, відповідальний за кольорову сліпоту, було локалізовано ще у 1911 році. Особливості успадкування дали змогу локалізувати понад 100 локусів у Х-хромосомі.