Геном людини

План:



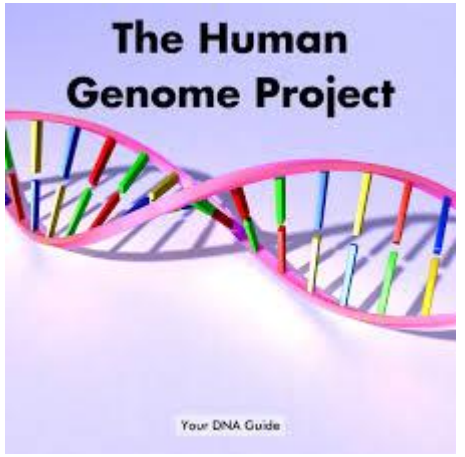
1. Загальна характеристика геному людини
2. Тандемно повторювані послідовності
3. Диспергувані повторювані послідовності
4. Картування генів людини

Загальна характеристика геному людини

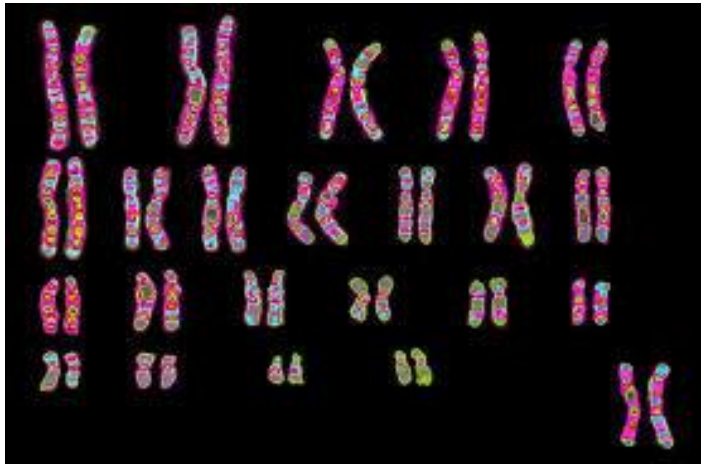








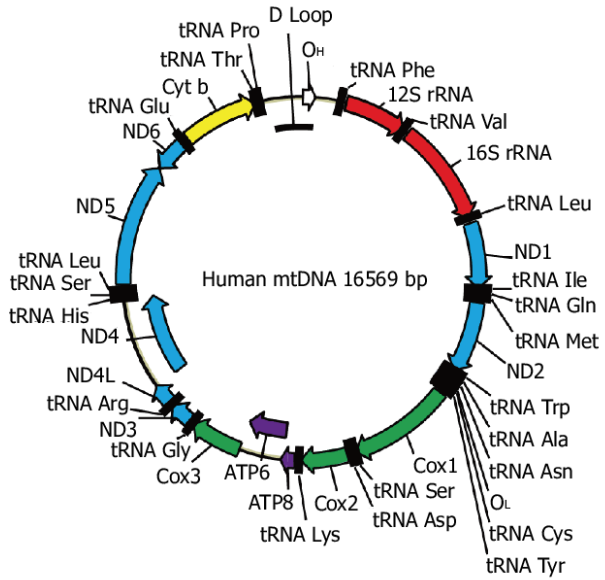
Проект розшифрування геному людини (Human Genome Project) був одним із найамбітніших наукових проектів XX століття. Він тривав 13 років і закінчився у 2003 р. Проект мав на меті побудувати достатньо точну генетичну карту (з кроком 2-5 сМ), отримати послідовність 3 млрд пар основ ДНК, ідентифікувати генні послідовності, визначити геномні варіації. У результаті виконання проекту побудована генетична карта з кроком у 1 сМ, отримані 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 є напевно заниженою). Загальна довжина кодуючих ділянок генів дорівнює приблизно 34 млн пар основ, тобто становить 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 сім 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. Найпоширенішим методом диференційного пофарбування є метод G-забарвлення, який дозволяє отримати хромосоми, нерівномірно пофарбовані по довжині. Малюнок чергування темних і світлих смуг є унікальним для кожної пари гомологічних хромосом. Цікаво, що темні та світлі смуги асоційовані з ділянками, збагаченими на AT- і GC-пари відповідно (тобто, відповідно, з геномними зонами, збагаченими на мобільні елементи та кодуючі послідовності). Практично таку саму картину чергування смуг дає метод Q-забарвлення (базується 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 останнім оцінкам, будь-які дві людини є ідентичними за нуклеотидними послідовностями на 99,5 %, тобто вся сукупність різноманітних фенотипів у людини зумовлюється варіаціями тільки 0,5 % геному. Першими, ще задовго до появи методів секвенування ДНК, були описані варіанти геному людини, які характеризуються зміною кількості й 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 були виявлені 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 спостерігаються також у мітохондріальному геномі.

На початку XXI сторіччя були розроблені нові підходи ефективного аналізу структури геному, які дозволили виявити варіації нуклеотидних послідовностей, що мають проміжну довжину 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.

При дослідженні 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 дії 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 в основному зосереджують на екзонних варіантах, особливо на несинонімічних нуклеотидних замі-нах. 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 виявлені 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 даних 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 та чергуються з унікальними послідовностями.
- ✓ Довгі дисперговані повтори (5000-7000 п.н.) обмежені з обох боків прямими повторами (300-600 п.н.).
- ✓ Короткі дисперговані 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 не менше третини з них є паліндромами, що входять до складу регуляторних елементів геному.
- ✓ Міні- та мікросателітні послідовності – це розсіяні по геному короткі тандемні повтори.



Тандемно повторювані послідовності

Основним класом тандемно повторюваних послідовностей є центромерна ДНК. Найбільш розповсюджений тип центромерної ДНК – це 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, які є складними утвореннями, що окрім ДНК містять білки кількох типів до яких приєднуються нитки веретена поділу під час поділу.

Тандемно повторювані послідовності знайдені також 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 для 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 реплікаціїДНК. 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 повторювальних послідовностей – це гени для некодуючих РНК. Найбільш відомими з них є гени для рибосомальної РНК. В геномі людини є п’ять груп таких генів, кожна з яких містить біля 60 копій. Вони розташовані в коротких плечах акроцентричних хромосом 13, 14, 15, 21 та 22. Ці кластери для рРНК разом з кількома додатковими ділянками ДНК отримали назву ядришкових організаторів, тому що кожен з них може утворити ядришко, що розташовується у ядрі. Існує більш ніж 80 типів рибосомальних білків. Вони синтезуються у цитоплазмі та мігрують в ядро, де зв’язуються з рРНК, яка утворюється в ядрі. Крім того у зборці рибосом приймають участь невеличкі неккодуючі РНК кількох типів. В хромосомі 1 розташован ще один кластер тандемних повторів для іншого важливого компоненту рибосом – 5S-рРНК.

Диспергування повторювані послідовності



Диспергованні повторювані послідовності зазвичай розташовуються в геномі окремо, а не кластерами. Згідно з 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 сам 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 500000копій та складає приблизно 15% всього геному. Більшість елементів L1 це скорочені копії повних одиниць, довжина яких складає біля 5000 п.н., проте кілька тисяч елементів L1 мають повні довжину. Тільки 40-50 з них функціонують, вони кодують кілька білків, які здатні викликати транспозицію або саме L1, або деяких інших мобільних елементів. Елементи L1 містять два гена, що мають назву відкриті рамки зчитування (ORF). ORF1 кодує білок який зв'язується з нуклеїновими кислотами, ORF2 кодує і зворотню транскріптазу (цей фермент використовує мРНК як матрицю для синтезу одноланцюгової ДНК, кДНК) та ендонуклеазу, яка робить розрізи геномної ДНК куди може вбудуватись кДНК.

Коли мРНК з елемента L1 транслюється, то білки що утворились можуть безпосередньо зв'язуватись зі своєю мРНК. Такий комплекс білка з РНК переміщується в ядро, де ендонуклеаза розрізає один з ланцюгів ДНК і утворюється вільний кінець. Зворотня транскриптаза використовує цю ділянку для синтезу затравки та утворєю ДНК-копію 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 в геномі людини складають 0,5-1%.

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 складає приблизно 300 п.н., але в класі Alu існують і інші послідовності. Вони знаходяться переважно між генами та в середині 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, існує поки що не підтверджена думка, що ферменти які продукують елементи L1 приймають участь в ретротранспозиції послідовностей Alu.

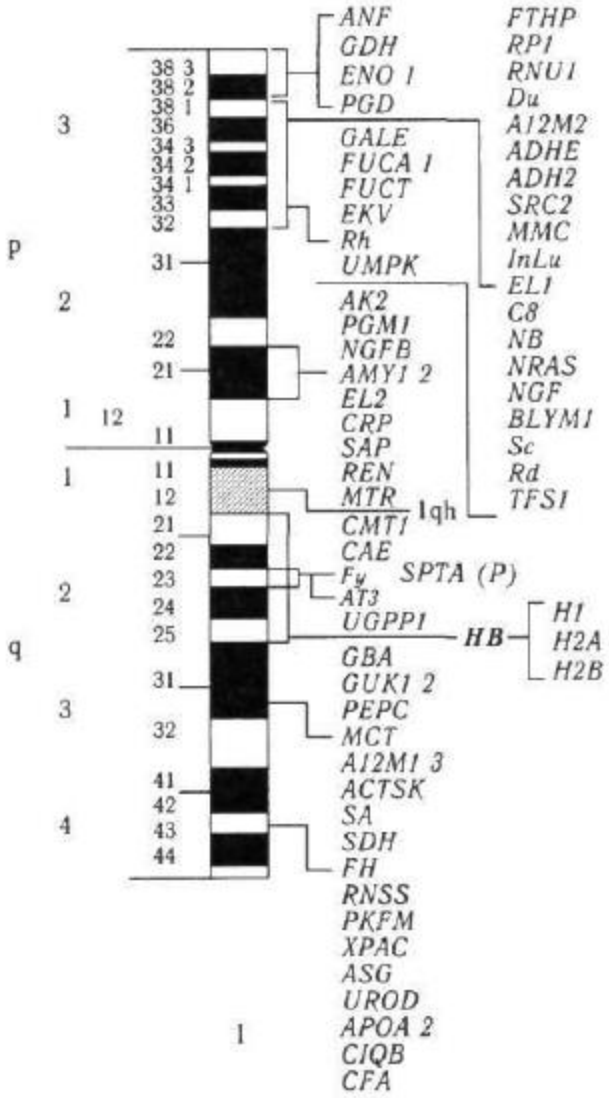
Вбудова нового мобільного елемента в ДНК потенційно здатна порушити функцію гену. У людини відомо біля 30 захворювань, що викликані ретротранспозицією. Більш того, наявність великої кількості споріднених послідовностей може призвести до втрати або дуплікації генетичного матеріалу під час мейозу та викликати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



Картування генів людини

Картування генів людини – це надзвичайно складне завдання, яке проте трохи полегшується через те, що деякі гени зібрані в кластери. Для картування генів людини використовують різні методи, проте особливо ефективними виявились 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 найбільш детальною.


ANF
GDH
ENO 1
PGD
GALE
FUCA 1
FUCT
EKV
Rh
UMPK
AK2
PGM1
NGFB
AMY1 2
EL2
CRP
SAP
REN
MTR
CMT1
CAE
Fy
AT3
UGPP1
GBA
GUK1 2
PEPC
MCT
A12M1 3
ACTSK
SA
SDH
FH
RNSS
PKFM
XPAC
ASG
UROD
APOA 2
CIQB
CFA
FTHP
RP1
RNU1
Du
A12M2
ADHE
ADH2
SRC2
MMC
InLu
EL1
C8
NB
NRAS
NGF
BLYM1
Sc
Rd
TFS1
1qh
SPTA (P)
HB
H1
H2A
H2B
p
q
38 3
38 2
38 1
36
34 3
34 2
34 1
33
32
31
22
21
12
11
23
24
25
41
42
43
44



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

Величина геному людини у сантиморганах складає 3000, тобто на олну 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.

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

Досить успішно відбувається картування генів Х-хромосоми. Ген відповідальний за кольорову сліпоту був локалізований ще у 1911 році. Особливості успадкування дали можливість локалізувати в Х-хромосомі понад 100 локусів.

Дякую

