Лекція 3 Особливості геному еукаріот

Розмір геному зростає на основних етапах прогресивної еволюції (збільшення складності організмів): у бактерій кількість ДНК на клітину становить 0.001-0.008 пкг, у дріжджів - 0.02-0.04 пкг, у безхребетних тварин - 0.1-12 пкг, у хребетних - 0.8-280 пкг. В іншій еволюційній гіці, у вищих рослин вміст ДНК в клітині знаходиться в межах 0.4-180 пкг. у вірусів розмір генома знижується до 0.000001- 0.00003 пкг ДНК.

Однак після досягнення певного рівня організації (хребетні, комахи, судинні рослини) розмір генома перестає корелювати зі складністю організму. Так, наприклад, серед хребетних найбільший розмір генома спостерігається у хвостатих амфібій і дводишніх риб. Навіть у безхвостих амфібій геном в середньому більше, ніж у амніот, а середній геном рептилій більший за геном птахів. У дрозофіли, представника одного з найбільш сучасних відділів комах (двокрилих), розмір генома набагато менший, ніж у цвіркуна, що належить до стародавнього відділу прямокрилих, і навіть менше, ніж у деяких кишковопорожнинних і поліхет, тобто представників більш примітивних типів тварин. У покритонасінних рослин розмір генома в середньому не більше, ніж у голонасінних. Крім того, розміри генома можуть значно відрізнятися навіть у філогенетично близьких видів. Ці факти дозволяють говорити про феномен розміру генома. Тобто прямої кореляції між кількістю ДНК і еволюційної просунутістю організму немає.

Особливості геному еукаріот:

1. «Надлишковість»

Підвищений вміст ДНК в геномі еукаріот можна пояснити одним лише збільшенням потреби цих організмів у додатковій генетичній інформації в зв'язку з ускладненням організації, оскільки велика частина їх геномної ДНК, як правило, представлена некодуючими послідовностями нуклеотидів.

На ~ 10^6 пар нуклеотидів у бактерій припадає ~ 5 тис. генів. На ~ 10^9 пар нуклеотидів у ссавців ~ 50 тис. генів.

Мінуси "надлишкової" ДНК:

- збільшення часу синтезу ДНК;
- складніше організовувати подвоєння ДНК;
- висока енергоємність - на 1 нуклеотид для включення в ланцюг ДНК потрібно затратити ~ 60 молекул АТФ.

Невизначений наслідок: - завдяки залежності розміру ядра від кількості ДНК відбувається збільшення розмірів клітини.

Плюси "надлишкової" ДНК:

- виникає можливість створення складного регуляторного апарату, що дозволяє підняти організм на більш високий еволюційний рівень.

Причини надлишковості:

1. Великий розмір генів (за рахунок наявності інтронів).

2.Наявність повторюваних послідовностей. Повторюються і гени, і не кодуючі ділянки. У еукаріот деякі послідовності повторені сотні і тисячі разів.

3. Наявність великої кількості некодуючих послідовностей, частина з яких виконує регуляторну функцію при транскрипції, а частина - необхідна для компактизації генома.

2. Компактність.

При середній різниці розмірів геномів на 3 порядки, лінійні розміри еукаріотичних хромосом порівняні з довжиною ДНК прокаріот. Виділяють, принаймні, 4 рівня компактизації ДНК. При цьому нитка ДНК "коротшає" в 10000 разів.

Два перших рівня компактизації еукаріотичного генома забезпечуються гістонами. Гістони - основні білки, збагачені лізином і аргініном - позитивно зарядженими амінокислотами. Виділяють 5 фракцій гістонів. Напрацьовується їх дуже багато - 60 млн. молекул кожної фракції на клітку.

Фракція	Лізин	Аргінін	Ліз/арг	Мол. вага (Да)
H1 (дуже багата на лізин)	9%	1%	>20	3000
H2B (помірно багата на лізин)	6%	6%	~2.5	3774
H2A (помірно багата на лізин та аргінін)	1%	9%	~1	3960
H4 (багата на аргінін та гліцин)	1%	14%	~0.8	1282
H3 (дуже багата на аргінін); в ній присутній цистеїн, а в інших - ні	0%	13%	~0.7	5348

Всі гістони, крім H1, надзвичайно консервативні в еволюційному відношенні (у корови і конюшини різниця в H2A всього в одну амінокислоту). Отже, ці білки виконують принципову функцію, яка у всіх еукаріот забезпечується однаково.

Будь-яка мутація в гістонових генах летальна.

H1 - дуже варіабельна фракція. Цей гістон різний не тільки у видів, але навіть у одного організму, в залежності від стадій онтогенезу.

Позитивно заряджені амінокислоти гістонів забезпечують електростатичні взаємодії з ДНК. Центральна частина (містить гідрофобні амінокислоти) необхідна для взаємодії гістонів між собою.

Чотири рівня компактизації ДНК

1. Нуклеосомний.

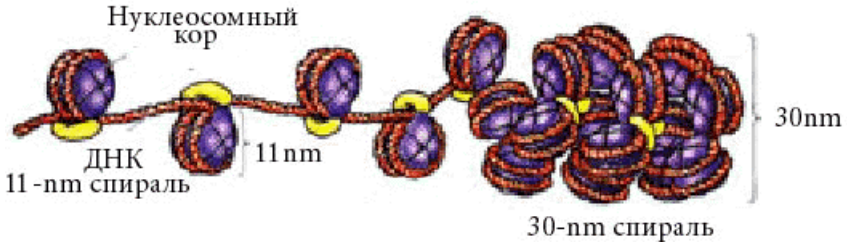
Нуклеосомою називається повторюваний структурний елемент хроматину, що містить гістоновий октамер і ~ 180 п.н. ДНК. В основі нуклеосоми лежить гістоновий октамер.

Розташування гістонів не випадкове. Кожна молекула представлена двічі. Вони утворюють кор (серцевину) нуклеосоми. На кор намотується ДНК - 1.75 лівих витки спіралі. Безпосередньо з октамер контактує 145 п.н. і 20-30-40 п.н. між нуклеосомними корами.

Нуклеосомний рівень упаковки властивий всій еукаріотичній ДНК, він дає вкорочення в 7 разів. Діаметр збільшується з 20 Å до 110 Å.

Гістонові октамери "ковзають" по ДНК. При реплікації знімається і цей рівень компактизації. При транскрипції нуклеосоми зберігаються.

2. Супербідний, або соленоїдний.



Фактично забезпечується Н1 гістоном.

Н1 взаємодіє з октамерами, зближує їх, і ще на нього намотується ДНК. Утворюється супербід.

Відбувається скорочення лінійного розміру ДНК в 6-10 разів. Діаметр збільшується до 300Å. Цей рівень компактизації, як і перший, не залежить від первинної структури ДНК.

3. Петльовий рівень.

Забезпечується негістоновими білками. Вони впізнають певні послідовності ДНК і зв'язуються з ними і один з одним, утворюючи петлі по 20-80 тис. Пар основ

Петля забезпечує експресію гена, тобто петля є не тільки структурним, але і функціональним утворенням.

Є ділянки, в яких немає петель. Скорочення за рахунок петель проходить в 20-30 разів. Утворюються і петльові домени. Діаметр збільшується до 700Å.

4. Метафазна хромосома.

Метафазна хромосома вже подвоєна. Вона складається з двох хроматид. Кожна з них містить одну молекулу ДНК.

Сюди входять білки ядерної ламіни, серія білкових ниток, сполучених з ядерною оболонкою і пронизуючих все ядро.

Модифікації гістонів дуже сильно впливають на компактизацію ДНК. Гістони можуть метилюватися, фосфорилюватися (по серину, треоніну, тирозину), тобто амінокислотні залишки легко модифікуються. Крім того, можливо алкілування і ацетилювання гістонів.

Геном вищих еукаріот **А-Т типу** (пари А-Т переважають), нижчих еукаріот - **Г-Ц типу**. У людини співвідношення (Г + Ц) / (А + Т) = 0.45. У різних типів бактерій діапазон співвідношення А-Т пар і Г-Ц пар великий. Чим більше в геномі А-Т пар, тим більше можливостей для зміни вторинної структури ДНК. При суперспіралізації ДНК А- Т багаті ділянки плавляться в першу чергу.

Фракціонування ДНК

В кінці 60-х років роботами американських вчених Р. Бріттена, Е. Девідсона і інших була відкрита фундаментальна особливість молекулярної структури генома еукаріот - нуклеотидні послідовності різного ступеня повторюваності. Це відкриття було зроблено за допомогою молекулярно-біологічного методу вивчення кінетики ренатурації денатурованої ДНК.

Основи методу ренатурації ДНК: ДНК обробляють ультразвуком. При цьому вона деградує на дволанцюгові шматки однакового розміру. Потім суміш денатурують і повільно охолоджують. При температурі на 20^0С нижче, ніж температура плавлення ДНК, йде відновлення вторинної структури (ренатурації). Якщо послідовності часто зустрічаються, то вони ренатурують швидше.

Розрізняють такі фракції в геномі еукаріотів.

1. *Часті повтори.* Це самокомплементарна ДНК (foldback DNA) і високо повторена ДНК (highly repetitive DNA). Самокомплементарна ДНК складається з паліндромних

послідовностей, здатних утворювати дволанцюгові шпильки, "замикаючись" самі на себе. Частота зустрічаємості на гаплоїдний геном більше 10^5. Зустрічальність паліндромів може бути набагато нижче.

Значення швидкості реассоціації у *паліндромів* таке ж, як і у швидких повторів, а зустрічальність, як у уніків або помірних повторів. У деяких організмів, наприклад, у черепах, 20% ДНК - паліндроми. В середньому у тварин від 2% до 12% генома припадає на паліндроми. У рослин - від 1% до 4% (у пшениці 3 млн. обернених повторів). Вони можуть містити від декількох десятків до десятків тисяч нуклеотидів. Найбільш часто паліндроми зустрічаються в регуляторних ділянках генів.

2. *Помірні повтори* (middle repetitive DNA) .Частота зустрічаємості на гаплоїдний геном більше 10, але менше 10^5.

3. *Унікальні послідовності* (single-copy DNA). Частота зустрічаємості менше 10 разів на геном.

Всі повтори можна поділити на: а) тандемно повторювані, тобто розташовані один за одним «голова до хвоста», і б) диспергованні по геному. Тандемні повтори зустрічаються між генами і в інтронах, але найбільше їх в центромерних і теломерних районах хромосом. При денатурації сумарною геномної ДНК і подальшої реассоціації довгі тандемні повтори становлять фракцію швидко ренатуруючої ДНК.

А) Добре вивченим представником перших є **сателітна ДНК**, яка складається з коротких тандемних повторів довжиною 5-200 п.о., організованих в довгі блоки (по 100-200 одиниць в блоці). Одними з перших серед повторюваних послідовностей ДНК еукаріот були відкриті сателітні ДНК тимуса телят. Свою назву вони отримали на підставі того, що при аналізі сумарної еукаріотичної ДНК центрифугуванням в градієнті щільності хлористого цезію вони супроводжували основний пік оптичної щільності у вигляді плеча (супутника, сателіта). Саме гомогенний нуклеотидний склад фракції сателітних ДНК, який визначається наявністю в ній численних коротких повторів, змінював її плавучу щільність, що легко виявлялося при центрифугуванні. У своєму класичному визначенні сателітних ДНК Р.Д. Бріттен відзначав, що сателіти - це мінорний компонент ДНК, що відділяється від основної ДНК при рівноважному ультрацентрифугуванні в градієнті щільності CsCl. Зміст сателітної ДНК в геномі еукаріот може досягати 5-50% від сумарної кількості ДНК.

Якщо розділяти геномну ДНК на фракції центрифугуванням в градієнті щільності, крім основної смуги (у людини 1,701 г\см$^{-3}$), утворюються ще кілька легших смуг (у людини 1,687, 1,693 і 1,697 г\см$^{-3}$). Хімічно ці смуги легші, так як містять нетиповий для геному в цілому вміст GC- послідовностей (для генома людини це 40,3%).

Послідовності повторюваних одиниць сателітної ДНК (сатДНК) видоспецифічні. Так у *Drosophila melanogaster* є чотири основні типи сателітних послідовностей, що складаються з 5, 7, 10 і більшого числа нуклеотидів, разом складаючих 16% геномної ДНК. ДНК *Drosophila virilis* має три основних сателіта, кожен з яких складається з 7 пар основ, не суттєво відрізняються між собою за складом азотистих основ.

У деяких видів ці повтори становлять більшість геномної ДНК. Наприклад, у кенгурового щура (*Dipodomys ordii*) більше 50% всього геному складається з трьох повторених послідовностей: AAG (2.4 млрд. копій), TTAGGG (2.2 млрд. копій) і ACAAGCGGG (1.2 млрд. копій).

Зазвичай в сателітних послідовностях відсутній один з нуклеотидів. Звідси випливає, що ця ДНК не може бути кодуючою, вона ніколи не транслюється.

У місцях розташування сателітної ДНК можлива максимальна компактизація. В конститутивному гетерохроматині всі чотири рівня упаковки ДНК представлено в інтерфазі. Вважається, що сателітні ДНК відіграють важливу роль у підтримці структур хромосом і, можливо, в їх спарюванні в процесі мейозу. За сателітною ДНК відбувається кросинговер між гомологічними хромосомами і працює в якості сайтів зв'язування для центромерних білків.

Сателітні ДНК людини поділяються на дві групи. Представники *першої групи* - це прості послідовності невеликих розмірів. Сателіти цього типу називають «класичними», вони розташовані в прицентромерних районах хромосом. До них відносяться сателіти 1, 2, 3. **Сателіт 1** - це відрізок з 42 пар основ, він виявлений в метацентричних хромосомах 3 і 4, а також в акроцентриках 13, 14, 15, 21, 22. **Сателіт 2** - малоконсервативний пентамерний повтор. Масиви сателіта цього типу присутні в хромосомах 1, 2, 10, і 16. **Сателіт 3** людини - пентамерний повтор, що чергується з декамером на хромосомах 1, 5, 9, 10, 13, 14, 15, 17, 20, 21, 22, Y.

До *другої групи* належать блоки сатДНК, що складаються з відрізків 100-200 п.н. Вони вбудовуються в області формування кінетохора в центромерному районі хромосом (α-сателітна (альфоїдна) ДНК, що складається зі слідуючих один за іншим мономерів довжиною 170-171 п.н). або впритул прилягають до нього (перицентромерний район).

Хромосом-специфічні послідовності сателітної альфоїдної ДНК знайшли широке застосування в молекулярній цитогенетиці як ДНК-зондів, зручних для маркування індивідуальних хромосом в метафазних і інтерфазних клітинах (багатокольоровий бендінг хромосом).

Мікро- (1-13 п.о. в основному повторювальному блоці дають 150 п.н.) і **мінісателітні** (з більшим ніж 25 п.о. в індивідуальному повторі, число яких до декількох тисяч) ДНК характеризуються високою варіабельністю за кількістю копій в геномах організмів навіть одного виду і в ряді випадків володіють генетичною нестабільністю як в нормі, так і при деяких патологічних станах організмів. Завдяки цій властивості міні- і мікросателіти часто називають тандемними повторами із змінним числом копій VNTR (variable number of tandem repeats).

З мікросателітних повторів найбільш часто в геномі людини зустрічається СА повтор, кількість копій якого в геномі - 140000 і мононуклеотидний повтор А - 120000 копій. Блоки з цих повторів можна виявити, приблизно через кожні 30 тис. пар основ Виникають мікросателітні повтори внаслідок помилок в репликаційному процесі і їх функція до кінця не з'ясована. Однак мікросателітні повтори є мінливими (поліморфними) - у різних представників біологічного виду число повторюваних одиниць в масиві різне. Ніякі дві особини (крім близнюків) не мають в точності однакової комбінації варіантів довжини всіх мікросателітів. Якщо дослідити достатню кількість мікросателітів, то для кожної людини може бути побудований унікальний **генетичний профіль**. Мікросателіти успішно використовують для встановлення еволюційних і родинних зв'язків в генетичному аналізі.

Мінісателітні повтори характеризуються більшою довжиною повторюваного елемента (більше 25 пар основ). До мінісателітних повторів формально можна віднести і теломерні повтори, проте вони, на відміну від інших, диспергованих по геному, розташовані кластером і, очевидно, грають певну структурну роль. Зазвичай кількість теломерних повторів становить від 250 до 1500 - загальною довжиною приблизно 9 т.п.н.

Повтори строго однакові, але можуть відрізнятися для різних видів. Так, у людини основний елемент теломерних повторів представлений гексамером 5-TTAGGG-3.

Б) Інший тип повторів - диспергованні повторювані послідовності ДНК, які не організовані у великі блоки, а розсіяні по геному. Відносяться вони до фракції помірних (по частоті) повторів.

Існує кілька родин широко поширених диспергованих повторів. Короткі диспергованні повтори **SINE** (від англ. Short interspersed nucleotide element) довжиною 100-500 п.н. мають по краях короткі прямі повтори і po1y(A) - послідовність на 3'-кінці. Найбільш відомий повтор цього типу - Alu-повтор. Його довжина - 300 п.н., він складається з двох мономерів і має poly(A) -хвіст на 3'-кінці. У людини на гаплоїдний геном доводиться 10^5-10^6 копій ALU-послідовностей. Частота повторів у людини - 1.5 * 10^6, родин - 3, становлять 13% генома. Ці повтори зустрічаються між генами, в інтронах, в сателітній ДНК і в будь-якій орієнтації відносно один одного.

Поряд з різними SINE-родинами повторів геноми ссавців містять родину **LINE**-повторів (від англ. Long interspersed nucleotide element). LINE- повтор являє собою один з варіантів ретротранспозонів, в структурі якого є ген зворотної транскриптази, але немає довгих кінцевих повторів LTR по флангах від кодуючої області. Одиниці повтору у різних членів родини мають розмір не більше 6-7 т.п.н .; вони відрізняються по нуклеотидній послідовності. Частота копій різних членів родини LINE людини 8.5 * 10^5, число родин 3, всього 21% генома людини.

Так само як і сателітні ДНК, SINE- і LINE-повтори характеризуються генетичною нестабільністю. Їх загальними рисами є здатність до транспозиції. Є експериментальні дані про те, що нові копії повторюваних елементів обох типів виникають в геномі в результаті функціонування механізму, названого *ретротранспозицією*, або *ретропозицією*. За участю подібного механізму під дією зворотної транскриптази спочатку утворюється кДНК на матриці РНК-транскрипту відповідного повтору, яка далі інтегрується в новий локус генома, як це має місце у ретровірусів. Такий механізм дає можливість локально змінювати число копій певних послідовностей нуклеотидів в еукаріотичному геномі.

В цілому, до *помірних повторів* відносять як ті що транскрибуються і транслюються, так і тільки ті що транскрибуються, але нетрансльовані послідовності ДНК і регуляторні ділянки.

Крім того, припускають, що в геномі ссавців деякі родини помірних повторів мають вірусну природу. У геномі еукаріотів знаходять також мобільні генетичні елементи, аналогічні прокаріотичним IS- і Tn-елементам. Велика частина вивчених до теперішнього часу генів у еукаріот не є чимось унікальним, а представлені або безліччю подібних копій, або обмеженим набором дуплікованих і дивергованиxx в різній мірі послідовностей, так званими мультигенними родинами. Повторюваність генів в мультигенних родинах необхідна або для посилення дози генів (дозова повторюваність), або для забезпечення синтезу різних варіантів родинних поліпептидів (варіантна повторюваність). Дозова повторюваність характерна для генів, що кодують тРНК і рРНК, а також гістонових генів. Варіабельна повторюваність властива таким генам, як глобінові, актинові, імуноглобулінові і багатьом іншим.

Помірні повтори	
Гени	*Регуляторні ділянки*

Транскрибуємі та транслюємі Гени білків рибосо, гістонові гени, гени мембранних білків, цитоскелетних білків, гени імуноглобулінів.	*Транскрибуємі, але нетранслюємі* Гени рРНК, тРНК (в середньому повторюються в геномі 5 тис. разів)	енхансерні модулі, ori реплікації, промотори та термінатори транскрипції

Унікальні гени

Багато генів еукаріот повторені в геномі від декількох одиниць до декількох сотень разів і утворюють мультигенні родини. Ці гени зазвичай згруповані в кластери в певних районах однієї, або декількох хромосом. У багатьох мультигенних родинах поряд з функціонально активними генами містяться псевдогени - мутаційно змінені послідовності, які не здатні транскрибуватися або продукують функціонально неактивний генний продукт. Прикладами мультигенних родин можуть служити гени рибосомних, транспортних і ядерних РНК, гени альфа- і бета-глобінів, тубулінів, міоглобіну, актину, інтерферону і багатьох інших. У ряді випадків, можлива виборча ампліфікація деяких родин генів в процесі їх експресії, як, наприклад, генів рибосомних РНК. При цьому число здатних транскрибуватися копій генів збільшується за рахунок їх виборчої ампліфікації в сотні і навіть тисячі разів, що супроводжується лавиноподібним наростанням частки відповідного генопродукту в клітинах.

Особливе місце серед мультигенних родин займають супергени - дуже великі кластери з сотень функціонально і структурно родинних генів, розташованих в сегментах окремих хромосом. Класичним прикладом супергена може служити HLA комплекс, контролюючий головні антигени гістосумісності. Він займає район більше 6000 кб на короткому плечі хромосоми 21 і складається з серії тісно зчеплених генів, відповідальних за синтез безлічі білків, що включають клітинні поверхневі антигени, молекули імунної відповіді і деякі компоненти комплементу. До супергенних родин відносяться три комплекси розташованих на різних хромосомах мультигенних, контролюючих синтез важких і легких ланцюгів імуноглобулінів. Цікаво, що в процесі диференціювання В лімфоцитів, які продукують імуноглобуліни, відбувається структурна перебудова цих родин. При цьому окремі послідовності ДНК елімінуються, тоді як інші зливаються, так що структура генів імуноглобулінів в зрілих В лімфоцитах значно відрізняється від вихідної, тобто від тієї, яка спостерігається в зародкових клітинах.

Однією з важливих структурних особливостей геному еукаріот є наявність так званих псевдогенів, унікальних послідовностей, дуже схожих за своєю структурою з певними нормальними генами, але в силу присутності в кодуючих послідовностях цілого ряду мутацій нездатних транскрибуватися або правильно транслюватися з утворенням структурно і функціонально активного продукту. Псевдогени виявлені для багатьох генів. Їх кількість варіює від однієї до кількох десятків копій на геном і в цьому випадку вони, як правило розташовані тандемно. Іноді псевдогени тісно зчеплені з нормальними генами, у багатьох випадках псевдогени і гени локалізовані в різних хромосомах. Для деяких моногенних захворювань ідентифіковані мутантні аллелі, подібні з мутаціями в відповідних псевдогенах. У цих випадках обговорюється можлива роль псевдогенів в спонтанному мутаційному процесі.

У геномі еукаріот присутні також нуклеотидні послідовності, гомологічні генам деяких вірусів. Вперше ці послідовності були ідентифіковані в геномі вірусів, які індукують розвиток пухлин у тварин і людини, і тому вони були названі онкогенами. Гомологічні послідовностям в геномі людини звуться протоонкогенами. В даний час вже ідентифіковано понад 100 протоонкогенів. Білкові продукти протоонкогенів, мабуть, грають важливу роль в нормальній проліферації клітин особливо на ранніх стадіях ембріонального розвитку, контролюючи клітинний цикл і вибір геномної програми розвитку клітини. При виникненні специфічних мутацій в протоонкогенах, а також при порушеннях регуляції їх роботи, що виражаються в гіперпродукції або в експресії в нетиповому місці або в невластивий момент життєдіяльності клітини, вони починають вести себе як онкогени, стимулюючи неконтрольоване розмноження і проліферацію певних клітинних клонів, що і може , в кінцевому рахунку, призвести до формування пухлини.

Близько 10-15% геному еукаріот (у рослин 3-10%) представлено унікальними транскрибуємими послідовностями, що складають основу структурних генів. В даний час в поняття "ген" включається не тільки його область, яка транскрибується - екзони + інтрони, але також фланкируючі послідовності - лідерна, що передує початку гена, і хвостова нетрансльована область, розташована на 3'кінці гена. На відміну від генів прокаріот гени еукаріот рідко представлені однієї безперервної послідовністю і в переважній більшості мають переривчасту структуру. Відносно короткі кодуючі ділянки - екзони, чергуються з довгими інтронами, що транскрибуються і входять до складу первинного РНК-продукту, але потім при процесингу первинного РНК-транскрипту вони вирізаються і не беруть участі в трансляції.

Згідно класичним уявленням ген - це локус на хромосомі, мутації в якому реалізуються на рівні фенотипу. У молекулярній біології ген трактується як асоційований з регуляторними послідовностями фрагмент ДНК, що відповідає певній одиниці транскрипції. Отже, уявлення про ген формальних генетиків далеко не повністю тотожний його фізичноій одиниці і співвідношення між цими двома поняттями досить заплутані. Відзначимо деякі причини цих суперечностей. Відомо, що мутації одного гена можуть призводити до абсолютно різних і навіть у ряді випадків до комплементарним фенотипам. Результати прямого секвенування генома свідчать про присутність в ньому значно більшого числа генів, ніж можна очікувати від результатів мутаційного аналізу. Одна і та ж послідовність ДНК в геномі може кодувати кілька різних білків, що досягається за рахунок так званого альтернативного сплайсингу (утворення різних мРНК з одного первинного РНК-транскрипту). У великих інтронах ряду генів виявлені смислові послідовності інших генів ("ген в гені"), що зчитуються в протилежному напрямку. Транскрипційні одиниці генома можуть перекриватися за рахунок наявності різних промоторів. Нарешті, завдяки соматичній рекомбінації структура транскрибуємих послідовностей деяких генів може бути різною в різних клонах клітин одного організму (Т-клітинні рецептори).

Ситуація з визначенням поняття "ген" ще більше ускладнюється, якщо в це поняття включати численні регуляторні послідовності. Виникає питання: "Як далеко від гена можуть розташовуватися ці послідовності, щоб їх можна було включати в структуру гена?". Для багатьох цілей виявляється зручним введене останнім часом поняття "рахований ген" - counting gene. Останній розглядається як окрема транскрибуєма одиниця ДНК або її частина, яка може транслюватися в одну або кілька взаємопов'язаних амінокислотних послідовностей. Тому послідовність, що дає дві транскрипційні одиниці за рахунок альтернативного сплайсингу і, як наслідок, два різних білки враховуються як один ген.

Однак, якщо ступінь гомології двох генопродуктів, що мають загальний транскрибуєму ділянку, невелика, то ці послідовності розцінюються як два різних гени.

За своїм функціональним призначенням гени можуть бути розділені на дві групи. Група I представлена генами, що кодують власне білки; група II - генами, контролюючими синтез рибосомних, транспортних і ядерних РНК. За характером експресії гени також можуть бути поділені на дві групи - гени "домашнього господарства" (housekeeping genes), продукти яких необхідні для забезпечення життєдіяльності будь-якого типу клітин, і тканинноспецифічні гени, гени "розкоші", яких свідомо більше в 2-3 рази, це гени, які експресуються в клітинах певних тканин і в певний час. забезпечують спеціалізовані функції клітин, тобто гени, функціонально активні тільки в певних типах клітин (тканин) і тільки на певних стадіях онтогенезу, так звані гени термінального диференціювання. Вважається, що середні розміри гена людини становлять, приблизно, від 10 до 30 кб. Однак, ця величина може коливатися від декількох десятків до мільйонів пар нуклеотидів. Згідно з останніми даними, найменший з відомих генів - MCC-7, має розміри всього 21 п.о., а найбільший - ген дистрофіну -2.2 мегабаз. Гени відокремлені один від одного протяжними проміжками - спейсерами, що містять в своєму складі велику кількість повторюваних послідовностей ДНК і унікальні послідовності, які не транскрибуються.

Розглянемо більш докладно сучасні дані про число генів у геномі людини.

1. Виходячи з розміру геному (близько 3 000 мільйонів п.о.), і середнього розміру одного гена порядку 10 - 30 тис.п.о, загальне число генів має бути близько 100 000.

2. При цьому розподіл генів на хромосомах вкрай нерівномірний. Встановлено, що понад 90% генів знаходиться в Гімза-негативних районах метафазних хромосом, в так званих R-бендах. Проведені недавно прямі дослідження методом секвенування показують, що в R-бендах їх число досягає 43-50 на один бенд, а в Гімза-позитивних районах хромосом (G-бендах), - тільки 1-2 на 75-80 кілобаз, тобто всього в геномі можна очікувати близько 70 000 генів.

3. Оцінка числа генів за часткою транскрибуємої частини генома (всього 12%) дає зовсім маленьку величину - 20 000 генів.

4. Методом дослідження кінетики реасоціації РНК в культурі клітин число генів оцінюється між 20 - 40 000. У той же час в клітинах мозку число різних мРНК за деякими даними досягає 97 000.

5. Найбільш точні підрахунки числа генів людини проведені останнім часом і засновані на оцінці числа CG острівців. Відомо, що в промоторной області всіх генів "домашнього господарства" і приблизно у 40% генів термінального диференціювання, що забезпечують спеціалізовані функції диференційованих клітин, знаходяться області коротких динуклеотидних CG повторів. Розроблено молекулярні методи точної реєстрації цих ділянок, які показали, що їх число в геномі близько 45 000, звідси число структурних генів оцінюється в 67-70 000.

6. Практично таке ж число генів (64 000) визначено недавно методом обліку маркерних послідовностей, які експресуються (expressed sequence tags).

Таким чином, підсумовуючи вищевикладене, можна зробити висновок, що в геномі людини міститься, в середньому, близько 70-80 000 окремих транскрибуємих ДНК послідовностей, тобто генів.