

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
МИКОЛАЇВСЬКИЙ НАЦІОНАЛЬНИЙ АГРАРНИЙ УНІВЕРСИТЕТ

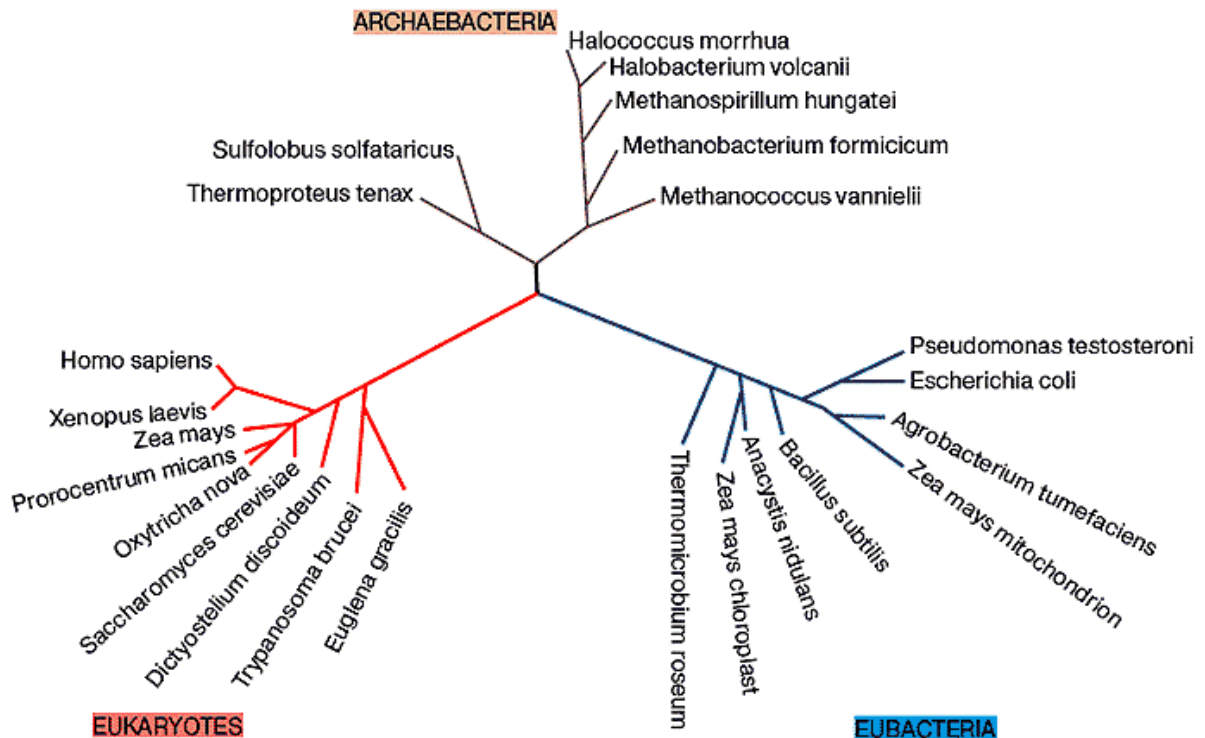
Факультет ТВПШТСБ

Кафедра генетики, годівлі тварин та біотехнології

МОЛЕКУЛЯРНА ФІЛОГЕНЕТИКА ТА БІОІНФОРМАТИКА

Методичні рекомендації

для виконання лабораторно-практичних робіт та вивчення
дисципліни для здобувачів вищої освіти СВО «Магістр»
освітньої спеціальності 162 – «Біотехнології та біоінженерія»
денної форми навчання



Миколаїв - 2020

УДК 575.852 : 577.21.06
М-75

Друкується за рішенням науково-методичної комісії факультету ТВПШТСБ Миколаївського національного аграрного університету від 22 травня 2020 р., протокол № 10.

Укладач:

С.С. Крамаренко – д-р біол. наук, професор, професор кафедри генетики, годівлі тварин та біотехнології Миколаївського національного аграрного університету.

Рецензенти:

О.В. Жуков – д-р біол. наук, доцент, професор кафедри зоології та екології Дніпропетровського національного університету імені Олеся Гончара;

Є.В. Баркаръ – кандидат с.-г. наук, доцент, доцент кафедри генетики, годівлі тварин та біотехнології Миколаївського НАУ.

З М І С Т

ВСТУП	4
Тема 1. Формування теорії еволюції. Передумови та фактори еволюції	5
Тема 2. Природний відбір та видоутворення	12
Тема 3. Штучний відбір та доместикація. Походження та еволюція сільськогосподарських тварин	16
Тема 4. Молекулярні основи еволюції	19
Тема 5. Генетичний поліморфізм та еволюція	23
Тема 6. Еволюція амінокислотних послідовностей	26
Тема 7. Синонімічні і несинонімічні нуклеотидні заміни	31
Тема 8. Філогенетичні дерева. Побудова та тестування філогенії ..	34
Тема 9. Біоінформатика	39
СПИСОК ВИКОРИСТАНОЇ ТА РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ	43

ВСТУП

Дисципліна «Молекулярна філогенетика та біоінформатика» розрахована на підготовку здобувачів вищої освіти (ЗВО) денної форми навчання 2-го РВО, СВО «Магістр», освітня спеціальність 162 – «Біотехнології та біоінженерія», спеціалізація «Сільськогосподарська біотехнологія».

Філогенетика, або філогенетична систематика, займається ідентифікацією і проясненням еволюційних взаємин серед різних видів життя на Землі, як сучасних, так і вимерлих. Еволюційна теорія стверджує, що схожість серед індивідуумів або видів часто вказує на загальне походження або загального предка. Тому взаємини, встановлені філогенетичною систематикою, часто описують еволюційну історію видів і, відтепер, його філогенез, історичні взаємини серед гілками організмів або їх частин, наприклад їх генів. Філогенетична таксономія, яка є відгалуженням, але не логічним продовженням, філогенетичної систематики, займається класифікацією груп організмів згідно зі ступенем їхніх еволюційних відносин.

Біоінформатика застосовує машинні алгоритми і статистичні методи для аналізу великих наборів біологічних даних, які, як правило, складаються з великого числа нуклеотидних (ДНК і РНК) та пептидних (білки) послідовностей і даних структури білків. Головні напрямки досліджень біоінформатики включають вирівнювання послідовностей, пошук генів, збірку геномів, вирівнювання структур білків, передбачення структури білків, передбачення експресії генів та білок-білкової взаємодії та реконструювання процесу еволюції. Великим напрямком досліджень біоінформатики – отримання високоякісних послідовностей геномів з фрагментів послідовностей, отриманих за допомогою традиційних методів секвенування ДНК та конструювання сигнальних мереж за даними ДНК-мікрочипів.

Основна мета навчальної дисципліни є ознайомлення з сучасними комп'ютерними і теоретичними методами аналізу структури генетичних макромолекул, які дозволяють вивчати основні закономірності та особливості їх функціонування та еволюції.

Завдання курсу:

- визначити специфіку комп'ютерного та теоретичного аналізу структури генетичних макромолекул;
- дати огляд стану сучасних методів аналізу структур, банків даних та обчислювальних ресурсів і програм структурної біології, звернувши особливу увагу на їх обмеження і особливості інтерпретації результатів;
- охарактеризувати основні напрямки досліджень у галузі структурної комп'ютерної біології, а також в галузі молекулярної еволюції білків.

Тема 1. Формування теорії еволюції. Передумови та фактори еволюції

Біологічна еволюція — незворотний та, частіше, неспрямований історичний розвиток живої природи, що супроводжується змінами генетичного складу популяцій, формуванням адаптацій (тобто пристосувань до умов існування), створенням та зниканням видів, перетворенням біогеоценозів та біосфери в цілому.

Головні ознаки біологічної еволюції:

1. Біологічна еволюція має незворотний характер (закон Долло).
2. Еволюція — це зміна біологічно корисних ознак.
3. Еволюція базується на гармонійному вбудовуванні нових елементів в стару організацію.
4. На відміну від онтогенезу, еволюція не запрограмована, тобто має випадковий характер.

Головна задача еволюційної теорії - опанування чинників та головних закономірностей історичного розвитку живої матерії. Її розв'язання складається з двох основних напрямків:

1. Широке експериментальне вивчення всіх етапів еволюційного процесу, починаючи з мінливості популяцій та закінчуючи видоутворенням.
2. Розвиток теоретичних досліджень основних проблем еволюційної науки.

Прикладне значення еволюційної теорії:

1. Створює наукову основу селекції тварин, рослин та мікроорганізмів.
2. Має широке використання у вирішенні медичних проблем та у дослідженнях чинників захворювань с.-г. тварин та рослин.
3. Приймає участь у розробці способів наукового управління еволюційним процесом в умовах, коли різко підвищилась роль людини у регуляції функцій біосфери.

Головні теорії виникнення життя на Землі:

1. Креаціонізм (архієпископ Ашер, 1650 р.).
2. Теорія спонтанного зародження життя (стародавній Китай, Вавилон, Єгипет).
3. Теорія панспермії.
4. Біохімічна еволюція (О.І.Опарін, 1923).

Головні етапи формування еволюційних поглядів:

Етап 1. *Зародження еволюційних ідей серед філософів*

Стародавнього Сходу та Античного світу:

1. Ідея єдності Природи (Єгипет, Індія, Китай; II-I тис. до н.е.): існують особливі “початки” чи “першооснови”, з яких формуються всі тіла та явища природи; поява атомістичних поглядів.

2. Ідея “дробини організмів” (Аристотель): усі тіла природи відрізняються один від одного за ступенем складності організації.

мінерали → рослини → зоофіти → нижчі тварини → вищі тварини → людина.

3. Ідея природного виникнення живого (Емпедокл): ніщо не існувало вічно. Не тільки неорганічна матерія, але й все живе (навіть, людина) з’явилося колись за рахунок розвитку єдиної “першооснови”.

4. Ідея чинності розвитку: всі тіла живої ті неживої природи досягають розвитку за рахунок своїх внутрішніх особливостей (наприклад, “ентелехії” Аристотеля) чи на підставі принципу відбору найбільш гармонійних форм (Емпедокла).

Етап 2. Розвиток еволюційних ідей в додарвінівський період:

1. Зародження систематики: введення поняття “вид” (Дж.Рей) та бінарної номенклатури (К.Лінней); на підставі цього - необхідність створення *природної* класифікації, що повинна відображати філогенетичні зв’язки між організмами.

2. Теорії ембріонального розвитку: преформізм (А.Левенгук) та епігенез (У.Гарвей).

3. Розвиток теорії трансформізму (Ж.Бюффон):

— види змінюються;

— пристосування до середовища не дано організму спочатку, а є результатом історичного розвитку видів.

4. Еволюційна теорія Ж.Б.Ламарка (“Філософія зоології”, 1809):

— “закон прямого пристосовування”;

— “закон використання та невикористання органів”;

— “закон успадкування набутих ознак”.

Етап 3. Створення вчення Ч.Дарвіна про штучний та природній відбір

Вчення про штучний відбір

Характерні особливості свійських тварин та рослин:

1. Кожна група свійських тварин (велика рогата худоба, коні, голуби) або культурних рослин (пшениця, капуста та ін.) дуже різноманітні за складом й включають багато форм (порід у тварин або сортів у рослин).

2. Породи всередині однієї групи різко відрізняються як проміж собою, так і по відношенню до свого дикого предка.

3. Все різноманіття свійських форм неможливо пояснити лише гібридизацією або “раптовим породженням”.

Мінливість — процес виникнення нових особливостей, що призводять до розвитку відмінностей серед особин; виникає (згідно уявлень Ч.Дарвіна) під впливом зовнішніх умов та природи самого організму.

Механізми впливу факторів довкілля на тварини або рослини:

— безпосередньо на всю організацію тварини чи рослини;

— непрямо, через “генеративну систему” організма.

Основні форми мінливості (за Ч.Дарвіном):

1. Невизначена (індивідуальна, спадкова) мінливість:
 - має поодинокий характер;
 - кожна ознака може змінюватися у різних напрямках;
 - обов'язково передається нащадкам.
2. Визначена (групова, не спадкова) мінливість:
 - має груповий характер;
 - ознака змінюється в певному напрямку;
 - набуті ознаки не передаються нащадкам.
3. Кореляційна (співвідносна) мінливість.
4. Комбінаційна мінливість.

Еволюція видів у природі

Матеріалом для еволюційного процесу для диких форм, також, є невизначена (спадкова) мінливість.

Найважливішою умовою відбору в природному середовищі (згідно Ч.Дарвіна), є перенаселення, яке виникає як наслідок геометричної прогресії розмноження.

Механізмом еволюції в природному середовищі Ч.Дарвін вважав природній відбір.

Природний відбір — виживання найбільш пристосованих і загибель (елімінація) найменш пристосованих форм.

Обставини, що прискорюють дію природного відбору:

1. Достатня частота виникнення невизначених спадкових змін.
2. Багаточисельність особин виду, що підвищує вірогідність прояву мінливості.
3. Неспоріднене (випадкове) схрещування — панміксія.
4. Ізоляція.
5. Розміри ареалу виду.

Результатом еволюції природних видів є дивергенція, тобто розходження й пристосованість до відмінних умов існування.

Популяція — елементарна одиниця еволюції

Головні вимоги до елементарної еволюційної одиниці:

- повинна виступати в просторі та часі як певна єдність;
- повинна бути здібною спадково змінюватися в ланцюзі біологічних поколінь (в часі);
- повинна реально існувати в природних умовах.

Популяція (за О.В.Яблоковим) — *це мінімальна самовідтворна група особин одного виду, що мешкає на певному просторі протягом еволюційно тривалого часу, утворює самостійну генетичну систему й формує власну екологічну нішу.*

Основні критерії популяції:

— *панміксія*, тобто можливість вільного схрещування двох випадково взятих дорослих організмів даної популяції;

— *територіальна та генетична ізоляція* від інших подібних груп протягом життя достатньо великого числа поколінь.

Основні характеристики популяції:

1. *Популяційний ареал:*

- трофічний ареал;
- репродукційний ареал.

2. *Чисельність популяції та її динаміка:*

- мінімальна чисельність;
- ефективна чисельність.

3. *Віковий склад популяції:*

- співвідношення груп особин різного віку;
- співвідношення різних поколінь;
- співвідношення передрепродуктивного, репродуктивного та пострепродуктивного періодів.

4. *Статевий склад популяції:*

- первинне співвідношення статей;
- вторинне співвідношення статей;
- третинне співвідношення статей.

5. *Генетична структура популяції:*

- генетична гетерогенність;
- генетичний поліморфізм (гетерозиготний та адаптаційний).

6. *Екологічна структура популяції:*

- угруповання за особливостями харчування;
- угруповання за статеві-віковими особливостями;
- угруповання за особливостями міграційної активності;
- угруповання за фенологією.

Особливості штучних популяцій:

1. Високий рівень міжпопуляційної мінливості.
2. Перевага особин жіночої статі.
3. Залежність вікового складу від технологічного напрямку господарства.
4. Відсутність панміксії.

Генетичні основи еволюції

Фенотип — це вся сукупність ознак та властивостей будь-якого індивідуума; є результатом взаємодії між генотипом та навколишнім середовищем.

Генотип — вся генетична інформація організму, генетична структура організму за одним або декількома генними локусами, що досліджуються.

Норма реакції — діапазон реакції розвитку особини на вплив факторів середовища.

Складові частини *фенотипової мінливості*:

— генотипова, або спадкова,

— паратипова, що пов'язана з дією факторів навколишнього середовища.

Модифікації — не спадкові зміни, що виникли під впливом факторів абіотичного чи біотичного середовища в межах реалізації однієї норми реакції.

Характер модифікацій:

1. Завжди суворо закономірні та специфічні.
2. Можуть відрізнятися у одного й того ж організму на різних етапах онтогенезу.
3. Спадково не передаються наступній генерації.

Типи модифікацій:

1. Індивідуальні.

2. Групові.

Головні вимоги до елементарного еволюційного матеріалу:

— це матеріальні одиниці, які повинні являти собою елементарні спадкові зміни й постійно виникати у живих організмів з певною, достатньою частотою;

— повинні торкатися всіх ознак і властивостей живих організмів (у тому числі й “біологічно важливих”) та викликати, при цьому, їх різнобічні відхилення від вихідних форм;

— повинні не тільки виникати, але й розповсюджуватися в межах видового ареалу та зустрічатися в різних популяціях у відмінних частотах.

Мутації — дискретні зміни спадкової інформації особини.

Типи мутацій:

1. Ядерні мутації:

— генні мутації;

— хромосомні мутації.

— геномні мутації.

2. Позаядерні мутації.

Еволюційні характеристики мутацій:

1. *Частота виникнення мутацій* (спонтанних) визначається кількістю гамет однієї генерації, що несуть дану мутацію, у відношенні до загального числа гамет (в середньому складає від 10^{-4} до 10^{-9}).

2. *Спектр мутантних ознак дуже широкий*; немає ознак чи властивостей організму, яких б не торкалися мутації.

3. *Всі природні популяції насичені мутаціями*, що приховуються у гетерозиготному стані (С.С.Четверіков, 1926), але практично немає двох

популяцій, що мали б однакові частоти зустрічальності та спектри мутантних ознак.

Еволюційне значення мутацій:

1. Мутації відносно шкідливі в кожний даний момент еволюції.
2. Дрібні мутації мають змогу вільно накопичуватися в популяції виду.
3. При схрещуванні відносна шкідливість мутації знижується через те, що вона “приховується” у гетерозиготному стані.
4. Еволюційний процес не можливо зводити лише до мутаційного процесу.

Елементарні фактори еволюції

Елементарне еволюційне явище — це *тривала, незворотна й спрямована зміна популяційного генофонду; при цьому здійснюється перехід від однієї генотипової рівноваги до другої.*

Кожне *елементарне еволюційне явище* виникає в популяції внаслідок впливу *елементарних еволюційних факторів* на *елементарний еволюційний матеріал* в межах *елементарної одиниці еволюції*.

Елементарні еволюційні фактори:

- повинні насичати популяції новим елементарним еволюційним матеріалом, тобто мутаціями;
- бути в змозі розподіляти одну “материнську” популяцію на дві або декілька нових;
- повинні обумовлювати появу нових адаптацій та зміну (ускладнення) організації живих організмів.

Елементарні фактори еволюції:

1. *Мутаційний процес* призводить до зміни в популяції частоти одної алелі по відношенню до іншої, тобто постійно підвищує її генетичне різноманіття.

Проява мутації у гомозиготному стані в певних умовах середовища може стати селекційно цінною, через що генетичний склад популяції може змінюватися дуже сильно, оскільки відбір сприятиме розповсюдженню її у всій популяції.

2. *Популяційні хвилі* — зміни чисельності особин однієї популяції, пов’язані з проявом абіотичних та біотичних факторів середовища. Впливають на випадкове коливання частот тих чи інших алелей в популяції, особливо при різкому зменшенні її чисельності.

Класифікація популяційних хвиль:

- періодичні коливання чисельності;
- неперіодичні коливання чисельності;
- спалахи чисельності;
- коливання чисельності, пов’язані з катастрофами.

3. *Ізоляція* — виникнення будь-яких бар'єрів, що обмежують панміксію. Вона закріплює та підсилює початкові етапи генотипової диференціації. Роз'єднані бар'єрами частини популяції та виду раніше чи пізніше потрапляють під різне спрямування відбору. Ізоляція має не спрямований характер.

Класифікація типів ізоляції:

А. *Просторова ізоляція*:

- ізоляція будь-якими бар'єрами;
- ізоляція відстанню;

Б. *Біологічна ізоляція*:

- а) *Екологічна* форма ізоляції обумовлюється умовами існування видів, що пов'язані з місцями та строками розмноження:
 - біотопова ізоляція;
 - сезонна (фенологічна) ізоляція.
- б) *Морфологічна* ізоляція обумовлюється особливостями будови та функціонуванням органів розмноження:
 - морфологічна ізоляція;
 - фізіологічна ізоляція.
- в) *Етологічна* форма ізоляції обумовлюється особливостями поведінки, що дозволяє парування тільки з особинами свого виду.
- г) *Генетична* форма ізоляції проявляється в несумісності статевих продуктів самця та самиці, наприклад, по числу та будові хромосом.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ:

1. Що таке біологічна еволюція? Які головні ознаки біологічної еволюції?
2. Яке значення має еволюційна теорія у селекційній роботі?
3. Які вимоги пред'являють до елементарної еволюційної одиниці?
4. Що таке популяція? Її основні критерії.
5. Що таке генетична гетерогенність? Які основні фактори підтримки генетичного поліморфізму?
6. Які основні характеристики модифікацій? Які є типи модифікацій?
7. Яке значення мають мутації в еволюційному та селекційному процесі?
8. Які є елементарні фактори еволюції?
9. Яке еволюційне значення популяційних хвиль?
10. Яке еволюційне значення має ізоляція?

Тема 2. Природний відбір та видоутворення

Передумови природного відбору:

1. *Гетерогенність особин*: всі особини в популяціях відрізняються між собою за багатьма спадковими особливостями, при цьому імовірність виживання особини повинна бути хоча би частково залежати від її фенотипу.

2. *Прогресія розмноження* обумовлює збиткову чисельність особин кожного виду, оскільки народжується значно більше особин, ніж їх доживає до дорослого стану.

3. *Боротьба за існування* є складний процес протилежних взаємодій особин одного або різних видів, який призводить до природного відбору через елімінацію менш пристосованих організмів. Вона пов'язана з протиріччям між тенденцією організмів до безмежного розмноженню та обмеженістю реальних умов для цього.

Форми боротьби за існування:

А. Конкуренція — провідна форма боротьби за існування через те, що через змагання проявляються головні протиріччя між організмами. Конкуренція обумовлена, з одного боку, однаковими біологічними потребами (їжа, умови розмноження та т.п.), з іншого — однаковими можливостями їх задоволення.

1. *Трофічна конкуренція*:

— внутрішньовидова трофічна конкуренція;

— міжвидова трофічна конкуренція.

2. *Топічна конкуренція*:

— індивідуальна топічна конкуренція;

— групова топічна конкуренція.

3. *Репродуктивна конкуренція*.

Б. Пряма боротьба за існування має своє вираження в безпосередніх взаємодіях організмів один з одним та з абіотичним середовищем.

1. *Пряма боротьба з біотичними факторами* :

— пряма трофічна боротьба;

— пряма репродуктивна боротьба.

2. *Пряма боротьба з абіотичними факторами*.

Природний відбір — вибіркове (диференціальне) розмноження генотипів (або генних комплексів).

Головні особливості природного відбору:

1. *Імовірнісний характер дії відбору* обумовлюється двома сторонами селекційного процесу: статистичністю та стохастичністю.

2. *Накопичувальна дія природного відбору* — поступове додання малих корисних спадкових змін ознак, що призводить до появи нових ознак або удосконалення вже існуючих.

3. *Інтегруюча дія природного відбору* — об'єднання (тобто, інтеграція) малих змін ознак в цілісну систему організму.

Головні форми елімінації:

А. Вибіркова елімінація особин або цілих груп обумовлена індивідуальними або груповими відмінностями організмів щодо рівня пристосованості.

1. *Пряма елімінація*. Особини знищуються безпосередньою дією біотичних або абіотичних факторів середовища:

- пряма елімінація абіотичними факторами;
- пряма елімінація біотичними факторами.

2. *Непряма елімінація* – ослаблення організму (наприклад, через хвороби або довготривале голодування), що тягне за собою його пряме знищення факторами середовища або усунення від розмноження.

Б. Невибіркова елімінація. Вживання особин та їх груп залежить від випадкових дій фізичних або біологічних чинників.

Результати природного відбору:

- створення нових адаптивних ознак;
- знищення ознак, що втратили своє пристосувальне значення;
- збереження пристосовань, що корисні в даних умовах середовища.

Типи та форми природного відбору:

А. Рушійний тип відбору призводить до виникнення еволюційно нових адаптивних ознак.

1. *Спрямований відбір* відображається у виживанні та розмноженні адаптивних відхилень від попередньої норми реакції в умовах середовища, що змінюються довгостроково та в одному напрямку.

2. *Дизруптивний відбір* призводить до виживання та розмноження більш адаптивних граничних відхилень від норми, через елімінацію середніх варіантів.

3. *Транзитивний (перехідний) відбір* відображається у виживанні та більш інтенсивному розмноженні на початку малочисельної форми, що має перевагу над іншою, багаточисельною формою.

Б. Стабілізуючий тип відбору зберігає норму, що має місце в даних умовах середовища, через елімінацію усіх виражених відхилень від неї.

1. *Каналізуючий відбір* означає виживання та розмноження організмів з більш сталим механізмом онтогенезу, які здатні протидіяти випадковим внутрішнім та зовнішнім діям, що порушують вже існуючу норму реакції, тобто нормальне протікання процесів онтогенезу.

2. При *нормалізуючому відборі* зберігається норма реакції внаслідок елімінації відхилень від неї, при цьому спадкова основа її не змінюється.

3. *Балансуєчий (зрівноважуючий) відбір* здійснюється у випадках виживання внутрішньовидових форм, які мають різну адаптивну цінність, але співіснують разом на одній території внаслідок корисності їх сумісного існування для виду в цілому.

Інші форми відбору:

1. *Дестабілізуючий відбір* — порушення корелятивних зв'язків в організмі при інтенсивному відборі в будь-якому визначеному напрямку.

2. *r/K-відбір* пов'язаний із вибиранням однієї з двох головних стратегій розмноження організмів.

3. *Статевий відбір* торкається ознак особин однієї статі; частіше обумовлюється боротьбою між самцями (рідше, самицями) за можливість до розмноження.

4. *Індивідуальний та груповий відбір*. Індивідуальний відбір — диференціальне розмноження деяких особин, які мають перевагу в боротьбі за існування в межах популяції. Він базується на змаганні особин у межах популяції. Груповий відбір — переважне розмноження особин якої-небудь групи.

Творча роль природного відбору має своє вираження в тому, що через елімінацію частини раніше існуючих форм, на підставі закономірностей успадкування та мінливості він призводить до створення зовсім нових форм.

Вид це *складна, інтегрована система внутрішньовидових одиниць — популяцій*. Між популяціями у середині виду здійснюється обмін генетичною інформацією через схрещування особин із різних популяцій.

Головні особливості виду:

1. Вид — якісно новий етап еволюції органічного світу.

2. Вид — сукупність особин, яким придатні загальні морфофізіологічні ознаки і що об'єднані можливістю схрещування один з одним. Вони формують систему популяцій та загальний, частково розірваний або суцільний, ареал.

3. Види в природних умовах звичайно відділені один від одного і представляють генетично стійкі системи.

Критерії виду:

1. *Морфологічний критерій*. Особини різних видів розрізняються за зовнішньою та внутрішньою будовою.

2. *Фізіолого-біохімічний критерій*. Послідовність та інтенсивність фізіологічних процесів, що специфічні для кожного виду. Ці відмінності звичайно менші між близькими видами, ніж між видами філогенетично більш далекими.

3. *Еколого-географічний критерій*. Кожен вид існує в конкретних умовах середовища; для кожного виду характерні свої межі придатних для життя умов, своя історія виникнення, що визначає обриси та розмір ареалу; свої специфічні взаємодії з видами-конкурентами та т.ін.

4. *Генетико-репродуктивний критерій*. Суттєва властивість виду — несхрещуваність його з іншими видами (тобто, генетична ізоляція).

Загальні ознаки виду:

— дискретність;

— чисельність;

- цілісність;
- сталість;
- історичність.

Видоутворення — це розподіл в часі або у просторі раніше єдиної видової форми на дві або декілька. Іншими словами, видоутворення — перетворення генетично відкритих систем (тобто, популяцій всередині виду) в генетично закриті (окремі види).

Головні типи видоутворення:

А. Алопатричне (географічне) видоутворення — це утворення видів із внутрішньовидових форм, територіально ізольованих одна від одної:

- при розподілі ареалу широко розповсюдженого “батьківського” виду;
- при розселенні особин “батьківського” виду, в результаті якого більш віддалені від центру розселення периферійні популяції, перетворюючись в нових умовах, започатковують нові види.

Б. Симпатричне видоутворення. При цьому типі новий вид виникає всередині вихідного ареалу:

- шляхом швидкої зміни каріотипу (наприклад, при поліплоїдії), тобто при подвоєнні, потроєнні і т.ін. основного набору хромосом предкового виду;
- шляхом гібридизації з послідуєчим подвоєнням числа хромосом (алоплоїдія);
- через формування екотипів — популяцій (або груп популяцій), що характеризуються спадково закріпленими адаптаціями до чітко виражених екологічних умов певного місця існування в межах ареалу виду (*екологічне видоутворення*).

В. Філетичне видоутворення. Вид змінюється в середовищі поколінь, перетворюється в новий вид, який чітко відрізняється від «батьківського» за всіма критеріями (частіше, за морфологічним).

Г. «Принцип засновника».

Фактори, що визначають швидкість (темп) видоутворення:

1. Численність популяції та її динаміка.
2. Ізоляція.
3. Просторова структура популяції.
4. Вікова структура популяції.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ:

1. Які головні форми боротьби за існування?
2. Що таке природній відбір та які його головні особливості?
3. Які головні типи природного відбору?
4. Що таке дестабілізуючий відбір та його роль у доместикаційному процесі?

5. Яке значення та творча роль природного відбору?
6. Що таке вид?
7. Які основні критерії виду? Які загальні ознаки виду?
8. Що таке видоутворення? Які фактори зумовлюють темпи видоутворення?

Тема 3. Штучний відбір та domestикація. Походження та еволюція сільськогосподарських тварин

Одомашнюванням тварин (доместикацією) називають процес перетворення диких тварин в свійські. Згідно Е.А.Богданова, він складається з приборкування, приручення та власно одомашнювання.

До *свійських* відносять тварин, які легко розмножуються в умовах господарського утримання, пристосовані до відповідної технології годівлі, утримання та розведення, а також до вимог людини і частіше диференційовані на внутрішньовидові породи.

Сільськогосподарськими називають свійських тварин, розведення яких є галуззю сільського господарства, що має напрямок на отримання від цих тварин того чи іншого виду продукції. До них відносять: велику рогату худобу, коней, верблюдів, віслуків, свиней, овець, кіз, кролів, домашніх птахів, деяких хутрових звірів та інших.

До *приручених* відносять таких диких тварин, які змалку потрапляють під підпорядкування до людини, звикають до неї, підпорядковуються його волі та привчаються робити щось корисне для нього. На відміну від свійських, вони за своїми морфологічними властивостями не відрізняються від диких родичів та частіше не розмножуються у неволі.

Головні центри одомашнювання с.-г. тварин:

1. *Китайський* (Індокитай, Малайський архіпелаг): свині, буйволи, качки, кури, гуси.
2. *Індійський* (Індія): буйволи, гаяли, зебу, павичі та бджоли.
3. *Південно-Західний Азіатський* (Мала Азія, Кавказ, Іран): велика рогата худоба, коні, вівці, свині, верблюди.
4. *Середземноморський* : велика рогата худоба, коні, вівці, кози, кролі, качки.
5. *Андійський* (Північні Анди, Південна Америка): альпака, мускусна качка, індик.
6. *Африканський* (Північно-Східна Африка): страус, осел, свиня, собака, кішка та цесарка.

Біологічні властивості тварин, що сприяють їх domestикації:

1. Філогенетична (еволюційна) "молодість" тварин.
2. Екологічна неспеціалізованість (еврибіонтність) тварин.

3. Особливості фізіології.
4. Особливості поведінки.

Наслідки доместикації:

1. *Підвищений рівень мінливості* є результатом розхитування спадковості під впливом змін умов навколишнього середовища, результатом порушення існуючих особливостей онтогенезу, руйнації внутрішніх кореляцій, що були сформовані у дикому стані та тих взаємовідносин між організмом та середовищем, що склалися під час дикого існування тварини.

2. *Зміна відтворювальної функції*, тобто більший розвиток відтворювальної системи, а також більш раннє статеве дозрівання, порушення сезонності розмноження та підвищення багатоплідності.

3. *Зміна розмірів та форми тіла* (у відповідності напрямку дії штучного відбору) призводить як до підвищення, так й до зниження розмірів організму у порівнянні із дикими предками.

4. *Масть та волосяний покрив*. Більшість диких тварин мають захисне забарвлення волосся (сіра, мишаста), але свійським тваринам притаманна більш висока мінливість по забарвленості волосся (упригол до повної відсутності пігменту — альбінізм).

5. *Шкіра, хвіст та вуха*. Зміна довжини хвоста (в результаті збільшення або зменшення кількості хвостових хребців), або його повна відсутність. Більший розвиток підшкірної клітковини. Розвиток капловухості у деяких порід кіз, собак та кролів.

6. *Голова, череп, роги та скелет в цілому*. Укорочування лицьової частини черепа (ахондроплазія), що пов'язане з порушенням функції ендокринної системи. Зменшення довжини рогів або їх повна відсутність.

7. *Внутрішні органи*. В зв'язку із зміною умов існування (насамперед, травлення) у свійських тварин змінюються органи травної, дихальної, кровоносної систем.

8. *Продуктивність* свійських тварин різко підвищилась як у кількісному, так і в якісному відношенні.

9. *Нервова система та поведінкові реакції*. Менший розвиток органів почуття та головного мозку. Заміна старих умовних рефлексів новими, що сформувалися в техногенному середовищі.

Штучний відбір – це вибір людиною та подальше розмноження організмів, що характеризуються господарсько цінними ознаками.

Порівняльний аналіз штучного та природного відбору:

Х а р а к т е р и с т и к а	Штучний відбір	Природний відбір
Елементарний еволюційний матеріал	індивідуальна спадкова мінливість	
Елементарна одиниця еволюції	особина	популяція
Елементарне еволюційне явище	накопичення змін генетичної структури	

Зміни обумовлені	дією людини	умовами середовища
Умови середовища	ідеальні	реальні
Відбір проводиться на підставі	якості нащадків	пристосування самих тварин
Наявність напрямку відбору	є	немає
Має місце диференційоване	розмноження	смертність
Пристосування до	інтересів людини	умов середовища
Час дії	лише деякі стадії онтогенезу	постійно
Швидкість	швидко	повільно
Рівень генетичної мінливості	поступово знижується	залишається високим
Рівень модифікаційної мінливості	поступово підвищується	залишається постійним
Рівень життєвої здатності	знижується	залишається постійним
Рівень ознак, що підлягають відбору	мають максимальний прояв	мають середній прояв
Результатом є	створення нових порід	створення нових видів

Основні форми штучного відбору:

1. *Несвідомий відбір* – видалення “поганих”, незадовільних відхилень; для нього характерна відсутність мети селекційної роботи.

2. Для *методичного відбору* характерно наявність мети селекції, а при виведенні нової породи (чи сорту) попереднє формування потреб до неї, тобто, “ідеалу породи”.

Принципи штучного відбору:

1. *Масовий відбір* характеризується тим, що його проводять звичайно тільки за фенотипами й підбір плідників відсутній.

2. *Індивідуальний відбір* характеризується націленим підбором пар плідників та їх оцінкою за нащадками.

Фактори, що впливають на ефективність штучного відбору:

1. Ступінь успадкування ознаки.
2. Характер мінливості та її інтенсивність.
3. Чисельність популяції чи породи.
4. Плодючість й частота зміни генерацій.
5. Умови утримання.

6. Характер корелятивних зв'язків.
7. Інтенсивність відбору.
8. Число ознак, що підлягають селекції.
9. Вік тварини.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ:

1. Яке значення мають одомашненні тварини?
2. Які тварини відносяться до свійських?
3. Які характерні особливості свійських тварин?
4. Що таке одомашнювання? Які його головні стадії?
5. Які є центри одомашнювання свійських тварин?
6. Які біологічні особливості сприяють одомашнюванню тварин?
7. Що таке доместикаційні зміни? Які доместикаційні зміни притаманні свійським тваринам?
8. Що таке штучний відбір за сучасними уявленнями?
9. Які основні етапи штучного відбору?
10. Які фактори впливають на ефективність штучного відбору?
11. Які особливості природного та штучного відбору?

Тема 4. Молекулярні основи еволюції

Молекулярна еволюція – це наука, що вивчає зміни генетичних макромолекул (ДНК, РНК, білків) у процесі еволюції, закономірності і механізми цих змін, а також реконструює еволюційну історію генів і організмів.

Об'єктами дослідження молекулярної еволюції є:

1. Послідовності нуклеїнових кислот (ДНК і РНК) як носіїв генетичної інформації.
2. Послідовності білків.
3. Структура білків.
4. Геноми організмів.

Основними задачами молекулярної еволюції є:

1. Виявлення закономірностей еволюції генетичних макромолекул.
2. Реконструкція еволюційної історії генів і організмів.

Молекулярна еволюція тісно зв'язана з такими галузями науки як палеонтологія (датування еволюційних подій), генетика (принципи організації і передачі спадкової інформації), молекулярна біологія (будова генетичних макромолекул), біофізика (механізми функціонування генетичних макромолекул), математика (побудова моделей еволюції), еволюція (загальні

еволюційні закономірності), інформатика (обробка й аналіз даних) і біохімія (хімічна природа і перетворення генетичних макромолекул).

Передумови виникнення і розвитку молекулярної еволюції:

- поява великої кількості експериментальних даних про послідовності нуклеїнових кислот і білків, а також розробка нових теоретичних методів їхнього аналізу. Перші амінокислотні послідовності були розшифровані в 50-і роки минулого сторіччя, нуклеотидні послідовності – у 60-і, а зараз найбільші міжнародні банки даних містять інформацію приблизно про один мільярд секвенованих нуклеотидів (~1,6 млн. генів і їхніх фрагментів) і про амінокислотні послідовності мільйонів білків;

- з часу відкриття будови нуклеїнових кислот стало можливо вивчати еволюційні зв'язки між організмами шляхом порівняння їх нуклеотидних послідовностей.

Такий підхід має деякі переваги над класичними морфологічними і фізіологічними дослідженнями, тому що:

1. ДНК складається з 4 типів нуклеотидів (А, Т, Ц, Г), і це може бути використано для порівняння будь-яких груп організмів, включаючи бактерій, протистів, грибів, рослин і тварин. При використанні класичного підходу таке порівняння неможливе.

2. У процесі еволюції зміни ДНК відбуваються більш-менш регулярно, що дозволяє використовувати математичні моделі для визначення змін і порівняння ДНК філогенетично віддалених організмів. Оцінити еволюційні зміни морфологічних ознак украй складно, особливо за короткий період історичного розвитку органічного світу.

3. Геноми всіх організмів складаються з об'ємної послідовності нуклеотидів і містять значно більше інформації, ніж морфологічні ознаки. З цих причин варто очікувати, що молекулярна філогенетика прояснить ситуацію по встановленню багатьох гілок філіпченкового древа, що неможливо було зробити при використанні класичного підходу. Це можливо в першу чергу при аналізі висококонсервативних послідовностей (наприклад, цитохрома С), які залишалися незмінними протягом тривалих відрізків геологічного часу.

В історії виникнення і розвитку молекулярної еволюції як науки можна виділити два етапи. На першому етапі молекулярна еволюція була підрозділом порівняльної біохімії, а на другому – відокремилася в самостійну науку. Можливість проводити кількісне порівняння амінокислотних послідовностей білків здавалася дуже багатообіцяючою для з'ясування еволюційного споріднення організмів.

Наприкінці 60-х років минулого сторіччя були висунуті дві фундаментальні теорії молекулярної еволюції – теорія М. Кімури і теорія Н. Суєокі. Теорія нейтральної молекулярної еволюції була запропонована в 1968 р. М. Кімурою на підставі даних популяційної генетики і на встановлених швидкостях заміщень амінокислотних залишків у ряді білків і

нуклеотидів у нуклеїнових кислотах. Суть цієї теорії можна викласти у виді п'яти постулатів, перші чотири з яких – емпіричні, а п'ятий – установлений теоретичним шляхом.

1. Швидкість еволюції будь-якого білку, виражена через число амінокислотних замін на сайт у рік, приблизно постійна й однакова у різних філіпченкових лініях, якщо функція і третинна структура цього білка залишаються в основному незмінними. Цей постулат формулює гіпотезу Э. Цукеркендла і Л. Полінга про **молекулярний еволюційний годинник**.

2. Функціонально менш важливі молекули чи їхні частини еволюціонують (накопичуючи мутаційні заміни) швидше, ніж більш важливі.

3. Мутаційні заміни, що приводять до менших порушень структури і функції молекули (консервативні заміни) у ході еволюції відбуваються частіше тих, котрі викликають більш істотне порушення структури і функції цієї молекули.

4. Появі нового у функціональному відношенні білку завжди повинна передувати дуплікація гена.

5. Селективна елімінація шкідливих мутацій і випадкова фіксація селективно-нейтральних чи дуже слабо шкідливих мутацій відбувається в ході еволюції набагато частіше, ніж позитивний дарвінівський відбір сприятливих мутацій.

Розділи молекулярної еволюції:

Молекулярна еволюція містить у собі два основних розділи: *еволюцію генетичних макромолекул* (власне молекулярну еволюцію) і *молекулярну філогенію*.

Еволюція макромолекул вивчає типи і швидкості змін, що відбуваються у генетичному матеріалі (ДНК, РНК), а також утворених на його основі білків, і механізми, відповідальні за ці зміни. Молекулярна філогенія вивчає еволюційну історію макромолекул і організмів, отриману на основі вивчення нуклеотидних і амінокислотних послідовностей.

Методи молекулярної еволюції:

Методи молекулярної еволюції можливо розділити на 2 групи:

I. Методи еволюції макромолекул:

1. Методи пошуку гомологічних послідовностей у базах даних (Blast search, PSI-Blast search і ін.). Ця група методів призначена для створення повної вибірки послідовностей, що досліджується чи білків нуклеїнової кислоти.

2. Методи вирівнювання послідовностей нуклеїнових кислот і білків (Clustal W Protein, Multalin DNA і ін.). Проведення вирівнювання необхідно для наступного вивчення змін порівнюваних нуклеотидних і амінокислотних послідовностей.

3. Методи визначення картини заміщень у вирівняних послідовностях (композиційна дистанція, індекс невідповідності, метод Монте-Карло й ін.). Знання картини заміщень у порівнюваних послідовностях потрібно для диференційованого використання різних методів і формул з метою обчислення еволюційних змін.

4. Методи вивчення еволюційних змін амінокислотних послідовностей (кількість амінокислотних розходжень, p -дистанція та ін.).

5. Методи визначення характеру амінокислотних замінів (дистанція Грентхема, індекс Танга та ін.) дозволяють встановити ступінь консервативності чи радикальності замінів.

6. Методи вивчення еволюційних змін нуклеотидних послідовностей (методи Джукса-Кантора, Кімури та ін.) більш інформативні, тому що враховують не лише ті заміщення, що змінюють амінокислоту (власне несинонімічні), але й ті, що мовчать (синонімічні) заміни.

7. Методи вивчення синонімічних і несинонімічних нуклеотидних замінів (методи Нея-Годжборі, Кумара та ін.) дозволяють проаналізувати ці заміни і використовувати отримані дані для проведення селекційних тестів.

8. Селекційні тести (диференціація дистанцій, ADAPSITE та ін.) служать для встановлення виду відбору в порівнюваних нуклеотидних і амінокислотних послідовностях.

9. Методи вивчення використання кодонів у послідовностях нуклеїнових кислот (RSCU, метод визначення коефіцієнта нейтральності Н. Суеокі та ін.) необхідні для детального вивчення еволюційних змін ДНК і РНК.

10. Допоміжні методи (GZ-gamma, розрахункове співвідношення транзицій і трансверсій та ін.) дозволяють одержати проміжні величини, використані іншими методами.

II. Методи молекулярної філогенії:

1. Дистанційно-матричні методи (метод зв'язування найближчих сусідів, метод мінімальної еволюції та ін.) використовують для створення філіпченкових дерев (дендрограм) різні коефіцієнти їх подібності.

2. Дискретні методи (метод максимальної парсимонії та ін.) будують дендрограми, використовуючи безпосередньо нуклеотидні чи амінокислотні послідовності.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ:

1. Що таке молекулярна еволюція?
2. Що є об'єктами дослідження молекулярної еволюції?
3. Що є основними задачами молекулярної еволюції?
4. Які положення має теорія нейтральної молекулярної еволюції?
5. Які розділи має молекулярна еволюція?
6. Які методи використовує молекулярна еволюція?

Тема 5. Генетичний поліморфізм та еволюція

Генетичний поліморфізм – співіснування в межах популяції двох або декількох різних спадкових форм, що знаходяться в динамічній рівновазі протягом декількох чи навіть багатьох поколінь.

У багатьох випадках внутрішньопопуляційний поліморфізм активно підтримується природним відбором різними способами, зокрема:

а) в деяких випадках гетерозиготи визначаються підвищеною пристосованістю, але внаслідок менделевського розщеплення вони постійно поповнюють популяцію породженими ними менш життєздатними гомозиготами.

б) інтенсивність відбору може змінюватися в межах деякого діапазону, причому на одній із його меж, відбір може сприяти одній формі (морфі), а на іншій межі – іншій. При проміжній інтенсивності відбору можуть виникати поліморфні популяції.

в) інколи відбір буває частотно-залежним: будь-яка внутрішньовидова форма більш життєздатніша тоді, коли вона виникає рідше. Вважають, що саме тому виживають незвично забарвлені форми: вони життєздатні, оскільки хижаки їх не розпізнають і не чіпають;

г) дрібномасштабна просторова структура популяції і її місцезростання бувають дуже складними, і в різних частинках цієї “латаної ковдри” (мозаїки) відбір може відбутися в різних напрямках. Збереження відповідності між організмом і середовищем у такій ситуації неминуче залежить від розсіювання численних нащадків: якщо чисельність їх достатньо велика, то достатньо велика і вірогідність того, що частина їх укорінюється саме в цьому “клаптику”, де відповідна форма найстійкіша. Можливий і інший варіант: організм може бути довговічним і здатним переміщатися, використовуючи при цьому умови і ресурси найпридатніших “клаптиків”.

Типи поліморфних генетичних маркерів:

ДНК-маркери або молекулярно-генетичні маркери, поліморфні ознаки, що виявляються методами молекулярної біології на рівні нуклеотидної послідовності ДНК, для певного гена або для будь-якої іншої ділянки хромосоми при порівнянні різних генотипів, особин, порід, сортів, ліній.

За останні роки накопичився великий масив даних про ефективність використання *молекулярно-генетичних маркерів*, як на рівні білків, так і ДНК, РНК, для вирішення багатьох завдань генетики, селекції, збереження біорізноманіття, вивчення механізмів еволюції, картування хромосом, а також для насінництва та племінної справи.

Найбільш широко використовувані молекулярно-генетичні маркери умовно можна поділити на такі типи:

- **маркери ділянок структурних генів**, що кодують амінокислотні послідовності білків (електрофоретичні варіанти білків), маркери некодуючих

ділянок структурних генів і маркери різних послідовностей ДНК, відношення яких до структурних генів, як правило, невідомо (це: **RAPD** – випадково поліморфна ДНК, що апліфікується; **ISSR** – інвертовані повторення; **AFLP** – поліморфізм у сайтах рестрикції);

- *мікросателітні локуси* (тандемні повторення з довжиною елементарної одиниці в 2-6 нуклеотида) .

Існує цілий набір сучасних технологій виявлення поліморфізму на рівні ДНК, серед яких можна виділити наступні:

- *аналіз поліморфізму довжин рестрикційних фрагментів ДНК (RFLP);*
- *аналіз поліморфізму за допомогою полімеразної ланцюгової реакції (ПЛР) та інші методи на основі ампліфікації ДНК між повторюваними послідовностями в геномній ДНК.*

Маркери на основі ДНК-зондів:

RFLP (ПДРФ-маркери). Поліморфізм довжин фрагментів. ПДРФ – оцінка поліморфізму довжин рестрикційних фрагментів ДНК може бути здійснена різними способами, але найбільш традиційний метод з використанням блот-гібридизації). Цей метод включає в себе виділення ДНК, отримання фрагментів рестрикції, їх електрофоретичної розподіл, перенесення на фільтри з подальшою гібридизацією специфічними ДНК-зондами з отриманими фрагментами ДНК. ДНК-зонд – відносно коротка послідовність клонованої ДНК з певним рівнем гомології і здатністю гібридизуватися з відповідною ділянкою геномної ДНК. Комбінації рестриктаз і зондів дають високо-відтворювані поліморфні спектри фрагментів ДНК, специфічні для кожного індивідуума. Відмінності між останніми можуть бути обумовлені, наприклад, мутаціями, змінами сайтів рестрикції. ПДРФ має ряд важливих переваг, в тому числі й висока відтворюваність спектрів в різних лабораторіях. ПДРФ ефективний при картуванні геному, маркуванні генів багатьох біологічних і економічно важливих ознак. RFLP-алельні варіанти мають кодомінантний характер успадкування .

VNTR (Variable Number Tandem Repeat), метод отримав назву ДНК фінгерпрінт (відбитки пальців). Тандемні повтори широко поширені в різних геномах і високополіморфні. У результаті високої варіабельності цих ділянок ДНК, ПДРФ-аналіз з зондами до мікро- і мінісателітних послідовностей дозволяє отримувати мультілокусні спектри на популяційному рівні. Завдяки дуже високому рівню поліморфізму цей підхід зараз є хорошим інструментом для аналізу внутрішньо- і межпопуляційної мінливості і визначення генетичних відстаней між групами організмів. VNTR-алельні варіанти також мають кодомінантний характер успадкування.

ПЛР-маркери:

Метод *полімеразної ланцюгової реакції* (ПЛР) припускає використання специфічних праймерів та отримання дискретних ДНК-продуктів ампліфікації окремих ділянок геномної ДНК. Велика кількість споріднених технологій побудовано на цьому принципі.

SSR (Simple Sequence Repeats). ПЛР з праймерами, що фланкують до короткого міні- або мікросателітного повтору дозволяє виявляти маркери з кодомінантним типом успадкування і, відповідно, зручний для виявлення гетерозигот по цьому локусу. Однак, одна пара праймерів для флангів в ПЛР дозволяє розглядати поліморфізм тільки одного локусу. Для багатьох мікросателітних локусів не вдається виявити будь-який поліморфізм. Як правило, фланкуючі послідовності для певного мікросателітного локусу виявляються видоспецифічними.

RAPD (Random Amplified Polymorphic DNA). Полімеразна ланцюгова реакція з використанням одиничних коротких, зазвичай, 10-членних праймерів, з довільною нуклеотидної послідовністю. Послідовність праймерів не абсолютно будь-яка, а обмежена в межах 40-70% GC і 50-100% лінгвістичної складності нуклеотидної послідовності. У RAPD можна використовувати як одиночний праймер, так і декілька RAPD-праймерів. Продукт RAPD утворюється в результаті ампліфікації фрагмента геномної ДНК, що фланкована інвертованою послідовністю використаного праймера. Метод універсальний для досліджень різних видів, при використанні однакових праймерів. Як правило, праймер, що виявляє високий поліморфізм для одного виду, буде також ефективний і для інших видів.

ISSR (Inter Simple Sequence Repeats). Спеціалізований варіант RAPD методу, в якому праймер складається з мікросателітної послідовності. У цьому методі, також як і в RAPD, використовується один або кілька праймерів, довжиною в 15-24 нуклеотида. Але в даному випадку, праймери складаються з тандемних коротких 2-4 нуклеотидних повтора, наприклад 5'-CACACACA CACAG-3' і одним або двома селективними нуклеотидами на 3'-кінці праймера. Продукти ISSR ампліфікації містять на флангах інвертовану мікросателітну послідовність праймера. Але в даному методі послідовність праймерів специфічна і підбирається більш суворо, ніж в RAPD, то температуру відпалу в ПЛР можна використовувати вище (55-60°C), ніж для RAPD методу, а тому фінгерпрінт, зазвичай, краще відтворюється.

AFLP (Amplified Fragment Length Polymorphism). Технологія являє собою комбінацію між ПДРФ і ПЛР методами. AFLP – складний метод і складається з декількох етапів: геномна ДНК одночасно розрізається двома рестриктазами (EcoRI і MseI), отримуючи фрагменти з виступаючими 3'-кінцями. Потім рестрикована геномна ДНК лігується з адаптором, що містить «липкі» кінці для рестрикційних сайтів (EcoRI та MseI). Після цього проводиться дві послідовні ПЛР. У першій ПЛР (пре-ампліфікація) використовуються праймери повністю комплементарні адапторам EcoRI та MseI. Після першої ПЛР утворюється велика кількість продуктів ампліфікації між EcoRI і MseI адапторами, які

важко диференціювати за допомогою електрофорезу. Тому, під час другої ПЛР, праймери з EcoRI та MseI адапторами містять на 3'-кінці додаткові і не комплементарні адаптори від 1 до 3 підстави, для селективної ампліфікації. Поділ фрагментів ДНК виконується в поліакриламідному гелі, з радіоактивною або флюоресцентною міткою, відповідно.

Однонуклеотидний поліморфізм (SNP). Гетерогенність первинної структури ДНК, що виявляється в однонуклеотидних (точкових) відмінностях алелей. Наприклад, дві послідовності ДНК різних людей – AAGCСТА і AAGСТТА – відрізняються на один нуклеотид. У такому разі говорять про існування двох алелей: С і Т. SNPs використовуються в якості молекулярно-генетичних маркерів (разом з RFLP та AFLP). Вивчення SNP знайшло широке використання в молекулярній систематиці – побудові біологічної систематики на основі дивергенції гомологічних ділянок ДНК у філогенезі. У даній області найчастіше використовуються спейсери генів рибосомної РНК. З огляду на те, що мутації в даних спейсерах не впливають на структури кінцевих продуктів гену (теоретично вони не впливають на життєздатність), в першому наближенні зазвичай постулюється пряма залежність між ступенем поліморфізму і відстанню філогенезу між організмами.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ:

1. Що таке генетичний поліморфізм?
2. Які є механізми підтримання внутрішньопопуляційного генетичного поліморфізму?
3. Які є типи поліморфних генетичних маркерів?
4. На якому принципі побудовані маркери на основі ДНК-зондів?
5. Які є ПЛР-маркери?

Тема 6. Еволюція амінокислотних послідовностей

Пептидна або амінокислотна послідовність – порядок, в якому амінокислотні залишки, сполучені пептидними зв'язками, знаходяться в поліпептидному ланцюгу білка або пептиду. Пептидна послідовність часто називається *білковою послідовністю*, якщо вона представляє первинну структуру білка. Загалом пептидна послідовність представляє собою первинну структуру всіх білків на момент синтезу, проте через посттрансляційну модифікацію первинна структура дещо змінюється з утворенням зрілого білка.

Характер амінокислотних замін визначається ступенем їх консервативності або радикальності. Консервативною заміною амінокислоти називається мутаційна заміна, яка не призводить до суттєвих змін структури та функції білка. В процесі еволюції консервативні заміни амінокислот відбуваються частіше, ніж радикальні. Ці заміни переважно зустрічаються в функціонально важливих ділянках білкової молекули (наприклад, сайтах

Основним параметром, що характеризує еволюцію амінокислотних послідовностей, є еволюційна дистанція між ними. Еволюційні дистанції використовуються для побудови філогенетичних дерев (дендрограм) та визначення часу дивергенції, що можливо завдяки існуванню однозначної відповідності між часом дивергенції та ступенем амінокислотних відмінностей при умові постійності швидкостей еволюції білків в різних філогенетичних лініях.

Методи обчислення еволюційних дистанцій між амінокислотними послідовностями:

1. Методи приблизної оцінки еволюційних дистанцій (кількість амінокислотних відмінностей, *p*-дистанція).
2. Методи коректованої оцінки еволюційних дистанцій (Пуассон-коректована дистанція, дистанція з поправкою Кімури, гама-дистанція, дистанція Грішина і т.ін.).
3. Методи, які базуються на використанні матриць замінів амінокислот (Дейхофф та Джонса-Тейлора-Торнтон).

Методи приблизної оцінки еволюційних дистанцій

Кількість амінокислотних відмінностей. Вивчення еволюційних змін білків та поліпептидів починається з порівняння та вирівнювання двох або більше амінокислотних послідовностей різних організмів.

Як вже зазначалося, вирівнюванням амінокислотних послідовностей називається процес співставлення порівнюваних послідовностей для такого їх взаєморозташування, при якому спостерігається максимальна кількість збігів амінокислотних залишків. До програм, призначених для вирівнювання амінокислотних послідовностей, відносяться **Clustal W Protein** та **Multalin Protein**, а також програмні пакети, які їх містять (наприклад, **MEGA**).

Найбільш простим методом визначення еволюційних відмінностей після проведення вирівнювання є підрахунок кількості амінокислотних відмінностей n_a між двома послідовностями. Однак, для використання цього параметра необхідно, щоб послідовності, які порівнюються, складались з однакової кількості амінокислотних залишків n , що спостерігається досить рідко. На практиці амінокислотні послідовності часто включають пробіли, які не повинні враховуватися при обчисленні n_a .

Таким чином аналіз еволюції амінокислотних послідовностей по кількості відмінностей є можливим лише при однаковій довжині послідовностей. В інших випадках цей параметр є лише проміжною величиною для подальших розрахунків.

P-дистанція. Більш точним методом визначення кількості відмінностей між послідовностями є обчислення долі різних амінокислот в цих послідовностях (позначається p , P_d , *p-distance* від англ. part of differences). Використовуючи цю характеристику, можна оцінити кількість відмінностей

між послідовностями навіть, якщо їх довжини значно відрізняються. p -дистанція визначається за формулою:

$$p_d = n_d / n.$$

Якщо заміни у всіх амінокислотних сайтах відбуваються з рівною ймовірністю однієї заміни на сайт p_1 , то ймовірність k замін на сайт розподілена за біноміальним розподілом:

$$P(k) = C_n^k p_1^k (1 - p_1)^{n-k},$$

де

$$C_n^k = \frac{n!}{k!(n-k)!},$$

k - кількість замін.

Дисперсія величини $P(k)$ визначається за формулою:

$$D(p) = np(1 - p).$$

Слід відмітити, що p -дистанція не є суворо пропорційною часу дивергенції таксономічних груп організмів t і тому описаний метод використовується для отримання приблизної оцінки еволюційних відмінностей амінокислотних послідовностей (таксономічна група - група організмів, що об'єднані певним ступенем спорідненості, спільними рисами будови та функціональних особливостей).

Методи корегованої оцінки еволюційних дистанцій

Пуассон-коректована дистанція. Однією з причин нелінійної залежності p -дистанції від t є поступове зростання відмінності між n_d та істинною кількістю замін амінокислот при множинних замінах на певних ділянках сайтів. Для більш точного визначення кількості замін слід використовувати корекцію Пуассона.

Позначимо швидкість амінокислотних замін в рік в певному сайті r і будемо для спрощення вважати, що r у всіх сайтах однакова. Далі ми покажемо, що похибка, яка з'являється внаслідок такого припущення, є малою при малих значеннях P_d . Добуток rt дорівнює середній кількості амінокислотних замін на сайт за період часу t . Нехай ймовірність $P(k)$ здійснення k амінокислотних замін в сайті описується розподілом Пуассона з інтенсивністю rt . Тоді

$$P(k) = \frac{(rt)^k}{k!} e^{-rt}$$

Отже, ймовірність відсутності амінокислотних замін в сайті дорівнює:

$$P(0) = e^{-rt}$$

а очікувана кількість консервативних (незмінних) амінокислотних сайтів дорівнює ne^{-rt} . На практиці амінокислотна предкова послідовність, як правило, невідома, що робить неможливим безпосереднє використання формули, яка описує розподіл Пуассона. Кількість амінокислотних замін оцінюється при порівнянні двох гомологічних послідовностей, які дивергували t років назад. Оскільки ймовірність відсутності заміни в амінокислотному сайті послідовності

дорівнює e^{-rt} , то ймовірність q того, що ні в одному з гомологічних сайтів двох послідовностей не відбудеться заміни, дорівнює

$$q = (e^{-rt})^2 = e^{-2rt}$$

Ймовірність того, що заміни відбудуться можна записати як

$$p = 1 - q.$$

Отримана формула для q є наближеною, оскільки не враховує зворотні та паралельні мутації (однакові мутації, які відбуваються в гомологічних амінокислотних сайтах в двох різних еволюційних лініях). Однак, внесок цих ефектів є малим в широкому діапазоні значень $p \geq 0,3$.

Використовуючи отримані результати знаходимо, що загальна кількість амінокислотних замін на сайт $d = 2rt$ для двох послідовностей дорівнює:

$$d = -\ln(1 - p).$$

Величина d має назву Пуассон-коректованої дистанції (Poisson correction distance, PC-distance).

Якщо, використовуючи додаткові джерела вдається визначити час дивергенції двох послідовностей одна від одної, то швидкість замін амінокислот дорівнює:

$$r = d/2t.$$

Знаходження в знаменнику величини $2t$ (а не t) викликано тим, що t – це час, що пройшов після еволюційної дивергенції двох ланцюгів від спільної для них предкового ланцюга і множник 2 в знаменнику відповідає двом гілкам філогенетичного дерева.

Дистанція Кімури. Еволюційна дистанція Кімури d_K є ще одним варіантом корекції р-дистанції. Її величина обчислюється по емпіричній формулі

$$d_K = -\ln(1 - p - 1/5p^2).$$

Кімура з'ясував, що ця формула справедлива при значеннях p , які не перевищують 0,7. Позначимо вектор-рядок відносних частот появи амінокислот поліпептиду в момент часу t_0 як p_0 . Тоді частоти замін амінокислот за час t (або для t РАМів) визначаються формулою

$$p_0 = pM^t.$$

Елемент M^t_{ab} матриці дорівнює ймовірності того, що амінокислота в рядку i_a заміниться на амінокислоту в стовпчику b за час t РАМ, а M^t_{aa} – ймовірності незмінності амінокислоти a за проміжок часу t РАМ. Ці ймовірності можна використати для співставлення долі різних амінокислот у гомологічних послідовностях p та числом амінокислотних замін на сайт (d_D - дистанція Дейхоффа), припустивши, що частоти амінокислот залишаються незмінними в процесі еволюції. Тоді p обчислюється як

$$p = 1 - \sum_a p(a)M^{2t}_{aa}$$

де $p(a)$ – це частота появи амінокислоти в послідовності. Використання M^{2t}_{aa} замість M^t_{aa} пов'язано з тим, що ми розглядаємо дві послідовності, які дивергували t РАМ часових одиниць тому назад.

Не зважаючи на те, що для різних білків частоти появи амінокислот g відрізняються, Дейхофф запропонувала використовувати середні частоти появи амінокислот для багатьох різних білків. Такий підхід не враховує специфічність кожного білка, але при цьому робить можливим застосування цього методу для багатьох білків. До того ж велика кількість білків має досить близькі частоти появи амінокислот, що робить цей метод досить точним.

Метод Джонса-Тейлора-Торнтона. У 1992 р. Джонс, Тейлор і Торнтон запропонували для побудови матриці замін новий метод (Jones-Taylor-Thornton matrix, JTT-matrix), який враховує більшу кількість замін для більшого числа білків. Також були зроблені спроби створення окремих матриць для мітохондріальних білків хребетних.

Оскільки різні білки мають різні матриці замін, було б доречно створити матриці замін для кожної групи білків, але для цього необхідно більше даних по їх амінокислотним послідовностям. Тому зараз для вимірювання еволюційної дистанції між послідовностями визначають параметр a та обчислюють відповідну гама-дистанцію.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ:

1. Який характер мають амінокислотні заміни?
2. Які параметри використовуються для оцінки характеру амінокислотних замін?
3. Які існують методи обчислення еволюційних дистанцій між амінокислотними послідовностями?
4. Які є методи приблизної оцінки еволюційних дистанцій?
5. Які є методи корегованої оцінки еволюційних дистанцій?

Тема 7. Синонімічні і несинонімічні нуклеотидні заміни

Еволюційна відстань (K) являє собою середнє число нуклеотидних замін, що приходяться на один нуклеотидний сайт двох порівнюваних гомологічних послідовностей генетичних макромолекул двох видів організмів.

Визначення еволюційної відстані між нуклеотидними послідовностями генетичних макромолекул полягає в знаходженні взаємозв'язку між розходженнями на рівні ДНК чи РНК і часом (геологічним) еволюції видів. Знання такої залежності дозволяє будувати *філіпченкові дерева* і датування подій молекулярної еволюції, наприклад, визначення часу дивергенції таксонів на основі порівняння послідовностей генетичних макромолекул.

У найпростішому випадку, коли заміни рідкісні (чи час еволюції замалий), можна припустити, що кількість замін у парі послідовностей прямо пропорційна часу їхньої еволюції. Відмітимо, що відстань між двома послідовностями відбиває подвоєний час, що пройшов з моменту дивергенції (за умови, що швидкості нагромадження замін у двох послідовностях були однакові).

Визначення еволюційних відстаней між нуклеотидними послідовностями можна здійснити, використовуючи наступні методи:

1. Метод Джукса-Кантора.
2. Метод Таджими-Нея.
3. Двопараметричний метод Кімури.
4. Метод Тамури.
5. Метод Тамури-Нея.

У дослідженнях еволюційних розходжень нуклеотидних послідовностей молекул РНК чи ДНК часто потрібно визначити не тільки загальне число і швидкість нуклеотидних замін, але й окремо встановити частку і швидкості *синонімічних* і *несинонімічних* замін.

Синонімічною заміною (чи заміною, що мовчить) прийнято вважати заміну нуклеотида, що не приводить до зміни амінокислоти.

Несинонімічною заміною називають заміну, що приводить до зміни амінокислоти, що кодується.

Таблиця генетичного коду вказує, що всі заміни в другому положенні нуклеотида в кодоні є **несинонімічними**, у той час як частина замін нуклеотидів у першому і третьому положеннях – **синонімічні**. Згідно гіпотезам про рівні частоти нуклеотидів і випадкових замін, ця частина приблизно складає 5% для першого положення і 72% для третього положення.

Через те, що на синонімічні заміни не має вплив природний відбір на рівні білків, швидкість синонімічних замін відповідає швидкості нейтральних нуклеотидних замін.

Встановлено, що швидкість синонімічних замін приблизно однакова для багатьох генів, якщо не диференціювати її відповідно до використання кодонів і іншими факторами. Швидкість несинонімічних замін, навпаки, звичайно набагато нижче такої для синонімічних і значно варіює для різних генів.

Але існують гени, у процесі еволюції яких несинонімічні заміни зустрічаються частіше, ніж синонімічні. Ці несинонімічні заміни, мабуть, викликані позитивним дарвінівським відбором, тому що відповідно до представлень про нейтральну селекцію швидкості синонімічних і несинонімічних замін повинні бути рівні. З цієї причини визначення швидкостей синонімічних і несинонімічних замін є важливим предметом досліджень у молекулярної еволюції.

Визначення швидкостей синонімічних і несинонімічних замін набагато складніше визначення загальної швидкості нуклеотидних замін. В багатьох нуклеотидних послідовностях міститься більше нуклеотидних сайтів, у яких можуть відбутися несинонімічні мутації, чим сайтів, у яких можливі синонімічні мутації, при чому в різних генах кількість синонімічних і несинонімічних сайтів відрізняється.

Швидкість синонімічних і несинонімічних замін необхідно визначати як число синонімічних замін на синонімічний сайт (r_S) і як число несинонімічних замін на несинонімічний сайт (r_N) у рік чи іншу одиницю часу (t), відповідно.

Однак тривалість дивергенції між двома послідовностями ДНК, що порівнюються, звичайно не відома і можливо лише розглядати число синонімічних замін на синонімічний сайт ($d_S = 2r_S t$) і число несинонімічних замін на несинонімічний сайт ($d_N = 2r_N t$) для цієї пари послідовностей.

Всі існуючі методи визначення d_S і d_N можна розділити на три групи:

1. Методи, засновані на еволюційних шляхах;
2. Методи, засновані на двохпараметричній моделі Кімури;
3. Методи максимальної подібності з моделями заміщень кодонів.

Але ці методи засновані на різних гіпотезах і саме тому вони не завжди дають однакові результати.

Існують два типи нуклеотидних замін, що призводять до зміни різних типів основ в будові ДНК: *транзиції* та *трансверсії*.

Транзиція – мутація, обумовлена заміною однієї пуринової основи на іншу ($A \leftrightarrow G$), або однієї піримідинової на іншу ($U, T \leftrightarrow C$). Транзиції – це прості заміни (не відбувається зміни орієнтації пурин-піримідин в мутантному сайті молекули ДНК).

Частота транзицій (P) обчислюється за формулою:

$$P = \frac{n_P}{L},$$

де n_P – число спостережених транзицій; L – загальне число нуклеотидних сайтів, за якими порівнюються послідовності.

Трансверсія – мутація, обумовлена заміною пуринової основи на піримідинову й навпаки ($A, G \leftrightarrow U, T, C$). Трансверсії – це складні або перехресні заміни (відбувається зміна орієнтації пурин-піримідин в мутантному сайті молекули ДНК).

Частота трансверсії (Q) визначається за формулою:

$$Q = \frac{n_Q}{L},$$

де n_Q – число спостережених трансверсій; L – загальне число нуклеотидних сайтів, за якими порівнюються послідовності.

Співвідношення спостережуваних трансверсій і транзицій (q) розраховується за формулою:

$$q = \frac{n_Q}{n_P}.$$

Наприклад, для двох вирівняних послідовностей ДНК:

ATGAAGTCCGCCTTA

ATGAAAACGGCACTA

спостерігається дві транзиції (G → A й T → C) та три трансверсії (T → A, C → G та C → A):

```

ATGAAAGTCCGCCTTA
    ↓ ↓ ↓ ↓ ↓
ATGAAAACGGCACTA
  
```

Селекційні тести (тести відбору) – це група методів, призначених для визначення виду відбору в амінокислотних послідовностях білків і/чи нуклеотидних послідовностях РНК і ДНК, а також у їхніх частинах.

До селекційних тестів, використовуваних при аналізі еволюції нуклеотидних послідовностей, відносять диференціацію дистанцій, Z-тест, ASL-тест, тест Фішера та тест Голдмана-Янга.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ:

1. Що таке еволюційна відстань?
2. Яка заміна нуклеотида вважається синонімічною, а яка несинонімічною?
3. Як визначається швидкість синонімічних і несинонімічних замін?
4. Що таке транзиція?
5. Що таке трансверсія?
6. Що таке селекційні тести?

Тема 8. Філогенетичні дерева. Побудова та тестування філогенії

Філогенетика або *філогенетична систематика* – галузь біологічної систематики, що займається ідентифікацією і проясненням еволюційних взаємин серед різних видів життя на Землі, як сучасних, так і вимерлих. Еволюційна теорія стверджує, що схожість серед індивідуумів або видів часто вказує на загальне походження або загального предка. Тому взаємини, встановлені філогенетичною систематикою, часто описують еволюційну історію видів і, відтепер, його філогенез, історичні взаємини серед гілками організмів або їх частин, наприклад їх генів. Філогенетична таксономія, яка є відгалуженням, але не логічним продовженням, філогенетичної систематики, займається класифікацією груп організмів згідно зі ступенем їхніх еволюційних відносин.

Філогенетичне древо (еволюційне дерево, дерево життя) – умовна схема, що відображає еволюційні взаємозв'язки між різними видами, іншими таксонами, генами або іншими об'єктами, що мають загального предка.

Вершини філогенетичного дерева діляться на три класи: *листя*, *вузли* і (максимум один) *корінь*. *Листя* – це кінцеві вершини, тобто ті, в які входять рівно по одному ребру; кожен лист відображає деякий вид живих організмів (або інший об'єкт, схильний до еволюції, наприклад білковий домен). Кожен вузол представляє еволюційну подію: розділення предкового виду на два або більше, які надалі еволюціонували незалежно. *Корінь* представляє загального предка всіх даних об'єктів. *Ребра* дерева філогенезу прийнято називати «гілками».

Типи філогенетичних дерев:

Вкорінене дерево – дерево, що містить виділену вершину – корінь. Вкорінене дерево можна вважати орієнтованим графом, оскільки воно має природну орієнтацію – від кореня до листя. Кожен вузол вкоріненого дерева відповідає останньому загальному предкові листя дерева, що лежить нижче.

Невкорінене дерево не містить кореня і відображає зв'язок листя без передбачуваного положення загального предка. Необхідність розглядати невикорінені дерева виникає через те, що часто зв'язки між вузлами відновити легше, ніж напрямок еволюції. На другому малюнку показано невикорінене філогенетичне дерево. Найдостовірнішим методом для перетворення невикоріненого дерева на вкорінене (для цього треба або оголосити коренем один з вузлів, або розбити одну з гілок дві, що виходять з кореня) є використання достовірної «зовнішньої групи» видів – достатньо близьких до набору видів, що цікавить нас, але в той же час явний є окремою групою. В інших випадках положення кореня можна встановити, виходячи з додаткових припущень про швидкість еволюції досліджуваних об'єктів.

Як вкорінені, так і невикорінені філогенетичні дерева можуть бути *біфуркаційними* або *мультифуркаційними*, а також *маркованими* або *немаркованими*. У біфуркаційному дереві до кожного вузла підходять рівно три гілки (у разі вкоріненого дерева – одна вхідна гілка і дві вихідні). Таким чином біфуркаційне дерево припускає, що всі еволюційні події сприяли в походженні від предкового об'єкта рівно двох нащадків. До вузла мультифуркаційного дерева можуть підходити чотири й більше гілок. Марковане дерево містить назви листків, тоді як немарковане просто відображає топологію.

Форми філогенетичних дерев:

Дендрограма – загальний термін, що позначає схематичне представлення філогенетичного дерева.

Кладограма – філогенетичне дерево, що не містить інформації про довжини гілок.

Філограма (або *фенограма*) — філогенетичне дерево, що містить інформацію про довжини гілок — ці довжини представляють зміну якоїсь характеристики.

Хронограма — філограма, довжини гілок в якій представляють еволюційний час.

Методи філогенетичного аналізу:

Фенетичні методи. Фенетика, також відома як числова таксономія, використовує різні заходи для визначення певної схожості видів, що розглядаються. Немає ніяких обмежень на число або тип особливостей (даних) які можуть використовуватися, хоча всі дані повинні бути вперше перетворені на числові значення, без будь-якого «нормування». Кожен організм потім порівнюється з кожним іншим по всіх особливостях, в результаті підраховується число схожостей (або відмінності). Організми потім групуються таким чином, що найподібніші були розташовані разом, а найвідмінніші – на максимальній відстані. В результаті отримують таксономічні кластери або фенोगрами, які не обов'язково відображають генетичну схожість або еволюційний зв'язок. Відсутність еволюційного значення в фенетиці привела до незначного впливу на класифікацію тварин, і внаслідку інтерес до використання фенетики значно зменшився.

Фенетичні методи включають в себе численні форми кластеризації та класифікації. За допомогою цих методів стає можливим витончене редукування варіативності організмів до вигляду, що підлягає логічній класифікації. На практиці це означає дослідження десятків варіацій деяких ознак, і їхню візуалізацію у вигляді дво- та тривимірних графіків. Найважливішим концептуальним завданням фенетики є знаходження коректного балансу між втратами інформації в процесі описаного спрощення та узагальнення, та легкості і очевидності інтерпретації вихідних графіків.

Фенетика не бере до уваги різницю між *плезіоморфами* – ознаками, що успадковані від організму, що є попередником (і тому філогенетично неінформативні) – та *апоморфами* – ознаками, які еволюціонували в одній чи кількох із груп організмів, що походять від загальної попередникової форми. Таким чином, фенетика потенційно зазнає небезпеку помилки при наявності конвергентної еволюції та адаптивної радіації в досліджуваних групах. Типовою помилкою при фенетичному аналізі є визначення як монофілетичних таких базових попередникових форм, що мають багато плезіоморф порівняно з еволюційно досконалішими нащадками.

Фенетичні методи можуть бути інформативнішими за кладистичні у тому випадку, коли тільки різниця між таксонами є важливою для аналізу, і при цьому ресурси для цифрових розрахунків обмежені. У випадку ж, коли важливою є еволюційна історія досліджуваних таксонів, у сучасній систематиці використовуються лише кладистичні методи.

Кладистичні методи. Альтернативний підхід до схематичного зображення взаємин між таксонами називається кладистикою. Основне припущення кладистики полягає в тому, що члени групи розділяють загальну еволюційну історію. Тому вони більш близько відносяться одна до одної, ніж вони до інших груп організмів. Зв'язані групи визначаються по наявності набору унікальних особливостей (апоморфій), які були відсутні у віддалених

предків, але які розділяються більшістю або всіма організмами в межах групи. Отримані характеристики, що розділяються членами групи, називаються синапоморфіями. Тому, на відміну від фенетичних, кладистичні групи не залежать від того, чи організми розділяють фізичні риси, але залежать від їх еволюційних взаємин. Дійсно, в кладистичних аналізах два організми можуть розділяти численні характеристики, але бути членами різних груп.

Кладистичний аналіз використовує ряд припущень. Наприклад, вважається що види з'являються тільки роздвоєнням, або відділенням, із спадкової групи. У випадку гібридизації (схрещування) або горизонтального переносу генетичної інформації види вважаються зниклими, а такі явища — рідкими або відсутніми. Крім того, кладистичні групи повинні мати такі характеристики: всі види в групі повинні розділяти загального предка і всі види, отримані від загального предка, повинні увійти до таксону.

Дотримання цих вимог приводить до наступних термінів, що використовуються для посилання на різні можливі способи складу груп:

Монофілетична група (або клада), в якій всі види розділяють загального предка і включаються всі види що походять від цього загального предка. Тільки така форма сприймається кладистами як «правильна».

Парафілетична група, в якій всі види розділяють загального предка, але не всі види, що походять від цього загального предка, включаються.

Поліфілетична група, в якій види, які не розділяють безпосереднього загального предка, складаються одну групу, виключаючи види, які б зв'язали їх.

Результатом кладистичного аналізу походження таксону є спеціальні діаграми, що називаються *кладограмами*, які віддзеркалюють гіпотетичні філогенетичні зв'язки. Кладистичний аналіз може ґрунтуватись на матеріалі різного обсягу, доступному для конкретного дослідника. Наприклад, стосовно викопних організмів, як правило, доступні тільки морфологічні дані, але для сучасних тварин використовується також аналіз послідовностей ДНК (так звані «молекулярні дані») та біохімічні дані.

У кладограмі всі таксони розташовані на кінцях гілок, і кожне розгалуження є дихотомічним (двійковим). Два таксони, розділені одним розгалуженням, називаються сестринськими таксонами або сестринськими групами. Кожна частина загального дерева, незалежно від того, налічує вона один або кілька тисяч таксонів, називається «клада». Клада включає всі таксони, які мають спільного унікального предка (такого, що не є попередником ніякого більше організму в діаграмі). Кожна клада характеризується набором ознак, притаманних її членам, але відсутніх у предків. Ці ознаки кладу називаються «синапоморфії» (спільні риси). Наприклад, затверділа передня пара крил (надкрила) є синапоморфіями жуків, а кільцеве листоутворення (або розвертання вайів) є синапоморфією папоротей.

Важливо також зауважити, що **ноди** («гілки») кладограми не обов'язково віддзеркалюють розходження еволюційних гілок, а натомість лише розходження станів ознак, що спостерігаються між цими гілками. Ознаки, що

полягають у різниці послідовностей ДНК, здатні розходитись після того, як генний дрейф між популяціями редукується до деякої порогової величини, в той час як помітні морфологічні зміни, зазвичай будучі епістатичними (тобто результатом взаємодії кількох генів) виявляються лише після того, як таксони, що розійшлись, відокремлено еволюційно розвивались протягом деякого (зазвичай досить значного) часу; так, біологічні підвиди найчастіше можуть бути розрізнені генетично, але не морфологічно (за будовою тіла чи внутрішньою анатомією).

З огляду на те, що секвенування ДНК стає все дешевшим та легшим у впровадженні, молекулярна систематика набуває все більшої популярності як спосіб реконструкції філогенезу. Використання критерію парсимонії є лише одним з методів знаходження філогенії таксону на базі молекулярно-біологічних даних; методи максимальної правдоподібності та Байєзівського виведення, котрі включають в себе докладно розроблені моделі еволюції послідовностей, є відомими не-геннігіанськими методами оцінки даних секвенування. Іншим потужним методом реконструкції філогенії є використання ретротранспозонних маркерів геному, котрі вважають такими, що менш піддаються впливу реверсії секвенованих даних, котра (реверсія) може зробити дані непридатними для побудови філогенетичної послідовності. Також перевагою ретротранспозонних маркерів є те, що вони мало піддаються гомоплазії через інтегрування в геном у повністю випадкових місцях (хоча така статистична випадковість, щонайменше подекуди, є сумнівною).

В ідеалі морфологічні, молекулярно-біологічні, та інші (поведінкові, екологічні, палеонтологічні, і т.ін.) філогенетичні дані мають бути узагальненими при опрацюванні підсумкового висновку; при цьому ніякий з методів не є більш доказовим за інші, але всі вони мають різні внутрішні джерела помилок. Наприклад, конвергенція розвитку ознак (гомоплазія) набагато частіше з'являється при аналізі морфологічних даних, ніж серед даних, отриманих при молекулярному секвенуванні, але реверсії ознак в них обох зустрічаються з приблизно однаковою частотою; зазвичай, морфологічні гомоплазії можуть бути знайдені при достатньо уважному та детальному аналізі характеристичних ознак.

Кладистика не визнає жодну з наявних теорій еволюції, беручи до уваги лише загальну інформацію про походження з (часом) істотними її модифікаціями. Завдяки цьому кладистичні методи можуть бути (і бувають) використаними в небіологічних дисциплінах, включаючи історичну лінгвістику та завдання з визначення авторства текстів.

Молекулярна філогенетика

Макромолекулярні дані, під якими мають на увазі послідовності генетичного матеріалу (ДНК) і білків, накопичуються все швидшими темпами завдяки успіхам молекулярної біології. Для еволюційної біології швидке накопичення даних послідовностей цілих геномів має значну цінність, тому що

сама природа ДНК дозволяє використовувати його як «документ» еволюційної історії. Порівняння послідовностей ДНК різних генів між різними організмами можуть сказати вченому багато нового про взаємини організмів, що не можуть інакше бути виявлено засновуючись на морфології, або зовнішньої формі організмів і їх внутрішньої структурі. Оскільки геноми еволюціонують поступовим накопиченням мутацій, кількість відмінностей послідовності нуклеотидів між парою геномів різних організмів повинна вказати, як недавно ті два геноми розділили загального предка. Два геноми, що відхилилися в недавньому минулому, повинні мати менші відмінності, ніж два геноми, чий загальний предок дуже давній. Тому, порівнюючи різні геноми один з одним, можливе отримання еволюційних взаємин між ними – є головним завданням молекулярної філогенетики.

Молекулярна філогенетика намагається визначити швидкість і відмінності змін в ДНК і білках щоб відновити еволюційну історію генів і організмів. Два загальні підходи можуть використовуватися щоб отримати цю інформацію. У першому підході, вчені використовують ДНК, щоб вивчати еволюцію організму. У другому підході, різні організми використовуються, щоб вивчати еволюцію ДНК. У будь-якому підході загальна мета — зробити висновок щодо процесу еволюційних змін організму по змінам ДНК і процесу молекулярної еволюції по картині змін ДНК.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ:

1. Що таке філогенетика?
2. Які є типи філогенетичних дерев?
3. Які є форми філогенетичних дерев?
4. Які особливості фенетичних методів аналізу?
5. Які особливості кладистичних методів аналізу?
6. Які завдання має молекулярна філогенетика?

Тема 9. Біоінформатика

Біоінформатика - наука, що займається вивченням організації та функціонування біологічних систем різного рівня (від молекулярного до популяційного) на основі методів і засобів інформатики.

Біоінформатику відносять до числа високих технологій сучасної біології, що забезпечує інформаційно-комп'ютерні та теоретичні основи генетики і селекції, молекулярної біології, генетичної та білкової інженерії, біотехнології, медичної генетики, генної діагностики та екології.

З точки зору біоінформатики молекули нуклеїнових кислот та білків представляють собою текст, що складається з обмеженої кількості алфавітних знаків (4 для нуклеїнових кислот та 20 для білків). Задачею біоінформатики є аналіз змісту цього тексту.

Можна виділити три **основних напрямки** біоінформатики:

- створення баз даних, які дозволяють зберігати велику кількість класифікованих біологічних даних і керувати ними;
- розробка алгоритмів і методів статистичного аналізу для визначення співвідношень між тими чи іншими даними;
- використання програмного забезпечення, алгоритмів і баз даних для аналізу біологічних даних різних типів, зокрема, послідовностей ДНК, РНК і білків, білкових структур, профілів експресії генів і біохімічних шляхів, та отримання на основі цих даних нових знань.

Біоінформатика у вузькому сенсі слова, а саме - застосування комп'ютерних методів для вирішення завдань молекулярної біології, в основному аналізу різних послідовностей (амінокислотних, нуклеотидних). Ця наука виникла в 1976-1978 роках, остаточно сформувалася в 1980 році із спеціальним випуском журналу «Nucleic Acid Research» (NAR).

Основні області досліджень у біоінформатиці:

- 1) Дослідження еволюції живої природи за допомогою методів інформатики та математики.
- 2) Комп'ютерне та математичне моделювання інформаційних процесів в біологічних системах.
- 3) Комп'ютерна генетика: розшифровка та моделювання структурної організації генів та геномів, а також кодуємих генами білків; аналіз мутацій та ін.
- 4) Комп'ютерна нейробіологія: моделювання природних нейронних систем, розробка нейромереж та ін.
- 5) Дослідження екологічних систем за допомогою інформаційних технологій.
- 6) Комп'ютерне моделювання біологічної дії ксенобіотиків.
- 7) Комп'ютерне моделювання процесів отримання, накопичення, обробки та систематизації біологічних та медичних даних.
- 8) Комп'ютерне розпізнавання та синтез зображень біологічних об'єктів.
- 9) Створення нових інформаційних технологій на основі результатів досліджень живої природи.
- 10) Організація та використання автоматизованих банків даних з біології та медицини, в тому числі банків міждисциплінарних даних.
- 11) Розробка інтелектуальних систем аналізу та прогнозування властивостей біологічних об'єктів на основі спеціалізованих баз та банків даних.
- 12) Створення систем інформаційного забезпечення та підтримки біологічних та медичних досліджень.

Біоінформатика включає в себе:

- бази даних, в яких зберігається біологічна інформація;
- набір інструментів для аналізу тих даних, які містяться в таких базах;
- правильне застосування комп'ютерних методів для правильного вирішення біологічних задач.

Біолог в біоінформатиці зазвичай має справу з базами даних та інструментами їх аналізу.

Тепер розберемося, які бази даних бувають залежно від того, що в них поміщають.

Перший тип - архівні бази даних, куди будь-хто може помістити все, що захоче. До таких баз відносяться:

- GeneBank & EMBL - тут зберігаються первинні послідовності;
- PDB - просторові структури білків;

Другий тип - бази даних, за достовірність яких відповідає господар бази даних. Туди інформацію ніхто не надсилає, її з архівних баз даних відбирають експерти, перевіряючи достовірність інформації - що записано в цих послідовностях, які є експериментальні підстави для того, щоб вважати, що ці послідовності виконують ту або іншу функцію.

До баз даних такого типу належать:

- Swiss-Prot - найбільш якісна база даних, яка містить амінокислотні послідовності білків;
- KEGG - інформація про метаболізм;
- FlyBase - інформація про *Drosophila*;
- COG - інформація про ортологічних генах.

Підтримка бази вимагає роботи кураторів або анотаторів.

Третій тип - похідні бази даних. Такі бази виходять у результаті обробки даних з архівних і підвідомчих баз даних. Сюди входять:

- SCOP - База даних структурної класифікації білків (описується структура білків);
- PFAM - База даних за домами білків;
- GO (Gene Ontology) - Класифікація генів (спроба створення набору термінів, упорядкування термінології, щоб один ген не називався по різному, і щоб різним генам не давали однакові назви);
- ProDom - білкові домени;
- AsMamDB - альтернативний сплайсинг у ссавців.

Інтегровані бази даних, в яких вся інформація звалена в купу, і ввівши ім'я гена, можна знайти всю пов'язану з ним, інформацію - в яких організмах зустрічається, в якому місці геному локалізована, які функції виконує і т.д.

- NCBI Entrez - доступ до інформації про нуклеотидні та амінокислотні послідовності і структури
- Ecosuc - все про *E. coli* - гени, білки, метаболізм і пр.

ПИТАННЯ ДЛЯ САМОПЕРЕВІРКИ:

1. Що таке біоінформатика?
2. Які основні напрямки біоінформатики?
3. Які основні області досліджень у біоінформатиці?
4. Які бази даних використовують у біоінформатиці?

СПИСОК ВИКОРИСТАНОЇ ТА РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. Лукашов В. В. Молекулярная эволюция и филогенетический анализ. Москва : БИНОМ, 2009. 256 с.
2. Ней М., Кумар С. Молекулярная эволюция и филогенетика. Киев : КВЦ, 2004. 418 с.
3. Картавцев Ю. Ф. Молекулярная эволюция и популяционная генетика. Владивосток : Изд-во Дальневост. ун-та, 2005. 234 с.
4. Айала Ф., Кайгер Дж. Современная генетика. Т. 3. Москва : Мир, 1988. 335 с.
5. Кимура М. Молекулярная эволюция: Теория нейтральности. Москва : Мир, 1985. 394 с.
6. Ратнер В. А. Молекулярная эволюция. *Соросовский образовательный журнал*. 1996. № 3. С. 41-47.
7. Левонтин Р. Генетические основы эволюции. Москва : Мир, 1978. 351 с.
8. Павлинов И. Я. Кладистический анализ. Москва : Изд-во Моск. ун-та, 1990. 158 с.
9. Хаубольд Б., Вие Т. Введение в вычислительную биологию : эволюционный подход. Москва – Ижевск : НИЦ «Регуляторная и хаотическая динамика», Ижевский институт компьютерных исследований, 2011. 456 с.

Навчальне видання

МОЛЕКУЛЯРНА ФІЛОГЕНЕТИКА ТА БІОІНФОРМАТИКА
Методичні рекомендації

Укладач: **Крамаренко Сергій Сергійович**

Формат 60×84.1/16. Ум. друк. арк. 0,9
Тираж ___ прим. Зам № _____

Надруковано у видавничому відділі
Миколаївського національного аграрного університету.
54020, м. Миколаїв, вул. Георгія Гонгадзе, 9
Свідоцтво суб'єкта видавничої справи ДК № 4490 від 20.02.2013 р.