

Лекція 3

Тема. Моделювання у дослідженнях.

План:

1. Основні поняття і терміни
2. Складні системи та системний аналіз
3. Системна біологія у біохімічних дослідженнях.
4. Моделювання, як напрям досліджень
5. Експеримент *in silico*

Література:

- 1) Ілюха Л.М. Системна біологія та системна регуляція фізіологічних процесів. *Вісник Черкаського університету*. 2022. №2. С. 32-37.
- 2) Оболенська М. Ю. Системна біологія і проект «Encode». *Ukr. Biochem. J.*, 2014, Vol. 86, N 4. P. 5-17. URL: http://ua.ukrbiochemjournal.org/wp-content/uploads/sites/3/2015/06/Obolenska_4_14.pdf

1. Основні поняття і терміни

Біотехнологія - дисципліна, що вивчає можливості використання живих організмів, їх систем або продуктів їх життєдіяльності для вирішення технологічних завдань, а також можливості створення живих організмів з необхідними властивостями методом генної інженерії.

Емерджентність – наявність у будь-якої системи особливих властивостей, не властивих її підсистемам і блокам, а також сумі елементів, не пов'язаних системоутвірними зв'язками; неможливість зведення властивостей системи до суми властивостей її компонентів.

Інформаційні технології – це використання комп'ютерів для створення, обробки, зберігання, вилучення та обміну всіма видами даних та інформації.

Когнітивні (пізнавальні) інформаційні технології – один із найбільш «інтелектуальних» розділів теорії штучного інтелекту. Вони описують основні розумові процеси людини.

Молекулярний докінг (або молекулярне стикування) — це метод молекулярного моделювання, який дозволяє передбачити найбільш вигідну для утворення стійкого комплексу орієнтацію та положення однієї молекули по відношенню до іншої.

Нанотехнології - це напрям науки, що спеціалізується на розробці та застосуванні об'єктів, розмір яких становить від одиниць до кількох сотень нанометрів.

Розміри біологічних макромолекул — нуклеїнових кислот (ДНК, РНК) та білків (антигени, антитіла, вірусні капсиди, ферменти та ін.) знаходяться в нанодіапазоні. Нанооб'єкти небіогенної природи (наприклад, наночастинки металів або квантові напівпровідникові точки) можуть бути носіями біомакромолекул, призначених для цільового впливу на певні біологічні мішені. З іншого боку, біологічні макромолекули можуть бути засобом доставки небіогенних наночастинок в орган-мішень для діагностичного або терапевтичного впливу.

Редукціонізм – методологічний принцип, згідно з яким складні явища можуть бути повністю пояснені за допомогою законів, властивих явищам простішим.

Система – це сукупність взаємозв'язаних елементів, уособлена від зовнішнього середовища, але яка діє з цим середовищем як єдине ціле.

Складна система – впорядкована множина структурно взаємозв'язаних та функціонально взаємодіючих різнотипних систем, які об'єднані структурно в цілісний об'єкт функціонально різнородними взаємозв'язками для досягнення заданих цілей в певних умовах.

Теорія динамічних систем - це область математики, яка використовується для опису поведінки складних динамічних систем, зазвичай за допомогою диференціальних рівнянь або різницевих рівнянь.

Холізм - ідея про те, що різні системи слід як єдине ціле, а чи не просто як набір частин.

2. Складні системи та системний аналіз

Бурхливий розвиток сучасних когнітивних, біо-, нано- та інформаційних технологій породив нові феномени в природі та суспільстві, що значною мірою визначають шляхи еволюції техносфери та біосфери.

Технологічне перетворення соціуму породило інноваційні зв'язки у біосфері та суспільстві. Ці зв'язки та визначальні їх інноваційні явища зажадали висунути нові наукові парадигми за допомогою **теорії систем та системотехніки**, які забезпечують адекватний опис цих нових феноменів. Нові фундаментальні явища, що зустрічаються у фізиці, техносфері, біології, економіці та екології, важко піддаються опису в рамках традиційної парадигми точних наук, отримали назву **складних систем** (це, наприклад, глобальний клімат Землі, організми, людський мозок, інфраструктура, така як електромережа, транспортні або комунікаційні системи, складне програмне забезпечення та електронні системи, космічні системи зв'язку, мегаполіси, екосистема, жива клітина і, зрештою, весь всесвіт).

Принципово новий підхід, що здатний забезпечити опис подібних явищ, представляє **теорія системного аналізу**. Вона народилася у тісному сприянні з теорією нерівноважних термодинамічних систем, теорією випадкових дій, теорією катастроф (динамічного хаосу).

Успіхи системного аналізу дозволили подолати низку криз, пов'язаних з розвитком сучасних технологій, і показали, що дослідження складних систем неможливо без використання системних методів.

Сутність системного підходу у тому, що досліджувані об'єкти (складні системи) представляються як ієрархічний набір взаємодій між компонентами системи, які описуються (з допомогою сучасних методів математики) як набір відносин між характеристиками компонентів системи і як кібернетичні моделі системи. Щоб підкреслити, що сама система та її компоненти (підсистеми) описуються набором характеристик (фізичних величин, змінних, параметрів, чинників) їх часто під час аналізу називають об'єктами.

Мається на увазі, що вони виступають як об'єкти експерименту і є «джерелами» експериментальних даних, які перетворюються (завдяки процедурам обробки даних) у вимірні значення параметрів. Часто деякі набори вдається об'єднати та запровадити узагальнені параметри чи чинники, що успішно характеризують об'єкт (складну систему) загалом.

3. Системна біологія у біохімічних дослідженнях.

Системна біологія – наукова дисципліна, що утворилася внаслідок перетину біології та теорії складних систем. Уперше термін був використаний у 1993 році.

Системна біологія є міждисциплінарною наукою про життя. Спрямована на вивчення складних взаємодій в живих системах. Використовує новий підхід в біології: холізм замість редукціонізму. Основна увага в системній біології приділяється емерджентним властивостям, тобто властивостями біологічних систем, які неможливо пояснити тільки з точки зору властивостей її компонентів. Таким чином завданнями системної біології є дослідження та моделювання властивостей складних біологічних систем, які не можна пояснити сумою властивостей її складових.

Широкого вжитку термін «системна біологія» набув після 2000 року. Системна біологія має зв'язок із математичною біологією.

До завдань системної біології можна віднести:

- інтеграцію та зберігання експериментальних даних та результатів їх аналіз;
- розробку методів, підходів та технологій для аналізу біологічних даних;
- аналіз біологічних даних великого обсягу – повних геномів, транскриптомів, протеомів тощо;
- математичне моделювання динаміки біологічних систем

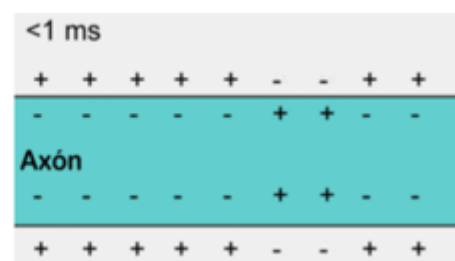
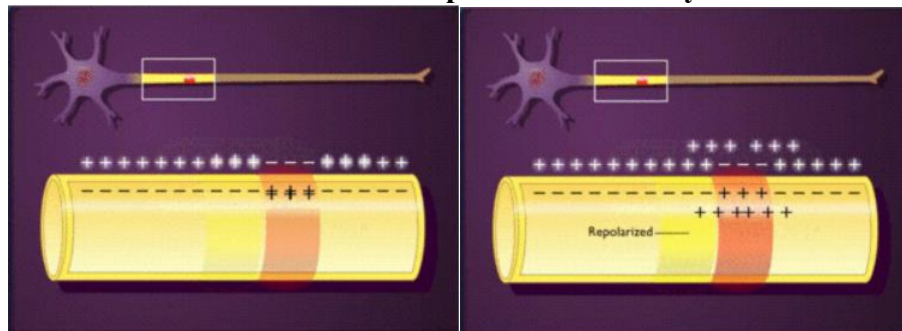
Історія виникнення системної біології

Передумовами появи системної біології:

- кількісне моделювання ферментативної кінетики, що почало формуватись як напрям у 1900 році;
- математичне моделювання зростання популяцій;
- моделювання в нейрофізіології;
- теорія динамічних систем і кібернетика.

Однією з перших чисельних моделей у біології є модель британських нейрофізіологів та лауреатів нобелівської премії Ходжкіна та Хакслі, опублікованої у 1952 році. Автори створили математичну модель, що пояснює поширення потенціалу дії вздовж аксона нейрона. Їхня модель описувала механізм поширення потенціалу як взаємодію між двома різними молекулярними компонентами: каналами для калію та натрію, що можна вважати початком обчислювальної системної біології.

Анімація поширення потенціалу дії:

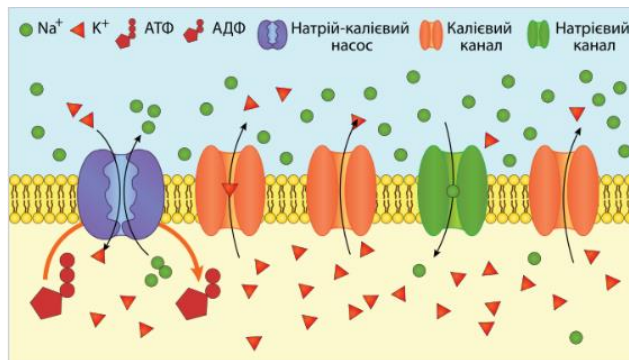


Молекулярні механізми виникнення потенціалу дії:

Активні властивості мембрани, що забезпечують виникнення потенціалу дії, ґрунтуються головним чином на поведінці потенціал-залежних натрієвих (Na^+) і калієвих (K^+) каналів. Початкова фаза ПД формується вхідним натрієвим струмом, пізніше відкриваються калієві канали і вихідний K^+ -струм повертає потенціал мембрани до вихідного рівня. Початкову концентрацію іонів потім відновлює натрій-калієвий насос.

По ходу ПД канали переходять зі стану в стан: у Na^+ -каналів основних стани три — закритий, відкритий і інактивований (в реальності все складніше), у K^+ каналів два — закритий і відкритий.

Поведінка каналів, що беруть участь у формуванні ПД, описується через провідність і розраховується через коефіцієнти переносу (трансферу).



Коефіцієнти переносу були виведені Аланом Ллойдом Ходжіном та Ендрю Хакслі.
Приклад провідності для калію G_K на одиницю площі [S/cm^2]:

$$G_K = G_{Kmax} n^4$$

$$dn/dt = \alpha_n (1 - n) - \beta_n n,$$

де:

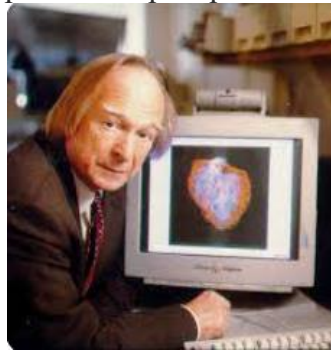
α_n - коефіцієнт трансферу з закритого у відкритий стан для K^+ каналів [1/s];

β_n - коефіцієнт трансферу з відкритого в закритий стан для K^+ каналів [1/s];

n - фракція K^+ каналів у відкритому стані;

$(1 - n)$ - фракція K^+ каналів в закритому стані

У 1960 році на основі моделі Ходжіна та Хакслі Деніс Нобл створив першу комп'ютерну модель серцевого ритму. Його робота дозволила теоретично інтерпретувати серцеві аритмії та розробити антиаритмічні препарати.



У другій половині двадцятого століття був розроблений ряд підходів для вивчення складних молекулярних систем, таких як теорія контролю метаболізму і теорія біохімічних систем. З успіхами молекулярної біології у 80-х роках відбулося падіння інтересу до моделювання біологічних систем.

На початку двадцятого століття, коли створювалися інститути системної біології в Сіетлі і Токіо, системна біологія залучилася в різні геномні проекти, допомагаючи в інтерпретації експериментів, включаючи біоінформатику.

Експериментальні методи системної біології

Для формулювання гіпотез і висновків системна біологія використовує дані, отримані в інших галузях біології: біохімії, біофізики, молекулярної біології.

Проте, існує низка специфічних методів, міцно асоційованих із системною біологією. Для цих методів характерна велика кількість експериментальних вимірювань, а також

одночасного детектування (виділення вихідного інформаційного сигналу з несучої хвилі) багатьох характеристик. Прикладами таких методів можуть бути:

- **Геноміка** – високопродуктивні методи секвенування ДНК, включаючи вивчення варіабельності у різних клітинах одного організму;
- **Епігеноміка**, Епігенетика — вивчення факторів, що впливають на експресію генів, які не визначаються послідовністю нуклеотидів у ДНК;
- **Транскриптоміка** — вимірювання експресії генів, використовуючи ДНК-мікрочіпи та інші методи;
- **Інтерфероміка** — вимір взаємодії транскрибованих РНК;
- **Протеоміка**, Транслатоміка: вимірювання рівня білків або пептидів з використанням двовимірного гель-електрофорезу, мас-спектрометрії або багатовимірних методик вимірювання білків;
- **Метаболоміка** — вимірювання концентрацій малих молекул, метаболітів;
- **Глікоміка** — вимірювання рівня вуглеводів;
- **Ліпідоміка** — вимірювання рівня ліпідів.

Крім представлених методів вимірювання рівня молекул, існують також складніші методи, що дозволяють визначити динаміку характеристик у часі та взаємодію між компонентами:

- **Інтерактоміка** — вимір взаємодій між молекулами;
- **Флаксоміка** — вимір динаміки потоків та концентрацій у годині;
- **Біоміка** – системний аналіз біому.

Перераховані методи все ще активно розвиваються як в напрямку збільшення точності й інформативності вимірювань, так і в способах чисельної обробки одержуваних даних.

Інструменти системної біології

Дослідження в галузі системної біології найчастіше полягають у розробці механічної моделі складної біологічної системи, тобто моделі, сконструйованої на основі кількісних даних про елементарні процеси, що складають систему.

Метаболічний чи сигнальний шлях може бути описаний математично на основі теорій ферментативної чи хімічної кінетики.

Сучасна системна біологія неможлива без використання комп'ютерних технологій. Комп'ютери використовуються для розв'язання систем нелінійних рівнянь, вивчення стійкості та чутливості системи, визначення невідомих параметрів рівнянь за експериментальними даними.

Нові комп'ютерні технології справляють значний вплив на розвиток системної біології. Зокрема, використання обчислювальних процесів, автоматичних засобів пошуку інформації у публікаціях, обчислювальної лінгвістики, розробка та наповнення загальнодоступних баз даних.

В рамках системної біології проводитиметься робота над створенням власних програмних засобів для моделювання та універсальних мов для зберігання та анотації моделей. Як приклад можна навести SBML, CellML – розширення XML для запису моделей, а також SBGN – мову графічного подання структури взаємодій елементів біологічних систем.

4. Моделювання, як напрям досліджень

Моделювання (англ. simulation) — подання різноманітних характеристик поведінки фізичної чи абстрактної системи за допомогою іншої системи.

Під моделлю розуміють уявну або матеріальну систему, яка, відображаючи або відтворюючи об'єкт дослідження, може замінити його так, що її вивчення дає нову інформацію про цей об'єкт.

Метод моделювання має таку структуру: а) постановка завдання; б) визначення аналога; в) створення або вибір моделі; г) розробка конструкту; д) дослідження моделі; е) переведення знань з моделі на оригінал.

За допомогою моделювання вивчаються ті процеси і явища, що не піддаються безпосередньому вивченню.

Метод моделювання зарекомендував себе як ефективний засіб виявлення суттєвих ознак явищ та процесів за допомогою моделі (концептуальної, вербальної, математичної, графічної, фізичної тощо).

Математичне моделювання в біології

Одним з основних напрямків системної біології є математичне моделювання динаміки біологічних процесів. Навіть якщо дослідник знає механізм протікання того чи іншого процесу, наприклад, біосинтезу амінокислоти, регуляції експресії сигнального пептиду, диференціювання стовбурової клітини та. т. д., дослідження цього процесу в динаміці дозволяє відповісти на безліч додаткових питань, таких як:

- Чи змінюється концентрація результуючих речовин при зміні умов перебігу реакції? Наприклад, як змінюється концентрація цукру у крові при стресі (рис. А);

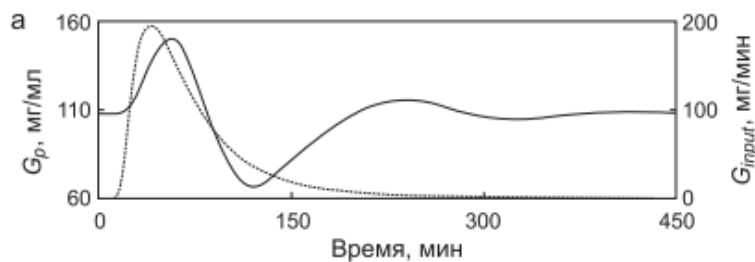


Рис. А. Моделювання динаміки зміни концентрації глюкози в крові (чорна крива) після вживання її в їжу (пунктирна лінія) (адаптовано з: (Kang et al., 2012. P. 84-93));

- Чи змінюється динаміка досліджуваного процесу при зміні зовнішніх умов та патології? Наприклад, чи змінюються періодичні коливання концентрацій гормонів росту при зміні часового поясу (рис. Б);

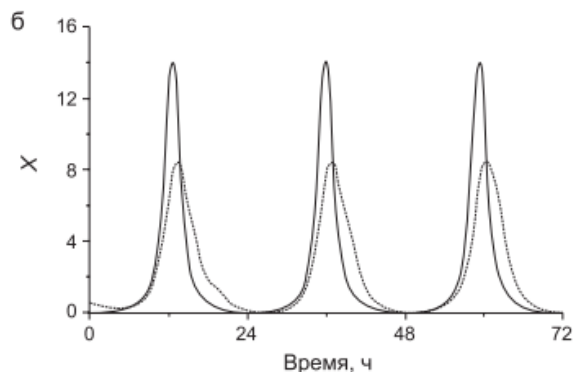


Рис.Б Моделювання динаміки активності нейронів у добових ритмах. Чорна лінія – зимовий час, пунктирна лінія – літній час (Bodenstein et al., 2012). P. 1633-1638);

- Який тип регуляції краще описує зміни концентрації речовин, що спостерігаються у різних експериментах? Наприклад, якщо відомо, що регуляція філлотаксису у рослини здійснюється фітогормоном ауксином, який регулює експресію своїх транспортерів PIN, який тип регуляції найбільш добре описує формування розподілу ауксину в меристемі пагона (рис. В).

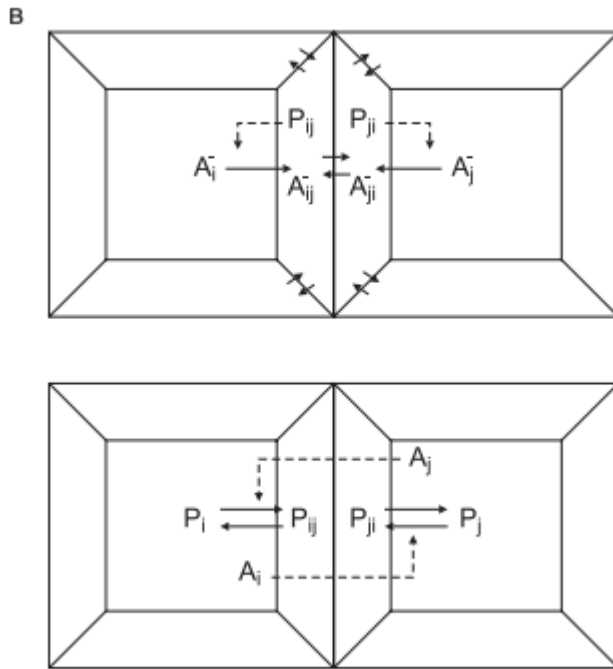


Рис. В. Два різних механізми регулювання транспорту ауксину A через його транспортер P . В результаті моделювання було доведено, що механізм на нижній схемі добре описує розподіл ауксину у меристемі пагону (Jönsson et al., 2006. P. 1633-1638).

Основною метою математичного моделювання є перевірка гіпотез про особливості регуляції біологічних процесів, а також генерація нових гіпотез. Крім того, математичні моделі дозволяють вирішувати практичні завдання у біології та медицині. Наприклад, актуальною є завдання оцінки концентрації поживних речовин, оптимальних для сталого розвитку бактерій у ферментерах – апаратах для промислового вирощування бактерій.

5. Експеримент *in silico*

За аналогією з експериментами *in vivo* (у живому організмі) та *in vitro* (у пробірці) біологічні експерименти, здійснені на комп'ютері, стали називати *in silico* (комп'ютерний експеримент).

Раніше комп'ютер розглядався лише як засіб обробки експериментальних даних. Наприклад, статистичні дослідження чисельності популяцій або обробка даних рентгеноструктурного аналізу, з часом почала розвиватися ідея повністю комп'ютерного експерименту. Сьогодні комп'ютер розглядається як середовище проведення експерименту. Маючи на вході певні вихідні параметри можна отримати деякі результати моделювання, що описують поведінку досліджуваної системи.

Підхід *in silico* багаторівневий і включає завдання моделювання поведінки окремих молекул, біохімічних процесів і навіть функціонування окремих фізіологічних систем.

Молекулярне моделювання – це один із аспектів поняття *in silico*, який є набором обчислювальних методів, що дозволяють з різним ступенем фізичності на атомарному рівні вивчати молекулярні системи різної складності.

Методи молекулярного моделювання можна розбити на певні групи, які поєднують різні підходи і дозволяють вирішувати завдання різного характеру.

1. Методики аналізу біологічних текстів (амінокислотних та нуклеотидних послідовностей), а також роботу з відповідними базами даних – біоінформатика.

2. Методики вивчення еволюції молекулярних систем з використанням емпіричних силових полів.
3. Моделювання на підставі гомології – визначення просторової структури білкових молекул, враховуючи наявні структурні шаблони білків з амінокислотними гомологічними послідовностями.
4. Молекулярний докінг – спеціальні підходи, що дозволяють вивчати взаємодії типу ліганд-рецептор або фермент-субстрат.

Основні етапи експерименту *in silico* з математичного моделювання у біології:

1. Збір та аналіз експериментальних даних. Будь-яка математична модель у біології повинна ґрунтуватися на доведених в експериментах фактах. У разі недостатності експериментальних даних та з метою спрощення розрахунку моделі в експерименті *in silico* можуть бути прийняті деякі припущення та спрощення. Їхня правомочність має бути добре аргументована

2. Формалізація досліджуваного процесу у вигляді математичних формул та рівнянь. При цьому враховуються основні фізичні закони, наприклад, закон збереження маси. Ключовими поняттями в математичній моделі є змінні та параметри. **Змінні** – біологічні об'єкти, динаміка яких вивчається в експерименті, наприклад транскрипційні фактори, біологічно активні речовини та метаболіти. Також це можуть бути концентрації речовин. Для їх розрахунку у моделі необхідно задати початкові дані – значення змінних у початковий момент розрахунку моделі. **Параметри** – константи швидкостей реакції, коефіцієнти та інші величини. Наприклад, часто параметрами виступають стаціонарні значення концентрації речовин, динаміка яких у даній математичній моделі не цікавить дослідників. Параметри можуть бути розмірними та безрозмірними.

3. Вибір початкових даних та параметрів моделі. Це найважливіший етап, який багато в чому залежить від біолога, що надає дані, необхідні оцінки параметрів.

4. Чисельне дослідження створеної моделі з вибраними набором параметрів та початковими даними. Залежно від розв'язуваної задачі в експерименті *in silico* аналізують стаціонарне рішення моделі або зміну значень змінних у часі. Наприклад, у моделі метаболізму жирів зменшення концентрації АТФ у початковий час може імітувати голодування чи стрес.

5. Перевірка запропонованих в результаті аналізу моделі гіпотез. Це можливо експериментальна перевірка. А також для перевірки гіпотез іноді проводять наступний раунд експериментів *in silico*, який починається із зміни вихідної моделі.

На останніх двох етапах експерименту *in silico* результати чисельних експериментів необхідно порівняти з наявними експериментальними даними. Необхідно критично підійти до таких питань. Чи адекватна постановка та формалізація моделі для вирішення цього завдання? Чи описує модель набір експериментальних даних, які можна досліджувати за допомогою даної моделі *in silico*? Чи дає модель нові передбачення порівняно з наявними експериментальними даними? Якщо відповідь на всі три запитання стверджувальна, то експеримент *in silico* вважатиметься успішним.