

ЕПОХА ВЕЛИКИХ ГЕНОМНИХ ВІДКРИТТІВ: КОРОТКИЙ НАРИС ІСТОРІЇ НАУКИ ПРО ГЕНОМНЕ РІЗНОМАНІТТЯ

Тарас ОЛЕКСИК¹, Владислав МІРУТЕНКО², Ярослава ГАСИНЕЦЬ¹, Роман КІШ¹, Андрій ВОШЧЕПИНЕЦЬ³

У динамічному світі сучасної біології, надзвичайно стрімкий розвиток нової науки геноміки, що поєднала у собі методи біології та комп'ютерні технології, відкриває безмежні можливості для розуміння таємниць живих організмів, від найпростіших бактерій до людини. Незважаючи на безліч відкриттів у царині структури та функції геномів, ми все ще знаходимося на початкових стадіях зародження і розвитку цього напрямку. Глобальні ініціативи секвенування починаючи від Проєкту Геном Людин (HGP), НарМар та Проєкту 1000 геномів та інших подібних проєктів розширили наше розуміння генетичної різноманітності людей у контексті популяційної історії та еволюції нашого виду. Геномні методи революціонізували медицину, дозволяючи розробляти персоналізовані лікувальні стратегії з урахуванням генетичних особливостей кожної людини. Сучасні технології, такі як CRISPR-Cas9 та одноклітинне секвенування, відкривають нові можливості для геномних досліджень, але також ставлять перед нами нові, раніше невідомі, етичні, правові та соціальні виклики. Геноміка значно вплинула на зростання сучасної наукомісткої економіки, особливо додаючи до розвитку біоінформатики, біомедицини, та біотехнології, пропонуючи нові терапії та діагностичні інструменти, а також сприяючи розвитку точної медицини. Перехід до технологій секвенування нового покоління (NGS) та вдосконалення в обчислювальній техніці зробили геномну інформацію дешевішою та доступнішою. Майбутнє геноміки обіцяє радикальні зміни в охороні здоров'я, сільському господарстві та збереженні біорізноманіття, пропонуючи точнішу діагностику та персоналізоване лікування. Прогрес у цій галузі є наслідком роботи великої когорти науковців в усьому світі та міжнародної співпраці між ними, що відкрила і продовжує відкривати нові горизонти для покращення життя людства.

Ключові слова: геноміка, геном людини, біоінформатика, філогенетика, бази даних, секвенування, генетичні проєкти.

¹ Кафедра ботаніки, Ужгородський національний університет, вул. Волошина, 32, Ужгород, 88000, Україна; e-mail: taras.oleksyk@uzhnu.edu.ua; yaroslava.hasynets@uzhnu.edu.ua; roman.kish@uzhnu.edu.ua

² Кафедра ентомології та збереження біорізноманіття, Ужгородський національний університет, вул. Волошина, 32, Ужгород, 88000, Україна; e-mail: vladyslav.mirutenko@uzhnu.edu.ua

³ Кафедра системного аналізу і теорії оптимізації, Ужгородський національний університет, вул. Університетська, 14, Ужгород, 88000, Україна; e-mail: andrii.voshchepynets@uzhnu.edu.ua

The era of great genomic discoveries: a brief outline of the history of genome diversity science

Oleksyk T.¹, Mirutenko V.², Hasynets Ya.¹, Kish R.¹, Voshchepynets A.³

In the dynamic world of modern biology, the rapid development of the new science of genomics – which merges biological methods with computer technologies – has opened up vast possibilities for understanding the mysteries of living organisms, from the simplest bacteria to humans. Despite numerous discoveries in the field of genome structure and function, we are still at the early stages of the emergence and development of this discipline. Global sequencing initiatives such as the Human Genome Project (HGP), HapMap, the 1000 Genomes Project, and others have expanded our understanding of human genetic diversity in the context of population history and the evolution of our species. Genomic methods have revolutionized medicine by enabling the development of personalized treatment strategies tailored to the genetic characteristics of each individual. Modern technologies such as CRISPR-Cas9 and single-cell sequencing are opening new frontiers for genomic research, while also presenting novel ethical, legal, and social challenges previously unknown. Genomics has significantly contributed to the growth of today's knowledge-based economy, particularly by advancing bioinformatics, biomedicine, and biotechnology. It offers new therapies and diagnostic tools and facilitates the rise of precision medicine. The shift to next-generation sequencing (NGS) technologies and advancements in computing have made genomic information more affordable and accessible. The future of genomics promises radical changes in healthcare, agriculture, and biodiversity conservation, offering more accurate diagnostics and personalized treatments. Progress in this field is the result

of the dedicated work of a large cohort of scientists worldwide and the international collaboration among them, which has opened – and continues to open – new horizons for improving human life.

Key words: genomics, human genome, bioinformatics, phylogenetics, databases, sequencing, genetic projects.

¹ Department of Botany, Uzhhorod National University, 32, Voloshyna Str., Uzhhorod, 88000, Ukraine; e-mail: taras.oleksyk@uzhnu.edu.ua; yaroslava.hasynets@uzhnu.edu.ua; roman.kish@uzhnu.edu.ua

² Department of Entomology and Biodiversity Conservation, Uzhhorod National University, 32, Voloshyna Str., Uzhhorod, 88000, Ukraine; e-mail: vladyslav.mirutenko@uzhnu.edu.ua

³ Department of Systems Analysis and Optimization Theory, Uzhhorod National University, 14, Universytetska Str., Uzhhorod, 88000, Ukraine; e-mail: andrii.voshchepynets@uzhnu.edu.ua

Дослідження у світі геноміки

У широкому та постійно змінюваному світі біології, геноміка змушує нас переосмислити традиційні поняття здоров'я та хвороби, розширюючи горизонти можливого в медичній науці. Ця, відносно новий науковий напрямок зміцнив містки між різними науковими галузями, такими як біотехнологія, біомедицина і біоінформатика. сприяючи міждисциплінарним дослідженням та відкриттям. Геном є сукупністю всієї генетичної інформації організму, закодованої у дезоксирибонуклеїновій кислоті (ДНК) (у випадку більшості організмів) або рибонуклеїновій кислоті (РНК – для деяких вірусів). Він включає як гени, які містять інструкції для синтезу білків та інших молекул, так і некодуючі послідовності, які регулюють активність генів та виконують інші важливі функції. Геном визначає унікальний набір характеристик організму та є ключем до розуміння його розвитку, функціонування та еволюції.

Геномні дані змушують нас переосмислити традиційні поняття здоров'я та хвороби, розширюючи горизонти можливого у медичній науці. Дослідження геномного різноманіття прискорює відкриття в області еволюційної біології, дозволяючи нам краще зрозуміти історію життя на Землі. У свою чергу, це обіцяє революцію в області збереження біорізноманіття, надаючи нам інструменти для захисту та відновлення унікальних генетичних ресурсів, започатковує нову еру в селекції рослин та тварин, відкриваючи шлях до створення більш стійких та продуктивних сортів. У той же час це ставить перед нами нові етичні, правові та соціальні виклики, які потребують глибокого розуміння та відповідального підходу. Щоб зрозуміти розвиток науки за останні три десятиліття, потрібно мати уявлення про історичний контекст технологічного прогресу та історію епохи великих геномних відкриттів, від ранніх днів секвенування одного геному, до новітніх ініціатив які готують підґрунтя для майбутнього.

Наука про геноми трансформувала наше розуміння людської генетики та заклала основу для багатьох інших відкриттів. Технологічний про-

грес відіграв ключову роль в її розвитку – від ранніх днів секвенування методом Сенгера до більш передових технологій наступного покоління (NGS) (International Human ... 2001; Venter et al. 2001; Mardis 2008; Shendure et al. 2008). Такі технологічні стрибки зробили геномні дані більш доступними, спричинивши зростання популярності і кількості геномних досліджень. Проекти на кшталт HGP (Проект Геном Людини) разом з іншими, такими як *HapMap* та Проект 1000 геномів, значно розширили наші знання про різноманіття генетики людини (International ... 2001; Nurk et al. 2022; The 1000 ... 2015). Національні геномні проекти у різних країнах продовжують поглиблювати і поширювати світові зусилля з вивчення генетичного різноманіття та його впливу на здоров'я та хвороби (Cavalli-Sforza 2003; UK10K ... 2015; Oleksyk et al. 2021; Oleksyk et al. 2022).

З огляду на майбутнє, ми стоїмо на порозі ще більш захоплюючих відкриттів. Нові технології обіцяють розширити межі можливого у охороні здоров'я, сільському господарстві та збереженні біорізноманіття (Doudna, Charpentier 2014; Human Cell ... 2017; Topol 2014; Church 2005). Однак зі швидким прогресом виникає необхідність розглядати етичні, правові та соціальні наслідки технологічного прогресу спричиненого доступністю інформації (Caulfield, McGuire 2012; Lander 2011; Knoppers, Chadwick 2005).

Мета цієї історично-оглядової роботи – не тільки констатація досягнень геноміки в минулому та її поточного стану, це також погляд у багатообіцяюче майбутнє, про те, як ці відкриття можуть покращити наше життя та світ навколо нас.

Секвенування геному людини

Проект Геном Людини (HGP) став віхою в історії біологічних наук, відкривши двері до нової ери генетичних досліджень. Започаткований у кінці 20-го століття, цей амбітний проект мав за мету не лише картографувати, але й послідовно визначити всю послідовність ДНК у людському геномі, що є фундаментальним кроком для глибшого розуміння генетичних основ життя. Основна мета HGP полягала в тому, щоб зрозуміти функції,

закодовані в нашій ДНК, від механізмів розвитку та функціонування організму до генетичних складових захворювань людини.

Ідея Проєкту Геном Людини виникла в середині 1980-х років і стала одним з найамбітніших викликів у світі науки. Вчені прагнули не просто розшифрувати послідовність 3 мільярдів хімічних пар основ, які утворюють ДНК людини, але й ідентифікувати всі гени, закодовані в ній, що становить приблизно 20 000–25 000 генів. Ця масштабна задача вимагала не лише новітніх технологій та методологій, але й глибокого розуміння етичних, правових та соціальних аспектів, пов'язаних з доступом та використанням генетичної інформації. Впровадження програми *ELSI (Ethical, Legal, and Social Implications)* в рамках HGP стало важливим кроком у вирішенні цих питань, забезпечуючи, щоб наукові відкриття йшли рука об руку з етичними міркуваннями.

Два роки потому, використовуючи найкращі доступні технології для секвенування ДНК та доводячи їх до абсолютних меж, проєкт представив надзвичайно високоякісну послідовність людського геному, яка була майже повною, охоплюючи більше, ніж 90 % людського геному.

Однією з фундаментальних цілей HGP було також створення культури відкритості та співпраці в науковому співтоваристві, що передбачало безперешкодний обмін даними та результатами досліджень. Проєкт стикався з викликами, такими як дебати щодо методів секвенування геному, занепокоєння щодо патентування генетичної інформації та етичні наслідки доступу до генетичних даних. Нарешті, у 2001 році HGP та альтернативний приватний Проєкт Крега Вентера (Venter et al. 2001) і компанії *Celera Genomics* кожен оголосили про створення «чорновиків» послідовності людського геному.

HGP був глобальним зусиллям з великими внесками від США, Великобританії та інших країн. Джеймс Вотсон один з вчених, які відкрили структуру ДНК (Watson, Crick 1953), відіграв ключову роль у запуску Проєкту та встановленні його цілей. Доктор Френсіс Коллінз, який змінив Вотсона, відіграв вирішальну роль у організації Проєкту та виступав за вільний доступ до даних для всіх дослідників.

Цей підхід не тільки сприяв прогресу в геномних дослідженнях, але й заклав основу для майбутніх міжнародних проєктів у різних галузях науки. Під керівництвом доктора Коллінза команда HGP досягла своєї мети до 2003 року, опублікувавши так званий «чорновик» повної послідовності люд-

ського геному – високоякісну послідовність людського геному, яка була майже повною, охоплюючи більше, ніж 90 % довжини хромосом (Collins et al. 2003).

Врешті решт, «чистовик» геному був опублікований майже два десятиліття потому – у 2022 році. Повний геном людини включає безпрогалинну збірку від теломери до теломери (T2T) для всіх 22 людських аутосом та хромосоми X, що складаються з 3 054 815 472 нуклеотидів ДНК, а також мітохондріальний геном довжиною 16 569 нуклеотидів. Ця публікація усунула перешкоди для подальшого аналізу всіх 100 % людського геному, включаючи всі центромерні та теломерні регіони хромосом.

Розвиток нових технологій секвенування ДНК

Завершення HGP стало переломним моментом для технологій секвенування ДНК. Потреба отримувати великі обсяги інформації цілих геномів призвела до розробки швидших, доступніших та більш точних методів секвенування ДНК. Таким чином, десятиліття після HGP відзначилося значними змінами у підходах, зокрема переходом від секвенування за методом Сенгера, яке використовувалося до цього часу включно з HGP і потребувало значних коштів, до секвенування нового покоління (NGS), яке повністю трансформувало галузь геноміки, зробивши секвенування набагато дешевшим і швидшим.

Традиційно секвенування ДНК було тісно пов'язане з розумінням структури ДНК. Ранні методи, такі як техніка Максама-Гілберта та секвенування за Сенгером. Останнє набуло популярності у 1970-х роках як основний метод і «золотий стандарт» якості даних на початкових етапах HGP. Цей метод і досі використовується як стандарт для верифікації результатів отриманих за допомогою інших технологій. Однак метод Сенгера був занадто повільним, відносно дорогим і не підходив для нових великих проєктів, таких як секвенування повних геномів.

Розробка методу піросеквенування Полом Нюреном (Ronaghi 2000) позначило початок ери секвенування нового покоління (NGS) – ери швидших і ефективніших технологій. Піросеквенування – це метод секвенування ДНК, який вимірює випромінювання світла при додаванні нуклеотиду до синтезованого ДНК ланцюга. Цей метод дозволяє секвенувати одразу повний геном, порізаний на мільярди фрагментів однакової довжини. Це означає що дані секвенування потрібно впорядкувати у геномні карти вздовж лінійних координат хро-

мосом. Цей процес збірки (асемблювання), вимагає інтенсивних біоінформатичних обчислювань для яких потрібні швидкі сервери, великі кластери і об'ємні блоки пам'яті. Ранні NGS методи, такі як піросеквенування мали кілька важливих недоліків, зокрема обмежену довжину зчитування фрагментів ДНК, що ускладнювала асемблювання геному, та труднощі з точним визначенням гомополімерних регіонів, де один і той самий нуклеотид повторювався багато разів підряд.

У 2010-х роках більш сучасні, дешеві і доступні підходи були розвинуті компаніями *Ion Torrent* та *Illumina*, поставши новими стандартами у галузі NGS. Платформа *Illumina* заснована на принципі секвенування синтезом з використанням реверсних термінаторів, що забезпечує високу пропускну здатність та точність. На противагу цьому *Ion Torrent* використав напівпровідникові сенсори для безпосереднього виявлення потоку атомів водню під час синтезу ДНК, що дозволило швидко та ефективно читати послідовність нуклеотидів без використання флуоресцентних міток, зменшуючи вартість та складність процесу. Обидві технології в змозі одночасно секвенувати мільйони фрагментів ДНК, забезпечуючи глибину покриття та можливість детального аналізу геномів. Перехід до нових платформ NGS значно збільшив швидкість та знизив вартість секвенування, роблячи геномні дослідження більш доступними для широкого кола дослідників.

Технології NGS дозволили проводити секвенування ДНК у величезних масштабах. У цю епоху були зроблені внески такими вченими як Джордж Черч та Крейг Вентер, які відіграли ключові ролі у підвищенні швидкості та доступності геномних досліджень, і внесли великий вклад у розуміння геноміки. Черч зробив значний внесок у розробку методів секвенування наступного покоління, а Вентер був одним з перших, хто використав ці технології для секвенування повного людського геному, відкриваючи нові можливості для генетичних досліджень. Ці досягнення мали глибокий вплив на геноміку, створивши умови до прогресу в персоналізованій медицині (Church 2005; Topol 2014), генетичних дослідженнях та розумінні генетичної складової складних і комплексних захворювань.

У 2020-х роках NGS стає центральним інструментом у різних наукових галузях, з тенденціями до розвитку ще більш передових технологій. Секвенування широко використовується у різних областях, включаючи клінічну діагностику, геноміку раку та еволюційну біологію. Інновації, такі

як секвенування окремих клітин та секвенування з довгими зчитуваннями, набувають поширення. Інтеграція великих даних, штучного інтелекту та машинного навчання для аналізу даних зростає, вимагаючи все більших знань і можливостей у сфері біоінформатики.

Вирішальна роль біоінформатики у геноміці

У контексті сучасних біологічних досліджень біоінформатика відіграє ключову роль у розшифруванні та аналізі масивних наборів даних, отриманих у результаті вивчення фундаментальних структур життя, зокрема ДНК і РНК. У сфері геноміки, яка зосереджена на всебічному дослідженні генетичного матеріалу організмів, значення біоінформатики неможливо переоцінити. Ця дисципліна надає дослідникам необхідні інструменти та методології для структурування та інтерпретації генетичної інформації, дозволяючи інтегрувати різноманітні дані в єдину когерентну картину геному (Mount 2007; Pevzner 2000). Через це вчені можуть визначати функціональність різних сегментів ДНК, розкриваючи механізми, що лежать в основі генетичних процесів та їх вплив на фенотип організму.

Біоінформатика є сумішшю біології, комп'ютерних наук та математики, і цей напрямок наразі активно розвивається. Першопочатково біоінформатика в основному займалася читанням та порівнянням послідовностей ДНК. HGP надав особливо сильного поштовху для розвитку цієї науки, тому що вченим потрібен був спосіб впоратися з усіма даними людського геному. Одне з ключових завдань у біоінформатиці – вирівнювання послідовностей ДНК, щоб знайти схожості та відмінності (Altschul et al. 1990). Це допомагає скласти геном з багатьох маленьких фрагментів ДНК та зрозуміти, як пов'язані різні гени між собою. Біоінформатики з HGP мусили створити спеціальні алгоритми для вирівнювання послідовностей або бази даних, які зберігають усю цю генетичну інформацію. Ці інструменти використовуються не тільки в академічних дослідженнях, вони також застосовуються у медицині для допомоги у розумінні захворювань та в інших галузях науки.

На сьогодні сфера біоінформатики стосується обробки та розуміння величезних обсягів різноманітних геномних даних, а завдяки прогресу в технологіях, таких як хмарні обчислення та машинне навчання, вона стала ще більш потужною. Зі збільшенням кількості генетичних даних біоінформатика стикається з великими викликами, такими як спосіб зберігання всієї цієї інформації та продовження вдосконалення інструментів,

щоб вони були доступні для всіх. Але з постійним прогресом у технологіях, біоінформатика готова відігравати ще більшу роль у допомозі розуміння життя на генетичному рівні та надавати поради щодо генетичних проблеми зі здоров'ям з великою точністю.

Швидкий розвиток технологій секвенування, особливо секвенування нового покоління (NGS), сприяв реалізації масштабних проєктів, значно прискорив процес секвенування та водночас знизив їхню вартість. Ціна секвенування людського геному значно зменшилася протягом останнього десятиліття, що часто ілюструється на графіку «вартість за геном», демонструючи стрімке падіння цін з появою технологій NGS. Таке зниження вартості відбувається значно швидше, ніж темпи технологічного розвитку в інших сферах, що часто порівнюють із законом Мура, роблячи секвенування геному більш доступним і широко використовуваним. З'явилися бази даних, які містять терабайти інформації про геноми, які поповнюються і зростають щодня. В результаті, наявність величезної кількості геномних даних перетворила аналіз цих даних, а саме біоінформатику, на критичну ділянку в науці про геноми. Проблема полягає не стільки в генерації даних, скільки в їх інтерпретації та аналізі і вимагає розширення можливостей та розвитку інструментів у галузі біоінформатики.

Бази даних геномного різноманіття

Бази даних геномів подібні до величезних бібліотек, які зберігають інформацію про послідовність генетичного матеріалу у цифровому форматі. Вони містять цілі геноми, окремі гени та варіанти, такі як SNP (одно-нуклеотидні поліморфізми), делеції (вирізки), інсерції (вставки), транслокації (переміщення фрагментів) та інші зміни в нашій ДНК, які роблять кожний організм унікальним. Ці бази даних полегшують доступ та використання величезної кількості генетичної інформації дослідникам по всьому світу і є важливими ресурсами для вчених, які намагаються зрозуміти генетичні відмінності між людьми та їх вплив на здоров'я носіїв, ризик захворювань та індивідуальну реакцію на ліки.

Одним з перших великих геномних проєктів у цій галузі був Міжнародний Проєкт Карти Гаплотипів (*International HapMap*) започаткований у 2002 році. Основна ідея проєкту полягала в тому, щоб зробити каталог SNP, та визначити, як ці варіації групуються разом у різних популяціях людей по всьому світу. У першій стадії HapMap було включено зразки від 269 осіб, які представ-

ляли чотири популяції з різних частин світу: по 30 тріо (дитина та обоє батьків) з Африки (YRI) і США (CEU), а також 44 зразки з Японії (JPT) та 45 з Китаю (CHB). Це дозволило зібрати важливу інформацію про генетичну різноманітність в різних етнічних групах та визначити комбінації генетичних варіацій розташовані у локусах вздовж хромосоми, тобто гаплотипи, що успадковуються разом у людському геномі. Гаплотипна мапа дозволила вченим краще розуміти зчеплення між генетичними маркерами в різних популяціях, щоб більш ефективно ідентифікувати локуси генів, асоційовані з конкретними захворюваннями, а також допомогла зрозуміти, як генетичні відмінності між людьми з різним походженням впливають на здоров'я. В той же час популяційні дані повних геномів дозволили виявити генетичні варіації під впливом природнього добору, які сприяли адаптації до специфічних умов середовища, вивчати еволюційні зміни в часі та аналізувати зв'язок між генетикою та фенотипами.

Розширення етнічної різноманітності включених в HapMap зразків дало можливість краще зрозуміти генетичну структуру та історію різних популяцій, а також сприяло виявленню нових генетичних варіацій, які можуть бути пов'язані з певними захворюваннями або фенотиповими ознаками. Остання, третя фаза HapMap аналізувала зразки від 1184 осіб, які представляли 11 популяцій з різних частин світу, включаючи нові популяції з Африки, Азії, Європи та Америки, які не були представлені у попередніх публікаціях. Після завершення HapMap, який зосереджувався на вивченні генетичних варіацій у різних популяціях та створенні гаплотипної карти людського геному, наукова спільнота прагнула отримати щонайбільш повну картину генетичного різноманіття людства.

Проєкт 1000 геномів (*1000 Genomes Project*), що був розпочатий у 2008 році з метою створення найбільш детальної карти генетичних варіацій людини, можна розглядати як логічне продовження та розширення HapMap. Проєкт прагнув каталогізувати всі генетичні варіації, що зустрічаються хоча б у 1 % у одній з популяцій. Проєкт зібрав і проаналізував геноми більш ніж 2500 осіб з різних етнічних груп, значно розширивши розуміння генетичної структури людських популяцій та генетичної варіабельності порівняно з HapMap.

Проєкт 1000 геномів зробив великий крок вперед, секвенуючи ДНК людей з усього світу, багато з яких ніколи раніше не були включені у геномні дослідження. Він був прикладом успішної міжнародної співпраці в галузі геноміки, де різні групи

ділилися даними, ресурсами та експертизою для досягнення спільної мети. Популяції включалися з метою забезпечення широкого географічного та етнічного охоплення, щоб створити детальну карту генетичних варіацій всього людського виду. По перше, в проєкт включили популяції з різних частин світу, щоб забезпечити глобальне представництво генетичної різноманітності. По друге, важливо було вибрати популяції, які добре представляють генетичну структуру відповідних регіонів, уникаючи надмірного зосередження на дуже специфічних або ізольованих групах, щоб результати були максимально репрезентативними. Практичність збору та аналізу зразків також відіграла роль. Популяції, з яких було легше отримати зразки з достатньою кількістю інформації про походження та здоров'я учасників, мали більшу ймовірність включення. Карлос Д. Бустаманте, вчений венесуельського походження наполіг на включенні в Проєкт 1000 геномів генетично змішані популяції з Центральної та Південної Америки, висвітлюючи, як різні групи людей еволюціонували та адаптувалися з часом. Ці підходи і принципи допомогли забезпечити Проєкт 1000 геномів широким та різноманітним представництвом генетичних варіантів, що є критично важливим для розуміння генетичної структури людського виду та взаємозв'язків між генетикою та захворюваннями.

Паралельно великий крок у розбудові картини нашої генетичної різноманітності був зроблений завдяки Проєкту різноманітності людського геному (HGDP), який базувався на попередній роботі Луїджі Лука Каваллі-Сфорца зі Стенфордського університету (Cavalli-Sforza, Feldman 2003), який був одним з ініціаторів проєкту, піонером у галузі генетики людини та зробив значний внесок у його розвиток. Він був одним з перших, хто використовував генетичні дані для відстеження історії та переміщень людських популяцій. Зразки для HGDP були зібрані різними дослідницькими групами та індивідуальними вченими протягом кількох років. Каваллі-Сфорца та його колеги працювали над збором зразків з різних куточків світу, особливо зосереджуючись на ендемічних, аборигенних та ізольованих популяціях. HGDP мав на меті створити детальну карту генетичної різноманітності корінних народів світу і зумів секвенувати повні геноми 1 043 осіб приблизно з 52 різних популяцій, значно збагативши наші знання про людську генетичну різноманітність на Землі.

Міжнародні геномні проєкти, такі як Проєкт 1000 геномів (G1K) та Проєкт різноманітності

людського геному (HGDP), стали переломними моментами у розумінні широкого спектру людської генетичної різноманітності. Вони допомогли вченим виявити генетичні корені захворювань, які можуть значно відрізнятися серед різних етнічних груп, пролили світло на еволюцію і міграції людей. Розуміння повної картини геномного різноманіття має вирішальне значення для розвитку персоналізованої медицини, гарантуючи ефективність підходів для людей з різних популяцій.

Глобальні та національні геномні проєкти: подолання генетичної прірви

При всіх можливостях 1000 геномів та Проєкт різноманітності людського геному не могли представити кожен популяцію світу. Нерівномірність зібраних зразків призвела до появи «геномних пустель», де генетичні дані про місцеве населення не представлені у геномних базах даних, а тому недоступні для дослідників. Щоб заповнити ці пробіли, з'явилися національні геномні проєкти, які зосередилися на генетичній різноманітності конкретних країн. Ці ініціативи є ключовими для створення каталогів унікального генетичного профілю місцевого населення. Такі проєкти як *Genome England* та *UK Biobank* у Великобританії та програма *All of Us* у Сполучених Штатах Америки, очолюють цей напрям, секвенуючи геноми десятків тисяч осіб у своїх країнах.

Національні геномні проєкти відкрили нові можливості для розуміння людської генетичної різноманітності, зосереджуючись на конкретних популяціях, які не були повністю представлені у попередніх міжнародних дослідженнях. Знаковим прикладом може слугувати Проєкт *deCODE* в Ісландії, де дослідники використовували детальні генетичні та медичні записи країни для вивчення генетичних основ різних захворювань, таких як рак та хвороба Альцгеймера. Цей проєкт не тільки просунув наше розуміння цих захворювань, але й відіграв роль у відповіді Ісландії на виклики пандемії COVID-19. Незважаючи на деякі етичні занепокоєння щодо конфіденційності, *deCODE* став прецедентом для майбутніх національних геномних проєктів.

У Великобританії Проєкт *UK10K* мав на меті секвенувати геноми 10 000 осіб для виявлення рідкісних генетичних варіантів, пов'язаних з захворюваннями та іншими фенотипами (UK10K... 2015). Аналогічно Проєкт 100 000 геномів, також у Великобританії, став одним з найбільших національних зусиль у світі, інтегруючи геномну інформацію в національну систему охорони здоров'я. Водночас, оскільки генетичні дані стають більш

доступними, забезпечення конфіденційності та безпеки цієї чутливої інформації є надзвичайно важливим.

Національні проекти доповнюють глобальні зусилля, заповнюючи прогалини в нашому генетичному знанні, особливо щодо рідкісних та регіон-специфічних генетичних варіантів. Незважаючи на всі зусилля, все ще існує значний дисбаланс у генетичних дослідженнях між різними регіонами. Багато країн, що розвиваються, особливо в Африці та Південній Америці, відстають через обмежені ресурси та брак наукового досвіду. Хоча по всьому світу триває багато геномних проєктів, як міжнародних, так і національних, досить значна частина світового населення залишається недостатньо представленою в генетичних дослідженнях і повна картина геномного різноманіття людства залишається недосяжною. Це залишає значну частину людства поза межами генетичних досліджень.

Секвенування давніх геномів

Секвенування давніх геномів відкриває нові горизонти для розуміння нашого минулого та впливу історичних подій на сучасну генетичну структуру, дозволяючи глибше зрозуміти, як минуле формує наше генетичне різноманіття. За останні 30 років, підходи до древньої ДНК пройшли через кілька значних етапів розвитку, які кардинально змінили наше розуміння людської еволюції, міграції та історії хвороб. Поєднання генетичних даних з археологічними та антропологічними знахідками дозволяє створювати більш повну картину історії та міграцій давніх людських популяцій (Krause, Trappe 2016; Reich 2018; Roberts 2015; Willerslev, Davison 2021).

Ранні спроби секвенування ДНК з давніх зразків у 1990-і роки часто були обмежені короткими фрагментами мітохондріальної ДНК. Це було пов'язано з технічними обмеженнями секвенування методом Сенгера та високим рівнем забруднення зразків сучасною ДНК. Завдяки секвенуванню нового покоління (NGS) стало можливим проводити більш широкомасштабні та детальні дослідження давніх геномів, включаючи секвенування повних геномів з невеликої кількості збереженої ДНК. Завдяки вдосконаленню методів екстракції ДНК, розробці стратегій для боротьби з забрудненням (контамінацією) зразка та появи секвенування NGS, стало можливим отримувати довші послідовності кращої якості ДНК з давніх зразків. Розвиток палеомікробіоміки дозволив досліджувати мікробіоми давніх зразків, надаючи інформацію про дієту, хвороби

та взаємодію між мікробами та їхніми носіями в далеку давнину.

Однією з ключових подій стало часткове секвенування геному неандертальця у 2010 році під керівництвом німецького вченого шведського походження Сванте Паабо (Paabo 2014). Це дозволило вперше порівняти геноми неандертальців із геномами сучасних людей, виявивши спільні генетичні елементи. Визнанням цієї роботи була Нобелівська премія, яку Паабо отримав у 2022 році. Це дослідження виявило, що сучасне населення за межами Африки має невеликий відсоток неандертальської ДНК, що свідчить про історичне змішування між двома видами.

Крім того, С. Паабо у співробітництві з Девідом Рейхом з Гарвардського Університету відіграли ключову роль у відкритті та аналізі геному денісовців – групи давніх людей, про існування яких стало відомо лише завдяки генетичним дослідженням. Денісовці – це група давніх людей, ідентифікованих виключно за їхнім генетичним матеріалом, отриманим з кількох невеликих фрагментів кісток та зубів, знайдених у печері Денісова на Алтаї. Аналіз ДНК показав, що денісовці були відокремленою групою, близькою до неандертальців, але зі своєю унікальною генетичною ідентичністю. Це відкриття вказало на складніші взаємодії між різними групами давніх людей, ніж вважалося раніше. Що найцікавіше, денісовська ДНК була виявлена не тільки у рештках, знайдених у Сибіру, але й у геномах сучасних людей, особливо у популяціях Меланезії та деяких інших азійських та океанських груп. Це свідчить про історичні змішування між денісовцями та анатомічно сучасними людьми.

Секвенування пангеномів

Пангеном описує повний набір генів усіх штамів, видів або індивідів у певній групі організмів, включаючи ядерний геном зі спільними генами та додатковий геном з унікальними генами для окремих індивідів або штамів. Він складається з двох основних компонентів. Ядерний геном (Core Genome) – набір генів, які є спільними для всіх індивідів або штамів у групі. Ці гени зазвичай відповідають за основні життєві функції організму. Додатковий геном (Accessory Genome) – набір генів, які присутні тільки в деяких індивідах або штамів. Ці гени можуть надавати специфічні адаптації або властивості, такі як стійкість до антибіотиків у мікроорганізмів або спеціалізовані метаболічні здібності.

Активні пангеномні проєкти, як-от Проєкт пангеному людини та Проєкт пангеному мікро-

біому, прагнуть створити більш комплексні та репрезентативні довідкові геноми, що відображають генетичну різноманітність та сприяють розумінню функціональної біології організмів (McInerney et al. 2017; Liao et al. 2023). Пангеномні аналізи можуть виявити раніше невідомі гени та метаболічні шляхи, що відкриває нові можливості для біотехнологічних застосувань. Пангеномні проекти дозволяють вченим досліджувати генетичну різноманітність на рівні виду або популяції, що є ключовим для розуміння еволюційних процесів, адаптацій та взаємодій між різними групами організмів. У медицині пангеномні проекти можуть допомогти ідентифікувати генетичні фактори, які пов'язані з хворобами, розумінням мікробіому людини та розробкою нових лікарських препаратів. У сільському господарстві пангеномні дані можуть сприяти відбору та розробці культур з покращеними властивостями, такими як стійкість до шкідників та хвороб, а також адаптація до змін клімату. Пангеномні проекти надають цінну інформацію для збереження біорізноманіття, дозволяючи оцінити генетичну варіабельність та стійкість популяцій до зовнішніх впливів.

Системна біологія

Під кінець 2010-х і на початок 2020-х років інтеграція великих наборів даних з різних омівних платформ (геноміка, транскриптоміка, протеоміка, метаболоміка) за допомогою найсучасніших біоінформатичних інструментів та алгоритмів стала ключовою для розуміння складних біологічних систем на системному рівні. Застосування штучного інтелекту та машинного навчання в функціональній геноміці почало відкривати нові шляхи для аналізу та інтерпретації складних генетичних даних, сприяючи відкриттю нових біологічних механізмів та потенційних терапевтичних мішеней.

Системна біологія інтегрує дані з різних біологічних дисциплін для побудови та аналізу комплексних біологічних систем (Kitano 2002; Alon 2020).

Функціональна геноміка включає аналіз експресії генів на рівні транскриптомів (РНК) за допомогою таких методів, як секвенування РНК (*RNA-seq*). Це дозволяє визначити, які гени активуються або пригнічуються в певних умовах або тканинах. Аналізуючи, як різні генетичні варіації впливають на фенотипи, функціональна геноміка допомагає з'ясувати молекулярні механізми, що лежать в основі розвитку захворювань, та ідентифікувати потенційні мішені для лікування.

Транскриптоміка займається вивченням всіх транскриптів (РНК молекул), включаючи інформаційну РНК (іРНК), рибосомну РНК (рРНК),

транспортну РНК (тРНК) та некодуючі РНК в клітині або організмі в певний момент часу. Транскриптоміка дає змогу аналізувати експресію генів та їх регуляцію, що допомагає зрозуміти функціональні аспекти клітинної активності та відповідь на різні зовнішні стимули. У свою чергу, протеоміка досліджує весь набір білків, які виробляються або модифікуються клітиною або організмом.

Протеоміка допомагає зрозуміти, як зміни на рівні геному впливають на функції білків та їх взаємодії, як гени та білки взаємодіють між собою та з іншими компонентами клітини, формуючи складні мережі, які регулюють функції на рівні клітин та організмів.

Функціональна геноміка і протеоміка є ключовою складовою системної біології, яка прагне інтегрувати різні типи даних (геномні, транскриптомні, протеомні та інші) для створення всеосяжної моделі живих систем.

На початок 2000-х років розвиток технології дозволив одночасно аналізувати експресію тисяч генів, що відкрило шлях для масштабних досліджень транскриптомів. Поява NGS технологій радикально змінила функціональну геноміку і транскриптоміку, зробивши секвенування геномів, транскриптомів та епігеномів швидшим і більш доступним. Розвиток *CRISPR-Cas9* та інших систем редагування геномів відкрив нові можливості для функціональних досліджень, дозволяючи вченим цілеспрямовано модифікувати геноми для вивчення функцій конкретних генів.

Одноклітинне секвенування стало новою віхою в революції біомедичних досліджень, надаючи глибоке розуміння біологічної гетерогенності та механізмів захворювань на рівні транскриптомів окремих клітин. Цей метод дозволяє виявляти варіації експресії генів на клітинному рівні, що є критично важливим для розуміння клітинної гетерогенності за нормальних та патологічних станів. Одноклітинне секвенування допомагає відстежувати шляхи розвитку та диференціації клітин, виявляючи, як генетичні та епігенетичні зміни впливають на долю окремих клітин.

Метаболоміка є важливою складовою системної біології, оскільки вона забезпечує глибоке розуміння метаболічних процесів у живих організмах. Ця дисципліна займається вивченням метаболітів – невеликих молекул, які є продуктами або проміжними речовинами метаболічних реакцій у клітинах, тканинах, організмах і біологічних рідинах. Метаболоміка відіграє ключову роль у системній біології, надаючи інструменти для глибокого аналізу метаболічних процесів та

їх впливу на функціонування організму на молекулярному рівні. Метаболоміка доповнює інші «омічні» дослідження, такі як геноміка, транскриптоміка та протеоміка, забезпечуючи комплексне розуміння біологічних систем. Така інтеграція дозволяє створити комплексне розуміння біологічних систем, їх регуляції та адаптації до змінних умов середовища.

Ключову роль у системній біології відіграє також **епігенетика**, наука про механізми, які регулюють активність генів без змін у послідовності ДНК. Епігенетика розширює наше розуміння генетики, показуючи, що фенотип організму визначається не тільки генами, але й регуляцією їх експресії, що відкриває нові перспективи для вивчення біології та лікування захворювань. Епігенетичні зміни можуть бути успадковані під час клітинного поділу та навіть передатися між поколіннями, але не впливають на первинну послідовність нуклеотидів у ДНК. Епігенетичні механізми включають ДНК-метилування, модифікації гістонів та некодуючі РНК, які разом впливають на структуру хроматину та доступ генів до транскрипційних процесів у клітині. Зростання інтересу до епігеноміки та вивчення пост-транскрипційних та пост-трансляційних модифікацій розширило розуміння регуляції генної експресії та її впливу на фенотип. Багато захворювань, включаючи рак, неврологічні розлади та метаболічні синдроми, пов'язані з епігенетичними змінами, що вказує на потенціал для розробки нових лікувальних стратегій.

Порівняльна геноміка та філогенетика

Початок порівняльної геноміки можна віднести до моменту, коли стало можливим секвенування ДНК. Геном *Haemophilus influenzae* став першим повністю просеквенованим геномом вільно живучого організму в 1995 році. Перші проекти з секвенування, такі як Проект геному людини та геноми модельних організмів (наприклад, шимпанзе, собаки, щура, миші, дрозофіли, нематоди, дріжджів, рижійки), забезпечили основу для порівняльних аналізів. Ці модельні організми були обрані для секвенування з різних причин, включаючи їх відносну простоту, короткі життєві цикли, легкість утримання в лабораторних умовах та велику кількість наявної генетичної інформації. Секвенування їхніх геномів відкрило нові можливості для генетичних та біологічних досліджень, забезпечивши фундаментальну інформацію для розуміння біологічних процесів на молекулярному рівні. Поглиблене вивчення геномних варіацій між різними видами дозволяє краще зрозуміти механізми еволюції, такі як горизонтальний пере-

нос генів, дуплікації генів, та їхній вплив на адаптацію та спеціалізацію.

Проект «Genome 10K» (G10K) та «Vertebrate Genomes Project» (VGP) значно розширили обсяг порівняльної геноміки, маючи на меті секвенування та аналіз геномів 10 000 та більше видів хребетних відповідно. Ці проекти покликані забезпечити глибоке розуміння еволюційних процесів та біорізноманіття. Такі ініціативи, як «Insect 10 000 Genomes» (i10K), проекти з геноміки птахів та збереження геноміки, розширюють межі порівняльної геноміки на інші групи організмів, включаючи комах та види, що знаходяться під загрозою зникнення. Це допомагає виявити ключові генетичні фактори, пов'язані з адаптацією, виживанням та збереженням видів.

Використовуючи нову інформацію про різноманітність між біологічними популяціями й видами, порівняльна геноміка стала потужним інструментом у вивченні генетичного коду, дозволяючи ідентифікувати генетичні варіанти, пов'язані з фенотипами притаманними людині у порівнянні з іншими видами. Порівняння геномів дозволяє виявити специфічні генетичні варіанти, які корелюють з певними фенотипами. Ці варіанти можуть включати *SNP*, інсерції, делеції та інші структурні варіації. Аналіз різноманіття геномів дозволяє не тільки ідентифікувати генетичні варіанти, але й досліджувати, як ці варіанти впливають на біологічні шляхи та процеси на молекулярному рівні. Це може включати вивчення впливу мутацій на експресію генів, білкову функцію та взаємодії між білками, що допомагає розкрити механізми розвитку захворювань.

Аналіз геномних послідовностей допомагає ідентифікувати ключові генетичні варіанти, які можуть впливати на адаптацію та виживання видів. Це включає варіанти, пов'язані зі стійкістю до хвороб, змінами в дієті, толерантністю до екстремальних умов середовища та іншими адаптивними рисами. Геномний аналіз дозволяє оцінити рівень генетичного різноманіття в межах та між популяціями видів. Високий рівень генетичного різноманіття зазвичай асоціюється з більшою стійкістю популяцій до змін у навколишньому середовищі та здатністю адаптуватися до нових умов. Визначення генетичних «вузьких місць» або зниження різноманіття може вказувати на популяції, які потребують негайних заходів зі збереження. На основі геномних даних можна розробляти цільові стратегії збереження, які враховують генетичну структуру популяцій, рівень різноманіття та потенційні адаптивні риси. Це може включати

в себе заходи зі збереження та відновлення середовища існування, створення генетичних резерватів, програми розведення з метою збереження генетичного різноманіття та інші заходи.

Порівняльна геноміка дозволяє визначити філогенетичні зв'язки між видами та популяціями, що може виявити історичні міграції, еволюційні розділення та гібридизацію (Delsuc et al. 2005; Ellegren 2008; Koonin 2012). Це розуміння допомагає визначити еволюційно значущі групи для збереження та розробки стратегій управління біорізноманіттям. Аналізуючи геномні послідовності різних популяцій та видів, можна ідентифікувати шляхи міграції та колонізації нових територій. Це допомагає зрозуміти, як географічні та кліматичні зміни впливали на розповсюдження та адаптацію видів. Геномні дані можуть вказувати на історичні події гібридизації між видами або популяціями, які сприяли обміну генетичним матеріалом та виникненню нових генетичних комбінацій. Це розуміння допомагає виявити механізми, за допомогою яких види адаптуються до умов середовища, які змінюються. Філогенетична інформація може слугувати основою для розробки стратегій збереження та управління біорізноманіттям, забезпечуючи збереження генетичного різноманіття та адаптивного потенціалу видів у світі, що швидко змінюється.

Висновки

Геноміка сприяла значному економічному зростанню, зокрема швидкому розширенню біотехнологічного сектору. Вона призвела до розробки нових терапій, діагностичних інструментів і стала рушійною силою в точній медицині. Перехід від секвенування за Сенгером до технологій секвенування нового покоління революціонував геноміку, зробивши її швидшою та більш економічно вигідною. Хмарні обчислення та вдосконалення у галузі біоінформатики також були вирішальними у керуванні та аналізі величезних наборів даних. Найбільш помітний вплив спостерігається у сфері охорони здоров'я, зокрема в персоналізованій медицині, а також у нашому розумінні людської еволюції та генетичної різноманітності.

Водночас цей швидкий розвиток породжує багато суперечливих питань, які потрібно вирішити, щоб повністю зрозуміти вплив геноміки на наше життя. Будуть тривати дебати про суспільні наслідки геномних технологій, включаючи їх вплив на медичне страхування, зайнятість та суспільні нерівності. Права інтелектуальної власності в геноміці, особливо що стосується патентів на гени, продовжуватимуть бути предметом

правових суперечок. Питання, такі як генетична дискримінація, згода на генетичне тестування та етичні наслідки технологій редагування геному, як-от CRISPR, будуть на передньому плані.

Вивчення геномного різноманіття людини почалося з амбітного Проєкту Геному Людини, який мав на меті картографувати весь людський геном. Це монументальне завдання, завершене у 2003 році, поклало основу сучасної геноміки, революціонувавши наше розуміння людської біології та хвороб. Ключові постаті, як-от Фредерік Сенгер, який розробив перший практичний метод секвенування ДНК, та Крейг Вентер, відомий своєю роллю в HGP та секвенуванні цілого геному, зіграли вирішальну роль у цих ранніх розробках.

Після HGP такі проєкти, як НарМар та Проєкт 1000 геномів, ще більше розширили наше розуміння людської генетичної різноманітності. Ці зусилля значною мірою були підтримані прогресом у технологіях секвенування, переходом від секвенування за Сенгером до більш ефективних методів NGS. Ця технологічна еволюція не тільки зробила секвенування швидшим та доступнішим, але й дозволила аналізувати складні генетичні дані, що призвело до відкриттів у походженні різних захворювань та розвитку персоналізованої медицини. Поява національних геномних ініціатив стала унікальною можливістю заповнити прогалини та представити популяції, які раніше були ігноровані міжнародними проєктами.

Глобальні та національні геномні проєкти, такі як Проєкт 1000 геномів та Проєкт різноманітності людського геному, зіткнулися з викликом неповного представлення всіх світових популяцій, що призвело до появи «геномних пустель». Відповіддю на це стали національні геномні проєкти, які зосереджуються на генетичній різноманітності конкретних країн, наприклад *Genome England* та *All of Us* (США). Ці ініціативи допомагають заповнити прогалини у генетичних даних, але все ще існує дисбаланс у генетичних дослідженнях між різними регіонами світу.

Секвенування давніх геномів відкриває нові можливості для розуміння історичних подій та їх впливу на сучасну генетичну структуру. Пангеномні проєкти, які описують повний набір генів усіх індивідів групи, розширюють наше розуміння генетичної різноманітності та функціональної біології. Інтеграція великих наборів даних з різних оміксних платформ за допомогою системної біології та передових біоінформатичних інструментів відкриває нові шляхи для аналізу та інтерпретації складних генетичних даних, спри-

яючи відкриттю нових біологічних механізмів та потенційних терапевтичних мішеней.

Сьогодні геноміка перебуває на захоплюючому перехресті, де нові технології, такі як редагування геному CRISPR-Cas9 та секвенування окремих клітин, відкривають нові шляхи для досліджень та терапії. Технологія CRISPR, зокрема, трансформувала генетичні дослідження, дозволяючи точно та ефективно редагувати ДНК, з потенційними застосуваннями від генної терапії до сільського господарства. Інтеграція штучного інтелекту (AI) та машинного навчання (ML) в біоінформатику – ще одна значна сучасна тенденція. Ці технології покращують нашу здатність обробляти та інтерпретувати величезні обсяги даних, отриманих у результаті геномних досліджень, що призводить до більш точного моделювання захворювань та стратегій лікування. Однак це майбутнє не позбавлене викликів. Етичні, правові та соціальні наслідки, такі як питання конфіденційності, генетична дискримінація та справедливий розподіл геномних технологій, будуть ключовими питаннями, які потрібно вирішити в наступному десятилітті.

Дивлячись у майбутнє, можна стверджувати, що потенційний вплив геноміки є величезним. У сфері охорони здоров'я вона обіцяє більш точну діагностику та персоналізоване лікування. У сільському господарстві геноміка може призвести до розробки більш стійких та продуктивних культур. Крім того, вона пропонує інструменти для збереження біорізноманіття, допомагаючи зрозуміти

та зберегти генетичну різноманітність зникаючих видів.

Майбутнє геноміки готується будуватися на своїх видатних досягненнях, де нові технології, такі як CRISPR та секвенування окремих клітин, сприятимуть подальшим проривам. Геноміка продовжить революціонізувати охорону здоров'я, дозволяючи здійснювати більш точну діагностику, цільові терапії та глибше розуміння генетичної основи захворювань. Її вплив буде глибоко відчуватися в охороні здоров'я, сільському господарстві та збереженні біорізноманіття. Однак це майбутнє також несе складні етичні, правові та соціальні виклики, які потребуватимуть уважного розгляду та проактивного управління.

Прогрес у геноміці та біоінформатиці є свідченням людського винахідництва та наполегливості. Від розшифрування геному людини до точного редагування генів ми розкрили таємниці життя, які колись здавалися непереборними. Ця подорож, відзначена неймовірними технологічними досягненнями та співпрацею між націями, не лише про наукові досягнення. Це маяк надії на здоровий, більш сталий світ. Стоячи на порозі нових відкриттів, обіцянка геноміки поглибити наше розуміння життя, покращити людське здоров'я та захистити біорізноманіття нашої планети є більш відчутною, ніж будь-коли. Майбутнє геноміки, засноване на інноваціях та кероване етичними міркуваннями, має безпрецедентний потенціал для покращення життя людства та природного світу, в якому ми живемо.

-
- ALON, U. (2020) *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC.
- ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W., LIPMAN, D. J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. DOI: 10.1016/S0022-2836(05)80360-2
- CAULFIELD, T., MCGUIRE, A.L. (2012) Direct-to-consumer genetic testing: Perceptions, problems, and policy responses. *Annual Review of Medicine*, 63, 23–33. DOI: 10.1146/annurev-med-062110-123753
- CAVALLI-SFORZA, L. L., FELDMAN, M. W. (2003) The application of molecular genetic approaches to the study of human evolution. *Nature Genetics*, 33, 266–275. DOI: 10.1038/ng1113
- CHURCH, G. M. (2005) The personal genome project. *Molecular Systems Biology*, 1 (1). DOI: 10.1038/msb4100040
- COLLINS, F. S., GREEN, E. D., GUTTMACHER, A. E., GUYER, M. S. (2003) A vision for the future of genomics research. *Nature*, 422 (6934), 835–847. DOI: 10.1038/nature01626
- DELSUC, F., BRINKMANN, H., PHILIPPE, H. (2005) Phylogenomics and the reconstruction of the tree of life. *Nature reviews. Genetics*, 6 (5), 361–375. DOI: 10.1038/nrg1603
- DOUDNA, J. A., CHARPENTIER, E. (2014) The new frontier of genome engineering with CRISPR-Cas9. *Science*, 346 (6213), 1258096. DOI: 10.1126/science.1258096
- ELLEGREN H. (2008) Comparative genomics and the study of evolution by natural selection. *Molecular ecology*, 17(21), 4586–4596. DOI: 10.1111/j.1365-294X.2008.03954.x
- INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM (2001) Initial sequencing and analysis of the human genome. *Nature*, 409, 860–921. DOI: 10.1038/35057062
- KITANO, H. (2002) Systems biology: a brief overview. *Science*, 295 (5560), 1662–1664. DOI: 10.1126/science.1069492
- KNOPPERS, B. M., CHADWICK, R. (2005) Human genetic research: Emerging trends in ethics. *Nature Reviews Genetics*, 6 (1), 75–79. DOI: /10.1038/nrg1505

- KOONIN, E. V. (2012) *The Logic of Chance: The Nature and Origin of Biological Evolution*. FT Press.
- KRAUSE, J., TRAPPE, T. (2016) *A Short History of Humanity: A New History of Old Europe*. Vintage.
- LANDER, E. S. (2011) Initial impact of the sequencing of the human genome. *Nature*, 470 (7333), 187–197. DOI: 10.1038/nature09792
- LIAO, WW., ASRI, M., EBLER, J., DOERR D., HAUKNES, M., PATEN, B. (2023) A draft human pangenome reference. *Nature*, 617, 312–324. DOI: 10.1038/s41586-023-05896-x
- MARDIS, E. R. (2008) Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics*, 9, 387–402. DOI: 10.1146/annurev.genom.9.081307.164359
- McINERNEY, J. O., McNALLY, A., O'CONNELL, M. J. (2017) Why prokaryotes have pangenomes. *Nature Microbiology*, 2(4), 1–5.
- MOUNT, D. W. (2007) *Bioinformatics: Sequence and genome analysis*. Cold Spring Harbor Laboratory Press.
- NURK, S., KOREN, S., RHIE, A., Rautiainen, M., Bzikadze, A., Phillippy, A. M. (2022) The complete sequence of a human genome. *Science*, 376(6588), 44–53. DOI: 10.1126/science.abj6987
- OLEKSYK, T. K., WOLFSBERGER, W. W., SCHUBELKA, K., MANGUL, S., O'BRIEN, S. J. (2022) The Pioneer Advantage: Filling the blank spots on the map of genome diversity in Europe. *GigaScience*, 11, giac081. DOI: 10.1093/gigascience/giac081
- OLEKSYK, T. K., WOLFSBERGER, W. W., WEBER, A. M., SHCHUBELKA, K., OLEKSYK, O. T., LEVCHUK, O., PATRUS, A., LAZAR, N., CASTROMARQUEZ, S. O., HASYNETS, Y., BOLDYZHAR, P., NEYMET, M., URBANOVYCH, A., STAKHOVSKA, V., MALYAR, K., CHERVYAKOVA, S., PODOROHA, O., KOVALCHUK, N., RODRIGUEZFLORES, J. L., ZHOU, W., MEDLEY, S., BATTISTUZZI, F., LIU, R., HOU, Y., CHEN, S., YANG, H., YEAGER, M., DEAN, M., MILLS, R. E., SMOLANKA, V. (2021) Genome diversity in Ukraine. *GigaScience*, 10 (1), giaa159. DOI: 10.1093/gigascience/giaa159
- PAABO, S. (2014) *Neanderthal Man: In Search of Lost Genomes*. Basic Books.
- PEVZNER, P. A. (2000) *Computational molecular biology: An algorithmic approach*. MIT Press.
- HUMAN CELL ATLAS MEETING PARTICIPANTS (2017) The Human Cell Atlas. *eLife*, 6, e27041. DOI: 10.7554/eLife.27041
- REICH, D. (2018) *Who We Are and How We Got Here: Ancient DNA and the New Science of the Human Past*. Pantheon Books.
- ROBERTS, A. (2015) *The Incredible Unlikelihood of Being: Evolution and the Making of Us*. Heron Books.
- RONAGHI, M. (2000) Improved Performance of Pyrosequencing Using Single-Stranded DNA-Binding Protein. *Analytical Biochemistry*, 286 (2), 282–288. DOI: 10.1006/abio.2000.4808
- SHENDURE, J., JI, H. (2008) Next-generation DNA sequencing. *Nature Biotechnology*, 26(10), 1135–1145. DOI: 10.1038/nbt1486
- THE 1000 GENOMES PROJECT CONSORTIUM (2015) A global reference for human genetic variation. *Nature*, 526 (7571), 68–74. DOI: 10.1038/nature15393
- TOPOL, E. J. (2014) Individualized medicine from prewomb to tomb. *Cell*, 157(1), 241–253. DOI: 10.1016/j.cell.2014.02.012
- UK10K CONSORTIUM (2015) The UK10K project identifies rare variants in health and disease. *Nature*, 526(7571), 82–90. DOI: 10.1038/nature14962
- VENTER, J. C., ADAMS, M. D., MYERS, E. W., LI, P. W., MURAL, R. J.,... & ZHU, X. (2001) The sequence of the human genome. *Science*, 291(5507), 1304–1351. DOI: 10.1126/science.1058040
- WATSON, J. D., CRICK, F. H. C. (1953) Molecular Structure of Nucleic Acids. *Nature*, 171, 737–738.
- WILLERSLEV, E., DAVISON, J. (2021) *Origins: A Genetic History of the Americas*. St. Martin's Press.