## Od pytania do wyniku – jak planować eksperymenty genomiczne, by dobrać właściwą technologię i uzyskać biologicznie istotne dane

Rafał Kazimierz Wóycicki, dr inż.



16 czerwca 2025 r. IZiBB, WB, UJ

# W skrócie moje doświadczenie w genetyce, biologii molekularnej, genomice i bioinformatyce

#### 1.Biologia molekularna i genetyka klasyczna (10 lat do 2008 roku):

- Generowanie markerów molekularnych genów rozwoju kwiatów (RAPD -> SCAR -> klonowanie pozycyjne).

### Promotor Prof. Zbigniew Przybecki - do 2011 roku



## Składanie *de-novo* genomu jądrowego ogórka (367 Mbp) oraz ganomika porównawcza między dwiema odmianami ogórka.

Process	Line B10	Line 9930	Envinormental conditions
Photosynthesis	+	-	Temper climate of Northern Europe: cold, low light intensity
Sugar metabolism	+	-	
Respiration	+	-	
Reg.of gene expreession	+	-	
Chlorophyll degradation	+	1	
Nitrogen binding as amonium ions	+	,	Continouos and higher emision of CO2 in Europe than in South-Eastern Asia, when counting from beginning of industrial era to 80's of XX century - lowered abillity for binding nitrogen ions
Oxidative stress resistance	-	+	Subtropical climate of South-Eastern China: high
High temperature resistance	-	+	sesonal intensity of sun light including UV-B radiation, together with high temperature



Bureau lab 2011-2012 - obróbka informatyczna sygnałów z kamer (RGB, IR, NIR, Fluoro, 3D) do automatycznego fenotypowania roślin w celu szybkiej analizy korelacji genotypów z fenotypami.



Kovalchuks & Metz labs 2012 - 2014 (epigenetyka roślin, chorób rzadkich i raka u ludzi, neuro-epigenetyka zwierząt) - microarrays, WGS, RNA-seq (miRNAs), Bisulfide-Seq



Nowacki lab 2014-2016 (Pośredniczone przez RNA dziedziczenie epigenetyczne u *Paramecium tetraurelia*) - WGBS, miRNA, de-novo PacBio, single-cell genomics



O UNIVERSITÄT BERN

Phillip Morris International 2016-2018 (dynamika ewolucji roślin allopoliploidalnych na przykładzie rodzaju Nicotiana) - składanie de-novo genomów genomów 5 gatunków (Illumina/PacBio), genomika porównawcza, składanie transkryptomów, profilowanie ekspresji



Libault lab 2020 - pionierskie analizy scRNA-seq & scATAC-seq u roślin od 100 Mpz do 18 Gpz (adnotacja typów komórek, wykrywanie specificznych różnic transkrypcyjnych w komórkach włośnikowych u soji (symbioza z bakteriami nitryfikacyjnymi)



Genentech/Roche 2021 - 2023 - kurator danych genomowych oraz analizy scRNA-seq w poszukiwaniu leków chorób genetycznych u ludzi



# W skrócie moje doświadczenie w genetyce, biologii molekularnej, genomice i bioinformatyce

#### 2. Genomika i bioinformatyka (16 lat od 2008 roku):

- Składanie, adnotacja i analiza genomów i transkryptomów de-novo gatunków niemodelowych.
- Profilowanie ekspresji genów i niekodujących małych RNA (RNA-seg, mikromacierze).
- Profilowanie metylacji (bisulfide, mikromacierze).
- Transkryptomika pojedynczych komórek (scRNA) zarówno u roślin ważnych gatunków uprawnych: kukurydza, soja, pszenica, jak i u ludzi (astma).
  - Analiza ewolucji genomów i transkryptomów gatunków diploidalnych i allotetraploidalych.

Organizmy: Rośliny, Zwierzęta, Ludzie (w tym komórki rakowe)

Technologie: Sanger, 454, PacBio, Illumina, Oxford Nanopore

Narzędzia: Perl, BASH, R, Python, BIOCONDA i inne.

Zainteresowania: Szybka adaptacja do stresów poprzez zmiany regulacji ekspresji genów i dziedziczenie pamięci stresu.

### Plan prezentacji

- Zmiana Paradygmatu w Genomice: Od wariantów punktowych do złożoności strukturalnej i regulacyjnej.
- 2. **Definiowanie Problemu Badawczego:** Kluczowe pytania determinujące strategię.
- 3. **Przegląd Technologii Sekwencjonowania (2025):** Ocena platform pod kątem konkretnych zastosowań.
  - o Poziom informacji genetycznej: genom, transkryptom, proteom, metabolom,
  - Poziom odkryć genomu: mapowanie vs de-novo, SNP/InDel vs SV/GChR, krótkie odczyty vs długie odczyty, exomy vs sekwencje niekodujące vs sekwencje powtarzalne, informacja liniowa vs 3D, consensus vs tyle sekwencji ile wynosi ploidalność
  - Poziom odkryć transkryptomu: mapowanie vs de-novo (znane transkrypty vs nowe transkrypty),
    krótkie odczyty vs długie odczyty (mix transkryptów vs rozdział na izoformy vs IncRNA)
  - Specyficzność komórkowa czy ogólna odpowiedź: bulk czy pojedyncza komórka
  - Epigenetyka: modyfikacje nukleotydów (natywne vs amplifikowane DNA/RNA), histonów, dostępność chromatyny
  - Metagenomika: klasyfikowanie mikroorganizmów vs bioreaktory
- 4. Fundamentalne Zasady Projektowania Eksperymentów.
- 5. **Podsumowanie i Dyskusja.**

### Ewolucja Postrzegania Zmienności Genetycznej

- Paradygmat Historyczny (era wczesnego NGS):
- o **Fokus:** Warianty punktowe (SNPs) i małe INDELs jako główne źródło zmienności.
- **Metodologia:** Mapowanie krótkich odczytów (np. Illumina) do genomu referencyjnego.
- Fundamentalne Ograniczenie: Detekcja wariantów głównie w regionach unikalnych (przede wszystkim w eksonach).
  Sekwencje powtarzalne, stanowiące znaczną część genomu i kluczowe dla regulacji, były w dużej mierze ignorowane lub niemożliwe do jednoznacznej analizy.
- Paradygmat Aktualny (era długich odczytów):
- Fokus: Warianty Strukturalne (SVs) i rearanżacje chromosomowe jako główne motory ewolucji i szybkiej adaptacji.
  Genomy tego samego gatunku nie są tylko wariantami tej samej sekwencji mogą mieć fundamentalnie różną architekturę.
- Kluczowa rola regionów niekodujących: Introny i sekwencje powtarzalne są centrami regulacji ekspresji genów.
  Zmienność w tych regionach ma krytyczne znaczenie funkcjonalne.
- Implikacja: Poszukiwanie przyczyn zmienności fenotypowej wyłącznie w regionach kodujących jest podejściem niewystarczającym i może prowadzić do pominięcia kluczowych mechanizmów biologicznych.

### Hierarchia Pytań Kształtujących Projekt

Zanim wybierzemy technologię, musimy precyzyjnie zdefiniować cel:

- 1. Poziom Analizy: Gdzie znajduje się poszukiwana informacja?
- o **Genom (DNA):** Analiza stabilnej, dziedzicznej podstawy cech, wariantów strukturalnych, ewolucji.
- Transkryptom (RNA): Badanie dynamicznej odpowiedzi komórki, aktywności transkrypcyjnej poszczególnych izoform, IncRNA.
- **Epigenom:** Analiza mechanizmów regulacyjnych modyfikacji DNA, RNA i histonów.

#### 2. Skala Zmienności:

- **Warianty Punktowe (SNPs/INDELs):** Gdy hipoteza dotyczy specyficznych, drobnych zmian w znanych regionach.
- Warianty Strukturalne (SVs) / Rearanżacje: Gdy interesuje nas architektura genomu, adaptacja, złożone choroby genetyczne.

#### 3. Rozdzielczość Komórkowa:

- o **Bulk:** Uśredniona odpowiedź całej tkanki lub populacji komórek.
- Single-Cell: Analiza heterogeniczności odpowiedzi, identyfikacja rzadkich subpopulacji, śledzenie trajektorii rozwojowych (np. w onkologii, neurologii).

### Cel: Identyfikacja wariantów typu SNP/INDEL

- **Pytanie naukowe:** "Czy w genomach pacjentów z daną chorobą występują nieznane dotąd warianty punktowe w porównaniu do populacji referencyjnej, np. w badaniach typu GWAS?"
- Standardowa Technologia: Resekwencjonowanie Krótkimi Odczytami (Illumina)
  - Zasada: Generowanie milionów krótkich (150-300 pz), wysoce dokładnych odczytów i ich mapowanie do genomu referencyjnego.

#### Zalety:

- Bardzo wysoka dokładność na poziomie pojedynczej zasady (Phred score Q>30, błąd 1 na 1000).
- Ekstremalnie niski koszt w przeliczeniu na zsekwencjonowaną bazę, co pozwala na analizę dużych kohort.
- Ugruntowane, standardowe i szybkie protokoły laboratoryjne i bioinformatyczne.

#### Ograniczenia:

- Niska efektywność w analizie regionów powtarzalnych i bogatych w GC.
- Niemożność wiarygodnej detekcji dużych wariantów strukturalnych (SVs).
- Problemy z fazowaniem haplotypów.

### Cel: Pełna Architektura Genomu i Warianty Strukturalne

- **Pytanie naukowe:** "Jaka jest kompletna sekwencja genomu niebadanego dotąd gatunku?" LUB "Czy u pacjenta występuje złożona rearanżacja chromosomowa, niewykrywalna standardowymi metodami?"
- Rozwiązanie: Sekwencjonowanie Długimi Odczytami
  - o PacBio HiFi:
    - Długość odczytu: 15-25 kbp; Dokładność: Q>30 (99.9%).
    - Zastosowanie: Złoty standard dla składania genomów de novo do poziomu T2T, precyzyjna detekcja SVs, fazowanie haplotypów.
  - Oxford Nanopore (ONT):
    - Długość odczytu: **20-50 kbp, z potencjałem >1 Mbp**; Dokładność: Niższa surowa dokładność (Q10-Q20).
    - Zastosowanie: Składanie najbardziej złożonych regionów powtarzalnych (np. centromerów), analiza ultraszybka i w terenie (MinION).
- Technologia Wspierająca: Hi-C
  - Cel: Mapowanie przestrzennych interakcji w chromatynie.
  - Zastosowanie w składaniu genomu: Umożliwia łączenie zsekwencjonowanych kontigów w scaffoldy na poziomie chromosomów.
  - Zastosowanie funkcjonalne: Identyfikacja pętli chromatynowych, których dynamika może korelować z odpowiedzią na stres i procesami adaptacyjnymi.

### Cel: Ilościowa Analiza Różnicowej Ekspresji Genów (DGE)

- **Pytanie naukowe:** "Które geny ulegają statystycznie istotnej zmianie ekspresji w komórkach po stymulacji w porównaniu do kontroli?"
- Standardowa Technologia: RNA-Seq Krótkimi Odczytami (Illumina)
- o Zasada: Sekwencjonowanie fragmentów cDNA i zliczanie odczytów mapowanych do poszczególnych genów.

#### Zalety:

- Niski koszt, wysoka przepustowość, co pozwala na analizę wielu próbek z dużą liczbą powtórzeń.
- Bardzo duża dokładność ilościowa.
- Ugruntowane i wiarygodne narzędzia do analizy statystycznej (np. DESeq2, edgeR).

#### Ograniczenia:

- Wynik jest sumaryczny dla wszystkich izoform danego genu. Nie dostarcza informacji o strukturze poszczególnych wariantów splicingowych.
- Może być nieefektywna w składaniu i kwantyfikacji nowych transkryptów lub lncRNA.

### Cel: Precyzyjna Analiza Ekspresji na Poziomie Izoform

- **Problem:** Mówienie o "ekspresji genu" jest uproszczeniem. Różne izoformy tego samego genu, powstające w wyniku alternatywnego splicingu, mogą mieć odmienne, a nawet przeciwstawne funkcje.
- Rozwiązanie: Sekwencjonowanie Długimi Odczytami
  - PacBio Iso-Seq:
    - Zasada: Sekwencjonowanie pełnej długości transkryptów cDNA.
    - Zaleta: Bardzo wysoka dokładność pozwala na bezbłędną identyfikację i kwantyfikację izoform. Wysoka przepustowość (Revio).
    - Wada: Wymaga etapu odwrotnej transkrypcji (RT-PCR), co eliminuje informację o modyfikacjach epigenetycznych na RNA.
  - Oxford Nanopore Direct RNA Sequencing:
    - Zasada: Bezpośrednie sekwencjonowanie natywnych cząsteczek RNA.
    - Unikalna zaleta: Zachowuje i pozwala na identyfikację modyfikacji epigenetycznych na RNA (np. m6A).
    - Wada: Niższa dokładność odczytu (Q-value) wymaga większego pokrycia lub strategii korekcji błędów.

### Analiza Warstw Regulacyjnych i Systemów Złożonych

#### • Epigenetyka – Detekcja Modyfikacji:

- o DNA: Bisulfite-Seq (WGBS) na Illumina to historyczny złoty standard, ale wykrywa tylko 5-mC i degraduje DNA.
- DNA: Sekwencjonowanie natywne (PacBio, ONT) pozwala na jednoczesną detekcję wielu typów modyfikacji (5-mC, 5-hmC, 6-mA itd.) bez amplifikacji PCR.
- o RNA: Jedyną powszechnie dostępną metodą detekcji modyfikacji jest ONT Direct RNA Sequencing.
- Chromatyna: ATAC-Seq (mapowanie dostępności chromatyny), ChIP-Seq (mapowanie miejsc wiązania białek i modyfikacji histonów).

#### • Single-Cell Genomics:

- o scRNA-Seq (Illumina): Standardem jest profilowanie ekspresji oparte na znacznikach na końcu 3' lub 5'.
- scRNA-Seq (PacBio/ONT): Pozwala na analizę pełnej długości izoform w pojedynczych komórkach.
  Wymaga jednak wyższych kosztów i przepustowości (np. platforma Revio, PromethION).

#### Metagenomika:

- Podejście przestarzałe: Analiza markerów 16S rRNA.
- Podejście współczesne: Shotgun Metagenomics długimi odczytami. Pozwala na składanie kompletnych genomów mikroorganizmów z złożonej mieszaniny lub pełnych genów 16s (~1500 bp).

### Od Teorii do Praktyki: Krytyczne Elementy Projektu

#### • 1. Jakość Materiału Wyjściowego:

- Warunek najistotniejszy: Niska jakość materiału uniemożliwia uzyskanie wiarygodnych danych.
- DNA (dla długich odczytów): Kluczowa jest wysoka masa cząsteczkowa (HMW DNA). Należy unikać fragmentacji mechanicznej, nadmiernego pipetowania i wielokrotnego zamrażania/rozmrażania.
- RNA: Kluczowa jest integralność (RIN > 8). Należy stosować inhibitory RNaz i standaryzować protokół izolacji.

#### • 2. Warunki Eksperymentalne i Przechowywanie:

- o Standaryzacja warunków hodowli/wzrostu jest niezbędna do minimalizacji zmienności niebiologicznej.
- Pobieranie i przechowywanie wszystkich próbek w identyczny, szybki sposób minimalizuje zmiany transkrypcyjne ex vivo.

#### • 3. Problem Modeli Komórkowych:

- Należy zachować szczególną ostrożność przy ekstrapolacji wyników z linii komórkowych na cały organizm.
- Linie komórkowe akumulują mutacje i rearanżacje chromosomowe z każdym pasażem nie są stabilne genetycznie. Ich profil epigenetyczny i transkrypcyjny może znacząco odbiegać od komórek pierwotnych in vivo.

### Równanie sukcesu: Budżet, Powtórzenia i Moc Statystyczna

#### • 1. Budżet i Dobór Platformy:

- Wybór technologii jest kompromisem między celem naukowym a dostępnymi zasobami.
- Przykład: Analiza ekspresji izoform na małą skalę może być wykonana na MinION (ONT), ale badanie typu single-cell full-length wymaga dostępu do platform o wysokiej przepustowości, jak PromethION (ONT) czy Revio (PacBio).

#### • 2. Powtórzenia Biologiczne:

- Są absolutnie kluczowe dla analiz statystycznych. Porównywanie pojedynczych próbek jest niedopuszczalne.
- o Minimum to trzy niezależne powtórzenia biologiczne na każdą grupę, aby oszacować wariancję.

#### • 3. Moc Statystyczna:

Należy mieć świadomość, że przy małej liczbie powtórzeń (np. n=3) i dużej liczbie testów (np. 20 000 genów), wykrywalne będą jedynie zmiany o dużej amplitudzie (np. > dwukrotne). Wykrycie subtelniejszych zmian wymaga zwiększenia liczby powtórzeń.

#### • 4. Głębokość Sekwencjonowania:

 W przypadku analizy izoform, wymagana liczba odczytów jest znacznie wyższa niż w analizie na poziomie genów, ponieważ sygnał jest rozproszony na wiele (czasem 5-10) transkryptów na gen.

### Planowanie: Fundament Wiarygodnych Wyników

#### Zasada #1: Jasno Zdefiniowane PYTANIE BADAWCZE!

- "Co konkretnie chcę zbadać? Jaka jest moja hipoteza?"
- Precyzyjne pytanie determinuje: wybór technologii, projekt eksperymentu, metody analizy.

#### Zasada #2: Odpowiednie KONTROLE!

- Bez nich wyniki są bezwartościowe! Kontrole pozwalają odróżnić efekt biologiczny od artefaktów technicznych.
- Rodzaje Kontroli (przykłady):
  - Negatywne: Próbki nie poddane działaniu czynnika (np. "mock treatment", placebo).
  - Pozytywne: Próbki, dla których oczekujemy znanego efektu (jeśli to możliwe).
  - Biologiczne: Niezależne próbki biologiczne (różne osobniki, różne linie komórkowe).
  - o *Techniczne:* Ta sama próbka biologiczna przetwarzana wielokrotnie (do oceny zmienności technicznej).
  - Środowiskowe/Czasowe: Kontrola warunków, pobieranie próbek w tych samych warunkach/punktach czasowych.

#### Zasada #3: POWTÓRZENIA BIOLOGICZNE!

- Dlaczego? Aby uchwycić naturalną zmienność biologiczną i odróżnić ją od zmienności technicznej. Pozwalają na wiarygodną analizę statystyczną.
- Ile?
  - o To zależy od: oczekiwanej wielkości efektu, zmienności w systemie, wybranej technologii, pożądanej mocy statystycznej.
  - RNA-Seq, Epigenetyka: Absolutne minimum to 3 powtórzenia biologiczne na grupę, ale zalecane jest 4-6 lub więcej, zwłaszcza dla subtelnych efektów lub dużej zmienności.
  - WGS (badania populacyjne): Zależy od projektu, może wymagać dziesiątek lub setek próbek.
- Jak? Muszą to być *prawdziwie niezależne* powtórzenia (np. różne zwierzęta, różne partie komórek hodowane niezależnie).

### Diabeł Tkwi w Szczegółach: Próbki i "Batch Effects"

#### "Garbage In, Garbage Out" (GIGO) – Jakość Materiału Wyjściowego jest KRYTYCZNA!

- RNA:
  - Ocena jakości: RIN (RNA Integrity Number) idealnie > 8 (dla większości aplikacji).
  - Unikać degradacji (szybkie przetwarzanie, inhibitory RNaz, odpowiednie przechowywanie).
  - Czystość: Brak zanieczyszczeń DNA, białkami, inhibitorami PCR.
- DNA:
  - Ocena jakości: Wysoka masa cząsteczkowa (brak degradacji), czystość (spektrofotometria A260/280, A260/230).
  - Unikać wielokrotnego zamrażania/rozmrażania.

#### Efekty Odrzutowe (Batch Effects) – Podstępny Wróg!

- **Co to?** Systematyczne różnice techniczne między próbkami, które są przetwarzane w różnych "partiach" (np. w różne dni, przez różne osoby, z różnymi partiami odczynników, na różnych urządzeniach).
- **Dlaczego są groźne?** Mogą maskować prawdziwe efekty biologiczne lub generować fałszywie pozytywne wyniki. Są głównym źródłem braku powtarzalności badań!
- Jak Minimalizować "Batch Effects"?
  - PLANOWANIE! Najważniejszy etap.
  - **Randomizacja:** Całkowicie losowe przypisanie próbek do poszczególnych partii przetwarzania (przygotowanie bibliotek, sekwencjonowanie).
  - **Równoważenie Grup:** Jeśli nie można przetworzyć wszystkich próbek naraz, upewnij się, że każda partia zawiera próbki ze wszystkich porównywanych grup (np. kontrola i traktowane).
  - Standaryzacja Protokołów: Używaj tych samych protokołów, odczynników (z tej samej partii, jeśli to możliwe) i sprzętu dla wszystkich próbek.
  - Dokładne Notatki: Zapisuj WSZYSTKO (daty, operatorów, numery partii odczynników, wszelkie odchylenia od protokołu). Te metadane są kluczowe dla późniejszej analizy i korekcji "batch effects".

### Analiza mocy statystycznej

Procedura statystyczna pozwalająca oszacować prawdopodobieństwo wykrycia rzeczywistego efektu biologicznego (o określonej wielkości) przy danej liczbie powtórzeń i poziomie zmienności. Innymi słowy: "Czy mój eksperyment ma wystarczającą 'siłę', aby zobaczyć to, czego szukam?"

- Kiedy ją przeprowadzić? PRZED rozpoczęciem eksperymentu, na etapie planowania!
- Co jest potrzebne do analizy mocy?
  - Oczekiwana wielkość efektu (np. 2-krotna zmiana ekspresji genu).
  - Szacowana zmienność danych (z badań pilotażowych lub literatury).
  - Pożądany poziom istotności statystycznej (α, np. 0.05).
  - Pożądana moc (1-β, np. 0.8, czyli 80% szansy na wykrycie efektu, jeśli istnieje).
- Wynik: Określenie minimalnej liczby powtórzeń biologicznych potrzebnych do osiągnięcia pożądanej mocy.
- Dlaczego to ważne? Pomaga uniknąć marnowania czasu i zasobów na eksperymenty, które są z góry skazane na niepowodzenie (zbyt mała moc).
- **Narzędzia:** Istnieją kalkulatory online i pakiety w R (np. pwr, RNASeqPower). Konsultacja z bioinformatykiem jest tu bardzo pomocna.
- Zasoby własne: np. <a href="https://rafalwoycicki.github.io/power\_calculator/index.html">https://rafalwoycicki.github.io/power\_calculator/index.html</a> dla analiz transkryptomicznych

### Głębokość Sekwencjonowania

#### Głębokość Sekwencjonowania (Sequencing Depth):

- Od czego zależy optymalna głębokość?
  - **Typ Technologii:** RNA-Seq do analizy ekspresji vs. WGS do składania genomu.
  - Rozmiar Genomu/Transkryptomu: Większe genomy/transkryptomy (geny a izoformy (geny \*10))
    wymagają większej głębokości.

#### Cel Badania:

- Detekcja rzadkich wariantów (WGS) lub transkryptów (RNA-Seq): Wymaga większej głębokości.
- Profilowanie ekspresji wysoko wyrażanych genów: Może wystarczyć mniejsza głębokość.
- Analiza splicingu, odkrywanie nowych izoform: Wymaga większej głębokości w RNA-Seq.
- Złożoność Próbki: Metagenomy wymagają bardzo dużej głębokości.

### Schemat procesu: od hipotezy do interpretacji

PYTANIE NAUKOWE / HIPOTEZA

1

 DEFINIOWANIE CELU (Poziom analizy: genom/transkryptom/epigenom; Skala zmiany: SNP/SV; Rozdzielczość: bulk/SC)

3. OCENA I WYBÓR TECHNOLOGII (np. Illumina vs. PacBio vs. ONT; WGS vs. Iso-Seq vs. ATAC-Seq)

1

- 4. RYGORYSTYCZNE PROJEKTOWANIE EKSPERYMENTU
  - Jakość HMW DNA / RNA (RIN>8)
  - Liczba powtórzeń biologicznych (analiza mocy)
  - Standaryzacja i randomizacja (kontrola batch effects)
  - Ocena budżetu i przepustowości

1

ANALIZA BIOINFORMATYCZNA I INTERPRETACJA BIOLOGICZNA

**Główne Przesłanie:** Sukces w badaniach genomicznych zależy od rygorystycznego procesu projektowego, w którym technologia jest świadomie dobranym narzędziem do weryfikacji precyzyjnie sformułowanej hipotezy.

### Dziękuję! Pytania? Dyskusja?

Chętnie odpowiem na pytania dotyczące:

- prezentowanych technologii,
- planowania eksperymentów,
- możliwości zastosowania genomiki w Państwa konkretnych projektach badawczych.

Przykładowe koszty sekwencjonowania w odniesieniu do rynku USA i cen w USD, uśredniony stan na rok 2025 wg Gemini & AI:

https://rafalwoycicki.github.io/genomics/costs.html