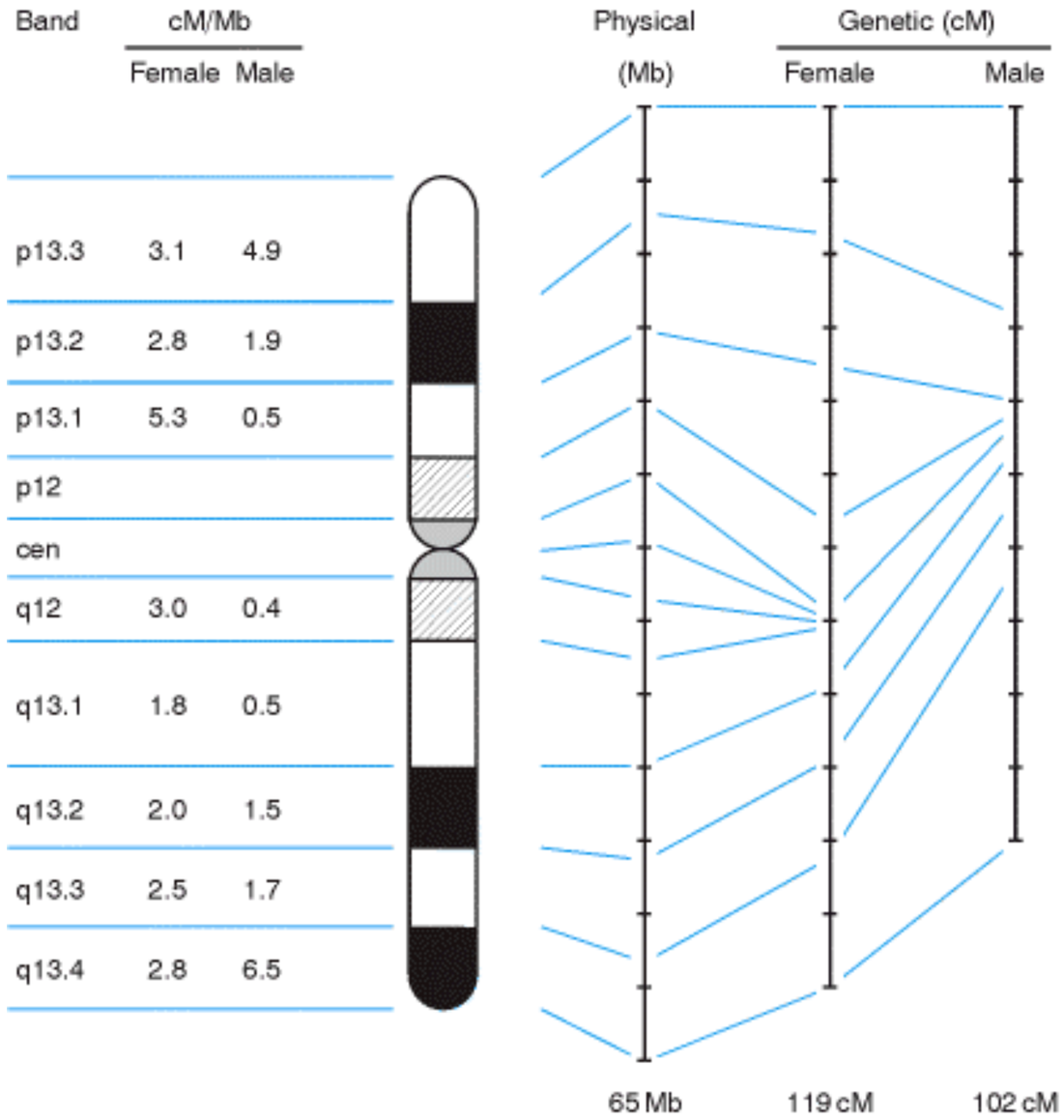
Klasyczne metody genetyki

Mapy genetyczne człowieka i innych organizmów

- Całkowita mapa mężczyzny = 2851cM
- Całkowita mapa kobiety = 4296cM (wyłączając X)
- Dla 3000Mb genomu autosomalnego
  - 1 cM u mężczyzny ≈ 1,05 Mb
  - 1 cM u kobiety ≈ 0,88Mb
- 1 cM u *Drosophila* ≈ 0,5 Mb
- 1cM u drożdży ≈ 3 kb



Jak mapować geny człowieka

- Bezpośrednie zliczanie klas potomstwa - niepraktyczne

Wiarygodność (*likelihood*)

- Wiarygodność: prawdopodobieństwo uzyskania zaobserwowanych danych przy założeniach modelu i jego określonych parametrach

Wiarygodność (*likelihood*)

- W rodowodzie w pełni informatywnym
  - dane: R=liczba rekombinantów; NR=liczba genotypów rodzicielskich
  - parametr: częstość (prawdopodobieństwo) rekombinacji θ
- Hipoteza zerowa – brak sprzężenia (θ=0,5)
- Stosunek wiarygodności dla danej wartości θ: L(θ)/L(θ=0,5)
- lod score (Z) = *logarithm of odds* – logarytm dziesiętny stosunku wiarygodności

Proste przykłady obliczeń *lod*

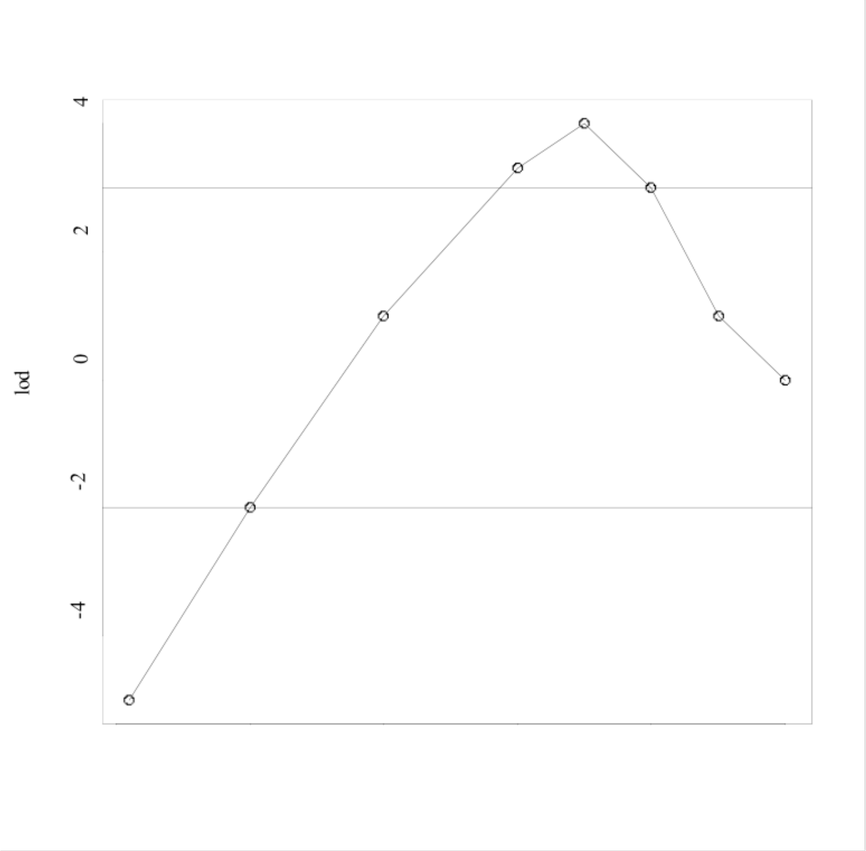
Dla danego rodowodu (i), *lod* dla danego θ wynosi:

$$Z_i(\theta) = \log_{10} \frac{L(rodowód / \theta)}{L(rodowód / \theta = 0{,}5)}$$

Dla danej wartości θ, sumuje się *lod-score* z różnych rodowodów (F):

$$Z(\theta) = \sum_{i=1}^{F} Z_i(\theta)$$

Analiza dwupunktowa



znaczące
(Z>3,
Z>2 dla sprzężonych z płcią)

wykluczone

Tabela

θ =	0.01,	0.10,	0.20,	0.30,	0.35,	0.40,	0.45,	0.50
lod=	-5.0,	-2.0,	1.0,	3.3,	4.0,	3.0,	1.0,	0.0

Markery w analizie sprzężeń u człowieka

- Sprzężenie dwóch genów o obserwowalnym fenotypie – praktycznie niespotykane
  - wyjątek – zespół paznokciowo-rzepkowy (NPS – *Nail Patella Syndrome*) i grupy krwi AB0
  - *Loci* w obrębie kompleksów MHC
- Markery molekularne
  - PCR, RFLP

Czy genetyka klasyczna ma dziś znaczenie?

- Wciąż aktualne metody:
  - Izolacja i charakterystyka mutantów
  - Test komplementacji
  - Interakcje genetyczne! – jedna z podstaw biologii systemów
  - Konstrukcje organizmów (głównie mikroorganizmy) przez odpowiednio dobrane krzyżówki
  - Dziedziczenie mendlowskie w medycynie – poradnictwo genetyczne. Metody probabilistyczne

Podstawy genetyki II

Metody badawcze i strategie genetyki i genomiki.

Podstawy genetyki II

Metody badawcze i strategie genetyki i genomiki.

Czym jest inżynieria genetyczna?

- Ang. *recombinant DNA* – manipulacje DNA *in vitro*
  - izolacja i amplifikacja DNA i cDNA
  - mapowanie i sekwencjonowanie DNA
  - tworzenie nowych cząsteczek DNA
    - przez rekombinację cząsteczek naturalnych
    - przez syntezę *de novo*
    - redagowanie genomów
  - wprowadzanie konstruktów DNA do komórek i organizmów
  - modyfikacje syntezy białek
  - ekspresja heterologiczna
  - bioinformatyka

A co nie jest inżynierią genetyczną?

- Inżynieria embrionalna (np. klonowanie)
- Tworzenie nowych form organizmów przez selekcję

Zastosowania

- Badania podstawowe
- Biotechnologia

Granica między badaniami podstawowymi a stosowanymi jest płynna, stosowane techniki są podobne, różnice dotyczą głównie skali.

Podstawowe techniki

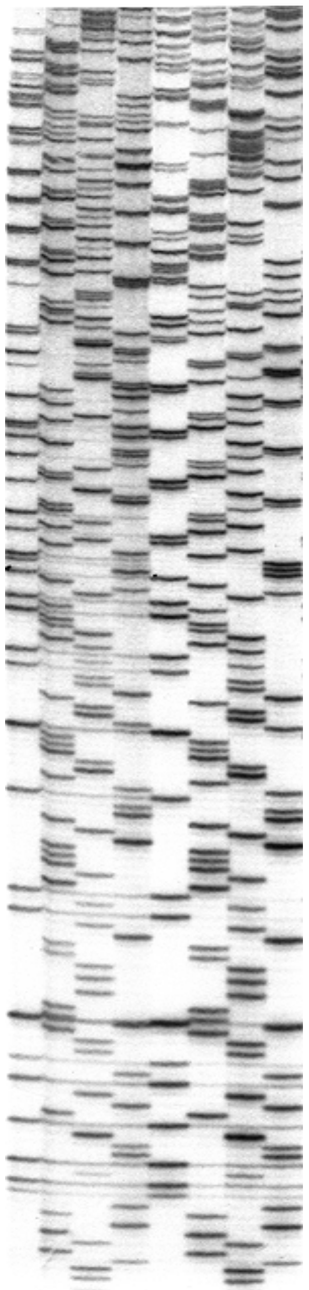
- Izolacja DNA lub RNA
  - cDNA – izolacja RNA i przepisanie na DNA
  - Chemiczna synteza DNA *de novo*
- PCR
- Klonowanie DNA
- Mutageneza losowa i ukierunkowana
  - w tym wprowadzanie modyfikacji do genomu
- Wykrywanie DNA, RNA i białek
- Sekwencjonowanie

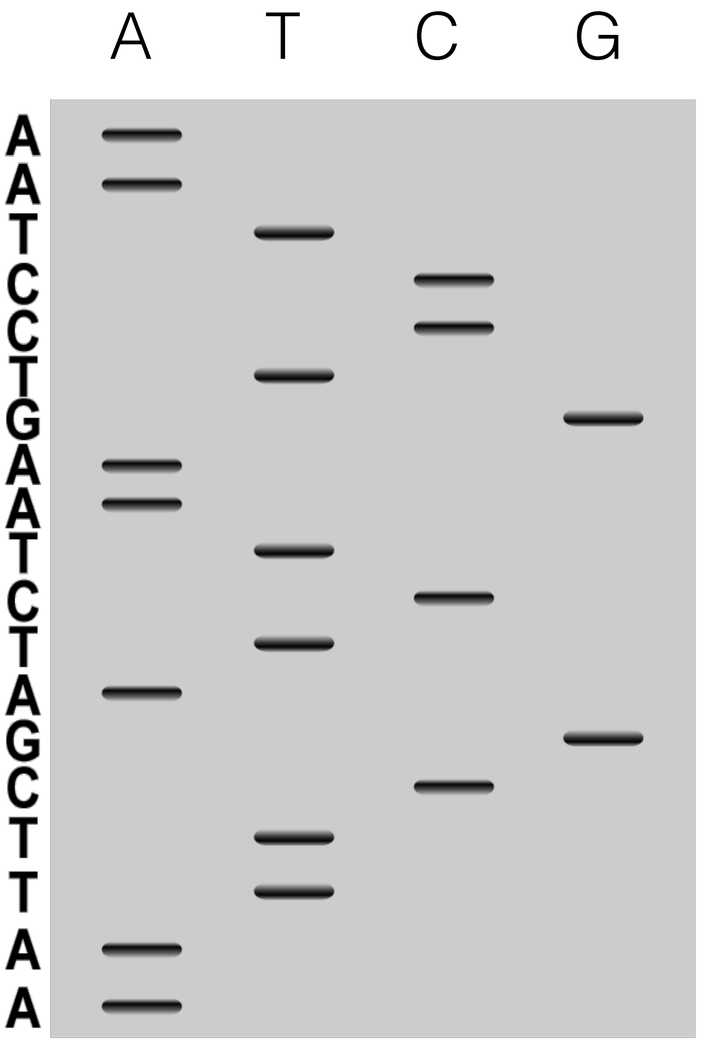
Lektura

- Allison “Podstawy biologii molekularnej”, rozdział 8 i 9

Tradycyjny odczyt sekwencji

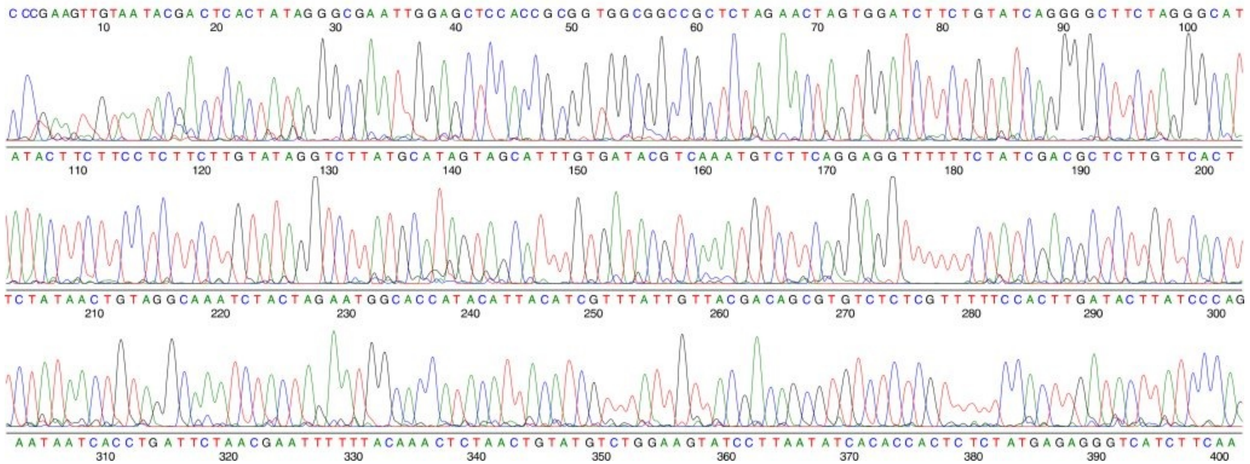
- Metoda Sangera
- Synteza DNA w obecności analogów nukleotydów (forma dideoksy) terminujących reakcję na określonym nukleotydzie
- Znakowanie radioaktywne, osobne reakcje





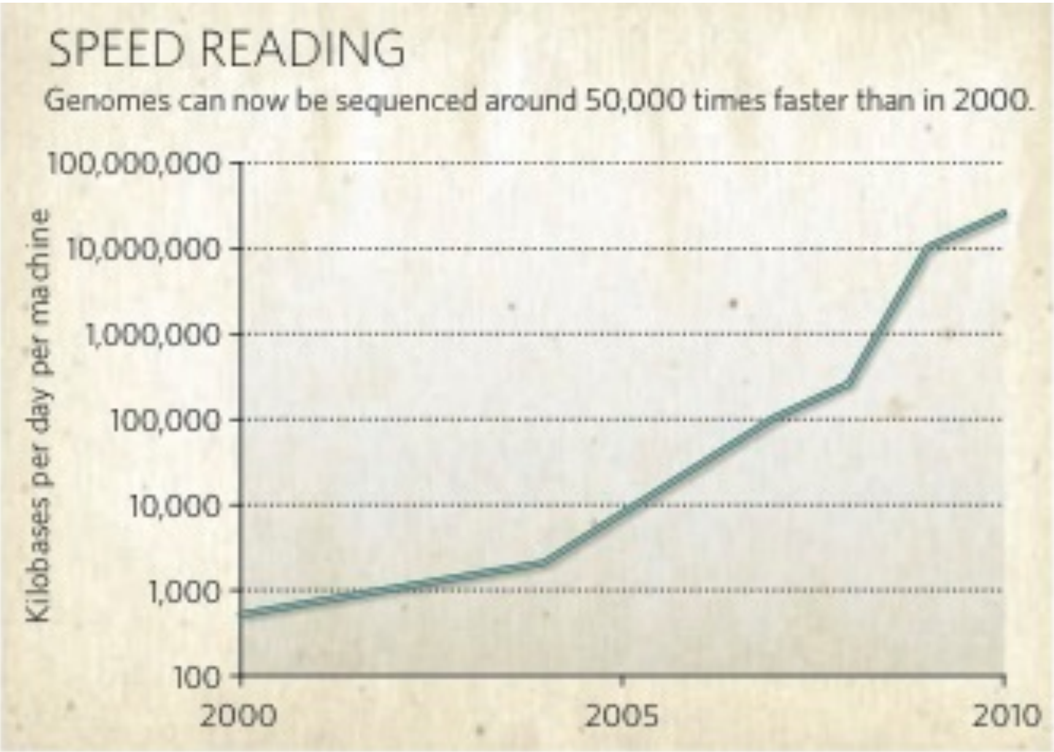
Sekwencjonowanie automatyczne

- Dideoksynukleotydy znakowane fluorescencyjnie (4 kolory)
- Elektroforeza kapilarna



Sekwencjonowanie - postęp techniczny

- Koszt sekwencjonowania między 1999 a 2009 obniżył sie 14 000 razy
- Prędkość odczytu sekwencji między 2000 a 2010 r. wzrosła 50 000 razy
- Cel: sekwencja genomu jednej osoby za 1000$ osiągnięty
- Im więcej znamy sekwencji, tym łatwiej poznajemy kolejne



© 2010 Macmillan Publishers Limited. All rights reserved

Sekwencjonowanie wysokoprzepustowe

- Tzw. *deep sequencing*, sekwencjonowanie nowej generacji (NGS)
- Generowanie w jednym przebiegu milionów niezależnych odczytów
- Pojedyncze odczyty krótkie (25-400 bp)
- Zastosowania
  - sekwencjonowanie nowych genomów
  - resekwencjonowanie
    - np. analiza zmienności
  - badanie ekspresji przez sekwencjonowanie cDNA

Sekwencjonowanie

- Głównym wyzwaniem w sekwencjonowaniu nie jest sam odczyt sekwencji
- Odczytywane fragmenty są krótkie
  - do ~700-800 nt (sekw. tradycyjne Sangera)
  - 50-400 nt (NGS)
  - Istnieją metody (PacBio, Nanopore) dające długie odczyty (>10 000 nt), ale z licznymi błędami
- Wyzwaniem jest złożenie długiej sekwencji z tych krótkich fragmentów

Genomika

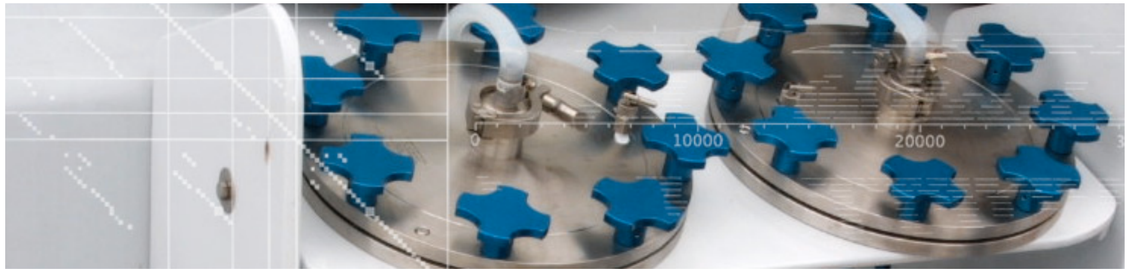
- Genomika jest dziedziną zajmującą się badaniem całych genomów (kompletu informacji genetycznej) różnych organizmów
- Techniki biologii molekularnej + robotyka + informatyka
- Sekwencjonowanie i charakteryzowanie genomów
- Badanie funkcji zawartych w nich genów - genomika funkcjonalna

Metagenomika

- Izolacja DNA ze środowiska i sekwencjonowanie
- Jedyny sposób badania mikroorganizmów, które nie dają się hodować

Metagenomika

Analiza sekwencji całości DNA
wyizolowanego ze zbiorowiska organizmów



Odkrycia dzięki sekwencjonowaniu

- Tajemnicza UCYN-A
  - Sinica (cyjanobakteria)
  - Niewielki genom (1,4mln par zasad, 1200 genów)
  - Brak zdolności fotosyntezy, cyklu Krebsa, syntezy niektórych aminokwasów
  - Zdolność asymilacji azotu
  - Symbioza (gospodarz - *Haptophyta, Primnesiophyta* - jednokomórkowe glony)
  - Jeden z głównych czynników asymilacji azotu w środowisku morskim

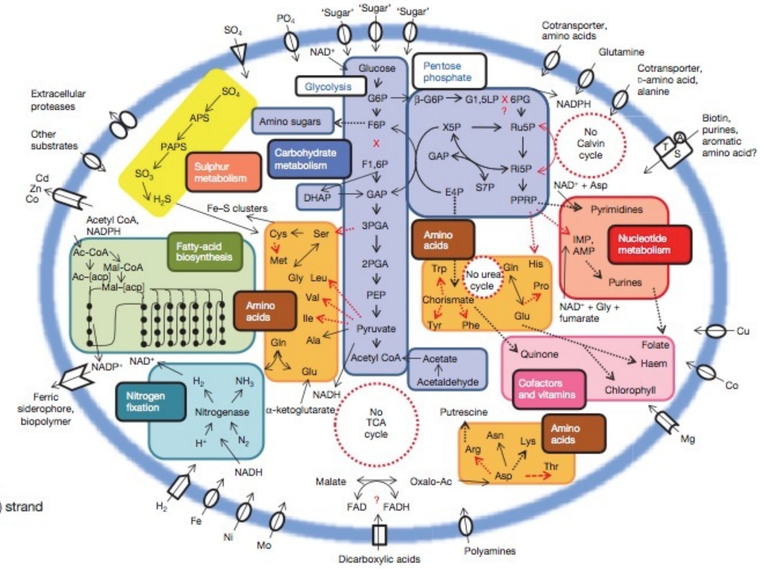
Candidatus Atelocyanobacterium thalassa

doi:10.1038/nature08786 nature

LETTERS

Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium

H. James Tripp[1], Shellie R. Bench[1], Kendra A. Turk[1], Rachel A. Foster[1], Brian A. Desany[2], Faheem Niazi[2], Jason P. Affourtit[2] & Jonathan P. Zehr[1]



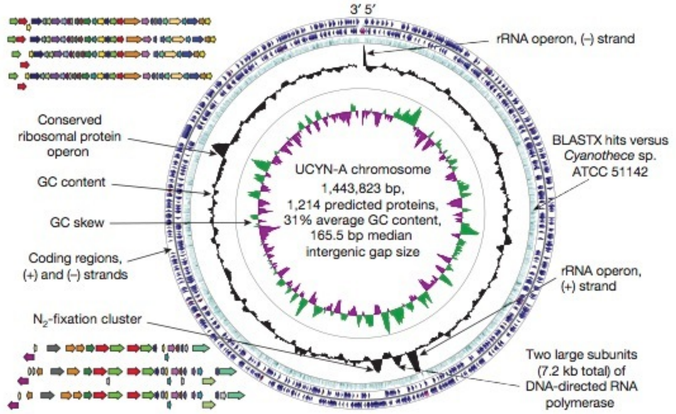
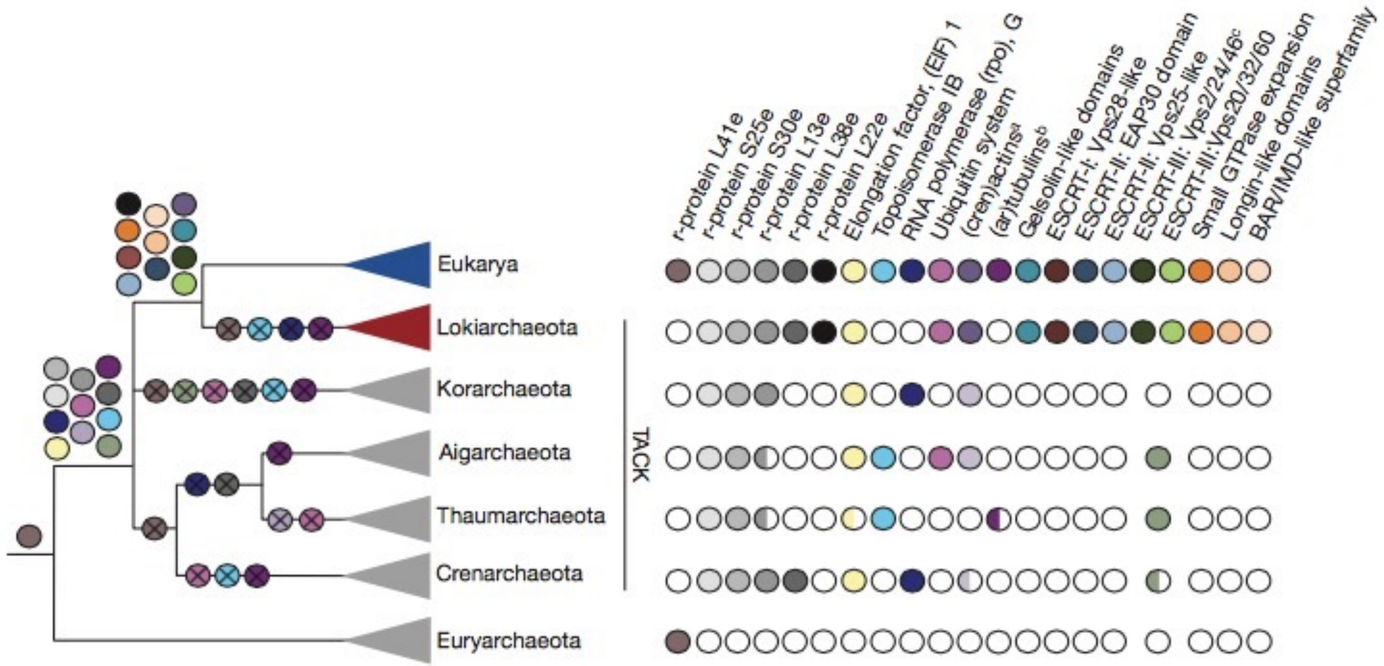


Figure 1 | Circular map of UCYN-A chromosome compared with *Cyanothece* sp. ATCC 51142. The outer two rings show the two strands of the

Lokiarcheota

- Typ w domenie Archaea
- Zidentyfikowany na podstawie analiz metagenomowych
- Żaden przedstawiciel nie został wyizolowany ani wyhodowany!
- Najbliżej spokrewniona z Eukaryota
- Posiada geny kodujące białka umożliwiające tworzenie złożonych struktur błonowych
- Tak mógł wyglądać gospodarz endosymbiozy, która dała początek Eukaryota



ARTICLE

doi:10.1038/nature14447

Complex archaea that bridge the gap between prokaryotes and eukaryotes

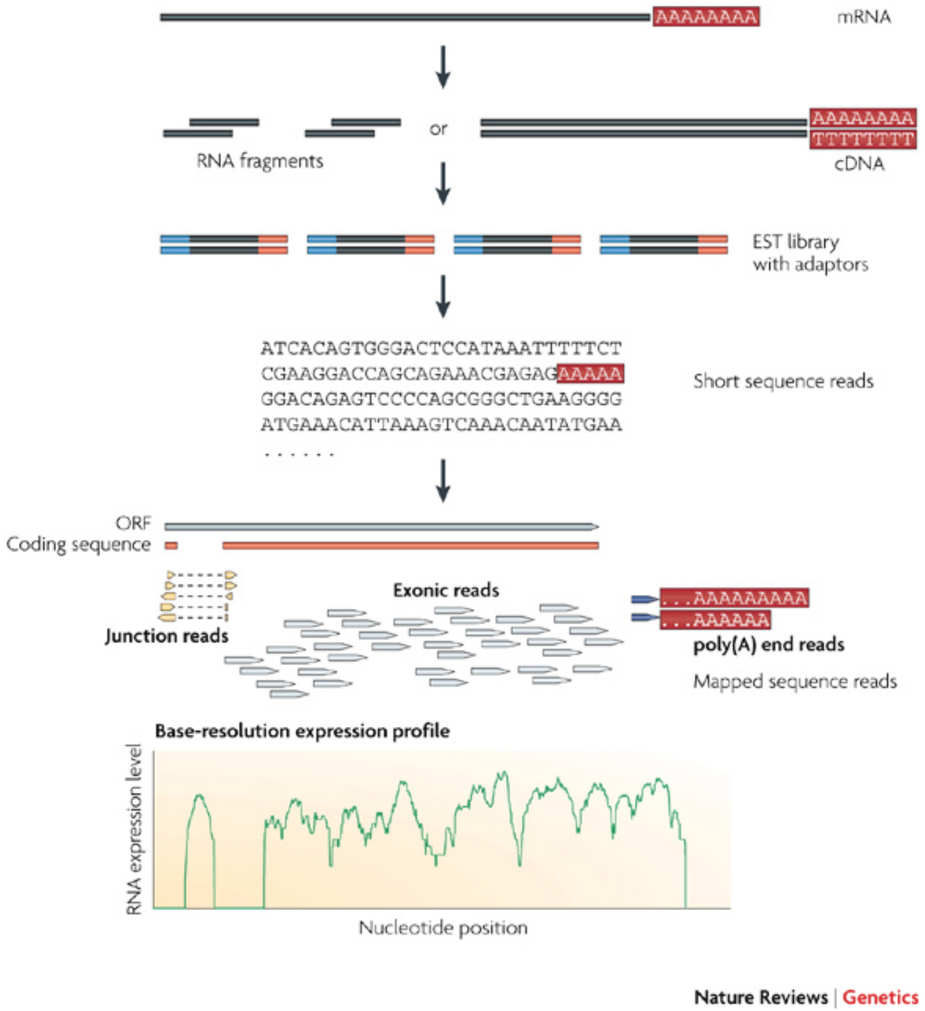
Anja Spang[1]*, Jimmy H. Saw[1]*, Steffen L. Jørgensen[2]*, Katarzyna Zaremba-Niedzwiedzka[1]*, Joran Martijn[1], Anders E. Lind[1], Roel van Eijk[1]†, Christa Schleper[2,3], I

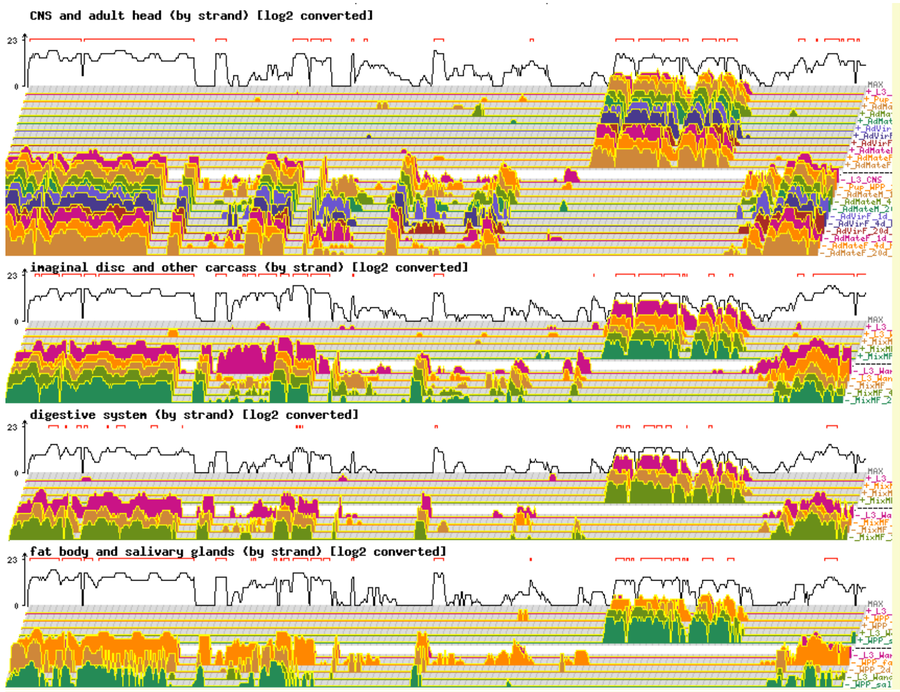
©2015 Macmillan Publishers Limited. All rights reserved

Wielkie projekty

- Projekt 1000 genomów - różnorodność genetyczna człowieka
- Metagenomika mikrobiomu przewodu pokarmowego człowieka
- Genomy wymarłych gatunków (np. Neandertalczyk)

RNA-seq

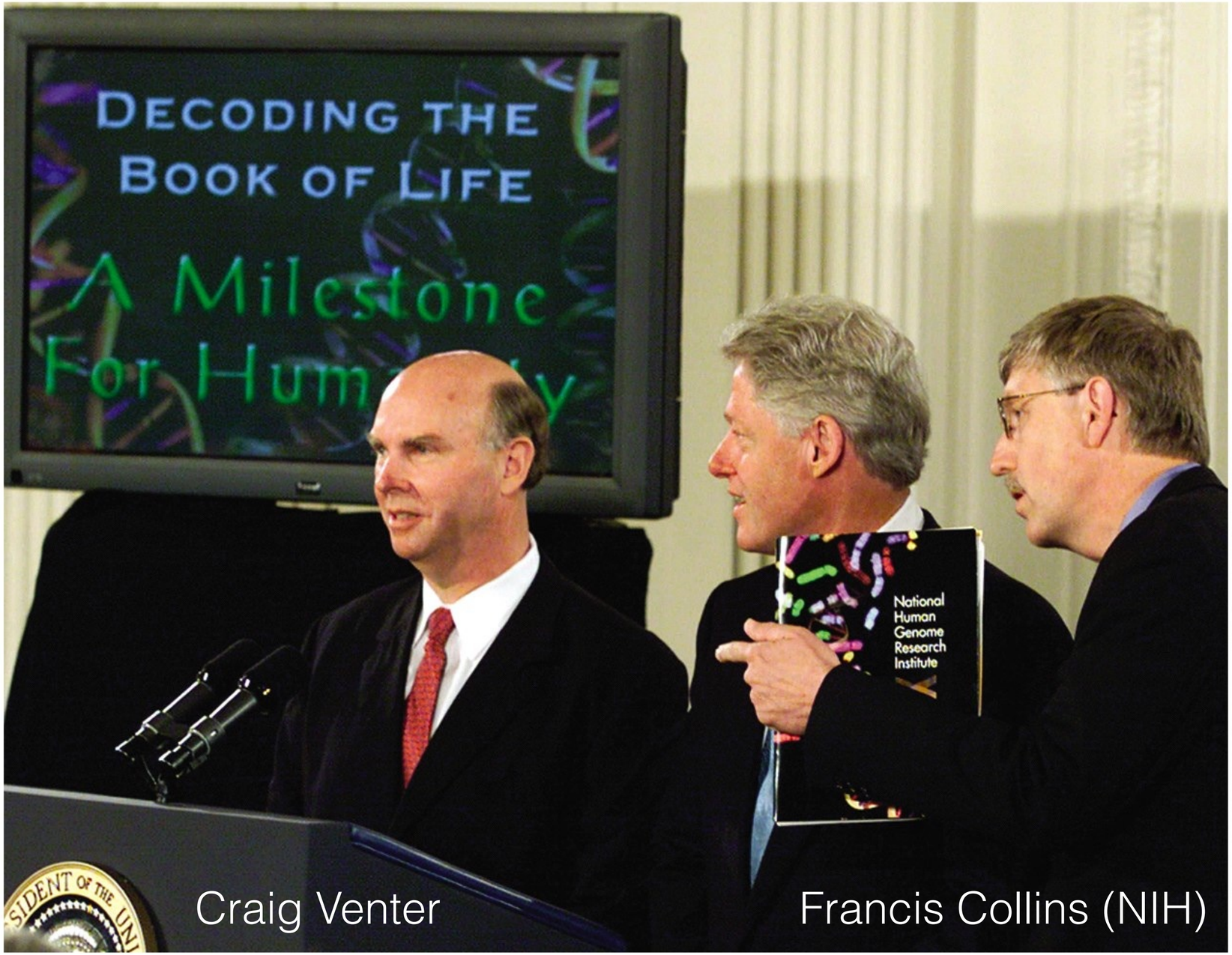

AAAAAAAA
mRNA
or
AAAAAAAA
TTTTTTTT
RNA fragments
cDNA
EST library
with adaptors
ATCACAGTGGGACTCCATAAATTTTTCT
CGAAGGACCAGCAGAAACGAGAGAAAAA
GGACAGAGTCCCCAGCGGGCTGAAGGGG
ATGAAACATTAAAGTCAAACAATATGAA
Short sequence reads
ORF
Coding sequence
Exonic reads
Junction reads
...AAAAAAAAA
...AAAAAA
poly(A) end reads
Mapped sequence reads
Base-resolution expression profile
RNA expression level
Nucleotide position
Nature Reviews | Genetics


CNS and adult head (by strand) [log2 converted]
imaginal disc and other carcass (by strand) [log2 converted]
digestive system (by strand) [log2 converted]
fat body and salivary glands (by strand) [log2 converted]

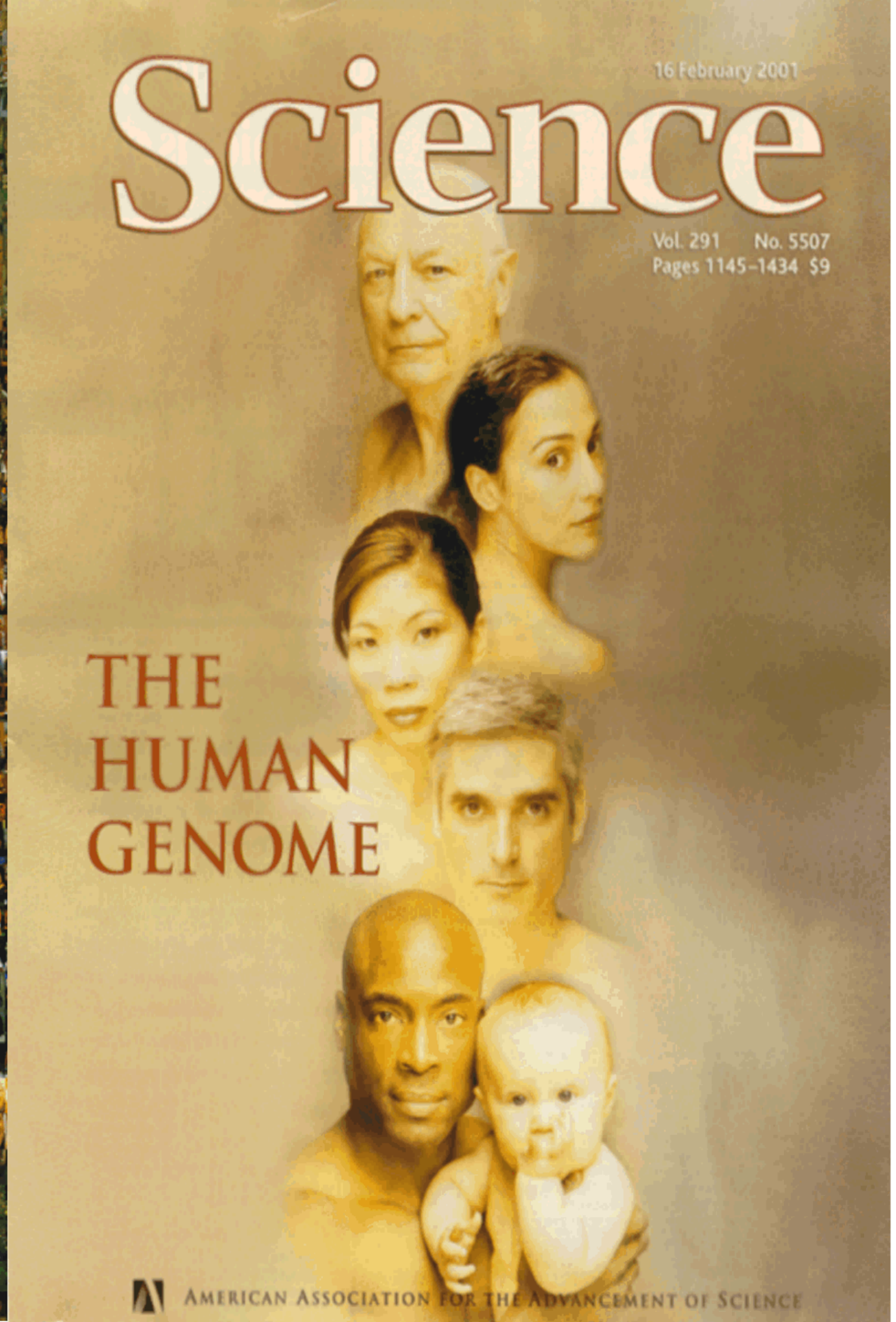
Sekwencjonowanie nowej generacji – wyzwanie dla bioinformatyki

- Krótkie odczyty (50-150 nt)
  - pojedyncze
  - “paired-end”
- Problem identyfikacji i składania sekwencji
- Indeksowanie i multipleks

Genomika funkcjonalna


DECODING THE
BOOK OF LIFE
A Milestone
National
Human
Genome
Research
Institute
Craig Venter
Francis Collins (NIH)

15 February 2001
nature
$10.00
www.nature.com
the human genome
Nuclear fission
Five-dimensional energy landscapes
Seafloor spreading
The view from under the Arctic ice
Career prospects
Sequence creates new opportunities
naturejobs
genomics special


16 February 2001
Science
Vol. 291 No. 5507
Pages 1145–1434 $9
THE HUMAN GENOME
AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE

Czym jest znajomość genomu

- Nie jest “odczytaniem księgi życia”
- Sama sekwencja nie daje jeszcze zrozumienia, jak funkcjonują komórki
  - Ale jest niezwykle cennym narzędziem w badaniach
- Sekwencja nie jest lekarstwem
  - Ale bardzo pomaga w zrozumieniu mechanizmów chorób i wynajdywaniu nowych terapii

Co genom już dał nauce

- Dużo ciekawych i zaskakujących odkryć
  - Dlaczego mamy tak mało genów
  - Jak ewoluował człowiek
- Narzędzie do badania funkcji genów

Czy warto badać genomy

- Nowoczesne techniki generują bardzo dużo danych
- Dwa podejścia
  - “*hypothesis driven*” – dane zbierane dla zweryfikowania jakiejś hipotezy. Klasyczna metoda naukowa.
  - “*data driven*” – dane zbierane bez wstępnych założeń, potem wyszukiwane w nich prawidłowości.

Koniec hipotezy?

- Podejście zakładające poszukiwanie prawidłowości w dużych zbiorach danych, zbieranych bez wstępnych założeń, może być produktywne
- Ale niesie też (dobrze znane w literaturze) ryzyko
  - Lem, S. Cyberiada, Wyprawa szósta: czyli jak Trurl i Klapaucjusz demona drugiego rodzaju stworzyli, aby zbójcę Gębona pokonać., 1965.

“[…] i pomału diabli go zaczynali brać, bo już mu świtało, że wszystkie owe całkiem prawdziwe i ze wszech miar sensowne informacje zupełnie nie są mu potrzebne, gdyż robił się z tego groch z kapustą, od którego głowa pękała, a nogi drżały.”

–Stanisław Lem, “Wyprawa szósta […]”

Co to znaczy? Genomika funkcjonalna.

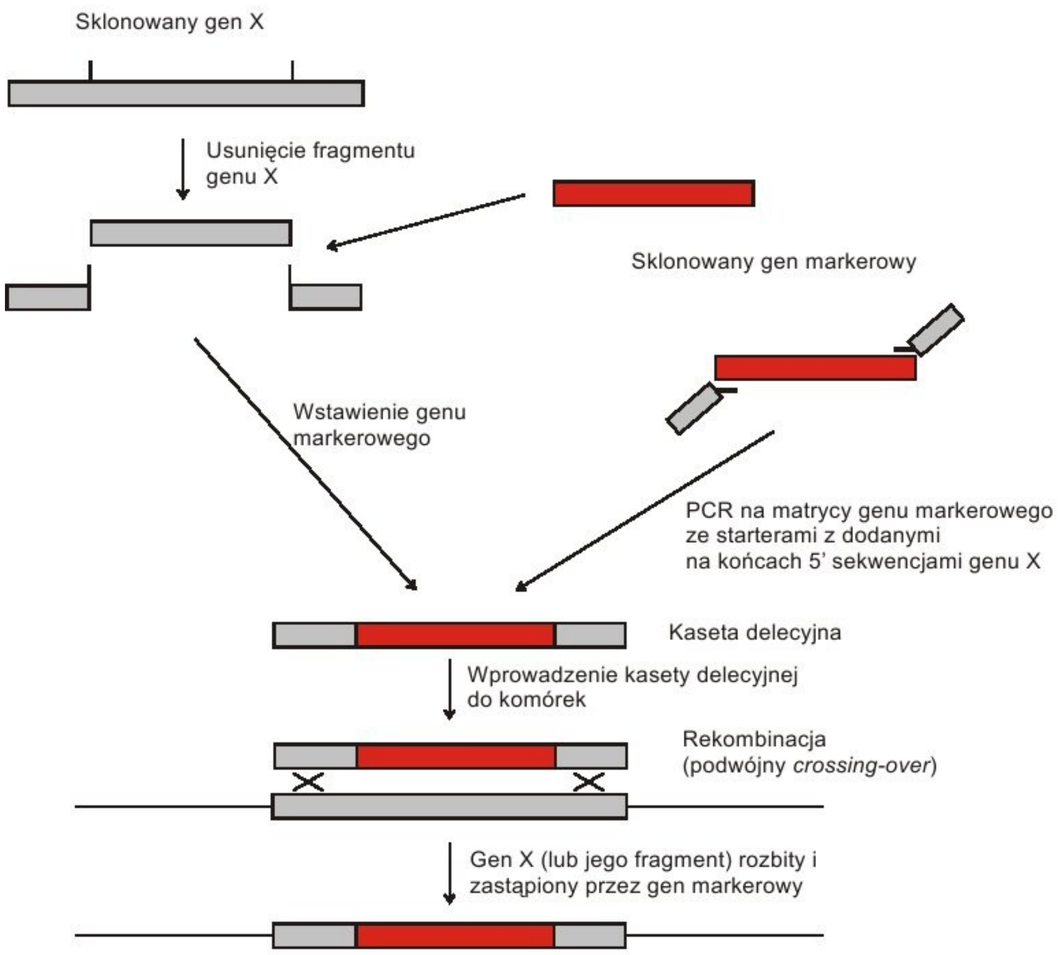
TCACAATTTAGACATCTAGTCTTCCACTTAAGCATATTTAGATTGTTTCCAGTTTTCAGCTTTTATGACTAAATCTTCTAAAATTGTTTTTCCCTAAATGTATATTTTAATTTG
TCTCAGGAGTAGAATTTCTGAGTCATAAAGCGGTCATATGTATAAATTTTAGGTGCCTCATAGCTCTTCAAATAGTCATCCCATTTTATACATCCAGGCAATATATGAGAG
TTCTTGGTGCTCCACATCTTAGCTAGGATTTGATGTCAACCAGTCTCTTTAATTTAGATATTCTAGTACATACAAAATAATACCTCAGTGTAACCTCTGTTTGTATTTCCCT
TGATTAACTGATGCTGAGCACATCTTCATGTGCTTATTGACCATTAATTAGTCTTATTTGTTAAATGTCTCAAATATTTTATACAGTTTTACATTGTGTTATTCATTTTTTAAA
AAATTCATTTTAGGTTATATGTATGTGTGTGTCAAAGTGTGTGTACATCTATTTGATATATGTATGTCTATATATTCTGGATACCATCTCTGTTTCATGCATTGCATATATATTT
GCCTATTTAGTGGTTTATCTTTTCATTTTCTTTTGGTATCTTTTCATTAGAAATGTTATTTATTTTGAGTAAGTAACATTTAATATATTCTGTAACATTTAATGAATCATTTTATG
TTATGTTTAGTATTAAATTTCTGAAAACATTCTATGTATTCTACTAGAATTGTCATAATTTTATCTTTTATATACATTGATATTTTTATGTCAAATATGTAGGTATGTGATATTATG
CACATGGTTTTAATTCAGTTAATTGTTCTTCCAGATGTTTGTACCATTCCAACATCATTTAAATCATTAAATGAAAAGCCTTTCCTTACTAGCTAGCCAGCTTTGAAAATC
CATTCATAGGGTTTGTGTTAATATATTTTTGTTCTTTTTTTTCCTTTCTACTGATCTCTTTATATTAATACCTACTGTGGCTTTATATGAAGTCATGGAATAATACGTAGTAAG
CCCTCTAACACTGTTCTGTTACTGTTGTTATTGTTTTCTCAGGGTACTTTGAAATATTCGAGATTTTATTATTTTTTAGTAGCCTAGATTTCAAGATTGTTTTGACGATCAAT
TTTTGAATCAATTGTCAATATTTTTAGTAATAAAATGATGATTTTTGATTGGAAATACATTAAATCTATAAGCCAAATTGGAGATTATTGATATATTAACAAAAATGAGTTTTCC
AGTCCATGAATGTATGCACATTATAAAATTCATTCTTAAGTATGTCATTTTTTAAGTTTTAGTTTCAGCAGTATATGTTTGTTACATAGGTAAACTCCTGTCATGGGGGTTA
GTTGTACAGGTTATTTTATCATCCAGGCATAAAGCCCAGTACCCAGTAGTTATCTTTTCTGCTCCTCTCCCTCCTGTCACCCTCCACTCTCAAGTAGACCCCAGTTTC
TGTTGTTCTCTTCTTTGCATTAATGACTTCTCATCATTTAGATTGCACTTGTAAGTGAGAACAGGACGTATGTGGTTTTCTACTCCTGTGTTAGTTTGCTAAGGATAACC
ACCTCCATCTCCATCCATGTTCCCACAAAAGACATGATCTCCTTTTTTATGGCTGCATATTATTCCATGGTATATATGTACCACATTTTCTTTATCCAATCTGTCATTGATG
GACATTTAGGTTGTTTCCACATCATTGCCGTTGTAAATACTGCTGCAGTGAATATTCGTGTGTATGTCTTTATGGTAGAATGATTTATATTCCTCTGGGTATATTTCCAAGT
AATGGGATGGTTGGGTCAAATGGTAATTCTGCTTTTAGCTTTTTGAGGAATTGCCATATTGCCTTTCACAACGGTTGAACTAATTTATACTCCCAAGAGTGTATAAGTTG
TTCCTTTTTCTCTGCAACCTCGACATCACCTGTTATTTATGACTTTTATATAATAGCCATTCTGCTGGTCTGAGATGGTATCTCATTATGATTTTGATTTGCATTTCTCTAAT
GCTCAGTGATATTGAGCTTGGCTGCATATATGTCTTCTTTTAAAAATATCTGTTCATGTCCTTTGCCTAATTTATAACGGGGTTGTTTGTTTTTCTCTTGTAAATTTGTTTAA
GTTCCTTATAGATTCTAGGTATTAAACCTTTTTTCAGAGGCGTGGCTTGCAAATATTTTCTCCCATTCTATAGGTTGTCTGTTTATTCTGTTGATAGTTTCCCTTGCTGTG
CAGAAGCTCTTAACTTTAATTAGATCCGACTTGTCAATTTTTGCTTTGGTCGCAATTGCTTTTGATGTTATTGTCGTGAAATCTTTGCTAGTTCTTAGGTCCAGGATGATA
TTGCCCAAGTTGTCTTCCAGGGCTTTTATAATTTTGGATTTTACATTTAAGTCTTAATATATTTATTAAATTTGTTAGGGTTTCAGGATACAAGGACAATATAGCAGCAAAC
AATGTAAAAGTAAAATCTGAAAAATAATAGAAAACAGTTTAATTGAACACTTTACCATTATGTAATGCCCTTCTTTGTCTTTCCTGATCTTTGTTGGTTTGAAGTTCAAAAA
AGACAAACTTAATGGTACAATAGGTATTGTAGATTTCAGGACTTTCTGTATAAAATATTTTGTATATATGAATAGATCATTTTTTATTTCCAGTCTTTAAACATTTTCTTAACAT
TTTCTTCTATTGCTTCACTTCACTCGCTAGGACCATCAGGACAGTGTTGAACAGAAATTGTCAGACTGATCATCACAACTTTTTCTAGATTTTAGAAGGAAATTTTTCTT
TATTTCAACATAAAGCAGCATGTTAATGCCAAGTTTTAATATGTGTTATCAGATTGAAATTTTTTTGTATATTTCTACATTACCAAGAATTTTTAGCAAGAGTTTTTGTTGAG
TTTTAATTTAAAAATCATTTGTTAATTTCATCTGATTTTTTTATTTCTCTTTTTACCTTAAGAGATTAAACTGACTACAGATTGAATATAAACAAACAAACAAACAAACAAAAA
CTCTAAAATGCTGTGGATCAACACCACTTAGTAATTTGTATACTTGGATTCAATTTGCTGAAATTTTGTTAGACATTTTTGCGTCGATATTTATGAGGGATGTTGATCTGT
AAAAGTATTAAAATGCCTTTGACAGATAGTGTCACCATATAAAAAACTTTGAACAAAATCAGATTATATCACTGTGGATATTTCTATTTTGAACTAACTTAGATGATAATTTT
AATCTATATCCTAGATGAACT

Mały fragment chromosomu 21

Odwrotna genetyka – od genu do funkcji

Genetyka tradycyjna	Genetyka „odwrotna”
Funkcja (mutacja, fenotyp)	Gen (z sekwencji całego genomu)
↓	↓
Klonowanie genu	Inaktywacja genu
↓	↓
Analiza sklonowanego genu	Analiza uzyskanego fenotypu

Odwrotna genetyka – inaktywacja przez rekombinację

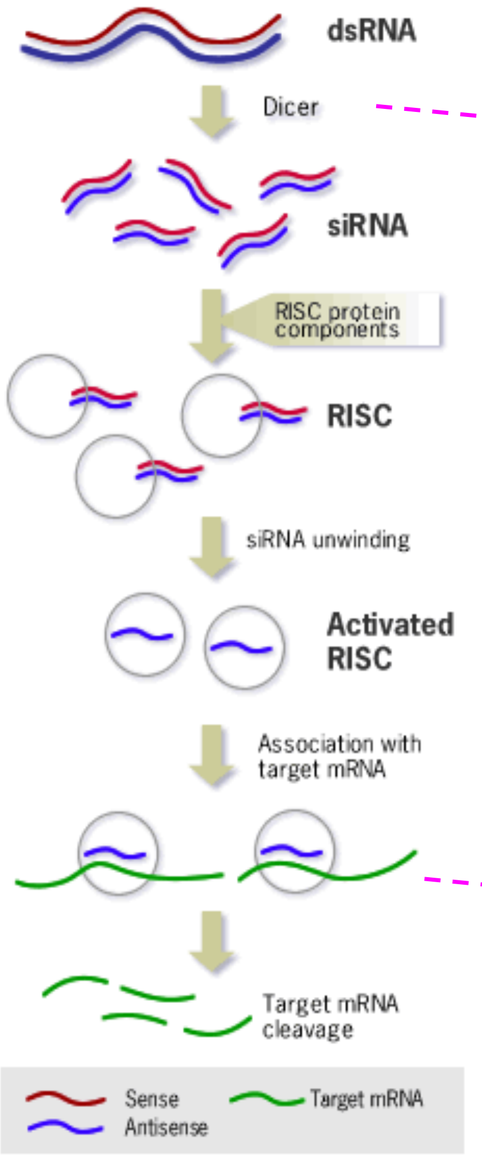


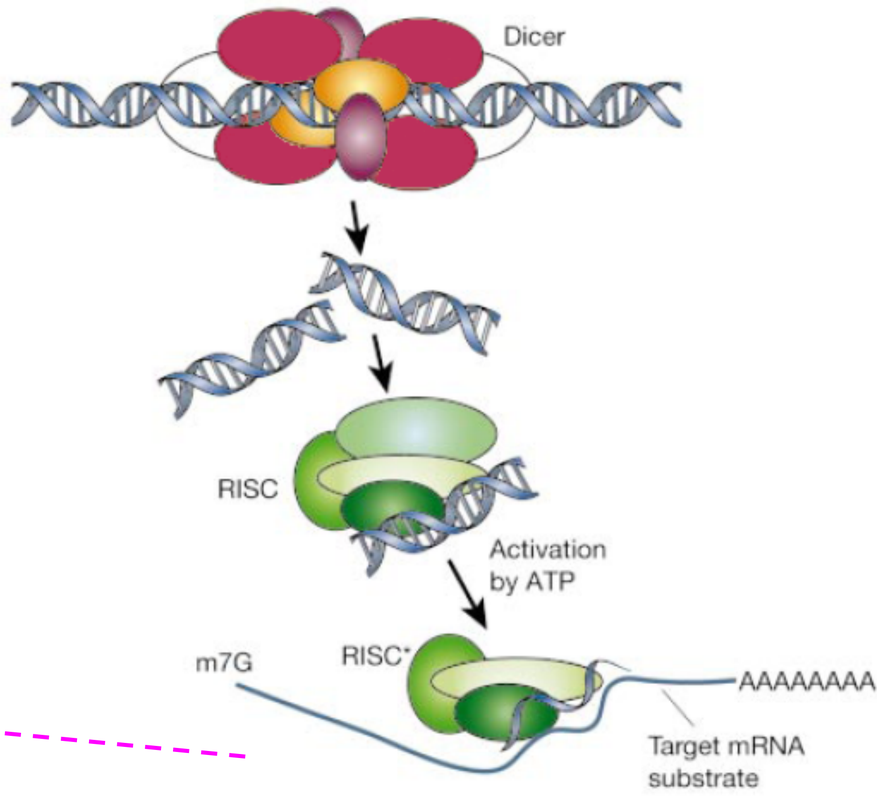
Odwrotna genetyka – interferencja RNA

Odkrycie roku 2002 –
regulacyjna rola małych RNA

Nagroda Nobla w dziedzinie medycyny 2006, za odkrycie mechanizmu interferencji RNA
A. Fire i C. Mello

siRNA - jak to działa?





Efekt – degradacja mRNA

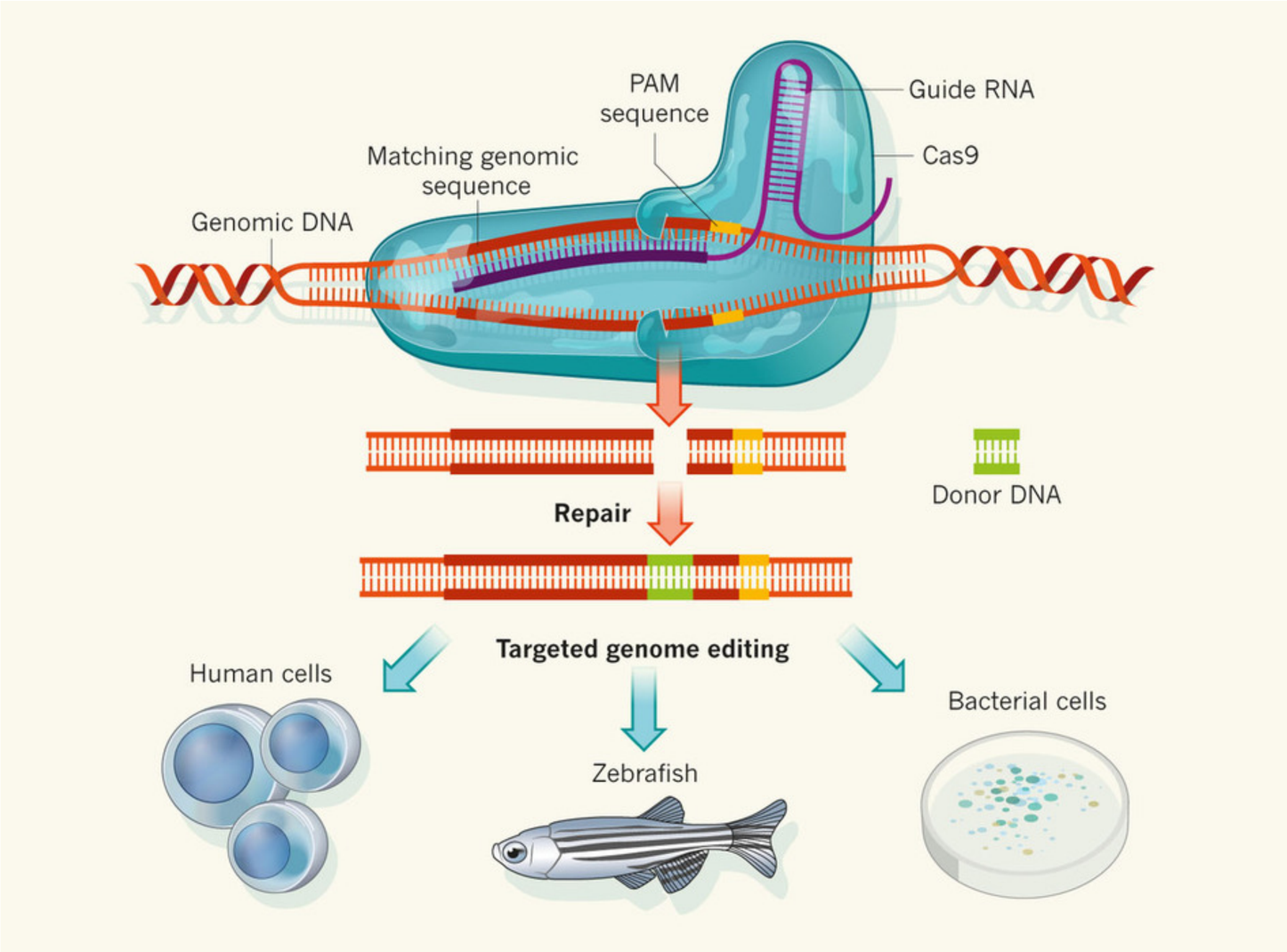
Hannon G.J.: ‘RNA interference’, Nature 418, July 11, 2002

Zastosowania siRNA

- Badanie funkcji genów (“odwrotna genetyka”) - szczególnie skuteczne u nicienia Caenorhabditis, ale działa też w komórkach owadów, ssaków i roślin

- Hamowanie wybranych genów jako metoda leczenia (np. zwalczania wirusów czy nowotworów) – przykład: obniżenie poziomu receptora LDL („zły cholesterol”) u myszy

Redagowanie genomu - system CRISPR/Cas9



System obronny bakterii przed fagami, zaadaptowany do edycji dowolnej sekwencji w genomie.

Działa także u organizmów, dla których nie istnieją stabilne wektory.

Nature 495, 50–51 (07 March 2013) doi:10.1038/495050a

Co to oznacza dla nauki i biotechnologii

- Można zmienić dowolny gen u roślin i zwierząt
- Można łatwo adaptować do kolejnych gatunków (w odróżnieniu od tradycyjnej inżynierii genetycznej)
- Modyfikacja nie pozostawia żadnego śladu, nie zostaje włączony obcy DNA
- Wg. prawa UE to nadal jest GMO

Common buckeye wings: normal (left) and with optix deleted (right) (Robert Reed)

A dla człowieka

- Możliwa jest zmiana sekwencji dowolnego genu w komórkach zarodkowych
- Skuteczność: około 50-80%
- W połączeniu z zapłodnieniem pozaustrojowym możliwość uzyskania zmodyfikowanego człowieka
- Pierwsze badania - na komórkach, z których nie mogły rozwinąć się zarodki (2015)
- Obecnie badania na komórkach zarodkowych człowieka

Anna Tärnhuvud

Fredrik Lanner, who will conduct gene-editing in human embryos at the Karolinska Institute in Sweden.

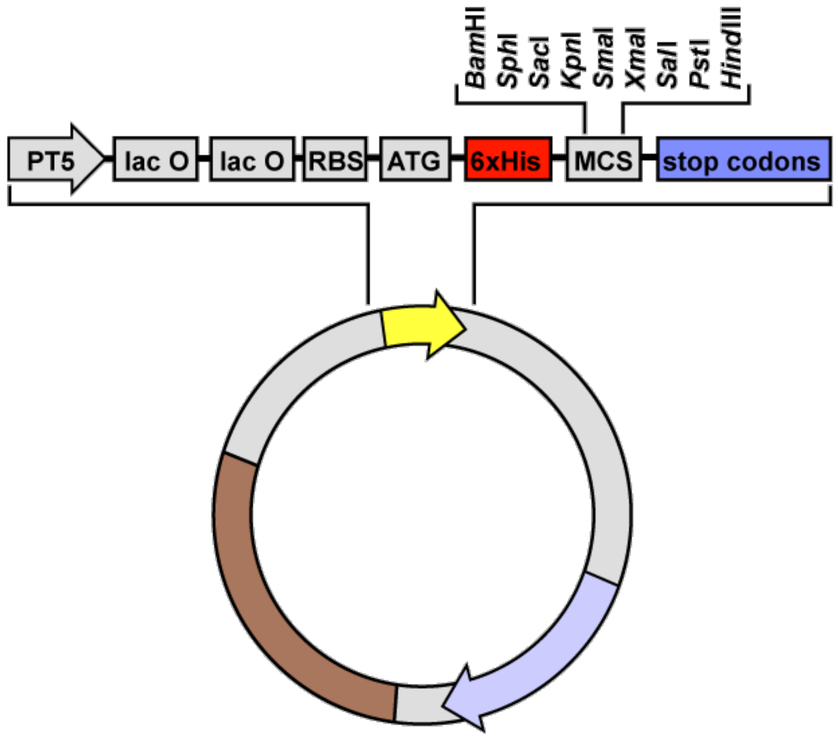
Dzieci na zamówienie?

Photosani/Shutterstock

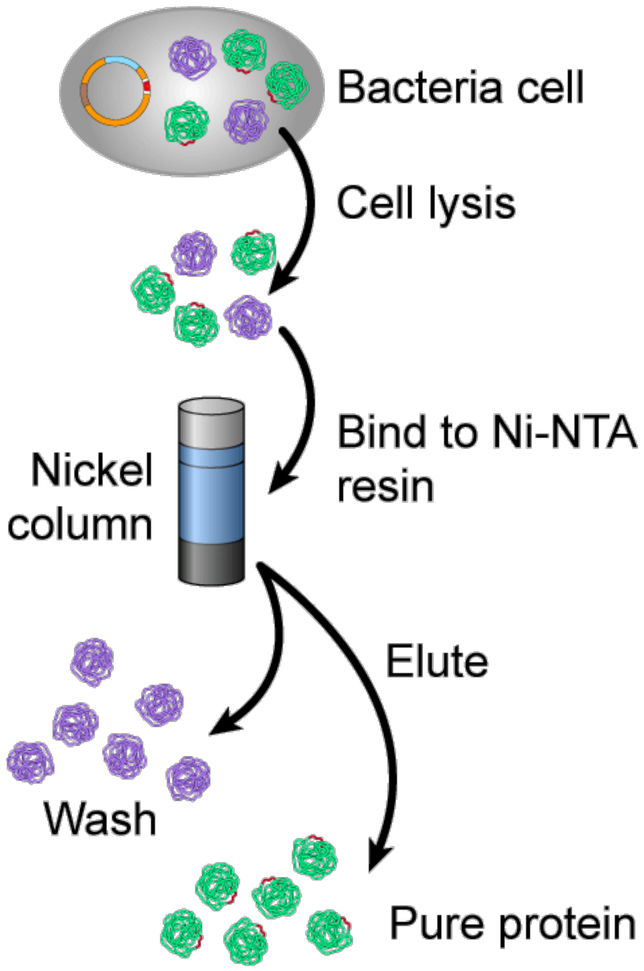
?

Badanie i inżynieria białek

Ekspresja heterologiczna



Wektor ekspresyjny

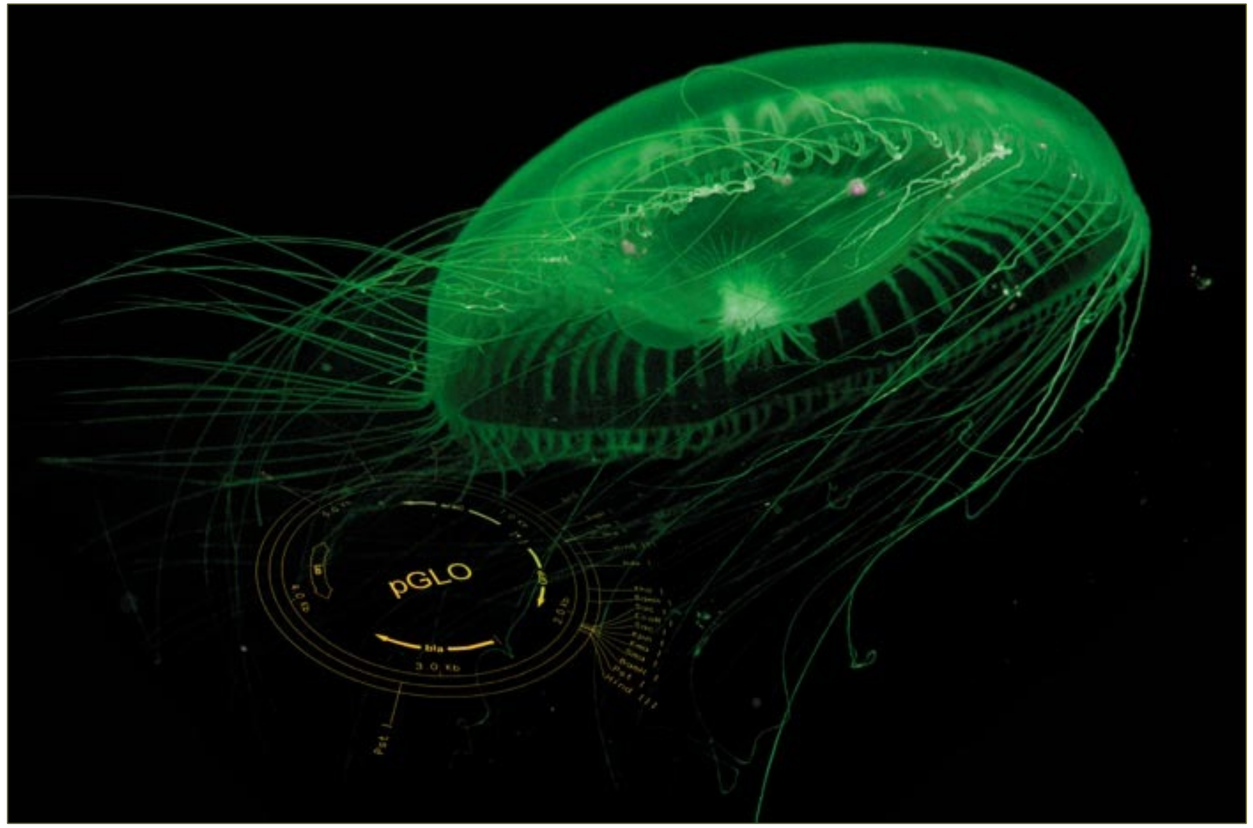


Oczyszczanie białka

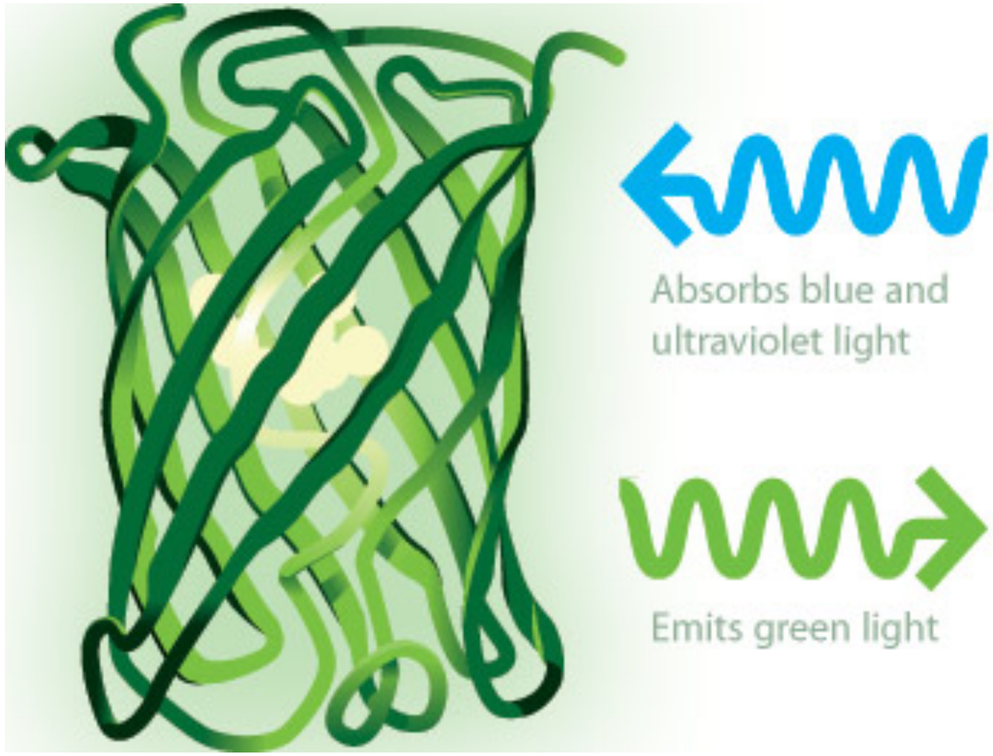
Fuzje białkowe

- Do sekwencji kodującej białko dołącza się inną domenę, np. ułatwiającą wykrycie i/lub oczyszczanie
  - znaczniki epitopowe
  - GFP (zielone białko fluorescencyjne)

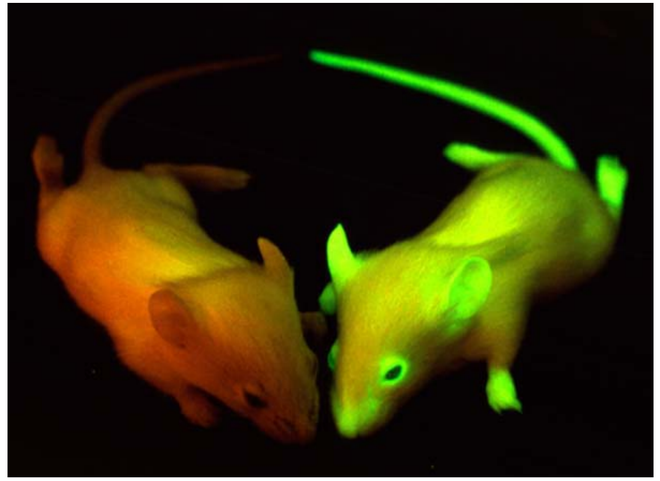
Green Fluorescent Protein-

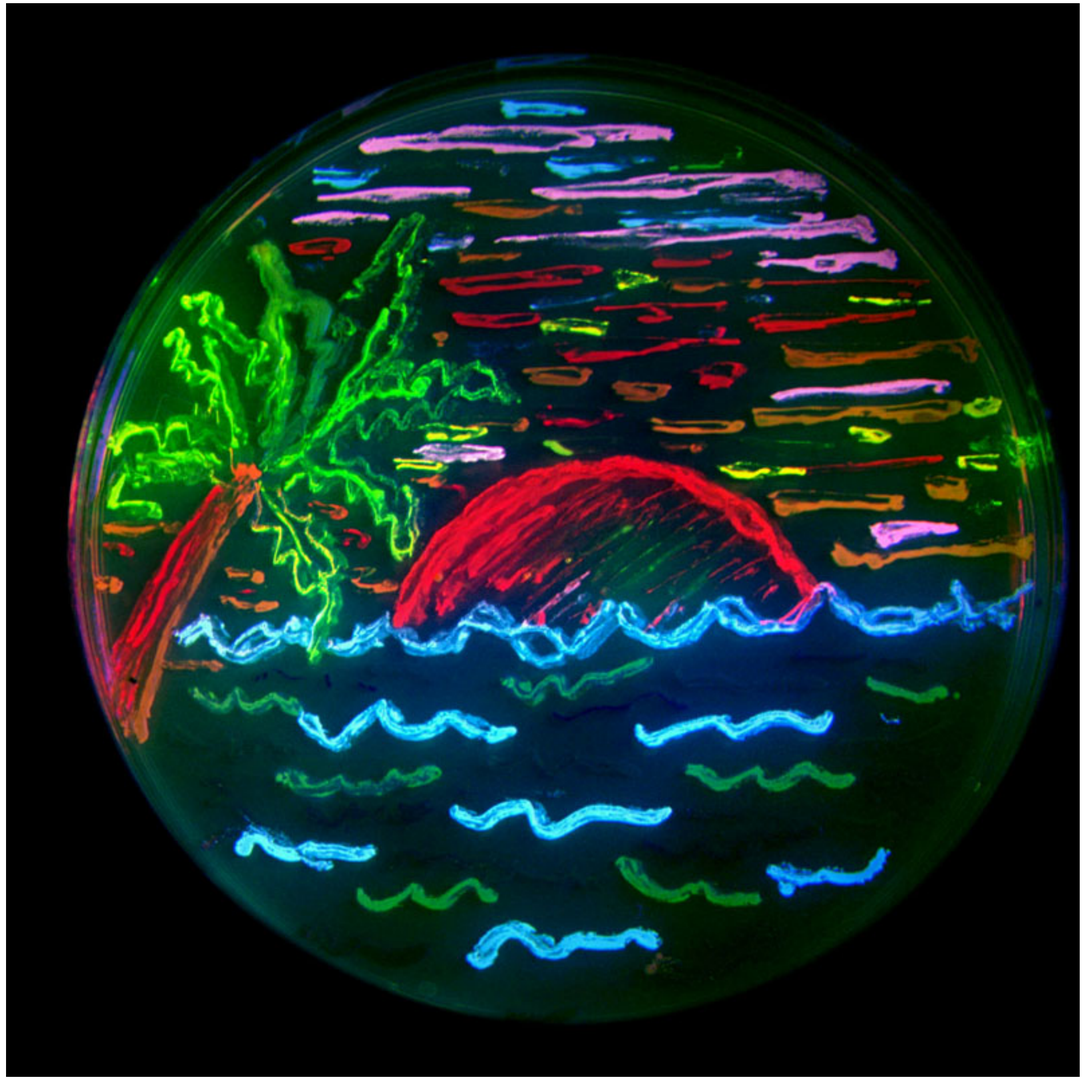


Aequorea victoria

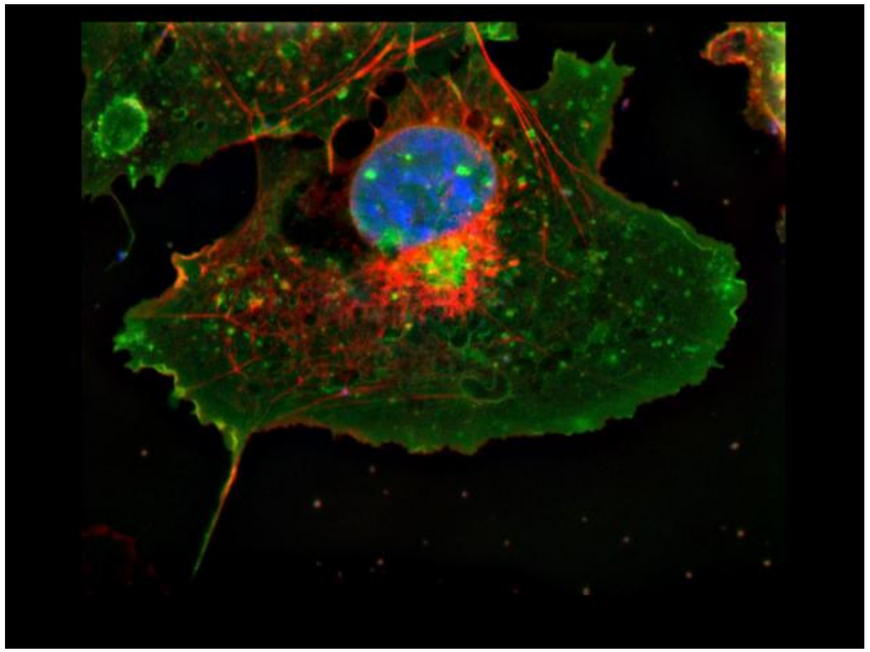


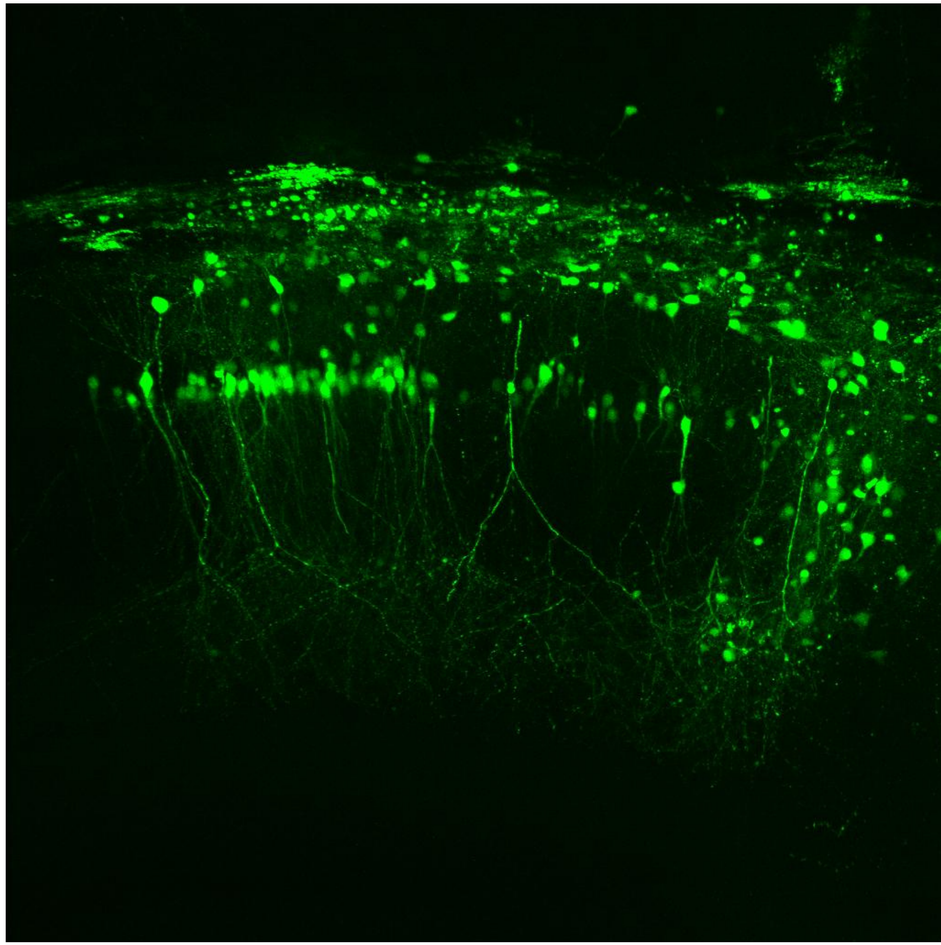
GFP



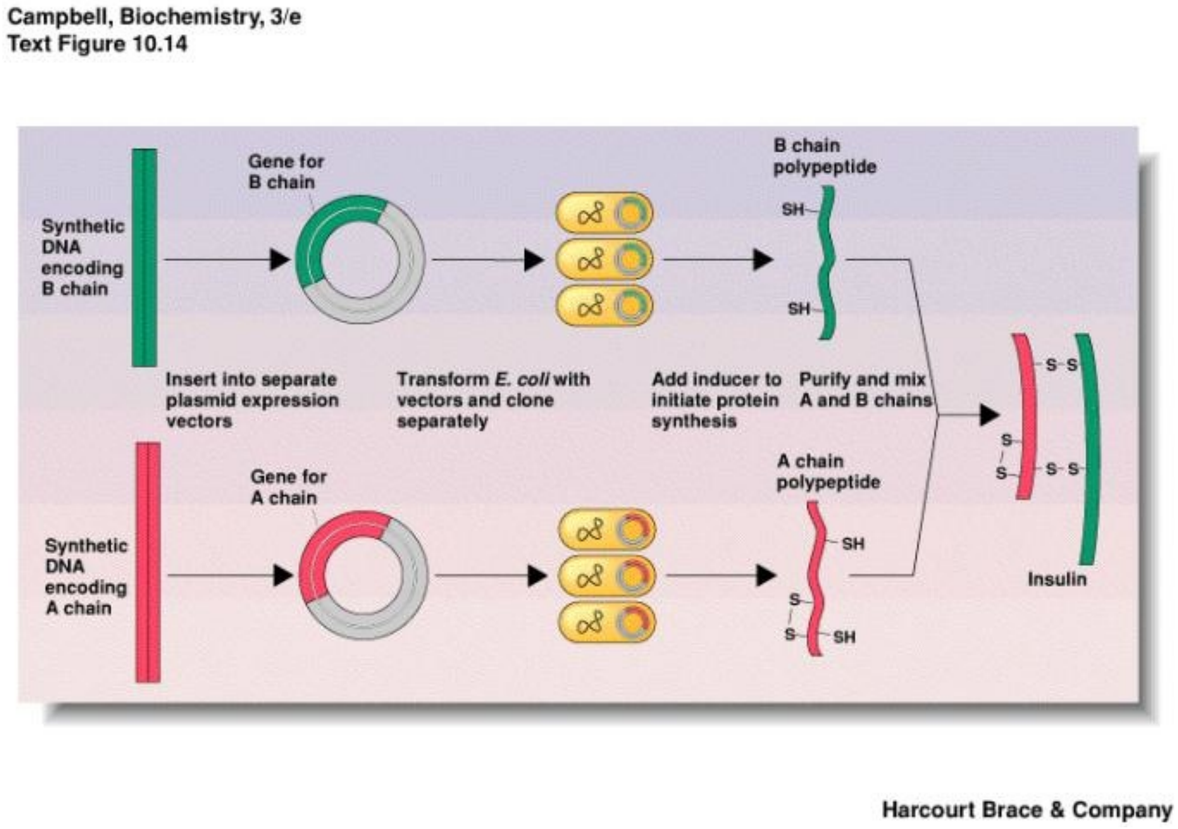


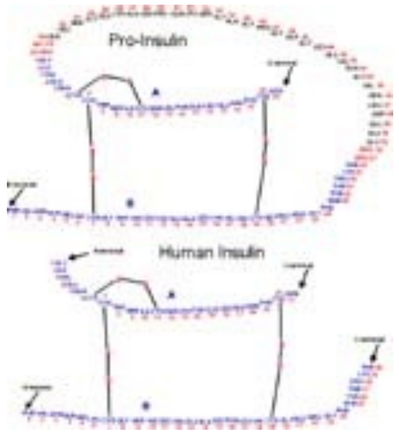
GFP

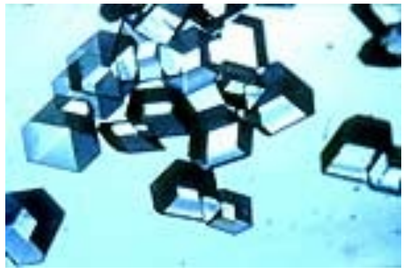




Ekspresja heterologiczna w biotechnologii

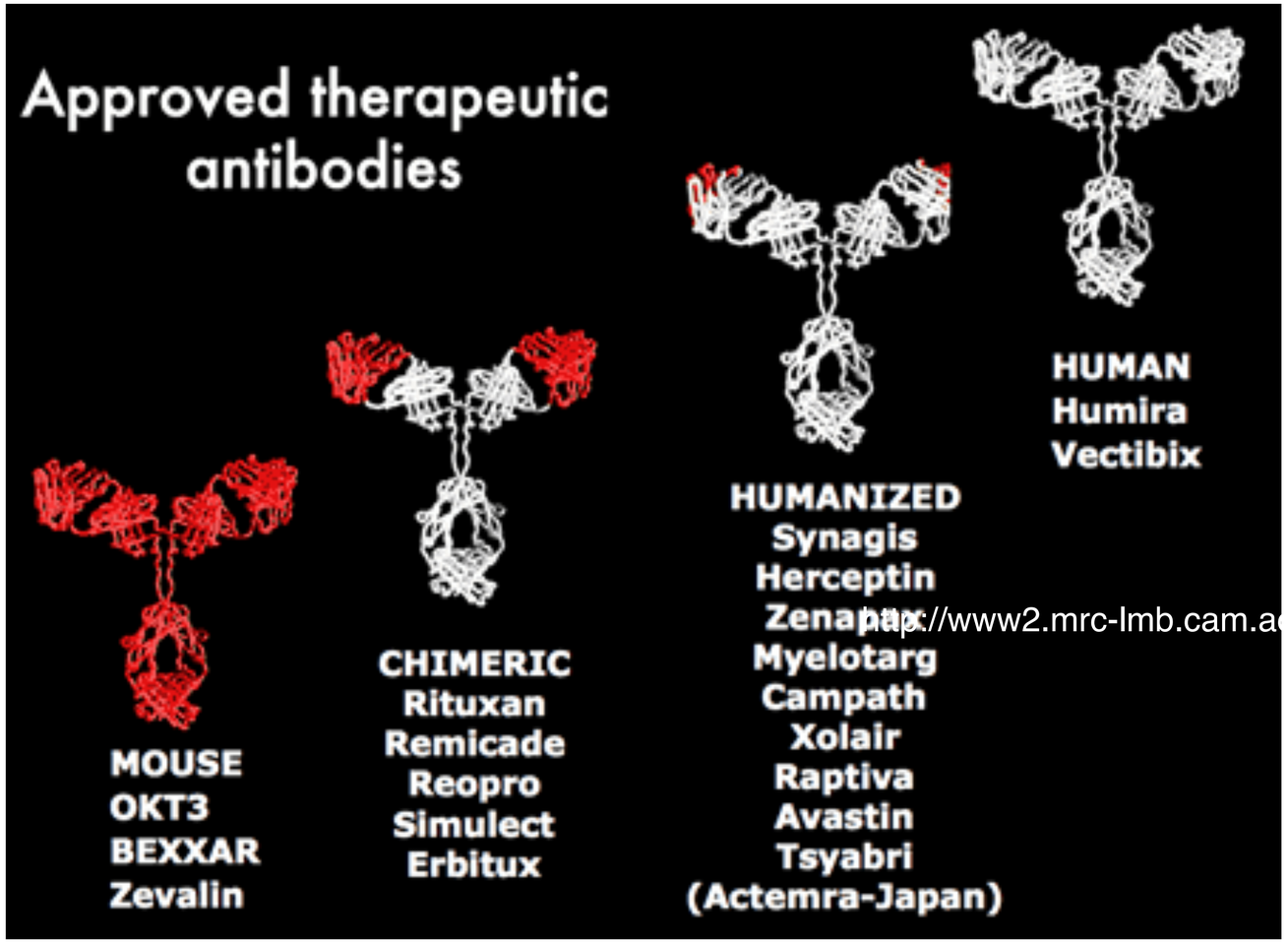






Ludzka insulina wytwarzana w bakteriach

Inżynieria przeciwciał



Humanizowane i rekombinowane przeciwciała są stosowane w terapii np. nowotworów

Inżynieria przeciwciał – leki przyszłości

- W łańcuchach immunoglobulin obszary zmienne odpowiadają za rozpoznawanie różnych antygenów, obszary stałe nadają specyficzność gatunkową
- W klasycznych metodach wytwarzania przeciwciał monoklonalnych uzyskiwano przeciwciała zwierzęce
- Klonowanie i ekspresja genów kodujących przeciwciała jest alternatywnym sposobem ich uzyskiwania
- Możliwe jest uzyskiwanie przeciwciał humanizowanych – obszary zmienne z genu przeciwciała zwierzęcego wstawione między obszary stałe przeciwciała ludzkiego
- Humanizowane i rekombinowane ludzkie przeciwciała są stosowane w terapii np. nowotworów

Biologia systemów i biologia syntetyczna

Biologia systemów

- Postęp biologii molekularnej - genomika
- Ogromne zbiory danych
- Ujawnienie złożoności interakcji genetycznych leżących u podstaw fenotypu
- Narzędzia teoretyczne do systemowego opisu życia

Biologia systemów - wyzwanie

A w biotechnologii?

- Współczesna biotechnologia molekularna bardzo sprawnie manipuluje pojedynczymi genami
  - ekspresja heterologiczna
  - transgeneza roślin

- A co z bardziej złożonymi, wieloczynnikowymi cechami?

Inżynieria ewolucyjna

Inżynieria ewolucyjna

Brassica oleracea var. *silvestris* (brzoskiew)

Brassica oleracea odmiany uprawne

A visual guide to identifying genetically modified corn

NON GM

GM

Biologia syntetyczna

- Współczesna inżynieria genetyczna ograniczona jest do prostych systemów, gdzie za pożądaną funkcję odpowiada jeden lub kilka genów

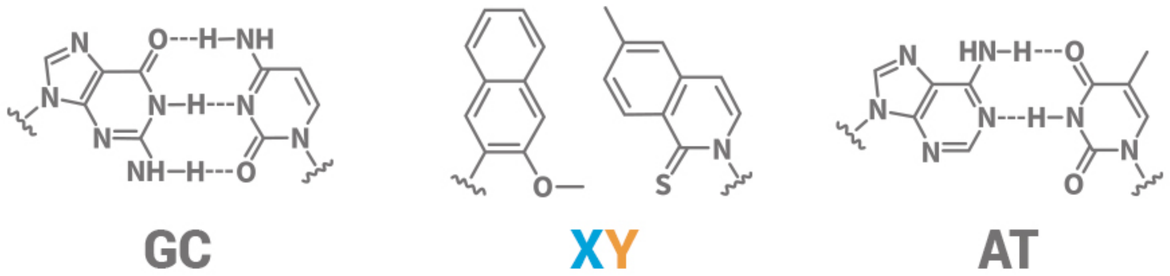
- Biologia syntetyczna - projektowanie nowych systemowych właściwości organizmów żywych

Biologia syntetyczna

- Dwa oblicza biologii syntetycznej:
  - Odtworzenie właściwości układów żywych (właściwości emergentne) za pomocą cząsteczek niewystępujących w naturze
    - E. Kool, R. Rawls, 2000 – podejście chemików
  - Wykorzystanie elementów układów biologicznych do konstrukcji nowych systemów o zaprojektowanych właściwościach
    - W. Szybalski, B. Hobom, 1980 – inżynieria genetyczna

Nowe zasady DNA

- Dodatkowa para zasad w DNA
- Utrzymuje się w genomie



synthorx.com

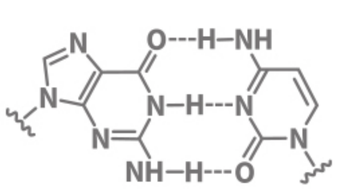
doi:10.1038/nature13314

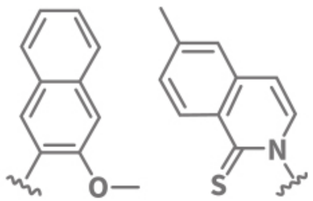
A semi-synthetic organism with an expanded genetic alphabet

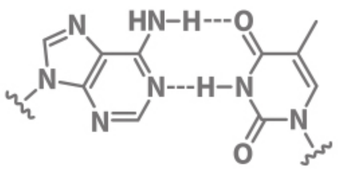
Denis A. Malyshev[1], Kirandeep Dhami[1], Thomas Lavergne[1], Tingjian Chen[1], Nan Dai[2], Jeremy M. Foster[2], Ivan R. Corrêa Jr[2] & Floyd E. Romesberg[1]

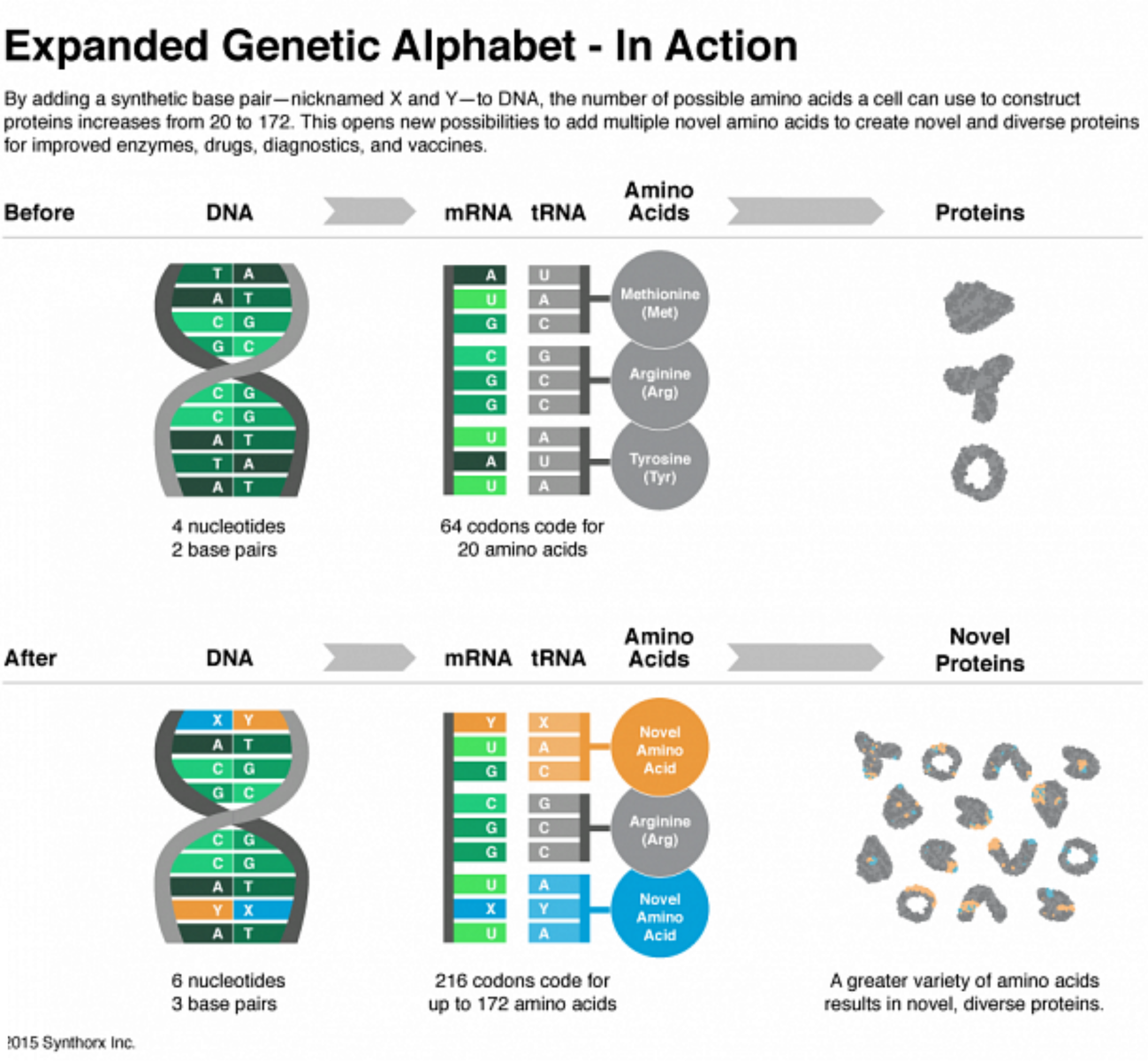
Syntetyczna biochemia

- Analogi cząsteczek biologicznych o nowych, rozszerzonych możliwościach
- DNA z dodatkowymi parami zasad
- Białka z nowymi aminokwasami









https://synthorx.com/

Biologia syntetyczna

- Wykorzystanie elementów układów biologicznych do konstrukcji nowych systemów o zaprojektowanych właściwościach

Biologia syntetyczna a inżynieria genetyczna

- Inżynieria genetyczna - manipulacje pojedynczymi genami - zmiany pojedynczych cech
  - bakterie wytwarzające insulinę
  - rośliny broniące się przed szkodliwymi owadami (wytwarzaniem jednego białka pochodzenia bakteryjnego)
- Biologia syntetyczna - konstrukcje organizmów o nowych właściwościach (“*biokonstruktorstwo*” - *S. Lem*)

Biologia syntetyczna

- Współczesna inżynieria genetyczna ograniczona jest do prostych układów, gdzie za pożądaną funkcję odpowiada jeden lub kilka genów
  - np. rośliny GMO oporne na herbicyd albo toksyczne dla owadów
    - nadal taka sama roślina, tylko 1-2 białka więcej
    - znacznie mniejsze różnice, niż między współczesnymi roślinami uprawnymi a ich dzikimi przodkami

- Biologia syntetyczna - projektowanie nowych systemowych właściwości organizmów żywych

Stan współczesny

- Możliwe są coraz bardziej zaawansowane manipulacje poszczególnymi składnikami życia
  - syntetyczne peptydy i białka, lipidy, DNA, RNA
  - syntetyczne geny i genomy
  - proste syntetyczne replikatory

- Nie udało się jeszcze stworzyć układu żywego w całości, bez udziału elementów uzyskanych z istniejących układów żywych

Podejścia biologii syntetycznej

- “od góry” (top-down) - głęboka modyfikacja istniejących systemów
  - minimalne genomy
  - syntetyczne genomy
  - przeprojektowane genomy

- Przykład - ortogonalny kod genetyczny

Manipulacje DNA

- DNA można manipulować w probówce, amplifikować, zmieniać itp.
- Ponownie wprowadzony do komórki funkcjonuje równie dobrze, jak “naturalny”

Synteza chemiczna DNA

- DNA można “skopiować” z istniejącego DNA lub RNA
  - nie więcej niż 10-20 tys. nukleotydów na raz
- DNA można syntetyzować chemicznie
  - Ograniczenie – nie więcej niż ~100 nukleotydów na raz
- Większe cząsteczki trzeba składać z tych małych fragmentów

Syntetyczne geny

- 1972, Khorana i wsp., gen kodujący tRNA drożdży, 77 par nukleotydów z 15 fragmentów (po 5-20 nt)
  - 1979 – pierwszy, który działał (207 pz)
- 1981, Edge i wsp., gen kodujący białko – interferon człowieka (514 pz), wprowadzony do bakterii

Pierwszy syntetyczny wirus

- 2002, Cello, Paul & Wimmer
- Matryca DNA wirusa polio złożona z syntetycznych fragmentów i przepisana na RNA
- Genom wprowadzony do komórek w hodowli namnaża się i produkuje cząstki wirusowe

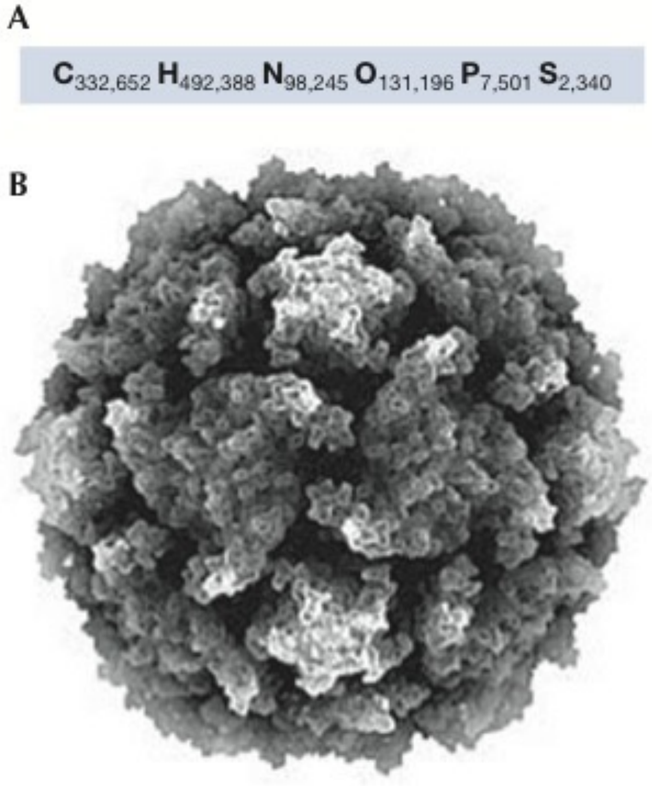
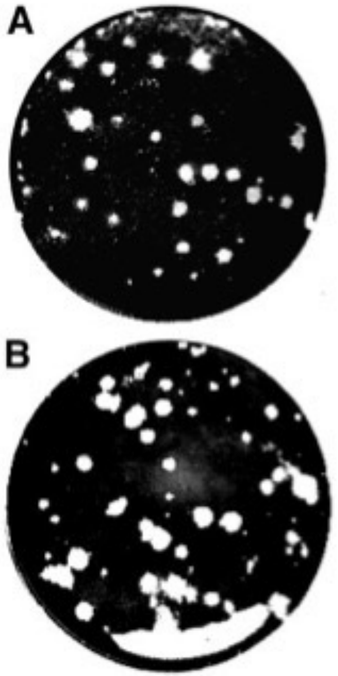
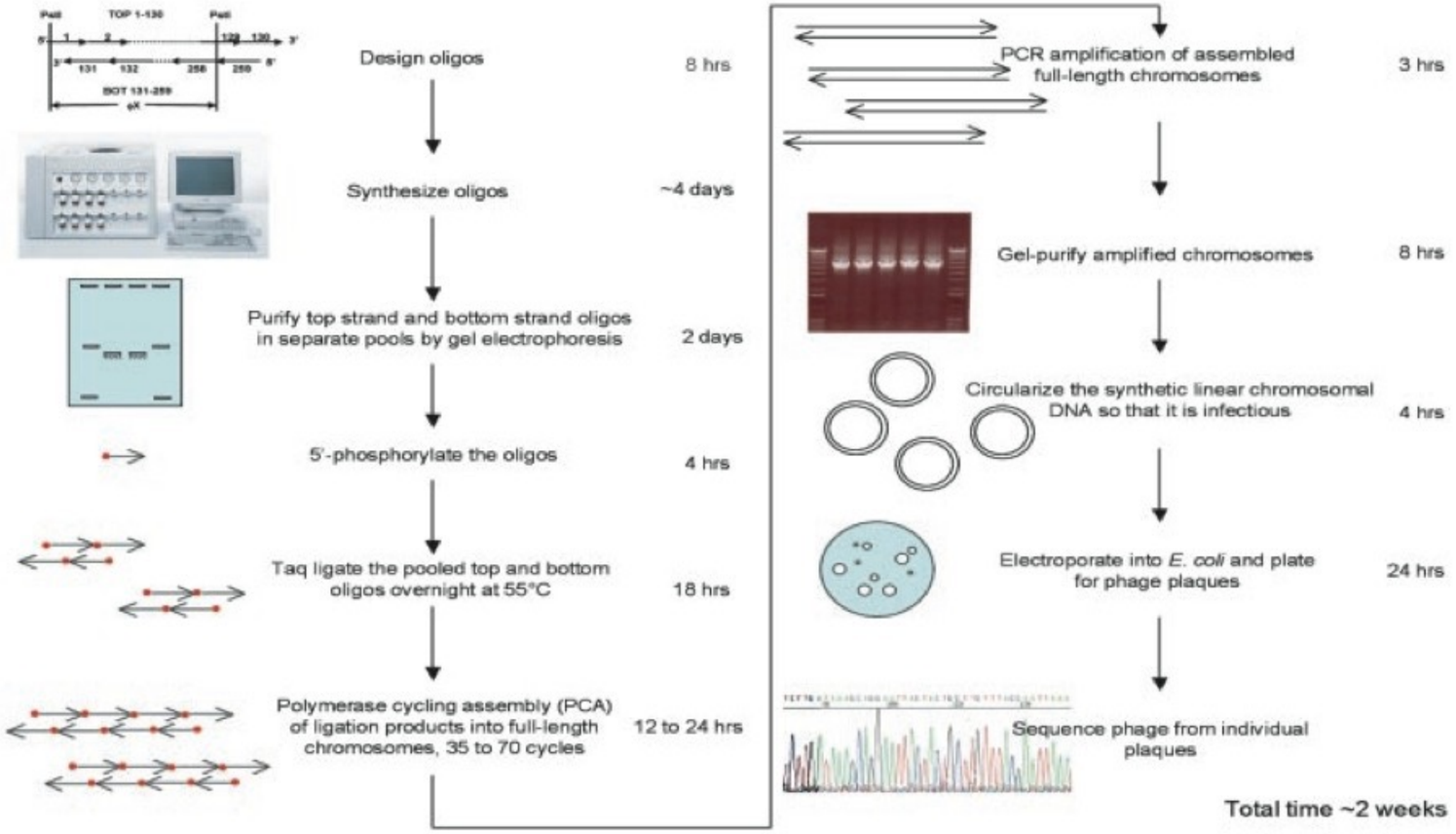


Fig 1 | Poliovirus and its empirical formula.



Nowe metody i fag ΦX174

- Składanie całego genomu faga w jednym etapie
- Cały projekt ~2 tygodnie (Smith i wsp. 2003)



Sukces

Fig. 4. Plaques of synϕX-A. There appear to be several plaque morphologies:

Smith i wsp. 2003

Syntetyczne wirusy

- Synteza genomów wirusowych stała się powszechnym narzędziem
- Odtwarzanie wirusów nieistniejących już w naturze

Grypa 1918

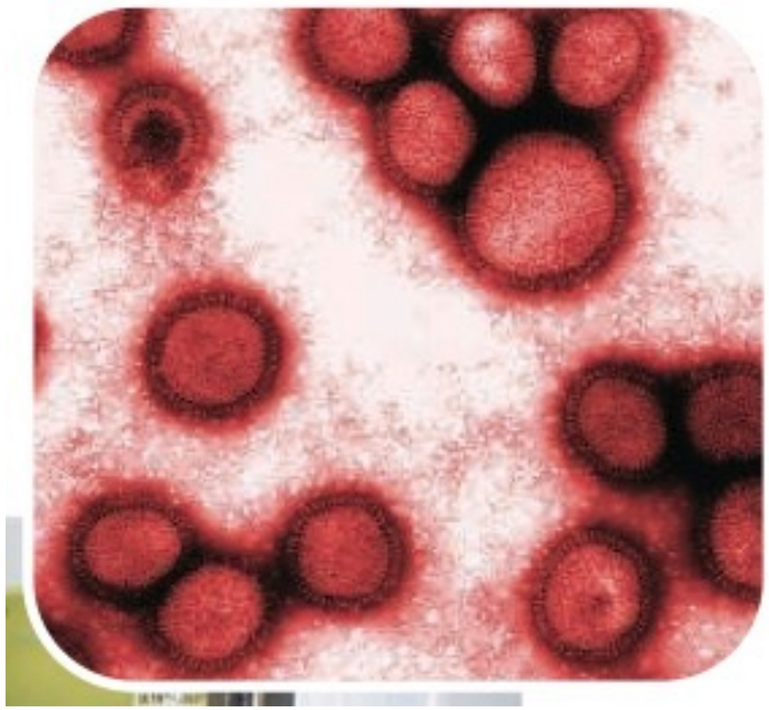
- Na podstawie sekwencji z tkanek ciał ofiar odtworzono sekwencję wirusa grypy epidemii 1918 (hiszpanki)
  - ~50 milionów ofiar śmiertelnych
- Odtworzono genom i uzyskano zdolne do infekcji wirusy (2005)

Characterization of the Reconstructed 1918 Spanish Influenza Pandemic Virus

Terrence M. Tumpey,[1*] Christopher F. Basler,[2] Patricia V. Aguilar,[2] Hui Zeng,[1] Alicia Solórzano,[2] David E. Swayne,[4] Nancy J. Cox,[1] Jacqueline M. Katz,[1] Jeffery K. Taubenberger,[3] Peter Palese,[2] Adolfo García-Sastre[2]

1918 – tajemnice zabójcy

- Odtworzony wirus jest bardzo zjadliwy
  - Infekuje i zabija myszy silniej niż współczesne wirusy grypy
  - Wiadomo, który gen (i która cząsteczka) wirusa odpowiada za taką wirulencję – hemaglutynina (HA)
  - Wiadomo, dlaczego ten wirus dobrze namnaża się nie tylko w komórkach płuc
  - Wiadomo, że może infekować ptaki i prawdopodobnie pochodzi od ptasiego wirusa



Grypa 1918 - bezpieczeństwo

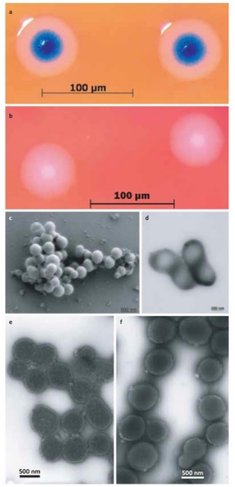
- Projekt wzbudził wiele kontrowersji
- Zagrożenie, ale i korzyści
  - Lepiej rozumiemy, dlaczego pewne szczepy wirusa są bardziej niebezpieczne od innych
  - Wielu ludzi wciąż ma częściową odporność na wirusa 1918
  - Wiadomo, które leki na niego działają
  - W razie problemu łatwo będzie można stworzyć szczepionkę

Pierwszy syntetyczny funkcjonujący genom

- 2010
- Syntetyczny genom *M. mycoides* JCVI-syn1.0 (~1 mln par zasad)
- Złożony z 1000 kaset po 1080 pz
- Składanie z pomocą drożdży

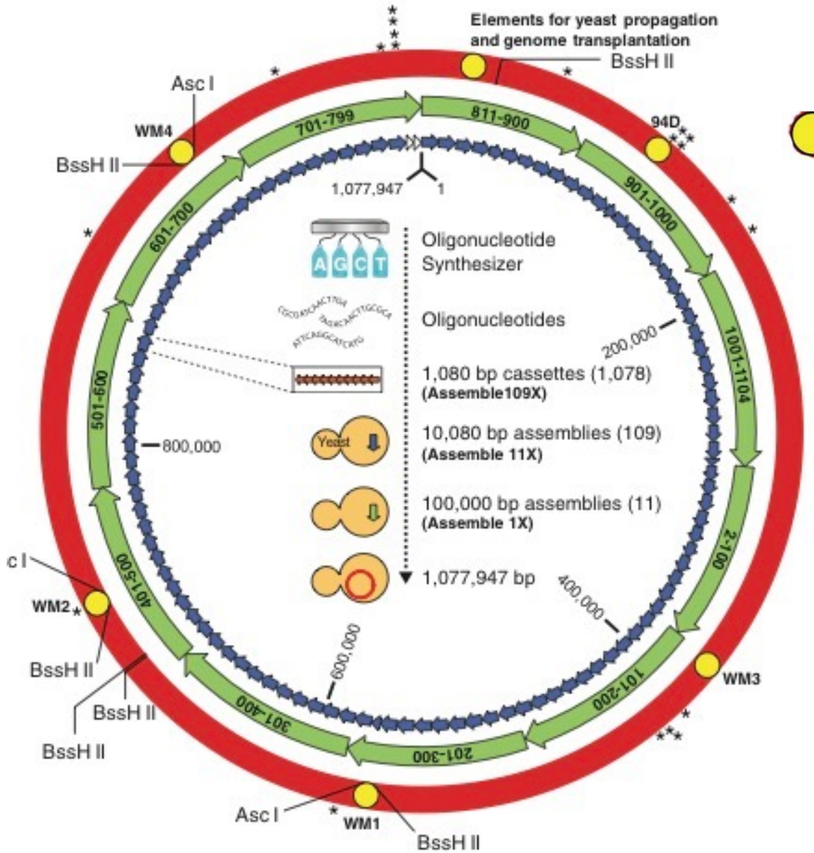
Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

Daniel G. Gibson,[1] John I. Glass,[1] Carole Lartigue,[1] Vladimir N. Noskov,[1] Ray-Yuan Chuang,[1] Mikkel A. Algire,[1] Gwynedd A. Benders,[2] Michael G. Montague,[1] Li Ma,[1] Monzia M. Moodie,[1] Chuck Merryman,[1] Sanjay Vashee,[1] Radha Krishnakumar,[1] Nacyra Assad-Garcia,[1] Cynthia Andrews-Pfannkoch,[1] Evgeniya A. Denisova,[1] Lei Young,[1] Zhi-Qing Qi,[1] Thomas H. Segall-Shapiro,[1] Christopher H. Calvey,[1] Prashanth P. Parmar,[1] Clyde A. Hutchison III,[2] Hamilton O. Smith,[2] J. Craig Venter[1,2]*



Gibson, D. G., J. I. Glass, C. Lartigue, V. N. Noskov, R.-Y. Chuang, M. A. Algire, G. A. Benders, M. G. Montague, L. Ma, M. M. Moodie, C. Merryman, S. Vashee, R. Krishnakumar, N. Assad-Garcia, C. Andrews-Pfannkoch, E. A. Denisova, L. Young, Z.-Q. Qi, T. H. Segall-Shapiro, C. H. Calvey, P. P. Parmar, C. A. Hutchison III, H. O. Smith, and J. C. Venter. 2010. Creation of a bacterial cell controlled by a chemically synthesized genome. Science, Published online May 20 2010.

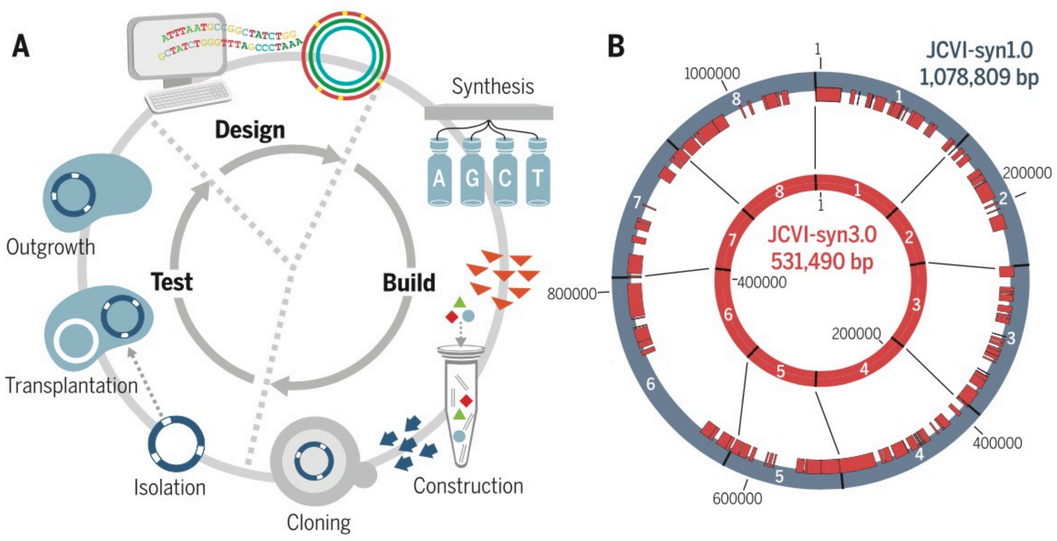
J. Craig Venter INSTITUTE

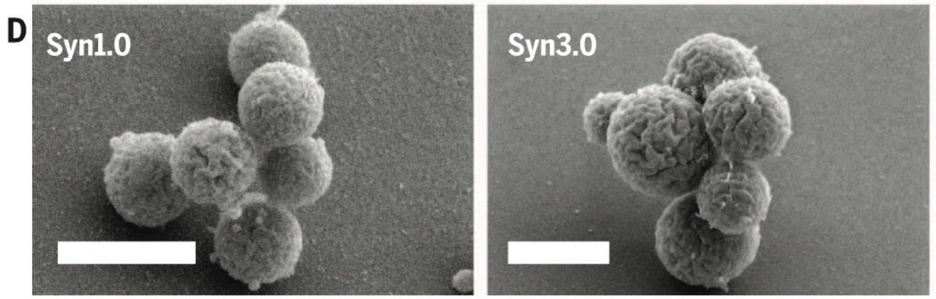


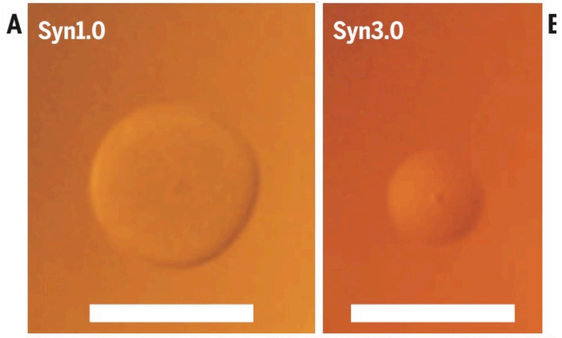
“znaki wodne”

Kolejny krok - syntetyczny genom minimalny

- Na podstawie JCVI-syn1.0
- Usunięte geny, które nie są niezbędne do życia (w warunkach laboratorium)
- 531 kb, 473 geny







Design and synthesis of a minimal bacterial genome

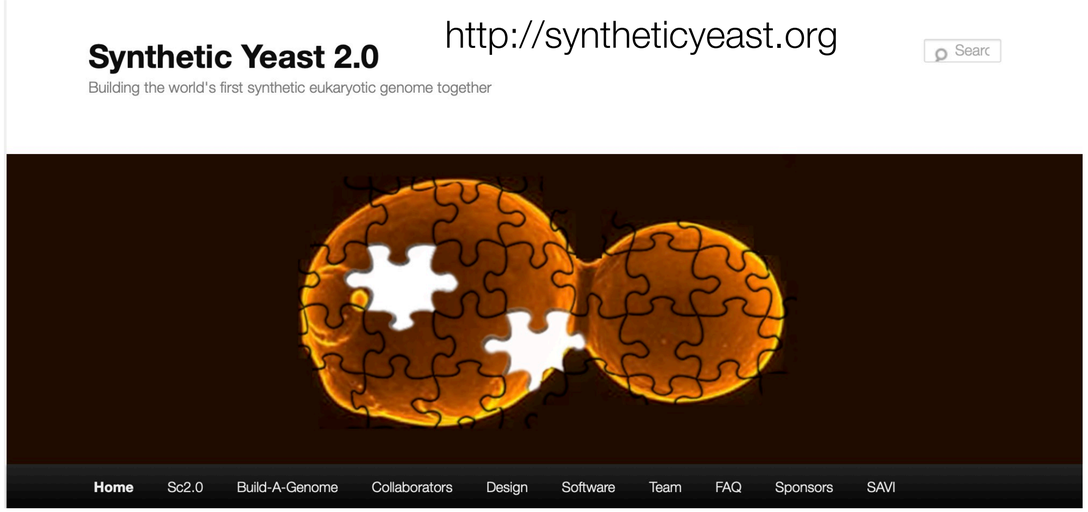
Clyde A. Hutchison III,*† Ray-Yuan Chuang,† Vladimir N. Noskov, Nacyra Assad-Garcia, Thomas J. Deerinck, Mark H. Ellisman, John Gill, Krishna Kannan, Bogumil J. Karas, Li Ma, James F. Pelletier, Zhi-Qing Qi, R. Alexander Richter, Elizabeth A. Strychalski, Lijie Sun, Yo Suzuki, Billyana Tsvetanova, Kim S. Wise, Hamilton O. Smith, John I. Glass, Chuck Merryman, Daniel G. Gibson, J. Craig Venter*

25 MARCH 2016 • VOL 351 ISSUE 6280

sciencemag.org SCIENCE

Drożdże 2.0 - syntetyczny genom eukariotyczny

- Projekt stworzenia syntetycznego genomu *S. cerevisiae*

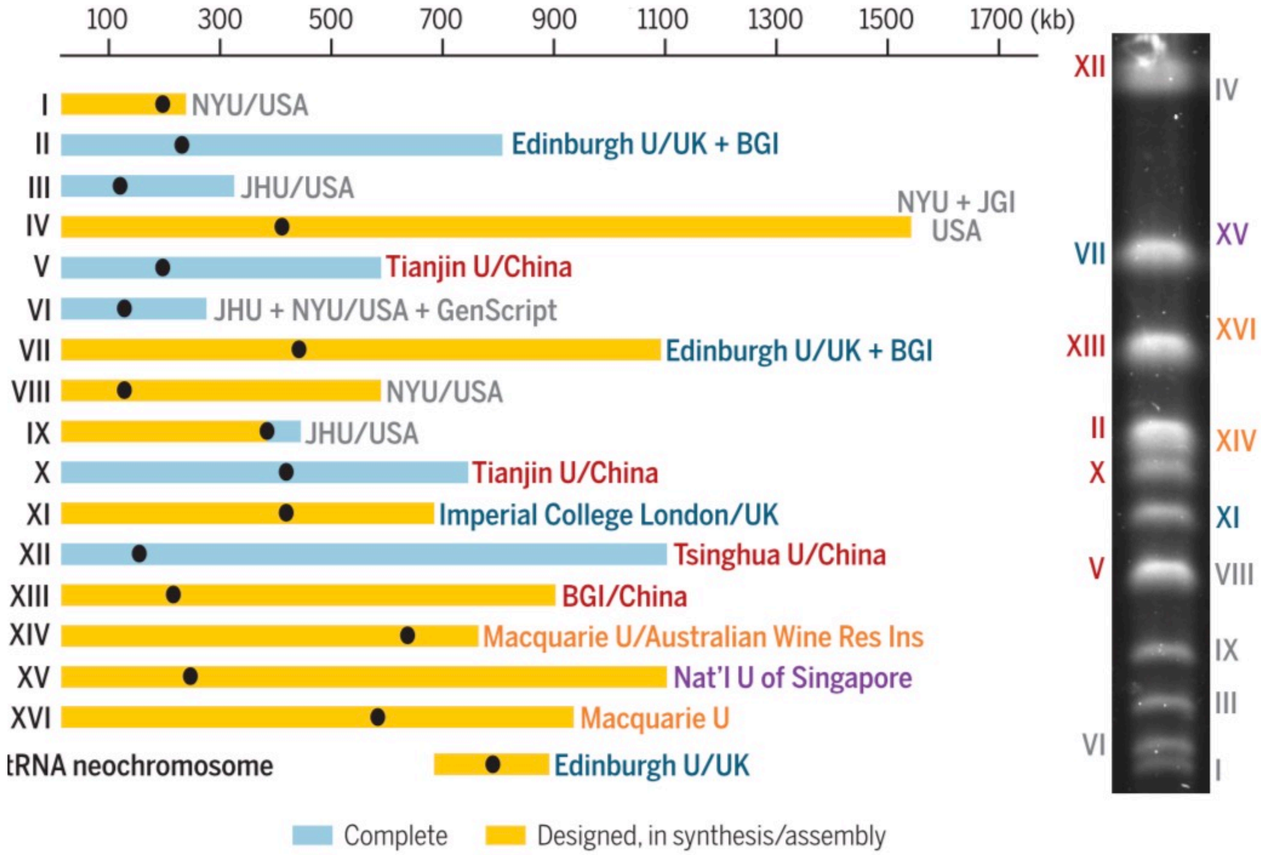


http://syntheticyeast.org

Design of a synthetic yeast genome

Sarah M. Richardson,[1,2]*‖ Leslie A. Mitchell,[2,3]‖ Giovanni Stracquadanio,[1,2,4]‖ Kun Yang,[1,2]‖ Jessica S. Dymond,[2]† James E. DiCarlo,[2]‡ Dongwon Lee,[1]§ Cheng Lai Victor Huang,[2] Srinivasan Chandrasegaran,[5] Yizhi Cai,[2,6] Jef D. Boeke,[2,3]# Joel S. Bader[1,2]#

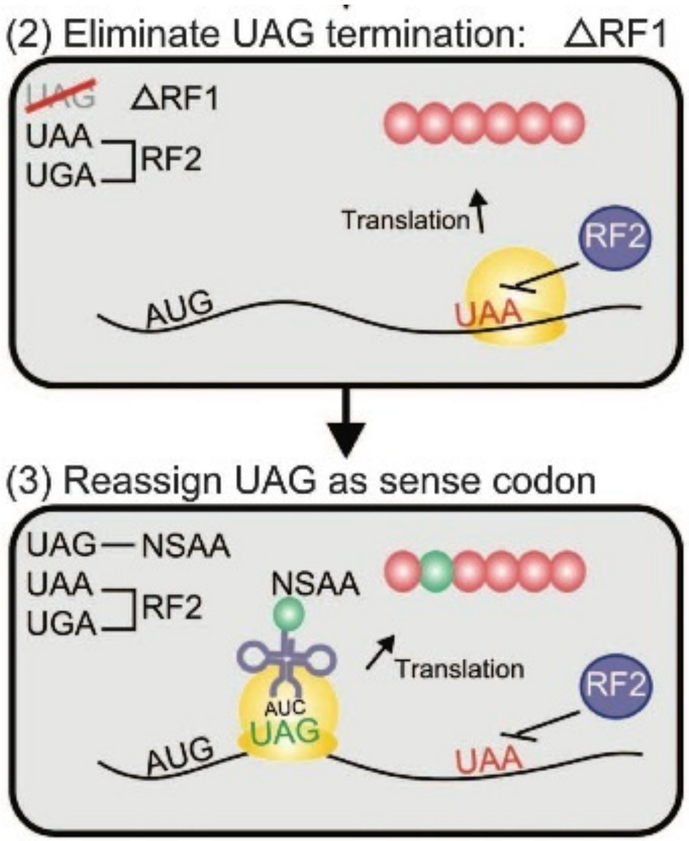
Richardson *et al.*, *Science* **355**, 1040–1044 (2017)



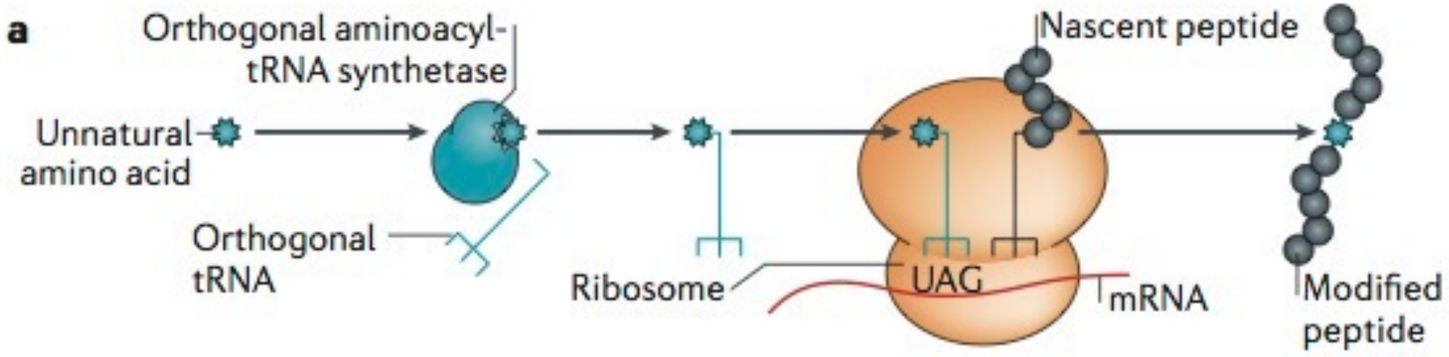
Inżynieria kodu genetycznego

- Zmiana kodonu stop na sensowny (może kodować niestandardowy aminokwas)

- Wprowadzenie równoległego kodu, np. czwórkowego, kodującego niestandradowe aminokwasy



Lajoie et al., 2013, Science 342: 357-342



Davis, L., and Chin, J.W. (2012). Nat Rev Mol Cell Biol 13, 168–182.

Podejście “od dołu” (bottom-up)

- Repertuar elementów i podstawowych obwodów
- Matematyczny model elementów
- Projektowanie i składanie systemów z elementów (cegiełek)

$$
\begin{cases}
\dfrac{\partial[mRNAcI]}{\partial t} = \dfrac{pcI\max transcription\,rate \times K^n_{mLacI}}{K^n_{mcI} + [LacProteinI]^n} - degradation\,rate[cImRNA] \\
\dfrac{\partial[ProteincI]}{\partial t} = \max translation\,rate\,[cImRNA] - degradation\,function \\
degradation\,function[cIProtein] \begin{cases} degradation\,rate\,[cIProtein] & if\,T < 40C \\ current\,cI\,protein\,concentration\,if\,T \geq 42C \end{cases}
\end{cases}
$$

When LacI dependent promoter is active:

$$
\begin{cases}
\dfrac{\partial[mRNALacI]}{\partial t} = \dfrac{pLac\max t\,ranscription\,rate\,pLac \times K^n_{mcI}}{K^n_{mcI} + [cIProtein\,]^n} - degradation\,rate[LacImRNA] \\
\dfrac{\partial[LacIProtein]}{\partial t} = \max translation\,rate\,[LacImRNA] - degradation\,rate[LacIProtein] \\
\qquad -kb[LacI][IPTG] + kd[IPTGLacIcomplex] \\
\dfrac{\partial[IPTG]}{\partial t} = -kb[LacI][IPTG] + kd[IPTGLacIcomplex] \\
\dfrac{\partial[IPTGLacIcomplex]}{\partial t} = kb[LacI][IPTG] - kd[IPTGLacIcomplex]
\end{cases}
$$

Obwody biologiczne

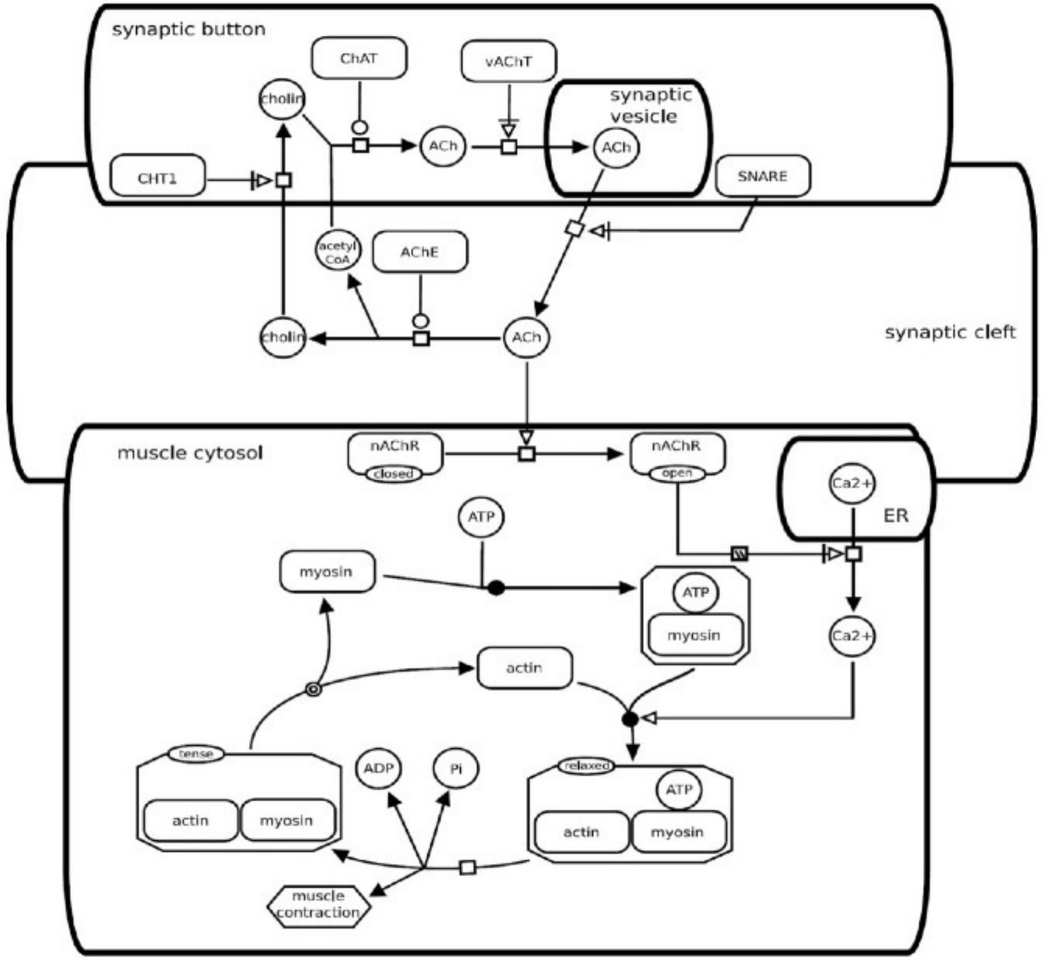
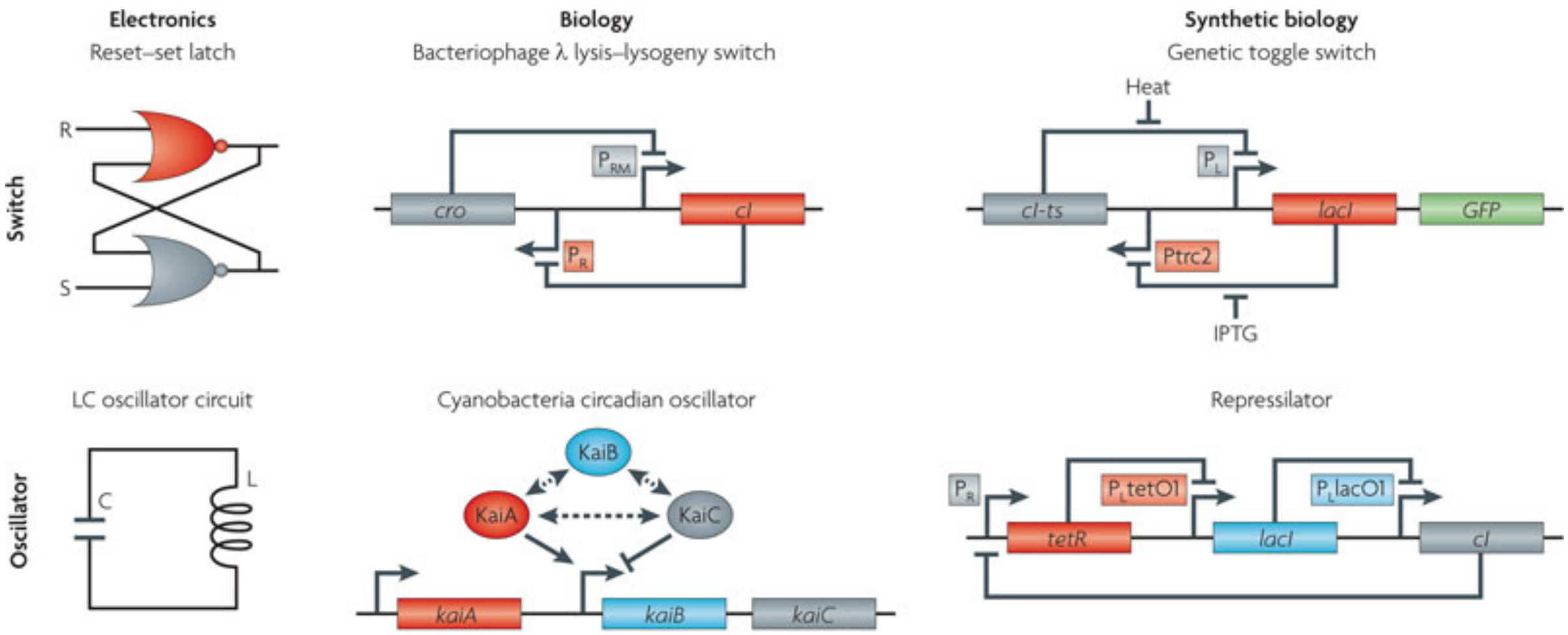


Figure 3: SBGN network example. of inter-cellular signaling near the neuromuscular junction [22, 23]. Biological

Alterovitz G, Muso T, Ramoni MF. The challenges of informatics in synthetic biology: from biomolecular networks to artificial organisms. Brief Bioinform. 2010 Jan;11(1):80-95

Metafora obwodu



Nature Reviews | Genetics

Biologia syntetyczna *open source*

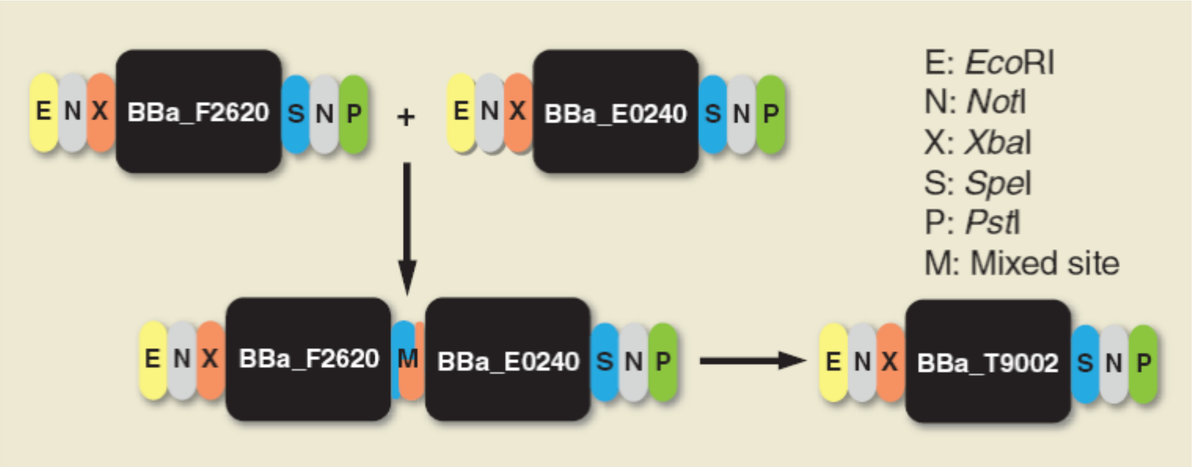
- Etos otwartości i wymiany informacji w biologii syntetycznej silniejszy, niż w tradycyjnej biologii molekularnej

- OpenWetware

Format BioBrick

BBa standard:

- prefix: EcoRI/NotI/XbaI
- suffix: BcuI/NotI/PstI



http://partsregistry.org/Main_Page

iGEM

- International Genetically Engineered Machine competition (iGEM)
- Studenci i doktoranci

Synthetic Biology
based on standard parts

“What I cannot build I cannot understand”

– Richard Feynman