

Podstawy ewolucji molekularnej

Ewolucja sekwencji DNA i białek

Ewolucja biologiczna

- Znaczenie ogólne:
 - proces zmian informacji genetycznej organizmów (częstości i rodzaju alleli w populacji),
 - które to zmiany są przekazywane z pokolenia na pokolenie
 - dotyczy populacji, nie pojedynczego osobnika
 - dotyczy zmian **dziedziczonych**
- Dziedziczenie z modyfikacją (*descent with modification* - Darwin)

Ewolucja biologiczna

- Zjawisko (fakt)
- Teoria ewolucji
- Historia zmian ewolucyjnych

Ewolucja biologiczna - teoria

- Teoria ewolucji
 - Wyjaśnienie mechanizmów zmian ewolucyjnych
 - darwinizm i neodarwinizm
 - Wyjaśnienie roli mechanizmów ewolucyjnych w kształtowaniu organizmów żywych
 - wspólne pochodzenie

Podstawy ewolucji

- Replikacja informacji genetycznej wprowadza zmienność
 - Losowe błędy w replikacji (nieuniknione)
- Wytworzone przez zmienność warianty nie są równocenne
 - Różne warianty mają różne dostosowanie (*fitness*) – różne prawdopodobieństwo przekazania informacji kolejnym pokoleniom w danych warunkach środowiska

Mechanizmy ewolucji

- Generujące zmienność
 - mutacje
 - rearanżacje genomu
 - horyzontalny transfer genów
- Działające na warianty wytworzone przez zmienność
 - dobór naturalny - nielosowy
 - dryf genetyczny - losowy

Główne założenie darwinizmu i neodarwinimu

- Podstawowym mechanizmem kształtującym proces ewolucji biologicznej jest dobór naturalny
 - dryf genetyczny i inne zjawiska też mają znaczenie
 - znaczenie doboru i dryfu jest różne na różnych poziomach zmian ewolucyjnych
 - na poziomie molekularnym (zmian sekwencji DNA i białek) dryf może być głównym mechanizmem zmian – teoria neutralna

Podobieństwo i homologia

- Homologia: podobieństwo wynikające ze wspólnego pochodzenia ewolucyjnego – cecha odziedziczona od wspólnego przodka
 - vs. homoplazja – podobieństwo powstałe niezależnie, nie odziedziczone po wspólnym przodku
- Homologia jest właściwością dyskretną, nie stopniuje się
 - cechy mogą być mniej lub bardziej podobne, ale albo są homologiczne, albo nie

Pierwsza synteza

- Ewolucja jako zmiany częstości alleli w populacji
- Mutacje jako źródło nowych alleli
- Dobór i dryf wpływają na utrwalanie się lub zanikanie alleli
 - Równowaga między dryfem a dobozem zależy od N_e , zwłaszcza wyraźnie przy małych wartościach s

Druga synteza - ewolucja molekularna

- Ewolucja jako zmiany sekwencji DNA (i białek)
- Mechanizmy ewolucji molekularnej
 - źródła zmienności - mutacje i rearanżacje DNA
 - mechanizmy kształtujące zmienność - dobór i dryf
- Początki - lata 60. i 70. XX w. (sekwencjonowanie DNA - 1977, wcześniej sekwencje peptydowe)

Zmiany genetyczne w ewolucji

- Mutacje
 - tworzą nowe allele genów
- Inwersje
 - zmieniają układ genów na chromosomach
 - mogą uniemożliwić rekombinację na danym odcinku i doprowadzić do utrwalenia haplotypu
- Duplikacje
 - dotyczą fragmentów DNA, w tym całych genów
 - lub całych chromosomów i całych genomów
 - główne źródło innowacji ewolucyjnej
- Transfer horyzontalny
 - w tym zdarzenia symbiotyczne

Mutacje

- Podstawienia (substytucje)
- Niewielkie delecje i insercje
 - niewielkie tzn. wpływające na sekwencję 1-2 genów

Substytucje

- Tranzycje zachodzą w naturze częściej od transwersji
- mimo tego, że możliwych transwersji jest więcej
- stosunek ts/tv od ~2 (nDNA) do ~15 (mtDNA człowieka)
- wyjątek – mtDNA roślin
- duże różnice ts/tv u różnych grup organizmów

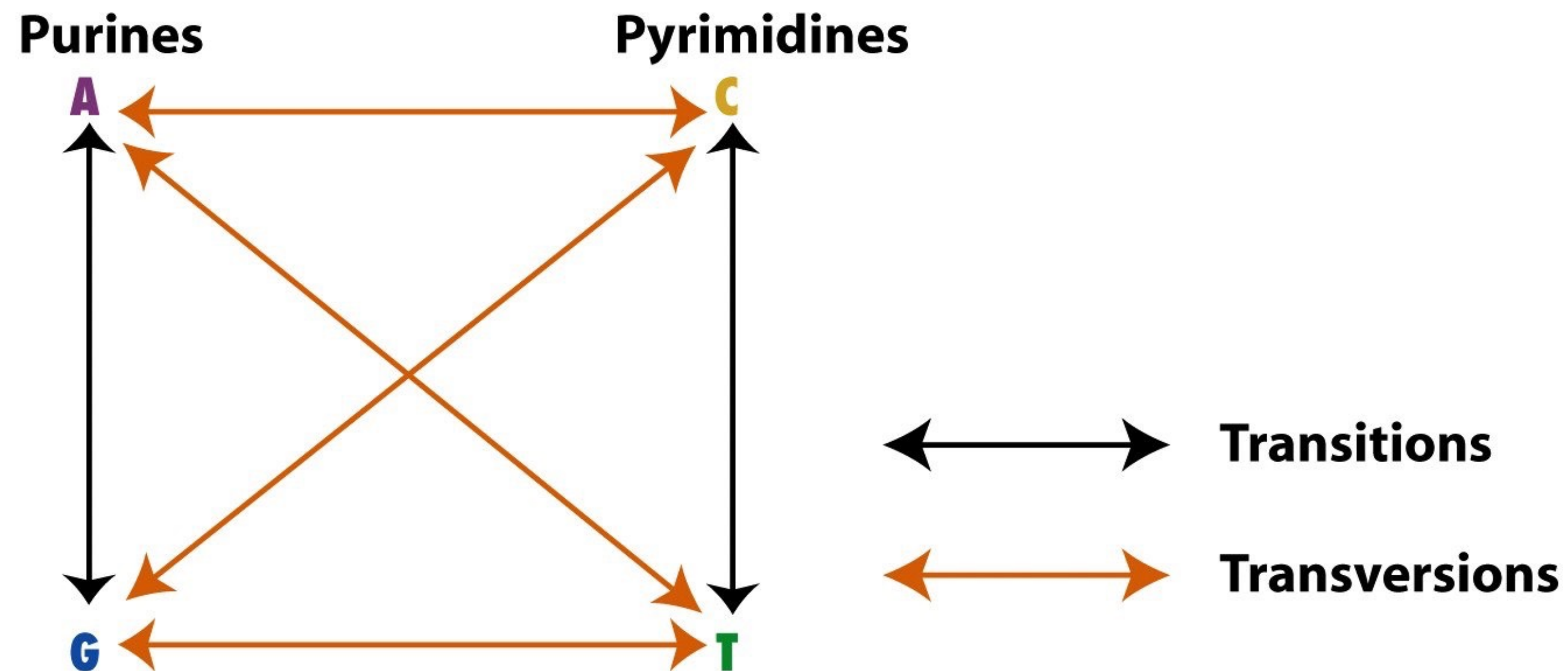


Figure 5-4 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

Tranzycje i transwersje

- Dlaczego tranzycje są częstsze?
 - Wyjaśnienia selekcyjne (tranzycje rzadziej zmieniają aminokwas i częściej są neutralne)
- Ale:
 - tranzycje są częstsze też w genach rRNA, pseudogenach i obszarach niekodujących
 - tranzycje są częstsze w pozycjach 4-krotnie zdegenerowanych (kodony typu CUN – Leu)

Tranzycje i transwersje

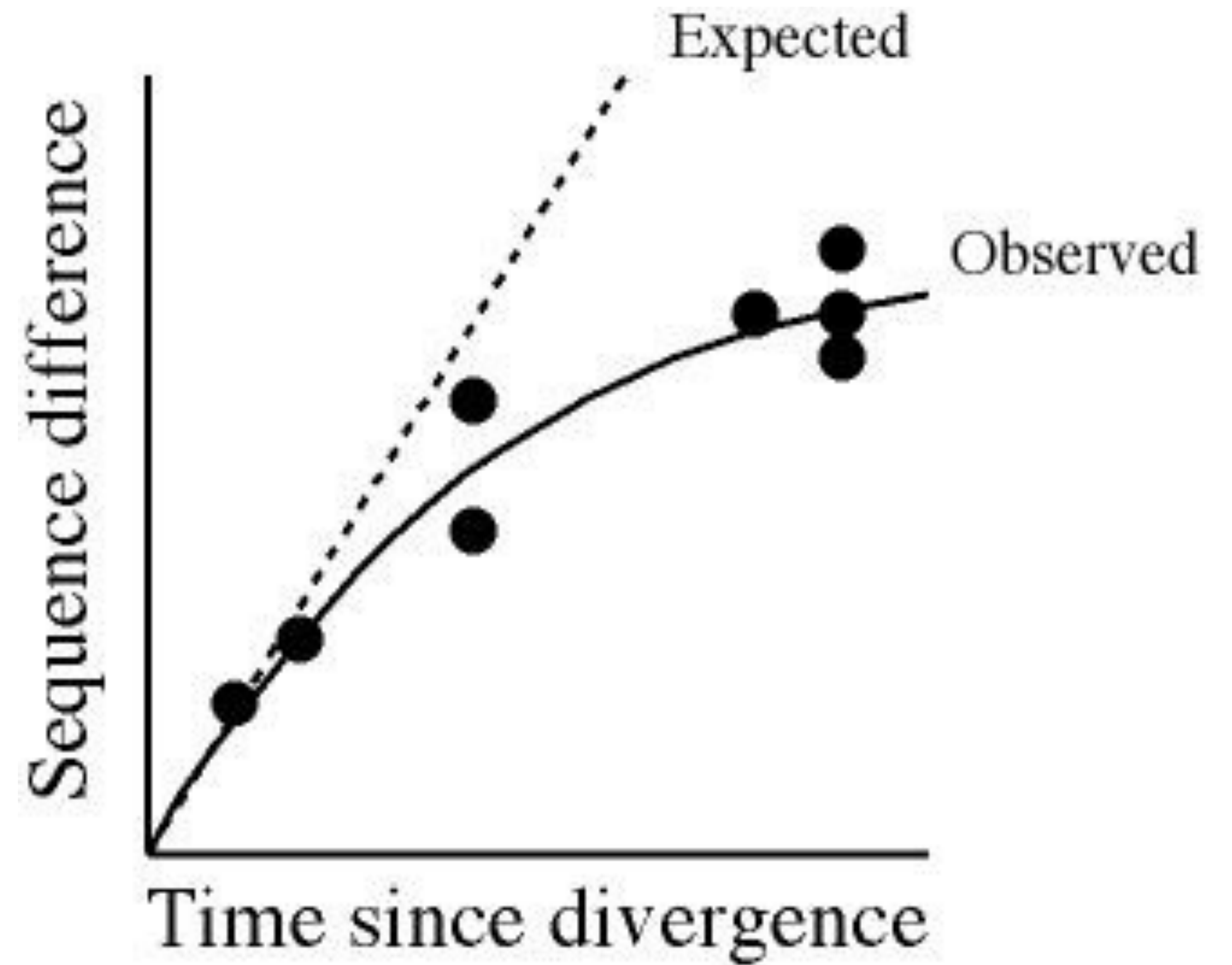
- Dlaczego tranzycje są częstsze?
 - Wyjaśnienia mechanistyczne – mechanizmy powstawania i naprawy mutacji
 - Tranzycje powstają w wyniku m. in.:
 - przejść tautomerycznych
 - deaminacji (np. oksydacyjnej)
 - Tranzycje w mniejszym stopniu zaburzają strukturę podwójnej helisy podczas replikacji
 - mniejsza wydajność naprawy przez system MMR

Modele ewolucji sekwencji

- Badając ewolucję nie dysponujemy z reguły sekwencją przodka
- Liczbę mutacji musimy oszacować na podstawie różnic między sekwencjami współczesnymi
- Konieczne jest uwzględnienie wielokrotnych mutacji w tej samej pozycji, zwłaszcza dla bardziej odległych sekwencji

Problem obliczania odległości (liczby zdarzeń mutacyjnych)

ACGGTGC
↓ ↓
C A
↓ ↓
GCGGTGA



Modele ewolucji sekwencji

- Modele Markova – stan w pokoleniu $n + 1$ zależy tylko od stanu w pokoleniu n i reguł przekształcenia (macierz prawdopodobieństw zmiany stanów)
- Modele o różnym stopniu skomplikowania
- Mogą uwzględniać:
 - mutacje wielokrotne w tej samej pozycji (poprawka Poissona)
 - różne prawdopodobieństwa zmian nukleotydowych (lub białkowych)
 - różne prawdopodobieństwo mutacji w różnych pozycjach sekwencji
 - różne częstości nukleotydów

Modele ewolucji DNA – model Jukesa-Cantora

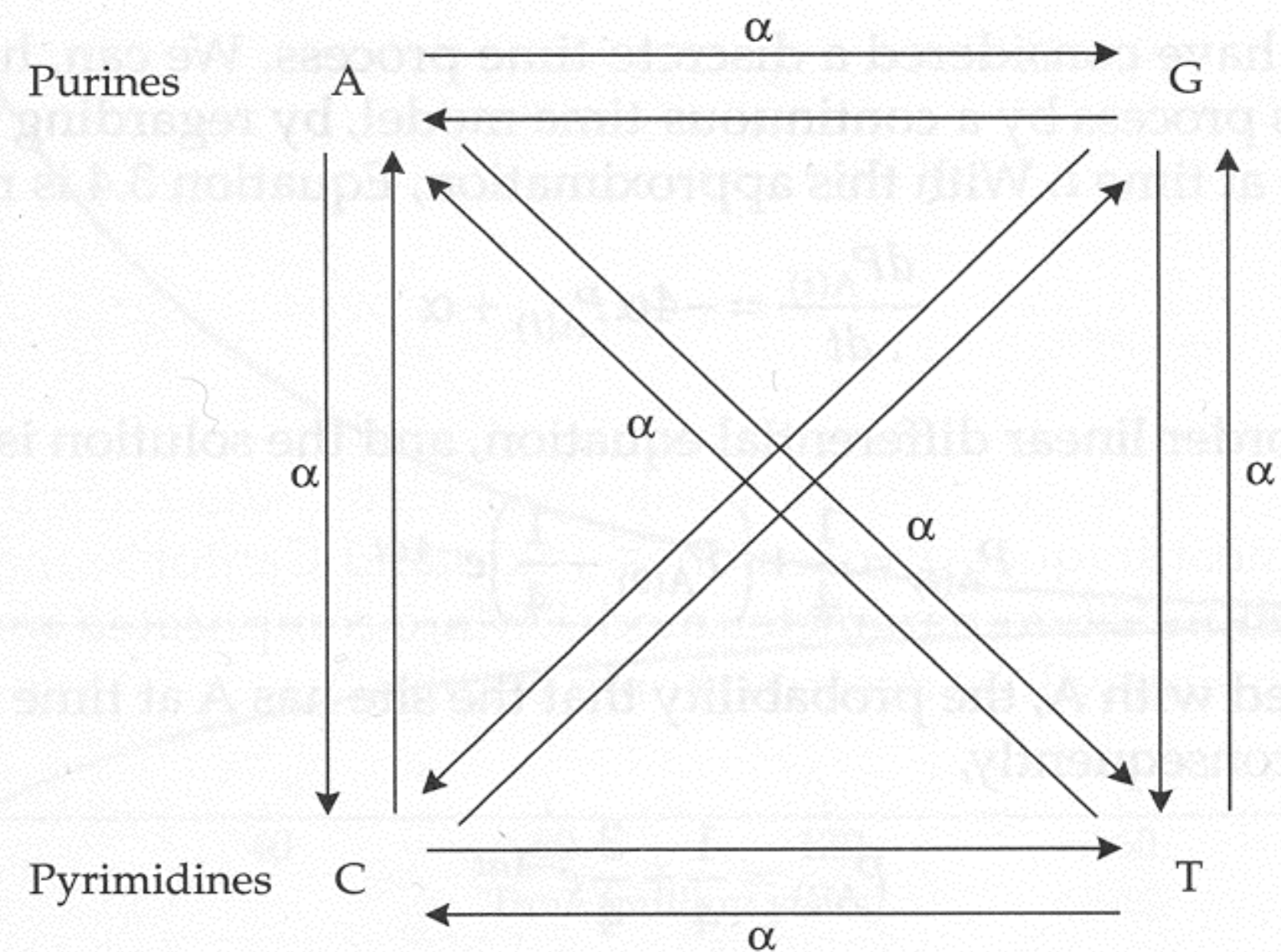


FIGURE 3.1 One-parameter model of nucleotide substitution. The rate of substitution in each direction is α .

| | A | C | G | T |
|---|--------------|--------------|--------------|--------------|
| A | 1-3 α | α | α | α |
| C | α | 1-3 α | α | α |
| G | α | α | 1-3 α | α |
| T | α | α | α | 1-3 α |

$$D_{JC} = -\frac{3}{4} \ln\left(1 - \frac{4}{3} D\right)$$

Inne modele

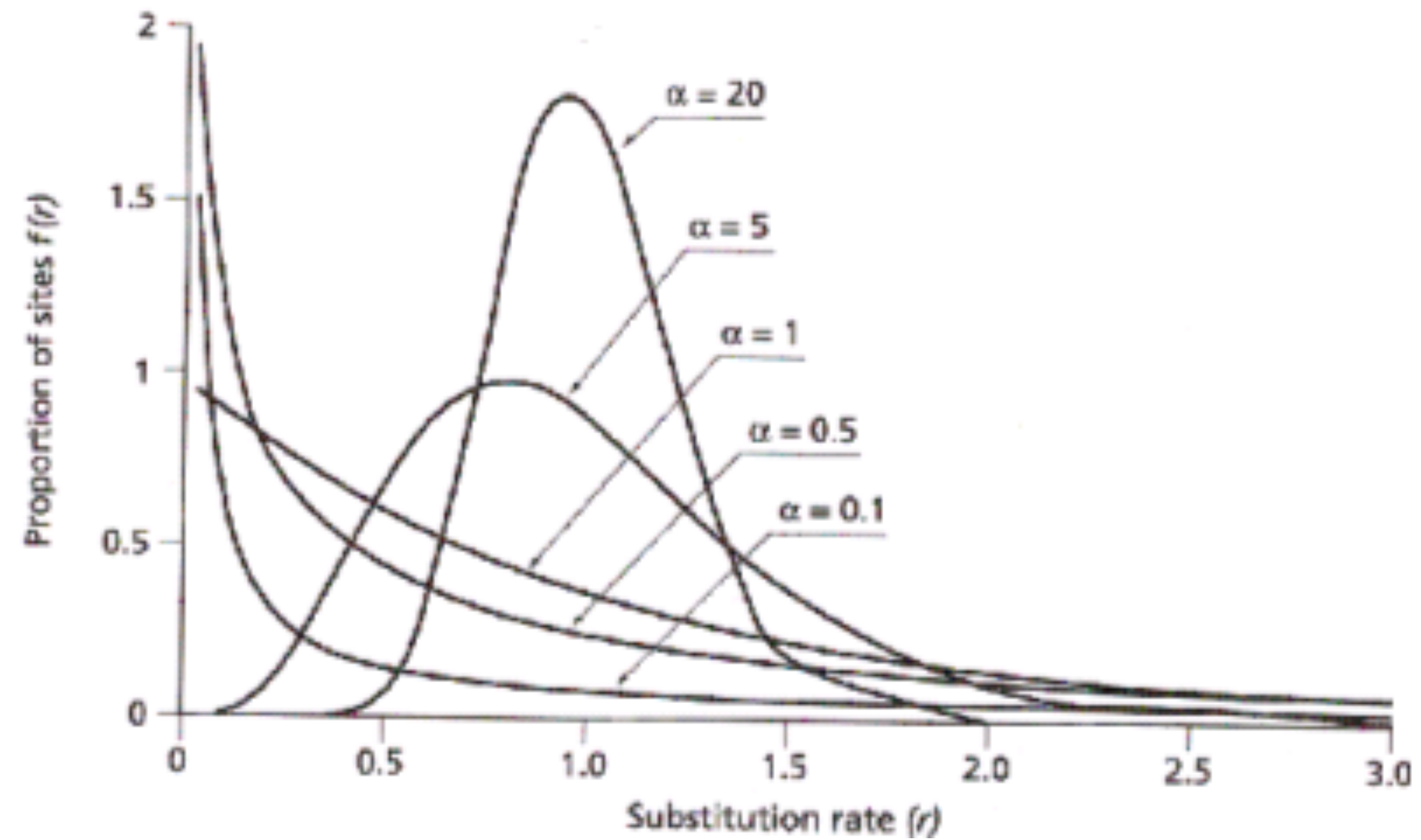
- Kimura (K80, dwuparametrowy) - różne prawdopodobieństwo tranzycji i transwersji
- Felsenstein (F81), Hasegawa-Kishino-Yano (HKY85) - różne częstości nukleotydów (F81) + różne prawd. tranzycji i transwersji (HKY85)
- GTR (General Time Reversible, Tavare '86)

Model GTR

- Różne prawdopodobieństwo każdej substytucji (ale symetrycznie, czyli np. $A \rightarrow T = T \rightarrow A$) - 6 parametrów
- Różne częstości nukleotydów - 4 parametry

Rozkład gamma

- Proste modele zakładają jednakowe prawdopodobieństwo zmiany w każdej pozycji - nierealistyczne
- Rozkład prawdopodobieństw zmian w różnych pozycjach – rozkład gamma

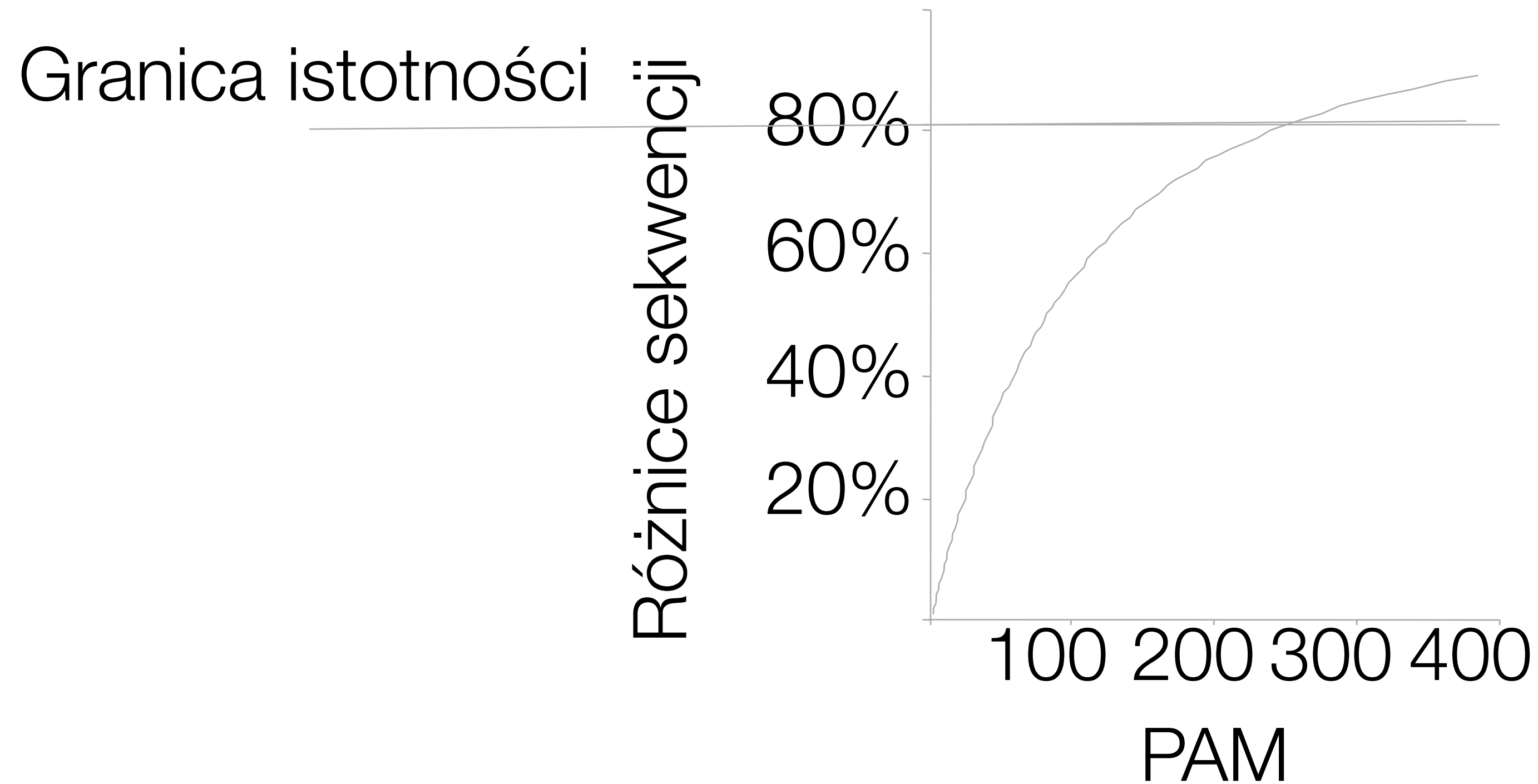


Ewolucja sekwencji aminokwasowych

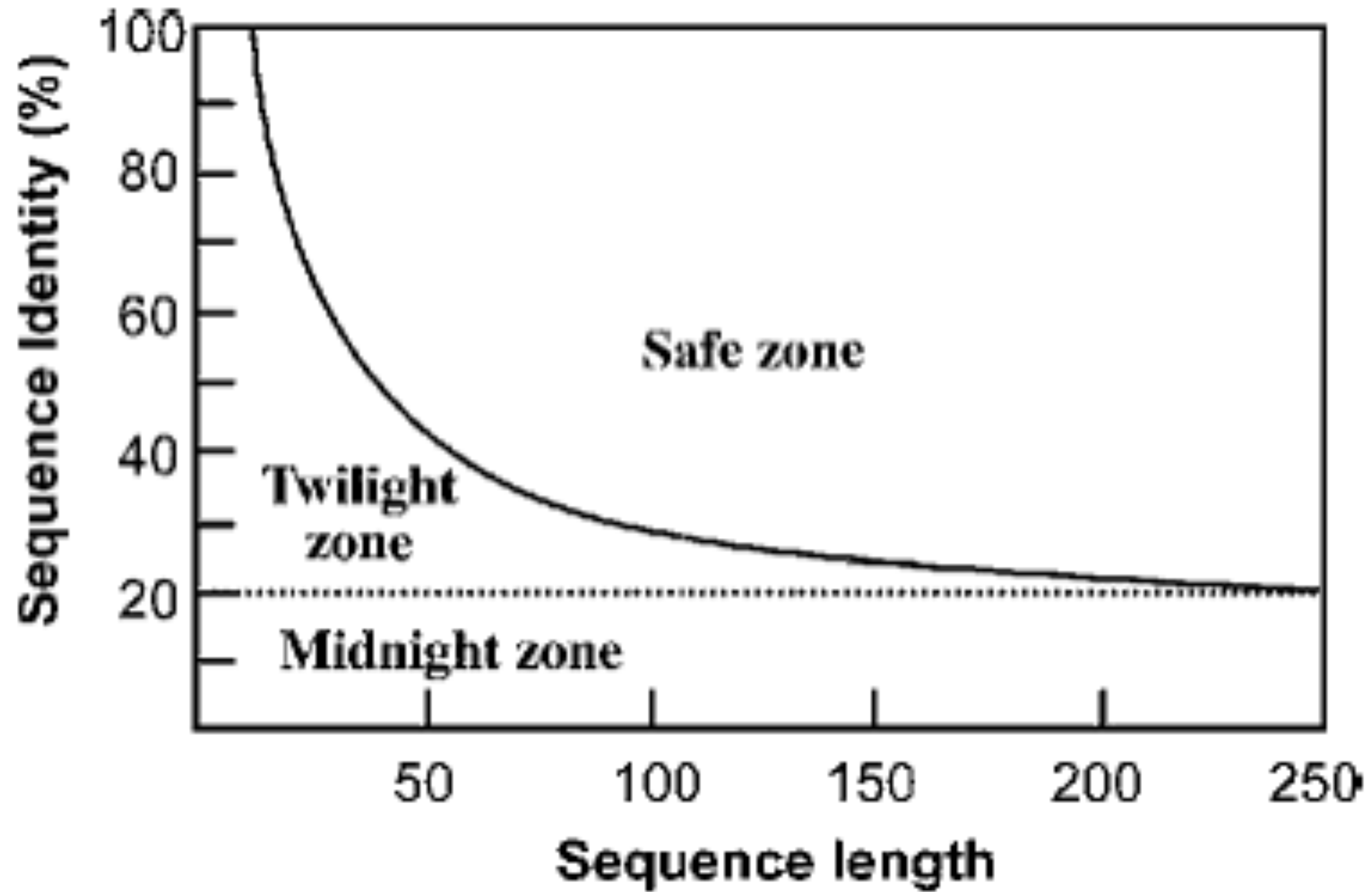
- Trudno stworzyć model analityczny
 - złożoność kodu
 - aminokwasy o różnych właściwościach - konieczna miara niepodobieństwa
- Stosuje się empirycznie uzyskiwane macierze prawdopodobieństwa zmiany danego aminokwasu w inny

Tempo zmian sekwencji białka

PAM - utrwalone mutacje punktowe/100 pozycji (od *Point Accepted Mutation*)



Istotność podobieństwa



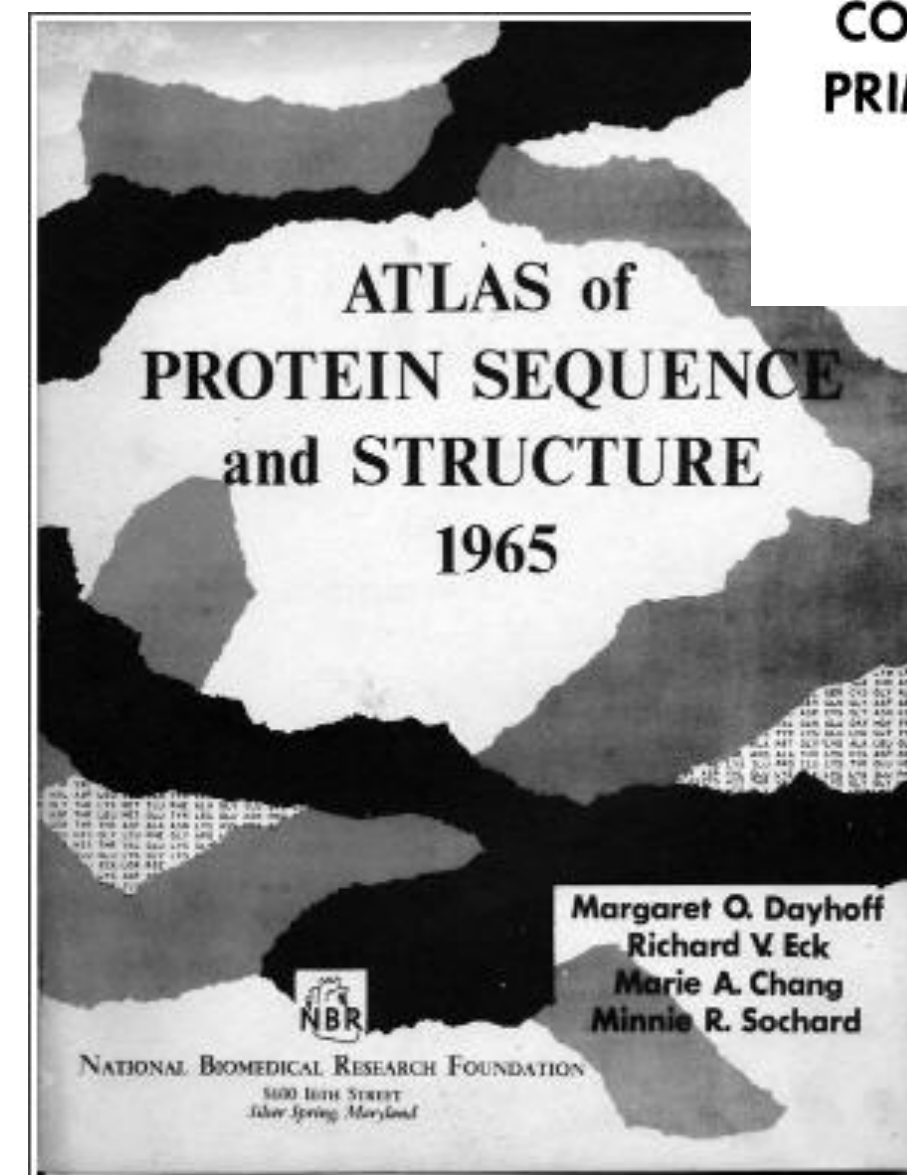
Porównywanie białek - macierze

- Macierze Dayhoff (PAM)
 - Na podstawie globalnych porównań sekwencji różniących się o 1 PAM ustalono prawdopodobieństwo zmiany każdego aminokwasu w inny → macierz PAM-1
 - Ekstrapolacja dla sekwencji bardziej odległych - mnożenie macierzy PAM-1 przez samą siebie odpowiednią liczbę razy → macierze PAM-20, PAM-40, PAM-250 itp. (proces Markova)

| | C | S | T | P | A | G | N | D | E | Q | H | R | K | M | I | L | V | F | Y | W | |
|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|----|----|---|
| C | 12 | | | | | | | | | | | | | | | | | | | | C |
| S | 0 | 2 | | | | | | | | | | | | | | | | | | | S |
| T | -2 | 1 | 3 | | | | | | | | | | | | | | | | | | T |
| P | -3 | 1 | 0 | 6 | | | | | | | | | | | | | | | | | P |
| A | -2 | 1 | 1 | 1 | 2 | | | | | | | | | | | | | | | | A |
| G | -3 | 1 | 0 | -1 | 1 | 3 | | | | | | | | | | | | | | | G |
| N | -4 | 1 | 0 | -1 | 0 | 0 | 2 | | | | | | | | | | | | | | N |
| D | -5 | 0 | 0 | -1 | 0 | 1 | 2 | 4 | | | | | | | | | | | | | D |
| E | -5 | 0 | 0 | -1 | 0 | 0 | 1 | 3 | 4 | | | | | | | | | | | | E |
| Q | -5 | -1 | -1 | 0 | 0 | -1 | 1 | 2 | 2 | 4 | | | | | | | | | | | Q |
| H | -3 | -1 | -1 | 0 | -1 | -2 | 2 | 1 | 1 | 3 | 6 | | | | | | | | | | H |
| R | -4 | 0 | -1 | 0 | -2 | -3 | 0 | -1 | -1 | 1 | 2 | 6 | | | | | | | | | R |
| K | -5 | 0 | 0 | -1 | -1 | -2 | 1 | 0 | 0 | 1 | 0 | 3 | 5 | | | | | | | | K |
| M | -5 | -2 | -1 | -2 | -1 | -3 | -2 | -3 | -2 | -1 | -2 | 0 | 0 | 6 | | | | | | | M |
| I | -2 | -1 | 0 | -2 | -1 | -3 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | 2 | 5 | | | | | | I |
| L | -6 | -3 | -2 | -3 | -2 | -4 | -3 | -4 | -3 | -2 | -2 | -3 | -3 | 4 | 2 | 6 | | | | | L |
| V | -2 | -1 | 0 | -1 | 0 | -1 | -2 | -2 | -2 | -2 | -2 | -2 | -2 | 2 | 4 | 2 | 4 | | | | V |
| F | -4 | -3 | -3 | -5 | -4 | -5 | -4 | -6 | -5 | -5 | -2 | -4 | -5 | 0 | 1 | 2 | -1 | 9 | | | F |
| Y | 0 | -3 | -3 | -5 | -3 | -5 | -2 | -4 | -4 | -4 | 0 | -4 | -4 | -2 | -1 | -1 | -2 | 7 | 10 | | Y |
| W | -8 | -2 | -5 | -6 | -6 | -7 | -4 | -7 | -7 | -5 | -3 | 2 | -3 | -4 | -5 | -2 | -6 | 0 | 0 | 17 | W |

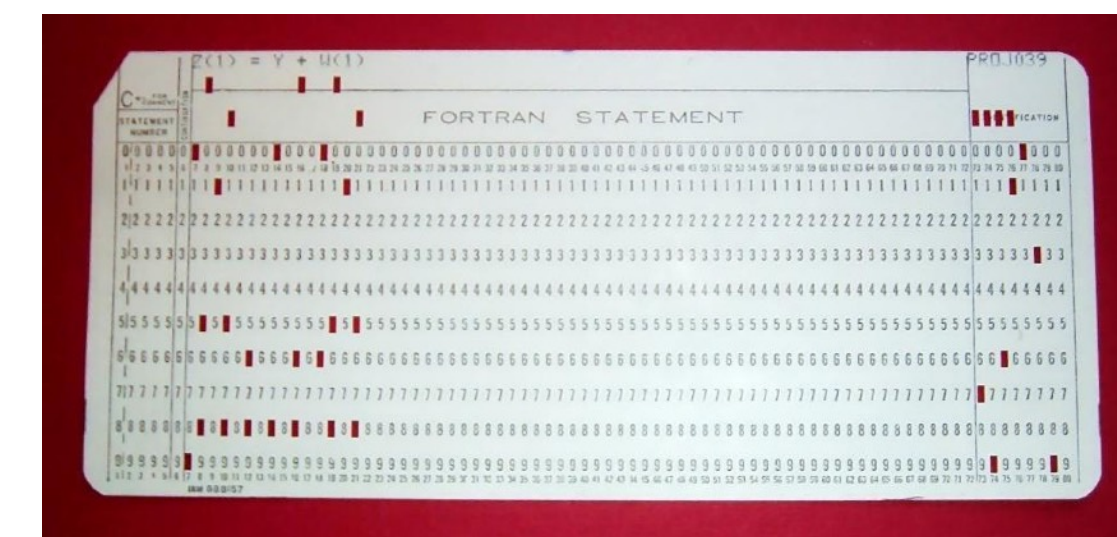


Margaret O. Dayhoff (1925-1983)



COMPROTEIN: A COMPUTER PROGRAM TO AID PRIMARY PROTEIN STRUCTURE DETERMINATION*

Margaret Oakley Dayhoff and Robert S. Ledley
National Biomedical Research Foundation
Silver Spring, Maryland



Porównywanie białek - macierze

- Macierze BLOSUM
 - Na podstawie prawdopodobieństwa zmiany każdego aminokwasu w inny w bloku lokalnego przyrównania sekwencji o n% identycznych aminokwasów (BLOSUM62 - 62% identycznych aa itp.)

| | A | C | D | E | F | G | H | → |
|---|----|----|----|----|----|----|----|---|
| A | 4 | 0 | -2 | -1 | -2 | 0 | -2 | |
| C | 0 | 9 | -3 | -4 | -2 | -3 | -3 | |
| D | -2 | -3 | 6 | 2 | -3 | -1 | -1 | |
| E | -1 | -4 | 2 | 5 | -3 | -2 | 0 | |
| F | -2 | -2 | -3 | -3 | 6 | -3 | | |
| G | 0 | -3 | -1 | -2 | -3 | | | |
| H | -2 | -3 | -1 | 0 | | | | |

BLOSUM 62

Mutacje na poziomie kodu

- Mutacje mogą:
 - zmienić kodon na inny, ale kodujący ten sam aminokwas
 - mutacje synonimiczne
 - zmienić kodon na kodujący inny aminokwas
 - mutacje niesynonimiczne
 - zmienić kodon na kodon STOP
 - mutacje nonsens
 - spowodować zmianę fazy odczytu
 - wpłynąć na ekspresję genu

Mutacje i dobór naturalny

- Efekty działania mutacji obserwujemy pośrednio
 - różnice sekwencji między populacjami (gatunkami)
 - polimorfizm sekwencji w obrębie populacji
- Na allele wytworzone przez mutacje może działać dobór
- Za zmiany częstości powstających alleli może odpowiadać dryf genetyczny
- Obserwujemy mutacje utrwalone całkowicie lub częściowo (polimorfizmy) w puli genowej

Podstawowe pytanie ewolucji molekularnej

- Jaka jest rola dryfu i doboru w wyjaśnieniu obserwowanego zróżnicowania sekwencji?
 - wewnątrzpopulacyjnego (polimorfizmy)
 - międzygatunkowego
- Pytanie dotyczy zróżnicowania ilościowego!
 - Nikt nie podaje w wątpliwość tego, że adaptacje w ewolucji powstają dzięki działaniu doboru!

Dobór czy dryf?

- **Selekcjonizm**

- większość utrwalonych mutacji została wyselekcjonowana przed dobór
- większość polimorfizmów jest utrzymywana przez dobór
 - dobór równoważący, naddominacja, dobór zależny od częstości

- **Neutralizm** (Kimura, 1968)

- większość utrwalonych mutacji została utrwalona przez dryf
- za większość polimorfizmów odpowiada dryf
- mutacje utrwalane przez dobór są rzadkie, nie mają wpływu na ilościową analizę zmienności molekularnej

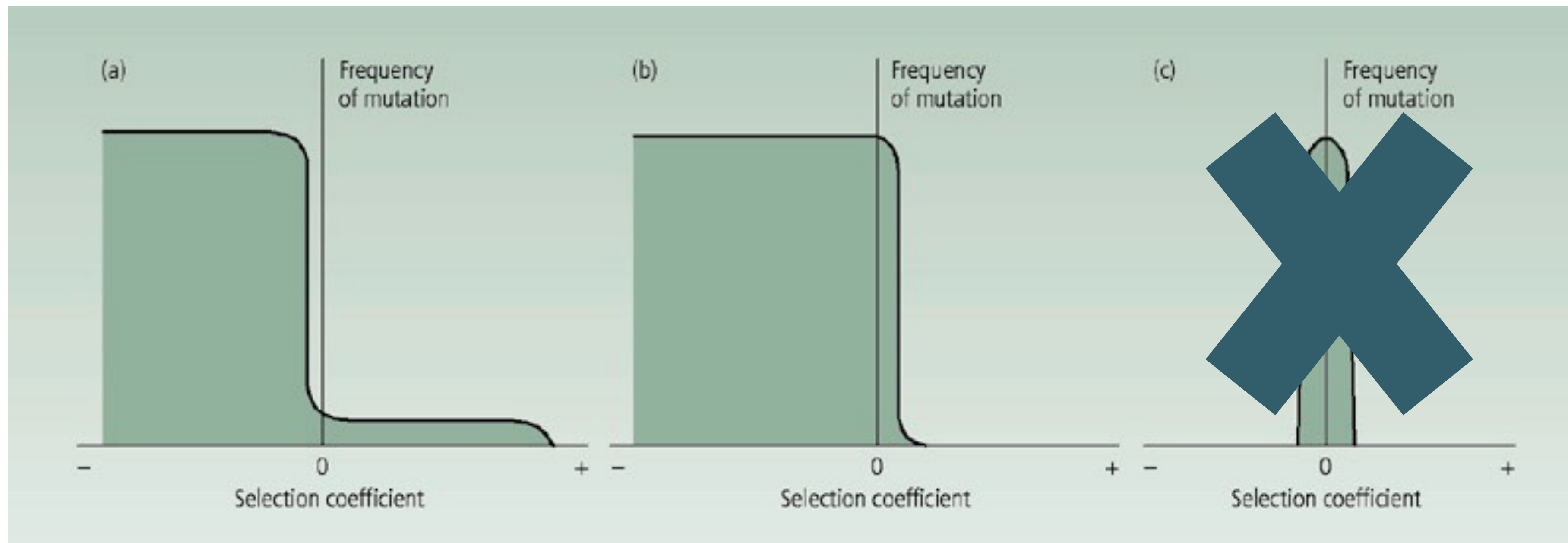
Mutacje i dobór

- **niekorzystne** (szkodliwe)
 - $s < 0$
 - eliminowane przez dobór (oczyszczający/negatywny)
- neutralne
 - $s \approx 0$ (a konkretniej, $s \leq 1/4N$)
 - utrwalane przez dryf
- **korzystne**
 - $s > 0$
 - utrwalane przez dobór (z udziałem dryfu dla niewielkich s)

Selekcjonizm i neutralizm

- Selekcjonizm:
 - większość mutacji jest niekorzystna lub korzystna
 - większość utrwalonych mutacji jest korzystna
 - mutacje neutralne są rzadkie (nie częstsze od korzystnych)
 - **dobór** jest głównym mechanizmem kształującym zmienność sekwencji
- Neutralizm
 - większość mutacji jest niekorzystna lub neutralna
 - większość utrwalonych mutacji jest neutralna
 - **dryf** jest głównym mechanizmem kształującym zmienność sekwencji
 - mutacje korzystne są rzadkie (znacznie rzadsze od neutralnych)

Selekcjonizm i neutralizm



selekcjonizm

neutralizm

pan-neutralizm

Neutralizm nie oznacza pan-neutralizmu, czyli negowania znaczenia selekcyjnego mutacji!

Przesłanki teorii neutralnej

- Tempo zmian sekwencji i polimorfizm są zbyt duże, by dały się wyjaśnić doborem
- Stałe tempo ewolucji molekularnej (zegar molekularny)
- Sekwencje o mniejszym znaczeniu funkcjonalnym (pseudogeny, mniej istotne obszary białek) ewoluują szybciej, niż obszary kluczowe dla funkcji

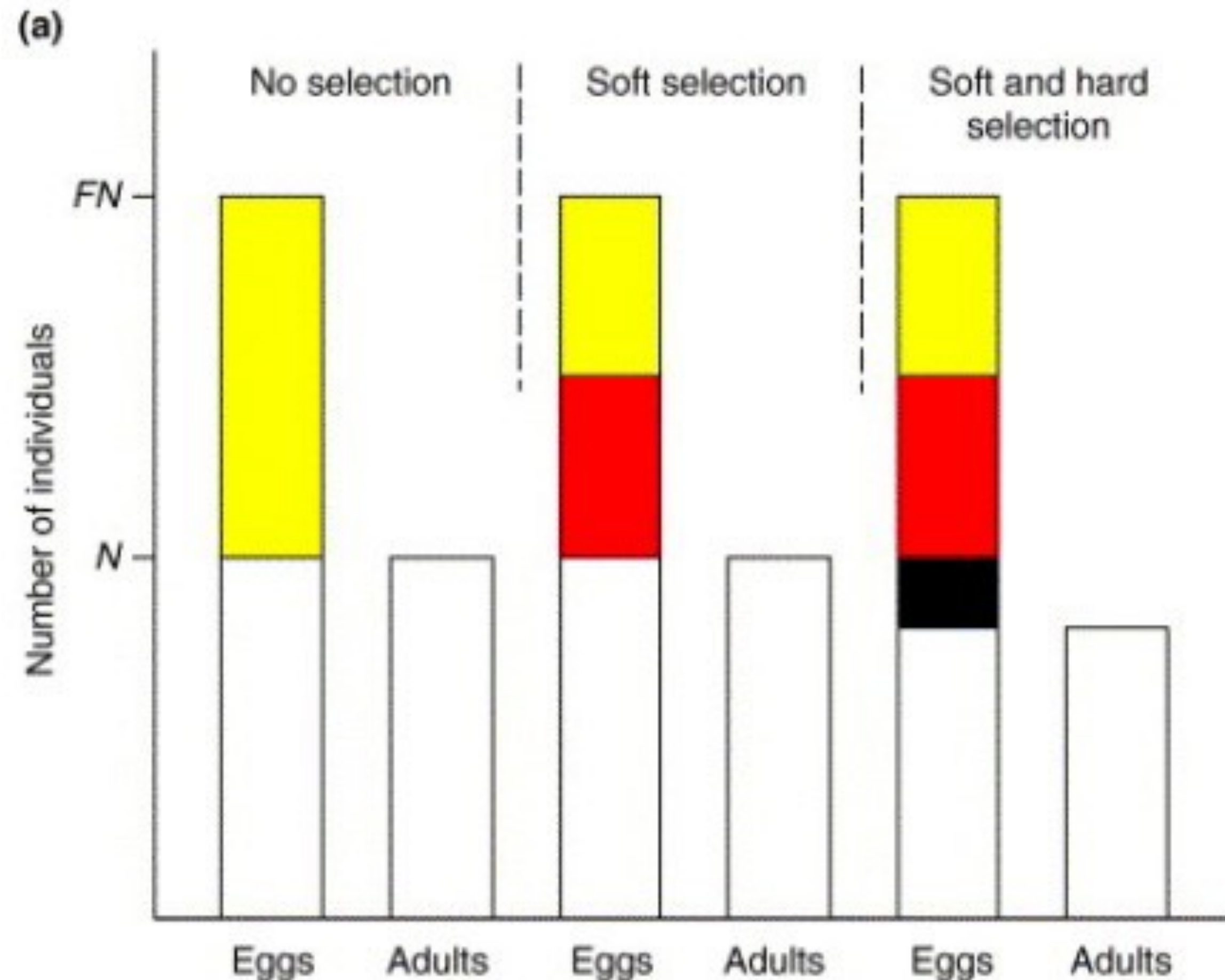
Tempo zmian

- Pojęcie obciążenia genetycznego – spadek średniego dostosowania populacji na skutek działania doboru
- Im silniej selekcjonowany nowy, korzystny allel, tym większy koszt u osobników go pozbawionych
- Silna selekcja to duży koszt dla populacji - spada średnie dostosowanie (*fitness*) populacji
- Haldane (1957) szacował maksymalne tempo ewolucji rzędu 1 mutacja na 300 pokoleń

Tempo zmian

- Rzeczywiste tempo zmian jest wyższe niż oszacowane przez Haldane'a
- Ale...
- Obliczenia Haldane'a i Kimury oparte były na założeniu tzw. twardego doboru
 - twardy dobór– zwiększona śmiertelność słabiej przystosowanych osobników, ponad typową (“ekologiczną”) śmiertelność w populacji
 - miękki dobór– działa w ramach stałej (“ekologicznej”) śmiertelności,
- Tylko dobór twardy znacząco ogranicza tempo mutacji

Dobór twardy i miękki



Dobór twardy –
więcej osobników
ginie (lub nie wydaje
potomstwa)

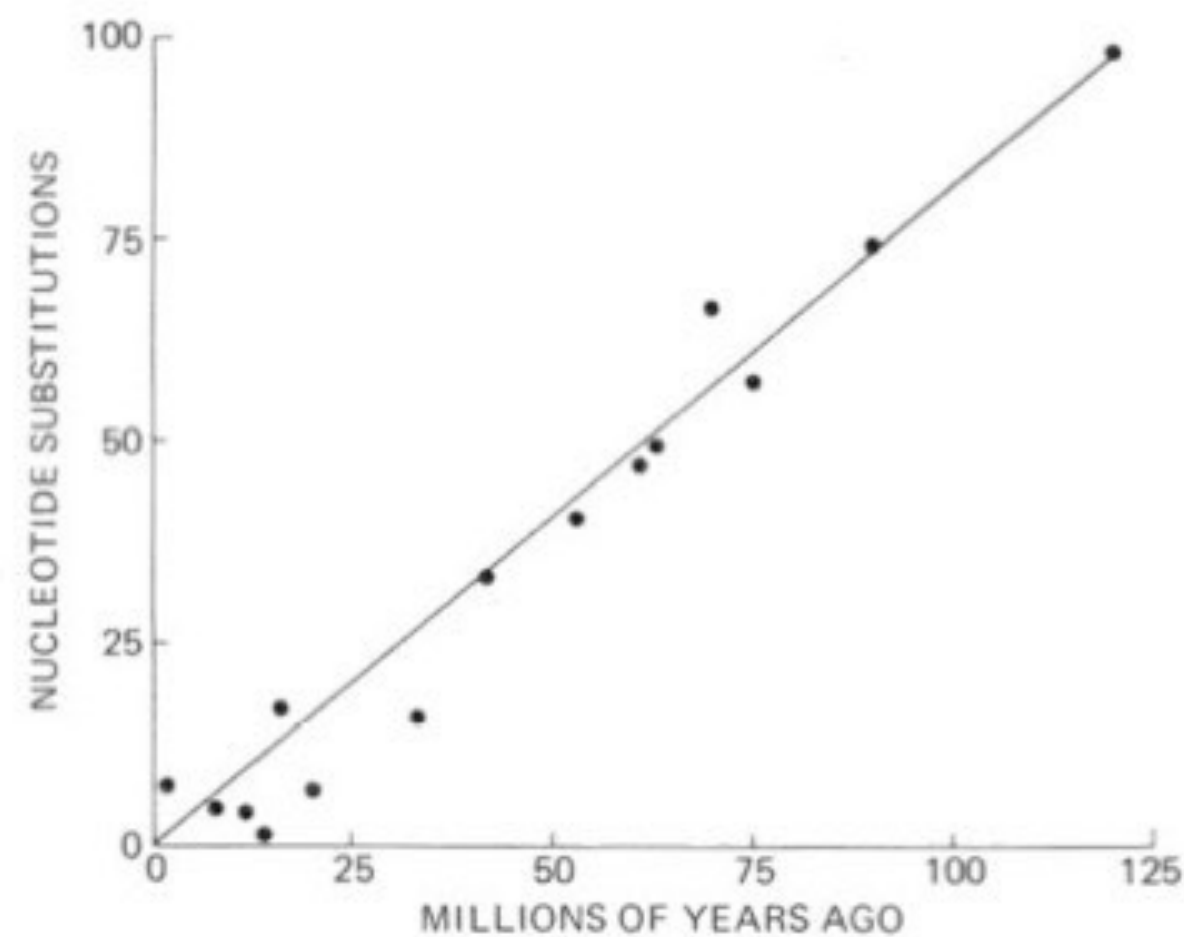
Dobór miękki – ginie
tyle samo
osobników, dobór
wpływa tylko na to,
które giną

Tempo zmian

- Obliczenia Haldane'a i Kimury oparte były na założeniu tzw. twardego doboru
 - twardy dobór– zwiększona śmiertelność słabiej przystosowanych osobników, ponad typową (“ekologiczną”) śmiertelność w populacji
 - miękki dobór– działa w ramach stałej (“ekologicznej”) śmiertelności,
- Tylko dobór twardy znacząco ogranicza tempo mutacji
- Dobór równoważący może utrzymywać zróżnicowanie i zwiększać tempo utrwalania mutacji

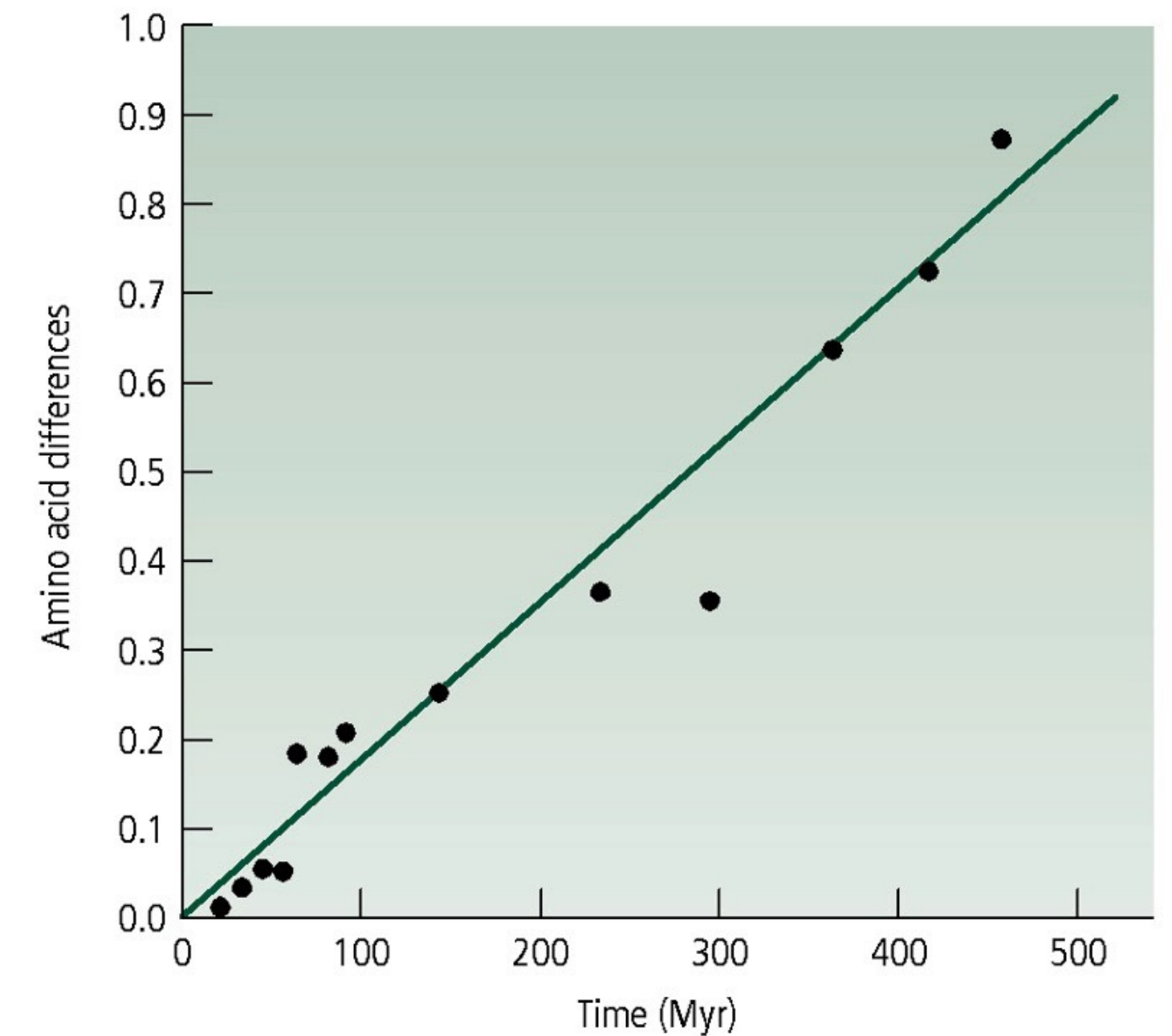
Stałe tempo ewolucji molekularnej

- Wiele sekwencji ewoluuje w stałym tempie
- Tempo to jest różne dla różnych sekwencji, ale stałe w czasie ewolucji dla danej sekwencji



Pairwise nucleotide differences among 17 mammals from 7 proteins, plotted against date of divergence as estimated from fossil record

- Tzw. zegar molekularny



Różnice sekwencji globin kręgowców

Tempo ewolucji i dryf

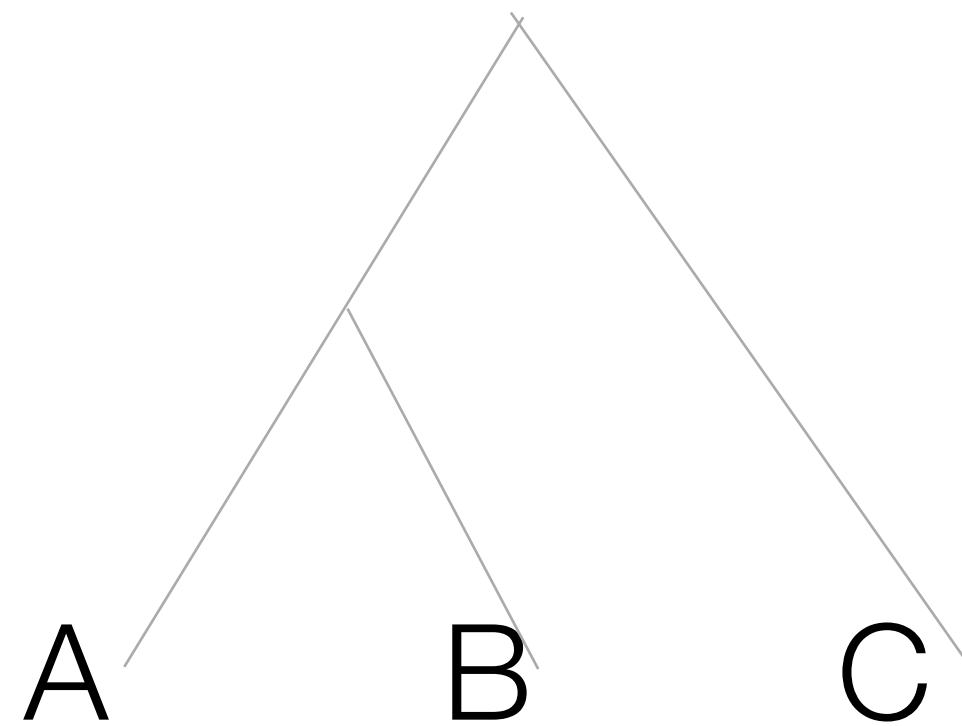
- Neutralny dryf jest procesem losowym, ale jego tempo będzie stałe w odpowiednio długim czasie
- Zależy tylko od częstości mutacji (jedna zmiana na $1/\mu$ pokoleń)

$$2N\mu \cdot \frac{1}{2N} = \mu$$

- Dla doboru stałe tempo zmian oznacza stałe tempo zmian środowiska
- Tempo zmian adaptacyjnych nie wydaje się być stałe

Zegar molekularny

- Jest konsekwencją neutralnego modelu ewolucji
- Tempo akumulacji zmian w danej sekwencji jest stałe
 - ale różne dla różnych sekwencji
- Weryfikacja – test względnego tempa



$$K_{AC} - K_{BC} = 0$$

- W rzeczywistości testuje stałość tempa pomiędzy gałęziami, ale nie w czasie

Zegar molekularny - problem

- W modelu neutralnym tempo utrwalania mutacji:

$$2N\mu \frac{1}{2N} = \mu$$

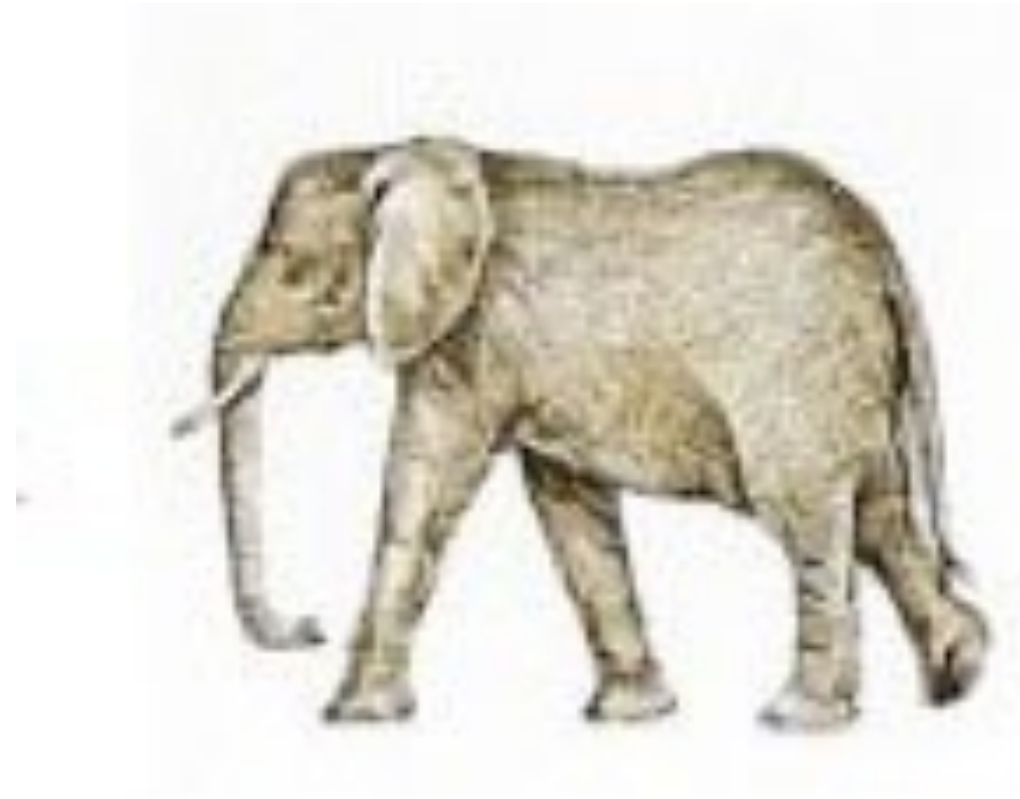
- Powinno być stałe w przeliczeniu na pokolenie
- Czas generacji jest różny u różnych organizmów
- Czyli nie powinna być obserwowana stałość tempa w czasie rzeczywistym

- A często jest (w tych sekwencjach, które zachowują zegar)

Problem czasu generacji

- Czas generacji różnych organizmów jest istotnie różny
- Dlaczego nie wpływa to na tempo utrwalania mutacji?

~0,03 pokolenia/rok



~3 pokolenia/rok



Zmiany prawie neutralne

- Model Kimury dotyczy zmian neutralnych ($s = 0$), takie nie są (w sekwencji białek) częste
- Mutacje zachowują się jak neutralne gdy spełnione jest:

$$|s| \leq \frac{1}{4N_e}$$

- Mutacje o niewielkim współczynniku doboru s będą zachowywały się jak neutralne w małych populacjach, a w większych populacjach będą podlegały doborowi

Zmiany prawie neutralne

- Istnieje odwrotna korelacja między czasem generacji a wielkością populacji

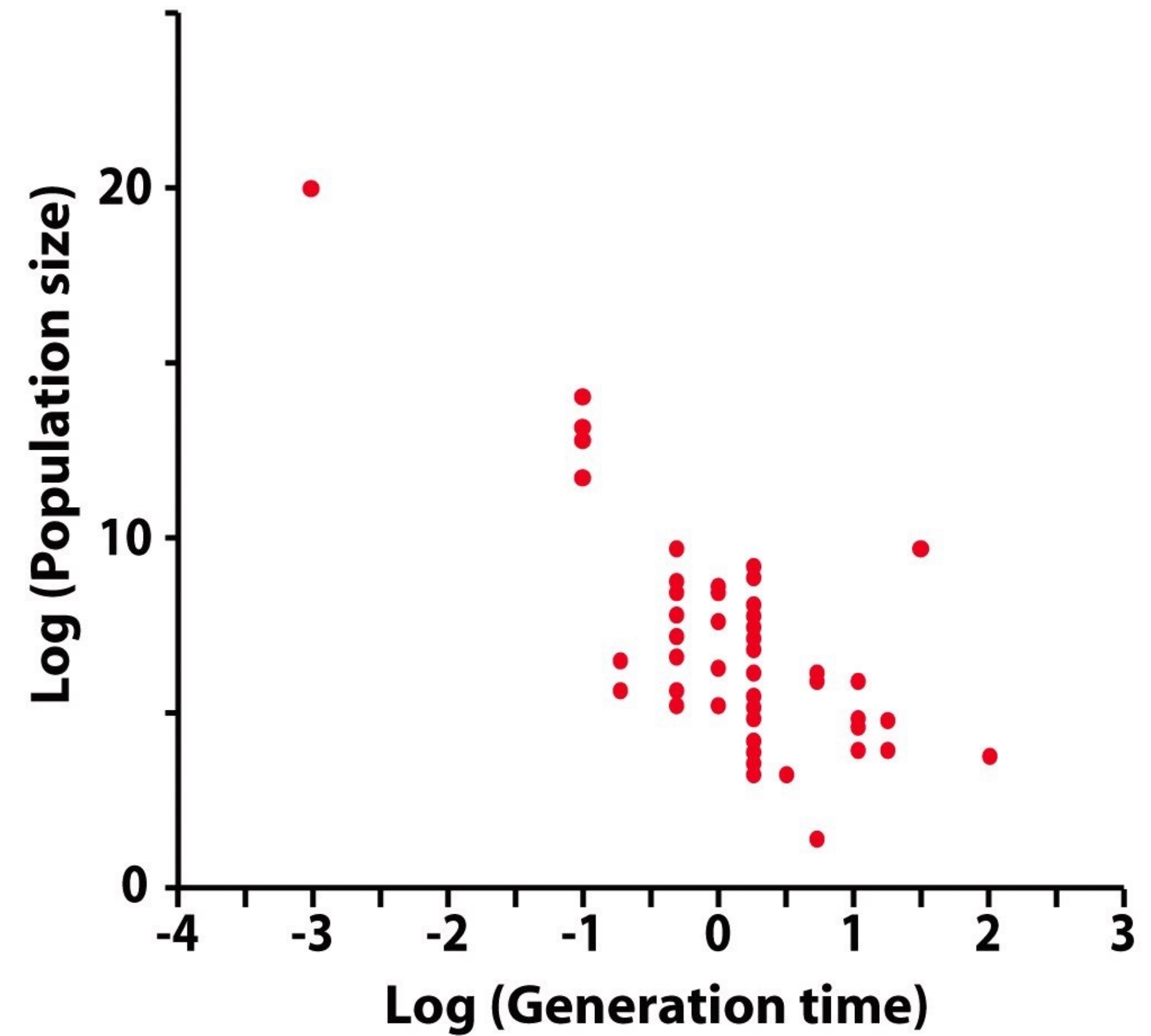
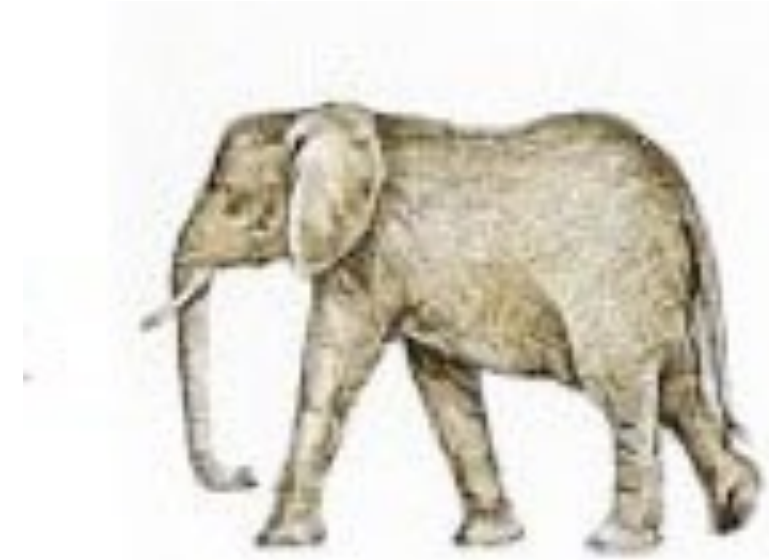


Figure 7-22a Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

Zmiany prawie neutralne

~0,03 pokolenia/rok



$$|s| \leq \frac{1}{4N_e}$$



~3 pokolenia/rok

Długi czas generacji

Krótki czas generacji

Mniej mutacji na rok

Więcej mutacji na rok

Populacja nieliczna (małe N_e)

Populacja liczna (duże N_e)

Więcej mutacji zachowuje się jak neutralne i utrwała przez dryf

Więcej mutacji podlega doborowi (i jest eliminowane przez dobór oczyszczający)

Efekty czasu generacji i wielkości populacji się znoszą, dając stałe tempo w czasie (Ohta & Kimura, 1971).

Zegar molekularny

- Dla sekwencji białek i zmian niesynonimicznych w DNA zmiany jednostajne w czasie
- Na poziomie DNA,
 - dla mutacji synonimicznych
 - pseudogenów
 - niektórych sekwencji niekodujących
- tempo ewolucji zależy od czasu generacji

Degeneracja w kodzie

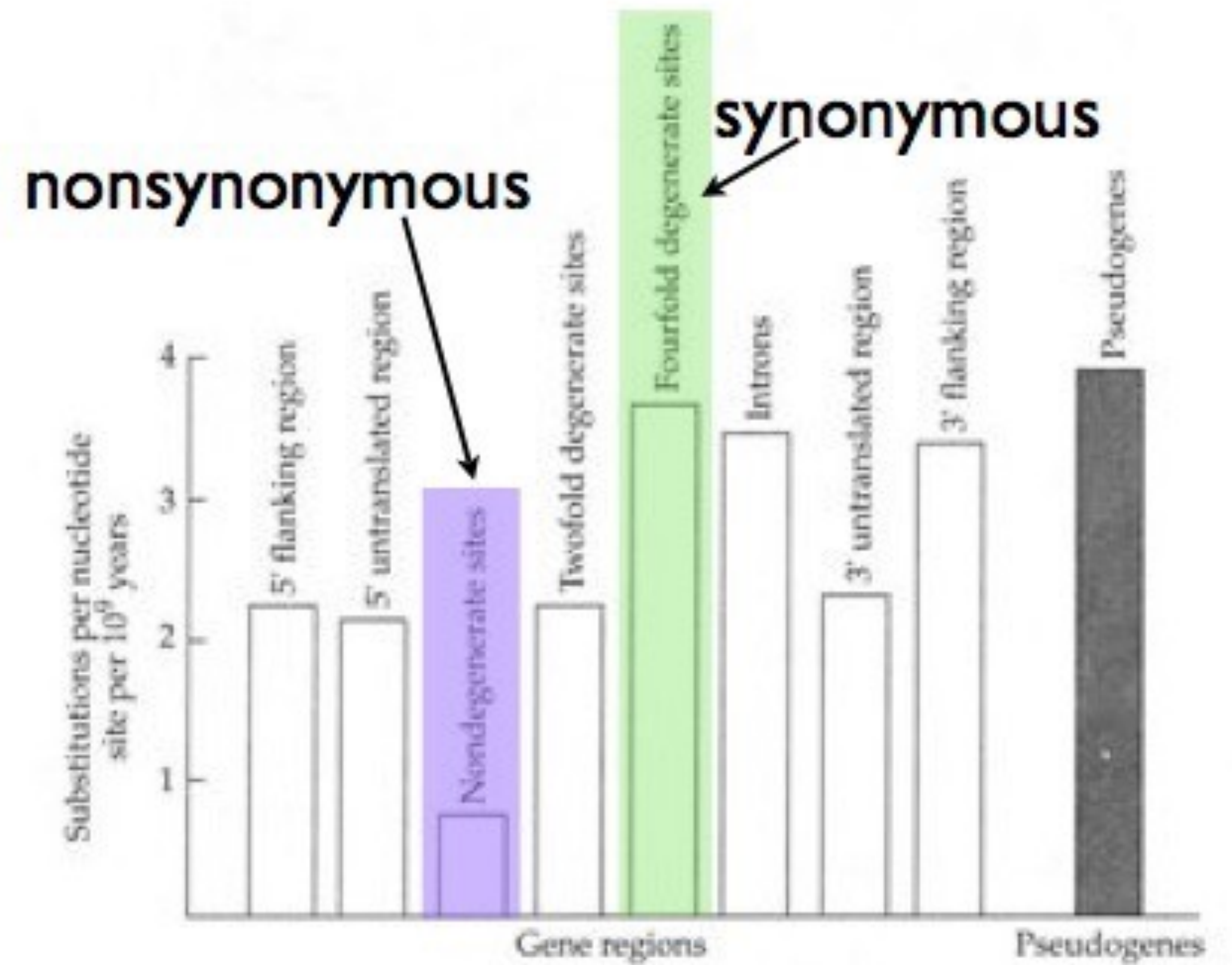
| | | Second Letter | | | | | |
|--------------|---|--|--------------------------------------|--|---|--------------|------------------|
| | | T | C | A | G | | |
| First Letter | T | TTT } Phe TTC } TTA } Leu TTG } | TCT } TCC } Ser TCA } TCG } | TAT } Tyr TAC } TAA Stop TAG Stop | TGT } Cys TGC } TGA Stop TGG Trp | Third Letter | T C A G |
| | C | CTT } CTC } Leu CTA } CTG } | CCT } CCC } Pro CCA } CCG } | CAT } His CAC } CAA } Gln CAG } | CGT } CGC } Arg CGA } CGG } | | T C A G |
| | A | ATT } ATC } Ile ATA } ATG Met | ACT } ACC } Thr ACA } ACG } | AAT } Asn AAC } AAA } Lys AAG } | AGT } Ser AGC } AGA } Arg AGG } | | T C A G |
| | G | GTT } GTC } Val GTA } GTG } | GCT } GCC } Ala GCA } GCG } | GAT } Asp GAC } GAA } Glu GAG } | GGT } GGC } Gly GGA } GGG } | | T C A G |

↑
miejsce 4-krotnie
zdegenerowane

↑
miejsce 2-krotnie
zdegenerowane

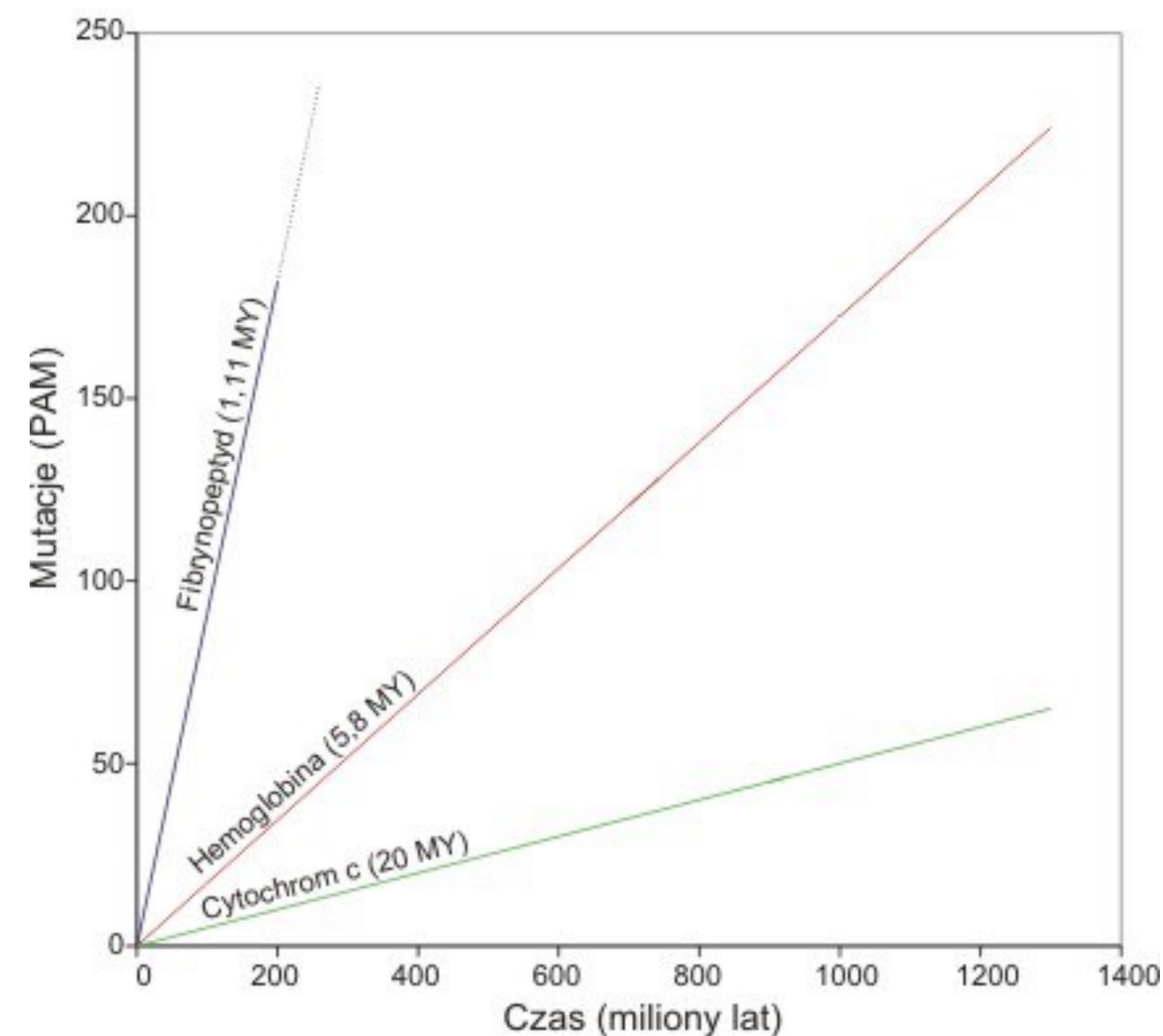
Tempo ewolucji sekwencji a funkcja

- Głównym czynnikiem determinującym ilościową zmienność sekwencji jest dobór negatywny (oczyszczający)
- Sekwencje o mniejszym znaczeniu funkcjonalnym (pseudogeny, mniej istotne obszary białek) ewoluują szybciej, niż obszary kluczowe dla funkcji
- **Konserwacja sekwencji świadczy o jej funkcji!**



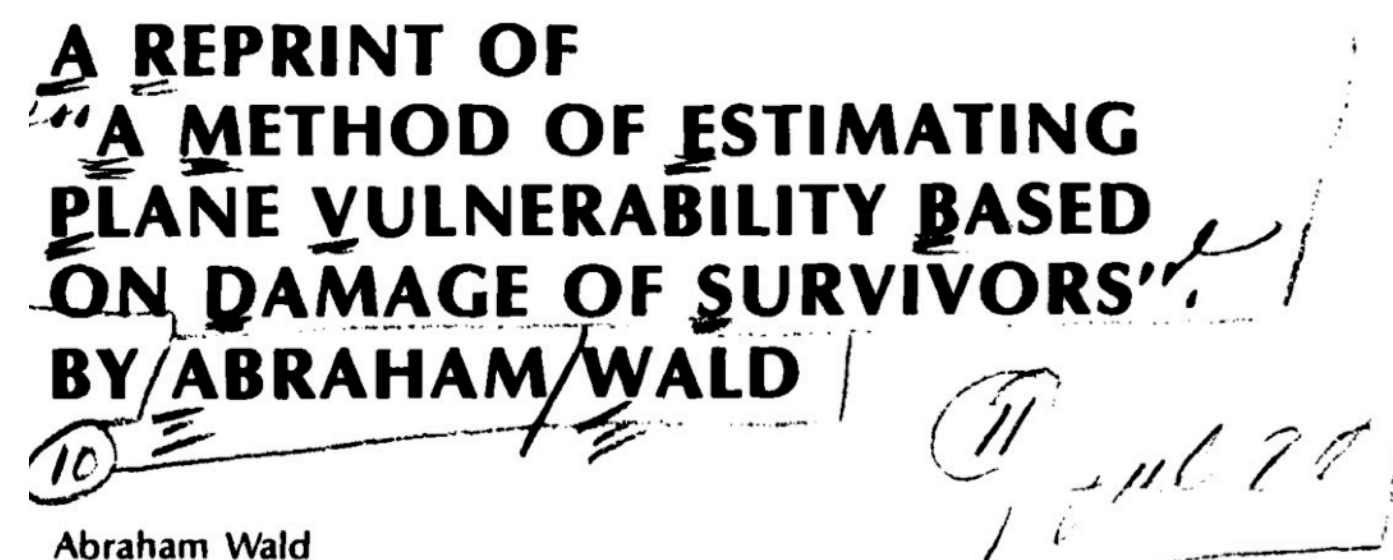
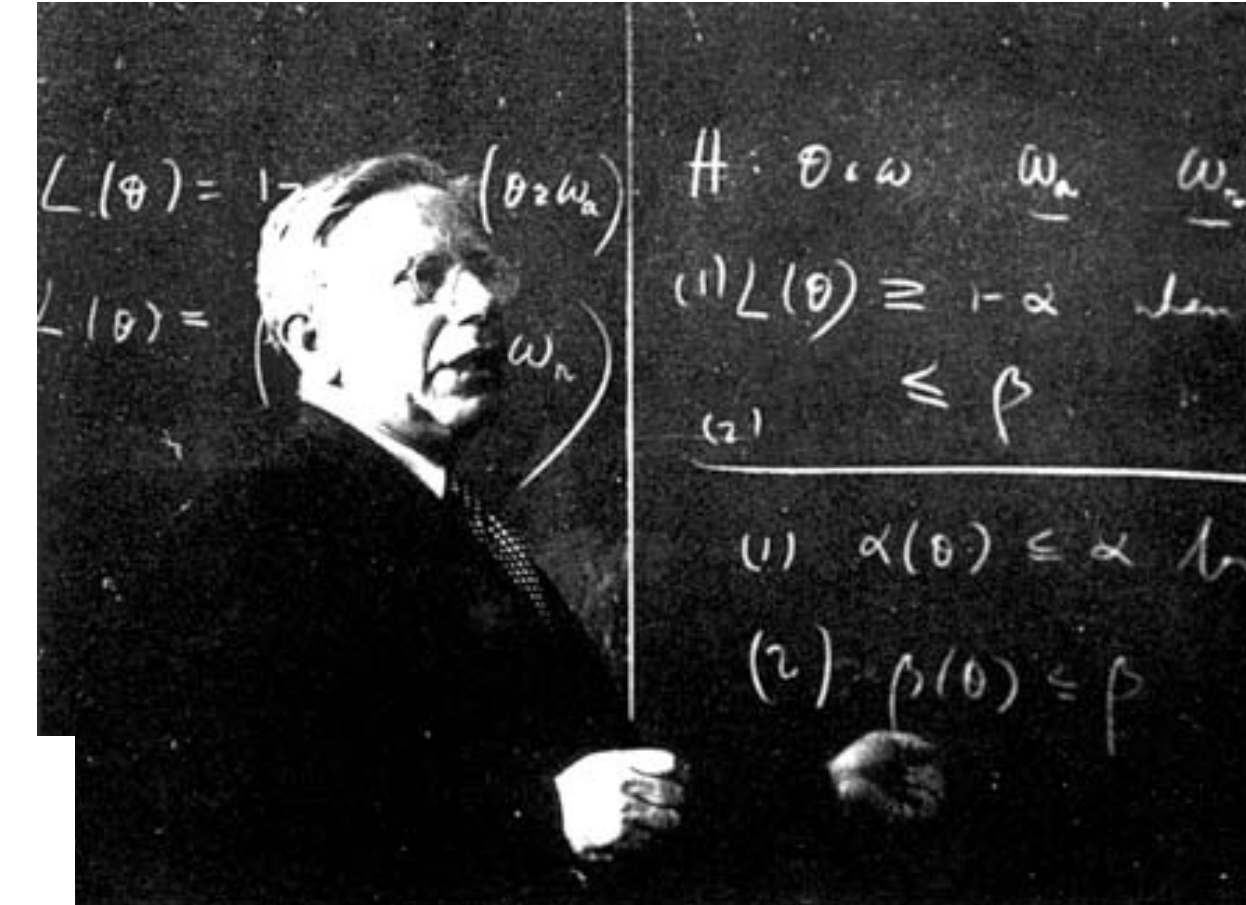
Tempo zmian

- Białka zaangażowane w podstawowe funkcje komórki ewoluują wolniej.
- W sekwencji białka obszary kluczowe dla funkcji ewoluują wolniej.
 - Jednostka: PAM/10⁸ lat
 - Jednostka czasu ewolucyjnego: ile lat (w milionach, 10⁶) potrzeba do utrwalenia 1 mutacji/100 aa (1 PAM)



Abraham Wald

- Pionier badań operacyjnych (teoria decyzji)
- Prace dla *Center for Naval Analyses* podczas II w. ś.
- Analiza rozmieszczenia przestrzelin w uszkodzonych samolotach
 - oryginalny plan: wzmocnić pancerz w miejscach, gdzie obserwuje się najwięcej przestrzelin
 - analiza Walda: wzmocnić tam, gdzie nie obserwuje się przestrzelin (samoloty tam trafione nie wróciły)



Abraham Wald

<http://www-history.mcs.st-andrews.ac.uk/PictDisplay/Wald.html>

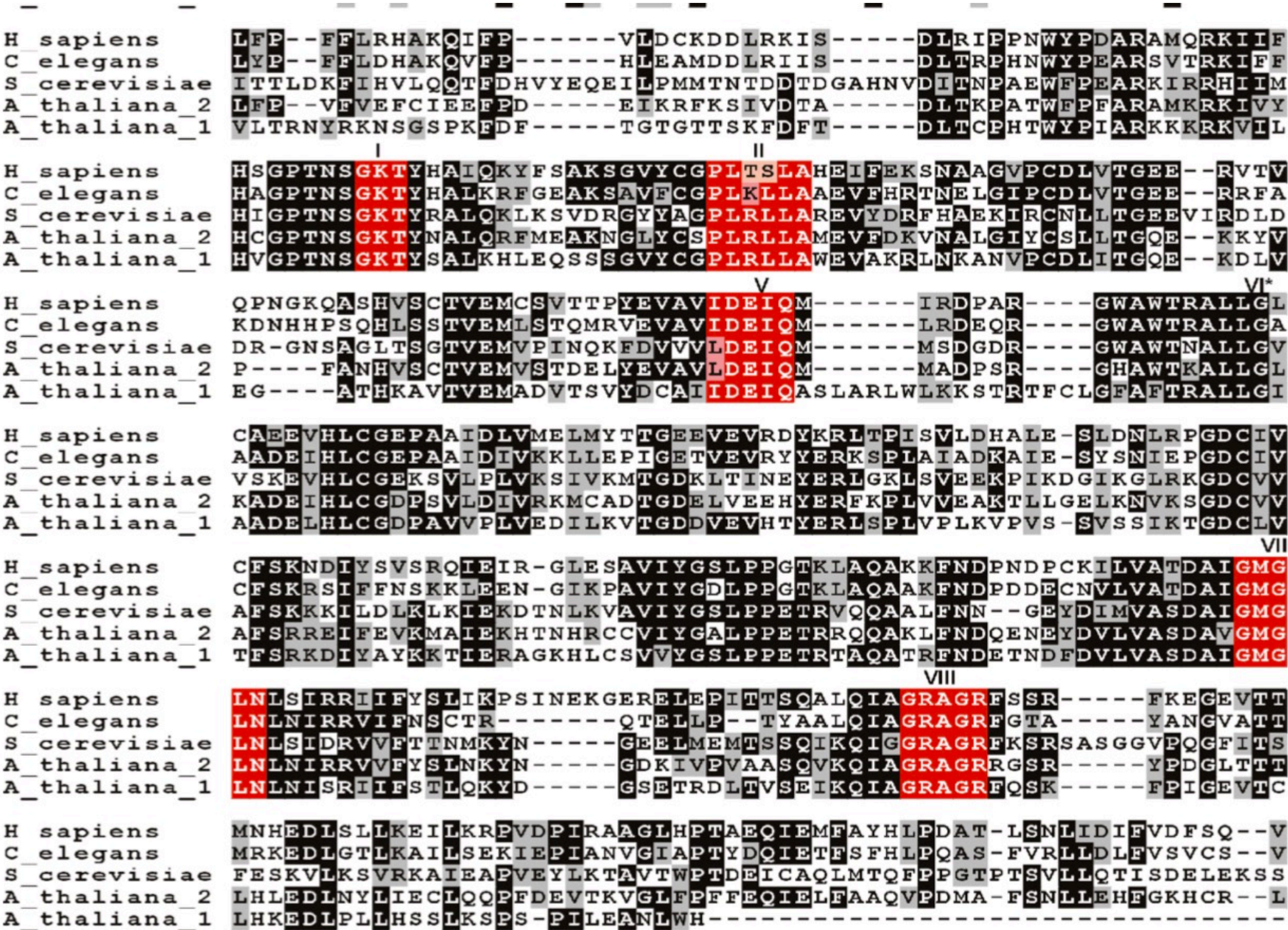


 **Operations Evaluation Group**
CENTER FOR NAVAL ANALYSES
2000 North Beauregard Street, Alexandria, Virginia 22311

<http://oai.dtic.mil/>

Zachowawczość sekwencji

- Obszary sekwencji białkowych najwolniej się zmieniające są zwykle kluczowe dla funkcji



Tempo zmian

- Czynnikiem decydującym o tempie zmian jest dobór oczyszczający (negatywny)
 - w “ważniejszych” sekwencjach więcej zmian będzie niekorzystnych (eliminacja przez dobór)
 - w mniej istotnych sekwencjach więcej zmian będzie neutralnych (utrwalanie przez dryf)
 - zmiany bez znaczenia dla funkcji będą neutralne
 - pseudogeny
 - niekodujące obszary międzygenowe?
 - podstawienia synonimiczne?

Spór wokół ENCODE

- ENCODE - projekt opisujący sekwencje w genomie (Encyclopedia of DNA Elements)
- Wiele sekwencji międzygenowych, niekodujących ulega transkrypcji
 - 80% genomu funkcjonalne
 - czy istnieje “śmieciowy DNA”?
- Czy to znaczy, że są funkcjonalne?
- **Jeżeli nie ma śladów działania doboru - nie ma funkcji!**
- Ślady działania doboru: 2-15% całego genomu

ARTICLE

doi:10.1038/nature11247

An integrated encyclopedia of DNA elements in the human genome

The ENCODE Project Consortium*

The human genome encodes the blueprint of life, but the function of the vast majority of its nearly three billion bases is unknown. The Encyclopedia of DNA Elements (ENCODE) project has systematically mapped regions of transcription, transcription factor association, chromatin structure and histone modification. These data enabled us to assign biochemical functions for 80% of the genome, in particular outside of the well-studied protein-coding regions. Many



GBE

GENOME BIOLOGY AND EVOLUTION

On the Immortality of Television Sets: “Function” in the Human Genome According to the Evolution-Free Gospel of ENCODE

Dan Graur^{1,*}, Yichen Zheng¹, Nicholas Price¹, Ricardo B.R. Azevedo¹, Rebecca A. Zufall¹, and Eran Elhaik²

¹Department of Biology and Biochemistry, University of Houston

²Department of Mental Health, Johns Hopkins University Bloomberg School of Public Health

*Corresponding author: E-mail: dgraur@uh.edu.

Accepted: February 16, 2013

Status neutralizmu

- Wyjaśnia wiele zjawisk obserwowanych w ewolucji molekularnej
 - wysoki polimorfizm sekwencji DNA i białek
 - zegar molekularny
 - ale jest wiele odstępstw, nie istnieje globalny zegar prawdziwy dla wszystkich gałęzi drzewa życia
 - wolniejsza ewolucja sekwencji o kluczowym znaczeniu
 - to też można wyjaśnić modelem, w którym większość mutacji jest albo niekorzystna, albo korzystna, ale niekorzystnych jest więcej
- Jest bardzo przydatny jako hipoteza zerowa do badania doboru naturalnego na poziomie sekwencji!

Status neutralizmu

- Dane molekularne, zwłaszcza genomowe, pozwoliły ocenić zgodność modelu neutralnego z obserwacją zmienności sekwencji
- Kimura: 1968 – nie były wtedy znane metody sekwencjonowania DNA!

Status neutralizmu

- Smith & Eyre-Walker 2002 – 45% podstawień aminokwasowych w ewolucji *Drosophila* sp. utwalonych przez dobór dodatni
- Andolfatto 2005 – pomiędzy *D. melanogaster* i *D. simulans* dobór dodatni odpowiada za utrwalenie:
 - 20% podstawień w DNA w intronach i obszarach międzygenowych
 - 60% podstawień w DNA w sekwencjach UTR

Status neutralizmu

- Głównym i nieprzemijającym osiągnięciem jest stworzenie matematycznego opisu współdziałania dryfu i doboru naturalnego (dodatniego i oczyszczającego) w ewolucji molekularnej
- Dzięki tym modelom opracowano testy poszukujące śladów doboru w sekwencjach (model neutralny jako hipoteza zerowa)
- Istnieje znacząca liczba pozycji i sekwencji ewoluujących według modelu neutralnego
 - można dobrać sekwencje tak, by uzyskać zegar molekularny

Status neutralizmu

- Dryf genetyczny ma w ewolucji molekularnej bardzo znaczącą, ale nie wyłączną rolę
 - znaczne obszary genomu ewoluują w sposób bliski neutralnemu

Badanie doboru

- Założenie: mutacje synonimiczne są neutralne, sekwencje porównywane są parami
- K_a (dN) – liczba mutacji niesynonimicznych na liczbę możliwych miejsc niesynonimicznych
- K_s (dS) – liczba mutacji synonimicznych na liczbę możliwych miejsc synonimicznych
- Stosunek K_a/K_s (ω) jest miarą działania doboru

Badanie doboru

- Wartość ω rzadko przekracza 1 dla całej sekwencji (wyjątek np. geny MHC)
- Średnia wartość ω w porównaniach między naczelnymi a gryzoniami wynosi 0,2, między człowiekiem a szympansem 0,4
- Odchylenie ω od średniej dla konkretnego genu w konkretnej linii ewolucyjnej może świadczyć o działaniu doboru
- W sekwencji mogą występować obszary o różnej wartości ω , wskazując na działanie doboru na poszczególne regiony a nawet pozycje aminokwasowe w białku

Badanie doboru II

- Porównanie zmian synonimicznych i niesynonimicznych w obrębie populacji danego gatunku i pomiędzy gatunkami.

Test McDonalda-Kreitmana

- Stosunek mutacji synonimicznych do niesynonimicznych w obrębie populacji vs. taki sam stosunek dla różnic między gatunkami
- Jeżeli zmiany są neutralne, wówczas stosunek ten powinien być w obu przypadkach taki sam

Przykład: gen ADH u trzech gatunków *Drosophila*

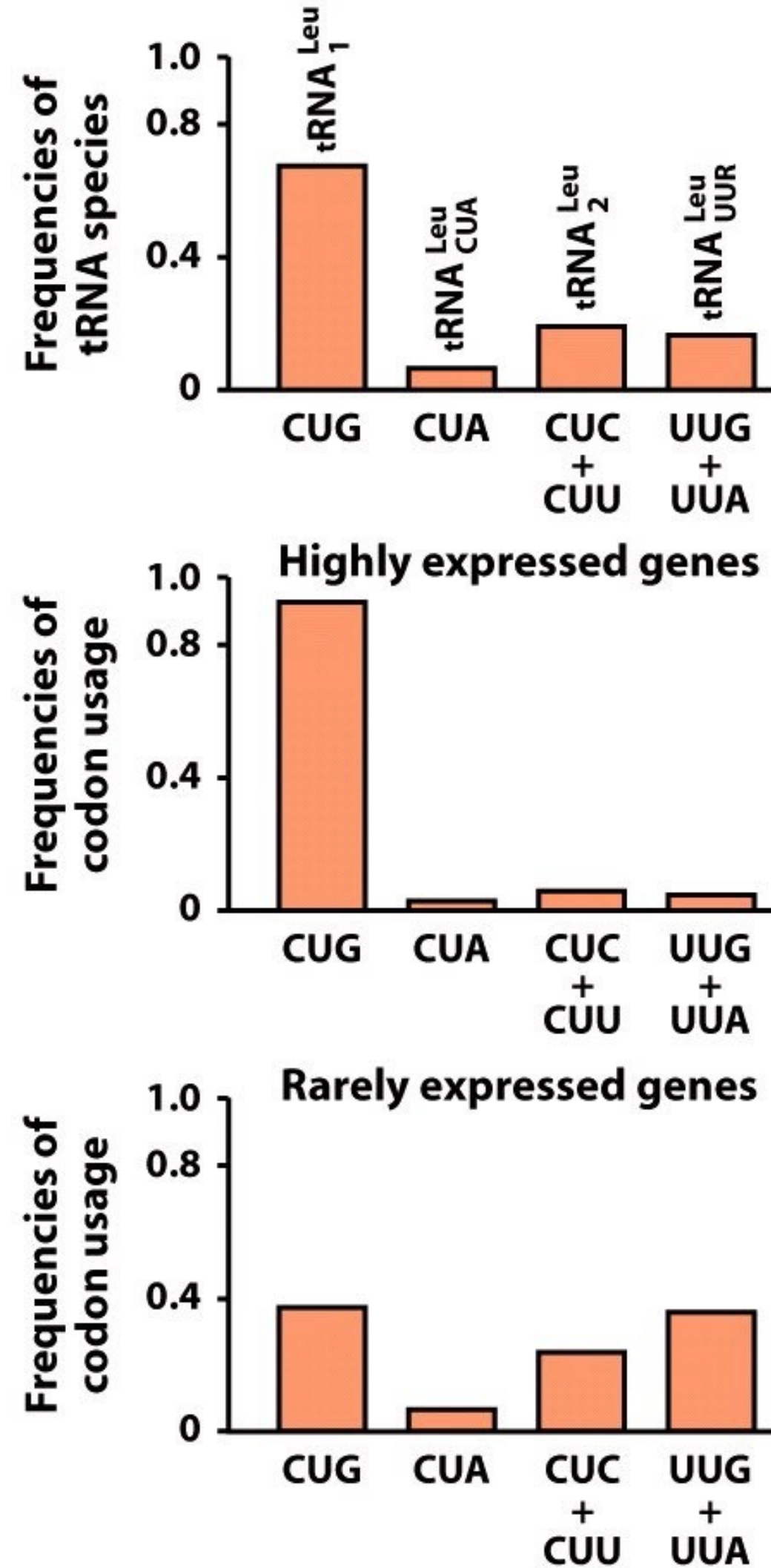
| | synonimiczne | niesynonimiczne | stosunek |
|---------------------|--------------|-----------------|----------|
| wewnątrzpopulacyjne | 42 | 2 | ~0,05 |
| międzygatunkowe | 17 | 7 | ~0,41 |

Wniosek: zmiany niesynonimiczne są szybko utrwalane w specjacji – nie są neutralne

Czy zmiany synonimiczne są neutralne

- Kodony synonimiczne nie są równocenne
- Zmiana kodonu częstego na rzadki może wpłynąć na poziom ekspresji i kinetykę translacji

(a) *Escherichia coli*



(b) Yeast

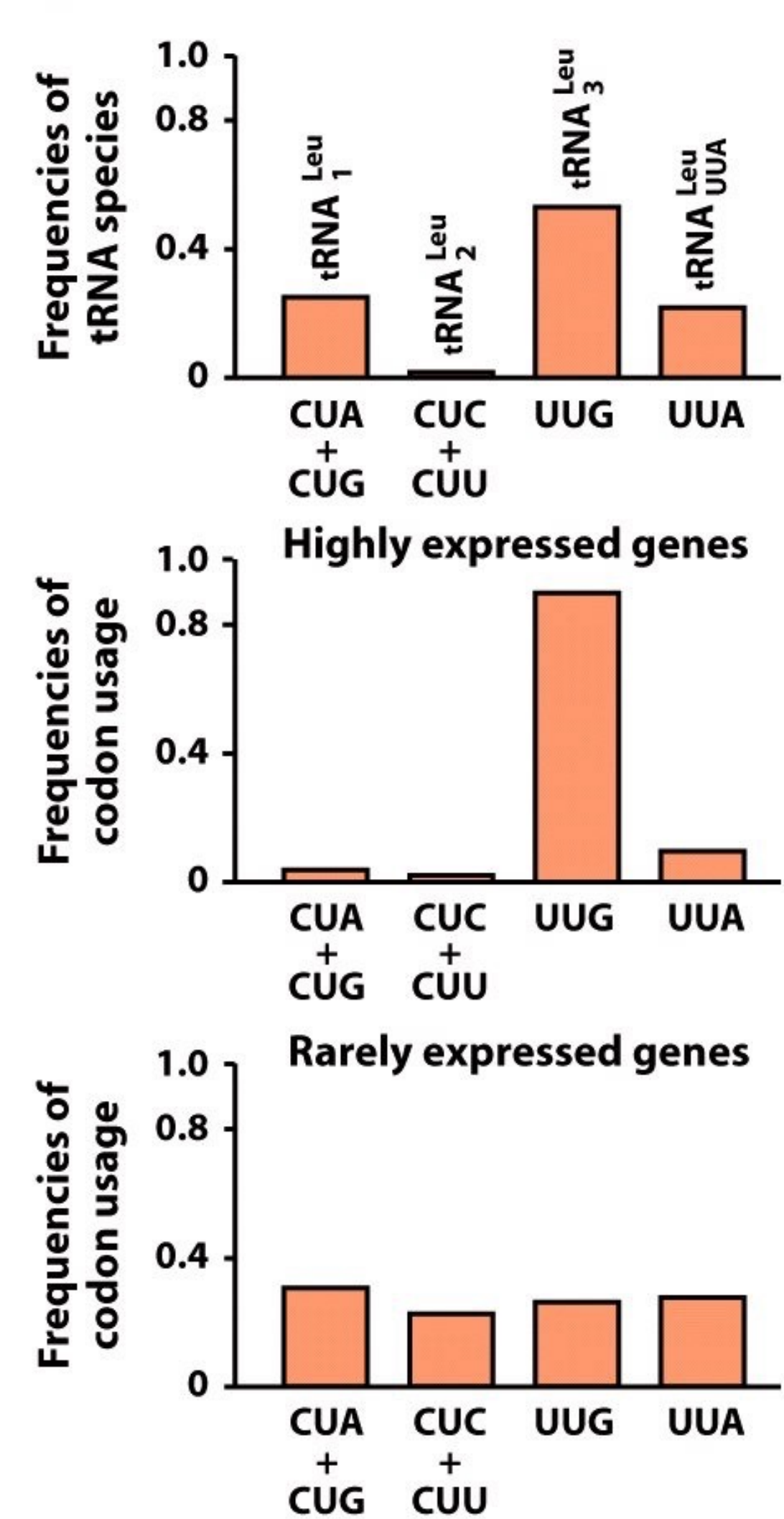


Figure 7-24 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

Czy zmiany synonimiczne są neutralne?

RESEARCH ARTICLE

AMERICAN JOURNAL OF
medical genetics 

A Synonymous Mutation in *TCOF1* Causes Treacher Collins Syndrome Due to Mis-Splicing of a Constitutive Exon

D. Macaya,¹ S.H. Katsanis,¹ T.W. Hefferon,² S. Audlin,¹ N.J. Mendelsohn,³ J. Roggenbuck,³ and G.R. Cutting^{1*}

¹DNA Diagnostic Laboratory, Institute of Genetic Medicine, Johns Hopkins University, Baltimore, Maryland

²Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland

³Children's Hospitals & Clinics of Minnesota, Minneapolis, Minnesota

Received 26 September 2008; Accepted 25 February 2009

Czy zmiany synonimiczne są neutralne?

A “Silent” Polymorphism in the *MDR1* Gene Changes Substrate Specificity

Chava Kimchi-Sarfaty,*† Jung Mi Oh,†‡ In-Wha Kim, Zuben E. Sauna, Anna Maria Calcagno, Suresh V. Ambudkar, Michael M. Gottesman†

Synonymous single-nucleotide polymorphisms (SNPs) do not produce altered coding sequences, and therefore they are not expected to change the function of the protein in which they occur. We report that a synonymous SNP in the *Multidrug Resistance 1 (MDR1)* gene, part of a haplotype previously linked to altered function of the *MDR1* gene product P-glycoprotein (P-gp), nonetheless results in P-gp with altered drug and inhibitor interactions. Similar mRNA and protein levels, but altered conformations, were found for wild-type and polymorphic P-gp. We hypothesize that the presence of a rare codon, marked by the synonymous polymorphism, affects the timing of cotranslational folding and insertion of P-gp into the membrane, thereby altering the structure of substrate and inhibitor interaction sites.

The *MDR1* gene product, the adenosine triphosphate (ATP)-binding cassette (ABC) transporter ABCB1 or P-gp, is an ATP-driven efflux pump contributing to the pharmacokinetics of drugs that are P-gp substrates and to the multidrug resistance of cancer cells (1, 2). To date, more than 50 single-nucleotide polymorphisms (SNPs) have been reported for *MDR1* (www.ncbi.nlm.nih.gov/SNP/GeneGt.cgi?geneID=5243). One of these, a synonymous SNP in exon 26 (C3435T), was

sometimes found to be associated with altered P-gp activity (3–6) and, when it appears in a haplotype, with reduced functionality (7). This association may be explained in different ways. Perhaps it is because C3435T is in linkage disequilibrium with other common functional non-synonymous polymorphisms such as G2677T. In fact, the C1236T (a synonymous SNP), G2677T, and C3435T polymorphisms are part of a common haplotype (8, 9). Another possible explanation is that allele-specific differences in

mRNA folding could influence splicing, processing, or translational control and regulation (10, 11). A third possibility is that the effect of the C3435T polymorphism on the levels of cell surface P-gp activity or its function is rather modest or drug-specific. Finally, numerous environmental factors are known to affect the expression and phenotypic activity of P-gp (12).

To determine whether the C3435T polymorphism actually does affect P-gp activity, we expressed wild-type and polymorphic P-gps in HeLa cells with the use of a transient expression system (13). The same experiments were carried out on BSC-1 (epithelial cells of African green monkey kidney origin), Vero-76 (monkey kidney cells), and 12E1 (CEM human cells) cell lines (14), with similar results, indicating that this phenomenon is not specific to HeLa cells.

Laboratory of Cell Biology, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892, USA.

*Present address: Center for Biologics Evaluation and Research, Food and Drug Administration, 29 Lincoln Drive, Room 316, Bethesda, MD 20892, USA.

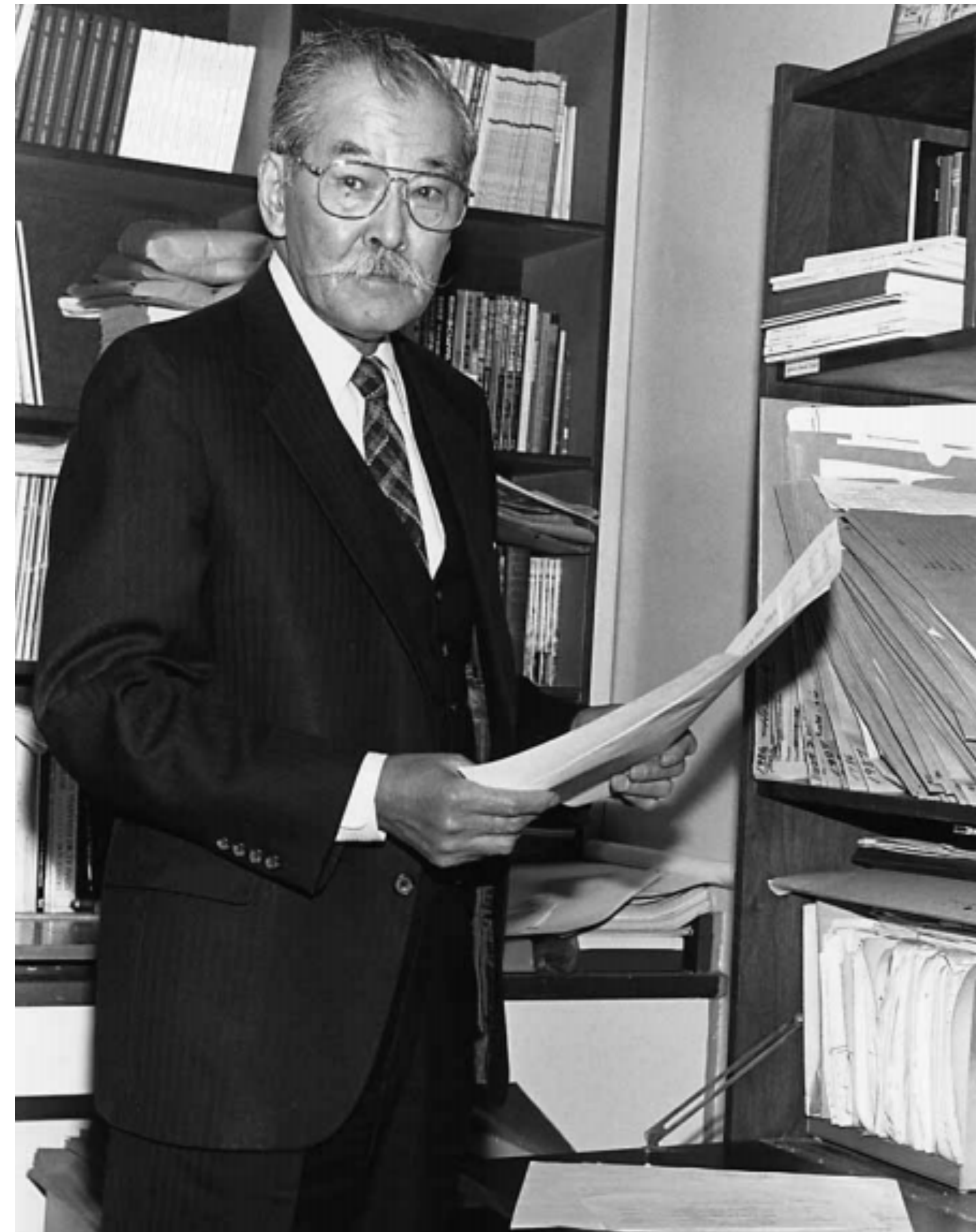
†To whom correspondence should be addressed. E-mail: mgottesman@nih.gov (M.M.G.); jmoh@snu.ac.kr (J.M.O.); kimchi@cber.fda.gov (C.K.-S.)

‡Present address: College of Pharmacy, Seoul National University, Seoul 151-742, South Korea.

Innowacje ewolucyjne w genomie

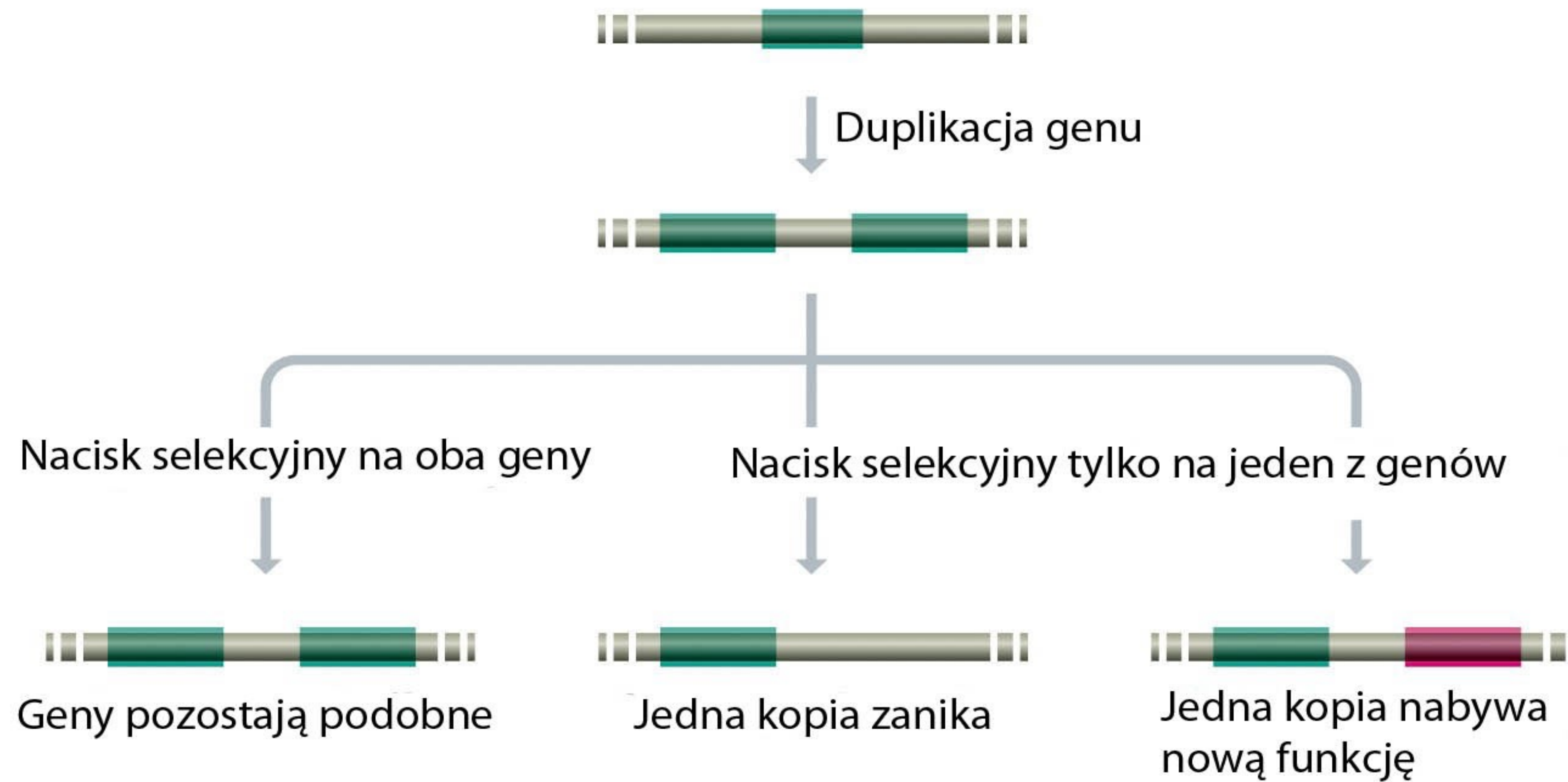
Skąd biorą się nowe funkcje (geny)

- Mutacje mogą zmienić funkcję genu, ale zwykle z utratą funkcji dotychczasowej
- Prawdopodobieństwo powstania nowego genu *de novo* (np. z sekwencji niekodującej) jest małe
- Rozwiązanie - duplikacje



Susumu Ohno (1928-2000)

Duplikacje



Liczba genów wzrastała w historii życia

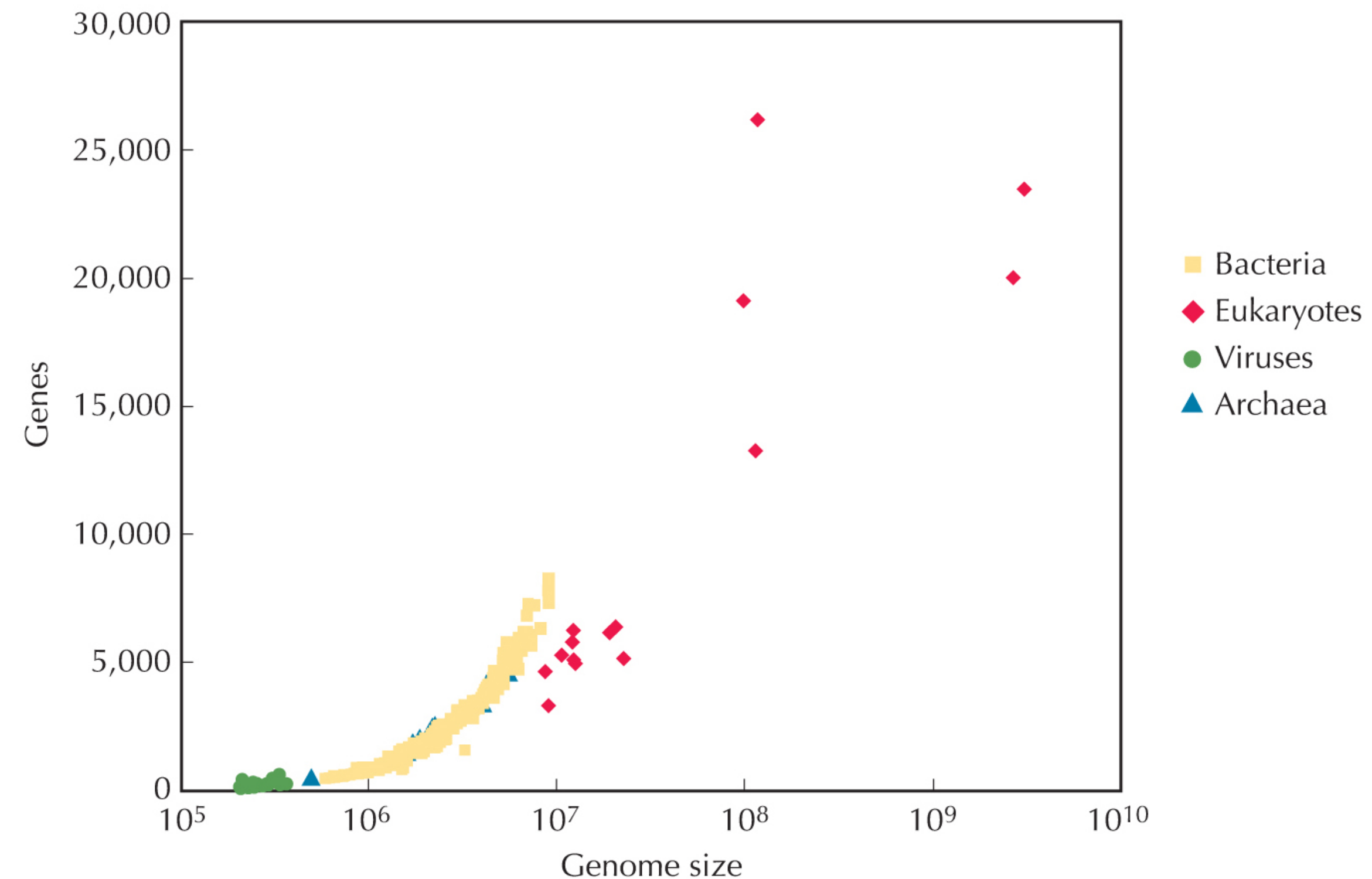
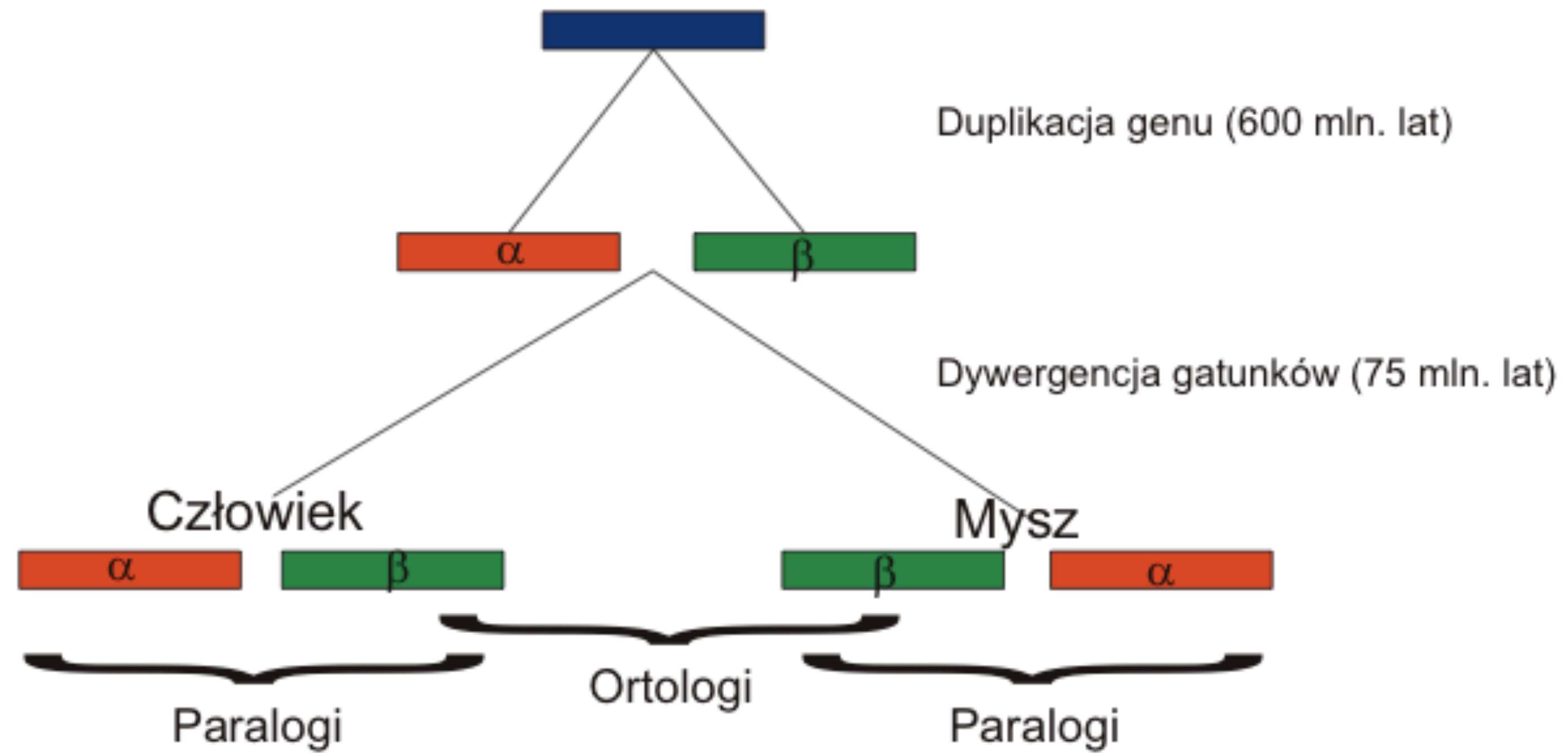


FIGURE 7.3. Genome size vs. number of protein-coding genes. The number of genes is highly correlated to genome size for bacteria, archaea, and viruses, but less so for eukaryotes. Many archaeal points (*blue triangles*) are hidden under bacterial ones (*yellow squares*).

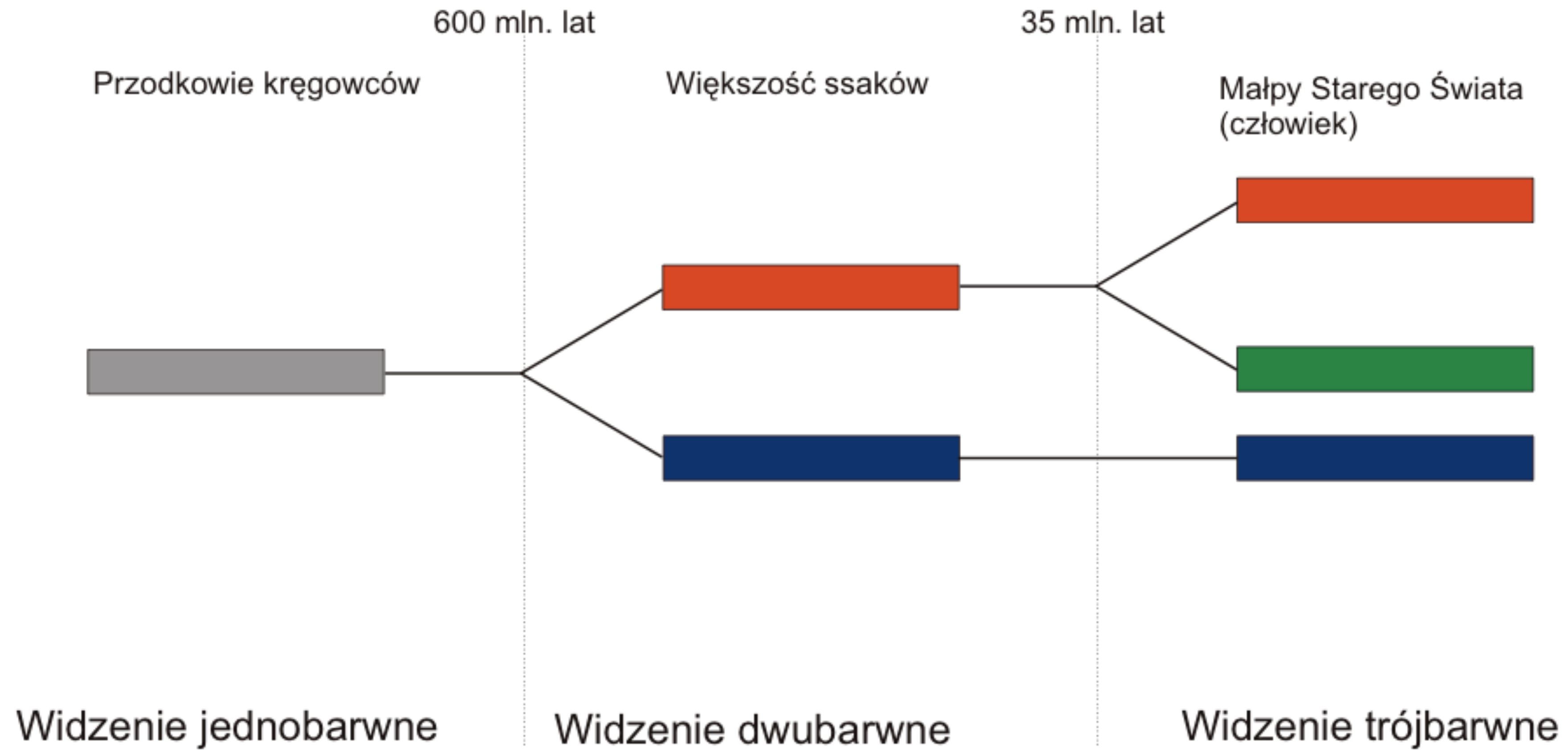
Ewolucja globin



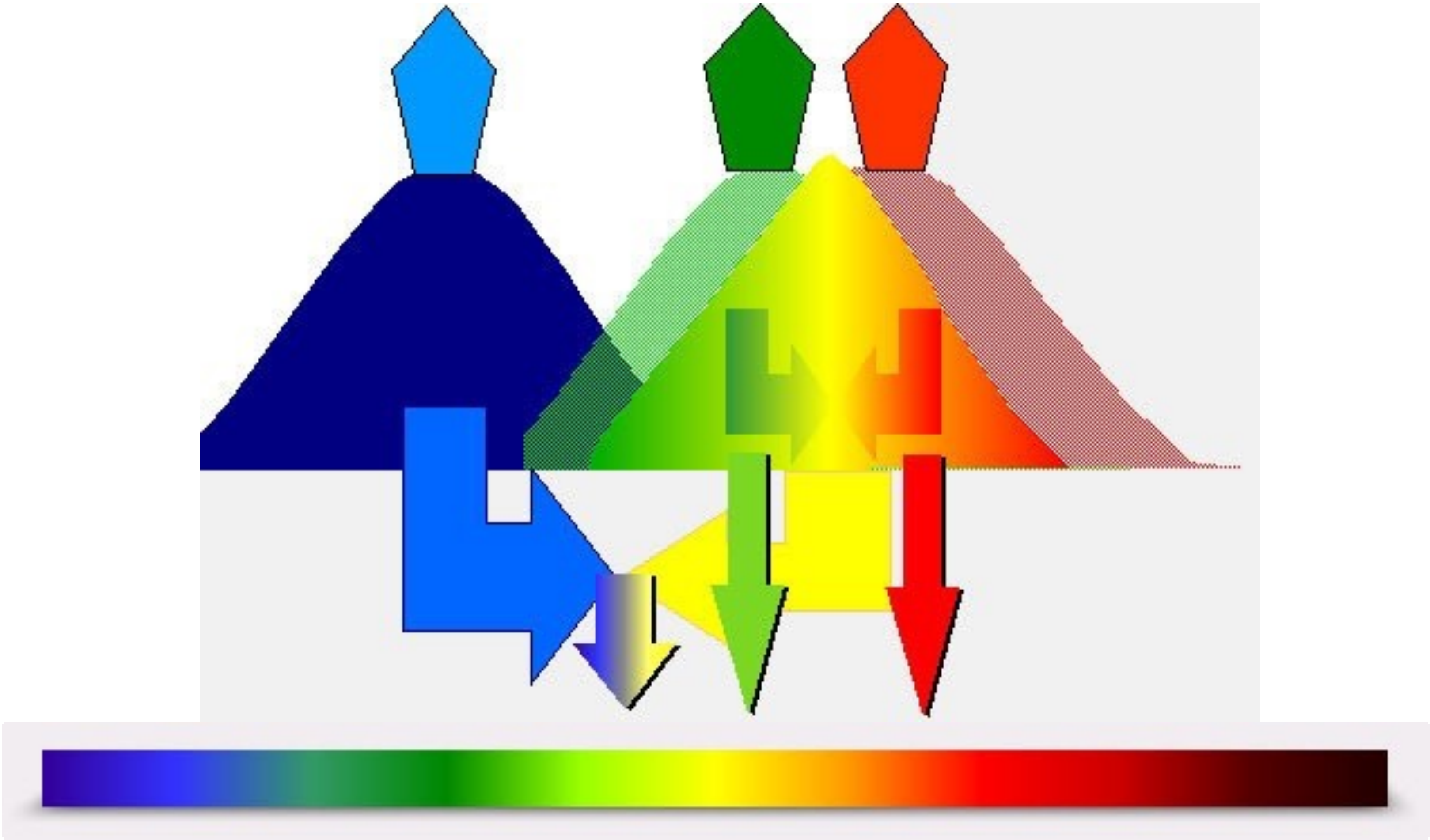
Paralogi i ortologu

- Paralogi – geny homologiczne w tym samym genomie, powstałe w wyniku duplikacji genu - np. α -globina i β -globina człowieka
- Ortologu – geny homologiczne powstałe w wyniku specjacji, pochodzące od genu u wspólnego przodka – np. α -globina człowieka i α -globina myszy

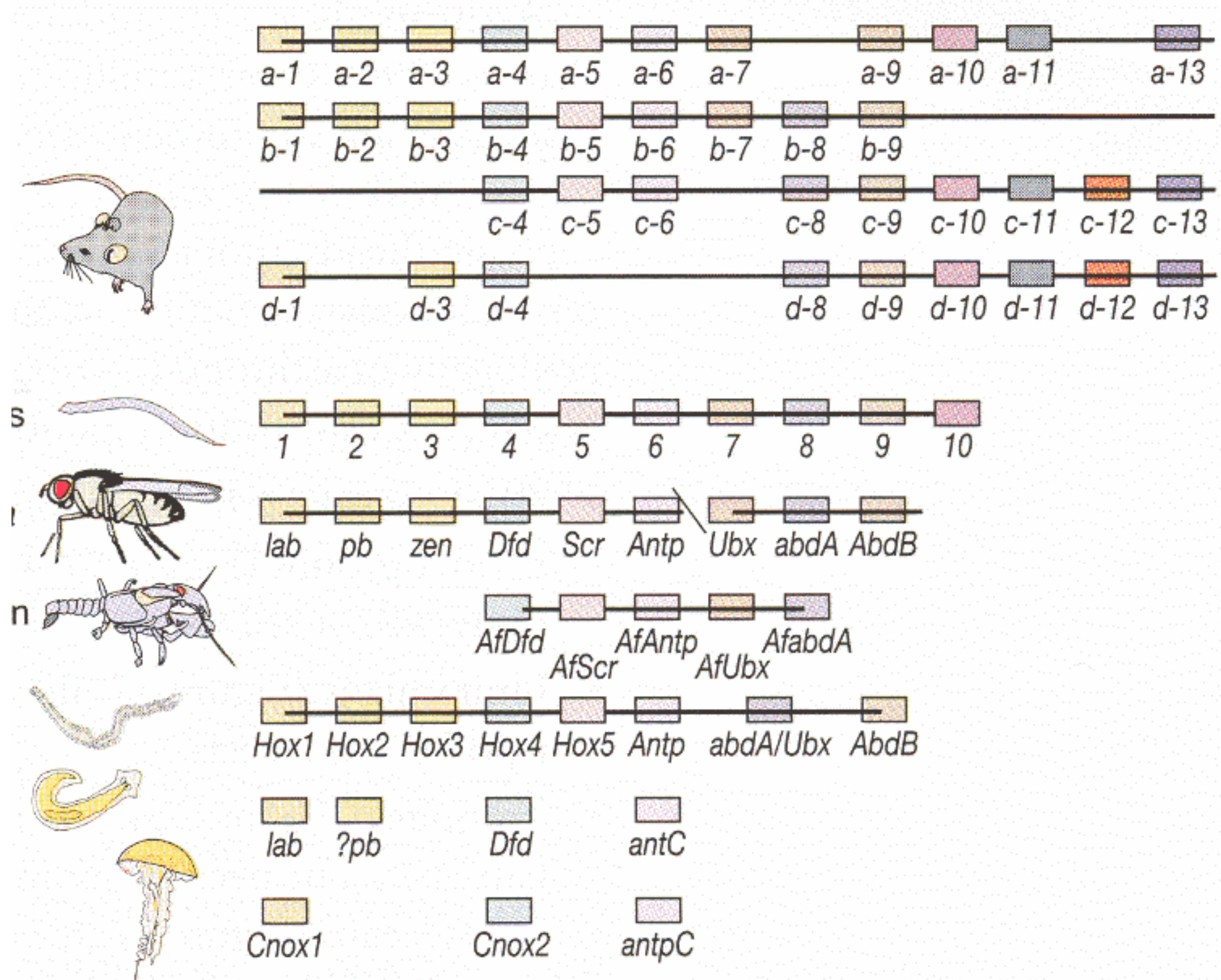
Ewolucja genów opsyn



Ewolucja widzenia barw

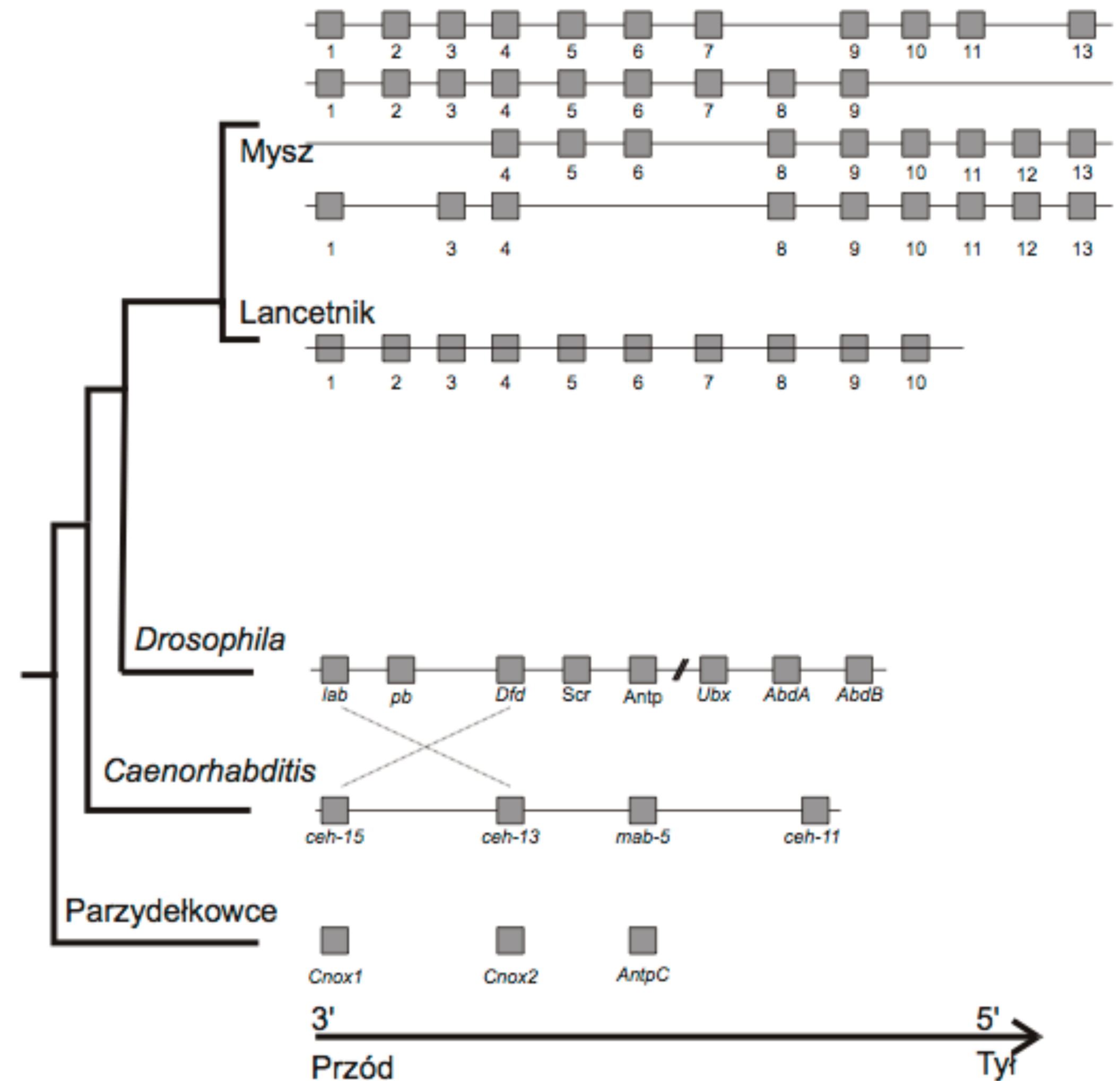


Geny HOX – regulatory rozwoju

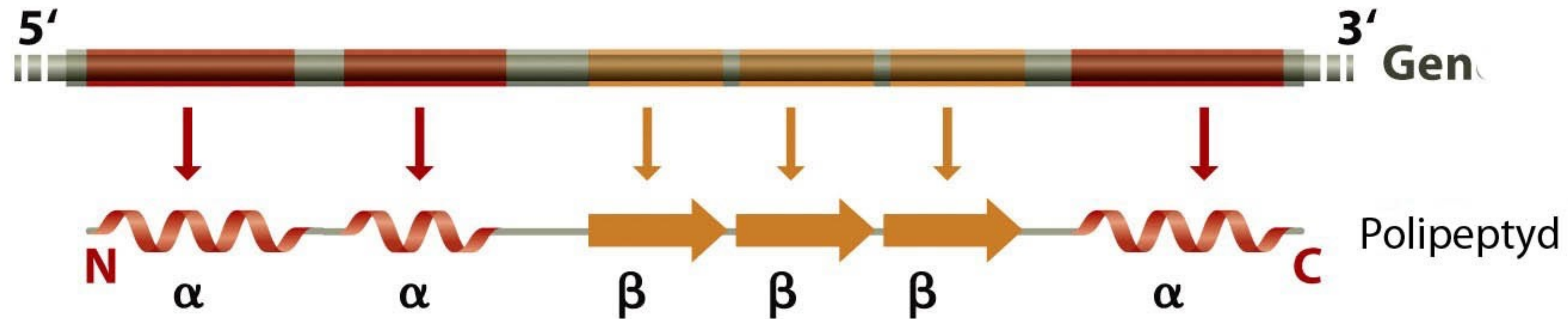


Duplikacje całych genomów

- Zmianie może ulec liczba chromosomów
- Podwojeniu może ulec cały genom
- Hipoteza 2R (hipoteza Ohno) – podwojenie genomu na początku ewolucji kręgowców
 - 2 rundy podwojenia
 - np. geny Hox



Białka składają się z domen

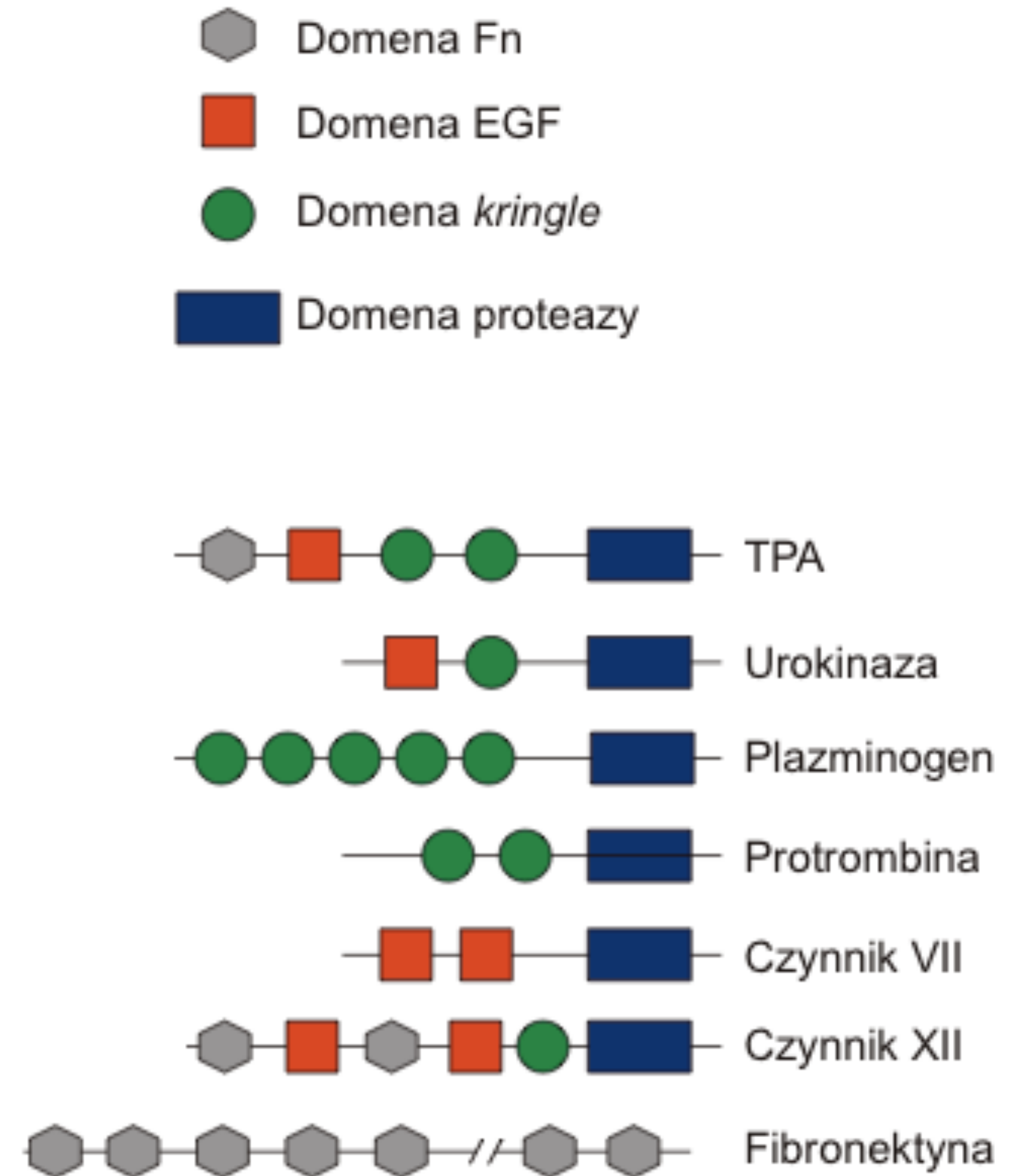


T.A. Brown. Genomy III, PWN 2009

Tasowanie domen – kombinatoryka w białkach.

Wspólne motywy w różnych genach

- Możemy stawiać hipotezy dotyczące funkcji nieznanymi białek na podstawie motywów znajdujących w sekwencji.
- Podstawa większości współczesnych badań biochemicznych!!



Ewolucyjne klocki

- Złożone sieci współzależności – złożoność budowana przez oddziaływania i kombinacje, a nie liczbę elementów składowych
- Nowe elementy przez duplikację istniejących



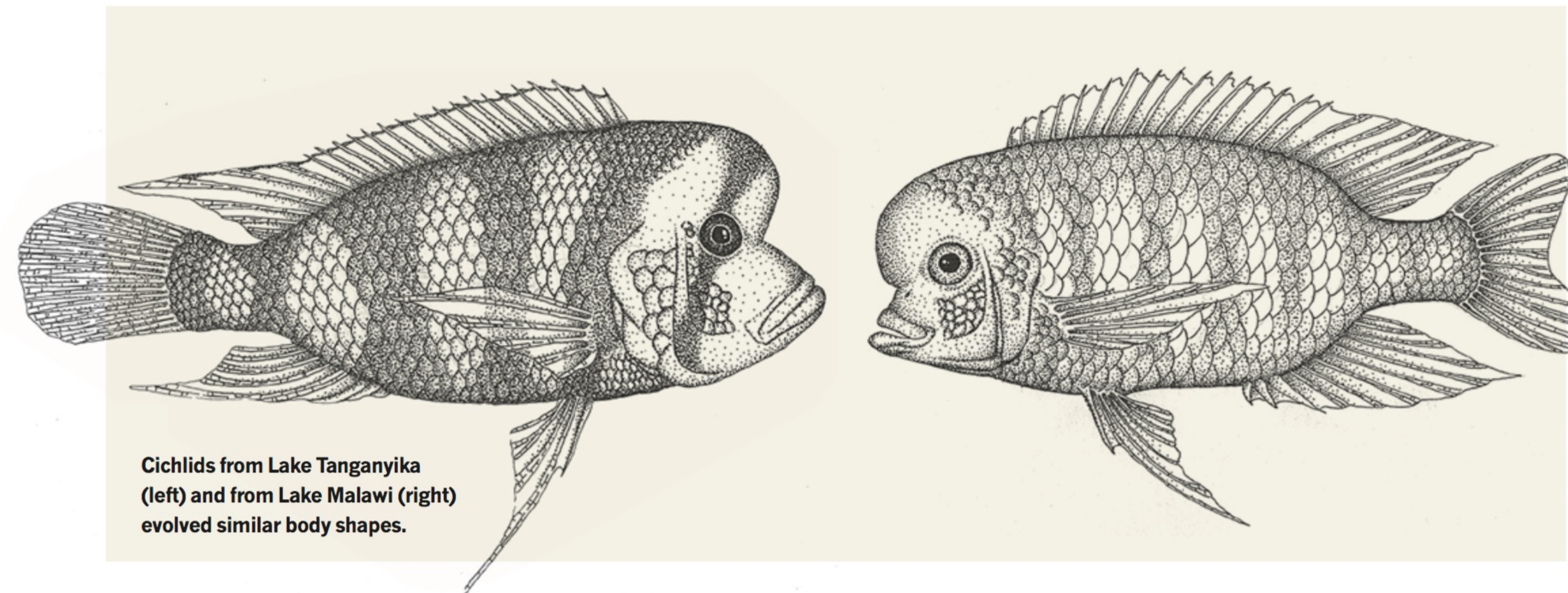
Ewolucyjna zmienność genomów

- U Eukaryota skład genomu zmienia się powoli
 - większość genów człowieka (>95%) ma odpowiedniki w genomie myszy, żaby itp.
 - za różnice odpowiadają subtelne zmiany regulacji i współdziałania genów
- U Prokaryota (bakterie, archeony) duże różnice w zestawie genów nawet u blisko spokrewnionych organizmów

Syntetyczna teoria ewolucji

- Tzw. “współczesna synteza” (*modern synthesis*)
- Połączenie teorii doboru naturalnego z genetyką
 - klasyczną (pierwsza połowa XX w.)
 - molekularną - ewolucja molekularna
- Neodarwinizm
 - niezbyt jasne definicje:
 - nurt STE kładący nacisk na rolę doboru
 - synonim STE

Aktualna dyskusja



Does evolutionary theory need a rethink?

Researchers are divided over what processes should be considered fundamental.

POINT

Yes, urgently

COUNTERPOINT

No, all is well

Status STE

- Trwają dyskusje dotyczące włączenia do teorii ewolucji takich zjawisk, jak:
 - epigenetyka
 - plastyczność rozwojowa i fenotypowa
 - kształtowanie niszy przez organizm (dobór jako proces dwukierunkowy)
 - symbioza na poziomie genomu
 - ogólnie - nie zawsze ewidentne przełożenie genotypu na fenotyp (dostosowanie)

Podstawowa wątpliwość

- Czy allelom pojedynczych genów można przypisywać określony współczynnik selekcji?
 - Oddziaływania genetyczne
 - Zdarza się, że mutacja jest korzystna w jednym tle genotypowym, a w innym - niekorzystna

Status STE

- Na gruncie nauki **nie** jest podważana idea ewolucji biologicznej, ani jej zasadniczo darwinowski mechanizm (losowa zmienność i dobór)
- Dyskusje w obrębie TE nie są podstawą do jej odrzucenia
- Odrzucanie TE zawsze ma motyw pozanaukowy (religijny), niezależnie od prób prezentacji na gruncie nauki
- “inteligentny projekt” - próba przedstawienia kreacjonizmu jako teorii naukowej

Powstanie i najwcześniejsze dzieje życia

Od abiogenezy do LUCA

Czym jest życie?

metabolizm

+

informacja
(replikacja)

(A)biogeneza

- Ewolucja jest właściwością organizmów żywych
 - Życie = ewolucja
- Powstanie życia z materii nieożywionej nie było zjawiskiem ewolucyjnym
 - trudności z wyjaśnieniem abiogenezy nie mogą być traktowane jako zarzut wobec teorii ewolucji
 - właściwe dziedziny:
 - fizyka (teoria złożoności, teoria samoorganizacji, termodynamika_
 - chemia
 - planetologia

Koncepcje abiogenezy

- Powstanie informacji
 - kluczowe powstanie zdolności (samo)replikacji
 - powiązanie genotypu z fenotypem – możliwość działania doboru
 - progenota
- Powstanie metabolizmu
 - kluczowe powstanie samoorganizującej się sieci metabolicznej
 - powielanie struktury nie na zasadzie replikacji matrycowej
 - replikacja “wynaleziona” później

Czas i scena

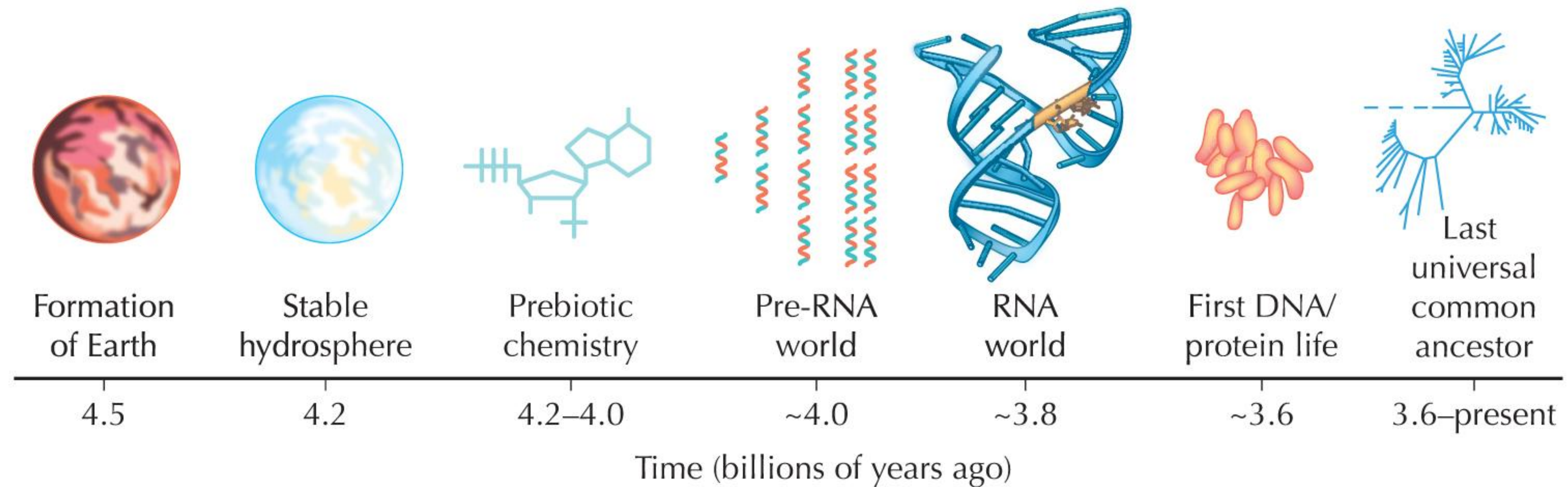


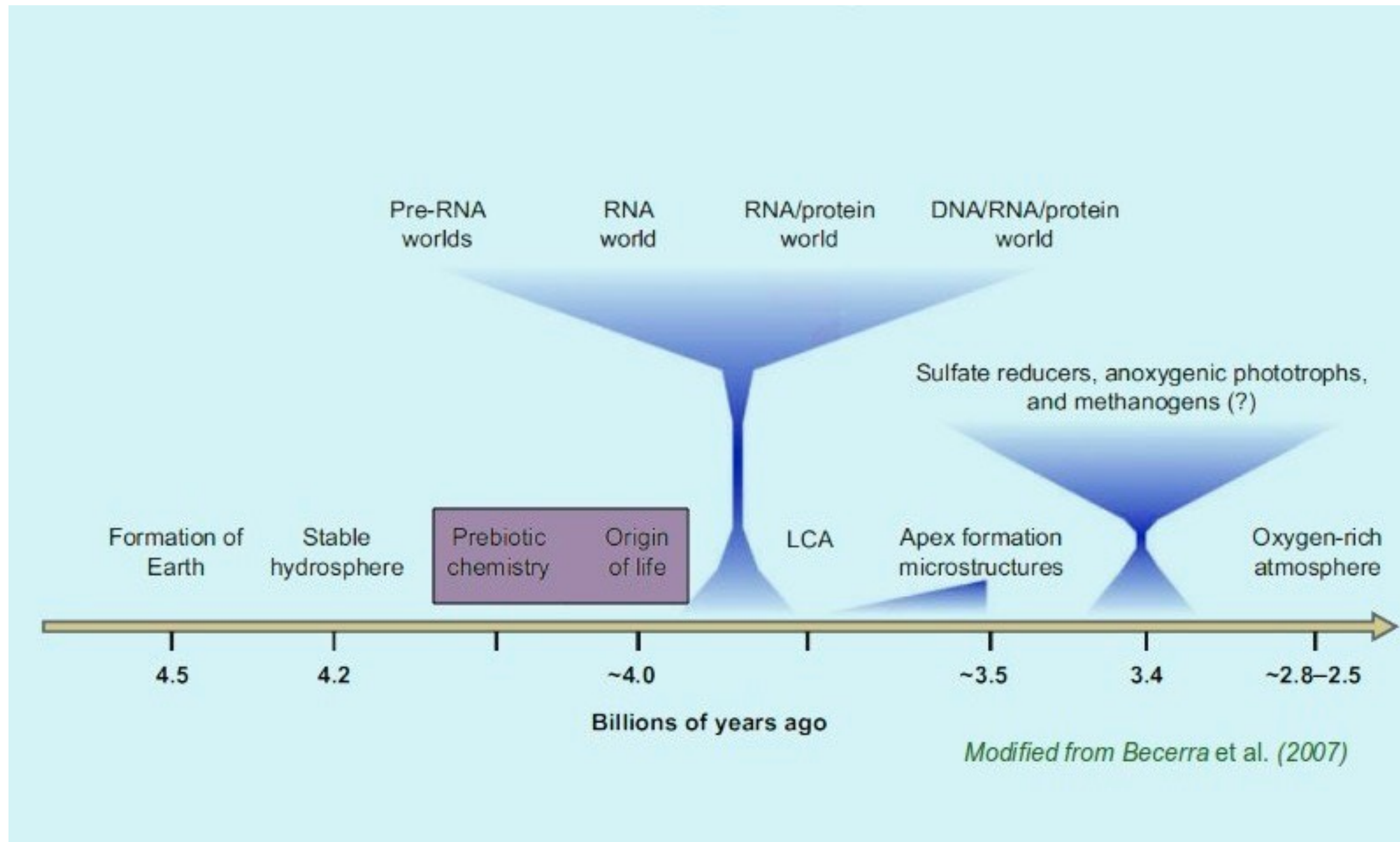
FIGURE 4.4. Steps in the origin of life.

4.4, modified from Joyce G.F., *Nature* **418**: 214–221, © 2002 Macmillan, www.nature.com

Evolution © 2007 Cold Spring Harbor Laboratory Press

Istnieją też koncepcje umieszczające część z tych etapów poza Ziemią

Prehistoria życia



Pierwsze ślady życia

Współczesne stromatolity



czert Strelley Pool, Australia – wiek 3,4 mld lat,
(Brasier et al. 2006)

Pilbara Craton - wiek 3,5 mld lat
(Van Kranendonk et al. 2008)



skamieniałe maty mikroorganizmów - stromatolity

Skamieniałości ze Strelley Pool



A close-up, cross-section view of the interior of a domical stromatolite. The black layers are the "cooked" organic remains of Early Archean microbial mats.
Credit: Abigail Allwood

Najstarsze ślady życia

- Odkryte w 2016 r. w skałach z Grenlandii
- Możliwe ślady stromatolitów
- Wiek: 3,7 miliarda lat

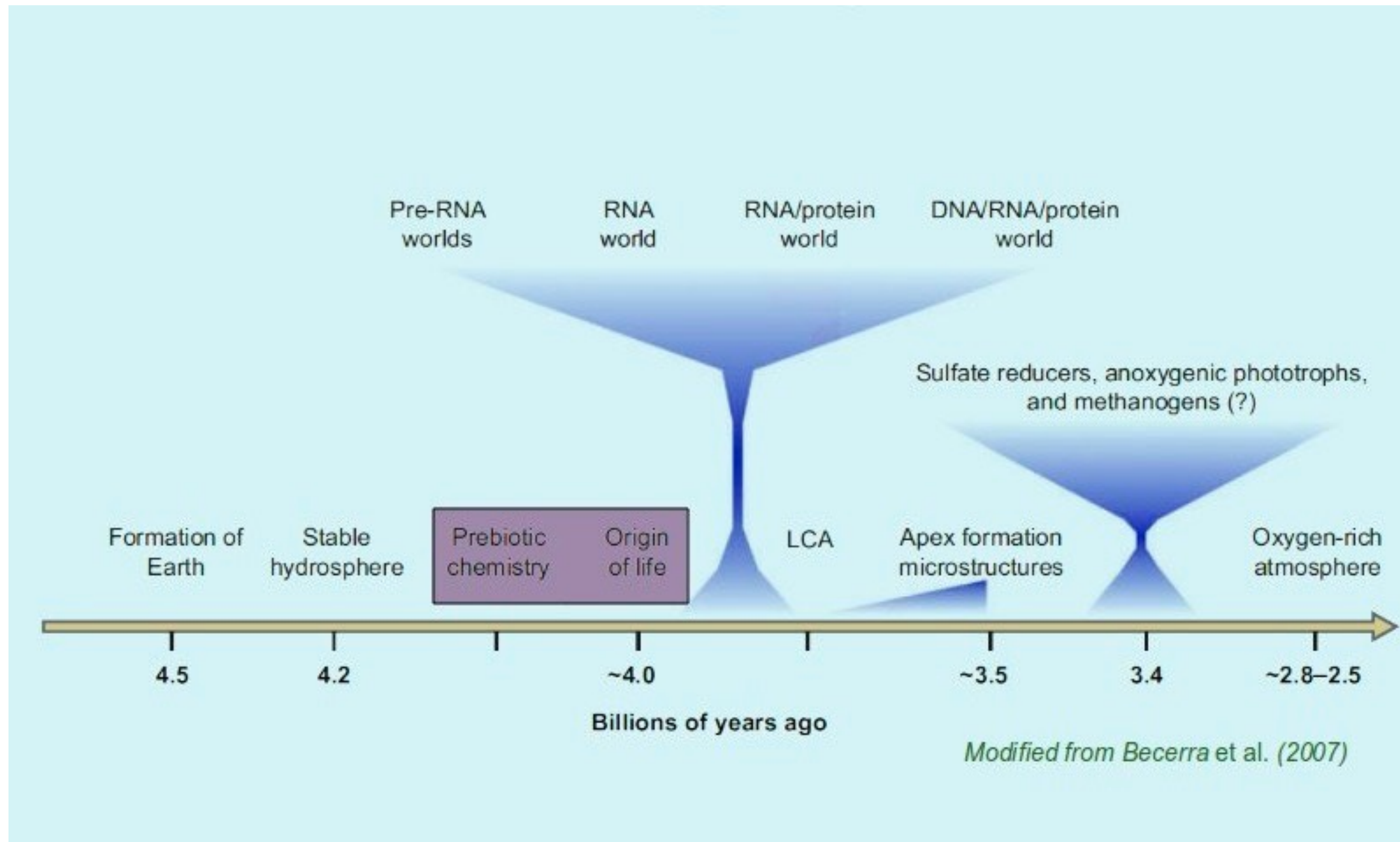
Geology: Evidence of life in Earth's oldest rocks

Abigail C. Allwood

Nature **537**, 500–501 (22 September 2016) | doi:10.1038/nature19429



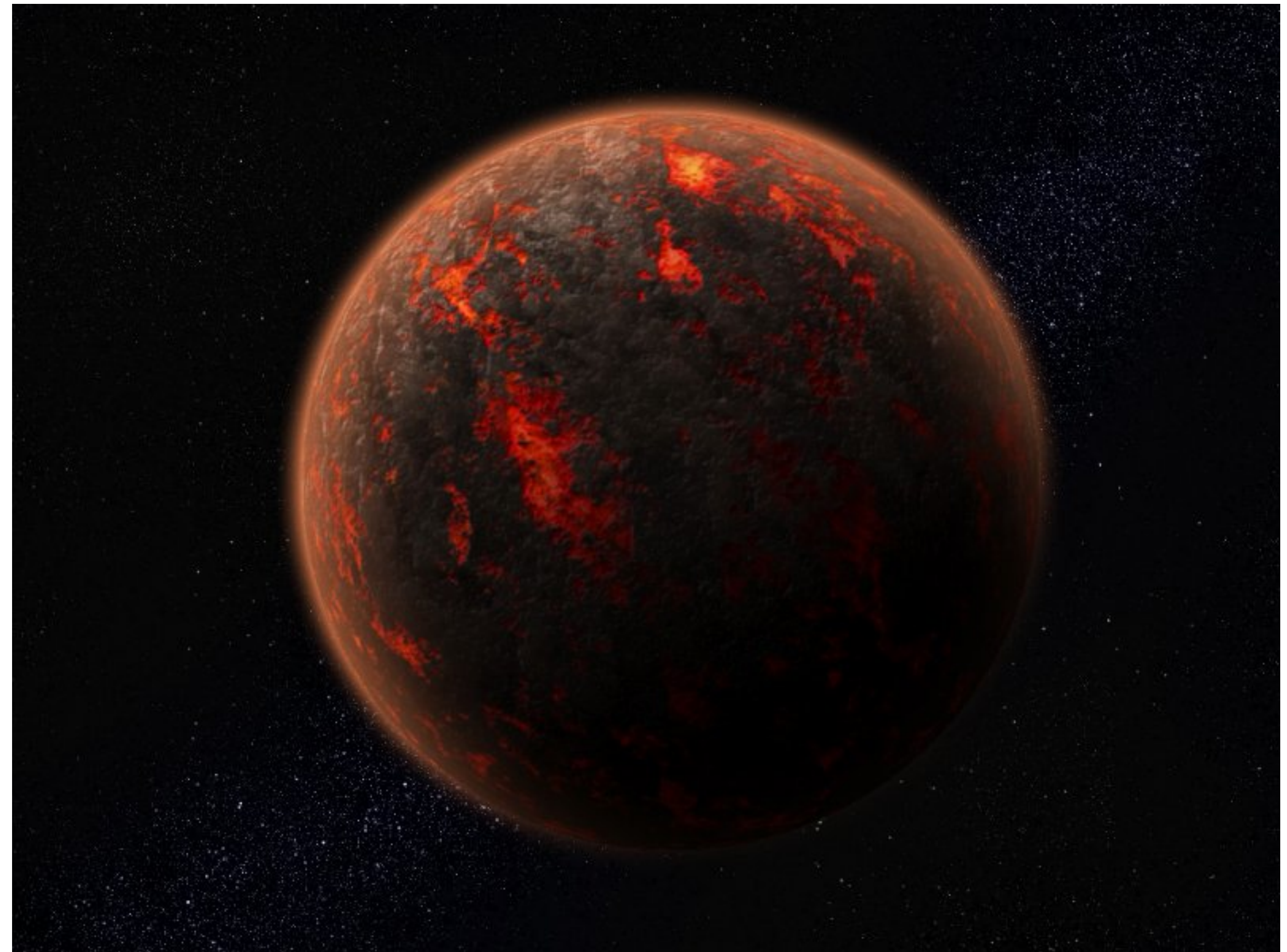
Prehistoria życia



Początki Ziemi

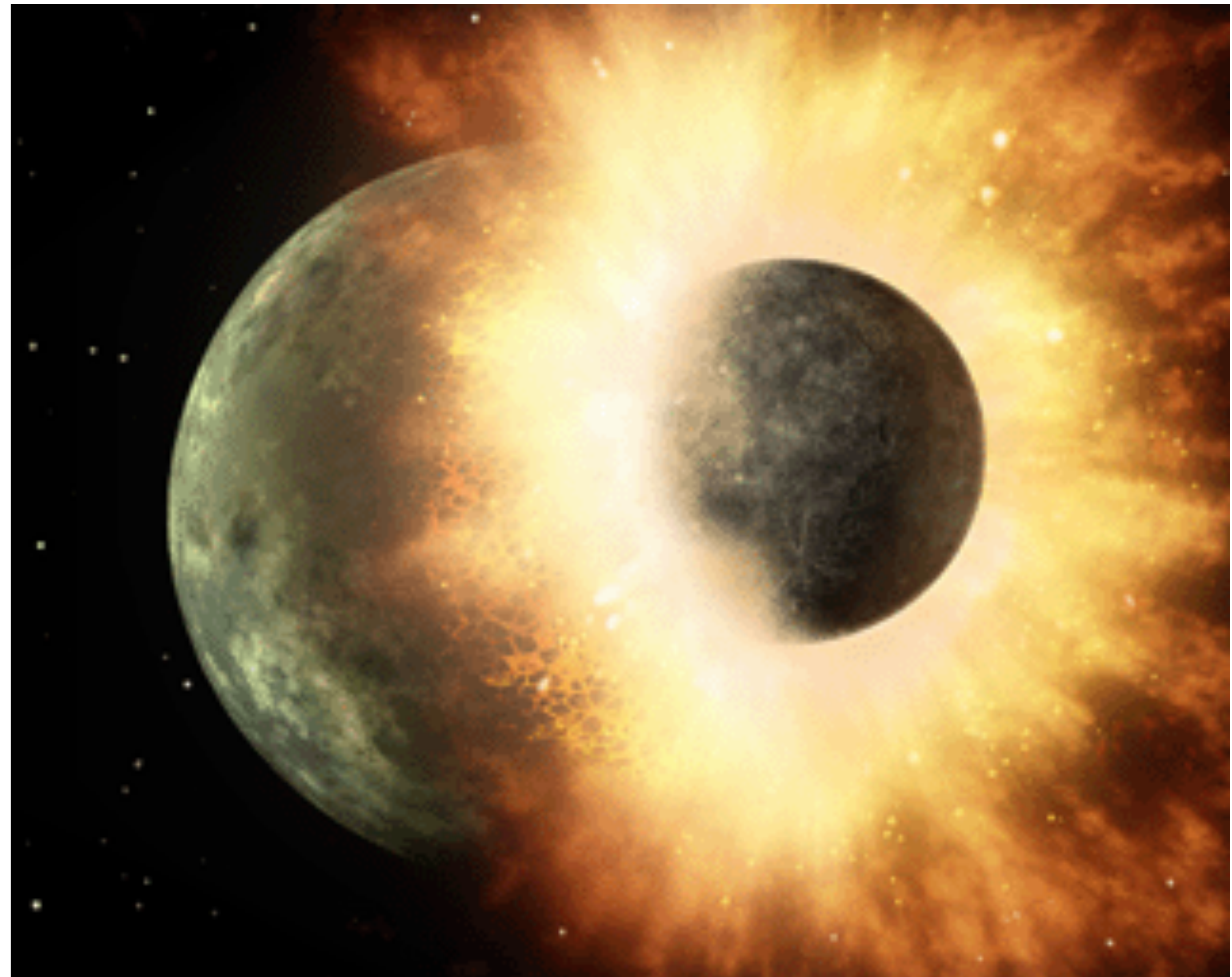
4,5 miliarda lat temu, ok milion lat po utworzeniu Układu Słonecznego

- brak atmosfery, powierzchnia – płynna skała
- bombardowanie meteorytami



Początki Ziemi

- Zderzenie Ziemi z jej satelitą (Theia) doprowadziło do:
- powstania Księżyca
- nachylenia osi obrotu Ziemi
- zainicjowania cyrkulacji magmy pod skorupą (astenosfera)
- początek zjawisk wulkanicznych, tektonika płyt



Początki atmosfery

- Ostudzenie powierzchni Ziemi
- Emisja gazów w wyniku działalności wulkanicznej



skład gazów wulkanicznych:

| | |
|------------------|-----------|
| H ₂ O | 95% |
| CO ₂ | 1-2% |
| SO ₂ | 1,5-2,5% |
| H ₂ S | 0,01-0,1% |
| HCl | 0,005% |

+ obecny wcześniej azot

Pierwotna atmosfera

- Po ochłodzeniu powierzchni woda przeszła do fazy ciekłej
- Część CO_2 osadziła się w minerałach
- Związki węgla, siarki, fosforu rozpuszczają się w wodzie
- Atmosfera beztlenowa, bogata w azot, dwutlenek węgla, związki siarki

Eksperyment Millera - Ureya

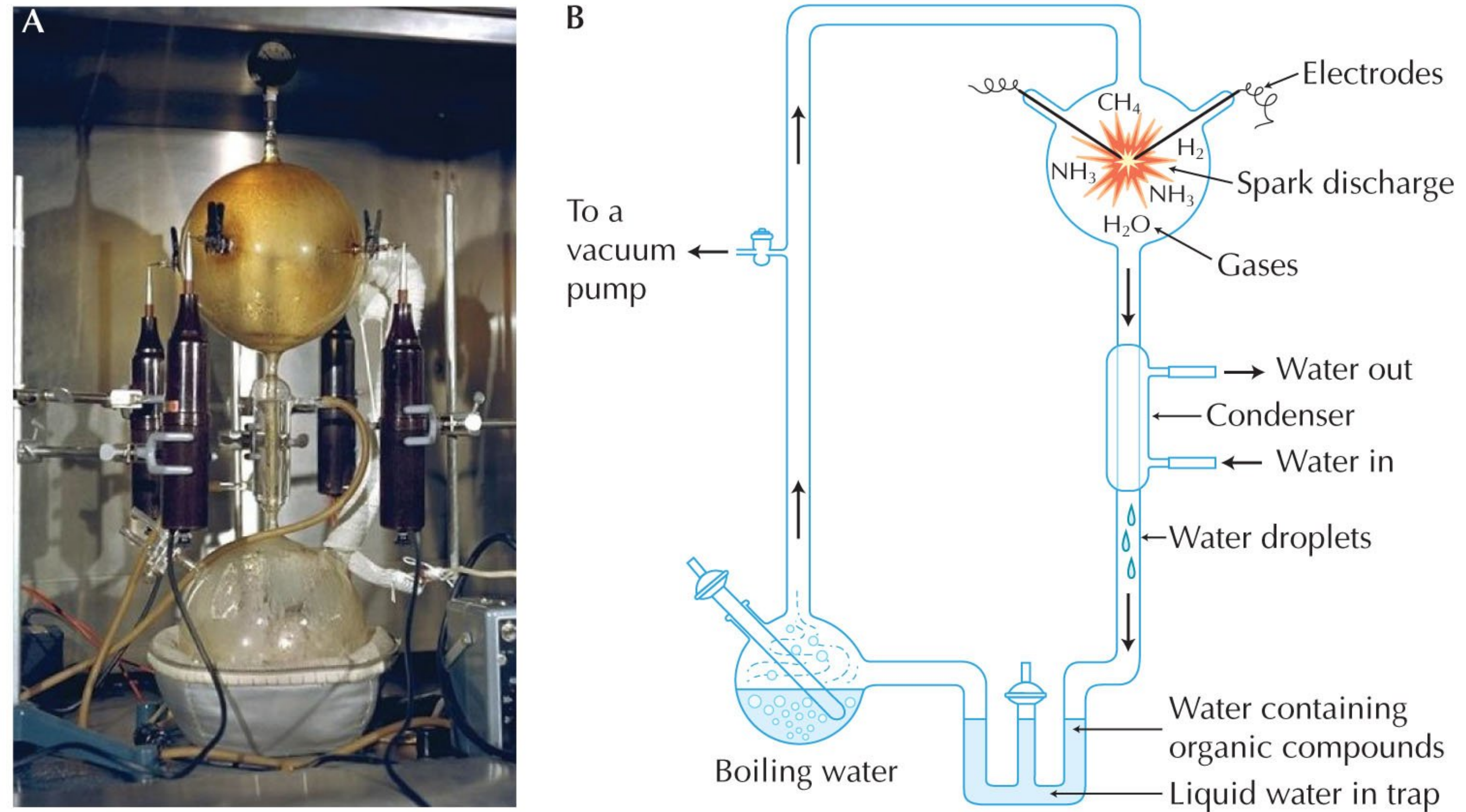
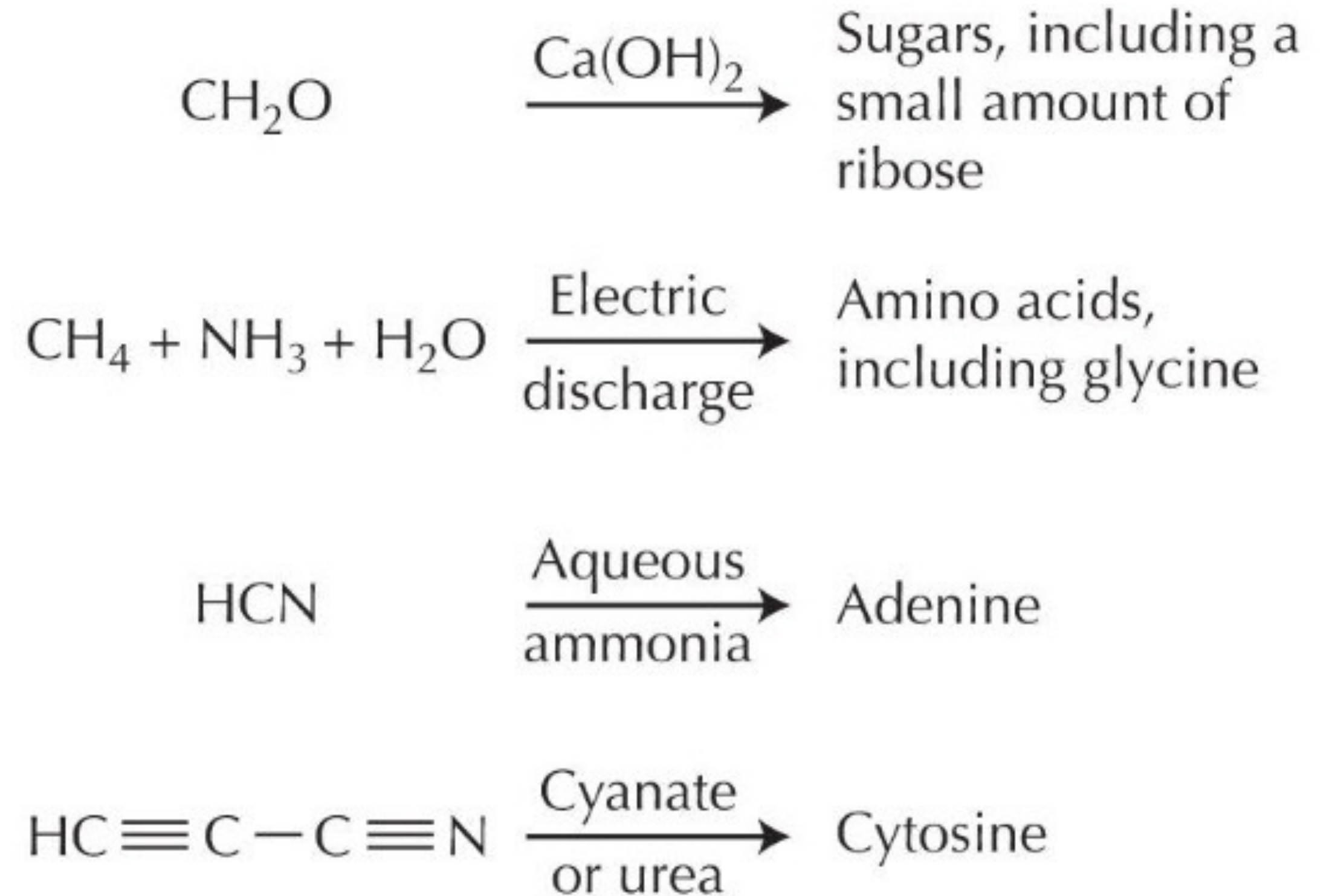


FIGURE 4.6. The apparatus used in the Miller–Urey experiments. (A) Recreation of the original apparatus. (B) Diagram of the apparatus.

4.6A, photo courtesy of NASA

Prebiotyczna synteza związków organicznych

- Liczne proste związki organiczne mogły powstać w warunkach wczesnej Ziemi



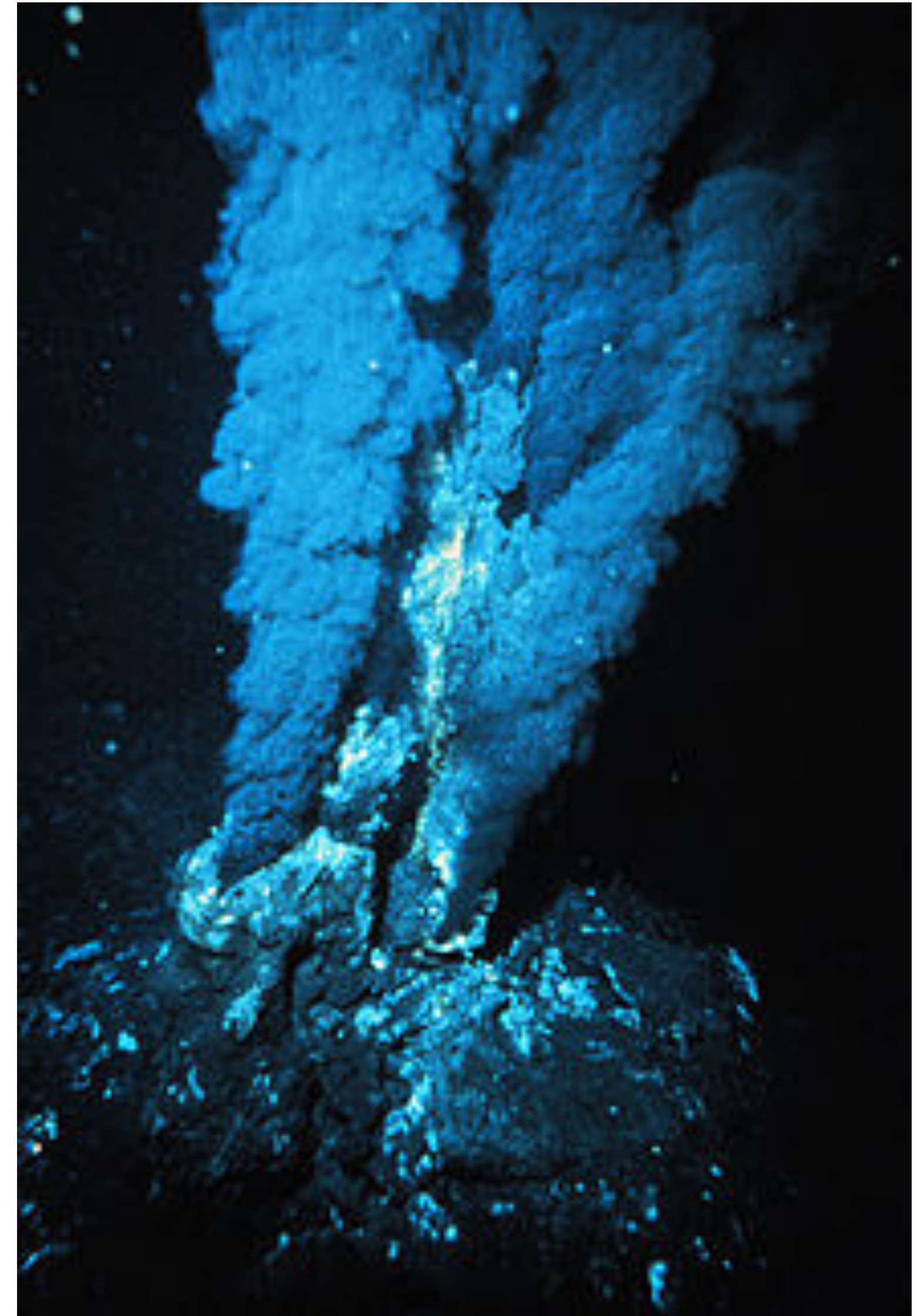
4.8, redrawn from Orgel L.E., *Trends Biochem. Sci.* **23**: 491–495, © 1998 Elsevier

Chemia prebiotyczna

- Tzw. “ciepły bulion” (Oparin, Haldane, Miller-Urey)
- Zwiększenie stężenia prekursorów (uzyskanych przez wyładowania elektryczne w atmosferze redukującej) przez cykliczne wysychanie płytkich zbiorników
- Problem: pierwotna atmosfera nie była aż tak silnie redukująca (dużo N_2 i CO_2)

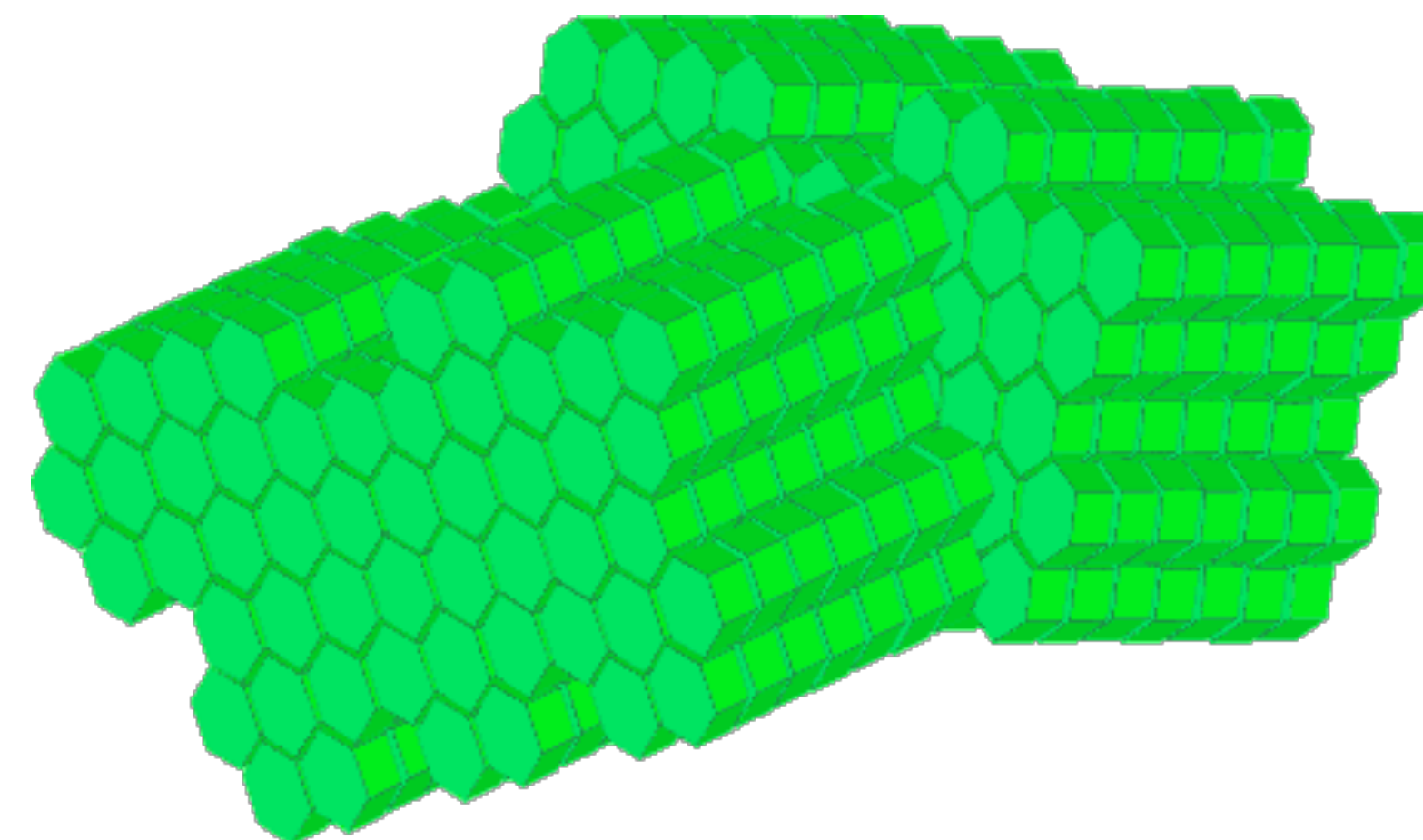
Chemia prebiotyczna

- kominy hydrotermalne
 - Günter Wächtershäuser – świat żelazowo-siarczkowy
- Minerale (iły) jako pierwotne replikatory
 - Alexander G. Cairns-Smith – “genetic takeover”



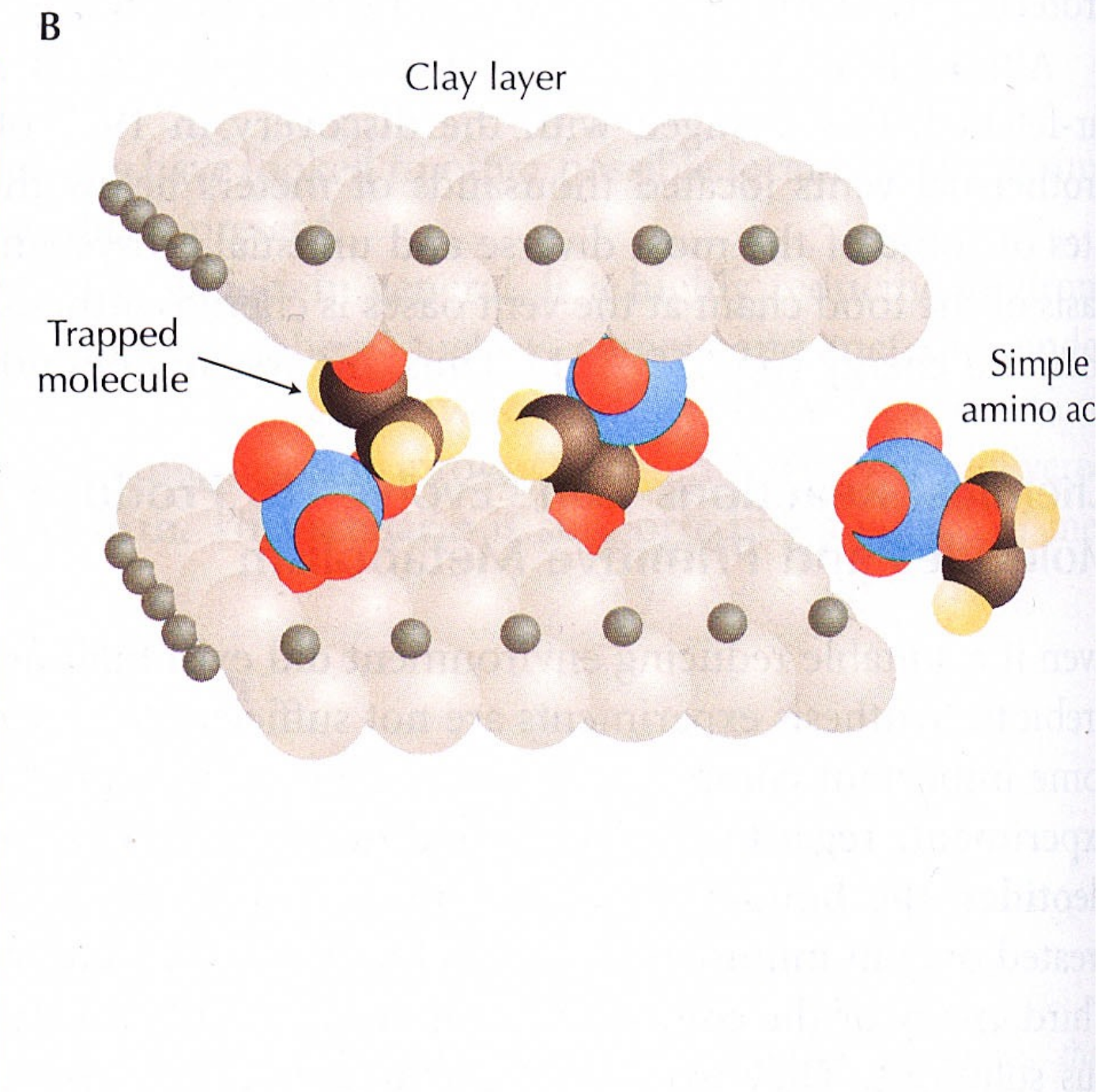
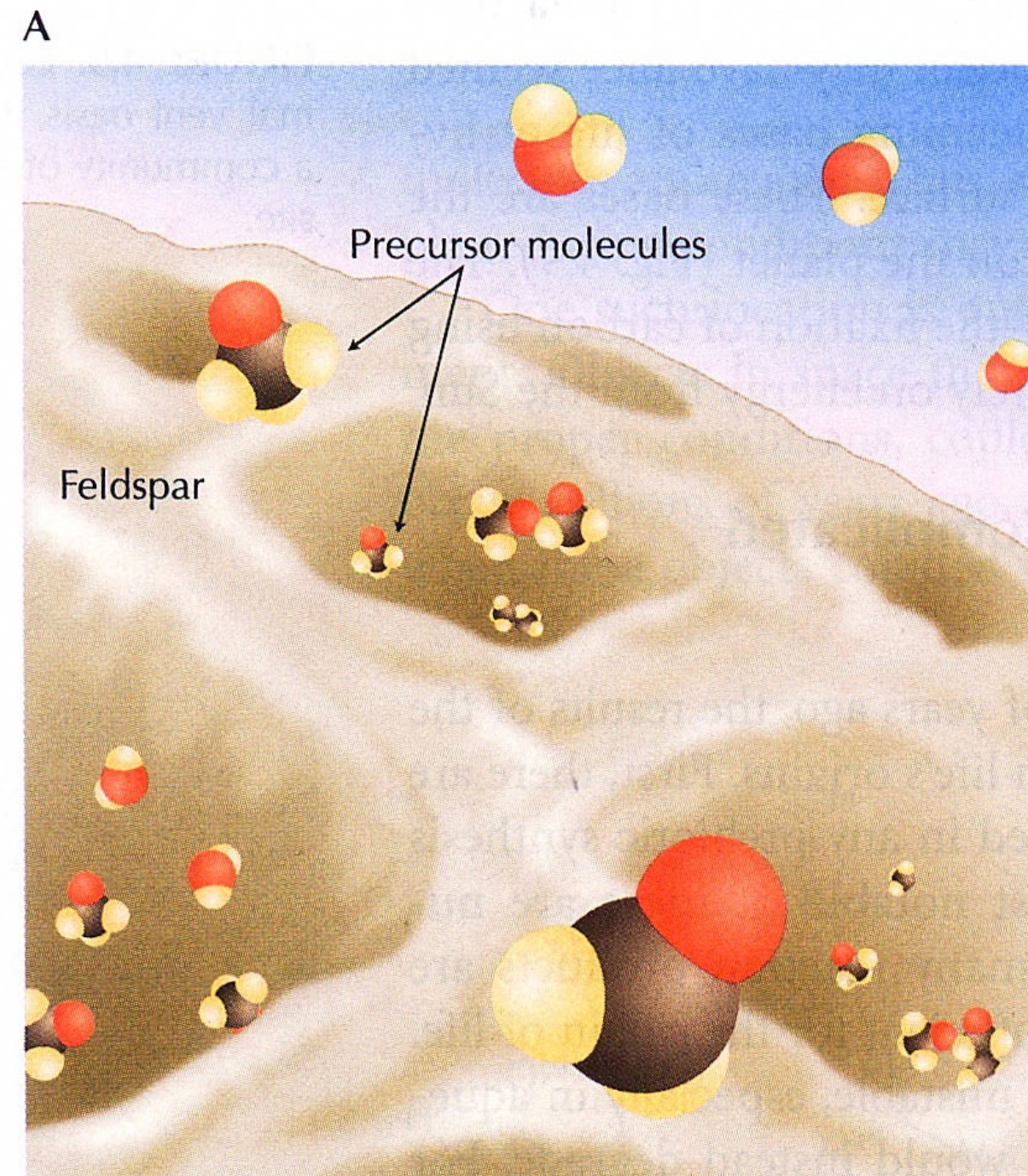
Życie mineralne?

- Alexander G. Cairns-Smith – “genetic takeover”
- Życie oparte na minerałach (glinokrzemiany, ility)
- Replikacja przez odtwarzanie zaburzeń struktury kryształu przy osadzaniu się na nim kolejnej warstwy
- Metabolizm – kataliza na powierzchni kryształu
- Później “przejęcie władzy” przez kwasy nukleinowe



Kataliza powierzchniowa

- Niezależnie od modelu, powierzchnia minerałów jest dobrym środowiskiem do syntezy
- mikronierówności, ładunki powierzchniowe
- Czy mogły tworzyć cykle proto-metaboliczne?



Abiotyczna synteza polinukleotydów

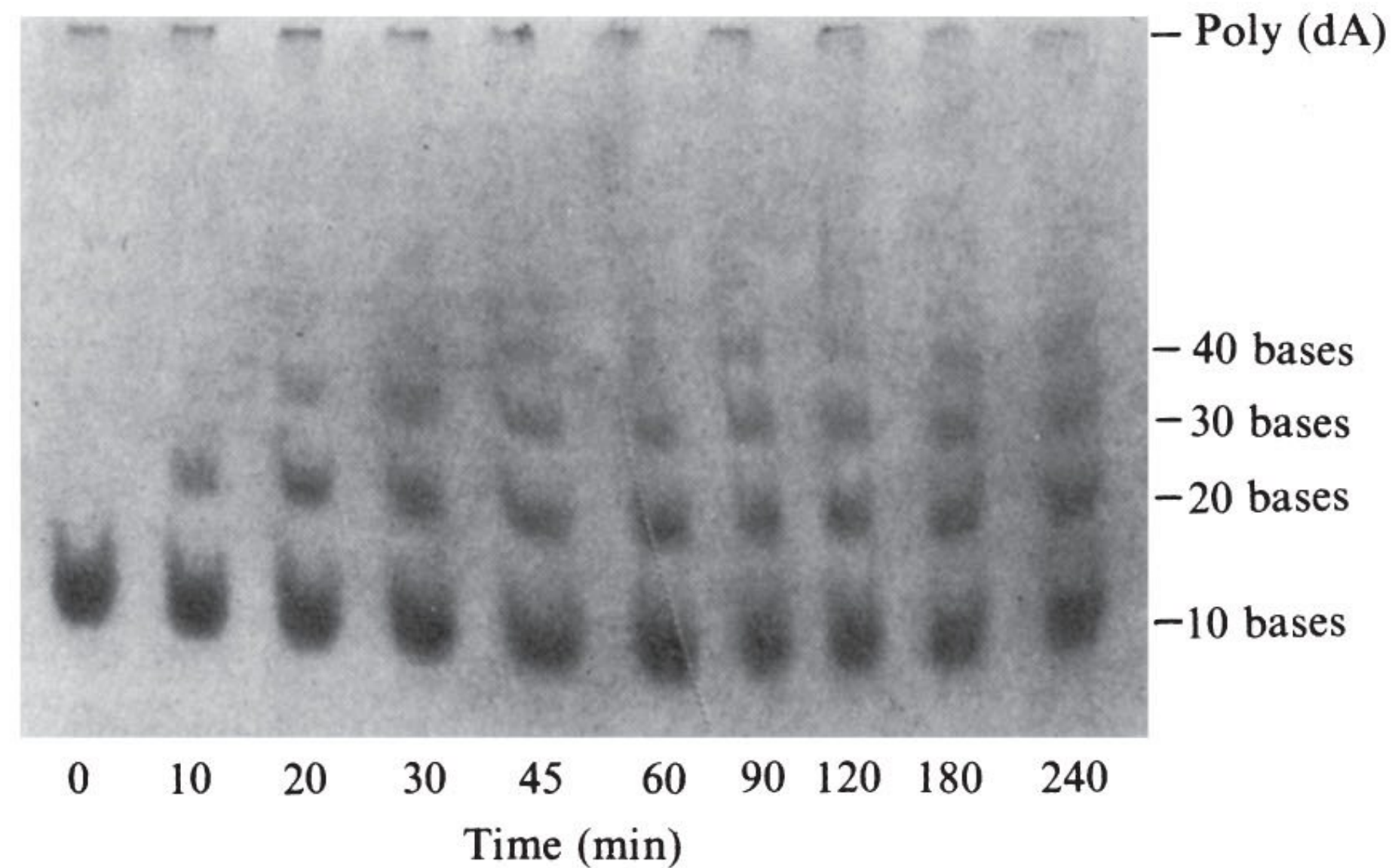


FIGURE 4.12. Polymerization of oligonucleotides by UV cross-linking. A solution of a DNA oligomer (10 bases long) was exposed to UV irradiation for different lengths of time. Longer exposures generated longer polymers.

4.12, reprinted from Lewis R.J. et al., *Nature* **298**: 393–396, © 1982 Macmillan, www.nature.com

Evolution © 2007 Cold Spring Harbor Laboratory Press

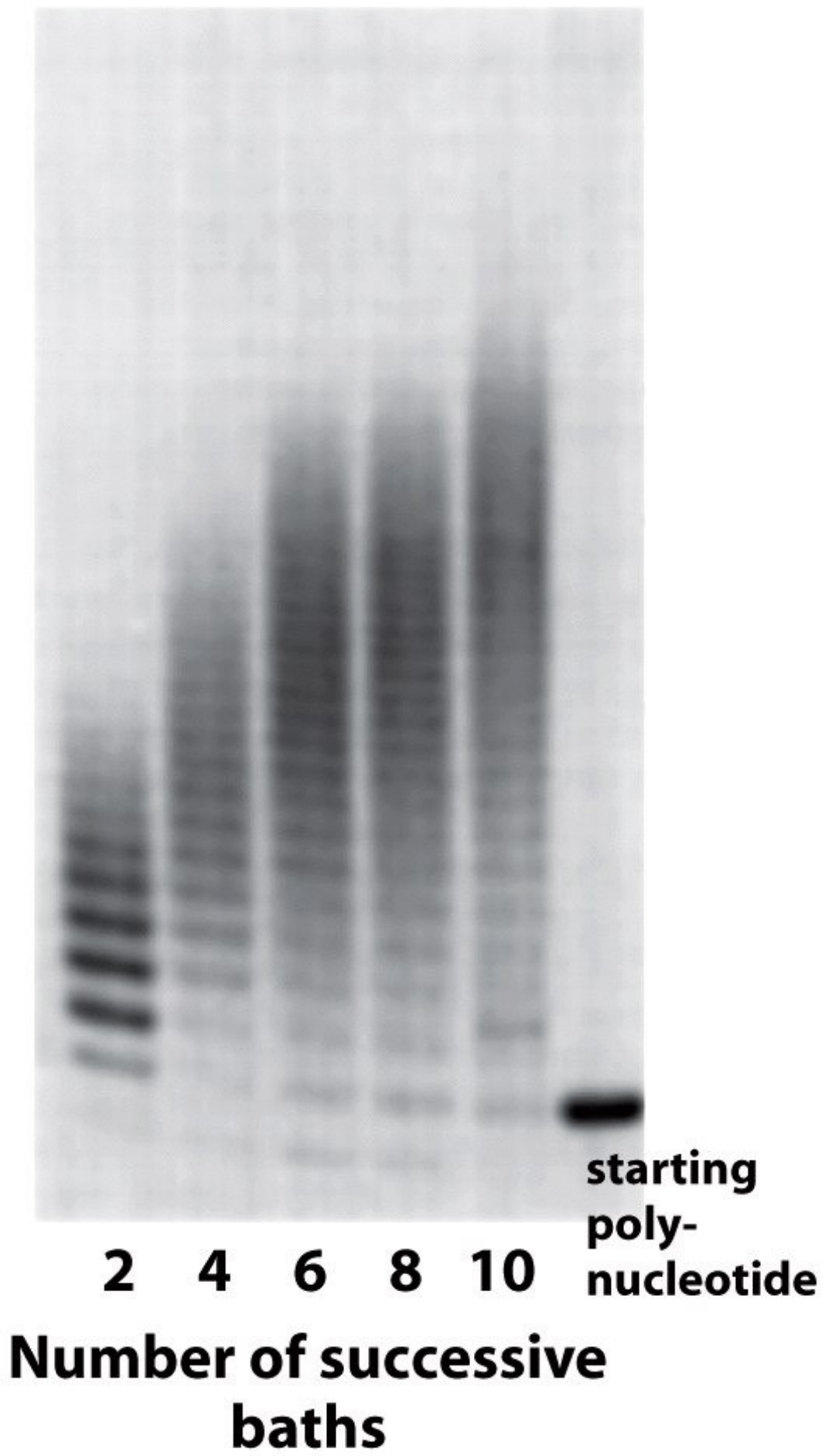


Figure 17-13 *Evolutionary Analysis, 4/e*
© 2007 Pearson Prentice Hall, Inc.

Ligacja polinukleotydów na katalizatorze
glinokrzemianowym

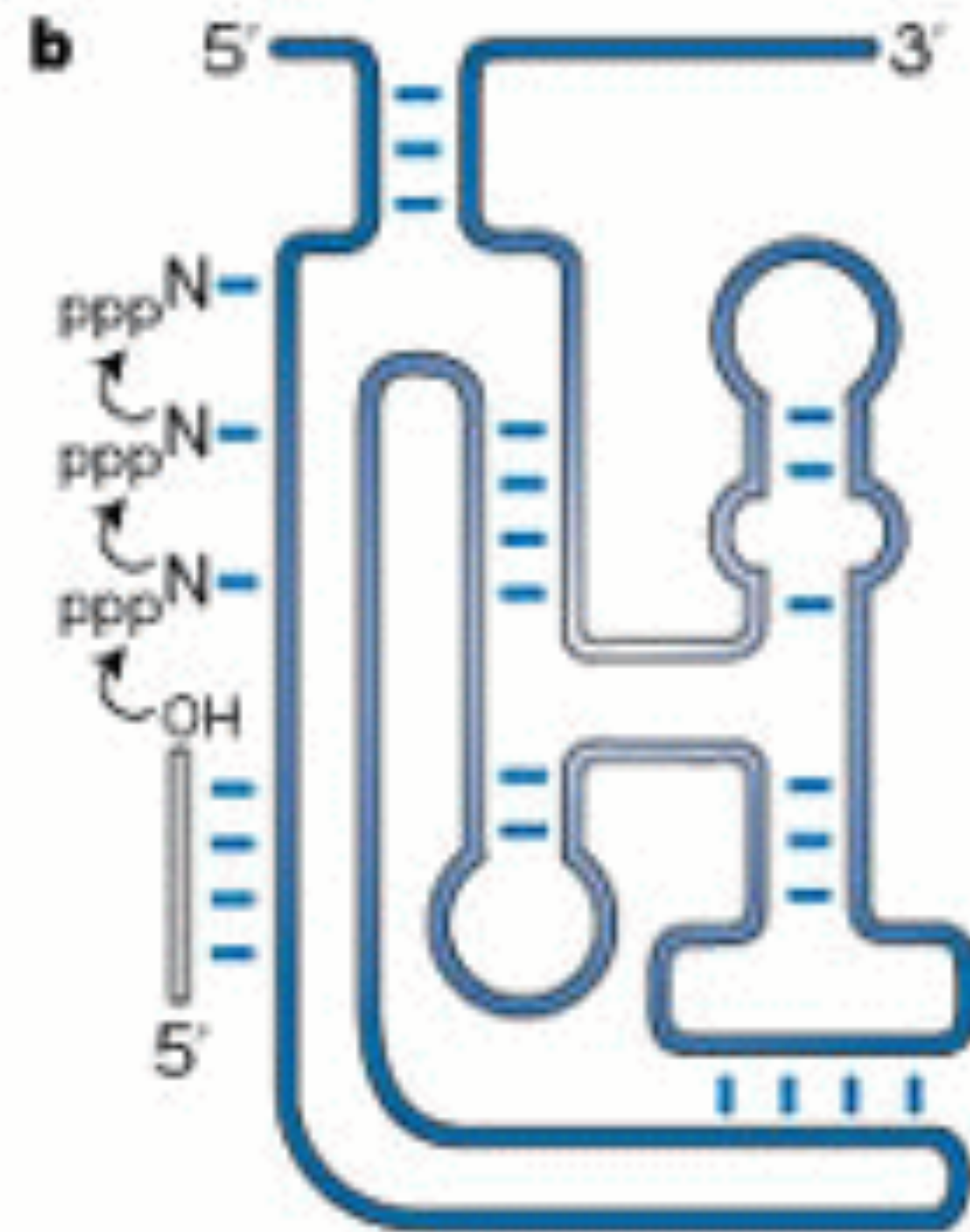
Co było najpierw?

- Metabolizm (Oparin, Dyson)
 - Zależny od informacji genetycznej (kodowane enzymy)
- Replikacja (Eigen)
 - Zależna od metabolizmu (enzymy replikujące DNA)

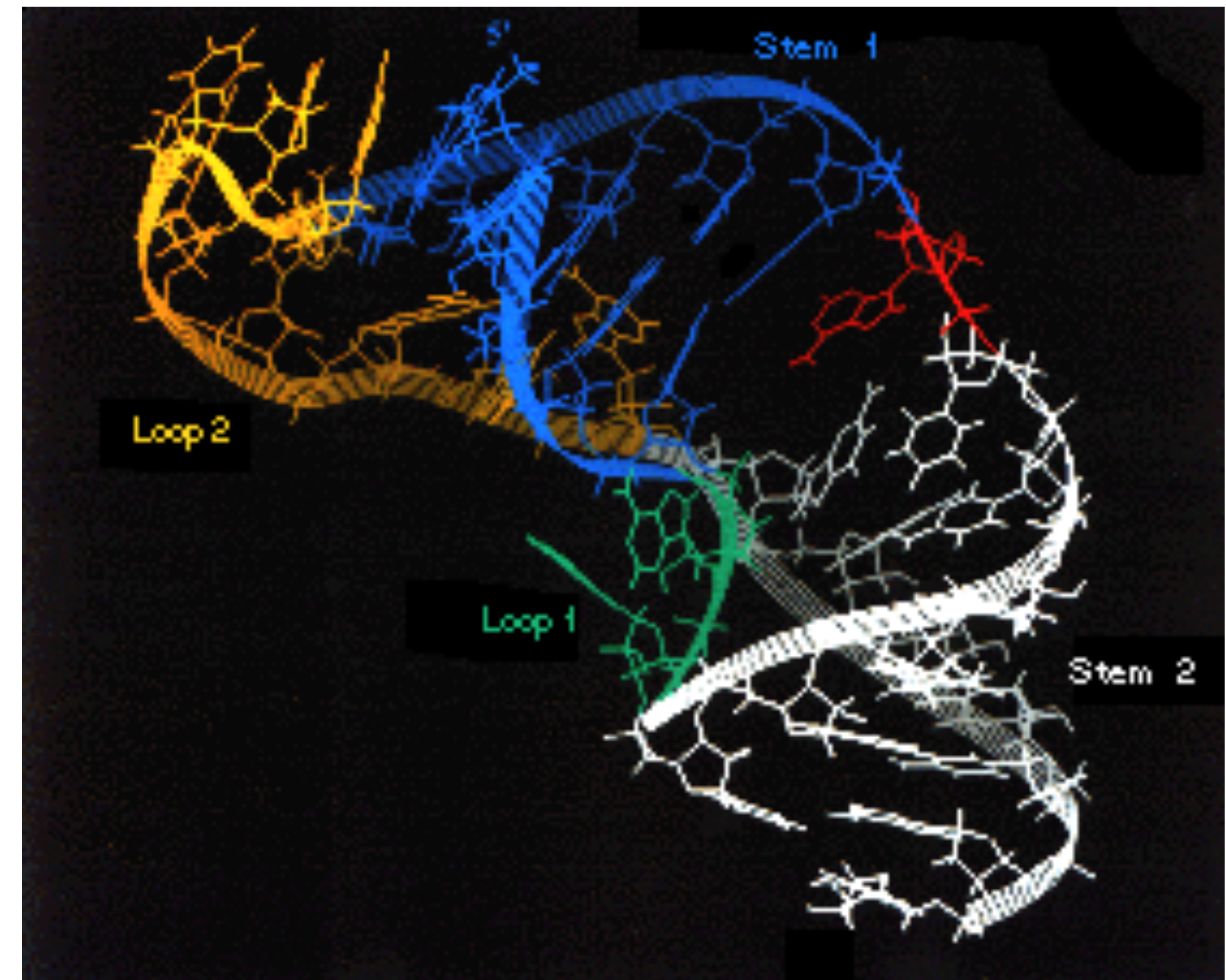
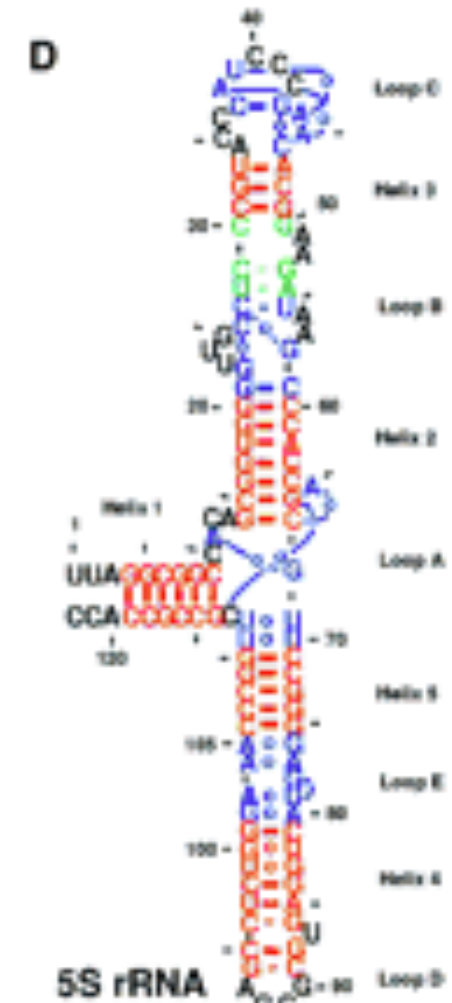
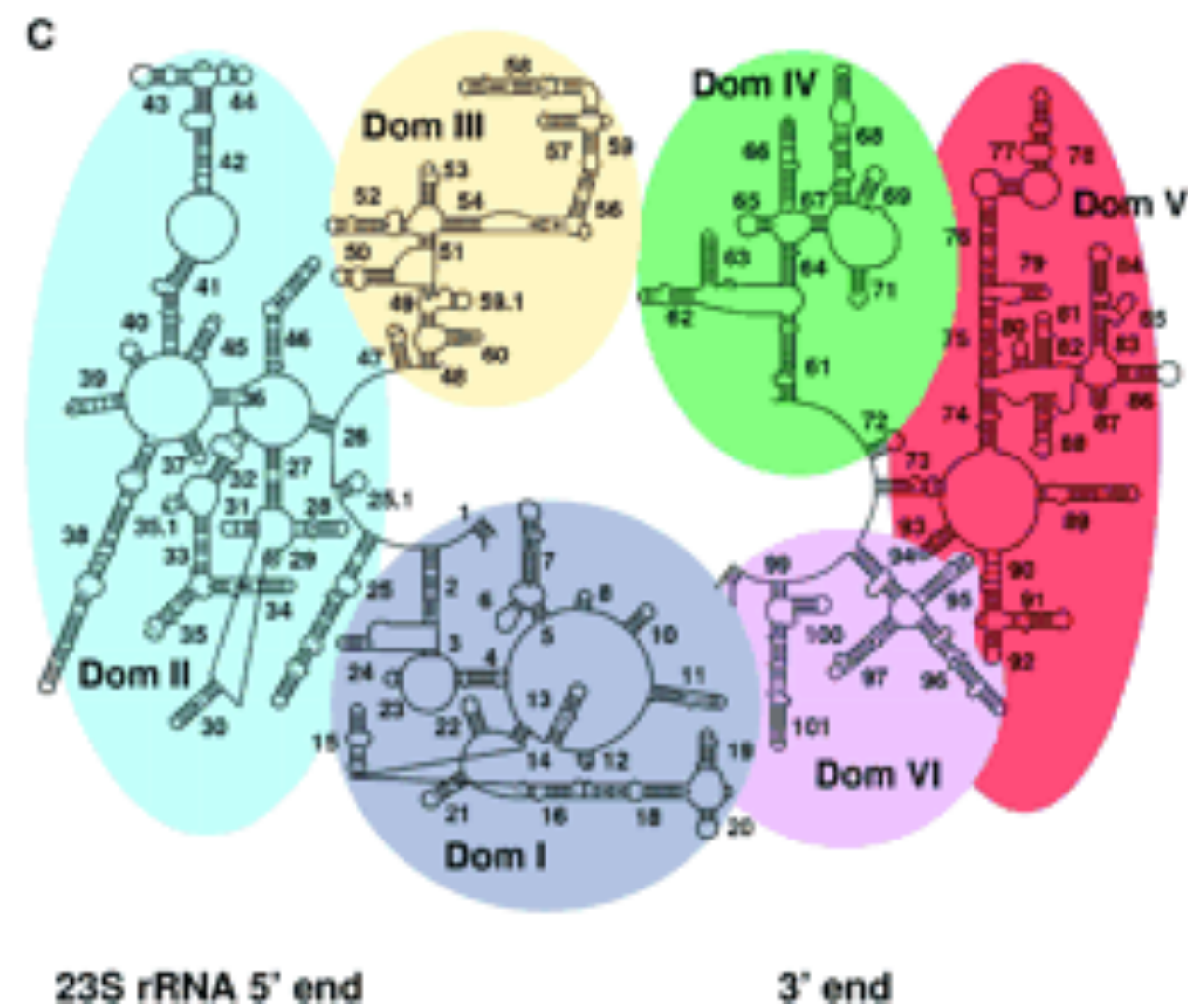
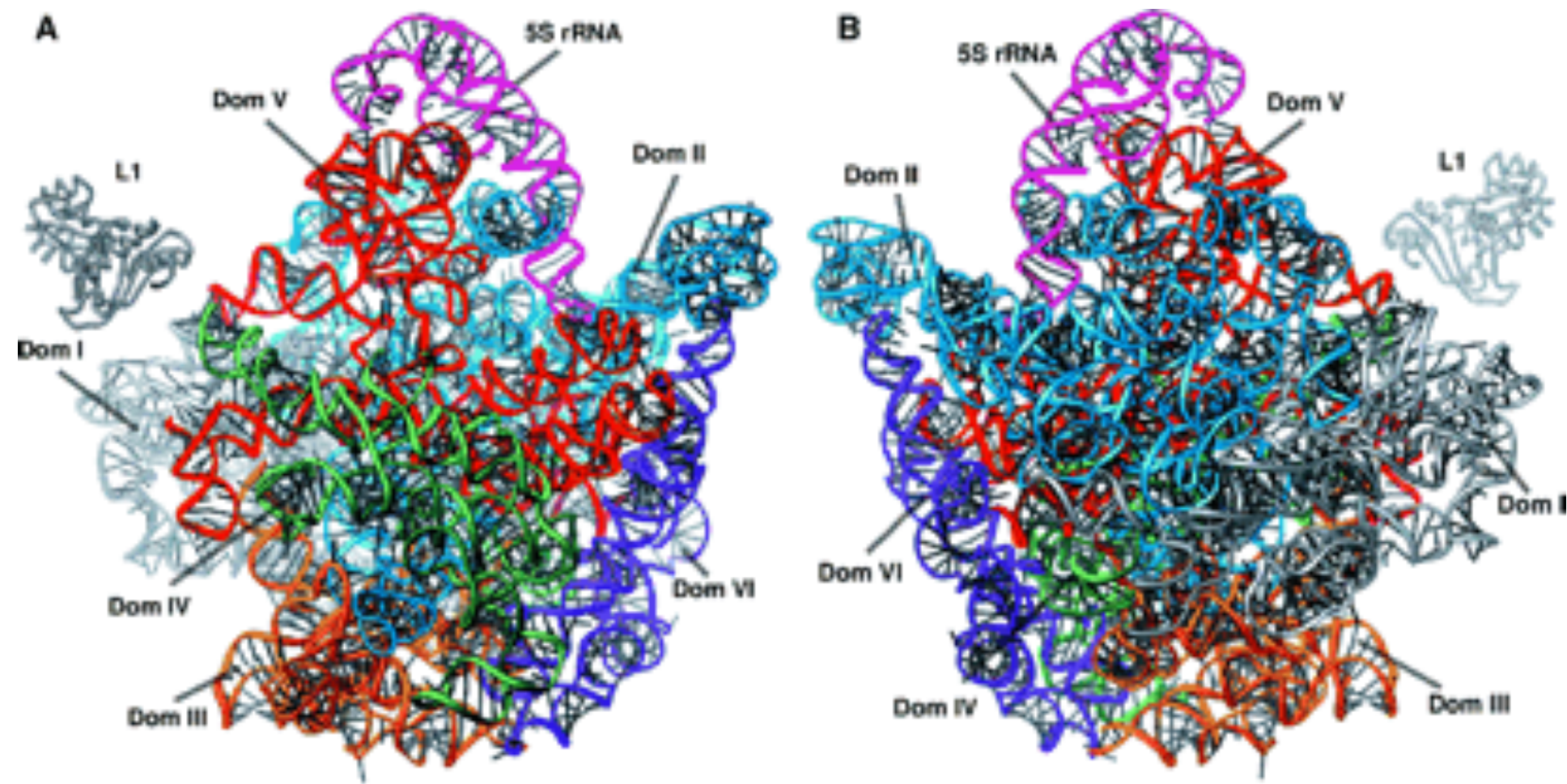


Świat RNA: metabolizm + replikacja

- RNA może wykazywać aktywność enzymatyczną (metabolizm)
- Rybozomy – enzymy o centrum katalitycznym zbudowanym z RNA
- W oparciu o jeden rodzaj cząsteczek można zbudować układ zdolny do replikacji (zasada komplementarności nukleotydów) i metabolizmu

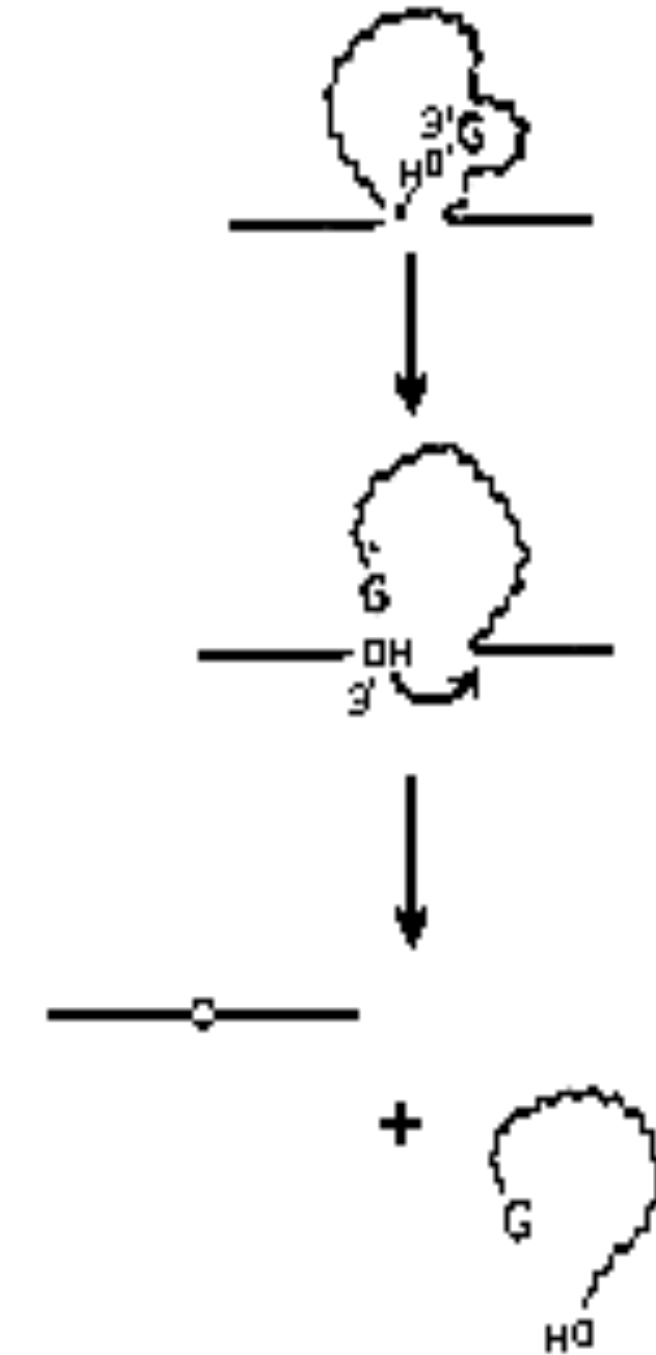
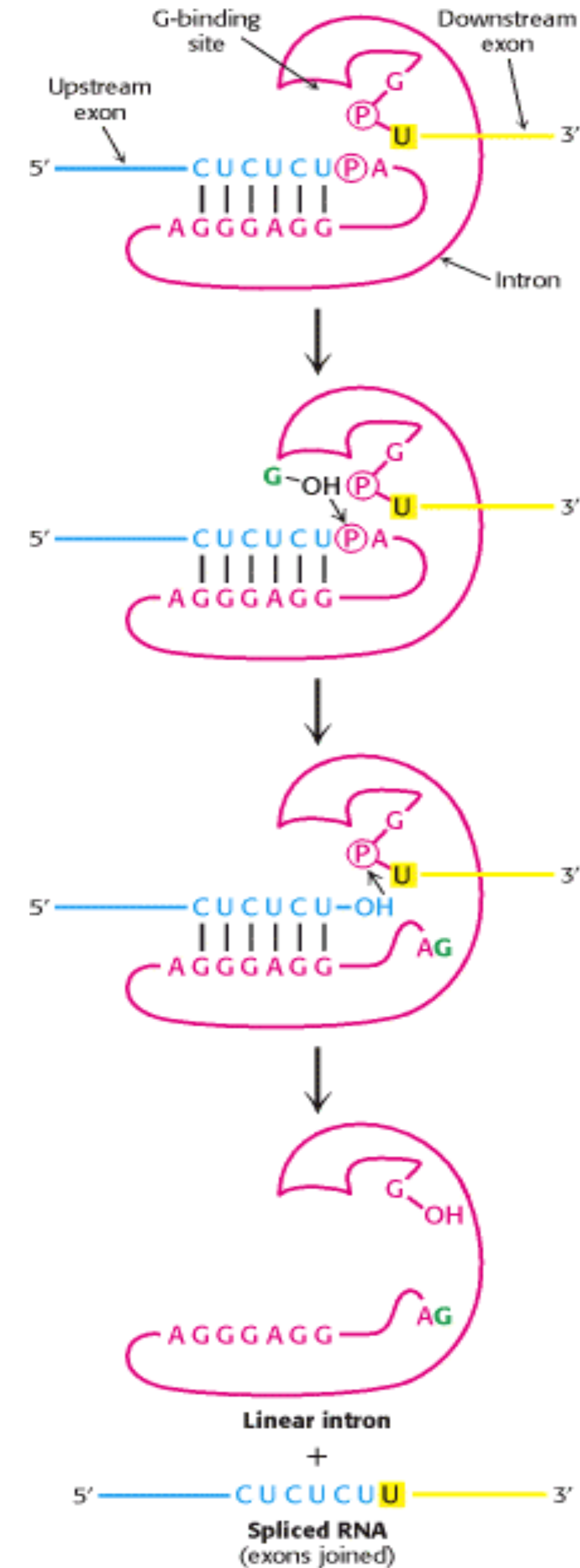


RNA może tworzyć różne struktury



RNA katalityczne

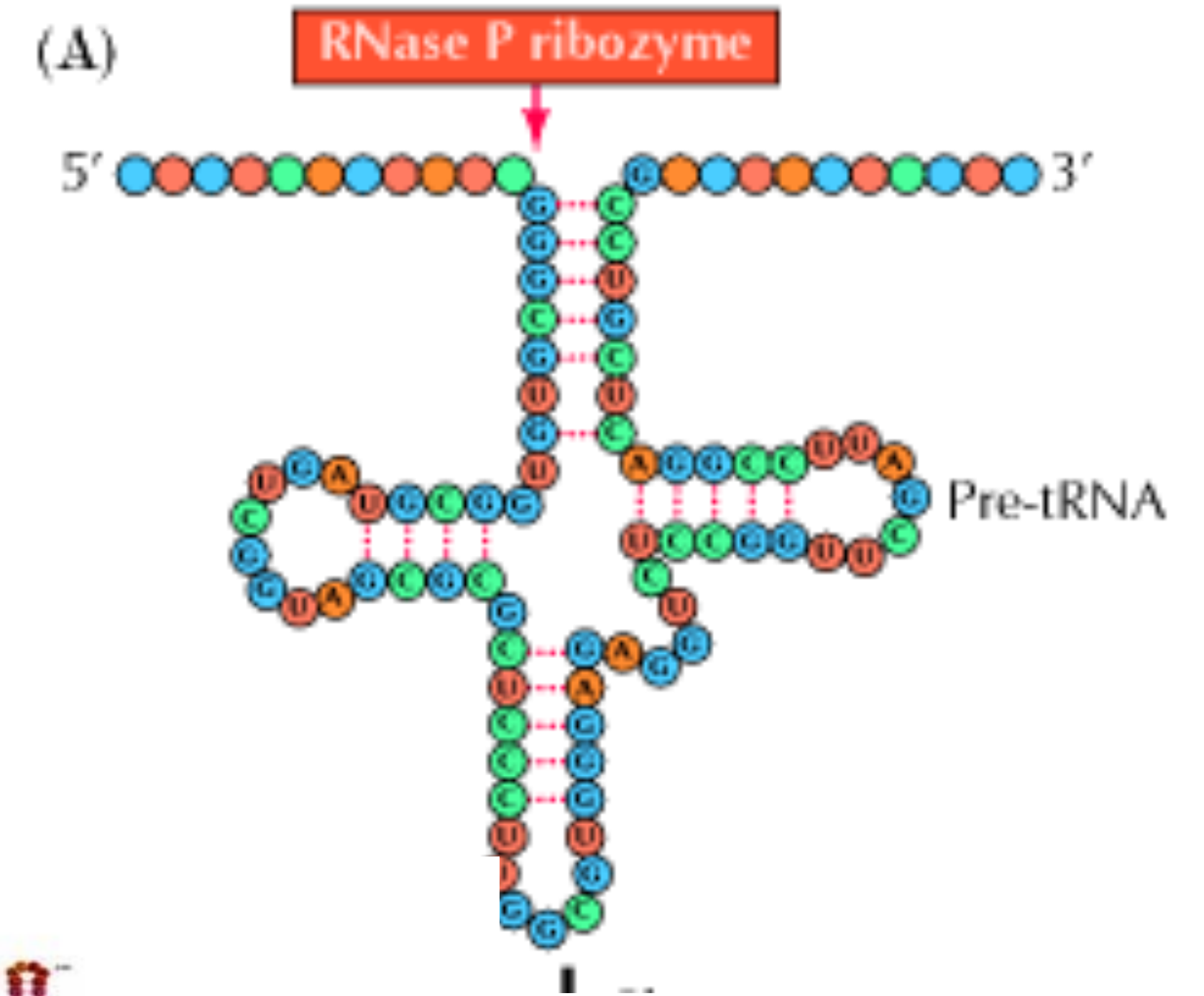
- Thomas Cech (1982) – intron w *Tetrahymena* sam się wycina



Nagroda Nobla 1989

RNA katalityczne

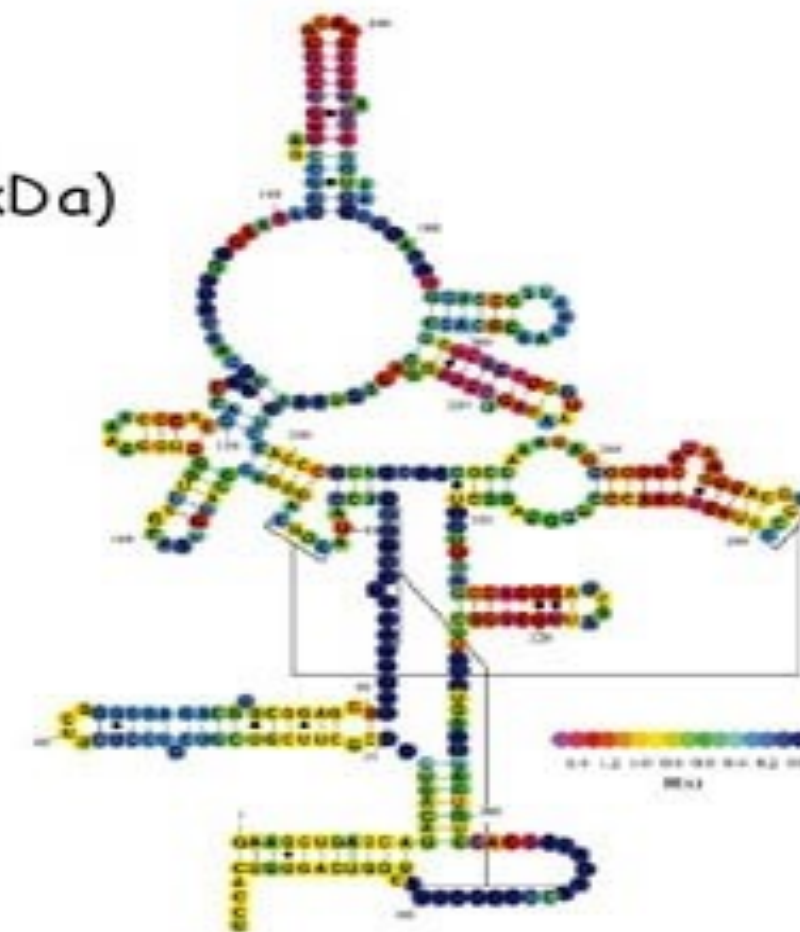
- Sidney Altman (1983) – RNaza P (enzym tnący prekursor tRNA) składa się z białka i RNA, to RNA jest katalizatorem



RNase P in Bacteria

1 large RNA
rnpB - 377nt (140kDa)

1 small protein
rnpA - 119aa (14kDa)

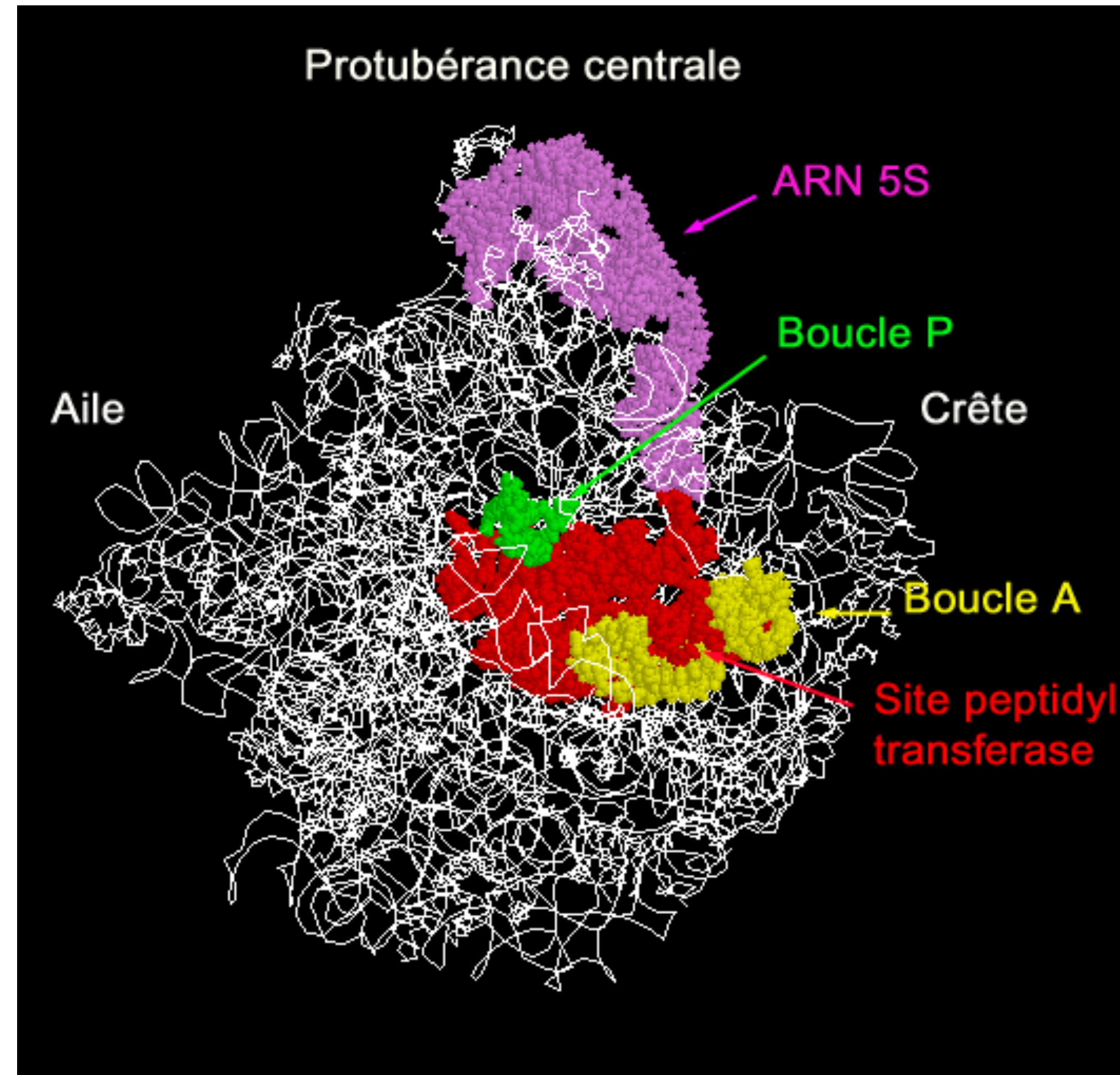
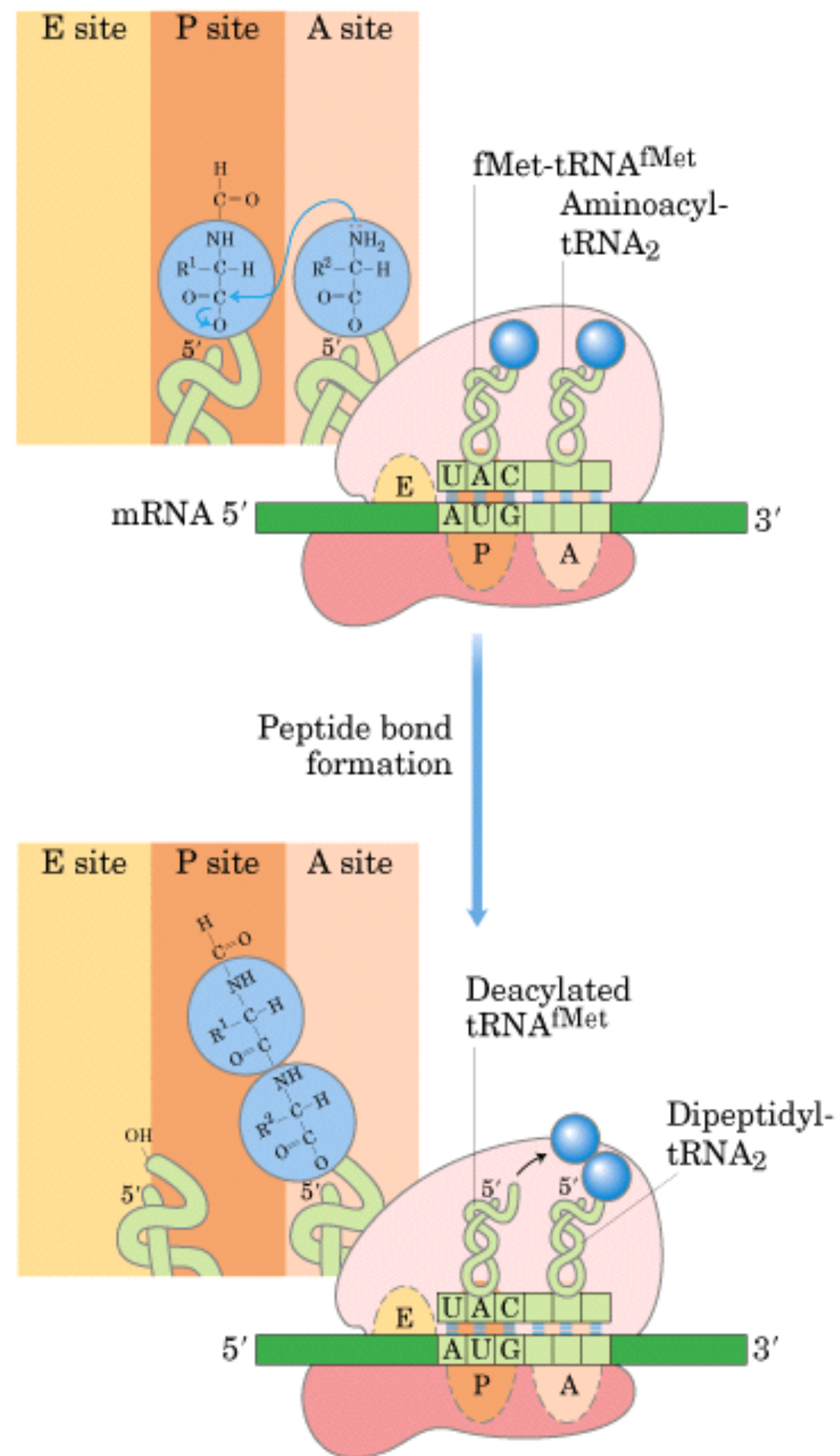


The RNA *by itself* is catalytically proficient *in vitro*



Nagroda Nobla 1989

RNA syntetyzuje białko



Aktywność peptydylotransferazy w rybosomie to rybozym

RNA wycina introny

doi:10.1038/nature12734

RNA catalyses nuclear pre-mRNA splicing

Sebastian M. Fica^{1,2*}, Nicole Tuttle^{3*}, Thaddeus Novak⁴, Nan-Sheng Li⁴, Jun Lu³, Prakash Koodathingal², Qing Dai³, Jonathan P. Staley² & Joseph A. Piccirilli^{3,4}

2013

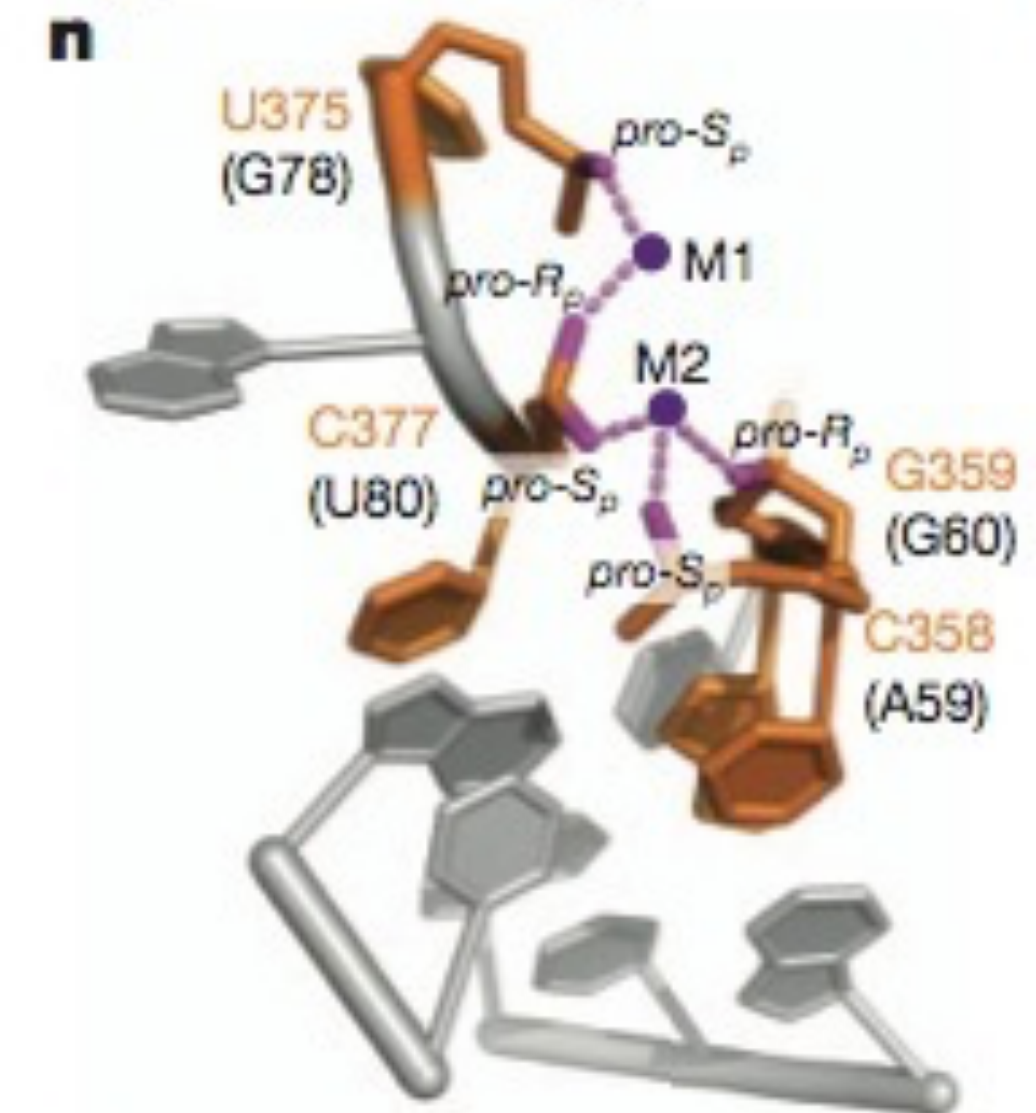
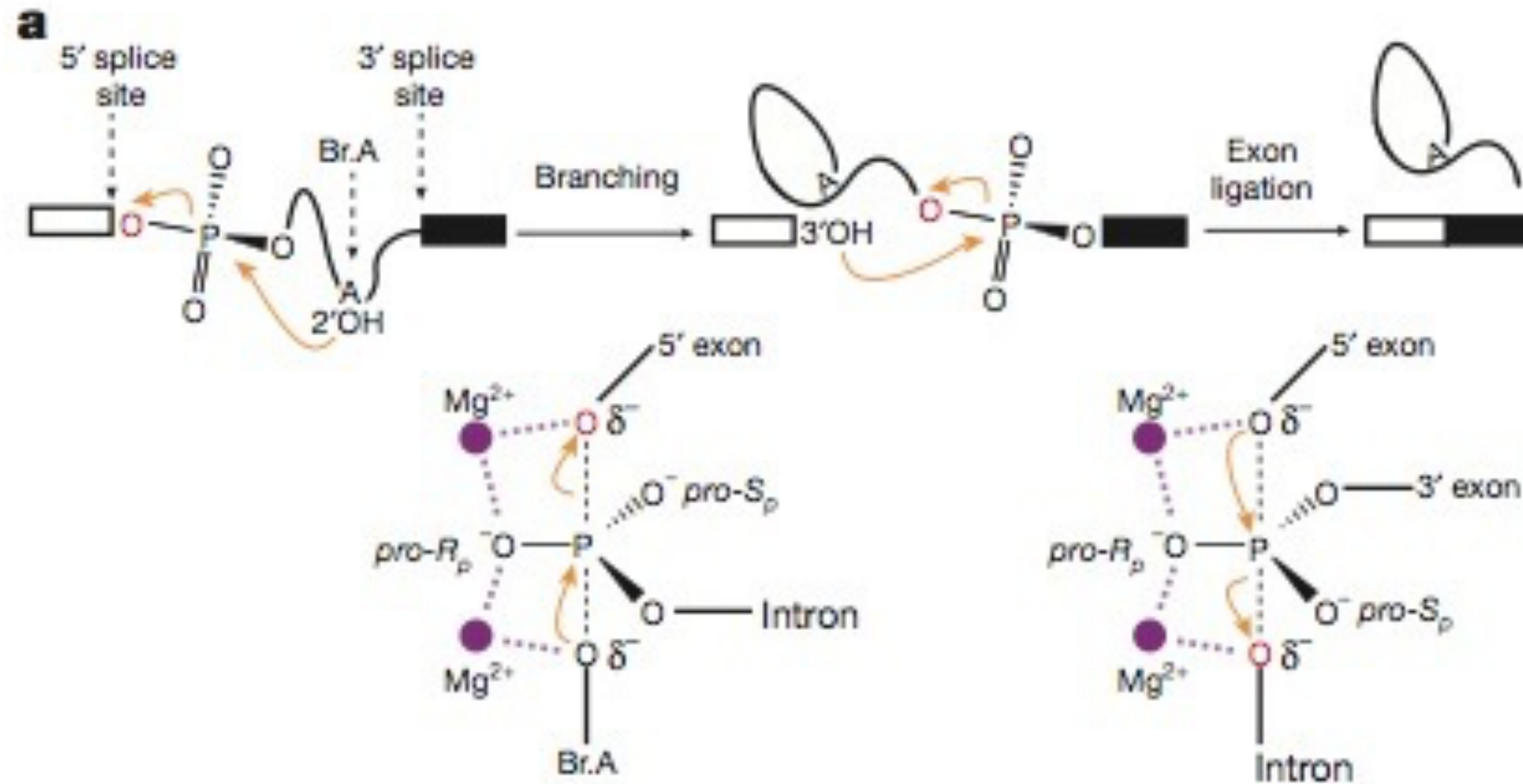
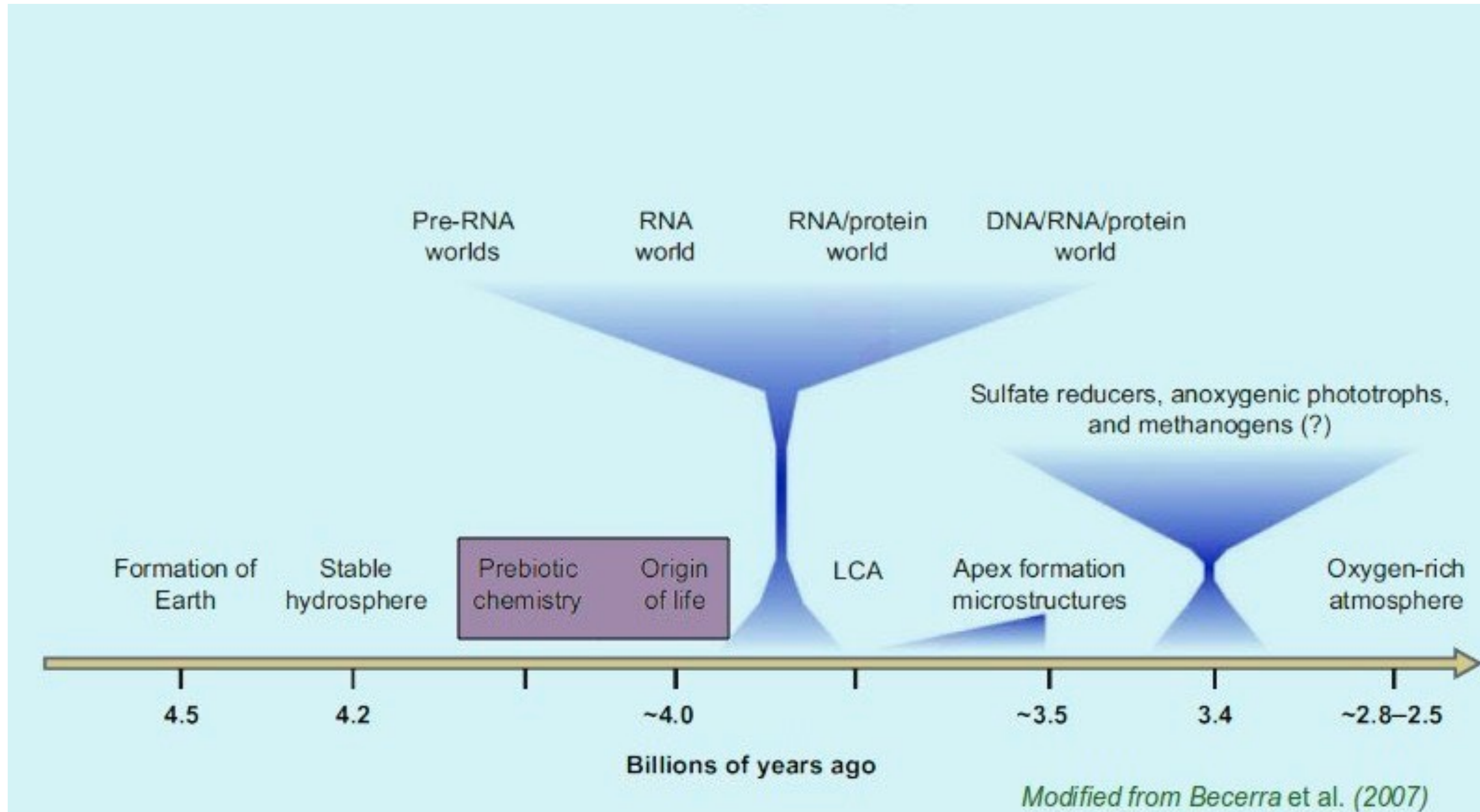


Figure 2 | U6 snRNA positions metals important for both steps of splicing.

Prehistoria życia



Świat RNA

- RNA pełniące rolę metaboliczną (enzymy) i informacyjną (matryca)
- Początki oddziaływań RNA-aminokwasy – początki kodu genetycznego

Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions

Matthew W. Powner¹, Béatrice Gerland¹ & John D. Sutherland¹

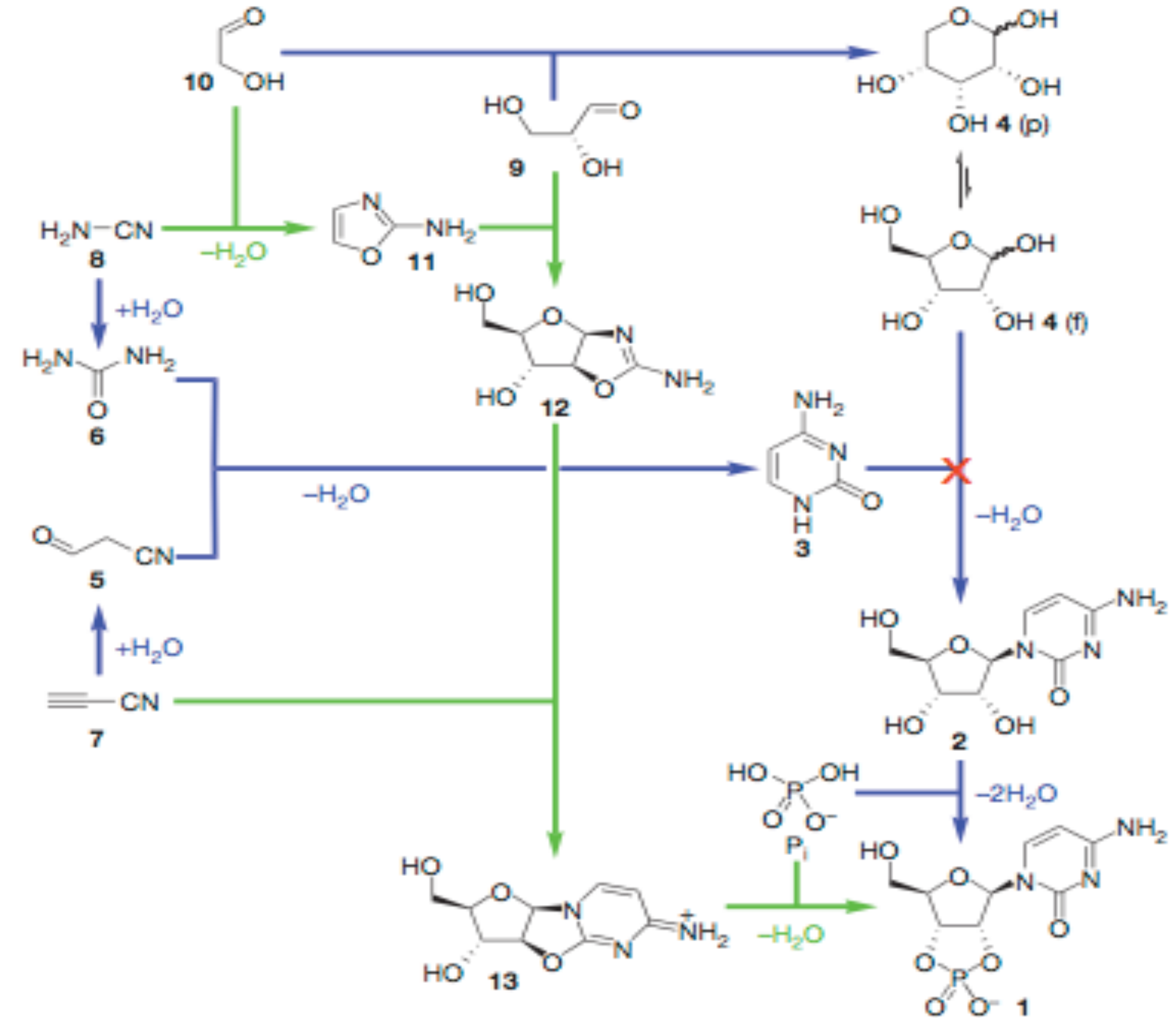


Figure 1 | Pyrimidine ribonucleotide assembly options. Previously assumed synthesis of β -ribocytidine-2',3'-cyclic phosphate **1** (blue; note the failure of the step in which cytosine **3** and ribose **4** are proposed to condense together) and the successful new synthesis described here (green). p, pyranose; f, furanose.

Rybozym zdolny do syntezy RNA

- Wyselekcjonowany w laboratorium
- Wciąż nie autonomiczna replikaza

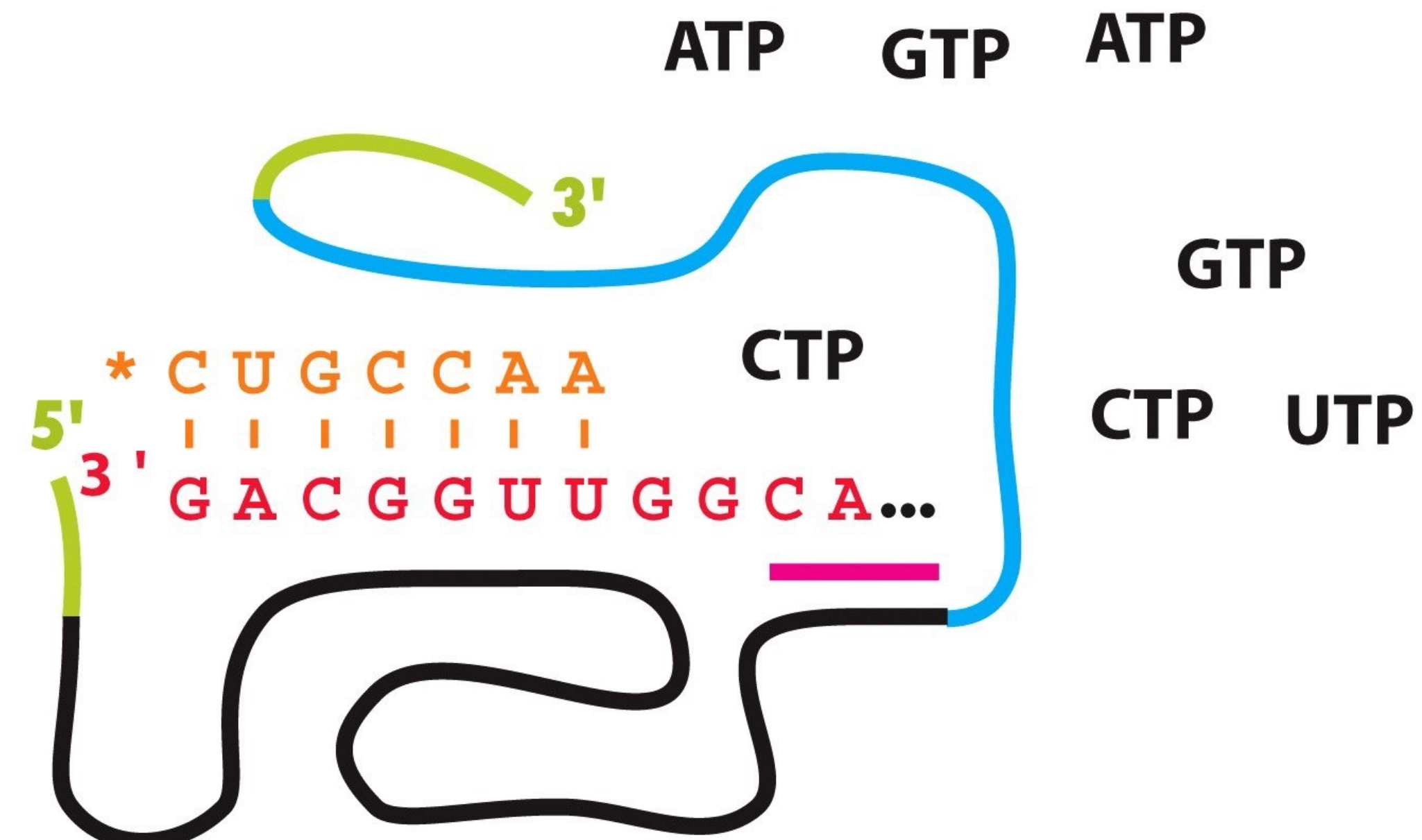


Figure 17-8a Evolutionary Analysis, 4/e © 2007 Pearson Prentice Hall, Inc.



Figure 17-8b Evolutionary Analysis, 4/e © 2007 Pearson Prentice Hall, Inc.

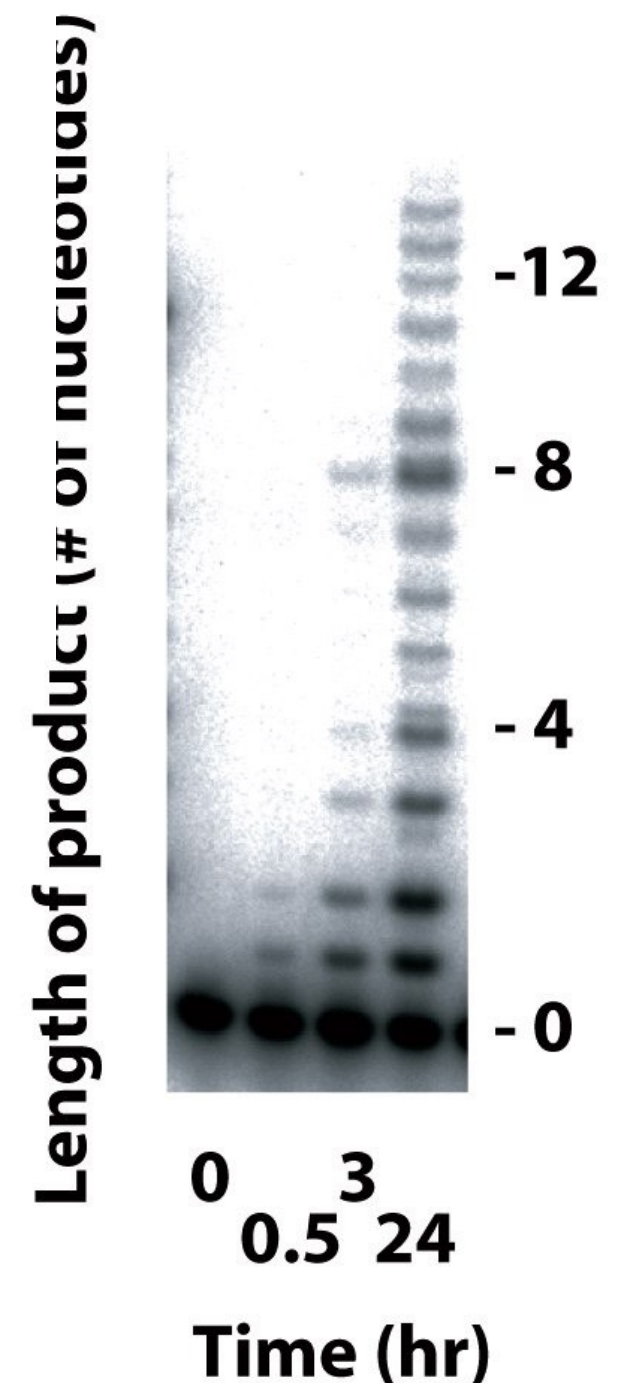
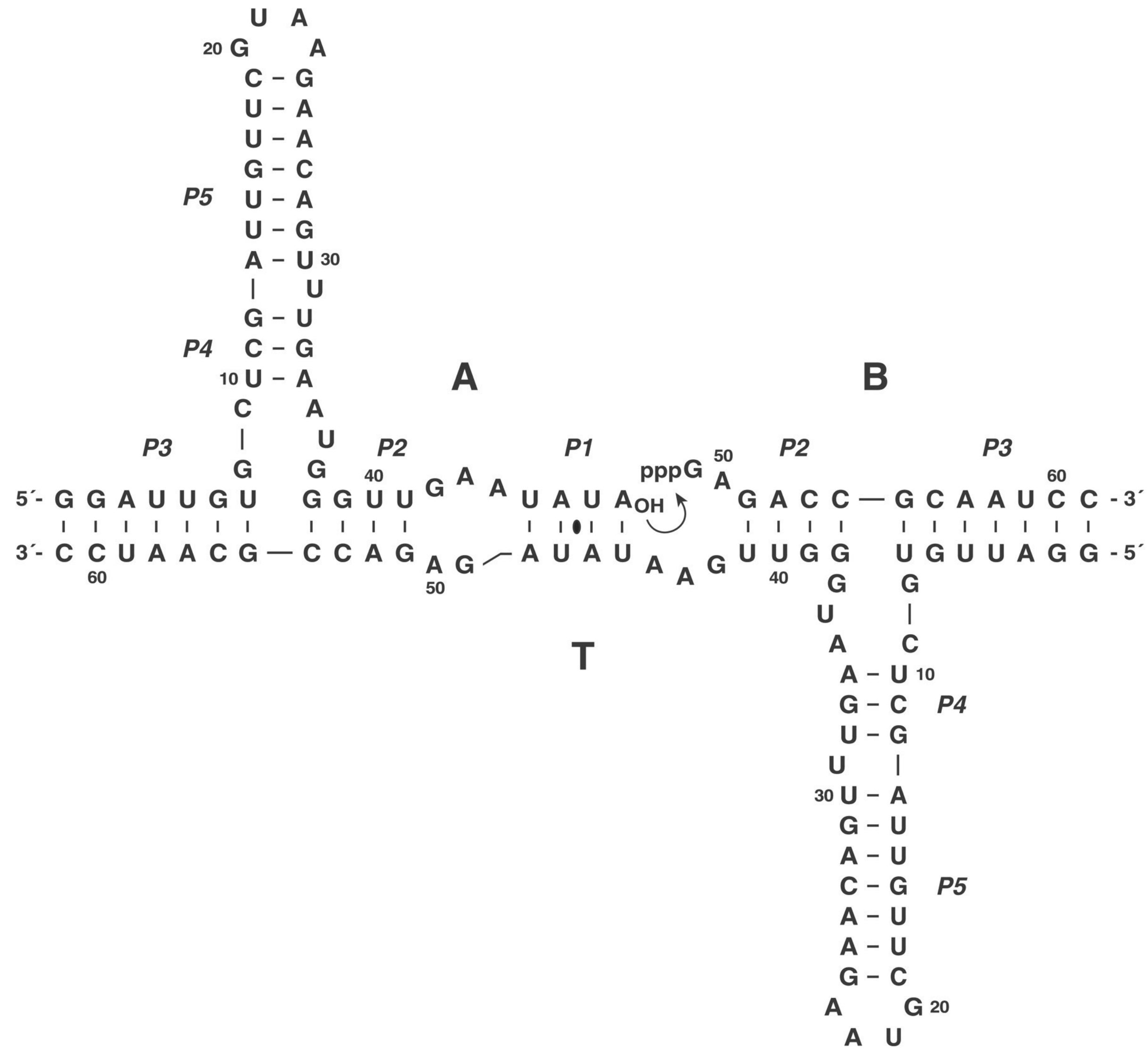


Figure 17-8c Evolutionary Analysis, 4/e © 2007 Pearson Prentice Hall, Inc.

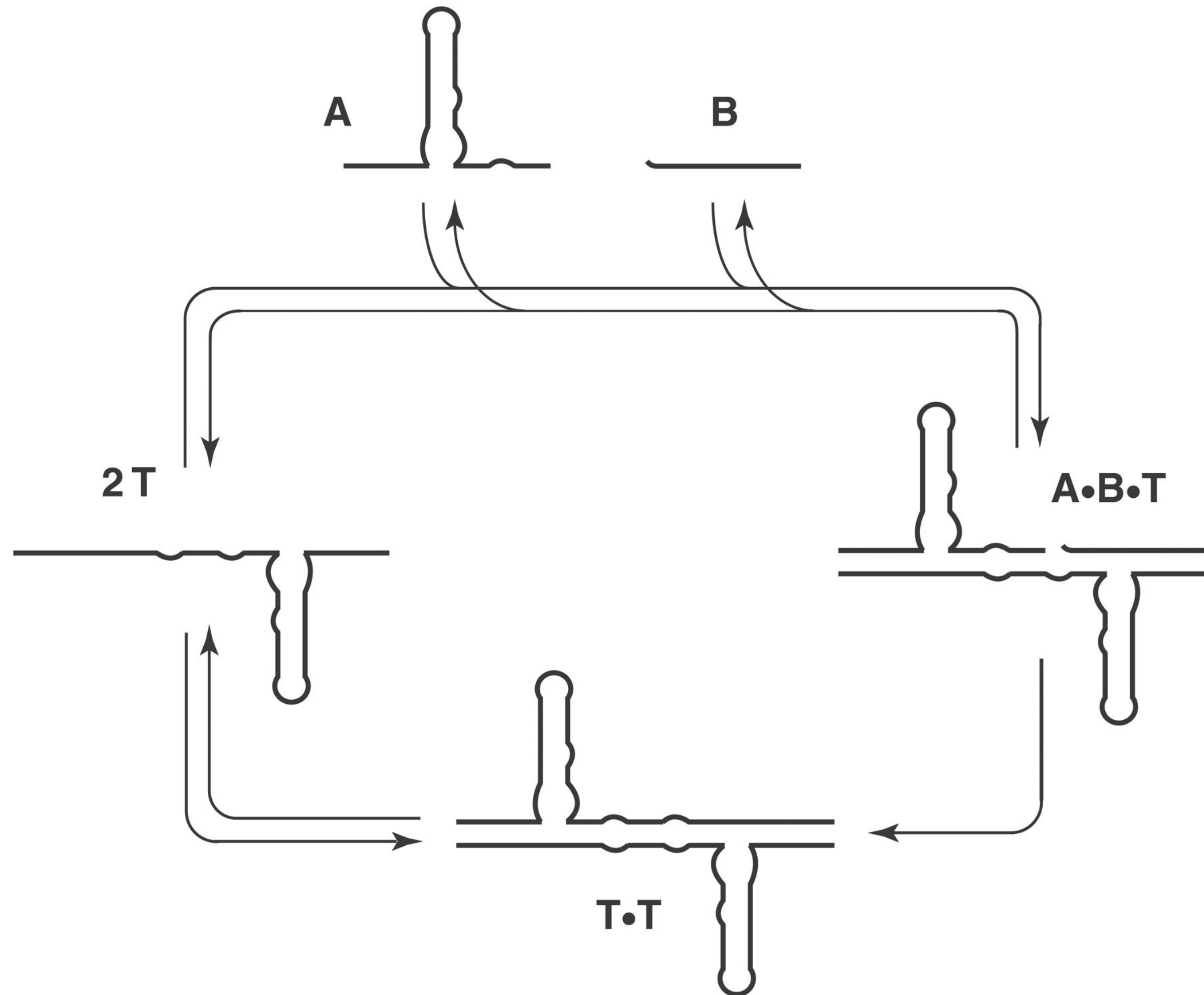
Autokatalityczna ligaza RNA



©2002 by National Academy of Sciences

Paul N , Joyce G F PNAS
2002;99:12733-12740

Autokatalityczna ligaza RNA

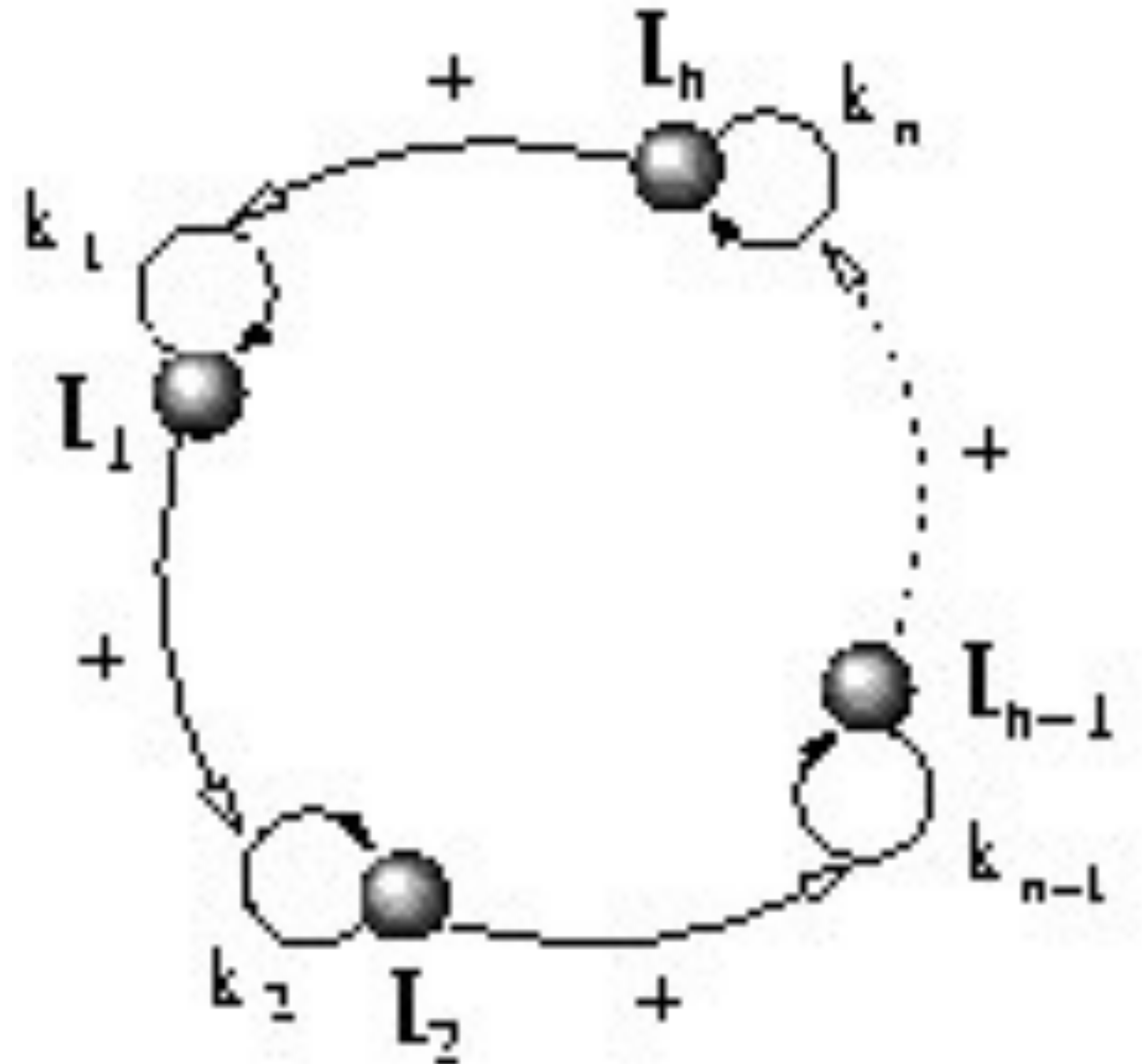


©2002 by National Academy of Sciences

Paul N , Joyce G F PNAS
2002;99:12733-12740

Problemy świata RNA

- Ograniczona zdolność magazynowania informacji w pojedynczym replikatorze (ilość informacji możliwej do zakodowania jest odwrotnie proporcjonalna do częstości błędów replikacji – **granica Eigena**)
- Rozwiązanie – sieci współdziałających replikatorów (**hipercykle**)
- „Samolubne RNA” w sieci replikatorów
- rozwiązanie – wydzielenie hipercyklu błoną i specjalizacja



Kooperacja i samoorganizacja

ARTICLE

doi:10.1038/nature11549

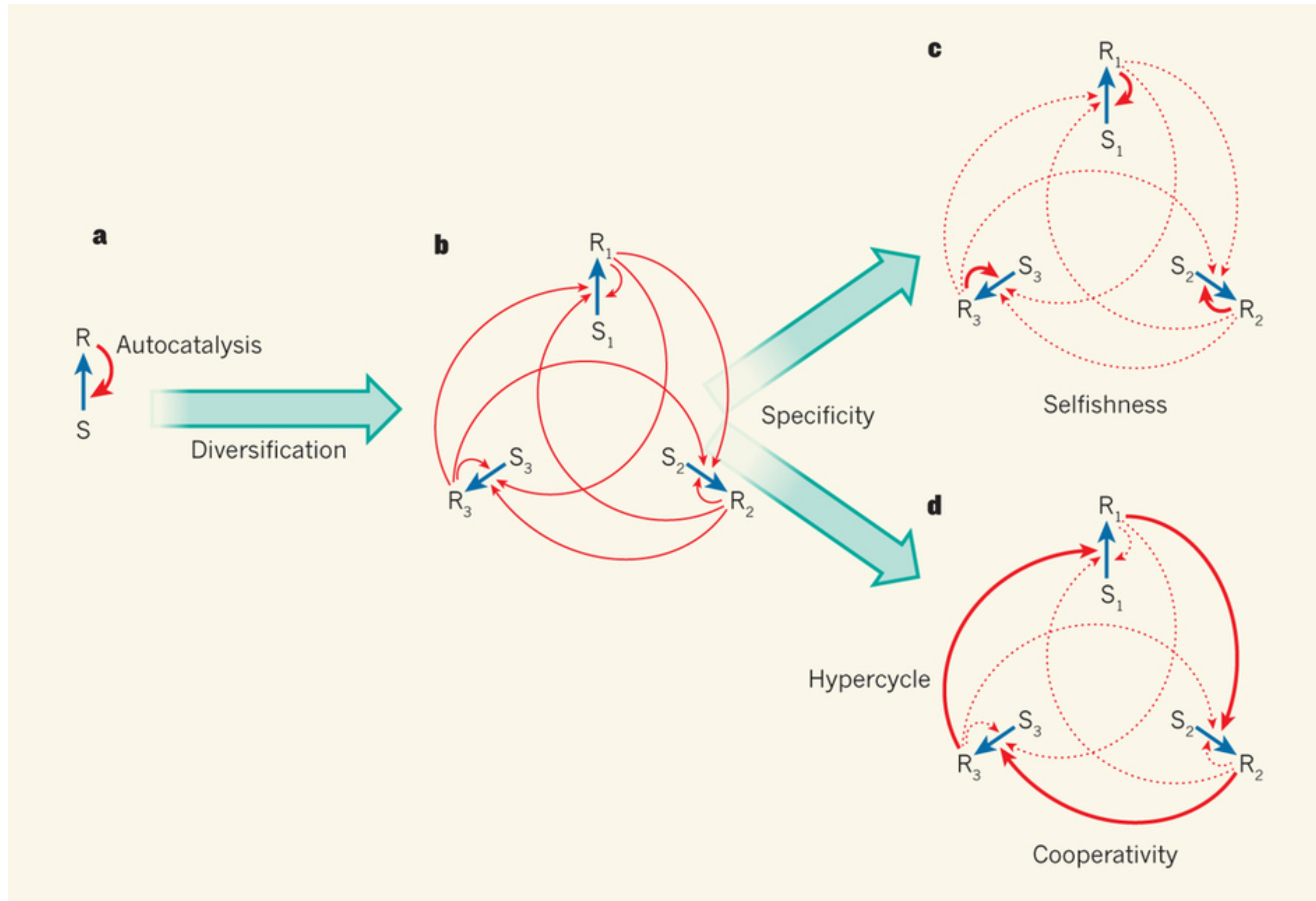
Spontaneous network formation among cooperative RNA replicators

Nilesh Vaidya¹, Michael L. Manapat², Irene A. Chen^{3†}, Ramon Xulvi-Brunet³, Eric J. Hayden⁴ & Niles Lehman¹

Nature, 2012 Nov 1;491(7422):72-7

Samoorganizujące się sieci rybozymów (2012) - potwierdzone doświadczalnie

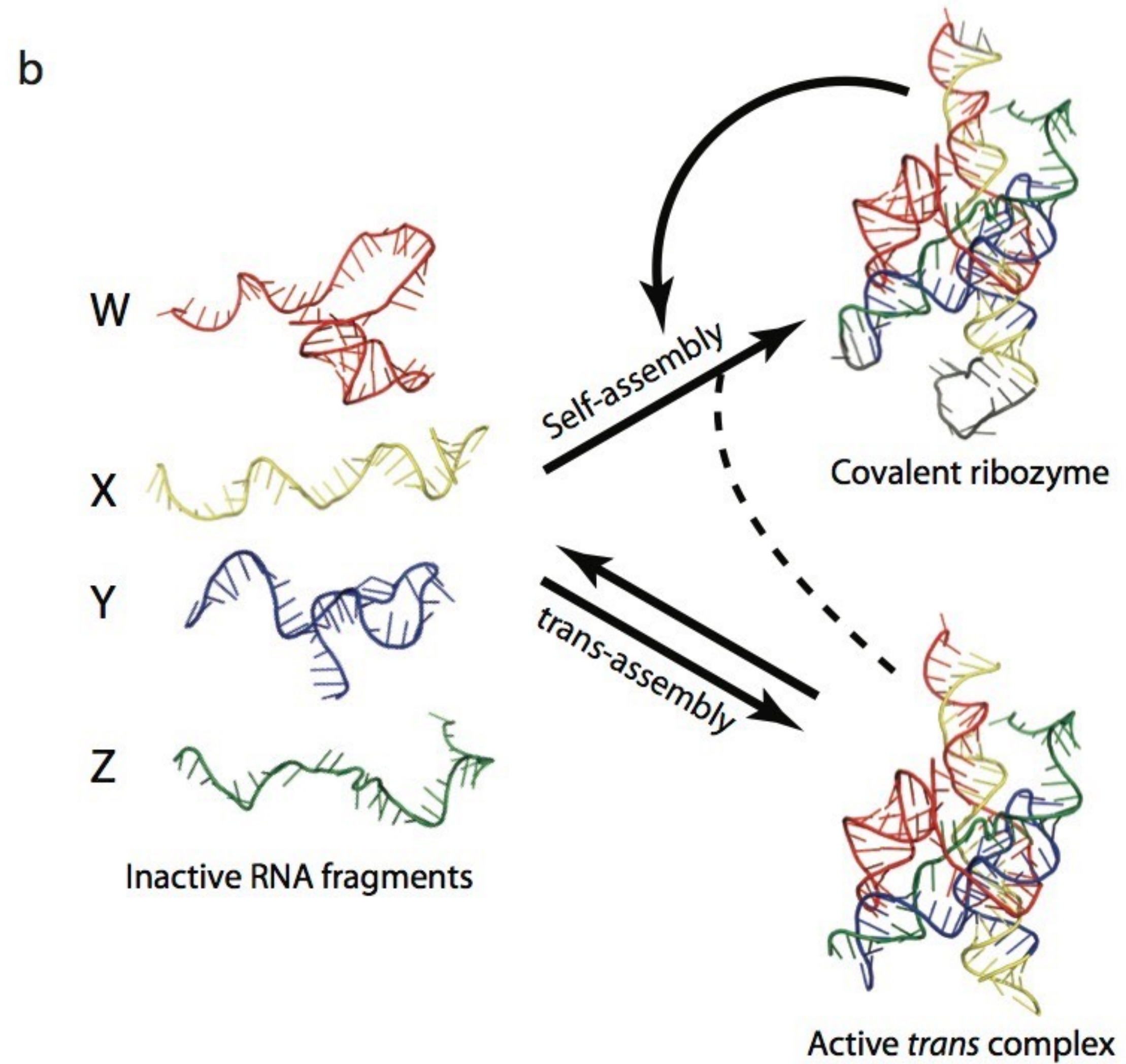
Kooperacja czy samolubność



Samorganizacja sieci RNA

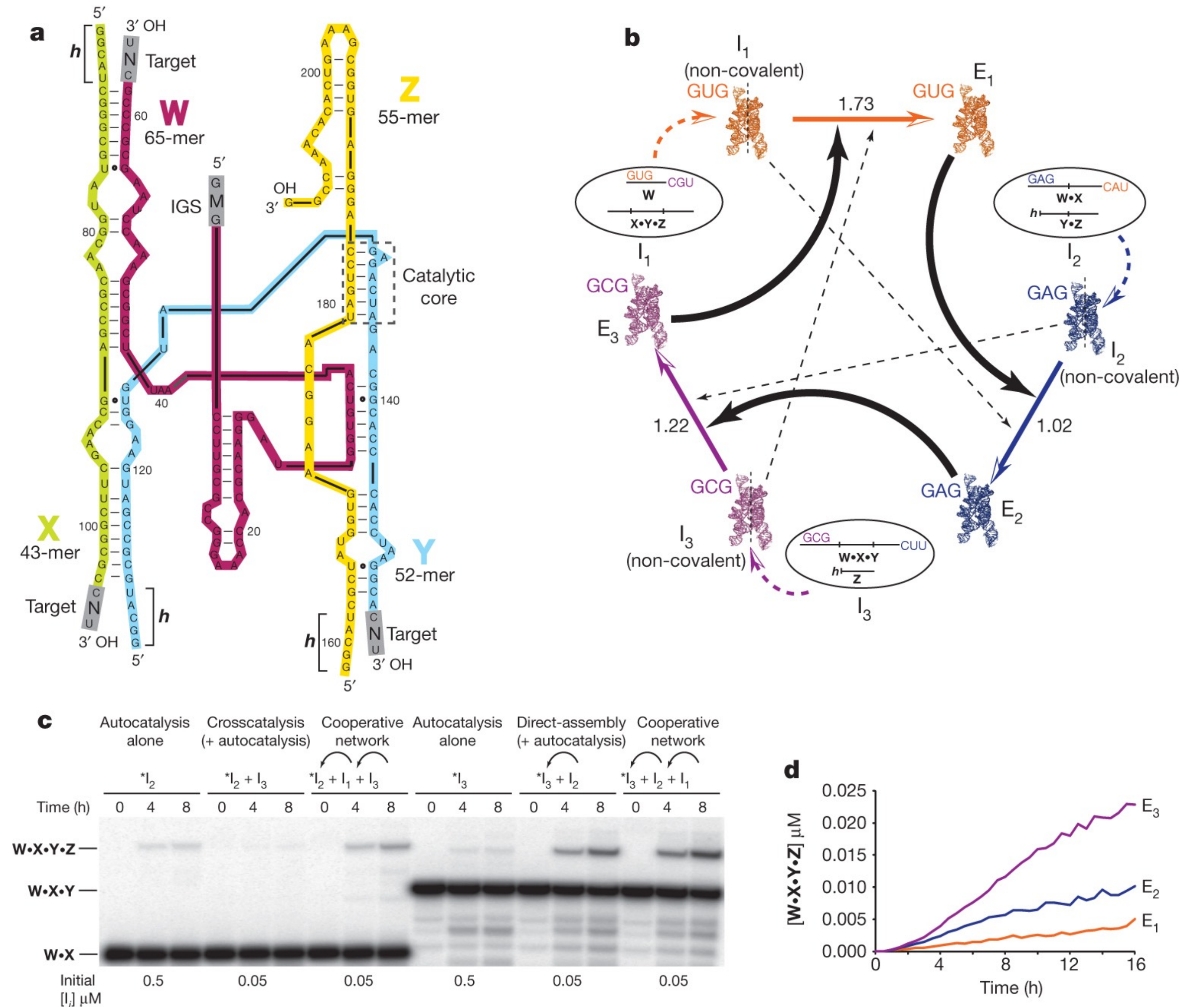
- Rybozym pochodzący z intronu *Azoarcus* (Proteobacteria)
- Pofragmentowany ma zdolność do autokatalitycznej ligacji fragmentów
- Stworzono warianty o różnej sekwencji i podobnych właściwościach
- Na ich podstawie stworzono warianty zdolne do ligowania innych wariantów

Rybozym Azoarcus

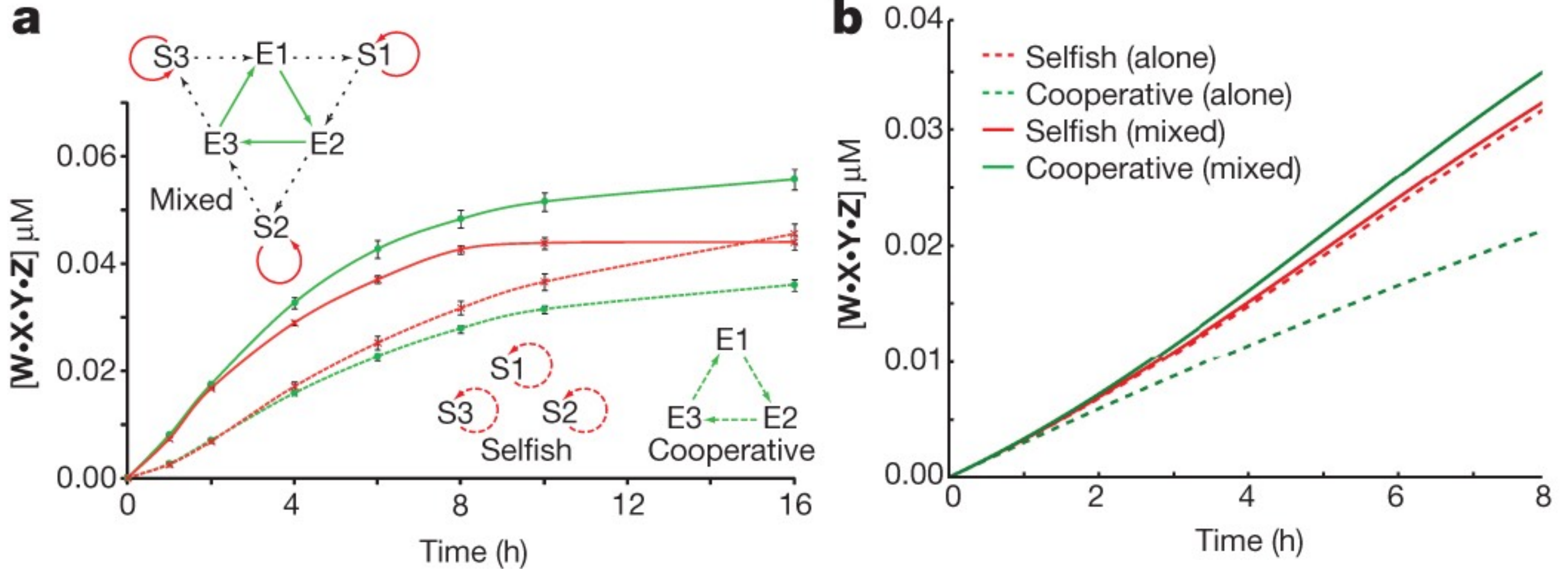


Samorganizacja sieci RNA

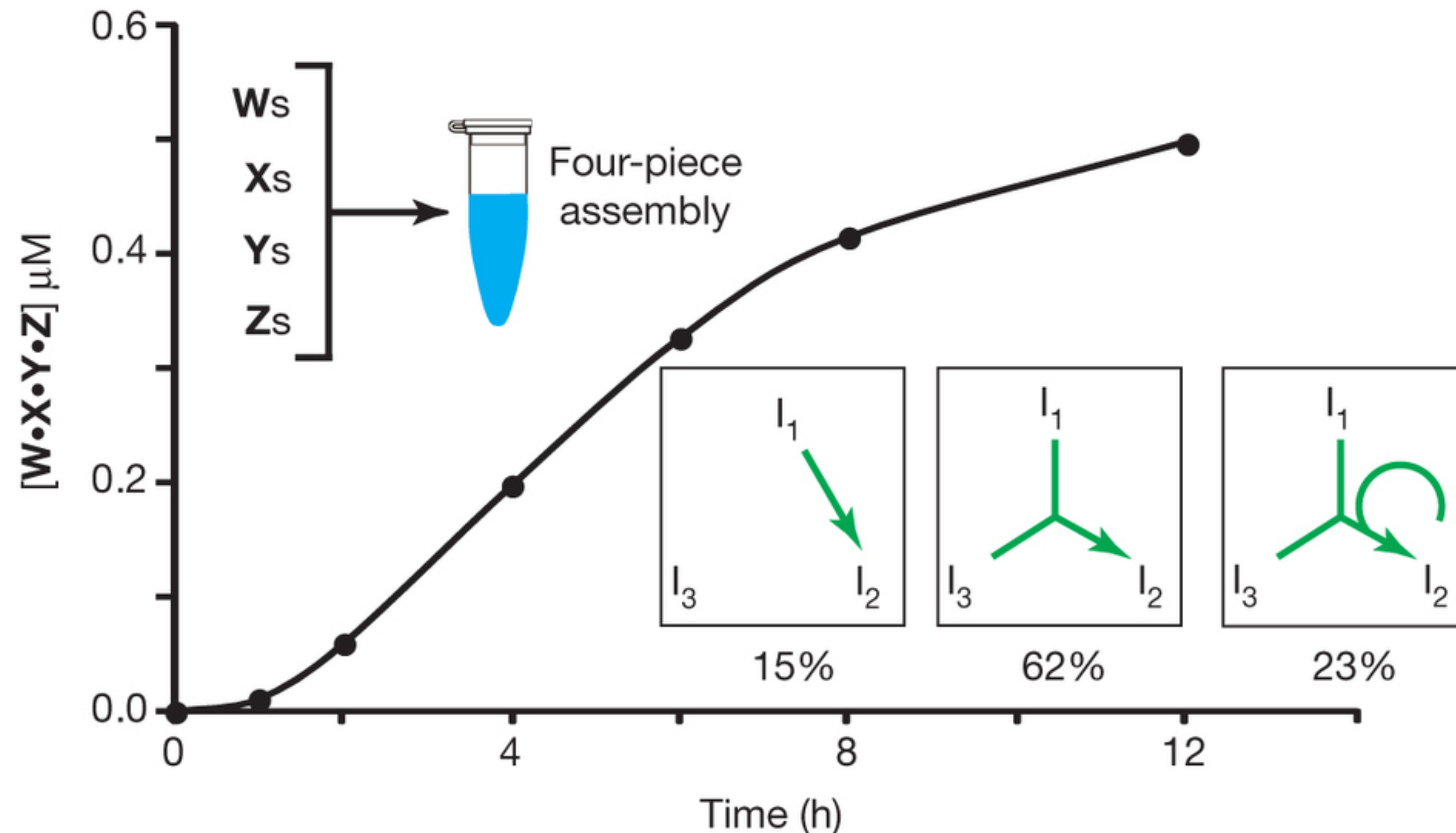
- Rybozomy zdolne do katalizy ligacji innych wariantów tworzą cykle autokatalizy
- Cykl jest wydajniejszy od pojedynczych “samolubnych” rybozymów



Kooperacja jest skuteczniejsza



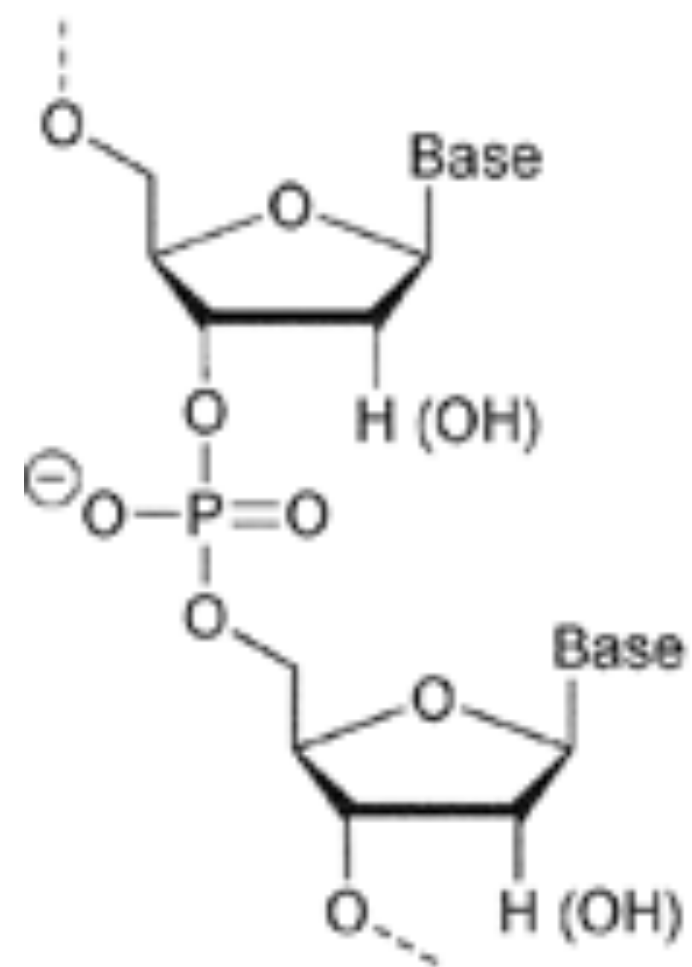
Zwiększenie liczby fragmentów



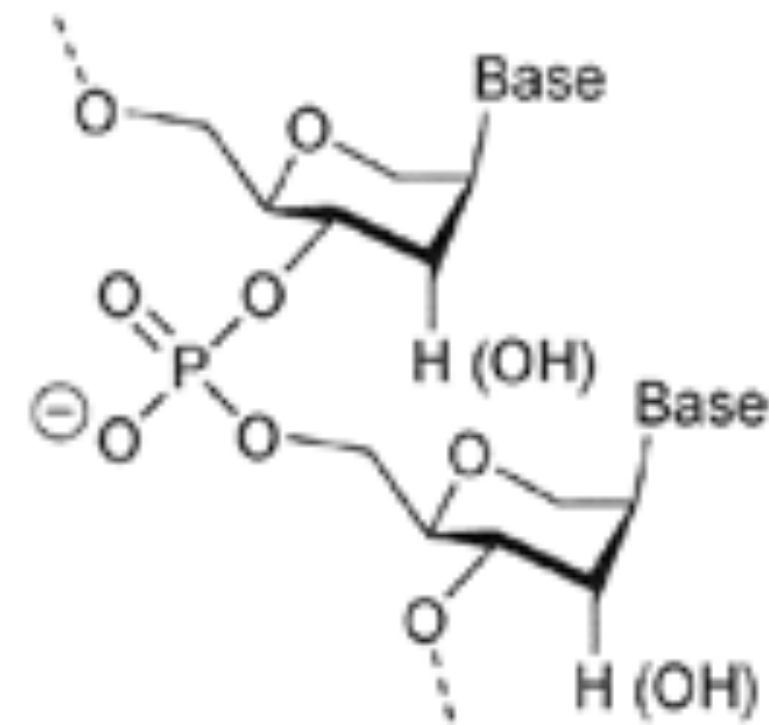
Tworzą się sieci, których większość wymaga kooperacji.
Złożoność w ewolucji *in vitro* przyrasta.

Alternatywy dla RNA

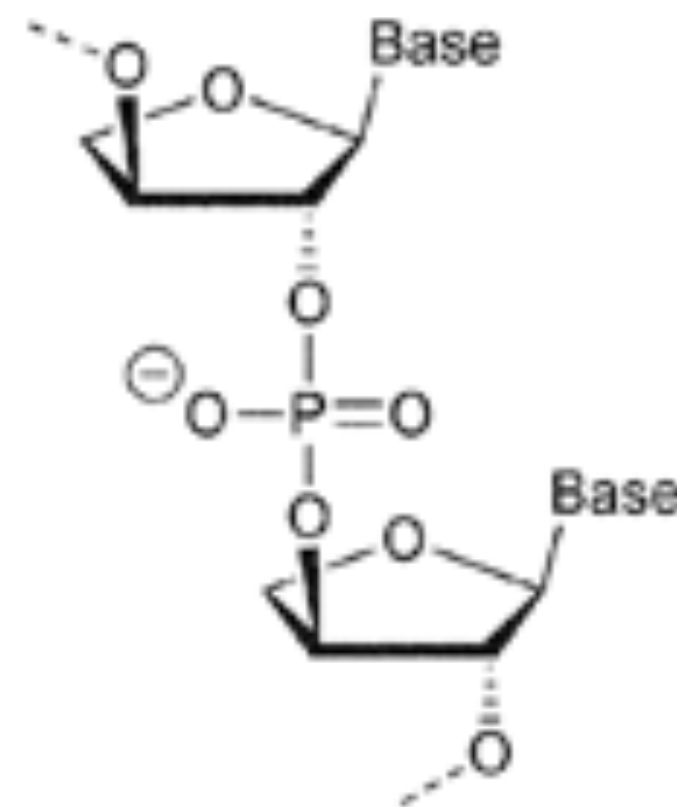
- RNA jest cząsteczką mało stabilną chemicznie
- Alternatywy – zastąpienie szkieletu rybozofosforanowego



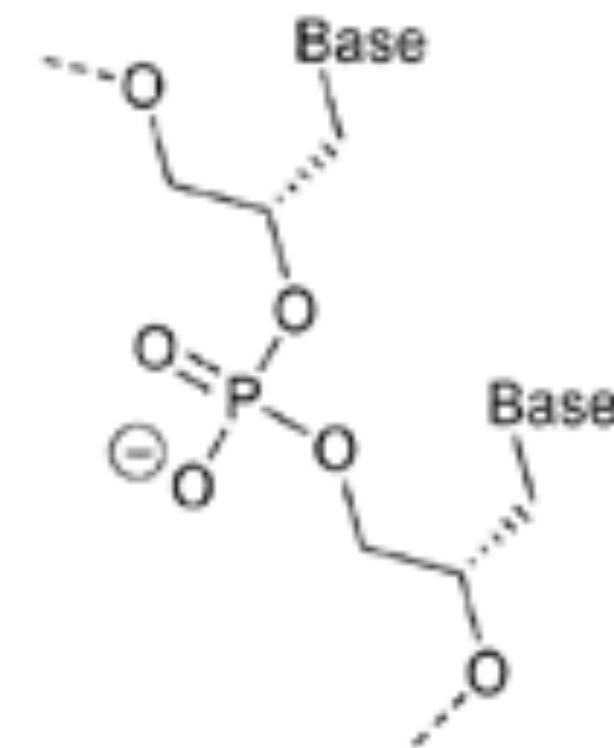
a) DNA (RNA)



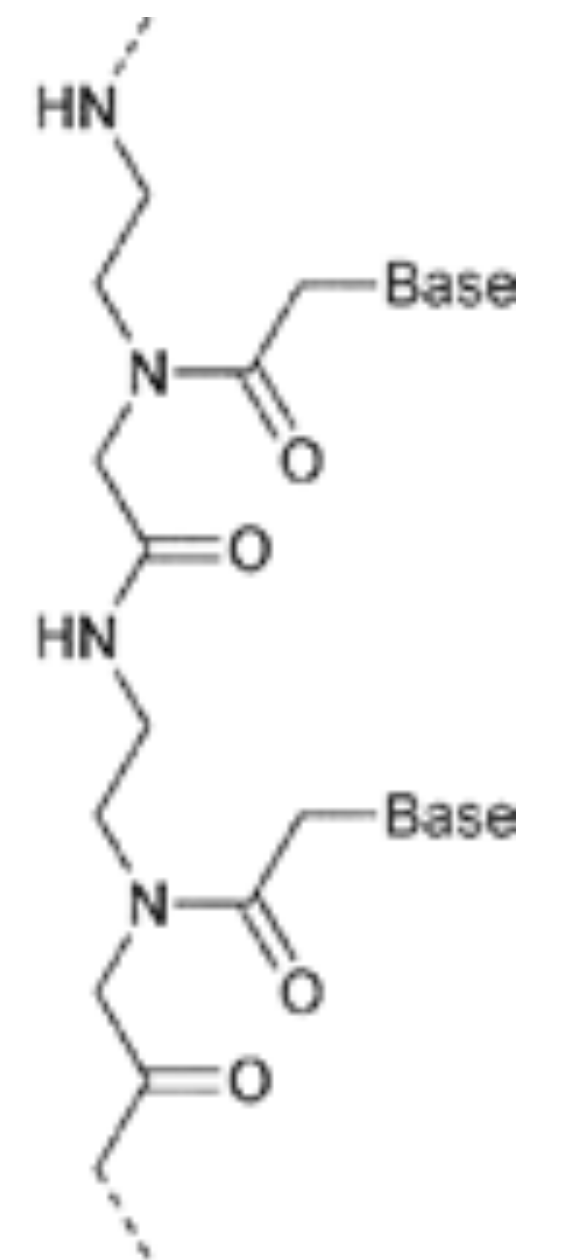
b) HNA (ANA)



c) TNA

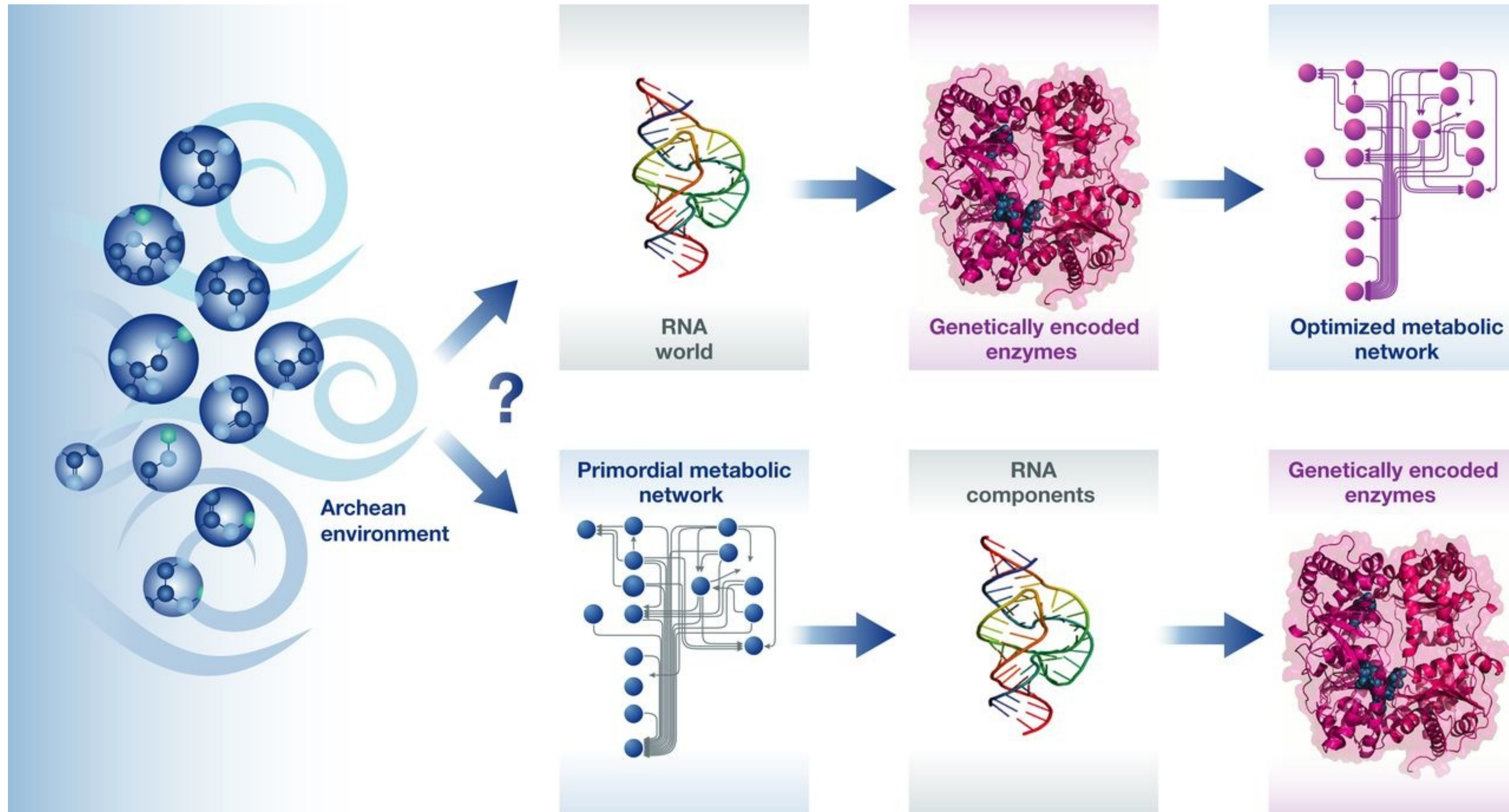


d) (S)-GNA



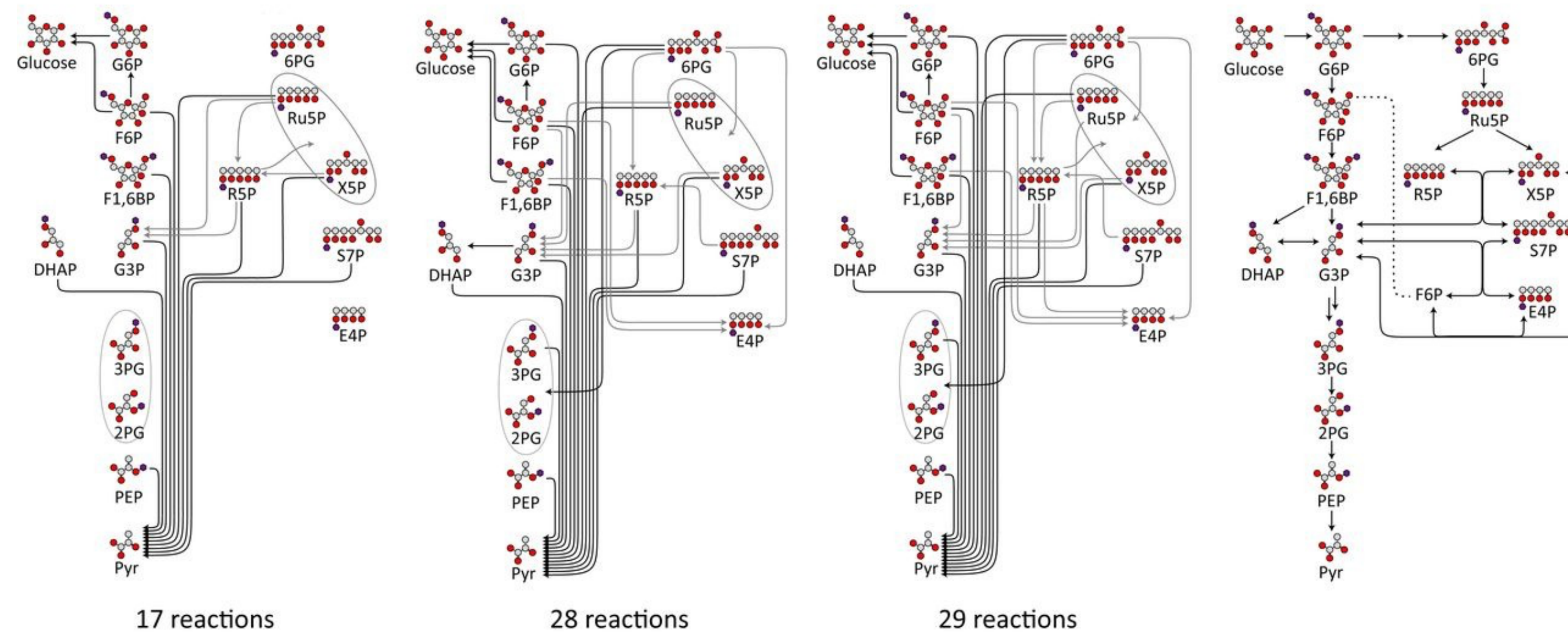
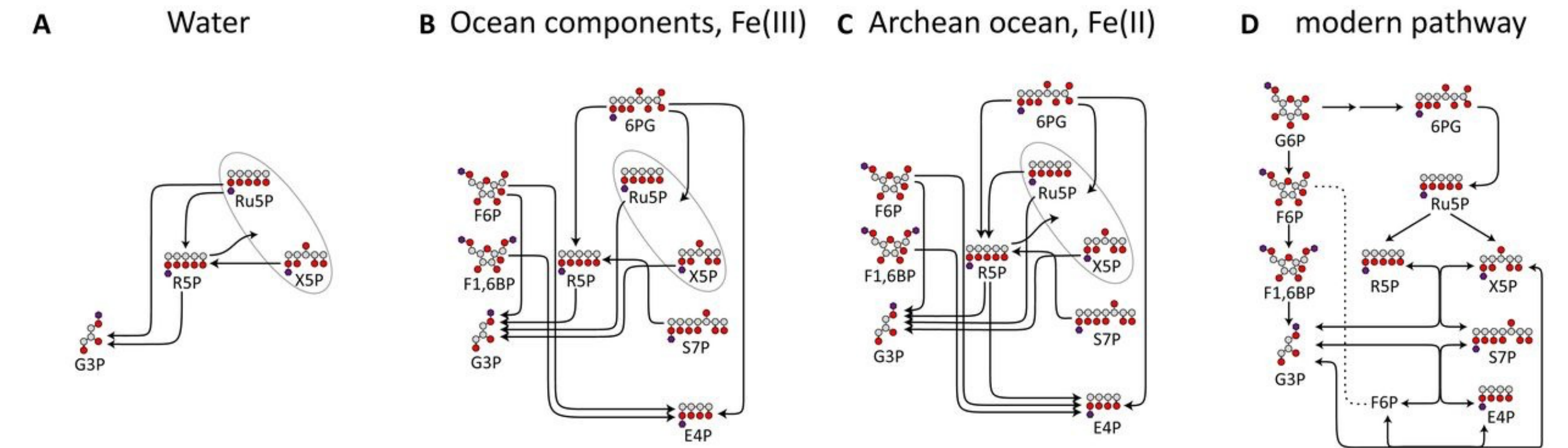
e) PNA

Prebiotyczna biochemia?



Prebiotyczna biochemia?

Cząsteczki niezbędne do syntezy RNA mogły powstać w pierwotnych sieciach metabolicznych



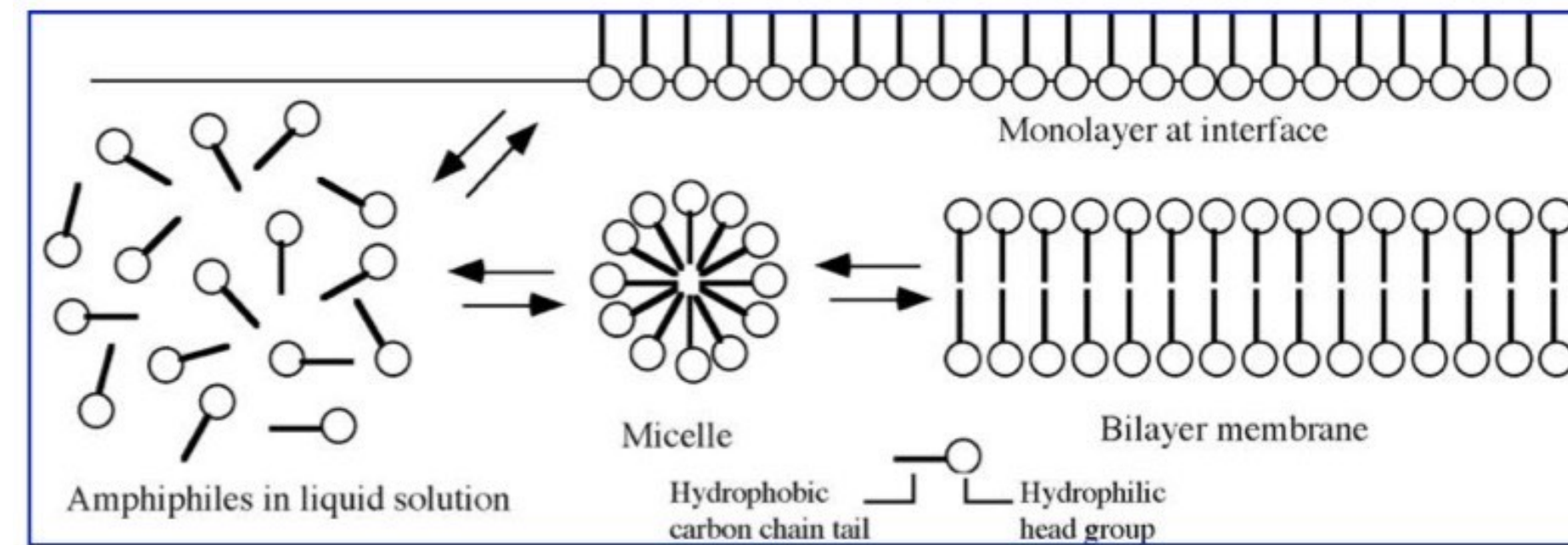
Non-enzymatic glycolysis and pentose phosphate pathway-like reactions in a plausible Archean ocean

Markus A Keller, Alexandra V Turchyn, Markus Ralser

DOI: 10.1002/msb.20145228 | Published 25.04.2014

Powstanie błon – pierwsze prakomórki

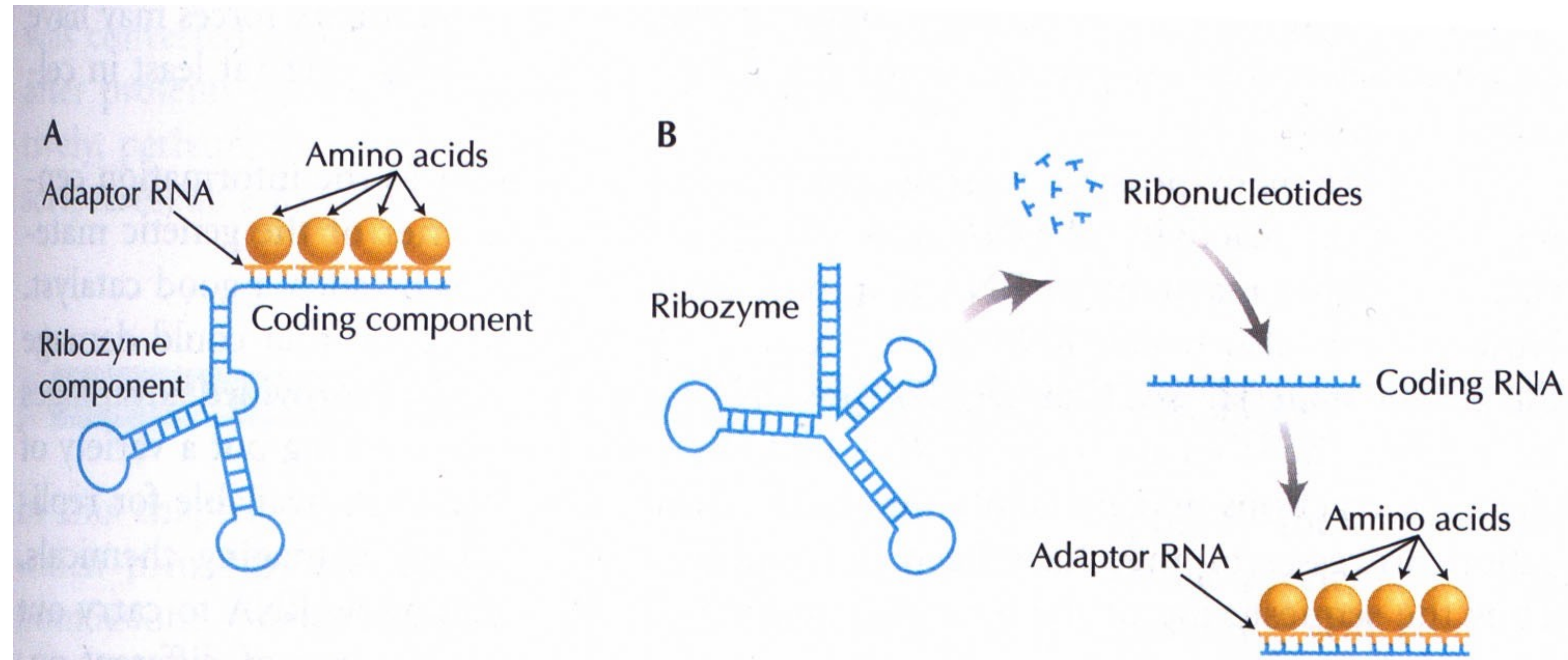
- Samoorganizacja lipidów amfipatycznych w struktury mogące otaczać prakomórki



- Takie lipidy mogły powstawać w warunkach prebiotycznych, a nawet w kosmosie
- Wyodrębnienie prakomórek błoną nastąpiło wcześnie w ewolucji

Ewolucja kodu

- Oddziaływania RNA – aminokwasy (prae-tRNA)
- rybozomy syntetyzujące
- uniwersalny rybosom pojawił się później



Aminoacylacja RNA przez prosty rybozym

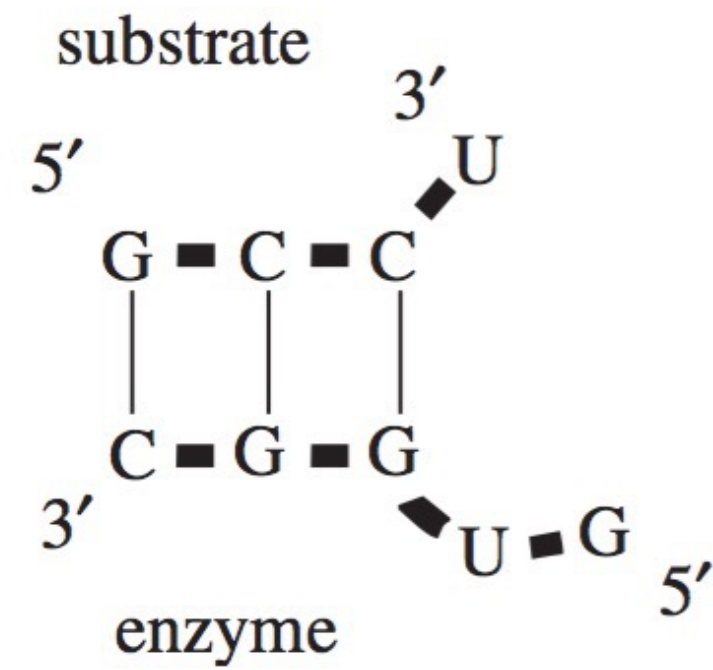
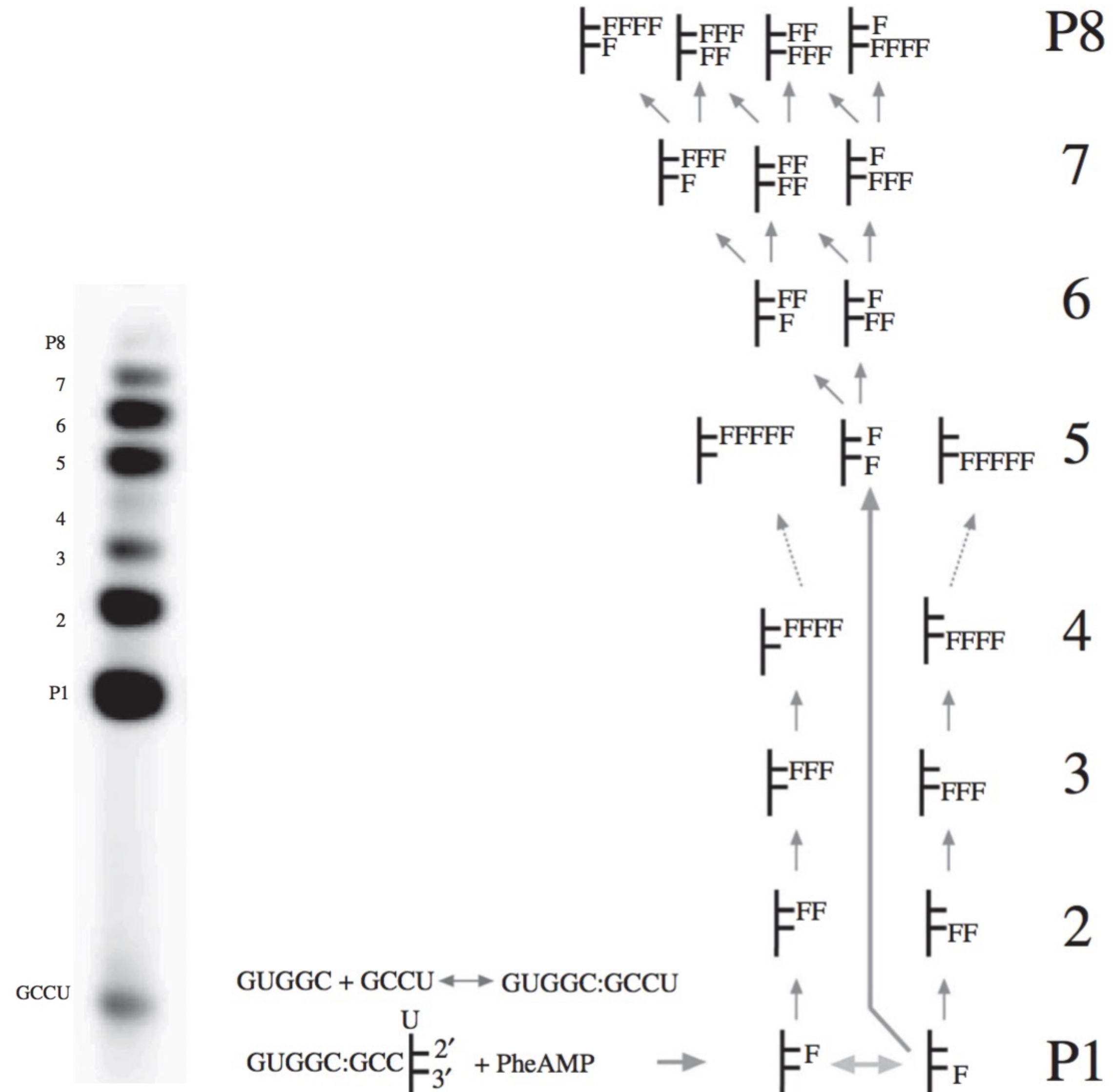


Figure 5. The minimal GCCU/GUGGC ribozyme system [10].



The meaning of a minuscule ribozyme

Michael Yarus

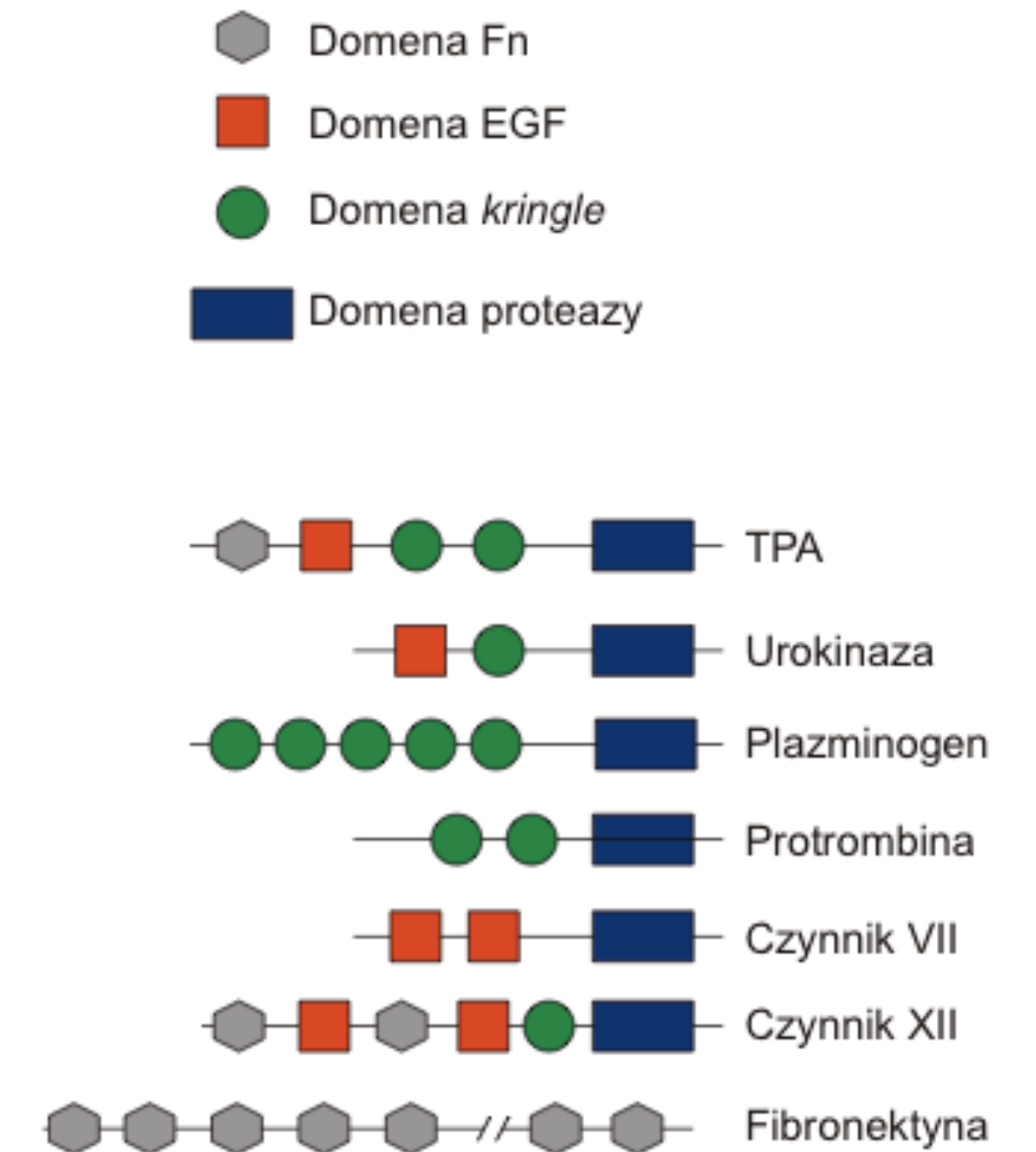
Phil. Trans. R. Soc. B 2011 **366**, doi: 10.1098/rstb.2011.0139

Ewolucja kodu

- Pierwszy kod był mniej specyficzny
 - rozróżniane mniej aminokwasów
 - rozróżniane grupy aminokwasów
- Ewolucja przez
 - zwiększanie liczby kodowanych aminokwasów
 - zwiększanie specyficzności
- Czy kod od początku był trójkowy?
 - może wyewoluował z dwójkowego, ale z przecinkiem (trzeci nukleotyd nieznaczący)

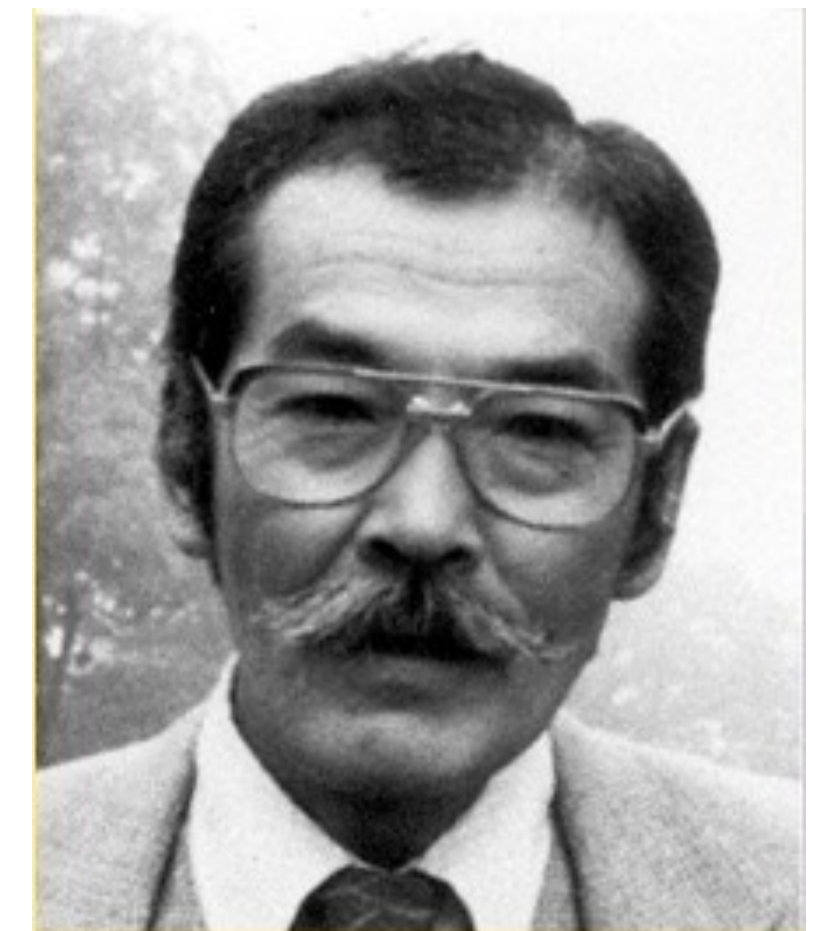
Ewolucja białek

- Pierwsze peptydy były bardzo krótkie
- Do dzisiaj w białkach ślady budowy z powtarzających się krótkich motywów

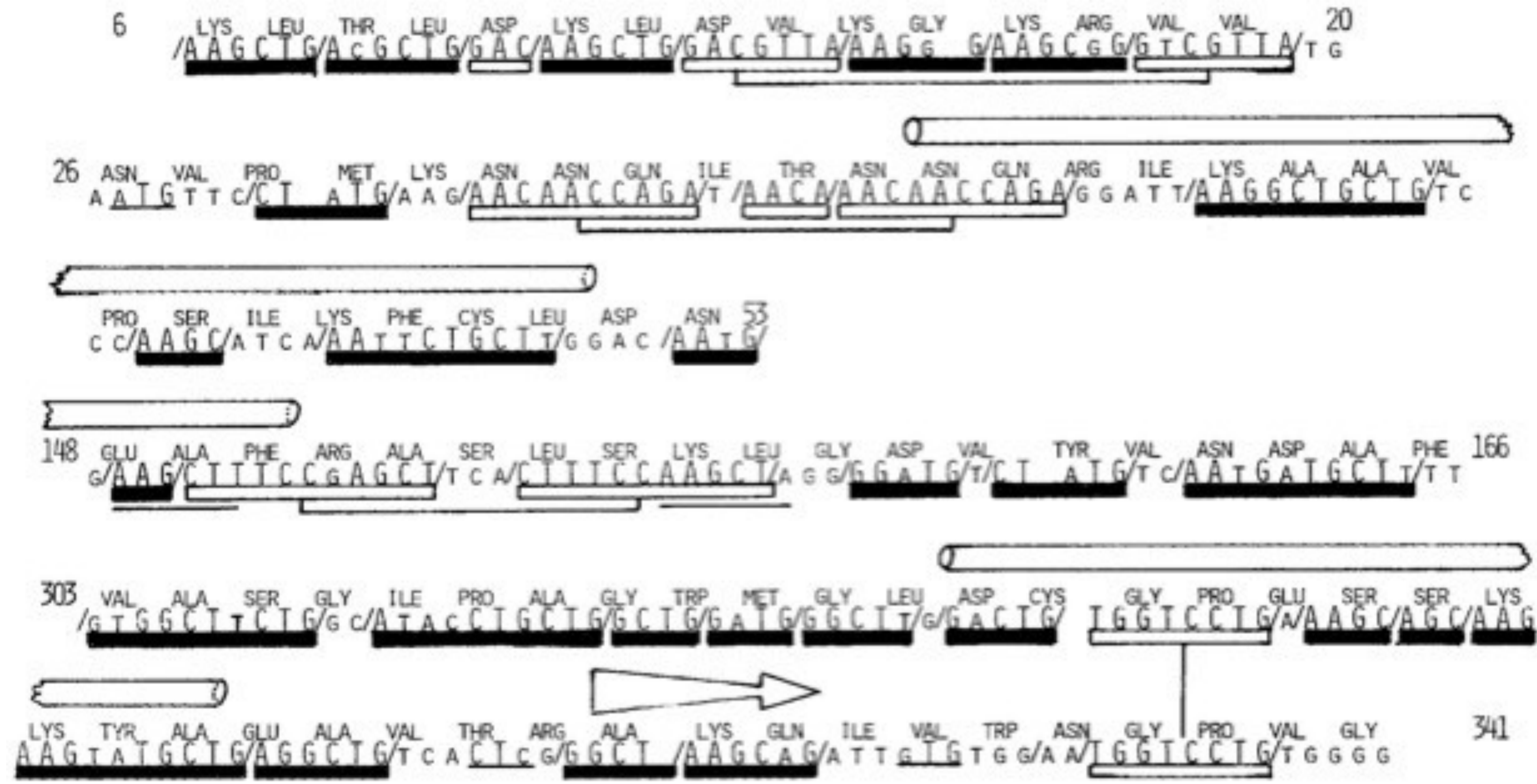


Geny i muzyka – Susumu Ohno

- Ohno S, Ohno M. 1986. The all pervasive principle of repetitious recurrence governs not only coding sequence construction but also human endeavor in musical composition. *Immunogenetics* 24:71-78
- Ohno S. 1987. Repetition as the essence of life on this earth: music and genes. *Haematol. Blood Transfus.* 31:511-518
- Ohno S. 1993. A song in praise of peptide palindromes. *Leukemia* 2S:157-159



Powtórzenia



Muzyka sekwencji

C / A A G G T T C C / A

HUMAN X-LINKED PHOSPHOGLYCERATE KINASE

1 MET SER LEU SER ASN LYS LEU THR LEU
 A T G T C G C T T T C T A A C A A G C T G A C G C T G

10 ASP LYS LEU ASP VAL LYS GLY LYS ARG VAL VAL
 G A C A A G C T G G A C G T T A A A G G G A A G C G G G T C G T T A

20 MET ARG VAL ASP PHE ASN VAL PRO MET LYS ASN ASN
 T G A G A G T C G A C T T C A A T G T T C C T A T G A A G A A C A A C

30 GLN ILE THR ASN ASN GLN ARG ILE LYS ALA ALA VAL PRO
 C A G A T A A C A A A C A A C C A G A A G G A T T A A G G C T G C T G T C C C

40 SER ILE LYS PHE CYS LEU ASP
 A A G C A T C A A A T T C T G C T T G G A

50

Fig. 3. Musical transformation of the first 52 codons of the human X-linked phosphoglycerate kinase coding sequence (Singer-Sam et al. 1983) in D minor and with a time signature of $9/8$. As the piece was written for the violin, only the treble clef musical score is given

Muzyka sekwencji

The image displays a musical score for the treble clef of Chopin's Nocturne, Op. 55, No. 1. The score is transcribed in a 2/4 time signature with a key signature of one flat (B-flat). The music is written on four staves. Above each staff, amino acid sequences are aligned with the notes. The sequences are: Staff 1: GLY LEU ALA PRO THR SER PRO GLU ILE THR HIS SER; Staff 2: GLN PRO PHE SER SER THR VAL PHE LYS ILE PRO; Staff 3: ARG CYS GLN THR PRO PHE VAL GLY VAL GLY PRO; Staff 4: ASN LEU THR GLN LYS SER PRO ILE HIS PRO THR PHE PHE HIS. The amino acid sequences are written in three-letter codes. The musical notation includes various note values, rests, and dynamic markings such as *pp* and *z*.

Fig. 4. Transcription of the initial portion of the treble clef musical score of the Nocturne, opus 55, no. 1 by Frederic Chopin to the base sequence according to the dictate specified in Figure 2. The amino acid sequence encodable by the longest open reading frame is also shown. The principal subject transcribable

Powstanie DNA

- Reduktaza rybonukleotydów
- Enzym obecny we wszystkich gałęziach drzewa życia – bardzo stary

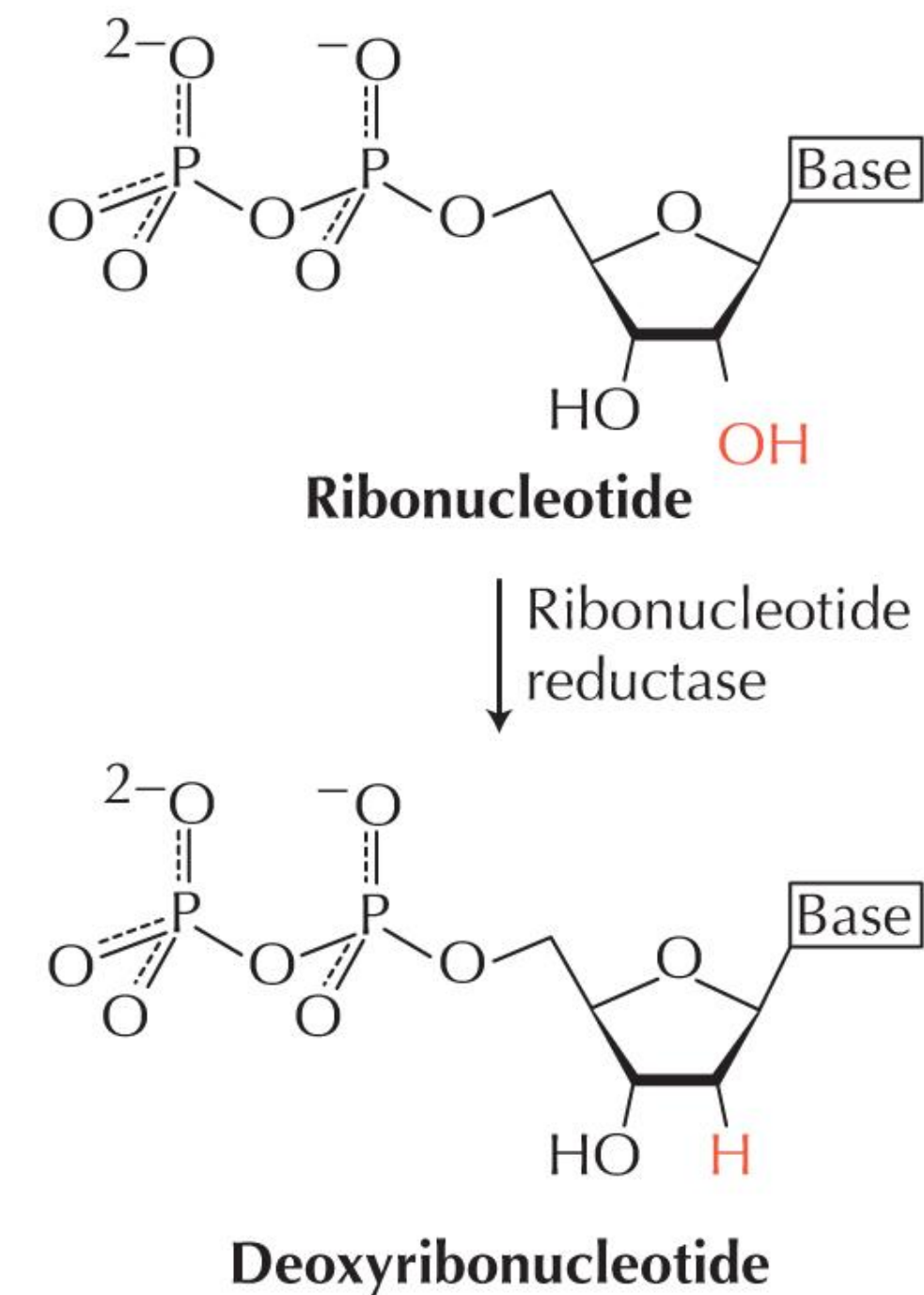
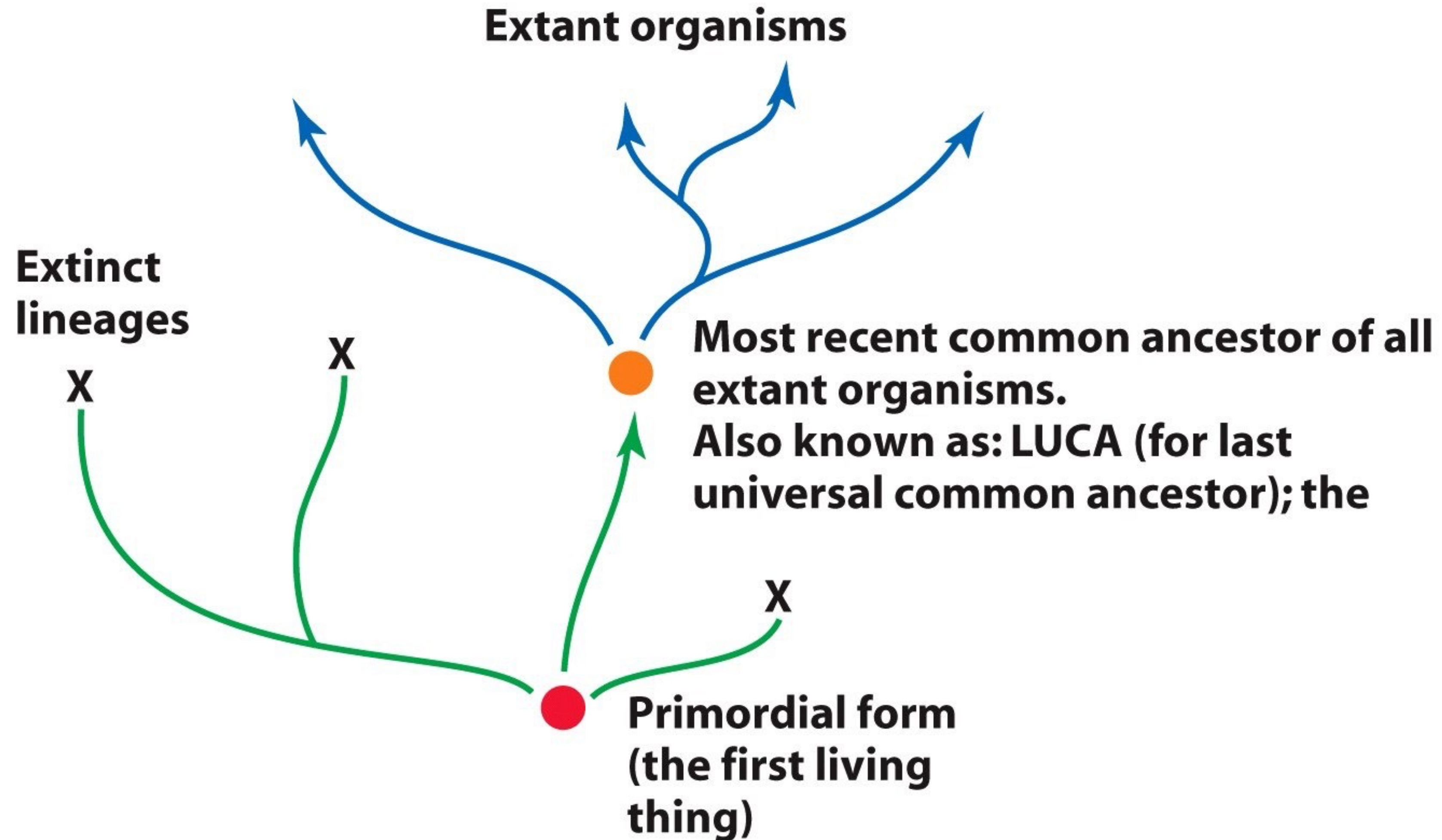


FIGURE 4.15. Ribonucleotide reductase. In modern organisms, deoxyribonucleotides are synthesized from ribonucleotides using the enzyme ribonucleotide reductase. This is consistent with deoxyribonucleotide synthesis being a more recent invention than ribonucleotide synthesis.

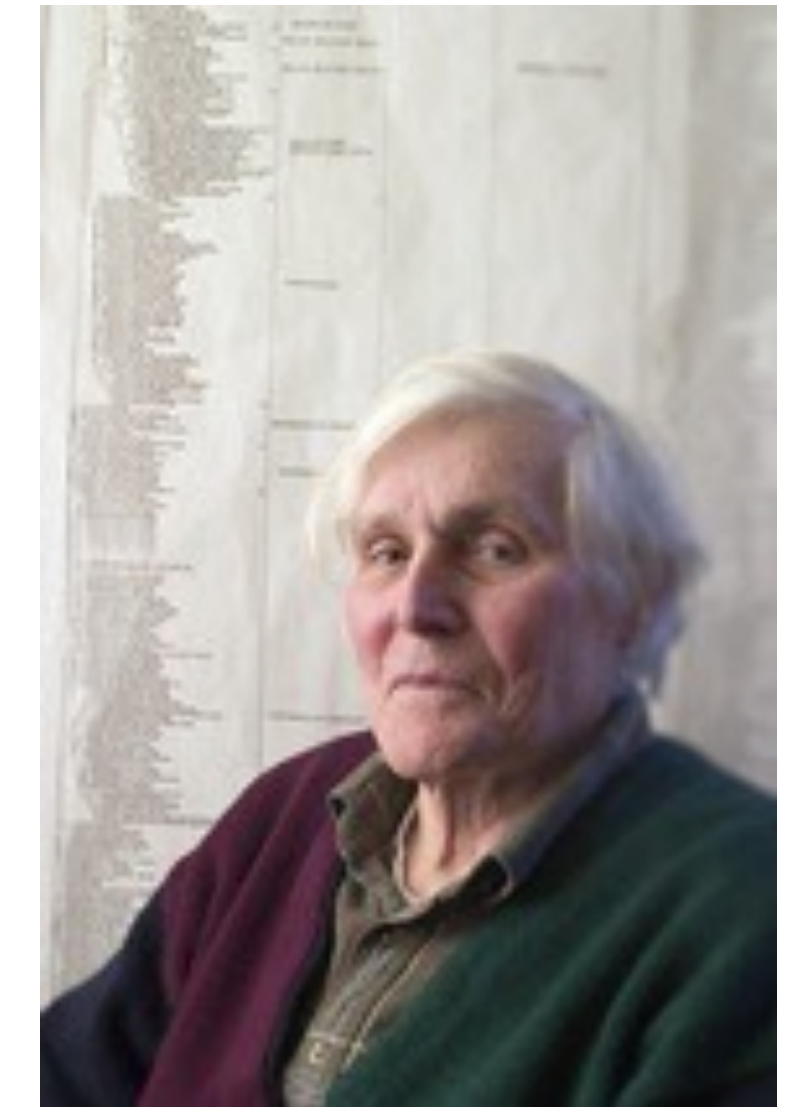
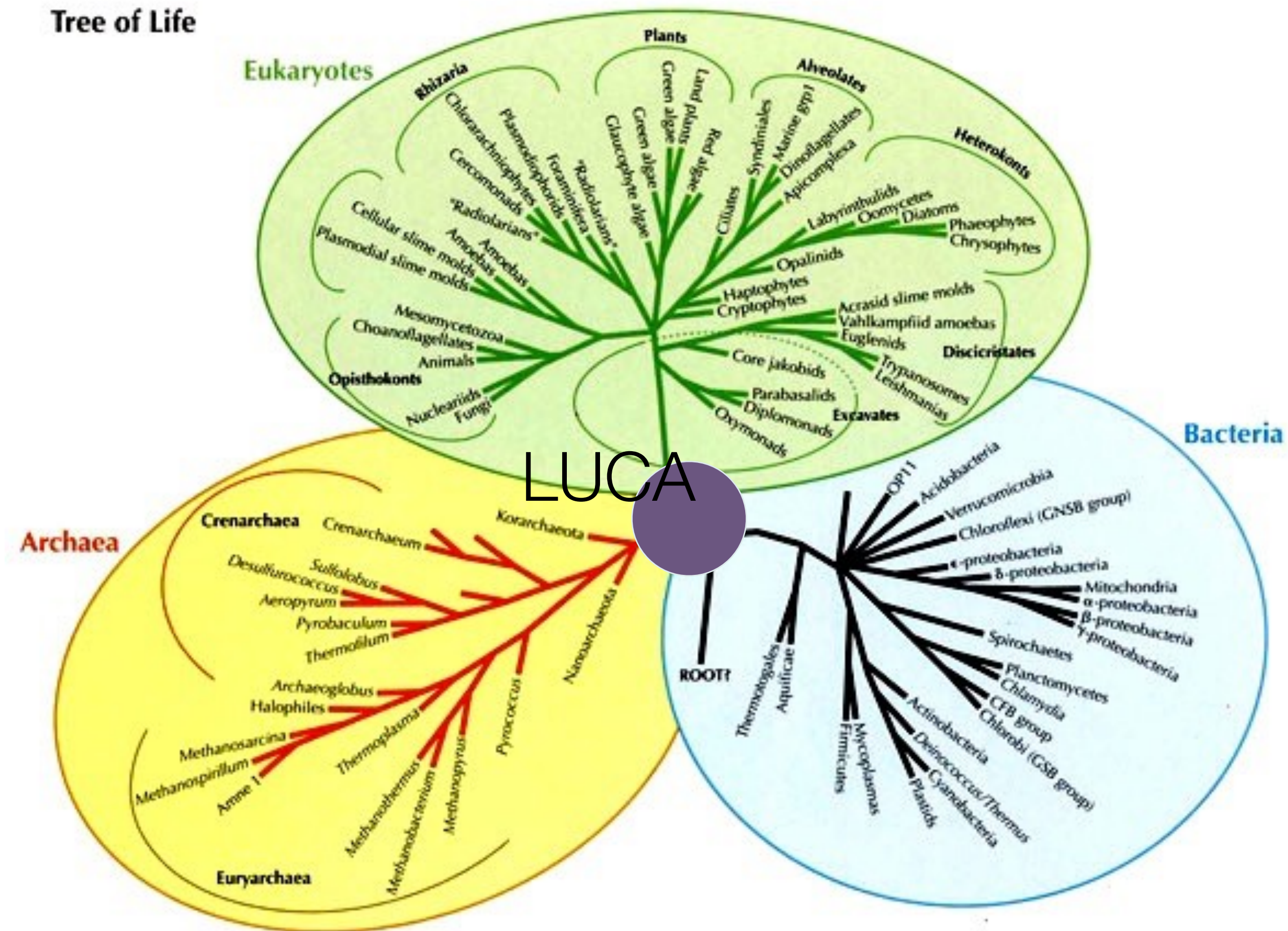
LUCA



Jaki był LUCA?

- Na podstawie zestawu genów obecnych we wszystkich gałęziach drzewa życia
 - problem – horyzontalny transfer genów (przeszacowanie zestawu genów pierwotnych)
- Organizm o budowie komórkowej zbliżony do współczesnych prokariotów
- Poprzedzał rozejście się linii Bacteria, Archaea i Eukarya

Prokaryota – nie są jedną grupą

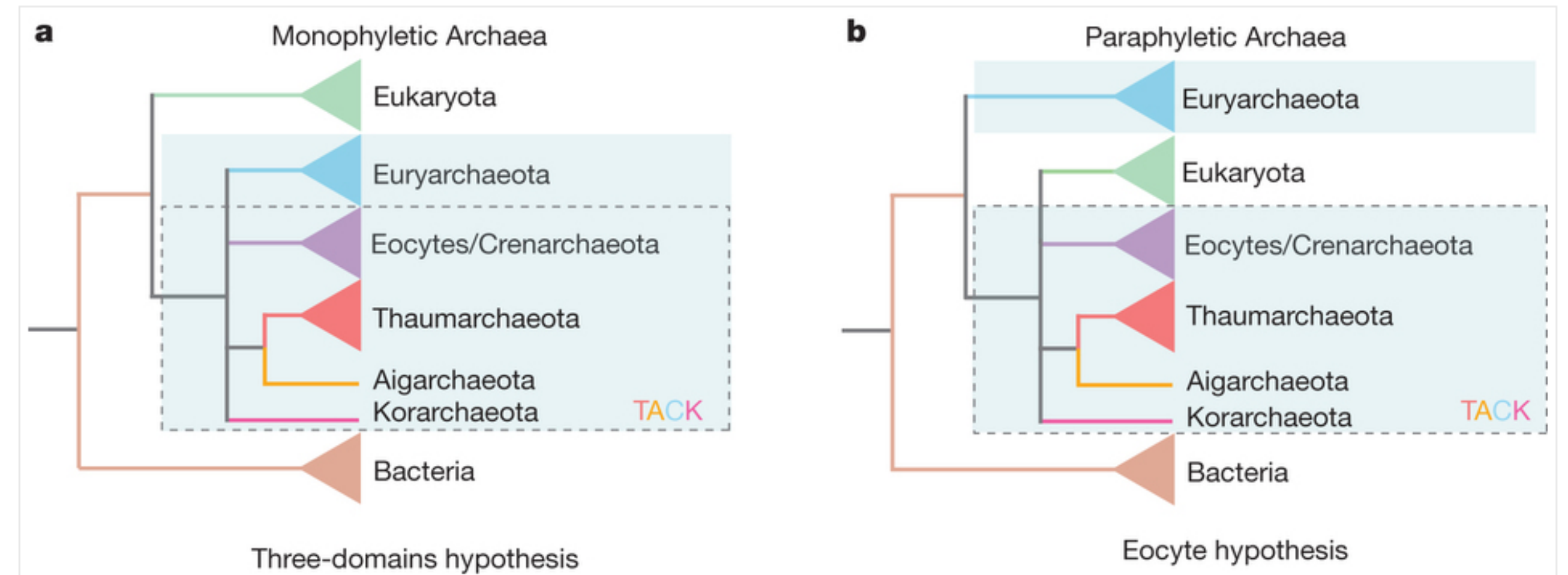


Carl Woese (1928 - 2012)

LUCA – Last Universal Common Ancestor

Hipoteza eocytów

- Być może Eukaryota (jądro) to jedna z gałęzi Archaea
- Dwie, a nie trzy domeny drzewa życia



An archaeal origin of eukaryotes supports only two primary domains of life

Tom A. Williams, Peter G. Foster, Cymon J. Cox & T. Martin Embley

[Affiliations](#) | [Contributions](#) | [Corresponding author](#)

Nature **504**, 231–236 (12 December 2013) | doi:10.1038/nature12779

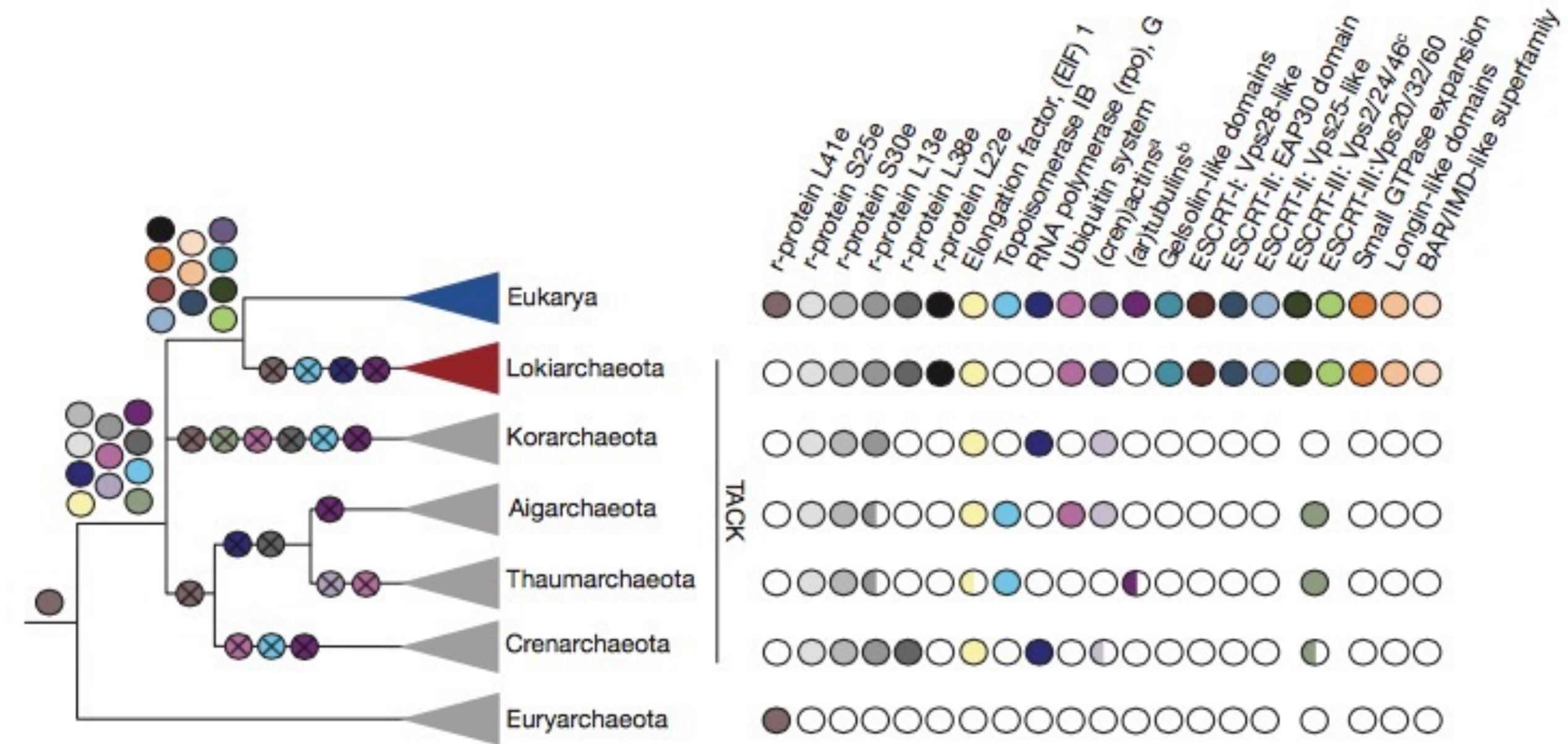
Received 17 June 2013 | Accepted 14 October 2013 | Published online 11 December 2013

[Citation](#) [Reprints](#) [Rights & permissions](#) [Article metrics](#)

The discovery of the Archaea and the proposal of the three-domains 'universal' tree, based on ribosomal RNA and core genes mainly involved in protein translation, catalysed new ideas for cellular evolution and eukaryotic origins. However, accumulating evidence suggests that the three-domains tree may be incorrect: evolutionary trees made using newer methods place eukaryotic core genes within the Archaea, supporting hypotheses in which an archaeon participated in eukaryotic origins by founding the host lineage for the mitochondrial endosymbiont. These results provide support for only two primary domains of life—Archaea and Bacteria—because eukaryotes arose through partnership between them.

Gospodarz endosymbiozy był archeonem

- Lokiarcheota - monofiletyczna grupa Archaea, odkryta w badaniach metagenomicznych (2105 r.)
- 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 Martijn¹, Anders E. Lind¹, Roel van Eijk[†], Christa Schleper^{2,3}, Lionel Guy^{1,4} & Thijs J. G. Ettema¹

©2015 Macmillan Publishers Limited. All rights reserved

Co potrafił LUCA

- transkrypcja, translacja (kod odpowiadający współczesnemu)
- metabolizm energetyczny oparty na ATP
- synteza długich łańcuchów DNA
- białka błonowe
- ok. 600 genów

- Ale nie wszyscy się zgadzają:
 - prosty, niewiele genów, bez współczesnego systemu replikacji genomu (Koonin 2003)

Genom LUCA: DNA czy RNA

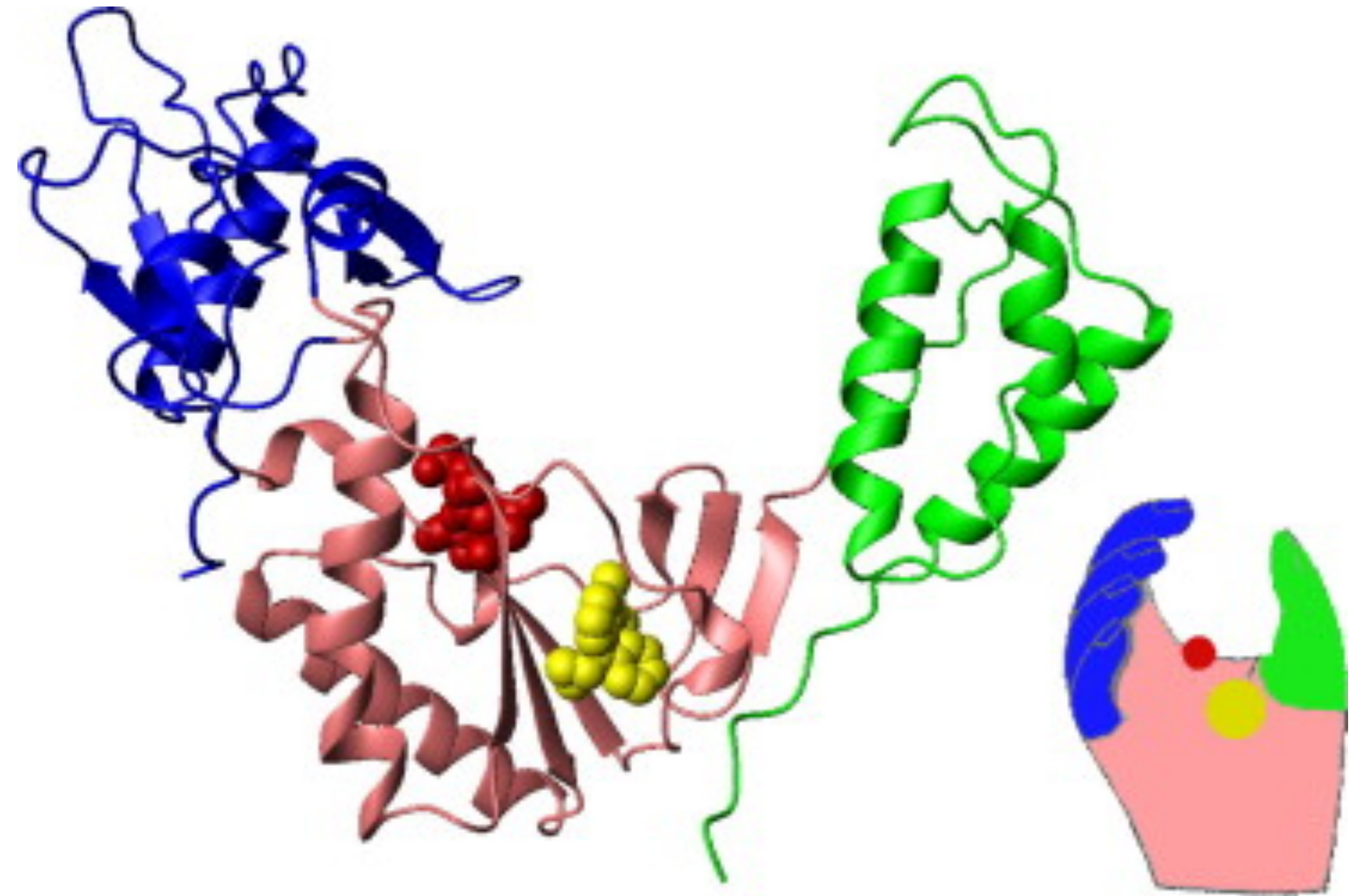
- Wiele białek metabolizmu DNA homologicznych w całym drzewie życia (gyrazy, topoizomerazy)
- Ale nie polimerazy
- Koonin – LUCA miał genom RNA
- Forterre - DNA jako materiał genetyczny pochodzi od wirusów

Genom LUCA: DNA czy RNA?

- Ale...
 - monofiletyczna reduktaza rybonukleotydów
 - mało prawdopodobne niezależne powstanie reduktazy rybonukleotydów – reakcja kosztowna termodynamicznie
 - zachowane podobieństwo sekwencji między trzema domenami sugeruje dosyć wierną replikację u LUCA – trudno uzyskać z RNA

Genom LUCA: DNA czy RNA?

- domena “dłoni” (palm domain) wszystkich polimeraz DNA i odwrotnych transkryptaz ma wspólny (homologiczny) fałd ~150 aminokwasów
- katalizuje tworzenie wiązań fosfodiesterowych
- mogła być replikazą i transkryptazą
- łatwe przejście od replikazy RNA do polimerazy DNA



LUCA jako wspólnota genów

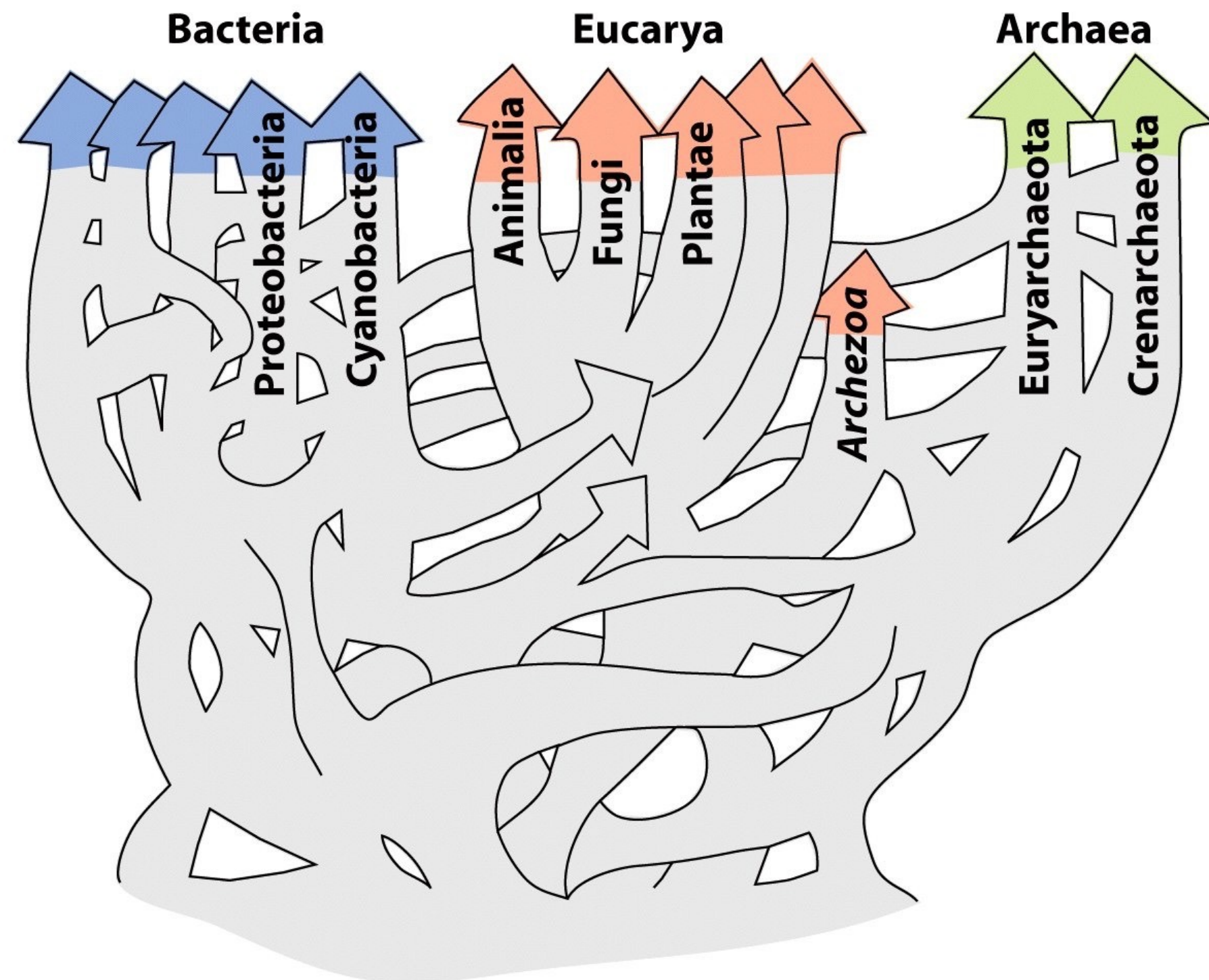


Figure 17-26 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

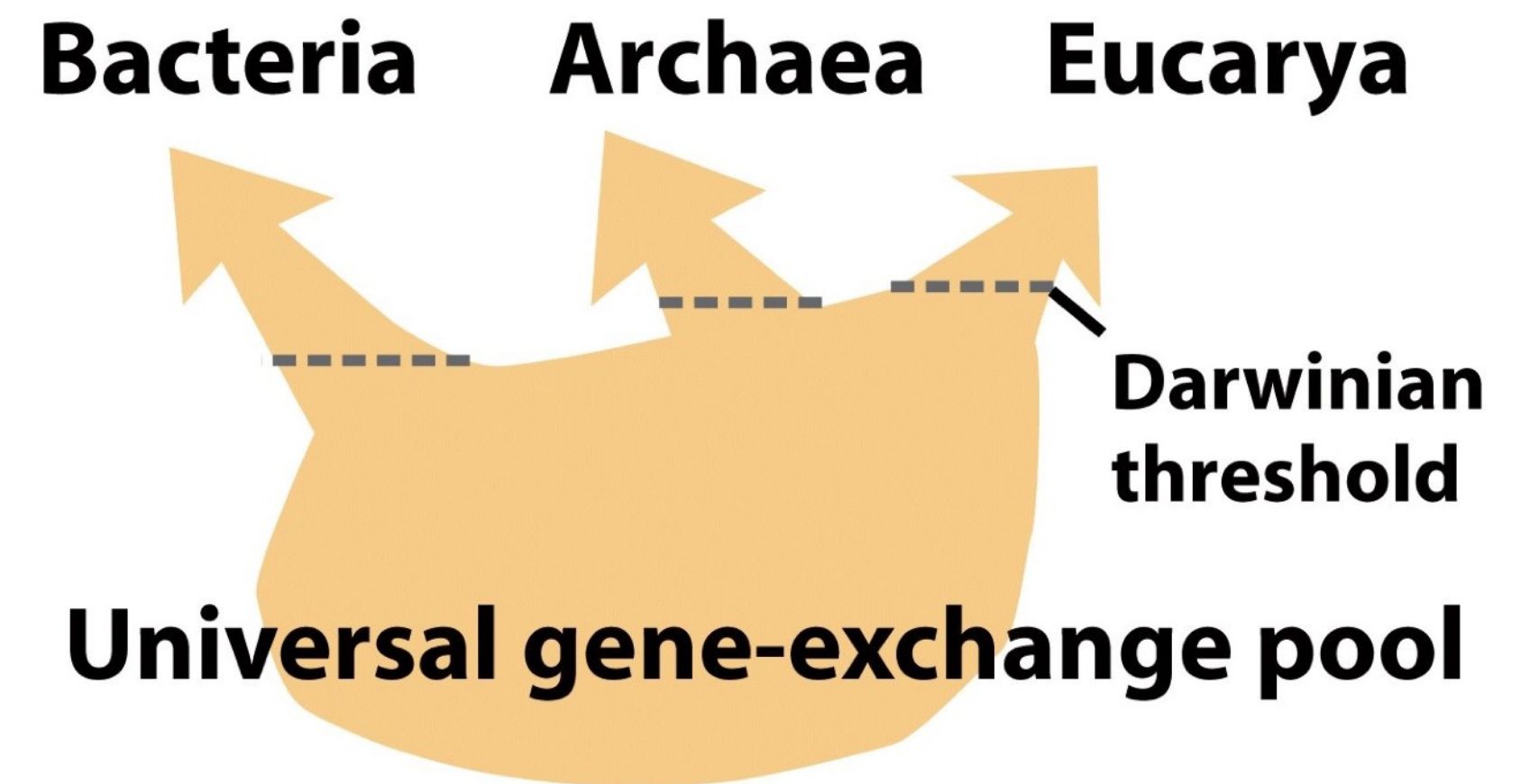


Figure 17-30 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

Niezależnie od tego, u podstawy był jednolity kod genetyczny i wspólne podstawowe mechanizmy genetyczne

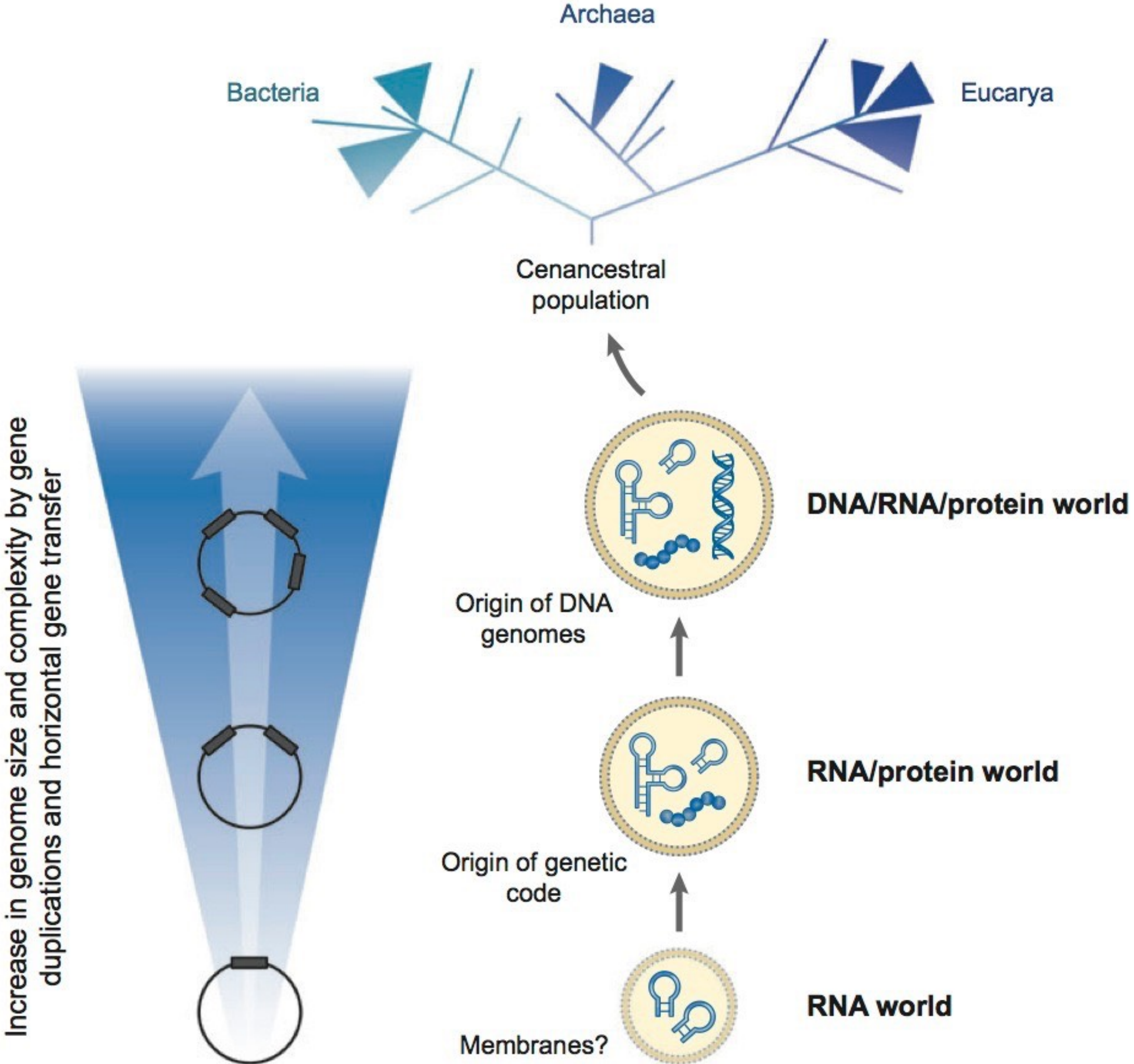
Czy LUCA był hipertermofilem?

- *Aquifex* i *Thermotoga* to wcześnie odgałęziające się grupy *Bacteria*
- Są termofilami i hipertermofilami, podobnie jak pierwotne *Archaea*
- Czy tolerancja wysokich temperatur była pierwotną cechą LUCA?
 - Niekoniecznie:
 - Pozycja tych linii w drzewie jest podważana
 - Cechy warunkujące termofilię mogły się przenosić w HGT
 - Przodkowie dzisiejszych *Bacteria* i *Archaea* przeszli epizod selekcji w wysokich temperaturach – wynikiem było uproszczenie struktury i funkcji (termoredukcja) – Forterre 1995

Termoredukcja

- Nie ma hipertermofilnych eukariontów
- RNA są niestabilne w wysokich temperaturach
- Eukarionty nie są w stanie zaadoptować się do wysokich temperatur ze względu na degradację mRNA
- Bakterie wyewoluowały swój model ekspresji (brak błony jądrowej, sprzęgnięcie translacji z transkrypcją) jako adaptację do wysokich temperatur

Podsumowanie



Hipotezy alternatywne

- Przesunięcie niektórych etapów prehistorii życia poza Ziemię
 - kosmiczne pochodzenie prostych cząsteczek organicznych
 - kosmiczne pochodzenie życia - panspermia

Panspermia

- Pierwsze cząsteczki biologiczne, a nawet organizmy nie powstały na Ziemi
 - Cząsteczki organiczne, aminokwasy w materiale kosmicznym
 - Problem ustalenia warunków początkowych
 - Jeżeli nie wiemy, gdzie powstawało życie, nie mamy możliwości formułowania hipotez
- “Panspermia ukierunowana” – życie celowo “zasiane” na Ziemi przez cywilizację (Orgel & Crick, 1973)



Cząsteczki organiczne z kosmosu

- Meteoryt z Murchinson
- liczne związki organiczne, w tym aminokwasy
- Komety (potwierdzona obecność związków organicznych)

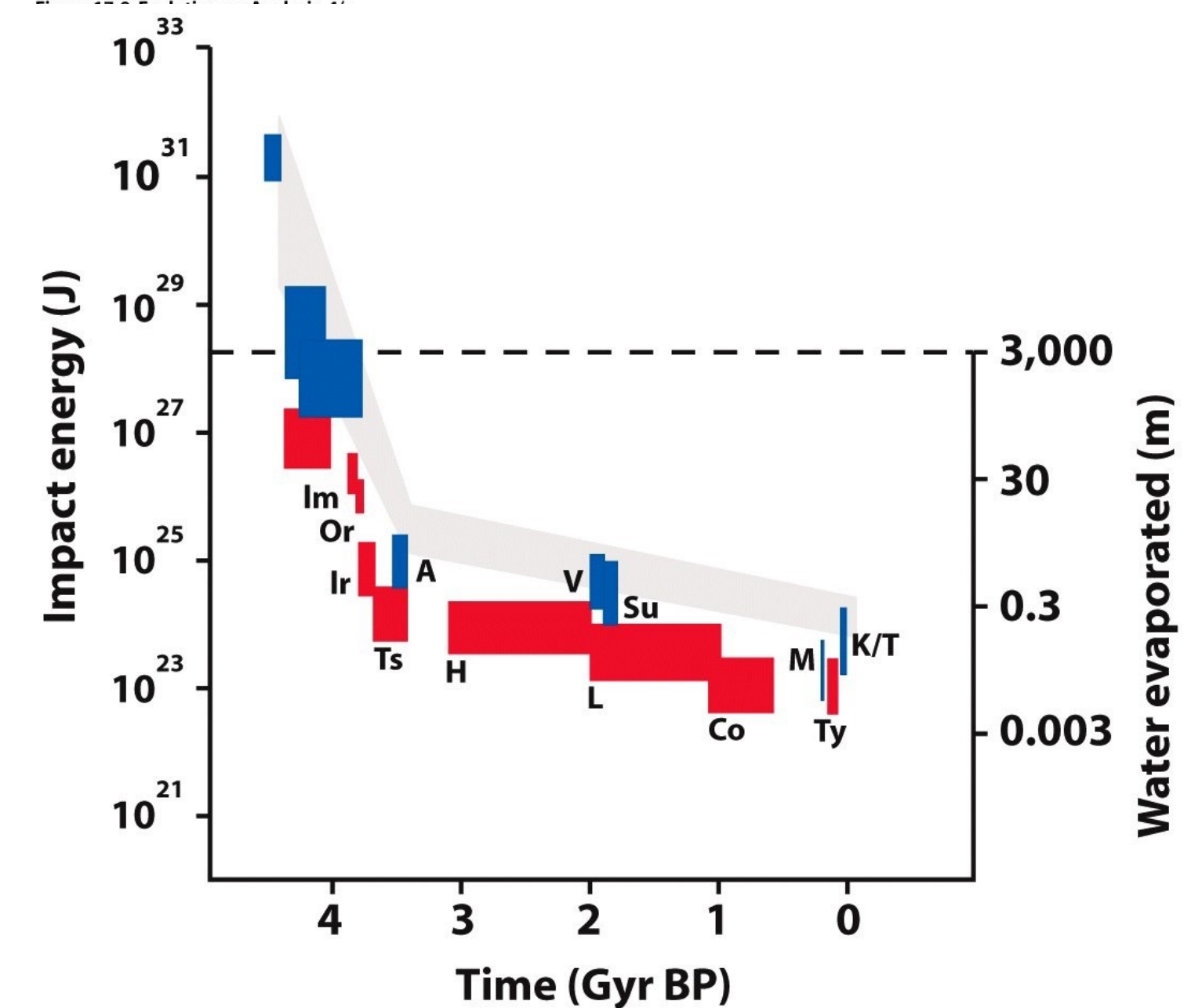


Figure 17-11 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

Błony z kosmosu

- Struktury lipidowe tworzone przez ekstrakty materiału z meteorytu z Murchinson (góra)
- Struktury tworzone w reakcjach fotochemicznych z materiału naśladującego lód z przestrzeni kosmicznej (środek)

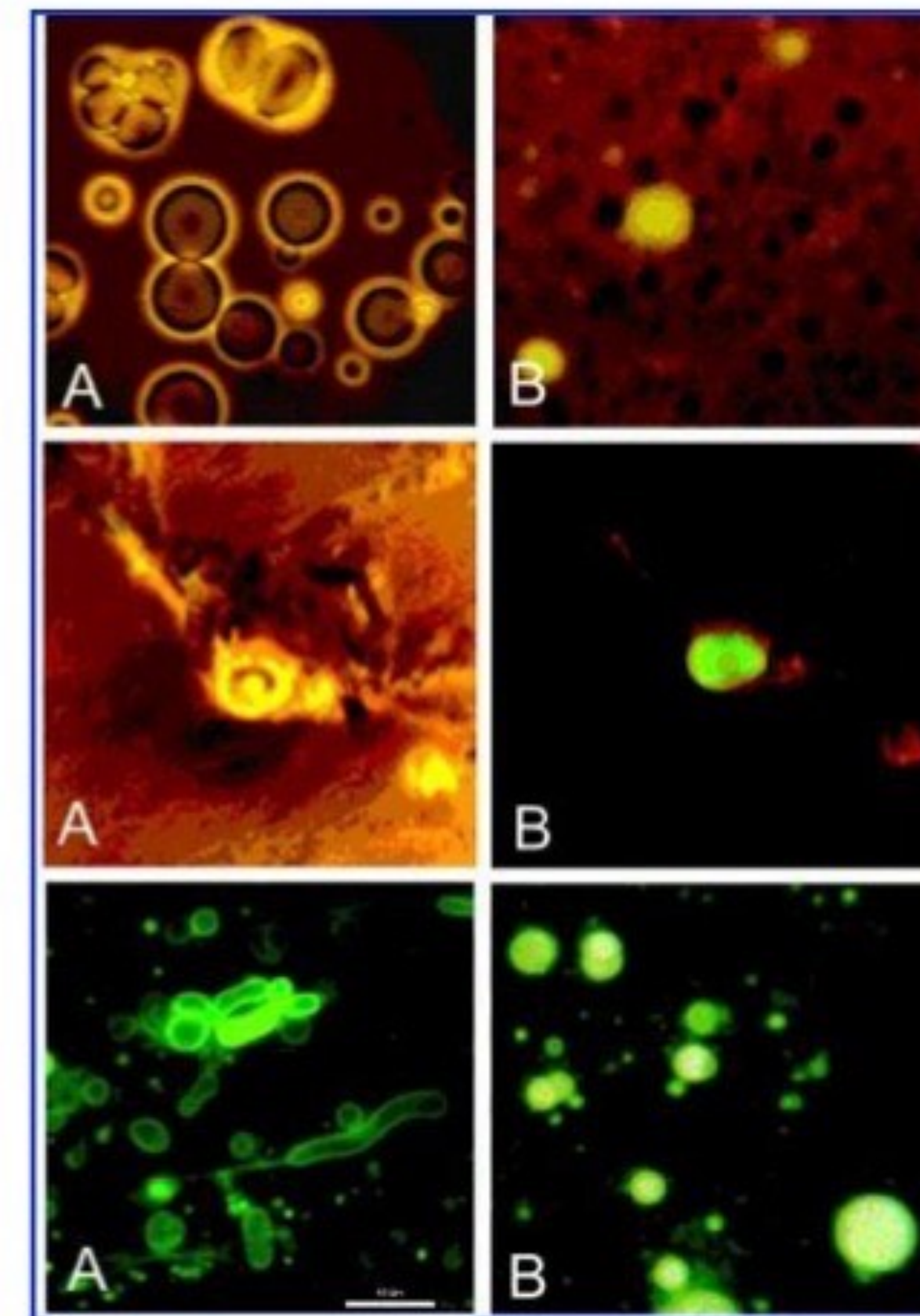
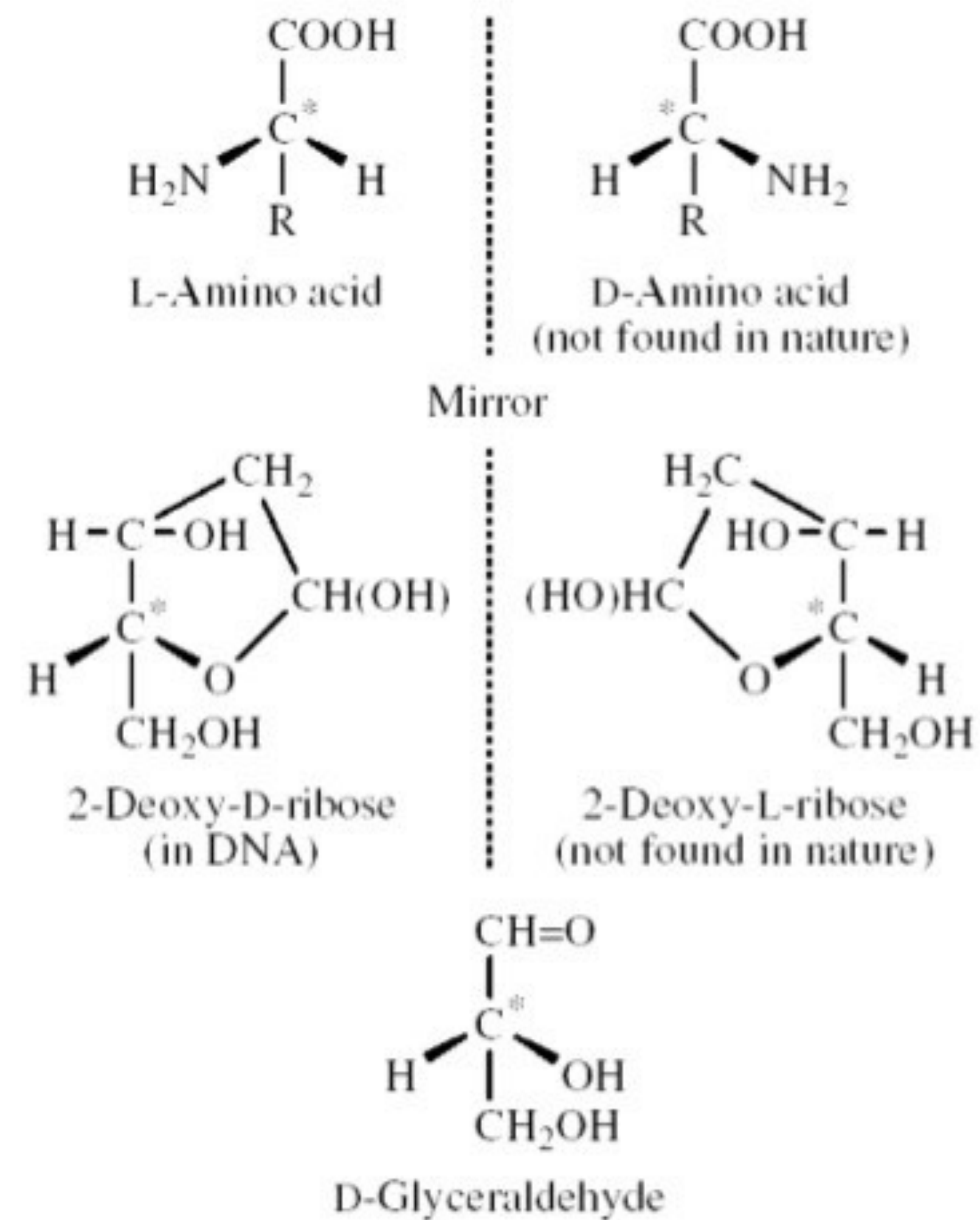


FIG. 3. Phase (A) and fluorescence (B) micrographs of membranous vesicular structures formed from a Murchison meteorite extract (Deamer, 1985; Deamer and Pashley, 1989, upper panel) compared with a vesicular structure produced in the extract from an interstellar ice analog (Dworkin *et al.*, 2001, center panel). Both the meteoritic vesicles and those synthesized photochemically from compounds in the ice analog have the ability to capture pyranine, a fluorescent anionic dye. **Lower panel:** Vesicles formed by a 20 mM decanoic acid/decanol mixture (1:1) at pH 8.0. Such vesicles readily encapsulate macromolecules as large as ~600-bp double-stranded DNA, shown on the right. The DNA has been stained with acridine orange.

Zagadka chiralności

- Życie zachowuje chiralność: L-aminokwasy, D-pentozy

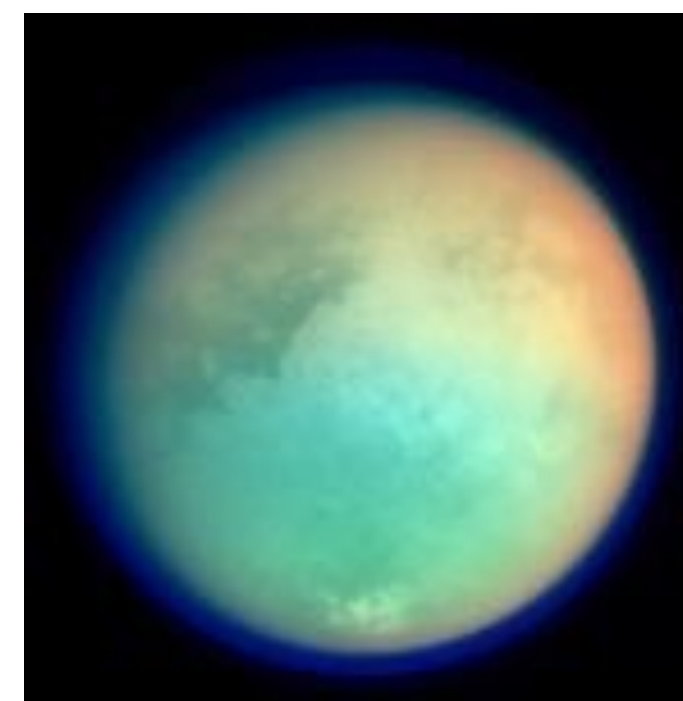


Astrobiologia

Aby szukać życia poza Ziemią należy zrozumieć, jak powstawało na Ziemi



Europa (księżyc Jowisza), © Wikimedia

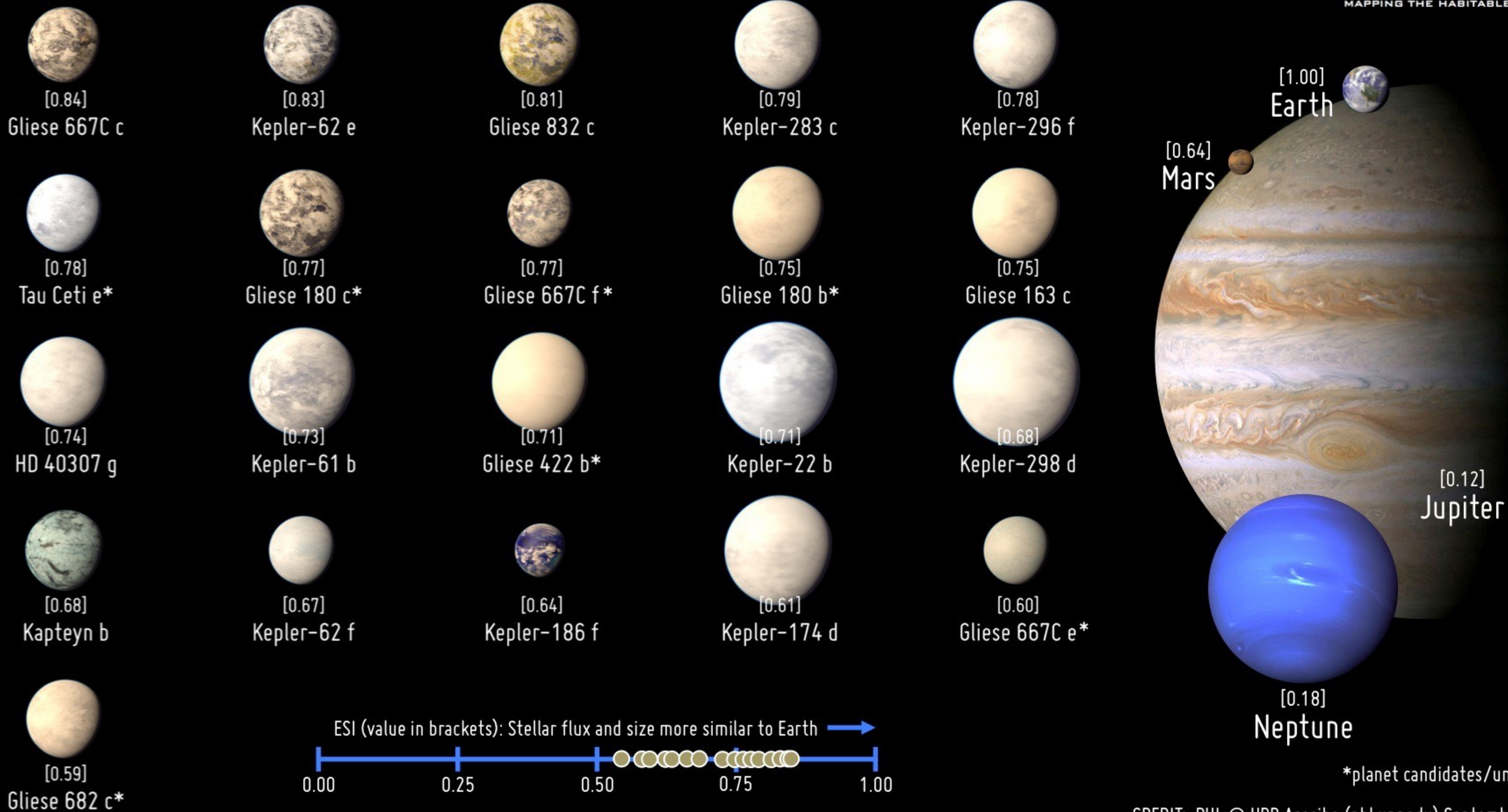


Tytan (księżyc Saturna), © Wikimedia, New Scientist



Current Potentially Habitable Exoplanets

Ranked by the Earth Similarity Index (ESI)



*planet candidates/unconfirmed



Earth

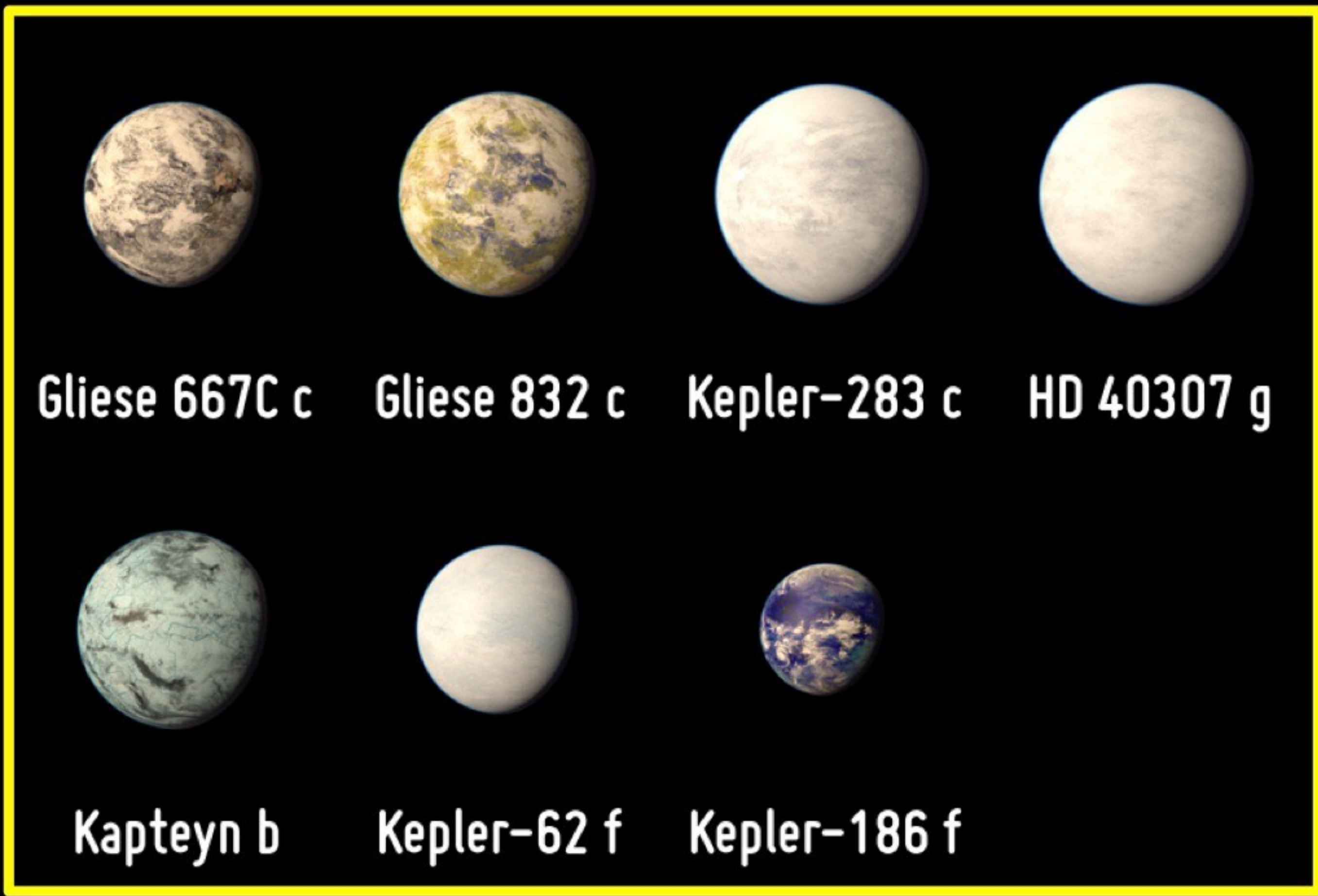
Conservative Selection of Potentially Habitable Exoplanets

0.1 to 10 Earth Masses or 0.5 to 2.0 Earth Radii, Conservative Habitable Zone for Super-Earth Planets




Confirmed

Unconfirmed



Gliese 667C c Gliese 832 c Kepler-283 c HD 40307 g

Kapteyn b Kepler-62 f Kepler-186 f



Gliese 180 c Gliese 422 b Gliese 682 c

Gliese 667C f Gliese 667C e

ARTISTIC REPRESENTATIONS