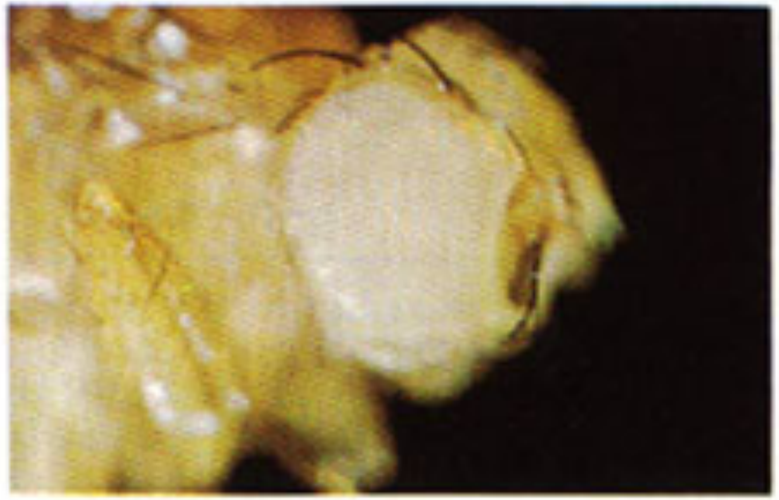
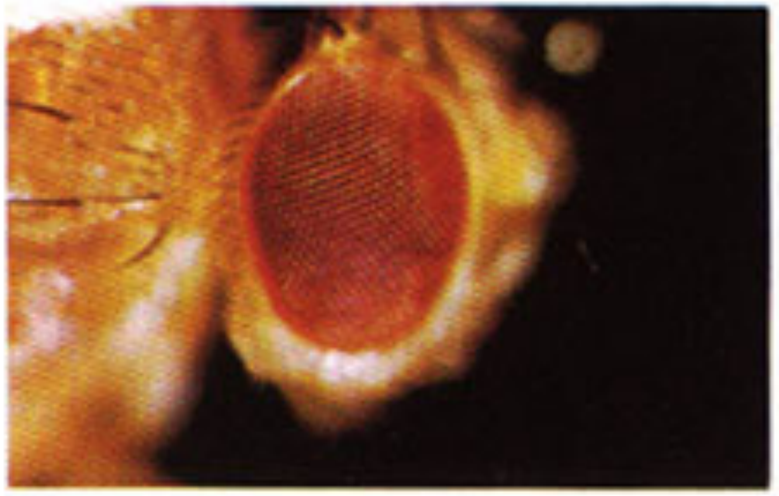
Metody i strategie genetyki i genomiki

Egzamin przełożony: 14.06.2019, 10:00

Przykłady mutantów *Drosophila*



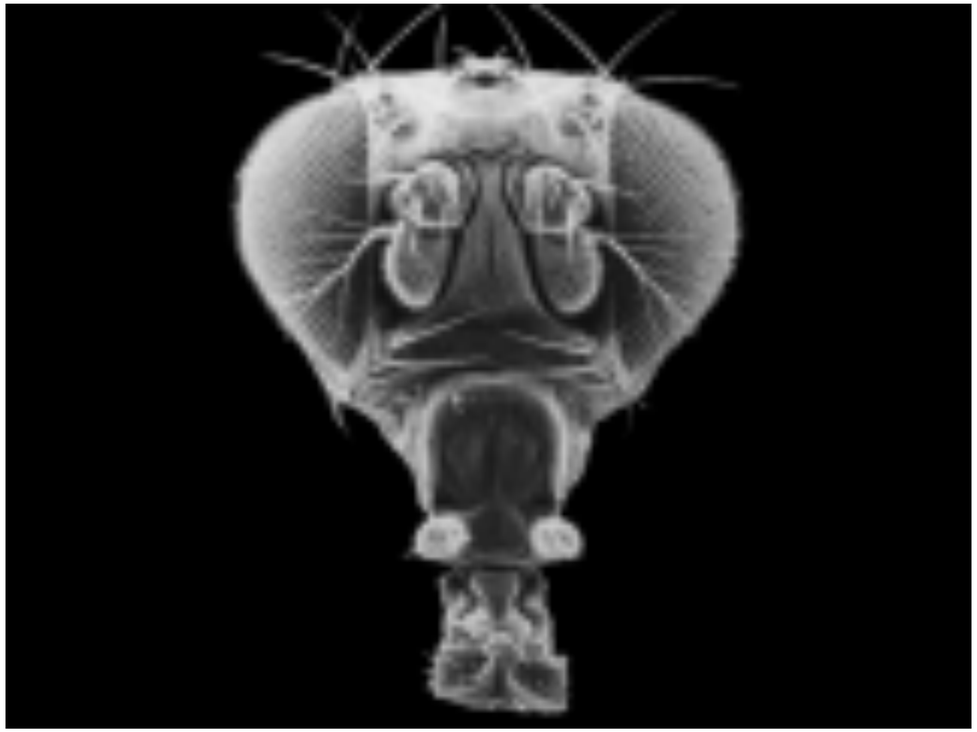
gen *white*: genotyp *w/w*



w^+/w^+ albo w^+/w

W. S Klug, M.R Cummings “Concepts of Genetics” 8th edition, Prentice Hall, 2005

Mutacje rozwojowe



Antennapedia

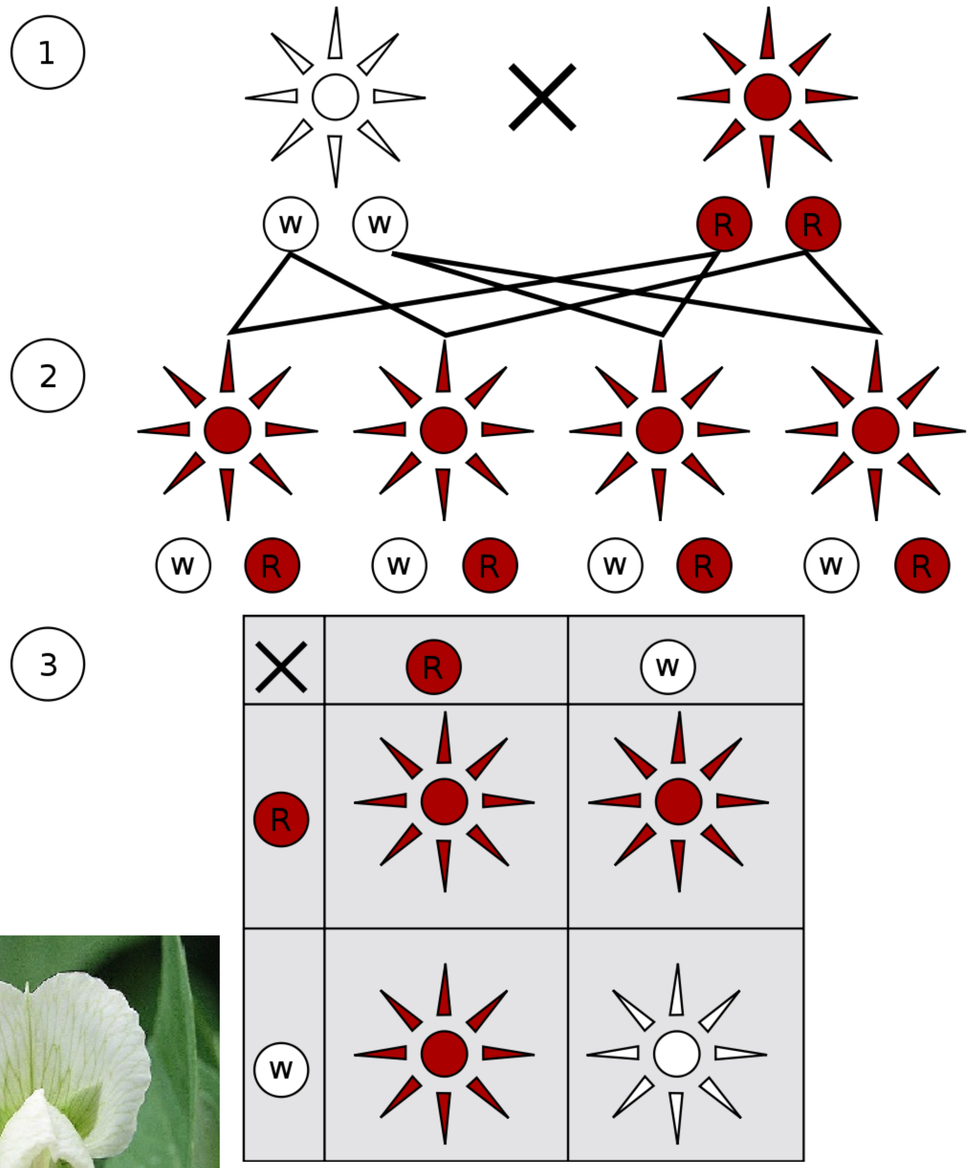
Mutacje rozwojowe

Ultrabithorax

Istota genetyki mendlowskiej (“szkolnej”)

- Jeden gen determinujący barwę kwiatu
- Allel “czerwony”, allel “biały”
- Jeden gen – jedna cecha

- Wystarczy, że poznamy wszystkie geny, a będziemy mogli opisać, zrozumieć i przewidzieć wszystkie cechy organizmu?



© wikimedia

Istota genetyki mendlowskiej

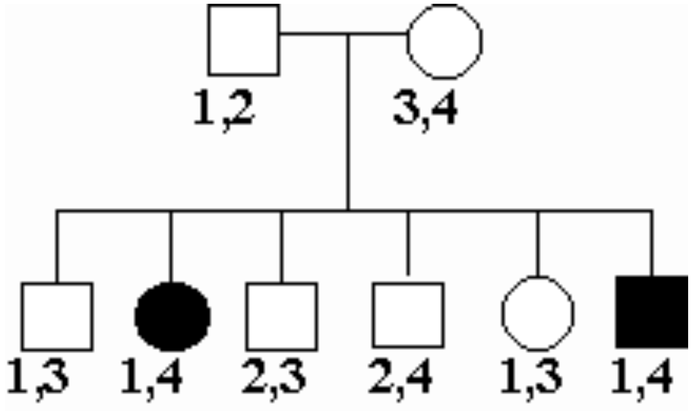
- **Za konkretną cechę fenotypową odpowiadają allele pojedynczego genu lub niewielkiej liczby genów**
- Prosta analiza genotypu pozwala przewidywać fenotyp
- Funkcję genu poznajemy dzięki analizie fenotypu mutanta utraty funkcji tego genu

Istota genetyki mendlowskiej

- Za konkretną cechę fenotypową odpowiadają allele pojedynczego genu lub niewielkiej liczby genów
  - np. barwa kwiatu groszku - allele pojedynczego genu
  - zasadniczo prawdziwe na poziomie molekularnym - geny i ich produkty (białka, RNA)
    - mutacja jednego genu wpływa na syntezę jednego białka (chyba, że nie)
    - ale jak często cecha fenotypowa zależy od aktywności pojedynczego białka (albo RNA)?

Człowiek jako obiekt w genetyce klasycznej

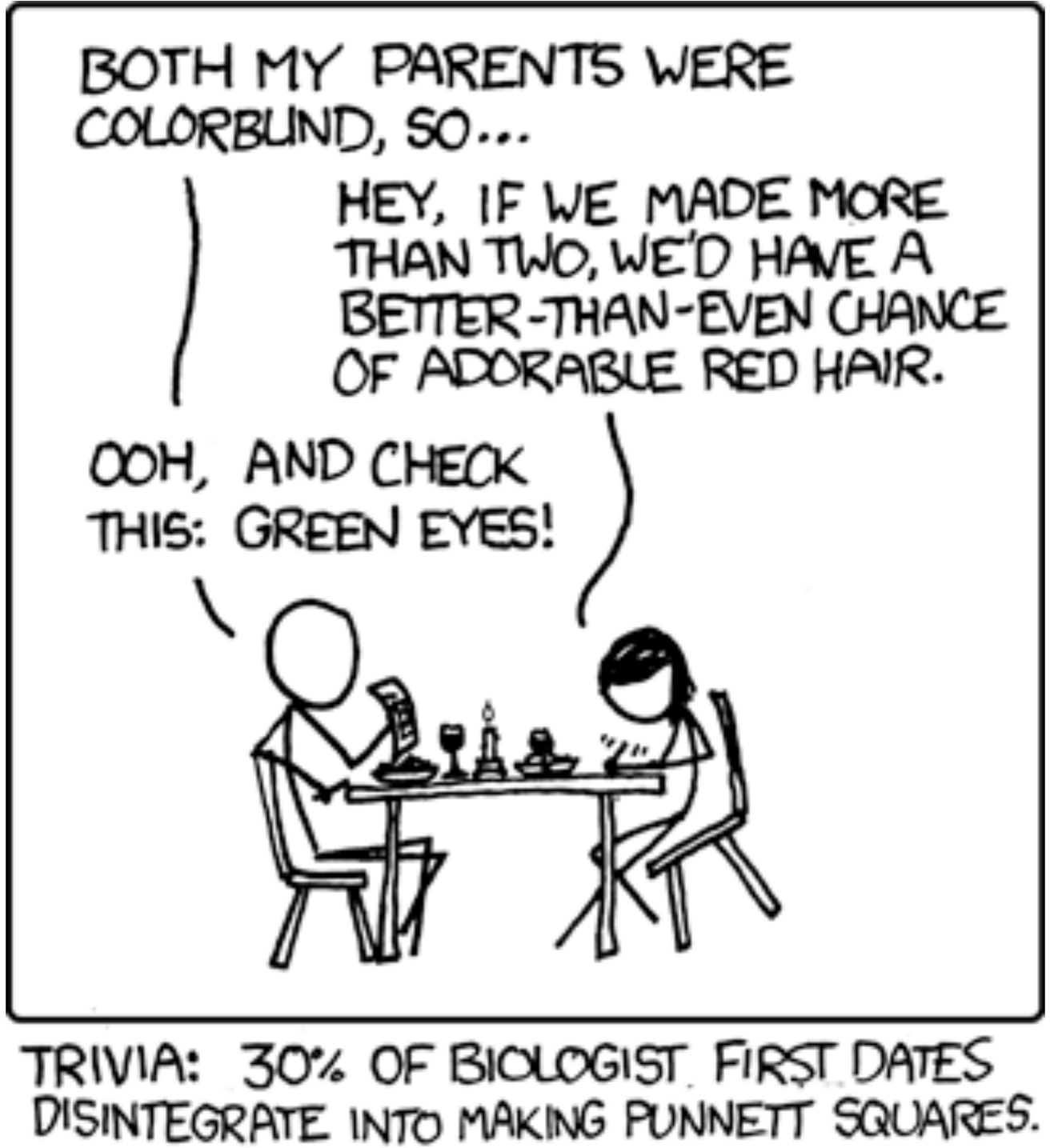
- ■ Mężczyzna (chory)
- □ Mężczyzna (zdrowy)
- ● Kobieta (chora)
- ○ Kobieta (zdrowa)
- ◇ Płeć nieokreślona



Cecha recesywna, autosomalna
Rodzice są heterozygotami (nosiciele)

Człowiek mendlowski?

- Na ile taka prosta genetyka pozwala na opisywanie zmienności fenotypowej człowieka?



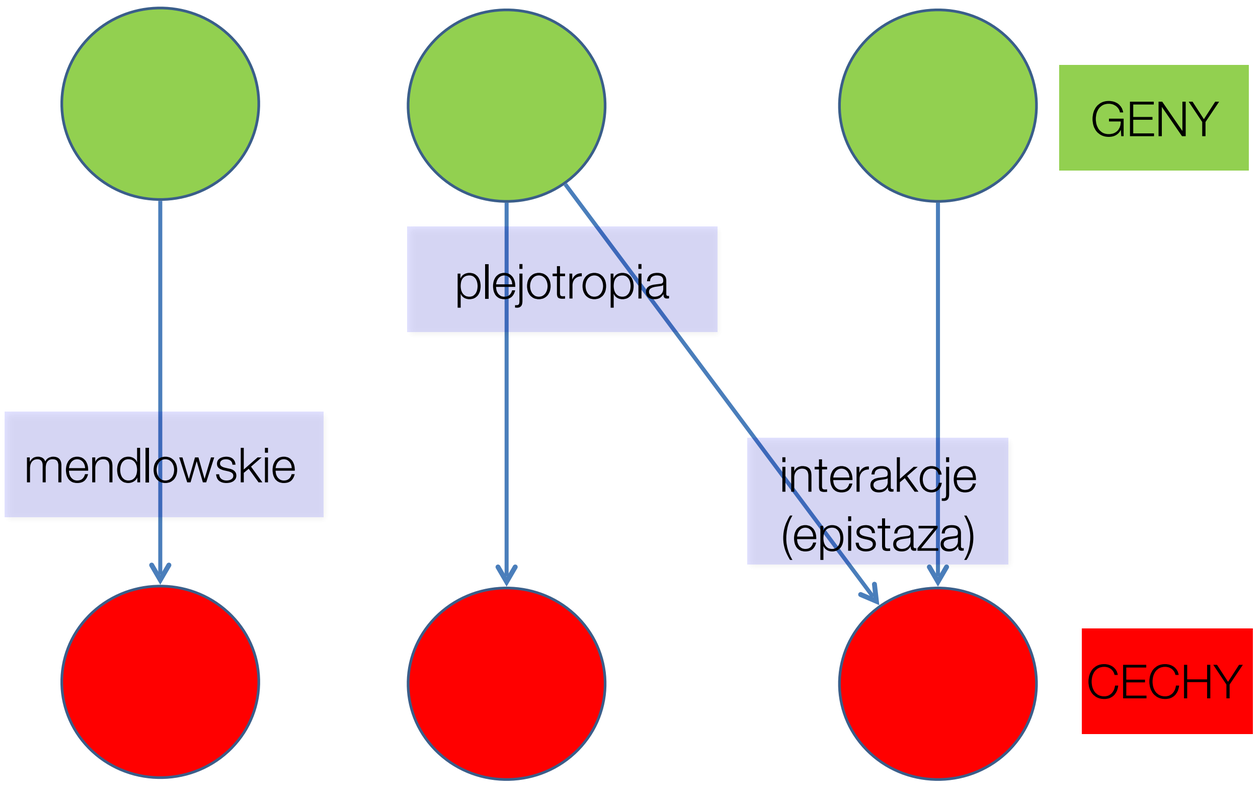
©xkcd.com

Istota genetyki mendlowskiej

- Prosta analiza genotypu pozwala przewidywać fenotyp
  - krzyżujemy groszek o kwiatach białych i czerwonych, jakie będzie potomstwo F1 i F2 itp.
  - w przypadku nietrywialnych cech zmienności prawidłowej taki wzór dziedziczenia jest rzadko spotykany
  - **cechy wieloczynnikowe** - na fenotyp wpływają interakcje wielu genów i czynników środowiskowych
  - największe wyzwanie współczesnej genetyki, np. w odniesieniu do człowieka

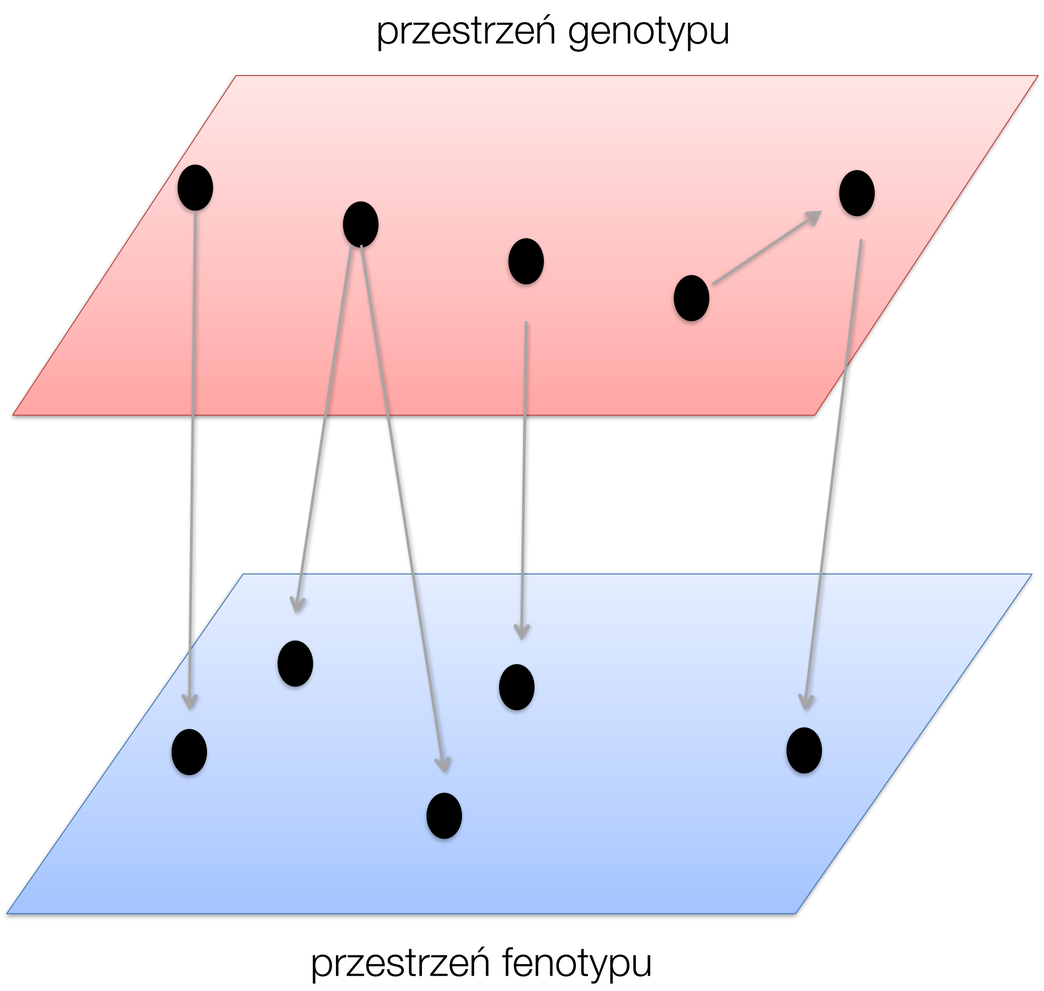
Plejotropia i interakcje genetyczne

- Plejotropia - jeden gen wpływa na wiele cech
  - rzadka u bakterii, częsta u złożonych eukariontów
- Interakcje genetyczne (epistaza) - wiele genów wpływa na jedną cechę

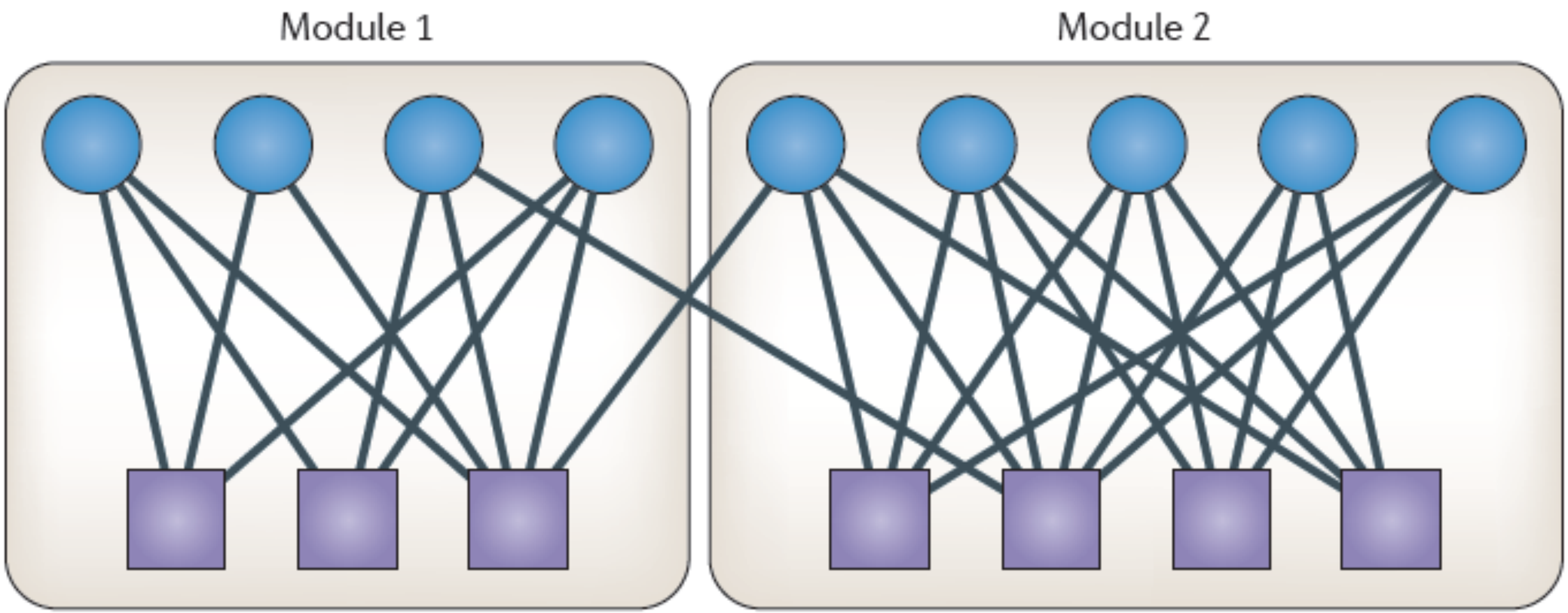


Mapa genotyp-fenotyp

- Genom człowieka: około 23 tys. genów, o wiele więcej różnych cech
- Plejotropia i interakcje genetyczne generują złożoność
- Koszt złożoności wg. Fishera - im bardziej złożone relacje genotyp-fenotyp, tym mniej prawdopodobne, że losowa mutacja będzie korzystna

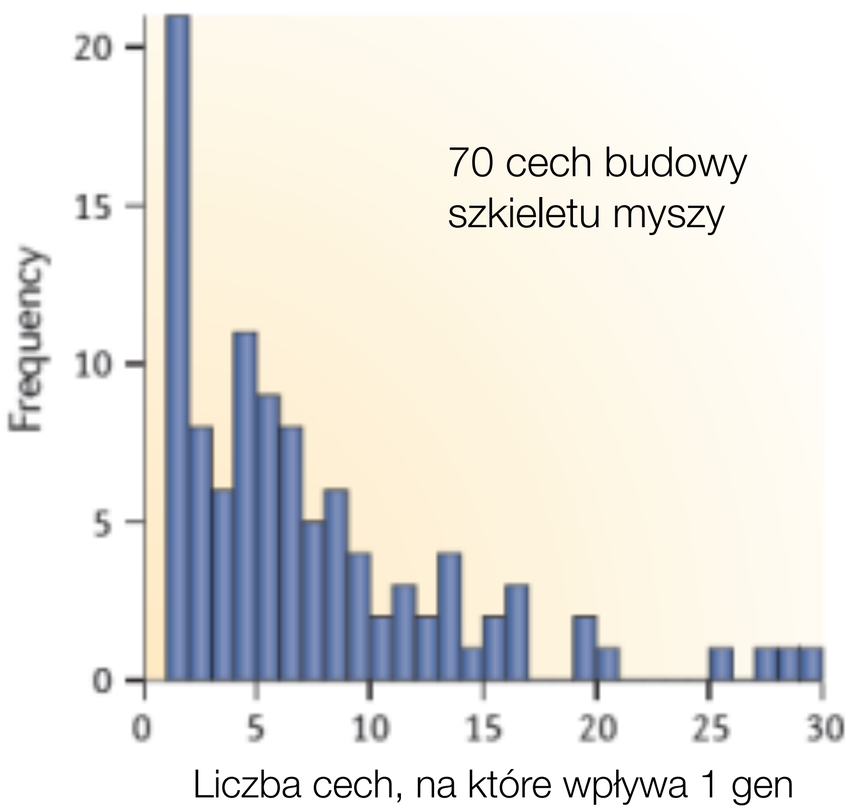


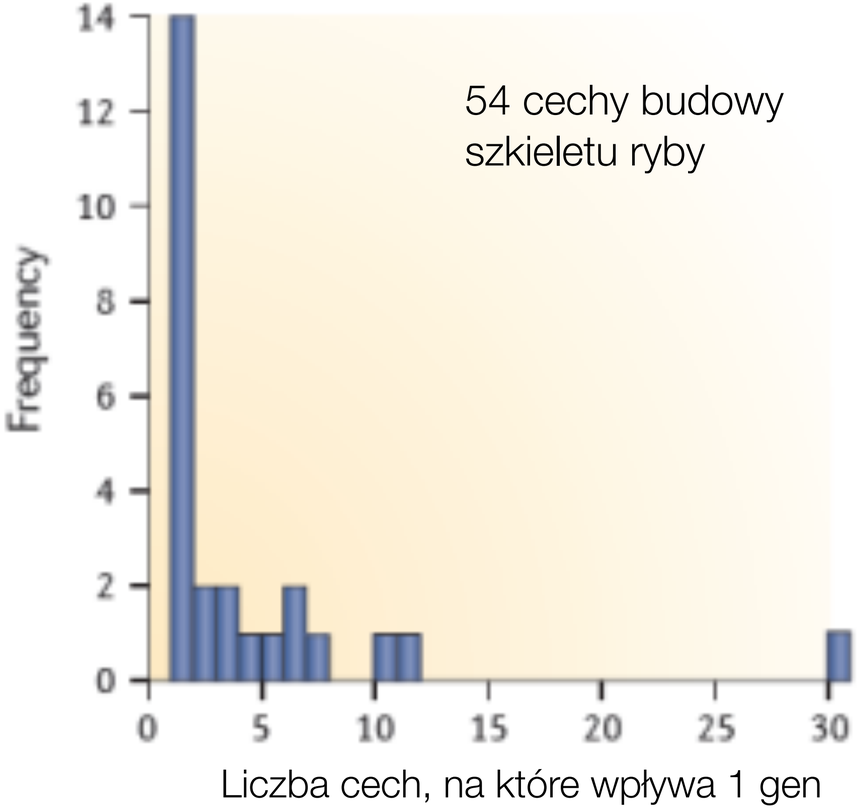
Modularność plejotropii



[Wagner and Zhang. Nature Reviews Genetics. 2011]

Jak częsta jest plejotropia?





[Wagner and Zhang. Nature Reviews Genetics. 2011]

Metoda poznawcza genetyki

- Mutant (np. bakterii) niezdolny do syntezy leucyny (*leu*) -> identyfikacja genu odpowiadającego za syntezę leucyny

- W przypadku bardziej złożonych relacji genotyp-fenotyp konieczna jest ostrożność w interpretacji

  - czy za wykształcenie cechy odpowiada pojedynczy gen?

  - czy za **zmienność** cechy odpowiada zmienność pojedynczego genu (to nie to samo)?

Analiza mutacji a zmienność prawidłowa

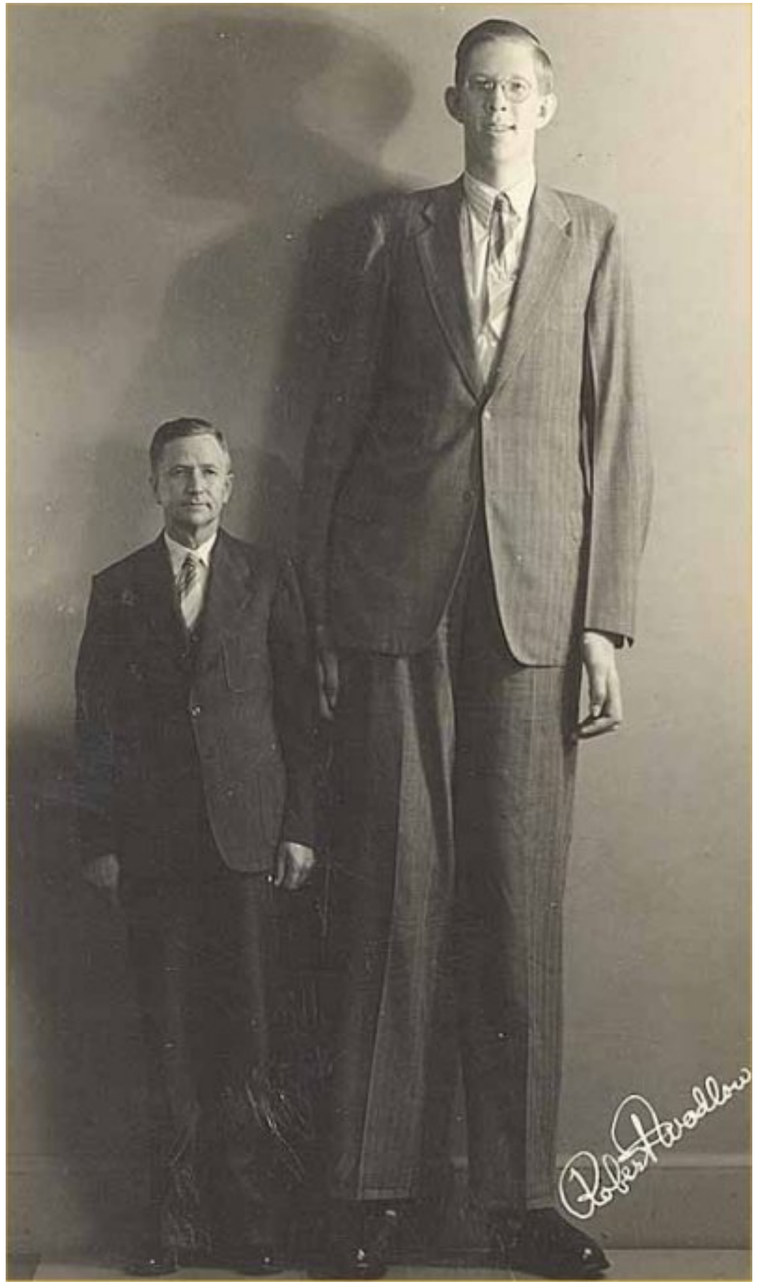
- Podstawa podejścia genetycznego - genotyp mutacji w genie daje informację o funkcji tego genu i jego produktu
- Mutacja pojedynczego genu może silnie zaburzyć określony proces
  - np. mutanty *wingless* u *D. melanogaster* - nie rozwijają się skrzydła
  - prosta analiza
  - ale w powstawaniu skrzydeł biorą udział produkty setek różnych genów
  - czy ten sam gen decyduje o prawidłowej zmienności tej cechy?

Defekty a zmienność prawidłowa

- Achondroplazja
- Mutacja pojedynczego genu *FGFR3* u człowieka
- cecha mendlowska, dominująca (letalna u homozygot)
- niski wzrost, nieproporcjonalnie krótkie kończyny

Defekty a zmienność prawidłowa

- Czy gen *FGFR3* wystarczy do zrozumienia genetyki różnic wzrostu u ludzi?
- Nie - to nie jest cecha jednogenowa
- W odziedziczalną zmienność wzrostu zaangażowane jest conajmniej 200 genów



GIANT (Genetic Investigation of Anthropometric Traits), Lango et al. Nature. 2010 467(7317):832-8.

Defekty a zmienność prawidłowa

- To, że mutacja jakiegoś pojedynczego genu całkowicie zaburza działanie jakiegoś systemu nie oznacza, że
  - jest on jedynym czynnikiem odpowiadającym za ten system
  - odgrywa on istotną rolę w prawidłowej (populacyjnej) zmienności tej cechy

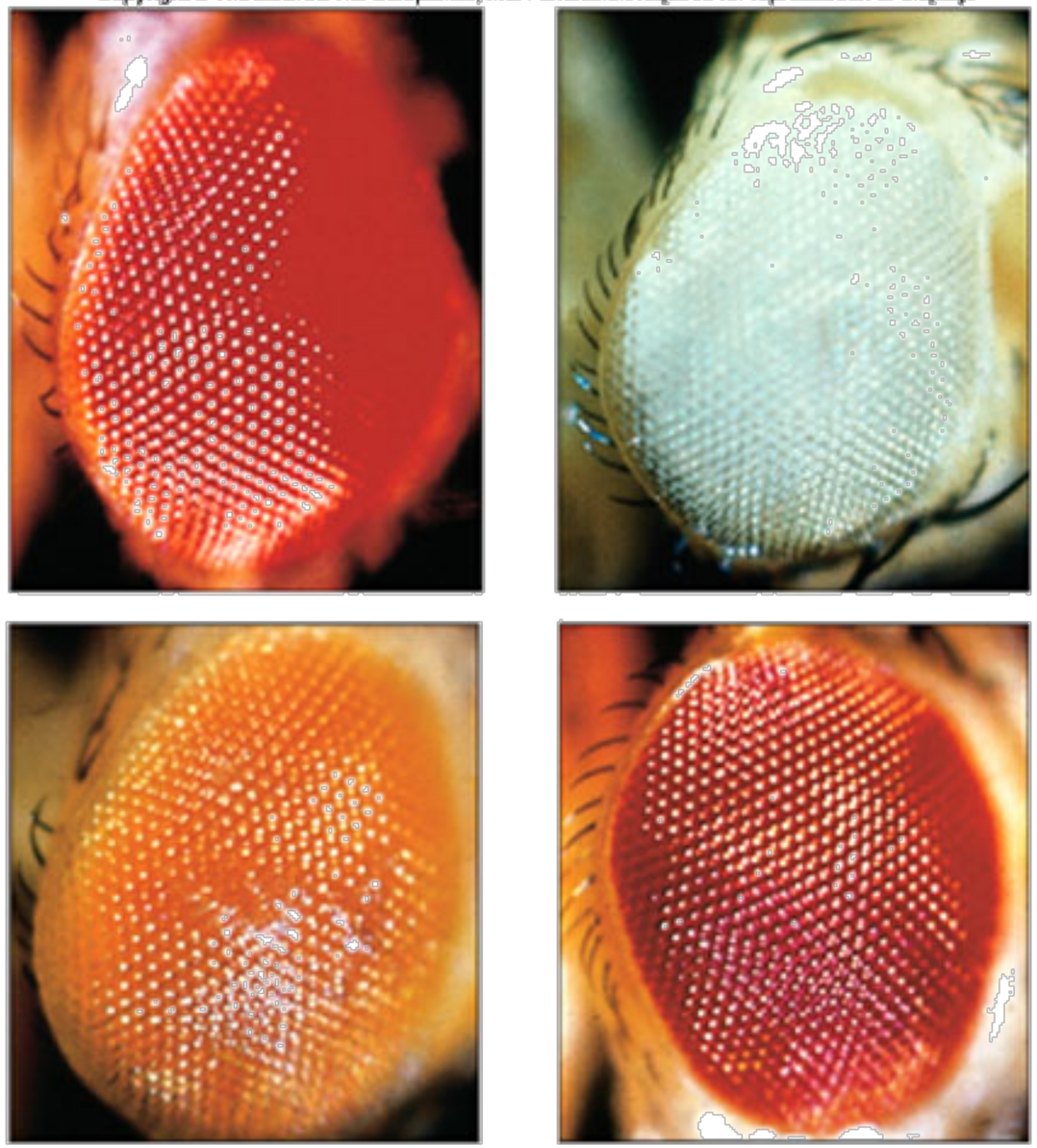
Podstawowe pytanie genetyki

- W jaki sposób genotyp determinuje fenotyp?
- Dla cech wieloczynnikowych i zmienności prawidłowej wciąż wiadomo niewiele
- Badanie efektów mutacji w pojedynczych genach pomaga identyfikować elementy, ale nie wystarcza do zrozumienia całego systemu

Komplementacja

- Wiele mutacji dających taki sam, lub podobny fenotyp
- Czy są to mutacje w tym samym genie, czy w różnych
- Ile genów związanych jest z daną funkcją?

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



Podwójne heterozygoty cis i trans

m1, m2 – mutacje (bez znaczenia, czy w tym samym genie, czy w różnych

m1	m2
+m1	+m2

Układ *cis*

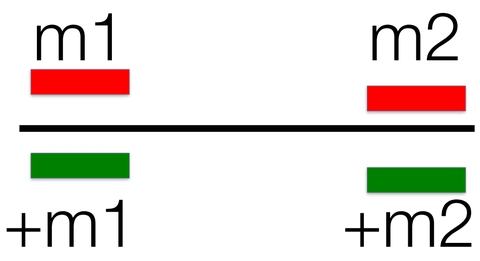
Otrzymywanie:
m1,m2 x wt (czyste linie)

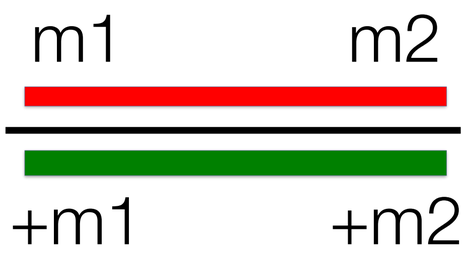
m1	+m2
+m1	m2

Układ *trans*

Otrzymywanie:
m1 x m2 (czyste linie)

Komplementacja

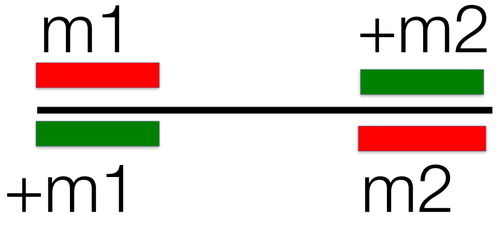




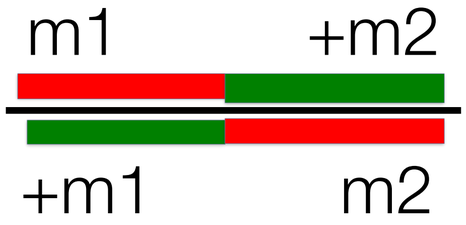
W układzie *cis* fenotyp zawsze dziki, niezależnie od tego, czy m1 i m2 są w tym samym genie, czy w różnych.

Warunek m1 i m2 recesywne.

Komplementacja



Jest funkcjonalny allel jednego i drugiego genu



Oba allele niefunkcjonalne

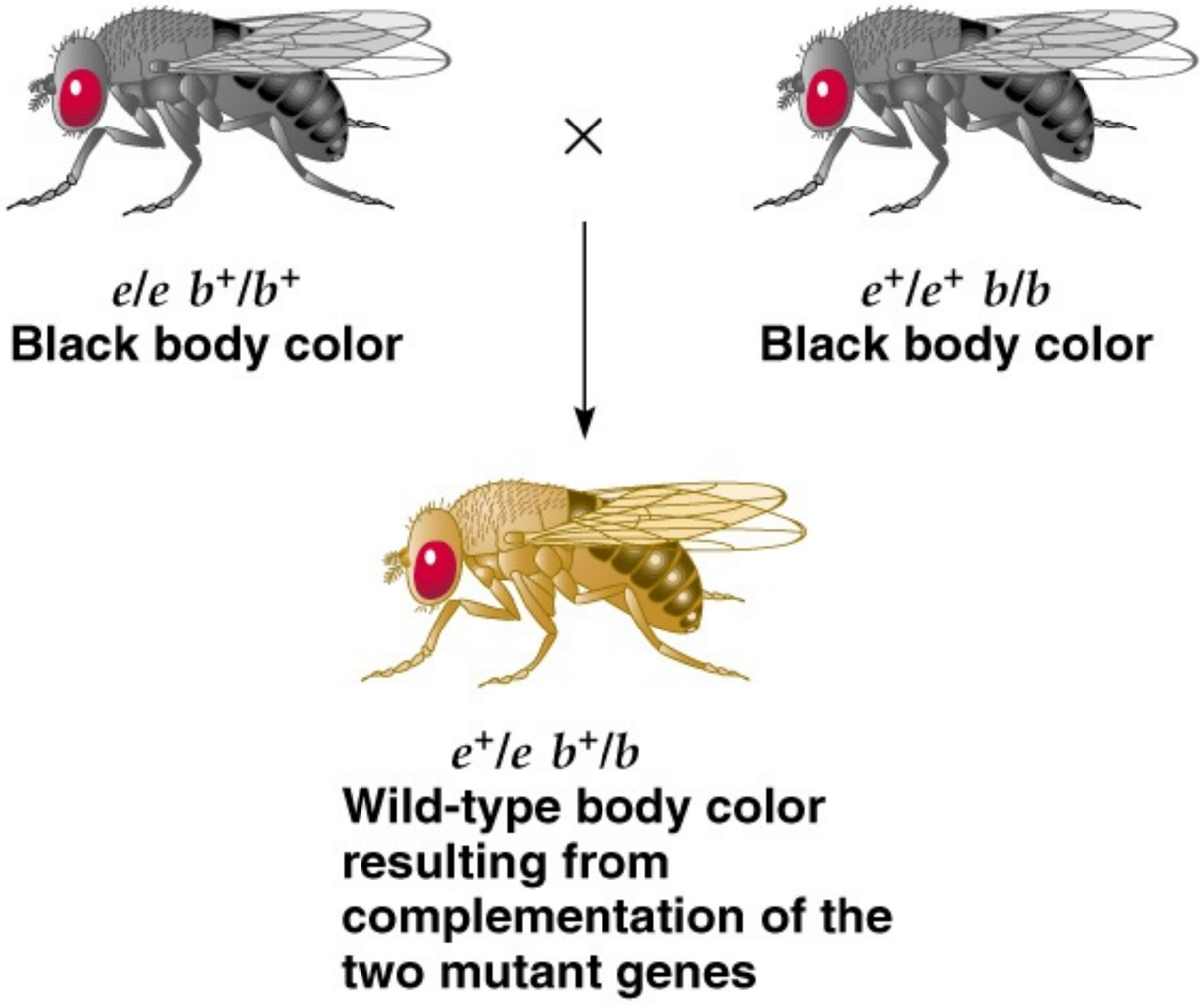
W układzie *trans* test daje odpowiedź

Warunek m1 i m2 recesywne.

Test komplementacji – wersja najprostsza

- Podwójna heterozygota trans
  - Fenotyp dziki – komplementacja, różne geny
  - Fenotyp mutanta – brak komplementacji – ten sam gen
- Tylko dla mutacji recesywnych

Komplementacja



Peter J. Russell, iGenetics: Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

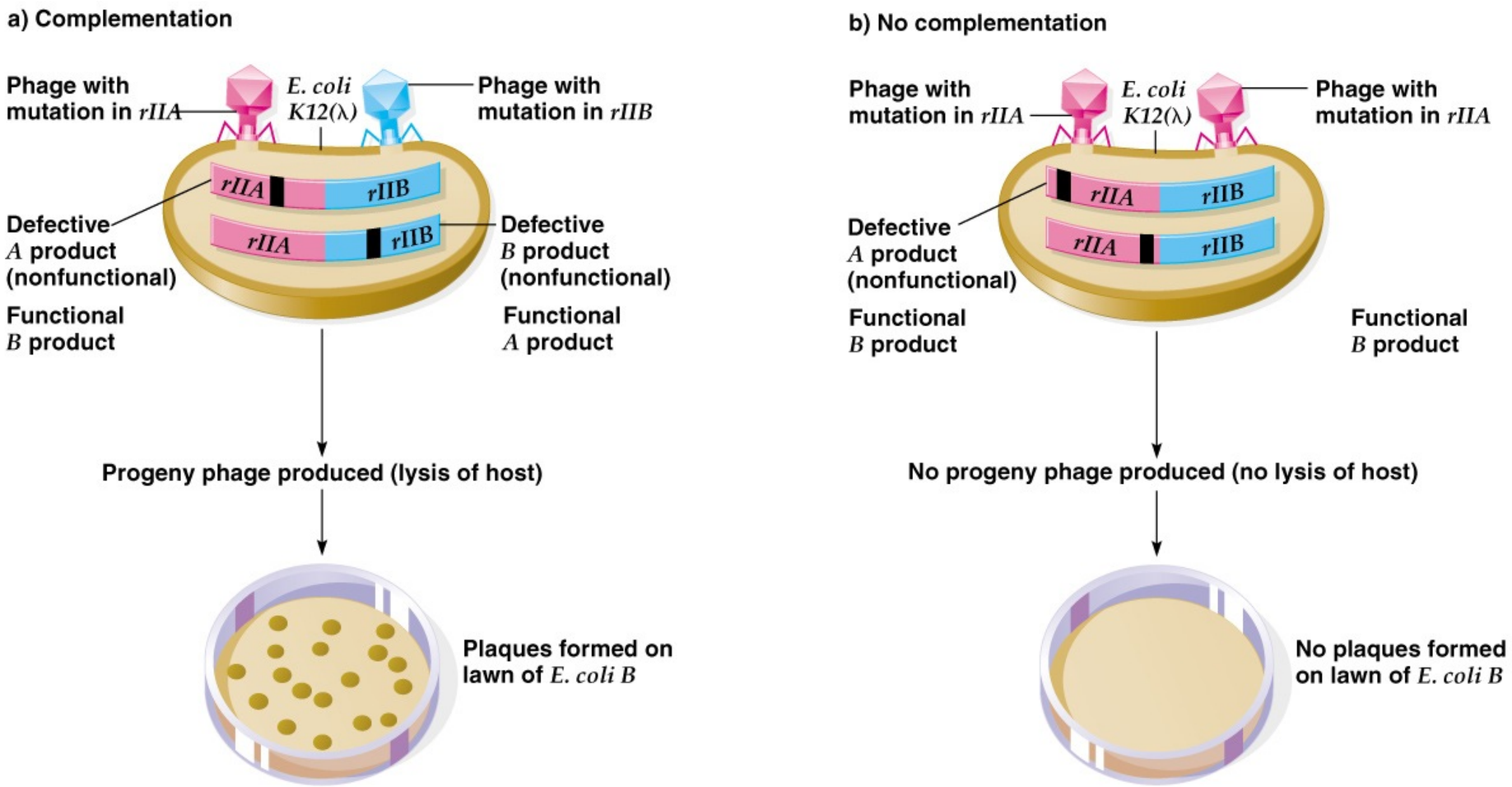
Grupy komplementacji

- *white* i *cherry* to allele tego samego genu
- *white* i *garnet* to allele w różnych genach
- Ile jest genów w tym doświadczeniu? Który gen ma wiele alleli?

Mutation	white	garnet	ruby	vermilion	cherry	coral	apricot	buff	carnation
white	–	+	+	+	–	–	–	–	+
garnet		–	+	+	+	+	+	+	+
ruby			–	+	+	+	+	+	+
vermilion				–	+	+	+	+	+
cherry					–	–	–	–	+
coral						–	–	–	+
apricot							–	–	+
buff								–	+
carnation									–

Cistron

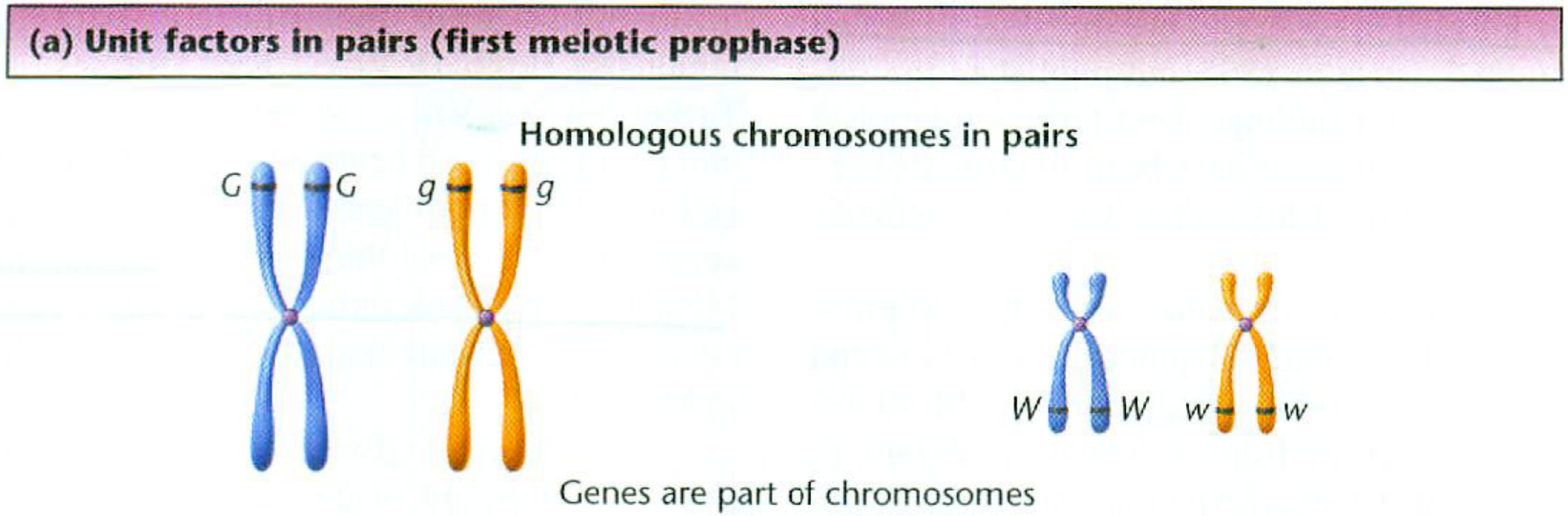
Mutacje w obrębie tego samego cistronu nie komplementują



Peter J. Russell, iGenetics: Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

Geny i chromosomy

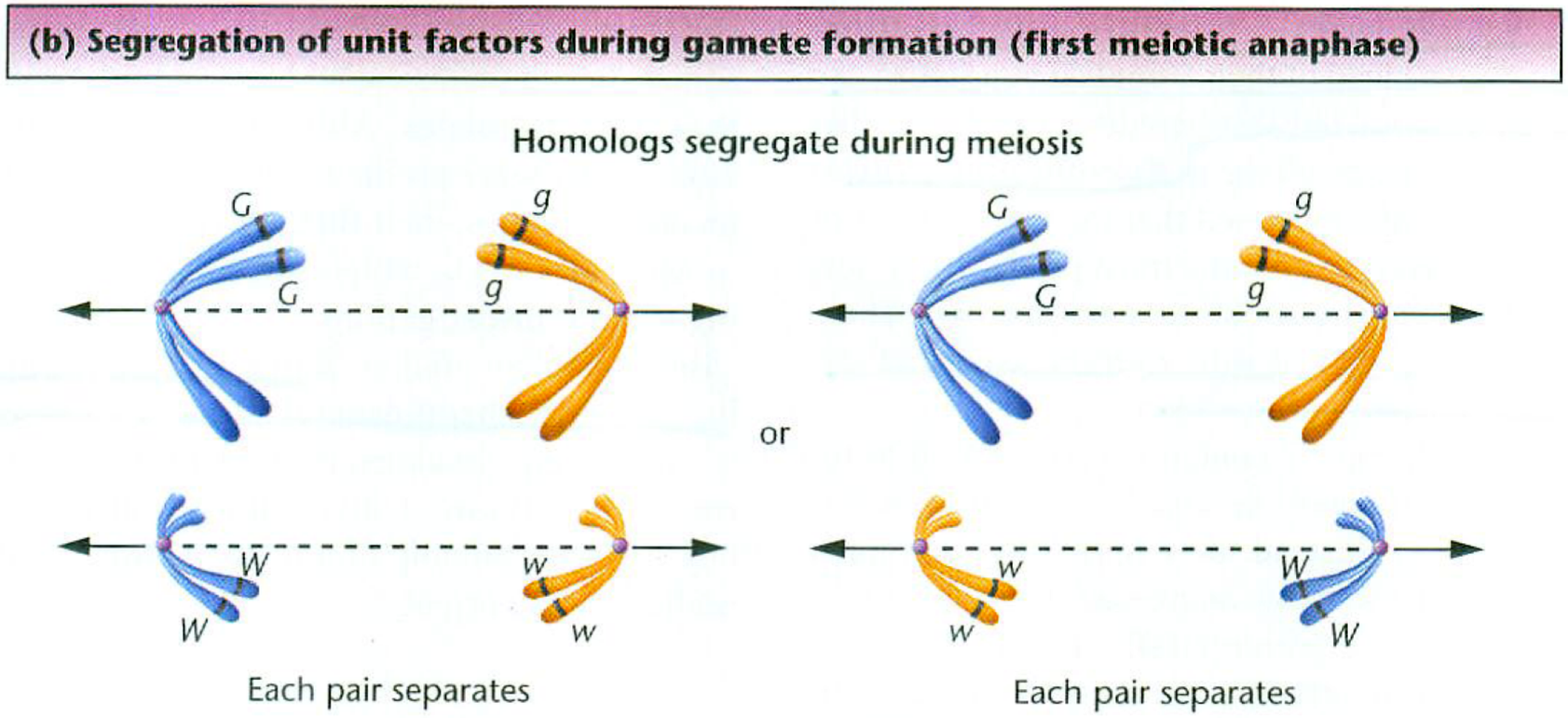
Dwa allele genu – dwa chromosomy homologiczne u organizmów diploidalnych



W. S Klug, M.R Cummings “Concepts of Genetics” 8th edition, Prentice Hall, 2005

Geny i chromosomy

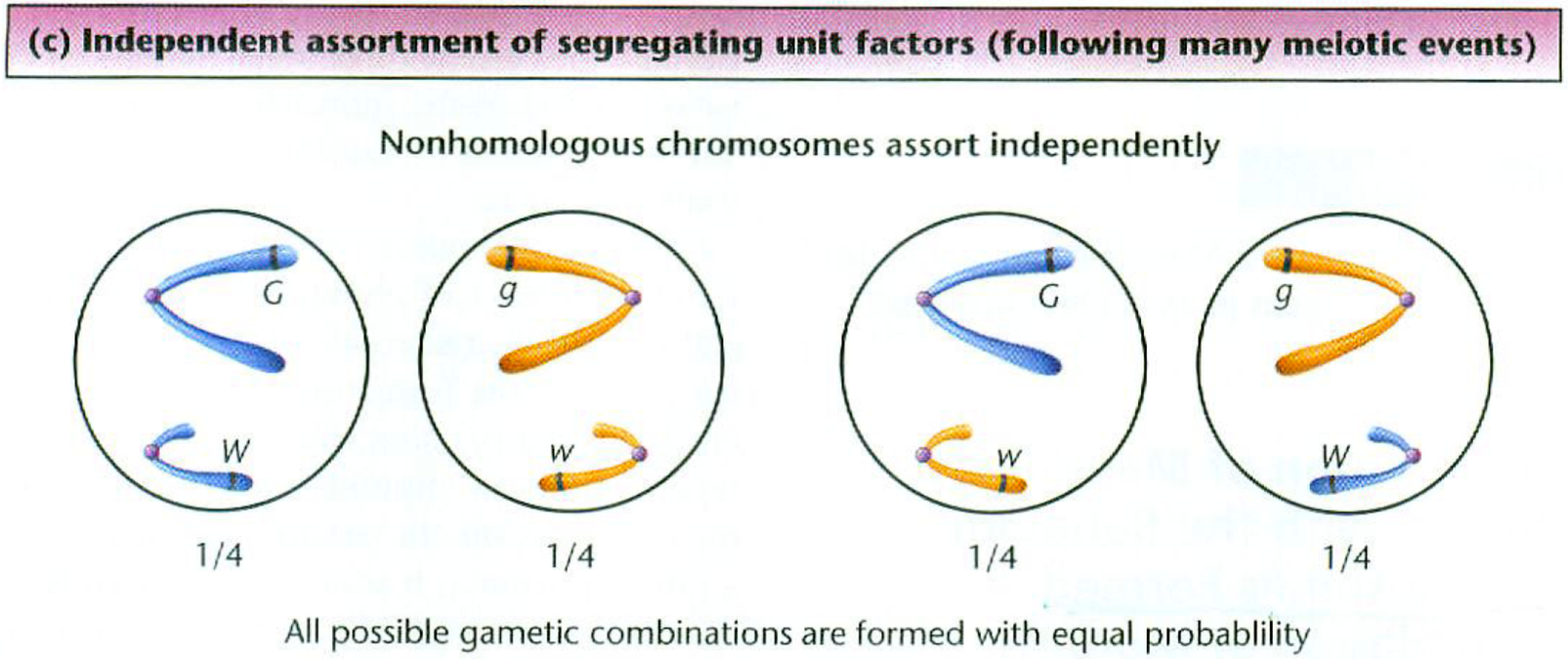
Segregacja alleli do gamet (I prawo Mendla) koreluje z zachowaniem chromosomów podczas mejozy



W. S Klug, M.R Cummings “Concepts of Genetics” 8th edition, Prentice Hall, 2005

Geny i chromosomy

Niezależne dziedziczenie alleli różnych genów – niezależna segregacja różnych chromosomów

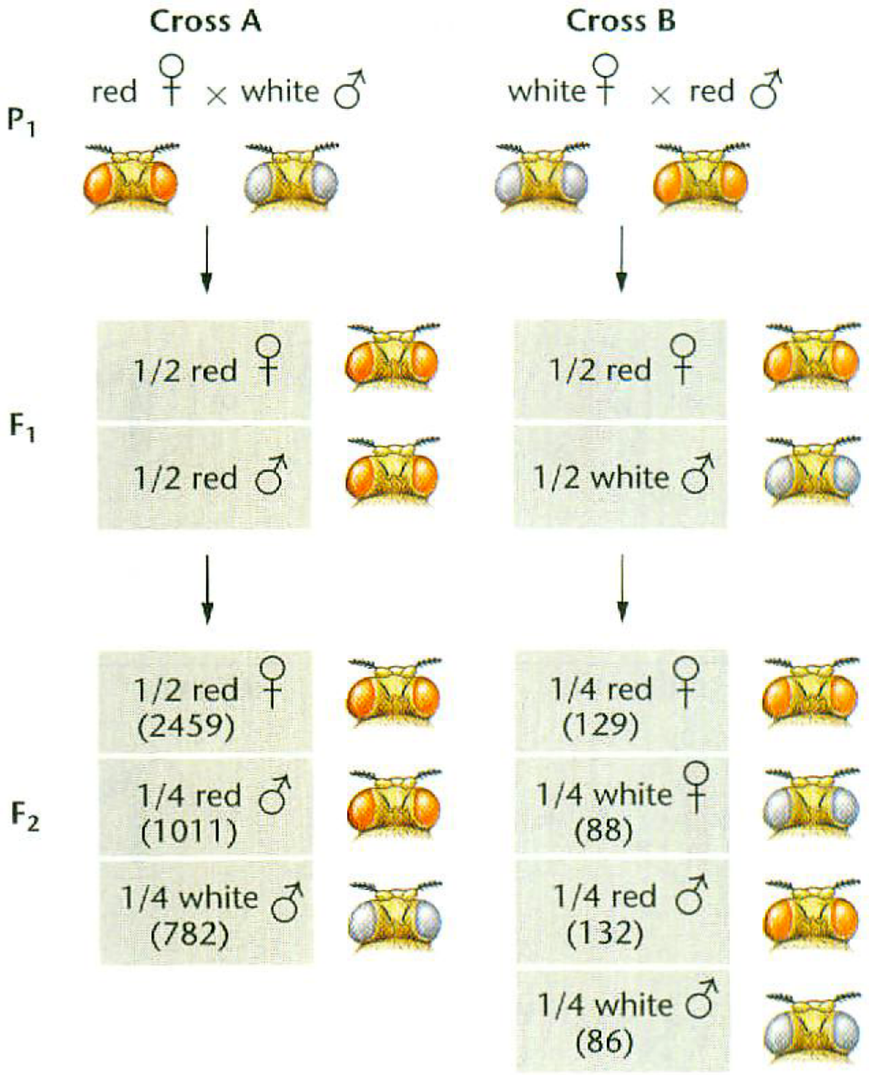


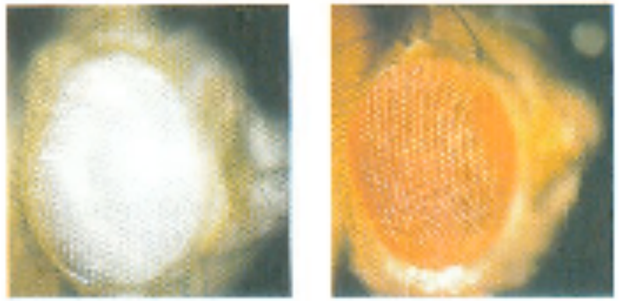
W. S Klug, M.R Cummings “Concepts of Genetics” 8th edition, Prentice Hall, 2005

Chromosomy płci

- U wielu (ale nie wszystkich) organizmów płeć jest determinowana przez specjalną parę chromosomów
- Ssaki łożyskowe
  - XX ♀; XY ♂
    - Y niezbędny do rozwoju fenotypu męskiego, X0 (zespół Turnera) fenotypowo kobiecy
- *Drosophila*
  - XX ♀; XY ♂
  - Fenotyp determinowany przez stosunek X do autosomów, X0 fenotypowo samiec (niepłodny u *D. melanogaster*)
- Ptaki, owady, niektóre jaszczurki
  - ZW ♀; ZZ ♂

Sprzężenie z płcią





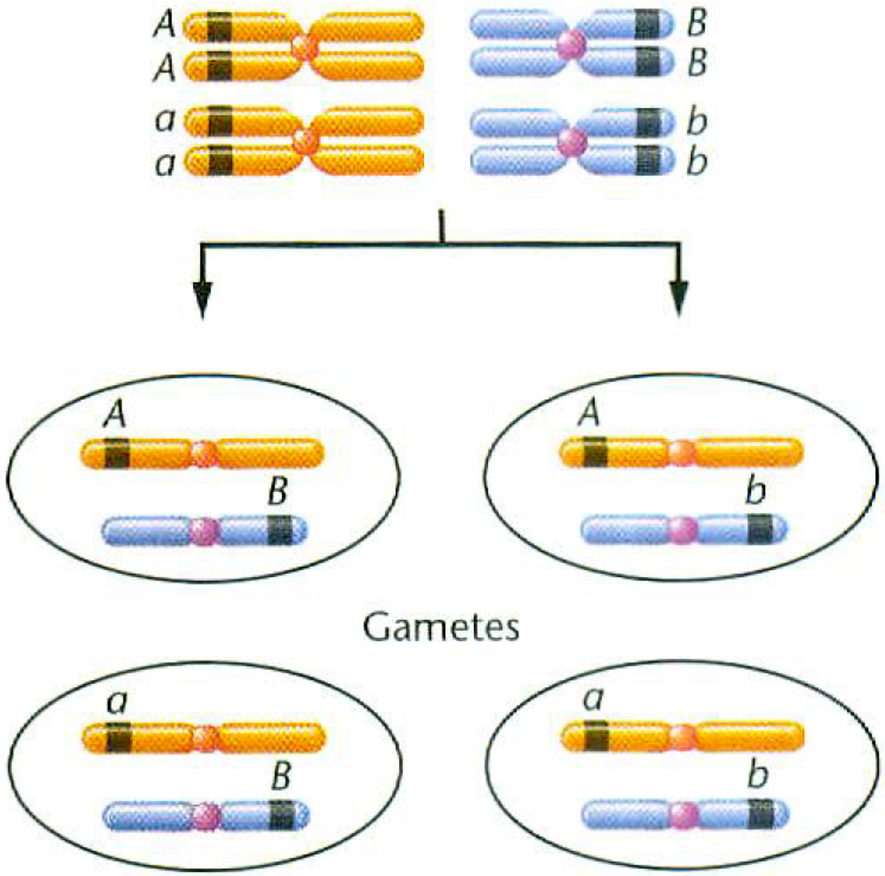
w wt (*w*+)

Thomas H. Morgan - 1910

W. S Klug, M.R Cummings “Concepts of Genetics” 8th edition, Prentice Hall, 2005

Sprzężenie

Geny leżące na różnych chromosomach spełniają II prawo Mendla

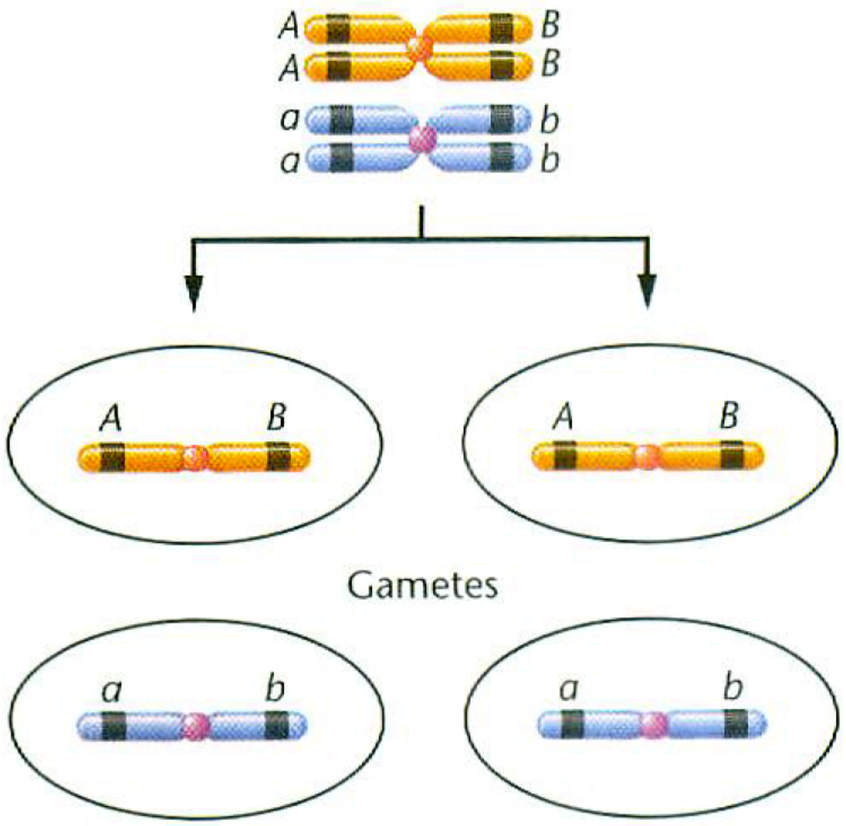


Dla 2 genów:
4 równoliczne klasy gamet

W. S Klug, M.R Cummings “Concepts of Genetics” 8th edition, Prentice Hall, 2005

Sprzężenie

Allele genów leżących na tym samym chromosomie dziedziczą się razem – sprzężenie

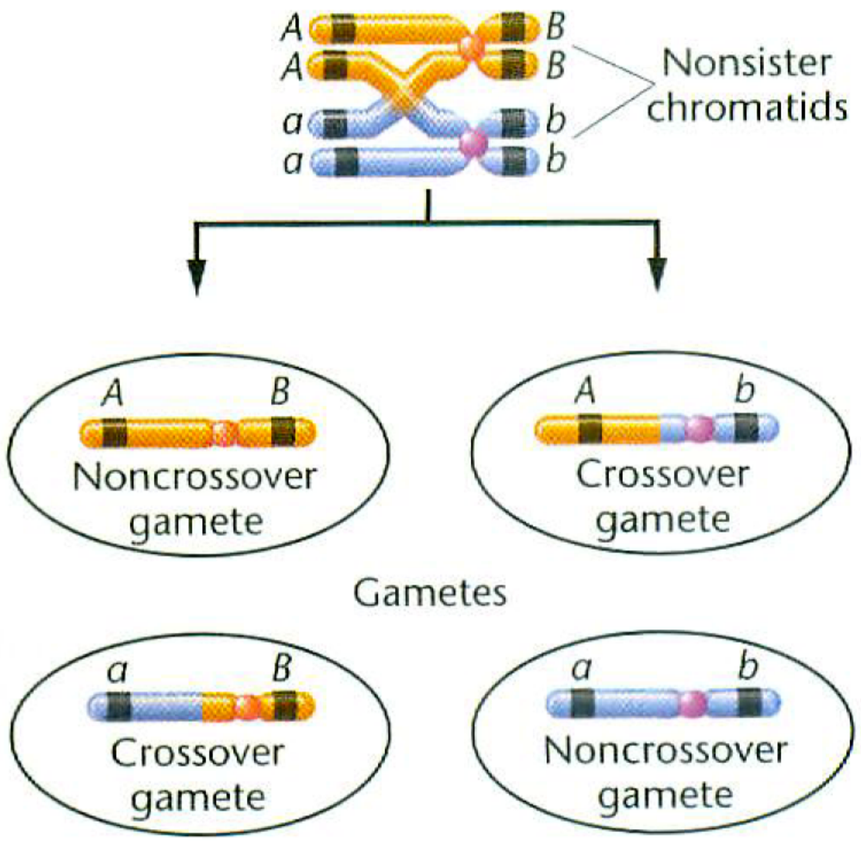


Dla 2 genów:
2 równoliczne klasy
gamet rodzicielskich

W. S Klug, M.R Cummings “Concepts of Genetics” 8th edition, Prentice Hall, 2005

Sprzężenie

Crossing-over (rekombinacja chromatyd niesiostrzanych)

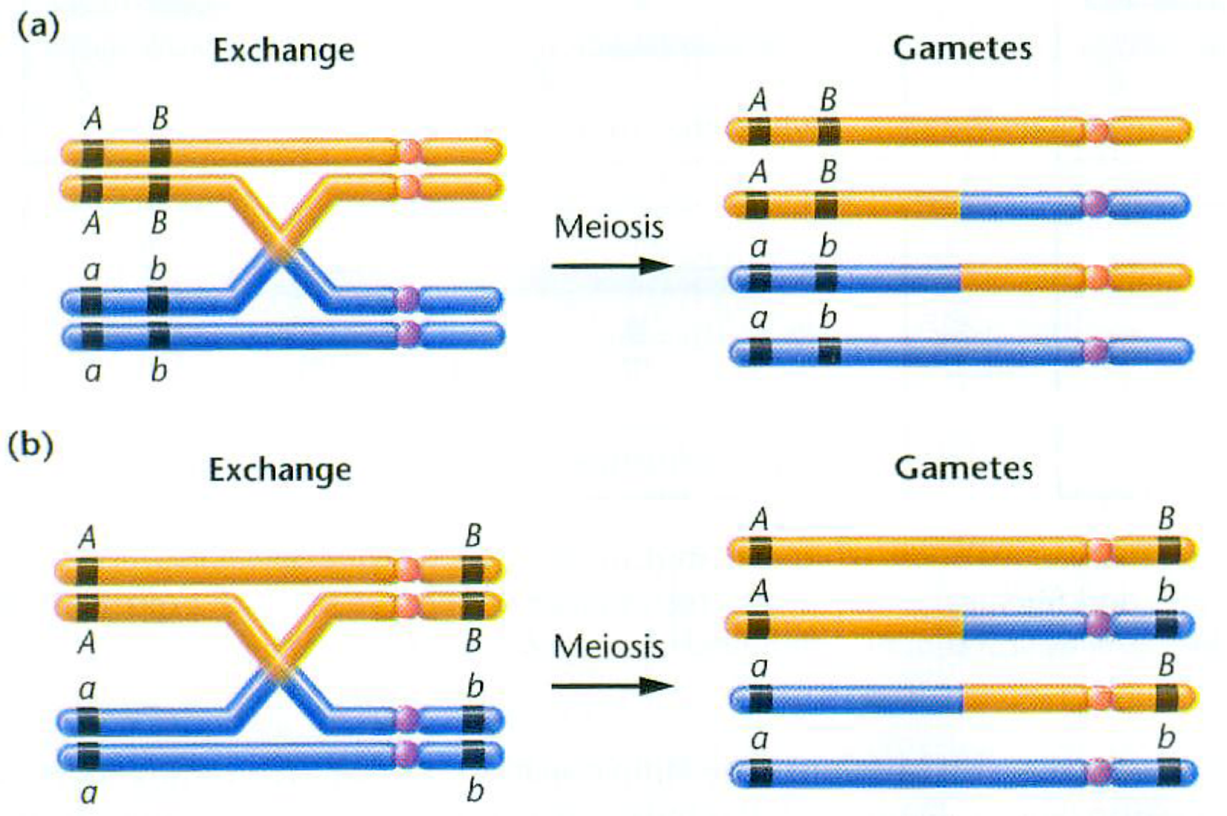


Dla 2 genów:
2 równoliczne klasy gamet rodzicielskich
2 równoliczne klasy gamet zrekombinowanych
Klasy zrekombinowane mniej liczne od rodzicielskich

W. S Klug, M.R Cummings “Concepts of Genetics” 8th edition, Prentice Hall, 2005

Mapowanie genów

Aby powstały gamety zrekombinowane, crossing-over musi zajść pomiędzy genami (loci)



powstają gamety zrekombinowane

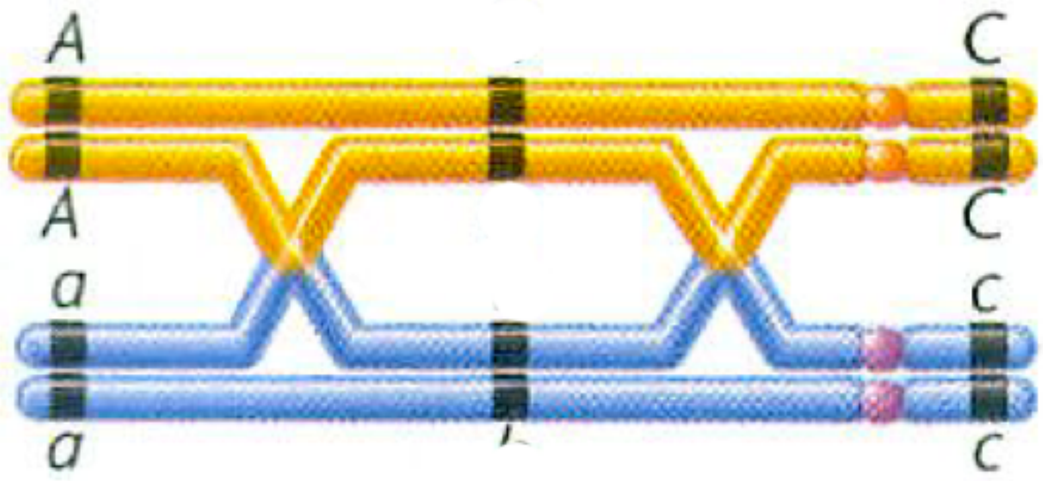
W. S Klug, M.R Cummings “Concepts of Genetics” 8th edition, Prentice Hall, 2005

Mapowanie genów

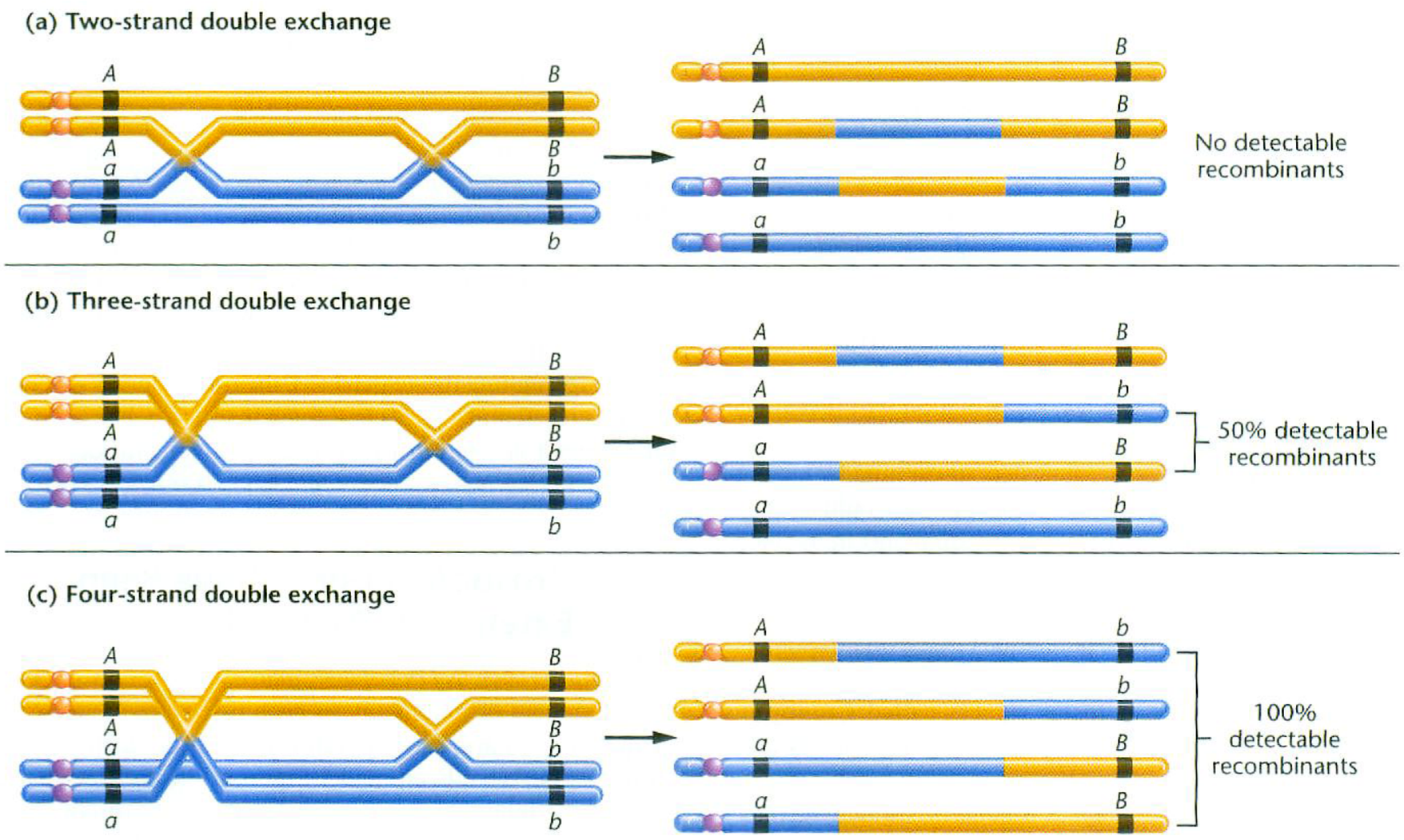
- Prawdopodobieństwo crossing-over pomiędzy genami jest proporcjonalne do odległości między nimi na chromosomie
- Liczebność klas rodzicielskich w potomstwie jest miarą odległości genetycznej

Mapowanie

- Jednostka cM (centymorgan) = 1% rekombinacji
- W rzeczywistości zależność nie jest liniowa
- Podwójny crossing-over – gamety typu rodzicielskiego
- Interferencja – zajście crossing-over w danym miejscu wpływa na prawdopodobieństwo zajścia kolejnego w pobliżu



Podwójny c-o – jeszcze bardziej złożony



W. S Klug, M.R Cummings “Concepts of Genetics” 8th edition, Prentice Hall, 2005

Funkcja mapowa

- Zależność odległości genetycznej od częstości rekombinacji
- Funkcja mapowa Haldane'a
  - wielokrotne c-o, bez interferencji

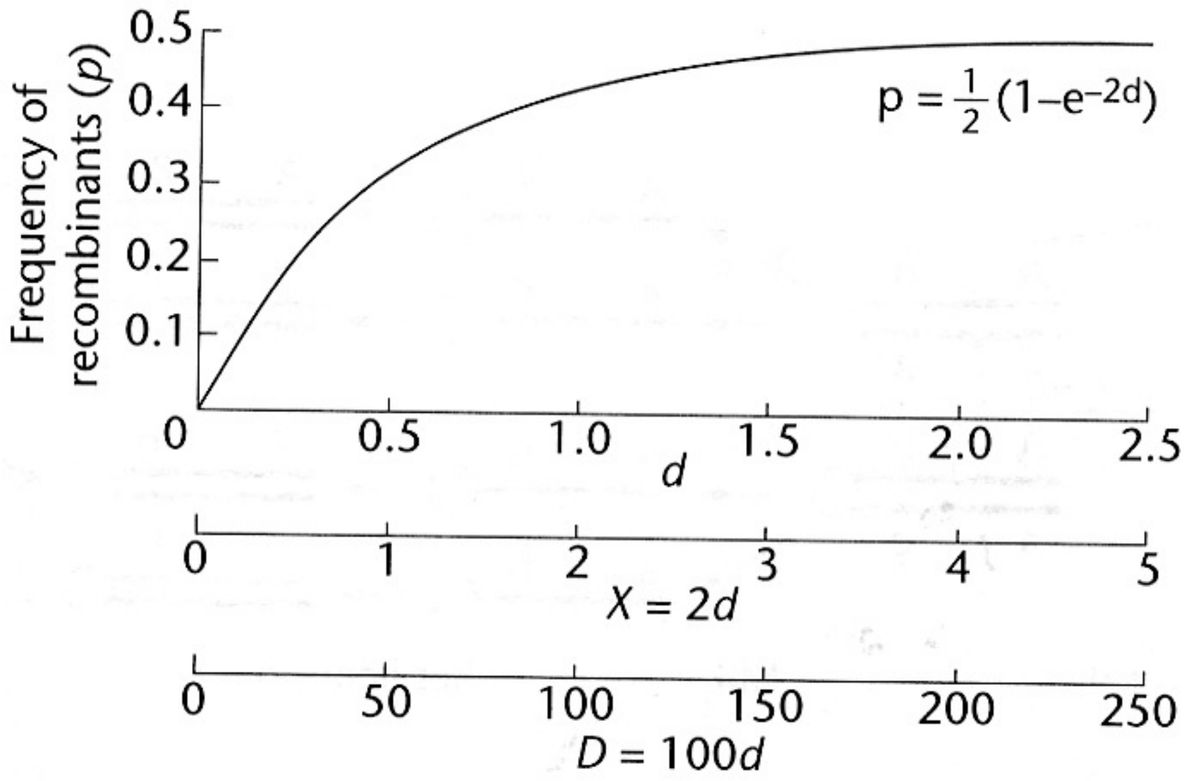
$$d = \frac{\ln(1-2\theta)}{2}$$

- Funkcja Kosambi'ego
  - uwzględnia też interferencję, szeroko stosowana
- Dla małych θ: d≈θ

$$d = \frac{\ln(\frac{1+2\theta}{1-2\theta})}{4}$$

Funkcja mapowa

- Wraz ze wzrostem odległości częstość obserwowanych c-o dąży do 0,5
- Dla genów niesprzężonych “rekombinantów” jest 50%, podobnie jak dla genów leżących w dostatecznie dużej odległości



Metody badawcze i strategie genetyki i genomiki

Lektura

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

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*
  - 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

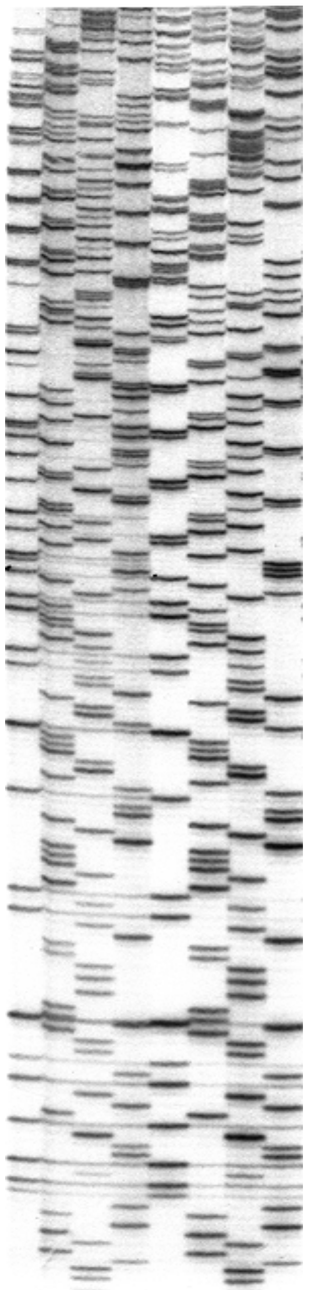
Granica między badaniami podstawowymi a stosowanymi jest płynna, stosowane techniki są podobne, różnice dotyczą głównie skali.

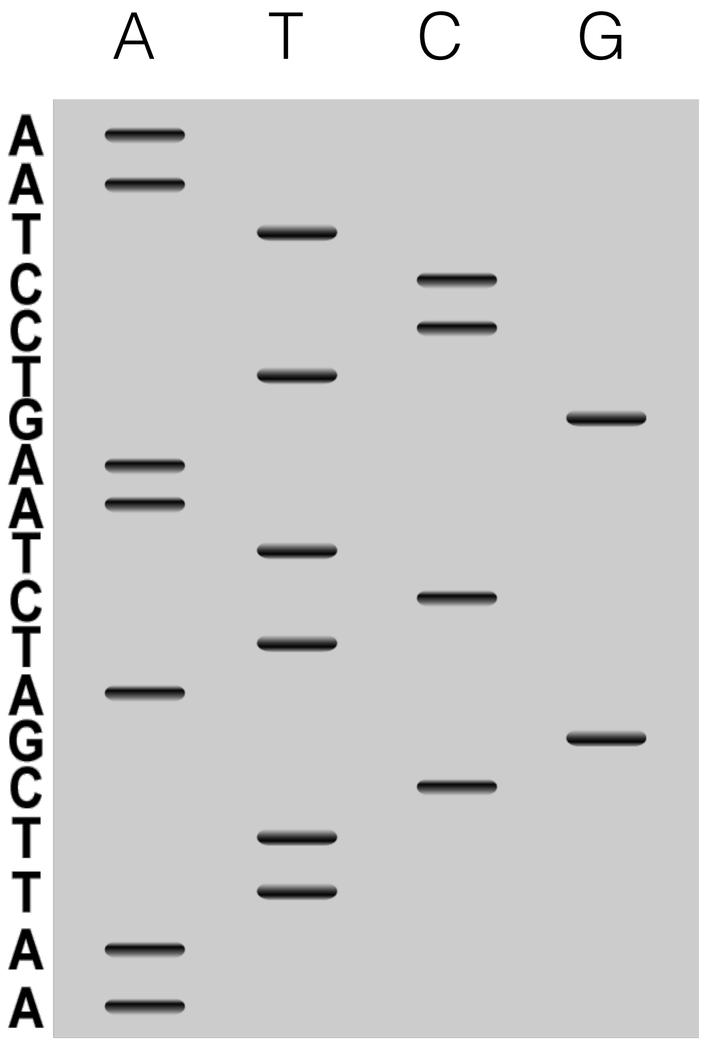
Podstawowe techniki

- 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

Tradycyjny odczyt sekwencji

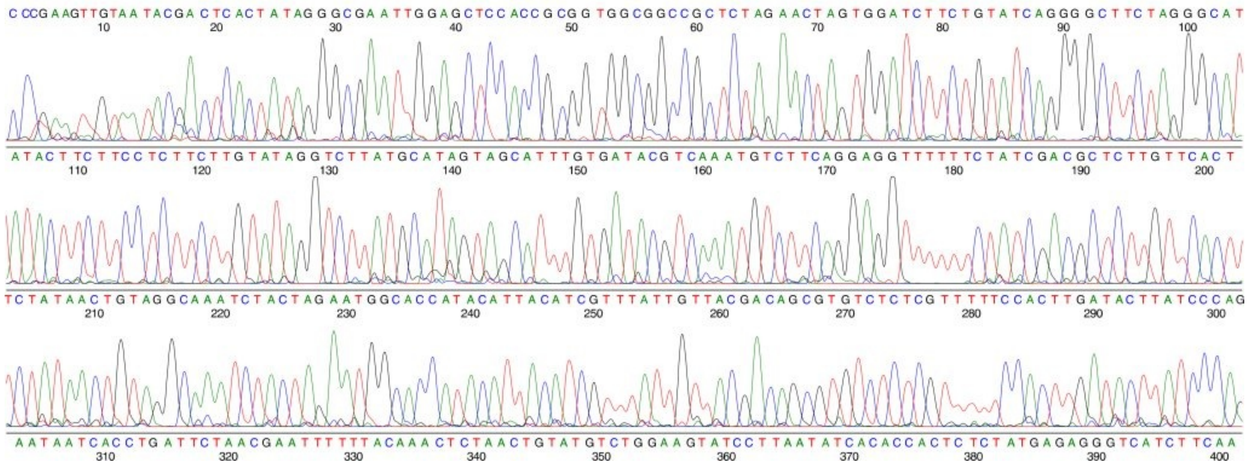
- Metoda Sangera
- Synteza DNA w obecności analogów nukleotydów (forma dideoksy) terminujących reakcję na określonym nukleotydzie
- Znakowanie radioaktywne, osobne reakcje





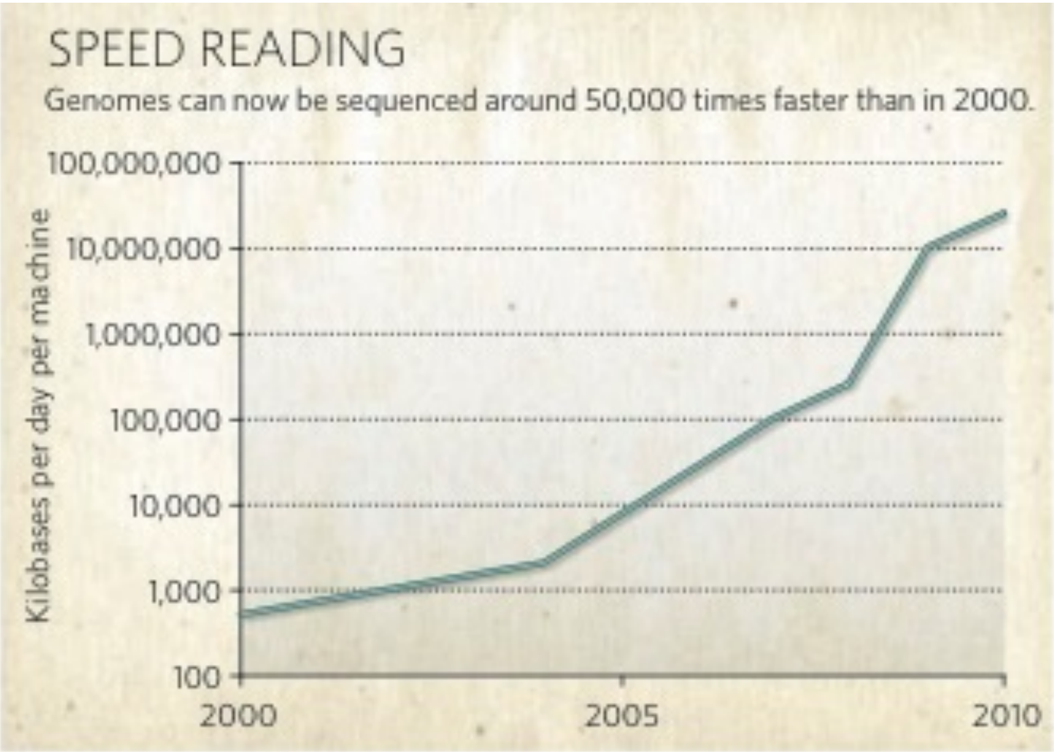
Sekwencjonowanie automatyczne

- 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

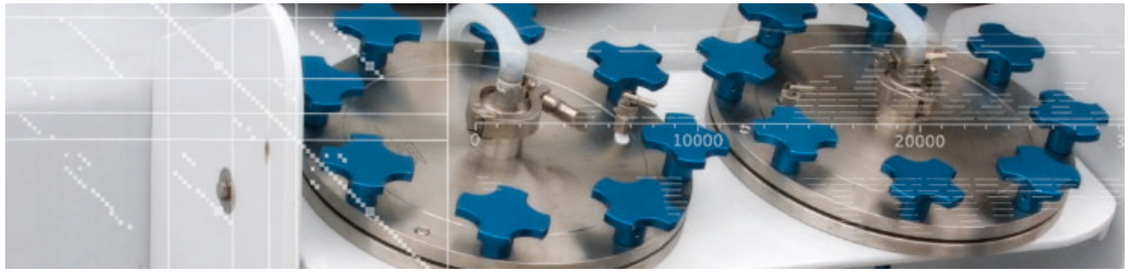
- Genomika jest dziedziną zajmującą się badaniem całych genomów (kompletu informacji genetycznej) różnych organizmów
- Techniki biologii molekularnej + robotyka + informatyka
- Sekwencjonowanie i charakteryzowanie genomów
- Badanie funkcji zawartych w nich genów - genomika funkcjonalna

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

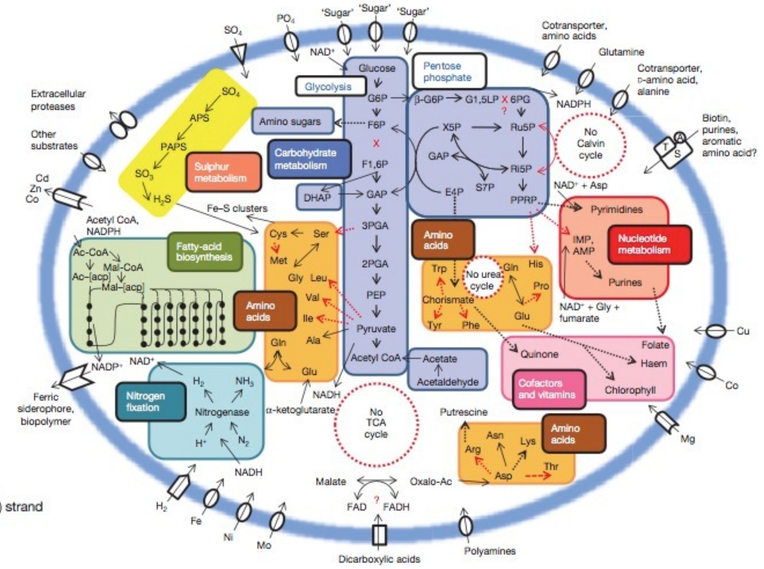
Candidatus Atelocyanobacterium thalassa

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]



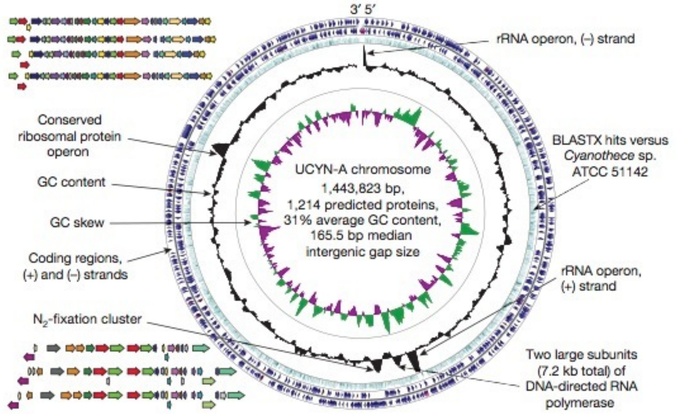
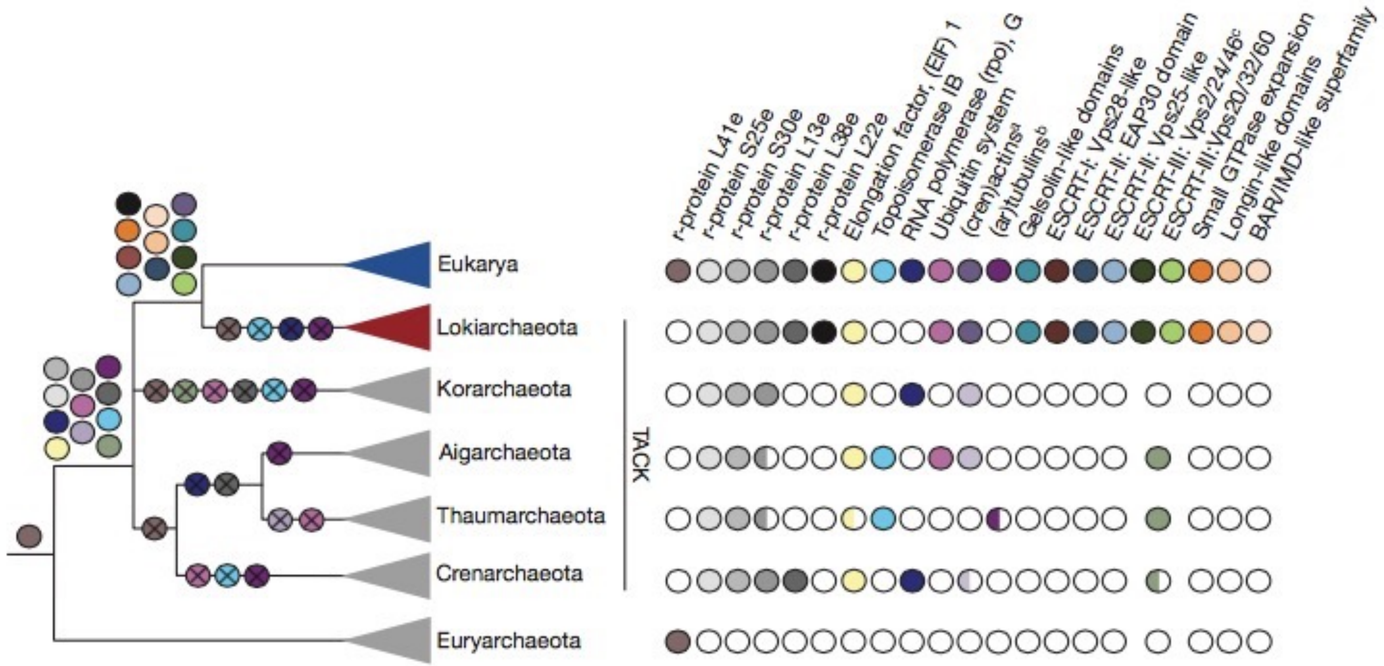


Figure 1 | Circular map of UCYN-A chromosome compared with *Cyanothece* sp. ATCC 51142. The outer two rings show the two strands of the

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

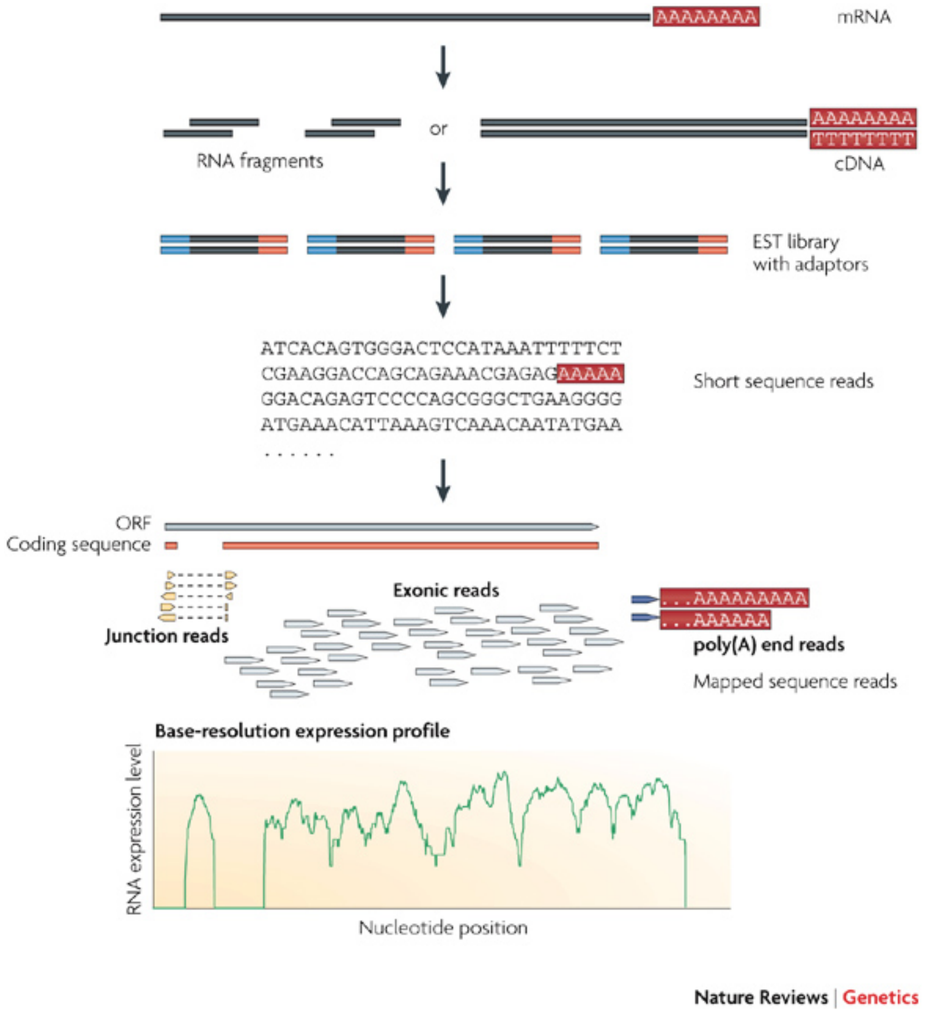
Anja Spang[1]*, Jimmy H. Saw[1]*, Steffen L. Jørgensen[2]*, Katarzyna Zaremba-Niedzwiedzka[1]*, Joran Martijn[1], Anders E. Lind[1], Roel van Eijk[1]†, Christa Schleper[2,3], I

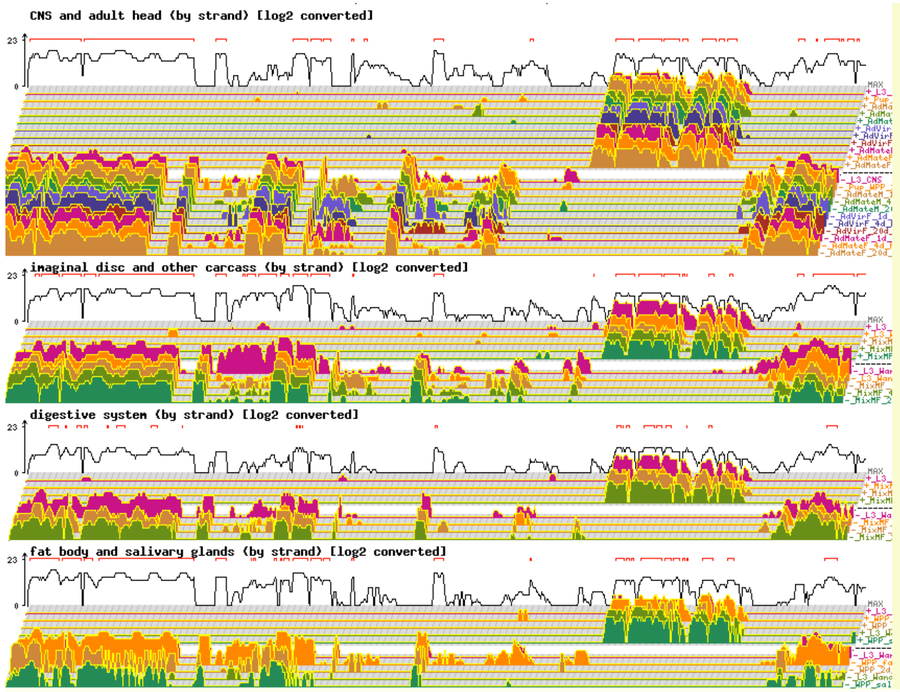
©2015 Macmillan Publishers Limited. All rights reserved

Wielkie projekty

- Projekt 1000 genomów - różnorodność genetyczna człowieka
  - obecnie dane całogenomowe dla dziesiątków tysięcy osób (np. UK BioBank)
- Metagenomika mikrobiomu przewodu pokarmowego człowieka
- Genomy wymarłych gatunków (np. Neandertalczyk)
- Sekwencje genomów wszystkich gatunków organizmów żywych

RNA-seq - badanie ekspresji genów





Sekwencjonowanie nowej generacji – wyzwanie dla bioinformatyki

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

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

Czy warto badać genomy

- Nowoczesne techniki generują bardzo dużo danych
- Dwa podejścia
  - “hypothesis driven” – dane zbierane dla zweryfikowania jakiejś hipotezy
  - “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 […]”