



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Organizacja genomu człowieka i sekwencjonowanie DNA

Materiały dydaktyczne współfinansowane ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego.

Organizacja genów na ludzkim chromosomie. Najkrótszy autosom - chromosom 22 zawierający ok. 1,5% ludzkiego genomu. Największą niespodzianką dla badaczy po zsekwencjonowaniu dużych porcji ludzkiego genomu było stwierdzenie, że tylko niewielki ułamek koduje użyteczne cząsteczki – pozostałą część nazwana „**junk DNA**” lub „**selfish DNA**”.

Ludzki genom w liczbach

Cecha	Liczba
Długość DNA	$3,2 \times 10^9$ par zasad
Liczba genów	ok. 25 000
Najdłuższy gen	$2,4 \times 10^6$ par zasad
Średnia długość genu	27 000 par zasad
Najmniejsza liczba eksonów na gen	1
Największa liczba eksonów na gen	178
Średnia liczba eksonów na gen	10,4
Rozmiar najdłuższego eksonu	17 106 par zasad
Średnia długość eksonu	145 par zasad
Liczba pseudogenów	ponad 20 000
Procent DNA zawartego w eksonach	1,5 %
Procent DNA zawartego w innych konserwowanych sekwencjach, np. odcinki regulatorowe, o. kodujące RNA.	3,5 %
Procent DNA zawartego w elementach powtórzonych w dużej liczbie kopii.	50 %

Na podstawie: *Molecular Biology of the Cell* (© Garland Science 2008)

Ludzki genom mitochondrialny

Małe (16.5 kb) koliste DNA

Geny kodują białka oraz rRNA, tRNA (37)

Mitochondria mają własne rybosomy przypominające rybosomy bakterii.

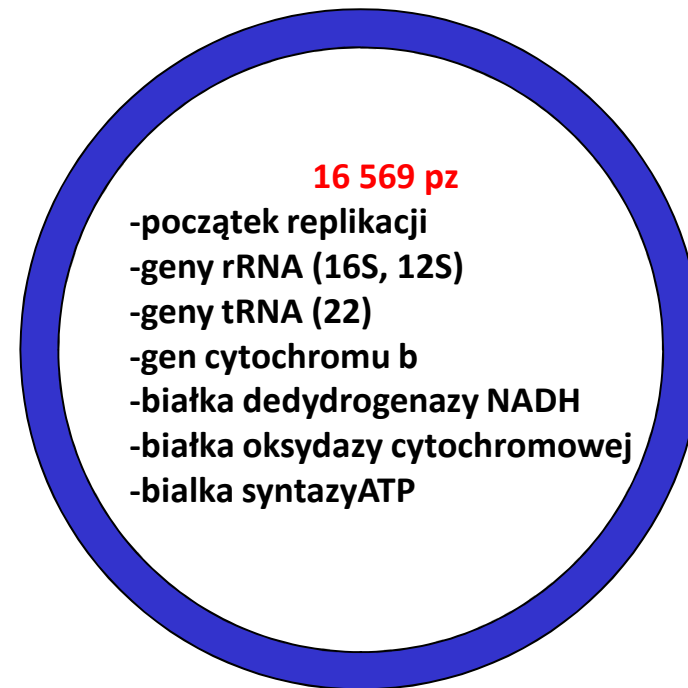
1 gen/0.45 kb

Niewiele powtórzeń

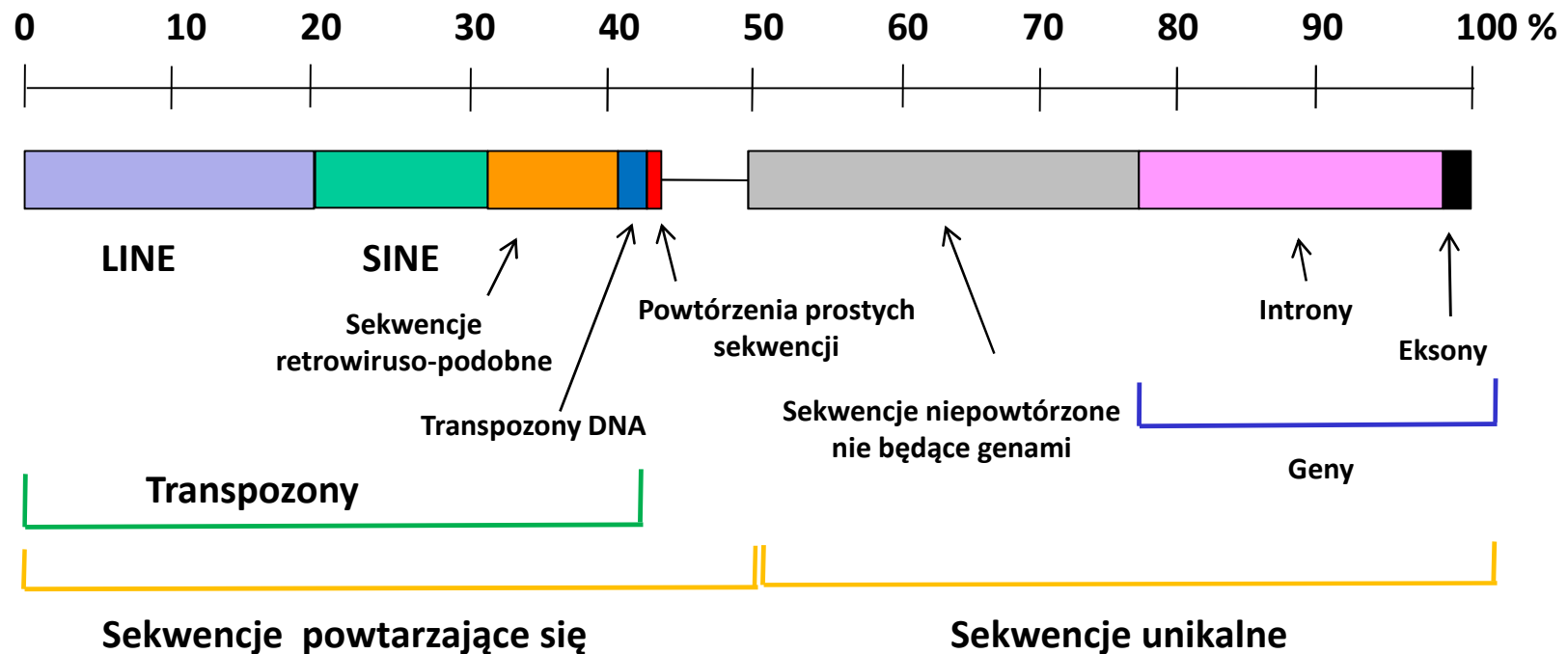
Brak intronów

93% to część kodująca;

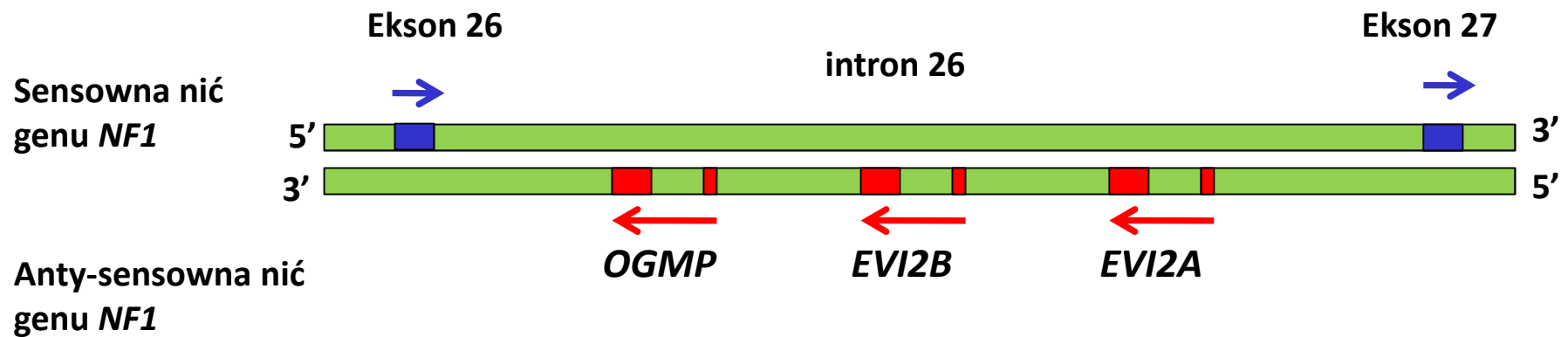
Dziedziczone tylko po matce.



Typy sekwencji (w procentach) występujących w ludzkim genomie. Ponad 40% genomu stanowią tzw. sekwencje mobilne (transpozony). Połowa ludzkiego genomu to sekwencje powtarzające się. Tylko niewielki ułamek genomu to sekwencje kodujące łańcuchy białek.



Geny w genach. Introny niektórych genów są tak obszerne, że mogą pomieścić wiele innych genów z własnymi intronami. Stwarza to konieczność bardzo precyzyjnej regulacji ekspresji genów (transkrypcji, wycinania intronów).



Gen neurofibromatozy (NF1) zawiera intron 26 kodujący trzy inne geny: *OGMP*, *EVI2A* i *EVI2B* zawierające po dwa eksony.

Dlaczego człowiek – „szczyt drabiny stworzenia” ma tak mało genów?

Człowiek	25 000 genów
Muszka owocowa	13 000
Robak obły	19 000

1. Genom – suma wszystkich genów oraz całość informacji genetycznej organizmu.
2. Transkryptom – suma cząsteczek RNA organizmu (komórki).
3. Proteom – suma wszystkich białek organizmu (komórki).

Potencjalna różnorodność genomu i transkryptomu jest tak duża, że najwyraźniej nie ma potrzeby zwiększania liczby genów

Powtarzające się sekwencje w genomie

- Powtórzenia rozproszone (*interspersed repeats*)
 - Powtórzenia tandemowe (*tandem repeats*)
 - Mikrosatelity
 - Minisatelity
 - Satelity
 - Powtórzenia tandemowe genów (g. rRNA, g.histonów)
-

Znaczenie powtórzeń rozproszonych.

- Mogę generować podczas transpozycji chorobotwórcze mutacje.
 - Przyspieszają tworzenie się bioróżnorodności (znaczenie ewolucyjne).
 - Inne nieznane funkcje ???
 - „Selfish DNA” ?
-

Trzy główne klasy elementów ruchliwych (transpozonów).

Opis klasy i budowa.	Enzym konieczny do przenoszenia	Sposób przenoszenia
Transpozony DNA  Krótkie, odwrócone Powtórzenia na końcu	Transpozaza	Porusza się wyłącznie w formie DNA
Retrotranspozony retrowirusopodobne  Flankowane sekwencjami LTR	Odwrotna transkryptaza i integraza	Porusza się za pośrednictwem RNA powstałego w transkrypcji z promotora w obrębie LTR
Retrotranspozony nie-retrowirusopodobne  Na końcu występuje ogon poliA	Odwrotna transkryptaza i iendonukleaza	Porusza się za pośrednictwem RNA powstałego w transkrypcji z promotora sąsiadującego genu

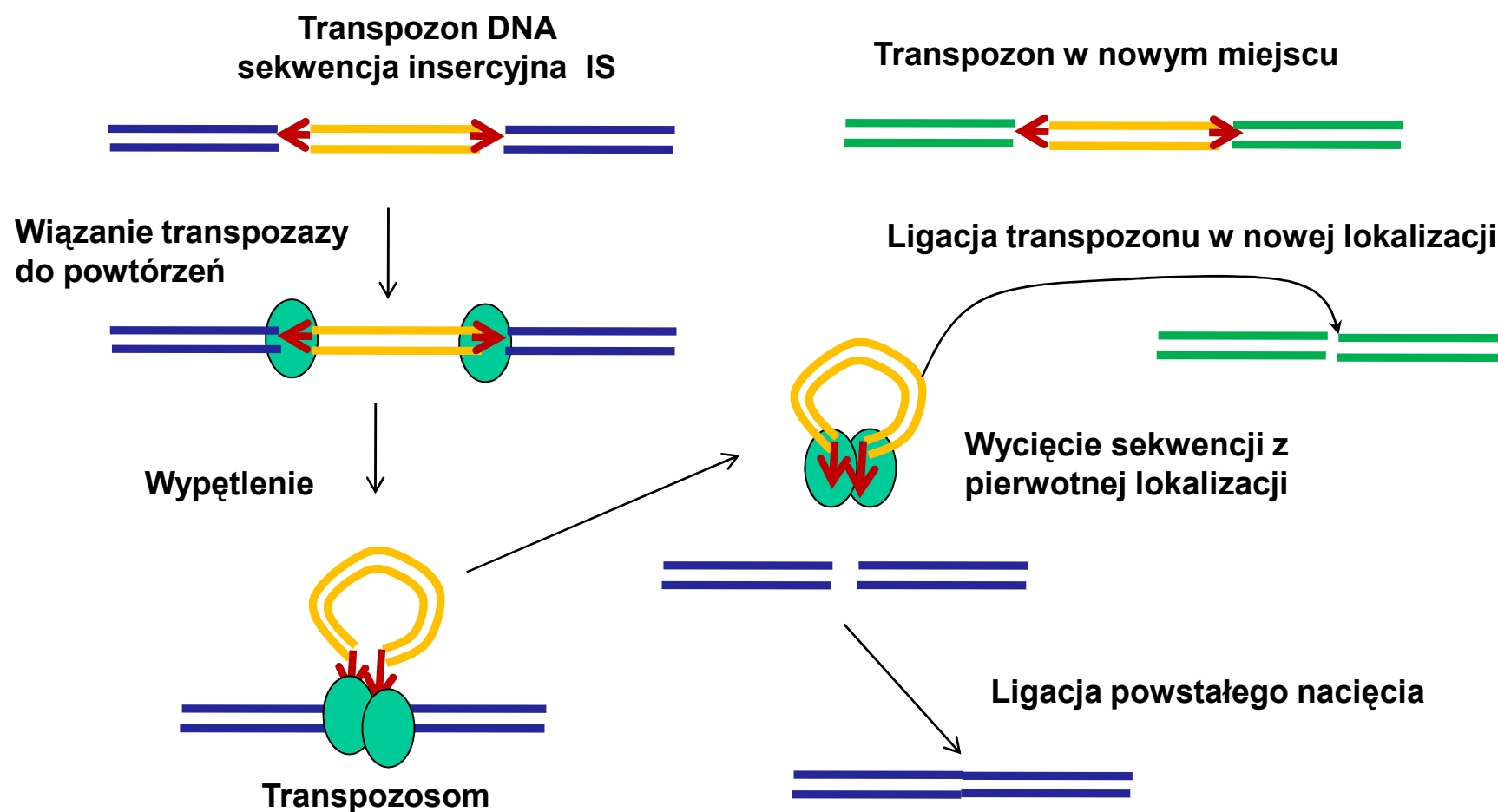
Duże powtórzenia: Transpozony

- Elementy ruchliwe, transpozony ang.
“Transposable elements” (TE’s) – sekwencje zdolne do przemieszczania się w obrębie genomu.

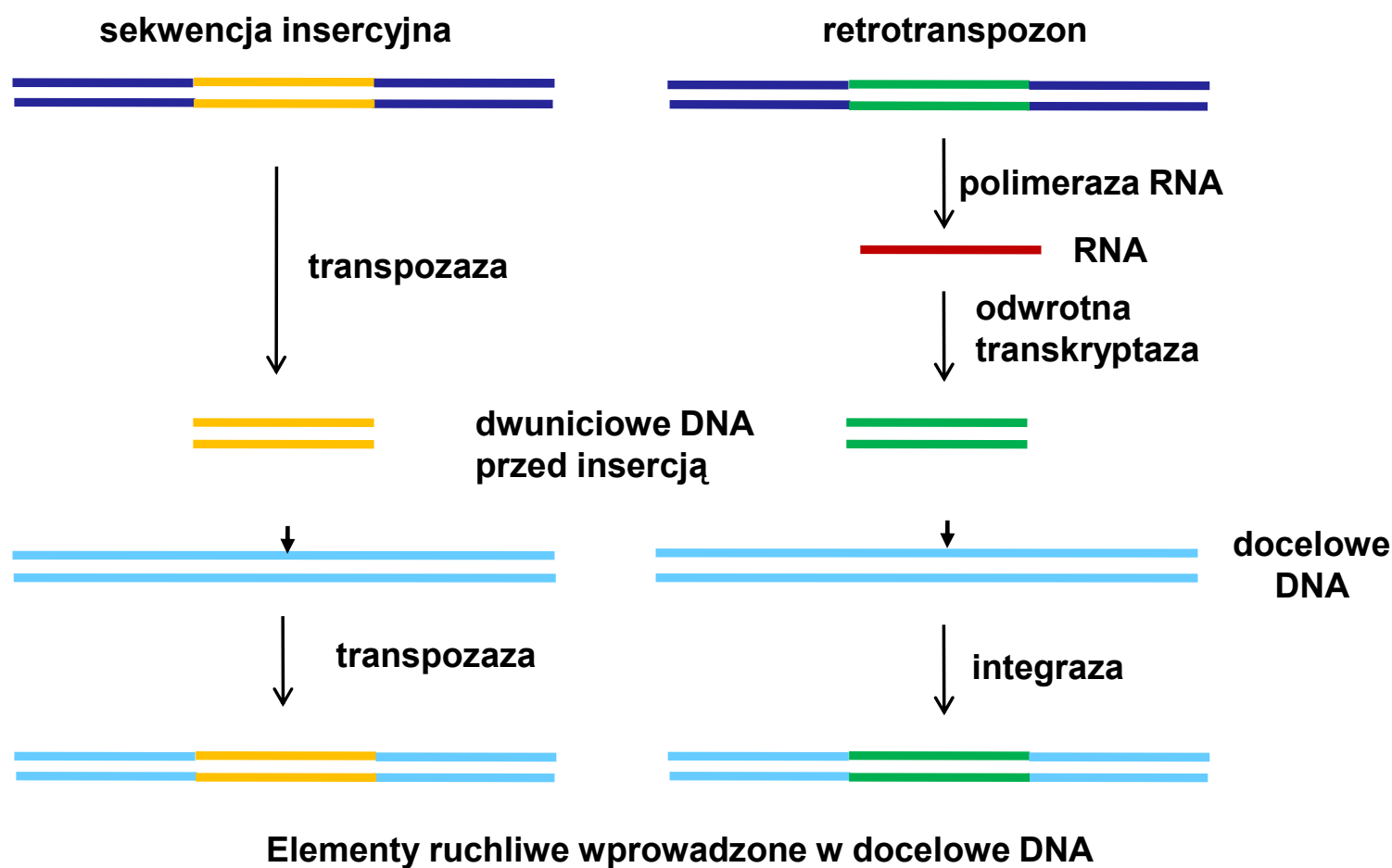
Dlaczego transpozony są istotne?

Zapamiętaj: ok. 50% genomu człowieka zawdzięcza swoje pochodzenie transpozonom.

Mechanizm transpozycji typu „wytnij-wklej”. Tak poruszają się transpozony bez pośrednictwa RNA. Przemieszczanie się tych elementów odbywa się za pośrednictwem kompleksu enzymatycznego zwanego transpozazą. Sekwencja transpozonu jest oflankowana krótkimi, odwróconymi powtórzeniami. W miejscu insercji powstają krótkie powtórzenia proste.



„**Retroelementy**” to transpozony przemieszczające się za pośrednictwem cząsteczki RNA i wymagające do przemieszczania enzymu zwanego odwrotną transkryptazą oraz integrazą. Podczas przemieszczania się retroelementu następuje duplikacja sekwencji ruchliwej – przemieszczanie odbywa się metodą „kopiuj-wklej”.



Rodzaje retroelementów

Retroelementy zawierające sekwencję LTR
ang. *long terminal repeat* (LTR)

- retrotranspozony

Retroelementy pozbawione sekwencji LTR

- LINES
 - SINES
 - Processed pseudogenes
-

SINES: short interspersed nuclear elements

SINES; np. sekwencja Alu

Nie potrafią poruszać się autonomicznie

Nie koduje odwrotnej transkryptazy

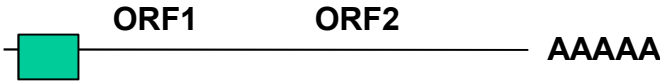
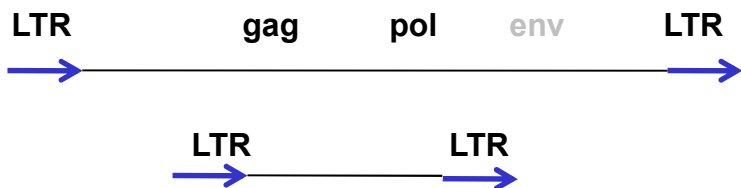
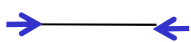
Transkrybowany przez RNA Pol III

W ludzkim genomie występuje ok. 1 milion kopii

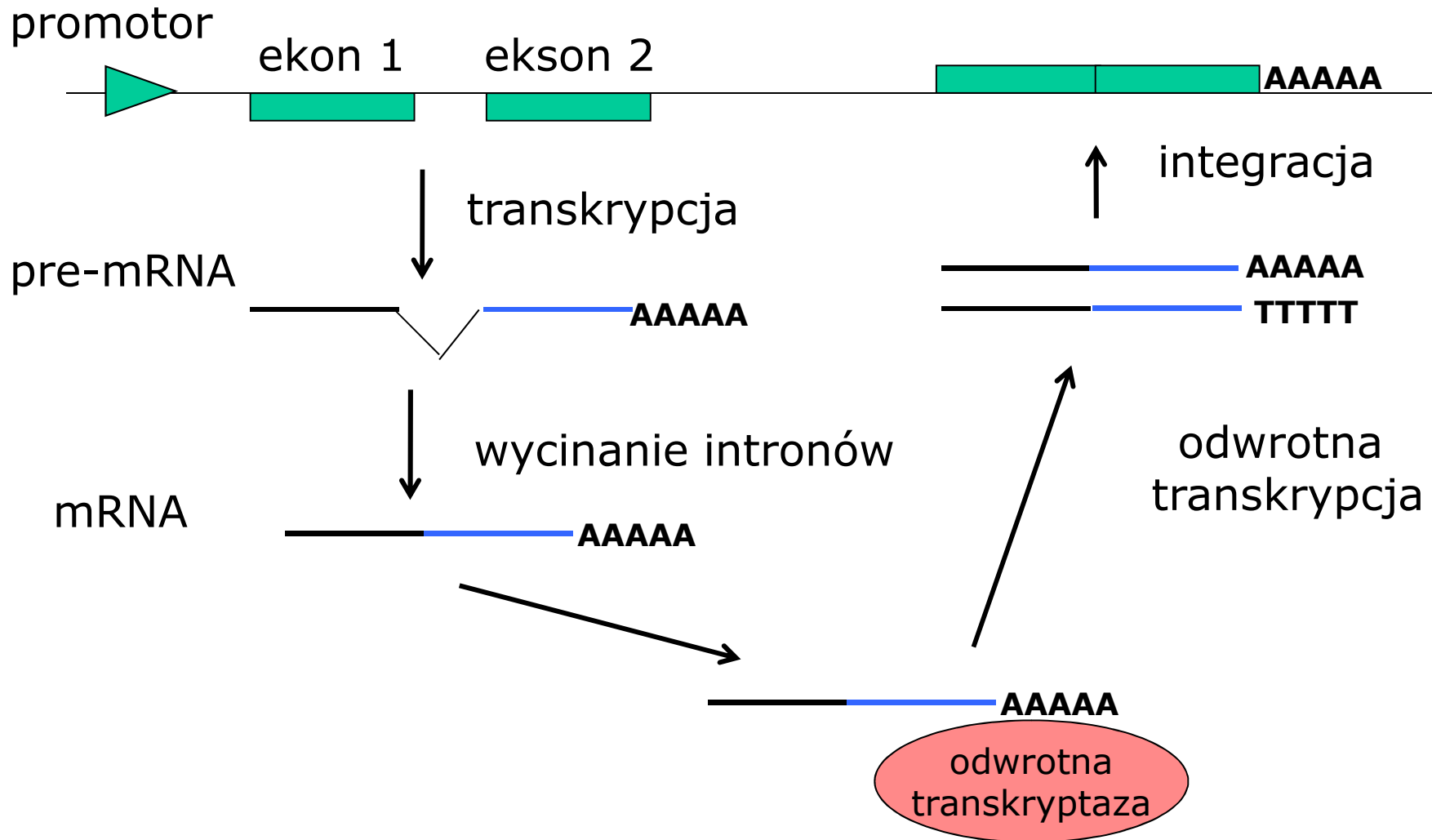


Element długości 150 pz, zawiera bogaty w A łącznik (linker).
Sekwencja zawiera miejsce rozpoznawane przez enzym *AluI*.

„Retroelementy” i transpozony występujące w genomie człowieka.

		Przeciętna długość	% genomu
LINE		8 000 pz	20 %
SINE		250 pz	13 % 1,5 mln kopii
Transpozony Retrowiruso- podobne		10 000 pz	9 %
Pozostałości po transpozonach DNA		3 000 pz	3 %
		70 - 3 000 pz	

Tworzenie się pseudogenów.



Sekwencje minisatelitarne

- 6-20 bp - długość powtarzającej się jednostki

TTGCAATGCGTCGTGCACCTATGCCAGGATTTT
TTAGCTTTGAGGATTTTAGAGGATATTAGGCAGC
AGGATTTTAAGGATATTAGGCAGC TAGTATTATG

Konsensus AGGATTTT

Sekwencje minisatelitarne

- Wysoce polimorficzne, pomocne w identyfikacji osobniczej na podstawie śladów DNA ang. „DNA fingerprinting”

Prawdopodobnie uczestniczą w regulacji ekspresji genów.

Sekwencje mikrosatelitarne

- 1-5bp - długość powtarzającej się jednostki
- W puli genowej populacji człowieka występują wersje mogące różnić się liczbą powtórzeń.

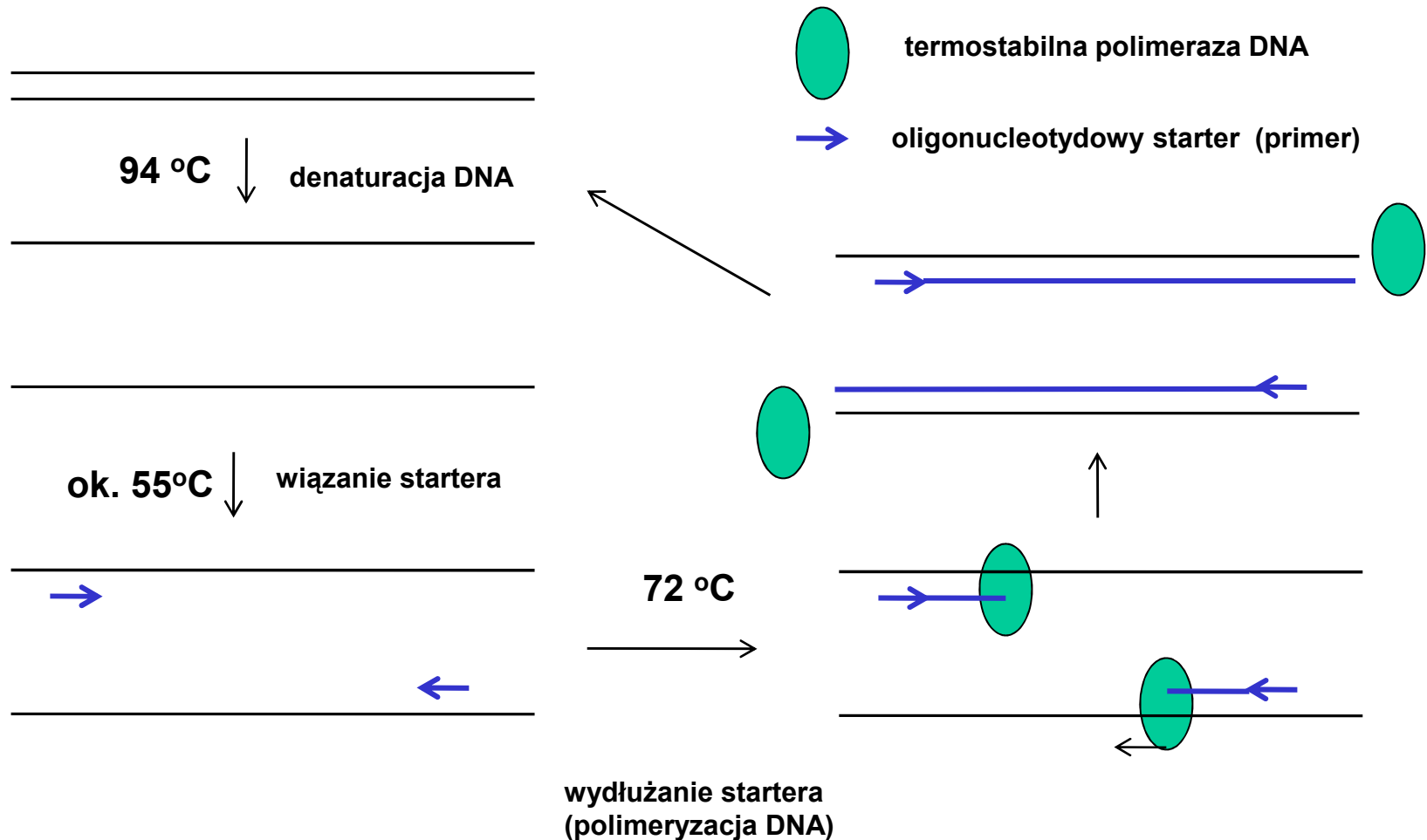
actgtagtcctaaag ctctctctctctctctctctctctctctct agcctaagggtcaagcttt

tctgccgtcctcccg gtggtggtggtggtggtggtggtggtggtggtg attcctaaaagtctggct

Sekwencje mikrosatelitarne

- Liczba powtórzeń wykazuje międzyosobnicze różnice.
 - Gwałtowna zmiana liczby kopii jest przyczyną wielu chorób:
 - Fragile X syndrome, Huntington's disease, Myotonic dystrophy
 - Ze względu na dużą zmienność są powszechnie wykorzystywane w sądownictwie i kryminalistyce np. do testów na ojcostwo.
-

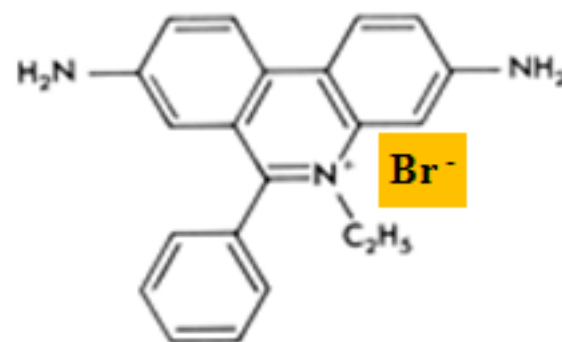
Główne etapy polimerazowej reakcji łańcuchowej (PCR ang. polymerase chain reaction).



Reakcja PCR zachodzi w urządzeniach zwanych termocyklerami, które są blokami grzejąco-chłodzącymi sterowanymi komputerowo. Grzejące pokrywy uniemożliwiają odparowywanie wody z roztworu reagentów.

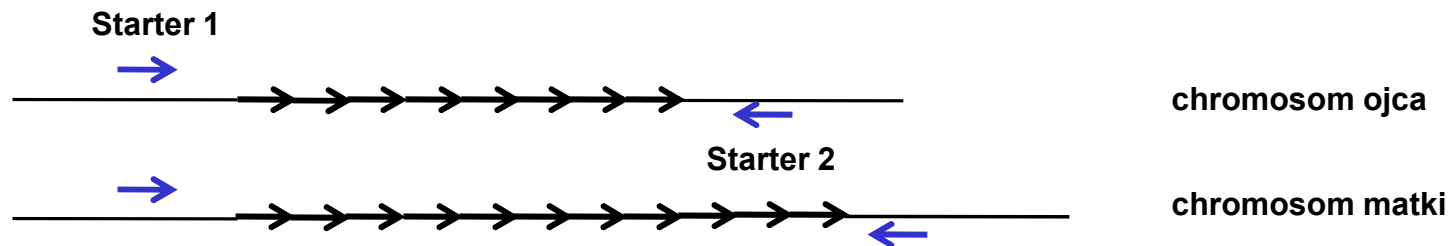


Produkty PCR są zwykle rozdzielane w celu uwidocznienia przy pomocy elektroforezy w żelu agarozowym. DNA jest wybarwiany interkalującą substancją – bromkiem etydyny.

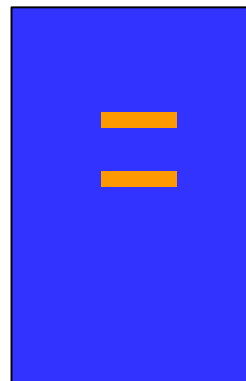


Wykorzystanie PCR do amplifikacji sekwencji mikrosatelitarnych i identyfikacji osobniczej.

Starter 1 →
actgtagtcctaaag ctctctctctctctctctctctctctctct agcctaaggggtcaagcttt
← Starter 2



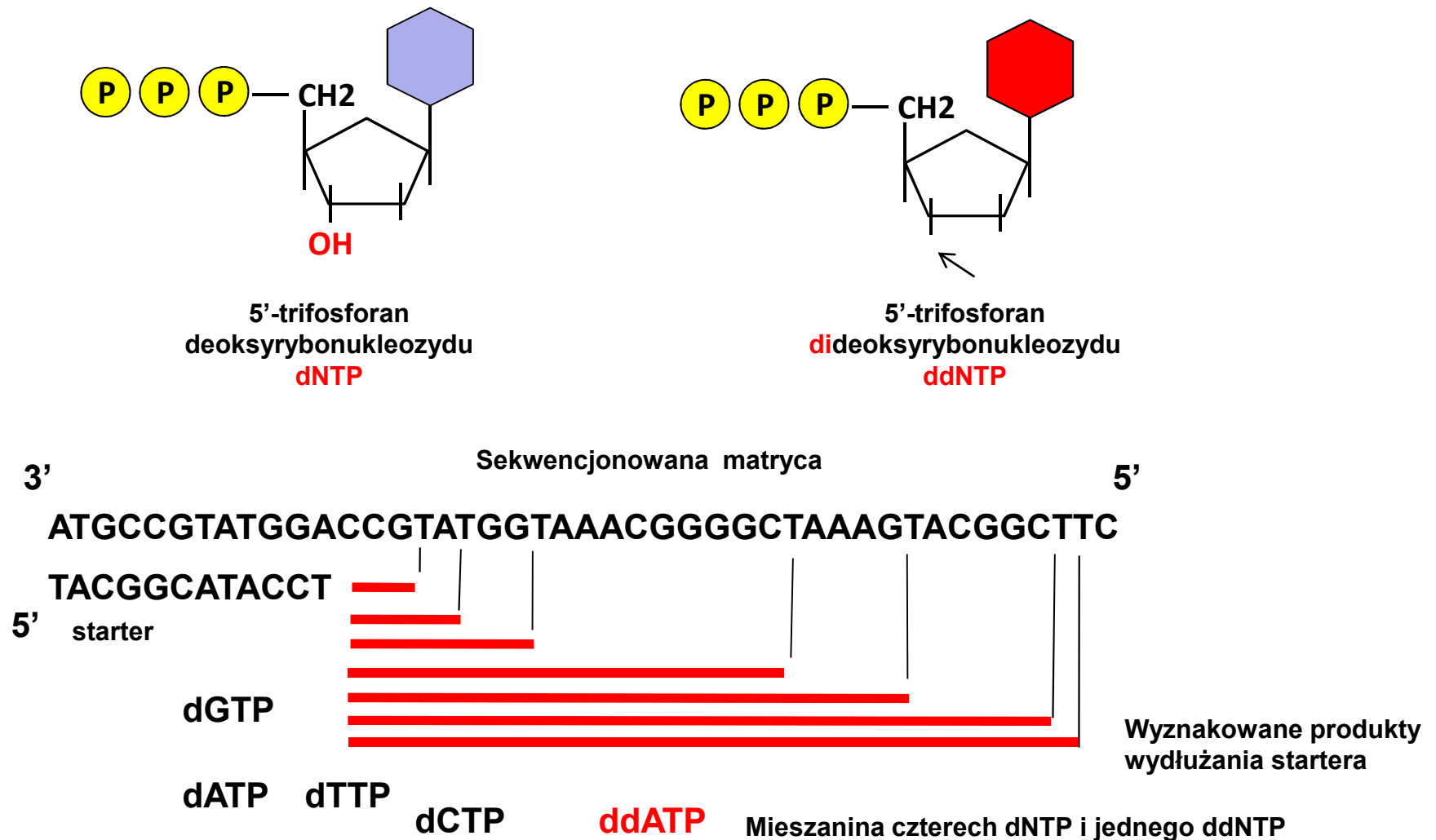
↓ PCR i rozdział produktów na żelu



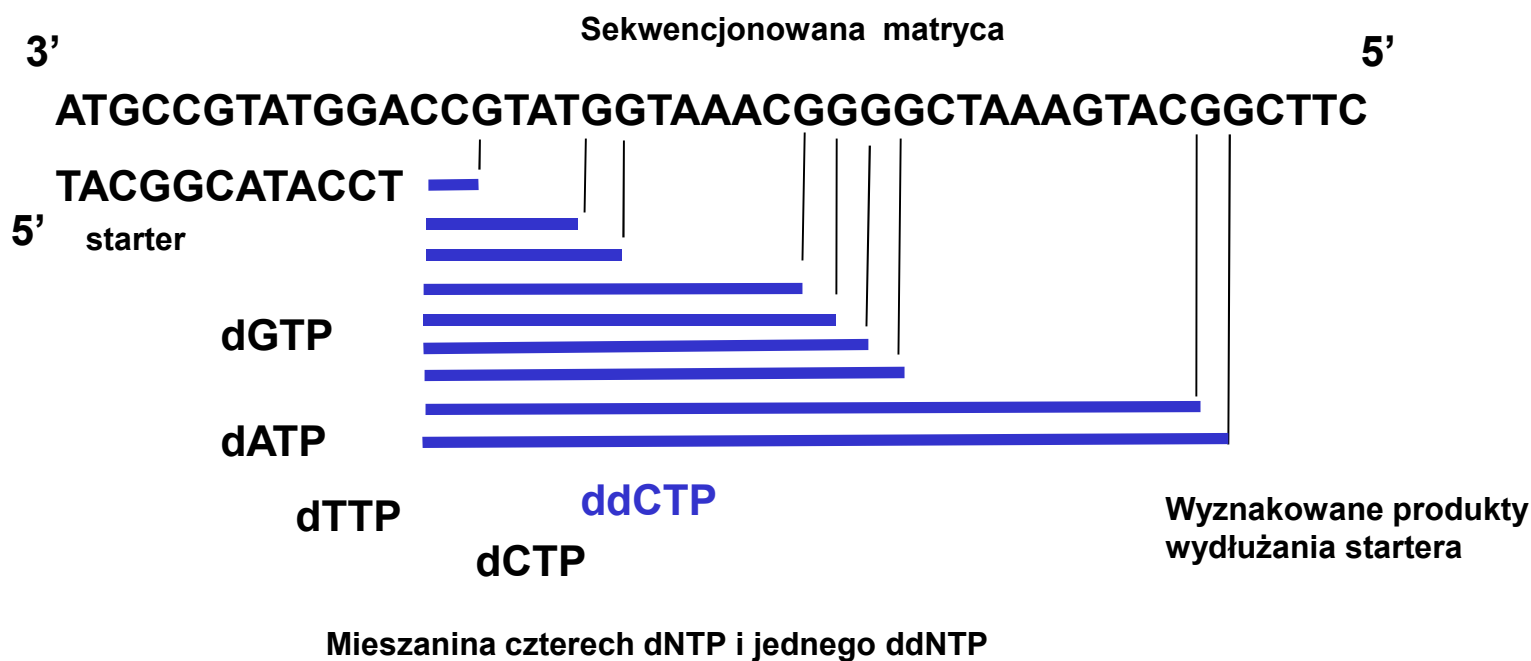
Fragment z 11 powtórzeń

Fragment z 8 powtórzeń

Zasada sekwencjonowania DNA metodą dideoksy. W reakcji polimeryzacji DNA in vitro biorą udział deoksynukleotydy i dideoksynukleotydy. Wbudowanie dideoksynukleotydu powoduje zakończenie syntezy łańcucha DNA. Metodę opracował Fred Sanger w drugiej połowie lat 70-tych ubiegłego wieku.



Reakcję wydłużania startera prowadzi się w osobnych probówkach, które zawierają identyczną matrycę starter i polimerazę i cztery dNTP ale w każdej znajduje się inny dideoksynukleotyd. Następnie produkty polimeryzacji rozdziela się w osobnych ścieżkach żelu poliakryloamidowego. Z powstałego obrazu produktów odczytuje się sekwencję DNA.



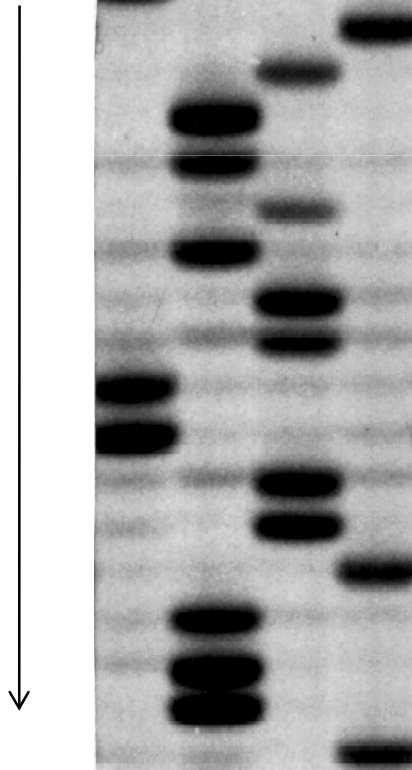
Odczyt sekwencji na podstawie fragmentów DNA znakowanych radioaktywnie.

Rozdział produktów reakcji w probówkach z:

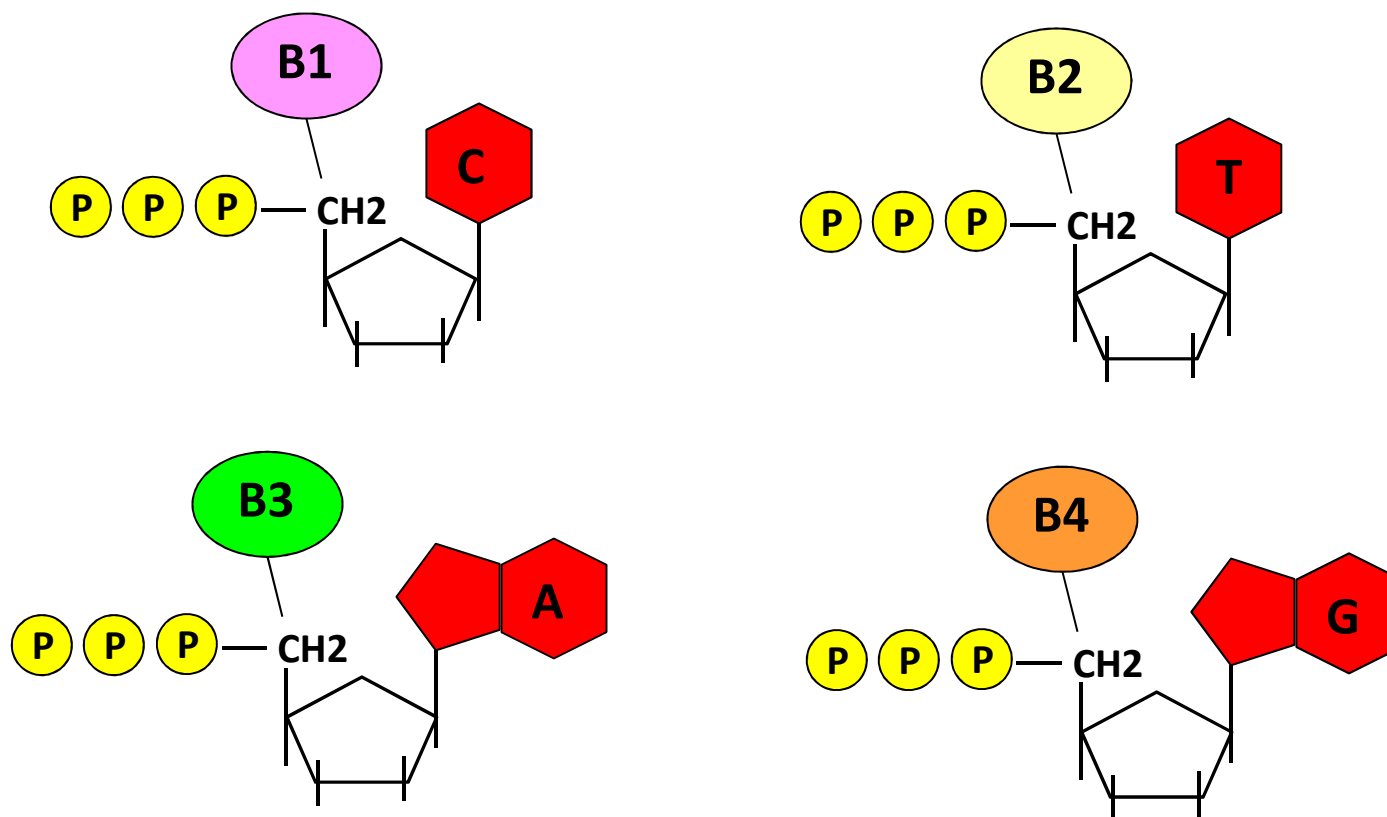
ddCTP ddTP
ddATP ddGTP

A C G T

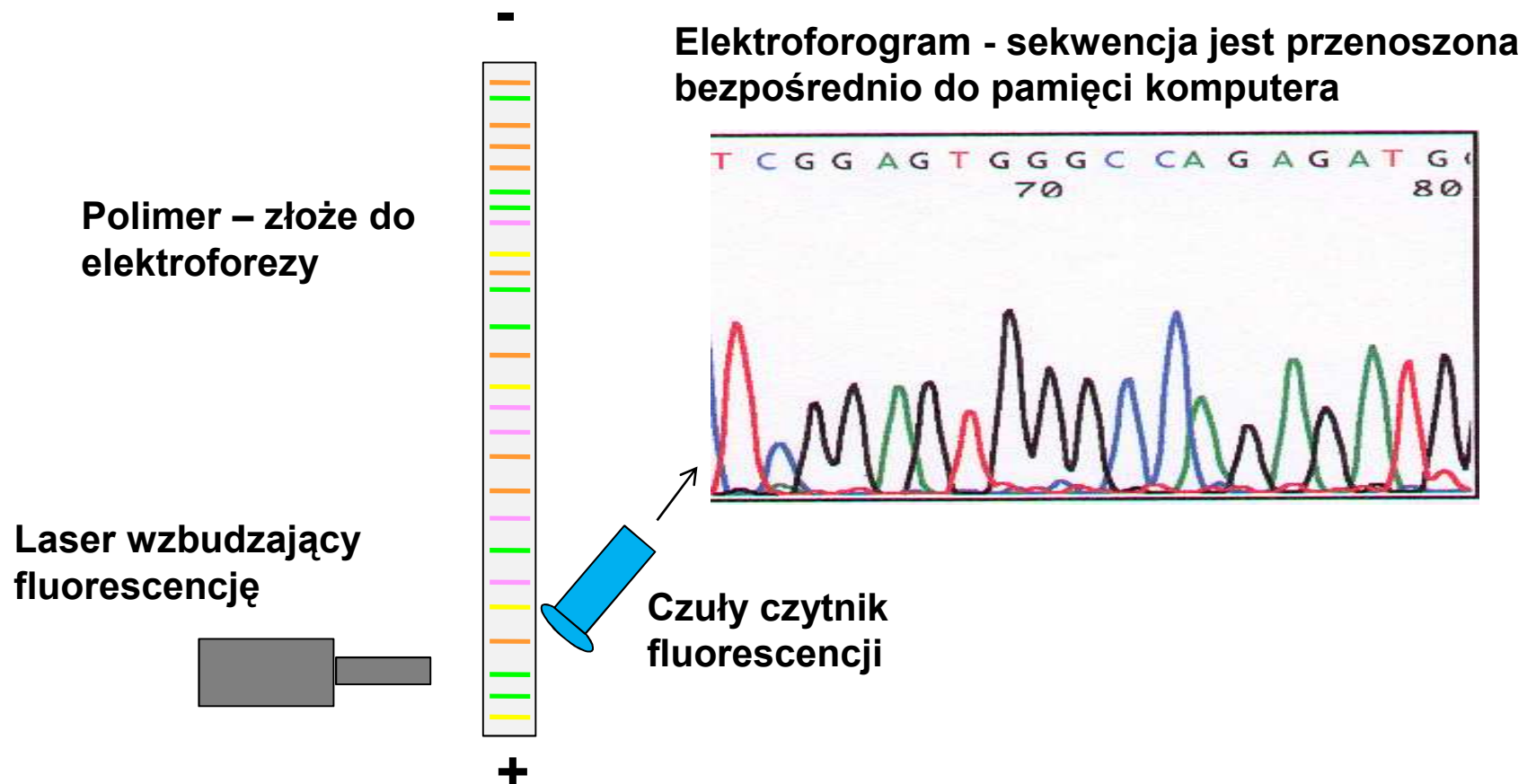
Kierunek rozdziálu



Sekwencjonowanie z użyciem znaczników radioaktywnych było kłopotliwe ze względu na stosunkowo krótkie odczyty sekwencji z jednego rozdzału (maks. 300 bp), konieczność rozdzału 1 próbki na czterech ścieżkach i długi czas oczekiwania na ekspozycję autoradiograficzną. W latach 80-tych powstały automatyczne sekwenatory DNA wykorzystujące barwniki fluorescencyjne do znakowania DNA i czułe czytniki do detekcji wyznakowanych fragmentów DNA. Każdy dideksynukleotyd ma dołączony inny barwnik fluorescencyjny.



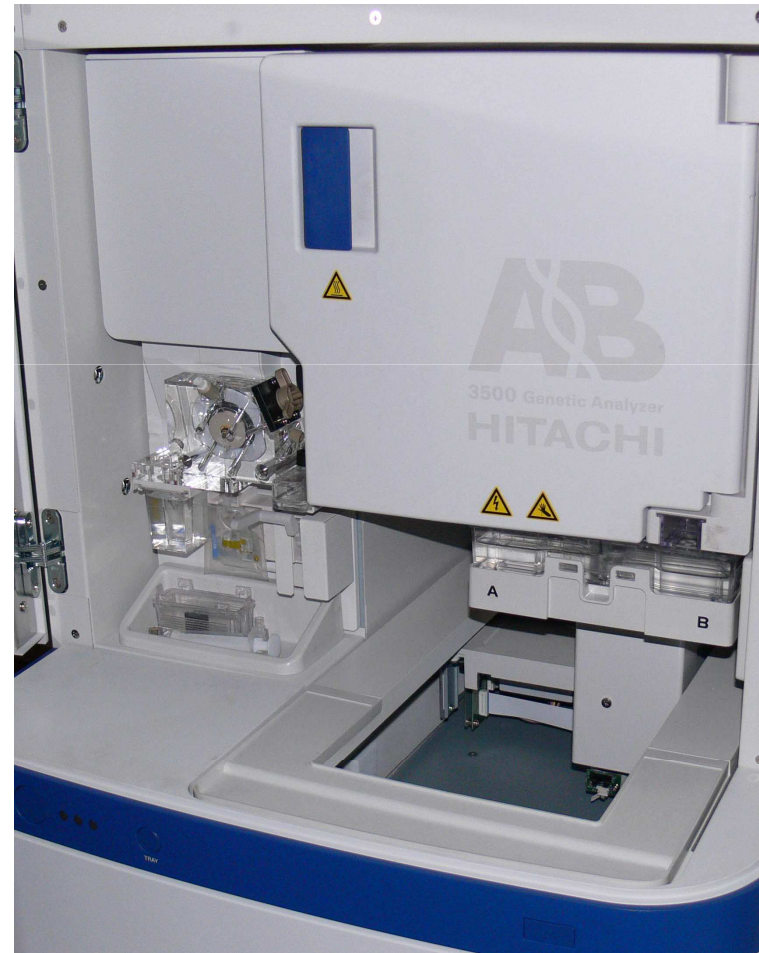
Zastosowanie różnych barwników fluorescencyjnych dołączonych do dideoksynukleotydów umożliwiło przeprowadzenie reakcji i rozdziłu dla jednej próbki w jednej ścieżce żelu. Ponadto, odczyt odbywa się w czasie rzeczywistym i dzięki temu próbki mogą być rozdzielane na całej długości żelu, co umożliwia odczyt nawet 900 nukleotydów.



Pierwsza generacja sekwenatorów automatycznych wykorzystywała elektroforezę płytową. Nowsze generacje korzystają z elektroforezy kapilarnej, którą można zautomatyzować.

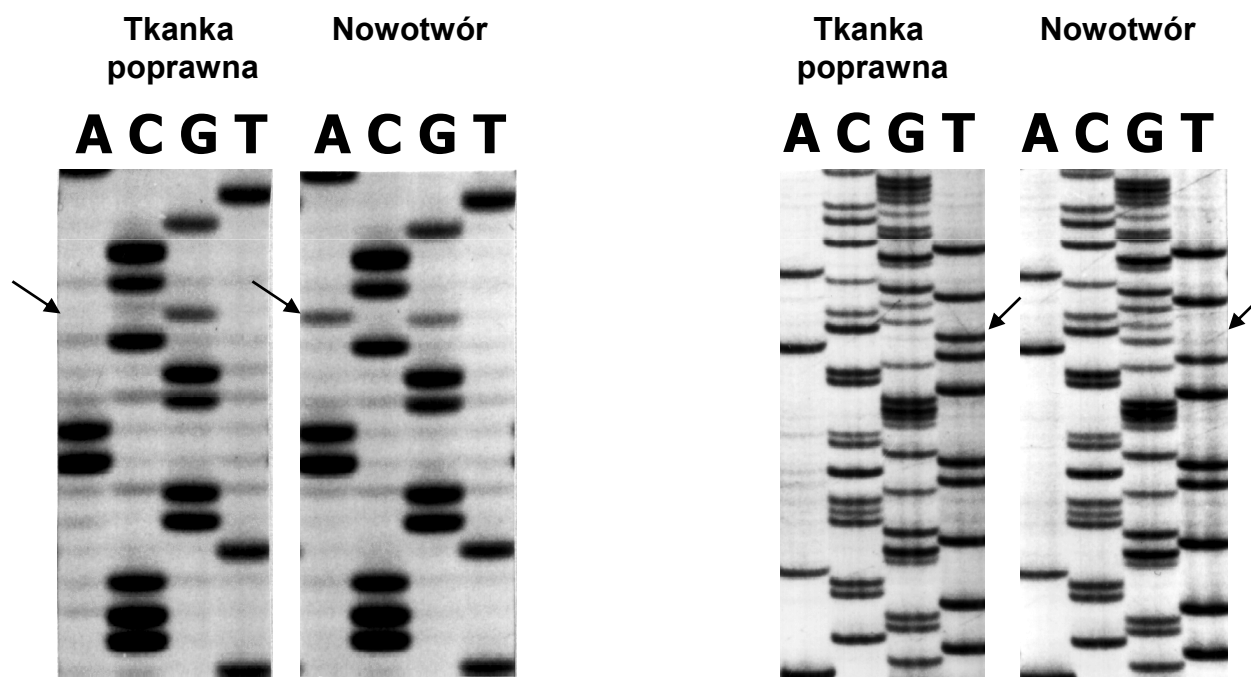


ABI 377 - płytowy



GA3500 - kapilarny

Sekwencjonowanie może być wykorzystane do poszukiwania mutacji. Na ilustracji przedstawiono wyniki sekwencjonowania z radioaktywnym znakowaniem. Poszukiwano mutacji genu *TP53* w DNA tkanki raka płuca.



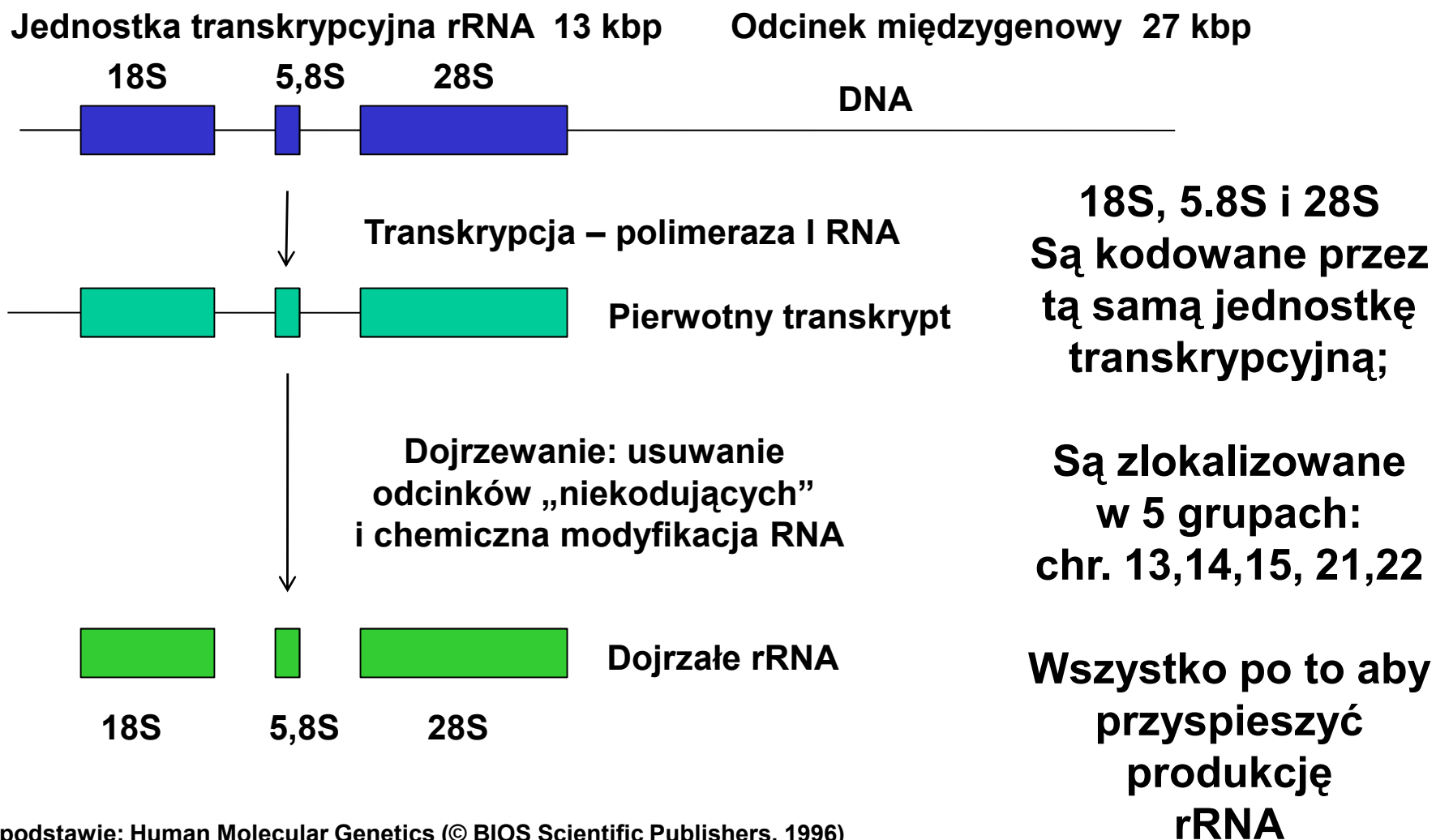
Wykorzystanie PCR do amplifikacji sekwencji mikrosatelitarnych i identyfikacji osobniczej.

Genotyp matki (ilość powtórzeń)		Genotyp badanych mężczyzn (ilość powtórzeń)							
Marker 1	16 20	<u>14</u>	<u>20</u>	<u>20</u>	<u>20</u>	<u>14</u>	<u>20</u>	16	16
Marker 2	8 10	8	8	8	10	<u>12</u>	<u>12</u>	<u>8</u>	<u>12</u>
		A		B		C		D	

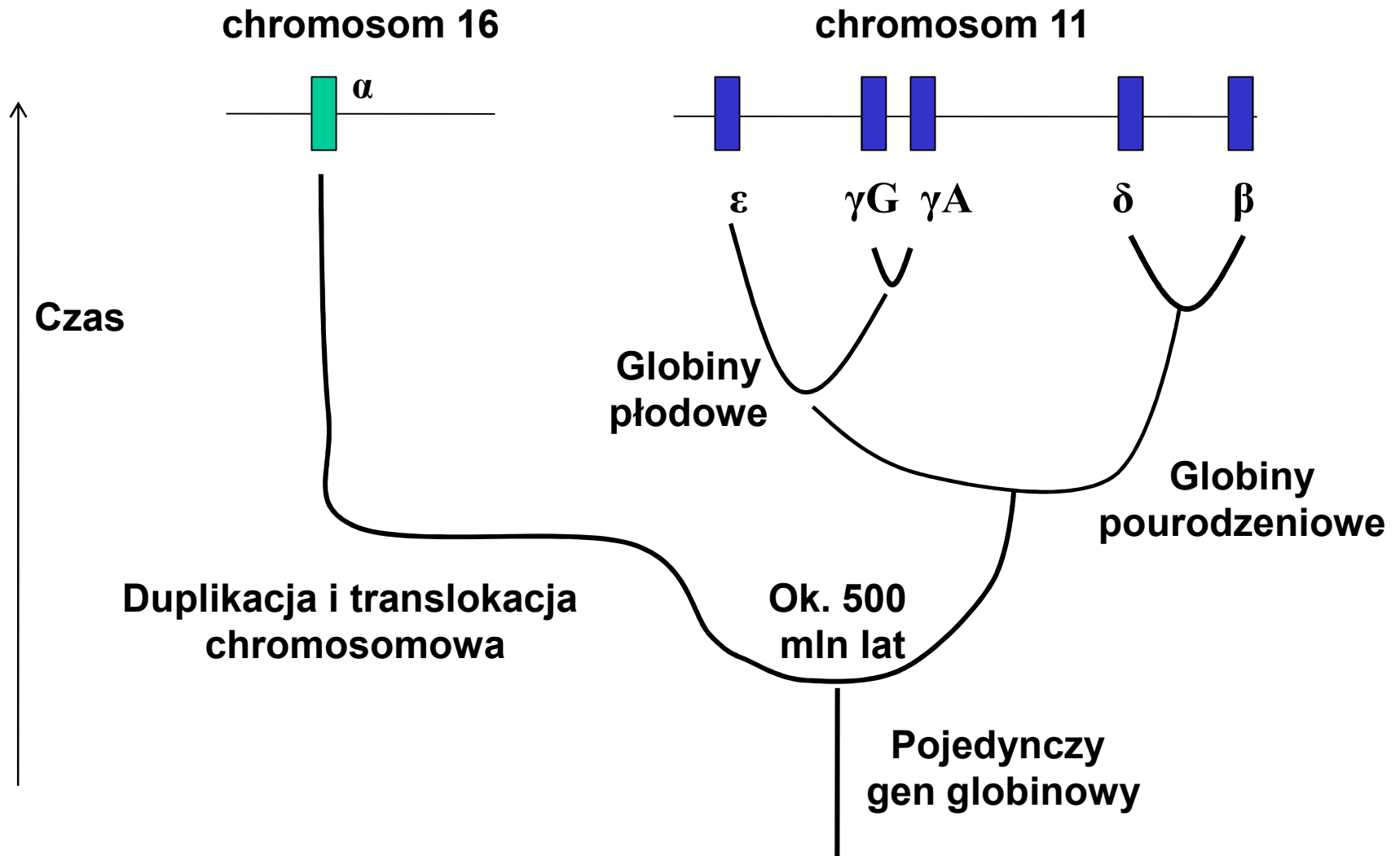
Genotyp dziecka
(ilość powtórzeń) 20 20
 8 12

Na podstawie analizy dziedziczenia markera 1 udało się wykluczyć ojcostwo mężczyzny D, na podstawie analizy markera 2 udało się wykluczyć ojcostwo mężczyzn A i B. Powtórzenie podobnej analizy dla większej liczby markerów daje prawie 100% pewność potwierdzenia ojcostwa.

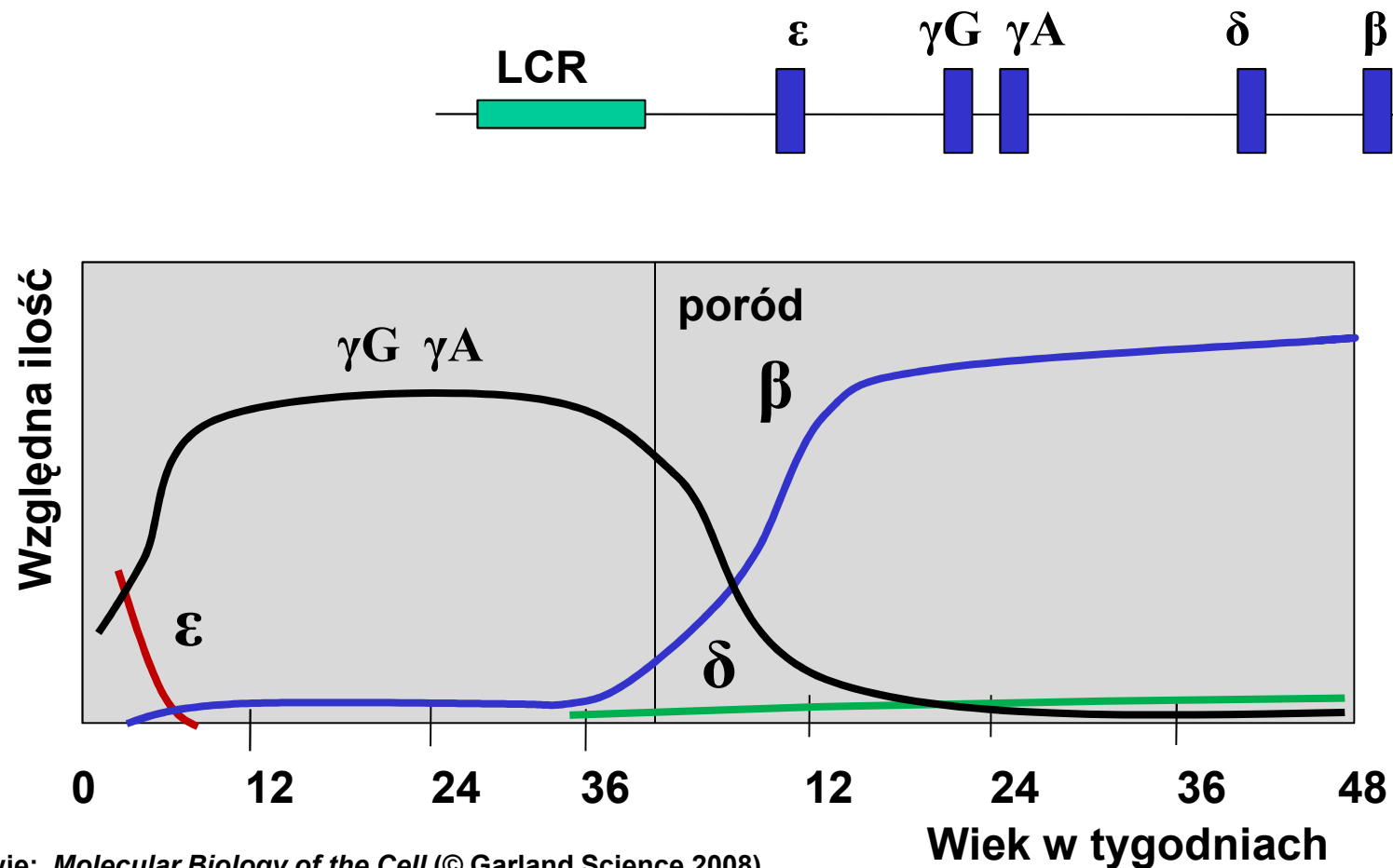
Struktura genu kodującego rRNA (ok. 1200 genów w ludzkim genomie).



Powielanie (duplikacja genów) oraz jej skutki. Drzewo genealogiczne genów łańcucha globiny – białka przenoszącego tlen we krwi zwierząt. Rozmieszczenie genów globiny w ludzkim genomie przedstawiono w górnej części ilustracji.



Grupa (klaster) genów globinowych na ludzkim chromosomie 11. Kluster ma długość ok. 100 tys. par zasad. U jego początku znajduje się tzw. sekwencja LCR kontrolująca ekspresję każdego z genów w grupie. Jej brak jest przyczyną groźnej anemii. Ekspresja poszczególnych genów globinowych podczas życia osobniczego jest zilustrowana u dołu. Globiny płodowe zanikają ok. 12 tygodnia po porodzie. Synchronicznie wzrasta ekspresja β -globiny. W czasie ewolucji pierwotny gen globinowy uległ powieleniu i spokrewnione geny pełnią nieco inne, ale podobne, funkcje.



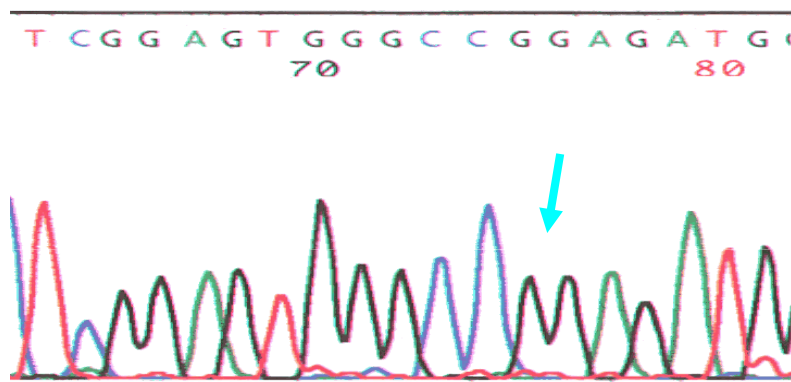
Na podstawie: *Molecular Biology of the Cell* (© Garland Science 2008)

Co to jest polimorfizm genetyczny?

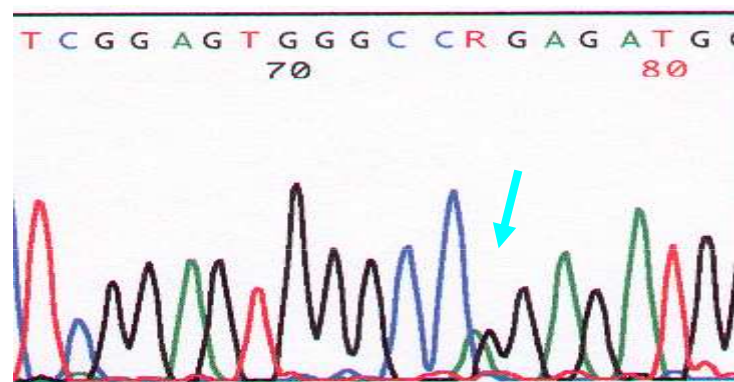
- Częste występowanie w populacji kilku odmian sekwencji genu (alleli). Rzadsza wersja występuje przynajmniej u jednego osobnika na 50.
 - Mutacja różni się od polimorfizmu funkcjonalnie – skutkuje ona obniżeniem płodności organizmu – dlatego występuje w populacji o wiele rzadziej niż polimorficzna wersja genu.
 - Polimorficzne odmiany genu (allele) mogą modulować funkcjonowanie produktu genu – białka, w niektórych warunkach może ono działać lepiej lub gorzej w porównaniu z najczęściej występującą wersją.
 - Duża część morfologicznych, fizjologicznych i psychologicznych różnic międzyosobniczych ma swoje źródło w polimorfizmie genetycznym.
-

Przykład polimorfizmu pojedynczego nukleotydu (SNP – ang. *single nucleotide polymorphism*). Oznaczenie, przy pomocy sekwencjonowania, polimorfizmu G23A genu XPA kodującego białko naprawy DNA.

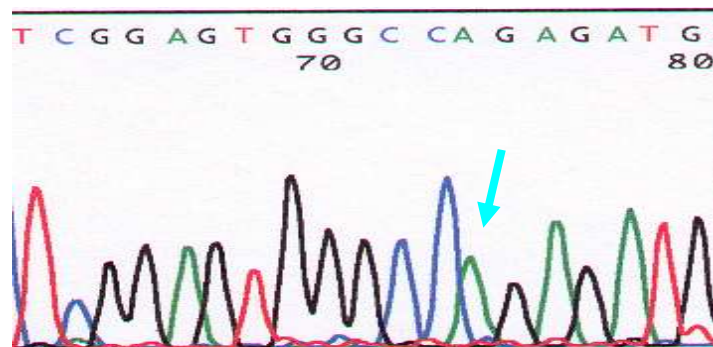
A. XPA (-4) G→A - homozygote GG



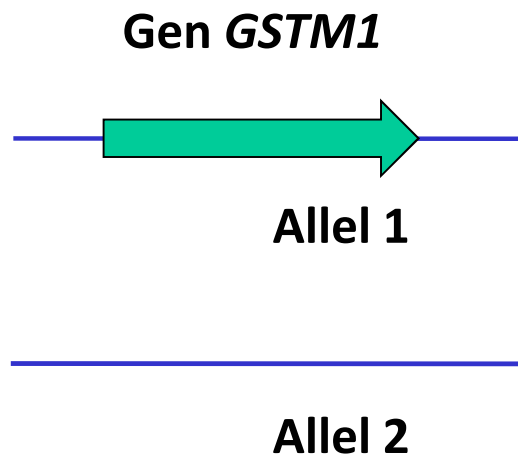
B. XPA (-4) G→A - heterozygote AG



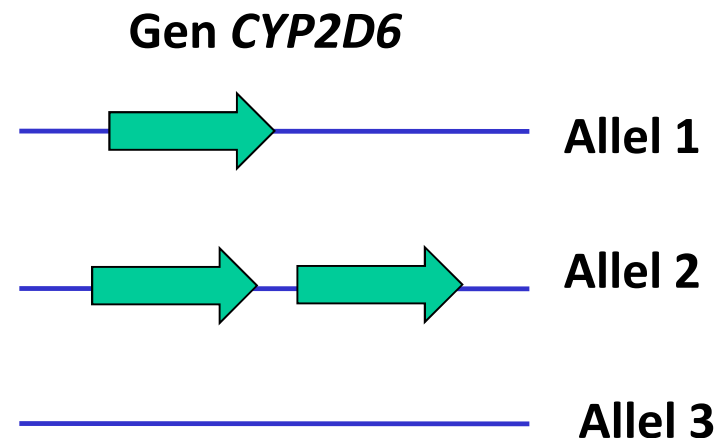
C. XPA (-4) G→A - homozygote AA



Występowanie polimorfizmów typu CNV (*copy number variations*). Ten polimorfizm polega na duplikacji lub delecji dużych segmentów DNA mogących zawierać nawet całe geny. Przez długi czas to źródło zmienności międzyosobniczej było niezauważane. Każda osoba ma średnio 100 polimorfizmów typu CNV.



Część osób, ok. 50% rasy białej,
nie posiada transferazy glutationowej μ
ze względu na allel z delecją genu.



Przykłady alleli genu *CYP2D6*
kodującego cytochrom
metabolizujący wiele leków:
tamoksyfen, β -blokery,
antydepresanty. Polimorfizm
może modulować reakcję na leki.



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Materiały dydaktyczne współfinansowane ze
środków Unii Europejskiej w ramach
Europejskiego Funduszu Społecznego.
