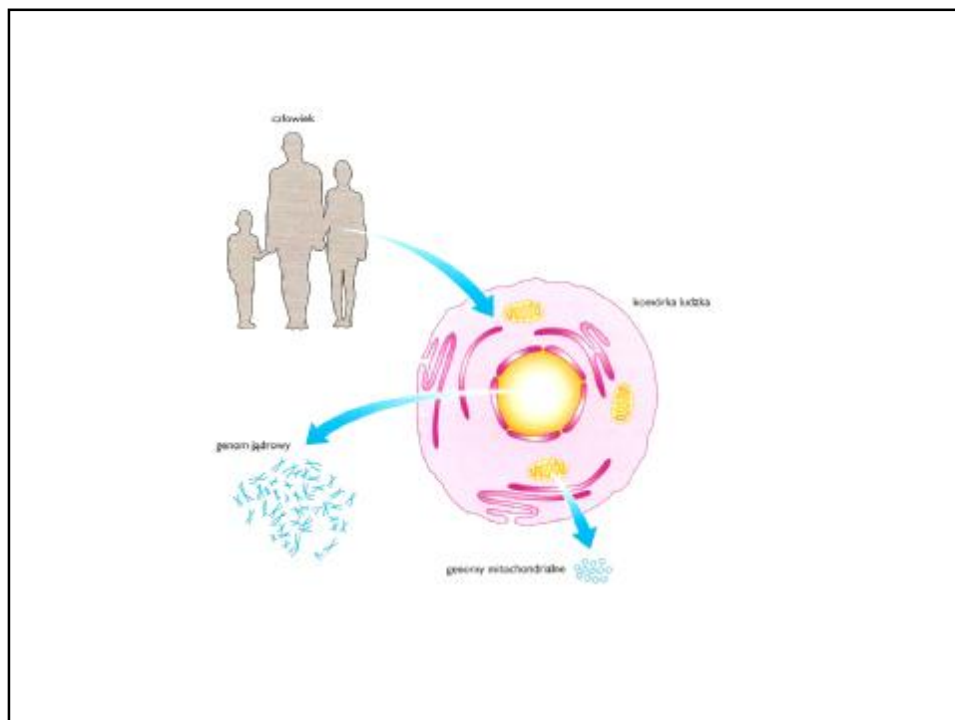


Wykład 3

✓ Budowa genomów a ewolucja

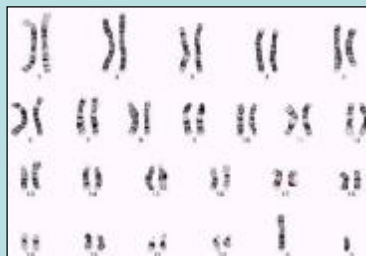
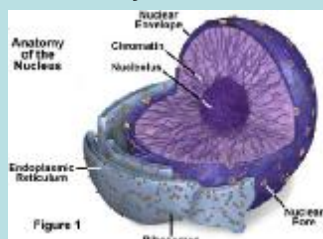


GENOM - to całkowity DNA znajdujący się w komórce, obejmujący geny oraz odcinki międzygenowe.

KILKA DANYCH O GENOMIE CZŁOWIEKA:

A. jądrowym:

- zawiera około 80 000 genów, ale odcinki kodujące stanowią tylko 3% całego genomu;
- genom jądrowy zawiera 3 000 000 000 pz;
- dzieli się na 24 różne liniowe cząsteczki DNA (22 autosomy i 2 chromosomy płci X i Y), z których najkrótsza ma 55 Mb (milionów pz), a najdłuższa 250 Mb i każda zawiera się w osobnym chromosomie;

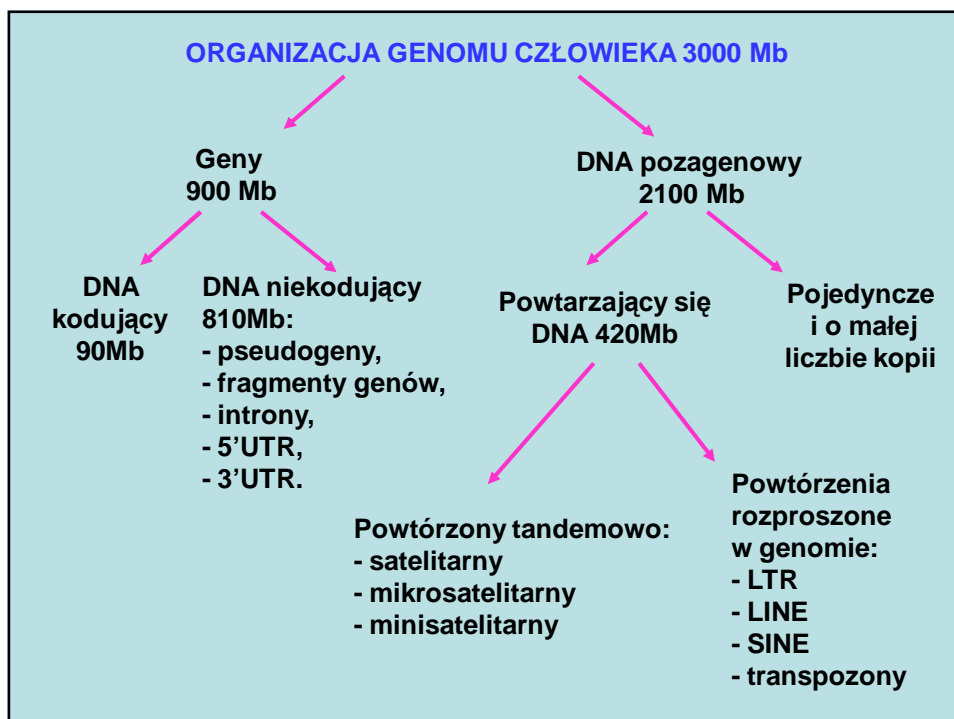
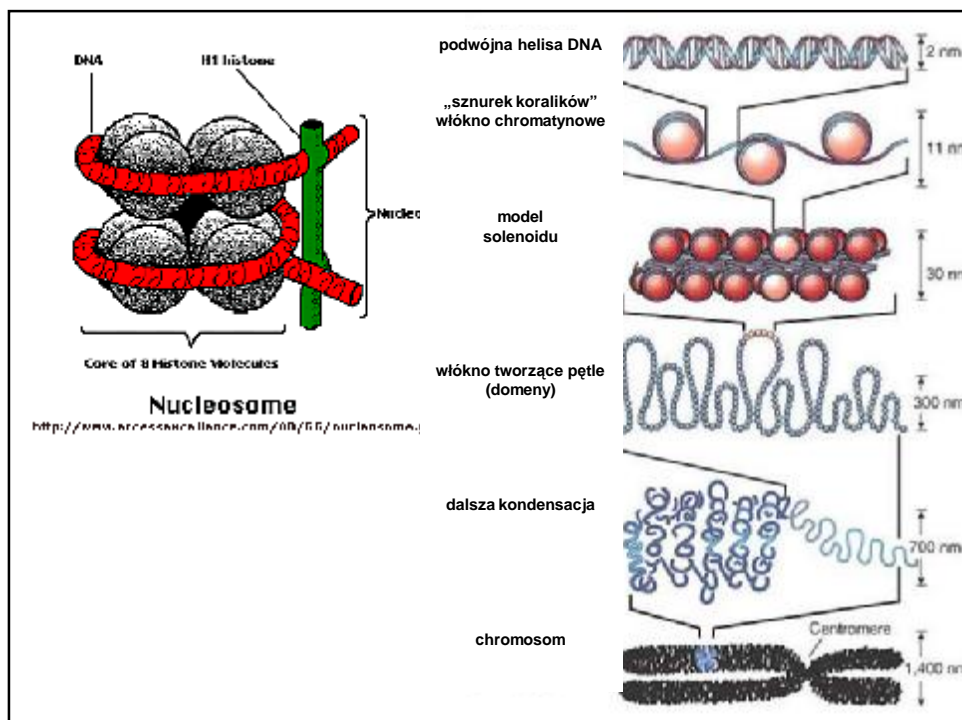


B. mitochondrialnym:

- jedna komórka ma około 8 000 kopii tego genomu, po około 10 kopii w każdym mitochondrium;
- kolista cząsteczka DNA o długości 16 569 pz;
- każda cząsteczka zawiera 37 genów:
 - 13 z nich koduje białka,
 - 22 - tRNA,
 - 2 - rRNA

Białka kodowane przez mtDNA to białka łańcucha oddechowego, ale większość białek wchodzących w skład łańcucha jest kodowana przez genom jądrowy.





Pseudogen – nie działająca kopia genu, zwykle zmutowana do tego stopnia, że informacja genetyczna staje się nie możliwa do odczytania.

Sekwencje powtarzające się rozproszone w genomie:

Ø **LINE** (*long interspersed nuclear elements*) – długie rozproszone sekwencje jądrowe

Ø **SINE** (*short*) – krótkie rozproszone

Ø **LTR** (*long terminal repeats*) – długie powtórzenia końcowe

Mikrosatellity – sekwencje składające się z tandemowo ułożonych, krótkich motywów powtarzających się

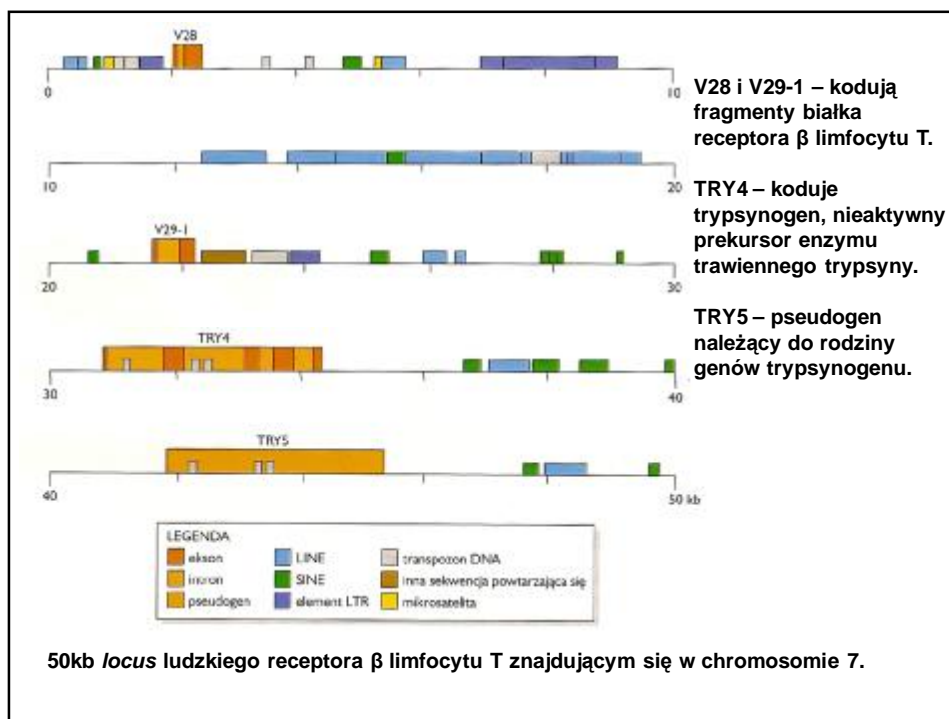
np. motyw GA powtórzony 16x:

GAGAGAGA GAGAGAGA GAGAGAGA GAGAGAGA
CTCTCTCT CTCTCTCT CTCTCTCT CTCTCTCT

lub 6 powtórzeń motywu TATT:

TATT TATT TATT TATT TATT TATT
ATAA ATAA ATAA ATAA ATAA ATAA

Wiele mikrosatelitów jest polimorficznych tzn. liczba powtórzeń różni się u poszczególnych osobników .



ZAWARTOŚĆ DNA W RÓŻNYCH GENOMACH:

- **haploidalny genom** – czyli jeden zespół chromosomów, jaki występuje w komórce haploidalnej (lub jednym chromosomie bakteryjnym) – „wartość C”
W komórce diploidalnej: przed replikacją 2C, po replikacji 4C
- **masa DNA** w pg (pikogramy)
- **długość** w mm (milimetry)
- **liczba par zasad** (par nukleotydów) - pz

PARADOKS C DNA

organizm	kb	mm	pg	% w genach kodujących białka
<i>Escherichia coli</i>	4 000	1,36	0,0036	100
<i>Saccharomyces</i>	13 500	4,6	0,009	69
<i>Drosophila</i>	165 000	56	0,18	33
<i>Homo sapiens</i>	2 900 000	990	3,5	9,27
<i>Protopterus</i>	102 000 000	34 700	142	0,4-1,2

PARADOKS C DNA

Flavell i wsp 1974

wykazali dodatnią korelację między rozmiarami genomu (w pg DNA) a zawartością sekwencji powtarzalnych (pg DNA) oraz stwierdzili, że w toku ewolucji, począwszy od 200 mln lat temu, genomy roślin w znacznej mierze, a w mniejszym stopniu genomy zwierząt ulegały kolejnym cyklom poliploidyzacji.

Zwiększanie się zawartości materiału genetycznego:

- duplikacje większych lub mniejszych sekwencji wyjściowych;
- poliploidyzacje.

Organizm może utrzymać tylko ograniczoną liczbę genów w stanie funkcjonalnym.

Genom człowieka zawierający 3×10^9 pz wystarczyłby do zakodowania 3 milionów białek średniej wielkości (330 aa – 1000 pz). Natomiast analiza poszczególnych frakcji mRNA produkowanych przez różne komórki wykazuje, że w rzeczywistości syntetyzowanych jest tylko ok. 30 000 – 40 000 różnych białek.

Rozmiary genomu roślin są dodatnio skorelowane z:

- ✓ czasem trwania cyklu życiowego,
- ✓ cykliów mitotycznych i meiotycznych,
- ✓ objętością jądra.

Gatunki z dużymi genomami, bogatymi w retroelementy:

- charakteryzują się niskim potencjałem adaptacyjnym i obniżoną dywergencją,
- są eliminowane ze środowisk o skrajnych warunkach,
- są zagrożone wyginięciem.

Te obserwacje wskazują na **ewolucyjną tendencję w kierunku redukcji rozmiarów genomów**.

- ✓ Redukcja rozmiarów genomów zachodzi u naturalnych allo- i autopoliploidów.

POWTARZALNE I NIEPOWTARZALNE SEKWENCJE DNA

U eukariotów (zwłaszcza wyższych) – przeważająca większość DNA to sekwencje nie niosące informacji o strukturze białek.

Część z nich odgrywa funkcje regulatorową np. w transkrypcji. Funkcje pozostałych nie są znane.

Wiele sekwencji DNA eukariotów powtarza się w genomie tysiące razy.

Sekwencje	Sekwencje powtarzalne DNA:	
	mikrosatelitarne	minisatelitarne
Charakterystyka		
Liczba powtórzeń	5 – 100	2 – 400
Motywy powtarzający się	1 pz – 6 pz	7 pz – 100 pz
Liczba loci	10^4 – 10^5 w genomie	ok. 4000 w genomie
Rozmieszczenie w genomie	równomiernie, często transkrybowane	rozproszone, często telomery
Występowanie	człowiek, zwierzęta, rośliny	człowiek, zwierzęta, rośliny, bakterie
Zastosowanie	mapowanie genów, diagnostyka molekularna, kontrola pochodzenia, szacowanie zmienności genetycznej	diagnostyka molekularna, kontrola pochodzenia, szacowanie zmienności genetycznej

ZNACZENIE NADMIARU DNA

Rola sekwencji niekodujących i „nadmiaru” DNA u organizmów eukariotycznych nie została wyjaśniona.

Część tego materiału może być związana z regulacją aktywności genów (promotory) lub utrzymaniem struktury i funkcji chromosomów np. taką rolę mogą pełnić skupienia satelitarnego DNA występujące w okolicy centromerów i telomerów.

Hipoteza Cavallera-Smitha (1978) główną funkcją DNA satelitarnego może być po prostu wypełnianie, wraz z innymi elementami, osłonki jądrowej i w ten sposób wpływanie na rozmiary jąder komórkowych. Z kolei zwiększenie rozmiarów jąder pociąga za sobą zwiększenie rozmiarów całych komórek, ze względu na konieczność utrzymania odpowiedniego stosunku jądrowo-cytoplazmatycznego, który jest na ogół stały dla danego typu tkanki.

Część pozostałego DNA jak sekwencje międzygenowe, pseudogeny prawdopodobnie pełnią jakąś funkcję i zostały błędnie zaliczone do „smietniska” genomowego.

ORGANIZACJA MATERIAŁU GENETYCZNEGO

Ü Wirusy:

bardzo zwarta informacja genetyczna, niekiedy może być dwukrotnie mniejsza niż potrzebna do zakodowania wszystkich białek danego wirusa.
Wyjaśnienie: cząsteczki RNA przepisane z tej samej sekwencji DNA mogą kodować dwa różne białka zależnie od punktu startu i fazy odczytu. Jest to układ „**zachodzących na siebie genów**”, zmniejsza on ich możliwości ewolucyjne, gdyż ta sama mutacja może wywołać w każdym z tych genów odmienne skutki.

Ü Prokariota:

informacja genetyczna jest ciasno upakowana, geny ułożone są bezpośrednio po sobie, sekwencje regulatorowe są wspólne dla całego operonu.

Bardzo rzadko występują odcinki DNA niekodującego (krótkie, nie więcej niż kilkadziesiąt pz) oraz bakteryjne transpozony. Tylko geny kodujące rRNA i tRNA powtarzają się w genomie, ale w niewielkiej liczbie kopii.

Ü Eukariota:

w chromosomach przeważają odcinki niekodujące.
Struktura genomu eukariotów stanowi mozaikę sekwencji:
kodujących i niekodujących, niepowtarzalnych na przemian
z powtarzającymi się.

Powtarzające się sekwencje w formie DNA satelitarnego są zgrupowane
głównie w heterochromatynie konstytutywnej (centromery i telomery)

Szacuje się, że u człowieka funkcjonalne geny są oddalone od siebie
średnio o około 35 000 pz.

Rozmieszczenie genów jest jednak bardzo nierównomierne,
gdyż w niektórych miejscach występują one w zagęszczeniu
np. blisko sprzężone geny kodujące globiny A γ i G γ dzieli odległość
3000pz, a główny region genów zgodności tkankowej MHC o długości
4 x 10⁶ pz zawiera 74 geny, które częściowo nawet zachodzą na siebie,
jak u wirusów.

Ü W genomie eukariotów rozsiane są **regiony bardzo zmienne**
(*hypervariable*) złożone z krótkich sekwencji powtarzających się
(minisatelitarnych) wielokrotnie o różnej liczbie kopii.

Zmienna liczba tych odcinków wynika z częstych rekombinacji np.
wskutek niesymetrycznego crossing-over.
Są to „**gorące miejsca rekombinacji**”.

Ü Np. u człowieka po obu stronach genu kodującego insulinę
(liczącego 1430 pz) występują tzw. fragmenty flankujące, w których
powtarza się krótki odcinek: ACA GGG GTG TGG GG.

Liczba powtórzeń u badanych osobników waha się od 364 do 2 926,
przy czym u ponad 60% stwierdzono układy heterozygotyczne tzn.
homologiczne chromosomy zawierały niejednakową liczbę kopii.

STRUKTURA GENU U ORGANIZMÓW EUKARIOTYCZNYCH

1977 - geny różnią się budową pomiędzy prokariotami a eukariotami.

U prokariota - struktura **ciągła** genu, w której liczba kodonów odpowiada liczbie aminokwasów w kodowanym białku.

Typowy gen eukariotyczny – gen **podzielony**, składa się z odcinków kodujących: eksonów i intronów.

Dopiero po precyzyjnym wycięciu intronów czyli tzw. składaniu RNA (**splicing**) i innych modyfikacjach, gotowy mRNA przechodzi do cytoplazmy. Introny są degradowane jeszcze w jądrze.

Mechanizm wycinania intronów okazał się bardzo znaczący dla rozważań nad ewolucją materiału genetycznego.

W przypadku większości jądrowych genów kodujących białka, introny są wycinane z pre-mRNA enzymatycznie przy pomocy dużych kompleksów białkowo-rybonukleinowych, rozpoznających specyficzne sekwencje graniczne na początku i końcu intronu.

Natomiast w niektórych przypadkach dotyczących np.
 - jądrowych genów kodujących rRNA u pierwotniaka *Tetrahymena*,
 - niektórych genów mitochondrialnych i chloroplastowych
 wycinanie intronów zachodzi **bez udziału enzymów**,
 gdyż sam RNA pełni rolę enzymu tzw. RYBOZYMU, powodując samowycinanie się intronów z własnej cząsteczki.

Jest to przypuszczalnie pozostałość pierwotnego ewolucyjnie mechanizmu, świadczącego o tym, że katalityczne właściwości RNA odgrywały rolę w funkcjonowaniu genów.

POCHODZENIE INTRONÓW

- brak jednoznacznej odpowiedzi

1. **TEORIA I - PÓŹNE POCHODZENIE INTRONÓW:**
czyli struktura genów początkowo była ciągła, niepodzielona.
2. **TEORIA II - WCZESNE POCHODZENIE:**
 - wycinanie intronów może odbywać się bez udziału enzymów (niekodujące sekwencje RNA muszą mieć długą historię, skoro istnieje taki pierwotny system ich usuwania),
 - introny występują już u archebakterii.
3. **TEORIA OBECNA to TEORIA II –** pierwotny zapis informacji genetycznej był **nieciągły**.
Nieciągłość i występowanie sekwencji powtarzających się u prymitywnych organizmów o mało sprawnych mechanizmach komórkowych mogło zwiększać szansę, że przynajmniej część tej informacji zostanie prawidłowo powielona i przepisana.

5. Prawdopodobnie więc eukariota wywodzą się od tych organizmów, na które presja w kierunku pozbycia się nadmiaru DNA, z jakiś nieznanych powodów nie zadziałała, co pozwoliło im utrzymać plastyczność genomów i otworzyło drogę do wielokierunkowej ewolucji.
6. Nieciągły zapis informacji genetycznej zwiększa również szansę rekombinacji, gdyż częstość wymiany między odcinkami jest proporcjonalna do odległości między nimi.

Ponadto crossing-over zachodzący w obrębie intronów może doprowadzić do przetasowania całych egzonów bez naruszania ich struktury.

W ten sposób z wyselekcjonowanych już poprzednio elementów zawartych w różnych genach mogły być zestawiane nowe geny o bardziej złożonej i wszechstronnej funkcji.

WYDŁUŻANIE SIĘ PRAGENÓW

- długość eksonów odpowiada przeciętnie odcinkom DNA kodującym polipeptydy zbudowane z około 20-40 aa.
- prageny były krótkimi odcinkami DNA, oflankowanymi sekwencjami niekodującymi, które „dały” dzisiejsze introny.
- z pragenów poprzez duplikacje i przetasowanie w różnych kombinacjach powstałyby dzisiejsze geny.
- o takim wydłużaniu się genów w trakcie ewolucji świadczy fakt, że w wielu białkach powtarzają się identyczne lub w wysokim stopniu podobne sekwencje aminokwasów.

OBECNOŚĆ INTRONÓW W GENACH NIE JEST TYLKO ŚLADEM PO ZAJŚCIU ZDARZEŃ EWOLUCYJNYCH

- struktura genu podlegała ewolucji, a dzisiejszy układ eksonów może mieć znaczenie funkcjonalne.
- w niektórych przypadkach stwierdzono, że poszczególne eksony wyznaczają w kodowanym białku tzw. domeny, czyli wyodrębnione przestrzennie i funkcjonalnie fragmenty.
- np. gen kodujący globinę.

W genie tym występują 3 eksony, z których środkowy koduje domenę globiny tworzącą kieszeń dla hemu. Badania wykazały, że w obrębie środkowego eksonu powinien być występować jeszcze dodatkowy intron, który przypuszczalnie został wyeliminowany przez dobór naturalny, gdyż zapewniało to większą stabilność kieszeni hemowej.

- ✓ **znane są przypadki zarówno utraty intronów jak i powstania nowych.**
- ✓ **najprostszy mechanizm „uwolnienia się” genu od intronów polega na włączeniu do chromosomu kopii DNA przepisanej z mRNA, który intronów już nie posiada.**
Potrzebnej do tego procesu odwrotnej transkryptazy mogą dostarczyć retrowirusy lub niektóre ruchome elementy DNA.
- ✓ **funkcjonowanie genów przepisanych z RNA wymaga jednak, aby zostały one włączone w pobliżu sekwencji regulatorowej, zdolnej uruchomić transkrypcję.**
W przeciwnym razie powstanie pseudogen.

ALTERNATYWNE SPOSOBY WYCINANIA INTRONÓW

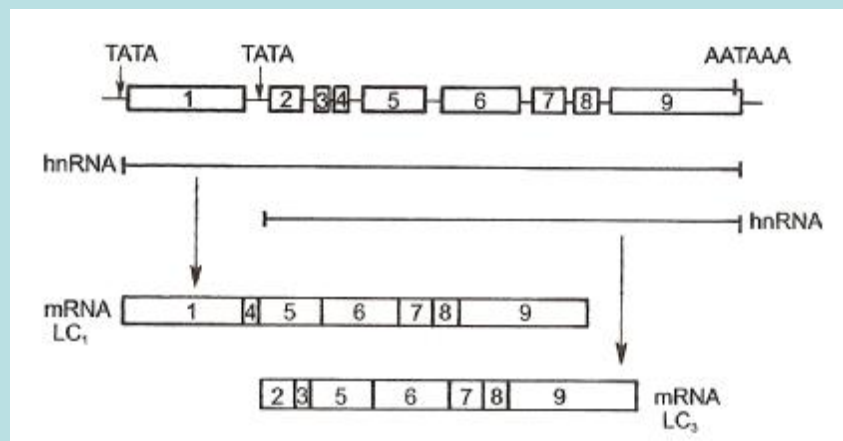
Podczas rozwoju organizmu i różnicowania się organów, w różnych tkankach i narządach na matrycy tego samego genu mogą być produkowane odmienne białka.

Decyduje o tym:

- wybór sygnału rozpoczynającego się i kończącego transkrypcję
- sposób wycinania intronów – alternatywne sposoby zwiększają plastyczność funkcjonowania genów.

Pojęcie egzonu i intronu jest względne, gdyż dany egzon może wejść w skład mRNA lub zostać wycięty razem z sąsiadującymi intronami, stanowiąc wtedy sekwencję niekodującą.

Gen kodujący łańcuchy lekkie miozyny kurczęcia.



REDAGOWANIE RNA (RNA editing)

- mechanizm zmian potranskrypcyjnych polegający na zamianie, dodaniu lub usunięciu pewnych zasad z pierwotnego transkryptu.

Edycja RNA typu (C → U) zachodzi u roślin i zwierząt. U roślin głównie dotyczy genomów mitochondrialnego i chloroplastowego.

np. apolipoproteina B (apo-B) – udział w metabolizmie tłuszczu. Polipeptyd (512 kDa) wytwarzany w komórkach wątroby jest dłuższy niż w komórkach nabłonka jelita cienkiego (u ludzi).

W komórkach jelita w mRNA jeden z kodonów CAA (glutamina) po deaminacji cytozyny staje się kodonem terminacyjnym (UAA), co powoduje skrócenie transkryptu (białko 241 kDa).

W tym przypadku zapisana w DNA informacja jest nieczytelna i staje się zrozumiała dopiero po zredagowaniu.

Edycja RNA apo-B zachodzi przy udziale deaminazy cytydynowej apobec-1, która ma zdolność do wiązania się do RNA.