



Transpozony

Transpozycja

- Transpozycja to przemieszczanie się materiału genetycznego z jednego miejsca na drugie
- Sekwencje DNA posiadające taką zdolność nazywane są elementami mobilnymi
- Konserwatywne elementy mobilne przenoszą się z jednego miejsca na drugie
- Replikujące się elementy są skopiowane i nowa kopia zostaje wbudowana w inne miejsce;
- Retropozony to elementy mobilne przenoszące poprzez transkrypcję na RNA, które następnie ulega transkrypcji odwrotnej w cDNA i dopiero cDNA zostaje wbudowane w nowe miejsce
- Liczba tych elementów powiększa się z czasem.

Odkrycie transpozonów



Barbara McClintock (1984)

Pierwszy transpozon został odkryty w 1947 roku przez Barbarę McClintock. Był to element *Activator (Ac)* w kukurydzy. Należy on do klasy transpozonów poruszających się poprzez „wycięcie i wklejenie” (“cut-and-paste”) produktu pośredniego DNA. Elementy te są obecne u wszystkich wyższych eukariotów od roślin po ssaki.

Charakterystyka transpozonów

- Występują w genomach wszystkich organizmów
- Transpozony są odpowiedzialne za zmiany genetyczne
- Istotnie wpływają na ewolucję genomów
 - Plastyczność genomów
 - tworzenie nowych genów
 - regulacja transkrypcji
- Narzędzia manipulacji genetycznej

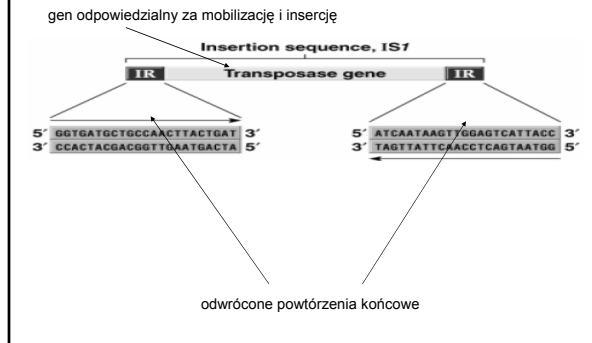
Klasyfikacja

- Dwie główne grupy u bakterii:
 - Sekwencje insercyjne - insertion sequence (IS)
 - Transpozony - Transposons (Tn)
- Trzy klasy u Eukariotów
 - retrotranspozony
 - LTRs, LINEs, SINEs, retro(pseudo)geny
 - transpozony DNA (konserwatywne transpozony)
 - Mariner
 - inne
 - np. helitron – replikują się przy użyciu formy kolistej (rolling circle replication)

Sekwencje Insercyjne (IS)

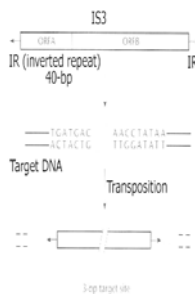
- Najprostsze z transpozonów występujące w chromosomach bakteryjnych i plazmidach.
- Kodują wyłącznie geny odpowiedzialne za mobilizację i insercję elementów.
- Wielkość od 768 pz do 5 kbp.
- IS1 (po raz pierwszy zidentyfikowany w operonie galaktozowym E. coli) jest długości 768 pz i występuje w 4-19 kopiach w chromosomie E. coli.
- Na końcach wszystkich elementów IS znajdują się odwrócone powtórzenia końcowe (inverted terminal repeats).

Struktura elementów IS



Sekwencje Insercyjne (IS)

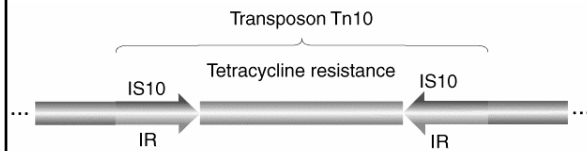
Podczas insercji krótki odcinek DNA w miejscu integracji zostaje zduplikowany.



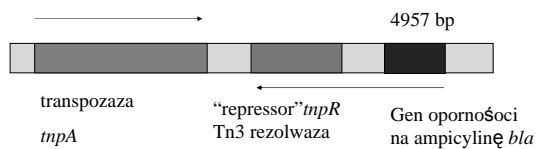
Złożone transpozony bakteryjne

- Powtórzone sekwencje na końcach, zazwyczaj odwrócone
- Powtórzone końce są same w sobie elementami IS i mogą się przenosić na nowe miejsce samodzielnie
- Końce mobilizują cały odcinek DNA znajdujący się pomiędzy nimi
- Często przenoszą geny odpowiedzialne za odporność antybiotykową, np. Tn3 (ampicylinę), Tn5 (kanamycynę), Tn10 (tetracyklinę)
- Często znajdują się w plazmidach

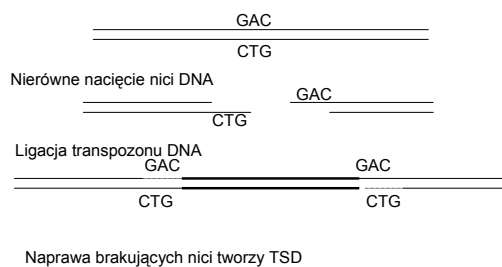
Transpozony bakteryjne



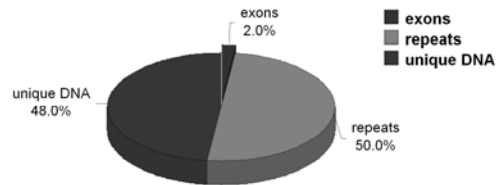
Struktura transpozonu Tn3



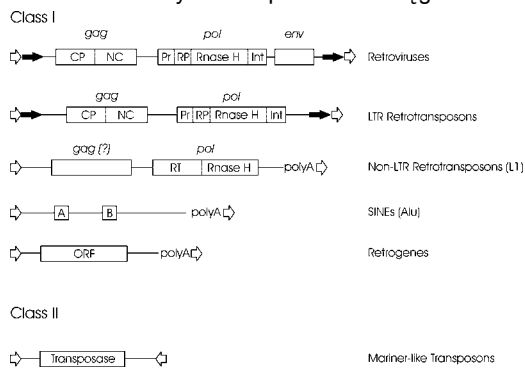
Mechanizm tworzenia duplikacji w miejscu integracji (target site duplication - TSD)



Różne typy sekwencji w genomach ssaków



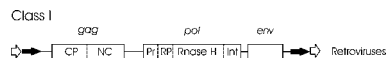
Struktura różnych transpozonów u kręgowców



Makalowski W. (2001) *Acta Biochimica Polonica* 48: 587-598.

Retrowirusy

- Wirusy podobne w swojej strukturze do transpozonów
- Po inwazji komórki 'gospodarza' RNA genomowe wironu ulega odwrotnej transkrypcji w wirusowe cDNA
- cDNA wironu może zintegrować się z genomem gospodarza i powstaje prowirus
- DNA pochodzące od prowirusa może następnie ulec transkrypcji w RNA, które może zostać użyte do syntezy białek wirusa lub jako genomowa sekwencja nowego wironu
- Sekwencje kodujące retrowirusów posiadają na końcach elementy LTR (long terminal repeats)



Charakterystyka elementów posiadających LTR

- Posiadają długie powtórzenia końcowe" (Long terminal repeats - LTR): istotne w cyklu replikacyjnym
- Geny: *gag*, *pol*
- Nie posiadają genu *env* i w związku z tym nie mogą samodzielnie przemieszczać się między komórkami
- *Pol* jest polipeptydem, prekursorem z którego tworzą się odwrotna transkryptaza (RT), integraza (IN), RNase H (RH) i proteaza (PR)
- Tworzą cząsteczki wirusopodobne (virus-like particles - VLPs)
- Integraza jest funkcjonalnie i strukturalnie podobna do transpozy transpozonów DNA
- Integracja pozostawia ślad w postaci TSD

Elementy bez LTR

- Cechy charakterystyczne
 - Brak powtórzeń końcowych
 - Ogon poliadenylowy na końcu 3'
 - Obecność zduplikowanego miejsca integracji na obu końcach (TSD)
 - Kilka superrodzin zostało wyodrębnionych na podstawie sekwencji białek kodowanych przez tę klasę elementów
 - Mogą być autonomiczne (same kodują maszynę umożliwiającą transpozycję) lub nie

Elementy bez LTR – Long Interspersed Elements (LINEs)

- Najbardziej poznanym elementem jest ludzki L1
 - Tworzy 21% ludzkiego DNA
 - 850,000 kopii ale tylko 3-5,000 jest pełnej długości – większość „obcięta” na końcu 5' – wynik niepełnej działalności odwrotnej polimerazy która często „odpada” za wcześniej od matrycy
 - Dwie ramki odczytu – I-sza o nieznanym potencjale kodującym (pozostałość po *gag?*), II-ga (*pol*) koduje maszynę transpozonoową
 - L1 jest dostarczycielem odwrotnej transkryptazy również dla niesamodzielnych retropozonów, np. SINEs, jak również jest odpowiedzialny za tworzenie retro(pseudo)genów

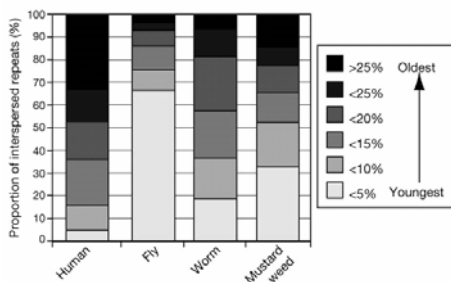
Elementy bez LTR – Short Interspersed Elements (SINEs)

- Z definicji są krótkie – do 1000 nt
- Nie kodują żadnego białka i są nieautonomiczne
- Ewolucyjnie pochodzą od genów RNA – tRNA lub 7SL RNA
- Najlepiej poznanym elementem jest Alu
 - Specyficzne dla ssaków naczelnych
 - Tworzy 10% ludzkiego DNA
 - Ponad 1 mln kopii, element pełnej długości ma ok. 300 nt
 - nierówny dimer, każda jednostka pochodzi niezależnie od 7SL RNA, są połączone sekwencją bogatą w adenozyne (A)
- Inne przykłady SINEs to: MIR (obecny w genomach wszystkich ssaków), B1 specyficzny dla gryzoni

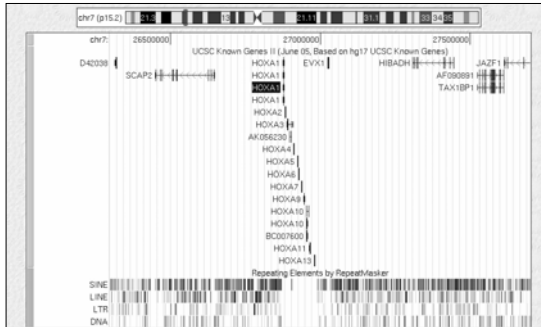
Elementy bez LTR – Retro(pseudo)geny

- Są produktami aktywności odwrotnej transkryptazy na dojrzałym mRNA
- Cechy charakterystyczne
 - Brak intronów i sekwencji promotorowych
 - Obecność TSD i ogona poliA
 - Większość z nich kończy jako pseudogeny głównie ze względu na brak promotorów
- W genomie ludzkim jest ok 20.000 pseudogenów z czego ponad 70% powstała przez retropozycję
- Większość z wykrywalnych pseudogenów jest wynikiem niedawnych retropozycji, które nastąpiły po rozdzieleniu się naczelnych i gryzoni

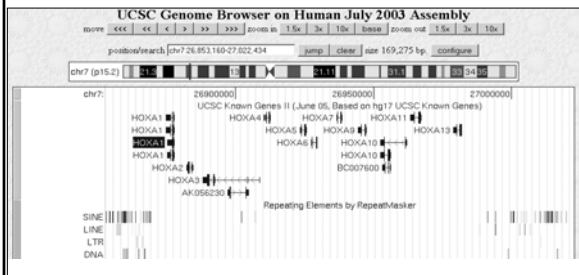
Rozkład transpozonów w różnych genomach ze względu na wiek



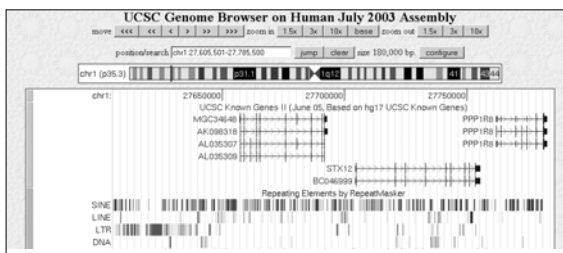
Dystrybucja transpozonów nie jest jednolita w genomie



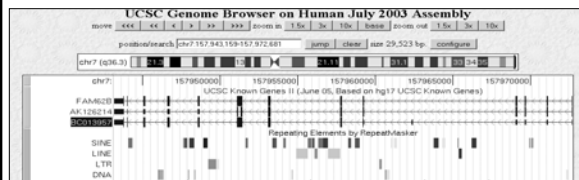
Dystrybucja transpozonów nie jest jednolita w genomie



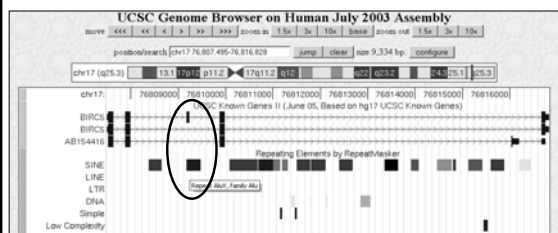
Dystrybucja transpozonów nie jest jednolita w genomie



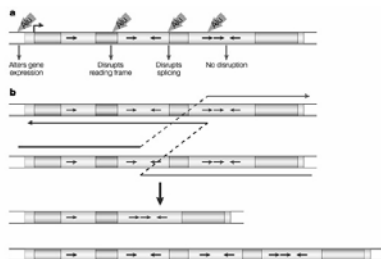
Transpozony zazwyczaj są
ulożone w intronach i
przestrzeniach międzygenowych



Rzadko można je znaleźć w
sekwencjach kodujących białko



Szkodliwy wpływ elementów Alu na
genom człowieka

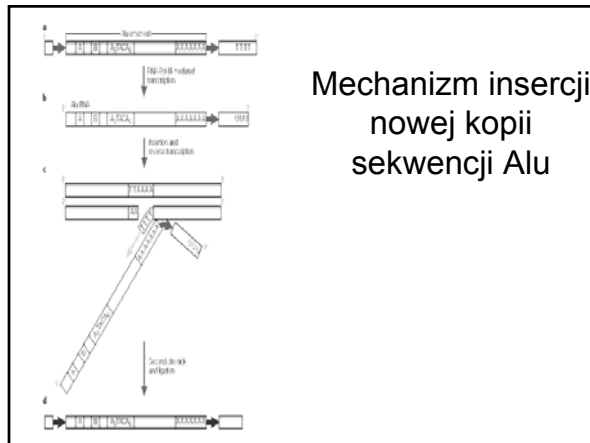


Batzer MA, Deininger PL. Nat Rev Genet. 2002 May;3(5):370-9.

Nature Reviews | Genetics

Hipotezy na temat roli transpozono- nów w genomie gospodarza

- Ewolucja regulacji genów
- Początek replikacji DNA
- Ewolucja struktury i funkcji genów
- Tworzenie wysp CpG
- Naprawa uszkodzeń chromosomowych
- Regulacja translacji podczas stresu komórkowego
- Davidson and Britten (1973) Quart Rev Biol
- Jelinek et al. (1981) PNAS
- Chesnokov et al. (1995) JBC
- Teng et al. (1996) Nature
- Chu et al. (1998) MCB

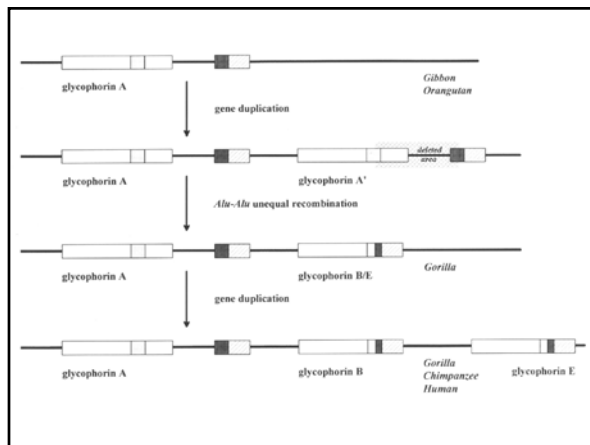


Dlaczego transpozony są tolerowane w genomach?

- f Ekspansja genomów poprzez akumulację transpozonów zmniejsza ich negatywny wpływ na gospodarza
- f Transpozony zwiększają zdolność ewolucyjną gospodarza
 - „gorące” miejsca rekombinacji
 - źródło gotowych do użycia motywów
 - mechanizm tasowania genomu

Gorące miejsca rekombinacji

Ewolucja glikoforyn u ssaków naczelnych

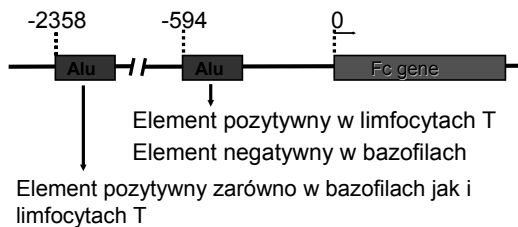


Typy motywów genomowych pochodzących z transpozonów

- f* Elementy regulujące transkrypcję
- f* Sygnały poliadenylacji
- f* Sekwencje kodujące białka
- f* retrogeny

Elementy regulujące transkrypcję

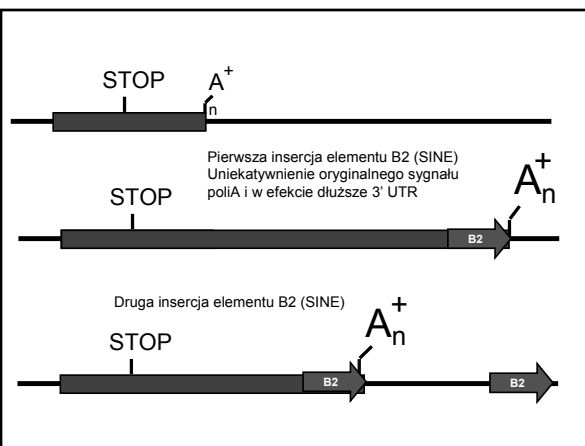
Komórkowo-specyficzna regulacja transkrypcji genu receptora IgE (FcεRI-g)



Hamdi et al. J Mol Biol. 2000 Jun 16;299(4):931-9.

Sygnały poliadenylacji

Ewolucja końca 3' mysiej kinazy fosforylanowej-g

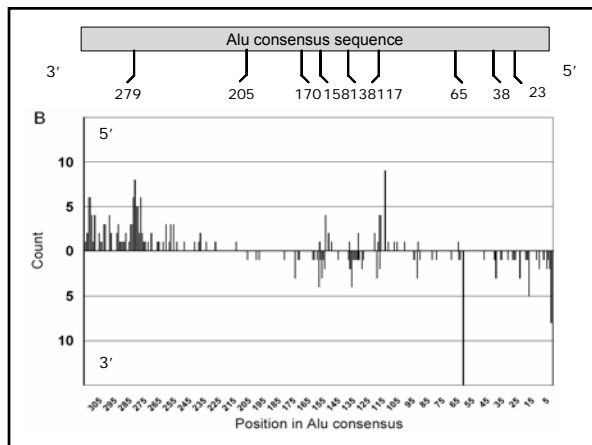


Sekwencje kodujące białka

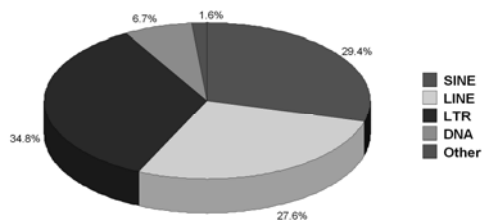
f Różne transpozony są znajduwane w otwartych ramach odczytu (ORF) ale ...

f Elementy Alu są predysponowane do włączenia do mRNA

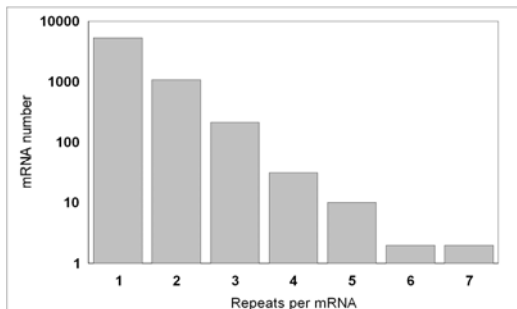
- Kryptyczne miejsca splicingowe istnieją w Alu
- ponad 2000 ludzkich mRNA zawierają fragmenty Alu w ORF
- elementy Alu w ORF mogą zmienić własności fizyko-chemiczne białek



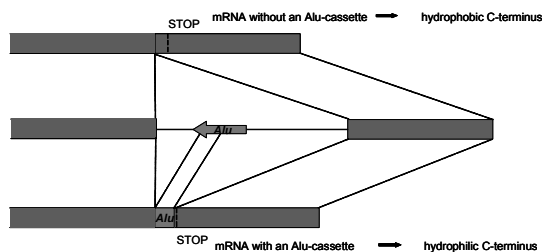
Udział różnych transpozonów w sekwencjach kodujących białka



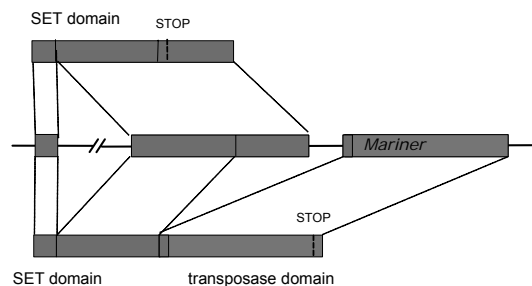
Liczba transpozonów znalezionych w pojedynczych mRNA



Decay Accelerating Factor – translacja mRNA zawierającego kasetę Alu prowadzi do rozpuszczalnej formy tego białka



Aktywacja kryptycznego (cryptic) miejsca 5'-splicingowego (donor) i eksaptacja części *Marinera* doprowadziła do powstania nowego białka u człowieka.



Kategorie białek z kasetami transpozonowymi w ORF

	human	rodent	other mammals	other vertebrates
proteins	594	115	36	6
proteins with assigned function/	59	12	7	5
nucleic acid binding	16			1
enzyme	15	3		1
signal transducer	11	6	4	1
structural_protein	7			
cell adhesion molecule	3		2	
apoptosis regulator				
enzyme_regulator				
transporter	5	2		
cell cycle regulator	2			1
ligand binding or carrier	10	1		
chaperone			1	
defense immunity protein	1			1
obsolete	1			
function unknown	535	103	29	1

Ciągle otwarte pytanie – czy rzeczywiście mRNA z kasetami transpozonowymi kodują funkcjonalne białka

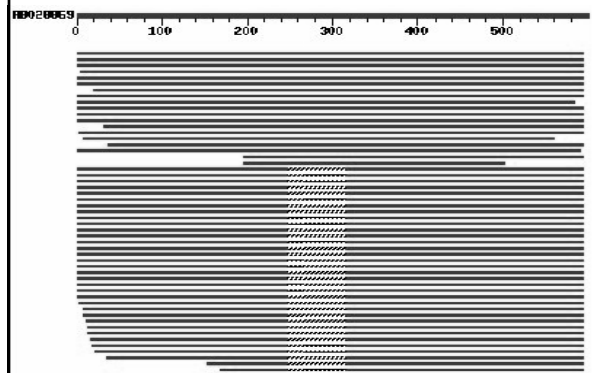
Fragmenty TE w białkach ze znaną strukturą trzeciorzędową

PDB ID	Długość insertu	Pochodzenie insertu	Różnica sekwencji (%)
1i49	21 aa	LTR27	22.8
1dmt	12 aa	L2	17.1
1g7f	34 aa	L3	27.7
1h9f	26 aa	L3	33.8
1rw2	41 aa	L3	31.7
1m1l	30 aa	L3	25.0

Przykład I: human survivin



Study case I: human survivin



Study case I: human survivin

```

      10      20      30      40      50      60      70
13096773 > 1  NGAPILPPAM QPFLKDHRIIS TFKNWPFLEQ CACTPERMAE AGFINCPTEN EPDLAQCFPC FKELEGWEPD 70
7416053  > 1  NGAPILPPAM QPFLKDHRIIS TFKNWPFLEQ CACTPERMAE AGFINCPTEN EPDLAQCFPC FKELEGWEPD 70

      80      90     100     110     120     130     140
13096773 > 71 DDPI-----EEH KKHSOGCAFL SVKQFREL T LGEFLKLDRK RAKNKIAKET 117
7416053  > 71  DDPIgggtva yaentstlgy rgyritrEEH KKHSOGCAFL SVKQFREL T LGEFLKLDRK RAKNKIAKET 140

```


**Przykład I: białko survivin u człowieka
– funkcja w apoptozie**



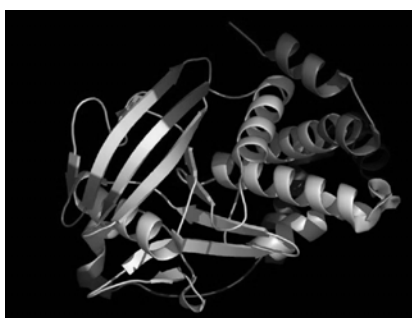
Kaseta Alu w drodze alternatywnego splicingu zostaje wprowadzona w środek helisy alfa, która jest odpowiedzialna za wiązanie atomu cynku

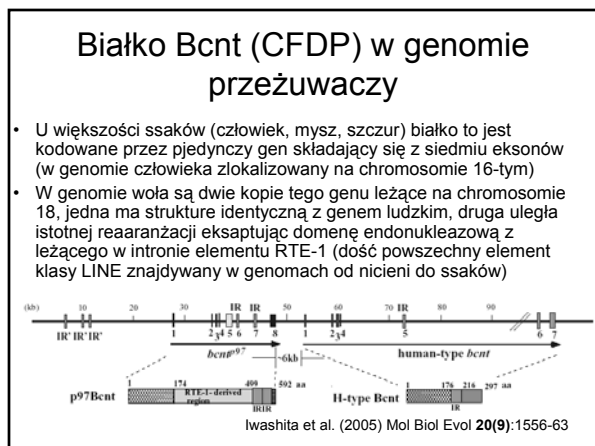
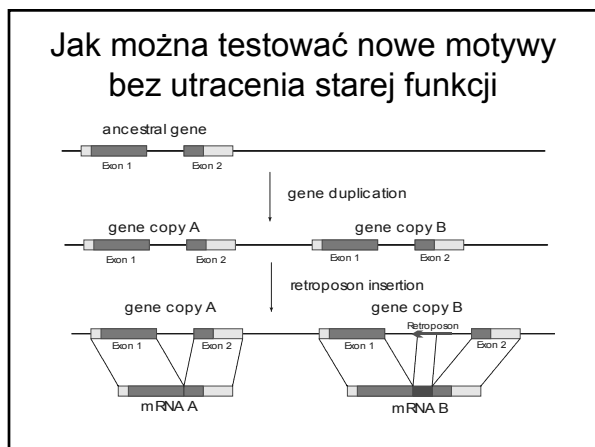
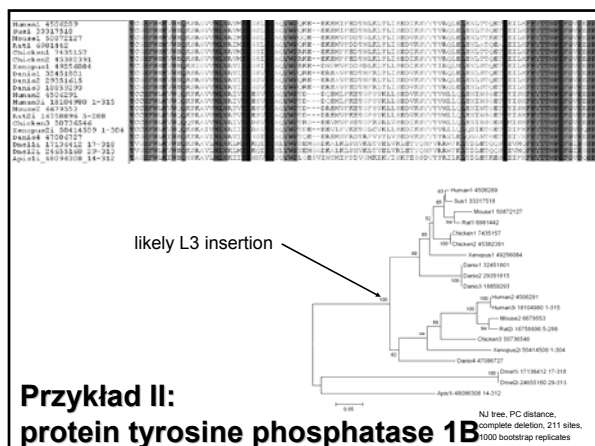
Przykład I: struktura III-rzędowa



Alternatywne białko nie tworzy zwartej struktury – kaseta Alu najprawdopodobniej burzy natywną strukturę funkcjonalnego białka. Tak więc ta forma jest albo nieaktywna albo posiada nową funkcję

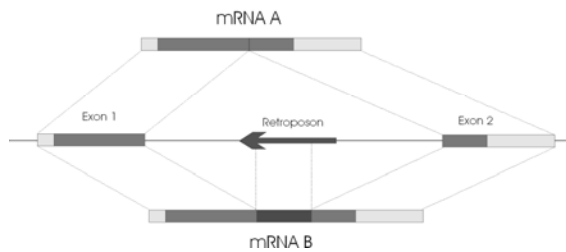
**Przykład II:
protein tyrosine phosphatase 1B**





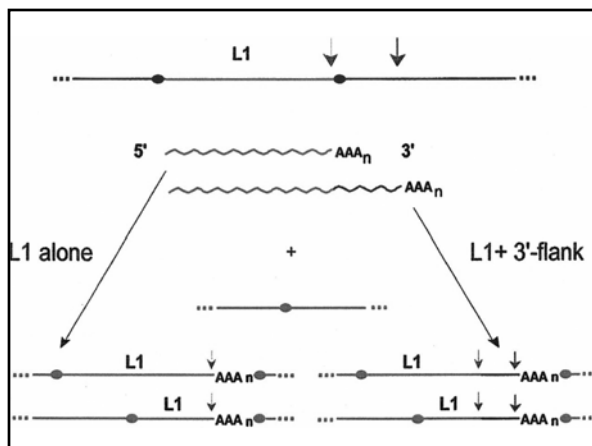
Jak można testować nowe motywy bez utracenia starej funkcji

Eksonifikacja elementów Alu, MIR i wielu sekwencji LTR



Tasowanie genomów

Transdukcja genomowego DNA przeprowadzana przez elementy L1



Transdukcja genomowego DNA przeprowadzana przez elementy L1 – statystyka

- f liczba elementów L1 w genomie człowieka 72,148
- f liczba elementów L1 które spowodowały transdukcję DNA o długości conajmniej 30 nt - 6,178
- f Częstość transdukcji - 8.6%
- f Najdłuższa ztransdukowany odcinek DNA - 2883 nt
- f Liczba eksonów w ztransdukowanych odcinkach - 0
