

B03 - Stavba a genetický význam informačních biomakromolekul. Centrální dogma molekulární biologie.

Lidský genom a typy sekvencí v něm obsažených. Mutace a polymorfismy. Příklady chorob způsobených mutacemi. Základní nástroje molekulární genetiky. (Genetika)

1 Stavba a genetický význam informačních biomakromolekul.



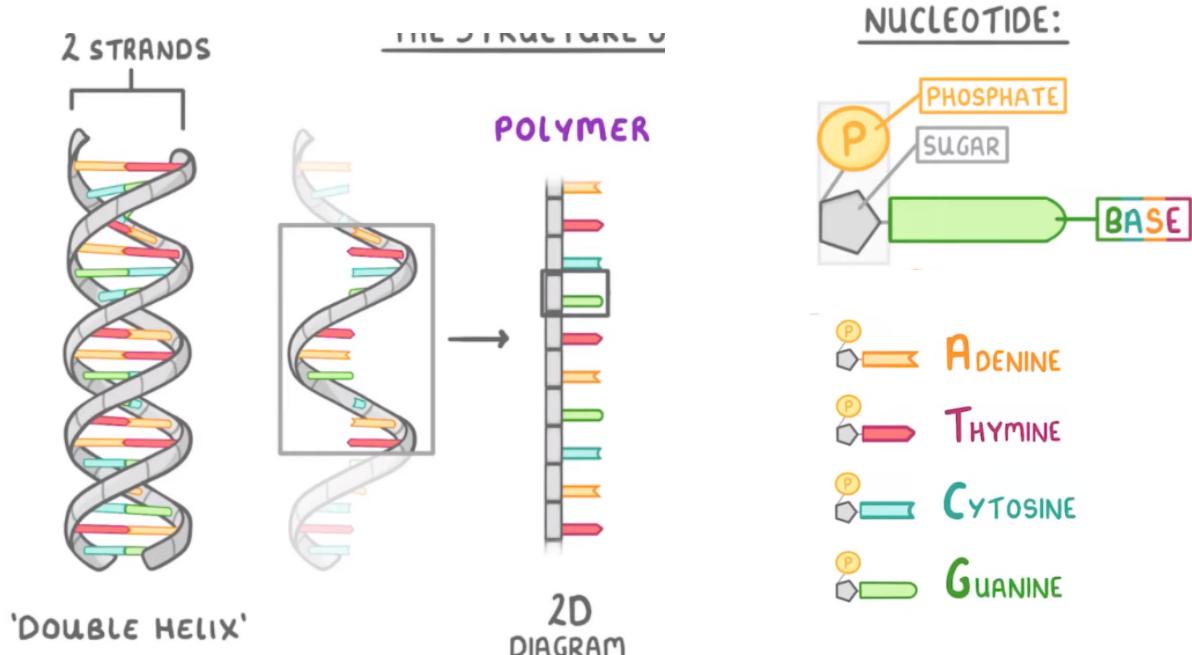
DNA

[GCSE Biology - What is DNA? \(Structure and Function of DNA\) #79 - YouTube](#)

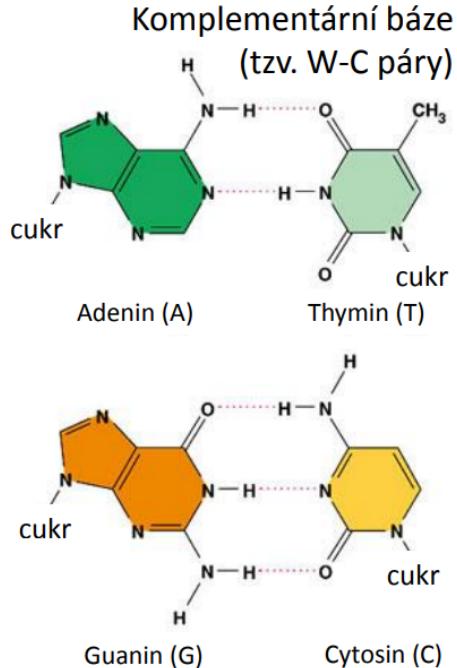
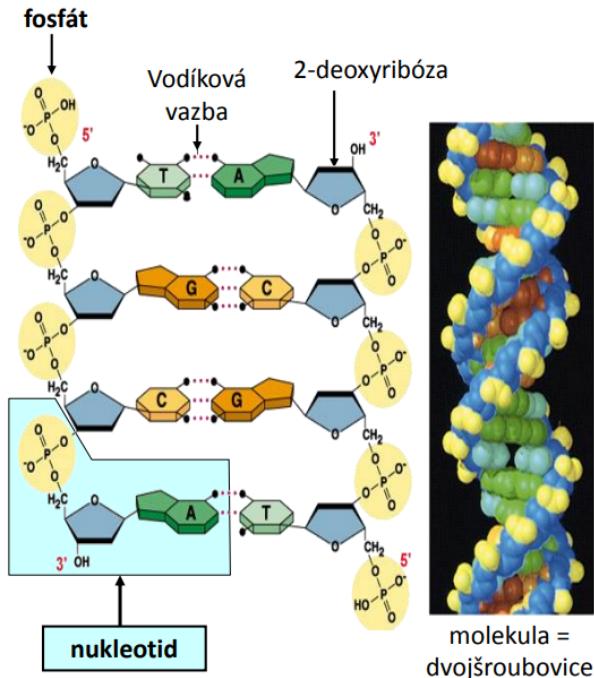
DNA je tvořena ze dvou vláken, která jsou "omotaná" kolem sebe a tvoří strukturu, která se nazývá **Double Helix**. Vláko jako takové je potom **polynukleotidový cukrfosfátový řetězec** v němž jsou částice deoxyribozy navzájem propojeny fosfátovými skupinami.

DNA je **Polymer** který je tvořen z hromady monomerů ty se jmenují **nukleotidy**. Každý nukleotid se tvořen ze 3 částí a to:

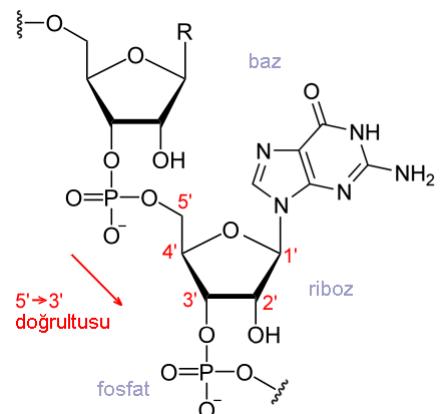
- Fosfátu
- Cukr (2 - deoxyriboza)
- Dusíkatá báze



Struktura DNA



Cukr (2 - deoxyriboza)



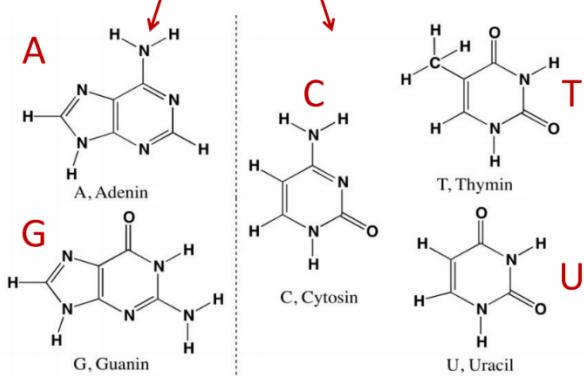
Dobré si všimnou, že směr 5->3 nebo 3->5 je odvozen Od číslování uhlíků v cukru

Dusíkaté báze

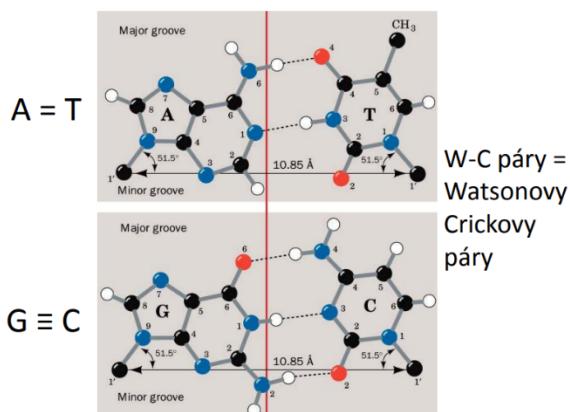
Báze jsou heterocyklické sloučeniny odvozené buď od purinu nebo pyrimidinu

Podle báze se také rozlišují nukleotidy na adeninové, guaninové, cytosinové a thyminové.

Báze = heterocyklické sloučeniny;
puriny a pyrimidiny



Komplementární párování bází



Nukleotidy jsou v DNA uspořádány za sebou v určitém pořadí neboli **sekvenci**. Každý **gen** je pak zapsán jako **charakteristická sekvence nukleotidů**. Sekvenci řetězce pak zapisujeme zkratkami bází jimiž jsou příslušné nukleotidy tvořeny.

Jak bylo uvedeno výše DNA je tvořena 2 polynukleotidovými řetězci, které jsou navzájem vodíkovými můstky mezi dusíkatými bázemi.

Adenin se vždy páruje s Thiminem (A-T) a cytosin s guaninem (C-G) proto jsou uvedené báze navzájem komplementární.

Polinukleotidové řetězce jsou navzájem antiparalelní to znamená, že jeden z nich je orientovaný ve směru 5-3 zatím co druhý ve směru 3-5.

Každá molekula obsahuje stejný počet purinových bází jako pirimidových.

Rychlé shrnutí

Nukleové kyseliny jsou heteropolymery. Základem je cukrfosfátový polynukleotidový řetězec. Nukleotidy jsou základní stavební jednotky molekul nukleových kyselin. Nukleotidy jsou tvořené z cukerné složky (u DNA 2-deoxyribóza), fosfátové skupiny a dusíkaté báze. Fosfát je zbytek kyseliny fosforečné. Báze jsou heterocyklické sloučeniny, pirminy(adenin,guanin) a primidiny (tymin,cytozin). Nukleotidy jsou vzájemně propojeny fosfodiesterovou vazbou.

Molekula DNA je tvořena dvěma polynukleotidovými řetězci. Řetězce jsou vůči sobě antiparalelní – jeden řetězec má směr 5'→3' a druhý 3'→5'. Na konci 3' je navázána -OH skupina, zatímco na 5' fosfátová skupina. Váží se spolu jen 2 specifické N-báze (vždy jedna pyrimidinová a + purinová) a to:

A-T (2 vodíkovými můstky), G-C (3 vodíkovými můstky).

Obě polynukletidová vlákna (primární struktura DNA) vytvářejí nejčastěji pravotočivou šroubovici označovanou jako double helix (sekundární struktura DNA), nejčastěji vyskytující se formou je B-forma DNA = pravotočivá.

Pořadí sekvence nukleotidů určuje DNA, každý gen má specifickou sekvenci. Sekvence se vždy zapisuje od 5' konce. Chargaffovo pravidlo – v DNA je stejný počet purinových a pirimidinových bázi

Replikace

[Replikace DNA - YouTube](#)

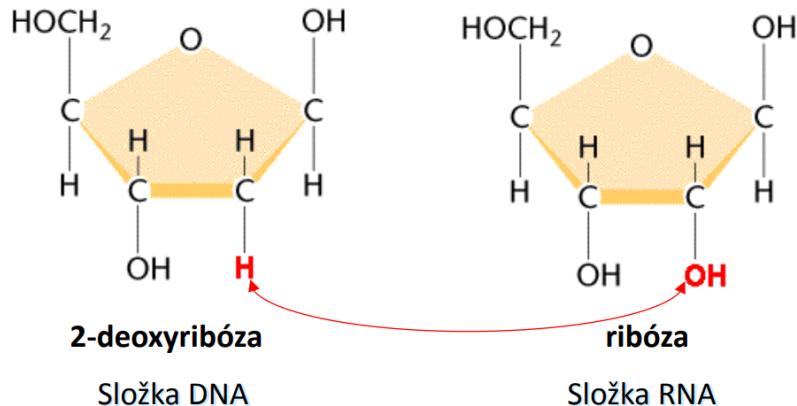
[DNA Replication - Leading Strand vs Lagging Strand & Okazaki Fragments - YouTube](#)

RNA

Nezbytná složka proteosyntézy. (proteosyntéza je základní proces, jímž se informace obsažená v DNA převede do podoby konkrétního znaku).

RNA je tvořena podobně jako DNA **polynukleotidovým řetězcem** ale místo deoxyribozy obsahuje **ribozu**. Jejíž jednotky jsou stejně jako u DNA propojeny přes fosfátové skupiny.

Cukr: 2-deoxyribóza × ribóza



od DNA se RNA odlišuje také tím, že místo pyrimidové báze Uracil místo Thiminu. Ten stejně jako thymin tvoří komplementární bázi k adeninu (A-U) ostatní báze zůstávají stejné.

RNA má jednořetězcové lineární struktury v některých strukturách se ale mohou uplatňovat dvouřetězcové struktury např.: tRNA

Proteosyntéza

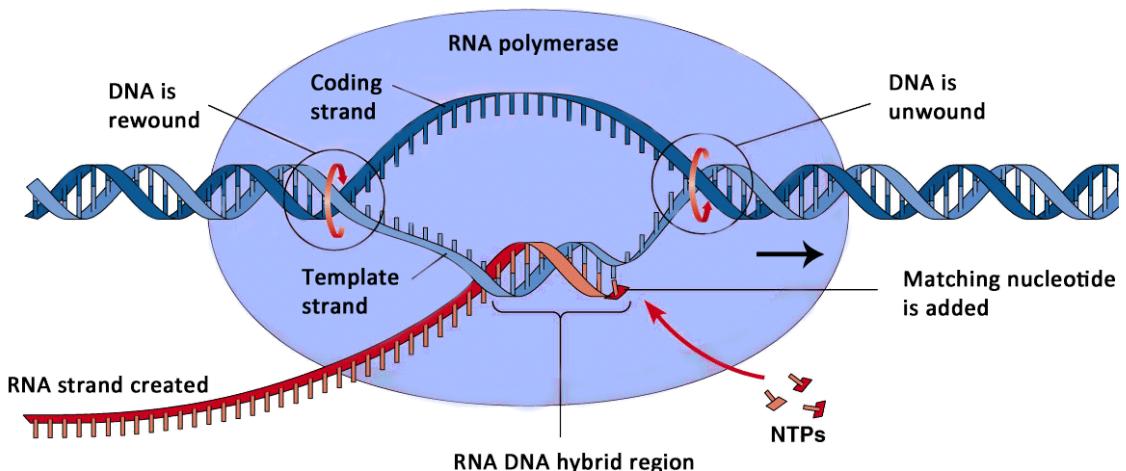
Transkripce – první fáze proteosyntézy

V této fázi se sekvence nukleotidů přepisuje do mediatorové RNA (mRNA)

Molekula mRNA je lineární. syntéza probíhá v buněčném jádře a katalyzuje ji enzym zvaný RNA polymeráza. Matricí pro tvorbu mRNA je jeden z řetězců DNA. DNA řetězce označujeme jako pozitivní (+) a negativní (-) (templátový). Negativní řetězec je ten co se kopíruje. V místě začátku genu nasedá RNA polymeráza na templátové vlákno DNA a k jeho deoxyribonukleotidům přiřazuje komplementární ribonukleotidy. Z nich

se vytváří nový řetězec

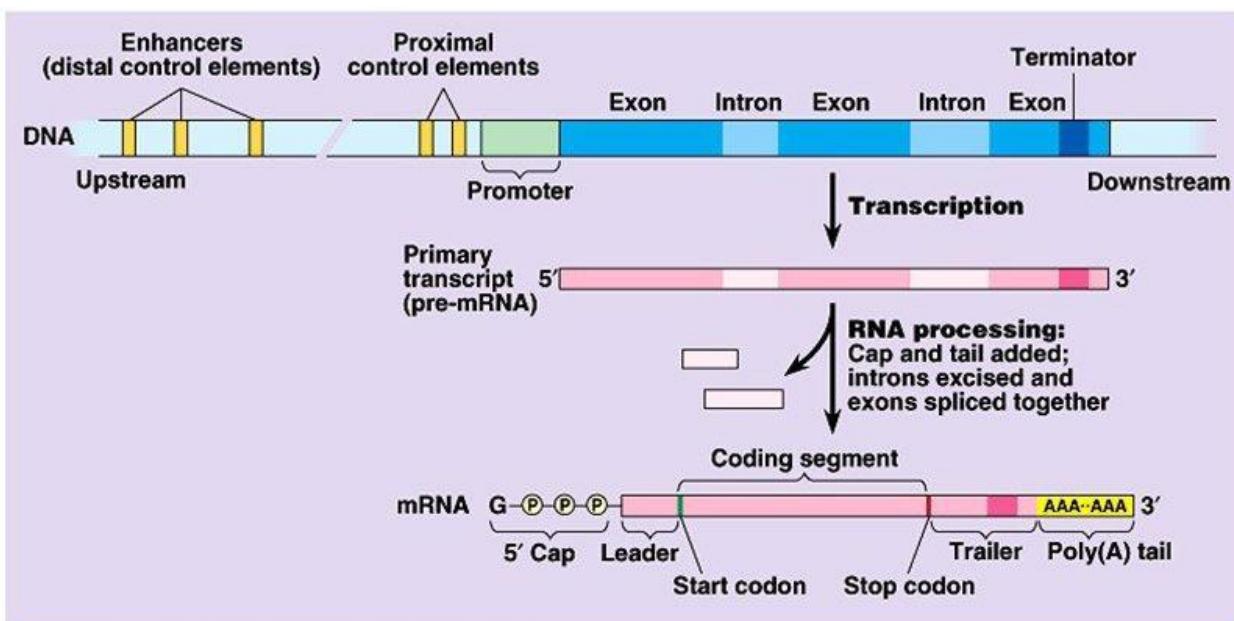
mRNA. Jako zdroje ribonukleotidů slouží molekuly nukleotidtrifosfatu (ATP)



Kde začíná a kde kočí transkripce

Gen na úrovni DNA je ohraničen regulačními oblastmi. Počáteční úsek je promotor. Ten není přepisován ale nasedá na něj RNA polymeráza za ním leží vlastní transkripční jednotka, která obsahuje informaci k syntéze proteinu a je zakončena terminátorem. V tomto bodě se RNA polymeráza odpojí. Terminátor zpravidla přechází v terminační kodon.

Nejprve dojde k transkripci genu. Jestli je gen jednoduchý (převážně bakterie) tak jeho transkripcí rovnou vznikne mRNA, když je ale gen složitý vznikne (u eukaryí) nejdříve heterogenní RNA(hnRNA). z ní se nejdříve při „sestřihu“ (splitingu) odštěpí pomocí snRNA a dalších látok, intronové sekvence. Zbylé části (exony) se pak spojí a vytvoří mRNA.



Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

Translace – překlad

V cytoplazmě buňky se k molekule mRNA připojují ribozomy. Tyto útvary jsou tvořeny molekulami ribozomové RNA (rRNA) a specifickými proteiny. Ribozom nasedá na 5` konec řetězce mRNA, který odpovídá počátku příslušného genu v DNA. Od tohoto místa probíhá syntéza proteinové molekuly podle matrice mRNA.

Informace z DNA přepsaná do mRNA udává pořadí aminokyselin v proteinovém řetězci.

Jednotlivé aminokyseliny jsou do ribozomu transportovány pomocí transferové RNA (tRNA)

Molekula tRNA má složitou stavbu. Části komplementární nukleotidy jsou místy propojeny vodíkovými můstky. Sekundární struktura tRNA připomíná jetelový trojlístek. Při proteosyntéze se uplatňuje akceptorové raménko tRNA. to obsahuje 3 konec polynukleotidového řetězce a na ten se před vstupem do ribozomu naváže příslušná aminokyselina.

Další důležitá část tRNA je antikodonové raménko s antikodonovou smyčkou. Ta obsahuje trojici nukleotidů označených, jako antikodon. Každý antikodon v tRNA je komplementární ke příslušné trojici v mRNA tato

trojice se pak nazývá kodon. Při nukleosyntéze se v ribozomu na chvíli mRNA a tRNA spojí na základě párování kodonů a antikodonů.

Proteosyntézy se účastní celkem 20 druhů tRNA to znamená dvacet druhů aminokyselin

Ribozom

Ribozom je složitý nukleoproteinový útvar skládající se z malé a velké podjednotky, které se při zahájení translace spojují. Nejprve se k malé jednotce naváže mRNA a pak se k nim připojí velká podjednotka. Uvnitř pak dochází ke kontaktu mRNA a tRNA a k následnému vytvoření peptidového řetězce.

Ribozom má několik vazebných míst:

aminoacylové místo (A-místo)

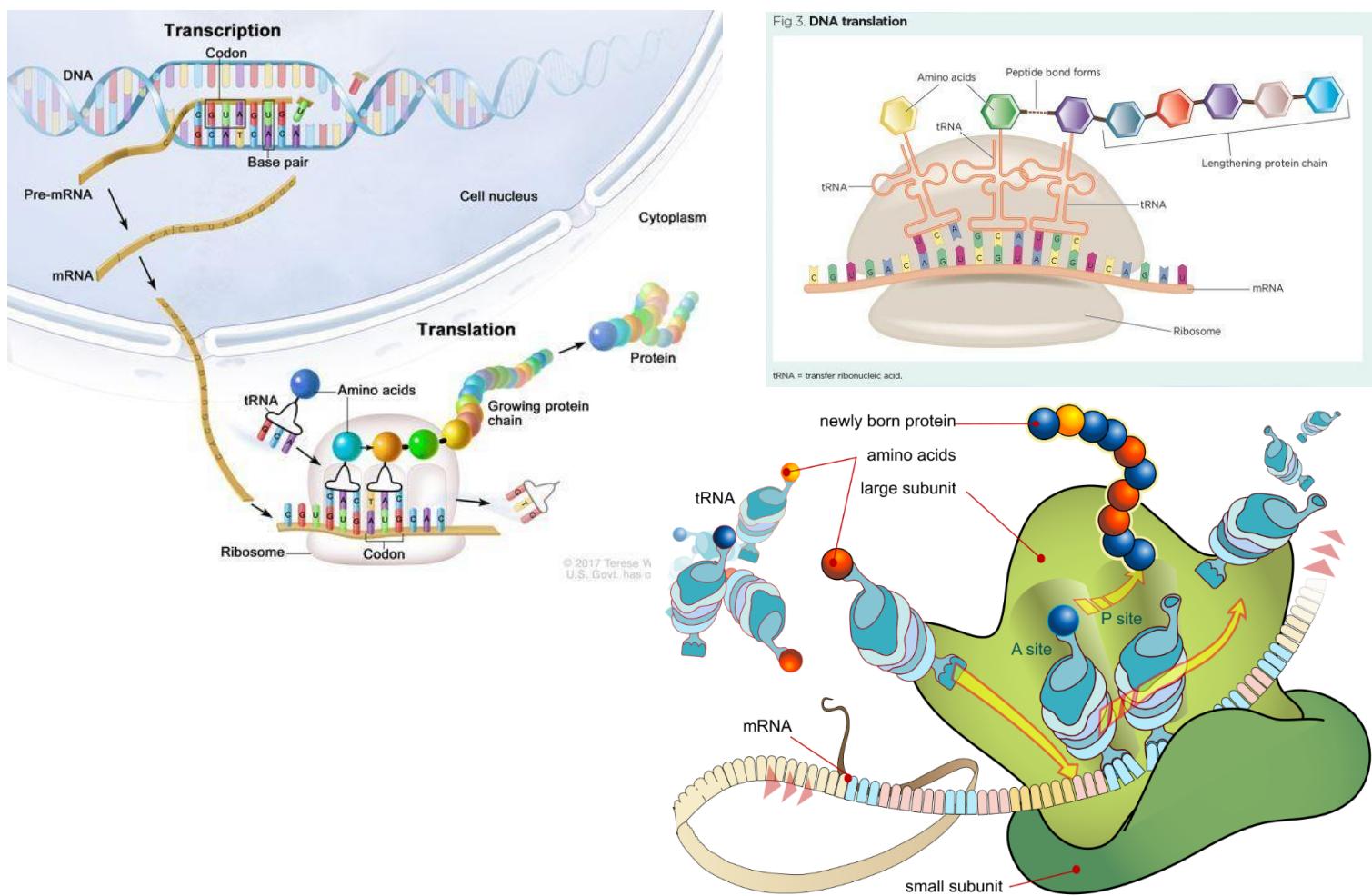
vstupuje tRNA s aminokyselinou přitom se tRNA spojuje s mRNA.

Peptidové místo (P-místo)

Obsahuje předchozí molekulu tRNA na níž je navázán už vzniklý peptidový řetězec

Peptidyltransferázové místo

dochází k připojení nové aminokyseliny ke stávajícímu peptidovému řetězci. tRNA ke které byl řetězec do teď připojen opouští ribozom a po připojení další aminokyseliny může být znova opět využita.



[Transcription and mRNA processing | Biomolecules | MCAT | Khan Academy - YouTube](#)

[How are Proteins Made? - Transcription and Translation Explained #80 - YouTube](#)

Sekvence mRNA je čtena po kodonech (tripletech). Jednotlivé tripletey určují zařazení konkrétních aminokyselin a jejich sekvence udává pořadí aminokyselin v polypeptidovém řetězci. tento systém přiřazení dané aminokyseliny k určitému tripletu se nazývá genetický kod.

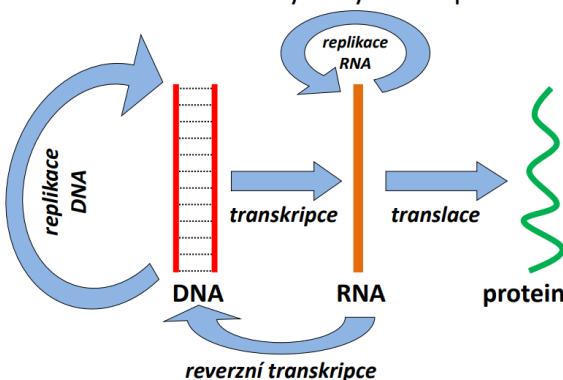
Triplety kodují aminokyseliny některé triplety ačkoliv odlišné mohou kodovat stejnou aminokyselinu to se nazývá degenerace genetického kodu(nejde o úpadek jen to znamená, že každou aminokyselinu jde kodovat vice než jedním způsobem), některé trojce kodují pro konec řetězce(stop kodon) po jeho přečtení se polypeptidový řetězec odpojí.

Translace začíná metioninem.

Ústřední dogma molekulární biologie říká:

Ústřední (centrální) dogma molekulární biologie

postuluje přenos genetické informace mezi molekulami nukleových kyselin a proteinů



Genetická informace se přenáší z DNA do RNA a odtud do molekuly proteinu. Opačný přenos z proteinu do molekuly nukleové kyseliny není možný.

Jde o tok informace

DNA<->DNA<->RNA->protein

Aminokyseliny a proteiny

Aminokyselina, molekula obsahující COOH (karboxylovou funkční skupinu) a NH₂ (aminovou funkční skupinu), při proteosyntéze se uplatňuje jen 20 aminokyselin(alanine, valine, leucine, izoleucine, proline, phenylalanine, tryptofan, methionine, glycine, serine, threonine, tyrosine, asparagine, glutamine, cysteine, selenocysteine, kyselina asparagová, kyselina glutamová, arginine, lysine, histidine). Biologické vlastnosti proteinů jsou dány druhem aminokyselin, jejich pořadím a jejich vzájemnými prostorovými vztahy. Aminokyseliny v proteinech jsou spojeny peptidovou vazbou. Peptidová vazba spojuje jednoduchou kovalentní vazbou aminoskupinu jedné aminokyseliny a karboxylovou skupinu druhé aminokyseliny. Polykondenzací vzniká libovolně dlouhý řetězec aminokyselin. Konec řetězce, který má volnou aminoskupinu, se nazývá N-konec. Na opačné straně řetězce nalezneme naopak volnou karboxylovou skupinu, tento konec se nazývá C-konec.

Proteiny jsou organické makromolekuly, polypeptidový řetězec. Základní stavební kameny jsou aminokyseliny. Typický protein jich obsahuje 200-300.

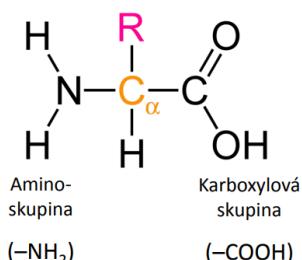
Struktura proteinů vychází z uspořádání aminokyselin v řetězci. Struktura proteinů je velmi důležitá pro jejich funkci. Primární struktura je definována přesným pořadím aminokyselin v řetězci. Sekundární strukturou rozumíme prostorové uspořádání aminokyselin v řetězci a stabilizace vodíkovými můstky. Existují dvě základní sekundární struktury: alfa-helix, řetězec je stočen do pravotočivé šroubovice; beta-skládaný list, dva rovnoběžně a antiparalelně uspořádané řetězce připomínající složený list papíru. Tertiární strukturu charakterizují další intramolekulární vazebné interakce. Např. disulfidické můstky, inotové vazby a van der Waalsovy síly. Kvarterní struktura vzniká u proteinů, které se skládají ze dvou a více polypeptidových řetězců.

Zápis sekvence proteinů, uvádíme jednotlivé aminokyseliny. Postupujeme od N-konce k C-konci. Pro zadávání sekvence do weboých aplikací, resp. databázi používáme jednopísmenné kódy aminokyselin stanovené IUPAC.

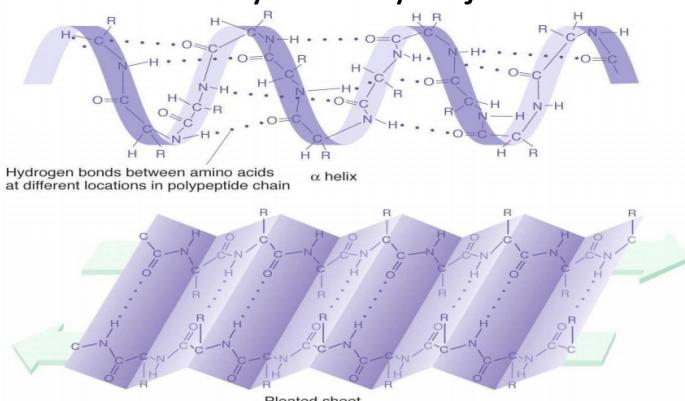
Proteiny

- Makromolekuly – polypeptidový řetězec
 - Základní stavební kameny = aminokyseliny

Známe 20 – 22 proteinogenních aminokyselin

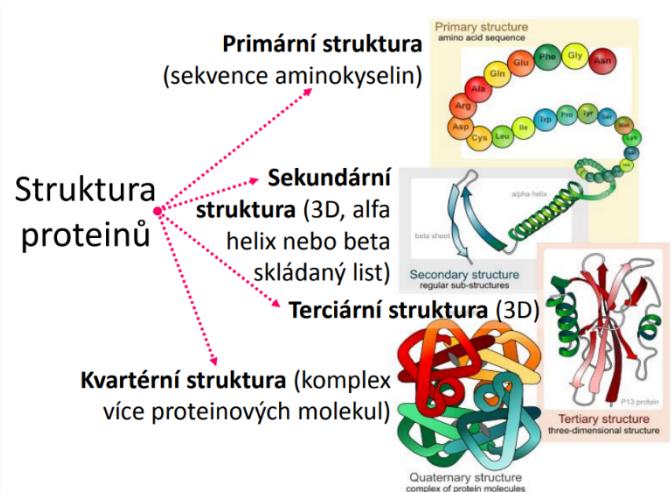


Struktury alfa-helix a beta-skládaný list
jsou udržovány vodíkovými vazbami
mezi aminokyselinovými jednotkami



Příklady zápisu sekvence proteinu

- Starý zápis:
Gly Cys Ala Thr Arg Tyr Asn Phe
 - Stejný zápis jednopísmenným kódem IUPAC:
GCATRYNFG
 - Pozor při zadávání do databází a programů:
jednopísmenné kódy jsou sice stručnější, ale nelze je dobře odlišit od sekvence nukleových kyselin. Je třeba specifické označení
 - Např. tzv. FASTA formát používá přípon „.aa“ (aminokyseliny, sekvence proteinů) a „.nt“ (nukleotidy, sekvence nukleových kyselin)
 - Jinde je údaj o typu sekvence součástí informace uvozující zápis sekvence.



Zápis sekvence proteinu

- Uvádíme jednotlivé aminokyseliny
 - Postupujeme od N-konce k C-konci (neboť protein je v tomto směru syntetizován)
 - Pro zadávání sekvence do webových aplikací, resp. databází používáme jednopísmenné kódy aminokyselin stanovené IUPAC (mezinárodní unie pro čistou a aplikovanou chemii)

Mutace a polymorfizmy příklady chorob jimi způsobené

[Mutace \(cvut.cz\)](http://Mutace.(cvut.cz))

[GCSE Biology - What are DNA Mutations? #81 - YouTube](https://www.youtube.com/watch?v=JyfjwvXWzqU)

Lidský genom a typy sekvencí v něm obsažených

Gen je úsek DNA kódující syntézu určitého genového produktu, tedy: proteinu a funkční RNA (která již nepodléhá translaci). Název "gen" zavedl v roce 1905 dánský genetik Wilhelm Johannsen. Délku genů, vzdálenost genů i jiných úseků DNA většinou vyjadřujeme v párech bází "base pairs" neboli bp.

Genom je soubor veškeré genetické informace konkrétního organismu. Tato informace je zapsána v DNA (pouze u některých virů v RNA). Obecně pod tento pojem zahrnujeme kódující i nekódující sekvence DNA. Genom řady organismů byl již v současné době kompletně osekvenován. Human Genome Project, zahájen r. 1990. Zúčastnilo se ho 20 univerzit a výzkumných pracovišť. K urychlení přispěla: automatizace analýz DNA, pokrok v informačních technologiích, kompetence se soukromou společností Celera Genomics, která si stanovila stejný cíl. R. 2003 – publikována finální verze sekvence lidského genomu.

V případě eukaryotních organismů lze rozlišovat jaderný genom (genomická DNA, tvořen lineárními chromozomy, obsahuje většinu genetické informace) a mimojaderný genom (tvořený DNA sekvencemi v semiautonomních organelách – mitochondriích či plastidech). U člověka v této souvislosti musíme uvažovat mitochondriální genom.

Složení lidského genomu. Jedinečné genomové sekvence, zejména geny kódující proteiny. Řada genů tvoří genové rodiny: skupiny sekvencně podobných genů (genové komplexy) vzniklé během evoluce opakovánou duplikací jednoho genu. Geny kódující protilátky (imunoglobuliny), histokompatibilní geny (main histocompatibility complex) kódují tzv. transplantační antigeny (u člověka Human Leukocyte Antigens). Pseudogeny nefunkční geny, „molekulární fosilie“. Možnost vzniku pseudogenu bud' duplikace a následná mutace vedoucí ke ztátě funkce; nebo reverzní transkripce z mRNA, během níž dojde k mutaci. Repetitivní sekvence, tandemové (jednotky tvořící repetece následují za sebou, tvoří "shluhy"), rozptýlené (interspersed repeats; na různých místech genomu).

Satelitní DNA. Satelity, resp. makrosatelity – délka jednotky řádově stovky bp - příklad alfa-satelity v centromerách chromozomů (součást trvale neaktivního – konstitutivního heterochromatinu). Minisatellity – délka jednotky řádově desítky bp – sekvence VNTR (Variable Number of Tandem Repeats).

Mikrosatellity – délka jednotky zpravidla 2-6 bp. Sekvence STR (Short tandem repeats). Nejčastěji v oblasti konců chromozomů – telomer (u člověka repetece TTAGGG). Menší část také v rámci genů (trinukleotidové a hexanukleotidové repetece).

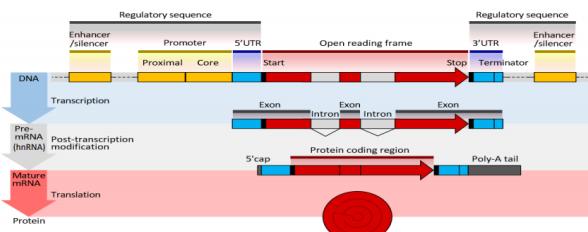
Lidský genom obsahuje značný podíl transpozonů. Transpozony – transpozabilní elementy "jumping genes". Sekvence, které mění pozici v genomu. Objeveny ve letech 1944-1950 americkou genetičkou Barbarou McClintockovou. Celkem asi 44% lidského genomu. Možná virový původ (apoň u některých). Retrotranspozony – amplifikují se pomocí reverzní transkripce. Reverzní transkripce je zpětný přepis z RNA do DNA. Syntéza cDNA (complementary DNA) podle matrice RNA, katalyzovaná reverzní transkriptázou. Poprvé prokázána u retrovirů, ale probíhá i u eukaryot. LTR – retrotranspozony, ohrazeny z obou stran sekvencemi LTR (Long Terminal Repeats). Patrně vznikly z retrovirů, zpravidla však již netvoří virové částice. Rozmnožují se stejně jako retroviry (RNA → cDNA → integrace na jiné místo genomu). Long interspersed nuclear elements (LINEs) – non-LTR retrotranspozony. 4-16 20-21% lidského genomu. Transkribovány do mRNA, která je překládána do reverzní transkriptázy. Reverzní transkriptáza vytváří podle mRNA kopie DNA, které se integrují na různá místa v genomu. Short interspersed nuclear elements (SINEs) – non-LTR retrotranspozony, které nekódují žádný protein. Přepisovány RNA-polymerázou III – vzniklá RNA je reverzně přepsána do DNA, která se může integrovat na různá místa v genomu. Příklad rodina Alu sekvení (=nejpočetnější transpozony v lidském genomu – asi 1 000 000 kopií).

Gen

- Úsek DNA (specifická sekvence nukleotidů) kódující syntézu určitého genového produktu, tedy:
 - proteinu
 - funkční RNA (která již nepodléhá translaci)
- Název „gen“ zavedl v roce 1905 dánský genetik Wilhelm Johannsen (předtím používán pro materiální nosič dědičných znaků pojmem „*pangen*“).



Wilhelm
Johannsen



Transkriptom

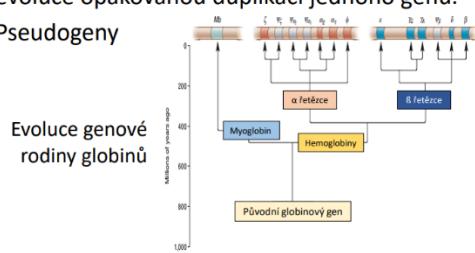
- Soubor všech molekul RNA přítomných v dané buňce, popř. tkáni
- Transkriptomika: sleduje na základě produkce RNA expresi příslušných **genových populací** v daných buňkách, popř. též rozdíl v genové exprese mezi jednotlivými buňkami nebo tkáněmi

Proteom

- Soubor všech proteinů, který se v daném čase a v daných podmínkách nachází v určité buňce; resp. buněčné populaci.
- Je komplexní a velmi dynamický, jeho skladba se neustále mění.
- Liší se od buňky k buňce a plynule se mění v souvislosti s biochemickými interakcemi s genomem a prostředím.
- Proteomika** se zabývá systematickou analýzou proteomu, tj. proteinů z hlediska jejich identity, množství a funkcí.

Jedinečné genomové sekvence

- Zejména geny kódující proteiny
- Řada genů tvoří **genové rodiny**: skupiny sekvenčně podobných genů (genové komplexy) vzniklé během evoluce opakovánou duplikací jednoho genu.
- Pseudogeny

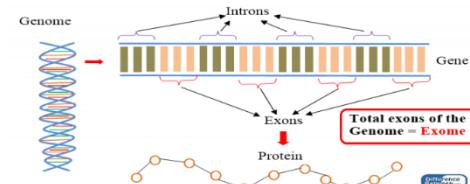


Evoluce genové rodiny globinů



Exom

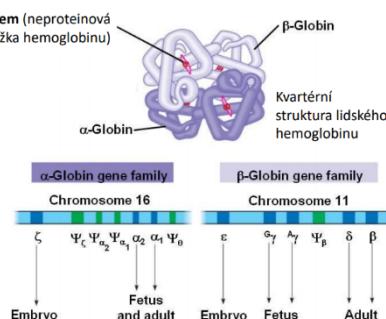
- Všechny sekvence DNA tvořící exony protein kódujících genů v dané buňce nebo organismu
 - Pozor! Exom není totožný s transkriptorem!
- Ne všechny RNA slouží k syntéze proteinů!
➤ Transkriptom se zpravidla týká příslušné buněčné populace, kde nemusí být nutně transkribovány všechny geny.



Další příklady genových rodin

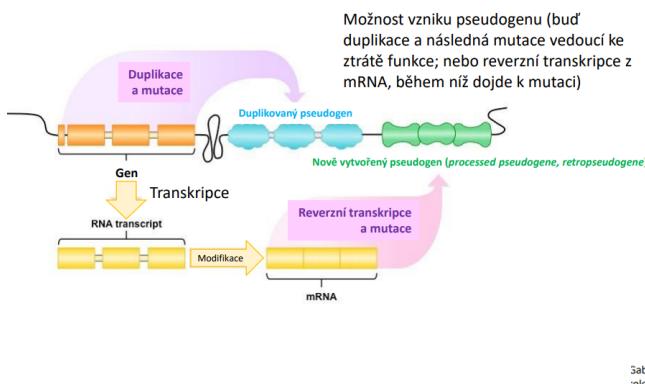
- Geny kódující protilátky (imunoglobuliny)
- Histokompatibilitní geny (*Main histocompatibility complex*, MHC) – kódují tzv. **transplantační antigeny** (u člověka *Human Leukocyte Antigens*, HLA)

Hemoglobinové podjednotky



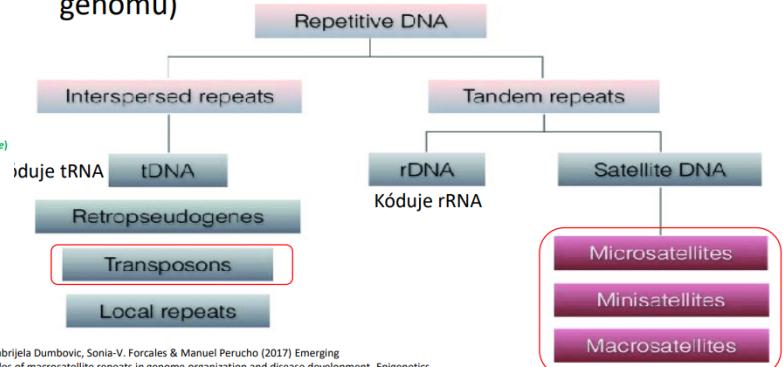
Pseudogeny

- Nefunkční geny, tzv. „molekulární fosilie“



Repetitivní sekvence (repetice, repeats)

- Tandemové (*tandem repeats*; jednotky tvořící repetice následují za sebou, tvoří „shluky“)
- Rozptylené (*interspersed repeats*; na různých místech genomu)



Satelitní DNA

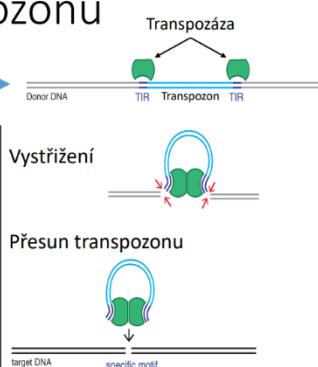
- Sateliity, resp. makrosateliity** – délka jednotky řádově stovky bp
 - Příklad: alfa-sateliity v centromerách chromozomů (součást trvale neaktivního – konstitutivního heterochromatinu)
- Minisateliity** – délka jednotky řádově desítky bp
 - Sekvence VNTR (Variable Number of Tandem Repeats)
- Mikrosateliity** – délka jednotky zpravidla 2-6 bp
 - Sekvence STR (Short tandem Repeats)
 - Nejčastěji v oblasti konců chromozomů – tzv. **telomer** (u člověka repetice TTAGGG)
 - Menší část také v rámci genů (trinukleotidové a hexanukleotidové repetice)

Typy transpozonů

DNA transpozony

- Retrotranspozony – amplifikují se pomocí **reverzní transkripce**

• V lidském genomu jsou daleko častější



Transpozony

- Transpozabilní elementy, „jumping genes“
- Sekvence, které mění pozici v genomu
- Objeveny v letech 1944-1950 americkou genetičkou Barbarou McClintockovou
- Celkem asi 44% lidského genomu
- Možná virový původ (aspoň u některých).

LTR-retrotranspozony

- Ohraničeny z obou stran sekvencemi LTR (Long Terminal Repeats)
- Patrně vznikly z retrovirů, zpravidla však již netvoří virové částice.
 - Endogenní retroviry
- Rozmnožují se stejně jako retroviry (RNA → cDNA → integrace na jiné místo genomu)

LTR-retrotransposon (5-20 kb)

