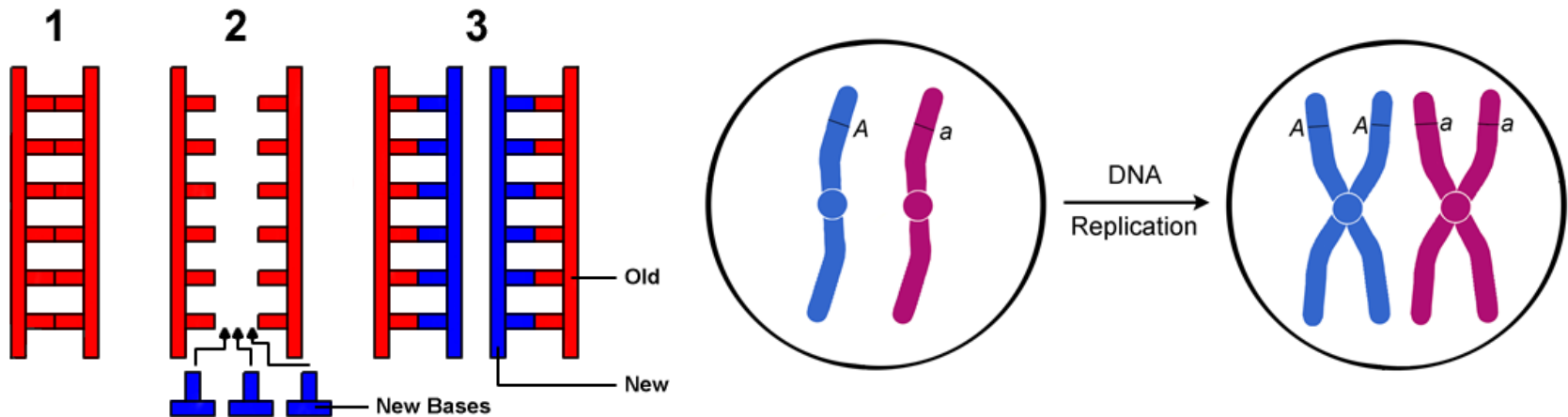


CURS 2

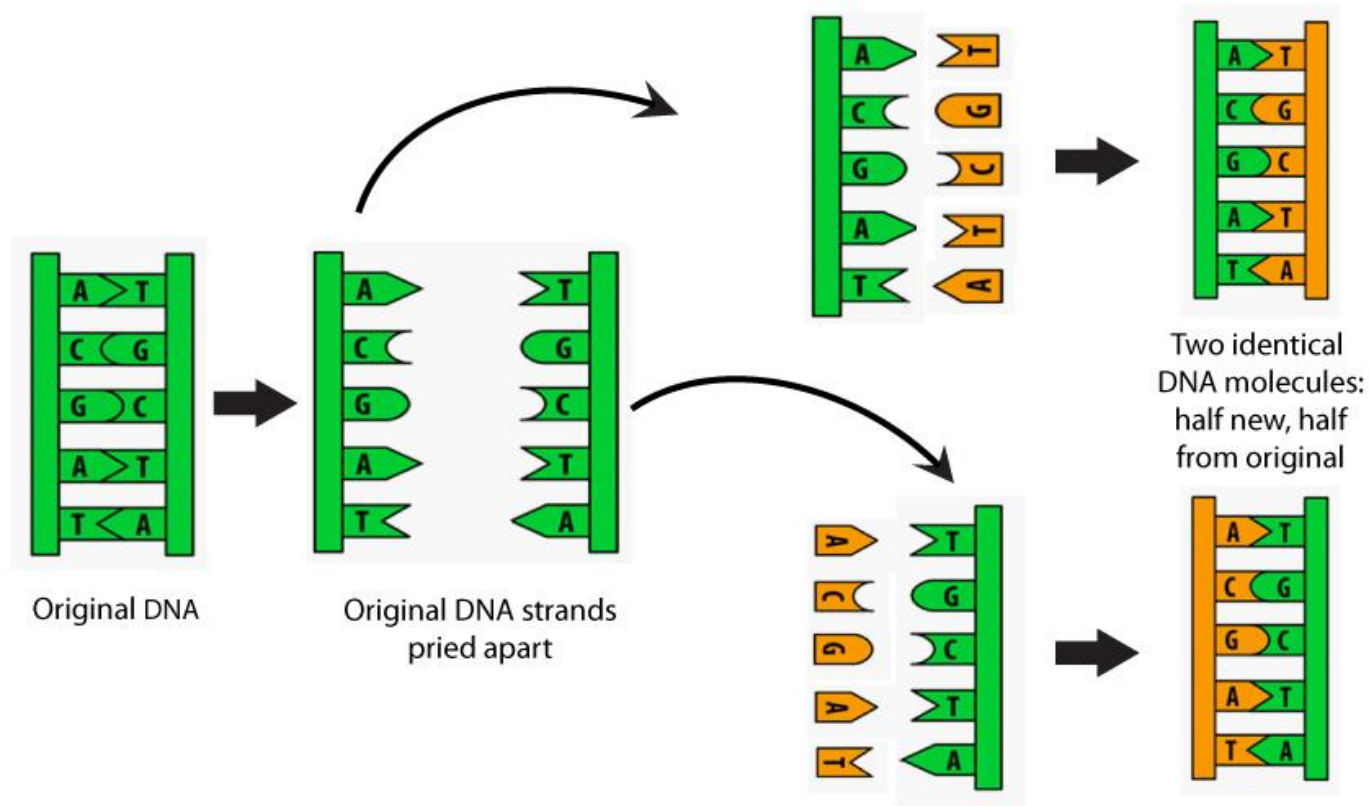
- **REPLICAREA ADN**
- **GENOMUL UMAN**
- **STRUCTURA GENEI**

Replicarea ADN

Replicarea ADN este un proces fundamental ce stă la baza transmiterii informației genetice (ereditare), prin care informația genetică este copiată identic și transmisă prin diviziunea celulară, în succesiunea generațiilor de celule.



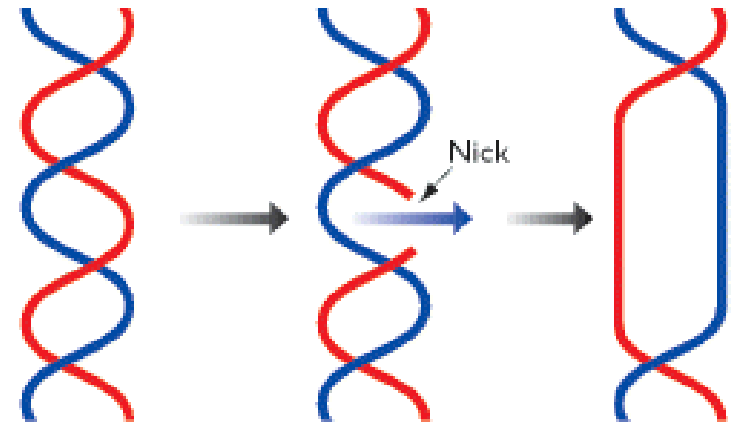
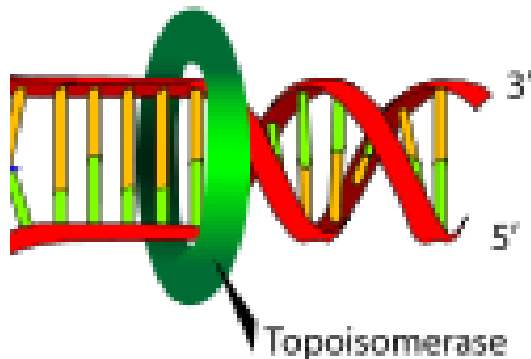
Prin procesul de replicare, dintr-o moleculă parentală de ADN bicatenar se vor forma 2 molecule de ADN identice între ele și cu moleculă parentală, fiecare din moleculele nou formate fiind alcătuită dintr-o catenă nou sintetizată și o catenă veche (matriță) - *replicare semiconservativă*.



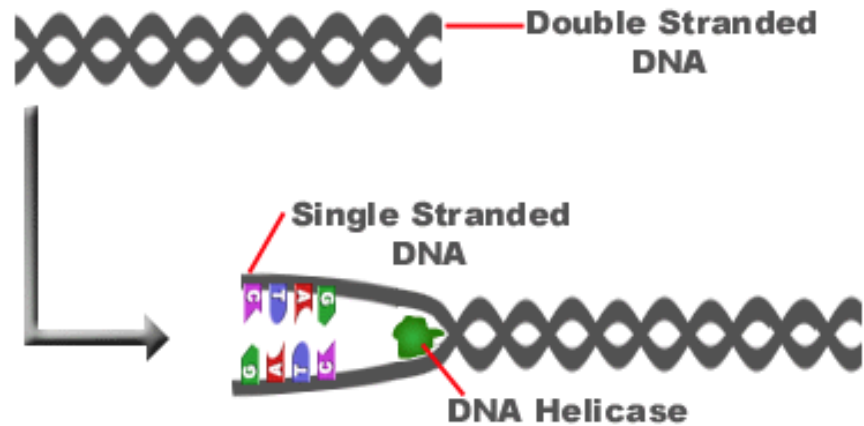
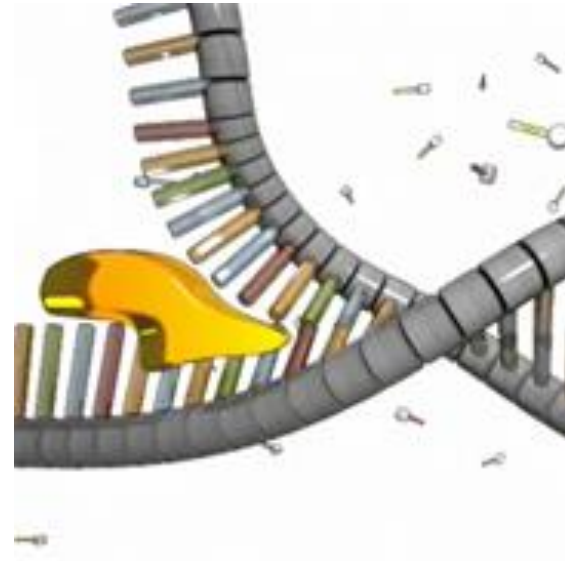
Procesul de replicare se realizează prin intermediul unui complex care cuprinde numeroase enzime și proteine.

Principalele componente ale aparatului replicativ sunt:

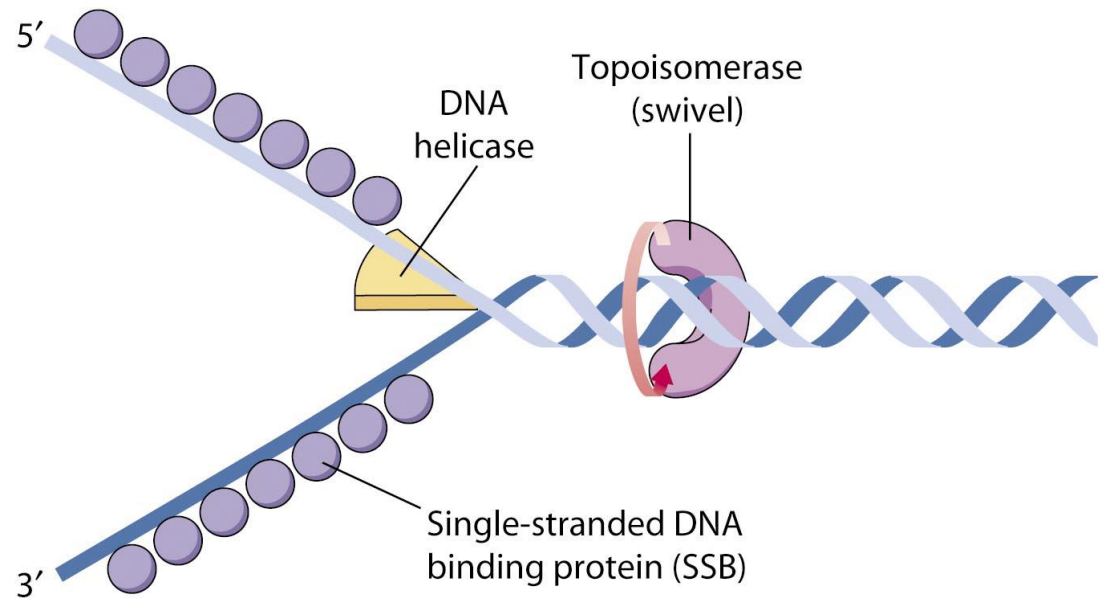
ADN topoisomeraza - secționează și despiralizează dublului helix, reducând astfel tensiunea din aval.



ADN helicaza - rupe legaturile de hidrogen între bazele azotate complementare și eliberează monocatenenele de ADN.



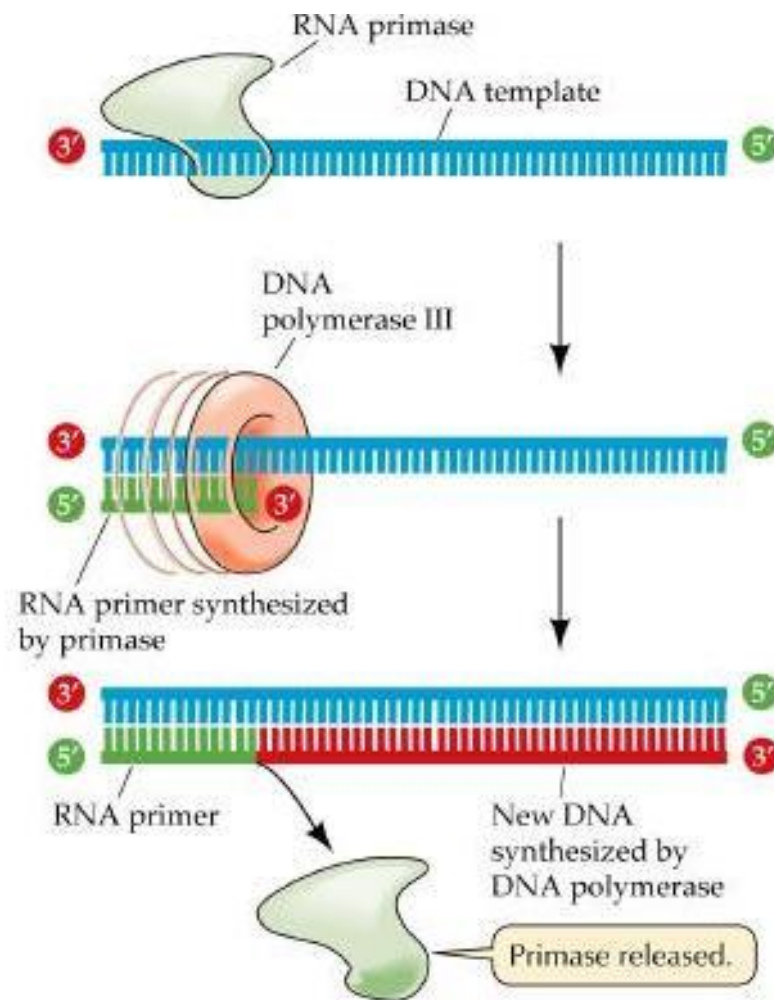
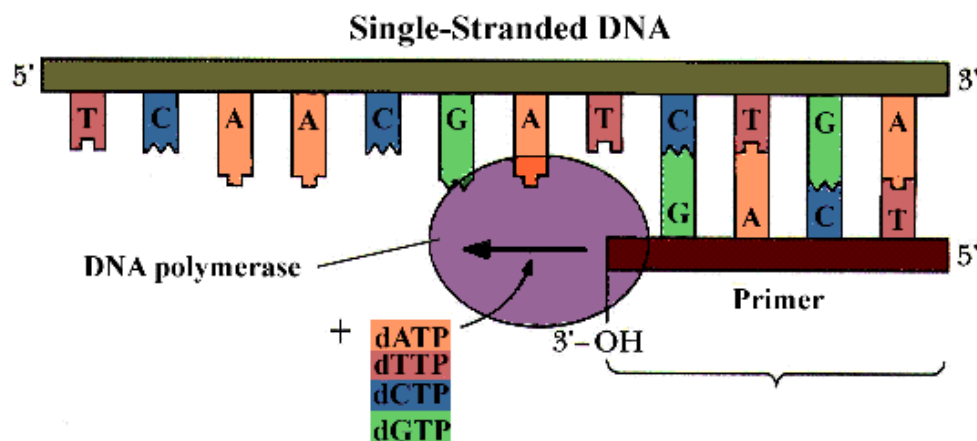
Proteinele RPA (SSB) - împiedică refacerea dublului helix prin atașarea la monocatenenele de ADN eliberate de helicaze



ARN primaza - sintetizează unui scurt fragment de ARN (10-15 nucleotide) denumit primer.

Fragmentul este extins de către polimeraza α până la 60 nucleotide.

Polimerazele nu știu să înceapă sinteza unui fragment nou de ADN și au nevoie de prezența primerului, ele pot doar să extindă o catenă, adăugând dezoxiribonucleotide activate la capatul 3'-OH.



ADN polimerazele alpha (α), delta (Δ) și epsilon (ϵ) – adaugă nucleotide complementare catenei matriță.

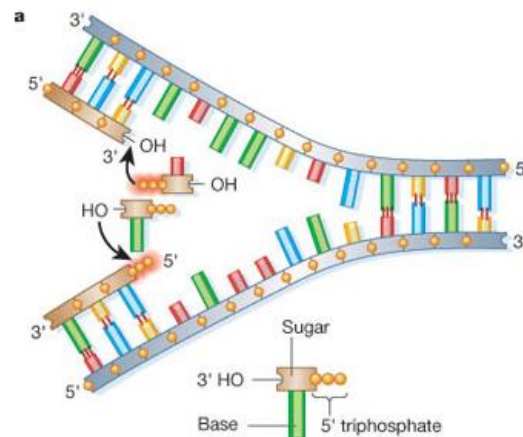
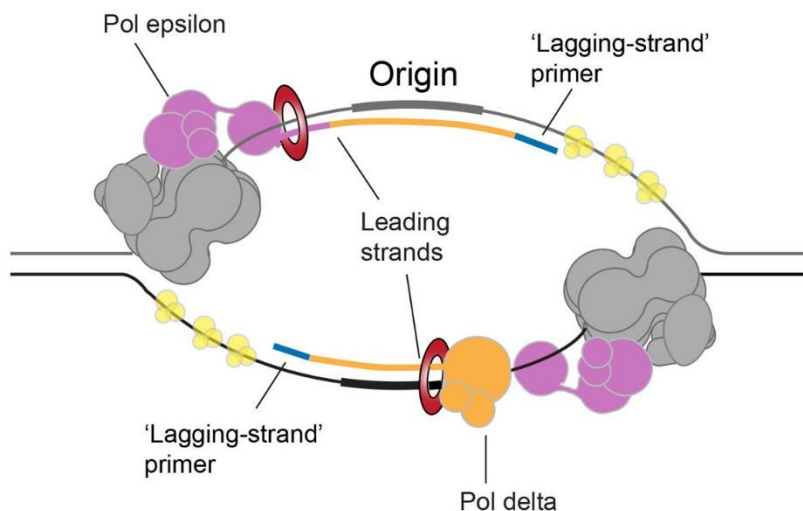
Polimeraza α extinde primerul până la 60 nucleotide.

Polimerazele δ și ϵ adaugă nucleotide complementare, sintetizând astfel noile catene.

ADN polimerazele δ și ϵ adaugă nucleotide numai în direcția 5'-3', încorporând nucleotide complementare catenei matriță.

La nivelul catenei matriță 3'-5', sinteza noii catene (5'-3') se realizează rapid și continuu de către polimeraza ϵ .

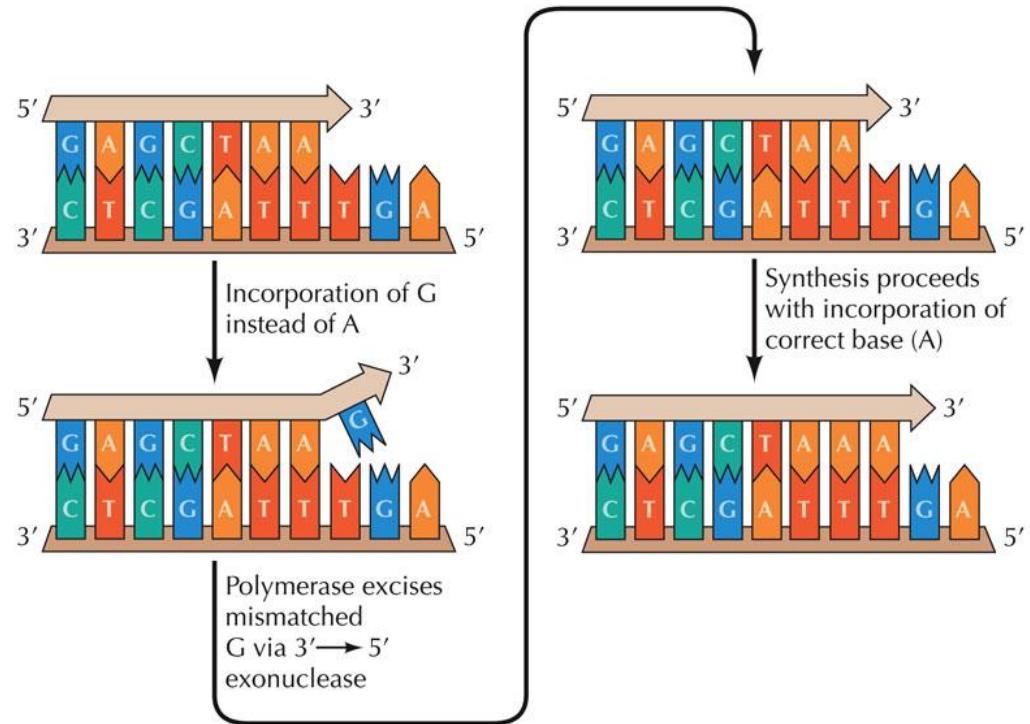
La nivelul catenei matriță 5'-3', sinteza se realizează lent, discontinuu, prin intermediul polimerazei δ care adaugă nucleotide în direcția 5'-3' sub forma unor fragmente de 100 până la 1000 nucleotide denumite fragmente Okazaki.



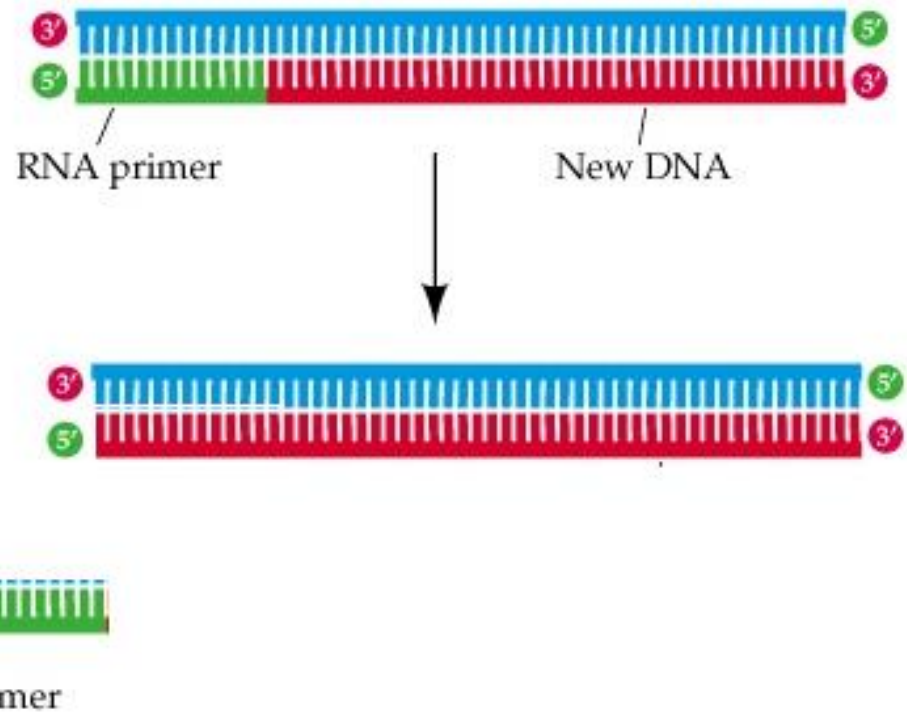
Polimerazele δ și ϵ operează cu mare fidelitate și au activitate de autocorecție.

Acestea recunosc bazele inserate greșit și datorită activității lor exonucleazice 3'-5', revin și corectează eroarea.

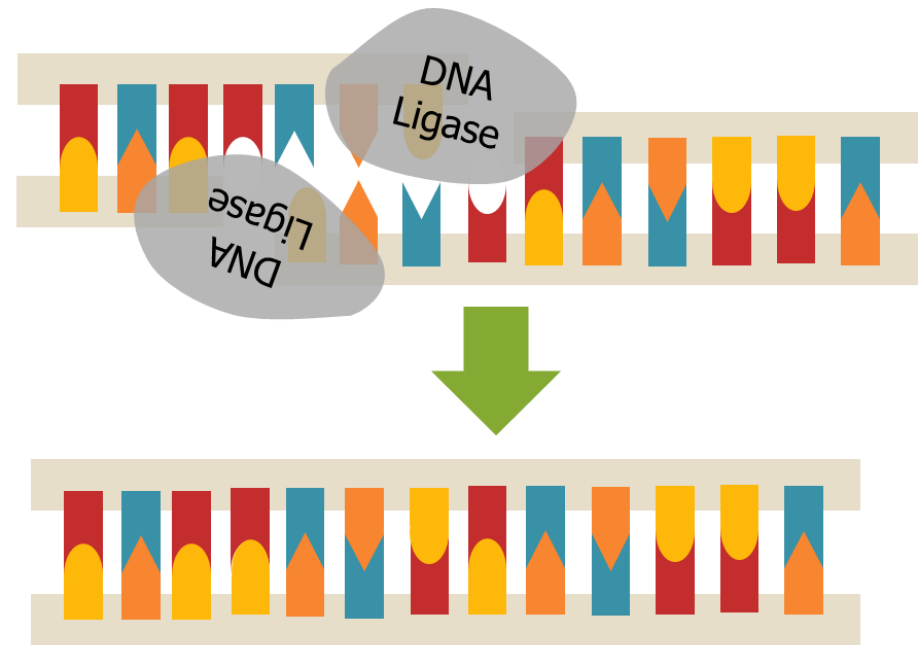
Astfel, polimerazele fac o eroare la un milion nucleotide. Prin intermediul mecanismelor de reparare a ADN, erorile se reduc la una la 1 miliard de nucleotide.

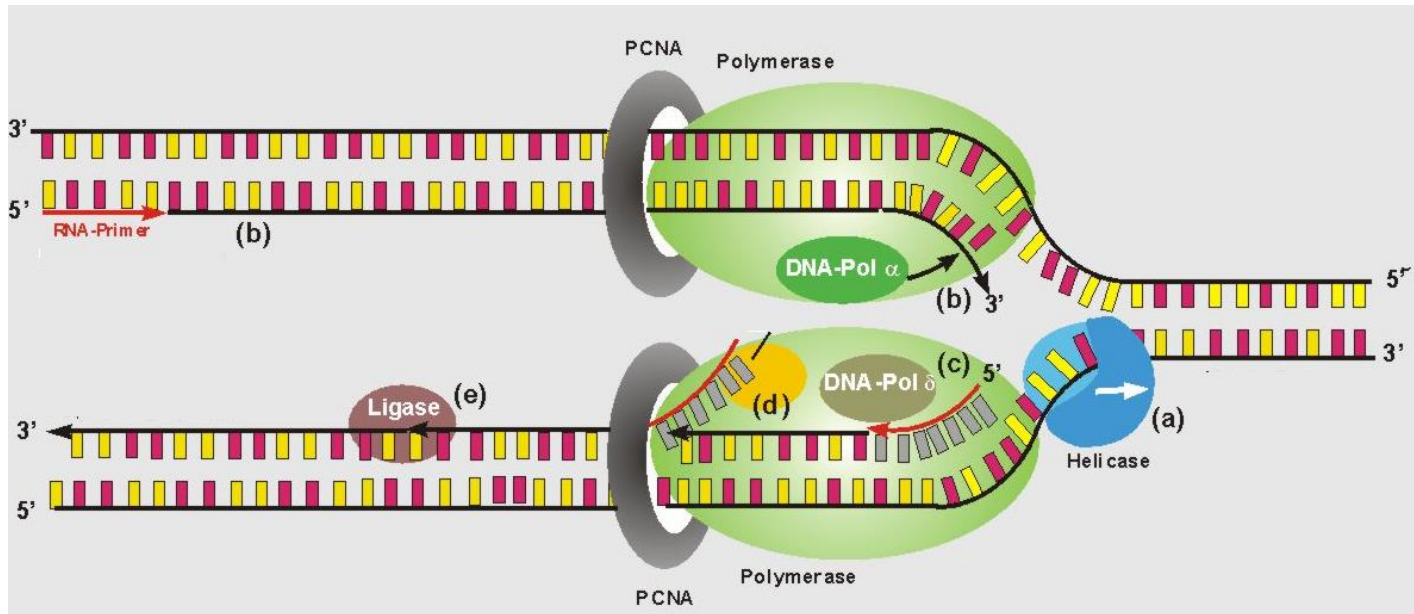


Enzima ribonucleaza HII -
degradează primerul



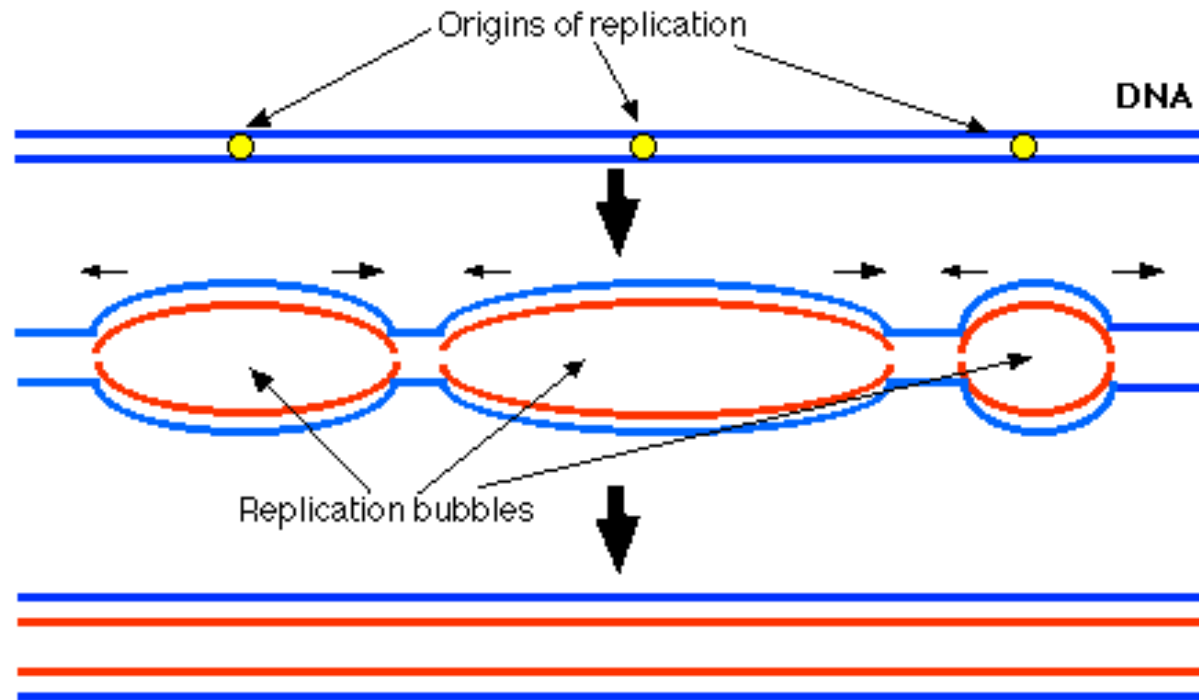
ADN ligaza - reface continuitatea catenei.





Proteinele PCNA - mențin mașinăria de replicare atașată la catenele de ADN

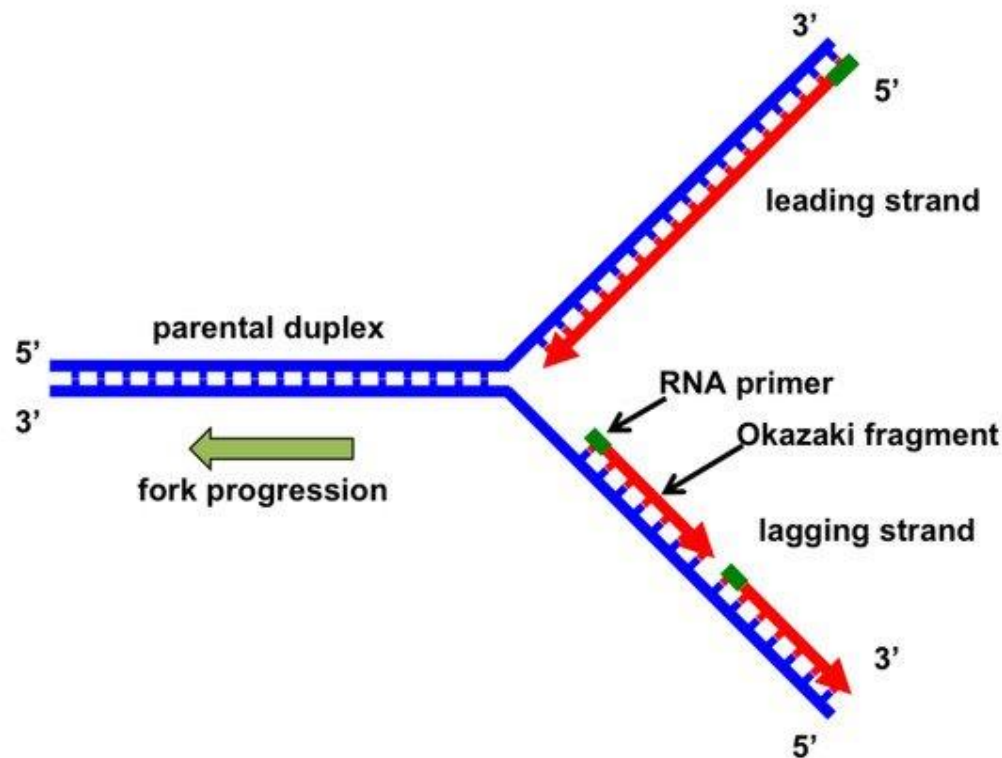
Etapele replicării



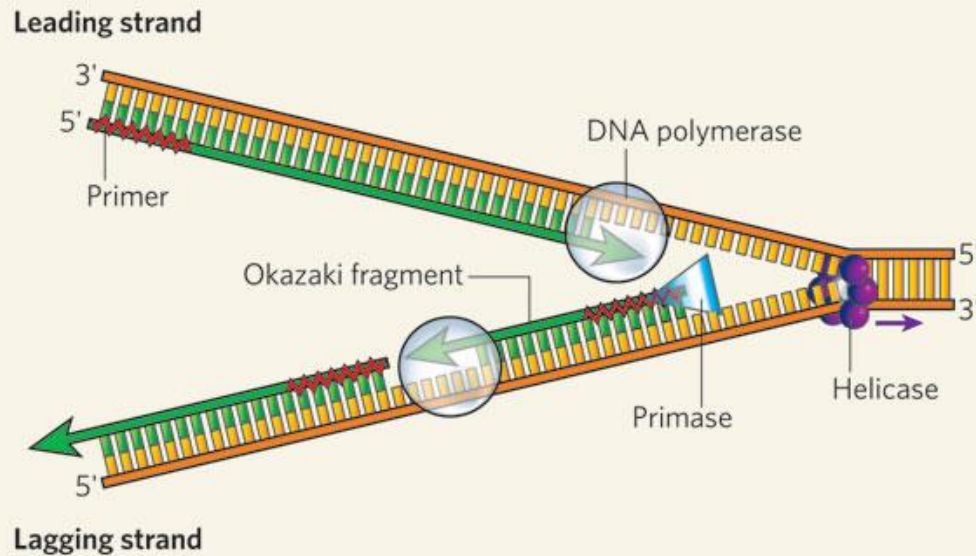
Procesul de replicare începe la nivelul unor regiuni specifice din genom (cca. 100 000) formate din secvențe de ADN specifice (secvențe de 3000 pb) denumite ORI - origini de replicare. La aceste secvențe se atașează proteine de inițiere și enzime de replicare formând complexul de pre-inițiere a replicării (pre-RC). Procesul odată început, se desfășoară în ambele direcții.

Prin intermediul topoizomerazelor și al helicazelor se despiralizează ADN-ul bicatenar și se separă cele 2 catene. Separarea monocatenelor este menținută prin intermediul proteinelor SSB, formându-se furca de replicare, structură în forma de Y.

În etapa următoare are loc sinteza unui scurt fragment de ARN (10-15 nucleotide) denumit primer, de către enzima ARN primaza.

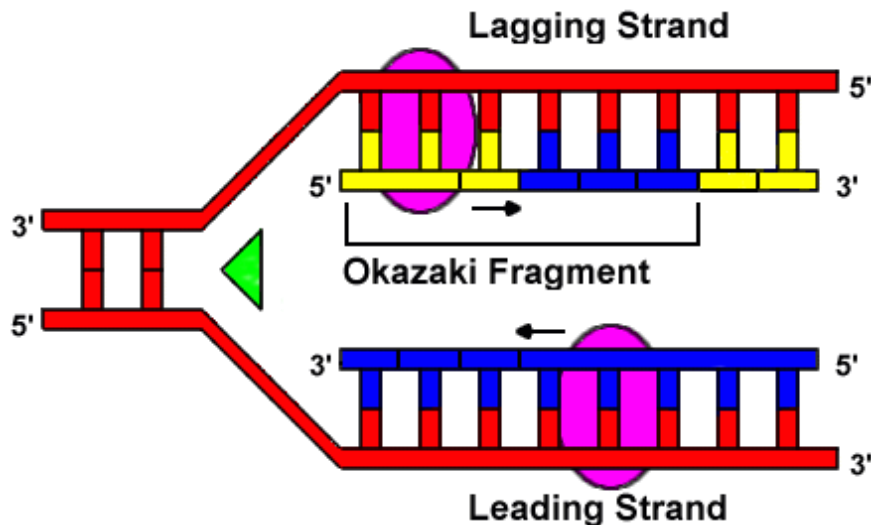


Sinteza de către ADN polimeraze se face întotdeauna în direcția 5'-3'.

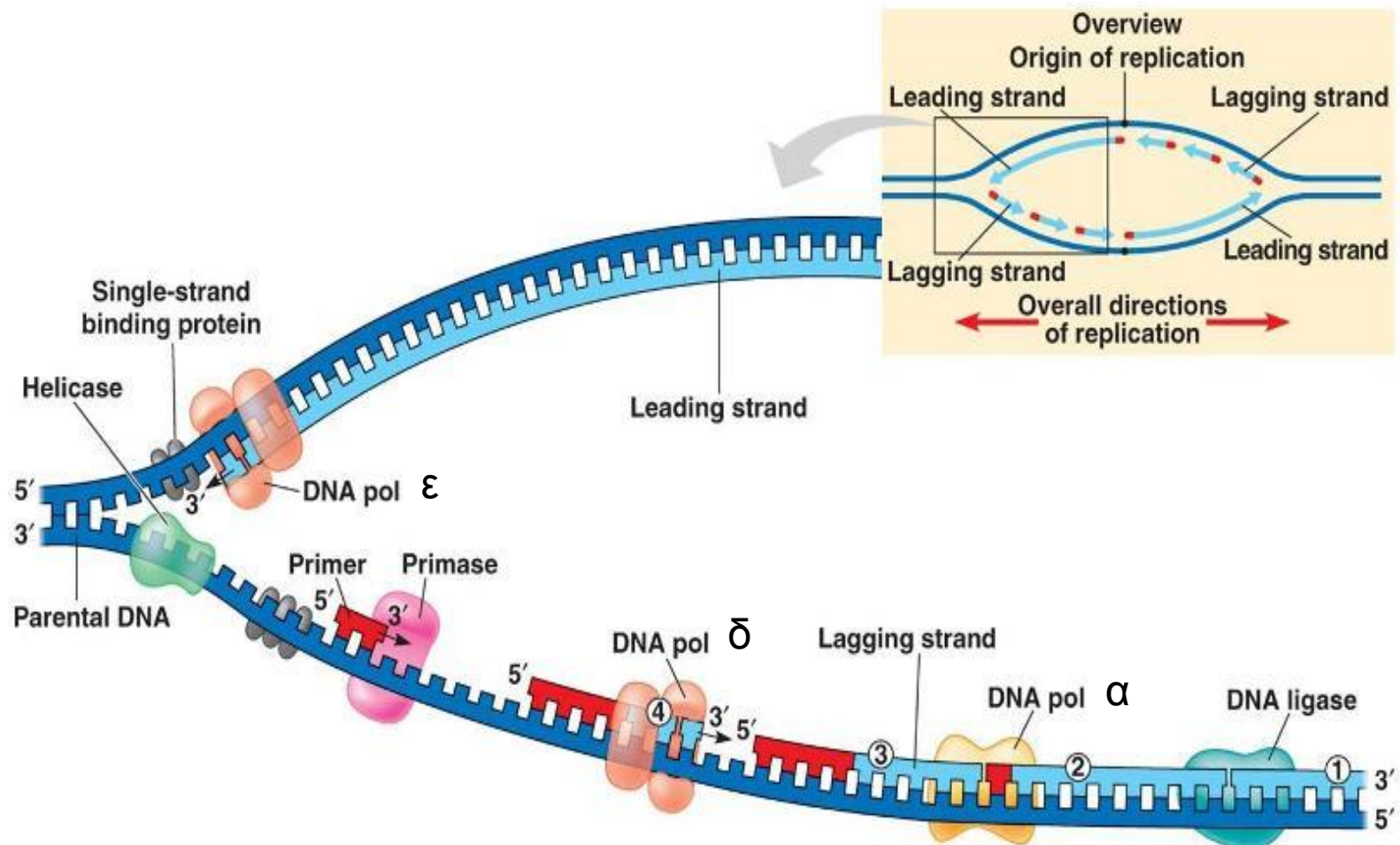


Polimerazelor δ și ϵ adaugă nucleotide complementare, sintetizând astfel noile catene. ADN polimerazele δ și ϵ adaugă nucleotide numai în direcția 5'-3', încorporând nucleotide catenei matriță.

La nivelul catenei matriță 3'-5', sinteza noii catene (5'-3') se realizează rapid și continuu de către polimeraza ϵ (catena se mai numeste “avansată” sau “rapidă” (*engl.* - leading strand)).

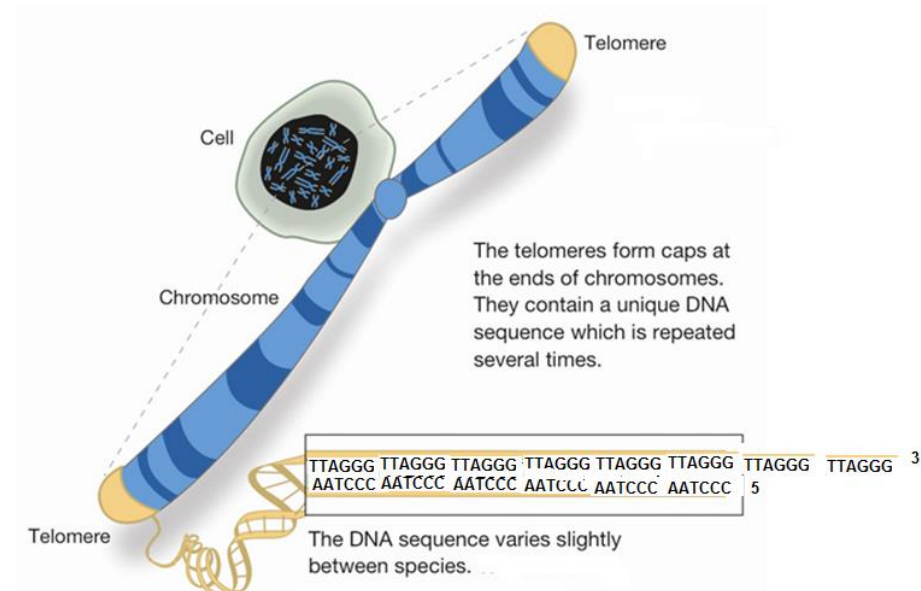


La nivelul catenei matriță 5'-3', sinteza se realizează lent, discontinuu, prin intermediul polimerazei δ care adaugă nucleotide în direcția 5'-3' sub forma unor fragmente de 100 pana la 1000 nucleotide denumite fragmente Okazaki (catena “lentă” “discontinuuă” - lagging strand).



Telomerele sunt structuri cromatiniene localizate la capetele cromozomilor

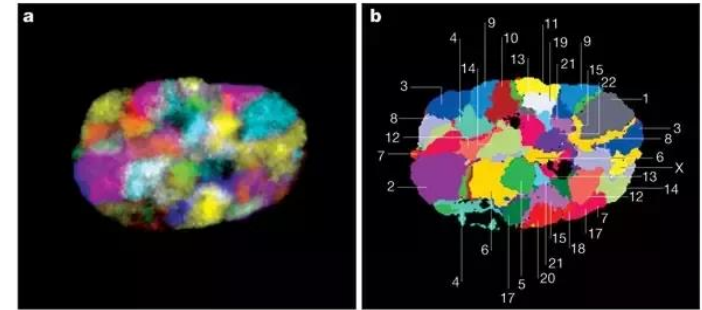
ADN telomeric este format, la toți cromozomii, prin repetiția în tandem a unui hexanucleotid 5'-TTAGGG-3' pe o catenă (catena bogată în G.) și a secvenței complementare 3'- AATCCC-5' pe cealaltă catenă (catena bogată în C.)



Telomerele au dimensiuni cuprinse între 5 si 20 de kb (20000 pb).

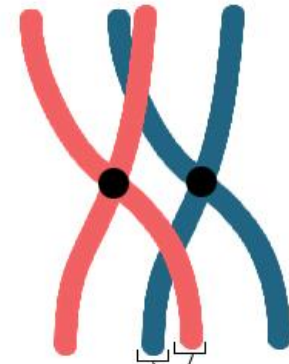
Roluri:

- 1) asigură stabilitatea cromozomilor, prevenind unirea cu alți cromozomi
- 2) intervin și în stabilirea arhitecturii tridimensionale a nucleului
- 3) procesul de recombinare cromozomială din meioză
- 4) asigurarea replicării complete a capetelor cromozomului. Procesul se realizează cu ajutorul telomerazei.



Copyright © 2005 Nature Publishing Group
Nature Reviews | Genetics

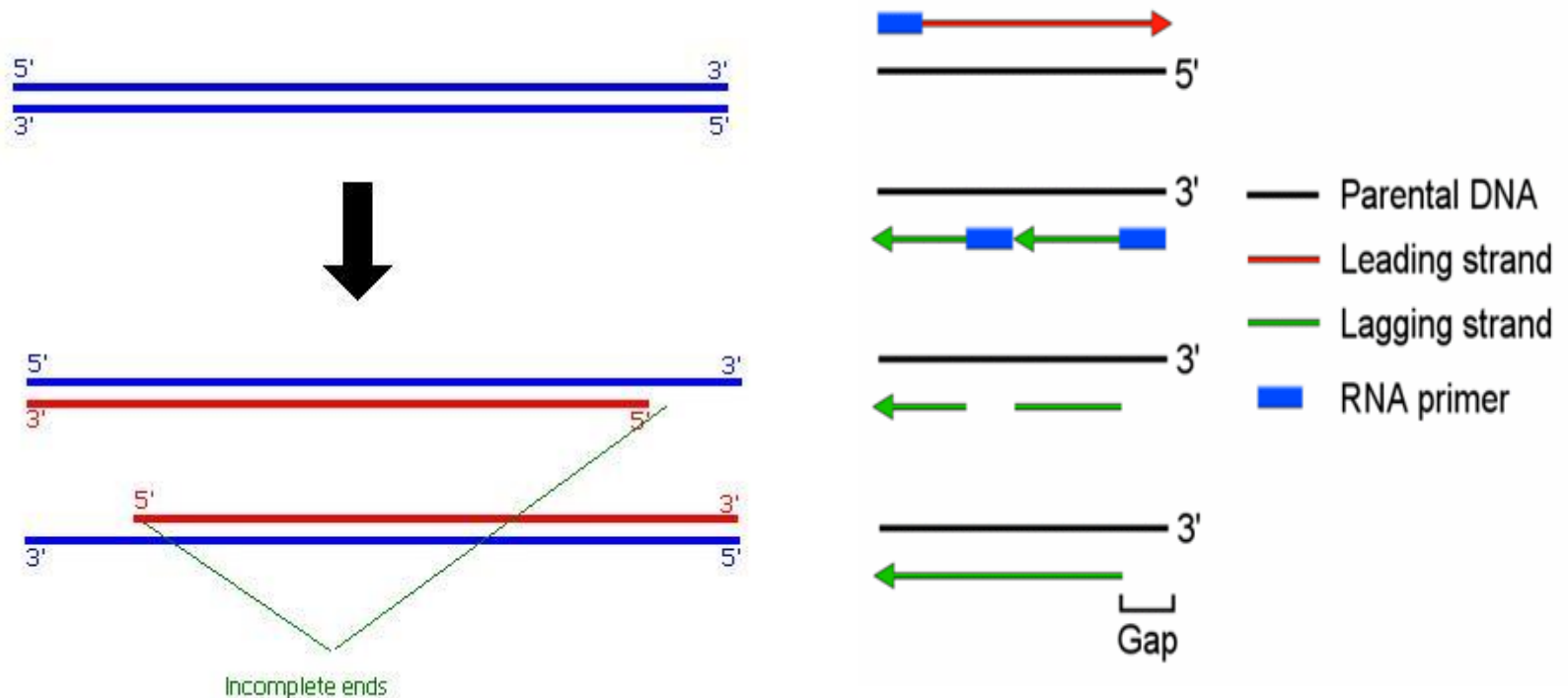
homologous
chromosomes



non-sister
chromatids

Datorită particularităților procesului de replicare (activității polimerazei), la nivelul capătului 5' al fiecărei catene nou-sintetizate rămâne un fragment care nu poate fi sintetizat (nereplicat).

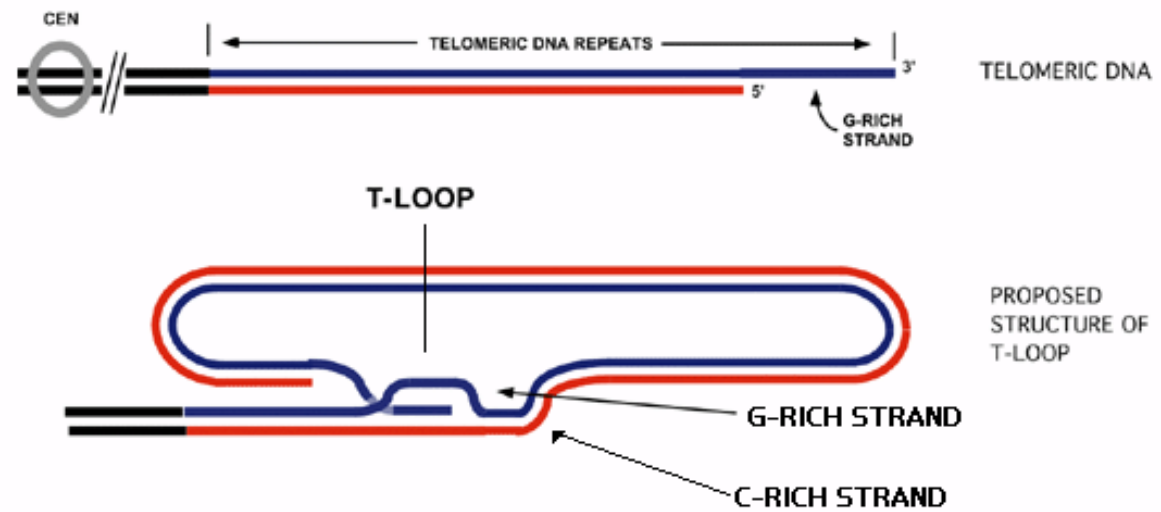
Molecula de ADN nou-sintetizată va avea un capăt 5' mai scurt, regiunea nereplicată fiind regiunea corespunzătoare primerului degradat. Aceste fragmente terminale sunt la nivelul telomerelor, fiecare telomer având o regiune bicatenară și o regiune monocatenară (200-300 nucleotide).



Protecția telomerului este realizată prin pătrunderea tractului monocatenar 3' între cele două catene ale regiuni bicatenare, formându-se o buclă, denumită bucla T.

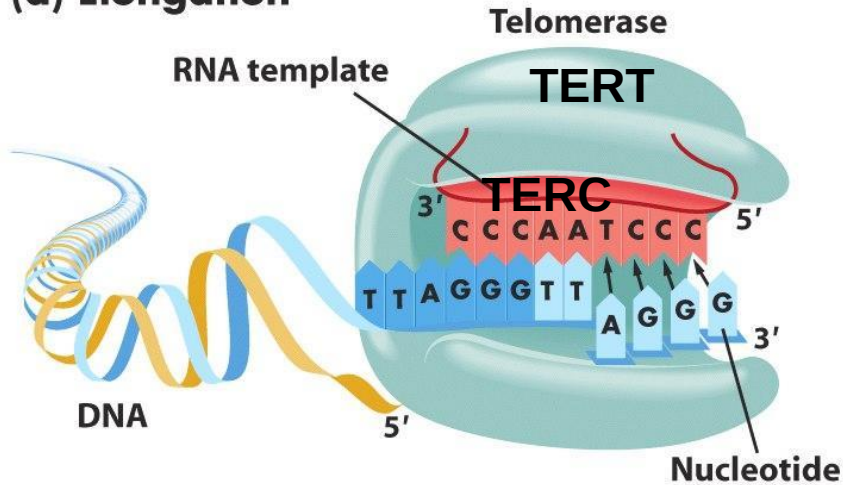
Buclele sunt structuri tranzitorii care se desfac în timpul replicării.

După fiecare ciclu de replicare, se pierde aproximativ 15-20 de nucleotide de la nivelul telomerelor, datorită particularităților procesului de replicare. În realitate scurtarea telomerelor este mai pronunțată (60-150 bp), diferența fiind cauzată de radicalii liberi.

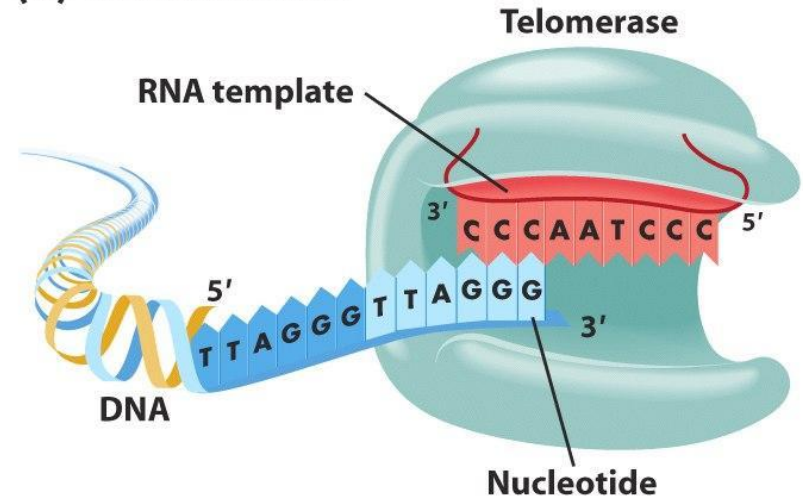


TELOMERAZA

(a) Elongation

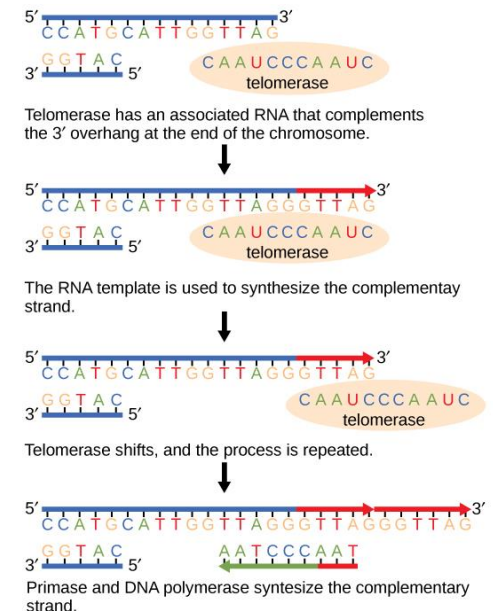


(b) Translocation



Scurtarea telomerelor este prevenită de acțiunea unei enzime specifice denumită telomerază, care asigură replicarea completa.

Enzima cu structura ribonucleoproteica formată dintr-o subunitate proteică “TERT”, cu activitate de transcriptază inversă (reverstranscripție) și dintr-o subunitate “TERC”, care e o subunitate de ARN TERC, care conține o secvență de ARN complementară secvenței telomerice (3'-CCCAATCCC-5').

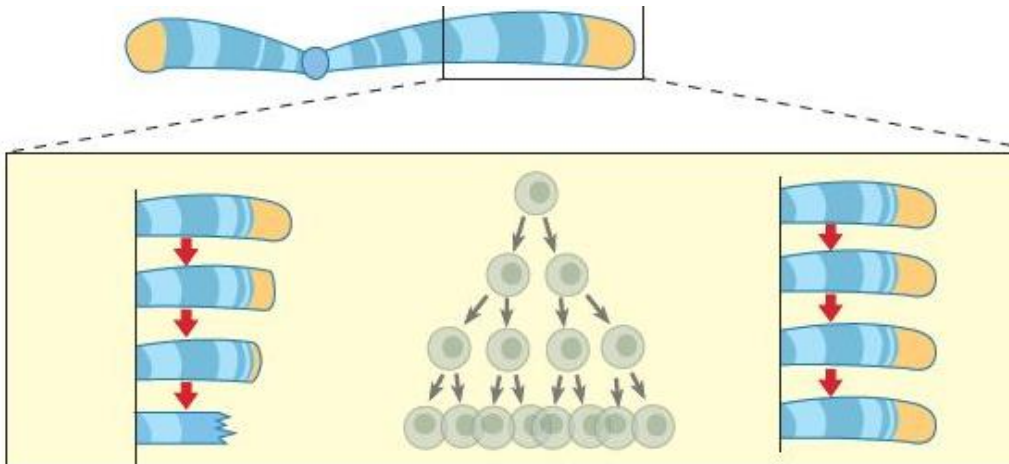




Telomeraza e o enzima activă mai ales în timpul vieții embrionare (celule germinate, stem, blastocit), după naștere, diminuându-și activitatea aproape complet, dar rămâne activă numai la nivelul celulelor stem.

Pe măsură ce celula se divide, se pierd secvențele telomerice, ajungând la un moment dat la o scurtare care conduce la oprirea diviziunii.

Astfel, telomerele sunt considerate ceasul molecular al diviziunii celulare, dând informații asupra numărului de diviziuni restante ale unei celule.



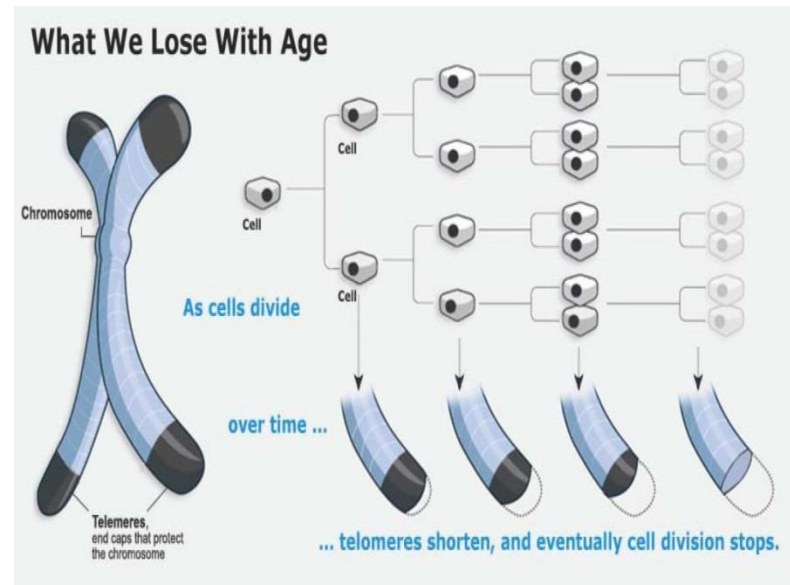


Hutchinson Gilford

Scurtarea telomerelor a fost asociată cu fenomenul de senescentă (unele persoane cu stări progeoride, ex Hutchinson Gilford **au o scurtare precoce a telomerelor** - telomerele sunt mult mai mici decât la persoane normale de aceeași vârstă)

În unele cancere, se produce o reactivare a telomerazei, care stabilizează lungimea telomerelor, adăugând secvențele hexamerice specifice, și celulele încep să se dividă activ.

Acest fenomen este, foarte probabil, implicat în imortalizarea celulară.



GENOMUL UMAN

Proiectul Genomului Uman (PGU)

SECVENȚIEREA GENOMULUI UMAN

IHGSC

Celera Genomics

International Human Genome Sequencing Consortium



- 1990 – începerea PGU
 - 2000- prima schiță de lucru
 - **2001** - primele date publicate ce acopereau eucromatina 96% (2,69 Gb)
 - **2004** - prima versiune finisată (2,85 GB, precizie 99%)
 - 2007 – versiunea terminată, după au apărut variante actualizate
- **determinarea secvenței genomului uman = stabilirea ordinii liniare a nucleotidelor în ADN**
 - **Secvențierea genomului haploid (22A+X+Y)**
 - **identificarea și localizarea genelor ce alcătuiesc genomul uman**

**Genom = ansamblul informatiei ereditare
din ADN-ul unei celule sau al unui organism.**

**= informatia din ADN-ul cromozomial (24
molecule de ADN)**

GENOMUL UMAN - GRCh38 (iunie 2016)
(datele sunt pentru cele 24 de molecule de
ADN diferite - 24 cromozomi)

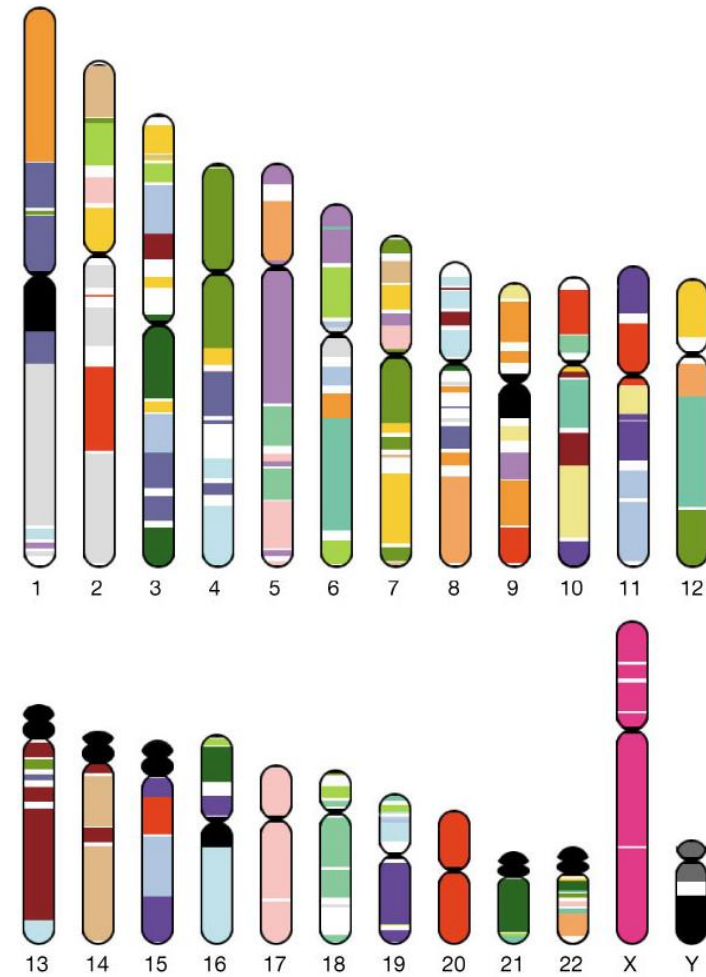
Lungime totala –3270 Mb -3,27 GB
- eucromatina – 3100 Mb
- heterocromatina -170 Mb

Gene ce codifica proteine - 20 441
ADN codant (exoni) 1,2%

Gene ARN necodant - 22 219

Pseudogene - 14 606

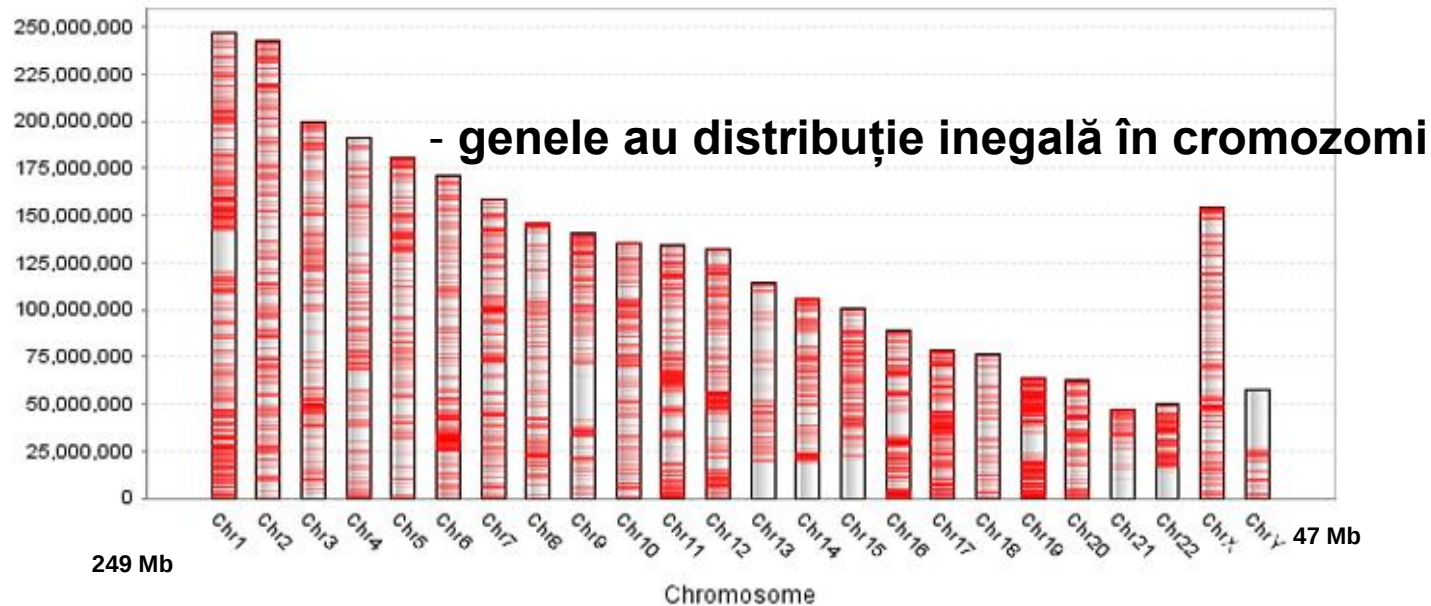
Transcripte genice - 198 002



Genomul a doua persoane neinrudite este comun 99,9% (1 la 1000 pb)

98% - este ADN necodant ce se găsește în principal în gene ce codifica ARNnc si diferite clase de ADN repetitiv

1,2% - ADN codant (exoni)



ADN necodant intră în structura genelor ce codifică proteine (elemente reglatoare, introni, regiuni UTR), genele ce codifica ARN necodant (ARNr, ARNt, ARNsn, ARNsno, ARN telomeraza, ARNmi, ARNa, ARNsi, ARNlnc ,ARNdi), elementelor genetice mobile – transpozoni, ADN inalt repetitiv

Pseudogene - gene care au devenit nefuncționale din cauza unor mutații inactivatoare

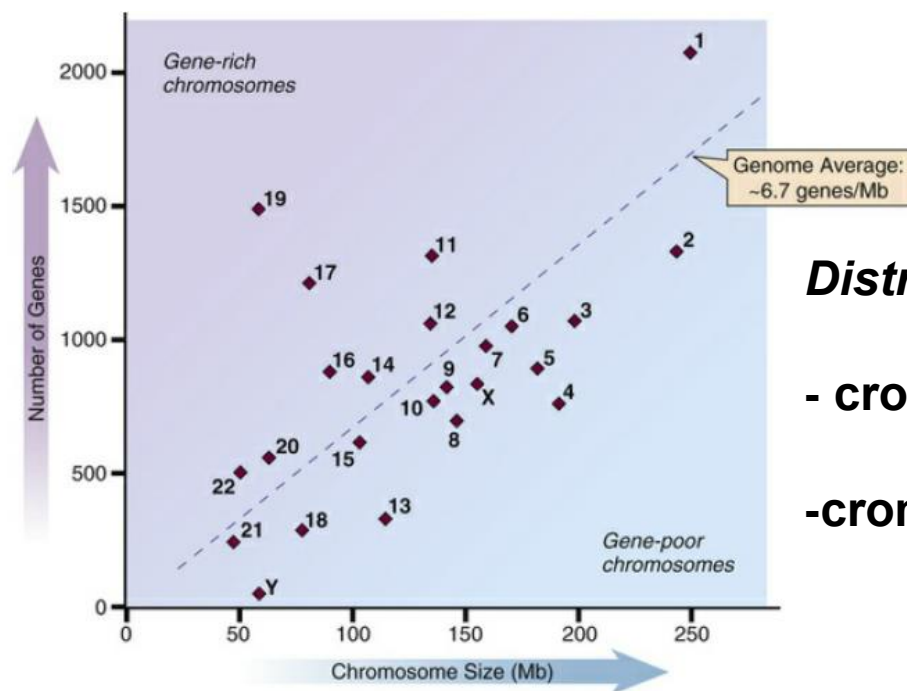
Double Helix 5'...GGATTTC TAGGTA ACTCAGTCGA... 3'
 3'...CCTAAAGATCCATTGAGTCAGCT... 5'

Reference Sequence ...GGATTTC TAGGTA ACTCAGTCGA...

Individual 1 ...GGATTTC TAGGTA ACTCAGTCGA...
 Individual 2 ...GGATTTC CAGGTA ACTCAGTCGA...
 Individual 3 ...GGATTTC CAGGTA ACTCAGTCGA...
 Individual 4 ...GGATTTC TAGGTA ACTCAGT AGA...
 Individual 5 ...GGAT - -CTAGGTA ACTCAGTCGA...

Genomul de referinta (nu include diversitatea și polimorfismul genetic), se referă la secvențele identice la toți indivizii.

Genomul a doua persoane neinrudite este comun 99,9% (diferențe 1 pb la 1000 pb).



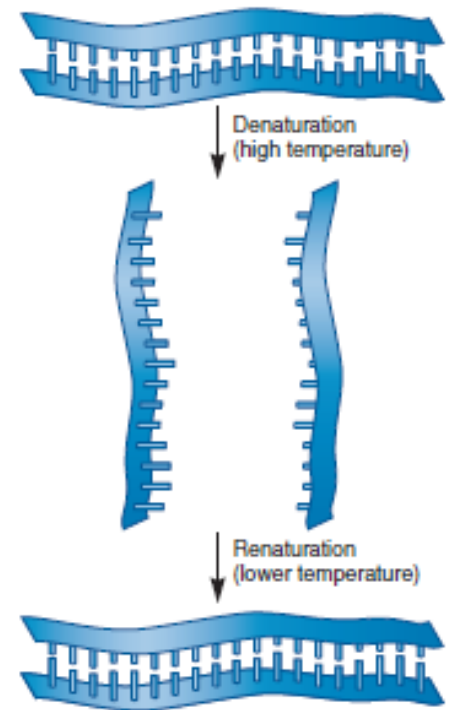
Distribuție inegală a genelor în cromozomi:

- cromozomi bogați în gene: ex.19, 22, 17

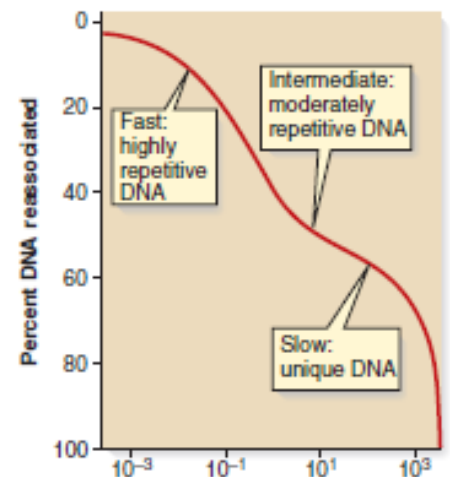
-cromozomi săraci în gene: ex. 13, 18, X

CLASE DE ADN

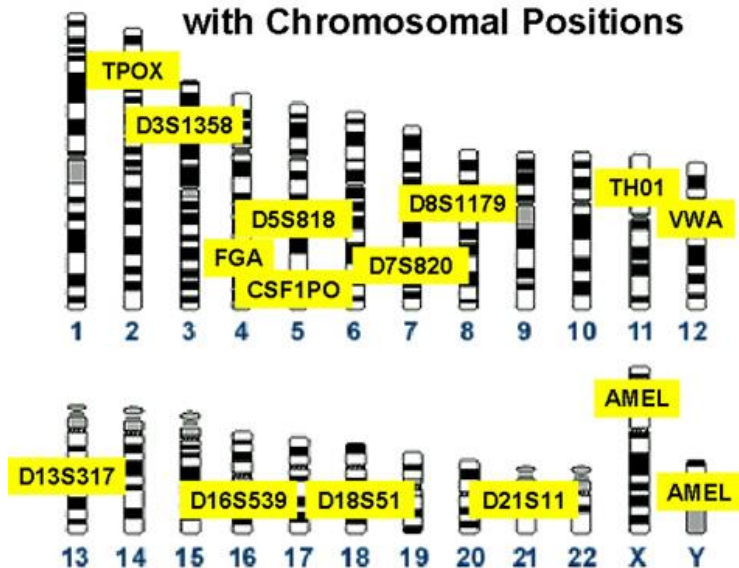
- **ADN nerepetitiv** (secvențe unice) se găsește în regiunile bogate în eucromatină, exoni, introni, secvențe de reglare, extragenic.
- **ADN moderat repetitiv** constă în secvențe de 300 – 1000 pb repetate de zeci sau sute de ori; se găsește la nivelul genelor ce codifică ARNr, ARNt, elemente genetice mobile (transpozoni) - dispersate în genom
- **ADN înalt repetitiv** - secvențe de 2-200 pb, repetate de mii de ori nu sunt transcrise în ARN, ce se găsește în:
 - ADN satelit ce alcătuiește blocuri mari de heterocromatină situate în anumite regiuni cromozomiale: centromer, telomere, constricții secundare, benzi G pozitive
 - minisateliti - repetitii de 10-60 bp
 - microsateliti - repetitii de 1-6 bp



(a) Renaturation of DNA strands



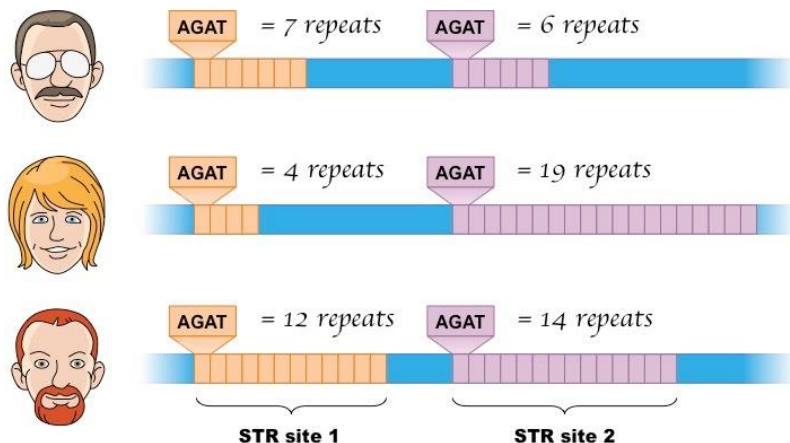
13 CODIS Core STR Loci with Chromosomal Positions



Profilul ADN (nr.repetiții- mamă, copil, tată)

	M		C		T	
D2S11	28	30	28	31	29	31
D7S820	9	10	10	11	11	12
TH01	14	15	14	16	15	16
D13S317	7	8	7	9	8	9
D19S433	14	6	14	15	15	17

Microsateliții și minisateliții sunt secvențe foarte variabile care au un număr de repetiții specific pentru fiecare persoană realizându-se un profil genetic specific = amprenta genetică specifică a unui individ



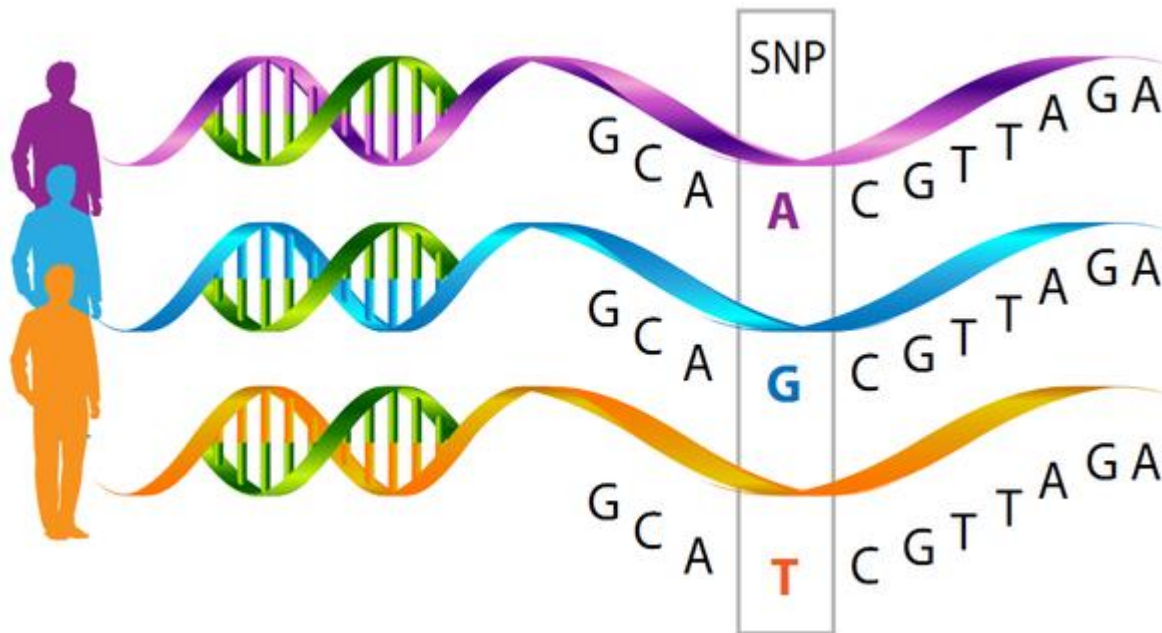
Aplicații: - medicina legală
- stabilirea paternității
- studii populaționale

VARIABILITATEA STRUCTURALĂ A GENOMULUI UMAN

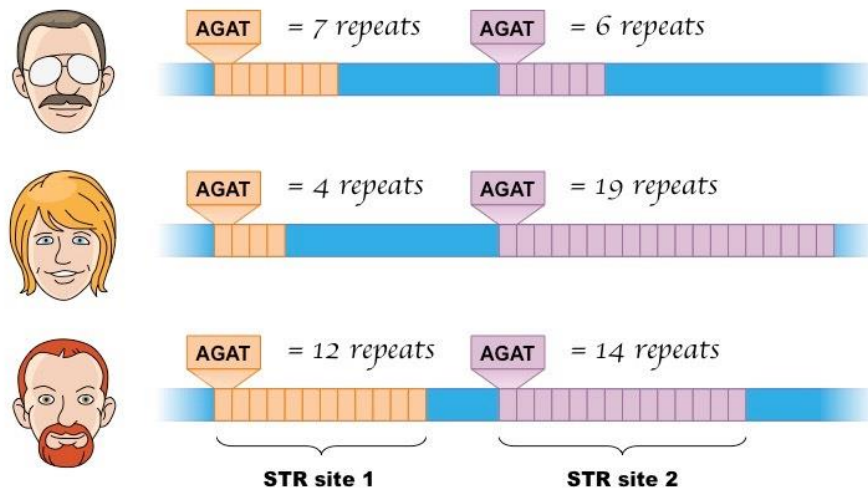
POLIMORFISME GENETICE

= variații individuale de secvență și structură a genomului fără modificări fenotipice, ce se găsesc >1% din populație

❑ Polimorfisme mononucleotidice – SNP (1 la 1000 bp)



❑ Polimorfisme ale repetițiilor de secvență scurtă (SSR) minisateliți și microsateliți – amprenta genetică



❑ Insertii și deleții scurte – indel (2-16 pb)

Indel examples

wild-type sequence

ATCTTCAGCCATAAAAGATGAAGTT

3 bp deletion

ATCTTCAGCCAAAGATGAAGTT

4 bp insertion (orange)

ATCTTCAGCCATATGTGAAAGATGAAGTT

Polimorfisme ce implică secvențe ADN mari (peste 100 nucleotide) = variații structurale

- ❑ Variante ale numărului de copii – CNV - CNP (duplicații/deleții) ADN > 1kb (mai mari de 1000 bp)

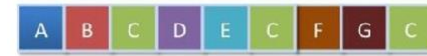
- ❑ Polimorfismul inversiilor și translocatiilor



REFERENCE SEQUENCE



COPY NUMBER VARIANT



SEGEMENTAL DUPLICATION



COMPLEX COPY NUMBER VARIANT



INVERSION



TRANSLOCATION



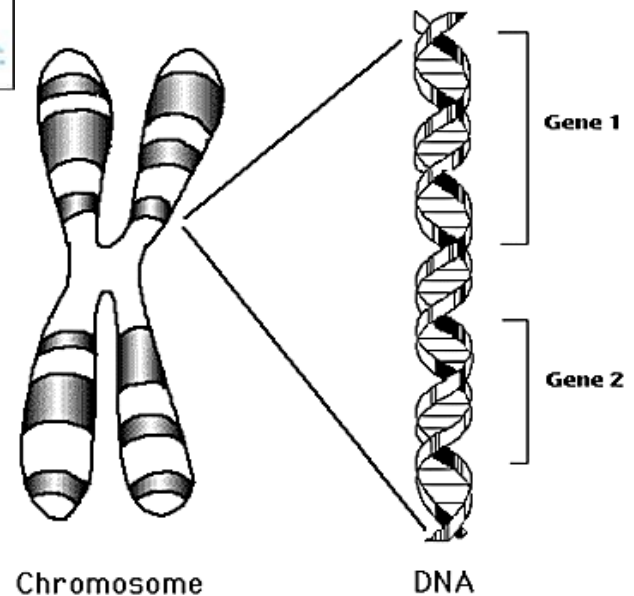
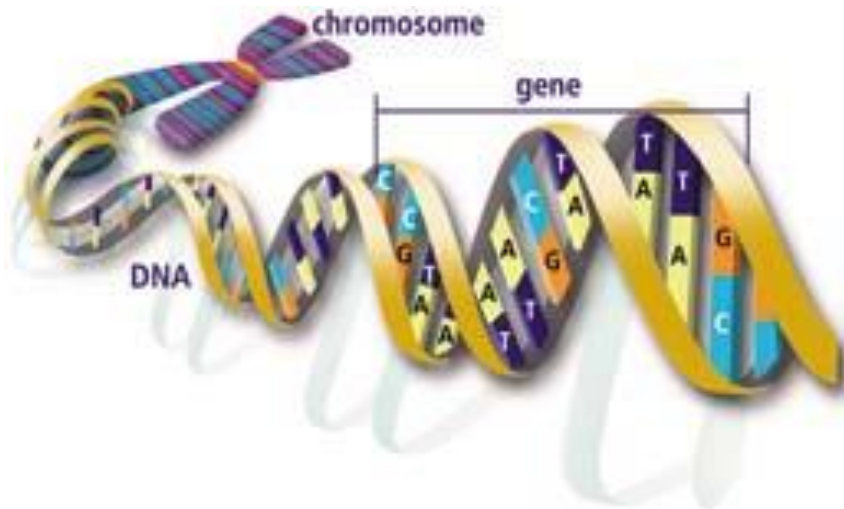
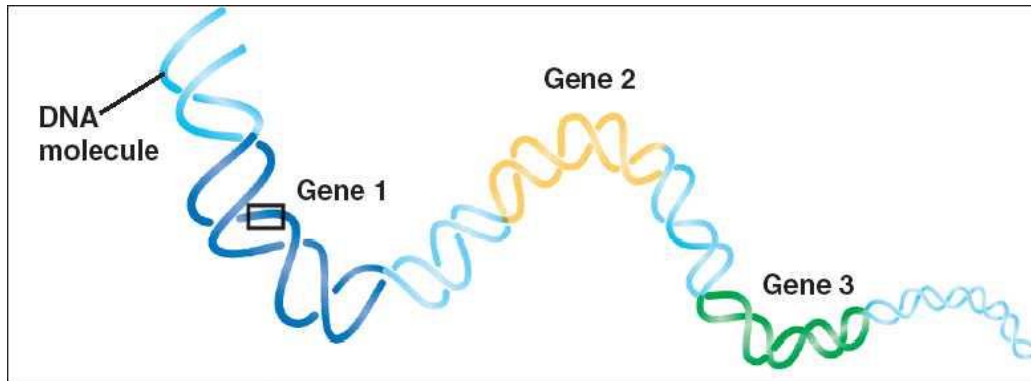
- ❑ Inserții de elemente mobile

GENA – STRUCTURA UNEI GENEI CE CODIFICĂ PROTEINE

GENA – STRUCTURA UNEI GENEI CE CODIFICĂ PROTEINE

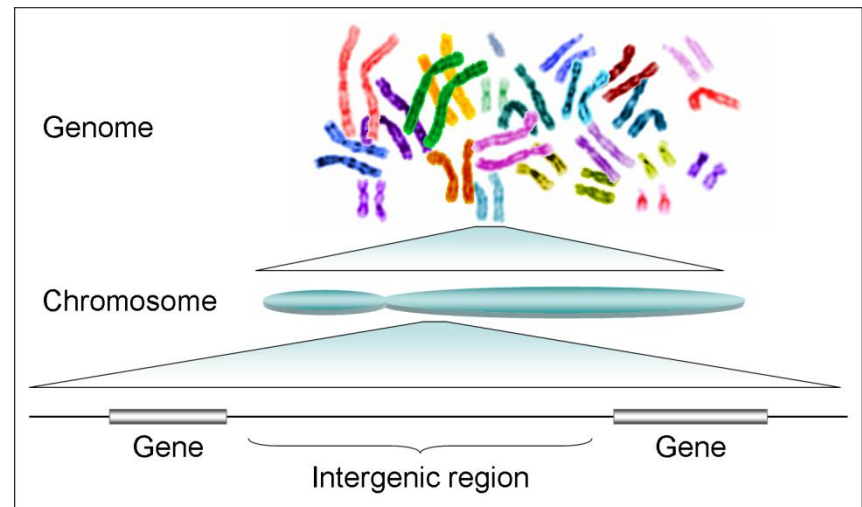
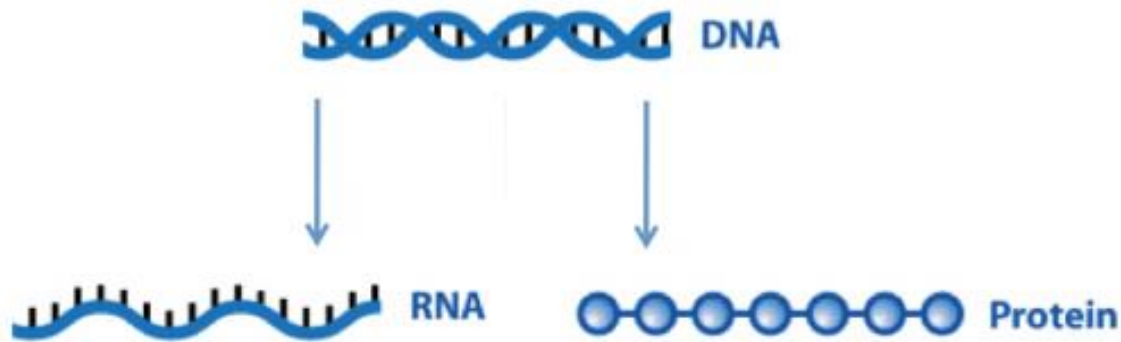
Număr gene ce codifică proteine în genomul uman - aprox. 20 000 (20441 - **GRCh38**)

Locus - poziția specifică a unei gene pe cromozom



Genes

GENA = un ansamblu liniar de secvențe nucleotidice, necesar pentru a produce un polipeptid sau o moleculă de ARN funcțional.



STRUCTURA GENEI

Genele ce codifică proteine conțin:

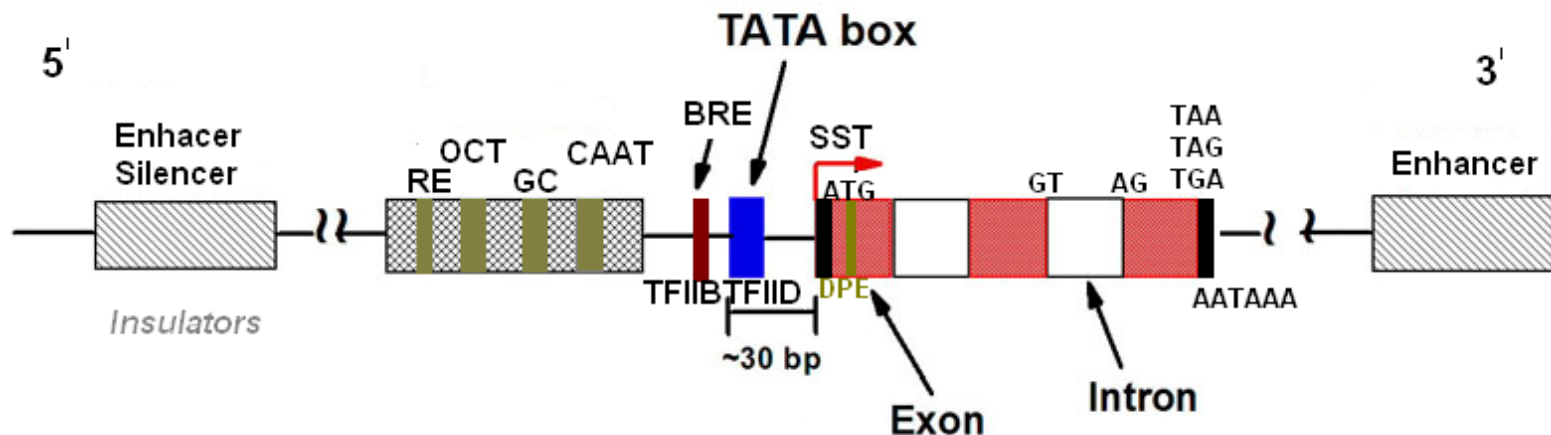
❑ **Regiune centrală = Cadrul de lectură (ORF)**, transcrisă (copiată) în ARN mesager precursor

❑ **2 Regiuni laterale :**

- regiunea laterală 5' (contine promotorul si alte elemente de reglare)

- regiunea laterală 3' (imprecis delimitată și funcție necunoscută

complet, conține secvențe semnal care afectează procesarea, stabilitatea și durata de viață a ARNm)



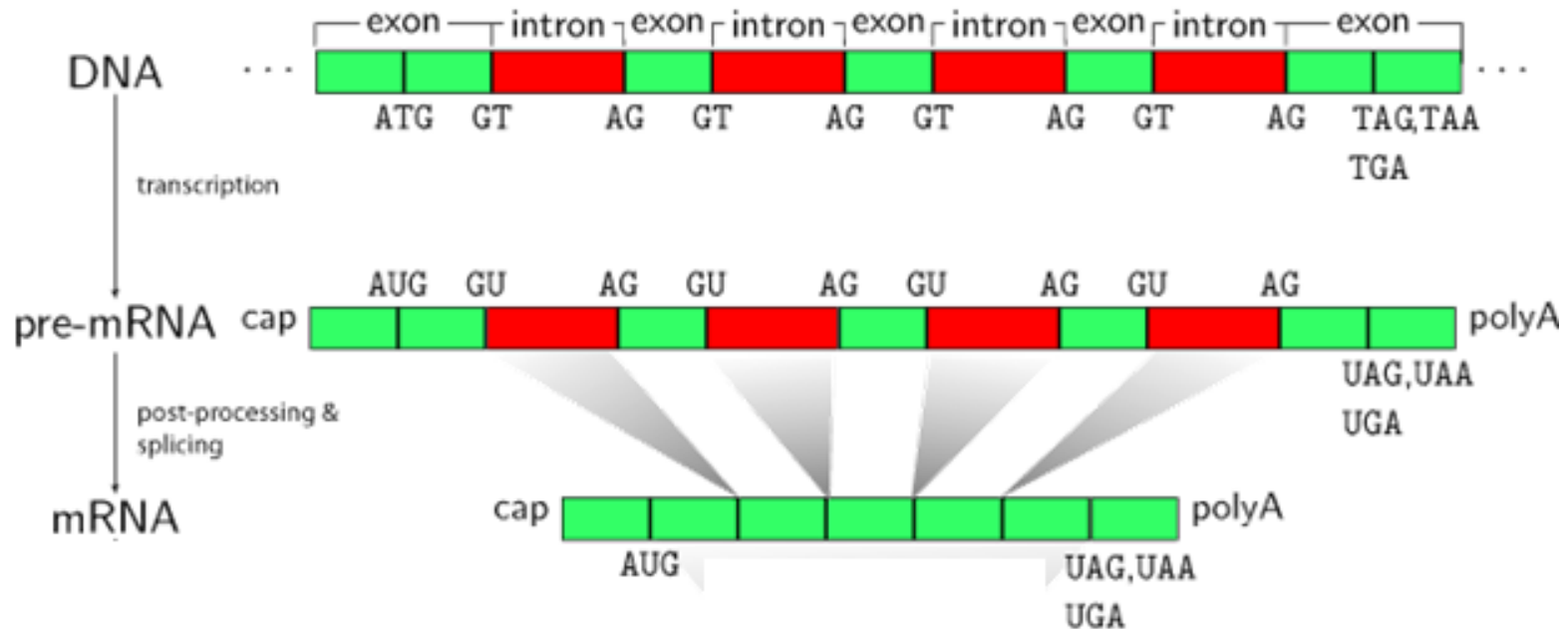
REGIUNEA CENTRALA – contine informația necesară sintezei proteinei



Regiunea centrala este formată din:

- **situsul strat al transcripției - SST** (locul de unde începe transcripția)
- **regiunea 5' UTR** (5' untranslated region) – regiune transcrisă, dar netranslată ce conține secvența consens (5' CCAGCCATG 3') care cuprinde codonul initiator (ATG – AUG in ARN)
- **EXONI** – secvențe codante (ultimul exon cuprinde un codon stop)
- **INTRONI** - secvențe necodificatoare
- **regiunea 3' UTR** (3' untranslated region) - regiune transcrisă, dar netranslată, conține situsul de poliadenilare AATAAA

EXONII

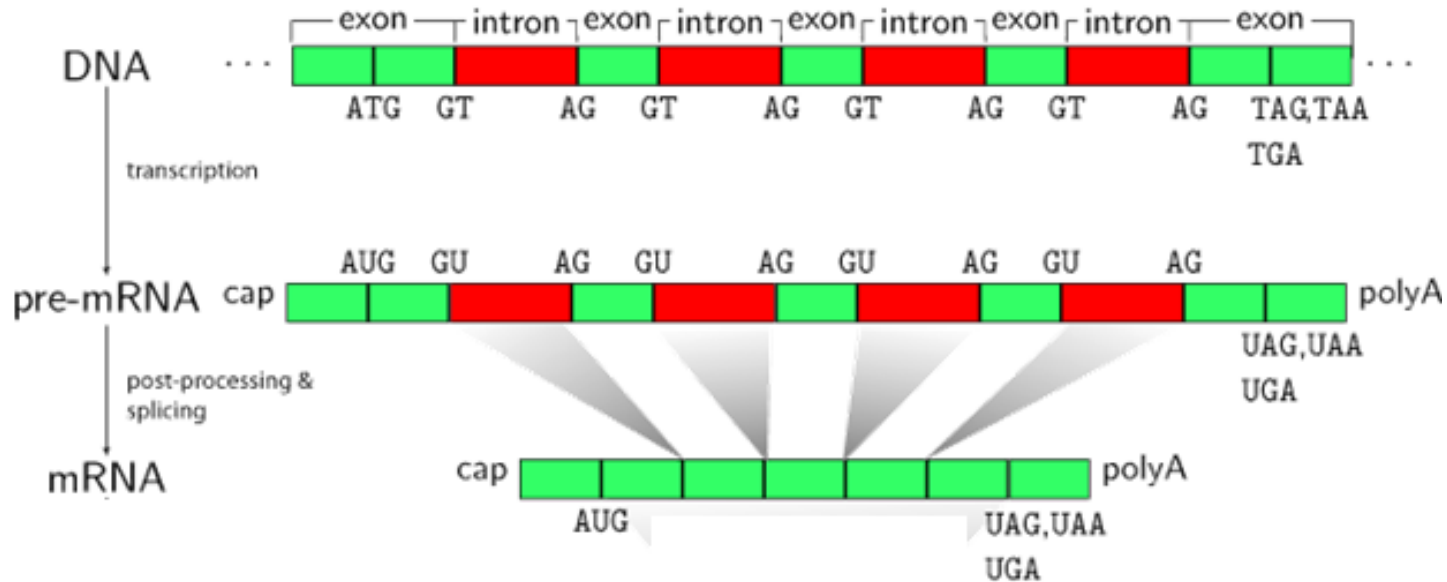


Exonii sunt secvențe codante transcrise în ARN mesager precursor și păstrate în ARNm matur.

Exonii sunt regiunile funcționale din structura genei, care codifică anumite părți structurale și/sau funcționale distincte ale proteinei.

Lungimea exonilor este mult mai mică decât cea a intronilor (lungime medie exoni 140 pb, lungimea medie a intronilor este 1000 pb).

INTRONII



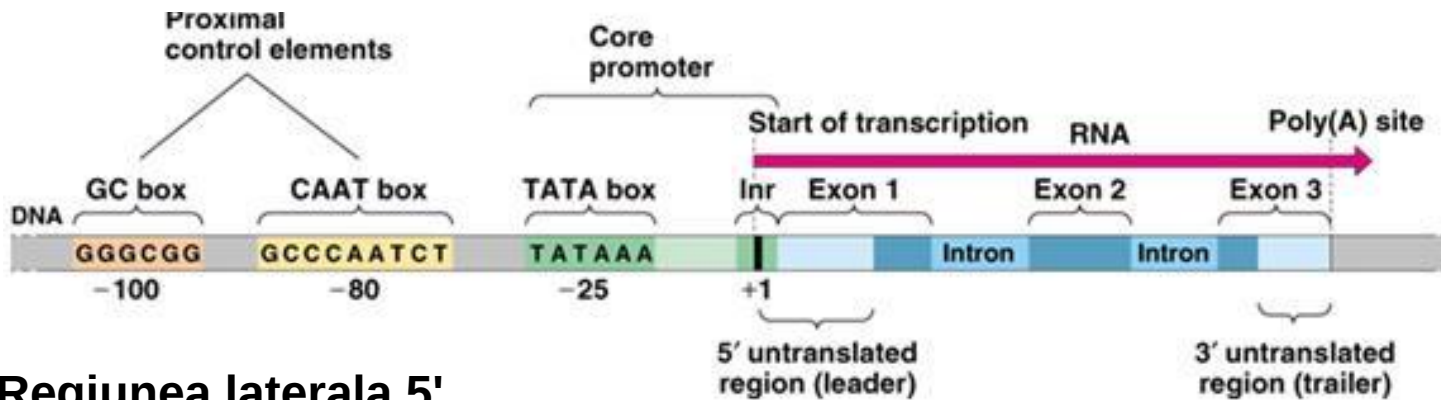
Intronii sunt sunt secvențe necodante, transcrise inițial în ARNm precursor (transcript primar) dar decupate precis și îndepărtate ulterior din ARNm matur, ce va fi alcătuit prin asamblarea exonilor.

Intronii sunt secvențe intercalate pentru că se interpun între exoni (nr.exonilor este mai mare cu unu decât nr intronilor), majoritatea intronilor încep cu dinucleotidele 5' GT și sfârșesc cu AG 3'.

Regiunile laterale netranscrise au rolul de a semnaliza inițierea sau terminarea transcripției de către ARN-polimerază și de a regla intensitatea ei.

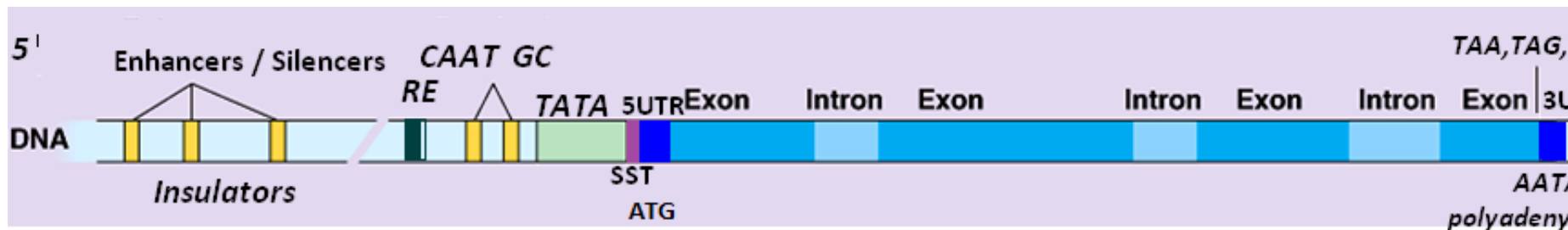
Regiunea laterală 5': cuprinde **promotorul** (alcătuit din miezul promotorului și promotorul proximal) și **alte elemente de reglare**

ROL IN REGLAREA EXPRESIEI GENICE

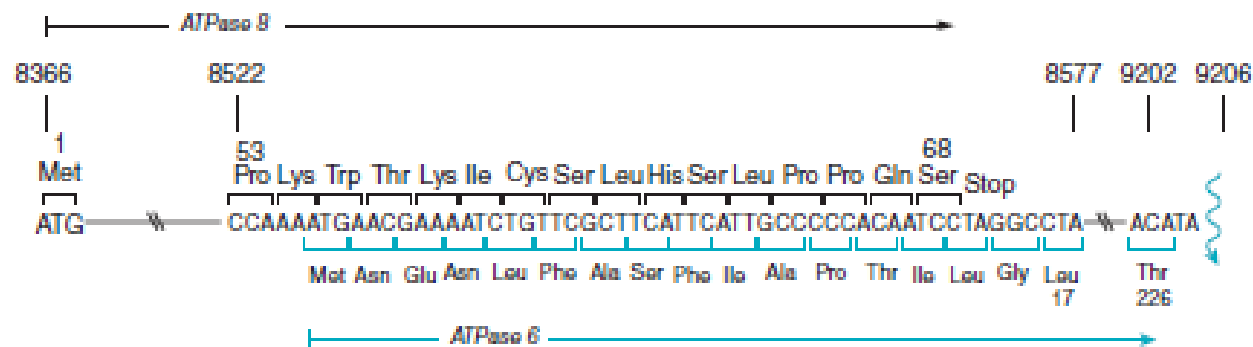


Regiunea laterală 5'

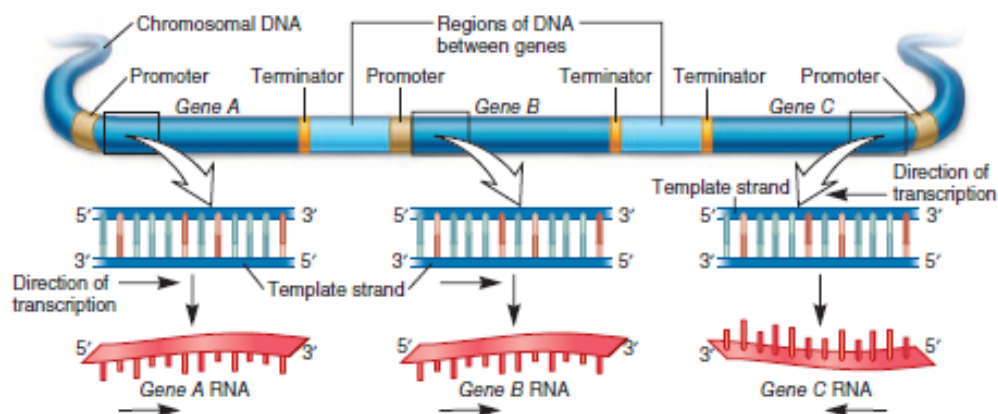
- Miezul promotorului: boxa TATA, boxa GC, boxa CAAT
- Promotor proximal : elemente care determină expresia specific-tisulară a anumitor gene
- Alte elemente : activatori + inhibitori + insulatori



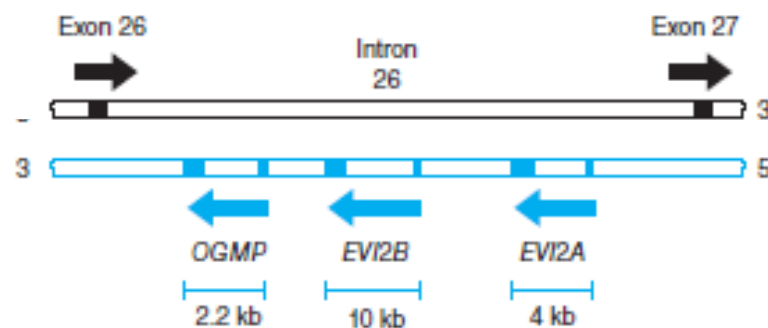
- **boxa TATA** - este locul în care se fixează ARN-polimeraza II la ADN (prin intermediul TFIID) – rol în transcripția genelor specifice de țesut sau care se exprimă în anumite momente
- **boxa GC** modulează transcripția bazală a genelor comune (housekeeping genes), funcționale în toate țesuturile
- **boxa CAAT** modulează transcripția, inițiată de secvența TATA
- **activatori** (enhancers) sunt elemente de reglare pozitivă care amplifică nivelul bazal al transcripției, inițiată de promotor
- **inhibitori** (silencers) reduc nivelul transcripției și uneori reprezintă funcția unor gene
- **insulatorii** – limitează acțiunea activatorilor și inhibitorilor



Gene suprapuse



Gene transcribe opus



Gene în gene

(intronul 26 al genei NF1 conține 3 gene transcribe opus)

GENE COMUNE ȘI GENE SPECIFICE

Gene comune/ubicuitare (housekeeping genes)

- sunt exprimate în toate celulele
- codifică proteine comune, indispensabile menținerii structurilor și funcțiilor celulare (ex. proteine membranare, proteine implicate în replicare, proteine implicate în producerea de energie)
- reprezintă 1/5 din numărul total de gene și 90% din genele exprimate în orice tip celular

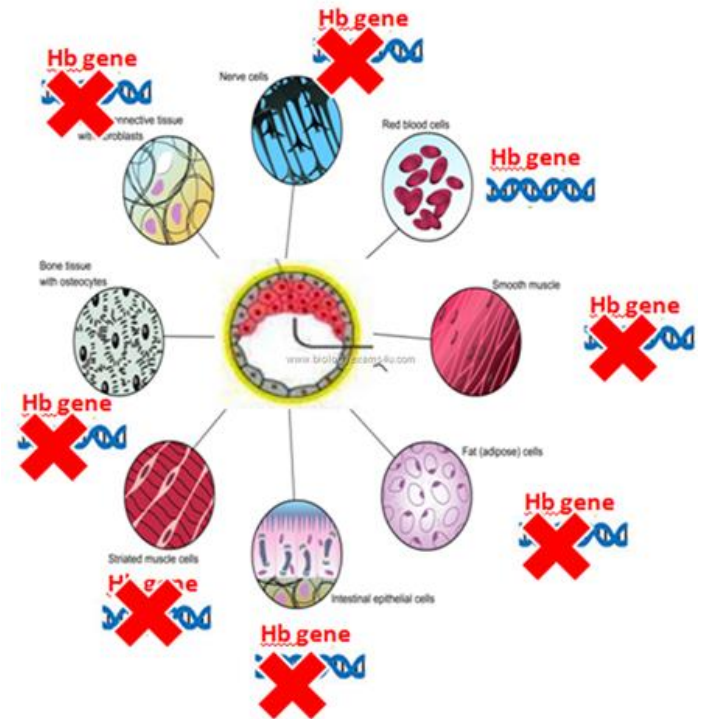


GENE SPECIFICE

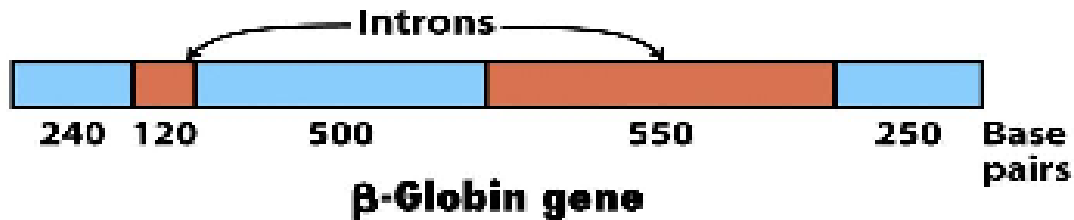
- sunt exprimate numai în anumite tipuri celulare și /sau în anumite momente specice
- reprezintă aprox. 4/5 dintre genele umane

Exprimare spațială = exprimare în anumite țesuturi/linii celulare

Exprimare temporală = expresia anumitor gene în anumite stadii ale dezvoltării, diferențierii sau ciclului celular



Familiile genelor alfa și beta globinei

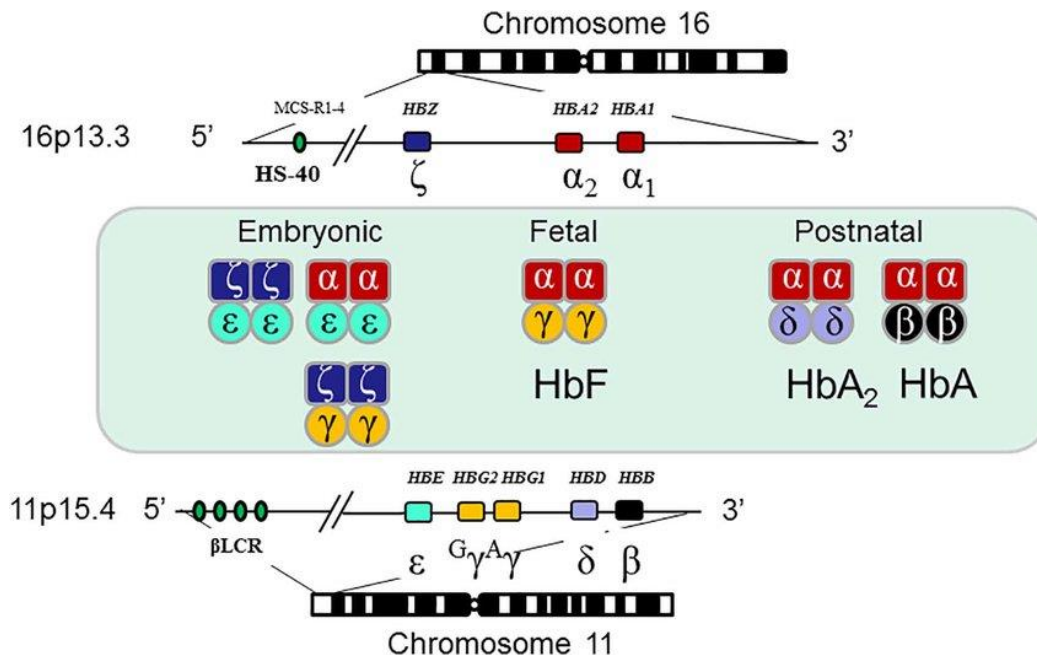


alpha - 16p13 - 141-AA
 beta - 11p15 - 147 AA



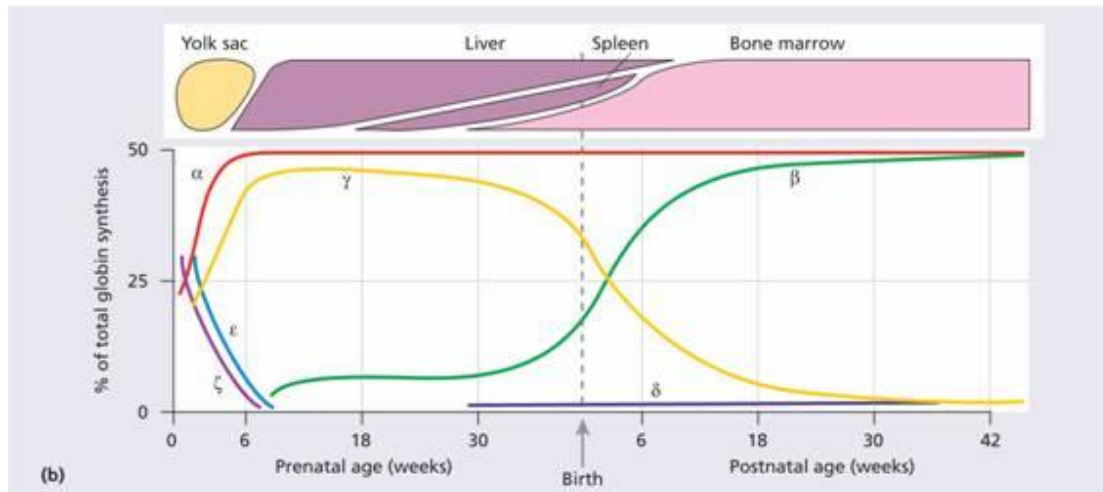
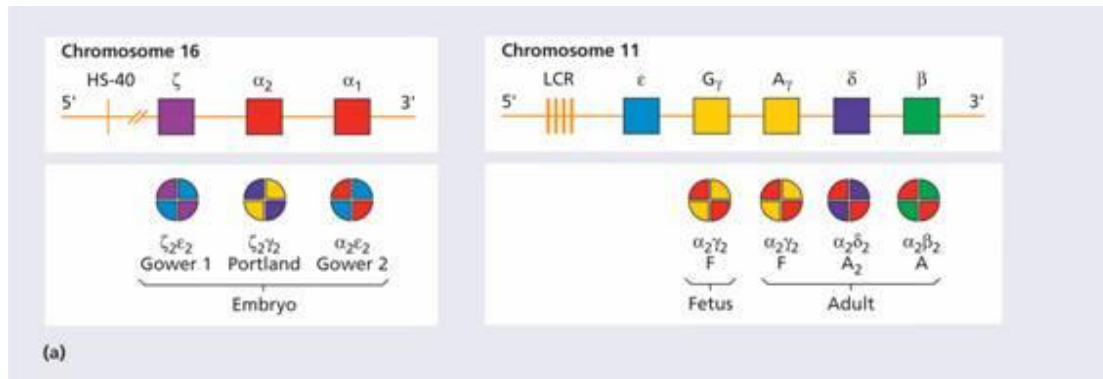
Power 1	$\zeta_2\epsilon_2$		$\alpha_2\beta_2$	HbA
Power 2	$\alpha_2\epsilon_2$	Hb F	$\alpha_2\gamma_2$	$(\alpha_2\delta_2)$ HBA2
	$\zeta_2\gamma_2$		$(\alpha_2\beta_2)$	$(\alpha_2\gamma_2)$
	Embryonic	Fetal	Adult	





Familia genelor alfa globinei: 4 gene funcționale: zeta – ζ (funcțională numai în perioada embrionară), alfa 2 – α_2 și alfa 1 - α_2 (ambele se exprimă din perioada fetală și rămân active toată viața), teta – θ (produsul său nu a fost identificat încă în condiții fiziologice)

Familia genelor beta globinei: 5 gene funcționale: epsilon – ϵ (activă numai în perioada embrionară), gamma G – γ^G și gamma A – γ^A (active în perioada fetală), delta – δ (activă după naștere) și beta – β (activă după naștere)



Valori normale Hb adult:

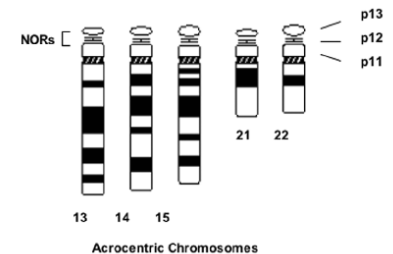
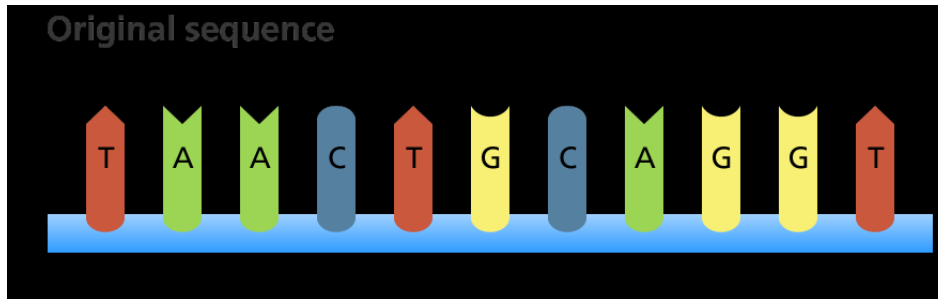
HbA = 97-98%

HbA2 = 2-3%

HbF < 1%.

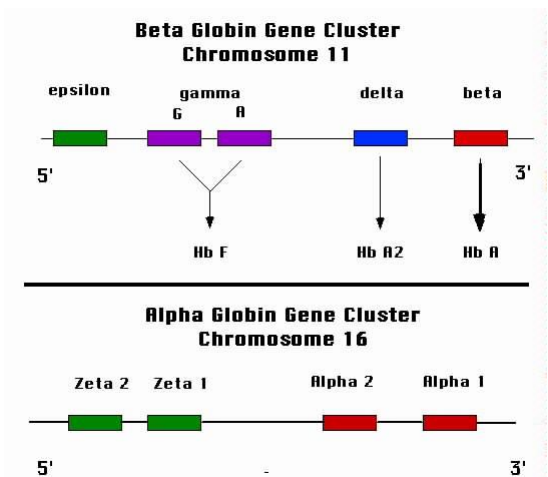
GENE UNICE ȘI FAMILII DE GENE

1) **Gene unice** - Sunt gene ce se găsesc în 2 copii în genomul diploid



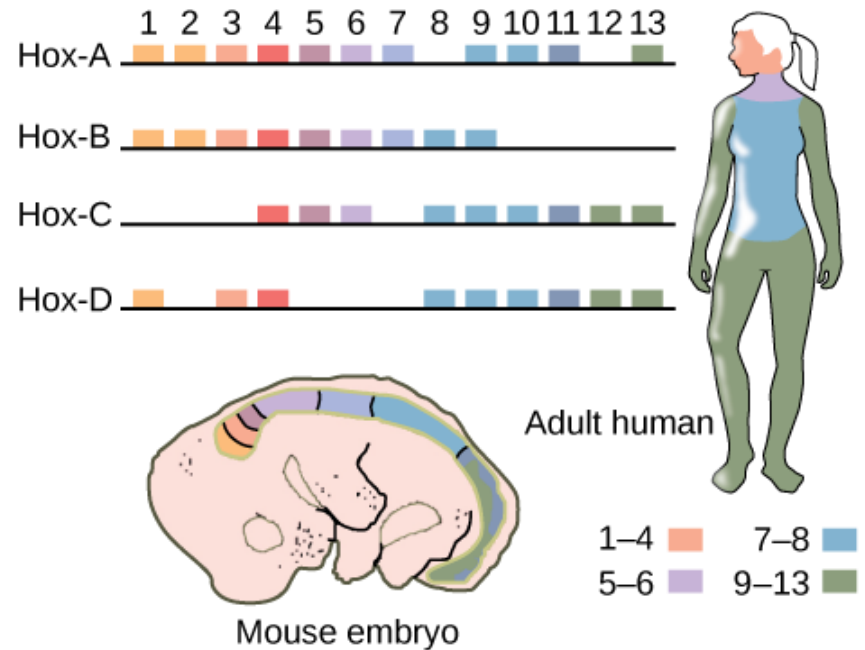
2) **Familii de gene** - gene cu structuri asemănătoare

- familii clasice- secvența cuasiidentică sau foarte asemănătoare, origine comună (familia genelor pentru histone și cea a genelor ARNr, familia alfa sau beta-globinei)



- familii de gene ce codifică proteine cu domenii mari conservate

(HOX,PAX)



- superfamilii de gene – număr mare de copii genice, structuri și funcții diferite intervin procese corelate, mai păstrează o secvență structurală identică (superfamilia genelor imunității)

Elemente genetice mobile

Transpozoni



Barbara McClintock



Transpozoni – sunt secvențe de ADN care sunt capabile să își schimbe localizarea în genom, fără să apară însă niciodată în formă liberă și individualizată în celulă.

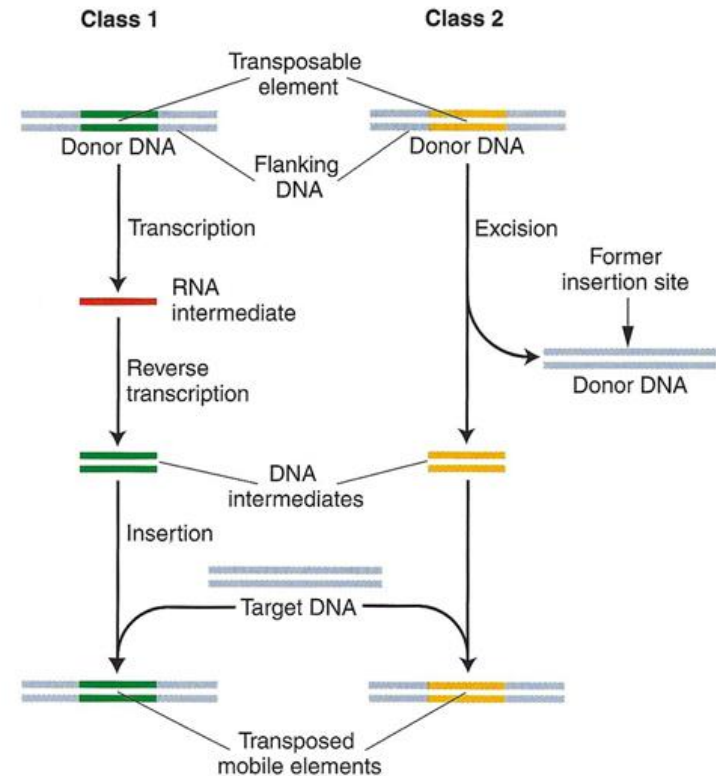
Deplasarea unor secvențe de ADN dintr-o poziție în alte poziții în genomul unei celule a fost numită transpoziție, iar secvențele respective elemente genetice mobile, elemente transpozabile sau transpozoni.

|Transpozoni (Tn) – 2 tipuri în funcție de mecanismul deplasării:

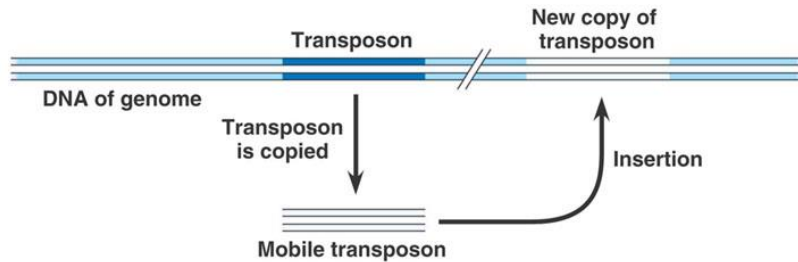
- Tn I sau retrotranspozoni
- Tn II

Tn de clasă I = retrotranspozoni, retroelemente care se copiază ei înșiși mai întâi sunt transcriși în molecule de ARN, apoi sub acțiunea unei **transcriptaze inverse** (codificată adesea de Tn respectiv) aceste molecule vor forma copii complementare (de tip ADNc) care se vor insera (sub acțiunea unei **endonucleaze**) într-un alt loc în genom.

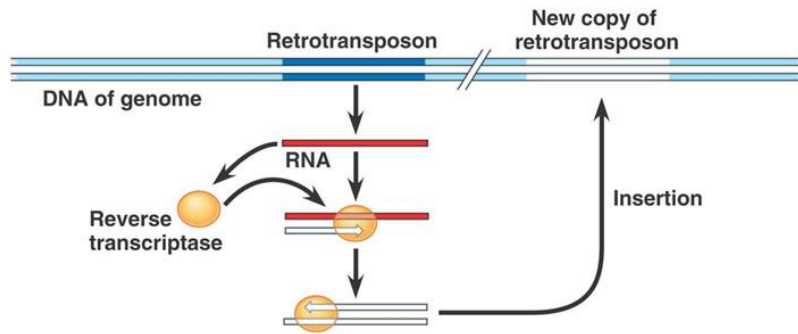
Two classes of mobile elements



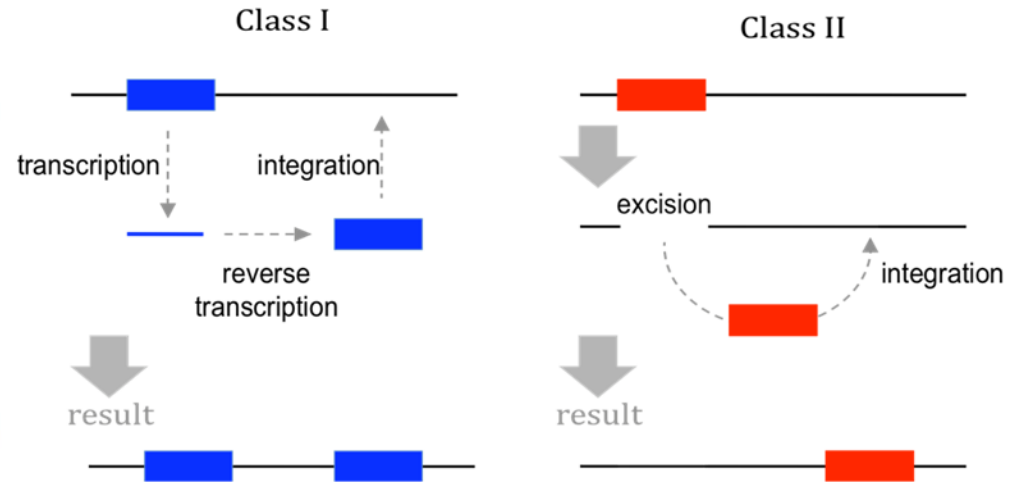
Tipuri de transpozoni



(a) Transposon movement ("copy-and-paste" mechanism)



(b) Retrotransposon movement

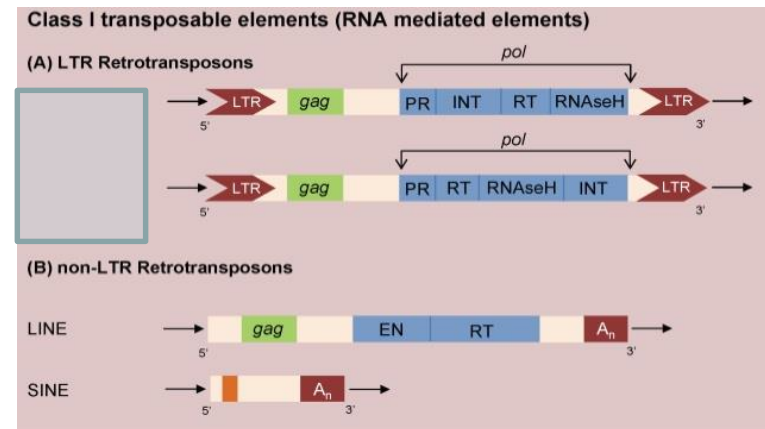


Tn de clasă II se deplasează direct dintr-o poziție în alta folosind o enzimă (**transpozaza**, codificată de Tn) care taie ADN în alt loc din genom; Tn se inseră în această breșă și apoi capetele sale se sudează la ADN (prin acțiunea **rezolvazei**, o altă enzimă codificată de Tn).

Transpozonii produc o clasă distinctă de ADN repetitiv care a determinat expansiunea genomului organismelor complexe (formând circa 45% din genomul uman).

Există 4 tipuri de elemente transpozabile

- **elemente nucleare dispersate lungi (LINEs)** – lungime 0,5 - 6kb, în număr de 850 000, 21%
- **elemente nucleare dispersate scurte (SINEs)** – lungime 100- 300 pb, în număr de 1.7 mil, 13% (exemplu familia Alu)
- **retrotranspozoni LTR** , asemănători retrovirusurilor, lungime 0,2-5 kb, în număr de 700 000 (8%)
- **transpozoni fosili (3%)**



Rol - modificarea structurii și funcției unei gene, transferul unor exoni sau copii de gene