

OBTENCIÓ DELS FITXERS NECESSARIS:

Els fitxers els obtindrem d'una banc de dades biològiques, concretament del "National Center for bioTechnology Information" (referència bibliografia [NCBI]).

De la base de dades, com s'observa a la figura 5.18, ens podrem baixar diferents tipus de fitxers, en el nostre cas ens interessaran els fitxers ".gbk". Aquests fitxers donen molts tipus d'informació sobre el cromosoma de l'espècie on ens trobem. Tot seguit veurem què ens interessa al nostre cas.

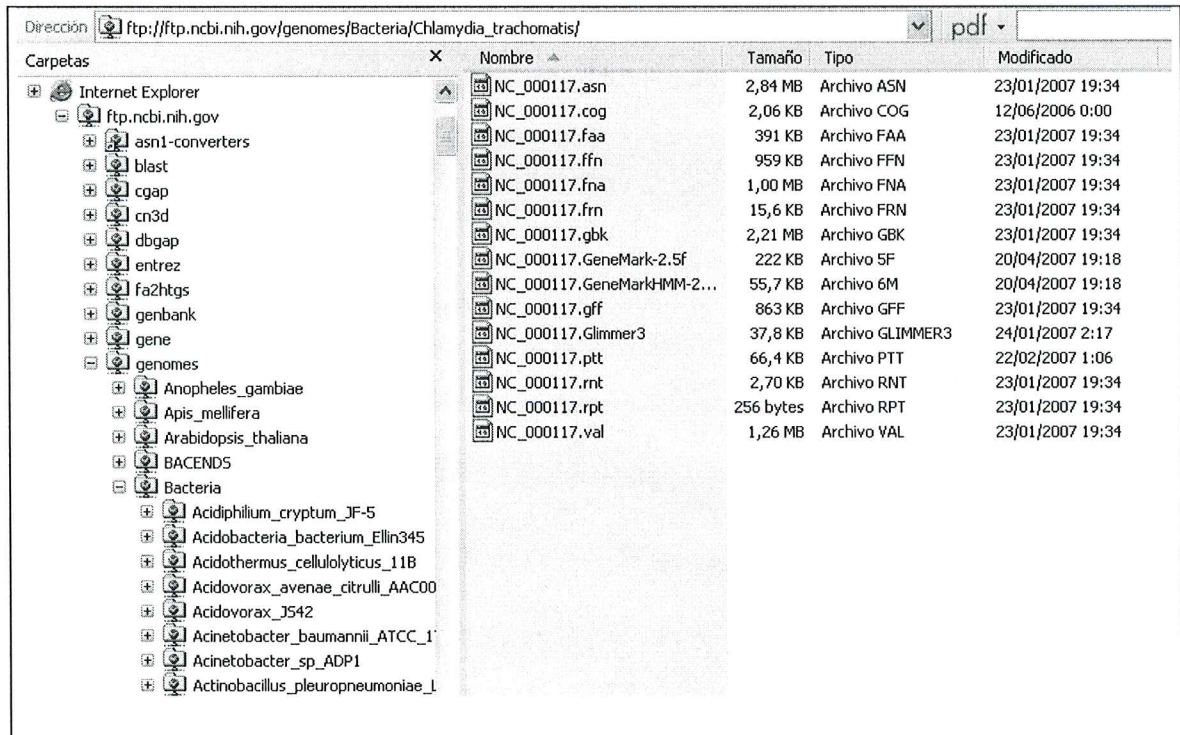


Figura 5.18 Obtenció de fitxers del NCBI

Podem observar a la figura 5.19 quin aspecte pot tenir un fitxer ".gbk".

```
LOCUS   NW_926573           26826 bp   DNA   linear   CON 29-AUG-2006
gene    complement(<2998..>3194)
        /gene="MRPL14P1"
CDS     join(3000..3010,3050..3100,3150..3190)

        gene    complement(<3998..4294)
                /gene="MRPL14P2"
                /pseudo
                /db_xref="GeneID:359735"
                /db_xref="HGNC:29696"
```



```

    gene      complement(4998<..5294>)
              /gene="MRPL14P3"
              /pseudo
ORIGIN
    26821 ctgtgt
//

LOCUS      NW_926584      15369037 bp  DNA   linear  CON 29-AUG-
2006
    geneTRET      5443..8936
    gene      5443<..8936>
              /gene="SCGB1C1"
    CDS      join(5463..5517,6075..6274,6781..6813)
              /gene="SCGB1C1"
    CDS      join(6463..6517,7075..7274,7781..7813)
              /gene="SCGB1C1"
    CDS      join(7463..7517,8075..8274,8781..8813)
              /gene="SCGB1C1"

    gene      complement(14403..54778)
              /gene="DOC2B"

    CDS      complement(join(15036..15172,15768..15864,16706..16787,
37468..37547,46046..46418))
              /gene="DOC2B"

ORIGIN
    15369001 gcactgcacc tggctggaag attttttttt tcttttt
//

LOCUS      NW_926595      24822 bp  DNA   linear  CON 29-AUG-2006

    gene      18563..24784
              /gene="TRIM16"

    CDS      join(18836..19394,20406..20470,24287..24405,24679..24784)
              /gene="TRIM16"

ORIGIN
    24781 ctgagttcta ttagggcata atgactcgca gctttaatg ca
//

(Exemple fet, amb coherència, però amb trossos)
[i mes del mateix ...]

```

Figura 5.19 Exemple (trossos) d'un fitxer ".gbk"

INFORMACIÓ QUE EMMAGATZEMAREM:

En el nostre cas ens interessarà el següent:

- El tamany del cromosoma (bp)
 - Marcat per un "LOCUS"
 - El fitxer ens dóna tamany de trossos, els comptabilitzarem tots
- El contingut de bases concretes del cromosoma
 - Ens les acostuma a donar també per trossos, segons els tamany esmentats al punt anterior.
 - Els anirem unint i en finalitzar el parseig conformaran el fitxer de sortida ".fa"
- Els gens que conté, i per a cadascun d'ells
 - el seu nom
 - la seva ubicació
 - rang relatiu dins el cromosoma, que usant els tamany dels trossos, passem adequadament a rang absolut
 - el rang serà del tipus:
posicio_inici_rang ... posicio_fi_rang
 - l'ubicació dels seus exons
 - identificats pel mot "CDS"
 - tindrem tant rangs com exons contingui el gen
 - si no tenim cap exó, el "CDS" estarà buit
 - podem tenir més d'un "CDS", en tal cas ens guardarem totes les variants.
- Tota la informació relativa als gens l'emmagatzemarem a un fitxer ".ptt"
- El fitxer ".ptt" serà usat quan calgui per la funcionalitat de l'apartat 5.3

A la figura 5.20 hi tenim un tros d'exemple on podem observar l'aspecte d'un fitxer ".ptt".

Cromosoma hs_chr17

NOM GEN	RANG GEN	RANG CDS
MRPL14P1	2998..3194	
SCGB1C1	32269..33762	32289..32343,32901..33100,33607..33639
LOC730597	35414..37511	35414..35549,35648..36022,37423..37511
DOC2B	41229..81604	41862..41998,42594..42690,43532..43613
LOC730605	89699..90865	89699..90865
LOC730612	98313..100773	98313..98631,98972..99254,99440..99536
FAM101B	336671..340362	338340..338670,340265..340362
FAM57A	592254..602489	592295..592416,592745..592828,597491..597692
GEMIN4	604272..611817	604520..607686,611787..611796
C17orf25	618984..642010	619827..619892,629575..629661,629750..629863
RNMTL1	641976..652203	
642082..642395,642785..643029,647667..647834		
NXN	659048..838917	660652..660834,663385..663509,664771..664950
LOC730649	817575..817967	
ABR	862428..1039300	864987..865076,865942..866089,868422..868527

Figura 5.20 Exemple (tros inicial) d'un fitxer ".ptt"

A la figura 5.21 hi tenim un tros d'exemple on podem observar l'aspecte d'un fitxer ".fa".

```
ttctcttgactcctgagagttccatgactaagacagtagtttaatatagttggttccct
ttgtcatcctgcatagtttagttttacgcatttcaaaccattcttcagagaaaggtccct
tggcttaccaggctgctgtgggtccaaggcacgaaagcatgaagagtctctgctcatcg
ccagaggactgtgagtccaaacaggctccctgtagctgaacaggcttcagccttctct
gggtaatcctaggctgtgggttacaacgggtggctttggctgactgggaggcaggagatct
gggtttttattctacctccgccactcccatgccagccctctgtgtcctcagctgtgaaat
gtaagggttgactagccctacgcacttttgattccttttttttttaagacagagctct
tgctctgtcaccaggctggagtgagtgacggccttggctcattgcaacttccaact
cccagggtcaagcaattctcctacctcagcctcccagtagctgggactacagggtccca
ccatcacgcctggctaatttttgatatttttagtacagacaagggttcatcatgttggcca
agctggtcttgaactcctgacctcaggcgatctgcctgtctcggcctccaaagtgtgg
gattacagggtgtgagccaccgacccagcattttctgtgttatatgtgaaatgtagcaa
tatgcaaacagcacatacaatgtctgggtttggtttactaaagactaacgagacggctg
```

Figura 5.21 Exemple (tros inicial) d'un fitxer ".fa"

ACTUALITZACIÓ D'INFORMACIÓ:

Finalment només ens quedarà actualitzar la informació seguint l'algorisme indicat a l'inici d'aquest apartat.

Per això caldrà que, des de la zona de la web on hem afegit la funcionalitat (des del "mrepatt.pl" que detectarà que estem en mode "actualització"), faci el següent:

- cridar el fitxer executable perl "compute.pl"
- amb els paràmetres necessaris perquè faci el "make_all" (recàlcul de tot)
- i ho faci amb tots els patrons bàsics (no li passarem cap), que seran, per defecte, totes les combinacions de màxim 4 bases.

5.3. Detecció de zones cromosòmiques:

La Detecció es basa en l'esquema que podem observar a la figura 5.22

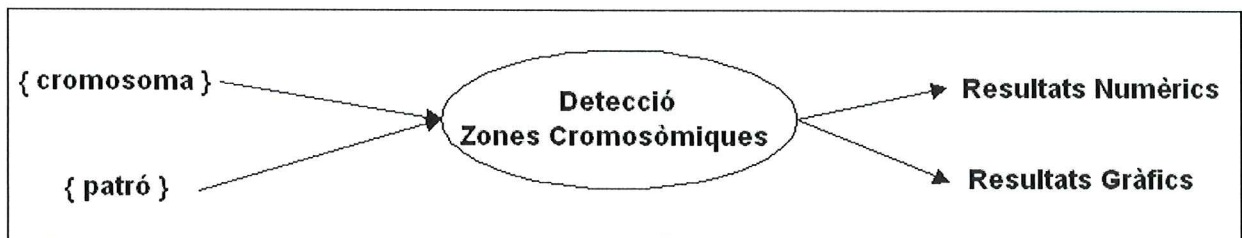


Figura 5.22 Esquema Bàsic Detecció de Zones Cromosòmiques

Com podem observar com a entrada hi tenir un o més cromosomes i un o més patrons. I la sortida seran resultats tant numèrics com gràfics de les mateixes dades, seguint la mateixa tipologia de format que explicàvem a l'apartat 3. sobre la versió anterior.

Els càlculs els fem individualitzats, seguint l'esquema de la figura 5.23 per a cada parella cromosoma-patró, i després els unim obtenint els resultats totals.

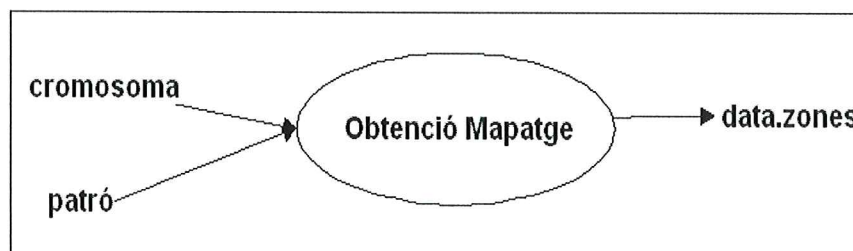


Figura 5.23 Esquema Bàsic d'Obtenció Informació per Mapatge Cromosòmic

Respecte als resultats per a cada parella individualitzada, s'emmagatzemen en fitxers de text com el que podem observar a la figura 5.24. Es tracta d'un fitxer generat.

El contingut del fitxer és:

- cada fila ens indica el nombre de repeticions consecutives
 - aniran de 1 fins al màxim que s'hagi trobat
- cada columna ens indica:
 - número de repetició actual (ex: '2', cerquem pel patró P, on hi ha PP)
 - total de repeticions
 - quantes es troben a zona exònica
 - quantes es troben a zona intrònica
 - quantes es troben a zona intergènica (fora de cap gen)
 - quantes es troben a zona gènica (dins un gen, però no podem saber si intró o exó)

1	26546	0	26000	546	0
2	346	220	100	0	26
3	2	0	0	2	0

Figura 5.24 Exemple fitxer "data.zones" generat per un cromosoma-patró

Per a obtenir aquestes dades ens cal:

- Mentre recorrem el fitxer d'entrada del cromosoma corresponent (observem figura 5.23), recorrem també el fitxer ".ptt" que conté els rangs on es troben els gens i els seus possibles exons/introns
 - tenint en compte que ambdós estan ordenats, i que anem guardant-nos l'última posició visitada del "ptt", ens assegurem que el recorregut és màxim de cost $n+m$ (on n : tamany del cromosoma, i m : tamany del fitxer ptt)
- Per a cada posició on es detecti un patró allò que fem és comprovar si aquesta posició pertany a una o altra zona del cromosoma, això ho fem segons l'esquema que es mostra a la figura 5.25

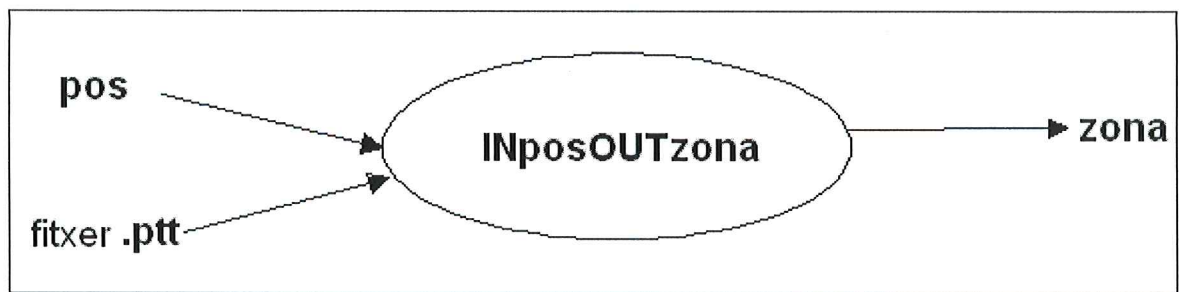


Figura 5.25 Esquema Bàsic obtenció de la zona per a una posició donada

Observem que és en aquest punt on usem el fitxer ".ptt" amb la informació per al marcatge del cromosoma donat.

6. Planificació i Estudi econòmic:

Tasques a desenvolupar. Fases:

El projecte l'hem dividit en les següents fases:

Fases inicials:

- **Fase 1:** Estudi de Bioinformàtica, centrant-nos en allò que ens caldria per al desenvolupament del projecte.
- **Fase 2:** Anàlisi de l'eina web generada i de l'estructura de funcionament de la versió *mRepatt* de la que partíem.

Fases centrals (totes inclouen: anàlisi + implementació + predocumentació)

- **Fase 3:** Desenvolupament de l'eina que obtindrà la informació necessària per a crear el mapa cromosòmic.
- **Fase 4:** Desenvolupament de l'eina que, donada una posició absoluta d'un cromosoma donat, en indicarà a quin tipus de zona pertany dins el mapa cromosòmic corresponent.
- **Fase 5:** Desenvolupament de la funcionalitat complerta que creua la informació del mapa cromosòmic (que s'obté de la fase 3) amb la informació generada per la versió anterior de l'eina (explicada a l'apartat 3).
- **Fase 6:** Desenvolupament de la funcionalitat complerta que permet l'actualització de cromosomes amb l'obtenció dels càlculs necessaris i emmagatzemament.

Fases Finals:

- **Documentació:** Desenvolupament de la documentació final del projecte, tenint en compte en tot moment les petites predocumentacions fetes durant les 2 a 6 del projecte.
- **Presentació:** Preparació de la presentació.

Planificació en el temps:

Tot seguit mostrem com ha estat el desenvolupament en hores invertides i tasques dutes a terme indicant-ne els mesos corresponents.

Comencem mostrant la distribució mensual d'hores als esquemes següents:

<u>Mes:</u>	Novembre	Desembre	Gener	Febrer
<u>Hores:</u>	40	30	70	80

<u>Mes:</u>	Març	Abril	Maig	Juny
<u>Hores:</u>	80	125	100	190

I finalment, indiquem les fases dutes a terme al llarg d'aquests mesos als següents esquemes:

<u>Mes:</u>	Novembre				Desembre				Gener				Febrer						
<u>Setmana:</u>	S6	S13	S20	S27	S4	S11	S18	S25	S1	S8	S15	S22	S29	S5	S12	S19	S26		
<u>Hores:</u>	10	10	10	10	10	10	5	5	10	15	15	15	15	20	20	20	20		
<u>Tasques:</u>																			
Fase 1	■																		
Fase 2					■														
Fase 3																			
Fase 4																			
Fase 5																			
Fase 6																			
Docum.								■										■	
Present.																			

<u>Mes:</u>	Març				Abril					Maig				Juny					
<u>Setmana:</u>	S5	S12	S19	S26	S2	S9	S16	S23	S30	S7	S14	S21	S28	S4	S11	S18	S25		
<u>Hores:</u>	20	20	20	20	25	25	25	25	25	25	25	25	25	40	40	50	60		
<u>Tasques:</u>																			
Fase 1																			
Fase 2								■										■	
Fase 3	■																		
Fase 4	■																		
Fase 5					■														
Fase 6														■					
Docum.														■					
Present.																			

Estudi econòmic:

Mostrem el total d'hores invertides per a cada fase, el cost econòmic per fase i, finalment, el cost total que suposa el desenvolupament del projecte.

Hem pres com a preu/hora, mitja entre el que pot ser d'analista i de programador, donat que ambdues tasques s'han anat duen a terme aproximadament en la mateixa proporció, el de 25 euros/hora.

El resultat el tenim a la taula següent:

	Hores invertides	Cost econòmic
Fase 1	60	1500
Fase 2	160	4000
Fase 3	50	1250
Fase 4	20	500
Fase 5	215	5375
Fase 6	80	2000
Documentació	70	1750
Presentació	30	750
TOTAL	685 h	17.125 euros

7. Glossari de termes:

Contingut glossari provinent de la bibliografia:

- Ref. per a definicions: [fBIO], [UDBM_glo], [DITM], [DEMed]
- Ref. per a definicions i terminologia català: [termCat]
- Ref. per a definicions i terminologia anglès: [TGGT], [OLDS], [MP_ME]

Adenina (*adenine*)

L'adenina és una de les quatre bases químiques de l' ADN, i també de l' ARN, i en el codi genètic es representa amb la lletra A. L'adenina sempre s'aparella amb la timina, en el cas de l' ARN amb l'uracil.

ADN Àcid desoxiribonucleic (*DNA deoxyribonucleic acid*)

Component químic dins el nucli de la cèl·lula que duu les instruccions per a elaborar els organismes vivents.

ADN polimerasa (*polymerase DNA*)

Enzim que duu a terme la síntesi de DNA a partir d'un motlle d' ARN.

Al·lel (*allele*)

Cadascuna de les formes alternatives d'un gen que ocupen el la mateixa posició a cromosomes homòlegs, però que es diferencien en quan a la seva seqüència nucleotídica i, per tant, poden donar lloc a proteïnes amb diferent funció, algunes de les quals poden estar implicades en processos patològics. Produeixen efectes diferents sobre un mateix caràcter hereditari.

Alineament de seqüències (*sequence alignment*)

Col·locació mútua de dues o més seqüències, que mostra on aquestes són similars i on difereixen. Un alineament òptim és aquell que ens mostra major quantitat de correspondències i menor quantitat de diferències.

ARN Àcid ribonucleic (*RNA ribonucleic acid*)

Component químic resultant de la transcripció de l'ADN. Les seves bases són adenina, guanina, citosina i uracil (aquesta última substitueix la timina de l'ADN). A les cèl·lules hi trobem tres tipus d' ARN: l'ARN missatger, l'ARN de transferència i l'ARN ribosomal.

ARN missatger, ARNm (*messenger RNA, mRNA*)

Molècules d'ARN que duen la informació continguda als gens fins al ribosoma, on té lloc la traducció. Es forma al nucli de la cèl·lula mitjançant el procés de transcripció de l'ADN.

ARN polimerasa (*polymerase RNA*)

Enzím que intervé en la síntesi d'ARN (concretament ARN missatger) a partir d'un motlle de ADN.

ARN ribosomal, ARNr (*ribosomal RNA, rRNA*)

Qualsevol de les molècules d'ARN que formen part de l'estructura del ribosoma.

ARN de transferència, ARNt (*transference RNA, tRNA*)

Qualsevol de les molècules (relativament petites) d'ARN que intervien en la correcta inserció d'un aminoàcid a la cadena peptídica naixent durant el procés de traducció.

Biosíntesis de proteïnes (*protein biosynthesis*)

Procés de síntesi de proteïnes a la cèl·lula. La biosíntesis de proteïnes avarca la transcripció de la informació continguda als gens estructurals (en la forma d'ADN) per a formar els ARN missatgers corresponents. Aquests posteriorment són usats a la síntesi de les cadenes polipeptídiques als ribosomes, com a part del procés de traducció.

Citosina (*cytosine*)

La citosina és una de les quatre bases químiques de l'ADN, i també de l'ARN, i en el codi genètic es representa amb la lletra C. La citosina sempre s'aparella amb la guanina.

Codi genètic (ACTG) (*genetic code*)

Les instruccions contingudes en un gen que li diuen a la cèl·lula com fer una proteïna específica. A, T, G, i C són les "lletres" del codi genètic i representen les bases nitrogenades adenina, timina, guanina i citosina, respectivament. Aquestes bases juntament amb un sucre i un enllaç fosfat constitueixen els nucleòtids que són la unitat fonamental de l'ADN. A cada gen es combinen les quatre bases en diverses formes, per a crear paraules de 3 lletres (tripletes anomenades també codons) que especifiquen quin aminoàcid és necessari en cada pas de l'elaboració de la proteïna.

Codó (codon)

Tripleta de nucleòtids en l'ARN missatger que codifica un aminoàcid particular o que assenjala el començament o la fi de la traducció.

Codó amb sentit

Codó que codifica un aminoàcid.

Codó d'inici

Codó que determina el primer aminoàcid d'una cadena polipeptídica.

Codó de terminació

Codó que assenjala la fi d'un missatge genètic en l'ARN.

Cromosoma (chromosome)

Estructura en forma de bastonet, resultat de l'empaquetament de l'ADN, observable en el nucli durant la divisió cel·lular, que constitueix el suport físic de la informació genètica heretada.

Detecció de Patrons (pattern matching)

Cerca d'un o més patrons determinats dins una seqüència (o seqüències) donades. P.ex. podríem detectar aparicions del patró "acct", dins la seqüència de nucleòtids: "aaaacggtggcagggttgcaaacctacctgggacct".

DNA

Terme anglès per a ADN.

Escissió (splicing)

Procés mitjançant el qual els introns son eliminats del ARN missatger, que ha estat transcrit a partir d'un gen discontinu. En alguns casos també s'eliminen exons (de forma controlada) donant lloc a diferents productes proteics que deriven del mateix gen, en un procés que es coneix com escissió o splicing alternatiu.

Enzim (enzyme)

Macromolècula, proteïna o ARN, que intervé en el metabolisme dels éssers vius accelerant la velocitat de les reaccions químiques cel·lulars (el que s'anomena catalitzar). Una de les característiques més importants dels enzims és l'especificitat, és a dir, la propietat que tenen d'actuar només sobre una substància o un grup de substàncies.

Exó (exon)

Cadascun dels segments d'un gen que està representat en l'ARN missatger madur i que, per tant, codifica aminoàcids.

FASTA (FASTA)

1. (algorisme, eina) Eina desenvolupada per Pearson i Lipman per a l'alineament de seqüències d'àcids nucleics i proteïnes.
2. (format) Un dels formats més simples utilitzats per a emmagatzemar seqüències nucleotídiques o aminoacídiques. Una entrada en format FASTA té dos blocs fonamentals, el primer està format per una sola línia que comença amb '>' i no és més que una descripció de la seqüència, i el segon està format per la seqüència en sí i implica tantes línies com sigui necessari. Per exemple:

```
>NADH dehydrogenase I chain J
MTFQLILFYIFAVIILYGAIKTVTAKNPVHAA
VGAVMVLTFCVSAMLWMLMQAEFLGVTLVVVY
LHLVLFVVMMLNIDIEEMRAGFWRHAPVAG
DIPAVVGTLLAVALILILVNPKTDLAAFGLMK
DYNNIRDLDGSRIYTDY
```

Gen (gene)

Unitat hereditària elemental formada per una seqüència de nucleòtids al llarg d'una molècula de ADN o d'ARN.

Es distingeixen tres tipus de gens: (1) els gens que son transcrits a ARN missatger i traduïts a cadena polipeptídica (coneguts com a codificants o estructurals), (2) els gens que només son transcrits a ARN (que codifiquen per a ARN ribosomal o ARN de transferència), i els gens reguladors de l'expressió de les dues classes anteriors. Els gens estructurals típics de les cèl·lules eucariotes són discontinus, les regions codificadores (exons) estan interrompudes per regions no codificadores (introns) que son eliminades com a part d'un procés conegut com a escissió.

Genoma (genome)

Contingut total d'ADN propi del conjunt de cromosomes d'una espècie.

Guanina (guanine)

La guanina és una de les quatre bases químiques de l'ADN, i també de l'ARN, i en el codi genètic es representa amb la lletra G. La guanina sempre s'aparella amb la citosina.

Informació genètica (genetic information)

Seqüència de nucleòtids d'una molècula de ADN o d'ARN

Intró (intron)

Cadascun dels segments d'un gen que no està representat en l'ARN missatger madur.

Mapatge cromosòmic (*chromosome mapping*)

Identificació de la posició que cada gen ocupa en el cromosoma.

Mutació (*mutation*)

Procés mitjançant el qual el material genètic pateix un canvi detectable i heretable, generalment en forma de variació a la seqüència d'ADN. Aquestes variacions poden ser puntuals (canvi d'un nucleòtid per un altre), d'inserció (guany) o deleció (pèrdua) d'un segment de nucleòtids. Les mutacions són la base de l'existència dels al·lels.

Nucleòtid (*nucleotide*)

Molècula complexa que resulta de la composició d'una molècula d'àcid fosfòric amb una pentosa i una molècula cíclica anomenada base nitrogenada, la qual pot ser: adenina (A), timina (T), citosina (C), guanina (G) o uracil (U).

Patró (*pattern*)

Petita seqüència, habitualment de nucleòtids. S'usa per a dur a terme el que anomenem detecció de patrons dins una seqüència donada. És a dir, un patró podria ser qualsevol seqüència que inclogués "A", "C", "T" o "G", p.ex.: "acct".

Plegament (*folding*)

El procés a través del qual una cadena polipeptídica (proteïna) adquireix una estructura tridimensional específica i única.

Proteïna (*protein*)

Macromolècula composta per una o més cadena polipeptídiques, cadascuna amb una seqüència característica d'aminoàcids enllaçats per enllaç peptídic. Les proteïnes són les principals macromolècules que formen part dels organismes vius i són crucials en pràcticament totes les funcions cel·lulars

Replicació d'ADN (*DNA replication*)

Procés pel qual la doble hèlix de l'ADN es desenrotlla i produeix una còpia exacta de sí mateixa.

Ribosoma (*ribosome*)

Agregat supramolecular intracel·lular, format per ARN i proteïnes. El ribosoma és el lloc de la traducció del ARN missatger en el procés de biosíntesis de proteïnes.

RNA

Terme anglès per a ARN.

Seqüenciació (*sequencing*)

Procés de determinació de la seqüència d'aminoàcids d'una determinada proteïna o bé de la seqüència de nucleòtids d'una molècula de ADN o d'ARN.

Seqüència (*sequence*)

Disposició ordinal dels monòmers que formen part de les biomolècules de tipus "polimèric" com els àcids nucleics i les proteïnes. La major part de la informació en el context dels sistemes biològics està representada en la forma de seqüències nucleotídiques (en els àcids nucleics) i aminoacídiques (en les proteïnes). Les variacions en la seqüència nucleotídica dels gens que codifiquen per a proteïnes, degudes a mutacions, poden produir variacions en la seqüència aminoacídica d'aquestes proteïnes. Mentre que, tals variacions en la seqüència aminoacídica de les proteïnes usualment impliquen variacions en la seva estructura tridimensional i per tant, en la seva funció.

Timina (*thymine*)

La timina és una de les quatre bases químiques de l'ADN, en el cas de l'ARN es substituïda per l'uracil, i en el codi genètic es representa amb la lletra T. La timina sempre s'aparella amb l'adenina.

Transcripció (*transcription*)

Procés mitjançant el qual la informació continguda en una de les dues branques entrelaçades que formen la doble hèlix d'una molècula de ADN és usada com motlle per a la síntesi d'una cadena d'ARN, amb una seqüència complementària a la de la cadena de ADN que serveix de motlle

Traducció (*translation*)

El procés mitjançant el qual la informació continguda a una molècula d'ARN missatger especifica la seqüència d'aminoàcids per a la síntesi d'una cadena polipeptídica, durant el procés de biosíntesis de les proteïnes.

Uracil (*uraci*)

L'uracil és una de les quatre bases químiques de l'ARN, a l'ADN el seu lloc l'ocupa la timina, i en el codi genètic es representa amb la lletra U. L'uracil sempre s'aparella amb l'adenina.

Biologia:

- [Aprot] Apunts de proteïnes
<http://es.geocities.com/tonicasany/proteines.html>
- [Aanuc] Apunts d'àcids nucleics
http://es.geocities.com/tonicasany/a_nucleics
- [aCel] Apunts sobre la cèl·lula
<http://www.xtec.es/~jgurrera/index.htm>
- [BgtoMB] Beginner's Guide to Molecular Biology
<http://www.rothamsted.bbsrc.ac.uk/notebook/courses/guide/>
- [bioI] Bioinfo: Bases de datos, tutoriales, glosarios
<http://www.cecalc.ula.ve/bioinformatica/BIOTUTOR/bioinformatica.html>
- [bioMod] BioModel: Pàgines de Bioquímica i Biología Molecular
<http://www2.uah.es/biomodel/biomodel-misc/anim/prot/estruc.html>
- [CB_Anim] Cell Biology Animation
<http://www.johnkyrk.com/>
- [DEMed] Diccionari enciclopèdic de medicina
<http://www.grec.net/home/cel/mdicc.htm>
- [DITM] Diccionario ilustrado de Términos Médicos
<http://www.iqb.es/diccio/diccio1.htm>
- [dnaftb] DNA From The Beginning
<http://www.dnaftb.org/dnaftb/1/concept/index.html>
- [fBIO] fBioInformatica:
<http://fbio.uh.cu/bioinfo/glosario.html>
- [GG] Graphics Gallery
<http://www.accessexcellence.org/RC/VL/GG/>
- [MP_ME] Medline Plus - Medical Encyclopedia.
<http://www.nlm.nih.gov/medlineplus/encyclopedia.html>
- [TGGT] National Human Genome Research Institute
Talking glossary of Genetic Terms
<http://www.genome.gov/glossary.cfm>
- [TPBQ] Diagramas Animados de Técnicas y Procesos Bioquímicos
<http://biomodel.uah.es/biomodel-misc/anim/inicio.htm#replic>
- [UCMP] University of California - UCMP Glossary
<http://www.ucmp.berkeley.edu/glossary/glossary.html>
- [UDBM] Unidad Didáctica de biología Molecular
<http://www.arrakis.es/~ibrabida/biologia.html>
- [UDBM_glo] Glosario Unidad Didáctica de biología Molecular
<http://www.arrakis.es/~ibrabida/igglosario.html>
- [UHMB] Harvard University. Molecular and Cellular Biology.
<http://www.iqb.es/diccio/diccio1.htm>
- [UT_MB] University of Texas. Medical Branch. Cell Biology.
<http://cellbio.utmb.edu/cellbio/>

Webs investigació i de Bases de Dades Biològiques:

- [CRG] Centre for Genomic Regulation
<http://pasteur.crg.es/portal/page/portal/Internet>
- [Imim] Institut Municipal d'Investigacions Mèdiques
<http://www.imim.es/>
- [INAB] INAB - Institut Nacional de Bioinformàtica
<http://www.inab.org/>
- [genomesNCBI] Genomes from NCBI
<ftp://ftp.ncbi.nih.gov/genomes/>
- [NCBI] NCBI HomePage - National Center for Biotechnology Information
<http://www.ncbi.nlm.nih.gov/>
- [ssr] www.ccmb.res.in/ssr
- [tandem] <http://tandem.bu.edu/cgi-bin/trdb/trdb.exe>

Útils:

- [Wiki] Wikipedia, la enciclopedia libre
<http://es.wikipedia.org/wiki/Genoma>
- [LlengCat] Recursos Lingüístics en català
<http://llengcat.com>
- [OLDS] One Look Dictionary Search
<http://www.onelook.com/>
- [softCat] Soft Català: traductors, correctors, etc.
<http://www.softcatala.org/>
- [termCat] TERMCAT, Centre de Terminologia
<http://www.termcat.cat/>

