

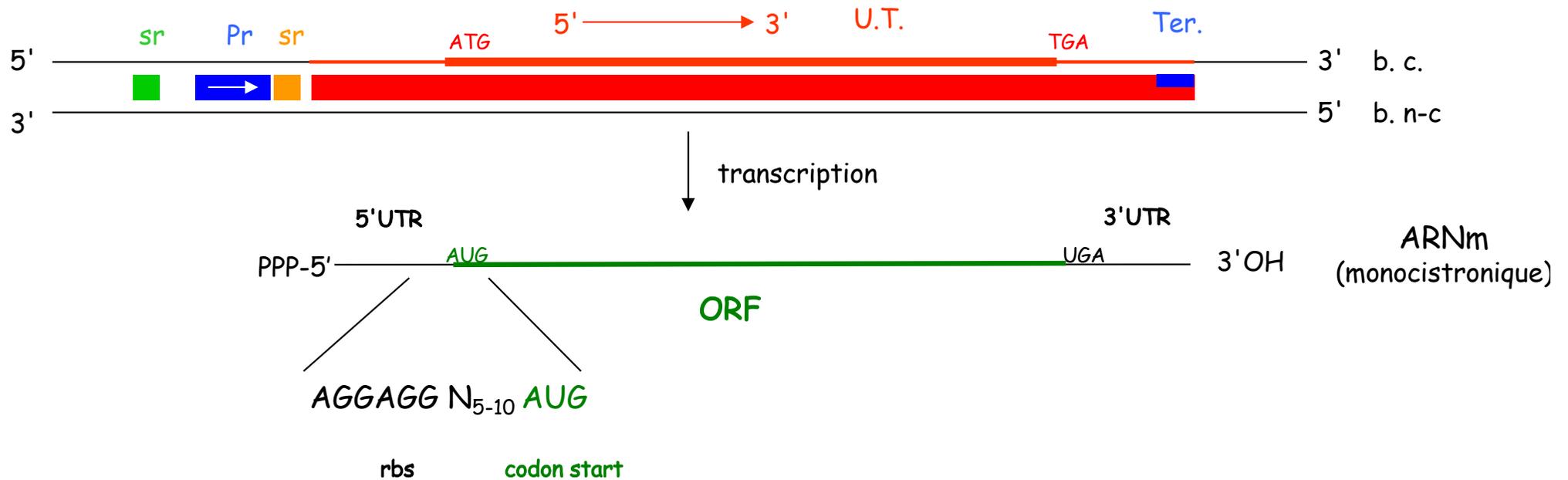
L'organisation des gènes chez les procaryotes et les eucaryotes

(diapositives en support de l'enregistrement podcast à écouter avant la réalisation du travail personnel de génétique)

1. L'organisation typique d'un gène bactérien
2. L'organisation typique d'un gène eucaryote
3. La notion de densité génique
4. Analyse de l'organisation des gènes via le site du **NCBI**
5. Complexité dans l'organisation des gènes chez les eucaryotes supérieurs

1. L'organisation typique d'un gène bactérien

RAPPEL

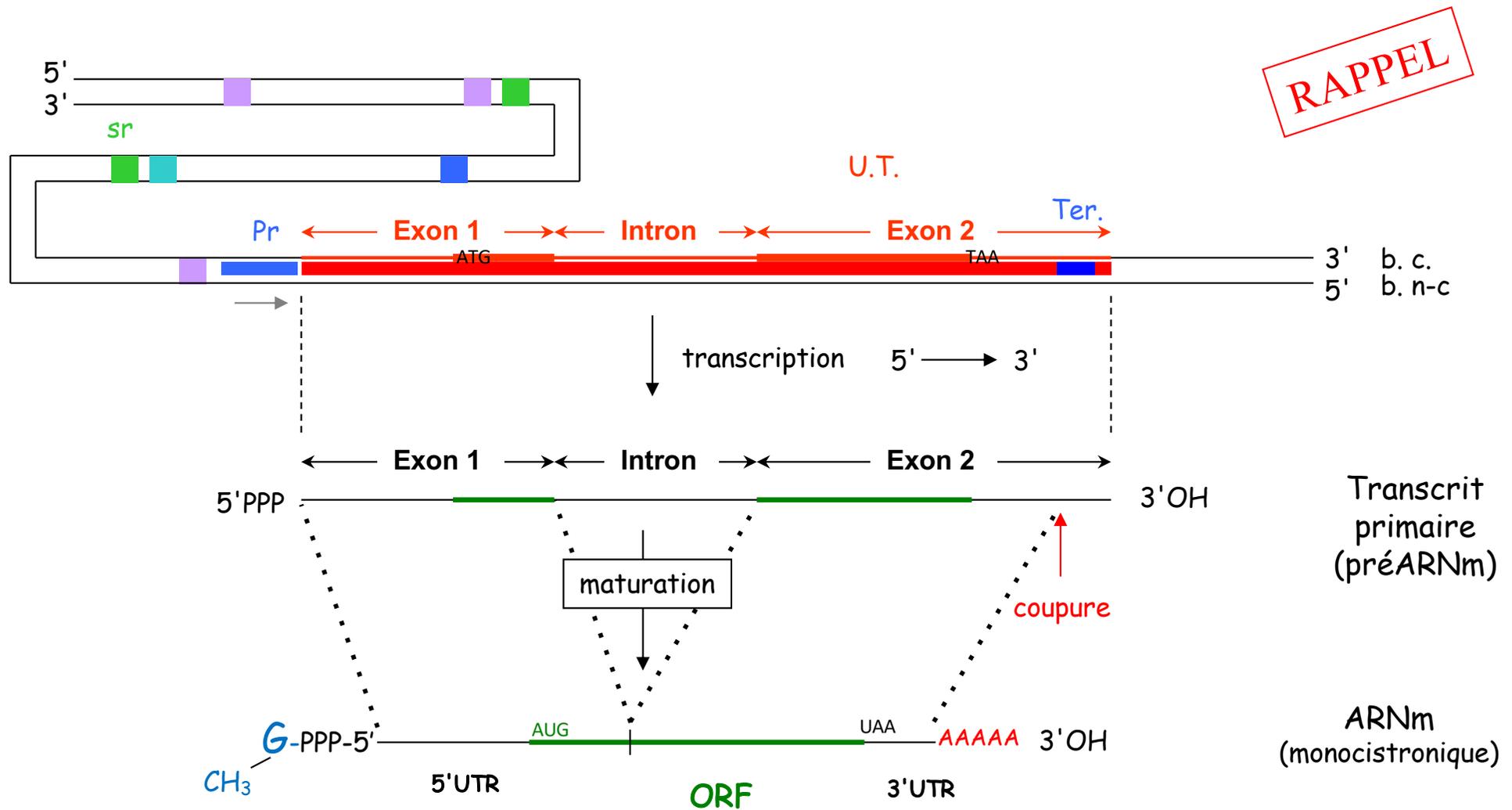


Transcription et traduction sont couplées :

la traduction de l'ARNm par les ribosomes (> polysomes) débute avant même que la transcription du gène ne soit achevée

| | |
|--------|--|
| U.T. | unité de transcription (du gène, de l'opéron) |
| b.c. | brin codant (sens), qui est copié en ARNm |
| b. n-c | brin non-codant (non sens), qui sert de modèle (matrice) pour assembler l'ARNm lors de la transcription |
| Pr | promoteur -> permet la liaison de l'ARN POL ^{ase} + facteurs accessoires |
| Ter. | terminateur -> séquence marquant l'arrêt de la transcription |
| sr | séquence régulatrice (position et nombre variables), permet la liaison des activateurs et répresseurs de transcription |
| ORF | région codante (open reading frame) |
| 5'UTR | région 5' non traduite (5' untranslated region) |
| 3'UTR | région 3' non traduite (3' untranslated region) |
| rbs | ribosome binding site (séq. de Shine-Dalgarno) |

2. L'organisation typique d'un gène eucaryote



L'ATG n'est pas nécessairement situé dans le 1^{er} exon - le 5' UTR peut donc couvrir plusieurs exons.

Maturation : **coiffage (capping)**
polyadénylation
épissage (splicing)
édition ?

3. Notion de densité génique : ex. de la levure et de *H. sapiens*

Densité génique
élevée

Levure :
- la majorité des gènes sont dépourvus d'introns
- les distances intergéniques sont quasi nulles
→ ~70% de l'ADN de levure est codant (→ ORF)

Densité génique
faible

Homme : - la majorité des gènes contiennent des introns

Exons : taille moyenne 290 pb (assez homogène)
nombre d'autant plus grand que la protéine codée par le gène est longue:

insuline : 104 aa 3 EX / 2 IN
dystrophine : 4569 aa 79 EX / 78 IN

Introns : taille très variable :

| Protéine | Taille | Nombre d' Exons / Introns | Taille du gène | Taille moyenne des introns |
|---------------|---------|------------------------------|----------------|-------------------------------|
| β-globine | 146 aa | 3 / 2 | 1.6 kpb | 0.5 kpb |
| myoglobine | 153 aa | 3 / 2 | 10.4 kpb | 4.7 kpb |
| dystrophine : | 4569 aa | 79 / 78 | 2500 kpb | 30kpb |

rem : parfois, des introns contiennent un ou plusieurs autres gènes

- les distances intergéniques sont généralement plus grandes
et certaines régions chromosomiques sont pauvres en gènes

→ seulement ~1,5 % de l'ADN humain
est codant (→ ORFs)

4. Analyse de l'organisation des gènes via le site du NCBI (National Center for Biotechnological Information) : ncbi.nlm.nih.gov

➢ menu à droite *Genome* ➢ *Human genome* → ex. *C7* (= code du gène)

NIH U.S. National Library of Medicine NCBI National Center for Biotechnology Information Login

Human Genome Resources at NCBI

Download Browse View Learn

Search for Human Genes

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y MT

Select a chromosome to access the [Genome Data Viewer](#)

Search results

Items: 1 to 20 of 34

<< First < Prev Page 1 of 2 Next > Last >>

 [See also 3 discontinued or replaced items.](#)

| Name/Gene ID | Description | Location | Aliases | MIM |
|--|---|--|--|--------|
| <input type="checkbox"/> C7 ID: 730 | complement C7 [<i>Homo sapiens</i> (human)] | Chromosome 5, NC_000005.10 (40909497..40982939) | | 217070 |
| <input type="checkbox"/> CXCL10 ID: 3627 | C-X-C motif chemokine ligand 10 [<i>Homo sapiens</i> (human)] | Chromosome 4, NC_000004.12 (76021116..76023536, complement) | C7, IFI10, INP10, IP-10, SCYB10, crg-2, gIP-10, mob-1 | 147310 |
| <input type="checkbox"/> ERVW-1 ID: 30816 | endogenous retrovirus group W member 1 [<i>Homo sapiens</i> (human)] | Chromosome 7, NC_000007.14 (92468380..92477915, complement) | ENV, ENVW, ERVWE1, HERV-7q, HERV-W-ENV, HERV7Q, HERVW, HERVWENV | 604659 |
| <input type="checkbox"/> IGLC7 ID: 28834 | immunoglobulin lambda constant 7 [<i>Homo sapiens</i> (human)] | Chromosome 22, NC_000022.11 (22922594..22922913) | C7 | |
| <input type="checkbox"/> ERVK-6 ID: 64006 | endogenous retrovirus group K member 6 [<i>Homo sapiens</i> (human)] | | ERVK6, HERV-K(C7), HERV-K108, K-Rev, c-orf, cORF | 605626 |



accès à une page d'information sur le gène C7

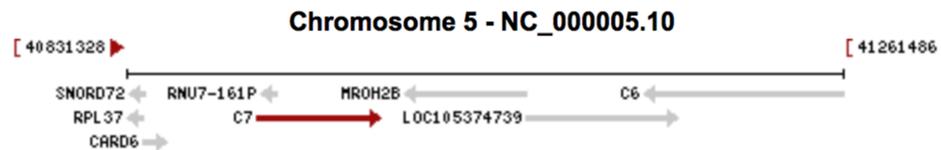
Genomic context

Location: 5p13.1

See C7 in [Genome Data Viewer](#) [Map Viewer](#)

Exon count: 18

| Annotation release | Status | Assembly | Chr | Location |
|---------------------|-------------------|---|-----|-----------------------------------|
| 108 | current | GRCh38.p7 (GCF_000001405.33) | 5 | NC_000005.10 (40909497..40982939) |
| 105 | previous assembly | GRCh37.p13 (GCF_000001405.25) | 5 | NC_000005.9 (40909599..40983041) |



Genomic regions, transcripts, and products

Go to [reference sequence details](#)

Genomic Sequence:

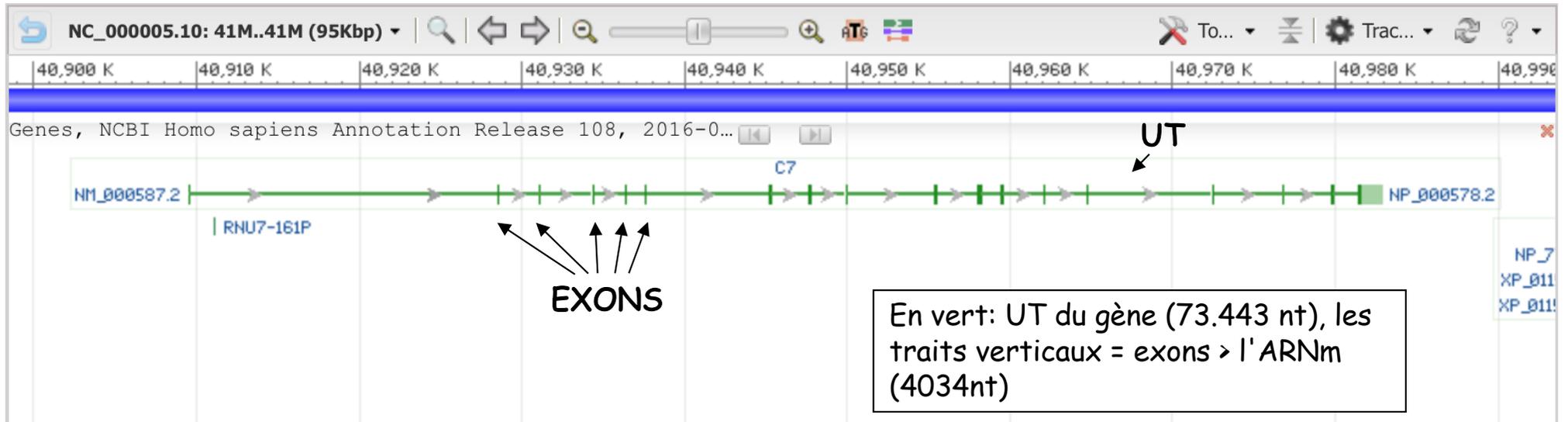
Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)

représentation graphique du gène

séquence du gène

Genomic Sequence: NC_000005.10 Chromosome 5 Reference GRCh38.p7 Primary Assembly

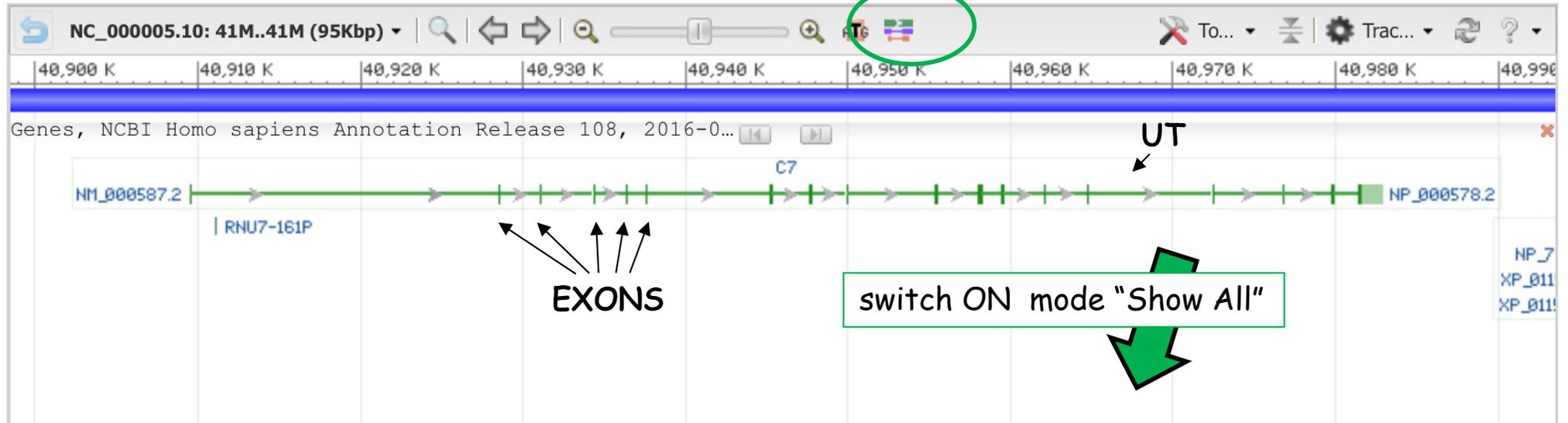
Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)



Genomic Sequence: NC_000005.10 Chromosome 5 Reference GRCh38.p7 Primary Assembly

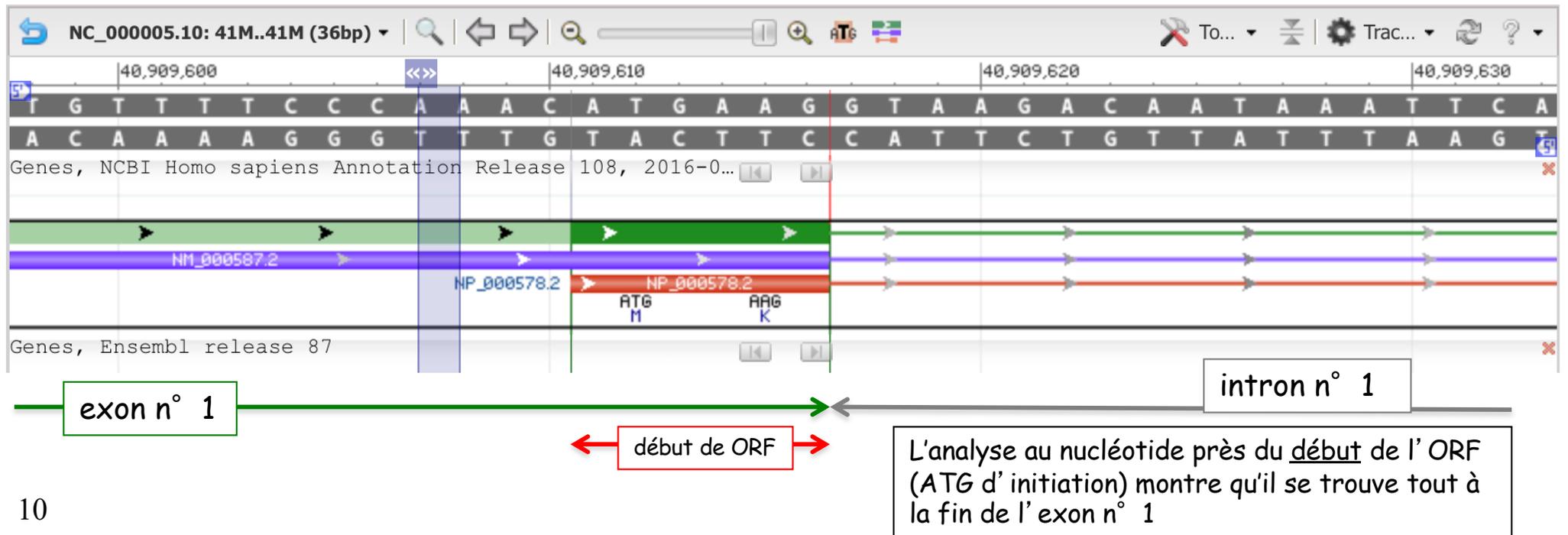
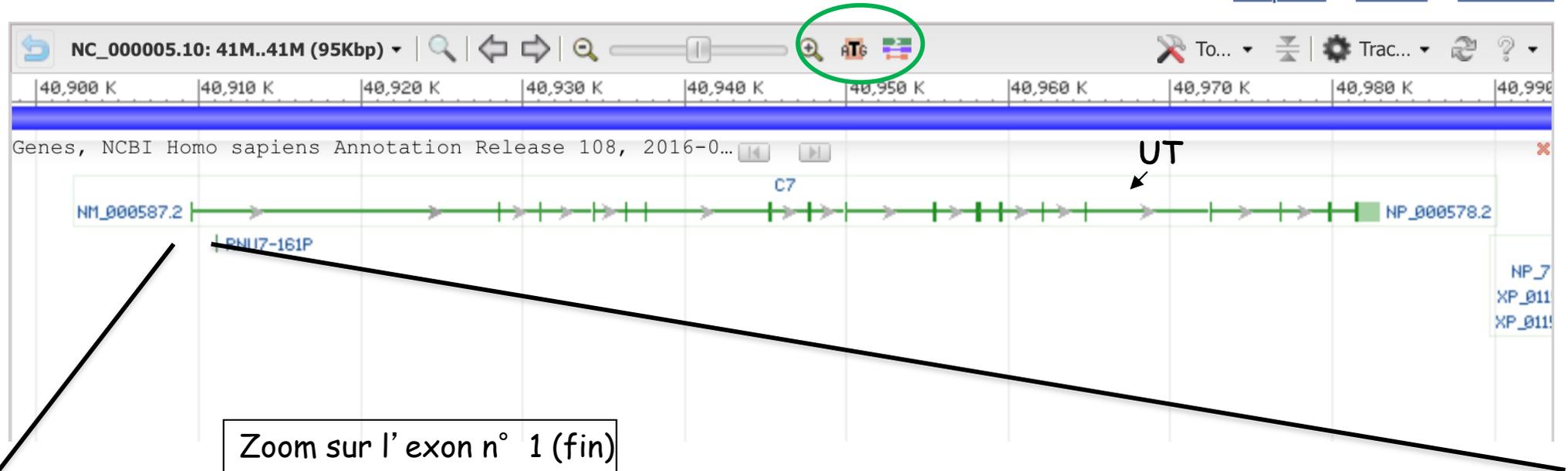
mode "Show All"

Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)

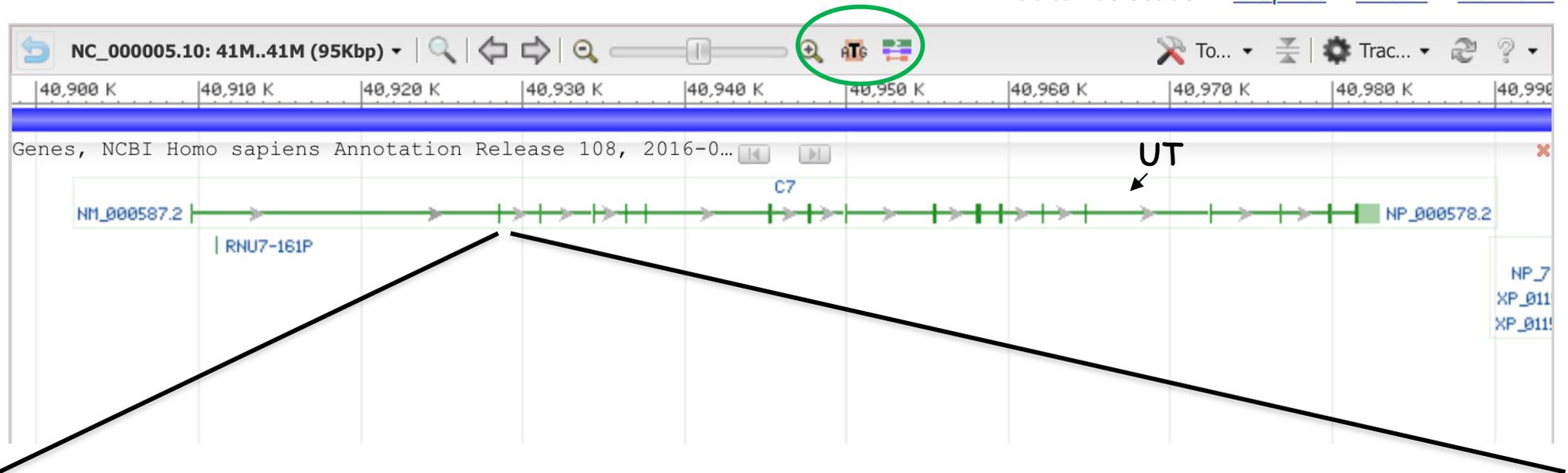


Genomic Sequence: NC_000005.10 Chromosome 5 Reference GRCh38.p7 Primary Assembly

Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)



Genomic Sequence: NC_000005.10 Chromosome 5 Reference GRCh38.p7 Primary Assembly

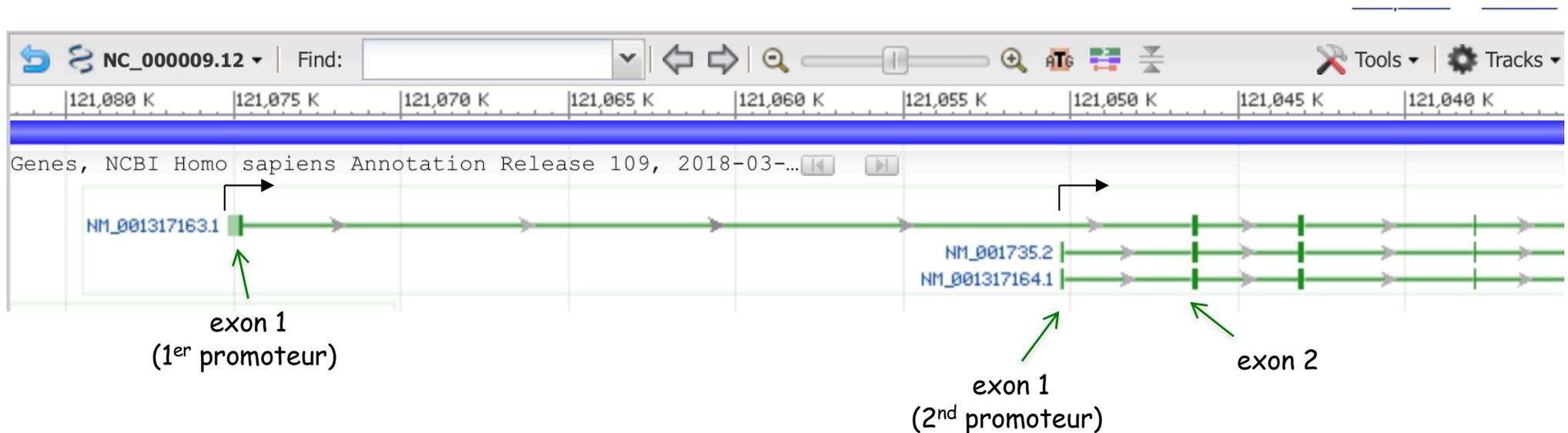
Go to nucleotide: [Graphics](#) [FASTA](#) [GenBank](#)

5. Complexité dans l'organisation des gènes chez les eucaryotes supérieurs

a) Utilisation de promoteurs alternatifs

Pour certains gènes, il y a plusieurs promoteurs possibles -> le choix du promoteur diffère selon les tissus et/ou les types cellulaires

Ex. gène C5 ⇒ 2 promoteurs différents pour ce gène :

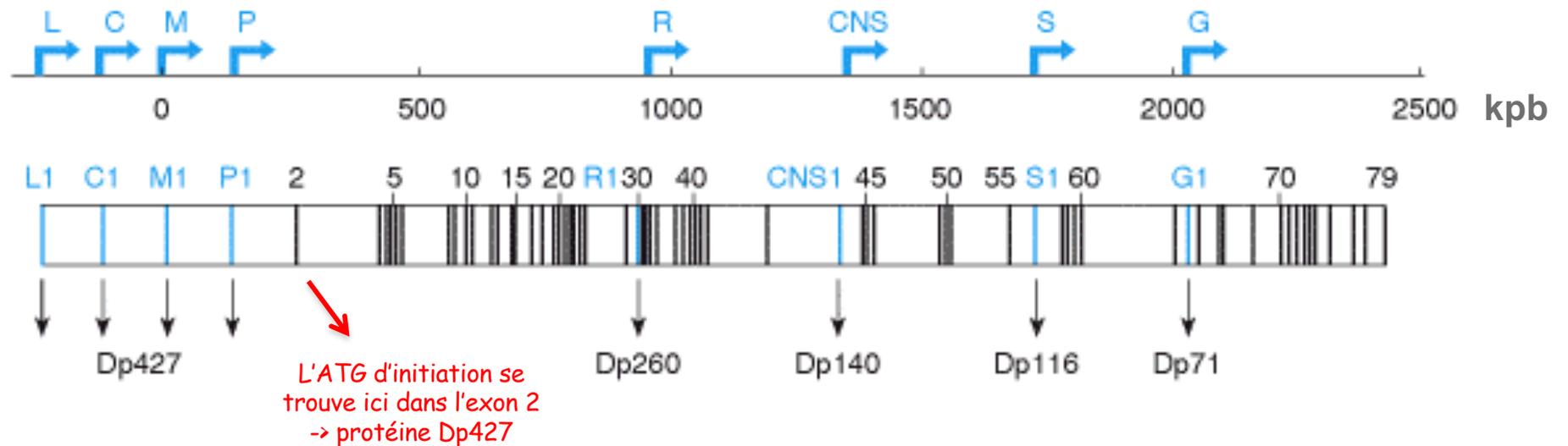


Un premier promoteur est directement suivi d'un exon 1. Cet exon 1, qui contient un ATG d'initiation, sera relié à l'exon 2. Le 2nd promoteur est situé dans l'intron qui sépare ces exons 1 et 2. Ce second promoteur est donc directement suivi d'un autre exon 1. Cet exon 1 alternatif contient lui aussi un ATG initiateur et sera relié à l'exon 2. Les 2 ARNm ainsi formés présenteront des 5' UTR différents, et le début de l'ORF sera aussi différent.

Ex. : gène de la dystrophine (un des plus grands gènes humains, UT = 2.220.382 pb, code = "DMD" dans NCBI)

⇒ 8 promoteurs différents (!)

Représentation simplifiée :



Pas moins de 8 promoteurs sont utilisés pour initier la transcription du gène de la dystrophine, en fonction des tissus ou des types cellulaires : L : lymphocyte; C : cellules corticales; M : muscles; P : cellule de Purkinje; R : rétine; CNS : système nerveux central; S : cellule de Schwann; G : autres tissus. Chaque promoteur est suivi de "son" exon 1 (en bleu : L1, C1, M1, ..) qui est relié à l'exon qui suit (en noir). Sans les tissus L, C, M, et P, chaque exon 1 (L1, C1, ..) est relié aux 78 exons (en noir) qui suivent et cela conduit à la synthèse de la même protéine (de 427 KDa, Dp427) car l'ATG d'initiation se trouve dans l'exon 2 (les exons 1 ne couvrent qu'une première partie du 5' UTR, la 2de est dans l'exon 2). Les autres promoteurs sont situés dans des zones introniques, par ex. celui utilisé dans les cellules de la rétine (R) est dans l'intron qui précède l'exon 30. L'exon R1 est donc relié à l'exon 30 et la dystrophine synthétisée est donc écourtée du côté N-terminal et ne fait que 260 Kda. Le même principe explique pourquoi dans le CNS la dystrophie fait 140 KDa, ... La dystrophine "générale" (G) fait 71 KDa, le promoteur dans ce cas est situé avant l'exon 63. La maturation des transcrits primaires est aussi associée à de l'épissage alternatif (non illustrée).

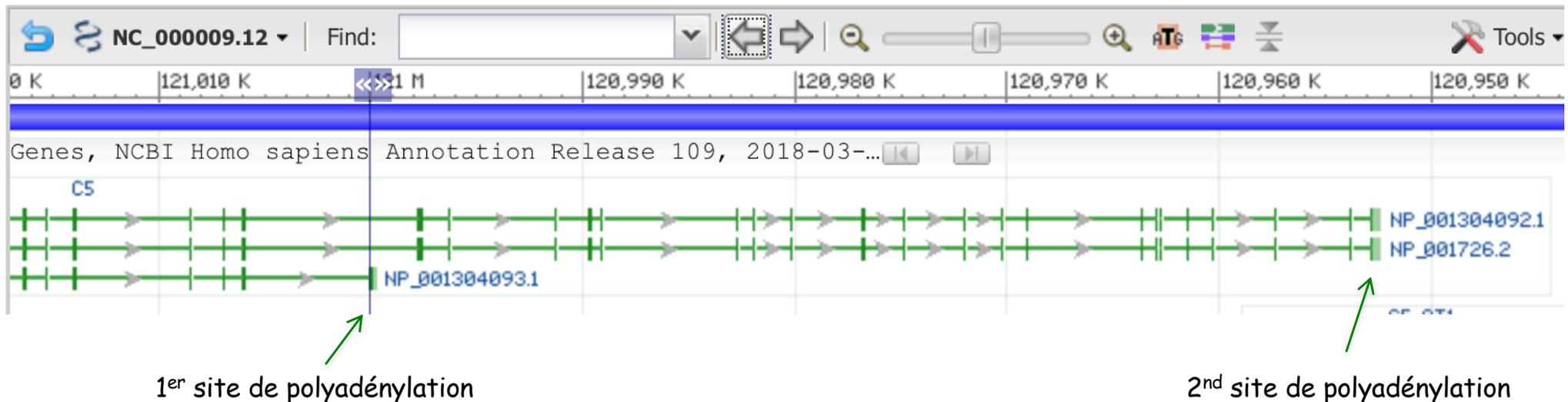
b) Epissage alternatif et polyadénylation alternative

RAPPEL

-> Epissage alternatif : selon les tissus / types cellulaires, l'épissage du transcrit primaire aboutit à des ARNm différents, par ex. un exon particulier sera inclus ou pas dans l'ARNm, c' est un exon alternatif -> voir ex. dia suivante

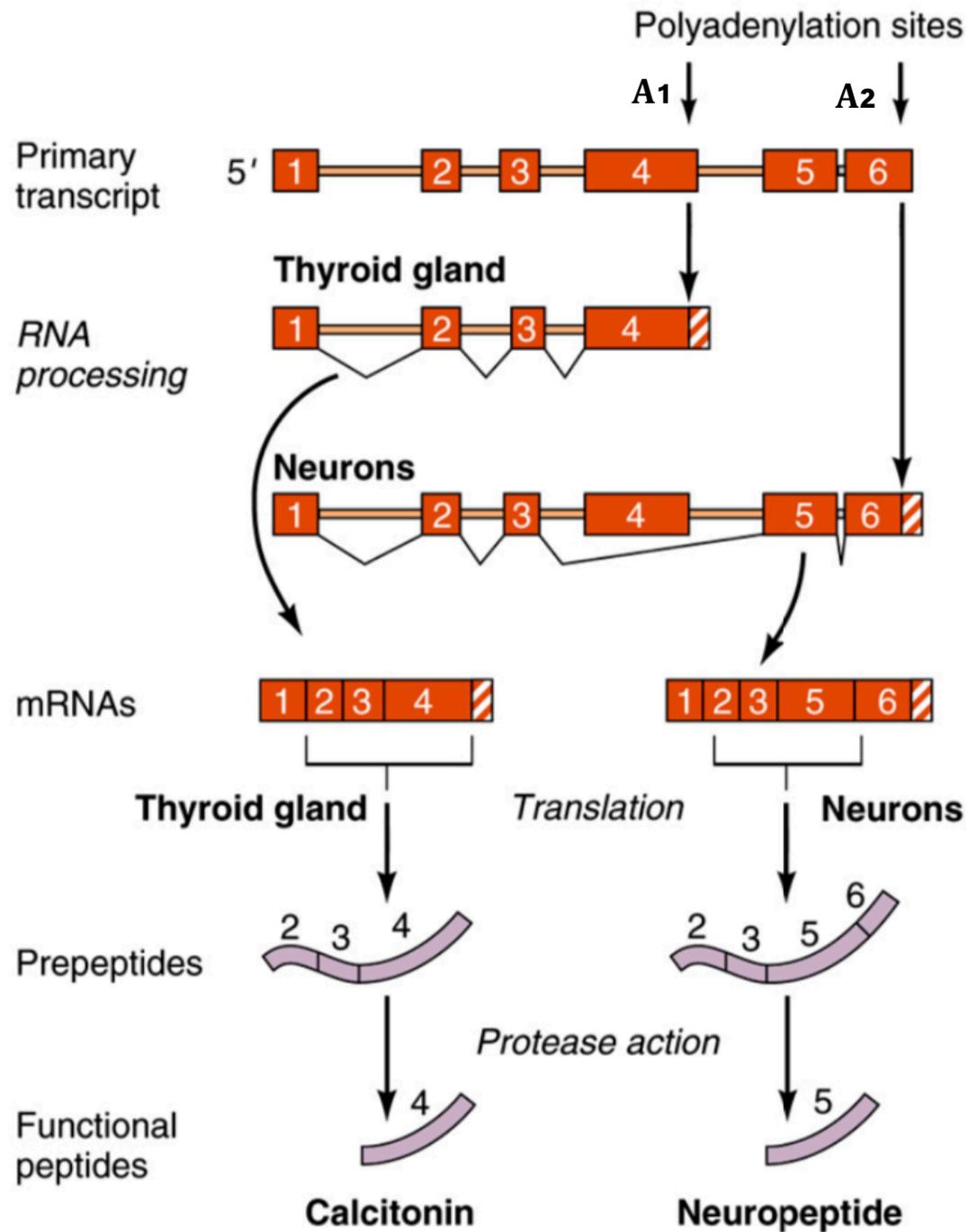
-> Polyadénylation alternative : selon les tissus / types cellulaires, la coupure en 3' qui précède la polyadénylation se produit à différents endroits possibles du transcrit primaire

Ex. gène C5 ⇒ 2 sites de polyadénylation alternative pour ce gène :



Gène humain de la calcitonine :

Le transcrit primaire du gène de la calcitonine subit une maturation qui varie selon le tissu où il s'exprime. Ainsi, pA1 et pA2 sont des signaux de polyadénylation alternative employés respectivement dans le tissu neural et le tissu thyroïdien. Un épissage alternatif peut conduire ou non à l'élimination de l'exon 4. L'exon 1 couvre le 5' UTR, l'ATG se trouve dans l'exon 2, l'exon 6 contient le 3' UTR. Les protéines sont synthétisées sous forme de précurseurs, ceux-ci sont ensuite clivés pour libérer soit la calcitonine (thyroïde), produit de l'exon 4, soit un neuropeptide (tissu nerveux), produit de l'exon 5.



RAPPEL