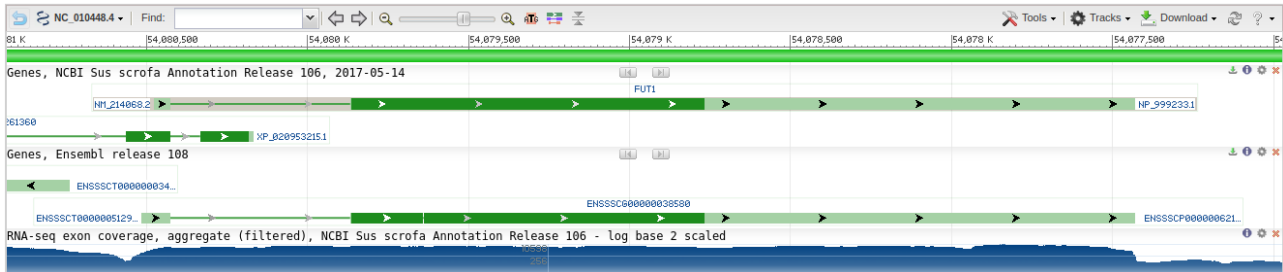


Fut1

Annotation actuelle

[Annotations ensembl et ncbi sur ncbi](#)



- Remarque : présence d'un intron de 2 nucléotides sur l'annotation ensembl dans la cds, absent chez ncbi
- NCBI : 2 exons, 1 intron (CDS exon 2)
- [Ensembl](#) : 3 exons, 2 introns (CDS exons 2 et 3)

Alignement des gènes (cf genes.clustal_num) & transcrits (mrna.clustal_num) ncbi et ensembl

ncbi_gene	CAACGCCTCCGATTCTGTCCCAAGCA	CCTGCCTCCTTT	CCGGGACCTGGACTATTTAC	861
ensembl_gene	CAACGCCTCCGATTCTGTCCCAAGCA	CCTGCCTCCTTT	CCGGGACCTGGACTATTTAC	900

NM_214068.2_FUT1_ncbi	TCCCAAGCA	CCTGCCTCCTTT	CCGGGACCTGGACTATTTACCCGGATGGCCGTTTGG	321
ENSSSCT0000000051297.2_cdna:protein_coding	TCCCAAGCA	CCTG--CCCTTT	CCGGGACCTGGACTATTTACCCGGATGGCCGTTTGG	357

Site de l'intron :

gène	CCTGCCTCCTTT
ARNm ensembl	CCTGCC..CTTT
ARNm ncbi	CCTGCCTCCTTT

=> Erreur dans la séquence du transcrit NCBI ?

Structure de la protéine

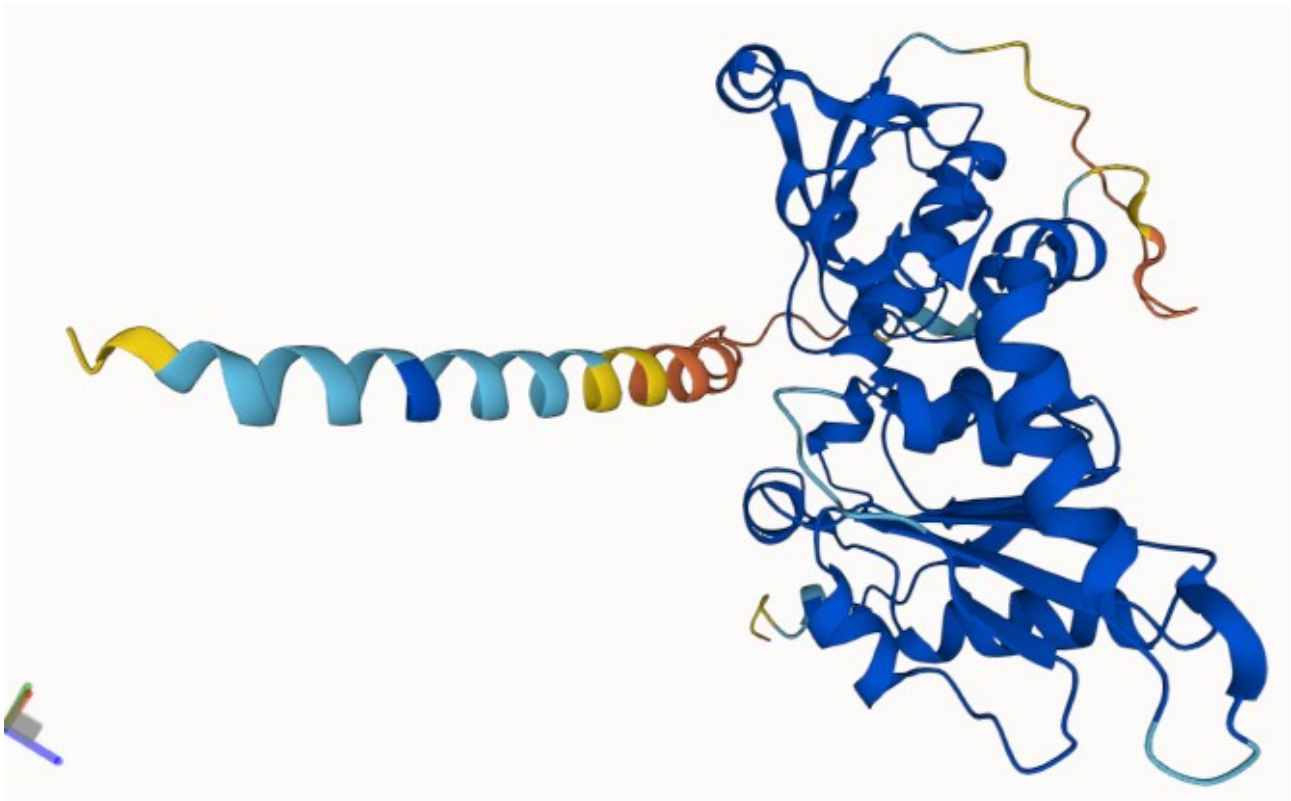
Protéine [humaine](#) (ID P19526) et [porcine](#) (ID Q29043)

2 domaines principaux :

- Peptide signal (AA 1-29) avec hélice transmembranaire (domaine 1)
- Domaine alpha-1,2-fucosyltransférase (AA 87-341) (domaine 2)

Remarque :

- Peptide signal dans exon 2 chez ensembl, domaine FUT dans exon 3
- Décalage de cadre de lecture entre exons 2 et 3 sur FUT1 ensembl



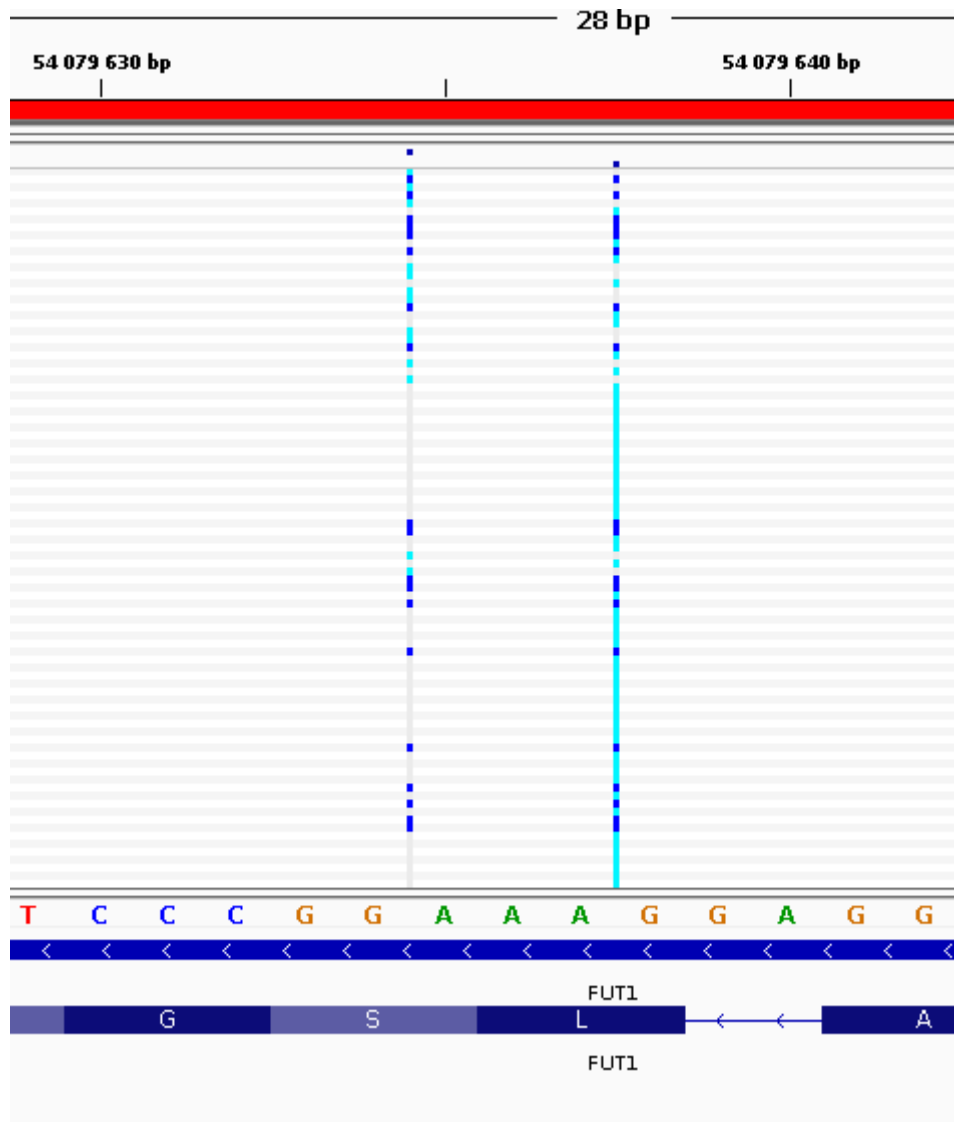
Alignement de FUT1 humain et des domaines

Alignement de séquences : FUT1 humain, domaine 1 (cadre de lecture 2) et domaine 2 (cadre de lecture 1)

	Peptide signal		
fut1_frame2	MWVPSRRHLCLTFLLVCLAAIFFLNVYQDLFYSGLDLLALCPDHNVVSSPVAIFCLAGT	60	
fut1_human	MWLSHRQLCLAFLLVCVLSVIFFLHIHQDSFPHGLGLSILCPDRRLVTPPVAIFCLPGT	60	
fut1_frame1	-----RGIIS--RGHILPGG	13	
		:	:::
	Exon 2 Exon 3		
fut1_frame2	PV-HPNASDSCPKHPASFPGPLFTRMAGLGTWDSMPRCWP-----	101	
fut1_human	AM-GPNASSSCPQHPASLSGTWTVYPNGRFGNQMGQYATLLALAQLNGRRAFILPAMHAA	119	
fut1_frame1	HAGTPQRLRLFSQASCLLSGTWTIYPDGRFGNQMGQYATLLALTQLNGRQAFIQPAMHAV	73	
	*:	:	: * *
		:	:::
fut1_frame2	-----	101	
fut1_human	LAPVFRITLPVLAPVEVDSRTPWRELQLHDWMSEFYADLRDPFLKLSGFPCSWTFFHHLRE	179	
fut1_frame1	LAPVFRITLPVLAPVEDRHAPWRELELHDWMSDYAHLKEPWLKLTGFPCSWTFFHHLRE	133	
		:	:::
	Domaine FUT		
fut1_frame2	-----	101	
fut1_human	QIRREFTLHDHLREEAQSVLGQLRLGRTGDRPRTFVGHVRRGDYLQVMPQRWKGVVGDG	239	
fut1_frame1	QIRSEFTLHDHLRQEAQVLSQFRLPRTGDRPSTFVGHVRRGDYLRVMPKRWKGVVGDG	193	
		:	:::
fut1_frame2	-----	101	
fut1_human	AYLRQAMDWFRARHEAPVFVVTNSGMEWCKENIDTSQGDVTFAGDQGEATPWKDFALLTQ	299	
fut1_frame1	AYLQQAMDWFRARYEAPVFVVTNSGMEWCKRKNIDTSRGDVI FAGDGREAAPARDFALLVQ	253	
		:	:::
fut1_frame2	-----	101	
fut1_human	CNHTIMTIGTFGFWAAYLAGGDTVYLANFTLPDSEFLKIFKPEAAFLPEWVGINADLSPL	359	
fut1_frame1	CNHTIMTIGTFGFWAAYLAGGDTIYLANFTLPTSSFLKIFKPEAAFLPEWVGINADLSPL	313	
		:	:::
fut1_frame2	-----	101	
fut1_human	WTLAKP	365	
fut1_frame1	QMLAGP	319	

Hypothèse :

- Intron entre ces deux domaines ?
- En examinant le vcf, on remarque que **tous** les individus ont exactement 1 insertion dans ce gène près de l'intron 2 (annotation ensembl) => redécalage du cadre de lecture



Haplotypes trouvés

Réf	CCT	GCC	TC	CTT	TCC	GGG
	P	A		L	S	G
Hap 1	CCT	GCC	TCC	C TT	TCC	GGG
	P	A	S	L	S	G
Hap 2	CCT	GCC	TCC	TTT	T CC	GGG
	P	A	S	F	S	G

- Remarque : Le premier haplotype correspond au transcrit FUT1 trouvé sur NCBI