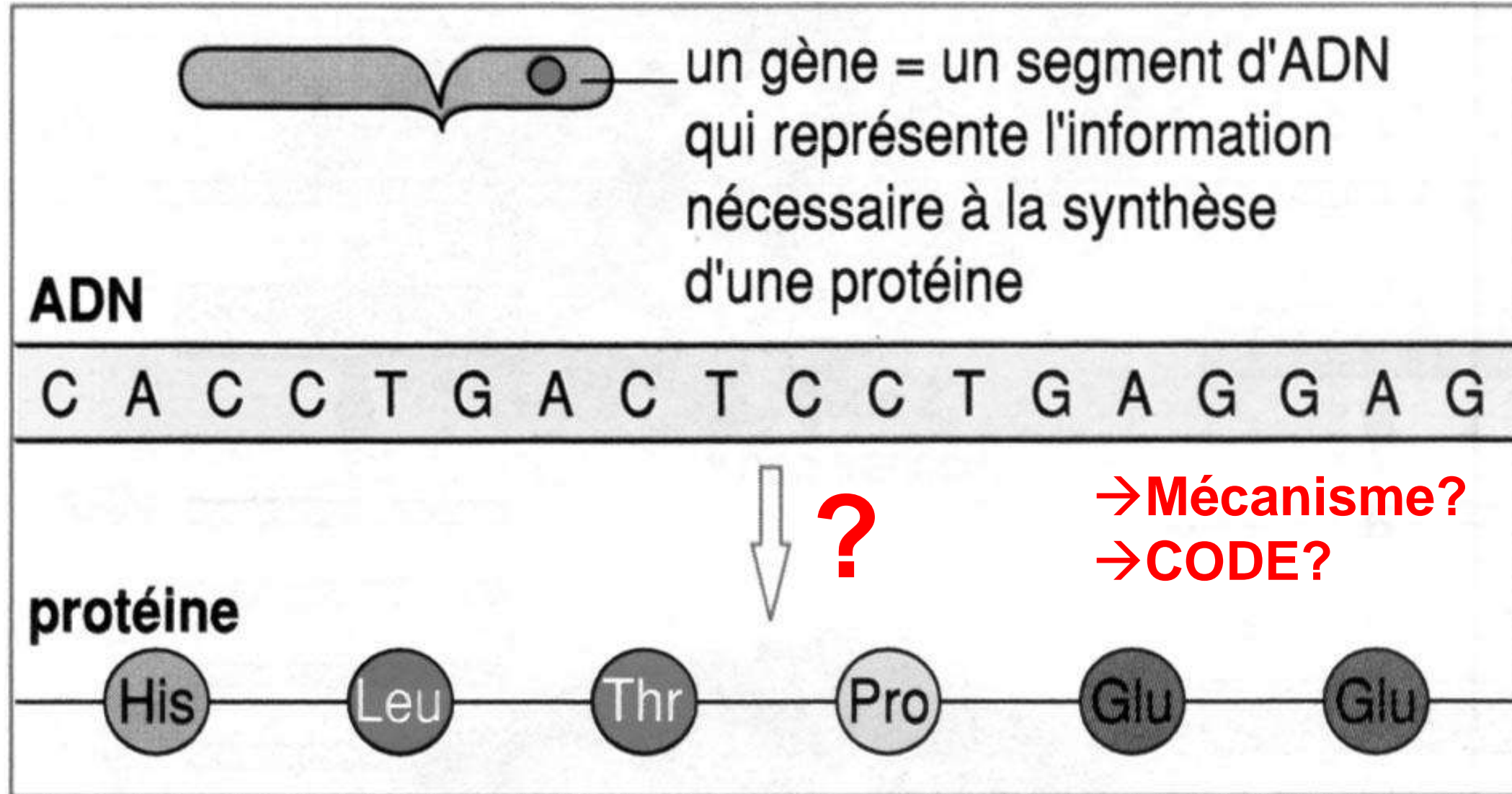


Correction du travail préparatoire

1958: *F. Crick* « *On protein synthesis* »: L'information permettant la synthèse des protéines est portée par l'ADN



Quelle est la correspondance entre la séquence des nucléotides de l'ARNm et la séquence des acides aminés de la protéine ?

<https://padlet.com/profsvt/codegenetiquespe1>

Document 1 :

Modification du nombre de nucléotides dans l'ADN viral	Séquence en acides aminés de la protéine de référence	Infection bactérienne
0	Normale	Oui
+1 ou -1	Mutée (nombreux acides aminés modifiés)	Non
+2 ou -2	Mutée (nombreux acides aminés modifiés)	Non
+3 ou -3	Mutée (1 acide aminé supplémentaire ou manquant , le reste de la séquence est identique)	Oui
+4 ou -4	Mutée (nombreux acides aminés modifiés)	Non
+5 ou -5	Mutée (nombreux acides aminés modifiés)	Non
+6 ou -6	Mutée (2 acides aminés supplémentaires ou manquant , le reste de la séquence est identique)	Oui

Seuls, l'addition ou le retrait de 3 nucléotides de l'ADN entraînent l'apparition ou la disparition d'un acide aminé dans la séquence de la protéine.

→ 3 nucléotides (= un codon) code pour un acide aminé

Document 2 : Résultats des expériences de Nirenberg & al. :

ARNm de synthèse	Séquence protéique obtenue
...UUUUUUUUUUUU... ..UUUUUUUUUUUU..	...Phe-Phe-Phe-Phe... ..Phe-Phe-Phe-Phe..
...AAAAAAAAAAAAA... ..AAAAAAAAAAAAA..	...Lys-Lys-Lys-Lys... ..Lys-Lys-Lys-Lys..
...UCUUUCUUUCUU... ..UCUUUCUUUCUU..	...Ser-Phe-Phe-Leu... ..Ser-Phe-Phe-Leu..

Les codons UUU, UUC codent pour la phénylalanine (Phe)

→ Plusieurs codons peuvent coder pour le même acide aminé

Le codon AAA code pour la lysine (Lys)

Le codon UCU code pour la sérine (Ser)

Le codon CUU code pour la leucine (Leu)

Capture d'écran de portions des séquences d'ARNm codant pour les globines (protéines) et de leurs séquences en acides aminés avec Anagène

Début des séquences :

Toujours le même codon au début de la séquence de l'ARNm → **AUG** = **codon d'initiation** codant pour l'acide aminé **Met** (**Méthionine**)

	1	10	20	30	40
alpha.pro	Met	Val	Leu	Ser	Pro
alphacod.arn	AUG	GUG	CUG	UCC	UGC
beta.pro	Met	Val	His	Leu	Thr
betacod.arn	AUG	GUG	CAC	CUG	ACU
gamma.pro	Met	Gly	His	Phe	Thr
gammacod.arn	AUG	GUC	AUU	CAC	AGG
delta.pro	Met	Val	His	Leu	Thr
deltacod.arn	AUG	GUC	AUC	UGA	CUC

Fin des séquences :

euThrSerLysTyrArg
UGACCUCCAAAUACCGUUA
lyValAlaAsnAlaLeuAlaHisLysTyrHis
GUGUGGCUAAUGCCUGGCCACAGUAUCACUAA
lyValAlaSerAlaLeuSerSerArgTyrHis
GAGUGGCCAGUGCCUGUCCUCCAGAUACCACUGA
lyValAlaAsnAlaLeuAlaHisLysTyrHis
GUGUGGCUAAUGCCUUGGCUCACAGUACCAUUGA

Le dernier codon de l'ARNm ne code pour aucun acide aminé → **UAA, UGA, UAG** sont appelés des **codons stop**

Le code génétique

Deuxième nucléotide

Premier nucléotide

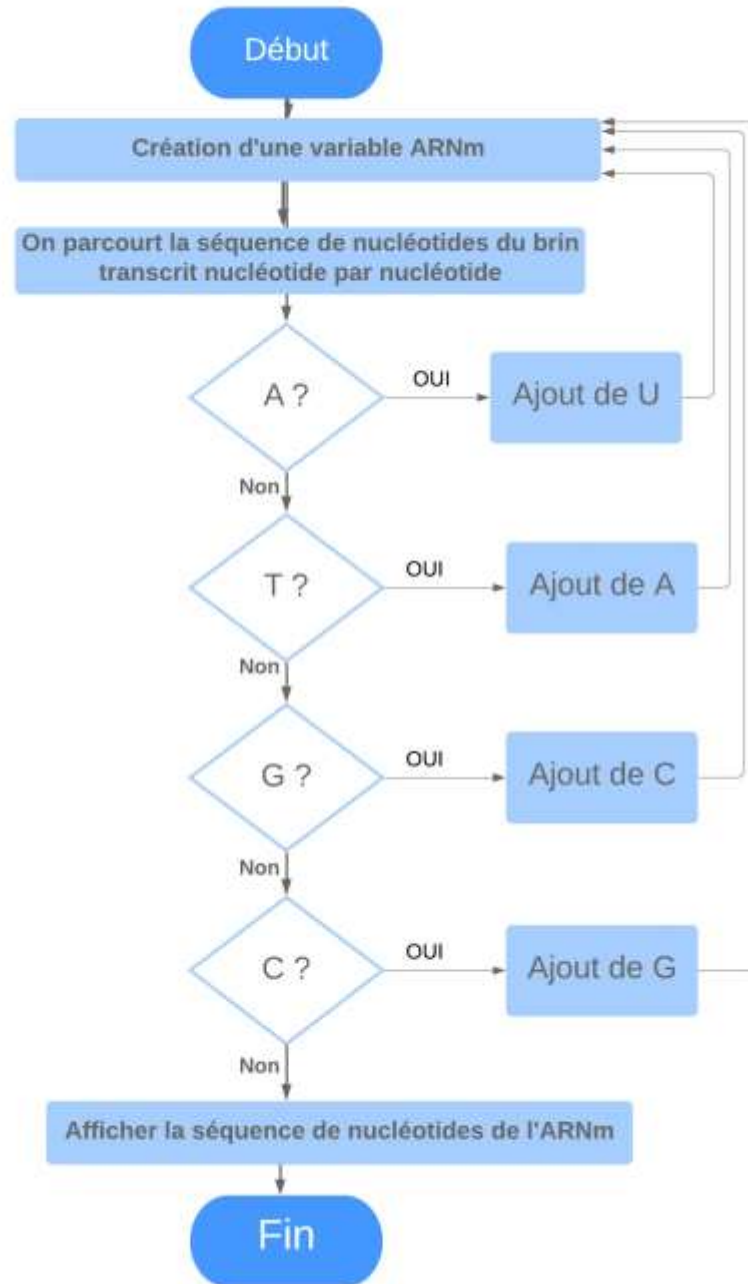
		U		C		A		G			
U	UUU	phényl- alanine	UCU	sérine	UAU	tyrosine	UGU	cystéine	U	C	A
	UUC		UCC		UAC		UGC				
	UUA	leucine	UCA		UAA	STOP	UGA	STOP			
	UUG		UCG		UAG		UGG				
C	CUU	leucine	CCU	proline	CAU	histidine	CGU	arginine	U	C	A
	CUC		CCC		CAC		CGC				
	CUA		CCA		CAA	glutamine	CGA				
	CUG		CCG		CAG		CGG				
A	AUU	isoleucine	ACU	thréonine	AAU	asparagine	AGU	sérine	U	C	A
	AUC		ACC		AAC		AGC				
	AUA		ACA		AAA	lysine	AGA	arginine			
	AUG	méthionine	ACG		AAG		AGG				
G	GUU	valine	GCU	alanine	GAU	acide	GGU	glycine	U	C	A
	GUC		GCC		GAC	aspartique	GGC				
	GUA		GCA		GAA	acide glutamique	GGA				
	GUG		GCG		GAG		GGG				

Troisième nucléotide

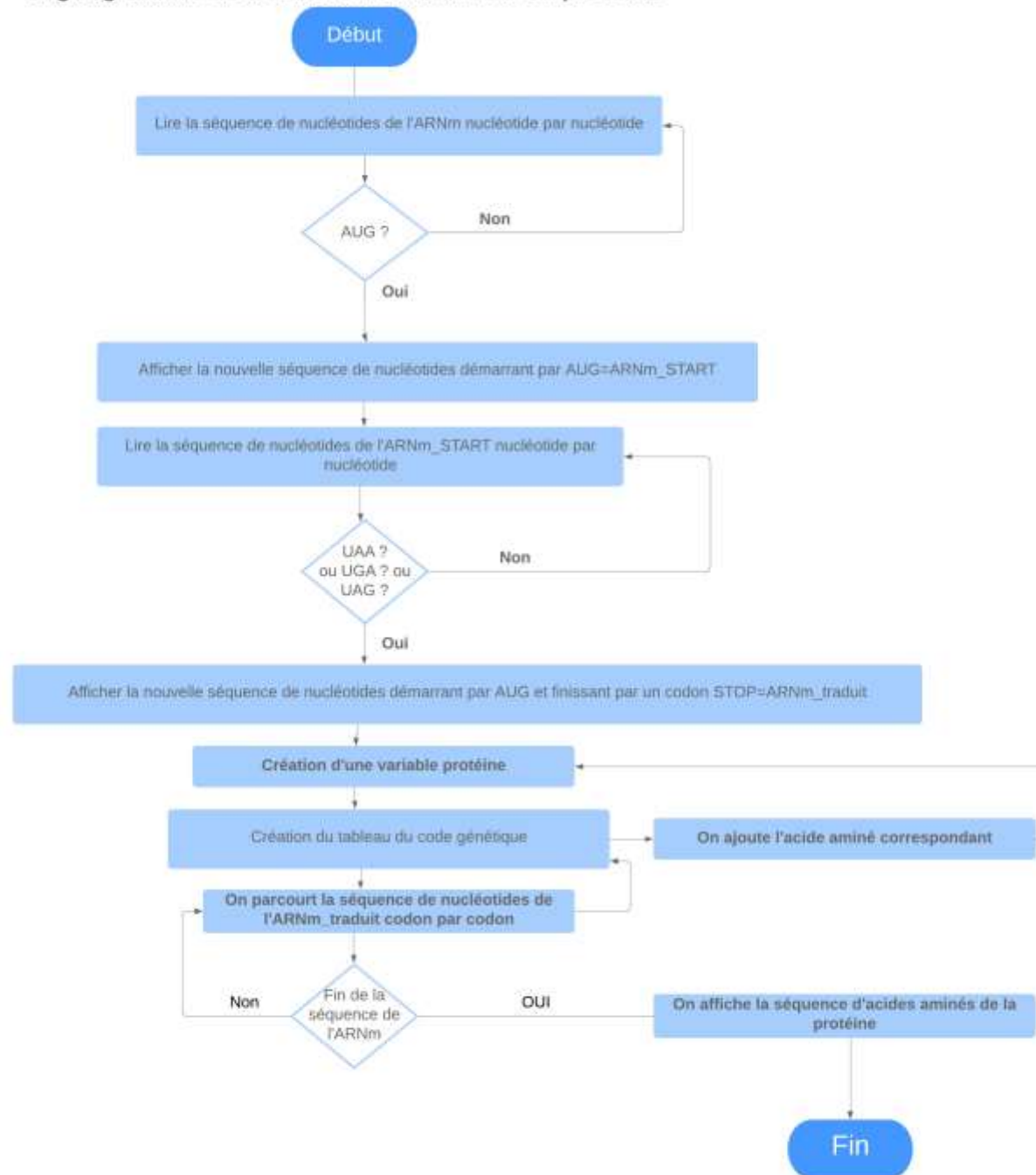
La production de la protéine (=traduction) démarre donc lorsque le codon d'initiation AUG est repéré et s'arrête lorsque qu'un des codons STOP (UAA, UGA,UAG) est repéré.

<http://svt.pages.ac-besancon.fr/codegenetique/>

Algorithme de transcription d'un brin transcrit d'ADN



Algorithme de traduction d'un ARNm en protéine



Consigne

Voici la portion codante de la séquence du gène TAS2R38 assurant la production de la protéine TAS2R38.

Cette protéine présente dans la membrane des cellules de la langue est un récepteur responsable de la sensibilité au goût amer des aliments.

En vous servant de votre travail maison sur le code génétique et de vos connaissances, **transcrire (=passer de l'ADN à l'ARNm) et traduire (= passer de l'ARNm à la protéine)** la portion de la séquence du gène TAS2R38 proposé ci-dessus afin d'obtenir la séquence d'acides aminés de la protéine TAS2R38.

	1	5	10	15	20	25	30	1003
Brin non transcrit	ATGTTGACTCTAACTCGCATCCGCACTGTG.....CTGTGCTGA							
Brin transcrit	TACAAGTGAAGATTGAGCGTAGGCGTGACAC.....GACACGACT							
ARNm	AUGUUGACUCUAACUCGCAUCCGCACUGUG.....CUCUGCUGA							
Protéine	MetLeuThrLeuThrArgIleArgThrVal..... LeuCys							

Mais vous n'aviez qu'une portion du gène... Il contient en réalité plus d'un millier de nucléotides ! Vous pouvez bien sûr continuer à le transcrire et à le traduire à la main mais cela prendrait beaucoup de temps... Une solution, automatiser la transcription et la traduction grâce à un programme informatique qui vous permettra de transcrire et de traduire cette séquence en un temps record mais aussi toutes autres séquences que vous pourriez disposer !

D'ailleurs des étudiants en bio-informatique l'ont bien compris... Afin de faciliter leur travail, ils ont développé un programme informatique de transcription et de traduction en langage Python. Ce programme doit leur permettre, à partir de n'importe quelle séquence de nucléotides d'un gène d'obtenir la séquence d'acides aminés de la protéine correspondante. Mais, les étudiants ne sont pas d'accord entre eux...

- a. Par groupe de 4, **mettre en commun** vos algorithmes de transcription et de traduction réalisés à la maison.
- b. À partir des ressources à votre disposition (fiches techniques + votre réponse à la question 1), **tester** les programmes de chaque étudiant afin de **déterminer** celui ou ceux qui sont corrects ainsi que celui ou ceux qui ne le sont pas. Vous justifierez chacun de vos choix en argumentant.

Les programmes Python de chaque étudiant (transcription_traduction1, transcription_traduction2...) sont à ouvrir avec l'éditeur d'Edupython présent sur les postes.

👉 Pour comprendre chaque étape des programmes Python, copier/coller des parties de programme sur le site Python tutor présent au lien suivant : <http://pythontutor.com/visualize.html#mode=edit>

brin_transcrit_TAS2R38 =

CTCTTCACTGTAGTACAACCTGAGATTGAGCGTAGGCGTGACACAGGATACTTCAGTCCTCATGTAAAGACAAGTAAAGTCAGGACCTCAA
ACGTCACCCCAAAGACTGGTTACGGAAGCAAAGAACCCTTAAAAACCCTACATCACTTCTCCGTCGGTGACTCGTTGTCACTAACACAC
GACGACACAGAGTCGTAGTCGGCCGAAAAGGACGTACCTGACGACAAGGACTCACGATAGGTCTGAATGGGTGAAGGTCTTCAACTCACT
TGGTGACTTGGTGTTCGATGGTTCGGTAGTAGTACGATACCTACTAACGTTTGGTTCGGTTGGAGACCGAACGACGGACGGAGTCGGACG
AAATGACGAGGTTTCGAGTAGGCAAAGAGAGTGTGGAAGGACTAGACGAACCGTTCGACCCAGAGGTCCTTCTAGAGGGTCTACGAGGA
CCCATAATAAGAAACGAGGACGTAGACGTGACAGGAGACACAAACCACGAAAAAATCGTCTGGAGTGAAGTGTTCAGTGTTGACACGATA
AGTACTTATTGTTATGTTCCGAGTTGACCGTCTAATTTCTAGAGTTAAATAAAATAAGGAAAGAGAAGACGATAGACACCAGACACGGAGG
AAAGGATAACAAAGACCAAAGAAGACCCTACGACTGACAGAGGGACCCTTCCGTGTACTIONCTGTTACTTCCAGATATGGTCTTTGAGAGC
ACTGGGGTCGGACCTCCGGGTGTAATTTCGGGAGTTCAGAGAACAGAGGAAAAAGACGAAGAAACACTATAGTAGGACACGACGGAA
GTAGAGACACGGGGATGACTAAGACACCGCGCTGTTTTATCCCCACTACCAAACACAACCCTATTACCGTCGAACAGGGAGACCCGTACG
TCGGCAGGACTAGAGTCCGTTACGGTTCAACTCCTCTCGACACTACTGGTAAGACGAGACCCGAGTCTCGTCGGACTTCCATTCTCGGCT
GGTGTTCGCTAAGGGCCTGTGACACGACTTATTAT

Nombre de nucléotides : 1021

Correction

Programme 1 : **ne commence pas par AUG** car la fonction codon_initiation a une erreur au niveau du codon d'initiation qui est AUC donc la protéine commence par Ile au lieu de Met

Programme 2 : c'est le juste

Programme 3 : **codon stop faux UGG au lieu de UGA**, la protéine est donc beaucoup plus courte

Programme 4 : **erreur dans le tableau du code génétique**

"UUU":"Phe", "UUC":"Phe", "UUA":"Phe", "UUG":"Phe",

Au lieu de :

"UUU":"Phe", "UUC":"Phe",

"CUU":"Leu", "CUC":"Leu", "CUA":"Leu", "CUG":"Leu", "UUA":"Leu", "UUG":"Leu",

→ D'où une séquence d'acides aminés fausse

Programme 5 : erreur dans le programme de transcription / **ne permet pas la transcription des G en C** d'où la création d'un codon stop UGA et un ARNm de longueur 66 (nucléotides) et une protéine de 22 acides aminés

Activités complémentaires pour les plus rapides et/ou les plus experts :

- **Déterminer** la longueur des différentes séquences en utilisant la console python et l'instruction `len`
- Le gène TAS2R38 existe sous 2 formes alléliques. Les 2 allèles diffèrent par 3 mutations par substitution (cf tableau ci-dessous). Les individus possédant 2 exemplaires de l'allèle muté sont insensibles au goût amer.

Créer votre propre programme sur Edupython en téléchargeant le fichier Mutations et l'ouvrir avec Edupython. Ce programme doit **permettre de modifier** la séquence d'ADN du gène (=mutations) ainsi que de **déterminer** les conséquences de ces 3 mutations sur la protéine TAS2R38.

 *Utiliser le fichier Mutations_secours si vous n'y arrivez vraiment pas.*

Compléter ensuite le tableau avec vos résultats.

Tableau de résultats :

Position du nucléotide dans la séquence du brin transcrit	Changement de nucléotide	Changement de codon	Changement d'acide aminé
145	G -> C	CCA->GCA	Pro->Ala
785	G -> A	GCU->GUU	Ala->Val
886	C -> T	GUC->AUC	Val->Ile