

Biologie

Fonctionnement du cours : explication d'un principe du cours précédent pendant le cours suivant (~ 10 min). Modalité d'évaluation ?

PCR

Principe de répllication d'une grande quantité d'ADN. Augmenter la température pour dénaturer (en fait séparer) les deux brins d'ADN. On peut fabriquer des amorces (petits bouts qui vont servir d'amorce à la répllication $5' \rightarrow 3'$). Quand on augmente la température, on sépare les brins, et quand on refroidit, les amorces vont se fixer, ce qui va permettre de répliquer sur deux fois plus d'ADN qu'au départ. On a perte de certaines régions dues au fait que les amorces sont l'extrémité $5'$. Beaucoup d'amorces pour éviter la réhybridation entre brins.

1 ARN

ADN into ARN into proteines. Les ARN sont transcrits au niveau du noyau de la cellule. Export vers le cytoplasme.

T into U. Simple brin. Sucre : ribose vs désoxyribose (un H à la place de OH).

ARNmtr. Micro-ARN.

ARNm : code pour les protéines.

ARNr : font partie du ribosome.

ARNt : apportent les acides aminés pendant la traduction.

ARN de transcription/maturation, voire de régulation.

ARN polymérase : pas besoin d'amorce contrairement aux ADN polymérase.

TF : facteurs de transcription.

Polymérisation de l'ARN. NTP : nucléotide triphosphate. C'est rapide (plus que pour l'ADN).

Qu'est-ce qui détermine le niveau de transcrption d'un gène ?

- les promoteurs : essentiel au démarrage de la transcription : au début du niveau de la transcription. TATA box. Les promoteurs permettent les sites de liaison des facteurs de transcription. Le choix du promoteur va changer le résultat de la transcription. Les co-activateurs sont également importants
- enhancer : module l'efficacité de la transcription, parfois à plusieurs kilobases du début du gène
- chromatine

Etude des facteurs de transcription.

Retard sur gel.

PCR. Amplification d'un produit. Analyse de pcr. pour cela on utilise un anticorps pour spot les protéines (notamment facteurs de transcription) qui interviennent lors du mécanisme. On immunoprécipite la protéine + l'ADN à laquelle elle est attachée, et on détermine ainsi où était la protéine sur l'ADN. phospholysation, acétylation, méthylation, etc.

chromatine : ADN et les protéines qui lui sont liées.

L'ADN est enroulée autour d'une protéine appelée histone. Le nucléosome est l'ensemble histone + ADN. et il ya plein d'histones comme ça. Les histones forment des structures, etc. Les histones ont des queues qui peuvent être modifiées. Il y a plusieurs types de chromatines : l'euchromatine (ouverte, la transcription est active) et l'hétérochromatine (transcription inactive).

il y a beaucoup de modifications possibles sur les queues des histones. la chromatine active est globalement acétylée. L'inactive est globalement méthylée (et inaccessible). Marques des histones. boucles de rétroaction. conséquence : en regardant les histones, on peut prédire si un gène va être transcrit. dinucléotide cpg. Il y a des îlots cpg.

→ en gros : marques épigénétiques. C'est une sorte de code des histones. épigénétique : transmissible à la génération suivante, mais pas tout le temps -.-

Il y a des protéines pour lire, écrire ou effacer des marques épigénétiques. enhancers → tosee. Reconnus par un facteur de transcription. il va stabiliser la polymérase par exemple. Sur l'enhancer, se fixent d'autres protéines qui vont aider.

Hybridation à l'ARNm. Boucles. L'ARNm est plus court que la portion d'ADN génomique dont il est issu, c'est l'épissage.

Coiffe.

Au fur et à mesure de la transcription, les bouts vont être coupés, etc. Tout n'est pas fait à priori.

C'est source d'une grande variation dans les transcrits. C'est l'épissage alternatif → ce qui donne plusieurs protéines différentes.

Queue poly-a. Exons mutuellement exclusifs

Tout est régulé. Modification des histones ou méthylation de l'ADN qui peuvent donner l'un ou l'autre.

1.1 transcriptomique

C'est l'étude quantitative et qualitative de l'ARN.

1.1.1 quantitative

Transcription inverse : partir de l'ARN et on produit l'ADN. Reverse-transcriptase.

On fait ça avec des rt (radioactifs) pour pouvoir spot les ADN.

Amorce polyt pour la queue poly-a.

Puces à ADN #landers #vieux. A chaque position on a fixé un oligo de séquence connue. Par conséquent, on a accès à la quantité d'ADNc sur les spots qui sont très représentés. On peut les spot en regardant la fluorescence sur la puce.

Actuellement quantification haut débit.

1.1.2 qualitative

Ligation et clonage (comme Landers, on fait plus trop). Maintenant, on fait haut débit, qui a pour désavantage de ne pas être trop sensible aux variations d'exons.

séquencage nanopore. On peut séquencer directement de l'ARN. permet de

séquencer un ARN entier.
Pré-messager : introns+exons.
Oligo-nucléotides : petits bouts d'ADN.

Facteurs TF II.

1.2 L'odyssée des protéines

Grande variété de longueurs. Alphabet encodant des acides aminés (code génétique).
Code dégénéré (dû à la redondance du code). AUG : initiation Met. UAA, UAG, UGA codons stop.

Open reading frame. ORF. Délétère. Cadre de lecture.

Initiation de la traduction. Pour traduire, il faut l'ARMm, mais aussi d'autres choses. Les ARN ont un rôle dans la traduction (ribosome). Coiffe. A la fin de l'ARNm, il y a une queue poly-A.

MET en position 1. Ensuite, les acides aminés (sous forme d'ARNt) reconnaissent les uns après les autres le bout d'ARN en cours. Puis, par hydrolyse, il y a liaison avec le début de la chaîne polypeptidique. Quand on arrive au poly-A, on récupère un eRF3 au lieu d'un peptide.

On peut faire les opérations en parallèle.

1.2.1 Régulations traductionnelles

Les Kinases. Ce sont des facteurs de régulation de la traduction. On régule la quantité de traduction en fonction des acides aminés par exemple. PAPB (poly-A binding protein). PAPB à la fin et 4G au début qui se fixent ensemble, de manière à circulariser l'ARN et favoriser la réutilisation des ribosomes.
Initiation indépendante de la coiffe.

Phosphorylation d'une protéine pour lui faire changer de conformation et ainsi l'inhiber.

Adressage des protéines. NLS Séquence de localisation cellulaire. Histoire du peptide signal. Il y a plein de moyens pour envoyer les protéines là où elles doivent aller. Protéines chaperonnes. Prions.

Modifications post-traductionnelles. On peut ajouter des tas de trucs sur les protéines.

Ubiquitine. Protéine absolument **vitale**. On peut attacher l'ubiquitine en construisant une chaîne basée sur l'ubiquitine. 1000 gènes pour l'ubiquitine (1/26ème du génome!)
Protéasome : broyeur à protéines.

Parenthèse : N-terminal : début de la protéine. C-terminal : fin de la protéine.

Anticorps. On peut séparer les protéines notamment en fonction de leur taille (expérimentalement, par électrophorèse, ou par tamis moléculaire, en

commençant par casser les ponts dissulfures \Rightarrow c'est de la chromatographie).

Immunodétection : détection par des anticorps qui viennent se fixer sur une protéine d'intérêt. Production d'anticorps on détourne une réponse immunitaire.

Spectrométrie de masse : exo de spé de physique : on a un champ magnétique qui fait tourner les particules et selon la masse on détecte à différents endroits. Masse/charge.

Purification par affinité des protéines (immunoaffinité ou affinité). La protéine a des affinités et le but est de récupérer les complexes qui sont souvent avec.

Production de protéines. Protéines recombinantes. Insuline : source de monnaie incroyable due au brevet. Le protéomique : l'ensemble des protéines d'un individu (plutôt d'une cellule).

Résumé du début du cours Remarque : le complexe ne peut se former que sous la condition qu'il y ait le premier MET. Micro ARN.

2 La biologie numérique

La bioinfo. 'in silico'. Bioinformatique \sim biologie des systèmes. Margaret Deyhoff #badass. Pauling et Zuckerkandl : retracer les arbres phylogénétiques.

2.1 Bioinformatique structurale

Déterminer la structure 3D des macromolécules biologiques à partir de leurs séquences. C'est un problème NP-complet et difficile. #foldit. Découverte de la double-hélice via de la cristallographie. Rosalind Franklin. Les 4 structures de la protéine. *Remarque : ce ne serait pas con de faire résoudre des problèmes intéressants tels que la prédiction de la conformation 3D d'une protéine pour miner du bitcoin au lieu de faire de la factorisation stupide.*

2.2 Analyse de réseaux

Étudier les interactions de tous les composants. Déterminer des états stables d'un réseau, tracer la transduction du signal ? Déterminer les gènes actifs, etc. C'est plutôt marrant en fait. Arbres de décision qui sont un peu complexes. Cancerologie et régulation.

Analyse de données. Machine learning, deep learning. Random forest. Votes d'arbres de décisions.

DL.

2.3 Analyse de séquence

Comment faire pour séquencer facilement. Méthode de Sanger.

Seconde génération sequencing (SGS). On va amplifier et récupérer des bouts, que l'on identifie par leur couleur. Genomic RNA ? (ça n'existe pas on dirait).

Third generation sequencing (Next Generation Sequencing, NGS). MinIon et PacBio. Long read (grands fragments). Problèmes de la probabilité d'erreurs (qui est bien plus importante). Islande DCode.

Elizabeth Warren.

Miicrobiome.

2.4 Comparaison des séquences pour l'évolution

Evolution des séquences. Parallèle entre les séquences et l'évolutoin des espèces.

Reconstitution de l'histoire des migrations. Ratio dN/dS. Synonymes et non synonymes (qui mènent au même acide aminé ou pas).

Aligement des séquences. Dinucléotides cPg (hypermutable quand il est méthylé ?). On peut faire des probabilités conditionnelles, qui sont plus performantes.