



CAPSULE 4 COMMENT LES INSTRUCTIONS GÉNÉTIQUES CODENT LES PROTÉINES DU CORONAVIRUS ?

PAR JULIAN VENABLES ET PHILIPPE FORT

Les instructions génétiques pour fabriquer et maintenir nos cellules demeurent sur deux types de longues molécules similaires, de l'ADN ou de l'ARN. Le coronavirus injecte de l'ARN qui code les instructions pour faire plus de virus. Comment sont-elles donc codées ?

Considérons un code simple, le code Morse. Ce code, inventé il y a presque deux siècles, était utilisé pendant la seconde guerre mondiale pour envoyer des messages par radio. Le message classique de détresse est le 'SOS' pour 'Save Our Souls' en anglais, ou 'sauver nos âmes' en français. L'alphabet Morse ne contient que deux lettres, les points et les tirets. A la radio, un point est un son court (ti), un tiret est un son long (tah). Le Morse peut coder l'ensemble des lettres de l'alphabet, qui contient 26 lettres chez nous, par une combinaison de points « . » et des tirets « - ». SOS est ainsi codé par trois « . » , trois « - » et trois « . ». On reconnaît très facilement SOS à la radio, ça fait ti-ti-ti, tah-tah-tah, ti-ti-ti !

Bien sûr, la cellule utilise un autre code que le Morse pour fabriquer les protéines, y compris la protéine Spike. Cette protéine est un assemblage très long, de 1273 'briques'. Les briques sont des molécules appelées acides aminés, et il y a 20 briques différentes dans l'alphabet des protéines. Pour fabriquer une protéine, chaque brique est ajoutée à la queue-leu-leu comme les maillons d'une chaîne par une machine de la cellule, appelée le Ribosome, capable de décoder l'information génétique.

Le support de l'information génétique est le même dans tout le monde vivant et il est composé de quatre lettres d'ARN : A, C, G et U. Pour décoder l'information, le Ribosome lit les lettres génétiques trois par trois, et comprend ainsi quelle brique il doit ajouter à la protéine en construction. Par exemple, les 5 premières briques de la protéine Spike sont MFVFL et sont codées par les 15 lettres génétiques, AUG UUU GUU UUU CUU. Le premier triplet, AUG, code le premier acide aminé, M. Le deuxième triplet, UUU, code le second acide aminé, F, et ainsi de suite. Notez que la deuxième et la quatrième brique de Spike sont toutes les deux 'F', et sont bien codées par le même triplet 'UUU'.

Nos cellules peuvent contenir jusqu'à 30 000 protéines différentes qui en constituent la matière. La protéine Spike est une des dix protéines codées par le Coronavirus. A seulement dix, ces protéines sont capables d'utiliser tout ce qu'il y a dans la cellule pour fabriquer autant de protéines et d'acides nucléiques viraux nécessaires à la production de milliers de virus, qui finiront par sortir de la cellule pour aller en infecter d'autres. On étudiera plus en détail le génome viral dans la prochaine capsule.