

Comment l'analyse du génome crée des virus fictifs



[Source : Recherches Covid-19 (cv19.fr)]

L'épidémie du nouveau coronavirus 2019 ou de la maladie à coronavirus (COVID-19) est revendiquée comme une menace pour le monde entier depuis début 2020. Les scientifiques travaillent jour et nuit pour comprendre l'origine du COVID-19. Peut-être avez-vous déjà entendu la nouvelle que le génome suspecté du SARS-CoV-2 a été publié^[1] ? Comment exactement les scientifiques ont-ils identifié le génome complet du SARS-CoV-2 ?

Cet article explique, de manière purement technique, comment ils ont fait cela. Sans aucun doute, vous vous rendrez compte que cette méthode ne convient ni pour déterminer l'origine de l'acide nucléique d'un échantillon ni pour déterminer un génome exact. Divers algorithmes sont utilisés, dont les calculs tentent de proposer une construction fictive basée sur des spécifications, qui à son tour sert de modèle pour tous les processus ultérieurs. Cependant, chaque séquençage individuel conduit à des résultats différents du séquençage précédent ; ces écarts sont souvent vendus comme des mutations^[2], et ce sont simplement des incohérences dans le calcul par les outils d'analyse du génome.

Du fait qu'une structure prétendument virale n'a jamais été isolée en culture pure et que la caractérisation biochimique a été réalisée directement à partir de celle-ci, tout séquençage est basé sur un mélange de matériel génétique dont l'origine ne peut être déterminée et sont toutes des constructions assistées par ordinateur à considérer comme de la pure spéculation.

[Voir aussi :

- ▶ 87 institutions sanitaires/scientifiques du monde entier n'ont pas réussi à citer un seul cas d'isolement et de purification du "SARS-CoV-2", par qui que ce soit, où que ce soit, jamais
- ▶ Comment créer votre propre "nouveau virus" généré par ordinateur
- ▶ Un article du Dr Stefan Lanka 2020 réfute l'idée fausse du virus
- ▶ Réfutation de la virologie par le Dr Stefan Lanka
- ▶ Stefan Lanka conduit les expériences témoins réfutant la virologie
- ▶ Le culte de la virologie – Documentaire
- ▶ Virus fantôme : à la recherche du SARS-CoV-2
- ▶ Virus Mania : Comment Big Pharma et les médias inventent les épidémies
- ▶ Pourquoi tout ce que vous avez appris sur les virus est faux

- Anthrax, Arsenic, Pasteur et ses moutons
- La théorie des germes : Une erreur fatale
- Pour en finir avec les virus et le Nouvel Ordre Mondial]

Comme un petit mais non moins important détail, qui vous fait parfaitement comprendre que la manipulation technique de l'alignement/assemblage (comme expliqué dans l'article) est purement fictive et manipulatrice :

70% des virologues ne précisent pas quelle technique d'assemblage ils utilisent !

Et le fait tacite est que 100% des virologues ne divulguent pas quel pourcentage du génome «viral» calculé a été modifié jusqu'à ce qu'il ait été publié en tant que génome fini.

Avec les données du génome, qui par le chinois Fan Wu et. al^[1] ont été publiés et mis à disposition en téléchargement, il ne s'agit en aucun cas de données brutes, mais plutôt de données préalablement étoffées ! 80% des séquences humaines ont disparu et surtout : Tout ce qui était publié ailleurs comme « viral » a été filtré ! Cela rend encore plus difficile la compréhension de la prétendue découverte d'un nouveau virus pathogène, puisque des expériences témoins sont également recherchées en vain dans les publications publiées.

Mais nous arrivons maintenant à l'analyse technique du génome.

Génome

Le génome est l'ensemble du matériel génétique, y compris tous les gènes d'un organisme. Le génome contient toutes les informations sur un organisme qui sont nécessaires à sa structure et à son maintien.

Les virus qui causent des maladies sont définis par le fait que leur séquence (génome) est unique et ne se produit pas dans des organismes sains.

Pour pouvoir détecter et déterminer la présence du matériel génétique d'un virus, ce virus doit être isolé et disponible à l'état pur conformément aux lois de la pensée et à la logique qui précède toute science comme règle fondamentale, afin que les propres séquences de gènes de la cellule ne soient pas interprétées à tort comme des composants d'un virus.

La séquence d'une substance génétique ne peut être déterminée que si elle se présente sous la forme d'un ADN.

Séquençage

Comment lire les informations présentes dans le génome ? C'est là qu'intervient le séquençage.

Le séquençage est utilisé pour déterminer la séquence de gènes individuels, de chromosomes entiers ou de génomes entiers d'un organisme.



Fig 1. A PacBio sequencing machine. PacBio is a third-generation sequencing technology which produces long reads. Image by KENNETH RODRIGUES from Pixabay (CC0)

Des dispositifs spéciaux appelés machines de séquençage sont utilisés pour extraire de courtes séquences aléatoires du génome qui nous intéressent. Les technologies de séquençage actuelles ne peuvent pas lire le génome entier à la fois. De petits morceaux d'une longueur moyenne comprise entre 50 et 300 bases (séquençage de nouvelle génération / lectures courtes) ou 10 000 à 20 000 bases (séquençage de troisième génération / lectures longues) sont lus, selon la technologie utilisée. Ces courtes pièces sont appelées lectures.

La présence et la longueur d'un matériel génétique sont déterminées en le divisant dans le sens de la longueur dans un champ électrique. Les pièces courtes voyagent plus rapidement, les pièces plus longues plus lentement. En même temps, afin de pouvoir déterminer la longueur du matériel génétique à examiner, des morceaux de matériel génétique de différentes longueurs avec une longueur connue sont ajoutés. Cette technique standard fiable pour détecter et déterminer la longueur du matériel génétique est connue sous le nom d'« électrophorèse sur gel ».

Si la concentration d'un certain matériel génétique est trop faible pour qu'il ne puisse pas être détecté avec la technique de "l'électrophorèse sur gel", cela peut être fait à volonté en utilisant la technique de réplique

illimitée de l'ADN, appelée amplification en chaîne par polymérase PCR. De cette manière, un ADN indétectable peut être rendu visible par électrophorèse sur gel. C'est une condition préalable pour rendre la substance génétique accessible à d'autres investigations, en particulier pour la détermination ultérieure et décisive de sa longueur et de sa séquence. Cette méthode est également connue sous le nom de PCR en abrégé.

Si vous recherchez des informations plus détaillées sur le fonctionnement exact du séquençage du génome viral à partir d'échantillons cliniques, consultez l'article suivant.

1. A complete protocol for whole-genome sequencing of virus from clinical samples: Application to coronavirus OC43
2. Specific Capture and Whole-Genome Sequencing of Viruses from Clinical Samples

Assemblage du génome

Une fois qu'il y a de petits morceaux du génome, nous devons les combiner en fonction de leurs informations de chevauchement et construire le génome complet. Ce processus est appelé assemblage et est comparable à la résolution d'un puzzle. Des outils logiciels spéciaux, appelés assembleurs, sont utilisés pour assembler ces lectures en fonction de leur chevauchement. Ceci est utilisé pour générer des chaînes continues, appelées contigs. Ces contigs peuvent être le génome entier lui-même ou juste une fraction de celui-ci (comme le montre la figure 2). Il est important de noter que la source du matériel utilisé est sans importance.

Les virologues chinois n'ont pas effectué d'expériences témoins pour écarter la possibilité que même avec un ARN humain/microbien provenant d'un lavage de poumon d'une personne saine, une personne atteinte d'une autre maladie pulmonaire, une personne qui a été testée négative au SRAS-CoV-2., ou à partir de cet ARN provenant d'échantillons de réserve de l'époque où le virus SRAS-CoV-2 était encore inconnu, exactement la même addition du génome d'un virus à partir de courts fragments d'ARN soit possible !

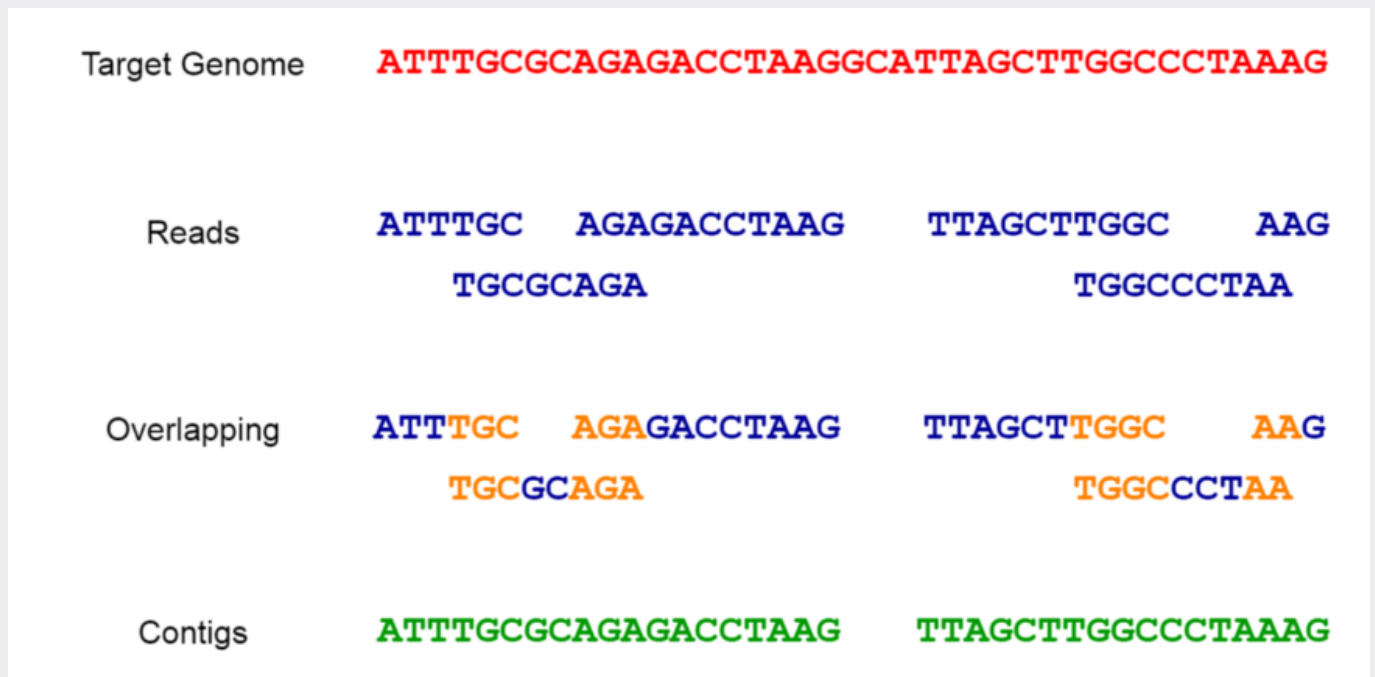


Fig. 2 Séquençage et assemblage

Les assembleurs sont divisés en deux catégories,

Assembleurs de novo : assemblent sans utiliser de génomes de référence (par exemple : SPAdes, SGA, MEGAHIT, Velvet, Canu et Flye).

Assembleurs guidés par des références : assemblent en assignant des séquences à des génomes de référence.

Deux grands types d'assembleurs

On trouve deux grands types d'assembleurs dans la littérature bioinformatique. Le premier type est la méthode des consensualités d'implantation par recouvrement (OLC). Dans la méthode OLC, tous les chevauchements entre les lectures sont d'abord déterminés. Ensuite, toutes les lectures et tous les chevauchements sont organisés sous la forme d'un graphique. Enfin, nous identifions la "séquence de consensus". SGA est un outil populaire basé sur la méthode OLC.

Le deuxième type d'assembleur est la méthode du graphe de-Bruijn (DBG)^[3]. Au lieu d'utiliser les lectures complètes telles quelles, la méthode DBG décompose les lectures en fragments encore plus courts appelés k-mers (de longueur k) et crée ensuite un graphe de-Bruijn en utilisant tous les k-mers. Enfin, les séquences génomiques sont déduites sur la base du graphe de-Bruijn. SPAdes est un assembleur populaire basé sur la méthode DBG.

Qu'est-ce qui peut mal tourner dans l'assemblage du

génomique ?

Les génomes contiennent des motifs d'acides nucléiques qui apparaissent de nombreuses fois dans le génome. Ces structures sont appelées répétitions. Ces répétitions peuvent compliquer le processus d'assemblage et entraîner des ambiguïtés.

Nous ne pouvons pas garantir que le séquenceur puisse produire des lectures couvrant l'ensemble du génome. Il se peut que le séquenceur omette certaines parties du génome et qu'aucune lecture ne couvre cette région. Cela affectera le processus d'assemblage et ces régions omises ne seront pas présentes dans l'assemblage final.

Les assembleurs de génomes doivent relever ces défis et essayer de minimiser les erreurs causées lors de l'assemblage.

Comment les assemblages sont-ils évalués ?

L'évaluation des assemblages est très importante car nous devons décider si l'assemblage résultant répond aux normes. L'un des outils d'évaluation des assemblages les plus connus et les plus utilisés est QAST. Vous trouverez ci-dessous quelques-uns des critères utilisés pour évaluer les assemblages.

N50 : longueur minimale des contigs nécessaire pour couvrir 50% de la longueur totale de l'assemblage.

L50 : nombre de contigs plus longs que N50

NG50 : longueur minimale des contigs requise pour couvrir 50 % de la longueur du génome de référence.

LG50 : nombre de contigs plus longs que NG50

NA50 : longueur minimale des blocs alignés nécessaire pour couvrir 50 % de la longueur totale de l'ensemble.

LA50 : nombre de contigs plus longs que NA50

Pourcentage du génome (%) : pourcentage de bases qui correspondent au génome de référence.

Se salir les mains

Commençons par les expériences. Je vais utiliser l'assembleur SPAdes pour assembler des lectures obtenues à partir d'échantillons de patients séquencés. SPAdes utilise des lectures de séquençage de nouvelle génération. Vous pouvez également télécharger QAST gratuitement. Vous pouvez obtenir le code et les programmes binaires sur les pages d'accueil correspondantes et exécuter ces outils.

Entrez les commandes suivantes et vérifiez que les outils fonctionnent correctement.

```
<your_path_to>/SPAdes-3.13.1/bin/spades.py -h  
<your_path_to>/quast-5.0.2/quast.py -h
```

```
<your_path_to>/SPAdes-3.13.1/bin/spades.py -h  
<your_path_to>/quast-5.0.2/quast.py -h
```

[Ndlr : ils s'agit ici de lignes de commande informatiques sous systèmes Linux ou Unix.]

Les logiciels Spades et quast utilisés ont apparemment été créés en langage Python.]

Télécharger les données

Je suppose que vous savez comment télécharger les données du National Center for Biotechnology Information (NCBI). Sinon, vous pouvez vous référer à ce lien.

Les lectures pour nos expériences peuvent être téléchargées à partir de NCBI en utilisant le numéro d'accèsion NCBI SRX7636886. Vous pouvez télécharger l'exécution SRR10971381, qui contient des lectures provenant d'une exécution MiniSeq d'Illumina. Assurez-vous de télécharger les données au format FASTQ. Vous pouvez trouver le fichier téléchargé sous le nom de "sra_data.fastq.gz". Vous pouvez extraire le fichier FASTQ avec gunzip.

Après l'extraction, vous pouvez exécuter la commande bash suivante pour compter le nombre de lectures dans notre jeu de données. Vous verrez qu'il y a 56 565 928 lectures.

```
grep '^@' sra_data.fastq | wc -l
```

```
grep ,^@' sra_data.fastq | wc -l
```

[Ndlr : autre ligne de commande sous Linux ou Unix.]

Vous pouvez télécharger le génome complet SARS Cov-2^[1] disponible publiquement sur le NCBI avec le numéro d'accèsion GenBank MN908947 (N'oubliez pas qu'il ne s'agit pas des véritables données brutes). Vous verrez un fichier au format FASTA. Ce sera notre génome de référence. Notez qu'il a été renommé "MN908947.fasta".

Assembler

Assemblons les données provenant de l'échantillon COVID-19. Exécutez la commande suivante pour assembler les lectures avec SPAdes. Vous pouvez

transmettre le fichier .gz compressé directement à SPAdes.

```
<your_path_to>/SPAdes-3.13.1/bin/spades.py --12 sra_data.fastq.gz  
-o Output -t 8
```

```
<your_path_to>/SPAdes-3.13.1/bin/spades.py --12 sra_data.fastq.gz -o Output -t  
8
```

Nous avons utilisé l'assembleur général SPAdes comme démonstration pour cet article. Cependant, étant donné que le jeu de données de lecture est constitué de données RNA-Seq, il est préférable d'utiliser l'option "-rna" dans SPAdes.

Dans le dossier de sortie, vous verrez un fichier appelé "contigs.fasta" qui contient nos contigs assemblés.

Évaluation des résultats de l'assemblage

Exécutez QUAST pour les assemblages avec la commande suivante.

```
<your_path_to>/quast-5.0.2/quast.py Output/contigs.fasta-l  
SPAdes_assembly -r MN908947.fasta -o quastResult
```

```
<your_path_to>/quast-5.0.2/quast.py Output/contigs.fasta-l SPAdes_assembly -r  
MN908947.fasta -o quastResult
```

Visualisation du résultat de l'évaluation

Lorsque QUAST a terminé, vous pouvez aller dans le dossier "quastResult" et voir les résultats de l'évaluation. Vous pouvez visualiser le rapport QUAST en ouvrant le fichier report.html dans votre navigateur web. Vous verrez un rapport similaire à celui de la figure 3. Vous pouvez cliquer sur "Rapport avancé" pour obtenir plus d'informations comme NG50 et LG50.

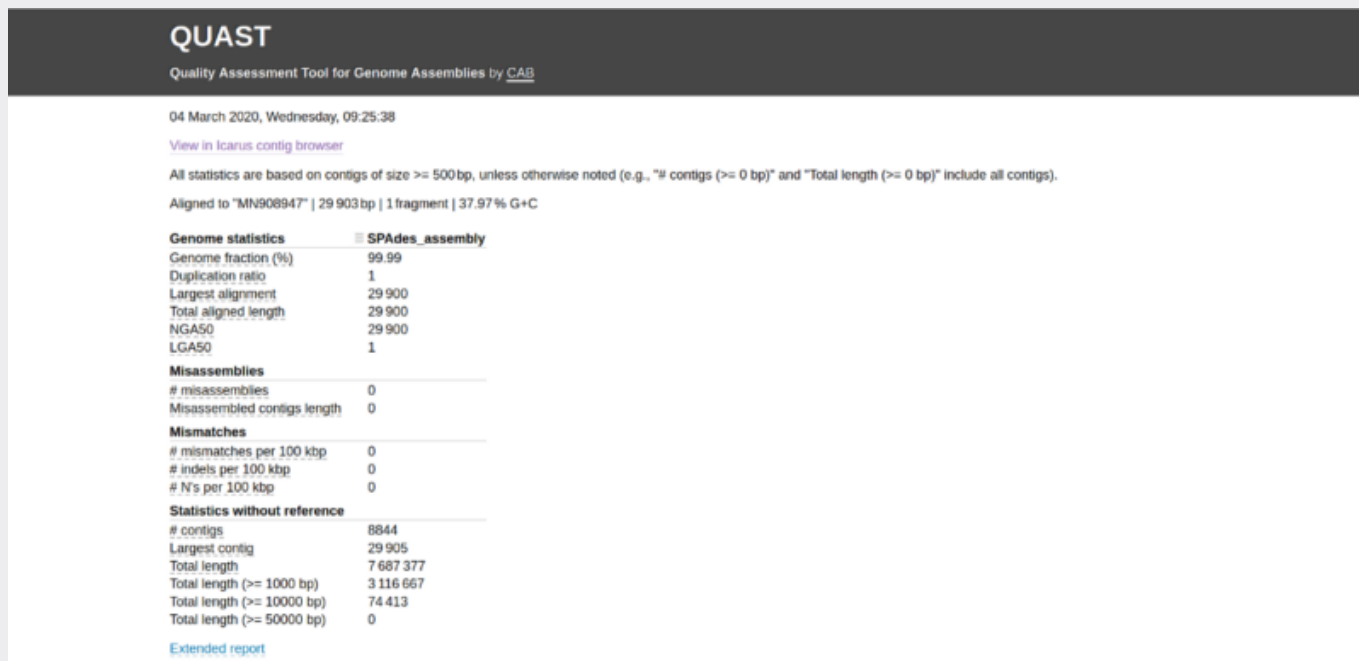


Fig. 3 : Rapport QUAST

Vous pouvez examiner les valeurs des différents critères de notation tels que la proportion de génome NG50, NA50, les mauvais assemblages et le nombre de contigs. Vous pouvez également visualiser l'alignement des contigs sur le génome de référence à l'aide du navigateur de contigs Icarus (cliquez sur "View in Icarus contig browser"), comme le montre la figure 4.

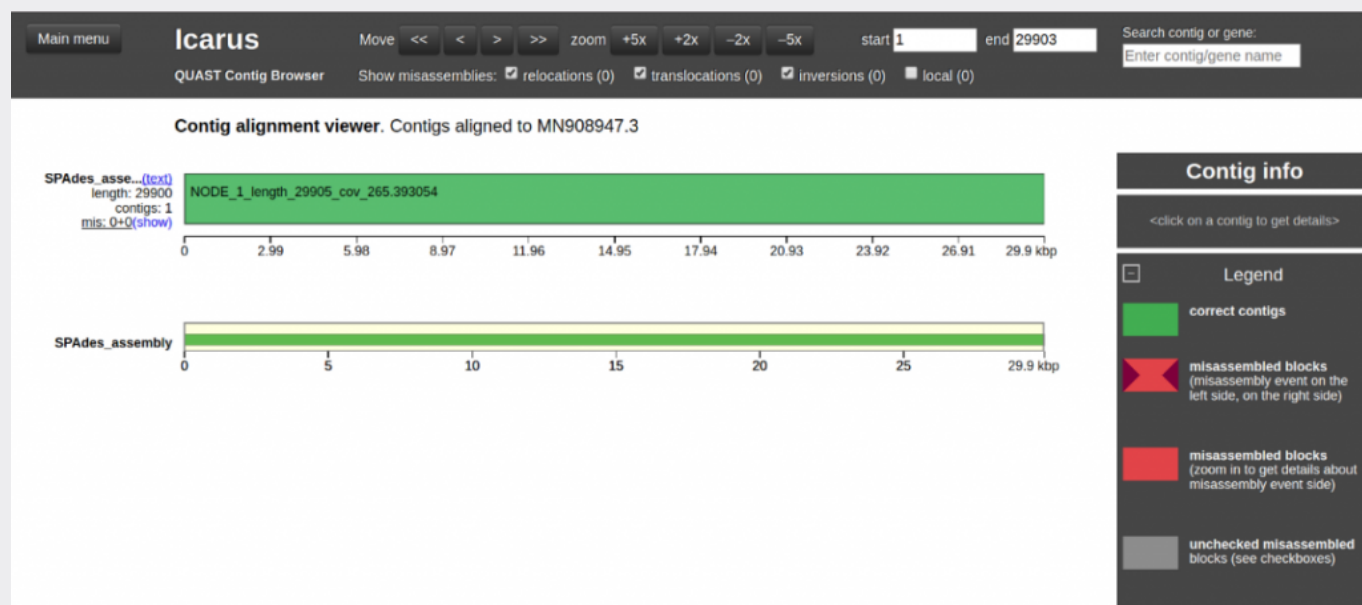


Fig. 4 : Rapport QUAST

Dans le navigateur de contigs d'Icarus, nous pouvons voir que le contig nommé "NODE_1" correspond très étroitement au génome de référence de COVID-19. Son contenu génomique est de 99,99 % (comme le montre la figure 3). En outre, la longueur totale alignée de 29 900 paires de bases est très proche de la longueur du génome de référence, qui est de 29 903 paires de bases.

Visualisation du graphique d'assemblage

Il existe un outil appelé Bandage qui vous permet de visualiser le graphique d'assemblage. Vous pouvez télécharger les exécutables précompilés depuis leur page d'accueil et exécuter l'outil. Vous pouvez charger le fichier graphique " assembly_graph_with_scaffolds.gfa " situé dans le dossier de sortie de SPAdes (allez dans Fichier → Charger le graphique → sélectionnez le fichier .gfa dans Sortie et ouvrez-le) dans Bandage et cliquez sur " Draw graph " pour visualiser comme indiqué dans la Figure 5. Notez que le long segment incurvé de couleur verte au milieu de la première rangée de segments de la figure 5 correspond à NODE_1 de notre assemblage SPAdes.

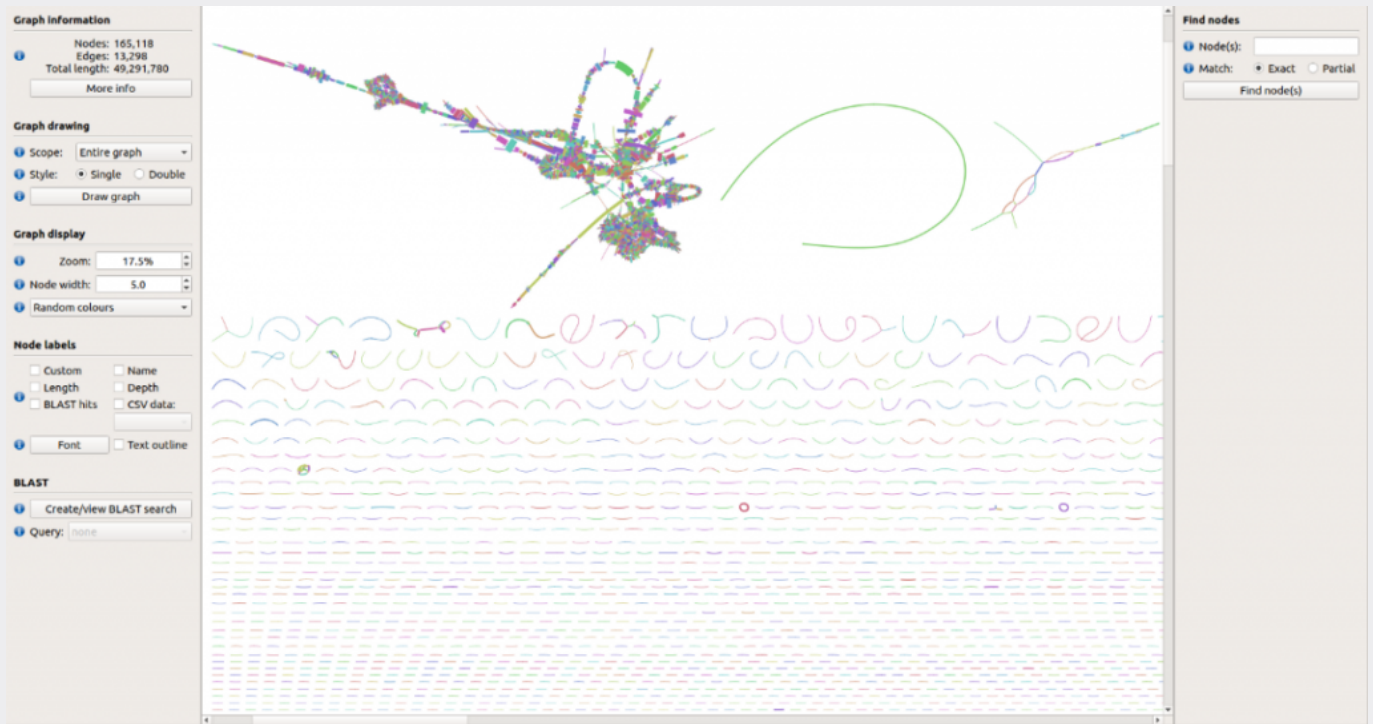


Fig. 5 : Partie du graphique d'assemblage SPAdes du jeu de données de lecture COVID-19 visualisée à l'aide de Bandage.

Exemple pratique de towardsdatascience

Comment ont-ils découvert le génome du SARS-CoV-2 ?

Maintenant que le génome de référence du SARS-CoV-2 est disponible, nous pouvons évaluer notre assemblage. Cependant, au départ, il n'existait pas de génome de référence exact pour le SARS-CoV-2, alors qu'ont fait les scientifiques pour le découvrir ? Comme expliqué dans cet article, l'analyse des génomes viraux relève de la métagénomique et il existe de nombreuses techniques pour y parvenir. Ils ont analysé la couverture des contigs (le nombre moyen de lectures couvrant chaque position de base dans un contig) et l'ont comparée à l'isolat de coronavirus (CoV) analogue au SRAS de la chauve-souris – bat SL-CoVZC45 (numéro d'accèsion GenBank MG772933)^[3] ^[6]. Les résultats ont montré que le contig le plus long assemblé avait une couverture "élevée" (à partir de notre assemblage, vous pouvez voir que NODE_1 a

également une couverture élevée) et était très étroitement lié à la chauve-souris SL-CoVZC45.

Les virologues chinois soulignent même explicitement que le brin de génome construit présente jusqu'à 90 % de similitude avec les brins de génome des coronavirus inoffensifs et revendiqués des chauves-souris, connus depuis des décennies.

Although 2019-nCoV is similar to some betacoronaviruses detected in bats (Figure 4), it is distinct from SARS-CoV and MERS-CoV. The three 2019-nCoV coronaviruses from Wuhan, together with two bat-derived SARS-like strains, ZC45 and ZXC21, form a distinct clade. SARS-CoV strains from humans and genetically similar SARS-like coronaviruses from bats collected from southwestern China formed another clade within the subgenus sarbecovirus. Since the sequence identity in conserved replicase domains (ORF1ab) is less than 90% between 2019-nCoV and other members of betacoronavirus, the 2019-nCoV — the likely causative agent of the viral pneumonia in Wuhan — is a novel betacoronavirus belonging to the sarbecovirus subgenus of Coronaviridae family.

Cette similitude de 90 % résulte du fait que c'est exactement un tel "génom" d'un "corona virus" de chauve-souris inoffensif qui a été utilisé comme modèle pour aligner les nombreuses séquences génétiques très courtes. Il convient toutefois de noter que ce génome de chauve-souris, comme tous les "génom" de "virus" pathogènes, n'a été que calculé, c'est-à-dire assemblé mentalement à partir de séquences génétiques endogènes très courtes et/ou de fragments de gènes de nombreux microbes, qui n'ont en réalité jamais été trouvés en tant que brin héréditaire entier et n'apparaissent nulle part dans la littérature scientifique comme un "génom viral" complet.

L'étude du 24.01.2020 indique sous la rubrique Discussion :

"Notre étude ne répond pas aux postulats de Koch".

Il est important de savoir ici que l'on ne connaît pas le pourcentage du génome "viral" calculé qui a été modifié avant d'être publié comme génome complet.

La publication de Fan Wu et al, dans Nature, Vol 579 du 3.2.2020, dans laquelle le génome (brin complet du génome) du SARS-CoV-2 a été présenté pour la première fois et avancé pour devenir le modèle de tous les alignements ultérieurs, a montré que la totalité de l'ARN obtenu à partir du lavage bronchique (BALF) d'un patient avait clairement été utilisée, sans aucun isolement ou enrichissement préalable des structures virales ou des acides nucléiques.

Dans cette publication, le professeur Zhang décrit comment il a calculé un génome de 29 903 nucléotides sur la base de courts segments de gènes d'une longueur de seulement 21 et 25 nucléotides (ce sont les paramètres par défaut des programmes d'alignement Megahit et Trinity utilisés) en utilisant une séquence donnée d'un génome (le virus inoffensif du coronavirus de chauve-

souris) selon sept méthodes différentes et très élaborées, y compris des méthodes statistiques.

Cet ARN a ensuite été converti en ADNc et des molécules d'une longueur de seulement 150 nucléotides ont été séquencées afin de construire le génome complet d'une longueur d'environ 30 000 nucléotides par simple calcul.

Il suppose – sans le dire explicitement – que les courtes séquences à partir desquelles il additionne la proposition de séquence du génome du virus SARS-CoV-2 sont de nature virale car il exclut les séquences plus longues qui résultent de la superposition (= contigs) des courtes pièces 21 et 25 et qui présentent des similitudes avec les séquences humaines issues de l'addition ultérieure au génome viral.

En termes simples, cela signifie que :

Comme les séquences humaines qui nous sont “connues” ont été calculées/supprimées du mélange de matériel génétique, les séquences restantes, nées de la réflexion irrationnelle des virologues (qui a été récompensée en 1954 par l'attribution du prix Nobel à John Franklin Enders pour une spéculation)^[7], sont supposées être précisément de nature virale.

Le pourcentage de l'ensemble du génome présentant des lacunes (de 1 % à la quasi-totalité ? ??) n'est pas indiqué.

Conséquence logique :

Ce qui a été créé artificiellement ici en toutes sortes d'étapes, toutes sous des “hypothèses” simplement crues, jamais vérifiées, n'a absolument RIEN à voir avec la réalité ! La méthode de séquençage ne peut ni dire de quelle source supposée provient le génome généré par de nombreuses étapes de calcul, ni affirmer qu'il est de nature purement virale. Il s'agit tout au plus d'un outil permettant de générer un nouveau brin de génome fictif construit à partir de nombreuses séquences génétiques très courtes, au moyen de divers algorithmes et de programmes dits de “gap-filling” (comblement des lacunes dans la construction du génome). Il est plus qu'effronté de prétendre que cette construction (du génome), qui a été faite à partir d'un échantillon via lavage bronchoalvéolaire d'un patient sans isoler une certaine structure, est virale juste parce que le patient a présenté des symptômes.

Les bioinformaticiens ne se soucient pas de la source de la séquence du gène.

Lorsque l'idée des brins génétiques des virus de la grippe a été conçue pour la première fois, des embryons de poulet ont été mécaniquement blessés et empoisonnés, et un modèle a été laborieusement créé à la main à partir des courtes séquences d'acide nucléique des tissus en décomposition. Aujourd'hui, cette tâche est effectuée par des programmes informatiques alimentés par les séquences que les virologues qualifient de “viraux”. L'origine réelle de ces

séquences est sans importance pour les bioinformaticiens qui créent les génomes des virus fictifs par "alignement". Erwin Chargaff avait d'ailleurs prévenu de cette évolution dès 1976 dans son livre "Das Feuer des Herakli" [L'incendie d'Herakli].

Sources :

[1] F. Wu, S. Zhao, B. Yu *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* (2020). <https://doi.org/10.1038/s41586-020-2008-3>

[2] Die behauptete SARS-CoV-2-Mutation aus England ist eine Mogelpackung

[3] Zhenyu Li *et al.* Comparison of the two major classes of assembly algorithms: overlap–layout–consensus and de-bruijn-graph, *Briefings in Functional Genomics*, Volume 11, Issue 1, January 2012, Pages 25–37. <https://doi.org/10.1093/bfgp/elr035>

[4] Jang-il Sohn and Jin-Wu Nam. The present and future of de novo whole-genome assembly. *Briefings in Bioinformatics*, Volume 19, Issue 1, January 2018, Pages 23–40. <https://doi.org/10.1093/bib/bbw096>

[5] S. Heerema and C. Dekker. Graphene nanodevices for DNA sequencing. *Nature Nanotech* 11, 127–136 (2016). <https://doi.org/10.1038/nnano.2015.307>

[6] Fan Wu: A new coronavirus associated with human respiratory disease in China

The viral genome organization of WHCV was determined by sequence alignment to two representative members of the genus *Betacoronavirus*: a coronavirus associated with humans (SARS-CoV Tor2, GenBank accession number [AY274119](#)) and a coronavirus associated with bats (bat SL-CoVZC45, GenBank accession number [MG772933](#)). The un-translational regions and open-reading frame (ORF) of WHCV

[7] Machtwerk – Einstieg in die Widerlegung der Virusbehauptung

Vous trouverez ci-dessous une liste de nos articles les plus importants [en allemand] :

<https://telegra.ph/Corona-Fakten-Liste-der-wichtigsten-Artikel-09-08>

Article original

: <https://www.samueleckert.net/durch-die-genom-analyse-werden-fiktive-viren-erschaffen/>