

SARS-CoV-2 et HIV : fake news ?

Extrait d'une activité de l'Atelier 5 : <http://education.expasy.org/bioinformatique/>
(Atelier 26 dans la liste <http://unige.ch/-/bioinformatique>)

Expasy Ateliers de bioinformatique

You are here: Home > Education > Ateliers de bioinformatique

00118926

Ateliers de bioinformatique

Une activité proposée par le **SIB Institut Suisse de Bioinformatique** (contact: [outreach\(at\)sib.swiss](mailto:outreach(at)sib.swiss)).

Définition: La **bioinformatique** est une discipline des sciences de la vie qui s'appuie sur des outils informatiques pour **stocker**, **analyser** et **visualiser** des données biologiques. C'est un domaine de recherche et d'expertise en pleine expansion et les applications sont très variées.

Exemples:

- Stocker la **séquence en acides nucléiques** du chromosome 1 humain
- Visualiser la **structure 3D** de l'insuline
- Prédire la **structure d'un gène** à partir d'une séquence ADN
- Choisir le meilleur traitement en fonction d'un **profil génétique**
- Découvrir **différentes cellules**

Vous allez découvrir des sites internet et des outils bioinformatiques utilisés par les biologistes du monde entier.
Le but de ces ateliers est de vous familiariser avec certains d'entre eux et surtout vous faire une idée de ce qu'on peut faire avec...

Une exposition virtuelle sur le **génomme humain**: www.ChromosomeWalk.ch (dossier pédagogique)
Comment les chercheurs conçoivent-ils **les médicaments de demain?**: www.atelier-drug-design.ch
NEW Un tour d'horizon pour comprendre **la médecine de précision**: www.PrecisionMed.ch

Atelier 1	Atelier 2
Des chromosomes et des gènes	Protéines et drug design
Atelier 3	Atelier 4
Phylogénie, biodiversité et pizza...	Médecine de précision et profil génétique
Atelier 5	Atelier 6
Coronavirus et protéines...	Découverte de BLAST...
Atelier 7	Quelques liens
L'insuline de A à Z	Pour les plus curieux.....



Recherche de similarité

Blast (recherche de similarité):

- <https://www.uniprot.org/blast/>
- <https://blast.ncbi.nlm.nih.gov/Blast.cgi>

La 'E value' est une probabilité de trouver le même résultat par hasard. Plus cette valeur est petite (< 0) plus le match est 'validé'.

SARS-CoV-2: un virus provenant d'une manipulation humaine ?


FAKE NEWS **FACTS**


- SARS-Cov-2 est très similaire à de nombreuses souches de coronavirus circulant dans la nature en Asie, avant et après la pandémie.
- Ces virus sont connus pour sauter d'une espèce à l'autre sans problème, donc rien de surprenant.
- L'analyse génétique de SARS-CoV-2 présente une organisation génomique et des protéines en tout point similaires aux autres virus sauvages.

Les thèses selon lesquelles une partie d'HIV a été insérée par l'homme dans le virus sont de mauvaises interprétation des résultats de recherche de similarité (BLAST)

Uncanny similarity of unique inserts in the 2019-nCoV spike protein to HIV-1 gp120 and Gag

Prashant Pradhan, Ashutosh Kumar Pandey, Akhilesh Mishra, Parul Gupta, Praveen Kumar Tripathi, Manoj Balakrishnan Menon, James Gomes, Perumal Vivekanandan, Bishwajit Kundu

doi: <https://doi.org/10.1101/2020.01.30.927871>

This article is a preprint and has not been certified by peer review [what does this mean?].

Abstract

Info/History

Metrics

 Preview PDF

Abstract

This paper has been withdrawn by its authors. They intend to revise it in response to comments received from the research community on their technical approach and their interpretation of the results. If you have any questions, please contact the corresponding author.

**FAKE NEWS
FACTS**

[Ce papier](#) a été démenti et retiré. Les auteurs ont utilisé des fragments de 6 à 10 acides aminés: il existe près de 1 millions de séquences d'HIV hypervariables, c'est inévitable qu'il y ait des similarités juste par hasard. Ce papier, même si il n'a jamais été publié, a fait le buzz...

Pour les experts (1):

[Emerg Microbes Infect.](#) 2020; 9(1): 378–381.

PMCID: PMC7033698

Published online 2020 Feb 14. doi: [10.1080/22221751.2020.1727299](https://doi.org/10.1080/22221751.2020.1727299)

PMID: [32056509](https://pubmed.ncbi.nlm.nih.gov/32056509/)

HIV-1 did not contribute to the 2019-nCoV genome

[Chuan Xiao](#),^{a,CONTACT} [Xiaojun Li](#),^b [Shuying Liu](#),^c [Yongming Sang](#),^d [Shou-Jiang Gao](#),^e and [Feng Gao](#)^{b,f}

▶ [Author information](#) ▶ [Article notes](#) ▶ [Copyright and License information](#) [Disclaimer](#)

Voici des exemples de séquences d'acides aminés retrouvées dans les protéines de SARS-CoV-2 utilisées pour 'démontrer' que SARS-CoV-2 contient des morceaux de génome de HIV:

TNGTKR
HKNNKS
RSYLTPGDSSSG
QTNSPRRA

La 'E value' est une probabilité de trouver le même résultat par hasard. Plus cette valeur est petite (< 0) plus le match est 'validé'.

Faire un Blast 'protein' au NCBI: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>

Sélectionner Organism: "Human immunodeficiency virus (taxid:12721)"

Cliquer sur 

Créer une séquence aléatoire avec les mêmes 'lettres' ([edit sequence]) et refaire un BLAST.

Que pouvez-vous conclure?

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7033698/>

Pour les experts (2):

Voici un morceau du gène codant pour la protéine Spike:

AATGGTACTAAGAGGTTTGATAACCCTG

Faire un Blast 'nucléotide' au NCBI: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>.

Sélectionner Organism: "Human immunodeficiency virus (taxid:12721)"

Cliquer sur **BLAST**

La 'E value' est une probabilité de trouver le même résultat par hasard. Plus cette valeur est petite (< 0) plus le match est 'validé'.

Créer une séquence aléatoire avec les mêmes 'lettres' ([edit sequence]) et refaire un BLAST.

Que pouvez-vous conclure?

A HIV-1 isolate 19828.PPH11 from Netherlands envelope glycoprotein (env) gene, partial cds				
Sequence ID:	HQ644953.1	Length: 1143	Number of Matches: 1	Range 1: 967 to 994
Score	Expect	Identities	Gaps	Strand
38.3 bits(41)	7.5	25/28(89%)	0/28(0%)	Plus/Plus
Query	86	AATGGTACTAAGAGGTTTGATAACCCTG	113	
Sbjct	967	AATGGTACTAAAAGGTTAGATAACACTG	994	

B HIV-1 isolate patient B clone 16.3 from Netherlands envelope glycoprotein (env) gene, complete cds				
Sequence ID:	HQ386166.1	Length: 2580	Number of Matches: 1	Range 1: 2493 to 2523
Score	Expect	Identities	Gaps	Strand
39.2 bits(42)	2.1	27/31(87%)	0/31(0%)	Plus/Minus
Query	351	CCTAAAAGTTCCTTGTGAATAACTGTATTATT	381	
Sbjct	2523	CCTAAAAGTTCCTTGTGAATAATTCTATAATT	2493	

la séquence aléatoire. On peut en conclure que la similarité entre la séquence codante de la protéine S et le génome du VIH n'est pas significative. Les alignements ont été réalisés sur le site BLAST du NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

Figure 6. Recherche de similarités entre les séquences codant pour la protéine spike de CoV2 et le génome de HIV. A. Alignement le plus significatif entre la séquence codant pour la protéine S de SARS-CoV-2 (query) et le génome du VIH (subject). B. Contrôle négatif : alignement le plus significatif entre une séquence aléatoire, obtenue en mélangeant les nucléotides de la séquence précédente, et le génome du VIH. Noter la valeur du score *expect*, qui indique le nombre de faux-positifs attendus au hasard. Ce score présente pour les deux alignements des valeurs supérieures à 1, et est même plus élevé pour l'alignement de la séquence de CoV que pour