

Compte-rendu de la journée du PHROGATON

Visioconférence et atelier du 04/06/2021

Groupe de travail "Génomique et biologie computationnelle"

La journée du PHROGATON s'est déroulée en visioconférence la 04/06/2021. Cette journée visait à présenter la base de données PHROGS développée par le LGME (CNRS – Université Clermont-Auvergne) et ses collaborateurs. Cette base de données est dédiée aux virus des bactéries et des archaea. Basée sur la création de groupes de protéines virales orthologues (les PHROGs), elle vise à fournir aux annotateurs de génomes viraux une annotation standardisée.

La matinée a regroupé une cinquantaine de participants et a été consacrée à la présentation par François ENAULT et Eric OLO NDELA de la base de données et de l'interface web qui permet son interrogation. Les participants ont pu interroger les auteurs et apporter des suggestions.

L'après-midi a regroupé une vingtaine de participants durant un visio-atelier au cours duquel Marie-Agnès PETIT, Julien LOSSOUARN et Nicolas GINET ont présenté en quelques diapositives des cas concrets d'utilisation de la base de données PHROGS et des possibilités d'amélioration experte de l'annotation prévue par l'interface web. Nous avons pu aussi discuter des points de vigilance inhérents à tout système d'annotation basé en partie sur l'annotation automatique. Clovis GALIEZ nous a précisé les principes de l'utilisation des outils proposés par la plateforme HHpred et la construction de profils HMMs (Hidden Markov Models pour la détection d'homologies lointaines).

En conclusion de la journée il a été fait appel à l'expertise des scientifiques du réseau pour rendre la base de données PHROGS encore plus pertinente et en faire un outil de référence pour l'annotation des génomes viraux.

Cette journée a mis en évidence l'intérêt des ateliers dédiés à l'annotation de génomes. C'est bien souvent le premier pas de sujets de recherche divers, allant du plus fondamental au plus appliqué. La diversité des affiliations des personnes qui ont assisté à cette journée le démontre (laboratoires académiques, entreprises privées). L'annotation fonctionnelle est en constante évolution du fait de l'augmentation des données de génomique constamment publiées et des méthodes de bio-informatiques pour faire sens de cette masse de données. Le réseau Phages.fr poursuivra ses efforts de "veille" pour proposer de nouveaux ateliers.



Grâce aux échanges entre les utilisateurs et les concepteurs de la base de données PHROGS durant cette journée, Eric OLO NDELA a pu rapidement corriger quelques bugs mineurs et intégrer des améliorations et de nouvelles fonctionnalités au site (<https://phrogs.lmge.uca.fr/>). Il a aussi apporté quelques précisions à la documentation associée. Nous pouvons remercier Eric pour sa réactivité ! Les modifications sont résumées ci-dessous.

Page d'accueil :

- Il est désormais possible d'aller chercher un phrog juste avec son numéro
- Il est maintenant possible de chercher un virus par son nom plutôt que son numéro accession
- Nous pouvons retrouver une protéine du NCBI grâce à son identifiant ou un petit morceau de sa séquence

Pages individuelles des phrogs :

- L'utilisateur dispose maintenant des liens vers les protéines du NCBI sur le premier grand tableau (*list of all proteins*)
- Sur ce même tableau les protéines apparaissent maintenant dans le même ordre que l'alignement multiple en bas

Pages individuelles des virus :

- Les liens vers les protéines du NCBI sont maintenant disponibles dans le tableau des protéines du virus
- Sur les cartes génomiques, il est désormais possible de cliquer sur les gènes pour aller voir le phrog correspondant (sauf si c'est un singleton)

Page de documentation :

- Une phrase a été ajoutée dans les manuels de comparaison :

"NOTE : If your query is not similar to any PHROG, you can try screening a larger database like UNICLUST or NR to build an enriched query set and have better results against PHROGs."

- Une phrase a été ajoutée à la section concernant HHsuite :

"Try running HHSearch if HHblits doesn't retrieve any hits."

