

Réunion du 19/11/2015 – Un projet Bioinformatique pour USPC ?

Présents : Franck Letourneur, Jean-Philippe Jais, Patrick Nitschke, Claudine Pique, Georges Bismuth, Stefano Marullo, Valentina Boeva, Bertrand Cosson.

Cette rencontre a permis d'échanger sur les pratiques et les besoins en bioinformatique. Le but est de recenser les besoins au sein d'USPC, et proposer des actions à prioriser à moyen terme. Cet échange est particulièrement intéressant dans le cadre de la préparation de la réunion biologie computationnelle / bio informatique du pôle SVS de l'USPC aura lieu le 23 novembre.

Discussion sur les accès à l'analyse en génomique (au sens large) :

Les besoins et les compétences au sein des équipes d'USPC sont très hétérogènes. Vu du biologiste, l'accès aux analyses peut être perçu selon une échelle de niveaux :

– Niveau I :

Accès à une plateforme de service pour les besoins ponctuels ou les demandes récurrentes

Activité de bioanalyse/biostatistique pour les biologistes qui récupèrent les données de la plateforme (ou réalisé sous forme collaborative si la compétence n'est pas suffisante au sein de l'équipe)

– Niveau II :

Accès à une plateforme de fouille et de gestion de données et d'archivage électronique destinée à rendre la bioinformatique accessible aux chercheurs n'ayant pas de compétence en programmation informatique.

– Niveau III :

Travail avec des outils en ligne de commande.

– Niveau IV

Accès à un serveur de calcul de grande capacité, utilisation de langage de programmation.

Besoins des équipes et plateformes d'USPC

- Besoins en formation

Une forte demande concerne la formation continue des personnels USPC, sur un format plus exigeant que les formations de courtes durées proposées aujourd'hui (souvent sur 1 à 3 jours), mais restant compatible avec une vie professionnelle.

Proposition : création d'un Diplôme Universitaire proposant des cours sur 2 à 3 jours par mois ; incluant une partie projet pouvant permettre de traiter des données de sa propre activité de recherche.

- Besoin en assistance / accompagnement suivant les formations

En complément des formations, une demande récurrente des biologistes désirant accéder à des compétences de bioanalyse/bioinformatique concerne l'accompagnement dans la mise en place des outils et l'assistance dans leur utilisation.

Proposition : réfléchir à un nouveau modèle permettant l'accès à des référents en informatique et en bioinformatique pour la veille technologique sur les outils, l'évolution de leur usage, et accompagner l'acquisition de compétences des personnels. Voir ci dessous la proposition dans le paragraphe 2) plateformes de service

- Besoins des plateformes :

1) Plateformes de séquençage : offre matérielle et logicielle pour répondre à l'ensemble des besoins de séquençage et réaliser les « premières » opérations de manipulations de données. La plateforme de Paris Descartes propose par exemple un ensemble d'outils bioinformatique pour réaliser certaines étapes : démultiplexage, mapping.

Proposition : Quel rôle la plateforme de séquençage de Paris Descartes souhaite t elle jouer à l'échelle de l'USPC ? Quels moyens seraient ils nécessaires ?

2) Plateformes de service : seule Paris Descartes dispose aujourd'hui d'une plateforme de service dédiée à l'analyse en génomique. Elle joue un rôle prépondérant pour fédérer un réseau de compétences, et représente pour certaines équipes l'unique accès à l'analyse bio-info. Quid de l'accès des équipes de recherche des autres établissements ? Du développement de nouveaux axes (mais basé sur le même fonctionnement, les mêmes structures ?) : l'étude de l'ARN (transcriptome, régulation transcriptionnelle, structure des ARN) et l'épigénétique, la bactériologie ?

Proposition : Une idée a été énoncée : disposer d'une plateforme bio-informatique proposant des services couvrant les besoins ponctuels ou les demandes récurrentes, et disposant d'une salle d'ordinateurs en accès libre pour accueillir les personnels USPC qui pourraient y effectuer leurs propres analyses, et bénéficier des conseils des bio-informaticiens de la plateforme. Où et comment l'organiser ? Avec des personnels appartenant aux différents établissements ?

3) Plateformes de fouille et de gestion de données et d'archivage électronique : rend service à une communauté de biologistes non programmeurs, non utilisé par les informaticiens. Permet travail collaboratif et traçabilité. Nécessite à minima un interlocuteur ingénieur de plateforme.

Proposition : organiser l'accès à une plateforme extérieure ou développer un serveur dédié au sein d'USPC (nécessite serveur + ingénieur pour sa gestion, sa maintenance et l'installation d'outils et base de données) ?

4) Plateformes de distribution de logiciels : de nombreux serveurs existent en accès libre (comme Migale à l'INRA par exemple).

Proposition : Evaluer la plus-value de développer un serveur « USPC » de ce type, qui se justifierait par exemple si une offre de service la complétait (voir ci dessous) ou si un nombre suffisant de développements des plateformes/équipes USPC était mis en accès libre ?

5) Plateformes de serveur de calcul et de stockage : l'accès aux grands serveurs de calculs nationaux est jugé délicat.

Proposition : réflexion sur l'accès aux serveurs existants ou sur l'opportunité de déployer un serveur USPC. Des travaux en cours visent à interconnecter les établissements de l'USPC par un réseau très haut débit.