



Données génomiques massives : applications, enjeux computationnels et perspectives

Eric Rivals

► **To cite this version:**

Eric Rivals. Données génomiques massives : applications, enjeux computationnels et perspectives. Open Science, Open Access, Open Data et données, Jun 2018, Strasbourg, France. lirmm-01951492

HAL Id: lirmm-01951492

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-01951492>

Submitted on 11 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Données génomiques massives : applications, enjeux computationnels et perspectives

Eric Rivals, LIRMM & Institut de Biologie Computationnelle (IBC)

CNRS et Université de Montpellier, France

Résumé pour l'école d'été "Open Science, Open Access, Open Data et données"
Université de Strasbourg, France. 19-21 juin 2018.

De nombreux aspects du fonctionnement biologique peuvent être étudiés grâce aux nouvelles technologies de séquençage à haut débit (SHD) : de l'échelle moléculaire (ex. épigénomique ou épitranscriptomique) jusqu'à celle d'un système écologique (structure et dynamique des interactions entre espèces d'un environnement). Le SHD est devenu indispensable aux avancées des recherches fondamentales aussi bien qu'aux développements industriels – par ex. en agronomie. Bien évidemment, la médecine de précision (aussi nommée médecine personnalisée) s'appuie sur ces technologies qui commencent à s'implanter dans les hôpitaux.

Les caractéristiques principales des données de séquençage sont d'une part le volume de données produites par séquençage (volume qui augmente exponentiellement), d'autre part la capacité d'exploration exhaustive (du génome, du transcriptome, des interactions dans leur ensemble). L'enjeu computationnel majeur devient alors d'analyser avec précision, efficacité et robustesse ces données génomiques. Après quelques exemples d'applications, j'aborderai le rôle clef de l'indexation de données dans les algorithmes d'analyse de SHD. Les données génomiques codent une part cruciale de ce qui détermine un individu et contiennent des informations sur sa vie et son histoire. Je conclurai sur les enjeux globaux pour la recherche et la société que posent la disponibilité de ces informations.

