

## Boîte à outils de bioinformatique de MATLAB

Bioinformatics Toolbox™ fournit des algorithmes et des applications pour le séquençage de nouvelle génération (NGS), l'analyse de microréseaux, la spectrométrie de masse et l'ontologie des gènes. À l'aide des fonctions de la boîte à outils, vous pouvez lire des données génomiques et protéomiques à partir de formats de fichiers standard tels que SAM, FASTA, CEL et CDF, ainsi que dans des bases de données en ligne telles que NCBI Gene Expression Omnibus et GenBank. Vous pouvez explorer et visualiser ces données avec des navigateurs de séquences, des cartes thermiques spatiales et des diagrammes de cluster. La boîte à outils fournit également des techniques statistiques pour détecter les pics, imputer des valeurs pour les données manquantes et sélectionner des entités.

Vous pouvez combiner les fonctions de la boîte à outils pour prendre en charge les workflows bioinformatiques courants. Vous pouvez utiliser les données ChIP-Seq pour identifier les facteurs de transcription; analyser les données RNA-Seq pour identifier les gènes différentiellement exprimés; identifier les variantes du nombre de copies et les SNP dans les données de microréseau; et classer les profils protéiques à l'aide de données de spectrométrie de masse.

## Analyse de séquence

### **Alignement de séquence**

La boîte à outils fournit des fonctions, des objets et des méthodes pour l'analyse de séquence, y compris la séquence par paire, le profil de séquence et l'alignement de séquences multiples. Ceux-ci inclus:

Implémentations MATLAB d'algorithmes standard pour l'alignement de séquence local et global, tels que :

- Les algorithmes de modèle de Needleman-Wunsch, Smith-Waterman et par profil
- Alignement multiple des séquences
- Représentations graphiques des matrices de résultats d'alignement
- Matrices de notation standard, telles que les familles de matrices PAM et BLOSUM
- Calcul de séquence de consensus et affichage du logo de séquence