

# Fiche UE MU5IN753-6 - Statistiques pour la classification et fouille de données en génomique (SPLEX) – Mutualisée avec le master BMC et le master informatique

<b>Responsables</b>	P.-H. Wuillemin, N. Sokolovska, J.-D. Zucker			
<b>Descriptif</b>	<b>Niveau</b>	<b>Semestre d'enseignement</b>	<b>ECTS</b>	<b>Effectif maximal</b>
	M2	S1	6	
<b>Modalités pédagogiques</b>	<b>Volume horaire Cours</b>	<b>Volume horaire TD</b>	<b>Volume horaire TP</b>	<b>Présentiel/Distanciel</b>
	10		20	Présentiel
<b>Objectifs</b>	<p>Introduire les méthodes d'analyse de données complexes, depuis l'analyse statistique classique jusqu'aux plus récentes techniques d'apprentissage automatique. Présenter les principales questions analytiques en génomique fonctionnelle et les approches méthodologiques permettant une analyse exploratoire ou la construction automatique de modèles prédictifs en intégrant des données hétérogènes: transcriptomiques, cliniques, génomiques, métabolomiques ou métagénomiques, etc. On abordera en particulier les données haut-débits rencontrées en transcriptomique et métagénomique. Différentes applications médicales seront présentées pour illustrer les problèmes posés par l'analyse du transcriptome.</p>			
<b>Thèmes abordés</b>	<ul style="list-style-type: none"> <li>• Statistique non paramétrique</li> <li>• Modèle linéaire et linéaire généralisé ; Analyse de données (analyse multivariée, statistique euclidienne, ACP, analyse des correspondances multiples, analyse canonique, etc.) ;</li> <li>• Méthodes à noyaux ;</li> <li>• L'étude du transcriptome par les données puces et séquençage (NGS), standards de représentation et stockage, la normalisation ;</li> <li>• Principales directions analytiques, ressources et approches utilisées, tests statistiques, ajustement multiples, l'échantillonnage et les techniques de validation utilisées pour l'analyse du transcriptome ;</li> <li>• Etude de l'expression différentielle, ressources et approches d'analyse fonctionnelle transcriptomique ;</li> <li>• Echantillonnage, classification supervisée (arbre de décision, règles de décision, k-PPV, Séparateurs à Vastes Marges (SVM), etc.) ;</li> <li>• Classification non supervisée (classification ascendante hiérarchique, k-means, règles d'associations, etc.) des données et applications aux données transcriptomiques dont cinétiques d'expression ;</li> <li>• L'analyse intégrative des réseaux transcriptionnels : utilité, approches conventionnelles et intégratives, applications ;</li> <li>• Directions futures de développement d'approches computationnelles intégratives pour la biologie des systèmes dont le "séquençage de nouvelle génération" (NGS).</li> </ul>			

# Fiche UE MU5IN753-6 - Statistiques pour la classification et fouille de données en génomique (SPLEX) – Mutualisée avec le master BMC et le master informatique

<b>Compétences acquises à l'issue de l'UE (concepts, méthodologie et outils)</b>	<ul style="list-style-type: none"> <li>• Compréhension de modèle avancées de physiologie</li> <li>• Modèles de fonctionnement tissulaires (tissu adipeux, intestinal)</li> <li>• Modèle dynamique spatialisé</li> <li>• Introduction aux équations au dérivées partielles (EDP)</li> <li>• Introduction aux méthodes individu-centrées (IBM)</li> </ul>			
<b>Modalités d'évaluation/100</b>	<i>Ecrit</i>	<i>Oral</i>	<i>CC</i>	<i>Autre</i>
<b>Langues utilisées</b>	<i>Dans les cours, TD, TP</i>		<i>Dans les documents, supports</i>	
	FR		EN	
<b>Localisation</b>	Campus Jussieu			