

FORMULAIRE STAGE Recherche-M2 BBSG
(du 5 janvier 2017 au 3 juillet 2017)

Titre du stage :

Benchmarking dimensionality reduction techniques for cancer pharmacogenomics data

Laboratoire (intitulé, adresse, site web) :

Centre de Recherche en Cancérologie de Marseille (CRCM), 27 Boulevard Leï Roure, 13009 – Marseille
<http://crcm.marseille.inserm.fr/en/>

Equipe :	Ballester
Maitre de stage :	Dr Pedro Ballester
E-mail :	pedro.ballester@inserm.fr
Téléphone :	04 86 97 74 21

Descriptif du stage :

Cancer pharmacogenomics modelling [1] is challenged by the high dimensionality of the molecular profiles that characterise the studied tumours. These profiles often include detected single nucleotide variants, copy number alterations and gene expression levels in the tumour exome. Consequently, there is a need to reduce the dimensionality of tumour profiles to gain interpretability while retaining the predictive power of these models.

This 6-month project aims at assessing the performance of existing dimensionality reduction techniques for genomic and transcriptomic data from public pharmacogenomics resources. These techniques range from those exploiting cancer pathways knowledge [2] to those entirely data driven [3].

This project is suitable for a self-motivated and diligent master student wishing to gain experience in advanced data analysis techniques and their application to cancer research. Ability to code in R and/or Python is required.

[1] Menden, M. P. et al. Machine Learning Prediction of Cancer Cell Sensitivity to Drugs Based on Genomic and Chemical Properties. PLoS ONE 8, e61318 (2013).

[2] Alcaraz, N. et al. KeyPathwayMiner 4.0: condition-specific pathway analysis by combining multiple omics studies and networks with Cytoscape. BMC Syst Biol 8, 99 (2014).

[3] Haury, A.-C., Gestraud, P. & Vert, J.-P. The Influence of Feature Selection Methods on Accuracy, Stability and Interpretability of Molecular Signatures. PLoS ONE 6, e28210 (2011).