

**CRUK Cambridge Centre MRes rotation project**

Rotation Project Title	Machine Learning for Mass Spectrometry Imaging of Tumours
Head of Laboratory (PI) Name	Jules Griffin
Second supervisor if applicable	
Programme	CMB
Supervisor's Email	Jlg40@cam.ac.uk
Laboratory Location	Department of Biochemistry

Project Outline	<p>Aims and objectives</p> <p>Using mass spectrometry imaging of tissue slices we have examined the distribution of lipids and metabolites across tumours in the liver (with Mike Allison and Michele Vacca, Liver Clinic, Addenbrooke's Hospital) and lung (with Gerard Evan and Trevor Littlewood, Biochemistry), using a laser to raster across the surface of the tissue at 40-50 um steps. Each pixel of the resultant image represents a mass spectrum and using current approaches measures over 300 metabolites. The aim of this rotation project is to explore the use of machine learning techniques to identify patterns in the mass spectral images. While we have written software to process the mass spectrometry images it relies on principal component analysis to process the data. We will explore machine learning approaches to examine what extra information can be mined in these datasets.</p>
Experimental plan	<p>This project will use statistical pattern recognition techniques that rely in part on machine learning approaches, and in particular classifier learning methods such as random forest, k-nearest neighbour and support vector machines, to process the complex, megavariable data obtained using mass spectrometry imaging. We will also explore how to fuse this data with conventional histology staining, clinical chemistry and conventional metabolomics to have a holistic picture of metabolism. All the data required for this project has already been acquired although the student could also acquire data of their own.</p>
Main Techniques	<ul style="list-style-type: none"> • Familiarity with R or Matlab to allow the student to write their own scripts. • Interest in machine learning techniques. • This project would best suit a systems biologist with interests in both wet and dry lab projects (I would hope they also acquire data as well as process it).
Key References	<p>Bond NJ, Koulman A, Griffin JL, Hall Z. <i>massPix</i>: an R package for annotation and interpretation of mass spectrometry imaging data for lipidomics. <i>Metabolomics</i>. 2017;13(11):128. doi: 10.1007/s11306-017-1252-5.</p> <p>Hall Z, Ament Z, Wilson CH, Burkhart DL, Ashmore T, Koulman A, Littlewood T, Evan GI, Griffin JL. Myc Expression Drives Aberrant Lipid Metabolism in Lung Cancer. <i>Cancer Res</i>. 2016 Aug 15;76(16):4608-18. doi: 10.1158/0008-5472.CAN-15-3403.</p>



Hall Z, Bond NJ, Ashmore T, Sanders F, Ament Z, Wang X, Murray AJ, Bellafante E, Virtue S, Vidal-Puig A, Allison M, Davies SE, Koulman A, Vacca M, Griffin JL. Lipid zonation and phospholipid remodeling in nonalcoholic fatty liver disease. *Hepatology*. 2017 Apr;65(4):1165-1180. doi: 10.1002/hep.28953.