Opportunity 2020045

Novel artificial intelligence methodology for predicting the most effective personalised drug treatment for patients

Drug Ranking Using Machine Learning (DRUML)

DRUML is an artificial intelligence tool for ranking drugs in order of their efficacy for treating a particular patient. DRUML uses proteomic, phosphoproteomic and transcriptomic data, improving accuracy vs traditional methods using genomic data. In principle, any largescale omics dataset can be used as the input of DRUML and it does not require a reference sample for comparison.

Background

When treating a cancer patient, selecting the most promising drug is a complex process. Tumours are highly heterogeneous, meaning that patients with the same pathological classifications may widely differ in their response to the same therapy. Personalised medicine aims to identify individualised therapeutic interventions by indicating the effectiveness of specified drugs or their combinations. The application of machine learning (ML) to predict drug responses has the potential to revolutionise cancer treatment. Stratifying patients for therapy has mainly relied on the association of genetic markers with drug response, however recently proteomic data has been suggested as a more effective predicter due to the high frequency of false positives and negatives seen in genomics-derived results. The use of large-scale proteomics and phosphoproteomics in ML has not yet been systematically applied. Recent advances in proteomic techniques and an increased number of drug response profiles means that this data can now be used as the input for predictive ML models of drug response, advancing the field of precision medicine.

The Problem

- Selecting the most effective drug treatment for individual patients is complex
- Tumour heterogeneity means patients respond differently to the same treatment
- Proteomics look to be more effective at predicting drug response than genomics
- Low sample throughout previously made it difficult to use proteomics for ML models

Invention: Benefits & Application

DRUML is a methodology for building and integrating ML models, using ensembles of proteomic, phosphoproteomic and transcriptomic features to generate lists of ranked drugs based on their efficacy. DRUML is able to predict drug rankings within a cancer cell population, without the need of reference samples for comparison; a crucial requirement for the clinical implementation of ML and a core aim of precision medicine. DRUML requires a dataset of expression values of biological markers from a sample taken from the patient. From this dataset, DRUML calculates a plurality of drug response distance values for each drug, from which they are ranked in order of predicted efficacy. The method predicts treatment for both human and animal subjects, with equal use in both human and/or veterinary medicine.

A patent claiming the use of DRUML has been filed. Queen Mary Innovation are now seeking partners who would be interested in licensing this technology for commercial development.

Patent Application Number: WO 2022/013562

Contact:
Monika Hamilton PhD CLP
Senior Commercialisation Manager
monika.hamilton@qmul.ac.uk

Project Development

To develop DRUML, 48 carcinoma cell lines (incl. AML, oesophageal, hepatocellular) were analysed in triplicate by LC-MS/MS, producing a phosphopeptide and protein dataset, generating over 4 million quantitative data-points. This approach was systemically applied to 466 drugs, with each being annotated for phosphorylation site markers of drug responses for AML and solid models respectively, and 40-50 protein markers of resistance or sensitivity. Importantly, the markers of drug response grouped drugs by their mode of action, reinforcing the notion that these markers are indicative of the biological mechanisms that determine drug response. DRUML was verified using data from publicly available label-free datasets, to test whether the models generated with the training datasets would predict the actual drug responses. Data from 53 diverse cancer cell models (cell lines and solid tumours) was analysed using DRUML, showing significant correlation between DRUML-derived drug response predictions and the actual responses. This was true across drugs with diverse modes of action and developmental phase (fig 1). 80% of the drugs could be ranked with accuracy ±15%, and 95% of the drugs with accuracy ±25%. The verification datasets consisted of cell lines derived from bone, brain, breast, cervix, colorectal, ovary and prostate cancers, showing remarkable accuracy of DRUML which was trained using oesophageal and liver cancers.

Overall these data indicate that proteomics data, collected using routine LC-MS/MS, may be used as the input of DRUML to accurately predict and rank the efficacy of drugs of diverse mode of action in cancer cells derived from different pathologies.

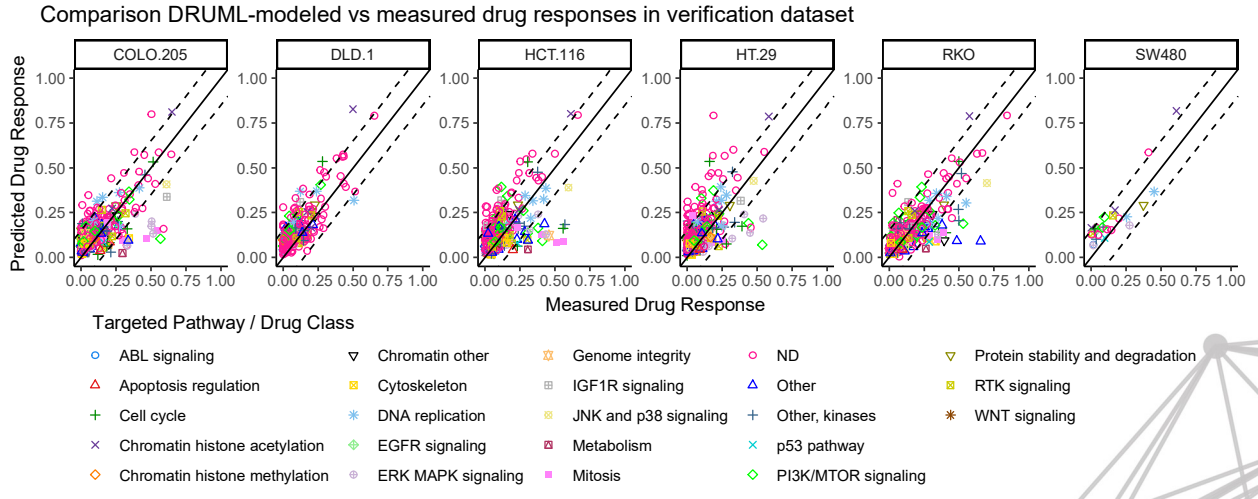

Figure 1. Comparison of measured and predicted drug responses within cellular models. Each data-point represents a drug prediction. Drugs are coloured according to mode of action and class. Dashed lines represent 10% error boundaries.

Inventor

Dr Pedro Cutillas, BSc, PhD.
Centre for Haemato-Oncology, Barts Institute of Cancer.

Inventor's Profile: https://www.bartscancer.london/staff/professor-pedro-cutillas/

Publication

Gerdes, H., Casado, P., Dokal, A. *et al.* Drug ranking using machine learning systematically predicts the efficacy of anti-cancer drugs. *Nat Commun* **12**, 1850 (2021). https://doi.org/10.1038/s41467-021-22170-8

Contact:
Monika Hamilton PhD CLP
Senior Commercialisation Manager
monika.hamilton@qmul.ac.uk