RAPID RESPONSE DATA ANALYSIS PUMP PRIME CALL

———————————————

*The Onco-Innovation Programme at the Milner Therapeutics Institute is seeking proposals from*

*CRUK Cambridge Centre members, for data analysis projects that could benefit from our multimodal data-driven AI/machine learning approaches.*

———————————————

Do you have clinical or experimental “multi-omics” data sets that would benefit from expert data integration and unsupervised network-based machine learning analysis? Our computational research team, led by Dr Namshik Han, use bespoke AI and machine learning methods to identify new signatures of disease and therapeutic targets, as well as network analysis to gain a deep understanding of the underlying causes of disease. Integration of data sets to provide systematic multi-omics analysis, with biological insight is the key focus in order to answer clinically relevant, scientific questions.

We are seeking applications for new collaborative projects from researchers and clinicians with multimodal data sets, who wish to address questions of clinical importance through the integration and interrogation of these data sets.

Interested academics are invited to submit an initial non-confidential expression of interest form (which can be downloaded [here](https://www.milner.cam.ac.uk/onco-innovation-programme/)) by **8am Monday 10th May 2021**, by e-mail to [contact@milner.cam.ac.uk](mailto:contact@milner.cam.ac.uk). Please contact [r.harris@milner.cam.ac.uk](mailto:%20r.harris@milner.cam.ac.uk) with informal queries ahead of the deadline.

**What we are looking for:**

Proposals should:

* Encompass multiple different types of data, with 2 different types(e.g. transcriptomics, genomics, epigenomics, proteomics) a minimum; strategic uplift will be given to proposals with >2 different types of data sets.
* Be based on data sets already available, and which have not previously been analysed for machine learning/network analysis; pre-processed data sets would be an advantage.
* Address clinical or therapeutically relevant questions and have a clear line of sight to clinical impact.
* Consider extended collaborative opportunities, both with computational team and experimental target validation teams within the Milner Therapeutics Institute.

**What we can provide:**

* We will provide 0.3 FTE resource of a postdoctoral researcher for approx 6 months to perform integrated analysis of multi-omics data sets, to answer a specific biological question(s) proposed by the researchers.
* Exact timelines and deliverables will be dependent on types of data, samples numbers etc
* Possible project types include but are not limited to:

1. Network Analysis Project

* Aim: To identify key proteins/pathways associated with a known protein of interest
* Data Required: Clinical, pre-processed multi-omics data sets
* Output: Network analysis will provide a list of statistically significant proteins (approx 200), scored and ranked using mathematical modelling based on omics data (such as expression levels, mutation frequency etc) to take into account biological context

Similar analysis can also be performed in an unsupervised way if there is no known key protein of interest.

1. Virtual Drug Simulation Project

* Aim: To identify small molecules that may target a known protein target and/or pathway through virtual drug screening methodology
* Data Required: Clinical, pre-processed multi-omics data sets, chemical compound database/list
* Output: A prioritised list of potential small molecules for targeting known target/pathway

*NB/ Output scale dependent on the small molecule databases being screened (publicly available database such as ChEMBL vs proprietary dedicated list/database)*

1. Machine Learning Project
   * Aim: To use machine learning approaches to predict novel proteins within a specific protein family based on molecular features of that protein family
   * Data Required: Clinical, pre-processed multi-omics data sets, a protein family name or a list of proteins of interest

*NB/ dependent on the size of data sets, publicly available large-scale data sets will be integrated to the data sets*

* + Output: A prioritised list of putative protein family members

Resource and funding is available for up to 2 projects initially.

*\*Examples of further case studies can be found at* [*https://www.milner.cam.ac.uk/machinelearning/*](https://www.milner.cam.ac.uk/machinelearning/)

**Important to note:**

Ownership of data and results (and associated IP) generated as part of this call will remain with the applicant; the MTI will own any computational methods developed as part of the project.

We hope to build on this call to develop extended collaborations for the future, both with the computational research team, but also with the experimental target discovery and validation team within the MTI.

Our expectation is co-authorship on publications where the data analysis is included, or which arise as a result of the collaboration.

**What happens next:**

We will be hosting a live “Introduction to the call and Q&A” session on **Monday 12th April** **12-1pm** to answer any queries you might have, with a second drop-in Q&A session on **Monday** **26th April 1-2pm**. Alternatively any informal enquiries can be sent to r.harris@milner.cam.ac.uk.

* Applications must be submitted by **8am Monday 10th May**. Interested academics are invited to submit an initial non-confidential expression of interest form (which can be downloaded [here](https://www.milner.cam.ac.uk/onco-innovation-programme/)) by e-mail to [contact@milner.cam.ac.uk](mailto:contact@milner.cam.ac.uk).

Initial proposals will be triaged by a review panel including senior members of the Milner Therapeutics Team, followed by additional review and sign off of shortlisted applications by Onco-Innovation programme lead and Milner Director Prof Tony Kouzarides.

* Those selected will be informed by **Wednesday 19th May** and will then work up the full research plan in discussion with scientists from the Milner Computational ResearchTeam. We anticipate that projects will be agreed and commence week beginning **Monday** **7th June**.

**Timeline Summary**

**Onco-Innovation Rapid Response Data Analysis Call: Template application form**

Please complete this non-confidential template form, using no more than two sides of A4 and font size 11. Please e-mail the completed form to [contact@milner.cam.ac.uk](mailto:contact@milner.cam.ac.uk) no later than **8am Monday 12th April.**

|  |  |
| --- | --- |
| Lead Researcher Name |  |
| CRUK Cambridge Centre Programme |  |
| Department and/or Institute |  |
| e-mail address |  |
| Co-applicants and Departments/Institutes: |  |

|  |  |
| --- | --- |
| Proposal title |  |
| Cancer type (2-3 keywords) |  |

**Background** (<300 words, including scientific and clinical research completed and progress leading to this proposal**):**

|  |
| --- |
|  |

**Data Sets** (Please provide as much information as possible on the data sets available for this project. Information provided should include clinical data sets vs experimental data sets, sample numbers and type (patient groups, cancer types, stage of disease, availability of tumour vs normal), data types such as transcriptomics, genomics, epigenomics, proteomics and information on data processing*.* Please also provide details of the file formats (e.g. SAM/BAM, GFF/GTF, BED, VCF) for each data type.)

|  |
| --- |
|  |

**Key Objectives** (a high-level summary of the scientific questions you would like to address through proposed data integration and analysis; <300 words):

|  |
| --- |
|  |

**Expected Outcomes** (3-4 bullet points):

|  |
| --- |
|  |

**Key publications** (please include 2-3 publications that are most relevant to this proposal, from your own group or others):

|  |
| --- |
|  |

**Potential clinical impact** (1-2 bullet points):

|  |
| --- |
|  |