CANCER RESEARCH UK | Cambridge Centre
UNIVERSITY OF CAMBRIDGE
NHS Cambridge University Hospitals NHS Foundation Trust

Elucidating the role of Micropeptides in B cell lymphoma

Principal supervisor's name: Dr **Dan Hodson**
Principal supervisor's email address: djh1002@cam.ac.uk

Principal supervisor's CRUK CC theme: Haematological Malignancies Programme

Department for student registration: Haematology
Department or institute where research will take place: Wellcome-MRC Cambridge Stem Cell Institute

Postgraduate scheme:

- **MRes + PhD (1 + 3-year non-clinical applicants)**
- **Part-time MRes + PhD (2 + 3-year clinical applicants)**

MRes project outline:

Micropeptides are small protein molecules generated within cells. They are encoded within the genome by small “open reading frames” (smORFs). A relatively small number of translated micropeptides have already been described in the literature but until recently, technical limitations have precluded the experimental identification of these micropeptides or their smORFs. New technology, including the so-called ribosome foot printing (Ribo-Seq) is beginning to reveal the extent of translated smORFs within the genome and biological importance of the micropeptides they encode.

Whilst probing the basis of oncogene dysregulation in lymphoma, our lab has generated a huge RiboSeq collection (>80 libraries) which has provided an unprecedented map of RNA translation in B cell lymphoma. The most striking finding has been the extent of non-canonical translation – much of this from supposedly “non-coding” RNA. Many of these regions appear to encode short (<100amino acids) Micropeptides. The function of this exciting new category of small proteins remains almost completely unknown. The number of apparently translated micropeptides is huge (>10,000). We hypothesise that a small number of these micropeptides play an important role in the biology of cancer and may even be specific to the biology of B cell lymphomas.

The aim of this project is to narrow the large list of potential micropeptides to a focused list of high-confidence, B cell-expressed micropeptides that show evidence of biological activity in B cell lymphoma.
This data will form the start point for a subsequent PhD project.

During the first half of the MRes project the student will design and clone a customised CRISPR screening library (18,000 guides).
The student will then deploy the library in B cell lymphoma lines to identify those micropeptides that play a functional role in B cell lymphoma

MRes experimental plan:

The student will design and build a customised guide RNA library. This will include internal controls that will allow us to distinguish the essentiality of the non-coding RNA itself from the

essentiality of the translated micropeptide. Approximately 3,000 micropeptides will be targeted with 6 guides per micropeptide (total 18,000 gRNAs). Synthetic oligos will then be cloned as a pooled library into a lentiviral expression construct.

The student will then use the cloned library to perform CRISPR screens in two B cell lymphoma cell lines.

The student will develop expertise int he following techniques:
1. Use of bioinformatic tools to design a large, customised CRISPR guide RNA library.
2. Advanced molecular biology techniques including CRISPR library cloning.
3. Mammalian cell culture
4. Lentiviral transduction
5. CRISPR screening and sequencing library preparation.

PhD project outline:

Diffuse large B-cell lymphoma (DLBCL) is the commonest of aggressive B cell lymphomas. Whilst curable with immunochemotherapy, up to a third of patients still die from their disease. Recent studies have revealed the complex biological pathogenesis of DLBCL and reinforce how only by resolving this underlying biological complexity will we be able to make progress with new and better treatments.

Recent advances in DLBCL biology have sought to identify genetic alterations that drive lymphoma. These studies have almost exclusively focused on mutations that affect the known protein-coding genome. The human genome is thought to contain approximately 20,000 protein-coding sequences. These so-called Open Reading Frames (ORFs) have been annotated based on rules that predict synthesis of stable and functional proteins. One of these rules is a rather arbitrary size threshold of 100 amino acids. The advent of new sequencing technologies such as ribosome foot printing (RiboSeq) now challenges these prior assumptions. The ability of RiboSeq to map the position of every translating ribosome in a cell has revealed the unexpected translation of large numbers of previously unrecognised small ORFs. Many of these appear to be translated from so called “untranslated” regions or from “non-coding” RNAs. This non-canonical translation is beginning to be elucidated in other cellular systems. These studies have identified small ORFs that encode novel bona fide micropeptides with potent biological functions 3-6. However, the vast majority of small ORFs and the potential micropeptides they encode remain entirely uncharacterised. This suggests there may be an entire tier of bioactive molecules within the cell that has been as yet largely unexplored.

The PhD project will identify the mechanistic role of one or more micropeptides in the driving the pathogenesis of human B cell lymphoma.

PhD experimental plan:

The function of micropeptides in B cell lymphoma or even cancer generally remains almost completely unexplored. Using the top hits from the CRISPR screen performed during the MRes project, the student will use RNA-sequencing to identify transcriptional signatures associated with knockout or over expression of individual micropeptides.

Larger numbers of micropeptides (50-100) may initially be screened by single cell perturbational transcriptomic approaches such as CROP-Seq. These experiments may be performed in lymphoma cell lines and in primary human B cells and will provide clues at to the biological processed associated with these micropeptides.

CANCER RESEARCH UK | Cambridge Centre
UNIVERSITY OF CAMBRIDGE
NHS Cambridge University Hospitals NHS Foundation Trust

The student will investigate the physical location of the most interesting micropeptides within the cell. Since no antibodies exist for these previously unknown peptides we will first need to use CRISPR editing to engineer small epitope tags or fluorescent tags into the endogenous loci of the smORFs. This will allow us to use standard epitope antibodies for confocal imaging.

These epitope-tagged lines will also allow us to conduct immunoprecipitation and mass spectrometry experiments to identify the proteins that interact with the micropeptide of interest.

In parallel, the student will use existing whole genome sequencing and transcriptome data from patients with DLBCL and other cancer types to identify ways in which the expression of these micropeptides is altered or mutated in the context malignancy.

It is expected that these experiments will be completed during the first 18 months. By identifying the physical intracellular location, the associated proteins and the transcriptional signatures associated with each of a small number of micropeptides we will be able to generate testable hypotheses about their biological function. This will be the focus of the remaining time. An exciting aspect of this project is that we do not know precisely what direction these experiments will take us into. Therefore, this project will be most suited to a student who is eager to explore uncharted areas of biology and keen to bring broad scientific curiosity and independent thinking to a cutting edge project.

We anticipate that this PhD project will reveal completely new aspects of biology that will in turn reveal new potential strategies for the targeted therapy of cancer.

Main techniques:

CRISPR Screening
CRISPR editing of mammalian cell lines
Culture and genetic manipulation of primary human B cells Next generation sequencing
Proteomics Imaging Bioinformatic analysis

Key references:

1. Chen J, Brunner AD, Cogan JZ, et al. Pervasive functional translation of noncanonical human open reading frames. Science. 2020;367(6482):1140-1146.
2. Jackson R, Kroehling L, Khitun A, et al. The translation of non-canonical open reading frames controls mucosal immunity. Nature. 2018;564(7736):434-438.