|  |  |  |
| --- | --- | --- |
| **The Institute of Cancer Research** PHD STUDENTSHIP PROJECT PROPOSAL: | | |
| **PROJECT DETAILS** | | |
| **Project Title:** | Systems biology of the motif-mediated cell cycle interactome | |
| **Short Project Title:** |  | |
| (If main project title is >120 characters including spaces) | Systems biology of the motif-mediated cell cycle interactome | |
| **SUPERVISORY TEAM** | | |
| **Primary Supervisor(s):** | | Norman Davey |
| **Associate Supervisor(s):** | | Jyoti Choudhary |
| **Backup Supervisor:**  (must have IRS status) | | Jon Pines |
| **IRS Partner :**  (only required where Primary is a CDF, Associate Honorary Faculty or an ICR Fellow – ***see above***) | |  |
| **Lead contact person for the project:** | | Norman Davey |
| **DIVISIONAL AFFILIATION** | | |
| **Primary Division:** | | Cancer Biology |
| **Primary Team:** | | Short linear Motifs |
| **PROJECT PROPOSAL** | | |
| **BACKGROUND TO THE PROJECT (up to 300 words)** | | |
| The proposed PhD project is the computational part of a larger CRUK-funded research proposal titled “Discovery, characterisation and therapeutic targeting of motif-mediated interactions regulating cell cycle progression”. Cell cycle proteins contain extensive evolutionarily constrained unstructured regions. However, despite representing up to half of the residues in the core cell cycle proteins, the vast majority of these regions have not been functionally characterised (Van Roey et al. 2014; Tompa et al. 2014). The PhD project will analyse the elusive SLiM-mediated interactome of the key cell cycle regulators in order to understand the functional role of IDRs in cell cycle progression. The resulting census of functional modules in the IDRs of cell cycle proteins will drive future fundamental discoveries in the field of cell cycle biology.  The CRUK proposal will apply complementary protein-protein interaction identification methods specifically designed to characterise SLiM-mediated interactions to detect undiscovered cell cycle functional modules, the interactions they mediate, their conditional regulation and their functional role. The proeject output will be an amino acid resolution map of the interaction interfaces within the IDRs of cell cycle proteins, an understanding of the cancer-related perturbation of these interfaces and an analysis of the cancer therapeutic relevance of inhibiting these interactions.  The PhD position will support the project by developing computational approaches to **(i)** pinpoint protein regions for experimental investigation and **(ii)** process the large volumes of diverse yet complementary data produced by the project. | | |
| **PROJECT AIMS** | | |
| * Develop and apply evolutionary footprinting methods to define a set of peptides for experimental characterisation of SLiM-mediated PPIs in the cell cycle. * Develop methods for the analysis of SLiM-mediated interactions in AP-MS data. * Develop methods for data integration of SLiM-mediated PPI data from complementary PPI methods and datasets * Define a high confidence network of SLiM-mediated interactions of cell cycle proteins * Predict the conditionality of PPIs within the SLiM-mediated interactions of cell cycle proteins | | |
| **RESEARCH PROPOSAL (max. 1000 words)** Please provide information on the approaches to be used and the expected outcomes**.** | | |
| The CRUK project will apply a high-throughput integrative *omics* approach todiscover and characterise the functional elements in the IDRs of cell cycle regulators on the molecular level. The proposal has several work packages. The PhD position will support *WP1, WP2* and *WP3* by performing tasks described in the bullet points in the previous section and in detail below.  **Project 1: Develop and apply evolutionary footprinting methods for SLiM discovery**  The PhD student will develop and apply approaches to identify candidate uncharacterised SLiMs in cell cycle regulators. In previous work, we developed SLiMPrints, an in silico evolutionary footprinting method, to discover a novel cell cycle motif class, the ABBA APC/C Cdc20–binding motif family (Di Fiore et al. 2015; Davey et al. 2012; Davey et al. 2015). The PhD student will update, optimise and benchmark the SLiMPrints tool and apply the updated method to cell cycle proteins to create a set of candidate motifs for testing by the experimentalists on the project.  The development of a novel version of SLiMPrints will include:   * evaluation of the metrics for residue conservation as discriminatory attributes for motif discovery * development of a statistical framework to quantify the likelihood of residue conservation and to assess the significance of groupings of conserved residues * development of software and web server for residue conservation metrics and the novel SLiMPrints tool   The output of this section of the project will be **(i)** a framework for the evolutionary analysis of the intrinsically disordered regions of the human proteome and **(ii) a set of tools to discovery novel functional motifs using their evolutionary fingerprint.**  **Project 2: Integration, annotation and prioritisation of SLiM-mediated PPI data**  The PhD student will develop and apply approaches to integrate curated PPI data and data collected by the proteomic screens performed by the group to define the motif-mediated cell cycle interactome. Several complementary approaches optimised for interaction interfaces in IDRs will be applied in the CRUK project. Each of the PPI methods has major advantages and disadvantages. The disadvantages are often non-overlapping allowing a framework to be built that will allow the medium confidence data of each method alone to be integrated into a high confidence data set. The PhD Student will benchmark the methods to quantify the level of confidence in each motif discovery method. They will integrate the data produced by the project with publicly available interactomics data and discriminatory information for motif functionality using machine learning methods to produce a list of high confidence cell cycle motifs [(Davey et al. 2017; Krystkowiak & Davey 2017; Hertz et al. 2016)](http://f1000.com/work/citation?ids=3198913,4167255,1747159&pre=&pre=&pre=&suf=&suf=&suf=&sa=0,0,0).    The development of tools for the integration of SLiM-mediated PPI data will include:   * assisting in the benchmarking and optimisation of the interaction screens produced during the project * integrating motif specificity determinant data derived from the curated data, structural information and phage display screens to identify SLiM instances in cross-linking AP-MS data * develop methods for the analysis and integration of SLiM-mediated PPI data from complementary PPI datasets * integrating information from the Cell Cycle protein interaction screens with data related to the major discriminators of motif functionality to build high confidence sets of human cell cycle motifs * applying machine learning approaches to train classifiers based on the experimentally validated cell cycle motif instances and apply these approaches to the interaction data to define high confidence cell cycle motifs   The output of this section of the project will be **(i) a framework for the integration of large complex biological datasets and (ii)** an amino acid resolution map of the motif-mediated interactions modulating the cell cycle of the human cell.  **Collaborative Projects:**  The PhD student will also contribute to two projects that are collaborations between the multiple members of the group  **Collaborative Project 1: Analysis of the conditionality of SLiM-mediated PPIs**  The PhD student will contribute to the multi-person grant deliverable to identify cell type-, cell state- and disease-dependent regulatory mechanisms controlling high confidence cell cycle motifs. These analyses will be based on previous work by the group collecting and categorising the commonly utilised design principles for the conditional regulation of motif-mediated protein-protein interactions and the tools developed and applied for motif switch prediction (switches.ELM tool -<http://switches.elm.eu.org/>) [(Van Roey et al. 2013; Van Roey et al. 2012)](http://f1000.com/work/citation?ids=2024668,1037600&pre=&pre=&suf=&suf=&sa=0,0).  The development of tools for the analysis of the conditionality of SLiM-mediated PPIs will include:   * integrating the available interactomic, proteomic and genomic data with the available motif information * predicting the spatiotemporal and cell state conditionality of cell cycle motif-mediated interactions that allow integrative decision-making in cell regulation. * probe the PPI conditionality map to predict the propagation of therapeutically relevant perturbations resulting from physiological or drug-induced changes to protein PTM state, protein expression abnormalities, and cancer-causing mutations on a system level * creating a condition-specific motif-mediated interactome of proteins to produce testable predictions regarding the gain and loss of interactions in response to cell state perturbations   The output of this section of the project will be: acondition-specific motif-mediated interactome describing physiological and disease-state dependent conditionality of cell cycle interactions.    **Collaborative Project 2: Data collection of cell cycle SLiMs**  The vast majority of data describing cell cycle SLiMs, the regulatory mechanisms conditionally controlling their function, and their dysregulation in cancer, is isolated in the text of articles and in poorly formatted supplementary tables [(Gouw et al. 2018)](http://f1000.com/work/citation?ids=4499730&pre=&suf=&sa=0). Consequently, this data created at great expense is significantly underutilised. All team members of the CRUK project, including the PhD student, will curate the cell cycle motif publications over the course of the project. This data will be made available to the cell cycle community as a web-based resource, the Cell Cycle Motif Repository. | | |
| **LITERATURE REFERENCES** (Please use the Harvard system of referencing and provide up to 10 key references) | | |
| Davey, N.E. et al., 2017. Discovery of short linear motif-mediated interactions through phage display of intrinsically disordered regions of the human proteome. The FEBS Journal, 284(3), pp.485–498.  Davey, N.E. et al., 2012. SLiMPrints: conservation-based discovery of functional motif fingerprints in intrinsically disordered protein regions. Nucleic Acids Research, 40(21), pp.10628–10641.  Davey, N.E., Cyert, M.S. & Moses, A.M., 2015. Short linear motifs - ex nihilo evolution of protein regulation. Cell Communication and Signaling, 13, p.43.  Di Fiore, B. et al., 2015. The ABBA motif binds APC/C activators and is shared by APC/C substrates and regulators. Developmental Cell, 32(3), pp.358–372.  Gouw, M. et al., 2018. The eukaryotic linear motif resource - 2018 update. Nucleic Acids Research, 46(D1), pp.D428–D434.  Hertz, E.P.T. et al., 2016. A Conserved Motif Provides Binding Specificity to the PP2A-B56 Phosphatase. Molecular Cell, 63(4), pp.686–695.  Krystkowiak, I. & Davey, N.E., 2017. SLiMSearch: a framework for proteome-wide discovery and annotation of functional modules in intrinsically disordered regions. Nucleic Acids Research, 45(W1), pp.W464–W469.  Tompa, P. et al., 2014. A million peptide motifs for the molecular biologist. Molecular Cell, 55(2), pp.161–169.  Van Roey, K. et al., 2014. Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. Chemical Reviews, 114(13), pp.6733–6778.  Van Roey, K. et al., 2013. The switches.ELM resource: a compendium of conditional regulatory interaction interfaces. Science Signaling, 6(269), p.rs7.  Van Roey, K., Gibson, T.J. & Davey, N.E., 2012. Motif switches: decision-making in cell regulation. Current Opinion in Structural Biology, 22(3), pp.378–385. | | |
| **CANDIDATE PROFILE** | | |
| Note: the ICR’s standard minimum entry requirement is a relevant undergraduate Honours degree (First or 2:1) | | |
| **Pre-requisite qualifications of applicants:**  e.g. BSc or equivalent in specific subject area(s) | | - BSc or equivalent in computer science or similar |
| **Intended learning outcomes:**  Please provide a bullet point list (maximum of seven) of the knowledge and skills you expect the student to have attained on completion of the project. | | * Evolutionary biology of proteins * Analysis of proteomics data * Biology of short linear motifs and the cell cycle * Best practices for biological software development * Full stack development for biological webservers * Big data analysis skills |
| **ADVERTISING DETAILS** | | |
| **Project suitable for a student with a background in:**  (Please tick all categories that apply – your project will be advertised under all selected categories) | | Biological Sciences  Physics or Engineering  Chemistry  Maths, Statistics or Epidemiology  Computer Science  Other (provide details) |
| **Keywords:**  Please provide 4-6 words/short phrases that potential students may type into search engines (e.g. Google) to search for PhDs similar to yours – e.g. ‘cancer predisposition genes’, ‘physics PhD London’ etc. | | **1. Short Linear Motifs** |
| **2. Protein Protein Interactions** |
| **3. Cell Cycle** |
| **4. Computational Biology** |
| **5. Evolutionary Biology** |
| **6. Big Data Analysis** |