

| Project Details |  |
|-----------------|--|
| Project Code    | MRC23NMHEX Hannon  |
| Title           | Statistical inference of cell type specific epigenetic profiles for the molecular epidemiology of brain disorders  |
| Research Theme  | Neuroscience and Mental Health   |
| Summary         | Genetic risk factors for brain disorders can have substantial impacts on the regulation of gene activity, but we only have a limited understanding of how this manifests across the different cell types of the brain. This data science project will use innovative methodologies to derive cell-level genomic profiles from heterogeneous brain tissue. It would suit a quantitative student interested in learning about a broad range of genomic data types and technologies.  |
| Description     | <p>The focus of this data science PhD project is understanding which neural cell types are affected by the genetic and epigenetic variation associated with brain disorders. Genome-wide association studies (GWAS) have facilitated the identification of thousands of genetic variants associated with neurodegenerative and neurodevelopmental disorders. However, for the vast majority of genetic associations it is unclear how, and in what cell-type, they exert their effect. A fundamental conclusion is that many genetic risk factors mediate their effects by influencing the regulation of gene expression. As a consequence, there is a need to generate epigenomic profiles to annotate gene regulatory states across the genome in individual cell types. Such profiling across a range of brain cell types is both time-consuming and expensive, prohibiting analysis at scale. Yet, there is an abundance of data from profiling of bulk brain tissue where gene regulatory state is surveyed across the population of cell types. This project will assess the extent to which we can use innovative mathematical methodologies to infer the constituent cell-level epigenomic profiles from these existing data profiled from bulk brain tissue. The proposed project consists of 3 objectives designed to provide the student with experience of a range of bioinformatics tools and genomic data types. Critically, we have matched epigenetic data from prefrontal cortex and three constituent neural cell types from the same individual permitting characterisation of the accuracy of the described computational methods. We have epigenomic data available from a range of experiments (DNA (hydroxy)methylation, ATAC-Seq, ChIP-Seq) and technologies (microarray, Illumina short read, Oxford Nanopore long read sequencing, 10x single cell RNA-Seq and ATAC-Seq), which the student can choose to incorporate into their project in order to tailor their research experience to their own personal objectives. In addition, the student would be expected to take ownership of the specific phenotype(s) we prioritise and identify the datasets for analysis. Specific objectives: 1. Characterise the cell specificity of epigenomic features detected from heterogeneous brain tissue. The student will use reference cell-specific profiles to reconstruct brain level profiles which will be compared to empirical profiles from bulk brain tissue enabling us determine to what extent bulk level profiles are a linear combination of existing cell level profiles. We can then deduce annotations for cell types not yet profiled, and provide a comprehensive map of gene regulatory states in all brain cell types. This will form the basis of GWAS enrichment analyses using methods such as LD score</p> |

|                         |   |
|-------------------------|---|
|                         | <p>regression. 2. Perform a cell-specific association analysis using epigenomic data generated from brain tissue. Taking a methodology developed for deconvoluting whole blood profiles, tensor composition analysis will be applied to epigenetic data generated from heterogeneous brain tissue to obtain genome-wide data profiles for neurons, oligodendrocytes and microglia. After validating the performance of this method, an epigenome-wide association analysis will be performed for a brain disorder of the student's choosing. We have data for schizophrenia, autism and Alzheimer's disease, but there exist data in public repositories on alternative disorders. 3. Perform a cell type specific genetic analysis of epigenomic variation using epigenomic data generated from brain tissue. The student will generate a catalogue of neural cell-type specific epigenetic quantitative trait loci in large sample cohorts by incorporating an interaction between cellular abundance and genotype in the quantitative trait loci analysis framework. These QTLs can then be integrated with GWAS of brain disorders using methods such as Bayesian co-localisation analysis and Summary data-based Mendelian Randomization to prioritise the affected genes.</p> |
| <b>Supervisory Team</b> |   |
| <b>Lead Supervisor</b>  |   |
| Name                    | Dr Eilis Hannon   |
| Affiliation             | Exeter  |
| College/Faculty         | University of Exeter Medical School   |
| Department/School       | Institute of Biomedical and Clinical Sciences   |
| Email Address           | E.J.Hannon@exeter.ac.uk   |
| <b>Co-Supervisor 1</b>  |   |
| Name                    | Dr Nick Owens   |
| Affiliation             | Exeter  |
| College/Faculty         | University of Exeter Medical School   |
| Department/School       | Institute of Biomedical and Clinical Sciences   |
| <b>Co-Supervisor 2</b>  |   |
| Name                    | Dr Antonio Pardinias  |
| Affiliation             | Cardiff   |
| College/Faculty         |   |
| Department/School       |   |
| <b>Co-Supervisor 3</b>  |   |
| Name                    |   |
| Affiliation             |   |
| College/Faculty         |   |
| Department/School       |   |