**University of Leicester**

British Heart Foundation Centre of Research Excellence

PhD studentship

|  |  |
| --- | --- |
| **First Supervisor** | Dr Tom Webb |
| **School/Department** | Cardiovascular Sciences |
| **Email** | [Tw126@leicester.ac.uk](mailto:Tw126@leicester.ac.uk) |

|  |  |
| --- | --- |
| **Second Supervisor** | Prof. Huiyu Zhou |
| **School/Department** | Computing and Mathematical Sciences |
| **Email** | hz143@leicester.ac.uk |

|  |  |
| --- | --- |
| **Additional Supervisor** | Dr David McVey |
|  | Dr Charles Solomon |
|  | Dr Matthew Baxter (Oxford) |

**Section 2 – *Project Information***

|  |  |
| --- | --- |
| **Project Title** | Integrating Disease-Relevant Epigenomics and AI to Identify Functional Genetic Variants in Cardiovascular Disease |
| **Project Summary** | |
| Cardiovascular diseases (CVD) are a leading cause of death and morbidity worldwide1, placing a major burden on patients and healthcare systems. Improving outcomes requires a better understanding of the biological mechanisms and risk factors that drive disease. Genetic variation is a major contributor to CVD risk. While genome-wide association studies have identified specific risk loci2-7, the challenge now lies in identifying causal variants, affected genes and understanding the molecular mechanisms through which these variants influence vascular cell behaviour.  CVD involves multiple cell types (which may vary, depending on the disease) and it is likely that the effect of a causal variant only manifests under specific conditions, such as cell type, stress or environment signal. It is therefore challenging to translate the findings from GWAS into further mechanistic insights into how genetic variation contributes to CVD.  Artificial intelligence (AI) offers powerful new tools to address this challenge. AI models, including deep learning and large language models, are increasingly used in genomics to predict regulatory activity and prioritise candidate variants8, but these models require better data to perform optimally. To reflect the complexity of CVD development and pathology, we need to generate data from disease-relevant cell types under disease-relevant conditions to maximise the accuracy of AI models and uncover likely candidate variants for further experimental validation.  This interdisciplinary project will generate epigenomic data from primary human vascular cells exposed to disease-associated conditions, including pro-inflammatory stimuli and flow. These data will be used to train and fine-tune AI models through transfer learning, enhancing their ability to predict variant function in a vascular context. The top candidate variants will then be prioritised for experimental validation using genome editing and other functional assays.  The overall goal is to develop AI-enhanced frameworks for variant prioritisation that account for the dynamic, context-specific nature of vascular biology. This work has the potential to uncover novel mechanisms of CVD and inform future strategies for personalised risk prediction and therapeutic targeting. | |
| **References** | |
| 1. McAloon, C.J., Boylan, L.M., Hamborg, T. *et al*. 2016. The Changing Face of Cardiovascular Disease 2000-2012: An Analysis of the World Health Organisation Global Health Estimates Data. International Journal of Cardiology 224, 256-264. 2. Schunkert, H., Erdmann, J. and Samani, N.J. 2010. Genetics of Myocardial Infarction: A Progress Report. European Heart Journal 31(8), 918-925. 3. Schunkert, H., König, I.R., Kathiresan, S. *et al*. 2011. Large-Scale Association Analysis Identifies 13 New Susceptibility Loci for Coronary Artery Disease. Nature Genetics 43(4), 333-338. 4. CARDIoGRAMplusC4D Consortium, Deloukas, P., Kanoni, S. *et al*. 2013. Large-Scale Association Analysis Identifies New Risk Loci for Coronary Artery Disease. Nature Genetics 45(1), 25-33. 5. Webb, T.R., Erdmann, J., Stirrups, K.E. *et al*. 2017. Systematic Evaluation of Pleiotropy Identifies 6 Further Loci Associated with Coronary Artery Disease. Journal of the American College of Cardiology 69(7), 823-836. 6. Aragam, K.G., Jiang, T., Goel, A. *et al*. 2022. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. Nature Genetics 54(12), 1803-1815. 7. Tcheandjieu, C., Zhu, X., Hilliard, A.T. *et al*. 2022. Large-Scale Genome-Wide Association Study of Coronary Artery Disease in Genetically Diverse Populations. Nature Medicine 28(8), 1679-1692. 8. Avsec, Z., Agarwal, V., Visentin, D. *et al.* 2021. Effective Gene Expression Prediction from Sequence by Integrating Long-Range Interactions. *Nature Methods* 18(10), 1196-1203. | |
| **Specific entry requirements**  UK 2:1 Bachelor degree or overseas equivalent in maths/computer science or biology/informatics | |