Project Details	
Project Code	MRCPHS25Br Min
Title	Uncovering the role of epigenetics in modifying disease risk.
Research Theme	Population Health Sciences
Summary	Genetic and environmental factors influencing our health outcomes. Independent effects of genetic and environmental factors on diseases are well known but studies where both genes and environment are included are lacking. DNA methylation is an epigenetic mark of gene regulation and can be used to understand how cells respond to environmental factors and how this influence can vary by genotype. This studentship will apply genetic epidemiological and population health data science approaches to identify genotype environmental interactions with DNA methylation and to identify the influence of these interactions on health outcomes.
Description	Genome wide association studies (GWAS) have discovered many genetic associations for health outcomes but pinpointing the exact genes and variants that are functional remains difficult. These GWAS variants likely exert their influence by modulating regulatory mechanisms. Molecular traits such as DNA methylation (DNAm), gene expression and protein levels have increasingly been used to provide insights on the gene regulatory mechanism of these GWAS variants. DNAm is a fundamental epigenetic modification process by which methyl groups are added to DNA molecules. This process is essential for regulating gene expression, maintaining cell identity, and responding to environmental factors. DNAm arrays have enabled measurement of hundreds of thousands DNAm sites across the genome in epidemiological datasets. DNAm variation among individuals can be influenced by various factors, including lifestyle factors, environmental exposures, and genetic factors. The contribution for each of these factors varies for each DNAm site. By combining genotype information across many healthy participants, studies have identified genetic variants associated with DNAm (mQTL: methylation quantitative trait loci). To this end we published a large mQTL study in >27,750 human blood samples from 36 European cohorts, identifying that >45% of tested DNAm sites exhibit a genetic basis. Using statistical approaches, we combined mQTL and GWA data to propose candidate gene regulatory mechanisms for genetic variants that are associated with disease risk(1). Similar to other studies we found that for most disease associated variants there was no detectable effects on DNAm (2). One explanation for this could be that there is an interaction with an environmental factor. For example, most mQTL studies have identified mQTLs in blood in healthy participants. However blood has a variety of cell-types and each cell-type can exhibit distinct DNAm patterns in their cellular environment. Therefore, studying DNAm changes in blood may obscure cell-type specific DNAm differences (3). In addition identification of mQTLs in individual cell-types is expensive and laborious. Computational approaches have shown that mQTLs that are specific to certain cell-types can be inferred from blood. This is achieved by estimating the proportions of the relevant cell-types in blood and then testing for interactions with genotype. This is similar as in gene-environment interaction models. The contribution of these cell-type mQTLs and other context-specific mQTLs to health outcomes is still

	lacking. We now need to generate catalogs of mQTLs that are dependent on environmental factors such as cell-type and study their influences on health outcomes. This PhD project offers an opportunity to address these knowledge gaps using large-scale DNAm and genetic datasets by collaborating with academic centres that participate in the Genetics of DNA Methylation Consortium (GoDMC, http://www.godmc.org.uk/). This studentship will provide cross-disciplinary training in state-of-the-art epigenetic, genetic and causal inference and population health data science analyses. Aims and objectives: The aim is to get a better understanding of context-specific gene regulation and to identify causal influences between context-specific mQTLs and disease. The following are examples of specific research questions that the student may wish to address as part of their PhD. The student and supervisors will use the prep period to tailor and finalize research plans to reflect the student's interests. In the example below we use cell-type as environmental factor (or context), but the student may wish to choose an different environmental factor. 1. Identify genetic factors for cell-type interacting DNAm variation in blood 2. Identify causal associations between cell-type specific mQTLs and health outcomes 3. Validate genetic and causal associations in cell-type specific datasets Methods The student will analyse genetic and DNAm data on cohorts that participate in the Genetics of DNA Methylation Consortium. The student will identify cell-type specific mQTLs by modelling genotype environmental interactions where cell-type can be taken as a proxy for the environment. GoDMC promotes a federated analysis protocol (https://github.com/genetics-of-dna-methylation-consortium/godmc_phase2/wiki) where scripts developed by the student will be shared with contributing cohorts. This means that the PhD student has an excellent platform to develop analysis skills for genetic and DNAm analysis and develop his/her own research questions. To identify whether DNAm changes are causes or consequences of the health outcome, the student will apply Mendelian randomization (MR) analysis (4). Mendelian Randomization is a genetic epidemiological approach that uses genetic variants as proxies to interrogate potential causal links between exposure (eg cell counts) and outcome (disease). The student will validate the cell-type interacting mQTLs in DNAm datasets of purified cell-types. 1. J. L. Min et al., Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. Nat Genet 53, 1311-1321 (2021). 2. B. D. Umans, A. Battle, Y. Gilad, Where Are the Disease-Associated eQTLs? Trends Genet 37, 109-124 (2021). 3. J. Bergstedt et al., The immune factors driving DNA methylation variation in human blood. Nat Commun 13, 5895 (2022).

	4. G. Davey Smith, G. Hemani, Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet 23, R89-98 (2014).
Supervisory Team	
Lead Supervisor	
Name	Dr Josine Min
Affiliation	Bristol
College/Faculty	Faculty of Health Sciences
Department/School	Population Health Sciences/ Bristol Medical School
Email Address	josine.min@bristol.ac.uk
Co-Supervisor 1	
Name	Dr Eilis Hannon
Affiliation	Exeter
College/Faculty	Faculty of Health and Life Sciences
Department/School	Department of Clinical and Biomedical Sciences
Co-Supervisor 2	
Name	Professor Jonathan Mill
Affiliation	Exeter
College/Faculty	Faculty of Health and Life Sciences
Department/School	Department of Clinical and Biomedical Sciences
Co-Supervisor 3	
Name	Dr Genevieve Leyden
Affiliation	Bristol
College/Faculty	Faculty of Health Sciences
Department/School	Population Health Sciences/ Bristol Medical School