ABSTRACT BOOKLET

MEG-UK Annual Meeting

Bristol 2019

Prediction & Causal Pathways Towards Disease – What Can Molecular Markers Tell Us?

#MEGUK2019

Contents

Welcome!

Welcome to the 2019 Molecular Epidemiology Group (MEGUK) meeting.

MEGUK was founded in 1996 as a special interest group within the United Kingdom Environmental Mutagen Society. Still going strong today, our aim is to encourage multi-disciplinary links between epidemiologists, molecular biologists, biochemists, geneticists, toxicologists, pathologists, nutritionists, clinician scientists, public health scientists and others studying the role of environmental and genetic factors in the aetiology of chronic disease.

We have a great programme this year including several talks and posters on the use of molecular markers in disease risk prediction and the monitoring of environmental exposures. We hope you enjoy it!

If you are interested in joining our group or want to find out more about us, please have a look at our website: meg-uk.org

If you would like to tweet about the meeting, please use our hashtag #MEGUK2019

Our Committee

Jill McKay

Chair

Andrew Collins
Treasurer

Jessica Timms
Secretary

Sabine Langie

Hannah Elliott
Bristol 2019
local organiser

Jonine Figueroa

Marc Gunter

For more details see meg-uk.org/our-committee

Programme

10.00-10.50	**Registration, poster mounting and viewing**	
10.50-12.00	**Session One**	*chair: Jill McKay*
	10.50	**Opening Remarks**
	11.00	**Invited Speaker:** *Metabolic profiling and cancer development* Marc Gunter, IARC
	11.30	**Selected abstract:** *Do circulating metabolites mediate the impact of obesity on colorectal cancer?* Emma Vincent, University of Bristol
	11.45	**Selected abstract:** *Molecular markers of postsurgical recurrence of cervical cancer* Akinyemi Ojesina, University of Alabama at Birmingham
12.00-12.30	**Poster Session**	
12.30-13.15	**Lunch**	
13.15-14.30	**Session Two**	*chair: Jonine Figueroa*
	13.15	**Invited Speaker:** *Breast cancer risk prediction* Doug Easton, University of Cambridge
	13.45	**Selected abstract:** *A systematic evaluation of shared genetic drivers of epigenetic and transcriptomic processes and their influence on health and disease* Tom Richardson, University of Bristol
	14.00	**Selected abstract:** *Lung cancer risk prediction using DNA methylation markers* Florence Guida, IARC
	14.15	**Selected abstract:** *Leveraging DNA methylation signatures in peripheral blood as predictors of impaired respiratory function and chronic obstructive pulmonary disease* Mairead Bermingham, University of Edinburgh
14.30-15.00	**Coffee and Poster viewing**	
15.00-15.45	**Session Three**	*chair: Hannah Elliott*
	15.00	**Invited Speaker:** *Using DNA methylation as a predictor for alcohol exposure* Paul Yousefi, University of Bristol
	15.30	**Selected abstract:** *DNA methylation: biomarkers of current and past smoking* Alexandria Andrayas, University of Essex
15.45-16.15	**Medal Winner Presentation – Hilary Powers, University of Sheffield**	*chair: Andrew Collins*
16.15-16.30	**Prizes and Closing Remarks**	

Useful Information

Wifi … Wifi network: EngineShed

password: FixedDesksHere

Venue … more information about Engine Shed can be found on their website: engine-shed.co.uk or by asking at reception

General information … Our registration desk is open between 10:00 and 11:00. If you need information or assistance outside of this time, please ask one of our committee

Invited Speakers

Marc Gunter
Metabolic profiling and cancer development

Dr. Gunter is the section and group head of the Nutritional Epidemiology Group at the International Agency for Research on Cancer (IARC). We welcome Dr. Gunter to discuss his research on metabolic profiling in relation to the development of cancer.

Douglas Easton
Predicting breast cancer risk

Professor Easton is currently Director of the Centre for Cancer Genetic Epidemiology at the University of Cambridge. Professor Easton's research group focuses on identification and characterisation of genetic variants associated with cancer risk. We welcome Prof. Easton to the meeting to discuss models to predict breast cancer risk.

Paul Yousefi
Using DNA methylation as a predictor for alcohol exposure

Dr. Yousefi is currently a senior research associate at the MRC Integrative Epidemiology Unit at the University of Bristol. Dr. Yousefi is an environmental health researcher with specialisation in environmental epigenetics, molecular epidemiology and biostatistical computing. In his presentation, Dr. Yousefi will demonstrate the utility of DNA methylation as an exposure predictor for alcohol intake.

MEGUK Medal Winner Award

Every 3 years, MEGUK honours a leading researcher with an award in recognition of outstanding achievement in the field of molecular epidemiology. We are delighted to present this year's award to **Professor Hilary Powers**. Professor Powers is currently professor of nutritional biochemistry in the department of oncology at the University of Sheffield.

As well as receiving her award, Professor Powers will give a review of her research career to date under the title **"Understanding vitamin function; a personal odyssey".**

Abstracts

Notes:

Presenters are shown in **bold**.

Names underlined are eligible for early career researcher prizes which will be awarded in the final session of the meeting.

Platform presentations

Do circulating metabolites mediate the impact of obesity on colorectal cancer?

C Bull[1,2], J Bell[1], A Ryk[2], N Timpson[1], M Gunter[3], **EE Vincent**[1,2]

1. MRC Integrative Epidemiology Unit, University of Bristol, UK
2. Cellular and Molecular Medicine, University of Bristol, UK
3. International Agency for Research on Cancer, France

Observational epidemiological studies show that obesity is associated with an increased risk of developing colorectal cancer (CRC). Here we use a genetic epidemiological approach which provides unbiased and unconfounded causal evidence for this association. In addition, combining this approach with cellular metabolic studies we show evidence that circulating metabolites may mediate this causal link.

Obese individuals have abnormal levels of circulating metabolites in their blood. Exposure to these may be linked to cancer development and dictate metabolic phenotype. However, whether obesity, and the circulating metabolites it disrupts, are causally associated with CRC has been unclear.

We have used Mendelian randomization (MR) to determine that obesity is causally associated with CRC. MR uses single nucleotides polymorphisms (SNPs) as instrumental variables to test the causal effect of an exposure (obesity) on an outcome (CRC). Increased adiposity not accompanied by insulin resistance provides an opposite and protective effect.

To discover whether circulating metabolites mediate the effect of obesity on CRC we used MR to determine which dysregulated metabolites are causally associated with CRC. We found that dysregulated levels of several circulating metabolites including fatty acids, amino acids and sugars are causally associated with CRC.

To investigate these causal links, we use a CRC cell line series representing an in vitro model of tumor progression (adenoma to carcinoma sequence of the same lineage). We have shown using MR that elevated levels of circulating valine (characteristic in people with obesity) is associated with CRC. In investigating the role of valine in CRC metabolism we have shown that insulin exposure (mimicking insulin resistance in obese individuals) promotes valine uptake in early adenoma cells (but not adenocarcinoma cells) which are sensitive to the growth promoting effect of insulin.

Through this approach we direct our investigations in the laboratory to causal associations identified in a population.

Molecular Markers of Post-Surgical Recurrence of Cervical Cancer

Dewey Brooke[1,4], Jessica Blair[2], Aishwarya Sundaresan[3], Vinodh Srinivasasainagendra[3], Jianqing Zhang[2,4], Hemant Tiwari[3], **Akinyemi Ojesina**[2,4]

1. Graduate Biomedical Sciences Program, University of Alabama at Birmingham, USA
2. Department of Epidemiology, University of Alabama at Birmingham, USA
3. Department of Biostatistics, University of Alabama at Birmingham, USA
4. O'Neal Comprehensive Cancer Center, University of Alabama at Birmingham, USA

While nearly all cervical tumors are infected with HPV, infection alone is not sufficient for tumor development. Although most cervical HPV infections are cleared by cell-mediated immunity, progression to malignancy is linked to an immunosuppressive tumor microenvironment comprising of a subset of protumorigenic lymphocytes and immunosuppressive stromal fibroblasts. While the tumor microenvironment directs the biology of many cancers, recent evidence has linked dysbiosis of the vaginal microbiome with the extensive reprogramming and remodeling of the cervical stroma.
Using expression-based cell-deconvolution methods on RNAseq from 372 cervical carcinomas, we performed hierarchical clustering on principle components to identify three patient clusters, which were identified as either immune-rich, stromal-rich, or an intermediate immune/stromal type. Patients with immune enriched tumors exhibited a favorable prognosis. However, both the intermediate and stromal enriched tumors had significantly worse overall and disease-free survival ($p = 0.038$ and 0.0021) , with the stromal type caring the worst prognosis. Gene Set Enrichment Analysis found that genes associated with epithelial-mesenchymal transition were more strongly associated with the stromal subtype, while the immune type was strongly associated with genes involved in p53 pathways and networks. Furthermore, we used microbial transcriptomics to identify microbes significantly associated for both patients with or without recurrence.
These results are significant in that they confirm that tumors with higher stromal invasion and marked immunosuppression exhibit the worst prognosis, while identifying for the first time to our knowledge microbes associated with prognosis. Furthermore, we identified possible genetic markers that would aid both in differentiating tumor types and predicting the likelihood of recurrence.

A systematic evaluation of shared genetic drivers of epigenetic and transcriptomic processes and their influence on health and disease

Tom G Richardson[1], Gibran Hemani[1], Josine L Min[1], The GoDMC consortium, George Davey Smith[1], Tom R Gaunt[1], Caroline L Relton[1]

1. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

The vast majority of genetic variants associated with complex traits reside in non-coding regions of the genome. This suggests that these variants are likely to influence regulatory elements, such as epigenetic and transcriptomic processes, along the causal pathway to disease. In this study, we have developed an analytical pipeline to uncover loci throughout the genome where DNA methylation, gene expression and complex traits share a common causal variant. This was undertaken by leveraging findings from the largest consortia to date concerning whole blood derived gene expression (eQTLGen: n=31,684) and DNA methylation (GoDMC: n=27,750).

As an illustration of our approach, we have investigated molecular drivers of breast cancer risk and identified evidence of association at 45 unique loci surviving multiple testing corrections. This suggests that these loci harbour genetic variants which may influence breast cancer susceptibility due to changes DNA methylation, as well as the expression of a nearby gene. Amongst these loci were established breast cancer genes including XBP1, GATAD2A, CHEK2 and L3MBTL4. Epigenetic processes, such as DNA methylation, which may mediate the effects of genetic variants on trait variation are particularly important for diseases such as breast cancer, as they may help improve early detection rates and thus improve successful prevention and treatment plans.

We have subsequently applied our pipeline to over 700 complex traits and developed a web application to visualise and disseminate findings. In this presentation I will demonstrate the value of this resource in terms of prioritising molecular drivers of disease across the human phenome. Furthermore, I will describe how findings can be evaluated in conjunction with results from a companion project to investigate the tissue-dependency of associations (http://mrcieu.mrsoftware.org/Tissue_MR_atlas/). Finally, I will provide examples of how this research can help to highlight adverse side-effects of therapeutic intervention, as well as scope for drug repositioning.

Lung cancer risk prediction using DNA methylation markers

Florence Guida[1], Therese H. Nøst[2], Caroline Relton[3], Paolo Vineis[4], Marc Chadeau-Hyam[4], Gianluca Severi[5], Torkjel M. Sandanger[2], Mattias Johansson[1]

1. Group of Genetic Epidemiology, International Agency for Research on Cancer (IARC), Lyon, France
2. Department of Community Medicine, The Arctic University of Norway (UiT), Tromso, Norway
3. MRC Integrative Epidemiology Unit,University of Bristol, Bristol, UK
4. MRC-PHE Centre for Environment and Health, Imperial College, London, UK
5. UMR 1018, INSERM, Paris, France

There is an urgent need to improve lung cancer (LC) risk assessment as current screening criteria miss a large proportion of cases and result in a high rate of false positives on CT-screening. A meta-analysis of 4 epigenome-wide association studies of LC revealed differential DNA methylation at 16 CpG sites. The current study aimed to evaluate the extent to which such methylation markers can improve upon smoking-based risk-discrimination among ever smokers.
We used data on ever smoking LC case-control pairs from 4 individual prospective cohorts that measured DNA methylation using the Illumina Infinium HumanMethylation450 BeadChip in blood samples before diagnosis: The Italian part of the European Prospective Investigation into Cancer (EPIC), the Melbourne Collaborative Cohort (MCCS), the Norwegian Women and Cancer cohort (NOWAC) and the Northern Sweden Health and Disease Study (NSHDS).
We adopted a training-testing design where the training was performed on MCCS and NSHDS (N=511 pairs), and the testing on EPIC-Italy and NOWAC (N=151 pairs). Logistic regressions with lasso penalties were performed in the training set to select the best set of CpGs jointly predicting LC. A methylation score was trained by fitting a logistic regression model including the 9 lasso-selected CpGs in the training set, and yielded an AUC under the ROC curves of 0.78[0.73-0.84] in the validation set. In comparison, a baseline-smoking score based on self-reported smoking information (duration, cigarettes/day and time since smoking cessation) yielded an AUC of 0.73[0.68-0.79]. The model integrating both scores yielded an AUC of 0.79[0.73-0.84], a notable 0.06-increase in AUC from using the smoking score alone (P=0.008 for difference in AUC).
In conclusion, specific methylation biomarkers have a strong potential to improve LC risk assessment and current USPSTF criteria for CT-screening. During the conference, we will present absolute risk estimates based on the integrated risk prediction model.

Leveraging DNA methylation signatures in peripheral blood as predictors of impaired respiratory function and chronic obstructive pulmonary disease

Mairead L Bermingham[1], Rosie M Walker[1,2], Riccardo E. Marioni[1,2], Stewart M Morris[1], Konrad Rawlik[3], Yanni Zeng[4], Archie Campbell[1,5], Paul Redmond[6], Heather C Whalley[6], Mark J Adams[6], Caroline Hayward[4], Ian J Deary [2,6], David J Porteous[1,2], Andrew M McIntosh[1,2,6], Kathryn L Evans[1,2]

1. Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
2. Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK
3. Division of Genetics and Genomics, The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Roslin, UK
4. Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
5. Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh UK
6. Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK

The causes of poor respiratory function and chronic obstructive pulmonary disease (COPD) are still incompletely understood. It is clear however, that both genes and the environment play an important role. DNA methylation is under both genetic and environmental control. We therefore hypothesised that investigation of differential methylation associated with these phenotypes would improve prediction of, and permit mechanistic insights into COPD.

We investigated genome-wide differential DNA methylation patterns using the recently released 850K Illumina EPIC array. Epigenome-wide association studies (EWAS) of respiratory function and COPD were performed in peripheral blood samples from the Generation Scotland: Scottish Family Health Study cohort (GS:SFHS; N=3,791; 274 COPD cases and 2,928 controls). In independent COPD incidence data (N=150), significantly differentially methylated sites ($p<3.6\times10-8$) were evaluated for their added predictive power to a model including the clinical variables, age, sex, height and smoking history using receiver operating characteristic analysis. The Lothian Birth Cohort 1936 (LBC1936) was used to replicate results from the EWAS (N=895) and prediction analyses (N=178).

To our knowledge, this is the largest single cohort EWAS of respiratory function and COPD to date. We identified 29 respiratory function and/or COPD associated differentially methylated sites. A significant improvement in discrimination between COPD cases and controls ($p<0.05$) in independent GS:SFHS ($p=0.014$) and LBC1936 ($p=0.018$) datasets was observed when differentially methylated sites were incorporated in a clinical model. The differentially methylated sites mapped to genes involved in alternative splicing, JAK-STAT signalling, and axon guidance.

Identification of novel differentially methylated sites has provided novel mechanistic insights and supported previous hypotheses into impaired respiratory function, and improved the prediction of COPD risk. However, future longitudinal studies, with serial measurements of DNA methylation are warranted to evaluate the causal significance of the identified associations, and to assess the utility of DNA methylation profiling in the clinical management of COPD.

DNA methylation: biomarkers of current and past smoking

Alexandria Andrayas[1], Tyler Gorrie-Stone[1], Melissa Smart[2], Eilis Hannon[3], Jonathan Mill[3], Meena Kumari[2], Leonard Schalkwyk[1]

1. University of Essex, Colchester, UK
2. Institute for Social and Economic Research, Colchester, UK
3. Exeter Medical School, Exeter, UK

Despite recent reductions in prevalence, tobacco smoking remains one of the main preventable causes of ill-health and premature death worldwide. The impact of smoking on health can be mediated through a number of biological pathways. Recently, it has become apparent that smoking is associated with extremely strong differences in DNA methylation, to the extent that genome wide studies of diverse health outcomes that correlate with smoking are dominated by the smoking signal. It is therefore necessary to comprehensively examine the association of smoking with DNA methylation to enable researchers to adequately control for smoking in their analyses. We use a sub-sample of the UK Household Longitudinal Study which was previously in the British Household Panel Survey to derive several parameters of smoking behaviour: duration and intensity profiles for over 1000 individuals and analyse their relationship with DNA methylation at more than 850,000 loci. We replicate previously described signals for smoking and supplement these with novel ones as expected with a more comprehensive platform. We observe that the effect sizes associated with duration of smoking are larger than intensity-relevant measures, which parallels the epidemiology of smoking related diseases. Further, we can model and potentially predict the duration of smoking exposure from blood samples and also observe that smoking signals at different loci appear to decay at varying rates. Residual markers of smoking are still present after[4]0 years of not smoking, and we can model and potentially predict the years of no longer smoking in former smokers as well as predicting years spent smoking in current smokers.

Poster Presentations

The EWAS Catalog: a database of epigenome-wide association studies

James Staley[1], **Thomas Battram[1]**, Paul Yousefi[1], Gemma Crawford[1], Claire Prince[1], Mahsa Babaei[1], Gibran Hemani[1], Tom Gaunt[1], Matthew Suderman[1], Caroline Relton[1]

1. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

DNA methylation is the process of adding a methyl group to a DNA molecule, often changing how the molecule interacts with other cellular factors. Methylation mainly occurs at cytosines in humans, often in the context of a cytosine followed by a guanine (CpG). Epigenome-wide association studies (EWAS) seek to understand the link between DNA methylation patterns at thousands or millions of CpG sites across the genome to various traits and exposures. In recent years, the increase in availability of DNA methylation measures in population-based cohorts and case-control studies has resulted in a dramatic increase in the number of EWAS being performed and published. To make this rich source of molecular data more accessible, we have manually curated a database of CpG-trait associations (with p<1x10-4) from published EWAS, each assaying over 100,000 CpGs in at least 100 individuals. The database currently contains over 500,000 CpG associations for more than 150 EWAS. It is accompanied by a web-based tool and R package that allow these associations to be easily queried. In the near future, this database will be extended to include genome-wide EWAS summary statistics, including over 200 million associations from over 500 EWAS of the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort (N~900). This database will give researchers the opportunity to quickly and easily query EWAS associations to gain insight into the molecular underpinnings of disease as well as the impact of traits and exposures on the DNA methylome. The EWAS Catalog is available at: http://www.ewascatalog.org.

DNA methylation signatures of impaired respiratory function predicts depression in chronic obstructive pulmonary disease.

Mairead L Bermingham[1], Rosie M Walker[1,2], Riccardo E. Marioni[1,2], Stewart M Morris[1], Konrad Rawlik[3], Yanni Zeng[4], Archie Campbell[1,5], Paul Redmond[6], Heather C Whalley[6], Mark J Adams[6], Caroline Hayward[4], Ian J Deary[2,6], David J Porteous[1,2], Andrew M McIntosh[1,2,6], Kathryn L Evans[1,2]

1. Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
2. Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK
3. Division of Genetics and Genomics, The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Roslin, UK
4. Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
5. Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh UK
6. Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK

Major depressive disorder (MDD) is a common comorbidity in chronic obstructive pulmonary disease (COPD) but is often undiagnosed. Current diagnostic tools use somatic symptoms, which overlap with symptoms of COPD. The identification of biomarkers is therefore of great importance to aid the diagnosis of MDD in COPD. DNA methylation profiling has allowed for the development of molecular predictors for the early diagnosis of many diseases. We hypothesised that differential methylation could underlie comorbid MDD in COPD. Here, we evaluated the predictive value of differentially methylated sites associated with respiratory function and COPD in the classification of comorbid MDD in COPD.

DNA methylation was profiled using the 850K Illumina EPIC array. We performed epigenome-wide association studies of respiratory function (3,364 participants) and COPD (73 cases; 2,738 controls) in peripheral blood samples from the Generation Scotland: Scottish Family Health Study cohort (GS: SFHS). In independent COPD case data (56 MDD cases; 56 controls), significantly differentially methylated sites ($p<3.6\times10-8$) associated with respiratory function and COPD were evaluated for their added power to predict MDD in COPD to a model including the variables age, sex and smoking history using receiver operating characteristic analysis.

We identified 16 respiratory function and/or COPD associated differentially methylated sites. The final model included 13 the differentially methylated sites in addition to ever smoker and pack years of smoking with an area under the curve (AUC) of 0.62 (95% CI 0.439, 0.801), an increase compared to an AUC of 0.59 (95% CI 0.403, 0.767) for a model with age, sex and smoking history alone.

This model may be of value in predicting comorbid MDD.

Hypothesis-free analysis of deep vein thrombosis aetiology: a Mendelian randomization study

Andrei-Emil Constantinescu[1], Caroline J Bull[1,2], Jie Zheng[2], Benjamin Elsworth[2], Ingeborg Hers[3], Nicholas J Timpson[2], Samantha F Moore[3], Emma E Vincent[1,2]

1. School of Cellular and Molecular Medicine, University of Bristol, Bristol, United Kingdom
2. Medical Research Council Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom
3. School of Physiology, Pharmacology and Neuroscience, University of Bristol, Bristol, United Kingdom

Background:
Deep vein thrombosis (DVT) is the formation of a thrombus/clot in the deep veins; when part of this clot breaks off it can travel to the lungs, resulting in pulmonary embolism. These two conditions together are known as venous thromboembolism (VTE), a leading cause of death and disability worldwide. Despite the prevalence of VTE, we do not fully understand what causes it and it is often overlooked as a major public health problem. Confirming and identifying risk factors associated with DVT is likely to lead to a reduction in the incidence, morbidity and mortality of VTE, especially where these risk factors are modifiable. We can do this, by exploiting the availability of summary genetic data from genome-wide association studies (GWAS) of numerous phenotypes, including DVT, which permits hypothesis-free causal inference.

Objectives:
To identify novel risk factors for DVT and to assess the causality of factors previously shown to be associated with DVT.

Methods:
Two-sample Mendelian randomization (MR) was performed using summarised genetic data. Inverse variance weighted (IVW) estimates were calculated and validated by additional methods (MR Egger, simple mode, weighted mode, and weighted median). Bidirectional, horizontal pleiotropy and heterogeneity sensitivity analyses were performed to further evaluate our findings.

Results:
Forty-seven exposures passed an exposure-exposure correlation-adjusted Bonferroni P-value threshold (5.43E-05). These included previously hypothesised risk factors for DVT (e.g. body mass index, varicose veins, height, hyperthyroidism) and novel associations (e.g. prospective memory, basal metabolic rate).

Conclusion:
Our analyses confirmed causal associations of risk factors previously associated with DVT and highlighted several novel risk factors for the disease. Our study demonstrates the utility of using a hypothesis free Mendelian Randomization approach for the identification of novel disease risk factors.

Leukaemic and Bone Marrow Stromal Cells Alter Cytarabine Genotoxicity and Cytotoxicity in Favour of Disease

Liana Gynn[1], Elizabeth Anderson[1], Gareth Robinson[1], Jennifer May[1]

1. University of the West of England, Bristol, UK

Resistance to therapy remains one of the greatest challenges in the management of the blood and bone marrow (BM) disorder, acute myeloid leukaemia (AML). This study aimed to investigate how leukaemic and BM stromal cells contribute to altered chemo-sensitivity; with particular interest in genotoxicity protection of leukaemic cells and sensitisation of the supportive BM.

AML cells (HL-60 and K562) were mono-cultured or co-cultured with a stromal (HS-5) cell line using trans-well inserts, prior to treatment with a physiological dose of cytarabine (25 µM); standard AML induction therapy. Toxicity protection and sensitisation were determined by genotoxicity (micronucleus/alkaline comet) assays, viability (dye exclusion), chemo-sensitivity (bacterial bioluminescent biosensor) and proliferation (CFSE fluorescence).

Leukaemic-stromal co-culture altered both cell types in favour of disease progression. HS-5 significantly reduced micronuclei formation (4.72/3.45%; $p<0.01$) and reduced comet tail intensity in leukaemic cells following cytarabine exposure. Interestingly, HS-5 genotoxicity to cytarabine significantly increased (2.67/4.57%; $p<0.02$) when co-cultured with leukaemic cells. These findings are concordant with cytotoxicity findings, where HS-5 significantly increased viability and reduced chemo-sensitivity of K562 and HL-60 cells. Leukaemic cells also reduced viability and increased chemo-sensitivity of HS-5 cells in co-culture. Data suggest an increase in leukaemic cell proliferation (35.0/6.38%) when in co-culture, however HS-5 stromal cell proliferation remained consistent regardless of co-culture or treatment condition.

The findings of this study demonstrate the extent for which leukaemic cells alter their environment in favour of disease. Future work will investigate the specific cytotoxic and genotoxic mechanisms of these interactions, continuing both in cell lines and patient samples. Parallel work is also focusing on cytokine secretion and the downstream effects on therapy response. Better understanding of genotoxic damage to the BM stroma may direct the modification of current treatments to reduce the incidence of therapy-related malignancies, and improve overall therapy outcomes for patients with AML.

Leveraging brain cortex-derived molecular data to elucidate epigenetic and transcriptomic drivers of neurological function and disease

Charlie Hatcher[1], Caroline Relton[1], Tom Gaunt[1], Tom Richardson[1]

1. MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol

Integrative approaches which harness large-scale molecular datasets can help develop mechanistic insight into findings from genome-wide association studies (GWAS). We have performed extensive analyses to uncover transcriptional and epigenetic processes which may play a role in neurological trait variation. This was undertaken by applying Bayesian multiple-trait colocalization systematically across the genome to identify genetic variants responsible for influencing intermediate molecular phenotypes as well as neurological traits. In this analysis we leveraged high dimensional quantitative trait loci data derived from prefrontal cortex tissue (concerning gene expression, DNA methylation and histone acetylation) and GWAS findings for 5 neurological traits (Neuroticism, Schizophrenia, Educational Attainment, Insomnia and Alzheimer's disease). There was evidence of colocalization for 118 associations suggesting that the same underlying genetic variant influenced both nearby gene expression as well as neurological trait variation. Of these, 73 associations provided evidence that the genetic variant also influenced proximal DNA methylation and/or histone acetylation. These findings support previous evidence at loci where epigenetic mechanisms may putatively mediate effects of genetic variants on traits, such as KLC1 and schizophrenia. We also uncovered evidence implicating novel loci in neurological disease susceptibility, including genes expressed predominantly in brain tissue such as MDGA1, KIRREL3 and SLC12A5. An inverse relationship between DNA methylation and gene expression was observed more than can be accounted for by chance, supporting previous findings implicating DNA methylation as a transcriptional repressor. Our study should prove valuable in helping future studies prioritise candidate genes and epigenetic mechanisms for in-depth functional follow-up analyses.

Epigenetic mediation of prognostic factors and oropharyngeal cancer survival

Ryan Langdon[1], Rebecca Richmond[1], Hannah Elliott[1], Caroline Relton[1]

1. University of Bristol, UK

Oropharyngeal cancer (OPC), a subtype of head and neck cancer (HNC), has shown a significant increase in incidence in the UK, affecting younger populations with greater frequency and possessing a distinct epidemiology to other HNC subtypes. Several lifestyle and dietary factors as well as viral infections have been implicated in altering both incidence and prognosis for OPC, notably smoking, alcohol consumption and human papillomavirus (HPV) infection. We assessed the impact of these epidemiological factors on epigenome-wide DNA methylation patterns to ascertain whether novel exposure or prognostic indicators could be derived for them, and whether DNA methylation plays a causal mediating role in the observed incidence and prognosis associated with them.

In a large prospective HNC cohort (Head and Neck 5000), we undertook epigenome-wide association study (EWAS) and differentially-methylated region (DMR) analyses, using MethylationEPIC data, of the aforementioned epidemiological factors and OPC mortality (~3 years), respectively. We then determined the extent to which any results mediated the pathway between epidemiological factor and OPC mortality using Mendelian randomization (MR).

We identified CpG sites and DMRs associated with smoking and alcohol consumption below our multiple-testing threshold, but none with HPV positivity. We also identified 7 CpGs associated with survival (~3 years post-diagnosis), independent of smoking, alcohol consumption and HPV positivity. 18 CpGs (3 DMRs) were identified in DMR analyses of both smoking and mortality; 5 CpGs (1 DMR) were identified in DMR analyses of both alcohol and mortality. We hypothesised that for these CpG sites there could be a mediation effect of DNA methylation, whereby DNA methylation mediates the association between risk factor and OPC mortality.

MR analysis was conducted to accumulate evidence to support this theory. We found that hypomethylation in CpGs located within a DMR associated with smoking (located at Chr2:220325443-220326041; annotating to the SPEG gene) showed evidence of association with increased mortality (HR: 1.15, 95% CI: 1.05 to 1.25, P: 1.12e-3) using an inverse-variance weighted (IVW) MR approach. DNA methylation at this locus could potentially mediate the association between smoking and OPC mortality.

Within the context of OPC, we found novel epigenetic biomarkers measured by the MethylationEPIC array to be associated with the epidemiological factors of smoking and alcohol, and with mortality, respectively. We also found preliminary evidence of a potential mediation effect of DNA methylation at the SPEG gene, between smoking and OPC mortality. However, longer follow up in Head and Neck 5000 and suitable replication data is needed to strengthen the validity of these findings.

Exploring the role of the Hygiene Hypothesis in the aetiology of Respiratory Allergy and Childhood Acute lymphoblastic leukaemia; Investigating DNA methylation as a common mediating mechanism

Sabine A.S. Langie[1,2], Jessica A. Timms[3], Patrick De Boever[1,2], **Jill A. McKay[3,4]**

1. VITO- Sustainable Health, Boeretang 200, 2400 Mol, Belgium
2. Centre for Environmental Sciences, Hasselt University, Diepenbeek, Belgium
3. Institute of Health & Society, Newcastle University, U.K
4. Department of Applied Sciences, Northumbria University, U.K

The hygiene hypothesis, which implies that lack of exposure to infection in early life suppresses proper immune system development, has been aetiologically linked to respiratory allergy (RA) and childhood acute lymphoblastic leukaemia (ALL). Whilst epidemiological evidence supports these associations, little is known about underlying mechanisms. Given the suggested common aetiology of RA and childhood ALL, it is plausible that there may be common mechanisms toward disease development. DNA methylation is altered in RA and ALL, and is altered in response to environmental stimuli such as infection. We hypothesised that if the hygiene hypothesis is aetiological for RA and childhood ALL then there are likely to be common methylation marks between these diseases, and methylation changes associated with proxy measures for infection should also be observed for each disease. We compared disease-associated DNA methylation signatures of RA and ALL, and related these signatures to exposure-to-infection-signatures (utilising day care attendance and cold symptoms as proxy measures of infection). We found a significant number of genes with altered methylation in both RA and ALL, which is unlikely to be due to chance ($p=0.0019$). Furthermore, biological processes and pathways influenced by these disease-specific methylation signatures overlapped. Whilst day care attendance was associated with variation of methylation in genes associated with ALL, cold symptoms during early life were associated with methylation variation in genes associated with RA. Evidence suggests that DNA methylation may be a common mediating mechanism by which the hygiene hypothesis is associated with RA and childhood ALL risk.

Metabolomic prediction of pregnancy-related disorders

Nancy McBride[1], Paul Yousefi[1], Matthew Suderman[1], Caroline Relton[1], Debbie Lawlor[1]

1. University of Bristol, Bristol UK

Pregnancy disorders such as gestational diabetes (GD), hypertensive disorders of pregnancy (HDP), pregnancy loss and small/large for gestational age (S/LGA) are common and associated with perinatal morbidity and mortality. Earlier, more accurate identification of women at high risk of these disorders will enable better management and distribution of maternal care resources. Levels of antenatal risk are currently assessed using risk factors such as antenatal history, age, smoking, BMI and parity. However, these have poor sensitivity and specificity. We look to improve upon existing predictors using metabolite measures for a wide range of pregnancy disorders.

We used 227 NMR-derived metabolite measures from Pakistani and white British mothers from the Born in Bradford (BiB) longitudinal birth cohort to generate prediction models for pregnancy-related disorders. All analyses were stratified by ethnicity (Pakistani or white British). Penalised regression was used to create predictive models for HDP (Pakistani case/control[4]63/3261; white British 822/ 2711) and GD (Pakistani 398/3326; white British 172/3361) from 90% of the cohort. Out-of-sample performance was conducted in a random subset of 10% of observations for both ethnicities withheld during training using receiver operating characteristic curves.

We found that the metabolites had good discrimination for GD in both ethnic groups, particularly Pakistani women (AUC 0.810 and 0.711 in Pakistani and white British, respectively). Discrimination was poorer for HDP (AUC 0.694 and 0.648). We found the NMR metabolites had better discrimination when compared to a model of existing predictors (BMI, age, smoking and parity) for GD (AUC 0.741 and 0.709) and HDP (AUC 0.6603 and 0.6705).

Next, we will assess whether certain metabolites were driving these associations and evaluate the predictive accuracy of this platform for other pregnancy-related disorders. We aim to replicate these findings in independent cohorts.

Exposure to chemotherapy induces epigenetic damage in normal cells in childhood cancer patients

Jill McKay[1,4], Linda Sharp[1], John Casement[2], Gordon Strathdee[3]

1. Institute of Health & Society, Newcastle University
2. Bioinformatics support unit, Newcastle University
3. Northern Institute for Cancer Research, Newcastle University
4. Department of Applied Sciences, Northumbria University, U.K

Childhood cancer survival rates have improved leading to more survivors. Long term follow up shows childhood cancer survivors (CCS) have increased risk of chronic disease and early death (not related to disease re-occurrence) compared to the population as a whole. Mechanisms underlying high illness and death rates in CCSs are not understood. One mechanism could involve epigenetic damage induced by anti-cancer therapies. Epigenetic changes, such as altered DNA methylation, are an attractive mechanism as these changes, once induced, can be stable and thus could contribute to health problems years after treatment.

To explore this, we assessed the impact of chemotherapy on DNA methylation at six genes in normal cells from ALL patients in early and late remission. Five of six candidates exhibited no increase in methylation between early and late remission. However, HOXA4 exhibited dramatically increased methylation in late remission. Furthermore the extent of the increased methylation was associated with treatment duration, implying increased chemotherapy exposure was associated with increased aberrant methylation.

We expanded these studies, performing genome-wide methylation analysis using Illumina MethylationEpic arrays, in 32 sets of paired peripheral blood DNA samples taken at initial remission (or diagnosis in solid tumours) and late remission (up to 2 years) in childhood cancer patients. This confirmed the highly significant increase in HOXA4 methylation and identified 145 additional differentially methylated regions in normal, non-cancer cells, in later compared with early remission. This demonstrates that chemotherapy exposure induces large, wide-spread DNA methylation changes across the genome in normal healthy cells. Furthermore, induced methylation changes were highly similar in solid tumour (primarily neuroblastoma) and leukaemia patients, suggesting altered DNA methylation occurs largely independently of specific chemotherapeutic regimes. Further studies are required to determine if this altered DNA methylation is sustained into later life and may mediate or predict long term health effects in CCSs.

Incidence trends for molecular subtypes of breast cancer in Scotland from 1997 to 2016

Ines Mesa-Eguiagaray[1], Sarah Wild[1], Jonine Figueroa[1]

1. University of Edinburgh, Edinburgh, UK

Background: Breast cancer is a model disease of aetiologic heterogeneity, where genetic and molecular markers have provided novel insights into risk prediction. Oestrogen receptor (ER) and Herceptin receptor (HER2) expression in breast tumours are key prognostic markers. Yet, reporting whether breast cancer incidence trends differ by hormone status has not been assessed in any UK national cancer registry data.
Methods: We used Scottish cancer registry data on 72,217 women diagnosed with breast cancer (ICD10-C50) from 1997 to 2016. Trends in breast cancer incidence by combinations of ER and HER2 status are reported, corrected for missing values of marker data using observed molecular marker counts by age and year of diagnosis (missingness=7% for ER status from 1997 to 2016 and 11% for HER2 from 2009 to 2016). We calculated age-specific and age-standardised incidence rates and, using joinpoint regression models, we described trends and estimated annual percentage changes by tumour molecular subtype.
Results: Incidence of ER+ tumours increased by 0.4% per year (95% CI: -0.1 to 1.0%/year) from 1997 to 2016, whereas ER- tumours decreased by 2.5% (95% CI: -3.9 to -1.1%/year). From 1997 to 2012, ER+ tumours showed the greatest rate of increase at 1.2% per year (95% CI: 0.8 to 1.5%/year), followed by a 2.2% (95% CI: -4.7 to 0.4%/year) decrease in incidence observed from 2012 to 2016. The incidence of HER2- tumours showed a similar pattern to the trend for ER+ tumours. Trends for the four combinations of ER/HER2 status showed ER-/HER2+ tumours to have the lowest incidence rate.
Conclusions: These data from Scotland, the first in the UK, show rising breast cancer incidence is limited to ER+ tumours. The decrease in the incidence of ER+ tumours after 2011 requires further investigation of underlying aetiologic and demographic factors such as obesity, reproductive patterns and screening.

DNA methylation biomarkers of early life exposures and subsequent obesity: Findings from the Newcastle Thousand Families Study and the Avon Longitudinal Study of Parents and Children

N. Robinson[1], H. Brown[1], M.S. Pearce[1], V.J. Albani[1], HM Byun[1], J.A. McKay[2]

1. Newcastle University, Newcastle-upon-Tyne, UK
2. Northumbria University, Newcastle-upon-Tyne, UK

There is increasing evidence that associations between early life factors and obesity in later life may be mediated through epigenetic mechanisms. We investigated if early life risk factors associated with subsequent obesity (including rapid weight gain, antibiotic exposure, adversity, socio-economic status, advanced maternal age), are also associated with variation in DNA methylation in childhood and adulthood.

Using Illumina[4]50K array data from the Avon Longitudinal Study of Parents and Children (ALSPAC), we conducted an epigenome-wide association study examining early life exposures and blood methylation (in childhood and late adolescence) at individual CpG loci. A significant change in methylation was observed for rapid weight gain (RWG, 0-12 months) only. RWG was associated with a 1% increase in methylation at an individual CpG loci (CG11531579) in childhood (age 7, n=116) in ALSPAC (Bonferroni corrected). Furthermore, the highest levels of methylation (+2%) were seen in those with RWG who were subsequently overweight/obese (age 17).

The CG11531579 loci was investigated further in an older population to examine whether the associated variation in blood methylation persisted into adulthood using the Newcastle Thousand Families study (age 50, n=134). Combined bisulphite modification and pyrosequencing was used to assess DNA methylation.

RWG was also associated with methylation changes in an adult population, although in adults this was a decrease in methylation (-2%, age 50) for those with RWG who were subsequently overweight/obese (age 60, n=132).

This study identified that RWG in infancy is associated with small variations in methylation. The loci was positively associated with blood methylation in childhood but negatively in adulthood, which could suggest it is an irregular, dynamic, RWG-related loci.

Triangulating -omics evidence to fine-map causal loci in atopic dermatitis – eczema

Maria K. Sobczyk[1], Tom R. Gaunt[1], Lavinia Paternoster[1]

1. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

Atopic dermatitis (AD), commonly known as eczema, is one of the most prevalent chronic skin diseases, affecting up to 30% of children and 10% of adults. AD aetiology has a strong genetic component, with heritability estimates at between 71-84% which predisposes it well to GWAS analysis using common variants. In a previous study of predominantly European-descent populations, we identified 31 risk loci, which we now follow up to identify the causal variants and the genes whose function they influence. We develop a pipeline integrating credible set interval SNPs from 3 statistical fine-mapping methods (JAM, Finemap and fastPAINTOR), regression on full summary statistics from eQTLs, pQTLs, mQTLs in relevant tissues: skin, whole blood and immune cell types with coloc and TWAS methods. We combine that with functional annotation based on regulatory variant prediction, and genomic features such as chromatin accessibility, promoter-enhancer interactions, splicing sites, non-coding RNA regions, and we also mine -omics studies on eczema patients. Amongst others, we prioritize candidate targets for modulating the disease risk, where we can pinpoint the variant and its associated gene: rs4809219 and LIME1, rs10791824 and OVOL1. LIME1, a novel candidate gene, is a transmembrane adaptor protein that relays the T and B-cell receptor stimulation to downstream signalling pathways via its association with the Src family kinases. OVOL1, previously associated with eczema, is known to be a transcriptional repressor binding to c-myc promoter and necessary for growth arrest of embryonic epidermal progenitor cells in mice.

Environment, DNA methylation and risk of childhood acute lymphoblastic leukaemia: a novel Mendelian randomization study

Jessica Timms[1,8], Lynsey Hall[2,3], Caroline Relton[4], Rebecca Richmond[4], Gemma Sharp[4], Gordon Strathdee[5], Heather Cordell[2], Kenen Onel[6], Linda Sharp[1], Jill McKay[1,7]

1. Institute of Health & Society, Newcastle University, U.K
2. Institute of Genetic Medicine, Newcastle University, U.K
3. Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, U.K
4. MRC Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol
5. Northern Institute for Cancer Research, Newcastle University, U.K
6. The Feinstein Institute for Medical Research, Northwell Health, NY, U.S.A
7. Department of Applied Sciences, Northumbria University, U.K., Research Oncology, King's College London, Guy's Hospital, London.

Epidemiological data suggests various environmental exposures may be associated with increased risk of acute lymphoblastic leukaemia (ALL). DNA methylation is a hallmark of many cancers, including ALL, and is also modifiable by the environment. It may therefore be a plausible mediating mechanism for the effect of environmental exposures on ALL risk. Mendelian randomization (MR) is an approach which can be used to infer causal relationships between environmental factors and outcomes. The approach uses genetic variation (e.g. single nucleotide polymorphisms (SNP)) as an instrumental variable (proxy) for an exposure of interest. We therefore used two-sample Mendelian randomization (MR) to determine whether DNA methylation may act as a mediating mechanism between early life environmental exposures and ALL risk. Previously, through epigenome-wide association studies (EWAS), we identified variation of methylation at individual CpG sites associated with ALL-risk exposures (e.g. maternal smoking and folic acid intake). The methylation quantitative loci database (mQTLdb; http://www.mqtldb.org/) was used to select specific cis-SNPs associated with the identified CpG sites of interest. Data from a genome-wide association study (GWAS) (848 childhood ALL cases and 1190 cancer-free control children) were used to investigate whether the cis-SNPs associated with methylation at CpG sites of interest have a causal effect on childhood ALL. Analysis was carried out in R using packages designed at the University of Bristol (MRInstruments and TwoSampleMR). We identified hypermethylation (associated with folic acid intake (p= 0.047)) at cg20495738 (within CACNA1C gene) as having a protective effect on ALL, which was anticipated, but unexpectedly, we also identified a reduced risk with hypermethylation (associated with smoking (p = 0.003, 0.046, 0.045, and 2.955 x 10-4)) of the CYP1A1 gene (across multiple CpG sites). These results show the potential of the two-sample MR approach in inferring causal effects in rare and multifactorial diseases such as ALL.

Impact of short-term traffic-related air pollution on the metabolome – results from two metabolome-wide experimental studies

Karin van Veldhoven[1], Agneta Kiss[2], Pekka Keski-Rahkonen[2], Nivonirina Robinot[2], Augustin Scalbert[2], Paul Cullinan[3,4], Kian Fan Chung[3,4], Peter Collins[3,4], Rudy Sinharay[3,4], Benjamin M. Barratt[5], Mark Nieuwenhuijsen[6], Albert Ambros Rodoreda[6], Glòria Carrasco-Turigas[6], Jelle Vlaanderen[7], Roel Vermeulen[7], Lützen Portengen[7], Soterios A. Kyrtopoulos[8], Erica Ponzi[1,9], Marc Chadeau-Hyam[1,7], Paolo Vineis[1,10]

1. MRC/PHE Centre for Environment and Health, Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom
2. International Agency for Research on Cancer (IARC), Lyon, France
3. National Heart & Lung Institute, Imperial College London, United Kingdom
4. Royal Brompton & Harefield NHS Trust, London, United Kingdom
5. King's College London, United Kingdom
6. Barcelona Institute for Global Health (ISGlobal), Barcelona, Spain
7. Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, the Netherlands
8. National Hellenic Research Foundation, Athens, Greece
9. Epidemiology, Biostatistics and Prevention Institute, University of Zurich, Switzerland
10. Italian Institute for Genomic Medicine (IIGM), Turin, Italy

Exposure to traffic-related air pollution (TRAP) has been associated with adverse health outcomes but underlying biological mechanisms remain poorly understood. Two randomized crossover trials were used here, the Oxford Street II (London) and the TAPAS II (Barcelona) studies, where volunteers were allocated to high or low air pollution exposures. The two locations represent different exposure scenarios, with Oxford Street characterized by diesel vehicles and Barcelona by normal mixed urban traffic. Levels of five and four pollutants were measured, respectively, using personal exposure monitoring devices. Serum samples were used for metabolomic profiling. The association between TRAP and levels of each metabolic feature was assessed. All pollutant levels were significantly higher at the high pollution sites. 29 and 77 metabolic features were associated with at least one pollutant in the Oxford Street II and TAPAS II studies, respectively, which related to 17 and 30 metabolic compounds. Little overlap was observed across pollutants for metabolic features, suggesting that different pollutants may affect levels of different metabolic features. After observing the annotated compounds, the main pathway suggested in Oxford Street II in association with NO2 was the acyl-carnitine pathway, previously found to be associated with cardio-respiratory disease. No overlap was found between the metabolic features identified in the two studies.