National Institute for Health Research

2019 Biomedical Science, Bioinformatics & Data Analysis MRes & PhD Programme

Research and Development Department incorporating the NIHR Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London.

National Institute for Health Research

Dear Student,

Please see the catalogue of projects available below for your rotations and/or final PhD project. Please be aware that projects are subject to change and may not all be available for a final PhD.

National Institute for Health Research

Development of a pancreatic cancer cell migration and invasion high content image analysis pipeline.

Supervisor 1
Name: Dr Matthias Krause
School/Directorate: King's College London, FOLSM, Randall Centre
Email: Matthias.krause@kcl.ac.uk

Supervisor 2
Name: Dr Debashis Sarker
School/Directorate: Cancer CAG, Imaging and biomedical engineering CAG, School of Cancer, Research Oncology
Email: debashis.sarker@kcl.ac.uk

Abstract

Pancreatic cancer is a very devastating disease with a very low survival rate due to the uncontrolled metastatic spread. The first step in metastasis is local invasion followed by intravasation into the blood stream. A prerequisite for these early steps is that pancreatic cancer cells gain the ability to migrate. This is driven by changes in the regulation of the actin cytoskeleton, which provides the force for cell migration.

Surprisingly little is known about the molecular details of how the regulation of the actin cytoskeleton is changed to allow pancreatic cancer cells to become motile and no drugs targeting the metastatic spread of pancreatic cancer cells are available. To better understand the function of individual regulators of the actin cytoskeleton in this process high resolution live imaging of cellular protrusions during migration and invasion is necessary.

The high volume of data ("big data") from this high content imaging requires a detailed automated analysis of a plethora of parameters to transform the data into "deep data". In this proposed project, you will primarily focus on using MATLAB to develop an image analysis pipeline for high content big data analysis of pancreatic cancer cell migration and invasion.

As this is a multidisciplinary project, you will be also trained in molecular biology and tissue culture techniques to use advanced microscopy to acquire the imaging data in collaboration with other members of the team. This will also allow you to better understand the entire imaging pipeline.

We have been studying a regulator of the actin cytoskeleton, Lamellipodin, for several years and recently found that Lamellipodin is required for breast cancer invasion and metastasis. Thus, the imaging pipeline you are developing will initially be used to uncover the function of Lamellipodin in pancreatic cancer cell migration and invasion.

National Institute for Health Research

Investigating novel therapeutic targets for fibrotic skin disease

Supervisor 1
Name: Professor Maddy Parsons
School/Directorate: BMB/FoLSM
Email: maddy.parsons@kcl.ac.uk

Supervisor 2
Name: Professor John McGrath
School/Directorate: BMB/FoLSM
Email: john.mcgrath@kcl.ac.uk

Abstract

The fibrotic response is characterized by increased deposition of extracellular matrix (ECM) proteins, and plays an important part of normal tissue repair but if uncontrolled can lead to scars, contractures, pain, loss-of-function, and considerable morbidity. Skin fibrosis can take a number of different forms, with varying degrees of severity; however, all forms are currently untreatable and in many susceptible patients, fibrosis recurs following surgical removal. Whilst some potential contributing factors have been described, the factors that perpetuate these fibro-proliferative disorders remain unknown. One common form of excessive fibrosis in the skin is keloid scarring, hard or rubbery nodules that follow injury and which extend beyond the sites of the original wound. Another common form of skin fibrosis is known as dermatofibroma, firm lumps in the skin that can follow penetrating skin trauma.

The cellular responses and signalling pathways that give rise to either of these forms of skin fibrosis are poorly characterised and there is an urgent unmet need to develop better targeted therapies for affected individuals. Most examples of keloids or dermatofibromas arise in individuals without a strong family history of similar lesions. However, we have identified several pedigrees in which there is a clear autosomal dominant Mendelian pattern of inheritance. Such families serve as extreme variants in being able to characterise specific genes/proteins and pathways that are linked to skin fibrosis. Using whole exome sequencing in these pedigrees, we found novel segregating mutations in a number of genes that linked to a fibrotic phenotype. However, potential functional links between these genes and the emergence of a pro-fibrotic phenotype remain unknown.

The goal of this project is to understand the molecular interplay between these genes, identify common shared pathways in these and other patient-derived skin fibrosis samples and define whether these molecular signatures promote chronic feedback signalling and sensitisation of cells leading to constitutive fibrotic phenotypes. The project will use computational approaches to define novel pathways that mediate the fibrotic response and test these predictions using experimental models. The specific aims are:

1) Use computational pathway analysis and modelling to determine shared signalling pathways between identified genes.
2) Generate knockout/knock-in fibroblasts for mutant genes using CRISPR/Cas9 for experimental analysis of cell proliferation and ECM synthesis; inducible knockouts will be used in some cases where gene deletion renders cell non-viable. Using 2D and 3D in vitro models, analyse activation status of the predicted pathways.

3) Feed experimental information into predictive models to refine pathway analysis. Target key pathways using small molecule/overexpression approaches to determine whether fibrotic phenotypes can be corrected.
4) Confirm key predictions based on experimental and computational data using mass imaging cytometry of tissue samples from patients with different dermal fibrotic disease to determine whether pro-fibrotic nodes defined from our *in vitro*/*ex-vivo* analyses correlate with clinical disease settings.

Data arising from this study will define novel genes and pathways that contribute to dermal homeostasis and fibrotic signalling and provide potential targets for future therapeutic intervention.

National Institute for Health Research

Developing clinically useful predictive models for lupus Sub-phenotypes

Supervisor 1
Name: Dr David Morris
School/Directorate: Life Sciences and Medicine
Email: david.l.morris@kcl.ac.uk

Supervisor 2
Name: Professor Timothy Vyse
School/Directorate: Life Sciences and Medicine
Email: timothy.vyse@kcl.ac.uk

Abstract

Systemic Lupus Erythematosus (SLE, Lupus) is an autoimmune disease that affects millions of people worldwide. SLE is a heterogeneous disease with patients displaying a wide variety of clinical features. A recent collaboration by the KCL lab with groups in China resulted in a better understanding of why the prevalence of SLE is higher in non-Europeans, in that this can be explained in part by genetics. This result suggests that genetics could be used for prediction. Much progress has been made in the last decade on the genetics of Lupus. Our current understating is that there are likely hundreds of independently acting causal variants across the genome. The disease can be seen then as polygenetic in nature and predictive models of disease will exploit this fact.

This project will develop a clinically useful predictive tool for SLE and SLE sub-phenotypes using statistical models together with utility scores derived from discussion with health professionals. The statistical model for prediction will make use of variants identified as associated that explain heritability. A polygenetic risk score can be generated from the genotype data for a subject. The difficulty then is setting a threshold for the risk function above which a subject is predicted to be at risk of suffering from SLE at some point in their life. This is where we introduce costs and benefits in order to determine the best threshold. For example, we need to define clinical utility in terms of positive utility (e.g. benefits of identifying at risk individuals early in their lives or the benefit of correctly identifying someone as low risk when they otherwise (from family history) may believe they are high risk) and negative (e.g. costs of missing those at risk who later get diagnosed with SLE or costs related to incorrectly predicting a high risk subject that then never develops disease). The predictive model will make use of utility functions that balance the costs of misclassifying the disease against the benefits of correct classification.

This project will be based in the leading laboratory in the genetics of SLE, and have access to the vast majority of the genetic data on SLE worldwide, which presents the best opportunity to explain the heritability of the disease. Our discovery analyses will be applied to one dataset for each major population (European, Asia, African American and Amerindian) and subsequently other independent GWAS datasets will be used for replication. Further GWAS data will be used to validate our predictive model. The breath of data across populations available to the applicant makes this separation of discovery, replication and prediction possible. And through contacts with leading physicians this project will be best placed to create a useful predictive tool based on the genetic data, over and above knowledge of family history. This will reduce costs to society in terms of health provision and maximise benefits in terms of health and disease management.

National Institute for Health Research

Delineating Signalling Pathways of Antibody-Mediated Activation of Immune Cells in Cancer Immunotherapy

Supervisor 1
Name: Dr Sophia N Karagiannis
School/Directorate: St John's Institute of Dermatology, School of Basic & Medical Biosciences
Email: sophia.karagiannis@kcl.ac.uk

Supervisor 2
Name: Dr Sophia Tsoka
School/Directorate: NMS/Informatics
Email: sophia.tsoka@kcl.ac.uk

Abstract
Monoclonal antibodies are established treatment modality in oncology. A key mechanism by which antibodies can confer immune protection is through binding cognate Fc-receptors on human immune cells via antibody Fc-regions (effector functions). Clinically-approved antibodies belong to only one of 5 classes, namely IgG. This means that the effector functions of only this class are employed in cancer immunotherapy. We demonstrated that IgE class antibodies recognising cancer cells can offer new options for cancer immunotherapy based on their ability to engage human monocytes and macrophages, normally found in tumours, to restrict cancer growth. IgE can engender effector functions by engaging its cognate IgE Fc-receptor FcεRI, the highest-affinity Fc-receptor in humans, expressed on immune cells. However, the mechanisms that control IgE anti-tumour properties are not well-understood.

Here we aim to gain an insight into the immunological and signalling pathways that regulate IgE effector functions against cancer. We will study the monocyte- and macrophage-activating properties of an in-house-generated IgE panel, including our first-in-class therapeutic candidate presently undergoing an early-phase trial in cancer patients. These agents can engender effector functions against cancer cells as part of their anti-tumour properties.

We will seek to undertake, for the first time, a multidisciplinary approach to elucidate the biological and molecular mechanisms employed upon engagement of IgE with Fc-receptors on monocytes and macrophages. We will bring together wet-lab experiments that characterise antibody-mediated effector cell activation against cancer cells (Karagiannis), with computational and mathematical approaches (Tsoka) that can point to specific activation signalling cascades. We will interrogate our monoclonal antibody panel recognising different tumour-associated antigens, to seek common pathways associated with anti-tumour functions. Preliminary phosphoprotein and cytokine profiling in the Karagiannis and Tsoka laboratories point to molecular and immunological signatures associated with antibody-mediated tumour-killing properties, strengthened by analyses of genes in the Fc-receptor pathway when transcriptomic data are overlaid on networks and logical modelling is applied. We will extend these early findings by integrating antibody functional data into prior knowledge networks, discovering missing links and generating novel network structures for elucidating signalling pathways in immune-relevant interactions.

MRes Rotation and Yr1: a) quantification of antibody effector functions (e.g. cytotoxicity, phagocytosis) against cancer cells using multi-colour flow cytometry; b) explore the underpinning mechanisms by studying monocyte intracellular phosphorylation, gene expression (transcriptomic analyses) and secreted cytokines, chemokines and mediators in functional assays.

Yr2: a) integration of wet-laboratory data with pathway analysis tools available in the Tsoka laboratory to identify specific signatures associated with effector functions and potency; b) blocking specific molecules in the pathways identified to validate their importance in effector mechanisms and validate the findings of our computational strategy.

Yr3: Interrogation of antibody-dependent activation signatures in patient/clinical settings: molecular and immune activation profiles of healthy volunteer and patient monocytes/macrophages in functional assays with our antibody panel and clinical lead antibody known to activate monocytes/macrophages against cancer.

This project will advance our understanding of the unique abilities of IgE class antibodies to unleash the cytotoxic potential of tumour-resident immune cells against cancer and help support this exciting immunotherapy approach to benefit patients.

National Institute for Health Research

Transcriptional profiling of adipose tissue in cardiovascular disease patients using machine learning

Supervisor 1
Name: Dr Kerrin Small
School/Directorate: FOLSM
Email: Kerrin.small@kcl.ac.uk

Supervisor 2
Name: Alan Hodgkinson
School/Directorate: FOLSM
Email: alan.hodgkinson@kcl.ac.uk

Abstract
Obesity is rapidly increasing worldwide and is an independent predictor for the development of coronary artery disease and cardiovascular events. Adipose tissue has limited capacity for the storage of triglycerides and glucose, and hypertrophic cells drive macrophage accumulation and inflammation in adipose tissue, leading to an increased risk of cardiovascular disease.

In this project we aim to characterise the transcriptome from the adipose tissue of ~600 patients with coronary artery disease and use advanced computational techniques including machine learning to identify key components of disease progression. This data will then be complemented with adipose gene expression profiles and associated intermediate phenotypes from 800 individuals from TwinsUK cohort that have been sampled at two time points across a 10 year period. Using information gained from the analysis of coronary artery patients, we will then develop models to predict risk and future events of cardiovascular disease, which can be validated in the TwinsUK cohort.

National Institute for Health Research

Optoacoustic and MRI evaluation of maternal and neonatal brown adipose tissue function and relationship to dyslipidaemia in pregnancies complicated by maternal metabolic disease.

Supervisor 1
Name: Professor Catherine Williamson
School/Directorate: School of Life Course Sciences
Email: catherine.williamson@kcl.ac.uk

Supervisor 2
Name: Dr Samuel Powell
School/Directorate: School of Biomedical Engineering and Imaging Sciences
Email: s.powell@kcl.ac.uk

Abstract

Background
Intrahepatic cholestasis of pregnancy (ICP) and gestational diabetes mellitus (GDM) are maternal metabolic disorders that occur commonly in the UK. It has been shown that the children of pregnancies affected by these disorders have increased rates of obesity and associated cardiometabolic risk factors, including impaired glucose tolerance and dyslipidaemia. The supervisors have unpublished data to show increased fetal hepatic lipids and umbilical cord blood lipids in pregnancies complicated by cholestasis and diabetes and murine data to show that progesterone-mediated loss of brown adipose tissue (BAT) function causes gestational dyslipidaemia. This project will quantify the size and function of BAT in mothers and babies of women with gestational diabetes and cholestasis using MRI and optoacoustic imaging scans. The student will also evaluate the relationship between BAT function and dyslipidaemia.

Hypothesis
Mothers with ICP/GDM and their new born babies have reduced BAT function secondary to elevations of progesterone and progesterone sulphates. This exacerbates maternal and neonatal dyslipidaemia.

Objectives
1. To perform maternal and neonatal MRI and optoacoustic imaging to establish the impact of maternal metabolic disease on BAT size and function.
2. To establish whether there is a relationship between BAT function and the concentration of specific lipid species in maternal/fetal serum.

Methods to be used
Study participants will be recruited from the antenatal clinics at St Thomas' Hospital. Maternal and fetal dyslipidaemia will be measured using MRI, optical, and optoacoustic imaging. Both conventional diffuse optical and optoacoustic methods exploit the absorption spectrum of lipids, oxy-, and deoxyhaemoglobin in order determine the presence and functional state of BAT using non-invasive, non-ionising near-infrared light. Raw data are collected at multiple wavelengths are reconstructed to form images prior to spectral unmixing to determine the concentration of each of the relevant chromophores. Activation of BAT is indicated by dynamic variations in the total blood volume and oxygen driven by increased tissue oxygen demand. MRI imaging will also be used to localise maternal and fetal BAT

depots in pregnant mothers and neonates. In parallel with the optoacoustic imaging, the student will investigate whether BAT function can be evaluated using fat T2* relaxation time mapping and signal-fat-fraction (SFF) analysis based on a commercially available modified 2-point-Dixon (mDixon) water-fat separation method.

Neonatal lipid profiles will be evaluated using Guthrie blood spots taken as part of routine clinical assessment. Lipids will be isolated from the dried blood spots. Following extraction from the disc of dried blood, together with addition of authentic internal standards, samples will be analysed by direct infusion high resolution mass spectroscopy (DIHRMS) and liquid chromatography mass spectroscopy (LC-MS). Following collection of spectra and processing with XCMS, univariate and multivariate statistical analysis will be performed for relative quantification of lipid species and comparisons between groups.

National Institute for Health Research

Integrative personalized modelling of immunome, microbiome and metabolome in cancer patients

Supervisor 1
Name: Dr Saeed Shoaie
School/Directorate: Centre for Host-Microbiome Interactions – Translational Systems Biology Group
Email: saeed.shoaie@kcl.ac.uk

Supervisor 2
Name: Dr Shahram Kordasti
School/Directorate: School of Cancer and Pharmaceutical Sciences
Email: shahram.kordasti@kcl.ac.uk

Abstract
Cancer is heterogenous disease and the precise and effective therapy varies between patients. Multi-omics data emerge that the microbiome and their metabolic mediator can influence cancer progression and response to different therapy such as checkpoint inhibitor. Existence of symbiotic microorganisms in the human gut provides different functions for the host such as conversion of nutrients, training of the immune system, and resistance to pathogens. Dysbiosis of microbiome has been associated to metabolic diseases and cancer.

Increased understanding of the interactions between the microbiota, diet and environmental effects has allowed us to design efficient treatment strategies for better cancer treatment, increasing its efficacy and addressing precision medicine. Inflammatory environment, which is likely to be the result of a combination of microbiome changes, age related increases in pro-inflammatory cytokines and activation of the inflammasome pathway in cancer pathophysiology.

Investigating the role of microbiome in an inflammatory related neoplasm which could potentially be used as a biomarker for response to therapy in different cancers. During this PhD Project, multi-omics data (transcriptomics, metagenomics, metabolomics and immunome) together with the clinical, physiological parameters and therapies will be utilized.

During this project the student will have access to different types of cancer data and samples through KCL cancer centre. The student will apply cutting-edge systems-level analysis platform to integrate the data with the objective to embark the genotype-phenotype relationships in response to therapy in lung, colon, kidney cancer, myelodysplastic syndromes and acute myeloid leukaemia. The integrated model of gene-protein-reaction network so called Genome-Scale Model (GEM), where different data on transcriptomics, metabolomics and metagenomics can be integrate together and linked to immunome.

The outcomes will also lead to a personalized approach and better stratification of patients based on multi-dimensional data integration. These findings will also provide valuable information to perform personalized predictive modelling to simulate and coupling treatment efficacy based on host-microbiome interactions and therefore suggest an improved microbiome-based personalized medicine pathway for cancer therapy.

National Institute for Health Research

Bioinformatic integration of multi-omics datasets to accelerate precision medicine approaches in inflammatory skin disease

Supervisor 1
Name: Dr Paola Di Meglio
School/Directorate: School of Basic and Medical Biosciences/ St John's Institute of Dermatology
Email:paola.dimeglio@kcl.ac.uk

Supervisor 2
Name: Dr Alessandra Vigilante
School/Directorate: School of Basic and Medical Biosciences/Center for Stem Cells and Regenerative Medicine
Email:alessandra.vigilante@kcl.ac.uk

Abstract
Inflammatory skin diseases psoriasis and atopic dermatitis (AD) are common, chronic conditions with a major impact on patient's quality of life, and posing a significant financial burden on the healthcare system. Both diseases are underpinned by the combination of genetic and environmental factors, leading to dysregulated immune responses. Significant advances in understanding their pathogenesis has been translated into the usage of biological treatments or "biologics", which targets key immune pathways, such as TNF, IL-12/23 and IL-17 in psoriasis and IL-4/IL-13 in AD.

However, as human genomes are complex and regulated at multiple levels, clinical response to biologics is heterogeneous and biologics are not effective in every patient, suggesting that additional cellular players and molecular pathways, may be involved in disease pathogenesis. Moreover, with an expanding range of biologics available for psoriasis targeting different cytokines, there is ample scope to implement precision medicine approaches, where the selection of the drug is not "by-trial-and-error", as it currently happens, but is guided by biomarkers that predicts patient response to therapy.

To this end, we have employed a range of high-throughput 'omics' platforms to obtain the genomic, metabolic and immune profile of a cohort of untreated atopic dermatitis and psoriasis patients, as well as healthy controls. Moreover, we have been monitoring and profiling 2 cohorts of psoriasis patients receiving different biologics to identify biomarkers of clinical response. While each of these assays offers a glimpse of the complex system, they all measure rather interactive but interdependent events. Integrative analysis of multiple layers of data points from different sources is thus essential to understand psoriasis and AD etiopathogenesis and to identify potential novel therapeutic targets.

We will implement bioinformatics pipelines that apply data integration methods such as highlight matrix factorization methods, Bayesian methods, and network-based methods. This will allow to elucidate specific pathway constituents, interactions and mechanisms in the immune system implicated in the response to the biologics and to identify clinically relevant biomarkers. Finally, most promising novel target will be validated in independent perspective cohorts of patients using ex vivo models such as whole skin explant cultures in the presence of blocking reagents followed by imaging and gene expression analysis. Putative biomarkers will also be validated in retrospective additional cohorts of patients receiving biologics.

National Institute for Health Research

Developing the new generation of mesenchymal stem cells for regenerative medicine

Supervisor 1
Name: Dr Anita Grigoriadis
School/Directorate: Cancer and Pharmacological Studies
Email: anita.grigoriadis@kcl.ac.uk

Supervisor 2
Name: Professor Francesco Dazzi
School/Directorate: Cancer and Pharmacological Studies
Email: Francesco.dazzi@kcl.ac.uk

Abstract

Mesenchymal stem cells (MSCs) have received center-stage attention because they exhibit potent anti-inflammatory activities that have been extensively tested in several medical conditions as well as in regenerative medicine. The broad clinical use has been justified by its non-specific mode of action that targets any cells of the immune system. We have recently shown that MSC anti-inflammatory properties are largely dependent on MSC undergoing in vivo apoptosis and consequently being phagocytosed by tissue macrophages. Macrophages, thus converted into immunosuppressors, re-programme the microenvironment and prompt tissue regeneration. Importantly, the ability of patients' immune cytotoxic cells to induce MSC apoptosis correlates with clinical responses to MSC infusions.

A yet unresolved issue is which subpopulation of MSC is more therapeutically efficacious. Our discovery that MSC must undergo apoptosis to become anti-inflammatory offers the opportunity to use this property to measure clinical efficacy. Therefore, we propose to identify which subsets is selectively sensitive to undergo apoptosis and characterise the molecular machinery involved in delivering the apoptotic signal in MSC and the signalling cascade leading to the anti-inflammatory activities.

Aim 1. MSC heterogeneity. In the first phase we will characterise the heterogeneity of MSC as far as its sensitivity to undergo apoptosis when in contact with activated cytotoxic cells is concerned. We are going to perform RNAseq in MSC obtained from different tissue sources and, based on the profile, we will identify a selection of markers to classify different cell clusters by mass cytometry. Each cluster will be sorted and tested for their ability to undergo apoptosis when exposed to mononuclear cells obtained from patients affected by conditions that we know are associated with increased levels of MSC-killer cells. These will include Crohn's disease and liver failure.

Aim 2. Receptor-ligand pair inducing MSC apoptosis. Cell-surface proteins of apo-sensitive MSC will be labelled with membrane-impermeable sulfo-NHS-SS-biotin from intact cells after a brief exposure to patients' mononuclear cells. Labelled cell-surface proteins will be captured and subjected to high-resolution mass spectrometry and the proteins identified using the Global Proteome Machine (GPMdb). In parallel, apo-sensitive MSC will be sorted after contact with PBMC and subject to gene array. The proteomics and transcriptomics data will be analysed in combination to identify the transmembrane proteins that are associated with the upregulated gene profile in apoMSC.
The results will lead to the identification of the receptor-ligand interaction(s) inducing MSC apoptosis. The extrinsic and intrinsic pathways of cell death as well as the inflammasome machinery will be taken into consideration in the analysis.

Aim 3. Apoptosis signalling and tissue regeneration. The transcriptomic profiles will provide a set of candidate molecules by which apoMSC deliver the anti-inflammatory and tissue repair activities. We will reproduce an *in vivo* system in which the apo-sensitive MSC subset will be exposed to patients' PBMC in the presence or absence of inhibitors of the candidate molecules and the supernatant of apoMSC tested for immunosuppressive, pro-proliferative and anti-apoptotic activities on primary cell cultures.

National Institute for Health Research

Data integration of endocardial cell transcriptomes -in search of the heart valve switch

Supervisor 1
Name: Professor Rebecca Oakey
School/Directorate: BMBS
Email: rebecca.oakey@kcl.ac.uk

Supervisor 2
Name: Professor Tim Hubbard
School/Directorate: BMBS
Email: tim.hubbard@kcl.ac.uk

Abstract
Development of the human heart involves multiple cell lineages forming the endocardium, myocardium and epicardium. The identification of multipotent cardiac progenitor cells has provided insight into myocardial lineage specification but much less is known about the origins of the endocardium. Endocardium is the inner most layer of cells that lines the valves and chambers of the heart and when these cells fail to form properly, a range of cardiovascular diseases result.

Little is known about how the endocardium is specified or the mechanisms by which endocardial cells differentiate at the molecular level during heart development. Gene regulation studies have been advanced through the development of high-throughput sequencing methods to interrogate the genome. Understanding cell-specific gene expression and chromatin architecture will aid the molecular definition of the endocardial cell and provide a basis for utilizing this cell type in therapeutic cell re-programming strategies to reduce the burden of cardiovascular disease.

We have generated pilot data consisting of the transcriptomes (RNA-seq) and methylomes (BS-seq) of endocardial cells and endothelial cells from a model organism that have identified a set of candidate genes that are specific to endocardial cells but absent in endothelial cells and their common progenitor cell. These candidate genes are being validated using a variety of assays for roles as key switches in cell fate decisions. We would like to validate these findings in other large genome wide data sets, especially those from human cells to improve the capacity of our study to identify both genes and regulators specifying endocardial cells.

This project will dovetail with our current studies and focus on the computational analysis of our own ‘omics data compared to those from the public domain in endocardial cells. Human sequencing data from large-scale single-cell RNA-Seq studies from populations of Human Induced Pluripotent Stem Cell derived endothelial cells (for example from Paik et al, 2018, Circulation Research) and studies aimed at identifying the key features of vascular endothelial cells compared to other cells (eg. Sabbagh et al, eLife 2018) will be mined using pipelines already in place in the lab and comparative analysis will direct the choice of top candidate genes. We have a tissue-culture based differentiation assay combined with CRISPR ablation in place for validating endocardial specific genes identified through this approach.

National Institute for Health Research

Why do some babies fail to grow *in utero*? Investigating a novel molecular diagnostic tool for identification and stratification of pregnancy complications.

Supervisor 1
Name: Dr Marika Charalambous
School/Directorate: FOLSM / Department of Medical & Molecular Genetics
Email: marika.charalambous@kcl.ac.uk

Name: Dr Ferdinand von Meyenn
School/Directorate: FOLSM / Department of Medical & Molecular Genetics
Email: ferdinand.vonmeyenn@kcl.ac.uk

Abstract
Babies who grow poorly in the womb, called small for gestational age (SGA), are at increased risk of a number of adverse outcomes, including health problems for both mother and baby. Screening pregnant women in an attempt to detect SGA babies has the potential to reduce the number of these adverse outcomes, which accounts for the death of about 4000 babies every year in the UK. Currently it is difficult to monitor growth of the baby in the womb, and even harder to predict if there will be health problems associated with this. While there are a number of scan and blood markers for SGA babies, none of these is good enough on its own to predict them.

We have recently found that by measuring a protein called DLK1 in the mother's blood we could improve the prediction of whether the baby would be small. In this first study we examined only a small selection of pregnant women, but it appeared that abnormal DLK1 levels were associated with failure to make an efficient placenta. The placenta is an extraordinary organ that is necessary for providing nutrition and oxygen to the baby from the mother's blood, and removing waste. Many of the complications of pregnancy are thought to occur because the placenta does not develop properly.

However, we still have much to discover about which genes are needed to make a healthy placenta, and how this goes wrong if the genes are mutated. Interestingly, DLK1 is made in the placenta. When we studied mice with a mutated DLK1 gene we found that they had smaller placentas and a reduced area for nutrient and oxygen exchange between the maternal and fetal blood. However, we currently don't understand why DLK1 is required to make a healthy placenta.

The aim of this project is firstly to see how measuring DLK1 in the blood of a large number of pregnant women can help us improve prediction of small babies. We have access to scan information, blood samples and placental samples from a large scale study of unselected women in their first pregnancy (The POPs cohort consisting of ~3000 women).

We will use samples from the POPs cohort to understand the causes of low DLK1 production in pathological pregnancies. The project will utilise data high-throughput sequencing technologies to interrogate the regulation of DLK1 in pregnancy, and determine if genomic sequence variation and/or changes in the methylation of the DLK1 locus can be used as a diagnostic for fetal growth restriction. The student will develop a methodology for the analysis of targeted high-throughput bisulphite sequencing data and relate these data to a richly phenotyped pregnancy cohort.

National Institute for Health Research

Use of 'Axon-like' Nanofibres to quantify Glioblastoma cells migration

Supervisor: Dr Claire Wells
School/Division & CAG: Cancer Sciences
KCL/KHP E-mail: claire.wells@kcl.ac.uk
KCL/KHP Website: https://kclpure.kcl.ac.uk/portal/claire.wells.html
http://wellslaboratory.blogspot.co.uk/

Supervisor 2
Name: Dr Andrea Serio
School/Directorate: TEB
Email: andrea.serio@kcl.ac.uk

Abstract
Background: Glioblastoma multiforme (GBM) is the most aggressive brain tumour in adults with an incidence of over 3 per 100,000 persons, a median age of 64 years and a great burden for society. Despite current advances, the current standard of care is ineffective and survival prognosis remains around just over a year from diagnosis. As tumour cells spread on axon bundles on white matter tracts, innovative therapies targeting migratory infiltrating resistant cancer cells are needed.

Scientific Hypothesis: We hypothesise that a migratory subpopulation of cancer-derived cells depend for their migration on combinations of chemicals and physical signals from the environment. We set to define these signals by studying the migration of GBM patient-derived cells onto axons in microgrooves and onto engineering Nanofibers with defined physical properties, bio-functionalised with specific chemical signals. This project focuses on elucidating migration as a permissive and inductive process regulating tumour progression. Importantly, it opens the route to explore future drug discovery screening assays and aligns with our current application to the CRUK pioneer awards scheme to explore the feasibility of a future medical device based on these Nanofibers.

Experimental Plan: We have previously developed high content imaging platforms to analyse cell behaviour and migration of GBM-derived stem cell-like cells from patients (DD). We have defined the conditions to set up microgrooves containing axons that could lodge Nanofibers and be analysed with imaging methods (AS). We here aim to: (A1) develop a robust workflow to characterise and quantify the migration of a panel of patient-derived glioblastoma cells; (A2) compare migration of cells on axons in microgrooves versus artificial Nanofibres; (A3) define variation in the migration properties in populations and subpopulations of several patients-derived GBM cultures and relations with molecular subtype of disease, e.g. proneural, mesenchymal. Main deliverable for this project is an assessment of the robustness of the platform as screening bed for chemical compounds specifically acting on the migration of glioblastoma cells. This project dissects the relevant tumour microenvironment for this disease offering precision-medicine approaches towards new therapies.

National Institute for Health Research

A therapeutic HPV vaccine, for adjuvant therapy of HPV positive Head and Neck Cancer

Supervisor 1
Name: Professor Mahvash Tavassoli
School/Directorate: Dental Sciences and Craniofacial
Email: Mahvash.tavassoli@kcl.ac.uk

Supervisor 2
Name: Professor Farzin Farzaneh
School/Directorate: School of Cancer & Pharmaceutical Sciences
Email: Farzin.farzaneh@kcl.ac.uk

Abstract

Objectives: To identify common antigens present in HPV associated malignancies, in order to generate a small (as large as necessary) library of synthetic peptides that would be suitable for vaccinations in combination with a CASAC adjuvant formulation, in order to develop a therapeutic vaccine for HPV related cancers. The specific objectives of the present study are:

- Database searches for HPV strains that are commonly present in HPV positive cervical and head and neck carcinomas (HPV-16, 18, 31 and 33, etc).
- In addition to HPV E6 and E7 genes expressed by these variants, identify which other HPV genes are commonly expressed in these malignancies.
- Which are the most conserved regions (amino acid/peptide sequences with least variation) both across the HPV strains that are associated with these malignancies and between different patients.
- Use the already available algorithms (e.g. SYGPEITHI) to predict a peptide library of HPV genes that are expressed in the associated malignancies. The objective being the identification of peptide sequences from the conserved areas of cancer associated HPV genes for MHC class-I presentation. The aim being to identify a library (as large or small as is necessary) to cover each gene, for each strain in greater than 90% of the Caucasian population.
- Use other algorithms that are rapidly becoming available to identify a panel of suitable MHC Class-II presented promiscuous peptides, in order to provide adequate help by activation of the CD4+ helper cells.
- Employ human in vitro studies using PBMCs isolated from healthy donors (obtained from the blood transfusion service) and HPV positive cancers, to assess the potential in vitro efficacy of the vaccine - with the readout being the induction of effective cellular immunity against HPV positive, but not negative, tumour cell lines, and primary tumour isolates.

The other component of pre-clinical studies will be assessment of safety and efficacy in murine models and this will be completed by our current postdoctoral fellow who is working on other aspects of vaccination induced cellular immunity in the Farzaneh laboratory.

National Institute for Health Research

Is the brain already programmed for later psychiatric disease at birth? Using machine learning with neuroimaging and genomic data to study gene-environment interactions.

Supervisor 1
Name: Prof David Edwards
School/Directorate: Bioengineering and Imaging Sciences
Email: AD.Edwards@kcl.ac.uk

Supervisor 2
Name: Dr Maria Deprez
School/Directorate: Bioengineering and Imaging Sciences
Email: maria.murgasova@kcl.ac.uk

Abstract

The development of serious psychiatric disease is influenced by the interplay of environmental factors and genetic risk, but the neural substrate for this interaction is poorly understood. A significant proportion of the genetic contribution to psychiatric risk is polygenic resulting from the additive effects of multiple alleles throughout the genome, each of which has an individually small effect. A wide range of environmental risks have been postulated, but one of the most pervasive appears to be biological stress during early development, and premature birth is an extreme perinatal stress which is associated with a significant increase in later mental health disorders. By combining advanced neuroimaging, particularly connectomics, with genomic data in large specialised patient cohorts we now have the opportunity to test specific hypotheses about the nature of the perinatal environment-gene interaction. Genomic imaging approaches have proved highly successful in the perinatal period, discovering critical genetic influences on abnormal brain development. (ref 1,2)

Our preliminary work has discovered a neuroanatomical correlate for increased polygenic risk of psychiatric disease. In infants subjected to the stress of preterm birth increased polygenic risk for psychiatric disease was associated with: reduced volume of the lentiform- (β=-0.24, p=8x10^{-4}) and subthalamus (β=-0.18, p=0.01) (Figure 1.). This anatomical substrate has not been found in adults. These results generate the testable hypothesis that: *increased polygenic risk for psychiatric disease is a risk factor for abnormal brain development in infants subject to perinatal stress*. If this hypothesis is correct the gene pathways involved may suggest novel biologic mechanisms behind mental health disorders.

Using the large specialised sample of advanced images and collateral genomic data provided by the Developing Human Connectome Project (dHCP; European Research Council €15m; Lead PI Edwards) we plan to test this hypothesis. If it is correct, the result will be replicated in preterm infants in the dHCP but not in the term born infants. If on the other hand these genes lead to abnormal brain development in the absence of perinatal stress, the same changes will be observed in normal term infants. The dHCP cohort has acquired state-of-the-art brain MR data for approaching 1000 infants and fetuses, with collateral maternal, clinical, genetic and neurocognitive data, making it the ideal dataset to address

this hypothesis, and to further explore the detailed functional and structural connectivity associated with perinatal stress and polygenic risk.

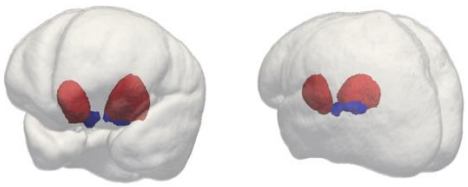

Figure 1. Subthalamic nucleus (blue) and Lentiform nucleus (red) within glass brain (top right and top left).

Having first addressed this specific hypothesis the project will further explore the relationship between genetic risk for a range of different psychiatric pathologies and both structural and functional MRI endophenotypes.

1: Krishnan ML et al. Machine learning shows association between genetic variability in PPARG and cerebral connectivity in preterm infants. **Proc Natl Acad Sci U S A.** 2017 Dec 26;114(52):13744-13749. 2: Krishnan ML et al . Integrative genomics of microglia implicates DLG4 (PSD95) in the white matter development of preterm infants. **Nature Commun.** 2017 Sep 5;8(1):428.

National Institute for Health Research

Predicting neurocognitive delays and behavioural disorders in children using MRI: a machine learning approach

Supervisor 1
Name: Professor Serena Counsell
School/Directorate: Biomedical Engineering & Imaging Sciences
Email: serena.counsell@kcl.ac.uk

Supervisor 2
Name: Dr Jonathan O'Muircheartaigh
School/Directorate: Institute of Psychiatry, Psychology and Neuroscience
Email: JonathanOM@kcl.ac.uk

Abstract
Neurodevelopmental difficulties, including inattention, hyperactivity, autism spectrum disorders, and learning difficulties are identified in early childhood in 15% of children in the UK. Some groups of children are at increased risk of developing such neurocognitive or behavioural deficits. Current approaches to predict which of these at-risk children will require additional support are limited in their effectiveness and specificity. Neuroimaging approaches such as MRI have focused on case-control studies, assessing specific groups of at-risk infants.

These approaches, while useful to see average effects in specific groups, can miss common brain changes from several distinct risk factors that may lead to the same phenotype (e.g. learning difficulties in later life). Moreover, in order to be useful in clinical practice, it is necessary to be able to predict developmental delay at an *individual* level, which requires a high discriminative power that has been lacking to date. With improving methods of brain morphometry and microstructure analysis, as well as the use of sophisticated and interpretable machine learning methods, individual based predictions may be feasible.
We hypothesise that neurocognitive and behavioural delays at school age are associated with altered brain development in utero and in the early neonatal period, and that brain development is modified by factors including gestational age at birth, sex, socioeconomic status, parental stress.

To test this hypothesis, we will use a large neuroimaging database of fetal and neonatal structural MRI data. We will model brain development using >600 fetal and 1200 neonatal brain MR images that have already been acquired. These data include both healthy controls and those at-risk including; preterm infants, infants with congenital heart disease, those at high genetic risk of developing ASD, intrauterine growth restriction, twin-to-twin transfusion syndrome, and offspring of obese women.

We will use parametric and non-parametric approaches to model normal brain development and use machine learning approaches to identify outliers and cluster cases based on imaging and cognitive phenotypes. Specifically, we will:

1. Using multimodal imaging measures, including cortical thickness, regional and global brain volume and white matter properties (fractional anisotropy, FA and mean diffusivity, MD), to model typical brain development across fetal and neonatal life in 50% of our infants who have normal outcome in early childhood (500 infants).

2. To validate this model, we will assess the ability of the model to predict early childhood outcome in the remaining cases; including over 800 cases with an atypical outcome and an additional 500 infants with normal development.
3. Assess the ability of our approach to identify clinical clusters (e.g. risk groups - preterm, CHD, IUGR) and clinical outcomes.
4. To demonstrate that the model is generalisable, we will assess the performance of the model to predict childhood outcome using independent datasets at other hospitals, provided by collaborators at other centres (Utrecht, Toronto, Edinburgh).

This project has the potential to stratify infants at risk of subsequent cognitive delay or behavioural disorders during development when there may be a window of opportunity for therapeutic intervention to improve outcome.