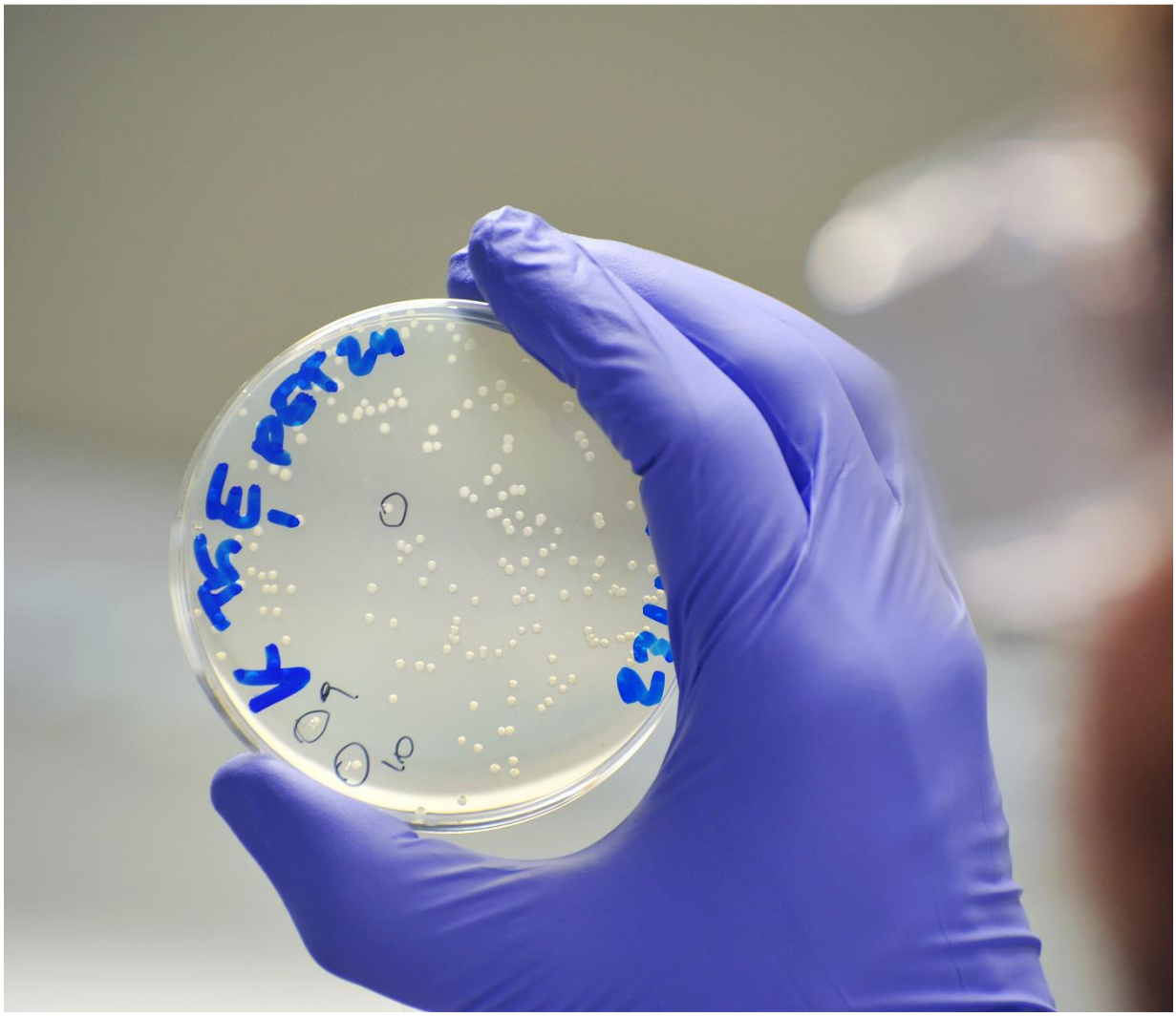
wellcome centre human genetics

UNIVERSITY OF OXFORD

NUFFIELD DEPARTMENT of MEDICINE



Away Day

6th September 2019

Maths Institute, ROQ, OX2 6GG

Table of Contents

Schedule

From 08.45	Registration
09.15	Welcome
09.30	Speaker 1
09.55	Speaker 2
10.20	Poster blitz session 1
10.45	Coffee & tea break
11.10	Speaker 3
11.35	Poster blitz session 2
12.30	Lunch and poster session
14.00	Keynote speaker: Professor Marcus du Sautoy, The Oxford Simonyi Professor for The Public Understanding of Science
14.45	Speaker 4
15.10	Speaker 5
15.35	Coffee & tea
15.55	Speaker 6
16.20	Prize giving for posters and public engagement
16.45	Close, followed by drinks and canapes in the foyer

The running order of the speakers will be determined on the day. The speakers will be:

Anjali Hinch, Donnelly Group
Title: This is how eggs and sperm do start: the tools that mend and the tools that part.

Luca Veccia, Jones Group
Title: The sweet taste of the Wnt pathway - structural insights into Wnt transport in the extracellular space.

Mark Boyce, Stuart Group
Title: Dissecting the assembly pathways of multi-layered viruses in situ using focused ion beam (FIB) milling coupled with cryo electron tomography (cryoET).

Peter Yeow, Chapman Group
Title: Exploiting TRIM37-driven centrosome dysfunction to eliminate 17q23-amplified breast cancer cells.

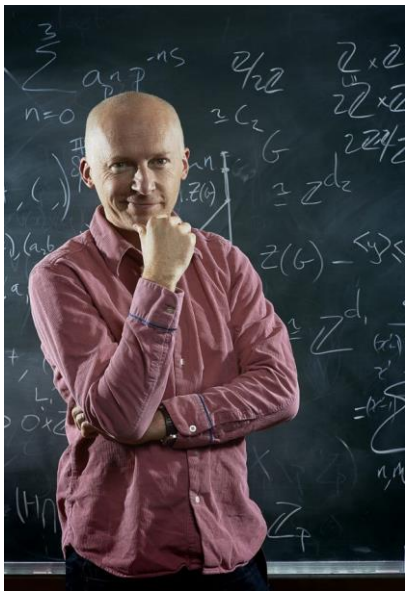
Vedanta Mehta, Tzima Group
Title: The Moonlighting Mechanosensor.

Victor Yeung, Dendrou Group
Title: Investigating pathophysiological mechanisms shared across immune-mediated diseases

Key Note Speaker

Marcus Du Sautoy

The Oxford Simonyi Professor for The Public Understanding Of Science



Marcus du Sautoy is the Charles Simonyi Professor for the Public Understanding of Science at the Oxford University, a chair he holds jointly at the Department of Continuing Education and the Mathematical Institute. He is also a Professor of Mathematics and a Fellow of New College. He was made a Fellow of the Royal Society in 2016. In 2001 he won the prestigious Berwick Prize of the London Mathematical Society awarded every two years to reward the best mathematical research made by a mathematician under 40.

In 2004 Esquire Magazine chose him as one of the 100 most influential people under 40 in Britain and in 2008 he was included in the prestigious directory Who's Who. In 2009 he was awarded the Royal Society's Faraday Prize, the UK's premier award for excellence in communicating science. He received an OBE for services to science in the 2010 New Year's Honours List. He also received the Joint Policy for Mathematics Board Communications Award for 2010 and the London Mathematical Society Zeeman Medal for 2014 for promotion of mathematics to the public.

Marcus will give a key note talk on Artificial Intelligence, based on his recent book: "The Creative Code: How AI is Learning to Write Paint and Think"

Abstract List

Group: Bashford-Rogers **Author:** Bo Sun **Poster no: 1**

Dichotomous pathophysiology of monoclonal antibodies against LGI1

Bo Sun MBBS,[1,2] Melanie Ramberger PhD,[1] Antonio Berretta PhD,[1] Sophia Michael MBChB,[1,2] Jeanne MM Tan MD PhD,[1] Tianrong Yeo MRCP,[3] Rachael Bashford-Rogers PhD,[4] Sofija Paneva,[1] Victoria O'Dowd MBiol,[5] Neesha Dedi PhD,[5] Robert Griffin MSc,[5] Maria I Leite DPhil,[1,2] Arjune Sen FRCP PhD,[1,2,6] Daniel Anthony PhD,[3] David McMillan PhD,[5] Diane Marshall PhD,[5] Daniel Lightwood PhD,[5] Patrick Waters PhD,[1] Sarosh R Irani FRCP DPhil[1,2,6]

1. *Oxford Autoimmune Neurology Group, Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK*
2. *John Radcliffe Hospital, Oxford University Hospitals, Oxford, UK*
3. *Experimental Neuropathology Group, Department of Pharmacology, University of Oxford, UK*
4. *Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK*
5. *UCB Pharma, 208-216 Bath Road, Berkshire, UK*
6. *Oxford Epilepsy Research Group, University of Oxford, Oxford, UK*

Objective

Leucine-rich glioma inactivated 1 (LGI1) autoantibodies are found in patients with limbic encephalitis and, most characteristically, faciobrachial dystonic seizures. We generated patient-derived monoclonal antibodies (mAbs) against LGI1 to study epitopes, genetic sequences and pathogenic molecular mechanisms.

Methods

LGI1-autoantibody positive patient sera (n = 30), CSFs (n = 10), longitudinal sera (n = 15), and, from two patients, peripheral B-cell-derived mAbs (n = 14), were studied using sequence analyses and live cell-based assays. Their functionality was characterised with LGI1-transfected HEK293T cells and rat hippocampal neurons.

Results

Consistent with polyclonality, all sera and 8/10 CSF bound to both the Leucine-rich repeat (LRR) and Epitempin-repeat (EPTP) domains of LGI1, with temporally stable LRR- and EPTP-specific antibody levels. In contrast to serum and CSF, the 14 patient-derived mAbs recognised either the LRR or EPTP domain. Sequence analysis revealed genetic heterogeneity of these mAbs and high mutation frequencies. LRR-specific mAbs recognized LGI1 bound to previously described interactants, a disintegrin and metalloproteinase domain-containing protein (ADAM) 22 and ADAM23, and induced LGI1-ADAM22/23 complex internalisation in HEK293T cells and live hippocampal neurons. In contrast, EPTP-specific mAbs inhibited the docking of LGI1 onto ADAM22/23.

Conclusion

Patient derived mAbs have discrete specificities and reveal dichotomous pathophysiological mechanisms. The dualistic action of polyclonal autoantibodies in LGI1 antibody-mediated disease may contribute differentially to the clinical spectra of disease.

Group: Bhattacharya **Author:** Christopher Lynch **Poster no: 2**

A knottin scaffold directs CXC-chemokine binding specificity of tick evasins

Christopher Lynch[1], Maud Deruaz[2], Angela W. Lee[1], Graham Davies[1], Kamayani Singh[1], Yara Alenazi[1], James R.O. Eaton[1,3], Akane Kawamura[1,3], Jeffrey Shaw[2], Amanda Proudfoot[2], João M. Dias[2] & Shoumo Bhattacharya[1]

[1]RDM Division of Cardiovascular Medicine, University of Oxford, Oxford, United Kingdom.
[2]Serono Pharmaceutical Research Institute, Geneva, Switzerland
[3]Department of Chemistry, University of Oxford, Oxford, United Kingdom.

Tick evasins bind either CC or CXC-chemokines by a poorly understood promiscuous or "one-to-many" mechanism to neutralize inflammation. Here we identify 27 CXC-binding evasins homologous to EVA3 and define two functional classes. The first, which includes EVA3, exclusively binds ELR+ CXC-chemokines, whereas the second binds both ELR+ and ELR- CXC-chemokines, including in several cases CXCL10, but, surprisingly, not CXCL8. The X-ray crystal structure of EVA3 reveals a single anti-parallel β-sheet with six conserved cysteine residues that form a disulfide-bonded knottin scaffold creating a contiguous solvent-accessible surface. Swapping and deletion analysis identifies distinct knottin scaffold segments responsible for different CXC-chemokine binding activities, indicating that differential ligand positioning, at least in part, plays a role in promiscuous binding. The solvent-accessible surfaces of these scaffold segments have distinctive shape and charge, which we suggest drives chemokine binding specificity. These studies provide insight into how CXC-binding tick evasins achieve class specificity but allow promiscuous binding

---

Group: Bowden **Author**: Rhys Price **Poster no: 3**

A molecular basis for unique ephrin-B1 utilisation by Cedar virus

Rhys Pryce, Bowden Group

The Bowden group utilises structural biology and biochemistry to delineate molecular-level details of the interactions between viral surface proteins and both host-cell receptors and antibodies. By describing aspects of virus entry and antibody-mediated neutralisation the group aims to provide a foundation for improved viral surveillance, vaccination and treatment strategies that can combat neglected but lethal human and animal pathogens. The spillover of viral pathogens from animal reservoirs into human populations accounts for a large proportion of world-wide emerging disease events and poses an indeterminate threat to human health, economy, and biodefence. As obligate intracellular parasites, viruses must identify and infect permissive host cells, typically utilising attachment-mediating proteins that recognise specific cellular receptors. The identity of viral entry receptors has a profound influence on host range, tissue- and cell-type tropism, and disease pathologies. Exploitation of highly species-conserved ephrin molecules for cellular entry by henipaviruses is intrinsically associated with pathological features that delineate these viruses as amongst the most lethal known. To interrogate functional diversity within the genus, we sought to determine the receptor repertoire utilised by a genetically distinct henipavirus, Cedar virus (CedV). We show that CedV possesses a unique cellular entry receptor repertoire which, in addition to utilisation of the common henipavirus receptor, ephrin-B2, includes the hitherto uncharacterised interaction with ephrin-B1. Crystal structures reveal that discrete regions of CedV-G are constrained to maintain high-affinity ephrin recognition and highlight a region of molecular specificity that is a key determinant of ephrin-B1 selectivity. These data set a precedent for ephrin-B1 utilisation and raise intriguing questions regarding the ecological and pathological

implications of expanded receptor repertoires in extant henipaviruses. Such studies enhance understanding of the determinants of receptor usage and therein facilitate assessment of the zoonotic and pathogenic potential of emerging viruses.

Group: OGC **Authors:** Amy Trebes **Poster no: 4**

Optimisations of Single Cell Transcriptomics Sequencing

Amy Trebes, Moustafa Attar, Ben Wright, Eshita Sharma, John Broxholme, James Docker, Tim Wright, Angie Green, Lorne Lonie, Helen Lockstone, Rory Bowden, David Buck

Single cell RNA-seq has become a robust and commonly used technology over the last 5 years. At the Oxford Genomics Centre, we utilise the high-throughput power of a dedicated single cell platform but also compliment this with lower-throughput, targeted workflows, starting from FACS sorted single cells.
Our dedicated single cell technology of choice, 10X Genomics, isolates individual cells from a mixed population for expression counting, however the platform can't target subtypes of cells. If a cell population contains a target cell type in low abundance it is possible to FACS sort these into a tube for loading, but to obtain a high enough cell count for the technology to work well, the number of cells wasted are sometimes too high a cost. As an alternative, we support RNA-seq of individual single cells that are FACS sorted directly into PCR-plates.
'SMART' (Switching Mechanism at 5' End of RNA Template) technology has been through a number of chemistry optimisations since its publication in 2012 and defines the backbone of many single cell RNA-seq protocols. SMARTer and SMART-seq2 have been popular components of the Oxford Genomics Centre portfolio for many years. We have recently undertaken an overhaul of these approaches, striving to improve the service and data quality our researchers receive. Key areas of consideration included throughput, gene detection coverage and sensitivity, as well as reduced consumable consumption.
Here we share an insight into the optimisations we have adopted and the improved data quality we are able to achieve with our new protocol, soon to be released as a standard service.
If you have a small cell population you want to understand better, chat with our project managers to discuss your best options.

Group: OGC **Authors**: Maria Lopopolo

Evaluation of PromethION long-read genome sequencing methodologies for somatic variant detection in a patient with B-cell lymphoma

Hannah E Roberts[1]*, Maria Lopopolo[1]*, Alistair T Pagnamenta[1,2]*, Lorne Lonie[1], Duncan Parkes[1], Colin Freeman[1], Samantha J L Knight[2], Helene Dreau[3,4], Helen Lockstone[1], Jenny C Taylor[1,2], Anna Schuh[3,5], Rory Bowden[1], David Buck[1]
1. Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK
2. National Institute for Health Research Oxford Biomedical Research Centre, Oxford, UK
3. Oxford University Hospitals NHS Trust, Oxford, UK
4. Department of Haematology, University of Oxford, Oxford, UK
5. Department of Oncology, University of Oxford, Oxford, UK
* These authors contributed equally

Technological progress in next-generation sequencing has led to increasing use of whole genome sequencing (WGS) as a comprehensive method for informing cancer diagnosis and

treatment. Recent advances in throughput, cost and accuracy mean that generating long read data with the Oxford Nanopore Technologies (ONT) PromethION is a now a viable solution for WGS. New bioinformatic methods have been developed to take advantage of this long read data, however, much of the validation of these tools has focussed on calling germline homozygous or heterozygous variants (both single nucleotide variants and structural variants). Detection of somatic variants is more challenging as their frequency can vary depending on tumour purity and mosaicism. In this study, we used the PromethION-beta device to sequence the tumour and germline genomes for a patient with diffuse B-cell lymphoma. Through the use of a PCR-based protocol we achieve high yields of up to 96 mapped Gb per flow cell with average read lengths of 5-6kb. We examine the capability of currently available tools for calling somatic variants in ONT data by comparing the data with results from 150bp short-read sequencing of the same samples. Whilst we are able to achieve good specificity and sensitivity for calling germline single nucleotide variants (SNVs) in ONT data, results of somatic SNV calling highlight the need for the development of sophisticated joint calling algorithms for this application. We then conduct a detailed analysis of the performance of multiple long-read mappers and structural variant callers for calling large, somatic structural variants (SVs) in ONT data. Our results show the comparative performance of different tools varies significantly between SV types, and suggest that overall, long reads are especially advantageous for calling large somatic deletions and duplications. Finally, we highlight the utility of long reads for phasing clinically relevant variants by using the ONT data to confirm that a 1.6Mb deletion and a somatic p.(Arg249Met) mutation involving *TP53* are oriented *in trans*.

Group: OGC **Authors:** Simon Engledow

The Anatomy of a Successful Genomics Core Facility
Simon Engledow, Christine Blancher, Angie Green, Lorna Witty, David Buck

Not all Genomics Core Facilities are created equally.
The Oxford Genomics Centre (OGC) has established itself as a commercial service entity working within the Wellcome Centre for Human Genetics (WHG). It has become recognized as a go-to provider of genomics services to users that require far more than the low prices offered by the large commercial providers.
OGC runs under a regime of total quality management and has a dedicated team of PhD level project managers, with decades of knowledge and experience in genomics, guiding users to the optimal experimental design.
This is coupled with a motivated operational team running highly automated pipelines and exploiting cutting edge technologies to generate high-quality data. Data quality is checked continuously through the running of sophisticated data QC pipelines. Project status is tracked using fully-integrated laboratory information management systems.
OGC has earned a reputation for service quality for both high throughput sequencing and array-based studies of any size. Beyond the internal support to WHG users, OGC has become a significant resource to the whole of the University of Oxford, providing a genomics boutique that is unique in its breadth and depth of experience and ability to support a broad range of applications.

Take the stress out of your High-Throughput Genomics and work with a team always giving you more.

Group: Bull - CCMP **Author:** Aneesha Bhandari **Poster no: 5**

Mutation in NUP93 Disrupts Podocyte Cytoskeletal Architecture

Aneesha Bhandari[1], Aga Bierzynska[2], Moin Saleem[2], Richard Cornall[1], Katherine Bull[1]

1 Nuffield Dept. of Medicine, University of Oxford
2 Bristol Renal, University of Bristol

Background

The renal glomerulus acts as a sieve, filtering water and electrolytes into the urinary space but keeping proteins in the blood. Protein leak into the urine occurs when the glomerular barrier fails, and is a hallmark of many kidney diseases, including diabetic nephropathy. Variants in over 70 genes have now been linked to monogenic forms of severe protein leak (nephrotic syndrome) and these genes highlight the key role of podocytes in glomerular function.

The nuclear pore complex consists of multiple copies of 34 nucleoporin proteins (NUPs), genetic variations in seven of these, including NUP93, have recently been linked to nephrotic syndrome, but the underlying mechanisms remain unknown. By modelling disease associated variation in NUP93 we aim to define cellular pathological consequences, and highlight potential therapeutic targets.

Results

By whole genome sequencing we identified a novel G to A mutation in *NUP93* exon 12 (Ala to Thr substitution), in a patient with early onset nephrotic syndrome and his affected brother. To model this *in vitro* we used virally delivered CRISPR/Cas9 in immortalised human podocytes. With Ben Davies in the transgenic core we then generated a mouse carrying the point mutation.

In vitro, *NUP93* mutation in exon 12 results in lower NUP93 expression, but NUP93 protein and the nuclear pore are correctly localised. We observe impaired survival, reduced adhesion and altered motility in a wound healing assay. RNA sequencing highlights ACTG2 (actin gamma 2) upregulation, oxidative stress and CDC42 pathway changes. Mutant lines show morphological changes, with shorter actin stress fibres in a dorsal rather than ventral pattern. In mice, both knock-out of *Nup93,* and compound heterozygosity for a null mutation and the human point mutation are embryonic lethal on the proteinuria prone FVB strain background. However, homozygosity for the point mutation does not result in overt renal failure but induces podocyte foot process effacement on electron microscopy. To induce a more overt phenotype we now plan to challenge with Doxorubicin, prior to defining the effects on the glomerular transcriptome.

Conclusions

CRISPR based allelic series are a powerful tool for the study of human disease genes, particularly when null mutation is incompatible with survival, as in the case for NUP93 both in cell lines and in mice.

Our results point to a role for NUP93 in podocyte actin cytoskeletal function and highlight potential pathways that may be linked to nuclear transport or regulation of gene expression.

Group: Casadei **Author:** Jillian Simon **Poster no: 6**

Oxidized PKARIα localizes to the endolysosome in cardiomyocytes where it regulates calcium release via two-pore channels

Jillian N. Simon*[1], Stefania Monterisi[2], Sandy Chu[1], Besarte Vrellaku[1], Oliver Lomas[1], Nadiia Rawlings[1], Margarida Ruas da Silva[3], Dominic Waithe[4], Gerard Marchal[1], Pawel Swietach[2], Manuela
Zaccolo[2], Phil Eaton[5], Barbara Casadei[1]
*presenting author, post-doctoral researcher
[1]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford. [2]Department of Physiology, Anatomy and Genetics, University of Oxford.
[3]Department of Pharmacology, University of Oxford. [4]Wolfson Imaging Centre, Weatherall Institute of Molecular Medicine, University of Oxford. [5]Cardiovascular Division, The Rayne Institute, King's College London.

Kinase oxidation is increasingly regarded as an important signalling mechanism through which changes in redox state alter cellular processes via phosphoregulation. Type-1 PKA (PKARIα) is redox sensitive and forms reversible disulfide bonds in response to oxidizing stimuli. Here we make use of a "redox dead" PKARIα mouse model (C16S) to investigate whether PKARIα oxidation impacts on kinase localization, activity, and downstream signalling.

Using FRET and FRAP-based imaging we demonstrate that disulphide bond formation enhances intracellular anchoring of PKARIα via A-kinase anchoring proteins (AKAP), without impacting their catalytic activation. Super-resolution STED microscopy reveals that this is associated with oxidant-dependent localization of PKARIα to the endolysosomal microdomain in cardiomyocytes, which is partly prevented by disruption of RIα-AKAP binding. In isolated cardiomyocytes, displacement of PKARIα from the endolysosomal microdomain leads to increased ryanodine receptor open probability and spontaneous Ca^{2+} oscillations from the sarcoplasmic reticulum which are prevented by endolysosomal Ca^{2+} depletion or by competitive inhibition of endosomal Ca^{2+}-releasing two-pore channels.
In summary, our findings have uncovered a novel pathway by which the intracellular redox state regulates calcium handling in cardiomyocytes, and demonstrates a role for PKARIα in stabilizing the endolysosome in response to oxidative stress. This regulatory mechanism may offer cardioprotection in conditions where endolysosomal calcium release is a significant contributor to injury, such as in ischemia-reperfusion injury.

Group: Channon **Author:** Gillian Douglas **Poster no: 7**

Risk Allele Of The Coronay Artery Disease Related Gene Jcad Is Associated With Attenuated Endothelial Cell Function

G Douglas, I Akoumianakis, E Drydale, A Antonopoulos, A Goel, H Watkins, C Antoniades and K, M. Channon*
BHF Centre of Research Excellence, Division of Cardiovascular Medicine, John Radcliffe Hospital, University of Oxford, UK,

Introduction: GWAS studies have consistently identified an association between coronary artery disease and a locus on chromosome 10 containing a single gene, *JCAD* (formally *KIAA1462*). However, as yet no study has investigated how *JCAD* genotype effects coronary artery disease burden and vascular function.

Methods: Patients undergoing elective cardiac surgery (n=443) were genotyped for the *JCAD* eQTL SNP rs2487928. Flow-mediated dilatations (FMD) and endothelium-independent vasodilatations (EID) of the brachial artery were assessed prior to surgery. *Ex vivo* vascular function was assessed in saphenous vein segments.

Results: The *JCAD* risk allele was associated with the presence of coronary artery disease β = 0.127 (standard error = 0.022), p = 0.004, and increased disease burden as assessed by the number of diseased vessels, (β = 0.136 (standard error = 0.066), p = 0.002. Carriers of the *JCAD* risk allele had significantly reduced FMD responses compared with carriers of the protective allele with no difference in responses to endothelial cell independent dilation to GTN. *Ex vivo* vascular function studies showed a similar trend with a significant decrease in sensitivity to bradykinin observed in carriers of risk allele with no difference observed in the response to the endothelium-independent dilator, sodium nitroprusside.

Conclusion
We show for the first time that carriers of the *JCAD* risk allele have increased diseased burden and attenuate vascular endothelial cell dependent vasomotor responses. Future studies investigating how loss of *JCAD* alters conduit and resistance artery function will be key to understand the mechanisms mediating these findings.

Group: Chapman **Author:** Peter Yeow **Poster no 8 & Talk**

Exploiting TRIM37-driven centrosome dysfunction to eliminate 17q23-amplified breast cancer cells

Genomic instability (GI) is a hallmark of cancer and plays a central role in breast cancer initiation and development. The success of Poly-ADP ribose polymerase inhibitors in the treatment of homologous recombination (HR)-deficient breast cancers exemplifies the utility of synthetic lethal drug-gene interactions in the treatment of GI-driven breast cancer. Given that HR-defects are present in only a subset of breast cancers, there is a need to identify additional GI-driver mechanisms, and targeted strategies to exploit these defects in cancer treatment. Here, we identify that centrosome-depletion induces synthetic lethality in cancer cells harbouring the 17q23 amplicon, a recurrent copy number aberration (CNA) that defines ~9% of all breast tumours and is associated with high GI. Specifically, small-molecule inhibition of Polo-like kinase 4 (PLK4) leads to centrosome depletion that triggers mitotic catastrophe in cells harbouring amplicon-directed overexpression of TRIM37. To explain this effect, we identify TRIM37 as a negative regulator of centrosomal pericentriolar material (PCM). In 17q23-amplified cells, elevated TRIM37 blocks the formation of non-centrosomal PCM foci, structures we show to be required for successful cell division in the absence of centrosomes. Lastly, we find TRIM37 overexpression causes GI by inducing cohesion fatigue, mitotic failure, and subsequent whole genome duplications. Collectively, these findings highlight TRIM37-dependent GI as a putative driver event in 17q23-amplified breast cancer and provide a rationale for centrosome-targeting therapeutics in their treatment.

Group: Choudhury **Author:** Naveen Akbar **Poster no: 9**

Endothelial Cell Derived Extracellular Vesicle Mediate Immune Cell Deployment from the Spleen and Transcriptional Programming Following Acute Myocardial Infarction
Naveed Akbar[1], Alastair Corbin[2], Adam Braithwaite[1], Eleanor Hogg[1], Mala Gunadasa-Rohling[3] Abhirup Banerjee[1], Charlotte Lee[1], Genevieve Melling[4], Laurienne Edgar[1], Rebecca

Dragovic[5], Dave Carter[4], Keith Channon[1], Paul Riley[3], Irina Udalova[2], Daniel Anthony[5], Robin Choudhury[1]

[1]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford. United Kingdom.
[2]Kennedy Institute of Rheumatology, University of Oxford, United Kingdom.
[3]Department of Physiology, Anatomy and Genetics, University of Oxford, United Kingdom.
4Department of Biological and Medical Sciences, Oxford Brookes University, United Kingdom.
[5]Department of Women's & Reproductive Health, University of Oxford, United Kingdom.
[6]Department of Pharmacology, University of Oxford, United Kingdom.

Background

We have previously shown that acute myocardial infarction (AMI) induces transcriptional activation of monocytes *en route* to the injured myocardium (Ruparelia et al. 2015), possibly through interactions involving plasma liberated endothelial cell (EC) derived extracellular vesicles (EVs), which are enriched for EC-vascular cell adhesion molecule-1 (VCAM-1) and EC-microRNA-126 (miRNA-126) (Akbar et al. 2017). Neutrophils are the first immune cells to arrive at sites of injury and mediate further damage to the ischemic myocardium. Here, we describe deployment of neutrophils from the spleen in AMI by EC-EVs. Moreover, we demonstrate that miRNA-126 mRNA targets are enriched within differential neutrophil transcriptomes following AMI and are associated with neutrophil cellular functions.

Methods

Patients presenting with AMI provided informed consent as part of the Oxford Acute Myocardial Infarction Study (OxAMI). Peripheral blood neutrophils were isolated at time of presentation with AMI and 1-month after AMI for RNA-sequencing. Plasma EVs were isolated at time of presentation and 6-months after AMI by differential ultra-centrifugation (2 hours at 120,000 x g) followed by washing and characterisation for: morphology using transmission electron microscopy (TEM), size and concentration profiling by Nanoparticle Tracking Analysis, EV markers (TSG101, ALIX, CD9, HSP70) by western blot and miRNAs by RT-qPCR. Mouse and human EC were used *in vitro to* derive EC-EVs under control conditions or after inflammatory stimulation with tumour necrosis factor-alpha (TNF-a) (10 ng/ml). EC-EVs were injected via tail vein into C57BL/6 wild-type mice and used to stimulate primary human peripheral blood neutrophils *in vitro*; EC-EV treated cells and tissues were analysed for changes in miRNA-126 mRNA targets.

Results

Patients presenting with AMI had differential activation of neutrophil transcriptomes and enrichment for gene sets involving TNF-α signalling via NF-kB ($p=0.001$), inflammatory responses ($p=0.001$) and interleukin-6 (IL-6) via JAK-STAT3 signalling ($p=0.001$). Patients presenting with AMI ($N=15$) had significantly more plasma EVs at time of injury vs. at a 6-month follow-up measurement (2.2-fold, $p=0.008$). The number of plasma EVs quantified at AMI presentation correlated significantly with the extent of ischemic injury quantified by T2-weighted MRI ($R=0.046$, $p=0.006$) and plasma neutrophils ($R=0.37$, $p=0.017$). Experimental AMI in wild-type mice induced a significant increase in peripheral blood neutrophils ($p<0.001$), and a corresponding reduction in splenic-neutrophils ($p<0.001$), suggesting splenic-neutrophil deployment.

We have previously shown that human plasma EVs are enriched for VCAM-1 and EC-associated miRNA-126 post-AMI (Akbar et al. 2017). Furthermore, miRNA-126 mRNA targets are significantly over-represented within neutrophil function-associated Gene Ontology terms including: degranulation ($p<0.001$), activation ($p<0.001$), chemotaxis ($p=0.008$) and migration ($p=0.008$). miRNA-126 mRNA targets were also significantly over-represented within differentially expressed genes of human neutrophil transcriptomes during AMI ($p=0.002$). Human and mouse EC release more EVs after inflammatory stimulation and show enrichment for miRNA-126. We found that exposure of primary human neutrophils to EC-EVs altered inflammatory gene expression for IL-6 ($p<0.05$), CCL7 ($p<0.001$) and CCL18 ($p<0.001$). EC-

EVs injected via the tail vein into wild type mice caused mobilisation of splenic-neutrophils to peripheral blood ($p<0.001$), possibly by down-regulating CXCL12/SDF-1 ($p=0.018$), a miRNA-126-mRNA target and neutrophil retention chemokine.

Conclusions
Neutrophil deployment from the spleen is a novel finding in acute injury and interactions with EC-EVs mediate their splenic liberation and transcriptional programming following AMI, *en route* to the injured myocardium. The splenic neutrophil reserve may be a novel therapeutic target in AMI to modulate the early inflammatory response before recruitment of monocytes to sites of injury.

Group: Church **Authors:** Ivy Xiong **Poster no: 10**

Inherited and somatic genetic variants in the p53 pathway interact to affect cancer risk and progression

Lingyun Xiong[1], Ping Zhang[1], Claire Palles[4], Isaac Kitchen-Smith[1], Siddhartha Kar[5], Paul Pharoah[5], Hannah Carter[6], Tim Maughan[2], Ian Tomlinson[3], Enric Domingo[2], David Church[3], Gareth Bond[1]

1. Ludwig Institute for Cancer Research, Nuffield Department of Clinical Medicine, University of Oxford, Oxford OX3 7DQ, UK; 2. Department of Oncology, University of Oxford, Oxford OX3 7DQ, UK; 3. Wellcome Trust Centre for Human Genetics, Nuffield Department of Clinical Medicine, University of Oxford, Oxford OX3 7BN, UK; 4. Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham B15 2TT, UK; 5. Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK; 6. Department of Medicine, University of California San Diego, La Jolla, CA 92093, USA;

Cancer is a complex disease affected by both inherited risk loci and somatic alterations. Genome-wide association studies (GWASs) have identified more than 3,000 inherited genetic variants associated with increased risk of cancer. However, our understanding of the complex genetic architecture of cancer susceptibility is far from complete, highlighted by the debate over the “missing heritability” in cancer. It has been shown that inherited genetic variants and somatic mutations converge in cancer driver genes and key signalling pathways to affect cancer development and progression, and emerging evidence supports the hypothesis that certain inherited genetic variants can influence the somatic cancer genomes. This interplay holds great promise in further personalizing cancer risk management and therapy development. Here, by integrating germline datasets relating to cancer susceptibility from multiple population-based case-control studies totalling more than 293,000 individuals with comprehensive tumour data capturing somatically-acquired genetic variations, we present evidence that single nucleotide polymorphisms (SNPs) and somatic mutations in the p53 tumour suppressor pathway can interact to influence cancer development, progression and response to treatment. Our experimental model also offers first-hand insight into the mechanisms how the inherited genome can influence the evolutionary trajectory of a tumour with focus on *TP53*, arguably the most important cancer gene involved in tumorigenesis. Together, the findings are instrumental that the inherited genome could hold much greater value than previously appreciated for cancer risk management as well as treatment prognosis in the clinics.

Group: Church **Authors:** Matthew Brown

Greater understanding of the molecular underpinnings of endometrial cancer has potential to improve its clinical management. The ubiquitin ligase component *FBXW7* regulates multiple cellular processes and is mutated in 10–30% of endometrial cancers. However, little is known about the role these mutations play in carcinogenesis. We modelled the cancer-associated *Fbxw7* R482Q mutation in the mouse uterus in isolation (i), and in combination with *Trp53* loss (ii) and *Pten* loss (iii). Uterine expression of $Fbxw7^{R482Q/+}$ did not cause endometrial neoplasia alone or in combination with *Trp53* loss; however, $Pten^{\Delta/\Delta}Fbxw7^{R482Q/+}$ females developed aggressive endometrial adenocarcinoma with significantly earlier onset than $Pten^{\Delta/\Delta}Fbxw7^{+/+}$ littermates (P<0.001). Comparison of $Pten^{\Delta/\Delta}Fbxw7^{R482Q/+}$ and $Pten^{\Delta/\Delta}$ uteri demonstrated several dysregulated pathways, including PI3K-mTOR, Hippo and Wnt; *in silico* prediction identified mTOR, YAP1/WWTR1 and LEF1 as putative Fbxw7 substrates responsible. Ongoing work seeks to determine the underpinning functional mechanisms, and its interaction with other endometrial cancer biomarkers.

Group: Church **Authors:** Yun Feng

Mutation signatures are the hallmarks of mutagenic processes in cancer that can provide clues about the biochemical mechanisms by which DNA is altered in cancer. The extraction of such signatures from next generation sequencing data has traditionally been formulated as an unsupervised learning problem and solved using non-negative matrix factorization. We present an entirely novel approach based on convolutional filtering, inspired by technologies used in computer vision and image processing for genomic data analysis. We show that our approach (convSig) has state-of-the-art performance compared to standard methods but also generalizes to allow consideration of longer sequence contexts using deep layering of convolutional networks providing a tool that could potentially reveal the impact of high-level genome structure on mutational density.

Group: Cornall - CCMP **Authors:** Rose Hodgson **Poster no: 11**

How does loss of Prolidase lead to autoimmune disease?

Rose Hodgson, Tanya Cheetham, Eleanor Cawthorne, Lucie Abeler-Dorner, Adrian Hayday, Katherine Bull, Richard Cornall

Introduction

Rare genetic variants provide a tractable way to discover novel genes and regulatory pathways involved in maintaining immunological tolerance, information that is crucial in advancing our understanding and ultimately treating autoimmune diseases. Individuals with loss-of-function mutations in the prolidase (*PEPD*) gene present with a highly variable spectrum of clinical manifestations including lower limb skin ulcerations, recurrent respiratory infections and, in some patients, systemic lupus erythematous (SLE) characterised by auto-antibodies against nuclear proteins. These phenotypes are concurrent with a loss of immune tolerance. The *PEPD* gene encodes prolidase, the metallopeptidase uniquely responsible for the hydrolysis of dipeptides containing proline or hydroxylproline in the C terminal position, which is required in the breakdown of collagen and other proline-rich substrates. Understanding how the loss of prolidase can lead to a breakdown of immune tolerance will

provide further insight into how biochemical pathways are dysregulated in autoimmune diseases.

Results

A high-throughput screen for autoimmune phenotypes at the Wellcome Trust Sanger Institute showed that *Pepd*$^{-/-}$ mice have higher incidence of class-switched anti-nuclear antibodies than wildtype counterparts. *Pepd*$^{-/-}$ mature B cells are abnormal in expressing elevated surface IgM levels, but have no accompanying defect in B cell tolerance at the pre-immune stage, as demonstrated using the hen egg lysozyme transgenic mouse model. In contrast, *Pepd*$^{-/-}$ $CD4^+$ and $CD8^+$ T cells have an activated effector phenotype, which is enhanced by immunisation. These findings suggest that the primary defect may be dysregulation of T cell function, with or without antigen stimulation. Multi-dimensional mass cytometry has been used to further classify the affected T cell sub-populations and paves the way for a single cell transcriptomic approach to pathway discovery. The use of T cell transgenic models will allow us to distinguish between the antigen-dependent and independent effects of PEPD on immune function and explore biochemical effects on groups of clonally distinct cells ex vivo.

Group: Davies **Authors:** Phalguni Rath **Poster no: 12**

Manipulation of gene expression using CRISPR-activation and CRISPR-inhibition systems in human iPS cells

Phalguni Rath, Adrià Dangla Valls, Marta Pérez Alcántara, Mark McCarthy, Noel Buckley and Ben Davies

In addition to achieving site-specific deletions and insertion, CRISPR systems have been adapted to allow experimental control of gene expression. Catalytically inactivated Cas9 (dCas9) fused to transcriptional activation / inhibition machinery can be recruited to the endogenous promoter regions of target genes, allowing up (CRISPRa) and down (CRISPRi) regulation of gene expression. Such systems can be used to perturb gene expression in more subtle and controlled ways than conventional techniques and might provide a better means of modelling disease-associated gene expression changes and their functional consequences.

We have engineered a control human induced pluripotent stem (iPS) cell line with the CRISPRa and CRISPRi machinery, by using Bxb1 integrase mediated cassette exchange to insert expression cassettes for the dCas9 transcriptional activators/inhibitors at the AAVS1 safe-habour locus. We have optimized the promoter used to drive the expression of the machinery to ensure consistent and reliable expression of the CRISPRa/i machinery in both undifferentiated and differentiated cell types.

Once activated by a Cre-recombinase switch, target gene expression within these stem cells can be manipulated by sgRNA transfection using optimized RNA transfection reagents. We demonstrate up- and down-regulation of several different target genes, at both the mRNA and the protein level by simple delivery of these small RNAs as the inducing agent and explore the kinetics of the gene expression change. We are also exploring the functionality of these systems for achieving gene expression changes in differentiated cells such as neurons, cardiomyocytes, macrophages and pancreatic endocrine precursors.

The engineered cell lines allow an alternative means of probing gene function using the iPS cell model system and could facilitate high-throughput screening approaches with sgRNA libraries.

Group: Davison - CCMP **Authors:** Marsha Wallace/Lucy Davison
Poster no: 13

Single-cell RNA sequencing reveals transcriptomic profiles of pancreatic islet and infiltrating immune cells before the onset of type 1 diabetes
Wallace MD, Knudsen JG, O'Callaghan CA, Davison LJ

Background: Type 1 diabetes is characterised by autoimmune pancreatic beta cell destruction resulting in insulin deficiency and hyperglycaemia. By the time of diagnosis, the autoimmune response has already destroyed many of the insulin-producing pancreatic beta cells, making it challenging to investigate the early phase of disease. This study used novel single cell transcriptomics technology in a spontaneous model of type 1 diabetes, the non-obese diabetic (NOD) mouse, to investigate the transcriptional profile of the wide variety of cells present in pancreatic islets prior to the onset of hyperglycaemia.

Methods: Pancreatic islets were purified from 10-12 week-old B6 mice (n=2) and from 10-12-week-old female non-obese NOD mice (n=10) prior to the onset of hyperglycaemia. After hand-picking and counting islets, a single cell suspension was prepared from islet tissue. Bead-based 10X Genomics technology was used to prepare a single-cell RNA-Seq islet library from each mouse, capturing between 500 and 1000 cells per mouse. High-throughput sequencing was undertaken using a HiSeq4000 machine at coverage of approximately 100k reads per cell.

Results: The transcriptome of multiple pancreatic and infiltrating immune cell populations during the early stages of beta cell destruction was established, with over 2,000 transcripts per cell being detected. Cells were grouped into distinct clusters based on their transcriptome, using a variety of single cell analysis software including CellRanger. In NOD mouse islets, the transcriptome of approximately 1 in 3 cells was consistent with an immunological origin, including activated CD8+ lymphocytes, macrophages, dendritic cells and plasma cells. We also identified a subset of atypical pancreatic beta cells expressing immune response genes. This work, which allowed analysis of both the immune response and the beta cell response to the immune attack, demonstrates the potential of single cell transcriptomics to reveal important new insights into the nature of early T1D.

Group: Calli Dendrou **Authors:** Melissa Grant-Peters & Victor Yeung
Poster no 14 & Talk

Investigating pathophysiological mechanisms shared across immune-mediated diseases
Victor Yeung, Melissa Grant-Peters, Calliope Dendrou

Immune-mediated diseases (IMDs) affect ~10% of the population worldwide and they have no cure. Currently available immunomodulatory drugs can alleviate symptoms, but can also have variable efficacy and can lead to severe adverse effects (Dendrou *et al*. 2015 *Nat Rev Immunol*; Haghikia *et al*. 2017 *Lancet Neurol*). A better understanding of pathophysiological mechanisms shared between IMDs and other clinical conditions can help to identify targets amenable to drug repositioning strategies, and to provide insights relating to likely drug efficacy and side effects for target prioritisation (Dendrou *et al.* 2016 *Sci Transl Med*; Dendrou *et al*. 2018 *Nat Rev Immunol*; Yeung and Dendrou 2019 *Annu Rev Genomics Hum Genet*). We are studying disease pathways common to multiple IMDs by (i) investigating the genetic

architecture of these conditions and determining the functional consequences of shared genetic variants, and by (ii) profiling immune cell types and activation states at the site of pathology of different IMDs through single-cell and spatially resolved transcriptomics approaches.

The UK Biobank provides an invaluable resource for interrogating the genetic contribution to over 16,000 clinical phenotypes in half a million individuals, and the development of new methods for analysing genetic associations with these phenotypes are revealing previously unappreciated genetic sharing across IMDs (Cortes *et al*. 2017 *Nat Genet*; Cortes *et al*. 2018 *bioRxiv*). For example, genetic variants in the tyrosine kinase 2 (TYK2) gene protect against at least 20 different IMDs – indicating TYK2 as a prime target for drug repositioning approaches. We are investigating the molecular mode of action of these variants and their consequences for cytokine-induced gene expression at the single-cell level to determine how TYK2 acts as an immunological hub for IMDs, and we are comparing the impact of these variants to novel allosteric TYK2 inhibitors in collaboration with Nimbus Therapeutics.

As part of the Human Cell Atlas we are involved in the single-cell profiling of immune cells obtained from biopsies from patients with gut-associated IMDs, and we are completing a study of liver immune cells from patients with primary sclerosing cholangitis with gut and liver pathology. We have analysed over 35,000 liver immune cells obtained by minimally invasive fine-needle aspiration and we can identify liver-specific cell types and disease-associated signatures of cell activation. In addition, we are using single-nuclear and spatially resolved transcriptomics methods to investigate gene signatures in lesional, peri-lesional, and normal-appearing central nervous system tissue from different neuroimmunological conditions, in conjunction with unbiased microspectroscopic definition of lesional and peri-lesional areas through high-resolution biochemical profiling at the B22 Beamline at Diamond.

Group: Donnelly **Authors:** Anjali Hinch **Poster no 15 & Talk**

Un ménage à trois: Novel insights into meiotic recombination revealed by a trio of single-stranded DNA binding proteins

Cells in diploid organisms have pairs of homologous chromosomes, which act independently of each other during most cellular functions. An exception occurs in meiosis, the process in which eggs and sperm are created. During meiosis each pair of chromosomes must locate each other in the nucleus and physically exchange genetic material via recombination and crossing over. Despite their critical importance, many aspects of the search for and engagement between homologous chromosomes remain mysterious.

Many species, including humans, have three key proteins, DMC1, RAD51, and RPA, each of which bind to single-stranded DNA and play essential, albeit poorly understood, roles in these processes. DMC1 is expressed only in meiosis, while RAD51 and RPA are essential to life and are also required in the response to DNA damage in somatic cells. Since meiosis initiates with programmed DNA double-strand breaks, understanding the roles and behaviour of these proteins sheds light not only on meiosis, but also on the universal mechanisms involved in DNA repair.
To understand these fundamental processes, we mapped the binding positions of RAD51, RPA, and DMC1 genome-wide in meiotic tissue from male mice. These are the first fine-scale maps in any meiotic system for these proteins, and reveal their localisation and timing with high sensitivity, specificity and spatial resolution. These maps answer long-standing questions

about the functions and dynamics of these proteins in meiosis, revealing key insights into the various stages of recombination: the formation and processing of DNA double-strand breaks, the search for homology, the exchange of strands between DNA duplexes, and the formation of crossovers. For example, we show, for the first time, that DMC1 preferentially occupies the cut end of single-stranded DNA filaments involved in homology search, while RAD51 occupies the double-stranded end, and that RPA binds single-stranded DNA on the chromosome that is used as template for repair, likely as part of a D-loop recombination intermediate.

Group: Farrall **Authors:** Anuj Goel **Poster no: 16**

Mining data to fine-map coronary artery disease associated loci

Goel A*, Grace C, Watkins H, Farrall M
Wellcome Centre for Human Genetics, University of Oxford, Oxford. Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford.

Rationale

With the availability of improved imputation reference panels and global collaboration (CARDIoGRAMplusC4D and other studies), we have even more power to detect genetic signals and this has given a major boost to coronary artery disease (CAD) genetics. However, genome wide association studies (GWAS) in isolation cannot identify causal genes or variants, which is its major limitation.

Methodology

The Roadmap Epigenomics project has catalogued several important functional elements in over one hundred human tissues and cell types. We have selected key functional chromatin marks from 10 CAD relevant cell types and implemented a publically available software (FGWAS) to integrate GWAS summary information (derived from a case-control sample size of over 1M) with these key chromatin marks to scan for plausible causal functional variants that define a 95% credible set of SNPs and INDELs.

Results

We find chromatin enrichment in 116 genome-wide significant CAD associated loci. Preliminary finding in a positive control locus (eg. *ABCG8*) suggests there are three functionally plausible variants in liver, with no such enrichment in any other tissue of interest, finding consistent with expectation.

Conclusions

Using this software, we are able to zoom into individual GWAS loci to reveal plausible functional variants in CAD relevant tissues/cell types, providing a strong focus for future downstream lab-based functional experiments.

Groups: Gilbert **Authors:** Robert Gilbert **Poster no: 17**

Structure and mechanism of bactericidal mammalian perforin-2, an ancient agent of innate immunity

Xiulian Yu, Tao Ni, Robert Gilbert

Perforin-2 (MPEG1) is thought to enable the killing of invading microbes engulfed by macrophages and other phagocytes, forming pores in their membranes. Loss of Perforin-2 renders individual phagocytes and whole organisms significantly more susceptible to bacterial pathogens than their wild-types. I will describe the mechanism of Perforin-2 activation and activity determined using a combination atomic structures of pre-pore and pore assemblies from cryo-EM and X-ray crystallography, high-speed atomic-force microscopy, and functional assays. Perforin-2 binds to negatively-charged lipid membranes such that its pore-forming domain points directly away from the membrane surface in the pre-pore state. Acidification triggers pore formation, via a 180° conformational change that results in the orientation of its membrane-inserting regions towards the target bilayer. This novel and unexpected mechanism prevents premature bactericidal attack, and may have played a key role in the evolution of all perforin-family proteins.

Group: Gloyn **Authors:** Claire Duff **Poster no: 18**

Modelling defects in human pancreas development due to loss of *HNF1A* in Human Induced Pluripotent Stem cells using Genome-Editing

Claire E. Duff[1], Nicole A.J. Krentz[2], Fernando Abaitua[2], Sameena Nawaz[1], Benoit Hastoy[1], Shailesh K. Gupta[3], Christian L. F. Honoré[4] & Anna L. Gloyn[1,2,5]

1. Oxford Centre for Diabetes, Endocrinology and Metabolism (OCDEM), University of Oxford, Oxford, UK, 2. Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK, 3. Discovery Biology, Discovery Sciences IMED Biotech Unit, AstraZeneca, Gothenburg, Sweden, 4. Novo Nordisk Department of Stem Cell Biology, Novo Nordisk A/S, Maaloev, Denmark, 5. Oxford NIHR Biomedical Research Centre, Churchill Hospital, Oxford, UK

Introduction: Understanding the impact of disease-causing mutations on human development usually relies on the use of model organisms which often do not faithfully recapitulate human phenotypes. Heterozygous loss of function mutations in *HNF1A,* encoding the transcription factor hepatocyte nuclear factor 1 alpha (HNF-1a) cause monogenic diabetes. HNF1-a is involved in both beta cell development and mature cell function. We used CRISPR-Cas9 genome editing in a well characterised human induced pluripotent stem (IPS) cell line to generate an *HNF1A* knockout (KO) line and characterised the impact of *HNF1A* loss on pancreatic beta cell development.

Methods: A NKX6.1-GFP IPS reporter line was transfected with CRISPR-Cas9 and a synthetic guide RNA with the aim of inducing double-strand breaks in the *HNF1A* locus, leading to frameshift mutations, premature stop codons and non-sense mediated decay. *HNF1A* KO cell lines were differentiated alongside two control lines, a NKX6.1-GFP reporter cell line and the parental line (SB), with each cell line differentiated (4 independent differentiations). Samples were collected at stage 1 (definitive endoderm) prior to *HNF1A* expression, stage 4 (pancreatic endoderm) at the beginning of *HNF1A* expression and stage 7 (beta-like cells) the end point of differentiation, and analysed using flow cytometry and RT-qPCR.

Results: Three compound heterozygous KO *HNF1A* clonal lines were successfully generated and frameshift mutations verified with genotyping. One clone was taken forward for differentiation after passing quality control for normal karyotype number and pluripotency markers. There were no gross differences in differentiation efficiency assessed by expression of PDX1, NKX6.1, glucagon or insulin using flow cytometry between the KO-cell line and the control lines. However, *HNF1A* transcript levels were significantly decreased in the KO

compared to the control lines ($p<0.005$). Similarly, there was a decrease in expression of both insulin and glucagon in the knockout cell line compared to controls ($p<0.05$).

Conclusions: We have generated compound heterozygous *HNF1A* knockout human IPS cell lines and have differentiated them along the endocrine lineage. Our results thus far indicate that *HNF1A* loss reduces expression of key hormones (glucagon and insulin) which are characteristic of mature endocrine function, suggesting that loss of *HNF1A* causes attenuated maturation of iSPCs towards beta-like cells.

Group: Green **Authors:** Cath Green/Lihao Wang **Poster no: 19**

Lihao Wang and Catherine M. Green, Chromosome Dynamics Group, Wellcome Centre for Human Genetics

Proliferating cell nuclear antigen (PCNA) is an essential cofactor for DNA replication and repair, recruiting multiple proteins required for chromosomal DNA synthesis to their sites of action. We identified an inherited syndrome, PCNA-associated DNA repair disorder (PARD) with clinical features including short stature, hearing loss, premature aging, telangiectasia, neurodegeneration, and photosensitivity, which results from a homozygous missense (p.Ser228Ile) alteration of PCNA. While the p.Ser228Ile alteration has no clear effect on protein levels or DNA replication, patient-derived cells exhibited marked abnormalities in response to UV irradiation, displaying substantial reductions in both UV survival and RNA synthesis recovery. These cells are also sensitive to the PCNA inhibitors T3 and T2AA, but not to ionising radiation. We determined the structure of the mutated PCNA protein and showed a dramatic but indirect effect on the conformation of the interdomain connecting loop. This profoundly altered PCNA's interaction with many of its interaction partners, including Fen-1, DNA Ligase 1, Cdt1, DNMT1, PolD3p66 and PolD4p12. In contrast p21 largely retains the ability to bind $PCNA^{S228I}$. We generated a $PCNA^{S228I}$ knock-in cell line in hTERT RPE-1 cells using CRISPR-Cas9. These $PCNA^{S228I}$ cells were more sensitive to treatment with the PARP inhibitor olaparib than wild-type controls, results that were reproduced in lymphoblastoid cells derived from PARD-affected individuals. By COMET assay as well as immunostaining of γH2AX, we also showed that olaparib also induced more DNA damage in cells with the $PCNA^{S228I}$ alteration, and FACS analysis showed perturbed S phase progression and checkpoint activation after PARP inhibition. By using an inhibitor of PARG, we detect high levels of PAR chains being formed at replication forks in $PCNA^{S228I}$ cells. Thus we propose that the PCNA S228I mutation, that causes PARD in humans, results in abnormal interactions between PCNA and key enzymes (included Fen1 and Ligase I) needed for normal Okazaki fragment maturation at replication forks. The resultant persistent single stranded breaks are processed via a back-up, PARP-dependent mechanism to ensure completion of lagging strand synthesis. These data have implications both for the etiology of PCNA-associated repair disorder as well as suggesting novel circumstances in which PARP-inhibition might be of use in the cancer clinic.

Group: Green /Chromosome Dynamics Core **Authors:** Cath Green/Daniela Moralli **Poster no: 20**

The Chromosome Dynamics Core is a facility employing molecular cytogenetic and fluorescence microscopy techniques in support of different research projects within the WCHG, the University of Oxford and across the UK.

Current techniques offered include:

- Karyotype analysis: chromosome counting, chromosome banding and M-FISH.
- Chromosome instability analysis: metaphase/anaphase segregation; fragile sites identification; double strand breaks (DSB) and radial formation. Sister chromatid exchange (SCE).
- Locus specific analysis: FISH and immunoFISH on metaphasic chromosomes, interphasic cells, chromatin fibres, FFPE sections., to evaluate copy number, nuclear localisation, interaction with specific cellular structures.
- Meiotic cell analysis: prophase characterisation of chromosome synapsis, and specific meiotic protein distribution.
- Immuno-hystochemistry, analysis of protein distribution in fixed cells.

We are happy to consult with any Centre members regarding new projects or techniques that we might be able to assist with.

Groups: Grimes **Authors:** Loic Carrique **Poster no: 21**

Influenza A virus RNA polymerase structures provide insights into viral genome replication

Haitian Fan[1,8], Alexander P Walker[1,8], Loïc Carrique[2,8], Jeremy R Keown[2,8], Itziar Serna Martin[1,2,6], Dimple Karia[2], Jane Sharps[1], Narin Hengrung[1,2,7], Els Pardon[3], Jan Steyaert[4], Jonathan M Grimes[2,5,9*] and Ervin Fodor[1,9*]

[1]Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, United Kingdom
[2]Division of Structural Biology, Henry Wellcome Building for Genomic Medicine, University of Oxford, Oxford OX3 7BN, United Kingdom
[3]VIB-VUB Center for Structural Biology, VIB, 1050 Brussels, Belgium
[4]Structural Biology Brussels, Vrije Universiteit Brussel, 1050 Brussels, Belgium
[5]Diamond Light Source Ltd, Harwell Science & Innovation Campus, Didcot OX11 0DE, United Kingdom
[6]Present address: Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Department of Chemistry, Faculty of Science, Utrecht University, Padualaan 8, 3584 CH, Utrecht, The Netherlands
[7]Present address: Francis Crick Institute, 1 Midland Road, Kings Cross, London NW1 1AT, United Kingdom
[8]These authors contributed equally: Haitian Fan, Alexander Walker, Loïc Carrique, Jeremy R Keown
[9]These authors jointly supervised this work: Jonathan M Grimes, Ervin Fodor

Influenza A viruses are responsible for seasonal epidemics, and pandemics can arise from novel zoonotic influenza A viruses transmitting to humans. Influenza A viruses contain a segmented negative sense RNA genome that is transcribed and replicated by the viral RNA-dependent RNA polymerase. We use X-ray crystallography and cryo-electron microscopy (cryo-EM) to obtain the structures of polymerases from human (H3N2) and avian (H5N1) influenza viruses. In combination to cell-based and *in vitro* assays we could demonstrate that

polymerase dimerisation is required for initiation of vRNA synthesis during viral genome replication. This study provides the first high-resolution structures of medically relevant influenza A virus polymerases and offers novel insights into the replication mechanisms of the viral RNA genome. Furthermore, it identifies novel sites of the influenza A virus polymerase that could be targeted for antiviral drug development.

Groups: Holmes - STOP-HCV **Authors:** Azim Ansari **Poster no: 22**

Interferon lambda 4 impacts on the genetic diversity of hepatitis C virus

M Azim Ansari[1], Elihu Aranday-Cortes[2], Camilla LC Ip[1], Ana da Silva Filipe[2], Lau Siu Hin[2], Connor G G Bamford[2], David Bonsall[3], Amy Trebes[1], Paolo Piazza[1], Vattipally Sreenu[2], Vanessa M Cowton[2], STOP-HCV Consortium, Emma Hudson[3], Rory Bowden[1], Arvind H Patel[2], Graham R Foster[4], William L Irving[5], Kosh Agarwal[6], Emma C Thomson[2], Peter Simmonds[3], Paul Klenerman[3], Chris Holmes[7], Eleanor Barnes[3], Chris CA Spencer[1], John McLauchlan[2], Vincent Pedergnana[1,8]

[1] Wellcome Centre Human Genetics, University of Oxford, Roosevelt Drive, Oxford, OX3 7BN, UK
[2] MRC-University of Glasgow Centre for Virus Research, Sir Michael Stoker Building, 464 Bearsden Road, Glasgow, G61 1QH, UK
[3] Nuffield Department of Medicine and the Oxford NIHR BRC, University of Oxford, Oxford, OX1 3SY, UK
[4] Blizard Institute, Queen Mary University, London, E1 2AT, UK
[5] National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and University of Nottingham, Nottingham, NG7 2RD, UK
[6] Institute of Liver Studies, King's College Hospital, London, United Kingdom
[7] Department of Statistics, University of Oxford, OX1 3LB
[8] Laboratoire MIVEGEC (UMR CNRS 5290, IRD, UM), 34000 Montpellier, France

Hepatitis C virus (HCV) is a highly variable pathogen that frequently establishes chronic infection. This genetic variability is affected by the adaptive immune response but the contribution of other host factors is unclear. Here, we examined the role played by interferon lambda-4 (IFN-λ4) on HCV diversity; IFN-λ4 plays a crucial role in spontaneous clearance or establishment of chronicity following acute infection. We performed viral genome-wide association studies using human and viral data from 485 patients of white ancestry infected with HCV genotype 3a. We demonstrate that combinations of host genetic variants, which determine IFN-λ4 protein production and activity, influence amino acid variation across the viral polyprotein - not restricted to specific viral proteins or HLA restricted epitopes - and modulate viral load. We also observed an association with viral di-nucleotide proportions. These results support a direct role for IFN-λ4 in exerting selective pressure across the viral genome, possibly by a novel mechanism.

Group: Jones **Authors:** Yuguang Zhao **Poster no: 23**

CACHD1, a Novel Wnt Signalling Player

Yuguang Zhao§, Gareth Powell#, Jingshan Ren§, Ana Faro#, Stephen W. Wilson#, E. Yvonne Jones§*

§ Division of Structural Biology, Oxford University, UK
# Department of Cell and Developmental Biology, University College London, UK

CACHD1 (VWFA and cache domain-containing protein 1) is evolutionarily conserved from insects to humans. It shares limited sequence homology to the $\alpha_2\delta1$ regulatory subunit of the neuronal calcium channel ($Ca_V2.2$) and can increase N-type calcium currents. Loss of function of CACHD1 in Zebrafish causes bilateral symmetry in the dorsal habenula. By using high throughput screen of reverse transfected expression cDNAs array, the FZD7 was identified as a novel ligand of CACHD1. The surface plasmon resonance (SPR) analysis shows CACHD1 interact with both WNT receptor FZDs and co-receptor LRP6 at high affinity. Crystal structure of CACHD1_FZD5_LRP6 complex reveals that CACHD1 shields the FZD WNT lipid binding site and blocks the proposed LRP6 WNT3a binding sites. The funding sheds new light into WNT calcium regulation pathways.

Group: Jones **Authors:** Vitul Jain **Poster no: 24**

The role of neuropilins in determining the outcome of the semaphorin-plexin signalling

Vitul Jain, Daniel Rozbesky, Karl Harlos and E. Yvonne Jones.

Division of Structural Biology, Wellcome Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, United Kingdom

The development of either the nervous or the immune system requires the establishment of cellular networks with the highest level of accuracy through a complex and changing environment. Members of several protein families act as guidance cues that can trigger localised changes in the cytoskeleton by binding of protein-ligand to the cell-surface receptor. A major cell-guidance system in mammalian organisms comprises of six classes of Semaphorin ligands, four classes of Plexin receptors and two classes of co-receptor Neuropilins. Only class 3 Semaphorins (Sema3A) depends on either Neuropilin (Nrp1 or Nrp2) co-receptors that associate with PlexinAs (PlxnAs) to generate a holoreceptor. Furthermore, recent reports suggest a class 4 Semaphorin (Sema4A) can initiate signalling in T-regulatory (Treg) cells using only Nrp1 as a receptor. This brings the tantalising hypothesis that Neuropilins might also function as receptors, next to their well-established canonical role as co-receptors. Hence, a complete structure-function insight into the differential mechanisms of Nrp1 interactions with Sema3 vs. Sema4 and with different plexins is an important unresolved question.

To this we expressed, purified, crystallized and solved the crystal structure of human Sema4A with the sema and psi domain. The structural superposition with Sema4D yielded striking differences between the two with regards to plexin binding interface and the homo-dimeric interface. These structural differences are being used to study the various Sema4A mutants and their altered interactions. Next, we performed binding lifetime measurements using FRET-FLIM with fluorescent proteins mRuby2 (red fluorescence) and mClover (green fluorescence) tagged Nrp1 on the membrane of COS-7 cells incubated with and without purified hSema4A-Fc (dimeric Sema4A). The lifetime data shows an increased clustering of the Nrp1 on the surface upon Sema4A binding indicating interaction between Sema4A and Nrp1. To further corroborate Sema4A interaction with Nrp1 and interactions with PlexinB1 and PlexinB2, Surface Plasmon Resonance measurements were performed using Biacore. The SPR data reveal strong interactions for hSema4A-Fc with hNrp1, hPlexinB1 and hPlexinB2 with binding constants of 12 µM, 7 µM and 1.7 µM respectively. Taken together, our data forms a solid

platform to build upon and understand in depth the versatile nature of the Neuropilins.

Group: Jones **Authors:** Daniel Rozbesky

Structural basis of semaphorin-plexin cis interaction

Daniel Rozbesky, Dimple Karia, Luis Alvarez, Karl Harlos, Sergi Padilla-Parra, Yvonne Jones

Division of Structural Biology, Wellcome Centre for Human Genetics, University of Oxford, Oxford, United Kingdom

One of the most fascinating processes during the nervous system development is the navigation of axons and dendrites towards their synaptic partner with an impressive level of accuracy. Semaphorins and plexins represent one of the major cell guidance factors that can trigger localised changes in the cytoskeleton of the growing axon. Semaphorins have been shown to form homodimers and upon the interaction with their plexin receptors, the semaphorin homodimer brings two plexins together to form a symmetrical complex, which is a prerequisite for cellular plexin signalling. We found that unlike all Drosophila semaphorins, which are disulfide-linked dimers, Sema1b is surprisingly a monomer due to a natural mutation C254F. Biophysical and cellular assays reveal that monomeric Sema1b binds PlexA but fails to mediate PlexA dimerization. We further report two different crystal structures of the semaphorin-binding region of PlexA with the ectodomain of Sema1b, plus the unliganded structure of Sema1b. The crystal structures along with biophysical assays show that monomeric Sema1b binds PlexA with two independent interaction modes. One mode is similar to the head-to-head interaction described before, while the other is formed by side-on interaction between the sema domains of PlexA and Sema1b. Furthermore, the latter interface indicates an interaction between Sema1b and PlexA is possible in cis, which we confirmed using FRET-FLIM in live cells. We further show that the full ectodomain of PlexA forms a ring-like structure in which the ring is closed by an interaction between the sema domain and IPT5 domain. Intriguingly, the surface of the PlexA sema domain that interacts with the IPT5 domain in the ectodomain ring structure is similar to that used for the side-on interaction in the complex between PlexA and Sema1b, suggesting a mechanism for opening the plexin ring by interaction with monomeric semaphorin in cis.

Group: Jones **Authors:** Luca Vecchia **Talk**

Structural basis for glypicans interaction with Wnt

Luca Vecchia, E. Yvonne Jones

The Wnt pathway is a cellular signalling route highly conserved in metazoans evolution, and pivotal during embryonic development, where it regulates tissue organization, patterning and organ growth; dysregulation of its control during adult life leads to multiple degenerative and proliferative abnormalities, including several types of cancers.

The Wnt proteins, which starts the signalling cascade engaging their membrane receptors, are lipid modified at a conserved serine position. This modification is essential for the morphogen biological activity; however, it also reduces protein solubility and hampers free diffusion in the extracellular space. Several models, including interaction with lipoprotein particles and lipocalin-like carrier proteins, have been therefore proposed over the years to explain Wnt spreading.

Unpublished genetic and biochemical data from our collaborators, the Vincent group at Francis Crick Institute (London), suggested a crucial role in Wnt trafficking for an additional family of proteins, GPI-anchored heparan sulfate proteoglycans or glypicans.

Here, we present the structure of the *Drosophila* glypican Dally-like protein (Dlp) solved at 2.21 Å resolution in complex with a palmitoylated human Wnt7a peptide, used as a surrogate for the Wnt protein. Upon lipid engagement, the structure undergoes a conformational rearrangement opening a hereto unsuspected binding cavity, where the palmitoleate can be accommodated and shielded from the aqueous environment. In combination with biophysical assays, we furthermore show that the interaction is mainly driven by the lipid moiety of the peptide, and that point mutations of the binding pocket significantly affect binding.
The identification of a subset of glypicans as novel interactors of Wnt, shaping its gradient and range of cell-to-cell signaling by sequestering its lipid portion, and the dissection of the molecular basis of binding, represent an important advance in our understanding of the modalities of Wnt trafficking in the extracellular space.

Group: Jones **Authors:** Tomas Malinauskas

Repulsive Guidance Molecules lock Growth Differentiation Factor 5 in an inhibitory complex

Tomas Malinauskas[1,*], Tina V. Peer[2], Alexander Kotzsch[2], Benjamin Bishop[1], Thomas D. Mueller[2,*], Christian Siebold[1,*]

[1]Division of Structural Biology, Wellcome Centre for Human Genetics, University of Oxford, Roosevelt Drive, OX3 7BN, Oxford, United Kingdom.
[2]Department of Molecular Plant Physiology and Biophysics, Julius-von-Sachs Institute, University of Würzburg, Julius-von-Sachs-Platz 2, 97082, Würzburg, Germany.

ABSTRACT
Repulsive Guidance Molecules (RGMs) are cell surface proteins that control neuron regeneration, migration, and iron homeostasis. RGMs signal via two fundamental pathways: the Neogenin (NEO1) and the Bone Morphogenetic Protein (BMP)/Growth Differentiation Factor (GDF). We determined structures of all RGMs in complex with GDF5, and ternary complexes of GDF5 with its type-1 and -2 receptor ectodomains, as well as with RGMB and NEO1. Our structural data, biophysical and cellular analyses illuminate the mechanisms of GDF5 signal activation. Surprisingly, we show that RGMs inhibit GDF5 signaling by competing with its type-1 receptor, which is in contrast to RGM-mediated enhancement of signaling by other BMPs. In the NEO1–RGMB–GDF5 complex, RGMB physically bridges NEO1 and GDF5, suggesting cross-talk between GDF5 and NEO1 signaling pathways. Our structural and functional studies provide a molecular map of the NEO1–RGM–BMP/GDF signaling network.

Group: Knight **Authors:** Yuxin Mi **Poster no: 25**

The sepsis plasma proteome: patient stratification and biomarker discovery

Yuxin Mi[1], Katie L Burnham[1], Raphael A Heilig[2], Marie-Laetitia Thézénas[2], Jonathan Webber[3], Paula Hutton[4], Eduardo M Svoren[5], Charles J Hinds[5], Benedikt Kessler[2], Roman Fischer[2], Julian C Knight[1]

[1]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK
[2]Target Discovery Institute, University of Oxford, Oxford, UK
[3]Kennedy Institute, University of Oxford, Oxford, UK
[4]Adult Intensive Care Unit, John Radcliffe Hospital, Oxford, UK
[5]William Harvey Research Institute, Bart's and The London School of Medicine, Queen Mary University, London, UK

Introduction:
Improvement in sepsis treatment has long been hampered by considerable disease heterogeneity. We have previously identified two disease endotypes from transcriptomics that differ in mortality rate and response to treatment. The aim of this study was to characterise variation in the individual host response at the proteome level, which could provide an opportunity to develop targeted therapies.

Methods:
We used mass spectrometry and Luminex assays to profile both the plasma proteome and 65 key cytokines in 144 samples from 96 sepsis patients admitted to intensive care, and 42 samples taken from 21 non-septic patients before and after cardiac surgery.

Results:
After quality control, 1123 proteins were quantified in 144 samples from patients with sepsis due to community acquired pneumonia. Unsupervised hierarchical clustering with cross-validation identified two distinct sample clusters, with Cluster 1 including 40% of sepsis samples and closely mixed with all post-surgery samples in principal component analysis. The 157 proteins up-regulated in Cluster 1 (FDR<0.01) were significantly enriched in signals of cell death, release of cellular components, and extracellular matrix organisation. Cluster 1 was also associated with more severe coagulation activation, impaired kidney and cardiovascular function, but not with mortality. Having accounted for this primary signal of variation, we then tested for evidence of our previously defined transcriptomic endotypes in the plasma proteome. These endotypes were associated with differential abundance of 7 cytokines and 51 proteins, enriched for innate immune response and inflammatory response annotations. A prediction model for endotype based on proteins gave correct prediction in 82% of test samples.

Conclusions:
Proteome-based sample clustering identified a group of sepsis patients with a more extreme inflammatory response and a greater impact of cell death. Transcriptomic disease endotypes are also supported by evidence at the proteome level, highlighting opportunities for patient stratification in sepsis.

Group: Kwiatkowski **Authors:** Gavin Band **Poster no: 26**

Evidence for interaction between human alleles and *P.falciparum* genetic variation in east and west Africa.

Genome-wide association studies have found a handful of human genetic variants that affect susceptibility to severe *P.falciparum* malaria. However, the impact of parasite genetic diversity on susceptibility has not been assessed. Here, we use whole-genome sequencing and selective whole genome amplification to study parasite genetics in 3,346 children from

The Gambia and Kenya, ascertained with severe symptoms of malaria in the period 1996-2010. We use these data to implement a test for correlation between host and parasite alleles across many pairs of variants. Perhaps surprisingly, we find evidence for an association between a well-known host protective allele and specific parasite genetic variants in these samples, that is not explained by parasite or host population structure or by other measured covariates in our data. A possible explanation is that genetic resistance to this host allele has arisen and remains segregating across multiple *P.falciparum* populations, with potential implications for disease outcomes and monitoring efforts in sub-Saharan Africa.

Group: Leedham **Authors:** Martijn Koppens **Poster no: 27**

Stromal upregulation of the BMP antagonist GREM1 contributes to colonic regeneration

Martijn Koppens, Hayley Davis, Nadia Nasreddin, Gabriel Valbuena, Lennard Lee, Viktor Koelzer, Matthias Friedrich, Fiona Powrie, Simon Leedham

Background

Inflammatory bowel disease (IBD) is characterized by chronic inflammation of the intestinal mucosa and insufficient epithelial regeneration. Patients with IBD are at an increased risk of developing bowel cancer. At homeostasis, the intestinal epithelium communicates with the surrounding stroma through cell-signalling networks, which includes the Bone Morphogenetic Protein (BMP) pathway, to ensure appropriate epithelial cell fate determination. BMP ligands are secreted by stromal cells neighbouring differentiated epithelial cells. Conversely, BMP antagonists such as GREM1 are secreted by stromal cells nearby epithelial stem cells, where they locally reduce BMP signalling and promote an epithelial stem/progenitor phenotype. However, the role of BMP signalling and its regulation during colonic regeneration are less clear.

Methods

Grem1 expression in response to colonic inflammation was quantified in two different colonic injury mouse models: the dextran sodium sulphate (DSS) model for chronic colitis, and a biopsy wounding model to assess spatiotemporal changes during wound repair. Both *Grem1* knock-in and *Bmp4* knock-in animal models were used to assess the functional role of BMP signalling in these two injury models.

Results

Within 5 hours after colonic injury, the muscle layers underneath the ulcer upregulated *Grem1* expression. Another population of GREM1-expressing cells emerged in the ulcer bed at two days after injury, and had properties of wound-associated fibroblasts. Importantly, *GREM1* upregulation was also found in the stroma of human colitis samples. While DSS-treated transgenic mice with epithelial *Grem1* overexpression had attenuated ulceration scores, *Bmp4* overexpression mice had higher ulceration scores as compared to DSS-treated wildtype mice. Genetic modulation of BMP signalling affected both the proliferative epithelium as well as the non-proliferative, migratory epithelium.

Conclusion

Intestinal stromal GREM1 expression is physiologically and temporarily dysregulated in intestinal injury and stimulates wound healing by altering overlying epithelial cell fate determination. Failure to downregulate BMP signalling in colitic mucosa hampers the epithelial capacity to regenerate. Our findings highlight the importance of intercompartmental cell

signalling in the control of epithelial cell fate in homeostasis and tissue regeneration. These pathways may represent attractive opportunities for therapeutic manipulation in inflammatory bowel disease and inflammation-associated bowel cancer

---

Group: Leedham **Authors:** Nadia Nasreddin

The WNT signalling pathway in colitis-associated colorectal cancer

Nadia Nasreddin[1], Ibrahim Al-Bakir[2], Salpie Nowinski[2], Gabriel Valbuena[1], Trevor Graham[2], Simon Leedham[1]
[1]Wellcome Centre for Human Genetics [2] Barts Cancer Institute

Colorectal cancer (CRC) is the most feared complication of inflammatory bowel disease. The aberrant activation of the wnt signalling pathway has long been established as key in sporadic CRC, with activating mutations being divided into ligand-independent (i.e. alterations in *APC* and *B-catenin*) and ligand-dependent (i.e. alternations in *RNF43* and *RSPOs*). Nevertheless, the source of wnt in colitis-associated CRC is not known. This study aims at identifying the source of aberrant wnt signalling in colitis-associated low-grade dysplasia (LGD). 3'RNAseq of FFPE colitis-associated LGDs and sporadic adenomas was undertaken. The expression of *RSPO2, RSPO3* and *RNF43* was analysed and compared with that of sporadic lesions. *In-situ* hybridisation was performed in colitis samples with high *RSPO2* and *RSPO3* expression, to validate such findings. In samples with low *RNF43*, loss of Ch17q, chromosomal location of *RNF43*, and expression *AHR*, a known *RNF43* regulator, were assessed. Results suggest aberrant *RSPO2, RSPO3* and *RNF43* as the source of aberrant wnt signalling in a proportion of the colitis-associated lesions, although molecular ground truth needs to be established by further fusion and mutational analysis. Where confirmed, this would provide an opportunity for treatment through wnt ligand inhibition, to which patients with ligand-dependent lesions remain sensitive.

Group: BSG core **Author:** Helen Lockstone **Poster no: 28**

Bioinformatics: Core Support and Training

The Core award from Wellcome enables the Bioinformatics Core to provide a wide range of support to WHG researchers that is both flexible and free of charge. We are a group of 7 experienced bioinformaticians, with a range of backgrounds and different areas of expertise, who enjoy helping researchers with their data analysis, to use bioinformatics tools and resources, programming problems, experimental design, or to develop their own bioinformatics skills.

You can contact us with a query or request for support at any time via bioinformatics@well.ox.ac.uk and we will put you in touch with a member of our team. We have extensive experience handling and analysing data from next-generation sequencing platforms including Illumina sequencing (whole genome, exome, transcriptomics, ChIP, ATAC), single cell applications using the 10X Genomics platform, and Oxford Nanopore Technology (ONT) for both DNA sequencing and transcriptomics profiling. We can run computationally intensive steps such as mapping (tailored to the specific lab protocol used), and offer sound advice on different processing and analysis steps – choice of tools and suitable parameters, data quality checks, setting up appropriate analysis models and understanding any unusual behaviour in the data or results.

Bioinformatics skills training is an increasingly important focus, as we seek to train more researchers to handle and analyse the genome-wide datasets routinely generated by current high-throughput technologies (genomics, proteomics and metabolomics). Courses in computational skills and programming (command line, R and Python) as well as bioinformatics data analysis (e.g. RNA-Seq, ChIP-Seq) have been developed and run by members of Bioinformatics Core over the past few years. We are liaising closely with Medical Sciences Division Skills Training and the IT Learning Centre to
expand and coordinate the training provision for the benefit of WHG researchers and others.

Group: Myers **Authors:** Daniel Wells **Poster no: 29**

ZCWPW1 is recruited to PRDM9 binding sites in an allele-specific manner, and is essential for fertility, proper synapsis, and recombination in male mice

Daniel Wells, Emmanuelle Bitoun, Daniela Moralli, Gang Zhang, Anjali Gupta Hinch, Catherine Green, Peter Donnelly, Simon R Myers

During meiosis homologous chromosomes pair (synapse) and exchange DNA. This enables balanced segregation and generates genetic diversity. In many vertebrates exhange is initiated by double-strand breaks (DSBs), at sites where PRDM9 binds and deposits both H3K4me3 and H3K36me3. However, how these marks are recognised remains unknown. Using single cell RNA-sequencing of mouse testis, we identified *Zcwpw1* as highly co-expressed with *Prdm9*. Zcwpw1 is specifically expressed in testis and fetal ovary, and possesses both CW and PWWP domains, known to bind H3K4me3 and H3K36me3 respectively. Here, we show that Zcwpw1 localises to sites bound by PRDM9 and is essential for fertility, proper synapsis, and double strand break repair in male mice. In human HEK293 cells with induction of ZCWPW1 and PRDM9 expression, ZCWPW1 is recruited to PRDM9-bound sites, more strongly than sites with H3K4me3 alone. Separately, ZCWPW1 is also able to bind CpG sites throughout the genome. Male Zcwpw1 homozygous knockout mice show normally positioned DSBs, but severe defects in DSB repair, and asynapsis (98%) resulting in pachytene arrest. In contrast, females show apparently normal fertility. Taken together, these results strongly support a functional interaction between PRDM9 and ZCWPW1, in a model where PRDM9-dependent H3K4 and H3K36 trimethylation enhances stable binding of ZCWPW1 to PRDM9 bound sites. These results also validate the utility of single cell co-expression as a source of candidate meiotic genes for further investigation.

Group: Naismith **Authors:** Micah Lee **Poster no: 30**

Oligosaccharides are components of the peptidoglycan cell wall, the serotype-determining O-antigen, and the protective extracellular capsule. Biochemical studies its synthetic pathway enzymes already led to elucidation of a novel fold and mechanism (Wza) and a deeper understanding their role as drug targets (MraY). Here, we present the crystalline structure of the inner membrane flippase Wzx. This represents the first lipidic cubic phase structure determined at Diamond Light Source’s long wavelength macromolecular crystallography beamline I23, where phases are determined by native sulfur single anomalous diffraction. The structure shows a typical fold, allowing identification of a putative substrate binding site.

Group: Naismith **Authors:** Sisi Gao

A synthetic biology toolbox for antiviral antibacterial C-nucleosides

There is an urgent need for new antibiotics against the so called ESKAPE pathogens and demands for potent antivirals against emerging filoviruses. The C-nucleotide class of natural products are known to possess both these activities. These promising molecules are likely to remain largely unused which given their known biological potency is a serious missed opportunity.

C-nucleotides, such as formycin, were first discovered as natural product antibiotics but due to o_ target activities, are little used and the complex synthetic chemistry surrounding their de novo synthesis has made them unattractive. However, recent work at Gilead has developed C-nucleotide drug to treat previously untreatable Ebola infections. This has driven a resurgence of medicinal chemical interest in C-nucleoside analogs. All nucleosides that have been licensed and commercialised for clinical use have a heterocyclic moiety linked to sugar through a C-N bond, hence termed ”N-nuleotides” This C-N linkage can be broken, however, when these compounds are either exposed to mild acid or react with cellular phosphate in a phosphorylase-catalysed transformation.
These problems are eliminated by replacing the C-N linkage by a C-C bond to give a class of compounds known as ”C-nucleotides”.

Forymycin A and its analogues have been shown to be a potent inhibitors of bacterial purine nucleoside phosphorylase, whereas they are inactive against mammalian enzymes. Therefore, formycin A is of interest for its potential in selective drug design. Furthermore, it has been reported that formycin A exhibits antiproliferative properties with respect to various cancer cells and antiviral activity against influenza virus A14 and human immunodeficiency virus type 1.

Although extensive research has been devoted to study the biological activities of formycins, investigation of their biosynthesis has been rare and slow. Early in vivo and in vitro studies have shown that the pyrazolopyrimidine ring of formycin A is derived from L-glutamate and L-lysine(Figure 2). However, the biosynthetic gene cluster and enzymes responsible for such transformations between these precursors and the final product have not yet been identified. The genes encoding for formycin have been identified by our collaborator and others, but the actual pathway remains elusive. There have been no systematic studies of their structures and chemical mechanisms, indeed the true substrates for many of these enzymes remain to be identified. I have already begun to study the entire formycin pathway and have determined stuctures of two key enzymes, a PLP dependent enzyme and a flavin dependent enzyme, with crystals of the third C-C bond forming enzyme.

Group: O'Callaghan - CCMP **Authors**: Thomas Hiron **Poster no: 31**

Defining molecular mechanisms for disease-risk loci to identify new therapeutic pathways.
Thomas Hiron, Da Lin, Michael Reschen, Esther Ng, Anil Chalisey, Chris O’Callaghan

The molecular mechanisms conferring disease risk are unknown for most GWAS-identified risk loci in common polygenic diseases. At such loci it is generally unclear which of the linked SNPs is (or are) the causative SNPs, in which cell type(s) these SNPs exert their disease-

modifying effect and at what stage(s) in the disease, or in the development of the individual they operate.
To address this bottleneck in the translation of genetic findings to new therapeutic approaches we have developed an approach we loosely term ‘pathonomics’. The essence of this approach is to take a component or stage in the disease pathogenesis, develop a genome-wide map of the chromatin changes that take place during this stage of the disease and then use this map to pinpoint potential GWAS-identified risk SNPs that may be responsible for the risk conferred by their genomic locus. Detailed downstream molecular and cellular biology are used to test and validate hypotheses generated by this approach.

We have used this approach in studying the uptake of ox-LDL by human macrophages, an integral stage of atherosclerosis. Using FAIRE-seq, ATAC-seq and ChIP-seq we generated genome-wide maps of the chromatin changes induced by ox-LDL uptake. We studied disease-associated SNPs in peaks of chromatin that underwent ox-LDL-induced change and identified an exemplar SNP that directly influences binding of the CEBP/B transcription factor and so the transcription of the PPAP2B gene, which encodes an enzyme that degrades inflammatory mediators. The role of CEBP/B was only apparent from identification of this precise SNP and its effects. Thus, this approach highlighted a new pathway in the disease
We are now testing this approach in other aspects of the disease pathogenesis involving different cell types and have preliminary results to present. This work dovetails with other work in the group on novel mechanisms of transcriptional regulation.

Group: O'Callaghan - CCMP **Authors:** Da Lin

Efficient hierarchical assembly of very long DNA constructs
Da Lin and Chris O’Callaghan

The assembly of large scale DNA constructs is important in a wide range of biomedical, bioindustrial and bioagricultural applications. As there are technical limitations to the length of DNA that can be synthesised, longer constructs are made by assembling multiple shorter DNA fragments, often in hierarchical stages. Current approaches have major limitations—in particular, they may leave unwanted sequence ‘scars’ between the assembled fragments and cannot be used to assemble DNA containing specific ‘forbidden’ sequences. With the use of type IIS restriction-enzyme based one-pot DNA assembly systems, the sequence to be assembled needs to be free of the recognition sequence of the type IIS restriction enzyme used for DNA assembly. Such sequence limitation places significant constraints on the sequence that can be assembled. As the length of the sequence to be assembled increases, so too does the probability that it contains such ‘forbidden’ sequences.

We have developed a type IIS restriction enzyme-based DNA assembly method that overcomes the above limitations. The approach uses site-specific DNA methylation to regulate the activity of type IIS restriction enzyme recognition motifs. It also uses methylation protection, whereby a DNA methylase is used to methylate the internal type IIS restriction sites within the DNA fragments to be assembled, and a DNA binding protein is used to selectively protect the flanking type IIS restriction sites from being methylated. This allows the assembled DNA fragment to be released intact. We have demonstrated efficient assembly of large DNA constructs in the hundreds of kilobase range and believe that this approach has potential to become the standard DNA assembly method for synthetic biology.

Group: Owens **Authors**: Ray Owens/Jiandong Huo **Poster no: 32**

Development of a nanobody production pipeline for the structural analysis of membrane proteins and macromolecular complexes

Jiandong Huo[1,2] and Ray Owens[1,2]

[1]Structural Biology Division, Wellcome Centre for Human Genetics, [2]Research Complex at Harwell and Rosalind Franklin Institute

Nanobodies are recombinant antigen-specific single domain antibodies derived from the heavy chain only subset of camelid immunoglobulins. Their small molecular size (15 KDa), high affinity and stability have combined to make them unique targeting reagents with numerous applications in the biomedical sciences. In structural biology, the use of nanobodies as co-crystallization chaperones has been transformative enabling the analysis of complex macromolecules including integral membrane proteins, by X-ray crystallography. Using nanobodies to build larger binding agents that facilitate imaging proteins by cryo-Electron Microscopy (cryoEM) is currently at the forefront of the field. Therefore, we have recently established a nanobody production pipeline at Oxford that involves immunisation of llamas, building of genetic libraries from the llama B cells, and identification of binders through phage-display based selections. To date we have successfully identified antigen-specific nanobodies for >90% (10 out of 11) of targets in our first two cohorts of antigen, for the purpose of structural analysis of membrane proteins and macromolecular complexes by Oxford groups[1]. We are currently working on new methods for generating nanobodies, combining the construction of a synthetic nanobody library with affinity maturation of selected binders *in vitro*.

[1]Collaborators supplying antigens: Tamnay Bharat, Thomas Bowden, Liz Carpenter, Opher Gileadi, Matt Higgins, Yvonne Jones, Susan Lea, Jim Naismith, Simon Newstead, Elena Seiradake, Christian Siebold.

Group: Padilla Parra **Authors:** Irene Carlon **Poster no: 33**

The Cellular Imaging group research focuses in two main axes. First, we develop advanced imaging and spectroscopic tools to quantify protein dynamics and interactions in live cells. These methodologies include-multi-colour single molecule localisation microscopy (SMLM) on three dimensions, fluorescence fluctuation spectroscopy (FFS) and Fluorescence Lifetime Imaging Microscopy (FLIM). During the years we have established several collaborations with academia and industry (including a close collaboration with Leica Microsystems and Abbelight (Paris)). The second axis of our research is the study of HIV entry and fusion in the context of developing a vaccine to eradicate AIDS. We apply tailored and home-made cutting-edge photonics methods in primary cells to test, with single molecule resolution, different broadly neutralising antibodies and host-restriction factors implicated in arresting HIV fusion. Recent discoveries of our group include the role of Dynanim-2 during HIV fusion, the exact stoichiometry of the HIV pre-fusion reaction or the importance of cholesterol during entry.

Group: Siebold **Authors:** Amalie Rudolf **Poster no: 34**

The morphogen Sonic hedgehog inhibits its receptor Patched by a pincer grasp mechanism

Amalie F. Rudolf[1,8], Maia Kinnebrew[2,8], Christiane Kowatsch[1,8], T. Bertie Ansell[3,8], Kamel El Omari[4,8], Benjamin Bishop[1], Els Pardon[5,6], Rebekka A. Schwab[1], Tomas Malinauskas[1], Mingxing Qian[7], Ramona Duman[4], Douglas F. Covey[7], Jan Steyaert[5,6], Armin Wagner[4], Mark S. P. Sansom[3], Rajat Rohatgi[2] & Christian Siebold[1]

[1]Division of Structural Biology, Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK
[2]Departments of Biochemistry and Medicine, Stanford University School of Medicine, Stanford, California, United States of America
[3]Department of Biochemistry, University of Oxford, Oxford, UK.
[4]Science Division, Diamond Light Source, Harwell Science and Innovation Campus, Didcot, UK
[5]Structural Biology Brussels, Vrije Universiteit Brussel (VUB), Brussels, Belgium
[6]VIB-VUB Center for Structural Biology, VIB, Brussels, Belgium
[7]Department of Developmental Biology, Washington University School of Medicine, St. Louis, Missouri, United States of America
[8]These authors contributed equally: Amalie F. Rudolf, Maia Kinnebrew, Christiane Kowatsch, T. Bertie Ansell, Kamel el Omari

Hedgehog (HH) ligands, classical morphogens that pattern embryonic tissues in all animals, are covalently coupled to two lipids - a palmitoyl group at the N-terminus and a cholesteroyl group at the C-terminus. While the palmitoyl group binds and inactivates Patched 1 (PTCH1), the main receptor for HH ligands, the function of the cholesterol modification has remained mysterious. Using structural and biochemical studies, along with the re-assessment of prior cryo-electron microscopy structures, we find that the C-terminal cholesterol attached to Sonic Hedgehog (SHH) binds the first extracellular domain of PTCH1 and promotes its inactivation, thus triggering HH signalling. Molecular dynamics simulations show that this interaction leads to the closure of a tunnel through PTCH1 that serves as the putative conduit for sterol transport. Thus, SHH inactivates PTCH1 by grasping its extracellular domain with two lipidic pincers, the N-terminal palmitate and the C-terminal cholesterol, which are both inserted into the PTCH1 protein core.

Group: Stuart **Authors:** Jingshen Ren & Yuguang Zhao **Poster no: 35**

Structures of receptor engagement of hand, foot and mouth disease (HFMD) causing enteroviruses

Jingshan Ren, Yuguang Zhao, Daming Zhou, Elizabeth E Fry, David I Stuart

HFMD infects mainly children under five-years of age and the symptoms are usually relatively mild, such as fever, oral ulcerations and swellings on the hands and feet. However, the disease is sometimes associated with cardiac and central nervous system complications and even death. It has been a health threat worldwide in recent years, especially in the Asia-Pacific region with around two million cases every year since 2010. More than 20 species of A and B enteroviruses (EVs) can cause HFMD. Whilst EV-A71 and Coxsackievirus A16 (CV-A16) are major etiological agents, infections by CV-A6 and CV-A10 have become increasingly common in recent years. HFMD viruses can be divided into 4 groups depending on their receptor usage for cell entry: SCARB2 (EV-A71, CV-A7, A14 and A1), KREMEN1 (KRM1) (CV-A2, A3-6, A8, A10, A12), CAR (CV-B1-3, B5) and DAF/FcRn (E3, E6, E7, E11, E12). We have determined cryo-EM structures of representative virus-receptor complexes from two HFMD virus groups, EV-A71/SCARB2 and CV-A10/KRM1.

Group: Stuart **Authors:** Mark Boyce/Geoff Sutton **Talk**

Dissection of the assembly pathway of Mammalian orthoreovirus using focused ion beam milling and in situ tomographic analysis of native assembly intermediates in infected cells

Contributors
Mark Boyce, Geoff Sutton, Corey Hecksel, Abhay Kotecha, Xiaofeng Fu, Peijun Zhang, David Stuart

The assembly pathways of viruses have been traditionally described using negative stain transmission EM of infected cells complemented by the analysis of purified stable intermediates using crystallographic or cryo EM analysis. The resolution achieved for cellular images has been limited, and the purification of assembly intermediates limits what can be analysed to abundant particles with selected physico-chemical properties (i.e. long-lived and stable).
Here we report an analysis of the assembly pathway of Mammalian orthoreovirus *in situ* in frozen infected cells using cryo focused ion beam milling of cells flash frozen at 12 hours post-infection to generate lamellae of ~150nm thickness. Assembly intermediates are captured in their native form during vitrification and imaged in the lamellae using cryo-electron tomography. Data were collected on a Gatan K2 detector attached to a Krios microscope. The data have been analysed by reconstruction of 3D tomograms, followed by selection of particles at different stages of assembly and averaging of these with appropriate symmetry (using emClarity). This not only reveals structures corresponding to those determined previously from purified particles, but also reveals assembly states previously inaccessible. A newly discovered star-shaped early assembly intermediate consisting of the inner layer protein Lambda1 will be described. Compared to later assembly stages, specific inter-subunit contacts are migrated by the diameter of one alpha helix analogous to a ratchet mechanism, rather than being made at the level of altered bond lengths. This molecular mechanism allows the assembly of the alternatively shaped indented particle early in the assembly/genome packaging pathway.
As far as we know this is the first time macromolecular structures have been captured inside living cells in sufficient detail to resolve the protein secondary structural elements.

Group: Stuart **Authors:** Liz Fry

Progress towards a virus-free FMDV vaccine

Elizabeth Fry[1], Ren Jingshan[1], Claudine Porta[2], Erwin van den Born[4], Silvia Loureir[3], Carina Leifeld[2], Carla Bravo-de-Rueda[2], Eva Perez[2], Sophia Hodgson[2], Amaya Serrano[4], Ruud Segers[4], Ian Jones[3], Bryan Charleston[2], Dave Stuart[1]

[1] Division of Structural Biology, The Henry Wellcome Building for Genomic Medicine, Headington, University of Oxford, Oxford, United Kingdom; [2] Animal and Microbial Sciences, University of Reading, Whiteknights, Reading, United Kingdom; [3] The Pirbright Institute, Pirbright, Woking, United Kingdom; [4] R&D Swine Biologicals, MSD Animal Health, Boxmeer, The Netherlands.

We have hitherto shown that FMDV VLP vaccines expressed in insect cells can elicit good neutralizing antibody titres in cattle. The challenge remains to provide a multi-valent vaccine product with good coverage that meets all the requirements of the Pharmacopeia. Our target is to generate stabilised capsids for 20 types of virus and to produce vaccine antigen sufficiently cheaply to be commercially viable. Here we present encouraging progress towards these aims.

Group: Taylor **Authors:** Alistair Pagnamenta **Poster no: 36**

Characterising a chromothripsis-like insertion of clinical relevance using nanopore genome sequencing

Alistair T Pagnamenta[1,2,*], Hannah E Roberts[2,*], Maria Lopopolo[2,*], Eduardo Calpena[3,*], Lorne Lonie[2], Duncan Parkes[2], Colin Freeman[2], Kerry A Miller[3], Simon J McGowan[3], Reza Maroofian[4], Samantha J L Knight[1,2], Helen Lockstone[2], Rory Bowden[2], Andrew O M Wilkie[1,3,*], David Buck[2,*], Jenny C Taylor[1,2,*]

1. National Institute for Health Research Oxford Biomedical Research Centre, Oxford, UK; 2. Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK; 3. MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK; 4. Genetics Research Centre, Molecular and Clinical Sciences Institute, St George's, University of London, London, UK; *These authors contributed equally to this work

Chromothripsis is a subtype of structural genomic rearrangement where multiple breakpoints are situated in close proximity. We present an Iranian kindred where multiple individuals had been diagnosed with maxillary prognathism and linkage analysis implicated chr17/19. Short-read genome/exome sequencing detected no plausibly pathogenic SNVs/indels in these regions. CNV analysis detected several duplicated segments clustering on chr1. Although rare, these were initially discounted as they lay outside the linkage regions. However, review of the breakpoints with IGV suggested that these five chr1 segments (7-647kb) had inserted into chr17, near a 137kb inversion.

To help determine the precise structure of this complex rearrangement, we sequenced a single affected individual using the PromethION device alongside 7 other clinical samples. A PCR-based protocol generated 96 mapped Gb from one flow cell with average read lengths of 5-6kb. Employing a PCR-free, fragmentation-free method yielded an extra 41Gb of mapped reads from across four flow cells, with 70,000 reads >50kb in length. Using a pipeline comprising minimap2 and Sniffles, and employing filtering including a 1kb length minimum, we detected 60 structural variants of potential interest across the genome. All 8 chr1/17 breakpoints observed using short-read data were present amongst this call-set; no additional complexity was picked up. Several nanopore reads spanned multiple breakpoints and this helped confirm that the duplicated segments and inversion all lay *in cis*, consistent with a single chromothripsis-like event. Of particular note was a 58kb read that captured both ends of a 18kb duplicated segment as well as one of the inversion breakpoints. Sanger sequencing confirmed co-segregation with disease and identified microdeletions/insertions at 3 of the breakpoints. The chr17 insertion site is close to two well characterised topologically associated domains. We hypothesize that the inserted chr1 sequence contains regulatory elements driving expression of neighbouring gene(s) in the face/maxilla, resulting in the observed phenotype.

Group: Taylor **Authors:** Alistair Pagnamenta

From WGS500 to the 100K Genomes Project; identification of a novel syndrome comprising structural brain abnormalities, cognitive deficiency, lower-limb hypertonia and defective tubulin detyrosination

Alistair T. Pagnamenta,[1] Pierre Heemeryck,[2] Hilary C. Martin,[3] Christophe Bosc,[2] Leticia Peris,[2] Ivy Uszynski,[2] Sylvie Gory-Fauré,[2] Simon Couly,[4] Charu Deshpande,[5] Ata Siddiqui,[6] Alaa A. Elmonairy,[7] WGS500 consortium, Genomics England Research Consortium, Sandeep Jayawant,[8] Sarada Murthy,[9] Ian Walker,[10] Lucy Loong,[11] Peter Bauer,[12] Frédérique Vossier,[2] Eric Denarier,[2] Tangui Maurice,[4] Emmanuel L. Barbier,[2] Jean-Christophe Deloulme,[2] Edward M. Blair,[11] Annie Andrieux,[2] Marie-Jo Moutin[2] and Jenny C. Taylor[1]

[1]NIHR Oxford BRC, Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK; [2]Univ. Grenoble Alpes, Inserm, U1216, CEA, CNRS, Grenoble Institut Neurosciences, 38000 Grenoble, France; [3]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK; [4]MMDN, Univ. Montpellier, INSERM, EPHE, UMR_S1198, Montpellier, France; [5]South East Thames Regional Genetics Unit, Guys and St Thomas NHS Trust, London, UK; [6]Department of Neuroradiology, Kings College Hospital, Denmark Hill, London SE5 9RS, UK; [7]Ministry of Health, Kuwait Medical Genetics Center, Sulaibikhat 80901, Kuwait; [8]Department of Paediatric Neurology, John Radcliffe Hospital, Oxford, UK; [9]Community Paediatrics, Upton Hospital, Slough, UK; [10]Clinical Biochemistry, Wexham Park Hospital, Slough, UK; [11]Oxford Centre for Genomic Medicine, Oxford University Hospitals NHS Foundation Trust, Oxford, UK; [12]Centogene AG, 18055 Rostock, Germany.

The WGS500 study was the first to pioneer the use of genome sequencing for mainstream clinical diagnosis in the UK. Initial efforts identified disease-causing variants in 21% of cases. For many other families, candidate variants were identified but additional evidence was deemed necessary for confirmation of pathogenicity. One such case involved a Pakistani kindred in which linkage analysis had yielded a single 12.7Mb region with a LOD score of 3.01. This candidate region harbored a rare c.82C>T; p.(Q28*) variant in *CCDC23* that was homozygous in all 4 affected individuals. Although genetic co-segregation was confirmed in June 2012, little was known about the function of the encoded 66 amino-acid protein. Several years later, the gene was renamed *SVBP* and its importance for tubulin detyrosination and neuron differentiation was demonstrated. This prompted the functional testing required to confirm p.(Q28*) as the causative allele. The protein-truncating change was associated with an absence of a catalytically active VASH-SVBP enzymatic complex. Results from HEK293T cells were recapitulated in brains and neuronal cultures derived from *Svbp* knockout mice. Reduced levels of detyrosinated tubulin were observed, with concomitant increases in tyrosinated tubulin. Defects of axonal differentiation and a global 7% decrease in brain volume were observed. Some white matter tracts were reduced by 30%. Knockout mice also displayed mild hyperactivity, lower anxiety and impaired social behavior. Thus, SVBP deficient mice recapitulate several features observed in human patients. Additional evidence supporting pathogenicity comes from the identification of biallelic *SVBP* mutations in 2 additional families, with affected individuals sharing significant phenotypic overlaps. Although we registered our findings on GeneMatcher, genetic replication came through an unconventional route; a second family was identified via the CentoMD mutation database whilst a third was ascertained through the 100K Genomes Project; this study represents one of the first novel genes to be reported from that programme. Altogether, our data demonstrate that inactivating variants in *SVBP* cause this neurodevelopmental pathology by leading to changes in brain tubulin tyrosination and alteration of microtubule dynamics and neuron physiology.

Group: Todd **Authors:** Dominik Trzupek **Poster no: 37**

Simultaneous mRNA and protein quantification at the single-cell level delineates trajectories of CD4+ T-cell differentiation

Dominik Trzupek, Melanie Dunstan, Antony J. Cutler, Mercede Lee, Leila Godfrey, Dominik Aschenbrenner, Holm H. Uhlig, Linda S. Wicker, John A. Todd, Ricardo C. Ferreira

The transcriptomic and proteomic characterisation of CD4+ T cells at the single-cell level has been performed traditionally by two largely exclusive types of technologies: single cell RNA-sequencing (scRNA-seq) technologies and antibody-based cytometry. Here we demonstrate that the simultaneous targeted quantification of mRNA and protein expression in single-cells provides a high-resolution map of human primary CD4+ T cells, and reveals precise trajectories of canonical T-cell lineage differentiation in blood and tissue. We report modest correlation between mRNA and protein in primary CD4+ T cells at the single-cell level, highlighting the importance of including quantitative protein expression data to characterise cell effector function. This approach provides a massively-parallel, cost-effective, solution to dissect the heterogeneity of immune cell populations and is ideally suited for the detailed immunophenotypic characterisation of rare and potentially pathogenic immune subsets.

Group: Todd **Authors:** Dan Rainbow **Poster no: 38**

Analysis of *IL2RA* traits associated with autoimmune disease

Daniel B. Rainbow[1], Jennifer Asimit[2], Marcin Pekalski[1], Jamie Inshaw[1], John A. Todd[1], Chris Wallace[2,3], Linda S. Wicker[1]
[1]JDRF/Wellcome Trust Diabetes and Inflammation Laboratory, University of Oxford, Oxford, UK. [2]MRC Biostatistics Unit, University of Cambridge, Cambridge, UK. [3]Department of Medicine, University of Cambridge, UK

IL2RA (the gene encoding the α chain of the IL-2 receptor, CD25) has been associated with multiple autoimmune diseases, including type 1 diabetes (T1D), multiple sclerosis (MS) and autoimmune thyroid disease (AITD). Many different variants in the *IL2RA* region, which are not in strong LD, have been identified as being associated with autoimmune disease. To understand the different associations, we first undertook a haplotype analysis using the SNPs identified by stepwise regression compared to a Bayesian stochastic search approach (GUESSFM) for T1D, MS and AITD. This indicated that the SNPs identified by stepwise for MS (rs2104286) and AITD (rs706779) are tagging the two signals identified by GUESSFM, for which one signal is shared between AITD and MS (rs61839660) and the second is unique to MS (rs62626317) or AITD (rs11597367). The results for T1D are more complex with four signals being identified; one signal being shared with MS and AITD (rs61839660), one shared with AITD (rs11597367) and two unique to T1D (rs6602437 and rs41295121).
Using allele-specific expression with genotype-selected individuals from the Cambridge BioResource we show that the tagging SNPs for MS and AITD do not explain the *IL2RA* phenotypes observed, whereas the two SNP model identified by GUESSFM does explain the phenotypes identified.
This work highlights that stepwise regression does not always identify the correct solution, especially when there is joint tagging, and a haplotype analysis is key to understanding the relationship of the associated SNPs. We provide functional evidence that supports the GUESSFM analysis that there are multiple variants within the *IL2RA* region that alter cellular phenotypes.

Group: Todd **Authors:** Matthew Baxter **Poster no: 39**

Cardiac mitochondrial function depends on BUD23 mediated ribosome programming

Efficient mitochondrial function is required in tissues with high energy demand such as the heart, and mitochondrial dysfunction is associated with cardiovascular disease. Expression of mitochondrial proteins is tightly regulated in response to internal and external stimuli. Here we identify a novel mechanism regulating mitochondrial content and function, through BUD23-dependent ribosome generation. BUD23 was required for ribosome maturation, normal 18S/28S stoichiometry and modulated the translation of mitochondrial transcripts. Deletion of BUD23 in cardiomyocytes reduced expression of mitochondrial proteins and, therefore, mitochondrial content and function, leading to severe cardiomyopathy and death. We discovered that BUD23 selectively promotes ribosomal interaction with low GC-content 5'UTRs. Taken together we identify a critical role for BUD23 in bioenergetics gene expression, by promoting efficient translation of mRNA transcripts with low 5'UTR GC content. BUD23 emerges as essential to mouse development, and to postnatal cardiac function.

Group: Todd **Authors:** Jamie Inshaw

ImmunoChip analysis of type 1 diabetes risk using >60,000 individuals identifies 22 novel regions associated with disease

Type 1 diabetes (T1D) is a complex disease, characterised by autoimmune attack on pancreatic beta cells and driven by diverse genetic and environmental factors. Over 60 chromosome regions are implicated in T1D, but causal variants and genes are not established at most T1D-associated loci. In addition, T1D genetic analyses to date have been in European populations only. To address these gaps, we analyzed a dataset of ancestrally diverse T1D cases (N=16,159), controls (N=25,386), and affected families (N=6143), twice the size of previous T1D genetic association studies. We genotyped participants with the Illumina ImmunoChip array and imputed additional variants using the Trans-Omics for Precision Medicine (TOPMed) multi-ethnic haplotype reference panel.
In a meta-analysis from five genetically-identified ancestry groups (clustering with Africans, Europeans, Finnish, Eastern and Admixed populations in the 1000 genomes reference data), we identified 74 regions associated with T1D, including 22 novel associations, 8 of which resided in densely genotyped ImmunoChip regions. We fine-mapped each associated region with European individuals only using GUESSFM and, if there was an association in other ancestry groups ($p<5\times10^{-3}$), we used a trans-ethnic fine-mapping method PAINTOR to reduce the number of variants in the set of most likely causal variants in some instances.

We examined enrichment in of the prioritised variants from the fine-mapping in open chromatin as measured by ATAC-seq in various different cell types, and found strong enrichment in immune cell types. We did not identify any enrichment in islets or in our control cell type, fibroblasts.
We then examined colocalisation between disease signal and chromatin accessibility quantitative trait loci (caQTL) in total CD4+ T cells from 123 donors without T1D. We identified three regions where the disease signal and caQTL colocalisaed, near the BACH2, ANKRD55 and CENPQ genes.
We have expanded the understanding of T1D genetics and pointed towards potential mechanism of three T1D associated regions by incorporating genetic data with ATAC-seq data.

Group: Todd **Authors:** Irina Stefana

Investigating the role of nuclear versus cytoplasmic hyperphosphorylated Tau in human beta cells during ageing and diabetes

M. Irina Stefana[1]*, Sarah J. Richardson[2]*, Hanna Tulmin[1], Darragh O'Brien[1,3], Marie-Louise Zeissler[2], John K. Chilton[2], Noel G. Morgan[2], John A. Todd[1]
*co-first authors

[1] Juvenile Diabetes Research Foundation (JDRF)/Wellcome Diabetes and Inflammation Laboratory (DIL), Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK
[2]Islet Biology Exeter (IBEx), Institute of Biomedical and Clinical Sciences, University of Exeter Medical School, Exeter, UK
[3]Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK

Aims: The microtubule-associated protein Tau, encoded by *MAPT* (Chr 17q21.31), has been extensively studied in neurodegenerative diseases, including Alzheimer's, and is thought of as a neuronal protein. However, it has also been associated with risk of developing type 1 diabetes. We asked whether Tau and its pathogenic form, hyperphosphorylated Tau (pTau), are expressed in β cells, and might be implicated in β-cell dysfunction.

Methods: A panel of total Tau and pTau antibodies were used in western blotting (WB), co-immunoprecipitation (co-IP), immunocytochemistry (ICC) and immunohistochemistry. Cultured human beta cell lines, EndoC-βH1 and EndoC-βH3, were studied, together with pancreas sections from 36 nPOD organ donors without diabetes (BM1 <31; ages 0 to 75 years) immunostained for Tau, pTau and insulin.

Results: Both total Tau and "pathogenic" pTau variants were present in human beta cells from donors without diabetes and in immortalised β-cells. pTau variants localised preferentially to the nucleus in cultured cells and in the beta cells of young organ donors. With increasing age, there was a dramatic increase in cytosolic pTau (R^2=0.4076; p<0.0001). Co-IP experiments revealed that Tau interacts with several partners in beta cells, notably a range of RNA-binding proteins and spliceosome components.

Summary: Tau as well as pTau variants, widely regarded as cytotoxic, are present in human pancreatic β-cells and a striking correlation exists between age and pTau sub-cellular localisation. We hypothesise that the loss of nuclear pTau and/or the accompanying increase in cytosolic pTau that occurs with age may be associated with reduced proliferative potential and a decline in β-cell function.

Group: Tzima **Authors:** Luke Payne **Poster no: 40**

Deconstructing the ribosomal monolith in angiogenesis

Luke Payne, Adam Keen, Vedanta Mehta, John Reader and Ellie Tzima

The growth of new blood vessels, angiogenesis, is a highly complex and dynamic process that requires the coordination of a range of key cellular processes, including mRNA transcription and microRNA regulation, metabolic signalling and cell cycle control. Protein synthesis is a spatially and temporally coordinated process involving a vast array of components that precisely tune the organismal proteome. Yet, to date, there have been limited studies on the

impact of components of the protein translation machinery in angiogenesis. One critical component of this machinery is the eukaryotic initiation factor 6 (eIF6), which binds the 60S large ribosomal subunit and prevents inappropriate formation of the 80S ribosome. Aortic ring assays from novel inducible endothelial specific eIF6 null mice show disruption and dysregulation of endothelial sprouting, whilst embryos display major developmental defects due to a vascular phenotype at around E11.5. Cellular experiments *in vitro* using Click-iT chemistry and a puromycin analogue surprisingly reveal no defects in nascent protein synthesis with depletion of eIF6. Instead, we find that loss of eIF6 activates the tRNA-dependent amino acid starvation response, which ultimately leads to defects in G1/S cell cycle progression. Our results point towards a role for eIF6 beyond translation and suggest additional mechanisms by which eIF6 regulates the complex angiogenic process.

Group: Tzima **Authors:** Adam Keen **Poster no: 41**

Coupling cellular mechanics and protein translation

Adam Keen, Lisa Simpson, Luke Payne, John Reader and Ellie Tzima

Vascular endothelial cells (ECs) that line blood vessels are highly mechanosensitive. A critical cellular structure that confers mechanosensitivity is the cytoskeleton: it responds to externally applied forces, but it also generates its own forces that ultimately play a central role in several fundamental cellular processes. Protein synthesis (or translation) is a highly complex process that requires the coordinated functions of various players, including ribosomes and their regulation by numerous translation factors. Since the discovery that specific mRNAs associate with the cytoskeleton, there has been increased interest in the complex cross-talk between protein translation and the cytoskeleton. One of the less studied members of the eukaryotic initiation factor family is eIF6. eIF6 binds to the large ribosomal subunit (60S) and in doing so prevents its binding to the small ribosomal subunit (40S) in the cytoplasm. eIF6 therefore regulates the formation of an active 80S ribosome capable of protein translation. We have identified that eIF6 co-localises with the actin cytoskeleton in ECs and that loss of eIF6 leads to alterations in the mechanical properties of the cells, including disorganisation of actin filaments, reduced number of focal adhesions and reduced stiffness. Alignment of ECs in the direction of flow, a hallmark response to shear stress, is also disrupted. Mechanistically, the cytoskeletal defects are not due to reduced expression levels of cytoskeletal proteins but associated with loss of molecular signalling complexes critical for cytoskeleton remodelling and force transduction. We have identified a physical/signalling link between the cytoskeletal and translational machinery that is required for dynamic remodelling of the cytoskeleton in response to mechanical stimulation by coordinating correct localisation of cellular signalling cascades.

Group: Tzima **Authors:** Vedanta Mehta **Talk**

The Guidance Receptor Plexin D1 Moonlights as an Endothelial Mechanosensor

Vedanta Mehta [1,2], Kar-Lai Pang[1,2], Daniel Rozbesky [2,3], Katrin Nather[1,2], Adam Keen[1,2], Dariusz Lachowski[4], Youxin Kong[2,3], Dimple Karia[2,3], Michael Ameismeier[2,3], Jianhua Huang[5], Yun Fang[6], Armando del Rio Hernandez[4], John S Reader[1,2], E. Yvonne Jones[2,3] and Ellie Tzima[1,2]

[1] Cardiovascular Medicine, Radcliffe Department of Medicine

[2] Wellcome Centre for Human Genetics

[3] Division of Structural Biology, Wellcome Centre for Human Genetics

University of Oxford, OX3 7BN, UK

[4] Cellular and Molecular Biomechanics Laboratory, Department of Bioengineering, Imperial College London, London, UK
[5] Department of Medicine, Duke University, Durham, USA
[6] Department of Medicine, University of Chicago, Chicago, IL60637, USA

Shear stress imparted by blood flow on arteries is critical for vascular development and homeostasis but can also be an instigator of atherosclerosis. Endothelial cells (ECs) lining the vasculature use molecular mechanosensors to directly detect shear stress profiles that will ultimately lead to atheroprotective or atherogenic responses. Plexins are key cell-surface receptors for the Semaphorin family of cell-guidance signalling proteins and can regulate cellular patterning by modulating the cytoskeleton and focal adhesion structures. However, a role for Plexins in mechanotransduction has not been examined. Here, we demonstrate a hitherto unrecognised role of Plexin D1 (PlxnD1) in mechanosensation and mechanically-induced disease pathogenesis. PlxnD1 is required for the EC response to shear stress *in vitro* and *in vivo* and regulates the site-specific distribution of atherosclerotic lesions. PlxnD1 is a direct force sensor in ECs that is necessary and sufficient for conferring mechanosensitivity upstream of the junctional complex and integrins. PlxnD1 achieves its binary functions as either a ligand or force receptor by populating two distinct molecular conformations. Our results establish a novel mechanosensor in ECs that regulates cardiovascular pathophysiology and provide a mechanism by which a single receptor can exhibit a binary biochemical nature.

Group: Tzima **Authors:** Lisa Simpson

Lost in translation: endothelial cytoskeleton 'out of line'

Lisa Simpson*, Adam Keen, John Reader, Ellie Tzima (*non-clinical DPhil student)

Departmental affiliations: Division of Cardiovascular Medicine, Wellcome Trust Centre for Human Genetics, University of Oxford

Haemodynamic forces such as fluid shear stress generated by blood flow are critical determinants of vascular function in health and disease. Endothelial cells lining straight regions of the vasculature experience laminar shear stress which induces cytoskeletal alignment parallel to the direction of blood flow. This hallmark shear stress response promotes a protective, anti-inflammatory endothelial phenotype regulated by transcription and translation. Eukaryotic translation initiation factors (eIFs) are key components of the protein translation machinery which have a close association with the cytoskeleton. However, the importance of this interaction in regulating the endothelial phenotype and the relationship between haemodynamics and translational control are largely unexplored. Here, we show an eIF co-localising with the endothelial cytoskeleton in aortic endothelial cells *in vitro* and *in vivo*. Loss of this initiation factor either by siRNA-mediated knockdown *in vitro* or inducible deletion in endothelial cells of mice *in vivo*, causes dramatic impairment of the cytoskeleton and alignment of cells in response to laminar shear stress. Mechanistic experiments show that shear stress-mediated activation of key cytoskeletal signalling pathways are disrupted in the absence of this eIF. Interestingly, preliminary results suggest that a small array of total cytoskeletal protein levels remain unaffected suggesting it is only their dynamic activation in response to shear stress which is reduced. Disturbed shear stress occurs at bends and branching points of blood vessels and in contrast to laminar flow, stimulates a pro-inflammatory endothelial phenotype characterised by poor cytoskeletal organisation and increased cell turnover. In the absence of this eIF under disturbed flow conditions, cell cycle signalling, and endothelial cell proliferation are altered. Our working model demonstrates a

crucial role for protein translation initiation factors in the endothelial cell response to fluid shear stress.

Group: Watkins **Authors:** Yingjie Wang **Poster no: 42**

Adaptive immunity limits fibrosis and cardiac remodelling in hypertrophic cardiomyopathy

Ying-Jie Wang[1]*, Julia Beglov[1], Matthew Kelly[1], Alex Stockenhuber[1], Sahar Ghaffari[1], Lee-Anne Stork[2], Mahon Maguire[2], Jurgen Schneider[3], Fiona Powrie[4], Vincenzo Cerundolo[5], Charles Redwood[1], Houman Ashrafian[1], Hugh Watkins[1]

Departmental affiliations: [1]RDM Cardiovascular Medicine, [2]Department of Oncology, [4]The Kennedy Institute of Rheumatology, [5]Weatherall Institute of Molecular Medicine - MRC Human Immunology Unit, University of Oxford; [3]Leeds Institute of Cardiovascular and Metabolic Medicine, University of Leeds, United Kingdom

Hypertrophic cardiomyopathy (HCM) is the commonest inherited cardiac disorder, characterised by left ventricular hypertrophy (LVH), myocardial fibrosis (MF) and sudden death. Resulting largely from sarcomeric mutations, HCM exhibits profound, yet poorly understood, heterogeneity in its phenotypic expression. We propose that sarcomeric mutations drive the initial onset and subsequent immune responses may determine the severity of subsequent remodelling in HCM via its impact on MF. Flow cytometric analysis showed a progressive increase of lymphocytes in a well-established animal model (i.e. cardiac actin *Actc*^{E99K}) of HCM. Notably, cardiac $CD4^+$ T cells were preferentially expanded and activated. Intriguingly, the number of $CD4^+CD25^+Foxp3^+$ regulatory T cells (Tregs) was also elevated in affected hearts but significantly reduced in the spleen. We then generated a double transgenic mouse model which carries an E99K transgene but no mature lymphocytes to investigate the role of lymphocytes in HCM pathogenesis. Cardiac phenotype analyses indicated i) markedly increased induction and distribution of interstitial fibrosis, ii) greater deficits in cardiac function measured by echocardiography and haemodynamic studies, and iii) elevated pericardial space in *ACTC*^{E99K}*Rag-1*$^{-/-}$ mice when compared to those in *ACTC*^{E99K} mice. These data suggest that immune suppressive Tregs may be a candidate cell type for this beneficial outcome.

Group: Zhang **Authors:** Tao Ni **Poster no: 43**

Title: CryoEM structure of HIV-1 capsid in complex with host factor Cyclophilin A

Tao Ni, Samuel Gerard, Jiying Ning, Peijun Zhang
Division of Structural Biology, University of Oxford

Mature HIV-1 capsid is a conical shaped protein shell which encloses HIV-1 genome and viral proteins, comprising a curved surface lattice of capsid protein (CA) hexamers and pentamers. but also provides binding platform for host cellular proteins Cyclophilin A (CypA), a host cell factor directly interacts with a flexible loop of CA. However, how CypA regulates HIV activity remains elusive. Here, we determined the structure of a curved capsid assembly at 3.6 Å resolution, and its complex with CypA at 4.5 Å resolution. Distinct from all existing structures of CA hexamer, the cryoEM CA hexamer is intrinsically highly curved and asymmetric. CypA bridges the curved hexamers through novel inter-molecular contacts between non-canonical

CypA binding site and two adjacent CA protomers, in addition to the canonical CypA binding site. Our structures demonstrate a mechanism of capsid pattern recognition by CypA and provide insights into the mechanism of CypA action.

Group: Zhang **Authors:** Luiza Mendonça

Cryo-electron tomography and sub-tomogram averaging of HIV-1 VLPs using emClarity

Luiza Mendonça[1], Benjamin Himes[2], Peijun Zhang[1,3]

[1]Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK
[2]Janelia Research Campus, Howard Hughes Medical Institute, Ashburn VA, 20147, USA
[3]Electron Bio-Imaging Centre, Diamond Light Source, Harwell Science and Innovation Campus, Didcot OX11 0DE, UK

Cryo-electron tomography (cryo-ET) allows the imaging of 3D objects in their native hydrated state, albeit at low resolution due to the low electron dose necessary for cryo-ET of biological specimens. Computational approaches such as sub-tomogram averaging proved to have the potential to circumvent this limitation and provide high resolution maps of protein complexes in-situ, provided multiple copies of the target protein is present. We have recently developed a new software package geared towards high-resolution subtomogram averaging, emClarity, a GPU accelerated image processing package for enhanced sub-tomogram classification and averaging. Using emClarity, we obtained a 3.1 Å map (Himes et al, 2018) of the immature HIV-1 Gag lattice stabilized by Bevirimat, an HIV-1 maturation inhibitor and commercial antiretroviral drug. In this map, the alpha-helices of capsid N-terminal and C-terminal subdomains and the continuation of spacer 1 peptide (SP1), are well resolved, but not the full SP1 6-helix bundle, a crucial region for HIV-1 maturation and infectivity. A mutation at position T8I has been shown to stabilize immature SP1 6-helix bundle (Fontana et al, 2015). We aim to determine the structure of T8I Gag lattice at near-atomic resolution by cryoET and sub-tomogram averaging. Our preliminary maps show a full SP1 6-helix bundle connecting the CA C-terminal domain and the nucleocapsid region of the Gag. Further improvement of the map with additional data would allow achieving near-atomic resolution, discerning the specific residues making up the important interactions that are critical for the conformational switch during maturation. Atomic models derived from the maps will further guide rational drug design to develop compounds capable of interfering with the Gag lattice stability and therefore, impairing viral infectivity.

Group: Zondervan **Authors:** Nilufer Rahmioglu **Poster no: 44**

Beyond GWAS: Fine-mapping, functional, and phenotypic follow-up of 44 genome-wide significant endometriosis associations in 61K cases and 711K controls

Nilufer Rahmioglu[1,2], Sally Mortlock[3], Peter Loof[4,5], Marzieh Ghiasi[6], Lilja Stefansdottir[7], Reedik Mägi[8], Ayush Giri[9], Genevieve Galarneau[10], Digna Velez Edwards[9], Piraye Yurttas Beim[10], Valgerdur Steinthorsdottir[7], Mette Nyegaard[4,5], Grant Montgomery[3], Stacey Missmer[6,11,12], Andrew P Morris[13], Krina T Zondervan[1,2] on behalf of the International Endometriosis Genomics Consortium (IEGC).

[1]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK.

[2] Endometriosis CaRe Centre Oxford, Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK.
[3] Institute for Molecular Bioscience, The University of Queensland, Brisbane, QLD, Australia.
[4] Department of Biomedicine, Aarhus University, Aarhus, Denmark.
[5] The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Denmark.
[6] Department of Obstetrics, Gynecology and reproductive biology, Michigan State University, Michigan, USA.
[7] deCODE genetics/Amgen, Reykjavik, Iceland.
[8] Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Estonia.
[9] Department of Obstetrics and Gynecology, Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, USA.
[10] Celmatix Inc., New York, NY, USA.
[11] Department of Epidemiology, Harvard University, Boston, USA.
[12] Boston Center for Endometriosis, Boston, MA, USA.
[13] School of Biological Sciences, University of Manchester, Manchester, UK.

Endometriosis is an oestrogen-dependent condition seen in 5-10% of women of reproductive age, associated with infertility and chronic pelvic pain. The heritability has been estimated to be ~50% with 26% due to common genetic variants. We have undertaken the largest genome-wide association study (GWAS) meta-analysis of endometriosis to date, including 61,077 cases and 711,319 controls of European and Japanese ancestry. We identified 44 genome-wide significant (GWS, $P<5x10^{-8}$) loci associated with endometriosis, 32 of which are novel with 22 showing larger effect sizes for severe (American Society of Reproductive Medicine (ASRM) stage III/IV) disease vs. mild (ASRM stage I/II). Conditional analyses revealed multiple distinct association signals for 11 GWS loci, including 7 mapping near the Estrogen Receptor 1 (*ESR1*) gene. A functional GWAS (fGWAS) analysis revealed enrichment of endometriosis association signals mapping to binding sites for 5 transcription factors (EGR1, GATA3, RELA, SUZ12, TFAP2C) and enhancers in the ovary (OVRY-EnhA). EGR1 mediates oestrogen activity to establish uterine receptivity for embryo implantation. GATA3 is induced by oestrogen in a dose/response manner promoting cytokine secretion in endometrial cells. TFAP2C is involved in epigenetic regulation of ESR1 expression in breast cancer. RELA and SUZ12 are involved in immune and inflammatory response pathways. Fine-mapping of endometriosis loci, informed by enriched annotations from the fGWAS analysis, uncovered 15 high-confidence variants (>50% probability of being causal for the association signal): rs9789525 in *GREB1*, rs6456259 near *ID4*, rs851983 in *ESR1* and rs71575922 in *SYNE1*, rs979165 near *miR148a* and rs6970537 in *HOXA10*, rs10757279 near *CDKN2-BAS1*, rs507666 in *ABO*, rs10828247 near *MLLT10*, rs4071558 near *FSHB*, rs3803042 near *HOXC10*, rs6538618 near *VEZT*, rs3742716 in *RIN3*, rs7183386 in *SRP14-AS1*, and rs73241342 near *PLAC1*. Furthermore, we have conducted analyses of RNAseq data from endometrium tissue collected from endometriosis cases (N=58) and controls (N=70), revealing 12 significantly differentially expressed genes mapping to GWS loci. Phenotypic stratified analyses of the GWS loci in deeply phenotyped datasets are ongoing. Taken together the results will provide better understanding of the underlying functional mechanisms of genetic risk variants on endometriosis and its subtypes.

Group: Zondervan **Authors:** Nina Shigesi

Unravelling the Association between Endometriosis and Immunological Diseases: A Study of 273,404 Women from UK Biobank

Nina Shigesi[1,2], Nilufer Rahmioglu[1,2], Holly Harris[3], Hai Fang[1], Andrew Morris[4], Stacey Missmer[5,6,7,8], Julian Knight[1], Christian Becker[2], Krina Zondervan[1,2].

1University of Oxford, Wellcome Centre for Human Genetics, Oxford, United Kingdom.
2University of Oxford, Nuffield Department of Women's & Reproductive Health, Oxford, United Kingdom.
3Fred Hutchinson Cancer Research Center, Seattle, Washington.
4University of Liverpool, Institute of Translational Medicine, Liverpool, United Kingdom.
5College of Human Medicine- Michigan State University, Department of Obstetrics-Gynecology- and Reproductive Biology, Grand Rapids, U.S.A..
6Harvard T.H. Chan School of Public Health, Department of Epidemiology, Boston, U.S.A..
7Boston Children's and Brigham and Women's Hospitals, Boston Center for Endometriosis, Boston, U.S.A..
8Boston Children's Hospital and Harvard Medical School, Division of Adolescent and Young Adult Medicine- Department of Medicine, Boston, U.S.A..

INTRODUCTION:

Endometriosis is a chronic gynaecological disorder with an estimated prevalence of 5-10% in women of reproductive age and as high as 20-30% in women suffering from infertility (1).
The aetiology of endometriosis is complex and under-explored, yet a hypothesis was formed that women affected by endometriosis have an immunity-associated disorder, allowing the implantation, survival and maintenance of ectopic endometrial cells. However, this association is uncertain due to the low quality of existing evidence (2). Robust quantification of the association between endometriosis and immunological diseases could facilitate understanding of both disorders, and help to provide reference for discovery of novel drug targets, such as immuno-modulators, and development of new diagnostic tools for endometriosis.

RESEARCH AIMS:

To investigating the risk for a range of immunological diseases by analyzing data collected from 273,404 women from the UK Biobank using a case-control and a cohort study approach, and to test robustness of the evidence with adjustment for confounders, stratification on effect modifiers, as well as evaluation of the influence of different disease ascertainment methods.

METHODS:

Study Population:

UK Biobank: 273,404 women aged 40 to 69 yrs old recruited from the UK with: 1) Retrospective and prospective data on disease history from hospital diagnosis records (generated since 1997) and self-reported information; 2) Information on lifestyle, medical history, family history; 3) Physical measurements, biological samples, genome-wide genotype data.

Statistical Analysis:

The association between endometriosis and immunological diseases was investigated through epidemiological analysis of UK Biobank data using both a nested case-control and a retrospective cohort study designs. The case-control analysis has been accounted for confounders such as age, lifestyle or hormonal factors using logistic regression modelling. The retrospective cohort analysis has been conducted using Cox proportional hazards

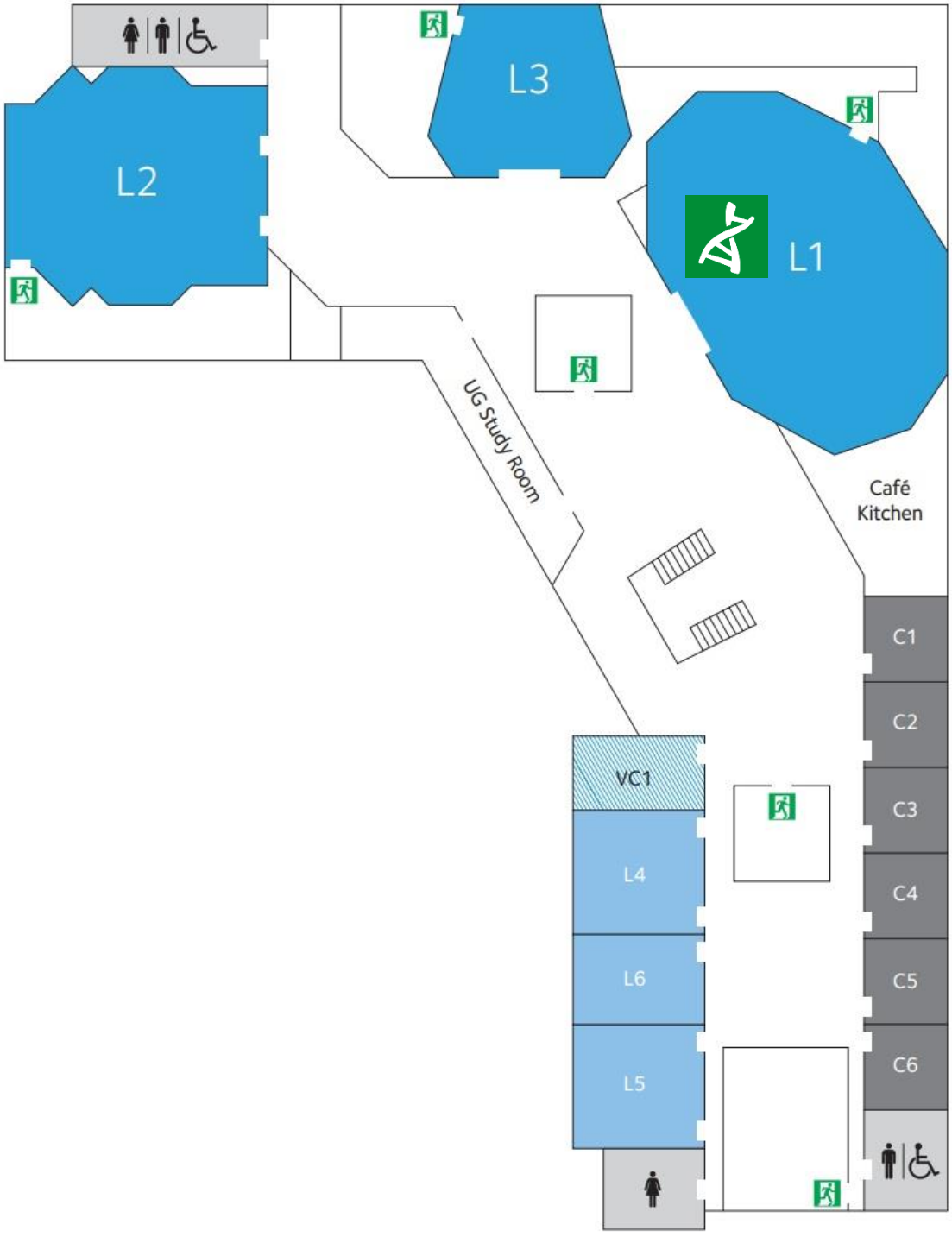
regression models of incident immunological diseases in association with endometriosis history with potential confounders addressed.

RESULTS:
The UK Biobank includes 273,404 women aged ≥40 years old at recruitment, linked to hospital records, of whom 8,223 are diagnosed with endometriosis and 64,620 with immunological diseases (52,027 autoinflammatory; 14,764 classic autoimmune; 4,379 mixed-pattern diseases, definitions and classification of immunological diseases see (3)). By analysing the UK biobank data, women with endometriosis was found to have an increased risk for one or more immunological diseases (OR=1.14 (1.08-1.21), P <0.001* by case-control study; HR=1.14 (1.03-1.26), P=0.012 by cohort study), classic autoimmune diseases (OR=1.12 (1.01-1.23), P=0.029; HR=1.21 (0.97-1.51), P=0.10), autoinflammatory diseases (OR=1.15 (1.08-1.22), P<0.001*; HR=1.13 (1.02-1.25), P=0.023*), mixed-pattern diseases (OR=1.14 (0.97-1.35), P=0.11; HR=1.73 (1.15-2.60), P=0.008*).

CONCLUSIONS:
Among the 273,404 women from the UK Biobank, women with endometriosis were found to have an increased risk of one or more immunological diseases by both a case-control analysis and a cohort analysis study approach. Possible explanation for the observed phenotypic association between endometriosis and immunological diseases might be explained hormonal/reproductive factors or the influence by the chronic pain from endometriosis that has affected the cortisol level.

REFERENCES:
(1) Zondervan, K.T., Becker, C.M., Koga, K., Missmer, S.A., Taylor, R.N. and Viganò, P., 2018. Endometriosis (Primer). Nature Reviews: Disease Primers.
(2) Shigesi, N., Kvaskoff, M., Kirtley, S., Feng, Q., Fang, H., Knight, J.C., Missmer, S.A., Rahmioglu, N., Zondervan, K.T. and Becker, C.M., 2019. The association between endometriosis and autoimmune diseases: a systematic review and meta-analysis. Human reproduction update, 25(4), pp.486-503.
(3) McGonagle, D. and McDermott, M.F., 2006. A proposed classification of the immunological diseases. PLoS medicine, 3(8), p.e297.

MATHEMATICAL INSTITUTE



wellcome centre human genetics