Cambridge Infectious Diseases 5th Annual Meeting

Friday 13th November 2015

Divinity School, St John's College, Cambridge

8.30	*Registration and Coffee*	
9.00	Introduction	Professor James Wood, Chair of Cambridge Infectious Diseases
Session 1: Data science and infectious disease		
9.15	The infectious disease funding landscape and Cambridge	Mike Head, University of Southampton
10.00	Text mining for infectious diseases	Nigel Collier, Department of Theoretical and Applied Linguistics
10.30	GI Infections and clinical informatics	Lydia Drumright, Department of Medicine
11.00	*Tea & Coffee*	
Session 2: Selected presentations		
11.30	Establishment of chronic infection in malaria involves a novel method of immune evasion	Adam Reid, Wellcome Trust Sanger Institute
11.50	Virulence and pathogen adaptation following host shifts	Ben Longdon, Department of Genetics
12.10	Recent advances in TB, Malaria and Kinetoplastids drug discovery at GSK Tres Cantos and new avenues for future Open Lab collaborations	John Haselden (Malaria DPU) David Barros (Tuberculosis DPU) José Fiandor (Kinetoplastids DPU) GSK Tres Cantos
1-2pm	*LUNCH*	
Session 3: Pests and Plagues		
2.00	Pandemic Vision: The third plague pandemic and the birth of epidemic photography	Christos Lynteris, Department of Sociology
2.30	Mathematical models of tree disease epidemics	Nik Cunniffe, Department of Plant Sciences
3.00	Of worms, germs and men: a role for the gut microbiota in helminth-induced suppression of inflammation	Cinzia Cantacessi, Department of Veterinary Medicine
3.30-4pm	*Tea & Coffee*	
Session 4: Selected presentations		
4.00	The therapeutic benefits of Poop in severe *Clostridium difficile* disease	Blessing Anonye, Wellcome Trust Sanger Institute
4.20	Investigating the potential costs and benefits of immunisation against neonatal GBS disease in the UK	Kyriaki Giorgakoudi, Department of Veterinary Medicine
4.40	On the Dynamics of Beliefs and Risky Sexual Behaviour	Flavio Toxvaerd, Department of Economics
5.00	Drinks Reception	

Poster Exhibition

A critical comparison of phylogenetic inference methods for bacterial populations

Lees JA, Harris SM, Parkhill J, Bentley SD
Submitted by: John Lees, Wellcome Trust Sanger Institute

Phylogenies are a critical first step in almost any modern analysis of a bacterial population. Whether summarising the evolution of a species, inferring transmission events or dating the emergence of a modern lineage in many ways they function as the Swiss Army knife of bacterial genomics. With over 5 years of CPU time used over the last month alone to construct these structures, they are clearly important.
What is less clear is what is the best way for researchers to construct them; the Wikipedia page alone lists 50 different methods. For something so central to the field it is surprising a practical guide to construct trees from NGS data of bacterial populations doesn't exist.
To answer this question we have used the recently published Colijin-Kendall metric to compare the performance of a large range of tree construction methods. As our testing framework we use simulation along a known tree, incorporating gene loss, recombination, SNP and INDEL events, and finally generating realistic Illumina reads.
We have quantified the performance of all the methods, and present recommendations for accuracy, speed and ease.

Digital Epidemiology: Modelling of Epidemic Spread using Contact Network Data

Yoneki E
Submitted by: Eiko Yoneki, Department of Computer Laboratory

Respiratory and other close-contact infectious diseases, such as TB, measles and pneumonia, are major killers in much of the developing world. Understanding how the diseases spread and for identifying how best to control them can be tackled by mathematically modelling the spread diseases. Although central to the models, few quantitative data are available on relevant contact patterns, and no study to measure these factors has yet been attempted in rural Africa. I will describe a desirable plan of pilot project to collect human mobility data using RFID sensors, Raspberry Pis and mobile-phones, recording proximity, to gather information on human interactions in rural and remote Amazon-Brazil communities. We have originally exploited device connectivity traces from the real world for modelling social network structure. The empirical study of contact networks shares many issues with network-based epidemiology, and our work has been extended towards understanding the epidemic spread of infectious diseases. Capturing human interactions will provide an empirical, quantitative measurement of social mixing patterns to underpin mathematical models of the spread of close-contact diseases.

Structural basis of clathrin recruitment by Hepatitis D virus

Muenzner J, Kelly BT, Graham SC
Submitted by: Julia Muenzner, Department of Pathology

Hepatitis D virus (HDV) is a small, enveloped RNA virus which can sustain infection only in cells simultaneously infected with Hepatitis B virus (HBV). Superinfection of chronic HBV carriers with HDV can cause fulminant hepatitis or severely aggravate hepatic diseases like cirrhosis and hepatocellular carcinoma. HDV encodes only one protein, the Hepatitis D antigen, which occurs in a small (HDAg-S) and large (HDAg-L) isoform. Previous studies demonstrated that clathrin, a cellular protein involved in endocytosis and intracellular trafficking, binds to HDAg-L. This interaction is

required for the release of enveloped virus-like particles from host cells. We identified the molecular basis of the clathrin:HDAg-L interaction and solved crystal structures of HDAg-L peptides from two HDV genotypes bound to the N-terminal domain of clathrin. Although the affinity of the interaction is comparatively low, the virus mimics the binding mode of cellular clathrin motifs. Our studies aim to enhance our understanding of the HDV egress pathway and to identify molecular targets for novel therapeutics to disrupt the virus life cycle.

The Phylogeny of Rickettsia Using Different Evolutionary Signatures: How Tree-Like is Bacterial Evolution?

Murray GGR, Weinert LA, Rhule EL and Welch JJ.
Submitted by: Gemma Murray, Department of Genetics

Rickettsia is a genus of intracellular bacteria whose hosts and transmission strategies are both impressively diverse, and this is reflected in a highly dynamic genome. It can be vectored to humans via blood-feeding arthropods, and causes diseases such as Typhus and Rocky Mountain spotted fever. Some previous studies have described the evolutionary history of Rickettsia as non-tree-like, due to incongruity between phylogenetic reconstructions using different portions of the genome. Here, we reconstruct the Rickettsia phylogeny using whole-genome data, including two new genomes from previously unsampled host groups. We find that a single topology, which is supported by multiple sources of phylogenetic signal, well describes the evolutionary history of the core genome. We do observe extensive incongruence between individual gene trees, but analyses of simulations over a single topology and interspersed partitions of sites, show that this is more plausibly attributed to systematic error than to horizontal gene transfer. Our results show that, even within a single genus, tests for gene exchange based on phylogenetic incongruence may be susceptible to false positives.

Boom and Bust: Can birth pulses drive pathogens extinct?

Restif O, Peel AJ
Submitted by: Olivier Restif, Department of Veterinary Medicine

The notion of a critical community size (CCS), or population size that is likely to result in long-term persistence of a communicable disease, has been developed based on the empirical observations of acute immunizing infections in human populations, and extended for use in wildlife populations. Seasonal birth pulses are frequently observed in wildlife and are expected to impact infection dynamics, yet their effect on pathogen persistence and CCS have not been considered. To investigate this issue theoretically, we use stochastic epidemiological models to ask how host life-history traits and infection parameters interact to determine pathogen persistence within a closed population. We fit seasonal birth pulse models to data from diverse mammalian species in order to identify realistic parameter ranges. When varying the synchrony of the birth pulse with all other parameters being constant, our model predicted that the CCS can vary by more than two orders of magnitude. Tighter birth pulses tended to drive pathogen extinction by creating large amplitude oscillations in prevalence, especially with high demographic turn- over and short infectious periods.

Quantitative proteomics of the Campylobacter jejuni secretome.
Scanlan E, Yu L, Choudhary J, Maskell DJ, Grant AJ.
Submitted by: Eoin Scanlan, Department of Veterinary Medicine

Campylobacter jejuni is the leading cause of bacterial gastroenteritis in the developed world. *C. jejuni* has been proposed to utilize the flagellar type 3 secretion system (F-T3SS) for delivery of non-flagellar proteins to the extracellular environment. Here SILAC analysis of *C. jejuni* proteomics has been achieved for the first time. This has enabled a hypothesis-free approach resulting in quantitative analysis of the *C. jejuni* strain M1 secretome. We have identified targets which may be highly aided by or dependent upon the F-T3SS for export to the extracellular environment. Mutant strains have been created for each of these targets, along with proteins previously shown to be dependent on the F-T3SS for export. Current work is identifying the phenotypic and virulence properties of each of these strains. Immunogenic tagging of these targets has also been carried out to complement the results of the SILAC screen and validate the value of SILAC as a tool to investigate *C. jejuni* pathogenesis. This work for the first time has revealed a quantitative view of the *C. jejuni* secretome which may prove highly informative in identifying the mechanisms by which *C. jejuni* causes human disease.

Identification of novel genes required for invasion of host epithelial cells by *Campylobacter jejuni*
Gupta S, de Vries SPW, Baig A, Wolanska DP, Maskell DJ, Grant AJ.
Submitted by: Srishti Gupta, Department of Veterinary Medicine

Campylobacter jejuni is the most common bacterial pathogen associated with diarrhoeal illness worldwide; mainly by consumption of contaminated poultry products. The ability of *C. jejuni* to adhere and invade gut epithelial cells is crucial during colonisation of reservoir host species including chickens, as well as for pathogenesis during human infection. Based on genome analyses, Campylobacters were found devoid of conventional adhesins, invasins, and/or a classical type-III secretion system. Despite its importance as an enteric pathogen, mechanism(s) underlying its virulence remain poorly understood. We used a genome-wide negative selection screening technology Tn-seq to identify factors required for adhesion and invasion of gut epithelial cells. A total of 57 genes were identified to influence cellular invasion of *C. jejuni* . Validation studies with 20 defined gene deletion mutants confirmed the role of novel candidates contributing to adherence and invasion properties of *C. jejuni* , with genes of the flagellar system playing a vital role.

From interaction to infection - Identifying social and demographic determinants of viral infection in Red Colobus monkeys in Kibale National Park
Lester JS, Simon DWF
Submitted by: James Lester, Department of Veterinary Medicine

Understanding disease dynamics within wildlife populations poses a major challenge. However by better understanding patterns of infection within wildlife populations, we can better understand, and perhaps even reduce, the risk of emergence of zoonotic pathogens.

Since the identification of the origin of the HIV-1 pandemic in non-human primates (NHPs), a great deal of attention has been paid to the many viruses known to circulate within NHP populations. However given the complex social behaviour of NHPs, it has been difficult to infer a great deal concerning possible patterns of transmission from isolated datasets. Here we present the analysis of a combination of observational social, viral sequencing and demographic data collected from red

colobus monkeys living in Kibale National Park. We identify demographic correlates of infection, alongside a pattern of increased sequence similarity associated with social linkage, and correlation in infection status between strong male-male social linkages. These results indicate the potential for social observations to provide novel insights into patterns of infection for several NHP viruses.

Genetic basis of within host and environmental survival of the foodborne pathogen Campylobacter jejuni

de Vries SPW, Gupta S, Baig A, Wright E, Lacharme-Lora L, MacLeod K, Wedley A, Nygaard-Jensen A, Everest P, Hald B, Humphrey T, Wigley P, Williams NJ, Maskell DJ, Grant AJ.
Submitted by: Stefan de Vries, Department of Veterinary Medicine

Campylobacter jejuni is the leading bacterial cause of foodborne gastroenteritis, with chickens being the main source of human infection. *C. jejuni* is capable of surviving through the food processing chain and has many environmental reservoirs including water, birds, and other domestic animals. To date, infection control by reducing *C. jejuni* prevalence in the food chain has been unsuccessful and vaccines are unavailable. To aid the development of efficient intervention strategies we set out to identify *C. jejuni* genes required for host colonisation (chickens) and infection (piglets and gut epithelial cell invasion), and for its survival under environmental and food-processing related stress conditions (housefly and cold-shock). For this, a *C. jejuni* gene inactivation library (~10,000 transposon mutants) was screened in the various models. To identify the *C. jejuni* genes of mutants that were attenuated in the models, the abundance of each mutant was analysed by massively parallel sequencing of transposon insertion sites (Tn-seq). Novel factors important for survival and host colonisation/infection were identified, with flagellar motility found to be key to *C. jejuni* pathogenesis.

The Influences of Aryl Hydrocarbon Receptor (AhR) on Gut Immunity against Enteric Viral Infections

Eisa O.E., Chettle J., Tonks P., Blacklaws B., Veldhoen M., Heeney J.
Submitted by: Osama Eisa, Department of Veterinary Medicine

AhR is an evolutionary conserved cytosolic transcription factor, which is highly expressed at the barrier sites, mainly the gut. It has been extensively studied for its role in mediating the responsiveness to environmental pollutants. However, recent findings suggest a more physiological role of the AhR system owing to its ability to sense certain dietary elements and to mediate the effects of healthy balanced diet in maintaining the function and homeostasis of the gut. A major impact of AhR deficiency is the reduction in the number and function of the Intraepithelial Lymphocytes (IELs). IELs are specialized T-cells with different subsets, including ab-TCR and gd-TCR IELs, which are located within the epithelial layer of the gut mucosa. Recent evidence suggests a powerful role of activated IELs in early immunity against enteric viral infections. The main aim of this project is to analyse the effects of manipulating the AhR system on the antiviral immunity against enteric viruses. The effects could be mediated directly or through the influences on IELs, therefore another aim of the project is to characterize the role of different IEL-subsets in the early gut immunity against viruses

Class II fusion proteins in endogenous retroviral elements of nematodes; below the fold, a novel case of gene exaption

Merchant MK
Submitted by: Monique Merchant, MRC Laboratory of Molecular Biology

The discovery of viral retroelements containing class II-like envelope fusion proteins is remarkable and calls for our reconsideration of viruses as simply disease-causing agents. A BLAST search with the Rift Vally Fever Virus glycoprotein, Gc, revealed homology throughout the phylum Nematoda. Hits include gag and pol containing viral retroelements in C. elegans and A. ceylanicum, a blood-feeding hookworm. While the viral retroelement of A. ceylanicum seems to be an intact endogenous retrovirus, decay of the gag and pol genes in the C. elegans retroelement render it inactive. Conversely, the env gene of the C. elegans retroelement has acquired a C-terminal extension and is up-regulated during developmental stages. This potential case of gene exaption and the additional novelty of a retrovirus with a class II envelope fusion protein (typically class I) have inspired a biostructural analysis of the potential envelope fusion proteins from the two species. X-ray crystallography will be supplemented with phylogenetic analysis of the endogenous retroviral elements to illuminate the history of this viral evolution.

Using regression trees to derive the genetic basis of antigenic variation in human influenza A/H3N2

James SL, Smith DL
Submitted by: Sarah James, Department of Zoology

It is estimated that seasonal influenza epidemics cause 3-5 million cases of severe illness and 250,000-500,000 deaths annually. According to biannual advice from the WHO (World Health Organisation), the viral components of the influenza vaccine are updated as the virus mutates to evade population immunity.
The antigenic properties of influenza virus are measured by haemagglutination assay and antigenic cartography is used to simplify these data into a map. Experimental work has identified 7 key amino acid positions. We probed the relationship between the antigenic coordinates and the genetic sequence of 253 A/H3N2 strains using regression trees, comparing two methods (one based on the absolute coordinates, the other analysing the distances between strains).
Using the coordinates method identified 6/7 of the key antigenic positions and the distance method identified 5/7; both methods identified an additional amino acid position that does not have a large antigenic effect. Training data was fit well, with significantly poorer performance on the test set. Further optimisation and experimental validation is needed to expand these techniques and fully exploit their potential